

Epigenetic Differences in Normal Colon Mucosa of Cancer Patients Suggest Altered Dietary Metabolic Pathways

Matthew L. Silveira^{1,3}, Brian P. Smith^{1,3}, Jasmine Powell³, and Carmen Sapienza^{2,3}

Abstract

We have compared DNA methylation in normal colon mucosa between patients with colon cancer and patients without cancer. We identified significant differences in methylation between the two groups at 114 to 874 genes. The majority of the differences are in pathways involved in the metabolism of carbohydrates, lipids, and amino acids. We also compared transcript levels of genes in the insulin signaling pathway. We found that the mucosa of patients with cancer had significantly higher transcript levels of several hormones regulating glucose metabolism and significantly lower transcript levels of a glycolytic enzyme and a key regulator of glucose and lipid homeostasis. These differences suggest that the normal colon mucosa of patients with cancer metabolizes dietary components differently than the colon mucosa of controls. Because the differences identified are present in morphologically normal tissue, they may be diagnostic of colon cancer and/or prognostic of colon cancer susceptibility. *Cancer Prev Res*; 5(3); 374–84. ©2012 AACR.

Introduction

One of the goals of human genome sequencing and whole genome association analyses is to identify genes involved in common human disease. Although there have been multiple successes [e.g., type 2 diabetes; (refs. 1, 2), asthma (refs. 3, 4), and other common diseases (refs. 5–7)], there has been little translational impact of identifying "common disease genes" because the relative risk of developing disease for carriers of each risk allele is small (ORs averaging 1.18 for single-nucleotide polymorphism at 10 loci with very strong association in type 2 diabetes; ref. 8, for example). Given the multifactorial nature of common disease, this fact is not surprising, but it begs the question of what additional information could be added to genetic risk data to increase the predictive power for any particular disease. In this vein, there is great interest in the potential for various measures of "epigenotype" to add predictive value to genetic risk data (9, 10).

The value of epigenetic information in this venture is potentially 3-fold. First, systemic epigenetic differences between individuals (i.e., those differences that result from

stochastic, environmental, or genetic factors that act very early in development) can help explain differences in gene expression between individuals of identical genotype at the affected locus. Second, systemic epigenetic differences that can be detected in an easily accessible tissue may serve as surrogate markers of gene activity in tissues that are inaccessible to analysis. Third, tissue-specific epigenetic differences between individuals may provide a mechanistic link between the genetic and environmental factors that contribute to disease risk.

Colon cancer accounts for more than 10% of all invasive cancer in the United States and more than 100,000 new cases are diagnosed annually (11). It is the third most common type of cancer in both men and women and is the third leading cause of cancer-related deaths (11). Only a small fraction (~5%) of all colon cancer is caused by highly penetrant inherited mutations (12) and only a minority of cases (up to 35%; ref. 13) seem to be influenced by heritable factors that have not yet been identified. Moreover, there is compelling evidence linking environmental influences such as Western diets and cigarette smoking with increased risk of colon cancer (14–17). Colon cancer is the ideal disease for studying both epigenetic differences between individuals and the epigenetic changes caused by the environment. Furthermore, epigenetic alterations have been associated with both increased risk of disease (18–20) and tumor progression (21, 22).

The genetic and epigenetic changes observed in colon tumors have been characterized in great detail by multiple laboratories (23–29). These differences have largely been characterized in patients with colon cancer, comparing colon tumors with adjacent normal colonic mucosa from the same patient. These studies lack true controls (i.e., patients without colon cancer) and do not reveal whether

Authors' Affiliations: Departments of ¹Surgery and ²Pathology, and ³Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, Philadelphia, Pennsylvania

Note: Supplementary data for this article are available at Cancer Prevention Research Online (<http://cancerprevres.aacrjournals.org>).

M.L. Silveira and B.P. Smith contributed equally to this work.

Corresponding Author: Carmen Sapienza, Fels Institute for Cancer Research and Molecular Biology, Temple University, 3307 N. Broad Street, Room 300, Philadelphia, PA 19140. Phone: 215-707-7373; Fax: 215-707-2102; E-mail: sapienza@temple.edu

doi: 10.1158/1940-6207.CAPR-11-0336

©2012 American Association for Cancer Research.

there is anything distinctive about the normal mucosa of patients with colon cancer compared with the colonic mucosa of patients without cancer. We hypothesize that the normal colonic mucosa of patients with cancer is, in fact, not "normal" but "epigenetically predisposed" to cancer because of the acquisition of multiple somatically heritable chromatin modifications, including differences in DNA methylation. The goal of our study was to identify DNA methylation differences that distinguish the "normal" colonic mucosa of patients with cancer from colon mucosa of individuals who do not have cancer and to determine whether these differences reflect an environmental interaction associated with colon cancer. The design of our current study allows us to identify epigenetic changes that may represent "field defects" that are found in the normal mucosa of patients with colon cancer that are unlikely to be a result of the aberrant cellular machinery of the tumor cells.

Methods

Tissue collection

The "normal" mucosa specimens from patients with colon cancer were collected from colon tissue removed in the operating room. Normal appearing colonic mucosa away from the tumor tissue was sharply removed and the samples were snap frozen prior to DNA and RNA isolation. Patients with a known history of familial adenomatous polyposis or hereditary nonpolyposis colon cancer were excluded.

Normal colon mucosa control specimens were collected from patients undergoing screening colonoscopy. Each patient was interviewed prior to the procedure by one of the investigators (M.L. Silveira or B.P. Smith). Patients who reported a personal or family history of colon cancer were excluded. Patients with a personal history of colon polyps or inflammatory bowel disease were also excluded. After providing informed consent, each patient underwent complete colonoscopy by a board-certified gastroenterologist. During that procedure, mucosal biopsies were obtained with a radial jaw large capacity biopsy forceps (Boston Scientific). Specimens were placed into RNALater RNA Stabilization Reagent (Ambion) and stored at 4°C prior to DNA and RNA isolation.

DNA and RNA isolation

Tissue samples were rinsed with sterile saline and blotted dry prior to nucleic acid extraction. DNA was extracted using standard phenol-chloroform techniques. The isolated DNA was dissolved in 10 mmol/L Tris-Cl (pH 8.0). Samples were quantified by spectrophotometry and stored at -80°C until ready for use. RNA was isolated using TRIzol Reagent (Invitrogen Corporation) according to the manufacturer's instructions. The RNA samples were purified using the Clean All RNA/DNA Clean Up Kit (Norgen Biotek Corporation). The isolated RNA was dissolved in Milli-Q water, quantified by spectrophotometry, and stored at -80°C until ready for use.

Bisulfite conversion and methylation assay

The EZ DNA Methylation-Gold Kit (Zymo Research) was used to convert unmethylated genomic DNA cytosine to uracil. Site-specific CpG methylation was analyzed in the converted DNA template (5 µL at 50 ng/µL) using the Infinium Assay (Illumina Inc.), the HumanMethylation27 BeadChip, and a BeadArray Reader according to the manufacturer's instructions. The HumanMethylation27 BeadChip targets 27,578 CpGs, the vast majority of which lie within the proximal promoter regions of transcription start sites of 14,475 consensus coding sequences in the National Center for Biotechnology Information (NCBI) database (Genome Build 36). Methylation data were analyzed on the GenomeStudio Data Analysis Software (Illumina Inc.) as well as SPSS version 16.0 (SPSS Inc.).

Pyrosequencing validation of methylation assay

Primers were designed for the genes of interest using PyroMark Assay Design Software version 2.0 (Qiagen). The PyroMark Gold Q96 Kit (Qiagen) was used to test 500 ng of bisulfite-converted DNA from samples and internal controls according to the manufacturer's recommendations. Analysis was conducted using the PSQ 96 HS Instrument and the PyroMark Q96 MD Software (Qiagen).

Transcriptome profiling

RNA integrity was tested using the 2100 Bioanalyzer 600 Nano RNA Chip (Agilent Technologies). The 6 RNA samples with the highest quality from patients with and without cancer were pooled. All 6 of the RNA samples from patients without cancer had RNA integrity number (RIN) values in excess of 8. Two of the RNA samples from patients with cancer had RIN values in excess of 8 but the remaining 4 samples had RIN values of 3.5, 4.0, 4.2, and 4.3. However, we tested whether the ($C_{t_housekeeping\ gene} - C_{t_gene\ of\ interest}$) was related to RIN value in the individual samples by comparing C_t values for *CEBPA*, *SLC2A1*, and *GAPDH* and showing that neither C_t values nor relative rank of transcript levels were related to RIN and therefore concluded that differences between groups could not be explained by RNA sample quality in these particular samples. In addition, we assayed transcript levels between cancer and control groups for 9 of the genes with the largest difference in an independent sample of cancer and control individuals by real-time reverse transcriptase (RT) PCR (see later). The Superscript III Reverse Transcriptase Protocol (Invitrogen Corporation) was used to create cDNA. One microgram of cDNA from each group was added to a 96-well RT Profiler PCA Array System: Human Insulin Signaling Pathway array plate (SABiosciences). Analysis was conducted using the Excel-based template provided by SABiosciences.

Quantitative real-time RT-PCR validation of transcript levels

Gene-specific TaqMan probes (Applied Biosystems) were used to quantify steady-state mRNA levels of *CEBPA*, *G6PC*, *IGFBP1*, *LEP*, *PTPRF*, *RETN*, *SERPINE1*, *SLC2A1*, and *INS* in 21 additional samples of normal colon mucosa from

patients with cancer and 21 additional samples of normal colon mucosa from patients without cancer or polyps. Glyceraldehyde—3—phosphate dehydrogenase (*GAPDH*) was used as the reference housekeeping gene. The cDNA and TaqMan mix were amplified under the following conditions: 50°C for 2 minutes, 95°C for 10 minutes, 45 cycles of 95°C for 15 seconds, and 60°C for 60 seconds. A melting curve analysis of the PCR products was conducted to verify their specificity and identity. Raw C_t values were used to compare relative gene expression levels using the $\Delta\Delta C_t$ method.

Results

DNA methylation profiling identifies an epigenetic signature of cancer in the normal colon mucosa of cancer patients

Normal colon mucosa (see Methods) from 30 patients with cancer and 18 controls was selected for quasi-genome-wide DNA methylation analysis. Approximately twice as many patients with cancer as controls were selected to gain statistical power (30). All of the colon mucosa specimens used in our analysis were from the right side of the colon (proximal to the hepatic flexure, see Methods). There was no difference in the mean age of patients with cancer and controls (65.6 ± 11.6 vs. 61.3 ± 11.6 , $P = 0.222$) or in the distribution of sex between the 2 groups (female: 55.6% of control vs. 56.7% of cancer, $P = 0.588$).

DNA was extracted by procedures that are standard for human tissues (see Methods) and 500 ng of each DNA sample was treated with sodium bisulfite and monitored for

conversion using a commercially available assay (see Methods). Site-specific DNA methylation was assayed using "HumanMethylation27 BeadChip" arrays, which contain probes for 27,578 CpG sites in 14,495 genes (see Methods).

Signals significantly above background were detected for more than 27,561 CpGs in all 48 samples. We compared mean β -values (" β -value" is the fraction of a particular CpG site that is methylated, which may range from 0 to 1; raw β -values were background normalized to correct for any differences in signal intensity between arrays) at each CpG site between the 30 patients with cancer and the 18 controls.

We used 3 different metrics to identify significant differences in site-specific mean methylation level between the normal colonic mucosa of patients with cancer and the colon mucosa of controls: (i) a Bonferroni-corrected P value of 1.8×10^{-6} (i.e., $0.05 / 27,578$ to correct for the number of individual CpG sites being tested) identified 119 sites in 114 genes (Table 1); (ii) a Benjamini-Hochberg false discovery rate (31) of 0.05 identified 909 sites in 873 genes (all of the genes identified in the Bonferroni-corrected data were also identified in the Benjamini-Hochberg false discovery rate screen; the additional 759 genes are shown in Supplementary Table S1; and (iii) a requirement that any candidate gene must show a significant difference ($P \leq 0.05$) between cancer and control groups for at least 3 CpGs in each gene identified 299 sites in 65 genes (an average of 4.5 CpGs per gene; Table 2).

We have used the latter *ad hoc* but "common sense" approach to identify candidate genes that were differentially methylated in children conceived through assisted reproduction (32, 33), as well as individuals with diabetic

Table 1. Genes in which mean colon mucosa methylation levels differ between patients with cancer and controls at the Bonferroni-corrected $P \leq 1.8 \times 10^{-6}$

<i>SEPT4</i>	<i>CD55</i>	<i>FBXO6</i>	<i>IL1B</i>	<i>NIP</i>	<i>SLC16A3</i>
<i>ACAD11</i>	<i>CDK9</i>	<i>FLJ14346</i>	<i>INPP5D</i>	<i>NT5E</i>	<i>SLC43A3</i>
<i>AHNAK</i>	<i>CGB5</i>	<i>FLJ20186</i>	<i>IRF5</i>	<i>PARVB</i>	<i>SLCO1C1</i>
<i>ALAS1</i>	<i>CGB8</i>	<i>FLJ30294</i>	<i>ISG20L2</i>	<i>PDCD1</i>	<i>SPRR2D</i>
<i>ALOXE3</i>	<i>CGI-69</i>	<i>FLJ39822</i>	<i>ITLN1</i>	<i>PDPK1</i>	<i>SULT1C2</i>
<i>ALPP</i>	<i>CHCHD1</i>	<i>FLJ43855</i>	<i>JAG2</i>	<i>PHLDB2</i>	<i>SUSD3</i>
<i>AP1S1</i>	<i>CHFR</i>	<i>FSD1NL</i>	<i>KCNQ1</i>	<i>PTD004</i>	<i>TBC1D5</i>
<i>APOA1</i>	<i>CHID1</i>	<i>FXYP7</i>	<i>KIAA1913</i>	<i>RAB11FIP5</i>	<i>TCF8</i>
<i>AQP7</i>	<i>COL13A1</i>	<i>GABRR2</i>	<i>KRTHB6</i>	<i>RAD23A</i>	<i>TCN2</i>
<i>ARFGAP1</i>	<i>CORO6</i>	<i>GATA2</i>	<i>KSP37</i>	<i>RAMP1</i>	<i>TEKT3</i>
<i>ARHGAP11A</i>	<i>CUGBP2</i>	<i>GCNT1</i>	<i>LTC4S</i>	<i>RASSF5</i>	<i>TFR2</i>
<i>ASAH1</i>	<i>DDX49</i>	<i>GGTLA1</i>	<i>MAPK10</i>	<i>RPS3A</i>	<i>TMEM55B</i>
<i>ATP9A</i>	<i>DLK1</i>	<i>GK2</i>	<i>MAPK15</i>	<i>RPUSD1</i>	<i>TNFRSF4</i>
<i>BCDIN3</i>	<i>DUSP5</i>	<i>GNRH2</i>	<i>MGC7036</i>	<i>RUFY3</i>	<i>TSSK6</i>
<i>BRF1</i>	<i>ELK4</i>	<i>GP1BB</i>	<i>MGC9712</i>	<i>S100A3</i>	<i>UNC13D</i>
<i>C12orf24</i>	<i>ENPEP</i>	<i>GSTP1</i>	<i>MUC5B</i>	<i>SEMA3B</i>	<i>USHBP1</i>
<i>C1QC</i>	<i>EPHA1</i>	<i>HIST1H2AJ</i>	<i>NDUFS2</i>	<i>SEMA6B</i>	<i>VAV1</i>
<i>CCL8</i>	<i>EXOC6</i>	<i>HIST2H4</i>	<i>NEK6</i>	<i>SERPING1</i>	<i>ZC3H11A</i>
<i>CCT6A</i>	<i>FAM105A</i>	<i>HS747E2A</i>	<i>NGFR</i>	<i>SGSH</i>	<i>ZNF248</i>

NOTE: Gene names in bold have largest mean differences between groups and were used to select individuals for analysis of insulin signaling pathway transcript levels.

Table 2. Genes in which mean colon mucosa methylation levels differ between patients with cancer and controls at 3 or more CpGs ($P < 0.05$ at each CpG)

<i>ABCB4</i>	<i>CCND2</i>	<i>FEN1</i>	<i>KLK10</i>	<i>PEG10</i>	<i>SLC22A18</i>
<i>ALX4</i>	<i>CDH13</i>	<i>GALR1</i>	<i>LOC129285</i>	<i>POLR2G</i>	<i>SMPD3</i>
<i>ATP10A</i>	<i>CDKN2A</i>	<i>GATA4</i>	<i>LOX</i>	<i>PPP1R9A</i>	<i>SNRPN</i>
<i>BCDIN3</i>	<i>CHFR</i>	<i>GNAS</i>	<i>MAGEL2</i>	<i>PSMB6</i>	<i>SYK</i>
<i>BCL2</i>	<i>CTS2</i>	<i>GNMT</i>	<i>MEST</i>	<i>PTPRO</i>	<i>THRB</i>
<i>BIK</i>	<i>DAPK1</i>	<i>GPX3</i>	<i>MGMT</i>	<i>PYCARD</i>	<i>TNFRSF10C</i>
<i>BRAF</i>	<i>DIRAS3</i>	<i>GRB10</i>	<i>MLH1</i>	<i>RAB32</i>	<i>UBE3A</i>
<i>C12orf24</i>	<i>DLX5</i>	<i>H19</i>	<i>MSX1</i>	<i>RB1</i>	<i>VHL</i>
<i>CALCA</i>	<i>DNAJC18</i>	<i>IGF2</i>	<i>NNAT</i>	<i>RUNX3</i>	<i>WT1</i>
<i>CASP8</i>	<i>EDNRB</i>	<i>INS</i>	<i>OBFC2B</i>	<i>SEMA3B</i>	<i>ZNF512</i>
<i>CCND1</i>	<i>ERBB2</i>	<i>KCNQ1</i>	<i>OSBPL5</i>	<i>SERPINB5</i>	

NOTE: Gene names in bold have largest mean differences between groups and were used to select individuals for analysis of insulin signaling pathway transcript levels.

nephropathy (34) and have shown that many of the methylation differences so identified are also correlated with differences in mean transcript level between groups (10, 32). We note that adopting the criterion of 3 or more differentially methylated CpGs distinguished the 65 candidate genes (Table 2) and 299 CpG sites from only 1,588 CpGs on the array that could fulfill the criterion of 3 or more CpGs per candidate gene. It is noteworthy that nearly 20% of the 1,588 CpGs that could have been different between patients with cancer and controls were found to be significantly different because these 1,588 CpGs are concentrated in genes selected on the basis of perceived importance in cancer or development. Furthermore, of the 114 genes that fulfill the Bonferroni-correction requirement in Table 1, only 5 are represented on the array by 3 or more CpGs (*BCDIN3*, *C12orf24*, *CHFR*, *KCNQ1*, *SEMA3B*). It is noteworthy that even though selection in Table 1 is for a single CpG to be different at the Bonferroni-corrected P value, all 5 genes exhibit significant differences at 3 or more CpGs, suggesting that the methylation differences between groups observed at single CpG sites in Table 1 are robust over greater distances. In fact, inspection of data on all of the CpGs interrogated in the 65 genes in Table 2 show numerous cases in which multiple CpGs, spread over hundreds to thousands of base pairs, are similarly and significantly differently methylated between cancer mucosa and controls (Supplementary Table S4).

As a measure of the magnitude of the difference in methylation levels between the mucosa of patients with cancer and the mucosa of controls and the discriminatory power of the approach, β -values for individual patients are graphed at 3 CpGs for 2 of the most interesting genes in Table 2 (from the standpoint of being cancer related and environment related), the tumor suppressor gene *VHL*, and the gene encoding insulin (Fig. 1). Similarly, individual β -values at 4 of the genes which each have 2 CpGs that are significantly different in the Bonferroni-corrected gene list

(Table 1), the oncogene *VAV1*, the oncogene *RASSF5*, the imprinted potassium channel gene *KCNQ1*, and the imprinted GABA receptor *GABRR2*, are shown in Fig. 2. It should be noted that the strong correlation between methylation levels at different but nearby (between 20 and 752 bp apart) CpGs that were also assayed on the array for 5 of the 6 genes shown in Figs. 1 and 2 (the *RASSF5* CpGs are 50 kb apart) suggests that the differences observed are representative of the actual level of methylation over the region (independently of external validation) and that the inter-individual differences observed are genuine. However, β -values for CpGs in 4 of the candidate genes (*SLC16A3*, *VAV1* from Table 1 and *INS*, *ZNF512* from Table 2) were validated independently by bisulfite pyrosequencing (see Methods and Supplementary Fig. S1. Supplementary Figure S2 shows β -values for individual patients at 2 CpGs in bisulfite pyrosequencing-validated candidates *SLC16A3* and *ZNF512*).

Functions of genes that are differentially methylated in cancer mucosa versus control mucosa

We conducted Ingenuity Pathway Analysis (Ingenuity Systems, Inc., see Methods) with the candidate genes identified using each of the 3 metrics (Bonferroni correction, Benjamini-Hochberg false discovery rate, 3 or more CpGs different in the same candidate gene) to identify potential functional pathway differences between the normal colon mucosa of patients with cancer and controls. Forty-nine of the 114 genes identified using the Bonferroni-corrected P value are found in pathways involved in carbohydrate and lipid metabolism and small molecule biochemistry (Supplementary Table S2). One of the top 3 networks (Fig. 3) is involved in both lipid metabolism and cell growth and proliferation. Interestingly, this network also has a link to vitamin D metabolism and high levels of vitamin D are suspected to be preventive of colon cancer (35). As expected, if the selection criteria are robust, the top genes identified using the Benjamini-Hochberg false discovery rate selection

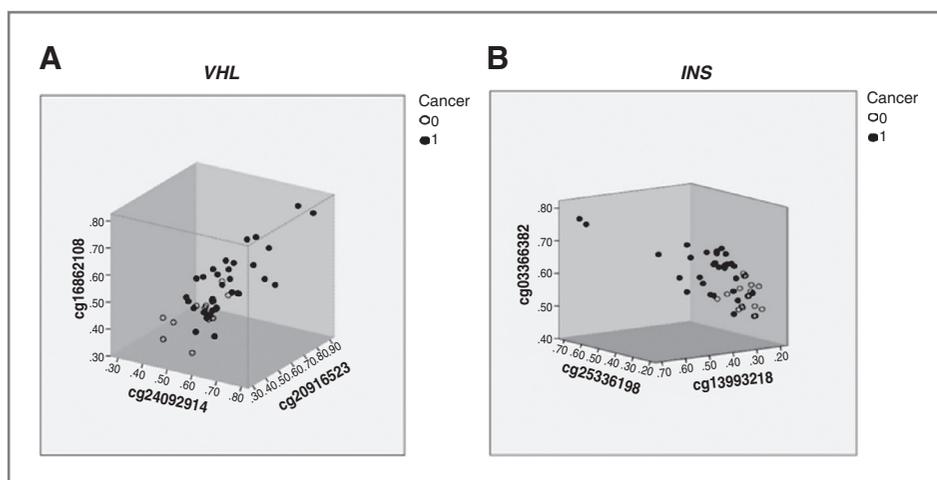


Figure 1. Methylation levels at 3 CpGs within the *VHL* gene (A) and the *INS* gene (B) in normal colon mucosa from patients with colon cancer (solid circles) and matched controls (open circles). CpGs were selected on the basis that mean methylation levels differed significantly between groups at $P < 0.05$.

yield similar pathways (we used only the top 114 genes, by P value, rather than all 873 genes, to determine whether the 2 metrics yielded comparable results; Supplementary Table S3). Of the top 4 networks obtained using each of the Bonferroni and Benjamini-Hochberg selected gene sets (Supplementary Tables S2 and S3), each of the 8 networks has between 13 and 22 of the input gene list present and a network score of greater than 20 (probability that the molecules are unrelated by function $< 10^{-20}$); the top function of 3 networks is carbohydrate metabolism, the top function of 2 networks is lipid metabolism, 3 networks are involved in small molecule

biochemistry, one network in amino acid metabolism and one of the 8 networks (Fig. 3) also has cell growth and proliferation as a top function.

Overlap with genes previously identified as differentially methylated in colon cancer

Relatively few of the colon mucosa differences identified in patients with cancer in our study are among the large number of genes that have been shown to be differentially methylated between colon tumors and matched colon mucosa of patients with cancer by others (23–26). In other words, the major methylation differences that we have

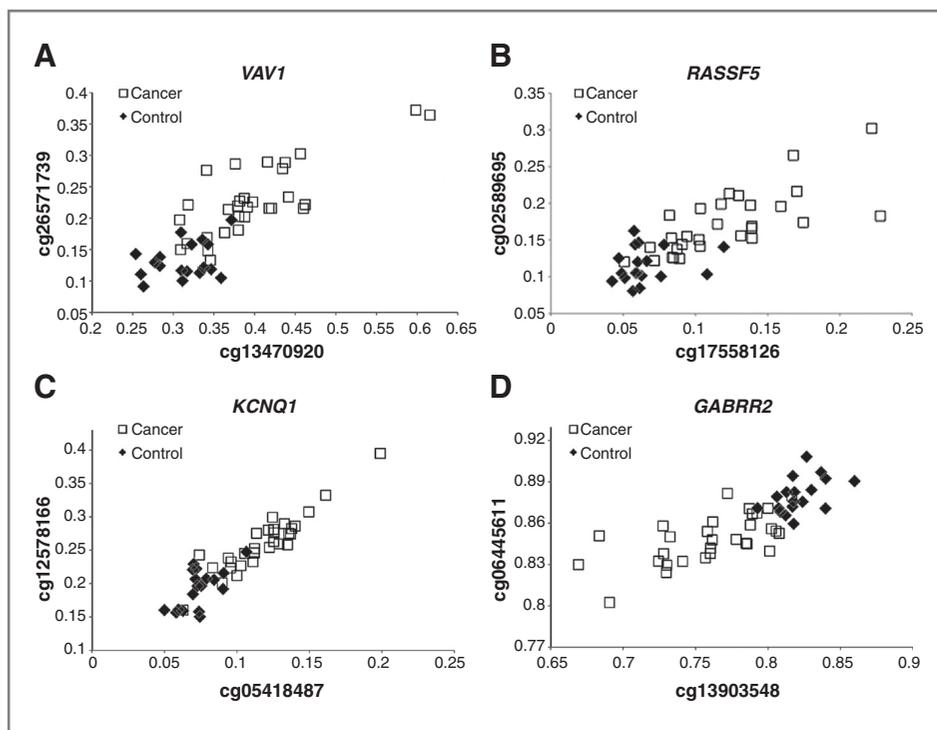


Figure 2. Methylation levels at 2 CpGs that differ significantly between groups at $P < 1.8 \times 10^{-6}$ within the *VAV1* gene (A), the *RASSF5* gene (B), and the imprinted genes *KCNQ1* (C) and *GABRR2* (D). Methylation levels plotted for normal colon mucosa from patients with colon cancer (open squares) and matched controls (filled diamonds).

Downloaded from <http://aacrjournals.org/cancerpreventionresearch/article-pdf/5/3/374/2251264/374.pdf> by guest on 28 September 2022

Table 3. Genes that differ in methylation between normal mucosa of patients with cancer and normal mucosa of controls and also differ between normal mucosa of patients with cancer and colon tumors

Bonferroni vs. Widschwendter and colleagues (23)	Benjamini-Hochberg vs. Widschwendter and colleagues (23)	Bonferroni vs. Irizarry and colleagues (26)	Genes with 3 CpGs different vs. Irizarry and colleagues (26)
<i>CHFR</i>	<i>BCL2</i>	<i>ALOXE3</i>	<i>ALX4</i>
<i>GSTP1</i>	<i>CHFR</i>	<i>CUGBP2</i>	<i>BCL2</i>
	<i>ESR1</i>	<i>DLK1</i>	<i>CALCA</i>
	<i>GATA4</i>	<i>DUSP54</i>	<i>DLX5</i>
	<i>GSTP1</i>	<i>GATA2</i>	<i>EDNRB</i>
	<i>IGF2</i>	<i>RASSF5</i>	<i>GALR1</i>
	<i>MSHR</i>	<i>SEMA6B</i>	<i>GATA4</i>
	<i>RUNX3</i>	<i>SUSD3</i>	<i>GNAS</i>
	<i>SFRP4</i>	<i>ZNF248</i>	<i>IGF2</i>
			<i>MEST</i>
			<i>MGMT</i>
			<i>MSX1</i>
			<i>PTPRO</i>
			<i>RB1</i>
			<i>THRB</i>
			<i>WT1</i>

"normal" mucosa from patients with colon cancer differs significantly at 3 or more CpGs are among those described by Irizarry and colleagues (26; Table 3, column 4).

Transcript levels of insulin signaling pathway genes are altered in normal colon mucosa from patients with cancer

If the DNA methylation differences we observe in genes involved in carbohydrate, lipid, and amino acid metabolism are indicative of altered metabolic function in the normal colon mucosa of patients with cancer, we might expect to observe differences in the expression of key components of important metabolic pathways. We used a commercially available PCR array (see Methods) to compare transcript levels of 89 genes in the insulin signaling pathway in the normal mucosa of 6 patients with cancer and 6 matched controls, pooled (Methods). The patients with cancer were selected on the basis of the greatest DNA methylation differences, compared with controls, in a selection of 10 genes from the Bonferroni candidates (Table 1) and "three CpG" candidates (Table 2). The insulin signaling pathway was selected for analysis because it is central to the metabolism of carbohydrates and Ingenuity Pathway Analysis suggested that methylation of genes in the insulin signaling pathway is altered in the normal mucosa of patients with cancer (Supplementary Tables S1 and S2).

Of the 89 genes in the insulin signaling pathway assayed for transcript level, 20 genes showing the greatest difference between cancer and control (higher or lower) are shown in Table 4. Among the genes showing the greatest increase in transcript level in cancer mucosa are the hormones *LEP* and *INS*. Among the genes showing the greatest decrease in transcript level in patients with cancer are a transcription

factor (*CEBPA*) that is intimately involved in glucose homeostasis, a protein tyrosine phosphatase receptor (*PTPRF*) involved in metabolic regulation, and an enzyme in the gluconeogenesis pathway (*G6PC*). Independent validation of the pooled sample array result was attempted for 9 of the genes in Table 4 (*LEP*, *SERPINE1*, *CEBPA*, *SLC2A*, *G6PC*, *IGFBP1*, *INS*, *RETN*, and *PTPRF*) using additional individuals from cancer and control groups (none of the individuals in the validation were analyzed on the array, see Methods) by real-time RT-PCR. Six (*LEP*, *G6PC*, *SERPINE1*, *CEBPA*, *PTPRF*, and *INS*) of the 9 candidate genes tested confirmed significant differences between groups of patients with cancer and control (Table 4) and individual transcript levels for 4 of these (2 in which transcript levels are higher in cancer mucosa and 2 in which transcript levels are lower) are illustrated in Fig. 4. The 3 genes for which significance was not reached also exhibited differences in the same direction as the pooled samples examined on the array (Table 4).

These results suggest that the gene-specific DNA methylation differences we observe between the normal mucosa of patients with cancer and the normal mucosa of controls result in differences in the ability of the 2 sources of normal mucosa to metabolize dietary components.

Discussion

Our findings indicate that there are major differences in DNA methylation between the normal mucosa of patients with cancer and the normal mucosa of controls. The major targets of these differences are genes involved in metabolism of carbohydrates, lipids, amino acids, and other small molecules. Our limited analysis of transcript levels in the

Table 4. Top 20 genes whose transcript level was highest or lowest in normal mucosa of patients with colon cancer compared with normal colon mucosa of patients without cancer, selected from the 89 genes profiled on the PCR array

Symbol	Fold change on array	Fold change in validation	P
Higher expression			
<i>LEP</i>	242	~Cancer specific	0.03
<i>PRKCG</i>	26.3		
<i>IRS4</i>	24.3		
<i>IGFBP1</i>	23.7	2.4	0.29
<i>GCK</i>	16.0		
<i>INS</i>	15.5	>1,000	<10 ⁻⁶
<i>PRL</i>	13.0		
<i>TG</i>	12.2		
<i>SERPINE1</i>	8.8	12.3	<10 ⁻⁶
<i>RETN</i>	7.6	1.6	0.34
Lower expression			
<i>PTPRF</i>	0.4	0.05	<10 ⁻⁵
<i>PPP1CA</i>	0.3		
<i>PDPK1</i>	0.3		
<i>PCK2</i>	0.3		
<i>GSK3A</i>	0.3		
<i>G6PC</i>	0.2	0.2	0.01
<i>ACOX1</i>	0.2		
<i>CEBPA</i>	0.2	0.35	<10 ⁻³
<i>HK2</i>	0.2		
<i>SLC2A1</i>	0.1	0.77	0.40

NOTE: Independent validation of 9 candidates by real-time RT-PCR was conducted for additional individuals from each group who were not analyzed on the PCR array and significant differences were confirmed for 6 genes. Significant *P* values are shown in bold.

insulin signaling pathway corroborate that such differences result in quantitative differences in gene expression in important metabolic pathways. The methylation differences observed between the normal mucosa of patients with cancer and the normal mucosa of controls are distinct from the differences found when the normal colon mucosa of patients with cancer is compared with the colon tumors of the same patients. These differences suggest that the normal colon mucosa of patients with cancer metabolizes dietary components differently than the colon mucosa of controls.

It is tempting to use the individual gene methylation differences observed to predict how they might affect transcription of each gene in each individual. Of the 20 genes profiled on the PCR array, only 2 exhibit significant between-group differences in methylation levels by the criteria considered in this study. Three CpGs in *INS* are significantly more methylated in cancer mucosa than control mucosa (Fig. 1 and Table 2). Two of the CpGs are located within 250 bp 5' to the transcription start site (but are not in a CpG island) and the third is in the first exon within 75 bp of the start site; however, *INS* is expressed at higher levels in cancer mucosa than control mucosa (Table 4 and Fig. 4). On the other hand, a CpG in an island 5' to the

PDPK1 transcription start site is significantly more methylated in cancer mucosa than in control mucosa (Table 1) and *PDPK1* is expressed at lower levels in cancer mucosa (Table 4). Of the 6 genes for which we validated significant differences in transcript level (Table 4), 2 (*LEP* and *SERPINE1*) have 2 CpGs in CpG islands adjacent to the start site that differ in the expected direction and one (*G6PC*) is interrogated by only a single CpG (that is not in an island) but this CpG also differs in the expected direction. One CpG in *PTPRF* is within 500 bp of the transcription start, but not in an island, and is more methylated in cancer mucosa than in controls and *PTPRF* is expressed at lower levels in cancer mucosa (Table 4 and Fig. 4). Both of the CpGs interrogated in *CEBPA* are in a CpG island but one is within the single exon and the other is 3' to the coding sequence. Both of these CpGs are less methylated in cancer mucosa than control mucosa but *CEBPA* is expressed at lower levels in cancer (Table 4 and Fig. 4). We have, however, observed a positive correlation between DNA methylation and transcript level at this gene in a previous study (32). Overall, we note that while approximately 50% of human genes exhibit an inverse correlation between transcription and DNA methylation (36), interindividual methylation differences, in *cis*, account for only a small fraction of interindividual

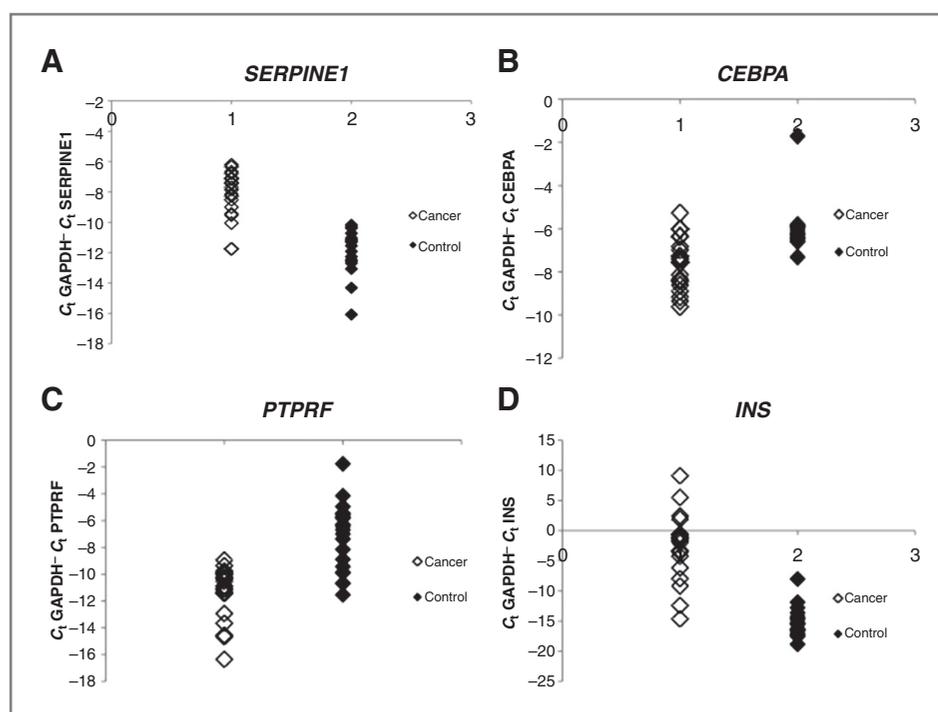


Figure 4. Real-time RT-PCR analysis of transcript level in normal colon mucosa of individual patients with cancer and control for *SERPINE1* (A; $n = 19$ cancer, 19 control; $P < 10^{-6}$), *CEBPA* (B; $n = 21$ cancer, 20 control; $P < 10^{-3}$), *PTPRF* (C; $n = 19$ cancer, 20 control; $P < 10^{-5}$), and *INS* (D; $n = 18$ cancer, 18 control; $P < 10^{-6}$).

variance in transcript level (approximately 10%–15%; ref.10, which is same fraction accounted for by genetic variation, in *cis* (reviewed in ref. 37), and cases of a positive correlation between methylation and transcript level also exist (e.g., ref. 32).

At this point, we cannot distinguish whether the epigenetic differences we observe are the result of preexisting differences between control individuals and individuals who later go on to develop cancer or are changes programmed by the tumor at distant sites in morphologically normal colon mucosa. In this regard, we note that at least some of the epigenetic differences we observe in the normal colon mucosa of patients with cancer may be associated with "field cancerization" epigenetic events. For example, promoter methylation of the *O*⁶-methylguanine-DNA methyltransferase (*MGMT*) has been observed in a significant fraction of normal mucosa samples from patients with cancer (38) and we also observed *MGMT* methylation differences at multiple CpG sites in our experiment (Table 2). However, at least some epigenetic differences between patients with cancer and controls are known to be preexisting, such as "loss of imprinting" at *IGF2/H19* (18–20). Moreover, it does not, *a priori*, seem obvious why the predominant epigenetic pathways reprogrammed by colon tumors should be involved in metabolism of lipids and carbohydrates. On the other hand, there are epidemiologic data that suggest a strong link between high fat diets and subsequent development of colon cancer (14–17), arguing that epigenetic reprogramming of lipid and carbohydrate pathways should occur before the development of cancer. If these metabolic

differences do preexist and predispose individuals toward further genetic and epigenetic changes that may lead to cancer, the identification of the pathways involved could allow for novel dietary or pharmaceutical interventions in those patients at highest risk.

Although none of the 8 examples of candidate gene methylation differences shown in Figs. 1 and 2 and Supplementary Fig. S2 completely discriminate the mucosa of patients with cancer from the mucosa of controls, it is apparent that a collection of such markers (the number would depend on the degree of overlap of cancer and control distributions for each marker) would have very high diagnostic power, in aggregate, to distinguish colon mucosa of patients with cancer from colon mucosa of controls and would fulfill the discriminatory demands of a clinical setting (39). Furthermore, if a significant fraction of the differences we observe are systemic, similar to the constitutional "loss of imprinting" found at *IGF2/H19* (18–20), then these markers may indicate a high probability that an individual will develop cancer and that this prediction could be made by assaying biomarker methylation levels in tissues such as peripheral blood or saliva. Even if all of the methylation differences observed are colon mucosa specific and these differences accumulate over the lifetime of the individual, they can serve as sentinel markers at which differences may occur prior to the appearance of colon polyps. The potential addition of an objective biochemical measure of cancer risk to an invasive screening test that relies entirely on the unaided eyes of the endoscopist to detect colon polyps would be an important diagnostic advance.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors thank Drs. John M. Daly and Daniel T. Dempsey for their ongoing support of our work and Dr. Benjamin Krevsky and the Department of Gastroenterology at Temple University Hospital and Drs. Elin Sigurdson and Andrew Godwin from Fox Chase Cancer Center for their assistance with colon biopsies and tissue acquisition.

Grant Support

This work was funded in part by an NIH training grant (T32 CA 103652-05) and the Fels Institute for Cancer Research and Molecular Biology.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 1, 2011; revised November 28, 2011; accepted December 19, 2011; published OnlineFirst February 1, 2012.

References

- Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007;445:881–5.
- Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 2006;38:320–3.
- Denham S, Koppelman GH, Blakey J, Wjst M, Ferreira MA, Hall IP, et al. Meta-analysis of genome-wide linkage studies of asthma and related traits. *Respir Res* 2008;9:38.
- Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KC, et al. A genome-wide association study of global gene expression. *Nat Genet* 2007;39:1202–7.
- Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;40:955–62.
- Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, Franchimont D, et al. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS Genet* 2007;3:e58.
- Zhang F, Sarginson J, Crombie C, Walker N, St Clair D, Shaw D. Genetic association between schizophrenia and the DISC1 gene in the Scottish population. *Am J Med Genet B Neuropsychiatr Genet* 2006;141B:155–9.
- Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007;316:1341–5.
- Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H, et al. Intra-individual change over time in DNA methylation with familial clustering. *JAMA* 2008;299:2877–83.
- Turan N, Katari S, Coutifaris C, Sapienza C. Explaining inter-individual variability in phenotype: is epigenetics up to the challenge? *Epigenetics* 2010;5:16–9.
- American Cancer Society. Colorectal cancer facts & figures 2011–2013. Atlanta, GA: American Cancer Society; 2011.
- Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology* 2010;138:2044–58.
- Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000;343:78–85.
- Willett WC. Diet and cancer. *Oncologist* 2000;5:393–404.
- Khan N, Afaq F, Mukhtar H. Lifestyle as risk factor for cancer: evidence from human studies. *Cancer Lett* 2010;293:133–43.
- Liang PS, Chen TY, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. *Int J Cancer* 2009;124:2406–15.
- Limsui D, Vierkant RA, Tillmans LS, Wang AH, Weisenberger DJ, Laird PW, et al. Cigarette smoking and colorectal cancer risk by molecularly defined subtypes. *J Natl Cancer Inst* 2010;102:1012–22.
- Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science* 2003;299:1753–5.
- Cui H, Horon IL, Ohlsson R, Hamilton SR, Feinberg AP. Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability. *Nat Med* 1998;4:1276–80.
- Nakagawa H, Chadwick RB, Peltomaki P, Plass C, Nakamura Y, de La Chapelle A. Loss of imprinting of the insulin-like growth factor II gene occurs by biallelic methylation in a core region of H19-associated CTCF-binding sites in colorectal cancer. *Proc Natl Acad Sci U S A* 2001;98:591–6.
- Hiraoka S, Kato J, Horii J, Saito S, Harada K, Fujita H, et al. Methylation status of normal background mucosa is correlated with occurrence and development of neoplasia in the distal colon. *Hum Pathol* 2010;41:38–47.
- Kim MS, Louwagie J, Carvalho B, Terhaar Sive Droste JS, Park HL, Chae YK, et al. Promoter DNA methylation of oncostatin m receptor-beta as a novel diagnostic and therapeutic marker in colon cancer. *PLoS One* 2009;4:e6555.
- Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, et al. Epigenetic stem cell signature in cancer. *Nat Genet* 2007;39:157–8.
- Hinoue T, Weisenberger DJ, Pan F, Campan M, Kim M, Young J, et al. Analysis of the association between CIMP and BRAF in colorectal cancer by DNA methylation profiling. *PLoS One* 2009;4:e8357.
- Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet* 2006;38:787–93.
- Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009;41:178–86.
- Zhang W, Press OA, Haiman CA, Yang DY, Gordon MA, Fazzone W, et al. Association of methylenetetrahydrofolate reductase gene polymorphisms and sex-specific survival in patients with metastatic colon cancer. *J Clin Oncol* 2007;25:3726–31.
- Estécio MR, Gallegos J, Vallot C, Castoro RJ, Chung W, Maegawa S, et al. Genome architecture marked by retrotransposons modulates predisposition to DNA methylation in cancer. *Genome Res* 2010;20:1369–82.
- Jones S, Chen WD, Parmigiani G, Diehl F, Beerewinkel N, Antal T, et al. Comparative lesion sequencing provides insights into tumor evolution. *Proc Natl Acad Sci U S A* 2008;105:4283–8.
- Andrieu N, Dondon MG, Goldstein AM. Increased power to detect gene-environment interaction using siblings controls. *Ann Epidemiol* 2005;15:705–11.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc* 1995;57:289–300.
- Katari S, Turan N, Bibikova M, Ernlie O, Chalian R, Foster M, et al. DNA methylation and gene expression differences in children conceived *in vitro* or *in vivo*. *Hum Mol Genet* 2009;18:3769–78.
- Turan N, Katari S, Gerson LF, Chalian R, Foster MW, Gaughan JP, et al. Inter- and intra-individual variation in allele-specific DNA methylation and gene expression in children conceived using assisted reproductive technology. *PLoS Genet* 2010;6:e1001033.

34. Sapienza C, Lee J, Powell J, Erinle O, Yafai F, Reichert J, et al. DNA methylation profiling identifies epigenetic differences between diabetes patients with ESRD and diabetes patients without nephropathy. *Epigenetics* 2011;6:20–8.
35. Peterlik M, Grant WB, Cross HS. Calcium, vitamin D and cancer. *Anticancer Res* 2009;29:3687–98.
36. Rakyan VK, Down TA, Thorne NP, Flicek P, Kulesha E, Gräf S, et al. An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res* 2008;18:1518–29.
37. Cheung VG, Spielman RS. Genetics of human gene expression: mapping DNA variants that influence gene expression. *Nat Rev Genet* 2009;10:595–604.
38. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, et al. MGMT promoter methylation and field defect in sporadic colon cancer. *J Natl Cancer Inst* 2005;97:1330–8.
39. Moonesinghe R, Liu T, Khoury MJ. Evaluation of the discriminative accuracy of genomic profiling in the prediction of common complex diseases. *Eur J Hum Genet* 2010;18:485–9.