Original Article

Development and preliminary validation of plasma cell-free DNA methylation-based diagnostic prediction model for colorectal cancer detection

Ying Qu[1,2#], Runheng Huang[1,2#], Yaqing Chen[1,2#], Peng Xia[1,3], Bin Huang[1,3], Jie Liu[4], Lingxiang Wu[1,2], Ruohan Zhang[1,2], Qianghu Wang[1,2,3,5]

[1]Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, China; [2]Collaborative Innovation Center for Personalized Cancer Medicine, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Nanjing Medical University, Nanjing, China; [3]School of Biological Science & Medical Engineering, Southeast University, Nanjing, China; [4]Onkocare Life Technology (Suzhou) Company, Suzhou, China; [5]Jiangsu Key Laboratory of Innovative Cancer Diagnosis & Therapeutics, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Nanjing, China

Contributions: (I) Conception and design: Q Wang, Y Qu; (II) Administrative support: L Wu, R Zhang, Q Wang; (III) Provision of study materials or patients: R Huang, Y Chen, R Zhang; (IV) Collection and assembly of data: R Huang, Y Chen, R Zhang; (V) Data analysis and interpretation: Y Qu, P Xia, B Huang, R Huang, Y Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Qianghu Wang, PhD. Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, 101 Longmian Avenue, Jiangning District, Nanjing 211166, China; Collaborative Innovation Center for Personalized Cancer Medicine, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Nanjing Medical University, Nanjing 211166, China; School of Biological Science & Medical Engineering, Southeast University, Nanjing 211189, China; Jiangsu Key Laboratory of Innovative Cancer Diagnosis & Therapeutics, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Nanjing 210029, China. Email: wangqh@njmu.edu.cn; Ruohan Zhang, PhD. Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, 101 Longmian Avenue, Jiangning District, Nanjing 211166, China; Collaborative Innovation Center for Personalized Cancer Medicine, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Nanjing Medical University, Nanjing, China. Email: ruohan.zhang@njmu.edu.cn.

Background: Colorectal cancer (CRC) is a common malignancy associated with genetic and epigenetic alterations. Several methylation biomarkers have been investigated for non-invasive CRC detection; however, their reported performance varies across clinical settings, and the detection of early-stage or precancerous disease and discrimination from non-malignant colorectal conditions remain challenging. This exploratory study aimed to identify reproducible CRC-associated plasma cell-free DNA (cfDNA) methylation regions and to develop and preliminarily evaluate diagnostic prediction model for distinguishing CRC from healthy controls and benign samples.

Methods: Public CRC tissue methylation datasets from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) were analyzed to identify reproducible CRC-associated methylation alterations. Plasma cfDNA methylation was profiled using methyl-CpG-binding-domain enrichment followed by paired-end sequencing in patients with CRC, patients with colorectal polyps, and healthy controls. After quality-control filtering, 30 CRC and healthy-control samples were randomly allocated at the participant level in a 7:3 ratio to a development set comprising 10 patients with CRC and 11 healthy controls and a held-out test set comprising 4 patients with CRC and 5 healthy controls. Hypermethylated regions were selected using least absolute shrinkage and selection operator (LASSO) logistic regression. The 12-region model was evaluated in the held-out test set and subsequently applied to 10 colorectal polyp samples without refitting or recalibration.

Results: Tissue methylation analysis identified reproducible CRC-associated alterations across independent datasets. In the plasma development set, 707 differentially methylated regions (DMRs) were identified between CRC and healthy-control samples, including 324 hypermethylated and 383 hypomethylated

regions. LASSO regression selected a 12-region hypermethylation signature. In the held-out test set, the model achieved an area under the curve (AUC) of 0.85 [95% confidence interval (CI): 0.579–1.000]. At the development-set-derived threshold, sensitivity was 75.0% (3/4), specificity was 60.0% (3/5), and accuracy was 66.7% (6/9). When the original model was applied to colorectal polyp samples, model scores were significantly higher in both CRC and polyp samples than in healthy controls, while CRC samples showed a tendency toward higher scores than polyp samples.

Conclusions: This exploratory study identified a 12-region plasma cfDNA hypermethylation signature associated with CRC and developed a LASSO-based diagnostic prediction model that showed preliminary discrimination between CRC and healthy controls in a small held-out test set. By integrating tissue methylation evidence with plasma cfDNA profiling, this study expands the repertoire of candidate region-level methylation markers for blood-based CRC detection.

Keywords: Colorectal cancer (CRC); methylation; plasma cell-free DNA (plasma cfDNA); biomarker

Submitted Jun 20, 2026. Accepted for publication Aug 13, 2026. Published online Aug 26, 2026.
doi: 10.21037/jgo-2026-0689
View this article at: https://dx.doi.org/10.21037/jgo-2026-0689

Introduction

Colorectal cancer (CRC) is among the most prevalent malignancies worldwide and is associated with substantial morbidity and mortality (1,2). Early detection and timely intervention are essential for improving outcomes in patients with CRC (3). Colonoscopy followed by histopathological assessment remains central to CRC screening and diagnosis; however, its invasive nature, associated discomfort, and resource requirements may limit its acceptability for population-wide and repeated screening (4). Therefore, identifying non-invasive biomarkers that can support preliminary risk identification and help select individuals for further diagnostic evaluation is important for complementing existing CRC screening strategies (5-7).

Liquid biopsy, which enables the detection of circulating tumor biomarkers such as circulating tumor DNA (ctDNA) and circulating tumor cells, has emerged as a potential non-invasive approach for cancer detection and monitoring (8-10). Among liquid-biopsy analytes, circulating cell-free DNA (cfDNA) has attracted considerable attention because of its accessibility and its ability to capture tumor-derived genomic and epigenomic alterations (11-13). Epigenetic modifications, particularly DNA methylation, are pivotal in the oncogenesis and advancement of cancers (14). Abnormal DNA methylation can drive tumorigenesis by silencing tumor suppressor genes or activating oncogenes, thereby promoting tumor initiation, cell proliferation, and metastasis (15). Widespread methylation alterations have been identified in CRC and have been associated with tumor development and clinical outcomes (16). Because tumor-associated methylation changes can be detected in circulating cfDNA, cfDNA methylation profiling has been increasingly investigated as a non-invasive strategy for CRC detection (17,18).

Highlight box

Key findings

- Integrated tissue and plasma methylation profiling identified reproducible colorectal cancer (CRC)-associated methylation regions in plasma cell-free DNA (cfDNA). A 12-region least absolute shrinkage and selection operator-based model showed preliminary discrimination between CRC and healthy controls in a small held-out test set.

What is known and what is new?

- Aberrant DNA methylation is an established feature of CRC, and several blood- or stool-based methylation markers have been investigated for non-invasive CRC detection.
- This study used a tissue-evidence-guided, region-level strategy to expand the repertoire of candidate cfDNA methylation markers beyond a small number of prespecified loci. Exploratory evaluation in colorectal polyp samples indicated that their disease specificity requires further confirmation.

What is the implication, and what should change now?

- These candidate regions require prospective evaluation in larger, independent cohorts that include early-stage CRC and clinically representative benign and precancerous colorectal conditions, together with direct comparison with established methylation assays, before clinical application can be considered.

Several DNA methylation biomarkers, including *SEPTIN9*, *SDC2*, and *BCAT1*, have shown potential for CRC detection, either individually or in multiplex assays (19,20). However, their reported performance varies across studies and clinical settings, and a recent meta-analysis of *SDC2* and *SEPT9* identified substantial between-study heterogeneity (21). Although a combined *SEPTIN9*/*SDC2*/*BCAT1* panel demonstrated favorable diagnostic performance, 51.9% of the patients with CRC had stage III or IV disease, all samples were collected at a single institution, and no separate independent external validation cohort was included. These characteristics may limit the generalizability of the reported performance to average-risk screening populations and patients with early-stage disease (19). This limitation is particularly relevant because the abundance of tumor-derived DNA is generally lower under conditions of low tumor burden, making early-stage CRC and precancerous colorectal lesions more difficult to detect using blood-based assays. Consistently, a prospective ctDNA screening study, although not methylation-specific, also reported limited sensitivity for advanced precancerous lesions (22). These findings support the continued identification of reproducible cfDNA methylation regions that may provide information complementary to established markers.

The present exploratory study aimed to develop and preliminarily evaluate plasma cfDNA methylation-based diagnostic prediction models for CRC. First, hypermethylated differentially methylated regions (hyper-DMRs) were identified in plasma samples from patients with CRC and healthy controls using methyl-CpG-binding-domain enrichment sequencing. A 12-region model was subsequently developed using least absolute shrinkage and selection operator logistic regression to explore its ability to detect CRC-associated methylation signals in a screening-oriented comparison between patients with CRC and healthy controls. The 12-region model was then evaluated in a held-out testing dataset and applied to colorectal polyp samples to examine whether its predictions could be extrapolated to benign colorectal condition. We present this article in accordance with the TRIPOD reporting checklist (available at https://jgo.amegroups.com/article/view/10.21037/jgo-2026-0689/rc).

Methods

Study design

This exploratory diagnostic prediction-model study aimed to develop and preliminarily evaluate plasma cfDNA methylation-based models for the identification of CRC-associated signals in a CRC-versus-healthy-control comparison. A total of 50 plasma samples were sequenced, including 19 CRC samples, 11 benign bowel disease samples, and 20 healthy control samples. After quality-control filtering, 30 samples from patients with CRC and healthy controls were randomly allocated at the participant level in a 7:3 ratio, with similar diagnostic-group proportions maintained between the two subsets. The model-development set comprised 10 patients with CRC and 11 healthy controls, and the held-out test set comprised 4 patients with CRC and 5 healthy controls. The held-out test set was not used for differential methylation analysis, candidate-feature selection, model fitting, penalty-parameter tuning, or classification-threshold determination. It was evaluated only after the model coefficients and classification threshold had been fixed in the development set.

Among the benign bowel disease samples, 10 samples passed the predefined quality-control criteria and were included in exploratory analyses. These samples were not used to develop the primary CRC-versus-healthy-control model.

The study was approved by the Institutional Review Board of The First Affiliated Hospital of Nanjing Medical University (Ethical review No. 2021-SR-189), which waived the requirement of written informed consent due to the retrospective nature of the analysis and the use of de-identified data. The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments.

Participants and diagnostic reference standards

Patients with CRC and patients with colorectal polyps were recruited from The First Affiliated Hospital of Nanjing Medical University between 2022 and 2023, and healthy controls were recruited from volunteer donors during the same period. Patients with CRC were eligible if a pretreatment plasma sample was available, sufficient cfDNA was obtained for methylation-enrichment sequencing, and the diagnosis of CRC was subsequently confirmed by postoperative histopathological examination of resected surgical specimens. Patients with colorectal polyps were eligible if polyps were diagnosed by colonoscopy and a pretreatment plasma sample with sufficient cfDNA was available. Plasma cfDNA methylation results were not used to establish the clinical diagnoses.

Healthy controls were eligible if they had no known history of CRC or other malignant tumors at the time

of blood collection. Samples were excluded from the corresponding analyses if plasma volume or cfDNA yield was insufficient, diagnostic-group information was unavailable, methylation-enrichment sequencing failed the predefined quality-control criteria, or methylation measurements required for model prediction were missing.

Outcome definition and reference standard

The diagnostic outcome for the primary model was CRC status, defined as CRC versus healthy control. Diagnostic-group status was established before plasma cfDNA methylation profiling and was not determined using methylation results. No formal blinding procedure was implemented during model development; however, the reference diagnoses had been established before methylation profiling, and the held-out test samples were not used during feature selection, model fitting, or threshold determination. All patient blood samples used for model development and evaluation were collected before treatment, and treatment information was not used as a predictor. The colorectal polyp samples were used only for supplementary evaluation of the original 12-region model and did not constitute a separate model-development outcome.

Plasma processing and cfDNA extraction

Peripheral blood was collected in 10-mL blood collection tubes (CWY056; CWBIO) before treatment and processed within 24 h of collection. Blood samples were maintained at room temperature before plasma separation. Plasma was separated by centrifugation at 1,500 ×g for 10 min to remove cellular material. cfDNA was extracted from 2 mL of plasma, using the Apostle MiniMax High-Efficiency Cell-Free DNA Isolation Kit, Standard Edition (Apostle Inc., Nanjing, China, catalog No. A17622-1536), according to the manufacturer's instructions. Extracted cfDNA was quantified and assessed for downstream methylation-enrichment sequencing. Samples with insufficient plasma volume or cfDNA yield were excluded from subsequent analyses.

Methyl-CpG binding domain (MBD)-sequencing library preparation and sequencing

For each sample, 10 ng of cfDNA underwent end repair, A-tailing, dephosphorylation, and adapter ligation. Methyl-CpG-binding-domain enrichment was performed using the EpiXplore Methylated DNA Enrichment Kit (Takara Bio Inc., Shiga, Japan; Cat. No. 631962) according to the manufacturer's instructions. The enriched libraries were subsequently subjected to paired-end sequencing on the NovaSeq 6000 Sequencing System (Illumina, Inc., San Diego, CA, USA). Samples were eligible for downstream analyses only if they simultaneously met all predefined quality-control criteria: a CpG relative enrichment score of ≥2, a methylation-capture specificity of ≥0.90, and a HIST1H2BA methylation level of ≥2. Samples failing any one of these criteria were excluded from the corresponding downstream analyses.

Read alignment

Raw MBD-seq FASTQ files were processed using fastp (version 0.20.0) (23) to remove adapter sequences and low-quality reads. Reads were discarded if more than 40% of bases had a Phred quality score below 15 or if they contained more than five ambiguous ("N") bases. The resulting clean reads were aligned to the hg19/GRCh37 human reference genome using Bowtie2 (version 0.7.17) (24). PCR duplicates were subsequently removed from the resulting binary alignment map (BAM) files using MarkDuplicates (version 2.18.25) (25), and properly paired reads with a mapping quality (MAPQ) score >20 were retained using samtools (version 1.3.1) (26). Fragments lacking CpG dinucleotides were further excluded to reduce the contribution of DNA fragments unlikely to have been specifically enriched through MBD-mediated capture. (27).

To assess the specificity of MBD-mediated methylated DNA capture, methylated and non-methylated lambda DNA (λDNA) were included as spike-in controls. Sequencing reads derived from the λDNA controls were aligned to the λDNA reference sequence using the same alignment and filtering pipeline described above. Reads corresponding to methylated and non-methylated λDNA were quantified using samtools, and capture specificity was calculated as follows:

$$S = \frac{M}{(M + nM)} \quad [1]$$

where S denotes the capture specificity, M is the number of reads derived from methylated λDNA , and nM is the number of reads derived from non-methylated λDNA.

450K methylation array data processing

Illumina Human Methylation 450K (HM450K) array datasets from The Cancer Genome Atlas (TCGA) Colon Adenocarcinoma (COAD) and READ projects (28) and Gene Expression Omnibus (GEO) (GSE193535) were used to identify reproducible CRC-associated tissue methylation alterations. The ChAMP R package facilitated the processing of TCGA HM450K array data (29). Probes that failed to meet the set criteria—such as non-CpG recognition, the presence of common single-nucleotide polymorphism (SNPs), ambiguous genomic mapping, or location on sex chromosomes—were excluded from further analysis.

Differentially methylated positions (DMPs) were computed using the champ.DMP function, with significant DMPs defined as $|\Delta\beta| \geq 0.15$ at a false discovery rate (FDR) <0.1, indicating significant methylation differences between the tumor and adjacent normal tissues. DMP β-value visualization was achieved through Z-score transformation, Euclidean distance metrics, and complete linkage clustering. The minfi R package annotated DMPs derived from Infinium HM450K data (30). Further, Enrichr (https://maayanlab.cloud/Enrichr/) was employed for the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, targeting genes associated with the identified DMPs (31).

MEDIPS analysis of MBD-sequencing datasets

The MEDIPS R package (version 1.46.0) was employed for the analysis of the BAM files, facilitating the computation of CpG relative enrichment scores (32). For each candidate tissue DMP, a 300-bp region centered on the CpG locus, comprising 150-bp flanking sequences on each side, was constructed for plasma analysis. MEDIPS and a custom R script were used to calculate normalized methylation-enrichment values for the candidate 300-bp regions in each plasma sample. The resulting methylation expression profile (MEP) matrix contained genomic regions in rows and individual samples in columns, and was quantified using the following formula:

$$FPKM = \frac{Total\ Region\ Fragments}{Mapped\ Fragments(Millions) * Region\ Length(KB)} \quad [2]$$

where FPKM denotes fragments per kilobase per million mapped fragments.

For diagnostic model development, the normalized MEP values of candidate 300-bp DMRs were treated as continuous predictors. Predictor values were measured from pretreatment plasma cfDNA and were not categorized before model development.

Differential methylation analysis of MBD sequencing

Differential methylation analyses were performed separately for the tissue and plasma datasets. For the tissue analysis, fold changes were calculated for each 300-bp region derived from the TCGA DMPs. A one-sided Wilcoxon rank-sum test was used to compare CRC tissues with adjacent non-tumor tissues. Regions with a fold change >1.5 and a nominal P value <0.1 were defined as tissue hypermethylated DMRs.

For the plasma analysis, only samples in the CRC-versus-healthy-control development set were included. The held-out test samples were not used for differential methylation analysis or candidate-feature selection. For each candidate 300-bp region, methylation-enrichment levels were compared between CRC and healthy-control plasma samples using a two-sided Wilcoxon rank-sum test. Regions with an absolute fold change >1.5 and a P value <0.05 were defined as plasma DMRs. Among these plasma DMRs, only regions showing higher methylation levels in CRC samples were retained as candidate predictors for subsequent least absolute shrinkage and selection operator (LASSO) model development. For visualization, the FPKM values of the identified DMRs were transformed into Z-scores, followed by hierarchical clustering using Euclidean distance and complete linkage. Tissue hypermethylated DMRs were annotated to cis-regulatory elements using the ChIPseeker R package. Functional enrichment analysis of genes associated with these DMRs was performed using Metascape (https://metascape.org/gp/index.html#/main/step1) (33).

LASSO model development and testing

LASSO-penalized logistic regression was used to develop an exploratory model for distinguishing patients with CRC from healthy controls. The development set included 10 patients with CRC and 11 healthy controls, and the 324 CRC-hypermethylated plasma DMRs were entered as candidate predictors. The diagnostic outcome was CRC status.

The penalty parameter was selected by 10-fold cross-validation in the development set, and 12 regions with

Table 1 Coefficients of the 12-marker LASSO model

Variable	Coefficient
(Intercept)	–9.811665666
chr5:178016445–178016745	5.818912076
chr17:4803954–4804254	4.740437437
chr22:50623537–50623837	3.410509061
chr6:392960–393260	2.197463652
chr13:79160932–79161232	1.469286459
chr7:157478360–157478660	1.225591602
chr11:32454814–32455114	1.170076187
chr18:44337514–44337814	1.081381757
chr1:91185272–91185572	0.621814345
chr3:147136754–147137054	0.355324709
chr8:99986025–99986325	0.20347615
chr22:50623542–50623842	0.005080651

LASSO, least absolute shrinkage and selection operator.

nonzero coefficients at lambda.min were retained in the final model.

For each sample, the fitted model linear predictor was calculated as follows:

$$\eta = \beta_0 + \sum_{i=1}^{12} \beta_i M_i \quad [3]$$

where β_0 is the model intercept, β_i is the coefficient for the i-th DMR, and M_i is the normalized methylation level of the corresponding DMR.

The fitted model response output was calculated from η. For diagnostic evaluation, a CRC-oriented model score was derived as follows:

$$S_{crc} = \frac{1}{1+e^{-\eta}} \quad [4]$$

where S_{crc} denotes the CRC-oriented model score. A higher S_{crc} indicates a higher likelihood of CRC. The full model intercept and coefficients are provided in *Table 1*.

The held-out test set included nine participants, comprising four patients with CRC and five healthy controls. The model intercept, regression coefficients, and classification threshold were determined exclusively in the development set and remained fixed when applied to the held-out test set. No differential analysis, predictor reselection, coefficient refitting, parameter retuning, threshold adjustment, or recalibration was performed using the test data.

Additionally, the 12-region model was applied to the 10 quality-control-passed polyp samples. Model coefficients, predictor definitions, and the classification threshold were retained without repeated feature selection, coefficient estimation, or threshold recalibration. This analysis was considered exploratory and was not used to update the original model.

Sample size and missing data

No formal a priori sample-size calculation was performed. The study size was determined by the availability of eligible pretreatment plasma samples with sufficient cfDNA yield and methylation-enrichment sequencing data meeting the predefined quality-control criteria. The present analysis was therefore considered exploratory, and model performance was evaluated as preliminary evidence requiring further confirmation in larger cohorts.

No missing-data imputation was performed. Complete-case analysis was used for model development and evaluation. Samples with missing diagnostic outcome labels, insufficient plasma volume or cfDNA yield, failed methylation-enrichment sequencing quality control, or missing predictor measurements were excluded from the

corresponding analysis.

Statistics analysis

The data processing and statistical analyses were conducted in R (version 4.1.3). The DMRs were identified using Wilcoxon rank-sum tests, as described above. Receiver operating characteristic (ROC) curves were generated using the pROC package in R (34). Model discrimination was assessed using the area under the ROC curve (AUC), with 95% confidence intervals (CIs) calculated using the DeLong method. The classification threshold for the 12-region model was determined exclusively in the development set by maximizing the Youden index and was subsequently applied unchanged to the held-out test set and colorectal polyp samples. Sensitivity, specificity, and accuracy were calculated from the corresponding true-positive, true-negative, false-positive, and false-negative classifications, with exact 95% binomial confidence intervals calculated using the Clopper-Pearson method. Differences in model-score distributions among CRC, colorectal polyp, and healthy-control samples were assessed using the Kruskal-Wallis test, followed by two-sided Wilcoxon rank-sum tests for pairwise comparisons. Unless otherwise specified, all statistical tests were two-sided, and $P<0.05$ was considered statistically significant.

Results

Identification of DMRs in CRC tissue

To identify reproducible CRC-associated methylation alterations in tissue, we analyzed paired CRC and adjacent non-tumor tissue methylation profiles from public datasets. In TCGA-COAD and TCGA-READ cohorts, 45 CRC tissue samples and 49 paired adjacent tissue samples were retained. Using a threshold of $|\Delta\beta| \geq 0.15$ and a FDR <0.1, 67,005 significant DMPs were identified, including 27,912 hypermethylated DMPs and 39,093 hypomethylated DMPs in CRC tissues compared with adjacent tissues (*Figure 1A*). An independent Illumina HM450K dataset from GEO, GSE193535, was also analyzed, which included 54 CRC tissues and 54 paired adjacent tissues (35). Applying the same criteria, 36,118 significant DMPs were identified, including 6,269 hypermethylated and 29,849 hypomethylated DMPs. The overlap between TCGA and GSE193535 resulted in 4,160 consistently hypermethylated DMPs and 18,003 consistently hypomethylated DMPs, demonstrating that a substantial subset of CRC-associated methylation alterations was reproducibly detected across independent tissue datasets.

A principal component analysis (PCA) (*Figure 1B*) and unsupervised hierarchical clustering (*Figure 1C*) based on the top 1,000 significant DMPs with $|\Delta\beta| \geq 0.2$ showed clear separation between CRC tissues and paired adjacent tissues, further supporting distinct CRC-associated tissue methylation patterns. A KEGG pathway enrichment analysis of the genes associated with the top 1,000 significant DMPs was then performed. The results revealed the top 10 pathways with a nominal P value <0.05 (*Figure 1D*). Notably, many of these pathways were related to tumor cell growth, adhesion, and migration, such as the Rap1 signaling pathway, cAMP signaling pathway, and cell adhesion molecules. These enrichment results indicated that genes associated with the identified DMPs were involved in several CRC-related biological processes.

Exploratory cfDNA methylation patterns in CRC, colorectal polyp, and healthy-control samples

A total of 50 plasma samples underwent methylation-enrichment sequencing, including samples from 19 patients with CRC, 11 patients with colorectal polyps, and 20 healthy controls. Sequencing-quality metrics for the individual samples are summarized in table online: https://cdn.amegroups.cn/static/public/jgo-2026-0689-1.xlsx. The quality of methylated cfDNA enrichment was evaluated using the CpG relative enrichment score, methylation-capture specificity, and methylation levels of the reference genes *HIST1H2BA* and *GAPDH* (*Figure 2A-2C*). After quality-control filtering, 40 samples were retained, including 14 CRC samples, 10 polyp samples, and 16 healthy-control samples. Detailed sample characteristics, quality-control results, cohort allocation, and reasons for exclusion are provided in table online: https://cdn.amegroups.cn/static/public/jgo-2026-0689-1.xlsx. No diagnostic group labels were missing among the retained samples.

To explore global variation in plasma cfDNA methylation patterns, unsupervised hierarchical clustering and PCA were performed using the top 500 regions identified by analysis of variance between the CRC and healthy-control groups (*Figure 2D,2E*). The resulting plots showed group-level variation in cfDNA methylation patterns; however, overlap among the three groups remained. These findings supported further evaluation of candidate plasma cfDNA methylation regions. Further, we investigated whether these

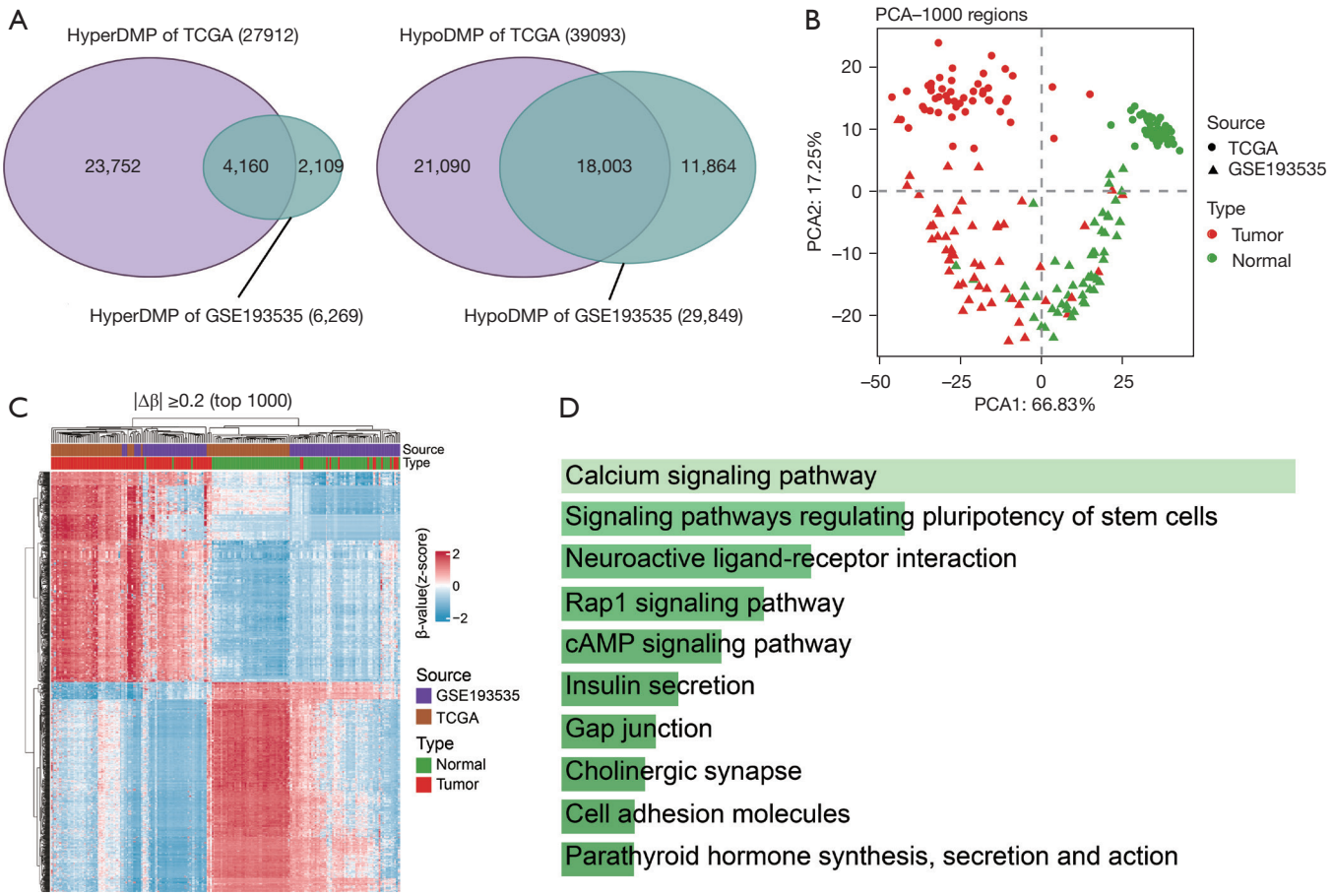

Figure 1 Identification of DMRs in TCGA and GEO datasets. (A) Intersection of hyper- and hypomethylated sites in CRC compared to adjacent tissue from TCGA and GEO datasets. (B,C) PCA plot (B) and unsupervised hierarchical clustering (C) of the top 1,000 most significant DMPs in CRC tissue and paired adjacent tissue. (D) KEGG pathway enrichment analysis of genes associated with the top 1,000 DMPs using Enrichr. CRC, colorectal cancer; DMP, differentially methylated position; DMR, differentially methylated region; GEO, Gene Expression Omnibus; KEGG, Kyoto Encyclopedia of Genes and Genomes; PCA, principal component analysis; TCGA, The Cancer Genome Atlas.

DMRs may contain CpG methylation events in various genomic elements.

Identification of candidate CRC-associated plasma cfDNA DMRs

To identify candidate CRC-associated plasma cfDNA regions, differential methylation analysis was performed using only the 21 samples in the CRC-versus-healthy-control development set, including 10 CRC samples and 11 healthy-control samples. Neither the held-out test samples nor the colorectal polyp samples were included in differential methylation analysis or candidate-feature selection. Using an absolute fold change >1.5 and a two-sided P value <0.05, we identified 707 significant plasma DMRs, including 324 regions that were hypermethylated and 383 regions that were hypomethylated in CRC relative to healthy controls (see table online: https://cdn.amegroups.cn/static/public/jgo-2026-0689-2.xlsx; *Figure 3A*,*3B*). We then performed PCA based on the 707 significant DMRs across 31 plasma samples, including 10 CRC samples, 10 colorectal polyp samples, and 11 healthy controls (*Figure 3C*). CRC and healthy-control samples showed distinct distribution patterns, whereas the polyp samples were distributed between the two groups. These observations provided additional descriptive evidence that the identified

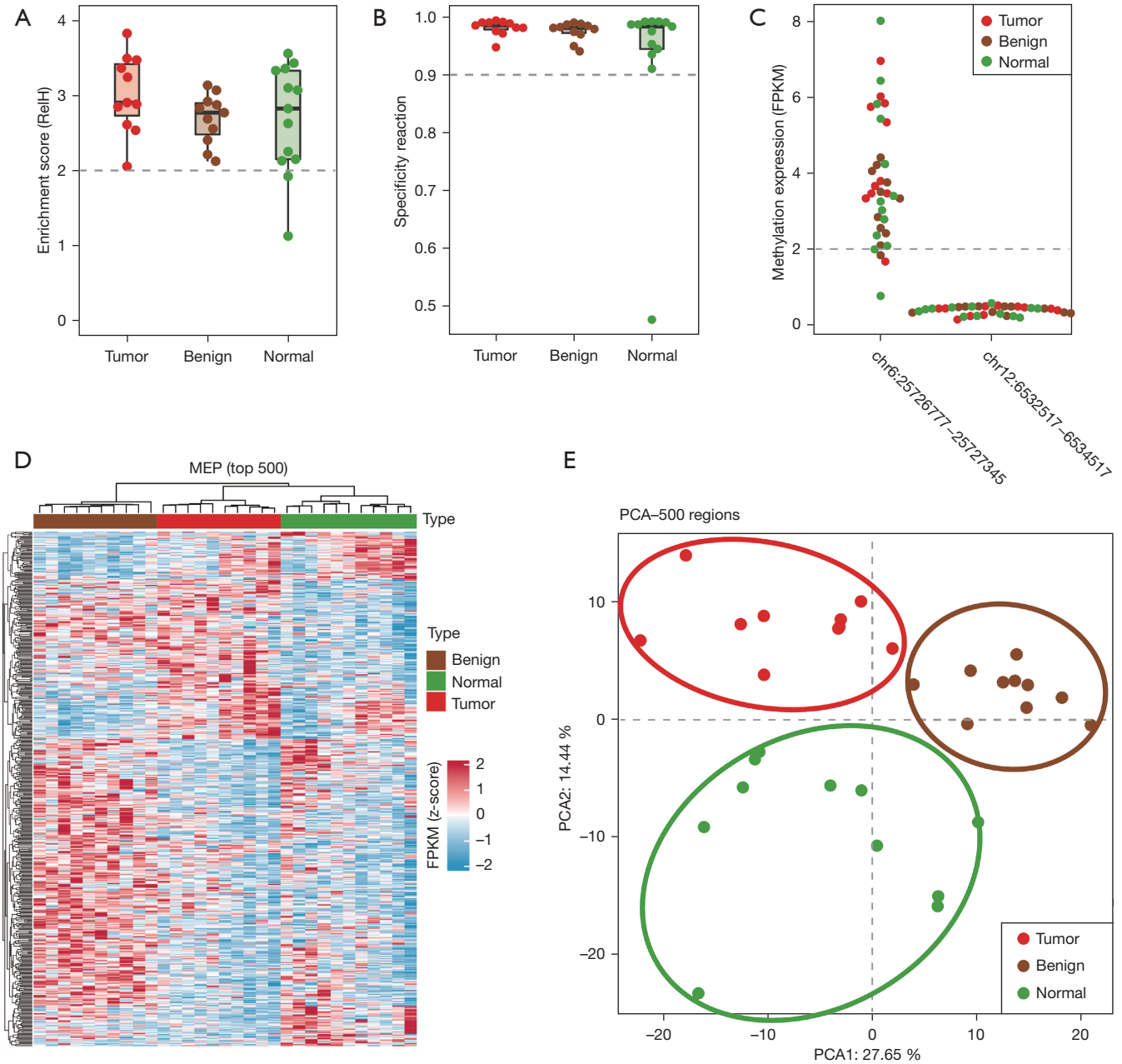

Figure 2 Unsupervised clustering analysis of plasma cfDNA methylation in bowel cancer, benign bowel disease, and healthy control samples. (A,B) CpG relative enrichment score (A) and the specificity of the reaction (B) comparison among CRC patients (n=11), benign intestinal disease patients (n=11), and healthy controls (n=13). (C) Methylation levels of the reference genes *HIST1H2BA* (chr6:25726777–25727345) and *GAPDH* (chr12:6532517–6534517). (D,E) Unsupervised hierarchical clustering (D) and PCA plot (E) of the top 500 significant DMRs using the ANOVA method to differentiate among CRC patients, benign bowel disease patients, and healthy controls. ANOVA, analysis of variance; cfDNA, cell-free DNA; CRC, colorectal cancer; DMR, differentially methylated region; FPKM, fragments per kilobase per million mapped fragments; MEP, methylation expression profile; PCA, principal component analysis.

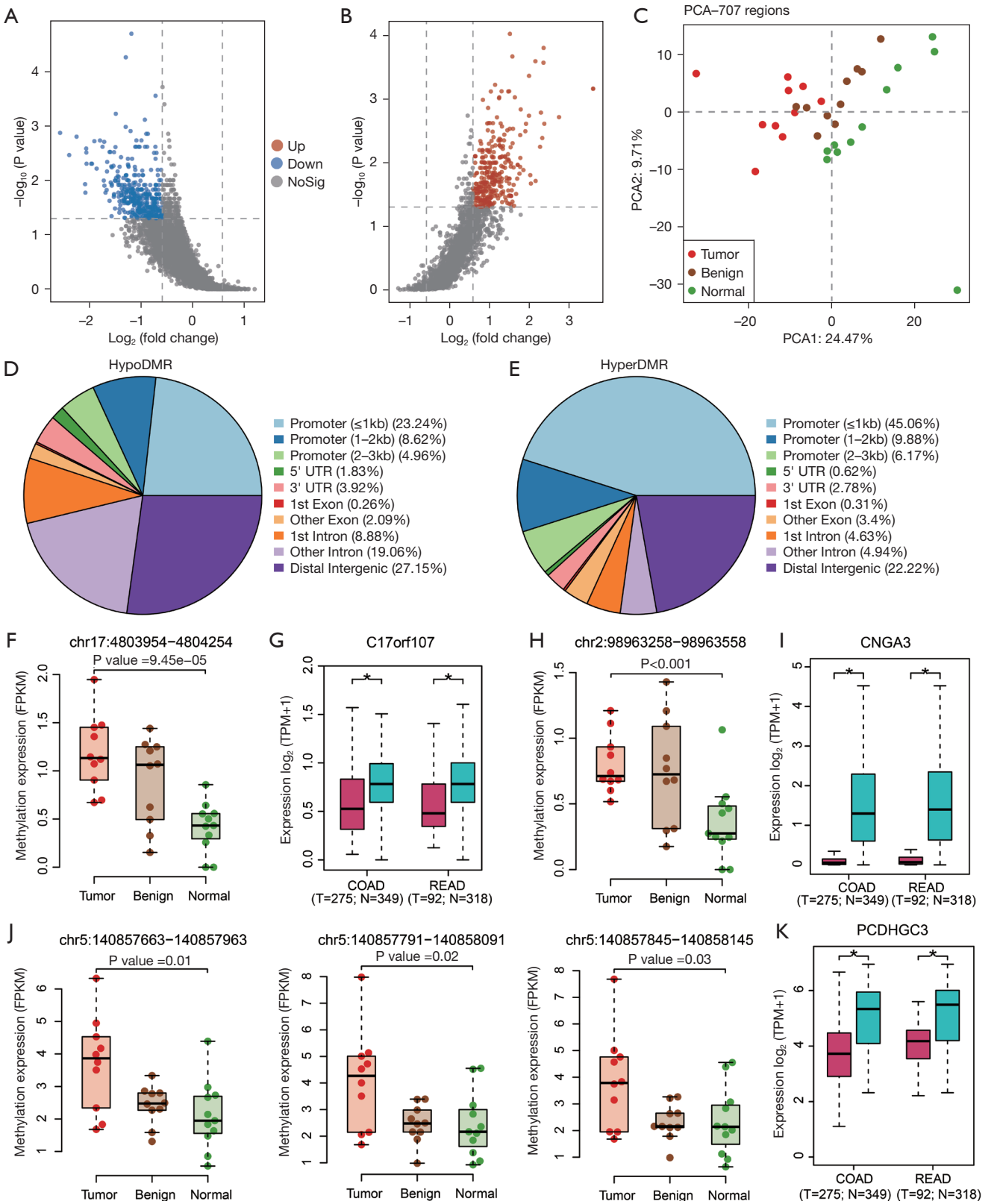

Figure 3 Identification of CRC-specific cfDNA methylation biomarkers. (A,B) Volcano plot highlighting consistent DMRs in tissue and

blood: blue dots (A) represent hypomethylated DMRs, red dots (B) represent hypermethylated DMRs; a total of 707 significant DMRs were detected. (C) PCA of 707 significant DMRs across samples from 10 CRC patients, 10 patients with benign lesions, and 11 healthy controls. (D,E) Pie chart of the proportional distribution of hyper- and hypomethylated DMRs across various cis-regulatory elements based on ChIPseeker analysis. (F) Boxplot displaying plasma cfDNA methylation levels of the promoter region of *C17orf107* among CRC, benign bowel disease, and healthy control plasma samples. (G) Boxplot displaying *C17orf107* mRNA expression levels in CRC tissue samples, with red indicating CRC tumor tissues and blue indicating adjacent normal tissues. (H) Boxplot illustrating plasma cfDNA methylation levels of the promoter region of *CNGA3* among CRC, benign bowel disease, and healthy control plasma samples. (I) Boxplot illustrating *CNGA3* mRNA expression levels in CRC tissue samples, with red indicating CRC tumor tissues and blue indicating adjacent normal tissues. (J) Boxplots illustrating plasma cfDNA methylation levels of three promoter regions of *PCDHGC3* among CRC, benign bowel disease, and healthy control plasma samples. (K) Boxplot illustrating *PCDHGC3* mRNA expression levels in CRC tissue samples, with red indicating CRC tumor tissues and blue indicating adjacent normal tissues. *, P<0.05. cfDNA, cell-free DNA; COAD, colon adenocarcinoma; CRC, colorectal cancer; DMR, differentially methylated region; N, normal; PCA, principal component analysis; READ, rectum adenocarcinoma; T, tumor.

plasma DMRs captured methylation variation across the three colorectal conditions and supported their further evaluation as candidate CRC-associated methylation regions.

We further annotated the 707 plasma DMRs using the ChIPseeker R package to investigate their distribution across cis-regulatory elements (see table online: https://cdn.amegroups.cn/static/public/jgo-2026-0689-3.xlsx). The results showed that the highest proportions of both hypomethylated and hypermethylated DMRs were located in promoter regions, accounting for 36.82% and 61.11%, respectively (*Figure 3D*,*3E*). These results indicated that promoter methylation changes may contribute to transcriptional dysregulation in CRC-related genes, such as *C17orf107*, *CNGA3*, and *PCDHGC3*. Specifically, we identified a hypermethylated DMR located in the promoter region of *C17orf107* (chr17:4803954-4804254) (*Figure 3F*). *C17orf107* messenger RNA (mRNA) expression was lower in the CRC tissues than in the adjacent tissues (*Figure 3G*). We also identified a hypermethylated DMR located in the promoter region of *CNGA3* (chr2:98963258-98963558) (*Figure 3H*), which was similarly associated with lower *CNGA3* mRNA expression in the CRC tissues (*Figure 3I*). Further, three hypermethylated DMRs were identified in the promoter region of *PCDHGC3*, including chr5:140857663-140857963, chr5:140857791-140858091, and chr5:140857845-140858145 (*Figure 3J*). *PCDHGC3* also showed lower mRNA expression in the CRC tissues than in the adjacent tissues (*Figure 3K*).

Development of the 12-region CRC-versus-healthy-control model

To develop the CRC-versus-healthy-control prediction model, the 324 plasma regions that were hypermethylated in CRC relative to healthy controls were entered as candidate predictors into a LASSO-penalized logistic regression model. Model development was performed using the 21-sample development set, which included 10 CRC samples and 11 healthy-control samples. The diagnostic outcome was CRC versus healthy control, and the candidate predictors were normalized methylation levels of DMRs. The penalty parameter λ was selected by 10-fold cross-validation, and DMRs with non-zero coefficients at the selected λ were retained (*Figure 4A*). The LASSO procedure selected 12 hypermethylated DMRs. Their relative methylation levels across the 21 model-development samples were shown in the heatmap (*Figure 4B*). All 12 regions showed higher methylation levels in the CRC samples than in the healthy controls in the model-development dataset (*Figure 4C*). These regions were annotated to or located near genes including *WT1*, *COL23A1*, *ST8SIA5*, *C17orf107*, *IRF4*, *OBI1-AS1*, *ZIC1*, and *PTPRN2*.

The model intercept, and regression coefficients of the 12-region model are provided in *Table 1*. After model development, the predictor definitions, intercept, coefficients, and classification threshold were fixed before the model was applied to the held-out test set.

The testing of the 12-marker model in CRC and healthy control samples

The 12-region model was evaluated in the held-out test set, which included four patients with CRC and five healthy controls. The held-out test samples were not used in differential methylation analysis, feature selection, model fitting, penalty-parameter selection, or classification-

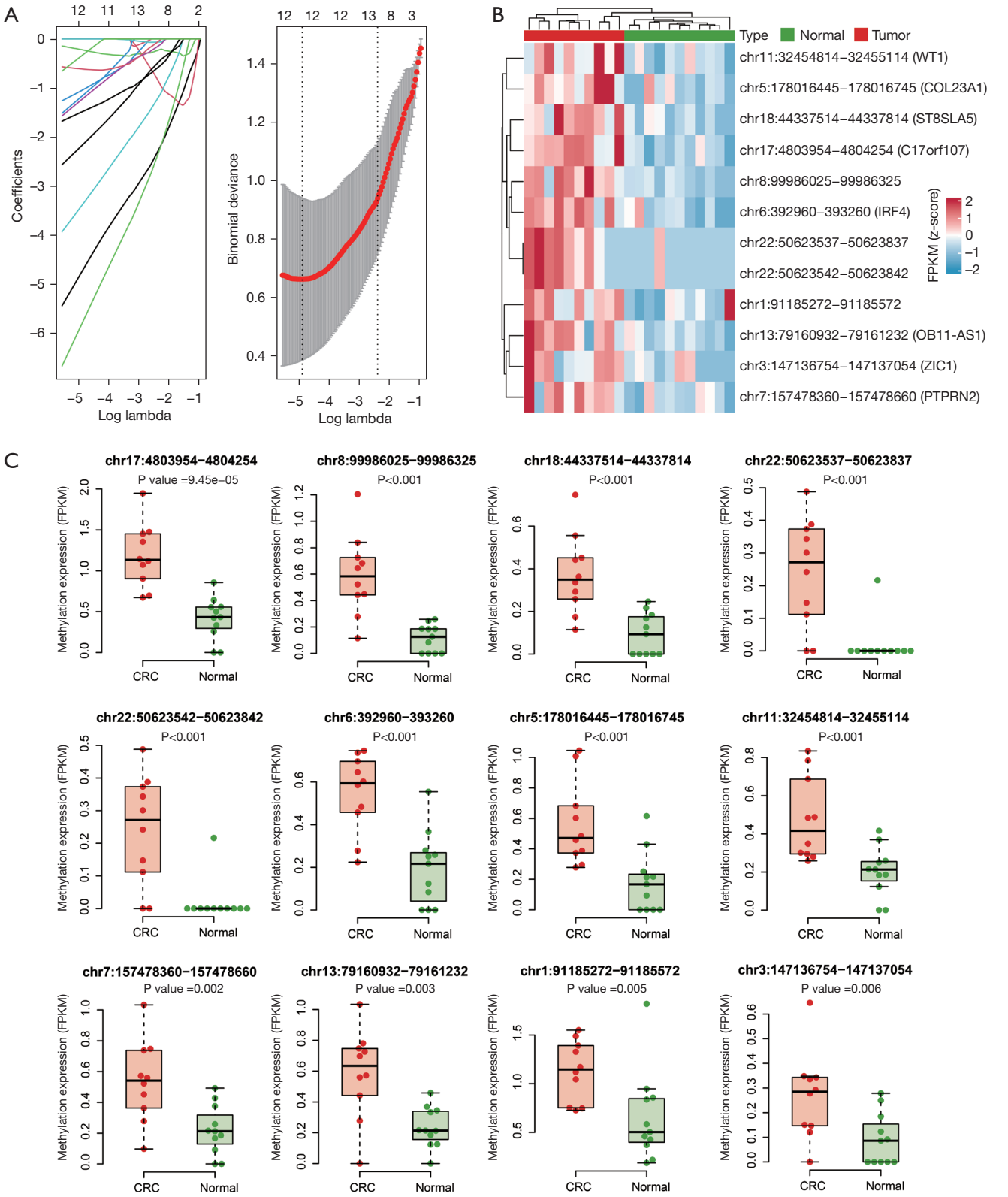

Figure 4 Building a LASSO model. (A) LASSO regression analysis of the hypermethylated DMRs, displaying coefficient profiles and binomial deviance across different values of the penalty parameter λ. (B) Heatmap of relative methylation levels in 12 regions across 21

samples, with red indicating high methylation levels, and blue indicating low methylation levels. (C) Boxplot analysis of methylation levels in 12 regions for the tumor and healthy control samples. CRC, colorectal cancer; DMR, differentially methylated region; FPKM, fragments per kilobase per million mapped fragments; LASSO, least absolute shrinkage and selection operator.

threshold determination. The predictor definitions, model coefficients, intercept, and classification threshold derived from the development set were applied without refitting or recalibration.

Our analysis revealed that the methylation levels of the 12 regions were consistently higher in the CRC patients than in the healthy controls in the held-out test set (Figure S1). Further, the LASSO scores of the 12 regions were higher in the CRC patients compared with the healthy controls in the held-out test set (*Figure 5A,5B*). The model achieved an area under the receiver operating characteristic curve (AUC) of 0.85, with a DeLong 95% CI of 0.579–1.000 (*Figure 5C*). At the development-set-derived classification threshold of 0.501, sensitivity was 75.0% (3/4; 95% CI: 19.4–99.4%), specificity was 60.0% (3/5; 95% CI: 14.7–94.7%), and accuracy was 66.7% (6/9; 95% CI: 29.9–92.5%).

To extend the evaluation of the CRC-versus-healthy-control model to a clinically relevant benign colorectal comparator, the original 12-region model was subsequently applied to the 10 quality-control-passed colorectal polyp samples. No feature reselection, coefficient refitting, threshold adjustment, or recalibration was performed. Model scores were significantly higher in both CRC and colorectal polyp samples than in healthy controls (Figure S2). CRC samples also showed a tendency toward higher CRC-oriented scores than polyp samples, although the difference between these two groups did not reach statistical significance. These findings extend the evaluation of the 12-region methylation signature beyond healthy controls and support further investigation of its performance across malignant and benign colorectal conditions.

Discussion

In this exploratory study, we identified CRC-associated methylation alterations in plasma cfDNA and developed a 12-region hypermethylation model for CRC detection. The selected regions showed consistently higher methylation levels in CRC plasma samples of the model-development dataset, and the locked model achieved an AUC of 0.85 in the held-out test set. When the 12-region model was subsequently applied unchanged to colorectal polyp samples, model scores were significantly higher in both CRC and polyp samples than in healthy controls, while CRC samples showed a tendency toward higher CRC-oriented scores than polyp samples. Together, these findings extend the evaluation of the identified methylation signature beyond the original CRC-versus-healthy-control setting and support further investigation of these regions as candidate plasma cfDNA biomarkers.

In our study, we identified 12 hypermethylated DMRs associated with genes including *WT1*, *COL23A1*, *ST8SIA5*, *C17orf107*, *IRF4*, *OB11-AS1*, *ZIC1*, and *PTPRN2*. Several of these genes have been implicated in CRC-related or tumor-associated biological processes. For example, *WT1* and *IRF4* have been associated with tumor progression and immune regulation, whereas dysregulation of *ZIC1*, *COL23A1*, and *PTPRN2* has been linked to CRC development, invasion, or metastasis (36-41). Other identified genes, including *C17orf107*, *OB11-AS1*, and *ST8SIA5*, have also been reported to participate in cancer-associated cellular processes (42). Collectively, these findings suggest that the identified plasma cfDNA methylation alterations may reflect CRC-associated epigenetic changes and provide biologically plausible candidate regions for further investigation (43-45).

DNA methylation has emerged as an important molecular basis for non-invasive CRC detection, and several methylation markers, including *SEPTIN9*, *SDC2*, and *BCAT1*, have demonstrated diagnostic potential in blood- or stool-based assays (19,20). These established markers and multiplex panels have been evaluated in substantially larger cohorts, and some have shown stronger diagnostic performance than that observed in the present study. However, their reported performance varies across studies and clinical populations, with substantial between-study heterogeneity reported for established markers (21). Some high-performing panels have also been evaluated in single-center cohorts containing a considerable proportion of patients with advanced-stage CRC and without independent average-risk screening validation (19). The principal contribution of the present study instead lies in its tissue-evidence-guided, region-level biomarker-discovery strategy. In contrast to targeted approaches centered on a small number of prespecified loci, we integrated independent CRC tissue methylation datasets with plasma cfDNA methylation profiling to screen CRC-associated methylation

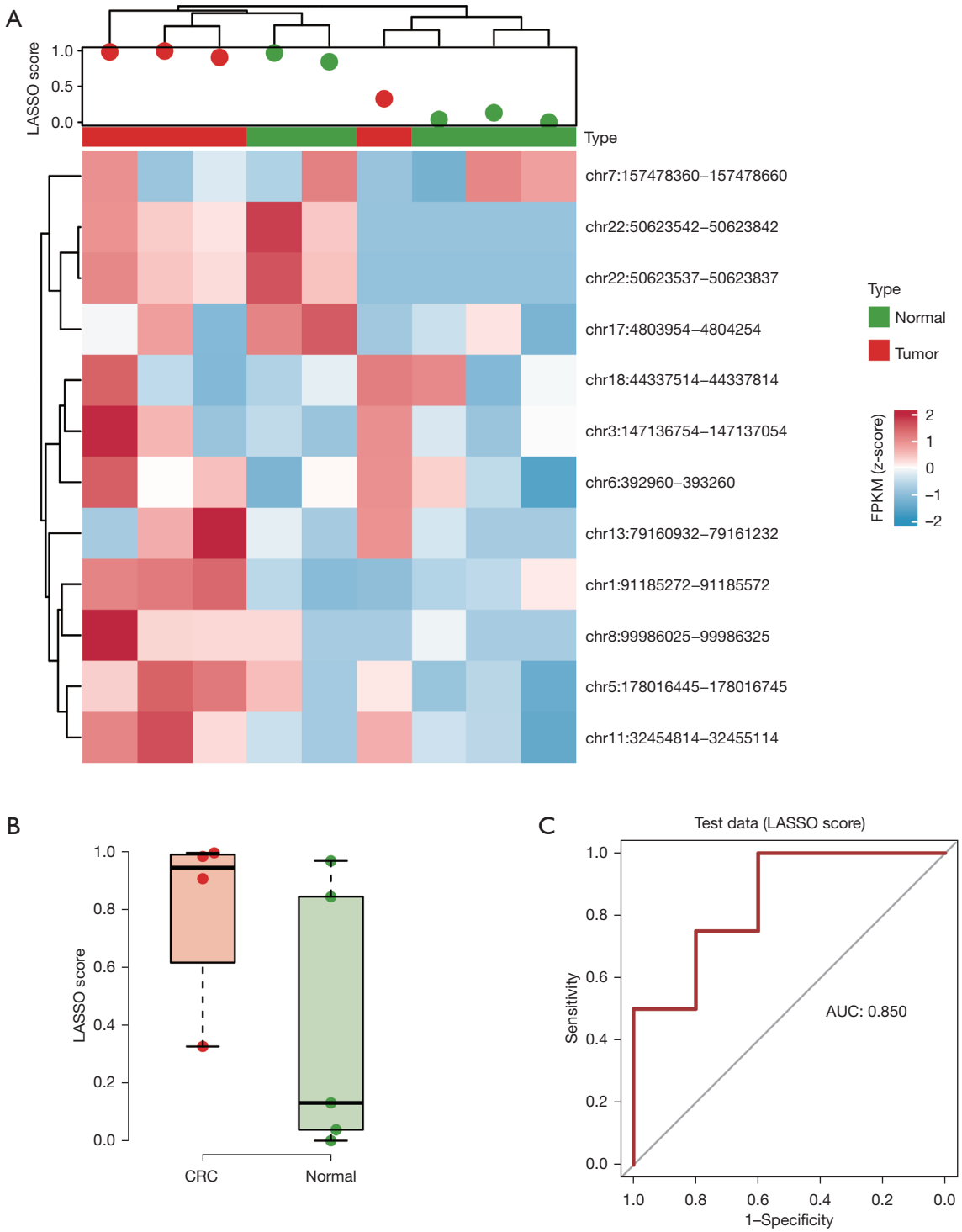

Figure 5 Validation of the discriminative ability of the 12-marker model. (A) A heatmap plot displaying methylation levels in 12 regions of CRC plasma samples (n=4) and healthy control plasma samples (n=5) in the independent testing dataset, and a scatter plot with LASSO scores for CRC patients displayed above the heatmap. (B) Boxplot analysis was conducted to examine the LASSO scores of 12 regions in two different plasma types (tumor and healthy control). (C) ROC analysis was performed for a 12-marker model, which achieved an AUC value of 0.85. AUC, area under the curve; CRC, colorectal cancer; FPKM, fragments per kilobase per million mapped fragments; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic.

alterations at the regional level. This strategy identified a distinct 12-region hypermethylation signature, several components of which were linked to genes with reported roles in tumor-related biological processes. Thus, the present findings expand the repertoire of candidate cfDNA methylation regions and provide additional molecular features for future evaluation. Further studies are required to determine whether these regions provide information complementary to established methylation markers.

The supplementary analysis in colorectal polyp samples further extended the evaluation of the original 12-region signature. When the locked model was applied without feature reselection, coefficient refitting, or recalibration, both CRC and polyp samples showed significantly higher model scores than healthy controls. CRC samples also showed a tendency toward higher CRC-oriented scores than polyp samples. These observations provide additional information on the behavior of the methylation signature across clinically relevant colorectal conditions and support further evaluation of its ability to distinguish malignant from benign colorectal lesions in larger cohorts.

Several limitations should be considered. First, no formal a priori sample-size calculation was performed, and the primary model was developed using the available quality-control-passed pretreatment plasma samples. Although candidate regions were supported by tissue methylation evidence and LASSO penalization was used for model development, the stability of the selected signature should be further evaluated in larger cohorts. Second, the held-out test set included nine participants, and the resulting performance estimates should therefore be regarded as preliminary. Third, the available benign comparator group consisted of colorectal polyps and did not represent the full spectrum of clinically relevant benign and precancerous colorectal conditions. Future studies should include larger prospective and clinically representative cohorts, independent validation, and direct comparison or combination with established methylation-based CRC markers.

Conclusions

This study identified a 12-region plasma cfDNA hypermethylation signature associated with CRC and developed a LASSO-based diagnostic prediction model that showed preliminary discrimination between CRC and healthy controls in a held-out test set. By integrating methylation evidence from independent CRC tissue datasets with plasma cfDNA profiling, we identified additional region-level methylation markers with potential value for CRC detection. Supplementary evaluation in colorectal polyp samples further extended the assessment of the 12-region methylation signature beyond the original CRC-versus-healthy-control setting and supported its continued evaluation across clinically relevant colorectal conditions. Further studies in larger prospective and clinically representative cohorts are warranted to validate these candidate methylation regions and assess their potential complementarity with established CRC methylation markers.

Acknowledgments

We sincerely thank Wu Wei for providing the data and for assistance with manuscript revision, and Professor Li Kening for her valuable guidance and constructive suggestions during the revision of the manuscript.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://jgo.amegroups.com/article/view/10.21037/jgo-2026-0689/rc

Data Sharing Statement: Available at https://jgo.amegroups.com/article/view/10.21037/jgo-2026-0689/dss

Peer Review File: Available at https://jgo.amegroups.com/article/view/10.21037/jgo-2026-0689/prf

Funding: This study was supported by funds to Q.W. from the National Natural Science Foundation of China (grant No. 82372897).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jgo.amegroups.com/article/view/10.21037/jgo-2026-0689/coif). Q.W. reports that this study was supported by funds from the National Natural Science Foundation of China (grant No. 82372897). The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was

conducted in accordance with the Declaration of Helsinki and its subsequent amendments. This study was approved by Institutional Review Board of The First Affiliated Hospital of Nanjing Medical University (Ethical review No. 2021-SR-189). The requirement of written informed consent for participation was waived for this retrospective analysis in accordance with the national legislation and the institutional requirements.

References

1. Pinheiro M, Moreira DN, Ghidini M. Colon and rectal cancer: An emergent public health problem. World J Gastroenterol 2024;30:644-51.
2. Ranasinghe R, Mathai M, Zulli A. A synopsis of modern - day colorectal cancer: Where we stand. Biochim Biophys Acta Rev Cancer 2022;1877:188699.
3. Kaminski MF, Robertson DJ, Senore C, et al. Optimizing the Quality of Colorectal Cancer Screening Worldwide. Gastroenterology 2020;158:404-17.
4. Ladabaum U, Dominitz JA, Kahi C, et al. Strategies for Colorectal Cancer Screening. Gastroenterology 2020;158:418-32.
5. Ciardiello F, Ciardiello D, Martini G, et al. Clinical management of metastatic colorectal cancer in the era of precision medicine. CA Cancer J Clin 2022;72:372-401.
6. Lee Y, Ni J, Beretov J, et al. Recent advances of small extracellular vesicle biomarkers in breast cancer diagnosis and prognosis. Mol Cancer 2023;22:33.
7. Zhao H, Ming T, Tang S, et al. Wnt signaling in colorectal cancer: pathogenic role and therapeutic target. Mol Cancer 2022;21:144.
8. Malla M, Loree JM, Kasi PM, et al. Using Circulating Tumor DNA in Colorectal Cancer: Current and Evolving Practices. J Clin Oncol 2022;40:2846-57.
9. Mo S, Ye L, Wang D, et al. Early Detection of Molecular Residual Disease and Risk Stratification for Stage I to III Colorectal Cancer via Circulating Tumor DNA Methylation. JAMA Oncol 2023;9:770-8.
10. Ye Q, Ling S, Zheng S, et al. Liquid biopsy in hepatocellular carcinoma: circulating tumor cells and circulating tumor DNA. Mol Cancer 2019;18:114.
11. Gao Q, Lin YP, Li BS, et al. Unintrusive multi-cancer detection by circulating cell-free DNA methylation sequencing (THUNDER): development and independent validation studies. Ann Oncol 2023;34:486-95.
12. Lo YMD, Han DSC, Jiang P, et al. Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies. Science 2021;372:eaaw3616.
13. Luo H, Wei W, Ye Z, et al. Liquid Biopsy of Methylation Biomarkers in Cell-Free DNA. Trends Mol Med 2021;27:482-500.
14. Xu X, Peng Q, Jiang X, et al. Metabolic reprogramming and epigenetic modifications in cancer: from the impacts and mechanisms to the treatment potential. Exp Mol Med 2023;55:1357-70.
15. Heeke S, Gay CM, Estecio MR, et al. Tumor- and circulating-free DNA methylation identifies clinically relevant small cell lung cancer subtypes. Cancer Cell 2024;42:225-237.e5.
16. Yasui K, Toshima T, Inada R, et al. Circulating cell-free DNA methylation patterns as non-invasive biomarkers to monitor colorectal cancer treatment efficacy without referencing primary site mutation profiles. Mol Cancer 2024;23:1.
17. Fang Q, Yuan Z, Hu H, et al. Genome-wide discovery of circulating cell-free DNA methylation biomarkers for colorectal cancer detection. Clin Epigenetics 2023;15:119.
18. Zhang K, Fu R, Liu R, et al. Circulating cell-free DNA-based multi-cancer early detection. Trends Cancer 2024;10:161-74.
19. Chen B, Zhao H, Hu H, et al. Circulating DNA methylation-based diagnostic, prognostic, and predictive biomarkers in colorectal cancer. Sci Rep 2025;15:13577.
20. Zhu D, Li J, Zhang W, et al. Highly specific multiplex DNA methylation detection for liquid biopsy of colorectal cancer. Clin Chim Acta 2025;565:120026.
21. Zhang J, Li C, An Y, et al. Comparative analysis of SDC2 and SEPT9 methylation tests in the early detection of colorectal cancer: a systematic review and meta-analysis. Front Med (Lausanne) 2024;11:1460233.
22. Shaukat A, Burke CA, Chan AT, et al. Clinical Validation of a Circulating Tumor DNA-Based Blood Test to Screen for Colorectal Cancer. JAMA 2025;334:56-63.
23. Chen S, Zhou Y, Chen Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;34:i884-90.

24. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357-9.
25. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491-8.
26. Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. Gigascience 2021;10:giab008.
27. Chemi F, Pearce SP, Clipson A, et al. cfDNA methylome profiling for detection and subtyping of small cell lung cancers. Nat Cancer 2022;3:1260-70.
28. Cancer Genome Atlas Research Network; Weinstein JN, Collisson EA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013;45:1113-20.
29. Tian Y, Morris TJ, Webster AP, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics 2017;33:3982-4.
30. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014;30:1363-9.
31. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016;44:W90-7.
32. Lienhard M, Grimm C, Morkel M, et al. MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. Bioinformatics 2014;30:284-6.
33. Wang Q, Li M, Wu T, et al. Exploring Epigenomic Datasets by ChIPseeker. Curr Protoc 2022;2:e585.
34. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.
35. Baharudin R, Ishak M, Muhamad Yusof A, et al. Epigenome-Wide DNA Methylation Profiling in Colorectal Cancer and Normal Adjacent Colon Using Infinium Human Methylation 450K. Diagnostics (Basel) 2022;12:198.
36. Aishanjiang K, Wei XD, Fu Y, et al. Circular RNAs and Hepatocellular Carcinoma: New Epigenetic Players With Diagnostic and Prognostic Roles. Front Oncol 2021;11:653717.
37. Gan L, Chen S, Zhong J, et al. ZIC1 is downregulated through promoter hypermethylation, and functions as a tumor suppressor gene in colorectal cancer. PLoS One 2011;6:e16916.
38. Loeb DM, Sukumar S. The role of WT1 in oncogenesis: tumor suppressor or oncogene? Int J Hematol 2002;76:117-26.
39. Necula L, Matei L, Dragu D, et al. Collagen Family as Promising Biomarkers and Therapeutic Targets in Cancer. Int J Mol Sci 2022;23:12415.
40. Sorokin AV, Nair BC, Wei Y, et al. Aberrant Expression of proPTPRN2 in Cancer Cells Confers Resistance to Apoptosis. Cancer Res 2015;75:1846-58.
41. Wang J, Li S, Li H, et al. IRF4 overexpression promotes the transdifferentiation of tregs into macrophage-like cells to inhibit the development of colon cancer. Cancer Cell Int 2021;21:58.
42. Du J, Su Y, Qian C, et al. Raman-guided subcellular pharmaco-metabolomics for metastatic melanoma cells. Nat Commun 2020;11:4830.
43. Liang L, Zhang Y, Li C, et al. Plasma cfDNA methylation markers for the detection and prognosis of ovarian cancer. EBioMedicine 2022;83:104222.
44. Liu MC, Oxnard GR, Klein EA, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol 2020;31:745-59.
45. Ye J, Zhang J, Ding W. DNA methylation modulates epigenetic regulation in colorectal cancer diagnosis, prognosis and precision medicine. Explor Target Antitumor Ther 2024;5:34-53.

(English Language Editor: L. Huleatt)

Cite this article as: Qu Y, Huang R, Chen Y, Xia P, Huang B, Liu J, Wu L, Zhang R, Wang Q. Development and preliminary validation of plasma cell-free DNA methylation-based diagnostic prediction model for colorectal cancer detection. J Gastrointest Oncol 2026;17(4):244. doi: 10.21037/jgo-2026-0689