Predictive Biomarkers and Personalized Medicine

The A/G Allele of Rs16906252 Predicts for MGMT Methylation and Is Selectively Silenced in Premalignant Lesions from Smokers and in Lung Adenocarcinomas

Shuguang Leng1, Amanda M. Bernauer1, Chibo Hong4, Kieu C. Do1, Christin M. Yingling1, Kristina G. Flores2, Mathewos Tessema1, Carmen S. Tellez1, Randall P. Willink1, Elizabeth A. Burki1, Maria A. Picchi1, Christine A. Stidley2, Michael D. Prados4, Joseph F. Costello4, Frank D. Gilliland5, Richard E. Crowell2, and Steven A. Belinsky1

Abstract

**Purpose:** To address the association between sequence variants within the MGMT (O6-methylguanine-DNA methyltransferase) promoter–enhancer region and methylation of MGMT in premalignant lesions from smokers and lung adenocarcinomas, their biological effects on gene regulation, and targeting MGMT for therapy.

**Experimental Design:** Single nucleotide polymorphisms (SNP) identified through sequencing a 1.9 kb fragment 5′ of MGMT were examined in relation to MGMT methylation in 169 lung adenocarcinomas and 1,731 sputum samples from smokers. The effect of promoter haplotypes on MGMT expression was tested using a luciferase reporter assay and cDNA expression analysis along with allele-specific sequencing for methylation. The response of MGMT methylated lung cancer cell lines to the alkylating agent temozolomide (TMZ) was assessed.

**Results:** The A allele of rs16906252 and the haplotype containing this SNP were strongly associated with increased risk for MGMT methylation in adenocarcinomas (ORs ≥ 94). This association was observed to a lesser extent in sputum samples in both smoker cohorts. The A allele was selectively methylated in primary lung tumors and cell lines heterozygous for rs16906252. With the most common haplotype as the reference, a 20 to 41% reduction in promoter activity was seen for the haplotype carrying the A allele that correlated with lower MGMT expression. The sensitivity of lung cancer cell lines to TMZ was strongly correlated with levels of MGMT methylation and expression.

**Conclusions:** These studies provide strong evidence that the A allele of a MGMT promoter–enhancer SNP is a key determinant for MGMT methylation in lung carcinogenesis. Moreover, TMZ treatment may benefit a subset of lung cancer patients methylated for MGMT. *Clin Cancer Res; 17(7); 2014–23. ©2011 AACR.*

Introduction

The fact that detection of gene promoter methylation of tumor suppressor genes in sputum that contains exfoliated cells from premalignant lesions from smokers’ lungs is a promising marker for early stage lung cancer makes identifying factors that influence the propensity for this epigenetic process throughout the respiratory epithelium a high priority. Such knowledge could impact strategies for screening to improve risk models, and identify persons who would benefit most from chemoprevention. The genes and the pathways they modulate underlying the interindividual susceptibility to DNA methylation remain largely elusive. Accumulating evidence from our group and others implicates DNA damage as an important step in the acquisition of de novo methylation (1–4). This premise was supported *in vivo* through a community-based study in which we showed a significant association between DNA repair capacity and germ line sequence variants within DNA repair genes and gene promoter methylation in sputum from smokers (5). The impact of this finding for prevention was addressed by asking whether dietary nutrients and vitamin supplements that may influence DNA damage and repair would modify the extent of gene methylation in the aerodigestive tract. Significant protection against methylation was observed in subjects consuming a diet rich in leafy green vegetables, or folate, or taking a...
Translational Relevance

Epigenetic silencing of MGMT (O\(^6\)-methylguanine-DNA methyltransferase) by promoter methylation is a biomarker for identifying persons at high risk for lung cancer and for predicting therapeutic response to alkylating agents in glioblastoma. The A/G allele of rs16906252 in MGMT promoter–enhancer region predicts for gene methylation in sputum from cancer free smokers. This association was increased substantially in primary lung tumors supporting this variant as a key determinant for epigenetic silencing of MGMT. Importantly, the sensitivity of lung cancer cell lines to temozolomide (TMZ) was strongly correlated with levels of MGMT methylation and expression, suggesting TMZ treatment may benefit a subset of lung cancer patients methylated for MGMT, a hypothesis that warrants further preclinical studies prior to initiating clinical trials.

Multivitamin, supporting the concept that diet and supplements can affect reprogramming of the epigenome and thus have the potential to modify risk for lung cancer (6). Specific sequence patterns within gene promoters and embryonic targets of polycomb-repressive complex 2 are predictive for identifying specific genes methylated in many cancers, but cannot discriminate individual susceptibility for gene silencing (7–10). However, other local sequence features that are polymorphic among people may contribute to the susceptibility of a specific CpG island to becoming methylated (11). A genome-wide survey in normal tissues found differential DNA methylation and expression between alleles at multiple nonimprinted loci that was sequence dependent (12). An additional study assessing 16 CpG rich gene promoters of chromosome 21 in leukocyte DNA identified 4 genes with allele-specific methylation and expression, suggesting TMZ treatment may benefit a subset of lung cancer patients methylated for MGMT, a hypothesis that warrants further preclinical studies prior to initiating clinical trials.

Methylation of MGMT is also a promising biomarker for lung cancer detection and its silencing is prognostic for response of glioblastoma patients to the alkylating agent temozolomide (TMZ; refs.18 and 19). Transcriptional regulation of MGMT occurs largely in a 1.2 kb fragment 5’ of the gene which contains the first untranslated exon (20). A 59 bp enhancer required for efficient MGMT promoter function was identified between the exon 1 and intron 1 boundary along with a SNP (rs16906252) located at the boundary between the enhancer and exon 1 (20, 21). Colorectal cancer patients heterozygous or homozygous for the A allele (vs. G allele) of this SNP had a 32-fold increased odds for methylation of the MGMT gene in their tumors, a finding replicated in a second, independent study (21, 22).

The purpose of this study was to address the association between sequence variants within the MGMT promoter–enhancer region and methylation of MGMT in lung adenocarcinomas and exfoliated epithelial cells within sputum from smokers, and their biological effects on gene regulation. Genetic and epigenetic variations that affect MGMT expression may impact personalized therapy for lung cancer. Therefore, the potential of TMZ as a therapeutic option for lung cancer was evaluated in lung cancer cell lines with different levels of methylation and expression of MGMT.

Materials and Methods

Study populations and sample collection

The Lovelace Smokers Cohort (LSC) began recruitment of female smokers in 2001 and expanded to include male smokers in 2004 (5). Enrollment is restricted to current and former smokers age 40 to 74 years with a minimum of 15 pack-years of smoking. At study entry, all cohort members completed a detailed questionnaire that collected information on medical, smoking, exposure history, socioeconomic status, and quality of life. Pulmonary function test was tested to assess lung function, blood was drawn for collection of plasma and lymphocytes, and sputum was collected by induction. The Veterans Smokers Cohort (VSC) began recruitment of smokers in 2000. Similar enrollment criteria and procedures as used in the LSC were followed. Most participants in the VSC are males and have over 20 pack-years smoking history. In total, 1,731 lung cancer-free subjects (1,256 subjects from LSC and 475 subjects from VSC) were included in this study.

Lung adenocarcinomas (n = 169) were obtained from the tumor banks at Johns Hopkins and the Mayo Clinic (23). Demographic characteristics of the patients have been described (23). Stage I lung cancer patients (n = 98) enrolled in the New Mexico lung cancer cohort were also used in this study because of the availability of frozen tissue for assessing MGMT expression. A detailed description of the enrollment and characteristics of the cohort members has been provided previously (24). All samples were collected with informed consent and the institutional review boards of the Lovelace Respiratory Research Institute (Western IRB) and New Mexico VA Health Care System.
approved all investigations using human tissues and clinical data.

Cell lines
Normal human bronchial epithelial cells (NHBECS) isolated from bronchoscopy of cancer-free smokers (n = 20), 19 lung cancer-derived cell lines (A549, Calu-3, Calu-6, HCC406, HCC827-1, H23, H358, H460, H522, H1299, H1435, H1568, H1795, H1993, H2009, H2023, H2085, H441, and H2228), 1 immortalized normal human bronchial epithelial cell line (HBEC2), and 2 glioblastoma cell lines (U251 and T98G) were used in this study. HBEC2 was received from Drs. Shay and Minna, Southwestern Medical Center, Dallas, TX. All the other cell lines were from American Type Culture Collection. Cells were maintained in recommended media in a 37°C incubator supplied with 5% CO2. All experiments were conducted using cell lines which were passaged in the laboratory for fewer than 6 months after resuscitation.

Sputum processing and MGMT promoter methylation
A detailed procedure for sputum processing has been introduced (18). Briefly, sputum samples were stored in Saccomanno’s fixative. Sputum adequacy defined as the presence of deep lung macrophages or Curschmann’s spiral (18) was assessed by cytotechnologists and subjects with cancer cells detected in sputum slides were removed from this study. An aliquot from sputum samples for each study subject was used for DNA isolation. Promoter methylation of MGMT was studied in NHBECS and cell lines using combined bisulfite modification and restriction analysis (COBRA) and methylation-specific PCR (MSP) as described (25). Promoter methylation of MGMT in primary adenocarcinoma and sputum samples was evaluated using MSP and nested MSP, respectively (25). The COBRA and MSP assays were designed to interrogate the methylation status of CpGs around the transcriptional start site of MGMT. COBRA and MSP primer sequences are listed in Supplementary Table S1, respectively.

SNP discovery and haplotype-tagging SNPs selection
A 1,871 bp fragment of the MGMT gene that contains the well-characterized maximal promoter, the entire CpG island, and part of intron 1 was carefully selected for SNP discovery because several major regulatory elements for MGMT transcription were identified to locate in this region (Supplementary Figure; refs. 20 and 26). Primers are listed in Supplementary Table S1. DNA samples from 49 cancer-free non-Hispanic Whites randomly selected from the LSC were used for discovery of SNPs. An additional 11 samples were selected for sequencing based on their rs16906252 genotypes (3 A/Gs and 8 A/As) to improve the probability of finding a SNP in perfect linkage disequilibrium (LD) with rs16906252. PCR products were gel-purified and were commercially sequenced to identify variants (Sequetech). Pairwise $r^2$ between any 2 SNPs was calculated to identify a SNP in perfect LD with rs16906252 ($r^2 = 1$). A Bayesian statistical method implemented in the program PHASE (Version 2.1) was used to construct the haplotypes using the variants discovered in the 60 subjects (27, 28). Haplotype-tagging SNPs (htSNP) were chosen using the TagSNPs program (29) based on the haplotype alleles and frequencies generated by the PHASE program in 49 randomly selected samples. In this process, the squared correlation ($R^2$) between the true haplotypes (h) and their estimates was calculated. An expectation maximization algorithm approach was then implemented to find the minimum set of SNPs (within the promoter–enhancer region of MGMT) that would have $R^2 > 0.80$ for all common haplotypes.

SNP genotyping
Genotyping was conducted in DNA samples isolated from peripheral lymphocytes in subjects from LSC and VSC. DNA isolated from tumor tissues was used for genotyping in 169 lung adenocarcinomas because no blood was collected from these patients. It is highly unlikely that DNA recovered from tumor tissue will generate false calls for germ line SNPs because stromal and inflammatory cells also comprise the tumor specimen. Primers amenable to detecting a PCR-restriction fragment length polymorphism (RFLP) were designed to genotype the SNPs in DNA samples isolated from the lung adenocarcinomas (n = 169), subjects (n = 475) from the VSC, NHBECS (n = 20), and cell lines (n = 22). The primers and restriction digestion enzymes are listed in Supplementary Table S1. The 60 DNA samples used for SNP discovery were regenotyped by PCR-RFLP and 100% concordance of genotypes was observed between this method and sequencing. Two DNA samples with wild homozygote and variant homozygote were included in each batch of PCR-RFLP assay for quality control. The Illumina Golden Gate Assay was used to genotype rs16906252 in subjects from the LSC (n = 1,256). Intraplate and interplate duplicates and 1 parent–child trio were included in each 96-well plate for quality control.

Selection of subjects and construction of MGMT promoter constructs
A 1,528-bp fragment covering the entire maximal promoter and CpG island of MGMT with different haplotypes was amplified by PCR from homozygotes, or heterozygotes when a homozygote for a haplotype was not present in the 60 subjects used in discovery of the SNPs (Supplementary Figure). The primers are listed in Supplementary Table S1. The promoter–enhancer fragment was directionally cloned into the pGL2-basic Luciferase Reporter Vector (Promega) upstream of the luciferase coding sequence. Four to 6 clones from each person were commercially sequenced to identify the correct construct with each specific haplotype (Sequetech).

Transient transfection and reporter gene assays
Transient transfections were conducted in Calu-6, H1299, and HBEC2. Cells (0.8–1.2 × 10⁶) were plated into 6-well plates and transfected the following day.

2016 Clin Cancer Res; 17(7) April 1, 2011 Clinical Cancer Research
Plasmid DNA (1 μg) and the pSV-β-Galactosidase control vector (0.2 μg; Promega) were cotransfected into cells with Lipofectamine 2000 transfection reagent (Invitrogen) at a Lipofectamine 2000 DNA ratio of 3:1. A promoter-less pGLO2-basic vector and the pGLO2-control vector that contains the SV40 promoter were used as a negative and positive control, respectively. Forty-eight hours after transfection, cells were harvested and lysed. Immediately after lysing, cell extracts were assayed in the Lumoskan Ascent luminometer (Thermo Electron) for luciferase activity using the Luciferase Assay System (Promega). β-galactosidase activity in cell lysates was measured using the GalactoStar Reporter Gene Assay System (Tropix). Promoter activity was calculated as the ratio of activities of luciferase and β-galactosidase. Transfections were done in triplicate in 3 independent experiments.

Cytotoxicity of TMZ in tumor cell lines

The stock TMZ (Sigma-Aldrich) solution (129 mmol/L) was prepared by dissolving 100 mg TMZ in 4 mL dimethyl sulfoxide and was frozen at −80°C in aliquots. Before treatment, TMZ stock was thawed and added to the culture medium at a series of final concentrations (0, 10, 30, 90, 270, 810, and 2,500 μmol/L). Cells (5 × 10^3/well) were seeded into 96-well plates and were incubated with TMZ in the culture medium for 72 hours. The MTT assay measured cytotoxicity and an IC_{50} (the half maximal inhibitory concentration) was calculated. Three independent experiments were conducted for each cell line and the average IC_{50} was presented.

Real-time PCR for measuring MGMT expression

Cells were harvested in TRI reagent (Sigma-Aldrich) and mRNA was isolated following TRI reagent instructions. cDNA was synthesized using the high capacity cDNA reverse transcription kit (Applied Biosystems). A real-time PCR using SYBR Green as the fluorescent dye was developed to measure the MGMT expression level using the ACT method. The primers are listed in Supplementary Table S1.

Statistical analysis

The two-sample t-test, Wilcoxon rank sum test, and χ² test were used to compare demographic variables and genotypes between MGMT methylated and unmethylated groups as appropriate. Hardy-Weinberg equilibrium (HWE) was tested for each SNP in the LSC, VSC, and the lung adenocarcinomas separately and no deviation from HWE was identified (not shown). The minor allele frequency (MAF) for rs16906252 is 7.6% in the 1,900 samples. This study is well powered to investigate an association between the A/A genotype and the 4 common haplotypes in 49 subjects. Under the additive inheritance model that assumes that the effect of the A/A genotype is twice that of the G/A genotype, we have 99% power to detect an OR of 2.3, 1.9, and 1.6 in lung adenocarcinoma samples, VA cohort, and Lovelace cohort, respectively (30). A logistic regression model was used to calculate the ORs and 95% CIs for each individual SNP with adjustment for age, sex, ethnicity, and smoking history selected a priori.

The PHASE program was used to reconstruct the haplotypes and calculate their estimated probabilities from the htSNPs data in the 5’ region of MGMT in lung adenocarcinomas (27, 28). All possible haplotypes with corresponding estimated probabilities (>0.01) for each individual were generated. The probabilities of the common haplotypes for each individual were used as explanatory variables in the logistic regression model with adjustment for non-genetic factors to assess the association between the haplotypes and the risk for MGMT methylation. Haplotypes with frequency less than 5% were combined into 1 group. The effect for each common haplotype was estimated using the most common haplotype as the reference group. The effects for the common haplotypes were also analyzed using all other haplotypes as the reference group and similar results were obtained. All statistical analyses were conducted in Statistical Analysis System 9.2.

Results

Association of SNPs to MGMT methylation

Several SNPs were discovered in the promoter–enhancer region of MGMT in a European population; however, the LD between those SNPs has not been characterized (31). Sequencing of a 1.9 kb fragment of MGMT identified 9 SNPs with 1 SNP detected for the first time (Supplementary Table S2). No 2 SNPs were in perfect LD (Fig. 1). The r² value between any 2 SNPs in this area was 0.52 or less.
except for rs1625649 and rs12356501 ($r^2 = 0.71$). Four SNPs had MAF greater than 0.04. A total of 4 haplotypes with frequency greater than 0.04 were constructed based on the 4 common SNPs. Three SNPs including rs1711646, rs1625649, and rs16906252 were selected as the htSNPs (29) that accounted for 87% of the haplotype variation in the promoter–enhancer region of *MGMT*.

The association between these 3 htSNPs and *MGMT* methylation was first tested in 169 lung adenocarcinomas. Rs1625649 was modestly associated with *MGMT* methylation (Table 1). In contrast, a striking and highly significant association was seen between rs16906252 and increased risk for *MGMT* methylation (OR = 97.0, $P < 0.0001$). Only rs16906252 remained significant after Bonferroni correction. The association between the genetic variation of the entire promoter–enhancer region and *MGMT* methylation was then analyzed using a haplotype-based approach that integrated the genetic information of multiple SNPs. A total of 7 haplotypes were reconstructed based on the genotype for the 3 htSNPs including rs1711646, rs1625649, and rs16906252 in the 169 lung adenocarcinomas (27, 28). Four of the 7 haplotypes including HAP1 (GCG), HAP3 (GAG), HAP4 (GAA), and HAP7 (TCG) had frequency greater than 5% with their cumulative frequency of 94% (Supplementary Table S3). Haplotype analysis found that only HAP4 defined solely by rs16906252 was associated with *MGMT* methylation (Supplementary Table S3).

*MGMT* is one of a panel of genes whose concomitant methylation detected in sputum collected from moderate and heavy smokers associates with increased risk for lung cancer in a high-risk cohort (18, 32). Therefore, the association between rs16906252 and *MGMT* methylation was assessed in sputum samples collected in 2 smoker cohorts. The variant homozygotes and heterozygotes were associated with 12.4- and 6.3-fold increased risk for *MGMT* methylation, respectively in the LSC (Table 1). The strong association between rs16906252 and *MGMT* methylation was also replicated in smokers from the VSC (Table 1).

The commonality for promoter sequence variants to affect other frequently methylated genes was addressed by studying the *p16* gene that is silenced by promoter methylation in greater than 50% of lung tumors (25). Only 1 common (MAF > 0.05) SNP (rs3814960) was identified within the *p16* basal promoter and CpG island by searching the Environmental Genome Project database (http://egp.gs.washington.edu/). No significant association ($P > 0.10$) was found between rs3814960 and *p16* methylation in adenocarcinomas or sputum samples from the LSC indicating that not all genes are silenced by sequence dependent ASM.

### Table 1. Association between SNPs and risk for *MGMT* methylation in lung adenocarcinoma and in sputum from smokers

<table>
<thead>
<tr>
<th>Samples/cohorts (n)</th>
<th>SNPs</th>
<th>Genotypes</th>
<th>Methylated/unmethylated</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung adenocarcinoma (n = 169)</td>
<td>rs1711646&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G/G</td>
<td>39/66</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G/T</td>
<td>15/35</td>
<td>0.7 (0.3–1.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T/T</td>
<td>5/9</td>
<td>0.9 (0.2–3.2)</td>
</tr>
<tr>
<td></td>
<td>rs1625649&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C/C</td>
<td>17/56</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C/A</td>
<td>32/42</td>
<td>2.4 (1.1–4.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A/A</td>
<td>10/12</td>
<td>2.7 (1.0–7.6)</td>
</tr>
<tr>
<td></td>
<td>rs16906252&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>G/G</td>
<td>36/109</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G/A</td>
<td>23/1</td>
<td>97.0 (12.1–778.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A/A</td>
<td>Not detected</td>
<td>–</td>
</tr>
<tr>
<td>Sputum LSC (n = 1,256)</td>
<td>rs16906252&lt;sup&gt;c&lt;/sup&gt;</td>
<td>G/G</td>
<td>207/853</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G/A</td>
<td>113/75</td>
<td>6.3 (4.5–8.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A/A</td>
<td>6/2</td>
<td>12.4 (2.5–62.8)</td>
</tr>
<tr>
<td>VSC (n = 475)</td>
<td>rs16906252&lt;sup&gt;d&lt;/sup&gt;</td>
<td>G/G</td>
<td>82/332</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G/A</td>
<td>39/21</td>
<td>8.0 (4.4–14.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A/A</td>
<td>1/0</td>
<td>Not calculated&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Age, gender, race (white vs. nonwhite), and smoking status (ever vs. never) were used as adjustment factors in logistic regression.

<sup>b</sup>Due to small sample size in one cell, exact logistic regression was used.

<sup>c</sup>Age, gender, race, current smoking status, and pack years were used as adjustment factors in logistic regression in the LSC.

<sup>d</sup>Age, gender, race, current smoking status, and pack years were used as adjustment factors in logistic regression in VA smokers cohort.

<sup>e</sup>OR for A/A genotype in the VA cohort was not calculated because no A/A genotype was identified in the unmethylated group.

Influence of the A/G genotype on *MGMT* transcription, expression, and methylation

The effect of sequence variants in the promoter–enhancer region on transcription of *MGMT* was assessed in HBEC2, Calu-6, and H1299. The lowest promoter activity...
was seen in HAP4 carrying the A allele. With the most common haplotype (HAP1) as the reference, a 20 to 40% reduction in promoter activity in 3 lung cell lines compared with HAP1. Values are mean ± SD from 3 experiments. B, a 153 bp fragment surrounding the rs16906252 was amplified from cDNA isolated from these 2 NHBECs and then digested by FspI. Densitometry analysis showed that the expression of the A allele was reduced 31 to 39% compared with the G allele, whereas rs16906252 dependent methylation of MGMT leads to selective loss of A allele expression in Calu-6 and H1993. As expected, expression of only the G allele is seen in cell lines (HCCC4006, H23, and H2009) with G/G genotype that are unmethylated for MGMT.

**Figure 2.** The A allele confers reduced transcription and expression of MGMT. A, haplotype 4 (HAP4) carrying the A allele of rs16906252 shows a 20 to 40% reduction in promoter activity in 3 lung cell lines compared with HAP1. Values are mean ± SD from 3 experiments. B, a 153 bp fragment surrounding the rs16906252 was amplified from cDNA isolated from these 2 NHBECs and then digested by FspI. Densitometry analysis showed that the expression of the A allele was reduced 31 to 39% compared with the G allele, whereas rs16906252 dependent methylation of MGMT leads to selective loss of A allele expression in Calu-6 and H1993. As expected, expression of only the G allele is seen in cell lines (HCCC4006, H23, and H2009) with G/G genotype that are unmethylated for MGMT.

Reduced gene transcription has been proposed as one mechanism predisposing a gene for epigenetic silencing via promoter methylation (35). Therefore, studies were conducted to test whether cell lines with the A/G genotype have preferential methylation of the A allele that in turn should lead to monoallelic expression of MGMT. Calu-6 and H1993 were heterozygous for rs16906252 and hemimethylated for MGMT. Bisulfite sequencing found that only the A allele is densely methylated and selectively silenced in Calu-6 and H1993 (Figs. 2B and 3A; not shown). The ASM of MGMT associated with rs16906252 was further confirmed in primary lung tumors. An MSP assay targeting the reverse strand of the MGMT promoter that distinguishes the G versus the A allele was developed. Six A/G tumors methylated for MGMT were sequenced and only the A allele was methylated (Fig. 3B).

**Targeting MGMT for lung cancer therapy**

TMZ is an alkylating chemotherapeutic agent that does not cause severe systemic toxicity and significantly extends survival in glioblastoma patients with methylation of MGMT (19). If a similar response to TMZ dependent on MGMT methylation status could be replicated in lung cancer, this agent could potentially benefit a subset of patients. This premise was tested by determining if lung tumor cell lines with differential MGMT methylation and expression would have different sensitivity toward TMZ treatment. Expression of MGMT was quantified in 9 G/G
lung cancer cell lines with unmethylated MGMT and compared to the average expression seen in 6 G/G NHBEC lines. Two unmethylated cell lines with high MGMT expression were selected to serve as controls (Table 2). All 3 of the G/G cancer cell lines with MGMT methylation showed hemimethylation based on the COBRA assay (not shown). Two G/G methylated cell lines and 1 A/G cell line were selected and expression levels of MGMT were quantified. Two cell lines derived from glioblastoma, 1 G/G fully methylated for MGMT showing no expression (U251) and the second (T98G) G/G was sparsely methylated showing no reduction in expression compared to NHBECs were included for comparison of sensitivity to TMZ. The IC50 for TMZ was similar in the unmethylated glioblastoma and lung cancer cell lines (1,590–2,500 μmol/L). As expected the fully methylated glioblastoma cell line showed heightened sensitivity to TMZ with an 83% reduction in IC50 compared to unmethylated cell lines (Table 2). The response of the hemimethylated lung cancer cell lines was intermediate between fully methylated and unmethylated tumor cell lines. A high correlation between TMZ IC50s and MGMT expression was observed for the lung cancer cell lines (r = 0.99; Table 2 and Fig. 4).

Discussion

This study provides strong evidence that the A allele of a MGMT promoter–enhancer SNP (rs16906252) is a key determinant in the acquisition of MGMT methylation in lung carcinogenesis. This association was found in primary adenocarcinoma and in exfoliated cells in sputum from cancer-free smokers, a finding that reflects the ongoing field cancerization within the aerodigestive tract. The magnitude of the association between rs16906252 and MGMT methylation in premalignant lesions from lung cancer-free smokers was much smaller than observed from the lung adenocarcinoma patients, consistent with the increase in prevalence of MGMT methylation of 25% in premalignant lesions from smokers to 43% in stage I lung cancer and 61% in stage II to IV lung cancer (36). The reduced transcription and expression seen for the A allele of MGMT may be one predisposing factor for methylation. This hypothesis is supported by the selective methylation and accompanied silencing of the A allele in tumors from A/G heterozygotes. Thus, some of the MGMT unmethylated lung cancer-free smokers who carry the A allele should be more prone to become methylated as their field cancerization progresses over time (32).

The lack of an association between rs16906252 with risk for lung cancer based on the similar percentage of subjects carrying the A allele in lung adenocarcinoma cases and cancer-free controls in the Lovelace and Veteran Cohorts (14.2%, 15.6%, and 12.8%, respectively) is not surprising. Our previous studies demonstrate that concomitant methylation of a panel of tumor suppressor genes including MGMT in sputum could prospectively predict risk for lung cancer in moderate and heavy smokers (18, 32). MGMT methylation alone in sputum only contributes to a portion of the predicted increased cancer risk. Thus, a large case-control study would be needed to more precisely estimate the association of this SNP to risk for lung cancer. Moreover, approximately 20% of subjects with G/G genotype also have MGMT methylated in sputum, suggesting that other genetic or environmental factors may also contribute to MGMT methylation in lung carcinogenesis, albeit probably to a lesser degree than rs16906252. Thus, the prediction for lung cancer risk by MGMT methylation in sputum could also vary by rs16906252 genotype. Testing this hypothesis will require long-term followup in our smoker cohorts to accumulate incident lung cancer cases with sputum samples collected before cancer diagnosis.

The sequence dependent ASM of MGMT is not lung tumor-specific because similar associations were also reported in colorectal cancer (21, 22). Studies in...
glioblastoma did not identify an association between rs16906252 and risk for methylation (37). However, the primers used for genotyping had several incorrect bases that likely influenced the accuracy for identifying the variant allele (37). Recently, methylation of \textit{GSTP1}, \textit{RIL}, and \textit{MSH2} were found to be allele-specific in several tumor types (11, 38, 39). Low-level methylation of \textit{MGMT} was also seen in normal colon mucosa and peripheral lymphocytes from heterozygote subjects, corroborating a likely important role for this SNP in the susceptibility of the A allele of \textit{MGMT} for silencing (22, 40). Zhang and colleagues studying methylation of gene promoters on chromosome 21 estimated that ASM may affect 10% of human genes in leukocytes (13). Thus, genetic differences can influence interindividual variation for DNA methylation of many genes in normal tissues and tumors.

Evidence is accumulating that reduced transcription may be one trigger for initiation and spreading of heterochromatin and methylation along a gene promoter (34). Song and colleagues (41) and Stirzaker and colleagues (42) showed that reduced transcription and seeds of methylation within the \textit{GSTP1} promoter served as a stimulus for the spread of methylation and chromatin remodeling. \textit{In vivo} studies found that breast cancer patients carrying the \textit{GSTP1} promoter haplotype E comprised of 4 SNPs and the shortest ATAAA repeat had increased methylation (38). The polymorphisms within haplotype E disrupted binding of c-Myb leading to reduced transcription. Our studies also support reduced transcription as a mechanism for \textit{de novo} promoter methylation based on the strong association between reduced promoter activity associated with the A allele in \textit{NHBEC} and the targeting of this allele for silencing in lung tumors from heterozygote patients. Other mechanisms independent of effects on gene transcription may also influence \textit{de novo} promoter methylation. For example, a 12 bp insertion polymorphism in the \textit{RIL} promoter creates an Sp1/Sp3 binding site that protects against methylation in cancer and does not change expression in lymphoblastoid cell lines (11). The Sp1/Sp3 binding site created by this polymorphism may through protein-DNA interactions prevent spreading of DNA methylation and heterochromatin from neighboring “methylation centers.” \textit{In silico} prediction of transcription factor-binding sites using TFSEARCH and MATCH programs did not identify any transcription factors with core motifs binding rs16906252. In addition, resequencing the 1.9 Kb 5’ region of the \textit{MGMT} gene did not identify any SNP in perfect LD with rs16906252. Therefore, rs16906252 may influence the binding of an unknown transcription factor or unlisted transcription factor in these 2 databases. Recently, Hawkins and colleagues examined 6 gene-associated CpG islands flanking \textit{MGMT} in colorectal tumors and found that the \textit{EBF3} gene located 500 kb downstream of \textit{MGMT} had a similar methylation pattern to \textit{MGMT} (22). Thus, it is possible that rs16906252 is in perfect LD with other polymorphisms present in the euchromatin and

### Table 2. \textit{MGMT} methylation status, mRNA expression, rs16906252 genotype, and TMZ IC50s

<table>
<thead>
<tr>
<th>Cell line</th>
<th>\textit{MGMT} methylation(^a)</th>
<th>Genotype</th>
<th>% Expression(^b)</th>
<th>TMZ IC50(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCC4006</td>
<td>Unmethylated</td>
<td>G/G</td>
<td>162</td>
<td>1,655</td>
</tr>
<tr>
<td>H2085</td>
<td>Unmethylated</td>
<td>G/G</td>
<td>232</td>
<td>2,500</td>
</tr>
<tr>
<td>Calu-6</td>
<td>Homemethylated</td>
<td>G/A</td>
<td>50</td>
<td>680</td>
</tr>
<tr>
<td>Calu3</td>
<td>Homemethylated</td>
<td>G/G</td>
<td>64</td>
<td>680</td>
</tr>
<tr>
<td>A549</td>
<td>Homemethylated</td>
<td>G/G</td>
<td>70</td>
<td>500</td>
</tr>
<tr>
<td>Brain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U251</td>
<td>Fully methylated</td>
<td>G/G</td>
<td>0</td>
<td>270</td>
</tr>
<tr>
<td>T98G</td>
<td>Sparsely methylated</td>
<td>G/G</td>
<td>100</td>
<td>1,590</td>
</tr>
</tbody>
</table>

\(^a\)\textit{MGMT} methylation status was measured by COBRA.

\(^b\)% expression was the average from 3 cDNAs independently made from 1 mRNA sample and was relative to the average expression seen in the \textit{NHBEC} lines.

\(^c\)TMZ IC50 was the average from 3 independent experiments.

---

**Figure 4.** \textit{MGMT} expression and TMZ IC50s in 5 lung tumor cell lines (\textbullet) is highly correlated (\(r = 0.99\)).
heterochromatin boundaries surrounding the MGMT and/or EBF3 which could affect gene regulation. Currently, testing this hypothesis will necessitate the completion of the 1,000 Genome Project that will provide a dense map of genetic variation by sequencing the genomes of 1,200 people (43).

The refractory nature of unresected lung cancer to chemotherapeutic combinations tested over 3 decades necessitates developing targeted therapies based on underlying genetic and epigenetic tumor profiles. Responses have been extremely positive for patients with mutations in the epidermal growth factor receptor receiving the tyrosine kinase inhibitors gefitinib or erlotinib (44). Alkylating agents such as 1,3-bis-(2-chloroethyl)-1-nitrosourea were tested as chemotherapeutics for lung cancer decades ago with response rates limited by systemic toxicity (35, 45, 46). TMZ has been evaluated in a phase II study for efficacy in pretreated lung cancer, melanoma, and breast cancer patients with brain metastasis. Treatment was provided without knowledge of MGMT methylation status, and 26% of patients had a clinical response with median overall survival greatest for lung cancer patients (47). Our in vitro studies in lung cancer cell lines demonstrating sensitivity to TMZ that is highly correlated with levels of MGMT methylation suggest that a subset of lung cancer patients with relatively low MGMT expression could respond to this drug, a hypothesis that requires further preclinical studies prior to initiating clinical trials.

Disclosure of Potential Conflicts of Interest

S.A. Belinsky is a consultant to Oncomelhyome Sciences. Under a licensing agreement between Lovelace Respiratory Research Institute and Oncomelhyome Sciences, the author is entitled to a share of the royalties received by the Institute from sales of the licensed technology. The Institute, in accordance with its conflict-of-interest policies, is managing the terms of these arrangements.

Grant Support

This work was supported by National Cancer Institute (R01 CA097356 and R01 ES008801 to S.A. Belinsky and I01CA128623-01 to K.G. Flores) and the State of New Mexico as a direct appropriation from the Tobacco Settlement Fund to S.A. Belinsky.

Received November 11, 2010; revised January 13, 2011; accepted February 4, 2011; published OnlineFirst February 25, 2011.

References

Selective Silencing of the A/G Allele of MGMT in Lung Cancer


The A/G Allele of Rs16906252 Predicts for MGMT Methylation and Is Selectively Silenced in Premalignant Lesions from Smokers and in Lung Adenocarcinomas

Shuguang Leng, Amanda M. Bernauer, Chibo Hong, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-10-3026

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2011/03/31/1078-0432.CCR-10-3026.DC1

Cited articles
This article cites 47 articles, 21 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/17/7/2014.full.html#ref-list-1

Citing articles
This article has been cited by 17 HighWire-hosted articles. Access the articles at:
/content/17/7/2014.full.html#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.