Int J Colorectal Dis (2005) 20: 353–362
DOI 10.1007/s00384-004-0664-7

ORIGINAL ARTICLE

Roland S. Croner
Thomas Foertsch
Wolfgang M. Brueckl
Klaus Guenther
Renate Siebenhaar
Christian Stremmel
Klaus E. Matzel
Thomas Papadopoulos
Thomas Kirchner
Jürgen Behrens
Ludger Klein-Hitpass
Michael Stuerzl
Werner Hohenberger
Bertram Reingruber

Common denominator genes that distinguish colorectal carcinoma from normal mucosa

Accepted: 25 August 2004
Published online: 22 December 2004

R. S. Croner (✉) · T. Foertsch · K. Guenther · R. Siebenhaar · C. Stremmel · K. E. Matzel · W. Hohenberger · B. Reingruber
Department of Surgery, University of Erlangen, Krankenhausstrasse 12, 91054 Erlangen, Germany
e-mail: Roland.Croner@chir.imed.uni-erlangen.de
Tel.: +49-9131-8533296
Fax: +49-9131-8535840

W. M. Brueckl
1st Department of Internal Medicine, University of Erlangen, Krankenhausstrasse 12, 91054 Erlangen, Germany

T. Papadopoulos · T. Kirchner
Institute of Pathology, University of Erlangen, Krankenhausstrasse 12, 91054 Erlangen, Germany

J. Behrens
Department of Experimental Medicine II, University of Erlangen, Krankenhausstrasse 12, 91054 Erlangen, Germany

L. Klein-Hitpass
Institute of Cell Biology, University of Essen, Germany

M. Stuerzl
Department of Molecular and Experimental Surgery, University of Erlangen, Krankenhausstrasse 12, 91054 Erlangen, Germany

Abstract *Purpose:* Microarray technology has been used by a growing number of investigators and several studies have been published that list hundreds of genes differentially expressed by colorectal carcinoma (CRC) and normal mucosa (MC). On the basis of our own and other investigators' microarray data, our goal was to identify a common denominator gene cluster distinguishing CRC from MC. *Methods:* Thirty GeneChips (HG-U133A, Affymetrix) were hybridized, 20 with RNA of CRC stages I–IV (UICC) and 10 with MC. Expression signals showing at least a 4-fold difference between CRC and MC ($p<0.01$) were identified as differentially expressed. In addition, in our integrative data analysis approach only those genes whose expression was altered simultaneously in at least 2 of 5 recently published studies were subjected to an unsupervised hierarchical cluster analysis. *Results:* We detected 168 up- and 283 down-regulated genes in CRC relative to MC. Twenty-three genes were filtered from the five articles reviewed. An unsupervised hierarchical cluster analysis of these 23 genes confirmed the high specificity of these genes to differentiate between CRC and MC in our microarray data. *Conclusions:* Colorectal cancer and mucosa could be clearly separated by 23 genes selected for being differentially expressed more than once in a recent literature review. These genes represent a common denominator gene cluster that can be used to distinguish colorectal MC from CRC.

Keywords Gene expression · Colorectal cancer · Mucosa · Microarrays · GeneChip

Introduction

Colorectal cancer (CRC) is a major cause of cancer-related deaths in the western world [1]. The development of CRC is considered a multistep process with an accumulation of different genetic alterations [2, 3]. For example, adenomas progress to invasive adenocarcinomas via mutations or methylations in the promoter regions of oncogenes and tumor-suppressor genes, or mutations of DNA repair genes [1, 2, 4]. Key players in CRC pathogenesis are the adenomatous polyposis coli gene (APC), β-catenin, k-ras, "deleted in colorectal carcinomas" (DCC), p53, as well as the wnt-signaling and smad-TGFβ signaling pathways [3]. However, the diagnosis and prognosis of CRC are still based on the conventional histopathological tumor/node/metastasis (TNM) staging system [5]. Early diagnosis in combination with standardized surgical therapy and stage-adapted multimodality treatment have contributed to an increasing cure rate. To improve diagnostic and prognostic accuracy further, gene expression profiling by DNA microarrays may provide a promising tool.

Recently, expression profiles have revealed a number of genes differentially expressed in normal mucosa (MC) and colorectal carcinoma [6–10]. It has not been evaluated how consistent the various sets of genes found in these studies were to discriminate between mucosa and cancer. Once identified, however, these 'common denominator' genes may represent the most significant and consistent molecular differences between cancer tissue and mucosa at the RNA level.

In the project presented, we performed microarray analyses using CRC and MC RNA. The aim was to use the previously published data regarding differential gene expression, extract a cluster of those genes that were quoted unanimously by a minimum of two authors in the past and test these genes for their ability to distinguish between carcinoma and normal mucosa in our own microarray data set of colorectal tissue samples.

Material and methods

Expression profiling

The Ethical Committee of our university approved the study, patient consent was obtained, and the research conformed to the principles of the Declaration of Helsinki. Twenty colorectal carcinomas of stages I–IV (UICC) and ten specimens of healthy mucosa were used for this study. Tissue samples were shock frozen in liquid nitrogen and embedded in TissueTek (Zakura, Zoeterwoude, Netherlands) immediately after surgery, and stored at −80°C until work-up by cryotomy after manual dissection (CMD) as recently described [11].

Ribonucleic acid (RNA) isolation was performed using commercial kits (RNeasy Kit; Qiagen, Hilden, Germany), following the manufacturer's protocol. Included in this procedure was a DNAse (Qiagen) digestion following the manufacturer's suggestions. RNA quality and quantity were determined by the 'Lab-on-a-Chip' method (Bioanalyzer 2100; Agilent Technologies, Palo Alto, CA, USA) following the manufacturer's instructions [12]. The 3′/5′ ratios for the housekeeping genes glycerinaldehyde-3-phosphatase (GAPDH) and β-actin present on the GeneChip were used as further parameters for RNA quality and to exclude partial degradation. A 3′/5′ ratio below 3 was regarded as an indicator of good RNA quality

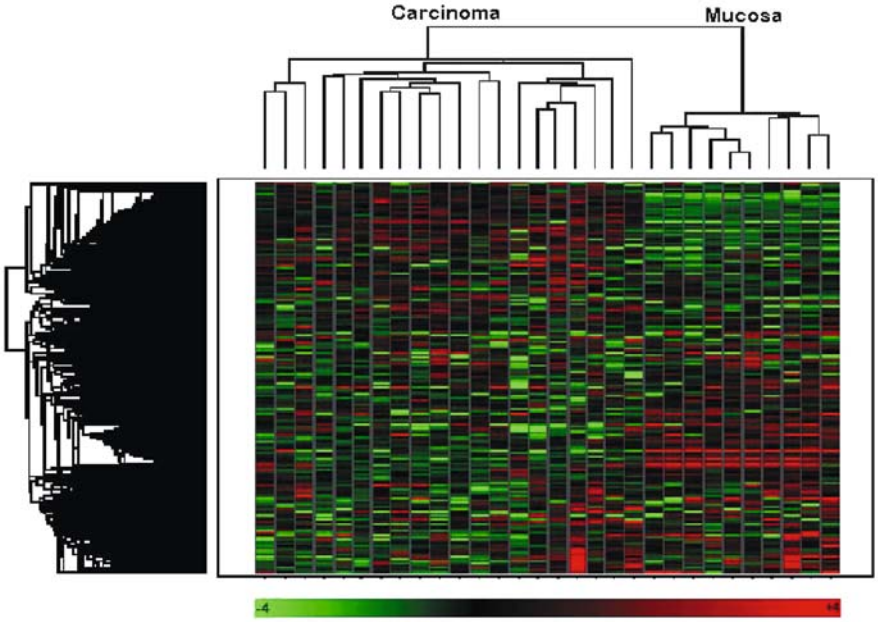

Fig. 1 Hierarchical cluster analysis of gene expression values from colorectal cancer and normal mucosa samples. All 22,284 probe sets (HG-U 133A; Affymetrix) per GeneChip are included. *−4* 4-fold down-regulation in colorectal cancer, *+4* 4-fold up-regulation in colorectal cancer

Table 1 Up-regulated genes in colorectal cancer (*CRC*; *n*=20) vs. normal mucosa (*n*=10) detected by microarray analyses

Accession number	Annotation	Fold change CRC vs. mucosa	Mann–Whitney *p* value
Extracellular matrix			
J04177	Alpha-1 type XI collagen (COL11A1)	46	<0.001
NM_002421.2	Matrix metalloproteinase 1 (MMP1)	21	0.002
NM_002423.2	Matrix metalloproteinase 7 (MMP7)	21	<0.001
NM_002422.2	Matrix metalloproteinase 3 (MMP3)	19	<0.001
X98568	Collagen, type X, alpha 1	14	0.002
M83248.1	Nephropontin	12	<0.001
NM_000089.1	Collagen, type I, alpha 2 (COL1A2)	10	<0.001
NM_000602.1	Nexin, PAI type 1 (SERPINE1)	7	<0.001
AW043713	KIAA1077 (sulfatase)	6	<0.001
NM_001711.1	Biglycan (BGN)	6	<0.001
NM_005940.2	Matrix metalloproteinase 11 (MMP11)	5	<0.001
NM_003247.1	Thrombospondin 2 (THBS2)	5	<0.001
NM_004353.1	Serine proteinase inhibitor (SERPINH1)	5	<0.001
AI433463	Enkephalinase CALLA, CD10	5	0.001
NM_000216.1	Kallmann syndrome 1 sequence (KAL1)	5	0.001
NM_000393.1	Collagen, type V, alpha 2 (COL5A2)	4	<0.001
NM_001845.1	Collagen, type IV, alpha 1 (COL4A1)	4	<0.001
H95960	**Osteonectin (SPARC)**	**4**	**<0.001**
Cell cycle/mitosis			
NM_006398.1	Diubiquitin (UBD)	9	0.006
D38553.1	HCAP-H	7	<0.001
NM_004749.1	Cell cycle progression 2 protein (CPR2)	5	<0.001
NM_003158.1	Serinethreonine kinase 6 (STK6)	5	<0.001
NM_005483.1	Chromatin assembly factor 1 (CHAF1A)	5	<0.001
NM_024094.1	MGC5528	6	<0.001
U33286	**Brain cellular apoptosis protein (CSE1)**	**5**	**<0.001**
NM_002467.1	Myelocytomatosis viral oncogene (MYC)	5	<0.001
NM_001168.1	Survivin (BIRC5)	4	<0.001
NM_013296.1	LGN protein (HSU54999)	4	<0.001
H50438	**Cell division cycle 25B (CDC25B)**	*4*	**<0.001**
U77949.1	Cdc6-related protein (HsCDC6)	4	<0.001
AI081175	**Interferon induced protein (IFITM1)**	**3**	**<0.001**
X57351	**Interferon induced protein (IFITM2)**	**3**	**<0.001**
Cell adhesion/cytoskeleton			
NM_001793.1	P-cadherin, placental (CDH3)	74	<0.001
NM_021101.1	Claudin 1 (CLDN1)	10	<0.001
NM_003474.2	Disintegrin, meltrin alpha (ADAM12)	9	<0.001
NM_006393.1	Nebulette (NEBL)	5	<0.001
Signal transduction			
M13436.1	**Ovarian beta-A inhibin**	**37**	**<0.001**
NM_000584.1	Interleukin 8 (IL8)	20	<0.001
AF030514.1	Small inducible cytokine 11 subfamily B	15	0.002
NM_002090.1	GRO3 oncogene	13	<0.001
AL524520	G protein-coupled receptor 49	9	<0.001
NM_001511.1	GRO1 oncogene	9	<0.001
NM_007079.1	Tyrosine phosphatase IVA (PTP4A3)	9	<0.001
M57731.1	GRO2 oncogene	8	<0.001
M80927.1	Chitinase 3-like 1	8	0.001
NM_021953.1	Forkhead box M1 (FOXM1)	7	<0.001
NM_003212.1	Teratocarcinoma-derived growth factor 1	7	<0.001

Table 1 (continued)

Accession number	Annotation	Fold change CRC vs. mucosa	Mann–Whitney *p* value
NM_004289.3	Nuclear factor-like 3 (NFE2L3)	6	<0.001
NM_004392.1	Dachshund (Drosophila) homolog (DACH)	6	0.001
NM_007019.1	Ubiquitin carrier protein E2-C (UBCH10)	5	0.001
BE550452	Homer, neuronal immediate gene, 1B	5	<0.001
M77349	**Transforming growth factor induced (TGFBI)**	**5**	**<0.001**
NM_003391.1	Wingless-type MMTV member 2 (WNT2)	5	<0.001
AU118882	Endothelin receptor type A	4	<0.001
NM_014467.1	Sushi-repeat protein (SRPUL)	4	<0.001
NM_001657.1	Amphiregulin (AREG)	4	0.003
Miscellaneous			
NM_001807.1	Carboxyl ester lipase (CEL)	31	0.002
U19869.1	Fatty acid binding protein 6 (FABP6)	30	0.001
NM_004413.1	Dipeptidase 1, renal (DPEP1)	23	<0.001
NM_021246.1	Megakaryocyte gene transcript 1 (MEGT1)	16	0.003
NM_012101.1	Ataxia-telangiectasia D protein (ATDC)	7	0.001
NM_005564.1	Lipocalin 2 (LCN2)	7	0.003
NM_001808.1	Carboxyl ester lipase-like (CELL)	6	0.002
NM_000483.2	Apolipoprotein C-II (APOC2)	5	0.002
BG165833	Fatty acid desaturase 1	5	0.001
M61832	**Homocysteine hydrolase (AHCY)**	**4**	**<0.001**
X04481	**Complement component 2 (C2)**	**3**	**0.001**

Fold change: gene expression values CRC/MC. Genes whose expression was altered simultaneously in at least two studies evaluated during the literature review are shown in bold

according to the manufacturer's protocol (Affymetrix, Santa Clara, CA, USA) [13].

Gene expression was examined using the GeneChip technology (Affymetrix). Biotin-labeled cRNA was generated by in vitro transcription as described previously and hybridized to the GeneChips (HG-U133A) following the manufacturer's instructions [14]. Fragmentation of cRNA, hybridization to GeneChips, washing and staining, as well as scanning of the arrays in the GeneArray scanner (Agilent) were performed as recommended by the Affymetrix Gene Expression Analysis Technical Manual. Signal intensities and detection calls for statistical analysis and hierarchical clustering were determined using the GeneChip 5.0 software.

Fluorescence staining

Stainings of colorectal carcinoma tissue sections were performed as described previously [15, 16]. Tumor biopsies were routinely processed, formalin-fixed, and paraffin-embedded. For staining, the MMP-1 antigen paraffin-embedded tumor sections were treated for 20 min at 80°C in Target Retrieval Solution pH 9.0 (Dako, Glostrup, Denmark). Slides were then incubated with a mouse anti-MMP-1 monoclonal antibody (1:100; IM352; Oncogene, Cambridge, MA, USA). Bound primary antibodies were detected with a goat anti-mouse antibody (1:500) coupled to the fluorochromes *AlexaFluor488* (Molecular Probes Europe, Leiden, The Netherlands). Evaluation of tissue sections was carried out with a Leica confocal fluorescence microscope.

Training set

Five previously published studies using cDNA or oligonucleotide microarrays for the comparison of colorectal cancer and mucosa were reviewed with respect to differentially expressed genes [6–10]. Accession numbers and annotations of our genes and those listed in tables in these publications were compared. Genes that were shown to be regulated in at least two were summarized (see Table 4). The corresponding probe sets on the GeneChip (HG-U 133A; Affymetrix) used in our study were determined and the list of resulting transcripts was used for further analysis.

Statistical analyses

Only genes with an average signal above 250 were included in our statistical analyses. Significance levels of microarray results in CRC vs. MC were calculated using the Mann–Whitney *U*-test. The average signal intensity of CRC/MC was calculated as fold change. Unsupervised

Table 2 Down-regulated genes in CRC (n=20) vs. normal mucosa (n=10) detected by microarray analyses

Accession number	Annotation	Fold change CRC vs. mucosa	Mann–Whitney p value
Extracellular matrix			
NM_001884.1	Cartilage linking protein 1 (CRTL1)	−6	<0.001
NM_021924.1	Mucin and cadherin-like (MUCDHL)	−5	<0.001
NM_004532.1	Mucin 4, tracheobronchial (MUC4)	−5	<0.001
NM_013363.1	Procollagen C-endopeptidase (PCOLCE2)	−5	<0.001
NM_021073.1	Bone morphogenetic protein 5 (BMP5)	−4	<0.001
Cell cycle/mitosis			
NM_004944.1	Deoxyribonuclease I-like 3 (DNASE1L3)	−7	<0.001
AF078844.1	Helicase-related protein	−7	<0.001
Cell adhesion/cytoskeleton			
AL049977.1	Claudin 8	−47	<0.001
NM_017717.2	Mucin and cadherin-like (MUCDHL)	−10	<0.001
NM_003277.1	Claudin 5 (CLDN5)	−6	<0.001
NM_021153.1	Cadherin 19, type 2 (CDH19)	−5	<0.001
NM_006614.1	Cell adhesion molecule homolog L1CAM	−4	<0.001
S67238.1	Smooth muscle myosin heavy chain SM2	−4	<0.001
Channels/transporter			
AF011390.1	**Sodium bicarbonate cotransporter**	**−37**	**<0.001**
NM_007168.1	ATP-binding cassette, (ABC1), member 8	−29	<0.001
NM_003759.1	Solute carrier family 4 (SLC4A4)	−25	<0.001
L02785	**Solute carrier family 26 (SLC26A3)**	**−13**	**<0.001**
AF069510.1	Sodium bicarbonate cotransporter (NBC)	−13	<0.001
NM_001442.1	Fatty acid binding protein 4 (FABP4)	−11	<0.001
U14528	**Sulfate transporter family 26**	**−9**	**<0.001**
NM_000336.1	Sodium channel (SCNN1B)	−9	<0.001
NM_005495.1	Solute carrier family 17 (SLC17A4)	−8	<0.001
AF127036.1	Chloride channel (CaCC1)	−7	<0.001
Signal transduction			
M97496	**Guanylate cyclase activator (GUCA1B)**	**−52**	**<0.001**
NM_004160.1	Peptide YY (PYY)	−32	<0.001
NM_020406.1	Polycythemia rubra vera 1 (PRV1)	−30	<0.001
NM_001275.2	Chromogranin A (CHGA)	−19	<0.001
NM_000114.1	Endothelin 3 (EDN3)	−10	<0.001
NM_005123.1	Nuclear receptor subfamily 1 (NR1H4)	−10	<0.001
NM_003381.1	Vasoactive intestinal peptide (VIP)	−8	<0.001
NM_000909.1	Neuropeptide Y receptor Y1 (NPY1R)	−7	<0.001
NM_003890.1	IgG Fc binding protein	−7	<0.001
U88321.1	Small inducible cytokine subfamily A	−7	<0.001
NM_000901.1	Nuclear receptor subfamily 3 (NR3C2)	−6	<0.001
NM_022969.1	Fibroblast growth receptor 2 (FGFR2)	−6	<0.001
NM_000072.1	CD36 antigen (thrombospondin receptor)	−5	<0.001
NM_000196.1	Hydroxysteroid dehydrogenase (HSD11B2)	−5	<0.001
NM_003265.1	Toll-like receptor 3 (TLR3)	−4	<0.001
U28249	**Ion transport regulator 3 (FXYD3)**	**−4**	**<0.001**
Metabolism			
NM_001738.1	Carbonic anhydrase I (CA1)	−59	<0.001
M83670	**Carbonic anhydrase IV (CA4)**	**−39**	**<0.001**
NM_001077.1	UDP glycosyltransferase (UGT2B17)	−30	0.001
J03037	**Carbonic anhydrase II**	**−26**	**<0.001**
M24317.1	Class I alcohol dehydrogenase (ADH2)	−22	<0.001
NM_020299.1	Aldo-keto reductase family 1(AKR1B10)	−12	<0.001

Table 2 (continued)

Accession number	Annotation	Fold change CRC vs. mucosa	Mann–Whitney *p* value
NM_020973.1	Cytosolic beta-glucosidase (GLUC)	−11	<0.001
M12272	**Alcohol dehydrogenase 3 (ADH3)**	**−10**	**<0.001**
NM_000055.1	Butyrylcholinesterase (BCHE)	−10	<0.001
M29644.1	Insulin-like growth factor (somatomedin C)	−4	<0.001
Y09616	**Carboxylesterase 2 (intestine, liver)**	**−4**	**<0.001**
M26393	**Acyl-CoA dehydrogenase (ACADS)**	**−3**	**<0.001**
M10050	**Fatty acid binding protein 1 (FABP1)**	**−3**	**<0.001**
M16364	**Creatine kinase, brain (CKB)**	**−3**	**0.001**
M19483	**ATP synthase, H+ transporting (ATP5B)**	**−1**	**0.001**
Miscellaneous			
NM_005478.2	Insulin-like 5 (INSL5)	−20	<0.001
NM_002153.1	Hydroxysteroid dehydrogenase (HSD17B2)	−18	<0.001
NM_001150.1	Microsomal aminopeptidase (ANPEP)	−13	<0.001
NM_002407.1	Mammaglobin 2 (MGB2)	−10	<0.001
NM_003278.1	Tetranectin (TNA)	−10	<0.001
NM_007072.1	HERV-H LTR-associating 2 (HHLA2)	−9	<0.001
U17077	**Clone MGC:4419**	**−5**	**0.001**
U29091	**Selenium binding protein 1 (SELENBP1)**	**−4**	**<0.001**
N80129	**Metallothionein 1L (MT1L)**	**−4**	**<0.001**

Fold change: gene expression values CRC/MC. Genes whose expression was altered simultaneously in at least two studies evaluated during the literature review are shown in bold

hierarchical cluster analysis of all 22,284 probe sets (HG-U 133A; Affymetrix) and unsupervised hierarchical cluster analysis of the selected probe sets extracted from the literature review ("training set") were performed.

Results

Expression profiling

The 3′/5′ ratio for GAPDH on the GeneChips was below 3 in all cases. This threshold value was suggested by the manufacturer as an indicator that no relevant RNA degradation occurred during the microdissection and the RNA processing for the microarray hybridization [13]. From all 30 samples we were able to isolate enough good-quality RNA for the oligonucleotide microarray hybridization.

In the hierarchical cluster analysis that was initially performed, which included all 22,284 probe sets of the HG-U 133A GeneChip, CRC and MC samples were clearly separated. A cluster of up- and down-regulated genes in CRC vs. MC was identified (Fig. 1). Using our selection criteria we detected 168 up- and 283 down-regulated genes in CRC compared with MC.

A list of selected genes up-regulated in CRC is presented in Table 1. Matrix metalloproteinases, several types of collagens, and other genes coding for components of the extracellular matrix were over-expressed in CRC. Another prominent group included genes related to cell cycle and mitosis such as survivin, CDC25B or diubiqui-

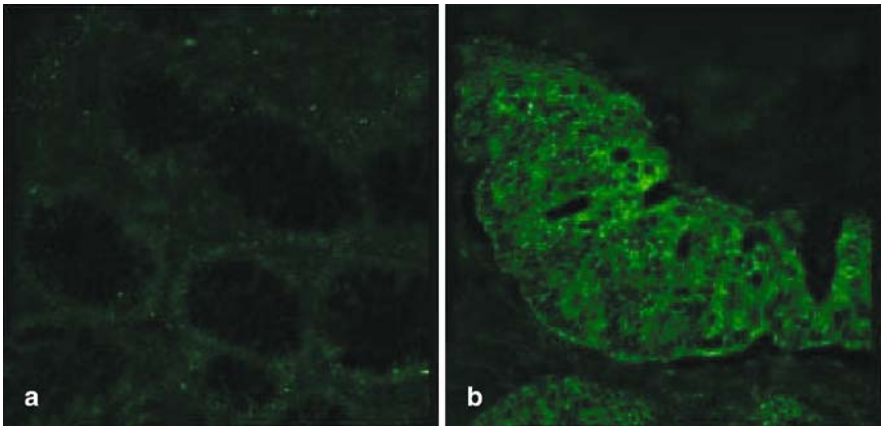

Fig. 2 Immunhistochemical fluorescence staining of MMP-1 in **a** normal mucosa and **b** colorectal cancer.

Table 3 Comparison of five previously published studies that evaluated differences in gene expression between colorectal carcinomas (*CRC*) and normal mucosa (*MC*) with the use of microarray techniques. *n.d.* not described, *LCM* laser capture microdissection, *s.a.* statistical algorithm, *HE* hematoxylin and eosin, *EST* expressed sequence taq

	Reference number				
	[8]	[7]	[6]	[9]	[10]
Total samples	33	16	36	32	40
Carcinoma	21	8	18	16	20
Mucosa	11	8	18	16	20
CRC ratio control	HE control	LCM	HE control	n.d.	n.d.
RNA isolation	RNAzol	Qiagen	n.d.	TRIzol	Qiagen
Microarray	HuGeneFL (EST A–D)[a]	cDNA	Hu6500[a]	cDNA	cDNA
Probe sets	6,800 (35,000 ESTs)	9,216	6,600	4,600	9,592
Change ratio	≤−3 and ≥3 (≤−5 and ≥5)[b]	s.a.	≤−4 and ≥4[b]	≤−2 and ≥2[b]	≤−2 and ≥2[b]
Genes CRC/MC	226 (157 ESTs)	235	107	82	2,632
Up	88	44	19	23	574
Down	138	191	88	59	2,058

[a]GeneChips (Affymetrix)
[b]Fold change

tin. Furthermore, several components of cell adhesion and the cytoskeleton were consistently up-regulated, e.g., P-cadherin, claudin 1, and nebulette. Genes involved in signal transduction such as TGFBI, GRO2 and GRO3 oncogene, and the tyrosine phosphatase IVA showed overexpression in CRC.

Table 4 List of differentially expressed genes in colorectal carcinomas vs. mucosa, discovered in our own data and five previously published studies working with microarray techniques

Accession number	Reference number					
	Present study	[8]	[7]	[6]	[9]	[10]
U33286	↑	↑	–	↑	–	–
L02785	↓	↓	–	↓	–	–
M77349	↑	↑	↑	–	–	–
M97496	↓	↓	–	↓	–	–
M61832	↑	↑	–	↑	–	–
M12272	↓	↓	–	↓	–	–
M83670	↓	↓	–	↓	–	–
L09708/X04481	↑	↑	↑	–	–	–
J03037	↓	–	↓	↓	↓	–
H50438/M81934	↑	–	↑	↑	–	–
H95960	↑	–	–	↑	↑	↑
AI081175/AA058323	↑	–	↑	–	–	↑
N80129	↓	↓	–	–	–	↓
M16364/L47647	↓	↓	↓	–	–	–
M26393/AA676663	↓	–	–	↓	–	↓
X57351	↑	–	↑	–	↑	–
U17077	↓	–	↓	↓	↓	–
M19483/M19482	↓	↓	–	–	↓	–
Y09616	↓	–	–	–	↓	–
M10050	↓	↓	–	–	↓	–
U14528	↓	↓	–	–	↓	–
U28249	↓	↓	–	–	↓	–
U29091	↓	↓	–	–	↓	–

↑ Up-regulated genes in colorectal carcinomas, ↓ down-regulated genes in colorectal carcinomas, – gene was not listed in the papers' tables

A list of selected genes down-regulated in CRC is presented in Table 2. Genes such as mucin 4 or smooth muscle myosin heavy chain SM2, and several others belonging to the extracellular matrix, were down-regulated in CRC. Only a few genes associated with the cell cycle and mitosis-like helicase-related protein or genes connected to cell adhesion and cytoskeleton such as claudin 5, 8, and CDH19 were under-expressed in CRC. Different transporters such as sodium or chloride channels and solute carriers were down-regulated in CRC. Genes responsible for signal transduction such as peptide YY, VIP or neuropeptide Y receptor showed an under-expression in CRC. A high number of genes involved in metabolism like carbonic anhydrases, alcohol dehydrogenase 2, 3, and somatomedin C were down-regulated in CRC.

Fluorescence staining

Ribonucleic acid (RNA) expression of MMP-1 was evaluated as significantly up-regulated in CRC vs. MC in our microarray data. Similarly, immunohistochemical staining for MMP-1 in tissue samples showed markedly increased levels in CRC compared with MC. The differences in MMP-1 mRNA expression correlated with the protein levels detected by immunohistochemical fluorescence staining (Fig. 2).

Training set

In five studies reviewed, 8–20 samples of CRC and MC were included. The investigators used different methods of tissue purification and RNA isolation. In two studies, Affymetrix GeneChips, and in three studies, non-commercial cDNA microarrays were used. Between 82 and 2,632 differentially expressed genes were described (Table 3). A total of 23 genes were detected to be similarly regulated in CRC vs MC in at least three studies including our own data. Overall, the genes most frequently listed were

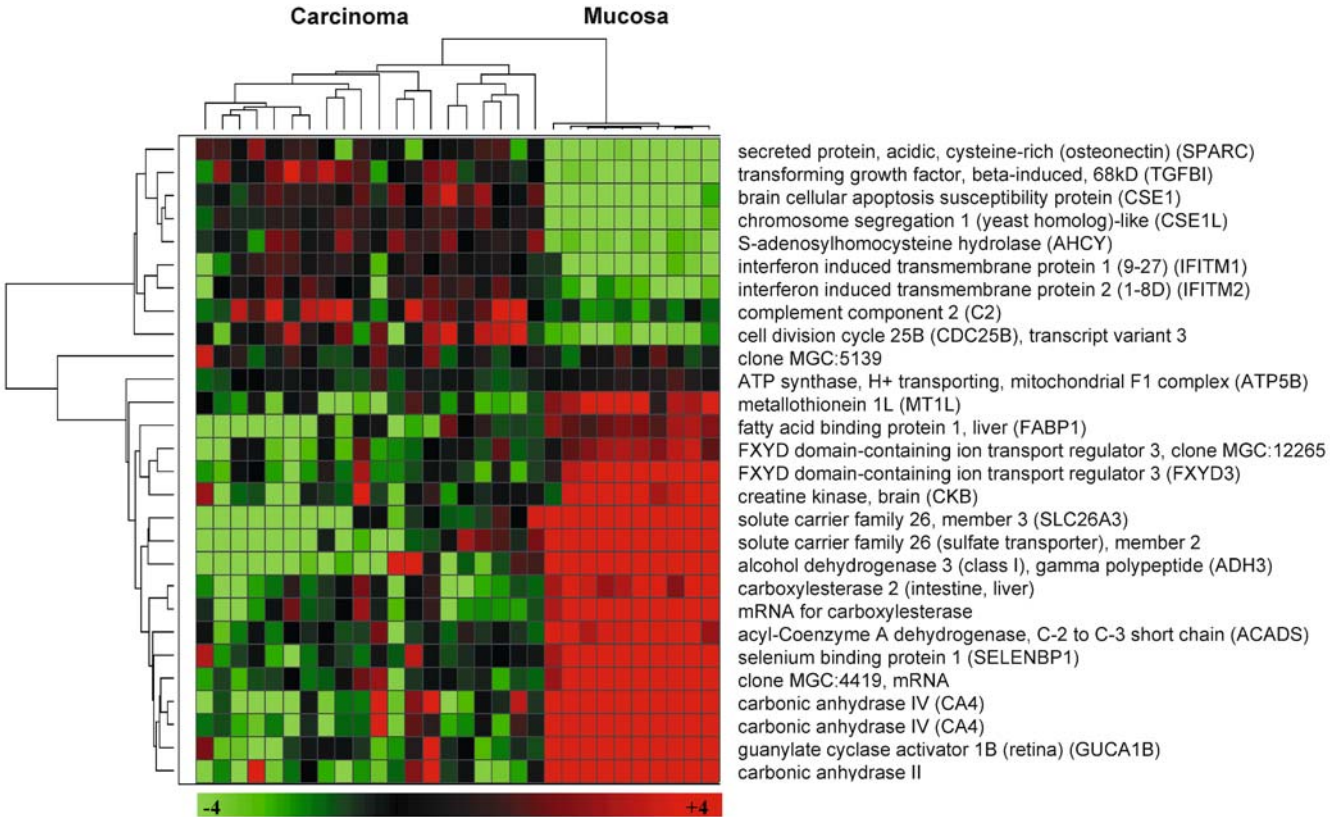

Fig. 3 Hierarchical cluster analysis of selected genes from our data, which were evaluated as "training set" during the literature review. *−4* 4-fold down-regulation in colorectal cancer, *+4* 4-fold up-regulation in colorectal cancer

human carbonic anhydrase II and homo sapiens clone MGC:4419. Not one single gene was common in all published series (Table 4). The most similarities were found between the studies of Birkenkamp-Demtroder et al. and Notterman et al., where eight common genes were found [6, 8]. For each of the accession numbers U33286, M26393, Y09616, and M83670 two probe sets were detected on our GeneChip (HG-U 133A). The 23 evaluated genes correspond to 28 probe sets on the HG-U133A Chip. When these probe sets were applied for hierarchical cluster analysis of our expression data, a clear separation between MC and CRC was achieved (Fig. 3), i.e., all 23 genes found in the literature review showed a differential expression in CRC vs. MC (Tables 1, 2). The fold change ratio (FC) CRC/MC of the average signals on the GeneChips was between 5.04 and −51.58. Eight genes were up- and 15 genes were down-regulated in CRC vs MC.

Discussion

We identified 168 up- and 283 down-regulated genes in CRC vs MC in accordance with our selection criteria. This number of differentially expressed genes in CRC vs. MC is within the range of the numbers given by the studies previously published (Table 3). Together with other investigators we found more genes under- than over-expressed in CRC vs. MC. The down-regulation in CRC of many genes responsible for tissue differentiation may indicate the regression of the tumors into a pluripotent state comparable to the fetal organism.

We detected several up- and down-regulated genes of the extracellular matrix in CRC. High levels of collagen, type I, and type IV in CRC tissue have been already reported [17]. Matrix metalloproteinases were identified as being responsible for tumor progression and metastasis in CRC. Especially MMP-1 is an important factor for tumor invasion [18]. MMP-1 showed very high expression values in our data and the immunohistochemical staining correlated with the microarray results. These findings confirm the reliability of our gene expression analysis. Survivin, a regulator of cell death and cell proliferation, which is usually not expressed in healthy tissue, has already been described in CRC pathogenesis and was detected as being up-regulated in CRC in our series as well [19]. Over-expression of CDC25 in CRC as found in our study correlated with a poor prognosis in CRC [20]. Claudin 1 is involved in the beta-catenin/Tcf signaling pathway and is frequently up-regulated in CRC, which was confirmed in our results as well [21]. The putative role of interleukin 8 (IL8) in the modulation of

cancer cell progression and metastasis was recently described [22]. Again, an over-expression of IL8 was detected in our data. In summary, the involvement of many previously described genes in carcinogenesis, tumor progression, and metastasis was confirmed by our investigations.

Several series of differentially expressed genes in colorectal cancer vs mucosa have been published by investigators using microarray technology [6–10]. The procedures of microarray technology include many steps in tissue handling and RNA preparation. Factors like tissue collection, ischemia, tissue preparation (e.g., laser microdissection), RNA pre-amplification and the type of microarray used can influence the results. Different procedures during these processes may lead to different results. In the five studies compared in our literature review only 23 genes were detected as being differentially expressed in CRC vs MC in at least two studies. Also, the numbers of differentially expressed genes vary between 82 and 2,632 genes in the series published. These very low similarities may be due to the various techniques that were used in the studies reviewed. In particular, the gene selection and amount during the use of non-commercial cDNA microarrays may be an important reason for these diversities. Furthermore, the complete data sets were not available in all studies. Only summarized versions of the gene lists were presented by the investigators. The highest correspondence of eight identical genes was found between Birkenkamp-Demtroder et al. and Notterman et al. [6, 8]. Both groups of authors monitored the carcinoma ratio in the tissue samples by haematoxylin and eosin (H&E) staining and used commercial oligonucleotide microarrays (Gene Chips, Affymetrix).

In the 'training set', genes over-expressed in CRC were those involved in cellular growth, tissue differentiation, cell cycle, and immunity [23–25]. Several of these genes, like CSE1L or SPARC, have been found in other carcinoma types [26, 27]. As expected, TGFBI was markedly over-expressed in CRC (FC 4.91). The lowest gene expression in CRC was found for GUACA1B (FC-51.58). This gene plays a key role in the recovery phase of phototransduction and is involved in retinal dystrophies [28]. Other genes under-expressed in CRC, like FXYD, SLC26A3, and carbonic anhydrase II, have a function during ion exchange and are related to diarrhea or osteoporosis [29–32]. SLC26A3 has been described as being down-regulated in colon adenomas and adenocarcinomas in a previous study that did not employ microarray technology [33]. Carboxylesterases are currently under discussion as becoming useful markers of clinical drug response in colorectal cancer [34]. However, measurement of gene expression values by microarray analysis provides no information about gene function. The biological valence of up- or down-regulated genes should be evaluated by further experiments, e.g., by functional protein analysis via zymography.

Our aim was to distil the large number of genes previously listed into a small group of the most significant genes. We found gene expression profiles, by literature review and in our own data, which clearly distinguish CRC from MC. Unfortunately, the studies reviewed were not fully comparable, even though they all used microarray technology. Standardized microarrays will be necessary to guarantee the comparability of future studies.

Acknowledgements This study was sponsored by the German Federal Ministry for Education and Science (BMBF) within the National Genome Science Network (NGFN).

References

1. Leslie A, Carey FA, Pratt NR, Steele RJ (2002) The colorectal adenoma–carcinoma sequence. Br J Surg 89:845–860
2. Cho KR, Vogelstein B (1992) Genetic alterations in the adenoma–carcinoma sequence. Cancer 70 [Suppl]:1727–1731
3. Arends JW (2000) Molecular interactions in the Vogelstein model of colorectal carcinoma. J Pathol 190:412–416
4. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Nakamura Y, White R, Smits AM, Bos JL (1988) Genetic alterations during colorectal-tumor development. N Engl J Med 319:525–532
5. Hermanek P, Sobin LH, Wittekind C (1999) How to improve the present TNM staging system. Cancer 86:2189–2191
6. Notterman DA, Alon U, Sierk AJ, Levine AJ (2001) Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. Cancer Res 61:3124–3130
7. Kitahara O, Furukawa Y, Tanaka T, Kihara C, Ono K, Yanagawa R, Nita ME, Takagi T, Nakamura Y, Tsunoda T (2001) 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
8. Birkenkamp-Demtroder K, Christensen LL, Olesen SH, Frederiksen CM, Laiho P, Aaltonen LA, Laurberg S, Sorensen FB, Hagemann R, Orntoft TF (2002) Gene expression in colorectal cancer. Cancer Res 62:4352–4363
9. Takemasa I, Higuchi H, Yamamoto H, Sekimoto M, Tomita N, Nakamori S, Matoba R, Monden M, Matsubara K (2001) Construction of preferential cDNA microarray specialized for human colorectal carcinoma: molecular sketch of colorectal cancer. Biochem Biophys Res Commun 285:1244–1249
10. Williams NS, Gaynor RB, Scoggin S, Verma U, Gokaslan T, Simmang C, Fleming J, Tavana D, Frenkel E, Becerra C (2003) Identification and validation of genes involved in the pathogenesis of colorectal cancer using cDNA microarrays and RNA interference. Clin Cancer Res 9:931–946

11. Croner RS, Guenther K, Foertsch T, Siebenhaar R, Brueckl WM, Stremmel C, Hlubek F, Hohenberger W, Reingruber B (2004) Tissue preparation for gene expression profiling of colorectal carcinoma: three alternatives to laser microdissection with preamplification. J Lab Clin Med 143:344–351
12. Nachamkin I, Panaro NJ, Li M, Ung H, Yuen PK, Kricka LJ, Wilding P (2001) Agilent 2100 bioanalyzer for restriction fragment length polymorphism analysis of the Campylobacter jejuni flagellin gene. J Clin Microbiol 39:754–757
13. Hsiao LL, Dangond F, Yoshida T, Hong R, Jensen RV, Misra J, Dillon W, Lee KF, Clark KE, Haverty P, Weng Z, Mutter GL, Frosch MP, Macdonald ME, Milford EL, Crum CP, Bueno R, Pratt RE, Mahadevappa M, Warrington JA, Stephanopoulos G, Stephanopoulos G, Gullans SR (2001) A compendium of gene expression in normal human tissues. Physiol Genomics 7:97–104
14. Durig J, Nuckel H, Huttmann A, Kruse E, Holter T, Halfmeyer K, Fuhrer A, Rudolph R, Kalhori N, Nusch A, Deaglio S, Malavasi F, Moroy T, Klein-Hitpass L, Duhrsen U (2003) Expression of ribosomal and translation-associated genes is correlated with a favorable clinical course in chronic lymphocytic leukemia. Blood 101:2748–2755
15. Guenzi E, Topolt K, Lubeseder-Martellato C, Jorg A, Naschberger E, Benelli R, Albini A, Sturzl M (2003) The guanylate binding protein-1 GTPase controls the invasive and angiogenic capability of endothelial cells through inhibition of MMP-1 expression. EMBO J 22:3772–3782
16. Lubeseder-Martellato C, Guenzi E, Jorg A, Topolt K, Naschberger E, Kremmer E, Zietz C, Tschachler E, Hutzler P, Schwemmle M, Matzen K, Grimm T, Ensoli B, Sturzl M (2002) Guanylate-binding protein-1 expression is selectively induced by inflammatory cytokines and is an activation marker of endothelial cells during inflammatory diseases. Am J Pathol 161:1749–1759
17. Hiki Y, Iyama K, Tsuruta J, Egami H, Kamio T, Suko S, Naito I, Sado Y, Ninomiya Y, Ogawa M (2002) Differential distribution of basement membrane type IV collagen alpha1(IV), alpha2(IV), alpha5(IV) and alpha6(IV) chains in colorectal epithelial tumors. Pathol Int 52:224–233
18. Shiozawa J, Ito M, Nakayama T, Nakashima M, Kohno S, Sekine I (2000) Expression of matrix metalloproteinase-1 in human colorectal carcinoma. Mod Pathol 13:925–933
19. Kim PJ, Plescia J, Clevers H, Fearon ER, Altieri DC (2003) Survivin and molecular pathogenesis of colorectal cancer. Lancet 362:205–209
20. Takemasa I, Yamamoto H, Sekimoto M, Ohue M, Noura S, Miyake Y, Matsumoto T, Aihara T, Tomita N, Tamaki Y, Sakita I, Kikkawa N, Matsuura N, Shiozaki H, Monden M (2000) Overexpression of CDC25B phosphatase as a novel marker of poor prognosis of human colorectal carcinoma. Cancer Res 60:3043–3050
21. Miwa N, Furuse M, Tsukita S, Niikawa N, Nakamura Y, Furukawa Y (2000) Involvement of claudin-1 in the beta-catenin/Tcf signaling pathway and its frequent upregulation in human colorectal cancers. Oncol Res 12:469–476
22. Li A, Varney ML, Singh RK (2001) Expression of interleukin 8 and its receptors in human colon carcinoma cells with different metastatic potentials. Clin Cancer Res 7:3298–3304
23. Mason IJ, Taylor A, Williams JG, Sage H, Hogan BL (1986) Evidence from molecular cloning that SPARC, a major product of mouse embryo parietal endoderm, is related to an endothelial cell 'culture shock' glycoprotein of Mr 43,000. EMBO J 5:1465–1472
24. Lewin AR, Reid LE, McMahon M, Stark GR, Kerr IM (1991) Molecular analysis of a human interferon-inducible gene family. Eur J Biochem 199:417–423
25. Lincoln AJ, Wickramasinghe D, Stein P, Schultz RM, Palko ME, De Miguel MP, Tessarollo L, Donovan PJ (2002) Cdc25b phosphatase is required for resumption of meiosis during oocyte maturation. Nat Genet 30:446–449
26. Dhanesuan N, Sharp JA, Blick T, Price JT, Thompson EW (2002) Doxycycline-inducible expression of SPARC/osteonectin/BM40 in MDA-MB-231 human breast cancer cells results in growth inhibition. Breast Cancer Res Treat 75:73–85
27. Behrens P, Brinkmann U, Wellmann A (2003) CSE1L/CAS: its role in proliferation and apoptosis. Apoptosis 8:39–44
28. Payne AM, Downes SM, Bessant DA, Plant C, Moore T, Bird AC, Bhattacharya SS (1999) Genetic analysis of the guanylate cyclase activator 1B (GUCA1B) gene in patients with autosomal dominant retinal dystrophies. J Med Genet 36:691–693
29. Puscas I, Coltau M, Baican M, Domuta G, Hecht A (2001) Calcium, carbonic anhydrase and gastric acid secretion. Physiol Res 50:359–364
30. Makela S, Kere J, Holmberg C, Hoglund P (2002) SLC26A3 mutations in congenital chloride diarrhea. Human Mutat 20:425–438
31. Sweadner KJ, Rael E (2000) The FXYD gene family of small ion transport regulators or channels: cDNA sequence, protein signature sequence, and expression. Genomics 68:41–56
32. Sly WS, Whyte MP, Sundaram V, Tashian RE, Hewett-Emmett D, Guibaud P, Vainsel M, Baluarte HJ, Gruskin A, Al-Mosawi Met al (1985) Carbonic anhydrase II deficiency in 12 families with the autosomal recessive syndrome of osteopetrosis with renal tubular acidosis and cerebral calcification. N Engl J Med 313:139–145
33. Schweinfest CW, Henderson KW, Suster S, Kondoh N, Papas TS (1993) Identification of a colon mucosa gene that is down-regulated in colon adenomas and adenocarcinomas. Proc Natl Acad Sci USA 90:4166–4170
34. Stoehlmacher J, Lenz HJ (2003) Implications of genetic testing in the management of colorectal cancer. Am J Pharmacogenomics 3:73–88