Thoracic Cancer
Open Access
Thoracic Cancer ISSN 1759-7706

ORIGINAL ARTICLE

HapMap-based study: *CYP2A13* may be a potential key metabolic enzyme gene in the carcinogenesis of lung cancer in non-smokers

Feng Hua[1], Yonglu Guo[2], Qiang Sun[3], Leizhou Yang[4] & Fang Gao[1]

1 Department of Thoracic Surgery, Shandong Cancer Hospital Affiliated to Shandong University, Shandong Academy of Medical Sciences, Jinan, China
2 Department of Respiratory, Jiuquan City People's Hospital, Jiuquan, China
3 Department of Infection, Jiuquan City People's Hospital, Jiuquan, China
4 Department of Internal Medicine, Jining City Yanzhou District Railway Hospital, Jining, China

Keywords
CYP2A13; genetic polymorphism; HapMap Project; lung cancer; susceptibility.

Correspondence
Fang Gao, Department of Thoracic Surgery, Shandong Cancer Hospital Affiliated to Shandong University, Shandong Academy of Medical Sciences, Jiyan Road 440, Jinan, Shandong Province 250117, China.
Tel: +86 531 6762 6882
Fax: +86 531 6762 6882
Email: gaofangjinan@aliyun.com

Received: 8 October 2018;
Accepted: 3 December 2018.

doi: 10.1111/1759-7714.12954

Thoracic Cancer **10** (2019) 601–606

Abstract

Background: The aim of this study was to evaluate the association between *CYP2A13* polymorphisms and lung cancer susceptibility using the HapMap database.

Methods: A case-control analysis of 532 subjects with lung cancer and 614 controls with no personal history of the disease was performed. The tag SNPs rs1645690 and rs8192789 for *CYP2A13* were selected, and the genetic polymorphisms were confirmed experimentally through real-time PCR, cloning, and sequencing assay.

Results: SNP frequency in this study was consistent with the HapMap Project database of Han-Chinese and lung cancer risk was associated with *CYP2A13* polymorphisms in non-smokers. *CYP2A13* shares a 93.5% identity with *CYP2A6* in the amino acid sequence and the homologous sequences may interfere with the study of SNPs of *CYP2A13*.

Conclusions: *CYP2A13* may be a potential key metabolic enzyme gene in the carcinogenesis of lung cancer in non-smokers. The common polymorphisms of *CYP2A13* may be candidate biomarkers for lung cancer susceptibility in Han-Chinese.

Introduction

Publications relevant to genetic polymorphisms and lung cancer risk have been important to investigative strategies in the past decades. The outcomes of these studies have been contradictory regarding most genes, except *CYP2A13*.[1] Several studies support the inference that *CYP2A13*, a member of the *CYP2A* subfamily, may play an important role in lung cancer susceptibility. Firstly, it is selectively expressed in the mucosa of the trachea and the lungs,[2,3] and *CYP2A13* expression is markedly increased in non-small cell lung carcinomas.[4,5] Secondly, *CYP2A13* plays important roles in xenobiotic toxicity and tumorigenesis in the human respiratory tract, such as 4-(methyl-nitrosamino)-1-(3-pyridyl)-1- butanone (NNK), hexamethylphosphoramide, N,N- dimethylaniline, and nitrosomethyl- phenylamine, aflatoxin B1 (AFB1).[6–11] in vitro studies have shown that *CYP2A13* significantly enhances the alpha-hydroxylation of NNK, and the alpha-hydroxylation activity of *CYP2A13* is significantly higher than that of *CYP2A6*.[12] Further vivo studies using a *CYP2A13*-humanized mouse model indicate that *CYP2A13* is a low Km enzyme for catalyzing NNK bioactivation and support the notion that genetic polymorphisms of *CYP2A13* can influence the risk of tobacco-induced lung tumorigenesis in humans.[13] A structure-activity relationship study indicated that all CYP2A13 mutant proteins showed a significant decrease in catalytic efficiency (Vmax/Km) for NNK alpha-hydroxylation.[14–16] The catalytic activity of *CYP2A13* on NNK by computational calculation was consistent with the experimental results.[17]

The HapMap Project (www.hapmap.org) provides single nucleotide polymorphism (SNP) and disequilibrium information of *CYP2A13* on Han-Chinese (phase 2 data released 2007; phase 3, Feb 2009). The common disease–common variant hypothesis states that the genetic risk factor that contributes most to the risk of disease is commonly occurring polymorphisms, which have only a minor influence. In light of this hypothesis, association analyses using linkage disequilibrium (LD) mapping seems a reasonable approach to narrow down the number of potential risk genes or variants for the disease. The international HapMap database can be used to select haplotype tagging SNPs (htSNPs) for the genome-wide association research of many samples around the world.[18,19]

In the present study, we tested the association between *CYP2A13* polymorphisms and the risk of lung cancer by using the tagging SNP approach according to the acquired information. Another aim of this study was to assess whether the tagging SNPs of *CYP2A13* selected from HapMap predict genetic variation in our Chinese population.

Methods

Study population and data collection

This case–control study was conducted at Shandong Cancer Hospital and Institute, Jinan, China, and was approved by the Research Ethic Committee of the Shandong Cancer Hospital. All participants were ethnic Han-Chinese from Shandong province and its surrounding regions, recruited from February 2008 to October 2009. Informed consent was obtained from all participants.

A total of 532 patients were histologically and cytologically diagnosed with lung cancer, including 240 (45.1%) squamous cell carcinoma, 198 (37.2%) adenocarcinoma, 46 (16.9%) small-cell carcinoma, and 48 (9.0%) other. Two senior pathologists determined all histological classifications. The control subjects were randomly selected from a pool of healthy volunteers who had visited the general health check-up center at Shandong Cancer Hospital during the same period. A detailed questionnaire including information on demographic data (e.g. gender, age, tobacco smoking, tumor history, environmental exposure, diet) was completed for each participant by a trained interviewer. Information was collected on the number of cigarettes smoked per day, the age at which the subjects started smoking, and the age at which ex-smokers stopped smoking. A person who had smoked at least 100 cigarettes during his or her lifetime was considered a smoker. The cumulative cigarette dose (pack-years) was calculated using the following formula: pack-years = (packs per day) × (years smoked). We further categorized the subjects according to smoking status: never smokers, light smokers (≤ 27 pack-years, the mean of the pack-year), and heavy smokers (> 27 pack-years).

Selection of haplotype-based tag single nucleotide polymorphisms (SNPs)

Genotypes for SNPs in *CYP2A13* representing Han-Chinese were downloaded from the HapMap database (http://www.hapmap.org, HapMap Data Rel#24/phase II on NCBI B36 assembly, dbSNP b126). LD and haploptype in *CYP2A13* were determined using Haploview software version 4.1 (Broad Institute, Cambridge, MA, USA) by default value. LD was estimated among all pairs of SNPs using the D'statistic. Haplotype block structure was determined using the confidence interval (CI) option. Two tag SNPs, rs1645690 and rs8192789, were chosen to capture the common variants within *CYP2A13*.

DNA extraction and genotyping

Genomic DNA was isolated from peripheral blood using a standard kit-based method (Axygen, Corning Life Sciences, Tewksbury, MA, USA). The DNA concentration was adjusted to 40 mM by Tris ethylene-diamine-tetraacetic acid buffer, and all DNA preparations were stored at −20°C until used for genotyping. A TaqMan genotyping assay (Applied Biosystems, Foster City, CA, USA) was employed to genotype all samples for the two selected tag SNPs. For each of the SNPs, primer-probe sets were made using the Applied Biosystems design service (Table 1). We performed real-time PCR on 10 ng genomic DNA using TaqMan universal PCR master mix (Applied Biosystems), forward and reverse primers, and fluorescein amidite and 2′-chloro-7′phenyl-1,4-dichloro-6-carboxy-fluorescein (VIC)-labeled probes. Real-time PCR was performed using 5.0 uL universal master mix (Applied Biosystems), 0.25 uL primer-probe, 2.25 uL RNase-free and DNase-free water, and 2.5 uL sample DNA (40 mM). Assay conditions were 10 minutes at 95°C, 40 cycles of 92°C for 15 seconds, and 58°C for 1 minute. The Real-time PCR 7500 system (SDS version 1.4, Applied Biosystems) was used to perform and analyze genotyping. For the purpose of quality control, more than two negative controls containing all reagents, but with water instead of the DNA template, were included in each amplification set. Genotyping was carried out blinded to case–control status. To verify genotyping results, 10% of random samples were repeated. Each genotype of the two SNPs was cloned and sequenced randomly, and all were concordant with the judgment according to the results of real-time PCR.

Table 1 Primers and probes used for genotyping and cloning sequencing of tagging NSPs

dbSNP	Variants	Primers or probes sequence for genotyping	Primers for cloning sequencing (5′-3′)
rs1645690	A/G	F: GGCACGCACGGTGAGTA	F: TTCTCCATCGCCACCCTAA
		R: CCCATCCCCAGGCAGAAAG	R: CCTGTTGAGCCGAATCCC
		P1:^{VIC}- CGCGGGTTCCCGC –^{NFQ}	—
		P2:$^{FAM-}$ CGCGGGTCCCCGC –^{NFQ}	—
Rs8192789	A/G	F: GGAGGACTTCATCGCCAAGAA	F: CTGATCCATGAGGTCCTAGCC
		R: CGGATGAGAAAGGAGTCGATGA	R: TCCCGCTCAGAGTCCCCCCAA
		P1:^{VIC}- AGCGTGCGCTGGTT -^{NFQ} (reverse)	—
		P2:$^{FAM-}$ AGCGTGCACTGGTT –^{NFQ}	—

The polymorphic site is shaded. F, forward primer; NFQ, nonfluorescent quencher; P1, probe 1; P2, probe 2; R, reverse primer; SNP, single nucleotide polymorphism.

Statistical analysis

Demographic and clinical information between cases and controls was compared using the χ^2 test for categorical variables and the Student's *t*-test for continuous variables, where appropriate. Hardy–Weinberg equilibrium was confirmed by χ^2 analysis. The asymptotic Pearson's chi-square test was used to assess genotype frequencies different from those expected under Hardy–Weinberg equilibrium among Chinese controls. Odds ratios (ORs) and corresponding 95% CIs were calculated using logistic regression analysis where log odds of lung cancer were adjusted for smoking (a categorical variable), age (a continuous variable), and gender (a categorical variable). In order to detect important differences in the population subgroups, stratification by subgroup analysis of clinically relevant factors (smoking status, histologic type) was performed. All tests were two-sided and a P value ≤ 0.05 was considered significant. All analyses were performed using SPSS version 11.5 (SPSS Inc., Chicago, IL, USA).

Results

The demographics of the cases and controls enrolled in this study are shown in Table 2. There were no significant differences between the cases and controls in mean age and gender distribution, suggesting that the matching based on these variables was adequate. There was a higher prevalence of smokers and a significantly higher number of pack-years in the cases than in the controls (both $P < 0.001$). These differences were controlled for in multivariate analyses. There was a strong dose–response relationship between lung cancer risk and pack-years of smoking ($P < 0.001$), especially in the squamous carcinoma subgroup. After adjusting for age at diagnosis and gender, smokers had a statistically increased lung cancer risk (OR 3.30, 95% CI 2.18–4.99; $P < 0.001$) compared to never smokers in the overall cases. Further stratified analysis according to smoking status showed that the association was predominant in the heavy smoker subgroup (> 27 pack-years, OR 8.98, 95% CI 5.40–14.96; $P < 0.001$). Stratified by histologic type, smokers

Table 2 Characteristics of the study subjects

Characteristic	Cancer (n = 532)	Control (n = 614)	χ^2/t†	P
Age (years)				
Mean	60.39 ± 9.58	60.38 ± 10.05	0.017	0.987
Range	37–84	34–84		
Gender, n (%)				
Male	382 (71.8%)	436 (71%)	0.044	0.834
Female	150 (28.2%)	178 (29%)		
Smoking‡				
Never	174 (32.7%)	332 (54.1%)	86.37	0.000
Light-smoking	88 (16.5%)	198 (32.2%)		
Heavy smoking	254 (47.7%)	82 (13.4%)		
Missing	16 (3%)	2 (0.3%)		
Histologic type, N (%)				
SCC	240 (45.1%)			
AC	198 (37.2%)			
SCLC	46 (8.6%)			
Other	48 (9.0%)			

†Two-sided χ^2 or *t*-test, cases versus controls. ‡Pack-years (a measure of cumulative smoking exposure) was defined as the average number of packs (20 cigarettes/pack) of cigarettes smoked per day multiplied by the number of years of smoking. AC, adenocarcinoma; SCC, squamous carcinoma; SCLC, small cell lung cancer.

Table 3 The relationship between rs1645690 and lung cancer susceptibility (not stratified)

Genotype	Case	Control	OR (95% CI)	*P* value	*P* trend
AA	428 (80.5)	466 (75.9)	Ref		
AG	94 (17.7)	142 (23.1)	0.642 (0.403–1.021)†	0.248†	0.094†
GG	10 (1.9)	6 (1.0)	—	—	

†Adjusted by age, gender and smoking status. CI, confidence interval; OR, odds ratio.

Table 4 The relationship between rs1645690 and lung cancer susceptibility (non-smokers)

Genotype	Case	Control	OR (95% CI)	*P* value	*P* trend
AA	144 (85.7)	252 (75.9)	Ref		
AG	20 (11.9)	76 (22.9)	0.448 (0.208~0.967)†	0.041†	0.098†
GG	4 (2.4)	4 (1.2)	—	—	

†Adjusted by age, gender and smoking status. CI, confidence interval; OR, odds ratio.

had a statistically significant increased lung cancer risk (OR 8.28, 95% CI 4.24–16.20; $P < 0.001$), particularly in the squamous carcinoma subgroup. The genotype distributions of rs1645690 in the 1146 subjects are consistent with the data from HapMap Project phase 3 (when genotyping was completed, the HapMap data had been updated to phase 3). The genotype distributions among the controls were in Hardy–Weinberg equilibrium. The distributions were not different between cases and controls in the overall population, but were different in the non-smoker subgroup (Table 3). The G allele of rs1645690 statistically reduced lung cancer susceptibility in the non-smoker subgroup, a trend that was also observed in the adenocarcinoma and female subgroups but was not statistically significant in the female subgroup (Table 4). This result is concordant with the conclusion that smoking tobacco has a major association with squamous carcinoma, while other pollution types are significantly associated with adenocarcinoma. The genotype distributions of rs8192789 showed that three subjects were TT and all others were CT. Homologous sequences (*CYP2A6*) interfered with further cloning research, which was thus abandoned.

Discussion

In this study, we directly analyzed the association between *CYP2A13* polymorphisms and the risk of lung cancer by using the tagging SNP approach according to the HapMap Project. When the study began in 2008, the HapMap Project provided two common SNPs, rs1645690 and rs8192789, for Han-Chinese, the Estonian Genome Project did not provide any, and the dbSNP provided 107 SNPs and 11 common SNPs (minor allele frequency > 5%). Submissions to dbSNP come from a variety of sources including individual laboratories, collaborative polymorphism discovery efforts, and large-scale genome sequencing centers, and these SNPs are not verified. The HapMap Project is the most important functional genetic database from which Han-Chinese SNP information can be acquired.[20] In 2007, HapMap phase 2 data was released and reached an average of < 1 KB per SNP; accuracy of 99.8%; and included a higher accuracy, multi-racial genetic polymorphism map. In this study, we used the tag SNP approach according to information retrieved from the HapMap Project to select rs1645690 and rs8192789.

Our genotyping results showed that the CT frequency is 100%, which is consistent with the dbSNP database. Further cloning and sequencing results showed that the interference was caused by *CYP2A6*, thus the TaqMan-probe test was abandoned.[21] The HapMap database was updated to delete the rs8192789 locus (phase 3), but we think this point of SNPs should not have been deleted.[22–24] Interference caused by homologous sequences may be the main reason causing false-positive SNPs, except in regard to genotyping errors. We believe that the 139 SNPs of *CYP2A13* in the dbSNP database partly comprised false-positive SNPs caused by homologous sequences. The HapMap database contains the SNP distribution of the Han-Chinese population, but because of the huge workload, the published data may have suffered from oversights.

Two approaches have been proposed for the case–control study design. The traditional, hypothesis-driven approach investigates SNPs in coding regions, as they are more likely to have a functional role and to directly influence the traits under study. Another indirect approach is to select a set of haplotype-tagging SNPs that serve as markers to detect associations between a particular region and diseases, whether or not the SNPs themselves have a functional effect. It is not necessary to genotype all possible polymorphisms because the alleles of SNPs are in LD, these SNPs are physically close, and tend to be correlated with each other.[25] Although the SNPs rs1645690 was not in a coding region, in contrast to the haplotype-tagging SNPs, it reflects the common polymorphisms of *CYP2A13*. In this study, the genotype distributions of rs1645690 are consistent with HapMap Project phase 3 data. The G allele of

rs1645690 statistically reduced lung cancer susceptibility in the non-smoker subgroup, and the trend was observed in the adenocarcinoma and female subgroups. This result is consistent with the results of several previous studies, indicating that the common polymorphisms of *CYP2A13* influence lung cancer susceptibility.[26,27]

Wang *et al.* investigated the association between *CYP2A13* polymorphisms and lung cancer. They estimated that the Arg257Cys polymorphism (rs8192789) was related to adenocarcinoma risk, and further stratification analysis showed that the reduced risk was limited to smokers, especially light smokers (OR 0.23, 95% CI, 0.08–0.68).[28] Timofeeva *et al.* studied another tag SNP with lung cancer susceptibility, rs1709084, and their results suggested that the genetic polymorphism of *CYP2A13* may contribute to individual susceptibility to early-onset lung cancer in women aged > 51 years (OR 1.64, 95% CI 1.00–2.70; $P = 0.05$).[29] Using the *CYP2A13* specific antibody, Fukami *et al.* performed immunohistochemical analysis of human lung carcinomas. Their results showed that *CYP2A13* expression was markedly increased in non-small cell lung carcinomas, especially adenocarcinoma.[4] The high expression of *CYP2A13* might be associated with tumor development and progression in non-small cell lung carcinomas.

China is rapidly industrializing and lung cancer incidence has become the most common malignant tumor. New epidemiological subtypes, including women and adenocarcinoma, have obviously increased recently and should be considered, as they might reflect new etiological factors.[30,31] Because lung adenocarcinoma among Chinese women is not strongly linked to tobacco smoking, it is not surprising to see a null association between the risk of lung cancer among non-smokers and the genetic polymorphism in *CYP2A13*, which is a metabolic enzyme that mainly activates tobacco-specific NNK. These findings might also indicate that other carcinogenic factors involved in the etiology of cancer among non-smokers or light-smokers in this population are likely to be substrates of *CYP2A13*, such as indoor air pollution derived from Chinese-style cooking and/or coal burning.[32] Another factor may be gender.[29,33–35] Clinical data and related assays have shown that there are differences in susceptibility to carcinogen, the mechanism of lung cancer development and etiology of disease between genders.[36,37] Studies of animal models have shown that gender and sex hormones play an essential role both in normal lung development and pathological processes in lung tissue.[38] Cigarette exposure might be another explanation for the association between risk of lung cancer among non-smokers and the genetic polymorphism in *CYP2A13*.

We report results for a set of htSNPs selected from HapMap of *CYP2A13*. Our results suggest that HapMap correctly predicts genetic variation in our Chinese population: allele frequencies in the participants were similar to those obtained from HapMap, thus strengthening the argument for the widespread use of the database for htSNP selection. Our results also provide evidence that *CYP2A13* gene polymorphisms may be candidate biomarkers of lung cancer susceptibility in Chinese that can be used in future genome-wide association studies.

Acknowledgments

This work was supported by Shandong Province Natural Science Foundation of China (ZR2011HL028) and the National Natural Science Foundation of China (No. 81502668). The authors thank the physicians, surgeons and pathologists at Shandong Cancer Hospital and Institute who endorsed the project and the patients and donors who participated in this study.

Disclosure

No authors report any conflict of interest.

References

1 Hua F, Zhou Q. [Research progress of lung cancer on single nucleotide polymorphism]. *Zhongguo Fei Ai Za Zhi* 2011; **14**: 156–64. (In Chinese.).
2 Leclerc J, Courcot-Ngoubo NE, Cauffiez C *et al.* Xenobiotic metabolism and disposition in human lung: Transcript profiling in non-tumoral and tumoral tissues. *Biochimie* 2011; **93**: 1012–27.
3 Chiang HC, Wang CK, Tsou TC. Differential distribution of CYP2A6 and CYP2A13 in the human respiratory tract. *Respiration* 2012; **84**: 319–26.
4 Fukami T, Nakajima M, Matsumoto I, Zen Y, Oda M, Yokoi T. Immunohistochemical analysis of CYP2A13 in various types of human lung cancers. *Cancer Sci* 2010; **101**: 1024–8.
5 Chiang HC, Lee H, Chao HR, Chiou YH, Tsou TC. Pulmonary CYP2A13 levels are associated with early occurrence of lung cancer-its implication in mutagenesis of non-small cell lung carcinoma. *Cancer Epidemiol* 2013; **37**: 653–9.
6 He XY, Tang L, Wang SL, Cai QS, Wang JS, Hong JY. Efficient activation of aflatoxin B1 by cytochrome P450 2A13, an enzyme predominantly expressed in human respiratory tract. *Int J Cancer* 2006; **118**: 2665–71.
7 Zhu LR, Thomas PE, Lu G *et al.* CYP2A13 in human respiratory tissues and lung cancers: An immunohistochemical study with a new peptide-specific antibody. *Drug Metab Dispos* 2006; **34**: 1672–6.
8 Zhang Z, Lu H, Huan F *et al.* Cytochrome P450 2A13 mediates the neoplastic transformation of human bronchial epithelial cells at a low concentration of aflatoxin B1. *Int J Cancer* 2014; **134**: 1539–48.
9 Su T, Bao Z, Zhang QY, Smith TJ, Hong JY, Ding X. Human cytochrome P450 CYP2A13: Predominant

expression in the respiratory tract and its high efficiency metabolic activation of a tobacco-specific carcinogen, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Cancer Res* 2000; **60**: 5074–9.
10 von Weymarn LB, Brown KM, Murphy SE. Inactivation of CYP2A6 and CYP2A13 during nicotine metabolism. *J Pharmacol Exp Ther* 2006; **316**: 295–303.
11 Zhang X, D'Agostino J, Wu H *et al.* CYP2A13: Variable expression and role in human lung microsomal metabolic activation of the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *J Pharmacol Exp Ther* 2007; **323**: 570–8.
12 Chiang HC, Wang CY, Lee HL, Tsou TC. Metabolic effects of CYP2A6 and CYP2A13 on 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK)-induced gene mutation--a mammalian cell-based mutagenesis approach. *Toxicol Appl Pharmacol* 2011; **253**: 145–52.
13 Megaraj V, Zhou X, Xie F, Liu Z, Yang W, Ding X. Role of CYP2A13 in the bioactivation and lung tumorigenicity of the tobacco-specific lung procarcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone: in vivo studies using a CYP2A13-humanized mouse model. *Carcinogenesis* 2014; **35**: 131–7.
14 He XY, Shen J, Ding X, Lu AY, Hong JY. Identification of critical amino acid residues of human CYP2A13 for the metabolic activation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, a tobacco-specific carcinogen. *Drug Metab Dispos* 2004; **32**: 1516–21.
15 Wang SL, He XY, Shen J, Wang JS, Hong JY. The missense genetic polymorphisms of human CYP2A13: Functional significance in carcinogen activation and identification of a null allelic variant. *Toxicol Sci* 2006; **94**: 38–45.
16 Schlicht KE, Michno N, Smith BD, Scott EE, Murphy SE. Functional characterization of CYP2A13 polymorphisms. *Xenobiotica* 2007; **37**: 1439–49.
17 Xu Y, Shen Z, Shen J, Liu G, Li W, Tang Y. Computational insights into the different catalytic activities of CYP2A13 and CYP2A6 on NNK. *J Mol Graph Model* 2011; **30**: 1–9.
18 Xing J, Witherspoon DJ, Watkins WS, Zhang Y, Tolpinrud W, Jorde LB. HapMap tagSNP transferability in multiple populations: General guidelines. *Genomics* 2008; **92**: 41–51.
19 Surakka I, Kristiansson K, Anttila V *et al.* Founder population-specific HapMap panel increases power in GWA studies through improved imputation accuracy and CNV tagging. *Genome Res* 2010; **20**: 1344–51.
20 Phillips C. Online resources for SNP analysis: A review and route map. *Mol Biotechnol* 2007; **35**: 65–97.
21 Hua F, Wan H, Mei C *et al.* [Interference of homologous sequences on the SNP study of CYP2A13 gene]. *Zhongguo Fei Ai Za Zhi* 2010; **13**: 94–7. (In Chinese.).
22 Zhang X, Su T, Zhang QY *et al.* Genetic polymorphisms of the human CYP2A13 gene: Identification of single-nucleotide polymorphisms and functional characterization of an Arg257Cys variant. *J Pharmacol Exp Ther* 2002; **302**: 416–23.
23 Zhang X, Chen Y, Liu Y *et al.* Single nucleotide polymorphisms of the human cyp2a13 gene: Evidence for a null allele. *Drug Metab Dispos* 2003; **31**: 1081–5.
24 Cheng XY, Chen GL, Zhang WX, Zhou G, Wang D, Zhou HH. Arg257Cys polymorphism of CYP2A13 in a Chinese population. *Clin Chim Acta* 2004; **343**: 213–6.
25 Liu CT, Lin H, Lin H. Functional analysis of HapMap SNPs. *Gene* 2012; **511**: 358–63.
26 D'Agostino J, Zhang X, Wu H *et al.* Characterization of CYP2A13*2, a variant cytochrome P450 allele previously found to be associated with decreased incidences of lung adenocarcinoma in smokers. *Drug Metab Dispos* 2008; **36**: 2316–23.
27 Xiang C, Wang J, Kou X *et al.* Pulmonary expression of CYP2A13 and ABCB1 is regulated by FOXA2, and their genetic interaction is associated with lung cancer. *FASEB J* 2015; **29**: 1986–98.
28 Wang H, Tan W, Hao B *et al.* Substantial reduction in risk of lung adenocarcinoma associated with genetic polymorphism in CYP2A13, the most active cytochrome P450 for the metabolic activation of tobacco-specific carcinogen NNK. *Cancer Res* 2003; **63**: 8057–61.
29 Timofeeva MN, Kropp S, Sauter W *et al.* CYP450 polymorphisms as risk factors for early-onset lung cancer: Gender-specific differences. *Carcinogenesis* 2009; **30**: 1161–9.
30 Chang S, Dai M, Ren JS, Chen YH, Guo LW. [Estimates and prediction on incidence, mortality and prevalence of lung cancer in China in 2008]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2012; **33**: 391–4. (In Chinese.).
31 Han X, Qiao P, Xie M *et al.* [The incidence and mortality of lung cancer among residents in Yangpu district of Shanghai from 2002 to 2010]. *Zhonghua Zhong Liu Za Zhi* 2012; **34**: 712–8. (In Chinese.).
32 Kligerman S, White C. Epidemiology of lung cancer in women: Risk factors, survival, and screening. *AJR Am J Roentgenol* 2011; **196**: 287–95.
33 Dougherty SM, Mazhawidza W, Bohn AR *et al.* Gender difference in the activity but not expression of estrogen receptors alpha and beta in human lung adenocarcinoma cells. *Endocr Relat Cancer* 2006; **13**: 113–34.
34 Carey MA, Card JW, Voltz JW *et al.* It's all about sex: Gender, lung development and lung disease. *Trends Endocrinol Metab* 2007; **18**: 308–13.
35 Chakraborty S, Ganti AK, Marr A, Batra SK. Lung cancer in women: Role of estrogens. *Expert Rev Respir Med* 2010; **4**: 509–18.
36 Marquez-Garban DC, Mah V, Alavi M *et al.* Progesterone and estrogen receptor expression and activity in human non-small cell lung cancer. *Steroids* 2011; **76**: 910–20.
37 Kiyohara C, Ohno Y. Sex differences in lung cancer susceptibility: A review. *Gend Med* 2010; **7**: 381–401.
38 Ninomiya F, Yokohira M, Kishi S *et al.* Gender-dependent effects of gonadectomy on lung carcinogenesis by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in female and male A/J mice. *Oncol Rep* 2013; **30**: 2632–8.

