DNMT3B polymorphisms and risk of primary lung cancer

Su Jeong Lee1,*, Hyo-Sung Jeon1,*, Jin-Sung Jang2, Sun Ha Park3, Ga Young Lee4, Byung-Heon Lee2, Chang Ho Kim3, Young Mo Kang3, Won Kee Lee4, Sin Kam2,5, Rang Woon Park2, In-San Kim2, Young Lae Cho2, Tae Hoon Jung6 and Jae Yong Park1,3,7

1Cancer Research Institute, Kyungpook National University Hospital, Samduk 2Ga 50, Daegu, 700-412, Korea, 2Department of Biochemistry and 3Department of Internal Medicine, School of Medicine, Kyungpook National University, Dong In 2Ga 101, Daegu, 700-422, Korea, 4Health Promotion Research Center, Kyungpook National University, Dong In 2Ga 101, Daegu, 700-422, Korea, 5Department of Preventive Medicine, School of Medicine, Kyungpook National University, Dong In 2Ga 101, Daegu, 700-422, Korea and 6Department of Obstetrics and Gynecology, School of Medicine, Kyungpook National University, Dong In 2Ga 101, Daegu, 700-422, Korea

These authors contributed equally to this paper.

DNA-methyltransferase-3B (DNMT3B) plays an important role in the generation of aberrant methylation in carcinogenesis. Polymorphisms and haplotypes of the DNMT3B gene may influence DNMT3B activity on DNA methylation, thereby modulating the susceptibility to lung cancer. To test this hypothesis, we investigated the association of the −283T>C (from exon 1A transcription start site) and −579G>T (from exon 1B transcription start site) polymorphisms in DNMT3B promoter, and their haplotypes with the risk of lung cancer in a Korean population. The DNMT3B genotype was determined in 432 lung cancer patients and 432 healthy controls that were frequency-matched for age and sex. Individuals with at least one −283T allele were at a significantly decreased risk of adenocarcinoma (AC) and small cell carcinoma (SM) [adjusted odds ratio (OR) = 0.48, 95% confidence interval (CI) = 0.28-0.82, P = 0.007; and adjusted OR = 0.47, 95% CI = 0.24-0.93, P = 0.03, respectively] compared with those harboring a −283CC genotype. Individuals with at least one −579G allele were also at a significantly decreased risk of AC and SM (adjusted OR = 0.47, 95% CI = 0.28-0.81, P = 0.006; and adjusted OR = 0.51, 95% CI = 0.26-0.99, P = 0.048, respectively) compared with those having a −579TT genotype. The −283T allele was linked with the −579G allele, and haplotype −283T/−579G was associated with a significantly decreased risk of AC (adjusted OR = 0.48, 95% CI = 0.29-0.81, P = 0.006) as compared with haplotype −283C/−579T. In a promoter assay, carriage of the −283T allele showed a significantly lower promoter activity (~50%) compared with the −283C allele (P < 0.001), but the −579G>T polymorphism did not have an effect on the DNMT3B promoter activity. These results suggest that the DNMT3B −283T>C polymorphism influences DNMT3B expression, thus contributing to the genetic susceptibility to lung cancer.

Introduction

Although cigarette smoking is the major cause of lung cancer, only a fraction of smokers develop lung cancer during their lifetime, suggesting that genetic and epigenetic factors are of importance in determining individuals’ susceptibility to lung cancer (1,2).

DNA methylation is a major epigenetic mechanism that regulates chromosomal stability and gene expression (3,4). DNA methylation is mediated by DNA methyltransferases (DNMTs). In human, three catalytically active DNMTs (DNMT1, DNMT3A and DNMT3B) have been identified (5–7). Although DNMTs act cooperatively for the establishment and maintenance of genomic methylation patterns (8,9), DNMT1 is thought to be primarily responsible for maintaining pre-existing methylation patterns after DNA replication because of its preference for hemimethylated DNA substrates and targeting to replication foci (10). DNMT3A and DNMT3B have equal preference for hemimethylated and unmethylated DNA substrates, and therefore they are believed to be principally required for de novo methylation (11).

Many human cancers including lung cancer have both global hypomethylation and regional hypermethylation of CpG islands (12–15). Such aberrant DNA methylation may contribute to carcinogenesis in several ways. Hypomethylation may lead to chromosomal instability, reactivation of transposable elements, and loss of imprinting (13,16). Methylation of CpG sequences may facilitate C-to-T transition mutations in tumor suppressor genes and/or oncogenes through deamination of 5-methylcytosine to thymine (17). Methylated CpG sequences may also increase susceptibility to attack by some environmental carcinogens like benzo[a]pyrene diol epoxide (18,19). Finally, de novo hypermethylation of promoter CpG islands may lead to silencing of tumor suppressor genes and DNA repair genes (4,13,16).

In various human cancers, DNMT genes are up-regulated (20–22). Several studies have linked the aberrant de novo hypermethylation of promoter CpG islands to the over-expression of DNMTs (23,24). Genetic disruption of both DNMT1 and DNMT3B in human cancer cells has been shown to result in global and gene-specific demethylation, and abrogation of silencing of tumor suppressor genes (8,25). Taken together, these results suggest that altered activities of DNMTs contribute to the generation of aberrant methylation in cancer.

Single nucleotide polymorphisms (SNPs) are the most common form of human genetic variation, and may contribute to

Abbreviations: AC, adenocarcinoma; CI, confidence interval; DNMT, DNA methyltransferase; OR, odds ratio; SCC, squamous cell carcinoma; SM, small cell carcinoma; SNP, single nucleotide polymorphism; SP1, simian virus-40 protein 1.

*These authors contributed equally to this paper.
an individuals’ susceptibility to cancer. Many studies have demonstrated that some variants affect either the expression or activities of enzymes and therefore are associated with cancer risk (1,26). Recently, several candidate SNPs in the DNMT3B gene have been deposited in public databases (http://www.ncbi.nlm.nih.gov/SNP). Although the functional effects of these polymorphisms have not been elucidated, we hypothesized that some of these variants, particularly their haplotypes, may influence DNMT3B activity on DNA methylation, thereby modulating the susceptibility to lung cancer. To test this hypothesis, a case control study was conducted to evaluate the association between DNMT3B genotypes/haplotypes and the risk of lung cancer. Among the candidate SNPs in DNMT3B, we have focused on amino acid substitution variants (Ala204Ser, 1541945G>T in exon 7, and Glu368Lys, 1547344G>A in exon 10; GenBank accession no. NT_028392), and variants in promoter regions –283T>C (from exon 1A transcription start site) and –579G>T (from exon 1B transcription start site); GenBank accession no. NT_028392) as these most probably affect the gene function. In addition, DNMT3B 46359C>T polymorphism (GenBank accession no. AL103507) was also evaluated since this polymorphism was associated with the risk for lung cancer in a previous study (27). In the present study, we evaluated the association of –283T>C and –579G>T polymorphisms and their haplotypes with lung cancer because Ala204Ser, Glu368Lys and 46359C>T polymorphisms were not detected in a preliminary study consisting of 50 lung cancer cases and 50 healthy controls.

Materials and methods

Study population

This case-control study included 432 lung cancer patients and 432 healthy controls. The method used for subject enrollment was the same as for our previous studies (28,29). Eligible cases included all patients (447 patients) newly diagnosed with primary lung cancer between January 2001 and February 2002 in Kyungpook National University Hospital, Daegu, Korea. There were no age, sex, histological or stage restrictions, but 15 patients with a prior history of cancers were excluded. Authors and a trained coordinator explained the objectives and importance of this study, and we could include all patients with histologically confirmed primary lung cancer. All patients signed an informed consent for blood sample collection. The cases included all patients with histologically confirmed primary lung cancer. All patients had been smoking. For smoking status, a person who had smoked at least once in his or her lifetime was regarded as a smoker. A former smoker was defined as one who had stopped smoking at least 1 year before diagnosis in the case of patients and 1 year before the date signed on an informed consent for blood sample collection in the case of controls. Cumulative cigarette dose (pack-years) was calculated by the following formula: pack-years = (packs per day) × (years smoked).

DNMT3B genotyping

Genomic DNA was extracted from peripheral blood lymphocytes by proteinase K digestion and phenol-chloroform extraction. The DNMT3B –283T>C and –579G>T polymorphisms were determined by a PCR–RFLP assay. PCR reactions were performed in a total of 20 μl containing 200 ng genomic DNA, 25 pM each primer, 0.2 mM dNTPs, 75 mM Tris–HCl (pH 9.0), 15 mM sodium sulfate, 0.1 μg/ml BSA, 2.5 mM MgCl2, and 1 U of Taq polymerase (Neurotics, Seoul, Korea). PCR cycle conditions consisted of an initial denaturation step at 95°C for 5 min followed by 35 cycles of 20 s at 94°C; 20 s at 54°C for –283T>C and 20 s at 56°C for –579G>T; 20 s at 72°C; and a final elongation at 72°C for 10 min.

PCR primers were designed based on the GenBank reference sequence (accession no. NT_028392). The –283T>C polymorphism was determined using the following primers: 5′-GAAAAGCCCCAGAACGG(mutated A → CG-3′ (bases 1 516 656–1 516 675 of DNMT3B) and 5′-GGCCG-GGACGAGGAGAATT-3′ (bases 1 516 839–1 516 820 of DNMT3B). The 184 bp PCR product was digested with BanI (New England BioLabs, Beverly, MA), and resolved on 6% acrylamide gel and stained with ethidium bromide for visualization under UV light. The wild-type T allele has only one band, while the polymorphic C allele has two bands (167 and 17 bp). The –579G>T polymorphism was determined using the following primers: 5′-GAGGCTCTATTGCTGCGAGG-3′ (bases 1 533 753–1 533 772 of DNMT3B); 5′-GGGACCTCACCTCTTACAGAAA-3′ (bases 1 533 977–1 533 958 of DNMT3B). The 225 bp PCR product was digested with AciI (New England BioLabs) and resolved on 6% acrylamide gel and stained with ethidium bromide for visualization under UV light. The wild-type G allele has only one band, while the polymorphic C allele has two bands (132 and 93 bp). For quality control, genotyping analysis was performed blind with respect to case/control status and repeated twice for all subjects. The results of genotyping were 100% concordant. To confirm the genotyping results, selected PCR-amplified DNA samples (n = 2, respectively, for each genotype) were examined by DNA sequencing, and the results were also 100% concordant.

Transcription activity analysis

The fragments of the DNMT3B exon 1A promoter region (from –1140 to +53, transcription start site of exon 1A counted as +1) and exon 1B promoter region (from –863 to +196, transcription start site of exon 1B counted as +1) were synthesized, respectively, by PCR using genomic DNA from donors carrying either the wild-type or polymorphic allele of each DNMT3B promoter region. The PCR primers used for the exon 1A and exon 1B promoter regions were 5′-TCCGGTGATGCTGTAGAAC-3′ (forward) and 5′-GGACGGACTGCTGCCAGGAA-3′ (reverse) and 5′-AGTAAACACTGCGAGCTAGC-3′ (forward) and 5′-TCTCAAGCTCAGACTGCTT-3′ (reverse), respectively. The PCR products were inserted upstream of the luciferase gene in the pGL3-basic plasmid (Promega, Madison, WI), and the correct sequence of all clones was verified by DNA sequencing. Promoter activity was measured using the Luciferase Reporter Assay System (Promega). Chinese hamster ovary cells were grown in minimal essential medium supplemented with 10% fetal bovine serum, 10% heat-inactivated fetal calf serum and 1% penicillin/streptomycin. The cells were plated in a 24-well plate at 105 cells/well 24 h before transfection, and cell lysates were prepared according to Promega’s instruction manual. Luciferase activity was measured using a Lumat LB953 luminometer (EG&G Berthold, Bad Wildbad, Germany), and normalized using the activity of Renilla luciferase. The experiment was performed four times in triplicates, and the results were reported as mean ± SD.

Statistical analysis

Cases and controls were compared using Student’s t-test for continuous variables and the χ² test for categorical variables. Hardy–Weinberg equilibrium was tested with a goodness-of-fit χ² test with one degree of freedom to compare the observed genotype frequencies with the expected frequencies among the subjects. We examined widely used measure of linkage disequilibrium (LD) between pairs of biallelic loci, Lewontin’s D’ (31). Haplotypes and their frequencies were estimated based on Bayesian algorithm using PHASE program (32), which is available at http://www.stat.washington.edu/stephens/phase.html. Unconditional logistic regression analysis was used to calculate odds ratios (ORs) and 95% confidence intervals (CIs), with adjustment for possible confounders (sex as a nominal variable; and age and pack-years of smoking, as continuous variables). To assess the potential interaction between genotype and smoking (smoking status or pack-years of smoking), we first included the interaction term in the logistic model. The interaction between genotype and smoking (smoking status or pack-years of smoking) was
not statistically significant, and this was removed from the model. The difference of luciferase activities between the wild-type allele and variant allele was assessed by Student’s t-test. All analyses were performed using Statistical Analysis Software for Windows, version 8.12 (SAS institute, Gary, NC).

**Results**

The demographics of the cases and controls enrolled in this study are shown in Table I. There were no significant differences in the mean age and sex distribution between cases and controls, suggesting that the matching based on these two variables was adequate. The cases had a higher prevalence of current smokers than the controls (P < 0.001). The number of pack-years in smokers was significantly higher in cases than in controls (39.9 ± 17.9 versus 34.4 ± 17.6 pack-years; P < 0.001). These differences were controlled for later by multivariate analyses.

The distributions of DNMT3B −283T>C and −579G>T genotypes among controls and cases are shown in Table II. The genotype distributions of both polymorphisms among the controls were in Hardy–Weinberg equilibrium. No significant deviation was observed for the genotype distributions of both polymorphisms between overall lung cancer cases and controls. When the cases were categorized by tumor histology, however, the distributions of −283T>C and −579G>T genotypes in the AC group (TT 0%, TC 14.2%, CC 85.8% and GG 0%, GT 13.7%, TT 86.3%) were significantly different from those among the controls (TT 1.6%, TC 24.5%, CC 73.8% and GG 1.6%, GT 25.0%, TT 73.4%; P = 0.009 and P = 0.007, respectively). The distributions of genotypes of both polymorphisms in the other histological subtypes were not significantly different from those among the controls.

Table III shows the lung cancer risk related to the DNMT3B −283T>C and −579G>T genotypes, respectively. ORs and their 95% CIs were calculated using the more common homozygous variant genotype as the reference group (−283 CC and

### Table I. Characteristics of the study population

<table>
<thead>
<tr>
<th>Variables</th>
<th>Case (n = 432)</th>
<th>Control (n = 432)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>61.6 ± 9.0</td>
<td>60.9 ± 9.3</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>352 (81.5)</td>
<td>352 (81.5)</td>
</tr>
<tr>
<td>Female</td>
<td>80 (18.5)</td>
<td>80 (18.5)</td>
</tr>
<tr>
<td>Smoking statusb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>317 (73.4)</td>
<td>229 (53.0)</td>
</tr>
<tr>
<td>Former</td>
<td>39 (9.0)</td>
<td>98 (22.7)</td>
</tr>
<tr>
<td>Never</td>
<td>76 (17.6)</td>
<td>105 (24.3)</td>
</tr>
<tr>
<td>Pack-yearsb</td>
<td>39.9 ± 17.9</td>
<td>34.4 ± 17.6</td>
</tr>
</tbody>
</table>

*aNumbers in parentheses, percentage.

*bP = 0.001.

In current and former smokers, P < 0.001.

### Table II. DNMT3B genotypes and allele frequency among controls and cases

<table>
<thead>
<tr>
<th>−283T&gt;C genotype</th>
<th>Control</th>
<th>TT</th>
<th>TC</th>
<th>CC</th>
<th>T allele frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>106</td>
<td>319</td>
<td>73.8%</td>
</tr>
<tr>
<td>Case</td>
<td>5 (1.1)</td>
<td>85</td>
<td>342</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td>Squamous</td>
<td>3 (1.4)</td>
<td>55</td>
<td>152</td>
<td>14.5</td>
<td></td>
</tr>
<tr>
<td>Adeno</td>
<td>0 (0.0)</td>
<td>20</td>
<td>121</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>Large</td>
<td>0 (0.0)</td>
<td>1</td>
<td>7</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>2 (2.7)</td>
<td>9</td>
<td>62</td>
<td>8.9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>−579G&gt;T genotype</th>
<th>Control</th>
<th>TT</th>
<th>GC</th>
<th>GT</th>
<th>T allele frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>108</td>
<td>317</td>
<td>73.4%</td>
</tr>
<tr>
<td>Case</td>
<td>4 (0.9)</td>
<td>87</td>
<td>341</td>
<td>11.0</td>
<td></td>
</tr>
<tr>
<td>Squamous</td>
<td>2 (0.9)</td>
<td>56</td>
<td>152</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>Adeno</td>
<td>0 (0.0)</td>
<td>18</td>
<td>113</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>Large</td>
<td>0 (0.0)</td>
<td>1</td>
<td>7</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>2 (2.7)</td>
<td>10</td>
<td>61</td>
<td>9.6</td>
<td></td>
</tr>
</tbody>
</table>

*aNumbers in parentheses, percentage.

*bP = 0.009.

*cP = 0.007, AC versus control.

### Table III. Crude and adjusted† ORs for lung cancer associated with DNMT3B genotypes

<table>
<thead>
<tr>
<th>Genotype</th>
<th>All cases</th>
<th>Squamous cell carcinoma</th>
<th>AC</th>
<th>SM</th>
</tr>
</thead>
<tbody>
<tr>
<td>−283T&gt;C genotype</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC (reference)</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>TC + TT crude</td>
<td>0.74 (0.54-1.02)</td>
<td>1.08 (0.74-1.56)</td>
<td>0.47 (0.28-0.78)</td>
<td>0.50 (0.25-0.99)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.75 (0.54-1.03)</td>
<td>1.04 (0.71-1.53)</td>
<td>0.48 (0.28-0.82)</td>
<td>0.47 (0.24-0.93)</td>
</tr>
<tr>
<td>−579G&gt;T genotype</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT (reference)</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>GT + GG crude</td>
<td>0.74 (0.54-1.01)</td>
<td>1.05 (0.73-1.52)</td>
<td>0.46 (0.27-0.77)</td>
<td>0.54 (0.28-1.04)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>0.75 (0.54-1.03)</td>
<td>1.04 (0.71-1.53)</td>
<td>0.47 (0.28-0.81)</td>
<td>0.51 (0.26-0.99)</td>
</tr>
</tbody>
</table>

*ORs (95% CIs) and corresponding P-values were calculated by logistic regression, adjusted for age, sex and pack-years of smoking.

†P = 0.07.

‡P = 0.003.

§P = 0.04.

¶P = 0.08.

‖P = 0.007.

*P = 0.03.

**P = 0.06.

***P = 0.006.

****P = 0.048.
−579 TT genotypes, respectively). Because of low prevalence of homozygous wild-type genotype, we combined this genotype with heterozygous genotype into one group and compared it with the reference group (a dominant model for the wild-type allele). As compared with the reference group, the combined −283 TT and TC genotype, and the combined −579 GG and GT genotype were associated with non-significant decreased risk of overall lung cancer, respectively (adjusted OR = 0.75, 95% CI = 0.54−1.03, P = 0.08; and adjusted OR = 0.75, 95% CI = 0.54−1.03, P = 0.08, respectively). When analyses were stratified by tumor histology, the combined −283 TT and TC genotype, and the combined −579 GG and GT genotype were associated with a significantly decreased risk of AC, respectively (adjusted OR = 0.48, 95% CI = 0.28−0.82, P = 0.007; and adjusted OR = 0.47, 95% CI = 0.28−0.81, P = 0.006, respectively), and these two combined genotypes were also associated with a significantly decreased risk of SM, respectively (adjusted OR = 0.47, 95% CI = 0.24−0.93, P = 0.03; and adjusted OR = 0.51, 95% CI = 0.26−0.99, P = 0.048, respectively). However, there was no significant association between these two polymorphisms and the risk of SCC.

DNMT3B −283 T>C and −579 G>T polymorphisms were in LD in the cases and controls (D' = 0.96 and 0.95, respectively). Two common haplotypes (−283T/−579G and −283C/−579T) accounted for 98% of the chromosome of the present study population. Two haplotypes (−283T/−579T and −283C/−579G) that had a frequency of <1% was excluded [controls, 5 (0.6%) and 7 (0.8%); and cases, 3 (0.4%) and 3 (0.4%), respectively] from further analysis to avoid possible errors in either the genotyping or the estimation process. The distributions of haplotypes −283T/−579G and −283C/−579T among the cases and controls are shown in Table IV. The distribution of the haplotypes among overall lung cancer cases was not significantly different from that among the controls (P = 0.08). When the cases were categorized by tumor histology, the haplotypic distribution in the AC group was significantly different from that among the controls (P = 0.002). To examine the effect of haplotypes on the risk of lung cancer, the −283C/−579T haplotype was used as the reference group. As compared with the reference haplotype, the −283T/−579G haplotype was associated with non-significant decreased risk of overall lung cancer (adjusted OR = 0.77, 95% CI = 0.57−1.04, P = 0.09). When analyses were stratified by tumor histology, the −283C/−579T haplotype was associated with a significantly decreased risk of AC (adjusted OR = 0.48, 95% CI = 0.29−0.81, P = 0.006). We investigated the effects of the −283 T>C and −579 G>T polymorphisms on the promoter activity of DNMT3B by luciferase assay. The promoter activity of the −283T allele was significantly lower (~50%) compared with the −283C allele (P < 0.001, Figure 1), but the −579G>T polymorphism did not affect the promoter activity (data not shown).

Discussion

We investigated the influence of DNMT3B polymorphisms on the risk of lung cancer in a hospital-based case-control study. The −283 T>C and −579 G>T polymorphisms in the DNMT3B promoter, and their haplotypes were significantly associated with the risk of AC of the lung. This finding suggests that the DNMT3B −283 T>C and −579 G>T polymorphisms could be used as markers of genetic susceptibility to AC. Of three major histological types of lung cancer, the proportion of AC is increasing world-wide. Thus, identification of genetic factors responsible for susceptibility to AC is indispensable to establishing novel and efficient ways of preventing the disease.

The DNMT3B gene has two transcriptional start sites, which exist in different exons (exon 1A and 1B) and the expression is regulated by different promoters. One promoter is nested within a CpG-rich area, whereas the other promoter is found in CpG poor (33). The DNMT3B −283 T>C (−283 bp from exon 1A transcription start site) and −579 G>T (−579 bp from exon 1B transcription site) polymorphisms are located in the CpG-rich and CpG-poor promoters, respectively. In the present study, we found that these two polymorphisms are in linkage disequilibrium.

In the present study, carriers with −283 T and −579 G alleles were at decreased risk of lung cancer as compared with individuals having −283 C and −579 T alleles. To determine whether the association between the DNMT3B −283 T>C and −579 G>T polymorphisms, and the risk of lung cancer.
is due to difference in the transcriptional activity of the DNMT promoter, we compared the promoter activity of the wild-type allele or polymorphic allele of these two polymorphisms by luciferase assay. Since the $-283T>C$ and $-579G>T$ polymorphisms were in LD, the functional effects of both polymorphisms might be haplotype-dependent. Therefore, it is reasonable to compare the promoter activity of haplotypes (haplotype $-283T/-579G$ versus haplotype $-283C/-579T$). However, it is difficult to synthesize the construct containing both $-283T>C$ and $-579G>T$ polymorphisms by PCR since the $-579G>T$ polymorphism is located at 17 171 bp from the $-283T>C$ polymorphism. Therefore, we investigated the effect of each polymorphism on the promoter activity, respectively. The in vitro promoter assay revealed that the $-283T$ allele had a significantly lower transcriptional activity than that of the $-283C$ allele, and the $-579G>T$ polymorphism did not effect on the transcriptional activity of the DNMT promoter. These results suggest that the genetic effects of DNMT3B polymorphisms on lung cancer risk are mainly attributed to the $-283T>C$ polymorphism, and also suggest that the $-283 T$ to $C$ change up-regulates DNMT3B expression, resulting in a predisposition towards aberrant de novo methylation of CpG islands in tumor suppressor genes and DNA repair genes. This explanation is comparable with the suggestion proposed by previous studies (34,35) that individuals with increased DNMT levels may have a greater susceptibility to lung cancer. The mechanism by which the DNMT3B $-283C$ allele led to higher promoter activity is currently unknown. Analysis of potential transcription factor binding sites by Alibaba2 program (36) showed that the $-283T$ to $C$ change leads to the creation of a simian virus-40 protein 1 (Sp1) binding site and eliminates a c-Jun binding site. Sp1 is a transcription factor enhancing expression of several viral and cellular genes by binding to specific enhancer sites. Yanagisawa et al. (33) reported that the deletion of the Sp1 binding site in the DNMT3B exon 1A promoter region results in a reduction of the promoter activity. Therefore, it is possible that the creation of the new Sp1 binding site, due to the $-283C$ allele, may lead to enhanced promoter activity. However, this hypothesis needs to be verified in future studies.

In the current study, the DNMT3B polymorphisms were significantly associated with the risk of AC, but not SCC. Although the reason for the observed histology-specific difference in the risk conferred by the DNMT3B polymorphisms remains to be elucidated, this difference may be attributable to the differences in the pathways of carcinogenesis among histological types of lung cancer. Certain genotypes could therefore confer a greater susceptibility to a particular histological type of lung cancer (37-39). Various lines of evidence suggest that DNMT activity and aberrant promoter methylation can be modulated differentially by specific carcinogens and cell types exposed to the carcinogen (40,41). Therefore, the observed histology-specific difference in the susceptibility conferred by the DNMT3B polymorphisms may result from a different DNMT3B activity by histological types of lung cancer. This explanation is comparable with the previous report (42) that the aberrant promoter methylation rates of several genes were significantly higher in AC than in SCC. Another possible explanation is that this finding may be due to the different effects of other genetic variants on a genetic predisposition toward aberrant DNA methylation by histological types of lung cancer. For example, the DNMT3A is also required for de novo DNA methylation and can compensate for lack of DNMT3B activity (11). Therefore, DNMT3A polymorphisms could effect on the occurrence of de novo promoter methylation. It should be also considered that polymorphic variants of genes involved in carcinogen metabolism and repair of DNA damage may affect the occurrence of de novo promoter methylation (43).

Recently, Shen et al. (27) reported an association between DNMT3B 46359C>T polymorphism and lung cancer risk in non-Hispanic whites. They found that the heterozygous CT genotype was associated with a significantly increased risk for lung cancer as compared with the homozygous CC genotype. These authors reported that the risk estimate related to the variant genotype was higher for SCC than for AC. In contrast to their finding, DNMT3B $-283T>C$ and $-579G>T$ polymorphisms were associated with the risk of AC, but not with SCC in the current study. The reason for the difference in lung cancer histological type associated with DNMT3B polymorphism between two studies is not clear but may be because the polymorphic site they analyzed was different from those analyzed in the current study. Given that the DNMT3B 46359C>T polymorphism was not detected in our study population it is also possible that the different results between two studies may be attributable to the genetic differences in different ethnic populations. Geographic differences in carcinogen exposure or molecular pathogenesis should be also considered. In the previous study (27), they reported that the DNMT3B 46359C>T polymorphism is located in the DNMT3B promoter and is linked with DNMT3B promoter activity. Based on the GenBank reference sequence (accession no. AL035071), however, this polymorphism is located at intron 5 (IVS6-49C>T). It is unclear how this intronic polymorphism effects on the promoter activity of the DNMT3B gene.

Genetic susceptibility to lung cancer may depend on the level of carcinogen exposure (44,45). However, in the current study, we did not find any evidence for a gene-smoking (smoking status or pack-years of smoking) interaction. The DNMT3B polymorphism may have an influence on disease progression. In the present study, however, no significant difference was observed in the genotype distributions of both polymorphisms according to the stage of lung cancer (data not shown). One must consider a number of limitations of this study. Since this study was a hospital-based case-control study, there might be some selection bias. Given that most lung cancer patients are treated at university hospital in Korea, the demographics and clinical characteristics of the cancer patients in the current study were compatible with those of a nationwide lung cancer survey (30). Furthermore, as all the lung cancer patients diagnosed at a national university hospital were included in this study, it is reasonable to assume that the case group represents the lung cancer cases in our community. All cases and controls were ethnically Koreans and residents of Daegu City or the surrounding regions. Another selection bias may derive from controls that did not participate in this study. However, because the age and sex distribution of non-participating controls were similar to those of the participating controls in the current study, a self-selection bias is unlikely. By matching on age and sex, potential confounding factors might be minimized. An inadequacy in matching on smoking exposure would be controlled in data analysis with additional adjustment.

In conclusion, we found that the $-283T>C$ and $-579G>T$ polymorphisms in the DNMT3B promoter, and their
haplotypes were significantly associated with the risk of lung cancer, particularly AC. These results suggest that the DNMT3B gene may be involved in the development of lung cancer although additional studies with larger sample sizes are required to confirm our findings. Future studies of other DNMT3B sequence variants and their biologic function are also needed to understand the role of DNMT3B polymorphisms in determining the risk of lung cancer. Moreover, since genetic polymorphisms often vary between ethnic groups, further studies are needed to clarify the association of the DNMT3B polymorphism with lung cancer in diverse ethnic populations.

Acknowledgements

This study was supported by a grant of National Cancer Control R&D Program 2003, Ministry of Health & Welfare, Republic of Korea.

References


Received June 29, 2004; revised September 30, 2004; accepted October 1, 2004.