Transcriptional Programs following Genetic Alterations in *p53*, *INK4A*, and *H-Ras* Genes along Defined Stages of Malignant Transformation

Michael Milyavsky,[1] Yuval Tabach,[1,2] Igor Shats,[1] Neta Erez,[1] Yehudit Cohen,[1] Xiaohu Tang,[1] Marina Kalis,[1] Ira Kogan,[1] Yosef Buganim,[1] Naomi Goldfinger,[1] Doron Ginsberg,[1] Curtis C. Harris,[3] Eytan Domany,[2] and Varda Rotter[1]

Departments of [1]Molecular Cell Biology and [2]Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel and [3]Laboratory of Human Carcinogenesis, National Cancer Institute, NIH, Bethesda, Maryland

Abstract

The difficulty to dissect a complex phenotype of established malignant cells to several critical transcriptional programs greatly impends our understanding of the malignant transformation. The genetic elements required to transform some primary human cells to a tumorigenic state were described in several recent studies. We took the advantage of the global genomic profiling approach and tried to go one step further in the dissection of the transformation network. We sought to identify the genetic signatures and key target genes, which underlie the genetic alterations in *p53*, *Ras*, *INK4A* locus, and *telomerase*, introduced in a stepwise manner into primary human fibroblasts. Here, we show that these are the minimally required genetic alterations for sarcomagenesis *in vivo*. A genome-wide expression profiling identified distinct genetic signatures corresponding to the genetic alterations listed above. Most importantly, unique transformation hallmarks, such as differentiation block, aberrant mitotic progression, increased angiogenesis, and invasiveness, were identified and coupled with genetic signatures assigned for the genetic alterations in the *p53*, *INK4A* locus, and *H-Ras*, respectively. Furthermore, a transcriptional program that defines the cellular response to p53 inactivation was an excellent predictor of metastasis development and bad prognosis in breast cancer patients. Deciphering these transformation fingerprints, which are affected by the most common oncogenic mutations, provides considerable insight into regulatory circuits controlling malignant transformation and will hopefully open new avenues for rational therapeutic decisions. (Cancer Res 2005; 65(11): 4530-43)

Introduction

Human carcinogenesis can be divided into defined clinicopathologic stages. For example, colon cancer progression has been divided into distinct histologic stages directly correlated with genetic alterations in key tumor suppressors and oncogenes (1). Over the last two decades, the molecular nature of genes frequently mutated in human neoplasia was elucidated. Functionally, those genes can be divided to many categories. The most studied genes include signaling molecules (*Ras*, *Src*, *Akt*, *tyrosine kinase receptors*, etc.), core cell cycle regulators (*pRb*, *p16^{INK4A}*, *cyclins*, etc.), and transcription factors (*p53*, *E2F*, *NF-κB*, etc.), This knowledge leads to the realization that neoplastic transformation involves aberrant signal transduction pathways intimately linked with the deregulated gene expression. Nevertheless, the underlying transcriptional changes, which arise as a consequence of sequential accumulation of genetic alterations and eventually drive the pathologic process, are still elusive.

We addressed this challenge by the microarray technology. Monitoring gene expression changes on a genome-wide scale has proven to be a powerful method to study transcriptional programs involved in carcinogenesis (2). Comparisons between normal tissues and corresponding tumors or between various tumor types revealed significant differences in their mRNA profiles, including hundreds of differentially expressed genes (2). By combining classic supervised statistical methods with unsupervised techniques, such as hierarchical clustering and its advanced variants (3), analysis of microarray data can potentially identify specific biological signatures that reflect profound alterations in cellular pathways and processes. Indeed, molecular signatures that correlate with diagnosis and prognosis were discovered (2, 4–6). Yet, associations of those signatures with specific biological processes and genetic alterations acquired *in vivo* along transformation are not obvious. The difficulties stem largely from different genetic backgrounds of patients, variable and uncharacterized mutations, and undefined contributions to a resulting expression pattern of several cell types, such as inflammatory, endothelial, and stroma cells in addition to the bona fide tumor cells. Those considerations made it almost impossible to use established malignant cell lines or naturally occurring tumors to dissect the contribution of individual tumor suppressors or oncogenes to the observed changes in gene expression.

Modeling of human carcinogenesis *in vitro* is an invaluable tool to examine the effects of individual oncogenes, tumor suppressors, and their combinations on the evolvement of the transformed phenotype. In this way, recently, the defined combinations of oncogenic events required to convert primary human cells into full-blown tumors were determined. Initially, full transformation was achieved by the combination of viral oncogenes, such as large and small T antigens together with cellular genes, such as mutant *Ras* and *telomerase* (7, 8). Later on, the cellular counterparts of viral oncogenes were showed to be sufficient to transform primary human fibroblasts. These include inactivation of p53 and either pRb or p16^{INK4A} tumor suppressors, overexpression of the catalytic

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

M. Milyavsky and Y. Tabach contributed equally to this work. V. Rotter holds the Norman and Helen Asher Professorial Chair in Cancer Research at the Weizmann Institute. E. Domany is the incumbent of the H.J. Leir Professorial Chair.

Requests for reprints: Varda Rotter, Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 76100, Israel. Phone: 972-8-9344501; Fax: 972-8-9465265; E-mail: varda.rotter@weizmann.ac.il.

subunit of telomerase (hTERT), inhibition of protein phosphatase 2A, and abnormal activation of Ras downstream pathways (9, 10). Importantly, it was noted that different cell types vary significantly in their susceptibility to the same combination of transforming elements (11, 12).

Thus, to obtain both novel and more reliable results, we set out to study a stepwise process of malignant conversion that makes use of human primary isogenic cells.

Here, we describe a transcriptional program involved in malignant transformation in a unique cellular model. The model consists of WI-38 human diploid fibroblasts in which replicative senescence was overcome by using hTERT, resulting in sustained proliferation of cells. The resulting extraordinary large number of divisions [150 population doublings (PDL) following hTERT introduction] eventually gave rise to the INK4A-deficient clones with a significantly higher rate of proliferation, defective contact inhibition checkpoint, and, perhaps most importantly, sensitivity to H-Ras-mediated transformation (13). Recent studies have denoted similar inactivation of $p16^{INK4A}$ and subsequent sensitivity to the H-Ras-mediated transformation in additional strains of human fibroblasts that overcome telomere-independent crisis during immortalization (14–16). Collectively, these studies, including our own, which model malignant transformation *in vitro*, created a framework of defined oncogenic aberrations that initiate and promote neoplastic process. Based on these analyses, we hypothesized that our *in vitro* cellular model could recapitulate, with the known limitations of cell transformation in culture, the distinct stages that characterize mesenchymal cell transformation initiation and progression.

We describe here the identification of specific "genetic signatures" associated with each of the genetic changes that lead from normal diploid cells to fully transformed and tumorigenic cells.

Materials and Methods

Cell culture, retroviral constructs, and infection. Primary human embryonic lung fibroblasts (WI-38), amphotropic and ecotropic Phoenix retrovirus-producing cells, and retroviral constructs and infection procedures have been described (13).

Subcutaneous tumorigenicity assay. Immunocompromised athymic nude mice (CD-1-nude; 6-8 weeks old) were irradiated with 4 Gy 24 hours before injection. WI-38/T^{fast} and its transformed derivatives (1×10^7 cells) were resuspended in 100 μL PBS. Immediately before injection, Matrigel (100 μL, Becton Dickinson, Franklin Lakes, NJ) was added to the cells. Tumor size was monitored every 5 days. Mice were sacrificed when the tumor reached a diameter of 1 to 1.5 cm or after 26 weeks of monitoring. Tumors were collected in a sterile field and minced. Tumor fragments were immediately frozen in liquid nitrogen for DNA, RNA, and protein extraction. Additional fragments were fixed in 10% formalin for histologic and immunohistochemical examinations. Finally, fragments were finely minced, washed in PBS, and plated in culture medium for isolation of tumor cells. All mouse procedures were done with the approval of the Animal Care and Use Committee of the Weizmann Institute of Science (Rehovot, Israel).

Isolation of total RNA. Total RNA for the microarray hybridization and quantitative real-time PCR (QRT-PCR) was isolated using RNeasy kit (Qiagen, Valencia, CA) according to the manufacturer's protocol.

Quantitative real-time PCR. A 2 μg aliquot of the total RNA was reverse transcribed using Moloney murine leukemia virus reverse transcriptase (Promega, Madison, WI) and random hexamer primers. QRT-PCR was done using SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA) on an ABI 7000 instrument (Applied Biosystems). The values for the specific genes were normalized to the *GAPDH* housekeeping control. Primer sequences for SYBR Green PCR were as follows: *GAPDH*, 5′-ACCCACTCCTCCACCTTTGA and 3′-CTGTTGCTGTAGCCAAATTCGT; *CDKN1C* ($p57^{KIP2}$), 5′-GAACGCCGAGGACCAGAAC and 3′-GGCATGTCCTGCTGGAAGTC; *LDOC1*, 5′-CGTGCAGACGGCGTCTTAC and 3′-GGCGTCGTTGCAGAATCG; *MAGEA1*, 5′-CCGCCTTTCCCACTACCA and 3′-CCTCACTGGGTT GCCTCTGT; *SSX1*, 5′-ACCGCAGGATTCAGGTTGAA and 3′-TGTGGAGC CTGCCGAAAG; *CXCL1*, 5′-AGTCATAGCCACACTCAAGAATGG and 3′-GATGCAGGATTGAGGCAGC; *IL1B*, 5′-GCCTGAAGCCCTTGCTGTAGT and 3′-GCGGCATCCAGCTACGAAT; *MMP3*, 5′-ACAAAGGATACAACAGGGACCAA and 3′-CAATTTCATGAGCAGCAACGA; *ACTA2*, 5′-TGTAAGGCCGGCTTTGCT and 3′-CGTAGCTGTCTTTTTGTCCCATT; *CNN1*, 5′-CCGTGAAGAAGATCAATGAGTCAA and 3′-CAGGTCGTTGGCCTCAAAA; and *CALD*, 5′-GGAGATGCGACTCGAAGCA and 3′-GTCACCTGTC CCAAGGATTC.

Induction of smooth muscle cell differentiation by transforming growth factor-β1. WI-38 cells and their hTERT derivatives were grown to visual confluence in MEM supplemented with 10% FCS, 1 mmol/L sodium pyruvate, and 2 mmol/L L-glutamine. Then, the cells were brought to quiescence in serum-free MEM for 24 hours and exposed to control medium (serum-free MEM) or medium containing 1 ng/mL transforming growth factor-β1 (TGF-β1; R&D Systems, Abingdon, United Kingdom) for 36 hours.

Western blot analysis. Cells were lysed in TLB buffer (50 mmol/L Tris-HCl, 100 mmol/L NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with protease inhibitor cocktail (Roche, Basel, Switzerland) and phosphatase inhibitor cocktails I and II (Sigma, St. Louis, MO) for 30 minutes on ice. Extracts were analyzed for protein concentration by Bradford assay. The following primary antibodies were used: mouse monoclonal anti-α-smooth muscle actin (α-SMA; clone 1A4, Sigma) and mouse monoclonal anti-tubulin (Sigma). The protein-antibody complexes were detected using horseradish peroxidase–conjugated secondary antibodies and the SuperSignal enhanced chemiluminescence system (Pierce, Rockford, IL).

Cell cycle analysis. First, 1.5×10^6 cells were plated directly into complete medium or complete medium containing 0.05 μg/mL nocodazole. After 72 hours, the cells were detached with trypsin, fixed in 70% ethanol/30% HBSS for at least 24 hours, washed, and resuspended in PBS containing 5 μg/mL propidium iodide and 0.1 mg/mL RNase A. Samples were analyzed by flow cytometry using a fluorescence-activated cell sorting (FACS) sorter machine (Becton Dickinson). Cells with more than 4N DNA were determined by gating the cells to the right of the 4N G_2-M peak on the histogram plots. At least 20,000 cells were analyzed per sample. Experiments were repeated at least thrice and similar results were obtained.

Immunofluorescence. Smooth muscle differentiation markers were detected by immunofluorescence as described essentially by Chambers et al. (17).

Microarray hybridization and processing. Double-stranded cDNA was generated from 15 μg total RNA using the SuperScript Choice System from Invitrogen (Carlsbad, CA), with an oligo$(dT)_{24}$ primer containing a T7 promoter site at the 3′ end (Genset, La Jolla, CA). cDNAs were purified via phenol/chloroform extraction followed by ethanol precipitation. Purified cDNA was used as template for *in vitro* transcription, using the Enzo BioArray High-Yield RNA Transcript Labeling kit (Enzo Diagnostics, New York, NY). Labeled *in vitro* transcripts were purified over RNeasy mini columns according to the manufacturer's instructions. The labeled cRNA was fragmented at 94°C for 35 minutes in fragmentation buffer [40 mmol/L Tris acetate (pH 8.1), 100 mmol/L potassium acetate, 30 mmol/L magnesium acetate], and a hybridization mixture was generated by addition of 0.1 mg/mL herring sperm DNA, 0.5 mg/mL acetylated bovine serum albumin (Invitrogen), 1 mol/L NaCl, 10 mmol/L Tris-acetate, and 0.0001% Tween 20. A mixture of four control bacterial and phage cRNA panels (1.5 pmol/L BioB, 5 pmol/L BioC, 25 pmol/L BioD, and 100 pmol/L Cre) was included as an internal control for hybridization efficiency. Aliquots of each sample (12 μg cRNA in 200 μL hybridization mix) were hybridized to a GeneChip Human Genome Focus Array (Affymetrix, Santa Clara, CA). After hybridization, each array was washed according to procedures developed by the manufacturer (Affymetrix) and stained with streptavidin-phycoerythrin conjugate (Molecular Probes, Eugene, OR). The hybridization signal was amplified using biotinylated anti-streptavidin

antibodies (Vector Laboratories, Burlingame, CA) followed by restaining with streptavidin-phycoerythrin. Arrays were scanned by the GeneArray scanner G2500A (Hewlett Packard, Palo Alto, CA), and scanned images were visually inspected for hybridization imperfections. Arrays were analyzed using Affymetrix Microarray Suite software version 5.0 by scaling to an average intensity of 250.

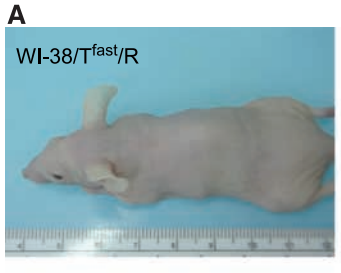
Data analysis. Gene expression analysis was done in duplicate on 12 data points. Each transcript represented on the array was designated by Microarray Suite software version 5.0 as either present, absent, or marginal, and only those genes that had a present call ($P < 0.05$) in both repeats of at least one data point were retained. The log-transformed expression values of each gene were mean centered (by subtracting the average) and normalized to generate the final expression matrix, of 5,581 rows (genes) and 24 columns (samples), which served as the input for the super-paramagnetic clustering analysis (SPC; ref. 18). Stable, statistically significant gene clusters (3) were identified using this algorithm. Fold change: to meet the condition of a "2-fold increase" in condition A versus condition B, the lower of the two repeats of A must have at least twice the value of the higher one of B.

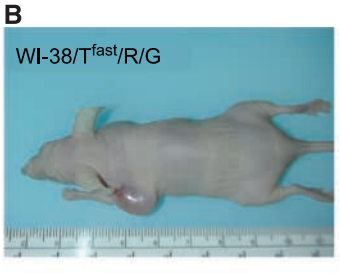
Results

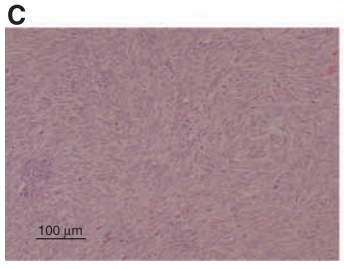
Conversion of the *INK4A* locus–deficient WI-38/hTERT fibroblasts into tumor cells requires inactivation of *p53* and *H-RasV12* expression. To define the possible requirements for tumorigenesis *in vivo*, we first determined which of the oncogenic changes in WI-38 cells results in cells capable of forming tumors in mice. We showed previously that hTERT-induced immortalization of WI-38 human diploid fibroblasts results in the spontaneous emergence of rapidly proliferating variants (WI-38/T^{fast}). Those clones do not express *INK4A* locus genes ($p16^{INK4A}$ and $p14^{ARF}$) and have elevated levels of the c-*myc* oncogene. In addition, these cells could be further transformed with a constitutively active *H-RasV12* gene, which confers them with anchorage-independent growth (WI-38/T^{fast}/R cells). Inactivation of wild-type p53, using the dominant-negative polypeptide GSE56 (19), concomitant with *H-RasV12* expression, resulted in a dramatic increase in anchorage-independent proliferation (WI-38/T^{fast}/R/G cells; ref. 13). We found that only WI-38/T^{fast}/R/G cells formed tumors in 30% of the mice into which they were injected (Fig. 1). This low frequency of tumor formation as well as a long latency period (between 68 and 108 days) suggested that additional genetic changes might have been required. None of the WI-38/T^{fast}, WI-38/T^{fast}/G, or WI-38/T^{fast}/R was tumorigenic after 6 months of monitoring. These results strongly suggest that a tumorigenic phenotype in the WI-38 fetal lung fibroblasts requires coexpression of constitutively active H-Ras with inactivation of wild-type p53 in the *INK4A* locus silenced cells.

Gene expression profiling along defined stages of malignant transformation *in vitro*. To obtain a comprehensive picture of changes in gene expression along defined stages of the mesenchymal malignant transformation, representative samples were selected. They include parental WI-38 fibroblasts in the young and senescent stages as well as the hTERT immortalized cells at the different time points. In addition, p53 was inactivated by a dominant-negative peptide GSE56 (19), and H-Ras expression was induced by infection at the indicated time points (Fig. 2; Table 1). Samples were taken in duplicates at 12 points and hybridized with the GeneChip Human Genome Focus Array; the relative mRNA abundance of ~8,500 human genes was monitored. After standard preprocessing steps (see Materials and Methods), >200,000 expression values were collected from 24 microarrays. Of these, 5,581 probe sets passed a filter (see Materials and Methods) and were analyzed by an unsupervised clustering algorithm, SPC (18), to identify genes with a correlated pattern of expression. Our working hypothesis was that genes forming a cluster represent unique genetic signatures, which are able to distinguish between the stages of transformation. In addition, the unique expression profiles over the samples may contain information about transcriptional programs initiated by tumor suppressor inactivation ($p16^{INK4A}$ and p53) versus oncogene activation (H-Ras). By using this approach, we were able to identify 10 predominant, stable, and significant gene clusters. We focus below on the most informative of them.

A defect in myogenic differentiation program characterizes early stages of transformation. By using SPC analysis, we identified a set of 397 genes that showed orchestrated down-regulation at the early stages of transformation (Fig. 3*A*;







D

Cell line	Number tumors/ Number injected
WI-38/T^{fast}	0/5
WI-38/T^{fast}/G	0/4
WI-38/T^{fast}/R	0/5
WI-38/T^{fast}/G/R	3/8

Figure 1. Conversion of WI-38/T^{fast} fibroblasts into tumor cells requires inactivation of *p53* and *H-RasV12* expression. *A* and *B,* nude mice injected with either 1×10^7 WI-38 cells expressing H-RasV12 (*A*) or H-RasV12 and the p53-inactivating peptide, GSE56 (*B*). S.c. tumor mass of ~1 cm diameter is clearly visible in (*B*). *C,* H&E staining of WI-38/T^{fast}/R/G tumor sample (cross-section magnification, ×100) revealing an unencapsulated but well-demarcated mass composed of densely packed spindle cells. This histologic pattern is most consistent with fibrosarcoma. *D,* summary of *in vivo* tumorigenicity assay. Representative of one experiment of three, which gave similar results.

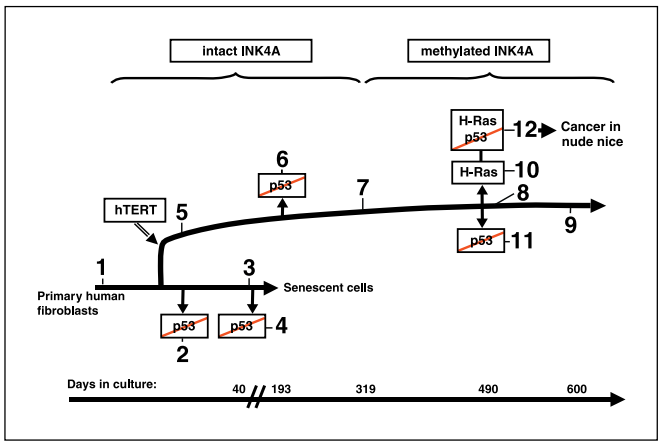


Figure 2. Outline of WI-38 primary human fibroblast malignant transformation process. Schematic representation of the physiologic (young, senescent, immortal, tumorigenic, *INK4A* methylation) and introduced (hTERT, H-Ras, p53 inactivation) modifications of the WI-38 cells along the process of malignant transformation. The stages chosen for microarray profiling are indicated by *boxes*. Time scale of the process is depicted by *horizontal axes*. Additional information about the samples is given in Table 1.

Supplementary Fig. S1). Importantly, this set of genes allowed us to discriminate between samples bearing an intact *INK4A* locus (WI-38 and T^{slow}), in which expression of these genes was relatively high, versus the *INK4A* locus–deficient samples and their derivatives (T^{fast}, T^{fast}/G, T^{fast}/R, and T^{fast}/R/G), all of which showed reduced expression of the same set of genes. By analyzing functional annotations of the down-regulated genes in this group, we found that a significant fraction of them (62 of 397 genes) is involved in various aspects of mesenchymal cell development and differentiation and, more specifically, in the differentiation of the smooth muscle lineage. These include extracellular signaling molecules (*BMP1*, *BMP2*, and *NOTCH3*) and their modulators (*IGFBP3*, *IGFBP4*, *LTBP1*, *LTBP2*, etc.), mesenchymal lineage-specific transcription factors (*TWSG1*, *PITX1*, *HHEX*, *MSX2*, *SOX11*, *PMX1*, and *OSR2*), and smooth muscle differentiation markers (*ACTA2*, *ACTG2*, *FHL1*, *CALD1*, *MYL6*, etc.; refs. 20–22). Coordinated reduction in the expression of numerous differentiation markers, as evidenced by the DNA microarrays results, led us to consider the possibility that WI-38 cells at this stage of transformation develop a defect in the smooth muscle terminal differentiation program.

To validate this hypothesis, we treated, in parallel, WI-38/$hTERT^{slow}$ (a representative of the normal cells) and WI-38/$hTERT^{fast}$ (*INK4A* locus–inactivated cells with a suspected defect in muscle differentiation) with TGF-β. This cytokine is a principal inducer of myogenic differentiation (23). We measured the changes in the expression of typical myofibroblast and smooth muscle cell markers, such as smooth muscle specific α isoform of actin (*ACTA2* or *α-SMA*), caldesmon (*CALD*), and calponin 1 (*CNN1*). WI-38/$hTERT^{fast}$ and their progeny expressed lower basal levels of all three markers compared with the parental WI-38 fibroblasts and WI-38/$TERT^{slow}$ as was predicted by the microarrays analysis and validated by QRT-PCR. More significantly, on TGF-β treatment, dramatic induction of these markers was evident in WI-38/$hTERT^{slow}$ cells, whereas only marginal changes were observed in WI-38/$hTERT^{fast}$ cells under the same treatment regimen (Fig. 3*B*). These results were further explored by α-SMA immunofluorescence

Table 1. Description of the samples selected for the microarray experiment

No.	Sample name	PDLs	Exogenously introduced genes	Karyotype	INK4A status	Tumor formation
1	WI-38/p, young	24	puro	46, XX	Intact	NA
2	WI-38/G, young	23	GSE56-puro	NA	Intact	NA
3	WI-38/p, senescent	40	puro	46, XX	Intact	NA
4	WI-38/G, senescent	42	GSE56-puro	NA	Intact	NA
5	WI-38/T^{slow}	40	hTERT-puro	46, XX	Intact	NA
6	WI-38/T^{slow}/G	140	hTERT-puro GSE56-Neo	NA	Intact	NA
7	WI-38/T^{slow}	215	hTERT-puro	NA	Intact	NA
8	WI-38/T^{fast}	355	hTERT-puro	46, X	Promoter region hypermethylation	NA
9	WI-38/T^{fast}	490	hTERT-puro	der(X)t(X;17) 46, X	Promoter region hypermethylation	—
10	WI-38/T^{fast}/R	360	hTERT-puro H-RasV12-hygro	der(X)t(X;17) 46, X	Promoter region hypermethylation	—
11	WI-38/T^{fast}/G	360	hTERT-puro GSE56-Neo	der(X)t(X;17) 46, X	Promoter region hypermethylation	—
12	WI-38/T^{fast}/R/G	365	hTERT-puro GSE56-Neo H-RasV12-hygro	der(X)t(X;17) Near tetraploid cells with 2X der(X)t(X;17) and also with different translocations in different cells	Promoter region hypermethylation	+

Abbreviation: NA, not assessed.

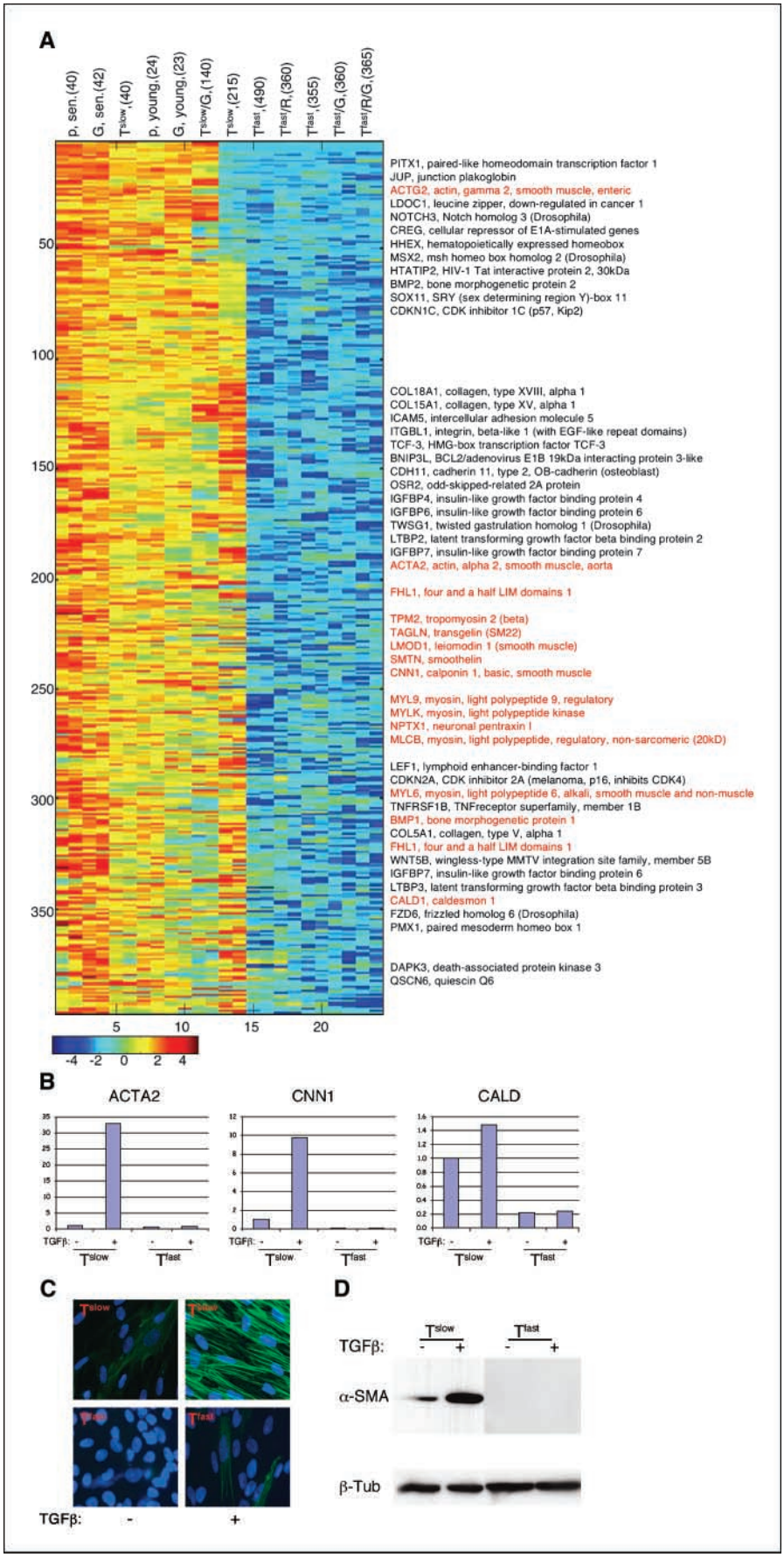


Figure 3. Identification of a genetic signature associated with the earliest detectable transformation stage. *A,* cluster of down-regulated genes correlating with *INK4A* locus expression. Set of 397 of 5,581 genes (that passed the filtering criteria using all 24 samples) identified by clustering all samples according to the similarity of their gene expression profiles. The cluster was identified by unsupervised analysis using the SPC clustering procedure. *Rows,* genes; *columns,* samples. The sample names are denoted at the *top* of the cluster and detailed in Table 1. Logarithmic color scale representing centered and normalized values (*bottom left*). All the samples with a silenced *INK4A* locus (T^{fast} and derivatives) showed reduced expression. Selected genes from the cluster, which are discussed in the text, are indicated at the *right* of the cluster. Smooth muscle differentiation markers are shown in *red*. A complete list of genes is available in Supplementary Fig. S1. *B,* validation of smooth muscle differentiation in WI-38/hTERT^{slow} and WI-38/hTERT^{fast} cells, representing *INK4A* locus intact and silenced stages, respectively. QRT-PCR of smooth muscle differentiation markers: α-SMA (*ACTA2*), calponin 1 (*CNN1*), and caldesmon (*CALD*). Subconfluent WI-38/hTERT^{slow} (65 PDLs, T^{slow}) and WI-38/hTERT^{fast} (560 PDLs, T^{fast}) were made quiescent by serum deprivation for 24 hours. Then, the medium was replaced with medium containing 1 ng/mL TGF-β1 for 36 hours. Gene expression was normalized to the *GAPDH* expression in the same sample. *C,* effect of TGF-β1 on the induction of α-SMA fibers in T^{slow} and T^{fast} cells exposed to serum-free medium or TGF-β1 for 36 hours and analyzed by immunofluorescence. Nuclei were stained with 4′,6-diamidino-2-phenylindole to show cell number in both control and TGF-β1-treated cells. Cells were prepared and treated as described in Materials and Methods in (*B*) above. Representative of several fields photographed at ×1,000 magnification. *D,* Western blot analysis showing the effect of TGF-β1 on the induction of α-SMA protein in the T^{slow} and T^{fast} cells. Blotting was done after exposure of cells to serum-free medium or 1 ng/mL TGF-β1 for 36 hours using antibodies specific to α-SMA and β-tubulin (*β-tub*; control).

analysis. Whereas a small fraction (~10%) of WI-38/hTERT^{slow} cells showed low levels of constitutive α-SMA staining, exposure to TGF-β induced marked cytoskeleton reorganization and development of a prominent network of brightly stained actin filaments in almost 100% of the cells, suggestive of myofibroblast differentiation. In stark contrast, WI-38/hTERT^{fast} cells showed no staining of actin fibers in the control sample, and only a few isolated cells developed slight staining of actin on TGF-β treatment (Fig. 3*C*). Western blotting analysis gave very similar results (Fig. 3*D*). These results provide compelling evidence that the myogenic differentiation program is impaired in WI-38/hTERT^{fast} cells.

In addition to the smooth muscle differentiation markers, several cell cycle regulators showed a marked reduction of their expression during the transition from the T^{slow} to the T^{fast} phenotype. Among these are several Rb regulators, such as *CDKN2A* (*p16^{INK4A}*; in agreement with our previous findings; ref. 13), *CDKN1C* (*p57^{KIP2}*), negative regulators of cell proliferation (*CREG* and *QSCN6*), and genes with known tumor suppressor activity (*LDOC1* and *FAT*). To confirm this trend, we measured the expression of two representative genes, *p57^{KIP2}* and *LDOC1*, by QRT-PCR. In agreement with the microarray results, both genes showed a dramatic reduction in their expression (Supplementary Fig. S2).

An additional functional group that showed decreased expression in the T^{fast} cells is composed of genes promoting apoptosis, such as *TNFRSF1B*, *TNFRS21*, *DAPK3*, *HTATIP2*, and *BNIP3L*, suggesting acquisition of increased resistance to some apoptotic stimuli starting from the premalignant stage.

Early stages of transformation: high biosynthetic activity and embryonic marker reexpression negatively correlate with *INK4A* locus status. In addition to the group of genes down-regulated during hTERT-mediated immortalization and *INK4A* inactivation, we identified a large cluster of genes that show the opposite pattern of expression; that is, their expression was relatively low in primary and *INK4A* locus–intact hTERT-immortalized samples and elevated in cells with inactivated *INK4A*, including both premalignant and malignant samples (Fig. 4*A*; Supplementary Fig. S3).

Among the 250 genes that comprise this cluster, 138 genes were found to be involved in various aspects of cell metabolism and to a large extent (93 of 138) to have a direct role in protein synthesis (Fig. 4*A*). For example, 47 genes encode ribosomal proteins, which belong to the 40S (17 genes) and 60S (30 genes) ribosomal subunits. In addition to the ribosomal proteins, multiple genes that participate in different steps of protein biosynthesis showed increased expression. This group consists of genes associated with amino acid metabolism, amino acid activation (*NARS*, *TARS*, *VARS2*, *FARSL*, *KHSRP*, *CARS*, and *QARS*), translational initiation (*EIF3S7*, *EIF3S5*, *EIF3S3*, *EIF2B1*, etc.), translational elongation (*EEF1B2* and *TUFM*), protein modification and protein folding (*CCT7*, *CCT4*, *PPIA*, *HSPBP1*, and *PFDN2*), and regulation of translation (*ETF1*, *EIF4EBP1*, etc.). A third group of genes up-regulated in these cells and associated with translation are nucleolar proteins (*U5-200KD*, *SNRPD2*, *PABPC4*, *SF3A3*, etc.) that were shown to regulate ribosome assembly and nucleocytoplasmic transport of mature ribosomal subunits (24).

Global up-regulation of genes associated with ribosomal biogenesis and translation could explain the increased proliferation rate of the WI-38/T^{fast} cells versus WI-38 and WI-38/T^{slow}. Indeed, the rate of protein synthesis was found to be a limiting factor in proliferation and growth of cells in several experimental models (24).

Another group of genes that was dramatically induced in all WI-38/T^{fast} samples was enriched for members of the cancer/testis–associated gene family. This group included the subfamilies *MAGE* (10 transcripts), *GAGE* (6 transcripts), *SPANXC*, and *SSX1*. Expression of *SSX1* and *MAGE* was validated by QRT-PCR (Fig. 4*B* and *C*). Indeed, their expression was almost undetectable in primary cells and in the WI-38/T^{slow} cells but was induced >1,000-fold in the WI-38/T^{fast} cells and became even higher in the tumor sample derived from the WI-38/T^{fast}/R/G cells. The expression of these genes was described in many cancers, including sarcomas (25). This similarity suggests that this *in vitro* model reflects some of the physiologic changes that occur during neoplastic initiation in the presumed cell of origin of sarcoma.

Expression of several genes involved in apoptosis was also up-regulated in the WI-38/T^{fast} cells. Among them are both antiapoptotic (*AATF* and *MCL1*) and proapoptotic (*TP53*, *TNFSF7*, *IL24*, *PDCD2*, and *SMAC*) genes. The increase in the expression of these proapototic genes represents the possible activation of a cellular anticancer response at this stage.

As evident from this cluster, the protein biosynthetic pathway and embryonic antigen expression constitute major transcriptional programs, which are abnormally activated during the transition from normal to premalignant cells.

Identification of p53 target genes along the process of malignant transformation: the "proliferation signature" emerges following p53 inactivation in WI-38/T^{fast} cells. p53 inactivation serves as a hallmark of the malignant transformation process. Its role in the onset of cell death, cell cycle regulation, and genome stability has been extensively studied (26). However, the effect of p53 inactivation on the pattern of gene expression at different stages of the transformation process was not addressed previously. In contrast to previous studies, which aimed to identify p53 target genes by p53 overexpression or activation by different stresses in the context of cancer cells (27–29), we inactivated endogenous wild-type p53 protein in the normal cells as well as at the different stages of the transformation process and then searched for down-regulated target genes. By doing a pairwise comparison between isogenic samples isolated at each stage of the transformation process (WI-38, young; WI-38, senescent; WI-38/hTERT^{slow}; WI-38/hTERT^{fast}; and WI-38/hTERT^{fast}/R) and differing only in their p53 status, we identified 210 transcripts that showed at least 1.6 reduction in at least two of six pairs (Table 2; data not shown). As shown in Table 2, well-known p53 direct transcriptional targets, such as *cyclin G1*, *p21^{WAF1}*, *PA26*, *DDB2*, *FDXR*, *WIG1*, and *TNFRS6* (*CD95* or *Fas antigen*), showed reduced expression in all pairs (score 6). When we classified p53-responsive genes into functional groups, we found that p53 in the nonstressed cells regulates genes involved in a plethora of physiologic processes. The most prominent group consists of genes participating in apoptosis, such as *TNFRS6* (*CD95*), *TNFRSF10B* (*KILLER/DR5*), *TNFSF7*, *TNFRSF10D*, *APLP1*, and *NLK*. Additional functional categories involve genes participating in cell cycle control (*cyclin G1*, *p21^{WAF1}*, and *WIG1*), signal transduction (*GRP51*, *PDE5A*, and *RGS20*), extracellular matrix (ECM) and cytoskeleton organization (*COLL11A1*, *TAGLN*, and *ACTA2*), and angiogenesis (*THBS3*). These results suggest that under regular cell culture conditions p53 constitutively transactivates many genes participating in a variety of physiologic processes.

The identification of many known p53 transactivation targets by this approach further supports the value of our experimental model in addressing the importance of p53 inactivation in the

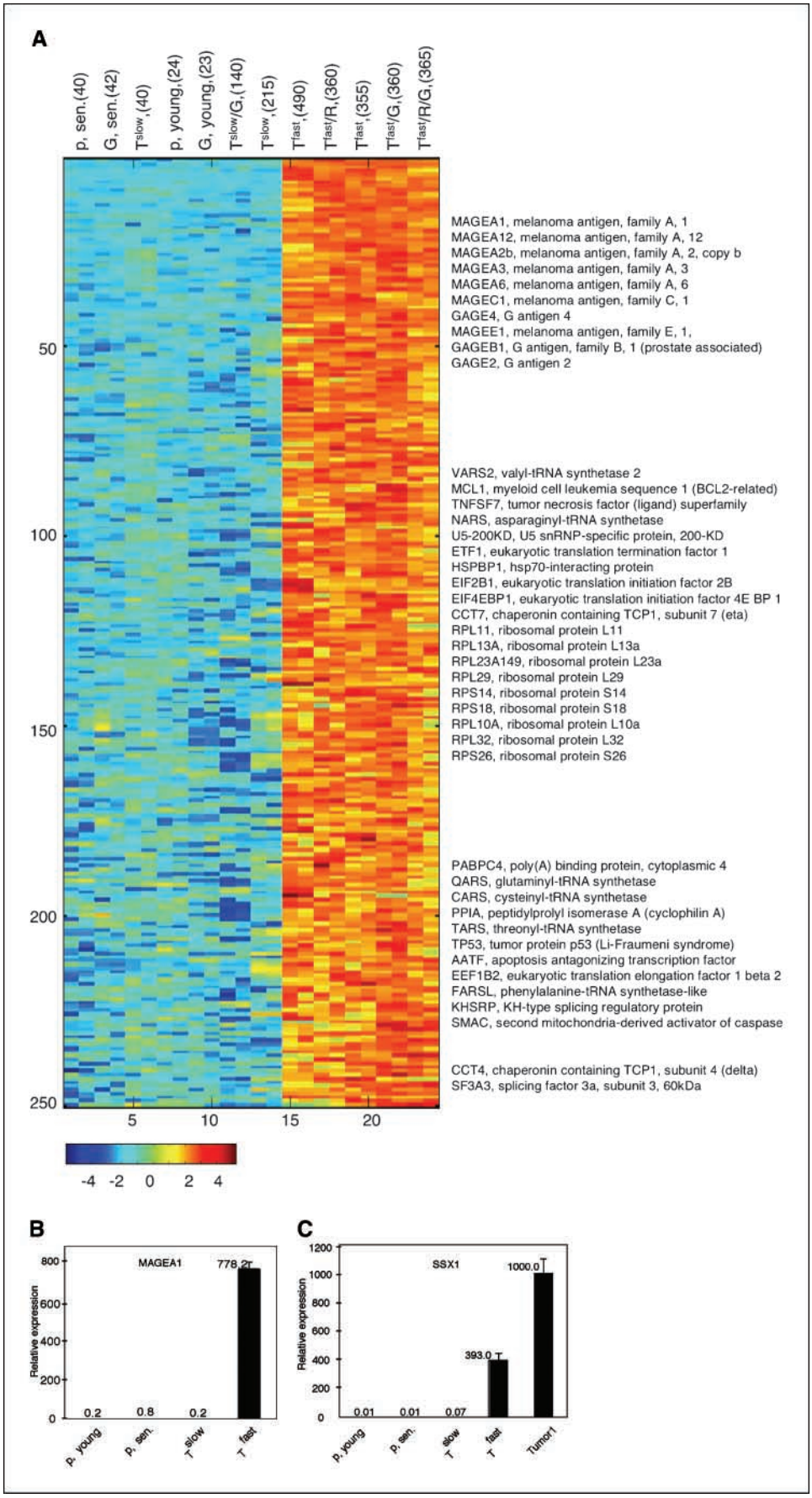


Figure 4. Identification of a genetic signature of up-regulated genes associated with accelerated growth of premalignant and malignant cells. *A,* cluster of up-regulated genes that correlate with *INK4A* locus expression. Set of 250 of 5,581 genes (that passed the filtering criteria using all 24 samples) clustering all samples according to the similarity in their gene expression profiles. This cluster was identified and organized as described in Fig. 3*A*. All the samples with a silenced *INK4A* locus (T^{fast} and its derivatives) showed high expression of this cluster. Selected genes from the cluster, which are discussed in the text, are indicated at the *right* of the cluster. A complete list of genes is available in Supplementary Fig. S3. *B* and *C,* QRT-PCR of *MAGEA1* (*B*) and *SSX1* (*C*) expression in young (*p, young*), senescent (*p, sen.*), T^{slow}, T^{fast}, and Tumor1 cells. Tumor1 cells were recovered from a tumor formed by T^{fast}/R/G cells following injection into nude mice. Gene expression was normalized to the *GAPDH* expression in the same sample. *Columns,* averages of duplicate QRT-PCR measurements.

transformation process. To further understand the downstream effects of p53, we next identified genes that are transrepressed by p53. We found that inactivation of p53 by the dominant-negative peptide GSE56, at every stage of transformation, resulted in a concerted up-regulation of 168 transcripts (Fig. 5*A*; Supplementary Fig. S4). The highest expression was observed when p53 function was ablated in WI-38/T^{fast} cells (T^{fast}/G and T^{fast}/R/G samples). The same group of genes had the lowest expression in senescent WI-38 cells (p, senescent) and was only slightly up-regulated by p53 inactivation (G, senescent). Importantly, their expression patterns are correlated with the proliferation rate of each sample (data not shown). A detailed examination of this gene cluster revealed that it includes mainly genes associated with various aspects of cell proliferation, such as DNA replication, cell cycle progression and its control, DNA repair, and metabolic demands of cell growth. Due to this functional similarity, we termed this group of genes "inactivated p53-associated proliferation signature." Genes that function in the G_2-M phase of the cell cycle represented the largest functional category. More specifically, cyclin-dependent kinase *CDC2* and its regulators such as *cyclin B2*, *cyclin A2*, *CKS1B*, *CKS2*, *CDC25A*, and *CDKN3*, whose function is critical for entrance to mitosis, showed marked up-regulation. In addition, genes with distinctive function in mitosis, including mitotic spindle organization (*TTK* and *kinesins*), mitotic spindle checkpoints (*BUB1*, *BUB1B*, *MAD2L1*, and *BIRC5*), and chromosome segregation (*CDC20*, *CENPF*, *ESPL1*, *UBE2C*, *PLK1*, and *STK12*), were also up-regulated. This cluster also includes genes that are responsible for DNA packaging (*HAT1*, *CHC1*, *SUV39H1*, and *TOP2A*) and chromosome organization (*H1FX*).

In addition to the enhanced expression of the *CDC2* kinase and its positive regulators, two inhibitors of *CDC2*, which are direct p53 transcriptional targets, *CDKN1A* and *GADD45*, showed marked down-regulation (data not shown).

p53 inactivation as well as deregulated expression of *PLK1*, *STK12*, and *CDC2* were shown to promote aneuploidy (30, 31). Furthermore, tetraploid cells were readily detectable in the T^{fast}/R/G as well as in the resulting tumors (ref. 13; data not shown). To analyze if this p53-dependent and transformation progression-dependent up-regulation in the expression of multiple mitotic genes correlates with the development of polyploidy, we exposed cells to nocodazole, a microtubule-depolymerizing drug. Then, the DNA content of the cells was quantitated by FACS analysis. As shown in Fig. 5*B*, normal parental WI-38 (WI-38, puro) fibroblasts were unable to progress through the cell cycle in the presence of nocodazole, whereas their p53-deficient counterpart (WI-38, GSE56) replicated their DNA, resulting in 9% of polyploid cells. In agreement with the increased expression of proliferation signature genes in the WI-38/$hTERT^{fast}$ cells, upon exposure to nocodazole, they produced 12.5% of polyploid cells. Inactivation of p53 in these cells (WI-38/$hTERT^{fast}$/G) resulted in the maximal expression of the proliferation signature, further promoting the fraction of the cells with more than 4N DNA content (20%). Less than 2% of polyploid cells were consistently seen in the untreated samples.

Fifty-one of the 168 transcripts form our inactivated p53-associated proliferation cluster (31% of the genes) belong also to the proliferation cluster found by Rosty et al. (32) in cervical cancer samples. They reported that the expression levels of the genes of their cluster were predictive of outcome in cervical cancer. To evaluate the clinical significance of our inactivated p53-associated proliferation signature genes, we tested its ability to predict the clinical outcome of cancer patients. To this end, we studied their expression levels in the large breast cancer gene expression data set of (33). We found that 71 genes from our proliferation cluster (42%) showed alteration in the 98 primary breast cancers included in the breast cancer study. More significantly, we were able to segregate all patients according to the expression of this signature into two groups (Fig. 5*C*). Those tumors that were characterized by a high level of expression had a significantly higher risk of death due to development of distant metastases within 5 years ($P = 4.9 \times 10^{-5}$). These results show that this inactivated p53-associated proliferation signature contributes to tetraploidy generation and has a strong predictive value regarding aggressive tumor behavior.

"Tumor-forming" genetic signature. Inactivation of wild-type p53 and expression of constitutively active H-Ras conferred WI-38/$hTERT^{fast}$ cells with tumorigenic potential (Fig. 1). Therefore, identification of unique transcriptional changes in the WI-38/T^{fast}/G/R cells is of great significance for understanding the transforming efficiency of both genetic hits. To this end, we carried out pairwise comparisons of WI-38/T^{fast}/Neo cells with their counterparts expressing dominant-negative p53 (GSE56), H-Ras, or a combination of both. Next, we identified a group of genes that showed the highest extent of Ras/GSE56 synergism; that is, their expression in the T^{fast}/G/R sample compared with the T^{fast}/Neo control had a higher fold change than the additive effects of H-Ras and GSE56 alone (Table 3). This tumor-forming genetic signature was composed mainly of secreted molecules that belong to the CXC chemokine family (*CXCL1*, *CXCL2*, *IL8*, *CXCL6*, and *CXCL10*), cytokines (*IL1B*, *IL6*, and *CSF2*), and modifiers of ECM (*TFPI2*, *MMP3*, *PRSS2*, *C1QTNF1*, *PRSS3*, and *ADAMTS8*).

The CXC chemokines were shown to play a role in processes essential for tumor growth, such as autocrine stimulation, angiogenesis, invasion and metastasis (34). High levels of expression and secretion of these chemokines was detected in melanoma, colon, pancreatic and breast tumors. In some cases, a direct correlation between metastatic potential and *CXCL8* (*IL8*) secretion was described (35).

Because we observed the strongest synergism between H-Ras and p53 inactivation on the expression of *CXCL1*, we used QRT-PCR to study its expression in independent samples derived from WI-38 at different stages of transformation (Fig. 6). We found that expression of *CXCL1* was extremely low in young WI-38 fibroblasts as well as in WI-38/T^{slow} and WI-38/T^{fast}/Neo cells. Inactivation of p53 or expression of constitutively active Ras up-regulated *CXCL1* ~10- and 36-fold, respectively, compared with the Neo control. However, the combination of p53 inactivation and constitutive Ras expression resulted in >10,000-fold induction. Importantly, expression of *CXCL1* was even further up-regulated in cells retrieved from a tumor sample originating from WI-38/T^{fast}/G/R, suggesting that this chemokine confers a selective advantage in tumor formation. We detected a similar pattern of expression of the *IL1B* gene (Fig. 6). The regulation of these secreted factors by H-Ras and p53 was not reported previously, to the best of our knowledge.

ECM turnover, which is governed mainly by matrix metalloproteinases (MMP), is essential for the ability of malignant cells to promote neovascularization, invasion, and metastasis (36). The activities of MMPs are under tight control of the physiologic tissue inhibitors of MMPs (TIMP). We observed profound changes in the expression pattern of several MMPs and TIMPs following

Table 2. Genes whose expression was down-regulated upon endogenous p53 inactivation

Gene symbol	Description	Down-regulation upon p53 inactivation at the selected stage of the transformation*						Sum score †
		1	2	3	4	5	6	
CCNG1	Cyclin G1	1	1	1	1	1	1	6
CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)	1	1	1	1	1	1	6
PA26	p53-regulated PA26 nuclear protein	1	1	1	1	1	1	6
DDB2	Damage-specific DNA binding protein 2, 48 kDa	1	1	1	1	1	1	6
FDXR	Ferridoxin reductase	1	1	1	1	1	1	6
WIG1	p53 target zinc finger protein	1	1	1	1	1	1	6
TNFRSF6	TNF receptor superfamily, member 6 (CD95, Fas antigen)	1	1	0	1	1	1	5
C12orf5	Chromosome 12 open reading frame 5	1	1	1	1	1	0	5
ANXA4	Annexin A4	1	1	0	1	0	1	4
TNFRSF10B	TNF receptor superfamily, member 10B (KILLER/DR5)	1	1	1	1	0	0	4
APLP1	Amyloid β (A4) precursor-like protein	0	1	1	1	1	0	4
GAMT	Guanidinoacetate *N*-methyltransferase	1	0	1	1	0	1	4
COL11A1	Collagen, type XI, α1	1	1	0	1	1	0	4
GPR51	G protein-coupled receptor 51	1	1	1	1	0	0	4
PDE5A	Phosphodiesterase 5A, cyclic guanosine 3′,5′-monophosphate specific	1	0	1	1	0	1	4
PSTPIP2	Proline-serine-threonine phosphatase interacting protein 2	0	1	1	0	1	1	4
TAGLN	Transgelin	0	0	1	0	1	1	3
ACTA2	Actin, α2, smooth muscle, aorta	1	0	0	0	1	1	3
MFGE8	Milk fat globule-epidermal growth factor 8 protein	1	0	1	1	0	0	3
OSF-2	Osteoblast-specific factor 2 (fasciclin I-like)	1	0	0	0	1	1	3
SERPINE1	Serine (or cysteine) proteinase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	0	0	1	1	1	0	3
NET-6	Transmembrane 4 superfamily member tetraspan NET-6	0	0	1	1	1	0	3
CUL4B	Cullin 4B	0	0	1	1	0	1	3
RNASE6PL	RNase 6 precursor	0	0	1	1	1	0	3
NLK	Nemo-like kinase	0	0	1	1	0	1	3
PSG5	Pregnancy-specific β1-glycoprotein 5	1	0	0	0	1	1	3
UGCG	UDP-glucose ceramide glucosyltransferase	0	0	0	1	1	1	3
THBS3	Thrombospondin 3	1	0	0	1	0	1	3
TM7SF2	Transmembrane 7 superfamily member 2	1	0	0	1	0	1	3
COL4A5	Collagen, type IV, α5 (Alport syndrome)	0	1	0	0	1	1	3
COPZ2	Coatomer protein complex, subunit ζ2	1	1	0	1	0	0	3
CYP24	Cytochrome *P*450, subfamily XXIV (vitamin D 24-hydroxylase)	0	1	0	0	1	1	3
APG-1	Heat shock protein (hsp110 family)	0	1	1	1	0	0	3
RGS20	Regulator of G-protein signaling 20	1	1	0	0	0	1	3
TNFSF7	TNF (ligand) superfamily, member 7	1	0	0	0	1	1	3
KCNN2	Potassium intermediate/small conductance calcium-activated channel, subfamily N, member 2	1	0	1	1	0	0	3
PRSC	Corin	1	1	0	1	0	0	3
TNFRSF10D	TNF receptor superfamily, member 10D, decoy with truncated death domain	1	0	1	1	0	0	3

NOTE: Pairwise comparison between normalized values of each gene expression was done between the above samples infected with either control virus or GSE56-containing one (G). Gene that showed at least 1.6-fold reduction in the GSE56-containing sample in the corresponding pairwise comparison received score 1; otherwise, the score was 0.

*Comparisons were listed in the following order: (1) WI-38/p, young versus WI-38/G, young; (2) WI-38/p, senescent versus WI-38/G, senescent; (3) WI-38/T^{slow} (40 PDLs) versus WI-38/T^{slow}/G (140 PDLs); (4) WI-38/T^{slow} (215 PDLs) versus WI-38/T^{slow}/G (140 PDLs); (5) WI-38/T^{fast} versus WI-38/T^{fast}/G; and (6) WI-38/T^{fast}/R versus WI-38/T^{fast}/R/G.

† Only genes that showed significant changes in at least three of six comparisons (sum score 3) were enlisted.

introduction of H-Ras and GSE56 (Table 3). The expression of MMP3 was further validated by QRT-PCR analysis (Fig. 6). Once again, we observed a synergistic effect of H-Ras expression and p53 inactivation on *MMP3* induction, and the high expression persisted during *in vivo* growth of the corresponding tumor. Of note, matrix *MMP1* and *MMP3* were reported to be targets of Ras transformation (37).

Thus, it seems that strong induction of proangiogenic and autocrine chemokines in concert with potent ECM modifiers is characteristic of the tumor-forming genetic signature. Those genes could form the basis of the *in vivo* tumorigenic potential conferred on cells by expression of H-Ras and by p53 ablation.

Discussion

Deregulated transcriptional programs resulting from the activation of oncogenes and inactivation of tumor suppressors underlie many important aspects of cancer. In this study, primary human fibroblasts were the cell of origin, gradually transformed to a tumorigenic state. Then, complex and multistep biological phenomenon of transformation was dissected to a distinct number of transcriptional programs that exhibit well-defined orderly temporal organization. This was achieved by the identification of stable clusters of gene expression along defined stages of the transformation process. The identification of these specific genetic signatures reflects the acquisition of specific physiologic features

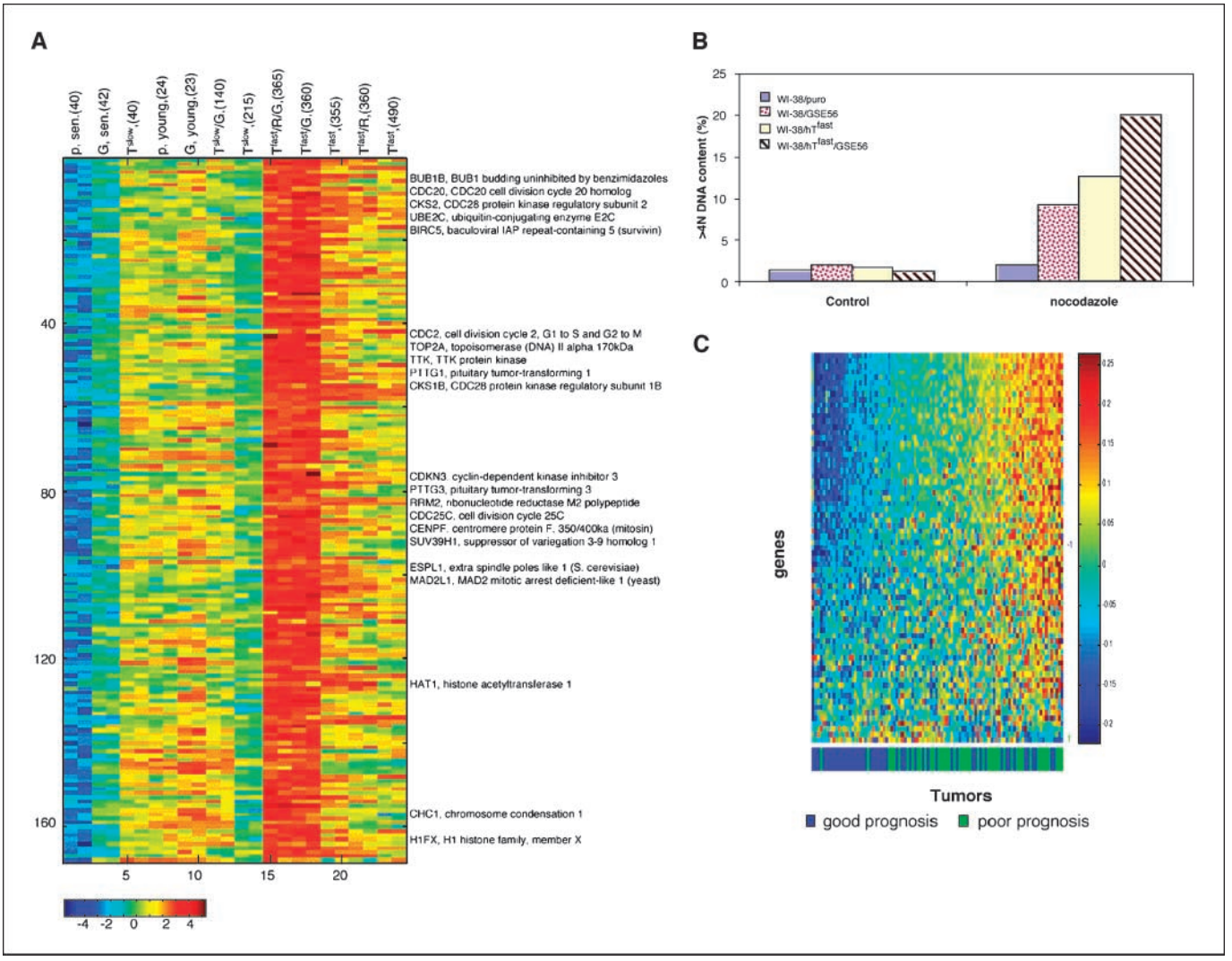


Figure 5. Identification of the inactivated p53-associated proliferation signature. *A,* a cluster of 168 genes, which showed similar patterns of expression and was associated with p53 inactivation and increased cell proliferation rate. A complete list of genes is available in Supplementary Fig. S4. Note that expression is highest in the T^{fast}/GSE56 (T^{fast}/G) and lowest in the WI-38/puro senescent cells. This cluster was identified and organized as described in Fig. 3*A*. Selected genes from the cluster, which are discussed in the text, are indicated. *B,* effect of p53 inactivation on the different stages of transformation on DNA rereplication in response to nocodazole. Asynchronous cultures of WI-38/puro, WI-38/GSE56, WI-38/hTERT^{fast}, and WI-38/hTERT^{fast}/GSE56 were seeded into medium with or without 0.05 μg/mL nocodazole. After 72 hours, the cells were fixed and stained with PI for flow cytometry (FACS) analysis. Cells with DNA content greater that 4N were considered as polyploid and were quantitated by gating cells to the right of the 4N peak on histogram plots. *C,* prognostic value of the inactivated p53-associated proliferation signature in breast cancer patients described in the study of van't Veer at al. (33). Gene symbols of the inactivated p53-associated proliferation signature (168 genes) were intersected with the gene symbols of the breast cancer profiling experiments. From our panel of 168 genes (Fig. 5*A*), 71 were also found in the study of van't Veer at al. Then, a new expression matrix of 71 × 96 was created, such that each *column* represented a patient sample and each *row* represented a gene. Each column was labeled according to the patient prognosis: *blue,* good prognosis (no distant metastasis were developed within 5 years); *green,* poor prognosis (distant metastasis developed within 5 years). The samples were then sorted in an unsupervised manner according to their expression using the SPIN algorithm (58). The *P* for the separation was calculated using the Wilkinson rank-sum test. Tumors with high inactivated p53-associated proliferation signature had worse prognosis than tumors with low inactivated p53-associated proliferation signature ($P = 4.9 \times 10^{-5}$).

Table 3. Tumor-forming genetic signature

Functional category, gene symbol and name	Fold induction relative to T^{fast}		
	GSE56	H-Ras	GSE/H-Ras
Chemokines and cytokines			
CXCL1, chemokine (C-X-C motif) ligand 1*	1.24	2.8	68.5
CXCL3, chemokine (C-X-C motif) ligand 3	0.65	2.26	21.3
CXCL2, chemokine (C-X-C motif) ligand 2	2.64	4.24	12.7
IL1B, IL-1β*	1.17	4.34	12
IL8, IL-8	2.90	4.19	11.3
CSF2, colony-stimulating factor 2 (granulocyte-macrophage)	0.32	8.70	9.93
IL6, interleukin-6 (IFN, β2)	2.11	2.54	9.83
CXCL6, chemokine (C-X-C motif) ligand 6	0.87	1.04	2.76
ECM modulators			
MMP3, matrix metalloproteinase 3 (stromelysin 1)*	1.16	3.74	7.73
PRSS3, protease, serine, 3 (mesotrypsin)	1.33	1.66	2.34
TIMP1, tissue inhibitor of metalloproteinase 1	1.60	1.52	2.25
PRSS2, protease, serine, 2 (trypsin 2)	2.42	2.13	3.77
ADAMTS8, a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 8	0.96	1.49	2.31
Cell death			
BCL2A1, BCL2-related protein A1	1.43	5.85	9.37
BIRC3, baculoviral IAP repeat-containing 3	0.67	1.17	2.79
CST1, cystatin SN	1.22	3.99	7.71

NOTE: Selected genes whose expression was affected in a cooperative manner by p53 inactivation and H-Ras overexpression. The fold change is calculated by pairwise comparison between normalized values of each gene expression in the corresponding samples.
*Genes for which the pattern of expression was validated by QRT-PCR (Fig. 5).

essential for initiation and progression of mesenchymal cell transformation (Fig. 7). It is important to stress that transformation of other lineages, such as epithelial or myeloid, by similar combination of genetic elements might lead to the discovery of different genetic signatures. Equally possible is that other combinations of transforming genes might result in additional transformation fingerprints. Nevertheless, recently, two studies (38, 39) found expression modules shared between many cancer types, suggesting common tumor progression mechanisms and essential transcriptional features that evolve along neoplastic transformation of different origins.

Disruption of the fine balance between differentiation and proliferation is one of the hallmarks common to many tumor types. However, molecular markers that distinguish between differentiation-proficient and differentiation-deficient cells are still elusive in many cases. We found here a particular genetic signature that describes the defect in the smooth muscle differentiation program. Furthermore, a correlated expression pattern between the Rb regulators $p16^{INK4A}$ and $p57^{KIP2}$ and several molecules associated with myogenic differentiation was found. Importantly, alterations in the expression of mesenchymal cell developmental and differentiation markers as well as inactivation of $p16^{INK4A}$ and $p57^{KIP2}$ were observed previously in tissues similar to WI-38 fibroblasts, such as synovial and other types of soft tissue sarcomas (40–45). Our results strongly suggest that this genetic signature provides a mechanistic link between the disruption of the cell cycle regulation and inability to properly differentiate. The fact that this genetic signature persisted in the increasingly transformed cell populations suggests its active contribution to the more aggressive phenotype as well. Taking into consideration that the cell of origin for fibrosarcoma is still obscure (46), acquisition of a differentiation defect relevant to fibroblasts, as we found, suggests that human lung embryonic fibroblasts (such as WI-38) could be a cell of origin for this tumor type.

Our gene expression array analysis allowed us to identify a list of genes down-regulated upon p53 inactivation in the context of nonstressed cells, among them are both novel and known p53 targets genes participating in a variety of physiologic processes. Those genes may be of primary importance in the understanding the role of p53 in the normal cell cycle. Keeping in mind that germ line mutations in the p53 gene (as in Li-Fraumeni syndrome) predispose to development of sarcoma (47), we suggest that at least some of the genes we identified as endogenous p53 targets may be crucial for this carcinogenic process. Constitutive transactivation of several tumor necrosis factor-α (TNF-α) apoptotic genes, such as *TNFRSF6*, *TNFRSF10*, *TNFSF7*, and *TNFRSF10D*, by basal p53 levels is particularly exciting. It could provide a testable hypothesis that p53 inactivation confers cells with the increased resistance to death stimuli mediated by TNF-α and its family members.

Acquisition of genomic instability is an additional hallmark of malignant transformation. We identified a unique genetic signature associated with p53 inactivation and development of increased polyploidy. This signature consists mainly of genes regulating cell cycle progression and mitosis. Importantly, maximum expression of these genes required the INK4A-deficient background, suggesting a novel level of cooperation between p53 and $p16^{INK4A}$ tumor suppressors. Phenotypically, our finding supports existence of this cooperation; that is, duplication of DNA in the presence of

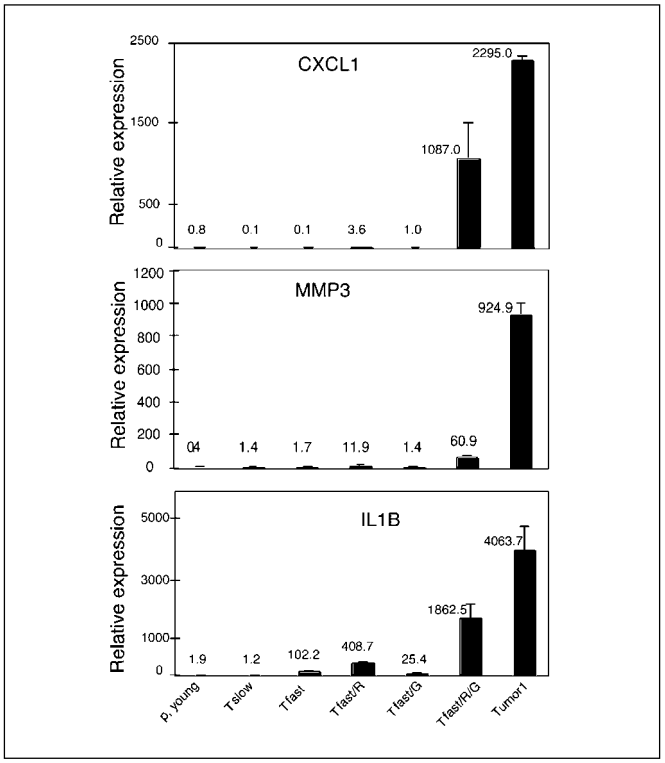


Figure 6. QRT-PCR validation of the tumor-forming genetic signature. Expression of *CXCL1* (*top*), *MMP3* (*midlle*), and *IL1B* (*bottom*) as representative genes from the tumor-forming genetic signature (see also Table 3) was tested at different stages of transformation. Sample stages are detailed in Table 1. Tumor1 cells were obtained from a tumor formed by T^{fast}/R/G cells on injection into nude mice. Gene expression was normalized to *GAPDH* expression in the same sample. *Columns,* averages of duplicate QRT-PCR measurements.

mitotic inhibitor attained maximal levels when both tumor suppressors were inactivated. Several reports showed that alterations in mitosis genes found in our cluster (31, 48) facilitate the acquisition of additional mutations, thereby further promoting cancer progression. Taking into account that tetraploid metaphases were readily identified in T^{fast}/G and T^{fast}/R/G cells and in the tumor samples retrieved from mice, we suggest that inactivation of p53 in the INK4A-deficient cells primarily leads to the acquisition of chromosomal instability, a key event in human tumorigenesis (49–51).

The ability to offer accurate survival prognosis of cancer patients is one of the critical issues in cancer medicine. According to our results, the inactive p53-associated proliferation signature is a strong predictor of a poor outcome in several common solid tumors, including breast (Fig. 5*C*) and prostate carcinomas (data not shown). Similar proliferation signatures were identified in several studies as a partial characteristic of highly proliferative and more aggressive tumors (52, 53). In addition, a similar genetic signature was found in the tetraploid progenitors of esophageal cancer, which contained mutant p53 (54). Based on these facts, an apparent link between a high rate of proliferation and tumor aggressiveness could be predicted. Finally, these observations suggest that this particular signature, which we have identified by defined genetic manipulations *in vitro*, represents an authentic physiologic genetic program common to many cancer types. This transformation fingerprint is mainly regulated by the p53 and $p16^{INK4A}$/pRb tumor suppressors and affects both genome destabilization and the malignant potential of cells.

Tumorigenicity in our strain of fibroblasts is strictly dependent on mutant H-Ras expression and p53 inactivation, suggesting that both oncogenic events are required for this malignant phenotype (55, 56). Indeed, we identified a specific group of genes, which was induced in a highly synergistic manner only when both H-Ras

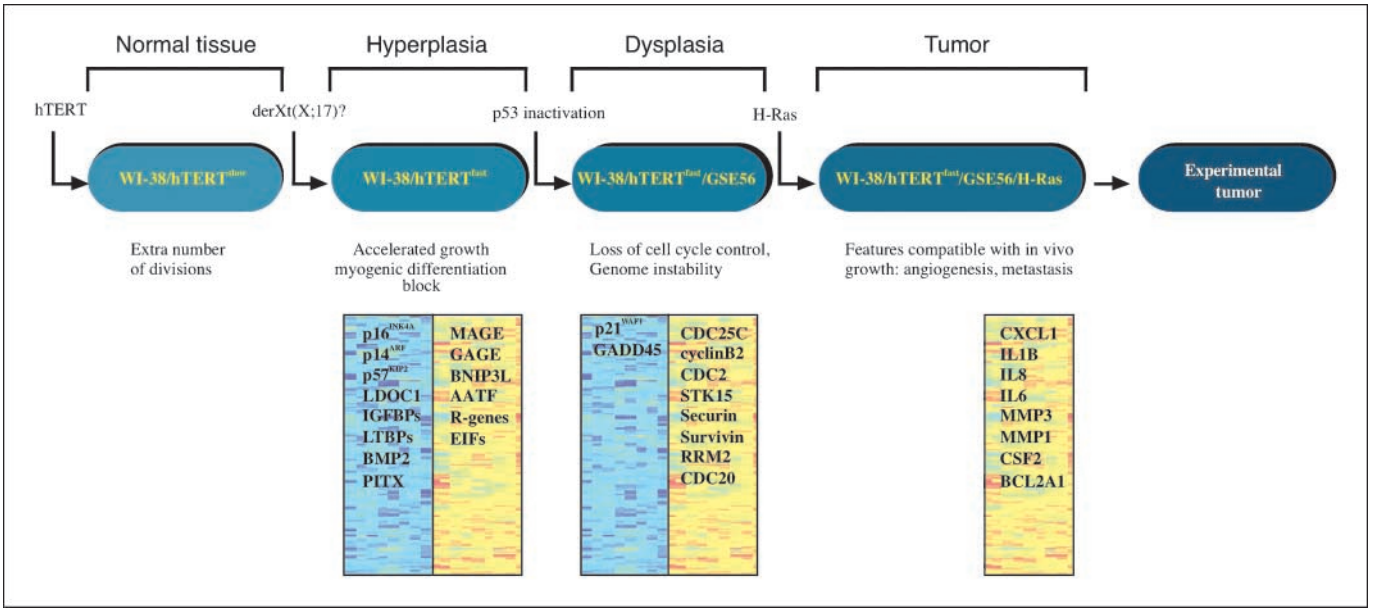


Figure 7. Outline of a stepwise malignant transformation process based on the transcriptional programs identified in this study. Microarray profiling revealed specific genetic signatures associated with the particular stages in our *in vitro* transformation model. (Selected genes are shown in the *boxes* colored according to their expression level: *yellow-red,* high expression; *blue,* low expression.) These alterations in gene expression reflect the biological features spontaneously acquired by cells [der(X)t(X;17) and *INK4A* locus silencing] or induced by engineered mutations (GSE56 and H-Ras) along the transformation process. We suggest that the genetic signatures identified in our study provide a conceptual framework for similar transcriptional alterations associated with the transition from normal tissue to hyperplasia, dysplasia, and then to cancer.

and p53 inactivation occurred (Table 3). This tumor-forming genetic signature involves important genes required for angiogenesis, autocrine stimulation, and metastasis. The expression of genes that enable neoplastic cells to modulate their stromal environment represents a critical stage in the tumorigenic process. There are several possible explanations for the transcriptional synergism we observed between H-Ras and p53 in inducing the tumor-forming genetic signature: it is possible that both activated H-Ras and p53 deficiency are required for maximal activation of a single key transcription factor, such as nuclear factor-κB (NF-κB). An alternative possibility could be comodulation of distinct transcriptional factors (e.g., activator protein-1, NF-κB, and CBP/p300) that are involved in the transcriptional regulation of those genes (34). It is also possible that similar transcriptional programs provide the basis for the long-known cooperative effect between pairs of oncogenes in transformation (56, 57). Furthermore, modulation of the signal transduction pathway elements downstream to p53 and Ras provide a promising avenue for future therapy.

Although many of the genes identified in our study were already known to be associated with cancer, the novelty of our findings lies in the fact that we identified specific clusters of genes that underlie the acquisition of specific transformation hallmarks. It seems that fully transformed cells contain a limited number of deregulated transcriptional programs. We hypothesize that the information obtained from our *in vitro* model of malignant transformation accurately reflects, to the extent one could modulate *in vitro*, the changes that occur during tumor initiation and progression *in vivo*. Such common features should be carefully evaluated in naturally occurring malignancies. More significantly, our data provide new knowledge essential for better understanding of the transcriptional programs that underlie the transformed state.

Acknowledgments

Received 10/28/2004; revised 2/13/2005; accepted 3/23/2005.

Grant support: Israel-USA Binational Science Foundation, Ridgefield Foundation, Flight Attendant Medical Research Institute, and NIH grant 5 POI CA 65930-06.

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 Dr. Shirley Horn-Saban (Head of the DNA Microarray Unit, Crown Human Genome Center, Weizmann Institute of Science) for microarray processing; Ezra Vadai, Raanan Shaked, and Dr. Ori Brenner (Veterinary Resources Department, Weizmann Institute of Science) for assistance with the *in vivo* experiments; and the members of V. Rotter's group for fruitful discussion.

References

1. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759–67.
2. Liotta L, Petricoin E. Molecular profiling of human cancer. Nat Rev Genet 2000;1:48–56.
3. Getz G, Levine E, Domany E. Coupled two-way clustering analysis of gene microarray data. Proc Natl Acad Sci U S A 2000;97:12079–84.
4. Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 2000;24:227–35.
5. Scherf U, Ross DT, Waltham M, et al. A gene expression database for the molecular pharmacology of cancer. Nat Genet 2000;24:236–44.
6. Rosenwald A, Wright G, Wiestner A, et al. The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. Cancer Cell 2003;3:185–97.
7. Elenbaas B, Spirio L, Koerner F, et al. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. Genes Dev 2001;15:50–65.
8. Hahn WC, Counter CM, Lundberg AS, Beijersbergen RL, Brooks MW, Weinberg RA. Creation of human tumour cells with defined genetic elements. Nature 1999;400:464–8.
9. Voorhoeve PM, Agami R. The tumor-suppressive functions of the human INK4A locus. Cancer Cell 2003;4:311–9.
10. Drayton S, Rowe J, Jones R, et al. Tumor suppressor $p16^{INK4a}$ determines sensitivity of human cells to transformation by cooperating cellular oncogenes. Cancer Cell 2003;4:301–10.
11. Akagi T, Sasai K, Hanafusa H. Refractory nature of normal human diploid fibroblasts with respect to oncogene-mediated transformation. Proc Natl Acad Sci U S A 2003;100:13567–72.
12. Rangarajan A, Hong SJ, Gifford A, Weinberg RA. Species- and cell type-specific requirements for cellular transformation. Cancer Cell 2004;6:171–83.
13. Milyavsky M, Shats I, Erez N, et al. Prolonged culture of telomerase-immortalized human fibroblasts leads to a premalignant phenotype. Cancer Res 2003;63:7147–57.
14. Taylor LM, James A, Schuller CE, Brce J, Lock RB, Mackenzie KL. Inactivation of $p16^{INK4a}$, with retention of pRB and $p53/p21^{cip1}$ function, in human MRC5 fibroblasts that overcome a telomere-independent crisis during immortalization. J Biol Chem 2004;279:43634–45.
15. Tsutsui T, Kumakura S, Yamamoto A, et al. Association of p16(INK4a) and pRb inactivation with immortalization of human cells. Carcinogenesis 2002;23:2111–7.
16. Noble JR, Zhong ZH, Neumann AA, Melki JR, Clark SJ, Reddel RR. Alterations in the p16(INK4a) and p53 tumor suppressor genes of hTERT-immortalized human fibroblasts. Oncogene 2004;23:3116–21.
17. Chambers RC, Leoni P, Kaminski N, Laurent GJ, Heller RA. Global expression profiling of fibroblast responses to transforming growth factor-β1 reveals the induction of inhibitor of differentiation-1 and provides evidence of smooth muscle cell phenotypic switching. Am J Pathol 2003;162:533–46.
18. Blatt M, Wiseman S, Domany E. Superparamagnetic clustering of data. Phys Rev Lett 1996;76:3251–4.
19. Ossovskaya VS, Mazo IA, Chernov MV, et al. Use of genetic suppressor elements to dissect distinct biological effects of separate p53 domains. Proc Natl Acad Sci U S A 1996;93:10309–14.
20. Oklu R, Hesketh R. The latent transforming growth factor β binding protein (LTBP) family. Biochem J 2000;352:601–10.
21. Grimberg A, Cohen P. Role of insulin-like growth factors and their binding proteins in growth control and carcinogenesis. J Cell Physiol 2000;183:1–9.
22. Abate-Shen C. Deregulated homeobox gene expression in cancer: cause or consequence? Nat Rev Cancer 2002;2:777–85.
23. Massague J, Blain SW, Lo RS. TGFβ signaling in growth control, cancer, and heritable disorders. Cell 2000;103:295–309.
24. Ruggero D, Pandolfi PP. Does the ribosome translate cancer? Nat Rev Cancer 2003;3:179–92.
25. Dalerba P, Frascella E, Macino B, et al. MAGE, BAGE and GAGE gene expression in human rhabdomyosarcomas. Int J Cancer 2001;93:85–90.
26. Oren M. Decision making by p53: life, death and cancer. Cell Death Differ 2003;10:431–42.
27. Zhao R, Gish K, Murphy M, et al. Analysis of p53-regulated gene expression patterns using oligonucleotide arrays. Genes Dev 2000;14:981–93.
28. Kannan K, Amariglio N, Rechavi G, et al. DNA microarrays identification of primary and secondary target genes regulated by p53. Oncogene 2001;20:2225–34.
29. Yoon H, Liyanarachchi S, Wright FA, et al. Gene expression profiling of isogenic cells with different TP53 gene dosage reveals numerous genes that are affected by TP53 dosage and identifies CSPG2 as a direct target of p53. Proc Natl Acad Sci U S A 2002;99:15632–7.
30. Di Leonardo A, Khan SH, Linke SP, Greco V, Seidita G, Wahl GM. DNA rereplication in the presence of mitotic spindle inhibitors in human and mouse fibroblasts lacking either p53 or pRb function. Cancer Res 1997;57:1013–9.
31. Jallepalli PV, Lengauer C. Chromosome segregation and cancer: cutting through the mystery. Nat Rev Cancer 2001;1:109–17.
32. Rosty C, Sheffer M, Tsafrir D, et al. Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. Oncogene. In press 2005.
33. 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.
34. Dhawan P, Richmond A. Role of CXCL1 in tumorigenesis of melanoma. J Leukoc Biol 2002;72:9–18.
35. Luca M, Huang S, Gershenwald JE, Singh RK, Reich R, Bar-Eli M. Expression of interleukin-8 by human melanoma cells up-regulates MMP-2 activity and increases tumor growth and metastasis. Am J Pathol 1997;151:1105–13.
36. Jiang Y, Goldberg ID, Shi YE. Complex roles of tissue inhibitors of metalloproteinases in cancer. Oncogene 2002;21:2245–52.
37. Zuber J, Tchernitsa OI, Hinzmann B, et al. A genome-wide survey of RAS transformation targets. Nat Genet 2000;24:144–52.
38. Segal E, Friedman N, Koller D, Regev A. A module map showing conditional activity of expression modules in cancer. Nat Genet 2004;36:1090–8.
39. Rhodes DR, Yu J, Shanker K, et al. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proc Natl Acad Sci U S A 2004;101:9309–14.

40. Segal NH, Pavlidis P, Antonescu CR, et al. Classification and subtype prediction of adult soft tissue sarcoma by functional genomics. Am J Pathol 2003;163:691–700.
41. Nielsen TO, West RB, Linn SC, et al. Molecular characterisation of soft tissue tumours: a gene expression study. Lancet 2002;359:1301–7.
42. Nagayama S, Katagiri T, Tsunoda T, et al. Genome-wide analysis of gene expression in synovial sarcomas using a cDNA microarray. Cancer Res 2002;62:5859–66.
43. Allander SV, Illei PB, Chen Y, et al. Expression profiling of synovial sarcoma by cDNA microarrays: association of ERBB2, IGFBP2, and ELF3 with epithelial differentiation. Am J Pathol 2002;161:1587–95.
44. Dauphinot L, De Oliveira C, Melot T, et al. Analysis of the expression of cell cycle regulators in Ewing cell lines: EWS-FLI-1 modulates p57KIP2and c-Myc expression. Oncogene 2001;20:3258–65.
45. Deneen B, Denny CT. Loss of p16 pathways stabilizes EWS/FLI1 expression and complements EWS/FLI1 mediated transformation. Oncogene 2001; 20:6731–41.
46. Helman LJ, Meltzer P. Mechanisms of sarcoma development. Nat Rev Cancer 2003;3:685–94.
47. Malkin D, Li FP, Strong LC, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 1990;250: 1233–8.
48. Taylor WR, Stark GR. Regulation of the G_2/M transition by p53. Oncogene 2001;20:1803–15.
49. Shih IM, Zhou W, Goodman SN, Lengauer C, Kinzler KW, Vogelstein B. Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis. Cancer Res 2001;61:818–22.
50. Barrett MT, Sanchez CA, Prevo LJ, et al. Evolution of neoplastic cell lineages in Barrett oesophagus. Nat Genet 1999;22:106–9.
51. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature 1998;396: 643–9.
52. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406: 747–52.
53. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000;403:503–11.
54. Barrett MT, Pritchard D, Palanca-Wessels C, Anderson J, Reid BJ, Rabinovitch PS. Molecular phenotype of spontaneously arising 4N (G2-tetraploid) intermediates of neoplastic progression in Barrett's esophagus. Cancer Res 2003;63:4211–7.
55. McCormick F. Signalling networks that cause cancer. Trends Cell Biol 1999;9:M53–6.
56. Hahn WC, Weinberg RA. Modelling the molecular circuitry of cancer. Nat Rev Cancer 2002;2: 331–41.
57. Land H, Parada LF, Weinberg RA. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature 1983;304: 596–602.
58. Tsafrir D, Tsafrir I, Em-Dor L, et al. Sorting points into neighborhoods (SPIN): data analysis and visualization by ordering distance matrices. Bioinformatics. In press 2005.