Am J Transl Res 2021;13(10):11597-11607
www.ajtr.org /ISSN:1943-8141/AJTR0135057

Original Article

Combined detection of stool-based methylation indicators for early screening of colorectal neoplasm

Xinyu Shao*, Huiyu Wang*, Yang Yu*, Chunli Zhou

*Department of Gastroenterology, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou, Jiangsu, China. *Equal contributors.*

Received May 9, 2021; Accepted August 13, 2021; Epub October 15, 2021; Published October 30, 2021

Abstract: In the past two decades, several methylated DNA targets, including gene promoters and other intronic markers have been explored in tumors and benign lesions. Therefore, it can be expected that a panel of stool-based biomarkers will become a screening method for colorectal cancer (CRC) and adenoma with better sensitivity and specificity, aiming to decrease the incidence and mortality of CRC. In this study, the methylation of secreted frizzled-related protein 1 (SFRP1), hyperplastic polyposis protein 1 (HPP1), α-internexin (INA), Wnt inhibitory factor 1 (WIF1), tissue factor pathway inhibitor 2 (TFPI2), ikaros family zinc finger protein 1 (IKZF1), and spastic paraplegia 20 (SPG20) were detected in stool samples from patients with CRC, adenoma, polyps, and healthy controls, respectively, and these biomarkers were used to establish a logistic regression model for classification. Receiver operating characteristic (ROC) curves were drawn to assess the importance of each biomarker. Subsequently, a biomarker or combination of biomarkers was analyzed for early screening of high-risk neoplasm. The data showed that when a single biomarker was used for CRC screening, the sensitivity ranged from 63.9% to 76.8%, the area under the curve (AUC) ranged from 0.821 to 0.875, and the accuracy ranged from 77.0% to 84.5%. Finally, the methylation of SFRP1, HPP1, TFPI2, and IKZF1 was selected using a backward stepwise method in the multivariate logistic analysis according to the Akaike Information Criterion. These findings indicate that stool DNA biomarkers have good diagnostic power in discriminating high-risk level of neoplasm from healthy population.

Keywords: DNA methylation, non-invasive diagnosis, colorectal neoplasms screening, stool based, combined application

Introduction

Colorectal cancer (CRC) is a common disease causing approximately million deaths each year, ranks the third in cancer incidence, and is the second leading cause of cancer-related death worldwide [1, 2]. Currently, fecal occult blood testing and colonoscopy are principal methods for CRC screening. However, these two tests have several shortcomings, including an uncomfortable feeling and troublesome bowel clearing and high costs [3, 4]. It has been shown that CRC is the result of accumulation of genetic and epigenetic changes, which can lead to a variation in fecal DNA. Tumorigenesis in CRC is associated with methylation of specific genes in cancer cells. In a recent study including 24 candidate genes, it was demonstrated that the regulation of the cell cycle, apoptosis, and angiogenesis are all related to hypomethylation and discrete hypermethylation at the promoter region of associated genes [5]. Therefore, the potential application of DNA methylation in the detection of CRC has attracted increased attention. In recent studies, it has been shown that screening of methylation is superior to a fecal occulted blood test in identifying the presence of high-grade dysplasia and sessile serrated polyps in one centimeter or larger, which can ultimately change the outcome of disease [6].

Despite extensive efforts and studies performed around the world, identifying molecular biomarkers of CRC remains a challenging, especially in the identification of predisposition, precancerous lesions, and early stage cancer [7]. With the development of human technology, many methods have been designed to detect DNA methylation in blood and

feces [8]. In fact, many studies were performed to identify a novel panel of methylation biomarkers, which are expected to provide highly sensitive and specific information for the diagnosis and prediction of CRC. A systematic review and quantitative evaluation demonstrated that methylation levels of secreted frizzled-related protein 2 (SFRP2), SFRP1, TFPI2, bone morphogenetic protein 3 (BMP3), N-Myc downstream-regulated gene 4 (NDRG4), SPG20, and BMP3 plus NDRG4 genes exceeded a sensitivity and a specificity of 80% for the CRC test. The diagnostic ratio (DOR) and AUC values of 7 candidate biomarkers indicate the good diagnostic and recognition ability between tumors and normal (healthy) tissues [9]. In 2016, the United States Preventive Services Task Force (USPSTF) included fecal DNA testing in published CRC screening guidelines [10]. Although DNA methylation biomarkers, such as ColoVantage®, Epi proColon® and ColoSure® have been approved by the Food and Drug Administration (FDA), they are not optimally used in clinical practice [11, 12].

Various genes were chosen as potential blood-based methylation markers for CRC with a sensitivity ranging from 34% to 90% and a specificity ranging from 69% to 100% [13, 14]. Previous studies have focused on collecting serum and plasma samples from patients to detect potential biomarkers of CRC, however, in recent years, several studies demonstrated that fecal samples could serve as a useful tool in screening CRC and precancerous lesions. Compared to colonoscopy, fecal DNA testing does not carry the risk of intestinal perforation and bleeding. Moreover, stool DNA is highly stable during collection, transportation, and storage [8, 15]. Hypermethylation promotes tumorigenesis by means of transcriptional silencing or downregulation of tumor suppressor genes, and over 600 hypermethylated candidate genes were successfully identified as associated with CRC, including hypermethylated adenomatous polyposis coli (APC), BMP3, SFRP2, GATA binding protein 4 (GATA4), glutathione S-transferase pi 1 (GSTP1), helicase like transcription factor (HLTF), mutl homolog 1 (MLH1), NDRG4, TFPI2, and WIF1 [13, 14].

Although many methylated DNA markers that are associated with CRC have been reported, few have been successfully used in clinical trials to detect CRC. In addition, the performance of commercially available epigenetic tests is poor. Therefore, it is necessary to identify better candidate biomarkers. Based on the literature, the following seven methylation markers were selected as candidate markers for the detection of CRC: SFRP1, HPP1, INA, WIF1, TFPI2, IKZF1, and SPG20. Various combinations were tried to achieve a better accuracy and a lower financial burden. Finally, a panel of stool-based biomarkers with a higher sensitivity and specificity served as a screening tool for CRC and adenoma to decrease the incidence and mortality of CRC.

Materials and methods

Samples and clinical data

All stool samples used in this study were collected at the Affiliated Suzhou Hospital of Nanjing Medical University (Suzhou, China). A total of 304 stool samples from 129 CRC patients, 65 colorectal adenoma patients, 63 patients with hyperplastic polyps, and 47 normal controls were recruited. All cancers, adenomas, and polyps were diagnosed based on histopathology. Tumor staging was determined according to the seventh edition of the American Joint Committee on Cancer (AJCC) staging system, which was revised in 2010. Participants were divided into two groups: CRC and colorectal adenoma patients were included in the high-risk group, while hyperplastic polyp, colonoscopies, and the negative control group were included in the low-risk group. Stool samples were collected before bowel preparation for colonoscopy. Clinical data of all study participants are listed in **Table 1**. All participants signed written informed consent before stool samples were collected. This study was approved by the Ethics Committee of the Affiliated Suzhou Hospital of Nanjing Medical University (IRB No. KL901070) (Suzhou, China). This study was conducted in compliance with the Declaration of Helsinki. All methods were performed in accordance with relevant guidelines and regulations.

Extraction of stool nucleic acid

Patients were instructed to collect the stool. From each patient, a total of 2-5 g of stool was collected using a disposable feces collection device, and immediately stored at -80°C.

Table 1. Clinical and pathological data of patients and negative controls

Variable	Sample Size	Percentage (%)
Age (years)		
≤65	142	46.7%
>65	162	53.3%
Gender		
Male	169	55.6%
Female	135	44.4%
Non-malignant tumor		
Negative colonoscopy	47	26.9%
Polyps	63	36.0%
Adenoma	65	37.1%
Cancer		
I stage	19	14.7%
II stage	42	32.6%
III stage	56	43.4%
IV stage	12	9.3%
Risk assessment		
Low risk	110	36.2%
High risk	194	63.8%

The high risk group includes cancer and adenoma, and the low risk group includes negative colonoscopy and polyps.

Before the experiment, stool samples to be tested were thawed at 2-8°C overnight. Sample tubes were removed and centrifuged at 5000 rpm for 10 minutes. The suspension was transferred to a new centrifuge tube, and DNA was extracted using a QIAamp Fast DNA Stool Mini Kit (QIAGEN). Investigators were blinded to clinical data and patient grouping.

Detection of DNA methylation

Stool samples were treated with a DNA methylation kit (Zymo Research) following the manufacturer's guidelines. MethyLight (Eads CA 2000) was selected to detect the methylation of SFRP1, HPP1, INA, WIF1, TFPI2, IKZF1, and SPG20 by singleplex PCR, and GAPDH served as an internal reference. The total volume of the real-time fluorescence quantitative PCR reaction system was 20 μL, consisting of 18 μL TaqMan Universal Master Mix II (no UNG), 1.0 μL of reaction mixture (primer, probe, RNase-Free ddH_2O) and 1.0 μL of bisulfte-treated DNA. For the results, the Ct value of GAPDH was be less than 35, thereby indicating the validity of the sample. A biomarker with a Ct value of less than 42 indicated that the methylation test result of the gene was positive.

Establishing a diagnostic combination of biomarkers

A minimal combination of biomarkers with high diagnostic potential was constructed by using the selected biomarkers. The input of DNA methylation values of all biomarkers was tested and recorded as high risk or low risk. By using the backward stepwise method according to the Akaike Information Criterion to analyze multiple combinations, we selected a panel of stool-based biomarkers with a higher sensitivity and specificity as screening methods for CRC and adenoma patients, so as to reduce the incidence and mortality of CRC in the population [15].

Statistical analysis

Clustering analysis is a type of non-constrained ranking analysis that reflects the similarity and difference in structure between samples. The Mann-Whitney U test was used to analyze the significance of clustering results according to human DNA quantity in stool samples. Single or multiple biomarkers were used to establish logistic regression models to distinguish between high-risk and low-risk groups. ROC curve analysis was used to assess the performance of classification models based on a single biomarker and multiple biomarkers. Because of the interactions between biomarkers, in the model building process, the backward stepwise method was applied to identify the optimal subset of biomarkers in the multivariate logistic analysis according to the Akaike Information Criterion (AIC). A nomogram including the best subset of biomarkers was used to predict malignant risk according to human DNA quantity in stool sample. Statistical analysis was performed using SPSS version 22.0, Prism 7, and R program. $P<0.05$ was considered statistically significant.

Results

Clinical and pathological information of patients and negative controls

A total of 304 stool samples from 129 CRC patients, 65 colorectal adenoma patients, 63 hyperplastic patients with polyps, and 47 nor-

mal controls were included in our study. Among them, patients over 65 years of age accounted for 53.3%, and patients ≤65 years of age accounted for 46.7%. There were 169 male patients and male negative controls, accounting for 55.6% of the total sample size. In the non-malignant tumor samples, there were a total of 47 negative control subjects, 63 patients with polyps and 65 patients with adenomas, accounting for 26.9%, 36.0%, and 37.1%, respectively. A total of 129 CRC samples were included in this study, of which staging TNM I, II, III, and IV accounted for 14.7%, 32.6%, 43.4%, and 3.9%, respectively. Moreover, the high-risk group included cancer and adenoma, and the low risk group included negative colonoscopy and polyps (**Table 1**).

A logistic regression model was established according to the Ct value of the methylation biomarkers. Samples were divided into two groups as follows: CRC and colorectal adenoma were included in the high-risk group, and polyps and a negative colonoscopy were included in the low-risk group. Next, principal coordinate analysis (PCoA) was performed to display changes in fecal methylation. The data indicated that the level of risk could drive changes of cluster in the fecal methylation of stool specimens (**Figure 1A**, $P<0.001$). Furthermore, we found that other variables, such as age could also affect the cluster result (**Figure 1B**, $P<0.001$). However, no significant differences were observed in the results of stool methylation according to gender or TNM stage (**Figure 1C**, **1D**, $P>0.05$).

The accuracy of methylation biomarkers for screening of colorectal neoplasms

Table 2 shows the performance of the 7 methylated biomarkers included in this study in screening of colorectal neoplasms, including sensitivity, specificity, and accuracy. Based on the sensitivity, the highest of all fecal markers was TFP12 (76.8%), and the lowest was SPG20 (63.9%). Based on the accuracy results, the best of all fecal biomarkers was TFP12 (84.5%), whereas the worst was SPG20 (77.0%). Further analysis showed that the sensitivity of included methylation biomarkers in CRC was similar, among which the sensitivity of SPG20 (86.1%) was the highest. However, the sensitivity of included methylation biomarkers in adenoma tissues varied greatly, among which IKZF1 (64.4%) was the most sensitive and SPG20 (20.0%) the least sensitive. In the analysis of specificity, the differences of methylation factors were small. Therefore, the accuracy of the final screening results may be greatly influenced by the sensitivity of the indicators of the adenoma tissue.

Subsequently, ROC curve analysis was performed for each single biomarker (**Figure 2A-G**), and the data revealed that the AUC area with the highest was TFP12 (AUC = 0.875) and the lowest was SPG20 (AUC = 0.820). These results were consistent with the accuracy of all indicators. Considering the relatively high proportion of CRC in the included high risk neoplasm samples, the AUC and accuracy of our methylation index were superior, thereby indicating a good diagnostic power in discriminating the high-risk level of neoplasm from normal tissues.

ROC curves of all the possible combinations of the 7 methylation biomarkers.

To improve the accuracy of methylation detection, the combination of multiple indicators is a feasible option. However, to consider increasing the accuracy of detection and decreasing the inspection cost to the greatest extent, it is important to find the best matching combination. In this study, for the 7 included methylation indicators, we calculated the AUC areas from all possible combinations for comparison (Supplementary Figure S1A-G and Supplementary Table 1). The results suggested that the AUC area elevated with the increase in the number of included methylation indicators. Furthermore, when more than 4 indicators were included, the AUC area increased less (Supplementary Figure 1H).

The performance of a novel stool-based methylation biomarker panel for the screening of colorectal neoplasms

To further clarify the advantages of combination detection, a logistic regression model was built with multiple biomarkers to predict the high risk of neoplasm. Due to the interactions between biomarkers, in the model building process, the backward stepwise method was applied to identify the optimal subset of biomarkers in the multivariate logistic analysis

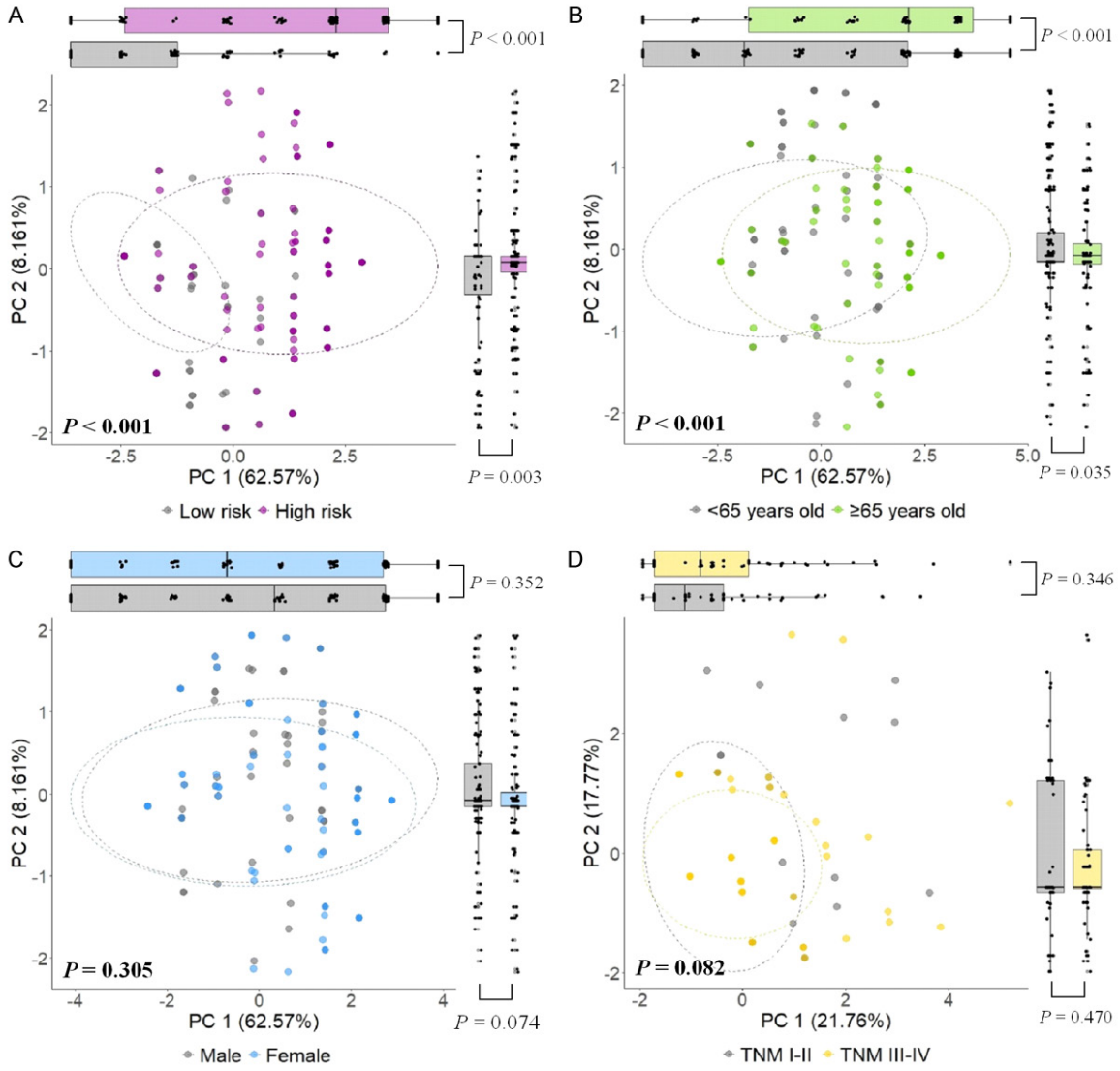


Figure 1. Principal coordinate analysis of multidimensional methylation data was performed to display changes in fecal methylation according to major variables retrieved to level of risk (A) age (B) gender (C) and stage TNM (D). The x- and y-axes represent the two most informative principal coordinates (PCs) of the principal coordinate analysis (PCoA), and marginal boxplots describe the distribution of those values for the different groups. Color legends represent the respective variables. The paired Wilcoxon rank-sum test was used to compare the PC1 or PC2 values between groups, and *p* values are displayed beside the edge box. The outcomes of the permutation-based test (PERMANOVA) of different indicators between samples are shown in the lower left corner of the image.

Table 2. Sensitivity, specificity, and accuracy of methylation biomarkers

Biomarker	Sensitivity (adenoma = 65)	Sensitivity (cancer = 129)	Sensitivity (high risk = 194)	Specificity (low risk = 110)	Accuracy (total = 304)
SFRP1	27 (41.5%)	108 (83.7%)	135 (69.6%)	104 (94.6%)	239 (78.6%)
HPP1	33 (50.8%)	103 (79.8%)	136 (70.1%)	110 (100.0%)	246 (80.9%)
INA	22 (33.9%)	109 (84.5%)	131 (67.5%)	108 (98.2%)	239 (78.6%)
WIF1	30 (46.2%)	106 (82.2%)	136 (70.1%)	110 (100.0%)	246 (80.9%)
TFP12	39 (60.0%)	110 (85.3%)	149 (76.8%)	108 (98.2%)	257 (84.5%)
IKZF1	42 (64.6%)	106 (82.2%)	148 (76.3%)	108 (98.2%)	256 (84.2%)
SPG20	13 (20.0%)	111 (86.1%)	124 (63.9%)	110 (100.0%)	234 (77.0%)

The high risk group includes cancer and adenoma, and the low risk group includes negative colonoscopy and polyps.

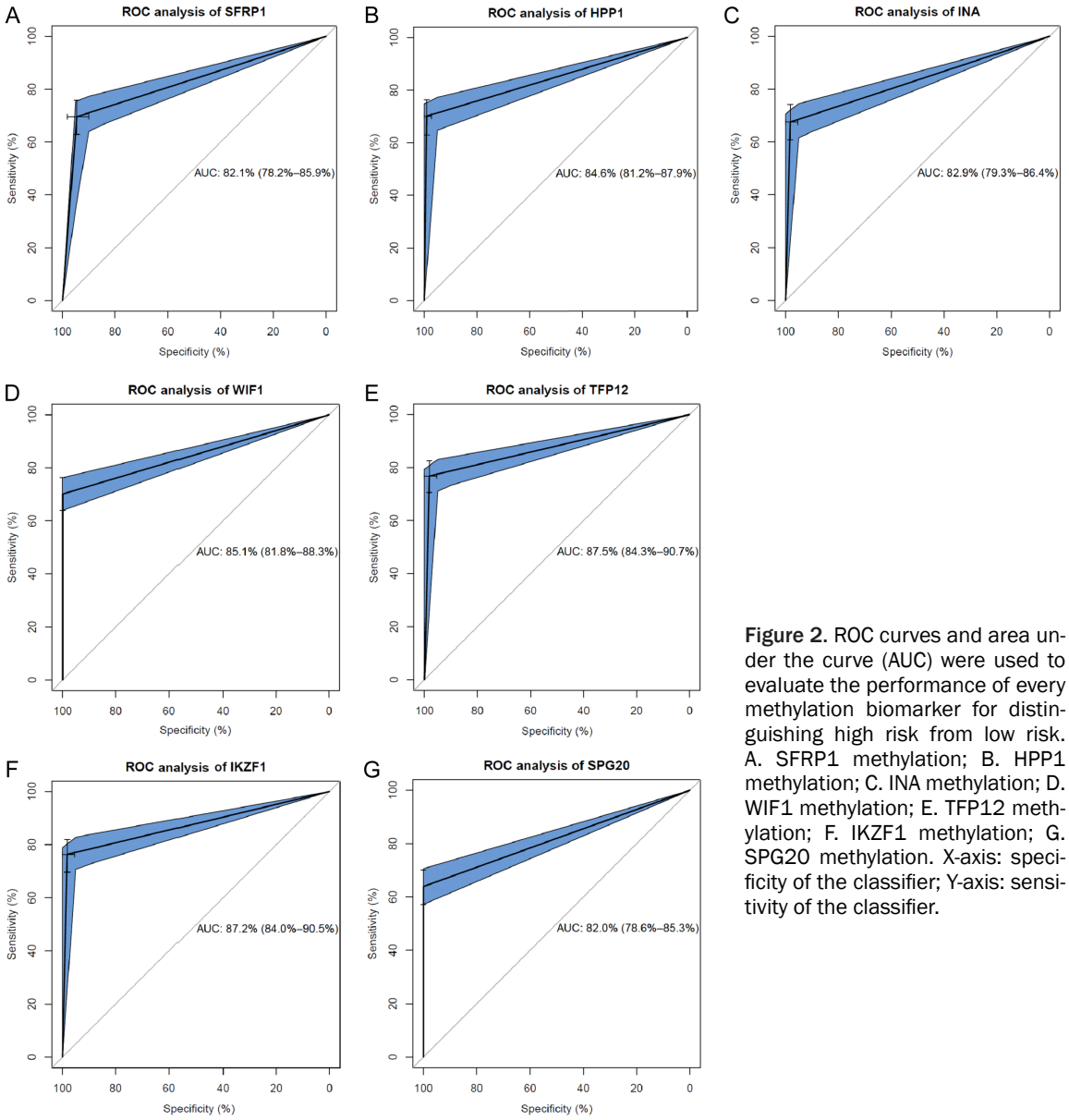


Figure 2. ROC curves and area under the curve (AUC) were used to evaluate the performance of every methylation biomarker for distinguishing high risk from low risk. A. SFRP1 methylation; B. HPP1 methylation; C. INA methylation; D. WIF1 methylation; E. TFP12 methylation; F. IKZF1 methylation; G. SPG20 methylation. X-axis: specificity of the classifier; Y-axis: sensitivity of the classifier.

according to the AIC. Finally, SFRP1, HPP1, TFP12, and IKZF1 methylation biomarkers were selected using the backward stepwise method in the multivariate logistic analysis. These four biomarkers were identified as the best subsets and were used to develop the risk nomogram (**Figure 3A**). In the nomogram, all four methylation biomarkers had obvious effects, among which positive HPP1 and IKZF1 methylation indicators had the greatest effect on the diagnosis of high risk of a neoplasm in patients.

Next, the fitting of ROC curve of the selected four methylation biomarkers was performed, and the AUC value (AUC = 0.982, 95% CI: 0.968-0.997) was significantly superior compared to a single methylation biomarker (**Figure 3B**). The internal calibration plot in the optimal logistic model showed a close distance between the fit line and the diagonal line, implying optimal consistency between the nomogram-predicted probability and the actual observation (**Figure 3C**).

Discussion

Because most CRC patients have no characteristic clinical symptoms in the early stage, at

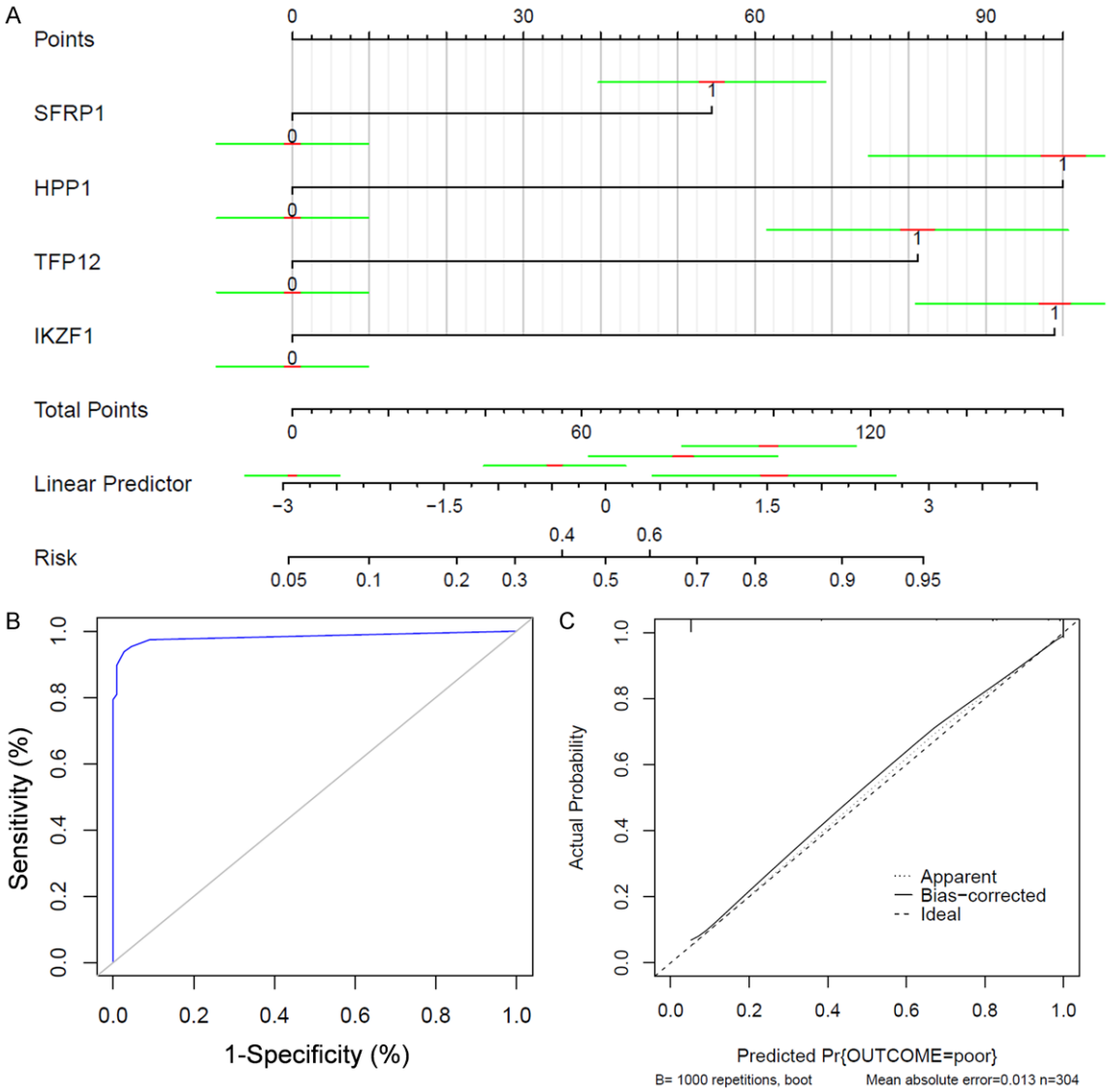


Figure 3. The novel panel comprised of stool DNA biomarkers selected according to the Akaike Information Criterion. A. Risk nomogram based on the detection of SFRP1, HPP1, TFP12, and IKZF1 methylation which distinguishes high malignant risk from low malignant risk. B. The fitting of ROC curve plotted according to the detection of SFRP1, HPP1, TFP12, and IKZF1 methylation. C. The calibration plot of risk nomogram. For an individual patient, the nomogram-predicted probability of malignant risk is plotted on the x-axis and the actual probability is plotted on the y-axis.

present, more than half of all CRC patients are clinically diagnosed at an advanced stage. Research statistics have shown that patients with stage I CRC have a 5-year survival rate of more than 90%, while that of stage IV CRC is less than 10%, indicating that early diagnosis of CRC is very important [16, 17]. To lower the incidence and mortality of CRC, many countries have implemented national CRC screening programs. Currently, the most commonly used methods for screening CRC clinically include a fecal occult blood test and endoscopy. Fecal occult blood is most used in clinical practice because of low cost and non-invasive nature; however, its detection accuracy is not optimal. Although colonoscopy is the gold standard for the final clinical diagnosis of CRC, it is not an ideal primary screening method because of significant discomfort to patients and the risk of bleeding and perforation during operation [15, 18]. Therefore, there is a need for an accurate and non-invasive method to screen average-risk populations. With advances in human technology, a DNA methylation-based test is expect-

ed to be a highly cost-effective approach in clinical practice. In recent years, an increased number of studies have revealed that compared with colonoscopy and the circulating DNA test, fecal DNA detection has the following advantages: simple operation, stable sample transportation, good safety profile, and low cost [19].

Epigenetic changes are an indispensable part of CRC pathogenesis. As an epigenetic marker in CRC, DNA methylation has attracted extensive attention around the world. In humans, the definition of DNA methylation is the addition of a methyl moiety from active methylene compounds to the 5' position of cytosine residues in CpG dinucleotides [20]. In 2008, the American Gastroenterological Association made a statement on the goal of CRC screening, whether through colonoscopy or fecal occult blood tests, the purpose was to reduce the mortality by decreasing the incidence of advanced disease. To achieve this goal, an ideal screening tool should be able to detect precancerous lesions, such as adenomas, with significant accuracy [21, 22].

The tumor suppressor protein serves as an extracellular Wnt inhibitor and is encoded by the SFRP1 gene. A lack of SFRP1 was found in many cancers. One study analyzed gene expression, promoter methylation, and survival data in more than 8,000 tumors and normal samples from 29 types of cancer, and it was found that SFRP1 was consistently associated with tumor suppressive function, and most studies using SFRP1 had a satisfactory differential performance with different levels of sensitivity and specificity [23]. DNA methylation of SFRP1, SFRP2, syndecan 2 (SDC2), and proline rich membrane anchor 1 (PRIMA1) gene promoters in 121 plasma and 32 biopsy samples was tested by methyLight polymerase chain reaction (PCR). It was found that the methylation of SFRP1, SFRP2, SDC2, and PRIMA1 promoter sequences in plasma samples of CRC patients was 85.1%, 72.3%, 89.4%, and 80.9%, respectively, and in plasma samples of adenoma patients, it was 89.2%, 83.8%, 81.1%, and 70.3%, respectively [24]. In comparison with normal intestinal mucosa, IKZF1 has been demonstrated to be down-regulated in CRC tissue. Moreover, promotor hypermethylation was considered to mediate the down-regulation of this gene, which revealed its suitability for further evaluation as a biomarker for CRC diagnosis [25]. A pilot study that focused on examining branched chain amino acid transaminase 1 (BCAT1) and IKZF1 for the detection of CRC was reported with an estimated sensitivity of 77% and specificity of 92.4%. Moreover, in a study that examined the methylation of BCAT1 and IKZF1 in the blood of CRC patients reported an estimated sensitivity of 77% and a specificity of 92.4%. In addition, it was found that methylated DNA levels increased with CRC progression. IKZF1 is considered a key factor in regulating cell proliferation and differentiation [26]. WIF1 is a type of secreted protein, the role of WIF1 involves the binding to Wnt proteins and inhibiting their activities. Hypermethylation of WIF1 has long been shown to release inhibition of the Wnt/β-catenin signaling pathway [27]. In two recent studies, it has been shown that WIF1 and neuropeptide Y (NPY) genes have a significantly higher level of hypermethylation in tumor tissue than that in healthy tissue. Changes in the concentration of circulating tumor DNA can be used not only to monitor the progression of the tumor but also serve as a marker for the evaluation of the therapeutic efficacy of patients with metastatic CRC [28, 29]. As for the role of TFPI2 in tumors, in previous studies, it was found that TFPI2, which encodes a broad-spectrum serine protease inhibitor, can protect the extracellular matrix of cancer cells from degradation and inhibit *in vitro* proliferation and colony formation, which plays an important role in tumor invasion and metastasis [30, 31]. In addition, the HPP1 gene encodes a transmembrane protein and frequent methylation is observed in primary CRC and liver metastases [32]. It is a tumor suppressor gene and was known to be downregulated by promoter hypermethylation in various tumor types, including those of the colorectum, esophagus, stomach, and gallbladder [33]. SPG20 encodes Spartin, a widely-expressed protein with an unknown function, which has been found to mediate intracellular epidermal growth factor receptor or trafficking [34]. Moreover, proteins encoded by the INA gene have been shown to regulate cell-cell interactions, which are essential for tumor growth and invasion. In a study involving 523 human samples, it was shown that the methylation rate of the INA gene was 66% in CRC patients and 42% in colorectal adenomas, but

no methylation was observed in normal mucosa [35]. The purpose of this study was to find promising methylation markers for CRC-based detection, and combine the use of multiple indicators to improve the detection efficiency and reduce the detection cost.

In this study, the methylations of SFRP1, HPP1, INA, WIF1, TFPI2, IKZF1, and SPG20 genes in stool samples were detected in CRC, adenoma, polyp patients and normal controls, and the ability of single or multiple biomarkers in early screening for high-risk level of neoplasm was evaluated. When a single biomarker was used for CRC screening, the sensitivity ranged from 63.9% to 76.8%, the AUC ranged from 0.821 to 0.875, and the accuracy ranged from 77.0% to 84.5%, thereby indicating a good diagnostic power in discriminating high-risk level of neoplasm from normal tissues. In a systematic meta-analysis, which consisted of 4867 individuals, the sensitivity and specificity of single-gene stool DNA methylation analysis for CRC were 56.5% and 93.2% respectively, and 32.6% and 93.2% for adenomas [36]. Moreover, the sensitivity of the methylation indicators included in this study was significantly superior to the previously reported results of the meta-analysis. Taken together, these findings indicated that DNA methylation could serve as a non-invasive biomarker for the clinical diagnosis of CRC and adenoma.

To further improve the accuracy of methylated biomarkers in CRC screening, a variety of methods can be used. Multi-target DNA assays using different methylated gene combinations have been reported to improve the sensitivity of the assay [37]. For example, in a double-blind, multicenter, and case-control study, Ahlquist et al. developed Cologuard for the combination detection of NRDG4, BMP3, vimentin (VIM), and TFPI2, which provided important stool-based trials for CRC screening [38]. The introduction of multigene fecal methylation panels may lead to better results. However, how to match the methylation indicators to maximize both the accuracy of the test and the cost of detection is challenging. In this study, the AUC of a panel of biomarkers was 0.888 and 0.993 for the ROC curve, which was much better than the single stool-based biomarker. Thus, the results suggested that the AUC area was elevated with the increase of the number of included methylation indicators. When more than 4 indicators were included, the increase of the AUC area was less than before.

To investigate the efficacy of a novel stool-based methylation biomarker panel for screening colorectal tumors, we established a logistic regression model with multiple biomarkers according to the AIC. SFRP1, HPP1, TFP12, and IKZF1 methylation biomarkers were selected using a backward stepwise method in the multivariate logistic analysis. These four biomarkers were identified as the best subsets. The fitting ROC curve of the four selected methylation biomarkers was significantly superior to the single methylation biomarker and the internal calibration plot in the optimal logistic model showed close distances between the fit line and the diagonal line.

In summary, although promising, the use of DNA methylation markers has not yet been recommended for wide use in the clinic. Identifying novel, more effective methylation markers for colon cancer remains an exciting goal. A combination of critical methylation markers will promote a new wave of clinical trials. The findings of this study further confirmed the use of a novel panel of stool-based biomarkers, including SFRP1, HPP1, TFP12, and IKZF1 in CRC and adenoma screening, which showed a higher sensitivity. Many studies have confirmed altered methylation levels in the DNA extracted from stool samples. However, due to the limited sample size, the findings of these studies need to be further validated. Therefore, continued research on the mechanisms of DNA methylation is critical.

Acknowledgements

This work was supported by Suzhou Special Project of Diagnosis and Treatment for key Clinical Disease (LCZX201814), Suzhou Science and Technology Development Project (SYSD2018144), Suzhou Integrated Traditional Chinese and Western Medicine Scientific Research Fund Project (SYSD2018208).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Chunli Zhou, Department of Gastroenterology, The Affiliated Suzhou

Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou, Jiangsu, China. Tel: +86-13962124345; E-mail: zhouchunli079@163.com

References

[1] Li K, Zeng L, Wei H, Hu J, Jiao L, Zhang J and Xiong Y. Identification of gene-specific DNA methylation signature for colorectal cancer. Cancer Genet 2018; 228-229: 5-11.

[2] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and 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: 394-424.

[3] Freitas M, Ferreira F, Carvalho S, Silva F, Lopes P, Antunes L, Salta S, Diniz F, Santos LL, Videira JF, Henrique R and Jeronimo C. A novel DNA methylation panel accurately detects colorectal cancer independently of molecular pathway. J Transl Med 2018; 16: 45.

[4] Imperiale TF and Ransohoff DF. Screening for colorectal neoplasia. N Engl J Med 2017; 376: 1598-1599.

[5] Silva TD, Vidigal VM, Felipe AV, DE Lima JM, Neto RA, Saad SS and Forones NM. DNA methylation as an epigenetic biomarker in colorectal cancer. Oncol Lett 2013; 6: 1687-1692.

[6] Robertson DJ and Imperiale TF. Stool testing for colorectal cancer screening. Gastroenterology 2015; 149: 1286-1293.

[7] Kel A, Boyarskikh U, Stegmaier P, Leskov LS, Sokolov AV, Yevshin I, Mandrik N, Stelmashenko D, Koschmann J, Kel-Margoulis O, Krull M, Martinez-Cardus A, Moran S, Esteller M, Kolpakov F, Filipenko M and Wingender E. Walking pathways with positive feedback loops reveal DNA methylation biomarkers of colorectal cancer. BMC Bioinformatics 2019; 20 Suppl 4: 119.

[8] Osborn NK and Ahlquist DA. Stool screening for colorectal cancer: molecular approaches. Gastroenterology 2005; 128: 192-206.

[9] Liu R, Su X, Long Y, Zhou D, Zhang X, Ye Z, Ma J, Tang T, Wang F and He C. A systematic review and quantitative assessment of methylation biomarkers in fecal DNA and colorectal cancer and its precursor, colorectal adenoma. Mutat Res Rev Mutat Res 2019; 779: 45-57.

[10] US Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Epling JW Jr, García FAR, Gillman MW, Harper DM, Kemper AR, Krist AH, Kurth AE, Landefeld CS, Mangione CM, Owens DK, Phillips WR, Phipps MG, Pignone MP and Siu AL. Screening for colorectal cancer: US preventive services task force recommendation statement. JAMA 2021; 315: 2564-2575.

[11] Warren JD, Xiong W, Bunker AM, Vaughn CP, Furtado LV, Roberts WL, Fang JC, Samowitz WS and Heichman KA. Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer. BMC Med 2011; 9: 133.

[12] Toth K, Sipos F, Kalmar A, Patai AV, Wichmann B, Stoehr R, Golcher H, Schellerer V, Tulassay Z and Molnar B. Detection of methylated SEPT9 in plasma is a reliable screening method for both left- and right-sided colon cancers. PLoS One 2012; 7: e46000.

[13] Ng JM and Yu J. Promoter hypermethylation of tumour suppressor genes as potential biomarkers in colorectal cancer. Int J Mol Sci 2015; 16: 2472-2496.

[14] Okugawa Y, Grady WM and Goel A. Epigenetic alterations in colorectal cancer: emerging biomarkers. Gastroenterology 2015; 149: 1204-1225, e12.

[15] Sun M, Liu J, Hu H, Guo P, Shan Z, Yang H, Wang J, Xiao W and Zhou X. A novel panel of stool-based DNA biomarkers for early screening of colorectal neoplasms in a Chinese population. J Cancer Res Clin Oncol 2019; 145: 2423-2432.

[16] Qian CN. At-home cancer screening: a solution for China and other developing countries with a large population and limited number of healthcare practitioners. Chin J Cancer 2017; 36: 68.

[17] Liu X, Wen J, Li C, Wang H, Wang J and Zou H. High-yield methylation markers for stool-based detection of colorectal cancer. Dig Dis Sci 2020; 65: 1710-1719.

[18] Rabeneck L, Paszat LF, Hilsden RJ, Saskin R, Leddin D, Grunfeld E, Wai E, Goldwasser M, Sutradhar R and Stukel TA. Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. Gastroenterology 2008; 135: 1899-1906, 1906, e1.

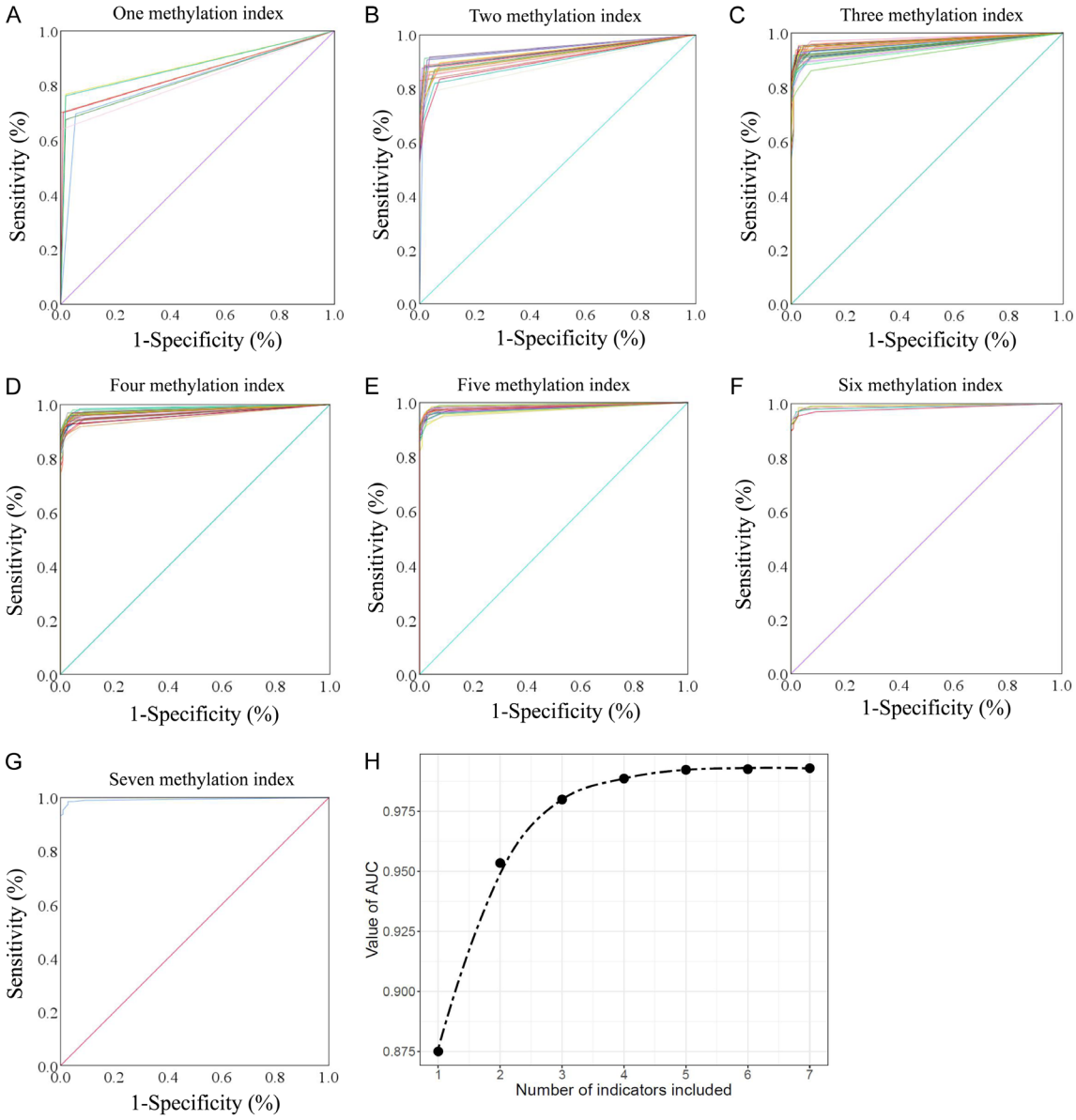
[19] Bailey JR, Aggarwal A and Imperiale TF. Colorectal cancer screening: stool DNA and other noninvasive modalities. Gut Liver 2016; 10: 204-211.

[20] Zamani M, Hosseini SV and Mokarram P. Epigenetic biomarkers in colorectal cancer: premises and prospects. Biomarkers 2018; 23: 105-114.

[21] Dominic OG, McGarrity T, Dignan M and Lengerich EJ. American College of Gastroenterology Guidelines for colorectal cancer screening 2008. Am J Gastroenterol 2009; 104: 2626-2627; author reply 2628-2629.

[22] Carethers JM and Jung BH. Genetics and genetic biomarkers in sporadic colorectal cancer. Gastroenterology 2015; 149: 1177-1190.

[23] Vincent KM and Postovit LM. A pan-cancer analysis of secreted Frizzled-related proteins: re-examining their proposed tumour suppressive function. Sci Rep 2017; 7: 42719.
[24] Bartak BK, Kalmar A, Peterfia B, Patai AV, Galamb O, Valcz G, Spisak S, Wichmann B, Nagy ZB, Toth K, Tulassay Z, Igaz P and Molnar B. Colorectal adenoma and cancer detection based on altered methylation pattern of SFRP1, SFRP2, SDC2, and PRIMA1 in plasma samples. Epigenetics 2017; 12: 751-763.
[25] Mitchell SM, Ross JP, Drew HR, Ho T, Brown GS, Saunders NF, Duesing KR, Buckley MJ, Dunne R, Beetson I, Rand KN, McEvoy A, Thomas ML, Baker RT, Wattchow DA, Young GP, Lockett TJ, Pedersen SK, Lapointe LC and Molloy PL. A panel of genes methylated with high frequency in colorectal cancer. BMC Cancer 2014; 14: 54.
[26] Pedersen SK, Baker RT, McEvoy A, Murray DH, Thomas M, Molloy PL, Mitchell S, Lockett T, Young GP and LaPointe LC. A two-gene blood test for methylated DNA sensitive for colorectal cancer. PLoS One 2015; 10: e0125041.
[27] Hsieh JC, Kodjabachian L, Rebbert ML, Rattner A, Smallwood PM, Samos CH, Nusse R, Dawid IB and Nathans J. A new secreted protein that binds to Wnt proteins and inhibits their activities. Nature 1999; 398: 431-436.
[28] Garrigou S, Perkins G, Garlan F, Normand C, Didelot A, Le Corre D, Peyvandi S, Mulot C, Niarra R, Aucouturier P, Chatellier G, Nizard P, Perez-Toralla K, Zonta E, Charpy C, Pujals A, Barau C, Bouche O, Emile JF, Pezet D, Bibeau F, Hutchison JB, Link DR, Zaanan A, Laurent-Puig P, Sobhani I and Taly V. A study of hypermethylated circulating tumor DNA as a universal colorectal cancer biomarker. Clin Chem 2016; 62: 1129-1139.
[29] Garlan F, Laurent-Puig P, Sefrioui D, Siauve N, Didelot A, Sarafan-Vasseur N, Michel P, Perkins G, Mulot C, Blons H, Taieb J, Di Fiore F, Taly V and Zaanan A. Early evaluation of circulating tumor DNA as marker of therapeutic efficacy in metastatic colorectal cancer patients (PLACOL study). Clin Cancer Res 2017; 23: 5416-5425.
[30] Sato N, Parker AR, Fukushima N, Miyagi Y, Iacobuzio-Donahue CA, Eshleman JR and Goggins M. Epigenetic inactivation of TFPI-2 as a common mechanism associated with growth and invasion of pancreatic ductal adenocarcinoma. Oncogene 2005; 24: 850-858.
[31] Wong CM, Ng YL, Lee JM, Wong CC, Cheung OF, Chan CY, Tung EK, Ching YP and Ng IO. Tissue factor pathway inhibitor-2 as a frequently silenced tumor suppressor gene in hepatocellular carcinoma. Hepatology 2007; 45: 1129-1138.
[32] Ebert MP, Mooney SH, Tonnes-Priddy L, Lograsso J, Hoffmann J, Chen J, Rocken C, Schulz HU, Malfertheiner P and Lofton-Day C. Hypermethylation of the TPEF/HPP1 gene in primary and metastatic colorectal cancers. Neoplasia 2005; 7: 771-778.
[33] Wang J, Elahi A, Ajidahun A, Clark W, Hernandez J, Achille A, Hao JH, Seto E and Shibata D. The interplay between histone deacetylases and c-Myc in the transcriptional suppression of HPP1 in colon cancer. Cancer Biol Ther 2014; 15: 1198-1207.
[34] Bakowska JC, Jupille H, Fatheddin P, Puertollano R and Blackstone C. Troyer syndrome protein spartin is mono-ubiquitinated and functions in EGF receptor trafficking. Mol Biol Cell 2007; 18: 1683-1692.
[35] Lind GE, Danielsen SA, Ahlquist T, Merok MA, Andresen K, Skotheim RI, Hektoen M, Rognum TO, Meling GI, Hoff G, Bretthauer M, Thiis-Evensen E, Nesbakken A and Lothe RA. Identification of an epigenetic biomarker panel with high sensitivity and specificity for colorectal cancer and adenomas. Mol Cancer 2011; 10: 85.
[36] Mojtabanezhad Shariatpanahi A, Yassi M, Nouraie M, Sahebkar A, Varshoee Tabrizi F and Kerachian MA. The importance of stool DNA methylation in colorectal cancer diagnosis: a meta-analysis. PLoS One 2018; 13: e0200735.
[37] Xue M, Lai SC, Xu ZP and Wang LJ. Noninvasive DNA methylation biomarkers in colorectal cancer: a systematic review. J Dig Dis 2015; 16: 699-712.
[38] Ahlquist DA, Zou H, Domanico M, Mahoney DW, Yab TC, Taylor WR, Butz ML, Thibodeau SN, Rabeneck L, Paszat LF, Kinzler KW, Vogelstein B, Bjerregaard NC, Laurberg S, Sorensen HT, Berger BM and Lidgard GP. Next-generation stool DNA test accurately detects colorectal cancer and large adenomas. Gastroenterology 2012; 142: 248-256; quiz e225-246.



Supplementary Figure 1. The ROC curves and AUC were used to investigate all the possible combinations of the 7 methylation biomarkers. A. ROC curve of one methylation biomarker alone. B. ROC curve of two methylation biomarkers combination. C. ROC curve of three methylation biomarkers combination. D. ROC curve of four methylation biomarkers combination. E. ROC curve of five methylation biomarkers combination. F. ROC curve of six methylation biomarkers combination. G. ROC curve of seven methylation biomarkers combination. H. The number of methylation biomarkers included in the combination analysis and the relevant highest AUC value.

Supplementary Table 1. Methylation index detection and multi-index fitting evaluation

Included index number	Multi methylation index	*P* value	AUC	95% CI
1	a	<0.001	0.821	0.773-0.869
	b	<0.001	0.846	0.803-0.889
	c	<0.001	0.829	0.783-0.874
	d	<0.001	0.851	0.808-0.893
	e	<0.001	0.875	0.835-0.915
	f	<0.001	0.872	0.832-0.913
	g	<0.001	0.820	0.773-0.866
2	a, b	<0.001	0.920	0.888-0.953
	a, c	<0.001	0.903	0.869-0.937
	a, d	<0.001	0.927	0.898-0.957
	a, e	<0.001	0.930	0.900-0.959
	a, f	<0.001	0.937	0.909-0.965
	a, g	<0.001	0.900	0.865-0.935
	b, c	<0.001	0.920	0.888-0.951
	b, d	<0.001	0.937	0.908-0.965
	b, e	<0.001	0.935	0.907-0.964
	b, f	<0.001	0.948	0.922-0.974
	b, g	<0.001	0.921	0.890-0.953
	c, d	<0.001	0.923	0.891-0.954
	c, e	<0.001	0.924	0.893-0.954
	c, f	<0.001	0.951	0.926-0.976
	c, g	<0.001	0.888	0.851-0.925
	d, e	<0.001	0.938	0.910-0.966
	d, f	<0.001	0.953	0.929-0.978
	d, g	<0.001	0.915	0.882-0.948
	e, f	<0.001	0.949	0.924-0.974
	e, g	<0.001	0.927	0.897-0.957
	f, g	<0.001	0.940	0.913-0.968
3	a, b, c	<0.001	0.947	0.921-0.973
	a, b, d	<0.001	0.962	0.940-0.984
	a, b, e	<0.001	0.958	0.936-0.981
	a, b, f	<0.001	0.972	0.953-0.990
	a, b, g	<0.001	0.945	0.919-0.971
	a, c, d	<0.001	0.947	0.922-0.973
	a, c, e	<0.001	0.952	0.928-0.976
	a, c, f	<0.001	0.965	0.945-0.986
	a, c, g	<0.001	0.921	0.890-0.953
	a, d, e	<0.001	0.961	0.939-0.983
	a, d, f	<0.001	0.980	0.965-0.995
	a, d, g	<0.001	0.946	0.920-0.972
	a, e, f	<0.001	0.968	0.948-0.987
	a, e, g	<0.001	0.950	0.925-0.975
	a, f, g	<0.001	0.963	0.942-0.985
	b, c, d	<0.001	0.954	0.930-0.978
	b, c, e	<0.001	0.952	0.927-0.976

3	b, c, f	<0.001	0.970	0.951-0.989
	b, c, g	<0.001	0.943	0.916-0.970
	b, d, e	<0.001	0.964	0.943-0.986
	b, d, f	<0.001	0.975	0.957-0.992
	b, d, g	<0.001	0.958	0.935-0.981
	b, e, f	<0.001	0.973	0.954-0.991
	b, e, g	<0.001	0.956	0.933-0.980
	b, f, g	<0.001	0.971	0.953-0.990
	c, d, e	<0.001	0.951	0.926-0.976
	c, d, f	<0.001	0.971	0.953-0.990
	c, d, g	<0.001	0.942	0.915-0.969
	c, e, f	<0.001	0.970	0.950-0.989
	c, e, g	<0.001	0.937	0.909-0.965
	c, f, g	<0.001	0.960	0.938-0.982
	d, e, f	<0.001	0.972	0.953-0.991
	d, e, g	<0.001	0.957	0.934-0.981
	d, f, g	<0.001	0.968	0.947-0.988
	e, f, g	<0.001	0.971	0.952-0.990
4	a, b, c, d	<0.001	0.970	0.950-0.989
	a, b, c, e	<0.001	0.967	0.948-0.987
	a, b, c, f	<0.001	0.979	0.964-0.995
	a, b, c, g	<0.001	0.958	0.935-0.981
	a, b, d, e	<0.001	0.975	0.958-0.993
	a, b, d, f	<0.001	0.989	0.978-1.000
	a, b, d, g	<0.001	0.971	0.952-0.990
	a, b, e, f	<0.001	0.982	0.968-0.997
	a, b, e, g	<0.001	0.967	0.947-0.987
	a, b, f, g	<0.001	0.980	0.965-0.995
	a, c, d, e	<0.001	0.966	0.946-0.987
	a, c, d, f	<0.001	0.984	0.970-0.998
	a, c, d, g	<0.001	0.954	0.930-0.978
	a, c, e, f	<0.001	0.977	0.961-0.994
	a, c, e, g	<0.001	0.957	0.934-0.980
	a, c, f, g	<0.001	0.969	0.950-0.989
	a, d, e, f	<0.001	0.984	0.970-0.998
	a, d, e, g	<0.001	0.968	0.948-0.988
	a, d, f, g	<0.001	0.986	0.974-0.999
	a, e, f, g	<0.001	0.978	0.962-0.994
	b, c, d, e	<0.001	0.969	0.950-0.989
	b, c, d, f	<0.001	0.980	0.964-0.996
	b, c, d, g	<0.001	0.968	0.947-0.988
	b, c, e, f	<0.001	0.981	0.966-0.996
	b, c, e, g	<0.001	0.961	0.939-0.983
	b, c, f, g	<0.001	0.977	0.960-0.994
	b, d, e, f	<0.001	0.983	0.968-0.997
	b, d, e, g	<0.001	0.975	0.958-0.993
	b, d, f, g	<0.001	0.983	0.968-0.998
	b, e, f, g	<0.001	0.987	0.975-0.999
	c, d, e, f	<0.001	0.977	0.960-0.994

4	c, d, e, g	<0.001	0.962	0.940-0.984
	c, d, f, g	<0.001	0.977	0.961-0.994
	c, e, f, g	<0.001	0.976	0.959-0.993
	d, e, f, g	<0.001	0.980	0.964-0.996
5	a, b, c, d, e	<0.001	0.978	0.962-0.994
	a, b, c, d, f	<0.001	0.989	0.979-1.000
	a, b, c, d, g	<0.001	0.976	0.959-0.993
	a, b, c, e, f	<0.001	0.986	0.973-0.999
	a, b, c, e, g	<0.001	0.972	0.953-0.990
	a, b, c, f, g	<0.001	0.983	0.969-0.997
	a, b, d, e, f	<0.001	0.990	0.979-1.000
	a, b, d, e, g	<0.001	0.979	0.963-0.995
	a, b, d, f, g	<0.001	0.992	0.983-1.000
	a, b, e, f, g	<0.001	0.989	0.978-1.000
	a, c, d, e, f	<0.001	0.985	0.971-0.999
	a, c, d, e, g	<0.001	0.970	0.951-0.990
	a, c, d, f, g	<0.001	0.987	0.975-1.000
	a, c, e, f, g	<0.001	0.981	0.966-0.996
	a, d, e, f, g	<0.001	0.988	0.975-1.000
	b, c, d, e, f	<0.001	0.985	0.971-0.999
	b, c, d, e, g	<0.001	0.978	0.961-0.995
	b, c, d, f, g	<0.001	0.986	0.972-0.999
	b, c, e, f, g	<0.001	0.987	0.975-1.000
	b, d, e, f, g	<0.001	0.991	0.980-1.000
	c, d, e, f, g	<0.001	0.983	0.968-0.998
6	a, b, c, d, e, f	<0.001	0.990	0.979-1.000
	a, b, c, d, e, g	<0.001	0.982	0.967-0.997
	a, b, c, d, f, g	<0.001	0.993	0.984-1.000
	a, b, c, e, f, g	<0.001	0.989	0.979-1.000
	a, b, d, e, f, g	<0.001	0.993	0.984-1.000
	a, c, d, e, f, g	<0.001	0.988	0.975-1.000
	b, c, d, e, f, g	<0.001	0.991	0.980-1.000
7	a, b, c, d, e, f, g	<0.001	0.993	0.984-1.000

The letter of a, b, c, d, e, f, g represents the related methylation index of SFRP1, HPP1, INA, WIF1, TFPI2, IKZF1 and SPG20, respectively.