[CANCER RESEARCH 64, 5973–5981, September 1, 2004]

Global Expression Profiling Identifies Signatures of Tumor Virulence in MMTV-PyMT-Transgenic Mice: Correlation to Human Disease

Ting Hu Qiu,[1,2] Gadisetti V. R. Chandramouli,[2,3] Kent W. Hunter,[4] Nawal W. Alkharouf,[2] Jeffrey E. Green,[1] and Edison T. Liu[2,5]

[1]Laboratory of Cell Regulation and Carcinogenesis, [2]Molecular Signaling and Oncogenesis Section, Department of Cancer and Cell Biology, [3]Laboratory of Systems Biology, and [4]Laboratory of Population Genetics, Cancer Research Center, National Cancer Institute, Bethesda, Maryland; and [5]Genome Institute of Singapore, Singapore

ABSTRACT

FVB/N-Tg (MMTV-PyMT)^{634Mul}-transgenic mice develop multifocal mammary tumors with a high incidence of pulmonary metastasis. We have demonstrated previously that mammary tumors derived from transgene-positive F1 progeny in particular inbred strains display altered latency, tumor growth rates, and metastatic rates when compared with the FVB/NJ homozygous parent. To identify genes with expression that might be critical in modifying the biological behavior of MMTV-PyMT tumors, we performed a detailed comparative analysis of expression profiles from mammary tumors arising in the parental FVB/NJ background and F1 progeny from crosses with I/LnJ, LP/J, MOLF/Ei, and NZB/B1NJ mice. Compared with normal mammary glands, gene expression profiles of tumors from all five strains exhibited up-regulation of genes involved in cell growth (*e.g.*, Cks1 and CDC25C) and down-regulation of cell adhesion molecules, with many genes associated previously with human breast cancer such as STAT2, CD24 antigen, gelsolin, and lipocalin2. To identify genes with significant variation in expression between the five different genotypes, significance analysis of microarrays (SAM) and one-way ANOVA were used. Three definable groupings of tumors were identified: (*a*) tumors derived in the LP/J F1 and MOLF/Ei F1 strains in which tumor growth and dissemination are suppressed and latency prolonged; (*b*) the most aggressive tumors from the FVB/NJ parental strain and I/LnJ F1 genomic backgrounds; and (*c*) an intermediate virulence phenotype with tumors from NZB/B1NJ–F1 crosses. These array based assessments correlated well with a composite phenotype ranking using a "virulence" index. The gene expression signature that is associated with a high metastatic rate in the mouse contains the same 17 genes described recently as the signature gene set predictive of metastasis in human tumors (1) with 16 of the 17 genes exhibiting the same directional change in expression associated with human metastases. These results demonstrate that the genetic analysis of mouse models of tumorigenesis may be highly relevant to human cancer and that the metastatic phenotype of a tumor may be affected by the germline genetic configuration of the host.

INTRODUCTION

FVB/N-Tg (MMTV-PyMT)^{634Mul}-transgenic mice express the mouse polyomavirus middle-T antigen (PyMT) under the control of the mouse mammary tumor virus (MMTV) long terminal repeat. These transgenic animals uniformly develop multifocal mammary tumors with a high incidence of pulmonary metastasis (2). Mammary transformation by PyMT appears to be mediated through the Src family of tyrosine kinases. PyMT associates with the Src family members Src, Fyn, and Yes to form stable complexes resulting in an increase in the specific activities of these Src family kinases (3–5). Activation of Src alone, however, does not appear sufficient for the rapid induction of mammary tumors, because expression of a constitutively active version of Src in mammary glands of transgenic mice results in the development of mammary epithelial hyperplasia that rarely progresses to invasive cancer (6). Thus, PyMT must recruit additional cellular signal pathways to effect the malignant transformation of the mammary epithelial cell.

In addition to Src, PyMT also acts through the phosphatidylinositol 3′-kinase (PI3K; refs. 7–9) which subsequently activates *Akt* (also named protein kinase B; ref.10) thus inhibiting proapoptotic proteins such as BAD, Forkhead transcription factor, and caspase 9 (11, 12). Lastly, PyMT is also involved in the activation of the Ras signaling pathway in mammary tumorigenesis (13–16).

Using the MMTV-PyMT mouse mammary carcinoma model, we previously identified mouse strains that harbored dominant genetic modifiers of metastasis by breeding MMTV-PyMT mice to 27 different inbred strains (17). Mammary tumors derived from the different transgene-positive F1 progeny displayed altered latency, metastasis, and tumor growth rates compared with the FVB/NJ homozygous parent (17). Our observations raised the intriguing possibility that the host genetic background has a significant impact on the biological behavior of subsequent tumors induced by a single oncogene. In this study, we sought to identify molecular signatures that might be involved with different tumor phenotypes given an identical genetic initiator of tumorigenesis. We compared gene expression profiles of MMTV-PyMT mammary tumors that developed in five different mouse genetic backgrounds: MMTV-PyMT in FVB/NJ and four F1 strain combinations (I/LnJ, LP/J, MOLF/Ei, and NZB/B1NJ).

Although the overall expression profiles from PyMT tumors emerging in various strain backgrounds are similar, a supervised approach identified approximately 200 genes that are differentially expressed between tumors from different genetic backgrounds. Using these subsets of genes that differentiate the genotypes, distinct gene expression patterns could separate the tumors from the different genetic backgrounds into two major clusters that correlated with a general virulence index of the tumors. The behavior of a core set of these virulence-associated genes match those identified previously as belonging to the metastatic signature set of genes for human cancer (1). These results suggest that host genetic factors can independently alter the transcriptional profiles of mammary cancers and may explain some of the differences in the virulence of the resultant tumors.

MATERIALS AND METHODS

Animals

FVB/N-TgN (MMTVPyMT)^{634Mul}-transgenic mice have been described previously and were obtained from the laboratory of W. Muller (2). Tumors initiated by MMTV-PyMT in the original FVB/NJ strain and four F1 strain combinations, including I/LnJ, LP/J, MOLF/Ei, and NZB/B1NJ, were selected for study. Tumor samples were generated as described previously (17). Mammary tumors were detected by palpation. After the initial identification of the primary tumor, the animal was aged for 40 days post-diagnosis to permit tumor development and potential metastatic dissemination. After 40 days, the animals were sacrificed by carbon dioxide inhalation. The details of the tumor growth and metastatic patterns in the parent FVB/Nj strain and in each F1 are described in Table 1. The overall

Received 1/24/04; revised 4/27/04; accepted 7/6/04.

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

Note: T. H. Qiu and G. V. R. Chandramouli contributed equally to this work. Supplemental data for this article can be found at Cancer Research Online (http://cancerres@aacrjournals.org).

Requests for reprints: Jeffrey E. Green, Laboratory of Cell Regulation and Carcinogenesis, Cancer Research Center, National Cancer Institute, Bethesda, Maryland 20892. E-mail: JEGreen@nih.gov.

Table 1 *Virulence index*

Strain	Latency*		Metaststic density		Tumor growth		Average virulence index†
	Days	(Rank)‡	No. of mets./area	(Rank)	Relative speed	(Rank)	
FVB/Nj	58	(3)	1.00	(1)	Fast	(1)	1.67
I/LnJ F1	37	(1)	0.08	(3)	Slow	(3)	2.33
NZB/B1Nj F1	60	(4)	0.07	(4)	Fast	(1)	3.00
MOLF/Ei F1	81	(5)	0.20	(2)	Slow	(3)	3.33
LP/j F1	51	(2)	0.00	(5)	Veryslow	(5)	4.00

* Latency is defined by the number of days after birth at diagnosis of first palpable tumor. Metastatic density is defined as the number of mets per square micron of pulmonary tissue, normalized to make the metastatic density of FVB/NJ equal to 1 (based upon data from refs. 17 and 33). Tumor growth is defined by the average change in tumor mass per day, based on change in tumor mass over forty day period.

† The virulence index for each parameter is the average of the relative rankings of latency, growth, and metastatic density (1–5, in parentheses) for each genetic background.

‡ Values in parentheses represent rank order for each phenotypic category.

behavior of the resultant tumors is based on three observed phenotypes: growth rate, the latency of tumor onset, and the metastatic potential (as measured by the number of tumors at the time of sacrifice). Because there was no definable match between the individual virulence phenotypes and each genetic background, we sought to derive a unified index of tumor virulence that would take into account these heterogenous biological metrics. Therefore, we derived an overall estimate of tumor virulence in the following manner. We ranked each phenotype (latency of tumor onset, metastatic density, and rate of tumor growth) for the tumors arising from each of the F1 crosses. The average of the three rankings for each genotype provided a qualitative measure of tumor virulence. The ranking from most virulent to least virulent appeared to be as follows: FVB/NJ > I/LnJ F1 > NZB/B1NJ F1 > MOLF/Ei F1 > LP/J F1. Mammary glands were collected from 10 to 11 week-old wild-type FVB/NJ female mice to generate normal mammary reference RNA (Charles River, Frederick, MD). All mice were housed and cared for in accordance with guidelines set up by the NIH.

cDNA Array Hybridization

Four tumor samples from the FVB/N-Tg (MMTV-PyMT)^{634Mul} and five tumors samples from each of the four F1 strain combinations were snap frozen and stored at −80°C for RNA isolation. Total RNA was extracted from frozen tissue using TRIzol Reagent (Life Technologies, Inc., Rockville, MD) according to the manufacturer's protocol. Reference RNA was extracted and pooled from 10 to 11 week-old FVB/NJ virgin mammary glands. Three micrograms of total RNA from the reference and tumor samples were amplified using the modified Eberwine method (18). Ten micrograms of linearly amplified RNA were used to generate Cy3-dUTP- or Cy5-dUTP- (Amersham Pharmacia Biotech) labeled first-strand cDNA by reverse transcription using random primers. The cDNA products synthesized from the samples and reference were hydrolyzed with NaOH and then purified with microcon YM-30 columns (Millipore Corp., Bedford, MA). Each tumor sample was reciprocally labeled with Cy3-dUTP or Cy5-dUTP and hybridized to microarrays. A total of 48 cDNA arrays were used in this study. cDNA clones (Incyte GEM1 set, ~8,700 elements) were purchased from Incyte Genomics Inc. (Wilmington, DE) and spotted on poly-lysine-coated glass slides by the microarray core at the National Cancer Institute Advanced Technology Center.

Microarray hybridizations were performed as described previously (19). The arrays were air-dried and scanned using the Axon GenePix4000A scanner (Axon, Union City, CA) and images were processed using GenePix-Pro3.0 software (Axon). Both image and signal intensity data were stored in the microarray database supported by the Center for Information Technology at the National Cancer Institute, NIH, Bethesda.[6] This data are available through the National Cancer Institute Director's Challenge website.[7]

Oligonucleotide Microarray Hybridization and Data Analysis

To extend the gene coverage of this analysis, we reanalyzed the tumors using a different array platform. The same preparation of total RNA used for cDNA arrays was used for oligonucleotide array studies on the Murine Genome U74Av2 chip containing 12,500 features (Affymetrix, Inc., Santa Clara, CA). Equal amounts of RNA isolated from each genomic background were pooled. Ten micrograms of total RNA were reverse transcribed using a T7 $(dThd)_{24}$ primer to synthesize the cDNA, followed by the incorporation of biotinylated ribonucleotides by *in vitro* transcription using T7 RNA polymerase. The biotinylated-labeled RNA was fragmented and hybridized to the Murine Genome U74Av2 chip (Affymetrix) according to the manufacturer's recommendations. Each mixed RNA sample was repeated twice by two independent labeling and hybridizations. Signal intensities for different chips were scaled to a trimmed mean signal of 150. Gene expression analysis was performed using the Microarray Suite version 5.0 software (Affymetrix) where pooled 10 to 11 week-old FVB/Nj mammary gland RNA served as the reference.

Gene Expression Analysis

cDNA Arrays. The mean spot intensities corrected for local background were used in calculating the expression ratio (Cy5 tumor: Cy3 reference and the converse for dye swap) for each clone (excluding the spots that were flagged as bad or not found during the image analysis). The ratios were then normalized to a median ratio value of 1 for each array, log transformed, and averaged for replicated dye-swap experiments. The resulting data matrix used for analysis consisted of the expression levels of 8,464 clones relative to the normal mammary gland. Two-tailed F tests and other statistical calculations were performed on log-transformed ratios. An overall error level of the averaged expression ratios was determined by computing the SE of each averaged expression and the 99th percentile of these SEs. This estimated value corresponds to a ratio of 1.43. This threshold level was set as a minimum requirement for the significance of expression change in addition to the *P* value of a statistical test.

Gene Expression in Tumor Tissue. The expressions were considered to be significant when the Student *t* test *P* value is <0.01 and 75% percent of replicates have a 2-fold change in the same direction. The number of genes up- and down-regulated in each strain were calculated, and the intersect of all five strains was determined.

Significance Analysis of Microarrays. The program SAM[8] was used to identify genes that are likely to exhibit a significant change in expression among the tumor groups. The relative difference in expression of gene *i* is defined by a score $d(i) = r(i)/[s(i) + s_0]$, where r(i) is Fisher's linear discriminant, s(i) is a measure of SD, and s_0 is a small positive constant (20). A set of 125 significant genes having d(i) scores above the threshold (Δ)level of 0.1 from expected values was determined as described by Tusher *et al.* (20). At this threshold level, the false discovery rate was estimated to be 0.8%.

Multidimensional Scaling and Hierarchical Clustering. Multivariate analysis of the correlations among gene distributions of samples was carried out by multidimensional scaling. The multidimensional information was reduced to three dimensions using $1-\rho$ as the distance metric where ρ is the Pearson correlation coefficient for pictorial representation. The analysis was carried out using all genes as well as a subset of genes having $P < 0.01$ by one-way ANOVA calculation. The calculations were performed using Partekpro 5.0 software (Partek Inc., St. Charles, MO). Hierarchical clustering of both genes and samples was performed using the Eisen Cluster program (21) with $1-\rho$ as the distance metric and the complete linkage option of the software for agglomerative clustering.

[6] http://nciarray.nci.nih.gov.

[7] URL: http://dc.nci.nih.gov/.

[8] http://www-stat-class.stanford.edu/SAM/SAMServlet.

Table 2 *Intersections (∩) of genes up- and down-regulated in tumors according to strains of mice*

Strain no.	Group symbol*	Number of up-regulated genes	Number of down-regulated genes
1	FVBN/NJ only	65	70
2	I/LnJ only	7	18
3	FVBN/NJ ∩ I/LnJ	26	21
4	LP/J only	19	62
5	FVBN/NJ ∩ LP/J	8	12
6	I/LN ∩ JLP/J	1	4
7	FVBN/NJ ∩ I/LnJ ∩ LP/J	7	11
8	MOLF/Ei only	17	17
9	FVBN/NJ ∩ MOLF/Ei	2	4
10	I/LnJ ∩ MOLF/Ei	3	1
11	FVBN/NJ ∩ I/LnJ ∩ MOLF/Ei	6	6
12	LP/J ∩ MOLF/Ei	4	32
13	FVBN/NJ ∩ LP/J ∩ MOLF/Ei	1	6
14	I/LnJ ∩ LP/J ∩ MOLF/Ei	1	6
15	FVBN/NJ ∩ I/LnJ ∩ LP/J ∩ MOLF/Ei	4	16
16	NZB/B1NJ only	35	48
17	FVBN/NJ ∩ NZB/B1NJ	19	23
18	I/LnJ ∩ NZB/B1NJ	5	19
19	FVBN/NJ ∩ I/LnJ ∩ NZB/B1NJ	21	58
20	LP/J ∩ NZB/B1NJ	4	12
21	FVBN/NJ ∩ LP/J ∩ NZB/B1NJ	9	7
22	I/LnJ ∩ LP/J ∩ NZB/B1NJ	2	7
23	FVBN/NJ ∩ I/LnJ ∩ LP/J ∩ NZB/B1NJ	16	60
24	MOLF/Ei ∩ NZB/B1NJ	3	5
25	FVBN/NJ ∩ MOLF/Ei ∩ NZB/B1NJ	3	2
26	I/LnJ ∩ MOLF/Ei ∩ NZB/B1NJ	3	1
27	FVBN/NJ ∩ I/LnJ ∩ MOLF/Ei ∩ NZB/B1NJ	8	16
28	LP/J ∩ MOLF/Ei ∩ NZB/B1NJ	3	19
29	FVBN/NJ ∩ LP/J ∩ MOLF/Ei ∩ NZB/B1NJ	6	18
30	I/LnJ ∩ LP/J ∩ MOLF/Ei ∩ NZB/B1NJ	4	10
31	FVBN/NJ ∩ I/LnJ ∩ LP/J ∩ MOLF/Ei ∩ NZB/B1NJ	49	116
	Total	361	707

* Virulence factor (as defined in Table 1) FVB/Nj > I/LnJ F1 > NZB/B1Nj F1 > MOLF/Ei F1 > LP/j F1.

RESULTS

Identification of Cancer-Related Genes

The global differential gene expression profiles of MMTV-PyMT tumors compared with normal mammary gland tissue were determined by identification of genes with expression that was altered by 2-fold in at least 75% of the replicates with a Student *t* test value of $P < 0.01$. A total of 1,068 genes meeting these criteria were found in the PyMT tumors from the five mouse strains of which 361 were up-regulated and 707 were down-regulated. Many of these genes were common to more than one strain, as summarized in Table 2. To define a minimal set of genes with expression levels that could distinguish the mammary tumors from the normal mammary gland, we selected those genes that were significantly dysregulated in all tumors. This resulted in a common gene set of 164 genes. Forty-eight genes were up-regulated and 116 genes were down-regulated (Supplementary Table 1*A*). This set of genes appears to differentiate tumor from normal gland in all host genetic backgrounds.

Validation of Changes in Gene Expression in Mammary Transformation

To validate this PyMT tumor-related gene set, we compared our results with those of transgenic mammary tumors generated from the overexpression of six different transgenes (22). This previous study identified 624 genes that are differentially expressed between tumors from the various transgenic mouse mammary tumors and normal virgin mammary gland (ref. 22; referred to as the Desai list). Because we used the same cDNA array (containing 8,700 features of the Incyte GEM1 collection) and the same normal virgin mammary gland as a reference, a direct comparison between the datasets could be performed. The only difference is that we used the linear amplified RNA in this study whereas the study by Desai *et al.* used total RNA. Gratifyingly, 83% of the 624 genes in the Desai *et al.* gene list were altered in the same manner despite differences in the method of quantitation. A scatter plot of the correlation between our current results and those of Desai *et al.* using the 624 gene list showed an excellent correlation ($r = 0.89$). Using the most stringent criteria, we compared our minimal list of 164 cancer-associated genes common to all strains with the 624 genes from this previous transformation-specific list and found 63 overlaps (Supplementary Table 1*B*). As expected, the correlation in expression between our data and those of Desai *et al.* using this 63 intersect gene list was extremely high ($r = 0.95$; Fig. 1). The comparison results also indicate that the gene expression data generated from either total RNA or linear amplified RNA are highly consistent, at least in our hands.

Gene Expression Profiles Generated by Oligonucleotide Arrays

The oligonucleotide array (Affymetrix) expression profiles of tumors from each strain were compared with the expression profiles of pooled normal mammary gland RNA using Affymetrix Microarray Suite 5.0 software. Our decision to use both Affymetrix and cDNA arrays was based on the fact that the genes represented in the two arrays only partially overlapped. The two platforms together would therefore cover the largest range of probes. Moreover, with the overlapping gene sets, we could assess the overall reproducibility of these experiments. The number of genes consistently up- or down-regulated associated with the cancer state in all strains using a 2-fold threshold for change in expression was found to be 99 and 89, respectively. These include 30 genes also identified by cDNA arrays (Supplementary Table 1*C*). The direction of expression of these genes was concordant within all strains as also noted in the cDNA array measurements. Because the oligonucleotide and cDNA data provided gene coverage of approximately 16,000 genetic elements, we developed an integration approach to combine both datasets for further analysis.

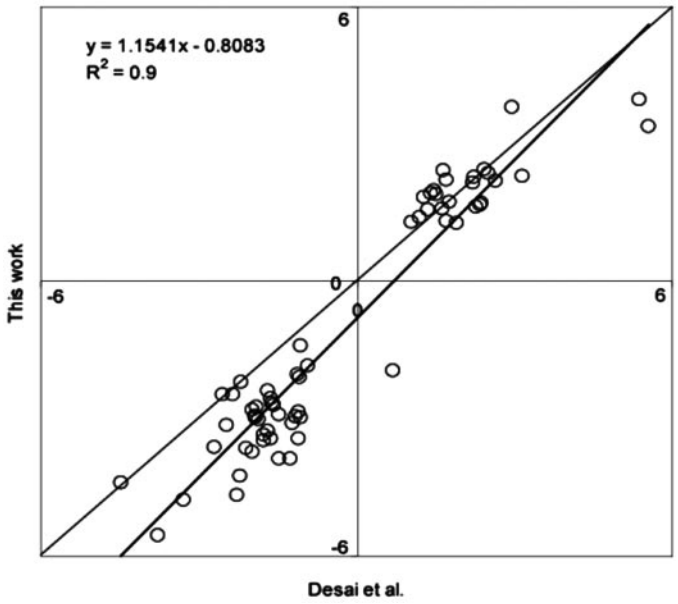

Fig. 1. Correlation of gene expression between data published by Desai *et al.* (22) and data from this study. Representative scatter plot shows the correlation between the two independent cDNA array experiments. The $\log_2$ (ratio) of 63 genes with 2-fold change in expression in tumors from all five strains examined in this study in common with those of Desai *et al.* (22) were used in this analysis.

Gene Expression Data Integration

Concordance between the Array Platforms. The overlap of genes represented on both the oligonucleotide and cDNA arrays is only about 27% of the 16,000 total features represented on both arrays. The correlation between the overlapping genes in the data sets from the cDNA and oligonucleotide arrays was excellent with $r = 0.9$ indicating that the data obtained by two independent methods are very consistent. Because 73% of all genes represented in both array formats were non-overlapping, the combination of these data points would not adversely bias the results of clustering. The key challenge was the harmonization of expression data of genes that were represented in both oligonucleotide and cDNA arrays. To accomplish this, we took the overlapping genes and modeled their association. From the overlapping genes, 3,219 average gene expressions above 1.5 threshold in both oligonucleotide and cDNA arrays of all strains were pooled to determine the degree of correlation. The regression line could not be determined accurately in the presence of significant noise because the error in the estimation of the slope is large. Two slopes, m_1 and m_2, were determined by fitting these average expressions into the equations $y = m_1 x$ and $x = m_2 y$ (where x and y represent oligonucleotide and cDNA expressions) and by checking for the equality of m_1 and $1/m_2$. Elimination of 33% noisy data estimated from residuals of trial fit resulted in good agreement of m_1 with $1/m_2$. The best fit was thus depicted by the eq. A.

$$\log_2 (\text{cDNA ratio}) = 0.865 \log_2 (\text{oligonucleotide ratio}; 1). \quad (A)$$

With this conversion, an additional 158 genes from oligonucleotide arrays could be used as a second gene set in determining the pathways.

The genes common to all five strains of MMTV-PyMT mammary tumor (164 genes identified from the cDNA arrays and 188 genes identified using the oligonucleotide arrays with 30 overlaps) are most likely representative of molecular features associated with PyMT mammary tumors. Included in this list are genes associated previously with human breast cancer, specifically, STAT3, CD24 antigen, lipocalin2, and procollagen Iα and IIIα. In addition, molecules associated with cellular growth, such as cyclin-dependent protein kinase CDC28 (CKS1), hairy (HEY-1), lactotransferrin, and transferrin were up-regulated in tumors. In contrast, cell adhesion molecules or membrane proteins such as junction cell adhesion molecule 1, integral membrane protein 2A, lipocalin2, and procollagen Iα and IIIα, caveolin, CD34, CD36 were all down-regulated. Taken together, this suggests that the cancer cells express a cassette of genes supporting cell growth and reducing cell adhesion. Transcription factors and regulators of cell metabolism were highly induced in the PyMT-induced tumor models. The transcription factors included hairy/enhancer-of-split related (hey-1), Cbp/p300-interacting Transactivator (*Cited2*), activating transcription factor 3 (*Atf3*), transcription factor AP-2β (*Tcfap2b*), CCAAT/enhancer-binding protein (C/EBP) β (*Cebpb*), ets variant gene1 (*Etv1*), and forkhead box C1 (*Foxc1*). RAB3D, a member of the Ras oncogene family (*Rab3d*), was also induced.

Although there did not appear to be specific signatures of signaling pathway activation, several important molecules involved in the PI3K and Ras signaling pathways were differentially regulated. The p85 regulatory subunit of PI3K was up-regulated in mammary tumors from all five strains biologically consistent with the involvement of PI3K in PyMT action. The expression of several Ras signaling-related molecules was also altered tumors from the different background strains as shown in Table 3.

Distinct Tumor Expression Profiles According to Strain Background. The expression data were further analyzed to identify genes that might be associated with the different tumor phenotypes and with expression that varied according to the strain background. Of the set of 1,068 genes showing significant differential expression in any single strain, only 164 were common to all strains. SAM identified 125 significant genes at δ of 0.1 with an estimated false detection rate of 0.8%. Analysis with the F test at significant $P < 0.01$ indicated 80

Table 3 *Ras signaling-related molecules differentially expressed among the five different strain backgrounds*

Gene	Up-/down-regulation*				
	FVB/NJ	I/LnJ F1	NZB/B1NJ F1	MOLF/Ei F1	LP/J F1
RAB3D (member RAS oncogene family)	NC	Up	NC	NC	NC
RAB6 (member RAS oncogene family)	NC	Up	NC	NC	Up
G3bp-pending (Ras-GTPase-activating protein SH3-domain binding protein)	NC	NC	Up	NC	NC
Rassf3 (Ras association domain family 3 protein)	Up	Up	Up	NC	Up
Arhu (ras homology gene family, member U)	Up	NC	Up	NC	Up
Spec1-pending (small protein effector 1 of Cdc42)	NC	NC	NC	Up	NC
Arhgef6 [Rac/Cdc42 guanine nucleotide exchange factor (GEF) 6]	Up	Up	Up	NC	NC
Stk39 [serine/threonine kinase 39, STE20/SPS1 homolog (yeast)]	NC	Up	Up	NC	Up
Caspase 1	Up	NC	NC	NC	NC
Tnfrsf19 (tumor necrosis factor receptor superfamily, member 19)	Up	NC	Up	NC	NC
Fos (general c-fos Mm.5043 FBJ osteosarcoma oncogene)	Up	NC	NC	NC	NC
Rrad (Ras-related associated with diabetes)	Down	NC	NC	NC	NC
Rasa3 (RAS p21 protein activator 3)	Down	Down	Down	NC	NC

* NC, no call.

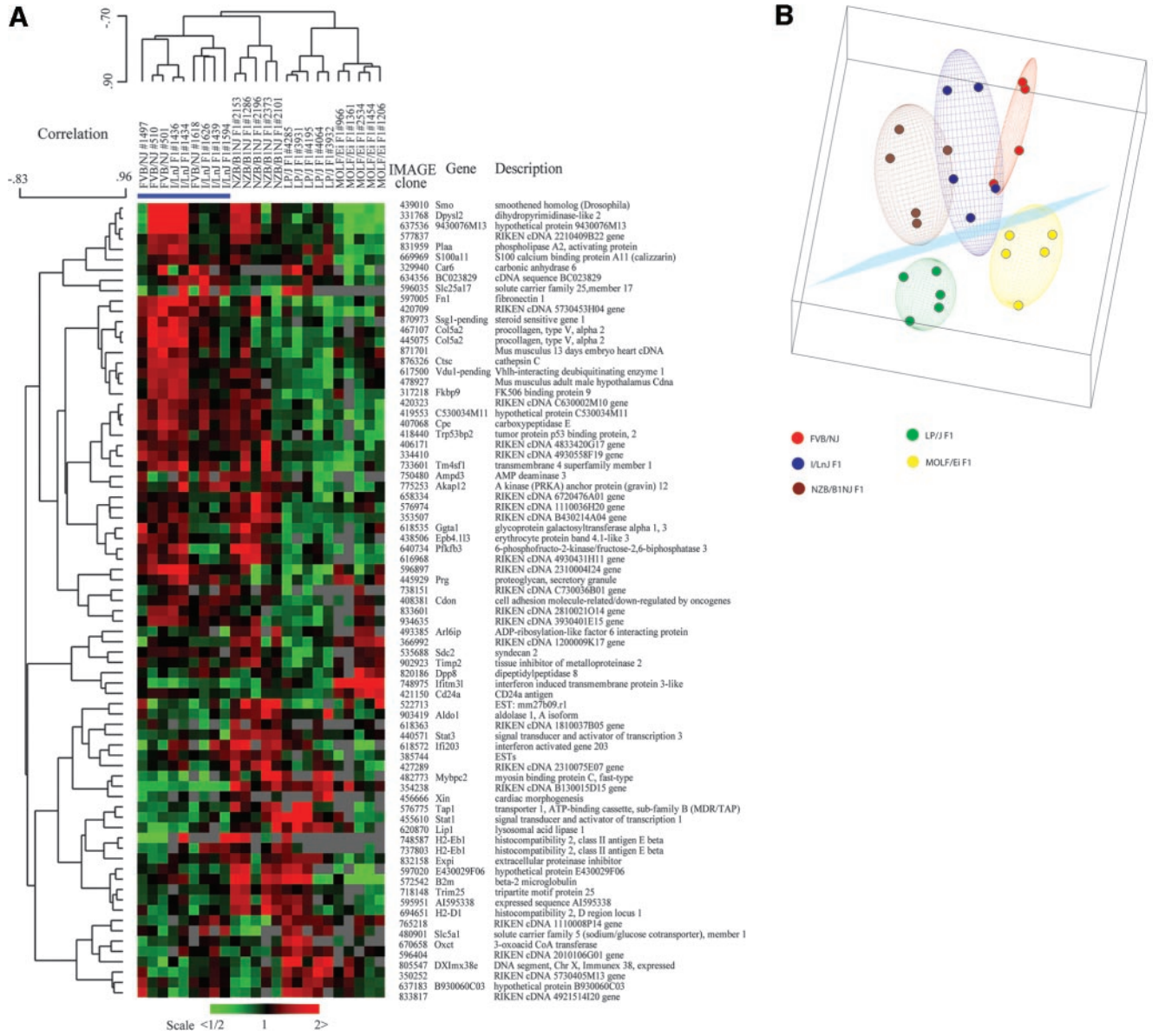

Fig. 2. *A*, hierarchical clustering of 80 genes determined by F test ($P < 0.01$) with at least a 2-fold difference between any two strains. *Blue line* indicates the most aggressive tumors in mouse background strains determined by virulence index. *B*, multidimensional scaling using a subset of genes determined by F test ($P < 0.01$) with a minimum of a 1.5-fold change in expression (275 genes). The distance metric used is 1-Pearson correlation. The *blue plane* represents the cluster separation identified in Fig. 2*A*.

genes with a 2-fold differential expression and 275 genes at a threshold of 1.5-fold (Supplementary Table 2). Using this set of 80 differentially expressed genes (2-fold, $P < 0.01$) selected by this stringent criteria, we applied agglomerative clustering (Fig. 2*A*) and multidimensional scaling (Fig. 2*B*) analyses in an unsupervised manner on the tumor expression data. This revealed that the tumors consistently separated into two major clusters according to the background strain: MOLF/Ei F1, LP/J F1 *versus* NZB/B1NJ F1, FVB/NJ, and I/LNJ F1, with NZB/B1NJ F1 tumors showing greatest overlap similarities between the two clusters. This separation by background strain correlated to the phenotypic behavior of the tumors (Table 1). The cluster shown in Fig. 2*A* identifies genes significantly up or down-regulated in a specific strain compared with other strains (*e.g.,* ALDO1, STAT3, IFI203 are up-regulated in NZB/B1NJ whereas these are down-regulated in all others).

To verify this clustering, we performed an analysis of the expression data from Affymetrix arrays. The pooled RNA samples of each strain were measured in duplicate, and the genes with a consistent 2-fold differential expression from normal mammary gland tissue in any strain were used for clustering and multidimensional scaling studies. For this analysis, we used only those genes with the highest variance and significant differential expression but not overlapping with the genes present in the cDNA array. In this manner, only new genes not studied with the cDNA arrays were analyzed. This filter identified a set of 118 genes with the highest variance present only in the Affymetrix arrays. Hierarchical clustering revealed the presence of two major tumor groups: those arising from MOLF/Ei F1, LP/J F1, and NZB/B1NJ F1 backgrounds *versus* those from FVB/NJ and I/LnJ F1 strains (Fig. 3).

Taken together, the two array studies clearly demarcated tumors arising from FVB/NJ and I/LNJ F1 strains as belonging to one expression cluster associated with high virulence (*i.e.,* qualitatively

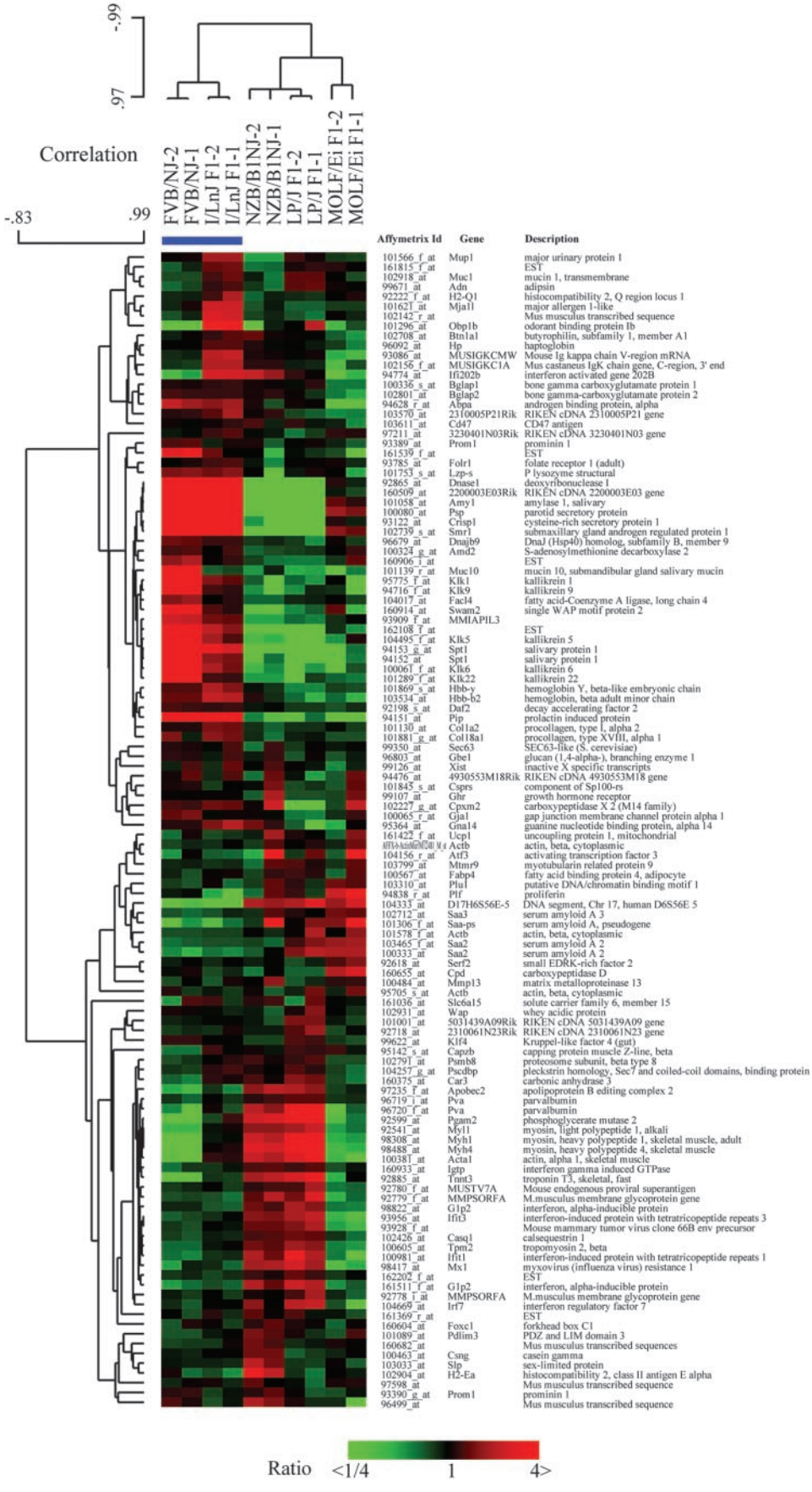

Fig. 3. Hierarchical clustering of 118 genes having highest variance and significant differential expression in Affymetrix arrays but not present on cDNA arrays. Mean-centered gene expressions are shown. Clustering was done using average linkage and uncentered correlation. Tumors from the FVB/NJ parental strain and the I/LnJ F1 cross-cluster together from the others (NZB/B1NJ, MOLF/Ei, and LP/J).

more malignant behavior or higher metastatic density, shorter latency period, and faster growth rate). Those tumors from MOLF/Ei F1 and LP/J F1 strains formed another cluster associated with a low virulence phenotype (Table 1). Tumors arising from NZB/B1NJ F1 mice appeared to be intermediate from the other two clusters because its classification was dependent on the array gene sets used. This intermediate appearance on arrays correlated well with the intermediate virulence index of the NZB/B1NJ F1. Thus, our expression results using two array platforms separate the MMTV-PyMT tumors into three virulence clusters based on their host genetic backgrounds.

Correlation of a Virulence Phenotype to Gene Expression. Our attempts to define specific gene lists that could distinguish the different individual phenotypes of metastases, growth rate, and latency were not successful, suggesting that the specific phenotype ascertainment is either imperfect or highly interconnected in a complex manner that precludes simple classification. For example, using the expression patterns from all five of the genotypes did not permit the unambiguous identification of a gene set predictive of any metric for metastatic potential. This strongly suggests that other factors, including tumor growth rate and tumor mass, which vary between the FVB/NJ and LP/J F1, MOLF/Ei F1, and I/LnJ F1 genotypes, influence the efficiency of tumor dissemination. Alternatively, covariates in addition to metastasis are likely to degrade the power of the analysis so that five genotypes may not have sufficient statistical power to identify a predictive gene set especially if these genes individually have minor effects on metastases.

Recently however, a small set of 17 genes was reported to predict metastatic potential for a variety of solid tumors by comparing the expression patterns of primary and metastatic samples from human samples (1). We therefore asked whether in our dataset, the murine orthologs of these putative human metastatic genes could also play a role in the metastatic phenotype in this model system. To test this, we examined expression data from two strains with the greatest difference in their metastatic potential (as ranked by how many metastases per square micrometer of lung tissue): high-metastasis (FVB/NJ) and low-metastasis (NZB/B1NJ F1). Orthologs of all 17 genes reported from the human analysis were identified on the Affymetrix mouse chips. Intriguingly, except small nuclear ribonucleoprotein polypeptide F, 16 of these genes showed the same expression direction as the human set (*i.e.*, murine tumors with the higher metastatic potential showed the same differential expression of 16 of 17 genes that distinguished human primary from metastases), 10 of which had expression differences of at least 1.4-fold. These results (Table 4; Fig. 4) suggest that a limited number of genes associated with metastases in human cancers appear also to be associated with metastatic potential of primary murine tumors and that the behavior of these genes may be modulated by host genetic background and not because of initiating oncogenes (23).

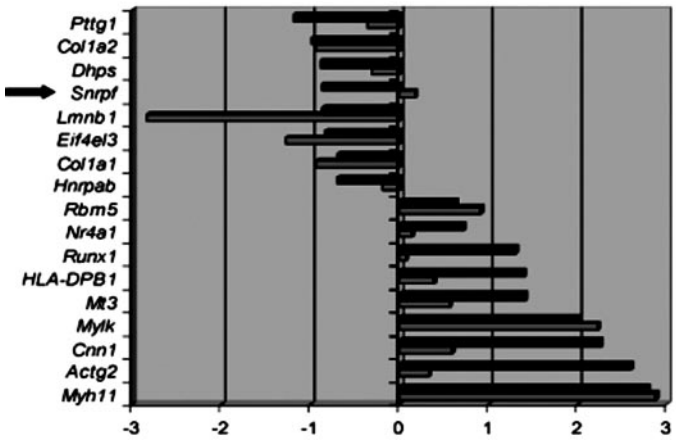

Fig. 4. Correlation of changes in gene expression between those observed for the 17 metastases-related genes identified by Ramaswamy *et al.* (1) in human cancer and orthologous mouse genes identified in the present study. Data are represented as the $\log_2$-transformed ratio of gene expression comparing primary tumor to metastatic tumor (human) [1] to the $\log_2$-transformed ratio of gene expression of the poorly metastatic NZB/B1NJ tumors to the highly metastatic FVB/NJ tumors. *Values to the right* of 0 indicate lower levels of expression of a gene in the metastasis-prone tumor compared with the poorly metastatic tumors, whereas *values to the left* of 0 indicate higher levels of gene expression in the metastasis-prone tumors compared with the poorly metastatic tumors. The *arrow* highlights the exception to the human-mouse correlation.

Table 4 *Comparison of expression of human metastasis genes with orthologous mouse genes*

Gene	Ratio NZB/B1NJ F1 to FVB/NJ (mouse)	Primary/Met (human)
Myh11	2.882063	2.786237
Actg2	0.331283	2.596367
Cnn1	0.587787	2.265409
Mylk	2.230014	2.010031
Mt3	0.570307	1.41755
HLA-DPB1	0.38637	1.397184
Runx1	0.074384	1.305966
Nr4a1	0.13072	0.72063
Rbm5	0.92	0.637207
Hnrpab	−0.19343	−0.69672
Col1a1	−0.92726	−0.7011
Eif4el3	−1.27297	−0.82898
Lmnb1	−2.82689	−0.86586
Snrpf	**0.18**	**−0.87507**
Dhps	−0.30111	−0.88902
Col1a2	−0.94036	−0.98035
Pttg1	−0.35154	−0.35154

NOTE. Mouse homologues were identified to the 17 genes identified by Ramaswamy *et al.* (1) as predictors of metastatic potential for human solid tumors. Sixteen of the 17 genes displayed similar changes in the direction of expression when comparing primary tumors with metastases; the exception is bold. Values represent the natural log of the ratio of expression values for the tumor samples being compared (*i.e.*, NZB/B1NJ F1 *versus* FVB/NJ or human primary *versus* metastatic tumors).

DISCUSSION

To better decipher the molecular mechanisms involved in mammary tumorigenesis including factors in the host background that may influence tumor behavior, we analyzed the global gene expression patterns of mammary tumors initiated by the identical oncogene PyMT but expressed in the context of five different genomic backgrounds. First, we compared the differences in gene expression profiles between the PyMT mammary tumors and the normal mammary gland to identify a general set of "cancer genes" irrespective of the strain backgrounds. Our results were consistent with and extended a study published previously (22), thus validating our methodological approach. Second, using ANOVA, we identified a set of genes with expression that varied according to the host background. These results were consistent using two different methodological approaches, providing strong evidence that this gene set was highly relevant.

A total of 1,068 and 1,241 "cancer genes" were found by cDNA and oligonucleotide (Affymetrix) array analyses, respectively. These genes are consistently up- or down-regulated in the mammary tumors from any one of the five strains. To define a minimal set of genes that could distinguish a mammary tumor from the normal mammary gland, we selected those elements that were significantly dysregulated in tumors derived from all five genomic backgrounds. This resulted in a common gene set of 164 genes (Supplementary Table 1*A*) using the cDNA microarrays and 188 genes (Supplementary Table 1*C*) from the oligonucleotide arrays.

Included in this list are genes associated previously with human breast cancer, such as, STAT3, CD24 antigen, lipocalin2, and procollagen Iα and IIIα. Genes associated with cellular growth were up-

regulated in tumors [CDC28, hairy (Hey-1), and lactotransferrin]. In contrast, cell adhesion molecules or membrane proteins were generally down-regulated (integral membrane protein 2A, procollagen Iα and IIIα, caveolin, actin α1, CD34, CD36). Several transcription factors and regulators of cell metabolism were highly induced in the PyMT-induced tumor models. Taken together, this suggests that the cancer cells express a cassette of genes supporting cell growth, cellular metabolism, and reducing cell adhesion.

Of particular interest is the fact that the p85 regulatory subunit of PI3K (P85α) was up-regulated in all five PyMT-induced tumor models. p85 has been shown previously to be a key signaling mediator of the Src tyrosine kinase family and the Shc-adaptor protein involved in PyMT transformation (24). Specifically, PyMT associates with one of the two SH2 domains of the p85 subunit of PI3K (25). The importance of the PI3K in mammary tumorigenesis has been established. Transgenic mice expressing a mutant PyMT unable to interact with PI3K develop widespread mammary hyperplasia with a high rate of apoptosis (26) in contrast to the rapid induction of metastatic mammary tumors observed in mice expressing wild-type PyMT. Mutant PyMT-transgenic mice do eventually develop tumors but with a greatly extended latency. Activated PI3K can also result in the activation of Akt, which can stimulate a number of antiapopototic signaling molecules (10) and can inhibit proapoptotic proteins such as Bad, Forkhead transcription factor, and caspase 9 (11, 12).

One hallmark of oncogene transformation is the deregulation of cell cycle control mechanisms. Two important cell cycle regulatory molecules, CDC kinase subunit (*Cks1*) and *Cdc25C* phosphatase, were up-regulated in PyMT tumors in all five mouse backgrounds.

Cks proteins interact tightly with cyclin-CDK (cyclin-dependent kinases) complexes. Cks proteins are essential for CDK function and cell division in yeast (27). *Cks1* has been shown recently to be an essential cofactor in the ubiquitination of the CDK inhibitor p27. p27 acts as a powerful negative regulator to the G_1-S transition by binding to cyclin E-Cdk2 and to cyclin A-Cdk2. Therefore, Cks1 appears to play an important role in regulating G_1-S transition through the cell cycle, which may be important for cancer progression (28, 29). Aggressive human tumors are associated with low p27 levels, and the absence of p27 expression is a powerful prognostic marker of poor survival in patients with breast, esophageal, lung and colorectal carcinomas (30). Given its role in promoting p27 degradation, Cks1 may function as a novel proto-oncogene. The discovery of elevated expression levels of Cks1 in all PyMT-initiated mammary tumors in our study suggests that the molecule might play a key role in the mammary epithelial cell transformation.

Both Cks1 and Cdc25C are up-regulated in PyMT-transgenic mammary tumors. The dual protein phosphatase Cdc25 family members regulate intracellular signaling through the mitogen-activated and stress-activated kinases and control the cell transition through G_1-S and G_2-M in cell circle checkpoints by affecting cyclin-dependent kinase activities. The Cdc25 phosphatases are overexpressed in many human tumors, and this increased expression is associated with a poor prognosis (31). Cdc25A acts on Cdk2 and Cdk4 at the G_1-S-phase and promotes cell entrance into S-phase (32), whereas Cdc25B and Cdc25C act on cdc2 at the G_2-M phase and play a role in the onset of mitosis. The biological significance of Cdc25A in the PyMT-transgenic mice has been studied recently. The epistatically interacting modifier loci in chromosomes 9 and 15 accelerate PyMT-induced mammary tumorigenesis in the I/LnJ background (33). Further analysis revealed that FVB/NJ chromosome 15 was associated with tumor acceleration and was conditioned on the presence of the I/LnJ allele on chromosome 9 (loci that have been designated as Apmt1 and Apmt2; ref. 17). A combined genetics, genomics, and bioinformatics approach identified c-myc and Cdc25A as strong candidates for the tumor latency modifier Apmt1 and Apmt2, respectively (34). Taken together, Cks1 plays an important role in controlling the G_1-S transition, whereasCdc25C is a key positive regulator of mitosis.

Lastly, our results demonstrated that gelsolin is significantly reduced in all PyMT mammary tumors. Gelsolin has been considered to be a prognostic factor for a variety of human tumors (35), especially for erbB2+ and EGFR+ breast cancer patients (36). Gelsolin is an actin-binding protein that participates in and regulates dynamic changes in the actin cytoskeleton and is widely expressed by normal cells and down-regulated with transformation of multiple cell types, including breast epithelium (37, 38). Recently, gelsolin has been shown to have tumor-suppressive activity (37, 39).

Uniquely, we have addressed genetic elements that underlie strain differences in tumor behavior. Hierarchical clustering or multidimensional scaling using these sets of cancer genes was unable to separate the tumor samples into particular subgroups based on the specific tumor phenotypes such as tumor latency, tumor growth rate, and metastatic potential. We suspect that this is because of phenotypic characteristics used to describe the tumors cannot be mechanistically well separated. We defined a virulence index (Table 1) that takes into account contributions from each tumor phenotype to a general virulence metric for the tumor. When we compared the array clusters to the virulence indices, we found a natural separation in that the least virulent (MOLF/Ei F1 and LP/J F1) constitutes a set that is distinct from the most virulent (FVB/NJ and I/LnJ F1) tumors, with NZB/B1NJ F1 showing a profile associated with an intermediate phenotype. Taken together, these results suggest that the standard division of tumor behavior into growth, latency, and metastatic potential may not reflect reality and that all three characteristics may be inextricably integrated in biology. We suggest that gene functions affecting metastatic behavior will be difficult to dissect from those influencing tumor growth and latency. More interestingly, however, we have identified genes that might influence tumor virulence modulated by host genetic factors.

Our replication of the human metastasis predictive expression pattern in the FVB/NJ *versus* NZB/B1NJ F1 comparison has several important implications. First, it suggests that many of the factors present in humans breast cancers that induce the metastatic expression signature may also be present in these inbred murine strains, further validating this model system to study metastasis in human breast cancers. Moreover, earlier models predicted that only a rare subpopulation of primary tumor cells would acquire the numerous genetic alterations required for metastatic spread (*i.e.,* a progression of secondary mutations). However, our findings suggest that the primary tumor may already clearly express a signature set of genes that increases the chance for metastatic dissemination and that secondary genetic mutations may be necessary but not sufficient for this metastatic progression. Finally, other genetic models for metastases suggest that the different metastatic potential of tumors is directly caused by a variety of genetic mutations within the tumor. Our observations, however, raise the importance of non-tumor factors that modulate metastatic potential (17) because all of the tumors studied here were initiated by the same oncogenic event. In our study, the association between metastatic efficiency and the metastatic signature gene expression profiles is most likely because of genetic background effects rather than different combinations of oncogenic mutations.

Thus, intriguingly, there may be human subpopulations that are "susceptible" to metastatic dissemination. Furthermore, if the predictive metastatic signature profile is the result of differences in background allelic composition, it predicts that a metastatic "signature" gene expression profile may be obtainable from virtually any tissue of the body. In support of this possibility, we have demonstrated recently that several of the human 17 gene metastasis signature genes are

differentially expressed in normal, transgene-negative mammary, and lung tissue from the high-metastatic FVB/NJ *versus* two low-metastatic F1 genotypes (NZB/B1NJ and DBA/2J[9]). These data suggest that easily accessible non-tumor tissue such as skin may be useful for identifying those patients at highest risk for latent disseminated disease. Such an approach could stratify patients who may benefit from aggressive neoadjuvant therapy to eliminate micrometastases or potential chemoprevention regimes to prevent development of clinically detectable disease.

ACKNOWLEDGMENTS

The authors are grateful to the National Cancer Institute Microarray Core at the Advanced Technology Center, Gaithersburg, MD, for providing cDNA microarrays and to the National Cancer Institute Mouse Models of Mammary Cancer Collective and Mouse Models of Human Cancer Consortium for support of this project.

REFERENCES

1. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. Nat Genet 2003;33:49–54.
2. 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.
3. Andrechek ER, Muller WJ. Tyrosine kinase signalling in breast cancer: tyrosine kinase-mediated signal transduction in transgenic mouse models of human breast cancer. Breast Cancer Res 2000;2:211–6.
4. Cheng SH, Harvey R, Espino PC, et al. Peptide antibodies to the human c-fyn gene product demonstrate pp59c-fyn is capable of complex formation with the middle-T antigen of polyomavirus. EMBO J 1988;7:3845–55.
5. Courtneidge SA, Smith AE. The complex of polyoma virus middle-T antigen and pp60c-Src. EMBO J 1984;3:585–91.
6. Webster MA, Cardiff RD, Muller WJ. Induction of mammary epithelial hyperplasias and mammary tumors in transgenic mice expressing a murine mammary tumor virus/activated c-Src fusion gene. Proc Natl Acad Sci USA 1995;92:7849–53.
7. Courtneidge SA, Heber A. An 81 kd protein complexed with middle T antigen and pp60c-Src: a possible phosphatidylinositol kinase. Cell 1987;50:1031–7.
8. Whitman M, Kaplan DR, Schaffhausen B, et al. Association of phosphatidylinositol kinase activity with polyoma middle-T competent for transformation. Nature (Lond) 1985;315:239–42.
9. Alessi DR, Deak M, Casamayor A, et al. 3-Phosphoinositide-dependent protein kinase-1 (PDK1): structural and functional homology with the Drosophila DSTPK61 kinase. Curr Biol 1997;7:776–89.
10. Kane LP, Shapiro VS, Stokoe D, Weiss A. Induction of NF-kappaB by the Akt/PKB kinase. Curr Biol 1999;9:601–4.
11. Cardone MH, Roy N, Stennicke HR, et al. Regulation of cell death protease caspase-9 by phosphorylation. Science (Wash D C) 1998;282:1318–21.
12. Datta SR, Dudek H, Tao X, et al. Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery. Cell 1997;91:231–41.
13. Blaikie PA, Fournier E, Dilworth SM, et al. The role of the Shc phosphotyrosine interaction/phosphotyrosine binding domain and tyrosine phosphorylation sites in polyoma middle T antigen-mediated cell transformation. J Biol Chem 1997;272:20671–7.
14. Campbell KS, Ogris E, Burke B, et al. Polyoma middle tumor antigen interacts with SHC protein via the NPTY (Asn-Pro-Thr-Tyr) motif in middle tumor antigen. Proc Natl Acad Sci USA 1994;91:6344–8.
15. Dilworth SM, Brewster CE, Jones MD, et al. Transformation by polyoma virus middle T-antigen involves the binding and tyrosine phosphorylation of Shc. Nature (Lond) 1994;367:87–90.
16. Srinivas S, Schonthal A, Eckhart W. Polyomavirus middle-sized tumor antigen modulates c-Jun phosphorylation and transcriptional activity. Proc Natl Acad Sci USA 1994;91:10064–8.
17. Lifsted T, Le Voyer T, Williams M, et al. Identification of inbred mouse strains harboring genetic modifiers of mammary tumor age of onset and metastatic progression. Int J Cancer 1998;77:640–4.
18. Eberwine J, Yeh H, Miyashiro K, et al. Analysis of gene expression in single live neurons. Proc Natl Acad Sci USA 1992;89:3010–4.
19. Chandrasekharan S, Qiu TH, Alkharouf N, et al. Characterization of mice deficient in the Src family nonreceptor tyrosine kinase Frk/rak. Mol Cell Biol 2002;22:5235–47.
20. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 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 USA 1998;95:14863–8.
22. Desai KV, Xiao N, Wang W, et al. Initiating oncogenic event determines gene-expression patterns of human breast cancer models. Proc Natl Acad Sci USA 2002;99:6967–72.
23. Hunter K, Welch DR, Liu ET. Genetic background is an important determinant of metastatic potential. Nat Genet 2003;34:23–4.
24. Andrechek ER, Muller WJ. Tyrosine kinase signalling in breast cancer: tyrosine kinase-mediated signal transduction in transgenic mouse models of human breast cancer. Breast Cancer Res 2000;2:211–6.
25. Soltoff SP, Carpenter CL, Auger KR, et al. Phosphatidylinositol-3 kinase and growth regulation. Cold Spring Harb Symp Quant Biol 1992;57:75–80.
26. Webster MA, Hutchinson JN, Rauh MJ, et al. Requirement for both Shc and phosphatidylinositol 3′ kinase signaling pathways in polyomavirus middle T-mediated mammary tumorigenesis. Mol Cell Biol 1998;18:2344–59.
27. Pines J. Cell cycle: reaching for a role for the Cks proteins. Curr Biol 1996;6:1399–402.
28. Spruck C, Strohmaier H, Watson M, et al. A CDK-independent function of mammalian Cks1: targeting of SCF(Skp2) to the CDK inhibitor p27Kip1. Mol Cell 2001;7:639–50.
29. Ganoth D, Bornstein G, Ko TK, et al. The cell-cycle regulatory protein Cks1 is required for SCF(Skp2)-mediated ubiquitinylation of p27. Nat Cell Biol 2001;3:321–4.
30. Slingerland J, Pagano M. Regulation of the cdk inhibitor p27 and its deregulation in cancer. J Cell Physiol 2000;183:10–7.
31. Pestell KE, Ducruet AP, Wipf P, Lazo JS. Small molecule inhibitors of dual specificity protein phosphatases. Oncogene 2000;19:6607–12.
32. Jinno S, Suto K, Nagata A, et al. Cdc25A is a novel phosphatase functioning early in the cell cycle. EMBO J 1994;13:1549–56.
33. Le Voyer T, Lu Z, Babb J, Lifsted T, Williams M, Hunter K. An epistatic interaction controls the latency of a transgene-induced mammary tumor. Mamm Genome 2000;11:883–9.
34. Cozma D, Lukes L, Rouse J, et al. A bioinformatics-based strategy identifies c-Myc and Cdc25A as candidates for the Apmt mammary tumor latency modifiers. Genome Res 2002;12:969–75.
35. Sazawa A, Watanabe T, Tanaka M, et al. Adenovirus mediated gelsolin gene therapy for orthotopic human bladder cancer in nude mice. J Urol 2002;168:1182–7.
36. Thor AD, Edgerton SM, Liu S, Moore DH, Kwiatkowski DJ. Gelsolin as a negative prognostic factor and effector of motility in erbB-2-positive epidermal growth factor receptor-positive breast cancers. Clin Cancer Res 2001;7:2415–24.
37. Asch HL, Head K, Dong Y, et al. Widespread loss of gelsolin in breast cancers of humans, mice, and rats. Cancer Res 1996;56:4841–5.
38. Winston JS, Asch HL, Zhang PJ, et al. Downregulation of gelsolin correlates with the progression to breast carcinoma. Breast Cancer Res Treat 2001;65:11–21.
39. Vandekerckhove J, Bauw G, Vancompernolle K, Honore B, Celis J. Comparative two-dimensional gel analysis and microsequencing identifies gelsolin as one of the most prominent downregulated markers of transformed human fibroblast and epithelial cells. J Cell Biol 1990;111:95–102.

[9] K. Hunter, unpublished results.