DNA hypermethylated status and gene expression of *PAX1/SOX1* in patients with colorectal carcinoma

Jin Huang[1,2]
Zhi-Rong Tan[1,2]
Jing Yu[1,2]
He Li[1,2]
Qiao-Li Lv[1,2]
Ying-Ying Shao[3]
Hong-Hao Zhou[1,2]

[1]Department of Clinical Pharmacology, Xiangya Hospital, Central South University, Changsha, Hunan, [2]Institute of Clinical Pharmacology, Central South University, Hunan Key Laboratory of Pharmacogenetics, Changsha, Hunan, [3]Institute of Life Science, Chongqing Medical University, Yuzhong District, Chongqing, China

Background: Colorectal cancer (CRC) is a widespread and aggressive carcinoma with poor prognosis. Hypermethylation of specific gene promoters is an important mechanism of CRC. In this study, we investigated the hypermethylation of paired boxed gene 1 (*PAX1*) and sex-determining region Y-related high-mobility group box 1 (*SOX1*) genes in CRC tissues.

Methods: DNA methylation at cg2,09,07,471 *PAX1* and cg0,66,75,478 *SOX1* from 166 cancer tissues and 37 normal tissues from CRC patients were compared using datasets downloaded from The Cancer Genome Atlas. Quantitative methylation-specific polymerase chain reaction and assay of *PAX1* and *SOX1* were performed in dissected tumor and paracancerous tissues by surgery from 41 CRC patients. Quantitative reverse transcription polymerase chain reaction and immunohistochemistry assay were performed in both CRC and paired normal tissues to detect mRNA and protein expression, respectively.

Results: Methylation levels of *PAX1*/*SOX1* genes were significantly higher in cancer tissues than in paired normal tissues. *PAX1* and *SOX1* genes were methylated in 28 (68.3%) of the 41 CRC samples but in 5 (12.2%) and 0 (0%) of the paired normal control samples (both $P<0.001$), respectively. Sensitivities and specificities of *PAX1* methylation for the detection of cancer were 68.3% and 87.8%, respectively, whereas the corresponding values for *SOX1* were 68.3% and 100%. However, the Kaplan–Meier analysis illustrated no significant difference in the overall survivals between patients with high and low methylation levels of *SOX1* or *PAX1* ($P>0.5$). In addition, the methylation level of *PAX1*/*SOX1* was significantly higher in CRC patients with high TNM stage (TNM stage III/IV, 3.11±2.43) than those with low TNM stage (TNM stage I/II, 1.26±2.94, $P<0.05$). Relative RNA and protein expression levels of *PAX1*/*SOX1* were both significantly lower in CRC tissues than in their paired normal tissue.

Conclusions: This study is the first analysis of the methylation of *PAX1*/*SOX1*, which may be new biomarkers for CRC screening.

Keywords: colorectal cancer, *PAX1*, *SOX1*, molecular biomarker, DNA methylation, epigenetic

Introduction

Colorectal cancer (CRC) is a widespread carcinoma and has more than 600,000 associated mortalities annually worldwide.[1,2] CRC is one of the most common cancers of the gastrointestinal tract. After lung and breast cancers, CRC ranks third among the most common cancers worldwide.[3,4] Colon cancer mortality ranks fourth among cancer-related mortalities. CRC incidence and mortality increase with age. In China, the incidence of CRC has been rising because of improved living standards and the shift to high-protein, high-calorie, and low-fiber diet. According to the cancer statistics in 2016, the incidence of CRC in China was 33.1 million, which ranked fourth in all malignant

Correspondence: Hong-Hao Zhou
Department of Clinical Pharmacology, Xiangya Hospital, Central South University, No 110 Xiangya Road, Changsha, Hunan 410078, China
Tel +86 731 8480 5380
Fax +86 731 8225 4476
Email hhzhou2003@163.com

tumors. The mortality of CRC in China was 15.9 million, which ranked fifth in cancer deaths.[5] Colon cancer mortality increased by 15.22% in China, among which female colon cancer mortality ranked third among malignant carcinomas, whereas that of males ranked fourth.[6] Paired boxed gene 1 (*PAX1*) and sex-determining region Y-related high-mobility group box 1 (*SOX1*) are tumor suppressor genes. Previous studies demonstrated that *PAX1* was silenced because of hypermethylated status in many carcinomas, especially in cervical, ovarian, and oral cancers.[7–10] *PAX1* methylation is a potential biomarker to screen cervical cancer, as well as oral squamous cell carcinoma, with high sensitivity and specificity.[8,11,12] *SOX1* is a tumor suppressor that was suppressed by the hypermethylation of its promoter region in various carcinomas, including cervical, hepatocellular, and nasopharyngeal cancers.[10,12,13] However, the methylation of *PAX1* or *SOX1* in CRC has not been analyzed yet.

CRC and other carcinomas are often caused by an accumulation of genetic and epigenetic alterations, among which DNA methylation is the most frequent epigenetic alteration. DNA methylation usually occurs at carbon 5 of the cytosine ring within the CpG dinucleotides, and this process is a major epigenetic modification in the human genome.[14–16] DNA hypermethylation is an important mechanism to inactivate tumor suppressor genes that could be involved in DNA repair, metabolism of cancer, cell cycle control, apoptosis, and angiogenesis.[15,17,18] However, a few hypermethylated genes have been associated with CRC, and a limited number of studies have focused on prognostic methylation biomarker in CRC. To the best of our knowledge, *APC*, *RASSF1A*, *FAS*, *ZNF582*, *GFRA1*, and *GSTM2* promoter hypermethylation has been reported to be involved in the development with CRC.[19–22] However, whether the promoter methylation status of the *PAX1* and *SOX1* genes is associated with CRC remains to be elucidated.

In this study, the methylation status of *PAX1* and *SOX1* in CRC tissues and paired paracarcinoma tissues was investigated using quantitative methylation-specific polymerase chain reaction (qMSP). The results were verified using datasets from The Cancer Genome Atlas (TCGA) database.

Materials and methods
Patients and samples
Samples of CRC tissues and adjacent normal tissues were collected from 41 patients with CRC. Written informed consents were signed by all patients before collection of specimens. This study was approved by the Institutional Review Board of Department of Clinical Pharmacology, Xiangya

Hospital, Central South University (registration number: CTXY-1,50,001–2) and the Chinese Clinical Trial Registry (registration number: ChiCTR-DCD-1,50,06,289). Tumor tissues were defined as the sections next to the tumor diagnosed by histopathology. Paracancerous tissues were taken from surgically dissected tissues ~2 cm away from the defined tumor area without tumor invasion. The dissected tumor and paracancerous tissues were further evaluated histologically.

Genomic DNA was extracted from the tissues with the iSat Nucleic Acid Extraction kit (iStat Biomedical Co., Ltd., New Taipei City, Taiwan) according to the manufacturer's manual. The concentration of DNA was determined using BioSpec-Nano spectrophotometer (Shimadzu Company, Kyoto, Japan). Samples with a DNA yield of more than 500 ng were considered for further testing. Total RNA was extracted using TRIZOL reagent (Invitrogen, San Diego, CA, USA) according to the manufacturer's protocol and dissolved in 30 µL of diethylpyrocarbonate-treated water.

Determination of DNA methylation
Sodium bisulfite conversion of 500 ng of genomic DNA was performed using the iStat Bisulfite Conversion Kit (iStat Biomedical Co., Ltd) according to the manufacturer's manual. The bisulfited DNA was stored at −80°C until used. The DNA methylation of *PAX1* and *SOX1* genes was analyzed using *PAX1* and *SOX1* DNA Detection Kits (iSat Biomedical Co., Ltd), which use simplified technology based on TaqMan technologies for qMSP. The qMSP reaction was conducted in a total volume of 20 µL containing 2 µL of bisulfite-converted template gDNA (50 ng), 1 µL of 2× custom detection mix, and 10 µL of 2× custom universal polymerase chain reaction (PCR) master mix. The reaction was performed using the LightCycler 480 Real-Time PCR System (Roche Applied Science, Penzberg, Germany) under the following conditions. The reaction systems were first subjected to an initial incubation at 95°C for 10 min, followed by 50 cycles at 95°C for 10 s, and annealing and extension at 60°C for 40 s. DNA methylation levels were assessed as the methylation index (meth-index) was calculated as follows: $10,000 \times 2^{(Cp \text{ value of gene-}Cp \text{ value of } COL2A)}$. Testing results with *Cp* values of *COL2A* >36 were defined as detection failures.

Quantitative reverse transcription-PCR
Single-stranded cDNA was prepared using a PrimeScript RT Reagent Kit (Takara, Otsu, Japan) according to the manufacturer's protocol. Quantitative reverse transcription-PCR (qRT-PCR) was performed using LightCycler 480 Real-Time PCR System (Roche Applied Science) by SYBR

Premix DimerErase™ kit (Takara) to detect the expression of *SOX1* and *PAX1*, with *PPIA* and *B2M* as normalizing controls. The sequences of the primers used were as follows: *SOX1*: forward, 5′-CAGTACAGCCCCATCTCCAAC-3′ and reverse, 5′-GCGGGCAAGTACATGCTGA-3′; *PAX1*: forward, 5′-TCGCTATGGAGCAGACG TATG-3′ and reverse, 5′-GCTGCCGACTGATGTCACA-3′; *PPIA*: forward, 5′-CC CACCGTGTTCTTCGACATT-3′ and reverse, 5′-GGACCCGTATGCTTTAGGATGA-3′; and *B2M*: forward, 5′-GAGGCTATCCAGCGTACTCCA-3′ and reverse, 5′-CG GCAGGCATACTCATCTTTT-3′. The qRT-PCR reaction systems were performed with incubation at 95°C for 10 min, followed by 45 cycles at 95°C for 10 s, and annealing and extension at 60°C for 60 s. Relative quantification results were determined by $2^{-\Delta\Delta Ct}$ relative to the normalizing control genes.

Validation using TCGA data

The dataset of DNA methylation was obtained from the TCGA database, which included clinical information from 166 CRC tissues and 37 normal tissues. Level 3 datasets were obtained through Infinium HumanMethylation 27 BeadChip and analyzed using R software (http://www.R-project.org/).

Statistical analysis

Paired-sampled *t*-test was performed to compare the methylation levels between the CRC and normal tissue samples. Correlations between the frequency of DNA methylation occurrence and clinicopathological characteristics of CRC patients were analyzed by χ^2 test. Associations between DNA methylation status and clinicopathological characteristics of CRC patients were analyzed by Mann–Whitney *U*-test. Wilcoxon signed-rank test was performed to compare the methylation levels between CRC and paired normal tissues. *P*-values <0.05 were considered statistically significant. Kaplan–Meier method was performed to evaluate the overall survival rates. Log-rank test was used to calculate the difference in survival. The association between methylation levels and relative mRNA expression was analyzed by calculating Pearson's correlation coefficient and *P*-values. All statistical analyses were performed using SPSS Statistics 19.0 software (IBM Corporation, Armonk, NY, USA).

Results
Analysis of DNA methylation of *PAX1/SOX1* in CRC and paracancer tissues

DNA methylation status of *PAX1/SOX1* in 166 CRC patients was analyzed using TCGA datasets. Figure 1A shows that

the methylation levels of *PAX1* and *SOX1* genes were significantly greater in the 166 CRC samples than in the 37 normal samples (both *P*<0.05). In addition, the methylation levels of the two genes were higher in the majority of the CRC samples than in their paired normal samples (Figure 1B). The areas under the curve (AUCs) of the receiver-operating characteristic (ROC) plots for the methylation of *SOX1* and *PAX1* in 166 CRC and 37 normal tissues were 0.515 and 0.805, respectively (Figure 1C). *SOX1* methylation testing achieved 78.3% sensitivity and 83.8% specificity, and *PAX1* methylation testing achieved 31.9% sensitivity and 100% specificity.

The results from the TCGA datasets were verified by detecting the methylation levels of *PAX1* and *SOX1* in the tissues of patients with CRC. A total of 41 patients with CRC were recruited from Xiangya Hospital in Changsha, China, between May 2015 and June 2016. Patient ages ranged from 24 to 84 years (median age, 57 years). The clinicopathological characteristics of these patients are presented in Tables 1 and 4. The CRC tissues showed a significantly greater level of hypermethylated *SOX1/PAX1* compared with their paired normal tissues as detected by qMSP (both *P*<0.05, Figure 2A and B). The methylation status of samples and frequencies were determined based on the best meth-index cut-off values calculated using the ROC curves (Figure 2C and D). Log (meth-index) ≥3.6 for *SOX1* or log (meth-index) ≥3.7 for *PAX1* were considered to be methylation positivity. *PAX1* and *SOX1* genes were both methylated in 28 (68.3%) of the 41 CRCs samples but in 5 (12.2%) and 0 (0%) of paired normal control samples (both *P*<0.001), respectively. The AUCs of the ROC plots in Figure 2C and D for the methylation of *PAX1* and *SOX1* to detect CRC from normal tissues were 0.743 (95% CI 0.627–0.858) and 0.798 (95% CI 0.689–0.907), respectively (Table 2). At meth-index of 1.56, *PAX1* methylation testing achieved 68.3% sensitivity and 87.8% specificity. At meth-index of 3.01, *SOX1* methylation testing achieved 68.3% sensitivity and 100% specificity.

Furthermore, the correlation between the methylation status of *SOX1* and *PAX1* in the same CRC samples is analyzed in Figure S1. Pearson's correlation analysis showed that there is no linear relationship between the methylation status of *PAX1* and *SOX1* in the same CRC samples.

PAX1/SOX1 methylation status and survival in CRC patients using TCGA datasets

We used the TCGA dataset to investigate whether the methylation status of *SOX1* or *PAX1* influenced prognosis.
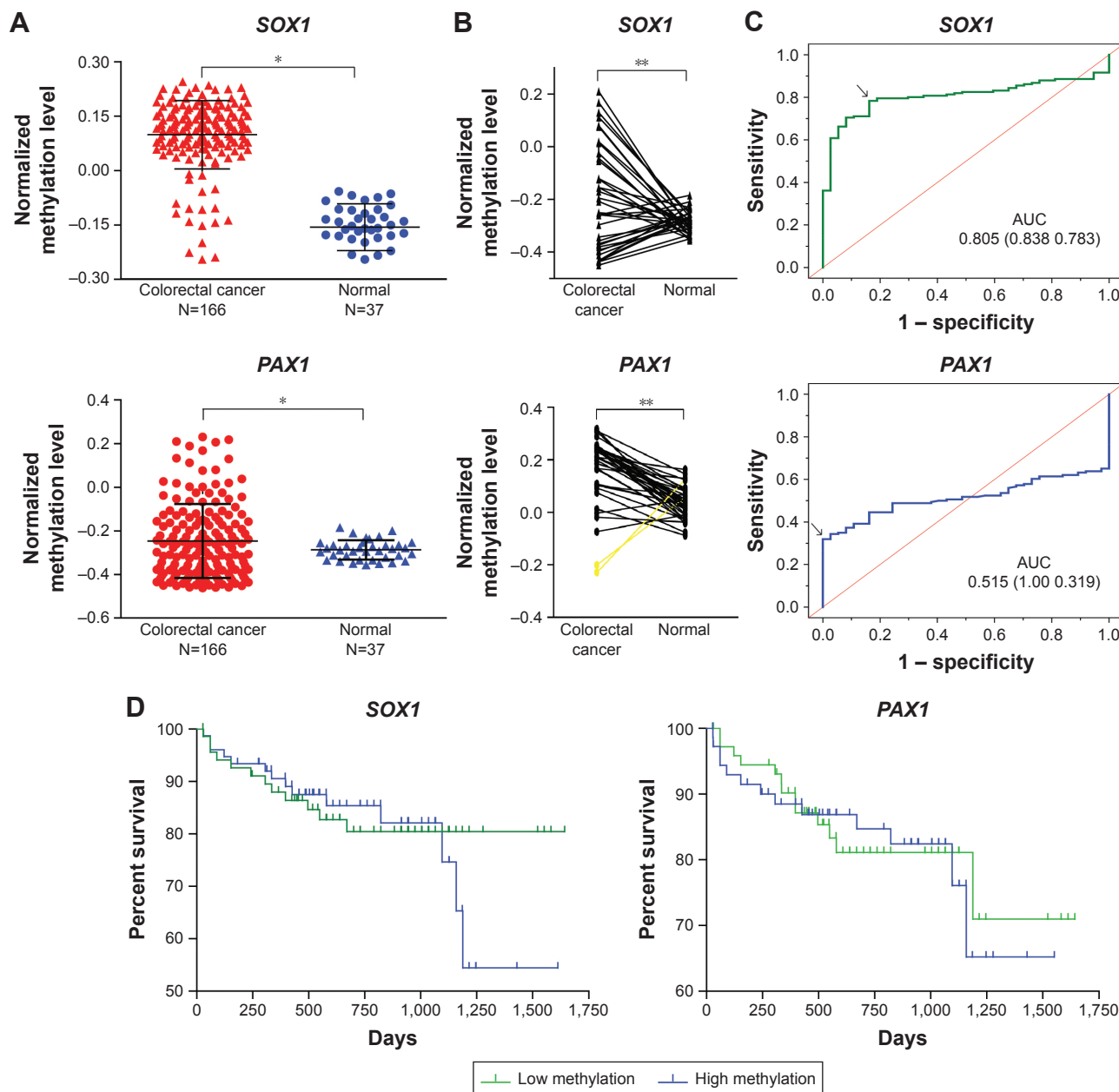
Figure 1 DNA methylation of *PAX1* and *SOX1* analyzed in CRC using TCGA dataset.
Notes: (**A**) Comparison of DNA methylation status of *SOX1* and *PAX1* in 166 CRC tissues and 37 normal tissues. $P<0.05$ was considered significant. (**B**) DNA methylation status of *SOX1* and *PAX1* in 37 CRC tissues and paired normal tissues. (**C** and **D**) Kaplan–Meier estimation of overall survival according to *PAX1* and *SOX1* methylation status in 166 patients. Median methylation status of each gene was defined as cutoff. *$P<0.05$ was calculated by independent sample *t*-test. **$P<0.05$ was calculated by Wilcoxon signed-rank test. AUC data shown as value (specificity sensitivity).
Abbreviations: CRC, colorectal cancer; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1; TCGA, The Cancer Genome Atlas.

The results from the Kaplan–Meier analysis illustrated no significant difference in the overall survival between patients with high and low methylation levels of *SOX1* or *PAX1* ($P>0.5$, Figure 1D). The mean survival time ± standard error (SE) of patients with CRC and low methylation level of *SOX1* was 619.0±42.2 days, while that of patients with CRC and high methylation level of *SOX1* was 602.2±49.8 days. The mean survival time ± SE of patients with CRC and low methylation level of *PAX1* was 603.4±44.9 days, while that

of patients with CRC and high methylation level of *PAX1* was 617.8±47.5 days.

Correlation between methylation status of *PAX1*/*SOX1* and clinicopathological characteristics of CRC patients

The association of patients' demographic and clinicopathological characteristics with *PAX1*/*SOX1* methylation status was analyzed using the TCGA dataset and clinical samples.

Table 1 Association of the frequency of methylation occurrence of *PAX1* and *SOX1* in cancer tissues to the clinical characteristics of 41 patients with CRC

| Characteristics	N	SOX1				PAX1			
		M	U	Methylation percentage	P-value*	M	U	Methylation percentage	P-value*
Age, years					0.460				0.460
Mean (SD) [range]	56	(12.65)	(27–84)						
<60	25	16	9	64.00		16	9	64.00	
>60	16	12	4	75.00		12	4	75.00	
Gender					0.658				0.025
Male	20	13	7	65.00		17	3	85.00	
Female	21	15	6	71.43		11	10	52.38	
Tumor location					0.987				0.987
Colon	19	13	6	68.42		13	6	68.42	
Rectum	22	15	7	68.18		15	7	68.18	
LN metastasis					0.203				0.072
N0	29	19	10	65.52		18	11	62.07	
N1	11	9	2	81.82		10	1	90.91	
N2	1	0	1	0.00		0	1	0.00	
Tumor invasion					0.548				0.661
T1	3	3	0	100.00		3	0	100.00	
T2	12	8	4	66.67		8	4	66.67	
T3	5	4	1	80.00		3	2	60.00	
T4	21	13	8	61.90		14	7	66.67	
TNM stage					0.666				0.646
I	9	7	2	77.78		7	2	77.78	
II	18	9	9	50.00		9	9	50.00	
III	12	10	2	83.33		10	2	83.33	
IV	2	2	0	100.00		2	0	100.00	
Tumor HG					0.432				0.432
Poor	7	6	1	85.71		6	1	85.71	
Moderate	20	12	8	60.00		12	8	60.00	
Well	14	10	4	71.43		10	4	71.43	
Tumor size					0.063				0.128
<3 cm	3	0	3	0.00		1	2	33.33	
3–5 cm	9	6	3	66.67		4	5	44.44	
>5 cm	5	4	1	80.00		4	1	80.00	
NA	24	18	6	75.00		19	5	79.17	

Note: *Calculated by χ^2 test.
Abbreviations: CRC, colorectal cancer; HG, histological grade; LN, lymph node; M, methylated; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1; U, unmethylated.

The relationship between methylation and clinicopathological characteristics in CRC is summarized in Table S1 for TCGA datasets and Table 3 for clinical samples. No significant association was found between the *PAX1* or *SOX1* methylation frequency and the clinicopathological characteristics of patients with CRC in TCGA datasets and clinical samples.

The correlation between methylation status of *PAX1/SOX1* in cancer tissues and clinical characteristics of CRC patients is summarized in Table 4. The methylated status of *SOX1* and *PAX1* was significantly associated with TNM stage. The methylation distributions in the CRC tissues

from the 41 patients with CRC at various TNM stages are illustrated in Figure 3. The methylation level log (meth-index) of *PAX1* in patients with CRC and high TNM stage (TNM stage III/IV) was 3.11±2.43, which was significantly higher than that of patients with low TNM stage (TNM stage I/II, 1.26±2.94, $P<0.05$). The AUCs under ROC plots for the methylation of *SOX1* and *PAX1* to detect high TNM stage of CRC were 0.759 (95% CI 0.515–0.869) and 0.692 (95% CI 0.588–0.930), respectively. At log (meth-index) of 3.84, *PAX1* methylation testing achieved 71.4% sensitivity and 70.4% specificity. At log (meth-index) of 3.82, *PAX1* methylation testing achieved 63.0% sensitivity and 69.2% specificity.
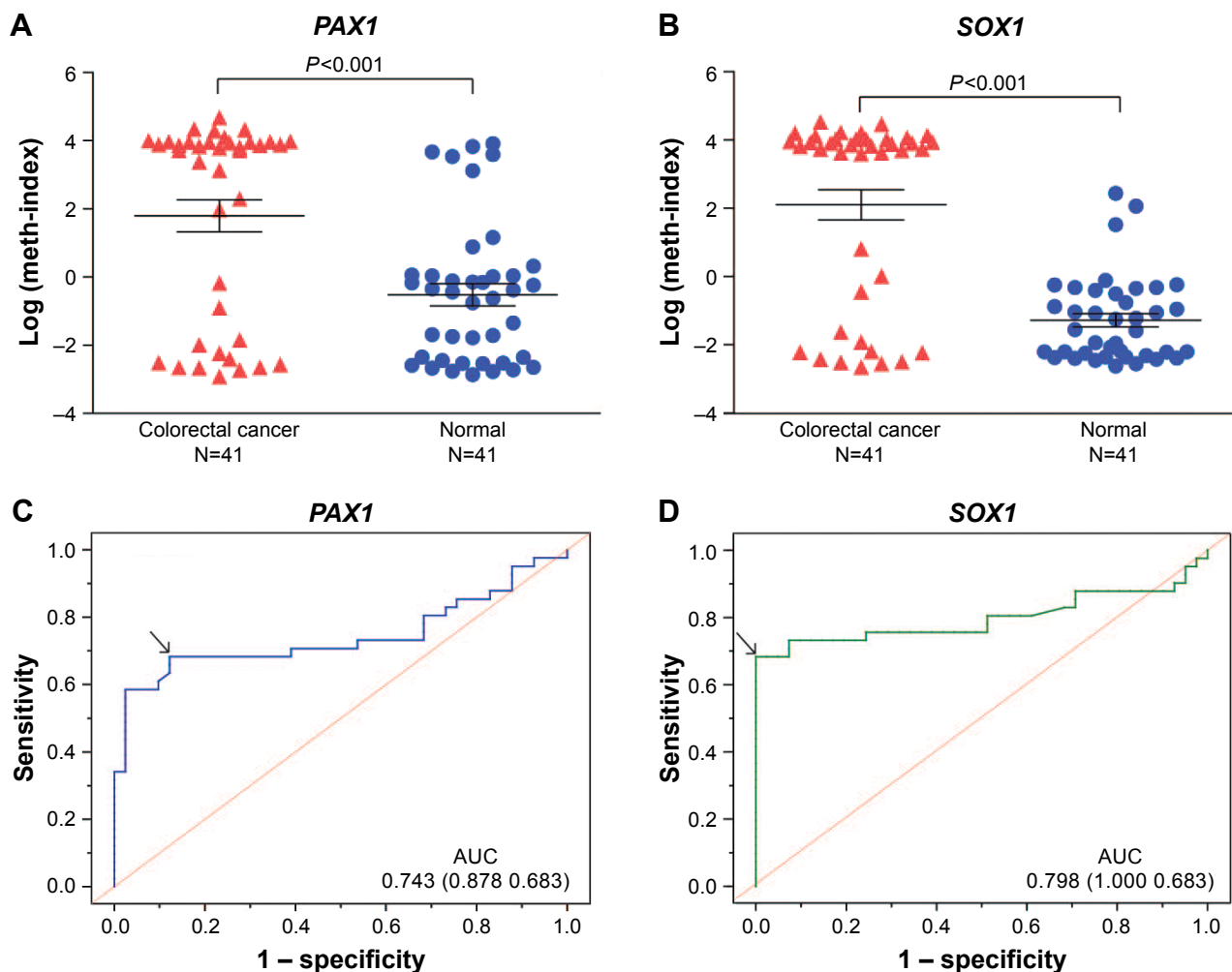
Figure 2 DNA methylation status of *PAX1* and *SOX1* genes in CRC tissues and paired normal tissues from 41 patients with CRC.
Notes: Comparison of DNA methylation status log (meth-index) of *PAX1* (**A**) and *SOX1* (**B**) genes in 41 CRC tissues and paired normal tissues. The AUC of the ROC curve was calculated to diagnose the CRC and normal tissues for *PAX1* (**C**) and *SOX1* (**D**). *P*-value was calculated by Wilcoxon signed-rank test.
Abbreviations: AUC, areas under the curve; CRC, colorectal cancer; *PAX1*, paired boxed gene 1; ROC, receiver-operating characteristic; *SOX1*, sex-determining region Y-related high-mobility group box 1.

Correlations between mRNA expression and DNA methylation of *PAX1/SOX1* in CRC samples

The mRNA expression level of *PAX1/SOX1* in the cancer and paired normal tissues of the 41 patients was quantified using quantitative RT-PCR to determine whether methylation changes in these two genes affected gene expression.

The results indicated significant difference in the mRNA levels of *PAX1/SOX1* between CRC and paired normal tissues (*P*<0.05, Figure 4A and B). The mRNA levels of both genes were lower in the CRC samples than in their paired normal samples.

Pearson's correlation analysis was also performed to analyze the relationship between mRNA expression levels

Table 2 Sensitivities, specificities, and AUC of *PAX1* and *SOX1* methylation in detecting CRC tumor (n=41)

Detection modality	Cancer		Paracancer		Sensitivity %	Specificity %	AUC	95% CI	*P*-value*
	U	M	U	M					
PAX1	13	28	36	5	68.3	87.8	0.743	0.627–0.858	<0.001
SOX1	13	28	41	0	68.3	100.0	0.798	0.689–0.907	<0.001

Note: *Calculated by χ² test.
Abbreviations: AUC, areas under the curve; CRC, colorectal cancer; M, methylated; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1; U, unmethylated.

Table 3 Sensitivities, specificities, and AUC of *PAX1* and *SOX1* methylation in detecting low (n=27) and high (n=14) TNM stages

Detection modality	Sensitivity %	Specificity %	AUC	95% CI	P-value*
PAX1	0.714	0.704	0.759	0.588–0.930	0.007
SOX1	0.714	0.630	0.692	0.515–0.869	0.046

Note: *Calculated by χ^2 test.
Abbreviations: AUC, areas under the curve; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1.

Immunohistochemical expression of *PAX1/SOX1*

Immunohistochemical examination of the 41 pairs of cancer and normal tissues of patients with CRC was performed to evaluate whether methylation changes in the two genes affected protein expression. High SOX1 expression was detected in 9 (22.0%) of 41 CRC tissues and in 28 (68.2%) of 41 paracancer tissues. High SOX1 expression was detected

Table 4 Correlation between methylation levels of *PAX1* and *SOX1* in the cancer tissues and clinical characteristics of 41 patients with CRC

Characteristics	*PAX1* methylation level		*SOX1* methylation level	
	Mean ± SD	P-value*	Mean ± SD	P-value*
Age, years		0.621		0.779
<60	1.59±3.10		1.94±2.83	
>60	2.37±2.53		2.46±2.69	
Gender		0.050		0.531
Male	2.92±2.07		2.00±2.79	
Female	0.92±3.24		2.28±2.78	
Tumor location		0.296		0.143
Colon	2.17±2.77		2.33±2.66	
Rectum	1.66±3.02		1.98±2.89	
TNM stage		**0.007**		**0.046**
I/II	1.26±2.94		1.77±2.88	
III/IV	3.11±2.43		2.87±2.43	
Tumor size		0.399		0.792
<5 cm	0.19±3.37		1.09±3.14	
>5 cm	2.60±2.90		2.57±2.84	
Tumor HG		0.273		0.080
W/M	1.52±2.99		1.86±2.82	
Poor	3.39±1.93		3.21±2.41	
Tumor invasion		0.665		0.807
T1/2	1.92±2.85		2.41±2.70	
T3/4	1.89±2.96		1.99±2.83	
LN metastasis		0.071		0.192
N0	1.46±2.93		1.94±2.85	
N1/2/3	2.93±2.60		2.65±2.56	

Notes: *Association analysis for methylation status by Mann–Whitney *U*-test. Significant values <0.05 were shown in bold font.
Abbreviations: CRC, colorectal carcinoma; LN, lymph node; HG, histological grade; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1; W/M, well/moderate.

in 9 (22.0%) of 41 CRC tissues and in 22 (53.6%) of 41 para-cancer tissues. The results indicated significantly lower PAX1 and SOX1 protein expression in CRC compared with their paired normal tissues ($P>0.05$). Typical immunohistochemical (IHC) staining images for SOX1 and PAX1 in CRC tissues and paracancer tissues are shown in Figure 5.

Discussion

DNA methylation is the most frequent epigenetic alteration. Hypermethylation of specific gene promoters is an important mechanism of carcinogenesis. High frequency of promoter methylation of *PAX1* and *SOX1* genes has been detected in cervical and oral cancers. In the present study, we demonstrated that the methylation levels of *PAX1* and *SOX1* are higher in colorectal tumor tissues compared with the paired normal tissues from the patients with CRC. DNA methylation at cg2,09,07,471 (*PAX1*) and cg0,66,75,478 (*SOX1*) from TCGA datasets supports our results.

PAX1 is a member of the PAX family, which plays a key role during fetal development. The PAX gene family is classified into four groups based on the gene structure.[7] *PAX1* is categorized into group 1, which is the most frequently methylated gene in the *PAX* family. *PAX1* may be involved in carcinogenesis and tumor progression in invasive or aggressive carcinomas. *PAX1*, as a tumor suppressed gene, was reported to inhibit cervical, oral, and ovarian carcinomas silenced by promoter hypermethylation. A study on head and neck squamous cell carcinoma illustrated that *PAX1* participated in the Notch pathway.[23] *PAX1* activation was blocked by the silence of *HOX* family of transcription factors, which are downstream targets of Notch signaling.
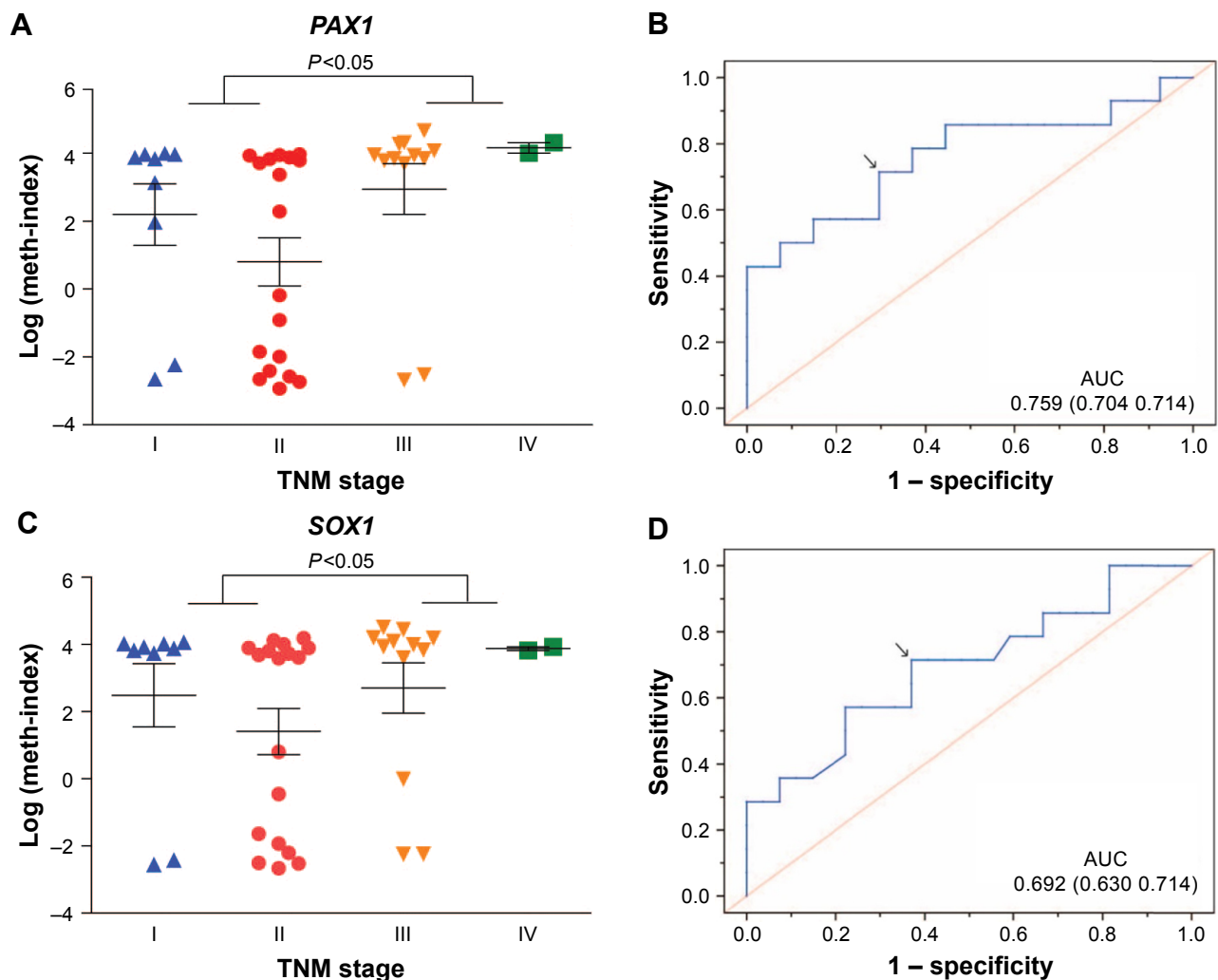
SOX1 is a member of the *SOX* transcription factor family, which plays crucial roles in embryonic and postnatal development and stem cell regulation.[24] Several members of the *SOX* family have been implicated in cancer development. *SOX10* is a novel oncogene in hepatocellular carcinoma, which could activate the Wnt/β-catenin signaling pathway.[25] *SOX9* is a transcription factor expressed in human fetal prostate epithelium and enhances prostate cancer invasion.[26] *SOX1* has been reported as a tumor suppressor to promoters in various carcinoma tissues depending on tumor status and type. *SOX1* was reported to inhibit cervical, liver, and

**Figure 3** Methylation distribution of CRC tissues at various TNM stages.
Notes: (**A**) Dot plots illustrate the log (meth-index) distributions of *PAX1* gene at various TNM stages. (**B**) The AUC of ROC curve was calculated to diagnose the TNM stage I/II and stage III/IV for *PAX1*. (**C**) Dot plots illustrate the log (meth-index) distributions of *SOX1* gene at various TNM stages. (**D**) The AUC of ROC curve was calculated to diagnose the TNM stage I/II and stage III/IV for *SOX1* gene. *P*<0.05 was determined by Mann–Whitney *U*-test. Arrows show the best specificity and sensitivity.
Abbreviations: AUC, areas under the curve; CRC, colorectal cancer; *PAX1*, paired boxed gene 1; ROC, receiver-operating characteristic; *SOX1*, sex-determining region Y-related high-mobility group box 1.
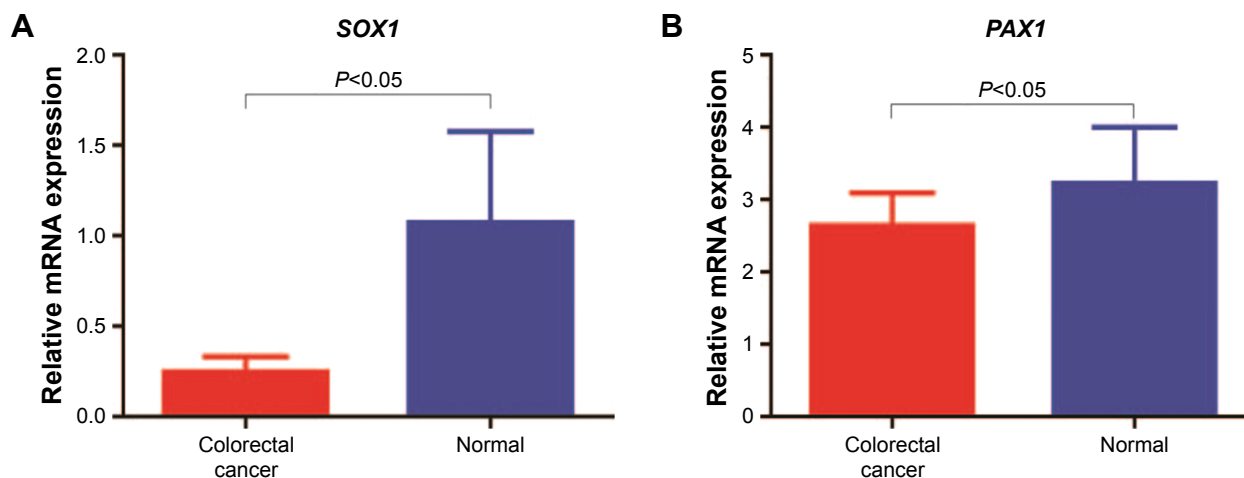

Figure 4 (*Continued*)
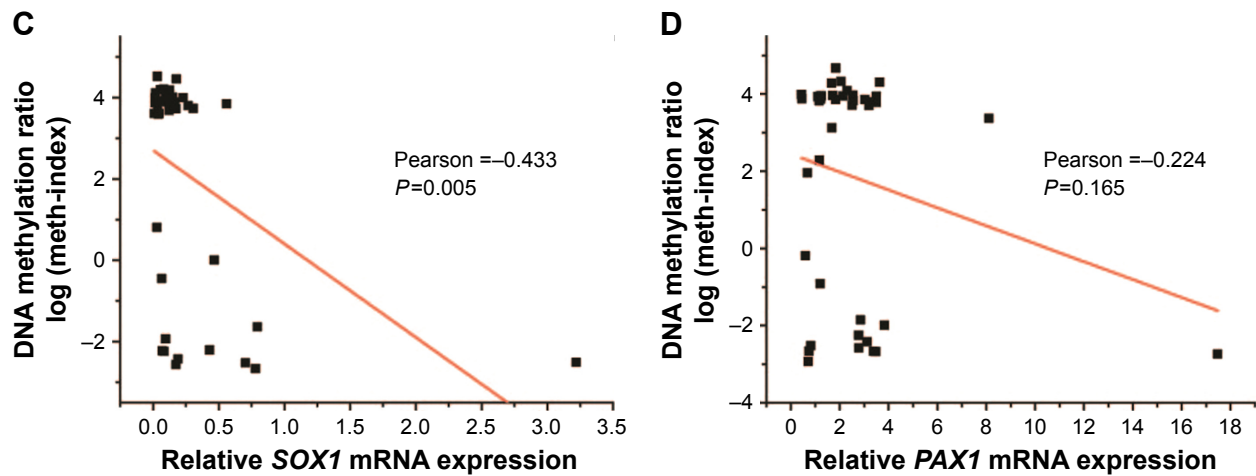
Figure 4 mRNA expression of *SOX1/PAX1* genes in CRC tissues of 41 patients with CRC.
Notes: (**A**) Relative *SOX1* and (**B**) *PAX1* mRNA expression levels ± SEM in CRC tissues and paired normal tissues. Pearson's correlation was used to measure the linear relationship between mRNA expression (*x*-axis) and methylation level log (meth-index) (*y*-axis) of (**C**) *SOX1* and (**D**) *PAX1* genes in 41 CRC tissues. *P*-values were performed by *t*-test.
Abbreviations: CRC, colorectal cancer; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1.

nasopharyngeal carcinomas via the Wnt/β-catenin signaling pathway.[13] Moreover, *SOX1* gene could regulate prostate cancer stem cell invasion by interfering with STAT3 signaling pathway.[27,28]

The association of *PAX1/SOX1* methylation to the clinicopathological characteristics of patients with CRC was analyzed to further evaluate the clinical significance of *PAX1* and *SOX1* methylation in CRC development and progression.
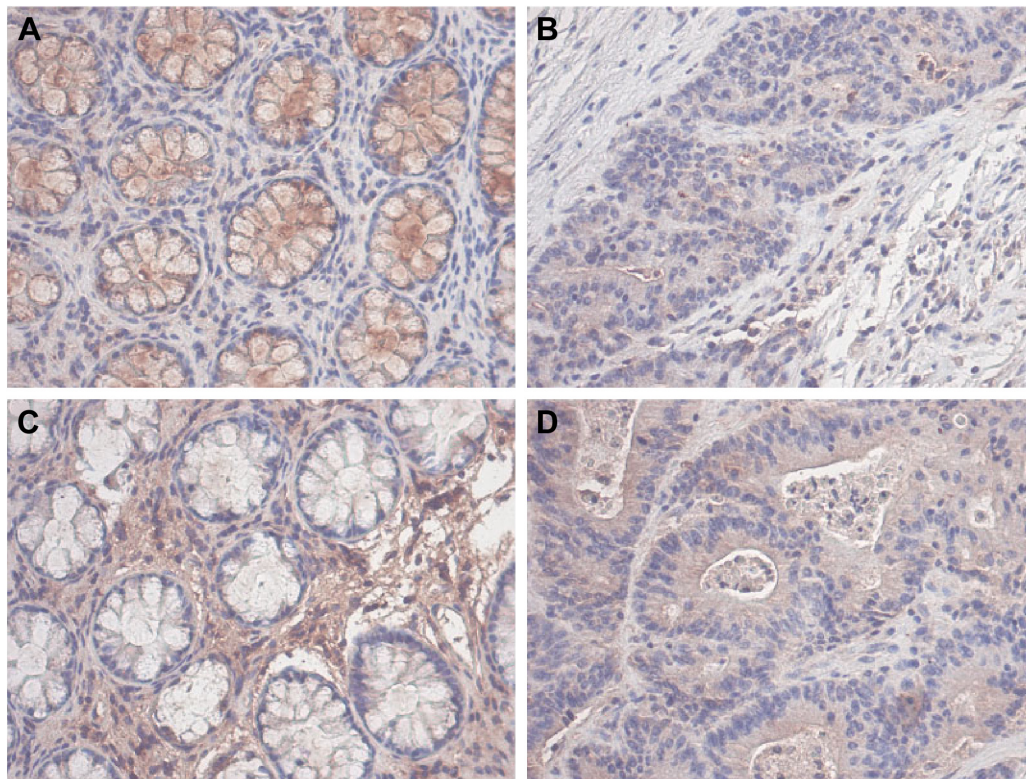
Figure 5 Representative expression of *PAX1* and *SOX1* protein in colorectal tissues compared with paired normal tissues.
Notes: (**A**) IHC staining of *SOX1* protein expression in normal tissue samples. (**B**) IHC staining of *SOX1* protein expression in CRC tissue samples. (**C**) IHC staining of *PAX1* protein expression in normal tissue samples. (**D**) IHC staining of *PAX1* protein expression in CRC tissue samples (×200).
Abbreviations: CRC, colorectal cancer; IHC, immunohistochemical; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1.

The results showed significant association between *PAX1/SOX1* and TNM stage of these patients. Besides, the sensitivity and specificity of *SOX1* and *PAX1* may play important roles in the poor prognosis and tumor development in CRCs. These results were consistent with a previous study on other carcinomas. A study on hepatocellular cancer has also indicated that *SOX1* is correlated with TNM stage of hepatocellular cancer because low *SOX1* expression in cancer tissues was significantly associated with advanced TNM stage.[29] The methylation of *PAX1* was significantly associated with the cervical intraepithelial neoplasia (CIN) grade, and the methylation level of *PAX1* increased as the grade of CIN increased.[7,30]

Hypermethylation was believed to play key roles in gene expression. The methylation status of *SOX1* and *PAX1* was observed to be significantly higher in cancer tissues compared with paired paracancer tissues, whereas the mRNA expression of these genes was lower in the cancer tissues. However, no evident linear relationship was found with methylation status and mRNA expression in the same sample of CRC. This phenomenon may be due to the fact that the methylation of specific CpG site in a promoter could not affect the expression of the whole gene. Nevertheless, mRNA expression of *SOX1* and *PAX1* genes was inhibited by promoter hypermethylation. This observation is in agreement with previous hepatocellular cancer studies.[29] In these studies, the mRNA expression of *SOX1* gene was significantly lower in hepatocellular cancer tissues than in the paired normal tissues based on the semiquantitative RT-PCR detection in 15 paired cancer tissues and paracancer tissues in hepatocellular cancer patients. In addition, mRNA expression of *SOX1* was downregulated in nasopharyngeal cancer tissues compared with their paired paracancer tissues because of the hypermethylation of *SOX1* in nasopharyngeal cancer tissues.[13]

We further detected the expression of *PAX1* and *SOX1* proteins in 41 paired CRC and paracancer tissues by IHC. The expression level of *PAX1/SOX1* protein was significantly lower in CRC tissues than in their paired normal tissues. It can be inferred that the hypermethylation may induce the low protein expression in CRC tissues. These results were in agreement with previous studies on hepatocellular cancer.[29] Previous results showed that the expression level of *SOX1* protein was significantly lower in the tissues of hepatocellular cancer than in their paired normal tissues as shown by IHC and Western blot assays.

Conclusions

For the first time, the hypermethylation levels of *PAX1/SOX1* genes in CRC are reported in this study. The methylation level of *PAX1/SOX1* was significantly higher in patients with CRC at high TNM stage than those at low TNM stage. Relative mRNA and protein expressions of *PAX1/SOX1* were downregulated by the aberrant methylation in CRC compared with their paired normal tissue. Thus, aberrant DNA methylation of *PAX1/SOX1* could be a potential molecular tumor biomarker for CRC screening.

Author contributions

HHZ designed the study. JH carried out the molecular experiments and data analysis and wrote the manuscript. ZRT and YYS contributed in data collection. QLL, JY, and HL provided technical assistance. All authors read and approved the final manuscript. All authors contributed toward data analysis, drafting and revising the paper and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Zhang ZM, Guan X, Li YJ, Zhu MC, Yang XJ, Zou X. HLA Class I Expressions on Peripheral Blood Mononuclear Cells in Colorectal Cancer Patients. *Chin J Cancer Res*. 2012;24(1):77–82.
2. Marley AR, Nan H. Epidemiology of colorectal cancer. *Int J Mol Epidemiol Genet*. 2016;7(3):105–114.
3. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. *CA Cancer J Clin*. 2008;58(2):71–96.
4. Hwang C, Lee S, Kim A, Kim YG, Ahn SJ, Park DY. Venous invasion in colorectal cancer: impact of morphologic findings on detection rate. *Cancer Res Treat*. 2016;48(4):1222–1228.
5. Jemal A, Siegel R, Xu J, Ward E. Cancer Statistics, 2010. *CA Cancer J Clin*. 2012;62(1):10–29.
6. Song D, Huang R, Ma T, Chen Y, Li J, Wang X. [Analysis of screening for colorectal cancer high-risk groups in Harbin, China 2012–2015]. *Zhonghua wei chang wai ke za zhi*. 2016;19(10):1139–1143. Chinese.
7. Chang CC, Huang RL, Wang HC, Liao YP, Yu MH, Lai HC. High methylation rate of LMX1A, NKX6–1, PAX1, PTPRR, SOX1, and ZNF582 genes in cervical adenocarcinoma. *Int J Gynecol Cancer*. 2014;24(2):201–209.
8. Huang YK, Peng BY, Wu CY, Su CT, Wang HC, Lai HC. DNA methylation of PAX1 as a biomarker for oral squamous cell carcinoma. *Clin Oral Investig*. 2014;18(3):801–808.

9. Chao TK, Ke FY, Liao YP, Wang HC, Yu CP, Lai HC. Triage of cervical cytological diagnoses of atypical squamous cells by DNA methylation of paired boxed gene 1 (PAX1). *Diagn Cytopathol*. 2013;41(1):41–46.

10. Su HY, Lai HC, Lin YW, Chou YC, Liu CY, Yu MH. An epigenetic marker panel for screening and prognostic prediction of ovarian cancer. *Int J Cancer*. 2009;124(2):387–393.

11. Kan YY, Liou YL, Wang HJ, et al. PAX1 methylation as a potential biomarker for cervical cancer screening. *Int J Gynecol Cancer*. 2014;24(5):928–934.

12. Lai HC, Ou YC, Chen TC, et al. PAX1/SOX1 DNA methylation and cervical neoplasia detection: a Taiwanese Gynecologic Oncology Group (TGOG) study. *Cancer Med*. 2014;3(4):1062–1074.

13. Guan Z, Zhang J, Wang J, et al. SOX1 down-regulates beta-catenin and reverses malignant phenotype in nasopharyngeal carcinoma. *Mol Cancer*. 2014;13:257.

14. Chela-Flores J, Migoni RL. CG methylation in DNA transcription. *Int J Theor Phys*. 1990;29(8):853–862.

15. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med*. 2003;349(21):2042–2054.

16. Teodoridis JM, Strathdee G, Brown R. Epigenetic silencing mediated by CpG island methylation: potential as a therapeutic target and as a biomarker. *Drug Resist Updat*. 2004;7(4–5):267–278.

17. Kulis M, Esteller M. DNA methylation and cancer. *Adv Genet*. 2010;70:27–56.

18. Dayeh T, Tuomi T, Almgren P, et al. DNA methylation of loci within ABCG1 and PHOSPHO1 in blood DNA is associated with future type 2 diabetes risk. *Epigenetics*. 2016;11(7):482–488.

19. Matthaios D, Balgkouranidou I, Karayiannakis A, et al. Methylation status of the APC and RASSF1A promoter in cell-free circulating DNA and its prognostic role in patients with colorectal cancer. *Oncol Lett*. 2016;12(1):748–756.

20. Manoochehri M, Borhani N, Karbasi A, Koochaki A, Kazemi B. Promoter hypermethylation and downregulation of the FAS gene may be involved in colorectal carcinogenesis. *Oncol Lett*. 2016;12(1):285–290.

21. Harada T, Yamamoto E, Yamano HO, et al. Analysis of DNA methylation in bowel lavage fluid for detection of colorectal cancer. *Cancer Prev Res (Phila)*. 2014;7(10):1002–1010.

22. Wei J, Li G, Zhang J, et al. Integrated analysis of genome-wide DNA methylation and gene expression profiles identifies potential novel biomarkers of rectal cancer. *Oncotarget*. 2016;7(38):62547–62558.

23. Guerrero-Preston R, Michailidi C, Marchionni L, et al. Key tumor suppressor genes inactivated by "greater promoter" methylation and somatic mutations in head and neck cancer. *Epigenetics*. 2014;9(7):1031–1046.

24. Sarkar A, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell*. 2013;12(1):15–30.

25. Zhou D, Bai F, Zhang X, et al. SOX10 is a novel oncogene in hepatocellular carcinoma through Wnt/beta-catenin/TCF4 cascade. *Tumour Biol*. 2014;35(10):9935–9940.

26. Wang H, Leav I, Ibaragi S, et al. SOX9 is expressed in human fetal prostate epithelium and enhances prostate cancer invasion. *Cancer Res*. 2008;68(6):1625–1630.

27. Tsao CM, Yan MD, Shih YL, et al. SOX1 functions as a tumor suppressor by antagonizing the WNT/beta-catenin signaling pathway in hepatocellular carcinoma. *Hepatology*. 2012;56(6):2277–2287.

28. Mathews LA, Hurt EM, Zhang X, Farrar WL. Epigenetic regulation of CpG promoter methylation in invasive prostate cancer cells. *Mol Cancer*. 2010;9:267.

29. Lou J, Zhang K, Chen J, Gao Y, Wang R, Chen LB. Prognostic significance of SOX-1 expression in human hepatocelluar cancer. *Int J Clin Exp Pathol*. 2015;8(5):5411–5418.

30. Lai HC, Lin YW, Huang RL, et al. Quantitative DNA methylation analysis detects cervical intraepithelial neoplasms type 3 and worse. *Cancer*. 2010;116(18):4266–4274.

Supplementary materials
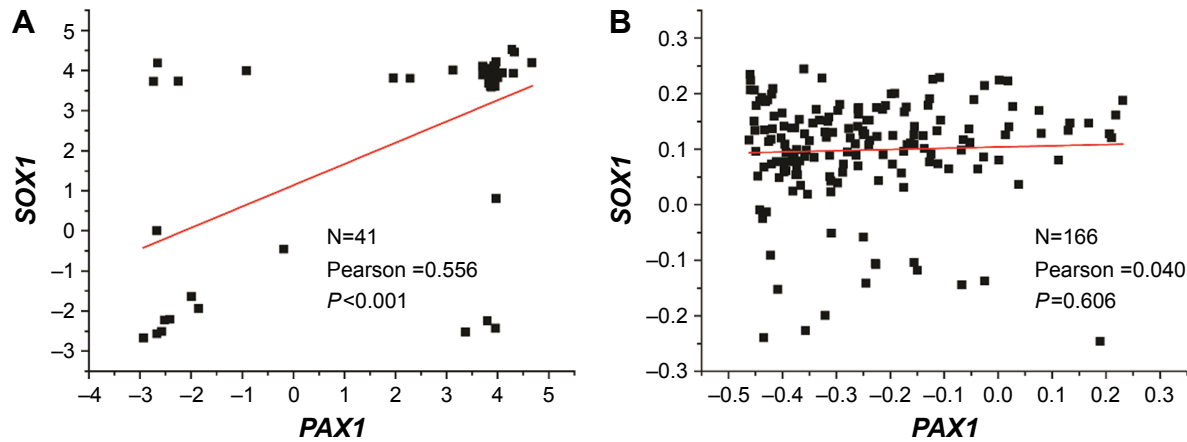
Figure S1 Correlation between the methylation status of *SOX1* and *PAX1* in the same CRC samples.
Notes: Pearson's correlation was used to measure the linear relationship between methylation log (meth-index) of *PAX1* (*x*-axis) and *SOX1* (*y*-axis) in 41 CRC tissues (**A**) and 166 CRCs downloaded from TCGA dataset (**B**). *P*-values were performed by independent sample *t*-test.
Abbreviations: CRC, colorectal cancer; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1; TCGA, The Cancer Genome Atlas.

Table S1 Association of the frequency of methylation occurrence of cg20907471 (*PAX1*) and average of cg22303211 and cg06675478 (*SOX1*) in cancer tissues to the clinical characteristics of 166 patients with CRC using TCGA datasets (n=166)

Characteristics	N	SOX1				PAX1			
		M	U	Methylated percentage	P-value	M	U	Methylated percentage	P-value*
Age, years					0.239				0.560
Mean (SD) [range]	70	(11.9)	(36–90)						
<50	15	7	8	46.67		3	12	20.00	
50–70	61	42	19	68.85		21	40	34.43	
>70	90	54	36	60.00		29	61	32.22	
Gender					0.631				0.617
Male	83	53	30	63.86		28	55	33.73	
Female	83	50	33	60.24		25	58	30.12	
M					0.576				0.191
M0	142	88	54	61.97		43	99	30.28	
M1	24	15	9	62.50		19	14	79.17	
N					0.329				0.209
N0	99	65	34	65.66		27	72	27.27	
N1	32	20	12	62.50		14	18	43.75	
N2	34	18	17	52.94		12	23	35.29	
T					0.310				0.834
T1	4	1	3	25.00		2	2	50.00	
T2	35	24	11	68.57		10	25	28.57	
T3	110	69	41	62.73		36	74	32.73	
T4	17	9	8	52.94		5	12	29.41	
Stage					0.666				0.646
I	32	22	10	68.75		9	23	28.13	
II	64	41	23	64.06		18	46	28.13	
III	45	25	20	55.56		16	29	35.56	
IV	24	15	10	62.50		10	15	41.67	
CEA					0.338				0.887
<9.7	78	46	32	58.97		26	52	33.33	
>9.7	24	13	11	54.17		8	16	33.33	
NA	64	44	20	68.75		19	45	29.69	

Note: *Calculated χ^2 test.
Abbreviations: CEA, carcinoembryonic antigen; CRC, colorectal cancer; M, methylated; *PAX1*, paired boxed gene 1; *SOX1*, sex-determining region Y-related high-mobility group box 1; TCGA, The Cancer Genome Atlas; U, unmethylated.