

Research Article

Discovery and Validation of Hypermethylated Markers for Colorectal Cancer

Jiufeng Wei,^{1,2} Guodong Li,^{1,2} Shuwei Dang,^{1,2} Yuhui Zhou,^{1,2} Kai Zeng,^{1,2} and Ming Liu^{1,2}

¹Department of General Surgery, The Fourth Affiliated Hospital of Harbin Medical University, Harbin 150001, China

²Bio-Bank of Department of General Surgery, The Fourth Affiliated Hospital of Harbin Medical University, Harbin 150001, China

Correspondence should be addressed to Ming Liu; mliu35@aliyun.com

Received 31 March 2016; Revised 6 June 2016; Accepted 15 June 2016

Academic Editor: Monica Neagu

Copyright © 2016 Jiufeng Wei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Colorectal carcinoma (CRC) is one of the most prevalent malignant tumors worldwide. Screening and early diagnosis are critical for the clinical management of this disease. DNA methylation changes have been regarded as promising biomarkers for CRC diagnosis. Here, we map DNA methylation profiling on CRC in six CRCs and paired normal samples using a 450 K bead array. Further analysis confirms the methylation status of candidates in two data sets from the Gene Expression Omnibus. Receiver operating characteristic (ROC) curves are calculated to determine the diagnostic performances. We identify 1549 differentially methylated regions (DMRs) showing differences in methylation between CRC and normal tissue. Two genes (*ADD2* and *AKR1B1*), related to the DMRs, are selected for further validation. ROC curves show that the areas under the curves of *ADD2* and *AKR1B1* are higher than that of *SEPT9*, which has been clinically used as a screening biomarker of CRC. Our data suggests that aberrant DNA methylation of *ADD2* and *AKR1B1* could be potential screening markers of CRC.

1. Introduction

Colorectal carcinoma (CRC) is one of the most prevalent malignant tumors worldwide. Global statistics showed that in 2012 alone, an estimated 1.36 million new cases were diagnosed with CRC, and approximately 694,000 people died from this disease [1]. Screening and early diagnosis are critical for the clinical management of CRC. Traditional screening tools include fecal occult blood test (FOBT) and colonoscopy. However, the effectiveness of FOBT is limited by the test performance, while colonoscopy is invasive, and it is therefore impractical to screen all patients for CRC in this manner. The identification of highly specific, noninvasive biomarkers is a top priority for screening and early diagnosis of CRC.

Aberrant DNA methylation is a well-recognized epigenetic feature of cancer, in general, and has been discovered in most tumors; it is thus gaining increasing attention as a potential biomarker [2–4]. Abnormally methylated genes can be used as biomarkers for early detection as well as tumor classification of CRC [5–7]. Some of these alterations have also been detected in stool or peripheral blood, suggesting

that they can be candidates for noninvasive biomarkers of CRC. Epi proColon®, a blood-based assay for measuring methylated *SEPT9*, has become available for clinical application and has been approved by China and Europe. However, the sensitivity and specificity are still not satisfactory [5]. Novel biomarkers are needed to improve the accuracy of diagnosis of CRC.

In this study, the genome-wide methylation pattern of CRC was compared with adjacent normal tissues using the Illumina 450 K microarray, thus revealing aberrantly differentially methylated regions (DMRs) in CRC. Among the list of DMRs that we identified, potential biomarkers were validated in two independent data sets. We also established the sensitivity and specificity of the new molecular markers, which showed a higher area under the curve (AUC) than *SEPT9*. These biomarkers could improve the accuracy of CRC screening and diagnosis.

2. Materials and Methods

2.1. Subjects. Six pairs of CRC and adjacent normal tissues were obtained from the Bio-Bank of the Department of

TABLE 1: The list of top 10 DMRs by the P value.

Number	Start	End	Size	CHR	Arm	Gene	Feature	cgi	P value	Average deltaBeta
1	29520841	29521887	1047	6	p	NA	IGR	Shore	$2.17E - 71$	0.297
2	130130327	130132504	2178	7	q	MESTIT1	Body	Shore	$4.20E - 63$	0.118
3	33140275	33148582	8308	6	p	COL11A2	Body	Open sea	$5.36E - 42$	-0.147
4	33130824	33138475	7652	6	p	COL11A2	3'UTR	Shore	$6.93E - 40$	-0.134
5	27139876	27142774	2899	7	p	HOXA2	Body	Shelf	$3.18E - 35$	0.232
6	133561800	133562545	746	6	q	EYA4	TSS1500	Shore	$2.09E - 34$	0.306
7	31938678	31939388	711	6	p	STK19	TSS1500	Shore	$9.84E - 32$	-0.149
8	32183994	32190096	6103	6	p	NOTCH4	Body	Open sea	$7.50E - 31$	-0.151
9	78493049	78493778	730	13	q	EDNRB	5'UTR	Island	$9.47E - 31$	0.296
10	30078146	30080891	2746	6	p	TRIM31	Body	Open sea	$1.47E - 26$	-0.196

General Surgery of the Fourth Affiliated Hospital of Harbin Medical University. Inclusion criteria were no cancer other than CRC, no indications of heredity, and no radio- or chemotherapy prior to surgical resection. This study was approved by the Medical Research Ethics Committee of the Fourth Affiliated Hospital of Harbin Medical University, and informed consent was obtained. The diagnosis of CRC tissues was acquired from pathology reports. Fresh tissue samples were collected within 30 min after resection surgery, frozen in liquid nitrogen, and stored at -80°C . Clinicopathological characteristics of CRC patients are shown in Table S1 in Supplementary Material available online at <http://dx.doi.org/10.1155/2016/2192853>.

2.2. Human Methylation 450 K Microarray. Genomic DNA was extracted using standard phenol-chloroform techniques and quantified using Nanodrop 2000c. Genomic DNA from all samples was treated with an EZ Methylation Kit (Zymo Labs, Irvine, CA). Bisulfite-converted DNA (500 ng) was hybridized onto the Infinium Human Methylation 450 K BeadChip according to the manufacturer's standard protocol.

2.3. Differential Methylation Region Analysis. Infinium Methylation data were processed with the Methylation Module of the GenomeStudio software. Methylation levels of CpG sites were calculated as β -values (0-1). We removed unreliable probes that were detected with a P value > 0.05 . In addition, CpG sites were removed on the X and Y chromosomes, containing single-nucleotide polymorphisms. The methylation data were deposited in the NCBI Gene Expression Omnibus (GEO): GSE75546. DMRs were analyzed using the ChAMP package, according to the instruction manual. To help identify regions of realistic length, the search was only conducted in regions where the distance between consecutive probes was less than 1 kb. The average β -values of the probes in the DMR were used as a representative of the DMR methylation levels. To screen the candidate DMR, the following criteria were used: β -difference > 0.4 , β -value in normal tissue < 0.15 , and P value $< 1E - 4$.

2.4. Data Set for Validation of Candidate Biomarkers. Methylation of candidate markers was evaluated in the data sets

GSE48684 (147 samples containing CRC, adenoma, and normal tissues) and GSE68060 (118 samples containing CRC and normal tissues) from the GEO. The methylation status of these samples was determined using the same version of the 450 K methylation array.

2.5. Statistical Analysis. Statistical analysis was conducted using the GraphPad Prism 6 software (La Jolla, CA, USA) and MedCalc version 10.1.6 (MedCalc Software, Mariakerke, Belgium). The Mann-Whitney U test was used to compare methylation levels between CRC, adenoma, and normal tissue. All reported P values were two-sided, with $P < 0.05$ being considered statistically significant. ROC analysis was performed by MedCalc.

3. Results

3.1. DMRs in Tumors versus Adjacent Normal Tissues. To identify DMRs related to colorectal carcinogenesis, whole genome DNA methylation analysis was performed using the ChAMP package with the Illumina 450 K bead array. Through this method, 1549 DMRs were identified with significant methylation differences between the six pairs of CRCs and adjacent normal samples. The top ten DMRs, according to P values, are shown in Table 1. A gene-based variant of the region-level test was performed, revealing 629 DMRs located in the promoter region, 207 DMRs located in the 5' Untranslated Regions (UTR), 117 DMRs located at the 1st Exon, 592 DMRs located in the gene body, 99 DMRs located in the 3'UTR, and 841 DMRs located in the intergenic region. A CpG island-based variant of the region-level test was conducted in the same manner, revealing 391 DMRs distributed in CpG islands, 585 DMRs distributed within the shores, 347 DMRs distributed in the shelves, and 1162 DMRs distributed in the open sea (Table S2).

We calculated the hypermethylated and hypomethylated DMRs as shown in Figure 1. Hypermethylated DMRs were mainly located in the promoter region and CpG islands (Figure 1(a)), while hypomethylated DMRs were mainly located in the intergenic region and open sea (Figure 1(b)). Interestingly, most of the hypermethylated DMRs were less than 400 bp, but most of the hypomethylated DMRs were greater than 1200 bp (Figure 1(c)).

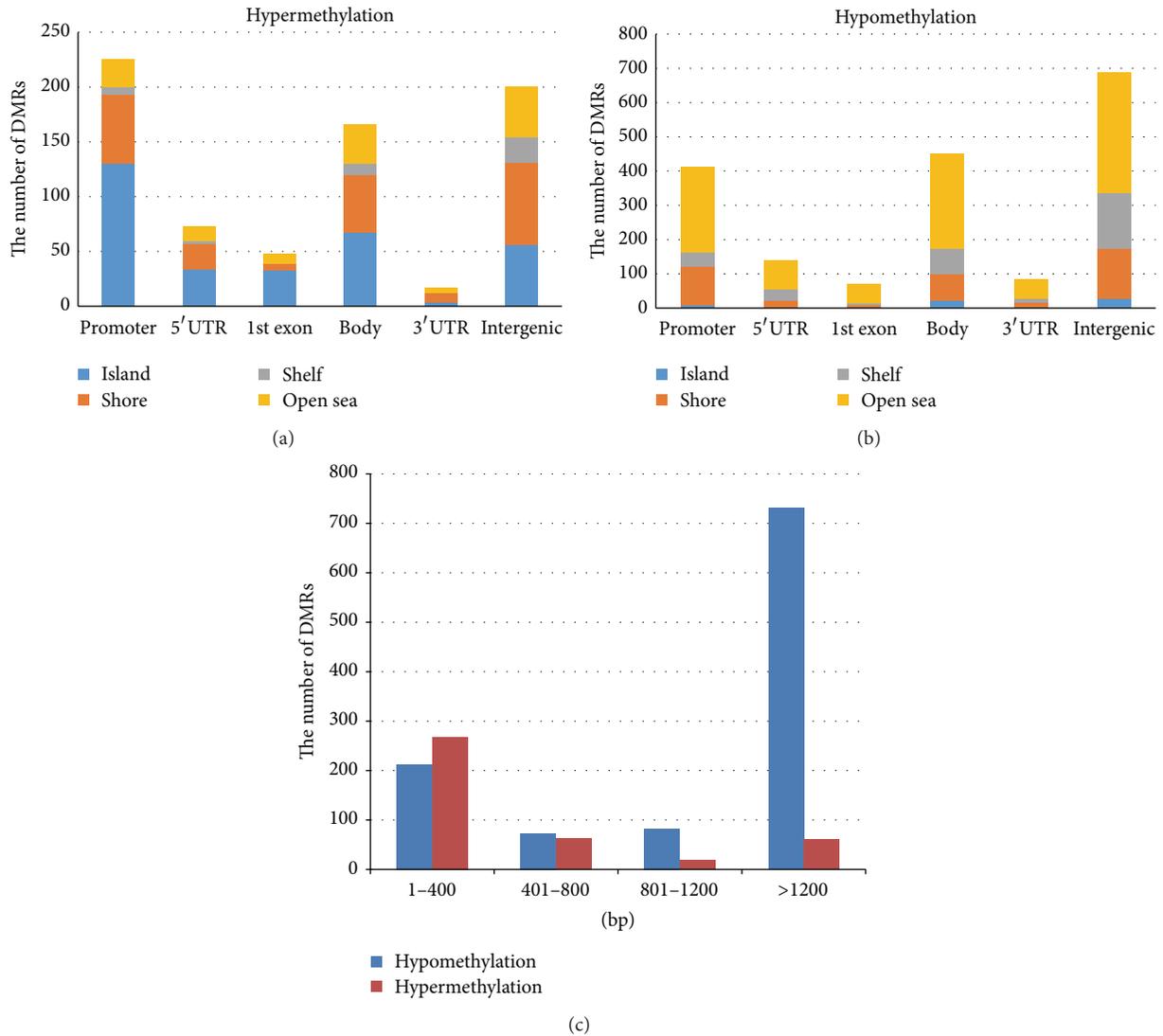


FIGURE 1: Distribution of DMRs in the human CRC genome. (a) Stacked bar charts showing the distribution of the hypermethylated DMRs over five gene categories: promoter, 5'UTR, 1st exon, gene body, 3'UTR, and intergenic regions. (b) Stacked bar charts showing the distribution of the hypomethylated DMRs over five gene categories: promoter, 5'UTR, 1st exon, gene body, 3'UTR, and intergenic regions. (c) Bar charts showing the distribution of the hypermethylated and hypomethylated DMRs considering four sizes: 1-400 bp, 401-800 bp, 801-1200 bp, and >1200 bp.

3.2. Identification of Candidate DNA Methylation Markers. To identify the candidate DNA methylation markers, the following criteria were used: β -difference > 0.4, β -value in normal tissue < 0.15, and P value < $1E - 4$. Identification was restricted to hypermethylated DMRs as these can be easily transferred to clinical application with Methylation-Specific Polymerase Chain Reaction (MSP). After evaluating all DMRs, three DMRs were identified that met all the criteria. Information regarding these three DMRs is shown in Table 2. Of these three DMRs-related genes, SEPT9 has been clinically used as a screening biomarker. In the present study, SEPT9 was used as a reference.

3.3. In Silico Validation of Selected Candidates. To investigate the selected DNA methylation candidates, we used two

independent data sets, namely, GSE48684 and GSE68060, from GEO. GSE48684 contained 41 normal tissues, 42 adenomas, and 64 CRCs. GSE68060 contained 36 normal tissues and 82 CRCs. The data set GSE48684 was generated from the 450 K methylation array. Technical and biological validation studies were conducted to demonstrate that the data were reproducible and robust [7]. The methylation levels of three DMRs in normal tissue, adenoma, and CRC in the two data sets are shown in Figure 2. In CRCs or adenomas, all candidates had significantly higher methylation levels compared to normal tissues ($P < 0.0001$). However, there were no significant differences in methylation levels of the three candidate DMRs between CRCs and adenomas.

3.4. Performance of Selected Aberrant DNA Methylation as Potential Diagnostic Markers. To determine whether these

TABLE 2: The list of candidate DMRs for validation.

Number	Start	End	Size	CHR	Arm	Gene	Feature	cgi	<i>P</i> value	Average deltaBeta
1	70995426	70995462	37	2	p	ADD2	TSS200	Shore	$2.22E - 08$	0.41
2	75369210	75369237	28	17	q	SEPT9	TSS200	Island	$3.74E - 05$	0.41
3	134144040	134144184	145	7	q	AKR1B1	TSS200	Island	$1.09E - 07$	0.41

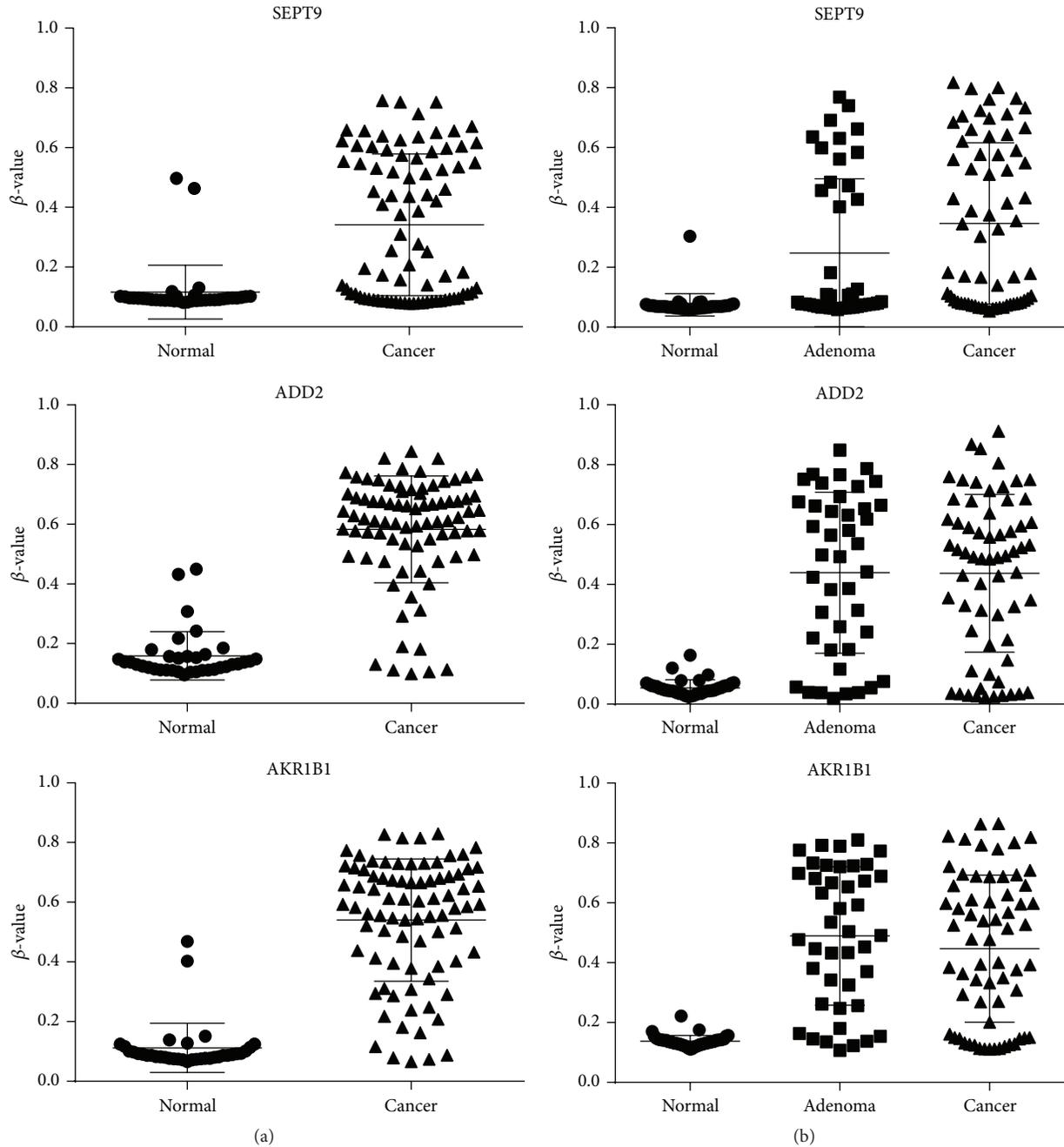


FIGURE 2: The different methylation levels of three genes in normal tissue, adenoma, and CRC. (a) The strip plot shows the different methylation levels of three genes (SEPT9, ADD2, and AKR1B1) in normal tissue and CRC with the data set GSE68060 from GEO. (b) The strip plot shows the different methylation level of three genes (SEPT9, ADD2, and AKR1B1) in normal tissue, adenoma, and CRC with the data set GSE48684 from GEO.

candidates could be potential biomarkers for use in diagnostics to distinguish between normal tissue and adenomas and CRC, we calculated AUC values for all three candidates in two data sets individually. In data set GSE48684, three markers had AUCs of 0.850 (*ADD2*, 95% CI: 0.767–0.912), 0.840 (*AKR1B1*, 95% CI: 0.756–0.904), and 0.877 (*SEPT9*, 95% CI: 0.798–0.933) between normal tissues and CRCs (Figure 3(a)). The same three markers had AUCs of 0.862 (*ADD2*, 95% CI: 0.796–0.913), 0.874 (*AKR1B1*, 95% CI: 0.809–0.923), and 0.840 (*SEPT9*, 95% CI: 0.770–0.895) between normal tissues and adenomas + CRCs (Figure 3(b)). In data set GSE68060, the three markers showed AUCs of 0.982 (*ADD2*, 95% CI: 0.935–0.997), 0.954 (*AKR1B1*, 95% CI: 0.895–0.985), and 0.752 (*SEPT9*, 95% CI: 0.659–0.831) between normal tissues and CRCs (Figure 3(c)). Two candidates, *ADD2* and *AKR1B1*, both showed better performances than *SEPT9*.

4. Discussion

Screening and early diagnosis is crucially important in the clinical management of CRC. Currently, colonoscopy and FOBT are the main approaches for CRC detection [8]. However, half of all CRCs are only detected at the advanced stages.

The widespread occurrence of modifications in CRC has major potential for being utilized as molecular markers, since alterations in DNA methylation in CRC was described by Fearon and Vogelstein over 20 years ago [9]. Compared with normal tissues, even adenomas showed apparent aberrant DNA methylation. Many aberrant DNA methylations have been reported as potential markers of CRC, such as *SEPT9*, *NDRG4*, and *VIM* [5, 10, 11]. To date, a blood-based assay named Epi proColon (Epigenomics AG, Berlin, Germany), which detects methylated *SEPT9*, has been applied clinically in several countries [12–14]. However, the sensitivity and specificity of *SEPT9* detection are still unsatisfactory. In a prospective clinical trial, sensitivity was 68% for all stages of CRC and 64% for CRC stages I–III, and much lower (22%) for advanced adenoma. In the present investigation, we performed a biomarker discovery and validation study to find new DNA methylation markers, which can be used for screening and diagnosing CRC.

Initially, we mapped the genome-wide methylation pattern of CRC compared with adjacent normal tissues using a 450 K bead chip and performed DMR analysis; this revealed that hypermethylation mainly occurred in CpG islands and promoter regions, while hypomethylation mainly occurred in the open sea and intergenic regions. These observations are in accordance with previous studies [15, 16]. We also found that most hypermethylated regions were short fragments (<400 bp), whereas most hypomethylated regions were long fragments (>1200 bp). These results suggest that hypermethylation occurs on a small scale and hypomethylation occurs on a large scale in CRC.

From the list of DMRs, we selected the candidates that most closely matched our criteria, which were set based on the premise that hypermethylation candidates are obviously better suited than hypomethylation ones for further clinical application. One of the candidates is *SEPT9*, which has

already been applied clinically. The protein *ADD2* is a subunit of adducin, a cytoskeletal protein, which caps and stabilizes the fast-growing end of actin filaments. *ADD2* is usually expressed in the nervous system and erythroid tissues [17–19]. For the first time, the present study describes the hypermethylation in the promoter region of the *ADD2* gene in malignancy. The aberrant methylation of *AKR1B1* in CRC has been previously reported [20, 21]. However, the role of *AKR1B1* as a potential biomarker has not yet been demonstrated. Therefore, the two candidates (*ADD2*, *AKR1B1*) were compared with *SEPT9*, and the performances of *ADD2* and *AKR1B1* were further evaluated for their potential as biomarkers.

The Infinium Methylation 450 K bead array is the new generation of the Methylation 27 K bead array, which contains high density methylation probes with a distribution over the entire genome. Many investigations have demonstrated the accuracy and reproducibility of this technology and have shown that the results of the Infinium Methylation 450 K Bead Chip had a good positive correlation with bisulfite sequencing [22, 23]. In the present study, two independent data sets of the 450 K bead array from GEO were used for *in silico* validation. ROC curves were performed to determine the performance of the selected candidates. In the data set GSE48684, *ADD2* and *AKR1B1* have similar AUCs to *SEPT9* when CRCs are compared to normal tissues. Comparing adenoma + CRC with normal tissues, *ADD2* and *AKR1B1* have higher AUCs than *SEPT9*. When comparing CRC and normal tissues in GSE68060, *ADD2* and *AKR1B1* also have higher AUCs than *SEPT9*. These results suggest aberrant methylations of *ADD2* and *AKR1B1* may have better screening and diagnostic performances in the early detection of CRC than *SEPT9* alone. These findings should be confirmed by additional studies, for example, by carrying out tests in stools or blood.

5. Conclusions

In summary, we conducted investigations into the discovery of tissue biomarkers to identify DNA methylation markers associated with CRC and then replicated the findings in two independent sets from GEO. Further studies are required to confirm these results and understand the role of these genes in colorectal carcinogenesis.

Abbreviations

AUC:	Area under curve
CRC:	Colorectal carcinoma
DMR:	Differentially methylated region
FOBT:	Fecal occult blood testing
GEO:	Gene Expression Omnibus
MSP:	Methylation-Specific Polymerase Chain Reaction
ROC:	Receiver operating characteristic
UTR:	Untranslated Regions.

Disclosure

Guodong Li is a co-first author.

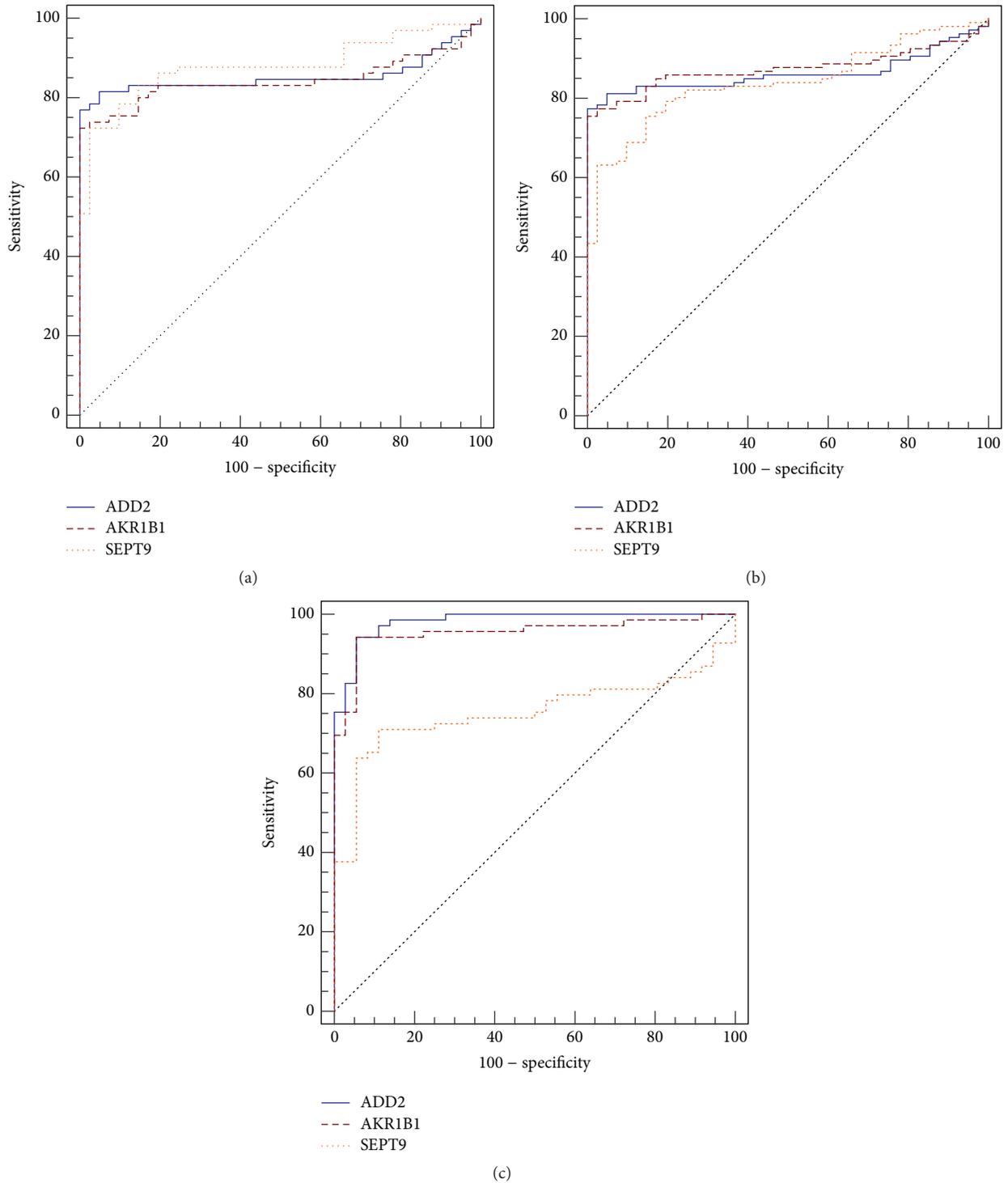


FIGURE 3: ROC curves of ADD2/AKR1B1/SEPT9 methylation using the data sets GSE48684 and GSE68060. (a) and (b) Receiver operating characteristic (ROC) curve showing accurate discrimination based on ADD2/AKR1B1/SEPT9 methylation using the data set GSE48684. (a) Between normal tissue and CRC tissue samples. (b) Between normal tissue and adenoma + CRC tissue samples. (c) ROC curve showing accurate discrimination based on ADD2/AKR1B1/SEPT9 methylation between normal tissue and CRC tissue samples using the data set GSE68060.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contributions

Jiufeng Wei and Guodong Li contributed equally as first authors.

Acknowledgments

This work is supported by National Natural Science Foundation of China (Grant no. 81372612); National Natural Science Foundation of China (Grant no. 81302059); Natural Science Foundation of Heilongjiang Province of China (Grant no. LC2013C35); Foundation of Educational Committee of Heilongjiang Province of China (Grant no. 12541300).

References

- [1] J. Ferlay, I. Soerjomataram, R. Dikshit et al., "Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012," *International Journal of Cancer*, vol. 136, no. 5, pp. E359–E386, 2015.
- [2] J. Shen, S. Wang, Y.-J. Zhang et al., "Genome-wide DNA methylation profiles in hepatocellular carcinoma," *Hepatology*, vol. 55, no. 6, pp. 1799–1808, 2012.
- [3] J. Botling, K. Edlund, M. Lohr et al., "Biomarker discovery in non-small cell lung cancer: integrating gene expression profiling, meta-analysis, and tissue microarray validation," *Clinical Cancer Research*, vol. 19, no. 1, pp. 194–204, 2013.
- [4] F. Haller, J. D. Zhang, E. A. Moskalev et al., "Combined DNA methylation and gene expression profiling in gastrointestinal stromal tumors reveals hypomethylation of SPP1 as an independent prognostic factor," *International Journal of Cancer*, vol. 136, no. 5, pp. 1013–1023, 2015.
- [5] N. T. Potter, P. Hurban, M. N. White et al., "Validation of a real-time PCR-based qualitative assay for the detection of methylated SEPT9 DNA in human plasma," *Clinical Chemistry*, vol. 60, no. 9, pp. 1183–1191, 2014.
- [6] F. J. Carmona, D. Azuara, A. Berenguer-Llergo et al., "DNA methylation biomarkers for noninvasive diagnosis of colorectal cancer," *Cancer Prevention Research*, vol. 6, no. 7, pp. 656–665, 2013.
- [7] Y. Luo, C.-J. Wong, A. M. Kaz et al., "Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer," *Gastroenterology*, vol. 147, no. 2, pp. 418–429.e8, 2014.
- [8] D. Moiel and J. Thompson, "Early detection of colon cancer—the kaiser permanente northwest 30-year history: how do we measure success? Is it the test, the number of tests, the stage, or the percentage of screen-detected patients?" *The Permanente Journal*, vol. 15, no. 4, pp. 30–38, 2011.
- [9] E. R. Fearon and B. Vogelstein, "A genetic model for colorectal tumorigenesis," *Cell*, vol. 61, no. 5, pp. 759–767, 1990.
- [10] X. Wang, Y.-Y. Kuang, and X.-T. Hu, "Advances in epigenetic biomarker research in colorectal cancer," *World Journal of Gastroenterology*, vol. 20, no. 15, pp. 4276–4287, 2014.
- [11] W.-D. Chen, Z. J. Han, J. Skoletsky et al., "Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene," *Journal of the National Cancer Institute*, vol. 97, no. 15, pp. 1124–1132, 2005.
- [12] M.-B. W. Ørntoft, H. J. Nielsen, T. F. Ørntoft, and C. L. Andersen, "Performance of the colorectal cancer screening marker Sept9 is influenced by age, diabetes and arthritis: a nested case-control study," *BMC Cancer*, vol. 15, no. 1, article 819, 2015.
- [13] P. Jin, Q. Kang, X. Wang et al., "Performance of a second-generation methylated SEPT9 test in detecting colorectal neoplasm," *Journal of Gastroenterology and Hepatology (Australia)*, vol. 30, no. 5, pp. 830–833, 2015.
- [14] A. Adler, S. Geiger, A. Keil et al., "Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany," *BMC Gastroenterology*, vol. 14, article 183, 2014.
- [15] V. A. Naumov, E. V. Generozov, N. B. Zaharjevskaya et al., "Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips," *Epigenetics*, vol. 8, no. 9, pp. 921–934, 2013.
- [16] T. Hinoue, D. J. Weisenberger, C. P. E. Lange et al., "Genome-scale analysis of aberrant DNA methylation in colorectal cancer," *Genome Research*, vol. 22, no. 2, pp. 271–282, 2012.
- [17] G. Tripodi, A. Piscione, G. Borsani et al., "Molecular cloning of an adducin-like protein: evidence of a polymorphism in the normotensive and hypertensive rats of the Milan strain," *Biochemical and Biophysical Research Communications*, vol. 177, no. 3, pp. 939–947, 1991.
- [18] A. F. Muro, M. L. Marro, S. Gajović, F. Porro, L. Luzzatto, and F. E. Baralle, "Mild spherocytic hereditary elliptocytosis and altered levels of α - and γ -adducins in β -adducin-deficient mice," *Blood*, vol. 95, no. 12, pp. 3978–3985, 2000.
- [19] D. M. Gilligan, L. Lozovatsky, B. Gwynn, C. Brugnara, N. Mohandas, and L. L. Peters, "Targeted disruption of the β adducin gene (Add2) causes red blood cell spherocytosis in mice," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 19, pp. 10717–10722, 1999.
- [20] J. W. Moon, S. K. Lee, J. O. Lee et al., "Identification of novel hypermethylated genes and demethylating effect of vincristine in colorectal cancer," *Journal of Experimental and Clinical Cancer Research*, vol. 33, no. 1, article 4, 2014.
- [21] I. Cebola, J. Custodio, M. Muñoz et al., "Epigenetics override pro-inflammatory PTGS transcriptomic signature towards selective hyperactivation of PGE2 in colorectal cancer," *Clinical Epigenetics*, vol. 7, no. 1, article 74, 2015.
- [22] S. Dedeurwaerder, M. Defrance, E. Calonne, H. Denis, C. Sotiriou, and F. Fuks, "Evaluation of the Infinium Methylation 450K technology," *Epigenomics*, vol. 3, no. 6, pp. 771–784, 2011.
- [23] C. Clark, P. Palta, C. J. Joyce et al., "A comparison of the whole genome approach of MeDIP-Seq to the targeted approach of the infinium humanmethylation450 BeadChip⁵ for methylome profiling," *PLoS ONE*, vol. 7, no. 11, Article ID e50233, 2012.



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

