

Novel Genetic Associations Between Lung Cancer and Indoor Radon Exposure

Jung Ran Choi^{1,*}, Sang-Baek Koh^{1,2,*}, Seong Yong Park³, Hye Run Kim⁴, Hyojin Lee⁵, Dae Ryong Kang⁶

¹Institute of Genomic Cohort and ²Department of Preventive Medicine, Yonsei University Wonju College of Medicine, Wonju, ³Department of Thoracic and Cardiovascular Surgery, Yonsei University College of Medicine, ⁴Department of Internal Medicine, Yonsei Cancer Center, Yonsei University College of Medicine, ⁵Macrogen Inc., Seoul, ⁶Center of Biomedical Data Science, Yonsei University Wonju College of Medicine, Wonju, Korea

Background: Lung cancer is the leading cause of cancer-related death worldwide, for which smoking is considered as the primary risk factor. The present study was conducted to determine whether genetic alterations induced by radon exposure are associated with the susceptible risk of lung cancer in never smokers.

Methods: To accurately identify mutations within individual tumors, next generation sequencing was conducted for 19 pairs of lung cancer tissue. The associations of germline and somatic variations with radon exposure were visualized using OncoPrint and heatmap graphs. Bioinformatic analysis was performed using various tools.

Results: Alterations in several genes were implicated in lung cancer resulting from exposure to radon indoors, namely those in epidermal growth factor receptor (*EGFR*), tumor protein p53 (*TP53*), NK2 homeobox 1 (*NKX2.1*), phosphatase and tensin homolog (*PTEN*), chromodomain helicase DNA binding protein 7 (*CHD7*), discoidin domain receptor tyrosine kinase 2 (*DDR2*), lysine methyltransferase 2C (*MLL3*), chromodomain helicase DNA binding protein 5 (*CHD5*), FAT atypical cadherin 1 (*FAT1*), and dual specificity phosphatase 27 (putative) (*DUSP27*).

Conclusions: While these genes might regulate the carcinogenic pathways of radioactivity, further analysis is needed to determine whether the genes are indeed completely responsible for causing lung cancer in never smokers exposed to residential radon.

(*J Cancer Prev* 2017;22:234-240)

Key Words: Lung cancer in never smokers, Radon exposure, Driver gene, Next generation sequencing

INTRODUCTION

Smoking has long been identified as a significant risk factor for lung cancers, and the global incidence and mortality rates of lung cancer continue to be correlated with tobacco smoking.¹ Meanwhile, however, approximately 25% of all lung cancer patients are lifelong never smokers and lung cancer in never smokers (LCINS) ranks as the seventh most common cause of cancer mortality worldwide.² Affecting lung carcinogenesis in never smokers, environmental tobacco smoke at home, radon, outdoor air pollution, cooking oil fumes, coal fumes, and asbestos

have been deemed to play important roles therein.^{1,3-6}

Recently, researchers outlined changes over the last 20 years in genes associated with lung cancer susceptibility.⁷ Among smokers, previous studies have revealed associations between the genes *GSTM1*, *GSTT1*, and *GSTP1* and higher ORs for lung cancer.^{1,8,9} However, among never smokers a recent investigation found no significant association between single or combined genotypes of *GSTM1*, *GSTT1*, or *GSTP1* and lung cancer risk.¹⁰ Additionally, *TP53* mutations have also been highlighted more frequently in lung carcinomas arising in smokers than in never smokers.¹¹⁻¹³ Moreover, other recent articles found *TP53* mutations

Received September 29, 2017, Revised November 21, 2017, Accepted November 23, 2017

Correspondence to: Dae Ryong Kang

Center of Biomedical Data Science, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju 26426, Korea

Tel: +82-33-742-0929, Fax: +82-33-747-0409, E-mail: dr.kang@yonsei.ac.kr

*These authors contributed equally to this work as co-first authors.

Copyright © 2017 Korean Society of Cancer Prevention

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

and methylation of the Ras association domain family 1A (*RASSF1A*) promoter to be more frequent in smokers with squamous cell lung cancer than in never smokers with adenocarcinoma.^{13,14}

Radon is the second leading cause of LCINS.^{15,16} Feasibly, biologic mechanisms by which radon emissions might elevate the risk for LCINS include genetic alterations, upregulation or downregulation of cytokines, and production of proteins related to the cell cycle.¹⁶ Among these, genetic alterations of pathways involved in detoxification of environmental carcinogens have been shown to heighten lung cancer risk.¹⁵ Recent articles have demonstrated an increased frequency of cytogenetic damage in people with DNA-repair gene variations related with chronic exposure to radon and have indicated that *ADPRT* and *NBS1* can be utilized as molecular genetic markers of increased radiosensitivity to long-term exposure to high concentrations of radon.^{16,17} Another study suggested that radon exposure in never smokers seems to be a risk factor for lung cancer and that LCINS subjects diagnosed at a younger age might have been exposed to higher indoor radon concentrations, indicating an accumulative effect for radon levels on lung cancer features.¹⁸

While several studies have outlined the role of many candidate genetic polymorphisms in LCINS and their interactions with smoking status, the genetic variations important in susceptibility to residential radon exposure among never smokers are still unclear.^{15,16} Accordingly, this study was designed to identify genetic alterations induced by radon exposure and their potential associations with the susceptible risk for LCINS.

MATERIALS AND METHODS

1. Ethics statement

We examined tumor tissue, normal tissue, and blood samples from 19 adults (1 male and 18 females) aged 41 to 80 years with lung cancer from 2015 to 2016. Peripheral blood from all patients was obtained from the Tumor Tissue Banking of Ajou University Medical Center, inclusion in which all participants provided written informed consent. All sample's histological types were adenocarcinoma. This study was approved by the Institutional Review Board of the Ajou University Medical Center according to the Helsinki Declaration (AJIRB-BMR-KSP-15-409).

2. Targeted next generation sequencing

Sufficient and good quality DNA from peripheral blood, normal tissue, and tumor tissues were collected from the 19 LCINS patients. DNA was extracted from peripheral blood

leukocytes and the tissues using standard protocols. To extract DNA, the Maxwell[®] 16 Tissue DNA Purification Kit (Promega, Madison, WI, USA) was used for tissue samples and the Maxwell[®] 16 LEV Blood DNA Kit (Promega) was used for blood samples. With 1 µg of input gDNA, we applied the Agilent SureSelect Target Enrichment protocol for Illumina paired-end sequencing (ver. B.3, June 2, 2015); in this experiment, the SureSelect Human All Exon V5 probe was used to generate standard exome capture libraries. PicoGreen and agarose gel electrophoresis was used to evaluate the quantity and quality of DNA samples. Diluted in EB buffer, 1 µg of DNA was sheared to 150 to 200 bp of target peak size using the Covaris LE220 focused-ultrasonicator (Covaris, Woburn, MA, USA), according to the manufacturer's recommendations. From the fragmented DNA, an 'A' was ligated to the 3' end, and then Agilent adapters were ligated to the fragments. After the ligation, the adapter ligated library went through PCR amplification. For exome capture, 5 µL of the SureSelect all exon capture library, hybridization buffers, blocking mixes, and RNase block were mixed with 250 ng of DNA library, according to the standard Agilent SureSelect Target Enrichment protocol. Then, by using the HiSeq[™] 2000 platform (Illumina, San Diego, CA, USA), captured libraries were sequenced with 101 base pair reads.

3. Sequence data analysis

Sequence data were mapped to the human genome, with the reference sequence UCSC assembly hg19 (NCBI build 37.1), using BWA aligner (ver. 0.5.9rc1). The output Sam files were converted to Bam files and were sorted with SAMtools (ver. 0.1.18). PCR duplicate reads were removed using Picard tools (ver. 1.5.9) before base substitution detection. Based on the BAM file previously generated, variant calling was conducted by SAMtools, SAMtools mpileup, bcftools view, and vcutils.pl. From vcf4 format files, the varFilter was applied with the maximum depth option '-D' set to 1,000; in this step, single nucleotide polymorphism (SNPs) and short indel candidates are detected at the nucleotide level. Variants were annotated by ANNOVAR (ver. November 2011) filtering with dbSNP version 135 and SNPs from the 1000 genome project. (Supplementary Table S1 and S2) Somatic variants were identified by VarScan (ver. 2.3.7).

4. Measurement of indoor radon levels

Between October 28, 2015 and May 30, 2016, indoor radon levels were measured at two sites in each of the study subjects' households. Alpha-track detectors (Raduet Model RSV-8; Radosys Ltd., Budapest, Hungary) were used as a passive radon measuring

device. The average concentration of radon in the indoor air was calculated from two points within the household. The measurement points were selected from the living room and a bedroom, spaces where residents of a household primarily spend most of their time. The measuring devices were positioned away from household electrical appliances, windows, and sealed drawers. The measurement period was 3 months.

5. Statistical analysis

Analysis was performed for patient characteristics between mutation positive and negative patient groups, and the percentage of mutation carriers in tumor tissue was compared with that of normal control tissue. VCF files from the SAMtools variant calling pipeline were merged to one vcf file, which was sorted by allele frequencies in the 1000 Genomes Project, and then, we filtered out variants with values of more than 0.5 to exclude effects from common SNPs.

RESULTS

1. Study population

Fifty-seven tissue and blood sample pairs from 19 individual patients were submitted for sequencing from May 2016 to July 2016. There were a total of 18 females and 1 male (Table 1). All tumors were diagnosed as non-small cell lung cancer; the vast majority were adenocarcinoma or poorly differentiated carcinoma.

2. Germline mutations

The tumor tissue, normal tissue, and blood sample pairs were successfully sequenced. Applying SAMtools, we identified 3,120 single-nucleotide variants (SNVs) in total DNA regions (data not shown) and 760 variants in exon regions (Fig. 1). We then plotted the distribution of minor allele frequencies across all identified variants. The minor allele frequencies demonstrated a clear bi-modal distribution, peaking at 0.5 and 1, a distribution expected for germline variants. Then, we selected 10 patients with common variants in exon regions, from which 49 variants in

37 genes were identified: *CHD5*, *RPS6KA1*, *DDR2*, and *PIK3C2B* exhibited nonsynonymous SNVs, while *FAT4* and *FAT1* showed both nonsynonymous and synonymous SNVs (Fig. 2).

3. Somatic mutations

In the 10 sample pairs with common variants, we discovered several genes with a median of two variants (range of 0 to 4) per sample. We also identified 68 somatic mutations in 38 genes, including unreported variants for lung cancer DNA. Six genes (*EGFR*, *TP53*, *NKX2.1*, *PTEN*, *CHD7*, and *PRB1*) were mutated in at least two independent lung cancer patients; variants were most commonly noted in *EGFR* (37.0%), *TP53* (21.0%), and *PTEN* (16.0%) (Fig. 3). In the 10 pairs, we analyzed genetic variations for both germline and somatic mutations, and 37 drive genes exhibited at least one or more variations (data not shown).

DISCUSSION

Although the genes responsible for radon-induced LCINS are unclear, screening for germline and somatic mutations in known tumor suppressor genes might provide more insights on predicting susceptibility to lung cancer. Herein, a customized panel was designed to capture all exons of 37 cancer susceptible genes related to LCINS. Using next-generation sequencing, we identified 68 variants in 10 of 19 LCINS patients. Remarkably, several germline mutations matched between sample pairs from the 10 LCINS patients, including mutations in *CHD5*, *RPS6KA1*, *EGFR*, *MLL3*, and *RPTOR* and deletions in *SMARCA2*, *DACHI*, and *MAP3K9*.

Studies suggest that impaired DNA repair capacity for double-strand breaks (DSBs) may confer inherent susceptibility to lung cancer in smokers. DSBs encompass the most noxious forms of DNA damage and, if not appropriately repaired, can provoke cell death or conversion to malignancy.¹⁹ Capable of

Table 1. Characteristics of the study populations (n=19)

Variable	Value
Age (yr)	64.31 ± 11.18
Sex (male/female)	1/18
Radon dose (Bq·m ⁻³)	61.47 ± 43.25
Maximum level	163.6
Minimum level	22.75

Values are presented as Mean ± SD or number only.

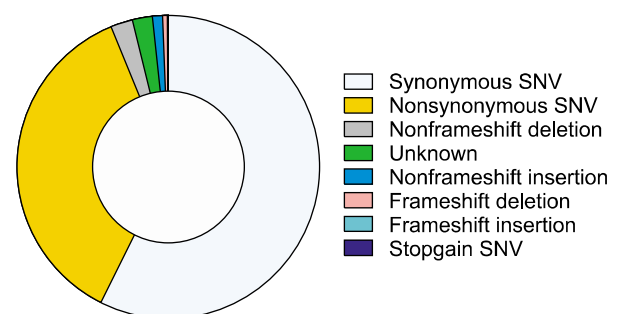


Figure 1. Distribution of germline variants in exon regions. SNV, single-nucleotide variant.

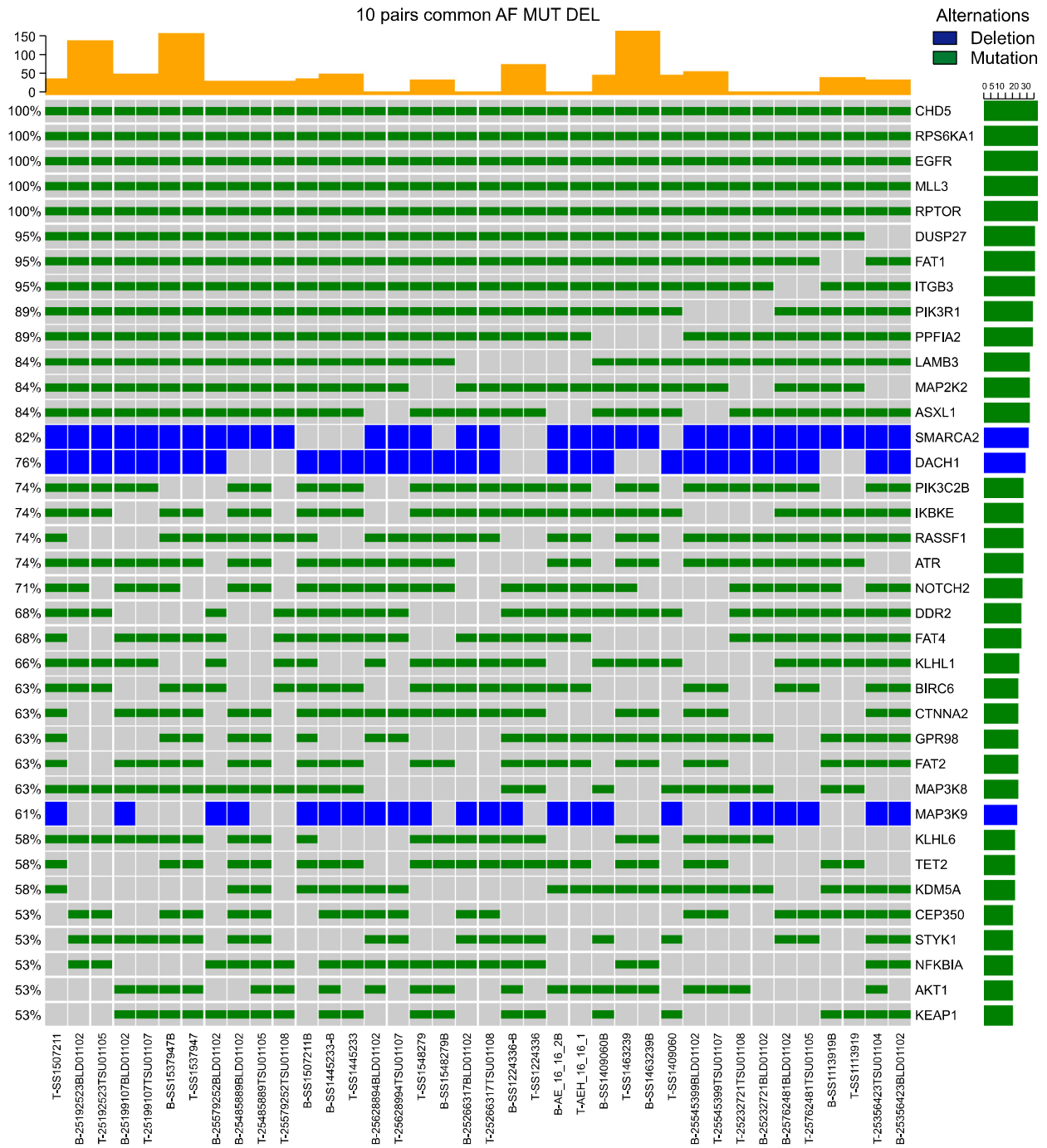


Figure 2. Germline alterations in 10 sample pairs.

generating DSBs in DNA, alpha particles radiated by radon and radon daughters can directly invade genomic DNA.^{20,21} Moreover, reactive oxygen species in the lungs arising from continuous radon exposure may bring about oxidative stress, resulting in pulmonary inflammation, tissue damage, and ultimately to chronic lung diseases, including chronic obstructive pulmonary disease, pulmonary fibrosis, and lung cancer.²²⁻²⁷ Genetic poly-

morphisms in genes important to DSBs repair and/or detoxification of environmental carcinogens, such as radon, can regulate lung cancer risk. Animal models have demonstrated that several gene polymorphisms may work together to increase an individual's risk for lung cancer.¹⁵ Ruano-Ravina et al.¹⁵ showed that deletions in *GSTM1* and *GSTT1* elevate the risk of lung cancer in subjects exposed to radon and suggested that these genes might

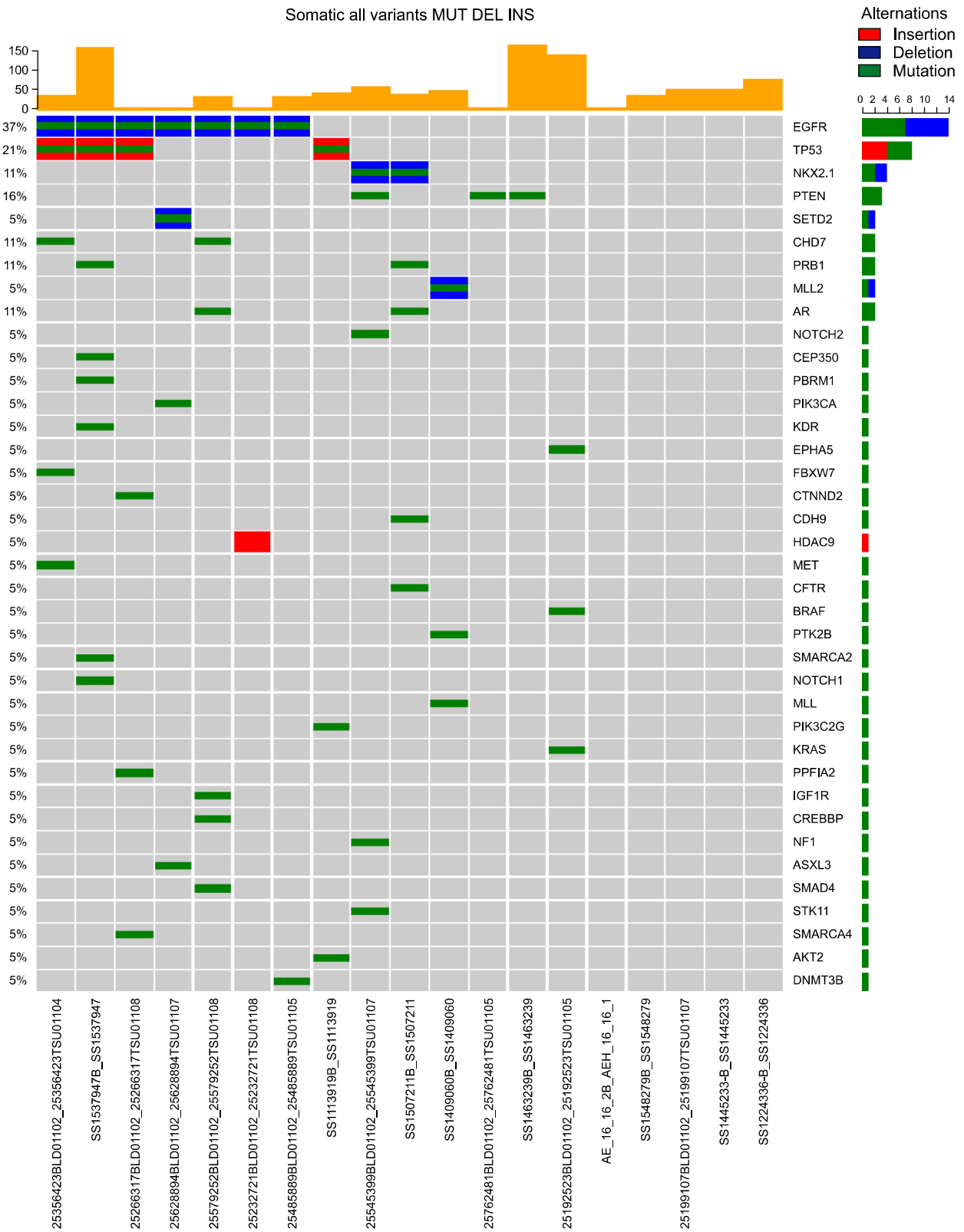


Figure 3. Somatic alterations in 19 patients.

control the carcinogenic pathway associated with alpha radiation. However, we did not find these genetic features in Korean LCINS exposed to high radon levels. Meanwhile, among former uranium miners, an association between genetic variations in the haplotype block of *SIRT1* and the risk for squamous cell carcinoma was described.¹⁹ Also, 16 genes involved in non-homologous end joining DNA repair, such as *PRKDC*, as well as histone acetylation and deacetylation, were identified.¹⁹ Again, however, we were unable to identify mutations in genes involved in DNA repair among the LCINS patients in the present study; this is likely because we included never smoking patients with non-small cell lung cancer exposed to high levels of indoor radon at their residence.

Among the cytokines and chemokines produced by tenacious pulmonary inflammation in response to constant radon exposure, interleukin-6 (IL-6) has been found to play an essential role in enhancing cancer development in *in vitro* and *in vivo* models of lung carcinogenesis. Leng et al.²⁸ revealed an association between four IL-6 promoter variants that influence binding of transcription factors and lung squamous cell carcinoma in former uranium miners exposed to high levels of radon. However, no variations in genes involved in cytokines and chemokines were discovered in our study.

Recently, carcinogenic exposure has been shown to play a part in the mutation of *TP53* (p53) in human cancers among workers exposed to occupational carcinogens.²⁹ Similar to previous studies, we noted somatic alterations in *TP53*. We presume that our findings may provide insight into how genetic variants within *TP53* can influence the function of tumor suppressors, such as p53.

In clinical practice, application of predictive biomarkers has enabled the selection of lung cancer patients for treatment with tyrosine-kinase inhibitors. For appropriate tyrosine kinase inhibitor (TKI) treatment, mutations in *EGFR* must be determined.³⁰⁻³³ Interestingly, we also highlighted variants in *EGFR* as common among LCINS patients. Further studies seeking to verify this association in LCINS are warranted.

A few limitations should be considered when interpreting the results of this study. First despite this study should be defined as an exploratory study, the sample size is extremely small. Second, genes were not found to differ according accumulated levels of indoor radon exposure. Residents in radon-prone areas are actually exposed to much larger amounts of radon than residents in areas with lower radon levels, if one were to consider accumulated levels of indoor and/or outdoor radon. As well, patients with lung cancer who resided in radon-prone areas might

have an advantage in the evaluation of dose-response relationships between indoor radon levels and lung cancer risk, because the radon exposure range is wider than that in areas with lower radon levels.³⁴ Third, we evaluated capture-based targeted DNA sequencing as a new approach for testing a broad spectrum of point mutations (SNVs) and short insertion-deletions (indels) possibly related to LCINS. However, there were no references with which to compare the noted genetic alterations induced by radon exposure and the risk of lung cancer in Korean never smokers.

Further studies are warranted to examine the associations between residential radon concentrations and LCINS.

ACKNOWLEDGMENTS

This study was supported by the Korean Ministry of Environment as part of the "Environmental Health Action Program" (grant number 2015001350002).

CONFLICTS OF INTEREST

No potential conflicts of interest were disclosed.

SUPPLEMENTARY MATERIALS

Supplementary Materials can be found via <https://doi.org/10.15430/JCP.2017.22.4.234>.

REFERENCES

1. Okazaki I, Ishikawa S, Ando W, Sohara Y. Lung adenocarcinoma in never smokers: problems of primary prevention from aspects of susceptible genes and carcinogens. *Anticancer Res* 2016;36: 6207-24.
2. Huang Y, Wang R, Pan Y, Zhang Y, Li H, Cheng C, et al. Clinical and genetic features of lung squamous cell cancer in never-smokers. *Oncotarget* 2016;7:35979-88.
3. Islami F, Torre LA, Jemal A. Global trends of lung cancer mortality and smoking prevalence. *Transl Lung Cancer Res* 2015;4:327-38.
4. Boffetta P. Involuntary smoking and lung cancer. *Scand J Work Environ Health* 2002;28 Suppl 2:30-40.
5. Kurahashi N, Inoue M, Liu Y, Iwasaki M, Sasazuki S, Sobue T, et al. Passive smoking and lung cancer in Japanese non-smoking women: a prospective study. *Int J Cancer* 2008;122:653-7.
6. Couraud S, Zalcmán G, Milleron B, Morin F, Souquet PJ. Lung cancer in never smokers—a review. *Eur J Cancer* 2012;48:1299-311.
7. Okazaki I, Sugita M, Matsuki H, Billah SM, Watanabe T. Additional candidates to conventional genes susceptible for lung cancer and changing trend in Japan. *Oncol Rep* 2010;23:1493-500.
8. Pliarchopoulou K, Voutsinas G, Papaxoinis G, Florou K, Skondra

- M, Kostaki K, et al. Correlation of CYP1A1, GSTP1 and GSTM1 gene polymorphisms and lung cancer risk among smokers. *Oncol Lett* 2012;3:1301-6.
9. Timofeeva M, Kropp S, Sauter W, Beckmann L, Rosenberger A, Illig T, et al. Genetic polymorphisms of MPO, GSTT1, GSTM1, GSTP1, EPHX1 and NQO1 as risk factors of early-onset lung cancer. *Int J Cancer* 2010;127:1547-61.
 10. Wenzlaff AS, Cote ML, Bock CH, Land SJ, Schwartz AG. GSTM1, GSTT1 and GSTP1 polymorphisms, environmental tobacco smoke exposure and risk of lung cancer among never smokers: a population-based study. *Carcinogenesis* 2005;26:395-401.
 11. Toyooka S, Tsuda T, Gazdar AF. The TP53 gene, tobacco exposure, and lung cancer. *Hum Mutat* 2003;21:229-39.
 12. Le Calvez F, Mukeria A, Hunt JD, Kelm O, Hung RJ, Tanière P, et al. 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.
 13. Okazaki I, Ishikawa S, Sahara Y. Genes associated with susceptibility to lung adenocarcinoma among never smokers suggest the mechanism of disease. *Anticancer Res* 2014;34:5229-40.
 14. 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.
 15. Ruano-Ravina A, Pereyra MF, Castro MT, Pérez-Ríos M, Abal-Arca J, Barros-Dios JM. Genetic susceptibility, residential radon, and lung cancer in a radon prone area. *J Thorac Oncol* 2014;9:1073-80.
 16. Choi JR, Park SY, Noh OK, Koh YW, Kang DR. Gene mutation discovery research of non-smoking lung cancer patients due to indoor radon exposure. *Ann Occup Environ Med* 2016;28:13.
 17. Sinitsky MY, Larionov AV, Asanov MA, Druzhinin VG. Associations of DNA-repair gene polymorphisms with a genetic susceptibility to ionizing radiation in residents of areas with high radon (²²²Rn) concentration. *Int J Radiat Biol* 2015;91:486-94.
 18. Ruano-Ravina A, Prini-Guadalupe L, Barros-Dios JM, Abal-Arca J, Leiro-Fernández V, González-Silva AI, et al. Exposure to residential radon and lung cancer in never-smokers: the preliminary results of the LCRINS study. *Arch Bronconeumol* 2012;48:405-9.
 19. Leng S, Picchi MA, Liu Y, Thomas CL, Willis DG, Bernauer AM, et al. Genetic variation in SIRT1 affects susceptibility of lung squamous cell carcinomas in former uranium miners from the Colorado plateau. *Carcinogenesis* 2013;34:1044-50.
 20. Prise KM, Pinto M, Newman HC, Michael BD. A review of studies of ionizing radiation-induced double-strand break clustering. *Radiat Res* 2001;156:572-6.
 21. Sethi TK, El-Ghamry MN, Kloecker GH. Radon and lung cancer. *Clin Adv Hematol Oncol* 2012;10:157-64.
 22. Archer VE, Renzetti AD, Doggett RS, Jarvis JQ, Colby TV. Chronic diffuse interstitial fibrosis of the lung in uranium miners. *J Occup Environ Med* 1998;40:460-74.
 23. Iyer R, Lehnert BE, Svensson R. Factors underlying the cell growth-related bystander responses to alpha particles. *Cancer Res* 2000;60:1290-8.
 24. Mapel DW, Coultas DB, James DS, Hunt WC, Stidley CA, Gilliland FD. Ethnic differences in the prevalence of nonmalignant respiratory disease among uranium miners. *Am J Public Health* 1997;87:833-8.
 25. Narayanan PK, Goodwin EH, Lehnert BE. Alpha particles initiate biological production of superoxide anions and hydrogen peroxide in human cells. *Cancer Res* 1997;57:3963-71.
 26. Rosanna DP, Salvatore C. Reactive oxygen species, inflammation, and lung diseases. *Curr Pharm Des* 2012;18:3889-900.
 27. Schubauer-Berigan MK, Daniels RD, Pinkerton LE. Radon exposure and mortality among white and American Indian uranium miners: an update of the Colorado Plateau cohort. *Am J Epidemiol* 2009;169:718-30.
 28. Leng S, Thomas CL, Snider AM, Picchi MA, Chen W, Willis DG, et al. Radon Exposure, IL-6 promoter variants, and lung squamous cell carcinoma in former uranium miners. *Environ Health Perspect* 2016;124:445-51.
 29. Vähäkangas K. TP53 mutations in workers exposed to occupational carcinogens. *Hum Mutat* 2003;21:240-51.
 30. Sequist LV, Waltman BA, Dias-Santagata D, Digumarthy S, Turke AB, Fidias P, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci Transl Med* 2011;3:75ra26.
 31. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res* 2013;19:2240-7.
 32. Walter AO, Sjin RT, Haringsma HJ, Ohashi K, Sun J, Lee K, et al. Discovery of a mutant-selective covalent inhibitor of EGFR that overcomes T790M-mediated resistance in NSCLC. *Cancer Discov* 2013;3:1404-15.
 33. Cross DA, Ashton SE, Ghiorghiu S, Eberlein C, Nebhan CA, Spitzler PJ, et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov* 2014;4:1046-61.
 34. Sheen S, Lee KS, Chung WY, Nam S, Kang DR. An updated review of case-control studies of lung cancer and indoor radon-Is indoor radon the risk factor for lung cancer? *Ann Occup Environ Med* 2016;28:9.