Clinicopathologic Correlation of Up-regulated Genes Identified Using cDNA Microarray and Real-time Reverse Transcription-PCR in Human Colorectal Cancer

Sou-Tyau Chiu,[1,3] Fon-Jou Hsieh,[4] Shi-Wen Chen,[1] Chun-Lieh Chen,[1] Hwei-Fan Shu,[2] and Hung Li[5]

Departments of [1]Surgery and [2]Pathology, Feng-Yuan Hospital, Feng-Yuan City, Taichung County; [3]Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University; [4]Department of Obstetrics and Gynecology, National Taiwan University Hospital; and [5]Institute of Molecular Biology, Academica Sinica, Taipei, Taiwan

Abstract

Purpose: **We hypothesize that changes in the transcription of up-regulated genes are biologically meaningful and may be linked to variations in tumor behavior and clinical features. This study aimed to find individual up-regulated genes responsible for clinicopathologic variations in human colorectal cancer.**

Experimental Design: **Genes up-regulated concurrently in four microarray experiments were taken as candidate genes; 20 candidate genes were verified using real-time reverse transcription-PCR in these four experiments, along with 27 new samples. The presence or absence of up-regulation of these genes was correlated with 10 clinicopathologic variables from 31 patients. The mRNA transcript levels of these 20 candidate genes in the 31 paired samples were also correlated with each other to disclose any expression relationship.**

Results: **Forty percent (8/20) of the candidate genes were verified by real-time reverse transcription-PCR to have a tumor/normal expression ratio > 2. Up-regulation of *THY1* and *PHLAD1* was associated with the presence of anemia in colon cancer patients ($P = 0.036$ and 0.009, respectively). Up-regulation of *HNRPA1* was more significant in cancer growing in the right-sided colon than the left side ($P = 0.027$). Up-regulated *GPX2* was related to a higher degree of tumor differentiation ($P = 0.019$). *c-MYC* was significantly over-expressed in specimens from male compared with female colon cancer patients ($P = 0.012$). *GRO1* was significantly up-regulated in patients younger than 65 years old ($P = 0.010$) and was found to be frequently over-expressed when cancers were less invasive. In addition, we found that normalized transcript levels of *HNRPA1* were tightly associated with that of *c-MYC* ($r = 0.948$).**

Conclusions: **Validation of microarray data using another independent laboratory approach is mandatory and statistical correlation between gene expression status and the patient's clinical features may reveal individual genes relevant to tumor behavior and clinicopathologic variations in human colorectal cancer.** (Cancer Epidemiol Biomarkers Prev 2005;14(2):437–43)

Introduction

Colorectal cancer is one of the leading causes of fatalities due to cancer in developed countries. Much effort has been devoted into researching its underlying molecular events. Understanding the differences in gene expression between cancerous and normal cells is important for cancer research. cDNA microarray technologies can perform analyses of such differences at a transcriptional level with high throughput. Using microarrays, researchers have unveiled gene expression profiles and signatures that characterize colorectal cancer and its phenotypes (1-3). However, the individual genes responsible for tumor behavior and clinicopathologic variations in colon cancer have not been reported thus far.

The goal of this study was to identify up-regulated genes in human colorectal cancer and determine if there are individual genes linked to tumor behavior and patient clinicopathologic factors. A simple and cost-saving strategy was adopted in this study to identify up-regulated genes: cDNA microarrays were taken as a genomic filter and real-time reverse transcription-PCR as a validation method for verifying array data; a large number tissue samples ($n > 30$) were used in post-array validation to minimize sample bias. To find the relevance between up-regulated genes and clinical factors, a variety of clinicopathologic parameters from 31 samples were correlated with the presence or absence of up-regulation of genes. In addition, the transcript levels of up-regulated genes were also analyzed in correlation with each other, using regression analysis, in an effort to determine if any relationship existed between these up-regulated genes.

We used fresh tissue samples instead of cultured cell lines as experimental material because cell lines may not express important surface molecules (antigens or receptors) and secretory molecules (matrix or signal molecules) related to the communication or interaction with the surrounding microenvironment as that found *in vivo*, and which has recently been considered crucial to tumor progression and metastasis (4, 5).

Materials and Methods

Patients and Tissue Samples. From September 2001 to December 2002, fresh paired tissue samples of colorectal cancer and corresponding normal mucosa were harvested from consecutive patients who underwent radical colectomy due to colorectal cancer in Feng-Yuan Hospital, Feng-Yuan City, Taichung County, Taiwan. The cancerous tissue was cut directly from the tumor mass and the normal mucosa was carefully dissected from the inner wall of the adjacent uninvolved colon with a sharp scalpel. The paired tissue

Received 5/31/04; revised 8/23/04; accepted 9/22/04.

Grant support: Supported in part by research grant AS92IMB3 from Academia Sinica, Taipei, Taiwan.

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.

Requests for reprints: Hung Li, Institute of Molecular Biology, Academica Sinica, Taipei 11529, Taiwan. Phone: 886-2-2788-0460; Fax: 886-2-2782-6085. E-mail: hungli@ccvax.sinica.edu.tw

samples were put in separate vials and snap-frozen with liquid nitrogen. Samples were placed in a −85°C refrigerator for storage. Samples with the following conditions were abandoned: tissue weight <500 mg, tumor tissue with ambiguous boundary to adjacent mucosa (easy contamination of normal mucosa), and tissue with warm ischemia (the time period between excision of surgical specimens from patients and the snap-freezing of tissue samples in liquid nitrogen) for >15 minutes. Informed consent was acquired from all patients. Data of clinicopathologic parameters were obtained from patients' clinical records, operative notes, and pathologic reports.

Isolation of RNA from Tissue Samples. Blocks of tissue were ground into powder in −196°C liquid nitrogen. Tissue powder was homogenized using TRIzol reagent (1 mL/50-100 mg; Life Technologies Inc., Gaithersburg, MD). Total RNA was isolated according to the manufacturer's instructions. Homogenates were incubated for 15 minutes at 20°C to permit complete dissociation of nucleoprotein complexes. Chloroform (1/5 TRIzol solution) was added. The homogenate was vigorously agitated and then centrifuged at 12,000 × *g* for 10 minutes at 4°C to allow for phase separation: RNA remained exclusively in the upper aqueous phase which could then be easily separated. Precipitation of RNA occurred after mixing of the separated aqueous phase with isopropanol (1/2 TRIzol solution), incubation for 10 minutes at 20°C, and centrifugation at 12,000 × *g* for 15 minutes at 4°C. The visible RNA precipitate pellet was washed with 75% ethanol, centrifuged at 7,500 × *g* for 5 minutes at 4°C, air-dried and redissolved in diethyl pyrocarbonate-treated H_2O. RNA solution was stored at −70°C. Total RNA with $OD_{260}/OD_{280} > 1.6$ was used for microarray experiments.

Microarray Experiments. Four microarray experiments were done using four paired (colon cancer and adjacent normal mucosa) tissue specimens from different patients. mRNA was isolated from total RNA with a Dynal MPC-s kit (Dynal Biotech, Lake Success, NY), and reverse-transcribed with Superscript II RNase H-reverse transcriptase (Life Technologies) to generate Cy5- and Cy3-labeled cDNA probes for cancer and normal samples, respectively. The labeled probes were hybridized to a commercial cDNA microarray, comprising a total of 8,000 immobilized cDNA fragments (ABC Human UniversoChip 8k-1; Asia BioInnovations Corporation, Taiwan). Fluorescence intensities of Cy5 and Cy3 targets were measured and scanned separately using a GenePix 4000B Array Scanner (Axon Instruments, Union City, CA). Data analysis was done using associated software, GenePix Pro 3.0.5.56 (Axon Instruments). The signal-to-noise ratios for the 635 and 532 nm channel were estimated by (F635Median-B635Median)/B635SD and (F532Median-B532Median)/B532SD. The signal-to-noise ratios were used to check the quality of the microarray hybridization process. The results were normalized for the labeling and detection efficiencies of the two fluorescence dyes, then used to determine differential gene expression between cancer and normal samples. The genes were considered to be up-regulated if the Cy5/Cy3 signal ratio was >2.

Real-time Reverse Transcription-PCR. A total of 31 paired samples were used for real-time reverse transcription-PCR, including the same samples studied in the initial array experiments, plus an additional 27 new samples. Two-step quantitative reverse transcription-PCR was done: 2 μg of total RNA was reverse-transcribed to cDNA using the Superscript preamplification system (Life Technologies), and quantitative real-time PCR was done on a Roche LightCycler system (Roche Molecular Biochemicals, Mannheim, Germany). For each PCR, the reaction was carried out in a reaction mixture (20 μL) consisting of 12.6 μL of H_2O, 2.4 μL of $MgCl_2$ (stock solution of kit), 0.5 μL (10 pmol) of forward primer, 0.5 μL (10 pmol) of reverse primer, 2 μL of cDNA, and 2 μL of LightCycler-FastStart DNA Master SYBR Green I. The gene-specific primers (Table 3) were designed using MaxVector software (Accelrys, Inc., San Diego, CA), and through genomic information obtained from the National Center for Biotechnology Information web site. The PCR protocol consisted of denaturation at 95°C for 10 minutes, followed by 40 to 60 cycles of PCR amplification (different annealing temperatures for specific primers, extension times, and fluorescence detection temperatures for different PCR products are listed in Table 3), with a subsequent melting curve analysis (continuous fluorescence detection from 65°C to 95°C with a temperature slope of 0.1°C/second). Relative mRNA quantification of each sample was calculated with reference to the standard curve constructed automatically by plotting the log number of 1-, 10-, 100-, 1,000-, and 10,000-fold serially diluted standard samples in each reaction. Most (95%) of the correlation coefficients of the standard curve in this study were $r = 1.0$ and the mean square errors were <0.2. The rest were $r = 0.99$ and the errors <0.4. Corrections for sample to sample differences were done by normalization to the reference gene (endogenous control). The constitutively expressed gene glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was taken as the reference gene in this study. The normalized expression of a target gene = mRNA level of the gene / mRNA level of *GAPDH* in the same sample.

Statistical Data Analysis. There were four parts to data analysis in this study. The first part consisted of correlating the various clinicopathologic factors with each other. We conducted this part of the analysis for the purpose of comparing the clinical features of our patients with common colorectal patients. From this comparison, we could inspect the population (patients) bias of patients enrolled in this study.

The second part of data analysis was to analyze the real-time PCR data generated from post-array validation in order to find genes that could be verified with regard to confirming the array results: for each candidate gene, its normalized mRNA expression in cancer and corresponding normal tissue were ''gene in tumor/*GAPDH* in tumor'' and ''gene in normal/*GAPDH* in normal,'' respectively. Wilcoxon signed ranks tests were used to determine the statistical significance of expression difference for each test gene in 31 paired samples.

The third part of data analysis was to analyze the association between confirmed genes and clinicopathologic factors: for each patient, a gene was defined as up-regulated if the normalized tumor/normal expression ratio = (gene in tumor / GAPDH in tumor) / (gene in normal / GAPDH in normal) >2. The presence or absence of up-regulation of confirmed genes in 31 patients was analyzed in correlation with important clinicopathologic factors, including age (≤65 versus >65 years old), gender (male versus female), tumor location (right side versus left side), carcinoembryonic antigen (CEA) level (≤4.3 versus >4.3 ng/mL), tumor stage (Duke's A, B, C, D), tumor differentiation (well, moderate, or poor), the presence or absence of anemia, lymph node metastasis, distal metastasis and concomitant polyp, using a χ^2 or Fisher's exact test.

The fourth part of our statistical analysis involved determining if there existed any expression relationship between the confirmed genes: 31 paired samples used in real-time reverse transcription-PCR provided data of varying transcript (expression) levels. For example, the mRNA transcript levels of gene A in 31 patients was analyzed in correlation with the levels of another gene, gene B, using regression analysis. The data were analyzed using Excel 2003 (Microsoft, Corp.) and the statistical package program Statistical Package for the Social Sciences 11.0 (SPSS Inc., Chicago, IL). Statistical significance was defined as $P < 0.05$.

Results

Part 1: Clinical Features of Patients in this Study. Table 1 illustrates the clinical characteristics of patients in this study. The average age (67.4 years old) was close to 70.

Tumors occurred with a mild predominance in males (males/females = 1.38:1); 45% of tumors occurred in the rectum and 55% in the colon. Tumors located in the rectum were largely in males, whereas colonic tumors had a female predominance (P = 0.067). Patients with cancers of the right-sided colon exhibited a preponderance to anemia, more so than those with cancers of the left-sided colon (P = 0.008). All these clinical features were consistent with the current epidemiologic data concerning colorectal cancer. The consistency indicated that the 31 patients in this study were unbiased samples from colorectal cancer populations. In addition, our data also showed significant elevation of serum CEA levels in patients with left-sided colonic tumors (P = 0.036), as well as in patients at a more advanced tumor stage (P = 0.043).

Part 2: Microarray Experiments and Real-time Reverse Transcription-PCR Validation. Cy5/Cy3 ratios of 8,000 genes showed a Gaussian distribution (Fig. 1). The moderate skew to the right in terms of distribution in Fig. 1 roughly indicates that the number of up-regulated genes was greater than the number of down-regulated genes in colorectal cancers. Those with Cy5/Cy3 signal ratios >2 were arbitrarily defined as up-regulated genes in this study. As summarized in Fig. 2, more than 400 genes were over-expressed in each array chip but only 29 genes (listed in Table 2) were concurrently up-regulated in all four array experiments. Of these 29 candidate genes, 20 genes were tested for post-array validation using real-time reverse transcription-PCR in 31 paired samples: 40% (8/20) of these genes were confirmed to have a tumor/normal expression ratio > 2 (Tables 2 and 3, Fig. 3).

Part 3: Correlation Between Up-regulated Genes and Clinicopathologic Factors. Among eight confirmed genes, six were linked to clinicopathologic factors. The statistical correlation between *GRO1* and the patients' clinicopathologic variables is shown in Table 4. Table 5 summarizes statistically significant relationships between the presence or absence of up-regulation of six genes and the presence or absence of 10 clinicopathologic factors. Up-regulation of *THY1* and *PHLAD1* were found to be associated with the presence of patients' anemia (P = 0.036 and 0.009). Up-regulation of *HNRPA1* was significant in cancers growing in the right-sided colon more so than the left side (P = 0.027). Up-regulated *GPX2* was related to a higher degree of tumor differentiation (P = 0.019). The *c-MYC* oncogene was up-regulated to a greater degree in specimens from male rather than female colon cancer patients (P = 0.012). *GRO1* was significantly up-regulated in patients younger than 65 years old (P = 0.010); tumor stage, lymph node metastasis, and serum CEA levels were found to be frequently over-expressed when analyzed in correlation with *GRO1* (P = 0.075, 0.060, 0.058, respectively). The data in Table 4 shows that the less invasive the tumor (less advanced in stage, lymph node metastasis, and CEA levels), the more prominent was *GRO1* over-expression.

Table 1. Patient characteristics

	n (%)
Total number of patients	31 (100%)
Age (years)	
Range	33-87
Mean ± SD	67.4 ± 12.4
Median	69
Mode	69
Gender	
Male	18 (58%)
Female	13 (42%)
Tumor location	
Colon (male/female)	17 (55%; 7:10)
Rectum (male/female)	14 (45%; 11:3)
Anemia	
Yes (right/left colon)	21 (68%; 10:11)
No (right/left colon)	10 (32%; 0:10)

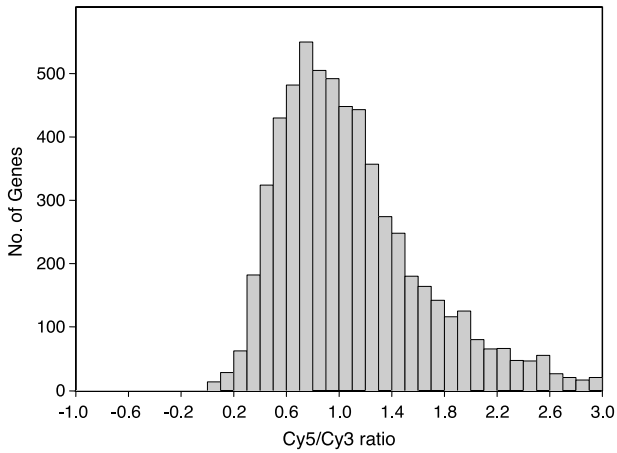

Figure 1. Cy5/Cy3 distribution profile of 8,000 genes in a microarray chip. Ratio > 2 defined as up-regulation.

Part 4: Expression Relationship Between Up-regulated Genes. Regression analysis of mRNA transcript levels between eight confirmed genes revealed a tight correlation between *c-MYC* and heterogeneous nuclear ribonucleoprotein A1 (*HNRPA1*) in tumor tissue (r = 0.948; Fig. 4). When transcript levels of these two genes in normal tissue were taken as the baseline, the adjusted mRNA levels of *c-MYC* and *HNRPA1* still showed a strong correlation (r = 0.871; Fig. 5). This finding suggested that the transcription of *HNRPA1* may be coupled to that of *c-MYC* in an unknown manner.

Discussion

By themselves, the long lists of data obtained from microarray experiments help little in the understanding of clinical characteristics. However, analysis of gene expression in correlation with clinical or phenotypic variations may indicate biologically meaningful changes at a transcriptional level. Prior to this study, other array-based studies had shown expression profiles or gene clusters associated with colorectal cancer (1-3); whereas any relationship between individual genes and clinicopathologic factors was never clarified. Singh et al. (6) had utilized microarrays to identify genes that might predict the

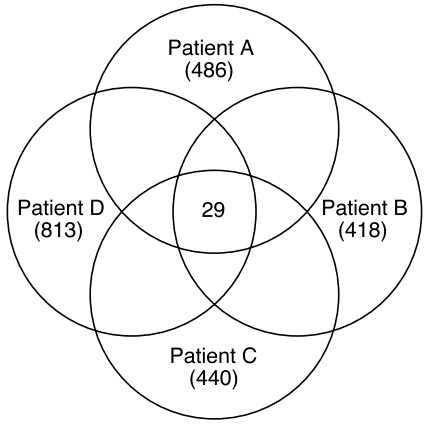

Figure 2. Number of up-regulated genes in four microarray experiments.

Table 2. Twenty-nine genes up-regulated in four microarray experiments

Gene name	Accession no.	Gene symbol	Function/category
GRO1 oncogene (melanoma growth-stimulating activity, α)	W46900	*GRO1**	oncogene/chemokine
B-factor, properdin	AA401441	*BF*†	m-protease
Thiosulfate sulfurtransferase (rhodanese)	AA461065	*MPST*†	m-transferase
IFN-inducible protein 1-8D	AA862371	*1-8D*†	surface molecule
V-*myc* avian myelocytomatosis viral oncogene homologue	AA464600	*MYC**	oncogene/transcription factor/cell cycle
Human γ-glutamyl hydrolase mRNA, complete cds	AA455800	*GGH**	m-hydrolase
Human ribosomal protein L10 mRNA, complete cds	AA434088	*RPL15*	structure protein/protein synthesis /RNA binding protein
Tropomyosin β chain (skeletal muscle)	AA477400	*TPM2*	structure protein
Pleckstrin homology-like domain, family A, member 1	AA258396	*PHLDA1**	binding protein/cytoskeleton/apoptosis
Human gene encoding prepro form of corticotropin releasing factor	R45054	*CRH*†	stress protein
Glutathione peroxidase 2, gastrointestinal	AA135289	*GPX2**	peroxidase
Heterogeneous nuclear ribonucleoprotein A1	AA127116	*HNRPA1**	RNA binding protein/transport
Ribosomal protein S5	AA456616	*RPS5*	structure protein/protein synthesis /RNA binding protein
Collagen, type IV, α 2	AA430540	*COL4A2*	structure protein
Human dystroglycan (DAG1) mRNA, complete cds	AA496691	*DAG1*†	receptor
Prostacyclin-stimulating factor (human, cultured diploid fibroblast cells, mRNA, 1,124 nucleotides)	T53298	*IGFBP7*	binding protein
H. sapiens HEK2 mRNA for protein tyrosine kinase receptor	AA455591	*EPHB3**	receptor/kinase
H. sapiens mRNA for testican	AA699317	*SPOCK*†	cell matrix protein/cell adhesion /cell proliferation
Nuclear factor (erythroid-derived 2)-like 3	W76339	*NFE2L3*	transcription factor
H. sapiens ERF-1 mRNA 3′ end	AA424743	*BRF1*	transcription factor
Nucleophosmin (nucleolar phosphoprotein B23, numatrin)	AA669758	*NPM1*†	RNA binding protein
Nucleoside diphosphate kinase A	AA644092	*NME1*†	transcription factor
Heat shock factor protein 1	AA449119	*UBE2M*†	apoptosis
Immunoglobulin γ 3 (Gm marker)	N92646	*CSF2RA*†	receptor
THY-1 membrane glycoprotein precursor	AA496283	*THY1**	antigen
H. sapiens mRNA for D1075-like gene	AA422058	*METTL1*†	m-transferase
Ribosomal protein S14	H73727	*RPS14*	structure protein/protein synthesis /RNA binding protein
Nonmetastatic cells 2, protein (NM23B) expressed in	AA496628	*NME2*†	kinase
Endoglin (Osler-Rendu-Weber syndrome 1)	AA446108	*ENG*	receptor

*Wilcoxon signed ranks test.
†Gene failed to pass post-array validation.

clinical behavior of a disease (prostate cancer), but there was no individual gene in their report whose expression correlated to the relevant clinical and pathologic parameters. In this study, we validated eight up-regulated genes in colorectal cancer tissue and found six of them to be linked to clinicopathologic variables.

In post-array validation of this study, we tested 20 candidate genes using real-time reverse transcription-PCR in 31 paired samples, including the same samples studied in the initial array experiments, plus an additional 27 new samples. The 20 candidate genes were up-regulated in all four chips; however, only 40% (8/20) of these candidate genes were post-array validated by real-time reverse transcription-PCR as being up-regulated in colorectal cancer. Previously, Rajeevan et al. (7) had used real-time reverse transcription-PCR to test 24 selected candidate genes from their array data and 71% (17/24) of those genes passed the post-array validation. Their high agreement between array data and real-time reverse

Table 3. Primers and experimental conditions of GAPDH and eight up-regulated genes confirmed with quantitative reverse transcription-PCR in 31 paired samples

Accession no.	Primer pair	Annealing temperatures (°C)	Elongation time (s)	Seg 4 detection (°C)*	*P*
GAPDH	(F) 5′-TCCTCTGACTTCAACAGCGACACC-3′ (R) 5′-GTCTCTCTCTTCCTCTTGTGCTCTTGC-3′	59	14	84	
AA464600	(F) 5′-TCTGGATCACCTTCTGCTGG-3′ (R) 5′-TGTTGCTGATCTGTCTCAGG-3′	60	12	85	<0.001
AA455800	(F) 5′-AGACTATTTTCCTGTGTGGGGC-3′ (R) 5′-CTCATAAGGTGCTTTCTCTGGATG-3′	61	17	82	<0.001
AA455591	(F) 5′-CAGAAGACCTGCTCCGTATTGG-3′ (R) 5′-TCACCCCCTCTCCTAATCCATC-3′	61	16	87	0.001
AA496283	(F) 5′-CGTTGCTACTAAGTGGTTGGGGAG-3′ (R) 5′-GGTGGTTCTTCCTGTTCTGTGACTG-3′	62	14	87	0.004
W46900	(F) 5′-CACATACATTCCCCTGCCTTCAC-3′ (R) 5′-CAACCCCAAGTTAGTTCAATCTGG-3′	63	10	81	<0.001
AA127116	(F) 5′-TTTGAGCAATGGGGAACGC-3′ (R) 5′-GGGCACCTGGTCTTTGAGAATC-3′	61	12	83	<0.001
AA258396	(F) 5′-GCTGGCACAACAATGAAAGTGTC-3′ (R) 5′-CTTGGAGAATAAAAAAGGGCGG-3′	61	6	74	<0.001
AA135289	(F) 5′-GCAGGAGAGACAGAAGTAGCAAACC-3′ (R) 5′-ACAAATGGCTTGGCTGGCTC-3′	64	8	82	0.001

NOTE: Wilcoxon signed ranks test.
*Seg 4 detection temperature = fluorescence detection temperature at the end of each PCR cycle.

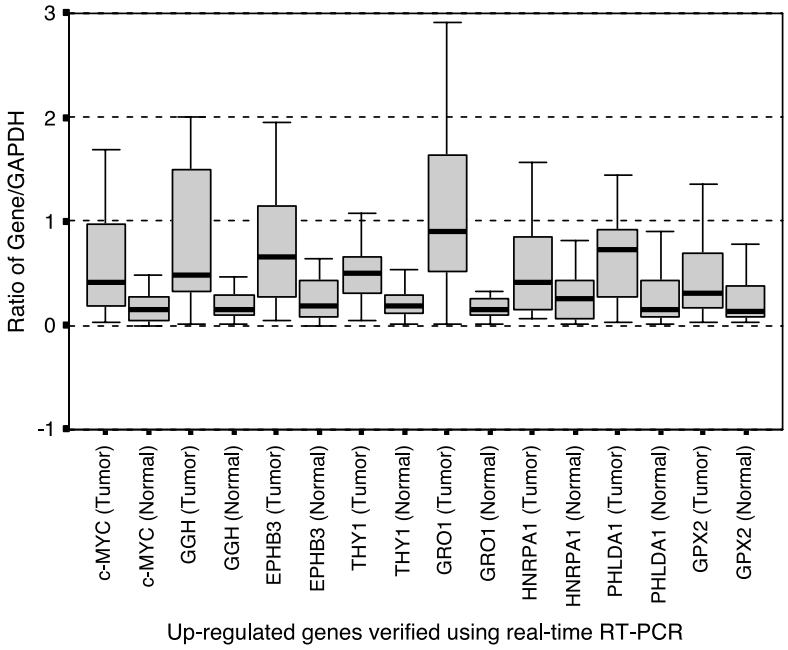

Figure 3. Eight genes confirmed to have normalized tumor/normal ratio> 2 in quantitative reverse transcription-PCR experiments.

transcription-PCR results was mainly due to the fact that they used only one pair of cultured cell line (two keratinocyte subclones) in both array experiments and post-array validation. In the current study, we used tissue samples from different patients; heterogeneity of tissue cells and sample-to-sample variation led to the low agreement between array data and real-time reverse transcription-PCR results. Therefore, validation of microarray results using another independent laboratory approach with additional tissue samples is mandatory if few microarrays are carried out and tissues are used as sources of experimental samples (8).

In combination with previous reports, some connections, either between up-regulated genes and carcinogenesis, or between gene expression and clinical factors, can be established from this study. Inferences based on our findings and evidence from others studies are discussed below.

Table 4. Correlation between clinicopathologic variables and expression status of the *GRO1* gene (*n* = 31)

Factors	Patients	No. of patients		*P*
		GRO1 (+)	*GRO1* (−)	
Age (years)*				
≤65	12	12	0	0.010 †
>65	19	11	8	
Gender*				
Male	18	15	3	0.170
Female	13	8	5	
Tumor location*				
Right	10	8	2	0.483
Left	21	15	6	
Anemia*				
Yes	21	16	5	0.517
No	10	7	3	
CEA level (ng/mL)*				
≤4.3	13	12	1	0.058 ‡
>4.3	18	11	7	
Duke's stage§				
A	0	0	0	0.075 ‡
B	18	16	2	
C	10	5	5	
D	3	2	1	
Lymph node metastasis*				
Yes	14	8	6	0.060 ‡
No	17	15	2	
Distal metastasis*				
Yes	4	3	1	0.732
No	27	20	7	
Other polyp*				
Yes	16	13	3	0.303
No	15	10	5	
Differentiation§				
Well to moderate	11	9	2	0.724
Moderate	16	11	5	
Moderate to poor	3	2	1	

NOTE: *GRO1* (+) = (GRO1 / GAPDH of tumor) / (GRO1 / GAPDH of normal) >2; *GRO1* (−) = (GRO1 / GAPDH of tumor) / (GRO1 / GAPDH of normal) <2.
*Fisher's exact test.
† Statistically significant ($P < 0.05$).
‡ Trend forward significance.
§ χ^2 test.

***MYC* versus Sex.** In this study, *c-MYC* was over-expressed in 94% (17/18) of samples from male colon cancer patients, but was over-expressed in only 54% (7/13) of females. This suggests a gender-related influence on *c-MYC* expression. The role of androgen in increasing *c-MYC* expression has been investigated and confirmed in prostatic cancer by many authors (9-12). Therefore, it is likely that androgen may have a similar effect in colorectal cancer. In fact, epidemiologic data concerning relative risk of colorectal cancer between males and females (1.25 in Taiwan), supports this postulation.

***MYC* versus *HNRPA1*.** Our data also disclosed a strong linear correlation between mRNA transcript levels of *c-MYC* and that of *HNRPA1* (correlation coefficient, 0.948). HNRPA1, the heterogeneous nuclear ribonucleoprotein (hnRNP) core protein A1, has a modular structure consisting of two conserved RNA binding domains (13) and functions as a carrier for RNA during export of RNA to the cytoplasm (14). The strong correlation between *c-MYC* and *HNRPA1* expression may indicate a tight association between the transcriptional factor and the RNA binding proteins trafficking in and out of the nuclear membrane. Many authors have reported using c-*myc* antisense oligonucleotides to inhibit the cellular proliferation of various cancers, including colon cancer (15). From our data, a combined antisense oligo of c-*myc* and hnRNP A1 could be a strategy for treating colon cancer more effectively than c-*myc* antisense oligos alone.

***GRO1* versus Age and Tumor Stage + Lymph Node Metastasis + CEA Level.** Up-regulation of the *GRO1* gene in colorectal cancer (1, 16) and that the GRO1 protein functions as a potent mediator of leukocyte recruitment and proliferation in inflammatory diseases (17-19) hint that the tumor growth of

Table 5. Summary of statistical correlation between up-regulated genes and clinicopathologic variables in 31 colorectal cancer patients

Gene symbol (function)	Variable	P (n = 31)
MYC (oncogene/transcription factor)	Gender	0.012
HNRPA1 (RNA binding protein/transport)	Tumor location	0.027
GRO1 (oncogene/chemokine)	Age	0.010
	Tumor stage	0.075*
	Lymph node metastasis	0.060*
	CEA elevation	0.058*
GPX2 (peroxidase)	Differentiation	0.019
THY1 (antigen)	Anemia	0.036
PHLDA1 (binding protein/cytoskeleton/apoptosis)	Anemia	0.009

NOTE: There was no statistical correlation between clinicopathologic factors and *GGH* (γ-glutamyl hydrolase) or *EPHB3* [erythropoietin-producing hepatoma amplified sequence B3 (*HEK2*, human embryo kinase 2)] in this study. *HNRPA1*, heterogeneous nuclear ribonucleoprotein core protein A1; *GRO1*, growth-related oncogene-α; *GPX2*, glutathione peroxidase 2; *THY1*, THY-1 T cell antigen (theta antigen); *PHLDA1*, pleckstrin homology-like domain, family A, member 1.
*Trend toward significance ($0.05 < P < 0.1$).

colon cancer might trigger an immune response together with *GRO1* over-expression. In this study, two findings indirectly support this speculation. First, our data showed that *GRO1* was frequently up-regulated in less invasive cancers (less advanced in stage, lymph node metastasis, and CEA levels). This implies that *GRO1* may have a protective effect (the immune response against tumors) in preventing the progression of colon cancer. Second, we found the *GRO1* gene was significantly up-regulated in patients younger than 65 years old; the immune response of young people is generally stronger than the elderly, hence, up-regulation of *GRO1* largely occurs in colon cancer patients younger than 65.

***GPX2* versus Differentiation.** *GPX2* expresses mainly in the epithelium of the gastrointestinal tract and its protein product, GPX-GI, functions as an intracellular selenium-dependent glutathione peroxidase that can reduce H_2O_2 and alkyl hydroperoxides (20). Previously, Chu et al. (21) reported that *GPX2* mRNA levels in the colon of mice relatively resistant to dimethylhydrazine-induced colon cancer were higher than those in mice sensitive to chemical-induced colon cancer. In this study, we found that an up-regulation of *GPX2* was related to a higher degree of tumor differentiation. Both Chu et al.'s study and ours suggest that *GPX2* gene expression has an adverse effect on tumor proliferation, maybe due to its differentiation-promoting effect. Chu et al. also identified three retinoic acid response elements in this gene and showed that retinoic acid, a vitamin with known prodifferentiation effects, could induce *GPX2* gene expression in a human breast cancer cell line (22). Therefore, if retinoic acid can induce *GPX2* over-expression in colon cancer cells just as it does in cultured breast cancer cells, vitamin A could possibly be used to clinically promote cell differentiation of colon cancer.

***THY1* and *PHLDA1* versus Anemia.** *THY1* encodes THY-1, a surface glycoprotein characteristic of T cells, hematopoietic stem cells, liver stem cells (23), and blast cells of acute myeloid leukemia (24). THY-1 is structurally the simplest of the T cell antigen receptors in the immunoglobulin supergene family (25), but its role in immune response is unclear. *PHLDA1*, known in mice as the T cell death-associated gene, is one of the gut-expressed proteins with high T cell epitope homology (26). In this study, we found that both *THY1* and *PHLDA1* were up-regulated in colon cancer tissue and both of them were statistically relevant to the patient's anemia. This finding together with other reports (26, 27) hint that *THY1* and *PHLDA1* products are two surface molecules responsible for crosstalk between colon cancer and the immune system.

Further work is required to determine whether the up-regulation of these two genes in colon cancer is related to tumor-infiltrating T cells (27) being recruited, leading to increased tumor cell necrosis and subsequent tumor mass bleeding and patients' anemia.

The clinical correlates of up-regulated genes in this study are essentially statistical inference. More scientific evidence confirming the correlation and associated postulations are still

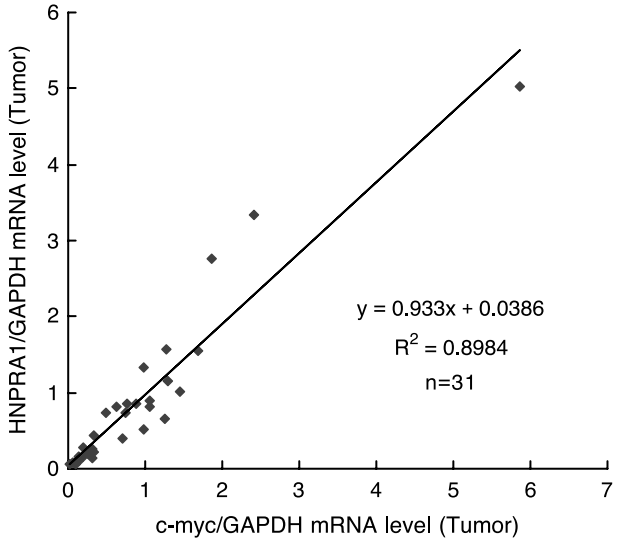

Figure 4. Correlation between the mRNA levels of *HNRPA1/GAPDH* and c-*MYC/GAPDH* in tumor tissue (r = 0.948; n = 31).

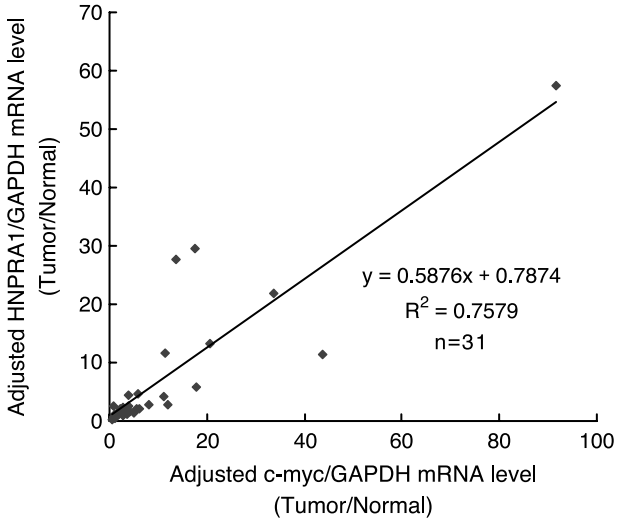

Figure 5. Correlation between the mRNA levels of adjusted *HNRPA1/GAPDH* and adjusted c-*MYC/GAPDH* (tumor/normal; r = 0.871; n = 31).

necessary. However, the findings in this study provide clues to molecular events related to the carcinogenesis or clinical features of human colorectal cancer and suggest possible therapeutic targets.

Acknowledgments

We thank Dr. K. Deen for his critical reading of this manuscript.

References

1. Zou TT, Selaru FM, Xu Y, et al. Application of cDNA microarrays to generate a molecular taxonomy capable of distinguishing between colon cancer and normal colon. Oncogene 2002;21:4855–62.
2. Lin YM, Furukawa Y, Tsunoda T, Yue CT, Yang KC, Nakamura Y. Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. Oncogene 2002;21: 4120–8.
3. Notterman DA, Alon U, Sierk AJ, Levine AJ. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. Cancer Res 2001;61:3124–30.
4. Matrisian LM, Cunha GR, Mohla S. Epithelial-stromal interactions and tumor progression: meeting summary and future directions. Cancer Res 2001;61:3844–6.
5. Liotta LA, Kohn EC. The microenvironment of the tumour-host interface. Nature 2001;411:375–9.
6. Singh D, Febbo PG, Ross K, et al. Gene expression correlates of clinical prostate cancer behavior. Cancer Cell 2002;1:203–9.
7. Rajeevan MS, Vernon SD, Taysavang N, Unger ER. Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. J Mol Diagn 2001;3:26–31.
8. Chuaqui RF, Bonner RF, Best CJ, et al. Post-analysis follow-up and validation of microarray experiments. Nat Genet 2002;32 Suppl:509–14.
9. Silva IS, Morsch DM, Urnauer L, Spritzer PM. Androgen-induced cell growth and c-*myc* expression in human non-transformed epithelial prostatic cells in primary culture. Endocr Res 2001;27:153–69.
10. Asadi FK, Sharifi R. Effects of sex steroids on cell growth and C-*myc* oncogene expression in LN-CaP and DU-145 prostatic carcinoma cell lines. Int Urol Nephrol 1995;27:67–80.
11. Quarmby VE, Beckman WC Jr, Wilson EM, French FS. Androgen regulation of c-*myc* messenger ribonucleic acid levels in rat ventral prostate. Mol Endocrinol 1987;1:865–74.
12. Grad JM, Dai JL, Wu S, Burnstein KL. Multiple androgen response elements and a Myc consensus site in the androgen receptor (AR) coding region are involved in androgen-mediated up-regulation of AR messenger RNA. Mol Endocrinol 1999;13:1896–911.
13. Biamonti G, Ruggiu M, Saccone S, et al. Two homologous genes, originated by duplication, encode the human hnRNP proteins A2 and A1. Nucleic Acids Res 1994;22:1996–2002.
14. Michael WM, Choi M, Dreyfuss G. A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. Cell 1995;83:415–22.
15. Li BD, Budnick RM, Russo CA, et al. Quantifying c-*myc* expression in c-*myc* antisense phosphorothioate oligodeoxynucleotide-treated leukemic and colon cancer cell lines. J Surg Res 1995;59:485–92.
16. Cuenca RE, Azizkhan RG, Haskill S. Characterization of GRO α, β and γ expression in human colonic tumours: potential significance of cytokine involvement. Surg Oncol 1992;1:323–9.
17. Kusugami K, Ando T, Ohsuga M, et al. Mucosal chemokine activity in Helicobacter pylori infection. J Clin Gastroenterol 1997;25 Suppl 1: S203–10.
18. Klein AM, Anderson K, Lafreniere D, et al. Growth-related oncogene-α expression in human nasal polyps. Otolaryngol Head Neck Surg 2000; 123:85–90.
19. Lawrance IC, Fiocchi C, Chakravarti S. Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes. Hum Mol Genet 2001;10:445–56.
20. Chu FF, Doroshow JH, Esworthy RS. Expression, characterization, and tissue distribution of a new cellular selenium-dependent glutathione peroxidase, GSHPx-GI. J Biol Chem 1993;268:2571–6.
21. Chu FF, Esworthy RS, Ho YS, et al. Expression and chromosomal mapping of mouse Gpx2 gene encoding the gastrointestinal form of glutathione peroxidase, GPX-GI. Biomed Environ Sci 1997;10:156–62.
22. Chu FF, Esworthy RS, Lee L, Wilczynski S. Retinoic acid induces Gpx2 gene expression in MCF-7 human breast cancer cells. J Nutr 1999;129:1846–54.
23. Petersen BE, Goff JP, Greenberger JS, Michalopoulos GK. Hepatic oval cells express the hematopoietic stem cell marker Thy-1 in the rat. Hepatology 1998;27:433–45.
24. Kozii R, Wilson J, Persichetti J, et al. Thy-1 expression on blast cells from adult patients with acute myeloid leukemia. Leuk Res 1997;21:381–5.
25. Hood L, Kronenberg M, Hunkapiller T. T cell antigen receptors and the immunoglobulin supergene family. Cell 1985;40:225–9.
26. Kumar R, Eastwood AL, Brown ML, Laurie GW. Human genome search in celiac disease: mutated gliadin T-cell-like epitope in two human proteins promotes T-cell activation. J Mol Biol 2002;319:593–602.
27. Golby SJ, Chinyama C, Spencer J. Proliferation of T-cell subsets that contact tumour cells in colorectal cancer. Clin Exp Immunol 2002;127:85–91.