A Novel Polymorphism in Human Cytosine DNA-Methyltransferase-3B Promoter Is Associated with an Increased Risk of Lung Cancer

Hongbing Shen, Luo Wang, Margaret R. Spitz, Waun K. Hong, Li Mao, and Qingyi Wei

Departments of Epidemiology [H. S., M. R. S., Q. W.] and Thoracic and Head and Neck Medical Oncology [L. W., W. K. H., L. M.], The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030

ABSTRACT

DNA repair is central to genomic integrity. Reduced expression of several nucleotide excision repair genes has been demonstrated to be associated with increased risk of lung cancer. Because methylation of gene promoters is one of the major regulatory mechanisms of gene expression and most nucleotide excision repair gene promoters have not been fully characterized, we hypothesized that genetic variants of the genes that are responsible for regulating genomic methylation are associated with increased risk of lung cancer. Recently, we identified a C→T transition at a novel promoter region of cytosine DNA-methyltransferase-3B (DNMT3B) and found that this polymorphic transition significantly increases the promoter activity. In this hospital-based case-control study of 319 patients with incident lung cancer and 340 healthy controls frequency-matched on age (±5 years), sex, ethnicity, and smoking status, we genotyped subjects for this DNMT3B promoter polymorphism to determine the association between this genetic variant and risk of lung cancer. Compared with CC homozygotes, CT heterozygotes had a >2-fold increased risk of lung cancer [adjusted odds ratio (OR), 2.13; 95% confidence interval (CI), 1.47–3.08] and TT homozygotes an OR of 1.42 (95% CI, 0.91–2.21). The combined variant genotype (CT + TT) was associated with nearly a 2-fold increased risk (adjusted OR, 1.88; 95% CI, 1.32–2.66). These results suggest that this novel variant of DNMT3B is associated with increased risk of lung cancer and may contribute to identifying individuals genetically susceptible to tobacco-induced cancers. Additional studies on the underlying molecular mechanism of this polymorphism are warranted.

INTRODUCTION

Lung cancer is the leading cause of cancer-related death for both men and women in the United States; ~154,900 lung cancer deaths will occur in the year 2002 (1). Although significant progress has been made toward understanding genetic changes associated with lung cancer, there are few useful biomarkers for identifying individuals at high risk in the general population (2). Therefore, it is important to investigate genetic and epigenetic variation in susceptibility to lung carcinogenesis and identify markers that will facilitate identification of individuals at risk of tobacco carcinogenesis.

Tobacco constituents such as B(α)P2 and especially its bioactivated form, BPDE, a polycyclic aromatic hydrocarbon, can damage DNA irreversibly by covalent binding or by oxidation to form BPDE-DNA adducts (3). The levels of polycyclic aromatic hydrocarbon-DNA adducts in tissues are higher in smokers than in nonsmokers (4), suggesting that tobacco smoke is a source of B(α)P that forms DNA adducts in humans. These BPDE-DNA adducts can block the transcription of essential genes and lead to carcinogenesis if not repaired efficiently by the NER pathway (5). Therefore, DNA repair is central to genomic integrity, and reduced DNA repair capacity is associated with increased risk of environmentally induced cancer.

In our recently reported molecular epidemiological study, we demonstrated that reduced NER capacity is associated with increased risk of lung cancer (6). Our other studies suggested that genetic variation (7) and altered expression levels (8) of several NER genes are also associated with increased risk of lung cancer. It is biologically plausible that epigenetic alterations in expression of NER genes could reduce DNA repair capacity (6) and hinder removal of DNA adducts induced by smoking-related carcinogens such as BPDE.

DNA methylation is a major epigenetic modification in humans, and aberrant DNA cytosine methylation may play an important role in carcinogenesis because methylation facilitates gene mutation through deamination of 5-methyl cytosine to thymine (9). Many tumors have genome-wide hypomethylation and CpG island-specific hypermethylation in the promoters of many tumor suppressor genes (10). De novo hypermethylation in promoter CpG islands has been identified as a possible mechanism for tumor suppressor gene inactivation in human cancer cells (10). Aberrant promoter methylation in various genes such as p16(ink4a) is also involved in human lung cancer (11). Recent studies have linked this aberrant de novo methylation of CpG islands to the overexpression of the newly cloned DNMT3 family (DNMT3A and DNMT3B; Refs. 12, 13). DNMT3A and DNMT3B are required for the establishment and maintenance of genomic methylation patterns and proper murine development (13). Both genes are up-regulated in some malignancies, including bladder, colon, kidney, and pancreas cancers, although to different degrees (14).

We previously identified a C→T single-base transition in a novel promoter of DNMT3B, −149 bp from the transcription start site and found that this polymorphism can significantly increase the promoter activity of DNMT3B gene (15). Because methylation of promoter regions is one of the major regulatory mechanisms of gene expression and most NER gene promoters have not been fully described, we hypothesized that genetic variants of genes such as the DNMTs that are responsible for regulating the methylation status of other genes are associated with increased risk of lung cancer. In this hospital-based case-control study of 319 incident lung cancer case patients and 340 healthy control subjects frequency-matched by age (±5 years), sex, ethnicity, and smoking status, we genotyped this DNMT3B polymorphism to investigate the association between this genetic variant and risk of lung cancer.

MATERIALS AND METHODS

Study Subjects. Lung cancer patients were selected consecutively from an ongoing molecular epidemiological study of lung cancer conducted in the Department of Epidemiology at The University of Texas M. D. Anderson Cancer Center (6). The 319 case subjects were patients with newly diagnosed, histopathologically confirmed, and previously untreated (by radiotherapy or chemotherapy) lung cancer at our institution. There were no age, stage, or histology restrictions, and all of the cases were non-Hispanic whites. The 340

Downloaded from cancerres.aacrjournals.org on July 13, 2017. © 2002 American Association for Cancer Research.
control subjects were selected from a pool of cancer-free subjects recruited from the largest multispecialty managed-care organization (Kelsey Seybold Clinics in the Houston metropolitan area and were frequency-matched to the case patients on age (±5 years), sex, ethnicity, and smoking status. The participation response rate of this case-control study was 77.4% among cases and 73.3% among control subjects (6). Those who had smoked <100 cigarettes in their lifetimes were defined as never smokers as traditionally used in epidemiological studies; those who had quit smoking for >1 year previously were considered former smokers, and the rest were considered current smokers. The exclusion criteria included previous radiotherapy or chemotherapy, previous cancer, and recent (in last 6 months) blood transfusion. Each subject was scheduled for an interview after informed consent was obtained, and a structured questionnaire was administered by interviewers to collect information on demographic data and risk factors such as smoking status. The study was approved by the Institutional Review Boards of M. D. Anderson Cancer Center and the Kelsey Seybold Clinics.

DNMT3B Genotyping. DNA was extracted from the leukocyte cell pellet obtained from each blood sample’s buffy coat by centrifugation of 1 ml of whole blood. The Qiagen DNA Blood Mini Kit (Qiagen Inc., Valencia, CA) was used according to the manufacturer’s instructions to obtain genomic DNA. We developed a PCR restriction fragment length polymorphism assay to identify this C for C transition at nt 46359 (C46359T, GenBank accession no. AB053071) in the novel promoter of DNMT3B (nt 46151–TGCTGTGACAGGCAGAGCAG-3′/H9273). This 380-bp target DNA fragment contains the upstream region and the first exon of DNMT3B gene. The 20-μl PCR mixture contained ~50 ng of genomic DNA, 12.5 pmol of each primer, 0.1 mM each deoxynucleotide triphosphate, 1× PCR buffer [50 mM KCl, 10 mM Tris-HCl, and 0.1% Triton X-100], 1.5 mM MgCl₂, and 1.0 units of Taq polymerase (Sigma Chemical Co., St. Louis, MO). The PCR profile consisted of an initial melting step of 95°C for 5 min; 35 cycles of 95°C for 30 s, 65°C for 30 s, and 72°C for 30 s; and a final extension step of 72°C for 10 min. The 380-bp fragment was then digested with AvrII (New England BioLabs, Inc., Beverly, MA) overnight at 37°C. The digested product was separated on a 2.5% NuSieve 3:1 agarose (FMC BioProducts, Rockland, ME) gel with ethidium bromide and photographed with a Polaroid film. The variant T allele had an AvrII restriction site that results in two bands (207 and 173 bp), and the wild-type C allele lacked the AvrII restriction site and therefore produced a single 380-bp band (Fig. 1). PCR was conducted and the results were evaluated by one of us (H. S.) without knowledge of the subjects’ case-control status. More than 10% of the samples were randomly selected for repeated assays, and the results were 100% concordant.

Statistical Analysis. Differences in select demographic variables, smoking status, pack-years smoked, and family history of any cancer between lung cancer case patients and control subjects were evaluated by using the χ² test. The associations between the DNMT3B variant and risk of lung cancer were estimated by computing the ORs and their 95% CIs from both univariate and multivariate logistic regression analyses. Stratification analysis was used to study subgroups of subjects by age, sex, and lung cancer histology. All of the statistical analyses were performed with Statistical Analysis System software (v.8.0; SAS Institute, Cary, NC). RESULTS

The select characteristics of the 319 lung cancer case patients and 340 control subjects are summarized in Table 1. The lung cancer case patients and control subjects appeared to be adequately matched on age and sex, although there were slightly higher percentages of female subjects among the case patients than among the control subjects. The mean age was 59.6 years (±10.2 years; range, 34–83 years) for the case patients and 59.8 years (±10.3 years; range, 32–84 years) for the control subjects. Because we frequency matched the subjects on smoking status, there was no significant difference in the smoking status of the case patients and control subjects. However, there were more heavy smokers (pack-years > 38) among the case patients (62.7%) than among the control subjects (50.9%). The case patients were significantly more likely than the control subjects to report a family history of any cancer (70.2 versus 59.4%; P = 0.004) in their first-degree relatives. Of the 319 lung cancer case patients, 161 (50.5%) were classified as adenocarcinomas, 60 (18.8%) as squamous cell carcinomas, 55 (17.2%) as other non-small cell carcinomas, 23 (7.2%) as small cell carcinomas, and 20 (6.3%) as other carcinomas, including large cell and giant cell carcinomas.

The DNMT3B T allele frequency and genotype distributions in the case patients and control subjects are summarized in Table 2. The distributions of the genotypes among the control subjects were in agreement with Hardy-Weinberg equilibrium (P = 0.129). We observed genotype frequencies of 22.3% (CC), 56.7% (CT), and 21.0% (TT) in the case patients and 35.0% (CC), 41.8% (CT), and 23.2% (TT) in the control subjects, respectively, and the differences between case patients and control subjects were statistically significant (χ² = 17.2; P = 0.0002). However, the variant T allele frequency was 0.494 for the case patients and 0.441 for the control subjects, and this difference was only borderline statistically significant (χ² = 3.7; P = 0.056).

By using logistic regression analysis, we evaluated the association between the T variant and risk of lung cancer. Compared with CC homozygotes, CT heterozygotes had a >2-fold increased risk (adjusted OR, 2.13; 95% CI, 1.47–3.08), and the TT homozygotes had a moderately elevated risk (adjusted OR, 1.42; 95% CI, 0.91–2.21). For the combined variant genotypes (CT + TT), there was a nearly 2-fold increased risk (adjusted OR, 1.88; 95% CI, 1.32–2.62; Table 2).

Associations between the DNMT3B genotype and lung cancer stratified on age, sex, and histopathological type of case patients are shown in Table 3. The risk associated with the combined CT + TT variant genotypes was more pronounced in subjects younger at diagnosis (<60 years; OR, 2.34; 95% CI, 1.34–4.00), in women (OR, 2.72; 95% CI, 1.57–4.71), and in patients with squamous cell carci-

![Fig. 1. PCR-based restriction fragment length polymorphism genotyping of DNMT3B promoter C46359T. Lanes 1, 4, and 5: CC wild-type; Lanes 2 and 3: CT heterozygotes; and Lanes 6 and 7: TT variant.](image-url)
polymorphism in the human case-control study, we first reported the association between a novel variant between either pack-years smoked or family history of cancer and the risk of lung cancer (OR, 2.81; 95% CI, 1.35–5.85). No interactions were observed between either pack-years smoked or family history of cancer and the variant DNMT3B genotypes (data not shown).

**DISCUSSION**

Single nucleotide polymorphisms are the most common form of human genetic variation. Single nucleotide polymorphisms may contribute to individual susceptibility to cancer but often the underlying molecular mechanism is unknown. Sequence analyses have suggested that some variants, especially those in the promoter regions of genes, may affect either the expression or activity levels of enzymes (16) and therefore may be mechanistically associated with cancer risk. In this case-control study, we first reported the association between a novel polymorphism in the human DNMT3B promoter and risk of lung cancer. We found that the T variant genotype was associated with a significantly increased risk of lung cancer. Although the mechanism of this association is unknown, one possible explanation is that this C→T transition, which increases the DNMT3b promoter activity, may up-regulate gene expression that involves an aberrant de novo methylation of CpG islands in some tumor suppressor genes (12, 14).

DNMT3B has many alternatively spliced forms and differential activity of DNMT3B variants that could have important implications in carcinogenesis (14). Different isoforms of the same enzyme may have altered catalytic activity or target-site specificity that may play a role in carcinogenesis (14). DNA methylation is essential for embryonic development and important for transcriptional repression of imprinted genes such as those on the inactive X chromosome, and critical genes are aberrantly silenced in many human cancers (17). This DNA modification is mediated by DNMTs (13). A recent study showed that three active DNMTs (DNMT1, DNMT3A, and DNMT3B) are required for the establishment and maintenance of genomic methylation patterns and proper murine development (18). However, the three DNMTs are located in different areas of the nucleus and are associated with different nuclear complexes during the cell cycle, indicating that the DNMTs have very different functions in mammalian cells (13). Overexpression of both the maintenance DNA methyltransferase DNMT1 and DNMT3B is involved in human carcinogenesis, probably at different stages and through different mechanisms (18).

In this study, we observed that the association between the DNMT3B variant and lung cancer risk was more pronounced in younger subjects (< 60 years), suggesting that genetic susceptibility, often associated with an early age of onset, may play a more important role in the development of lung cancer among younger patients than among relatively older subjects. We have observed similar findings with polymorphisms in the DNA repair gene XRCC1 in head and neck cancer (19). We also observed a significantly higher risk for lung cancer associated with the DNMT3B polymorphism in women than in men. A suggestive explanation for this is that women with DNMT3B variant genotypes are more sensitive to tobacco carcinogens than men are, which is consistent with previous findings that women tend to have a higher risk of lung cancer than men with the same level of tobacco exposure (20) and that women have significantly lower DNA repair capacity than do men (6). In addition, we found that risk associated with the combined variant genotypes was higher for lung cancer patients with squamous cell carcinoma, suggesting that different histopathological types may have different etiologies, not only in relation to environmental risk factors but also in genetic susceptibility. However, the small sample size in this case group did not allow testing for gene-environment interactions, therefore, this finding needs to be confirmed by larger studies.

Because this is a hospital-based case-control study and the controls were not selected from the same population from which the cases arose, we could not rule out possible selection bias. However, by matching on age, sex, ethnicity, and smoking status, potential

### Table 2

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Case patients (n = 319)</th>
<th>Control subjects (n = 340)</th>
<th>Crude OR (95% CI)</th>
<th>Adjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC (ref.)</td>
<td>71 (22.3)</td>
<td>119 (35.0)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>CT</td>
<td>181 (56.7)</td>
<td>142 (41.8)</td>
<td>2.14 (1.48–3.08)</td>
<td>2.13 (1.47–3.08)</td>
</tr>
<tr>
<td>TT</td>
<td>67 (21.0)</td>
<td>79 (23.2)</td>
<td>1.42 (0.92–2.20)</td>
<td>1.42 (0.91–2.21)</td>
</tr>
<tr>
<td>CT + TT</td>
<td>248 (77.7)</td>
<td>221 (65.0)</td>
<td>1.88 (1.33–2.66)</td>
<td>1.88 (1.32–2.66)</td>
</tr>
</tbody>
</table>

*χ² = 17.2, P = 0.0002 (genotype); χ² = 3.7, P = 0.056 (allele); the observed genotype frequency in the control subjects was in agreement with Hardy-Weinberg equilibrium (p² + 2pq + q² = 1) (χ² = 4.10, P = 0.129).

ORs were adjusted for age, sex, pack-years, and family history of cancer in a logistic regression model.

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>CT (OR (95% CI))</th>
<th>TT (OR (95% CI))</th>
<th>CT/TT (OR (95% CI))</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC (ref.)</td>
<td>2.13 (1.47–3.08)</td>
<td>1.42 (0.91–2.21)</td>
<td>1.88 (1.32–2.66)</td>
</tr>
<tr>
<td>Age (yr)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;60</td>
<td>2.90 (1.64–5.13)</td>
<td>1.50 (0.77–2.91)</td>
<td>2.34 (1.37–4.00)</td>
</tr>
<tr>
<td>≥60</td>
<td>1.72 (1.05–2.82)</td>
<td>1.41 (0.77–2.60)</td>
<td>1.62 (1.02–2.59)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.54 (0.95–2.50)</td>
<td>1.18 (0.64–2.16)</td>
<td>1.42 (0.90–2.25)</td>
</tr>
<tr>
<td>Female</td>
<td>3.30 (1.84–5.93)</td>
<td>1.88 (0.96–3.65)</td>
<td>2.72 (1.57–4.71)</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>1.84 (1.17–2.91)</td>
<td>1.41 (0.82–2.42)</td>
<td>1.69 (1.10–2.60)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>3.21 (1.50–6.83)</td>
<td>2.03 (0.82–5.03)</td>
<td>2.81 (1.35–5.85)</td>
</tr>
<tr>
<td>Non-small cell carcinoma</td>
<td>2.09 (1.04–4.23)</td>
<td>1.14 (0.46–2.85)</td>
<td>1.77 (0.90–3.50)</td>
</tr>
<tr>
<td>Small cell carcinoma</td>
<td>1.67 (0.59–4.72)</td>
<td>1.33 (0.38–4.63)</td>
<td>1.55 (0.58–4.14)</td>
</tr>
<tr>
<td>Others</td>
<td>2.35 (0.70–7.92)</td>
<td>1.24 (0.29–5.33)</td>
<td>1.90 (0.60–6.03)</td>
</tr>
</tbody>
</table>

* Adjusted for age, sex, pack-years smoked, and family history of cancer within the strata.
confounding factors might be minimized. Any inadequacy in matching would be controlled in data analysis with additional adjustment.

In conclusion, our study provides the first evidence that this newly identified polymorphism of the DNMT3B promoter, particularly for the variant heterozygous genotype, is significantly associated with increased risk of lung cancer in this non-Hispanic white population. This association is especially noteworthy in women, in younger individuals, and in squamous cell carcinoma. Although the reason for a higher risk associated with the variant heterozygotes remains unknown, it is possible that this heterozygous genotype may be in linkage disequilibrium with other susceptibility loci. Another possible explanation is that the CT heterozygotes may have impaired function because of imbalance of the protein structure of DNMT3B. Larger studies are needed to verify these findings. To further explore the role of the DNMT3B gene in the etiology of lung cancer, we are currently testing the hypothesis that this C→T polymorphism is associated with increased risk of lung cancer in this non-Hispanic white population.

ACKNOWLEDGMENTS

We thank Dr. Maureen Goode (Department of Scientific Publications, The University of Texas M. D. Anderson Cancer Center, Houston, TX) for editing the manuscript; Dr. Carol Bosken for tumor pathological classification; Kristina Dahlstrom and Li-E Wang for laboratory assistance; Susan Honn for assistance in recruiting study subjects; Wayne Gosbee for assistance with data management; and Joanne Sider for assistance in preparing the manuscript.

REFERENCES

A Novel Polymorphism in Human Cytosine DNA-Methyltransferase-3B Promoter Is Associated with an Increased Risk of Lung Cancer

Hongbing Shen, Luo Wang, Margaret R. Spitz, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/62/17/4992

Cited articles
This article cites 17 articles, 13 of which you can access for free at:
http://cancerres.aacrjournals.org/content/62/17/4992.full#ref-list-1

Citing articles
This article has been cited by 26 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/62/17/4992.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.