Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The association between polymorphisms in long non-coding RNA and lung cancer in non-smoking women

Guanghui Tong
China Medical University&Shengjing Hospital of China Medical University
Sixuan Li
China Medical University
Weiwei Tong
Shengjing Hospital of China Medical University
Min Gao
China Medical University
Baosen Zhou
China Medical University
Zhihua Yin (zhyin@cmu.edu.cn)
China Medical University https://orcid.org/0000-0003-0850-680X

Primary research

Keywords: LncRNA, Single nucleotide polymorphism, Lung cancer,Non-smoking

Posted Date: April 6th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-391055/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background

The relationship between long non-coding RNA and lung cancer has become a research hotspot.

Methods

Four polymorphisms, including rs10188946, rs11246867, rs2288947, and rs8105637 were evaluated in 556 patients with lung cancer and 395 age-matched controls in the present study.

Results

This study showed that the associations of the four polymorphisms of long non-coding RNA with the risks of lung cancer were not statistically significant. In the age stratification study, AG of rs2288947 was associated with the reduce risk of both lung cancer and adenocarcinoma (OR = 0.597,P = 0.017 and OR = 0.506, P = 0.005,respectively), and AG of rs8105637 was also protected both lung cancer and adenocarcinoma (OR = 0.636,P = 0.037and OR = 0.577, P = 0.023,respectively). We found that when the risk genotypes of the three SNPs rs10188946, rs11246867, rs2288947 and oil exposure worked together, the risk of lung cancer was higher than either of these two risk factors acted alone.

Conclusions

Four polymorphisms (including rs10188946, rs11246867, rs2288947, and rs8105637) were not associated with lung cancer risks in the present study. After age stratification, rs2288947 and rs8105637 were associated with the risks of lung cancer and adenocarcinoma among the individuals older than 60.

Introduction

Globally, according to statistics from 2008, lung cancer accounts for 13% (1.6 million) of the total cases and 18% (1.4 million) of deaths. Lung cancer is one of the most common cancers in the world. Among men, lung cancer is the leading cause of cancer death. Among women, lung cancer is the second leading cause of cancer death [1–2].In China, lung cancer is the most common cancer in both cities and rural areas. Despite the low smoking rate, the incidence of lung cancer in Chinese women is higher than that in women in some European countries [3].

Non-smoking related lung cancer is classified as an independent disease entity because it differs from smoking-associated lung cancer. Although smoking is the main cause of lung cancer, global statistics estimate that 15% of men and 53% of women who had lung cancer in 2000 were not attributed to smoking [4].Compared with men, the incidence of non-smoking related lung cancer in women is high, especially adenocarcinoma rather than squamous cell carcinoma [5–7].In one study, Asian women had a lower proportion of smokers than non-Hispanic white female smokers, but those women had a higher risk of lung cancer than expected, especially adenocarcinoma and Large cell undifferentiated carcinomas(6-fold and 4-fold, respectively).Female Chinese residents in the western United States (as Asian Chinese female residents) have a higher risk of lung cancer than their tobacco use [8].

According to the ENCODE project, about 87.3 % of the human genome is transcribed. Even though less than 3% of the human genome encodes proteins, most transcription consist of non-coding RNA. This discovery turned the focus to the long non-coding RNA (lncRNA) [9]. LncRNAs, which are polyadenylated, are larger than 200-nucleotide-long that are involved in gene expression [10–12]. Many studies have shown that long non-coding RNA was associated with the progression of various cancers [13–16]. Although lncRNAs in lung cancer are an emerging field, HOTAIR, H19, ANRIL and MALAT1 have been shown to be involved in lung cancer [17–20]. TINCR has been shown to be related with gastric cancer [21–22]. TINCR also affected lung cacner by interacting with BRAF to affect the activity of the oncogenic MAPK pathway [23]. Two polymorphism s in our study were from TINCR. Through this study, we will explore the relationship between single nucleotide polymorphisms (SNPs) in lncRNAs and the risk of lung cancer in non-smoking women.

Materials And Methods

The subject

This is a molecular epidemiological study of lung cancer. In this hospital-based case-control study, 556 patients with lung cancer and 395 controls were included. Both the case group and the control group were from four hospitals in Shenyang city, which locate in the northeast of China. The inclusion criteria for the case group were:(1) with a clear histological diagnosis of lung cancer, (2) new cases, (3) without receiving chemotherapy or radiotherapy, (4) non-smoking women.The exclusion criteria for the case group were:(1) previous cancer, (2) other parts of the tumor metastasized to the lungs.All participants were Han non-smoking women. The study was approved by the Institutional Review Board of China Medical University and informed consent was obtained from each participant.Each subject contributed 10 ml of venous blood and collected relevant baseline data when they were admitted to hospital.We defined people who have less than 100 cigarettes in their lifetime as non-smokers and the rest as smokers.

Genotyping

We isolated genomic DNA from venous blood samples using the phenol-chloroform method. We performed SNP genotyping using the previously described method [24].

Statistical analysis

We used the Pearson test and the T test to compare the categorical and continuous variables in the case and control groups, respectively. A Chi-square goodness-of-fit test was used to analyze Hardy Weinberg equilibrium. Binary Logistic regression and its 95% confidence interval were used to analyze the relationship between SNP genotypes and the risk of lung cancer. Binary Logistic regression was also used to analyze the interaction between genotypes and the environment. All statistical analyses were performed using SPSS20.0 software. The additive model of environmental gene interaction uses cross-analysis, and the multiplicative model uses multivariate logistic regression. In the analysis, we used a combination of protective genotypes and non-oil exposures as a reference. In the additive model, according to Anderson's report, we statistically analyzed the three biological indicators of RERI, AP, and S, and the 95% confidence interval of these three indicators [25].In order to obtain these three biological indicators, we first obtained three relative risk values and the corresponding covariance matrix obtained from a binary logistic regression.

Results

Subject characteristics

The study consisted of 556 case groups and 395 controls, all of whom were Non-exposure-smoking women. The average ages of the case group and the control group were 56.74±11.695 and 56.13±11.642, respectively. The difference in age between the case group and the control group was not statistically significant, and there was no difference in age between the case group and the control group (t=-0.797, P=0.905).Of the 556 patients in the case group, 371 (66.7%) were adenocarcinomas, 96 (17.3%) were squamous cell carcinomas, and 89 (16.0%) were small cell carcinomas. We observed that the polymorphisms of the four genotypes in the control group all conformed to Hardy-Weinberg's law (P=0.427 for rs10188946,P=1 for rs11246867, P=0.759 for rs2288947, P=0.608 for rs8105637).

The relationship between lncRNA polymorphisms and lung cancer, lung adenocarcinoma and lung squamous cell carcinoma

Table 1 shows that there is no statistically significant association between four SNPs and lung cancer risk. Table 2 shows that there is no statistically significant association between the four SNPs and the risk of adenocarcinoma and squamous cell carcinoma. Due to various factors such as immune aging, the incidence of cancer in older people is higher than that of young people [26-27]. So we made a stratified analysis of age. Tables 3 and 4 reflect the relationship between these four SNPs and the risk of lung cancer and adenocarcinoma. When the age is greater than 60, we found that rs2288947 is associated with the risk of lung cancer, the risk of heterozygous AG was lower than that of AA (OR=0.597, P=0.017). The recessive model was associated with the risk of lung cancer, that is, the risk of genotype GG was higher than that of AA+AG (OR=2.887, P=0.038).We also found that rs2288947 is associated with the risk of adenocarcinoma, and the risk of heterozygous AG is lower than that of AA (OR=0.506, P=0.005). The dominant model is associated with the risk of adenocarcinoma, that is, the risk of genotype GG+AG lower than AA (OR=0.598.P=0.025). We also found that rs8105637 was associated with the risk of lung cancer and the risk of heterozygous AG was lower than that of GG (OR=0.636, P=0.037). We found that rs8105637 was associated with the risk of adenocarcinoma, and the risk of heterozygous AG was lower than that of GG (OR=0.577, P=0.023).

Tables 5 and 6 reflect the interaction of oil exposure with these four SNPs. In our study, the number of people exposed to oil in the case group was 100 (37.3%), and the number of oil exposed in the control group was 66 (24.8%). The number of oil exposed in the case group was more than that of the control group (= 9.739, P=0.002). The risk of lung cancer in exposed was higher than in those without exposure (OR=1.804, 95%CI=1.243-2.618, P=0.002). We also found that for the AG+AA genotype of the gene rs10188946,the risk of lung cancer from exposure to oil was higher than that from Non-exposure-oil exposure (OR=1.912, P=0.009). Furthermore, the risk of lung cancer was higher in patients whose genotypes were GG and exposed to oil than those whose genotypes were AG+AA but not exposed to oil (OR=2.000, P=0.014).

For the GG genotype of rs11246867, the risk of lung cancer from exposure to oil was higher than that from non-exposure-oil exposure (OR=1.736, P=0.005). The risk of lung cancer in AG+AA genotypes exposed to oil was higher than that in GG genotypes that were not exposed to oil (OR=3.325, P=0.043). For the AG+GG genotype of the gene rs2288947, the risk of lung cancer was higher in oil exposure than non-exposure-oil exposure (OR=2.782, P=0.000). Moreover, the risk of lung cancer was higher in patients whose AA genotypes were exposed to oil at the same time than those whose AG+GG genotypes were not exposed to oil (OR=1.778, P=0.034). For the AG+AA genotype of the gene rs8105637, exposure to oil was more likely to have lung cancer than those without exposure to oil (OR=2.783, P=0.000). We did not find that the additive models of gene environment interactions make sense.

In the multiplication model of gene-environment interaction, we found that the negative multiplication effect of the SNP rs2288947 and rs8105637 were statistically significant (OR=0.459, 95%CI=0.216-0.977, P=0.043; OR=0.416, 95%CI=0.196-0.883, P=0.022 respectively). For the remaining SNPs, we did not find that the multiplication models of gene environment interactions make sense [OR, 95% CI, P values were: 0.860 (0.403, 1.836) and 0.697 for rs10188946;1.584 (0.389, 6.454) and 0.521 for rs11246867].

Discussion

As far as we know, the factors that affect lung cancer are very complicated. The influence of smoking on lung cancer is more deeply rooted in many factors. And it is estimated that the relative risk of lung cancer in long-term smokers is 10–30 times larger than that of non-smokers [28]. Even so, we still find about 5% of lung cancer cases among non-smokers[29], and the smoking rate of Chinese women is not high (only 2.4% of women older than 15 years old) [30], so we need to study the factors affecting lung cancer in non-smoking women. The same, factors that affect lung cancer in non-smokers are also complex. In our study, we investigated the relationship between four genotypes and lung cancer in female non-smokers.

The study by Yongbin Zheng et al. showed that the rs2288947 and rs8105637 were associated with the susceptibility of colorectal cancer, and allele G was a protective factor for rs2288947 (OR = 0.77, 95% CI = 0.67–0.88, P = 0.00012) and allele A was a risk factor for rs8105637 (OR = 1.22, 95% CI = 1.09–1.37, P = 0.00062). Rs2288947 and rs8105637 were associated with the occurrence of lymphatic metastasis in colorectal cancer, while allele G decreased the risk of lymphatic metastasis (OR = 0.77, 95%CI = 0.63–0.94, P = 0.011) and allele A increased the risk of lymphatic metastasis (OR = 1.22, 95%CI = 1.03–1.43, P = 0.019)[31].Ma et al.'s study showed that rs2288947 was associated with the risk of gastric cancer in Chinese. Compared with genotypes AG, GG, AG + GG, and allele G, AA genotype and allele A increased the occurrence of gastric cancer .Genotype polymorphisms are associated with gastric cancer most in young people, men, and non-smokers [21].In our study, rs2288947 and rs8105637 were associated with lung cancer and adenocarcinoma. When the age was greater than 60 years, We found that rs2288947 and rs8105637 were associated with the risk of lung cancer, while the risk of heterozygous AG for rs2288947 was lower than that of AA (OR = 0.597, P = 0.017) and the risk of heterozygous AG for rs2288947 was lower than that of GG (OR = 0.636, P = 0.037) .We also found that rs2288947 and rs8105637 were associated with the risk of adenocarcinoma, while the risk of heterozygous AG for rs2288947 was lower than that of AA (OR = 0.506, P = 0.005) and the risk of heterozygous AG for rs8105637 was lower than that of GG (OR = 0.577, P = 0.023). The dominant model of rs2288947 was associated with the risk of lung cancer, that is, the risk of genotype GG + AG was lower than that of AA (OR = 0.598. P = 0.025).The population in our study was older than 60, women, and non-smokers. The AG genotype of the rs2288947 gene reduced the incidence of gastric cancer among young people, men, and non-smokers, and decreased the incidence of lung cancer in people older than 60 years, women, and nonsmokers. The AG genotype of the rs8105637 was a risk factor, while in our study AG genotype was a protective factor. Zheng Y et al. found that TINCR can bind with STAU1 in order to influence the stability of CDKN2B mRNA. That means the transcription of TINCR accelerated gastric cancer[22].

Studies have shown that the air pollution of Chinese cooking was associated with lung cancer [32, 33]. Zhong L et al.'s population-based case-control study confirmed that people who exposed to indoor air pollution in Chinese cooking may increase the risk of lung cancer [34–36]. Our study also explored the interaction between the four genotypes and oil exposure in non-smoking females in China. In our study, we found that the negative multiplication models between the rs2288947, rs8105637 and the environment were meaningful and we did not find the additive models of four SNPs interacting with the environment are meaningful.

Our study also had some limitations. All of our subjects came from hospitals. Even if the subjects came from different hospitals, the Berkson bias still existed. Since we were researching non-smoking women, our sample size may not be particularly large because of this conditional restrictions. Therefore, we hope to increase the sample size for more powerful statistical analysis. This study was just a statistical analysis, the specific mechanism needs further experimental study.

Conclusions

Our study did not find any of these four SNPs associated with lung cancer. According to age stratification, we found that rs2288947 and rs8105637 were associated with lung cancer and adenocarcinoma when the age was older than 60.

Abbreviations

LncRNA long non-coding RNA

SNP Single-nucleotide polymorphism

Declarations

Ethics approval and consent to participate

The study was approved by the Institutional Review Board of China Medical University and informed consent was obtained from each participant.Each subject contributed 10 ml of venous blood and collected relevant baseline data when they were admitted to hospital.

Consent for publication

All of the authors have read and approved the content, and agree to submit the whole article in your journal.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

Not applicable.

Authors' contributions

YZ and ZB: conceptualization. TG, LS and TW: methodology. GM and YZ: validation. TG, LS, GM and YZ: formal analysis. TG, LS and TW: writing—original draft preparation. YZ: writing—review and editing. YZ and ZB: project administration.

Acknowledgements

Not applicable.

Footnotes

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1.Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A.Global cancer statistics, 2012.CA Cancer J Clin.2015;65(2):87-108.

2.Siegel RL, Fedewa SA, Miller KD, Goding-Sauer A, Pinheiro PS, Martinez-Tyson D, *et al*.Cancer statistics for Hispanics/Latinos, 2015.CA Cancer J Clin.2015;65(6):457-80.

3.Hong QY, Wu GM, Qian GS, Hu CP, Zhou JY, Chen LA, *et al*.Prevention and management of lung cancer in China.Cancer.2015;121 Suppl 17:3080-8.

4.Parkin DM, Pisani P, Ferlay J.Global cancer statistics.CA Cancer J Clin.1999;49(1):33-64, 1.

5.Gou LY, Niu FY, Wu YL, Zhong WZ.Differences in driver genes between smoking-related and non-smoking-related lung cancer in the Chinese population.Cancer.2015;121 Suppl 17:3069-79.

6.North CM, Christiani DC.Women and lung cancer: what is new?Semin Thorac Cardiovasc Surg.2013;25(2):87-94.

7.Ben Aissa A, Mach N.[Is lung cancer in women different?].Rev Med Suisse.2012;8(342):1108-11.

8.Epplein M, Schwartz SM, Potter JD, Weiss NS.Smoking-adjusted lung cancer incidence among Asian-Americans (United States).Cancer Causes Control.2005;16(9):1085-90.

9.Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A*, et al.*Landscape of transcription in human cells.Nature.2012;489(7414):101-8.

10.Enfield KS, Pikor LA, Martinez VD, Lam WL.Mechanistic Roles of Noncoding RNAs in Lung Cancer Biology and Their Clinical Implications.Genet Res Int.2012;2012:737416.

11.Huang Y, Liu N, Wang JP, Wang YQ, Yu XL, Wang ZB*, et al.*Regulatory long non-coding RNA and its functions.J Physiol Biochem.2012;68(4):611-8.

12.Mercer TR, Dinger ME, Mattick JS.Long non-coding RNAs: insights into functions.Nat Rev Genet.2009;10(3):155-9.

13.Martens-Uzunova ES, Bottcher R, Croce CM, Jenster G, Visakorpi T, Calin GA.Long noncoding RNA in prostate, bladder, and kidney cancer.Eur Urol.2014;65(6):1140-51.

14.Zhang H, Chen Z, Wang X, Huang Z, He Z, Chen Y.Long non-coding RNA: a new player in cancer.J Hematol Oncol.2013;6:37.

15.Zhong R, Tian J, Fu M, Ma S, Liu L, Li J*, et al.*LINC01149 variant modulates MICA expression that facilitates hepatitis B virus spontaneous recovery but increases hepatocellular carcinoma risk.Oncogene.2020;39(9):1944-56.

16.Tian J, Lou J, Cai Y, Rao M, Lu Z, Zhu Y*, et al.*Risk SNP-Mediated Enhancer-Promoter Interaction Drives Colorectal Cancer through Both FADS2 and AP002754.2.Cancer Res.2020;80(9):1804-18.

17.Ricciuti B, Mencaroni C, Paglialunga L, Paciullo F, Crino L, Chiari R*, et al.*Long noncoding RNAs: new insights into non-small cell lung cancer biology, diagnosis and therapy.Med Oncol.2016;33(2):18.

18.Cui J, Mo J, Luo M, Yu Q, Zhou S, Li T*, et al.*c-Myc-activated long non-coding RNA H19 downregulates miR-107 and promotes cell cycle progression of non-small cell lung cancer.Int J Clin Exp Pathol.2015;8(10):12400-9.

19.Nie FQ, Sun M, Yang JS, Xie M, Xu TP, Xia R*, et al.*Long noncoding RNA ANRIL promotes non-small cell lung cancer cell proliferation and inhibits apoptosis by silencing KLF2 and P21 expression.Mol Cancer Ther.2015;14(1):268-77.

20.Shen L, Chen L, Wang Y, Jiang X, Xia H, Zhuang Z.Long noncoding RNA MALAT1 promotes brain metastasis by inducing epithelial-mesenchymal transition in lung cancer.J Neurooncol.2015;121(1):101-8.

21.Ma X, Huang C, Luo D, Wang Y, Tang R, Huan X*, et al.*Tag SNPs of long non-coding RNA TINCR affect the genetic susceptibility to gastric cancer in a Chinese population.Oncotarget.2016;7(52):87114-23.

22.Xu TP, Wang YF, Xiong WL, Ma P, Wang WY, Chen WM*, et al.*E2F1 induces TINCR transcriptional activity and accelerates gastric cancer progression via activation of TINCR/STAU1/CDKN2B signaling axis.Cell Death Dis.2017;8(6):e2837.

23.Zhu ZJ, He JK.TINCR facilitates non-small cell lung cancer progression through BRAF-activated MAPK pathway.Biochem Biophys Res Commun.2018;497(4):971-7.

24.Lan Q, Hsiung CA, Matsuo K, Hong YC, Seow A, Wang Z*, et al.*Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia.Nat Genet.2012;44(12):1330-5.

25.Andersson T, Alfredsson L, Kallberg H, Zdravkovic S, Ahlbom A.Calculating measures of biological interaction.Eur J Epidemiol.2005;20(7):575-9.

26.Bueno V, Sant'Anna OA, Lord JM.Ageing and myeloid-derived suppressor cells: possible involvement in immunosenescence and age-related disease.Age (Dordr).2014;36(6):9729.

27.Gore E, Movsas B, Santana-Davila R, Langer C.Evaluation and management of elderly patients with lung cancer.Semin Radiat Oncol.2012;22(4):304-10.

28.Samet JM.Health benefits of smoking cessation.Clin Chest Med.1991;12(4):669-79.

29.Pallis AG, Syrigos KN.Lung cancer in never smokers: disease characteristics and risk factors.Crit Rev Oncol Hematol.2013;88(3):494-503.

30.Yang GH, Li Q, Wang CX, Hsia J, Yang Y, Xiao L, *et al*.Findings from 2010 Global Adult Tobacco Survey: implementation of MPOWER policy in China.Biomed Environ Sci.2010;23(6):422-9.

31.Zheng Y, Yang C, Tong S, Ding Y, Deng W, Song D, *et al*.Genetic variation of long non-coding RNA TINCR contribute to the susceptibility and progression of colorectal cancer.Oncotarget.2017;8(20):33536-43.

32.Wang GX.[Multivariate analyses of causal factors included cooking oil fume and others in matched case-control study of lung cancer].Zhonghua Yu Fang Yi Xue Za Zhi.1992;26(2):89-91.

33.Lee T, Gany F.Cooking oil fumes and lung cancer: a review of the literature in the context of the U.S. population.J Immigr Minor Health.2013;15(3):646-52.

34.Zhong L, Goldberg MS, Gao YT, Jin F.Lung cancer and indoor air pollution arising from Chinese-style cooking among nonsmoking women living in Shanghai, China.Epidemiology.1999;10(5):488-94.

35.Ko YC, Cheng LS, Lee CH, Huang JJ, Huang MS, Kao EL, *et al*.Chinese food cooking and lung cancer in women nonsmokers.Am J Epidemiol.2000;151(2):140-7.

36.Xue Y, Jiang Y, Jin S, Li Y.Association between cooking oil fume exposure and lung cancer among Chinese nonsmoking women: a meta-analysis.Onco Targets Ther.2016;9:2987-92.

Tables

Table 1 Relationship between four SNPs and lung cancer risk

SNP	Cases (%)	Controls (%)	OR (95%CI)	P value
rs10188946				
GG(ref)	233(41.9%)	154(39.0%)	1.00(ref)	-
AG	255(45.9%)	195(49.4%)	0.864(0.656,1.139)	0.300
AA	68(12.2%)	46 (11.6%)	0.977(0.638 ,1.496)	0.915
Dominant model AG+AA vs GG			0.886(0.681,1.153)	0.367
Recessive model AA vs GG+AG			1.057(0.710,1.575)	0.784
G allele(ref)	721(64.8%)	503(63.7%)	1.00(ref)	-
A allele	391(35.2%)	287(36.3%)	0.950(0.786,1.150)	0.600
rs11246867				
GG(ref)	496(89.2%)	359(90.9%)	1.00(ref)	-
AG	58(10.4%)	35(8.9%)	1.199(0.772,1.864)	0.419
AA	2(0.4%)	1(0.2%)	1.448(0.131,16.026)	0.763
Dominant model AG+AA vs GG			1.206(0.781,1.864)	0.398
Recessive model AA vs GG+AG			1.422(0.129,15.741)	0.774
G allele(ref)	1050(94.4%)	753(95.3%)	1.00(ref)	-
A allele	62(5.6%)	37(4.7%)	1.202(0.791,1.825)	0.389
rs2288947				
AA(ref)	319(57.4%)	217(55.0%)	1.00(ref)	-
AG	193(34.7%)	155(39.2%)	0.847(0.645,1.113)	0.233
GG	44(7.9%)	23(5.8%)	1.301(0.764,2.218)	0.333
Dominant model AG+GG vs AA			0.906(0.698,1.175)	0.455
Recessive model GG vs AA+AG			1.390(0.825,2.342)	0.216
A allele (ref)	831(74.7%)	589(74.6%)	1.00(ref)	-
G allele	281(25.3%)	201(25.4%)	0.991(0.804,1.222)	0.932
rs8105637				
GG(ref)	298(53.6%)	201(50.9%)	1.00(ref)	-
AG	210(37.8%)	167(42.3%)	0.848(0.647,1.112)	0.233
AA	48(8.6%)	27(6.8%)	1.199(0.724,1.985)	0.480
Dominant model AG+AA vs GG			0.897(0.693,1.161)	0.409
Recessive model AA vs GG+AG			1.288(0.789,2.103)	0.312
G allele(ref)	806(72.5%)	569(72.0%)	1.00(ref)	-
A allele	306(27.5%)	221(28.0%)	0.977(0.797,1.198)	0.826

Table 2 Relationship between four SNPs and adenocarcinoma and squamous cell cancer risk

SNP	adenocarcinoma				squamous cell			
	cases	Controls	OR(95%CI)	P	cases	controls	OR(95%CI)	P
rs10188946								
GG(ref)	152(41.0%)	154(39.0%)	1.00(ref)	-	43(44.8%)	154(39.0%)	1.00(ref)	-
AG	175(47.2%)	195(49.4%)	0.909(0.672,1.231)	0.538	41(42.7%)	195(49.4%)	0.753(0.467,1.213)	0.244
AA	44(11.9%)	46(11.6%)	0.969(0.606,1.551)	0.896	12(12.5%)	46(11.6%)	0.934(0.455,1.919)	0.853
Dominant model AG+AA vs GG			0.921(0.689,1.230)	0.576			0.788(0.502,1.235)	0.299
Recessive model AA vs GG+AG			1.021(0.658,1.585)	0.927			1.084(0.550,2.136)	0.816
G allele (ref)	479(64.6%)	503(63.7%)	1.00(ref)	-	127(66.1%)	503(63.7%)	1.00(ref)	-
A allele	263(35.4%)	287(36.3%)	0.962(0.781,1.186)	0.718	65(33.9%)	287(36.3%)	0.897(0.643,1.250)	0.521
rs11246867								
GG(ref)	334(90.0%)	359(90.9%)	1.00(ref)	-	88(91.7%)	359(90.0%)	1.00(ref)	-
AG	37(10.0%)	35(8.9%)	1.136(0.699,1.847)	0.606	7(7.3%)	35(8.9%)	0.816(0.351,1.898)	0.637
AA	0(0.0%)	1(0.3%)	-*	1.000	1(1.0%)	1(0.3%)	4.080(0.253,65.863)	0.322
Dominant model AG+AA vs GG			1.105(0.682,1.790)	0.686			0.907(0.407,2.019)	0.810
Recessive model AA vs GG+AG			-*	1.000			4.147(0.257,66.908)	0.316
G allele (ref)	705(95.0%)	753(95.3%)	1.00(ref)	-	183(95.3%)	753(95.3%)	1.00(ref)	-
A allele	37(5.0%)	37(4.7%)	1.068(0.669,1.704)	0.782	9(4.7%)	37(4.7%)	1.001(0.475,2.111)	0.998
rs2288947								
AA(ref)	217(58.5%)	217(54.9%)	1.00(ref)	-	59(61.5%)	217(54.9%)	1.00(ref)	-
AG	128(34.5%)	155(39.2%)	0.826(0.612,1.115)	0.212	31(32.3%)	155(39.2%)	0.736(0.455,1.190)	0.211
GG	26 (7.0%)	23(5.8%)	1.130(0.626,2.043)	0.685	6(6.2%)	23(5.8%)	0.959(0.374,2.465)	0.932
Dominant model AG+GG vs AA			0.865(0.650,1.152)	0.321			0.765(0.484,1.207)	0.249
Recessive model GG vs AG+AA			1.219(0.683,2.177)	0.503			1.078(0.426,2.726)	0.873
A allele(ref)	562(75.7%)	589(74.6%)	1.00(ref)	-	149(77.6%)	589(74.6%)	1.00(ref)	-
G allele	180(24.3%)	201(25.4%)	0.939(0.744,1.184)	0.592	43(22.4%)	201(25.4%)	0.846(0.581,1.231)	0.381
rs8105637								
GG(ref)	196(52.8%)	201(50.9%)	1.00(ref)	-	54(56.2%)	201(50.9%)	1.00(ref)	-
AG	142(38.3%)	167(42.3%)	0.872(0.647,1.175)	0.368	36(37.5%)	167(42.3%)	0.802(0.502,1.283)	0.358
AA	33(8.9%)	27(6.8%)	1.253(0.727,2.162)	0.417	6(6.2%)	27(6.8%)	0.827(0.325,2.105)	0.691
Dominant model AG+AA vs GG			0.925(0.697,1.228)	0.590			0.806(0.514,1.262)	0.346
Recessive model AA vs GG+AG			1.331(0.784,2.260)	0.290			0.909(0.364,2.267)	0.837
G allele (ref)	534(72.0%)	569(72.0%)	1.00(ref)	-	144(75.0%)	569(72.0%)	1.00(ref)	-
A allele	208(28.0%)	221(28.0%)	1.003(0.802,1.254)	0.980	48(25.0%)	221(28.0%)	0.858(0.598,1.232)	0.407

-*: OR could not be calculated.

Table 3 Relationship between these four SNPs and lung cancer based on age stratification

SNP	>60				<=60			
	cases	controls	OR(95%CI)	P	cases	controls	OR(95%CI)	P
rs10188946								
GG(ref)	100(43.1%)	55(34.8%)	1.00(ref)	-	133(41.0%)	99(41.8%)	1.00(ref)	-
AG	109(47.0%)	84(53.2%)	0.714(0.462,1.103)	0.129	146(45.1%)	111(46.8%)	0.979(0.684,1.401)	0.908
AA	23(9.9%)	19(12.0%)	0.666(0.334,1.329)	0.249	45(13.9%)	27(11.4%)	1.241(0.720,2.136)	0.437
Dominant model AG+AA vs GG			0.705(0.464,1.071)	0.101			1.030(0.733,1.448)	0.864
Recessive model AA vs GG+AG			0.805(0.423,1.534)	0.510			1.254(0.754,2.088)	0.383
G allele(ref)	309(66.6%)	194(61.4%)	1.00(ref)	-	412(63.6%)	309(65.2%)	1.00(ref)	-
A allele	155(33.4%)	122(38.6%)	0.798(0.592,1.074)	0.136	236(36.4%)	165(34.8%)	1.073(0.837,1.374)	0.578
rs11246867								
GG(ref)	211(90.9%)	142(89.9%)	1.00(ref)	-	285(88.0%)	217(91.6%)	1.00(ref)	-
AG	20(8.6%)	15(9.5%)	0.897(0.445,1.811)	0.762	38(11.7%)	20(8.4%)	1.447(0.819,2.557)	0.204
AA	1 (0.4%)	1(0.6%)	0.673(0.042,10.847)	0.780	1(0.3%)	0(0.0%)	-*	1.000
Dominant model AG+AA vs GG			0.883(0.446,1.751)	0.722			1.485(0.842,2.618)	0.172
Recessive model AA vs GG+AG			0.680(0.042,10.947)	0.785			-*	1.000
G allele(ref)	442(95.3%)	299(94.6%)	1.00(ref)	-	608(93.8%)	454(95.8%)	1.00(ref)	-
A allele	22(4.7%)	17(5.4%)	0.875(0.457,1.676)	0.688	40(6.2%)	20(4.2%)	1.493(0.861,2.589)	0.153
rs2288947								
AA(ref)	136(58.6%)	79(50.0%)	1.00(ref)	-	183(56.5%)	138(58.2%)	1.00(ref)	—
AG	76(32.8%)	74 (46.8%)	0.597(0.391,0.911)	0.017	117(36.1%)	81(34.2%)	1.089(0.760,1.560)	0.641
GG	20(8.6%)	5(3.2%)	2.324(0.839,6.434)	0.105	24(7.4%)	18(7.6%)	1.005(0.525,1.926)	0.987
Dominant model AG+GG vs AA			0.706(0.470,1.060)	0.093			1.074(0.765,1.507)	0.680
Recessive model GG vs AA+AG			2.887(1.060,7.861)	0.038			0.973(0.516,1.838)	0.934
A allele(ref)	348(75.0%)	232(73.4%)	1.00(ref)	-	483(74.5%)	357(75.3%)	1.00(ref)	-
G allele	116(25.0%)	84(26.6%)	0.921(0.664,1.276)	0.619	165(25.5%)	117(24.7%)	1.042(0.793,1.370)	0.766
rs8105637								
GG(ref)	127(54.7%)	73(46.2%)	1.00(ref)	-	171(52.8%)	128(54.0%)	1.00(ref)	-
AG	83(35.8%)	75(47.5%)	0.636(0.416,0.973)	0.037	127(39.2%)	92(38.8%)	1.033(0.726,1.470)	0.856
AA	22(9.5%)	10(6.3%)	1.265(0.568,2.817)	0.566	26(8.0%)	17(7.2%)	1.145(0.596,2.199)	0.685
Dominant model AG+AA vs GG			0.710(0.473,1.065)	0.098			1.051(0.751,1.470)	0.773
Recessive model AA vs GG+AG			1.550(0.713,3.371)	0.268			1.129(0.598,2.132)	0.708
G allele(ref)	337(72.6%)	221(69.9%)	1.00(ref)	-	469(72.4%)	348(73.4%)	1.00(ref)	-
A allele	127(27.4%)	95(30.1%)	0.877(0.640,1.202)	0.413	179(27.6%)	126(26.6%)	1.054(0.807,1.376)	0.699

*: OR could not be calculated.

Table 4 Relationship between these four SNPs and adenocarcinoma based on age stratification

SNP	>60				<=60			
	cases	controls	OR(95%CI)	P	cases	controls	OR(95%CI)	P
rs10188946								
GG(ref)	66(42.6%)	55(34.8%)	1.00(ref)	-	86(39.8%)	99(41.8%)	1.00(ref)	-
AG	75(48.4%)	84(53.2%)	0.744(0.463,1.196)	0.222	100(46.3%)	111(46.8%)	1.037(0.698,1.540)	0.857
AA	14(9.0%)	19(12.0%)	0.614(0.282,1.336)	0.219	30(13.9%)	27(11.4%)	1.279(0.706,2.319)	0.417
Dominant model AG+AA vs GG			0.720(0.456,1.137)	0.159			1.084(0.745,1.578)	0.672
Recessive model AA vs GG+AG			0.726(0.350,1.506)	0.390			1.254(0.719,2.188)	0.424
G allele(ref)	207(66.8%)	194(61.4%)	1.00(ref)	-	272(63.0%)	309(65.2%)	1.00(ref)	-
A allele	103(33.2%)	122(38.6%)	0.791(0.570,1.098)	0.161	160(37.0%)	165(34.8%)	1.102(0.839,1.446)	0.485
rs11246867								
GG(ref)	141(91.0%)	142(89.9%)	1.00(ref)	-	193(89.4%)	217(91.6%)	1.00(ref)	-
AG	14(9.0%)	15(9.5%)	0.940(0.438,2.019)	0.874	23(10.6%)	20(8.4%)	1.293(0.689,2.427)	0.424
AA	0(0.0%)	1(0.6%)	-*	1.000	0(0.0%)	0(0.0%)	-*	-*
Dominant model AG+AA vs GG			0.881(0.415,1.873)	0.742			1.293(0.689,2.427)	0.424
Recessive model AA vs GG+AG			-*	1.000			-*	-*
G allele(ref)	296(95.5%)	299(94.6%)	1.00(ref)	-	409(94.7%)	454(95.8%)	1.00(ref)	-
A allele	14(4.5%)	17(5.4%)	0.832(0.403,1.718)	0.619	23(5.3%)	20(4.2%)	1.277(0.691,2.359)	0.436
rs2288947								
AA(ref)	97(62.6%)	79(50.0%)	1.00(ref)	-	120(55.6%)	138(58.2%)	1.00(ref)	-
AG	46(29.7%)	74(46.8%)	0.506(0.315,0.812)	0.005	82(38.0%)	81(34.2%)	1.164(0.786,1.724)	0.448
GG	12(7.7%)	5(3.2%)	1.955(0.661,5.783)	0.226	14(6.5%)	18(7.6%)	0.894(0.427,1.875)	0.768
Dominant model AG+GG vs AA			0.598(0.381,0.938)	0.025			1.115(0.768,1.618)	0.566
Recessive model GG vs AA+AG			2.568(0.883,7.470)	0.083			0.843(0.409,1.740)	0.644
A allele(ref)	240(77.4%)	232(73.4%)	1.00(ref)	-	322(74.5%)	357(75.3%)	1.00(ref)	-
G allele	70(22.6%)	84(26.6%)	0.806(0.559,1.160)	0.246	110(25.5%)	117(24.7%)	1.042(0.772,1.408)	0.787
rs8105637								
GG(ref)	86(55.5%)	73(46.2%)	1.00(ref)	-	110(50.9%)	128(54.0%)	1.00(ref)	-
AG	51(32.9%)	75(47.5%)	0.577(0.360,0.926)	0.023	91(42.1%)	92(38.8%)	1.151(0.783,1.693)	0.475
AA	18(11.6%)	10(6.3%)	1.528(0.664,3.517)	0.319	15(6.9%)	17(7.2%)	1.027(0.490,2.151)	0.944
Dominant model AG+AA vs GG			0.689(0.441,1.075)	0.101			1.132(0.782,1.637)	0.512
Recessive model AA vs GG+AG			1.945(0.867,4.359)	0.106			0.966(0.470,1.984)	0.924
G allele(ref)	223(71.9%)	221(69.9%)	1.00(ref)	-	311(72.0%)	348(73.4%)	1.00(ref)	-
A allele	87(28.1%)	95(30.1%)	0.908(0.643,1.282)	0.582	121(28.0%)	126(26.6%)	1.075(0.802,1.440)	0.630

*: OR could not be calculated.

Table 5 Cross-over Analysis of Oil Fume Exposure and the four SNPs

SNP	oil	controls	cases	OR(95%CI)	P
rs10188946					
AG+AA	Non-exposure	126	98	1.00(ref)	-
GG	Non-exposure	74	70	1.216(0.799,1.851)	0.361
AG+AA	Exposure	39	58	1.912(1.178,3.103)	0.009
GG	Exposure	27	42	2.000(1.153,3.469)	0.014
rs11246867					
GG	Non-exposure	185	153	1.00(ref)	-
AG+AA	Non-exposure	15	15	1.209(0.573,2.552)	0.618
GG	Exposure	62	89	1.736(1.177,2.560)	0.005
AG+AA	Exposure	4	11	3.325(1.038,10.652)	0.043
rs2288947					
AG+GG	Non-exposure	91	63	1.00(ref)	-
AA	Non-exposure	109	105	1.391(0.916,2.114)	0.122
AG+GG	Exposure	27	52	2.782(1.581,4.895)	0.000
AA	Exposure	39	48	1.778(1.046,3.023)	0.034
rs8105637					
AG+AA	Non-exposure	100	77	1.00(ref)	—
GG	Non-exposure	100	91	1.182(0.784,1.783)	0.426
AG+AA	Exposure	28	60	2.783(1.625,4.766)	0.000
GG	Exposure	38	40	1.367(0.801,2.332)	0.251

Table 6 Three indicators and their confidence intervals for additive interactions between oil exposure and these four SNPs

SNP	Measure	Estimate	95%CI
rs10188946	RERI	-0.128	-1.464,1.207
	AP	-0.064	-0.751,0.623
	S	0.886	0.253,3.107
rs11246867	RERI	1.380	-2.563,5.324
	AP	0.415	-0.326,1.156
	S	2.461	0.333,18.194
rs2288947	RERI	-1.396	-3.118,0.327
	AP	-0.785	-1.869,0.299
	S	0.358	0.113,1.134
rs8105637	RERI	-1.598	-3.221,0.026
	AP	-1.169	-2.536,0.198
	S	0.187	0.029,1.189