

RESEARCH

Open Access



Genome-wide methylation profiling identified novel differentially hypermethylated biomarker *MPPED2* in colorectal cancer

Simeng Gu¹, Shujuan Lin¹, Ding Ye^{1,2}, Sangni Qian¹, Danjie Jiang¹, Xiacong Zhang¹, Qilong Li³, Jinhua Yang³, Xiaojiang Ying⁴, Zhenjun Li⁴, Mengling Tang¹, Jianbing Wang¹, Mingjuan Jin^{1*} and Kun Chen^{1,5*}

Abstract

Background: Epigenetic alternation is a common contributing factor to neoplastic transformation. Although previous studies have reported a cluster of aberrant promoter methylation changes associated with silencing of tumor suppressor genes, little is known concerning their sequential DNA methylation changes during the carcinogenic process. The aim of the present study was to address a genome-wide search for identifying potentially important methylated changes and investigate the onset and pattern of methylation changes during the progression of colorectal neoplasia.

Methods: A three-phase design was employed in this study. In the screening phase, DNA methylation profile of 12 pairs of colorectal cancer (CRC) and adjacent normal tissues was analyzed by using the Illumina MethylationEPIC BeadChip. Significant CpG sites were selected based on a cross-validation analysis from The Cancer Genome Atlas (TCGA) database. Methylation levels of candidate CpGs were assessed using pyrosequencing in the training dataset (tumor lesions and adjacent normal tissues from 46 CRCs) and the validation dataset (tumor lesions and paired normal tissues from 13 hyperplastic polyps, 129 adenomas, and 256 CRCs). A linear mixed-effects model was used to examine the incremental changes of DNA methylation during the progression of colorectal neoplasia.

Results: The comparisons between normal and tumor samples in the screening phase revealed an extensive CRC-specific methylomic pattern with 174,006 (21%) methylated CpG sites, of which 22,232 (13%) were hypermethylated and 151,774 (87%) were hypomethylated. Hypermethylation mostly occurred in CpG islands with an overlap of gene promoters, while hypomethylation tended to be mapped far away from functional regions. Further cross validation analysis from TCGA dataset confirmed 265 hypermethylated promoters coupling with downregulated gene expression. Among which, hypermethylated changes in *MPPED2* promoter was successfully replicated in both training and validation phase. Significant hypermethylation appeared since precursor lesions with an extensive modification in CRCs. The linear mixed-effects modeling analysis found that a cumulative pattern of *MPPED2* methylation changes from normal mucosa to hyperplastic polyp to adenoma, and to carcinoma ($P < 0.001$).

Conclusions: Our findings indicate that epigenetic alterations of *MPPED2* promoter region appear sequentially during the colorectal neoplastic progression. It might be able to serve as a promising biomarker for early diagnosis and stage surveillance of colorectal tumorigenesis.

Keywords: Epigenetics, DNA methylation, Colorectal cancer, EPIC, MPPED2

* Correspondence: [jinmj@zju.edu.cn](mailto:jinj@zju.edu.cn); ck@zju.edu.cn

¹Department of Epidemiology and Biostatistics, Zhejiang University School of Public Health, 866 Yuhangtang Road, Hangzhou 310058, China
Full list of author information is available at the end of the article



Introduction

Colorectal cancer (CRC) is the third most common diagnosed cancer and the fourth leading cause of cancer death worldwide [1]. It results from a series of changes at both genetic and epigenetic levels. Despite a large number of genetic alterations have so far been described, little has found their way in clinical practice. An in-depth understanding of epigenetic regulatory network might do work in screening, diagnosis, and treatment of CRC.

At epigenetic level, DNA methylation changes is one of the hallmark events in carcinogenesis, characterized by global hypomethylation and paradoxical gene-specific hypermethylation [2, 3]. Hypomethylation is primarily involved in chromosomal instability and global loss of imprinting, while gene-specific hypermethylation anchors to promoters, it causes transcriptional silencing of tumor suppressor genes and consequently sets the stage for neoplastic transformation. Deregulated DNA methylation is well known to be associated with CRC [4, 5]. The first evidence of DNA methylation contributing to CRC was presented by Goelz et al. [6], in which a global loss in DNA methylation was identified. Then, extensive efforts have been made in identifying aberrant hypermethylation in promoters of CRC-related suppressor genes, such as *APC*, *SFRP2*, *SEPT*, and *CDH1* [7–9]. Remarkably, a near universal phenomenon that gene-specific hypermethylation occurred in both pre-neoplastic and neoplastic phase of colorectal cancer was observed [10, 11]. This might shed light on the process of tumorigenesis. However, few studies have defined precisely the hierarchy of methylation events during the transformation from normal epithelial cells to malignant cells in colorectal carcinogenesis.

It is becoming increasingly apparent that the occurrence of molecular alterations could be found not only in tumor tissue but also in histological normal-appearing tissue adjacent to the tumor [12–14]. The presence of such molecular alterations in histological normal-appearing tissue is commonly known as field cancerization or field effect [15]. This has been thought to constitute the earliest clone in the carcinogenesis process. In colorectal carcinogenesis, a few number of gene-specific hypermethylation events have been reported in normal-appearing colonic mucosa from CRC patients (such as *APC*, *DKKI*, *MGMT*, *CDKN2A*, and *SFRP4*) [16–18]. In addition, previous studies have also demonstrated the effect of aberrant DNA methylation on suppressor genes such as *MINT1*, *MINT31*, *SLC5A8*, and *MGMT*, during adenoma-carcinoma sequence [19–21]. These findings suggest that the presence of field effect in methylation might be a useful intermediate biomarker in etiologic studies. A better understanding of when these epigenetic tags occur and how they take part in colorectal progression may represent a practical opportunity for colorectal cancer risk assessment.

Therefore, the aim of the present study was to investigate the altered DNA methylome for identification of methylation biomarkers in CRC. Additionally, we hypothesized that field effect due to DNA methylation might appear since the normal-appearing colonic mucosa adjacent to tumor tissue. A quantitative model was introduced to give a full scope of changes in epigenetic pattern during the carcinogenic process.

Methods

Study design and study population

This study has been processed through three phases. In the screening phase, a genome-wide methylation scan on cancerous and paired normal tissues from 12 CRC patients was performed, followed by a cross-validation analysis including transcriptome and DNA methylome data from the publicly available database. Then, in the training phase, 46 pairs of CRC tissue samples were tested for a given list of candidates to assess the reproducibility of Illumina MethylationEPIC BeadChip (EPIC) platform. Thirteen hyperplastic polyps, 129 adenomas, and 256 CRCs were collected into the validation phase, in which only potential genes with a high discriminative performance in the training phase were further testified. In order to evaluate the onset of colorectal neoplastic progression, we defined and evaluated methylation-changing patterns in the transition from normal mucosa related to low risk and high risk of CRC, to hyperplastic polyp, and to adenoma and carcinoma in the validation phase. The definitions for low- and high-risk normal mucosa were as follows: normal colon mucosa from individuals who had no history of CRC was considered as low-risk normal mucosa; normal colon mucosa from those who have primary solid CRCs, which might be at an increased risk of metachronous CRCs, was considered as high-risk normal mucosa. A flowchart is shown in Fig. 1.

The enrollment of CRC cases was based on Shaoxing People's Hospital between 2015 and 2017. Participants with colorectal polyps (hyperplastic polyp or adenoma) were drawn from a population-based cohort since 1989 in Jiashan County, Zhejiang Province, China. It has been described in detail previously [22, 23]. For each participant, a histologically confirmed colorectal lesion (hyperplastic polyp, adenoma, or carcinoma) and a distant normal mucosa sample were obtained. Individuals were excluded according to the following criteria: (1) familial adenomatous polyposis (FAP), (2) a history of CRC, and (3) preoperative anticancer treatment. Tumor stage was determined according to the American Joint Committee on Cancer (AJCC) TNM staging system of CRC. A complete list of participants included in each phase is shown in Table 1.

Every participant signed an informed consent prior to study participation. The study was approved by the

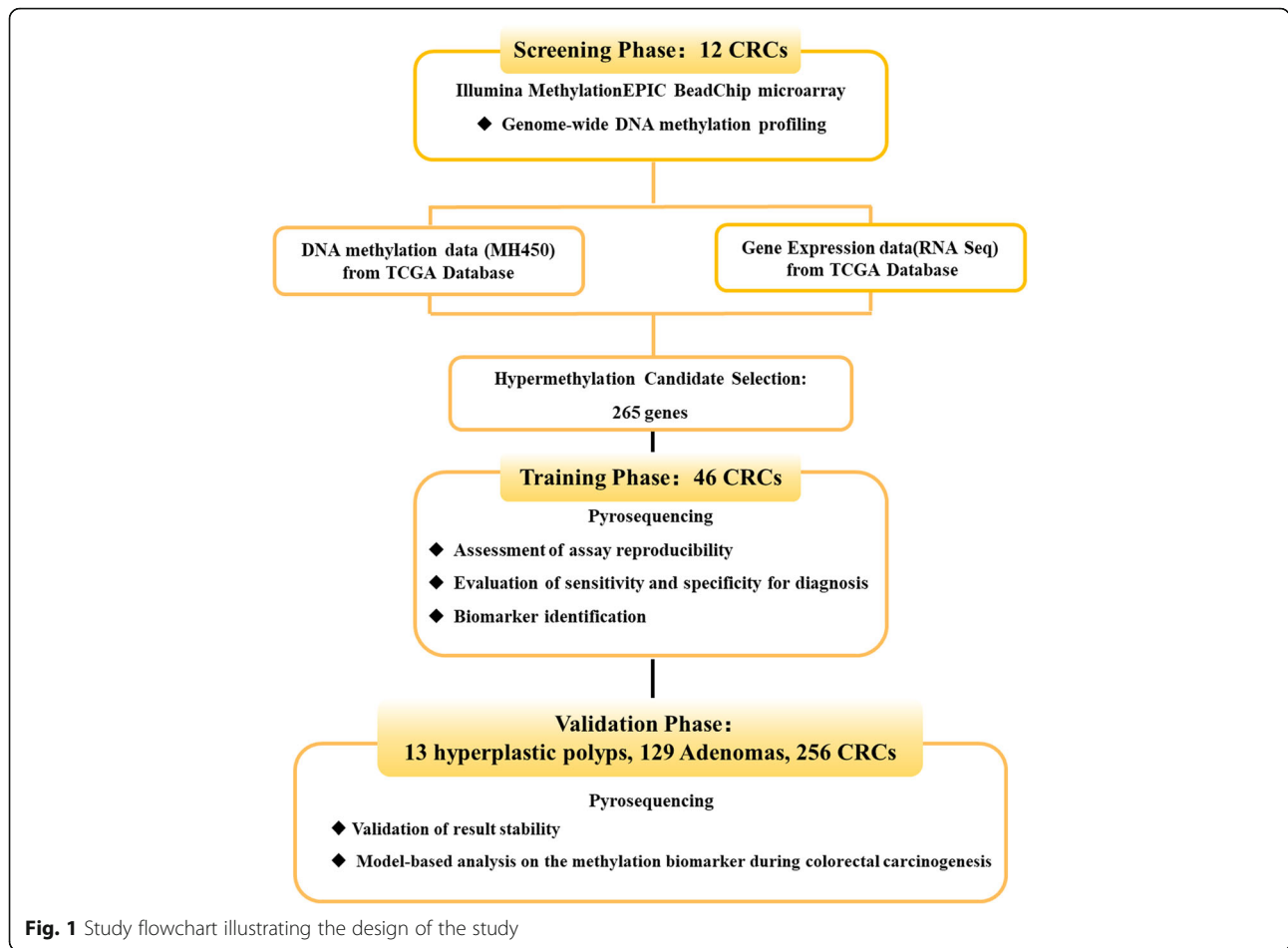


Table 1 Characteristics of the study population

Characteristics	Screening phase	Training phase	Validation phase		
	CRCs (n = 12)	CRCs (n = 46)	Hyperplastic polyps (n = 13)	Adenomas (n = 129)	CRCs (n = 256)
Age, mean ± SD	63.50 ± 4.15	64.11 ± 9.03	59.62 ± 5.99	62.14 ± 7.21	63.51 ± 9.55
Gender					
Male	7	23	6	79	159
Female	5	23	7	50	97
Location					
Colon	6	21	8	110	114
Rectal	6	25	5	19	142
Stage					
I	3	11			28
II	3	12			123
III	3	12			95
IV	3	11			10

SD standard deviation

Medical Ethics Committee of Zhejiang University School of Medicine.

DNA extraction and bisulfite conversion

DNA from fresh-frozen samples was extracted using DNA tissue Kit (Omega Bio-Tek, Norcross, GA, USA). DNA was used for bisulfite conversion, which was performed using EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA). All procedures were conducted according to the manufacturer’s instructions.

Genome-wide DNA methylation analysis

Genome-wide DNA methylation analysis was performed using a recently developed EPIC BeadChip (Illumina, San Diego, CA, USA), which covers more than 850,000 CpG sites for methylation study. Raw fluorescence intensities were loaded in BeadStudio software to field β values, which represent the methylation score of each CpG site. They range from 0 (non-methylated) to 1 (fully methylated). Prior to identification of differentially methylated probes, preprocessing steps, including data filtering, correction, and normalization, were implemented. The probe call rate at least 95% coverage per sample and

probes detection P value < 0.01 were required, otherwise they should be excluded. In addition, probes on X and Y chromosomes were also removed. Background correction, dye-bias, and beta-mixture quantile normalization (BMIQ) procedures were performed using Bioconductor packages [24, 25]. All 12 paired CRC samples included into the screening phase passed the quality control from EPIC platform with a probe call rate $> 99\%$, and 844,711 probes out of 853,307 were included in the following analysis.

Comparisons between cancerous and paired normal tissues were performed using the Illumina Methylation Analyzer (IMA) package. Differentially methylated CpG sites were identified when β difference > 0.2 and adjusted P value (Benjamini-Hochberg method, FDR) < 0.05 .

External validation of the candidate biomarker genes

We downloaded data from the Cancer Genome Atlas (TCGA) project (<https://cancergenome.nih.gov>) to validate our results. DNA methylation data consisting of 351 samples (38 normal, 313 tumor) was generated using Illumina HumanMethylation450 (HM450) BeadChip. The same processing procedures and filtering criteria as described above were applied to identify methylation differences between cancerous and normal tissues. Meanwhile, gene expression data, including 51 normal and 641 tumor tissue samples, was obtained from Illumina RNA sequencing (RNA-Seq) platform. After averaging the within-array replicate spots and removing genes that are zero in all libraries, preprocessing data was normalized using trimmed mean of M-values (TMM) method. We tested the differentially expressed genes between cancerous and normal tissues with a literature-based filtering criterion (FDR < 0.05 and fold change > 2) [26–28]. Two packages named edgeR and limma were used for data preprocessing, normalization, and differential expression analysis of TCGA samples.

Validation of methylation status by pyrosequencing

DNA methylation level of particular CpG sites located in the promoter region of candidate genes was quantified by pyrosequencing. Median of individual CpG values represented the DNA methylation status of each gene. A primer set is shown in Additional file 1: Table S1. Pyrosequencing reactions and quantification of DNA methylation were ran on the Pyromark Q96 MD pyrosequencing system (QIAGEN, Valencia, CA, USA).

In order to ensure accuracy and authenticity of our pyrosequencing results, quality control measures were implemented as follows: (1) 1% agarose gel Electrophoresis was chose to test the quality of the extracted DNA. (2) A sample set that included serial dilutions of fully methylated and non-methylated DNAs (Human Methylated & Non-methylated DNA Set, Zymo Research, Freiburg im Breisgau, Germany) (0%, 25%, 50%, 75%, and

100%) was used to control DNA standards. (3) No template controls were included in each experimental run. In addition, all samples were mixed across plates.

Statistical analysis

Data were summarized as mean, standard deviation (SD), median, 5th and 95th percentiles for continuous variables, and absolute and relative frequencies for categorical variables. Differences in the distribution of genome methylation pattern were tested by Pearson χ^2 test.

Differences in methylation and gene expression between cancerous and normal tissues were tested using paired Student's t test and empirical analysis based on the negative binomial distributions [29], respectively. A three-step approach was used to identify biomarkers associated with colorectal carcinogenesis. In the screening phase, we reanalyzed these differentially methylated probes with a set of strict filtering criteria, including (1) β difference > 0.35 and FDR < 0.05 , (2) hypermethylated CpG sites were annotated in both traditional promoter regions (TSS1500, TSS200, 5'UTR and first exon) and CpG islands (CGIs). After the cross-validation analysis using DNA methylation microarray and RNA-Seq data from TCGA database, five top genes showing consistently hypermethylated changes in promoters were selected into the training phase. To verify the accuracy and specificity of these five candidates as a signature, the discriminative performance of selected candidates was assessed by receiver operating characteristic (ROC) curves, and the area under the ROC curve (AUC), sensitivity, and specificity at the optimal cut-offs were calculated. Only potential genes with a high discriminative performance (sensitivity > 0.80 ; specificity > 0.90) [30] were further selected into the validation phase. Based on the conception of field effect, linear mixed-effects modeling analysis was used to describe the methylation changes in the transition from colorectal normal mucosa from individuals without a history of cancer ($n = 142$) to normal mucosa surrounding CRC tissue ($n = 256$), to hyperplastic polyp ($n = 13$), to adenoma ($n = 129$), and to carcinoma ($n = 256$). Correlated errors resulting from within-individual comparisons were assessed by setting up with a random-effect variable for non-independent measurements from matched lesion and normal samples. All these tests were two-sided. P value < 0.05 was considered statistically significant. Statistical analyses were performed using R software (Version 3.3.2).

Results

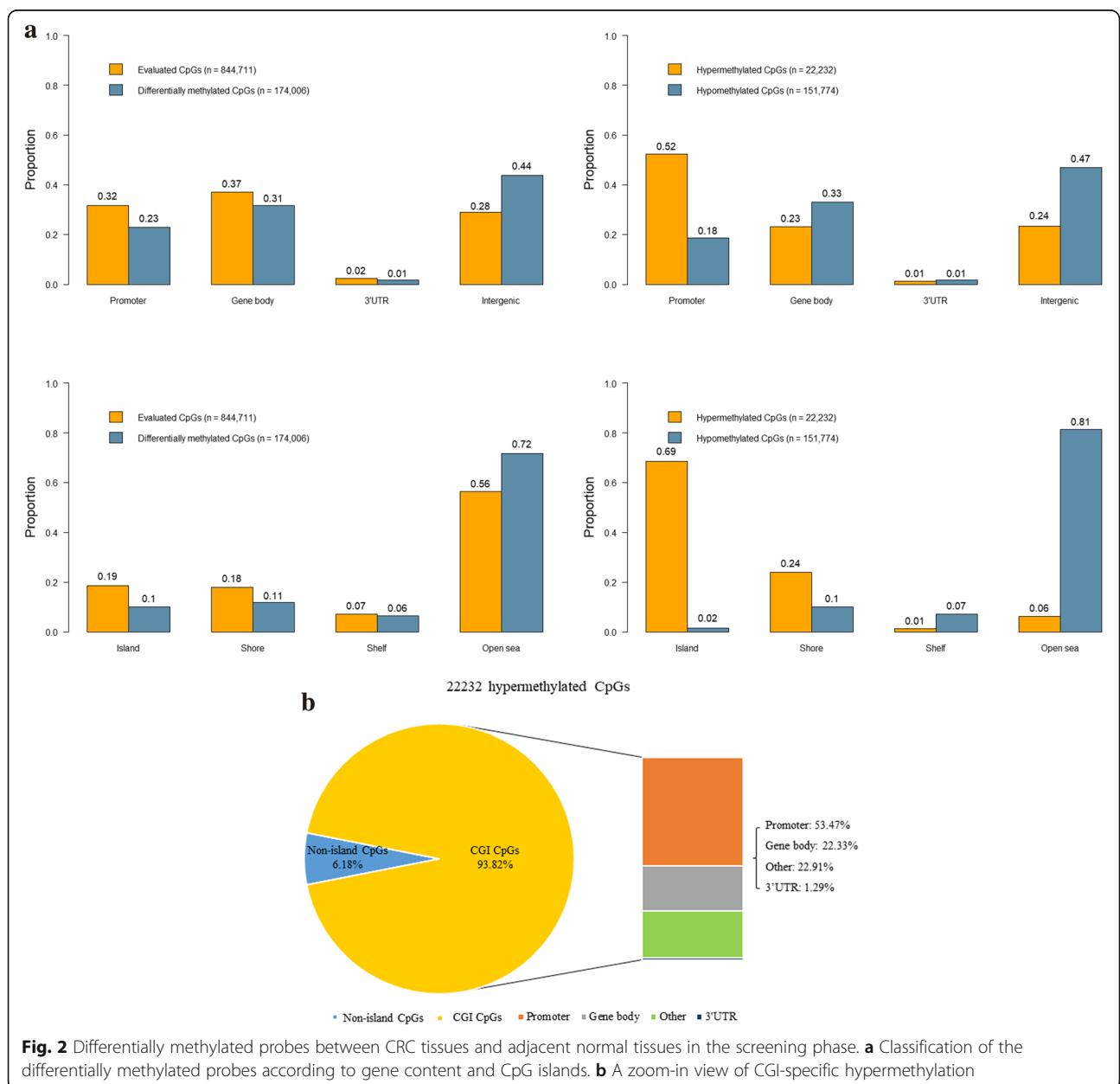
Analysis of global methylation profiles in colorectal cancer

In the screening phase, analysis of the differential methylation between 12 pairs of tumor tissues and adjacent

normal tissues from CRC patients identified a total of 174,006 (21%) CpG sites to be differentially methylated. They were mainly detected in the region of low CpGs (124,917/174,006 CpG sites 72% vs. 477,028/844,711 CpG sites, 56%, $P < 0.0001$), commonly named “open sea,” and far from gene promoters (134,228/174,006 CpG sites, 77% vs. 576,610/844,711 CpG sites, 68%, $P < 0.0001$), in comparison with the reference distribution of probes on EPIC platform (Fig. 2a left panel).

Among the identified differentially methylated CpG sites, 22,232 (13%) were significantly hypermethylated

and 151,774 (87%) were significantly hypomethylated. An overview of the frequency of all hyper- and hypomethylated CpG sites classified by CGI and gene content is shown in Fig. 2a (right panel). Significant hypomethylation was predominantly observed in open sea (123,542/151,774 CpG sites, 81%) and intergenic regions (71,100/151,774 CpG sites, 47%), while most significant hypermethylation was specifically located around CGIs (20,857/22,232 CpG sites, 94%) and gene promoters (11,636/22,232 CpG sites, 52%). Then, the overlap of CGI hypermethylation and genomic locations was examined in detail. Approximately, 53% (11,151 CpG sites) of CGI



hypermethylation overlapped specifically with gene promoters (Fig. 2b).

Cross-validation analysis with TCGA data

Considering that promoter-associated hypermethylation can trigger transcriptional silencing of the target gene, we merged the promoter methylation with RNA-Seq gene expression data by using data from TCGA and further identified a list of 265 genes undergoing promoter hypermethylation during colorectal tumor progression. As shown in Additional file 2: Table S2, this collection of genes showed an overlap between hypermethylation and downregulated expression in TCGA database.

Among the list of CpG sites, we focused on five sites mapping the promoter regions of their own genes for following training analysis (Fig. 3). Three of them were novel hypermethylated genes (cg11855526 in *MPPED2*, cg25437410 in *COL23A1*, and cg15093079 in *EPHA6*) ranking on the top, which were considered as candidate biomarkers. The remaining two (cg20078466 in *IKZF1* and cg07279933 in *RSPO3*), were chosen based on the publications, which were considered as the reference. Figure 4 presents the position of candidate gene promoters in relation to the respective CGIs.

Validation of promoter methylation status using pyrosequencing

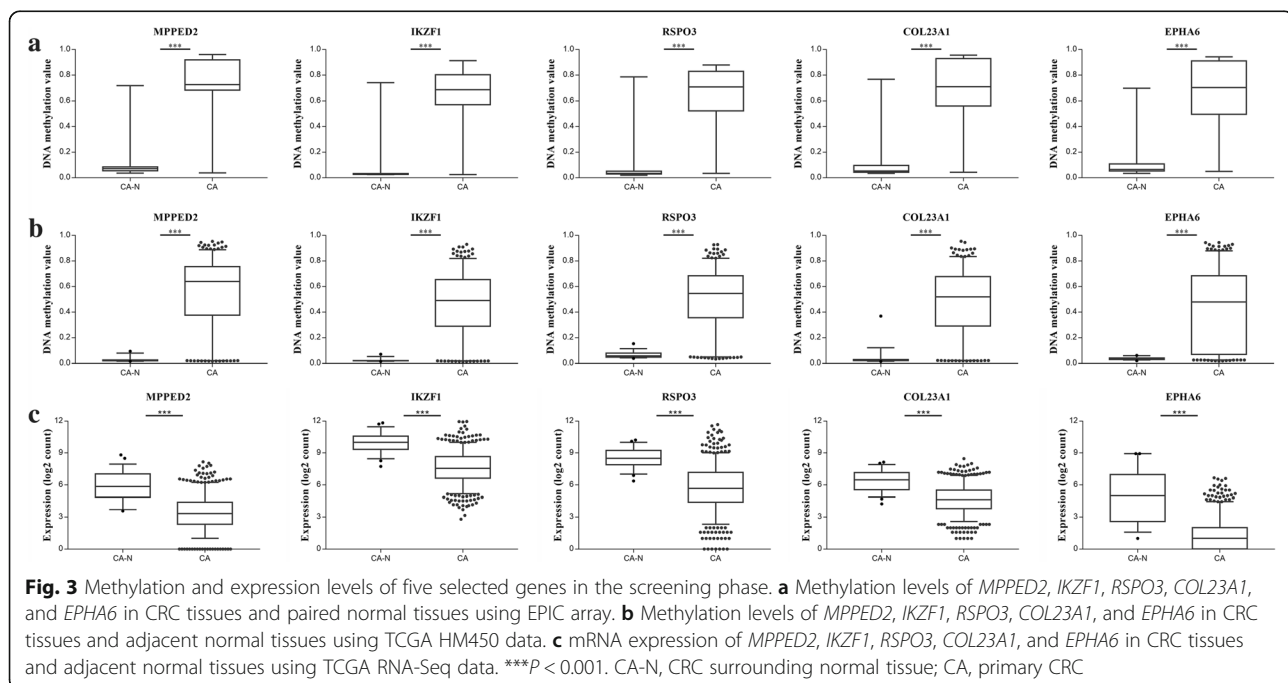
In the training phase, DNA methylation status of the above five candidate genes in 46 CRCs and paired normal tissues was detected by pyrosequencing. Box plots showing the distribution of β values of these genes'

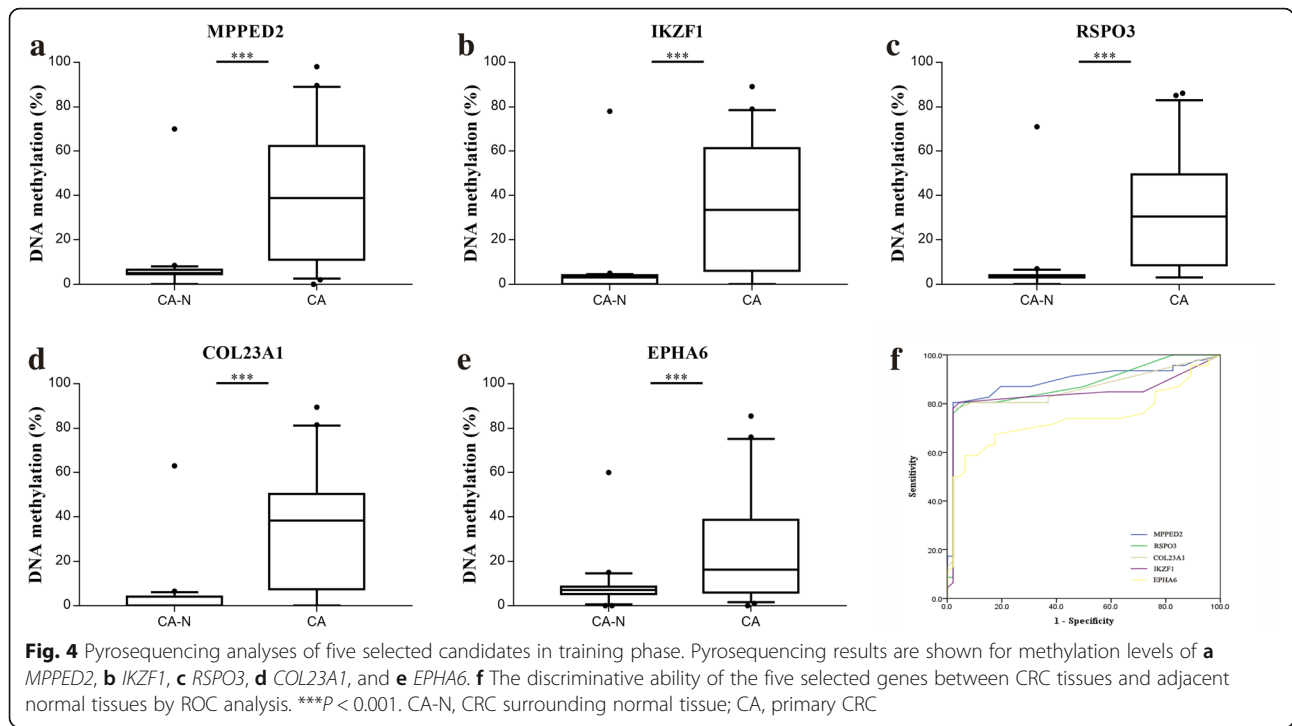
methylation status are shown in Fig. 5a–e. All five genes demonstrated significantly differential hypermethylation. The increase in methylation status between cancerous and paired normal mucosa was 33.35%, 30.54%, 27.28%, 31.62%, and 18.18% for *MPPED2*, *IKZF1*, *RSPO3*, *COL23A1*, and *EPHA6*, respectively (all $P < 0.001$). Further Pearson correlation analysis showed a strong correlation between the methylation level of *MPPED2* and that of *RSPO3* ($R = 0.81$, $P < 0.001$) (Additional file 3: Figure S1b).

ROC curve analyses revealed that methylation status of each individual genes significantly distinguished primary carcinoma from normal mucosa, as measured by sensitivity, specificity, and AUC value (*MPPED2*: sensitivity 0.804, specificity 0.978, AUC 0.890; *IKZF1*: sensitivity 0.761, specificity 0.978, AUC 0.875; *RSPO3*: sensitivity 0.783, specificity 0.978, AUC 0.858; *COL23A1*: sensitivity 0.783, specificity 0.978, AUC 0.840; *EPHA6*: sensitivity 0.587, specificity 0.935, AUC 0.736) (Fig. 5f, Additional file 4: Table S3). Among them, *MPPED2* presented a high discriminative performance. Therefore, it was chosen for further validation study.

Analysis of aberrant methylation pattern of *MPPED2* during neoplastic progression of colorectal cancer

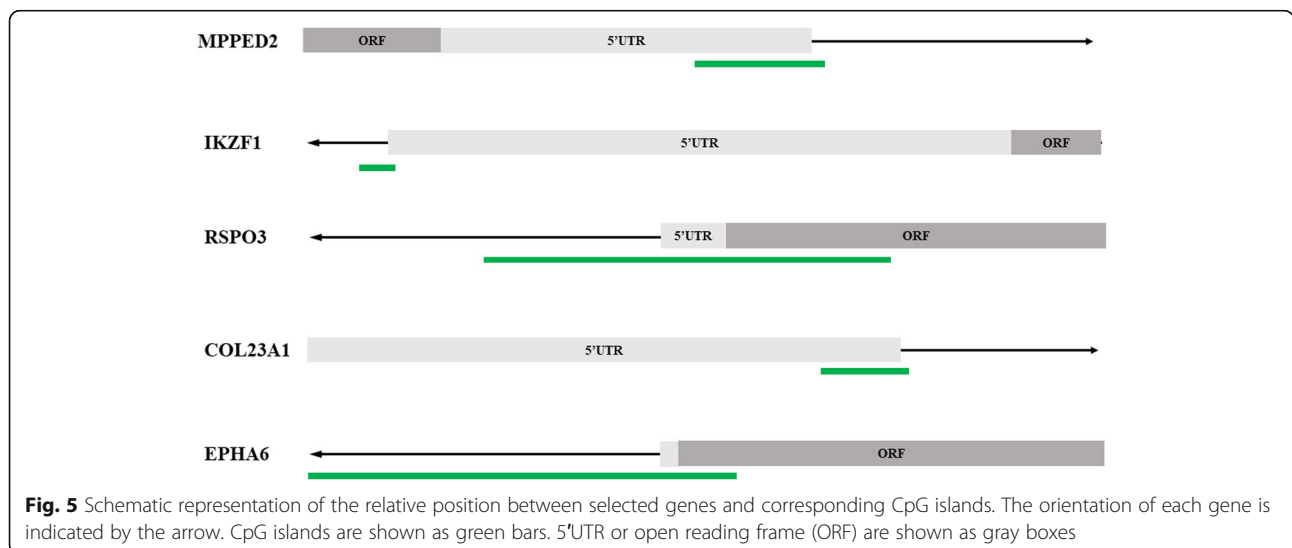
In order to define the *MPPED2* methylation changes during colorectal neoplastic progression, DNA methylation status among 796 colorectal tissue samples, including colorectal lesions and paired normal tissues from 13 hyperplastic polyps, 129 adenomas, and 256 primary carcinomas, was assessed during the validation phase. As

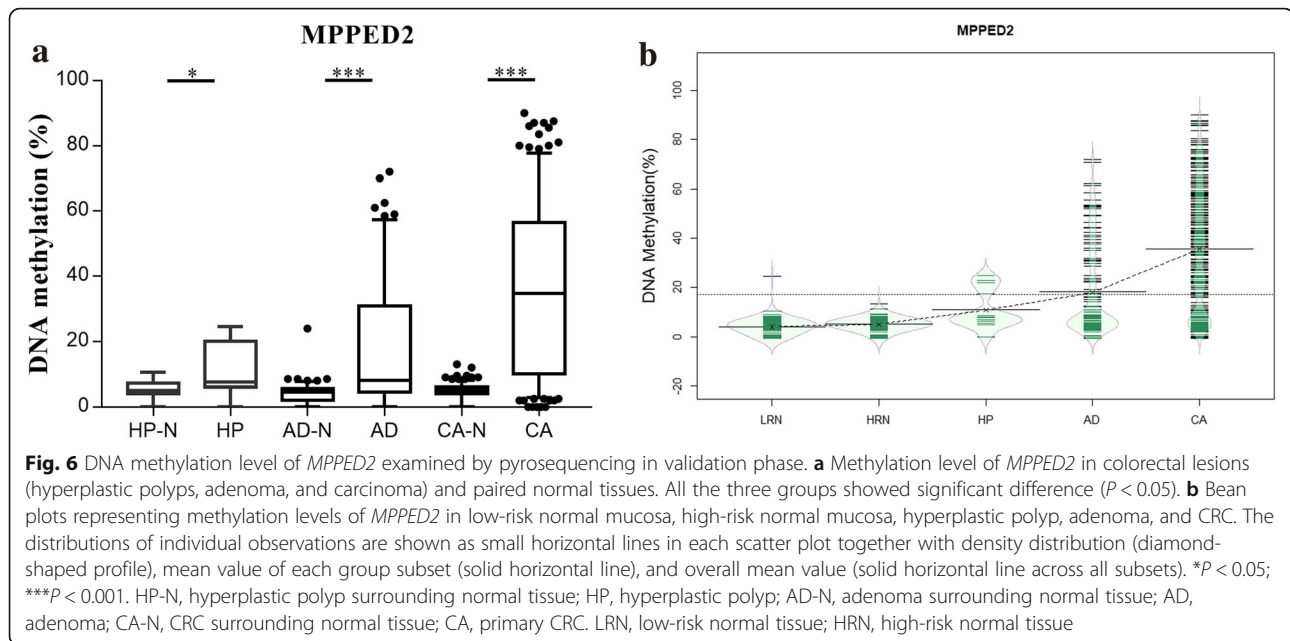




shown in Fig. 6a, the methylation status of *MPPED2* increased significantly in each colorectal lesion groups (all $P < 0.05$), in comparison with their respective normal groups. The methylation changes also increased quantitatively with the neoplastic progression (5.58% for hyperplastic polyps; 14.5% for adenomas; 30.56% for carcinomas). To examine the hierarchy of methylation events during colorectal neoplastic progression, the linear mixed-effects modelling analysis was performed. We

found that *MPPED2* methylation pattern showed a step-wise increase from low-risk normal mucosa to high-risk normal mucosa, to hyperplastic polyp, to adenoma, and to carcinoma ($P < 0.001$). A direct overview of this tendency was noted by intensity coefficients (4.07 for low-risk normal mucosa; 5.02 for high-risk normal mucosa; 10.86 for hyperplastic polyp; 17.95 for adenoma; 35.58 for carcinoma, respectively) (Fig. 6b). We also examined the interaction between methylation status and





clinical characteristics. However, these results did not present any statistical significance (Additional file 3: Figure S1).

Discussion

In this study, we performed a comprehensive epigenomic profiling assessment in CRC patients for screening promising biomarkers and carried out a model-based analysis to determine the hierarchy of DNA methylation changes during the carcinogenic process. Our results showed global alterations in DNA methylation patterns in CRC tissues with a large number of differentially methylated CpG sites, compared to their adjacent normal tissues. We also found that the pattern of differential methylation showed opposite characteristics between hypomethylation and hypermethylation across gene and island distribution. Analyses on both TCGA data and internal datasets provided the confirmatory evidence for our findings. Linear mixed-effects modeling analysis revealed that DNA methylation levels of *MPPED2* increased sequentially from normal to hyperplastic polyp, to adenoma, and to carcinoma.

As observed in many cancers [31–33], we found a common methylated pattern with 174,006 (21%) differentially methylated CpG sites while comparing CRC tissues to matched normal tissues, suggesting the widespread aberrations in methylation status during cancer development. Remarkably, we found that a majority of the identified CpG sites were hypomethylation rather than hypermethylation, which were different from what have been shown in general array-based studies [34–37]. This could be explained by the increased genome

coverage of the EPIC array. To be specific, the Illumina 27 K BeadChip (27 K) array targeted about 27,000 CpG sites and interrogated at least one site per gene with a bias towards CGIs. The second-generation product, HM450 array, provided more than 485,000 probes focusing on 99% of RefSeq genes, while a large number of new probes in EPIC array solve the limited interrogation in non-CGI regions before. Therefore, our findings with EPIC platform might provide a more unbiased and reasonable map of CRC epigenome.

CGI hypermethylation in promoter of specific genes has been linked to epigenetic transcriptional silencing, which appears to be crucial in the early stage of carcinogenesis. It is worthwhile to understand cancer-specific promoter hypermethylation with transcriptional repression in CRC. Therefore, we extended this analysis by performing cross-validation analysis using DNA methylation microarray and RNA-Seq data from TCGA database. Within 649 genes, 1666 CGI promoter CpG sites were further confirmed to be hypermethylated. However, not all the expression of these genes were significantly downregulated in CRCs compared with normal mucosa. The result indicates the fact that many epigenetic alterations in cancers are passengers that may occur as a consequence of or in association with carcinogenesis [38]. Moreover, multiple regulation mechanisms may implicate in gene expression, which could obscure a direct view of correlations between DNA methylation and gene expression. Although much remains unknown concerning the involvement of multiple regulation mechanisms in human cancers, it is feasible to pinpoint crucial genes that are suspectful for methylation-associated gene inactivation.

Among the cluster of candidate genes characterized by epigenetic silencing, five genes, named *MPPED2*, *COL23A1*, *EPHA6*, *RSPO3*, and *IKZF1*, were identified as of potential interest for CRC. They all have been reported to be associated with human cancers, which support our findings a possible role in the pathogenesis of CRC.

MPPED2 gene encodes a new metallophosphoesterase protein in mammals, and it normally regulates many essential cellular functions including differentiation, proliferation, and apoptosis [39, 40]. It is expressed ubiquitously in most human tissues, and the downregulated expression of *MPPED2* has been reported in several malignant tumors, including oral squamous cell carcinoma [40], papillary thyroid carcinoma [41], and cervical cancer [42]. Notably, recent evidence indicates that the tumor-suppressing activity of *MPPED2* could be epigenetically modified by DNA methylation [43]. Liguori et al. found that *MPPED2* expression significantly increased after the treatment with demethylation agent. In vivo, *MPPED2* could serve a function of anti-proliferation, which showed that the biochemical activity of *MPPED2* was beneficial for cancer suppression [39]. The above findings support the hypothesis that hypermethylation-induced down-expression is the most likely mechanism of *MPPED2* silencing during carcinogenesis.

Investigation of the other genes of possible biological significance in CRC identified that *EPHA6* was described as the tumor suppressor gene according to its role in the regulation of angiogenesis process. It has been found to be downregulated in colorectal cancer [44, 45]. *COL23A1* is known to be one of the transmembrane collagens. Survival analysis showed significant prognostic implications of *COL23A1* expression in non-small cell lung cancer [46] and prostate cancer [47]. As for *RSPO3*, our result is consistent with previous findings that *RSPO3* could be transcriptionally downregulated by promoter hypermethylation in CRC [35]. *IKZF1* was described as a tumor suppressor gene based on its role in regulation of cellular proliferation [48, 49]. It has been shown to be hypermethylated in CRC [50] and acute lymphoblastic leukemia [51]. Though there existed a strong correlation of methylation levels between *MPPED2* and *RSPO3*, no definite mechanism could explain this phenomenon. One recent computational study [52] showed that both *MPPED2* and *RSPO3* play a role in carcinogenic process across cancer types, which suggests that they might be part of the active processes underlying colorectal carcinogenesis.

It should be mentioned that some of the current hypotheses highlight that the regulatory role of epigenetic events could be implicated in histologically normal tissues surrounding cancer lesions, which is known as the field cancerization effect [53, 54]. DNA methylation has been shown to be involved in field effect in a variety

of tissues, including esophageal mucosa in Barrett's esophagus [55], normal-appearing gastric mucosa [56], and colonic mucosa effected by ulcerative colitis [57], highlighting the possible role of epigenetic alterations in cancer development. To answer the question when methylation levels of candidates occur significant changes during colorectal cancer progression, a linear mixed-effects model was used to combine the type of tissue, within- and between-individual comparisons, and methylation changes of candidate biomarkers [58]. The model allowed us to quantify the methylation changes in the transition from normal colonic mucosa to hyperplastic polyp, to adenoma, and to carcinoma, while taking into account within-individual comparisons. In this study, cumulative methylation alternations of *MPPED2* promoter showed a significant association with the progression of colorectal neoplasia, suggesting that aberrant methylation is an early event and occurs sequentially during colorectal neoplastic progression. It might be a useful biomarker for early detection and risk assessment in colorectal cancer.

Array-based analysis is a useful tool for genome-wide DNA methylation screening. The new-generation microarray, EPIC platform provides substantially increased genomic coverage than prior studies, therefore allowing identification of novel methylated CpG sites that have not been previously revealed. Although reports using EPIC platform are still scarce, evaluation analysis performed by Pidsley et al. [59], Solomon et al. [60], and Kling et al. [61] found that data from HM450, EPIC, and whole-genome bisulphite sequencing (WGBS) were highly reproducible across technical and biological replicates. However, one recent study [62] demonstrated that data from HM450 platform showed general congruence with data obtained from the same probes using the EPIC platform. Nevertheless, the combined application of data from EPIC and existed HM450 platform are encouraged for identifying new important insights in genomic regulation in disease states.

There are several limitations in this study. First, matched expression analysis was unavailable, which did not allow us to give a directly causal inference between DNA methylation and gene expression. Nevertheless, integration from the published database pinpointed the potential negative associations between DNA methylation and gene expression in CRC tissues. Second, for hyperplastic polyps, the sample size in validation phase was modest, resulting in a less widespread density shape in linear mixed-effects model analysis. Last, although we found that aberrant methylation of *MPPED2* could serve as an amplifiable signal in colorectal tumorigenesis, external validations consisting of one or more dataset from different institutions are still needed to evaluate the performance and potential clinical value of *MPPED2*

in CRC. Furthermore, it might be instructive to use different sample types, such as stool or plasma as well. Taken together, although we successfully assessed epigenetic alterations of *MPPED2* promoter hypermethylation during the colorectal neoplastic progression, further studies with the diversity of study population and sample types, and larger sample size are warranted to validate the above findings and elucidate the corresponding mechanism in colorectal carcinogenesis.

Conclusion

In summary, our work gave a detailed assessment of DNA methylation pattern with over 850,000 CpG probes for CRCs and revealed epigenetically regulated candidate genes in colorectal carcinogenesis. Specifically, our results provide the first evidence that tumor-specific hypermethylation in *MPPED2* promoter occurs frequently in CRCs and its increasing accumulation means the further development of colorectal neoplasia. The findings may offer an instructive clue for understanding the role of DNA methylation in the development of early colorectal neoplasia and provide promising signature for the use of epigenetic profiling in CRC detection and therapy decisions.

Additional files

Additional file 1: Table S1. Pyrosequencing methylation assay Primers. (DOCX 13 kb)

Additional file 2: Table S2. A list of differentially hypermethylated and expressed genes identified by both EPIC array and TCGA database. (DOCX 173 kb)

Additional file 3: Figure S1. a Pearson correlation analysis between age and methylation level of *MPPED2*. b Pearson correlation analysis between methylation levels of five selected genes c Comparisons of methylation level of *MPPED2* in different stages and locations of CRC tissues in the validation phase. (TIF 628 kb)

Additional file 4: Table S3. Results of ROC curve analysis between CRC tissues and adjacent normal tissues. (DOCX 12 kb)

Abbreviations

27 K: Illumina 27 K BeadChip; AD: Adenoma; AD-N: Adenoma surrounding normal tissue; AJCC: American Joint Committee on Cancer; AUC: Area under curve; BMIQ: Beta-mixture quantile normalization; CA: Primary CRC; CA-N: CRC surrounding normal tissue; CGI: CpG island; CRC: Colorectal cancer; EPIC: Illumina MethylationEPIC BeadChip; FAP: Familial adenomatous polyposis; FDR: Adjusted *P* value with Benjamini-Hochberg method; HM450: Illumina HumanMethylation450 BeadChip; HP: Hyperplastic polyp; HP-N: Hyperplastic polyp surrounding normal tissue; HRN: High-risk normal tissue; IMA: Illumina Methylation Analyzer; LRN: Low-risk normal tissue; ORF: Open reading frame; RNA-Seq: RNA sequencing; ROC: Receiver operating characteristic; SD: Standard deviation; TCGA: The Cancer Genome Atlas

Acknowledgments

We sincerely thank all participants in the study.

Funding

This study was supported by National Basic Research Program of China (973 Program No.2015CB554003) and National Science Foundation of China (NSFC No.81673262).

Authors' contributions

SG, MJ, and KC conceived and designed the study. SL, DY, and SQ conducted the experiments. SG, DJ, and XZ made contributions to statistical and bioinformatics analyses. QL, JY, XY, ZL, MT, and JW recruited the participants and performed the histopathological evaluation. SG completed the manuscript with intellectual input from MJ and KC. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This study was approved by Medical Ethics Committee of Zhejiang University School of Medicine. All participants were recruited after providing a signed informed consent.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Epidemiology and Biostatistics, Zhejiang University School of Public Health, 866 Yuhangtang Road, Hangzhou 310058, China. ²Department of Epidemiology and Biostatistics, Zhejiang Chinese Medical University School of Public Health, 548 Binwen Road, Hangzhou 310053, China. ³Jiashan Institute of Cancer Prevention and Treatment, 345 Jiefangdong Road, Jiashan 314100, China. ⁴Department of Anorectal Surgery, Shaoxing People's Hospital, 568 Zhongxingbei Road, Shaoxing 312000, China. ⁵Cancer Institute, the Second Affiliated Hospital, Zhejiang University School of Medicine, 88 Jiefang Road, Hangzhou 310009, China.

Received: 29 November 2018 Accepted: 4 February 2019

Published online: 07 March 2019

References

- Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut*. 2017; 66(4):683–91. <https://doi.org/10.1136/gutjnl-2015-310912>.
- Esteller M, Herman JG. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *J Pathol*. 2002;196(1):1–7. <https://doi.org/10.1002/path.1024>.
- Schnekenburger M, Florean C, Dicato M, Diederich M. Epigenetic alterations as a universal feature of cancer hallmarks and a promising target for personalized treatments. *Curr Top Med Chem*. 2016;16(7):745–76.
- Lao W, Grady WM. Epigenetics and colorectal cancer. *Nat Rev Gastroenterol Hepatol*. 2011;8(12):686–700. <https://doi.org/10.1038/nrgastro.2011.173>.
- Okugawa Y, Grady WM, Goel A. Epigenetic alterations in colorectal cancer: emerging biomarkers. *Gastroenterology*. 2015;149(5):1204–25. <https://doi.org/10.1053/j.gastro.2015.07.011>.
- Goelz SE, Vogelstein B, Hamilton SR, Feinberg AP. Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science*. 1985; 228(4696):187–90.
- Chang E, Park DJ, Kim YJ, Kim BK, Park JH, Kim HJ, et al. Detection of colorectal neoplasm using promoter methylation of ITGA4, SFRP2, and p16 in stool samples: a preliminary report in Korean patients. *Hepato-Gastroenterology*. 2010;57(101):720–7.
- Ng JM, Yu J. Promoter hypermethylation of tumour suppressor genes as potential biomarkers in colorectal cancer. *Int J Mol Sci*. 2015;16(2):2472–96. <https://doi.org/10.3390/ijms16022472>.
- Li X, Yao X, Wang Y, Hu F, Wang F, Jiang L, et al. MLH1 promoter methylation frequency in colorectal cancer patients and related clinicopathological and molecular features. *Plos One*. 2013;8(3):e59064. <https://doi.org/10.1371/journal.pone.0059064>.
- Kalmar A, Peterfia B, Hollosi P, Galamb O, Spisak S, Wichmann B, et al. DNA hypermethylation and decreased mRNA expression of MAL, PRIMA1, PTGDR and SFRP1 in colorectal adenoma and cancer. *Bmc Cancer*. 2015;15:736. <https://doi.org/10.1186/s12885-015-1687-x>.

11. Beggs AD, Jones A, El-Bahrawy M, Abulafi M, Hodgson SV, Tomlinson IP. Whole-genome methylation analysis of benign and malignant colorectal tumours. *J Pathol*. 2013;229(5):697–704. <https://doi.org/10.1002/path.4132>.
12. Teschendorff AE, Gao Y, Jones A, Ruebner M, Beckmann MW, Wachter DL, et al. DNA methylation outliers in normal breast tissue identify field defects that are enriched in cancer. *Nat Commun*. 2016;7:10478. <https://doi.org/10.1038/ncomms10478>.
13. van Houten VM, Leemans CR, Kummer JA, Dijkstra J, Kuik DJ, van den Brekel MW, et al. Molecular diagnosis of surgical margins and local recurrence in head and neck cancer patients: a prospective study. *Clin Cancer Res*. 2004; 10(11):3614–20. <https://doi.org/10.1158/1078-0432.CCR-03-0631>.
14. Giovannucci E, Ogino S. DNA methylation, field effects, and colorectal cancer. *J Natl Cancer Inst*. 2005;97(18):1317–9. <https://doi.org/10.1093/jnci/dji305>.
15. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer Am Cancer Soc*. 1953;6(5):963–8.
16. Belshaw NJ, Elliott GO, Foxall RJ, Dainty JR, Pal N, Coupe A, et al. Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa. *Br J Cancer*. 2008;99(1):136–42. <https://doi.org/10.1038/sj.bjc.6604432>.
17. Belshaw NJ, Pal N, Tapp HS, Dainty JR, Lewis MP, Williams MR, et al. Patterns of DNA methylation in individual colonic crypts reveal aging and cancer-related field defects in the morphologically normal mucosa. *Carcinogenesis*. 2010;31(6):1158–63. <https://doi.org/10.1093/carcin/bgq077>.
18. Alonso S, Dai Y, Yamashita K, Horiuchi S, Dai T, Matsunaga A, et al. Methylation of MGMT and ADAMTS14 in normal colon mucosa: Biomarkers of a field defect for cancerization preferentially targeting elder African-Americans. *Oncotarget*. 2015;6(5):3420–31. <https://doi.org/10.18632/oncotarget.2852>.
19. Chan AO, Broaddus RR, Houlihan PS, Issa JP, Hamilton SR, Rashid A. CpG island methylation in aberrant crypt foci of the colorectum. *Am J Pathol*. 2002;160(5):1823–30. [https://doi.org/10.1016/S0002-9440\(10\)61128-5](https://doi.org/10.1016/S0002-9440(10)61128-5).
20. Li H, Myeroff L, Smiraglia D, Romero MF, Pretlow TP, Kasturi L, et al. SLC5A8, a sodium transporter, is a tumor suppressor gene silenced by methylation in human colon aberrant crypt foci and cancers. *Proc Natl Acad Sci U S A*. 2003;100(14):8412–7. <https://doi.org/10.1073/pnas.1430846100>.
21. Rashid A, Shen L, Morris JS, Issa JP, Hamilton SR. CpG island methylation in colorectal adenomas. *Am J Pathol*. 2001;159(3):1129–35. [https://doi.org/10.1016/S0002-9440\(10\)61789-0](https://doi.org/10.1016/S0002-9440(10)61789-0).
22. Jin M, Gu S, Ye D, Li Y, Jing F, Li Q, et al. Association between genetic variants in the promoter region of a novel antisense long noncoding RNA RP11-392P7.6 and colorectal cancer risk. *Environ Mol Mutagen*. 2017;58(6): 434–42. <https://doi.org/10.1002/em.22100>.
23. Ye D, Hu Y, Jing F, Li Y, Gu S, Jiang X, et al. A novel SNP in promoter region of RP11-3N2.1 is associated with reduced risk of colorectal cancer. *J Hum Genet*. 2018;63(1):47–54. <https://doi.org/10.1038/s10038-017-0361-3>.
24. Liu J, Siegmund KD. An evaluation of processing methods for HumanMethylation450 BeadChip data. *BMC Genomics*. 2016;17:469. <https://doi.org/10.1186/s12864-016-2819-7>.
25. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*. 2013;29(2):189–96. <https://doi.org/10.1093/bioinformatics/bts680>.
26. Li J, Han L, Roebuck P, Diao L, Liu L, Yuan Y, et al. TANRIC: an interactive open platform to explore the function of lncRNAs in cancer. *Cancer Res*. 2015;75(18):3728–37.
27. Xu X, Zhang Y, Williams J, Antoniou E, Mccombie WR, Wu S, et al. Parallel comparison of Illumina RNA-Seq and Affymetrix microarray platforms on transcriptomic profiles generated from 5-aza-deoxy-cytidine treated HT-29 colon cancer cells and simulated datasets. *Bmc Bioinformatics*. 2013;14(S9):S1.
28. Rampal R, Alshahrour F, Abdelwahab O, Patel JP, Brunel JP, Mermel CH, et al. Integrated genomic analysis illustrates the central role of JAK-STAT pathway activation in myeloproliferative neoplasm pathogenesis. *Blood*. 2014;123(22):123–33.
29. Robinson MD, McCarthy DJ, Smyth GK. EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
30. McDermott J. Requirements for diagnostic tests to improve field investigations into the epidemiology and control of trypanosomiasis. In: Morzaria S, Masake R, Rowlands J, Musoke T, editors. *Antigen ELISAs for trypanosomes*. Nairobi: ILRI; 1996.
31. Jones PA, Baylin SB. The epigenomics of cancer. *Cell*. 2007;128(4):683–92. <https://doi.org/10.1016/j.cell.2007.01.029>.
32. Shi YX, Wang Y, Li X, Zhang W, Zhou HH, Yin JY, et al. Genome-wide DNA methylation profiling reveals novel epigenetic signatures in squamous cell lung cancer. *Bmc Genomics*. 2017;18(1):901. <https://doi.org/10.1186/s12864-017-4223-3>.
33. Zheng Y, Huang Q, Ding Z, Liu T, Xue C, Sang X, et al. Genome-wide DNA methylation analysis identifies candidate epigenetic markers and drivers of hepatocellular carcinoma. *Brief Bioinform*. 2018;19(1):101–8. <https://doi.org/10.1093/bib/bbw094>.
34. McInnes T, Zou D, Rao DS, Munro FM, Phillips VL, JL MC, et al. Genome-wide methylation analysis identifies a core set of hypermethylated genes in CIMP-H colorectal cancer. *Bmc Cancer*. 2017;17(1):228. <https://doi.org/10.1186/s12885-017-3226-4>.
35. Kok-Sin T, Mokhtar NM, Ali HN, Sagap I, Mohamed RI, Harun R, et al. Identification of diagnostic markers in colorectal cancer via integrative epigenomics and genomics data. *Oncol Rep*. 2015;34(1):22–32. <https://doi.org/10.3892/or.2015.3993>.
36. Vymetalkova V, Vodicka P, Pardini B, Rosa F, Levy M, Schneiderova M, et al. Epigenome-wide analysis of DNA methylation reveals a rectal cancer-specific epigenomic signature. *Epigenomics-Uk*. 2016;8(9):1193–207. <https://doi.org/10.2217/epi-2016-0044>.
37. Naumov VA, Generozov EV, Zaharjevskaya NB, Matushkina DS, Larin AK, Chernyshov SV, et al. Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips. *Epigenetics-U.S.* 2013;8(9):921–34. <https://doi.org/10.4161/epi.25577>.
38. Kalari S, Pfeifer GP. Identification of driver and passenger DNA methylation in cancer by epigenomic analysis. *Adv Genet*. 2010;70:277–308. <https://doi.org/10.1016/B978-0-12-380866-0.60010-1>.
39. Liguori L, Andolfo I, de Antonellis P, Aglio V, di Dato V, Marino Net al. The metallophosphodiesterase Mpped2 impairs tumorigenesis in neuroblastoma. *Cell Cycle*. 2012;11(3):569–81. <https://doi.org/10.4161/cc.11.3.19063>.
40. Shen L, Liu L, Ge L, Xie L, Liu S, Sang Let al. MiR-448 downregulates MPPED2 to promote cancer proliferation and inhibit apoptosis in oral squamous cell carcinoma. *Exp Ther Med*. 2016;12(4):2747–52. <https://doi.org/10.3892/etm.2016.3659>.
41. Sepe R, Pellecchia S, Serra P, D'Angelo D, Federico A, Raia Met al. The long Non-Coding RNA RP5-1024C24.1 and its Associated-Gene MPPED2 are Down-Regulated in human thyroid neoplasias and act as tumour suppressors. *Cancers (Basel)*. 2018;10(5) <https://doi.org/10.3390/cancers10050146>.
42. Zhang R, Shen C, Zhao L, Wang J, McCrae M, Chen Xet al. Dysregulation of host cellular genes targeted by human papillomavirus (HPV) integration contributes to HPV-related cervical carcinogenesis. *Int J Cancer*. 2016;138(5): 1163–74. <https://doi.org/10.1002/ijc.29872>.
43. Rahmatpanah FB, Carstens S, Hooshmand SI, Welsh EC, Sjahputera O, Taylor KHet al. Large-scale analysis of DNA methylation in chronic lymphocytic leukemia. *Epigenomics-Uk*. 2009;1(1):39–61. <https://doi.org/10.2217/epi.09.10>.
44. Zuo S, Dai G, Ren X. Identification of a 6-gene signature predicting prognosis for colorectal cancer. *Cancer Cell Int*. 2019;19:6. <https://doi.org/10.1186/s12935-018-0724-7>.
45. Hafner C, Schmitz G, Meyer S, Bataille F, Hau P, Langmann T, et al. Differential gene expression of Eph receptors and ephrins in benign human tissues and cancers. *Clin Chem*. 2004;50(3):490–9. <https://doi.org/10.1373/clinchem.2003.026849>.
46. Spivey KA, Banyard J, Solis LM, Wistuba II, Barletta JA, Gandhi L, et al. Collagen XXIII: a potential biomarker for the detection of primary and recurrent non-small cell lung cancer. *Cancer Epidemiol Biomark Prev*. 2010; 19(5):1362–72. <https://doi.org/10.1158/1055-9965.EPI-09-1095>.
47. J. Banyard, L. Bao, M. D. Hofer, D. Zurakowski, K. A. Spivey, A. S. Feldman, L. M. Hutchinson, R. Kuefer, M. A. Rubin, B. R. Zetter, (2007) Collagen XXIII Expression Is Associated with Prostate Cancer Recurrence and Distant Metastases. *Clinical Cancer Research*. 13(9):2634–42..
48. Marke R, Havinga J, Cloos J, Demkes M, Poelmans G, Yuniati Let al. Tumor suppressor IKZF1 mediates glucocorticoid resistance in B-cell precursor acute lymphoblastic leukemia. *Leukemia*. 2016;30(7):1599–603. <https://doi.org/10.1038/leu.2015.359>.
49. Scheijen B, Boer JM, Marke R, Tijchon E, van Ingen SD, Waanders Eet al. Tumor suppressors BTG1 and IKZF1 cooperate during mouse leukemia development and increase relapse risk in B-cell precursor acute

- lymphoblastic leukemia patients. *Haematologica*. 2017;102(3):541-51. <https://doi.org/10.3324/haematol.2016.153023>.
50. Jedi M, Young GP, Pedersen SK, Symonds EL. Methylation and gene expression of BCAT1 and IKZF1 in colorectal cancer tissues. *Clin Med Insights Oncol*. 2018;12: 2014496488. <https://doi.org/10.1177/1179554918775064>.
 51. Rahmani M, Talebi M, Hagh MF, Feizi A, Solali S. Aberrant DNA methylation of key genes and Acute Lymphoblastic Leukemia. *Biomed Pharmacother*. 2018;97:1493-500. <https://doi.org/10.1016/j.biopha.2017.11.033>.
 52. Wang Y, Liu Q, Huang S, Yuan B. Learning a structural and functional representation for gene expressions: To systematically dissect complex cancer phenotypes. *IEEE/ACM Trans Comput Biol Bioinform*. 2017; <https://doi.org/10.1109/TCBB.2017.2702161>.
 53. Sugai T, Yoshida M, Eizuka M, Uesugii N, Habano W, Otsuka Ket al. Analysis of the DNA methylation level of cancer-related genes in colorectal cancer and the surrounding normal mucosa. *Clin Epigenetics*. 2017;9:55. <https://doi.org/10.1186/s13148-017-0352-4>.
 54. Cesaroni M, Powell J, Sapienza C. Validation of methylation biomarkers that distinguish normal colon mucosa of cancer patients from normal colon mucosa of patients without cancer. *Cancer Prev Res (Phila)*. 2014;7(7):717-26. <https://doi.org/10.1158/1940-6207.CAPR-13-0407>.
 55. Eads CA, Lord RV, Kurumboor SK, Wickramasinghe K, Skinner ML, Long Tlet al. Fields of aberrant CpG island hypermethylation in Barrett's esophagus and associated adenocarcinoma. *Cancer Res*. 2000;60(18):5021-6.
 56. Tamura G. Promoter methylation status of tumor suppressor and tumor-related genes in neoplastic and non-neoplastic gastric epithelia. *Histol Histopathol*. 2004;19(1):221-8. <https://doi.org/10.14670/HH-19.221>.
 57. Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in ulcerative colitis. *Cancer Res*. 2001;61(9):3573-7.
 58. A. E. K. Ibrahim, M. J. Arends, A.-L. Silva, A. H. Wyllie, L. Greger, Y. Ito, S. L. Vowler, T. H.-M. Huang, S. Tavare, A. Murrell, J. D. Brenton, Sequential DNA methylation changes are associated with DNMT3B overexpression in colorectal neoplastic progression. *Gut* 2011;60(4):499-508
 59. Ruth Pidsley, Elena Zotenko, Timothy J. Peters, Mitchell G. Lawrence, Gail P. Risbridger, Peter Molloy, Susan Van Dijk, Beverly Muhlhauser, Clare Stirzaker, Susan J. Clark, Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology* 2016;17(1)
 60. Olivia Solomon, Julie Maclsaac, Hong Quach, Gwen Tindula, Michael S. Kobar, Karen Huen, Michael J. Meaney, Brenda Eskenazi, Lisa F. Barcellos, Nina Holland, Comparison of DNA methylation measured by Illumina 450K and EPIC BeadChips in blood of newborns and 14-year-old children. *Epigenetics* 2018;13(6):655-64
 61. Teresia Kling, Anna Wenger, Stephan Beck, Helena Carén, Validation of the MethylationEPIC BeadChip for fresh-frozen and formalin-fixed paraffin-embedded tumours. *Clinical Epigenetics* 2017;9(1)
 62. Thomas R. Pisanic, Leslie M. Cope, Shiou-Fu Lin, Ting-Tai Yen, Pornpat Athamanolap, Ryoichi Asaka, Kentaro Nakayama, Amanda N. Fader, Tza-Huei Wang, Je-Ming Shih, Tian-Li Wang, Methylomic Analysis of Ovarian Cancers Identifies Tumor-Specific Alterations Readily Detectable in Early Precursor Lesions. *Clinical Cancer Research* 2018;24(24):6536-47

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

