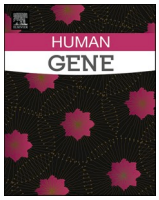
Human Gene 38 (2023) 201226

Contents lists available at ScienceDirect

Human Gene

journal homepage: www.journals.elsevier.com/human-gene



The significant role of a functional polymorphism rs9637231 in long non-coding RNA, *LINC02892* in colorectal cancer: Evidence from an Iranian cohort

Seyyed Reza Hashemi [a,b,c], Maryam M. Matin [d,i,j], Amir Avan [a,b], Saba Farrokhi [e], Farideh Rigi [f], Reza Khayami [b,g], Marjan Azghandi [c], Manizhe Esmaeilnia [c,h], Mohammad Amin Kerachian [a,b,c,*]

[a] Medical Genetics Research Center, Mashhad University of Medical Sciences, Mashhad, Iran
[b] Department of Medical Genetics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
[c] Cancer Genetics Research Unit, Reza Radiotherapy, and Oncology Center, Mashhad, Iran
[d] Stem Cell and Regenerative Medicine Research Group, Academic Center for Education, Culture and Research (ACECR)-Khorasan Razavi, Mashhad, Iran
[e] Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
[f] Razavi Cancer Research Center, Razavi Hospital, Imam Reza International University, Mashhad, Iran
[g] Cancer Research Center, Mashhad University of Medical Sciences, Mashhad, Iran
[h] Department of Biology, Faculty of Sciences, Mashhad Branch, Islamic Azad University, Mashhad, Iran
[i] Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Vakilabad Blvd, Mashhad, Iran
[j] Novel Diagnostics and Therapeutics Research Group, Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran

ARTICLE INFO

Edited by Mehdi Sadeghi

Keywords:
Colon
Single nucleotide polymorphism
LncRNA
Association study

ABSTRACT

Background: Long noncoding RNAs (lncRNAs) are known to be involved in cancer development and progression. Single nucleotide polymorphisms (SNPs) in lncRNAs may affect the structure, expression, and function of lncRNAs, influencing the cancer risk. In the present case-control study, we investigated the association of LINC02892 SNP rs963723l with colorectal cancer (CRC).
Methods: A total of 1130 samples (530 cases and 600 controls), were collected to determine the association between rs9637231 C > T and CRC. Tetra Primer ARMS-PCR was used to determine the genotypes. Statistical analysis was performed employing logistic regression.
Results: The analysis results indicated that individuals with TT genotype had 89% (95%CI = 1.40–2.57, $P <$ 1.00E-3) and 53% (OR = 1.53, 95%CI = 1.12–2.1, P = 8.00E-3) higher risks of CRC compared to wild-type homozygotes and C allele carriers, respectively. Furthermore, T allele carriers demonstrated a 72% (OR = 1.72, 95%CI = 1.31–2.25, $P < 0.001$) increased risk of CRC in comparison with wild-type homozygotes. The outcomes were adjusted for factors such as age, gender, and smoking status.
Conclusions: SNP rs9637231 might be a risk factor for CRC.

1. Background

Colorectal cancer (CRC) is the third most common cancer and the second cause of cancer-related death globally. According to GLOBOCAN 2020 estimation, the incidence of CRC in the world is 19.5 per 10,000 (Siegel et al., 2021; Siegel et al., 2020). The mortality is attributed to modifiable risk factors such as smoking, unhealthy diet, high alcohol consumption, physical inactivity, and excess body weight (Arani and Kerachian, 2017). In the past decade, diagnosis of CRC has been highly improved; however, distant metastasis, especially liver metastasis, leads

Abbreviations: CRC, Colorectal cancer; lncRNA, Long non-coding RNA; SNP, Single nucleotide polymorphism; MAF, Minor allele frequency; OS, Overall survival; EMT, Epithelial-mesenchymal transition; GTEx, The Genotype-Tissue Expression; eQTl, Expression quantitative trait locus; ESCs, Embryonic stem cells; iPSCs, induced pluripotent stem cells.

* Corresponding author at: Department of Medical Genetics, Faculty of Medicine, Mashhad University of Medical Sciences, Azadi Square, Mashhad 917794-8564, Iran.

E-mail addresses: sayedrezahashemi20@gmail.com (S.R. Hashemi), amin.kerachian@mail.mcgill.ca, kerachianma@mums.ac.ir (M.A. Kerachian).

https://doi.org/10.1016/j.humgen.2023.201226
Received 12 July 2023; Received in revised form 3 September 2023; Accepted 27 September 2023
Available online 30 September 2023

to poor prognosis and high mortality (Fearnhead et al., 2002; Talebi et al., 2021). Therefore, the early detection of CRC is key in the diagnosis of patients with acute symptoms and the adverse courses of disease (Ćwik et al., 2012; Kerachian et al., 2020).

Different molecular mechanisms have been attributed to the development and progression of CRC. Notably, emerging research highlights that long noncoding RNAs (lncRNAs) could play crucial roles in carcinogenesis through several mechanisms, such as the cell cycle, cell differentiation, cell proliferation and gene expression. LncRNAs could serve as tumor suppressor or oncogenic agents or even have dual effects, adding complexity to their role in CRC (Poursheikhani et al., 2021; Prensner and Chinnaiyan, 2011). Some oncogenic lncRNAs, including *CACS1, CCAT* family, and *PVT1,* have been implicated in potentially driving CRC through interactions with *Myc* (a family of regulator genes and proto-oncogenes that code for transcription factors), stimulating proteins, or other genes associated with the *Wnt* pathway at the post-transcriptional level. This engagement often results in promoting cell proliferation, differentiation, and stem cell self-renewal. In addition, lncRNAs could contribute to invasion, migration, and metastasis in CRC. For instance, lncRNA *SNHG1* expedites tumor invasion by promoting Epithelial-mesenchymal transition (EMT) phenomenon through the sequestration of miR497-5p and miR195-5p in colon adenocarcinoma cells (Shen et al., 2017).

Within the spectrum of colorectal cancer (CRC) risk factors, single nucleotide polymorphisms (SNPs) stand out as contributors that can increase the likelihood of cancer occurrence (Tenesa and Dunlop, 2009). The impact of mutations and variations, including SNPs, in non-coding genomic regions, particularly lncRNAs, has been well established in the structure, expression, and function of lncRNAs, consequently influencing the cancer susceptibility (Zhang et al., 2017; Plassais et al., 2016; Li et al., 2017; Peng et al., 2017; Korniłłowicz-Kowalska and Bohacz, 2011). For example, it has been revealed that 15% of all transcribed single nucleotide variants (SNVs) in a trio family (comparing father, mother, and child) were capable of modifying RNA structure (Wan et al., 2014). A study by Jin et al., unveiled an association between rs10845671 and RP11-392P7.6 lncRNA, suggesting a potential novel biomarker for CRC (Jin et al., 2017). Furthermore, a correlation was identified between rs55829688 and rs1951625 in the *GAS5* promoter and CRC (Wang et al., 2019). In another study by Wu et al., rs664589 in *MALAT1* increased the risk of CRC (Wu et al., 2019). Similarly, the genetic variants in lncRNA *h19* showed an association with CRC risk, where the A allele of rs2839698 was found to inrease the risk of CRC compared to the G allele (Li et al., 2016). In our previous investigation, we identified a novel long non-coding RNA (lncRNA), LINC02892, which exhibited notably elevated expression levels in tumors when compared to normal tissue (Kerachian and Azghandi, 2022). In the current study, our focus was on examining the potential link between rs9637231, located within the second exon of all three transcripts of LINC02892, and CRC.

2. Methods

2.1. Study participants

In total, 1130 samples, including 530 CRC cases and 600 controls, were acquired for this study. Blood samples were recruited from Reza Radiotherapy and Oncology Center, Razavi Hospital and Ghaem Hospital of Mashhad, Iran. Samples of the healthy controls were collected from patients referred to Reza Radiotherapy and Oncology Center for screening and colonoscopy. Ethical approval was obtained from the ethical committee of Mashhad University of Medical Sciences, Mashhad, Iran (ethics code no.: 990383). The written informed consent was obtained from all participants. Inclusion criteria for the CRC patient group included patients who had positive colonoscopy reports for CRC and no first-relative family history of cancer. For the control group, inclusion criteria included individuals who had negative colonoscopy reports of CRC or any other colorectal diseases with no first-relative family history of cancer. Exclusion criteria for the patient and control groups were to identify a secondary GI disease during the study. The treatment history of the patients was disregarded in order to concentrate solely on genetic variations and eliminate any potential ascertainment biases.

2.2. Genotyping

The rs9637231 polymorphism was detected in DNA extracted from whole blood using Tetra Primer Amplification Refractory Mutation System PCR (ARMS-PCR). The DNA was extracted from 300 ul of blood using the standard salting-out method (Mohammadpour, 2018). Primers were designed using the online tool primer1 (http://cedar.genetics.soton.ac.uk/public_html/primer1.html). The sequences of the primers were as follows, inner primer forward: 5'-CAATTGATAGGCATCCACACAGT-3′; inner primer reverse: 5'-TGCATCATTCTCCCCAACG-3′; outer primer forward: 5'-AGAATCCTGCACCCTCACAA-3′; outer primer reverse: 5'-CAGCCATGAAAGCCCACATT-3′. Genotyping was performed twice by two different individuals (Peng et al., 2017).

2.3. Functional in silico analysis of SNP

To predict functional effects of rs9637231, HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) (Ward and Kellis, 2016) and regulomeDB 2.0 (https://regulomedb.org/regulome-search/) (Boyle et al., 2012) were used. In addition, eQTLs of rs9637231 were obtained from GTEx database V8 (https://gtexportal.org/home/) (Lonsdale et al., 2013). STRING version 12.0 was used to create a network of interacting proteins (Szklarczyk et al., 2023).

2.4. Statistical analysis

The association of the SNP with CRC in different genetic models was calculated using unconditional logistic regression. The Chi-square test was applied to evaluate the Hardy–Weinberg equilibrium. Missing data were handled via Multiple Imputation by Chained Equations (MICE) using five predictor matrices and 50 iterations (Azur et al., 2011) by R programming. *P*-value <0.05 was defined as the significant value.

3. Results

3.1. Functional annotation and eQTL analysis of rs9637231

Data on minor allele frequency (MAF) of rs9637231 and its chromosomal location is summarized in Supplementary Table 1. According to RegulomeDB data, rs9637231 exhibits potential effects on transcription factor binding or DNase peak (rank = 4, score = 0.60906). Furthermore, Haploreg analysis reveals its potential localization within enhancer and promoter histone marks across various cell types such as ESCs and iPSCs. Notably, an alteration in the NRSF motif (The repressor element 1 (RE1) silencing transcription factor) was observed (Supplementary Table 2).

To explore potential regulatory impacts, eQTL analysis was conducted using the GTEx portal. This analysis identified significant correlations between the expression levels of *SPATC1L*, *FTCD, PRMT2* and rs9637231 were identified (Table 1, Fig. 1). Furthermore, DNA extraction results are illustrated in Fig. 2, while Fig. 3 displays the ARMs PCR outcomes. Also, STRING showed that *DIPA2* is co-expressed with *SPATC1L*, *PRMT2*. (See Fig. 4.)

3.2. Association study of rs9637231

The feature distribution details, and demographic characteristics of samples are shown in Table 2. The genotype frequency of rs9637231 did not conform to Hardy–Weinberg equilibrium ($P < 0.05$, Supplementary Table 3). Hence, all samples were genotyped twice by two individuals

Table 1
Genes with significant correlation on GTEx.

Gencode Id	Gene Symbol	P- Value	NES[a]	Tissue
ENSG00000160284.14	*SPATC1L*	1.40E-07	0.36	Colon - Transverse
ENSG00000160284.14	*SPATC1L*	7.4E-06	0.3	Colon - Sigmoid
ENSG00000160282.13	*FTCD*	0.000016	0.26	Colon - Sigmoid
ENSG00000160310.17	*PRMT2*	0.00016	0.09	Colon - Transverse
ENSG00000160305.17	*DIP2A*	9.20E-08	−0.11	Whole Blood
ENSG00000160284.14	*SPATC1L*	3.10E-07	0.2	Whole Blood
ENSG00000182362.13	*YBEY*	7.10E-07	0.21	Whole Blood
ENSG00000160285.14	*LSS*	0.000035	−0.12	Whole Blood
ENSG00000160310.17	*PRMT2*	0.000076	0.086	Whole Blood

[a] Normalized effect size (NES), "previously known as the effect size on the portal, is defined as the slope of the linear regression and is computed as the effect of the alternative allele (ALT) relative to the reference allele (REF) in the human genome reference (www.gtexportal.org).

using ARMS-PCR and any technical errors were ruled out. In addition, 10% of samples were genotyped in each group by Sanger sequencing.

A significant association between rs9637231 and CRC was found in codominant, dominant, recessive and over dominant models. The allele and genotype frequency and association of rs9637231 with CRC risk are shown in Table 3.

As shown in Supplementary Table 4, among the CRC patients aged 31 to 60, when considering the rs9637231 CC genotype as the reference, individuals carrying the TT genotype exhibited a substantial increase in the risk of CRC by 104% in the codominant model (OR = 2.04, 95%CI = 1.41–2.96, P = 1.60E-04). Female and male individuals with TT genotype exhibited a notably higher risk of colorectal cancer (CRC), with increases of 74% and 95% respectively, in comparison to individuals with the CC genotype (female: OR = 1.74, 95%CI = 1.15–2.62, P = 8.46E-03; male: OR = 1.95, 95%CI = 1.3–2.93, P = 1.34E-03). Furthermore, addicted individuals had more amplified risk of CRC in TT genotypes in comparison with CC genotype individuals (addicted: OR = 2.93, 95%CI = 1.37–6.28, P = 6.02E-03; not addicted: OR = 1.64, 95% CI = 1.19–2.26, P = 2.75E-03). A comprehensive analysis was conducted to assess the impact of each genotype under various environmental variables, revealing that advanced age increased the risk of CRC across all groups (Supplementary Table 5).

4. Discussion

The primary objective of the current study was to investigate, for the first time in an Iranian population, the link between rs9637231 and CRC. This particular SNP is situated within the second exon of a gene that codes for a lncRNA. This lncRNA has been previously identified and documented by our laboratory under the identifiers: Banklt2400105, LINC02892, MW248922; Banklt2400122, LINC02892, MW248923; Banklt2400131, LINC02892, MW248924; Banklt2400132, LINC02892,

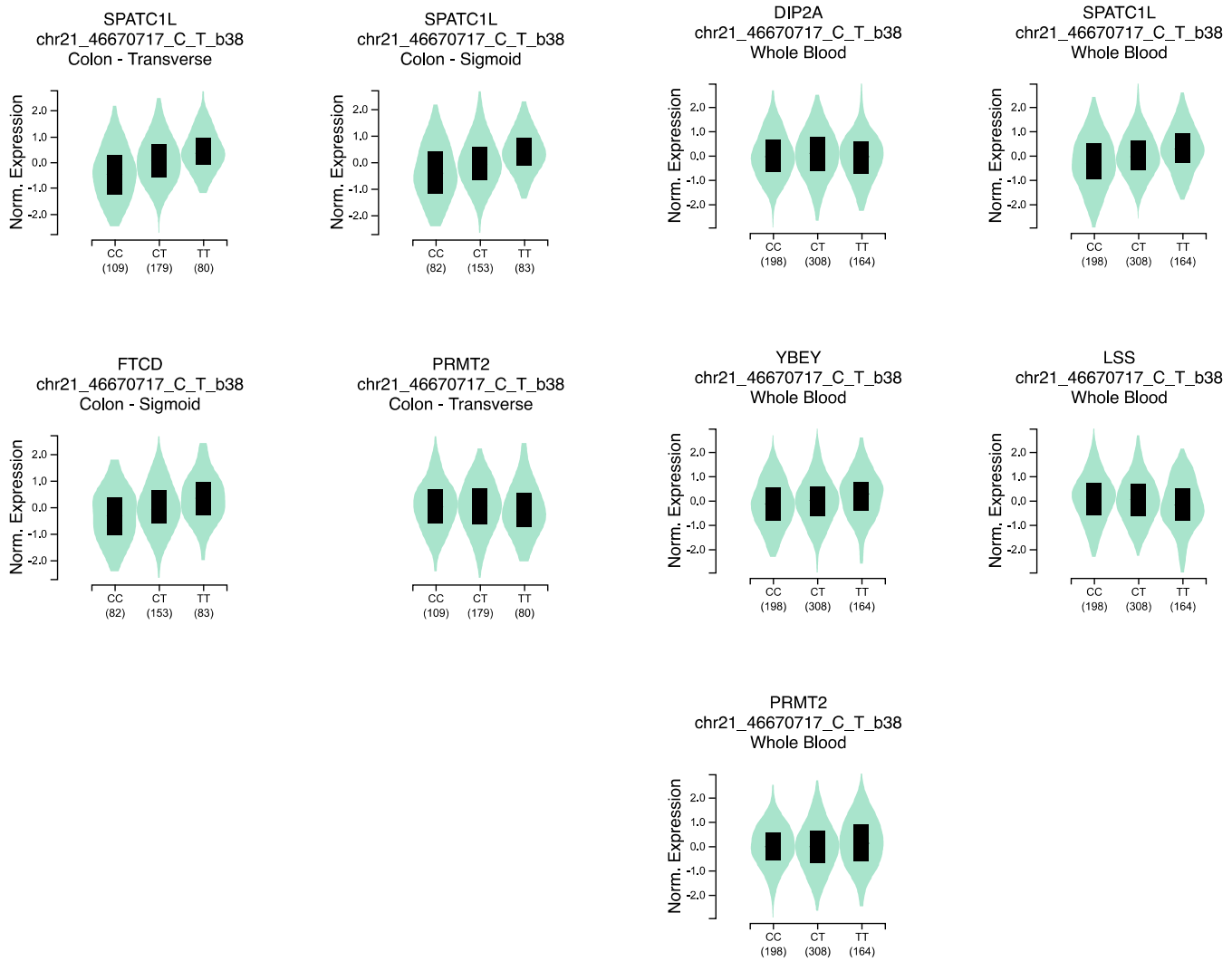


Fig. 1. Correlations between the expression of *SPATC1L*, *FTCD, PRMT2* genes and rs9637231.

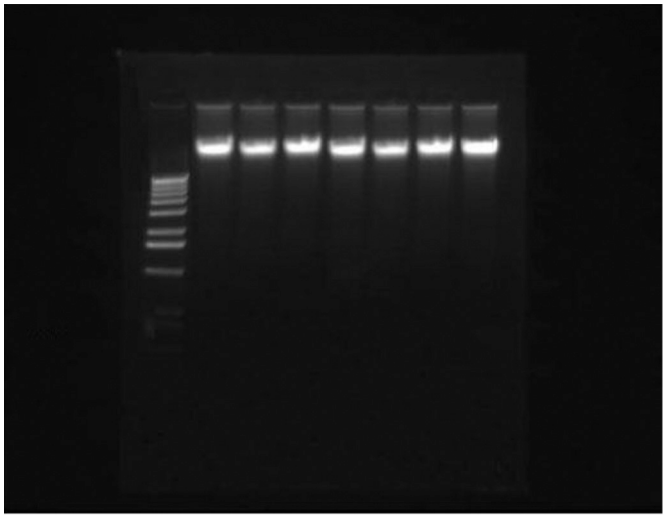


Fig. 2. The result of DNA extraction.

MW248925. Among a sample size of 1130 individuals, the frequency of the T allele (the risk allele) was observed to be 58% in the patient group, and 48% in the control group (Supplementary Table 2). We found that people with the TT genotype were more susceptible to CRC (OR = 1.89, 95%CI = 1.4–2.57, $P < 1.00E-03$). Furthermore, the T allele-carriers had a significantly higher risk of CRC than the wild-type homozygotes (OR = 1.72, 95%CI = 1.31–2.25, $P < 1.00E-03$, Supplementary Table 5).

Certain studies have reported that the activation of deaminases could play a role in CRC development and progression by elevating cytidine deaminase levels up to 1.34-fold (p value = 0.0001) (Han et al., 2008). In a study by Meshcheryakova et al., evidence was presented that Activation-Induced Cytidine Deaminase (AID)-positive ectopic lymphoid structures were gathered within the metastatic margin in CRC patients (Meshcheryakova et al., 2014), suggesting a role for AID in the development of CRC (Endo et al., 2008). Cytosine, particularly 5-methylcytosine (m5C), is chemically less stable compared to other nucleobases, and can undergo deamination, leading to the formation of uracil and thymine through the action of deaminase enzymes (Stier and Kiss, 2013). Given that DNA repair pathways are dysregulated in cancer (Li et al., 2020), it is hypothesized that the conversion of thymine back to cytosine might be hindered, providing a possible explanation for the observed elevation of the T allele of rs9637231 in patients with CRC.

T allele carriers between the ages of 31 and 60 were more vulnerable to CRC risk than CC homozygotes (OR = 1.91, 95%CI =1.37–2.66, $P < 1.00E-03$, Supplementary Table 4). As indicated the Supplementary Table 5, across all three genotypes (CC, CT and TT), individuals aged 61–90 exhibited a higher susceptibility to CRC compared to those aged 31–60, underscoring the influence of aging in cancer. Both female and male T allele carriers had a higher risk than their control counterparts did, with males having a slightly higher risk than females for cancer. In addition, it was observed that T carriers had a higher risk in both addicted to nicotine and non-addicted to nicotine groups. Remarkably, T allele carriers with addiction showed a more pronounced risk elevation than T allele carriers without addiction (Supplementary Table 5).

The effect of SNPs on CRC risk has been previously investigated. Notably, the lncRNAs *HOTAIR, HOTTIP*, and *MALAT1* have emerged as prominent examples of mutated lncRNAs in malignancies, exerting substantial influence on healthcare practices. The presence of SNPs within these lncRNAs have been found across numerous studies, suggesting their potential involvement in critical cancer-related processes. Elevated levels of serum *HOTTIP* in conjunction with the presence of the rs1859168 A > C have been related to distant metastasis, lymph node metastasis, and grade III CRC. Both rs18159168 and high *HOTTIP* have been associated with increased risk for CRC development (Ali et al., 2020). Moreover, Lv et al. found that rs3807598, rs2067087 and rs17427960 within lncRNA *HOTTIP* were linked with increased CRC risk, with the strength of this association being more pronounced in the stratified analysis. Additionally, they suggested that rs17501292 might contribute to enhancement of the overall survival (OS) of CRC patients with ulcerative/invasive tumors (Lv et al., 2019). Furthermore, Kim et al. found that the *HOTAIR* gene variant rs7958904 G > C is related to CRC prevalence and death (Kim et al., 2020). Similarly, there was a significant association between polymorphisms in the *SNHG16* gene (rs7353, rs8038, and rs15288) with the risk of developing CRC. The expression of the lncRNA *SNHG16* was linked with tumor development in patients with CRC. There was a substantial drop in plasma *SNHG16* levels due to a variation at rs7353 site (A > G), whereas rs8038 (G > A) and rs15278 (A > G) variations resulted in higher levels (Zhou et al., 2020). According to Ming-li Yang et al., the *PCAT1* rs2632159 SNP was associated with an increased CRC risk of 1.37-fold in the dominant model and 2.19-fold in the recessive model (Yang et al., 2019). Additional investigations highlighting the association of SNPs in lncRNAs with CRC are listed in Table 4.

According to our GTEx results (Table 1), individuals with the rs9637231 risk allele (T) exhibited considerably higher levels of *SPATC1L* expression in the transverse and sigmoid colons, as well as peripheral blood samples. Recent research has indicated that methylation in a cluster of six CpG sites at *SPATC1L* is associated with a reduced risk of CRC (Van Baak et al., 2018). In addition, the T allele demonstrated an association with higher *PRMT2* expression levels in transverse colon and blood samples. Su et al. reported that the silencing of the *PRMT2* gene could effectively reduce Wnt signaling, a signaling pathway contributing to carcinogenesis, in a CRC cell line (Su et al., 2014). As a result, a plausible explanation for the heightened risk observed in T allele carriers might be attributed to this upregulation in PRMT2

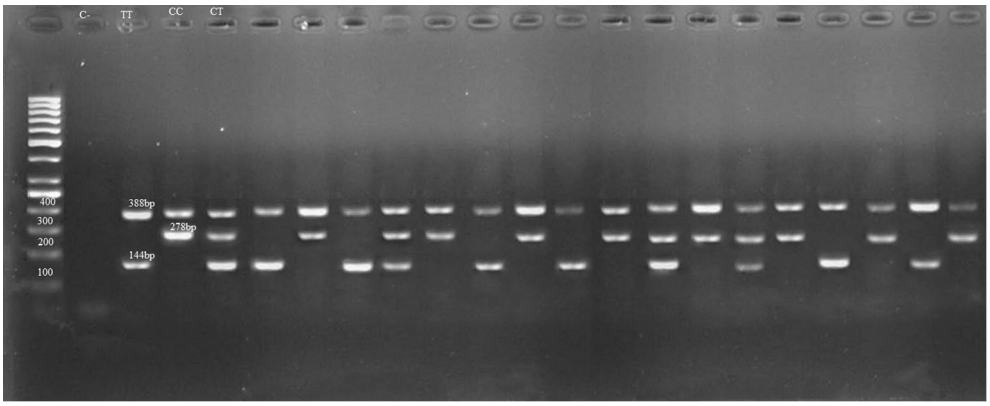


Fig. 3. The result of tetra primer ARMs PCR.

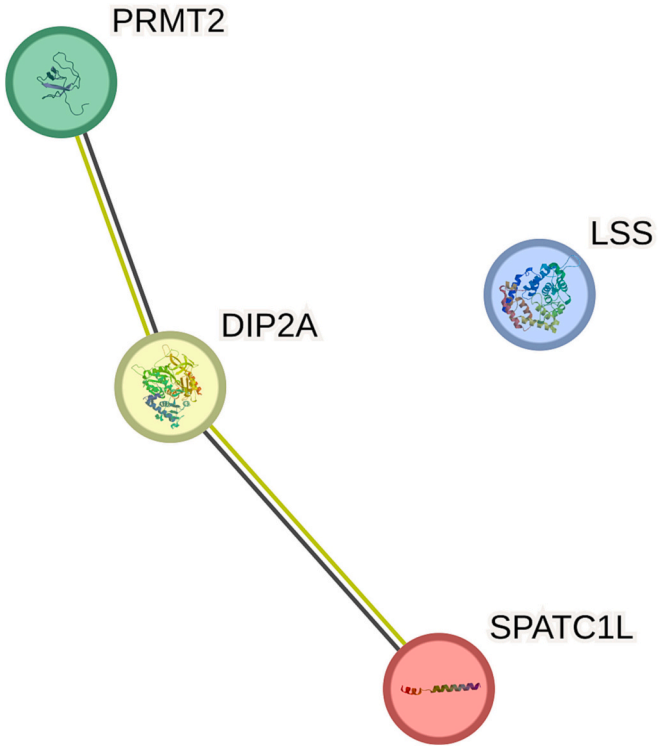


Fig. 4. Interaction network of *SPATC1L*, *FTCD, PRMT2* and *LSS*.

Table 2
Demographic data and characteristic distribution of samples.

	Control (N = 600)	Case (N = 530)
Age		
≤30	85.0 (14.2%)	12.0 (2.3%)
31-60	389 (64.8%)	288 (54.3%)
61-90	101 (16.8%)	203 (38.3%)
Missing	25.0 (4.2%)	27.0 (5.1%)
Gender		
Female	312 (52.0%)	248 (46.8%)
Male	281 (46.8%)	279 (52.6%)
Missing	7.00 (1.2%)	3.00 (0.6%)
Addiction		
No	482 (80.3%)	315 (59.4%)
Yes	117 (19.5%)	76.0 (14.3%)
Missing	1.00 (0.2%)	139 (26.2%)

expression. In addition, the GTEx analysis showed that *DIP2A* and *LSS* were less expressed in TT genotypes. Kudo-Saito et al. demonstrated that FSTL1-induced $CD11b^{+}DIP2A^{+}LAG3^{+}$ cells have a critical role in immune dysfunction in CRC (Kudo-Saito et al., 2021). In addition, *FSTL1* was significantly upregulated in tumor tissues and peripheral blood of CRC patients (Kudo-Saito et al., 2021). Furthermore, reduced expression of *LSS* has been shown in HT-29 CRC cell line compared to normal colon cells (Xu et al., 2021).

5. Conclusions

The findings of this study suggest that rs9637231 could increase susceptibility of patients to CRC. According to the results, T allele carriers exhibited increased risks across codominant, dominant, recessive and log-additive models. Consequently, the SNP rs9637231 emerges as a potential risk factor for CRC within our population, potentially offering insights into CRC diagnosis. While the present study benefited from a collaborative effort across multiple research centers and a relatively large sample size, it is important to note that our data lacked comprehensive demographic and anthropometric information, preventing us from conducting a comprehensive risk factor analysis.

Authors' contributions

SRH, RK were responsible for writing the paper and the original draft. MAK and MMM were responsible for the study's conception. SRH, MA and ME were responsible for the investigation and conducting the experiments. MAK, MMM, AA, SF and RF provided the samples and analyzed the data. All authors were responsible for reviewing and editing the final version of the paper. All authors read and approved the final manuscript.

Funding

This study was funded by Mashhad University of Medical Science (Grant number: 990383).

Ethics approval and consent to participate

This study was ethically approved by the ethics committee approval of Mashhad University of Medical Sciences, Mashhad, Iran (Grant number: 990383) and all methods were performed in accordance with the relevant guidelines and regulations of The Declaration of Helsinki (DoH). Informed written consent had been obtained from all participants in this study.

Consent to publish

Not applicable.

Declaration of Competing Interest

The authors declare no conflict of interest for this research.

Table 3
The allele and genotype frequency and association of rs9637231 with the risk of colorectal cancer in codominant, dominant, recessive, and over dominant models.

Model	Genotype	Cases (%)	Controls	OR (95%CI)	P	OR[a] (95% CI)	P[a]
Codominant	CC	135 (25%)	218 (36%)	Ref		Ref	
	CT	172 (32%)	185 (31%)	1.5 (1.11–2.03)	7.65E-03	1.53 (1.12–2.1)	0.008E-03
	TT	223 (42%)	197 (33%)	1.83 (1.37–2.44)	<1.00E-03	1.89 (1.4–2.57)	<1.00E-03
Dominant	CC	135 (25%)	218 (36%)	Ref		Ref	
	CT-TT	395 (75%)	382 (64%)	1.67 (1.29–2.16)	<1.00E-03	1.72 (1.31–2.25)	<1.00E-03
Recessive	CC-CT	307 (58%)	403 (67%)	Ref		Ref	
	TT	223 (42%)	197 (33%)	1.49 (1.17–1.89)	1.36E-03	1.53 (1.18–1.97)	1.25E-03
Over dominant	CC-TT	358 (68%)	415 (69%)	Ref		Ref	
	CT	172 (32%)	185 (31%)	1.08 (0.84–1.39)	558E-03	1.08 (0.83–1.41)	575E-03
Log Additive	0,1,2	530 (100%)	600 (100%)	1.35 (1.17–1.56)	<1.00E-03	1.37 (1.18–1.6)	<1.00E-03

[a] Adjusted for age, gender, and addiction.

Table 4
Characteristics of recently studied single nucleotide polymorphisms in lncRNAs and the risk of colorectal cancer.

Author (Ref)	LncRNA	Year	Ethnicity	Case	Control	Method	SNP ID	Allele	Association with CRC
(Zhou et al., 2020)	HOTTIP	2019	China	884	964	Kompetitive allele specific PCR(KASP)	rs3807598 rs17501292 rs2067087 rs17427960	G/C T/G C/G A/C	GG vs. CG + CC [OR = 1.4,P = 0.02] Negative CC vs. GC + GG [OR = 1.52,P < 0.001] AA+CA vs. CC [OR = 1.49,P = 0.001]
(Kim et al., 2020)	pcat1	2019	China	436	510	Real-time PCR	rs2632159	C/T	TT VS TC + CC[OR = 1.51, *P* = 0.022]
(Wang et al., 2019)	GAS5	2019	China	1078	1175	TaqMan	rs55829688	C/T	TT vs CC + CT [OR = 1.64 *P* = 0.001]
(Lv et al., 2019)	SNHG16	2020	China	361	360	Sanger sequencing	rs7353 rs8038 rs15278	G/A A/G A/G	Negative GA vs AA+GA [OR = 2.9, P < 0.01] GA vs GG + GA [OR = 3.60, P = 0.02]
(Yang et al., 2019)	H19	2020	China	315	441	PCR-RFLP	rs2839698	G/A	Negative
(Van Baak et al., 2018)	MALAT1	2020	China	340	340	Real-time PCR	rs664589	C/G	CC vs CG + GG [OR = 3.82,*P* < 0.001]
(Su et al., 2014)	HOTAIR	2020	China	245	245	Sequencing	rs17720428	T/G	TT vs TT + G [OR = 3.28,*P* < 0.01]
(Kudo-Saito et al., 2021)	PCAT1	2020	China	580	820	TaqMan	rs710886	A/G	AA vs AG + GG[[OR = 1.84,*P* = 0.005]
(Stier and Kiss, 2013)	HOTTIP	2020	Egypt	140	150	Real-time PCR	rs8159168	A/C	CC vs CA + AA[OR = 2.50,*P* = 0.038
(Xu et al., 2021)	MAGI2-AS3	2020	China	1078	1175	Real-time PCR	rs7783388	G/A	GG vs AA+AG[OR = 1.90, *P* < 0.001].
Gargallo-Puyuelo et al. (2021)	LINC00709	2021	Spain	715	715	MassArray (Sequenom) platform	rs11255841	T/A	AA vs AT+TT[OR = 2.04 P < 0.001]
Hennig et al. (2021)	LINC02006	2021	Poland	456	1548	TaqMan	rs10935945	C/T	TTvs.CC[OR = 2.11,P < 0.001]
Zhan et al. (2022)	C5orf66	2021	China	512	513	Real-time PCR	rs4976270 rs639933	C/T G/T	CC vs CT + TT [OR = 1.69, *p* = 0.021] TT vs GT + GG[OR = 1.67, *p* = 0.024]

Data availability

The datasets created during the current study are not publicly accessible due to the possibility of compromising the privacy of individuals. According to the written approval forms accepted by the Ethics Committee of the Mashhad University of Medical Sciences (MUMS), the data will only be available to researchers within the project. The data could be available upon request from the corresponding author, Dr. Mohammad Amin Kerachian (according to the MUMS rules and regulations).

Acknowledgments

We thank Mashhad University of Medical Sciences, Reza Radiotherapy and Oncology Center, Ghaem and Razavi Hospitals for supporting this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humgen.2023.201226.

References

Ali, M.A., Shaker, O.G., Ezzat, E.M., Gaber, S.N., Hassan, E.A., Abdelwahed, M.Y., AbdelHafez, M.N., Khalil, M.A., Abouelseoud, S., 2020. Association between rs1859168/HOTTIP expression level and colorectal cancer and adenomatous polyposis risk in Egyptians. J. Interf. Cytokine Res. 40 (6), 279–291.

Arani, S.H., Kerachian, M.A., 2017. Rising rates of colorectal cancer among younger Iranians: is diet to blame? Curr. Oncol. 24 (2), 131–137.

Azur, M.J., Stuart, E.A., Frangakis, C., Leaf, P.J., 2011. Multiple imputation by chained equations: what is it and how does it work? Int. J. Methods Psychiatr. Res. 20 (1), 40–49.

Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., 2012. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 22 (9), 1790–1797.

Ćwik, G., Solecki, M., Ciechański, A., Prystupa, A., Pedowski, T., Wallner, G., 2012. Perforacje w obrębie jelita grubego. Prz. Gastroenterol. 7, 125–132.

Endo, Y., Marusawa, H., Kou, T., Nakase, H., Fujii, S., Fujimori, T., Kinoshita, K., Honjo, T., Chiba, T., 2008. Activation-induced cytidine deaminase links between inflammation and the development of colitis-associated colorectal cancers. Gastroenterology 135 (3), 889–898 e883.

Fearnhead, N.S., Wilding, J.L., Bodmer, W.F., 2002. Genetics of colorectal cancer: hereditary aspects and overview of colorectal tumorigenesis. Br. Med. Bull. 64 (1), 27–43.

Gargallo-Puyuelo, C.J., Lanas, Á., Carrera-Lasfuentes, P., Ferrández, Á., Quintero, E., Carrillo, M., Alonso-Abreu, I., García-González, M.A., 2021. Familial Colorectal Cancer and Genetic Susceptibility: Colorectal Risk Variants in First-Degree Relatives of Patients With Colorectal Cancer. Clin Transl Gastroenterol 12 (2), e00301. https://doi.org/10.14309/ctg.0000000000000301. PMID: 33534415; PMCID: PMC7861964.

Han, M., Liew, C.T., Zhang, H.W., Chao, S., Zheng, R., Yip, K.T., Song, Z.-Y., Li, H.M., Geng, X.P., Zhu, L.X., 2008. Novel blood-based, five-gene biomarker set for the detection of colorectal cancer. Clin. Cancer Res. 14 (2), 455–460.

Hennig, E.E., Kluska, A., Piątkowska, M., Kulecka, M, Bałabas, A., Zeber-Lubecka, N., Goryca, K., Ambrożkiewicz, F., Karczmarski, J., Olesińśki, T., Zyskowski, Ł., Ostrowski, J., 2021. GWAS Links New Variant in Long Non-Coding RNA *LINC02006* with Colorectal Cancer Susceptibility. Biology (Basel) 10 (6), 465. https://doi.org/10.3390/biology10060465. PMID: 34070617; PMCID: PMC8229782.

Jin, M., Gu, S., Ye, D., Li, Y., Jing, F., Li, Q., Chen, K., 2017. Association between genetic variants in the promoter region of a novel antisense long noncoding RNA RP11-392P7. 6 and colorectal cancer risk. Environ. Mol. Mutagen. 58 (6), 434–442.

Kerachian, M.A., Azghandi, M., 2022. Identification of long non-coding RNA using single nucleotide epimutation analysis: a novel gene discovery approach. Cancer Cell Int. 22 (1), 337.

Kerachian, M.A., Javadmanesh, A., Azghandi, M., Shariatpanahi, A.M., Yassi, M., Davodly, E.S., Talebi, A., Khadangi, F., Soltani, G., Hayatbakhsh, A., 2020. Crosstalk between DNA methylation and gene expression in colorectal cancer, a potential plasma biomarker for tracing this tumor. Sci. Rep. 10 (1), 1–13.

Kim, J.O., Jun, H.H., Kim, E.J., Lee, J.Y., Park, H.S., Ryu, C.S., Kim, S., Oh, D., Kim, J.W., Kim, N.K., 2020. Genetic variants of HOTAIR associated with colorectal cancer susceptibility and mortality. Front. Oncol. 10, 72.

Korniłłowicz-Kowalska, T., Bohacz, J., 2011. Biodegradation of keratin waste: theory and practical aspects. Waste Manag. 31 (8), 1689–1701.

Kudo-Saito, C., Ogiwara, Y., Imazeki, H., Boku, N., Uemura, Y., Zhang, R., Kawano-Nagatsuma, A., Kojima, M., Ochiai, A., 2021. CD11b+ DIP2A+ LAG3+ cells facilitate immune dysfunction in colorectal cancer. Am. J. Cancer Res. 11 (11), 5428.
Li, S., Hua, Y., Jin, J., Wang, H., Du, M., Zhu, L., Chu, H., Zhang, Z., Wang, M., 2016. Association of genetic variants in lncRNA H19 with risk of colorectal cancer in a Chinese population. Oncotarget 7 (18), 25470.
Li, S., Shuch, B.M., Gerstein, M.B., 2017. Whole-genome analysis of papillary kidney cancer finds significant noncoding alterations. PLoS Genet. 13 (3), e1006685.
Li, L.Y., Guan, Y.D., Chen, X.S., Yang, J.M., Cheng, Y., 2020. DNA repair pathways in cancer therapy and resistance. Front. Pharmacol. 11, 629266.
Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., 2013. The genotype-tissue expression (GTEx) project. Nat. Genet. 45 (6), 580–585.
Lv, Z., Xu, Q., Sun, L., Wen, J., Fang, X., Xing, C., Yuan, Y., 2019. Four novel polymorphisms in long non-coding RNA HOTTIP are associated with the risk and prognosis of colorectal cancer. Biosci. Rep. 39 (5).
Meshcheryakova, A., Tamandl, D., Bajna, E., Stift, J., Mittlboeck, M., Svoboda, M., Heiden, D., Stremitzer, S., Jensen-Jarolim, E., Grünberger, T., 2014. B cells and ectopic follicular structures: novel players in anti-tumor programming with prognostic power for patients with metastatic colorectal cancer. PLoS One 9 (6), e99008.
Mohammadpour, A., 2018. Evaluation of a modified salt-out method for DNA extraction from whole blood lymphocytes: a simple and economical method for gene polymorphism. Pharm. Biomed. Res. 4 (2), 28–32.
Peng, W.-X., Koirala, P., Mo, Y.-Y., 2017. LncRNA-mediated regulation of cell signaling in cancer. Oncogene 36 (41), 5661–5667.
Plassais, J., Lagoutte, L., Correard, S., Paradis, M., Guaguère, E., Hédan, B., Pommier, A., Botherel, N., Cadiergues, M.-C., Pilorge, P., 2016. A point mutation in a lincRNA upstream of GDNF is associated to a canine insensitivity to pain: a spontaneous model for human sensory neuropathies. PLoS Genet. 12 (12), e1006482.
Poursheikhani, A., Abbaszadegan, M.R., Kerachian, M.A., 2021. Mechanisms of long non-coding RNA function in colorectal cancer tumorigenesis. Asia Pac. J. Clin. Oncol. 17 (1), 7–23.
Prensner, J.R., Chinnaiyan, A.M., 2011. The emergence of lncRNAs in cancer biology. Cancer Discov. 1 (5), 391–407.
Shen, P., Pichler, M., Chen, M., Calin, G.A., Ling, H., 2017. To Wnt or lose: the missing non-coding linc in colorectal cancer. Int. J. Mol. Sci. 18 (9), 2003.
Siegel, R.L., Miller, K.D., Jemal, A., 2020. Cancer statistics, 2020. CA Cancer J. Clin. 70 (1), 7–30.
Siegel, R.L., Miller, K.D., Fuchs, H.E., Jemal, A., 2021. Cancer statistics, 2021. CA Cancer J. Clin. 71 (1), 7–33.
Stier, I., Kiss, A., 2013. Cytosine-to-uracil deamination by SssI DNA methyltransferase. PLoS One 8 (10), e79003.
Su, X., Zhu, G., Ding, X., Lee, S.Y., Dou, Y., Zhu, B., Wu, W., Li, H., 2014. Molecular basis underlying histone H3 lysine-arginine methylation pattern readout by spin/Ssty repeats of Spindlin1. Genes Dev. 28 (6), 622–636.
Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva Nadezhda, T., Pyysalo, S., et al., 2023. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 51 (D1), D638–D646.
Talebi, A., Thiery, J.P., Kerachian, M.A., 2021. Fusion transcript discovery using RNA sequencing in formalin-fixed paraffin-embedded specimen. Crit. Rev. Oncol. Hematol. 103303.
Tenesa, A., Dunlop, M.G., 2009. New insights into the aetiology of colorectal cancer from genome-wide association studies. Nat. Rev. Genet. 10 (6), 353–358.
Van Baak, T.E., Coarfa, C., Dugue, P.A., Fiorito, G., Laritsky, E., Baker, M.S., Kessler, N.J., Dong, J., Duryea, J.D., Silver, M.J., et al., 2018. Epigenetic supersimilarity of monozygotic twin pairs. Genome Biol. 19 (1), 2.
Wan, Y., Qu, K., Zhang, Q.C., Flynn, R.A., Manor, O., Ouyang, Z., Zhang, J., Spitale, R.C., Snyder, M.P., Segal, E., 2014. Landscape and variation of RNA secondary structure across the human transcriptome. Nature 505 (7485), 706–709.
Wang, Y., Wu, S., Yang, X., Li, X., Chen, R., 2019. Association between polymorphism in the promoter region of lncRNA GAS5 and the risk of colorectal cancer. Biosci. Rep. 39 (4), BSR20190091.
Ward, L.D., Kellis, M., 2016. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Res. 44 (D1), D877–D881.
Wu, S., Sun, H., Wang, Y., Yang, X., Meng, Q., Yang, H., Zhu, H., Tang, W., Li, X., Aschner, M., 2019. MALAT1 rs664589 polymorphism inhibits binding to miR-194-5p, contributing to colorectal cancer risk, growth, and metastasis. Cancer Res. 79 (20), 5432–5441.
Xu, X.Y., Choi, H.S., Park, S.Y., Kim, J.K., Seo, K.H., Kim, H., Kim, Y.J., 2021. *Hibiscus syriacus* L. cultivated in callus culture exerts cytotoxicity in colorectal cancer via Notch signaling-mediated cholesterol biosynthesis suppression. Phytomedicine 95, 153870.
Yang, M.-l., Huang, Z., Wu L-n, Wu R., H-x, Ding, Wang, B.-G., 2019. lncRNA-PCAT1 rs2632159 polymorphism could be a biomarker for colorectal cancer susceptibility. Biosci. Rep. 39(7).
Zhan, X., Bai, F., Lin, S., Feng, T., Tang, X., Ding, Y., Jin, T., 2022. C5orf66 rs4976270/rs639933 Are Associated with Colorectal Cancer Risk in Southern Chinese Han Population: A Case-Control Study. Digestion 103 (2), 103–115. https://doi.org/10.1159/000518519.
Zhang, D.-D., Wang, W.-T., Xiong, J., Xie, X.-M., Cui, S.-S., Zhao, Z.-G., Li, M.J., Zhang, Z.-Q., Hao, D.-L., Zhao, X., 2017. Long noncoding RNA LINC00305 promotes inflammation by activating the AHRR-NF-κB pathway in human monocytes. Sci. Rep. 7 (1), 1–12.
Zhou, L., Zhang, Y., Jin, J., Gu, X., 2020. Correlation between lncRNA SNHG16 gene polymorphism and its interaction with environmental factors and susceptibility to colorectal cancer. Medicine 99 (48).