

## Research Article

DNA Methylation Biomarkers for Noninvasive Diagnosis of Colorectal Cancer 

F. Javier Carmona<sup>1</sup>, Daniel Azuara<sup>2</sup>, Antonio Berenguer-Llargo<sup>3,9</sup>, Agustin F. Fernández<sup>1,10</sup>, Sebastiano Biondo<sup>5,7</sup>, Javier de Oca<sup>5,7</sup>, Francisco Rodriguez-Moranta<sup>6</sup>, Ramón Salazar<sup>2,9</sup>, Alberto Villanueva<sup>2</sup>, Mario F. Fraga<sup>10,11</sup>, Jordi Guardiola<sup>6</sup>, Gabriel Capellá<sup>2,4</sup>, Manel Esteller<sup>1,7,8</sup>, and Victor Moreno<sup>3,7,9</sup>

## Abstract

DNA methylation biomarkers for noninvasive diagnosis of colorectal cancer (CRC) and precursor lesions have been extensively studied. Different panels have been reported attempting to improve current protocols in clinical practice, although no definite biomarkers have been established.

In the present study, we have examined patient biopsies starting from a comprehensive analysis of DNA methylation differences between paired normal and tumor samples in known cancer-related genes aiming to select the best performing candidates informative for CRC diagnosis in stool samples.

Five selected markers were considered for subsequent analyses in independent biologic cohorts and *in silico* data sets. Among the five selected genes, three of them (*AGTR1*, *WNT2* and *SLIT2*) were validated in stool DNA of affected patients with a detection sensitivity of 78% [95% confidence interval (CI), 56%–89%]. As a reference, DNA methylation of *VIM* and *SEPT9* was evaluated in a subset of stool samples yielding sensitivities of 55% and 20%, respectively. Moreover, our panel may complement histologic and endoscopic diagnosis of inflammatory bowel disease (IBD)-associated neoplasia, as it was also efficient detecting aberrant DNA methylation in non-neoplastic tissue samples from affected patients.

This novel panel of specific methylation markers can be useful for early diagnosis of CRC using stool DNA and may help in the follow-up of high-risk patients with IBD. *Cancer Prev Res*; 1–10. ©2013 AACR.

## Introduction

Colorectal cancer (CRC) is a leading cause of cancer-related death in developed countries (1). Several screening options are available for the early detection, fecal occult blood test (FOBT) being a cost-effective procedure to reduce cancer-related mortality when used program-

matically (2). However, this test still needs substantial improvement to detect early-stage lesions. Population screening based on flexible sigmoidoscopy or colonoscopy are more sensitive alternatives (3–5), yet the invasive nature of these procedures and the high cost associated have prompted the development of alternative noninvasive screening methods such as stool- and serum-based screening tests.

Sporadic CRC arises as a consequence of the accumulation of genetic and epigenetic alterations that transform colonic epithelial cells into colon adenocarcinoma cells (6). Stool DNA testing has emerged as a biologically rational and user-friendly strategy for the noninvasive detection of both CRC and critical precursor lesions (7–9). Stool DNA testing offers a biologically rational approach based on the fact that neoplastic cells are continuously shed into the colonic lumen and mixed with stool. This method is noninvasive, requires no diet or medication restriction, and has potential for high accuracy of neoplasm detection.

Hypermethylation of an increasing number of genes has been associated with human colorectal tumorigenesis (10, 11), and the detection of CpG island methylation in human DNA isolated from stool has been proposed as a promising approach for noninvasive screening and early diagnosis of colorectal neoplasia (7, 12–14). Aberrantly methylated genes are attractive tumor markers because of their relatively

**Authors' Affiliations:** <sup>1</sup>Cancer Epigenetics and Biology Program (PEBC), <sup>2</sup>Translational Research Laboratory, <sup>3</sup>Biomarkers and Susceptibility Unit, <sup>4</sup>Hereditary Cancer Program, Catalan Institute of Oncology (ICO-IDIBELL), <sup>5</sup>Department of General and Digestive Surgery, Colorectal Unit, <sup>6</sup>Department of Gastroenterology, University Hospital of Bellvitge (HUB-IDIBELL), L'Hospitalet de Llobregat; <sup>7</sup>School of Medicine, University of Barcelona; <sup>8</sup>Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona; <sup>9</sup>CIBER de Epidemiología y Salud Pública (CIBERESP), Instituto de Salud Carlos III; <sup>10</sup>Cancer Epigenetics Laboratory, Instituto Universitario de Oncología del Principado de Asturias (IUOPA), HUCA, Universidad de Oviedo, Oviedo; and <sup>11</sup>Department of Immunology and Oncology, Centro Nacional de Biotecnología/CNB-CSIC, Cantoblanco, Madrid, Spain

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

F.J. Carmona, D. Azuara, and A. Berenguer-Llargo contributed equally to the study.

**Corresponding Author:** Victor Moreno, Biomarkers and Susceptibility Unit, Catalan Institute of Oncology (ICO-IDIBELL), Av. Gran Via 199-203, Hospitalet de Llobregat, Barcelona 08908, Spain. Phone: 0034-932607434; Fax: 0034-932607188; E-mail: v.moreno@iconcologia.net

doi: 10.1158/1940-6207.CAPR-12-0501

©2013 American Association for Cancer Research.

high frequency in neoplasms (15, 16). Ideally, PCR-based methods for detecting methylated markers in bodily fluids would be able to detect minute amounts of methylated gene copies derived from early altered cells against an essentially unmethylated background from normal tissue.

The selection of the correct genes to analyze is crucial to improve the sensitivity and specificity of methylation DNA tests. Studies carried so far attempting to improve the diagnostic accuracy for tumor detection have mainly been based on candidate gene approaches. Current microarray technology such as the GoldenGate Methylation BeadArray (Illumina, Inc.) provides the opportunity for high-throughput unbiased methylation analysis of a large number of CpG sites (17).

The aim of the present study was to identify a set of DNA methylation markers useful for early diagnosis of CRC starting from a comprehensive analysis of the DNA methylation profile differences between tumors and paired normal mucosa in known cancer-related genes. The best predictor was validated in an independent set of samples, before the assessment of its performance in stool DNA. For comparisons purposes, 2 previously described and promising methylation-based markers were evaluated in a subset of the same samples: *VIM* (15, 18) and *SEPT9* (19–21). In addition, we implemented our panel to evaluate their

performance as markers for CRC in non-neoplastic mucosa from patients with inflammatory bowel disease (IBD).

## Material and Methods

### Patients and samples

All patients gave informed written consent to participate and to have their biologic specimens analyzed. The study was cleared by the Ethical Committee of the University Hospital of Bellvitge, Barcelona, Spain.

For the *discovery phase*, we assessed DNA methylation levels using the Illumina GoldenGate Methylation Cancer Panel I microarray in 92 patients with sporadic CRC randomly selected from participants in a previous case–control study in our Institution (22). The original study included consecutive newly diagnosed and pathology-verified sporadic CRC cases that had undergone surgical removal of the tumor. Samples from resected tumors and paired normal adjacent mucosa (>10 cm away from tumor) were collected and immediately frozen at  $-80^{\circ}\text{C}$  until they were processed. After microarray preprocessing and quality control, 88 patients remained suitable for further analysis (Table 1, see Supplementary Methods for details of microarray preprocessing). These data were submitted in NCBI's Gene Expression Omnibus (23, 24) and are accessible through GEO series accession number GSE43369.

**Table 1.** Characteristics of patients

	Discovery phase		Validation phase		
	Sporadic CRC (n = 88)	Healthy controls (n = 39)	Sporadic CRC (n = 126)	IBD-associated neoplasias (n = 25)	
				Dysplasias (n = 8)	Cancers (n = 17)
Sex					
Male	51 (58%)	20 (51%)	83 (66%)	8 (100%)	13 (76%)
Female	37 (42%)	19 (41%)	43 (34%)	0 (0%)	4 (24%)
Mean age $\pm$ SD, y	63.2 $\pm$ 10.3	58 $\pm$ 7	66 $\pm$ 11	57 $\pm$ 20	65 $\pm$ 17
Location					
Distal	71 (81%)		84 (67%)	4 (50%)	7 (41%)
Proximal	16 (10%)		42 (33%)	4 (50%)	10 (59%)
Missing data	1 (1%)		—	—	—
Stage					
I–II	23 (26%)		65 (52%)		11 (65%)
III–IV	75 (85%)		61 (48%)		6 (35%)
Extent of IBD					
Extensive				7 (88%)	5 (31%)
Left-sided				1 (12%)	11 (69%)
Missing data				0 (0%)	1 (6%)
Duration of IBD, y				9 $\pm$ 8.7	8.8 $\pm$ 8.7
Histologic grade <sup>a</sup>					
HGD				0 (0%)	
LGD				8 (100%)	

NOTE: No significant differences were observed comparing IBD-associated neoplasias (dysplasias and cancers).

Abbreviations: HGD, high-grade dysplasia; LGD, low-grade dysplasia.

An *in silico* validation was conducted for the selected probes using public data sets GSE17648 (25) and GSE29490 (26) downloaded from Gene Expression Omnibus (GEO) genomics data repository. These data sets provided methylation from sporadic CRC tissue samples and paired adjacent normal mucosa (22 and 24 individuals, respectively) hybridized in the Illumina Infinium Human-Methylation27 BeadChip array (27). Preprocessed series matrixes originally provided by the authors were used for analysis. Given that this platform includes different CpG loci than those used in our study, the nearest loci to a given marker was considered as a good proxy for validation purposes.

The *biologic validation* of the selected methylation biomarkers was conducted by pyrosequencing in 3 independent groups of samples: (i) 62 tumors, 32 normal-appearing adjacent mucosa, and 87 stool samples were tested from 126 patients with sporadic CRC. Paired tumor and normal mucosa was available for 26 patients and paired tumor and stool for 27 patients. Twelve patients contributed to the study with the 3 types of samples. Tumor and normal mucosa were obtained at surgery and immediately stored at  $-80^{\circ}\text{C}$  until use; (ii) 25 patients with IBD-associated neoplasia (17 cancers and 8 mild dysplasia) stored as formalin-fixed, paraffin-embedded blocks and obtained from surgical resections (partial or complete colectomy). Tumor, adjacent normal-appearing mucosa samples, and, if available, actively inflamed mucosa were collected; and (iii) colonic biopsy specimens from 39 non-IBD patients without abnormalities at colonoscopy and histories of gastrointestinal disease served as controls. Methylation levels of *VIM* and *SEPT9* were evaluated in a subset of these samples comprising 26 mucosa from healthy individuals and 14 neoplastic tissue and 35 stool samples from patients with sporadic CRC.

The main clinicopathologic characteristics of the patients are described in Table 1. Diagnoses were verified after examination of sections stained with hematoxylin and eosin. Carcinomas were classified according to the Union for International Cancer Control (UICC) TNM classification system.

#### DNA extraction and bisulfite treatment

DNA was extracted from paraffin-embedded samples using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's instructions. DNA from fresh-frozen tissues and stools was extracted using phenol chloroform (28). For stool samples, a final purification step was added using Wizard DNA Clean-Up System (Promega) to remove PCR inhibitors. In all cases 500 ng of DNA was chemically modified to convert all unmethylated cytosine to uracil by the EZ DNA Methylation-Gold Kit (Zymo Research) according to the manufacturer's protocol.

#### Microarray DNA methylation assay

Illumina GoldenGate Methylation Cancer Panel I was developed to assay 1,505 CpG sites selected from 807 genes, including oncogenes and tumor suppressor genes, previ-

ously reported differentially methylated or differentially expressed genes, imprinted genes, genes involved in various signaling pathways and those responsible for DNA repair, cell-cycle control, metastasis, cell migration and invasion, differentiation, and apoptosis. After bisulfite treatment of DNA, the GoldenGate genotyping assay protocol was followed using Illumina-supplied reagents and conditions (29).

The methylation status at a CpG locus was estimated by  $\beta$ -values computed as the ratio between the fluorescent signal from the methylated allele and the sum of the fluorescent signals from both methylated and unmethylated alleles.  $\beta$ -values provide a continuous measure of DNA methylation and are interpreted as the proportion of methylated cells at the locus being interrogated; 0 indicates unmethylation whereas 1 represents a totally methylated state.  $\beta$ -values are averaged over 30 perfect replicates to reduce experimental variability (see Supplementary Methods for details of microarray preprocessing).

#### Pyrosequencing assay

Pyrosequencing assays were designed to analyze and validate the results obtained from the array. Primers were designed for 28 CpGs located in the promoter region of the selected candidate genes. Additional primers were designed to interrogate 4 and 5 CpG sites within the promoter region of *VIM* and *SEPT9*, respectively. Bisulfite-treated DNA was used as a template for a PCR. The primers for PCR amplification and sequencing were designed with PyroMark assay design software version 2.0.01.15. Primer sequences (Supplementary Table S1) were designed, when possible, to hybridize with CpG-free sites to ensure methylation-independent amplification. PCR was carried out with primers biotinylated to convert the PCR product to single-stranded DNA templates. We used the Vacuum Prep Tool (Biotage) to prepare single-stranded PCR products according to the manufacturer's instructions. Pyrosequencing reactions and quantification of methylation were conducted in a PyroMark Q24 System version 2.0.6 (Qiagen). To rule out the presence of technical bias in the quantification of DNA methylation values, internal sequence-specific and bisulfite conversion controls were considered in the interpretation of the results. Representative programs for the markers on a healthy colonic mucosa are represented in Supplementary Fig. S1.

#### Statistical analysis

For exploratory purposes and to identify global methylation patterns, array  $\beta$ -values were analyzed using a Principal Component Analysis (PCA) and a hierarchical clustering with Euclidean distance and Ward agglomeration method. Clustering analysis was also conducted on paired differences tumor-normal within samples. Association between clusters and demographic and clinical covariates was assessed using an exact Fisher test.

To maximize the likelihood of finding useful markers for stool testing, candidates in the discovery phase were required to be hypermethylated in tumor and to show

methylation mean  $\beta$ -values difference larger than 0.2. To further increase their potential sensitivity, special attention was paid to loci that were consistently unmethylated in adjacent normal mucosa (low mean and low variability, see Supplementary Methods). Candidates were prioritized according to their  $P$  value derived from a  $t$  test for paired data. To assess the discriminant value of multiple loci, simple unweighted mean  $\beta$ -values were used as summary. The area under the receiver operator characteristic (ROC) curve (AUC) and sensitivity at 80% specificity were used to assess the discriminant ability.

For pyrosequencing data, the 99th percentile was computed on all methylation values obtained in healthy mucosa for all CpG sites being interrogated. This threshold value was found at 18% methylation and was used as decision cutoff for all markers. A CpG was considered methylated when its value exceeded this threshold. At gene level, a gene was considered methylated when a least one of its CpG sites was methylated. When multiple genes were jointly tested, a sample was predicted as positive (altered) when at least one gene was methylated. Statistical association was assessed by Fisher exact test applied to the contingency table of predicted and actual outcome. In each case, sensitivity was computed to measure the accuracy of the markers for CRC detection. An estimation of specificity was obtained using the colonic mucosa from the healthy patients. In this analysis, only samples with complete data and those showing at least one methylated marker were included (see Supplementary Methods).

For each of accuracy measures used, 95% confidence intervals (CI) were calculated using the Bias Corrected and Accelerated bootstrap (BCa) method with 1,000 resamples and outcome stratification. Empirical influence values were estimated by the usual jackknife method. These calculations were conducted using packages *boot* (30) and *ROCR* (31) within the R language and environment for statistical computing (32).

## Results

### Methylation patterns in tumor and normal colorectal mucosa

Clustering results on paired differences of our sample set showed large methylation changes between tumors and normal mucosa for the CpGs analyzed in the GoldenGate array (Supplementary Fig. S2). Hypermethylation in tumor was predominant in most of the patients being analyzed. The first principal component derived from a PCA on the array  $\beta$ -values clearly discriminated between normal and tumor samples (37% variability explained, Supplementary Fig. S3). A hierarchical cluster analysis of samples identified 2 main groups, which generally separated normal and tumors, with the exception of 12 misclassified tumor samples (Supplementary Fig. S4). These results indicated a large degree of methylation differences between tumor and normal mucosa that was captured easily by these unsupervised classification techniques. Clustering results of CpG loci in the vertical axis also showed that methylation changes are typically consistent among samples for large sets of genes.

No association was found among normal or tumor samples between age, sex, or tumor location and the main clusters derived from this analysis

### Selection of candidate methylation biomarkers

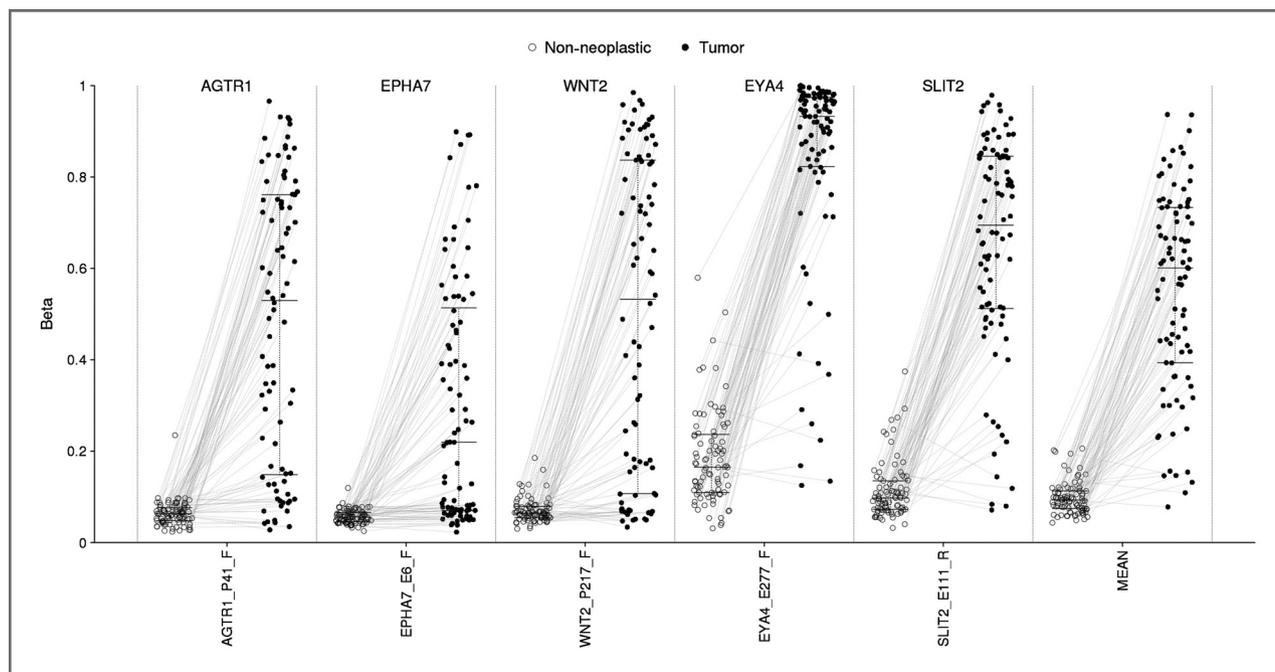
An unbiased selection of candidate genes resulted from the analysis of the methylation data provided by the methylation array (Supplementary Table S2). Of the 873 loci being analyzed, 147 were found to be highly hypermethylated in tumor samples according to the required criteria (difference  $> 0.20$  in mean  $\beta$ -values), thereby being informative indicators of aberrant methylation associated to tumorigenesis. The differences ranged from 0.20 to 0.65 and were highly significant as highlighted by their paired  $t$  test ( $P < 4.9E-10$ , Supplementary Table S3). Of them, 33 CpGs were identified as consistently unmethylated in most of the normal samples as described in the Materials Methods, a trait which could potentially increase sensitivity if included in a cancer panel for population screening (Supplementary Table S4 and Supplementary Methods). Candidates for subsequent validation were selected from these 2 ranked lists of markers and the selection process took into account statistical criteria ( $p$ -value of  $t$ -test) and technical criteria as primer design for validation. The final selection included the 2 most significant probes from the 147 markers list, which were located in genes *EYA4* and *SLIT2*, and 3 probes from the 33 markers list located in genes *AGTR1*, *WNT2*, and *EPHA7* which ranked in positions 8, 11, and 25, respectively (Supplementary Table S2).

Figure 1 gives a first insight into the promising discrimination potential contained in the selected loci. When evaluated formally, all markers showed AUCs that ranged between 0.85 (*EPHA7*) and 0.97 (*EYA4*). Sensitivity at 80% specificity for tumor discrimination was at least 78% (*EPHA7*) for all loci and reached their maximum at 96% (*EYA4* and *SLIT2*). A combination of all markers using their mean as summary achieved an AUC of 0.98 and nearly 98% sensitivity at 80% specificity (Table 2).

The GoldenGate array included 4 additional loci in genes *SLIT2* and *EPHA7* and 2 probes in *EYA4*. These loci were characterized and showed a similar methylation pattern to the loci that prompted the selection of the genes (Supplementary Fig. S5).

### In silico validation of selected genes

Selected markers were externally validated in public independent data sets. Evaluation at the locus level resulted in AUCs which ranged from 0.71 to 1.00, whereas loci combined together achieved a perfect classification of all samples in GSE17648 and an excellent performance in GSE29490 (AUCs: 1.00 and 0.95, respectively; Table 2). These data sets used the Infinium HumanMethylation27 array that has good overlap with the GoldenGate cancer panel I but not perfect, adding extra value to the markers identified in our analysis. Distances from the original marker to the nearest Infinium loci in which markers were evaluated ranged from 14 (*EYA4*) to 316 (*EPHA7*) pairs of



**Figure 1.** Methylation values in microarray experiments: microarray  $\beta$ -values for selected markers and their mean values in adjacent non-neoplastic and tumor samples. Segments inside the scatter plots indicate median and percentiles 25 and 75. Dot lines join paired samples from the same patient. In the graphics, marker's ability to discriminate between tissue types become apparent and it notably seems to increase when they are combined using a simple mean as summary.

bases (Reference Sequence 36.1), and methylation values displayed similar patterns to those observed in our platform (Supplementary Fig. S6).

#### Validation of selected genes by pyrosequencing

To further test the prediction ability of the panel, we conducted a validation step in an independent cohort using a different technique, aiming to confirm the differential methylation of the biomarkers in tumorigenesis. Primers were designed for 28 CpGs located in the promoter region of the 5 selected genes, and their methylation level

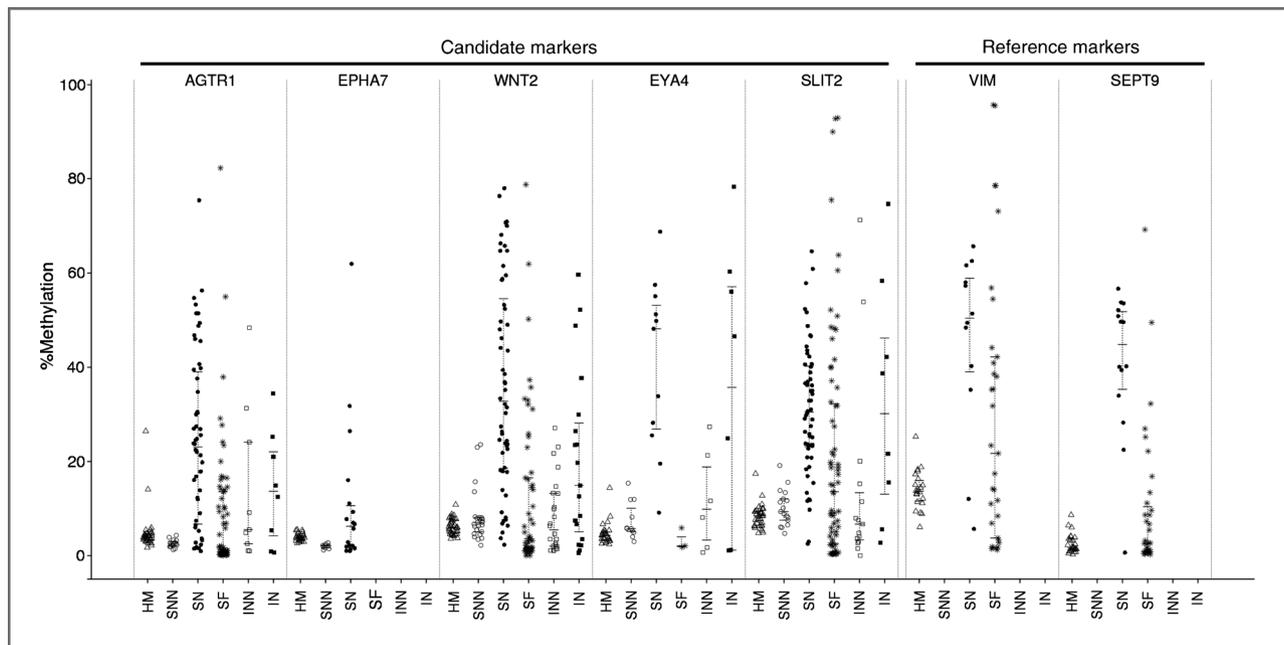
was quantitatively estimated in 267 normal mucosa, primary tumor, and stool samples from a total of 190 individuals. As expected, in normal biopsies of healthy controls ( $n = 39$ ), gene promoters showed methylation in a minority of cases (0%–7% samples). An increased methylation was observed in sporadic primary tumors, stool, and IBD samples.

Furthermore and for comparisons purposes, methylation levels of *VIM* and *SEPT9* were also estimated in a subset of these samples. Four and 5 CpG sites were interrogated within the promoter region of *VIM* and *SEPT9*, respectively,

**Table 2.** Prediction results in microarray data

	Discovery phase		<i>In silico</i> validation			
	GoldenGate data		GSE17648		GSE29490	
	AUC	SENS80 (%)	AUC	SENS80 (%)	AUC	SENS80 (%)
AGTR1	0.92 (0.85–0.96)	89 (78–93)	0.91 (0.75–0.98)	86 (59–96)	0.76 (0.59–0.88)	67 (42–83)
EPHA7	0.85 (0.78–0.90)	78 (66–88)	0.80 (0.61–0.93)	77 (50–86)	0.71 (0.52–0.85)	63 (33–75)
WNT2	0.88 (0.81–0.93)	80 (68–86)	0.93 (0.79–0.99)	91 (68–96)	0.85 (0.69–0.95)	79 (52–92)
EYA4	0.97 (0.93–0.99)	96 (86–98)	1.00 (1.00–1.00)	100 (100–100)	0.96 (0.82–1.00)	96 (83–100)
SLIT2	0.97 (0.92–0.99)	96 (89–99)	0.89 (0.68–0.96)	86 (55–96)	0.83 (0.68–0.93)	79 (42–92)
All loci	0.98 (0.95–0.99)	98 (91–99)	1.00 (1.00–1.00)	100 (100–100)	0.95 (0.86–0.99)	96 (58–100)

NOTE: For each loci and their combination, results from AUC and %sensitivity at 80% specificity (SENS80) computed in the microarray data used for markers selection (GoldenGate) and public data sets used for *in silico* validation. Brackets contain the corresponding intervals at 95% CI.



**Figure 2.** Methylation values in pyrosequencing experiments: percentage of methylation in genes analyzed by pyrosequencing in healthy mucosa (HM), tumor (SN), non-neoplastic adjacent mucosa (SNN), and stool (SF) from patients with sporadic CRC, and tumor (IN) and non-neoplastic mucosa (INN) from patients with IBD-associated CRC. Segments inside the scatter plots indicate median and percentiles 25 and 75. Methylation patterns are similar to those obtained in microarray analysis although stool and IBD differences are not as high as in tumor analysis.

which showed a similar methylation pattern to that observed in the panel markers. Figure 2 shows the distribution of the methylation values for all the CpG sites evaluated in the pyrosequencing assay.

### Methylation status in stool DNA of sporadic CRC

Aiming to define the minimal yet best-informative methylation signature, we used the 3 best-performing markers available for this analysis (*AGTR1*, *WNT2*, and *SLIT2*). Pyrosequencing of stool DNA detected 17 of 24 (71%) of carcinomas that scored positive for at least one methylation marker in tumor tissue. No stool methylation was detected in 7 of 24 tumors (29%) that were hypermethylated.

When the total stool DNA samples ( $n = 87$ ) were considered, the panel succeeded to detect 50 out of the 64 samples with evaluable data, which resulted in sensitivity for CRC detection of 78% (95% CI, 56%–89%). The relative contribution of every marker to the detection of carcinoma was: 21% (14 of 68) for *AGTR1*, 40% (21 of 52) for *WNT2*, and 52% (37 of 71) for *SLIT2* (Table 3). Whenever methylation was detected in stool, it was also detected in the corresponding tumor tissue except in 2 cases. In tissue DNA, the sensitivity of the panel of 3 markers for sporadic CRC detection was 95% (95% CI, 73%–100%) and the specificity was 89% (95% CI, 82%–92%; Table 3).

As a reference, *VIM* and *SEPT9* showed similar performances when evaluated in a subset of these samples: *SEPT9* reached a sensitivity of 93% in tumor tissue (13 of 14; 95% CI, 50%–100%) and 20% in stool samples (7 of 35; 95% CI, 6%–31%); *VIM* obtained sensitivities of 83% in tumor

samples (10 of 12; 95% CI, 33%–92%) and 55% in stools (18 of 33; 95% CI, 33%–70%). Estimations of specificity based on healthy mucosa were 100% for *SEPT9* and 86% for *VIM* (Table 3). When evaluated in this subset of samples, the 3-gene panel showed similar performance to that obtained in the full data set: 74% sensitivity (20 of 27; 95% CI, 50%–90%) and 88% specificity (22 of 25; 95% CI, 77%–93%).

### Methylation status in IBD-associated neoplasia

Methylation of the panel of genes analyzed was a common phenomenon observed in IBD-associated neoplasia (93%). *AGTR1*, *WNT2*, and *SLIT2* markers were methylated between 50 and 86% for every single marker. The non-neoplastic mucosa of patients with IBD-CRC showed higher methylation levels in *AGTR1*, *WNT2*, and *SLIT2* than in colorectal mucosa of healthy controls (82% vs. 11%, respectively,  $P = 4E-07$ ), indicating the existence of underlying molecular alterations in pre-altered mucosa as a prior step to malignant transformation. The proportion of samples methylated in non-neoplastic mucosa in IBD-associated neoplasia was more than 3 times the observed in sporadic cancer-adjacent mucosa (82% vs. 22%, respectively; Table 3).

### Discussion

Early detection of CRC is a major challenge to improve patient's survival and widen the window of therapeutic intervention. In this study, we followed an unbiased approach for candidate selection starting from a large candidate gene set and conducted several validation steps to

Table 3. Prediction results in pyrosequencing data

Panel markers	Non-neoplastic tissue			Neoplastic tissue			Stool				
	Healthy controls (n = 39)	Sporadic CRC (n = 32)	Fisher's P	IBD - CRC (n = 27)	Fisher's P	Sporadic CRC (n = 62)	Fisher's P	IBD - CRC (n = 20)	Fisher's P	Sporadic CRC (n = 87)	Fisher's P
AGTR1											
Methylated, n (%)	2/39 (5%)	0/15 (0%)	1	3/9 (33%)	0.0393	35/54 (65%)	2E-09	4/8 (50%)	0.0050	14/68 (21%)	0.0463
Unmethylated, n (%)	37/39 (95%)	15/15 (100%)		6/9 (67%)		19/54 (35%)		4/8 (50%)		54/68 (79%)	
WNT2											
Methylated, n (%)	1/39 (3%)	3/25 (12%)	0.2905	8/24 (33%)	0.0013	46/56 (82%)	5E-16	11/19 (58%)	3E-06	21/52 (40%)	1E-05
Unmethylated, n (%)	38/39 (97%)	22/25 (88%)		16/24 (67%)		10/56 (18%)		8/19 (42%)		31/52 (60%)	
SLIT2											
Methylated, n (%)	2/37 (5%)	1/17 (6%)	1	7/15 (47%)	0.0012	48/56 (86%)	2E-15	6/7 (86%)	3E-05	37/71 (52%)	5E-07
Unmethylated, n (%)	35/37 (95%)	16/17 (94%)		8/15 (53%)		8/56 (14%)		1/7 (14%)		34/71 (48%)	
≥1 methylated gene											
Yes	4/38 (11%)	4/18 (22%)	0.2544	14/17 (82%)	4E-07	57/60 (95%)	< 2.2E-16	13/14 (93%)	5E-08	50/64 (78%)	2E-11
No	34/38 (89%)	14/18 (78%)		3/17 (18%)		3/60 (5%)		1/14 (7%)		14/64 (22%)	
Reference markers	Healthy controls (n = 26)	Sporadic CRC (n = 14)	Fisher's P	Sporadic CRC (n = 35)	Fisher's P						
VIM											
Methylated, n (%)	3/22 (14%)	10/12 (83%)	0.0001	18/33 (55%)	0.0040	18/33 (55%)	0.0040	15/33 (45%)	0.0040	15/33 (45%)	0.0040
Unmethylated, n (%)	19/22 (86%)	2/12 (17%)		7/35 (20%)	0.0169	7/35 (20%)	0.0169	28/35 (80%)	0.0169	28/35 (80%)	0.0169
SEPT9											
Methylated, n (%)	0/26 (0%)	13/14 (93%)	1E-09	1/14 (7%)		1/14 (7%)					
Unmethylated, n (%)	26/26 (100%)	1/14 (7%)									

NOTE: Performance of individual genes and the 3-gene panel using methylation estimates from pyrosequencing data.

define a consistent panel of DNA methylation biomarkers useful for CRC early detection. The analysis of this panel in stool samples identified the presence of colorectal carcinomas with 78% sensitivity. An estimate of specificity was 89%, derived from the analysis of normal tissue, because no stool samples from healthy subjects were available. These estimates are encouraging about the potential use of these markers in clinical practice.

Several studies have focused in the detection of aberrant DNA methylation of specific loci in body fluids that can be informative of tumor initiation and progression and easily conducted by friendly and inexpensive PCR-based methods. A good example is the aberrant methylation of *SEPT9* in CRC that was first reported by Lofton-Day and colleagues in 2008 (19) and later identified as an informative screening marker in blood samples from affected patients (20, 21), allowing the detection of CRC at diverse stages and colonic locations. Another remarkable example is *VIM* (15, 18), which is the cornerstone of *ColoSure*<sup>TM</sup>, the only currently commercially or clinically available fecal DNA test marketed for CRC screening in the United States (33). Other studies have attempted to define comprehensive panels of DNA methylation biomarkers to test in stool samples, being frequently based on candidate gene approaches investigating frequent targets of aberrant hypermethylation in cancer (34). A compiler meta-analysis evaluating several studies reporting aberrantly methylated genes as biomarkers for CRC diagnosis reported an overall sensitivity and specificity of 0.54 and 0.88, respectively. The conclusion derived from this study inferred that panels reported up to date did not provide additional accuracy for the detection of colorectal neoplasia and underscored the need for additional biomarkers with better performance (35).

The panel of biomarkers identified in the present study equals or improves the detection power of previously reported genes including *PHACTR3* (55%–66%; ref. 36) and *VIM*. About the later, several works have been published reporting quite diverse sensitivity estimations of *VIM* for CRC detection in stool samples, ranging from 38% to 81% (reviewed in ref. 33). In the present work, *VIM* showed an intermediate value in relation to this range, which represented, however, the highest sensitivity obtained among all the markers when evaluated individually (55%). Specificity, although, was lower than those observed in the rest of markers (86%), which could be related to the relatively higher methylation levels observed for *VIM* in healthy mucosa. The better performance achieved by the 3-gene panel underscores the advantages of using a combination of markers identified from an agnostic procedure in front of the candidate gene approach.

This panel also increases the sensitivity provided by the <sup>m</sup>*SEPT9* assay in plasma, which is capable of identifying around 70% of patients with CRC at high specificity (21). Furthermore, we aimed to evaluate the diagnostic power of *SEPT9* in stool-based CRC diagnosis, as it had never been evaluated in this particular scenario. The sensitivity of this marker alone in stool DNA, however, was not as high as that reported in plasma (20%).

Our panel has similar diagnostic characteristics for CRC than others previously published, but we have proven that it may help in the early identification of colorectal dysplasia or cancer in high-risk patients with IBD. Patients with IBD, both ulcerative colitis (UC) and Crohn disease (CD) of the colon, are at an increased risk of developing CRC (37–39). Disease duration, extent, and severity of the inflammation are associated with an increased risk of CRC. Therefore, biomarkers that anticipate the development of full-blown tumors are needed to improve the follow-up and prognosis for these patients.

In sporadic CRC, the precursor lesion is the adenoma that is typically removed by simple endoscopic polypectomy. In contrast, dysplasia in patients with IBD can be polypoid or flat, localized, diffuse, or multifocal invisible for the endoscopist and difficult to detect for the pathologist, especially in areas of active inflammation (40, 41). In ulcerative colitis, promoter methylation seems to precede dysplasia and occurs throughout the mucosa of colitis (42–44) reinforcing the link between chronic inflammation and DNA methylation (45, 46) and could be a potential biomarker for testing to detect precancerous changes in routine surveillance biopsies. In line with these observations, we have been able to show that our panel of markers is hypermethylated in non-cancerous mucosa of patients with IBD-associated neoplasia being more than 3 times of that observed in sporadic cancer-adjacent mucosa. Surveillance colonoscopy (including careful inspection of the mucosa and multiple random biopsy sampling) in intervals of 1 to 2 years is recommended for these patients, and histologic assessment of dysplasia in mucosal biopsy specimens is considered the best marker of CRC risk in patients with IBD. However, its interpretation is subject to a high level of interobserver variability. Evaluation of our panel in these non-neoplastic biopsy specimens may contribute in a better follow-up and early diagnosis of neoplasia in these high-risk patients.

Among the 5 genes in our panel, abnormal methylation has been previously tested for *SLIT2* in the detection of neoplasia in high-risk ulcerative colitis and Crohn colitis patients by our group (47). *EYA4* has been proposed as biomarker for pancreatic cancer (48). The other 3 genes have been previously investigated in relation to several types of cancer but not in this particular scenario of diagnosis biomarker search.

One limitation of this study is that the series of patients analyzed did not include colorectal adenomas. We acknowledge that it is very important to analyze this group of patients to use the panel of biomarkers for the early detection of colorectal neoplasia. Also in our study, solid stools samples were not subject to a strict protocol regarding immediate storage temperature and collection buffer which may have affected the sensitivity of our approach. Despite that, our data provide evidence showing that the analysis of methylation in stool DNA may be useful for the diagnosis of colorectal carcinomas.

This panel of biomarkers should be tested in other studies and extended to larger cohorts of patients to confirm the

robustness of the findings. In addition, given the availability of platforms able to investigate DNA methylation in a genome-wide manner, a similar approach to the one followed here should be undertaken. The consecution of similar studies with increasing resolution and well-annotated clinical specimens would allow to define an optimal DNA methylation signature that refines the early detection of CRC, improving the management of the disease in the clinic and increasing the survival of the patients.

In conclusion, we provide a novel panel of specific methylation markers that can be assessed in stools and may complement currently applied protocols for the early detection of sporadic CRC. Moreover, our panel may contribute to improve the follow-up and early diagnosis of high-risk patients with IBD when assessed in non-neoplastic tissues obtained by surveillance colonoscopy.

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

### Authors' Contributions

**Conception and design:** F.J. Carmona, D. Azuara, F. Rodriguez-Moranta, V. Moreno

**Development of methodology:** F.J. Carmona, D. Azuara, A. Berenguer-Llargo, F. Rodriguez-Moranta, G. Capellá

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** F.J. Carmona, D. Azuara, A.F. Fernandez,

S. Biondo, J. de Oca, F. Rodriguez-Moranta, R. Salazar, A. Villanueva, M.F. Fraga, J. Guardiola, G. Capellá

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** D. Azuara, A. Berenguer-Llargo, A.F. Fernandez, F. Rodriguez-Moranta, R. Salazar, M.F. Fraga, J. Guardiola, G. Capellá, M. Esteller, V. Moreno

**Writing, review, and/or revision of the manuscript:** F.J. Carmona, D. Azuara, A. Berenguer-Llargo, S. Biondo, J. de Oca, F. Rodriguez-Moranta, R. Salazar, J. Guardiola, G. Capellá, M. Esteller, V. Moreno

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** F.J. Carmona, D. Azuara, F. Rodriguez-Moranta

**Study supervision:** G. Capellá, V. Moreno

### Acknowledgments

The authors thank Cristina Mangas and Rocío G. Urdinguio for technical support and helpful discussion.

### Grant Support

This study was supported by the Catalan Institute of Oncology, the Private Foundation of the Biomedical Research Institute of Bellvitge (IDIBELL), the Instituto de Salud Carlos III grants (FIS PI08-1635, PI08-1359, PI09-01037, PI11-01439, PI12-01420), CIBERESP CB07/02/2005, the Spanish Association Against Cancer (AECC) Scientific Foundation, the Catalan Government DURSI grant 2009SGR1489, and the European Commission FP7-COOP-Health-2007-B HiPerDART.

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 December 26, 2012; revised May 3, 2013; accepted May 7, 2013; published OnlineFirst May 21, 2013.

### References

- Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;127:2893-917.
- Mandel JS, Church TR, Bond JH, Ederer F, Geisser MS, Mongin SJ, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med* 2000;343:1603-7.
- Schoen RE, Pinsky PF, Weissfeld JL, Yokochi LA, Church T, Laiyemo AO, et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. *N Engl J Med* 2012;366:2345-57.
- Hol L, van Leerdam ME, van Ballegooijen M, van Vuuren AJ, van Dekken H, Reijerink JC, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. *Gut* 2010;59:62-8.
- Quintero E, Castells A, Bujanda L, Cubiella J, Salas D, Lanas A, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. *N Engl J Med* 2012;366:697-706.
- Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* 2008;135:1079-99.
- Ahlquist DA. Molecular detection of colorectal neoplasia. *Gastroenterology* 2010;138:2127-39.
- Berger BM, Ahlquist DA. Stool DNA screening for colorectal neoplasia: biological and technical basis for high detection rates. *Pathology* 2012;44:80-8.
- Zou H, Taylor WR, Harrington JJ, Hussain FT, Cao X, Loprinzi CL, et al. High detection rates of colorectal neoplasia by stool DNA testing with a novel digital melt curve assay. *Gastroenterology* 2009;136:459-70.
- Matsubara N. Epigenetic regulation and colorectal cancer. *Dis Colon Rectum* 2012;55:96-104.
- Lao VV, Grady WM. Epigenetics and colorectal cancer. *Nat Rev* 2011;8:686-700.
- Azuara D, Rodriguez-Moranta F, de Oca J, Soriano-Izquierdo A, Mora J, Guardiola J, et al. Novel methylation panel for the early detection of colorectal tumors in stool DNA. *Clin Colorectal Cancer* 2010;9:168-76.
- Osborn NK, Ahlquist DA. Stool screening for colorectal cancer: molecular approaches. *Gastroenterology* 2005;128:192-206.
- Ahlquist DA, Zou H, Domanico M, Mahoney DW, Yab TC, Taylor WR, et al. Next-generation stool DNA test accurately detects colorectal cancer and large adenomas. *Gastroenterology* 2012;142:248-56; quiz e25-6.
- Zou H, Harrington JJ, Shire AM, Rego RL, Wang L, Campbell ME, et al. Highly methylated genes in colorectal neoplasia: implications for screening. *Cancer Epidemiol Biomarkers Prev* 2007;16:2686-96.
- Glockner SC, Dhir M, Yi JM, McGarvey KE, Van Neste L, Louwagie J, et al. Methylation of TFP12 in stool DNA: a potential novel biomarker for the detection of colorectal cancer. *Cancer Res* 2009;69:4691-9.
- Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, et al. High-throughput DNA methylation profiling using universal bead arrays. *Genome Res* 2006;16:383-93.
- Chen WD, Han ZJ, Skoletsky J, Olson J, Sah J, Myeroff L, et al. Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. *J Natl Cancer Inst* 2005;97:1124-32.
- Lofton-Day C, Model F, Devos T, Tetzner R, Distler J, Schuster M, et al. DNA methylation biomarkers for blood-based colorectal cancer screening. *Clin Chem* 2008;54:414-23.
- Grutzmann R, Molnar B, Pilarsky C, Habermann JK, Schlag PM, Saeger HD, et al. Sensitive detection of colorectal cancer in peripheral blood by septin 9 DNA methylation assay. *PLoS ONE* 2008;3:e3759.
- deVos T, Tetzner R, Model F, Weiss G, Schuster M, Distler J, et al. Circulating methylated SEPT9 DNA in plasma is a biomarker for colorectal cancer. *Clin Chem* 2009;55:1337-46.
- Moreno V, Gemignani F, Landi S, Gioia-Patricola L, Chabrier A, Blanco I, et al. Polymorphisms in genes of nucleotide and base excision repair: risk and prognosis of colorectal cancer. *Clin Cancer Res* 2006;12:2101-8.
- Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002;30:207-10.

24. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, et al. NCBI GEO: archive for high-throughput functional genomic data. *Nucleic Acids Res* 2009;37:D885–90.
25. Kim YH, Lee HC, Kim SY, Yeom YI, Ryu KJ, Min BH, et al. Epigenomic analysis of aberrantly methylated genes in colorectal cancer identifies genes commonly affected by epigenetic alterations. *Ann Surg Oncol* 2011;18:2338–47.
26. Kibriya MG, Raza M, Jasmine F, Roy S, Paul-Brutus R, Rahaman R, et al. A genome-wide DNA methylation study in colorectal carcinoma. *BMC Med Genomics* 2011;4:50.
27. Bibikova M, Le J, Barnes B, Saedinia-Melnyk S, Zhou L, Shen R, et al. Genome-wide DNA methylation profiling using Infinium(R) assay. *Epigenomics* 2009;1:177–200.
28. Puig P, Urgell E, Capella G, Villanueva A, Grau M, Sancho FJ, et al. Improved detection of K-ras codon 12 mutations in fecal exfoliated cells. *Lab Invest* 1999;79:617–8.
29. Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL, et al. Highly parallel SNP genotyping. *Cold Spring Harb Symp Quant Biol* 2003;68:69–78.
30. Canty A, Ripley B. boot: Bootstrap R (S-Plus) Functions. R package version 1.3-2. Vienna, Austria: R Foundation for Statistical Computing; 2011.
31. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: Visualizing the performance of scoring classifiers. R package version 1.0-4. Vienna, Austria: R Foundation for Statistical Computing; 2009.
32. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011.
33. Ned RM, Melillo S, Marrone M. Fecal DNA testing for colorectal cancer screening: the ColoSure™ test. *PLoS Curr* 2011;3:RRN1220.
34. Kamimae S, Yamamoto E, Yamano HO, Nojima M, Suzuki H, Ashida M, et al. Epigenetic alteration of DNA in mucosal wash fluid predicts invasiveness of colorectal tumors. *Cancer Prev Res* 2011;4:674–83.
35. Luo YX, Chen DK, Song SX, Wang L, Wang JP. Aberrant methylation of genes in stool samples as diagnostic biomarkers for colorectal cancer or adenomas: a meta-analysis. *Int J Clin Prac* 2011;65:1313–20.
36. Bosch LJ, Oort FA, Neerincx M, Khalid-de Bakker CA, Terhaar sive Droste JS, Melotte V, et al. DNA methylation of phosphatase and actin regulator 3 detects colorectal cancer in stool and complements FIT. *Cancer Prev Res* 2012;5:464–72.
37. Choi PM, Zelig MP. Similarity of colorectal cancer in Crohn's disease and ulcerative colitis: implications for carcinogenesis and prevention. *Gut* 1994;35:950–4.
38. Ekobom A, Helmick C, Zack M, Adami HO. Ulcerative colitis and colorectal cancer. A population-based study. *N Engl J Med* 1990;323:1228–33.
39. Langholz E, Munkholm P, Davidsen M, Binder V. Colorectal cancer risk and mortality in patients with ulcerative colitis. *Gastroenterology* 1992;103:1444–51.
40. Itzkowitz SH, Present DH. Consensus conference: colorectal cancer screening and surveillance in inflammatory bowel disease. *Inflamm Bowel Dis* 2005;11:314–21.
41. Itzkowitz SH, Yio X. Inflammation and cancer IV. Colorectal cancer in inflammatory bowel disease: the role of inflammation. *Am J Physiol* 2004;287:G7–17.
42. Fujii S, Tominaga K, Kitajima K, Takeda J, Kusaka T, Fujita M, et al. Methylation of the oestrogen receptor gene in non-neoplastic epithelium as a marker of colorectal neoplasia risk in longstanding and extensive ulcerative colitis. *Gut* 2005;54:1287–92.
43. Garrity-Park MM, Loftus EV Jr, Sandborn WJ, Bryant SC, Smyrk TC. Methylation status of genes in non-neoplastic mucosa from patients with ulcerative colitis-associated colorectal cancer. *Am J Gastroenterol* 2010;105:1610–9.
44. Wang FY, Arisawa T, Tahara T, Takahama K, Watanabe M, Hirata I, et al. Aberrant DNA methylation in ulcerative colitis without neoplasia. *Hepatogastroenterology* 2008;55:62–5.
45. Nakajima T, Enomoto S, Yamashita S, Ando T, Nakanishi Y, Nakazawa K, et al. Persistence of a component of DNA methylation in gastric mucosae after *Helicobacter pylori* eradication. *J Gastroenterol* 2010;45:37–44.
46. Saito S, Kato J, Hiraoka S, Horii J, Suzuki H, Higashi R, et al. DNA methylation of colon mucosa in ulcerative colitis patients: correlation with inflammatory status. *Inflamm Bowel Dis* 2011;17:1955–65.
47. Azuara D, Rodriguez-Moranta F, de Oca J, Sanjuan X, Guardiola J, Lobaton T, et al. Novel methylation panel for the early detection of neoplasia in high-risk ulcerative colitis and Crohn's colitis patients. *Inflamm Bowel Dis* 2013;19:165–73.
48. Kisiel JB, Yab TC, Taylor WR, Chari ST, Petersen GM, Mahoney DW, et al. Stool DNA testing for the detection of pancreatic cancer: assessment of methylation marker candidates. *Cancer* 2012;118:2623–31.

# Cancer Prevention Research

## DNA Methylation Biomarkers for Noninvasive Diagnosis of Colorectal Cancer

F. Javier Carmona, Daniel Azuara, Antonio Berenguer-Llargo, et al.

*Cancer Prev Res* Published OnlineFirst May 21, 2013.

<b>Updated version</b>	Access the most recent version of this article at: doi: <a href="https://doi.org/10.1158/1940-6207.CAPR-12-0501">10.1158/1940-6207.CAPR-12-0501</a>
<b>Supplementary Material</b>	Access the most recent supplemental material at: <a href="http://cancerpreventionresearch.aacrjournals.org/content/suppl/2013/05/22/1940-6207.CAPR-12-0501.DC1">http://cancerpreventionresearch.aacrjournals.org/content/suppl/2013/05/22/1940-6207.CAPR-12-0501.DC1</a>

<b>E-mail alerts</b>	<a href="#">Sign up to receive free email-alerts</a> related to this article or journal.
----------------------	------------------------------------------------------------------------------------------

<b>Reprints and Subscriptions</b>	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at <a href="mailto:pubs@aacr.org">pubs@aacr.org</a> .
-----------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------

<b>Permissions</b>	To request permission to re-use all or part of this article, use this link <a href="http://cancerpreventionresearch.aacrjournals.org/content/early/2013/06/17/1940-6207.CAPR-12-0501">http://cancerpreventionresearch.aacrjournals.org/content/early/2013/06/17/1940-6207.CAPR-12-0501</a> . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.
--------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------