

Identification of Serial DNA Methylation Changes in the Blood Samples of Patients with Lung Cancer

Da Hye Moon, M.D.¹, Sung Ok Kwon, Ph.D.², Woo Jin Kim, M.D.^{1,3} and Yoonki Hong, M.D.^{1,3}

¹Department of Internal Medicine, ²Biomedical Research Institute, Kangwon National University Hospital, Chuncheon,

³Department of Internal Medicine, Kangwon National University School of Medicine, Chuncheon, Korea

Background: The development of lung cancer results from the interaction between genetic mutations and dynamic epigenetic alterations, although the exact mechanisms are not completely understood. Changes in DNA methylation may be a promising biomarker for early detection and prognosis of lung cancer. We evaluated the serial changes in genome-wide DNA methylation patterns in blood samples of lung cancer patients.

Methods: Blood samples were obtained for three consecutive years from three patients (2 years before, 1 year before, and after lung cancer detection) and from three control subjects (without lung cancer). We used the MethylationEPIC BeadChip method, which covers the 850,000 bp cytosine-phosphate-guanine (CpG) site, to conduct an epigenome-wide analysis. Significant differentially methylated regions (DMRs) were identified using p-values <0.05 in a correlation test identifying serial methylation changes and serial increase or decrease in β value above 0.1 for three consecutive years.

Results: We found three significant CpG sites with differentially methylated β values and 7,105 CpG sites with significant correlation from control patients without lung cancer. However, there were no significant DMRs. In contrast, we found 11 significant CpG sites with differentially methylated β values and 10,562 CpG sites with significant correlation from patients with lung cancer. There were two significant DMRs: cg21126229 (*RNF212*) and cg27098574 (*BCAR1*).

Conclusion: This study revealed DNA methylation changes that might be implicated in lung cancer development. The DNA methylation changes may be the possible candidate target regions for the early detection and prevention of lung cancer.

Keywords: DNA Methylation; Lung Neoplasms; Biomarkers

Address for correspondence: Yoonki Hong, M.D.

Department of Internal Medicine, Kangwon National University School of Medicine, 1 Gangwondaehak-gil, Chuncheon 24341, Korea

Phone: 82-32-258-9364, **Fax:** 82-32-258-2404

E-mail: h-doctor@hanmail.net

*The abstract of this paper was presented at the IASLC 18th World Conference on Lung Cancer, which was held 15–18 Oct 2017 at the Conference Center, Pacifico Yokohama, Japan.

Received: May. 6, 2018

Revised: Jun. 21, 2018

Accepted: Jul. 31, 2018

Published online: Sep. 28, 2018

©It is identical to the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>).



Copyright © 2019
The Korean Academy of Tuberculosis and Respiratory Diseases.

Introduction

Worldwide, lung cancer is the leading cause of cancer death¹. Although the survival rate trends have increased in recent years (due to early detection by low-dose computed tomography [CT] screening and new therapeutic agents, such as molecular targeted agents or immunotherapy), the 5-year survival rate remains poor, at about 20%–30%^{2,3}. Therefore, in the area of lung cancer research, it is an important task still to identify biomarkers that may be useful for early diagnosis.

The genetic changes found in lung cancer can mainly be explained by the interaction between permanent genetic mutations and dynamic epigenetic changes⁴. Since permanent genetic alterations in lung cancer have been studied extensively there has been substantial progress in molecular-targeted

therapeutics in the treatment of patients with non-small-cell lung cancer (NSCLC)⁵. However, the overall survival rate reduction effect of the molecular targeted therapeutics for lung cancer is insignificant. Although somatic genetic mutations might play a major role in cancer development, epigenetic changes are known to play more variable roles, since they can either repress the expression of tumor suppressor genes or activate the expression of oncogenes^{4,6}. Therefore, epigenetic changes could be a target to uncover biomarkers for the early detection of lung cancer.

Epigenetic changes are defined as heritable modifications in gene expression which do not alter the DNA coding sequence directly⁷. Over the past decade, they have been studied increasingly as biomarkers for the early detection of cancer and therapeutic targets^{8,9}. Epigenetic changes observed in oncogenesis consist of aberrant DNA methylation patterns, histone modifications, and regulation by noncoding RNAs⁷. DNA methylation is the most studied epigenetic change and is known to repress gene expression and maintain genomic stability⁹. Therefore, DNA methylation study may give a potential clue for identifying biomarkers for early detection of lung cancer or to predict tumor progression.

We evaluated serial changes in genome-wide DNA methylation patterns in blood samples from lung cancer patients.

Materials and Methods

1. Study subjects and the preparation of tissue samples

For this study, we used blood samples obtained from a Korean chronic obstructive pulmonary disease (COPD) cohort¹⁰. They were prospectively recruited from 2012 to 2015 to observe clinical outcomes of Koreans living near cement plants. They will be followed up for 10 years and will be monitored for clinical information, blood/urine samples, and surveyed via questionnaire. The cohort recruited 445 patients from 2012 till

December 2015. Among them, five patients developed lung cancer till the end of 2016, two of small cell lung cancer and three of NSCLC. Among them, blood samples were obtained for three consecutive years from three patients who had developed NSCLC (2 years before, 1 year before, and after lung cancer detection) and from three control subjects who did not have lung cancer and were matched for age and smoking history (Figure 1). Appropriate informed consent was obtained from all the patients, and the Institutional Review Board of Kangwon National University Hospital approved the study (IRB no. KNUH2012-06-007). All the COPD cohort study conduct adheres to Good Clinical Practice Guidelines and the tenets of the Declaration of Helsinki.

2. Genomic DNA preparation and DNA methylation profiling

We conducted an epigenome-wide analysis using a MethylationEPIC BeadChip kit, which covers the 850,000 bp cytosine-phosphate-guanine (CpG) site. The methylation value (β), a ratio between methylated probe intensity and total probe intensity, was interpreted as the proportion of methylation; β values range from 0 (unmethylated) to 1 (methylated).

The DNA quality was checked with a spectrophotometer (NanoDrop ND-1000 UV-vis; NanoDrop Technologies, Wilmington, DE, USA), and genomic DNA was diluted to 50 ng/ μ L using a Quant-iT PicoGreen quantitation assay (Invitrogen, Carlsbad, CA, USA). Bisulfite-conversion using an EZ DNA methylation kit (Zymo Research, Irvine, CA, USA) was carried out according to the manufacturer's protocols (Supplementary Methods).

3. Statistical analysis

We used methylation β values because they are more easily interpretable as methylation changes than M values are, the \log_2 ratio of methylated probe intensity and unmethylated

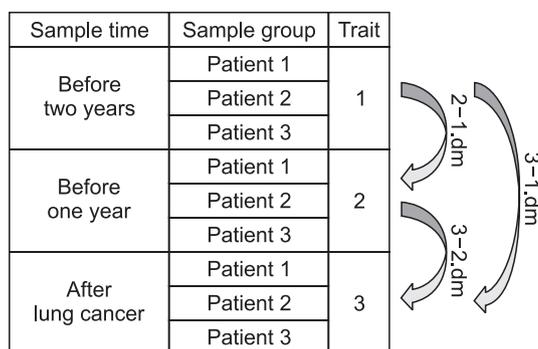


Figure 1. Blood samples were obtained for three consecutive years from three patients with non-small-cell lung cancer or control subjects and analyzed by delta mean and correlation test. Significant β changes were defined as when the delta mean of trait 2–1 and 3–2 was increased (≥ 0.1) or decreased (≤ -0.1), and significant correlation was defined when the p-value for the correlation test was > 0.05 . When both conditions were satisfied, significant differentially-methylated regions were obtained. dm: delta mean.

probe intensity.

In three patients who had developed lung cancer, we assigned continuous labels to each patient in the order of the sample collection year (1, 2 years before lung cancer detection; 2, 1 year before lung cancer detection; and 3, after lung cancer detection). The mean differentially-methylated level (β value) of each trait and correlation test, according to continuous sample collection order, was evaluated. Significant β changes were defined when the delta mean of trait 2–1 and 3–2 was increased or decreased by a factor of at least 0.1. Significant correlation was defined as when the p-value for the correlation test was <0.05 . Significant differentially methylated regions were obtained when both conditions were satisfied (Figure 1).

In addition, in the three control subjects, we assigned the continuous labels 1 to 3 in the order of sample collection year (1, first year; 2, second year; and 3, third year). The mean differentially methylated level (β values) of each trait and correlation test, according to continuous sample collection order, were evaluated. Significant β changes were defined when the delta mean of trait 2–1 and 3–2 was increased or decreased by a factor of at least 0.1, and significant correlation was defined when the p-value for the correlation test was >0.05 . When both conditions were satisfied, significant differentially methylated regions were obtained.

All the methylation data processing, statistical analyses, and visualizations were conducted in R 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results

The baseline characteristics of patients and control were summarized on Table 1.

We identified 11 regions with significant methylation changes in the three patients with lung cancer. The 11 regions were as follows: cg02740582, cg05130642, cg07850982, cg16057201, cg16298462, cg19536810, cg19565594, cg21126229, cg25304620, cg25994476, and cg27098574. In addition, 10,562 regions were identified as having significant correlation for continuous years. Significant differentially-methylated regions satisfying both conditions were found in two regions: cg21126229 and cg27098574 (Figure 2). The related genes for cg21126229 and cg27098574 were *RNF212* and *BCARI*. Figure 3 presents the clustering heatmaps.

Three regions with significant methylation changes were identified in the three control subjects. In addition, 7,105 regions were identified as having significant correlation for continuous years. However, there were no significant differentially-methylated regions which satisfied both conditions.

Meanwhile, to observe DNA methylation changes from another perspective, we analyzed cancer and control subjects by grouping them with same sample collection years. We sought

to select a position that showed the difference between the cancer and the control samples at the same time, ultimately to show the difference between the cancer samples and the controls, and, at the same time, to show the position indicating the serial change. However, we did not find significant differentially-methylated regions with significant methylation changes and correlation for continuous years.

Discussion

In this study, we identified differentially methylated regions associated with the development of NSCLC in 3-year serial blood samples obtained from patients with and without NSCLC. Two genes, *RNF212* and *BCARI*, that were associated with occurrence or early detection of NSCLC were found. Our study results may be used as a resource for finding biomarkers for the early detection of lung cancer.

Lung cancer is clinically manifested in various presentations and has few symptoms at the early stages. Early symptoms are subtle and non-specific and may lead to initial diagnosis being at an advanced stage for many patients with lung cancer. Although low dose CT screening was introduced as a promising strategy for the early detection of lung cancer^{11,12}, because it has several limitations, such as the possibility of lead time bias, a high rate of false-positive detection, and the as yet unanswered questions of radiation hazard effects, the debate continues¹³. The early detection of lung cancer may be more urgently needed. Epigenetic biomarkers such as DNA methylation and non-coding RNA may become alternative diagnostic methods to detect lung cancer in its earlier stages. The incidence of aberrant DNA methylation in cancer is higher than that of somatic mutation, and, since global DNA demethylation has been found in most lung cancers, it has the possibility to be a diagnostic marker for lung cancer^{4,9}. Recent studies have suggested that gene-specific promoter DNA

Table 1. Baseline characteristics of patients with lung cancer and matched control subjects

	Patient (n=3)	Control (n=3)
Age, yr	69 (60–74)	66 (57–68)
Smoking pack year, yr	20 (3–50)	20 (0.2–50)
Post bronchodilator FEV ₁	92 (46–106)	76 (76–77)
Histological type		
Adenocarcinoma	1 (33)	-
Squamous cell carcinoma	2 (67)	-
Stage II	2 (67)	-
Stage III	1 (33)	-

Values are presented as median (range) or number (%). FEV₁: forced expiratory volume on one second.

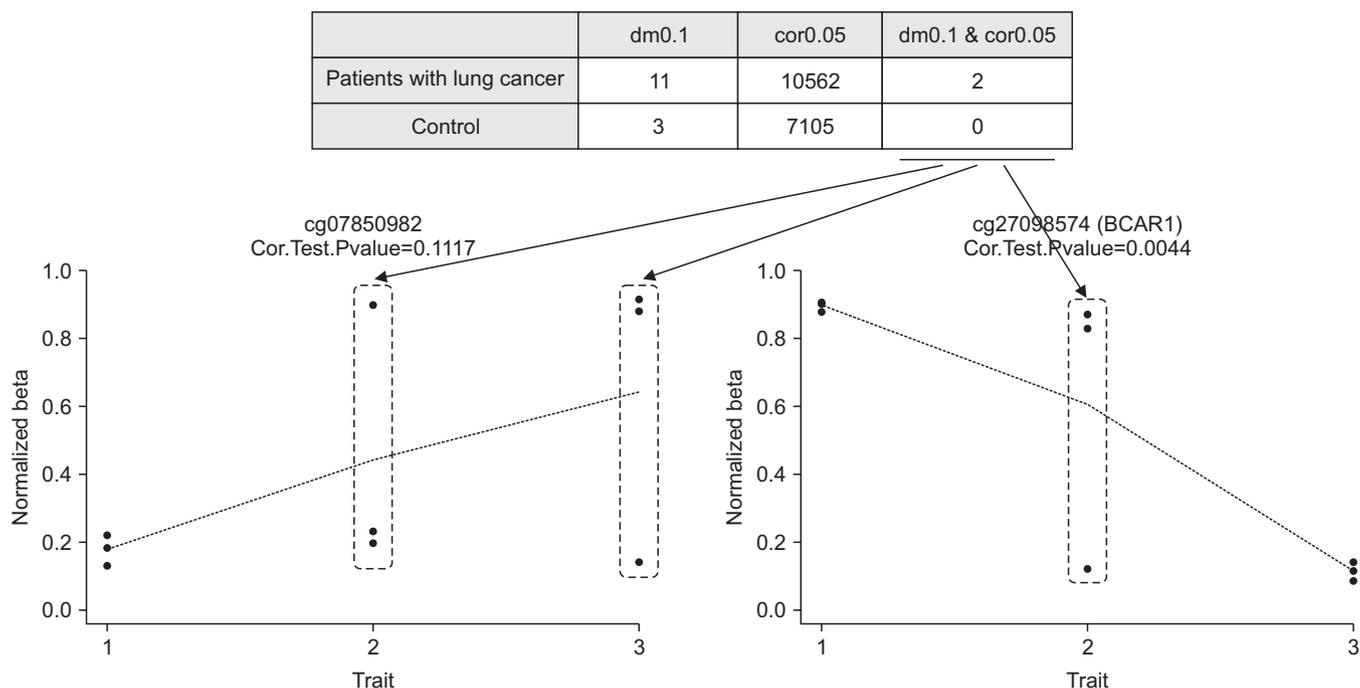


Figure 2. Significant differentially methylated regions with significant β changes and correlation. dm0.1: significant β changes of the delta mean with increased (≥ 0.1) or decreased (≤ -0.1); cor0.05: significant correlation with p-value of > 0.05 .

methylation may be a useful biomarker for the early detection of lung cancer or for the prediction of lung cancer outcomes^{4,9}.

The aberrant promoter methylation changes in NSCLC have been identified in several genes from various samples. *RASSF1A* in sputum¹⁴, *SHOX2* in bronchial aspirates¹⁵, and *CDKN2A* in bronchoalveolar lavage and exhaled breath condensates^{16,17} have been found. In tumor tissues, Brock et al.¹⁸ showed that tumor recurrence was associated with differentially methylated regions of four gene promoters (*p16*, *CDH13*, *APC*, and *RASSF1A*); and Sandoval et al.¹⁹ identified hypermethylations of genes (*HIST1H4F*, *PCDHGB6*, *NPBWR1*, *ALX1*, and *HOXA9*) were associated with shorter survival. These studies have shown the possibility that DNA methylation biomarkers might improve diagnostic accessibility or prognosis prediction beyond standard staging. However, biomarkers for the early detection of lung cancer are in development only and it is expected that they are unlikely to be clinically applied within the next few years⁸. For a biomarker in clinical practice, it may be needed to be detected easily and early in blood. Promotor hypermethylation of six genes (*RASSF1A*, *CDKN2A*, *RARB*, *CDH13*, *FHIT*, and *BLU*) showed approximately 80% concordance between serum and tissue samples in NSCLC²⁰. However, promoter methylation of *CDKN2A*, *DAPK*, *PAX5b*, and *GATA5* in blood was much lower than in tissue samples²¹. These studies show that DNA promoter methylation in blood already has limitations for the early detection of lung cancer, because it assumes that DNA methylation changes in serum or plasma were found

in DNA from apoptotic tumor cells, reflecting an advanced stage. Therefore, using an epigenome-wide associated study to identify new methylation changes of normal cells mediated by cancer cells, as immune cells, may open a new pathway for finding a biomarker for the early detection of lung cancer.

To understand DNA methylation changes induced by cancer, methylation across the whole genome should be measured simultaneously. However, whole-genome bisulphite sequencing may not be the most appropriate method because of its high cost and the technical expertise required to conduct it although it is considered a gold standard tool²². However, a microarray might be used as a popular and informative alternative. The Illumina Infinium BeadChips, the HumanMethylation27K and 450K BeadChip, have been introduced as tools for genome-wide DNA methylation profiling and have been affirmed as having easy and effective characteristics²³. Recently, a more advanced tool containing over 850,000 probes, the Infinium MethylationEPIC (EPIC) BeadChip, with an increased genome coverage of regulatory regions, was introduced. With high reproducibility and reliability, the EPIC array is a significant improvement over the 27K and 450K microarrays²⁴. Therefore, it may allow new valuable insights into the role of DNA methylation changes in cancer. This study, to our knowledge, is the first that uses the EPIC array to show DNA methylation changes in the blood samples of lung cancer patients.

Our study showed two genes associated with DNA methylation changes in lung cancer, *RNF212* and *BCRA1*. Stud-

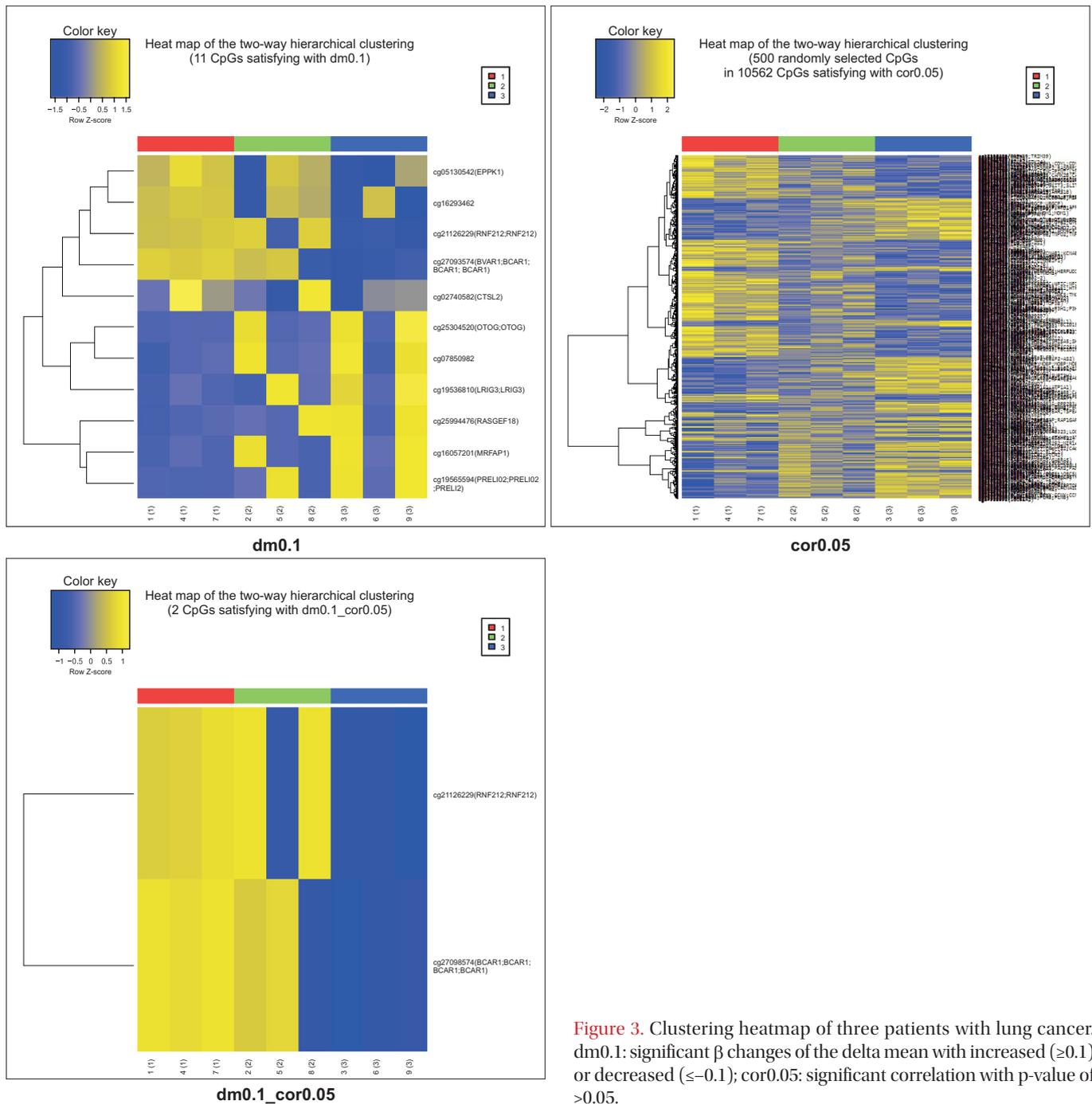


Figure 3. Clustering heatmap of three patients with lung cancer. dm0.1: significant β changes of the delta mean with increased (≥ 0.1) or decreased (≤ -0.1); cor0.05: significant correlation with p-value of >0.05 .

ies have reported that the *RNF212* gene is associated with variation in the genome-side recombination rate and acts to stabilize meiosis-specific recombination factors^{25,26}. There has been no reported association between *RNF212* and lung cancer. Further research is needed. Meanwhile, *BCRA* is a well-known gene whose role in lung cancer has been well studied. The *BRCA* gene acts as a tumor suppressor; its germ line mutation increases the risk of breast and ovarian cancer²⁷. In

early stage NSCLC, the DNA methylation of *BRCA1* and high mRNA expression of *BRCA1* is associated with poorer survival^{28,29}. However, to date, there is low-level evidence to support the use of *BRCA* as a clinically relevant biomarker in NSCLC. More precise and specific mechanism studies are needed.

This genome-wide DNA methylation study showed DNA methylation changes that might be implicated in the development of lung cancer. The DNA methylation changes may be

candidate target regions for the early detection and prevention of lung cancer. Further investigation of these DNA methylation changes and related genes may enhance development of a biomarker and/or clinical applications for lung cancer.

Authors' Contributions

Conceptualization: Kim WJ, Hong Y. Methodology: Moon DH, Kim WJ, Hong Y. Formal analysis: Kwon SO, Kim WJ, Hong Y. Data curation: Kwon SO, Software: Kwon SO. Validation: Kwon SO, Kim WJ, Hong Y. Investigation: Moon DH, Kim WJ, Hong Y. Writing - original draft preparation: Moon DH, Hong Y. Writing - review and editing: all authors. Approval of final manuscript: all authors.

Conflicts 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 for Health and Welfare, Republic of Korea (1631210).

Supplementary Material

Supplementary material can be found in the journal homepage (<http://www.e-trd.org>).

Supplementary Methods. Genomic DNA preparation and DNA methylation profiling.

References

- Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, et al. The global burden of cancer 2013. *JAMA Oncol* 2015;1:505-27.
- Park JY, Jang SH. Epidemiology of lung cancer in Korea: recent trends. *Tuberc Respir Dis* 2016;79:58-69.
- Shin A, Oh CM, Kim BW, Woo H, Won YJ, Lee JS. Lung cancer epidemiology in Korea. *Cancer Res Treat* 2017;49:616-26.
- Balgkouranidou I, Liloglou T, Lianidou ES. Lung cancer epigenetics: emerging biomarkers. *Biomark Med* 2013;7:49-58.
- Devarakonda S, Morgensztern D, Govindan R. Genomic alterations in lung adenocarcinoma. *Lancet Oncol* 2015;16:e342-51.
- Esteller M. Non-coding RNAs in human disease. *Nat Rev Genet* 2011;12:861-74.
- Zhang Y. Recent progress in the epigenetics and chromatin field. *Cell Res* 2011;21:373-4.
- Ansari J, Shackelford RE, El-Osta H. Epigenetics in non-small cell lung cancer: from basics to therapeutics. *Transl Lung Cancer Res* 2016;5:155-71.
- Mehta A, Dobersch S, Romero-Olmedo AJ, Barreto G. Epigenetics in lung cancer diagnosis and therapy. *Cancer Metastasis Rev* 2015;34:229-41.
- Hong Y, Ji W, An S, Han SS, Lee SJ, Kim WJ. Sex differences of COPD phenotypes in nonsmoking patients. *Int J Chron Obstruct Pulmon Dis* 2016;11:1657-62.
- Oken MM, Hocking WG, Kvale PA, Andriole GL, Buys SS, Church TR, et al. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *JAMA* 2011;306:1865-73.
- National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395-409.
- van der Aalst CM, Ten Haaf K, de Koning HJ. Lung cancer screening: latest developments and unanswered questions. *Lancet Respir Med* 2016;4:749-61.
- Hubers AJ, Brinkman P, Boksem RJ, Rhodius RJ, Witte BI, Zwinderman AH, et al. Combined sputum hypermethylation and eNose analysis for lung cancer diagnosis. *J Clin Pathol* 2014;67:707-11.
- Kneip C, Schmidt B, Seegebarth A, Weickmann S, Fleischhacker M, Liebenberg V, et al. *SHOX2* DNA methylation is a biomarker for the diagnosis of lung cancer in plasma. *J Thorac Oncol* 2011;6:1632-8.
- Ahrendt SA, Chow JT, Xu LH, Yang SC, Eisenberger CF, Esteller M, et al. Molecular detection of tumor cells in bronchoalveolar lavage fluid from patients with early stage lung cancer. *J Natl Cancer Inst* 1999;91:332-9.
- Xiao P, Chen JR, Zhou F, Lu CX, Yang Q, Tao GH, et al. Methylation of P16 in exhaled breath condensate for diagnosis of non-small cell lung cancer. *Lung Cancer* 2014;83:56-60.
- Brock MV, Hooker CM, Ota-Machida E, Han Y, Guo M, Ames S, et al. DNA methylation markers and early recurrence in stage I lung cancer. *N Engl J Med* 2008;358:1118-28.
- Sandoval J, Mendez-Gonzalez J, Nadal E, Chen G, Carmona FJ, Sayols S, et al. A prognostic DNA methylation signature for stage I non-small-cell lung cancer. *J Clin Oncol* 2013;31:4140-7.
- Hsu HS, Chen TP, Hung CH, Wen CK, Lin RK, Lee HC, et al. Characterization of a multiple epigenetic marker panel for lung cancer detection and risk assessment in plasma. *Cancer* 2007;110:2019-26.
- Belinsky SA, Grimes MJ, Casas E, Stidley CA, Franklin WA, Bocklage TJ, et al. Predicting gene promoter methylation in non-small-cell lung cancer by evaluating sputum and serum. *Br J Cancer* 2007;96:1278-83.

22. Stirzaker C, Taberlay PC, Statham AL, Clark SJ. Mining cancer methylomes: prospects and challenges. *Trends Genet* 2014;30:75-84.
23. Bibikova M, Le J, Barnes B, Saedinia-Melnyk S, Zhou L, Shen R, et al. Genome-wide DNA methylation profiling using Infinium(R) assay. *Epigenomics* 2009;1:177-200.
24. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina Methylation-EPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol* 2016;17:208.
25. Qiao H, Prasada Rao HB, Yang Y, Fong JH, Cloutier JM, Deacon DC, et al. Antagonistic roles of ubiquitin ligase HEI10 and SUMO ligase RNF212 regulate meiotic recombination. *Nat Genet* 2014;46:194-9.
26. Sandor C, Li W, Coppieters W, Druet T, Charlier C, Georges M. Genetic variants in *REC8*, *RNF212*, and *PRDM9* influence male recombination in cattle. *PLoS Genet* 2012;8:e1002854.
27. Foulkes WD. Inherited susceptibility to common cancers. *N Engl J Med* 2008;359:2143-53.
28. Marsit CJ, Liu M, Nelson HH, Posner M, Suzuki M, Kelsey KT. Inactivation of the Fanconi anemia/BRCA pathway in lung and oral cancers: implications for treatment and survival. *Oncogene* 2004;23:1000-4.
29. Rosell R, Skrzypski M, Jassem E, Taron M, Bartolucci R, Sanchez JJ, et al. BRCA1: a novel prognostic factor in resected non-small-cell lung cancer. *PLoS One* 2007;2:e1129.

Supplementary Methods

Genomic DNA preparation and DNA methylation profiling

1. Genomic DNA quantitation

DNA samples were checked the quality using NanoDrop ND-1000 UV-Vis spectrophotometer (Wilmington, DE, USA). Then samples were electrophoresed on agarose gels and samples with intact genomic DNA showing no smearing on agarose gel electrophoresis were selected for experiment. Intact genomic DNA was diluted to 50 ng/ μ L concentration based on Quant-iT Picogreen (Invitrogen, Carlsbad, CA, USA) quantitation. Concentrations were adjusted based on these results. All prepared samples were bisulfite-converted according to the Zymo EZ DNA methylation kit protocols (Zymo Research, Orange, CA, USA).

2. Bisulfite conversion (Zymo EZ DNA methylation kit)

Five hundred nanograms of input gDNA was required for the bisulfite conversion. Add conversion reagent and incubate in a thermocycler to denature. CT converted DNA was washing and desulphonating with desulphonation buffer. After desulphonation, the DNA was washing again and eluting with 12 μ L elution buffer.

3. Sample amplification and hybridization for BeadChips (MethylationEPIC DNA Analysis BeadChip Kit; Illumina, San Diego, CA, USA)

The whole-genome amplification process requires 200 ng of input bisulfite-converted DNA, MA1 and creates a sufficient quantity of DNA (1,000 \times amplification) to be used on a single BeadChip in the Infinium methylation assay (Illumina RPM and MSM). After amplification, the product is fragmented using a proprietary reagent (FMS), precipitated with 2-propanol (plus precipitating reagent; PM1), and resuspended in formamide-containing hybridization buffer (RA1). The DNA samples are denatured at 95°C for 20 minutes, then placed in a humidified container for a minimum of 16 hours at 48°C allowing CpG loci to hybridize to the 50mer capture probes.

4. Allele specific single-base extension and staining on BeadChips

Following hybridization, the BeadChip/Te-Flow chamber assembly was placed on the temperature-controlled Tecan Flowthrough Chamber Rack, and all subsequent washing, extension, and staining were performed by addition of reagents to the Te-Flow chamber.

For the allele specific single-base extension assay, primers were extended with a polymerase and labeled nucleotide mix (TEM), and stained with repeated application of STM (staining reagent) and ATM (anti-staining reagent). After staining was complete, the slides were washed with low salt wash buffer

(PB1), immediately coated with XC4, and then imaged on the Illumina iScan Reader.

5. Imaging the BeadChip and data analysis

The Illumina iScan Reader is a two-color (532 nm/658 nm) confocal fluorescent scanner with 0.53- μ m pixel resolution. The scanner excites the fluorophors generated during signal amplification/staining of the allele-specific (one color) extension products on the BeadChips. The image intensities are extracted using Illumina's iScan Control software.

6. Raw data preparation and Statistical analysis

The quality of hybridization and overall chip performance were monitored by visual inspection of both internal quality control checks and the raw scanned data. Raw data were extracted as beta values for each CpG for each sample using R watermelon package. Beta values were calculated by subtracting background using negative controls on the array and taking the ratio of the methylated signal intensity against the sum of both methylated and unmethylated signals. A beta value of 1–1.0 was reported significant percent methylation, from 0% to 100%, respectively, for each CpG site.

Array CpG probes that have detection p-value ≥ 0.05 (similar to signal to noise) in over 25% samples were filtered out. (We applied a filtering criterion for data analysis; good signal value was required to obtain a detection p-value < 0.05). And then filtered data was background correction & dye bias equalization by R methylumi & lumi package. To reduce Infinium I and Infinium II assay bias, corrected signal value was normalized by BMIQ (Beta Mixture Quantile) method.

Differentially expressed methylation list was determined using $|\Delta \text{mean}| \geq 0.2$ (the difference of methylation signal, avg beta of Case-avg beta of Control) and p-value < 0.05 of independent t-test in which the null hypothesis was that no difference exists among two groups.

All data analysis and visualization of differentially expressed genes was conducted using R 3.0.2 (<http://www.r-project.org>).

References

1. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A* 2002;99:3740-5.
2. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddleloh JA, et al. Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res* 2008;18:780-90.
3. Bjornsson HT, Brown LJ, Fallin MD, Rongione MA, Bibikova M, Wickham E, et al. Epigenetic specificity of loss of imprinting of the IGF2 gene in Wilms tumors. *J Natl Cancer Inst* 2007;99:1270-3.
4. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, et al.

High-throughput DNA methylation profiling using universal bead arrays. *Genome Res* 2006;16:383-93.

5. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA,

Bibikova M, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* 2011;6:692-702.