Influence of CMTM8 polymorphisms on Lung Cancer Susceptibility in the Chinese Han Population

Jiamin Wu
Northwest University
Yao Sun
Northwest University
Zichao Xiong
Northwest University
Fanglin Niu
Northwest University
Yuanwei Liu
Northwest University
Haiyue Li
Northwest University
Jianfeng Liu
Northwest University
Qiufang Liu
Northwest University
Tianbo Jin (tianbojin_69@163.com)
Northwest University https://orcid.org/0000-0001-7551-5760

Research

Keywords: Polymorphisms, CMTM8, Lung cancer, Chinese Han population, case-control study

Posted Date: June 11th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-34577/v1

Version of Record: A version of this preprint was published at Pharmacogenetics and Genomics on December 30th, 2020. See the published version at https://doi.org/10.1097/FPC.0000000000000426.

Abstract

Objective: Lung cancer is the leading cause of cancer-related mortality worldwide and *CMTM8* is a potential tumor suppressor gene, which is down-regulated in lung cancer. The objective of this research was to assess the association of *CMTM8* genetic polymorphisms with lung cancer risk in Chinese Han population.

Methods: To evaluate the correlation between *CMTM8* polymorphisms and lung cancer risk, Agena MassArray platform was used for genotype determination among 509 lung cancer patients and 506 controls. Multiple genetic models, stratification analysis and haploview analysis was used by calculating odds ratio (OR) and 95% confidence intervals (CIs).

Results: Significant associations were detected between *CMTM8* rs6771238 and an increased lung cancer risk in codominant (adjusted OR = 1.57, 95%CI: 1.01-2.42, p = 0.044) and dominant (adjusted OR = 1.54, 95%CI: 1.01-2.36, p = 0.047)models. After gender stratification analysis,weobserved that rs6771238 was related to an increased risk of lung squamous cell carcinoma, while rs6771238 was associated with anincreased risk of lung adenocarcinoma. Rs9835916 and rs1077868 were correlated with lung cancer staging.Rs9835916 was linked to increased risk of lymph node metastasis in lung cancer patients.

Conclusions: Our study firstly reported that the *CMTM8* polymorphisms were a risk factors for lung cancer, which suggested the potential roles of *CMTM8* in the development of lung cancerin Chinese Han population.

Introduction

Lung cancer is the leading cause of cancer-related mortality worldwide. Lung cancer is also a major health problem in China, where 605,946 new cases of lung cancer (416,333 men and 189,613 women) and 486,555 lung cancer-related deaths were reported in 2010. Despite the advance of therapeutic strategies, the prognosis of lung cancer patients remains poor, and the 5-year survival rate remains less than 10% in most parts of the world [1, 2]. The reason for this lack of improvement may be due to the high invasion and recurrence rate of lung cancer. Lung cancer development appears to result from a complex interaction between environmental exposures and genetic factors. And genetic factors may play a fundamental role in the development of lung cancer, such as the fact that some families are more likely to develop cancer than others in the general population. [3]. Therefore, novel biomarkers for predicting the tumor progression of lung cancer are urgently needed.

CMTM8 (CKLF-like MARVEL transmembrane domain containing 8), also known as *CKLFSF8*, belongs to the chemokine-like factor gene superfamily, a novel family that was first cloned by the Peking University Human Disease Genomics Research Center in 2003 [4–6]. The encoding product of this family gene has a special structure, which is between chemokines and four transmembrane proteins. Studies have shown that *CMTM8* is a potential tumor suppressor that can participate in various signal transduction pathways to control tumor occurrence and development, thereby affecting tumor formation, development and metastasis [7]. *CMTM8* is widely expressed in many normal human tissues and is often downregulated or absent in multiple solid tumors, including the liver, lung, colon, rectum, esophagus, stomach [8]. The over-expression of *CMTM8* can inhibit the proliferation, migration, and invasion of carcinoma cells [9]. Hence, *CMTM8* is a potential marker of early tumor detection in many cancers, including lung cancer.

Single nucleotide polymorphism (SNP), as a natural sequence variation, may affect the expression level of *CMTM8*. As for genetic variation, many SNPs sites affecting the occurrence of lung squamous cell carcinoma have been reported in GWAS studies in Chinese Han population [10, 11].However, the risk of *CMTM8* polymorphisms on lung cancer has been unexplored to date. Therefore, we performed a case-control study to determine whether *CMTM8* SNPs impact susceptibility to lung cancer in Chinese Han population. We also performed stratification analysis of lung cancer to evaluate the relationship between *CMTM8* polymorphism and different stratification indexes of lung cancer.

Materials And Methods

Study participants

The present hospital-based case control study included a total of 509 lung cancer cases and 506 healthy controls. Lung cancer subjects were recruited from the Tumor Hospital of Shaanxi province, China. Controls were taken from the people who visited the hospital for routine check-up with no history of cancer and any diseases associated with vital organs. All lung cancer cases were newly diagnosed and histopathological confirmed.

Clinical data and demographic information

We use a standardized epidemiological questionnaire including residential region, age, gender, smoking status, alcohol use, ethnicity, education status, and family history of cancer to collect personal data in an in-person interview. 5 ml of venous blood sample was drawn from each subject and used for DNA extraction and genotyping. All volunteers signed an informed consent form explaining the research purpose of the blood withdrawal.

SNPs selection and genotyping

We screened the SNPs of *CMTM8* with minor allele frequencies > 5% in 1,000 genome project (http://www.internationalgenome.org/). In addition, the genotype distributions of the SNPs in control group were in accordance with Hardy-Weinberg equilibrium (HWE) (p > 0.05). We used Agena MassARRAY RS1000 to genotype and we retained the SNPs with call rate over 95.0%. Then, the Haploview software package (version 4.2) was used to estimate pairwise linkage disequilibrium (LD) at *CMTM8* polymorphism loci. When r^2 (the measure value of LD) > 0.8, the SNP can represent all the polymorphisms

in a block. According to the above selection criteria, we selected the six SNPs (rs9853415, rs6796318, rs6771238, rs9835916, rs1077868 and rs6802418) in *CMTM8* as the gene variation to study. Genomic DNA was isolated from peripheral whole blood employing the Gold Mag - Mini Whole Blood Genomic DNA Purification Kit (Gold Mag Co. Ltd., Xi'an, China) following the manufacturer's instructions and quantified by Nano Drop spectrophotometer 2000C (Thermo Scientific, Waltham, Massachusetts, USA). Polymerase chain reaction (PCR) extension primers were designed for these SNPs by MassARRAY Assay Design 3.0 software (Agena). Primers in this study were listed in **Supplementary Table S1**. SNPs genotyping analysis was performed on Agena MassARRAY RS1000 instrument (Shanghai, China) system according to the standard scheme recommended by the manufacturer, and data were managed and analyzed by Agena Typer 4.0 software [12, 13].

Statistical analyses

Hardy-Weinberg equilibrium (HWE) of each SNP in control group was tested by Fisher s exact test [5]. Allele frequencies and genotype frequencies for each SNP of case and control subjects were compared using the Chi squared test. Odds ratios (ORs) and 95% confidence intervals (CIs) were determined using unconditional logistic regression analysis with adjustments for age, gender, smoking, and drinking [14–16]. Associations between genotypes and lung cancer risk were tested in different genetic models (co-dominant, dominant, recessive, and log-additive) by SNPStats website software (http://bioinfo.iconcologia.net/snpstats/start.htm) [17, 18]. All statistical analysis were performed using SPSS statistical package, version 19.0 (SPSS Inc., Chicago, IL, USA). Haploview software version 4.2 was used to analyze the association between haplotypes and the lung cancer [19]. The power of the significant difference was calculated by Power and Sample Size (PS) Calculation software (http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize) [20]. All *p* values in this study were two-sided, and *p*-value of less than 0.05 as the cutoff value for statistical significance.

SNP functional annotation and Gene expression analysis

HaploReg v4.1 database (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) was applied for exploring functional annotations of the candidate SNPs. Through GEPIA database (http://gepia.cancer-pku.cn/index.html) to predict *CMTM8* gene expression in lung tissue.

Results

General characteristics

Basic characteristics of the case and control groups were depicted in Table 1. This study involved 1015 subjects, including 509 patients (354 males and 155 females; age at diagnosis: 58.53 ± 10.12 years) and 506 healthy controls (355 males and 151 females; age: 61.43 ± 9.47 years). There were no significant differences in terms of age, sex or smoking status between lung cancer patients and healthy controls, but there were significant differences in alcohol consumption.

Table 1
Demographic characteristics of cases and controls in this study

Variable (s)	Case, n (%)	Control, n (%)	*p* - value
	(n = 509)	(n = 506)	
Age, N (%)			
Age, year (mean ± SD)	58.53 ± 10.12	61.43 ± 9.47	0.477
≤ 60	236 (45.0%)	225 (50.6%)	
> 60	273 (55.0%)	281 (49.4%)	
Gender, N (%)			0.838
Male	354 (55.6%)	355 (55.7%)	
Female	155 (44.4%)	151 (44.3%)	
Smoking			0.168
Yes	252 (49.5%)	206 (40.7%)	
No	251 (49.3%)	268 (53.0%)	
Unavailable	6 (1.2%)	32 (6.3%)	
Drinking			< 0.05
Yes	114 (22.3%)	209 (20.1%)	
No	358 (70.3%)	270 (19.8%)	
Unavailable	37 (7.2%)	27 (60.0%)	
BMI (kg/m2)			
< 24	321 (63.0%)	140 (27.7%)	
≥ 24	178 (35.0%)	117 (23.1%)	
Unavailable	10 (1.9%)	249 (49.2%)	
BMI, body mass index;			
P < 0.05 indicates statistical significance.			

Hardy-Weinberg equilibrium and SNPs alleles

The MAF distribution of selected six SNPs among all subjects was summarized in Table 2. In our study, the allele frequency of each SNP in controls was consistent with the CHB population (Han Chinese in Beijing, China) in the 1,000 genome project. Furthermore, all six SNP locus in the control subjects conformed to Hardy-Weinberg equilibrium ($p > 0.05$). By chi-square test, we found no SNPs sites associated with lung cancer risk.

Table 2
Basic characteristics and allele frequencies among SNPs

SNP	Chr	Position	Gene(s)	Role	Alleles	Frequency (MAF)		*p* - HWE	OR (95% CI)	*p*-value
						Cases	Controls			
rs9853415	3	32249089	*CMTM8*	intronic	A/G	0.480	0.478	0.790	1.01 (0.85–1.20)	0.925
rs6796318	3	32260294	*CMTM8*	intronic	A/G	0.090	0.082	0.764	1.11 (0.81–1.51)	0.511
rs6771238	3	32289870	*CMTM8*	intronic	A/C	0.059	0.041	0.192	1.48 (0.98–2.23)	0.056
rs9835916	3	32302348	*CMTM8*	intronic	C/T	0.430	0.400	0.397	1.13 (0.95–1.35)	0.173
rs1077868	3	32347212	*CMTM8*	intronic	C/T	0.173	0.166	0.520	1.05 (0.83–1.32)	0.680
rs6802418	3	32355312	*CMTM8*	intronic	A/G	0.324	0.303	0.600	1.10 (0.91–1.33)	0.313
SNP: Single nucleotide polymorphism; HWE: Hardy-Weinberg equilibrium; OR: Odds ratio; 95% CI: 95% confidence interval;										
p - HWE obtained from Fisher's exact test;										
p - value obtained from Wald test;										
* *p* - value < 0.05 indicates statistical significance.										

Association of SNPs with lung cancer risk

Four genetic analysis models (co-dominant, dominant, recessive and log-additive) were applied to analyze and assess the association between each variant and lung cancer risks. In Table 3, our analysis revealed the genotype "A/C" of rs6771238 was correlated with an increased the risk of lung cancer under the co-dominant model (OR = 1.57, 95% CI = 1.01–2.42, $p = 0.044$), the genotype "C/A-A/A" of rs6771238 was correlated with an enhanced lung cancer risk in the dominant mode (OR = 1.68, 95% CI = 1.05–2.68, $p = 0.031$), with power values of 0.534 and 0.684, respectively. Rs6771238 also reduced lung cancer risk in an additive model (OR = 1.66, 95% CI = 1.06–2.58, $p = 0.026$), with power values of 0.500.

Table 3
Relationship between SNPs in *CMTM8* gene and the risk of lung cancer

SNP	Model	Genotype	Control	Case	Without Adjustment		With Adjustment		Study power
					OR (95% CI)	p^a-value	OR (95% CI)	p^b-value	
rs9853415	Codominant	A/A	136 (26.9%)	128 (25.1%)	1		1		
		A/G	256 (50.6%)	273 (53.6%)	1.13 (0.84–1.52)	0.407	1.16(0.84–1.59)	0.369	
		G/G	114 (22.5%)	108 (21.2%)	1.01 (0.7–1.44)	0.971	1.13(0.77–1.65)	0.545	
	Dominant	A/A	136 (26.9%)	128 (25.1%)	1		1		
		A/G-G/G	370 (73.1%)	381 (74.9%)	1.09 (0.83–1.45)	0.53	1.15(0.85–1.55)	0.371	
	Recessive	A/A-A/G	392 (77.5%)	401 (78.8%)	1		1		
		G/G	114 (22.5%)	108 (21.2%)	0.93 (0.69–1.25)	0.613	1.02(0.74–1.41)	0.894	
	Log-additive	—	—	—	1.01 (0.84–1.21)	0.923	1.07(0.88–1.29)	0.518	
rs6796318	Co-dominant	A/A	426 (84.4%)	423 (83.1%)	1		1		
		A/G	75 (14.9%)	80 (15.7%)	1.07 (0.76–1.51)	0.682	1.06(0.73–1.53)	0.753	
		G/G	4 (0.8%)	6 (1.2%)	1.51 (0.42–5.39)	0.525	1.90(0.46–7.8)	0.372	
	Dominant	A/A	426 (84.4%)	423 (83.1%)	1		1		
		A/G-G/G	79 (15.6%)	86 (16.9%)	1.1 (0.79–1.53)	0.589	1.1(0.77–1.57)	0.61	
	Recessive	A/A-A/G	501 (99.2%)	503 (98.8%)	1		1		
		G/G	4 (0.8%)	6 (1.2%)	1.49 (0.42–5.33)	0.536	1.89(0.46–7.72)	0.378	
	Log-additive	—	—	—	1.11 (0.81–1.5)	0.518	1.12(0.81–1.56)	0.492	
rs6771238	Co-dominant	C/C	467 (92.3%)	451 (88.6%)	1		1		
		C/A	37 (7.3%)	56 (11.0%)	**1.57 (1.01–2.42)**	**0.043***	**1.57 (1.01–2.42)**	**0.044***	**0.534**
		A/A	2 (0.4%)	2 (0.4%)	1.04 (0.15–7.38)	0.972	3.29(0.29–36.87)	0.333	
	Dominant	C/C	467 (92.3%)	451 (88.6%)	1		1		
		C/A-A/A	39 (7.7%)	58 (11.4%)	**1.54 (1.01–2.36)**	**0.047***	**1.68(1.05–2.68)**	**0.031***	**0.684**
	Recessive	C/C-C/A	504 (99.6%)	507 (99.6%)	1		1		
		A/A	2 (0.4%)	2 (0.4%)	0.99 (0.14–7.08)	0.995	3.16(0.28–35.38)	0.35	

SNP: Single nucleotide polymorphism; OR: Odds ratio; 95% CI: 95% confidence interval;

p^a-value calculated by logistic regression analysis;

p^b-value calculated by logistic regression analysis with adjustments for gender, age, smoking, and drinking);

*p-value < 0.05 indicates statistical significance.

SNP	Model	Genotype	Control	Case	Without Adjustment		With Adjustment		Study power
					OR (95% CI)	p^a-value	OR (95% CI)	p^b-value	
	Log-additive	—	—	—	1.47 (0.98–2.19)	0.062	**1.66(1.06–2.58)**	**0.026***	**0.500**
rs9835916	Co-dominant	T/T	181 (36.9%)	161 (32.1%)	1		1		
		T/C	226 (46.1%)	249 (49.7%)	1.24 (0.94–1.64)	0.132	1.19(0.88–1.62)	0.254	
		C/C	83 (16.9%)	91 (18.2%)	1.23 (0.86–1.78)	0.262	1.07(0.72–1.59)	0.722	
	Dominant	T/T	181 (36.9%)	161 (32.1%)	1		1		
		T/C-C/C	309 (63.1%)	340 (67.9%)	1.24 (0.95–1.61)	0.112	1.16(0.87–1.54)	0.311	
	Recessive	T/T-T/C	407 (83.1%)	410 (81.8%)	1		1		
		C/C	83 (16.9%)	91 (18.2%)	1.09 (0.78–1.51)	0.612	0.97(0.68–1.38)	0.858	
	Log-additive	—	—	—	1.13 (0.95–1.35)	0.176	1.06(0.87–1.28)	0.559	
rs1077868	Co-dominant	A/A	354 (70%)	344 (67.6%)	1		1		
		A/G	136 (26.9%)	154 (30.3%)	1.17 (0.89–1.53)	0.274	0.66(0.29–1.48)	0.315	
		G/G	16 (3.2%)	11 (2.2%)	0.71 (0.32–1.55)	0.386	1.24(0.92–1.67)	0.154	
	Dominant	A/A	354 (70%)	344 (67.6%)	1		1		
		A/G-G/G	152 (30.0%)	165 (32.4%)	1.12 (0.86–1.46)	0.414	1.17(0.88–1.56)	0.28	
	Recessive	A/A-A/G	490 (96.8%)	498 (97.8%)	1		1		
		G/G	16 (3.2%)	11 (2.2%)	0.68 (0.31–1.47)	0.325	0.62(0.28–1.38)	0.243	
	Log-additive	—	—	—	1.05 (0.83–1.33)	0.677	1.08(0.84–1.38)	0.568	
rs6802418	Co-dominant	G/G	248 (49%)	230 (45.2%)	1		1		
		G/A	209 (41.3%)	228 (44.8%)	1.18 (0.91–1.53)	0.22	1.15(0.87–1.53)	0.32	
		A/A	49 (9.7%)	51 (10%)	1.12 (0.73–1.73)	0.6	1.06(0.66–1.68)	0.822	
	Dominant	G/G	248 (49%)	230 (45.2%)	1		1		
		G/A-A/A	258 (51.0%)	279 (54.8%)	1.17 (0.91–1.49)	0.222	1.13(0.87–1.48)	0.355	
	Recessive	G/G-G/T	457 (90.3%)	458 (90.0%)	1		1		

SNP: Single nucleotide polymorphism; OR: Odds ratio; 95% CI: 95% confidence interval;

p^a-value calculated by logistic regression analysis;

p^b-value calculated by logistic regression analysis with adjustments for gender, age, smoking, and drinking);

*p-value < 0.05 indicates statistical significance.

SNP	Model	Genotype	Control	Case	Without Adjustment		With Adjustment		Study power
					OR (95% CI)	p^a-value	OR (95% CI)	p^b-value	
		A/A	49 (9.7%)	51 (10.0%)	1.04 (0.69–1.57)	0.858	0.99(0.63–1.55)	0.954	
	Log-additive	—	—	—	1.10 (0.91–1.33)	0.313	1.07(0.88–1.31)	0.497	
SNP: Single nucleotide polymorphism; OR: Odds ratio; 95% CI: 95% confidence interval;									
p^a-value calculated by logistic regression analysis;									
p^b-value calculated by logistic regression analysis with adjustments for gender, age, smoking, and drinking);									
*p-value < 0.05 indicates statistical significance.									

Further, we stratified the samples according to pathological classification, clinical stage, lymph node metastasis and other characteristics. Within the subgroups of lung squamous cell carcinoma and lung adenocarcinoma, the allele "A" of rs6771238 (OR = 1.90, 95% CI = 1.07–3.38, p = 0.025, power = 0.642) showed an increased risk of lung squamous cell carcinoma in the allele model. Rs6771238 also was significantly associated with an increased lung squamous cell carcinoma risk under the log addition model (OR = 1.98, 95% CI = 1.01–3.87, p = 0.045, power = 0.704). Rs6771238 also was significantly correlated with an enhanced adenocarcinoma risk under log-additive model (OR = 1.79, 95% CI = 1.01–3.18, p = 0.047, power = 0.549) (Table 4).

Table 4
Stratified analysis of *CMTM8* polymorphism and lung squamous cell carcinoma and lung adenocarcinoma risk in lung cancer case types

SNP	Model	Genotype	Squamous cell carcinoma					Adenocarcinoma				
			Control	Case	OR (95% CI)	p^a-value	Study power	Control	Case	OR (95% CI)	p^b-value	Study power
rs6771238	Alleles	C	971 (95.9%)	224 (92.6%)	1			971 (95.9%)	363 (94.0%)	1		
		A	41 (4.1%)	18 (7.4%)	**1.90(1.07–3.38)**	**0.025***	**0.642**	41 (4.1%)	23 (6.0%)	1.50 (0.89–2.54)	0.151	
	Co-dominant	C/C	467 (92.3%)	104 (86.0%)	1			467 (92.3%)	171 (88.6%)	1		
		A/C	37 (7.3%)	16 (13.2%)	1.95(0.93–4.12)	0.078		37 (7.3%)	21 (10.9%)	1.66(0.89–3.08)	0.109	
		A/A	2 (0.4%)	1 (0.8%)	4.35(0.24–77.44)	0.317		2 (0.4%)	1 (0.5%)	7.54(0.45–125.1)	0.159	
	Dominant	C/C	467 (92.3%)	104 (86.0%)	1			467 (92.3%)	171 (88.6%)	1		
		A/C-A/A	39 (7.7%)	17 (14.0%)	2.04(0.99–4.21)	0.054		39 (7.7%)	22 (11.4%)	1.75(0.95–3.22)	0.07	
	Recessive	C/C-A/C	504 (99.6%)	120 (99.2%)	1			504 (99.6%)	192 (99.5%)	1		
		A/A	2 (0.4%)	1 (0.8%)	4.06(0.23–72.03)	0.34		2 (0.4%)	1 (0.5%)	7.24(0.44–120)	0.167	
	Log-addition	—	—	—	**1.98(1.01–3.87)**	**0.045***	**0.704**	—	—	**1.79(1.01–3.18)**	**0.047***	**0.549**
SNP: Single nucleotide polymorphism; OR: Odds ratio; 95% CI: 95% confidence interval;												
p^a-value calculated by logistic regression analysis;												
p^b-value calculated by logistic regression analysis with adjustments for gender, age, smoking, and drinking);												
*p-value < 0.05 indicates statistical significance.												

Stratified analysis based on clinical staging showed that "A/G" and "A/G-G/G" genotypes of rs1077868 were significantly correlated with lung cancer staging in codominant (OR = 1.96, 95% CI = 1.05–3.64, p = 0.034, power = 0.998) and dominant (OR = 2.03, 95% CI = 1.11–3.73, p = 0.022, power = 0.993)

models, respectively. Rs1077868 was also significantly correlated with lung cancer staging in the additive model (OR = 1.92, 95% CI = 1.10–3.35, *p* = 0.021, power = 0.970) (Table 5).

Table 5
Stratified analysis of *CMTM8* polymorphism and staging risk in patients with lung cancer

SNP	Model	Genotype	Control	Case	OR (95% CI)	*p*-value	Study power
rs1077868	Alleles	A	147 (87.5%)	423 (81.0%)	1		
		G	21 (12.5%)	99 (19.0%)	1.64 (0.99–2.72)	0.054	
	Co-dominant	A/A	64 (76.2%)	170 (64.9%)	1		
		A/G	19 (22.6%)	83 (31.7%)	**1.96(1.05–3.64)**	**0.034***	**0.998**
		G/G	1 (1.2%)	8 (3.1%)	3.28(0.39–27.22)	0.272	
	Dominant	A/A	64 (76.2%)	170 (64.9%)	1		
		A/G-G/G	20 (23.8%)	91 (34.8%)	**2.03(1.11–3.73)**	**0.022***	**0.993**
	Recessive	A/A-A/G	83 (98.8%)	253 (96.6%)	1	0.348	
		G/G	1 (1.2%)	8 (3.1%)	2.69(0.33–22.16)	0.358	
	Log-addition	—	—	—	**1.92(1.10–3.35)**	**0.021***	**0.970**
ORs, odds ratios; CI, confidence interval;							
p-value calculated by logistic regression analysis with adjustments for gender and age, smoking, and drinking);							
*p value < 0.05 indicates statistical significance.							

Stratified subgroup in the case of lymph node metastasis, rs9835916 was found to be associated with lymph node metastasis risk in patients with lung cancer. For rs9835916, allele “C” increased the risk of lymphatic metastasis based on the allele model (OR = 1.56, 95% CI = 1.08–2.26, *p* = 0.018, power = 0.940), the "T/C" genotype increased the risk of lymphatic metastasis under the co-dominant model (OR = 2.49, 95% CI = 1.36–4.55, *p* = 0.003, power = 0.956), the "T/C-C/C" genotype was related to an increased lymphatic metastasis risk in the dominant model (OR = 2.40, 95% CI = 1.37–4.21, *p* = 0.002, power = 0.998). Rs9835916 also was significantly associated with an increased the risk of lymphatic metastasis based on the additive model (OR = 1.66, 95% CI = 1.11–2.48, *p* = 0.014, power = 0.978) (Table 6).

Table 6
Stratified analysis of the polymorphisms of *CMTM8* and the risk of lymph node metastasis in patients with lung cancer

SNP	Model	Genotype	Control	Case	OR (95% CI)	*p*-value	Study power
rs9835916	Alleles	T	103 (61.3%)	293 (57.2%)	1		
		C	65 (38.7%)	219 (42.8%)	**1.56 (1.08–2.26)**	**0.018***	**0.940**
	Co-dominant	T/T	38 (45.2%)	56 (26.3%)	1		
		T/C	33 (39.3%)	119 (55.9%)	**2.49(1.36–4.55)**	**0.003***	**0.956**
		C/C	13 (15.5%)	38 (17.8%)	2.18(0.99–4.83)	0.054	
	Dominant	T/T	38 (45.2%)	56 (26.3%)	1		
		T/C-C/C	46 (54.8%)	157 (73.7%)	**2.40(1.37–4.21)**	**0.002***	**0.998**
	Recessive	T/T-C/T	71 (84.5%)	175 (82.2%)	1		
		C/C	13 (15.5%)	38 (17.8%)	1.3(0.63–2.69)	0.473	
	Log-additive	—	—	—	**1.66(1.11–2.48)**	**0.014***	**0.978**
ORs, odds ratios; CI, confidence interval;							
p-value calculated by logistic regression analysis with adjustments for gender and age, smoking, and drinking);							
*p value < 0.05 indicates statistical significance.							

Association of haplotypes with lung cancer risk

A haplotype-based association study was performed to show the association between *CMTM8* haplotype and risk of lung cancer. Among the subpopulation (staging), two SNPs (rs1077868 and rs6802418) form an LD block (Fig. 1). The frequencies’ distribution of haplotypes in case and control

group is presented in Table 7. The haplotypes "GG" and “AG” was found to prominently increase the risk of lung cancer staging (OR = 1.71; 95% CI = 1.02–2.88; $p = 0.043$).

Table 7
Four SNPs in *CMTM8* haplotypes and their associations with lung cancer risk

Variable	SNP	Haplotype	Freq (case)	Freq (control)	Without adjusted		With adjusted	
					OR (95% CI)	p^a-value	OR (95% CI)	p^b-value
Overall	rs9853415|rs6796318	GG	0.087	0.082	1.07(0.78–1.46)	0.675	1.09(0.78–1.52)	0.631
	rs9853415|rs6796318	GA	0.393	0.397	0.98(0.82–1.18)	0.844	1.04(0.85–1.27)	0.726
	rs9853415|rs6796318	AA	0.483	0.479	1.02(0.85–1.22)	0.85	1.08(0.89–1.3)	0.462
	rs1077868|rs6802418	AA	0.324	0.302	1.11(0.92–1.34)	0.289	1.08(0.88–1.32)	0.471
	rs1077868|rs6802418	GG	0.173	0.165	1.06(0.84–1.34)	0.634	1.08(0.84–1.39)	0.534
	rs1077868|rs6802418	AG	0.497	0.468	1.12(0.94–1.33)	0.198	1.11(0.92–1.34)	0.275
OR: odds ratio, 95% CI: 95% confidence interval;								
p^a-value calculated by Wald test;								
p^b-value calculated by Wald test and adjusted by age, gender, smoking, and drinking;								
*p-value < 0.05 indicates statistical significance.								

SNP functional evaluation

In order to evaluate the possible function of the six selected variants in the *CMTM8* gene, we performed a bioinformatics analysis using the HaploReg v4.1 database. The results showed that all the variants were predicted to be regulatory SNPs with different biological functions (**Supplementary table S2**).

GEPIA database analysis on gene expression

Furthermore, GEPIA database analyzed the expression of *CMTM8* gene in lung cancer and found that the expression level of *CMTM8* gene in lung adenocarcinoma was lower than that in normal tissues, which indicates that this gene has a certain relationship with the occurrence of lung cancer (**Supplementary Figure S1**)

Discussion

In recent years, a growing number of studies have found that the *CMTM8* gene plays an important role in the tumor s formation, development and metastasis, and the expression of *CMTM8* is down-regulated in lung cancer. In this study, we genotyped six polymorphisms of *CMTM8* and evaluated their correlations with the risk of lung cancer in a Chinese Han population. Our results firstly showed that rs6771238 was associated with an increased lung cancer susceptibility in Chinese Han Population. Stratified analysis showed that rs6771238 was related to an increased the risk of lung squamous cell carcinoma, rs6771238 was associated with an increased the risk of lung adenocarcinoma, rs9835916 and rs1077868 were correlated with lung cancer staging, and rs9835916 was correlated with an increased the risk of lymph node metastasis in lung cancer patients. Haplotype analysis illuminated that “GG” and “AG” were closely correlated with lung cancer staging, and “AG” was correlated with an increased lung cancer risk among individuals older than 50 years. To our knowledge, this is the first study that to explore the association between *CMTM8* gene polymorphism and lung cancer risk in Chinese Han Population.

Human *CMTM8* localizes to chromosome 3p22.3, where other known tumor suppressor genes that are frequently deleted or methylated in tumors reside [21, 22]. *CMTM8* may be silenced or down-regulated in a similar manner during tumorigenesis. Previous studies demonstrated that *CMTM8* induces caspase-dependent and caspase-independent apoptosis in multiple tumor cell lines [4]. Downregulation of *CMTM8* in epithelial cells induces epithelial-mesenchymal transformation (EMT) through MEK-ERK signaling [7]. Overexpression of *CMTM8* can accelerate the rate of epidermal growth factor receptor internalization, attenuates epidermal growth factor receptor mediated signaling, and inhibits tumor cell growth [6]. At present, indications for tumor suppressive function of *CMTM8* gene products have been found in several tumor types. In osteosarcoma, it was confirmed that *CMTM8* was identified as a candidate tumor suppressor gene, which inhibited the EGFR signaling pathway and affected the occurrence of osteosarcoma [6, 23]. *CMTM8* underexpression may result in upregulation of EGFR signaling. In bladder cancer, *CMTM8* is also an important tumor suppressor gene and a useful prognostic indicator for patients with bladder cancer [9, 24]. It is inferred that *CMTM8* overexpression blocks c-MET signaling in vivo model of bladder cancer. Research have also demonstrated that the downregulation of *CMTM8* induced epithelial-to-mesenchymal transition-like changes via c-MET/extracellular signal-regulated kinase (ERK) signaling in HepG2 hepatocellular carcinoma cells (14), thereby affecting the cancer process [7]. In addition, *CMTM8* was negatively correlated with the tumorigenesis and development of clear-cell renal cell carcinoma, and the location and intensity of expression were significantly correlated with prognosis [25].

However, the expression of *CMTM8* in lung cancer is only known to be down-regulated, and other relevant reports are relatively few. In present study, we investigated the relationship between *CTMT8* and lung cancer susceptibility for the first time. Our results suggested that rs6771238 was associated with an increased lung cancer susceptibility in Chinese Han population. Stratified analysis showed that rs6771238 was associated with the risk of lung squamous cell carcinoma, rs6771238 was associated with an increased the risk of lung adenocarcinoma, rs9835916 and rs1077868 were associated with lung cancer stage, and rs9835916 was associated with lymph node risk in lung cancer patients. In addition, it should be noted that the incidence of lung cancer significantly different according to the different stratification. Considering the potential function of the selected SNPs in our study, we speculated that SNPs may affect the carcinogenic process by changing the protein expression and this process may be influenced by individual background, thus leading to different outcomes on lung cancer risk.

Nevertheless, there are limitations that need to be noticed in the present study. First, because subjects are enrolled from the same hospital, inherent selection bias and information bias are inevitable problems. Second, our current research is fundamental, and further functional studies and larger population based prospective studies are required to illuminate the genetic factors underlying lung cancer. Despite the limitations mentioned above, our current findings provide scientific evidence for future studies of gene *CMTM8* with the risk of lung cancer.

Conclusion

To sum up, our study revealed a novel association between *CMTM8* polymorphisms and the risk of lung cancer among Chinese Han population. These results may help elucidate the underlying mechanisms for *CMTM8* polymorphisms in lung cancer. Larger sample sizes and mechanism studies are necessary to further explore and confirm the role of these variants in increasing lung cancer risk, which will help in better understanding the genetic heterogeneity in complex diseases like lung cancer.

Abbreviations

CMTM8
CKLF-like MARVEL transmembrane domain containing 8; SNP:Single nucleotide polymorphism; HWE:Hardy-Weinberg equilibrium; OR:Odds ratio; 95% CI:95% confidence intervals; LD:Linkage disequilibrium; MAF:Minor allele frequency.

Declarations

Ethics approval and consent to participate

The use of human blood sample and the protocol in this study were strictly comply with the criterions of the Declaration of Helsinki and were approved by the Ethics Committee of the Northwest University, Xi'an, China. Written informed consent was received from each participant.

Consent to publish

Not applicable.

Availability of data and materials

The datasets used/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable.

Authors' Contributions

TB J and JM W conceived and designed the project.

FL N and YW Lrecruited and collected study samples.

HY Land JF L selected the SNPs and designed primers.

Y S, JM W and QF L performed the experiments.

Y S and ZC X performed and analyzed the data.

JM W wrote and revised the manuscript.

Acknowledgments

We thank all the patients and individuals for their participation, as well as the clinicians and other hospital staff of the Tumor Hospital of Shaanxi province. Meanwhile,we are very grateful to the editors and reviewers for theirpatience and valuable comments to this work.

References

1. Bilello KS, Murin S, Matthay RA: **Epidemiology, etiology, and prevention of lung cancer.***Clin Chest Med* 2002, **23:**1-25.
2. Lin H, Ning B, Li J, Ho SC, Huss A, Vermeulen R, Tian L: **Lung cancer mortality among women in Xuan Wei, China: a comparison of spatial clustering detection methods.***Asia Pac J Public Health* 2015, **27:**Np392-401.
3. Hu QY, Jin TB, Wang L, Zhang L, Geng T, Liang G, Kang LL: **Genetic variation in the TP63 gene is associated with lung cancer risk in the Han population.***Tumour Biol* 2014, **35:**1863-1866.
4. Jin C, Wang Y, Han W, Zhang Y, He Q, Li D, Yin C, Tian L, Liu D, Song Q, Ma D: **CMTM8 induces caspase-dependent and -independent apoptosis through a mitochondria-mediated pathway.***J Cell Physiol* 2007, **211:**112-120.
5. Gao DH, Hu H, Fang ZW, Huo F, Wang HR, Xu KX, Wang XF: **Research Advances in Chemokine-like Factor Super Family Member 8.***Zhongguo Yi Xue Ke Xue Yuan Xue Bao* 2016, **38:**746-749.
6. Jin C, Ding P, Wang Y, Ma D: **Regulation of EGF receptor signaling by the MARVEL domain-containing protein CKLFSF8.***FEBS Lett* 2005, **579:**6375-6382.
7. Zhang W, Mendoza MC, Pei X, Ilter D, Mahoney SJ, Zhang Y, Ma D, Blenis J, Wang Y: **Down-regulation of CMTM8 induces epithelial-to-mesenchymal transition-like changes via c-MET/extracellular signal-regulated kinase (ERK) signaling.***J Biol Chem* 2012, **287:**11850-11858.
8. Zhang W, Qi H, Mo X, Sun Q, Li T, Song Q, Xu K, Hu H, Ma D, Wang Y: **CMTM8 is Frequently Downregulated in Multiple Solid Tumors.***Appl Immunohistochem Mol Morphol* 2017, **25:**122-128.
9. Gao D, Hu H, Wang Y, Yu W, Zhou J, Wang X, Wang W, Zhou C, Xu K: **CMTM8 inhibits the carcinogenesis and progression of bladder cancer.***Oncol Rep* 2015, **34:**2853-2863.
10. Jin G, Ma H, Wu C, Dai J, Zhang R, Shi Y, Lu J, Miao X, Wang M, Zhou Y, et al: **Genetic variants at 6p21.1 and 7p15.3 are associated with risk of multiple cancers in Han Chinese.***Am J Hum Genet* 2012, **91:**928-934.
11. Dong J, Jin G, Wu C, Guo H, Zhou B, Lv J, Lu D, Shi Y, Shu Y, Xu L, et al: **Genome-wide association study identifies a novel susceptibility locus at 12q23.1 for lung squamous cell carcinoma in han chinese.***PLoS Genet* 2013, **9:**e1003190.
12. Chen J, Liu W, Cao Y, Zhang X, Guo Y, Zhu Y, Li J, Du J, Jin T, Wang G, Wang J: **MMP-3 and MMP-8 single-nucleotide polymorphisms are related to alcohol-induced osteonecrosis of the femoral head in Chinese males.***Oncotarget* 2017, **8:**25177-25188.
13. Gabriel S, Ziaugra L, Tabbaa D: **SNP genotyping using the Sequenom MassARRAY iPLEX platform.***Curr Protoc Hum Genet* 2009, **Chapter 2:**Unit 2.12.
14. Wang T, Chen T, Thakur A, Liang Y, Gao L, Zhang S, Tian Y, Jin T, Liu JJ, Chen M: **Association of PSMA4 polymorphisms with lung cancer susceptibility and response to cisplatin-based chemotherapy in a Chinese Han population.***Clin Transl Oncol* 2015, **17:**564-569.
15. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT: **Data quality control in genetic case-control association studies.***Nat Protoc* 2010, **5:**1564-1573.
16. Zhang S, Thakur A, Liang Y, Wang T: **Polymorphisms in C-reactive protein and Glypican-5 are associated with lung cancer risk and Gartrokine-1 influences Cisplatin-based chemotherapy response in a Chinese Han population.** 2015, **2015:**824304.
17. Gao L, Thakur A, Liang Y, Zhang S, Wang T, Chen T, Meng J, Wang L, Wu F, Jin T, et al: **Polymorphisms in the TERT gene are associated with lung cancer risk in the Chinese Han population.***Eur J Cancer Prev* 2014, **23:**497-501.
18. Sole X, Guino E, Valls J, Iniesta R, V: **SNPStats: a web tool for the analysis of association studies.***Bioinformatics* 2006, **22:**1928-1929.
19. Barrett JC, Fry B, Maller J, Daly MJ: **Haploview: analysis and visualization of LD and haplotype maps.***Bioinformatics* 2005, **21:**263-265.
20. Dupont WD, Plummer WD, Jr.: **Power and sample size calculations for studies involving linear regression.***Control Clin Trials* 1998, **19:**589-601.
21. Imreh S, Klein G, Zabarovsky ER: **Search for unknown tumor-antagonizing genes.***Genes Chromosomes Cancer* 2003, **38:**307-321.
22. Ying J, Poon F, J, Geng H, Wong A, Qiu G, Goh H, Rha S, Tian L, Chan A, Sung J: **DLEC1 is a functional 3p22.3 tumour suppressor silenced by promoter CpG methylation in colon and gastric cancers.***British Journal of Cancer* 2009, **100:**663-669.
23. Both J, Krijgsman O, Bras J, Schaap GR, Baas F, Ylstra B, Hulsebos TJ: **Focal chromosomal copy number aberrations identify CMTM8 and GPR177 as new candidate driver genes in osteosarcoma.***PLoS One* 2014, **9:**e115835.
24. Zhang S, Pei X, Hu H, Zhang W, Mo X, Song Q, Zhang Y, Xu K, Wang Y, Na Y: **Functional characterization of the tumor suppressor CMTM8 and its association with prognosis in bladder cancer.***Tumour Biol* 2016, **37:**6217-6225.
25. Hu H, Chen JW, Xu KX, Wang D, Wang Y, Wang GW, Zhang SY, Wang XF: **[Expressions of CMTM8 and E-cadherin in primary and metastatic clear cell renal cell carcinoma].***Beijing Da Xue Xue Bao Yi Xue Ban* 2013, **45:**537-541.

Figures

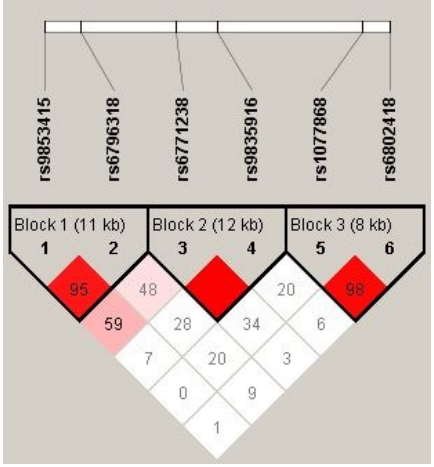

Figure 1

Haplotype block map for SNPs in the CMTM8 gene. D value has been rendered in the lattice.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfiles.docx