

Transcriptomic Molecular Markers for Screening Human Colon Cancer in Stool and Tissue

FARID E. AHMED^{1*}, PAUL VOS², STEPHANIE IJAMES³, DONALD T. LYSLE³,
RON R. ALLISON¹, GORDON FLAKE⁴, DENNIS R. SINAR⁵, WADE NAZIRI⁶,
STEFAN P. MARCUARD⁷ and RODNEY PENNINGTON⁸

Departments of ¹Radiation Oncology, LSB 003, Leo W. Jenkins Cancer Center and

⁵Internal Medicine, The Brody School of Medicine (BSOM)

at East Carolina University (ECU), Greenville, NC 27858;

²Department of Biostatistics, School of Allied Health Sciences, ECU, Greenville, NC 27858;

³Department of Psychology, University of North Carolina, Chapel Hill, NC 27599;

*⁴Laboratory of Experimental Pathology, National Institute of Environmental
Health Sciences, Research Triangle Park, NC 27709;*

⁶Carolina Physicians PA, Greenville, NC 27834;

⁷Carolina Digestive Disease, Greenville, NC 27858;

⁸Roche Applied Science, Indianapolis, IN 46250, U.S.A.

Abstract. *There is a need for sensitive and specific diagnostic molecular markers that can be used to monitor early patterns of gene expression in non-invasive exfoliated colonocytes shed in the stool, and in situ in adenoma-carcinoma epithelium of the colon. RNA-based detection methods are more comprehensive than either DNA-, protein- or methylation-based*

screening methods. By routinely and systematically being able to perform quantitative gene expression studies on these samples using less than ten colon cancer genes selected by the enormous resources of the National Cancer Institute's Cancer Genome Anatomy Project, we were able to monitor changes at various stages in the neoplastic process, allowing for reliable diagnostic screening of colon cancer particularly at the early, pre-malignant stages. Although the expression of some of the genes tested in tissue showed less variability in normal or cancerous patients than in stool, the stool by itself is suitable for screening. Thus, a transcriptomic approach using stool or tissue samples promises to offer more sensitivity and specificity than currently used molecular screening methods for colon cancer. A larger prospective clinical study utilizing stool and tissue samples derived from many control and colon cancer patients, to allow for a statistically valid analysis, is now urgently required to determine the true sensitivity and specificity of the transcriptomic screening approach for this preventable cancer.

Abbreviations: ATCC, American Type Culture Collection; cDNA, copy deoxyribonucleic acid; CD, Crohn's disease; CEA, carcinoembryonic antigen; CGAP, Cancer Genome Anatomy Project; CP, comparative cross point; CRC, colorectal cancer; CT, computed tomography; DEPC, Diethyl pyrocarbonate; DGED, Digital Gene Expression Display; E-Method, also referred to as Second Derivative Maximum or CP method; EST, expressed sequence tag; FOBT, fecal occult blood test; GI, gastrointestinal; GLS, Gene Library Summarizer; H&E, Hematoxylin and Eosin staining; IBD, inflammatory bowel disease; LCM, laser capture microdissection; mRNA, messenger ribonucleic acid; NCI, National Cancer Institute; OR, odd ratio; QC, quality control; RT-qPCR, reverse transcriptase quantitative polymerase chain reaction; rRNA, ribosomal ribonucleic acid; SAGE, Serial Analysis of Gene Expression; ss, single stranded; UC, ulcerative colitis; UDG, uracil-DNA glycosylase.

Correspondence to: Dr. Farid E. Ahmed, Clinical Professor of Molecular Oncology, The BSOM at ECU, Greenville, NC 27858, U.S.A. Tel: +252 744 4636, Fax: +252 744 3775, e-mail: ahmedf@ecu.edu

Key Words: Adenocarcinoma, colonocyte, diagnosis, IBD, LCM, genes, QC, RT-qPCR, RNA.

Colorectal cancer (CRC) is the second and third most common malignancy in men and women, respectively, in developed and developing countries, including the United States of America (USA) (1, 2). In the USA, an estimated 106,680 cases of colon and 41,930 cases of rectal cancer are expected to occur in 2006, of which 55,170 deaths are estimated to materialize, which account for ~10% of all cancer deaths. Incidence rates decreased by 1.8% per year during 1998-2002 in the USA (Table I), partially reflecting increased screening and polyp removal during colonoscopy screening. CRC incidence rates have been steadily

Table I. *Estimated new and Existing CRC cases, deaths and incidence rates in the ten States in the USA with the greatest prevalence[†].*

State	New CRC cases in 2006#	Estimated CRC deaths in 2006*	CRC Incidence rates, 1998-2002**		CRC Death rates, 1998-2002***	
			Male	Female	Male	Female
California	14,820	5,500	57.2	42.0	21.0	15.2
Florida	9,970	3,700	64.6	47.9	22.4	15.8
New York	9,540	3,540	71.9	52.8	26.2	18.4
Texas	9,510	3,530	58.5	41.1	23.7	16.2
Pennsylvania	8,000	2,970	74.6	52.3	28.0	19.4
Illinois	6,760	2,510	72.1	51.0	28.0	19.1
Ohio	6,730	2,500	67.1	94.2	27.6	19.2
Michigan	4,930	1,830	64.6	47.9	24.7	17.0
New Jersey	4,850	1,800	75.5	53.7	27.4	19.4
Virginia	3,690	1,370	58.5	41.1	25.0	18.2

[†]Source: Modified from Reference (2). #Estimated new colon cancer cases by sex for all locations in the USA in 2006 are: 106,680 for both sexes (49,220 for males and 57,460 for females). Estimated new colon cancer deaths by sex for all sites in the USA in 2006 are 55,170 for both sexes (27,870 for males and 27,300 for females). All estimates are rounded to the nearest 10. *Rounded to the nearest ten. **Per 100,000, age adjusted to the 2000 USA standard population. ***Per 100,000, age-adjusted to the 2000 US standard population.

decreasing since 1985, from 66 to 52 per 100,000 in 2002 (3). Mortality rates from CRC have also decreased in both genders over the past two decades at an average of 1.8% per year, reflecting declining incidence rates and improved survival. The 1- and 5-year relative survival for CRC for all stages combined is 83% and 64%, respectively. Survival continues to decline beyond 5 years to 5% at 10 years after diagnosis (2); thus, early detection contributes significantly to the prevention of death from this cancer (4-7).

The most commonly used screening tests in the USA for colon adenomas in men and women, aged ≥ 50 years old, are the fecal occult blood test (FOBT) and colonoscopy. The former, although convenient and relatively inexpensive, suffers from low sensitivity, whereas the latter – considered the gold standard for CRC screening – is expensive and requires cathartic preparation and patient sedation, which has resulted in a low rate of compliance (8). Computed tomographical (CT) colonography (virtual colonoscopy), which does not require bowel preparation may be considered an adequate alternative only for asymptomatic, not at risk individuals, and only if it can effectively improve the detection of small lesions (9, 10).

Cells are continuously shed by colon tumors in the lumen of the gastro intestine (GI) (*i.e.*, approximately 10^{10} normal adult colonic epithelial cells, each having a lifespan of 3-4 days, are shed daily from the lower two thirds of colon crypts) (11, 12), and their detection in the stool has allowed for the employment of mutation or other functional genomic techniques in their study (1, 6, 13-15). Since CRCs exhibit genetic heterogeneity, multitarget approaches employing mutations in *K-ras*, *APC* and *p53*, the microsatellite instability

marker Bat-26 and "long" DNA (representing DNA of nonapoptotic colonocytes characteristic of cancer cells exfoliated from neoplasms, but not normal apoptotic colonocytes) have been examined and undergone clinical testing (13, 16). However, DNA alterations were disappointingly detected in only 16 of 31 (51.6%) invasive cancer, 29 of 71 (40.8%) invasive cancer plus adenoma with high-grade dysplasia, and 76 of 418 (18.2%) in patients with advanced neoplasia (tubular adenoma ≥ 1 cm in diameter, polyps with high grade dysplasia, or cancer) (17). Protein-based methods are not suited for screening and early diagnosis because they are generally not specific, although they may be of more value as prognostic markers (18). More recently, by employing commercial preparations, we have overcome RNA's lability by stabilizing it within a short period of time after samples (*e.g.*, stool, tissue or blood) were removed from the body, resulting in a total RNA that was readily reverse-transcribable by another commercial preparation making a high quality single-stranded (ss) copy (c) DNA suitable for expression profiling (19-21). Therefore, the identification of new transcriptomic molecular markers with high sensitivities and specificities in exfoliated stool samples is now possible.

Materials and Methods

Adenocarcinoma cell line and culture conditions. Adenocarcinoma cell line HT-29 is used for validating the range of gene expression measurements in stool spiking studies. Cells were obtained from the American Type Culture Collection (ATCC), Rockville, MD, USA. The cells were propagated in Iscore's Modified Dulbeccos medium (IMDM) (Sigma, St. Louis, MO, USA), supplemented with 100 ml/L fetal calf serum, 105 IU/L penicillin and 0.1 g/L

streptomycin in an atmosphere of 5 % CO₂ in a humidified incubator kept at 37°C. Cultures were passed twice per week as per ATCC recommendations.

Acquisition of clinical specimens. Stool and tissue samples were obtained from twenty control subjects and thirty patients with various stages of colon adenocarcinoma (Dukes' stages 0 to 3), five patients with ulcerative colitis (UC) and 5 patients with Crohn's disease (CD) according to an approved ECU Medical Center Institutional Review Board (IRB) protocol. All laboratory work was carried out and standardized under blind conditions and in accordance with the guidelines for handling biohazardous material established by ECU's Biological Safety and Hazardous Substance Committee.

Control stool and tissue samples.

(i) *Fecal specimens.* Control stool samples were collected from consenting individuals visiting our GI Clinic/Endoscopy Lab who did not show any polyps or inflammatory bowel diseases, such as colitis or diverticulitis. Stool samples were either processed immediately to extract RNA, or stored overnight at 4°C in a bacteriostatic preservative S.T.A.R medium (Roche Diagnostics, Indianapolis, IN, USA), and RNA was extracted within a few days. In situations where longer storage of fecal specimens was desired, the preservative RNALater®-ICE (Ambion, Austin, TX, USA) was added at 2.5 ml per 1 g of stool, followed by freezing of the stool sample at -70°C.

(ii) *Tissue specimens.* Normal tissues were usually obtained from a small piece of colon tissue (about 0.5 cm³) removed >10 cm away from diseased patient tissue at surgery (22), or from biopsies taken during colonoscopy from non-diseased areas of consenting individuals. For UC or CD patients, a small piece of tissue taken further away from the inflamed or diseased tissue was considered normal. Tissues were flash frozen in liquid nitrogen and stored at -70°C for subsequent laser capture microdissection (LCM) work. Longitudinal sectioning of the tissue before LCM use was employed in order to pick up only the epithelial cells that would eventually be shed as colonocytes into the lamina propria from the bottom of epithelial cells among the proliferative enterocyte crypt lineage.

Experimental stool and tissue samples from cancerous or inflamed patients.

(i) *Fecal specimens.* A 10 g sample of feces (bowel movement) was collected the night before surgery or earlier, before administering any bowel preparation, in a plastic container containing either: a) a bacteriostatic preservative S.T.A.R. medium, which was then covered and either processed immediately to extract total RNA or stored overnight at 4°C then processed the next day for RNA extraction, or b) RNALater®-ICE to allow longer storage at -70°C, then the sample followed either RT-PCR processing or storage of the extracted RNA at -70°C until further manipulation and PCR analysis. Stool processing was standardized for all samples by scraping and employing the surface mucinous layer, which is usually rich in colonocytes (13).

(ii) *Tissue specimens.* A small piece of tissue sample (about 0.5 cm³) was obtained after colonoscopy for adenoma, or at surgery for carcinoma. Samples were processed after flash freezing in liquid N₂ and storage at -70°C for subsequent microdissection. Longitudinal LCM sectioning was performed (Figure 1A, B), and the marked areas of the crypt indicated where the transformed cells (*i.e.*, adenoma, carcinoma) were to be captured by laser microdissection (Figure 1C-E) for subsequent RNA extraction.

For the current study, stool and normal tissue samples were obtained from 20 non-cancerous non IBD control individuals, 20 patients having adenomatous polyps ≥1 cm with high grade dysplasia (stage 0-1), 5 patients with stage 2 carcinoma, five patients with stage 3 carcinoma, five non-cancerous patients with severe UC, and five non-cancerous patients with severe CD. Each subject provided a stool sample for a total of 60 stool samples. Tissue samples were obtained from only one of the UC patients and only one of the CD patients, but were obtained for each of the remaining patients, for a total of 52 tissue samples.

Selection of cancerous or inflamed cells from colon tissue of patients by LCM. LCM was employed as an enrichment technique for tumors isolated from colon adenocarcinoma patients to separate transformed cells from nonneoplastic stromal and inflammatory cells (23). The frozen tissues, embedded in Tissue Tek OCT compound (Sakura, MI, USA) were transported in cold packs to Laboratory of Experimental Pathology, National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, North Carolina. There, LCM was performed using an Arcturus PixCell II system (Arcturus Engineering, Inc., Mountain View, CA, USA), which employed a 15 µm diameter infrared (IR) laser pulse (220 mV, 49 mW) with a duration of 2.2 ms to microdissect only the tumor cells (24). Approximately 20,000 cells were captured for each preparation. The LCM samples, adhering to the thermoplastic polymer film on the plastic cap (Figure 1C-E), were sectioned at 6-µm in a cryostat and picked up on non-charged microscopic slides (Fisher Scientific, Pittsburgh, PA, USA). The slides were kept in a slide holder on dry ice, fixed for 30 sec in 70% ethanol, dipped in distilled water for 15 sec, stained in Mayer's Hematoxylin for 15 sec, rinsed for 15 sec in 1X automation buffer, pH 7.5 (Biomedica Corp. Foster City, CA, USA), again rinsed in distilled water followed by 70% ethanol for 30 sec each, counterstained in Eosin Y (Cell Point Scientific, Gaithersburg, MD, USA) for 30 sec, followed by dehydration in graded ethanol solutions (95, 100 and 100%), 30 sec each, cleared by two rinses in xylene, 1 min each, air dried for 5 min, and stored in a slide box in a dessiccator for up to 3 h before LCM. Captured cells were fitted to a 0.5 ml sterile microcentrifuge tube, and returned to ECU in cold packs for RNA extraction.

Manual extraction of total RNA from LCM cells and ss-cDNA preparation. This manual procedure was used for extracting RNA from a small number of LCM captured cells was carried out according to manufacture's specifications using the RNeasy isolation Kit® from Qiagen, Valencia, CA, USA, as previously described (19, 21). The quality of RNA was determined on an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc, Palo Alto, CA, USA) utilizing the RNA 6000 Nano LabChip®, or by electrophoresis on Superload (Viagen, Austin, TX, USA) agarose gels (25) and RNA quantitated with RiboGreen quantitation reagent (26) (Molecular Probes, Eugene, OR, USA). The "Sensiscript RT Kit®" from Qiagen was then employed for making a copy of ss-DNA, resulting in 40 µl of ss-cDNA, of which 2-3 µl was subsequently amplified by PCR. One hundred thousand captured cells on 5 plastic LCM caps (each accommodating 20,000 cells) was enough to test all the 11 genes of interest, considering that each cell contains ~20 pg total RNA or 0.4 pg mRNA (equivalent to 0.36 pg ss-cDNA), as only a few picograms of cDNA are needed per PCR reaction (27).

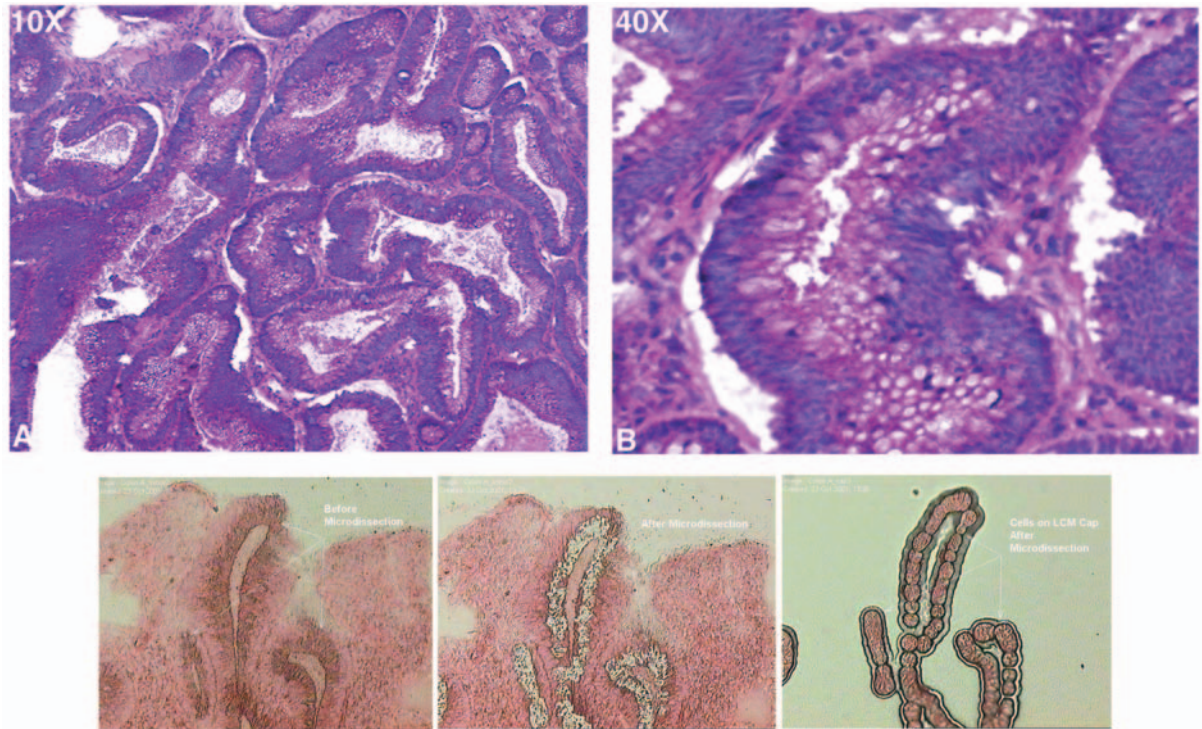


Figure 1. A) Longitudinal H & E cryostat section of colon adenoma exhibiting high grade dysplasia (i.e., carcinoma in situ, stage 0), 10X. B) As A x 40. C-E: LCM displaying dysplastic cells from the above section before being pulsed by an IR laser (C); the middle panel shows same area with dysplastic cells removed (D); and the right panel shows removed dysplastic regions on a film cap (E).

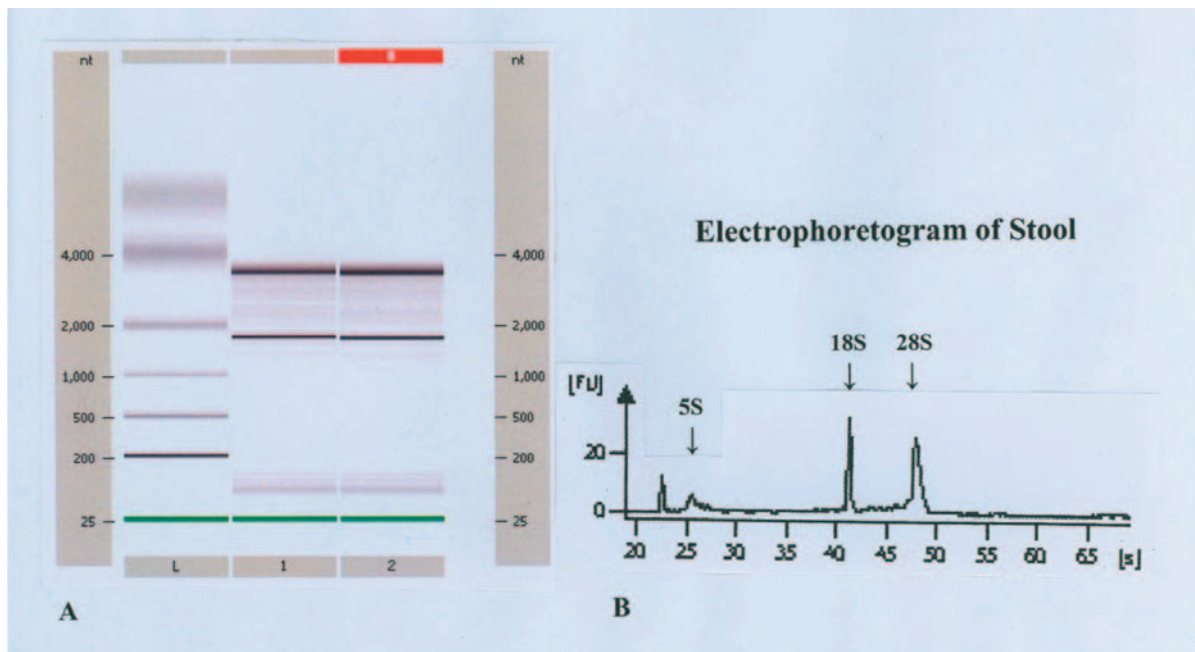


Figure 2. A) Representative Superload agarose gel protocol for a stool sample showing a nondegraded RNA. Lane L is molecular marker, and lanes 1 and 2 are replicates of a 4 ng total RNA of the same stool sample. B) An Agilent 2100 electrophoretogram showing the 28S, 18S and tRNA, 5.8S and 5S bands for the same stool sample as in panel A.

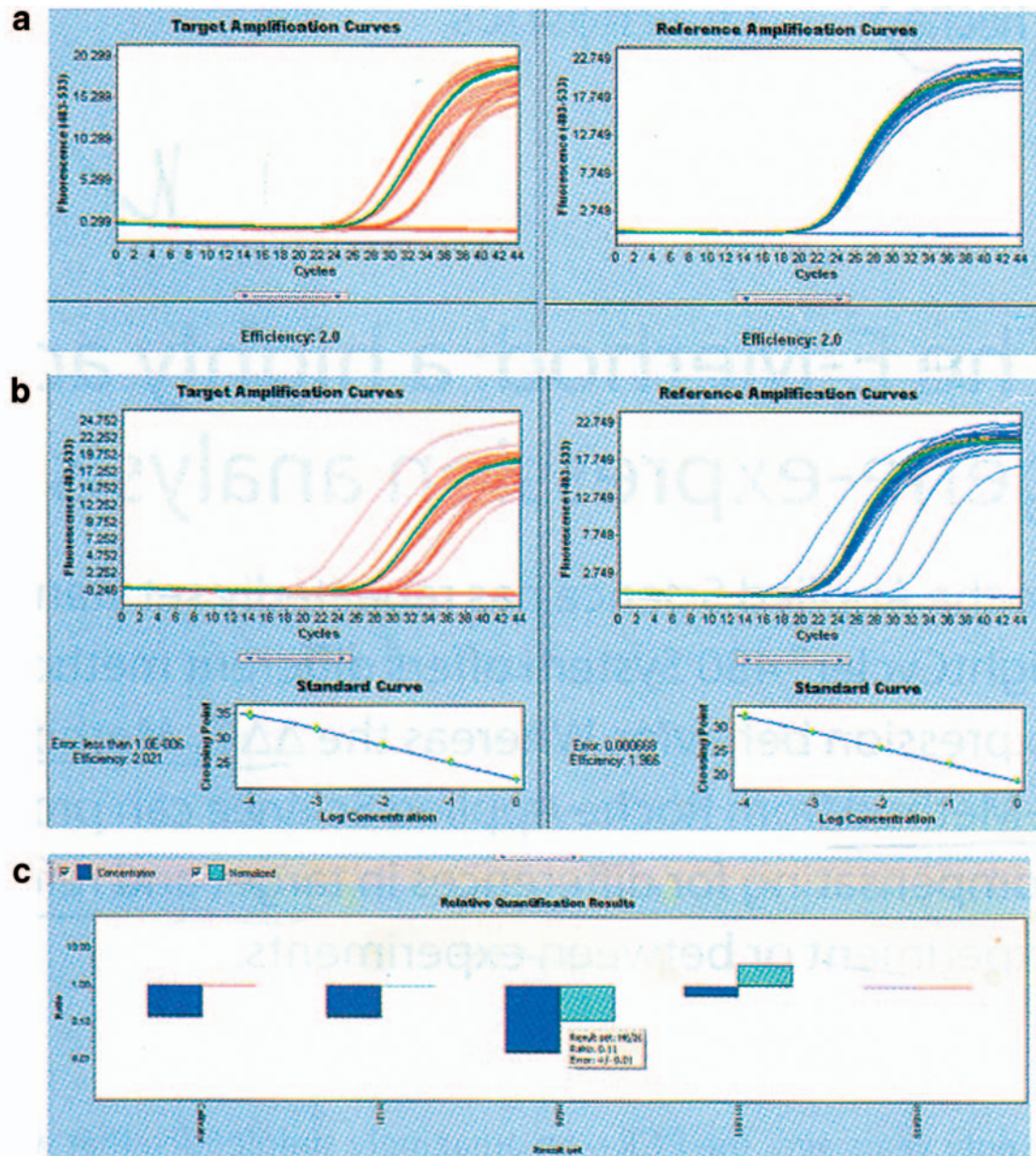


Figure 3. Two different relative quantification analysis of the same run for the $2^{-\Delta\Delta C_T}$ method, upper panel (a) or E-method (b). The final quantification results are automatically calculated from the crossing point (CP) values of the target and the reference gene (unknowns and calibrators) as shown in the bottom panel (c). Adapted from reference 29.

Automatic extraction of total RNA from stool samples. In collaboration with colleagues at Roche Applied Science (Indianapolis, IN, USA), total RNA from stool was automatically extracted by the "MagNA pure LC™" automated system. The MagNA Pure LC™ is a compact benchtop robotic workstation programmed to automatically perform separation of nucleic acids. Proprietary magnetic glass particles are used for separation of RNA, followed by transfer of eluted RNA (in a 100 µl volume) into a storage cartridge, which keeps samples cool at 4 °C until removed.

The Roche RNA Isolation Kit II™ was used. The machine can automatically pipette purified RNA into borosilicate capillaries of a LightCycler's PCR instrument in 2-3 µl volume (28) for high throughput qPCR analyses.

Stool samples preserved in bacteriostatic S.T.A.R. medium were shipped overnight to Roche in cold packs, and the RNA extracted from the samples (100 µl eluates kept in 0.5 ml sterile Eppendorf tubes) were returned to us overnight in cold packs. The quality and yield of the RNA extraction for each sample was determined in an

Table II. Primers/universal probes tested in PCR gene amplifications for colon cancer detection in stool and tissue

Primer	Sequence (Sense/Antisense)	Length	Position	Tm*	%GC	Amplicon	Probe	GenBank Accession #
A. Housekeeping gene standard								
HPRT	F 5'caacaggcttttctcggtt 3'	22	114-135	59	41	74 nt	#6 ttcctctg	M21641.1
	R 5'ggctactctgcccatgaaga 3'	21	167-187	59	48			
B. Genes highly expressed in SAGE data set								
IGF2	F 5' gctggcagaggagtgtcc 3'	18	653-670	59	67	109 nt	#10 ggaggtgg	NM_000612.3
	R 5' gattccattgtgtctgga 3'	20	742-761	60	67			
FLNA	F 5' tcgctctcaggaaacagca 3'	18	38-55	59	56	78 nt	#32 ctgctccc	NM_001456.1
	R 5' ttaattaagtcgcaggcaccta 3'	23	93-115	60	39			
TGF \leq -igh F	5' gacacctttgagaccctcg 3'	20	798-817	59	55	61 nt	#5 tgtggctg	NM_000358.1
	R 5' ctcaagcatcgtgttgagc 3'	20	839-858	59	50			
C. Genes highly expressed in microarray data sets								
CKS2	F 5' ttcgacgaactacgagtacc 3'	22	132-153	59	50	124 nt	#25 tggaggag	NM_001827.1
	R 5' agcctagactctgttgacacc 3'	22	234-255	59	55			
CSE1L	F 5' agattctgctaacaacctttcaa 3'	25	1786-1810	59	32	96 nt	#72 ttcctggc	NM_001316.2
	R 5'ggagagaaaaacttctcatgatagc 3'	25	1857-1881	59	32			
CXCL3	F 5' aaaatcatcgaaaagatactgaaca 3'	26	445-470	59	27	110 nt	#4 cttcctgc	NM_002090.2
	R 5' ggtaagggcagggaccac 3'	18	537-554	60	27			
D. Genes showing increased expression in CGAP cDNA DGED & SAGE DEGD databases								
DPEP1	F 5' gaggtactcgggacctgtgc 3'	20	161-180	60	65	68 nt	#53 tggcagag	NM_004413.1
	R 5' gcagaaggatgagcttcagg 3'	20	209-228	60	55			
KLK10	F 5' actgggagaagcctgtattcc 3'	21	133-153	59	52	92 nt	#15 gagcagga	NM_002776.3
	R 5' ctcattggccaggatctgc 3'	18	207-224	60	61			
E. Gene underexpressed in both SAGE and microarray data sets								
GUCA2B	F 5' ggaccttcagcctgtctgc 3'	19	251-269	60	63	76 nt	#11 cttccagc	NM_007102.1
	R 5' gtcgtctgttagcgtatggtc 3'	9	308-326	59	58			
F. Inflammatory gene								
IL-12	F 5' cactcccaaaacctgctgag 3'	20	435-454	60	55	88 nt	#50 gctccaga	NM_000882
	R 5' tctcttcagaagtccaagggtta 3'	22	501-522	59	45			

*Melting temperature.

Agilent 2100 Bioanalyzer and quantitated using RiboGreen RNA reagent. Good preparations showed two sharp ribosomal 18S and 28S rRNA bands, which corresponded well to the expected sizes of 1.9 kb and 3.7 kb, respectively, when electrophoresed on Superload formaldehyde-agarose gels (Viagen), in addition to smaller peaks of 5S rRNA and other micro rRNAs below 0.2 kb (Figure 2A, B).

Extraction of total RNA from HT-29 Cells and ss-cDNA preparation.

The extraction of total RNA from HT-29 cells to validate the detection ability of qPCR was carried out by adding 1 ml of cold (4°C) TRI REAGENT (TR-118, from Molecular Research Center, Inc., Cincinnati, OH, USA) to 10⁶ cells, extracting the total RNA according to the manufacturer's specifications. The extracted RNA was divided into several aliquots and stored at -20°C in diethyl pyrocarbonate (DEPC)-treated water, and screened for RNA integrity with the Agilent 2100 Bioanalyzer. RNA was quantitated using RiboGreen RNA quantitation reagent.

Single-stranded cDNA from RNA extracted from HT-29 cells was made by heating 54 µl of a solution containing 5 µg total RNA

in DEPC-treated water at 94°C for 5 min, followed by placing the sample on ice and adding 3 µl of oligo dT₁₂₋₁₈ primer (0.5 µg/µl), 18 µl of first-strand buffer (Invitrogen Life Technologies, Carlsbad, CA, USA), 7.5 µl of ribonuclease inhibitor (10 µg/µl) (Invitrogen), 4.5 µl of dNTP mix (2.5 mM), and 7.5 µl of Superscript III RNase H⁻ reverse transcriptase (200 U/µl) (Invitrogen), total volume 90 µl. The reaction was incubated at 37°C for 1 h and stopped by heating at 94°C for 5 min (20). The QIAquick® cleanup kit (Qiagen) was used to remove the RT enzyme and cDNA was eluted with two washes of 50 µl of kit buffer.

Two-step polymerase chain reaction on ss-cDNA. Both conventional (qualitative end-point PCR) and real-time qPCR were used to study the expression of selected genes in a two-step RT-PCR, as this method is preferable to the one-step RT-PCR for experiments requiring the same RT product to be used for analysis of multiple transcripts (19, 24).

(i) **Qualitative end-point PCR.** Qualitative endpoint PCR on representative samples was carried out in an Applied Biosystem

Table III. *Genes tested for aberrant expression in stool and tissue samples.*

Gene	Full Name	Gene (cancer)*	Total (cancer)*	Gene (normal)*	Total (normal)*	OR	p-value**	GenBank #
A. Housekeeping gene standard								
HPRT***	Hypoxanthine phosphoribosyl transferase							M21641
B. Genes highly expressed in SAGE data set (from reference 62)								
IGF2	Insulin-like growth factor II	0	98089	179	643586	Infinity	0.00	NM_000612
FLNA	Filamin A, α	0	98089	131	643586	Infinity	0.00	NM_001456
TGF β -igh	Transforming growth factor β -igh induced	0	98089	59	643586	Infinity	0.00	NM_000358
C. Genes highly expressed in microarray data sets (from references 66 to 81)								
CKS2	CDC28 protein kinase regulatory subunit	0 2	98089	49	643586	Infinity	0.00	NM_001827
CSE1L	Chromosome segregation 1-like	0	98089	41	643586	Infinity	0.00	NM_001316
CXCL3	Chemokine (C-X-C motif) Ligand 3	0	98089	15	643586	Infinity	0.12	NM_002090
D. Genes showing increased expression in CGAP cDNA DGED & SAGE DEGD databases								
DPEP1	Dipeptidase 1	0	98089	24	643586	Infinity	0.00	NM_004413
KLK10	Kallikrein 10	0	98089	23	643586	Infinity	0.00	NM_002776
E. Gene underexpressed in both SAGE and microarray data sets (from references 64 and 68)								
GUCA2B	Uroguanylin	57	98089	0	643586	0	0.00	NM_007102
F. Inflammatory gene								
IL-12	Interleukin-12	0	98089	0	643586	Infinity	–	NM_000882

*Parameters employed in the DGED gene selection tool (<http://cgap.nci.nih.gov/Tissues/Significance>) to calculate the odd ratio (OR): Total (cancer)/Gene (cancer) (89). Total (normal)/Gene (normal). **Obtained from virtual Northern blot of CGAP database (89). ***Low copy number housekeeping gene standard.

9600 thermocycler (Foster City, CA, USA) to validate the amplified products. A master mix was used containing final concentrations of 1X high fidelity PCR buffer, 0.2 mM dNTP, 2 mM MgSO₄, 0.4 μ M forward and reverse primers, 0.1 ng ss-cDNA template and 1 U of "hot start" Platinum High Taq DNA polymerase (Invitrogen) in a final volume of 25 μ l in a 100 μ l plastic PCR tube. Running conditions were: one cycle at 94°C for 3 min to activate the hot start Taq, 35 cycles of 94°C denaturation for 45 sec, 55°C annealing for 1 min and 72°C elongation for 1 min each, followed by one elongation/extension cycle at 72°C for 7 min. Reactions were placed in wells of a 1% agarose gel immersed with 1X Tris-acetate EDTA (TAE) gel running buffer in an electrophoresis apparatus (5 V per cm), stained with ethidium bromide (0.25 μ g/ml final concentrations) and visualized using an Alpha Innotech charge-coupled device (CCD) based imaging system (San Leandro, CA, USA).

(ii) *Semi-quantitative real-time PCR.* The comparative cross point (CP) method, also called the E-method (29), for semi-quantitative PCR analysis was carried out using the Roche's LightCycler (LC™), model 2.0 PCR instrument, utilizing the LC Relative Quantification Software™ (30, 31). The method uses standard curves, in which the standard concentrations are plotted *versus* the threshold cycles, to calculate the unknown samples automatically without user input (Figure 3) (29, 32-34).

Selection of primer, probes, genes and PCR conditions. A web-based assay design software was used for selecting target-specific PCR primers. This employs the Primer3 software (http://frodo.wi.mit.edu/primer3/primer3_code.html) with compatible sterically-modified short probes, and contains 90 pre-validated dual-labeled detection probes that target 98% of all human mRNA transcripts annotated in the RefSeq database at the USA's National Center for Biotechnology Information (NCBI) (35). The probes are short (only 8-9 nucleotide long) and the four nucleotide bases have been substituted with high affinity nucleotide analogs (*e.g.*, locked nucleic acid, LNA, which are conformationally locked in a C3'-endo/N-type sugar conformation that leads to reduced conformation flexibility). These conformational restraints raise melting temperatures to ~80°C under standard hybridization conditions, which assures duplex stability and specificity in real-time RT-qPCR assays, as even a single mismatch with the potential target affects binding and prevents generation of signal (36, 37). The ProbeFinder web-based assay design software is accessible at Exiqon site (<http://www.universalprobelibrary.com>), or via the Roche Applied Science home page (www.roche-applied-science.com).

Table II shows the probe and primer sequences generated for use in this study employing the above websites. Whenever non-intron-spanning assays were performed, DNase-treated RNA

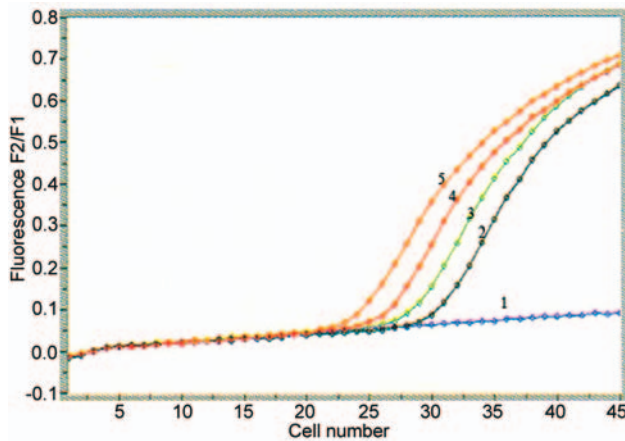


Figure 4. Relationship between fluorescence (F_2/F_1) versus cycle number for stool having HT-29 cells added at 1 cell (curve #1), 10 cells (curve #2), 100 cells (curve #3), 1,000 cells (curve #4), and 10,000 (curve #5) cells, thoroughly mixed and kept for 4 h at 4°C with 1 g of stool devoid of any colonocytes, before total RNA extraction. In this real-time RT-qPCR validation experiment that measures the amplification of human *HPRT* gene, the concentrations of calculated mRNA employed corresponding to the above number of cells were estimated to be ~0.4 pg, 4 pg, 40 pg, 0.4 ng and 4 ng, respectively. These values also correspond to ss-cDNAs of 0.36 pg, 3.6 pg, 36 pg, 0.36 ng and 3.6 ng, respectively.

preparations were used. The selected primers were also validated using the Basic Local Alignment Search Tool (BLAST) from the National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/BLAST/>). It was also important to determine whether folding of the mRNA might interfere with primer access during the RT step using the Mfold program (<http://www.bioinfo.rpi.edu/applications/mfold/old/rna>). Validated HPLC purified primers for this study were obtained from Invitrogen.

Primers were prepared for eleven genes: nine genes showing either increased expression by the serial analysis of gene expression (SAGE) method or microarrays (Insulin-like growth factor II (*IGF2*), Filamin A α (*FLNA*), transforming growth factor β -*igh* induced (*TGF β -igh*), CDC28 protein kinase regulatory subunit 2 (*CKS2*), chromosome segregation 1-like (*CSE1L*), chemokine (C-X-C motif) ligand 3 (*CXCL3*), dipeptidase 1 (*DPEP1*), kallikrein 10 (*KLK10*)); one gene showing decreased expression by SAGE or microarrays (uroguanylin, *GUCY2B*); an inflammatory gene (Interleukin-12, *IL-12*) that showed increased expression in IBD (UC and CD); in addition to a standard housekeeping gene (hypoxanthine phosphoribosyl transferase, *HPRT*). Table III shows the employed genes and some selection parameters. The rationale for gene selection is detailed in the discussion section below.

The PCR Mix for thirty-three, 20- μ l reactions was prepared by adding to a 1.5 ml reaction tube on ice: 10.4 μ l of water; primers at a final concentration of 200 nM in a volume of 0.2 μ l each for the forward and reverse primers; probe concentration of 100 nM in a volume of 0.2 μ l; 4 μ l of ready-to-use Roche LightCycler® FastStartTaqMan® Probe Master and 5 μ l of cDNA template at 7 μ g/ml (50 ng template). The PCR running conditions were: one cycle at 95°C for 10 min to activate the polymerase, 25-40 cycles of denaturation at 95°C for 10 sec, annealing at 60°C for 30 sec

and extension at 72°C for 1 sec. Control samples (negative control) to exclude contamination, in which cDNA was replaced by H₂O, were run in parallel with each experiment.

Detection sensitivity of gene expression. A validation study was carried out to establish the lower limits of detection (sensitivity) of the expression of the *HPRT* gene in HT-29 cells added to stool of a normal subject that was collected four days earlier and stored at 4°C. As colonocytes are precipitously lost after more than one day of fecal storage (13), any nucleic acid found in the employed stool will be due to bacteria and that extracted from spiked HT-29 cells. In that study, HT-29 cells were spiked and thoroughly mixed in a mortar using a pestle with 1 g of human stool at 10⁰, 10¹, 10², 10³ and 10⁴ cells, and because on average 2% of the total RNA is mRNA, the number of cells present in 1 g of stool corresponded to ~0.4 pg, 4 pg, 40 pg, 0.4 ng and 4 ng of mRNA, respectively, considering that each cell contains ~20 pg total RNA (27). The cells were allowed to remain in the harsh stool environment (containing PCR inhibitors such as heme derivatives, bile acids, and complex polysaccharides) for 4 h at 4°C, before extraction of RNA, thereby mimicking the real stool collection conditions. DNase (Invitrogen) treatment of the stool was then carried out to guard against any amplification of genomic DNA contaminant in stool, followed by RT and real-time qPCR detection of the *HPRT* gene. Various concentrations of an initial RNA, instead of a cDNA, were employed to guard against errors due to different cDNA synthetic reactions (24). Each point was run in triplicate.

Amplification specificity of studied genes. The amplification specificity on all eleven genes studied was evaluated by running 1% agarose gels on products of endpoint PCR in parallel with real time PCR to: (i) confirm and determine the analytical specificity of the RT-PCR reaction, and (ii) verify the ability of the Universal probes – specific for studied genes – to bind the PCR product. We performed a conventional 25 μ l qualitative endpoint PCR reaction, running 10 μ l of the reaction product on an agarose gel, followed by transfer the DNA into a Biotran™ Nylon membrane (ICN, Irvine, CA, USA) using a downward capillary transfer. After crosslinking the DNA to membranes by UV at 100 mJ/cm², a short hybridization probe – specific for the internal sequence of the PCR product end-labeled with digoxigenin – using terminal deoxynucleotidyl transeferase (Promega Corporation, Madison, WI, USA) was prepared and hybridized. The signal was detected by chemiluminescence using alkaline phosphatase-conjugated anti-digoxigenin antibody and CDP-star substrate (Roche Diagnostics). Digital capture of light emission was carried out using Alpha Innotech chemiluminescent imaging instrument (San Leandro, CA, USA).

Quality control (QC) considerations and optimization for the PCR. QC procedures were employed to ensure the uniformity, reproducibility and reliability of the PCR reaction. Random variability was minimized by running triplicate samples and averaging the data. Variability due to operator error was minimized by using a cocktail of reagents (*i.e.*, master mix). To compensate for various variabilities, we implemented QA methods that minimized intra- and inter-assay and intra-subject variations, intralaboratory differences, statistical variations; analytical sensitivity for minimum detection level and normal control population variation for both stool and tissue samples (38). To prevent suspect carry-over, we used a method for preamplification inactivation of amplified DNA

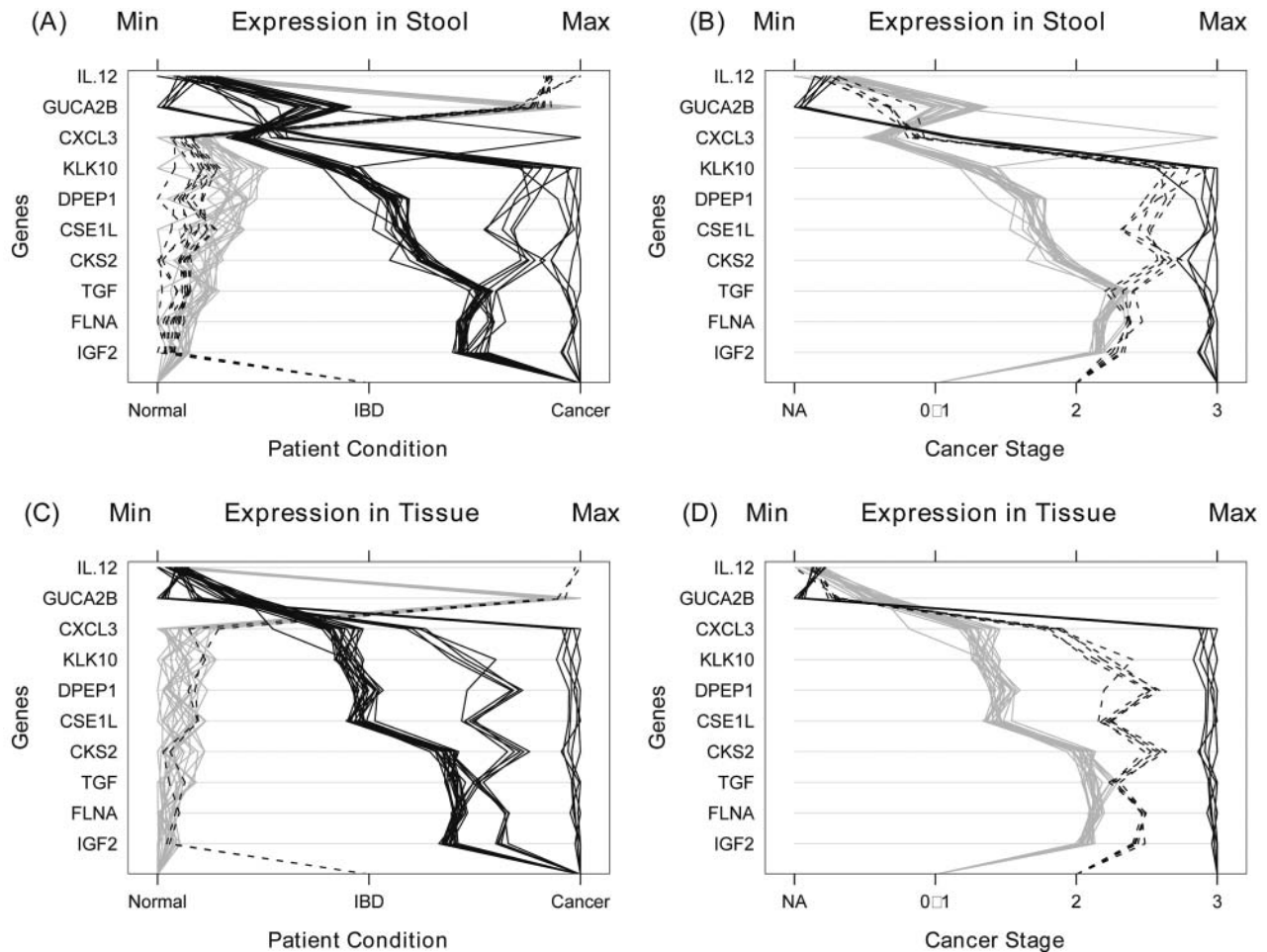


Figure 5. A) Gene expression for stool samples taken from 60 patients. The condition of the patient is indicated by the bottom row of the panel and by the type of line. There were 20 normal patients (gray lines), 10 with IBD (dashed lines), and 30 with cancer (black lines). Instances of high expression appear on the right and those with low expression on the left. Expression was measured by CP and scales were chosen so that minimum values line up on the Min mark labeled at the top of the panel. The same is true for the maximum values which line up under the mark labeled Max. B) This panel displays gene expression for stool samples taken from 30 cancer patients. Stage of cancer is indicated by the bottom row of the panel and by the type of line. There were 20 patients with stage 0 or 1 (gray lines), 5 with stage 2 (dashed lines), and 5 with stage 3 (black lines) cancer. The 30 noncancerous patients (stage NA) are not shown. C) Gene expression for tissue samples taken from 60 patients. Conditions of the patient are the same as in panel A. D) This panel displays gene expression for tissue samples taken from 30 cancer patients. Stages of cancer are indicated as in panel B.

that allowed for the selective destruction of previously amplified DNA using Roche LightCycler® Uracil-DNA-glycosylase (UDG) (39). The UDG inhibitor protein Ugi was used to inactivate residual UDG activity and protect dU-containing products during benchtop manipulations (40).

Results

Results from the validation study using *HPRT* in HT-29 cells added to stool are shown in Figure 4. They demonstrate that expression of the *HPRT* gene was detected in as few as 10 adenocarcinoma cells (or roughly 0.4 pg of mRNA) per gram of stool or 0.001% detection. The

genome of the bacteria present in the GI does not contain the *HPRT* gene; therefore, any amplification of that gene is because of the presence of spiked HT-29 cells, especially since stool was treated with DNase that removed any potential genomic DNA contamination. When only one adenocarcinoma cell was used, the curve approximated to that of an untreated control, indicating that the limit of our detection sensitivity was 10 adenocarcinoma cells thoroughly mixed with one gram of human stool.

After inspection of generated qPCR data (Table IV and Figure 5), several genes by themselves appear to offer distinct and clear separation between control and cancer cases in either stool or tissue samples. A gene expression

Table IV. Comparative crossing points of gene expression in stool and tissue from normal, cancer, and IBD patients.

Group	HPRT	IGF2	FLNA	TGFβ _{high}	CKS2	CSE1L	DPEP1	KLK10	CXCL3	GUCA2B	IL12
negcont ^a	35.02	35.90	35.74	35.84	35.20	35.64	35.06	35.14	35.56	35.14	35.10
N _S 1	36.00	36.12	35.66	35.86	35.66	35.46	36.10	35.66	35.40	35.08	36.12
2	35.86	35.24	34.98	36.08	35.16	35.78	35.42	36.04	35.38	36.02	35.90
3	35.88	35.94	36.02	35.52	35.80	36.20	35.10	35.22	35.48	35.52	35.18
4	35.52	35.42	35.92	35.16	36.02	35.80	35.42	35.68	36.72	36.04	35.66
5	35.76	35.38	35.28	35.26	35.34	35.16	36.28	35.54	35.08	35.52	36.02
6	36.80	35.76	35.10	35.24	35.16	36.02	35.42	35.24	35.42	35.08	35.86
7	35.88	35.40	35.68	36.34	35.90	35.40	35.14	35.86	36.34	35.90	36.02
8	35.48	35.18	35.27	35.08	36.26	36.78	35.40	35.98	35.16	35.64	35.88
9	35.58	35.82	36.38	36.14	35.06	35.80	35.44	35.24	35.33	35.16	36.42
10	35.04	35.62	35.42	35.46	35.26	35.27	35.14	36.68	35.24	35.54	35.58
11	36.07	36.32	35.22	35.26	35.70	35.65	35.02	35.30	35.66	35.68	35.14
12	35.26	35.58	36.44	36.68	35.14	35.28	35.22	35.86	35.40	35.62	36.26
13	34.96	35.66	35.22	35.76	35.34	35.76	35.18	35.92	35.16	36.07	35.66
14	35.30	35.32	35.92	35.40	36.04	35.24	35.38	35.16	35.82	35.76	35.10
15	35.38	35.80	36.08	35.32	35.42	35.86	35.46	35.62	35.18	35.50	36.04
16	36.78	35.16	35.02	35.36	35.78	36.38	35.12	35.36	35.30	35.86	35.22
17	36.14	35.28	35.36	35.78	35.38	35.50	35.18	36.22	36.36	36.38	36.46
18	35.38	35.14	35.28	35.42	35.84	35.16	35.72	35.44	35.28	35.76	35.36
19	35.24	35.12	35.56	35.36	35.46	36.22	35.38	35.52	35.68	35.64	35.24
20	35.84	36.12	36.44	35.16	35.20	35.30	35.42	35.66	35.69	35.18	35.56
N _T 1	34.80	35.02	35.08	35.16	35.08	34.82	35.14	34.26	35.18	34.90	34.46
2	34.56	34.54	34.96	35.01	34.98	34.66	35.17	34.77	35.14	35.26	34.58
3	34.18	34.76	35.22	34.09	35.06	35.16	34.96	35.32	34.88	35.16	35.26
4	35.06	35.16	35.22	34.52	34.82	34.14	35.04	35.20	34.60	35.30	34.86
5	34.92	34.88	34.96	34.16	35.02	34.46	34.66	34.80	35.10	34.36	35.12
6	34.88	34.76	34.80	35.08	34.70	35.06	34.62	34.24	34.84	34.10	34.64
7	34.90	35.16	34.70	34.64	34.18	34.50	35.16	34.48	34.54	35.00	34.76
8	34.76	34.80	34.96	35.06	34.84	34.92	34.88	34.78	35.08	34.46	34.32
9	34.38	34.54	34.60	34.68	35.06	34.46	34.38	34.90	35.12	35.04	34.68
10	35.04	34.16	35.30	34.16	34.32	35.18	35.26	35.22	34.68	34.36	35.06
11	34.16	35.22	35.28	35.22	34.40	35.16	34.70	34.76	35.14	34.66	34.70
12	34.52	34.18	34.64	34.26	34.18	34.68	34.88	35.08	34.92	35.16	34.86
13	34.56	35.16	35.06	35.12	35.16	34.78	34.54	34.82	34.68	34.56	35.14
14	35.14	35.26	35.10	34.86	34.90	34.44	35.02	35.06	35.08	35.12	34.80
15	34.78	34.84	34.66	34.54	34.48	34.56	34.66	34.98	34.88	35.00	34.60
16	34.58	34.68	34.46	34.48	34.86	34.36	35.08	34.86	34.92	34.98	34.72
17	34.54	34.34	34.80	35.06	34.64	35.14	34.64	35.24	34.66	34.76	34.56
18	35.06	34.48	35.14	35.02	34.88	34.16	35.08	34.86	34.48	34.50	34.38
19	34.24	34.38	35.20	34.78	34.60	35.28	35.16	34.68	35.04	35.00	34.90
20	34.64	34.66	34.74	34.56	34.86	34.42	34.54	34.36	34.92	34.66	34.70
S ₀₋₁ 1	35.90	23.18	25.06	28.08	31.06	32.04	33.02	34.00	34.22	40.04	35.87
2	35.76	23.44	25.26	28.18	30.98	32.14	32.88	34.08	34.12	39.92	34.98
3	35.40	22.92	24.88	28.06	31.00	32.16	33.06	34.32	34.18	40.10	35.16
4	35.56	23.34	24.92	28.36	31.16	32.32	33.16	34.28	34.44	40.16	35.36
5	35.84	23.56	25.14	28.56	31.37	32.42	32.96	34.20	34.28	41.00	35.68
6	36.02	22.88	24.94	28.10	31.18	32.28	33.08	34.08	34.18	40.12	35.66
7	35.78	23.12	25.14	28.00	31.04	32.76	33.28	34.16	34.34	39.94	36.76
8	35.48	22.88	25.00	28.24	31.18	32.52	33.16	34.06	34.36	40.56	35.46
9	35.64	23.10	24.88	27.96	31.14	32.16	33.10	34.28	34.04	40.32	35.44
10	35.56	23.16	25.24	28.26	31.22	32.08	32.90	34.36	33.88	41.04	35.72
11	35.72	23.34	24.78	28.14	31.06	32.14	33.08	34.14	34.02	40.78	36.18
12	35.42	22.96	23.96	28.04	30.90	32.26	33.18	33.90	33.76	41.06	35.64
13	35.30	23.08	24.10	28.16	31.06	32.34	33.20	34.10	24.22	40.82	35.68
14	35.74	23.14	25.06	28.10	31.08	32.16	33.08	34.06	34.20	41.04	36.26
15	35.38	23.26	25.14	28.02	31.06	32.26	33.26	34.18	34.52	40.66	35.34
16	36.06	23.24	24.88	28.12	31.76	32.08	33.12	33.87	34.26	41.08	35.60
17	35.88	23.22	25.06	28.34	30.84	32.18	33.04	34.12	34.24	40.56	35.88
18	35.96	23.28	25.18	28.32	31.16	32.10	33.14	34.26	34.38	40.68	35.26

continued

Table IV. *continued*

Group	HPRT	IGF2	FLNA	TGF β high	CKS2	CSE1L	DPEP1	KLK10	CXCL3	GUCA2B	IL12
19	35.44	23.18	24.87	28.36	31.04	32.22	33.68	34.40	34.52	40.88	35.68
20	36.34	23.08	24.78	28.28	31.08	32.12	32.86	34.04	34.68	40.26	35.46
S ₂ 21	36.10	22.02	23.92	28.60	29.16	30.12	31.08	31.24	33.04	42.34	35.43
22	36.08	22.56	23.88	28.14	29.00	30.68	31.28	30.98	33.36	42.00	36.12
23	35.86	22.36	24.02	27.88	28.88	30.34	31.36	31.40	33.14	41.96	35.68
24	35.48	22.78	23.44	27.86	28.68	30.64	31.46	31.52	32.88	41.46	35.48
25	35.90	22.16	23.86	28.44	29.04	30.18	31.16	31.66	32.98	42.02	36.14
S ₃ 26	35.76	18.72	20.86	26.04	28.78	29.00	30.66	31.02	32.04	44.00	35.88
27	35.62	17.94	20.64	25.86	28.56	29.08	30.54	30.90	31.78	43.86	36.14
28	36.28	18.12	21.00	25.68	27.98	28.96	30.44	30.86	31.88	43.90	35.66
29	35.46	18.80	20.76	26.02	28.14	28.88	30.50	31.06	32.12	43.80	35.42
30	35.88	18.50	20.76	26.12	28.56	29.10	30.76	31.70	32.06	43.78	35.78
T ₀₋₁ 1	35.84	21.16	23.06	26.82	29.16	30.80	31.71	31.84	33.18	44.56	35.46
2	34.78	21.34	22.86	26.66	29.22	30.74	31.56	31.90	33.10	44.16	34.50
3	34.44	20.87	23.12	25.98	29.18	30.60	31.78	31.88	33.06	44.32	34.62
4	35.04	21.44	22.86	26.60	28.98	30.44	31.62	31.76	33.04	44.44	35.06
5	35.12	21.36	23.08	26.78	29.36	30.52	31.54	31.68	32.98	43.94	34.90
6	35.32	21.48	22.66	26.36	29.04	30.40	31.76	31.52	32.88	43.88	34.82
7	34.16	20.78	23.08	26.46	29.48	30.64	31.44	31.80	33.78	43.78	34.78
8	34.66	21.14	23.00	26.44	29.16	30.56	31.34	31.66	33.14	43.92	34.66
9	35.36	21.22	22.78	26.52	29.32	30.84	31.28	31.72	33.12	44.20	35.20
10	34.80	20.98	22.94	25.88	29.08	30.76	31.58	31.56	32.96	43.96	34.82
11	34.52	20.94	23.24	25.98	29.08	30.82	31.42	32.02	32.90	44.24	34.60
12	34.60	21.14	22.58	26.36	29.14	30.50	31.74	31.96	32.94	44.18	34.64
13	35.18	21.56	22.68	26.56	29.34	30.44	31.82	31.92	32.78	43.76	34.80
14	35.68	21.14	22.90	26.56	28.96	30.74	31.58	31.86	32.80	43.68	34.44
15	34.38	20.88	23.10	26.74	28.88	30.16	31.38	32.06	32.70	43.46	34.70
16	34.46	21.46	22.88	25.78	29.08	30.36	31.38	31.80	33.06	44.02	34.54
17	35.38	21.48	23.12	26.22	29.34	30.72	31.63	32.04	33.08	43.82	35.16
18	35.68	21.70	22.78	26.38	29.22	30.54	31.42	31.78	32.68	43.56	35.08
19	34.66	21.22	22.86	26.66	28.92	30.78	31.58	31.68	32.74	44.20	34.46
20	34.90	21.56	23.16	26.78	29.04	30.60	31.80	32.00	33.04	43.90	35.02
T ₂ 21	35.16	18.96	21.00	25.92	27.86	28.04	29.02	29.90	32.08	45.02	35.52
22	34.90	18.88	20.86	26.02	27.60	28.12	29.80	29.08	32.00	45.22	34.90
23	34.88	18.36	21.06	25.82	27.72	27.80	28.98	29.76	32.16	45.34	34.72
24	35.16	18.70	20.90	25.70	27.52	27.94	28.88	29.48	31.98	45.12	34.84
25	35.36	18.82	20.88	25.86	27.40	27.86	28.82	29.64	31.92	45.14	34.66
T ₃ 26	35.26	15.02	18.16	23.10	26.70	25.60	28.00	27.96	30.20	46.10	34.80
27	34.78	14.88	18.36	22.90	26.74	25.32	27.92	27.52	30.12	46.18	34.42
28	34.66	15.14	17.94	23.06	26.52	25.40	27.86	27.84	30.16	46.32	34.36
29	35.08	14.90	18.22	23.12	26.34	25.36	27.80	27.68	30.04	46.02	34.86
30	35.36	14.88	18.04	22.86	26.60	25.56	27.96	27.70	30.26	46.22	34.54
UC _S 1	36.06	35.80	35.60	35.88	35.94	36.02	35.80	35.96	35.72	35.84	25.10
2	36.12	35.86	35.44	36.04	35.90	35.84	36.02	36.06	35.66	35.86	24.88
3	35.82	35.92	35.66	35.96	35.84	35.72	35.94	35.82	35.80	35.70	24.96
4	35.96	35.78	35.76	35.82	35.76	35.86	35.78	35.86	35.92	36.12	24.78
5	36.22	35.98	35.84	35.90	35.82	35.68	36.08	35.90	36.04	36.00	25.02
UC _T 1	34.74	34.82	34.56	34.42	34.90	34.32	34.60	34.52	34.44	34.54	20.02
CD _S 1	36.16	36.16	36.44	36.56	36.32	36.52	36.18	35.98	36.04	36.36	24.16
2	35.80	35.96	36.12	35.88	35.92	36.14	36.12	36.04	36.12	35.90	24.86
3	35.98	36.16	36.22	36.16	36.36	36.02	35.78	36.10	35.92	36.06	24.92
4	35.88	36.24	35.92	35.84	35.70	36.26	36.06	36.18	36.32	36.14	24.96
5	36.04	36.54	36.08	36.18	36.26	36.22	36.42	36.44	36.22	36.50	24.00
CD _T 1	34.52	34.68	34.40	34.88	35.04	34.34	34.72	34.44	34.80	34.76	19.82

^aNo DNA added to reaction (negative control). This table is obtained from 60 patients: 20 non-cancerous controls, 30 patients with colon cancer, 5 patients with severe ulcerative colitis (UC), and 5 patients with Crohn's disease (CD). N_S1 (N_T1) to N_S20 (N_T20) = Stool (Tissue) from normal individuals 1 to 20; S₀₋₁1 (T₀₋₁1) to S₀₋₁20 (T₀₋₁20) = Stool (LCM Tissue) from patients 1 to 20 with adenomatous colon polyp 1 cm (stage 0-1); S₂21 (T₂21) to S₂25 (T₂25) = Stool (LCM Tissue) from colon carcinoma patients 21 to 25 (stage 2); S₃26 (T₃26) to S₃30 (T₃30) = Stool (LCM Tissue) from colon carcinoma patients 26 and 30 (stage 3); UC_S1 (UC_T1) = Stool (LCM Tissue) from first patient with severe UC; UC_S2 to UC_S5 = Stool from patients 2 to 5 with severe UC; CD_S1 (CD_T1) = Stool (LCM Tissue) from first patient with CD; CDS2 to CDS5 = Stool from patients 2 to 5 with CD.

index is therefore not necessary for reaching conclusions from these data. However, a larger sample may include patients that respond differently for some genes in which case there may be a meaningful way to define a gene expression index. The level of gene expression is displayed using parallel coordinate plots (41) produced by the lattice package in R (version 2.4.0, <http://cran.r-project.org>).

For a gene like *IGF2*, a 4 ng total non-degraded RNA extracted from a fresh colon stool of adenocarcinoma patients gives an average CP value of about 20 (see Table IV), which translates to a median total RNA value of about 7×10^9 copy numbers/ μ g RNA. RT reactions employing oligo-dT₁₂₋₁₈ and using Superscript III Reverse Transcriptase (Invitrogen), which result in about 90% conversion efficiency, give ss-cDNA values of 0.36 pg, 3.6 pg, 36 pg, 0.36 ng and 3.6 ng, which correspond to the above mRNA values of ~0.4 pg, 4 pg, 40 pg, 0.4 ng and 4 ng of mRNA, respectively.

Figure 5 panel (A) displays gene expression for stool samples taken from 60 patients. Condition of the patient appears at the bottom of the panel and is indicated by the type of line: 20 normal patients (gray lines), 10 with IBD (dashed lines), and 30 with cancer (black lines). Instances of high expression appear on the right and those with low expression on the left. Expression was measured by CP and scales were chosen so that minimum values line up on the Min mark labeled at the top of the panel. The same is true for the maximum values which line up under the mark labeled Max. For *IGF2*, IBD samples are indistinguishable from normal samples. While there is some variability in *IGF2* expression among these patients, it is small compared to the distance each of the cancer patients are from the normal samples. *IGF2* distinguishes between two groups of cancer patients: about 5 near the maximum and 25 with lower expression but still much higher expression than normal and IBD patients. Genes *FLNA* through *KLK10* show a similar relationship among the three groups although the variability among normal and IBD patients increases as we follow this list of genes up the panel. Also, there is a group of 20 cancer patients whose expressions become lower going from *TGF β -igh* through *KLK10* but are still distinct from normal controls. The separation shown by *CXCL3* would be better were it not for one extremely highly expressed cancer sample. It is significant that this sample was not an outlier for other gene expressions indicating that a combination of genes could be helpful in identifying cancer even when one gene provides an anomalous value. *GUCA2B* shows a reversal with normal and IBD samples showing high expression and cancer showing low expression. Note that expression within cancer is also reversed, with the group that had the highest expression for the previous genes, now has the lowest expression. *IL-2* shows that IBD samples are highly expressed and distinct from cancer and normal individuals that have low expression.

Figure 5 panel (B) displays gene expression for stool samples taken from 30 cancer patients. Stage of cancer is indicated by the bottom row of the panel and by the type of line: 20 patients with stage 0 or 1 (gray lines), 5 with stage 2 (dashed lines), and 5 with stage 3 (black lines). The 30 non-cancerous patients are not shown. Stage 3 is well separated from the other stages for *IGF2*, *FLNA* and *TGF-igh*. Stage 0-1 is well separated from the other stages for *CKS2*, and *DPEP1*. These stages form three distinct groups for *CSE1L*. These genes indicate the potential for differentiating among cancer stages rather than just the presence of cancer.

Panels (C) and (D) display the same information as panels (A) and (B), respectively, except that now tissue samples are used instead of stool samples. Overall, the relationships among patient groups, panels (A) and (C), and among cancer stages, panels (B) and (D), are very similar using tissue samples. Tissue samples differ from stool samples in that the variability within patient groups and within cancer stages are smaller for tissue samples. Also, there is no outlier in the *CXCL3* gene for the tissue sample.

Discussion

Rationale for using RNA markers for colon cancer screening. The expression of individual genes may be altered either by mutations in the DNA, or by a change in their regulation at the RNA or protein levels (24, 42). Epigenetic silencing is also recognized as an important mechanism contributing to gene inactivation in CRC (43).

Epigenetic alterations have been studied in stool from CRC patients. MethylLight analysis of fecal DNA from three independent sets of patients recognized *SFRP2* methylation as a sensitive single DNA-based marker for identification of CRC in stool samples with 90% sensitivity and 77% specificity in the training set and 77% sensitivity and 77% specificity in an independent test set (44). Analysis of promoter methylation of hypermethylated in cancer 1 (*HIC1*) gene in stool showed it to be highly specific (98%) for colon adenoma and carcinoma, but sensitivity was quite low (31% for adenoma and 42% for all cancer) (45), which suggested that a combination of genetic and epigenetic markers may be needed to identify CRC at an early stage.

Working with DNA is relatively easier than working with the easily degradable messenger RNA. However, mutation detection in oncogenes and suppressor genes suffers from: (i) the detection of mutations in these genes in fewer than half of large adenomas and carcinomas, (ii) the detection of gene mutations in non-neoplastic tissues, (iii) mutations being found only in a portion of the tumor, and (iv) mutations often produce changes in the expression of many other genes. A major disadvantage of a DNA-based test in stool is that it does not detect alterations in gene expression.

Since activation of protooncogenes can occur by various mechanisms, including overexpression of normal mRNA protein products or expression of genes altered through point mutation, truncation or translocation, an mRNA-based methodology has a broader application and better predictive value in monitoring the detection of CRC (46).

Protein-based methods are not suited for screening and early diagnosis because proteins are not specific to one tumor or tissue type (*e.g.*, *CEA*), they are susceptible to proteases, there is a current lack of means to amplify proteins, and most importantly because detection of these markers in blood often signifies the presence of an advanced tumor stage (20, 47). Furthermore, the human proteome is complex, ranging from 100,000 to several million different protein molecules, and more complications arise because no function is known for more than 75% of predicted proteins of multicellular organisms. Additionally, the dynamic range of protein expression is as large as 10^7 . Moreover, mRNA levels do not necessarily correlate with protein expressions due to variable translation efficiencies in protein half-lives (48). Tissue microarray studies revealed that protein expression vastly exceeds RNA levels and only phosphorylated proteins are involved in signal transduction pathways leading to tumorigenesis (47). Transcript analysis offers many technical advantages over protein analysis because mRNA molecules possess high affinity and specificity binding patterns. Furthermore, mRNA molecules exhibit equivalent biochemical properties and can be amplified by PCR. In contrast, there is no PCR equivalent for proteins. In addition, proteins do not possess straightforward binding patterns and exhibit diverse biochemical features, and there is not always a direct correlation between protein abundance and activity (47).

Differences in gene expression have been detected immunocytochemically at the protein level in cells isolated from stool samples of asymptomatic and CRC patients (42). Fluorescence multiplexing technique of mixed cell populations captured on an antibody microarray – used for analysis of plasma membrane proteome of CRC cells – enabled expression profiling and immunophenotyping of multiple subpopulations of carcinoma cells within the tumor sample (49). There is currently no well-documented molecular protein test that has shown in clinical trials to be a sensitive and specific indicator of colon neoplasia, especially at the early stages. Proteomic research is still a relatively new discipline, so it will take considerable time to identify which proteins will be suitable clinical markers and resolve issues of bias and validations (50). On the other hand, we know that much of the phenotypic variation between organisms is related to changes in gene expression and not to alterations in protein sequences (1, 24, 46); hence, for screening, an RNA marker is preferable to either a DNA- or a protein-based marker (47-51).

Advantages of stool testing for colon cancer screens. Stool testing has several important advantages over other colon cancer screening methods as it is uniquely noninvasive and requires no unpleasant cathartic preparation, formal health care visits, or time away from routine activities. Unlike sigmoidoscopy, it reflects the full length of the colorectum. It is believed that colonocytes are released continuously into the fecal stream, unlike blood, which is lost intermittently, as observed in FOBT. Moreover, colonocytes from the colon cancer are 4- to 5-fold greater in number than from normal colonic mucosa (13); therefore, these colonocytes are more abundant than their normal counterpart and, unlike blood, are present in all stool samples from colon cancer patients, therefore partially obviating the use of an enrichment technique to separate tumorigenic from normal colonocytes. Add to that that tumorigenic cells contain more RNA than normal ones, and it becomes obvious that stool from colon cancer patients contains RNA that can be adequately measured by qPCR in spite of the dilution by normal colonocytes. Furthermore, because testing can be performed on mail-in-specimens, geographic access to stool screening is essentially unimpeded. However, CRC stool screening has historically relied on the nonsensitive FOBT, and many physicians and their patients have been reluctant to fully embrace stool testing (52). Nevertheless, application of improved molecular markers could potentially enhance the effectiveness and efficiency of the screening outcome. As a result of these considerations, the American Cancer Society has recognized stool-based molecular testing as a promising basis for future screening methods (53).

Stabilizing RNA and normalization of PCR. Our experience working with RNA has dictated that to prevent its degradation, stool or dissected tissue samples should not be allowed to stay unprocessed after removal from their normal environment for more than 20 min. Freezing and thawing of stool samples for subsequent RNA extraction resulted in rupture of the colonocytes due to ice crystal formation, leading to degradation of RNA and failed PCR amplification because of exposure of the RNA to the hostile stool environment. Hence, all our RNA stool extractions used either freshly collected stool, or stool that had been fixed in a preservative such as RNeasy[®]-Lysis Buffer to prevent destruction of colonocytes when frozen at -70°C . Extracted total RNA from both stool and tissue was stably stored at -70°C until further processing. Commercial kits were employed to extract stable total RNA from all samples: cells in culture, tissue and exfoliated colonocytes in stool (19, 21).

Quantitative real-time RT-PCR-specific errors in the quantification of mRNA transcripts could be easily compounded by variations in the amount of starting material between the samples that may be caused by sample-to-sample fluctuation, variation in RNA integrity,

RT efficiency differences, cDNA sample loading variations and cellular input. This is especially relevant when samples are obtained from different individuals or sources (tissue *versus* stool), or different time courses (run-to-run variation), which can result in the misinterpretation of the derived expression value. Therefore, normalization of the target gene expression level must be performed to compensate for intra- or inter-kinetic RT-PCR variations by measuring simultaneously the levels of a single universal cellular RNA standard that is present at constant levels in the cells, and its steady-state expression level should preferably be similar to a target gene (30, 37).

In this study, data normalization was carried out using an endogenous housekeeping gene (*HPRT*) with a low copy number. We previously employed the major histocompatibility complex β -2-Microglobulin ($\beta 2M$) gene (GenBank accession # NM_004048.2), a pseudogene-free housekeeping standard with an intermediate copy number (19-21). Although $\beta 2M$ appeared to be a suitable internal control in quantitative serum-simulation studies (54), it was not a suitable normalization standard in several subsequent studies (55-57), and was even found to be one of the worst in a well controlled study (58). Attempts were made to employ multiple housekeeping standards (56, 57). In a recent study that employed 13 genes, *HPRT* was found to most accurately represent the mean expression patterns of all other genes (58). Moreover, a low-copy housekeeping gene such as *HPRT* controls best for RNA isolation efficiency, RNA quality and RT-efficiency (58). Although the "perfect" standard gene does not exist, the choice of the reference gene utilized to normalize the expression of the target gene of interest is critical for the interpretation of results. Therefore, we decided to employ *HPRT* alone in the present study because the efficiencies of amplification of the control and experimental genes must be equal, as judged by similar slopes (24, 32); in this study, the *HPRT* gene met this criterion. Additionally, the use of one reference standard facilitates comparison.

Another recent study of the effect of RNA quality on reference housekeeping gene stability found major differences in reference gene expression stability between intact and degraded RNA samples from the same tissue (59). For example, while *HPRT1* was ranked as relatively stable in intact RNA nasal polyposis and chronic rhinosinusitis samples, but was quite unstable in degraded samples. Therefore, in our study, we measured RNA integrity on every sample tested (Figure 2) and discarded/replaced degraded ones, since we did not wish to miss the small changes in expression encountered with some of the selected genes.

Percentage of neoplastic cells as a fraction of total human cells in stool. Estimation of this percentage would depend on tumor stage and efficiency of the RT and/or PCR reaction. Traverso

et al. (15) showed that DNA could be detected in all stool samples, but the percentage of samples with APC mutations – presumably derived from cancer cells – varied. However, no mutation was detected in 43% of stool samples, probably due to lack of sensitivity. Work by Vogelstein and Kinzler (60) using digital PCR showed that 4% of stool samples of CRC patients had *K-ras* mutations; the work of Dong *et al.* (61) using three genetic markers (*TP53*, *BAT26* and *K-ras*) in stool DNA detected 71% of patients with CRC; and the work by Imperiale and colleagues (17) showed that adenomas were detected in only 18% of patients' stools. So averaging the four studies results in around 37% mutation detection in stool from patients with CRC at various Dukes' stages. Since our validation study in stool spiked with HT-29 adenocarcinoma cells has a sensitivity of 0.001% detection for mRNA expression, we hypothesize that we should be able to detect a percentage much smaller than that achieved by mutation detection, considering that gene expression is more prevalent and easier to measure than mutation by high throughput cyclers, and we will be measuring the expression of more than one gene, which provides increased sensitivity. Indeed, our experimental results presented in Table IV and Figure 5 for both stool and tissue samples from various stages of colon adenocarcinoma validate our hypotheses and assumptions.

Advantages of the E-method for qPCR analysis. In the CP or E-method, relative target concentration is a function of the difference between crossing points (or cycle numbers) as calculated by the Second Derivative Maximum, in which Cyclex's software algorithm (v4.0 software for the LC™ model 2.0) identifies the first turning point of the fluorescent curve in the graph showing fluorescence *versus* cycle number (31). This turning point corresponds to the second derivative curve (29, 31, 32).

The CP or E-method, produces more accurate results than the $2^{-\Delta\Delta C_T}$ method (33) because it can compensate for differences in target and reference gene amplification efficiency [$E = 10^{-1/\text{slope}}$] (34) either within an experiment, or between experiments as it analyzes the amplification efficiency of target and reference gene using relative standards of serial dilutions of a single normal sample (*e.g.*, undiluted, 1:10, 1:100, *etc.*) whose amplification efficiencies are very similar to those of the unknown samples, thereby avoiding the time-consuming preparation of artificial or cloned standards and the determination of their absolute values (29). Furthermore, the E-method normalizes for run-to-run differences, such as those caused by variations in reagent chemistry using one sample of the relative standards designated a "calibrator" for the target and for the reference genes. These calibrators can then be used repeatedly in subsequent runs during the entire analysis thereby guaranteeing a common reference point for comparison of all experiments within the series (32). Figure 3 illustrates graphically the relative quantification of the same run by the E- compared to the $2^{-\Delta\Delta C_T}$ -method.

Rationale for the selection of genes for the study. Looking at the ONCOMINE™ database at <http://oncomine.org> we could not find a gene with expression exclusive to colon cancer; the closest were a few genes that are expressed in colon and other closely-related cancers (*e.g.*, gastroesophageal). We also looked at SAGE data (62-65), oligonucleotide- and cDNA-array databases (66-80), combination of cDNA microarray and RNA interference (81), and RT-qPCR (82). Additionally, supporting reports in the peer-reviewed literature showed that some of these genes exhibited aberrant gene expression in pre-neoplastic stages (83, 84) and other stages of colon cancer (85-87). We carried out a meta-analysis to discriminate colon cancer genes from normal genes in the ONCOMINE™ database, but again the ones we found were not exclusive to colon cancer. One gene showed increased protein expression only in serum of colon cancer patients (*CCSP2*) (88), but no increased RNA expression was reported in colon cancer tissue.

Gene selection for this study was facilitated using the enormous resources of the Cancer Genome Anatomy Project (CGAP). SAGEmap (<http://www.ncbi.nlm.nih.gov/SAGE>) is a public database created as a component of the CGAP to provide a central location for depositing, retrieving and analyzing human gene expression data through the use of custom tools. A Digital Gene Expression Displayer (DGED) is a tool that compares gene expression between two pools of libraries by comparing the "degree" of the presence of a gene in pool A with its "degree" of presence in pool B from a "Virtual Northern" function that calculates the fractional representation of a particular gene in all of the posted libraries through the use of: (i) the sequence odd ratio, OR, [Seqs A/Total Sequences in A ÷ Seqs B/Total Sequences in B] (<http://cgap.nci.nih.gov/Tissues/Significance>), and (ii) a measure of significance (*p* value). The lower the OR and *p*, the higher the probability that the tag occurs more frequently in one pool as compared to the other (89). Using these tools we were able select the nine genes for this study as shown in Tables II and III.

The Gene Library Summarizer (GLS) tool explains how the tool works (<http://cgap.nci.nih.gov/Tissues/GLSHoTo>). The limitation on the GLS tool is that it analyzes gene expression solely on the basis of sequences drawn from cDNA libraries (*i.e.*, at least one sequence from the gene appearing in colon cancer libraries, and no sequences appearing in any other library). We used two expression analysis tools on the CGAP site: the cDNA Digital DGED (<http://cgap.nci.nih.gov/Tissues/GXS>) and the SAGE DGED (<http://cgap.nci.nih.gov/SAGE/SDGED?METHOD=SS10,L S10&ORG=Hs>) (89). These tools allowed us to create arbitrary pools of libraries and then find genes that are overexpressed in one pool in contrast to the other pool (*i.e.*, measure relative gene expression). Because of great differences in the underlying specimens and in the method

of mapping tags/sequences to genes, the two tools often give different results.

For each of these tools we ran three "analyses": (i) colon cancer *versus* all other libraries, (ii) colon cancer vs. normal colon, and (iii) colon cancer vs. other cancers. In case of the cDNA DGED, we kept the default limitation of including libraries that have a minimum of 1000 sequences. For SAGE DGED, we used the 14 bp SAGE tags in SAGE colon libraries. For both the cDNA DGED and the SAGE DGED, we kept the default parameters for fold change at 2x (*i.e.*, looking for genes that are expressed twice as much in colon cancer as in the contrasting pool) and a level of significance *p*-value of 0.05. We identified 354 genes in cDNA DGED and 954 in SAGE DGED. Only 41 genes appeared in both database lists. We then used another tool to provide Monochromatic Virtual Northern analysis based on the cDNA data and SAGE data in the entire CGAP website. From these data we calculated the OR on these genes (89). If the gene is not found at all in normal (0 value), calculation will give "infinity", which is better for example than an OR of 20. The advantage of using the above tools is that the SAGE technology gives absolute quantification of the tags, and the gene expression results derived from that technique allows for a semi-quantitative measure of the degree of gene expression (90, 91).

We selected 9 genes for our present study that represent a broad choice of genes showing aberrant expression using the above tool and that also have a biological relevance to cancer, based on an extensive search of the literature. Like mutations, expression of gene(s) is not only confined to a particular organ or tissue, but occurs in all cancer. The degree of gene expression, not just which genes are specific to a particular cancer or restricted to a particular tissue, is an important criterion for a transcriptomic marker development by the logic presented above as clearly supported by our data. From a literature review on the expression of some genes showing aberrant expression in inflammatory bowel diseases (IBD) (92-94), we selected one gene (*IL-12*) for preliminary testing on pre-neoplastic and cancerous tissue and stool of the colon, as compared to IBD. To our satisfaction, the *IL-12* gene showed increased expression in stool and tissue of IBD patients, but not colon cancer patients.

Although many other genes have been suggested to participate in colon cancer development in the literature, our in depth review and discussions with other investigators, and analyses using the SAGEmap tools has provided us with reassurance that the goals of developing high sensitivity and specificity transcriptomic approach will be met using a few selected genes with the above approach to screen for colon cancer. Naturally, other promising genes according to the criteria specified above would be added to eventually select a few genes that are indicative of colon cancer in stool and tissue in a future planned larger prospective clinical study.

Table V. Comparison of tests employed for preneoplastic colon cancer screening[†].

Test specifications	FOBT ¹ GuaicImmunoI		Methylated genes ²	Promoter methylation ³	Mutated DNA markers ⁴	Colonoscopy examination ⁵	Transcriptomic approach ⁶
Noninvasive	Yes	Yes	Yes	Yes	Yes	No	Yes
Sensitivity	10.8%	16.3%	Undetermined	31%	18.2%	87%	>95%
Specificity	95%	94.5%	Undetermined	95%	94.4	100%	>95%
Cost ⁷	\$20	\$22	\$75 ⁸	\$75 ⁸	\$695	\$820	\$100 ⁸

[†]For polyps ≥ 1 cm in diameter; ¹References 102 to 104; ²Based on one gene (44); ³Based on one gene (45); ⁴Reference 16; ⁵Reference 105;

⁶Based on our data; ⁷References 103 and 104; ⁸Estimates derived based on our experience with our assay and contact with other test developers.

Rationale for using real-time qPCR for transcriptomic marker development. Real-time amplification allows for a broad dynamic range of target molecule determination, and provides a means to precisely determine the product during the log/linear amplification phase, making the method very sensitive to quantifying low copy number transcripts. Although semi-quantitative PCR methods were thought to be inferior to quantitative competitive ones, side-by-side comparisons showed both assays produced equivalent measures of template abundance (95). This is because methods employing real-time qPCR measure amplification in the logarithmic phase, whereas quantitative competitive measurements determine amplification during the linear phase (24). Roche's LC 2.0™ performed consistently and accurately as it is the only thermocycler on the market that uses the reliable air technology for instantaneous heating/cooling. The newest LC 480™ model in a 96- or 384-well format has solved the problem of rapid heat transfer to blocks by employing an advanced Peltier technology, making it particularly suitable for sensitive and reliable high throughput applications (37, 96).

The choice of valid markers for colon cancer using a transcriptomic approach has been difficult because there are no colon cancer specific markers, and carrying out a reproducible qRT-PCR requires numerous validations due to issues of RNA degradation, PCR inhibitors, template quality, biological replicas and issues of sensitivity and specificity (97). However we and other groups have shown that it is possible by use of RT-qPCR to explore developing diagnostic and prognostic markers by such an inclusive approach (19-21, 51, 82-84, 90, 98, 99).

One gene or groups of genes for colon cancer screening by a transcriptomic approach. The data presented in Figure 5 A, B shows that the genes *IGF2*, *FLNA* and *TGF β -igh* are very good for screening colon cancer in both stool and tissue samples. On the other hand, the genes *CKS2* to *CXCL3* get progressively worse for several reasons: (a) an increase in variability of normal controls, (b) an increase in variability of cancer patients, and c) a smaller difference between the two groups. The *GUC42B* gene shows underexpression. The

IL-12 gene does not distinguish between cancer and controls, but does distinguish between inflammatory and non-inflammatory cases. Looking at various stages of cancer (0 to 3) in Figure 5 C, D, the *CSEIL* and *DPEP1* genes are particularly good for separating all three stages. The *IGF2*, *FLNA* and *TGF β -igh* genes are good for distinguishing stage 3 from the earlier stages. On the other hand, the *CXCL3* gene has a low outlier; however, this patient stays with his/her group on the other genes.

In conclusion, although our findings demonstrate that gene expression derived from carefully selected markers in stool or tissue is a more sensitive/specific diagnostic screening tool for colon cancer than any other molecular markers currently on the market, a large prospective clinical study employing patients at various stages of colon cancer development (from 0 to 4) is needed: (i) using several marker genes in a larger sample; (ii) utilizing several carefully selected housekeeping gene standards to ascertain similar slopes (indicating similar efficiencies) between the target gene and reference gene(s); (iii) correlating the molecular findings with various clinicopathological parameters in different genders and races as clinical usefulness of markers varies with stage, ethnicity and anatomic location of CRC (83, 90); and (iv) considering any other molecular parameters studied (mutations in oncogenes or tumor suppressor genes, protein abundance or truncations, DNA methylation) (42-46, 47). Results from these investigations will lead to selection of a group of fewer than 10 genes for stool and tissue screening, which is considered a more reliable indicator of tumorigenesis than the expression of a single gene.

Clinical application of the transcriptomic data to improve colon cancer screening. Targeting the preneoplastic stages of colon cancer (adenomas, stage 0 or 1) and their successful screening would result in lower rates and reduce the need for surgical intervention and in-patient treatment. As a lengthy period is required for colon cancer to develop from an adenoma, an effective adenoma screening test would need to be performed less frequently than a test for early cancer. However, because only small minorities of adenomas are destined to progress to malignancy, their detection would

involve gross over treatment of patients, which would be costly and harmful to the patient (both physically and emotionally). Therefore, it could be argued that the optimal test would be one that accurately detects advanced adenomas having a high chance of malignant progression (100).

Clinical management of adenomas entails removing them at the time of detection (which is carried out by colonoscopy of the entire colon). However, many patients do not wish to undertake this test because of the need of bowel preparation, dietary restrictions a few days before the test, and potential perforation of the colon and death in some patients. Moreover, colonoscopy tests requiring highly trained staff are likely to be expensive and impractical for wide scale adoption as a screen for the entire eligible population (101).

Given the desirability of using a noninvasive test, acceptable to the target population as an initial screen, many investigators have turned to developing molecular screening approaches, as the inexpensive FOBT (both guaiac and immunological) have quite a low sensitivity for polyp detection (see Table V). Important considerations for developing sound molecular tests include the number of stool specimens required (one sample, or multiple samples over consecutive dates), size of the specimen (whole stool or a sample only) and how the specimens are stored and delivered (*e.g.*, room temperature or frozen, posted or collected). Processing of samples must also be accepted by the laboratory staff. Any test that can be automated would be easier and cheaper to apply on a large scale (100).

We are of the opinion that our transcriptomic approach meets the above criteria for acceptability. For example, it is non invasive, requires only 1 g of stool, does not need sampling on consecutive dates, can be sent by mail in cold packs, is able to differentiate between normal and cancer patients, has a high sensitivity and specificity for detecting polyps, and can be automated, which makes it relatively inexpensive when compared to a test such as one using mutated DNA markers. This screening test has also the potential to develop into a diagnostic assay after we extend results from this study to a larger number of samples and determine the sensitivity and specificity of the test at all Dukes' stages, compared to today's gold standard "colonoscopy". It is envisioned a sensitivity of >95% and a specificity of >95% will be achieved with this transcriptomic test if several genes are included. Since this test is amenable to automation for sample preparation and for performing RT-qPCR quantitation with a real-time cycler, like the Roche LightCycler 480™ in either a 96- or 384-well format, or by having <10 genes printed on a silane slide to produce a "gene expression chip" for colon cancer containing a few genes to be detected on a gene reader in less than 32 h from collecting and processing a stool, as is currently used in quantification of genetically modified organisms (106), it will be easily accepted by laboratory personnel. Automation will also reduce the cost of assay implementation. It is also possible to combine this test

with other noninvasive tests to increase its screening potential. As of today, the promoter methylation test (45) is the only other test showing promise to detect adenomatous polyps with a sensitivity of only 31%, and specificity of about 95%, based on one gene. The sensitivity of that test for detecting polyps may improve as the German investigators are planning to test more marker genes (Kolligs FT, Personal communication). Improvements and the promise of virtual colonoscopy (107), as a less invasive procedure, may also allow combining our transcriptomic screening with CT in the future.

Acknowledgements

We express our gratitude to Sarah Smith Wright of ECU's College of Arts and Science for assistance with graphics; Larry J. Dobbs, Jr. of ECU's Department of Pathology for histopathological insight; Patricia Stockton for LCM work; and to many GI physicians and surgeons from our medical school and nearby clinics for patient accrual and sample acquisition. We are also indebted to Michelle A Moore of Roche Applied Sciences for automatic RNA extraction from stool samples, Carl F Schafer of the NCI's Center for Bioinformatics for compiling CGAP database, and Daniel Rhodes of Cancer Bioinformatics Program at the University of Michigan Medical School, Ann Arbor, for advice on the analysis of ONCOMINE™ database.

References

- 1 Ahmed FE: Colon cancer, prevalence, screening, gene expression and mutation, and risk factors and assessment. *J Env Sci Health C21*: 65-131, 2003.
- 2 American Cancer Society (ACS): Cancer Facts and Figures 2006. ACS, Atlanta, Georgia, USA, 2006.
- 3 American Cancer Society (ACS) Cancer in North America: 1998-2002, Volume One: Incidence, NAACR, based on data collected by cancer registries participating in NCI's SEER Program and CRC'S National Program of Cancer Registries. ACS, Atlanta, Georgia, USA, 2003.
- 4 Ahmed FE: Molecular markers for colon cancer therapy. *Exp Rev Mol Diag 5*: 353-375, 2006.
- 5 Ahmed FE: Effect of diet, life style and other environmental/chemopreventive factors on colorectal cancer development, and assessment of the risks. *J Env Sci Health C22*: 91-147, 2004.
- 6 Ahmed FE: Gene-gene, gene-environmental and multiple interactions in colorectal cancer. *J Env Sci Health C24*: 1-101, 2006.
- 7 Ahmed FE: Role of genes, the environment and their interactions in the etiology of inflammatory bowel diseases. *Exp Rev Mol Diag 6*: 345-363, 2006.
- 8 Atkin W: Options for screening for colorectal cancer. *Scand J Gastroenterol S237*: 13-16, 2003.
- 9 Mullhall BP, Veerappan GR and Jackson JL: Meta-analysis: Computed tomographic colonography. *Ann Intern Med 142*: 635-650, 2005.
- 10 Kealey SM, Dodd JD, MacEneaney PM, Gibney, RG and Malone DE: Minimal preparation computed tomography instead of barium enema/colonoscopy for suspected colon cancer in frail elderly patients: an outcome analysis study. *Clinical Radiol 59*: 44-52, 2004.

- 11 Ahlquist DA, Harrington JJ, Burgart LJ and Roche PC: Morphometric analysis of the "mucocellular layer" overlying colorectal cancer and normal mucosa: relevance to exfoliation and stool screening. *Hum Pathol* 31: 51-57, 2000.
- 12 Nair P, Lagerhola S, Dutta S, Shami S, Davis K, Ma S and Malayeri M: On the nature of cellular elements for stool in the pathophysiology of colonic diseases. *J Clin Gastroenterol* 36(Suppl 1): S44-S39, 2003.
- 13 Osborn NK and Ahlquist DA: Stool screening for colorectal cancer: molecular approaches. *Gastroenterology* 128: 192-206, 2005.
- 14 Ahlquist DA and Shuber AP: Stool screening for colorectal cancer: evolution from occult blood to molecular markers. *Clin Chim Acta* 31: 157-168, 2002.
- 15 Travieso G, Shuber A, Levin B, Johnson C, Olsson L, Schoetz DJ Jr, Hamilton S-R, Boynton K, Kinzler KW and Vogestein B: Detection of APC mutations in fecal and DNA from patients with colorectal tumors. *New Eng J Med* 346: 311-320, 2002.
- 16 Ahlquist DA, Skoletsky JE, Boynton KA, Harrington JJ, Mahoney DW, Pierceall WE, Thibodeau SN and Shuber AP: Colorectal cancer screening by detection of altered human DNA in stool: feasibility of a multitarget assay panel. *Gastroenterology* 119: 1219-1227, 2000.
- 17 Imperiale TF, Ransohoff DF, Itzkowitz SH, Turnbull BA and Ross MA: Fecal DNA *versus* fecal occult blood for colorectal cancer screening in an average-risk population. *New Eng J Med* 351: 2704-2714, 2004.
- 18 Sidransky D: Nucleic acids methods for the detection of cancer. *Science* 278: 1054-1058, 1997.
- 19 Ahmed FE, iJames S, Lysle DT, Dobbs LJ Jr, Johnke RM, Flake G, Stockton P, Sinar DR, Naziri W, Evans MJ, Kovacs CJ and Allison RR: Improved methods for extracting RNA from exfoliated human colonocytes from stool and RT-PCR analysis. *Dig Dis Sci* 49: 2650-2659, 2004.
- 20 Ahmed FE, Dobbs LJ Jr, Johnke RM, iJames S, Lysle DT, Sinar DR, Naziri W, Evans MJ, Kovacs CJ, Daly BM and Allison RR: Isolation of circulating colon carcinoma cells for RT-PCR. *Anal Biochem* 291: 301-304, 2004.
- 21 Ahmed FE and Vos P: Molecular markers for human colon cancer in stool and blood identified by RT-PCR. *Anticancer Res* 24: 4127-4134, 2004.
- 22 Deschner EE: Early proliferative changes in gastrointestinal neoplasia. *Am J Gastroenterol* 77: 207-211, 1982.
- 23 Ahmed FE: Laser microdissection: application to carcinogenesis. *Cancer Genom Proteom* 3: 217-226, 2006.
- 24 Ahmed FE: Molecular methods for studying differential gene expression in carcinogenesis. *J Env Sci Health C20*: 77-116, 2002.
- 25 Gregg K, Zhou W and Davis S: Enhanced sensitivity RNA gel loading buffer that enables efficient RNA separation on native gels. *BioTechniques* 36: 334-336, 2004.
- 26 Hashimoto JG, Beadles-Bohling AS and Wiren KM: Comparison of RiboGreen and 18S rRNA quantitation for monitoring real-time RT-PCR expression analysis. *BioTechniques* 36: 54-60, 2004.
- 27 Kuecker SJ, Jin L, Kulig E, Oudrago GL, Roche PC and Lloyd RV: Analysis of *PRL*, *PRL-R*, *TGFβ1* and *TGFβ-RII* gene expression in normal neoplastic breast tissue after laser capture microdissection. *Appl Immunol Molec Morph* 7: 193-200, 1999.
- 28 Roche Molecular Biochemicals: "The MagNA pure LC™" Automated System, Mannheim, Germany, 2002.
- 29 Tellman G: The E-method: a highly accurate technique for gene-expression analysis. *Nature Methods* 3: i-ii, 2006.
- 30 http://www.lightcycler-online.com/lc_sys/soft_nd.htm#quant.
- 31 LightCycler Software®, Version 3.5. Roche Molecular Biochemicals, Mannheim, Germany, pp. 64-79, 2001.
- 32 Pfaffl MW: Quantification strategies in real-time PCR. *In: A-Z of Quantitative PCR*. Bustin SA (ed.). International University Line, La Jolla, CA, USA, pp. 89-120, 2004.
- 33 Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} method. *Methods* 25: 402-408, 2001.
- 34 Tichopad A, Dilger M, Schwarz G and Pfaffl MW: Standardised determination of real-time PCR efficiency from a single reaction setup. *Nucleic Acids Res* 31: 21-33, 2003.
- 35 Mouritzen P, Nielsen PS, Jacobsen N, Noerholm M, Lomholt C, Pfundheller HM, Ramsing NB, Kauppinen S and Tolstrup N: The ProbeLibrary™ – expression profiling 99% of all human genes using only 90 dual-labeled real-time PCR probes. *BioTechniques* 37: 492-494, 2004.
- 36 Kennedy B, Arar K, Reja V and Henry RJ: Locked nucleic acids for optimizing displacement probes for quantitative real-time PCR. *Anal Biochem* 348: 294-299, 2006.
- 37 Ahmed FE: qRT-PCR: application to carcinogenesis. *Cancer Genom Proteom* 2: 317-332, 2005.
- 38 Bustin SA (ed.): A-Z of Quantitative PCR. International University Line, La Jolla, CA, USA, 2004.
- 39 Longo MC, Berninger MS and Hanley JL: Use of uracil DNA glycosylase to control carryover contamination in polymerase chain reaction. *Gene* 93: 125-128, 1990.
- 40 Thornton CG, Hartley JL and Rashtchian A: Utilizing uracil DNA glycosylase to control carryover contamination in PCR: characterization of residual UDG activity following thermal cycling. *BioTechniques* 13: 181-183, 1992.
- 41 Wegman E: Hyperdimensional data analysis using parallel coordinates. *J Am Stat Assoc* 85: 644-675, 1990.
- 42 Davies RJ, Freeman A, Morris LS, Bingham S, Dilworth S, Scott I, Laskey RA, Miller R and Coleman N: Analysis of minichromosome maintenance proteins as a novel method for detection of colorectal cancer in stool. *Lancet* 359: 1917-1919, 2002.
- 43 Iacopetta B: Aberrant DNA methylation: have we entered the era of more than one type of colorectal cancer? *Am J Pathol* 162: 1043-1045, 2003.
- 44 Müller HM, Oberwalder M, Fiegl H, Morandell M, Goebel G, Zitt M, Mühlthaler M, Öfner D, Margreiter R and Widschwendter M: Methylation changes in faecal DNA: a marker for colorectal cancer screening. *Lancet* 363: 1283-1285, 2004.
- 45 Lenhard K, Bommer GT, Asutay S, Schauer R, Brabletz T, Göke B, Lamerz R and Kolligs T: Analysis of promoter methylation in stool: a novel method for the detection of colorectal cancer. *Clin Gastroenterol Hepatol* 3: 142-149, 2005.
- 46 Sager R: Expression genetics in cancer: shifting the focus from DNA to RNA. *Proc Natl Acad Sci USA* 94: 952-955, 1997.
- 47 Ahmed FE: Expression microarray proteomics and the search for cancer biomarkers. *Current Genomics* 7: 399-426, 2006.
- 48 Greenbaum D, Colangelo C, Williams K and Gerstein M: Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol* 4: 117-126, 2003.
- 49 Ellmark P, Belov L, Huang P, Lee CS, Solomon MJ, Morgan DK and Christopherson RI: Multiplex detection of surface molecules on colorectal cancers. *Proteomics* 6: 1791-1802, 2006.

- 50 Ransohoff DF: Lessons from controversy: ovarian cancer screening and serum proteomics. *J Nat Cancer Inst* 79: 315-319, 2005.
- 51 Giordano TJ, Shedder KA, Schwartz DR, Kuick R, Taylor JMG, Misek DE, Greenson JK, Kardia SLR, Beer DG, Rennert G, Cho KR, Gruber SB, Fearon ER and Hanash S: Organ-specific molecular classification of primary lung, colon and ovarian adenocarcinomas using gene expression profiles. *Am J Pathol* 159: 1231-1238, 2001.
- 52 Rex D: Current colorectal cancer screening strategies: overview and obstacles to implementation. *Rev Gastroenterol Dis* 5: S2-S11, 2002.
- 53 Levin B, Brooks D, Smith RA and Stone A: Emerging technologies in screening for colorectal cancer: CT colonography, immunochemical fecal occult blood test, and stool screening using molecular markers. *CA Cancer J Clin* 53: 44-55, 2003.
- 54 Schmittgen TD and Zakrajsek BA: Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. *J Biochem Biophys Methods* 46: 69-81, 2000.
- 55 Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G and Zumla A: Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *BioTechniques* 37: 112-119, 2004.
- 56 Tricarico C, Pinzani P, Bianchi S, Paglierani M, Distanto V, Pazzagli M, Bustin SA and Orlando C: Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. *Anal Biochem* 309: 293-300, 2002.
- 57 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F: Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3: research 0034,1-0034,11, 2002.
- 58 de Kok JB, Roelofs RW, Giesendorf BA, Pennings JL, Waas ET, Feuth T, Swinkels DW and Span PN: Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. *Lab Invest* 85: 154-159, 2005.
- 59 Pérez-Novo CA, Claeys C, Speleman F, Van Cauwenberge P, Bachert C and Vandesompele J: Impact of RNA quality on reference gene expression stability. *BioTechniques* 39: 52-56, 2005.
- 60 Vogelstein B and Kinzler KW: Digital PCR. *Proc Natl Acad Sci USA* 96: 9236-9241, 1999.
- 61 Dong SM, Traverso G, Johnson C, Geng L, Favis R, Boynton K, Hibi K, Goodman SN, D'Allesio M, Paty P, Hamilton SR, Sidransky D, Barany F, Levin B, Shuber A, Kinzler KW, Vogelstein B and Jen J: Detecting colorectal cancer in stool with the use of multiple genetic targets. *J Natl Cancer Inst* 93: 858-865, 2001.
- 62 Zhang L, Zhou W, Velculescu VE, Kern SE, Hruban RH, Hamilton SR, Vogelstein B and Kinzler KW: Gene expression profiles in normal and cancer cells. *Science* 276: 1268-1272, 1997.
- 63 Zhou W, Sokoll LJ, Bruze K, Zhang L, Velculescu VE, Goldin SB, Hruban RH, Kern SE, Hamilton SR, Chan DW, Vogelstein B and Kinzler KW: Identifying markers for pancreatic cancer by gene expression analysis. *Cancer Epidemiol Biomarkers Prevent* 7: 109-112, 1998.
- 64 Velculescu V E, Zhang L, Vogelstein B and Kinzler KW: Serial analysis of gene expression. *Science* 270: 484-487, 1995.
- 65 Buckhaults P, Rago C, St Croix B, Romans KE, Saha S, Zhang L, Vogelstein B and Kinzler KW: Secreted and cell surface genes expressed in benign and malignant colorectal tumors. *Cancer Res* 61: 6996-7001, 2001.
- 66 Birkenkamp-Demtroder K, Lotte C, Olesen SH, Frederiksen CM, Laiho P, Aaltonen LA, Laurberg S, Sørensen B, Hagemann R and Ørntoft TF: Gene expression in colorectal cancer. *Cancer Res* 62: 4352-4363, 2002.
- 67 Backert S, Gelos M, Mann B, Löven N, Gratchev A, Kobalz U, Hanski M-L, Böhm C, Mansmann U, Moyer MP, Riecken E-O and Hanski C: Differential expression in colon carcinoma cells and tissues detected with cDNA array. *Int J Cancer* 52: 868-874, 1999.
- 68 Notterman DA, Alon U, Sierk AJ and Levine AJ: Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide array. *Cancer Res* 61: 3124-3130, 2001.
- 69 Kitahara O, Furukawa Y, Tanaka T, Kihara C, Ono K, Yanagawa R, Nita ME, Takagi T, Nakamura Y and Tsunoda T: Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissue and normal epithelia. *Cancer Res* 61: 3544-3549, 2001.
- 70 Nishizuka S, Chen S-T, Gwadry FG, Alexander J, Major SM, Scherf U, Reinhold WC, Waltham M, Charboneau L, Young L, Bussey KJ, Kim S, Lababidi S, Lee JK, Pittaluga S, Scudiero DA, Sausville EA, Munson PJ, Petricoin EF 3rd, Liotta LA, Hewitt SM, Raffeld M and Weinstein JN: Diagnostic markers that distinguish colon and ovarian adenocarcinomas: identification by genomic, proteomic and tissue array profiling. *Cancer Res* 63: 5243-5250, 2003.
- 71 Yanagawa R, Furukawa Y, Tsunoda T, Kitahara O, Kameyama M, Murata K, Ishikawa S and Nakamura Y: Genome-wide screening of genes showing altered expression in liver metastasis of human colorectal cancer by cDNA microarray. *Neoplasia* 3: 395-401, 2001.
- 72 Wang Y, Jatke T, Zhang Y, Mutchim G, Talantov D, Jiang J, McLeod H and Alkins I: Gene expression profile and molecular marker to predict the recurrence of Duke B colon cancer. *J Clin Oncol* 22: 1564-1571, 2004.
- 73 Zou T-T, Selaru FM, Xu Y, Shustova V, Yin J, Mori Y, Shibata D, Sato F, Wang S, Olaru A, Deacu E, Kiu TC, Abraham JM and Meltzer S: Application of cDNA microarrays to generate a molecular taxonomy capable of distinguishing between colon cancer and normal colon. *Oncogene* 21: 4855-4862, 2002.
- 74 Lin Y-M, Furukawa Y, Tsunoda T, Yue CT, Yang KC and Nakamura Y: Molecular diagnosis of colorectal tumor by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. *Oncogene* 21: 4120-4128, 2002.
- 75 Frederiksen, CM, Knudsen S, Laurberg S, Ørntoft TF: Classification of Dukes' B and C colorectal cancer using expression arrays. *J Cancer Res Clin Oncol* 129: 263-271, 2003.
- 76 Bertucci F, Salas S, Eysteris S, Nasser V, Finetti P, Rougemont J, Charafe-Jauffret E, Nasser V, Liorid B, Camerlo J, Tagett R, Tarpin C, Houvenaeghel G, Nguyen C, Maraninchi D, Jacquemier J, Houlgatte R, Birnbaum D and Viens P: Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters. *Oncogene* 23: 1377-1391, 2004.

- 77 Stremmel C, Wein A, Hohenberger W and Reingruber B: DNA microarrays: a new diagnostic tool and its implications in colorectal cancer. *Int J Colorectal Dis* 17: 131-136, 2002.
- 78 Ahmed S, Banerjee A, Hands RE, Bustin S and Dorudi S: Microarray profiling of colorectal cancer in Bangladeshi patients. *Colorectal Dis* 7: 571-575, 2004.
- 79 Hegde P, Qi R, Gaspard R, Abernathy K, Dharap S, Earle-Hughes J, Gay C, Nwokekeh NU, Chen T, Saeed AI, Sharov V, Lee NH, Yeatman TJ and Quackenbush J: Identification of tumor markers in models of human colorectal cancer using a 19,200-element complementary DNA microarray. *Cancer Res* 61: 7792-7797, 2001.
- 80 Takemasa I, Higuchi H, Yamamoto H, Sekimoto M, Tomita N, Nakamori S, Matoba R, Monden M and Matsubara K: Construction of preferential cDNA microarray specialized for human colorectal carcinoma: molecular sketch of colorectal cancer. *Biochem Biophys Res Commun* 285: 1244-1249, 2001.
- 81 Williams NS, Gaynor RB, Scoggin S, Verma U, Gokaslan T, Simmang C, Fleming J, Tavana D, Frenkel E and Beccerra C: Identification and validation of genes involved in the pathogenesis of colorectal cancer using cDNA microarrays and RNA interference. *Clin Cancer Res* 9: 931-946, 2003.
- 82 Davidson LA, Lupton JR, Miskovsky E, Fields AP and Chapkin RS: Quantification of human intestinal gene expression profiling using exfoliated colonocytes: a pilot study. *Biomarkers* 8: 51-61, 2003.
- 83 Ahmed FE: Colon cancer: molecular staging models that characterize risk and enhance progression. *Cancer Genomics Proteomics* 1: 295-310, 2004.
- 84 Alexander R and Raicht RE: Purification of total RNA from human stool samples. *Dig Dis Sci* 43: 2652-2658, 1998.
- 85 Coppola D, Hyacinthe M and Fu L: *CD44v6* expression in human colorectal carcinoma. *Hum Pathol* 29: 627-635, 1998.
- 86 Allgayer H: Molecular regulation of urokinase-receptor gene expression as one potential concept for molecular staging and therapy. *Rec Results Cancer Res* 162: 15-30, 2003.
- 87 Schutte PM, Glémet HC, Fiebig AA and Powers DA: Adaptive variation in lactate dehydrogenase B gene expression: role of a stress-responsive regulatory element. *Proc Natl Acad Sci USA* 97: 6597-6602, 2000.
- 88 Xin B, Platzer P, Fink SP, Reese L, Nosrati A, Wilson JKV, Wilson K and Markowitz S: Colon cancer secreted protein-2 (CCSP-2) a novel candidate serological marker of colon neoplasia. *Oncogene* 24: 724-731, 2005.
- 89 Lal A, Lash AE, Allschul SF, Velculescu V, Zhang L, McLendon RE, Marra M, Prange C, Morin PJ, Polyak K, Papadopoulos N, Vogelstein B, Kinzler KW, Strausberg RL and Riggins GJ: A public database of gene expression in human cancers. *Cancer Res* 59: 5403-5407, 1999.
- 90 Ahmed FE: Development of novel diagnostic and prognostic molecular markers for sporadic colon cancer. *Exp Rev Mol Diag* 5: 337-352, 2005.
- 91 Ahmed FE: Serial analysis of gene expression (SAGE) technology. In: *Encyclopedia of Diagnostic Genomics and Proteomics*. Fuchs J, Podda M (eds.). Marcell Dekker, NY, USA, pp. 1191-1196, 2004.
- 92 Nelsen OH, Kirman I, Rüdiger N, Hendel J and Vainer B: Upregulation of interleukin-12 and -17 in active inflammatory bowel disease. *Scand J Gastroenterol* 38: 180-185, 2003.
- 93 Dimberg J, Samuelsson A, Hugander A and Söderkvist P: Differential expression of cyclooxygenase 2 in human colorectal cancer. *Gut* 45: 730-732, 1999.
- 94 Berrebi D, Maudinas R, Hugot J-P, Chamaillard M, Chareyne F, Lagausie P, Yang C, Desreumaux P, Giovannini M, Cezard J-P, Zouali H, Emilie D and Peuchmaur M: Card 15 gene overexpression in mononuclear epithelial cells of the inflamed Crohn's disease colon. *Gut* 52: 640-646, 2004.
- 95 Haberbaugen G, Pinsl J, Kuha C-C and Markert-Hahn C: Comparative study of different standardization concepts in quantitative competitive reverse transcription PCR assay. *J Clin Microbiol* 36: 628-633, 1998.
- 96 Ahmed FE: Molecular markers for early cancer detection. *J Env Sci Health C18*: 75-126, 2000.
- 97 Nolan T, Hands RE, Ogunkolade W and Bustin SA: SPUD: A quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. *Anal Biochem* 351: 308-310, 2006.
- 98 Lloyd JM, McIver CM, Stephenson S-A, Hewett PJ, Rieger N and Hardingham JE: Identification of early-stage colorectal cancer patients at risk of relapse post-resection by reverse transcription-PCR analysis of peritoneal lavage fluid for malignant cells. *Clin Cancer Res* 12: 417-423, 2006.
- 99 Chen G, McIver CM, Texler M, Lloyd JM, Rieger N, Hewett PJ, Wan DS and Hardingham JE: Detection of occult metastasis in lymph nodes from colorectal cancer patients: a multiple-marker reverse transcriptase-polymerase chain reaction study. *Dis Colon Cancer* 47: 679-686, 2004.
- 100 Davies RJ, Miller R and Coleman N: Colorectal cancer screening: prospects for molecular stool analysis. *Nature Rev Cancer* 5: 199-209, 2005.
- 101 Gatto NM, Frucht H, Sundararajan V, Jacobson JS, Grann VR and Neugut AI: Risk of perforation after colonoscopy and sigmoidoscopy: a population based study. *J Natl Cancer Inst* 95: 230-236, 2002.
- 102 Morikawa T, Kato J, Yamaji T, Wada R, Miksushima T and Shiratori Y: Comparison of the immunochemical fecal occult blood test and total colonoscopy in the asymptomatic population. *Gastroenterology* 125: 422-428, 2006.
- 103 Song K, Fendrick AM and Ladabaum U: Fecal DNA testing compared with commercial colorectal cancer screening methods: a decision analysis. *Gastroenterology* 126: 1270-1279, 2004.
- 104 Allison JE: Colon cancer screening guidelines 2005: the Fecal Occult Blood Test option has become a better FITT. *Gastroenterology* 125: 745-748, 2005.
- 105 Bressler B, Paszat L F, Vinden C, Li C, He J and Rabeneck L: Colonoscopic miss rates for right-sided colon cancer: population-based study. *Gastroenterology* 127: 452-456, 2004.
- 106 Ahmed FE (ed.): *Testing of Genetically Modified Organisms in Food*. Haworth Press, Binghamton, NY, USA, 2004.
- 107 Imperiale TF: Can computed tomographic colonography become a "good" screening test? *Ann Intern Med* 142: 669-670, 2005.

Received November 11, 2006

Revised December 11, 2006

Accepted December 14, 2006