CELL CYCLE
2023, VOL. 22, NOS. 21–22, 2424–2435
https://doi.org/10.1080/15384101.2023.2296210

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

SMAD7 gene polymorphisms and their influence on patients with colorectal cancer

Yongsheng Wu[a,b*], Jue Xu[c*], Biaobin Tan[b], Ting Yi[b], Su Liu[b], Guang Yang[b], Kai Li[b], and Xinhan Zhao[a]

[a]Department of Oncology, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China; [b]The Second Ward of Oncology and Hematology Department, The People's Hospital of XiangXiang, Xiangxiang, China; [c]Department of Intrarenal Rheumatology and Immunology, The People's Hospital of XiangXiang, Xiangxiang, China

ABSTRACT

Colorectal cancer (CRC) is a prevalent malignant tumor, and its pathogenesis is still not fully understood. Studies have shown that *SMAD7* gene polymorphisms can affect CRC susceptibility, but the results have been inconsistent and require additional confirmation. Our study aimed to evaluate the effect of *SMAD7* variants on the risk of CRC in the Chinese Han population. A total of five single nucleotide polymorphisms (SNPs) in *SMAD7* were genotyped among 696 CRC patients and 696 healthy participants using the MassARRAY iPLEX platform. SNPs were evaluated for their associations with CRC using logistic regression analysis under multiple genetic models. The false-positive report probability (FPRP) analysis was used to validate the positive findings. Our study indicated that rs11874392 showed an increased association with CRC risk (odds ratio, 1.31; 95% confidence interval, 1.04–1.67; $p = 0.024$). Stratified analysis showed that rs11874392 might increase the risk of CRC in females (OR = 1.70, $p = 0.028$), individuals with smoking (OR = 1.87, $p = 0.026$), and drinking (OR = 1.38, $p = 0.027$). The rs11874392 was found to be related to an elevated risk of rectal cancer (OR = 1.73, $p = 0.003$), but not with colon cancer. FPRP analysis demonstrated that all of these associations were statistically significant (FPRP <0.2). Additionally, rs11874392 was the strongest predictive model for CRC. This study provides evidence that the *SMAD7* rs11874392 is related to an increased susceptibility to CRC.

ARTICLE HISTORY
Received 17 October 2023
Revised 12 December 2023
Accepted 13 December 2023

KEYWORDS
Colorectal cancer; *SMAD7*; polymorphisms; association study

Introduction

Colorectal cancer (CRC) is a malignant tumor that originates from the rectum or colon.

CRC is the third most prevalent cancer and ranks as the second leading cause of cancer-related mortality worldwide [1,2]. The estimated number of new CRC cases is projected to exceed 3.2 million in 2040 [3]. Metastasis is considered the primary cause of death in patients with CRC, and approximately 20% of newly diagnosed CRC patients have metastasis [4,5]. In addition, studies have demonstrated that the 5-year overall survival rate for patients with metastatic CRC is only 4–12% [6]. Immune cell PD-L1 expression is significantly higher in MMR-deficient (MSI-H) CRC as compared to MMR-proficient (MSI-L) tumors, with no differences among the different MSI-H molecular subtypes. The recommended screening for defective, DNA mismatch repair includes immunohistochemistry (IHC) and/or MSI test. However, there are challenges in distilling the biological and technical heterogeneity of MSI testing down to usable data. It has been reported by Adeleke et al that IHC testing of the mismatch repair machinery may give different results for a given germline mutation has been suggested that this may be due to somatic mutations [7]. Therefore, the early diagnosis and prevention for CRC are of great significance. The pathological mechanism of CRC is a complex process involving the interaction of multiple genetic and environmental factors. Recent studies have shown that various genetic changes affecting TGF-β signaling pathway, Wnt/β-catenin signaling pathway, DNA repair defects, inflammatory reactions, and angiogenesis have been proven to regulate the carcinogenesis of CRC [8–12].

CONTACT Xinhan Zhao zhaoxinhan@mail.xjtu.edu.cn Department of Oncology, The First Affiliated Hospital of Xi'an Jiaotong University, 277 West Yanta Road, Xi'an 710061, China
*Yongsheng Wu and Jue Xu contributed equally to this work.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/15384101.2023.2296210.

Additionally, Genome-wide association studies (GWAS) have identified genetic loci related to CRC risk, some of these genetic genes participate in the TGF-β signaling pathway [13]. The *SMAD7* gene contains the most prevalent genetic loci associated with CRC risk in this pathway [14].

The transforming growth factor β (TGF-β) signaling pathway regulates several biological functions. It is involved in cell proliferation, differentiation, apoptosis, migration, as well as the immune response, and the occurrence and progression of cancer, including CRC [8,15,16]. *SMAD7* is an important negative regulator in the TGF-β signaling pathway. The imbalance of this pathway is closely related to the pathogenesis of CRC [17]. Among them, changes in the function of *SMAD7*, including gene polymorphism and gene expression, may results in excessive inhibition of the TGF-β signaling pathway and facilitate the invasion and proliferation of cancer cells [18–21]. In recent years, several studies have demonstrated a correlation between *SMAD7* gene polymorphism and CRC. For example, Yin et al. [22] have identified a risk locus at 18q21.1 (rs6507854) in the European population through a post-GWAS analysis. An association study involving 109 CRC cases and 109 healthy individuals in the Iranian population reveals no correlation between *SMAD7* rs6507854 and the risk of CRC under allele, dominant, and additive genetic models [23]. Jiang et al. [24] indicates that *SMAD7* rs11874392 (OR = 0.8) and rs12953717 (OR = 1.29) are significantly correlated with CRC susceptibility in the North Carolina, whereas no significant association is found with rs4939827. *SMAD7* rs11874392 is associated with the susceptibility to CRC (OR = 1.16) in the California and Hawaii [25]. There is no significant association between rs12953717 and colon cancer in a population-based association study (561 cases and 721 controls) conducted in Kentucky [26]. *SMAD7* rs4939827 has been reported to be related to an enhanced susceptibility to CRC in the Hong Kong Chinese population [27]. However, in Utah, rs4939827 has been shown to decrease the risk of colon cancer [27]. A study has shown that rs7226855 is associated with CRC risk in Singapore Chinese [28]. To summarize, the findings of previous studies for the impact of *SMAD7* SNPs on CRC risk have yielded inconclusive. Therefore, further study is needed to clarify the possible roles of the *SMAD7* SNPs in CRC.

In this case-control study, we aimed to evaluate the association between specific polymorphisms (rs6507874, rs11874392, rs4939827, rs12953717, and rs7226855) of *SMAD7* and the risk of CRC susceptibility in the Chinese Han population. Our study will help to elucidate the mechanism of *SMAD7* gene polymorphism in the occurrence and progression of CRC, and provides new targets and strategies for the early diagnosis, prevention, and individualized treatment for CRC. The flow chart of this study is shown in Figure 1.

Materials and methods

Study population

In the present study, a total of 1392 unrelated Chinese Han individuals, including 696 CRC patients and 696 healthy participants, were randomly selected from The People's Hospital of XiangXiang. All patients were initially diagnosed with CRC and their diagnoses were confirmed by two experienced oncologists through colonoscopy and pathological examination of a malignant tumor in the colon or rectum. The exclusion criteria for the patients are as follows: 1) age ≤18 years old; 2) patients with genetic diseases; 3) patients with a family history of gastrointestinal defects; 4) patients with a history of any cancers, including gastrointestinal tumors. Controls were the healthy individuals who underwent routine medical checkup at the same hospital during the same time period. Exclusion criteria for the controls included with a history of any tumors, a family history of gastrointestinal defects, and colonoscopy reports indicating malignancy, inflammatory ulcers, or polyps were excluded from the study. Participant characteristics, such as age, gender, body mass index (BMI), smoking status, alcohol intake, and tumor-node-metastasis (TNM) stage, were collected from both a standardized questionnaire and medical records. Our study was approved by the ethics committee of The People's Hospital of XiangXiang. Each participant was informed about the study's purpose, and their informed consent was obtained prior to the start of the study.

SNP selection and genotyping

In this study, we assessed the relationship of CRC susceptibility with five selected SNPs in the *SMAD7* gene. These SNPs, namely rs6507874, rs11874392, rs4939827, rs12953717, and rs7226855, were specifically selected from the Chinese Han population in the Beijing (CHB) dataset of the 1000 Genomes Project. The inclusion criteria for SNP are as follows: 1) minor allele frequency (MAF) > 0.05, Hardy-Weinberg equilibrium (HWE) > 0.05), min genotype > 0.75, and r^2 < 0.8. 2) The call rate of each SNP is greater than 0.95. 3) It has not been reported among the Han population in China. Peripheral blood samples were collected, and genomic DNA was extracted from these samples using a Whole Blood Genomic DNA Isolation kit (GoldMag, Xi'an, China) according to the manufacturer's instruction. The purity and concentration of genomic DNA were measured using a NanoDrop 2000C spectrophotometer (Thermo Scientific, Waltham, USA), and the samples were subsequently stored at −20°C for further analysis. Primers for PCR amplification were designed using the Agena Bioscience Assay Design Suite V2.0 software (https://agenacx.com/online-tools/). Genotyping for each SNP was performed using the Agena MassARRAY iPLEX platform (Agena Bioscience Inc., CA, USA). The PCR reaction was performed using 1 μL of genomic DNA (10 ng/μL) and a 4 μL PCR mixture consisting of 1.8 μL of water, 0.5 μL of 10× PCR buffer, 0.4 μL of 25 mM $MgCl_2$, 0.1 μL of 25 mM dNTP, 1 μL of PCR Primer mix, and 0.2 μL of 5 U/μL PCR Taq (Agena Bioscience Inc., CA, USA). The PCR conditions were as follows: initial denaturation at 95°C for 2 min, followed by 45 cycles of denaturation at 95°C for 30 s, annealing at 56°C for 30 s, and a final extension at 72°C for 60 s. The reaction was then cooled to 25°C and kept indefinitely. SNP alleles from different quality extension primers were identified using matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry following alkaline phosphatase reaction, single group extension, and resin desalination reaction. Subsequently, the genotyping data obtained from the Agena MassARRAY iPLEX platform was analyzed through Agena Bioscience TYPER software (version 4.0).

Bioinformatic information analysis

A HaploReg v4.2 online software (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) was used to predict the potential roles of the five SNPs. Additionally, the mRNA expression of *SMAD7* in CRC was analyzed using the UALCAN online software (https://ualcan.path.uab.edu/analysis.html).

Statistical analyses

The statistical analysis was carried out with SPSS software, version 20.0. A Student's *t*-test was performed to compare the ages between the two groups. The statistical differences in gender, smoking status, alcohol consumption, and BMI were assessed by Pearson's chi-square test. The allele frequencies in controls and HWE were determined by Fisher's exact test. Logistic regression analysis was performed to assess the relationship between *SMAD7* SNPs and CRC risk. The analyses were adjusted for age, gender, smoking, drinking, and BMI. Various genetic models, including allele, codominant, dominant, recessive, and log-additive models, were applied in this analysis. Additionally, the correlation between SNPs and the risk of CRC was confirmed using false-positive report probability (FPRP) analysis. Furthermore, the multifactor dimensionality reduction (MDR) method was carried out to assess the impact of SNP interactions on CRC susceptibility.

Results

Characteristics of the study participants

Table 1 lists the demographic data for the healthy controls (390 men and 306 women) and CRC patients (386 men and 310 women). The study included 269 cases of colon cancer (CC) and 376 cases of rectal cancer (RC). The mean age of the case was 59.06 ± 10.49 years, and 59.11 ± 8.96 years in the controls. There were no significant differences in age ($p = 0.919$), gender ($p = 0.829$), smoking status ($p = 0.577$), alcohol consumption ($p = 1.000$), and BMI ($p = 0.086$) between the two groups.

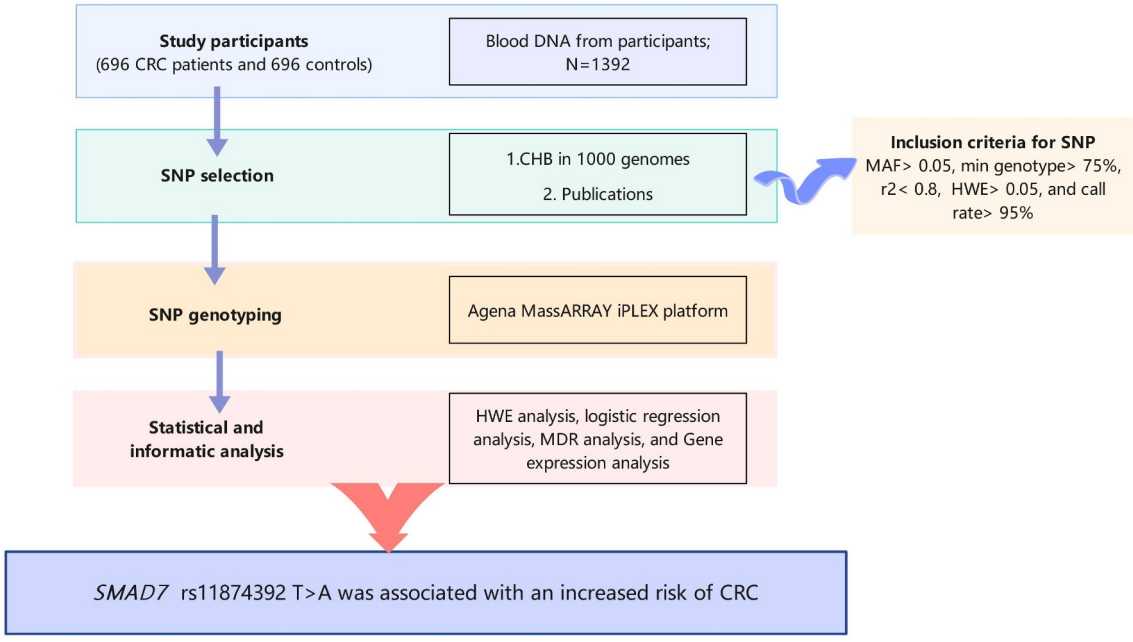

Figure 1. The flowchart of this study.

Table 1. Characteristics of the participants included in this study.

Variables	Case ($n = 696$)	Control ($n = 696$)	p
Gender			0.829[a]
Male	386 (55.5%)	390 (56.0%)	
Female	310 (44.5%)	306 (44.0%)	
Age	59.06 ± 10.49	59.11 ± 8.96	0.919[b]
>60 years	321 (46.1%)	318 (45.7%)	
≤60 years	375 (53.9%)	378 (54.3%)	
Smoking status			0.577[a]
Yes	257 (36.9%)	247 (35.5%)	
No	439 (63.1%)	449 (64.5%)	
Alcohol consumption			1.000[a]
Yes	203 (29.2%)	203 (29.2%)	
No	493 (70.8%)	493 (70.8%)	
BMI (kg/m^2)			0.086[a]
≥24	365 (52.4%)	333 (47.8%)	
<24	331 (47.6%)	363 (52.2%)	
Colon cancer	269 (38.6%)		
Rectal cancer	376 (54.0%)		
Grade			
I, II	46 (6.6%)		
III, IV	94 (13.5%)		
Lymph node metastasis			
Yes	72 (10.3%)		
No	140 (20.1%)		

[a]p value was calculated from two-sided Chi-squared test; [b]p value was calculated from Student's t test. $p < 0.05$ indicates statistical significance.

The basic characteristics of *SMAD7* gene polymorphisms

A total of five SNPs (including rs6507874, rs11874392, rs4939827, rs12953717, and rs7226855) were successfully genotyped in this case-control study. Table 2 shows the characteristics and potential functions of the selected SNPs. The MAF values of all SNP were greater than 0.05, and the allele frequency of each SNP in the controls conformed to the HWE (all $p > 0.05$). Additionally, the distribution of alleles (A/T) and genotypes (AA/AT/TT) of rs11874392 was remarkedly diverse between the two groups ($p < 0.05$), but no significant differences were found in other SNPs ($p > 0.05$), as shown in Figure 2.

Table 2. Basic information on *SMAD7* SNPs in this study.

SNP-ID	Chr.	Position	Allele (minor/major)	MAF Case	MAF Control	p^a-HWE	Haploreg4.2
rs6507874	18	48922435	T/C	0.251	0.246	0.222	Promoter histone marks; Enhancer histone marks; DNAse; Motifs changed
rs11874392	18	48926786	A/T	0.443	0.400	0.386	Promoter histone marks; Enhancer histone marks; DNAse; Proteins bound
rs4939827	18	48927093	T/C	0.253	0.246	0.309	Promoter histone marks; Enhancer histone marks; DNAse; Motifs changed; NHGRI/EBI GWAS hits
rs12953717	18	48927559	T/C	0.231	0.229	0.390	Promoter histone marks; Enhancer histone marks; DNAse; Motifs changed; GRASP QTL hits
rs7226855	18	48927678	A/G	0.247	0.245	0.304	Promoter histone marks; Enhancer histone marks; DNAse; Motifs changed; GRASP QTL hits

SNP, single-nucleotide polymorphism; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; OR, odds ratio; CI, confidence interval. $p < 0.05$ indicates statistical significance.

SMAD7 gene polymorphisms and the risk of CRC

The influence of *SMAD7* SNPs on CRC risk is shown in Figure 3. The result demonstrated that rs1187392 was observably associated with an enhanced susceptibility to CRC under allele (T > A, OR = 1.19, 95% CI = 1.03–1.39, $p = 0.022$), codominant (TA: OR = 1.31, 95% CI = 1.04–1.67, $p = 0.024$, AA: OR = 1.37, 95% CI = 1.01–1.87, $p = 0.044$), dominant (OR = 1.33, 95% CI = 1.06–1.66, $p = 0.013$), and log-additive (OR = 1.19, 95% CI = 1.03–1.39, $p = 0.022$) models.

SMAD7 gene polymorphisms related to the risk factors for CRC

We then investigated the association of SNPs with risk factors (such as gender, age, smoking, and drinking status). As displayed in Table 3, rs11874392 was linked to an promoted susceptibility to CRC in females under allele (T > A, OR = 1.36, 95% CI = 1.08–1.71, $p = 0.008$), codominant (TA: OR = 1.68, 95% CI = 1.16–2.42, $p = 0.006$, AA: OR = 1.70, 95% CI = 1.06–2.71, $p = 0.028$), dominant (OR = 1.68, 95% CI = 1.19–2.37, $p = 0.003$),

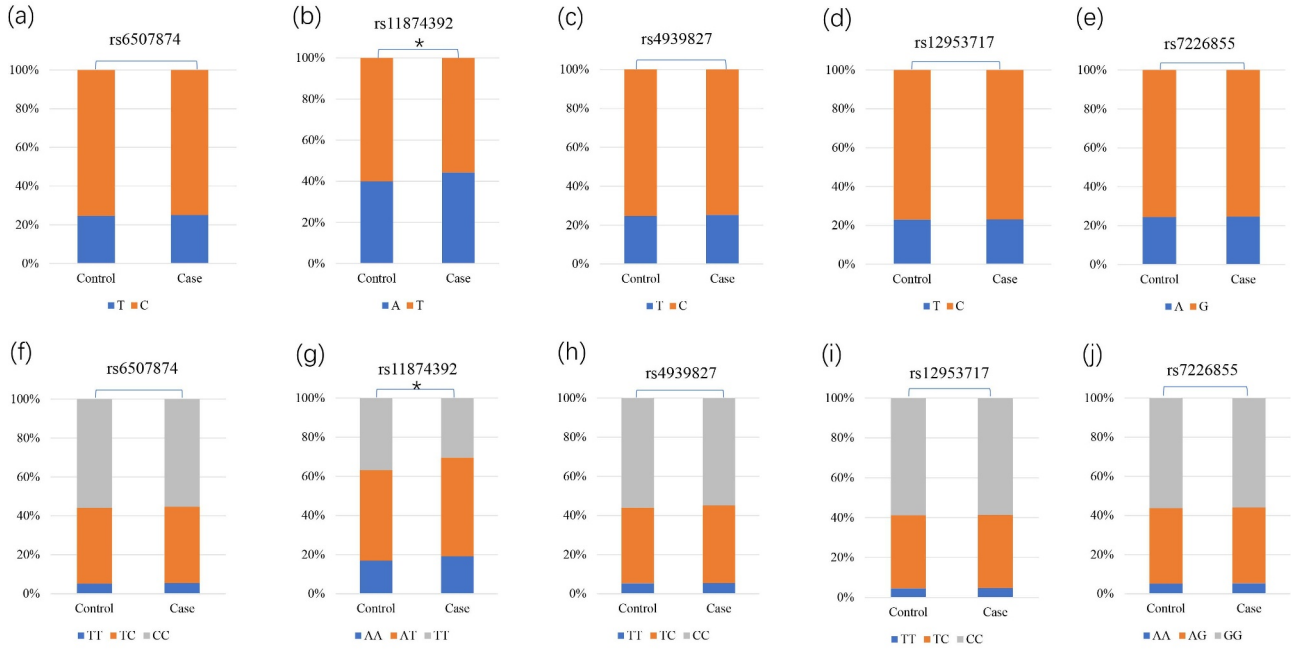

Figure 2. The frequencies of *SMAD7* SNPs genotype distributions in colorectal cancer patients and controls.

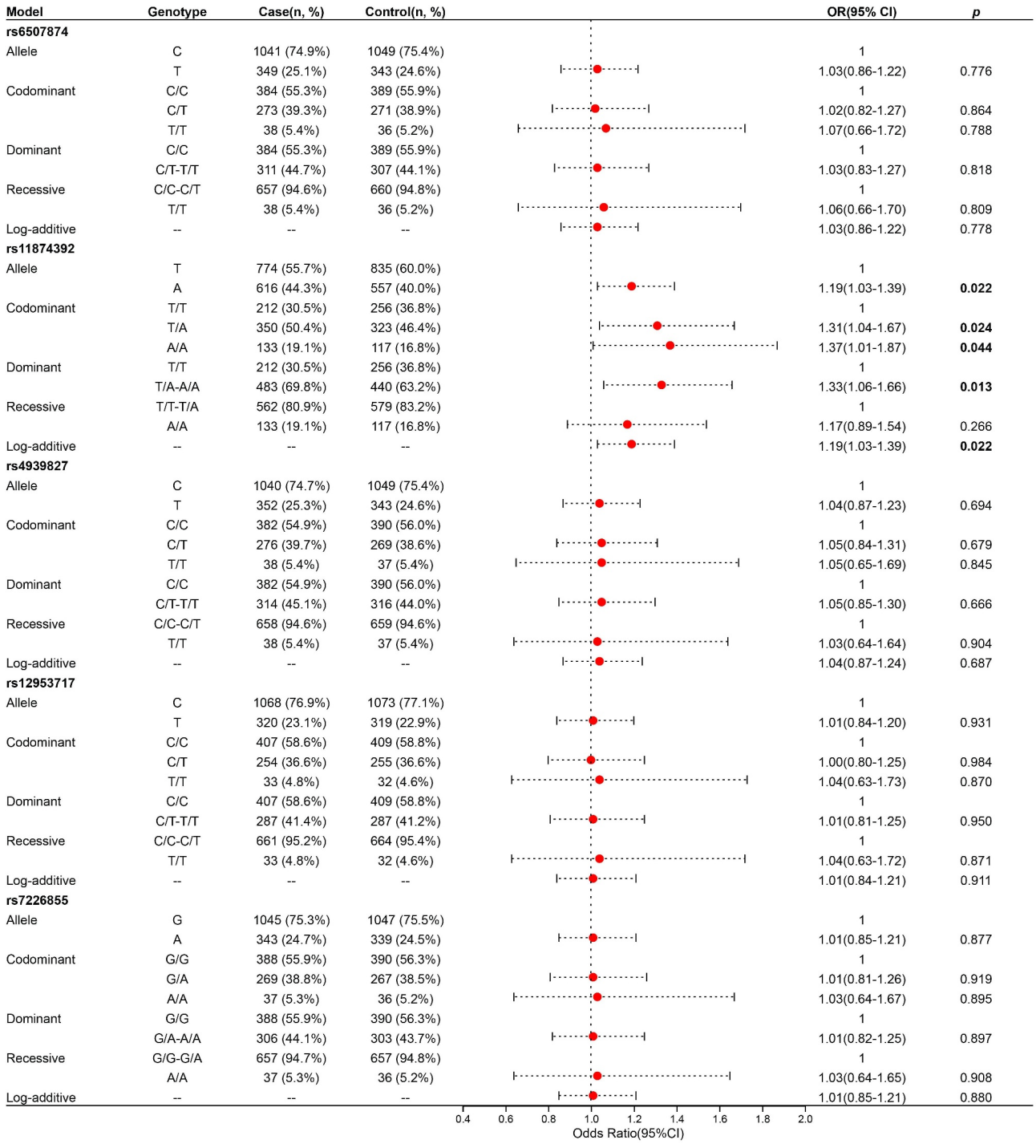

Figure 3. Association of *SMAD7* gene polymorphisms with colorectal cancer risk.

and log-additive (OR = 1.35, 95% CI = 1.08–1.70, p = 0.010) models. When stratified by smoking and drinking status (Table 4), it was discovered that rs11874392 indicated a higher risk of susceptibility to CRC in participants who smoked (allele: OR = 1.34, 95% CI = 1.04–1.72, p = 0.023, Codominant: TA, OR = 1.62, 95% CI = 1.07–2.47, p = 0.023, AA, OR = 1.87, 95% CI = 1.08–3.24, p = 0.026, dominant: OR = 1.69, 95% CI = 1.14–2.50, p = 0.010, and log-additive: OR = 1.41, 95% CI = 1.08–1.84, p = 0.013) and drank (T > A, OR = 1.38, 95% CI = 1.04–1.82, p = 0.027).

The influence of SNPs on CC and RC

Table 5 shows the results of our investigation into the relationship between SNPs and susceptibility

Table 3. The association between *SMAD7* rs11874392 and colorectal cancer susceptibility stratified by gender and age.

		Male		Female		> 60 years		≤ 60 years	
Model	Allele/Genotype	OR (95% CI)	p^a	OR (95% CI)	p^a	OR (95% CI)	p^b	OR (95% CI)	p^b
Allele	T	1		1		1		1	
	A	1.08 (0.88–1.32)	0.463	1.36 (1.08–1.71)	**0.008**	1.16 (0.92–1.44)	0.205	1.23 (0.99–1.51)	0.050
Codominant	T/T	1		1		1		1	
	T/A	1.08 (0.78–1.49)	0.643	1.68 (1.16–2.42)	**0.006**	1.56 (1.10–2.23)	**0.014**	1.12 (0.81–1.55)	0.497
	A/A	1.10 (0.72–1.69)	0.649	1.70 (1.06–2.71)	**0.028**	1.21 (0.76–2.23)	0.424	1.55 (1.02–2.34)	**0.038**
Dominant	T/T	1		1		1		1	
	T/A-A/A	1.09 (0.80–1.48)	0.600	1.68 (1.19–2.37)	**0.003**	1.47 (1.05–2.06)	**0.026**	1.23 (0.91–1.67)	0.181
Recessive	T/T-T/A	1		1		1		1	
	A/A	1.05 (0.72–1.53)	0.789	1.27 (0.83–1.93)	0.268	0.92 (0.61–1.40)	0.709	1.45 (1.00–2.10)	**0.048**
Log-additive	–	1.06 (0.86–1.30)	0.615	1.35 (1.08–1.70)	**0.010**	1.17 (0.93–1.47)	0.189	1.23 (1.00–1.50)	**0.048**

OR, odds ratio, 95% CI; 95% confidence intervals.
The p^a values were calculated by logistic regression adjusted by age, BMI, smoking, and drinking. The p^b values were calculated by logistic regression adjusted by gender, BMI, smoking, and drinking.
$p < 0.05$ indicates statistical significance.

Table 4. The association of rs11874392 and colorectal cancer susceptibility under smoking- and drinking-based stratification.

		Smoking		Non-smoking		Drinking		Non-drinking	
Model	Allele/Genotype	OR (95% CI)	p^a	OR (95% CI)	p^a	OR (95% CI)	p^b	OR (95% CI)	p^b
Allele	T	1		1		1		1	
	A	1.34 (1.04–1.72)	**0.023**	1.12 (0.93–1.35)	0.243	1.38 (1.04–1.82)	**0.027**	1.13 (0.94–1.35)	0.187
Codominant	T/T	1		1		1		1	
	T/A	1.62 (1.07–2.47)	**0.023**	1.16 (0.85–1.58)	0.349	1.41 (0.87–2.30)	0.164	1.25 (0.93–1.67)	0.135
	A/A	1.87 (1.08–3.24)	**0.026**	1.03 (0.69–1.53)	0.885	1.70 (0.86–3.36)	0.127	1.14 (0.79–1.65)	0.484
Dominant	T/T	1		1		1		1	
	T/A-A/A	1.69 (1.14–2.50)	**0.010**	1.12 (0.84–1.50)	0.440	1.48 (0.93–2.35)	0.099	1.22 (0.93–1.60)	0.160
Recessive	T/T-T/A	1		1		1		1	
	A/A	1.40 (0.86–2.27)	0.179	0.94 (0.66–1.34)	0.746	1.39 (0.75–2.58)	0.297	1.00 (0.72–1.39)	0.997
Log-additive	–	1.41 (1.08–1.84)	**0.013**	1.03 (0.85–1.26)	0.737	1.33 (0.96–1.84)	0.088	1.09 (0.91–1.31)	0.352

OR, odds ratio, 95% CI; 95% confidence intervals.
The p^a values were calculated by logistic regression adjusted by age, gender, BMI, and drinking. The p^b values were calculated by logistic regression adjusted by age, gender, BMI, and smoking.
$p < 0.05$ indicates statistical significance.

Table 5. rs11874392 associated with the risk of colon cancer and rectal cancer.

		Colon cancer		Rectal cancer	
Model	Allele/Genotype	OR (95% CI)	p	OR (95% CI)	p
Allele	T	1		1	
	A	1.02 (0.83–1.25)	0.850	1.33 (1.11–1.59)	**0.002**
Codominant	T/T	1		1	
	T/A	1.15 (0.84–1.58)	0.379	1.44 (1.08–1.93)	**0.014**
	A/A	0.99 (0.65–1.52)	0.964	1.73 (1.20–2.49)	**0.003**
Dominant	T/T	1		1	
	T/A-A/A	1.11 (0.82–1.49)	0.498	1.52 (1.15–2.00)	**0.003**
Recessive	T/T-T/A	1		1	
	A/A	0.91 (0.62–1.34)	0.646	1.39 (1.01–1.90)	**0.042**
Log-additive	–	1.02 (0.84–1.25)	0.825	1.33 (1.11–1.59)	**0.002**

OR, odds ratio, 95% CI; 95% confidence intervals.
The p values were calculated by logistic regression adjusted by age, gender, BMI, smoking, and drinking.
$p < 0.05$ indicates statistical significance.

to CC and RC. The findings indicated that rs11874392 may enhance the risk of RC under allele (OR = 1.33, 95% CI = 1.11–1.59, $p = 0.002$), codominant (TA, OR = 1.44, 95% CI = 1.08–1.93, $p = 0.014$, AA, OR = 1.73, 95% CI = 1.20–2.49, $p = 0.003$), dominant (OR = 1.52, 95% CI = 1.15–2.00, $p = 0.003$), recessive (OR = 1.39, 95% CI = 1.01–1.90, $p = 0.042$), and log-additive (OR = 1.33, 95% CI = 1.11–1.59, $p = 0.002$) models, but not with CC.

FPRP results

The FPRP analysis was used to validate the study's positive findings. Significant correlations between the *SMAD7* gene polymorphism and CRC were discovered in both the entire group and subgroup, as indicated in Table S1. All of these associations were statistically significant (FPRP < 0.2).

Table 6. Best models to predict colorectal cancer by MDR.

Model	Testing Bal. Acc.	CVC
rs11874392	0.5316	10/10
rs11874392, rs7226855	0.5201	5/10
rs6507874, rs11874392, rs12953717	0.523	7/10
rs6507874, rs11874392, rs12953717, rs7226855	0.5223	9/10
rs6507874, rs11874392, rs4939827, rs12953717, rs7226855	0.5244	10/10

Bal. Acc., balanced accuracy; CVC, cross-validation consistently; MDR, multifactor dimensionality reduction.
The model with the maximum testing accuracy and maximum CVC was considered the best model.

The influence of SNP-SNP interactions on CRC

MDR approaches were utilized to investigate the relationship between SNP-SNP interactions and CRC. According to Table 6, the strongest predictive model for CRC was rs11874392, with the perfect CVC (10/10) and highest testing accuracy (0.5316). As illustrated in Figure 4, the interaction map revealed a positive synergistic interaction in rs7226855 × rs6507874 (0.08%), and rs7226855 × rs11874392 (0.01%), whereas the map with negative percent entropy suggested the redundancy or independence of each pair of SNP combinations.

SMAD7 mRNA expression in CRC

Through bioinformatics analysis, it was found that *SMAD7* was significantly lower expressed in colon (Figure 5(a)) and rectum adenocarcinoma (Figure 5(b)) compared to normal tissues.

Discussion

CRC is a complex illness influenced by a complex combination of environmental and genetic variables. Furthermore, it is crucial to

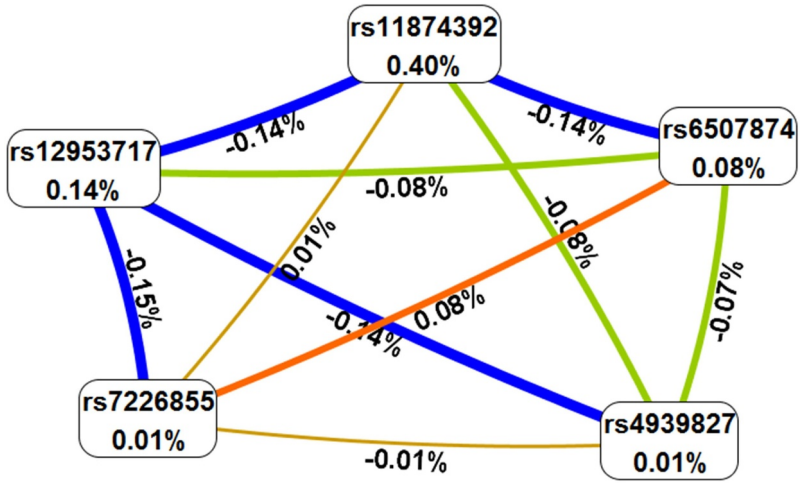

Figure 4. Fruchterman-Reingold of SNP-SNP interactions. Values in nodes represent the information gains of individual attributes (main effects). Values between nodes are information gains of each pair of attributes (interaction effects). Red and brown with positive percent entropy indicate synergistic interaction. Green, blue, and light brown with negative percent entropy indicates redundancy or independence.

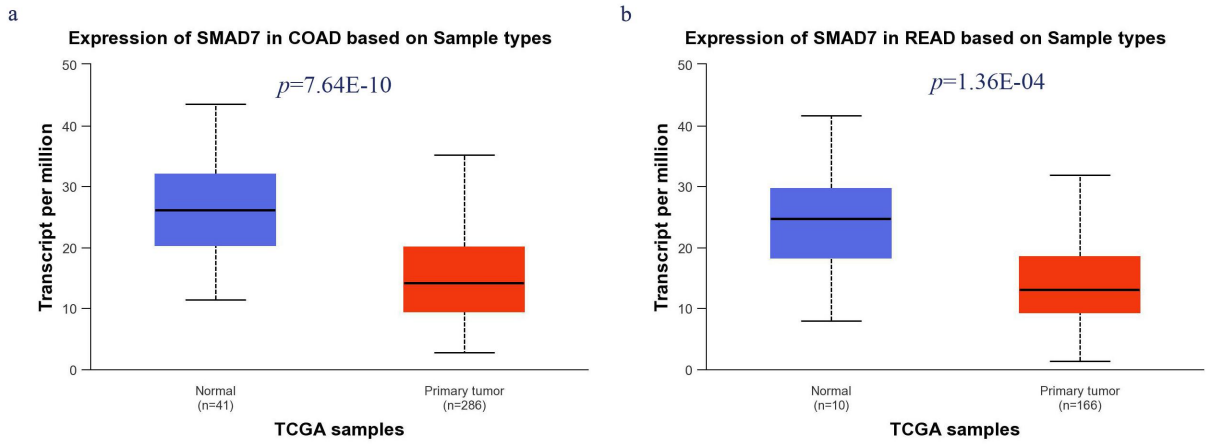

Figure 5. Expression of the *SMAD7* gene in CRC. (a) Expression of the *SMAD7* gene in COAD (primary tumor vs. 41 normal tissues = 286 vs. 41). (b) Expression of the *SMAD7* gene in AEAD (primary tumor vs. normal tissues = 166 vs 10). COAD, colon adenocarcinoma. READ, rectum adenocarcinoma.

take into account the elderly population when evaluating the prognosis of CRC. Even though older patients are more prone to severe postoperative complications, there is no consensus that age affects survival outcomes. The prognosis of older patients may be confounded by differences in stage at presentation, tumor site, preexisting comorbidities, and type of treatment received [29]. Previous studies have also explored the role of genetic factors in CRC risk, such as *SMAD7* gene polymorphisms, but the results are contradictory. This study was conducted to further investigate the roles of *SMAD7* gene polymorphisms in CRC risk. Our study indicated that *SMAD7* rs11874392 is related to an increased susceptibility to CRC. Furthermore, this genetic link was exclusively detected in RC, and no correlation was found in CC. This discovery provides us with a new perspective and helps us better understand the pathogenesis of CRC.

The rs11874392 A/T is located in the intron region of the *SMAD7* gene. An association study based on the population of Wuhan, China, shows that rs11874392 is associated with an increased risk of CRC [20]. The same correlation was found in another Japanese Americans of a multi-ethnic cohort [25]. However, Jiang et al. [24] find that this locus is related to a lower risk of CRC. In this study, rs11874392 showed an increased association with the risk of CRC. The contradiction in the results may be attributed to the heterogeneity of study samples, variations in research design, and disparities in research methods. Although there is currently no report on the biological function of rs11874392, we used Haploreg4.2 online software to predict its potential functions and found that rs11874392 could influence the regulation of promoter histone marks, enhancer histone marks, DNAse, and bound proteins. Taken together, we hypothesize that rs1187392 may enhance CRC susceptibility by affecting the aforementioned functions of *SMAD7*, which needs further experimental verification.

Tobacco smoking and alcohol drinking are modifiable risk factors for CRC. Compared to individuals who do not smoke or drink, those who engage in smoking and drinking have a 1.14–1.25 times higher relative risk of developing CRC [30,31]. In this study, we further determined the association between *SMAD7* SNPs and CRC susceptibility under smoking- and drinking status-based stratification. We found that rs11874392 T > A was able to increase the risk of CRC in subjects with smoking and drinking. Consistent with our findings, Zhong et al. [20] have shown that individuals who carry the variant genotype of rs11874392 among smokers are associated with an increased risk of CRC. These results suggest that individuals who carry the variant genotype of the rs11874392 should strictly control their drinking and smoking habits to reduce their risk of CRC.

SNPs in the *SMAD7* gene have different effects on tumor location of CRC. Curtin et al.'s research shows a significant correlation between two *SMAD7* SNPs and distal colon tumors, but no correlation was found with rectal and proximal colon tumors [32].

Similarly, Slattery et al. reported a stronger association between the *SMAD7* gene polymorphisms and distal CC compared to proximal CC [33]. Additionally, another study also showed a significant correlation between rs12953717 and distal colon tumor, but no significant correlation with proximal colon tumor and rectal tumor [24]. Similar to previous study results, our study demonstrated that rs11874392 was associated with the risk of RC, but not with CC, indicating that *SMAD7* rs11874392 has a specific influence on the anatomical site in CRC progression.

There are some limitations in this study. First, our research only includes Han people in China, so our results may not be applicable to other people or other ethnic groups in China. A larger and more diverse sample study is needed to fully understand the role of *SMAD7* gene polymorphism in the risk of CRC in future. Second, although previous studies show that gene polymorphism may be related to the expression of *SMAD7*, further experiments are needed to confirm this relationship. This will contribute to a deeper understanding of the function of *SMAD7* gene polymorphisms and its mechanism in CRC. Despite its shortcomings, our study provides a new perspective and helps us better understand the pathogenesis of CRC.

Conclusion

In brief, our study demonstrated that *SMAD7* rs11874392 may enhance the risk of CRC in the Chinese population. Future research will focus on exploring the function of SNPs and its relationship with *SMAD7* expression, so as to fully understand the mechanism of *SMAD7* gene polymorphism in the occurrence and development of CRC.

Acknowledgements

The authors thank all participants and volunteers in this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Author contributions

Yongsheng Wu contributed to design the study and write the manuscript. Jue Xu, Biaobin Tan, Ting Yi, and Su Liu contributed to recruit and collect study samples. Guang Yang and Kai Li contributed to analyze the data. Xinhan Zhao contributed to conceive the study and revise manuscript. All authors read and approved the final manuscript.

Consent to publish statement

All authors agree to publish this manuscript in this journal.

Data availability statement

Participant informed consent statements did not seek consent for data to be made publicly available; however, data will be made available to individual researchers upon reasonable request. In this case, please contact the corresponding author.

Ethics approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of The People's Hospital of XiangXiang and the 1964 Helsinki declaration.

Informed consent from participants

Informed written consent was obtained from each participant before the research.

References

[1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and

mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424. doi: 10.3322/caac.21492
[2] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7–34. doi: 10.3322/caac.21551
[3] Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl Oncol. 2021;14 (10):101174. doi: 10.1016/j.tranon.2021.101174
[4] Riihimäki M, Hemminki A, Sundquist J, et al. Patterns of metastasis in colon and rectal cancer. Sci Rep. 2016;6 (1):29765. doi: 10.1038/srep29765
[5] Zarour LR, Anand, S, Billingsley, KG et al. Colorectal cancer liver metastasis: evolving paradigms and future directions. Cell Mol Gastroenterol Hepatol. 2017;3 (2):163–173. doi: 10.1016/j.jcmgh.2017.01.006
[6] Brouwer NPM, Bos ACRK, Lemmens VEPP, et al. An overview of 25 years of incidence, treatment and outcome of colorectal cancer patients. Int J Cancer. 2018;143(11):2758–2766. doi: 10.1002/ijc.31785
[7] Adeleke S, Haslam A, Choy A, et al. Microsatellite instability testing in colorectal patients with lynch syndrome: lessons learned from a case report and how to avoid such pitfalls. Per Med. 2022;19(4):277–286. doi: 10.2217/pme-2021-0128
[8] Itatani Y, Kawada K, Sakai Y. Transforming growth factor-β signaling pathway in colorectal cancer and its tumor microenvironment. Int J Mol Sci. 2019;20(23). doi: 10.3390/ijms20235822
[9] Aghabozorgi AS, Ebrahimi R, Bahiraee A, et al. The genetic factors associated with Wnt signaling pathway in colorectal cancer. Life Sci. 2020;256:118006. doi: 10.1016/j.lfs.2020.118006
[10] AlDubayan SH, Giannakis M, Moore ND, et al. Inherited DNA-Repair defects in colorectal cancer. Am J Hum Genet. 2018;102(3):401–414. doi: 10.1016/j.ajhg.2018.01.018
[11] Ryan BM, Zanetti KA, Robles AI, et al. Germline variation in NCF4, an innate immunity gene, is associated with an increased risk of colorectal cancer. Int J Cancer. 2014;134(6):1399–407. doi: 10.1002/ijc.28457
[12] Uthoff SM, Duchrow M, Schmidt MHH, et al. VEGF isoforms and mutations in human colorectal cancer. Int J Cancer. 2002;101(1):32–36. doi: 10.1002/ijc.10552
[13] Cheng X, Zhang F, Gong J, et al. Identification of potential functional variants and genes at 18q21.1 associated with the carcinogenesis of colorectal cancer. PLoS Genet. 2022;18(2):e1010050. doi: 10.1371/journal.pgen.1010050
[14] Dolatkhah R, Somi MH, Kermani IA, et al. Increased colorectal cancer incidence in Iran: a systematic review and meta-analysis. BMC Public Health. 2015;15(1):997. doi: 10.1186/s12889-015-2342-9
[15] Zhao M, Mishra L, Deng CX. The role of TGF-β/SMAD4 signaling in cancer. Int J Biol Sci. 2018;14 (2):111–123. doi: 10.7150/ijbs.23230
[16] Yu ZL, Liu, J, Ning, ZK et al. The TGF-β/Smad(2/3) signaling pathway is involved in Musashi2-induced invasion and metastasis of colorectal cancer. Mol Carcinog. 2023;62(2):261–276. doi: 10.1002/mc.23484
[17] Halder SK, Beauchamp RD, Datta PK. Smad7 induces tumorigenicity by blocking TGF-beta-induced growth inhibition and apoptosis. Exp Cell Res. 2005;307 (1):231–246. doi: 10.1016/j.yexcr.2005.03.009
[18] Bayat Z, Ghaemi Z, Behmanesh M, et al. Hsa-miR-186-5p regulates TGFβ signaling pathway through expression suppression of SMAD6 and SMAD7 genes in colorectal cancer. Biol Chem. 2021;402(4):469–480. doi: 10.1515/hsz-2019-0407
[19] Li J, Zou, L, Zhou, Y et al. A low-frequency variant in SMAD7 modulates TGF-β signaling and confers risk for colorectal cancer in Chinese population. Mol Carcinog. 2017;56(7):1798–1807. doi: 10.1002/mc.22637
[20] Zhong R, Liu, L, Zou, L et al. Genetic variations in the TGFβ signaling pathway, smoking and risk of colorectal cancer in a Chinese population. Carcinogenesis. 2013;34(4):936–942. doi: 10.1093/carcin/bgs395
[21] Fukushima T, Mashiko M, Takita K, et al. Mutational analysis of TGF-beta type II receptor, Smad2, Smad3, Smad4, Smad6 and Smad7 genes in colorectal cancer. J Exp Clin Cancer Res. 2003;22(2):315–20.
[22] Yin R, Song, B, Wang, J et al. Genome-wide association and transcriptome-wide association studies identify novel susceptibility genes contributing to colorectal cancer. J Immunol Res. 2022;2022:5794055. doi: 10.1155/2022/5794055
[23] Akbari Z, Safari-Alighiarloo N, Asadzadeh Aghdaei H, et al. The association between SMAD7 polymorphisms and colorectal cancer susceptibility as well as clinicopathological features in the Iranian population. Gastroenterol Hepatol Bed Bench. 2020;13(1):23–30.
[24] Jiang X, Castelao JE, Vandenberg D, et al. Genetic variations in SMAD7 are associated with colorectal cancer risk in the colon cancer family registry. PLoS One. 2013;8(4): e60464. doi: 10.1371/journal.pone.0060464
[25] Cologne J, Loo L, Shvetsov YB, et al. Stepwise approach to SNP-set analysis illustrated with the metabochip and colorectal cancer in Japanese Americans of the multiethnic cohort. BMC Genomics. 2018;19(1):524. doi: 10.1186/s12864-018-4910-8
[26] Thompson CL, Plummer SJ, Acheson LS, et al. Association of common genetic variants in SMAD7 and risk of colon cancer. Carcinogenesis. 2009;30 (6):982–6. doi: 10.1093/carcin/bgp086
[27] Ho JW, Choi S-C, Lee Y-F, et al. Replication study of SNP associations for colorectal cancer in Hong Kong Chinese. Br J Cancer. 2011;104(2):369–75. doi: 10.1038/sj.bjc.6605977
[28] Thean LF, Li HH, Teo YY, et al. Association of Caucasian-identified variants with colorectal cancer risk in Singapore Chinese. PLoS One. 2012;7(8): e42407. doi: 10.1371/journal.pone.0042407
[29] Osseis M, Nehmeh, WA, Rassy, N et al. Surgery for T4 colorectal cancer in older patients: determinants of outcomes. J Pers Med. 2022;12(9):1534. doi: 10.3390/jpm12091534

[30] McNabb S, Harrison TA, Albanes D, et al. Meta-analysis of 16 studies of the association of alcohol with colorectal cancer. Int J Cancer. 2020;146 (3):861–873. doi: 10.1002/ijc.32377

[31] Botteri E, Borroni E, Sloan EK, et al. Smoking and colorectal cancer risk, overall and by molecular subtypes: a meta-analysis. Am J Gastroenterol. 2020;115 (12):1940–1949. doi: 10.14309/ajg.0000000000000803

[32] Curtin K, Lin W-Y, George R, et al. Meta association of colorectal cancer confirms risk alleles at 8q24 and 18q21. Cancer Epidemiol Biomarkers Prev. 2009;18 (2):616–21. doi: 10.1158/1055-9965.EPI-08-0690

[33] Slattery ML, Herrick, J, Curtin, K et al. Increased risk of colon cancer associated with a genetic polymorphism of SMAD7. Cancer Res. 2010;70(4):1479–1485. doi: 10.1158/0008-5472.CAN-08-1792