Received: 2 January 2019 | Revised: 9 May 2019 | Accepted: 16 May 2019

DOI: 10.1002/mgg3.778

ORIGINAL ARTICLE

Molecular Genetics & Genomic Medicine Open Access WILEY

Analysis of *MIR155HG* variants and colorectal cancer susceptibility in Han Chinese population

Huangfu Wu[*,†] | **Guisheng He**[*,†] | **Hua Han** | **Wei Xiong** | **Tao Song** | **Huamin Chen** | **Xiuxiu Chen** | **Xiaoming Wu** | **Guangyue Huang** | **Yazhen Zhang** | **Chuanwei Sun** | **Chaoyang Zhao** | **Yunjing Chen**

Surgical Oncology, The Second Affiliated Hospital of Hainan Medical College, Haikou, Hainan Province, China

Correspondence
Huangfu Wu, #48 Baishuitang Road, Longhua District 570100, Haikou, Hainan Province, China.
Email: wuhuangfu1103@163.com

Funding information
The Hainan Provincial Key Research and Development Program Project Fund (No. ZDYF2017087).

Abstract

Background: *MIR155HG* plays an important role in malignant tumors, but it is rarely reported in the occurrence and development of colorectal cancer (CRC). This study investigated the effects of M*IR155HG* polymorphisms on CRC susceptibility from the perspective of molecular genetics.

Methods: Eight SNPs in *MIR155HG* were selected and genotyped among 514 CRC cases and 510 healthy controls using the Agena MassARRAY platform. The associations between these SNPs and the CRC risk were evaluated under genetic models using conditional logistic regression analysis. The HaploReg v4.1 database was used for SNPs functional prediction.

Results: The allele "C" of rs12482371 ($p = 0.047$), allele "C" of rs1893650 ($p = 0.025$), and the allele "A" of rs928883 ($p = 0.037$) in *MIR155HG* were significantly associated with CRC risk. Genetic model analysis revealed that rs12482371 and rs1893650 increased CRC risk; whereas rs928883 was associated with reduced CRC risk. Stratification analysis showed that rs9383938 was a protective factor in CRC patients under 60 years old. Rs12482371 and rs1893650 were associated with the CRC risk in females. Rs11911469 and rs34904192 may affect the clinical stage and lymph node metastasis. Moreover, the haplotypes CTT and GTC of LD block rs4143370|rs77218221|rs12482371, and the haplotypes CATGA and CACGG of LD block rs77699734|rs11911469|rs1893650|rs34904192|rs928883 were significantly associated with CRC risk.

Conclusion: This study revealed that *MIR155HG* SNPs were associated with CRC susceptibility and could be predictive biomarkers for CRC risk.

KEYWORDS
Chinese Han population, colorectal cancer risk, *MIR155HG*, single-nucleotide polymorphisms

* Joint first authors.

[†]Co-first authors.

1 | INTRODUCTION

Colorectal cancer (CRC) is one of the most common gastrointestinal cancers worldwide with higher morbidity and mortality caused by biological interactions of various signaling pathways (Li & Gu, 2005). During the past few decades, with the change in modern lifestyle and dietary habit, the incidence and mortality of CRC have been increasing rapidly in China. Currently, CRC is the fifth commonest malignancy in China (Chen, Zheng, Zeng, & Zhang, 2015). The exact mechanisms underlying the development and progression of CRC remain generally unknown. According to a genetic model for colorectal tumorigenesis, genetic mutation is a critical accelerant contributing to malignant progression of CRC (Fearon & Vogelstein, 1990). Cancer-analyses of cohorts of twins have shown that inherited genetic factors accounts for about 35% of the variance in CRC susceptibility (Lichtenstein et al., 2000). These findings suggest that genetic factors significantly influence CRC susceptibility.

Long noncoding RNA (lncRNAs) refers to noncoding RNAs consisting of more than 200 nucleotides in length with important biological functions. The abnormal expression of lncRNAs plays an important role in different physiological and pathological processes, especially to the occurrence and development of tumors (Beckedorff, Amaral, Deocesano-Pereira, & Verjovski-Almeida, 2013; Prensner & Chinnaiyan, 2011; Rossi & Antonangeli, 2014). The MIR155 host gene (*MIR155HG*) is an lncRNA that has identified to be an important regulator of many physiological and pathological processes involved in hematopoiesis, inflammation, immunity, and tumorigenesis (Chang et al., 2011; Hu, Fong, Largman, & Shen, 2010; Kohlhaas et al., 2009; Van den Berg et al., 2003; Vargova et al., 2011). Studies have confirmed that *MIR155HG* overexpressed in acute myeloid leukemia, Hodgkin's lymphoma and chronic lymphoblastic leukemia (Elton, Selemon, Elton, & Parinandi, 2013). In addition, experimental results demonstrated that downregulation of *MIR155HG* inhibited the growth of gliomas in vitro and in vivo (Wu et al., 2017). Previous studies have shown lncRNA *MIR155HG* hold potential to be used as prognostic biomarkers in CRC patients (Thiele et al., 2018).

Single-nucleotide polymorphisms (SNPs) are the most common types of genetic variations that affect disease risk by altering the expression of related genes. As far as we know, there were three studies have reported the association of *MIR155HG*/miR-155 SNPs with multiple sclerosis (Paraboschi et al., 2011), atopic eczema (Sääf et al., 2013), and epilepsy (Tao et al., 2015). Only three SNPs (rs2829803, rs2282471, and rs2829806) in multiple sclerosis and two SNPs (rs969885 and rs987195) in epilepsy were identified as the genetic susceptibility factors. However, the exact relation between *MIR155HG* SNPs and CRC risk is still undetermined.

In the current study, we performed a case–control study among the Chinese Han population aimed to explore the association between *MIR155HG* SNPs and CRC risk from the perspective of molecular genetics. Our results will help to understand the pathological factors affecting the development of CRC, and thus provide a theoretical basis for the prevention and early detection of CRC.

2 | MATERIALS AND METHODS

2.1 | Subjects

In the present ongoing case–control study, a total of 514 CRC patients and 510 age- and sex-matched healthy controls were consecutively recruited at the Shaanxi Provincial Cancer Hospital. All the individuals were genetically unrelated ethnic northwest Han Chinese and all patients were diagnosed with colorectal cancer by pathological analysis. We excluded the patients who had undergone chemotherapy or irradiation from the current study. Subjects had no history of hereditary or malignant diseases. The clinical features of patients were collected from the patients' medical records provided, including age, gender, clinical stage, and lymph node metastasis (LNM). A standardized epidemiological questionnaire was used to collect subject's demographic and personal information and all individuals were informed of the purpose and experimental procedures of the study.

2.2 | DNA extraction and genotyping

Genomic DNA was extracted from blood samples using the commercially available GoldMag-Mini Purification Kit (GoldMag Co. Ltd, Xi'an, China) according to the manufacturer's instructions and the concentration and quality of the DNA were measured on a Nanodrop 2000 spectrophotometer (Thermo Scientific, USA). Eight SNPs with minor allele frequencies (MAFs) >5%, referenced from the 1,000 Genomes database (http://www.internationalgenome.org/), were selected for further analysis. Agena MassARRAY Assay Design Software (version 3.0, Agena Bioscience, San Diego, CA) was used to design multiplexed SNP MassEXTEND assay (Gabriel, Ziaugra, & Tabbaa, 2009). And Agena MassARRAY RS1000 was used to detect SNP genotyping. Data management and analysis were conducted using Agena Software (version 4.0, Agena, San Diego, CA) (Duan et al., 2014; Geng et al., 2015).

2.3 | Bioinformatics analysis

HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) database can be used to study the annotation of noncoding genomes on haplotype blocks, including candidate regulatory SNPS of disease-related sites. We used HaploReg v4.1 to predict the potential functions of candidate SNPs.

2.4 | Statistical analyses

The Hardy–Weinberg Equilibrium (HWE) was tested by using chi-square test in the healthy control group (Table S1). The odds ratios (ORs), 95% confidence intervals (CIs) and *p* values were calculated by the conditional logistic regression model with the PLINK software version 1.07 (http://www.cog-genomics.org/plink2/) to assess the risk for CRC altered by a certain allele and genotype. Four models (genotype, dominant, recessive, and additive) were used to assess the association between each genotype and the risk of CRC (Zhang et al., 2014). Moreover, the Haploview software package (version 4.2) and SHEsis software platform were used to analyze the pairwise linkage disequilibrium (LD), haplotype construction, and genetic association of polymorphism loci. In the case–control study, the SPSS 20.0 software was used to calculate the difference in basic information between patients and healthy control group, including age, gender, etc. The adjusted *p* value of less than 0.05 was deemed as statistically significant.

3 | RESULTS

3.1 | Characteristic of the study participates

We recruited 514 cases (285 males and 229 females; mean age, 60.27 ± 11.81) and 510 controls (286 males and 224 females; mean age, 60.13 ± 10.61) for this study. The characteristics of cases and controls were showed in Table 1. Statistical analysis results showed that there was no significant difference in age and gender distribution between cases and controls ($p > 0.05$).

TABLE 1 Characteristics of the study population

	Cases	Controls	*p*
Total	514	510	
Age			0.092[a]
Mean ± *SD*	60.27 ± 11.81	60.13 ± 10.61	
Sex			0.850[b]
Male	285 (55.4%)	286 (56.1%)	
Female	229 (44.6%)	224 (43.9%)	
Absence	38 (7.4%)		
Stage			
I-II	146 (28.4%)		
III-IV	212 (41.2%)		
Absence	156 (30.4%)		
LNM			
Negative	170 (33.1%)		
Positive	188 (36.6%)		
Absence	156 (30.3%)		

Abbreviations: *SD*: standard deviation; BMI: body mass index; LNM: lymph node metastasis.
[a]*p* values were calculated from independent sample *t* test.
[‡]*p* values were calculated from two-sided χ^2 test.

3.2 | Basic information and potential functions of the selected SNPs

Table 2 displays the basic information of eight SNPs in *MIR155HG,* including gene, chromosome, position, role, alleles, minor allele frequency (MAF), and functional effects. All SNPs were consistent with Hardy–Weinberg equilibrium (HWE) in the controls ($p > 0.05$). Additionally, we used HaploRegv4.1 to annotate the functional elements containing these selected SNPs. The results showed that these SNPs were involved in the regulation related to DNAase, marked promoter histone, changed motifs, GRASP QTL hits, selected eQTL hits enhancer histones, bound proteins, suggesting they might influence the related genes expression in this way.

3.3 | SNPs and the risk of CRC

In Table 2, we identified three significant SNPs variants associated with the CRC risk by the χ^2 test. They were the allele "C" of rs12482371 (OR = 1.21, 95% CI = 1.00–1.47, $p = 0.047$), the allele "C" of rs1893650 (OR = 1.29, 95% CI = 1.03–1.62, $p = 0.025$), and the allele "A" of rs928883 (OR = 0.83, 95% CI = 0.70–0.99, $p = 0.037$). Results of the genetic model analysis are showed in Table 3. The result indicated that rs12482371 moderately increased the risk of CRC in genotype model (OR = 1.35, 95% CI = 1.04–1.75, $p = 0.024$) and dominant model (OR = 1.33, 95% CI = 1.04–1.70, $p = 0.022$), respectively. We also determined that rs1893650 increased risk of CRC (TC-CC vs. TT, OR = 1.32, 95% CI: 1.02–1.71, $p = 0.036$), conversely, that rs928883 was associated with reduced susceptibility of CRC (AA vs. GG, OR = 0.70, 95% CI: 0.49–0.99, $p = 0.046$). In additive model, the rs1893650 and rs928883 also related with CRC risk ($p < 0.05$).

Stratified analysis revealed the relationships of *MIR155HG* SNPs with CRC risk and the results were exhibited in Table 4. The results showed that the allele "A" of rs9383938 was associated with a reduced risk of CRC in individuals under 60 years old (OR = 0.77, 95% CI = 0.59–0.99, $p = 0.047$), but the rs12482371 increased CRC risk in dominant model only among individuals over 60 years old (OR = 0.77, 95% CI = 0.59–0.99, $p = 0.047$). Rs12482371 and rs1893650 were associated with CRC risk in female under allele model, dominant model and additive model ($p < 0.05$). In addition, the rs1893650 played a risk role only in colon cancer under allele model, dominant model and additive model ($p < 0.05$). The genotypes distribution of rs11911469 and rs34904192 were significantly different between grades I-II and III-IV patients, as well as between lymph node metastasis and without metastasis patients ($p < 0.05$).

TABLE 2 Basic information of candidate SNPs in the study

SNP ID	Genes	Chr.	Position	Alleles A/B	MAF		*p*-HWE[a]	*p*[b]	OR (95% CI)	HaploReg
					Case	Control				
rs4143370	MIR155HG	chr21	25,564,661	C/G	0.175	0.165	0.053	0.518	1.08 (0.86–1.36)	Promoter histone marks, Enhancer histone marks, DNAse, Proteins bound, Motifs changed, Selected eQTL hits
rs77218221	MIR155HG	chr21	25,565,063	C/T	0.051	0.050	1.000	0.952	1.01 (0.68–1.51)	Promoter histone marks, Enhancer histone marks, Motifs changed
rs12482371	MIR155HG	chr21	25,566,041	C/T	0.318	0.278	0.151	**0.047**	1.21 (1.00–1.47)	Promoter histone marks, Enhancer histone marks, DNAse, Proteins bound, Motifs changed, Selected eQTL hits
rs77699734	MIR155HG	chr21	25,566,995	C/G	0.083	0.091	1.000	0.823	0.97 (0.71–1.31)	Promoter histone marks, Enhancer histone marks, DNAse, Proteins bound
rs11911469	MIR155HG	chr21	25,567,971	A/C	0.122	0.116	0.522	0.691	1.06 (0.81–1.38)	Promoter histone marks, Enhancer histone marks, DNAse, Proteins bound, Motifs changed
rs1893650	MIR155HG	chr21	25,568,503	C/T	0.205	0.167	0.751	**0.025**	1.29 (1.03–1.62)	Promoter histone marks, Enhancer histone marks, DNAse, Proteins bound, Motifs changed, GRASP QTL hits, Selected eQTL hits
rs34904192	MIR155HG	chr21	25,569,623	A/G	0.265	0.255	0.817	0.636	1.05 (0.86–1.28)	Promoter histone marks, Enhancer histone marks, DNAse, Proteins bound, Motifs changed, Selected eQTL hits
rs928883	MIR155HG	chr21	25,571,713	A/G	0.414	0.460	1.000	**0.037**	0.83 (0.70–0.99)	Enhancer histone marks, DNAse, Proteins bound, Motifs changed, Selected eQTL hits

Note: Bold values indicate statistical significance ($p < 0.05$).

Abbreviations: SNP: single-nucleotide polymorphism; MAF: minor allele frequency; HWE: Hardy–Weinberg equilibrium; OR: odds ratio; 95% CI: 95% confidence interval.

[a]HWE *p* values were calculated from Fisher's exact test.

[b]*p* values were calculated from two-sided χ^2 test.

TABLE 3 Analysis of association between rs12482371, rs1893650, and rs928883 polymorphism and risk of CRC

					Crude		Adjusted[a]	
SNP ID	**Model**	**Genotype**	**Case (*N*)**	**Control (*N*)**	**OR (95% CI)**	***P*[b]**	**OR (95% CI)**	***P*[b]**
rs12482371	Genotype	CC	238	272	1			
		CT	225	191	1.34 (1.04–1.75)	**0.025**	1.35 (1.04–1.75)	**0.024**
		TT	51	46	1.27 (0.82–1.96)	0.286	1.27 (0.82–1.96)	0.288
	Dominant	CC	238	272	1		1	
		CT-TT	276	237	1.33 (1.04–1.70)	**0.023**	1.33 (1.04–1.70)	**0.022**
	Recessive	CC-CT	463	463	1		1	
		TT	51	46	1.11 (0.73–0.69)	0.629	1.11 (0.73–0.68)	0.633
	Additive	—	—	—	1.21 (0.99–1.45)	0.051	1.21 (0.99–1.46)	0.051
rs1893650	Genotype	TT	326	355	1			
		TC	165	140	1.28 (0.98–1.68)	0.071	1.28 (0.98–1.68)	0.071
		CC	23	15	1.67 (0.86–3.26)	0.132	1.67 (0.86–3.25)	0.134
	Dominant	TT	326	355	1		1	
		TC-CC	188	155	1.32 (1.02–1.71)	**0.036**	1.32 (1.02–1.71)	**0.036**
	Recessive	TT-TC	491	495	1		1	
		CC	23	15	1.55 (0.80–3.00)	0.197	1.54 (0.80–2.99)	0.199
	Additive	—	—	—	1.29 (1.03–1.61)	**0.026**	1.29 (1.03–1.61)	**0.027**
rs928883	Genotype	GG	175	148	1			
		GA	242	254	0.81 (0.61–1.07)	0.132	0.81 (0.61–1.07)	0.132
		AA	88	107	0.70 (0.49–0.99)	**0.046**	0.70 (0.49–0.99)	**0.046**
	Dominant	GG	175	148	1		1	
		GA-AA	330	361	0.77 (0.59–1.01)	0.057	0.77 (0.59–1.01)	0.057
	Recessive	GG-GA	417	402	1		1	
		AA	88	107	0.79 (0.58–1.09)	0.147	0.79 (0.58–1.09)	0.147
	Additive	—	—	—	0.83 (0.70–0.99)	**0.038**	0.83 (0.70–0.99)	**0.038**

Note: Bold values indicate statistical significance ($p < 0.05$).

[a]Adjusted for age and sex in a conditional logistic regression model.

[b]*p* values were calculated from wald test.

3.4 | Haplotype analysis

Two blocks were detected in *MIR155HG* SNPs by haplotype analyses (Figure 1). Block 1 contained three SNPs (rs4143370|rs77218221|rs12482371). Block 2 contained five SNPs (rs77699734|rs11911469|rs1893650|rs34904192|rs928883). The results of the relationships between the haplotypes and the CRC risk were listed in Table 5. In block 1, haplotype "CTT" and haplotype "GTC" were more prevalent in case group than control group and were significantly associated with a reduced the risk of CRC ("CTT": OR = 0.83, 95% CI = 0.69–0.99, $p = 0.047$; "GTC": OR = 0.82, 95% CI = 0.69–0.97, $p = 0.021$). In block 2, haplotype "CATGA" and haplotype "CACGG" were also associated with a reduced the risk of CRC. ("CATGA": OR = 0.81, 95% CI = 0.68–0.97, $p = 0.021$; "CACGG": OR = 0.78, 95% CI = 0.62–0.97, $p = 0.028$).

4 | DISCUSSION

To date, several studies have investigated the association between genetic polymorphisms and risk of CRC, and identified a variety of susceptibility genes and SNPs for CRC in different populations. However, studies based on variants in the *MIR155HG* gene are infrequent in the occurrence and development of colorectal cancer. In this case–control study, we first time evaluated the relationships between eight SNPs in lncRNA *MIR155HG* and the CRC risk in the Chinese population. Five novel SNPs (rs12482371, rs1893650, rs92888, rs11911469, and rs34904192) in *MIR155HG* were found significantly influence the CRC susceptibility.

The lncRNA *MIR155HG* was found to be a marker of early stage cancer development (Thiele et al., 2018). *MIR155HG* gene is activated by MYB transcription factor and thus upregulated, which in turn leaded to downregulation of many

TABLE 4 Stratification analysis of the association of *MIR155HG* polymorphisms with CRC under genetic models

		P[a], OR (95% CI)					
	SNP ID	**Allele**	**Homozygote**	**Heterozygote**	**Dominant**	**Recessive**	**Additive**
Age							
≤60	rs12482371	0.344, 1.14 (0.87–1.51)	0.606, 1.23 (0.63–2.40)	0.320, 1.21 (0.83–1.77)	0.292, 1.21 (0.85–1.74)	0.724, 1.12 (0.59–2.14)	0.329, 1.15 (0.87–1.53)
	rs928883	**0.047**, 0.77 (0.59–0.99)	**0.037**, 0.55 (0.31– 0.96)	0.242, 0.78 (0.51–1.18)	0.102, 0.72 (0.49–1.07)	0.080, 0.64 (0.39–1.05)	**0.037**, 0.75 (0.57–0.98)
>60	rs12482371	0.068, 1.27 (0.98–1.65)	0.346, 1.32 (0.74–2.35)	**0.033**,1.48 (1.03–2.12)	**0.033**,1.45 (1.03–2.03)	0.702, 1.11 (0.64–1.94)	0.077, 1.26 (0.98–1.62)
	rs928883	0.327, 0.89 (0.70–1.13)	0.391, 0.81 (0.51–1.30)	0.272, 0.80 (0.54–1.19)	0.247, 0.81 (0.56–1.16)	0.709, 0.92 (0.61–1.39)	0.340, 0.89 (0.71–1.13)
Sex							
Male	rs12482371	0.463, 1.10 (0.85–1.42)	0.977, 0.99 (0.56–1.77)	0.154, 1.29 (0.91–1.82)	0.226, 1.23 (0.88–1.70)	0.663, 0.88 (0.51–1.54)	0.467, 1.10 (0.85–1.41)
	rs1893650	0.447, 1.12 (0.83–1.51)	0.472, 1.39 (0.57–3.36)	0.651, 1.09 (0.76–1.56)	0.536, 1.12 (0.79–1.58)	0.505, 1.35 (0.56–3.26)	0.451, 1.12 (0.83–1.51)
	rs928883	0.214, 0.86 (0.68–1.09)	0.208, 0.74 (0.47–1.18)	0.687, 0.93 (0.63–1.35)	0.416, 0.86 (0.61–1.23)	0.225, 0.78 (0.52–1.17)	0.226, 0.87 (0.69–1.09)
Female	rs12482371	**0.030**, 1.38 (1.03–1.84)	0.103, 1.74 (0.89–3.40)	0.075, 1.43 (0.96–2.11)	**0.038**, 1.48 (1.02–2.15)	0.224, 1.49 (0.78–2.85)	**0.034**, 1.36 (1.02–1.81)
	rs1893650	**0.012**, 1.54 (1.10–2.17)	0.152, 2.11 (0.76–5.85)	**0.028**, 1.58 (1.05–2.39)	**0.015**, 1.63 (1.10–2.42)	0.243, 1.83 (0.66–5.03)	**0.014**, 1.53 (1.09–2.15)
	rs928883	0.083, 0.79 (0.61–1.03)	0.112, 0.64 (0.36–1.11)	0.069, 0.67 (0.44–1.03)	**0.047**, 0.67 (0.44–0.99)	0.413, 0.81 (0.50–1.33)	0.072, 0.78 (0.59–1.02)
Tumor types							
Colon cancer	rs1893650	**0.031**, 1.36 (1.03–1.81)	0.189, 1.74 (0.76–3.97)	0.067, 1.38 (0.97–1.95)	**0.041**,1.42 (1.02–1.98)	0.280, 1.57 (0.69–3.55)	**0.035**, 1.36 (1.02–1.80)
Rectal cancer	rs1893650	0.393, 1.13 (0.85–1.50)	0.949, 1.03 (0.41–2.58)	0.295, 1.20 (0.86–1.67)	0.317, 1.18 (0.85–1.63)	0.960, 0.98 (0.39–2.43)	0.392, 1.13 (0.85–1.50)
Stage							
III-IV versus I-II	rs11911469	**0.001**, 0.47 (0.30–0.75)	0.329, 0.47 (0.10–2.15)	**0.002**, 0.43 (0.25–0.72)	**0.001**, 0.43 (0.26–0.72)	0.455, 0.56 (0.12–2.56)	**0.002**, 0.49 (0.31–0.77)
	rs34904192	0.225, 1.24 (0.88–1.75)	0.731, 0.88 (0.41–1.87)	**0.011**, 1.85 (1.15–2.97)	**0.043**, 1.57 (1.01–2.42)	0.357, 0.71 (0.34–1.48)	0.253, 1.21 (0.87–1.69)
LNM							
(+) versus (-)	rs11911469	0.064, 0.66 (0.43–1.03)	0.842, 0.87 (0.21–3.57)	**0.039**, 0.58 (0.35–0.97)	**0.046**, 0.61 (0.37–0.99)	0.973, 0.98 (0.24–4.01)	0.079, 0.68 (0.44–1.05)
	rs34904192	0.184, 1.26 (0.90–1.76)	0.594, 0.82 (0.39–1.72)	**0.002**, 2.07 (1.30–3.30)	**0.017**, 1.68 (1.10–2.57)	0.216, 0.63 (0.31–1.31)	0.212, 1.23 (0.89–1.68)

Note: Bold values indicate statistical significance ($p < 0.05$).

Abbreviations: BMI: body mass index; LNM: lymph node metastasis.

[a]*p* values were calculated from wald test.

tumor suppressor genes (Nielsen et al., 2007). Moreover, the *MIR155HG* transcript is processed into microRNA-155 (miR-155). MiR-155 is a multifunctional miRNA that plays varied roles in immune, inflammatory, and cardiovascular diseases (Elton et al., 2013). MiR-155 has been identified as a negative regulator of tumor protein 53 and DNA mismatch repair genes (Valeri et al., 2010). Especially, miR-155-5p was found to be overexpressed in solid tumors of diverse origin, including colon cancer (Zhang et al., 2013). Recent studies have shown that miR-155 may work as a tumor suppressor (Kim et al., 2016), and Liu et al. found that overexpressed miR-155 lead to apoptosis and suppresses cell proliferation in CRC (Liu, Chen, Xiang, & Gu, 2018). Therefore, we hypothesized that the *MIR155HG* may play a critical role in promoting CRC progression, and its mechanism may be achieved by producing miR-155. Further functional researches are still required to explore the underlying mechanisms of *MIR155HG*.

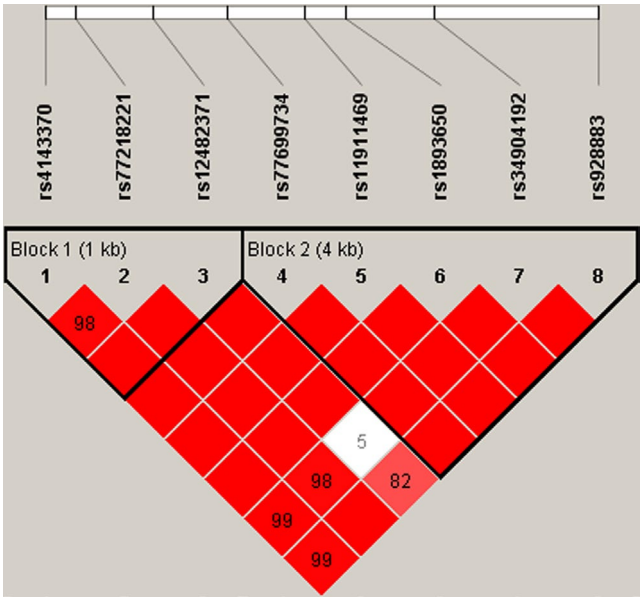

FIGURE 1 Haplotype block map for the SNPs in *MIR155HG* gene. The numbers inside the diamonds indicate the D′ for pairwise analyses. Block 1 indicates that there is a strong linkage disequilibrium between rs4143370, rs77218221, and rs12482371. Block 2 indicates that there is a strong linkage disequilibrium between rs77699734, rs11911469, rs1893650, rs34904192, and rs928883

For many years, intron sequences have been considered essentially nonfunctional. However, subsequent study showed that intron-containing genes presented higher levels of transcription when compared to intron-less genes in mammalian cells, suggesting that introns may be enhancers of transcription (Vaz-Drago, Custódio, & Carmo-Fonseca, 2017). Rs12482371, rs1893650, rs92888, rs11911469, and rs34904192 were all located in the intron region of *MIR155HG*, therefore combining the predicted functions of the database, we hypothesized that these SNPs may enhance the translation of *MIR155HG* gene thereby affect the risk of CRC. The results also provide new evidence for the early detection of CRC.

Some limitations need to be considered in our study. First, this study was only performed in Chinese Han populations. Due to differences in genetic background, the role of these SNPs in other ethnic groups remains to be revealed. Second, CRC is a complex disease affected by a variety of genetic and environmental factors. The limitations of the data make it difficult to explore the interaction between genetic polymorphism and environmental factors.

TABLE 5 Haplotype frequencies of *MIR155HG* SNPs and the association with CRC

Block	Haplotype	Haplotype Frequency		OR(95% CI)	P^a
		Case	Control		
Block 1	rs4143370|rs77218221|rs12482371				
	CTT	0.681	0.722	0.83 (0.69–0.99)	**0.047**
	CCC	0.950	0.950	1.01 (0.68–1.49)	0.972
	CTC	0.126	0.115	1.10 (0.85–1.41)	0.471
	GTC	0.505	0.557	0.82 (0.69–0.97)	**0.021**
Block 2	rs77699734|rs11911469|rs1893650|rs34904192 |rs928883				
	CATGA	0.409	0.459	0.81 (0.68–0.97)	**0.021**
	GATAG	0.912	0.909	1.04 (0.76–1.40)	0.826
	CATAG	0.176	0.164	1.08 (0.87–1.35)	0.483
	CACGG	0.795	0.833	0.78 (0.62–0.97)	**0.028**
	CCTGG	0.878	0.882	0.97 (0.74–1.26)	0.806

Note: Bold values indicate statistical significance ($p < 0.05$).
OR: odds ratio; 95% CIs: 95% confidence intervals.
[a]p values were calculated by conditional logistic regression with adjustment for age and gender.

In conclusion, our study of SNPs in *MIR155HG* confirmed the relationships between genetic polymorphisms and CRC in the Chinese Han population. We hope this research provides a new direction for the detection and diagnosis of CRC and permits early prevention of this disease. To better understand the relationship, further analyses should increase sample size and clinical data, as well as functional experiments, and require repeated studies in different ethnic populations.

ACKNOWLEDGMENTS

We thank the physicians and medical staff of the Shaanxi Provincial Cancer Hospital for their contributions and support of this work. We are grateful to all participants for providing blood samples.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All procedures performed in studies involving human participants were in accordance with the ethical standards of Shaanxi Provincial Cancer Hospital committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

INFORMED CONSENT

All participants were informed in writing and verbally of the procedures and purpose of this study, and signed written informed consent forms.

DATA AVAILABILITY STATEMENT

The datasets used and analyzed in the current study are available from the corresponding author on reasonable request.

ORCID

Huangfu Wu https://orcid.org/0000-0001-9503-9773

REFERENCES

Beckedorff, F. C., Amaral, M. S., Deocesano-Pereira, C., & Verjovski-Almeida, S. (2013). *Long non-coding RNAs and their implications in cancer epigenetics*. Portland: Press Limited.

Chang, S., Wang, R.-H., Akagi, K., Kim, K.-A., Martin, B. K., Cavallone, L., … Sharan, S. K. (2011). Tumor suppressor BRCA1 epigenetically controls oncogenic microRNA-155. *Nature Medicine*, *17*(10), 1275. https://doi.org/10.1038/nm.2459

Chen, W., Zheng, R., Zeng, H., & Zhang, S. (2015). The updated incidences and mortalities of major cancers in China, 2011. *Chinese Journal of Cancer*, *34*(3), 53. https://doi.org/10.1186/s40880-015-0042-6

Duan, X., Li, X., Lou, H., Geng, T., Jin, T., Liang, P., … Chen, C. (2014). Genetic association of PLCE1, C11orf92-C11orf93, and NOC3L with colorectal cancer risk in the Han population. *Tumor Biology*, *35*(3), 1813–1817. https://doi.org/10.1007/s13277-013-1242-9

Elton, T. S., Selemon, H., Elton, S. M., & Parinandi, N. L. (2013). Regulation of the MIR155 host gene in physiological and pathological processes. *Gene*, *532*(1), 1–12. https://doi.org/10.1016/j.gene.2012.12.009

Fearon, E. R., & Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*, *61*(5), 759–767. https://doi.org/10.1016/0092-8674(90)90186-I

Gabriel, S., Ziaugra, L., & Tabbaa, D. (2009). SNP Genotyping Using the Sequenom MassARRAY iPLEX Platform. *Current Protocols in Human Genetics*, *60*(1), 2.12. 11–12.12. 18.

Geng, T.-T., Xun, X.-J., Li, S., Feng, T., Wang, L.-P., Jin, T.-B., & Hou, P. (2015). Association of colorectal cancer susceptibility variants with esophageal cancer in a Chinese population. *World Journal of Gastroenterology: WJG*, *21*(22), 6898. https://doi.org/10.3748/wjg.v21.i22.6898

Hu, Y.-L., Fong, S., Largman, C., & Shen, W.-F. (2010). HOXA9 regulates miR-155 in hematopoietic cells. *Nucleic Acids Research*, *38*(16), 5472–5478. https://doi.org/10.1093/nar/gkq337

Kim, S., Song, J. H., Kim, S., Qu, P., Martin, B. K., Sehareen, W. S., … Chang, S. (2016). Loss of oncogenic miR-155 in tumor cells promotes tumor growth by enhancing C/EBP-β-mediated MDSC infiltration. *Oncotarget*, *7*(10), 11094.

Kohlhaas, S., Garden, O. A., Scudamore, C., Turner, M., Okkenhaug, K., & Vigorito, E. (2009). Cutting edge: The Foxp3 target miR-155 contributes to the development of regulatory T cells. *The Journal of Immunology*, *182*(5), 2578–2582. https://doi.org/10.4049/jimmunol.0803162

Li, M., & Gu, J. (2005). Changing patterns of colorectal cancer in China over a period of 20 years. *World Journal of Gastroenterology*, *11*(30), 4685. https://doi.org/10.3748/wjg.v11.i30.4685

Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iliadou, A., Kaprio, J., Koskenvuo, M., … Hemminki, K. (2000). Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England Journal of Medicine*, *343*(2), 78–85. https://doi.org/10.1056/NEJM200007133430201

Liu, J., Chen, Z., Xiang, J., & Gu, X. (2018). MicroRNA-155 acts as a tumor suppressor in colorectal cancer by targeting CTHRC1 in vitro. *Oncology Letters*, *15*(4), 5561–5568.

Nielsen, C. B., Shomron, N., Sandberg, R., Hornstein, E., Kitzman, J., & Burge, C. B. (2007). Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA*, *13*(11), 1894–1910. https://doi.org/10.1261/rna.768207

Paraboschi, E. M., Soldà, G., Gemmati, D., Orioli, E., Zeri, G., Benedetti, M. D., … Asselta, R. (2011). Genetic association and altered gene expression of mir-155 in multiple sclerosis patients. *International Journal of Molecular Sciences*, *12*(12), 8695–8712. https://doi.org/10.3390/ijms12128695

Prensner, J. R., & Chinnaiyan, A. M. (2011). The emergence of lncRNAs in cancer biology. *Cancer Discovery*, *1*(5), 391–407. https://doi.org/10.1158/2159-8290.CD-11-0209

Rossi, M. N., & Antonangeli, F. (2014). LncRNAs: New players in apoptosis control. *International Journal of Cell Biology*, *2014*, 1–7. https://doi.org/10.1155/2014/473857

Sääf, A., Kockum, I., Wahlgren, C., Xu, N., Sonkoly, E., Ståhle, M., … Pivarcsi, A. (2013). Are BIC (miR-155) polymorphisms associated with eczema susceptibility? *Acta dermato-venereologica*, *93*(3), 366–367. https://doi.org/10.2340/00015555-1466

Tao, H., Cui, L., Li, Y., Zhou, X., Ma, G., Yao, L., Zhou, H. (2015). Association of tag SNPs and rare CNVs of the MIR155HG/miR-155 gene with epilepsy in the Chinese Han population. *BioMed Research International*, *2015*, 1–8. https://doi.org/10.1155/2015/837213

Thiele, J.-A., Hosek, P., Kralovcova, E., Ostasov, P., Liska, V., Bruha, J., … Pitule, P. (2018). lncRNAs in non-malignant tissue have prognostic value in colorectal cancer. *International Journal of Molecular Sciences*, *19*(9), 2672. https://doi.org/10.3390/ijms19092672

Valeri, N., Gasparini, P., Fabbri, M., Braconi, C., Veronese, A., Lovat, F., Bottoni, A. (2010). Modulation of mismatch repair and genomic stability by miR-155. *Proceedings of the National Academy of Sciences*, *107*(15), 6982–6987. https://doi.org/10.1073/pnas.1002472107

van den Berg, A., Kroesen, B.-J., Kooistra, K., de Jong, D., Briggs, J., Blokzijl, T., … Poppema, S. (2003). High expression of B-cell receptor inducible gene BIC in all subtypes of Hodgkin lymphoma. *Genes, Chromosomes and Cancer*, *37*(1), 20–28. https://doi.org/10.1002/gcc.10186

Vargova, K., Curik, N., Burda, P., Basova, P., Kulvait, V., Pospisil, V., … Stopka, T. (2011). MYB transcriptionally regulates the miR-155 host gene in chronic lymphocytic leukemia. *Blood*, *117*(14), 3816–3825. https://doi.org/10.1182/blood-2010-05-285064

Vaz-Drago, R., Custódio, N., & Carmo-Fonseca, M. (2017). Deep intronic mutations and human disease. *Human Genetics*, *136*(9), 1093–1111. https://doi.org/10.1007/s00439-017-1809-4

Wu, X., Wang, Y., Yu, T., Nie, E. R., Hu, Q. I., Wu, W., … You, Y. (2017). Blocking MIR155HG/miR-155 axis inhibits mesenchymal transition in glioma. *Neuro-oncology*, *19*(9), 1195–1205. https://doi.org/10.1093/neuonc/nox017

Zhang, G.-J., Xiao, H.-X., Tian, H.-P., Liu, Z.-L., Xia, S.-S., & Zhou, T. (2013). Upregulation of microRNA-155 promotes the migration and invasion of colorectal cancer cells through the regulation of claudin-1 expression. *International Journal of Molecular Medicine*, *31*(6), 1375–1380. https://doi.org/10.3892/ijmm.2013.1348

Zhang, T., Li, X., Du, Q., Gong, S., Wu, M., Mao, Z., … Chen, C. (2014). DUSP10 gene polymorphism and risk of colorectal cancer in the Han Chinese population. *European Journal of Cancer Prevention*, *23*(3), 173–176. https://doi.org/10.1097/CEJ.0b013e3283647408

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Wu H, He G, Han H, et al. Analysis of *MIR155HG* variants and colorectal cancer susceptibility in Han Chinese population. *Mol Genet Genomic Med*. 2019;7:e778. https://doi.org/10.1002/mgg3.778