Association of Genetic Variants of O^6-Methylguanine-DNA Methyltransferase with Risk of Lung Cancer in Non-Hispanic Whites

Luo Wang, Hongji Liu, Zhengdong Zhang, Margaret R. Spitz, and Qingyi Wei

Department of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas

Abstract

O^6-methylguanine, a methylated damage lesion in DNA, correlates with spontaneous G:C → A:T transition mutations and leads to activation of oncogene *K-ras* or dysfunction of the tumor suppressor gene *p53*. O^6-methylguanine-DNA methyltransferase (MGMT) is critical for repairing damage to the O^6-position of guanine. Therefore, we tested our hypothesis that genetic variants of *MGMT* are associated with increased lung cancer risk in a Caucasian population of 1,121 lung cancer patients and 1,163 matched cancer-free controls. We genotyped four potentially functional single nucleotide polymorphisms (SNPs) of *MGMT*: exon 3 codon 84C → T (L84F), exon 5 codon 143A → G (I143V), and two promoter SNPs 135G → T and 485C → A. The allele frequency distributions of the SNPs of codon 84C → T and the promoter 135G → T in the cases were borderline different from that in the controls. After defining the minor allele (T for codon 84C → T and G for codon 143A → G) as the variant allele, we categorized the *MGMT* genotypes as either 0 variants (84CC-143AA) or 1-4 variants. Compared with 0 variants, those with 1-4 variants showed a statistically significantly increased risk of lung cancer (P = 0.040). Further stratification analysis showed that this increased risk was more pronounced in women, current smokers, and non–small cell lung cancer. We did not find any association between the *MGMT* promoter SNPs and lung cancer risk. Our findings suggest that non-synonymous SNPs in *MGMT* are associated with modestly increased risk of lung cancer in Caucasians and need to be further investigated. (Cancer Epidemiol Biomarkers Prev 2006;15(12):2364–9)

Introduction

DNA repair maintains the genomic integrity of human cells against damage induced by environmental hazards, including tobacco carcinogens (1, 2). The G:C → A:T transition is a frequent mutation observed in tumors and cancer cell lines, which may result in activation of the *K-ras* oncogene or dysfunction of the *p53* tumor suppressor gene (3, 4). Although spontaneous deamination of methylcytosine is a source of the G:C → A:T mutation (5), the O^6-methylguanine lesions induced by the nicotine-derived 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone may also lead to the G → A mutation, and persistence of O^6-methylguanine is one of the critical events during lung tumorigenesis in the A/J mice model (6, 7).

O^6-methylguanine-DNA methyltransferase (MGMT) is a DNA repair enzyme that specifically repairs damage to the O^6 position of guanine in DNA through transferring the O^6-methyl group to an internal cysteine residue of the protein (8). Interestingly, the G:C → A:T mutation in both *K-ras* and *p53* has been observed to be correlated with inactivation of MGMT (9, 10). In *MGMT*-defective mice, numerous lung adenomas and thymic lymphomas developed after exposure to low-dose methylnitrosourea (11). In contrast, transgenic mice carrying extra copies of the foreign *MGMT* gene showed decreased susceptibility to carcinogenesis (12). These results suggest a central role of the *MGMT* gene in protecting genomic integrity through repairing the O^6-methylguanine induced by tobacco-specific nitrosamines, including 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone.

The *MGMT* gene is located at chromosome 10q26, and its mRNA contains only five exons. A single 207-amino-acid peptide, coded by only four exons with the translation starting codon ATG in the second exon, has been shown to have basic methyl-transferring activity (13). Twelve single nucleotide polymorphisms (SNPs) in *MGMT* exons have been identified in different populations, of which only three are non-synonymous SNPs with a variant allele frequency >0.05 in Caucasians. These three SNPs are codon 84C → T (L84F), codon 143A → G (I143V), and codon 178A → G (K178R) (14-16). They seem to have an effect on the function of MGMT, possibly leading to increased cancer risk. Therefore, there have been several case-control studies that evaluated the associations of these three SNPs with risk of bladder, lung, brain, and breast cancers (17-24). Although modest risks were reported in some studies for *MGMT* variants, there are no conclusive findings regarding the role of these *MGMT* genetic variants in cancer etiology because most studies included relatively small sample sizes of cases and controls with mixed ethnic backgrounds and cancer types in different populations.

A typical CpG island exists in the *MGMT* promoter region, and hypermethylation of CpG sites in this region is common in many cancers (25, 26), indicating that epigenetic events through DNA cytosine methylation are one of the important mechanisms regulating *MGMT* mRNA expression in specific targeted tissues (27). In addition, there are five SNPs found in the *MGMT* promoter region, with allele frequencies of 135G → T and 485C → A reported as 0.15 and 0.37, respectively (14). It is likely that genetic variation in the promoter region may also play a role in the MGMT-associated susceptibility to lung cancer (23). To test the hypothesis that genetic variants in exons and the promoter of *MGMT* are associated with increased lung cancer risk, we selected and genotyped two non-synonymous SNPs [i.e., exon 3 codon 84 C → T (L84F) and

Received 5/31/06; revised 9/2/06; accepted 9/25/06.

Grant support: NIH grants CA55769 (M.R. Spitz), ES11740 (Q. Wei), and CA16672 (The University of Texas M.D. Anderson Cancer Center).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Qingyi Wei, Department of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Unit 1365, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-792-3020; Fax: 713-792-0807. E-mail: qwei@mdanderson.org

doi:10.1158/1055-9965.EPI-06-0437

Table 1. Primer information of four SNPs used in this study

SNP*	dbSNP	PCR primers (5′ → 3′)	Annealing temp. (°C)	Note †
Promoter 135G → T	rs1711646	Forward, TGTGTTAGGATCCTGCTCCC Reverse, AGTATGACTCGTGTCCTTATAGGGCG	58	Newly designed
Promoter 485C → A	rs1625649	Forward, TGAGTCAGGCTCTGGCAGTGTCT Reverse, CTTAGTGAGAATCCCGGTCCTGC	64	Newly designed
Codon 84C → T (L84F)	rs12917	Forward, AGGCTATCGAAGAGTTCCCC Reverse, TAAGTCAAGCTCCCCAAAGG	60	Ref. (17)
Codon 143A → G (I143V)	rs2308321	Forward, CTGTCTTCCAGGTCCCCCTCC Reverse, TCCCTTGAGCCAGTACCCTGC	62	Newly designed

*Sequence accession nos. are AL355531 for promoter SNPs 135G → T and 485C → A and NM_002412 for codon 84C → T (L84F) and codon 143A → G (I143V).
†We used the mismatch approach to design the primers for 135G → T and codon 143A → G.

exon 5 codon 143A → G (I143V) that is in complete linkage disequilibrium with codon 178A → G (K178R) in Caucasian population (20)] and two promoter SNPs (i.e., 135G → T and 485C → A) in a large lung cancer case-control study.

Materials and Methods

Subjects. The recruitment of study subjects has been described previously (28). Briefly, the subjects in the final analysis included 1,121 patients with lung cancer and 1,163 cancer-free control subjects recruited consecutively during the period of August 1995 to December 2004. The case subjects were patients with newly diagnosed and histopathologically confirmed lung cancer. There were no restrictions on age, sex, disease stage, or histology, but all of them were non-Hispanic Whites. The controls were selected from a pool of cancer-free individuals recruited through the Kelsey-Seybold multispecialty physician practice with multiple clinics throughout the Houston metropolitan area and were frequency matched with the patients according to age (±5 years), sex, ethnicity, and smoking status (i.e., current, former, and never). The exclusion criteria for both cases and controls, when appropriate, included previous cancer treatment (i.e., radiotherapy or chemotherapy), previous cancer except non-melanoma skin cancer, and recent blood transfusion (within previous 6 months). After providing a written informed consent, each participant in this study was interviewed by an interviewer with a structured questionnaire, and the blood samples were collected. The study was approved by the institutional review boards of the M.D. Anderson Cancer Center and the Kelsey-Seybold Foundation.

Selection of SNPs in the *MGMT* Gene. We chose those SNPs with allele frequencies >0.05 and those that are likely to have biologically functional relevance (i.e., non-synonymous SNPs and the SNPs in the promoter region; ref. 29). *MGMT* mRNA only has five exons, and there is no evidence that alternative splicing involves in regulation of *MGMT* mRNA. Therefore, we did not include intron SNPs. Three non-synonymous SNPs [codon 84C → T (L84F), codon 143A → G (I143V), and codon 178A → G (K178R)] were chosen, but codon 178A → G (K178R) was subsequently excluded because it is in strong linkage disequilibrium with codon 143A → G (I143V) ($D' = 0.96$, $r^2 = 0.89$, $P < 0.01$; refs. 20, 22). We also included two SNPs (i.e., 135G → T and 485C → A) in the *MGMT* promoter region with reported allele frequencies of 0.15 and 0.37, respectively (14).

Genotyping. Genomic DNA was extracted from blood leukocyte cell pellet by using the Qiagen DNA blood mini kit (Qiagen, Valencia, CA) according to the manufacturer's instruction.

The PCR-RFLP method was used for genotyping. Briefly, a 10-μL PCR reaction mixture was assembled with 40 ng genomic DNA and cycled in a PTC-200 DNA Engine (Peltier Thermal Cycler, MJ Research, Inc., Watertown, MA). The information on primers used to detect the SNPs is listed in Table 1. The lengths of PCR fragments for 135G → T, 485C → A, codon 84C → T (L84F), and codon 143A → G (I143V) were 168, 212, 102, and 179 bp, respectively. The restriction enzymes *Hha*I, *Ban*I, *Ear*I, and *Bst*F5I (New England Biolabs, Beverly, MA) were used to distinguish these SNPs by generating the 144- and 24-bp fragments in the presence of the 135G → T G allele; 139- and 73-bp in the presence of the 485C → A A allele; 68-, 30-, and 4-bp in the presence of codon 84C → T (L84F) T allele; and 157- and 22-bp for codon 143A → G (I143V) A allele. The PCR products were visualized on a 3% agarose gel containing 0.25 μg/mL ethidium bromide and photographed. The genomic structure of the *MGMT* gene and the PCR-RFLP patterns of four SNPs are shown in Fig. 1. We also randomly selected >10% of all genotyped DNA samples and repeated PCR-RFLP genotyping for validation of results. The repeated data were 100% concordant with the original genotyping results.

Statistical Analysis. The demographic characteristics (e.g., age and sex), smoking status, pack-years smoked, and tumor histology were included in our analysis. Those who had smoked fewer than 100 cigarettes in their lifetimes were defined as never smokers, and the others were defined as ever smokers. Of the ever smokers, those who quit smoking a year before interview were considered former smokers, and the remainder were considered as current smokers. We used the χ^2 test to detect the differences in the frequency distributions of categorical variables as well as the genotypes and alleles of the selected *MGMT* SNPs between the cases and controls. In addition, the Student's *t* test was used to detect the differences in numerical variables, such as age and number of pack-years smoked.

To assess risk of lung cancer for individuals who carried variant alleles of *MGMT*, we defined the minor alleles of *MGMT* exon 3 codon 84C → T (L84F) (T allele) and exon 5 codon 143A → G (I143V) (G allele) as the variant alleles based on their higher frequencies in cases than that in controls. Finally, we created a summary variable based on the numbers of variant alleles (0 versus 1-4 variants). Unconditional multivariate analyses were done to calculate adjusted odds ratios and 95% confidence intervals for the defined *MGMT* variant genotypes with adjustment for age, sex, smoking status, and pack-years in logistic regression models when appropriate. We further stratified frequency distributions of the defined *MGMT* variant genotypes by age, sex, smoking status, pack-years, and tumor histologies as well as the *MGMT* promoter SNPs 135G → T and 485C → As and assessed their potential interactions by fitting a product of two interactive variables in a multiplicative model with adjustment for their main effects and other covariates. All tests were two sided with a significance level of <0.05 and done with the Statistical Analysis System Software program (SAS software, version 9.1; SAS Institute, Inc., Cary, NC).

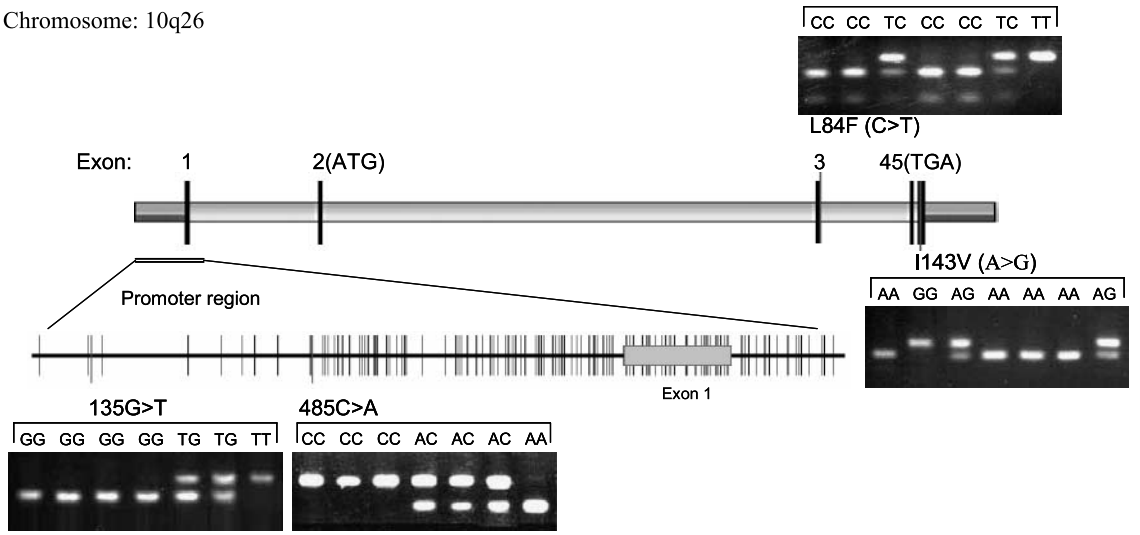

Figure 1. The genomic structure of the *MGMT* gene, locations of four interesting SNPs, and the corresponding genotyping patterns. Vertical bar in the promoter region indicates CpG sites.

Results

Selected characteristics of the study population are shown in Table 2. All subjects were non-Hispanic Whites. The cases and controls were well matched on age with a mean age of 61.4 years in the cases and 61.1 years in the controls. Similarly, the cases and controls were well matched on sex. Although the cases and controls were frequency matched on smoking status, there were more current smokers among the cases than the controls (41.0% versus 35.3%, $P = 0.017$). Cases were also more likely to be heavier smokers, and their mean number of pack-years smoked was significantly higher (44.7 ± 35.3) than that of the controls (35.8 ± 31.1; $P < 0.001$). For any residual effect of the frequency matching, we further adjusted for these variables in the multivariate logistic regression analysis.

Table 2. Frequency distributions of selected characteristics in the patients and controls

Characteristic	No. patients (%)	No. controls (%)	*P**
Total	1,121	1,163	
Age (mean ± SD), y	61.4 ± 10.7	61.1 ± 9.9	0.551†
≤62	568 (50.7)	617 (53.0)	0.254
>62	553 (49.3)	546 (47.0)	
Sex			
Male	590 (52.6)	574 (49.4)	0.117
Female	531 (47.4)	589 (50.6)	
Smoking status			
Never	179 (16.0)	199 (17.1)	0.017
Former	482 (43.0)	554 (47.6)	
Current	460 (41.0)	410 (35.3)	
Pack-years smoked (mean ± SD)	44.7 ± 35.3	35.8 ± 31.1	<0.001†
Tumor histology			
Non–small cell lung cancer	1,016 (90.6)	—	—
Adenocarcinoma	590 (58.1)	—	—
Squamous cell carcinoma	252 (24.8)	—	—
Large cell carcinoma	38 (3.7)	—	—
Other‡	136 (13.4)	—	—
Small cell lung cancer	93 (8.3)	—	—
Unclassified	12 (1.1)	—	—

*Two-sided χ^2 test unless otherwise specified.
†Student's t test.
‡Other non–small cell lung cancer not clearly classified into the above three major histologic types.

We evaluated the observed genotypes in the cancer-free controls by performing the Hardy-Weinberg equilibrium ($p^2 + 2pq + q^2 = 1$) test, and the Ps in controls were 0.708 for the promoter SNP 135G → T, 0.850 for the promoter SNP 485C → A, 0.674 for codon 84C → T (L84F), and 0.916 for codon 143A → G (I143V). We then compared the overall genotype frequency distribution of these four *MGMT* SNPs between the cases and controls (Table 3). Although the differences were not statistically significant for any single SNP tested, the adjusted odds ratios showed that homozygotes for the promoter SNP 135T or 485A minor allele had slightly reduced risk of lung cancer, compared with their homozygotes of the common allele, whereas homozygotes for either the 84T or 143G allele were associated with a moderately increased risk, compared with their homozygotes of the common allele (Table 3). By combining two non-synonymous SNPs [codon 84C → T (L84F) and codon 143A → G (I143V)], we created a summary variable of variant alleles (i.e., T in codon 84C → T and G in codon 143A → G) and categorized the codon 84CC and codon 143AA as the reference group (0 variants), and the others as 1-4 variants (Table 4). The overall adjusted odds ratio associated with the 1-4 *MGMT* variants was 1.19 (95% confidence interval, 1.01-1.41; $P = 0.040$) compared with 0 variants.

We further examined the effects of the *MGMT* variants on lung cancer risk in stratification analyses by selected variables (Table 5). Significantly increased risk of lung cancer was more evident for those carrying 1-4 variants who were women, current smokers, and non–small cell lung cancer. There was a borderline significant trend that risk associated with 1-4 *MGMT* variants increased as the age decreased ($P = 0.055$).

Finally, we assessed the effects of the two promoter SNPs on the *MGMT* variants defined from the two non-synonymous SNPs (Table 5). Among those in the 1-4 *MGMT* variant subgroup, there was an increased risk of lung cancer in the subgroup who carried either the T allele of promoter SNP 135G → T or C allele of 485C → A with borderline statistical significance ($P = 0.060$ and 0.067, respectively). However, we did not detect any interaction between the *MGMT* variants of the two non-synonymous SNPs and age, smoking status, or presence of the promoter SNPs (data not shown).

Discussion

As one of the most important events during tumorigenesis, mutations in DNA need to be repaired to maintain genomic

Table 3. Frequency distribution of *MGMT* genotypes and association with risk of lung cancer

Genotype	No. patients (%)	No. controls (%)	P*	Adjusted OR (95% CI)†
Total	1,121 (100.0)	1,163 (100.0)		
Promoter SNP 135G → T				
GG	757 (67.5)	813 (69.9)	0.125	1.00
TG	339 (30.3)	314 (27.0)		1.15 (0.96-1.38)
TT	25 (2.2)	36 (3.1)		0.74 (0.44-1.25)
T allele	389 (17.4)	386 (16.6)	0.496	
Promoter SNP 485C → A				
CC	526 (46.9)	536 (46.1)	0.812	1.00
AC	486 (43.4)	505 (43.4)		0.99 (0.83-1.17)
AA	109 (9.7)	122 (10.5)		0.93 (0.70-1.24)
A allele	704 (31.4)	749 (32.2)	0.561	
Exon 3 codon 84C → T (L84F)				
CC (84LL)	832 (74.2)	872 (75.0)	0.228	1.00
TC (84LF)	259 (23.1)	272 (23.4)		1.00 (0.82-1.21)
TT (84FF)	30 (2.7)	19 (1.6)		1.66 (0.92-2.99)
T allele	319 (14.2)	310 (13.3)	0.377	
Exon 5 codon 143A → G (I143V)				
AA (143II)	836 (74.6)	903 (77.6)	0.220	1.00
AG (143IV)	269 (24.0)	244 (21.0)		1.21 (0.99-1.47)
GG (143VV)	16 (1.4)	16 (1.4)		0.94 (0.46-1.91)
G allele	301 (13.4)	276 (11.9)	0.113	

Abbreviations: OR, odds ratio; 95% CI, 95% confidence interval.
*Two-sided χ^2 test.
†Adjusted for age, sex, smoking status, and number of pack-years smoked in logistic regression models.

integrity of the cells. MGMT is of particular interest for its role in sporadic lung carcinogenesis because inactivation of MGMT correlates with mutations in *K-ras* and *p53* (9, 10). In this study, we investigated the association between genetic variants of *MGMT* and risk of lung cancer. We found that the allele frequency distributions of the codon 84T, codon 143G, and promoter SNP 135T variant genotypes were associated with a non-statistically significantly increased risk of lung cancer. By creating a new variable of *MGMT* variant genotype based on two non-synonymous SNPs (codon 84 and codon 143), we found that 1-4 variants were associated with increased risk of lung cancer compared with the homozygous wild-type of 0 variants (84CC-143AA). The risk was more pronounced in women, current smokers, and non–small cell lung cancer. These findings support our a priori hypothesis that genetic variants of *MGMT* are associated with increased lung cancer risk. It also compliments our previous report that epigenetic DNA cytosine methylation involves in the etiology of lung cancer, as shown by the finding that a functional C46359 C → T SNP in a novel DNMT3B promoter was associated with lung cancer risk (27, 30).

Several studies have investigated the role of MGMT/AGT variants in the etiology of lung cancer (19, 20, 23, 24). However, most of the studies had included relatively small numbers of subjects of different ethnic background. One early study found that the 178A → G SNP was associated with risk of lung cancer in Caucasians (20), but this was not substantiated by other two later studies (19, 24). The most recent and largest study of a Korean population with 432 cases and 432 controls (23) found that among the SNPs (i.e., 485C → A, Leu53Leu, and Leu84Phe) tested, only the promoter 485C>A was associated with risk of lung cancer. However, we did not find any evidence for an association of the promoter 135G → T and 485C → A SNPs with risk of lung cancer in the present study of

Table 4. Frequency distribution of the *MGMT* exonic variants and association with risk of lung cancer

No. variant alleles of *MGMT* codon 84-codon 143*	No. individuals (%)		P†	Adjusted OR (95% CI)‡	P
	Patients	Controls			
Total	1,121 (100.0)	1,163 (100.0)			
(0) CC-AA (84LL-143II)	604 (53.9)	676 (58.1)	0.171	1.00	
(1) CT-AA (84LF-143II)	209 (18.6)	211 (18.1)		1.11 (0.89-1.39)	0.353
(1) CC-AG (84LL-143IV)	213 (19.0)	183 (15.7)		1.32 (1.05-1.66)	0.017
(2) TT-AA (84FF-143II)	23 (2.1)	16 (1.4)		1.57 (0.82-3.03)	0.174
(2) CC-GG (84LL-143VV)	15 (1.3)	13 (1.1)		1.15 (0.53-2.47)	0.726
(2) CT-AG (84LF-143IV)	49 (4.4)	58 (5.0)		0.95 (0.64-1.41)	0.792
(3) TT-AG (84FF-143IV)	7 (0.6)	3 (0.3)		2.90 (0.74-11.34)	0.126
(3) CT-GG (84LF-143VV)	1 (0.1)	3 (0.3)		0.25 (0.03-2.45)	0.232
(4) TT-GG (84FF-143VV)	0 (0.0)	0 (0.0)		NA	NA
				$P_{trend} = 0.085$	
0 variant	604 (53.9)	676 (58.1)	0.119	1.00	
1 variant	422 (37.6)	394 (33.9)		1.21 (1.01-1.44)	0.036
2-4 variants	95 (8.5)	93 (8.0)		1.12 (0.82-1.52)	0.480
				$P_{trend} = 0.092$	
0 variant	604 (53.9)	676 (58.1)	0.041	1.00	
1-4 variants	517 (46.1)	487 (41.9)		1.19 (1.01-1.41)	0.040

Abbreviations: OR, odds ratio; 95% CI, 95% confidence interval; NA, not available.
*The minor allele was defined as the variant allele (T in codon 84 and G in codon 143).
†Two-sided χ^2 test.
‡Adjusted for age, sex, smoking status, and number of pack-years smoked in logistic regression models.

Table 5. Frequency distribution and adjusted ORs (95% CIs) for the *MGMT* exonic variants stratified by selected variables in the patients and controls

Variables	No. patients (%)			No. controls (%)			Adjusted OR (95% CI)*		*P*
	Total	0 variants	1-4 variants	Total	0 variants	1-4 variants	0 variants	1-4 variants	
Age, y									
≤62	568	302 (53.2)	266 (46.8)	617	360 (58.3)	257 (41.7)	1.00	1.21 (0.96-1.54)	0.105
>62	553	302 (54.6)	251 (45.4)	546	316 (57.9)	230 (42.1)	1.00	1.14 (0.90-1.46)	0.277
Sex									
Male	590	319 (54.1)	271 (45.9)	574	322 (56.1)	252 (43.9)	1.00	1.09 (0.86-1.38)	0.463
Female	531	285 (53.7)	246 (46.3)	589	354 (60.1)	235 (39.9)	1.00	1.31 (1.03-1.66)	0.028
Smoking status									
Never	179	102 (57.0)	77 (43.0)	199	112 (56.3)	87 (43.7)	1.00	0.96 (0.64-1.45)	0.860
Former	482	257 (53.3)	225 (46.7)	554	317 (57.2)	237 (42.8)	1.00	1.16 (0.91-1.49)	0.237
Current	460	245 (53.3)	215 (46.7)	410	247 (60.2)	163 (39.8)	1.00	1.39 (1.06-1.84)	0.019
								P_{trend} = 0.122	
Tumor histology									
NSCLC	1,016	545 (53.6)	471 (46.4)				1.00	1.20 (1.01-1.43)	0.035
AC	590	314 (53.2)	276 (46.8)				1.00	1.22 (1.00-1.49)	0.055
SCC	252	143 (56.7)	109 (43.3)				1.00	1.06 (0.80-1.42)	0.672
LCC	38	20 (52.6)	18 (47.4)	1,163	676 (58.1)	487 (41.9)	1.00	1.24 (0.65-2.39)	0.514
Other	136	68 (50.0)	68 (50.0)				1.00	1.36 (0.94-1.94)	0.099
SCLC	93	52 (55.9)	41 (40.1)				1.00	1.12 (0.72-1.73)	0.618
NA	12	7 (58.3)	5 (41.9)				1.00	1.02 (0.32-3.26)	0.967
Promoter SNP 135G → T									
GG	757	409 (54.0)	348 (46.0)	813	466 (57.3)	347 (42.7)	1.00	1.13 (0.93-1.39)	0.227
TG	339	184 (54.3)	155 (45.7)	314	193 (61.5)	121 (38.5)	1.00	1.39 (1.01-1.91)	0.042
TT	25	11 (44.0)	14 (56.0)	36	17 (47.2)	19 (52.8)	1.00	1.21 (0.39-3.76)	0.742
TG + TT	364	195 (53.6)	169 (46.4)	350	210 (60.0)	140 (40.0)	1.00	1.34 (0.99-1.81)	0.060
Promoter SNP 485C → A									
CC	526	282 (53.6)	244 (46.4)	536	319 (59.5)	217 (40.5)	1.00	1.29 (1.01-1.65)	0.044
AC	486	263 (54.1)	223 (45.9)	505	283 (56.0)	222 (44.0)	1.00	1.09 (0.84-1.40)	0.516
AA	109	59 (54.1)	50 (45.9)	122	74 (60.7)	48 (39.3)	1.00	1.32 (0.78-2.24)	0.308
CC + AC	1,012	545 (53.8)	467 (46.2)	1,041	602 (57.8)	439 (42.2)	1.00	1.18 (0.99-1.41)	0.067

Abbreviations: NSCLC, non-small cell lung cancer; AC, adenocarcinoma; SCC, squamous cell carcinoma; LCC, large cell carcinoma; SCLC, small cell lung cancer; NA, unclassified; OR, odds ratio; 95% CI, 95% confidence interval.
*Adjusted for age, sex, smoking status, and number of pack-years smoked in logistic regression models when appropriate.

1,121 cases and 1,163 controls, the largest study to the best of our knowledge.

MGMT is one of the most conserved genes that protect genomic integrity against carcinogenic damage. To date, a total of 662 SNPs in the human *MGMT* gene have been reported in the dbSNP database of the National Center for Biotechnology Information. Most of the SNPs are located in introns or noncoding regions, and there are only 12 exonic SNPs, 8 nonsynonymous, and 4 synonymous. In general, those SNPs altering the conserved amino acids are more likely to be associated with cancer susceptibility (29), and in the two common *MGMT* non-synonymous SNPs, codon 178A → G (K178R) is in strong linkage disequilibrium with codon 143A → G (I143V) (20, 22). Therefore, the codon 84C → T and codon 143A → G of *MGMT* are two relevant candidate SNPs in our study. This choice was also supported by a finding that the cells from individuals carrying either the codon 84T (84F) or 143G (143V) alleles were found to be associated with significantly increased frequencies of chromosome aberrations induced by the tobacco carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (31).

Two functional domains of human MGMT have been interpreted by its crystal structure and enzymatic-kinetic study (32, 33). The NH_2-terminal residues 1 to 91 of MGMT contains a domain that can bind with zinc, and zinc binding will greatly enhance the methyl transferring activity of the intact MGMT protein. Meanwhile, the COOH-terminal residues 92 to 207 contain all residues thus far implicated in methyl-transferring activity, including a cysteine acceptor site (Cys^{145}), an O^6-methylguanine–binding pocket, and a DNA-binding domain (33). It is interesting that codon 84C → T (L84F) is located at the NH_2-terminal domain, and codon 143A → G (I143V) is located at the COOH-terminal domain. Therefore, it is likely that codon 84T (84F) and codon 143G (143V) variants may present different zinc-binding and methyl-transferring activities, respectively. Loss of such functionality in the variants is consistent with our finding that these variants were associated with increased risk of sporadic lung cancer in Caucasians. However, validation studies are needed in the populations with well-defined ethnic backgrounds because studies have shown that there was no codon 143A → G (I143V) variant in Asians (18) and Koreans (23).

The observed increased risk in women is of interest, but this kind of association study cannot provide a clear underlying mechanism. However, studies have shown that women seem to be more sensitive to tobacco-induced DNA damage (34). This may lead to increased O^6-methylguanine, the repair of which may be impaired in the presence of certain *MGMT* variants. Moreover, it seems that women tend to have a suboptimal DNA repair capacity to remove tobacco-induced DNA adducts (28). However, more evidence is needed to interpret the functionality of *MGMT* variants in maintaining the genomic integrity in humans, particularly in women whose DNA repair capacity may be under the influence of altered hormone status (35).

The strength of this study is that we only selected those SNPs that are most likely to be functional. Although we also included two *MGMT* promoter SNPs in this study, we did not find any risk associated with these variant genotypes, which may imply that epigenetic modification may play a more important role than sequence variation in the *MGMT* promoter in risk of lung cancer. This result is consistent with our previous finding that epigenetic modulation through *de novo* DNA-cytosine methylation was associated with increased risk of lung cancer (35).

In summary, we tested our hypothesis that genetic variants of *MGMT* are associated with increased lung cancer risk. We found that in non-Hispanic Whites, individuals having 1-4 *MGMT* variants showed a moderately increased risk of sporadic lung cancer compared with those having 0 variants. Stratification analysis further showed that this increased risk was more pronounced in women, ever smokers, and non–small cell lung cancer. Our findings suggest that nonsynonymous SNPs in *MGMT* are associated with modestly increased risk of lung cancer in Caucasians, but these findings need to be further validated.

Acknowledgments

We thank Susan Honn and the team for subject recruitment and database maintenance; Li-E Wang, Zhaozheng Guo, Yawei Qiao, Jianzhong He, and Kejing Xu for technical assistance; Joanne Sider for article preparation; and Donald R. Norwood for article editing.

References

1. Wei Q, Spitz MR. The role of DNA repair capacity in susceptibility to lung cancer: a review. Cancer Metastasis Rev 1997;16:295–307.
2. Spitz MR, Wei Q, Dong Q, Amos CI, Wu X. Genetic susceptibility to lung cancer: the role of DNA damage and repair. Cancer Epidemiol Biomarkers Prev 2003;12:689–98.
3. Guerrero I, Villasante A, Corces V, Pellicer A. Activation of a c-K-ras oncogene by somatic mutation in mouse lymphomas induced by gamma radiation. Science 1984;225:1159–62.
4. Levine AJ, Momand J, Finlay CA. The p53 tumour suppressor gene. Nature 1991;51:453–6.
5. Pfeifer GP. p53 mutational spectra and the role of methylated CpG sequences. Mutat Res 2000;450:155–66.
6. Peterson LA, Hecht SS. O^6-methylguanine is a critical determinant of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone tumorigenesis in A/J mouse lung. Cancer Res 1991;51:5557–64.
7. Peterson LA, Thomson NM, Crankshaw DL, Donaldson EE, Kenney PJ. Interactions between methylating and pyridyloxobutylating agents in A/J mouse lungs: implications for 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone-induced lung tumorigenesis. Cancer Res 2001;61:5757–63.
8. Pegg AE, Byers TL. Repair of DNA containing O^6-alkylguanine. FASEB J 1992;6:2302–10.
9. Esteller M, Toyota M, Sanchez-Cespedes M, et al. Inactivation of the DNA repair gene O^6-methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis. Cancer Res 2000;60:2368–71.
10. Esteller M, Risques RA, Toyota M, et al. Promoter hypermethylation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis. Cancer Res 2001;61:4689–92.
11. Sakumi K, Shiraishi A, Shimizu S, Tsuzuki T, Ishikawa T, Sekiguchi M. Methylnitrosourea-induced tumorigenesis in MGMT gene knockout mice. Cancer Res 1997;57:2415–8.
12. Nakatsuru Y, Matsukuma S, Sekiguchi M, Ishikawa T. Characterization of O^6-methylguanine-DNA methyltransferase in transgenic mice introduced with the E. coli ada gene. Mutat Res 1991;254:225–30.
13. Daniels DS, Woo TT, Luu KX, et al. DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat Struct Mol Biol 2004;11:714–20.
14. Egyhazi S, Ma S, Smoczynski K, Hansson J, Platz A, Ringborg U. Novel O^6-methylguanine-DNA methyltransferase SNPs: a frequency comparison of patients with familial melanoma and healthy individuals in Sweden. Hum Mutat 2002;20:408–9.
15. Ford BN, Ruttan CC, Kyle VL, Brackley ME, Glickman BW. Identification of single nucleotide polymorphisms in human DNA repair genes. Carcinogenesis 2000;21:1977–81.
16. Packer BR, Yeager M, Burdett L, et al. SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. Nucleic Acids Res 2006;34 Database issue: D617–21.
17. Li C, Liu J, Li A, et al. Exon 3 polymorphisms and haplotypes of O^6-methylguanine-DNA methyltransferase and risk of bladder cancer in southern China: a case-control analysis. Cancer Lett 2005;227:49–57.
18. Kaur TB, Travaline JM, Gaughan JP, Richie JP, Jr., Stellman SD, Lazarus P. Role of polymorphisms in codons 143 and 160 of the O^6-alkylguanine DNA alkyltransferase gene in lung cancer risk. Cancer Epidemiol Biomarkers Prev 2000;9:339–42.
19. Yang M, Coles BF, Caporaso NE, Choi Y, Lang NP, Kadlubar FF. Lack of association between Caucasian lung cancer risk and O^6-methylguanine-DNA methyltransferase-codon 178 genetic polymorphism. Lung Cancer 2004;44:281–6.
20. Cohet C, Borel S, Nyberg F, et al. Exon 5 polymorphisms in the O^6-alkylguanine DNA alkyltransferase gene and lung cancer risk in non-smokers exposed to second-hand smoke. Cancer Epidemiol Biomarkers Prev 2004;13:320–3.
21. Wiencke JK, Aldape K, McMillan A, et al. Molecular features of adult glioma associated with patient race/ethnicity, age, and a polymorphism in O^6-methylguanine-DNA-methyltransferase. Cancer Epidemiol Biomarkers Prev 2005;14:1774–83.
22. Shen J, Terry MB, Gammon MD, et al. MGMT genotype modulates the associations between cigarette smoking, dietary antioxidants and breast cancer risk. Carcinogenesis 2005;26:2131–7.
23. Chae MH, Jang JS, Kang HG, et al. O^6-alkylguanine-DNA alkyltransferase gene polymorphisms and the risk of primary lung cancer. Mol Carcinog 2006;45:239–49.
24. Zienolddiny S, Campa D, Lind H, et al. Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. Carcinogenesis 2006;27:560–7.
25. Esteller M, Hamilton SR, Burger PC, Baylin SB, Herman JG. Inactivation of the DNA repair gene O^6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. Cancer Res 1999;59:793–7.
26. Virmani AK, Tsou JA, Siegmund KD, et al. Hierarchical clustering of lung cancer cell lines using DNA methylation markers. Cancer Epidemiol Biomarkers Prev 2002;11:291–7.
27. Wang L, Liu Z, Mao L, Spitz MR, Wei Q. Functional relevance of the C/T polymorphism (C46359T) in the promoter region of human DNA-cytosine methyltransferase-3B6. Proc Am Assoc Cancer Res 2004;45:#2913–1.
28. Wei Q, Cheng L, Amos CI, et al. Repair of tobacco carcinogen-induced DNA adducts and lung cancer risk: a molecular epidemiologic study. J Natl Cancer Inst 2000;92:1764–72.
29. Zhu Y, Spitz MR, Amos CI, Lin J, Schabath MB, Wu X. An evolutionary perspective on single-nucleotide polymorphism screening in molecular cancer epidemiology. Cancer Res 2004;64:2251–7.
30. Shen H, Wang L, Spitz MR, Hong WK, Mao L, Wei Q. A novel polymorphism in human cytosine dna-methyltransferase-3b promoter is associated with an increased risk of lung cancer. Cancer Res 2002;62:4992–5.
31. Hill CE, Wickliffe JK, Wolfe KJ, Kinslow CJ, Lopez MS, Abdel-Rahman SZ. The L84F and the I143V polymorphisms in the O^6-methylguanine-DNA-methyltransferase (MGMT) gene increase human sensitivity to the genotoxic effects of the tobacco-specific nitrosamine carcinogen NNK. Pharmacogenomics 2005;15:571–8.
32. Daniels DS, Mol CD, Arvai AS, Kanugula S, Pegg AE, Tainer JA. Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. EMBO J 2000;19:1719–30.
33. Fang Q, Kanugula S, Pegg AE. Function of domains of human O^6-alkylguanine-DNA alkyltransferase. Biochemistry 2005;44:15396–405.
34. Zang EA, Wynder EL. Differences in lung cancer risk between men and women: examination of the evidence. J Natl Cancer Inst 1996;88:183–92.
35. Wei Q, Matanoski GM, Farmer ER, Hedayati MA, Grossman L. DNA repair related to multiple skin cancers and drug use. Cancer Res 1994;54:437–40.