

# Identification by a Digital Gene Expression Displayer (DGED) and test by RT-PCR analysis of new mRNA candidate markers for colorectal cancer in peripheral blood

MATTIA LAURIOLA<sup>1,2\*</sup>, GIAMPAOLO UGOLINI<sup>3\*</sup>, GIANCARLO ROSATI<sup>3</sup>, SIMONE ZANOTTI<sup>3</sup>,  
ISACCO MONTRONI<sup>3</sup>, ALESSIO MANARESI<sup>3</sup>, DAVIDE ZATTONI<sup>3</sup>, STEFANO RIVETTI<sup>1,2</sup>,  
GABRIELLA MATTEI<sup>1,2</sup>, DOMENICO COPPOLA<sup>4</sup>, PIERLUIGI STRIPPOLI<sup>1,2</sup>,  
MARIO TAFFURELLI<sup>3</sup> and ROSSELLA SOLMI<sup>1,2</sup>

<sup>1</sup>Dipartimento di Istologia, Embriologia e Biologia Applicata, Università di Bologna, Via Belmeloro 8, I-40126 Bologna;

<sup>2</sup>Centro di Ricerca in Genetica Molecolare 'Fondazione CARISBO', Bologna; <sup>3</sup>Dipartimento Emergenza/Urgenza, Chirurgia Generale e dei Trapianti, Università di Bologna, Via Massarenti 9, I-40138 Bologna, Italy;

<sup>4</sup>Department of Anatomic Pathology, H. Lee Moffitt Cancer Center and Research Institute, Magnolia Drive 12902, Tampa, FL 33612, USA

Received April 21, 2010; Accepted May 27, 2010

DOI: 10.3892/ijo\_00000701

**Abstract.** Evidence from the literature widely supports the efficacy of screening for colorectal cancer (CRC) in reducing mortality. A blood-based assay, potentially, represents a more accessible early detection tool for the identification of circulating tumour cells originating from a primary tumour site in the body. The present work aimed at identifying a set of specific mRNAs expressed in colon tissue but not in blood cells. These mRNAs may represent useful markers for early detection of circulating colon cancer cells by a simple, qualitative RT-PCR assay, following RNA extraction from peripheral blood samples. Using a data-mining tool called cDNA digital gene expression displayer (DGED), based on serial analysis of gene expression (SAGE) from the Cancer Genome Anatomy Project (CGAP) database, 4-colon and 14-blood cDNA libraries were analyzed. We selected 7 genes expressed in colon tissue but not in blood and were able to test 6 of them by RT-PCR in peripheral blood of CRC patients and healthy controls. We present a relatively easy and highly reproducible technique for the detection of mRNA expression of genes as candidate markers of malignancy in blood samples of patients with colon cancer. SAGE DGED provided a list of the best candidate mRNAs predicted to detect colon cells in

the blood, namely those encoding the following proteins: hypothetical protein LOC644844 (*LOC644844*, whose cDNA was not amplifiable), fatty acid binding protein 1 (*FABP1*), carcinoembryonic antigen-related cell adhesion molecule 5 (*CEACAM5*), mucin 13 cell surface associated (*MUC13*), guanylate cyclase activator 2A (*GUCA2A*), amiloride binding protein 1 (*ABP1*), galactoside-binding, solute carrier family 26, member 3 (*SLC26A3*). The mRNA expression of these genes was evaluated in 8 samples from subjects diagnosed with CRC and 9 from healthy controls. We observed the expression of 2 of the 6 investigated genes in the blood samples of the vast majority of patients considered, but also in a subset of the controls. Our data confirm the extreme sensitivity of RT-PCR, making this technique able to detect minimal amounts of mRNA expressed in a non-tissue-specific manner. Moreover, DGED remains a powerful tool to identify candidate epithelial markers in blood, such as colon related mRNAs. However, to date, none of these qualified as tumour markers.

## Introduction

Colorectal cancer remains a leading cause of mortality worldwide despite the well-characterized molecular events in the adenoma-to-carcinoma sequence (1). Screening for CRC has been shown to reduce disease incidence and cancer mortality. Several screening methods are available, but none is ideal (2).

The diagnostic stage determines survival but, unfortunately, many colorectal cancers are recognized after regional or distant metastasis has occurred (3). Treatment options for the advanced stages of the disease are limited, but favorable chances are greatly increased in early disease, therefore, the most effective means of reducing colorectal cancer mortality is early detection. Current modalities for screening, which include fecal occult blood testing (FOBT), flexible sigmoidoscopy and colonoscopy, have been available for decades, but their

---

Correspondence to: Dr Rossella Solmi, Dipartimento di Istologia, Embriologia e Biologia Applicata, Università di Bologna, Via Belmeloro 8, I-40126 Bologna, Italy  
E-mail: rossella.solmi@unibo.it

\*Contributed equally

**Key words:** digital gene expression display, circulating tumor cells, colorectal cancer

use in the general population is far below the recommended levels. Colonoscopy is currently the most effective means of detecting colonic lesions. Patient discomfort and distaste for this procedure, costs, lack of physician training and equipment, are some of the obstacles to the wide dissemination of this screening test (2).

A simple blood test capable of early CRC detection and/or of monitoring CRC progression would mean significant reduction of costs and higher compliance from the CRC patient population, often reluctant to undergo colonoscopy. A blood-based assay for the detection of circulating tumour cells identified by a selected panel of CRC specific mRNAs potentially represents a more accessible screening tool.

Modest improvements in patient survival from CRC have been achieved in the last few decades, and the overall 5-year survival rate remains at 40-45%. Surgical resection is the treatment of choice for CRC; however, nearly one-half of all patients who undergo a potentially curative resection will relapse because of undetected micrometastasis. This suggests that tumour cell dissemination occurs early in the disease process and emphasizes the need for easily accessible diagnostic methods with sufficient sensitivity and specificity (3). The spontaneous circulation of tumour cells and/or tumour microemboli is the hallmark of the 'invasive and metastatic behaviour' of a subset of cancer cells. Their detection is expected to provide a powerful tool for cancer prognosis, diagnosis of minimal residual disease, assessment of tumour sensitivity to anticancer drugs, and personalization of anti-cancer therapy (4,5). A highly sensitive and specific identification of circulating tumour cells (CTC) and circulating tumour microemboli (CTM) could also be helpful for an early diagnosis of invasive cancers (6).

From simulation models, it has been estimated that from each gram of tumour tissue around  $1 \times 10^6$  tumour cells can be introduced daily into the bloodstream (7). Solid tumours as small as 2 mm in diameter typically display active angiogenesis and are hence capable of releasing tumour cells into peripheral blood. In the earlier stages of disease, disseminated cells are not capable of forming metastases, but they may represent a source for cancer detection.

Epithelial cancer cells have very low survival rates into the circulation (8). The fate of CTC includes a rapid phase of intravascular cancer cell disappearance, which is completed in <5 min and accounts for 85% of the circulating cancer cells (9,10).

This process has been related to 'anoikis' (apoptosis induced by the disruption of cell attachment and cell-matrix interactions processes). Many cancer cell types with increased metastatic potential are resistant to anoikis compared to their parental cells, and the tumour phenotype appears to be related to the expression of apoptosis inhibitors (9). Based on evidence from animal studies, where tumour local invasion and intravasation are bypassed (because of technical difficulties) and tumour cells are directly introduced into the systemic circulation, it has been established that around 1/40 CTC give rise to micrometastases and only ~0.01% proliferates into macrometastasis (11).

The number of tumour circulating cells may be very small; therefore, methods for detecting them need to be both sensitive

and specific. RT-PCR is potentially a very sensitive technique that has been proven to detect minute amounts of specific mRNA species in the total amount of RNA. The success of this technique depends on the availability of a specific target, consistently distinguishing tumour cells from blood cells (12).

Detection of circulating neoplastic cells in patients with CRC has hitherto been carried out using a small number of different RNA targets. Carcinoembryonic antigen (CEA) has been chosen as a target (12-16) since this cell surface molecule is constitutively expressed in the colonic epithelial tissue (12) and its expression is maintained in almost all colorectal carcinomas. Cytokeratins, a multigene family of proteins with differentiation-associated patterns of expression, have also been used to characterize neoplastic cells of epithelial origin in the peripheral blood (15,17-21). Epidermal growth factor receptor (EGFR) mRNA is an alleged specific marker of circulating epithelial cells (16,22,23). Furthermore, Guanylate cyclase C (24), tumour splice variants of CD44 (25), MUC1 and MUC2 (26), matrilysin MMP7 and HeLa metastatic gene (HLM) (22), Apolipoprotein A-I (27), and MAGE (28) mRNAs have also been tested for the same purpose. However, in most of these studies the expression of the targeted mRNAs was also detected in the peripheral blood of control subjects, free of neoplastic disease. In other cases, low sensitivity has hampered the usefulness of the assay.

At present, therefore, no available assay allows for early molecular detection of CRC in clinical practice. The aim of our research was to identify new potential molecular markers of CRC cell detection in peripheral blood, using bioinformatics investigation tools. This approach was used to recognize an optimal subset of genes suitable for early and highly specific detection of circulating CRC cells by a simple, qualitative reverse transcriptase-polymerase chain reaction assay. Our results include the identification of new colon-specific genes.

## Patients and methods

**Patients.** The study was approved by the ethics committee of Sant'Orsola-Malpighi General Hospital and complied with the Ethical Principles for Medical Research Involving Human Subjects of the Helsinki Declaration. Written, informed consent was obtained from all patients and healthy control before study entry. Eight unrelated patients with a histologically confirmed diagnosis of CRC and 9 healthy subjects were enrolled in the study. Peripheral blood samples were obtained from each patient and control. The main clinical data of the patients are summarized in Table I.

**RNA sources.** A peripheral blood sample (5 ml) was obtained from each patient just before surgery. To reduce contamination of samples with skin epithelial cells from the needle stick, the first ml of blood was discarded. The whole blood was then drawn into an EDTA tube and RNA extracted within one hour of being drawn, as described below. In addition, peripheral blood samples from 9 healthy donors were collected and processed with the same method and used as control. Total RNA samples from CRC commercially obtained from Stratagene were used as an additional positive control.

Table I. The main clinical data of the patients.

Subject	Age	Sex	Therapy	Stage <sup>a</sup>
CRC1	80	F	RH	III (T3 N1 Mx)
CRC2	73	M	RH	Adenoma <sup>b</sup>
CRC3	84	F	RH	II (T3 N0 Mx)
CRC4	86	F	RH	II (T3 N0 Mx)
CRC5	81	M	RR	III (T3 N2 Mx)
CRC6	73	M	NO	IV (Tx Nx M1)
CRC7	71	M	LH	IV(T3 N0 M1)
CRC8	44	F	NO	IV(T3 N0 M1)
N1	72	M		
N2	35	M		
N3	56	F		
N4	41	F		
N5	64	M		
N6	44	F		
N7	35	M		
N8	62	F		
N9	65	M		

CRC, patients with colorectal cancer; N, normal subject. <sup>a</sup>Staging according to the American Joint Committee on Cancer (AJCC) and International Union Against Cancer. <sup>b</sup>Adenoma, low-grade tubulo-villous adenoma. NO, not operable; RR, rectal resection; RH, right hemi-colectomy; LH, left hemicolectomy.

**Identification of potential molecular markers.** The identification of potential molecular markers for CRC cell detection was performed using a systemic approach by bioinformatics tools. SAGE DGED (<http://cgap.nci.nih.gov/SAGE/>) is an original computer approach for statistical analysis of

differential gene expression patterns that identifies those genes expressed at significantly different levels (as defined by the user) in two pools of human cDNA libraries, based on SAGE (29) tag analysis. The algorithm takes into account the differences in sample size between Pools A and B, which can be large. The user selects a value for statistical significance (P-value) and a value for the difference in the level of expression (F-value) between the two pools. The results are based on the sequence odds ratio and measures of significance.

For this study two pools were created and analyzed, the first encompassing all cDNA libraries from colon tissue (normal as well as neoplastic and tumour associated), while the second included all cDNA libraries from white blood cells (normal as well as neoplastic and tumour associated), to highlight colon-specific genes, having ideally a level of expression equal to zero in the blood. The formula for the sequence odds ratio is: (Seqs A / Total Sequences in A) / (Seqs B / Total Sequences in B), where A and B are the selected pools. We selected genes with more than 100 tags for colon libraries (Table II).

**RNA extraction.** Blood samples were lysed within one hour of being drawn, diluting 1 ml of whole blood with 1 ml of DEPC water and adding 6 ml TRIzol Ls (Invitrogen). Then, 6 ml of lysed cells was added to the sample suspension, passing the suspension several times through a pipette.

The lysate was stored at -80°C until RNA extraction. Extraction was performed according to the manufacturer's instructions. Total RNA extracted from 1 ml of peripheral blood was then subjected to standard ethanol precipitation, and the pellet was resuspended in sterile bidistilled water to a final concentration of up to 0.5 µg/µl, 1 U of RNAase enzyme was added, and then the sample was stored at -20°C.

**Reverse transcription-polymerase chain reaction (RT-PCR) amplification.** Total RNA from peripheral blood (~1 µg) or total RNA from commercial human colon was reverse transcribed into a final volume of 25 µl in the presence of 500 µM each of deoxynucleotide triphosphate (dNTP), 200 units reverse transcriptase (Murine Moloney leukemia virus) with

Table II. Candidate markers selected by DGED.

Tag	Gene	Libraries		Tag		P-value
		A	B	A	B	
GCCCAGGTCA	<i>LOC644844</i>	4	0	1046	0	0.00
ACATTGGGTG	<i>FABP1</i>	4	0	840	0	0.00
ATTGGAGTGC	<i>CEACAM5</i>	4	0	277	0	0.00
CCCTCCCGAA	<i>MUC13</i>	4	0	168	0	0.00
ATACTCCACT	<i>GUCA2A</i>	4	0	154	0	0.00
GCCGACCAGG	<i>ABP1</i>	4	0	121	0	0.00
CATAGGTTTA	<i>SLC26A3</i>	3	0	110	0	0.00

Total tags in Pool A, 195,160; total tags in Pool B, 946,843; total libraries in Pool A, 4 (colon); total libraries in Pool B, 14 (white blood cells); F (expression factor) 2x ; P=0.05 (significance filter).

Table III. The primers used.

Primer	Sequences 5'→3'
<i>CEACAM5</i> L	CACCCTGGATGTCCTCTATGGGCCG
<i>CEACAM5</i> R	TGGCCCCAGCTGAGAGACCAGGA
<i>ABP1</i> L	CCGTGGTCTTTGAGCAGTTT
<i>ABP1</i> R	GCCACACGATCACAGTGTCT
<i>SLC26A3</i> L	CTGGCGCCTCTACAAAAGTC
<i>SLC26A3</i> R	ACACTAGCTGCCAGGCCTAA
<i>MUC13</i> L	TCCCTGGGAAGATTTTCAGTG
<i>MUC13</i> R	AACCATTGAGGCAGTCATCC
<i>FABP1</i> L	GCAGAGCCAGGAAAACCTTTG
<i>FABP1</i> R	TCTCCCCTGTCATTGTCTCC
<i>GUCA2A</i> L	GGGGTCACCGTGCAGGATGGAA
<i>GUCA2A</i> R	GTCCCTGCTGCGGAGGGGAGG
<i>B2M</i> L	GTGGGATCGAGACATGTAAGCAGC
<i>B2M</i> R	CTCCTAGAGCTACCTGTGGAGCAA

L, left (forward) primer; R, right (reverse) primer. Source sequence for primer design were: #AB021288 (*B2M*), #NM\_004363.2 (*CEACAM5*), #NM\_001091.2 (*ABP1*), #NM\_000111.2 (*SLC26A3*), #NM\_033049.2 (*MUC13*), #NM\_001443 (*FABP1*), #NM\_033553.2 (*GUCA2A*).

companion buffer 1X (Promega, Madison, WI, USA) and oligo dT-15 5  $\mu$ M. The incubation was performed at 42°C for 1 h, followed by denaturation at 95°C for 5 min. The primers for amplification were designed using the software Amplify, following standard criteria. Primer sequences are given in Table III. Primers were designed on different exons for each gene. PCR experiments were performed in a final volume of 25  $\mu$ l, containing 10  $\mu$ l RT mix, 1 U Taq polymerase (Takara, Shiga, Japan) with companion reagents (0.2 mM each dNTP, 1.5 mM MgCl<sub>2</sub>, 1X PCR buffer), and 0.4 mM of each primer. An initial denaturation step of 2 min at 94°C was followed by amplification for 40 cycles (30 sec at 94°C, 30 sec at 61°C, 30 sec at 72°C) and final extension for 7 min at 72°C. All the RT-PCR products obtained were analysed by gel electrophoresis following standard methods (30): ethidium bromide incorporation in the gel (0.5  $\mu$ g/ml final concentration), and photography by GelDoc 2000 system (Bio-Rad, Hercules, CA).

Each potential marker was tested by RT-PCR on 8 CRC patients, 9 healthy controls, and on 1  $\mu$ g of RNA from CRC tissue from Stratogene. In some cases, genes expressed in at least three normal subjects were not further analyzed, as they were obviously unsuitable as cancer markers. *B2M* was used as positive control of RT-PCR product.

## Results

**Identification of potential molecular markers.** Using a systematic bioinformatics approach we were able to identify

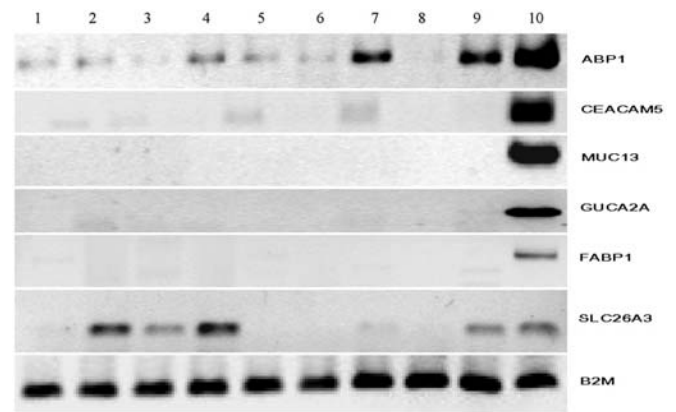


Figure 1. Candidate marker expression in blood of healthy subjects. Expression of *ABP1*, *CEACAM5*, *MUC13*, *GUCA2A*, *FABP1* and *SLC26A3*. *B2M* was used as positive control. Ethidium-bromide stained 1.5% agarose gel of RT-PCR amplification products. Lanes 1-9, peripheral blood of healthy subjects; lane 10, positive control (CRC tissue from Stratogene).

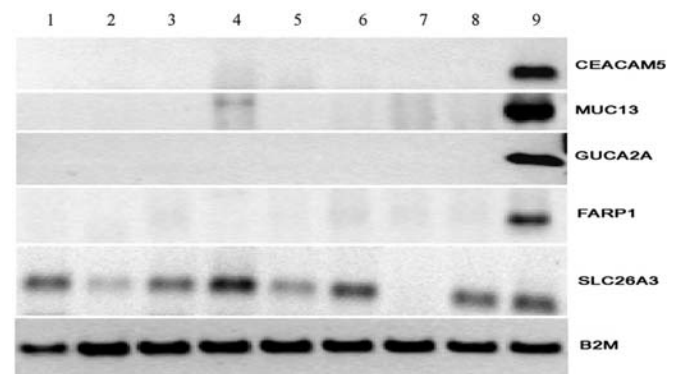


Figure 2. Candidate marker expression in blood of CRC patients. Expression of *CEACAM5*, *MUC13*, *GUCA2A*, *FABP1*, *SLC26A3*. *B2M* was used as a positive control. Ethidium-bromide stained 1.5% agarose gel of RT-PCR amplification products. Lanes 1-8, CRC patients peripheral blood; lane 9, positive control (CRC tissue from Stratogene).

a first series of potential molecular markers for CTC detection. In particular, DGED analysis produced a list of genes that are differentially expressed ( $P < 0.001$ ) between a pool of 4 colon libraries and a pool of 14 blood libraries. This powerful tool was able to compare 195,160 tags from colon tissue to 946,843 tags from blood. We selected the best genes significantly expressed in the colon (with a fraction of tags present in the colon pool  $> 100$ ) and virtually not expressed in the blood (with a fraction of tags present in the blood pool equal to 0.00000), as shown in Table II. The genes selected were: *LOC644844*, *FABP1*, *CEACAM5*, *MUC13*, *GUCA2A*, *ABP1*, and *SLC26A3*.

**Reverse transcription-polymerase chain reaction (RT-PCR) amplification.** RT-PCR products of the expected size of 6 of the 7 candidates markers were obtained from all 9 healthy samples and 5 of 7 were also obtained from the CRC samples. The genes were tested in healthy controls (Fig. 1) and in patients (Fig. 2), and a clear cut-off between the two groups



was not detected. *ABP1*, *CEACAM5*, *MUC13*, *GUCA2A*, *FABP1* and *SLC26A3* were first tested on 9 normal subjects.

*ABP1* was found to be clearly positive in 7 of the 9 control subjects while *SLC26A3* was positive in 4.

*CEACAM5*, *MUC13*, *GUCA2A* appeared to be favorable candidate colon markers because they were almost absent in the normal subjects tested and were highly expressed in CRC tissue from Stratagene. However, no positivity for these markers was found in the patient blood samples (Fig. 2).

*FABP1* was also negative in normal blood, but its expression level in CRC tissue from Stratagene was too low to qualify as an effective marker, and it was also negative in the blood samples of CRC patients. We were not able to amplify *LOC644844* cDNA, as its expression value was too low also in the positive control (data not shown). Therefore, this gene was removed from the study.

## Discussion

Early detection is one of the most effective means of reducing CRC related mortality. According to the World Health Organization, the criteria for developing and/or applying a screening test for cancer include acceptability, practicability, high specificity and high sensitivity (31).

The RT-PCR assay for detection of circulating tumour cells in peripheral blood potentially represents a sensitive and accessible screening tool, able to detect minute amounts of specific RNA species amid the total amount of RNA. However, this technique has been, to date, unsuccessfully employed, due to the difficulty in identifying specific molecular markers able to distinguish CRC cells from blood cells.

The candidate genes, identified without any *a priori* knowledge in the present study as markers of colonic circulating cells in the peripheral blood of patients with different stages CRC, have already been reported to be specific for intestinal expression. For example, the *MUC13* is a member of genes encoding complex mucosal glycoproteins that can be highly expressed by adenocarcinomas. These glycoproteins have diagnostic, therapeutic, and biological significance. *MUC13* encodes a cell surface membrane-anchored mucin, expressed in the normal gastrointestinal tract and trachea as well as in colorectal, esophageal, gastric, pancreatic, and lung cancers. *MUC13* protein expression was determined immunohisto-chemically in 99 sporadic CRC, by assessing the proportion of tumour cells stained, the staining intensity, and their localization. In normal colon, intense apical membrane and variable cytoplasmic *MUC13* staining was present in both goblet and columnar cells, with strongest reactivity in the upper crypts and surface epithelium. *MUC13* was expressed abundantly in all colorectal cancers, with the highest expression seen in poorly differentiated tumours (32). *GUCA2A* is a bioactive intestinal peptide involved in the cystic fibrosis transmembrane conductance regulator (CFTR)-regulated electrolyte/water secretion. Interestingly, a previous study, using microarray analysis, found *SLC26A3* to be downregulated in colon cancer tissue in comparison to normal tissue. The authors suggested that *SLC26A3* might function as a tumour suppressor gene in CRC (33).

We expected that, within our highly selected set of colon tissue specific genes, we could identify genes with a cut-off

able to distinguish control subjects and patients with colon cancer at different stages. Surprisingly, 2 of the 6 genes studied (*ABP1* and *SLC26A3*) were expressed in the vast majority of CRC patients considered, but also in most of the control subjects. Some of the genes analysed had previously been considered epithelial-specific. This finding confirms that the task of identifying a useful mRNA cancer marker is not trivial, and that the extreme sensitivity of RT-PCR is able to detect minimal amounts of mRNA expressed in a non-tissue-specific manner.

The concept that any gene may be expressed in any cell type, albeit at a low level, is known as 'illegitimate transcription'. It has been speculated that gene expression in the absence of specific transcription factors may be very low, but not equal to zero. Expression of DNA that is part of condensed chromatin may be impossible under most conditions, but these inactive states of DNA may change when replication is taking place. Although these expression levels are very low (it is estimated that one transcript of any gene is present in any collection of 500-1,000 cells), this amount is however detectable by RT-PCR (33,34). These results are similar to those previously obtained by our group (35), for another set of candidate markers, identified by expressed sequence tag (EST) based approach, rather than by SAGE DGED.

However, the finding that the expression of *MUC13* and *FABP1* is detectable in human colon tissues but not in blood samples rules out the hypothesis that illegitimate transcription is a universal phenomenon, which would exclude all possibility of finding a useful mRNA marker (36).

Interestingly, we were able to identify by DGED 2 candidates, *CEACAM5* and *FABP1*, that were already described as CTC associated genes by gene expression profiling, and specifically up-regulated in the metastatic CRC patient's blood (37). The same authors selected *SLC2A10* as CTC associate gene, belonging to the same family of our candidate, *SLC26A3*, which seems to be a useful marker because it is localized in the mucosa of the lower intestinal tract, particularly in the apical membrane of columnar epithelium.

Our research is the first one testing *ABP1*, *GUCA2A* and *MUC13* as blood marker in CRC patients. *MUC13* seemed to be promising candidate because is known as strong differentiation marker for gastrointestinal mucosa. Differently from other authors, that used *MUC1* and *MUC2* as blood markers, our member of the mucin family seems to be more specific for colon rather than epithelial tissue (32).

Alternative approaches to circumvent the need for a clear cut-off between positive and negative samples could be quantitative RT-PCR, or the use of a combination of several markers to achieve statistical significance. The consistency of results obtained by quantitative RT-PCR among different laboratories is often questioned, and it is known that this test is commonly difficult to optimize and interpret. Assays incorporating a combination of multiple markers would be a reasonable approach, although we feel that the existence of a major locus functioning as a master marker could be a crucial feature for the applicability of the test.

In conclusion, to date the search for a useful mRNA marker for detection of colon cancer cells in peripheral blood has involved repeated trials often limited to the same small set of genes. By employing a systematic, open approach such

as the one described here, there is a concrete possibility to find few genes or at least a single major locus suitable for diagnostic RT-PCR, based on a total ~30,000 genes available in the human genome. DGED is a powerful tool to identify differentially expressed candidate genes.

The test of a marker is best conducted in patients with clearly diagnosed colon cancer, and is possibly the basis for the development of a further prospective study targeting early diagnosis in population screening. However, among the gene set tested in this study, there was no clear cut-off between healthy subjects and patients with colon cancer. Further investigations, aimed at selecting the best candidate markers using large-scale approaches and at validating them with a simple blood test by RT-PCR, are required.

## Acknowledgements

We are grateful to Dr Maria Carmen Biffoni and Dr Daniela Solmi (Laboratorio Analisi Cliniche S. Antonio - Bologna) for the precious collaboration in supplying blood samples from healthy donors. This work was supported by grants from Fondazione Elio Bisulli (Cesena, Italy) and Fondazione Enzo Piccinini (Modena, Italy).

## References

- Huerta S: Recent advances in the molecular diagnosis and prognosis of colorectal cancer. *Expert Rev Mol Diagn* 8: 277-288, 2008.
- Kronborg O and Regula J: Population screening for colorectal cancer: advantages and drawbacks. *Dig Dis* 25: 270-273, 2007.
- Luzzi KJ, MacDonald IC, Schmidt EE, Kerkvliet N, Morris VL, Chambers AF and Groom AC: Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *Am J Pathol* 153: 865-873, 1998.
- Mocellin S, Keilholz U, Rossi CR and Nitti D: Circulating tumor cells: the 'leukemic phase' of solid cancers. *Trends Mol Med* 12: 130-139, 2006.
- Paterlini-Brechot P and Benali NL: Circulating tumor cells (CTC) detection: clinical impact and future directions. *Cancer Lett* 253: 180-204, 2007.
- Lambrechts AC, Vanit Veer LJ and Rodenhuis S: The detection of minimal numbers of contaminating epithelial tumor cells in blood or bone marrow: use, limitations and future of RNA-based methods. *Ann Oncol* 9: 1269-1276, 1998.
- Chang YS, di Tomaso E, McDonald DM, Jones R, Jain RK and Munn LL: Mosaic blood vessels in tumors: frequency of cancer cells in contact with flowing blood. *Proc Natl Acad Sci USA* 97: 14608-14613, 2000.
- Glinsky VV, Glinsky GV, Glinskii OV, Huxley VH, Turk JR, Mossine VV, Deutscher SL, Pienta KJ and Quinn TP: Intravascular metastatic cancer cell homotypic aggregation at the sites of primary attachment to the endothelium. *Cancer Res* 63: 3805-3811, 2003.
- Berezovskaya O, Schimmer AD, Glinskii AB, Pinilla C, Hoffman RM, Reed JC and Glinsky GV: Increased expression of apoptosis inhibitor protein XIAP contributes to anoikis resistance of circulating human prostate cancer metastasis precursor cells. *Cancer Res* 65: 2378-2386, 2005.
- Heyder C, Gloria-Maercker E, Hatzmann W, Zaenker KS and Dittmar T: Visualization of tumor cell extravasation. *Contrib Microbiol* 13: 200-208, 2006.
- Yang J, Mani SA and Weinberg RA: Exploring a new twist on tumor metastasis. *Cancer Res* 66: 4549-4552, 2006.
- Jonas S, Windeatt S, O-Boateng A, Fordy C and Allen-Mersh TG: Identification of carcinoembryonic antigen-producing cells circulating in the blood of patients with colorectal carcinoma by reverse transcriptase polymerase chain reaction. *Gut* 39: 717-721, 1996.
- Castells A, Boix L, Bessa X, Gargallo L and Pique JM: Detection of colonic cells in peripheral blood of colorectal cancer patients by means of reverse transcriptase and polymerase chain reaction. *Br J Cancer* 78: 1368-1372, 1998.
- Guadagni F, Kantor J and Aloe S: Detection of blood-borne cells in colorectal cancer patients by nested reverse transcription-polymerase chain reaction for carcinoembryonic antigen messenger RNA: longitudinal analyses and demonstration of its potential importance as an adjunct to multiple serum markers. *Cancer Res* 61: 2523-2532, 2001.
- Huang P, Wang J, Guo Y and Xie W: Molecular detection of disseminated tumor cells in the peripheral blood in patients with gastrointestinal cancer. *J Cancer Res Clin Oncol* 129: 192-198, 2003.
- Vlems FA, Diepstra JH and Cornelissen IM: Investigations for a multi-marker RT-PCR to improve sensitivity of disseminated tumor cell detection. *Anticancer Res* 23: 179-186, 2003.
- Burchill SA, Bradbury MF, Pittman K, Southgate J, Smith B and Selby P: Detection of epithelial cancer cells in peripheral blood by reverse transcriptase-polymerase chain reaction. *Br J Cancer* 71: 278-281, 1995.
- Wyld DK, Selby P, Perren TJ, Jonas SK, Allen-Mersh TG, Wheeldon J and Burchill SA: Detection of colorectal cancer cells in peripheral blood by reverse-transcriptase polymerase chain reaction for cytokeratin 20. *Int J Cancer* 79: 288-293, 1998.
- Chausovsky G, Luchansky M and Figer A: Expression of cytokeratin 20 in the blood of patients with disseminated carcinoma of the pancreas, colon, stomach, and lung. *Cancer* 86: 2398-2405, 1999.
- Gradilone A, Gazzaniga P and Silvestri I: Detection of CK19, CK20 and EGFR mRNAs in peripheral blood of carcinoma patients: correlation with clinical stage of disease. *Oncol Rep* 10: 217-222, 2003.
- Vlems FA, Diepstra JH and Cornelissen IM: Limitations of cytokeratin 20 RT-PCR to detect disseminated tumour cells in blood and bone marrow of patients with colorectal cancer: expression in controls and downregulation in tumour tissue. *Mol Pathol* 55: 156-163, 2002.
- Clarke LE, Leitzel K, Smith J, Ali SM and Lipton A: Epidermal growth factor receptor mRNA in peripheral blood of patients with pancreatic, lung, and colon carcinomas detected by RT-PCR. *Int J Oncol* 22: 425-430, 2003.
- De Luca A, Pignata S, Casamassimi A, DiAntonio A, Gridelli C, Rossi A, Cremona F, Parisi V, De Matteis A and Normanno N: Detection of circulating tumor cells in carcinoma patients by a novel epidermal growth factor receptor reverse transcription-PCR assay. *Clin Cancer Res* 6: 1439-1444, 2000.
- Bustin SA, Gyselman VG, Williams NS and Dorudi S: Detection of cytokeratins 19/20 and guanylyl cyclase C in peripheral blood of colorectal cancer patients. *Br J Cancer* 79: 1813-1820, 1999.
- Wong LS, Cantrill JE, Morris AG and Fraser IA: Expression of CD44 splice variants in colorectal cancer. *Br J Surg* 84: 363-367, 1997.
- Hardingham JE, Hewett PJ, Sage RE, Finch JL, Nuttall JD, Kotasek D and Dobrovic A: Molecular detection of blood-borne epithelial cells in colorectal cancer patients and in patients with benign bowel disease. *Int J Cancer* 89: 8-13, 2000.
- Normanno N, De Luca A and Castaldo A: Apolipoprotein A-I reverse transcriptase-polymerase chain reaction analysis for detection of hematogenous colon cancer dissemination. *Int J Oncol* 13: 443-447, 1998.
- Miyashiro I, Kuo C and Huynh K: Molecular strategy for detecting metastatic cancers with use of multiple tumor-specific MAGE-A genes. *Clin Chem* 47: 505-512, 2001.
- Velculescu VE, Zhang L, Vogelstein B and Kinzler KW: Serial analysis of gene expression. *Science* 270: 484-487, 1995.
- Davis LG, Kuehl WM and Battey JF: *Basic Methods in Molecular Biology*. 2nd edition. Appleton & Lange, Norwalk, 1994.
- Therasse P and European Organisation for Research and Treatment of Cancer Data Center: Evaluation of response: new and standard criteria. *Ann Oncol* 4: 127-129, 2002.
- Walsh MD, Young JP, Leggett BA, Williams SH, Jass JR and McGuckin MA: The MUC13 cell surface mucin is highly expressed by human colorectal carcinomas. *Hum Pathol* 38: 883-892, 2007.
- Mlakar V, Berginc G, Volavsek M, Stor Z, Rems M and Glavac D: Presence of activating KRAS mutations correlates significantly with expression of tumour suppressor genes DCN and TPM1 in colorectal cancer. *BMC Cancer* 9: 282, 2009.

34. Chelly J, Kaplan JC, Maire P, Gautron S and Kahn A: Transcription of the dystrophin gene in human muscle and non-muscle tissue. *Nature* 333: 858-860, 1988.
35. Solmi R, De Sanctis P, Zucchini C, Ugolini G, Rosati G, Del Governatore M, Coppola D, Yeatman TJ, Lenzi L, Caira A, Zanotti S, Taffurelli M, Carinci P, Valvassori L and Strippoli P: Search for epithelial-specific mRNAs in peripheral blood of patients with colon cancer by RT-PCR. *Int J Oncol* 25: 1049-1056, 2004.
36. Chelly J, Concordet JP, Kaplan JC and Kahn A: Illegitimate transcription: transcription of any gene in any cell type. *Proc Natl Acad Sci USA* 86: 2617-2621, 1989.
37. Smirnov DA, Zweitzig DR, Foulk BW, Miller MC, Doyle GV, Pienta KJ, Meropol NJ, Weiner LM, Cohen SJ, Moreno JG, Connelly MC, Terstappen LWMM and O'Hara SM: Global gene expression profiling of circulating tumor cells. *Cancer Res* 65: 4993-4997, 2005.