



# Targeted Bisulfite Sequencing Reveals DNA Methylation Changes in Zinc Finger Family Genes Associated With KRAS Mutated Colorectal Cancer

## OPEN ACCESS

### Edited by:

Ata Abbas,  
Case Western Reserve University,  
United States

### Reviewed by:

Matteo Pellegrini,  
University of California, Los Angeles,  
United States  
Sabiha Khatoun,  
Texas Tech University Health  
Sciences Center, United States

### \*Correspondence:

Weifeng Ding  
dwf@ntu.edu.cn  
Yanyun Ma  
mayymail@163.com  
Shaoqing Ju  
jsq814@hotmail.com

† These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

**Received:** 17 August 2021

**Accepted:** 11 October 2021

**Published:** 28 October 2021

### Citation:

Pu W, Qian F, Liu J, Shao K,  
Xiao F, Jin Q, Liu Q, Jiang S, Zhang R,  
Zhang J, Guo S, Zhang J, Ma Y, Ju S  
and Ding W (2021) Targeted Bisulfite  
Sequencing Reveals DNA Methylation  
Changes in Zinc Finger Family Genes  
Associated With KRAS Mutated  
Colorectal Cancer.  
*Front. Cell Dev. Biol.* 9:759813.  
doi: 10.3389/fcell.2021.759813

Weilin Pu<sup>1,2,3†</sup>, Fei Qian<sup>4†</sup>, Jing Liu<sup>2</sup>, Keke Shao<sup>5</sup>, Feng Xiao<sup>6</sup>, Qin Jin<sup>7</sup>, Qingmei Liu<sup>8</sup>,  
Shuai Jiang<sup>2</sup>, Rui Zhang<sup>2</sup>, Jun Zhang<sup>9</sup>, Shicheng Guo<sup>10</sup>, Jianfeng Zhang<sup>11</sup>,  
Yanyun Ma<sup>3,12\*</sup>, Shaoqing Ju<sup>1\*</sup> and Weifeng Ding<sup>1\*</sup>

<sup>1</sup> Department of Laboratory Medicine, Affiliated Hospital of Nantong University, Nantong, China, <sup>2</sup> State Key Laboratory of Genetic Engineering, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China, <sup>3</sup> Human Phenome Institute, Fudan University, Shanghai, China, <sup>4</sup> Department of Gastrointestinal Surgery, Affiliated Hospital of Nantong University, Nantong, China, <sup>5</sup> Department of Laboratory Medicine, The First People's Hospital of Yancheng City, Yancheng, China, <sup>6</sup> Department of Pathology, The Third People's Hospital of Nantong City, Nantong, China, <sup>7</sup> Department of Pathology, Affiliated Hospital of Nantong University, Nantong, China, <sup>8</sup> Department of Dermatology, Huashan Hospital, Fudan University, Shanghai, China, <sup>9</sup> Department of Gastroenterology, Huashan Hospital, Fudan University, Shanghai, China, <sup>10</sup> Center for Precision Medicine Research, Marshfield Clinic Research Institute, Marshfield, WI, United States, <sup>11</sup> Department of Gastroenterology, Affiliated Hospital of Nantong University, Nantong, China, <sup>12</sup> Six Industrial Research Institute, Fudan University, Shanghai, China

**Background:** Colorectal cancer (CRC) is a leading cause of cancer death, and early diagnosis of CRC could significantly reduce its mortality rate. Previous studies suggest that the DNA methylation status of zinc finger genes (ZFGs) could be of potential in CRC early diagnosis. However, the comprehensive evaluation of ZFGs in CRC is still lacking.

**Methods:** We first collected 1,426 public samples on genome-wide DNA methylation, including 1,104 cases of CRC tumors, 54 adenomas, and 268 para-tumors. Next, the most differentially methylated ZFGs were identified and validated in two replication cohorts comprising 218 CRC patients. Finally, we compared the prediction capabilities between the ZFGs and the SEPT9 in all CRC patients and the KRAS + and KRAS- subgroup.

**Results:** Five candidate ZFGs were selected: *ESR1*, *ZNF132*, *ZNF229*, *ZNF542*, and *ZNF677*. In particular, *ESR1* [area under the curve (AUC) = 0.91] and *ZNF132* (AUC = 0.93) showed equivalent or better diagnostic capability for CRC than *SEPT9* (AUC = 0.91) in the validation dataset, suggesting that these two ZFGs might be of potential for CRC diagnosis in the future. Furthermore, we performed subgroup analysis and found a significantly higher diagnostic capability in KRAS + (AUC ranged from 0.97 to 1) than that in KRAS- patients (AUC ranged from 0.74 to 0.86) for all these

five ZFGs, suggesting that these ZFGs could be ideal diagnostic markers for KRAS mutated CRC patients.

**Conclusion:** The methylation profiles of the candidate ZFGs could be potential biomarkers for the early diagnosis of CRC, especially for patients carrying KRAS mutations.

**Keywords:** colorectal cancer, DNA methylation, zinc finger family, KRAS, diagnosis

## INTRODUCTION

Colorectal cancer (CRC) is the second most common cause of cancer in the United States and accounts for 10% of cancer incidences and 9.4% of cancer deaths worldwide in 2020 (Siegel et al., 2020; Sung et al., 2021). Previous studies have found that the accumulations of both genetic and epigenetic alterations lead to CRC carcinogenesis. The 5-year survival rate of CRC ranges from 90% in its early stages (when it is localized and regional), but it decreases significantly to about 14% when detected at a distant stage, highlighting the importance of early detection methods (Siegel et al., 2020).

Recently, the application of screening modalities, including colonoscopy and image-based detection, significantly decreased the mortality of CRC (Siegel et al., 2017). However, these screening methods are not widely used across populations due to abdominal pain, discomfort, and other contraindications. Thus, the development of non-invasive and precise diagnostic methods for CRC is needed.

DNA methylation is a crucial epigenetic modification in the human genome and plays vital roles in embryonic development, transcription regulation (He et al., 2011; Jiang et al., 2018), and genomic imprinting (Schubeler, 2015). Moreover, the dynamic changes of DNA methylation in different tissues and disease courses make it a promising tool to develop the tissue-of-origin test (Guo et al., 2017a) and disease prediction (Guo et al., 2015), especially for cancers (Koch et al., 2018) and immune diseases (Guo et al., 2017b; Ding et al., 2018). Till now, A series of DNA methylation-based biomarkers have been found in CRC, including *SEPT9* (Barault et al., 2018; Freitas et al., 2018; Wills et al., 2018; Sun et al., 2019). However, the performance of *SEPT9* is not as good as that of the stool DNA test (Ahlquist et al., 2012; Church et al., 2014; Song et al., 2017). Therefore, identifying better DNA-methylation-based biomarkers with high accuracy would be beneficial in liquid biopsy in CRC.

Zinc finger proteins (ZFPs) are a prominent component of transcriptional factors in eukaryotes. The ZFP family can be divided into eight categories according to their distinct motifs: Cys2His2 (C2H2)-like, Gag knuckle, Treble clef, Zinc ribbon, Zn2/Cys6, TAZ2 domain-like, zinc-binding loops, and metallothionein (Jen and Wang, 2016). The C2H2-type zinc finger motifs form the largest class. Currently, growing bodies of evidence suggest that ZFPs could contribute to tumor progression or suppress it *via* transcriptional regulation. The DNA methylation alterations of multiple ZFPs have been recognized as promising biomarkers for tumor diagnosis, prognosis,

and drug response due to their vital roles in cancers. However, a comprehensive and systemic assessment of the DNA methylation profiles of zinc finger genes (ZFGs) in CRC is lacking.

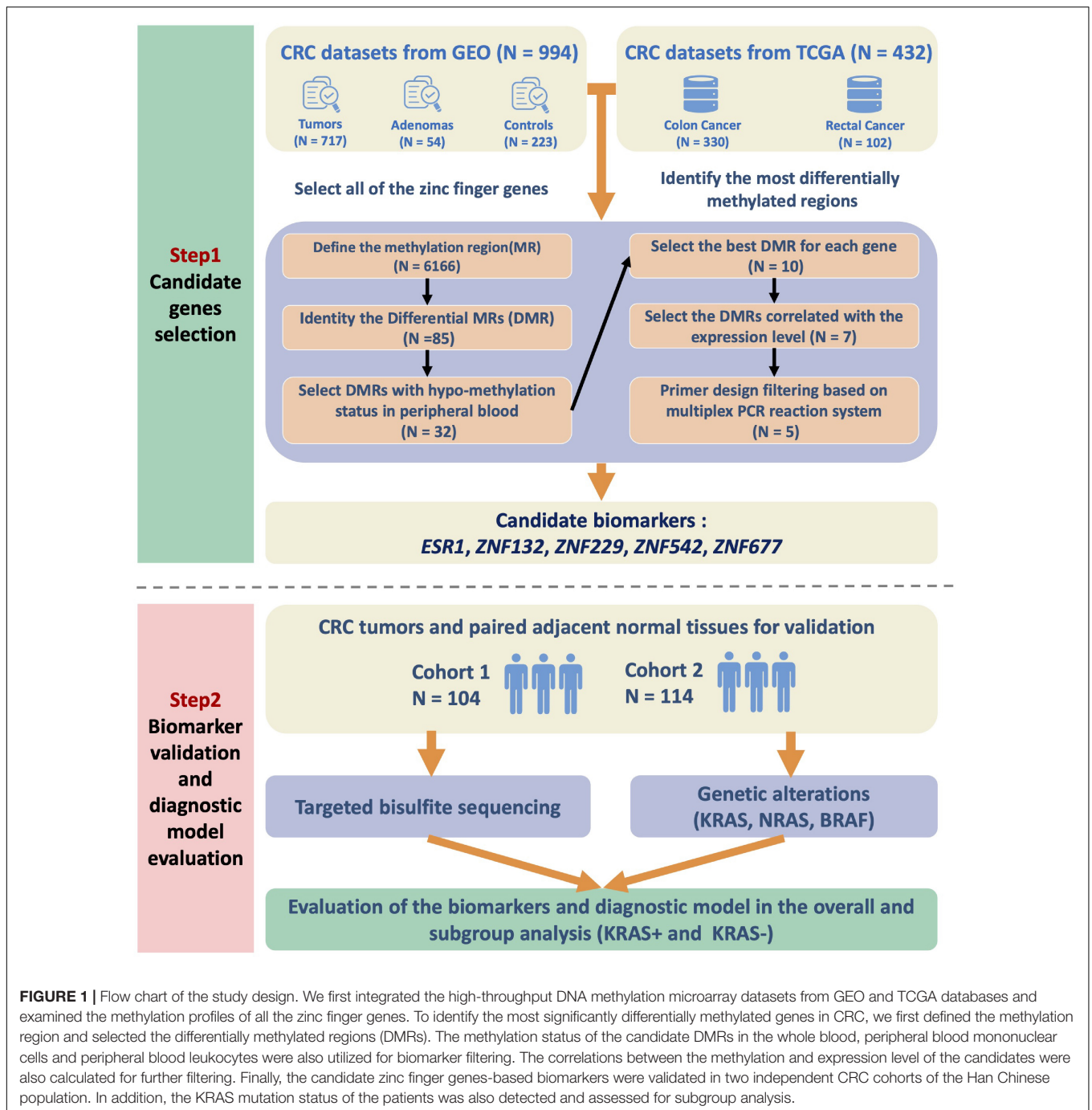
In this study, we exhaustively searched and combined public microarray datasets on high-throughput DNA methylation for the first time, including 1,104 CRC tumors, 54 adenomas, and 268 para-tumors. We identified seven candidate zinc finger genes using comprehensive filtering procedures, and five of them were successfully validated in 104 CRC Han Chinese patients, especially in the CRC tumors carrying the KRAS mutation. To confirm the findings, we recruited another 114 CRC patients of Han Chinese descent and yielded consistent results. In particular, two of these ZFGs, *ESR1*, and *ZNF132*, showed a significantly higher diagnostic capability than *SEPT9*, suggesting that they might be promising biomarkers for CRC diagnosis, especially for KRAS mutated patients.

## RESULTS

### Five Zinc Finger Genes Were Identified as Candidate Colorectal Cancer Diagnostic Biomarkers

To identify robust DNA methylation-based biomarkers, we searched the public datasets for the DNA methylation status in CRC cases and collected 1,104 CRC tumors, 268 para-tumor samples, and 54 adenomas for further analysis (See section “Materials and Methods,” **Supplementary Table 1**). We also obtained the complete list of genes belonging to the ZFP family (**Supplementary Table 2**). Based on the feature selection procedures described in the section “Materials and Methods,” we finally identified five candidate genes: *ESR1*, *ZNF132*, *ZNF229*, *ZNF542*, and *ZNF677* (**Figure 1**).

These five genes were significantly hypermethylated in both CRC and adenoma tissues compared to the para-tumors (**Supplementary Figure 1**). Consistently, the expression levels of these genes were also significantly down-regulated in CRC tumors compared to para-tumors in the TCGA dataset (**Figure 2**). To characterize the abilities of these biomarkers quantitatively in the combined discovery dataset, we constructed a univariate logistic regression model for each gene and obtained robust discrimination between CRC tumors and para-tumor

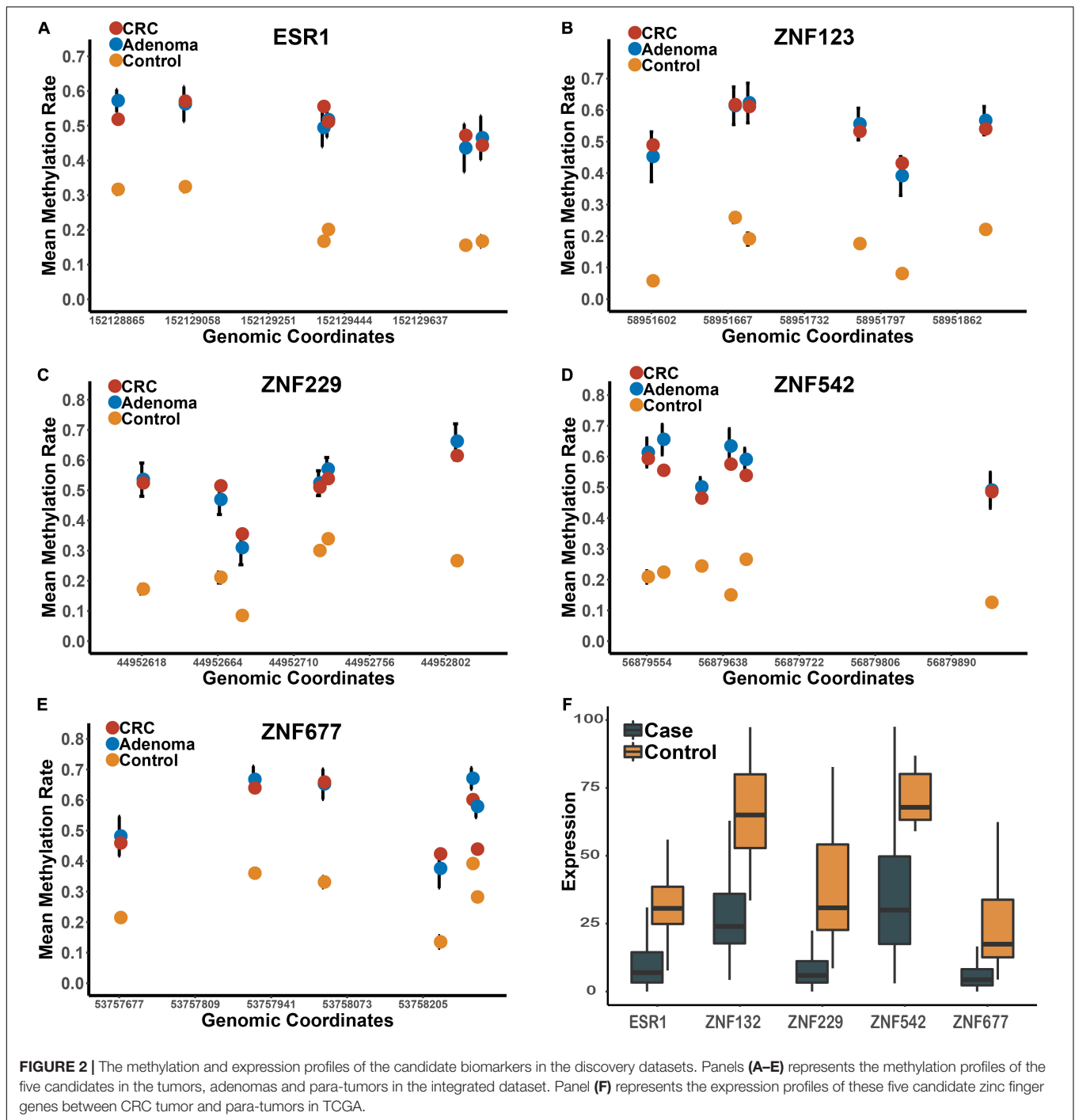


tissues (sensitivity = 0.82–0.90, specificity = 0.88–0.97, AUC = 0.93–0.97).

### Targeted Bisulfite Sequencing in the Han Chinese Population Confirmed the Efficacy of the Candidate Zinc Finger Genes-Based Biomarkers

To further verify these DNA methylation-based biomarkers, we recruited 104 CRC patients of the Han Chinese population.

The characteristics of these CRC patients (replication cohort 1) are shown in **Table 1**. The CRC tumors and their matched para-tumors were obtained for targeted bisulfite sequencing. The methylation profiles of the candidates and the known biomarker (*SEPT9*) were examined using targeted bisulfite sequencing (See section “Materials and Methods”). The bisulfite conversion rate (C to T) was high (>99%) in both tumors and para-tumors, and no significant difference in the read mapping rate was found between groups (**Supplementary Figure 2**). After quality control, 187 samples were retained



for further analysis, including 98 CRC tumors and 89 para-tumors.

Based on the methylation profiles of these five ZFGs, we performed principal component analysis (PCA) and found a significant distinction between CRC tumors and para-tumors (Supplementary Figure 3). A differential methylation analysis was also conducted for these candidates, and we found that all candidate genes were significantly hypermethylated in CRC tumors (Figure 3 and Supplementary Figure 4). Furthermore,

we performed a univariate logistic regression analysis without adjusting for covariates, and created the receiver operating characteristic (ROC) curve to reveal the diagnostic ability of each candidate gene. As shown in Supplementary Table 3, the area under the curve (AUCs) of these candidates ranged from 0.85 to 0.93. In particular, we found that the diagnostic capability of *ZNF132* (AUC = 0.91) and *ESR1* (AUC = 0.93) was equal to or better than that of *SEPT9* (AUC = 0.91), indicating that these ZFGs might have great potential for CRC diagnosis.



## The Diagnostic Abilities of the Candidates Were Significantly Affected by the KRAS Mutation Status of Colorectal Cancer Patients

In addition to the overall differential methylation analysis, we also evaluated the effects of age, gender, tumor stage, location of the tumor (colon or rectum), and the mutation status (KRAS + vs. KRAS-) of the CRC samples (Supplementary Tables 4–8). No significant differences were found in the diagnostic capability between CRC patients in the young/old, male/female, early/late stage, and colon/rectum subgroups (Supplementary Table 9, see section “Materials and Methods”). However, the diagnostic capability for CRC patients carrying the KRAS mutation

**TABLE 1** | Characteristics of the CRC patients included in this study.

Characteristics	Patient distribution	
	Cohort 1 (N = 104)	Cohort 2 (N = 114)
Age	66 (IQR = 62 to 74)	68 (IQR = 60 to 75)
<b>Gender</b>		
Male	71	75
Female	33	39
<b>Subtype<sup>a</sup></b>		
Colon	55	68
Rectum	49	45
<b>UICC Stage<sup>b</sup></b>		
I	18	28
II	35	38
III	40	38
IV	11	9
<b>Tumor invasion depth<sup>c</sup></b>		
T1	5	7
T2	21	22
T3	70	69
T4	6	15
Tx	2	1
<b>Lymph node involvement<sup>c</sup></b>		
N0	57	72
N1	29	25
N2	11	15
N3	5	2
Nx	2	
<b>Distant metastasis<sup>c</sup></b>		
M0	93	105
M1	11	9
<b>KRAS mutation<sup>d</sup></b>		
Positive	50	52
Negative	52	62

<sup>a</sup>The tumors were classified into colon or rectum based on the location of the tumor.

<sup>b</sup>The UICC stage was determined after surgical intervention and histological examination of the specimen.

<sup>c</sup>TNM Stages were assessed by the seventh edition of the TNM classification criteria.

<sup>d</sup>The KRAS mutation status was examined using the FastTarget next-generation sequencing method.

(KRAS +) was significantly better than that for KRAS- samples (Supplementary Tables 8, 9). In the KRAS + group, the sensitivity of each gene ranged from 0.86 to 0.98, the specificity range was 0.89 to 1.00, and the AUC range was 0.97 to 1.00, which is significantly higher than the sensitivity (0.52 to 0.82), specificity (0.80 to 0.98) and AUC (0.74 to 0.86) in the KRAS- group (Supplementary Table 8). Moreover, the hierarchical clustering analysis revealed that all CRC tumors misclassified into the control group were KRAS- samples (Figure 4). In addition, we further examined if other factors may be associated with these misclassified CRC tumors ( $n = 17$ ). However, none of the age, gender, tumor stage, and tumor locations showed significant differences in our study's overall CRC tumors. Furthermore, we examined the diagnostic abilities of the ZFGs in the KRAS + and KRAS- subgroup from the TCGA dataset for verification. Consistently, the diagnostic capability of the ZFGs in the KRAS + subgroup (AUC = 0.96–1) was significantly higher than in the KRAS- subgroup (AUC = 0.93–0.97), suggesting the significant methylation changes between KRAS + and KRAS- subgroup (Supplementary Table 10).

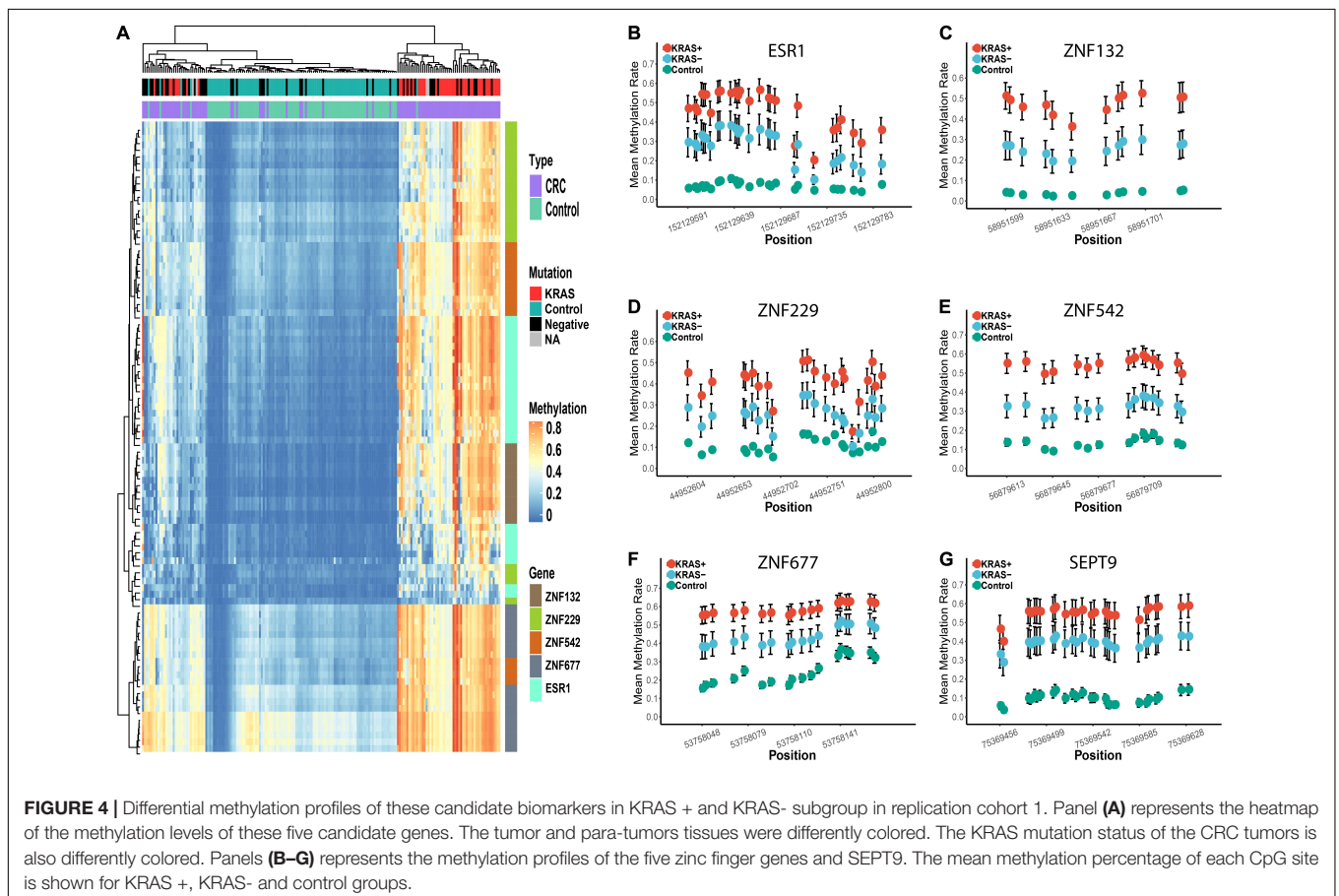
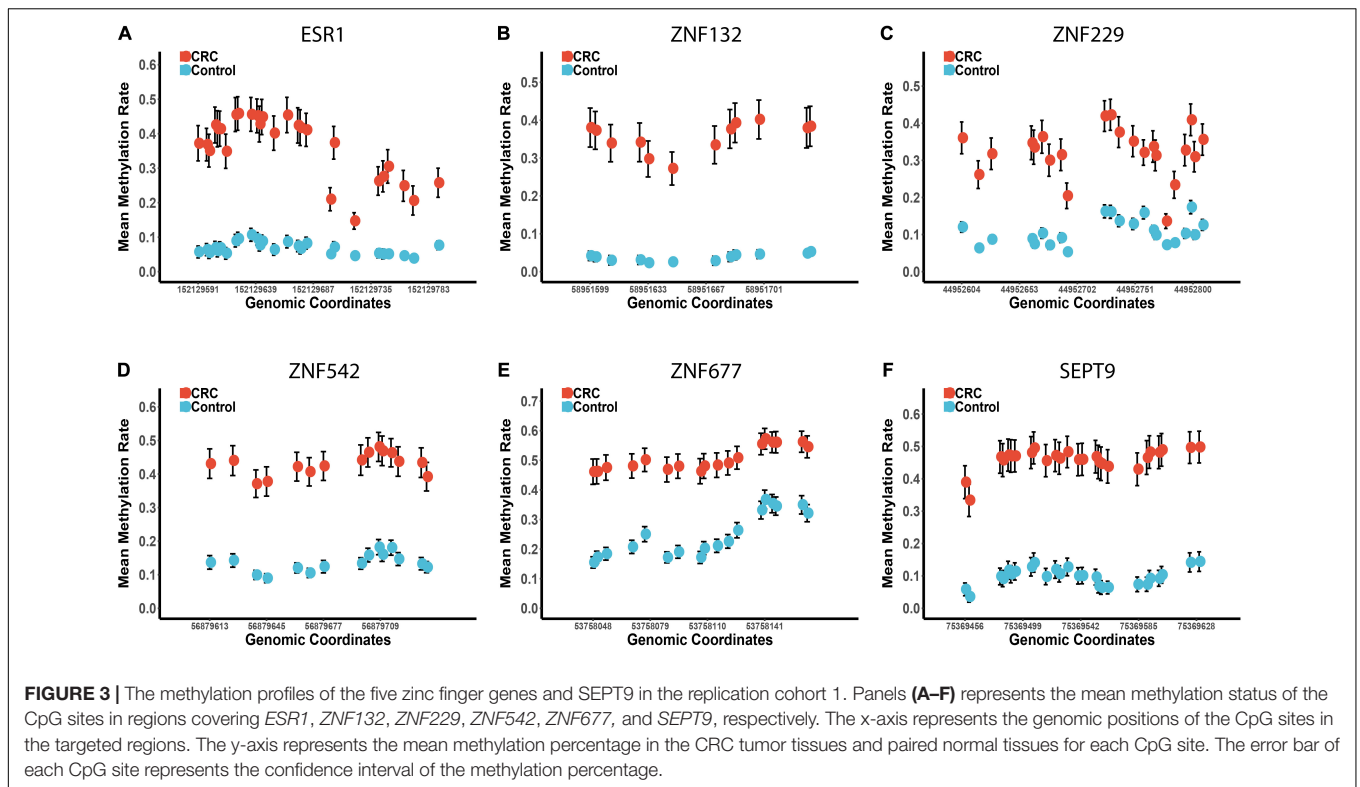
## An Independent Replication Cohort Validated the Association Between KRAS Mutation and Diagnostic Capabilities of the Zinc Finger Genes

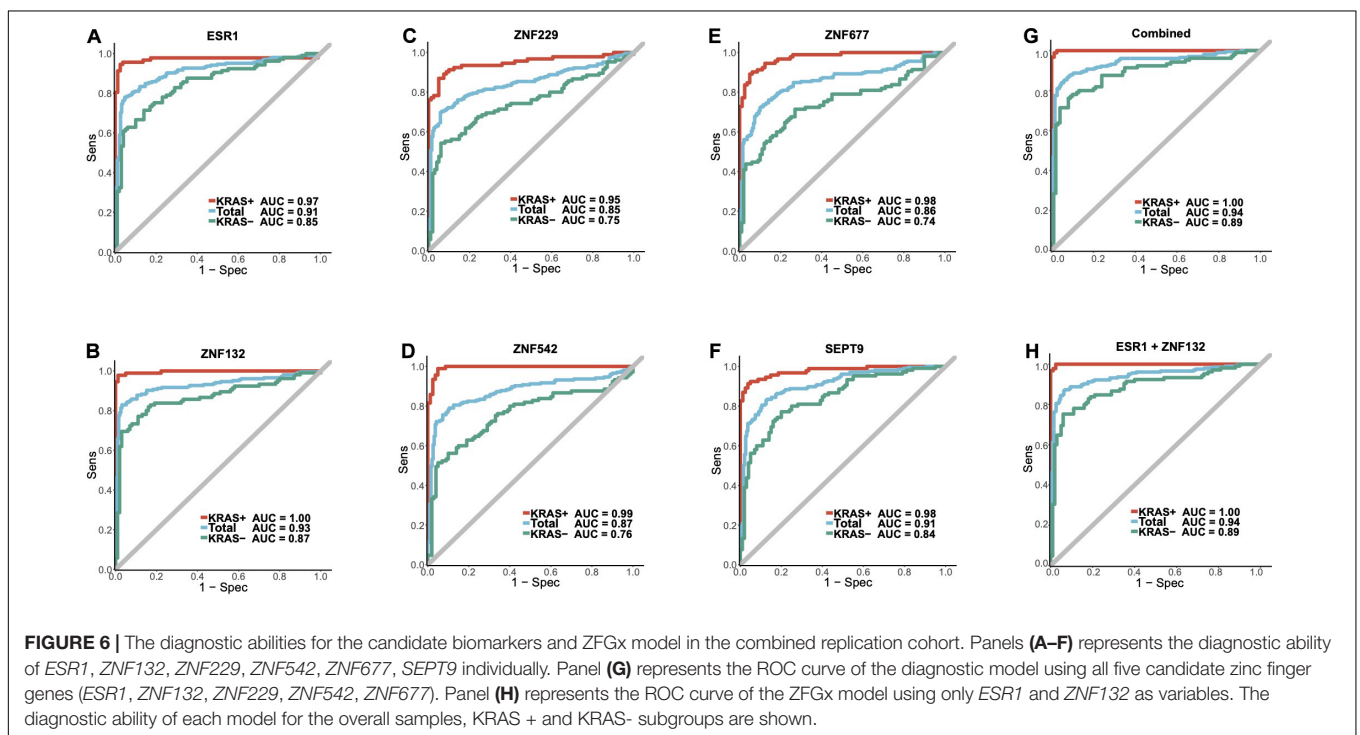
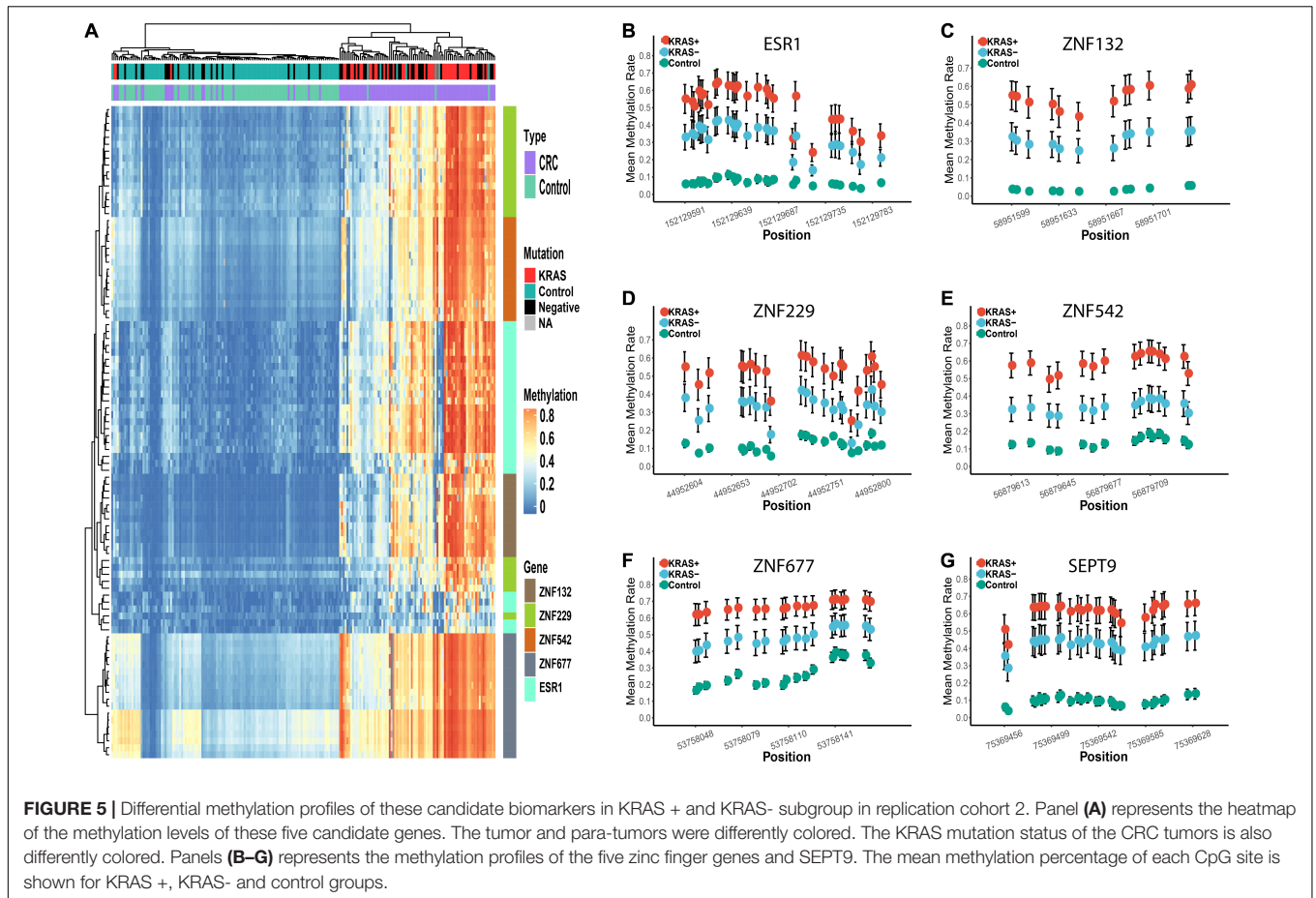
Due to the limited sample size in replication cohort 1, we recruited another sample group (replication cohort 2), including 114 CRC tumors and matched para-tumors from patients of Han Chinese descent (Table 1) for further validation. The methylation profiles of these ZFGs were measured in replication cohort 2 (Supplementary Figures 5, 6). Consistently, we found that *ESR1* (AUC = 0.90) and *ZNF132* (AUC = 0.94) still showed the best diagnostic capability for CRC (Supplementary Table 11).

Furthermore, the significant differences of methylation profiles between the KRAS + and KRAS- subgroups were also validated (Supplementary Figure 7). In the KRAS + subgroup, the sensitivity (0.90 to 1.00), specificity (0.91 to 0.98), and the AUC (0.92 to 1.00) of each gene was significantly higher than that in the KRAS- subgroup (sensitivity: 0.58 to 0.85, specificity: 0.78 to 0.96, AUC: 0.71 to 0.88) (Supplementary Table 12). Similarly, the CRC tumors misclassified as the para-tumors were in the KRAS- subgroup, confirming that the KRAS + CRC samples were more epigenetically homogeneous than the KRAS- CRC samples (Figure 5).

## The Diagnostic Abilities of Zinc Finger Genes Could Achieve Superior Performance Than That of *SEPT9* for Colorectal Cancer Diagnosis Using All Replication Samples

To further identify the robust diagnostic biomarker for CRC diagnosis, we combined the datasets from replication cohort 1 and replication cohort 2. As shown in Figure 6, *ZNF132* had the highest diagnostic capability (sensitivity = 0.83, specificity = 0.97, AUC = 0.93) than other genes, including





**TABLE 2** | The mean methylation status of the five genomic regions in the KRAS + and KRAS- samples of replication cohort 1 and cohort 2.

	Genomic Region <sup>a</sup>	Gene <sup>b</sup>	McaM <sup>c</sup>	McoM <sup>c</sup>	P value <sup>d</sup>	FDR	OR	95% CI <sup>e</sup>	Sens <sup>f</sup>	Spec <sup>f</sup>	AUC <sup>f</sup>
KRAS +	6:152129591-152129791	<i>ESR1</i>	0.49	0.06	$8.00 \times 10^{-27}$	$9.60 \times 10^{-27}$	11.00	7.53–15.60	0.96	0.98	0.97
	19:58951599-58951728	<i>ZNF132</i>	0.51	0.03	$2.10 \times 10^{-29}$	$1.20 \times 10^{-28}$	29.00	17.6–47.90	0.98	0.99	1.00
	19:44952604-44952808	<i>ZNF229</i>	0.47	0.11	$1.70 \times 10^{-24}$	$1.70 \times 10^{-24}$	10.20	7.19–14.00	0.90	0.93	0.95
	19:56879613-56879735	<i>ZNF542</i>	0.57	0.13	$4.00 \times 10^{-29}$	$1.20 \times 10^{-28}$	14.30	9.60–21.50	0.99	0.95	0.99
	19:53758048-53758164	<i>ZNF677</i>	0.63	0.25	$3.40 \times 10^{-27}$	$5.10 \times 10^{-27}$	10.90	7.64–15.30	0.90	0.95	0.98
	17:75369456-75369630	<i>SEPT-9</i>	0.58	0.08	$2.40 \times 10^{-27}$	$4.70 \times 10^{-27}$	7.85	5.55–11.00	0.91	0.96	0.98
KRAS-	6:152129591-152129791	<i>ESR1</i>	0.31	0.08	$2.60 \times 10^{-18}$	$7.90 \times 10^{-18}$	4.63	3.21–6.39	0.75	0.83	0.85
	19:58951599-58951728	<i>ZNF132</i>	0.28	0.04	$1.30 \times 10^{-19}$	$7.70 \times 10^{-19}$	6.10	4.09–8.73	0.70	0.97	0.87
	19:44952604-44952808	<i>ZNF229</i>	0.29	0.12	$6.90 \times 10^{-10}$	$8.20 \times 10^{-10}$	3.30	2.20–4.60	0.54	0.94	0.75
	19:56879613-56879735	<i>ZNF542</i>	0.33	0.15	$1.50 \times 10^{-10}$	$2.30 \times 10^{-10}$	2.78	1.87–3.84	0.50	0.96	0.76
	19:53758048-53758164	<i>ZNF677</i>	0.46	0.27	$1.80 \times 10^{-09}$	$1.80 \times 10^{-09}$	2.38	1.62–3.25	0.71	0.73	0.74
	17:75369456-75369630	<i>SEPT-9</i>	0.41	0.11	$2.30 \times 10^{-17}$	$4.70 \times 10^{-17}$	2.94	2.11–3.92	0.77	0.80	0.84
Total	6:152129591-152129791	<i>ESR1</i>	0.39	0.07	$1.20 \times 10^{-45}$	$3.50 \times 10^{-45}$	6.06	4.81–7.52	0.78	0.95	0.91
	19:58951599-58951728	<i>ZNF132</i>	0.39	0.04	$6.30 \times 10^{-49}$	$3.80 \times 10^{-48}$	8.67	6.60–11.20	0.83	0.97	0.93
	19:44952604-44952808	<i>ZNF229</i>	0.37	0.11	$3.50 \times 10^{-33}$	$3.50 \times 10^{-33}$	5.04	3.99–6.24	0.70	0.94	0.85
	19:56879613-56879735	<i>ZNF542</i>	0.44	0.14	$1.00 \times 10^{-37}$	$1.50 \times 10^{-37}$	4.42	3.56–5.39	0.79	0.90	0.87
	19:53758048-53758164	<i>ZNF677</i>	0.54	0.26	$1.60 \times 10^{-34}$	$1.90 \times 10^{-34}$	3.85	3.11–4.67	0.73	0.90	0.86
	17:75369456-75369630	<i>SEPT-9</i>	0.49	0.10	$5.50 \times 10^{-45}$	$1.10 \times 10^{-45}$	4.02	3.27–4.02	0.83	0.87	0.91

<sup>a</sup>Genomic region represents the genomic coverage of the reads with targeted bisulfite sequencing, and the genomic coordinates shown here is based on the hg19 version of the genome.

<sup>b</sup>The gene name of the genomic region.

<sup>c</sup>McaM represents the mean methylation percentage of the cases in each region, which consisting of several CpG sites, while the McoM represents the mean methylation percentage of the controls in each region.

<sup>d</sup>P value is calculated through the wilcoxon rank-sum test following with FDR (false discovery rate) adjustment for multiple corrections.

<sup>e</sup>OR and 95% CI were conducted through logistic regression.

<sup>f</sup>Sens, sensitivity; while Spec, specificity; AUC, area under curve.

<sup>f</sup>The sensitivity, specificity, and the AUC were calculated through a logistic regression prediction model without adjustment for gender, age and smoking status and alcohol status.

*SEPT9* (sensitivity = 0.83, specificity = 0.87, AUC = 0.91). In addition, *ESR1* also achieved comparable diagnostic capability (sensitivity = 0.78, specificity = 0.97, AUC = 0.91) to *SEPT9* (Table 2), suggesting that these two genes have great potential in the liquid biopsy of CRC. Simultaneously, we assessed the diagnostic abilities of these ZFGs in the KRAS + subgroups and KRAS- group using all samples. Concordantly, all candidate biomarkers achieved better performance (AUC  $\geq$  0.95) in the KRAS + group than that in the KRAS- group, further indicating that KRAS- CRC patients are more epigenetically homogeneous.

To further improve the robustness of the diagnostic abilities of these ZFGs, we performed multiple machine-learning algorithms following the fivefold cross-validation method, which split the dataset into training and validation dataset randomly to obtain an unbiased estimation (see section “Materials and Methods,” Supplementary Table 13). As shown in Supplementary Table 13, the random forest (RF) model achieved the best accuracy (accuracy = 0.89) in the test dataset. In the KRAS + subgroup, the Naïve Bayes algorithm performed best on the test dataset (sensitivity = 0.99, specificity = 0.97, accuracy = 0.98), while the Neural Network model had the best performance in the KRAS- subgroup (sensitivity = 0.75, specificity = 0.89, accuracy = 0.82). Therefore, we suggested that the diagnostic abilities of these ZFGs are reliable for CRC diagnosis in test dataset.

## DISCUSSION

To our knowledge, few studies have extensively explored the methylation alterations of ZFGs in CRC. Herein, we integrated datasets from TCGA and GEO databases to identify robust and statistically powerful biomarkers. In total, we identified five hyper-methylated ZFGs as candidate biomarkers and validated them using Han Chinese CRC patients. Two ZFGs, *ZNF132*, and *ESR1*, were recognized as promising diagnostic biomarkers for CRC. Moreover, we found significantly higher diagnostic abilities of these ZFGs in the KRAS + group than the KRAS- group, suggesting the significant association between somatic mutations and DNA methylation alterations. Our results highlighted the importance of combining genetic mutations and epigenetic alterations for CRC diagnosis in further studies.

The interaction between genetic mutations and epigenetic alterations in the tumorigenesis of CRC has been reported previously. Gazin et al. (2007) performed a genome-wide RNA interference (RNAi) screening of K-ras-transformed NIH 3T3 cells and identified 28 genes that are required for Ras-mediated epigenetic silencing of the pro-apoptotic Fas gene. It was suggested that Ras-mediated epigenetic silencing could lead to CRC oncogenesis through the epigenetic inactivation of key genes. Nagasaka et al. also found that both *KRAS* and *BRAF* mutation could contribute to the global hypermethylation



phenotype of CIMP genes in colon cancer. Furthermore, Serra et al. (2014) revealed that *KRAS* mutation could result in the hypermethylation and transcriptional silencing of CIMP genes through *ZNF304*, indicating the importance of ZFPs in the carcinogenesis of CRC. In our study, the methylation profiles of the five ZFGs were significantly associated with the *KRAS* mutation status, suggesting that *KRAS* mutation may alter the downstream pathways through the epigenetic regulation of these ZFGs and required further verification.

Among the five ZFGs, *ZNF132* was identified as the most promising biomarker for CRC diagnosis in our analysis. *ZNF132* is located at 19q13.4 and belongs to the C2H2 ZFP family (Tommerup and Vissing, 1995). Previous studies have identified the DNA methylation alterations of *ZNF132* in breast cancer, esophageal squamous cell carcinoma (ESCC), oropharyngeal squamous cell carcinoma, and prostate cancer (Lleras et al., 2011; Abildgaard et al., 2012; Stefansson et al., 2015; Jiang et al., 2018). It is reported that *ZNF132* hypermethylation could reduce the Sp1 transcript factor activity and decrease the growth, migration, invasion, and tumorigenicity capabilities of cells in a nude mouse model of ESCC (Jiang et al., 2018). Our study identified the hypermethylation and down-regulation of *ZNF132* in CRC, especially in *KRAS*-mutated samples, suggesting its biological implications in CRC tumorigenesis.

*ESR1* (estrogen receptor alpha 1) has been recognized as a tumor-suppressor gene and an estrogen receptor gene. It encodes the main mediator of estrogen effects in breast epithelial cells and has also been shown to be activated by epidermal growth factor (EGF). The hyper-methylation status of *ESR1* has been reported previously in lung adenocarcinoma, breast cancer, prostate cancer, squamous cell cervical cancer, and CRC (Li et al., 2004; Lin et al., 2009; Kim et al., 2010; Elliott et al., 2013; Kirn et al., 2014; Martinez-Galan et al., 2014). *ESR1* hypermethylation is also correlated with poor prognosis and drug response in breast cancer (Ramos et al., 2010; Mastoraki et al., 2018). Additionally, the *ESR1* promoter hypermethylation has been associated with *KRAS* mutation, which is consistent with our results (Horii et al., 2009).

*ZNF229* is a protein-coding gene, and few studies have suggested the hypermethylation status of *ZNF229* in cancer diagnosis. The biological functions and their role in CRC tumorigenesis should be explored further. *ZNF542* may involve in the epigenetic regulation of puberty through transcriptional repression (Lomniczi et al., 2015). Moreover, a CpG site located at *ZNF542* has been found to be a promising biomarker for ESCC (Pu et al., 2017). A pan-cancer study revealed that *ZNF542* was significantly hypermethylated in 10 kinds of cancers (Gevaert et al., 2015). *ZNF677* is located at chromosomal region 19q13 and was found to regulate the putative tumor cell growth suppressor in non-small cell lung cancers through hypermethylation (Heller et al., 2015). In addition, *ZNF677* is frequently down-regulated by promoter methylation in primary papillary thyroid cancers (PTC), and the decreased expression of *ZNF677* is significantly correlated with poor survival (Li et al., 2018).

Currently, most CRC patients are still diagnosed at later stages, especially in developing countries, and there is an urgent need for better diagnostic biomarkers. DNA methylation alterations

may occur before mRNA and protein changes and could thus be ideal for early diagnosis. *SEPT9* has been approved as a DNA methylation-based diagnostic biomarker for CRC. The first release of the cfDNA *SEPT9* DNA methylation assay achieved considerable sensitivity (72%) and specificity (86%) in CRC detection using plasma (deVos et al., 2009). Since then, an updated version of the assay (Epi proColon 2.0) has shown better sensitivities (68–95%) and specificities (80–99%) in CRC diagnosis (Payne, 2010). However, the significant heterogeneity of CRC makes it challenging to use a single DNA methylation-based biomarker to diagnose CRC accurately. Therefore, the present study showed that *ZNF132* and *ESR1* have comparable or even better diagnostic capability than *SEPT9*, suggesting that the panel integrating *ZNF132*, *ESR1*, and *SEPT9* may better serve as a robust non-invasive diagnostic tool for CRC. In addition, the diagnostic ability of our model and the *SEPT9* model were both significantly affected by the *KRAS* mutation status in patients, and none of them achieved satisfactory sensitivities and specificities in the diagnosis of *KRAS*-CRC patients, suggesting that the mutation landscape of the patients should be taken into account to diagnose CRC accurately.

We acknowledge that our study has several limitations. First, our study's sample size and patient diversity are limited, and studies with larger sample size and diverse populations are required to identify better CRC diagnostic markers. Second, we identified the candidate ZFGs by analyzing public DNA methylation datasets using Illumina HumanMethylation 450K microarray, which only covered a small proportion of the genome while a large amount of the genome remains undetected. Further studies using whole-genome bisulfite sequencing may be required to identify better DNA methylation-based biomarkers for CRC. Finally, our study only detected the methylation profiles of these ZFGs in tumor or para-tumor tissues, while the efficacy of these ZFGs for CRC non-invasive diagnosis using plasma or stool is elusive. We will try to validate the diagnostic abilities of these ZFGs in our future studies.

## MATERIALS AND METHODS

### Integration of Public Datasets and Biomarker Discovery

Public high-throughput DNA methylation microarray datasets (Illumina HumanMethylation 450K) were searched exhaustively from TCGA and GEO databases. Two datasets from TCGA and nine datasets from GEO were included, yielding a total of 1,104 CRC tumor samples, 268 control samples, and 54 adenoma samples (**Supplementary Table 1**). The comprehensive list of genes in the zinc finger family ( $n = 1,594$ ) was obtained from HGNC (**Supplementary Table 2**).

As shown in a previous study, the methylation linkage equilibrium decreased significantly when the block was longer than 1,000 bp (Guo et al., 2017a). Therefore, we defined the methylation region (MR) to have at least 6 CpG sites with less than 1,000 bp. We then arranged all the CpG sites in the high-throughput microarray according to their genomic coordinates and applied the sliding window method to identify



all MRs. In total, we identified 6,166 MRs and performed differential methylation analyses (**Supplementary Table 14**). The 85 DMRs ( $M_{caM} > 0.50$ ,  $M_{adM} > 0.50$  and  $M_{coM} < 0.30$ ) were found for further analysis. To correct the noise from DNA originating from mixed-in peripheral blood, we suggested that the methylation rate of the DMRs need to be extremely low in the peripheral blood. Thus, we integrated the public high-throughput microarray datasets of the whole blood (WB,  $n = 1438$ ), peripheral blood mononuclear cells (PBMC,  $n = 111$ ), and peripheral blood leukocytes (PBL,  $n = 529$ ) as references for DMR identification (**Supplementary Table 15**). The 32 DMRs were retained due to their low methylation rates in WB, PBMC, or PBL (mean methylation rate  $< 0.10$ ).

Several DMRs were located at the same gene, and we selected the DMR with the most significant differences between CRC and control tissues for each gene ( $N = 10$ ). To obtain the DMRs that may regulate the expression of neighboring genes, we further selected the DMRs with transcription factor binding sites (TFBSs) that correlate significantly with the expression of neighboring genes (**Supplementary Tables 16, 17**). In total, seven out of the ten candidate DMRs were selected for validation. However, due to the difficulties in the primer design due to CG percent, PolyT, and the number of SNPs, two candidate DMRs were removed (*SALL1* and *ZSCAN23*). Finally, we obtained the top five candidate DMRs for further validation (*ESR1*, *ZNF132*, *ZNF229*, *ZNF542*, and *ZNF677*).

## Patients, Samples, and DNA

Colorectal cancer patients in both replication cohort 1 and cohort 2 were recruited from the Affiliated Hospital of Nantong University between 2016 and 2018. The earlier recruited samples ( $N = 104$ ) were included in replication cohort 1, while the others belonged to replication cohort 2. The recruited patients with available CRC tumors and matched para-tumors (the corresponding adjacent normal tissue at least 5 cm distant from the tumor tissue) were selected in this study. The patients recruited had not been treated with any neoadjuvant therapy before. Tumors were classified according to TNM (Tumor Node Metastasis)/UICC (Union for International Cancer Control) criteria following the histopathological examination. At least two professional pathologists evaluated all tumor samples carefully. All procedures performed in this study were in accordance with the ethical standards of the institutional research committee, as well as the 1964 Declaration of Helsinki and its later amendments. The study was approved by the institutional review boards of the Affiliated Hospital of Nantong University. Written informed consent was obtained from each participant of the study. All tumors and para-tumors were immediately frozen at  $-80^{\circ}\text{C}$  after surgical resection.

## Targeted Bisulfite Sequencing Assay and Detection of *KRAS*, *NRAS*, and *BRAF* Mutation Status

DNA extraction was performed using the AIIperp DNA/RNA Mini Kit (Qiagen, Duesseldorf, Germany) according to the manufacturer's protocols. The EpiTect Fast DNA Bisulfite Kit

(Qiagen, Duesseldorf, Germany) was further used for bisulfite conversion. After carefully evaluating CG percent, PolyT, and the occurrence of SNPs in the targeted regions of the candidate DMRs, we designed primers to detect them in a panel for NGS sequencing (**Supplementary Table 18**). The PCR amplicons were then diluted and amplified using these primers, and the products (170–270 bp) were further separated and purified by QIAquick Gel Extraction kit (Qiagen, Duesseldorf, Germany). Finally, the sequencing libraries from different samples were pooled together and sequenced using the Illumina HiSeq 2000 platform according to the manufacturer's protocols.

The BSseeker2 was applied for read mapping and methylation calling. After that, we removed the samples with a bisulfite conversion rate  $< 98\%$ . The methylation level of each CpG site for each sample has been provided in **Supplementary Tables 19,20**. The average coverage and missing rate for each CpG site were calculated and utilized for quality control (average coverage  $> 20\times$ , missing rate  $< 20\%$ ). In addition, the sample whose missing rate  $> 30\%$  were also filtered out. FastTarget next-generation sequencing was used to detect tumor DNA for the mutations in codons 12, 13, 59, 61, 117, and 146 of the *KRAS* and *NRAS* genes, as well as the mutation in codon 600 of the *BRAF* gene (Liao et al., 2019).

## Statistical Analysis and Machine Learning Methods

The Wilcoxon rank-sum test was performed in the discovery stage to identify the differential methylation sites and regions between CRC tumors, adenomas, and para-tumors. Moreover, the differential methylation status (odds ratios) between tumors and para-tumors of the DMRs were calculated with logistic regression. Benjamini-Hochburg correction was utilized for multiple test correction.

To identify the factors that may affect the diagnostic abilities of these ZFGs in CRC, we first split the dataset based on these covariates, including young (age  $\leq$  median age) vs. old (age  $>$  median age), male vs. female, early (Stage I/II) vs. late (Stage III/IV), colon vs. rectum and *KRAS* + vs. *KRAS*- subgroups. Second, we performed the univariate logistic regression model using each gene in these subgroups separately. The cut-off value was determined as the value which maximized the sum of both sensitivity and specificity of the model. After that, the predicted sample type for each sample was obtained, and was used to calculate the number of accurate predictions (true positive + true negative) and inaccurate predictions (false positive + false negative) of the logistic model. Finally, we utilized the Fisher exact test to explore if there are significant differences between the prediction outcomes between these subgroups.

To further give an unbiased estimation of the diagnostic ability of these ZFGs, we performed fivefold cross-validation of our dataset using the logistic regression (Package stats), support vector machine (SVM, Package e1071), random forest (Package randomForest), Naïve Bayes (Package e1071), neural network (Package nnet), linear discriminant analysis (LDA, Package mda),

mixture discriminant analysis (MDA, Package *mda*), flexible discriminant analysis (FDA, Package *mda*), gradient boosting machine (Package *gbm*), *catboost* (Package *catboost*), and *XGBoost* (Package *xgboost*) methods. We randomly selected 80% of our samples as the training dataset, and the remaining 20% of samples were used as the test dataset. After that, we optimized these machine learning methods using the training dataset and validate these optimized models in the test dataset. We repeated these procedures 1,000 times, and obtained the sensitivities, specificities and accuracies for both training and test datasets. The averaged sensitivities, specificities and accuracies for both training and test datasets were calculated and shown in **Supplementary Table 13**. All statistical analyses were conducted using R (v3.4.3).

## DATA AVAILABILITY STATEMENT

Datasets related to this article can be found at Zenodo (doi: 10.5281/zenodo.5214274) and in **Supplementary Tables 19, 20**. The analysis steps, functions, and parameters used are described in detail in section “Materials and Methods”. The custom R scripts used to analyze data and generate figures are available at a GitHub repository: [https://github.com/puweilin/CRC\\_MethylationAnalysis.git](https://github.com/puweilin/CRC_MethylationAnalysis.git).

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the institutional review boards of the Affiliated Hospital of Nantong University. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

WD, WP, SG, YM, and SJ contributed to the conception and design of the study. WD, FQ, KS, FX, QJ, SJ, RZ, JZ, JFZ, and SQJ collected the samples of the study. FQ, WD, and SQJ performed the experiments of this study. WP, WD, FQ, JL, YM, and SJ performed the statistical analysis. WP and WD wrote the first draft of the manuscript. SJ, YM, and SG revised the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

## REFERENCES

- Abildgaard, M. O., Borre, M., Mortensen, M. M., Ulhøi, B. P., Tørring, N., Wild, P., et al. (2012). Downregulation of zinc finger protein 132 in prostate cancer is associated with aberrant promoter hypermethylation and poor prognosis. *Int. J. Cancer* 130, 885–895. doi: 10.1002/ijc.26097
- Ahlquist, D. A., Taylor, W. R., Mahoney, D. W., Zou, H., Domanico, M., Thibodeau, S. N., et al. (2012). The stool DNA test is more accurate than the plasma septin 9 test in detecting colorectal neoplasia. *Clin. Gastroenterol. Hepatol.* 10, 272–277. doi: 10.1016/j.cgh.2011.10.008

## FUNDING

This study was supported by the National Natural Science Foundation of China (81974313 and 82003360), Postdoctoral Science Foundation of China (2018M641919 and 2019M651930), the Natural Science Foundation of Jiangsu Commission of Health (M2020065), the Natural Science Foundation of Higher Education of Jiangsu Province (20KJB320001), and Nantong People’s Livelihood Science and Technology Plan (MS12018032 and JC2019066).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.759813/full#supplementary-material>

**Supplementary Figure 1** | The Mean methylation rate of the candidate zinc finger genes in the CRC tumors, adenomas and control tissues. The y-axis represents the mean methylation percentage in the CRC tumor tissues and paired normal tissues for each CpG site. The error bar of each CpG site represents the confidence interval of the methylation percentage.

**Supplementary Figure 2** | Bisulfite conversion rate and mapping rate between CRC tumors and para-tumors. Panels (A,B) represent the bisulfite conversion rate and the reads mapping rate of the samples in replication cohort 1, respectively.

**Supplementary Figure 3** | Principal Component Analysis of the methylation profiles of the samples in replication cohort 1. All of the CpG sites were preprocessed and simplified into the two-variable space through principal component analysis. The green and red dots represented the CRC tumors and paired controls in replication cohort 1.

**Supplementary Figure 4** | Boxplot of the methylation rate of the KRAS +, KRAS- and control samples in replication cohort 1. The mean methylation rate of all the CpG sites in each candidate gene of each sample was depicted as one dot in the boxplot.

**Supplementary Figure 5** | Bisulfite conversion rate and mapping rate between CRC tumors and para-tumors. Panels (A,B) represent the bisulfite conversion rate and the reads mapping rate of the samples in replication cohort 2, respectively.

**Supplementary Figure 6** | Principal Component Analysis of the methylation profiles of the samples in replication cohort 1. All of the CpG sites were preprocessed and simplified into the two-variable space through principal component analysis. The green and red dots represented the CRC tumors and paired controls in replication cohort 2.

**Supplementary Figure 7** | Boxplot of the methylation rate of the KRAS +, KRAS- and control samples in replication cohort 2. The mean methylation rate of all the CpG sites in each candidate gene of each sample was depicted as one dot in the boxplot.

- Barault, L., Amatu, A., Siravegna, G., Ponzetti, A., Moran, S., Cassingena, A., et al. (2018). Discovery of methylated circulating DNA biomarkers for comprehensive non-invasive monitoring of treatment response in metastatic colorectal cancer. *Gut* 67, 1995–2005. doi: 10.1136/gutjnl-2016-313372
- Church, T. R., Wandell, M., Lofton-Day, C., Mongin, S. J., Burger, M., Payne, S. R., et al. (2014). Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut* 63, 317–325. doi: 10.1136/gutjnl-2012-304149
- deVos, T., Tetzner, R., Model, F., Weiss, G., Schuster, M., Distler, J., et al. (2009). Circulating methylated SEPT9 DNA in plasma is a biomarker for

- colorectal cancer. *Clin. Chem.* 55, 1337–1346. doi: 10.1373/clinchem.2008.115808
- Ding, W., Pu, W., Wang, L., Jiang, S., Zhou, X., Tu, W., et al. (2018). Genome-wide DNA methylation analysis in systemic sclerosis reveals hypomethylation of IFN-Associated Genes in CD4(+) and CD8(+) T Cells. *J. Invest. Dermatol.* 138, 1069–1077. doi: 10.1016/j.jid.2017.12.003
- Elliott, G. O., Johnson, I. T., Scarll, J., Dainty, J., Williams, E. A., Garg, D., et al. (2013). Quantitative profiling of CpG island methylation in human stool for colorectal cancer detection. *Int. J. Colorectal. Dis.* 28, 35–42. doi: 10.1007/s00384-012-1532-5
- Freitas, M., Ferreira, F., Carvalho, S., Silva, F., Lopes, P., Antunes, L., et al. (2018). A novel DNA methylation panel accurately detects colorectal cancer independently of molecular pathway. *J. Transl. Med.* 16:45. doi: 10.1186/s12967-018-1415-9
- Gazin, C., Wajapeyee, N., Gobeil, S., Virbasius, C. M., and Green, M. R. (2007). An elaborate pathway required for Ras-mediated epigenetic silencing. *Nature* 449, 1073–1077. doi: 10.1038/nature06251
- Gevaert, O., Tibshirani, R., and Plevritis, S. K. (2015). Pancancer analysis of DNA methylation-driven genes using MethylMix. *Genome Biol.* 16:17. doi: 10.1186/s13059-014-0579-8
- Guo, S., Diep, D., Plongthongkum, N., Fung, H. L., Zhang, K., and Zhang, K. (2017a). Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. *Nat. Genet.* 49, 635–642. doi: 10.1038/ng.3805
- Guo, S., Zhu, Q., Jiang, T., Wang, R., Shen, Y., Zhu, X., et al. (2017b). Genome-wide DNA methylation patterns in CD4+ T cells from Chinese Han patients with rheumatoid arthritis. *Mod. Rheumatol.* 27, 441–447. doi: 10.1080/14397595.2016.1218595
- Guo, S., Yan, F., Xu, J., Bao, Y., Zhu, J., Wang, X., et al. (2015). Identification and validation of the methylation biomarkers of non-small cell lung cancer (NSCLC). *Clin. Epigenetics* 7:3. doi: 10.1186/s13148-014-0035-3
- He, Y., Cui, Y., Wang, W., Gu, J., Guo, S., Ma, K., et al. (2011). Hypomethylation of the hsa-miR-191 locus causes high expression of hsa-mir-191 and promotes the epithelial-to-mesenchymal transition in hepatocellular carcinoma. *Neoplasia* 13, 841–853. doi: 10.1593/neo.11698
- Heller, G., Altenberger, C., Schmid, B., Marhold, M., Tomasich, E., Ziegler, B., et al. (2015). DNA methylation transcriptionally regulates the putative tumor cell growth suppressor ZNF677 in non-small cell lung cancers. *Oncotarget* 6, 394–408. doi: 10.18632/oncotarget.2697
- Hori, J., Hiraoka, S., Kato, J., Saito, S., Harada, K., Fujita, H., et al. (2009). Methylation of estrogen receptor 1 in colorectal adenomas is not age-dependent, but is correlated with K-ras mutation. *Cancer Sci.* 100, 1005–1011. doi: 10.1111/j.1349-7006.2009.01140.x
- Jen, J., and Wang, Y. C. (2016). Zinc finger proteins in cancer progression. *J. Biomed. Sci.* 23:53. doi: 10.1186/s12929-016-0269-9
- Jiang, D., He, Z., Wang, C., Zhou, Y., Li, F., Pu, W., et al. (2018). Epigenetic silencing of ZNF132 mediated by methylation-sensitive Sp1 binding promotes cancer progression in esophageal squamous cell carcinoma. *Cell Death Dis.* 10:1. doi: 10.1038/s41419-018-1236-z
- Kim, M. S., Lee, J., and Sidransky, D. (2010). DNA methylation markers in colorectal cancer. *Cancer Metastasis Rev.* 29, 181–206. doi: 10.1007/s10555-010-9207-6
- Kirn, V., Zaharieva, I., Heublein, S., Thangarajah, F., Friese, K., Mayr, D., et al. (2014). ESR1 promoter methylation in squamous cell cervical cancer. *Anticancer Res.* 34, 723–727.
- Koch, A., Joosten, S. C., Feng, Z., de Ruijter, T. C., Draht, M. X., Melotte, V., et al. (2018). Analysis of DNA methylation in cancer: location revisited (vol 15, pg 459, 2018). *Nat. Rev. Clin. Oncol.* 15, 467–467. doi: 10.1038/s41571-018-0028-9
- Li, L. C., Shiina, H., Deguchi, M., Zhao, H., Okino, S. T., Kane, C. J., et al. (2004). Age-dependent methylation of ESR1 gene in prostate cancer. *Biochem. Biophys. Res. Commun.* 321, 455–461. doi: 10.1016/j.bbrc.2004.06.164
- Li, Y., Yang, Q., Guan, H., Shi, B., Ji, M., and Hou, P. (2018). ZNF677 suppresses Akt phosphorylation and tumorigenesis in thyroid cancer. *Cancer Res.* 78, 5216–5228. doi: 10.1158/0008-5472.CAN-18-0003
- Liao, X., Zhang, T., Li, B., Hu, S., Liu, J., Deng, J., et al. (2019). Rare RNF213 variants and the risk of intracranial artery stenosis/occlusion disease in Chinese population: a case-control study. *BMC Med. Genet.* 20:55. doi: 10.1186/s12881-019-0788-9
- Lin, Q., Geng, J., Ma, K., Yu, J., Sun, J., Shen, Z., et al. (2009). RASSF1A, APC, ESR1, ABCB1 and HOXC9, but not p16INK4A, DAPK1, PTEN and MT1G genes were frequently methylated in the stage I non-small cell lung cancer in China. *J. Cancer Res. Clin. Oncol.* 135, 1675–1684. doi: 10.1007/s00432-009-0614-4
- Lleras, R. A., Adrien, L. R., Smith, R. V., Brown, B., Jivraj, N., Keller, C., et al. (2011). Hypermethylation of a cluster of Kruppel-type zinc finger protein genes on chromosome 19q13 in oropharyngeal squamous cell carcinoma. *Am. J. Pathol.* 178, 1965–1974. doi: 10.1016/j.ajpath.2011.01.049
- Lomniczi, A., Wright, H., Castellano, J. M., Matagne, V., Toro, C. A., Ramaswamy, S., et al. (2015). Epigenetic regulation of puberty via Zinc finger protein-mediated transcriptional repression. *Nat. Commun.* 6:10195. doi: 10.1038/ncomms10195
- Martinez-Galan, J., Torres-Torres, B., Nunez, M. I., Lopez-Penalver, J., Del Moral, R., Ruiz De Almodovar, J. M., et al. (2014). ESR1 gene promoter region methylation in free circulating DNA and its correlation with estrogen receptor protein expression in tumor tissue in breast cancer patients. *BMC Cancer* 14:59. doi: 10.1186/1471-2407-14-59
- Mastoraki, S., Strati, A., Tzanikou, E., Chimonidou, M., Politaki, E., Voutsina, A., et al. (2018). ESR1 Methylation: A liquid biopsy-based epigenetic assay for the follow-up of patients with metastatic breast cancer receiving endocrine treatment. *Clin. Cancer Res.* 24, 1500–1510. doi: 10.1158/1078-0432.CCR-17-1181
- Payne, S. R. (2010). From discovery to the clinic: the novel DNA methylation biomarker (m)SEPT9 for the detection of colorectal cancer in blood. *Epigenomics* 2, 575–585. doi: 10.2217/epi.10.35
- Pu, W., Wang, C., Chen, S., Zhao, D., Zhou, Y., Ma, Y., et al. (2017). Targeted bisulfite sequencing identified a panel of DNA methylation-based biomarkers for esophageal squamous cell carcinoma (ESCC). *Clin. Epigenetics* 9:129. doi: 10.1186/s13148-017-0430-7
- Ramos, E. A., Camargo, A. A., Braun, K., Slowik, R., Cavalli, I. J., Ribeiro, E. M., et al. (2010). Simultaneous CXCL12 and ESR1 CpG island hypermethylation correlates with poor prognosis in sporadic breast cancer. *BMC Cancer* 10:23. doi: 10.1186/1471-2407-10-23
- Schubeler, D. (2015). Function and information content of DNA methylation. *Nature* 517, 321–326. doi: 10.1038/nature14192
- Serra, R. W., Fang, M., Park, S. M., Hutchinson, L., and Green, M. R. (2014). A KRAS-directed transcriptional silencing pathway that mediates the CpG island methylator phenotype. *Elife* 3:e02313. doi: 10.7554/eLife.02313
- Siegel, R. L., Miller, K. D., Fedewa, S. A., Ahnen, D. J., Meester, R. G. S., Barzi, A., et al. (2017). Colorectal cancer statistics, 2017. *CA Cancer J. Clin.* 67, 177–193. doi: 10.3322/caac.21395
- Siegel, R. L., Miller, K. D., Goding Sauer, A., Fedewa, S. A., Butterly, L. F., Anderson, J. C., et al. (2020). Colorectal cancer statistics, 2020. *CA Cancer J. Clin.* 70, 145–164. doi: 10.3322/caac.21601
- Song, L., Jia, J., Peng, X., Xiao, W., and Li, Y. (2017). The performance of the SEPT9 gene methylation assay and a comparison with other CRC screening tests: A meta-analysis. *Sci. Rep.* 7:3032. doi: 10.1038/s41598-017-0321-8
- Stefansson, O. A., Moran, S., Gomez, A., Sayols, S., Arribas-Jorba, C., Sandoval, J., et al. (2015). A DNA methylation-based definition of biologically distinct breast cancer subtypes. *Mol. Oncol.* 9, 555–568. doi: 10.1016/j.molonc.2014.10.012
- Sun, J., Fei, F., Zhang, M., Li, Y., Zhang, X., Zhu, S., et al. (2019). The role of (m)SEPT9 in screening, diagnosis, and recurrence monitoring of colorectal cancer. *BMC Cancer* 19:450. doi: 10.1186/s12885-019-5663-8
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71, 209–249. doi: 10.3322/caac.21660
- Tommerup, N., and Vissing, H. (1995). Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identify putative candidate genes for developmental and malignant disorders. *Genomics* 27, 259–264. doi: 10.1006/geno.1995.1040

Wills, B., Gorse, E., and Lee, V. (2018). Role of liquid biopsies in colorectal cancer. *Curr. Probl. Cancer* 42, 593–600. doi: 10.1016/j.currprobcancer.2018.08.004

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Pu, Qian, Liu, Shao, Xiao, Jin, Liu, Jiang, Zhang, Zhang, Guo, Zhang, Ma, Ju and Ding. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*