



Analysis of Fifty Hotspot Mutations of Lung Squamous Cell Carcinoma in Never-smokers

Ha Youn Lee,^{1,2} Se-Hoon Lee,^{2,3}
Jae-Kyung Won,⁴ Dong Soo Lee,⁵
Nak-Jung Kwon,⁶ Sun Mi Choi,^{1,2}
Jinwoo Lee,^{1,2} Chang-Hoon Lee,^{1,2}
Sang-Min Lee,^{1,2} Jae-Joon Yim,^{1,2}
Chul-Gyu Yoo,^{1,2} Young Whan Kim,^{1,2}
Sung Koo Han,^{1,2} and Young Sik Park^{1,2}

¹Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Seoul National University Hospital, Seoul, Korea; ²Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Korea; ³Cancer Research Institute, Seoul National University College of Medicine, Seoul, Korea; ⁴Department of Pathology, Seoul National University College of Medicine, Seoul, Korea; ⁵Department of Nuclear Medicine, Seoul National University College of Medicine, Seoul, Korea; ⁶Macrogen Inc., Seoul, Korea

Received: 11 April 2016
Accepted: 11 November 2016

Address for Correspondence:
Young Sik Park, MD
Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine and Lung Institute, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Korea
E-mail: mdyspark@gmail.com

Funding: This study was supported by the National Research Fund (Grant No. H113C-1299-020013) of the Korea Health Industry Development Institute (KHIDI) and the Ministry of Health and Welfare of Korea.

INTRODUCTION

Smoking is a major risk factor for lung cancer; about 80% of lung cancers are associated with smoking in men and more than 50% are associated with smoking in women (1). Chronic exposure of tobacco smoke to lung epithelium causes diverse oncogenic mutations, so smoking-related lung cancer shows high burden of mutations such as *TP53* (2). According to The Cancer Genome Atlas (TCGA) data, the mutational burden gradually increases in adenocarcinoma with increasing exposure to smoke, and the highest burden of mutations are found in current smokers with either adenocarcinoma or squamous cell carcinoma (SCC) (3).

Lung SCC accounts for approximately 30% of all lung cancers and is the most common histologic type of smoking-related non-small cell lung cancer (4). However, lung SCC also occurs in a

small number of never-smokers. We hypothesized that second-hand smoke exposure might be the most significant risk factor for lung SCC in never smoker, because SCC is considered as the typical histologic type of smoking-related lung cancer. If this is correct then we would expect that SCC mutation profiles in never-smokers would be similar to those in smokers. Therefore, we compared 50 hotspot mutations of lung SCCs between never-smokers and smokers.

Smoking is the major risk factor for lung squamous cell carcinoma (SCC), although a small number of lung SCCs occurs in never-smokers. The purpose of this study was to compare 50 hotspot mutations of lung SCCs between never-smokers and smokers. We retrospectively reviewed the medical records of patients newly diagnosed with lung SCC between January 1, 2011 and December 31, 2013 in the Seoul National University Hospital. Formalin-fixed, paraffin-embedded tumor samples were used for analysis of hotspot mutations. Fifty cancer-related genes in never-smokers were compared to those in ever-smokers. Of 379 lung SCC patients, 19 (5.0%) were never-smokers. The median age of these 19 patients was 67 years (interquartile range 57–73 years), and 10 of these patients were women (52.5%). The incidence rates of stage I, II, III, and IV disease in this group were 26.4%, 5.3%, 31.6%, and 36.8%, respectively, and sequencing was performed successfully in 14 cases. In the 26 lung SCC tumor samples (12 from never-smokers and 14 from ever-smokers) sequenced using personal genome machine, the most common mutations were in *TP53* (75.0%), *RAS* (66.7%), and *STK11* (33.3%), but mutations were also found in *EGFR*, *KIT*, and *PTEN*. The distribution of hotspot mutations in never-smokers was similar to that in ever-smokers. There was no significant difference in overall survival between the 2 groups. The 50 hotspot mutations of lung SCC in never-smokers were similar to those of ever-smokers.

Keywords: Lung Cancer; Squamous Cell Carcinoma; Never-Smoker; Mutation; *TP53*; *RAS*; *STK11*

small number of never-smokers. We hypothesized that second-hand smoke exposure might be the most significant risk factor for lung SCC in never smoker, because SCC is considered as the typical histologic type of smoking-related lung cancer. If this is correct then we would expect that SCC mutation profiles in never-smokers would be similar to those in smokers. Therefore, we compared 50 hotspot mutations of lung SCCs between never-smokers and smokers.

MATERIALS AND METHODS

Study population

Between January 2011 and December 2013, consecutive patients who were newly diagnosed with lung SCC at Seoul National University Hospital were enrolled in this study. We retrospectively collected information including age, sex, stage, and

smoking status. The patients were divided into the never-smoker group (< 100 cigarettes during their life-time) and the ever-smoker group. In order to compare these groups, we matched patients based on age and final stage. Mortality data were obtained from the database of the National Health and Safety Executive of the Republic of Korea Government.

Definition of final disease stage

Clinical and pathologic stages were evaluated based on the 7th Tumor, node, metastasis (TNM) staging system. Surgical stage is more accurate than clinical stage in patients who have undergone surgical resection, and hence the former was used for those patients who had undergone this procedure. However, patients with advanced stage disease are not usually treated using surgery, so clinical stage was used in these cases. We defined the final disease stage as the combined surgical and clinical stage, according to whether or not surgical resection was performed.

Tumor samples and DNA extraction

Archived formalin-fixed, paraffin-embedded (FFPE) tumor tissues were collected for DNA extraction. Resected lung tissue was used when available; otherwise, small bronchoscopic or needle biopsy samples were used. DNA was isolated using Promega Maxwell 16 MDX (Promega, Mannheim, Germany) according to the manufacturer's instructions. We identified DNA quality on agarose gel whether it is degraded, and extracted DNA was assessed for quantity and quality using Invitrogen Qubit® 2.0 quantitation assays (Invitrogen, Grand Island, NY, USA).

Personal genome machine (PGM) for sequencing 50 hotspot mutations

For sequencing of the 50 hotspot mutations, genomic DNA purity was assessed by electrophoresis on a 1% agarose gel followed by visualization using a Qubit® 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Purified genomic DNA was used for library construction with the Ion AmpliSeq™ Cancer hotspot panel v2 (Life Technologies) that targets mutations in the following 50 genes: *ABLI*, *AKT1*, *ALK*, *ATM*, *APC*, *BRAF*, *CDH1*, *CDKN2A*, *CSF1R*, *CTNNB1*, *EGFR*, *ERBB2*, *ERBB4*, *EZH2*, *FBXW7*, *FGFR1*, *FGFR2*, *FGFR3*, *FLT3*, *GNAQ*, *GNAS*, *GNA11*, *HNF1A*, *HRAS*, *IDH1*, *IDH2*, *JAK2*, *JAK3*, *KDR*, *KIT*, *KRAS*, *MET*, *MLH1*, *MPL*, *NOTCH1*, *NPM1*, *NRAS*, *PDGFRA*, *PIK3CA*, *PTEN*, *PTPN11*, *RBI*, *SMAD4*, *SMARCB1*, *SMO*, *SRC*, *STK11*, *TP53*, *RET*, and *VHL*. Briefly, about 192 amplicons were isolated using the Ion AmpliSeq™ Cancer Panel (Thermo Fisher Scientific, Fremont, CA, USA) from 10–20 ng of tumor genomic DNA from lung cancer patients, and then purified using the Agencourt® AMPure® XP system (Beckman Coulter, Miami, FL, USA). The amplicons were end-repaired and ligated with Ion Xpress barcode adapters, and then detected using the BioAnalyzer High Sensitivity Chip (Agilent, Santa Clara, CA, USA). The final library was generated by

diluting it to 10 pM in low Tris-EDTA buffer, and 5 µL of the library was used for emulsion polymerase chain reaction (PCR) reactions using the Onetouch™ reagent kit (Life Technologies). The emulsion PCR products were then enriched using Dynabeads® MyOne™ Streptavidin C1 beads (Invitrogen). The Ion sphere bead-enriched DNA was sequenced on an Ion Torrent PGM 316 chip, and base calling, alignment, variants calling, and an analysis report were generated using Torrent Suite 4.0 software (Ion Torrent PGM server; Thermo Fisher Scientific) with tmap-f3.

Statistical analysis

Fisher's exact test was used to analyze differences between categorical variables. The non-parametric Wilcoxon rank sum test and the Mann-Whitney test were also applied. The Kaplan-Meier method was used to estimate survival as a function of time, and significance was estimated using the log-rank test. All statistical tests were 2-sided, and differences were considered significant if $P < 0.05$. All analyses were performed using SPSS version 21.0 (SPSS Inc., Chicago, IL, USA).

Ethics statement

The Institutional Review Board of Seoul National University Hospital approved the study protocol (IRB No. H-1401-037-548). The board waived the requirement for written consent.

RESULTS

Patient characteristics

A total of 379 patients were newly diagnosed with lung SCC, 19 (50.0%) of whom (5.0%) were never-smokers. The baseline char-

Table 1. Baseline clinical characteristics of 379 patients with lung SCC

Parameters	Never-smoker (n = 19)	Ever-smoker (n = 360)	P value
Age, yr	67 (57–73)	70 (65–75)	0.055*
Male	9 (47.4)	349 (96.9)	< 0.001†
BMI, kg/m ²	21.9 (21.0–24.3)	22.8 (20.6–25.0)	0.644*
Smoking status			
Pack-years	0	40 (30–50)	-
Final stage			0.618†
I	5 (26.4)	91 (25.2)	
II	1 (5.3)	68 (18.9)	
III	6 (31.6)	114 (31.7)	
IV	7 (36.8)	87 (24.2)	
Tumor size, cm	3.9 (2.9–4.4)	4.0 (3.0–5.7)	0.552*
SUVmax	15.0 (11.6–17.0)	16.0 (11.5–20.8)	0.193*
CEA	2.9 (1.8–8.5)	2.1 (1.5–3.9)	0.230*
CYFRA 21-1	3.8 (2.4–6.1)	4.5 (2.5–10.2)	0.517*
NSE	19.9 (13.0–36.7)	17.5 (13.9–26.1)	0.571*

Values are presented as median (IQR) or number (%).

SCC = squamous cell carcinoma, BMI = body mass index, SUVmax = maximum standardized uptake value, CEA = carcinoembryonic antigen, CYFRA 21-1 = cytokeratin 19-fragments, NSE = neuron specific enolase, IQR = interquartile range.

*P value from Mann-Whitney test; †P value from χ^2 test.

acteristics of the patients according to smoking history are shown in Table 1. The median age of the never-smoker SCC patients was 67 years (interquartile range [IQR] 57–73 years), and 10 (52.5%) of these patients were women. There was a higher proportion of women in the never-smokers group than in the ever-smokers group. The average number of pack-year for the ever smokers was 40 pack-year (median, IQR 30–50). The distribution of final stages in never smokers was similar to those of ever smokers. Tumor markers and imaging findings did not differ between 2 groups.

Comparative mutational analysis of lung SCCs according to smoking history

In total, 14 tumor samples could be obtained from the 19 never-smokers. After sample matching at a 1:1 ratio based on age and stage, a total of 28 tumor samples were used for DNA extraction and sequencing. Two of these samples failed to yield sufficient DNA; thus, a total of 26 samples (12 from never-smokers and 14

from ever-smokers) were sequenced successfully and 2 samples failed the sequencing. Eleven (42.3%) of these successfully sequenced samples were obtained from small biopsy specimens.

In the 26 lung SCC tumor samples sequenced using PGM, 1,078 sequence variants and 211 hotspot mutations were detected. Most variants were missense mutations, with the exception of 2 deletions. The most common mutations were in the *TP53* (84.6%), *RAS* (69.2%), and *KIT* (34.5%) genes (Fig. 1).

As no comparison had previously been reported for lung SCC according to smoking history, we compared the hotspot mutations of never-smokers with those of ever-smokers, and found no significant differences in their distributions (Table 2). Both groups displayed a similarly high rate of mutations in *TP53* (75.0% in never-smokers and 92.9% in ever-smokers; Fisher’s exact $P = 0.306$) and *RAS* (66.7% in never-smokers and 71.4% in ever-smokers; Fisher’s exact $P > 0.99$), and mutations in *HRAS*, *NRAS*, *KRAS*, *EGFR*, *KIT*, and *PTEN* were also observed at comparable frequencies. *EGFR* mutations were present in 14.3% and 25.0% of

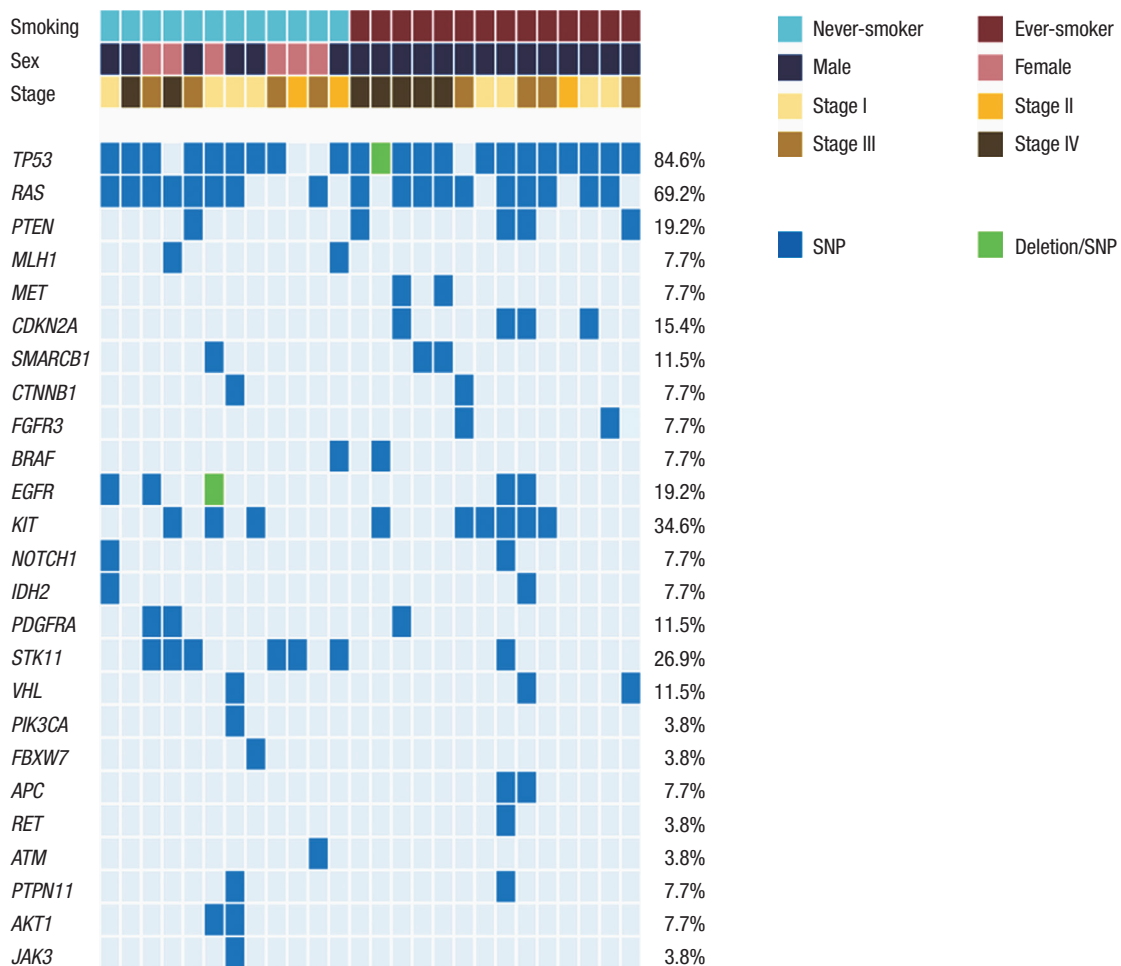


Fig. 1. Mutational profiles of lung SCC in never-smokers and ever-smokers. Heatmap representation of individual mutations present in 26 lung SCC samples (12 from never-smokers and 14 from ever-smokers). Percentages refer to the proportion of samples that carried at least 1 mutation in the listed gene (right). *RAS* is a conglomerated representation of *HRAS* and *KRAS*.

SCC = squamous cell carcinoma, SNP = single-nucleotide polymorphism.

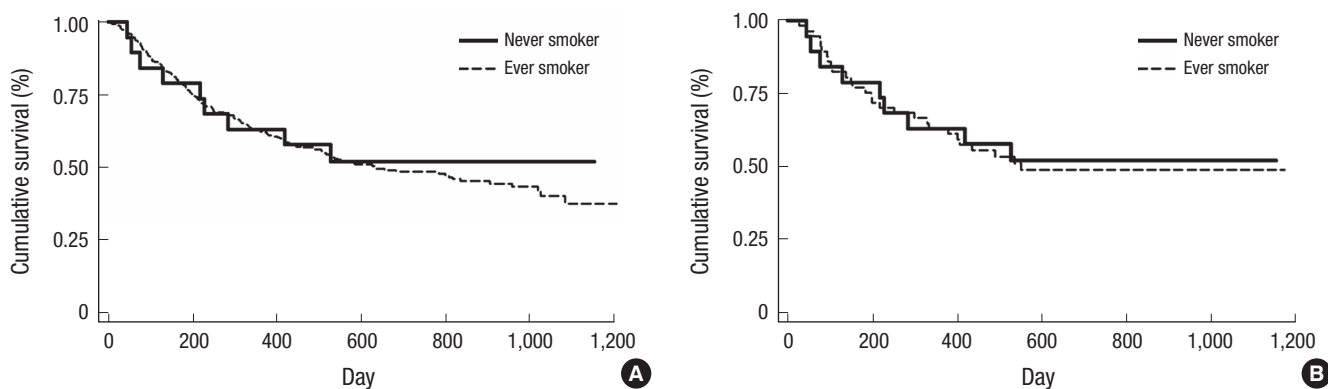


Fig. 2. OS of patients with lung SCC based on smoking history. (A) OS of all lung SCC patients ($n = 379$). (B) OS of patients matched by age and stage at a 1:3 ratio ($n = 76$). The P value was calculated using the log-rank test. OS = overall survival, SCC = squamous cell carcinoma.

Table 2. Comparison of genetic mutational profiles of lung SCC based on smoking status

Genes	Never-smoker ($n = 12$)	Ever-smoker ($n = 14$)	P value
<i>TP53</i>	9 (75.0)	13 (92.9)	0.306
<i>PTEN</i>	1 (8.3)	4 (28.6)	0.330
<i>HRAS</i>	7 (58.3)	10 (71.4)	0.683
<i>KRAS</i>	2 (16.7)	2 (14.3)	1.000
<i>MLH1</i>	2 (16.7)	0 (0.0)	0.203
<i>MET</i>	0 (0.0)	2 (14.3)	0.483
<i>CDKN2A</i>	0 (0.0)	4 (28.6)	0.100
<i>SMARCB1</i>	1 (8.3)	2 (14.3)	1.000
<i>CTNNB1</i>	1 (8.3)	1 (7.1)	1.000
<i>FGFR3</i>	0 (0.0)	2 (14.3)	0.483
<i>BRAF</i>	2 (16.7)	0 (0.0)	0.203
<i>EGFR</i>	3 (25.0)	2 (14.3)	0.635
<i>KIT</i>	3 (25.0)	6 (42.9)	0.429
<i>NOTCH1</i>	1 (8.3)	1 (7.1)	1.000
<i>IDH2</i>	1 (8.3)	1 (7.1)	1.000
<i>PDGFRA</i>	2 (16.7)	1 (7.1)	0.580
<i>STK11</i>	6 (50.0)	1 (7.1)	0.026
<i>VHL</i>	1 (8.3)	2 (14.3)	1.000
<i>PIK3A</i>	1 (8.3)	0 (0.0)	0.462
<i>FBXW7</i>	1 (8.3)	0 (0.0)	0.462
<i>APC</i>	0 (0.0)	2 (7.1)	0.483
<i>RET</i>	0 (0.0)	1 (8.3)	1.000
<i>ATM</i>	1 (8.3)	0 (0.0)	0.462
<i>JAK3</i>	1 (8.3)	0 (0.0)	0.462
<i>AKT1</i>	2 (16.7)	0 (0.0)	0.203
<i>PTPN11</i>	1 (8.3)	1 (7.1)	1.000

Values are presented as number (%).
SCC = squamous cell carcinoma.

tumors from ever-smokers and never-smokers, respectively, and this difference was also not statistically significant ($P = 0.637$).

Survival analysis

At the data cut-off point (August 2014), the median follow-up period was 766 days (IQR 523–972 days). The overall survival (OS) did not differ significantly between the 2 groups (log rank P value = 0.725) (Fig. 2). The mortality rates of never-smokers

and ever-smokers were 47.4% (9/19) and 50.0% (180/360), respectively ($P > 0.99$). Likewise, in an analysis of subgroups matched at a ratio of 1:3 for age and stage, neither OS nor mortality rates differed significantly (log rank P value = 0.872).

DISCUSSION

In an analysis of 50 hotspot mutations in lung SCC of never-smokers and ever-smokers, we found that there was a similar mutational profile in tumors regardless of smoking history. The analyzed mutations included those in the *TP53* and *RAS* genes, which have been found to occur more frequently in lung carcinomas from smokers compared to never-smokers (5,6). Furthermore, a number of clinical characteristics, including OS, did not differ significantly between never-smokers and smokers. Our findings support the hypothesis that lung SCCs follow the same oncogenic pathway regardless of the patient's smoking history.

Although data from TCGA revealed genetic aberrations in 178 SCC samples, never smokers were only 8 patients, which was too small to analyse a difference (7). In addition, TCGA investigated genetic mutations using fresh specimens obtained by surgical resection in order to give sufficient material for analysis, unlike the current study in which mutations in advanced stage SCC were detected using DNA from FFPE specimens. In order to improve the reliability of analysis, specimens were sequenced using Ion Torrent PGM, which was previously validated for use with FFPE samples (8).

Numerous studies have found that genomic alterations in lung SCC, particularly mutations in *TP53*, *PIK3CA*, *PTEN*, and *FGFR1* amplification are common in SCC but not in adenocarcinoma (9,10). In addition, advances in next-generation-sequencing have recently made it possible to perform a detailed comprehensive genomic characterization of lung SCC as well as adenocarcinoma. A whole genome sequencing study by TCGA reported recurrent mutations in 18 genes, including *TP53* and genes in the *CDKN2A/RB1*, *NFE2L2/KEAP1/CUL3*, *PI3K/AKT*, and *SOX2/*

TP63/NOTCH1 pathways (7). A comparative mutational analysis of lung SCC patients in East Asia also demonstrated a high rate of mutations in *TP53*, *RBI*, *PTEN*, *NFE2L2*, *KEAP1*, *MLL2*, and *PIK3CA* (11). We found a similar rate of mutations in genes, including *TP53*, *PTEN*, and *FGFR*. With respect to smoking status, a 10-fold higher mutation rate has been reported in smokers compared to never-smokers (12), and Ras association domain family 1A (*RASSF1A*) methylation and *FGFR1* amplification occur more frequently in ever-smokers than in never-smokers (13,14). Interestingly, the tumors from smokers more often showed DNA hyper-methylation compared to those from never-smokers, but tumors from never-smokers with second-hand smoke exposure had a similar hyper-methylation rate to those from smokers (15).

The mutational analysis of the current study focused on SCC tumors from never-smokers. Only 2 small subsets of never-smokers with SCC have previously been reported, without clinical implications (7,11). In addition, our analysis included the mutational profile of both early and advanced stage disease, in contrast to previous studies using tumor samples from resected lung specimens. We also performed a survival analysis of never-smokers with SCC.

This study has several limitations. The number of samples used for mutational analysis was quite small because of the low incidence of SCC in never-smokers, which was only 5.0% in our institution and 3.6%–10.2% in previous reports (16–18). Second, we could not survey the extent to which patients were exposed to environmental factors, including second-hand smoke. Lastly, our mutational profiles for these samples could not be validated using different sequencing technologies due to the limited amount of tumor tissue. However, the performance of Ion PGM technology has already been proven to have extremely high sensitivity in non-small-cell lung cancer, compared with Sanger sequencing (19,20). Lastly, we could not survey the extent to which patients were exposed to environmental factors, including second-hand smoke.

In conclusion, we found no significant differences in 50 hotspot mutations in lung SCCs between never-smokers and ever-smokers.

DISCLOSURE

The authors have no potential conflicts of interest to disclose.

AUTHOR CONTRIBUTION

Conceptualization: Park YS. Data curation: Choi SM, Lee J, Lee SM, Yim JJ, Yoo CG, Kim YW, Han SK. Investigation: Lee HY, Won JK, Kwon NJ, Lee CH, Park YS. Writing - original draft: Lee HY, Kwon NJ, Park YS. Writing - review & editing: Lee SH, Lee DS.

ORCID

Ha Youn Lee <http://orcid.org/0000-0002-7664-3120>
 Se-Hoon Lee <http://orcid.org/0000-0002-9219-3350>
 Jae-Kyung Won <http://orcid.org/0000-0003-1459-8093>
 Dong Soo Lee <http://orcid.org/0000-0001-9013-4835>
 Nak-Jung Kwon <http://orcid.org/0000-0003-0874-4146>
 Sun Mi Choi <http://orcid.org/0000-0002-0742-6085>
 Jinwoo Lee <http://orcid.org/0000-0003-0958-106X>
 Chang-Hoon Lee <http://orcid.org/0000-0001-9960-1524>
 Sang-Min Lee <http://orcid.org/0000-0002-1388-9318>
 Jae-Joon Yim <http://orcid.org/0000-0002-9605-0074>
 Chul-Gyu Yoo <http://orcid.org/0000-0003-1321-7834>
 Young Whan Kim <http://orcid.org/0000-0002-2768-2422>
 Sung Koo Han <http://orcid.org/0000-0003-0876-9781>
 Young Sik Park <http://orcid.org/0000-0003-0235-6943>

REFERENCES

- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011; 61: 69-90.
- Pfeifer GP, Denissenko ME, Olivier M, Tretyakova N, Hecht SS, Hainaut P. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* 2002; 21: 7435-51.
- Gibbons DL, Byers LA, Kurie JM. Smoking, p53 mutation, and lung cancer. *Mol Cancer Res* 2014; 12: 3-13.
- Kenfield SA, Wei EK, Stampfer MJ, Rosner BA, Colditz GA. Comparison of aspects of smoking among the four histological types of lung cancer. *Tob Control* 2008; 17: 198-204.
- Okazaki I, Ishikawa S, Sohara Y. Genes associated with susceptibility to lung adenocarcinoma among never smokers suggest the mechanism of disease. *Anticancer Res* 2014; 34: 5229-40.
- Le Calvez F, Mukeria A, Hunt JD, Kelm O, Hung RJ, Tanière P, Brennan P, Boffetta P, Zaridze DG, Hainaut P. TP53 and KRAS mutation load and types in lung cancers in relation to tobacco smoke: distinct patterns in never, former, and current smokers. *Cancer Res* 2005; 65: 5076-83.
- Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 2012; 489: 519-25.
- Singh RR, Patel KP, Routbort MJ, Reddy NG, Barkoh BA, Handal B, Kana-gal-Shamanna R, Greaves WO, Medeiros LJ, Aldape KD, et al. Clinical validation of a next-generation sequencing screen for mutational hotspots in 46 cancer-related genes. *J Mol Diagn* 2013; 15: 607-22.
- Cooper WA, Lam DC, O'Toole SA, Minna JD. Molecular biology of lung cancer. *J Thorac Dis* 2013; 5 Suppl 5: S479-90.
- Weiss J, Sos ML, Seidel D, Peifer M, Zander T, Heuckmann JM, Ullrich RT, Menon R, Maier S, Soltermann A, et al. Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Sci Transl Med* 2010; 2: 62ra93.
- Kim Y, Hammerman PS, Kim J, Yoon JA, Lee Y, Sun JM, Wilkerson MD, Pedamallu CS, Cibulskis K, Yoo YK, et al. Integrative and comparative genomic analysis of lung squamous cell carcinomas in East Asian patients. *J Clin Oncol* 2014; 32: 121-8.
- Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, Ma-

- her CA, Fulton R, Fulton L, Wallis J, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* 2012; 150: 1121-34.
13. Lee SM, Lee WK, Kim DS, Park JY. Quantitative promoter hypermethylation analysis of RASSF1A in lung cancer: comparison with methylation-specific PCR technique and clinical significance. *Mol Med Rep* 2012; 5: 239-44.
14. Seo AN, Jin Y, Lee HJ, Sun PL, Kim H, Jheon S, Kim K, Lee CT, Chung JH. FGFR1 amplification is associated with poor prognosis and smoking in non-small-cell lung cancer. *Virchows Arch* 2014; 465: 547-58.
15. Scesnaite A, Jarmalaite S, Mutanen P, Anttila S, Nyberg F, Benhamou S, Boffetta P, Husgafvel-Pursiainen K. Similar DNA methylation pattern in lung tumours from smokers and never-smokers with second-hand tobacco smoke exposure. *Mutagenesis* 2012; 27: 423-9.
16. Kawaguchi T, Takada M, Kubo A, Matsumura A, Fukai S, Tamura A, Saito R, Kawahara M, Maruyama Y. Gender, histology, and time of diagnosis are important factors for prognosis: analysis of 1499 never-smokers with advanced non-small cell lung cancer in Japan. *J Thorac Oncol* 2010; 5: 1011-7.
17. Kawaguchi T, Takada M, Kubo A, Matsumura A, Fukai S, Tamura A, Saito R, Maruyama Y, Kawahara M, Ignatius Ou SH. Performance status and smoking status are independent favorable prognostic factors for survival in non-small cell lung cancer: a comprehensive analysis of 26,957 patients with NSCLC. *J Thorac Oncol* 2010; 5: 620-30.
18. Yun YH, Lim MK, Jung KW, Bae JM, Park SM, Shin SA, Lee JS, Park JG. Relative and absolute risks of cigarette smoking on major histologic types of lung cancer in Korean men. *Cancer Epidemiol Biomarkers Prev* 2005; 14: 2125-30.
19. Scarpa A, Sikora K, Fassan M, Rachiglio AM, Cappellesso R, Antonello D, Amato E, Mafficini A, Lambiase M, Esposito C, et al. Molecular typing of lung adenocarcinoma on cytological samples using a multigene next generation sequencing panel. *PLoS One* 2013; 8: e80478.
20. Kim HS, Sung JS, Yang SJ, Kwon NJ, Jin L, Kim ST, Park KH, Shin SW, Kim HK, Kang JH, et al. Predictive efficacy of low burden EGFR mutation detected by next-generation sequencing on response to EGFR tyrosine kinase inhibitors in non-small-cell lung carcinoma. *PLoS One* 2013; 8: e81975.