Radon Exposure-induced Genetic Variations in Lung Cancers among Never Smokers

Jung Ran Choi, Sang Baek Koh, Hye Ryun Kim, Hyejin Lee, and Dae Ryong Kang

1Institute of Genomic Cohort, Yonsei University Wonju College of Medicine, Wonju, Korea
2Department of Preventive Medicine, Yonsei University Wonju College of Medicine, Wonju, Korea
3Department of Internal Medicine, Yonsei Cancer Center, Yonsei University College of Medicine, Seoul, Korea
4Macrogen Inc., Seoul, Korea

ABSTRACT

Background: Lung cancer in never smokers (LCINS) differs etiologically and clinically from lung cancer attributed to smoking. After smoking, radon exposure is the second leading cause and the primary risk factor of lung cancer among never smokers. Exposure to radon can lead to genetic and epigenetic alterations in tumor genomes affecting genes and pathways involved in lung cancer development. The present study sought to explore genetic alterations associated with LCINS exposed to radon gas indoors.

Methods: Genetic associations were assessed via a case-control study of LCINS (39 cases and 30 controls) using next generation sequencing. Associations between genetic mutations and high exposure to radon were investigated by OncoPrint and heatmap graphs. Bioinformatic analysis was conducted using various tools. According radon exposure levels, we divided subjects in two groups of cases and controls.

Results: We found that ABL2 rs117218074, SMARCA4 rs2288845, PIK3R2 rs142933317, MAPK1 rs1803545, and androgen receptor (AR) rs66766400 were associated with LCINS exposed to high radon levels. Among these, Chromodomain helicase DNA-binding protein 4 (CHD4) rs74790047, TSC2 rs2121870, and AR rs66766408 were identified as common exonic mutations in both lung cancer patients and normal individuals exposed to high levels of radon indoor.

Conclusion: We identified that CHD4 rs74790047, TSC2 rs2121870, and AR rs66766408 are found to be common exonic mutations in both lung cancer patients and normal individuals exposed to radon indoors. Further analysis is needed to determine whether these genes are completely responsible for LCINS exposed to residential radon.

Keywords: Lung Cancer in Never Smokers; Radon Exposure; Genetic Alteration; Next Generation Sequencing

INTRODUCTION

Lung cancer is generally associated with tobacco smoking. However, lung cancer in never smokers (LCINS) has been recognized as a disease independent of smoking-associated lung cancer. According to the World Health Organization, the incidence of LCINS is almost 25%, ranking as the seventh most common cause of cancer-related death. Recently, remarkable variance in the proportion of LCINS among all lung cancers has been noted, ranging from nearly
10% in Western males and up to approximately 40% in Asian females. However, ethnic/genetic features and/or environmental characteristics of LCINS remain unclear.

Radon is the most important natural source ionizing radiation to which humans are exposed and is the second leading cause of LCINS. Exposure to radon is estimated to be associated with more than 20,000 lung cancer deaths per year in the United States. Studies suggest that up to 30% of all lung cancer deaths among non-occupationally exposed never-smokers might be linked to indoor radon. The Environmental Protection Agency (EPA) action level is 148 Bq/m³, and the World Health Organization has recently lowered the recommended radon exposure levels to below 100 Bq/m³.

Possible biologic mechanisms by which radon exposure might increase the risk of LCINS are gene mutations and chromosome alterations affecting the production of proteins associated with cell cycle control. Recently, genome-wide association studies have investigated specific chromosomal loci (15q24-15q25 and 5p15.33) as regions associated with LCINS. The 5p15.33 locus contains two candidate genes, TERT and CLPTM1, and Hsiung et al. suggested that the 5p15.33 region is associated with the risk of LCINS in Asian females. Another study recommended that GSTM1 and GSTT1 deletion increases the risk of lung cancer and that these genes might regulate the carcinogenic pathway via radon radiation exposure.

Although several studies have been demonstrated the role of candidate genes for developing LCINS, the genetic alterations for susceptibility to LCINS upon residential radon exposure are still unclear, and further studies are needed to analyze associations between exposure to radon and LCINS. Accordingly, the present study aimed to explore genetic alterations associated with LCINS exposed to radon indoors.

**METHODS**

**Cases and controls**
We designed a hospital-based case-control study and examined blood samples from 69 adults (39 cases and 30 controls) aged 41–80 years with lung cancer in 2015–2016. The lung cancer cases included 39 patients treated at Ajou University Medical Center, Sinchon Severance Hospital, and Seoul St. Mary’s Hospital. The control group included 30 unrelated volunteers who had no history of cancer. We excluded participants who were diagnosed with lung cancer-related disease and current smokers, which left 69 participants as the final samples. All participants were subjected to thorough medical history checks, including smoking history, clinical examination, and routine laboratory investigations. Diagnosis of lung cancer was confirmed by cytological or histopathological examination of tumor biopsies.

**Measurement of indoor radon levels**
Between October 28, 2015 and May 30, 2016, indoor radon levels were measured at two sites in each household of the subjects of the study population. 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 the 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. After estimating indoor radon levels...
levels, we divided the participants into high radon and low radon level groups based on indoor radon levels greater or less than 100 Bq/m².

**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 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 a target peak size of 150–200 bp using a Covaris LE220 focused-ultrasonicator (Covaris, Woburn, MA, USA), according to the manufacturer’s instructions. 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 polymerase chain reaction (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.

**Sequence data analysis**

Sequence data were mapped to the human genome, with the reference sequence UCSC assembly hg19 (National Center for Biotechnology Information [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 vcfutils.pl. From vcf4 format files, the varFilter was applied with the maximum depth option ‘-D’ set to 1,000: in this step, SNPs and short insertion-deletions (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. Somatic variants were identified by VarScan (ver. 2.3.7).

**Statistical analysis**

Patient characteristics were analyzed between mutation positive and negative patient groups, and the percentage of mutation carriers in tumor was compared with that in normal control. 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 greater than 0.5 to exclude effects from common SNPs. Based on radon exposure of 100 Bq/m², we divided 69 subjects into two groups and defined as group 1 (under 100 Bq/m²) and group 2 (over 100 Bq/m²). Then, we compared the exonic variants of each group according to allele frequency, location, amino acid change etc. and filtered out genetic variants existed in greater than 100 Bq/m², group 2, compared to that of group 1. In 39 lung cancer patients, we analyzed the variants which existed in exonic region in subjects with high radon level compared to that in patients had low radon level. In 30 normal control, we filtered out the variants with values lower than 0.05 of allele frequencies and compared the variants in normal controls who have high radon exposure.
**Ethics statement**
This study was approved by the Institutional Review Board of Ajou University Medical Center according to the Helsinki Declaration (AJIRB-MED-SUR-16-220). All participants provided their written informed consent.

**RESULTS**

**Study populations**
There was no significant difference among cases and controls in terms of mean age; however, there was a significant difference among cases and controls regarding the distribution of males and females ($P < 0.001$) (data not shown). Radon levels were $67.2 \pm 38.9$ Bq/m$^3$ in patients and $81.2 \pm 51.0$ Bq/m$^3$ in controls. All tumors were diagnosed as non-small cell lung cancer: the vast majority had adenocarcinoma or poorly differentiated carcinoma.

**Somatic mutations in populations**
In 69 participants, 39 cases and 30 controls had complete radon measurements, and the 69 individuals did not have information on smoking history that would allow for the calculation of lifetime tobacco consumption. Dividing subjects into two groups based on radon exposure of 100 Bq/m$^3$, we noted that several genetic variations, including ABL2 rs117218074, CHD4 rs74790047, SMARCA4 rs2288845, TSC2 rs2121870, and androgen receptor (AR) rs66766400, were associated with high radon exposure. The subjects with high radon levels had significant associations with nonsynonymous alteration and non-frameshift mutation (Fig. 1). We discovered several variants which existed in exonic region in subjects with high radon level and these variants were EPHA2, DUSP27, XPO1, and CASP8 (Fig. 1).

In lung cancer patients, we compared the exonic variants in subjects with high radon level compared with that in patients who have low radon level. We found that single nucleotide polymorphism, such as ABL2 rs117218074, SMARCA4 rs2288845, PIK3R2 rs142933317, MAPK1 rs1803545, and AR rs66766400, were associated with LCNIS exposed to high radon levels (Fig. 2). The patients who have high radon level had significant associations with insertion and deletion mutation in AR gene (Fig. 2). In controls with high radon exposure levels, there were various exonic alterations, such as ATR rs146405935 and TRIO rs55920001. In contrast, the normal control with high radon level had significant associations with non-frameshift and deletion mutation in AR gene (Fig. 3). Among these, CHD4 rs74790047, TSC2 rs2121870, and AR rs66766408 were identified common exonic mutations in both of lung cancer patients and normal individuals with high exposure to radon (Figs. 2 and 3). In the high radon group, an interaction between radon exposure and genetic variations for the risk of LCINS was identified.

**DISCUSSION**
In order to analyze driver mutations among LCINS who have been exposed to radon, we identified several genetic alterations associated therewith. Single nucleotide variations, including CHD4 rs74790047, TSC2 rs2121870, and AR rs66766408, were found to be common exonic mutations in both lung cancer patients and normal individuals exposed to high levels of radon indoors.

Radon is a radioactive gas generated naturally by the collapse of uranium from soils and rocks.\(^4\) Radon decays into active progenies with an electrical charge and can be inhaled when combined
with natural aerosols, ultimately arriving at lung epithelial cells. In lung tissue, cumulative radon progeny decays to produce alpha-particles that lead to DNA damage through double-strand breaks (DSBs) and large chromosomal alterations. Reports have shown that impaired DNA repair ability of DSBs may be responsible for distinct susceptibility to lung cancer in smokers. If not repaired adequately, DSBs can induce cell death or conversion to malignancy. In the process of repairing DSBs, alpha particles radiated by radon and radon daughters can directly invade genomic DNA. Moreover, reactive oxygen species in the lungs developing from continuous radon exposure may bring about oxidative stress, leading to pulmonary inflammation and ultimately to chronic lung diseases, including lung cancer.

The association between radon and lung cancer has mainly been derived from epidemiologic studies of miners. In further analysis, it has been observed that up to 30% of lung cancer deaths among non-occupationally exposed never-smokers might be connected with indoor radon. Recently, genome-wide association studies have identified specific chromosomal loci as regions associated with LCINS: Iwamoto et al. indicated that EPAS1 rs4953354 may be an essential susceptible biomarker for the development of lung adenocarcinoma, especially in

Fig. 1. Exonic variants according to exposure to radon in the study population. In subjects with high on radon exposure of 100 Bq/m³, there were several exonic genetic variations, including ABL2 rs117218074, CHD4 rs74790047, SMARC4 rs2288845, TSC2 rs2121870, and AR rs66766400. The subjects with high radon levels had significant associations with nonsynonymous alteration and non-frameshift mutation.
female never-smokers. In this study, we discovered ABL2 rs117218074, SMARCA4 rs2288845, PIK3R2 rs142933317, MAPK1 rs1803545, and AR rs66766400 in LCINS exposed to high radon levels. Also, we identified common genetic alterations, including CHD4 rs74790047, TSC2 rs2121870, and AR rs66766408, in both LCINS and normal individuals.
Fig. 3. Exonic alterations in normal individuals according to radon exposure levels. In controls with high radon exposure levels, there were various exonic alterations, such as ATR rs146405935 and TRIO rs55920001. The normal control with high radon level had significant associations with non-frameshift and deletion mutation in AR gene.
Chromodomain helicase DNA-binding protein 4 (CHD4) plays an important role in chromatin remodeling, and has been implicated in the development of cancer.33 Also, CHD4 is one of the key enzymes in the DNA-damage response and regulates the cell cycle. Its dysfunction can lead to the development of cancer. Yamada et al.33 suggested that nonsynonymous single nucleotide variants (SNVs) in CHD4 (rs74790047) were associated with cancer and might connect with smoking habits to increase cancer risk. Similar to this previous study, we identified that an SNV (rs74790047) in the CHD4 gene contributes to the risk of cancer in both LCINS and normal people with high radon levels. In result, we assume that CHD4 nonsynonymous polymorphisms is associated with high radon levels in individuals who never smoke and normal subjects. Recently, it has been suggested that TSC1 loss synergizes with Kras mutation to enhance lung tumorigenesis in mice, however, this mutation is a rare event in human lung cancer.35 In our study, we discovered that TSC1 mutation is associated with high radon levels in LCINS; however, we did not identify an association with LCINS due to insufficient data. Previous studies regarding the association between AR gene CAG repeat length and ovarian cancer risk reported inconsistent results.36 Overall, there was no association between the AR CAG repeat polymorphism and ovarian cancer risk. Meanwhile, short CAG repeat polymorphism was associated with increased ovarian cancer risk in African Americans and Chinese under the dominant model.36 However, there were no studies to investigate an association of AR single nucleotide polymorphisms with lung cancer in previous studies. Further well-designed epidemiological and functional studies are needed to elucidate the role of AR in lung carcinogenesis.

A few limitations should be considered when interpreting the results of this study. Despite an exploratory study, the sample size is extremely small. Our findings may not be generalizable to other populations, especially those of different ethnicity. Also, we evaluated capture-based targeted DNA sequencing as a new approach for testing a broad spectrum of point mutations (SNVs) and short 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.

In this study, CHD4 rs74790047, TSC2 rs2121870, and AR rs66766408, were found to be common exonic mutations in both lung cancer patients and normal individuals exposed to high levels of radon indoors. The relative significance of radon exposure and genetic polymorphisms in the development of LCINS with high radon levels has not been well investigated in Korean adults. Further analysis is needed to determine whether several notable genes are completely responsible for LCINS exposed to residential radon.

REFERENCES


