The effect of susceptibility variants, identified in never-smoking female lung cancer cases, on male smokers

Seung Soo Yoo1,2,*, Hyo-Gyoung Kang3,4,*, Jin Eun Choi3,4, Mi Jeong Hong3,4, Sook Kyung Do3,5, Jang Hyuck Lee3,5, Won Kee Lee6, Shin Yup Lee1,2, Jaehee Lee7, Seung Ick Cha1, Chang Ho Kim4, Eung Baek Lee7, and Jae Yong Park1,2,3,4,5

1Department of Internal Medicine, Kyungpook National University Hospital, Daegu; 2Lung Cancer Center, Kyungpook National University Chilgok Hospital, Daegu; 3Department of Biochemistry and Cell Biology, 4Cell and Matrix Research Institute, School of Medicine, Kyungpook National University, Daegu; 5BK21 Plus KNU Biomedical Convergence Program, Department of Biomedical Science, 6Biostatistics, Medical Research Collaboration Center, Kyungpook National University, Daegu; 7Department of Thoracic Surgery, School of Medicine, Kyungpook National University, Daegu, Korea

Received: November 22, 2018
Revised: April 30, 2019
Accepted: May 13, 2019

Correspondence to
Jae Yong Park, M.D.
Lung Cancer Center, Kyungpook National University Chilgok Hospital, 807 Hoguk-ro, Buk-gu, Daegu 41404, Korea
Tel: +82-53-200-2631
Fax: +82-53-200-2627
E-mail: jaeyong@knu.ac.kr

*These authors contributed equally to this work.

INTRODUCTION
Lung cancer is the leading cause of cancer-related death. Although the majority of lung cancer cases are related to tobacco smoking, approximately 20% of lung cancer patients have never smoked [1,2]. Radon, asbestos, and...
indoor air pollution, such as from cooking fumes, are thought to cause lung cancer in never smokers (LCINS) [1].

LCINS exhibits different epidemiological, clinical, and molecular characteristics to those in lung cancer in smokers [2]. The proportion of LCINS is higher in women than in men [3]. Epidermal growth factor receptor (EGFR) mutations and anaplastic lymphoma kinase rearrangements are more commonly found in LCINS, while KRAS proto-oncogene mutations are more frequent in lung cancer in smokers than in never smokers. Genome-wide association studies (GWAS) have identified several genetic variants that have been associated with the susceptibility to LCINS [4-9]. In a previous study, we also found three polymorphisms that may contribute to lung cancer susceptibility in never-smoking Korean females [10].

Although these variants influenced the risk of lung cancer in never-smoking females, the association of these variants with the lung cancer risk in male smokers has not been studied. The pathogenesis of lung cancer in female never smokers and male smokers can be different. We wondered whether these polymorphisms were confined to never smokers and females. So, we investigated the association with lung cancer risk among male smokers, on the other side of female-never smokers. If these polymorphisms also affect the lung cancer risk in male smokers, this suggests that they may be associated with etiology of lung cancer other than smoking. We conducted a case-control study to investigate the association between the variants identified in the GWAS or candidate gene association studies in never-smoking females and the lung cancer risk in male smokers.

METHODOLOGY

Study population
This study included 714 lung cancer patients and 626 healthy controls. All of the patients and control subjects were male smokers (current or former). The patients were histologically confirmed to have lung cancer between September 1998 and June 2012 at Kyungpook National University Hospital (KNUH) in Daegu, Korea. Patient blood samples were provided by the National Biobank of Korea–KNUH, which is supported by the Ministry of Health, Welfare, and Family Affairs of Korea. The blood samples for the 626 healthy controls were randomly selected from the pool of healthy volunteers and provided as follows: 456 from the National Biobank of Korea–KNUH, and 190 from the Korean Biobank Project (4851-307, KBP-2011-24) and the Korean Genome and Epidemiology study (4845-302), which is supported by the Korea Centers for Disease Control & Prevention, Republic of Korea. All materials derived from the National Biobank were obtained under protocols that were approved by Institutional Review Boards. This study was approved by the Institutional Review Board of KNUH (Approval No.: KNUHBIO_09-1018), and written informed consent was obtained from all participants.

Polymorphism selection and genotyping
Eight polymorphisms were selected that were identified in never smoker GWAS [4-9]. Three additional polymorphisms: rs10079250 A > G, rs7631358 G > A, and rs13009079 T > C that contributed to lung cancer susceptibility in never-smoking Korean females in our previous study were also included [10]. Among the 11 polymorphisms studied, 10 were genotyped using SEQUENOM’s MassARRAY iPLEX assay (SEQUENOM Inc., San Diego, CA, USA). The rs7631358 G > A polymorphism was genotyped using the Taq-Man assay (Applied Biosystems, Foster City, CA, USA), following the manufacturer’s instructions. The genotyping analysis was performed in a blinded manner with respect to the case/control status. For quality control, approximately 5% of the samples were randomly selected to be genotyped again using a restriction fragment length polymorphism assay by a different investigator, and the results were 100% concordant.

Statistical analysis
Continuous variables, such as age and pack-years, were analyzed using the Student’s t test. A chi-square test was used to compare the differences between the categorical variables, such as the frequencies of the genotypes between the cases and controls. Unconditional logistic regression analysis was used to calculate the odds ratio (OR) and 95% confidence interval (CI), by adjusting for age and pack-years of smoking. In the stratified analysis, individuals were grouped according to age (≤ 61 vs. > 61), pack-years of smoking (≤ 35 vs. > 35), and histology.
(squamous cell carcinoma vs. adenocarcinoma vs. small cell carcinoma). The Statistical Analysis System for Windows version 9.4 (SAS Institute, Cary, NC, USA), was used for the statistical analyses. A \( p \) value of less than 0.05 was considered statistically significant.

**RESULTS**

The characteristics of the cases and controls are shown in Table 1. All of the participants were current or former smokers and of the male gender. The proportions of current smokers and pack-years of smoking were higher in the case group than in the control group (\( p = 0.01 \) and \( p < 0.001 \), respectively). These differences were controlled for in the multivariate analyses by adjusting for the pack-years of smoking.

The genotype frequencies of the 11 polymorphisms in the cases and controls are shown in Supplementary Table 1. Among the 11 polymorphisms, three (rs4975616, rs9387478, and rs13009079) showed different genotype frequencies between the cases and controls. After adjusting for age and pack-years of smoking, the rs4975616 A > G polymorphism was significantly associated with the risk of lung cancer. Individuals carrying the rs4975616 variant allele had a significantly decreased risk of lung cancer (in a codominant model: OR, 0.77; 95% CI, 0.61 to 0.96; \( p = 0.02 \) (Table 2). The rs9387478 C > A polymorphism was also associated with a decreased lung cancer risk (in a codominant model: OR, 0.85; 95% CI, 0.73 to 0.997; \( p = 0.046 \) (Table 2). The rs13009079 T > C polymorphism was related to a risk of lung cancer with only marginal significance (under a dominant model: \( p = 0.07 \)) (Supplementary Table 2).

In this study, the median value of smoking was 35 pack-years. So we classified cases with less than 35 pack-years as lighter smokers and those with more than 35 pack-years as heavier smokers. When stratified by pack-years of smoking, the effect of rs4975616 A > G on the lung cancer risk was significant in lighter smokers (OR, 0.65; 95% CI, 0.47 to 0.90; \( p = 0.01 \), but not in heavier smokers (OR, 0.90; 95% CI, 0.65 to 1.24, \( p = 0.52 \) (Table 3).
The effect of rs9387478 C > A on the lung cancer risk was also observed in lighter smokers (p = 0.02). Upon stratification by histological subtype, the association between rs4975616 A > G or rs9387478 C > A and the risk of lung cancer was predominant in adenocarcinoma (p = 0.03 and p = 0.05, respectively) (Table 3). In small cell lung cancer, there was no evidence that either polymorphism was associated with a risk of lung cancer.

DISCUSSION

In this study, we investigated the impact of 11 polymorphisms, which influenced the risk of LCINS, on the risk of lung cancer in male smokers. The rs4975616 A > G and rs9387478 C > A polymorphisms were associated with a decreased lung cancer risk in male smokers.

The rs4975616 polymorphism is located near telomerase reverse transcriptase (TERT) and cleft lip and palate transmembrane protein 1 (CLPTM1L) on chromosome 5p15.33. TERT is one of the major functional subunits of the telomerase enzyme, which plays a critical role in the maintenance of telomere length [11]. CLPTM1L encodes a membrane protein that causes apoptosis when over-expressed in cisplatin-sensitive cells [12]. A number of studies have demonstrated that the TERT-CLPTM1L region of chromosome 5p15.33 is a common susceptibility locus for multiple cancers, including lung cancer [13-17]. A GWAS has reported that rs4975616 A > G is associated with a decreased lung cancer risk in never smokers [6]. In this study, we found that rs4975616 A > G was also associated with decreased lung cancer risk in male smokers. When stratified by pack-years of smoking, the effect of rs4975616 A > G on the risk of lung cancer was significant in lighter smokers, but not in heavier smokers. These results suggest that the rs4975616 variant allele may be associated with a risk of lung cancer, but only in never or lighter smokers.

The rs4975616 A > G polymorphism was not associated with the risk of small cell lung cancer. The effect of rs4975616 A > G on the lung cancer risk was predominant in the patients with adenocarcinoma. A varied effect of rs4975616 A > G, based on the lung cancer histological type, has also been reported in a GWAS of never smokers [6]. In that study, the rs4975616 A > G polymorphism decreased the lung cancer risk in adenocarcinoma, but
Table 3. Association between polymorphisms and lung cancer risk according to age, pack-years of smoking and histological subtype

| Gene/polyorphism | Variable | Cases | Controls | Adjusted OR (95% CI) | p value<sup>b</sup> | Codominant model | p value<sup>c</sup> |
|------------------|----------|-------|----------|----------------------|-------------------|------------------|------------------|
| **TERT-CLPTM1L**<br>rs4975616 | Age, yr |       |          |                      |                   |                  |                  |
| < 61             | 248 (75.8) 74 (22.6) 5 (1.6) | 227 (71.2) 83 (26.0) 9 (2.8) | 0.13 | 1.00 0.79 (0.58–1.07) 0.13 |                  |                  |
| > 61             | 293 (77.1) 80 (21.1) 7 (1.8) | 202 (71.1) 78 (27.5) 4 (1.4) | 0.14 | 1.00 0.78 (0.57–1.07) 0.13 |                  |                  |
| Pack-years       | < 35     | 236 (79.5) 57 (19.2) 4 (1.3) | 271 (70.4) 107 (27.8) 7 (1.8) | 0.03 | 1.00 0.65 (0.47–0.90) 0.01 |                  |                  |
|                  | > 35     | 305 (74.4) 97 (23.7) 8 (2.0) | 158 (72.5) 54 (24.8) 6 (2.8) | 0.76 | 1.00 0.90 (0.65–1.24) 0.52 |                  |                  |
| Histology        | SQCC     | 268 (76.8) 75 (21.5) 6 (1.7) | 429 (71.1) 161 (26.7) 13 (2.2) | 0.07 | 1.00 0.78 (0.59–1.02) 0.07 |                  |                  |
|                  | AC       | 162 (77.5) 46 (22.0) 1 (0.5) | 429 (71.1) 161 (26.7) 13 (2.2) | 0.04 | 1.00 0.68 (0.48–0.97) 0.03 |                  |                  |
|                  | SCLC     | 86 (72.9) 28 (23.7) 4 (3.4) | 429 (71.1) 161 (26.7) 13 (2.2) | 0.92 | 1.00 0.98 (0.66–1.44) 0.91 |                  |                  |
| **ROS-DCBLD1**<br>rs9387478 | Age, yr |       |          |                      |                   |                  |                  |
| < 61             | 103 (32.0) 155 (48.1) 64 (19.9) | 93 (28.6) 155 (47.7) 77 (23.7) | 0.20 | 1.00 0.87 (0.70–1.08) 0.20 |                  |                  |
| > 61             | 128 (34.0) 171 (45.4) 78 (20.7) | 79 (28.2) 134 (47.9) 67 (23.9) | 0.12 | 1.00 0.82 (0.66–1.02) 0.08 |                  |                  |
| Pack-years       | < 35     | 102 (34.7) 135 (45.9) 57 (19.4) | 108 (27.8) 184 (47.4) 96 (24.7) | 0.09 | 1.00 0.78 (0.63–0.97) 0.02 |                  |                  |
|                  | > 35     | 129 (31.9) 191 (47.2) 85 (21.0) | 64 (29.5) 105 (48.4) 48 (22.1) | 0.83 | 1.00 0.93 (0.74–1.17) 0.55 |                  |                  |
| Histology        | SQCC     | 114 (33.1) 198 (45.9) 72 (20.9) | 172 (28.4) 289 (47.8) 144 (23.8) | 0.12 | 1.00 0.87 (0.72–1.04) 0.13 |                  |                  |
|                  | AC       | 73 (35.3) 93 (44.9) 41 (19.8) | 172 (28.4) 289 (47.8) 144 (23.8) | 0.06 | 1.00 0.80 (0.64–1.00) 0.05 |                  |                  |
|                  | SCLC     | 32 (27.6) 60 (51.7) 24 (20.7) | 172 (28.4) 289 (47.8) 144 (23.8) | 0.75 | 1.00 0.96 (0.72–1.27) 0.76 |                  |                  |

Values are presented as number (%).

OR, odds ratio; CI, confidence interval; TERT, telomerase reverse transcriptase; CLPTM1L, cleft lip and palate transmembrane protein 1; SQCC, squamous cell carcinoma; AC, adenocarcinoma; SCLC, small cell lung cancer; ROS1, ROS proto-oncogene 1, receptor tyrosine kinase; DCBLD1, discoidin, CUB and LCCL domain containing 1.

<sup>a</sup>1 denotes a wild allele and 2 denotes a variant allele.

<sup>b</sup>p value for chi-square test.

<sup>c</sup>OR, 95% CI and their corresponding p value were calculated using unconditional logistic regression analysis.
not in small cell lung cancer [6]. The similar result for the effect of rs4975616 A > G on the lung cancer risk in male smokers, as in non-smokers, suggests that the effect of this variant on the lung cancer risk may not be specific to non-smokers. It was reported that TERT-CLPTM1 locus polymorphisms were associated with telomere length [18]. So, there is a possibility that the TERT-CLPTM1L rs4975616 A > G may influence on telomere length and affect lung cancer susceptibility. There may be in linkage disequilibrium between rs4975616 and other functional polymorphisms affecting telomere length. Additional studies are needed to understand the biological mechanism of rs4975616 A > G on lung cancer development.

The rs9387478 polymorphism is located in an interval between ROS proto-oncogene 1, receptor tyrosine kinase (ROS1) and discoidin, CUB and LCCL domain containing 1 (DCBLD1) on chromosome 6p21.2. ROS1 functions as a receptor tyrosine kinase. Increased expression or rearrangement of ROS1 has been reported in non-small cell lung cancer [19]. Little is known regarding the functional role of DCBLD1. A GWAS has identified that rs9387478 C > A is associated with a decreased lung cancer risk in never-smoking women in Asia [4]. In that study, rs9387478 was associated with a risk of adenocarcinoma, when analyzed by the histological subtype of lung cancer [4]. In this study, rs9387478 was also associated with a decreased lung cancer risk in male smokers, and the association was significant for adenocarcinoma only. Different EGFR mutation status has been reported, based on the rs9387478 genotypes. In never-smoking female lung cancer cases, the frequency of the rs9387478 AA genotype was significantly higher in the EGFR mutation-negative group than in the EGFR mutation-positive group [20]. Therefore, the varied effects of rs9387478 A > G on the risk of lung cancer, according to the histological subtype, may have been due to the different EGFR mutations. In this study, the association between the rs9387478 genotypes and the EGFR mutation status was not evaluated.

There are several limitations in this study. First, the p value of ROS1-DCBLD1 rs9387478 was less than 0.05 in the codominant model. However, the p value has marginal significance when considering multiple tests. Second, the relationship with survival outcomes is not investigated. For example, the ROS1-DCBLD1 rs9387478 was reported to correlate with overall survival in female nonsmoking patients with lung cancer [20]. However, we could not analyze survival outcomes of male smokers due to lack of survival data. And a lack of functional study may also be a weakness of this study.

Our study has shown that the effects of polymorphisms are not confined to never smokers. Two polymorphisms were associated with lung cancer risk in both never-smoking females and male smokers. Since there are little known about the functional roles of these polymorphisms, the reasons why they are related to lung cancer risk are unclear. Additional studies on the polymorphisms may reveal the etiology of lung cancer other than smoking.

In summary, polymorphisms that are known to be associated with the risk of LCINS were not specific to never smokers. The TERT-CLPTM1L rs4975616 A > G and ROS1-DCBLD1 rs9387478, which were identified in GWAS of never-smoking females, were also associated with the risk of lung cancer in male smokers.

KEY MESSAGE

1. This study was conducted to evaluate the association between 11 polymorphisms identified in female never smokers and the lung cancer risk in male smokers.
2. The telomerase reverse transcriptase (TERT)-cleft lip and palate transmembrane protein 1 (CLPTM1L) rs4975616 and ROS proto-oncogene 1, receptor tyrosine kinase (ROS1)-discoidin, CUB and LCCL domain containing 1 (DCBLD1) rs9387478, which were identified in genome-wide association studies of never-smoking females, were also associated with the risk of lung cancer in male smokers.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

This study was supported by a grant from the National R&D Program for Cancer Control, Ministry of Health.
and Welfare, Republic of Korea (grant number: 1720040); the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2017R1A1A3B03032114).

REFERENCES

1. Subramanian J, Govindan R. Lung cancer in never smokers: a review. J Clin Oncol 2007;25:561-570.
2. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers: a different disease. Nat Rev Cancer 2007;7:778-790.
3. Torok S, Hegedus B, Laszlo V, et al. Lung cancer in never smokers. Future Oncol 2011;7:1195-1211.
4. Lan Q, Hsiung CA, Matsuo K, et al. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. Nat Genet 2012;44:1330-1335.
5. Hsiung CA, Lan Q, Hong YC, et al. The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. PLoS Genet 2010;6:e1001051.
6. Wang Y, Broderick P, Matakidou A, Eisen T, Houlston RS. Role of 5p15.33 (TERT-CLPTM1L), 6p21.33 and 15q25.1 (CHRNA5-CHRNA3) variation and lung cancer risk in never-smokers. Carcinogenesis 2010;31:234-238.
7. Li Y, Sheu CC, Ye Y, et al. Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. Lancet Oncol 2010;11:321-330.
8. Ahn MJ, Won HH, Lee J, et al. The 18p11.22 locus is associated with never smoker non-small cell lung cancer susceptibility in Korean populations. Hum Genet 2012;131:365-372.
9. Hosgood HD 3rd, Wang WC, Hong YC, et al. Genetic variant in TP69 on locus 3q28 is associated with risk of lung adenocarcinoma among never-smoking females in Asia. Hum Genet 2012;131:1197-1203.
10. Kang HG, Lee SY, Jeon HS, et al. A functional polymorphism in CSF1R gene is a novel susceptibility marker for lung cancer among never-smoking females. J Thorac Oncol 2014;9:1647-1655.
11. Cohen SB, Graham ME, Lovrecz GO, Bache N, Robinson PJ, Reddel RR. Protein composition of catalytically active human telomerase from immortal cells. Science 2007;315:1850-1853.
12. Yamamoto K, Okamoto A, Isonishi S, Ochiai K, Ohtake Y. A novel gene, CRR9, which was up-regulated in CD-DP-resistant ovarian tumor cell line, was associated with apoptosis. Biochem Biophys Res Commun 2001;280:1148-1154.
13. McKay JD, Hung RJ, Gaborieau V, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet 2008;40:1404-1406.
14. Broderick P, Wang Y, Vijayakrishnan J, et al. Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. Cancer Res 2009;69:6633-6641.
15. Petersen GM, Amundadottir L, Fuchs CS, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 13q2.1 and 5p15.33. Nat Genet 2010;42:224-228.
16. Kote-Jarai Z, Olama AA, Giles GG, et al. Seven prostate cancer susceptibility loci identified by a multi-stage genome-wide association study. Nat Genet 2011;43:785-791.
17. Beesley J, Pickett HA, Johnatty SE, et al. Functional polymorphisms in the TERT promoter are associated with risk of serous epithelial ovarian and breast cancers. PLoS One 2011;6:e24987.
18. Llorca-Cardenosa MJ, Pena-Chilet M, Mayor M, et al. Long telomere length and a TERT-CLPTM1 locus polymorphism association with melanoma risk. Eur J Cancer 2014;50:3168-3177.
19. Berghethon K, Shaw AT, Ou SH, et al. ROS1 rearrangements define a unique molecular class of lung cancers. J Clin Oncol 2012;30:863-870.
20. Han JF, An SJ, Su WM, et al. Polymorphism of rs9387478 correlates with overall survival in female nonsmoking patients with lung cancer. Int J Biol Markers 2016;31:e144-e152.
**Supplementary Table 1. Genotype frequencies of 11 polymorphisms in lung cancer cases and controls**

| Polymorphism | Gene          | Genotype | Cases<sup>a</sup> | Controls<sup>a</sup> | p value<sup>b</sup> |
|--------------|---------------|----------|-------------------|----------------------|---------------------|
| rs4975616    | *TERT*        | AA       | 541 (76.5)        | 429 (71.4)           | 0.09                |
|              |               | AG       | 154 (21.8)        | 161 (26.7)           |                     |
|              |               | GG       | 12 (1.7)          | 13 (2.2)             |                     |
|              |               |          |                   |                      | Dominant 0.03       |
|              |               |          |                   |                      | Recessive 0.55      |
|              |               |          |                   |                      | Codominant 0.03     |
| rs9387478    | *ROS1*        | CC       | 231 (33.0)        | 172 (28.4)           | 0.13                |
|              |               | CA       | 326 (46.6)        | 289 (47.8)           |                     |
|              |               | AA       | 142 (20.3)        | 144 (23.8)           |                     |
|              |               |          |                   |                      | Dominant 0.07       |
|              |               |          |                   |                      | Recessive 0.13      |
|              |               |          |                   |                      | Codominant 0.05     |
| rs13009079   | **CIR1**      | TT       | 437 (62.4)        | 349 (56.4)           | 0.08                |
|              |               | TC       | 229 (32.7)        | 237 (38.3)           |                     |
|              |               | CC       | 34 (4.9)          | 33 (5.3)             |                     |
|              |               |          |                   |                      | Dominant 0.03       |
|              |               |          |                   |                      | Recessive 0.70      |
|              |               |          |                   |                      | Codominant 0.05     |
| rs10079250   | *CSF1R*       | AA       | 321 (45.7)        | 310 (50.4)           | 0.03                |
|              |               | AG       | 321 (45.7)        | 237 (38.5)           |                     |
|              |               | GG       | 61 (8.7)          | 68 (11.1)            |                     |
|              |               |          |                   |                      | Dominant 0.09       |
|              |               |          |                   |                      | Recessive 0.15      |
|              |               |          |                   |                      | Codominant 0.51     |
| rs2352028    | **GPC5**      | CC       | 385 (54.4)        | 330 (53.6)           | 0.78                |
|              |               | CT       | 275 (38.8)        | 238 (38.6)           |                     |
|              |               | TT       | 48 (6.8)          | 48 (7.8)             |                     |
|              |               |          |                   |                      | Dominant 0.77       |
|              |               |          |                   |                      | Recessive 0.48      |
|              |               |          |                   |                      | Codominant 0.60     |
| rs2395185    | *HLA class II*| GG       | 259 (36.7)        | 231 (38.8)           | 0.72                |
|              |               | GT       | 326 (46.2)        | 269 (45.2)           |                     |
|              |               | TT       | 120 (17.0)        | 95 (16.0)            |                     |
|              |               |          |                   |                      | Dominant 0.44       |
|              |               |          |                   |                      | Recessive 0.61      |
|              |               |          |                   |                      | Codominant 0.42     |
| rs11080466   | **PIEO2**     | AA       | 478 (68.7)        | 415 (67.3)           | 0.19                |
|              |               | AG       | 207 (29.7)        | 183 (29.7)           |                     |
|              |               | GG       | 1 (0.1)           | 19 (3.1)             |                     |
|              |               |          |                   |                      | Dominant 0.58       |
|              |               |          |                   |                      | Recessive 0.07      |
|              |               |          |                   |                      | Codominant 0.31     |
## Supplementary Table 1. Continued

| Polymorphism  | Gene     | Genotype | Cases<sup>a</sup> | Controls<sup>a</sup> | p value<sup>b</sup> |
|--------------|----------|----------|-------------------|----------------------|--------------------|
| rs2736100    | TERT     | TT       | 269 (38.5)        | 241 (39.8)           | 0.57               |
|              |          | TG       | 321 (45.9)        | 283 (46.7)           |                    |
|              |          | GG       | 109 (15.6)        | 82 (13.5)            |                    |
|              |          |          |                   |                      |                    |
|              |          |          | Dominant          | 0.64                 |                    |
|              |          |          | Recessive         | 0.29                 |                    |
|              |          |          | Codominant        | 0.38                 |                    |
| rs7631358    | TP63     | GG       | 214 (30.4)        | 192 (31.1)           | 0.32               |
|              |          | GA       | 338 (48.0)        | 312 (50.6)           |                    |
|              |          | AA       | 152 (21.6)        | 113 (18.3)           |                    |
|              |          |          | Dominant          | 0.78                 |                    |
|              |          |          | Recessive         | 0.14                 |                    |
|              |          |          | Codominant        | 0.30                 |                    |
| rs10937405   | TP63     | CC       | 343 (49.6)        | 280 (45.8)           | 0.23               |
|              |          | CT       | 294 (42.5)        | 270 (44.1)           |                    |
|              |          | TT       | 55 (7.9)          | 62 (10.1)            |                    |
|              |          |          | Dominant          | 0.17                 |                    |
|              |          |          | Recessive         | 0.17                 |                    |
|              |          |          | Codominant        | 0.09                 |                    |
| rs70868c3    | VTI1A    | GG       | 395 (56.9)        | 368 (59.6)           | 0.35               |
|              |          | GA       | 259 (37.3)        | 223 (36.1)           |                    |
|              |          | AA       | 40 (5.8)          | 26 (4.2)             |                    |
|              |          |          | Dominant          | 0.32                 |                    |
|              |          |          | Recessive         | 0.20                 |                    |
|              |          |          | Codominant        | 0.19                 |                    |

Values are presented as number (%).

*TERT*, telomerase reverse transcriptase; *CLPTM1L*, cleft lip and palate transmembrane protein 1; *ROS1*, ROS proto-oncogene 1, receptor tyrosine kinase; *DCBLD1*, discoidin, CUB and LCCL domain containing 1; *CIR1*, corepressor interacting with RBPJ, 1; *CSF1R*, colony stimulating factor 1 receptor; *GPC5*, glypican 5; *HLA*, major histocompatibility complex; *PIEZ2*, piezo type mechanosensitive ion channel component 2; *TP63*, tumor protein p63; *VTI1A*, vesicle transport through interaction with t-SNAREs 1A.

<sup>a</sup>Number may not add up total number of patients due to genotype failures.

<sup>b</sup>p value for chi-square test.
### Supplementary Table 2. Genotype frequency of rs13009079 in male smokers lung cancer cases and controls

| Polymorphism     | Gene     | Allele | Cases\(^a\) | Controls\(^a\) | \(p\) value\(^b\) | OR (95% CI)\(^c\) | \(p\) value\(^c\) |
|------------------|----------|--------|-------------|----------------|------------------|-------------------|------------------|
| rs13009079       | CIR1     | TT     | 437 (62.4)  | 340 (56.4)     | 1.00             |                   |                  |
|                  |          | TC     | 229 (32.7)  | 237 (38.3)     | 0.81 (0.64–1.03) | 0.08              |                  |
|                  |          | CC     | 34 (4.9)    | 33 (5.3)       | 0.82 (0.49–1.38) | 0.46              |                  |
| Dominant         |          |        | 0.03        | 0.81 (0.65–1.02) | 0.07             |                  |
| Recessive        |          |        | 0.70        | 0.89 (0.54–1.48) | 0.65             |                  |
| Codominant       |          |        | 0.05        | 0.85 (0.71–1.03) | 0.10             |                  |

Values are presented as number (%).

OR, odds ratio; CI, confidence interval; CIR1, corepressor interacting with RBPJ, 1.

\(^a\)Number may not add up total number of patients due to genotype failures.

\(^b\)\(p\) value for chi-square test.

\(^c\)OR, 95% CI and their corresponding \(p\) values were calculated using unconditional logistic regression analysis, adjusted for age and pack-years.