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
xi'an Jiaotong University
tianbo jin (jintianbo66@163.com)
Northwest University https://orcid.org/0000-0003-0037-7900

Research article

Keywords: Polymorphisms, CMTM8, Lung cancer, Chinese Han population

Posted Date: October 28th, 2019

DOI: https://doi.org/10.21203/rs.2.12045/v2

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

Abstract Background: 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 were 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 ($p < 0.05$). In stratified analysis, rs6771238 was related to an increased risk of lung squamous cell carcinoma ($p < 0.05$), rs6771238 was associated with an increased risk of lung adenocarcinoma ($p < 0.05$), rs9835916 and rs1077868 were correlated with lung cancer staging ($p < 0.05$), and rs9835916 was correlated with an increased risk of lymph node metastasis in lung cancer patients ($p < 0.05$). Additionally, Haplotype analysis illuminated that haplotypes GG and AG were closely correlated with lung cancer staging, and haplotype AG was correlated with an increased lung cancer risk among individuals older than 50 years ($p < 0.05$). Conclusions: Our study firstiy reported that the CMTM8 polymorphisms were risk factors for lung cancer in Chinese Han population. These findings also suggested the potential roles of CMTM8 in the development of lung cancer.

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 [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*. With respect to genetic variation, SNPs have been utilized in lung cancer research [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 histopathologically 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 and gender [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]. 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 was 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 difference between the lung cancer patients and healthy controls in terms of age, gender and smoking status.

Hardy-Weinberg equilibrium and SNPs alleles

The MAF distribution of selected six SNPs among all subjects were summarized in **Table 2**. In our study, the frequency of alleles of each SNP in controls was consistent with the HapMap CHB population. 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.

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 = 1.54, 95% CI = 1.01-2.36, p = 0.047).

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 genotype “A/C” of rs6771238 (OR = 2.07, 95% CI = 1.08-3.97, p = 0.028) showed an increased risk of lung squamous cell carcinoma in the co-dominant model. The genotype “A/C-A/A” of rs6771238 also was significantly associated with an increased lung squamous cell carcinoma risk under the dominant model (OR = 2.07; 95% CI = 1.10-3.89; p = 0.025). Rs6771238 also was significantly correlated with an enhanced lung squamous cell carcinoma risk under log-additive model (OR = 1.90; 95% CI = 1.06 - 3.38; p = 0.030). While the genotype “A/G” of rs6802418 (OR = 1.43, 95% CI = 1.00-2.03, p = 0.049) may increase the risk of lung adenocarcinoma under the co-dominant model (**Table 4**).

Stratified analysis was performed according to clinical stages, it was found the genotype “T/C” of rs9835916 and rs1077868 were significantly correlated with an enhanced the risk of lung cancer staging under the co-dominant model (OR = 1.75; 95 % CI = 1.00-3.05; p = 0.031) and log-additive model (OR = 1.71; 95 % CI = 1.02-2.88; p = 0.043), respectively (**Table 5**).

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” also may increase the risk of lymphatic metastasis based on the allele model (OR = 1.56; 95% CI = 1.08 - 2.26; p = 0.018), the "T/C" genotype may increase the risk of lymphatic metastasis under the co-dominant model (OR = 2.51; 95% CI = 1.42 - 4.44; p = 0.002), the "T/C-C/C" genotype was related to an increased lymphatic metastasis risk in the dominant model (OR = 2.39; 95% CI = 1.40 - 4.07; p = 0.001), rs9835916 may increase the risk of lymphatic metastasis based on the log-additive model (OR = 1.64; 95% CI = 1.11 - 2.41; p = 0.013), (**Table 6)**.

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

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 [20, 21]. *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, 22]. *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, 23]. It is inferred that *CMTM8* overexpression blocks c-MET signaling in vivo model of bladder cancer. Researchs 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 tumorgenesis and development of clear-cell renal cell carcinoma, and the location and intensity of expression were significantly correlated with prognosis [24].

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 statement

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.

Consent to publish

Not applicable.

Availability of data and materials

The datasets used/or analyzed 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

TBJ and QFL: conceived and designed the experiments;

ZIX and FLN: performed the experiments;

YWL and HYL: analyzed the data;

JFL: contributed reagents/materials/analysis tools;

JMW and YS: drafted the work or revised it critically for important content.

Acknowledgments

We thank all the patients and individuals for their participation, as well as the clinicians and other hospital staff of the Shaanxi Provincial Cancer Hospital Affiliated to Medical College. Meanwhile, we are very grateful to the editors and reviewers for their patience and valuable comments to this work.

References

1. Bilello KS, Murin S, Matthay RA: **Epidemiology, etiology, and prevention of lung cancer**. *Clinics in chest medicine* 2002, **23**(1):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-Pacific journal of public health* 2015, **27**(2):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 biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* 2014, **35**(3):1863-1866.
4. Jin C, Wang Y, Han W, Zhang Y, He Q, Li D, Yin C, Tian L, Liu D, Song Q *et al*: **CMTM8 induces caspase-dependent and -independent apoptosis through a mitochondria-mediated pathway**. *Journal of cellular physiology* 2007, **211**(1):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 Acta Academiae Medicinae Sinicae* 2016, **38**(6):746-749.
6. Jin C, Ding P, Wang Y, Ma D: **Regulation of EGF receptor signaling by the MARVEL domain-containing protein CKLFSF8**. *FEBS letters* 2005, **579**(28):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**. *The Journal of biological chemistry* 2012, **287**(15):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**. *Applied immunohistochemistry & molecular morphology : AIMM* 2017, **25**(2):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**. *Oncology reports* 2015, **34**(6):2853-2863.
10. Li TY, Zhang F: **Screening of lung cancer related SNPs and CNVs with SNP microarrays**. *European review for medical and pharmacological sciences* 2015, **19**(2):225-234.
11. Erichsen HC, Chanock SJ: **SNPs in cancer research and treatment**. *British Journal of Cancer* 2004, **90**(4):747-751.
12. Chen J, Liu W, Cao Y, Zhang X, Guo Y, Zhu Y, Li J, Du J, Jin T, Wang G *et al*: **MMP-3 and MMP-8 single-nucleotide polymorphisms are related to alcohol-induced osteonecrosis of the femoral head in Chinese males**. *Oncotarget* 2017, **8**(15):25177-25188.
13. Gabriel S, Ziaugra L, Tabbaa D: **SNP genotyping using the Sequenom MassARRAY iPLEX platform**. *Current protocols in human genetics* 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**. *Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico* 2015, **17**(7):564-569.
15. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT: **Data quality control in genetic case-control association studies**. *Nature protocols* 2010, **5**(9):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**. *European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP)* 2014, **23**(6):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**(15):1928-1929.
19. Barrett JC, Fry B, Maller J, Daly MJ: **Haploview: analysis and visualization of LD and haplotype maps**. *Bioinformatics* 2005, **21**(2):263-265.
20. Imreh S, Klein G, Zabarovsky ER: **Search for unknown tumor-antagonizing genes**. *Genes, chromosomes & cancer* 2003, **38**(4):307-321.
21. 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**(4):663-669.
22. 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**(12):e115835.
23. 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 biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* 2016, **37**(5):6217-6225.
24. 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 = Journal of Peking University Health sciences* 2013, **45**(4):537-541.

Tables

Table 1: Demographic characteristics of cases and controls in this study

Variable (s)	Case, n (%) (n = 509)	Control, n (%) (n = 506)	*p* - value
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%)	102 (20.1%)	
No	358 (70.3%)	100 (19.8%)	
Unavailable	37 (7.2%)	304 (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.

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.

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

SNP	Model	Genotype	Control	Case	Without Adjustment		With Adjustment	
					OR (95% CI)	p^a-value	OR (95% CI)	p^b-value
rs9853415	Codominant	A/A	136 (26.9%)	128 (25.1%)	1.00		1.00	
		A/G	256 (50.6%)	273 (53.6%)	1.13 (0.84-1.52)	0.407	1.14 (0.84-1.53)	0.400
		G/G	114 (22.5%)	108 (21.2%)	1.01 (0.7-1.44)	0.971	1.01 (0.7-1.44)	0.969
	Dominant	A/A	136 (26.9%)	128 (25.1%)	1.00		1.00	
		A/G-G/G	370 (73.1%)	381 (74.9%)	1.09 (0.83-1.45)	0.530	1.1 (0.83-1.45)	0.523
	Recessive	A/A-A/G	392 (77.5%)	401 (78.8%)	1.00		1.00	
		G/G	114 (22.5%)	108 (21.2%)	0.93 (0.69-1.25)	0.613	0.93 (0.69-1.25)	0.609
	Log-additive	---	---	---	1.01 (0.84-1.21)	0.923	1.01 (0.84-1.21)	0.921
rs6796318	Co-dominant	A/A	426 (84.4%)	423 (83.1%)	1.00		1.00	
		A/G	75 (14.9%)	80 (15.7%)	1.07 (0.76-1.51)	0.682	1.07 (0.76-1.51)	0.698
		G/G	4 (0.8%)	6 (1.2%)	1.51 (0.42-5.39)	0.525	1.53 (0.43-5.47)	0.513
	Dominant	A/A	426 (84.4%)	423 (83.1%)	1.00		1.00	
		A/G-G/G	79 (15.6%)	86 (16.9%)	1.1 (0.79-1.53)	0.589	1.09 (0.78-1.53)	0.600
	Recessive	A/A-A/G	501 (99.2%)	503 (98.8%)	1.00		1.00	
		G/G	4 (0.8%)	6 (1.2%)	1.49 (0.42-5.33)	0.536	1.51 (0.42-5.40)	0.523
	Log-additive	---	---	---	1.11 (0.81-1.5)	0.518	1.11 (0.81-1.50)	0.524
rs6771238	Co-dominant	C/C	467 (92.3%)	451 (88.6%)	1.00		1.00	
		C/A	37 (7.3%)	56 (11.0%)	1.57 (1.01-2.42)	**0.043***	1.57 (1.01-2.42)	**0.044***
		A/A	2 (0.4%)	2 (0.4%)	1.04 (0.15-7.38)	0.972	1.07 (0.15-7.63)	0.949
	Dominant	C/C	467 (92.3%)	451 (88.6%)	1.00		1.00	
		C/A-A/A	39 (7.7%)	58 (11.4%)	1.54 (1.01-2.36)	**0.047***	1.54 (1.01-2.36)	**0.047***
	Recessive	C/C-C/A	504 (99.6%)	507 (99.6%)	1.00		1.00	
		A/A	2 (0.4%)	2 (0.4%)	0.99 (0.14-7.08)	0.995	1.03 (0.14-7.33)	0.981
	Log-additive	---	---	---	1.47 (0.98-2.19)	0.062	1.47 (0.98-2.20)	0.061
rs9835916	Co-dominant	T/T	181 (36.9%)	161 (32.1%)	1.00		1.00	
		T/C	226 (46.1%)	249 (49.7%)	1.24 (0.94-1.64)	0.132	1.25 (0.94-1.65)	0.124
		C/C	83 (16.9%)	91 (18.2%)	1.23 (0.86-1.78)	0.262	1.24 (0.86-1.78)	0.256
	Dominant	T/T	181 (36.9%)	161 (32.1%)	1.00		1.00	

		T/C-C/C	309 (63.1%)	340 (67.9%)	1.24 (0.95-1.61)	0.112	1.24 (0.96-1.62)	0.105
	Recessive	T/T-T/C	407 (83.1%)	410 (81.8%)	1.00		1.00	
		C/C	83 (16.9%)	91 (18.2%)	1.09 (0.78-1.51)	0.612	1.09 (0.78-1.51)	0.611
	Log-additive	---	---	---	1.13 (0.95-1.35)	0.176	1.13 (0.95-1.35)	0.170
rs1077868	Co-dominant	A/A	354 (70%)	344 (67.6%)	1.00		1.00	
		A/G	136 (26.9%)	154 (30.3%)	1.17 (0.89-1.53)	0.274	1.16 (0.88-1.53)	0.283
		G/G	16 (3.2%)	11 (2.2%)	0.71 (0.32-1.55)	0.386	0.7 (0.32-1.54)	0.380
	Dominant	A/A	354 (70%)	344 (67.6%)	1.00		1.00	
		A/G-G/G	152 (30.0%)	165 (32.4%)	1.12 (0.86-1.46)	0.414	1.11 (0.85-1.45)	0.427
	Recessive	A/A-A/G	490 (96.8%)	498 (97.8%)	1.00		1.00	
		G/G	16 (3.2%)	11 (2.2%)	0.68 (0.31-1.47)	0.325	0.67 (0.31-1.47)	0.320
	Log-additive	---	---	---	1.05 (0.83-1.33)	0.677	1.05 (0.83-1.32)	0.694
rs6802418	Co-dominant	G/G	248 (49%)	230 (45.2%)	1.00		1.00	
		G/A	209 (41.3%)	228 (44.8%)	1.18 (0.91-1.53)	0.220	1.18 (0.91-1.52)	0.225
		A/A	49 (9.7%)	51 (10%)	1.12 (0.73-1.73)	0.600	1.12 (0.72-1.72)	0.618
	Dominant	G/G	248 (49%)	230 (45.2%)	1.00		1.00	
		G/A-A/A	258 (51.0%)	279 (54.8%)	1.17 (0.91-1.49)	0.222	1.16 (0.91-1.49)	0.229
	Recessive	G/G-G/T	457 (90.3%)	458 (90.0%)	1.00		1.00	
		A/A	49 (9.7%)	51 (10.0%)	1.04 (0.69-1.57)	0.858	1.03 (0.68-1.56)	0.877
	Log-additive	---	---	---	1.10 (0.91-1.33)	0.313	1.10 (0.91-1.33)	0.325

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 and age;

*p-value < 0.05 indicates statistical significance.

types (adjusted for sex and age)

SNP	Model	Genotype	Squamous cell carcinoma				Adenocarcinoma			
			Control	Case	OR (95% CI)	*p* value	Control	Case	OR (95% CI)	*p* value
rs6771238	Co-dominant	C/C	467 (92.3%)	104 (86.0%)	1.00		467 (92.3%)	171 (88.6%)	1.00	
		A/C	37 (7.3%)	16 (13.2%)	2.07 (1.08-3.97)	**0.028***	37 (7.3%)	21 (10.9%)	1.60 (0.90-2.82)	0.107
		A/A	2 (0.4%)	1 (0.8%)	1.97 (0.17-22.45)	0.586	2 (0.4%)	1 (0.5%)	1.70 (0.15-19.02)	0.668
	Dominant	C/C	467 (92.3%)	104 (86.0%)	1.00	**0.025***	467 (92.3%)	171 (88.6%)	1.00	0.098
		A/C-A/A	39 (7.7%)	17 (14.0%)	2.07 (1.10-3.89)		39 (7.7%)	22 (11.4%)	1.60 (0.92-2.80)	
	Recessive	C/C-A/C	504 (99.6%)	120 (99.2%)	1.00	0.627	504 (99.6%)	192 (99.5%)	1.00	0.694
		A/A	2 (0.4%)	1 (0.8%)	1.83 (0.16-20.81)		2 (0.4%)	1 (0.5%)	1.63 (1.45-18.18)	
	Log-addition	---	---	---	1.90 (1.06-3.38)	**0.030***	---	---	1.54 (0.92-2.58)	0.103
rs6802418	Codominant	G/G	248 (49.0%)	53 (43.8%)	1.00		248 (49.0%)	81 (42%)	1.00	
		A/G	209 (41.3%)	55 (45.5%)	1.21 (0.79-1.88)	0.369	209 (41.3%)	94 (48.7%)	1.43 (1.00-2.03)	**0.049***
		A/A	49 (9.7%)	13 (10.7%)	1.24 (0.61-2.49)	0.551	49 (9.7%)	18 (9.3%)	1.13 (0.62-2.06)	0.700
	Dominant	G/G	248 (49.0%)	53 (43.8%)	1.00	0.338	248 (49.0%)	81 (42%)	1.00	0.070
		A/G-A/A	258 (51%)	68 (56.2%)	1.22 (0.81-1.84)		258 (51%)	112 (58.0%)	1.37 (0.97-1.92)	
	Recessive	G/G-A/G	457 (90.3%)	108 (89.3%)	1.00	0.731	457 (90.3%)	175 (90.7%)	1.00	0.850
		A/A	49 (9.7%)	13 (10.7%)	1.12 (0.58-2.18)		49 (9.7%)	18 (9.3%)	0.95 (0.53-1.68)	
	Log-addition	---	---	---	1.15 (0.85-1.56)	0.377	---	---	1.18 (0.92-1.53)	0.194

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

p - value calculated by logistic regression analysis with adjustments for gender and age;

*p-value < 0.05 indicates statistical significance.

Table 5: Stratified analysis of *CMTM8* polymorphism and staging risk in patients with lung cancer (adjusted for gender and age)

SNP ID	Model	Genotype	control	case	OR (95% CI)	*p* value
rs9835916	Co-dominant	T/T	34 (40.5%)	77 (24.8%)	1.00	
		T/C	35 (41.7%)	193 (62.3%)	1.75 (1.00-3.05)	**0.047***
		C/C	15 (17.9%)	40 (12.9%)	1.19 (0.58-2.44)	0.637
	Dominant	T/T	34 (40.5%)	77 (24.8%)	1.00	0.081
		T/C-C/C	50 (59.6%)	233 (75.2%)	1.58 (0.94-2.65)	
	Recessive	T/T-C/T	69 (82.2%)	270 (87.1%)	1.00	0.654
		C/C	15 (17.9%)	40 (12.9%)	0.86 (0.45-1.66)	
	Log-addition	---	---	---	1.20 (0.83-1.73)	0.338
rs1077868	Co-dominant	A/A	64 (76.2%)	170 (64.9%)	1.00	
		A/G	19 (22.6%)	83 (31.7%)	1.69 (0.95-3.01)	0.076
		G/G	1 (1.2%)	8 (3.1%)	3.21 (0.39-26.52)	0.278
	Dominant	A/A	64 (76.2%)	170 (64.9%)	1.00	0.051
		A/G-G/G	20 (23.8%)	91 (34.8%)	1.76 (1.00-3.11)	
	Recessive	A/A-A/G	83 (98.8%)	253 (96.6%)	1.00	0.348
		G/G	1 (1.2%)	8 (3.1%)	2.74 (0.33-22.43)	
	Log-addition	---	---	---	1.71 (1.02-2.88)	**0.043***

ORs, odds ratios; CI, confidence interval; *p value < 0.05 indicates statistical significance.

Table 6: Stratified analysis of the polymorphisms of *CMTM8* and the risk of lymph node metastasis in patients with lung cancer (adjusted by sex, age)

SNP	Model	Genotype	control	case	OR (95% CI)	[a] *p* value
rs9835916	Co-dominant	T/T	38 (45.2%)	56 (26.3%)	1.00	
		T/C	33 (39.3%)	119 (55.9%)	2.51 (1.42-4.44)	**0.002***
		C/C	13 (15.5%)	38 (17.8%)	2.08 (0.98-4.45)	0.058
	Dominant	T/T	38 (45.2%)	56 (26.3%)	1.00	**0.001***
		T/C-C/C	46 (54.8%)	157 (73.7%)	2.39 (1.40-4.07)	
	Recessive	T/T-C/T	71 (84.5%)	175 (82.2%)	1.00	0.554
		C/C	13 (15.5%)	38 (17.8%)	1.23 (0.62-2.47)	
	Log-additive	---	---	---	1.64 (1.11-2.41)	**0.013***
	Alleles	T	103 (61.3%)	293 (57.2%)	1.00	
		C	65 (38.7%)	219 (42.8%)	1.56 (1.08-2.26)	**0.018***

ORs, odds ratios; CI, confidence interval□*p value < 0.05 indicates statistical significance.

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

Variable	SNP	Haplotype	Freq (case)	Freq (control)	Without adjusted		With adjusted	
					OR (95% CI)	*p*	OR (95% CI)	*p*
Staging	rs1077868|rs6802418	AA	0.349	0.292	1.30 (0.89-1.91)	0.173	1.31 (0.89-1.92)	0.168
		GG	0.190	0.125	1.66 (0.99-2.79)	0.053	1.71 (1.02-2.88)	**0.043***
		AG	0.5383	0.417	1.59(1.13-2.25)	0.008*	1.62 (1.15-2.30)	**0.007***

Block comprised of the three closely linked SNPs rs9440302 and rs1198574;

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 gender and age;

*P-value < 0.05 indicates statistical significance.

Figures

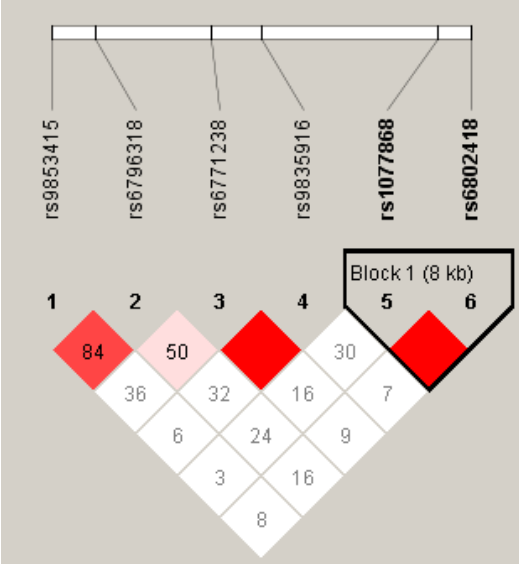

Figure 1

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

Supplementary Files

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

- Tables.pdf
- Supplementaryfiles.pdf