



## RESEARCH ARTICLE

# Genome-wide DNA methylation analysis of colorectal adenomas with and without recurrence reveals an association between cytosine-phosphate-guanine methylation and histological subtypes

David Fiedler<sup>1</sup> | Daniela Hirsch<sup>1</sup> | Nady El Hajj<sup>2,3</sup> | Howard H. Yang<sup>4</sup> | Yue Hu<sup>5</sup> | Carsten Sticht<sup>6</sup> | Indrajit Nanda<sup>2</sup> | Sebastian Belle<sup>7,8</sup> | Josef Rueschoff<sup>9</sup> | Maxwell P. Lee<sup>4</sup> | Thomas Ried<sup>5</sup>  | Thomas Haaf<sup>2</sup> | Timo Gaiser<sup>1</sup> 

<sup>1</sup>Institute of Pathology, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

<sup>2</sup>Institute of Human Genetics, Julius Maximilians University, Würzburg, Germany

<sup>3</sup>College of Health and Life Sciences, Hamad Bin Khalifa University, Education City, Doha, Qatar

<sup>4</sup>High Dimension Data Analysis Group, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

<sup>5</sup>Genetics Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

<sup>6</sup>Center for Medical Research, Bioinformatic and Statistic, Medical Faculty Mannheim, Mannheim, Germany

<sup>7</sup>Department of Internal Medicine II, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

<sup>8</sup>Central Interdisciplinary Endoscopy Unit, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

<sup>9</sup>Institute of Pathology Nordhessen, Kassel, Germany

## Correspondence

Timo Gaiser, MD, Institute of Pathology, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany.  
Email: timo.gaiser@umm.de

## Abstract

Aberrant methylation of DNA is supposed to be a major and early driver of colonic adenoma development, which may result in colorectal cancer (CRC). Although gene methylation assays are used already for CRC screening, differential epigenetic alterations of recurring and nonrecurring colorectal adenomas have yet not been systematically investigated. Here, we collected a sample set of formalin-fixed paraffin-embedded colorectal low-grade adenomas (n = 72) consisting of primary adenomas without and with recurrence (n = 59), recurrent adenomas (n = 10), and normal mucosa specimens (n = 3). We aimed to unveil differentially methylated CpG positions (DMPs) across the methylome comparing not only primary adenomas without recurrence vs primary adenomas with recurrence but also primary adenomas vs recurrent adenomas using the Illumina Human Methylation 450K BeadChip array. Unsupervised hierarchical clustering exhibited a significant association of methylation patterns with histological adenoma subtypes. No significant DMPs were identified comparing primary adenomas with and without recurrence. Despite that, a total of 5094 DMPs (false discovery rate <0.05; fold change >10%) were identified in the comparisons of recurrent adenomas vs primary adenomas with recurrence (674; 98% hypermethylated), recurrent adenomas vs primary adenomas with and without recurrence (241; 99% hypermethylated) and colorectal adenomas vs normal mucosa (4179; 46% hypermethylated). DMPs in cytosine-phosphate-guanine (CpG) islands were frequently hypermethylated, whereas open sea- and shelf-regions exhibited hypomethylation. Gene ontology analysis revealed enrichment of genes associated with the immune system, inflammatory processes, and cancer pathways. In conclusion, our methylation data could assist in establishing a more robust and reproducible

**ABBREVIATIONS:** BMP, bone morphogenetic protein; CGI, CpG island; CIMP, CpG island mismatch repair phenotype; CIN, chromosomal instability; CNA, copy number alteration; CpG, cytosine-phosphate-guanine; CRC, colorectal cancer; DMP, differentially methylated CpG position; ECM, extracellular matrix; FDR, false discovery rate; FFPE, formalin-fixed paraffin-embedded; GSEA, gene set enrichment analysis; H&E, hematoxylin and eosin; LGD, low-grade dysplasia; PCA, principal component analysis; SNP, single nucleotide polymorphism; TSS, transcription start site; UTR, untranslated region.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Genes, Chromosomes and Cancer* published by Wiley Periodicals, Inc.

**Funding information**

Translational Physician Scientist Program, Medical Faculty Mannheim, Heidelberg University, Grant/Award Number: Intramural Research Scholarship; University of Heidelberg; Boehringer Ingelheim Fonds, Grant/Award Number: Travel Grant

histological adenoma classification, which is a prerequisite for improving surveillance guidelines.

**KEYWORDS**

adenoma, DNA methylation, epigenetics, histological subtype, recurrence

## 1 | INTRODUCTION

Colorectal cancer (CRC) has the third highest cancer incidence worldwide with approximately 1.8 million new cases in 2018.<sup>1</sup> As CRC develops mainly via the adenoma-carcinoma sequence, endoscopic resection of adenomas reduces CRC risk significantly.<sup>2</sup> However, recurrences of initially resected polyps are common and contribute to the risk of developing CRC.<sup>3</sup> Many guidelines define postpolypectomy colonoscopy surveillance intervals,<sup>4-6</sup> but the scientific rationale for the different intervals is vague. Therefore, biomarkers for the prediction of recurrences could help to identify patients at risk and to define improved surveillance intervals.

As CRC develops over several years, molecular evidence has accumulated that the disease is driven by the acquisition of genetic and epigenetic alterations. Several distinct molecular pathways have been described: (a) the chromosomal instability (CIN) pathway that is associated with aneuploidy, copy number alterations (CNAs), and mutations most frequently affecting *APC*, *TP53*, and *KRAS*<sup>7,8</sup>; (b) the microsatellite instability pathway that is caused by DNA hypermutation<sup>9</sup>; (c) the CpG island mismatch repair phenotype (CIMP), which is defined by extensive methylation of CpG islands.<sup>10</sup> Abnormal DNA methylation is supposed to happen very early and affects up to 85% of adenomas and CRCs.<sup>11</sup> These tumors exhibit pathological methylation patterns predominantly affecting canonical and non-canonical WNT-pathway genes (eg, *AXIN2*, *CTNNB1*, *SFRP1*, and *SFRP2*) and Polycomb group genes.<sup>12,13</sup> Thus, it is reasonable to assume that adenoma recurrence after polypectomy might at least partially be triggered by methylation pattern changes.

The evolution from colorectal adenoma to CRC is associated with increasing hypermethylation of CpG islands (CGI) in promoter regions of tumor suppressor genes and concurrent global hypomethylation of CpG sites.<sup>14</sup> DNA hypomethylation might contribute to CIN because of the reactivation of transposable elements and chromosomal rearrangements.<sup>15,16</sup> Additionally, it is known that aberrant methylation of CpGs in gene bodies contributes to dysregulation of gene expression.<sup>17,18</sup> Recent advances in analytical technologies have permitted the assessment of genome-wide methylation patterns.<sup>19</sup> Consequently, multiple studies have shed light on the methylome landscapes of CRCs and determined the contribution of altered methylation to the development and progression of CRC.<sup>13,20,21</sup> The aberrant methylation of DNA also affects adenomas.<sup>22</sup> However, differential epigenetic alterations of recurring and nonrecurring colorectal adenomas have not yet been systematically and comprehensively investigated. In this study, we aimed to unveil differentially methylated CpG positions (DMPs) across the

genome of formalin-fixed paraffin-embedded (FFPE) colorectal specimens comparing primary adenomas with recurrence vs primary adenomas without recurrence, primary adenomas vs recurrent adenomas, and colorectal adenomas vs normal mucosa. We utilized the Illumina Infinium Human Methylation 450K BeadChip array (HM450K), covering more than 485 000 genome-wide distributed CpG sites, to analyze the distribution of DMPs across the genome. A pathway enrichment analysis was conducted to uncover the most frequently altered pathways and hence to assess the biological functions of the DMPs. Selected markers were subsequently quantitatively analyzed by pyrosequencing to validate the array results.

## 2 | MATERIALS AND METHODS

### 2.1 | Clinical samples

The study was approved by the local ethics committee of the Medical Faculty Mannheim of the University of Heidelberg, Germany (2012-608R-MA). All adenoma and normal colonic mucosa samples were collected between 2002 and 2014 and retrieved from the FFPE tissue archive from the Institute of Pathology of the University Medical Center, Mannheim. A total of 69 colorectal adenoma and 3 normal colonic mucosa specimens were selected. Although pathological evaluation for in toto polypectomy was not possible because of the fragmentation of the tissue, endoscopic removal was indicated as complete based on thorough clinical assessment for all lesions. Recurrent adenomas were defined by the endoscopist (SB) if a scar was present in connection with the new adenoma formation and/or if the same anatomical position as described in centimeters from the anus was ensured. The median observation time after polypectomy was 21.0 months (IQR, 10.9-36.3 months). The analyzed study cohort consisted of four sample groups (Table 1): (a) primary adenomas that did not unveil a tumor in the observation period ( $n = 30$ ; *primary adenoma without recurrence*); (b) primary adenomas that exhibited a documented recurrence of the adenoma at the same location during the follow-up period ( $n = 29$ , *primary adenoma with recurrence*); (c) recurrent adenoma ( $n = 10$ ; *matched pairs corresponding to 10 of the primary adenomas with recurrence*); and (d) normal colonic mucosa ( $n = 3$ ). Pathological classification was done, in accordance with the current WHO classification from 2010,<sup>23</sup> by two board-certified pathologists blinded to all data (TG/JR). Histological classification discerned tubular, tubulovillous, or villous adenomas (Table 2). The adenomas exhibited cytological and histological alterations fitting the WHO definition of low-grade dysplasia (LGD).

## 2.2 | Tumor sample preparation

Consecutive sectioning was performed on archived FFPE tissue blocks. The first tissue slide was stained by hematoxylin and eosin (H&E) and the region of interest (>70% tumor content) was marked. Two consecutive FFPE sections (10  $\mu$ m thick) were deparaffinized in xylene for 10 minutes and rehydrated in an ethanol series for 10 minutes. The tumor area was macrodissected with a scalpel under the guidance of the marked H&E slide. Genomic DNA was extracted using the Gentra Puregene Tissue Kit (Qiagen, Hilden, Germany) and subsequently spectrophotometrically quantified by NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA). Eluted DNAs were purified by applying the DNA Clean & Concentrator Kit (Zymo Research, Irvine, CA) before a quality check was performed using the Infinium HD FFPE QC Real-time PCR assay (Illumina, San Diego, CA) as indicated in the manufacturer's protocol.

## 2.3 | DNA methylation microarray analysis

DNAs (250 ng) of 69 colorectal adenomas and three normal mucosae were subjected to bisulfite conversion using the EZ-96 DNA Methylation Kit (Zymo Research). Converted DNAs were hybridized to six HM450K arrays (Illumina) by sequential whole-genome amplification,

enzymatic fragmentation, precipitation, and resuspension according to the manufacturer's instructions. Samples were randomly distributed on the arrays and hybridized for 20 hours at 48°C in a hybridization oven. Subsequently, arrays were treated with a primer extension, an immunohistochemical staining, and a coating as indicated by the supplier's protocol (Illumina).<sup>19</sup> Probe signals were detected by an iScan array scanner (Illumina) and microarray data (NCBI GEO Accession No. GSE129364) were exported as idat files. Quality control metrics for samples and probes were assessed by the GenomeStudio software version 1.8 (Illumina). Signal intensity normalization, background level correction, and color adjustments were conducted by applying the SWAN normalization method of the Bioconductor R-package minfi.<sup>24,25</sup> Starting from a total of 485 577 array probes, 156 004 probes were excluded because of one of the following criteria: (a) cross-reactive polymorphic CpG sites<sup>26</sup>; (b) nonspecific probes<sup>26</sup>; (c) age-dependent CpG positions<sup>27</sup>; and (d) single nucleotide polymorphisms (SNPs) near the targeted CpG site.<sup>28</sup> Filtering left a total of 329 573 array probes for analysis. Methylation levels were calculated as  $\beta$  values according to the following formula<sup>29</sup>:

$$\beta = \frac{\text{Intensity(methylated)}}{\text{Intensity(methylated)} + \text{intensity(unmethylated)} + 100} \quad (1)$$

**TABLE 1** Distribution of age at diagnosis, gender, adenoma location, histology, size, and observation time of patients with primary adenomas without recurrence and with recurrence

Variable	Primary adenomas without recurrence (n = 30)	Primary adenomas with recurrence (n = 29)	P value
Age at diagnosis (y)			
Mean $\pm$ SD	67.3 $\pm$ 7.5	64.9 $\pm$ 11.7	.350 <sup>a</sup>
Median (IQR)	67.3 (64.4-70.2)	64.3 (55.9-70.7)	
Gender			
Female	14	14	1.000 <sup>b</sup>
Male	16	15	
Location			
Right hemicolon	16	17	.405 <sup>c</sup>
Left hemicolon	7	3	
Rectum	7	9	
Histology			
Tubular	11	4	.072 <sup>b</sup>
Tubulovillous/villous	19	25	
Size (mm)			
Mean $\pm$ SD	19.6 $\pm$ 8.8	32.8 $\pm$ 20.8	.003 <sup>a</sup>
Median (IQR)	20.0 (12.0-20.0)	30.0 (20.0-30.0)	
Observation time/ recurrence-free time (m)			
Mean $\pm$ SD	26.9 $\pm$ 15.2	22.5 $\pm$ 16.6	.380 <sup>a</sup>
Median (IQR)	22.9 (13.7-22.9)	19.9 (8.0-19.9)	

<sup>a</sup>Student's *t* test.

<sup>b</sup>Fisher's exact test.

<sup>c</sup>Freeman-Halton test.

**TABLE 2** Clinicopathological features of the primary adenomas and recurrent lesions

	Sample ID	Gender	Age at diagnosis (y)	Grade of dysplasia	Recurrence	Observation time/recurrence (m)	Histology	Location	Size (mm)	
Primary adenoma w/o recurrence	A 1	M	66.7	Low grade	No	27.1	Tubulovillous	Ascending	20	
	A 2	M	66.0	Low grade	No	12.1	Tubular	Cecum	18	
	A 3	M	50.3	Low grade	No	14.1	Tubulovillous	Rectum	35	
	A 4	M	67.2	Low grade	No	38.0	Tubular	Descending	20	
	A 5	M	67.3	Low grade	No	47.2	Tubulovillous	Hepatic flexure	20	
	A 6	M	68.6	Low grade	No	19.6	Villous	Rectum	40	
	A 7	F	63.9	Low grade	No	23.9	Tubulovillous	Cecum	25	
	A 8	M	70.3	Low grade	No	45.4	Tubulovillous	Sigmoid	12	
	A 9	F	68.2	Low grade	No	21.9	Tubulovillous	Ascending	20	
	A 10	M	62.8	Low grade	No	20.4	Villous	Rectum	10	
	A 11	F	68.5	Low grade	No	47.5	Tubulovillous	Sigmoid	30	
	A 12	M	66.3	Low grade	No	3.4	Tubular	Cecum	40	
	A 13	F	71.8	Low grade	No	12.5	Tubulovillous	Ascending	12	
	A 14	M	50.9	Low grade	No	10.9	Tubular	Transverse	15	
	A 15	M	66.6	Low grade	No	36.5	Tubular	Ascending	25	
	A 31	F	66.3	Low grade	No	ND	Tubulovillous	Rectum	20	
	A 32	F	83.6	Low grade	No	50.6	Tubulovillous	Transverse	20	
	A 33	F	69.7	Low grade	No	ND	Tubulovillous	Cecum	30	
	A 34	M	67.3	Low grade	No	ND	Tubular	Rectum	10	
	A 35	F	75.4	Low grade	No	ND	Tubular	Sigmoid	8	
	A 36	M	76.0	Low grade	No	ND	Tubulovillous	Sigmoid	5	
	A 37	M	67.9	Low grade	No	ND	Tubular	Rectum	12	
	A 38	M	74.1	Low grade	No	ND	Tubulovillous	Ascending	20	
	A 39	F	84.4	Low grade	No	ND	Tubulovillous	Cecum	20	
	A 40	F	62.9	Low grade	No	ND	Tubulovillous	Ascending	15	
	A 41	F	63.4	Low grade	No	ND	Tubulovillous	Hepatic flexure	20	
	A 42	F	58.2	Low grade	No	ND	Tubular	Cecum	12	
	A 43	F	55.7	Low grade	No	ND	Tubulovillous	Rectum	23	
	A 44	F	71.0	Low grade	No	ND	Tubular	Sigmoid	12	
	A 45	M	69.0	Low grade	No	ND	Tubular	Descending	20	
	Primary adenoma w/ recurrence	A 16	M	69.2	Low grade	Primary	24.2	Tubulovillous	Ascending	15
		A 17	M	68.5	Low grade	Primary	7.7	Tubulovillous	Rectum	20
		A 18	M	55.1	Low grade	Primary	30.4	Tubulovillous	Sigmoid	42
		A 19	M	64.9	Low grade	Primary	6.6	Tubulovillous	Cecum	20
		A 20	F	70.2	Low grade	Primary	41.8	Tubulovillous	Hepatic flexure	40
		A 21	M	62.7	Low grade	Primary	21.0	Tubular	Ascending	20
		A 22	M	73.0	Low grade	Primary	8.0	Tubulovillous	Rectum	25
		A 23	F	64.9	Low grade	Primary	17.2	Tubular	Transverse	6
		A 24 <sup>a</sup>	F	45.3	Low grade	Primary	19.9	Tubulovillous	Ascending	25
		A 25	M	61.9	Low grade	Primary	5.3	Tubulovillous	Transverse	ND
		A 26	F	79.0	Low grade	Primary	60.1	Tubulovillous	Transverse	20
		A 27 <sup>a</sup>	M	87.9	Low grade	Primary	19.1	Tubulovillous	Rectum	94
		A 28	F	55.9	Low grade	Primary	8.4	Tubulovillous	Rectum	35
		A 29 <sup>a</sup>	M	37.9	Low grade	Primary	24.9	Tubulovillous	Cecum	40
		A 30	M	78.0	Low grade	Primary	14.5	Tubulovillous	Rectum	30
A 46		F	52.3	Low grade	Primary	3.1	Tubulovillous	Rectum	63	
A 47		F	58.8	Low grade	Primary	54.2	Tubulovillous	Cecum	40	
A 48		M	52.0	Low grade	Primary	7.0	Tubulovillous	Ascending	ND	
A 49		M	63.6	Low grade	Primary	52.7	Tubular	Ascending	7	
A 50		F	72.9	Low grade	Primary	26.6	Tubulovillous	Cecum	12	

(Continues)

TABLE 2 (Continued)

	Sample ID	Gender	Age at diagnosis (y)	Grade of dysplasia	Recurrence	Observation time/recurrence (m)	Histology	Location	Size (mm)
	A 51	F	70.7	Low grade	Primary	36.3	Tubular	Cecum	18
	A 52	M	63.5	Low grade	Primary	3.2	Tubulovillous	Rectum	40
	A 53 <sup>a</sup>	F	69.8	Low grade	Primary	32.6	Tubulovillous	Cecum	35
	A 54 <sup>a</sup>	F	64.3	Low grade	Primary	1.9	Tubulovillous	Rectum	20
	P 1a <sup>a</sup>	F	81.9	Low grade	Primary	26.4	Tubulovillous	Rectum	84
	P 2a <sup>a</sup>	F	53.6	Low grade	Primary	29.1	Tubulovillous	Sigmoid	45
	P 3a <sup>a</sup>	F	54.4	Low grade	Primary	12.8	Tubulovillous	Sigmoid	20
	P 4a <sup>a</sup>	M	63.8	Low grade	Primary	47.4	Tubulovillous	Cecum	30
	P 6a <sup>a</sup>	M	87.1	Low grade	Primary	9.5	Tubulovillous	Cecum	40
Recurrent adenoma	A 24b	F	46.8	Low grade	Recurrent	19.9	Tubular	Ascending	10
	A 27b	M	89.3	Low grade	Recurrent	19.1	Tubulovillous	Rectum	ND
	A 29b	M	39.8	Low grade	Recurrent	24.9	Tubular	Cecum	ND
	A 53b	F	72.5	Low grade	Recurrent	32.6	Tubulovillous	Cecum	15
	A 54b	F	64.4	Low grade	Recurrent	1.9	Tubulovillous	Rectum	ND
	P 1b	F	84.1	Low grade	Recurrent	26.4	Tubulovillous	Rectum	20
	P 2b	F	56.0	Low grade	Recurrent	29.1	Tubulovillous	Sigmoid	ND
	P 3b	F	55.4	Low grade	Recurrent	12.8	Tubulovillous	Sigmoid	ND
	P 4b	M	67.6	Low grade	Recurrent	47.4	Tubulovillous	Cecum	30
	P 6b	M	87.9	Low grade	Recurrent	9.5	Tubulovillous	Cecum	19
Normal	N 1	M	51.7	Low grade	Normal	ND	ND	Descending	ND
	N 2	M	52.6	Low grade	Normal	ND	ND	Transverse	ND
	N 3	M	69.1	Low grade	Normal	ND	ND	Transverse	ND

Abbreviations: F, female; M, male; ND, not determined.

<sup>a</sup>Matched recurrent adenoma was analyzed.

## 2.4 | Microarray data analysis

Principal component analysis (PCA) and unsupervised hierarchical clustering of the samples were conducted using  $P < .05$  in a two-way cluster analysis (Euclidean distance and Ward's group linkage). Clinico-pathological data were correlated via Spearman's correlation coefficient, whereas categorical variables were tested using the Freeman-Halton test.

Methylation levels expressed as  $\beta$  values were evaluated to identify significant DMPs applying the Benjamini Hochberg-corrected false discovery rates (FDR adjusted) of  $q < 0.05$  and  $q < 0.01$  using the R limma package. Significant DMPs were sought with a fold change of  $>10\%$  by comparing sample sets of primary adenomas without recurrence, primary adenomas with recurrence, recurrent adenomas, and normal mucosa. Preselected CpG sites were filtered for *top genes*, that is, the most differentially hyper- or hypomethylated CpG positions that harbor the greatest potential as informative biomarkers. We considered probes with  $\Delta\beta \geq |0.1|$  and which are not located on sex chromosomes to eliminate a gender-specific bias.

Overlapping DMPs across the comparisons were identified using the Venn diagram plotter software version 1.5.5 (<https://omics.pnl.gov>). The distribution of DMPs was evaluated by categorizing the probes' positions according to their linear location to the nearest CpG island: CGI (stretch 0.5-2 kb), CpG shores ( $<2$  kb from CGI), CpG shelves (2-4 kb from CGI), and open sea ( $>4$  kb from CGI).<sup>30</sup>

Furthermore, DMP distribution was investigated clustering the probes into gene regions comprising the proximal promoter [transcription start sites (TSS-1500, TSS-200), 5' untranslated region (5'UTR), first exon (1st exon)], gene body, and 3' untranslated region (3'UTR). The chromosomal distribution of DMPs was assessed by plotting the number of hyper- and hypomethylated probes per chromosome.

A gene set enrichment analysis (GSEA) of the *top gene* DMPs was conducted by correlating (applying an FDR of  $q < 0.001$ ) the CpG positions with the associated biological pathways of curated gene sets using the Molecular Signatures Database (MSigDB; <http://software.broadinstitute.org/gsea/msigdb>).<sup>31</sup>

## 2.5 | Bisulfite pyrosequencing of selected CpG loci

The obtained methylation array results were orthogonally validated via bisulfite pyrosequencing targeting eight CpG loci within a distinct gene body region of *GREM2* (Figure S1). The gene was selected for validation as two of these CpG loci were listed among the most significant DMPs (cg.id = cg01809217; cg.id = cg02577267). Customized forward and reverse assays were designed using the PyroMark Assay Design software (Qiagen) (Table S1).

Bisulfite conversion of up to 1  $\mu\text{g}$  of DNA was conducted by the EpiTect Plus DNA Bisulfite Kit (Qiagen). Pre-PCR for initial fragment amplification of 100 ng bisulfite-DNA was performed using the PyroMark PCR Kit (Qiagen) according to the manufacturer's protocol.

Additionally, biotinylated primers (Sigma-Aldrich, St. Louis, MO) were used for forward and reverse assays, respectively. The thermal profile comprised 95°C for 15 minutes, accompanied by 45 cycles of 94°C for 30 seconds, 56°C for 30 seconds, and 72°C for 30 seconds prior to a final extension at 72°C for 10 minutes.

Next, the bead solution (5  $\mu$ L Sepharose beads added to 35  $\mu$ L PyroMark binding buffer) and the sequencing primer solution (0.8  $\mu$ L primer [20  $\mu$ M] added to 39.2  $\mu$ L PyroMark annealing buffer) (Table S1) were set up. The primer solution was applied into the wells of the sequencing plate while the bead-solution was mixed with the pre-PCR product (40  $\mu$ L) and implemented in the PyroMark Q24 vacuum workstation (Qiagen) as recommended by the supplier. Prior to initiation of the sequencing reaction, the components of the PyroMark Gold Q96 SQA Reagents Kit (Qiagen) were applied onto the PyroMark Q24 cartridge (Qiagen) according to the manufacturer's protocol. The cartridge and the sequencing plate were loaded into the PyroMark Q24 device (Qiagen) and the reaction was started. The results were analyzed by the PyroMark-Q-CpG software (Qiagen) as mean methylation percentage per CpG loci. Sequencing reactions included bisulfite modification controls. A technical quality check for robustness and reliability was routinely performed assessing internal control DNAs. Methylation ratios were trichotomized (hypomethylated <33%, medium 33-66%, and hypermethylated >66%), color coded, and plotted per sample.<sup>32</sup> Comparisons of methylation ratios of the CpG positions across various groups were tested via the Mann-Whitney *U* test and Kruskal-Wallis test, respectively.

### 3 | RESULTS

#### 3.1 | Methylation array analysis

A total of 69 colorectal adenomas and 3 normal mucosa specimens were analyzed for DNA methylation using the HM450K array. A PCA of the average  $\beta$ -value per sample exhibited a widespread distribution that did not allow the separation of the sample groups according to recurrence (Figure 1A). In addition, we performed an unsupervised hierarchical cluster analysis based on the average  $\beta$  values of the probes resulting in three major clusters (Figure 1B). Cluster 3 showed a stronger divergence compared with clusters 1 and 2. The sample annotation revealed a very weak tendency to link cluster 3 to the status of recurrence ( $P = .171$ ; Figure 2A). With the objective to find an explanation for the cluster affiliation of the samples, we dichotomized the clusters (ie, clusters 1 and 2 vs 3) and investigated for correlations with the clinical parameters: neither patient age ( $P = .496$ ) nor adenoma localization ( $P = .837$ ), nor right- and left-sided location ( $P = .869$ ), nor adenoma size ( $P = .671$ ), nor patient gender ( $P = .511$ ) exhibited a correlation. Interestingly, dichotomized clustering revealed a significant correlation to the histological subtype ( $P = .008$ ; Figure 2B). Although clusters 1 and 2 were dominated by 83.3% (40/48) of adenomas with tubulovillous/villous histology, cluster 3 harbored a balanced number of tubulovillous/villous (50%; 12/24), and tubular adenomas (41.7%; 10/24), respectively. The comparison of the clusters 1 + 2 vs 3 by mean  $\beta$  values per sample provided a

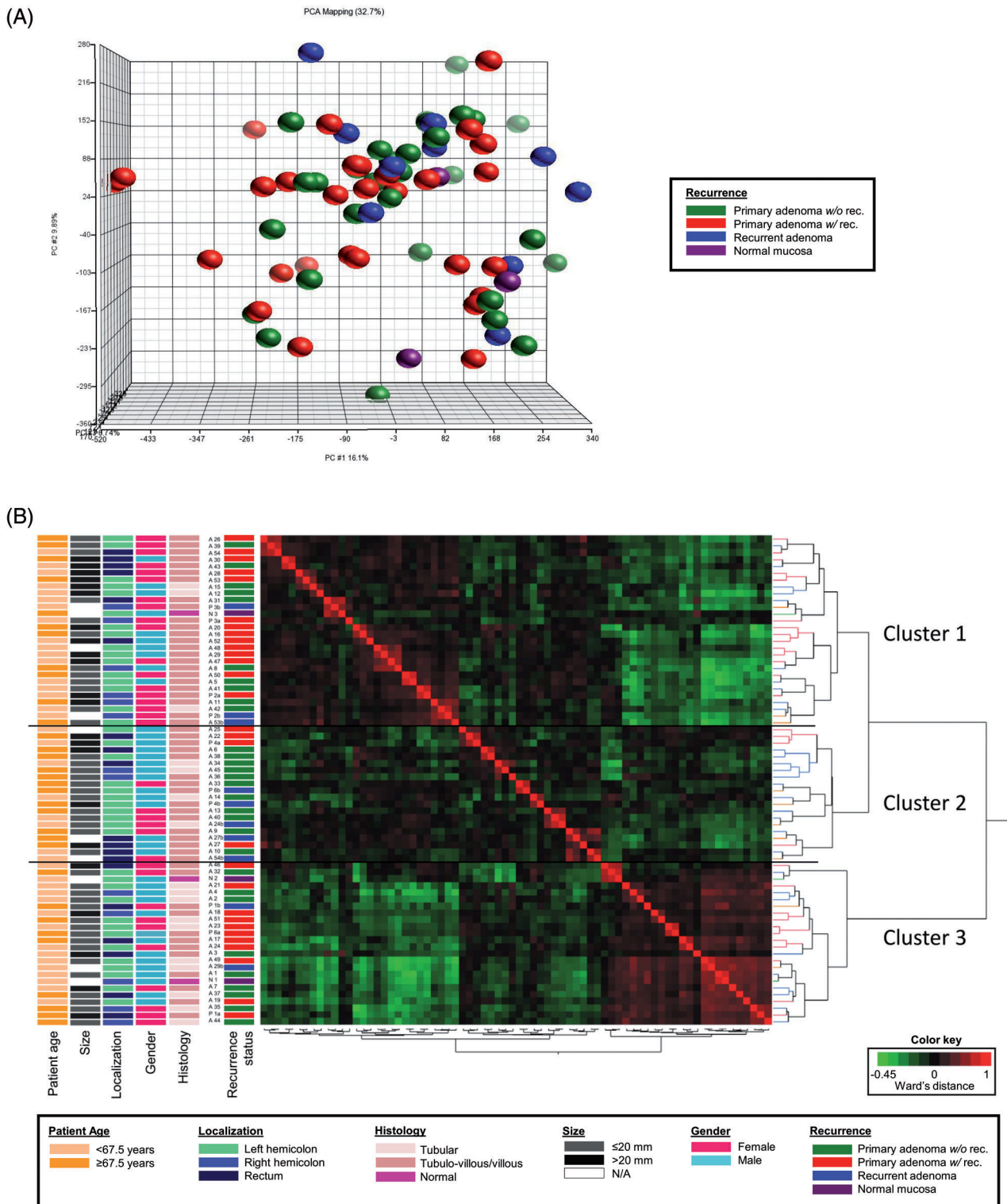
differential methylation of  $\Delta\text{mean}_\beta = 0.017$  ( $0.676 \pm 0.036$  vs  $0.693 \pm 0.028$ ;  $P = .053$ ).

#### 3.2 | Identification of DMPs

The filtered probes of the HM450K were used to identify DMPs via the FDR across the adenoma subgroups (Table 3). Neither the comparison of primary adenomas without recurrence vs primary adenomas with recurrence, nor the comparison of primary adenomas vs the corresponding recurrent adenomas (matched pairs), nor the comparison of primary adenomas without recurrence vs adenomas associated with recurrence (primary and recurrent) showed any significant DMPs (Table 3). However, comparison of recurrent adenomas vs primary adenomas with recurrence revealed 2415 (0.7%) DMPs of 329 573 assessed CpG sites (comparison A,  $\text{FDR} \leq 0.05$ ; Table 3). Interestingly, 99% of DMPs were hypermethylated in recurrent adenomas compared with the primary adenomas with recurrence (Figure S2A). Moreover, comparison of recurrent adenomas vs all primary adenomas exhibited 575 (0.2%) DMPs (comparison B,  $\text{FDR} \leq 0.05$ ; Table 3). The DMPs present in the recurrent adenomas were nearly completely (99%) hypermethylated (Figure S2B). As expected, colorectal adenomas displayed 9266 (2.8%) DMPs in comparison with normal mucosa samples (comparison C,  $\text{FDR} \leq 0.05$ ; Table 3). More than two-thirds of the DMPs (73%) were hypermethylated in colorectal adenomas compared with normal mucosa (Figure S2C).

The total of 12 256 DMPs ( $\text{FDR} \leq 0.05$ ) was tested for overlaps across the comparisons A-C, which unveiled 17 CpG sites in 14 genes (Figure 3A): *MPPED1*, *AACS*, *FARP1*, *SMARCA4*, *TTC25*, *SIDT2*, *NUMBL*, *ABCC12*, *TBC1D12*, *ZNF655*, *ABCC5*, *MUSTN1*, *AMZ1*, and *RASL11A*.

To identify the most diverging DMPs and biologically relevant CpG positions, we filtered the DMPs once again by excluding probes with a fold change <10%. This resulted in 5094 DMPs composed of comparison A (674; 98% hypermethylated), comparison B (241; 99% hypermethylated), and comparison C (4,179; 46% hypermethylated) (Figure 3B). The CpG probes were separated into 2824 (55.4%) hypermethylated and 2270 (44.6%) hypomethylated DMPs. The stringent filtering via  $\Delta\beta \geq |0.1|$  provided 35 *top gene*-DMPs for comparison A (Table S2), 7 *top gene*-DMPs for comparison B (Table S3), and 347 *top gene*-DMPs for comparison C (Table S4). The *top gene*-DMPs of comparison A were further used to distinguish the matched pair cases of primary adenomas and the corresponding recurrent adenomas (Figure S3). The *top gene*-table of comparison A uniquely listed 30 of 35 (85.7%) DMP-associated genes (Table S2). However, two DMPs were in the open sea of *TMEM85* (transmembrane protein 85), whereas three DMPs were found in a CGI of the gene body of *GREM2*. The gene codes for Gremlin2, a member of the DAN family (differential screening-selected gene in neuroblastoma), of BMP (bone morphogenetic protein) antagonists and TGF- $\beta$  modulators. The relevance of the TGF- $\beta$ /BMP axis was further underpinned by *BMP3* being listed among the *top genes*. Therefore, we selected the *GREM2*-associated CpGs cg01809217 ( $\Delta\beta = 0.103$ ;  $q = 0.026$ ) and cg02577267 ( $\Delta\beta = 0.101$ ;  $q = 0.048$ ) as validation targets.

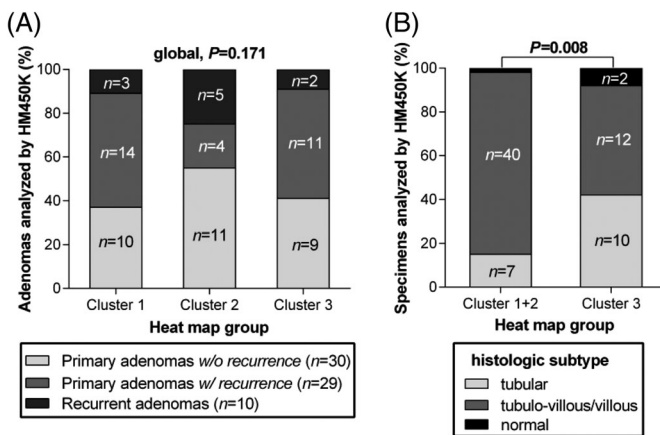


**FIGURE 1** DNA methylation in colorectal adenomas. A, The principal component analysis (PCA) of the sample set analyzed by HM450K (mean  $\beta$  values) plots the variance of the specimens in the orthogonal coordinate system. An association of sample groups based on the status of recurrence was not discovered. B, Unsupervised hierarchical clustering of the sample set performed using the  $\beta$  values and shows the Ward's distance across the samples. Cluster analysis exhibited three cluster groups of which cluster 3 diverged stronger compared with both others [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

### 3.3 | Genome-wide distribution of DMPs

The distribution of DNA methylation across the genome was investigated by classifying the genomic localization of the DMPs concerning the nearest CGI and the linear direction of either up- or downstream (ie, north or south), respectively. The total count of 5094 DMPs (Figure 3B) was filtered for probes, which were shared by two or more comparisons leaving 4917 unique CpG sites. The overview exhibited an unequal distribution of DMPs with an enrichment of hypermethylated DMPs in CGIs (33.7%; 1655/4917), whereas DMPs in open sea regions were predominately hypomethylated (24.1%; 1186/4917) (Figure 4A). Next, the localization of DMPs was analyzed by ratios of hypermethylation vs hypomethylation, which demonstrated an excess of hypomethylation in open sea and shelf regions (83% and 79%, respectively), while, in contrast, DMPs in CGIs were hypermethylated in 86% (Figure 4B). The genomic distribution of DMPs classified by individual comparisons showed that differences

among the DNA methylation occurred predominately in CGIs (comparison of A and B), whereas comparison C exhibited a balanced distribution of DMPs across the genome (Figure 4C). An analysis of the functional genomic regions unveiled that DMPs localized in the proximal promoter (ie, upstream of TSS, 5'UTR, and first exon) were most frequently hypermethylated (Figure 4D). DMPs in the gene body were most abundant and frequently hypomethylated (Figure 4D) although the average numbers of probes per functional region on the HM450K varied for the proximal promoter (15.2), gene body (9.9), and 3'UTR (1.5), respectively. The unequal distribution of hyper- and hypomethylated probes was further confirmed by plotting the ratio of hypermethylated vs hypomethylated DMPs, which demonstrated frequent hypermethylation within the proximal promoter and an excess of hypomethylation in gene body and 3'UTR (Figure 4E). The distribution of functional gene regions per comparison confirmed an apparent excess of DMPs localized in proximal promoter regions (Figure 4F). The distribution of DMPs across the autosomal chromosomes demonstrated the highest frequency for chromosome 1 (Figure 4G), as this is the largest chromosome with a high gene density. Notably, the ratio of hyper- and hypomethylated DMPs was unbalanced for chromosomes 8, 14, and 19, respectively, while the ratios were equalized for other chromosomes.



**FIGURE 2** A, Cluster groups displayed a weak linkage to an unequal distribution associated with recurrence (Freeman-Halton test). B, Dichotomized clustering of the specimens was significantly associated with the histologic subtype (Freeman-Halton test). Cluster 1 + 2 was dominated by colorectal adenomas with tubulovillous/villous histology

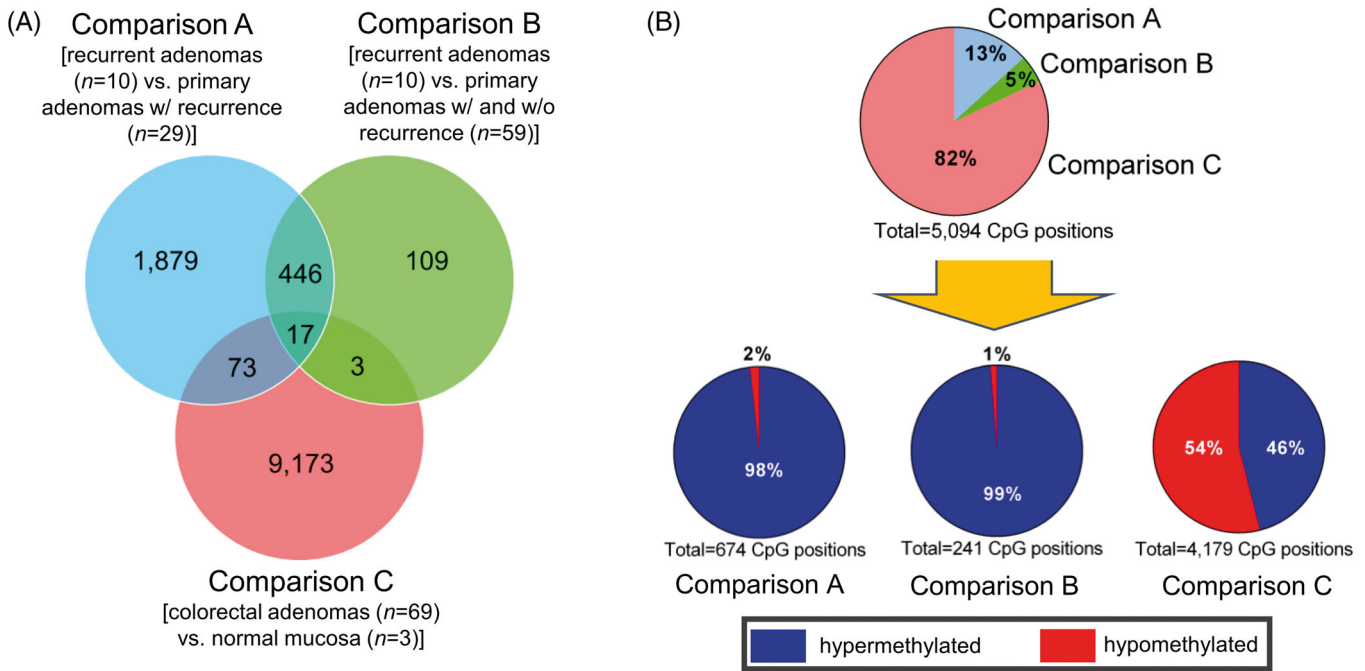
### 3.4 | Functional pathway enrichment of DMPs

The gene annotations of the DMPs identified by individual comparisons (Figure 3B) were analyzed via GSEA for common pathways and biological functions. Comparison A revealed a highly significant ( $q \leq 0.001$ ) association comprising 340 genes of 25 pathways and gene ontologies (Figure 5A). The most common genes were involved in the immune system ( $n = 44$ ;  $q = 2.9 \times 10^{-9}$ ), the adaptive immune system ( $n = 32$ ;  $4.8 \times 10^{-9}$ ) and inflammation-mediated chemokine and cytokine pathways, for example, IL2 pathway ( $n = 9$ ;  $q = 2.1 \times 10^{-5}$ ) (Figure 5A). Adenoma recurrence might be further associated with cellular survival and growth as indicated by involvement of the mTOR pathway ( $n = 9$ ;  $q = 6.9 \times 10^{-5}$ ) and cell cycle pathways ( $n = 19$ ;  $q = 7.3 \times 10^{-4}$ ). Comparison B provided two pathways

**TABLE 3** Significant DMPs in adenoma groups

#	Specimen groups	Sample number	Description	Significant DMPs (n)	
				FDR $\leq 0.01$	FDR $\leq 0.05$
1	Primary adenomas without recurrence vs primary adenomas with recurrence	30 vs 29		0	0
2	Matched pairs: Primary adenomas with recurrence vs recurrent adenomas	10 vs 10		0	0
3	Primary adenomas without recurrence vs primary adenomas with recurrence and recurrent adenomas	30 vs 39		0	0
4	Recurrent adenomas vs primary adenomas with recurrence	10 vs 29	Comparison A	22	2415
5	Recurrent adenomas vs primary adenomas with and without recurrence	10 vs 59	Comparison B	54	575
6	Colorectal adenomas vs normal mucosa	69 vs 3	Comparison C	5457	9266





**FIGURE 3** Distribution of DMPs. A, The Venn diagram displays the overlapping DMPs across the comparisons A, B, and C, respectively. B, The upper pie chart shows the distribution of the DMPs among the three comparisons. The portion of hyper- and hypomethylated DMPs per comparison is detailed in the pie charts [Color figure can be viewed at wileyonlinelibrary.com]

( $q \leq 0.001$ ) based on 31 genes: the immune system ( $n = 21$ ;  $q = 2.1 \times 10^{-5}$ ) and cytokine signaling in the immune system ( $n = 10$ ;  $q = 9.3 \times 10^{-4}$ ) (Figure 5B). Comparison C provided 34 highly significant pathways associated with 1022 genes (Figure 5C). Differential genes between adenomas vs normal mucosa encoded, notably, for proteins of the extracellular matrix (ECM) ( $n = 84$ ;  $q = 5.8 \times 10^{-13}$ ) and core ECM proteins ( $n = 32$ ;  $q = 2.4 \times 10^{-8}$ ). However, targets of tumorigenesis were identified such as pathways in cancer ( $n = 39$ ;  $q = 3.4 \times 10^{-10}$ ), MAPK signaling ( $n = 34$ ;  $q = 1.0 \times 10^{-9}$ ), and p53 signaling ( $n = 13$ ;  $q = 1.3 \times 10^{-5}$ ) (Figure 5C). As before, genes involved in the immune system and adaptive immune system were enriched.

### 3.5 | CpG methylation by pyrosequencing

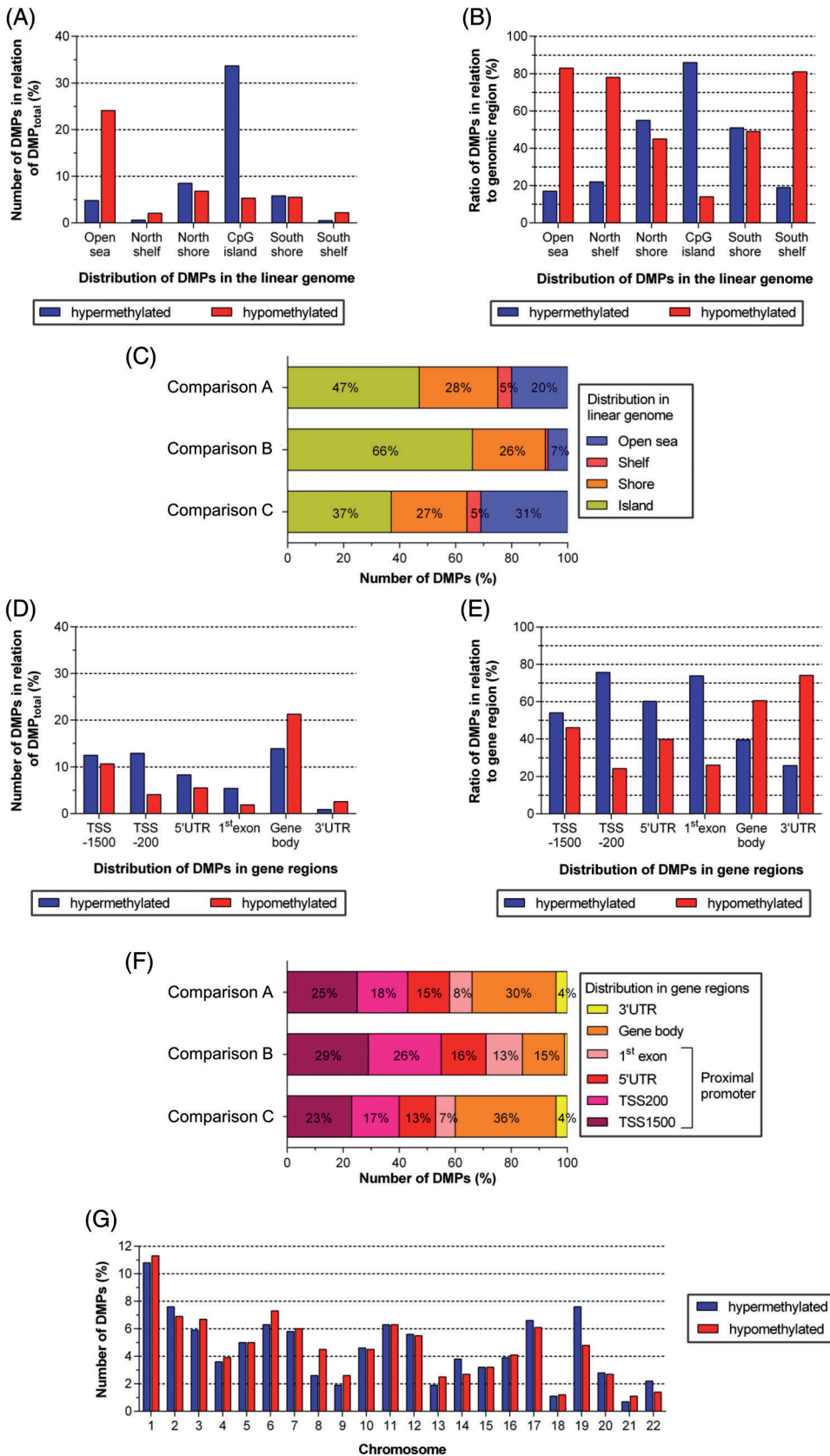
Locus-specific methylation alterations of eight CpG dinucleotides in the gene body of the validation candidate *GREM2* were evaluated by pyrosequencing (Figure S1). Bisulfite-converted DNAs of primary adenomas with ( $n = 15$ ) and without ( $n = 9$ ) recurrence, and recurrent adenomas ( $n = 4$ ) were analyzed for *GREM2* gene body methylation frequencies (Figure 6A). Detected *GREM2* methylation frequencies displayed a high sample variance as 14.6% (28/192) of CpG sites were hypomethylated ( $\leq 33\%$ ), 49.5% (95/192) of CpGs were medium-methylated ( $>33\%$ -66%), and 35.9% (69/192) were hypermethylated ( $>66\%$ ). The mean methylation frequency per CpG sites 1-4 was calculated and showed similar ratios for primary adenomas with and without recurrence while recurrent adenomas exhibited slightly higher values (Figure 6B). On the contrary, methylation frequencies of CpG sites 12-15 were rather heterogeneous with recurrent adenomas,

displaying lower frequencies than primary adenomas with and without recurrence (Figure 6C). Although CpG 2 demonstrated the highest methylation ratio (mean,  $70\% \pm 18\%$ ), CpG 13 showed the lowest ratio (mean,  $38\% \pm 15\%$ ). The HM450K data of the average  $\beta$ -methylation of CpG 1 of *GREM2* (cg01809217) was validated by pyrosequencing for recurrent adenomas [( $\bar{\beta}_{\text{HM450K}} = 0.716$ ;  $n = 10$ ) vs ( $\bar{\beta}_{\text{Pyros}} = 0.753$ ;  $n = 4$ )] and primary adenomas with recurrence [( $\bar{\beta}_{\text{HM450K}} = 0.613$ ;  $n = 29$ ) vs ( $\bar{\beta}_{\text{Pyros}} = 0.687$ ;  $n = 14$ )].

The median methylation frequency by individual CpG sites 1-4 sorted by adenoma groups did not show a difference ( $P = .590$ ) (Figure 6D). However, recurrent adenomas tended to be slightly more hypermethylated compared with primary adenomas with ( $P = .312$ ) and without ( $P = .406$ ) recurrence. Comparing the median methylation levels of the individual CpG sites 12-15 sorted by groups demonstrated a highly significant difference among the adenoma groups ( $P = .002$ ) (Figure 6E). Moreover, recurrent adenomas were hypomethylated compared with primary adenomas with ( $P = .003$ ) and without ( $P = .004$ ) recurrence.

## 4 | DISCUSSION

Methylation of CGI is a mechanism for suppressing gene transcription in cancer including colorectal neoplasia.<sup>33</sup> Many tumor suppressor genes are hypermethylated at CpG-dense promoters, which are, therefore, silenced in colorectal tumors.<sup>16,34</sup> It is also now accepted that methylation assays can be applied for CRC screening. The FDA approved a *SEPT9* gene methylation assay in 2016 as an additional CRC screening option to be considered for patients who have a

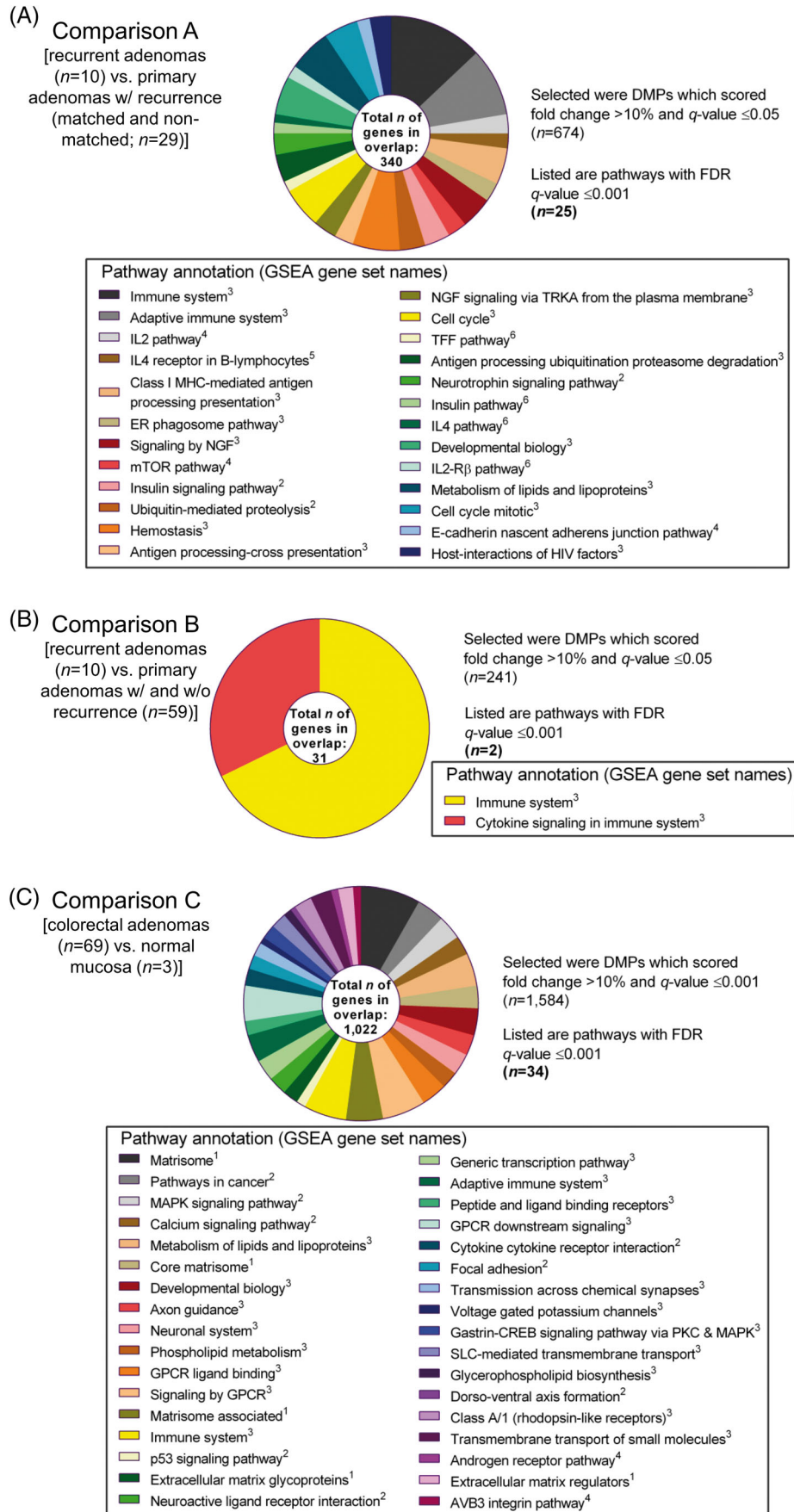


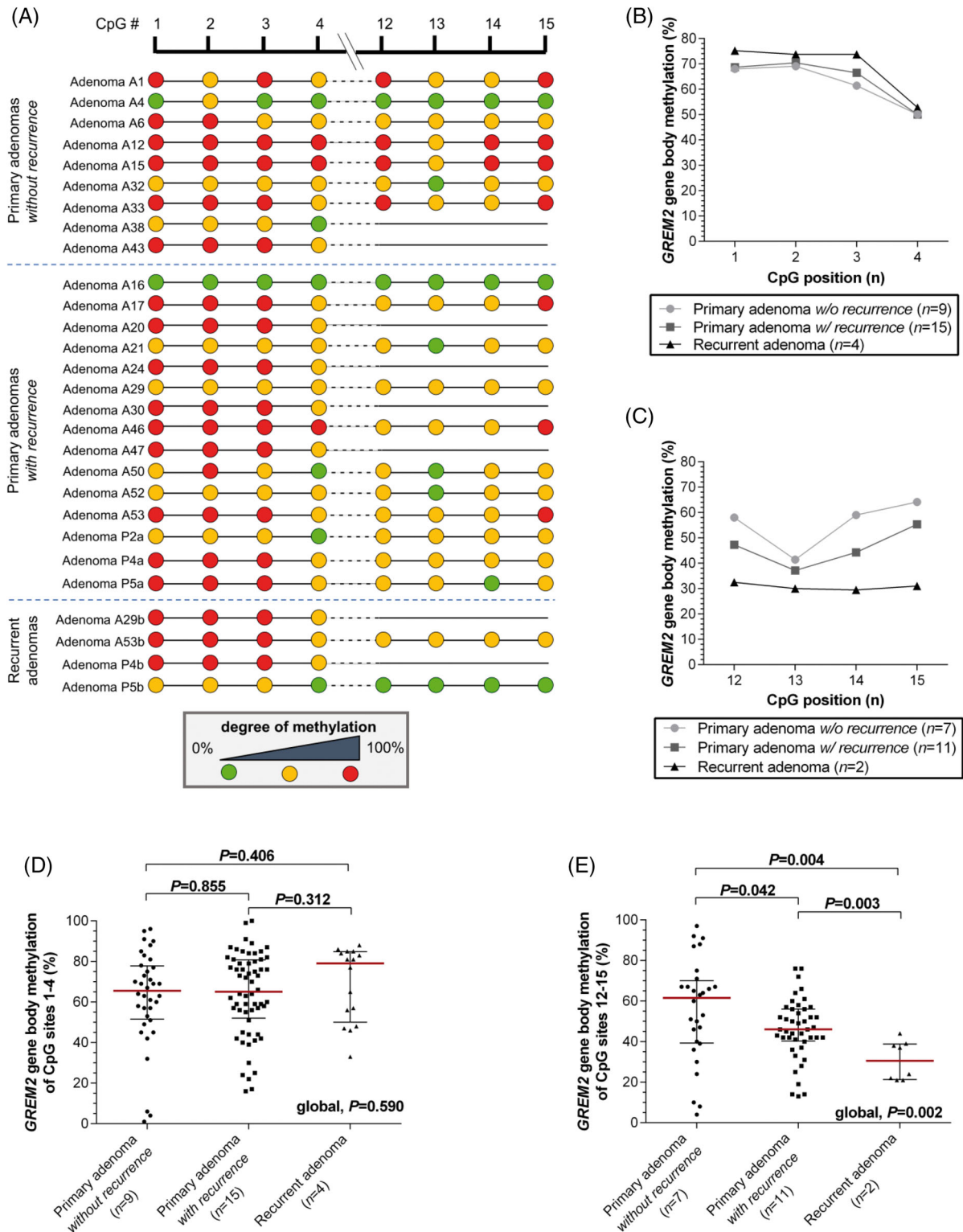
**FIGURE 4** Genomic distribution of DMPs. A, Distribution of DMPs based on the CpG content across the linear genome. Shown percentages are computed with the total number of DMPs ( $n_{\text{total}} = 4917$ ). B, Depicted is the ratio of hyper- vs hypomethylated DMPs in multiple genomic regions defined by CpG content. C, Distribution of DMPs based on the CpG content and sorted by individual comparisons. D, Distribution of DMPs in gene-associated regions defined as 200-1500 bp upstream from transcription start site (TSS1500), up to 200 bp upstream from transcription start site (TSS200), 5' untranslated region (5' UTR), first exon, gene body, and 3' untranslated region (3' UTR). E, Ratio of hyper- vs hypomethylated DMPs in different gene-associated regions. F, Distribution of DMPs based on gene-associated regions and sorted by individual comparisons. G, Chromosomal distribution of DMPs categorized as hyper- or hypomethylated, respectively [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

history of an intermittent completion of colonoscopy and fecal occult blood tests.<sup>35</sup> Thus, this study sought to discover aberrantly methylated CpG dinucleotides in the methylome of primary colorectal

adenomas for prediction of adenoma recurrence. We surmised that a genome-wide CpG analysis using the HM450K array should unveil a methylation pattern associated with adenoma recurrence. However,

**FIGURE 5** Distribution of DMPs according to their gene ontology. The circle plots show the frequency of gene-associated DMPs and their pathway/gene ontology for (A) comparison A, (B) comparison B, and (C) comparison C, respectively. The pathways were ordered by significance and not by the frequency of genes involved. Pathway annotations relate to the database entry: 1, NABA; 2, KEGG; 3, REACTOME; 4, PID; 5, SIG; 6, BioCarta [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]





**FIGURE 6** GREM2 CpG methylation by pyrosequencing. A, Detected methylation values are depicted as lollipops resembling the color-coded degree of methylation from low to high (ie, hypo- to hypermethylation) in trichotomized order (*green* < *orange* < *red*). Not determined (N.D.) methylation levels of REV-assay CpG sites are shown as straight lines. B, Quantitative GREM2 gene body methylation analysis of four CpG dinucleotides (FOR-assay) in primary adenomas with and without recurrence, and recurrent adenomas. Data points represent the mean methylation frequency per CpG and subgroup. C, Quantitative GREM2 gene body methylation analysis of four CpG dinucleotides (REV-assay) in primary adenomas with and without recurrence and recurrent adenomas. D, Plotting the individual methylation frequencies of CpG sites 1-4 did not show differences (median with IQR; Mann-Whitney *U* test; global, Kruskal-Wallis test). E, The individual methylation frequencies of CpG sites 12-15 showed a significant difference across primary adenomas with and without recurrence (median with IQR; Mann-Whitney *U* test; global, Kruskal-Wallis test) [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

PCA and unsupervised hierarchical clustering of the mean  $\beta$ -values based on 329 573 CpG sites did not separate the specimens by recurrence groups. Neither were primary adenomas without recurrence discernable from primary adenomas with recurrence nor were there any significant changes in primary adenomas vs the corresponding recurrent adenomas (matched pairs). Instead, the individual methylation status is so variable that such a complicated process like recurrence prediction could not be attributed to a specific methylation pattern. Even within a single group, for example, primary adenomas without recurrence, we observed a high degree of variability among DMPs. This seems contradictory to the findings of Rengucci and colleagues who identified the promoters of *MLH1*, *ATM*, and *FHIT* to be hypermethylated in adenomas with high recurrence risk.<sup>36</sup> However, this risk assessment was based on assumptions and not on real clinical follow up data, which could explain the discrepancies to our findings.

Our analysis of recurrent adenomas showed DMPs compared with primary adenomas with and without recurrence. Even though these differences cannot predict recurrence, they suggest that recurrent adenomas develop through significant changes in the methylation pattern. The DMPs of the comparison between recurrent adenomas vs primary adenomas with recurrence were almost entirely (98%) hypermethylated. The pathway analysis and GSEA discovered that altered DMPs of recurrent adenomas were predominantly associated with the immune system and cytokine signaling in the immune system. This observation is consistent with many reports claiming that endoscopic resection initiates inflammatory processes, which are a hallmark of tumor development and recurrence.<sup>37-41</sup> However, it remains unclear whether the increased methylation of genes involved in immune system pathways is directly correlated with activated inflammatory processes as protein expression patterns were not investigated.

The comparison of colorectal adenomas vs normal mucosa revealed 4179 DMPs. Several of the observed targets of differential methylation, for example, *AXIN2*, *DKK3*, *GRIA4*, *MLH1*, *SEPT9*, *SFRP1*, *SLC8A1*, and *SYN3*, were also reported in recent studies.<sup>21,42,43</sup> Additionally, our study confirmed earlier publications that DMPs are not randomly distributed within the genome<sup>13,22,34</sup>: DMPs in CGIs were frequently hypermethylated while open sea and shelf regions were hypomethylated; proximal promoter regions were more frequently affected by hypermethylation compared with gene body and 3'UTR, which were found to be targeted by hypomethylation. This fact is linked to the current understanding of alterations in tumors suggesting an association of gene silencing with promoter hypermethylation, whereas gene transcription is associated with gene body hypomethylation.<sup>16</sup> However, DNA methylation of gene bodies might also be involved in differential promoter usage, alternative splicing, or the prevention of transcription initiation.<sup>17</sup> Thus, hypermethylation within the gene body of *GREM2* in recurrent adenomas vs primary adenomas with recurrence could indicate a gene silencing leading to deregulation of the BMP axis and WNT-pathway disruption.<sup>44</sup> We discovered a discrepancy of the mean  $\beta$  values of recurrent adenomas in probe cg02577267 when comparing the results obtained via pyrosequencing and HM450K. This is because of the less representative number of samples analyzed by pyrosequencing.

Furthermore, we could detect that the hierarchical clustering (methylation  $\beta$  values) of specimens showed a significant difference in the distribution of adenomas with tubular histology compared with adenomas with tubulovillous/villous histology. It is known that the frequency of methylated genes is higher in histologically advanced adenomas compared with hyperplastic polyps and low-grade adenomas.<sup>45</sup> However, the differentiation of histological subgroups based on their methylation pattern is novel. A Dutch register study demonstrated a considerable interlaboratory variation in evaluating colorectal adenomas.<sup>46</sup> Specialized gastrointestinal histopathologists showed only moderate ( $\kappa = 0.49$ ) concordance in discriminating the different polyp forms of the group of serrated adenomas.<sup>47</sup> Thus, an expert panel proved that morphological criteria alone are not sufficient but can be improved by including molecular data like *BRAF* or *KRAS* mutation status.<sup>48</sup> Additionally, a meta-analysis indicated that hypermethylation of *p16* might be an unfavorable biomarker for CRC patients.<sup>49</sup> Recent studies about brain tumor classification showed that the analysis of methylation patterns could be highly relevant for precise diagnostics and treatment suggestions.<sup>50,51</sup> The same holds true for sarcoma diagnostics and allows a more precise classification of undifferentiated and small blue round cell tumors with array-based DNA-methylation profiling.<sup>52</sup>

The here presented finding that adenoma morphology differs by DNA methylation might provide the opportunity for enhancing the accuracy of the up to now mainly morphology-based classification of colonic adenomas.

## ACKNOWLEDGMENTS

The authors would like to thank Romina Laegel and Alexandra Eichhorn for technical assistance in retrieving the archived tissue material and Dr. Christian Sauer for aid in performing the pyrosequencing assay. DF received a travel grant from the Boehringer Ingelheim Fonds (BIF), Germany. DH was supported by an intramural research scholarship from the University of Heidelberg, Medical Faculty Mannheim (Translational Physician Scientist Program).

## DATA ACCESSIBILITY

The array data that support the findings of this study are openly available in NCBI Gene Expression Omnibus (GEO) at <https://www.ncbi.nlm.nih.gov/geo/>, accession number GSE129364.

## ORCID

Thomas Ried  <https://orcid.org/0000-0002-0767-6009>

Timo Gaiser  <https://orcid.org/0000-0002-6022-073X>

## REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.

2. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med*. 2012;366(8):687-696.
3. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med*. 2014;370(14):1298-1306.
4. Hassan C, Quintero E, Dumonceau J-M, et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) guideline. *Endoscopy*. 2013;45(10):842-864.
5. Pox C, Aretz S, Bischoff SC, et al. *S3-Leitlinie Kolorektales Karzinom Version 1.0 - Juni 2013 AWMF-Registernummer: 021/007OL*. Vol 51. DKG und DKH: Leitlinienprogramm Onkologie (AWMF; 2013.
6. Winawer SJ, Zauber AG, Fletcher RH, et al. Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US multi-society task force on colorectal cancer and the American Cancer Society. *CA Cancer J Clin*. 2006;56(3):143-159. quiz 184-5.
7. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487(7407):330-337.
8. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. *N Engl J Med*. 1988;319(9):525-532.
9. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*. 1993;363(6429):558-561.
10. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JPJ. CpG Island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A*. 1999;96(15):8681-8686.
11. Patai ÁV, Valcz G, Hollósi P, et al. Comprehensive DNA methylation analysis reveals a common ten-gene methylation signature in colorectal adenomas and carcinomas. *PLoS One*. 2015;10(8):e0133836.
12. Galamb O, Kalmár A, Péterfia B, et al. Aberrant DNA methylation of WNT pathway genes in the development and progression of CIMP-negative colorectal cancer. *Epigenetics*. 2016;11(8):588-602.
13. Luo Y, Wong C, Kaz AM, et al. Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. *Gastroenterology*. 2014;147(2):418-429.
14. Jones PA, Baylin SB. The epigenomics of cancer. *Cell*. 2007;128(4):683-692.
15. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science (80- )*. 2003;300(5618):455.
16. Esteller M. Epigenetics in cancer. *N Engl J Med*. 2008;358(11):1148-1159.
17. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat rev Genet*. 2012;13(7):484-492.
18. Yang X, Han H, De Carvalho DD, et al. Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell*. 2014;26(4):577-590.
19. Bibikova M, Barnes B, Tsan C, et al. High density DNA methylation array with single CpG site resolution. *Genomics*. 2011;98(4):288-295.
20. Dumenil TD, Wockner LF, Bettington M, et al. Genome-wide DNA methylation analysis of formalin-fixed paraffin embedded colorectal cancer tissue. *Genes Chromosomes Cancer*. 2014;53(7):537-548.
21. Fadda A, Gentilini D, Moi L, et al. Colorectal cancer early methylation alterations affect the crosstalk between cell and surrounding environment, tracing a biomarker signature specific for this tumor. *Int J Cancer*. 2018;143(4):907-920.
22. Koestler DC, Li J, Baron JA, et al. Distinct patterns of DNA methylation in conventional adenomas involving the right and left colon. *Mod Pathol*. 2014;27(1):145-155.
23. Bosman F, Carneiro F, Hruban R, Theise ND. *WHO Classification of Tumors of the Digestive System*. 4th edn. Lyon: IARC Press; 2010.
24. Maksimovic J, Gordon L, Oshlack A. SWAN: subset-quantile within Array normalization for Illumina Infinium HumanMethylation450 BeadChips. *Genome Biol*. 2012;13(6):R44.
25. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363-1369.
26. Chen Y, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium Human-Methylation450 microarray. *Epigenetics*. 2013;8(2):203-209.
27. Rakyan VK, Down TA, Maslau S, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res*. 2010;20(4):434-439.
28. Wang D, Yan L, Hu Q, et al. IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data. *Bioinformatics*. 2012;28(5):729-730.
29. Martin-Subero JI, Ammerpohl O, Bibikova M, et al. A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms. *PLoS One*. 2009;4(9):e6986.
30. Schneider E, Dittrich M, Böck J, et al. CpG sites with continuously increasing or decreasing methylation from early to late human fetal brain development. *Gene*. 2016;592(1):110-118.
31. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
32. Ashktorab H, Daremipouran M, Goel A, et al. DNA methylome profiling identifies novel methylated genes in African American patients with colorectal neoplasia. *Epigenetics*. 2014;9(4):503-512.
33. Park S-J, Rashid A, Lee J-H, Kim SG, Hamilton SR, Wu TT. Frequent CpG Island methylation in serrated adenomas of the colorectum. *Am J Pathol*. 2003;162(3):815-822.
34. Andrew AS, Baron JA, Butterly LF, et al. Hyper-methylated loci persisting from sessile serrated polyps to serrated cancers. *Int J Mol Sci*. 2017;18(535):1-12.
35. Song L, Li Y. Progress on the clinical application of the SEPT9 gene methylation assay in the past 5 years. *Biomark Med*. 2017;11(6):415-418.
36. Rengucci C, De Maio G, Casadei Gardini A, et al. Promoter methylation of tumor suppressor genes in pre-neoplastic lesions; potential marker of disease recurrence. *J Exp Clin Cancer Res*. 2014;33:65.
37. Evans C, Galustian C, Kumar D, et al. Impact of surgery on immunologic function: comparison between minimally invasive techniques and conventional laparotomy for surgical resection of colorectal tumors. *Am J Surg*. 2009;197(2):238-245.
38. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature*. 2008;454(7203):436-444.
39. McLean MH, Murray GI, Stewart KN, et al. The inflammatory micro-environment in colorectal neoplasia. *PLoS One*. 2011;6(1):e15366.
40. Terzić J, Grivennikov S, Karin E, Karin M. Inflammation and colon cancer. *Gastroenterology*. 2010;138(6):2101-2114.e5.
41. Tohme S, Simmons RL, Tsung A. Surgery for cancer: a trigger for metastases. *Cancer Res*. 2017;77(7):1548-1552.
42. Gyparaki M-T, Basdra EK, Papavassiliou AG. DNA methylation biomarkers as diagnostic and prognostic tools in colorectal cancer. *J Mol Med*. 2013;91(11):1249-1256.
43. Sugai T, Yoshida M, Eizuka M, et al. Analysis of the DNA methylation level of cancer-related genes in colorectal cancer and the surrounding normal mucosa. *Clin Epigenetics*. 2017;9:55.
44. Kosinski C, Li VSW, Chan ASY, et al. Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. *Proc Natl Acad Sci U S A*. 2007;104(39):15418-15423.
45. Petko Z, Ghiassi M, Shuber A, et al. Aberrantly methylated CDKN2A, MGMT, and MLH1 in colon polyps and in fecal DNA from patients with colorectal polyps. *Clin Cancer Res*. 2005;11(3):1203-1209.
46. Kuijpers CCHJ, Sluiter CE, von der Thüsen JH, et al. Interlaboratory variability in the grading of dysplasia in a nationwide cohort of colorectal adenomas. *Histopathology*. 2016;69(2):187-197.

47. Wong NACS, Hunt LP, Novelli MR, Shepherd NA, Warren BF. Observer agreement in the diagnosis of serrated polyps of the large bowel. *Histopathology*. 2009;55(1):63-66.
48. Rau TT, Agaimy A, Gehoff A, et al. Defined morphological criteria allow reliable diagnosis of colorectal serrated polyps and predict polyp genetics. *Virchows Arch*. 2014;464(6):663-672.
49. Xing X, Cai W, Shi H, et al. The prognostic value of CDKN2A hypermethylation in colorectal cancer: a meta-analysis. *Br J Cancer*. 2013;108(12):2542-2548.
50. Capper D, Stichel D, Sahm F, et al. Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience. *Acta Neuropathol*. 2018;136(2): 181-210.
51. Sahm F, Schrimpf D, Stichel D, et al. DNA methylation-based classification and grading system for meningioma: a multicentre, retrospective analysis. *Lancet Oncol*. 2017;18(5):682-694.
52. Koelsche C, Hartmann W, Schrimpf D, et al. Array-based DNA-methylation profiling in sarcomas with small blue round cell histology

provides valuable diagnostic information. *Mod Pathol*. 2018;31(8): 1246-1256.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Fiedler D, Hirsch D, El Hajj N, et al. Genome-wide DNA methylation analysis of colorectal adenomas with and without recurrence reveals an association between cytosine-phosphate-guanine methylation and histological subtypes. *Genes Chromosomes Cancer*. 2019;58: 783–797. <https://doi.org/10.1002/gcc.22787>