

A homologue of the *Drosophila* headcase protein is a novel tumor marker for early-stage colorectal cancer

CHIH-CHENG CHIEN^{1,2*}, CHUN-CHAO CHANG^{3,4*}, SHUNG-HAUR YANG^{5,6},
SHU-HUNG CHEN⁷ and CHI-JUNG HUANG^{2,8}

¹Department of Anesthesiology, Cathay General Hospital, Neihu, Taipei 11449; ²School of Medicine, Fu Jen Catholic University, Taipei 24255; ³Division of Gastroenterology, Department of Internal Medicine, Taipei Medical University Hospital; ⁴Digestive Disease Research Center, Taipei Medical University, Taipei 11031; ⁵Division of Colon and Rectal Surgery, Department of Surgery, Taipei Veterans General Hospital, Taipei 11217; ⁶Institute of Clinical Medicine, National Yang-Ming University, Taipei 11221; ⁷Department of Surgery, Cathay General Hospital, Hsinchu, 30060; ⁸Molecular Genetics and Biochemistry Laboratory, Cathay Medical Research Institute, Cathay General Hospital, Taipei 22174, Taiwan, R.O.C.

Received October 31, 2005; Accepted December 22, 2005

Abstract. The altered expression of certain genes is frequently detected as a hallmark of malignant tumors. These differentially expressed genes have the potential to become molecular markers. We purified the fecal mRNA from patients with colorectal cancer to identify novel candidates by using oligonucleotide microarray hybridization. We identified 21 upregulated and 22 downregulated candidates with significantly altered expression in patient fecal samples that have not been previously characterized. A gene encoding a homologue of the *Drosophila* headcase protein (HECA) was further examined due to its high ranking and possible importance in some human cancers. A tendency towards increased expression of HECA was dependent upon the clinicopathological stage. In this report, healthy individuals expressed less HECA, either in their blood samples or feces. Moreover, we detected upregulated HECA in blood and fecal samples of patients with colorectal cancer, and its expression level was shown to be significantly correlated with disease status. Our data show that HECA may be an early-stage classifier of colorectal cancer that can discriminate between late- and early-stage disease. In conclusion, this study is the first to analyze

differentially expressed genes in the feces of colorectal cancer patients using oligonucleotide microarrays. The data suggest that HECA expression levels in feces may be an effective classifier for early-stage colorectal cancer.

Introduction

Human cancers frequently result from multiple genetic alterations (1). The altered expression of certain genes has been detected in a variety of malignant tumors (2,3). Colorectal cancer has shown a dramatic increase in incidence to become one of the most common malignancies worldwide (4). Because it is a molecular genetic disease, cancer genes and their ultimate expression could be keys to the development of successful therapies (5). Thus, it is not surprising that many genes have been characterized with aberrant expression and mutations in colorectal cancer (6). However, most of the differentially expressed genes were identified using array-based studies with cancerous and corresponding noncancerous colonic epithelia (7,8).

Studies using human feces have been the target for noninvasive colorectal cancer detection in the last decade (9). The feasibility of isolating intact colonic cells from feces has been well demonstrated (10). Moreover, the analysis of molecular biomarkers has been used for some specific medical diagnoses. For examples, fecal protein serves as a sample for enzyme-linked immunosorbent assays (ELISA) (11), and fecal DNA has been examined for the identification of intestinal bacteria (12) and viruses causing gastroenteritis (13). For the purpose of analyzing gene expression, fecal total RNA must first be purified from frozen feces (14). Upon examination, many genes were shown to change their expression in colorectal cancer patients (15,16). We previously established an organic solvent-free method of extracting fecal total RNA and revealed the aberrant presence of fecal CK19 to be a potential biomarker of colorectal cancer progression (17). However, there are likely to be

Correspondence to: Dr Chi-Jung Huang, Molecular Genetics and Biochemistry Laboratory, Cathay Medical Research Institute, Cathay General Hospital, No. 32, Ln. 160, Jian-Cheng Road, Hsi-Chih, Taipei 22174, Taiwan, R.O.C.
E-mail: aaronhuang@cgh.org.tw

*Contributed equally

Key words: colorectal cancer, fecal differential expressed genes, oligonucleotide microarray, homologue of the *Drosophila* headcase protein, stage classifier

Table I. Sequence, expected fragment size and annealing temperature of primers used in the RT-PCR analysis of the upregulated candidate in patients with colorectal cancer.

Gene	Sequences (5' to 3')	Expected fragment size (bp)	Annealing temperature (°C)
HECA ^a			
Outer ^b			
Forward ^c	TTTCATTGGAGGTTGGGAAG	323	58
Reverse	CATTTACAAAACAGAAACCCTCT		
Inner			
Forward	TGGATGGGCTCCATTAAAAC	157	66 ^d
Reverse ^b	CCACACACAATTCCACAACC		

^aHECA, NM_016217 (accession number) and Hs. 197644 (UniGene number). ^bIn fecal gene amplifications, the first round was performed with 0.1 μ M outer primers and subsequent amplification with 0.4 μ M inner primers. ^cForward, forward primer; Reverse, reverse primer. ^dAmplifications from cell lines, multiple tissue panel, tissue, and blood samples were carried out with only inner primers (0.1 μ M), and the annealing temperature was set at 60°C.

many other molecular biomarkers with different expression levels in the feces of patients with colorectal cancer.

Advanced microarray technologies were used to define global changes in gene expression (18). Many researchers have reported expression array-based studies analyzing colorectal cancer with corresponding noncancerous colonic epithelia (7,8,19). With a view toward colorectal cancer screening across the general population, we purified the fecal mRNA from patients with this tumor. Such fecal mRNA served as targets for microarray hybridization to identify differentially expressed genes in feces. Further validation was performed with samples of feces, blood, and paired tissue using reverse transcription-PCR (RT-PCR). The aim of the present study was to investigate the uncharacterized candidates that correlated with colorectal cancer and may lead to a possible fecal molecular marker of colorectal cancer.

Materials and methods

Patients, tissue, blood, and fecal samples. Patients with colorectal cancer were classified by Dukes' staging, and their initial tumor stages were categorized. All clinical samples in this study were obtained from the Cathay General Hospital or Taipei Veterans General Hospital with the informed consent of the patient or their legal representative. Colorectal cancer and corresponding normal tissues were snap-immersed into RNAlater and stored according to the manufacturer's instructions (Ambion, Austin, TX). Blood in EDTA tubes and feces in specific preservation buffer were collected from healthy volunteers and patients with colorectal cancer (17).

Cancer cell lines and human multiple tissue cDNA. Seven colon cancer cell lines (DLD-1, SW480, LS174T, HT29, LoVo, Caco-2, and HCT116) with epithelial morphology were purchased either from the American Type Culture Collection (ATCC; Manassas, VA) or Bioresource Collection and Research Center (BCRC; Hsinchu, Taiwan). All cultured cells were maintained in Dulbecco's modified Eagle's medium with 5 mM glutamine, 1:500 antibiotics/

antimitotics, and 10% fetal calf serum, according to routine culture procedures. The human multiple tissue cDNA panel (spleen, thymus, prostate, testis, ovary, small intestine, colon, and peripheral blood leukocytes), Human MTC Panel II (BD Biosciences Clontech, Mountain View, CA), was used as a PCR template according to the manufacturer's protocol.

Total RNA extraction. Total RNA from tissues and cultured cells were extracted using Trizol reagent, according to the manufacturer's procedures (Invitrogen, Carlsbad, CA). Blood samples were prepared using the QIAamp RNA Blood Mini kit according to the manufacturer's instructions (Qiagen, Valencia, CA). Fecal total RNA was purified according to our previous report (17).

Fecal mRNA preparation and linear amplification for labeling. For oligonucleotide array hybridization, approximately 0.1 mg of fecal total RNA was further purified using the Oligotex mRNA kit (Qiagen) according to the manufacturer's protocol. A normal pool was made from equal amounts of total fecal RNA from two healthy adult males. Approximately 5 ng of each fecal mRNA was first amplified with a RiboAmp OA RNA Amplification kit (Arcturus, Mountain View, CA) to yield 0.2 to 1.5 μ g of amplified RNA. Fluorescently labeled cRNA was then synthesized by *in vitro* transcription of fecal amplified RNA (0.2-1.5 μ g) using the Agilent Fluorescent Linear Amplification kit (Agilent Technologies, Palo Alto, CA) in the presence of Cy5-CTP (Perking-Elmer, Wellesley, MA) according to the manufacturer's procedures. An appropriate universal human reference RNA (272 ng) (Stratagene, La Jolla, CA) was also amplified and labeled as described above except that the fluorescent dye was changed to Cy3-CTP (Perking-Elmer).

Oligonucleotide arrays and data extraction. For each hybridization experiment, 4 μ g of Cy-labeled cRNA (2 μ g Cy5-labeled fecal target and 2 μ g Cy3-labeled reference target) was fragmented to an average size of approximately 50-100 nucleotides by incubation with fragmentation buffer

Table II. Novel candidates demonstrating a significant upregulated change in the feces of patients with early-stage colorectal cancer (>3-fold difference, p<0.01).

Accession number	Description	Fold change	p-value
NM_021199	Moderate similarity to sulfide: quinone oxidoreductase (SQRDL)	9.71	0.0038
NM_017721	Putative NF- κ B activating (CC2D1A)	8.84	<0.0001
NM_032671	Hypothetical protein MGC10814	8.61	0.0079
NM_016217	Protein of unknown function (HECA)	7.75	0.0038
NM_024829	Hypothetical protein FLJ22662	6.45	<0.0001
NM_030627	Containing two RNA recognition motifs (CPEB4)	5.74	<0.0001
XM_209073	Protein of unknown function (METRNL)	5.69	0.0057
NM_018169	Hypothetical protein FLJ10652	5.52	0.0090
XM_496802	Hypothetical protein FLJ45422	5.12	0.0002
NM_016470	Chromosome 20 open reading frame 111 (C20orf11)	4.65	0.0012
NM_01818	Hypothetical protein FLJ10700	4.63	0.0035
NMJ45293	Hypothetical protein FLJ20897 (LOC196549)	4.47	0.0053
NM_005410	Strong similarity to human SEPP1	4.18	0.0021
AB040934	mRNA for KIAA1501 protein	4.08	0.0003
NM_033180	Olfactory receptor, family 51, subfamily B, member 2 (OR51B2)	4.03	<0.0001
NM_030759	Strong similarity to nuclear receptor binding factor 2 (NRBF2)	4.00	0.0033
NM_018022	Transmembrane protein 51 (TMEM51)	3.61	0.0004
NM_152834	Transmembrane protein 18 (TMEM 18)	3.46	0.0028
NM_024069	Hypothetical protein MGC2749	3.29	0.0030
NM_016053	Homo sapiens CGI-116 protein	3.22	0.0009
NM_032342	Chromosome 9 open reading frame 125 (CC9orf125)	3.04	0.0005

at 60°C for 30 min. Fragmented labeled cRNA was then hybridized to the Agilent Human 1A oligo microarray (version 2) (Agilent Technologies) at 60°C for 17 h (20,21). After washing and drying under nitrogen, microarrays were scanned with an Agilent microarray scanner (Agilent Technologies) at 535 nm for Cy3 and 625 nm for Cy5. Scanned images were analyzed by Agilent feature extraction software (version 7.5) (Agilent Technologies), and the quantified signal and background intensities for each feature were substantially normalized by the rank consistency filtering Lowess method (22). Signal information from patients with colorectal cancer was compared with the normal pool.

Reverse transcription-polymerase chain reaction (RT-PCR). We performed RT-PCR to evaluate candidates from the microarray. Clinical samples (tissue, blood and feces) and colorectal cancer cell lines were subjected to RT-PCR analyses. Total RNA (1 μ g) was reverse transcribed for single-stranded cDNAs using an oligo(dT)12 primer with the PowerScript® Reverse Transcriptase kit (BD Biosciences Clontech) according to the manufacturer's protocol. Synthesized cDNA could be used directly in the PCR. A housekeeping gene, glyceraldehydes-3-phosphate dehydrogenase (G3PDH; for tissue and blood samples) or cytokeratin-18 (CK18; for fecal

samples), was used as an internal control. Primer sequences and PCR programs for these two housekeeping genes were as previously described (17,23). Primer sequences and optimal PCR annealing temperatures for target genes are listed in Table I. Except for the amplifications from fecal total RNA, each PCR was carried out in one round in a 20 μ l reaction volume. In fecal gene amplifications, subsequent amplifications with nested primers used 2 μ l of a 50-fold dilution of the first round product as the template with a final reaction volume of 20 μ l. All final PCR products were separated on 2% agarose gel, stained with ethidium bromide, and visualized under UV illumination. PCR bands of the predicted size were isolated and subjected to sequencing (ABI 3100; Applied Biosystems, Foster City, CA) to confirm gene identities. Expression levels of the target gene were quantified using ImageQuant software (Amersham Biosciences, Piscataway, NJ) and normalized to G3PDH levels (24).

Statistical analysis. The Student's t-test was used, with statistical significance set at p<0.01 to extract data from the hybridization of oligonucleotide array, and p<0.05 to compare differences between low and high expression of target genes. Statistical comparisons were performed with the highest mean intensity.

Table III. Novel candidates demonstrating a significant downregulated change in the feces of patients with early-stage colorectal cancer (<1/3-fold difference, $p < 0.01$).

Accession number	Description	Fold change	p-value
NM_032259	WD repeat domain 24 (WDR24)	0.02	0.0089
NM_023071	Spermatogenesis associated, serine-rich 2 (SPATS2)	0.07	0.0002
NM_018390	Phosphatidylinositol-specific phospholipase C, X domain containing 1 (PLCXD1)	0.07	0.0005
NM_024899	Chromosome 18 open reading frame 9 (C18orf9)	0.14	0.0070
AK095284	Homo sapiens cDNA FLJ37965 fis	0.14	<0.0001
NM_024633	Chromosome 14 open reading frame 139 (C14orf139)	0.14	0.0053
NMJ52753	Signal peptide, CUB domain, EGF-like 3 (SCUBE30)	0.14	<0.0001
AF000560	TTF-I interacting peptide 20 mRNA	0.16	0.0039
NMJ73831	Zinc finger protein 707 (ZNF707)	0.17	0.0005
NM_032334	Chromosome 8 open reading frame 53 (C8orf53)	0.21	0.0017
BC032249	Hypothetical protein MGC27165	0.21	0.0077
NMJ52358	SCRL protein (SCRL)	0.24	<0.0001
NM_152511	Dual specificity phosphatase 18 (DUSP18)	0.26	0.0065
NM_022451	Nucleolar complex associated 3 homologous (NOC3L)	0.26	0.0098
NM_014679	Translokine (PIG8)	0.27	0.0040
NM_014661	Family with sequence similarity 53, member B (FAM53B)	0.28	0.0038
NMJ38704	Necdin-like 2 (NDNL2)	0.29	0.0034
AB065662	Gene for seven transmembrane helix receptor	0.29	<0.0001
NM_020247	Chaperone, ABC 1 activity of bc 1 complex like (CABC1)	0.29	0.0034
NM_018118	MCM3 minichromosome maintenance deficient 3 associated protein, antisense (MCM3APAS)	0.29	0.0017
AK057531	Homo sapiens cDNA FLJ32969 fis	0.30	0.0079
NM_020685	Chromosome 3 open reading frame 14 (C3orf14)	0.31	0.0037

Results

Novel candidates significantly changed in the feces of patients with early-stage colorectal cancer. In this study, we identified 21 upregulated (>3-fold difference, $p < 0.01$) and 22 downregulated (<1/3-fold difference, $p < 0.01$) novel candidates from the Agilent oligonucleotide microarray system based on intensity and p-value (Tables II and III). These selected candidates showed homology to other well-known proteins, hypothetical proteins, open reading frames, or proteins with unknown function.

A homologue of the Drosophila headcase protein is upregulated and related to early-stage colorectal cancer. To examine the reliability of our microarray data and identify upregulated molecules in patients with colorectal cancer, a gene encoding a homologue of the *Drosophila* headcase protein (HECA; NM_016217) was chosen from the list of

upregulated genes for additional validation (Table II). HECA was chosen based on its top 4 ranking in the list of upregulated novel candidates and its possible importance in some human cancers (25).

Expression of HECA in colorectal cancer cell lines and the multiple tissue panel. The expression of HECA was assessed for its suitability as a specific tumor marker for colorectal cancer. Total RNA was extracted from seven colorectal cancer cell lines, reverse-transcribed into cDNA and then PCR amplified with HECA-specific primers. Although other researchers have shown only low-level expression of HECA in pancreatic and renal cancer cell lines (25), all of the colorectal cancer cell lines in this study expressed HECA (Fig. 1A). Compared with the expression of HECA in LoVo cells (1.2-fold to G3PDH, Dukes' stage C), the relative expression of HECA in two cell lines, SW480 and LS174T (both at Dukes' stage B), was increased 12.0-fold and

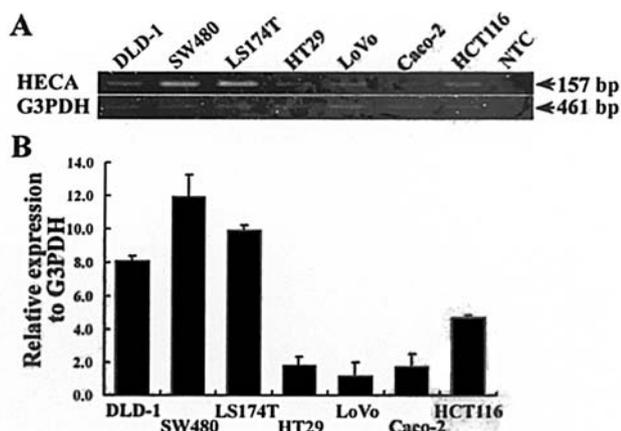


Figure 1. Comparison and relative quantitation of HECA expression in seven colorectal cancer cell lines with variant Dukes' stage using RT-PCR. (A) Expression of HECA and G3PDH in colorectal cancer cell lines, DLD-1, SW480, LS174T, HT29, LoVo, Caco-2, and HCT116. For each cell line, total RNA (1 μ g) was reverse transcribed using an oligo(dT)12 primer, and aliquots of cDNA were amplified with primers specific to G3PDH (25 cycles) and HECA (32 cycles) after establishing the dilution at which amplification remained within the linear range. Sizes of the amplified products are shown on the right (bp). NTC (non-template control) was performed as negative controls for PCR amplification. (B) Relative quantitation of HECA expression normalized to G3PDH. Triplicate amplifications of HECA and G3PDH for each cell line were analyzed to determine the relative levels of HECA expression. Values are shown as mean \pm SD.

10.0-fold, respectively. In addition, HCT116 (Dukes' stage unknown) showed a moderate increase in the expression of HECA to 4.8-fold, and DLD-1 (Dukes' stage C) showed a 8.1-fold upregulated expression. However, the levels of HECA expression in HT29 (1.8-fold) and Caco-2 (1.7-fold) were closer to LoVo that possessed the lowest expression of HECA in this study (Fig. 1B). Moreover, we also employed PCR to analyze the tissue distribution of HECA from multiple human tissue cDNA samples. Likewise PCR products of HECA were readily visible in the spleen, thymus, prostate, testis, small intestine, colon, and peripheral blood leukocytes, but less so in the ovary (Fig. 2A). When comparing the HECA expression levels of colon tissue (1.5-fold to G3PDH) with other tissues, except that of the ovary (0.8-fold), the

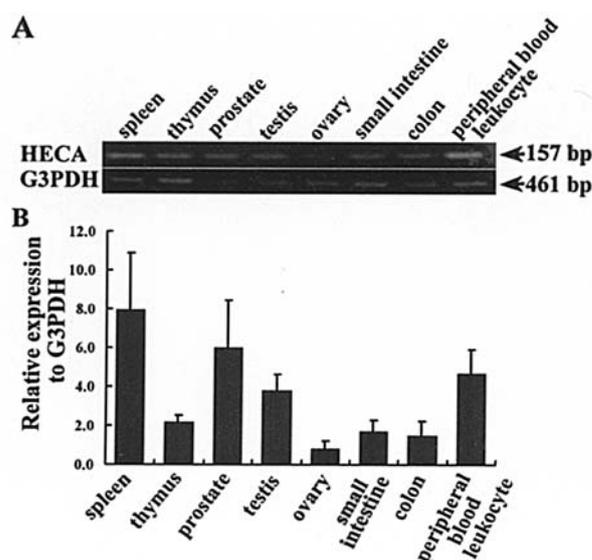


Figure 2. Comparison and relative quantitation of the tissue expression pattern of HECA in multiple human tissues. (A) Expression of HECA and G3PDH in Human MTC Panel II: spleen, thymus, prostate, testis, ovary, small intestine, colon, and peripheral blood leukocytes. Aliquots of cDNA from the listed human tissue sources were used as templates for PCR with primers specific to G3PDH (25 cycles) and HECA (29 cycles) to screen for the expression of transcripts. The size of the amplified products is shown on the right (bp). (B) Relative quantitation of HECA expression normalized to G3PDH. Triplicate amplifications of HECA and G3PDH for each tissue were analyzed to determine the relative levels of HECA expression. Values are shown as mean \pm SD.

spleen (7.9-fold), prostate (6.0-fold), peripheral blood leukocytes (4.7-fold), testis (3.8-fold), and small intestine (1.7-fold) expressed higher levels of HECA (Fig. 2B).

Alteration in the gene expression of HECA in colonic cancerous tissue. To further validate the change in HECA expression detected by the microarray, total RNA was extracted from eight cancers and their matched noncancerous colonic tissue of patients with various Dukes' stages (one in Dukes' stage A, two in Dukes' stage B, three in stage C, and two in stage D). cDNA from these paired tissues was amplified by PCR with HECA-specific primers after establishing the

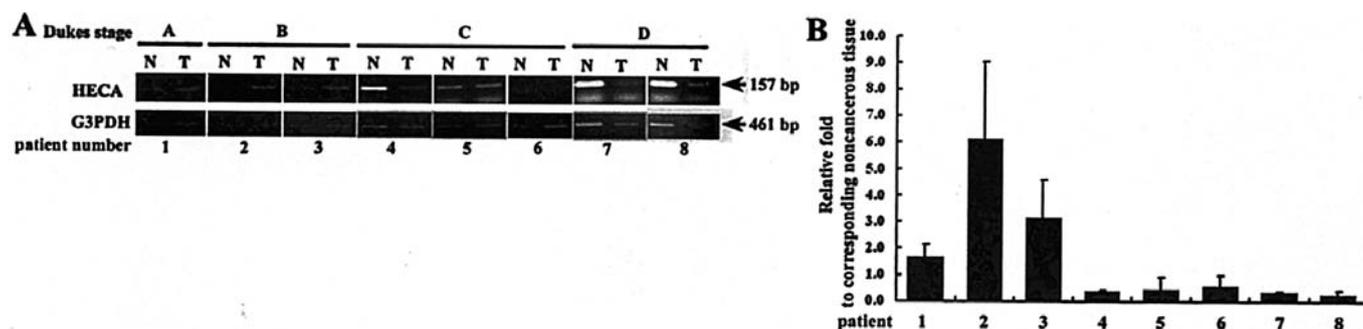


Figure 3. Alteration in the gene expression of HECA in colonic cancerous tissue. (A) Expression of HECA and G3PDH in eight paired colonic tissues from patients with various Dukes' stages. For each pair of tissue (T, cancerous tissue; N, corresponding noncancerous tissue), total RNA (1 μ g) was reverse transcribed using an oligo(dT)12 primer, and aliquots of cDNA were amplified with primers specific to G3PDH (25 cycles) and HECA (25 cycles) after establishing the dilution at which amplification remained within the linear range. Patient 1 was in Dukes' stage A, patients 2 and 3 in Dukes' stage B, patients 4-6 in stage C, and patients 7 and 8 in stage D. The size of the amplified products is shown on the right (bp). (B) Relative quantitation of HECA expression normalized to G3PDH. Triplicate amplifications of HECA and G3PDH for each cell line were analyzed to determine the relative levels of HECA expression. Each patient's HECA:G3PDH ratio of cancerous tissue (T) was relative to data from corresponding noncancerous tissue (N).

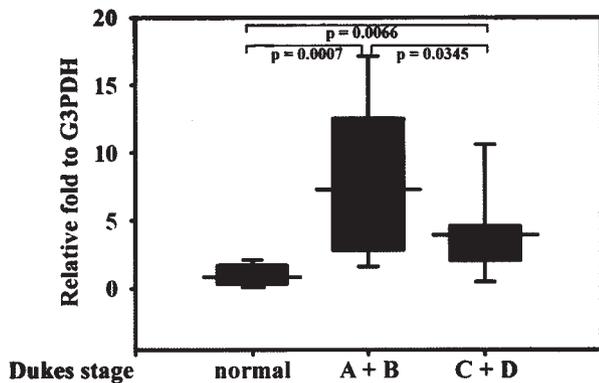


Figure 4. Box plot presentation of HECA expression in blood samples. Expression of HECA in blood samples was determined by RT-PCR analysis from healthy individuals (n=6), Dukes' stage A or B (n=19) and Dukes' stage C or D (n=12) patients, then summarized by box plot. The limits of the box are the 25th and 75th percentiles of the data sets, and the whiskers indicate the 10th and 90th percentiles. The black horizontal line inside the box marks the median value. A significant difference ($p < 0.05$) as determined by the Student's t-test is directly indicated as a p-value.

dilution at which amplification remained within the linear range (Fig. 3). To establish an internal standard for each sample, G3PDH-specific amplifications were also performed with HECA. Relative quantitation of cancerous HECA expression is displayed in Fig. 3. Evaluation of RT-PCR results revealed that the altered expression of HECA could be detected in colorectal cancer tissue from four early-stage patients. They showed a 1.7- to 6.1-fold change in the relative expression of HECA. At Dukes' stages C and D, colonic cancerous tissue had a tendency to decrease the expression of HECA and showed a cancerous:noncancerous ratio of 0.3- to 0.6-fold.

Expression of HECA in blood samples. To validate the upregulated HECA in patients with early-stage colorectal cancer, blood samples from healthy individuals, and Dukes' stage A or B (n=21) and Dukes' stage C or D (n=12) patients were studied by RT-PCR analysis. For evaluation of variation and central tendency of data, this processed data was presented as a box plot (Fig. 4). The boxes are defined by the 25th and 75th percentiles and divided by the median values. The whiskers indicate the 10th and 90th percentiles. HECA was significantly upregulated ($p=0.0019$) in blood samples from all patients (n=34). Patients with Dukes' stage A or B ($p=0.0007$) and Dukes' stage C or D ($p=0.0066$) showed a significant increase in the expression of HECA when compared with those of the healthy individuals. A comparison between the various stages showed that patients with Dukes' stage A or B expressed significantly more HECA in blood than Dukes' stage C or D patients ($p=0.0345$). These findings were in partial agreement with the data from cancer cell lines in which cells at an earlier stage expressed more HECA than cells at a later stage (Fig. 1).

Expression of HECA in feces. Upregulated HECA was also confirmed in feces from patients with colorectal cancer. Fecal total RNA was purified from healthy individuals (n=17), Dukes' stage A or B (n=25) and Dukes' stage C or D (n=19) patients and then analyzed by nested RT-PCR. As

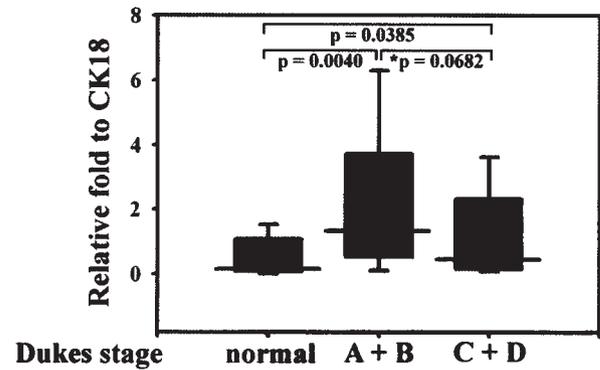


Figure 5. Box plot presentation of HECA expression in fecal samples. Expression of HECA in fecal samples was determined by RT-PCR analysis from healthy individuals (n=17), Dukes' stage A or B (n=25) and Dukes' stage C or D (n=19) patients, then summarized by this box plot. A significant difference ($p < 0.05$) as determined by the Student's t-test is directly indicated as a p-value. Specifics of the box plot are described in Fig. 4. *No statistical significance.

with the data from the blood samples, this processed data was also presented as a box plot (Fig. 5). A significant increase ($p=0.0089$) was shown by comparing healthy individuals (n=17) and patients (n=44). Patients with Dukes' stage A or B ($p=0.0040$) and Dukes' stage C or D ($p=0.0385$) also significantly expressed more HECA in their feces when compared to healthy individuals.

Discussion

Colorectal cancer is a leading cause of malignant death, and better preventive strategies are needed. Research into the possible pathways leading to colorectal cancer has revealed a range of biological intermediates, which could be used to identify high-risk populations and aid the early diagnosis of cancer (26). Further, the evaluation of molecules with significant expression patterns in early- and late-stage tumors should provide useful clinicopathological tools for the management of colorectal cancer (27).

Human feces are a heterogeneous mixture of materials, including luxuriant populations of viable colonocytes (16,28). Because of the clinical availability of feces, we had previously optimized a method to extract fecal total RNA (17). With this study, we were the first to use purified fecal mRNA as the subject for oligonucleotide microarray hybridization. The aim of this study was the high-throughput monitoring of gene expression in exfoliated colonocytes (9,29), quite different to that used previously by Shih *et al.* (30).

Microarrays are an emerging technology in the study of colorectal cancer (30). We previously identified 2248 candidates, including 316 novel proteins, in terms of their significant differential expression between the feces of patients and control subjects (fold change ≥ 2 ; $p < 0.01$) (data not shown). Thus, great strides have been made in analyzing fecal expressed genes in our laboratory and will logically progress into a noninvasive screening tool for colorectal cancer (29).

To stratify results and study uncharacterized candidates, we listed the 45 candidates (21 up- and 22 downregulated) with a 3-fold change. In this study, a gene encoding a homologue

of the *Drosophila* headcase protein (HECA; NM_016217) was chosen from the list of upregulated genes for additional validation. HECA was first isolated and characterized as a human gene homologous to the *Drosophila* headcase gene (25). Headcase protein was found to regulate the re-entry of imaginal cells into the mitotic cycle during adult morphogenesis (31). The gene for HECA encodes a 543 amino acid protein located on the long arm of chromosome 6 (6q23-24). HECA has been suggested to be a tumor suppressor because of its low expression and chromosomal location, being a region of common deletion in human pancreatic cancer (25).

Of the tissue tested, our results suggest that HECA appeared to be normally expressed in many tissues (including colon tissue) but not in that of the ovary. Moreover, we found that the expression of HECA in a healthy colon was not high when compared with other tissues. For this reason, we hypothesized that increased expression may relate to the alteration of HECA in colorectal cancer.

To study the expression levels of HECA in late- and early-stage colorectal cancer, different cell lines were assayed. From the results of relative quantitation of electrophoretic gels, we found that Dukes' stage C LoVo cells expressed less HECA than other late-stage cells including DLD-1 (Dukes' stage C) and two assessed late-stage colorectal cancer cells, Caco-2 and HT29 (32,33). HECA was detectable in our selected colorectal cancer cell lines but with differential expression levels. Two early-stage (Dukes' stage B) cell lines, SW480 and LS174T, and the HCT116 cell line close to Dukes' stage B (33) were shown to have high HECA levels. In this study, we found that only one late-stage cell line, DLD-1, expressed more HECA but the other three late-stage cells (LoVo, Caco-2, and HT29) did not. These findings suggest that HECA may be an early-stage classifier for colorectal cancer and is able to discriminate between late and early stages.

To explore this possibility, we compared the expression patterns of HECA in paired samples of cancerous and corresponding noncancerous colonic tissue from eight patients with a pathological diagnosis of colorectal cancer. The studies on paired tissue from four early-stage patients (Dukes' stage A or B) showed higher levels of HECA expression in cancerous tissue. Such an increase in HECA expression was not found in cancerous tissue from late-stage patients (Dukes' stage C or D). These findings coincided with the results from cell lines that were defined as late stage and relatively expressed less HECA.

To assess the expression of HECA in healthy individuals, we collected 9 blood and 17 feces samples from healthy donors. These samples expressed less HECA. When the patients were subdivided into early- (Dukes' stage A or B) and late-stage (Dukes' stage C or D) groups, both samples from each patient group showed higher expression of HECA than the healthy controls. Moreover, the difference in the HECA expression found in blood samples between early- and late-stage patients was significant. It should be noted that the same trend towards reduced HECA levels from early-to-late stage existed in feces samples, although statistical significance was not achieved ($p=0.0682$).

Results from the analyses of tissue and blood samples suggest that HECA is expressed more often in patients with

colorectal cancer, their clinicopathological stages can be determined by the expression levels of HECA, and the highest expression of HECA is observed in early-stage patients. However, the significance of the altered HECA levels is not clearly understood, nor is it clear whether these increases are a cause or consequence of colorectal cancer. In addition, other than HECA, genes for sulfide dehydrogenase-like (34), coiled-coil and C2 domain containing 1A (CC2D1A) that can activate the NF- κ B and MAPK signaling pathway (35) and the hypothetical proteins, MGC10814 (36) and FLJ22662 (36), were also upregulated and detected in feces. For the purpose of molecular staging and screening in a noninvasive manner, the data for HECA support further evaluation of other novel candidates. The clinical assessment for these differentially expressed genes remains to be undertaken.

In conclusion, the molecular biomarker, HECA, a homologue of the *Drosophila* headcase protein, was detected in feces and blood samples and upregulated in patients with colorectal cancer, and its expression level showed significant correlation with disease status. These results suggest that HECA from feces may be an effective molecular marker for colorectal cancer and disease progression.

Acknowledgements

This work was supported by grant 94CGH-TMU-18 from the Cathay General Hospital and Taipei Medical University.

References

1. Sugimura T, Terada M, Yokota J, Hirohashi S and Wakabayashi K: Multiple genetic alterations in human carcinogenesis. *Environ Health Perspect* 98: 5-12, 1992.
2. Jones PA: DNA methylation and cancer. *Oncogene* 21: 5358-5360, 2002.
3. Fenech M: Biomarkers of genetic damage for cancer epidemiology. *Toxicology* 181-182: 411-416, 2002.
4. Jemal A, Tiwari RC, Murray T, *et al*: Cancer statistics, 2004. *CA Cancer J Clin* 54: 8-29, 2004.
5. Allegra C and Sargent D: Molecular diagnostics: assays, tissues, progress, and pitfalls. *J Clin Oncol* 21: 395-396, 2003.
6. Ahmed FE: Colon cancer: prevalence, screening, gene expression and mutation, and risk factors and assessment. *J Environ Sci Health C Environ Carcinog Ecotoxicol Rev* 21: 65-131, 2003.
7. Birkenkamp-Demtroder K, Christensen LL, Olesen SH, *et al*: Gene expression in colorectal cancer. *Cancer Res* 62: 4352-4363, 2002.
8. Williams NS, Gaynor RB, Scoggin S, *et al*: 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.
9. Osborn NK and Ahlquist DA: Stool screening for colorectal cancer: molecular approaches. *Gastroenterology* 128: 192-206, 2005.
10. Iyengar V, Albaugh GP, Lohani A and Nair PP: Human stools as a source of viable colonic epithelial cells. *FASEB J* 5: 2856-2859, 1991.
11. Huang CJ, Jeng JY, Lee CL, Wu CH, Chen CW and Chien CC: An efficient sandwich enzyme-linked immunosorbent assay (ELISA) detects adenomatous polyposis coli (APC) protein from human stool. *Fu-Jen J Medicine* 2: 225-230, 2004.
12. Lopez-Saucedo C, Cerna JF, Villegas-Sepulveda N, *et al*: Single multiplex polymerase chain reaction to detect diverse loci associated with diarrheagenic *Escherichia coli*. *Emerg Infect Dis* 9: 127-131, 2003.
13. Rasool NB, Monroe SS and Glass RI: Determination of a universal nucleic acid extraction procedure for PCR detection of gastroenteritis viruses in faecal specimens. *J Virol Methods* 100: 1-16, 2002.
14. Alexander RJ and Raicht RF: Purification of total RNA from human stool samples. *Dig Dis Sci* 43: 2652-2658, 1998.

15. Kawada M, Mizuno M, Nasu J, *et al*: Release of decay-accelerating factor into stools of patients with colorectal cancer by means of cleavage at the site of glycosylphosphatidylinositol anchor. *J Lab Clin Med* 142: 306-312, 2003.
16. Kanaoka S, Yoshida K, Miura N, Sugimura H and Kajimura M: Potential usefulness of detecting cyclooxygenase 2 messenger RNA in feces for colorectal cancer screening. *Gastroenterology* 127: 422-427, 2004.
17. Yang SH, Chien CC, Chen CW, Li SY and Huang CJ: Potential of faecal RNA in diagnosing colorectal cancer. *Cancer Lett* 226: 55-63, 2005.
18. Schulze A and Downward J: Navigating gene expression using microarrays – technology review. *Nat Cell Biol* 3: E190-E195, 2001.
19. Lin YM, Furukawa Y, Tsunoda T, Yue CT, Yang KC and Nakamura Y: Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. *Oncogene* 21: 4120-4128, 2002.
20. Mecham BH, Klus GT, Strovel J, *et al*: Sequence-matched probes produce increased cross-platform consistency and more reproducible biological results in microarray-based gene expression measurements. *Nucleic Acids Res* 32: e74, 2004.
21. Jurata LW, Bukhman YV, Charles V, *et al*: Comparison of microarray-based mRNA profiling technologies for identification of psychiatric disease and drug signatures. *J Neurosci Methods* 138: 173-188, 2004.
22. Charbonnier Y, Gettler B, Francois P, *et al*: A generic approach for the design of whole-genome oligoarrays, validated for genotyping, deletion mapping and gene expression analysis on *Staphylococcus aureus*. *BMC Genomics* 6: 95-106, 2005.
23. Blaeser F, Bryce PJ, Ho N, *et al*: Targeted inactivation of the IL-4 receptor alpha chain I4R motif promotes allergic airway inflammation. *J Exp Med* 198: 1189-1200, 2003.
24. Richter TM, Tong BD and Scholnick SB: Epigenetic inactivation and aberrant transcription of CSMD1 in squamous cell carcinoma cell lines. *Cancer Cell Int* 5: 29-39, 2005.
25. Makino N, Yamato T, Inoue H, *et al*: Isolation and characterization of the human gene homologous to the *Drosophila* headcase (hdc) gene in chromosome bands 6q23-q24, a region of common deletion in human pancreatic cancer. *DNA Seq* 11: 547-553, 2001.
26. Garcea G, Sharma RA, Dennison A, Steward WP, Gescher A and Berry DP: Molecular biomarkers of colorectal carcinogenesis and their role in surveillance and early intervention. *Eur J Cancer* 39: 1041-1052, 2003.
27. Koehler A, Bataille F, Schmid C, *et al*: Gene expression profiling of colorectal cancer and metastases divides tumours according to their clinicopathological stage. *J Pathol* 204: 65-74, 2004.
28. Albaugh GP, Iyengar V, Lohani A, Malayeri M, Bala S and Nair PP: Isolation of exfoliated colonic epithelial cells, a novel, non-invasive approach to the study of cellular markers. *Int J Cancer* 52: 347-350, 1992.
29. 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.
30. Shih W, Chetty R and Tsao MS: Expression profiling by microarrays in colorectal cancer (Review). *Oncol Rep* 13: 517-524, 2005.
31. Weaver TA and White RA: headcase, an imaginal specific gene required for adult morphogenesis in *Drosophila melanogaster*. *Development* 121: 4149-4160, 1995.
32. Halder SK, Beauchamp RD and Datta PK: A specific inhibitor of TGF-beta receptor kinase, SB-431542, as a potent antitumor agent for human cancers. *Neoplasia* 7: 509-521, 2005.
33. Van Erk MJ, Krul CA, Caldenhoven E, Stierum RH, Peters WH, Woutersen RA and van Ommen B: Expression profiling of colon cancer cell lines and colon biopsies: towards a screening system for potential cancer-preventive compounds. *Eur J Cancer Prev* 14: 439-457, 2005.
34. Vande Weghe JG and Ow DW: A fission yeast gene for mitochondrial sulfide oxidation. *J Biol Chem* 274: 13250-13257, 1999.
35. Matsuda A, Suzuki Y, Honda G, *et al*: Large-scale identification and characterization of human genes that activate NF-kappaB and MAPK signaling pathways. *Oncogene* 22: 3307-3318, 2003.
36. Strausberg RL, Feingold EA, Grouse LH, *et al*: Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci USA* 99: 16899-16903, 2002.