

Identification of an Integrated SV40 T/t-Antigen Cancer Signature in Aggressive Human Breast, Prostate, and Lung Carcinomas with Poor Prognosis

Kristin K. Deeb,¹ Aleksandra M. Michalowska,¹ Cheol-yong Yoon,¹ Scott M. Krummey,¹ Mark J. Hoenerhoff,¹ Claudine Kavanaugh,¹ Ming-chung Li,⁵ Francesco J. Demayo,⁶ Ilona Linnoila,² Chu-xia Deng,⁴ Eva Y-H.P. Lee,⁸ Daniel Medina,⁷ Joanna H. Shih,³ and Jeffrey E. Green¹

¹Laboratory of Cell Regulation and Carcinogenesis, ²Cell and Cancer Biology Branch, and ³Biometrics Research Branch, Center for Cancer Research, National Cancer Institute, and ⁴Genetics of Development and Disease Branch, National Institutes of Diabetes, Digestive and Kidney Diseases, NIH, Bethesda, Maryland; ⁵The EMMES Corp., Rockville, Maryland; ⁶Department of Molecular and Cellular Biology, Baylor College of Medicine; ⁷Baylor University Cancer Center, Houston, Texas; and ⁸University of California Irvine, Irvine, California

Abstract

Understanding the genetic architecture of cancer pathways that distinguishes subsets of human cancer is critical to developing new therapies that better target tumors based on their molecular expression profiles. In this study, we identify an integrated gene signature from multiple transgenic models of epithelial cancers intrinsic to the functions of the Simian virus 40 T/t-antigen that is associated with the biological behavior and prognosis for several human epithelial tumors. This genetic signature, composed primarily of genes regulating cell replication, proliferation, DNA repair, and apoptosis, is not a general cancer signature. Rather, it is uniquely activated primarily in tumors with aberrant p53, Rb, or BRCA1 expression but not in tumors initiated through the overexpression of myc, ras, her2/neu, or polyoma middle T oncogenes. Importantly, human breast, lung, and prostate tumors expressing this set of genes represent subsets of tumors with the most aggressive phenotype and with poor prognosis. The T/t-antigen signature is highly predictive of human breast cancer prognosis. Because this class of epithelial tumors is generally intractable to currently existing standard therapies, this genetic signature identifies potential targets for novel therapies directed against these lethal forms of cancer. Because these genetic targets have been discovered using mammary, prostate, and lung T/t-antigen mouse cancer models, these models are rationale candidates for use in preclinical testing of therapies focused on these biologically important targets. [Cancer Res 2007;67(17):8065–80]

Introduction

The misexpression or activation of oncogenes, and loss of tumor suppressor gene function are major determinants of the phenotypic behavior of tumors that significantly influence patient morbidity

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

K.K. Deeb and A.M. Michalowska contributed equally to this work.

Current address for K.K. Deeb: Department of Pharmacology and Therapeutics, Roswell Park Cancer Institute, Buffalo, NY; current address for C-y. Yoon: Biology Department, State University of New York, Stony Brook, NY; current address for C. Kavanaugh: Food and Drug Administration, College Park, MD.

Requests for reprints: Jeffrey E. Green, Transgenic Oncogenesis and Genomics Section, Laboratory of Cancer Biology and Genetics, National Cancer Institute, Room 4054, Building 37, 37 Convent Dr., Bethesda, MD 20892. Phone: 301-435-5193; Fax: 301-496-8709; E-mail: JEGreen@nih.gov.

©2007 American Association for Cancer Research.

doi:10.1158/0008-5472.CAN-07-1515

and mortality. Understanding the genetic networks through which these genes operate is critical to comprehensively defining mechanisms of tumorigenesis and identifying critical genetic nodes that could be important new targets for therapies.

The directed expression of SV40 T/t-antigen has led to the development of several important transgenic models with spontaneous epithelial tumor formation, including carcinomas of pancreatic β -islet cells (1) and acinar cells, ovary (2), choroid plexus, lung (3), mammary (4), and prostate glands (4, 5). The oncogenic relevance of SV40 T/t-antigen to human cancer is that the T-antigen oncoprotein binds to and functionally inactivates two major tumor suppressor genes, *Rb* and *p53* (6, 7), which are often involved in many human tumors, whereas t-antigen dysregulates the protein phosphatase 2A family of serine/threonine phosphatases (8). Unlike many transgenic cancer models induced by the overexpression of oncogenes, such as *ras* and *myc*, T/t-antigen induces significant genomic instability leading to aneuploidy and chromosomal gains and losses that are often seen in human solid tumors (9). Additionally, T/t-antigen can function as one of three minimal cooperating elements along with hTERT and oncogenic *H-ras* to fully transform epithelial cells (10).

Unlike cancer arising in the human population, tumors in GEM models arise in mice with well-defined genetic backgrounds, where genetic variability can be minimized. This offers significant advantages for studying tumor pathogenesis and molecular mechanisms of oncogenesis caused by a single initiating oncogenic event introduced through the mouse germline (11).

To comprehensively define an intrinsic SV40 T/t-antigen signature embedded in multiple epithelial tumors *in vivo* and determine its relevance to human cancers, we did an extensive analysis of tumor gene expression profiles from mouse models of the three most prevalent human epithelial cancers, the breast [C3(1)/Tag transgenic mice; ref. 4], lung (Clara cell secretory protein CC10-Tag mice; ref. 3), and prostate [probasin/Tag [transgenic adenocarcinoma of mouse prostate (TRAMP)] mice; ref. 5]. Although the expression of thousands of genes are altered in these tumors compared with their normal tissue counterparts, we have determined that the three mouse models share an “intrinsic” SV40 T/t-antigen gene signature composed of approximately 150 genes many connected to functional nodes related to p53, pRb, E2F, *myc* DNA damage/repair, replication, and apoptosis pathways.

The intrinsic SV40 T/t-antigen gene signature defines a specific genetic signature that is quite distinctive compared with the expression patterns of tumors arising through the overexpression

of other oncogenes, including *Ras*, *Her2/neu*, *Myc*, and *PyMT*. Approximately half of the named genes in the T/t-antigen signature are involved in cell cycle, DNA repair/binding/metabolism, chromosome assembly, and cytoskeletal/microtubule assembly. Although many of the genes identified in the T/t-antigen signature have also been found in the wound-related (12) and proliferation gene (13) signatures that predict tumor outcome, a substantial number of additional genes related to replicative function, metabolism, and other processes are contained in the T/t-antigen signature. Importantly, when used to cluster human breast, prostate, or lung tumors based on their expression profiles, the intrinsic T/t-antigen gene signature identifies the most aggressive forms of human breast, prostate, and lung cancers with poor prognosis. This gene set should provide further biological insights into human tumors with an aggressive phenotype and poor prognosis.

We propose that the T/t-antigen signature may be informative in classifying patients with poor prognosis, provides additional potential targets for anticancer therapies for epithelial tumors with poor prognosis, and credentials particular SV40 Tag transgenic cancer models for preclinical testing of such targeted therapies.

Materials and Methods

Animals

C3(1)/Tag transgenic mice (4) and Clara cell secretory protein CC10-Tag mice (3) were carried in the FVB background strain. Probasin/Tag (TRAMP) mice (5), obtained from The Jackson Laboratory, were in the C57BL/6 background. Animals were housed and cared for in accordance with NIH guidelines and palpitated for tumors twice weekly. Mice were euthanized by CO₂ narcosis. Normal tissue and tumors from at least four to five mice were analyzed for each mouse tumor model. Tumors (0.6–0.8 cm) were removed, with portions fixed in 4% (w/v) paraformaldehyde for histologic analyses, and the remainder was snapped frozen in liquid nitrogen. Normal dorsal, lateral, and ventral lobes of the prostate gland were microdissected and pooled from the same C57BL/6 mouse. Twenty mammary glands at various stages of estrous cycle and lung tissues were collected from randomly selected wild-type FVB female mice and used for the normal tissue reference sample. During the course of this study, TRAMP mice in the C57BL/6 strain were found to develop seminal vesicle phylloides-like tumors, which were also collected for microarray analysis. Universal Reference RNA for arrays was purchased from Stratagene.

RNA Extraction for Microarray Data Analysis

Total RNA from normal prostate and mammary glands, lung tissues, and tumor samples was obtained by a two-step extraction method using Trizol reagent (Invitrogen) and RNeasy Mini protocol for RNA cleanup by Qiagen. Briefly, 30 mg of tissues were homogenized in 1 mL Trizol reagent using a polytron power homogenizer. After 5-min dissociation of homogenate at room temperature, phase separation of RNA was done by the addition of 200 μ L chloroform, vigorous mixing of the sample, and incubation for 10 min before centrifugation at $14,000 \times g$ for 15 min. The aqueous phase containing the RNA was precipitated with 500 μ L isopropanol, mixed, and incubated for 5 min in room temperature before centrifugation at $14,000 \times g$ for 10 min. The glossy white RNA pellet was washed with 1 mL of 70% ethanol and spun for an additional 5 min at $14,000 \times g$. After the air-dried pellet was hydrated with 100 μ L diethylpyrocarbonate-water, the RNeasy Mini kit (Qiagen) was used to clean up RNA according to the manufacturer's protocol. High-quality total RNA was obtained by elution with 30 μ L diethylpyrocarbonate-water. Twenty micrograms of total RNA from each sample were labeled and hybridized as described (11, 14) using the Incyte GEMII 10K cDNA array spotted on poly-L-lysine-coated glass slides provided by the National Cancer Institute (NCI) Advanced Technology Center.

Statistical Analysis

Microarray data preprocessing and normalization

SV40 T/t-antigen mouse models. cDNA microarray data from four or five independent tumor samples and their respective normal tissue samples were collected using the Incyte GEMII 10K cDNA array (9,984 features). Microarray image analysis and calculation of the average foreground signal adjusted for the local median background was done using GenePix software. Low-quality spots flagged by GenePix were excluded but low-intensity signals were processed as follows. If the signal was less than 100 in both channels, the spots were flagged as not reliable. If the signal was less than 100 in one channel but greater than 500 in the other channel, the signal value in that channel was set to 100. This quality filtering excluded only a small portion of all spots on an array (on average, 0.5%). Feature data were used if present in at least two specimens in each tumor or normal group. Features (9,574) were available for analysis. For each spot, log base 2 ratio of the green (target RNA) and red (Stratagene Universal RNA) signal was calculated. Within each array, the log ratios were normalized with intensity-dependent procedure (Lowess smoothing).

Other transgenic mouse models. Labeled probe from four to five tumor specimens from each transgenic model of breast cancer were cohybridized with labeled probe generated from normal mammary tissue from the same background strain, using the Incyte 10k GEMII chips and GenePix software to extract the raw channel intensities. The preprocessing and normalization steps were done as described above. For each array, feature log base 2 ratios of the intensity from the tumor (red label) and normal (green label) tissue were calculated. An array feature was excluded if there were fewer than three samples with nonmissing log ratios for that array feature in any tumor model, leaving 7,948 array features for the analysis.

All mouse array data are available at the National Center for Biotechnology Information array database.⁹

Human cancer microarray data sets. Human data sets were selected based on their having a high representation of the T/t-antigen gene set, relatively large sample size, and information related to metastases or clinical outcome. The publicly available data sets of breast cancer patients using Agilent oligonucleotide microarrays (NKI/Rosetta Inpharmatics-Merc) were obtained from Chang et al. (15). The downloaded raw ratios (individual tumor tissue cRNA versus all patient pool cRNA) were globally normalized shifting on an array median of log base 2 ratios to 0. The array slide includes over 24,000 features. Three published Affymetrix probe level data, which profile the prostate (16), lung (17), and mammary (18) cancer patients, were processed with the Robust Multiarray Average algorithm and quantile normalization to obtain log base 2 gene summary measures (19). Both the prostate and lung specimens were hybridized to HG-U95 GeneChips, providing over 12,000 oligonucleotide probe sets for the analysis. The mammary samples were assessed on HG-U133A GeneChip with over 22,000 probe sets (16, 17, 19, 20).

Mouse and human gene mapping. Human orthologues of the mouse SV40 T/t-antigen-specific gene signature were identified using Entrez Gene ID and JAX homology as provided in the NCI/Center for Cancer Research (CCR) mAdb database.¹⁰ Entrez IDs with multiple array probes were represented by average expression from the two most correlated probes if the pairwise Pearson correlation coefficient >0.5.

Analysis of Variance

Gene expression profiles from the SV40 T/t-antigen mouse models were compared with respect to specimen type (normal versus tumor tissue), location of tumor (mammary, lung, prostate, and seminal vesicle), and background strain of mice (FVB versus C57BL/6). A three-way ANOVA model with one interaction effect (type X location) was fitted. Cancer genes that differ among all four tumor locations were identified, as those that had a significant interaction effect at the 0.001 level and showed at least a 2-fold change between the maximal and minimal mean tumor/normal ratio over

⁹ <http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE8666>

¹⁰ <http://madb.nci.nih.gov>

the different locations (2,638 cDNA probes). Based on the ANOVA model, differentially expressed genes between normal and tumor specimens within each tumor location were also identified, as those genes whose expression were significant at the 0.001 level and were at least 2-fold different compared with the mean expression ratio. To limit the number of false-positive findings, a stringent criterion of 0.001 was used as a cutoff point for the unadjusted *P* values. We expect that the average number of false-positive findings will be 10 or less. In our analysis, the number of genes reaching this significance level far exceeds the expected false positives. Using the procedure of Benjamin and Hochberg (21), the false discovery rate (FDR) was calculated for each ANOVA testing. The FDRs were all <0.5% (see also Supplementary Table S1). Among the genes that were significant at the 0.001 level, we only reported genes that showed at least a 2-fold change for each comparison considered. Under the conservative assumption that the chance of a false-positive finding is the same for a gene with at least 2-fold change and a gene with less than 2-fold change, the FDRs would be no greater than those reported (see Supplementary Table S1).

Further selection was applied based on identification of differentially expressed genes between normal and tumor tissue for the three epithelial tumors (mammary, lung, and prostate). The SV40 T/t-antigen oncogene-specific signature included genes similarly differentially expressed in each epithelial tumor (153 cDNA clones). In contrast, genes were included in a tissue-specific SV40 T/t-antigen tumor signature if they were found to be differentially expressed between the tumor and normal samples exclusively for one location. Two hundred and eighty-three, 220, and 999 cDNA clones were identified as specifically dysregulated in mammary, lung, and prostate tumors, respectively. Overall, 3,004 unique array features were selected using ANOVA.

Unsupervised Learning

We used agglomerative hierarchical clustering and multidimensional scaling to inspect the global grouping of expression profiles from tumors and normal tissue for all four SV40 T/t-antigen mouse models. Both of these analyses used median-centered gene expression and the distance metric of one-minus the Pearson correlation coefficient. We applied two-way agglomerative hierarchical clustering with image display to visualize expression patterns of SV40 T/t-antigen-specific and tissue-specific tumor gene signatures in three SV40 mouse models (mammary, lung, and prostate) and other independent data sets of mouse and human cancers. The one-minus Pearson correlation coefficient distance metric and average linkage algorithm were used for each clustering analysis. Before clustering for each gene and tumor sample in the SV40 T/t-antigen mouse models, the mean of the respective normal samples was subtracted. For the other mouse models, the normalized log₂ ratio of the tumor/normal gene expression is displayed. For the clustering of human data, expression of each gene was median centered across all specimens.

Survival Analysis

The end point for the study was overall survival defined as the time until death from any cause. Patients alive at the date of last follow-up were censored at that date. To examine the prognostic value of the SV40 T/t-antigen-specific signature, we conducted both unsupervised and supervised methods for predicting survival. Unsupervised methods involved classifying the breast (15) and lung (17) cancer data set samples by hierarchical clustering analysis followed by relating the cluster identity to survival. The survival difference between the two clusters of samples was assessed by the log-rank test.

Two complementary supervised methods were used to establish whether the SV40 T/t-antigen-specific signature predicted survival. First, a leave-one-out cross-validation procedure was used to assess the ability of gene expression to predict survival using the data of 295 breast cancer patients from Chang et al. (15). Each training set was obtained by omitting one patient who was treated as the test patient. For each training set, genes in the SV40 T/t-antigen-specific signature, which were associated with survival, were identified by the Cox proportional hazards model. Using these genes, a classifier was constructed with the linear compound covariate algorithm (22), where the risk score was defined as a linear combination of gene expression values weighted by their estimated

Cox model regression coefficients. If the risk score of the test patient was higher than the median of risk scores obtained in a training set, the test patient was classified to the high-risk group, otherwise to the low-risk group. The procedure was repeated for every patient. A permutation log-rank test was used to assess the significance of prediction.

To further show the predictive ability of the SV40 T/t-antigen-specific signature, a predictor was built by the entire data of 295 breast cancer patients from Chang et al. (15) and validated by the gene expression profiles of 159 breast cancer patients reported by Pawitan et al. (18). Genes in the SV40 T/t-antigen-specific signature, which were associated with the survival in the training set, were used to build the classifier, and each of the 159 test patients was classified by this classifier into the high- or low-risk group. The Supplementary Materials and Methods provides variables needed to predict the prognosis for a new patient. The survival difference between the two risk groups in the test set was assessed by the regular log-rank test. Because the training and test sets were generated with different platforms under different designs (single versus dual channel), to make the data comparable with each other, the expression measurements were standardized to *z* score (i.e., each gene expression was converted to zero mean and unit variance).

Computational Software

Preprocessing of cDNA microarray data and ANOVA were done using Insightful S-Plus 6.0 (Insightful Corp.). The Affymetrix data were preprocessed with R 2.3.1¹¹ and Bioconductor Affy project.¹² Hierarchical clustering and image plots were done using R 2.3.1 and custom Microarr package. Multidimensional scaling was carried out in BRB ArrayTools,¹³ developed by Dr. Richard Simon and Amy Peng Lam. Survival analysis was done using R 2.3.1 and Survival package.

Pathway Analysis

Ingenuity Pathway Analysis (IPA) version 2.0 was used to identify patterns of biological interactions and relationships between the genes of interest identified by the microarray analyses. This algorithm identifies known gene, protein, and regulatory interactions based on existing reports in the literature and schematically depicts the interactions with relative expression levels collected from the microarray analyses. A detailed description of IPA can be found online.¹⁴ Data sets containing Gene ID identifiers and their corresponding expression log₂ fold changes (at least 2-fold mean difference of tumor/normal ratio) from three-way ANOVA analysis were uploaded into IPA. A total of 3,004 genes that showed at least a 2-fold mean difference in the tumor/normal expression ratio (*P* < 0.001) were used for the IPA analysis.

Validation of Observed Gene Expression Profiles by Semiquantitative Reverse Transcription-PCR

To confirm the microarray results, semiquantitative reverse transcription-PCR (RT-PCR) was done on a subset of genes to verify gene expression changes (selected from Table 1). Total RNA (1 μg) was reverse transcribed according to the manufacturer's instructions with oligo(dT) using SuperScript RT III (Invitrogen Life Technologies). One microliter of the cDNA was used for each PCR. The PCR primer information and thermocycling conditions for each gene are summarized in Supplementary Table S2. The PCRs were done using Platinum Taq polymerase according to the manufacturer's protocol (Invitrogen Life Technologies) with varying concentrations of MgCl₂ for specific primer sets (specified in Supplementary Table S2). The PCR thermocycling variables consisted of denaturation at 94°C for 30 s, followed by annealing at either 55°C or 58°C for 1.5 min for 24 to 35 cycles depending on the gene (indicated in Supplementary Table S2), and extension at 72°C for 1 min. PCR products were resolved and visualized on a 2.0% agarose gel containing 0.1 μg/mL ethidium bromide.

¹¹ <http://www.r-project.org/>

¹² <http://www.bioconductor.org>

¹³ <http://linus.nci.nih.gov/BRB-ArrayTools.html>

¹⁴ <http://www.ingenuity.com>

Table 1. Intrinsic SV40 T/t-antigen signature

Gene	Description	LocusLink ID	Gene expression (log 2 tumor/normal ratio)		
			TRAMP prostate	C3(1)-Tag mammary	CC10-Tag lung
Cytoskeleton- and microtubule-related proteins					
<i>Gsn</i>	<i>Gelsolin</i>	227753	-3.19	-2.60	-2.54
<i>Kif10</i>	<i>Kinesin family member 10</i>	229841	1.74	1.39	2.16
<i>Kif11</i>	<i>Kinesin family member 11</i>	16551	3.30	2.35	3.27
<i>Kif20a</i>	<i>Kinesin family member 20A</i>	52563	1.73	1.50	1.84
<i>Kif22-ps</i>	<i>Kinesin family member 22, pseudogene</i>	110033	2.61	1.25	1.84
<i>Kif23</i>	<i>Kinesin family member 23</i>	71819	2.37	1.28	2.03
<i>Kif4</i>	<i>Kinesin family member 4</i>	16571	2.88	1.67	1.87
<i>Mcm10</i>	<i>Minichromosome maintenance protein 10</i>	70024	1.89	1.07	1.49
<i>Mtap1b</i>	<i>Microtubule-associated protein 1 B</i>	17755	3.19	1.47	1.33
<i>Np95</i>	<i>Nuclear protein 95</i>	18140	2.50	1.70	1.40
<i>Spag5</i>	<i>Sperm-associated antigen 5</i>	54141	2.95	1.64	1.97
<i>Stmn1</i>	<i>Stathmin 1</i>	16765	5.97	3.07	1.25
<i>Tubb5</i>	<i>Tubulin, β5</i>	22154	4.37	1.13	1.58
<i>Vil1</i>	<i>Villin</i>	22349	-1.73	-1.45	-1.22
Cell cycle					
<i>Ask</i>	<i>Activator of S phase kinase</i>	27214	1.12	1.14	1.44
<i>Bra1</i>	<i>Breast cancer 1</i>	12189	3.11	1.90	2.41
<i>Bub1b</i>	<i>Budding uninhibited bybenzimidazoles 1 homologue, β (<i>S. cerevisiae</i>)</i>	12236	2.22	1.06	1.44
<i>Ccna2</i>	<i>Cyclin A2</i>	12428	4.06	2.77	3.26
<i>Ccnd2</i>	<i>Cyclin D2</i>	12444	-1.41	-1.77	-1.29
<i>Ccnf</i>	<i>Cyclin F</i>	12449	1.54	1.27	1.30
<i>Cdc2a</i>	<i>Cell division cycle 2 homologue A (<i>S. pombe</i>)</i>	12534	2.15	1.54	1.94
<i>Chek1</i>	<i>Checkpoint kinase 1 homologue (<i>S. pombe</i>)</i>	12649	1.73	1.78	1.93
<i>Cks1</i>	<i>CDC28 protein kinase 1</i>	54124	1.99	1.28	1.60
<i>Ect2</i>	<i>Ect2 oncogene</i>	13605	4.05	2.98	4.10
<i>Gmn</i>	<i>Geminin</i>	57441	2.42	1.83	2.01
<i>Lmo2</i>	<i>LIM only 2</i>	16909	-1.53	-1.04	-1.18
<i>Mad2l1</i>	<i>MAD2 (mitotic arrest deficient, homologue)-like 1 (yeast)</i>	56150	2.41	1.93	1.78
<i>Mtbp</i>	<i>Expressed sequence A1429604</i>	105837	1.42	1.01	1.25
<i>Ndr2</i>	<i>N-myc downstream regulated 2</i>	29811	-2.65	-1.42	-1.01
<i>Nek2</i>	<i>NIMA (never h mitosis gene a)-related expressed kinase 2</i>	18005	2.50	2.07	1.58
<i>PCNA</i>	<i>Proliferating cell nuclear antigen</i>	18538	3.27	1.81	2.40
<i>Rbbp6</i>	<i>Retinoblastoma binding protein 6</i>	19647	2.32	1.09	1.06
<i>Rbbp7</i>	<i>Retinoblastoma binding protein 7</i>	245688	1.20	1.03	1.02
<i>Rbl1</i>	<i>Retinoblastoma-like 1 (p107)</i>	19650	1.17	1.27	1.80
<i>Skp2</i>	<i>S-phase kinase-associated protein 2 (p45)</i>	27401	2.15	1.39	1.40
Cell metabolism					
<i>Acadsb</i>	<i>Acyl-CoA dehydrogenase, short/branched chain</i>	66885	-1.12	-1.63	-1.00
<i>Aldh2</i>	<i>Aldehyde dehydrogenase 2, mitochondrial</i>	11669	-2.55	-2.12	-1.81
<i>Car4</i>	<i>Carbonic anhydrase 4</i>	12351	-2.06	-1.28	-1.73
<i>Cbr2</i>	<i>Carbonyl reductase 2</i>	12409	-1.43	-1.16	-2.54
<i>Dpep1</i>	<i>Dipeptidase 1 (rel)</i>	13479	-1.50	-1.21	-1.76
<i>Ech1</i>	<i>Enoyl CoA hydratase 1, peroxisomal</i>	51798	-1.43	-1.76	-1.24
<i>Fmo1</i>	<i>Flavin containing monooxygenase 1</i>	14261	-2.13	-1.14	-2.42
<i>Gstm1</i>	<i>GST, μ1</i>	14862	-1.59	-1.56	-1.10
<i>Lpl</i>	<i>Lipoprotein lipase</i>	16956	-1.68	-2.28	-1.34
<i>Mlycd</i>	<i>Malonyl-CoA decarboxylase</i>	56690	-1.73	-2.45	-1.42
<i>Paics</i>	<i>phosphoribosylaminoimidazole carboxylase, Phosphoribosylaminoribosylaminoimidazole, succinocarboxamide synthetase</i>	67054	1.55	1.24	1.30
DNA binding/repair/recombination					
<i>Fmr1</i>	<i>Fragile X mental retardation syndrome 1 homologue</i>	14265	1.93	1.00	1.14
<i>Hells</i>	<i>Helicase, lymphoid specific</i>	15201	2.46	2.35	2.97

(Continued on the following page)

Table 1. Intrinsic SV40 T/t-antigen signature (Cont'd)

Gene	Description	LocusLink ID	Gene expression (log 2 tumor/normal ratio)		
			TRAMP prostate	C3(1)-Tag mammary	CC10-Tag lung
<i>Mcm3</i>	<i>Mini chromosome maintenance deficient (S. cerevisiae)</i>	17215	1.31	1.05	1.36
<i>Mcmd2</i>	<i>Mini chromosome maintenance deficient 2 (S. cerevisiae)</i>	17216	2.30	2.04	1.74
<i>Mcmd4</i>	<i>Mini chromosome maintenance deficient 4 homologue (S. cerevisiae)</i>	17217	1.77	1.43	1.27
<i>Mcmd6</i>	<i>Mini chromosome maintenance deficient 6 (S. cerevisiae)</i>	17219	3.13	2.41	2.01
<i>Msh6</i>	<i>MutS homolog 6 (E. coli)</i>	17688	1.74	1.50	1.47
<i>Pola1</i>	<i>Polymerase (D directed), α1</i>	18968	2.80	2.57	2.27
<i>Pole</i>	<i>Polymerase (D directed), ϵ</i>	18973	1.48	1.48	1.29
<i>Pole2</i>	<i>Polymerase (D directed), ϵ2 (p59 subunit)</i>	18974	1.39	1.01	1.07
<i>Prim1</i>	<i>Dprimase, p49 subunit</i>	19075	2.43	1.84	1.70
<i>Prim2</i>	<i>Dprimase, p58 subunit</i>	19076	1.83	1.07	1.04
<i>Rad21</i>	<i>RAD21 homolog (S. pombe)</i>	19357	2.73	1.48	1.30
<i>Rad51ap1</i>	<i>RAD51-associated protein 1</i>	19362	2.35	1.28	1.36
<i>Rfc5</i>	<i>Replication factor C (activator 1) 5 (36.5 kDa)</i>	72151	1.38	1.03	1.06
<i>Rrm1</i>	<i>Ribonucleotide reductase M1</i>	20133	1.88	1.15	1.62
DNA metabolism					
<i>Dhfr</i>	<i>Dihydrofolate reductase</i>	13361	2.76	1.84	2.52
<i>Tk1</i>	<i>Thymidine kinase 1</i>	21877	2.06	1.28	1.09
<i>Top2a</i>	<i>Topoisomerase (D) IIα</i>	21973	3.71	1.83	3.02
<i>Tyms</i>	<i>Thymidylate synthase</i>	22171	2.06	1.75	1.65
Chromosome assembly					
<i>Cenpa</i>	<i>Centromere autoantigen A</i>	12615	2.09	1.54	1.38
<i>Cenpf</i>	<i>leucine, glutamic acid, lysine family 1 protein</i>	108000	2.63	1.69	2.00
<i>H2afz</i>	<i>H2A histone family, member Z</i>	51788	3.10	1.74	2.40
<i>Hmgn2</i>	<i>High mobility group nucleosomal binding domain 2</i>	15331	3.30	1.10	1.70
<i>Mki67</i>	<i>Antigen identified by monoclonal antibody Ki 67 SWI/SNF related, matrix associated, actin</i>	17345	2.61	1.63	2.63
<i>Smarca5</i>	<i>Dependent regulator of chromatin, subfamily a, member 5</i>	93762	1.25	1.43	1.20
<i>Suv39h1</i>	<i>Suppressor of variegation 3-9 homologue 1 (Drosophila)</i>	20937	2.13	1.37	1.22
Signal transduction					
<i>Arhgap11a</i>	<i>Mus musculus 10 days neote cerebellum cDNA, RIKEN full-length enriched library, insert sequence</i>	228482	1.92	1.33	1.70
<i>AurkB</i>	<i>Serine/threonine kinase 12</i>	20877	2.30	1.33	1.47
<i>Melk</i>	<i>Materl embryonic leucine zipper kinase</i>	17279	2.25	1.40	2.42
<i>Plk4</i>	<i>Serine/threonine kinase 18 or polo-like kinase 4</i>	20873	1.67	1.40	1.49
<i>Shcblp1</i>	<i>Shc SH2-domain binding protein 1</i>	20419	2.19	1.63	1.53
RNA binding/RNA splicing					
<i>Rbm3</i>	<i>R binding motif protein 3</i>	5935	1.63	1.80	1.16
<i>Slbp</i>	<i>Stem-loop binding protein</i>	20492	1.70	1.71	1.79
Hormone/growth factors/receptors/effectors					
<i>Grin1</i>	<i>Glutamate receptor, ionotropic, NMDA1 (ζ1)</i>	14810	-1.36	-1.08	-2.35
<i>Hmnr</i>	<i>Hyaluron-mediated motility receptor (RHAMM)</i>	15366	2.20	1.19	1.09
Apoptosis					
<i>Birc5</i>	<i>Baculoviral IAP repeat-containing 5</i>	11799	2.27	1.08	1.65
<i>Casp2</i>	<i>Caspase-2</i>	12366	1.43	1.21	1.29
<i>Casp3</i>	<i>Caspase-3, apoptosis-related cysteine protease</i>	12367	1.80	1.02	1.23
Membrane proteins/transporters					
<i>Cav1</i>	<i>Caveolin, caveolae protein</i>	12389	-1.89	-2.84	-2.36
<i>Cd36</i>	<i>CD36 antigen</i>	12491	-2.19	-2.30	-1.58
<i>Cd59a</i>	<i>CD59a antigen</i>	12509	-1.91	-1.36	-1.14
<i>CyP39a1</i>	<i>Cytochrome P450, 39a1 (oxysterol 7α-hydroxylase)</i>	56050	1.31	1.22	1.29
<i>Elov1</i>	<i>Elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3,yeast)-like 1</i>	54325	-1.02	-1.35	-1.73
<i>Fxyd1</i>	<i>FXYD domain-containing ion transport regulator 1</i>	56188	-2.85	-2.24	-1.26
<i>Kpna2</i>	<i>Karyopherin (importin) α2</i>	16647	2.35	2.04	2.16

(Continued on the following page)

Table 1. Intrinsic SV40 T/t-antigen signature (Cont'd)

Gene	Description	LocusLink ID	Gene expression (log 2 tumor/normal ratio)		
			TRAMP prostate	C3(1)-Tag mammary	CC10-Tag lung
<i>Kpna4</i>	<i>ESTs, highly similar to IMA4_MOUSE importin α-4 subunit (karyopherin α-4 subunit) (importin α Q1) [M. musculus]</i>	16649	-1.04	-1.15	-1.33
<i>Mup1</i>	<i>Major urinary protein 1</i>	17840	-3.57	-1.62	-1.60
<i>Mup2</i>	<i>Major urinary protein 2</i>	17841	-4.89	-2.50	-2.93
<i>Ppap2b</i>	<i>Phosphatidic acid phosphatase type 2B</i>	67916	-2.42	-1.90	-2.57
<i>Ptprb</i>	<i>Protein tyrosine phosphatase, receptor type, B</i>	19263	-1.38	-1.44	-1.29
<i>Tcn2</i>	<i>Transcobalamin 2</i>	17841	-2.67	-1.24	-1.63
<i>Uqcrcf1</i>	<i>Ubiquinol-cytochrome C reductase Rieske iron-sulfur polypeptide 1</i>	66694	2.36	1.35	1.38
Protein translation/metabolism					
<i>Mrpl39</i>	<i>Mitochondrial ribosomal protein L39</i>	27393	-1.79	-2.12	-1.04
Transcription factors/regulators					
<i>Cbfb</i>	<i>Core binding factor β</i>	12400	1.29	1.88	1.21
<i>Ezh2</i>	<i>Enhancer of zeste homologue 2 (Drosophila)</i>	14056	2.44	1.90	1.68
<i>FoxM1</i>	<i>Forkhead box M1</i>	14235	2.96	1.62	1.81
<i>Hoxc8</i>	<i>Homeo box C8</i>	15426	3.19	1.93	2.23
<i>Meox2</i>	<i>Mesenchyme homeobox 2</i>	17286	-1.49	-1.77	-1.32
<i>Tacc3</i>	<i>Transforming, acidic coiled-coil containing protein 3</i>	21335	2.53	1.71	1.97
<i>Tfdp1</i>	<i>Transcription factor Dp 1</i>	21781	1.59	1.17	1.20
Calcium binding/signaling					
<i>Calmbp1</i>	<i>Calmodulin binding protein 1</i>	12316	3.34	1.83	2.43
Immune defense response					
<i>C3</i>	<i>Complement component 3</i>	12266	-1.46	-1.47	-1.84
<i>Cxcl12</i>	<i>Chemokine (C-X-C motif) ligand 12</i>	20315	-1.24	-1.62	-1.23
<i>Lta</i>	<i>Lymphotoxin A</i>	16992	2.10	1.97	2.30
Others					
<i>Brp44l</i>	<i>Brain protein 44-like</i>	55951	-2.01	-2.47	-1.10
<i>Crym</i>	<i>Crystallin μ</i>	12971	-1.17	-1.20	-1.69
<i>Dleu2</i>	<i>Deleted in lymphocytic leukemia, 2</i>	328425	1.21	1.35	1.38
<i>Grcc8</i>	<i>Gene rich cluster, C8 gene</i>	14793	2.97	1.23	1.87
<i>PHF6</i>	<i>RIKEN cDNA 4931428F02 gene</i>	70998	2.22	1.13	1.23
<i>Serpinf1</i>	<i>Serine (or cysteine) protease inhibitor, clade F, member 1</i>	20317	-1.37	-1.11	-1.42
<i>Sgol1</i>	<i>RIKEN cD 3300001M08 gene</i>	72415	2.56	2.51	2.26
<i>Tnxb</i>	<i>Tescin XB</i>	81877	-1.56	-1.16	-1.07

Immunohistochemical Staining for Ki67

Ki67 was identified using a rabbit polyclonal anti-Ki67 antibody (Novocastra) and the avidin-biotin-peroxidase (avidin-biotin complex) technique using the Vectastain ABC Rabbit Elite kit (Vector Laboratories, Inc.). Unrelated mouse immunoglobulin was used as the negative control at the same concentration as the specific antibody.

Results

Initial analyses of microarray data from tumor samples from the TRAMP model clustered the tumors into two very distinct groups. Histologic examination revealed that in the C57BL/6 background, phylloides (epithelial-stromal) tumors originating in the seminal vesicles frequently developed in the TRAMP mice (data not shown) in addition to adenocarcinomas of the prostate glands. Because the seminal vesicle phylloides tumors are composed primarily of Tag-positive mesenchymal cells with T-antigen-negative epithelial cells, cluster separately from the epithelial tumors as visualized by hierarchical clustering and multidimensional scaling (Fig. 1), and are a rare human tumor type, they were excluded to identify a T/t-antigen signature of epithelial tumors.

To further identify relations between the four SV40 T/t-antigen tumor types, two exploratory analyses (unsupervised hierarchical clustering and multidimensional scaling) were done using microarray gene expression from 2,638 cDNA clones, for which a statistically significant difference between the tumor versus normal tissue and the tumor location was identified by means of the multifactor ANOVA model (see Materials and Methods). This gene set represents cancer transformation-related genes that distinguish the four tumor types. As shown in Fig. 1, epithelial tumors of the lung, prostate, and mammary glands cluster separately from their respective normal tissues, with the largest separation observed between prostate tumors and normal prostate. Phylloides seminal vesicle tumors and normal seminal vesicle samples group relatively close to each other but cluster independently from tissues from other locations. Because the cellular composition, histologic features, and gene expression pattern of the seminal vesicle phylloides tumors are quite distinct from the epithelial tumors, this provided additional justification to focus the remainder of the study on only the epithelial breast, lung, and prostate tumors.

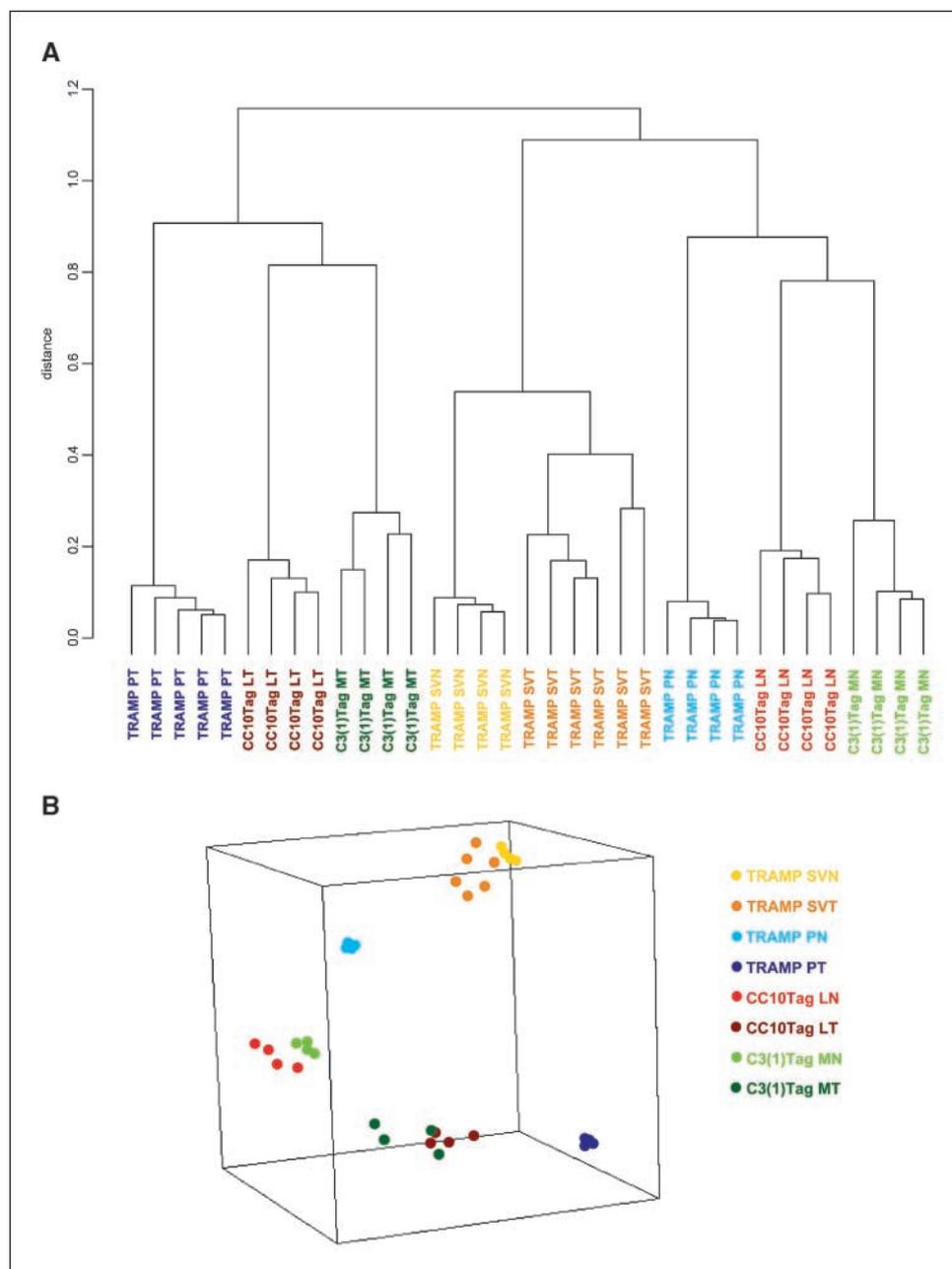
Of the 3,004 genes identified by the full ANOVA analysis (see the Materials and Methods and Supplementary Table S3), global functional analysis could be done on 1,601 genes using IPA. The majority of these genes were classified as cancer related in the disease and disorders functional category with the most significant functional classes, including cell cycle, DNA replication, recombination and repair, cellular growth and proliferation, cell death, and cellular assembly and organization (data not shown).

Identification of a SV40 T/t-antigen intrinsic signature. Array data from transgenic mammary, prostate, and lung tumors induced by the same SV40 viral oncoproteins were analyzed to identify genes that were similarly dysregulated in all of the tumor types. These genes represent an intrinsic set of genes whose expression is highly conserved in all three tumor types, as depicted in a representative heat-map in Supplementary Fig. S1. The SV40 T/t-antigen intrinsic gene signature includes a total of 153 genes of

which 117 (76%) genes were up-regulated and 34 (23%) genes were down-regulated. Only twenty-nine of these genes were similarly dysregulated in the phyllodes seminal vesicle tumors, further showing that the phyllodes tumors are quite distinct from the epithelial tumors (data not shown).

The intrinsic SV40 T/t-antigen signature is specific. Our previous work showed that the gene expression signature for T/t-antigen-induced mammary tumors was quite distinct from gene signatures identified for other transgenic mammary tumors (11). However, to further address whether the newly identified T/t-antigen signature is specific for SV40 T/t-antigen-induced epithelial tumors or whether similar changes are observed in tumors initiated by other oncogenic pathways, we analyzed microarray expression data from other mammary tumor models induced through the loss of p53 (Wap-cre;p53ko^{fllox/fllox} mammary mouse tumor model; ref. 23), loss of BRCA1 in combination with

Figure 1. Comparative cDNA microarray analysis of SV40 T/t-antigen mouse models of human cancers carried out for 2,638 array probes selected by the ANOVA interaction as described in Materials and Methods. *A*, average linkage dendrogram from the hierarchical clustering. *B*, three-dimensional representation of the gene expression profiles. Note that the microarray data sets from each group clustered closely to each other, indicating that all the replicates showed similar gene expression patterns and that the tumor groups are quite dissimilar from the normal tissues. The phyllodes tumor (*orange*) and normal tissue (*yellow*) of the seminal vesicle cluster together and are quite dissimilar to the epithelial tumors and normal tissues, respectively.



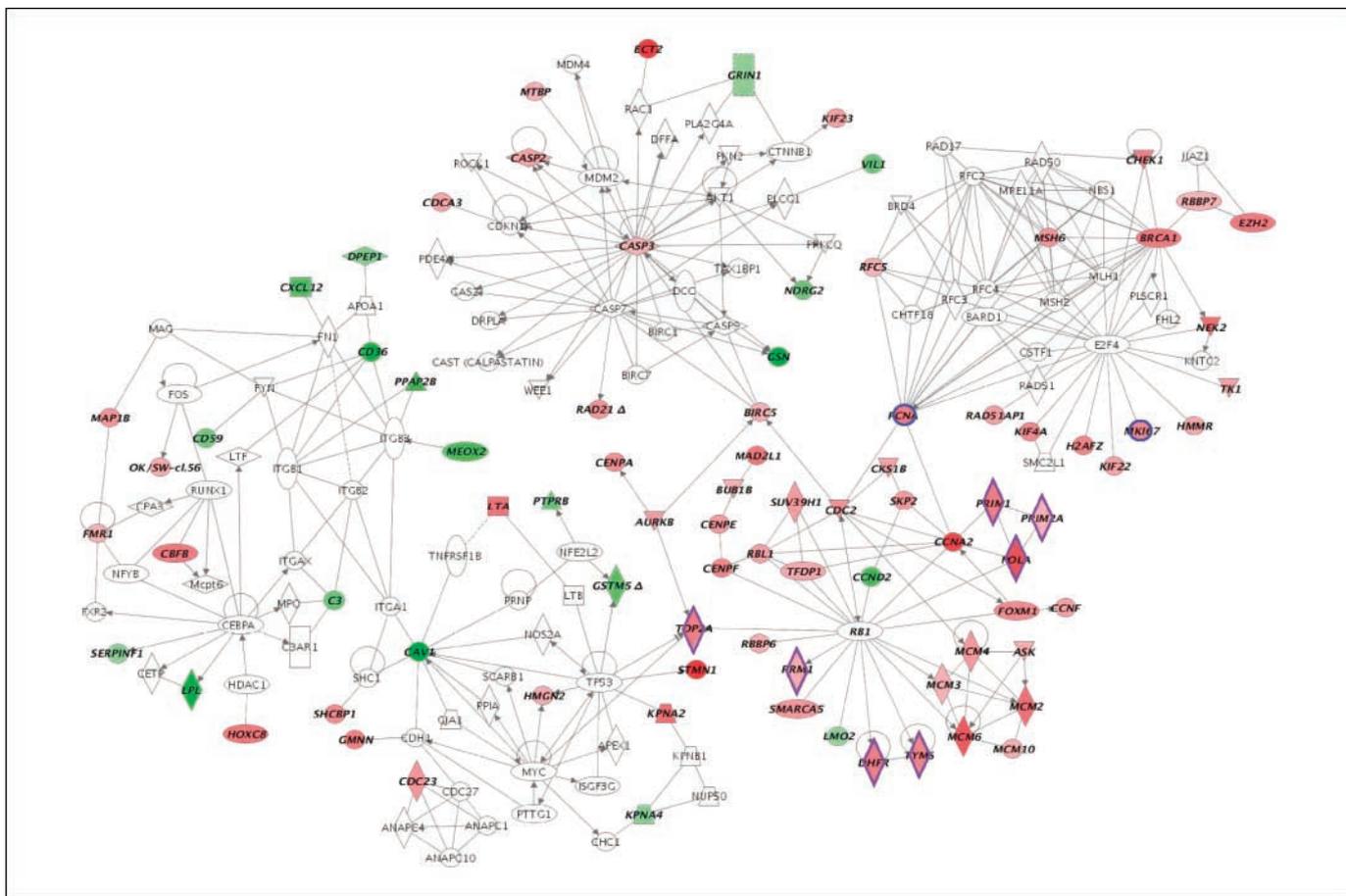


Figure 2. An intrinsic biological network associated with expression of the SV40 T/t-antigen oncoproteins in the GEM models. Of the 120 known genes in the SV40 T/t-antigen gene cluster, 85 of the genes formed biological networks that could be related to the tumor suppressor genes *p53* and *pRB* nodes. Gene symbol and description, LocusLink ID, and expression values for individual nodes for each tumor are listed in Table 1. Up-regulated (red) and down-regulated (green) genes. Genes highlighted in blue are proliferation markers, Ki67 and PCNA, and those that are highlighted in purple are potential chemotherapeutic targets.

p53 (24), the overexpression of *her2/neu* [mouse mammary tumor virus (MMTV)-*her2/neu*; ref. 25], *ras* (MMTV-*ras*; ref. 26), *myc* (MMTV-*c-myc*; ref. 27), or the polyoma middle T viral oncogene (MMTV-PyMT; ref. 28). As depicted in Supplementary Fig. S2, the SV40 T/t-antigen signature is characteristic of the C3(1)-T/t-antigen tumors and is also highly represented in the *p53*^{-/-} and *BRCA1*^{-/-}; *p53*^{+/-} compound mutant tumors that cluster together. However, the expression signature is not characteristic of tumors induced by *myc*, *ras*, *her2/neu*, or PyMT.

A relatively few number of genes in the SV40 T/t-antigen proliferation signature shared similar expression patterns in all of the tumor models. These included caveolin 1 (*Cav1*), enhancer of zeste homologue 2 (*Ezh2*), *Cdca3*, topoisomerase II α (*Top2a*), and a few of the chromosome maintenance factors, suggesting that these genes may be more ubiquitously involved in transformation and not oncogene specific.

The SV40 T/t-antigen intrinsic signature represents an integrated genetic network. The 120 named genes comprising the SV40 T/t-antigen intrinsic gene set is enriched for genes involved in several important cellular pathways that have been implicated in many human cancers (Table 1), including cell cycle, DNA replication/repair/metabolism/maintenance, cytokinesis and cytoskeletal structure, metabolism, transcription, and signal transduction pathways. This set of genes was queried for known functional interactions using IPA (Fig. 2). This algorithm identifies

biological relationships between proteins produced by the genes that directly interact with each other at various functional levels, including physical binding and interaction, protein activation, enzyme substrate association, and transcriptional regulation. The interactions are graphically displayed as networks composed of nodes (individual proteins) and edges (biological relationship between nodes).

Eighty-five of the 120 named genes in the SV40 T/t-antigen intrinsic gene signature are closely related to *p53*, *pRB*, and *E2F* genetic networks. The majority of these genes involve DNA-modifying pathways involving DNA binding, synthesis, and repair, including dihydrofolate reductase (*Dhfr*), DNA polymerases (*Pole*, *Pole2*, and *Pola1*), and primases (*Prim1* and *Prim2*), a family of chromosome maintenance factors (*Mcm*'s) and geminin. Furthermore, G₂ phase and mitosis cell cycle genes, *Ect2*, *Nek2*, *cyclin F*, *cyclin A2*, *Bub1*, and *cenpa*; DNA damage checkpoint regulators, *Brca1*, *cdc2a*, *Chek1*, and *Top2a*; and microtubule-associated process proteins, stathmin 1, kinesin family members, and tubulin, were found to be overexpressed in all three tumor types. Transcription factors, such as forkhead box M1 (*FoxM1*) and homeo box C8, and regulators, *Ezh2*, were up-regulated.

The overexpression of these genes as well as the proliferation markers proliferating cell nuclear antigen (*Pcna*) and *Ki67* clearly shows that SV40 T/t-antigen expression results in enhanced

progression through cell cycle checkpoints leading to rapid cell division in these tumors. Caspase-2 and caspase-3, regulators of apoptosis, were also overexpressed, consistent with the paradoxical increased rate of apoptosis observed in the tumors (29). A few genes involved in cellular metabolism of aldehydes by alcohol dehydrogenase 2 (*Aldh2*) and glutathione by glutathione *S*-transferase (*GST*), μ 1 (*Gstm1*) were down-regulated. Genes involved in cell signaling, adhesion, and cytoskeleton-related proteins (*Cav1*, *cxcl12*, *gelsolin*, and *villin*) were also down-regulated in the tumors, suggesting a loss in specific intracellular signaling pathways and cell-cell adhesion during the tumorigenic process in these models.

Less than half of the genes identified in the T/t-antigen signature are represented in the 70-gene (30), metastasis (31), wound response (12), proliferation (13), and invasive gene (32) signatures that have been reported for classification and prognosis of human tumors (Supplementary Table S1).

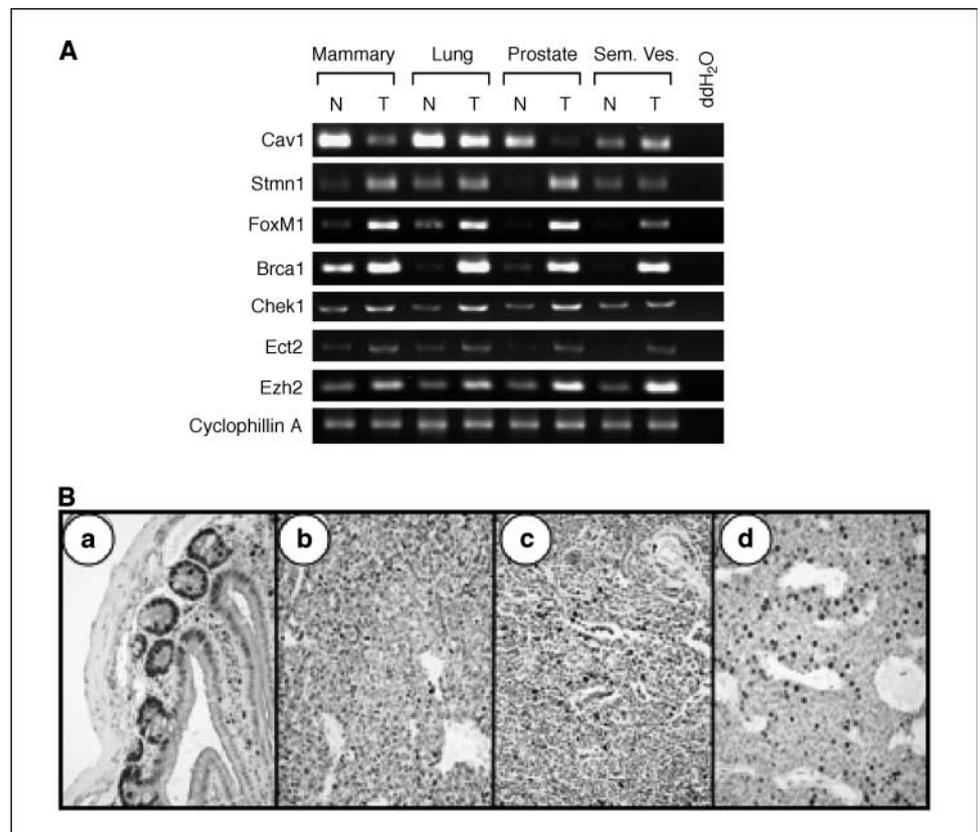
Validation of the T/t-antigen expression profiles. Semiquantitative RT-PCRs were done for several genes of interest identified by our microarray studies and discussed above. The majority of the genes that were in the intrinsic SV40 T/t-antigen signature were confirmed to be similarly up-regulated or down-regulated by RT-PCR analysis (Fig. 3A). Ki67 immunohistochemical staining correlated with the gene microarray data across the different GEM mammary tumors (Fig. 3B). Strong positive Ki67 nuclear immunoreactivity is present in mammary tumors from the C3(1)/Tag model (Fig. 3B, *b*). Weak and multifocal immunoreactivity of Ki67 is found in sections from BRCA1 tumors (Fig. 3B, *c*) and moderately diffused staining in sections of *her2/neu* tumors (Fig. 3B, *d*), further showing that the expression pattern of genes related to cellular proliferation correlates *in vivo* tumor proliferation that is highest

for the SV40 T/t-antigen-induced tumors and not as pronounced in the tumors initiated by other oncogenic pathways (Supplementary Fig. S2).

The intrinsic SV40 T/t-antigen oncogene signature identifies highly aggressive forms of human cancers with poor prognosis. Because several transgenic cancer models induced by T/t-antigen expression share important similarities to subsets of human tumors and are quite aggressive, we examined whether the intrinsic T/t-antigen signature was represented in specific classes of human tumors and whether it correlated with biological behavior and prognosis. Hierarchical clustering using the 120 named genes from the T/t-antigen signature was done using global expression profiles from human breast (15), prostate (16), and lung cancers (17).

One hundred eleven genes (~93% of the intrinsic T/t-antigen gene signature) were available in the data set from Chang et al. (15). Notably, the basal-like and majority of luminal type B subgroups of breast cancers share a remarkably similar expression pattern to the intrinsic T/t-antigen gene profile (Fig. 4A). Additional analyses done with data from Sorlie et al. (20) indicated that the T/t-antigen genes are consistently overexpressed in the basal and luminal type C breast tumors (however, less than half of the T/t-antigen genes were available in this human data set; data not shown). Furthermore, as displayed in Fig. 4A, the majority of breast tumors classified as luminal type-A and normal-like type exhibit lower expression of the T/t-antigen-specific signature. The basal-like subtype is associated with a higher frequency of *TP53* mutations and the worst prognosis, whereas luminal types B and C tumors are associated with considerably poorer overall and relapse-free survival than luminal subtype A as reported previously (20). The ERBB2⁺ tumors do

Figure 3. Validation of SV40 T/t-antigen signature genes in mammary, lung, prostate, and seminal vesicle (*Sem. Ves.*) tissues and tumors. *A*, semiquantitative RT-PCRs were done to confirm changes in transcript expression for FoxM1, Brca1, Chk1, Ect2, Cav1, and Stmn1 as described in Materials and Methods. The lanes listed on top represent normal tissue (*N*) and tumor sample (*T*) from various tissue types. Gene names are listed on the left. All genes were elevated in transcript expression for tumor sample when compared with respective normal tissue. Cyclophilin A served as internal control. *B*, Ki67 immunohistochemical staining of GEM tumors correlates with gene microarray data. *a*, control tissue (small intestine) showing positive nuclear immunoreactivity to Ki67 in crypt epithelia. Strong positive nuclear immunoreactivity to Ki67 is present in mammary tumors from the C3(1)/Tag model (*b*), whereas there is weak and multifocal immunoreactivity in sections from BRCA1 tumors (*c*) and moderate diffuse staining in sections of *her2/neu* tumors (*d*). Anti-Ki67 antibody, Mayer's hematoxylin counterstain. Magnification, $\times 20$.



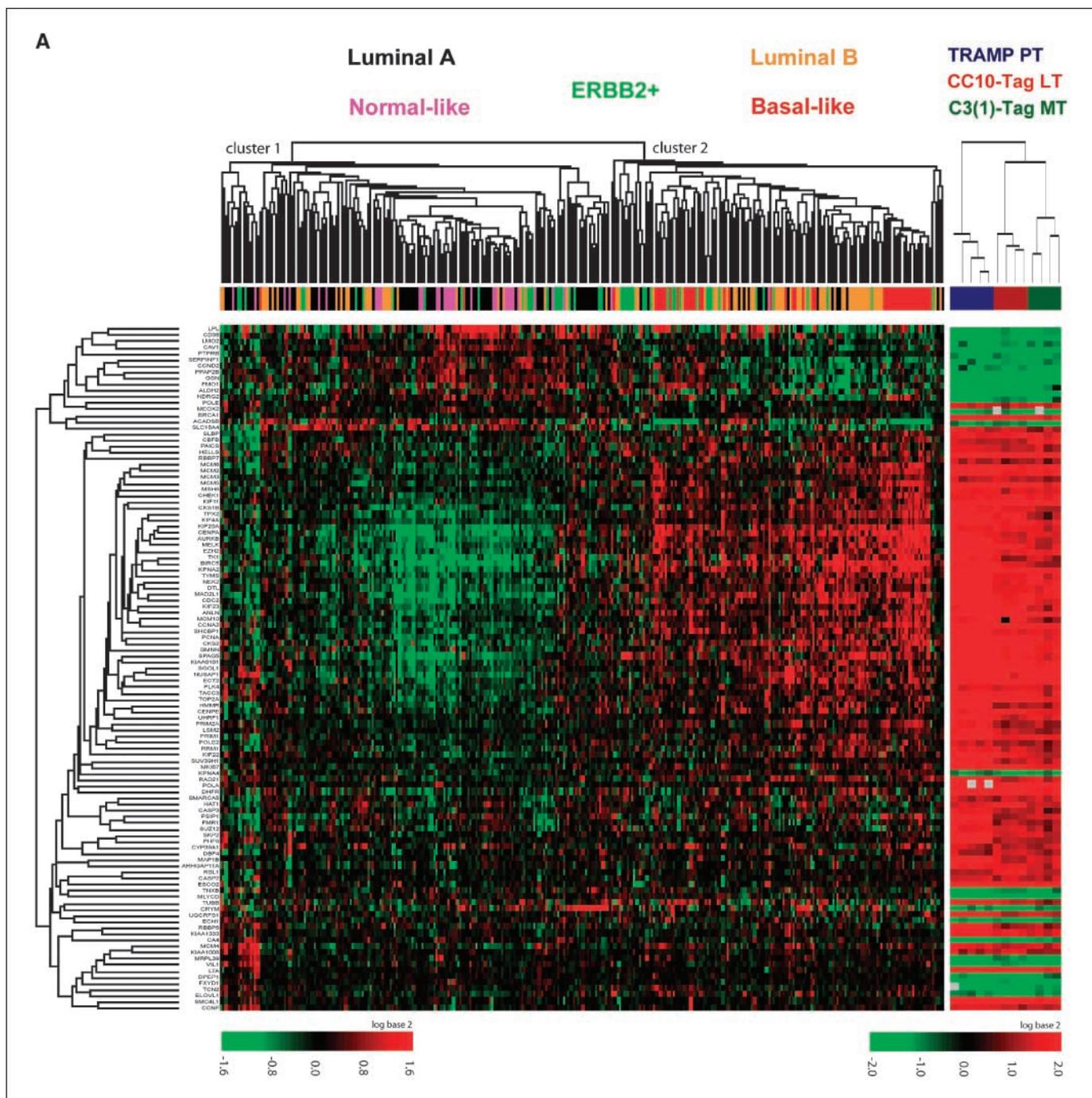


Figure 4. Gene expression patterns of the intrinsic SV40 T/t-antigen signature in human cancer patients. Two-way hierarchical clustering and image plots using available human orthologues of the SV40 oncogene-specific signature in the published human cancer microarray data sets as described in Materials and Methods. In addition, image plots of the SV40 T/t-antigen mouse models are displayed with gene ordering based on the hierarchical clustering of the respective human data set. A, 111 orthologues were found for the subclasses of human breast carcinoma reported by Chang et al. (15).

not seem to segregate into a specific cluster based on the T/t-antigen signature.

Eighty-seven gene orthologues (~72% of the 120 known genes from the intrinsic T/t-antigen gene signature) were identified in the gene expression data set representing 23 primary and 9 metastatic human prostate carcinomas from LaTulippe et al. (16). As depicted in Fig. 4B, metastatic prostate carcinomas cluster separately from primary prostate tumors based on the T/t-antigen signature.

Eighty-three orthologues of the T/t-antigen signature were found in the microarray data of the 186 human lung carcinomas reported by Bhattacharjee et al. (17). Hierarchical clustering showed that all of the small cell and squamous cell lung carcinomas and a subset of adenocarcinomas harbored the intrinsic T/t-antigen signature (Fig. 4C).

When the information on patients' survival was available with the data sets, we calculated the Kaplan-Meier survival curves for those classified by hierarchical clustering into two groups whose

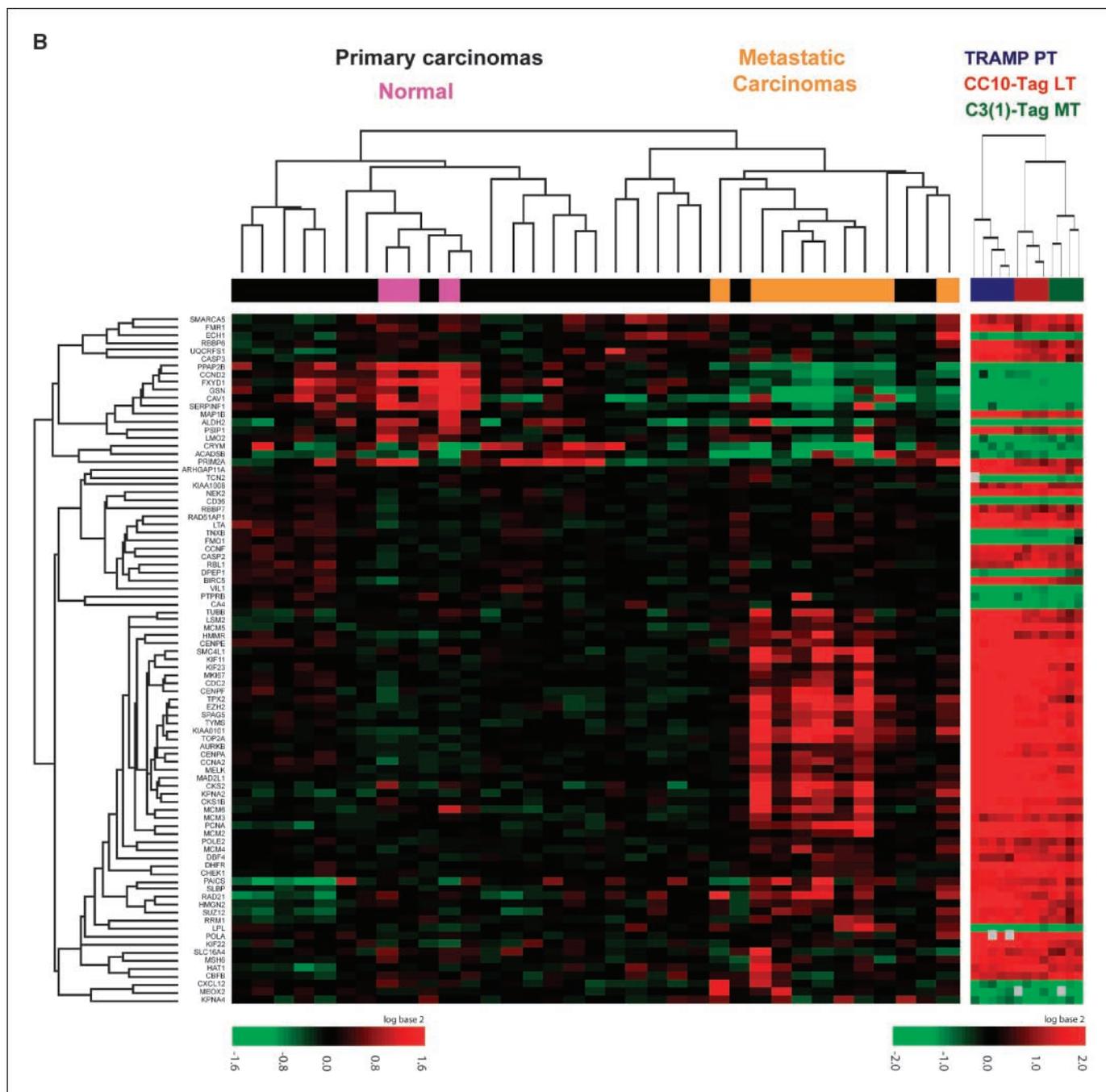


Figure 4 Continued. B, 87 orthologues were available for the expression profiles of primary and metastatic prostate carcinomas published by LaTulippe et al. (16).

expression profile does or does not share the T/t-antigen-specific pattern. The comparison of overall survival between the two groups show that breast cancer patients (Fig. 5A) with tumors expressing the T/t-antigen profile have a very significantly worse prognosis ($P = 7e-10$). The comparison of the lung cancer patient survival was available only for the adenocarcinoma subtype. As shown in Fig. 5B, the adenocarcinoma patients expressing the T/t-antigen signature also have a significantly poorer overall survival ($P = 0.0479$).

These findings strongly suggest that the intrinsic T/t-antigen signature reflects critical molecular derangements that contribute to aggressive clinical behaviors of several types of human epithelial cancers and is associated with poor prognosis.

The SV40 T/t-antigen oncogene signature is highly predictive of survival in breast cancer patients. In the cohort of breast cancer patients analyzed, our clustering analysis indicated a strong association between the T/t-antigen expression profile and overall survival. Therefore, we examined whether the T/t-antigen signature would be useful in predicting the clinical outcome. For this purpose, a supervised analysis was done using two independent population-based data sets of expression profiles of breast cancer patients [i.e., the group of 295 patients from Chang et al. (15) and another 159 patients from Pawitan et al. (18)]. Ninety-eight orthologues (85%) of the T/t-antigen signature were available in both of the microarray data sets. Using a univariate Cox proportional hazards model, we

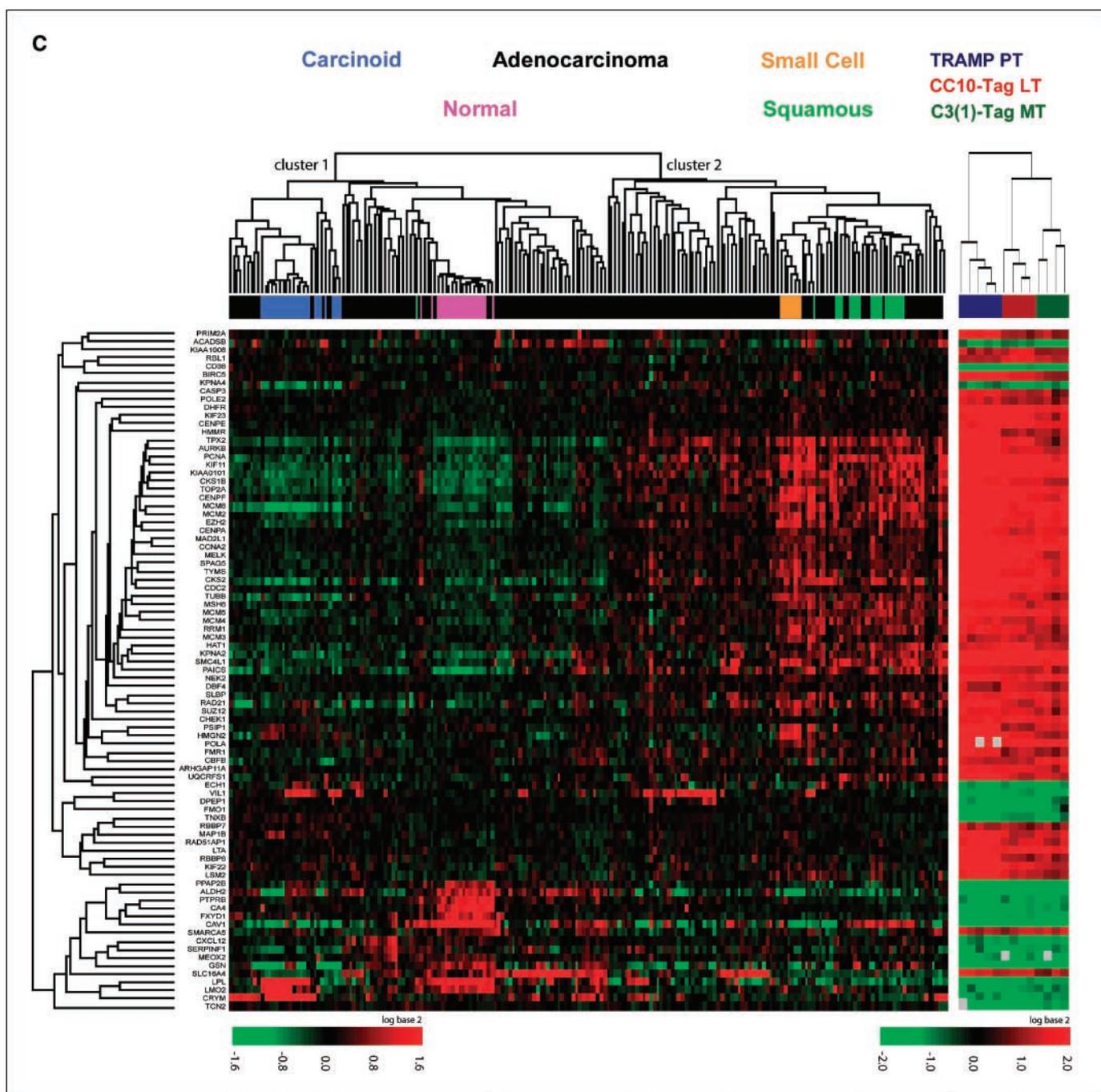


Figure 4 Continued. C, 83 orthologues of the SV40 T/t-antigen gene signature are shown in the subtypes of human lung carcinomas reported by Bhattacharjee et al. (17).

identified a subset of 61 genes significantly associated with overall survival ($P < 0.05$) in the Chang data set. The leave-one-out cross-validated Kaplan-Meier survival curves (Supplementary Fig. S3) for the predicted low- and high-risk groups in the Chang data set (15) exhibited a very clear separation of the groups (permutation $P < 0.001$). The classifier built on the 61 genes from the Chang data set was then applied to the external set of 159 breast cancer patients (28). The expression pattern of the 61 genes in the external test set ordered by increasing risk score is shown in the Fig. 5C. The resulting Kaplan-Meier curves (Fig. 5C) confirm that the T/t-antigen-specific expression profile is highly associated with a poor clinical outcome ($P = 0.00218$).

Discussion

In this study, we have discovered a conserved gene transcriptional signature in multiple epithelial tumors induced by the inactivation of p53 and Rb through the targeted expression of the early region of SV40. Although T/t-antigens are not etiologically involved in human epithelial tumors, the fact that a significant portion of human tumors contain dysfunctional p53 or Rb suggests that T/t-antigen may induce similar oncogenic alterations. The expression pattern identified in this study intrinsic to the expression of the SV40 T/t-antigens in transgenic mouse mammary, prostate, and lung tumors is significant for several reasons: (a) it shows that a core oncogenic mechanism is

operative in all of these epithelial tumor types involving the dysregulation of genes controlling the cell cycle, cell proliferation, replication, DNA damage responses, and apoptosis; (b) the intrinsic signature is quite specific for SV40 T/t-antigen-induced tumors when compared with tumors arising from other (ras, PyMT, and her2/neu) oncogenic pathways and, therefore, is not simply a general signature for cancer; (c) the SV40 T/t-antigen signature is composed of genes that form a highly integrated genetic network based on known functional and physical interactions; and (d) the T/t-antigen signature is predictive of poor prognosis for human breast, prostate, and lung tumors. Therefore, it is likely that this signature contains genes whose dysregulated function might be altered through therapeutic targeting to improve outcome.

Approximately 3,000 genes were found to be differentially expressed between the tumor types compared with their respective normal tissues. Genes expressed in a tumor type-specific manner were primarily related to cell-cell signaling/interaction, extracellular matrix components, growth factors and receptors, tissue structure, and cellular metabolism, suggesting that each tissue type has different requirements for tumorigenesis, especially related to the stromal environments of the breast, lung, and prostate.¹⁵ However, only about 120 named genes were found to be similarly expressed in all three tumor types, suggesting that this core intrinsic T/t-antigen set of genes are likely critical to the transforming function of T/t-antigen.

Importantly, the intrinsic SV40 T/t-antigen gene signature is not a feature of tumors initiated by other oncogenes or inactivation of suppressor genes but is most specific to tumors induced by T/t-antigen. By interrogating the gene expression profile of various mouse mammary tumor models carrying the transgenes MMTV-myc, MMTV-ras, MMTV-her2/neu, MMTV-PyMT, and C3(1)-Tag, and knockout of tumor suppressor genes Wap-cre;p53^{flp/flp} or Brca1^{Co/Co}, and MMTV-Cre; p53^{+/-}, we found that these tumors did not fully express the intrinsic SV40 T/t-antigen proliferation signature, although various degrees of alterations in the expression of some of these genes were present in tumors induced by other oncogenic processes. Gene expression of Wap-cre; Brca1^{Co/Co} tumors showed the greatest similarity to that of the C3(1)-Tag tumors followed by the two targeted p53 knockout models (Wap-cre;p53^{flp/flp} and MMTV-Cre; p53^{+/-}). However, transgenic overexpression of MMTV-myc, MMTV-ras, MMTV-her2/neu, and MMTV-PyMT oncogenes were most dissimilar to the SV40 T/t-antigen proliferation cluster. This analysis shows that the SV40 T/t-antigen proliferation signature represents a distinct state of gene expression related to cell cycle and DNA damage response that is not typically observed in tumors that have been transformed by other molecular mechanisms.

When analyzed in the context of gene interactive networks, 85 of the 120 named genes contained within the SV40 T/t-antigen signature are related through several nodes, including Rb and p53, as expected, because it has been well documented that T-antigen binds to and functionally inactivates both of these critical tumor suppressor genes. Based on the disruption of Rb function in the Tag-transformed cells, the identification of an E2F node and genes related to its function is also consistent with the regulation of E2F by the Rb pathway (6).

Interestingly, analysis of the SV40 T/t-antigen signature also revealed that the tumor suppressor *Brca1* was overexpressed in conjunction with a network of genes related to Brca1 function in all three epithelial tumor types, not only mammary tumors. Dysregulation of Brca1 pathways by SV40 T/t-antigen has not been reported previously. Brca1 was only overexpressed in the three T/t-antigen transgenic mouse models and not in the other mammary tumor models studied. The BRCA1 tumor suppressor gene has many functional roles, including DNA damage response, DNA repair and recombination, cell cycle checkpoint control, protein ubiquitylation, and chromatin remodeling (33). We found that Brca1-interacting genes, such as *Chek1*, *Nek2*, *Rbbp7*, and *Msh6*, fall within the Brca1 cluster, suggesting an integrative role for Brca1 in regulating G₂-M and G₁-S checkpoints. However, it is not clear whether expression of SV40 T/t-antigen oncoprotein directly affects Brca1 overexpression or whether this is a secondary result related to enhanced cell proliferation resulting in an increase in S and G₂ phase genes that could up-regulate *Brca1* expression.

The intrinsic SV40 T/t-antigen signature also includes many other genes associated with DNA replication, damage repair, cytokinesis, and chromosome maintenance. Our analysis revealed the overexpression of a set of genes that included replication-initiation complex minichromosome maintenance proteins (*Mcm*'s) and inhibitor of preinitiation complex assembly, geminin (*Gmnn*), suggesting that their dysregulation leads to inappropriate binding of these proteins to chromatin during the cell cycle (particularly at S, G₂, and early mitosis phases) leading to aberrant DNA replication through illegitimate origin firing. These effects could lead to significant genomic instability, which is characteristic of SV40 T/t-antigen transgenic tumors (9). Meta-analyses have shown a strong association between the overexpression of *Mcm*'s and *Gmnn* with aggressive epithelial tumor behavior and poor cancer prognosis, further suggesting that these cell cycle biomarkers may be potentially important predictors for routine clinical screening.

Contained within the T/t-antigen signature are many overexpressed genes related to proliferation, such as *Ki67* and *Pcna*. Immunohistochemical staining confirmed that Ki67 expression was highest in the T/t-antigen tumors compared with tumors arising in other tumor models. Additionally, many other genes related to replication and proliferation were identified in the T/t-antigen signature, including the cell cycle-regulated genes *Cdc28 protein kinase 2*, *Chek1*, *Cdc2a*, *Ccnd2*, *Ccna2*, and *Ccnf* (genes that encode cyclin D2, cyclin A2, and cyclin F, respectively); chromosomal instability genes, such as *polo-like kinase (Plk4)*, *serine/threonine kinase 12 (Stk12)*, *Bub1b*, and *Mad2 mitotic arrest deficient-like 1 (Mad2l1)*; and transcription regulators *FoxM1* and *Ezh2*. Elevated levels of the FoxM1 transcription factor has been shown to accelerate development and progression of prostate carcinomas in TRAMP/Rosa26-FoxM1b double transgenic mice compared with TRAMP mice (34). Furthermore, the T/t-antigen proliferation cluster includes *Top2a*, dihydrofolate reductase (*Dhfr*), thymidylate synthase (*Tyms*), and ribonucleotide reductase M1 polypeptide (*Rrm1*; Fig. 2, purple), potential targets for drug therapies.

The importance of genes related to cell cycle and proliferation in identifying tumors with poor prognosis has been recognized previously (13, 20, 35). However, only about one third of the genes contained in the intrinsic SV40 T/t-antigen signature have been identified previously as part of a "proliferation cluster" expressed in

¹⁵ K. Deeb et al., in preparation.

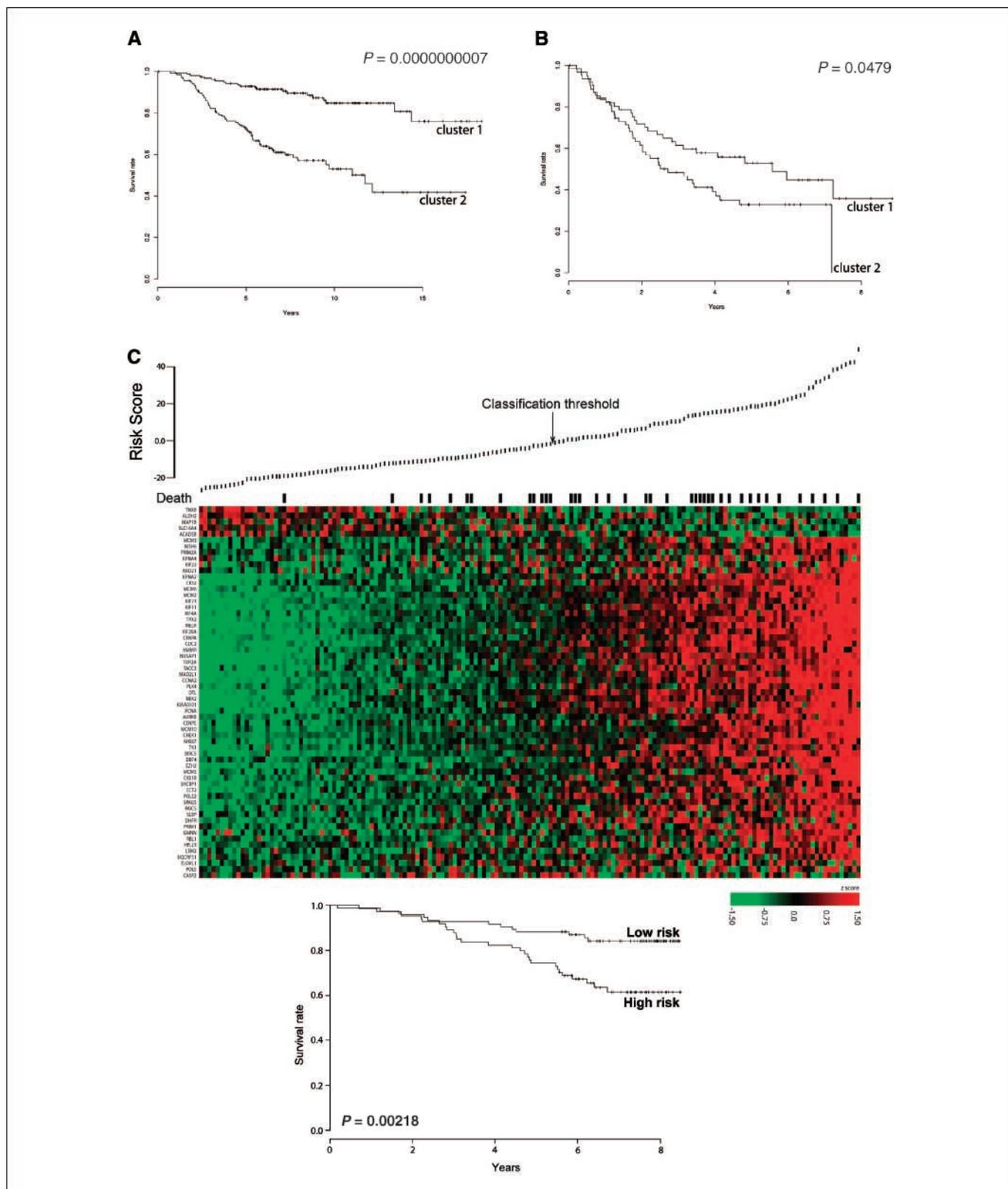


Figure 5. Survival and the T/t-antigen signature in human breast and lung cancer patients. **A**, Kaplan-Meier overall survival curves were calculated for every patient grouped in the cluster 1 or 2 from the Chang breast cancer data set depicted in Fig. 4A. **B**, Kaplan-Meier overall survival curves were calculated only for the lung adenocarcinoma patients from the Bhattacharjee data set grouped in cluster 1 or 2 (survival information is not available for the other lung cancer subtypes) as depicted in Fig. 4B. **C**, prediction of overall survival using the SV40 T/t-antigen signature. The image plot was generated with the external test set of 159 breast cancer patients from Pawitan et al. (18) and 61 survival associated genes selected in the training set as described in Materials and Methods. The samples are ordered by increasing risk scores computed with the compound covariate algorithm. *Black bars*, patient death. The Kaplan-Meier overall survival curves are shown for the high- and low-risk groups identified with the prognostic classifier.

the basal epithelial-like subgroup of breast carcinomas (13; reviewed in ref. 36). Therefore, the T/t-antigen signature identifies a significantly larger gene set that can be integrated into a highly robust genetic network based on known functions. Examples of genes previously not identified include cyclin regulators cyclin kinase subunit 1 and *cdc7/Dbf4*, claspin, a regulator of DNA damage repair, and aurora kinase B, essential for chromosome segregation and cytokinesis.

Other gene signatures have been reported that predict disease outcome. Thirty-seven genes identified in the wound response signature are also contained in the T/t-antigen signature, many of which are related to cell cycle and DNA replication. Eighteen genes in the T/t-antigen signature are found in common with the wound, cell cycle, and proliferation signatures, suggesting that these genes may be most often dysregulated in the most aggressive tumors. Interestingly, only three genes (*Ect2*, *Mcm6*, and *Cenpa*) identified in the 70-gene prognostic signature (30), one gene (*Ki67*) identified in the BMI-1 prognostic signature (37) and one gene (*Tubb*) found in the invasive gene signature (32), were also contained in the T/t-antigen signature. It is not clear exactly why these differences in gene content between the signatures exist, although each signature was identified based on different biological questions.

This study shows an important value of applying genetically engineered mouse models for the study of oncogenic pathways. Variance in gene expression, which is often a limitation of studies using relatively small cohorts of human samples, can be greatly minimized in studies using animals with similar genetic backgrounds (11).¹⁶ This provides a means of enhancing the identification of more robust gene expression signatures that may not be apparent in human data sets due to variations in gene expression resulting from heterogeneous genetic backgrounds. Previous studies have also used mouse models to identify important gene signatures related to particular oncogenic pathways relevant to human cancers, including prostate (37, 38), liver (39), lung (40), and breast (41).

We also did a reverse analysis by clustering tumors from the various mouse mammary cancer models using the "proliferation cluster" reported by Perou et al. (13) based on 49 available genes. This similarly shows that the expression patterns of the C3(1)-T/t-antigen and BRCA1^{-/-};p53^{+/-} compound tumors were the most similar to the expression pattern characteristic of the basal epithelial-like subgroup of human breast cancers (data not shown).

Interestingly, tumors from the T/t-antigen mouse models expressed lower levels of *Gstm1*, a member of a family of enzymes that play a major role in cellular detoxification system and likely protect cells against reactive oxygen metabolites or carcinogens. Lack of *Gstm1* or *GST* polymorphisms have been reported to correlate with an increase susceptibility to lung cancer (42), breast cancer (43), and prostate cancer (44). Similarities shared across tumors from the different oncogene models and p53-knockout models included up-regulation of *Cdc3a* and H2A histone family member z (*H2afz*) and down-regulation of genes, such as *lipoprotein lipase (Lpl)*, *complement component 3 (C3)*, *crystallin μ (Crym)*, and *Cav1*.

Crym and *Cav1* were consistently decreased in all the tumor types as well as in other mouse models of breast cancer. Reduced

Crym expression has been associated with in hormone-refractory prostate cancer and thus may play an important role in clinical progression of the disease (45). These results suggest that the T/t-antigen models may be useful for deciphering the functional properties of the *Crym* gene for growth and survival of tumor cells.

Recently, *Cav1* has emerged as an important gene in oncogenic transformation and may negatively regulate cell proliferation and metastasis in human cancers and mouse models of cancer (reviewed in ref. 46). Because it has been reported that transcription of *Cav1* is regulated by p53 (47), the targeted inactivation of p53 by T-antigen may result in the reduced *Cav1* expression in the multiple tumor models. *Cav1* expression negatively regulates cell cycle progression by inducing G(0)/G(1) arrest via a p53/p21(WAF1/Cip1)-dependent mechanism.

Because the T/t-antigen signature contains many genes controlling key cellular processes related to replication, proliferation, cell cycle regulation, DNA damage repair, and apoptosis, we wished to determine the relevance of this signature to human breast, prostate, and lung cancers. When applied to large expression profile data sets of these human tumors, hierarchical clustering using the T/t-antigen signature clearly distinguished aggressive tumors with poor prognosis from less aggressive tumors. Kaplan-Meier analyses of tumors clustered into groups either exhibiting or not exhibiting the T/t-antigen signature showed a highly significant difference in patient survival. This finding indicates that the genes and integrated network discovered in this study are extremely important in understanding molecular features that distinguish tumors with favorable versus unfavorable prognosis.

The T/t-antigen signature was sufficient to categorize human breast tumors into the basal or luminal types B/C categories (20), which generally are unresponsive to hormone and chemotherapies and have poor prognosis. Similarly, the T/t-antigen signature distinguished metastatic from nonmetastatic prostate tumors. Therefore, because the genetic network represented by this signature is operative in tumors often intractable to existing therapies with the worst prognosis, therapies that target key nodes within this network may be critical to improving survival in patients with tumors expressing the T/t-antigen signature.

The study of various SV40 T/t-antigen transgenic mouse models of cancer has allowed us to identify both common and tissue-specific oncogenic responses to these viral oncoproteins with several important similarities to the corresponding human cancers. This is the first study to show that the SV40 T/t-antigen viral oncoproteins can cause an intrinsic gene expression profile that recapitulates the aggressive phenotypes of aggressive human cancers. Whereas some variation in the complement of genes in the proliferation signature are likely to be found in different tumors, increased expression of the core set of proliferation genes in tumors is often associated with aggressive cancers and poor prognosis in patients with breast cancer (13, 30, 48), lung carcinoma (17, 49), and prostate cancer (16, 50). Genes from the intrinsic SV40 T/t-antigen signature clustered commonly with the aggressive basal-like tumors of the breast, metastatic prostate carcinomas, and small cell and squamous lung carcinomas, further supporting that intrinsic SV40 T/t-antigen gene cluster is representative of human epithelial tumors, which contributes to aggressive clinical pathologies and poor prognosis; thus, these models may be highly relevant for studying those subtypes of

¹⁶ A.M. Michalowska et al., submitted for publication.

human cancers and developing therapeutic agents that could target this proliferation network. Furthermore, identifying the expression changes during the course of tumor evolution in these models could significantly enhance our knowledge of genetic changes during tumor progression to identify potential biomarkers and therapeutic targets.

Acknowledgments

Received 4/25/2007; revised 5/31/2007; accepted 6/20/2007.

Grant support: Intramural Research program of the NIH, NCI, CCR.

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.

We thank Lisa Riffle for excellent assistance with animal handling.

References

- Hanahan D. Heritable formation of pancreatic β -cell tumours in transgenic mice expressing recombinant insulin/simian virus 40 oncogenes. *Nature* 1985;315:115–22.
- Connolly DC, Bao R, Nikitin AY, et al. Female mice chimeric for expression of the simian virus 40 TAG under control of the MISHR promoter develop epithelial ovarian cancer. *Cancer Res* 2003;63:1389–97.
- Magdaleno SM, Wang G, Mireles VL, Ray MK, Finegold MJ, DeMayo FJ. Cyclin-dependent kinase inhibitor expression in pulmonary Clara cells transformed with SV40 large T antigen in transgenic mice. *Cell Growth Differ* 1997;8:145–55.
- Maroulakou IG, Anver M, Garrett L, Green JE. Prostate and mammary adenocarcinoma in transgenic mice carrying a rat C3(1) simian virus 40 large tumor antigen fusion gene. *Proc Natl Acad Sci U S A* 1994;91:11236–40.
- Greenberg NM, DeMayo F, Finegold MJ, et al. Prostate cancer in a transgenic mouse. *Proc Natl Acad Sci U S A* 1995;92:3439–43.
- Ahuja D, Saenz-Robles MT, Pipas JM. SV40 large T antigen targets multiple cellular pathways to elicit cellular transformation. *Oncogene* 2005;24:7729–45.
- Ali SH, DeCaprio JA. Cellular transformation by SV40 large T antigen: interaction with host proteins. *Semin Cancer Biol* 2001;11:15–23.
- Arroyo JD, Hahn WC. Involvement of PP2A in viral and cellular transformation. *Oncogene* 2005;24:7746–55.
- Barkan D, Montagna C, Ried T, Green JE. Mammary gland cancer. In: Holland E, editor. *Mouse models of human cancer*. NY: Wiley; 2004. p. 103–31.
- Hahn WC, Dessain SK, Brooks MW, et al. Enumeration of the simian virus 40 early region elements necessary for human cell transformation. *Mol Cell Biol* 2002;22:2111–23.
- Desai KV, Xiao N, Wang W, et al. Initiating oncogenic event determines gene-expression patterns of human breast cancer models. *Proc Natl Acad Sci U S A* 2002;99:6967–72.
- Chang HY, Sneddon JB, Alizadeh AA, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol* 2004;2:E7.
- Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
- Hegde P, Qi R, Abernathy K, et al. A concise guide to cDNA microarray analysis. *Biotechniques* 2000;29:548–50, 52–4, 56 passim.
- Chang HY, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci U S A* 2005;102:3738–43.
- LaTulippe E, Satagopan J, Smith A, et al. Comprehensive gene expression analysis of prostate cancer reveals distinct transcriptional programs associated with metastatic disease. *Cancer Res* 2002;62:4499–506.
- Bhattacharjee A, Richards WG, Staunton J, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 2001;98:13790–5.
- Pawitan Y, Bjohle J, Amler L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005;7:R953–64.
- Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;4:249–64.
- Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–74.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc* 1995;57:289–300.
- Radmacher MD, McShane LM, Simon R. A paradigm for class prediction using gene expression profiles. *J Comput Biol* 2002;9:505–11.
- Lin SC, Lee KF, Nikitin AY, et al. Somatic mutation of p53 leads to estrogen receptor α -positive and -negative mouse mammary tumors with high frequency of metastasis. *Cancer Res* 2004;64:3525–32.
- Xu X, Wagner KU, Larson D, et al. Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. *Nat Genet* 1999;22:37–43.
- Bouchard L, Lamarre L, Tremblay PJ, Jolicoeur P. Stochastic appearance of mammary tumors in transgenic mice carrying the MMTV/c-neu oncogene. *Cell* 1989;57:931–6.
- Tremblay PJ, Pothier F, Hoang T, et al. Transgenic mice carrying the mouse mammary tumor virus ras fusion gene: distinct effects in various tissues. *Mol Cell Biol* 1989;9:854–9.
- Leder A, Pattengale PK, Kuo A, Stewart TA, Leder P. Consequences of widespread deregulation of the c-myc gene in transgenic mice: multiple neoplasms and normal development. *Cell* 1986;45:485–95.
- Guy CT, Cardiff RD, Muller WJ. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol Cell Biol* 1992;12:954–61.
- Shibata MA, Maroulakou IG, Jorcyk CL, Gold LG, Ward JM, Green JE. p53-independent apoptosis during mammary tumor progression in C3(1)/SV40 large T antigen transgenic mice: suppression of apoptosis during the transition from preneoplasia to carcinoma. *Cancer Res* 1996;56:2998–3003.
- van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530–6.
- Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. *Nat Genet* 2003;33:49–54.
- Liu R, Wang X, Chen GY, et al. The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* 2007;356:217–26.
- Deng CX. BRCA1: cell cycle checkpoint, genetic instability, DNA damage response, and cancer evolution. *Nucleic Acids Res* 2006;34:1416–26.
- Kalin TV, Wang IC, Ackerson TJ, et al. Increased levels of the FoxM1 transcription factor accelerate development and progression of prostate carcinomas in both TRAMP and LADY transgenic mice. *Cancer Res* 2006;66:1712–20.
- Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003;100:8418–23.
- Whitfield ML, George LK, Grant GD, Perou CM. Common markers of proliferation. *Nat Rev Cancer* 2006;6:99–106.
- Glinsky GV, Berezovska O, Glinskii AB. Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. *J Clin Invest* 2005;115:1503–21.
- Ellwood-Yen K, Graeber TG, Wongvipat J, et al. Myc-driven murine prostate cancer shares molecular features with human prostate tumors. *Cancer Cell* 2003;4:223–38.
- Kaposi-Novak P, Lee JS, Gomez-Quiroz L, Coulouarn C, Factor VM, Thorgerisson SS. Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. *J Clin Invest* 2006;116:1582–95.
- Sweet-Cordero A, Mukherjee S, Subramanian A, et al. An oncogenic KRAS2 expression signature identified by cross-species gene-expression analysis. *Nat Genet* 2005;37:48–55.
- Qiu TH, Chandramouli GV, Hunter KW, Alkharouf NW, Green JE, Liu ET. Global expression profiling identifies signatures of tumor virulence in MMTV-PyMT-transgenic mice: correlation to human disease. *Cancer Res* 2004;64:5973–81.
- Nakajima T, Elovaara E, Anttila S, et al. Expression and polymorphism of glutathione S-transferase in human lungs: risk factors in smoking-related lung cancer. *Carcinogenesis* 1995;16:707–11.
- Gudmundsdottir K, Tryggvadottir L, Eyfjord JE. GSTM1, GSTT1, and GSTP1 genotypes in relation to breast cancer risk and frequency of mutations in the p53 gene. *Cancer Epidemiol Biomarkers Prev* 2001;10:1169–73.
- Acevedo C, Opazo JL, Huidobro C, Cabezas J, Iturrieta J, Quinones Sepulveda L. Positive correlation between single or combined genotypes of CYP1A1 and GSTM1 in relation to prostate cancer in Chilean people. *Prostate* 2003;57:111–7.
- Mousses S, Bubendorf L, Wagner U, et al. Clinical validation of candidate genes associated with prostate cancer progression in the CWR22 model system using tissue microarrays. *Cancer Res* 2002;62:1256–60.
- Williams TM, Lisanti MP. Caveolin-1 in oncogenic transformation, cancer, and metastasis. *Am J Physiol Cell Physiol* 2005;288:C494–506.
- Bist A, Fielding CJ, Fielding PE. p53 regulates caveolin gene transcription, cell cholesterol, and growth by a novel mechanism. *Biochemistry* 2000;39:1966–72.
- 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 U S A* 1999;96:9212–7.
- Garber ME, Troyanskaya OG, Schluens K, et al. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A* 2001;98:13784–9.
- Lapointe J, Li C, Higgins JP, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci U S A* 2004;101:811–6.

Correction: T/t-Antigen Signature in Aggressive Human Cancers

In the article on T/t-antigen signature in aggressive human cancers in the September 1, 2007 issue of *Cancer Research* (1), the callout for Supplementary Table S1 in the second paragraph of page 8073 should be Supplementary Table S4.

1. Deeb KK, Michalowska AM, Yoon C-y, Krummey SM, Hoenerhoff MJ, Kavanaugh C, Li M-c, Demayo FJ, Linnoila I, Deng C-x, Lee EY-HP, Medina D, Shih JH, Green JE. Identification of an integrated SV40 T/t-antigen cancer signature in aggressive human breast, prostate, and lung carcinomas with poor prognosis. *Cancer Res* 2007; 67:8065–80.

Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

Identification of an Integrated SV40 T/t-Antigen Cancer Signature in Aggressive Human Breast, Prostate, and Lung Carcinomas with Poor Prognosis

Kristin K. Deeb, Aleksandra M. Michalowska, Cheol-yong Yoon, et al.

Cancer Res 2007;67:8065-8080.

Updated version	Access the most recent version of this article at: http://cancerres.aacrjournals.org/content/67/17/8065
Supplementary Material	Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2007/09/12/67.17.8065.DC1

Cited articles	This article cites 49 articles, 26 of which you can access for free at: http://cancerres.aacrjournals.org/content/67/17/8065.full.html#ref-list-1
Citing articles	This article has been cited by 17 HighWire-hosted articles. Access the articles at: /content/67/17/8065.full.html#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org .
Permissions	To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org .