DNA methylotype analysis in colorectal cancer

Jinsheng Yu
Washington University School of Medicine in St. Louis

Robert R. Freimuth
Washington University School of Medicine in St. Louis

Robert Culverhouse
Washington University School of Medicine in St. Louis

Sharon Marsh
Washington University School of Medicine in St. Louis

Mark A. Watson
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
Yu, Jinsheng; Freimuth, Robert R.; Culverhouse, Robert; Marsh, Sharon; Watson, Mark A.; and McLeod, Howard L., "DNA methylotype analysis in colorectal cancer." Oncology Reports 20,4. 921-927. (2008). http://digitalcommons.wustl.edu/open_access_pubs/2095

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
DNA methylotype analysis in colorectal cancer

JINSHENG YU1,6,7, ROBERT R. FREIMUTH2,6,7, ROBERT CULVERHOUSE5,7, SHARON MARSH1,6,7, MARK A. WATSON2,6,7 and HOWARD L. McLEOD1,3,4,6,7

Departments of 1Medicine, 2Pathology, 3Genetics, 4Molecular Biology and Pharmacology and the 5Division of Biostatistics, Washington University School of Medicine; the 6Siteman Cancer Center and the 7CREATE Pharmacogenetic Research Network, Saint Louis, MO 63110, USA

Received September 28, 2007; Accepted November 15, 2007

DOI: 10.3892/or_00000092

Abstract. The methylation status of a gene promoter is considered to be an important mechanism for the development of many tumors, including colorectal cancer. Recent studies have shown that specific patterns of DNA methylation across multiple CpG loci in some human tumors are more informative than the detection of one single CpG locus in tumor genomes. In the present study, multiple CpG methylations of three genes (CDKN2A, DPYD and MLH1) were detected in DNA samples from patients with colorectal cancer using Pyrosequencing® technology. The bisulfite-converted DNA was amplified with a nested PCR and five or six CpG loci of each gene were assessed to determine DNA methylotype. Our data showed that 10/49 (20.4%), 6/48 (12.5%) and 14/49 (28.6%) of tumors were methylated with a DNA methylation level >0.2 in CDKN2A, DPYD and MLH1, respectively. Our study indicated a similar DNA methylotype distribution across the multiple CpG loci for all three genes in the methylated tumor DNA samples, demonstrating a dichotomous trait in DNA methylation. The tumor DNA samples had unique DNA methylation patterns, which were high-degree and multiple-site methylation, but the normal DNA samples had no or a low-degree and dispersed single-site methylation. In addition, the inverse correlation in those methylated tumors was observed between DNA methylation and RNA expression (1,4,5), but these were described with studies of one single CpG locus in the gene promoters. In conclusion, distinctive DNA methylotypes in colorectal cancer (6).

Just as genetic variation across a gene shows distinct haplotypes, recent studies demonstrated that specific patterns of DNA methylation across multiple CpG loci (methylotype) in some human tumors are more informative than the detection of one single CpG locus in tumor genomes (7-9). For instance, Yegnasubramanian et al (7) noted that by using various combinations of several genes such as GSTP1, APC and ABCB1, the CpG island hypermethylation can distinguish primary prostate cancer from benign prostate tissue, demonstrating the tissue-specificity in DNA methylation. Furthermore, the different CpG loci in the promoter region can exhibit a different methylation status (8), suggesting that a multiple CpG investigation in DNA methylation studies is important (9). This study assessed DNA methylation at multiple CpG loci and demonstrated the presence of multiple methylotypes in colorectal cancer.

Materials and methods

Patients and samples. Tumor specimens and paired normal colon tissues used in this study were from 52 Dukes’ C colorectal cancer patients [29 male/23 female, age range: 32-96 (median 69.5) years], who had not received any chemotherapeutic agents before surgery. Samples were snap-frozen in liquid nitrogen immediately after surgery and stored at -80°C. The tumor specimens selected for DNA and RNA isolation had high tumor cellularities. Written informed consent was obtained from all patients to bank tumor tissue.

Key words: DNA methylation, methylotype, pyrosequencing, colorectal cancer
and perform genomic analysis. This study was approved by the Washington University Human Subjects Committee.

Bisulfite conversion of genomic DNA. Genomic DNA was extracted from the colon tumor and normal tissues using the Qiagen DNA isolation kit and converted with sodium bisulfite as previously described (10). After treatment, sodium bisulfite unmethylated cytosine residues were converted to thymine, whereas methylated cytosine residues were retained as cytosine. In addition, a human genomic DNA sample purchased from Promega (Catalog No. G304A) was included either as a negative control (unmethylated) when it was converted only by sodium bisulfite or as a positive control (high methylation) when it was treated first by the SssI methylase (New England Biolab, Beverly, MA) and then converted by sodium bisulfite.

Amplification of the promoter region. A nested PCR approach was used to amplify the bisulfite-converted DNAs. A scheme for the methylation assay is shown in Fig. 1. The regular PCR primers were designed using Primer Express v1.5 (Applied Biosystems, Foster City, CA) and the Pyrosequencing primers were designed using the single nucleotide polymorphism primer design online software at www.biotage.com.

All methylation primers are listed in Table I. The first-round external PCR reaction was carried out for 40 cycles using an AmpliTaq Gold PCR master mix (Applied Biosystems), 5 pmole of each primer and 10 ng of the bisulfite-converted genomic DNA in a 20-μl reaction. The external reaction was run at 94, 55 and 72˚C for 1 min each, while the second-round nested reaction was run for 55 cycles at 94, 60 and 72˚C for 1 min each for MLH1 and 30 sec each for CDKN2A and DPYD amplification. The nested reaction was carried out with 5'-biotinylated reverse or forward primers in a 20-μl reaction containing 2 μl of 1:5 diluted first-round PCR products and 4 pmole of each primer.

The Pyrosequencing reaction. Pyrosequencing was performed as previously described (11) using the Pyrosequencing® PSQ HS96A instrument and allele quantification software (Pyrosequencing, Uppsala, Sweden). The Pyrosequencing reaction contained 5'-biotinylated single-strand PCR fragments (purified from 10 μl of the second-round PCR products) and 2.4 pmole Pyrosequencing primers, followed by the sequential addition of enzymes, substrate and dNTPs. In order to validate specificity of the Pyrosequencing reaction, negative controls were set with a fragment only, Pyrosequencing primer only and no primer and fragment.

<table>
<thead>
<tr>
<th>Primers (all 5’ to 3’)</th>
<th>CDKN2A</th>
<th>DPYD</th>
<th>MLH1</th>
</tr>
</thead>
<tbody>
<tr>
<td>First round forward</td>
<td>gTTTTTTgaggttagaggataggg</td>
<td>ttatAttagattaaaggTTagt</td>
<td>ttTaaTTTTgggTgggg</td>
</tr>
<tr>
<td>First round reverse</td>
<td>tacctAactccctacctAc</td>
<td>AAACCattACCAAACCCACCAa</td>
<td>AAACCaaccaAAACCAAAACCA</td>
</tr>
<tr>
<td>Second round forward</td>
<td>gTtgTgggTggTgggTggg</td>
<td>ggTgaatTgggggaaag</td>
<td>TgTgTgTgTgtgatatag</td>
</tr>
<tr>
<td>Second round reverse</td>
<td>cAactAacccctAcccctaccct</td>
<td>aAttAatAatAatAaaaccctc</td>
<td>tctActctcttAAActAAatctt</td>
</tr>
<tr>
<td>Pyrosequencing</td>
<td>gagggagtagTaggTag</td>
<td>ggTgatTgggggaagg</td>
<td>TggTggTggTggg</td>
</tr>
</tbody>
</table>

Figure 1. Scheme of DNA methylation assay for CDKN2A, DPYD and MLH1. The dashed arrow denotes the external primers, the one-line arrow denotes the nested primers, the one-line arrow with a ball denotes the biotinylated primers and the double-line arrow denotes the Pyrosequencing primers.
Determination of the DNA methylation level. By applying the allele quantification function in the Pyrosequencing system, the percentage of the C allele in a C to T polymorphism (i.e., the DNA methylation level) was determined for each CpG locus according to a theoretical histogram and peak height adjustment factor (Fig. 2). An average percentage of the C allele (i.e., C/T ratio) from three experiments was used for the DNA methylation status at each CpG locus. For the overall DNA methylation status of a gene, the mean C/T ratio of all CpG loci tested for that gene was calculated. When the mean C/T ratio is <0.05, it is defined as unmethylated (UM). Otherwise, 0.05-0.20 is denoted as low (LM), 0.21-0.50 as medium (MM) and >0.50 as high methylation (HM).

Measurement of RNA expression. Real-time quantitative PCR (qPCR) was employed to measure relative RNA expression level of the three genes as previously described (12). Briefly, a 10-μl reaction mixture for qPCR was composed of 5 μl of 2× TaqMan universal PCR master mix (Applied Biosystems), 3 μl of primer and probe mix (600 nM each forward and reverse primers, 100 nM specific TaqMan probe) and 2 μl of cDNA (20 ng). All real-time PCR assays were performed in triplicate on an ABI Prism 7900HT Sequence Detector System (Applied Biosystems). An internal reference gene called amyloid β precursor protein, which had nearly an identical expression between colon tumor and normal tissues, was used to control variation in RNA concentration across individual samples. Primers and TaqMan® probes used in this study were designed using the Primer Express version 1.5 (Applied Biosystems) and the qPCR primers and probes are shown in Table II.

Statistical analysis. Statistical analyses were performed with the software Statistica (StatSoft, Inc., Tulsa, OK, USA). The significance of difference of the DNA methylation level between the multiple CpG loci of each gene was evaluated.
by the Kruskal-Wallis median ANOVA test. A correlation between variables was observed with the Spearman rank order test. The significance level was set at P<0.05.

**Results**

**Level of DNA methylation.** Overall there were 17/49 (34.7%), 19/48 (39.6%) and 20/49 (40.8%) of tumors which were defined as DNA-methylated (C/T ratio ≥0.05) in CDKN2A, DPYD and MLH1, respectively. Amongst these methylated tumors, MM and HM were observed in 9 and 1 tumors (20.4%) for CDKN2A, 2 and 4 (12.5%) for DPYD and 5 and 9 (28.6%) for MLH1, respectively.

**The dichotomous trait of DNA methylation.** The Pyro-sequencing reactions in this study evaluated 5 or 6 CpG loci per gene of interest. The degree of methylation at a single CpG locus in methylated tumors varied from 5.2 to 100%. A similar degree of DNA methylation was evaluated across the multiple CpG loci for each of the 3 genes (P=0.10-0.73, Fig. 3). It suggests that the DNA methylation of the multiple CpG loci within a CpG island is a dichotomous event, i.e., either all or none of the multiple CpG loci were methylated.

**Specific pattern of DNA methylation.** By analyzing the methylation level of the multiple CpG loci, five specific patterns of DNA methylation were apparent: i) unmethylated at all loci, ii) low methylated at any of the loci, iii) low-medium methylated at any of the loci, iv) medium-high methylated at all loci and v) high methylated at all loci. As shown in Table III, the tumor DNA samples had diverse patterns of DNA methylation for all three genes in this study. In the highly methylated tumor DNA samples (types 4 and 5), type 5 was predominant for MLH1 (8/49, 16.3%), type 4 for CDKN2A (9/49, 18.4%), but neither type 4 nor 5 for DPYD (8.4 and 2.1%, respectively). However, the normal DNA samples had only low levels of DNA methylation (types 1 and 2) for all three genes (Table III). In addition, the 5 methylation types were each assigned a score of 1, 2, 3, 4 and 5 accordingly and then a sum score of the 3 genes was given for each tumor.

**Table III. DNA methylation patterns of the 3 genes in colorectal cancer.**

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Gene</th>
<th>Type 1 (None)</th>
<th>Type 2 (Low)</th>
<th>Type 3 (Low-median)</th>
<th>Type 4 (Median-high)</th>
<th>Type 5 (High)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean methylaiton</td>
<td>n</td>
<td>Mean methylaiton</td>
<td>n</td>
<td>Mean methylaiton</td>
</tr>
<tr>
<td>Tumor</td>
<td>CDKN2A</td>
<td>16 0.03</td>
<td>7</td>
<td>0.17</td>
<td>9</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>DPYD</td>
<td>21 0.05</td>
<td>7</td>
<td>0.18</td>
<td>4</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>MLH1</td>
<td>16 0.04</td>
<td>4</td>
<td>0.25</td>
<td>3</td>
<td>0.47</td>
</tr>
<tr>
<td>Normal</td>
<td>CDKN2A</td>
<td>20 0.03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>DPYD</td>
<td>25 0.08</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>MLH1</td>
<td>14 0.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3. Similar DNA methylation level across the multiple CpG loci of the 3 genes (CDKN2A, DPYD and MLH1).
(DNA methylotype). The distribution of the tumor DNA methylotype (Fig. 4) showed that the majority of the tumors had methylotypes with no or a low level of DNA methylation (sum score 3-8: 37/45, 82%) and 18% of tumors were median-highly methylated (sum score 9-13: 8/45).

Correlation of DNA methylation with RNA expression. The tumor RNA expression was significantly correlated inversely with the average DNA methylation level for \textit{MLH1} (RS=-0.62, \(P=0.003\)) but not for \textit{CDKN2A} and \textit{DPYD} (RS=0.17 and 0.36 and \(P=0.52\) and 0.13, respectively) in those methylated tumors. The \textit{MLH1} and \textit{DPYD} methylation data were previously described in brief within published articles (13,14).

Discussion

As more comprehensive and quantitative methods in assessing DNA methylation emerge, it has been suggested that epigenetics will play a more important role in understanding the development and progression of human diseases such as cancer (15-18). A rapid, accurate screening technique is essential in assessing DNA methylation in multiple CpG sites (9). Below we discuss some of the technical aspects of Pyrosequencing in the application of DNA methylation assessment in conjunction with the results of this study.

Many of the previous DNA methylation studies were conducted at a single CpG site, which gave no information for DNA methylation patterns of gene promoter regions. On the other hand, although global genomic DNA methylation content (proportion of 5-methylcytosine in the entire genomic bases) may have an important role in carcinogenesis, its measurement in cancer cells has little to offer as a molecular marker, either in sensitivity or informational content (9,19). In contrast, the DNA methylation pattern/methylotype, built upon the measurement of multiple CpG methylation, can provide detailed information on the characterizations of the 5-methylcytosine distribution along a stretch of DNA, thereby providing a unique approach to genome-wide molecular profiling (9,20,21). In the present study, we have defined 5 different types of DNA methylation patterns with multiple CpG loci for those three genes. Notably, the tumor DNA samples have distinctive methylation patterns from the normal DNA samples; i.e., for the methylated DNAs, the tumor samples are characterized by multiple-site and high-degree CpG methylation and the normal samples are characterized by a dispersed single-site and low-degree CpG methylation (Table III), suggesting the importance of the DNA methylation pattern as a potential marker in molecular diagnostics. In addition, the tumor DNA methylotype has the potential to identify patient groups with either high or low methylation according to the dichotomous trait of the DNA methylation within the CpG islands. However, in the distribution of the tumor DNA methylotype our data did not clearly display the two patient groups: high and low methylation groups. Instead it showed a continuous distribution of the tumor DNA methylotype across the spectrum of possible configurations. Further studies with a larger number of genes are needed to clarify the importance of tumor DNA methylotype in identifying patient groups based on the DNA methylation patterns.

Pyrosequencing can also generate quantitative data for the DNA methylation level (21). However, the concept of the DNA methylation level can be confused in the literature, because this level may be representing either how many molecules with a gene sequence of interest in a DNA sample or how many CpG loci in a gene sequence for a certain DNA sample have been methylated. The C/T ratio in this study served as the DNA methylation level for a given gene, meaning how many molecules with a gene sequence of interest in a DNA sample have been methylated. However, the Pyrosequencing protocol used in this study also presents how many CpG loci have been methylated in a single reaction.
Therefore, when we assess a gene for hypermethylation, it is better to describe the DNA methylation level in at least two dimensions: how many molecules in a DNA sample and how many CpG loci in a gene sequence have been methylated. Most DNA samples from tissue specimens have mixed genomes from a variety of cell-types including cells not of interest to the study (e.g., stromal cells) and different cells or tissues that can be methylated differentially at a specific CpG locus. The DNA mixture from different types of cells can mask the true level of DNA methylation at a specific CpG site for a specific type of cell. This is particularly important for quantitative DNA methylation studies and laser capture microdissection, while fluorescence-activated cell-sorting can provide highly pure cells of specific types.

Moreover, the detection potential of multiple CpG loci with Pyrosequencing allows the comparison of differential levels of DNA methylation between multiple CpG loci. As a result, a similar DNA methylation level is found in our study at multiple CpG loci for all three genes in the methylated tumor DNA samples; i.e., a tumor DNA sample either has no methylation for a single CpG locus or has a similar methylation level across the multiple CpG loci in the tumors. It suggests that DNA methylation is a dichotomous event. Taken together the dichotomous traits and specific patterns identified by Pyrosequencing and DNA methylation in cancer cells can be differentiated from that in normal cells both qualitatively and quantitatively. However, Pyrosequencing has several disadvantages (15,21). First of all, it has a reading length limit in a single reaction (currently less than 75 bp and 10 CpG loci) and cannot be used to detect an entire CpG island. Due to difficulties in primer design, it cannot always be used to detect certain regions in a CpG island. Moreover, it cannot be used to detect a specific haplotypic pattern of DNA methylation because it utilizes double-stranded templates.

In addition, our study demonstrates an inverse correlation between DNA methylation and RNA expression for MLH1. It is notable that the region of the gene promoter detected for MLH1 methylation in our study is indicated in a previous study (22) showing that the methylation status invariably correlates with the lack of MLH1 expression (-248 to -178, relative to the transcriptional start site). But for CDKN2A and DPYD in our study, the RNA level is not associated with the DNA methylation level in the promoter region indicated in Fig. 1. It has been noted that the CpG island of the CDKN2A promoter extends to exon 1 and 2 and the hypermethylation of CDKN2A exon 2 is accompanied by an increased but not decreased expression via a so-called transcriptional-coupled methylation (5). Our results do not indicate a correlation between the DNA methylation level and RNA expression but indicate a positive methylation-expression correlation for CDKN2A even though it does not reach a statistically significant level of correlation. Meanwhile, we have noted a recent study (23) showing two transcriptional regulatory elements close to the transcription start site (-72 to -23) of DPYD. The CpG loci in DPYD detected in our study (-119 to -86) do not locate in either binding site of the two transcriptional regulatory elements. Our results indicate that the CpG loci of DPYD tested in our study do not hold the potential to regulate the RNA expression. Even though further studies are needed to define the relationship between the DNA methylation and RNA level for CDKN2A and DPYD, our study otherwise suggests that the specific patterns and locations of DNA methylation as well as the tissue-specific methylation can be more important in cancer epigenetics. As such, DNA methylation in the promoter regions has a substantial effect on gene transcription but not in the coding region (19).

In conclusion, our data suggest that distinctive DNA methylotypes exist in colorectal cancer and may depict a distinct biology in apparently homogeneous tumors.

Acknowledgements

This study was supported by NIH grants U01 GM63340 and P30 CA091842.

References


