Systematic identification and validation of candidate genes for detection of circulating tumor cells in peripheral blood specimens of colorectal cancer patients

PETER FINDEISEN[1], MATTHIAS RÖCKEL[1], MATTHIAS NEES[2,5], CHRISTIAN RÖDER[3], PETER KIENLE[4], MAGNUS VON KNEBEL DOEBERITZ[2], HOLGER KALTHOFF[3] and MICHAEL NEUMAIER[1]

[1]Institute for Clinical Chemistry, Medical Faculty Mannheim of the University of Heidelberg; [2]Department of Applied Tumor Biology, Institute of Pathology, University of Heidelberg; [3]Department of General Surgery and Thoracic Surgery, University Hospital Schleswig-Holstein, Campus Kiel, Kiel; [4]Department of Surgery, Medical Faculty Mannheim of the University of Heidelberg, Germany

Received May 14, 2008; Accepted July 30, 2008

DOI: 10.3892/ijo_00000088

Abstract. The presence of tumor cells in peripheral blood is being regarded increasingly as a clinically relevant prognostic factor for colorectal cancer patients. Current molecular methods are very sensitive but due to low specificity their diagnostic value is limited. This study was undertaken in order to systematically identify and validate new colorectal cancer (CRC) marker genes for improved detection of minimal residual disease in peripheral blood mononuclear cells of colorectal cancer patients. Marker genes with upregulated gene expression in colorectal cancer tissue and cell lines were identified using microarray experiments and publicly available gene expression data. A systematic iterative approach was used to reduce a set of 346 candidate genes, reportedly associated with CRC to a selection of candidate genes that were then further validated by relative quantitative real-time RT-PCR. Analytical sensitivity of RT-PCR assays was determined by spiking experiments with CRC cells. Diagnostic sensitivity as well as specificity was tested on a control group consisting of 18 CRC patients compared to 12 individuals without malignant disease. From a total of 346-screened genes only serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5 (*SERPINB5*) showed significantly elevated transcript levels in peripheral venous blood specimens of tumor patients when compared to the nonmalignant control group. These results were confirmed by analysis of an enlarged collective consisting of 63 CRC patients and 36 control individuals without malignant disease. In conclusion *SERPINB5* seems to be a promising marker for detection of circulating tumor cells in peripheral blood of colorectal cancer patients.

Correspondence to: Dr Peter Findeisen, Institute for Clinical Chemistry, Medical Faculty Mannheim of the University of Heidelberg, Theodor Kutzer Ufer 1-3, 68167 Mannheim, Germany
E-mail: peter.findeisen@ikc.ma.uni-heidelberg.de

Present address: [5]Medical Biotechnology Center, University of Turku, Itäinen Pitkäkatu 4 C, 20521 Turku, Finland

Abbreviations: CRC, colorectal cancer; CTC, circulating tumor cells; PBMC, peripeheral blood mononuclear cells; RT-PCR, reverse transcriptase polymerase chain reaction; RNA, ribonucleic acid; BM, bone marrow; PBMC, peripheral blood mononuclear cells; LN, lymph node; *SERPINB5*, serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5; VSNL1, visinin-like 1; STC1, stanniocalcin 1; DPEP1, dipeptidase 1 (renal); ROC-curve, receiver operating characteristic-curve

Key words: colorectal cancer, serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5, real-time RT-PCR, circulating tumor cells

Introduction

In colorectal cancer (CRC), a subgroup of patients with early stage disease still die of metastasis or recurrent disease within 5 years after having undergone surgery with curative intent (R0-resection). The prognosis of CRC patients as well as the choice of adjuvant therapy is related to the stage of disease (1,2). Staging of CRC is routinely performed using clinical and histopathological criteria (3). Recent years have seen the development of more sensitive methods such as molecular detection of tumor cells in blood, bone marrow or lymph nodes that may help to improve current staging strategies. For example, using RT-PCR to amplify carcinoembryonic antigen (CEA) messenger RNA, Liefers *et al* were able to retrospectively identify a subgroup of patients with micrometastatic disease but histologically negative lymph nodes who might have benefited from adjuvant chemotherapy (4). Similarly, the detection of circulating tumor cells (CTC) in peripheral venous blood has been suggested to contribute to improved staging of CRC patients, early diagnosis of relapse, monitoring of adjuvant therapy and as independent prognostic

factor (5,6). Particularly, the analysis of peripheral venous blood samples is appealing, as this biological material can be obtained easily and thus allows repetitive sampling for convenient monitoring of cancer patients. RNA-based assays are problematic due to lack of disease-specific marker genes and main drawbacks are related to the almost universal presence of background signals resulting in false positive results (7). Furthermore inflammation increases γ-Interferon levels, which in turn can induce the transcription of tissue-specific messenger RNA in blood leucocytes (8).

This has repeatedly been observed for 'first generation' RT-PCR markers such as cytokeratin (CK) and CEA that are deduced from serum markers (7-10) and studies concerning the specificity of RT-PCR based assays for the detection of disseminated tumor cells in peripheral venous blood samples are conflicting (5,11). Comparison of different study results is hardly possible as different techniques concerning RNA extraction, reverse transcription, PCR and signal detection are applied and standardization of methods for the detection of disseminated tumor cells has still not been agreed upon (12,13). The introduction of real-time PCR for quantification of gene expression is considered a step towards standardization and might also improve diagnostic specificity by distinguishing low-level background transcription from 'real signals' when defining a 'cut off value' for expression of marker genes in peripheral venous blood (14,15).

Circulating tumor cells are known to be highly heterogeneous concerning their genetic profiles (14,16) and using a panel of multiple tumor-specific transcript markers can increase diagnostic sensitivity for detection of CTC as well (17,18).

The aim of this study was to identify tumor-specific transcript markers that solely or in combination might be valuable for the specific as well as sensitive detection of disseminated tumor cells in peripheral blood of colorectal cancer patients. In total, 346 genes extracted from own and publicly available microarray data were found to display tumor-specific upregulation of gene expression.

Subsequently this gene list was submitted to a systematic selection process in order to reduce the amount of marker genes and select the most promising candidates. Finally four markers were tested on a small set of clinical samples to determine their diagnostic specificity and sensitivity by relative quantitative real-time RT-PCR.

Table I. Patients characteristics.

Cancer group			
Patient ID	Tumor stage (TNM)	Gender	Age
T1	T2, N0, M0	M	39
T2	T2, N0, M0	M	79
T3	T2, N0, M0	F	71
T4	T3, N0, M0	M	84
T5	T3, N0, M0	M	62
T6	T3, N2, M1 (Liver)	F	62
T7	T3, N2, M0	M	61
T8	T3, N2, M1 (Liver)	F	57
T9	T2, N0, M0	M	45
T10	T3, N1, M0	M	70
T11	T3, N2, M1 (Omentum)	F	64
T12	T3, N0, M0	F	79
T13	T3, N0, M0	M	73
T14	T3, N0, M0	M	80
T15	T4, N1, M0	M	60
T16	T3, N0, M0	F	79
T17	T3, N1, M0	M	69
T18	T3, N0, M0	M	64
Control group			
N1	Healthy	F	62
N2	Healthy	M	71
N3	IBD	F	19
N4	Healthy	F	48
N5	Healthy	M	55
N6	Healthy	M	46
N7	IBD	M	22
N8	Healthy	F	63
N9	Healthy	F	74
N10	Healthy	M	68
N11	IBD	F	30
N12	IBD	M	42

Patients' identification number (Patient ID), Tumor stage (TNM), gender and age are listed for the 30 individuals that were chosen for initial selection of markers. IBD, inflammatory bowel disease; F, female and M, male.

Materials and methods

Patient selection and processing of blood samples. The protocol was approved by the local Ethics Committee and informed consent was obtained from all patients. Blood specimens from CRC patients (n=18) were obtained directly before surgery and histopathological tumor classification was done according to the TNM guidelines (3). In addition, blood specimens were drawn from a nonmalignant control collective (n=12) (Table I). *SERPINB5* was further validated using an enlarged independent sample set consisting of 63 CRC patients (mean age 72±10 years) and 36 control individuals without malignant disease (mean age 66±12 years). Venous blood samples of 10 ml were collected in EDTA containing tubes (Sarstedt, Nümbrecht, Germany). In order to prevent contamination with skin cells, the first 5 ml of each venous blood sample were discarded. PBMC were isolated by density-gradient centrifugation through Ficoll-Paque (Amersham Pharmacia Biotech, Freiburg, Germany). Peripheral blood mononuclear cells (PBMC) were washed twice with phosphate-buffered saline and the cell pellets were snap-frozen in liquid nitrogen and stored at -80˚C until further use.

Selection of marker genes. In total, 346 genes were identified as being upregulated in colorectal cancer tissue and CRC cell lines by own cDNA microarry experiments (Nees, unpublished

Table II. Publicly available sources of gene expression data.

Publications	
First author	Title
Alon *et al* (19)	Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays
Kitahara *et al* (20)	Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia
Notterman *et al* (21)	Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma and normal tissue examined by oligonucleotide arrays
Williams *et al* (22)	Identification and validation of genes involved in the pathogenesis of colorectal cancer cDNA using microarrays and RNA interference
Zou *et al* (23)	Application of cDNA microarrays to generate a molecular taxonomy capable of distinguishing between colon cancer and normal colon
Su *et al* (24)	Molecular classification of human carcinomas by use of gene expression signatures
Ross *et al* (25)	Systematic variation in gene expression patterns in human cancer cell lines
WEB sources	
Institution	URL
National Center for Biotechnology Information (NCBI)	http://www.ncbi.nlm.nih.gov
SAGE Digital Gene Expression Displayer (DGED)	http://cgap.nci.nih.gov/SAGE
Stanford Genomic Resources (SOURCE)	http://genome-www5.stanford.edu/cgi-bin/source/sourceSearch
ONCOMINE (26)	http://www.oncomine.org/main/index.jsp

Candidate genes with upregulated gene expression in CRC tissue and cell lines were extracted from published data files or identified by own cDNA micoarray analysis (Nees, unpublished data), as described in Materials and methods.

data) and by screening publicly available sources of microarray data (Table II) (19-26).

Genes have been described as being upregulated by comparing expression profiles of cancer tissue and corresponding normal mucosa (19-23). Threshold setting for definition of 'upregulation' was heterogeneous reaching from expression changes of >2-fold in ~one-third of the patients (n=20) (22) to 4-fold difference or greater ($p>0.001$) in expression intensity between tumor and normal (21). Furthermore colon cancer-specific genes were identified throughout comparison of different cancer entities (24) or cell lines (25). Only genes that could be identified unambiguously on the basis of the given characteristics (accession number, short name, primer sequence) were further considered.

This approach rendered finally a list of 346 marker genes that were systematically evaluated for their feasibility to be used as a diagnostic tool for detection of CTC according to the strategy displayed in Fig. 1. All 346 genes were systematically investigated for their cDNA sources (screening step A in Fig. 1) using the NCBI- and SOURCE databases (Table II). Those genes with cDNA sources related to whole blood or single components of blood such as thrombocytes and nucleated blood cells were excluded from further investigations. The remaining genes with no evidence of any cDNA expression in the blood compartment were tested for gene expression by RT-PCR with template derived from mononuclear cells of one healthy individual (screening step B in Fig. 1). In the subsequent screening step C (Fig. 1) nested RT-PCR reactions were performed for the remaining candidate genes on cDNA from PBMC preparations of three healthy individuals. Marker genes that still displayed no positive signal were included in the final validation step. Relative quantitative real-time RT-PCR was performed on PBMC of blood

Table III. Oligonucleotide primers for nested real-time RT-PCR.

First PCR					
Primer name	Sequence (5' → 3')	Product length (bp)	Ta (˚C)	Gene symbol	Accession no.
RDP1_r	TGCACTGCTGGTGACATACA	460	55	*DPEP1*	NM_004413
RDP1_f	GCCTGAAGCTCATCCTTCTG				
SERPINB5 f	TGTGAACGACCAGACCAAAA	406	55	*SERPINB5*	U04313
SERPINB5 r	CAAGCCTTGGGATCAATCAT				
STC1 f	AAGGATGATTGCTGAGGTGC	193	55	*STC1*	U25997
STC1 r	CAGGCTGTCTCTGATTGTGC				
VSNL1 f	AGTTCTTTCCTTATGGAGACGC	391	55	*VSNL1*	NM_003385
VSNL1 r	AGTAATACAATGGAAGGGTCGC				
Second PCR					
Primer name	Sequence (5' → 3')	Product length (bp)	Ta (˚C)	Gene symbol	Accession no.
HMBS_N f	GGATGGGCAACTGTACCTGACTGGA	139	65	*HMBS*	NM_000190
HMBS_N r	TGCCTACCAACTGTGGGTCATCCTC				
RDP1_N r	CAGCTGCCAGGGGAGGTCATTGT	129	65	*DPEP1*	NM_004413
RDP1_N f	ATGTGGAGCGGATGGTGGCTGT				
SERPINB5_N f	TGCTGCCTACTTTGTTGGCAAGTGG	132	65	*SERPINB5*	U04313
SERPINB5_N r	CCCATACAGAACGTGGCCTCCA				
STC1_N f	CAGCAAGCTGAATGTGTGCAGCATC	128	65	*STC1*	U25997
STC1_N r	ACATTCCAGCAGGCTTCGGACAAG				
VSNL1_N f	AGCAGAAGCTGAACTGGGCCTTCAA	144	65	*VSNL1*	NM_003385
VSNL1_N r	TCAGGCCATCCTCATTCATTTTCATCA				

Relative quantification of gene expression for *DPEP1*, *SERPINB5*, *STC1* and *VSNL1* was performed as described in Materials and methods. *HMBS* was chosen as housekeeping gene. The primer sequences, length of amplified PCR products in base pairs (bp) and annealing temperatures (Ta) are indicated in the table.

specimens derived from 18 colorectal cancer patients before tumor surgery and 12 nonmalignant control individuals (Fig. 1).

RNA extraction and cDNA synthesis. Starting from TRIzol lysates (Gibco-BRL, Karlsruhe, Germany) total RNA was extracted using chloroform, precipitated with isopropanol and washed with 70% ethanol (Merck, Darmstadt, Germany). The resulting pellet was redissolved in 100 μl nuclease-free water (Sigma-Aldrich, München, Germany). RNA concentration was measured using a spectrophotometer (Pharmacia-Biotech, Genequant II, Freiburg, Germany). For cDNA synthesis, 15 μl total RNA was treated with DNase I (Gibco-BRL) and reversely transcribed for 1 h at 42˚C in 60 μl reaction volume containing final concentrations of 1X RT buffer, 500 μM dNTP, 2 U/μl RNAsin (Promega, Mannheim, Germany), 25 ng/μl oligo-$(dT)_{15}$ primer (Invitrogen, Karlsruhe, Germany) and 10 U/μl Superscript II Reverse Transcriptase (Invitrogen). The constitutively expressed gene Hydroxymethylbilane synthase (HBMS) was chosen as housekeeping gene (27) and a 137 bp fragment was amplified after each cDNA synthesis to check quality of the RNA preparation. The sequence of primers is shown in Table III.

Reverse transcription and polymerase chain reaction (RT-PCR). All oligonucleotides (German Cancer Research Center, Heidelberg, Germany) were designed with the PRIMER 3 software (28) and HPLC purified. The primer sets were constructed to span at least one intron to avoid amplification of contaminating genomic DNA. All primers were confirmed to show no significant homology with other known genes by using the BLAST Sequence Similarity Search tool (NCBI).

Reverse Transcription as well as PCR reactions were carried out on a Mastercycler (Eppendorf, Wesseling-Berzdorf, Germany). For RT-PCR, a total volume was adjusted to 25 μl using a final concentration of 200 μM dNTPs, 125 pM of each primer, 1X PCR buffer (20 mM Tris-Cl, pH 8.4; 50 mM KCl), 1.5 mM $MgCl_2$ and 0.75 units of Taq DNA polymerase (Invitrogen) and 5 μl of cDNA as template.

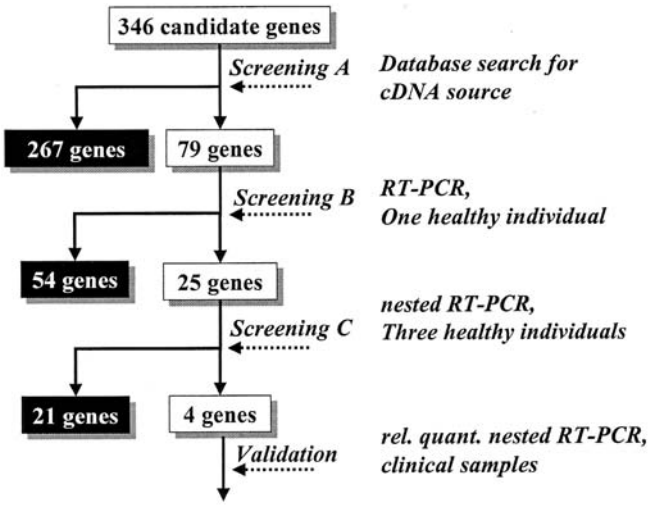

Figure 1. Selection of candidate genes. A list of 346 candidate genes was systematically screened to identify tumor-specific marker genes that can be used for detection of circulating tumor cells in peripheral venous blood of colorectal cancer patients. Each screening step (A-C) resulted in exclusion of several genes from further analysis due to background expression in PBMC (black boxes).

PCR was performed using the following conditions: initial denaturation at 94˚C for 5 min followed by 35 cycles of denaturation at 94˚C for 30 sec, annealing for 30 sec at 55˚C and primer extension at 72˚C, 1 min. The final extension step was carried out at 72˚C for 7 min. PCR products were analysed by agarose gel electrophoresis and ethidiumbromide staining.

Nested real-time RT-PCR. The first PCR was conducted as described above with the exception that only 15 PCR cycles were performed. The second PCR was conducted with 20 μl reaction volumes consisting of 10 μl SYBR-Green PCR master mix (QuantiTec SYBR, Qiagen, Hilden, Germany), 2 μl PCR product of the first PCR, 2 μl of each primer (10 μM) and 4 μl of water. The thermal cycling conditions comprised an initial denaturation step at 95˚C for 15 min and 35 cycles at 95˚C for 15 sec, 65˚C for 20 sec and terminal elongation at 72˚C for 10 min. All PCR reactions were performed using a LightCycler (Roche, Mannheim, Germany). Specificity of amplified products was checked by melting curve analysis and agarose gelelectrophoresis with ethidiumbromide staining and only primer-dimer-free reactions were considered valid. Negative RT-PCR samples, which were exactly generated in parallel to cDNA synthesis solely without adding reverse transcriptase, were investigated independently to exclude genomic DNA contamination and pseudogene amplification. Data analysis was carried out using the LightCycler software (version 3.5) supplied with the LightCycler (Roche).

Nested relative quantitative real-time RT-PCR. For marker genes *SERPINB5*, *VSNL1*, *STC1* and *DPEP1* nested real-time RT-PCR was performed on PBMC preparations of 18 CRC patients and 12 control individuals as described above with the only modification that PCR cycle number of the nested PCR was extended to 45 cycles. The relative amount of transcript in each sample was calculated using the comparative ΔCt-Method (29). Briefly, the cycle threshold (Ct) value was defined as the number of PCR cycles required for the fluorescent intensities to exceed a threshold just above background. Ct values were measured for each marker gene (MG) and the housekeeping gene (HKG) hydroxymethylbilane synthase (*HMBS*) in every sample. Ct values were subtracted to obtain ΔCt (ΔCt) using the following formula: ΔCt = Ct (MG) - Ct (HKG). The ΔCt value of each sample was calculated respectively. The mean ΔCt value of all specimens from the nonmalignant control group was designated 'calibrator'. The ΔCt values for each sample (S) and calibrator (C) were subtracted to obtain ΔΔCt [ΔΔCt = ΔCt (S) - ΔCt (C)]. The amount of marker gene, normalized to an endogenous reference (*HMBS*) and relative to the calibrator was calculated using the following formula: Relative amount of marker gene = $2^{-\Delta\Delta Ct}$.

Each experiment was performed in duplicate including separately performed first and second RT-PCR. The average value of both duplicates was used for calculation of relative transcript quantity according to the Δ-Ct method, as described above.

To reduce the risk of contamination, thermocycling and post-PCR steps on one hand and RNA extraction, cDNA synthesis and preparation of the PCR mixtures on the other hand were performed in separate laboratories. All samples were analyzed in duplicate.

Processing of cell lines, blood spiking experiments. Colorectal cancer cell lines SW480 and T84 were obtained from the 'Tumorbank', German Cancer Research Centre, Heidelberg, and the American Tissue and Cell Culture Collection (ATCC, Rockville, MD, USA), respectively.

Cells were grown in RPMI medium supplemented with 10% heat inactivated fetal bovine serum at 37˚C in a 5% CO_2 air environment. For blood spiking experiments, cells were harvested at 80% confluency with trypsin-EDTA (Sigma-Aldrich). Cell suspension was cautiously pipetted up and down several times in order to obtain singly dispersed cells; efficiency of disaggregation was checked microscopically. Cells were counted in a Neubauer chamber and were then serially diluted with blood from a healthy volunteer to obtain concentrations of tumor cells in the range of 10^4 to one cell per 1 ml blood (Fig. 2). The spiked blood samples were processed by Ficoll density gradient centrifugation as already described to obtain the PBMC fraction.

Statistical analysis. The Mann-Whitney test was performed as a two-tailed test to detect differences between blood samples from CRC patients and the control group concerning the expression levels of *SERPINB5*, *VSNL1*, *STC1* and *DPEP1* respectively. A significance level of $p<0.05$ was chosen as cut off. Receiver operating characteristic (ROC) curves were calculated using the MedCalc® statistical software package (Version 8.0, MedCalc Software, Belgium).

Results

Candidate gene selection. A total of 346 candidate marker genes that have been described to be upregulated in CRC tissue and cell lines were systematically screened for their feasibility to detect circulating tumor cells in the blood of

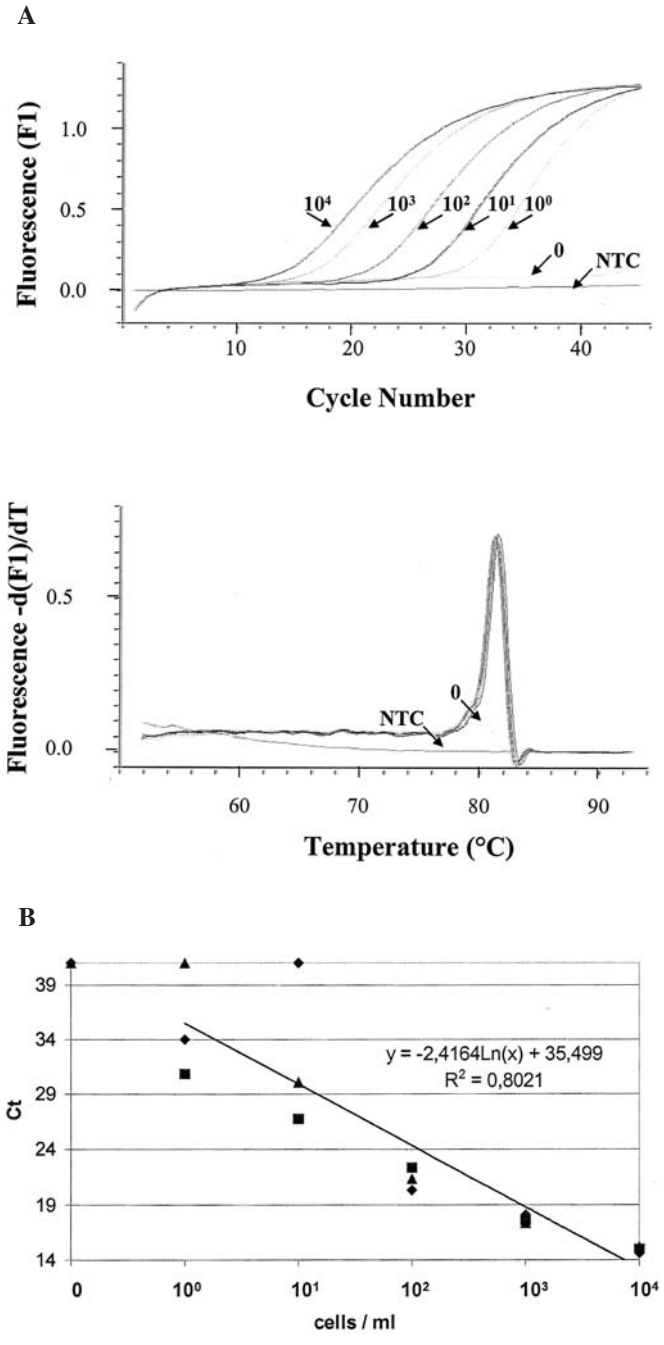

Figure 2. (A) Determination of analytical sensitivity. Colorectal tumor cells SW480 were serially diluted with blood of one healthy donor, as described in Materials and methods. The real-time fluorescent signals (upper graph) and result of melting curve analysis (lower graph) are shown for samples spiked with 10^4 to 10^0 cells per ml. Amplification was performed with *SERPINB5* primers. Blood samples with no tumor cell contamination as well as a no template control (NTC) were included in the run. (B) Reproducibility of Ct values. Blood specimens of one healthy donor were spiked with tumor cells (SW480) by serial dilution, as described in Materials and methods. PBMC were isolated by density gradient centrifugation (Ficoll). Nested RT-PCR and relative quantification of *SERPINB5* transcript was performed in triplicate. Linear regression equation and R^2 coefficient are displayed in the graph.

CRC patients by means of RT-PCR (Fig. 1). An initial screening step resulted in the reduction of the number of candidate markers by 267 genes as systematic database search revealed that their cDNA source was related to blood or cellular components of blood. The remaining 79 genes were submitted to 35 cycles of RT-PCR using a PBMC preparation of one healthy individual as template. Of these, 54 candidate markers gave positive signals and therefore were excluded from further analysis. For the remaining 25 genes a nested RT-PCR assay was set up to test if any background signals might arise from PBMC preparations of three healthy individuals. Finally 4 candidate genes *SERPINB5*, *DPEP1*, *VSNL1* and *STC1* were left over that consistently showed no signal in any of the described screening steps and therefore were subsequently submitted to a final validation step.

Determination of analytical sensitivity for detection of colorectal tumor cells in peripheral venous blood samples. A dilution of one CRC-cell in 1 ml peripheral venous blood gave positive results for *SERPINB5* and *VSNL1* using SW480 cells or T84 cells for detection of *STC1* and *DPEP1* transcripts respectively (Fig. 2A). These findings are in accordance with the reported analytical sensitivity of RT-PCR-based assays for detection of disseminated tumor cells in malignant disease (30). Poor reproducibility was observed for highly diluted samples. While cell density of 10^4, 10^3 10^2 cells/ml showed good reproducibility of Ct values (SD ± 0.3, 0.37 and 1.1), the highly diluted samples with ten and one cell per ml, respectively, displayed gross variations (SD ± 7.5 and 5.2) concerning the results of repetitive experiments (Fig. 2B). These results are caused by stochastic effects that play a major role at the limit of analytical sensitivity (31).

Relative quantification of SERPINB5, VSNL1, STC1, DPEP1 expression. To study diagnostic sensitivity as well as specificity for the detection of disseminated tumor cells, the relative expression levels of the four selected marker genes (*SERPINB5*, *VSNL1*, *STC1*, *DPEP1*) were analyzed in blood of 18 CRC patients and in 12 blood specimens of the control group. The expression of *HMBS* as housekeeping gene was comparable in all samples included in the study with a mean Ct value of 23.7 and a standard deviation of 1.23 indicating that the nucleic acid qualities were comparable between the samples. The relative amount of transcript level for specimens of the CRC patients and control group is shown in Fig. 3A. For the three markers *VSNL1*, *DPEP1* and *STC1* the median transcript levels are near the value of one and do not differ significantly. Only for *SERPINB5* a statistically significant difference of transcript levels was observed in the tumor group compared to the control group (Man-Whitney rank-sum test, $p<0.05$).

The relative expression fold of the four marker genes *SERPINB5*, *VSNL1*, *DPEP1* and *STC1* was visualized using the TreeView software (32): Clear differences of *SERPINB5* expression were found in specimens of tumor patients but the other markers (*VSNL1*, *DPEP1* and *STC1*) displayed a rather heterogeneous expression pattern (Fig. 3B).

For ROC curve assembly different cut off values were tested for their feasibility to correctly classify between patients and individuals of the control group. Highest classification accuracy was obtained for *SERPINB5* and integrated ROC curve areas gave highest values (0.801) for this marker (Fig. 3C) when compared to *DPEP1*, *STC1* and *VSNL1* with values of 0.637, 0.560 and 0.593, respectively.

Finally *SERPINB5* as most promising marker was further validated using a collective consisting of 63 CRC patients

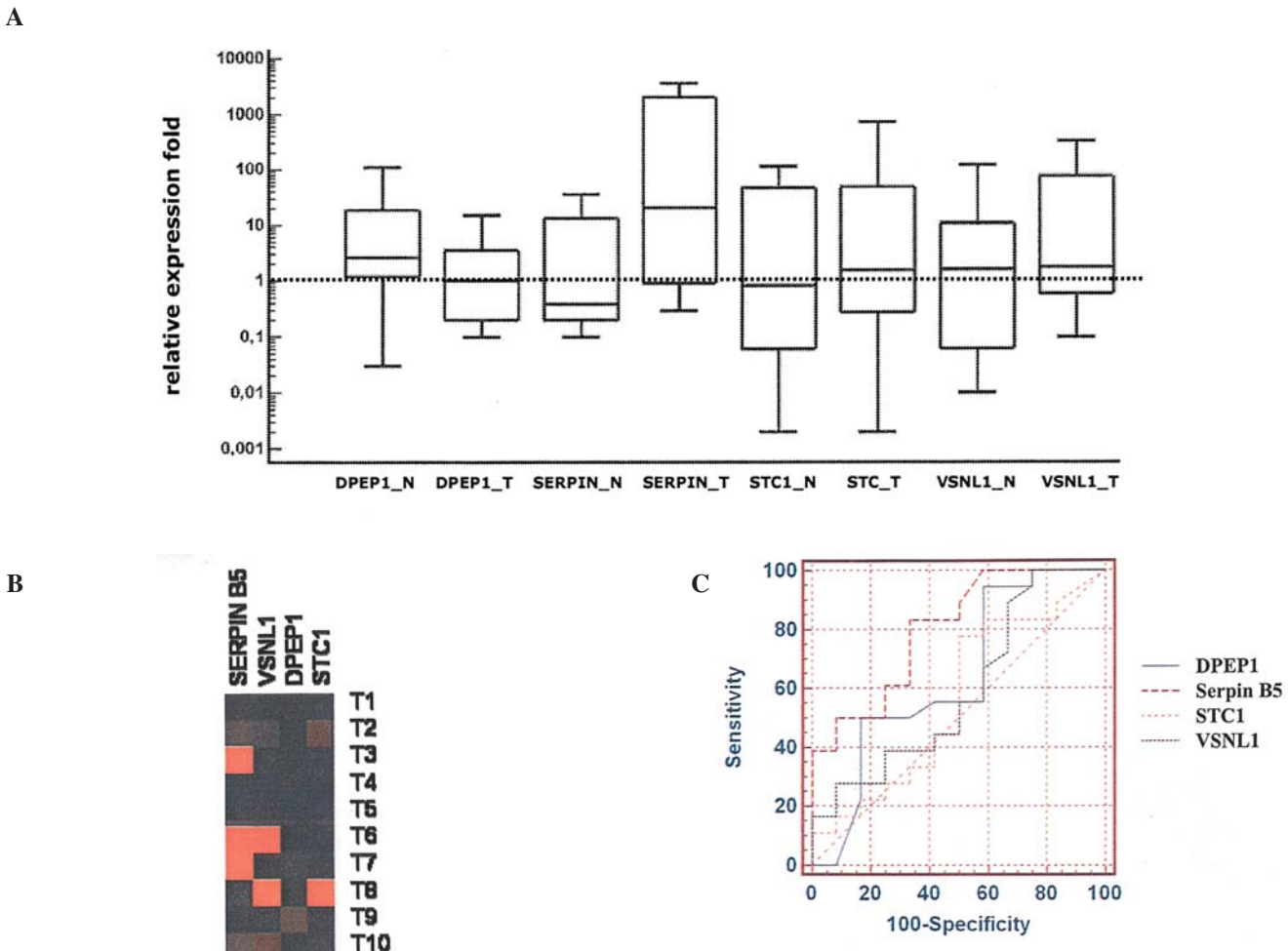

Figure 3. (A) Relative quantification of marker genes within tumor and control collectives. Comparison of nonmalignant (N) individuals and tumor patients (T) concerning relative expression fold of the marker genes *DPEP1*, *SERPINB5*, *STC1* and *VSNL1*. The Box-Whisker plot indicates median values, 95% intervals and minimal as well as maximal values of transcript levels. The only statistically significant difference of expression fold between the two groups (N/T) could be observer for *SERPINB5* (p=0.0059). The other marker genes (*DPEP1*, *STC1* and *VSNL1* respectively) displayed p-values >0.2. (B) Expression profiles of individual PBMC samples. The fold expression over baseline (mean of control group) is displayed for each blood specimen and any gene (*SERPINB5*, *VSNL1*, *DPEP1* and *STC1*) using the TreeView software (32). Tumor patients (T1-18, upper half) and the non-malignant control group (N1-12, lower half) are separated by the dashed line. Fold changes of gene expression are color-coded and red color indicates overexpression of respective genes in peripheral blood mononuclear cells. (C) Receiver operating characteristic (ROC) analysis. ROC curve analysis for discrimination of tumor patients from nonmalignant control individuals by marker gene expression of *DPEP1*, *SERPINB5*, *STC1* and *VSNL1*, respectively. Only the marker gene *SERPINB5* showed acceptable classification accuracy. The areas under the ROC curves were the following: *DPEP1*, 0.637; *SERPINB5*, 0.801; *STC1*, 0.560 and *VSNL1*, 0.593.

and 36 nonmalignant control individuals. The nonmalignant control collective showed a mean relative expression of 6.95±17.08 with values ranging from 0.0045 to 52.65. The tumor collective displayed a significantly elevated (p=0.0278) mean relative expression of 122.45±309 with values ranging from 0.0022 to 2045.26. (Fig. 4). When choosing a cut off value with maximum specificity 23 out of 63 (36%) tumor patients were classified correctly.

Discussion

The molecular detection of tumor cells in bone marrow, lymph nodes and blood of tumor patients has greater analytical sensitivity when compared to immunohistochemical detection. However limited diagnostic specificity of commonly used tissue-specific transcript markers is circumventing broad applications of RT-PCR-based assays for the diagnosis of

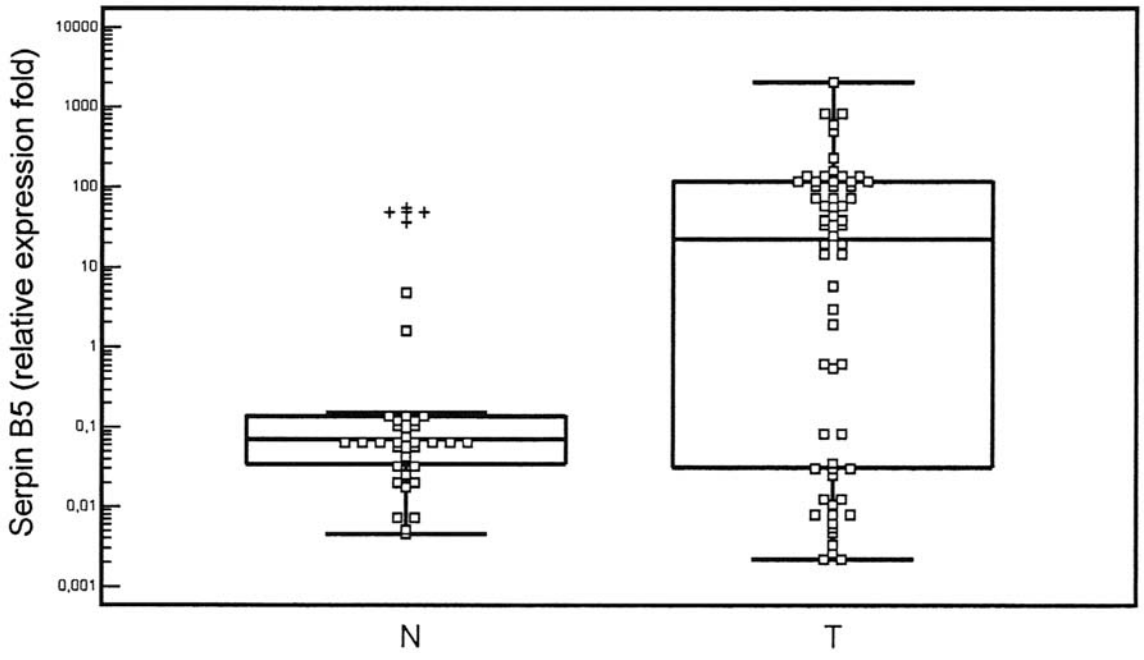

Figure 4. Relative quantification of *SERPINB5* in an enlarged tumor and control collective. Relative expression fold of *SERPINB5* in peripheral blood of 36 non-malignant control individuals (N) and 63 CRC patients (T). In the Box-and-whisker plot, the central box represents the values from the lower to upper quartile (25 to 75 percentile). The middle line represents the median ranging from 0.07 (N) to 22.6 (T). The horizontal line extends from the minimum to the maximum value, excluding outside (square symbol) and far out (cross symbol) values, which are displayed separately. An outside/far outside value is defined as a value that is larger than the upper quartile plus 1.5/3 times the interquartile range (inner fences).

minimal residual disease using peripheral venous blood samples (7,8,14).

In order to develop an assay system with improved sensitivity and specificity for the detection of circulating CRC cells, we screened a list of 346 marker genes that have been described as being overexpressed in CRC tissue samples. After a stepwise screening process only four marker genes *SERPINB5*, *DPEP1*, *VSNL1* and *STC1* were left over for validation of their feasibility to detect CTC in peripheral venous blood samples of CRC patients. To determine diagnostic specificity control individuals with no malignant disease were also investigated. Despite the transcript quantification, we found it difficult to define a cut off for the tested marker genes that clearly separated the tumor patients from the nonmalignant control group without reducing sensitivity. When applying the highest threshold values (>1000-fold overexpression), only *SERPINB5* could correctly classify 33% (6/18) of the tumor patients with 100% specificity (Fig. 3B); interestingly the *SERPINB5* expression was not correlated to the stage of disease (data not shown). As circulation of cancer cells seems to be a rare event, even in advanced diseases, the analysis of multiple blood samples has been proposed to circumvent this problem (33).

For quantitative real-time PCR, the use of fluorogenic probes seems to be advantageous when compared to SYBR-Green if low level transcripts are going to be detected (34). Nevertheless, SYBR-Green technology has the advantage of being easily adapted and this technology has already been applied successfully for the detection of CTC in peripheral blood of cancer patients (35). Baker *et al* (35) specifically isolated poly-A+ RNA instead of total RNA and this has recently been described to increase sensitivity and reproducibility of tumor cell detection in peripheral blood (36) allowing them to perform only one PCR with 50 thermocycles. In contrast, we extracted total RNA and subsequently had to perform a nested PCR protocol to obtain sufficient analytical sensitivity. These technical differences might contribute to decreased reproducibility that we observed when analyzing samples with low concentration of target transcript. Nevertheless, stochastic effects are crucial in real-time PCR if low numbers of disseminated tumor cells are going to be detected. In case of low concentrated target transcripts, amplification is determined by chance, which results in inconsistent positive results (31) and low reproducibility of quantitative results (37). Never-theless the definition of cut off values with emphasis on good specificity rather than sensitivity seems to be beneficial (17).

For any chosen threshold the classification accuracy of *SERPINB5* was superior to that of the other markers as shown by ROC curve analysis (Fig. 3C). Interestingly, *DPEP1* has recently been proposed for detection of circulating tumor cells in CRC patients (38) but proved to be of no diagnostic value within our setting. The authors used an immunobead RT-PCR protocol and superior sensitivity and specificity of immunobead RT-PCR over regular RT-PCR for the detection of tumor cells in peripheral venous blood specimens has already been demonstrated (9). Concerning *DPEP1*, *VSNL1* and *STC1* we have to conclude that either copy number of analyzed markers in circulating tumor cells is too low, or peripheral blood, due to background signals, is not the appropriate compartment for detection of circulating tumor cells when using the applied method.

In contrast *SERPINB5* showed up to be applicable for the detection of circulating tumor cells in peripheral blood samples of CRC patients. *SERPINB5* has recently been described as being upregulated in colorectal adenomas (39) and it is

synonym to *Maspin* that has been proposed for the detection of circulating tumor cells in bone marrow and peripheral blood of breast cancer patients (40). Although *Maspin* has been characterized as a tumor suppressor, the overexpression seems to be involved in tumor progression and metastatic spread of intestinal cancer as well (41-43). In colorectal cancer, the *Maspin* expression is related to microsatellite instability and typically displays a nuclear staining pattern (42). Previous studies have suggested that active *Maspin* must be located at the cytoplasmic membrane in order to conduct its tumor-suppressing activities (44) but the detailed role of *Maspin* in the tumorigenesis and progression of CRC still remains unclear.

According to our knowledge, this is the first study describing *SERPINB5* as a specific marker for the detection of circulating tumor cells in CRC patients. Diagnostic specificity of *SERPINB5* might be limited by induction of gene expression in normal leucocytes under influence of chemokines, growth factors and especially proinflammatory cytokines (45), a well known phenomenon that has been reported for other transcript markers as well (8).

However, quantitative results are enabling definition of cut off values and adjusting 100% specificity, 23 out of 63 cancer patients (36%) were identified with elevated *SERPINB5* expression.

Further studies have to clarify, if overexpression of *SERPINB5* in peripheral blood specimens of CRC patients is of long-term prognostic significance or might be beneficial for therapeutic decisions.

Acknowledgements

This project was supported in part by the Lesser-Loewe Foundation e.V.

References

1. Moertel C, Fleming T, MacDonald J, *et al*: Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. N Engl J Med 322: 352-358, 1990.
2. Wichmann M, Müller C, Hornung H, *et al*: Results of long-term follow-up after curative resection of Dukes A colorectal cancer. World J Surg 26: 732-736, 2002.
3. Sobin L and Fleming I: TNM classification of malignant tumors, fifth edition (1997). Cancer 80: 1803-1804, 1997.
4. Liefers G, Cleton-Jansen A, van de Velde C, *et al*: Micrometastases and survival in stage II colorectal cancer. N Engl J Med 339: 223-228, 1998.
5. Vogel I and Kalthoff H: Disseminated tumour cells. Their detection and significance for prognosis of gastrointestinal and pancreatic carcinomas. Virchows Archiv 439: 109-117, 2001.
6. Vlems F, Wobbes T, Punt C, *et al*: Detection and clinical relevance of tumor cells in blood and bone marrow of patients with colorectal cancer. Anticancer Res 23: 523-530, 2003.
7. Jung R, Petersen K, Krüger W, *et al*: Detection of micrometastasis by cytokeratin 20 RT-PCR is limited due to stable background transcription in granulocytes. Br J Cancer 81: 870-873, 1999.
8. Jung R, Krüger W, Hosch S, *et al*: Specificity of reverse transcriptase polymerase chain reaction assays designed for the detection of circulating cancer cells is influenced by cytokines in vivo and in vitro. Br J Cancer 78: 1194-1198, 1998.
9. Park S, Lee B, Kim I, *et al*: Immunobead RT-PCR versus regular RT-PCR amplification of CEA mRNA in peripheral blood. J Cancer Res Clin Oncol 127: 489-494, 2001.
10. Zippelius A, Kufer P, Honold G, *et al*: Limitations of reverse-transcriptase polymerase chain reaction analyses for detection of micrometastatic epithelial cancer cells in bone marrow. J Clin Oncol 15: 2701-2708, 1997.
11. Tsavellas G, Patel H and Allen-Mersh T: Detection and clinical significance of occult tumour cells in colorectal cancer. Br J Surg 88: 1307-1320, 2001.
12. Lacroix J and Doeberitz M: Technical aspects of minimal residual disease detection in carcinoma patients. Semin Surg Oncol 20: 252-264, 2001.
13. Krüger W, Jung R, Krüger N, *et al*: Sensitivity of assays designed for the detection of disseminated epithelial tumor cells is influenced by cell separation methods. Clin Chem 46: 435-436, 2000.
14. Vlems F, Diepstra J, Cornelissen I, *et al*: 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.
15. Champelovier P, Mongelard F and Seigneurin D: CK20 gene expression: technical limits for the detection of circulating tumor cells. Anticancer Res 19: 2073-2078, 1999.
16. Klein C, Blankenstein T, Schmidt-Kittler O, *et al*: Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. Lancet 360: 683-689, 2002.
17. Mitas M, Mikhitarian K, Walters C, *et al*: Quantitative real-time RT-PCR detection of breast cancer micrometastasis using a multigene marker panel. Int J Cancer 93: 162-171, 2001.
18. Bosma A, Weigelt B, Lambrechts A, *et al*: Detection of circulating breast tumor cells by differential expression of marker genes. Clin Cancer Res 8: 1871-1877, 2002.
19. Alon U, Barkai N, Notterman D, *et al*: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci USA 96: 6745-6750, 1999.
20. Kitahara O, Furukawa Y, Tanaka T, *et al*: Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia. Cancer Res 61: 3544-3549, 2001.
21. Notterman D, Alon U, Sierk A, *et al*: Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. Cancer Res 61: 3124-3130, 2001.
22. Williams N, Gaynor R, 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.
23. Zou T-T, Selaru F, Xu Y, *et al*: Application of cDNA microarrays to generate a molecular taxonomy capable of distinguishing between colon cancer and normal colon. Oncogene 21: 4855-4862, 2002.
24. Su A, Welsh J, Sapinoso L, *et al*: Molecular classification of human carcinomas by use of gene expression signatures. Cancer Res 61: 7388-7393, 2001.
25. Ross D, Scherf U, Eisen M, *et al*: Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 24: 227-235, 2000.
26. Rhodes D, Yu J, Shanker K, *et al*: ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 6: 1-6, 2004.
27. Warrington J, Nair A, Mahadevappa M, *et al*: Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. Physiol Genom 2: 143-147, 2000.
28. Rozen S and Skaletsky H: Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132: 365-386, 2000.
29. Livak K and Schmittgen T: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408, 2001.
30. Ghossein R, Bhattacharya S and Rosai J: Molecular detection of micrometastases and circulating tumor cells in solid tumors. Clin Cancer Res 5: 1950-1960, 1999.
31. Jung R, Ahmad-Nejad P, Wimmer M, *et al*: Quality management and influential factors for the detection of single metastatic cancer cells by reverse transcriptase polymerase chain reaction. Eur J Clin Chem Clin Biochem 35: 3-10, 1997.
32. Eisen MB, Spellman PT, Brown PO, *et al*: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95: 14863-14868, 1998.
33. Wharton R, Jonas S, Glover C, *et al*: Increased detection of circulating tumor cells in the blood of colorectal carcinoma patients using two reverse transcription-PCR assays and multiple blood samples. Clin Cancer Res 5: 4158-4163, 1999.

34. Yin J, Shackel N, Zekry A, *et al*: Real-time reverse transcriptase-polymerase chain reaction (RT-PCR) for measurement of cytokine and growth factor mRNA expression with fluorogenic probes or SYBR Green I. Immunol Cell Biol 79: 213-221, 2001.
35. Baker M, Mikhitarian K, Osta W, *et al*: Molecular detection of breast cancer cells in the peripheral blood of advanced-stage breast cancer patients using multimarker real-time reverse transcription-polymerase chain reaction and a novel porous barrier density gradient centrifugation technology. Clin Cancer Res 9: 4865-4871, 2003.
36. Burchill S, Lewis I and Selby P: Improved methods using the reverse transcriptase polymerase chain reaction to detect tumour cells. Br J Cancer 79: 971-977, 1999.
37. Vlems F, Diepstra J, Cornelissen I, *et al*: Investigations for a multi-marker RT-PCR to improve sensitivity of disseminated tumor cell detection. Anticancer Res 23: 179-186, 2003.
38. McIver C, Lloyd J, Hewett P, *et al*: Dipeptidase 1: a candidate tumor-specific molecular marker in colorectal carcinoma. Cancer Lett 209: 67-74, 2004.
39. Wentzensen N, Wilz B, Findeisen P, *et al*: Identification of differentially expressed genes in colorectal adenoma compared to normal tissue by suppression subtractive hybridization. Int J Oncol 24: 987-994, 2004.
40. Luppi M, Morselli M, Bandieri E, *et al*: Sensitive detection of circulating breast cancer cells by reverse-transcriptase polymerase chain reaction of maspin gene. Annal Oncol 7: 619-624, 1996.
41. Umekita Y, Souda M and Yoshida H: Expression of maspin in colorectal cancer. In Vivo 20: 797-800, 2006.
42. Bettstetter M, Woenckhaus M, Wild P, *et al*: Elevated nuclear maspin expression is associated with microsatellite instability and high tumour grade in colorectal cancer. J Pathol 205: 606-614, 2005.
43. Zheng H, Tsuneyama K, Cheng C, *et al*: Maspin expression was involved in colorectal adenoma-adenocarcinoma sequence and liver metastasis of tumors. Anticancer Res 27: 259-265, 2007.
44. Sheng S, Carey J, Seftor E, *et al*: Maspin acts at the cell membrane to inhibit invasion and motility of mammary and prostatic cancer cells. Proc Natl Acad Sci USA 93: 11669-11674, 1996.
45. Ballestrero A, Garuti A, Bertolotto M, *et al*: Effect of different cytokines on mammaglobin and maspin gene expression in normal leukocytes: possible relevance to the assays for the detection of micrometastatic breast cancer. Br J Cancer 92: 1948-1952, 2005.