OncoTargets and Therapy

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Variants of *MIR137HG* Genes are Associated with Liver Cancer Risk in Chinese Li Population

This article was published in the following Dove Press journal:
OncoTargets and Therapy

Chaoying Wang[1,*]
Xiaohong Zhuang[1,*]
Junnv Xu[1,2,*]
Zhisheng Dai[3]
Weixiong Wu[4]
Chengsheng Zhang[1]
Shu Lin[1]
Sehong Chen[3]
Haifeng Lin[1]
Wenjun Tang[1]

[1]Department of Medical Oncology, The Second Affiliated Hospital of Hainan Medical University, Haikou, Hainan, 570311, People's Republic of China; [2]Hainan Medical University, Haikou, Hainan 571199, People's Republic of China; [3]Department of Medical Oncology, The Second People's Hospital of Hainan Province, Wuzhishan, Hainan 572200, People's Republic of China; [4]Intensive Care Medicine 1 District, The Second Affiliated Hospital of Hainan Medical College, Haikou, Hainan 570311, People's Republic of China

*These authors contributed equally to this work

Correspondence: Haifeng Lin; Wenjun Tang
#48, Baishuitang Road, Haikou 570311, Hainan, People's Republic of China
Tel/Fax +8613322060949; +8613322060949
Email 13322060949@163.com; 304200886@qq.com

Background: Liver cancer (LC) is the sixth most common cancer and the second leading cause of cancer mortality worldwide, and its incidence rate is high in China.

Methods: In this study, we aimed to investigate the contribution of *MIR137HG* (MIR137 Host Gene) polymorphisms to LC risk in a case–control study with 432 LC patients and 430 healthy controls. A logistic recession model was used to evaluate the effects of candidate single nucleotide polymorphisms (SNPs) on LC risk. HaploReg v 4.1 database was conducted to predict the potential functionality of SNPs.

Results: The results revealed that rs17371457 and rs7554283 in the *MIR137HG* gene were correlated with an enhanced LC risk under the allele ($P = 0.001$ and $P = 0.043$, respectively) and genetic models ($P < 0.05$). When the sample was stratified by gender and age, statistically significant associations were found. Rs9440302, rs17371457 and rs7554283 were associated with an increased the risk of LC among individuals aged >55 years ($P < 0.05$); rs17371457 was related to higher LC risk in males ($P < 0.05$). Similarly, the haplotype AG constituted by rs12333983 and rs3735451 significantly increased LC risk in Chinese Li population ($P = 0.043$). Six SNPs distributed in *MIR137HG* were successfully predicted as regulatory SNPs with different biological functions.

Conclusion: Our research firstly showed that *MIR137HG* gene polymorphisms were implicated in LC susceptibility among Chinese Li population.

Keywords: liver cancer, genetics polymorphisms, *MIR137HG*, susceptibility

Background

Liver cancer (LC) is the sixth most common cancer and the second largest cause of cancer mortality worldwide, with nearly 800,000 cases in 2012.[1] The highest incidence rates of LC were observed in Eastern Asia, South-Eastern Asia, Northern Africa and Southern Africa, with China accounting for about 50% of all cases.[2] LC is usually caused by underlying diseases, including hepatitis B virus (HBV), mycotoxins, poor diet and inactivity.[1] However, there is a startling lack of treatments on LC.

MiRNA is regarded as the critical elements in cancer development and involved in major cellular processes of cancer, such as cell differentiation, proliferation, migration, and apoptosis.[3–5] Further studies have shown that miRNAs can be used as tumor initiation factors, invasion and metastasis factors, tumor suppressor factors, etc. Hence, MiRNA may have a potential role in the development of new cancer control methods.[4,6]

MIR137HG (*MIR137* Host Gene) is an RNA Gene and is affiliated with the miRNA class. Currently, researches have shown that the promoter of *MIR137HG* is highly methylated in colorectal cancer,[7] malignant pleural mesothelioma,[8] and oral

submit your manuscript | www.dovepress.com
DovePress
http://doi.org/10.2147/OTT.S225669

squamous cell carcinoma.[9] *MIR137* is known to be a tumor suppressor, and the increased methylation level of *MIR137HG* in malignant pleural mesothelioma cell lines and colorectal cancer affects the occurrence and development of corresponding cancers.[7] However, there are no previous studies have investigated the association of LC risk and *MIR137HG* polymorphisms.

Hence, a case-control study was performed in 432 LC patients and 430 healthy controls to evaluate the possible association of LC and *MIR137HG* polymorphisms among Chinese Li population. The investigation was expected to further deepen our understanding regarding the pathogenesis of LC and serve as non-invasive biomarkers for LC.

Materials and Methods

Ethics Statement

Our research was approved by the Ethics Committee of the Second Affiliated Hospital of Hainan Medical College (Ethics number: HZ2015-11), and all procedures were conducted in accordance with the ethical standards of the Declaration of Helsinki and following national and international guidelines. Written informed consent was obtained from each participant after the study was fully explained.

*Study subjects*862 individuals, consisting of 430 healthy people and 432 patients with LC, were recruited from the Second Affiliated Hospital of Hainan Medical College, Hainan Province, China. All LC patients were newly diagnosed and had not received relevant antitumor therapy, such as surgical treatment, radiotherapy, chemotherapy and biotherapy. They were diagnosed with liver cancer by routine examination, imaging (such as B-ultrasonic wave, CT and nuclear magnetic resonance) and histopathological examination. Our case group also excluded patients with the following diseases: liver diseases like hepatitis and liver cirrhosis induced by diabetes, fatty liver, metabolic disorder and vascular disease, primary and secondary biliary cirrhosis, recurrent liver cancer, metastatic liver cancer, hepatitis A, C, D, and E, autoimmune hepatitis, etc. Control subjects, without any history of tumor or chronic diseases, were healthy individuals selected from the physical examination center of the Second Affiliated Hospital of Hainan Medical College, Hainan Province, China. All the participants were Li nationality from Hainan Province of China and they have no relationship with each other.

SNP Selection and Genotyping

According to the 1000 Genomes Project (http://www.1000genomes.org/) and dbSNP (https://www.ncbi.nlm.nih.gov/projects/SNP/) database, five label single nucleotide polymorphisms (SNPs) (rs12138817, rs9440302, rs1198574, rs17371457 and rs7554283) with minor allele frequency (MAF) > 5% were finally selected to evaluate the effect of *MIR137HG* polymorphisms on LC susceptibility.

Peripheral blood samples from all participants were collected in tubes coated with EDTA and were stored at 80°C. Following the manufacturer's guidelines, genomic DNA was extracted from participants' peripheral blood samples using the GoldMag whole blood genomic DNA purification kit (GoldMag Co. Ltd., Xi an, China). DNA concentration was measured with NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA).[10] The primers for the amplification reactions were designed using the Agena Bioscience Assay Design Suite V2.0 software to perform (Agena Bioscience, San Diego, CA, USA, https://agenacx.com/online-tools/).[11] Primers used for this study were listed in Supplementary Table S1. The MassARRAY iPLEX platform and Agena Bioscience TYPER version 4.0 software were used for SNP genotyping and data analysis, respectively.

SNP Functional Evaluation

HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) was used to predict the potential functions of SNPs in *MIR137HG* gene.

Statistical Analysis

SPSS software package (version 20.0; SPSS Inc., Chicago, IL, USA) was adopted to process data, and $P < 0.05$ was regarded as statistically significant.[12,13] Hardy–Weinberg equilibrium (HWE) P values obtained from exact tests[14] were used to evaluate whether the control group meets HWE. The distributions of genotype and allele frequencies between cases and controls were compared by χ^2 tests. The genotype-specific risks were estimated as odds ratios (ORs) and 95% confidence interval (CI) based on logistic regression model analysis.[15] Stratified analyses were also performed to assess the relationship between each SNP and the risk of LC in different subgroups. Finally, linkage disequilibrium (LD) analyses and haplotype analyses were performed using the SHEsis software (http://analysis.bio-x.cn/myAnalysis.php).

Results

Participant Characteristics

The basic demographic characteristics of study population are described in Table 1. A total of 432 LC patients (344 males and 88 females, age at diagnosis: 55.09 ± 11.59 years) and 430 controls (342 males and 88 females, age at diagnosis: 55.22 ± 10.73) were enrolled in our present study. There was no statistical difference in gender and age distribution between cases and control groups ($P > 0.05$).

Basic Information and Preliminary Statistics of the Selected SNPs

Basic information containing SNP ID, chromosomal position, alleles, MAF distribution, *P*-HWE, call rate, ORs and 95% CIs of all the candidate SNPs were demonstrated in Table 2. The call rate for all SNPs was >99.00% among the LC cases and the controls. All candidate SNPs were in accordance with HWE (*P*-HWE > 0.05), indicating good sample selection. From the five SNPs, the minor allele T of rs17371457 and the minor allele G of rs7554283 in the *MIR137HG* were significantly correlated with an enhanced the risk of LC (OR = 1.60, 95% CI = 1.01–1.47, $P = 0.001$; OR = 1.22, 95% CI = 1.01–1.47, $P = 0.043$, respectively).

Table 1 Distributions of Age and Gender in LC Cases and Healthy Controls

Variables	Case (%)	Control (%)	Total	*P*-value
Total	432	430	862	
Age				> 0.05[a]
Mean age ± SD	55.09 ± 11.59	55.22 ± 10.73		
> 55	209 (48%)	185 (43%)	394	
≤ 55	223 (52%)	245 (57%)	468	
Gender				> 0.05[b]
Male	344 (80%)	342 (80%)	686	
Female	88 (20%)	88 (20%)	176	

Notes: P^a-value obtained from independent sample *t*-test; P^b-value obtained from Pearson's χ^2 test.

Associations Between Genotype Frequencies and LC Risk

Furthermore, the correlations between polymorphisms and LC susceptibility were analyzed based on multiple inheritance modes (codominant, dominant, recessive, and log-additive models) using logistic tests. As shown in Table 3, rs17371457 was found to be correlated with an increased the risk of LC in the codominant model (C/T vs C/C: OR = 1.41, 95% CI = 1.01–1.96, $P = 0.044$; T/T vs C/C: OR = 5.14, 95% CI = 1.46–18.04, $P = 0.011$), dominant model (OR = 1.54, 95% CI = 1.12–2.13, $P = 0.008$), recessive model (OR = 4.78, 95% CI = 1.36–16.76, $P = 0.015$), and log-additive model (OR = 1.58, 95% CI = 1.18–2.11, $P = 0.002$). Rs7554283 was significantly correlated with an enhanced LC risk in the recessive model (OR = 1.39, 95% CI = 1.02–1.90, $P = 0.038$) and log-additive model (OR = 1.22, 95% CI = 1.01–1.47, $P = 0.042$). No significant correlation was observed between the remaining SNPs (rs12333983, rs3735451, rs4646437 and rs2246709) and LC risk.

Stratification Analysis by Age and Gender

According to the average age, stratified analyses regarding the impact of SNPs on the risk of LC are summarized in Table 4. Among those aged over 55 years, the genotype A/G-G/G of rs17371457 was associated with an increased LC risk in the codominant model (OR = 1.56, 95% CI = 1.01–2.42, $P = 0.047$). Rs17371457 significantly increased LC risk in the codominant model (CT vs CC: OR = 1.76, 95% CI = 1.05–2.93, $P = 0.031$; TT vs CC: OR = 8.32, 95% CI = 1.03–67.41, $P = 0.047$), dominant model (OR = 1.98, 95% CI = 1.20–3.25, $P = 0.007$), and log-additive model (OR = 1.97, 95% CI = 1.26–3.08, $P = 0.003$). We also found

Table 2 Basic Information and Allele Frequency of the Selected SNPs in *MIR137HG* Gene

SNP	Chr	Position	Gene(s)	Role	Alleles	Frequency (MAF)		*P* - HWE	Call Rate (%)	OR (95% CI)	*P* value
						Cases	Controls				
rs12138817	1	98001230	*MIR137HG*	ncRNA_intronic	C/T	140	127	0.564	99.20%	1.12 (0.86–1.45)	0.398
rs9440302	1	98027024	*MIR137HG*	ncRNA_intronic	G/A	392	360	1.000	99.40%	1.17 (0.97–1.41)	0.110
rs1198574	1	98029050	*MIR137HG*	ncRNA_intronic	A/G	57	47	0.123	100.00%	1.23 (0.82–1.82)	0.318
rs17371457	1	98032783	*MIR137HG*	ncRNA_intronic	T/C	129	85	0.785	100.00%	1.60 (1.20–2.15)	0.001*
rs7554283	1	98034489	*MIR137HG*	ncRNA_intronic	G/C	453	408	0.628	99.80%	1.22 (1.01–1.47)	0.043*

Notes: *P*-HWE obtained from Fisher's exact test; *P*-value obtained from Wald test; *P-value < 0.05 indicates statistical significance.
Abbreviations: SNP, single nucleotide polymorphism; HWE, Hardy–Weinberg equilibrium; OR, odds ratio; 95% CI, 95% confidence interval.

Table 3 Significant Genetic Variants in *MIR137HG* Gene Associated with the Susceptibility of LC in Chinese Li Population

SNP	Model	Genotype	Control	Case	Without Adjustment		With Adjustment	
					OR (95% CI)	P^a-value	OR (95% CI)	P^b-value
rs12138817	Codominant	T/T T/C C/C	313 (73%) 105 (24.5%) 11 (2.6%)	307 (71.4%) 106 (24.7%) 17 (4.0%)	1 1.03 (0.75–1.41) 1.58 (0.73–3.42)	 0.857 0.250	1 1.03 (0.75–1.41) 1.58 (0.73–3.42)	 0.851 0.251
	Dominant	T/T T/C-C/C	313 (73%) 116 (27.0%)	307 (71.4%) 123 (28.6%)	1 1.08 (0.80–1.46)	0.609	1 1.08 (0.8–1.46)	0.603
	Recessive	T/T-T/C C/C	418 (87.4%) 11 (2.6%)	413 (99.6%) 17 (4.0%)	1 1.56 (0.72–3.38)	0.255	1 1.56 (0.72–3.38)	0.256
	Log-additive	–	–	–	1.11 (0.86–1.43)	0.413	1.11 (0.86–1.43)	0.409
rs9440302	Co-dominant	A/A A/G G/G	145 (33.7%) 210 (48.8%) 75 (17.4%)	127 (29.6%) 212 (49.4%) 90 (21.0%)	1 1.15 (0.85–1.56) 1.37 (0.93–2.02)	 0.362 0.112	1 1.15 (0.85–1.57) 1.38 (0.93–2.04)	 0.358 0.107
	Dominant	A/A A/G-G/G	145 (33.7%) 285 (66.3%)	127 (29.6%) 302 (70.4%)	1 1.21 (0.91–1.61)	0.195	1 1.21 (0.91–1.62)	0.191
	Recessive	A/A-A/G G/G	355 (82.6%) 75 (17.4%)	339 (79.0%) 90 (21.0%)	1 1.26 (0.89–1.77)	0.189	1 1.26 (0.9–1.78)	0.182
	Log-additive	–	–	–	1.17 (0.97–1.41)	0.111	1.17 (0.97–1.42)	0.106
rs1198574	Co-dominant	G/G G/A A/A	387 (89.8%) 41 (9.5%) 3 (0.7%)	377 (87.3%) 53 (12.3%) 2 (0.5%)	1 1.33 (0.86–2.04) 0.68 (0.11–4.12)	 0.199 0.679	1 1.33 (0.86–2.04) 0.68 (0.11–4.11)	 0.202 0.675
	Dominant	G/G G/A-A/A	387 (89.8%) 44 (10.2%)	377 (87.3%) 55 (12.7%)	1 1.28 (0.84–1.96)	0.246	1 1.28 (0.84–1.96)	0.249
	Recessive	G/G-G/T A/A	428 (99.3%) 3 (0.7%)	430 (99.5%) 2 (0.5%)	1 0.66 (0.11–3.99)	0.654	1 0.66 (0.11–3.97)	0.649
	Log-additive	–	–	–	1.22 (0.82–1.8)	0.328	1.21 (0.82–1.8)	0.333
rs17371457	Co-dominant	C/C C/T T/T	349 (81%) 79 (18.3%) 3 (0.7%)	317 (73.4%) 101 (23.4%) 14 (3.2%)	1 1.41 (1.01–1.96) 5.14 (1.46–18.04)	 0.043* 0.011*	1 1.41 (1.01–1.96) 5.14 (1.46–18.04)	 0.044* 0.011*
	Dominant	C/C C/T-T/T	349 (81.0%) 82 (19.0%)	317 (73.4%) 115 (26.6%)	1 1.54 (1.12–2.13)	0.008*	1 1.54 (1.12–2.13)	0.008*
	Recessive	C/C-C/T T/T	428 (99.3%) 3 (0.7%)	418 (96.8%) 14 (3.2%)	1 4.78 (1.36–16.75)	0.015*	1 4.78 (1.36–16.76)	0.015*
	Log-additive	–	–	–	1.58 (1.18–2.11)	0.002*	1.58 (1.18–2.11)	0.002*
rs7554283	Co-dominant	C/C G/C G/G	115 (26.8%) 220 (51.3%) 94 (21.9%)	100 (23.1%) 211 (48.8%) 121 (28%)	1 1.1 (0.79–1.53) 1.48 (1.01–2.17)	 0.558 0.043	1 1.10 (0.79–1.53) 1.48 (1.01–2.17)	 0.556 0.042
	Dominant	C/C C/G-G/G	115 (26.8%) 314 (73.2%)	100 (23.1%) 332 (76.9%)	1 1.22 (0.89–1.66)	0.215	1 1.22 (0.89–1.66)	0.214
	Recessive	C/C-G/C G/G	335 (78.1%) 94 (21.9%)	311 (72.0%) 121 (28.0%)	1 1.39 (1.02–1.89)	0.039*	1 1.39 (1.02–1.90)	0.038*
	Log-additive	–	–	–	1.22 (1.01–1.47)	0.043*	1.22 (1.01–1.47)	0.042*

Notes: P^a-value calculated by logistic regression analysis; P^b-value calculated by logistic regression analysis with adjustments for gender and age; *P-value < 0.05 indicates statistical significance.

Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval.

Table 4 Relationship of *MIR137HG* Polymorphisms with LC Risk Stratified by Age

SNP	Model	Genotype	Age (Years) ≤ 55				Age (Years) > 55			
			Control	Case	OR (95% CI)	*P* value	Control	Case	OR (95% CI)	*P* value
rs12138817	Codominant	T/T T/C C/C	182 (74.0%) 59 (24.0%) 5 (2.0%)	158 (70.9%) 54 (24.2%) 11 (4.9%)	1 1.05 (0.68–1.61) 2.78 (0.94–8.24)	 0.831 0.065	131 (71.6%) 46 (25.1%) 6 (3.3%)	149 (72.0%) 52 (25.1%) 6 (2.9%)	1 1.00 (0.63–1.59) 0.86 (0.27–2.75)	 0.990 0.803
	Dominant	T/T T/C-C/C	182 (74.0%) 64 (26.0%)	158 (70.9%) 65 (29.1%)	1 1.18 (0.78–1.78)	0.429	131 (71.6%) 52 (28.4%)	149 (72.0%) 58 (28.0%)	1 0.99 (0.63–1.54)	0.952
	Recessive	T/T-T/C C/C	241 (98.0%) 5 (2.0%)	212 (95.1%) 11 (4.9%)	1 2.75 (0.93–8.11)	0.067	177 (96.7%) 6 (3.3%)	201 (97.1%) 6 (2.9%)	1 0.86 (0.27–2.73)	0.801
	Log-additive	–	–	–	1.26 (0.89–1.78)	0.190	–	–	0.97 (0.67–1.42)	0.893
	Allele	T C	423 (86.0%) 69 (14.0%)	370 (83.0%) 76 (17.0%)	1 1.26 (0.88–1.8)	0.202	308 (84.2%) 58 (15.8%)	350 (84.5%) 64 (15.5%)	1 0.97 (0.66–1.43)	0.882
rs9440302	Co-dominant	A/A A/G G/G	82 (33.3%) 128 (52.0%) 36 (14.6%)	75 (33.8%) 106 (47.7%) 41 (18.5%)	1 0.92 (0.61–1.38) 1.28 (0.74–2.22)	 0.686 0.381	63 (34.2%) 82 (44.6%) 39 (21.2%)	52 (25.1%) 106 (51.2%) 49 (23.7%)	1 1.57 (0.98–2.5) 1.55 (0.88–2.72)	 0.060 0.126
	Dominant	A/A A/G-G/G	82 (33.3%) 164 (66.7%)	75 (33.8%) 147 (66.2%)	1 1 (0.68–1.47)	0.991	63 (34.2%) 121 (65.8%)	52 (25.1%) 155 (74.9%)	1 1.56 (1.01–2.42)	0.047*
	Recessive	A/A-A/G G/G	210 (85.4%) 36 (14.6%)	181 (81.5%) 41 (18.5%)	1 1.35 (0.82–2.2)	0.239	145 (78.8%) 39 (21.2%)	158 (76.3%) 49 (23.7%)	1 1.18 (0.73–1.9)	0.509
	Log-additive	–	–	–	1.09 (0.83–1.42)	0.529	–	–	1.27 (0.96–1.67)	0.099
	Allele	A G	292 (59.3%) 200 (40.7%)	256 (57.7%) 188 (42.3%)	1 1.07 (0.83–1.39)	0.600	208 (56.5%) 160 (43.5%)	210 (50.7%) 204 (49.3%)	1 1.26 (0.95–1.67)	0.105
rs17371457	Co-dominant	C/C C/T T/T	194 (78.9%) 50 (20.3%) 2 (0.8%)	166 (74.4%) 51 (22.9%) 6 (2.7%)	1 1.17 (0.75–1.82) 3.26 (0.65–16.42)	 0.493 0.153	155 (83.8%) 29 (15.7%) 1 (0.5%)	151 (72.2%) 50 (23.9%) 8 (3.8%)	1 1.76 (1.05–2.93) 8.32 (1.03–67.41)	 0.031* 0.047*
	Dominant	C/C C/T-T/T	194 (78.9%) 52 (21.1%)	166 (74.4%) 57 (25.6%)	1 1.25 (0.81–1.93)	0.310	155 (83.8%) 30 (16.2%)	151 (72.2%) 58 (27.8%)	1 1.98 (1.20–3.25)	0.007*
	Recessive	C/C-C/T T/T	244 (99.2%) 2 (0.8%)	217 (97.3%) 6 (2.7%)	1 3.15 (0.63–15.81)	0.164	184 (99.5%) 1 (0.5%)	201 (96.2%) 8 (3.8%)	1 7.46 (0.92–60.34)	0.060
	Log-additive	–	–	–	1.3 (0.88–1.92)	0.190	–	–	1.97 (1.26–3.08)	0.003*
	Allele	C	438 (89.0%)	383 (85.9%)	1	0.145	339 (91.6%)	352 (84.2%)	1	0.002

(*Continued*)

Table 4 (Continued).

SNP	Model	Genotype	Age (Years) ≤ 55				Age (Years) > 55			
			Control	Case	OR (95% CI)	P value	Control	Case	OR (95% CI)	P value
		T	54 (11.0%)	63 (14.1%)	1.33 (0.9–1.97)		31 (8.4%)	66 (15.8%)	2.05 (1.31–3.22)	
rs7554283	Co-dominant	C/C	65 (26.5%)	59 (26.5%)	1		50 (27.2%)	41 (19.6%)	1	
		G/C	127 (51.8%)	103 (46.2%)	0.89 (0.57–1.38)	0.595	93 (50.5%)	108 (51.7%)	1.42 (0.86–2.34)	0.166
		G/G	53 (21.6%)	61 (27.4%)	1.31 (0.79–2.19)	0.300	41 (22.3%)	60 (28.7%)	1.82 (1.02–3.23)	0.042*
	Dominant	C/C	65 (26.5%)	59 (26.5%)	1	0.961	50 (27.2%)	41 (19.6%)	1	0.072
		C/G-G/G	180 (73.5%)	164 (73.5%)	1.01 (0.67–1.53)		134 (72.8%)	168 (80.4%)	1.54 (0.96–2.48)	
	Recessive	C/C-G/C	192 (78.4%)	162 (72.6%)	1	0.109	143 (77.7%)	149 (71.3%)	1	0.131
		G/G	53 (21.6%)	61 (27.4%)	1.42 (0.93–2.17)		41 (22.3%)	60 (28.7%)	1.43 (0.9–2.26)	
	Log-additive	–	–	–	1.14 (0.88–1.47)	0.319	–	–	1.35 (1.01–1.80)	0.042*
	Allele	C	257 (52.4%)	221 (49.6%)	1	0.376	193 (59.3%)	190 (57.7%)	1	0.050
		G	233 (47.6%)	225 (50.4%)	1.12 (0.87–1.45)		175 (40.7%)	228 (42.3%)	1.32 (1–1.75)	

Notes: *P*-value calculated by logistic regression analysis with adjustments for gender and age; *P-value < 0.05 indicates statistical significance.
Abbreviations: SNP, single nucleotide polymorphism; OR, odds ratio; 95% CI, 95% confidence interval.

that rs7554283 was correlated with an increased risk of LC in the codominant model (G/G vs C/C: OR = 1.82, 95% CI = 1.02–3.23, $P = 0.042$) and log-additive model (OR = 1.35, 95% CI = 1.01–1.80, $P = 0.042$). No significant association was found between the five SNPs loci in *MIR137HG* and LC risk at age ≤50 years.

Stratified analyses by gender were also conducted. As shown in Table 5, the results revealed that rs17371457 was correlated with an enhanced risk of LC in males under the codominant model (C/T vs C/C: OR = 1.52, 95% CI = 1.01–1.80, $P = 0.042$; T/T vs C/C: OR = 14.61, 95% CI = 1.90–112.50, $P = 0.010$), dominant model (OR = 1.73, 95% CI = 1.20–2.50), $P = 0.003$), recessive model (OR = 13.41, 95% CI = 1.75–103.10, $P = 0.013$), log-additive model (OR = 1.81, 95% CI = 1.30–2.52, $P < 0.001$) and allele model (OR = 1.85, 95% CI = 1.33–2.58, $P < 0.001$). In the subgroup of females, we did not find a significant association between candidate SNPs in *MIR137HG* and LC risk.

Associations Between Haplotype Analyses and LC Risk

Finally, we studied the LD and haplotype analysis of candidate SNPs. As shown in Figure 1, the LD block in *MIR137HG* gene was constituted of two SNPs (rs9440302 and rs1198574). The frequency distribution of haplotypes in two groups and its association with LC risk were presented in Table 6. Haplotype AG in the *MIR137HG* gene was found to prominently increase the risk of LC (OR = 1.22, 95% CI = 1.01–1.47; $P = 0.043$), other haplotypes did not display the correlativity.

SNP Functional Evaluation

In order to explore the regulatory roles of the statistically significant variants, we predicted the function of them using HaploReg v4.1 database. Six SNPs in *MIR137HG* were successfully predicted as regulatory SNPs with different biological functions (Supplementary Table S2).

Discussion

The miRNA variants have been recognized as major contributors to the pathogenesis and development of cancers. In this study, we aimed to elucidate the relationship between *MIR137HG* gene polymorphisms and LC susceptibility in Chinese Li population. The results suggested that rs17371457 and rs7554283 in the *MIR137HG* gene were correlated with an enhanced LC risk. After stratifying the

Table 5 Relationship of *MIR137HG* Polymorphisms with LC Risk Stratified by Gender

SNP	Model	Genotype	Male				Female			
			Control	Case	OR (95% CI)	*P*-value	Control	Case	OR (95% CI)	*P*-value
rs17371457	Co-dominant	C/C	281 (82.2%)	250 (72.7%)	1		67 (76.1%)	67 (76.1%)	1	
		C/T	60 (17.5%)	81 (23.5%)	1.52 (1.04–2.21)	0.030*	19 (21.6%)	20 (22.7%)	1.05 (0.51–2.15)	0.895
		T/T	1 (0.3%)	13 (3.8%)	14.61 (1.90–112.50)	0.010*	2 (2.3%)	1 (1.1%)	0.5 (0.04–5.64)	0.574
	Dominant	C/C	281 (82.2%)	250 (72.7%)	1	0.003*	67 (76.1%)	67 (76.1%)	1	0.993
		C/T-T/T	61 (17.8%)	94 (27.3%)	1.73 (1.20–2.50)		21 (23.9%)	21 (23.9%)	1 (0.50–2.00)	
	Recessive	C/C-C/T	341 (99.7%)	331 (96.2%)	1	0.013*	86 (97.7%)	87 (98.9%)	1	0.567
		T/T	1 (0.3%)	13 (3.8%)	13.41 (1.75–103.10)		2 (2.3%)	1 (1.1%)	0.49 (0.04–5.54)	
	Log-additive	–	–	–	1.81 (1.30–2.52)	<0.001*	–	–	0.95 (0.51–1.77)	0.867
	Allele	C	622 (90.9%)	581 (84.4%)	1	<0.001*	154 (50.2%)	23 (51.1%)	1	0.873
		T	62 (9.1%)	107 (15.6%)	1.85 (1.33–2.58)		153 (49.8%)	22 (48.9%)	0.95 (23.00–0.51)	
rs7554283	Co-dominant	C/C	90 (26.5%)	77 (22.4%)	1		25 (28.4%)	23 (26.1%)	1	
		G/C	175 (51.5%)	173 (50.3%)	1.16 (0.80–1.68)	0.434	44 (50.0%)	38 (43.2%)	0.94 (0.46–1.92)	0.862
		G/G	75 (22.1%)	94 (27.3%)	1.47 (0.96–2.27)	0.078	19 (21.6%)	27 (30.7%)	1.55 (0.68–3.49)	0.297
	Dominant	C/C	90 (26.5%)	77 (22.4%)	1	0.208	25 (28.4%)	23 (26.1%)	1	0.735
		C/G-G/G	250 (73.5%)	267	1.25 (0.88–1.78)		63 (71.6%)	65 (73.9%)	1.12 (0.58–2.18)	
	Recessive	C/C-G/C	265 (77.9%)	250 (72.7%)	1	0.108	69 (78.4%)	61 (99.3%)	1	0.173
		G/G	75 (22.1%)	94 (27.3%)	1.33 (0.94–1.89)		19 (21.6%)	27 (30.7%)	1.61 (0.81–3.18)	
	Log-additive	–	–	–	1.21 (0.98–1.51)	0.078	–	–	1.24 (0.82–1.86)	0.305
	Allele	C	355 (52.2%)	327 (47.5%)	1	0.084	84 (47.2%)	82 (47.1%)	1	0.286
		G	325 (47.8%)	361 (52.5%)	1.21 (0.98–1.49)		94(52.8%)	92(52.9%)	1.26 (82.00–0.83)	

Notes: *P*-value calculated by logistic regression analysis with adjustments for gender and age; **P*-value < 0.05 indicates statistical significance.
Abbreviations: SNP, single nucleotide polymorphism; OR, Odds ratio; 95% CI, 95% confidence interval.

sample by gender and age, correlation were found between higher risk of LC and *MIR137HG* polymorphisms (rs9440302, rs17371457 and rs7554283) at age >55 years. Rs17371457 was related to higher LC risk in males. Similarly, the haplotype AG significantly increased LC risk in Chinese Li population. As far as we know, this is the first study to show the association of *MIR137HG* polymorphisms and LC susceptibility in Chinese Li population.

MiRNAs are involved in the major cellular processes of cancer, and *MIR137HG* belongs to the miRNA class. Previous studies showed that *MIR137HG* gene polymorphisms affect the gray matter structure in patients with schizophrenia, and it was also related to the volume of corpus callosum.[16,17] At the same time, it has been reported that the methylation level of *MIR137HG* promoter was increased in a variety of cancers and was related to the expression level of MIR137. Morandi et al believed that methylation level was the key factor for early detection of oral squamous cell carcinoma (OSCC), and found that *MIR137HG* was hypermethylated in OSCC and high-grade squamous intraepithelial lesion (HGSIL).[9] Johnson et al observed that *MIR137HG* was methylated in malignant pleural mesothelioma (MPM) cell lines. Besides, *MIR137* showed tumor inhibition function in MPM by targeting *YBX1*, which was dysregulated in MPM because of Copy number variations (CNV) and hypermethylation of the *MIR137HG* promoter.[8] Li et al showed that *HSF1* recruited *DNMT3a* to inhibit the promoter methylation of lncRNA *MIR137HG* by targeting *GLS1* mRNA, and further reduced the expression of *MIR137HG-MIR137*, and *HSF1* stimulated *GLS1*-dependent mTOR activation to promote colorectal carcinogenesis.[7]

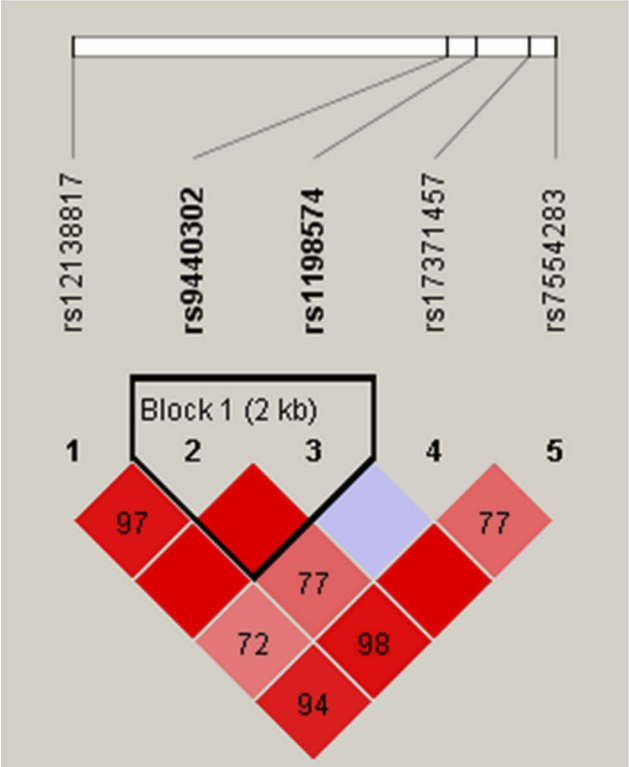

Figure 1 Haplotype block map for six SNPs in *MIR137HG* Gene.

Furthermore, *MIR137HG* is the host gene of *MIR137*. It is reported that *MIR137* was a tumor suppressor by inhibiting cancer proliferation, cell-cycle progression and migration,[18,19] and it was lowly expressed in various cancers, such as hepatocellular carcinoma, lung cancer, colorectal cancer, gastric cancer, and cervical cancer.[20–22] Additionally, studies have shown that *MIR137* may serve as a prognostic marker in hepatocellular carcinoma (HCC) patients,[23] and miR-137/miR-133a-TINCR pathway can be used as a promising target for tumor recurrence and prognosis in HCC patients.[24] Researches have shown that *MIR137* inhibits tumor growth and metastasis in HCC by targeting AKT serine/threonine kinase2[25] and the suppressive effects of *MIR137* on HCC cell proliferation and metastasis were regulated by cell division cycle.[26] *MIR137* was down-regulated in HCC and inhibits HCC cell migration and invasion by targeting EZH2-STAT3 signaling pathway.[27] Therefore, *MIR137HG* may also affect the expression of *MIR137* through promoter hypermethylation, thereby affecting the metastasis and growth of LC.

Significantly, the major allele of *MIR137HG* risk SNPs, including rs1625579, lead to lower levels of mature *MIR137* than those with the minor allele.[28] A post-mortem tissue analysis also showed that the biogenesis of this miRNA may be disrupted in carriers of the rs1625579 risk genotypes.[29] In our case–control study, we investigated the correlation between *MIR137HG* SNPs and LC risk, and the result showed that rs17371457 and rs7554283 in *MIR137HG* gene exhibited an increased risk of LC. Rs9440302, rs17371457 and rs7554283 also correlation with an increased LC risk at age >55 years. The rs17371457 was related to higher LC risk in males. Based on these results, we hypothesized that the allele of *MIR137HG* risk SNPs (s9440302, rs17371457 and rs7554283) reduced the level of mature *MIR137*, thereby increasing the risk of LC. Additionally, the case group in our study included patients with confirmed liver cancer, excluding patients with liver cirrhosis, chronic hepatitis B and C, NAFLD, Wilson's disease and other liver diseases. Thus, the distribution of MIR137HG SNPs in other liver diseases (such as chronic B and C hepatitis, cirrhosis, NAFLD and Wilson's disease), and the relationship with liver-related diseases are also unknown. The specific mechanism of MIR137HG SNPs on LC risk and the relationship between MIR137HG SNPs and other liver diseases still need further study. Several limitations of this study should not be ignored. Firstly, current study may be limited by the source of samples, which may limit statistical capacity. Secondly, the size of sample was not large enough to rule out false negative results. So, larger sample size and more ethnic groups are needed for further verification. Despite the

Table 6 Haplotype Frequencies of *MIR137HG* Gene and the Association with LC Risk

SNP	Chr	Position	Gene(s)	Role	Alleles	Frequency (MAF)		*P* - HWE	Call Rate (%)	OR (95% CI)	*P* value
						Cases	Controls				
rs12138817	1	98001230	*MIR137HG*	ncRNA_intronic	C/T	140	127	0.564	99.20%	1.12 (0.86–1.45)	0.398
rs9440302	1	98027024	*MIR137HG*	ncRNA_intronic	G/A	392	360	1	99.40%	1.17 (0.97–1.41)	0.11
rs1198574	1	98029050	*MIR137HG*	ncRNA_intronic	A/G	57	47	0.123	100.00%	1.23 (0.82–1.82)	0.318
rs17371457	1	98032783	*MIR137HG*	ncRNA_intronic	T/C	129	85	0.785	100.00%	1.60 (1.20–2.15)	0.001*
rs7554283	1	98034489	*MIR137HG*	ncRNA_intronic	G/C	453	408	0.628	99.80%	1.22 (1.01–1.47)	0.043*

Notes: Block comprised of the three closely linked SNPs rs9440302 and rs1198574; *P* value calculated by Wald test and adjusted by gender and age; *P-value < 0.05 indicates statistical significance.

Abbreviations: OR, odds ratio; 95% CI, 95% confidence interval.

above limitations, the results of this study provide a scientific basis for future studies on the relationship between LC and *MIR137HG* gene SNPs.

In general, this study firstly revealed the potential role of *MIR137HG* on LC susceptibility in the Chinese Li population, which provides new insights into the pathogenesis of LC. Further studies with larger samples and well-designed methods in diverse populations are necessary to confirm our results.

Acknowledgments

This work is supported by the Project of Hainan Health and Family Planning Science and Education (Grant No. 15A200102). We are grateful to the individuals for their participation in this study. We also thank the clinicians and hospital staff who contributed to the sample and data collection for this study. We would also like to thank all participants for this manuscript.

Disclosure

The authors report that they have no conflicts of interest in this work.

References

1. Gravitz L. Liver cancer. *Nature*. 2014;516(7529):S1. doi:10.1038/516S1a
2. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359–E386. doi:10.1002/ijc.29210
3. Tam C, Wong JH, Tsui SKW, Zuo T, Chan TF, Ng TB. LncRNAs with miRNAs in regulation of gastric, liver, and colorectal cancers: updates in recent years. *Appl Microbiol Biotechnol*. 2019;103(12):4649–4677.
4. Price C, Chen J. MicroRNAs in cancer biology and therapy: current status and perspectives. *Genes Dis*. 2014;1(1):53–63. doi:10.1016/j.gendis.2014.06.004
5. Tutar L, Ozgur A, Tutar Y. Involvement of miRNAs and Pseudogenes in cancer. *Methods Mol Biol*. 2018;1699:45–66.
6. Liu M, Jiang L, Guan XY. The genetic and epigenetic alterations in human hepatocellular carcinoma: a recent update. *Protein Cell*. 2014;5(9):673–691. doi:10.1007/s13238-014-0065-9
7. Li J, Song P, Jiang T, et al. Heat shock factor 1 epigenetically stimulates glutaminase-1-dependent mTOR activation to promote colorectal carcinogenesis. *Mol Ther*. 2018;26(7):1828–1839. doi:10.1016/j.ymthe.2018.04.014
8. Johnson TG, Schelch K, Cheng YY, et al. Dysregulated expression of the MicroRNA miR-137 and its target YBX1 contribute to the invasive characteristics of malignant pleural mesothelioma. *J Thorac Oncol*. 2018;13(2):258–272. doi:10.1016/j.jtho.2017.10.016
9. Morandi L, Gissi D, Tarsitano A, et al. CpG location and methylation level are crucial factors for the early detection of oral squamous cell carcinoma in brushing samples using bisulfite sequencing of a 13-gene panel. *Clin Epigenetics*. 2017;9:85. doi:10.1186/s13148-017-0386-7
10. Geng TT, Xun XJ, Li S, et al. Association of colorectal cancer susceptibility variants with esophageal cancer in a Chinese population. *World J Gastroenterol*. 2015;21(22):6898–6904.
11. Wang T, Chen T, Thakur A, et al. Association of PSMA4 polymorphisms with lung cancer susceptibility and response to cisplatin-based chemotherapy in a Chinese Han population. *Clin Transl Oncol*. 2015;17(7):564–569. doi:10.1007/s12094-015-1279-x
12. Huang CY, Xun XJ, Wang AJ, et al. CHRNA5 polymorphisms and risk of lung cancer in Chinese Han smokers. *Am J Cancer Res*. 2015;5(10):3241–3248.
13. Zhang T, Li X, Du Q, et al. DUSP10 gene polymorphism and risk of colorectal cancer in the Han Chinese population. *Eur J Cancer Prev*. 2014;23(3):173–176. doi:10.1097/CEJ.0b013e3283647408
14. Lorenzl S, De Pasquale G, Segal AZ, Beal MF. Dysregulation of the levels of matrix metalloproteinases and tissue inhibitors of matrix metalloproteinases in the early phase of cerebral ischemia. *Stroke*. 2003;34(6):e37–e38; author reply e37–e38. doi:10.1161/01.STR.0000075563.45920.24
15. Newby AC. Matrix metalloproteinase inhibition therapy for vascular diseases. *Vascul Pharmacol*. 2012;56(5–6):232–244. doi:10.1016/j.vph.2012.01.007
16. He E, Lozano MAG, Stringer S, et al. MIR137 schizophrenia-associated locus controls synaptic function by regulating synaptogenesis, synapse maturation and synaptic transmission. *Hum Mol Genet*. 2018;27(11):1879–1891. doi:10.1093/hmg/ddy089
17. Wright C, Gupta CN, Chen J, et al. Polymorphisms in MIR137HG and microRNA-137-regulated genes influence gray matter structure in schizophrenia. *Transl Psychiatry*. 2016;6:e724. doi:10.1038/tp.2015.211
18. Shen H, Wang L, Ge X, et al. MicroRNA-137 inhibits tumor growth and sensitizes chemosensitivity to paclitaxel and cisplatin in lung cancer. *Oncotarget*. 2016;7(15):20728–20742. doi:10.18632/oncotarget.8011
19. Liang ML, Hsieh TH, Ng KH, et al. Downregulation of miR-137 and miR-6500-3p promotes cell proliferation in pediatric high-grade gliomas. *Oncotarget*. 2016;7(15):19723–19737. doi:10.18632/oncotarget.7736
20. Zhang H, Yan T, Liu Z, et al. MicroRNA-137 is negatively associated with clinical outcome and regulates tumor development through EZH2 in cervical cancer. *J Cell Biochem*. 2018;119(1):938–947. doi:10.1002/jcb.26259
21. Yang YR, Li YX, Gao XY, Zhao SS, Zang SZ, Zhang ZQ. MicroRNA-137 inhibits cell migration and invasion by targeting bone morphogenetic protein-7 (BMP7) in non-small cell lung cancer cells. *Int J Clin Exp Pathol*. 2015;8(9):10847–10853.
22. Smith AR, Marquez RT, Tsao WC, et al. Tumor suppressive microRNA-137 negatively regulates Musashi-1 and colorectal cancer progression. *Oncotarget*. 2015;6(14):12558–12573. doi:10.18632/oncotarget.3726
23. Sakabe T, Azumi J, Umekita Y, et al. Prognostic relevance of miR-137 in patients with hepatocellular carcinoma. *Liver Int*. 2017;37(2):271–279. doi:10.1111/liv.2017.37.issue-2
24. Li J, Gao J, Tian W, Li Y, Zhang J. Long non-coding RNA MALAT1 drives gastric cancer progression by regulating HMGB2 modulating the miR-1297. *Cancer Cell Int*. 2017;17:44. doi:10.1186/s12935-017-0408-8
25. Liu LL, Lu SX, Li M, et al. FoxD3-regulated microRNA-137 suppresses tumour growth and metastasis in human hepatocellular carcinoma by targeting AKT2. *Oncotarget*. 2014;5(13):5113–5124. doi:10.18632/oncotarget.2089
26. Gao M, Liu L, Li S, Zhang X, Chang Z, Zhang M. Inhibition of cell proliferation and metastasis of human hepatocellular carcinoma by miR-137 is regulated by CDC42. *Oncol Rep*. 2015;34(5):2523–2532. doi:10.3892/or.2015.4261
27. Huang B, Huang M, Li Q. MiR-137 suppresses migration and invasion by targeting EZH2-STAT3 signaling in human hepatocellular carcinoma. *Pathol Res Pract*. 2018;214(12):1980–1986. doi:10.1016/j.prp.2018.08.005

28. Siegert S, Seo J, Kwon EJ, et al. The schizophrenia risk gene product miR-137 alters presynaptic plasticity. *Nat Neurosci*. 2015;18(7):1008–1016. doi:10.1038/nn.4023

29. Guella I, Sequeira A, Rollins B, et al. Analysis of miR-137 expression and rs1625579 in dorsolateral prefrontal cortex. *J Psychiatr Res*. 2013;47(9):1215–1221. doi:10.1016/j.jpsychires.2013.05.021