

Secreted and Cell Surface Genes Expressed in Benign and Malignant Colorectal Tumors¹

Phillip Buckhaults, Carlo Rago, Brad St. Croix, Katharine E. Romans, Saurabh Saha, Lin Zhang, Bert Vogelstein, and Kenneth W. Kinzler²

Howard Hughes Medical Institute [P. B., C. R., B. V.] and Johns Hopkins Oncology Center [B. S. C., K. E. R., S. S., L. Z., B. V., K. W. K.], Johns Hopkins Medical Institutions, Baltimore, Maryland 21231

Abstract

Serial analysis of gene expression was used to identify transcripts encoding secreted or cell surface proteins that were expressed in benign and malignant tumors of the colorectum. A total of 290,394 tags were analyzed from normal, adenomatous, and cancerous colonic epithelium. Of the 21,343 different transcripts observed, 957 were found to be differentially expressed between normal tissue and adenoma or between normal tissue and cancer. Forty-nine transcripts were elevated ≥ 20 -fold in adenomas, 40 transcripts were elevated ≥ 20 -fold in cancers, and 9 transcripts were elevated ≥ 20 -fold in both. Products of six of these nine transcripts (*TGFBI*, *LYS*, *RDP*, *MIC-1*, *REGA*, and *DEHL*) were predicted to be secreted or to reside on the cell surface, and these were analyzed in more detail. The abnormal expression levels predicted by serial analysis of gene expression were confirmed by quantitative PCR analyses of each of these six genes. Moreover, the cell types responsible for the elevated expression were identified by *in situ* hybridization and by PCR analyses of epithelial cells immunoaffinity purified from primary tumors. This study extends knowledge of the differences in gene expression that underlie various stages of neoplasia and suggests specific diagnostic approaches that may be useful for the early detection of colorectal neoplasia.

Introduction

Colorectal cancer is the second leading cause of cancer death in the United States, with $\sim 135,000$ patients diagnosed each year and $\sim 55,000$ ultimately succumbing to the disease (1). Most colorectal cancers develop slowly, beginning as small benign colorectal adenomas that progress over several decades to larger and more dysplastic lesions that eventually become malignant. This gradual progression provides multiple opportunities for prevention and intervention. Indeed, benign adenomas can be detected and removed by simple colonoscopy and polypectomy, precluding the need for radical surgical and adjuvant treatments. It is therefore believed that early detection and removal of these benign neoplasms provide the best hope for minimizing morbidity and mortality from colorectal cancer. Various screening methods for detecting early colorectal tumors are available, such as fecal occult blood testing, sigmoidoscopy, and colonoscopy (reviewed in Ref. 2). However, none of these methods are optimal,

and new approaches are needed. In an effort to identify potential molecular markers of early colorectal tumors, we have here analyzed gene expression in benign and malignant colorectal tumors in an unbiased and comprehensive fashion. Our results, which are reported below, demonstrate that several genes are expressed at markedly higher levels in both benign and malignant tumors compared with normal colonic epithelium. Furthermore, several of the most differentially expressed genes are extracellular, thereby providing opportunities for clinical application.

Materials and Methods

SAGE.³ For the initial SAGE of benign tumors, fresh adenomas were obtained from surgical specimens derived from familial adenomatous polyposis patients. Adenomas from familial adenomatous polyposis patients were used because of the ready availability of small lesions and the certainty of inactivation of the adenomatous polyposis coli pathway that initiates the formation of the majority of sporadic tumors. After histopathological verification of the neoplastic nature of the lesion ($>70\%$ neoplastic cells), total RNA was isolated by solubilizing the tissue in RNagents Lysis Buffer (Promega, Madison, WI) followed by ultracentrifugation over a cesium chloride gradient. mRNA selection was performed from the purified total RNA using oligo(dT) cellulose (Life Technologies, Inc., Gaithersburg, MD). Two adenoma SAGE libraries were prepared as described previously (3, 4) and sequenced to a total depth of over 90,000 transcript tags. For SAGE of normal and malignant tissues, four previously described normal (NC-1 and NC-2) and primary cancer (Tu-98 and Tu-102) SAGE libraries were used (4). In collaboration with the Cancer Genome Anatomy Project (5), the analyses of these libraries were extended from a total of 123,046 transcripts in the previously published work to 195,160 transcripts in the current work. Tags were extracted from the raw sequence data, and after excluding repeated ditags, linker sequences, and tags from the polymorphic major histocompatibility loci, the resulting tag libraries were compared, and statistical analysis was performed using SAGE software, version 4.0. Data from the libraries are publicly available,⁴ as are detailed SAGE protocols.⁵

Quantitative PCR. Tumors were collected, snap frozen, and stored at -80°C . They were verified to be composed predominantly of neoplastic cells by histopathological analysis. mRNA was isolated from tumors and patient-matched normal colonic mucosa using QuickPrep reagents (Amersham Pharmacia Biotech UK, Buckinghamshire, United Kingdom), and single-stranded cDNA was synthesized using Superscript II (Life Technologies, Inc.). Quantitative PCR was performed using an iCycler (Bio-Rad, Hercules, CA), and threshold cycle numbers were determined using iCycler software, version 2.1. Reactions were performed in triplicate, and threshold cycle numbers were averaged. All genes examined were normalized to a control gene (β -amyloid precursor protein, shown by SAGE to be expressed at equivalent levels in all colorectal samples), and fold induction was calculated according to the formula $2^{(Rt-Et)}/2^{(Rn-En)}$ where Rt is the threshold cycle number for the reference gene observed in the tumor, Et is the threshold cycle number for the experimental gene observed in the tumor, Rn is the threshold cycle number for the reference

Received 6/22/01; accepted 8/2/01.

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.

¹ Supported by NIH Grants CA43460, CA57345, and CA62924. K. W. K. received research funding from Genzyme Molecular Oncology (Genzyme), and K. W. K. and B. V. are consultants to Genzyme. Under a licensing agreement between the Johns Hopkins University and Genzyme, the SAGE technology was licensed to Genzyme, and K. W. K. and B. V. are entitled to a share of royalty received by the Johns Hopkins University from sales of the licensed technology. The SAGE technology is freely available to academia for research purposes. The Johns Hopkins University, K. W. K., and B. V. own Genzyme stock, which is subject to certain restrictions under university policy. The terms of this arrangement are being managed by the Johns Hopkins University in accordance with its conflict of interest policies.

² To whom correspondence should be addressed, at The Johns Hopkins Oncology Center, Cancer Research Building Room 588, 1650 Orleans Street, Baltimore, MD 21231. Phone: (410) 955-2928; Fax: (410) 955-0548; E-mail: kinzlike@jhmi.edu.

³ The abbreviations used are: SAGE, serial analysis of gene expression; TGF, transforming growth factor.

⁴ www.ncbi.nlm.nih.gov/sage.

⁵ www.sagenet.org/sage_protocol.htm.

Table 1 Summary of SAGE data

SAGE library	Total no. of tags observed	No. of different transcripts observed ^a
Normal colorectal epithelium		
NC-1	49,610	9,359
NC-2	48,479	9,610
Adenomas		
Ad-A	52,573	11,167
Ad-B	42,661	9,483
Cancers		
Tu-98	41,371	9,780
Tu-102	55,700	11,039
Total	290,394	21,343

^a To minimize the effect of potential sequencing errors, only tags observed more than once in a given SAGE library were counted to give a conservative estimate of the minimum number of different transcripts analyzed.

gene observed in the normal tissue, and En is the threshold cycle number for the experimental gene observed in the normal tissue. The primers used for quantitative PCR were obtained from GeneLink (Hawthorne, NY), and their sequences are available on request.

Epithelial Cell Immunoaffinity Purification. Tumor epithelial cells were purified using a modification of the procedure developed previously for the isolation of tumor endothelial cells (6). In brief, fresh surgical specimens of tumor and matched normal tissue were obtained and digested with collagenase, and the resulting material was filtered through a nylon mesh to obtain single cell suspensions. The cells were then bound to a mixture of anti-CD14 and anti-CD45 immunomagnetic beads (DynaL, Oslo, Norway) to deplete the population of hematopoietic cells (negative selection). The remaining cell suspension was then incubated with anti-Ber-EP4 immunomagnetic beads to isolate epithelial cells (positive selection). Purified cells were lysed directly on the beads, and mRNA was purified using QuickPrep reagents (Amersham Pharmacia Biotech UK).

In Situ Hybridization. Nonradioactive *in situ* hybridization was performed as described previously (6). For each gene analyzed, a mixture of antisense probes made through *in vitro* transcription was used to increase sensitivity. The primers used to generate templates for the synthesis of the *in situ* riboprobes were obtained from GeneLink, and their sequences are available on request.

Results and Discussion

We used SAGE to analyze global gene expression in normal, benign, and malignant colorectal tissue. SAGE is a gene expression profiling method that associates individual mRNA transcripts with 15-base tags derived from specific positions near their 3' termini (3). The abundance of each tag provides a quantitative measure of the transcript level present within the mRNA population studied. SAGE is not dependent on preexisting databases of expressed genes and therefore provides an unbiased view of gene expression profiles. For the current study, SAGE libraries derived from two samples of normal colonic epithelium, two colorectal adenomas, and two colorectal cancers were analyzed. These libraries contained a combined total of 290,394 transcript tags representing 21,343 different transcripts (Table 1).

Two comparisons were performed, one between the adenoma and normal samples, and one between the cancer and normal samples.

These comparisons revealed 957 transcript tags that were differentially expressed >2-fold between normal and tumor tissue (Table 2). A comparison of the fold change in adenomas *versus* cancers revealed that many transcripts were similarly elevated or repressed in both adenomas and cancers, although the magnitude often varied (Fig. 1A). Indeed, the majority (79%) of tags were in quadrants of the plot indicative of concordant elevation or repression.

From both practical and biological perspectives, those changes showing the greatest magnitude were deemed the most interesting. In this regard, 49 tags were identified to be elevated by ≥ 20 -fold in the adenomas, and 40 were elevated by ≥ 20 -fold in the cancers (Table 2). Conversely, there were 72 transcripts that were decreased by ≥ 20 -fold in adenomas, and 52 transcripts were decreased by ≥ 20 -fold in the cancers (Table 2).

There were 9 transcripts that were elevated by ≥ 20 -fold in both adenomas and cancers (Fig. 1B and Table 3) and 23 that were repressed by ≥ 20 -fold (Table 4). We were especially interested in genes whose products were predicted to be secreted or displayed on the cell surface because these would be particularly suitable for the development of serologic or imaging tests for presymptomatic neoplasia, respectively. We were able to identify six such genes (*TGFBI*, *LYS*, *RDP*, *MIC-1*, *REGA*, and *DEHL*) from among those whose transcript tags were elevated in both adenoma and carcinoma SAGE libraries. These genes were analyzed further, as described below.

To verify the increased expression of these six genes, we used quantitative reverse transcription-PCR techniques to analyze the expression in seven colorectal neoplasms (three sporadic adenomas and four sporadic cancers) and matched normal colonic mucosa. For these assays, specific primers were developed that resulted in amplification from cDNA but not genomic DNA. Controls were provided by similar quantitative PCR assays of a gene whose expression was found to be very similar in the SAGE libraries of normal and neoplastic colon (β -amyloid precursor protein). The quantitative PCR experiments verified that five of the six selected genes (*TGFBI*, *LYS*, *RDP*, *MIC-1*, and *REGA*) were expressed at significantly higher levels in every neoplastic sample analyzed compared with patient-matched normal mucosa (Fig. 2). Several tumors exhibited ≥ 20 -fold higher levels of the studied transcripts compared with their patient-matched normal colonic mucosa, as predicted by SAGE. Another control was provided by the quantitative PCR analysis of four genes whose expression was observed to be reduced in the SAGE libraries prepared from adenomas and cancers compared with those from normal colonic mucosa. As shown in Fig. 3, the quantitative PCR confirmed the lower levels of expression of each of these genes, emphasizing that the dramatic elevations in expression observed in Fig. 2 represented gene-specific phenomena.

The quantitative PCR data obtained from mRNA isolated from whole tumors provided independent evidence that SAGE provided an accurate indication of gene expression changes in colorectal neoplasia. However, neither analysis identified the cell types responsible for

Table 2 Differentially expressed transcripts in benign and malignant tumor colorectal tissue

Fold change in expression	Elevated in adenomas ^a	Elevated in cancers ^a	Elevated in both adenomas and cancer ^a	Repressed in adenomas ^b	Repressed in cancers ^b	Repressed in both adenomas and cancers ^b	Total no. of differentially expressed transcripts
2	346	170	50	313	380	192	957
4	263	119	23	225	270	117	735
10	160	79	10	134	157	58	462
20	49	40	9	72	52	23	181

^a Elevated transcripts showed a significantly different ($P < 0.05$) tag count between normal and tumor tissue, were expressed in both tumor tissues analyzed, and had an expression level that was higher in the tumors than in the normal tissues by the fold indicated in column one. For the purposes of calculation, 0.5 was substituted for the denominator when no tags were detected in the normal samples.

^b Repressed transcripts showed a significantly different ($P < 0.05$) tag count between normal and tumor tissue, were expressed in both normal tissues analyzed, and had an expression level that was lower in the tumors than in the normal tissues by the fold indicated in column one.

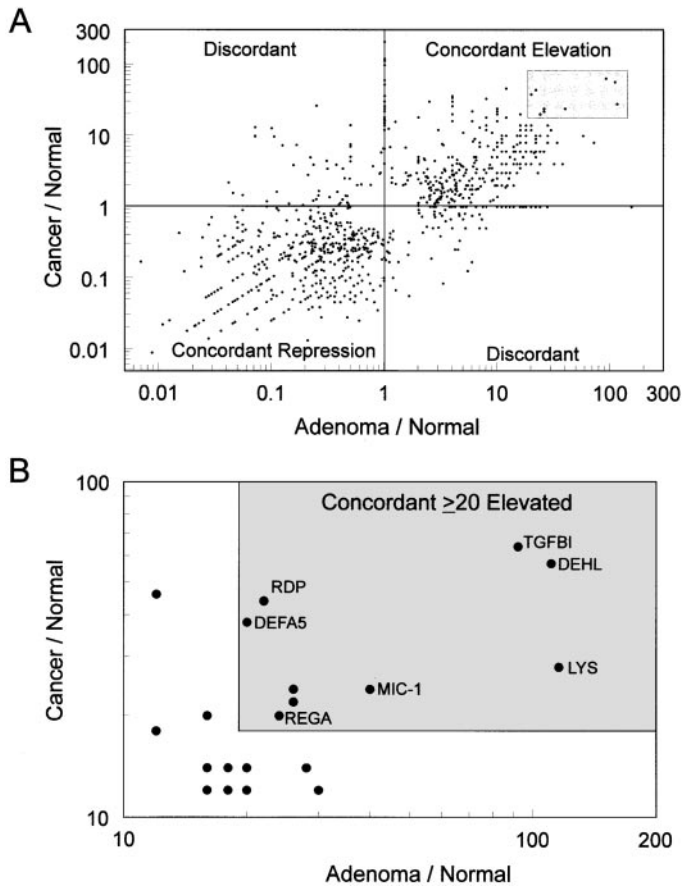


Fig. 1. A, distribution of the fold changes of differentially expressed transcript tags. Transcripts in which the significance criterion was met ($P < 0.05$, a total of 957 tags) in the comparisons between normal tissue and adenoma or normal tissue and cancer are plotted in the figure. The ratios of adenoma to normal and cancer to normal were plotted on a log scale. The shaded box shown in A and enlarged in B encloses the transcript tags detailed in Table 3. The two unlabeled dots in B correspond to tags whose differential expression could not be confirmed by quantitative PCR, suggesting that the tags were derived from different transcripts than the ones indicated in Table 3.

the increased expression. Nonneoplastic stromal cells within tumors may be considerably different than those in normal colonic mucosa (6), and the epithelial derivation of gene expression differences cannot reliably be concluded without direct supporting evidence. We therefore sought to determine whether the epithelial cells of cancers express elevated levels of the six genes depicted in Fig. 2. First, we affinity-purified cancerous and patient-matched normal epithelial cells from fresh surgical specimens using immunomagnetic beads directed to the panepithelial marker Ber-EP4, prepared cDNA, and performed quantitative PCR analysis to determine the expression

levels of the elevated genes as described above. Elevated expression was observed in the purified tumor epithelial cells for each of the six genes examined (Fig. 4), demonstrating that at least some of the increased expression was derived from epithelial cells. However, relative expression of *LYS* was not as prominent or reproducible in the purified epithelial cells as it was in the mRNA from the unfractionated tumors, suggesting that other cell types might have contributed transcripts from this gene.

Second, we performed *in situ* hybridization to RNA in frozen sections of tumors for five of the genes showing the most consistent elevation. *DEHL* was found to be elevated in only five of the nine tumors examined and was not investigated further. To increase the sensitivity of detection, we generated several RNA probes for each tested gene using *in vitro* transcription techniques. The results obtained are discussed below in conjunction with brief overviews of each of the five genes of interest.

Regenerating islet-derived pancreatic stone protein, encoded by the *REGA* gene, is a secreted polypeptide first found in pancreatic precipitates and stones from patients suffering from chronic pancreatitis (7). The cDNA encoding this protein was isolated from a random screen of genes highly expressed in a regenerating islet-derived cDNA library (8) and subsequently shown to be elevated in colorectal cancers (9). More recently, *REGA* was isolated in a hybridization-based screen for genes elevated in colorectal cancers and shown to be elevated in many colorectal adenocarcinomas (10). Consistent with these published observations, we observed a strong elevation in expression of *REGA* in unpurified tumors and a similar elevation in one purified tumor. *In situ* hybridization experiments demonstrated *REGA* to be strongly expressed in the epithelial cells of the tumors, with no expression evident in the stroma (Fig. 5A).

TGF- β -induced gene (*TGFBI*) encodes a small polypeptide of unknown function initially isolated through a differential display screen for genes induced in response to treatment with TGF- β (11). The protein is expressed in the keratinocytes of the cornea, (12) and, interestingly, germ-line mutations of this gene cause familial corneal dystrophies (13). *TGFBI* was previously shown to be among the most significantly elevated genes in colorectal cancers (4), and our new data show that it is expressed at high levels in adenomas as well. Quantitative PCR results demonstrated strong elevation in both unpurified tumors and purified tumor epithelial cells. Accordingly, *in situ* hybridization experiments revealed *TGFBI* to be expressed in many cell types, in both the stromal and epithelial compartments (Fig. 5B).

Lysozyme (*LYS*, 1,4- β -N-acetylmuramidase; EC 3.2.1.17) is an enzyme with bacteriolytic activity (14) capable of cleaving β 1,4 glycosidic bonds found in the cell walls of Gram-positive bacteria. The enzyme is expressed in the secretory granules of monocytes,

Table 3 Transcripts most elevated in adenomas and cancers^a

Tag sequence	Normal		Adenomas		Cancers		Transcript name
	NC-1	NC-2	AD-A1	AD-B2	Tu-98	Tu-102	
ATGTAAGAAA	0	0	26	32	2	12	Lysozyme (<i>LYS</i>)
TAATTTTGC	0	1	99	12	20	37	Differentially expressed in hematopoietic lineages (<i>DEHL</i>)
GTGTGTTGT	0	0	17	29	17	15	Transforming growth factor β -induced (<i>TGFBI</i>)
GTGTCATTC	0	0	13	7	2	10	Macrophage-inhibitory cytokine, 1 (<i>MIC-1</i>)
TTCCAGCTC	0	0	7	6	2	9	Adaptor-related protein complex 2, α 2 subunit ^b
ACCATTGGAT	0	0	3	10	3	9	Interferon-induced transmembrane protein 1 (9-27) ^b
TTTCCACTAA	0	0	8	4	4	6	Regenerating islet-derived 1 α (<i>REGA</i>)
CAAGGACCAG	0	0	5	6	10	12	Renal dipeptidase (<i>RDP</i>)
AGGACCATCG	0	0	8	2	1	18	Defensin, α 5, Paneth cell-specific ^c

^a These tags displayed at least 20-fold elevation in both neoplastic states. The numbers given are the raw tag counts for each tag observed in each library. Transcript name provides a description of matching UniGene cluster (Build 3/13/01). Rows shown in bold are genes confirmed by quantitative PCR to be differentially expressed.

^b Differential expression could not be confirmed by quantitative PCR, suggesting that the tag was derived from a different transcript than the one indicated.

^c Not tested.

Table 4 Transcripts most repressed in adenomas and cancers^a

Tag sequence	Normal		Adenoma		Cancer		UniGene ID	Transcript name
	NC-1	NC2	AD-A1	AD-B	Tu-981	Tu-102		
GTCATCACCA	35	22	0	0	0	0	32966	Guanylate cyclase activator 2B
CCTTCAAATC	29	17	0	0	1	0	23118	Carbonic anhydrase I
TCTGAATTAT	24	16	0	0	1	0	50964	Carcinoembryonic antigen-related cell adhesion molecule 1
TTATGGTGTG	11	17	0	0	0	0	271499	ESTs ^b
CTGGCAAAGG	14	22	1	0	0	0	72789	Hypothetical protein FLJ20217
AGGTGACTGG	10	14	0	0	0	0		No match
CTTATGGTCC	36	11	0	1	1	0	179608	Retinol dehydrogenase homologue
ATGATGGCAC	12	32	1	0	1	0	84072	Transmembrane 4 superfamily member 3
GTCCGAGTGC	17	3	0	0	0	0	3337	Transmembrane 4 superfamily member 1
ATTTCAAGAT	35	21	0	2	1	0	155097	Carbonic anhydrase II (CA2)
CAAGAGTTTC	14	2	0	0	0	0	183617	ESTs
GCCATCCTCC	9	13	0	1	0	0		No match
ACCCAAGTGC	12	3	0	0	0	0	232604	<i>Homo sapiens</i> cDNA: FLJ22675 fis, clone HSI10553
GCCCACGTCA	7	8	0	0	0	0		No match
TTTGGTTTCA	2	13	0	0	0	0		No match
CTCAGAAGTTC	18	3	1	0	0	0	194710	<i>N</i> -Acetylglucosaminyl transferase 3, mucin type
CCAACACCAG	9	19	1	0	0	1	181165	Eukaryotic translation elongation factor 1 α 1
GCCACATACT	3	9	0	0	0	0	4984	KIAA0828 protein
GTATTGGGGC	5	7	0	0	0	0		No match
CCGGCTTGAG	7	4	0	0	0	0	2722	Inositol 1,4,5-trisphosphate 3-kinase A
GATATGTAAA	1	10	0	0	0	0	227059	Chloride channel, calcium-activated, family member 4
CATAGGTTTA	66	39	4	1	5	0	1650	Solute carrier family 26, member 3 (DRA)
GTCCTGAACA	7	3	0	0	0	0	78546	ATPase, Ca ²⁺ transporting, plasma membrane 1

^a These tags displayed at least 20-fold decrease in both neoplastic states. The numbers given are the raw tag counts for each tag observed in each library. Transcript name provides a description of matching UniGene cluster (Build 3/13/01). Rows shown in bold are genes that were tested and confirmed by quantitative PCR to be differentially expressed.

^b ESTs, expressed sequence tags.

macrophages, and leukocytes, as well as in the Paneth cells of the gastrointestinal tract. Fecal lysozyme levels are dramatically elevated in patients with inflammatory bowel disease (15, 16), and serum lysozyme activity is significantly elevated in patients with sarcoidosis (17), both of which are diseases characterized by aberrant chronic inflammation. Furthermore, lysozyme immunoreactivity has been observed in the epithelial cells of both adenomas and carcinomas of the large intestine (18). In our study, the degree of elevation of expression of *LYS* varied from 4-fold to 55-fold in the unpurified samples. In contrast, the degree of elevation of expression of *LYS* observed in

purified epithelial cells was only 2–5 fold. This suggested that a substantial portion of the expression for this gene in the tumors could have been derived from nonepithelial cells. Consistent with this hypothesis, *in situ* hybridization experiments revealed that the majority of *LYS* mRNA was present in a stromal component that appeared to be macrophages (Fig. 5C). The expression of *LYS* in the macrophage compartment of colorectal tumors was also supported by its high representation in a SAGE library constructed from hematopoietic cells (CD45+, CD64+, and CD14+) purified from colorectal tumors (602 *LYS* tags/56,643 total tags; Ref. 6).

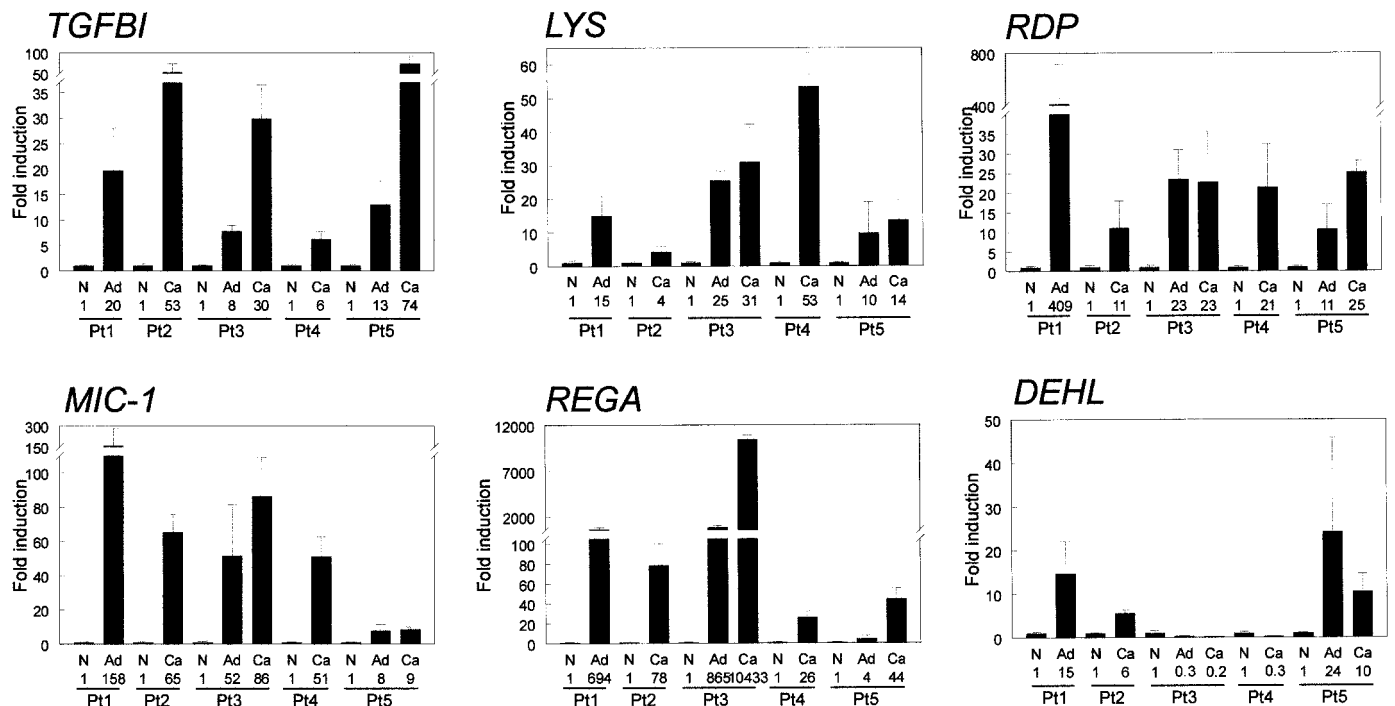
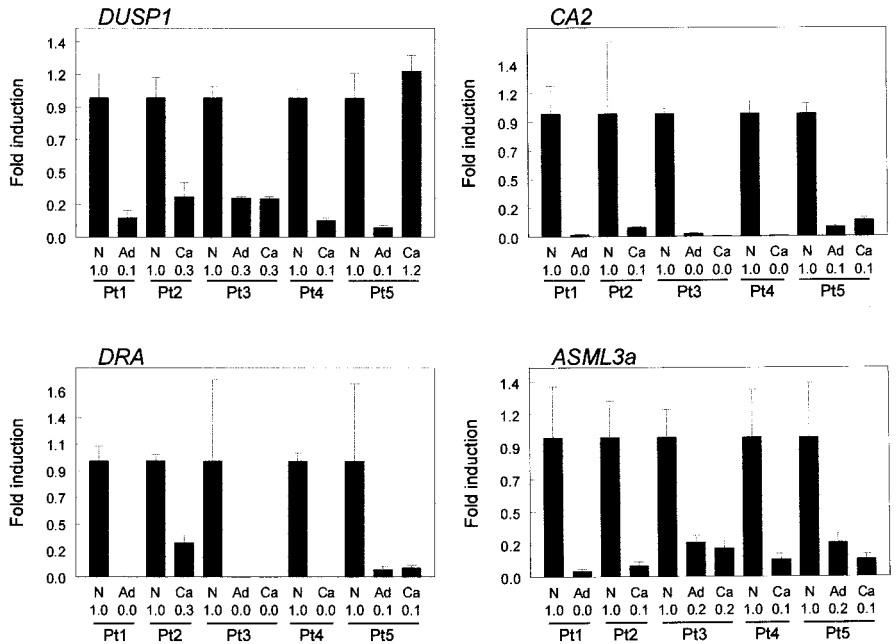


Fig. 2. Quantitative PCR analysis of genes elevated in both adenomas and cancers. Quantitation of expression of genes in tumors and matched normal tissues from five patients (Pt) is shown as fold elevation over that in matched normal colonic mucosa. Each bar represents the average of three independent measurements. *TGFBI*, *LYS*, *RDP*, *MIC-1*, *REGA*, and *DEHL* are as described in Table 3.

Fig. 3. Quantitative PCR analysis of genes decreased in both adenomas and cancers. Quantitation of expression of genes in tumors and matched normal tissue from five patients (Pt) is shown as a fraction of matched normal. Each bar represents the average of three independent measurements. CA2 and DRA are described in Table 4. Dual specificity phosphatase (*DUSP1*) and acid sphingomyelinase-like phosphodiesterase (*ASML3a*) represent transcripts that were repressed but did not meet the stringent criteria required for inclusion in Table 4. SAGE data indicated that *DUSP1* was 5- and 76-fold repressed in adenomas and cancers, respectively. *ASML3a* was 15-fold repressed in both adenoma and cancer.



One interesting gene identified in the current study is renal dipeptidase (*RDP*). *RDP* is a glycosylphosphatidylinositol-anchored enzyme whose major site of expression is the epithelial cells of the proximal tubules of the kidney (reviewed in Ref. 19). The enzyme has

been extensively analyzed with respect to its catalytic mechanism and inhibition kinetics by a variety of synthetic inhibitors. *RDP* is unique among the dipeptidases in that it can cleave amide bonds in which the COOH-terminal partner is a D-amino acid, providing an excellent opportunity for the development of specific probes for its detection *in vivo*. Quantitative PCR revealed *RDP* to be markedly elevated in both unpurified and purified tumor epithelial cells, and *in situ* hybridization experiments showed that *RDP* was localized exclusively to epithelial cells of colorectal tumors (Fig. 5D).

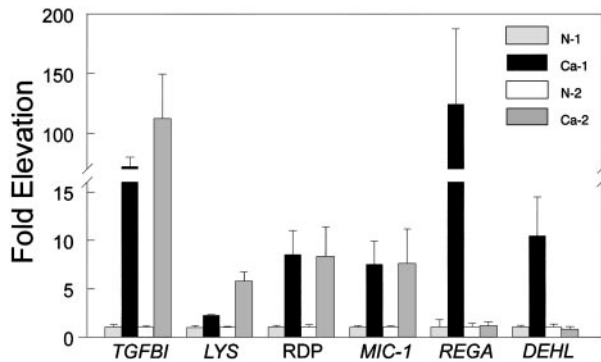


Fig. 4. Quantitative PCR analysis of mRNA from purified epithelial cells of genes elevated in both adenomas and cancers. Quantitation of expression of genes in the purified normal (N) or cancer (Ca) epithelial cells taken from two patients is shown as fold elevation over matched normal. Genes examined were the same as those in Fig. 2.

Macrophage inhibitory cytokine (*MIC-1*) is a small polypeptide of M_r 16,000 first isolated from a differential screen for genes that were induced on macrophage activation (20). Concurrently, it was identified by a search for molecules homologous to the bone morphogenic protein/TGF- β family of growth and differentiation factors (21). In addition to being highly expressed in activated macrophages, *MIC-1* has been noted to be highly expressed in placenta and the epithelial cells of normal prostate. In the current study, we found *MIC-1* expression to be elevated between 7- and 133-fold in the unpurified tumors. As observed for *LYS*, the purified tumor cells demonstrated significant but less elevation of expression of *MIC-1* (5-7-fold), indirectly implicating stromal expression to be partly responsible for

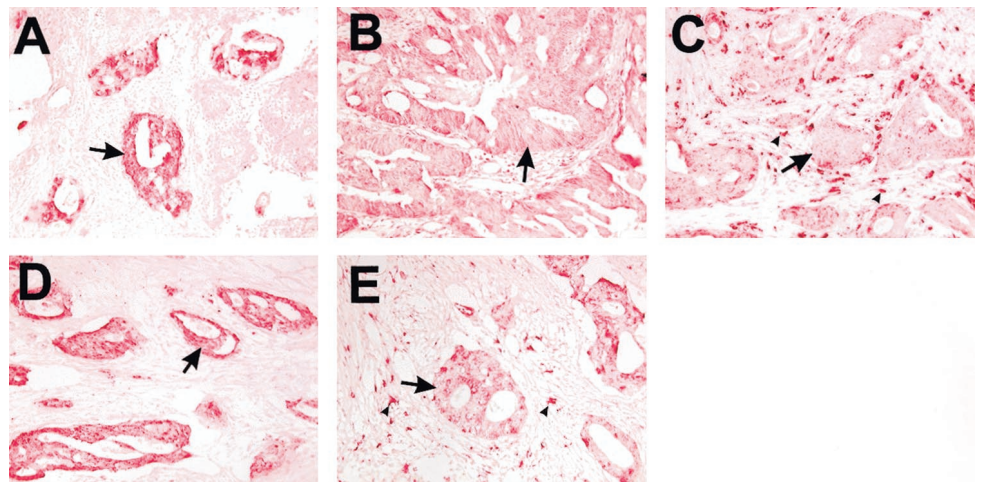


Fig. 5. *In situ* hybridization analyses of elevated genes. Genes examined were *REGA* (A), *TGFBI* (B), *LYS* (C), *RDP* (D), and *MIC-1* (E). Positive cells appear red, arrows point to clusters of malignant epithelial cells, and arrowheads point to macrophages.

the dramatic elevation seen in some tumors. Consistent with this hypothesis, *in situ* hybridization experiments revealed expression in both the epithelium of the tumor and a cell type resembling infiltrating macrophages (Fig. 5E).

The results summarized above show that although a large number of tags are observed in the colorectal tissues analyzed, only a small fraction (957 of 21,343; <5%) were expressed differentially in benign or malignant neoplastic tissues. A similarly small fraction of genes [66 of 4000 (1.7%)] were found to be aberrantly expressed in colorectal neoplasms using oligonucleotide arrays (22). Analysis of these differentially expressed genes not only has the potential to provide insights into the biology of human neoplasia but may also have clinically useful applications. One of the most exciting potential applications concerns the identification of genes whose products provide cellular and serum markers for colorectal neoplasia. The ideal tumor marker would be expected to have several characteristics. First, it should be expressed at high levels in tumors and at greatly reduced levels in normal tissues. Second, the elevated expression should occur early and remain elevated during the neoplastic process. Third, such a marker should be elevated in the majority of clinical samples. Fourth, the marker should be expressed on the cell surface or secreted to facilitate its detection. Previous application of a subset of these criteria led to the identification of a marker that was elevated in the serum of patients with pancreatic cancer (23). In the current study, we identified several genes that appeared to meet all of these criteria and may therefore be especially useful for the development of diagnostic tools for the early detection of presymptomatic colorectal neoplasia. Indeed, the product of one of these genes (*MIC-1*) has recently been found to be elevated in the serum of patients with colorectal and other cancers, providing further validation of this approach.⁶

References

1. Cancer Facts & Figures 1998. American Cancer Society, 1998.
2. Inger, D. B. Colorectal cancer screening. *Prim. Care*, 26: 179–187, 1999.
3. Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. Serial analysis of gene expression. *Science* (Wash. DC), 270: 484–487, 1995.
4. Zhang, L., Zhou, W., Velculescu, V. E., Kern, S. E., Hruban, R. H., Hamilton, S. R., Vogelstein, B., and Kinzler, K. W. Gene expression profiles in normal and cancer cells. *Science* (Wash. DC), 276: 1268–1272, 1997.
5. Lal, A., Lash, A. E., Altschul, S. F., Velculescu, V., Zhang, L., McLendon, R. E., Marra, M. A., Prange, C., Morin, P. J., Polyak, K., Papadopoulos, N., Vogelstein, B., Kinzler, K. W., Strausberg, R. L., and Riggins, G. J. A public database for gene expression in human cancers. *Cancer Res.*, 59: 5403–5407, 1999.
6. D. A. Brown, T. Liu, R. L. Ward, N. J. Hawkins, W. D. Fairlie, A. R. Bauskin, P. J. Russell, D. I. Quinn, J. J. Grygiel, A. G. Moore, R. L. Sutherland, J. Turner, E. A. Kingsley, and S. N. Breit. Macrophage inhibitory cytokine-1 (MIC-1) in epithelial neoplasia, submitted for publication.
6. St Croix, B., Rago, C., Velculescu, V., Traverso, G., Romans, K. E., Montgomery, E., Lal, A., Riggins, G. J., Lengauer, C., Vogelstein, B., and Kinzler, K. W. Genes expressed in human tumor endothelium. *Science* (Wash. DC), 289: 1197–1202, 2000.
7. Guy-Crotte, O., Amouric, M., and Figarella, C. Characterization and N-terminal sequence of a degradation product of 14,000 molecular weight isolated from human pancreatic juice. *Biochem. Biophys. Res. Commun.*, 125: 516–523, 1984.
8. Terazono, K., Yamamoto, H., Takasawa, S., Shiga, K., Yonemura, Y., Tochino, Y., and Okamoto, H. A novel gene activated in regenerating islets. *J. Biol. Chem.*, 263: 2111–2114, 1988.
9. Watanabe, T., Yonekura, H., Terazono, K., Yamamoto, H., and Okamoto, H. Complete nucleotide sequence of human *reg* gene and its expression in normal and tumoral tissues. The *reg* protein, pancreatic stone protein, and pancreatic thread protein are one and the same product of the gene. *J. Biol. Chem.*, 265: 7432–7439, 1990.
10. Rechreche, H., Montalto, G., Mallo, G. V., Vasseur, S., Marasa, L., Soubeyran, P., Dagorn, J. C., and Iovanna, J. L. *pap*, *reg I α* and *reg I β* mRNAs are concomitantly up-regulated during human colorectal carcinogenesis. *Int. J. Cancer*, 81: 688–694, 1999.
11. Skonier, J., Neubauer, M., Madisen, L., Bennett, K., Plowman, G. D., and Purchio, A. F. cDNA cloning and sequence analysis of β ig-h3, a novel gene induced in a human adenocarcinoma cell line after treatment with transforming growth factor- β . *DNA Cell Biol.*, 11: 511–522, 1992.
12. Escribano, J., Hernando, N., Ghosh, S., Crabb, J., and Coca-Prados, M. cDNA from human ocular ciliary epithelium homologous to β ig-h3 is preferentially expressed as an extracellular protein in the corneal epithelium. *J. Cell. Physiol.*, 160: 511–521, 1994.
13. Munier, F. L., Korvatska, E., Djemai, A., Le Paslier, D., Zografos, L., Pescia, G., and Schorderet, D. F. Kerato-epithelin mutations in four 5q31-linked corneal dystrophies. *Nat. Genet.*, 15: 247–251, 1997.
14. Fleming, A. On a remarkable bacteriolytic element found in tissues and secretions. *Proc. R. Soc.*, 93: 306–317, 1922.
15. Sugi, K., Saitoh, O., Hirata, I., and Katsu, K. Fecal lactoferrin as a marker for disease activity in inflammatory bowel disease: comparison with other neutrophil-derived proteins. *Am. J. Gastroenterol.*, 91: 927–934, 1996.
16. van der Sluys Veer, A., Brouwer, J., Biemond, I., Bohbouth, G. E., Verspaget, H. W., and Lamers, C. B. Fecal lysozyme in assessment of disease activity in inflammatory bowel disease. *Dig. Dis. Sci.*, 43: 590–595, 1998.
17. Tomita, H., Sato, S., Matsuda, R., Sugiura, Y., Kawaguchi, H., Niimi, T., Yoshida, S., and Morishita, M. Serum lysozyme levels and clinical features of sarcoidosis. *Lung*, 177: 161–167, 1999.
18. Ho, S. B., Itzkowitz, S. H., Frieri, A. M., Jiang, S. H., and Kim, Y. S. Cell lineage markers in premalignant and malignant colonic mucosa. *Gastroenterology*, 97: 392–404, 1989.
19. Yamada, R., and Kera, Y. D-amino acid hydrolysing enzymes. *EXS* (Basel), 85: 145–155, 1998.
20. Bootcov, M. R., Bauskin, A. R., Valenzuela, S. M., Moore, A. G., Bansal, M., He, X. Y., Zhang, H. P., Donnellan, M., Mahler, S., Pryor, K., Walsh, B. J., Nicholson, R. C., Fairlie, W. D., Por, S. B., Robbins, J. M., and Breit, S. N. MIC-1, a novel macrophage inhibitory cytokine, is a divergent member of the TGF- β superfamily. *Proc. Natl. Acad. Sci. USA*, 94: 11514–11519, 1997.
21. Paralkar, V. M., Vail, A. L., Grasser, W. A., Brown, T. A., Xu, H., Vukicevic, S., Ke, H. Z., Qi, H., Owen, T. A., and Thompson, D. D. Cloning and characterization of a novel member of the transforming growth factor- β /bone morphogenetic protein family. *J. Biol. Chem.*, 273: 13760–13767, 1998.
22. Notterman, D. A., Alon, U., Sierk, A. J., and Levine, A. J. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. *Cancer Res.* 61: 3124–3130, 2001.
23. Zhou, W., Sokoll, L. J., Brusek, D. J., Zhang, L., Velculescu, V., Goldin, S. B., Hruban, R. H., Kern, S. E., Hamilton, S. R., Chan, D. W., Vogelstein, B., and Kinzler, K. W. Identifying markers for pancreatic cancer by gene expression analysis. *Cancer Epidemiol. Biomark. Prev.*, 7: 109–112, 1998.

Cancer Research

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

Secreted and Cell Surface Genes Expressed in Benign and Malignant Colorectal Tumors

Phillip Buckhaults, Carlo Rago, Brad St. Croix, et al.

Cancer Res 2001;61:6996-7001.

Updated version Access the most recent version of this article at:
<http://cancerres.aacrjournals.org/content/61/19/6996>

Cited articles This article cites 22 articles, 10 of which you can access for free at:
<http://cancerres.aacrjournals.org/content/61/19/6996.full#ref-list-1>

Citing articles This article has been cited by 29 HighWire-hosted articles. Access the articles at:
<http://cancerres.aacrjournals.org/content/61/19/6996.full#related-urls>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link
<http://cancerres.aacrjournals.org/content/61/19/6996>.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.