의학박사학위 논문

인간 직장암 조직에서 Microarray 기법을 이용한 유전자 발현에 관한 연구

Identification of Genes Differentially Expressed in Human Rectal Carcinoma with Microarray Analysis

2011년 2월

인하대학교 대학원

의학과 (외과 전공)

최 진욱
의학박사학위 논문

인간 직장암 조직에서 Microarray 기법을 이용한 유전자 발현에 관한 연구

Identification of Genes Differentially Expressed in Human Rectal Carcinoma with Microarray Analysis

지도교수 김 경 래

이 논문을 박사학위 논문으로 제출함
Identification of Genes Differentially Expressed in Human Rectal Carcinoma with Microarray Analysis

by

Jinwook Choi

THESIS
Submitted to the faculty of
INHA UNIVERSITY
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY

Medical School
February 2011
ACKNOWLEDGEMENTS

I would like to acknowledge the advice and guidance of Prof. Kyung-Rae Kim, the supervisor of this dissertation. I also thank the members of my graduate committee for their counsel and suggestions, especially Prof. Jun-Hyeog Jang for all his input and encouragement. I also thank the lab staff of Dept. of Biochemistry, Medical School of Inha University. Without their knowledge and assistance, this study would not have been successful.

I would like to thank my parents, my parents-in-law, especially my wife, Christine, and my beloved son, Terry for supporting and encouraging me to pursue this degree. Without my wife’s reassurance, I would not have finished the degree.
# TABLE OF CONTENTS

Acknowledgements .............................................................................. iv

Table of Contents ................................................................................ v

List of Figures ....................................................................................... vi

List of Tables ........................................................................................ vii

Abstract ................................................................................................ viii

Introduction ........................................................................................ 1

Methods ............................................................................................... 4

Patients and Total RNA Isolation ......................................................... 4

Microarray Analysis ............................................................................. 4

RNA Extraction and cDNA Synthesis ............................................... 6

Quantitative Real-Time RT-PCR ......................................................... 6

Cloning and Sequencing ..................................................................... 7

Statistical Analysis ............................................................................. 8

Results ................................................................................................ 11

Identification of Differentially Expressed Genes

in Rectal Carcinoma Based on Gene Expression Profiles .............. 11

Real-Time RT-PCR Validation of Microarray Data ......................... 12

Discussion .......................................................................................... 25

References .......................................................................................... 39
LIST OF FIGURES

Fig. 1 A schematic diagram of Affymatrix Gene Chip arrays

Analysis procedure ......................................................... 16

Fig. 2A REG1 on rectal tumor vs. matching normal rectal tissue

....................................................................................... 17

Fig. 2B MMP3 on rectal tumor vs. matching normal rectal tissue

....................................................................................... 18

Fig. 2C MMP1 on rectal tumor vs. matching normal rectal tissue

....................................................................................... 19

Fig. 2D DSC3 on rectal tumor vs. matching normal rectal tissue

....................................................................................... 20

Fig. 2E DPEP1 on rectal tumor vs. matching normal rectal tissue

....................................................................................... 21

Fig. 2F CLDN1 on rectal tumor vs. matching normal rectal tissue

....................................................................................... 22

Fig. 3A AQP8 on rectal tumor vs. matching normal rectal tissue

....................................................................................... 23

Fig. 3B OGN on rectal tumor vs. matching normal rectal tissue

....................................................................................... 24
LIST OF TABLES

Table 1  Clinical and histopathological data for the 13 rectal cancer patients..................................................9

Table 2  Primer sequences for real-time RT-PCR experiments

..................................................................................................................10

Table 3  Highly up-regulated genes in human rectal carcinoma

..................................................................................................................14

Table 4  Highly down-regulated genes in human rectal carcinoma

..................................................................................................................15
Abstract

Colon cancer and rectal cancer share many features and are often referred to as ‘colorectal cancer’. However, the issue whether colon and rectal cancer should be considered as a single or two distinct entities is still debated, and few studies have addressed genetic alterations of rectal cancer. Gene expression profile with use of microarray technology allows the investigation of the cell status on a molecular, genome-wide scale. The purpose of this study was to find novel gene(s) involved in the development of human rectal carcinoma. Using DNA microarrays, we identified 17 up-regulated and 14 down-regulated genes with at least 5-fold between the human rectal carcinomas and the adjacent normal rectal tissues. These included genes involved in extracellular matrix (ECM), breakdown of ECM, cell proliferation, differentiation, cell adhesion, metabolism, cell signaling, and transmembrane transport/channel. Furthermore, we validated the highly differential expression of 8 genes including 6 highly up-regulated genes (REG1, MMP3, MMP1, DSC3, DPEP1, and CLDN1) and 2 highly down-regulated genes (AQP8 and OGN) in human rectal tumor samples compared with their corresponding normal tissues. Our results showed the utility of gene expression analysis to study human rectal carcinoma, and we identified several genes that may play a
role as potential therapeutic targets in the future.
Introduction

People often regard colon cancer and rectal cancer as one disease, because one is the continuation of the other, and they are similar in morphology and configuration. However, there are some differences between colon and rectal cancer [1-6]. Despite clinical advances, rectal cancer remains a significant cause of cancer-related death [7], and few studies have addressed possible genetic alterations of rectal cancer. Currently tumor classification which is an important step in determining treatment and prognosis relies on pattern recognition, based on the morphology of the tumor and surrounding tissues. However, this methodology is not sufficient to classify all varieties of tumors accurately, and patients diagnosed with the same stage of cancer by conventional clinical and histopathological criteria may have a completely different course of disease.

Since cancer is fundamentally a malfunction of gene expression that gives rise to malignant growth, the most direct classification approach would be to analyse gene expression patterns. In the last two decades these mechanisms have been intensively studied [8, 9]. It is generally accepted that one of the initiating steps in colorectal carcinogenesis is mutation in APC tumor suppressor gene [10, 11]. APC is known to bind to a cell signaling/transcription factor β catenin. Activation of the Wnt pathway...
normally signals the association of β-catenin with members of the T-cek factor/lymphocyte-enhancer factor-1 (TCF/LEF-1) family [12, 13]. This complex can activate the transcription of a variety of target genes [14-16]. TGF-β and its binding to TFG-β receptors cause phosphorylation of SMAD2 and SMAD3, two transcription factors that then bind to and activate SMAD4 [17, 18]. These complexes then activate a series of TGF-β responsive genes, typically the cell-cycle checkpoint genes [19]. Therefore the TGF-β pathway acts as a tumor suppressor pathway. Traditional methods of identifying novel targets that are involved in colon cancer progression are based on the studies of individual genes. To find the relatively small number of genes that are characteristically deregulated in a given cancer cell among thousands of genes that are normally expressed requires high-throughput technologies and sophisticated computational tools. The first high density microarrays were developed to analyse gene expression by quantitating thousands of mRNAs present in a cell or tissue sample (DNA arrays). DNA arrays are the most advanced and most commonly used [20]. Recent studies suggest that DNA microarray profiling performed on clinical specimens may provide information directly applicable to cancer diagnosis [21, 22].

We applied a microarray-based gene expression profiling approach on 13 human rectal carcinoma samples to identify molecular signatures that
distinguish rectal carcinoma. Elucidation of such molecular expression
signatures may be useful in predicting the clinical behavior of rectal
carcinoma as well as identifying candidate cellular pathways that can be
targets for future therapeutic approaches.
Methods

Patients and total RNA isolation

Adult human rectal tissues were obtained from surgically removed, discarded tissue after diagnostic tests were completed. The tumors were from 13 adult patients (8 males and 5 females, mean age at operation = 65.4 years, range = 35-77) who were operated at the department of Surgery, Inha University Hospital, Incheon, Korea (Table 1). The request for acquisition of the rectal tissues was approved by the Institutional Review Board (IRB), Human Subjects Protection Office at the Inha University Hospital. Total RNA was extracted from frozen tissue samples using Easy-spin RNA Extraction kit (iNtRON, Seoul, Korea) according to the manufacturer’s instruction. RNA quality was assessed by Agilent 2100 bioanalyser using the RNA 6000 Nano Chip (Agilent Technologies, Amstelveen, Netherlands), and quantity was determined by ND-1000 Spectrophotometer (NanoDrop Technologies, Inc., DE, USA).

Microarray analysis

Total RNA was repurified with RNeasy MinElute kit per manufacturer’s
instructions (Qiagen, Valencia, CA). The total RNA (5 μg) was then used for GeneChip analysis. Preparation of cRNA, hybridization to human GeneChip® Human Gene 1.0 ST arrays (Affymetrix, Santa Clara, CA, USA) and scanning of the arrays was carried out according to manufacturer’s protocols (https://www.affymetrix.com). Briefly, 300 ng of total RNA from each sample was converted to double-strand cDNA. Using a random hexamer incorporating a T7 promoter, amplified RNA (cRNA) was generated from the double-stranded cDNA template through an IVT( in-vitro transcription) reaction and purified with the Affymetrix sample cleanup module. cDNA was regenerated through a random-primed reverse transcription using a dNTP mix containing dUTP. The cDNA was then fragmented by UDG and APE 1 restriction endonucleases and end-labeled by terminal transferase reaction incorporating a biotinylated dideoxynucleotide. Fragmented end-labeled cDNA was hybridized to the GeneChip® Human Gene 1.0 ST arrays for 16 hours at 45 °C and 60 rpm as described in the Gene Chip Whole Transcript (WT) Sense Target Labeling Assay Manual (Affymetrix, Santa Clara, CA, USA). After hybridization, the chips were stained and washed in a Genechip Fluidics Station 450(Affymetrix, Santa Clara, CA, USA) and scanned by using a Genechip Array scanner 3000 7G (Affymetrix, Santa Clara, CA, USA).
RNA extraction and cDNA synthesis

Total RNA was extracted using Easy-spin RNA Extraction kit (iNtRON, Seoul, Korea) according to the manufacturer’s instruction. The purity of RNA was assessed by absorption at 260 and 280 nm (values of the ratio of $A_{260}/A_{280}$ of 1.9–2.1 were considered acceptable) and by ethidium bromide staining of 18 S and 28 S RNA on gel electrophoresis. RNA concentrations were determined from the $A_{260}$. Two micrograms of total RNA was reverse-transcribed in a 20-μl reaction mixture containing 50 units of SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA, USA), 5 μM DTT, 40 units of RNaseOUT recombinant ribonuclease inhibitor, 0.5 μM of random hexanucleotide primers, and 500 μM of dNTP mixture. The reverse transcription reaction was carried out at 50 °C for 60 min. Heating the reaction mixture at 70 °C for 15 min was subsequently performed to terminate the reaction, and the cDNA was stored at -20 °C.

Quantitative real-time RT-PCR

Overexpression of the obtained genes was confirmed by quantitative real-time RT-PCR. All real-time PCR analyses were performed on an ABI Step One realtime PCR system. Each reaction included a 20-μl reaction mixture
containing 0.1 μM of each primer, 10 μl of 2× SYBR Green PCR master mix (Applied Biosystem, including AmpliTaq Gold DNA polymerase with buffer, dNTPs mix, SYBR Green I dye, Rox dye, and 10 mM MgCl$_2$), and 1 μl of the template cDNA. The typical amplification program included activation of the enzyme at 94 °C for 10 min, followed by 40 cycles of denaturation at 94 °C for 15 s, and annealing and extension at 60 °C for 1 min. The $C_T$ (cycle threshold) value for each gene was determined by automated threshold analysis function of the ABI instrument and normalized to $C_{T(G3PDH)}$ to obtain $dC_T (= C_{T(G3PDH)} - C_{T(test)})$. The difference of $n$ between two $C_T$ or $dC_T$ values indicates a $2^n$-fold difference in amount of the target sequence between the two cDNA samples being compared. The primers used in the quantitative PCR were shown in Table 2.

**Cloning and sequencing**

10 ng of PCR products were cloned into plasmids pGEM-T Easy Vector™ (Promega, Madison, WI, USA) and transformed to competent *E. coli* DH5α competent cells. 64 colonies were randomly picked up and sequenced using the PRISM dye termination kit™ (Applied Biosystems, Foster City, CA, USA). BLAST Search 2.0 (www.ncbi.nlm.nih.gov/blast/blast.cgi) was used to analyse sequence
homologies in the gene database.

**Statistical analysis**

The Student’s *t* test was used to establish the statistical significance of differences in the expression of each target gene between the human rectal carcinomas and the adjacent normal rectal tissues. Results with *P* values of <0.05 were considered statistically significant.

The detection (Present/Absent) call was generated by the Affymetrix microarray suite 5(MAS5) algorithm. The scanned raw files were imported into the statistical programming environment R (Version 2.3) for further analysis with tools available from the Bioconductor Project (http://www.bioconductor.org).

Expression data were normalized and log2 transformed using the robust multichip average (RMA) method implemented in the Bioconductor package RMA (M2, M3). To reduce noise for the significance analysis, probe sets that did not show detection call rate at least 50% of the samples in the comparison were filtered out. Highly expressed genes that showed a 5-fold change in expression were selected. The results were classified using hierarchical clustering algorithms (Eisen et al., 1998) implemented in TMEV software 4.0. (www.tigr.org/software) (M4).
Table 1

Clinical and histopathological data for the 13 rectal cancer patients

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Age (yr)</th>
<th>Sex</th>
<th>AJCC TNM</th>
<th>Tumor size (cm)</th>
<th>Tumor differentiation</th>
<th>Lymphatic invasion</th>
<th>Venous invasion</th>
<th>Perineural invasion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65</td>
<td>M</td>
<td>T3N0M0</td>
<td>3.2</td>
<td>moderate</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>77</td>
<td>M</td>
<td>T3N0N0</td>
<td>4.5</td>
<td>moderate</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>65</td>
<td>F</td>
<td>T2N0M0</td>
<td>1.2</td>
<td>moderate</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>71</td>
<td>M</td>
<td>T2N0M0</td>
<td>5.4</td>
<td>moderate</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>F</td>
<td>T3N1M1</td>
<td>5.0</td>
<td>moderate</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>79</td>
<td>M</td>
<td>T2N0M0</td>
<td>5.5</td>
<td>moderate</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>69</td>
<td>F</td>
<td>T2N0M0</td>
<td>3.5</td>
<td>moderate</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>61</td>
<td>F</td>
<td>T3N0M0</td>
<td>3.5</td>
<td>moderate</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>9</td>
<td>62</td>
<td>M</td>
<td>T3N0M0</td>
<td>4.0</td>
<td>moderate</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>73</td>
<td>M</td>
<td>T2N0M0</td>
<td>6.0</td>
<td>moderate</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>35</td>
<td>M</td>
<td>T4bN0M0</td>
<td>7.5</td>
<td>moderate</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>74</td>
<td>F</td>
<td>T3N0M0</td>
<td>10.5</td>
<td>moderate</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>75</td>
<td>M</td>
<td>T3N1M1</td>
<td>6.5</td>
<td>moderate</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
**Table 2**

*Primer sequences for real-time RT-PCR experiments*

<table>
<thead>
<tr>
<th>Genes</th>
<th>Sense (5’-3’)</th>
<th>Anti-sense (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>REG1</td>
<td>TCCTCCCTGATGTTCTGTC</td>
<td>ATAGGCATTTGTCCTTCTG</td>
</tr>
<tr>
<td>MMP3</td>
<td>GAAGCTGGACTCCGACACTC</td>
<td>TGCCAGGAAAGTTCTGAAG</td>
</tr>
<tr>
<td>MMP1</td>
<td>GGTCTCTGAGGGTCAAGCAG</td>
<td>CAAGATTTCTCCAGGTCAG</td>
</tr>
<tr>
<td>DSC3</td>
<td>AGACCTCATCCGGTCAAGTG</td>
<td>TTATCAGACAGCGCAACAG</td>
</tr>
<tr>
<td>DPEP1</td>
<td>ACCCGGAGACCTCTGGTTAT</td>
<td>CCAACTGCTTGCAATGGAG</td>
</tr>
<tr>
<td>CLDN1</td>
<td>CCCTCTGGGAGTGATAGCA</td>
<td>CCAACTGTCCGTCATGGAG</td>
</tr>
<tr>
<td>AQP8</td>
<td>TTTGTGCCATCTGATCTGGA</td>
<td>AACGTTGTCCTCAGGACAC</td>
</tr>
<tr>
<td>OGN</td>
<td>TGGAATCCGTGCCTTTAAT</td>
<td>TGGTGTCATTAGCCTTGAC</td>
</tr>
<tr>
<td>GAPDH</td>
<td>TGGAAGGACTCATGACCACA</td>
<td>TCCAGCTCAGGGATGACCTT</td>
</tr>
</tbody>
</table>
Results

Identification of differentially expressed genes in rectal carcinoma based on gene expression profiles

To clarify relevant alterations of gene expression associated with human rectal carcinogenesis, we analyzed the gene expression profiles of human rectal carcinomas compared with the adjacent normal rectal tissue samples by an Affymetrix Gene Chip arrays containing more than 28,869 genes. From the results of Affymatrix Gene Chip arrays analysis, we detected 31 genes differentially expressed with at least 5-fold between the human rectal carcinomas and the adjacent normal rectal tissues. Of these, 17 were significantly up-regulated and 14 significantly down-regulated in human rectal carcinoma versus adjacent normal rectal tissues (Table 3 and 4).

Several genes involved in the extracellular matrix (ECM) or breakdown of ECM (MMP3, MMP1, MMP10, THBS2, TNC, COL11A1, MMP7) were expressed in higher levels in human rectal carcinoma. Genes involved in regulation of cell proliferation and differentiation (REG1, REG3A, and INHBA) were also expressed in increased levels. Furthermore, several genes implicated in cell adhesion
(DSC3, CLDN1), metabolism (DPEP1, DUOX2, and SFRP) and cell signaling (SPINK4 and IFITM1) had increased expression in human rectal carcinomas compared with the adjacent normal rectal tissue samples.

In contrast, 14 genes had reduced expression in human rectal carcinomas compared with the adjacent normal rectal tissue samples. There was an overrepresentation of members of genes involved in transmembrane transport/channel (AQP8, BEST4, SLC15) and metabolism (CA7, CPB1, TPH1, and MT1M). Table 4 lists the names and biological functions of genes expressed in reduced levels with a fold difference > 5. This group of highly differentially expressed genes may be useful as molecular tumor markers that can potentially be used for more accurate diagnosis, prognosis and possibly can serve as drug targets for effective therapies.

**Real-time RT-PCR validation of microarray data**

To confirm the results obtained using microarrays, the expression levels of mRNA of 8 known genes were examined by quantitative real-time RT-PCR analysis in 13 pairs of human rectal tumor samples and their corresponding normal tissues. These genes included 6 highly up-
regulated genes (REG1, MMP3, MMP1, DSC3, DPEP1, and CLDN1) and 2 highly down-regulated genes (AQP8 and OGN) in human rectal tumor samples. To accurately quantify the expression of the eight genes, GAPDH was amplified and the mean data in each case was used to normalize the result. As shown in Fig. 2 and 3, RT-PCR analysis performed on 13 pairs of human rectal tumor samples and their corresponding normal tissues demonstrated significant expression differences, of which REG1 ($P \leq 0.043$), MMP3 ($P \leq 0.001$), MMP1 ($P \leq 0.004$), DSC3 ($P \leq 0.0025$), DPEP1 ($P \leq 0.0024$), and CLDN1 ($P \leq 0.0004$) were up-regulated in human rectal tumor tissues compared with their corresponding normal tissues, whereas AQP8 ($P \leq 0.0005$), and OGN ($P \leq 0.00028$) were down-regulated in human rectal tumor tissues compared with their corresponding normal tissues. These results indicated that the real-time RT-PCR results were highly consistent with the microarray data.
### Table 3

*Highly up-regulated genes in human rectal carcinoma*

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene description</th>
<th>Fold Change</th>
<th>GenBank ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>REG1</td>
<td>Regenerating islet-derived 1</td>
<td>178.5</td>
<td>NM006507</td>
</tr>
<tr>
<td>MMP3</td>
<td>Matrix metallopeptidase 3</td>
<td>63.5</td>
<td>NM002422</td>
</tr>
<tr>
<td>MMP1</td>
<td>Matrix metallopeptidase 1</td>
<td>25.1</td>
<td>NM002421</td>
</tr>
<tr>
<td>DSC3</td>
<td>Desmocollin 3</td>
<td>21.1</td>
<td>NM024423</td>
</tr>
<tr>
<td>DPEP1</td>
<td>Dipeptidase 1</td>
<td>14.6</td>
<td>NM004413</td>
</tr>
<tr>
<td>CLDN1</td>
<td>Claudin 1</td>
<td>14.5</td>
<td>NM021101</td>
</tr>
<tr>
<td>SFRP4</td>
<td>Secreted frizzled-related protein 4</td>
<td>13.8</td>
<td>NM003014</td>
</tr>
<tr>
<td>DUOX2</td>
<td>Dual oxidase 2</td>
<td>12.0</td>
<td>NM014080</td>
</tr>
<tr>
<td>MMP10</td>
<td>Matrix metallopeptidase 10</td>
<td>11.9</td>
<td>NM002425</td>
</tr>
<tr>
<td>SPINK4</td>
<td>Serine peptidase inhibitor, Kazal type 4</td>
<td>10.9</td>
<td>NM014471</td>
</tr>
<tr>
<td>INHBA</td>
<td>Inhibin, beta A</td>
<td>8.3</td>
<td>NM002192</td>
</tr>
<tr>
<td>COL11A1</td>
<td>Collagen, type XI, alpha 1</td>
<td>7.1</td>
<td>NM001854</td>
</tr>
<tr>
<td>THBS2</td>
<td>Thrombospondin 2</td>
<td>6.6</td>
<td>NM003247</td>
</tr>
<tr>
<td>TNC</td>
<td>Tenascin C</td>
<td>6.5</td>
<td>NM002160</td>
</tr>
<tr>
<td>REG3A</td>
<td>Regenerating islet-derived 3 alpha</td>
<td>6.3</td>
<td>NM138938</td>
</tr>
<tr>
<td>IFITM1</td>
<td>interferon induced transmembrane protein 1</td>
<td>5.1</td>
<td>NM003641</td>
</tr>
<tr>
<td>MMP7</td>
<td>Matrix metallopeptidase 7</td>
<td>5.0</td>
<td>NM002423</td>
</tr>
</tbody>
</table>
Table 4

*Highly down-regulated genes in human rectal carcinoma*

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene description</th>
<th>Fold Change</th>
<th>GenBank ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQP8</td>
<td>Aquaporin 8</td>
<td>-22.6</td>
<td>NM001169</td>
</tr>
<tr>
<td>OGN</td>
<td>osteoglycin</td>
<td>-9.5</td>
<td>NM033014</td>
</tr>
<tr>
<td>CA7</td>
<td>carbonic anhydrase VII</td>
<td>-8.0</td>
<td>NM005182</td>
</tr>
<tr>
<td>BEST4</td>
<td>bestrophin 4</td>
<td>-8.0</td>
<td>NM153274</td>
</tr>
<tr>
<td>UGT2B15</td>
<td>UDP glucuronosyltransferase 2 family polypeptide B15</td>
<td>-8.0</td>
<td>NM001076</td>
</tr>
<tr>
<td>CPB1</td>
<td>carboxypeptidase B1</td>
<td>-7.4</td>
<td>NM001871</td>
</tr>
<tr>
<td>SST</td>
<td>somatostatin</td>
<td>-6.9</td>
<td>NM001048</td>
</tr>
<tr>
<td>PLP1</td>
<td>proteolipid protein 1</td>
<td>-6.9</td>
<td>NM000523</td>
</tr>
<tr>
<td>HOXD13</td>
<td>Homeobox D13</td>
<td>-6.5</td>
<td>NM002425</td>
</tr>
<tr>
<td>SLC15</td>
<td>Solute carrier family 15</td>
<td>-6.5</td>
<td>NM005073</td>
</tr>
<tr>
<td>TPH1</td>
<td>Tryptophan hydroxylase 1</td>
<td>-6.5</td>
<td>NM004179</td>
</tr>
<tr>
<td>ANGPTL1</td>
<td>Angiopoietin-like 1</td>
<td>-6.1</td>
<td>NM004673</td>
</tr>
<tr>
<td>MT1M</td>
<td>Metallothionein 1M</td>
<td>-5.6</td>
<td>NM176870</td>
</tr>
<tr>
<td>KIF5C</td>
<td>Kinesin family member 5C</td>
<td>-5.2</td>
<td>NM004522</td>
</tr>
</tbody>
</table>
Fig. 1  A schematic diagram of Affymatrix Gene Chip arrays analysis procedure (as given in the protocols of Affymatrix, Inc., Santa Clara, USA)
Fig. 2  Quantitative real-time RT-PCR analysis of the highly up-regulated genes (REG1, MMP3, MMP1, DSC3, DPEP1, and CLDN1) on 13 rectal carcinoma samples and matching normal rectal samples. The box plot analysis shows the median, 25th, and 75th percentiles. Differences in gene expression levels between the two groups were analysed by Student’s $t$ test.

**Fig. 2A** REG1 on rectal tumor vs. matching normal rectal tissue
Fig. 2B  MMP3 on rectal tumor vs. matching normal rectal tissue
Fig. 2C  MMP1 on rectal tumor vs. matching normal rectal tissue
Fig. 2D  DSC3 on rectal tumor vs. matching normal rectal tissue
Fig. 2E  DPEP1 on rectal tumor vs. matching normal rectal tissue
Fig. 2F  CLDN1 on rectal tumor vs. matching normal rectal tissue

CLDN1

p=0.0004

Relative expression of CLDN1

Normal  Tumor
Fig. 3  Quantitative real-time RT-PCR analysis of the highly down-regulated genes (AQP8 and OGN) on 13 rectal carcinoma samples and matching normal rectal samples. The box plot analysis shows the median, 25th, and 75th percentiles. Differences in gene expression levels between the two groups were analysed by Student’s t test.

Fig. 3A  AQP8 on rectal tumor vs. matching normal rectal tissue
Fig. 3B  OGN on rectal tumor vs. matching normal rectal tissue
Discussion

At a molecular level, much progress has been made in the last two decades in the identification and characterization of the genetic changes involved in the malignant colorectal transformation process. The model proposed [23] foresees a colon cancer where the temporary progression from healthy mucosa to carcinoma *in situ* is supported by mutations in APC, K-ras, TP53, and DCC genes. The model, originally formulated for sporadic colorectal cancer development, is also valid for familial adenomatous polyposis patients, who carry an APC germline mutation [24, 25]. A second pathway of colorectal tumorigenesis has been depicted in cases with a normal karyotype but carrying genetic instability at microsatellite loci attributable to alterations in the DNA mismatch repair genes. The latter, when present in germinal cells, are responsible for the familial syndrome hereditary nonpolyposis colorectal carcinoma [24, 26, 27] and when affecting somatic cells may cause microsatellite instability (MIN) in a subset of sporadic colorectal tumors [24, 28]. According to these data, the stepwise progression postulated by the aforesaid model [23] seems to be representative of colorectal tumor development in about 90% of sporadic colorectal cancers and indirectly supported by a number of epidemiological,
clinical, histopathological, and genetic studies [29, 30]. However, the subsequent discovery of several other genes involved in colorectal tumor development supports the notion that the model of Kinzler and Vogelstein could be more complex than originally proposed [31]. In a recent study [32], multiple alternative genetic pathways were shown to lead to tumor progression.

The issue of whether colon and rectal cancer should be considered as a single entity or two distinct entities is still debated. Despite clinical evidence of differences between colon and rectum cancers in terms of metastatic sites, treatment modalities, and outcome, few studies have addressed to molecular and/or biological differences between the two diseases and generally have focused on a single marker [33, 34]. cDNA microarray is the most commonly used technique now for surveying many samples with thousands of genes. Gene expression data from DNA microarray provide a snapshot of the molecular status of a sample of cells in a given tissue, returning the expression levels of thousands of genes simultaneously. They make it possible to analyze the genes involved in a particular type of cancer [35] as well as the classification of tumor specimens in different categories [36]. In an effort to use gene expression profiles to find the molecular marker of rectal cancer, we used cDNA microarray and found 17 up-regulated and 14 down-
regulated genes.

The canonical Wnt signaling cascade controls cell behavior by steering the transcriptional properties of DNA-binding proteins of the TCF/LEF-1 family. At the heart of the canonical Wnt pathway is the stabilization of cytosolic $\beta$-catenin, which activates target genes by binding to TCF/LEF-1 family. In the absence of Wnts, $\beta$-catenin is phosphorylated by casein kinases 1$\alpha$ at serine residue 45; this in turn enables glycogen synthase kinase 3$\beta$ (GSK3$\beta$) to phosphorylate serine/threonine residues 41, 37, and 33. Phosphorylation of these last two amino acids triggers ubiquitination of $\beta$-catenin and degradation in proteasomes. Phosphorylation of $\beta$-catenin occurs in a multiprotein complex containing the scaffold protein axin and/or its homologue axin2/conductin, the tumor suppressor gene product APC, GSK3$\beta$, and diversin. In the presence of Wnts, the cytoplasmic component dishevelled blocks $\beta$-catenin degradation by a largely unknown mechanism. Stabilized $\beta$-catenin enters the cell nucleus and associates with TCF transcription factors, leading to the transcription of Wnt target genes. A variety of Wnt/$\beta$-catenin target genes have been identified which include regulators of cell proliferation, developmental control genes, and genes implicated in tumor progression. In tumors, either overexpression of Wnts (which is uncommon in human tumors)
or mutations in one the components responsible for the degradation of β-catenin, leads to stabilization of β-catenin and activation of target genes. Of note, β-catenin is also involved in the control of cell-cell adhesion by binding to cadherin cell adhesion molecules and providing a link to the actin cytoskeleton. It remains unsolved as to what degree this cell adhesion function of β-catenin also plays a role in Wnt signaling. Up-regulation of COL11A1, DSC3, CLDN1 and IFITM1 are known to be related to Wnt pathway. The up-regulation of COL11A1 has been regarded as a consequence of Wnt pathway activation [37, 38]. Furthermore, Kahler et al. [38] suggest an indirect role of LEF-1 in regulating COL11A1 expression. Khan K et al. [39] showed that DSC2 protein expression is reduced in colorectal cancer, and that this is accompanied by de novo expression of DSC1 and DSC3. Alterations in the expression patterns of desmosomal cadherins in cancer could result in the release of plakoglobin from desmosomes, subsequent displacement of β-catenin from adherens junctions and increased Wnt/β-catenin signalling [40]. These data raise the possibility that modulation of desmosomal cadherin expression in cancer could stimulate transcription of β-catenin target genes. Another possible outcome of the changes is that release of plakoglobin from desmosomes could result in its translocation to the nucleus. Whether
this would promote proliferation (by stimulating transcription of genes such as c-myc and Bcl-2) or have the opposite effect (by acting as a negative regulator of Wnt/β-catenin signalling) remains uncertain. In general, the overall down-regulation of cell-cell adhesion molecules occurs during carcinoma development [41], and evidence of either a reduced expression or a loss of expression of the CLDN family members has been found to promote cell invasion and metastasis in malignant tumors, including cancer of the gastrointestinal tract [42, 43], pancreas [44], breast [45] and colorectal cancer [46]. However, the up-regulated expression of CLDN1 in colorectal cancer has been reported [47-50], suggesting that the CLDN family, especially CLDN1, plays a causal role in the process of cellular transformation and invasion in colorectal cancer. It is possible that CLDN protein interacts with signaling pathways, including TGFβ/SMAD and β-catenin [50, 51], in a manner that is separate from their effects on the barrier function. The previous observation that CLDN1 is regulated by β-catenin/TCF/LEF-1 signaling [50, 52-54] suggests that the increased CLDN1 expression observed may be due to β-catenin activation. IFITM1 is commonly considered as a protein induced by interferon, and it is involved in signal transduction in our immune system. However, with the improved understanding of the natural immune response to cancer, the
roles of interferons in carcinogenesis are different from previously recognized suggesting oncogenic function. IFITM1 has been shown to be overexpressed in gastric cancer cells, which resulted in tumor cells being more resistant to natural killer cells and produced a more invasive phenotype [55]. Andreu et al. [56] demonstrated up-regulated expression of IFITM genes (IFITM1, IFITM2, and IFITM3) on adenomas in a murine model as well as a human colorectal carcinoma cell line, revealing that IFITM gene expression was regulated by the β-catenin signaling.

sFRPs are antagonists that bind directly to Wnts and prevent them from binding to their receptors [57]. Hypermethylation of the promoter region of sFRP genes occurs frequently in colorectal tumors and is associated with the transcriptional silencing of these genes [58, 59]. Tumor cells may gain a selective advantage by shutting down the expression of normally proapoptotic sFRPs [60, 61]. Like its sFRP relatives, sFRP4 binds to both the frizzled receptor and Wnt ligands to antagonise the Wnt signalling pathway [62]. Later on, sFRP4 overexpression was found in primary prostate carcinomas [63], endometrial stromal sarcomas [64] and colorectal carcinomas [65]. The molecular mechanisms responsible for the overexpression of sFRP4 and the effect of this overexpression on tumors are not well studied.
However, increased levels of sFRP4 in tumor samples are evidence against the hypothesis that sFRP4 functions as a tumor suppressor in these tumor models. Like sFRP4, sFRP1 overexpression has also been reported in uterine leiomyomas, where it appears to have stimulated cell growth [66].

TGF-β signalling is initiated via the binding of TGF-β to type II TGF-β receptors (TGFBR2). This, then leads to recruitment via phosphorylation of the type I TGF-β receptor (TGFBR1) leading to activation of TGFBR1 protein kinase. This causes phosphorylation of SMAD2 and SMAD3, two transcription factors that then bind to and activate SMAD4. These complexes then migrate to the nucleus where they activate a series of TGF-β responsive genes [17, 18]. These genes typically include the cell-cycle checkpoint genes CDKN1A (p21), CDKN1B (p27) and CDKN2B (p15), which when activated cause cell cycle arrest [19]. Therefore the TGF-β pathway acts as a tumor suppressor pathway in normal colonic epithelium. In studies of colorectal carcinomas, mutations of the SMAD4 and SMAD2 genes have been observed [67-69]. Conversely, it has been found that at the later stages of colorectal cancer development, the TGF-β pathway actually acts to promote invasion and metastasis, shown in several experimental models where colonic epithelial cells were exposed to
high levels of TGF-β which induced malignant transformation [70] as well as invasion causing metastasis [71]. This is thought to be due to the fact that TGF-β regulates the production of growth factors including TGF-β, FGF and EGF [18], as well as the fact that tumor cells at an advanced state of development become resistant to the inhibitory effect of the TGF-β pathway. Activin A, the disulfide-linked homodimer of INHBA, is a ligand in the TGFβ superfamily [72, 73]. The possibility that increased activin expression may be tumorigenic under certain circumstances is supported by the findings that pancreatic, prostate, ovarian, and colon cancers overexpress activin A [74-77], and that patients with endometrial and cervical cancers have high serum levels of activin A [78].

Human colon tumors have been shown to express MMP1, 2, 3, 7, 8, 9, 10, 11, 12, and 14 [79]. Both the number of different MMPs and the expression level of each MMP are increased as the tumors become more advanced. It has been known that MMP3 and MMP11 are expressed in the stromal component of several carcinomas and are not present in adenomatous tissue [79]. The same authors noted that MMP7 gene expression is an early event in colorectal tumorigenesis and that the expression of MMP3, MMP11, and gelatinase A is primarily a late event. Experimental studies have provided confirmation of the role of
MMP7 in colorectal cancer development and progression [80-83], finding that the induction of MMP7 expression in tumor cells increased tumorigenicity and metastases, while inhibition of MMP7 led to decreased metastasis. The clinical impact of MMP7 expression has been assessed. Analysis of MMP7 as a prognostic indicator revealed that positive MMP7 expression correlates with depth of invasion, lymph node metastasis, lymphatic invasion, advanced Dukes’ stage, and poor outcome [84, 85]. MMP1 expression has been correlated with tumor progression, grade, presence of lymph node metastasis, and poor prognosis [86]. Normal colorectal epithelium and colorectal adenomas do not express MMP1, however, immunoreactivity for MMP1 is detected in 76% of colorectal adenocarcinomas [86]. TNC is a ECM glycoprotein which is transiently expressed in various developing tissues [87]. In malignancies, its role is presumed to be connected with cell adhesion and detachment, cell growth and migration, promotion of angiogenesis and inhibition of the immune response [88-91]. In colonic tissue, TNC immunoreactivity is almost undetectable in normal adult mucosa [92]. An increased TNC is usually seen in inflammatory and hyperplastic conditions of the colorectum, while an intense staining has been reported in invasive colorectal carcinoma [93-97]. According to data from cell culture experiments [89, 90], TNC enhances tumor cell
proliferation and promotes angiogenesis, and it may support escape from tumor immunosurveillance. However, in contrast to these cancer-promoting functions, it has been proposed that TNC may also act to prevent tumor invasion and metastasis and consequently confer a better clinical outcome [98-101]. The prognostic value of TNC expression in colon cancer is controversial. Some reports indicated that strong TNC expression in colon cancer may be associated with a good prognosis [93, 102]. However, high expression of TNC in colorectal cancer has directly or indirectly been correlated with a poor patient prognosis as well [103, 104].

DPEP1 is a zinc-dependent metallopeptidase that hydrolyses a variety of dipeptides and is involved in glutathione metabolism [105]. Several previous reports demonstrated an association between DPEP1 gene expression and colorectal cancer, showing that DPEP1 gene expression in colorectal cancer was significantly higher than that in normal colorectal mucosa [106, 107]. However, the relationship between clinicopathological findings and DPEP1 gene expression in colorectal cancer was not clarified. In fact, there was no statistical difference in the expression of DPEP1 at each stage classification, but a tendency for decreased DPEP1 expression as the disease progressed was observed [107]. Two hypotheses can account for the phenomenon
that the expression of DPEP1 is attenuated in advanced colorectal cancer. First, DPEP1 is implicated in the metabolism of glutathione [108], an important antioxidant. An attenuated expression of DPEP1 may result in decreased production of glutathione, causing increased oxidative stress which increases tumor progression [109]. Second, since DPEP1 is a membrane-bound dipeptidase, it may play a role in the degradation of surrounding ECM components, which would aid in the ability of tumor cells to invade from the primary site through the extracellular matrix [107].

We found that AQP8 gene expression was down-regulated in the recent study. Water crosses the plasma membrane in one of two ways; directly through the lipid bilayer (a slow, unregulated process) or via protein water channels termed AQPs. 13 mammalian homologs (AQPs 0–12) have been identified and are important during rapid water movement in cells [110, 111]. During cell apoptosis, one of the earliest morphological events is pronounced cell shrinkage termed the apoptotic volume decrease (AVD) [112]. During the AVD decreased intracellular K+ is required for the activation of the apoptotic caspase cascade [113, 114]. Changes in ion concentration in apoptotic cells also create an osmotic gradient for water to follow resulting in the characteristic cell shrinkage. Water movement during the AVD is
mediated primarily via AQP8s and, inhibition of AQP-dependent water movement inhibits the AVD and downstream apoptotic cascades while AQP over-expression increases plasma membrane water permeability and the rate of apoptosis [115]. Jablonski et al. [116] demonstrated that AQP8 and 9 expression is decreased in hepatocellular carcinoma versus normal liver. They confirmed a lack of water movement across the cell membrane via AQPs in isolated hepatic tumor cells and identified a resistance to apoptotic stimuli. Reduced AQP8 expression in human colorectal cancer was reported by Fisher et al. [117].

CAs have an important role in maintaining the pH homeostasis by catalysing the reversible hydration of carbon dioxide. Five of the isozymes are cytosolic (CA 1, 2, 3, 7 and 13), four are membrane associated (CA 4, 9, 12 and 14), two are mitochondrial (CA VA and VB) and one is a secretory form (CA6) [118-122]. Previous studies have indicated that CA9 and 12 are overexpressed in certain tumors, including colorectal adenomas and carcinomas [123, 124]. Expression of CA9 has been associated with poor prognosis of lung, breast, head and neck and bladder cancer [125-128]. Cytosolic CAs could be linked to a hypothetical tumor suppressor function [124, 129]. In the previous studies [130, 131], several isozymes of carbonic anhydrases including 1, 2, 4, 7 and 12 were down-regulated implying a pathogenic role in
cancer development or progression.

OGN played a role in collagen fibrillogenesis, a process essential in development, tissue repair, and metastasis [132, 133]. OGN might contribute to lymphatic metastasis suppression through regulation of gelatinase activity which is suggested to play critical roles in tumor invasion and metastasis [134, 135]. It was shown that the tumor suppressor protein p53 activated transcription of OGN genes, and OGN expression was absent in different cancer cell lines and tumors where p53 was inactivated/mutated [136, 137].

SST inhibits multiple functions, including exocrine and endocrine secretions, inflammation, and angiogenesis, as well as cell proliferation and tumorigenesis [138-141]. The mechanisms of the inhibition in tumor cells are the combined interaction of SST and its analogs with SST1-5R, either directly inhibiting division and proliferation of tumor cells or inhibiting the activities of growth factors such as vascular endothelial growth factor (VEGF), insulinlike growth factor (IGF), etc [142-144], thus counteracting tumorigenesis and tumor cell proliferation. In addition, the ability of SST and its stable analogues to promote tumor cell apoptosis has been demonstrated in various cell types [143-145], however, the underlying mechanisms of SST induced cell apoptosis are still poorly understood.
In this study, we have found several genes which are up- or down-regulated in rectal cancer. Nowadays the molecular biology represents an endless source of data that may be filtered, elaborated and transformed into useful clinical information. The comprehensive knowledge on molecular background of cancers should be considered one of the main challenges of medical oncology guiding the oncologists in patient’s clinical management. Gene expression profile with microarray technologies is thought to have several potential clinical applications in rectal cancer in the future, ranging from study on mechanisms involved in tumor development to the identification of gene signatures.
References


26. Laurent-Puig, P., H. Blons, and P.H. Cugnenc, *Sequence of*


43. Ueda, J., et al., *Heterogeneous expression of claudin-4 in human...*


58. Suzuki, H., et al., *A genomic screen for genes upregulated by


66. Fukuhara, K., et al., Secreted frizzled related protein 1 is


80. Wilson, C.L., et al., *Intestinal tumorigenesis is suppressed in mice*


88. Gulubova, M. and T. Vlaykova, Immunohistochemical assessment of fibronectin and tenascin and their integrin receptors alpha5beta1 and


96. Natali, P.G., et al., Comparative analysis of the expression of the


104. Emoto, K., et al., Annexin II overexpression correlates with stromal


112. Maeno, E., et al., Normotonic cell shrinkage because of disordered volume regulation is an early prerequisite to apoptosis. Proc Natl Acad Sci U


141. Pichler, R., et al., *Somatostatin-receptor positive brain stem glioma*


