Implication of a Chromosome 15q15.2 Locus in Regulating UBR1 and Predisposing Smokers to MGMT Methylation in Lung

Shuguang Leng1, Guodong Wu1, Leonard B. Collins2, Cynthia L. Thomas1, Carmen S. Tellez1, Andrew R. Jauregui3, Maria A. Picchi1, Xiequn Zhang1, Daniel E. Juri1, Dhimant Desai4, Shantu G. Amin4, Richard E. Crowell5, Christine A. Stidley5, Yushi Liu1, James A. Swenberg2, Yong Lin1, Marc G. Wathelet3, Frank D. Gilliland6, and Steven A. Belinsky1

Abstract

O6-Methylguanine-DNA methyltransferase (MGMT) is a DNA repair enzyme that protects cells from carcinogenic effects of alkylating agents; however, MGMT is silenced by promoter hypermethylation during carcinogenesis. A single-nucleotide polymorphism (SNP) in an enhancer in the MGMT promoter was previously identified to be highly significantly associated with risk for MGMT methylation in lung cancer and sputum from smokers. To further genetic investigations, a genome-wide association and replication study was conducted in two smoker cohorts to identify novel loci for MGMT methylation in sputum that were independent of the MGMT enhancer polymorphism. Two novel trans-acting loci (15q15.2 and 17q24.3) that were identified acted together with the enhancer SNP to empower risk prediction for MGMT methylation. We found that the predisposition to MGMT methylation arising from the 15q15.2 locus involved regulation of the ubiquitin protein ligase E3 component UBR1. UBR1 attenuation reduced turnover of MGMT protein and increased repair of O6-methylguanine in nitrosomethylurea-treated human bronchial epithelial cells, while also reducing MGMT promoter activity and abolishing MGMT induction. Overall, our results substantiate reduced gene transcription as a major mechanism for predisposition to MGMT methylation in the lungs of smokers, and support the importance of UBR1 in regulating MGMT homeostasis and DNA repair of alkylated DNA adducts in cells. Cancer Res; 75(15): 1–10. ©2015 AACR.

Introduction

O6-Methylguanine-DNA methyltransferase (MGMT) is a DNA repair enzyme that protects cells from carcinogenic effects of alkylating agents by removing adducts from the O6 position of guanine (1). This repair mechanism is conserved in all organisms and involves transfer of methyl and other alkyl groups from the O6 position of guanine in DNA to a cysteine residue within enzymatic active site of MGMT protein. The irreversible binding of the alkyl group to the MGMT protein functionally inactivates the enzyme and leads to a structural alteration resulting in the recognition by ubiquitin ligases and degradation by the proteasome (1, 2). The recovery of MGMT activity after its inactivation with alkylating agents is slow and results entirely from de novo protein synthesis (3–5).

The predominant mechanism for MGMT gene inactivation in carcinogenesis is hypermethylation of a CpG island within its promoter–enhancer region that leads to transcriptional silencing in lymphoma and several solid tumors, including brain, colon, lung, and head and neck cancers (6). Silencing of this gene is associated with an increased prevalence for G-C to A-T transition mutations in K-ras and p53 genes in colorectal and brain tumors respectively, consistent with the important role for this gene in removing the highly promutagenic adduct O6-methylguanine (O6MG; refs. 7 and 8). Although alkylating adducts such as O6MG can be derived from endogenous methylation by S-adenosylmethionine, knockouts of MGMT failed to show enhanced tumor formation unless treated with exogenous alkylating carcinogens (9). MGMT methylation is a common event in lung adenocarcinomas and associated with tumor progression (10). Concomitant methylation of a panel of cancer-relevant genes, including MGMT, detected in lung epithelial cells present in sputum predicts risk for subsequent lung cancer incidence in moderate and heavy smokers (11, 12). In addition, MGMT methylation is also a prognostic biomarker for response of glioblastoma patients to the alkylating agent temozolomide (13, 14).

Genome-wide profiling of allele-specific methylation (ASM) using a methylation-sensitive single nucleotide polymorphism (SNP) array identified >35,000 ASM sites across the genomes of multiple tissue types that were likely mediated by parental-origin
of short distance cis-acting SNPs (15–18). This effort was further expanded by running methylation analysis of cis-acting SNPs in cancer tissues. We included lung, adipose, peripheral blood lymphocyte, and brain tumors (19–23). In normal lung, 34,304 cis- and trans-acting germine meQTL were identified and enriched in CTCCT-binding factor (CTCF)-binding sites, DNaI hypersensitivity regions, and histone marks (19). Four of five established lung cancer risk loci among persons with European ancestry are cis-acting meQTL (19).

GWAS in normal lung tissues were performed using ENCODE-annotated regulatory regions (19). Four of these regions at chromosome 15q15.2 were enriched in the genomic loci that act as imprinting or short distance cis-acting SNPs (15–18). This effort was further expanded by running methylation analysis of cis-acting SNPs in cancer tissues. We included lung, adipose, peripheral blood lymphocyte, and brain tumors (19–23). In normal lung, 34,304 cis- and trans-acting germine meQTL were identified and enriched in CTCCT-binding factor (CTCF)-binding sites, DNaI hypersensitivity regions, and histone marks (19). Four of five established lung cancer risk loci among persons with European ancestry are cis-acting meQTL (19).

In addition to germine meQTL, somatic meQTL were also identified that affected the predisposition of cancer-relevant genes for promoter hypermethylation acquired in carcinogenesis for multiple tumor types (13, 26–32). This is illustrated by our studies and others for MGMT methylation. An enhancer SNP rs16906252 located at the first exon–intron boundary of MGMT was highly significantly associated with risk for MGMT methylation in colorectal cancer, pleural mesothelioma, lung adenocarcinoma, glioblastoma, and diffuse large B-cell lymphoma (13, 27, 29–32). Significant association was also seen for MGMT methylation in sputum from smokers (27). Functional characterization confirmed the AS pattern in heterozygotes for rs16906252 with the “A” allele preferentially silenced by methylation in primary lung tumors and lung cancer cell lines (27). Although rs16906252 is a major determinant for acquisition of MGMT methylation in lung carcinogenesis, 20% of current and former smokers carrying wild homozygote (G/G) of rs16906252 are methylated for MGMT in their sputum (27), suggesting that other genetic or environmental factors may contribute to MGMT methylation. To address this issue, a GWAS discovery and replication study was conducted in two smoker cohorts from New Mexico to identify novel loci whose association with risk for MGMT methylation detected in sputum was independent of the enhancer SNP. Subsequent hypothesis-driven studies were based on the genetic associations detected to explore the underlying biological mechanisms.

Materials and Methods

Study populations

Longitudinal cohorts that enrolled current and former smokers were used for the GWAS discovery (the Lovelace smokers cohort, LSC) and the replication (the veterans smokers cohort, VSC). The LSC began recruitment of female smokers in 2001 and expanded to include male smokers in 2004. Enrollment is restricted to current and former smokers ages 40 to 74 years with a minimum of 15 pack-years of smoking. The VSC began recruitment of smokers in 2000 using enrollment criteria and procedures similar to the LSC. Most participants in the VSC are males with greater than 20 pack-years smoking history. In total, 1,675 lung cancer-free subjects (1,200 subjects from LSC and 430 subjects from VSC) were included in this study. A detailed description of subject enrollment and follow has been provided previously (27). All samples were collected with informed consent. The study was approved by the Western Institutional Review Board and New Mexico VA Health Care System.

Sputum collection and measurement of MGMT promoter methylation

Induced and spontaneous morning sputum samples were collected in LSC and VSC, respectively. Sputum samples were washed multiple times using Saccamanno’s fixative prior to long-term storage at −80°C. Sputum adequacy defined as the presence of deep lung macrophages or Curschmann’s spiral (12) was assessed by cytotechnologists and subjects with cancer cells detected in sputum slides were removed from this study. Sputum DNA was isolated and bisulfite converted. Methylation status for CpGs around the transcriptional start site of MGMT was assessed using two stage methylation-specific PCR (33). Methylation status was scored as 0 (unmethylated) or 1 (methylated). Our assay can reproducibly detect one methylated allele in a background of 10,000 unmethylated alleles (12).

GWAS genotyping and quality assurance

The HumanOmni2.5-4v1-H BeadChip (Illumina) was used to genotype 2,450,000 SNPs in 1,200 Caucasian smokers from the LSC. We removed 37 subjects due to low call rate (<95%, n = 7), low heterozygosity (n = 1), low Caucasian ancestry (<85%, n = 2), and high relatedness with other samples (n = 27). Furthermore, SNPs were excluded if they had a call rate of <90%, a minor allele frequency (MAF) <0.008, or P < 10−8 for Hardy–Weinberg equilibrium test, or were on Y or pseudo-autosomal region of X chromosomes. The MAF cutoff is a technical one to identify at least 20 heterozygotes for accurate genotype clustering required by GenomeStudio. After quality assessment, 1,163 subjects with 1,599,980 SNPs remained in the genetic association analysis (Supplementary Table S1). The genomic coordinate is based on hg18 unless mentioned otherwise.

SNP selection for replication

Four SNPs (rs72887860, rs62287262, rs16957091, and rs997781) associated with risk for MGMT methylation in LSC were genotyped as the tag SNPs at chromosome 15q15.2 locus for replication. These SNPs were genotyped using the TaqMan assay (Life Technologies) in the VSC.

Imputation of chromosome 15

Imputation of chromosome 15 in the LSC was conducted using BEAGLE (version 3.3.2) with phased haplotype data of EUR populations (n = 758, EUR.chr15.pha-sel_release_v3.20101123) from 1,000 genomes project pilot 1 study as the reference panel. Masked analyses on 20% of SNPs on chromosome 15 identified a high correlation of the observed versus expected allele frequencies (Pearson correlation coefficient = 0.98). The estimated allelic dosages for SNPs (n = 165,451) with dosage R2 = 0.3 and MAF ≥ 0.05 in the LSC and EUR reference populations were included for assessing genetic association with risk for gene methylation.

Enrichment of SNPs in ENCODE-annotated regulatory regions

The enrichment of SNPs in ENCODE-annotated regulatory regions within the 472-kb haplotype block at chromosome 15q15.2 locus was assessed using FuncSNP (34). Peak files included CTCF-binding sites, DNaI hypersensitivity locations,
DNase digital genomic foot printing, and regions enriched in active (H3K4me3 and H3K36me3) and repressive (H3K27me3) histone modification markers from small airway epithelial cells (SAEC). Known FAIRE-Seq (Formaldehyde-Assisted Isolation of Regulatory Elements), DNase hypersensitivity location, and gene promoter were included as well.

Stable siRNA-mediated knockdown of UBR1 in human BEC lines

Stable UBR1 knockdown (KD) cell lines were established using three immortalized human bronchial epithelial cell lines (HBEC1, HBEC4, and HBEC26; ref. 35). The pSUPER.puro (Oligo Engine) was modified to encode resistance to hygromycin (pSUPER.purohyg). A pair of oligonucleotides were annealed and inserted into the pSUPER.purohyg. The two oligonucleotide sequences are as follows: 5'–GATCCCG-GATCCGAATCTATTAAGATTCAAGAGATCTTAATAGATTCCGATCCGGG-3' and 5'–AGCTTTAAAAAGGATCGGAATCTATTAAGATTCCGATCCGGG-3'). The resulting construct uses the H1 RNA promoter to express shRNAs that are converted by endogenous enzymes to siRNAs targeting UBR1 mRNA. These constructs were verified by sequencing and transfected in GP2-293 cells along with a plasmid encoding the VSV-G protein. The resultant pseudotyped retroviral particles were used to transduce HBECs and drug-resistant cells or clones were obtained after selection with hygromycin. Control cells or clones were established following the same procedure except pSUPER.purohyg plasmid backbone was used. On average, 70% to 80% KD was achieved at the mRNA level in all HBECs. Western blots suggested that minimal residual UBR1 protein remained in KD HBECs compared with control lines (Fig. 1A, not shown).

Promoter activity assays

A 1,528-bp fragment containing the entire promoter and CpG island of MGMT with different haplotype alleles was previously cloned into the pGL2-basic Luciferase Reporter Vector (Promega) upstream of the luciferase coding sequence (27). Luciferase reporter constructs that contain HAP1 (wild haplotype allele) or HAP4 (haplotype allele carrying A allele of rs16906252) were transiently transfected in control HBEC4 and HBEC4 with UBR1 KD using Lipofectamine 3000. Transfection in HBEC1 and HBEC26 was conducted using the Neon electroporation system. Cells were harvested 48 hours posttransfection and reporter activity was measured using the Dual Luciferase Assay System (Promega).

Real-time PCR for quantifying gene expression

RNA was isolated from primary human BEC cultures (n = 54), normal lungs (n = 6), and cell lines. TaqMan real-time PCR was conducted to quantify gene expression in cDNA using the delta threshold cycle method with β-actin as the endogenous control.

Western blot and protein expression quantification

Western blot was conducted to quantify protein levels of UBR1 and MGMT using antibodies from Abcam (anti-UBR1 antibody, ab42420) and Invitrogen (mouse anti-MGMT, ab23331). Western blot and protein expression quantification were conducted to quantify protein levels of UBR1 and MGMT using antibodies from Abcam (anti-UBR1 antibody, ab42420) and Invitrogen (mouse anti-MGMT, ab23331). Western blot and protein expression quantification were conducted to quantify protein levels of UBR1 and MGMT using antibodies from Abcam (anti-UBR1 antibody, ab42420) and Invitrogen (mouse anti-MGMT, ab23331).

Quantification of O6MG

HBECs were treated with 1 mmol/L nitrosomethylurea (MNU) in the dark for 30 minutes. Cells were harvested at 0, 24, 48, and 72 hours post-MNU treatment. O6MG was quantified in genomic DNA using HPLC-MS (37). The average amount recovered for 50 fmol positive controls ([15N5-13C5]-O6-methyl-2'-deoxyguanosine) was 47 ± 8 fmol (n = 4). The limit of detection was 1 fmol, whereas limit of quantitation was 2.5 fmol per injection. Genomic O6MG level was expressed as the ratio between the amount of O6MG and dG within each DNA sample. The percentage of O6MG remained in KD HBECs compared with control lines (Fig. 1A, not shown).

Figure 1. Effect of UBR1 KD on MGMT turnover and DNA repair. A, Western blot for UBR1 KD in HBEC1. Stable transfection with pSUPER.purohyg expressing siRNA targeting UBR1 led to almost undetectable UBR1 protein in HBEC1. B, the ratios of naïve versus suicide MGMT were higher in control HBECs compared with UBR1 KD lines (0.95 ± 0.29 vs. 0.71 ± 0.30, paired t test, P = 0.037), with largest difference seen in HBEC26 (0.36). C, Western blot for MGMT protein in control HBEC26 and UBR1 KD line with and without GluC digest. D, DNA repair for O6MG is slower in control HBECs than UBR1 KD lines (percentage of O6MG left unrepaired at 24 hours post-MNU treatment; 0.22 ± 0.0049 vs. 0.11 ± 0.02, paired t test, P = 0.011). Results are summarized with mean and standard deviation based on experiments from three HBEC lines.
O6MG left unrepaired at 24 hours post-MNU treatment was further standardized by taking into account the cell growth rate.

The cytokinesis-block micronucleus assay
Phytohemagglutinin-stimulated exponentially growing lymphocyte cultures established from LSC members (n = 107) were treated with 1 mmol/L MNU for 4 hours and then incubated with 4.5 μg/mL cytochalasin B for an additional 24 hours before harvest (38). The dose selected was within the linear dose–response range and caused obvious genotoxicity, but minimal cytotoxicity. The micronuclei and necleoplasmic bridge were scored to allow the calculation of the percentage of cells with chromosomal damages.

Statistical analysis
In the GWAS discovery, the genetic association was assessed in 1,163 subjects using logistic regression with adjustment for age and sex. The genotype of rs16906252 was further included for covariate adjustment in models that would identify novel SNPs whose association with risk for MGMT methylation was independent of rs16906252. SNPs were assessed under an additive genetic model. Odds ratio (OR) and 95% confidence intervals (95% CI) were calculated to quantify the magnitude of the association per allele.

In the VSC replication, logistic regression was applied to assess the association between each individual SNP and risk for MGMT methylation in 430 subjects with adjustment for age, ethnicity (Hispanic versus non-Hispanic White), and sex. The genotype of rs16906252 was included in the logistic regression model to allow for identification of novel SNPs with associations independent of rs16906252 and this identified 12 common SNPs with P < 10−5 (Supplementary Table S2). SNPs located around the MGMT promoter became nonsignificant (Ps > 0.94; Supplementary Table S2 and Fig. S2), suggesting the associations seen for these SNPs reflected some degree of linkage disequilibrium (LD) with rs16906252. However, multiple associations across chromosome 15q15.2 locus remained statistically significant and thus likely independent of rs16906252.

Replication in VSC
Four SNPs (rs72887860, rs62287262, rs16957091, and rs997781) associated with risk for MGMT methylation in LSC with P < 10−5 and MAF > 0.1 were selected for replication in 430 members from the VSC with MGMT methylation detected in 107 subjects (24.9%). Rs16957091 was genotyped as the tag SNP on chromosome 15q15.2. Associations were replicated on chromosome 15q15.2 (rs16957091, P = 0.016) and 17q24.3 loci (rs997781, P = 0.016). Meta-analysis combining the results for these two SNPs from LSC and VSC identified associations of near genome-wide significance (Table 1). Data from the LSC and VSC cohorts were pooled to assess change in the area under the curve (AUC) of the receiver operating characteristic (ROC) by adding rs16957091, rs997781, and rs16906252 to the basic model containing age, sex, race, and cohort identifier only (Supplementary Table S4). Adding all three SNPs into model increased the ROC AUC by 19.6% (from 54.2% to 73.8%, P = 2.97 × 10−5).

Association based on imputed genotypes at chromosome 17q24.3 and 15q15.2 loci
Rs997781 is located within a 131-kb haplotype block (chr17:64593108-64724190, hg18) on chromosome 17q24.3 containing ABCA6 and ABCA10 (not shown). Rs16957091 is located within a 472-kb haplotype block (chr15:40709948-41181608, Supplementary Table S5).

Table 1. Associations between four SNPs and risk for MGMT methylation in LSC and VSC

<table>
<thead>
<tr>
<th>SNP</th>
<th>Location</th>
<th>Allele*</th>
<th>MAF*</th>
<th>LSC OR (95% CI)</th>
<th>P</th>
<th>VSC OR (95% CI)</th>
<th>P</th>
<th>Meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs72887860</td>
<td>Chr3:61999431</td>
<td>G/A</td>
<td>0.14</td>
<td>1.87 (1.42–2.47)</td>
<td>7.5 × 10−6</td>
<td>0.91 (0.57–1.45)</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>rs62287262</td>
<td>Chr4:660119</td>
<td>A/G</td>
<td>0.40</td>
<td>0.58 (0.47–0.72)</td>
<td>1.4 × 10−4</td>
<td>1.20 (0.84–1.73)</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>rs16957091</td>
<td>Chr5:4080478</td>
<td>T/G</td>
<td>0.23</td>
<td>1.72 (1.36–2.17)</td>
<td>4.5 × 10−6</td>
<td>1.63 (1.30–2.04)</td>
<td>0.016</td>
<td>1.70 (1.39–2.07)</td>
</tr>
<tr>
<td>rs997781</td>
<td>Chr17:664629283</td>
<td>G/A</td>
<td>0.29</td>
<td>1.64 (1.33–2.03)</td>
<td>4.2 × 10−6</td>
<td>1.57 (1.09–2.25)</td>
<td>0.016</td>
<td>1.63 (1.36–1.95)</td>
</tr>
</tbody>
</table>

*Second allele is the minor allele.
*MAF was assessed in LSC.
Logistic regression with adjustment for age, sex, and rs16906252.
Logistic regression with adjustment for age, ethnicity (Hispanic vs. non-Hispanic White), and rs16906252.
hg18) on chromosome 15q15.2 that contains STARD9, CDAN1, TTBK2, and UBR1 (Fig. 2). Genotype imputation successfully (dosage $R^2 > 0.3$) estimated the allelic dosage for 126,145 common SNPs (MAF $\geq 0.05$) on chromosome 15 that were not genotyped by the Illumina HumanOmni2.5M chip. Association analyses based on imputation results on chromosome 15 identified five SNPs associated with risk for MGMT methylation with $P$-values slightly lower than that seen for rs16957091, although all five SNPs are in almost perfect LD ($R^2 > 0.95$; Fig. 2). Imputation at chromosome 17q24.3 did not find any associations that were more significant than that seen for rs997781.

Rs16957091 affects UBR1 expression

The enrichment of SNPs in ENCODE-annotated regulatory regions within the 472-kb haplotype block (chr15:40709948-41181608, hg18) at chromosome 15q15.2 locus was assessed using FunciSNP with data from SAEC. Multiple chromatin features surrounding STARD9, CDAN1, TTBK2, and UBR1 were identified to overlap with SNPs in LD with rs16957091 ($R^2 > 0.95$; Fig. 2). Imputation at chromosome 17q24.3 did not find any associations that were more significant than that seen for rs997781.

Reduced UBR1 expression affects MGMT turnover and DNA repair

UBR1 contains binding sites recognizing internal degrons of MGMT protein and the UBR1/RAD6-dependent N-end rule

Figure 2.
Imputation of chromosome 15q15.2. Imputation was conducted using phased genotype data of EUR populations from 1,000 Genomes project. Genetic associations are plotted for 941 variants (MAF $\geq 0.05$) within the region of Chr15: 40304748–41304718 (hg18) that include 758 imputed variants (dosage $R^2 > 0.3$). Circles, genotyped SNPs; squares, imputed SNPs; purple circle, rs16957091. Degree of LD with rs16957091 is schemed as the gradient of red to dark blue color, with red as high LD ($R^2 = 0.8–1$) and blue as low LD ($R^2 = 0–0.2$). No SNPs in LD ($R^2 > 0.2$) with rs16957091 can be detected outside of this region on chromosome 15.
pathway is a major mechanism targeting suicide MGMT protein by alkylating agents and also unmodified MGMT for subsequent degradation via the 26S proteasome in eukaryotes (40, 41). Recovery of MGMT activity after its inactivation results entirely from de novo protein synthesis (4). Our previous studies demonstrated reduced gene transcription associated with the enhancer SNP rs16906252 as a major mechanism predisposing MGMT for promoter hypermethylation in the lungs of smokers (27). Thus, the lower expression of UBR1 associated with rs16957091 variant allele could result in slower turnover of the MGMT protein under normal physiological conditions and/or following repair in response to challenge by the alkylating agents, that in turn may lead to reduced MGMT transcription.

UBR1 stable KD cell lines were established from three immortalized HBECs (HBEC1, HBEC4, and HBEC26) to test this hypothesis. A protease V8 digestion method (36) was used to distinguish naïve and suicide MGMT in HBEC1 treated with different MNU concentrations (0, 0.5, 1, and 2.5 mmol/L) for 30 minutes. The ratio of naïve versus suicide MGMT decreased with increasing concentrations of MNU treatment with almost no active MGMT left at the highest MNU concentration (not shown). Interestingly, without MNU treatment, protease V8 cutting is still detected supporting formation of endogenous alkylating DNA adducts as a part of normal cellular metabolism in these cells (1, 42), although the assay specificity for detecting suicide MGMT may not be perfect due to its potential binding to other proteins. The ratio of naïve versus suicide MGMT was higher in control HBECS compared with UBR1 KD lines (0.95 vs. 0.71, paired t test P = 0.037; Fig. 1B) with the largest difference seen in HBEC26 (0.36; Fig. 1C), suggesting accumulation of suicide MGMT in UBR1 KD lines. UBR1 knockout cells are resistant to alkylating agent (MNNG) induced genotoxicity probably because UBR1 targets both alkylated and unmodified MGMT (41). Thus, HBECs were treated with 1 mmol/L MNU and O6MG was quantified in genomic DNA recovered from cells at 0, 24, 48, and 72 hours post-MNU treatment. On average, 78% of O6MG adducts were repaired within 24 hours in control HBECs. In contrast, 89% O6MG were repaired within 24 hours in UBR1 KD lines (Fig. 1D). Rates of repair within 24 hours differed significantly between control and UBR1 KD (paired t test, P = 0.011; Fig. 1D). The modest but statistically significant effect of UBR1 KD on MGMT degradation and repair of O6-methylguanine may reflect the complexity of constitutive and exogenous alkylating agent-accelerated degradation of MGMT. In addition to the UBR1/RAD6-dependent N-end rule pathway, the UFD4/UBC4-dependent ubiquitin fusion degradation pathway functions as an enhancer of the processivity of MGMT polyubiquitylation by the N-end rule pathway through physical interaction between the HECT-type UFD4 E3 and the RING-type UBR1 E3 (40, 41).

Because rs16957091 reduced UBR1 expression and UBR1 KD was associated with faster DNA repair of O6MG induced by MNU, we further assessed whether rs16957091 is associated with better repair in cultured lymphocytes collected from LSC subjects (n = 107). A significant inverse association (P = 0.036; Fig. 3B) between the number of rs16957091 G alleles served, suggesting that the G allele associated with reduced UBR1 expression is associated with greater DNA repair of MNU-induced DNA damage.

UBR1 expression levels affect MGMT transcription

Reduced gene transcription is a major mechanism predisposing MGMT for promoter hypermethylation in multiple tissue types (13, 27); therefore, a luciferase reporter assay was used to assess whether UBR1 KD would result in reduced MGMT promoter activity. Consistent with our previous studies and others, HAP4 (rs16906252 carrier) was associated with a 35% (P < 0.0001) and 22% (P = 0.0019) reduction in promoter activity compared with HAP1 (wild type) in control and UBR1 KD HBEC1 lines, respectively (Fig. 4A). Excitingly, UBR1 KD reduced HAP1 promoter activity by 25% compared with control line (P = 0.0009). Reduction of promoter activity (12%) in UBR1 KD line was also observed for HAP4 (P = 0.089), although to a lesser degree that likely stemmed from the inefficient transcription regulatory mechanism due to carrying rs16906252. The combination of UBR1 KD and HAP4 reduced promoter activity by 50% (P < 0.0001) compared with wild-type UBR1 and HAP1. The reduction of HAP1 promoter activity in the UBR1 KD HBEC4 line was
replicated in HBEC1 (reduced by 23%; \( P = 0.0046 \)) and HBEC26 (reduced by 28%; \( P = 2.5 \times 10^{-8} \)), B, MNU treatment at a noncytotoxic dose (1 mmol/L) induced MGMT transcription in control HBEC4 and HBEC26 at 24 and 48 hours post-MNU treatment. UBR1 KD largely abolished this induction in HBEC4 and HBEC26. Although no induction was observed in HBEC1, control line still had higher MGMT expression than UBR1 KD line (\( P = 0.086 \)). Results are summarized with mean and standard deviation.

**Discussion**

Our GWAS study identified cis- and trans-acting SNPs that affect silencing of a DNA repair gene (MGMT) by promoter hypermethylation acquired during lung carcinogenesis. Two novel trans-acting loci at chromosome 15q15.2 and 17q24.3 were identified as being associated with risk for MGMT methylation detected in sputum and their effects were independent of the enhancer SNP rs16906252. These three SNPs together improved the ROC AUC by 20% in a logistic regression that explored determinants for risk for MGMT methylation in current and former smokers. Thus, SNPs affecting the predisposition of MGMT silencing by promoter hypermethylation, together with SNPs affecting methylation of other important cancer-relevant genes acquired during carcinogenesis may lead to the establishment of a polygenic marker that could improve current risk prediction models for these cancers (18).

MGMT methylation status is a validated prognostic biomarker for the response of glioblastoma patients to the alkylating agent temozolomide (14). Our previous study and others also suggested that MGMT methylation status may affect the response of lung tumors to temozolomide treatment (27, 44). The assessment of MGMT promoter activity using HBEC4 identified a 50% reduction in luciferase reporter activity for the combination of UBR1 KD and HAP4 (rs16906252 carrier) compared with wild-type UBR1 and HAP1. The reduction of HAP1 promoter activity in UBR1 KD HBEC4 lines was replicated in HBEC1 (reduced by 23%; \( P = 0.0046 \)) and in HBEC26 (reduced by 28%; \( P = 2.5 \times 10^{-8} \)).
Ontology classification of genes (n = 115) in the region of the proximal CpG sites mediating the trans-acting SNP rs16957091 located 223 kb from the UBR1 gene. This SNP was found to be associated with reduced MGMT transcription in normal lungs, and this association was supported by a search for trans-acting meQTL that regulate MGMT transcription in the region of the proximal CpG sites. The mechanism for how rs16957091, located 223 kb from the UBR1 gene, affects MGMT transcription is still unknown. This study also verified MGMT enhancer SNP (rs169062652) as the driver cis-acting SNP for MGMT methylation through reducing MGMT transcription. Thus, both cis- and trans-acting SNPs could affect the MGMT transcriptional machinery that predisposes MGMT for promoter hypermethylation in lungs of smokers.
rs68062403 that in turn may affect the binding of a transcription factor.

Our studies provide a proof-of-concept for a unique mechanism underlying the predisposition of MGMT methylation in lungs of smokers that involves the regulation of MGMT homeostasis by the UBR1/RAD6-mediated N-end rule pathway (Fig. 5). The cis- and trans-acting SNPs affecting MGMT methylation should have strong clinical translational significance in terms of their utility in cancer risk assessment and patient stratification for alkylating agent chemotherapy.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: S. Leng, C.S. Tellez, D. Desai, C.A. Stidley, Y. Liu, Y. Lin, M.G. Wathelet, F.D. Gilliland, S.A. Belinsky

Writing, review, and/or revision of the manuscript: S. Leng, C.S. Tellez, X. Zhang, C.A. Stidley, Y. Liu, S.A. Belinsky

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): S. Leng, L.B. Collins, C.S. Tellez, R.E. Crowell, F.D. Gilliland, S.A. Belinsky

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S. Leng, C.S. Tellez, X. Zhang, C.A. Stidley, Y. Liu, S.A. Belinsky

Other (primarily worked with Dr. Belinsky and Shuguang Leng on the DNA adduct data, in collaboration with Leonard Collins in the mass spectrometry facility): J.A. Swenberg

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C.I. Thomas, M.A. Piech, D.E. Juri, D. Desai, S.A. Belinsky

Study supervision: S. Leng, S.A. Belinsky

Other (statistical analysis): G. Wu

References


Implication of a Chromosome 15q15.2 Locus in Regulating UBR1 and Predisposing Smokers to MGMT Methylation in Lung


Cancer Res  Published OnlineFirst July 16, 2015.

Updated version  Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-15-0243

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, use this link http://cancerres.aacrjournals.org/content/early/2015/07/16/0008-5472.CAN-15-0243. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.