Characterization of Clinical and Molecular Features Related to the Down-regulated Expression of \( \mu \)-Protocadherin in Colorectal Cancer

Wybrich R Cnoosen Lorenza Losi\(^1\), Jean Benhattar\(^2\), Silvia Pizzini\(^3\), Andrea Bisognin\(^3\), Sandra Parenti\(^1\), Lucia Montorsi\(^1\), Claudia Gemelli\(^1\), Tommaso Zanocco-Marani\(^1\), Paola Zanovello\(^4,5\), Fabrizio Ferrari\(^6\), Sergio Ferrari\(^1\), Stefania Bortoluzzi\(^4\)\* and Alexis Grande\(^1\)*

\(^1\)Department of Life Sciences, Pathological Anatomy, University of Modena and Reggio Emilia, Modena, Italy
\(^2\)Institute of Pathology, University of Lausanne, Switzerland
\(^3\)Department of Surgery, Oncology and Gastroenterology, University of Padova, Italy
\(^4\)Istituto Oncologico Veneto (IOV), IRCCS, Padova, Italy

Abstract

Mu-protocadherin is a membrane protein belonging to the cadherin superfamily that, as the other members, promotes inter-cellular adhesion and proliferation arrest. Notably, both these onco-suppressive activities are mediated by its capacity to inhibit the \( \beta \)-catenin signaling pathway, that is constitutively activated in colorectal cancer (CRC) and, not surprisingly, its expression undergoes a remarkable down-regulation in CRC, although this finding has been to date only achieved in a limited set of patients. Based on this premise, the aims of our study were: (1) to confirm the down-regulated expression of \( \mu \)-protocadherin in a larger cohort of CRC patients; (2) to assess the possible relationship existing between the expression levels of this protein and the clinical-pathological parameters of the disease; (3) to identify the molecular mechanism underlying the decrease of its expression occurring in CRC. To address these issues, we developed a database containing more than one thousand CRC expression profiles, recovered by GEO (Gene Expression Omnibus) and supplied with a significant set of clinical information, that was subsequently used to perform a bioinformatic analysis of cadherin genes. A methylation analysis of \( \mu \)-protocadherin gene promoter was also carried out, using the bisulfite pyro-sequencing procedure, in an independent series of CRC samples. The results obtained confirmed the down-regulation of \( \mu \)-protocadherin expression in CRC, suggested a possible prognostic role for this alteration and indicated the promoter hyper-methylation of its gene as the responsible mechanism.

Keywords: Cadherin; Colorectal cancer; Hyper-methylation; Molecular mechanism

Introduction

Cell-cell adhesion is a fundamental activity in order to assure the maintenance of epithelial integrity and defects impairing this process participate to the formation of tumors, such as colorectal cancer (CRC). In this regard, a crucial role is played by adhesion molecules named cadherins that, at least in part, exert their functions through the inhibition of the \( \beta \)-catenin proliferation signaling pathway, constitutively activated in CRC. In fact, the E-cadherin/\( \beta \)-catenin complex has the capacity to suppress either the proliferation or the invasion of epithelial cells in normal conditions \(^1\) and down-regulation of E-cadherin expression facilitates the acquisition of these properties by cancer cells \(^2\).

Several reports, published over the last decade, have highlighted the existence of a novel cadherin family member that has been called \( \mu \)-protocadherin, to underline the hybrid nature of its extra-cellular region, characterized by the presence of both cadherin- and mucin-like domains \(^3\). Unlike other protocadherins, exhibiting a preferential expression in the nervous system, this protein is detectable in a number of different tissues including colorectal mucosa, where it is down-regulated during CRC tumor progression \(^4\) and up-regulated upon treatment with chemoprevention agents against this neoplastic condition \(^5\). These findings suggested that \( \mu \)-protocadherin determines an onco-suppressive effect on normal colon enterocytes, also supported by the observation that, as other cadherins, it is able to sequester and inhibit \( \beta \)-catenin \(^5-7\). In spite of the efforts made to characterize the role of \( \mu \)-protocadherin in CRC, many issues are still open: firstly, expression of \( \mu \)-protocadherin, so far evaluated in a limited set of samples, requires to be confirmed in a larger cohort of patients; secondly, the possible relationship existing between the expression levels of this protein and a number of clinical-pathological parameters needs to be investigated; finally, the molecular mechanism underlying the decrease of its expression, that is detected in CRC remains to be identified. In this regard, it is worth mentioning that previous reports published on other cadherins indicated the promoter hyper-methylation as the mechanism responsible for the loss of their expression \(^8\).

To clarify these aspects, we developed a database containing the microarray expression profiles of a considerable number of colorectal carcinoma, adenoma and normal mucosa samples, that had been recovered by GEO (Gene Expression Omnibus) and presented the advantage to be supplied with a significant set of clinical data such as stage, grade, survival and relapse. Bioinformatic analysis of this database, performed to assess all the available cadherin genes, confirmed that \( \mu \)-protocadherin undergoes a down-regulated

*Corresponding authors: Alexis Grande, Department of Life Sciences, University of Modena and Reggio Emilia, Via Campi 287, 41125, Modena, Italy, Tel: +39/0592055409; Fax +39/0592055410; E-mail: alexis.grande@unimore.it
Stefania Bortoluzzi, Department of Biology, University of Padova, Via G. Colombo 3, 35131, Padova, Italy, Tel: +39/0498276502; Fax: +39/0498276209; E-mail: stefania.bortoluzzi@unipd.it

Received November 30, 2013; Accepted January 13, 2014; Published January 14, 2014


Copyright: © 2014 Losi L, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
expression in CRC and highlighted, at the same time, that clinical cases experiencing a 5-year relapse are those in which this alteration is more pronounced. Methylation analysis, carried out in a distinct series of CRC samples using the bisulfite pyro-sequencing procedure, allowed the identification of two different CpG islands, placed inside the μ-protocadherin gene promoter and exhibiting a hyper-methylated pattern.

**Materials and Methods**

**Retrieval and annotation of microarray expression profile data**

A-MADMAN, a web application that allows the retrieval, annotation, organization and analysis of public available gene expression datasets [9], with the considerable advantage of integrating the signals deriving from different probes into custom probesets, providing a single and robust expression value for each analysed gene, was used to facilitate the meta-analysis of Affymetrix data contained in GEO. Using this tool, we downloaded from the NCBI database GEO, 1,795 experiments pertaining to 27 distinct series performed on biopsies of normal colon mucosa (Normal), colorectal adenoma (Adenoma) and CRC. The considered data were obtained from two different generations of Affymetrix gene chip technology that are not directly comparable, HG-U133A and HG-U133 plus2.0. Data belonging to the GEO series (CEL files and related metadata) were collected, and samples were manually assigned to the corresponding individual IDs. The raw data and metadata consisting of clinical and pathological information were used to build a local database. Metadata were associated to each sample as descriptive labels (tags) to subset samples into homogeneous groups. A strict quality control flow was applied to all samples raw data using a customized version of the ArrayQualityMetrics Bioconductor package.

**Platforms integration and gene expression matrix reconstruction**

Signal reconstruction was carried out applying the A-MADMAN with a workflow optimized to minimize inter-series and inter-studies variability biases. This is based on Robust Multichip Average (RMA) method [10] implemented in the Bioconductor R package affy, using GeneAnnot-based custom CDF [10,11]. Oligonucleotides common to all platforms are selected, and each chip is transformed into a virtual chip containing only the common probes. Of the three phases of RMA, background correction was applied for each platform before the selection of oligonucleotides, while summarization and quantile normalization were performed only on common probes.

**Differential expression analysis of microarray data**

Statistical analysis was performed using open source packages available from BioConductor for the R statistics environment (BioConductor, www.bioconductor.org).

**Immunohistochemical analysis of colorectal adenomas and carcinomas**

An independent series of 20 sporadic CRCs and 10 colorectal adenomas was selected through the files of the Pathological Anatomy of Modena. Clinical-pathological features of patients are summarized in Table 1. Immunoperoxidase staining was run with the Benchmark XT Automatic Staining System (Ventana, Strasbourg, France) using a rabbit polyclonal anti-human μ-protocadherin antibody (MUPCDH) (Sigma, St. Louis, USA) at dilution 1:100. The presence of immunoreactive neoplastic cells stained for μ-protocadherin was graded as previously described [4].

**Preparation of samples for methylation analysis**

CaCo2, HT29, HCT116 and Lovo cell lines were obtained from ATCC (Rockville, MD) and cultured as already described [5]. Normal colorectal mucosa samples were recovered from human admitted to the hospital to undergo surgical resection of CRC. Colorectal adenomas and carcinomas were obtained by formaldehyde-fixed and paraffin-embedded samples submitted to the manual microdissection, a laborious procedure (12) carried out to limit the analysis to the neoplastic tissue, always achieving a >90% homogeneity. DNA extraction was carried out with the DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA, USA).

**DNA methylation analysis**

Using the USCS Genome Browser we identified, inside the upstream region of the MUCDH1 gene, three putative CpG islands that were named CpG1, CpG2 and CpG3 and were respectively placed: around the translational start site; -2600 to -1800 bp from ATG; and -3200 to -2700 bp from ATG. Analysis of methylation was performed using the bisulfite pyrosequencing procedure that was carried out within small regions of the indicated Cpg islands as described [12,13]. PCR and pyrosequencing were performed at least in duplicates. The level of methylation observed at the analysed DNA region was graded as follows: 0-15%, absence of methylation; 16-35%, low methylation; 36-75%, moderate methylation; 76-100%, high methylation. The percentage of methylated CpG sites that was used as cutoff to decide whether a given DNA region was hypo- or hyper-methylated was 35%. Statistical analysis of methylation results, considered as mean percentage of methylated CpG nucleotides, was carried out according to the one-way Anova procedure followed by paired Tukey HSD test.

**Results**

**Data retrieval and creation of the GEO-derived CRC expression database**

The initial database, containing 1,795 experiments downloaded and annotated with the A-MADMAN web-tool, was subjected to the quality control step, obtaining 1,556 samples. Among the latter, we subsequently excluded undesired samples (cell lines, treated patients, metastatic tissues) thus selecting a number of 791 expression profiles obtained from 16 different GEO series (Table 2). This dataset included 670 primary untreated CRC samples associated to different clinical-pathological information that had been collected at the onset of the disease, such as stage (early [I or II] or late [III or IV]) and grade (low or high), or 5 years later, such as survival (alive, dead) and outcome (disease-free, relapse) (series GSE17536, GSE17537 [14], GSE12945 [15], GSE14333 [16]). A robustness test was carried out, in spite of their numerical variability, all the retrieved sample groups consisted of sizeable and non-redundant CRC subsets. Furthermore, datasets used for prognostic evaluations, exhibited a homogeneous composition in terms of patient staging and treatment.

**Analysis of MUPCDH and CDH1 expression in the CRC expression database**

As already anticipated in the Introduction section, by using a combined approach of microarray meta-analysis and immune-
normal to 7.87 of adenoma and to 7.04 of CRC and all the changes were
expression value of MUPCDH, reported as Log2, varied from 9.37 of
(Adenoma and CRC) as compared to normal. In fact, the median
conditions on x-axis and the expression levels on y-axis.
1 (MUPCDH) and 2 (CDH1) as bar plots, representing the analyzed
categories listed above. The results obtained are presented in Figures
described above. This analysis was conducted in Normal, used as a
database that we have
down-regulated expression in CRC, although this observation was
MUPCDH (the gene encoding for μ-protocadherin) undergoes a
expression of MUPCDH undergoes a
regulated Expression of μ-Protocadherin in Colorectal Cancer. J Carcinog Mutagen 5: 154. doi:
As shown in Figure 1, expression of MUPCDH undergoes a
significance in all the investigated neoplastic conditions
(Adenoma and CRC) as compared to normal. In fact, the median
expression value of MUPCDH, reported as Log2, varied from 9.37 of
normal to 7.87 of adenoma and to 7.04 of CRC and all the changes were
statistically significant (p<0.001).
Analysis of MUPCDH expression in the CRC subsets associated to
specific clinical-pathological categories revealed a decrease of signal
in all the considered conditions, the only exception being represented
by stage. More in detail, the Log, median expression value was 7.04 for
early stage compared to 7.02 of late stage, 7.50 for low grade compared to
6.99 of high grade, 7.99 for alive compared to 7.67 of dead patients
and 7.66 in disease-free patients compared to 7.17 of those with
relapse. This means that the most pronounced variation of MUPCDH
eexpression was detected in the outcome (disease-free/relapse) category,
where it also resulted statistically significant (p<0.05).
The same analysis, performed on CDH1 and presented in Figure 2,
showed that, although with high statistical significance (p<0.01, at least),
the expression of this gene was substantially unvaried from Normal to
Adenoma and exhibited only a slight decrease from Adenoma to CRC.
In fact, the Log2 median expression values were 10.95 in Normal, 10.89
in Adenoma and 10.61 in CRC.
Evaluation of CDH1 expression in the clinical-pathological
categories disclosed a Log, median expression value of 10.70 for early
stage compared to 10.47 of late stage, 9.93 for low grade compared to
9.43 of high grade, 10.61 for alive patients compared to 10.19 of dead
dones and 10.58 for disease-free patients compared to 10.30 of those
with relapse. All these comparisons resulted statistically significant
(p<0.05, at least).
These results confirmed our previous observations and, more
importantly, revealed the existence of a statistically significant
association between lower expression levels of μ-protocadherin and the
probability to develop a 5-year relapse of CRC.
Analysis of cadherin gene expression in the CRC expression
database
To further corroborate these findings, bioinformatic analysis was
extended to the entire cadherin gene superfamily. This evaluation was
performed on the same sample groups considered so far and the results
obtained were presented as cascade histograms (Figures 3 and 4) in
which cadherin genes were reported on x-axis whereas the expression
values were represented on y-axis as Log2 Fold Change mean value.
All plotted values were presented in each histogram according to an increasing order from the left (negative values, down-regulated genes) to the right (positive values, up-regulated genes), excluding data that did not reach the statistical significance (p<0.05).
Under these analysis conditions a considerable number of cadherin
genes, 35 out of the 48 genes included in the database, were detected
in the CRC-Normal comparison even if only 6 of them, equally distributed
between down- and up-regulated, exhibited an appreciable variation
of their expression levels (Log2 Fold Change mean value<-1 or >+1
respectively) (Figure 3A). Interestingly, among the former, MUPCDH
occupied the first position with a Log2 Fold Change mean value of
-2.05 and was immediately followed by PCDH24 and CDH19 whereas,
among the latter, we found CDH3, CDH11 and PCDH19.
Analysis of the Normal to Adenoma and Adenoma to CRC
categories disclosed a Log, median expression value of 10.16. To evaluate
variations among the latter, we found CDH3, CDH11 and PCDH19.
Analysis of MUPCDH expression in the CRC subsets associated to
specific clinical-pathological categories revealed a decrease of signal
in all the considered conditions, the only exception being represented
by stage. More in detail, the Log, median expression value was 7.04 for
early stage compared to 7.02 of late stage, 7.50 for low grade compared to
6.99 of high grade, 7.99 for alive compared to 7.67 of dead patients
and 7.66 in disease-free patients compared to 7.17 of those with
relapse. This means that the most pronounced variation of MUPCDH
eexpression was detected in the outcome (disease-free/relapse) category,
where it also resulted statistically significant (p<0.05).

Table 1. Clinico-pathological features of patients affected by CRC or Adenoma and selected for immune-histochemical examination of μ-protocadherin expression and methylation analysis of its upstream control region

<table>
<thead>
<tr>
<th>Gender</th>
<th>CRC</th>
<th>ADENOMAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Female</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Age: mean (range)</td>
<td>67.9 (54-85)</td>
<td>62.6 (51-81)</td>
</tr>
<tr>
<td>Site</td>
<td>Proximal</td>
<td>7</td>
</tr>
<tr>
<td>Distal</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Rectum</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Stage</td>
<td>I</td>
<td>2</td>
</tr>
<tr>
<td>II</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Differentiation</td>
<td>Well</td>
<td>4</td>
</tr>
<tr>
<td>Moderate</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Dysplasia</td>
<td>Low Grade</td>
<td>5</td>
</tr>
<tr>
<td>High Grade</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Summary of samples contained in the CRC expression profile database that have been used for bioinformatic analysis of cadherin gene expression.

<table>
<thead>
<tr>
<th>Stage at diagnosis</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>321</td>
</tr>
<tr>
<td>Late</td>
<td>662</td>
</tr>
<tr>
<td>Early</td>
<td>341</td>
</tr>
<tr>
<td>Grade at diagnosis</td>
<td>Total</td>
</tr>
<tr>
<td>Low</td>
<td>294</td>
</tr>
<tr>
<td>Total</td>
<td>367</td>
</tr>
<tr>
<td>High</td>
<td>73</td>
</tr>
<tr>
<td>5-year survival</td>
<td>Total</td>
</tr>
<tr>
<td>Alive</td>
<td>157</td>
</tr>
<tr>
<td>Total</td>
<td>222</td>
</tr>
<tr>
<td>Dead</td>
<td>65</td>
</tr>
<tr>
<td>5-year outcome</td>
<td>Total</td>
</tr>
<tr>
<td>Disease-free</td>
<td>183</td>
</tr>
<tr>
<td>Total</td>
<td>389</td>
</tr>
<tr>
<td>Relapse</td>
<td>206</td>
</tr>
</tbody>
</table>

Table 3: Summary of the CRC subsets, associated to the various considered clinical-pathological conditions, indicating the corresponding numerical entities.

the presence of only 21 genes (out of 48), only 4 of which exhibited a down-regulated expression with CDH1 and MUPCDH occupying an "ex aequo" first position with an identical Log2 Fold Change mean value of -0.23 (Figure 4). All the other 17 genes detected in this analysis exhibited an up-regulated profile.

These results provided further support to the other data presented so far since, in comparison with other cadherin genes, MUPCDH underwent the most pronounced variations of expression in all the analyzed conditions.

Immunohistochemical examination of colorectal carcinomas and adenomas selected for promoter methylation analysis

To elucidate the molecular mechanism responsible for the down-regulated expression of μ-protocadherin, we decided to assess the methylation status of its promoter region. This analysis was carried out in an independent series of sporadic colorectal carcinomas and adenomas that were preliminarily screened by immunohistochemical examination of μ-protocadherin expression, selected on the basis of the results obtained and subjected to the manual microdissection procedure (see Materials and Methods and Table 1 for details) (Figure 5).
**Figure 2:** Log2 median expression values of CDH1 (gene encoding for E-cadherin) in colorectal adenoma (Adenoma), carcinoma (CRC) and normal mucosa (Normal). Correlation with stage, grade, survival and outcome.
Figure 3: Log2 mean fold change of cadherin and protocadherin gene expression reaching statistical significance (t-test p-value<0.05) in different sample type comparisons between: Panel (A) normal colorectal mucosa (Normal) and carcinoma (CRC); Panel (B) Normal and adenoma (Adenoma); Panel (C) Adenoma and CRC

Figure 4: Log2 mean fold change of cadherin and protocadherin gene expression reaching statistical significance (t-test pvalue<0.05) in the comparison between relapsed (Relapse) and disease-free (Disease-free) patients.

Figure 5: Representative immune-histochemical analysis of μ-protocadherin in colorectal carcinomas. Left panel, a CRC μ-protocadherin positive case; right panel, a CRC μ-protocadherin negative case.

Table 4: Statistical analysis of the results obtained by methylation assay performed on μ-protocadherin promoter region. The mean percentage of methylated CpG nucleotides, detected in all sample groups, were subjected to the paired Tukey HSD Test. This evaluation was carried out separately for CpG1 and CpG2 sequences, respectively presented in the upper and lower lane of each table row. CRC MUPCDH+ and -, CRC positive and negative for μ-protocadherin expression respectively.

Methylation analysis of MUPCDH promoter region in colorectal adenomas and carcinomas

Methylation analysis of MUPCDH promoter region was performed using the bisulphite pyrosequencing procedure, able to indicate the precise methylation condition of each CpG site inside the investigated DNA regions, in 10 colorectal adenomas, 20 CRC cases, 5 normal colorectal mucosa samples and 4 CRC cell lines. This assay was conducted on three distinct CpG islands named 1, 2 and 3 (see Materials and Methods for details).

The CpG3 island appeared fully methylated in virtually all the tested samples (data not shown), without any correlation with the pattern of μ-protocadherin expression, indicating that it is probably not implicated in the phenomenon under consideration.

The CpG2 island, conversely, resulted hyper-methylated in cell lines and μ-protocadherin negative CRCs whereas it disclosed a completely unmethylated aspect in normal colorectal mucosa and μ-protocadherin positive CRCs (Figure 6).

The most interesting results were observed in the CpG1 island. In fact, in CRC cases, this DNA region exhibited a methylation pattern that was substantially superimposable to that observed in the CpG2 island whereas in CRC cell lines it appeared unmethylated in CaCo2 and HT29 cells, characterized by a detectable μ-protocadherin expression (although exclusively with RT-PCR) and a less transformed tumor phenotype, and strongly methylated in HCT116 and LoVo cells, carrying the opposite features [4] (Figure 6).

As expected, colorectal adenomas exhibited a methylation status that was intermediate between normal and cancerous colorectal mucosa.

Statistical analysis, performed on the results of methylation assay using the one-way ANOVA procedure, indicated that the differences
observed among the considered sample groups were statistically significant (p<0.001 for either CpG1 or CpG2 sequences). A subsequent paired Tukey HSD test, carried out on the mean percentage of methylated CpG nucleotides detected in the CpG1 and CpG2 islands, put in evidence that all the comparisons with μ-protocadherin negative CRCs were statistically significant (see Table 4).

These results clearly indicated that the molecular mechanism able to explain the down-regulation of μ-protocadherin expression observed in CRC is represented by the hyper-methylation of its promoter region.

**Discussion**

It is well recognized that down-regulation of cadherin expression is implicated in the tumour progression of a number of epithelial cancers, promoting a combined proliferation and invasion effect [17]. Recently, data obtained in our laboratory have allowed us to demonstrate the involvement of a cadherin superfamily protein, named μ-protocadherin, in CRC, where we observed a complete loss of its expression in the majority of analyzed cases [4].

To confirm this finding we employed a GEO-derived database containing more than one thousand microarray expression profiles of CRC, colorectal adenoma and normal colorectal mucosa samples, supplied with several clinical-pathological information. The subsequent bioinformatic analysis, carried out on this database using both a single gene- and a gene family-based approach, confirmed the down-regulated expression of μ-protocadherin, that resulted to begin at the initial steps of tumor transformation (Normal to Adenoma transition), maintaining the same trend in the later phases of the disease (Adenoma to CRC transition). In addition, we were also able to report, for the first time, that the entity of μ-protocadherin expression is directly correlated with the probability to achieve a 5-year disease-free survival, virtually coinciding with a definitive healing from CRC. These results clearly indicated that the molecular mechanism able to explain the down-regulation of μ-protocadherin expression observed in CRC can be conceivably ascribed to a hyper-methylated status of its promoter region.

These findings confirmed the data that have been previously published by our laboratory [11] and demonstrate the involvement of a cadherin superfamily protein, namely μ-protocadherin, in CRC, where it has been already observed by other authors [21,22], demonstrating the hyper-methylated status of its promoter region. Expression studies, performed in CRC cell lines upon treatment with de-methylating agents, will help to better characterize this finding.

**Conflict of Interest**

Any financial/commercial conflict of interests has been disclosed.

**References**