Short Communication

The genetic variant rs401681C/T is associated with the risk of non-small cell lung cancer in a Chinese mainland population

H. Wang[1,2], Y. Zhao[3], J. Ma[4], G. Zhang[2], Y. Mu[2], G. Qi[2], Z. Fang[5], L. Wang[5], Q. Fan[1] and Z. Ma[2]

[1]Department of Oncology, First Affiliated Hospital, Zhengzhou University, Zhengzhou, Henan, P.R. China
[2]Department of Medicine, Affiliated Cancer Hospital, Zhengzhou University, Zhengzhou, Henan, P.R. China
[3]Department of Surgery, Affiliated Cancer Hospital, Zhengzhou University, Zhengzhou, Henan, P.R. China
[4]Department of Pathology, Affiliated Cancer Hospital, Zhengzhou University, Zhengzhou, Henan, P.R. China
[5]Biomedical Research Institute, Shenzhen-PKU-HKUST Medical Center, Shenzhen, Guangdong, P.R. China

Corresponding authors: Q. Fan / Z. Ma
E-mail: fqx2243@yahoo.com.cn / mazhiyong@csco.org.cn

ABSTRACT. Although lung cancer (LC) is a highly environmentally associated disease, genetic factors are also thought to play a role in this disease. In recent years, genome-wide association studies have identified various susceptible genetic regions for LC. Herein, we used high-resolution melting analysis to genotype 2 significant single nucleotide polymorphisms previously reported in Caucasians, that is, rs401681 at 5p15.33 and rs8034191 at 15q25, in a case-control study with 492 LC cases and 486 cancer-free controls in a Chinese population. We found

that the rs401681C/T allele in the *TERT-CLPTM1L* gene was associated with the risk of non-small cell lung cancer [NSCLC; P = 0.012, odds ratio (OR) = 1.29, 95% confidence interval (95%CI) = 1.09-1.50], but was not associated with the risk of small cell lung cancer (P = 0.571, OR = 1.15, 95%CI = 0.82-1.47). However, no significant association was found between rs8034191T/C and LC risk. These results suggest that genetic variants in the *TERT-CLPTM1L* gene may predispose individuals to be susceptible to LC, particularly NSCLC, in the Chinese population.

Key words: Lung cancer; Single nucleotide polymorphism; Risk factor; Non-small cell lung cancer; High-resolution melting analysis

INTRODUCTION

Lung cancer (LC) represents the second most common cancer in both men and women and accounts for approximately 15% of all cancer diagnoses (Kollarova et al., 2002; Ginsberg, 2005; Thill et al., 2011). The most common etiological factors of LC are environment-related, including smoking, exposure to polycyclic aromatic hydrocarbons, asbestos, metals, welding fumes, air pollution, and malnutrition or poor diet (Haugen et al., 2000; Lam, 2005). However, only a small percentage of individuals exposed to hazardous environmental factors develop LC disease, which suggests that factors other than the environment, such as host (e.g., chronic lung diseases) or genetic factors (e.g., gene variation) may predispose individuals to develop LC (Sanchez-Cespedes, 2009; Sugimura et al., 2011). Cancer researchers have long been interested in investigating genetic susceptibility to LC. The findings in this area were achieved by using comprehensive approaches such as genomic, transcriptomic, or proteomic studies.

Despite new developments in molecular techniques and better knowledge of LC tumor biology, many genetic alterations associated with the development and progression of lung carcinogenesis still remain unclear. LC has a very high mortality rate, even at early stages. Thus, many studies have investigated LC susceptibility based on the presence of low-penetrance, high-frequency single nucleotide polymorphisms (SNPs). Identifying susceptibility SNPs may lead to the development of tests that allow for a more focused follow-up of high-risk individuals (Kiyohara et al., 2007).

Recently, genome-wide association studies have identified rs401681 (at the *TERT-CLPTM1L* locus on 5p15.33) and rs8034191 (at *CHRNA3/5* on 15q25) as the susceptible regions for LC in American and European populations (Hung et al., 2008; McKay et al., 2008). The rs401681 locus can influence the risk of several types of human solid carcinomas, including melanoma (Law et al., 2012), breast cancer (Haiman et al., 2011), and head and neck carcinomas (Liu et al., 2010). The rs8034191 locus has been shown to correlate with chronic obstructive pulmonary disease (Hardin et al., 2012) and nicotine dependence. However, polymorphisms of these 2 loci have not been widely examined in LC patients in a Chinese population. In the present study, we collected a relatively larger number of LC samples from Chinese mainland patients and tested the association between genetic variants rs401681 and rs8034191 and the risk of LC in a Chinese Han population through a recently developed method, unlabeled probe-based high-resolution melting analysis (HRMA).

MATERIAL AND METHODS

Patients

A total of 492 patients (380 males and 112 females; median age, 60 years; range, 16-79) and 486 ethnically matched healthy controls (378 males and 108 females; median age, 61 years; range 21-82) were recruited from the Affiliated Cancer Hospital of Zhengzhou University. The study was approved by the institutional review board of the Affiliated Cancer Hospital of Zhengzhou University, and written informed consent was obtained from all patients.

Genotyping

Genomic DNA from the patients was isolated from formalin-fixed/paraffin-embedded cancer tissues by using the Allprep DNA/RNA FFPE Kit (Qiagen, Gaithersburg, MA, USA) following the manufacturer protocol. Genomic DNA from the healthy controls was isolated from peripheral blood cells using the Innogent genomic DNA extraction kit (Innogent, Shenzhen, Guangdong, China) following the manufacturer protocol. Genotyping was assayed by HRMA with an unlabeled probe as previously described. Briefly, asymmetric polymerase chain reaction (PCR) was performed in a volume of 20 μL containing 20 ng genomic DNA, 1X PCR buffer (Takara, Otsu, Shigaa, Japan), 200 μM dNTPs, 0.5 U rTaq DNA polymerase (Takara), 0.05 μM forward primer, 0.5 μM excess reverse primer, and 0.5 μM C3-block probe. PCR was performed in an S1000 Thermal Cycler (Bio-Rad, Hercules, CA, USA). The conditions included an initial denaturation at 94°C for 2 min, followed by 45 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 20 s; with a final extension at 72°C for 5 min. The 10-μL volume of PCR products was supplemented with 1 μL LC Green (Idaho Technology, Salt Lake City, UT, USA) and then subjected to HRM in an HR-1 instrument (Idaho Technology). The samples were 1st denatured at 95°C for 30 s and rapidly cooled to 40°C for 30 s; then, were melted from 55° to 90°C with a -0.3°C/s ramp rate. Melting curves were analyzed with the Light Scanner Service software (Roche, Indianapolis, IN, USA). The sequences of the primers for rs401681 were as follows: forward 5'-AGA AAA CAA GGT CTG CTA TCC A-3', reverse 5'-GAC TCT TGA TAA ACT TAC CAG CC-3'. The sequence of the unlabeled C3-blocked probe was as follows: 5'-ACA ACT TCA GAG TCT ATC ATG GTG TGA AG-3'.

Statistical analysis

The SNPs were analyzed for association with LC by means of comparison of the minor allele frequency in patients and controls as well as the constancy of Hardy-Weinberg equilibrium using the chi-square test or the Fisher exact test. The magnitude of association was expressed as odds ratio with a 95% confidence interval. P values less than 0.05 were considered to be statistically significant.

RESULTS

High-resolution melting analysis using unlabeled probe

We performed HRMA with an unlabeled C3-blocked probe for genotyping. As shown in

Figure 1, 3 genotypes (CC, CT, and TT) of SNP rs401681 (C>T) were accurately distinguished by the derivative melting curves in the probe region. The derivative melting curves obtained directly from the PCR products were not able to distinguish these genotypes (Figure 1, right panel), and the sensitivity and accuracy of HRMA were dramatically improved using unlabeled probes (Figure 1, left panel). The data for SNP rs8034191were similar and not shown here.
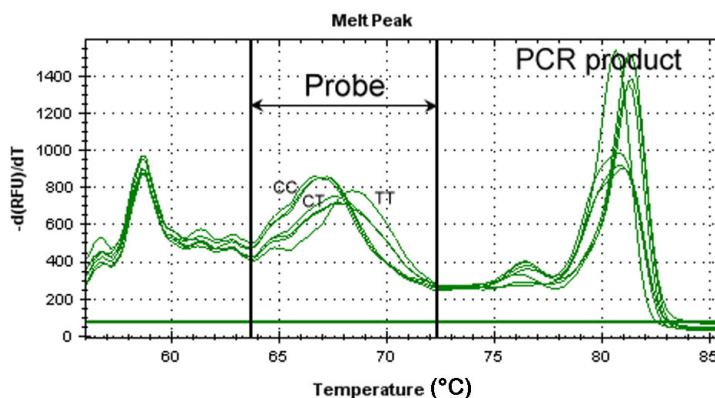

Figure 1. Derivative melting curves of unlabeled probes and amplicon for genotyping of single nucleotide polymorphism rs401681. Three genotypes (CC, CT, and TT) were discriminated as indicated in the probe region.

We could clearly discriminate all the genotypes after the normalized melting unlabeled probe region was analyzed (Figure 2). The difference curves generated by the subtraction of one heterozygote curve displayed a clearer view for genotype discrimination. Then, this method was employed to screen the LC samples in the present study.
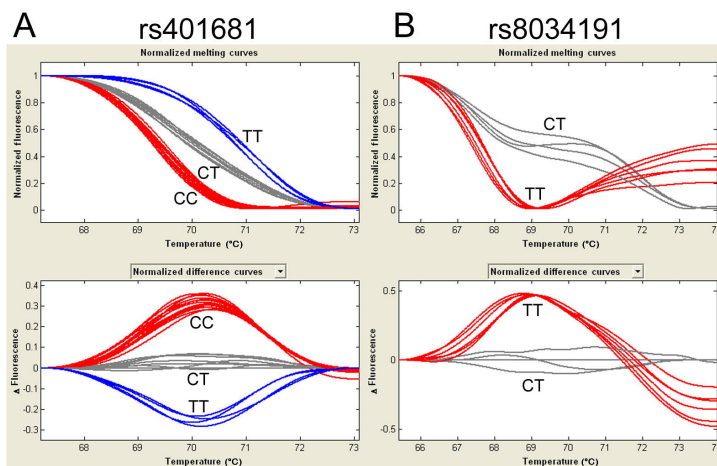

Figure 2. Normalized melting curves and difference curves of unlabeled probe region for genotyping of single nucleotide polymorphisms rs401681 and rs8034191. **A.** After the amplification of the rs401681 region, the melting curve was calculated and normalized. The difference curves were obtained by subtracting each curve from one heterozygote (CT) curve. Three genotypes (CC, CT, and TT) were shown as indicated. **B.** The melting curve of the amplification product of the rs8034191 region was presented after normalization. The difference curves were obtained by subtracting each curve from one heterozygote (TC) curve. Two main genotypes (TT, CT) of rs401681 were shown as indicated.

Association between SNP rs401681 and LC

Table 1 shows the *TERT-CLPTM1L* rs401681 and *CHRNA3/5* rs8034191genotypes and allele frequencies in 492 LC patients and 486 healthy controls. Genotype frequencies were in Hardy-Weinberg equilibrium in both the patients and controls. The C/C genotype of rs401681 was more prevalent among patients with LC (245: 492; 49.8%) than among healthy controls (215: 486; 44.2%). The frequencies of the rs401681 T allele were associated with LC (P = 0.045).

Neither the genotype nor allele frequencies of rs8034191 showed statistically significant differences between LC patients and controls.

Table 1. Genotype and allele frequencies of single nucleotide polymorphisms (SNP) rs401681 and rs8034191 in lung cancer (LC) cases and controls.

SNP population	No. of subjects	Genotype frequency [N (%)]				Allele frequency			OR (95%CI)
		Major homozygote	Heterozygote	Minor homozygote	P	Major allele	Minor allele	P	
Genotype or allele (for rs401681)		CC	CT	TT		C	T		
LC cases	492	245 (49.8)	201 (40.9)	46 (9.3)	0.045	691	293	0.016	1.26 (1.07-1.45)
Controls	486	215 (44.2)	203 (41.8)	68 (14.0)		633	339		
Genotype or allele (for rs8034191)		TT	TC	CC		T	C		
LC cases	381	350 (91.9)	29 (7.6)	2 (0.5)	0.234	729	33	0.175	0.69 (0.17-1.22)
Controls	410	385	25	0		795	25		

OR = odds ratio; 95%CI = 95% confidence interval.

Association between SNP rs401681 and non-small cell lung cancers

In order to further analyze the possible significance of *TERT-CLPTM1L* SNPs in different pathological types of LC, we compared the genotype and allele frequencies of rs401681 between non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) cases. As shown in Table 2, the C/C genotype of rs401681 was more prevalent in patients with NSCLC (200: 393; 50.9%) than in healthy controls (44.2%; P = 0.0479). The minor allele (T) of rs401681 was significantly associated with the incidence of NSCLC. On the other hand, no obvious association was observed between the rs401681 T allele and SCLC (T allele: 68.2 *vs* 65.1%, P = 0.482). Thus, rs401681 primarily influenced the risk of NSCLC, but not SCLC, in the selected population.

Table 2. Genotype and allele frequencies of single nucleotide polymorphism (SNP) rs401681 in non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) cases.

SNP population	No. of subjects	Genotype frequency [N (%)]				Allele frequency			OR (95%CI)
		Major homozygote	Heterozygote	Minor homozygote	P (vs control)	Major allele	Minor allele	P (vs control)	
Genotype or allele		CC	CT	TT		C	T		
NSCLC cases	393	200 (50.9)	156 (39.7)	37 (9.4)	0.0479	556	230	0.012	1.29 (1.09-1.50)
SCLC cases	99	45 (45.5)	45 (45.5)	9 (9.0)	0.571	135	63	0.482	1.15 (0.82-1.47)
Normal controls	486	215 (44.2)	203 (41.8)	68 (14.0)		633	339		

OR = odds ratio; 95%CI = 95% confidence interval.

DISCUSSION

HRMA is a powerful molecular biology technique for the detection of mutations, polymorphisms, and epigenetic differences in double-stranded DNA samples. Based on its ease of use, simplicity, flexibility, low cost, nondestructive nature, superb sensitivity, and specificity, HRMA is quickly becoming the tool of choice for screening patients for pathogenic variants (Vossen et al., 2009). However, it is sometimes difficult to distinguish SNPs of A/T or G/C, which make up about 16% of human SNPs, as these SNPs cause an almost undetectable melting temperature (Tm) shift (<0.4°C). Unlabeled probe melting analysis is a modified HRMA technique in which an approximately 30-bp C3-blocked probe is used to target the SNP of interest during the melting process. A single base pair difference in such a short probe could result in a significant Tm shift (e.g., 3-4°C). As shown in our study, we can discriminate almost all of the SNPs in a more accurate and reliable manner using unlabeled probe melting analysis. Thus, the technology of HRMA with an unlabeled probe is strongly recommended for studying genetic variants of tumors such as LC.

Our results revealed that rs401681 was significantly associated with LC cases in the Chinese population, which was consistent with the results obtained from Caucasian (Rafnar et al., 2009) and Korean (Bae et al., 2012) populations. However, rs8034191 on 15q25.1 was not significantly correlated with LC cases in the Chinese population. Similar results have been previously reported by another group in China (Wu et al., 2009). In their study, 4 common 15q25 variants were associated with LC, while rs8034191 was excluded (Wu et al., 2009). Other groups have reported similar results in western countries (Girard et al., 2010). The discrepancy of the obtained data might be due to the different races and populations used in the studies.

The underlying mechanism of rs401681 in LC or other types of carcinomas remains largely unknown. In recent years, TERT has been considered to be a more plausible candidate at 5p15.33 for LC risk due to its definitive role in telomerase activity as a reverse transcriptional catalytic subunit (Weinrich et al., 1997). Telomeres are specialized structures at eukaryotic chromosome ends, which role is to prevent them from degradation, end-to-end fusion and rearrangement. High expression or activation of telomerase has always been observed in tumor samples, including lung cancer, suggesting the critical role of TERT in tumorigenesis.

Genetic variant rs401681 in the *TERT-CLPTM1L* locus was strongly correlated with NSCLC but not SCLC cases in the Chinese population, which was a novel finding. SCLC accounts for 15-25% of all lung cancer cases. This type of LC is almost always caused by smoking, secondhand smoke, and exposure to radon or asbestos. Thus, genetic variants might play a minor role in the tumorigenesis of SCLC. NSCLC accounts for the majority of all LCs, and is divided into 3 subtypes: squamous cell carcinoma, adenocarcinoma, and large cell carcinoma. Smoking increases the risk of all histological subtypes, but is most strongly associated with squamous cell and small cell disease. Adenocarcinoma is more common in women than in men and more common in non-smokers with disease than smokers. Besides smoking, inherited cancer risk and exposure to toxins and viruses also contribute to tumorigenesis in NSCLC. This might be a reason as to why genetic susceptibility due to rs401681 mainly influences the risk of NSCLC but not SCLC.

Genotyping with unlabeled probes and PCR product melting has been successfully applied for the detection of disease-related SNPs such as factor V Leiden or hot-spot mutations in cystic fibrosis and even for detection and typing of herpes simplex virus. Unlabeled probe

HRM is a high-throughput, inexpensive, closed-tube genotyping method that may substantially reduce the burden of sequencing. In summary, our data suggest that a *TERT-CLPTM1L* gene polymorphism is associated with the incidence of LC, particularly NSCLC, in a Chinese population.

ACKNOWLEDGMENTS

REFERENCES

Bae EY, Lee SY, Kang BK, Lee EJ, et al. (2012). Replication of results of genome-wide association studies on lung cancer susceptibility loci in a Korean population. *Respirology* 17: 699-706.

Ginsberg MS (2005). Epidemiology of lung cancer. *Semin. Roentgenol.* 40: 83-89.

Girard N, Lou E, Azzoli CG, Reddy R, et al. (2010). Analysis of genetic variants in never-smokers with lung cancer facilitated by an Internet-based blood collection protocol: a preliminary report. *Clin. Cancer Res.* 16: 755-763.

Haiman CA, Chen GK, Vachon CM, Canzian F, et al. (2011). A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. *Nat. Genet.* 43: 1210-1214.

Hardin M, Zielinski J, Wan ES, Hersh CP, et al. (2012). CHRNA3/5, IREB2, and ADCY2 are associated with Severe COPD in Poland. *Am. J. Respir. Cell Mol. Biol.* [Epub ahead of print].

Haugen A, Ryberg D, Mollerup S, Zienolddiny S, et al. (2000). Gene-environment interactions in human lung cancer. *Toxicol. Lett.* 112-113: 233-237.

Hung RJ, McKay JD, Gaborieau V, Boffetta P, et al. (2008). A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 452: 633-637.

Kiyohara C, Yoshimasu K, Takayama K and Nakanishi Y (2007). Lung cancer susceptibility: are we on our way to identifying a high-risk group? *Future Oncol.* 3: 617-627.

Kollarova H, Janout V and Cizek L (2002). Epidemiology of lung cancer. *Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech. Repub.* 146: 103-114.

Lam WK (2005). Lung cancer in Asian women-the environment and genes. *Respirology* 10: 408-417.

Law MH, Montgomery GW, Brown KM, Martin NG, et al. (2012). Meta-analysis combining new and existing data sets confirms that the TERT-CLPTM1L locus influences melanoma risk. *J. Invest. Dermatol.* 132: 485-487.

Liu Z, Li G, Wei S, Niu J, et al. (2010). Genetic variations in TERT-CLPTM1L genes and risk of squamous cell carcinoma of the head and neck. *Carcinogenesis* 31: 1977-1981.

McKay JD, Hung RJ, Gaborieau V, Boffetta P, et al. (2008). Lung cancer susceptibility locus at 5p15.33. *Nat. Genet.* 40: 1404-1406.

Rafnar T, Sulem P, Stacey SN, Geller F, et al. (2009). Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat. Genet.* 41: 221-227.

Sanchez-Cespedes M (2009). Lung cancer biology: a genetic and genomic perspective. *Clin. Transl. Oncol.* 11: 263-269.

Sugimura H, Tao H, Suzuki M, Mori H, et al. (2011). Genetic susceptibility to lung cancer. *Front Biosci.* 3: 1463-1477.

Thill PG, Goswami P, Berchem G and Domon B (2011). Lung cancer statistics in Luxembourg from 1981 to 2008. *Bull. Soc. Sci. Med. Grand Duche Luxemb.* 43-55.

Vossen RH, Aten E, Roos A and den Dunnen JT (2009). High-resolution melting analysis (HRMA): more than just sequence variant screening. *Hum. Mutat.* 30: 860-866.

Weinrich SL, Pruzan R, Ma L, Ouellette M, et al. (1997). Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nat. Genet.* 17: 498-502.

Wu C, Hu Z, Yu D, Huang L, et al. (2009). Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. *Cancer Res.* 69: 5065-5072.