Carcinogenesis vol.29 no.7 pp.1380–1385, 2008
doi:10.1093/carcin/bgn121
Advance Access publication May 21, 2008

Haplotypes of *DNMT1* and *DNMT3B* are associated with mutagen sensitivity induced by benzo[*a*]pyrene diol epoxide among smokers

Shuguang Leng, Christine A.Stidley[1], Amanda M.Bernauer, Maria A.Picchi, Xin Sheng[2], Melissa A.Frasco[2], David Van Den Berg[2], Frank D.Gilliland[2], Richard E.Crowell[1,3] and Steven A.Belinsky*

Lung Cancer Program, Lovelace Respiratory Research Institute, 2425 Ridgecrest Drive Southeast, Albuquerque, NM 87108, USA, [1]Department of Internal Medicine, University of New Mexico, 2500 Lomas Boulevard, Albuquerque, NM 87131, USA, [2]Division of Enviromental and Occupational Health, Keck School of Medicine, University of Southern California, 1540 Alcazar Street, Los Angeles, CA 90033, USA and [3]Department of Internal Medicine, New Mexico VA Health Care System, 1501 San Pedro Southeast, Albuquerque, NM 87108, USA

*To whom correspondence should be addressed. Tel: +1 505 348 9465; Fax: +1 505 348 8567;
Email: sbelinsk@lrri.org

The mutagen sensitivity assay is an *in vitro* measure of DNA repair capacity used to evaluate intrinsic susceptibility for cancer. The high heritability of mutagen sensitivity to different mutagens validates the use of this phenotype to predict cancer susceptibility. However, genetic determinants of mutagen sensitivity have not been fully characterized. Recently, several studies found that three major cytosine *DNA methyltransferases* (*DNMTs*), especially *DNMT1*, have a direct role in the DNA damage response, independent of their methyltransferase activity. This study evaluated the hypothesis that sequence variants in *DNMT1*, *DNMT3A* and *DNMT3B* are associated with mutagen sensitivity induced by the tobacco carcinogen benzo[*a*]pyrene diol epoxide (BPDE) in 278 cancer-free smokers. Single-nucleotide polymorphisms ($n = 134$) dispersed over the entire gene and regulatory regions of these *DNMTs* were genotyped by the Illumina Golden Gate Assay. DNA sequence variation in the *DNMT1* and *DNMT3B* loci was globally associated with breaks per cell ($P < 0.04$ for both). No global association between *DNMT3A* and breaks per cell was seen ($P = 0.09$). Two haplotypes in block1 of *DNMT1* (H284) and 3B (H70) were associated with 16 and 24% increase in breaks per cell, respectively. Subjects with three or four adverse haplotypes of both *DNMT1* and *3B* had a 50% elevation in mean level of breaks per cell compared with persons without adverse alleles ($P = 0.004$). The association between sequence variants of *DNMT1* and *3B* and mutagen sensitivity induced by BPDE supports the involvement of these *DNMTs* in protecting the cell from DNA damage.

Introduction

The mutagen sensitivity assay is an *in vitro* measure of DNA repair capacity expressed as chromatid breaks per cell in short-term cultured lymphocytes challenged by mutagens. Numerous epidemiological studies have consistently found significant associations between mutagen sensitivity and risk for several environmental cancers (reviewed in refs 1–3). Different mutagens used in the mutagen sensitivity assay provide information about the involvement of specific DNA damage and repair pathways. For example, using the well-known carcinogens benzo[*a*]pyrene diol epoxide (BPDE)-1 (4), gamma radiation (5) and ultraviolet light (6) as the test mutagens can help address the susceptibility to cancers with a specific etiology. Wu *et al.* (7) systemically evaluated the heritability of mutagen sensitivity to four different mutagens among 255 twins and found that up to 40–60% of variability of this phenotype can be explained by genetics, validating the use of mutagen sensitivity as a cancer susceptibility factor. Several functional single-nucleotide polymorphisms (SNPs) in *XPA*, *XPC* and *RAD23B* in the nucleotide excision repair pathway modified the mutagen sensitivity induced by BPDE individually and jointly (8). Furthermore, the association between functional SNPs in *XPC* and mutagen sensitivity induced by BPDE was validated in another population with the comet assay as the readout of DNA damage and repair (9). Thus, sequence variations in DNA repair genes that modify protein activity clearly impact mutagen sensitivity, and alterations in sequence of these genes could cause suboptimal mutagen sensitivity to increase cancer risk. However, emerging studies also implicate a new class of genes that participate in DNA damage recognition to initiate repair. The *cytosine DNA methyltransferases* (*DNMTs*) represent one class of genes responding to DNA damage.

DNA methylation in mammals is critically involved in controlling gene expression, genomic imprinting, X-chromosome inactivation, embryonic development and neoplastic transformation (reviewed in ref. 10). DNA methylation patterns are established and maintained by three *DNMTs*: *DNMT1*, *DNMT3A* and *DNMT3B* (10). *DNMT1* acts primarily as a maintenance *DNMT* by copying existing 5-methylcytosine patterns in a proliferating cell nuclear antigen-dependent or independent manner following DNA replication and repair (10–12). *DNMT3A* and *DNMT3B* exhibit *de novo* methyltransferase activity (10). Recently, several studies found that *DNMTs*, especially *DNMT1*, have a direct role in the DNA damage response independent of their methyltransferase activity. Mortusewicz *et al.* (13) reported that *DNMT1* was rapidly recruited to the sites of DNA damage after ultraviolet radiation in a manner dependent on its proliferating cell nuclear antigen interaction domain. *DNMT1* strongly colocalized with *γH2AX* in 5-aza-2′-deoxycytidine (5-azadC)-treated cells, and *DNMT1* hypomorph HCT116 cells that express decreased amount of a mutant DNMT1 protein (12) demonstrated profound defects in DNA damage response toward 5-azadC and several other DNA-damaging agents, including doxorubicin, hydroxyurea, bleomycin and exposure to ultraviolet light (14). Furthermore, the defective DNA damage response in *DNMT1* hypomorph HCT116 cells was partially due to altered subcellular localization of *CHK1* (14). *DNMT1* also interacts with *p53* to repress transcription of the *survivin*, *cdc2* and *cdc25* genes in cells treated with doxorubicin (15–17) and may be integral to some checkpoint signaling mechanisms in the DNA damage response. Several studies support a role for *DNMT3A* and *DNMT3B* in modulating the DNA damage response. Palii *et al.* (14) found that *DNMT3B*-deficient cells showed a reduced DNA damage response after 5-azadC treatment. *DNMT3A* can repress *p53*-mediated *p21* transcription following DNA damage, and this repression does not require methyltransferase activity (18).

With the identification of additional functions for *DNMTs* in the DNA damage response, we hypothesized that the sequence variants in *DNMTs* may influence DNA repair capacity. This hypothesis was tested by determining whether haplotypes constructed from a panel of comprehensively selected SNPs within these genes were associated with mutagen sensitivity induced by BPDE in peripheral lymphocytes from 278 cancer-free smokers.

Abbreviations: 5-azadC, 5-aza-2′-deoxycytidine; BPDE, benzo[a]pyrene diol epoxide; DNMT, cytosine DNA methyltransferase; DSB, double-strand break; LD, linkage disequilibrium; LRT, likelihood ratio test; LSC, Lovelace Smokers Cohort; SCC, Spearman correlation coefficient; SNP, single-nucleotide polymorphism.

Materials and methods

Study population and sample collection

Subjects in this study were selected from participants of two smoker cohorts: the Lovelace Smokers Cohort (LSC) and the Veteran Smokers Cohort, which were established in 2001 and 2000, respectively, to conduct longitudinal studies on molecular markers of respiratory carcinogenesis in biological fluids such

as sputum and blood from people at risk for lung cancer. Descriptions of these cohorts have been detailed previously (19,20). Briefly, the cohort participants were all current or former smokers between 40 and 75 years old and were mainly residents in the Albuquerque metropolitan area. All participants signed the consent form. Whole blood was processed within 2 h after blood draw to isolate lymphocytes and plasma. Cryopreservation of lymphocytes began in 2005. Cryopreserved lymphocytes were available from 278 subjects, including 215 subjects from the LSC and 63 subjects from the Veteran Smokers Cohort. This study was approved by the institutional review boards of the Lovelace Respiratory Research Institute and the University of New Mexico.

Mutagen sensitivity assay

Phytohemagglutinin (M form)-stimulated lymphocytes were treated with BPDE to evaluate the generation of chromosome aberrations as an index of mutagen sensitivity. Briefly, cryopreserved lymphocytes were thawed and cultured in RPMI 1640 medium supplemented with fetal bovine serum (20%) and phytohemagglutinin (1.5%). Cell density was adjusted to $<0.5 \times 10^6$ per ml and 72 h after phytohemagglutinin stimulation, the culture was split into two T25 flasks. Cells were treated with BPDE or vehicle for 24 h because BPDE is an S-phase-dependent clastogen. The final concentration for BPDE in culture medium was 0.3 μM, a concentration defined through dose–response studies using isolated lymphocytes from cohort subjects and three lymphoblastoid cells lines: GM02345 (mutant *XPA*), GM16024 (mutant *XPG*) and GM00131 (wild-type *XPA* and *XPG*) (data not shown). The dose selected was within the linear dose–response range and caused obvious genotoxicity but minimal cytotoxicity. One hour before harvest, colcemid was added to the cultures at a final concentration of 0.06 mg/ml. Slides were prepared according to conventional procedures, and 100 well-spread metaphases were examined for chromatid breaks or exchanges. Samples from the LSC and Veteran Smokers Cohort were randomly mixed and assayed as a batch. Each simple chromatid break was scored as one break, whereas each isochromatid break set and each exchange event (including interstitial deletion) were considered as two breaks (21). Occasionally, a metaphase with >12 breaks was observed on a slide with BPDE treatment and was recorded as 12. Mutagen sensitivity was expressed as the mean number of chromatid breaks per cell.

SNP selection and genotyping by the Illumina platform

A comprehensive procedure was followed in which SNPs (n = 182) were selected from the single-nucleotide polymorphism database based on a SNP density of one to three SNPs/kb, validation status, Illumina design score and functional potential of the SNPs. These 182 SNPs covered the entire coding area of the three *DNMTs* together with the areas 20 kb upstream and 10 kb downstream of each gene. The inclusion of SNPs in the upstream and downstream regions improved the accuracy for defining the block structure surrounding the coding areas. The SNP density was predetermined based on the haplotype block structure of the expanded region of *DNMT1*, *3A* and *3B* that were generated using the genotype data of 30 U.S. trios in the HapMap project (http://www.hapmap.org/). One non-synonymous SNP (rs8111085) in *DNMT1* and two SNPs [rs6087990 and rs2424913; (22,23)] that change basal promoter activity of *DNMT3B* and Δ*DNMT3B* (a newly discovered *DNMT3B* alternative transcript in tumor tissues) (23) were forced in the final list of SNPs for genotyping. These 182 SNPs were genotyped by the Illumina Golden Gate Assay (the SNPs list is available on request).

Statistical analysis

The call rate for each SNP was assessed prior to data analysis in two major ethnicities in the study population: non-Hispanic whites and Hispanics. SNPs were deemed unsuitable for analysis if they were monomorphic, had minor allele frequency <0.05 in both non-Hispanic whites and Hispanics, had a genotyping failure rate >5% or showed significant deviation from Hardy–Weinberg equilibrium. Pair-wise linkage disequilibrium (LD) between the common SNPs in each gene was assessed by $|D'|$ and r^2 statistics. The haplotype block structure of the *DNMTs* in non-Hispanics whites and Hispanics were characterized by Gabriel's algorithm implemented in the HaploView software (24).

A global test based on principal component analysis was employed to test whether the sequence variants in each gene as a single locus were associated with mutagen sensitivity induced by BPDE (25). The number of principal components needed to describe common variation across a locus was defined a priori as the number of principal components needed to explain at least 80% of the sequence variance. These components were used as predictors in a linear regression, with natural log mutagen sensitivity as the outcome and globally assessed using a likelihood ratio test (LRT). Adjustment for sex, current smoking status and seeding count of lymphocytes per culture was included in the genetic association analysis because they were the non-genetic variables associated with mutagen sensitivity in this study.

Table I. Demographic characteristics of study subjects

Variables	Study subjects (n = 278)
Age, mean ± SD	58.0 ± 9.9
Gender (%)	
Female	52.9
Male	47.1
Race (%)	
Non-Hispanic white	74.5
Hispanic	19.1
Others	6.5
Smoking history	
Current (%)	46
Pack-years, mean ± SD	40.9 ± 25.5
Duration, mean ± SD	32.7 ± 10.6
Cryopreservation duration (days, mean ± SD)	368 ± 190
Seeding count ($\times 10^6$ per culture, mean ± SD)	1.7 ± 0.6
Mutagen sensitivity (breaks per cell, mean ± SD)	0.28 ± 0.16

A Bayesian statistical method implemented in the program PHASE (version 2.1) was used to reconstruct the haplotypes and calculate their estimated probabilities from the SNP data in each haplotype block in non-Hispanic whites (n = 207), Hispanics (n = 53) and subjects with other ethnicities (n = 18) (26,27). All possible haplotypes with corresponding probabilities (>0.01) for each individual were generated. The probabilities of the common haplotypes in each gene for each individual were used as explanatory variables in multivariate linear regression models with adjustment for non-genetic factors to assess the association between mutagen sensitivity and the haplotypes. Haplotypes with frequency <5% were combined into one group. The effect for each common haplotype was estimated using all other haplotypes as the reference group. The effects for the common haplotypes were also analyzed using the most common haplotypes in each gene as the reference group and similar results were obtained (data not shown).

A global test of the haplotypes within a gene was conducted using a LRT from a linear regression model that assessed all haplotypes simultaneously. A minimal P-value-based permutation test also was employed to assess whether a gene globally associated with mutagen sensitivity from the principal component analysis was associated with mutagen sensitivity induced by BPDE (28). This gene-based test provides a mean to avoid the high false-positive rate resulting from multiple comparisons. The mutagen sensitivity data were randomly permuted for the 278 subjects 10 000 times. Association tests for each common haplotype in each gene were performed using the permuted data. A distribution of the minimum P-values for the haplotype effects in each gene from each permutation was generated, and the minimal observed P-values for the haplotype effects in each gene were reevaluated in relation to this distribution. The permutation-based P-value was calculated as the proportion of the minimum P-values from each permutation that were equal to or smaller than the observed minimal P-value.

Analyses were run including only non-Hispanic white participants (74%) and including only participants in the LSC (77%) in order to assess the effects of ethnicity and differences in cohorts on the results. Only results from the analysis including the entire study group of 278 participants are presented because results were similar. All statistical analyses were performed with SAS/STAT 9.1.3 and PHASE (version 2.1).

Results

Population characteristics and mutagen sensitivity

Demographics and smoking habits of the 278 subjects are described in Table I. The average age was 58.0 ± 9.9. These subjects included 207 (74.5%) non-Hispanic whites, 53 (19.1%) Hispanics and 18 (6.5%) subjects of other ethnicities. Most cryopreserved samples were used within 2 years of sample collection and had >95% viability of lymphocytes at the start of culturing. The mean number of breaks per cell was 0.28 with a median of 0.24 (interquartile range of 0.16–0.35). Three non-genetic variables gender, current smoking status and seeding count of lymphocytes were associated with mutagen sensitivity induced by BPDE (Table II). Most of the previous mutagen sensitivity studies, except for the one that assessed ultraviolet light-induced breaks per cell in skin cancer (6), used whole blood to evaluate the

chromosomal aberrations induced by the mutagens and did not report an association between seeding count of lymphocytes and mutagen sensitivity induced by BPDE. However, our finding is not unexpected because increased cell number and the associated increase in size of lymphocyte aggregates in suspension in the cell culture will probably be protective against the genotoxicity of BPDE that is highly reactive in culture medium, because the cells are exposed to a lower concentration of BPDE. Indeed, the dose of BPDE added to the whole-blood culture was ~10 times higher than used for the lymphoblastoid cell lines to induce similar levels of chromatid breaks (4,29). Age, race, pack-years and cryopreservation duration did not influence mutagen sensitivity induced by BPDE (Table II).

Characterization of genetic variation at the three DNMTs loci

Of 182 SNPs assayed, 48 were deemed unsuitable because they were monomorphic or had minor allele frequency <0.05 in both non-Hispanic whites and Hispanics. The genotyping failure rate for the remaining 134 SNPs was <5%, and these SNPs showed no deviation from Hardy–Weinberg equilibrium ($P \geq 0.01$). Thus, these 134 SNPs were used for analysis. The final SNP densities for the expanded region were one SNP per 3.4, 2.3 and 1.5 kb for *DNMT1*, *3A* and *3B*, respectively (supplementary Table I and Figures 1–3, available at *Carcinogenesis* Online). These 134 SNPs were distributed relatively evenly across the expanded region of these genes (supplementary Figures 1–3, available at *Carcinogenesis* Online). Haplotype block structure was generated in non-Hispanic whites and Hispanics separately (supplementary Figures 1–3, available at *Carcinogenesis* Online). Similar haplotype block structures were found between non-Hispanic whites and Hispanics in our study population for all DNMTs. Thus, the haplotype block structure was redefined based on all subjects. The entire coding regions of *DNMT1* or *DNMT3B* were each located in one large block (block1 in supplementary Figure 1A or Figure 2A, available at *Carcinogenesis* Online). For *DNMT3B*, the large block extended ~20 kb upstream of the coding area. In contrast, three large blocks (≥17 kb) were found in *DNMT3A* separated by several small gaps and blocks (supplemental Figure 3, available at *Carcinogenesis* Online). Further comparison of the LD statistics of any two SNPs in each gene between the two major ethnicities in this study showed a high correlation for both non-Hispanic whites and Hispanics for the three *DNMTs* [*DNMT1*: r^2, Spearman correlation coefficient (SCC) = 0.81, $|D'|$, SCC = 0.47; *DNMT3A*: r^2, SCC = 0.58, $|D'|$, SCC = 0.51 and *DNMT3B*: r^2, SCC = 0.88, $|D'|$, SCC = 0.29; $P < 0.0001$ for all comparisons].

The association between mutagen sensitivity and sequence variants

A principal component-based approach was used to test whether the sequence variants in each of these three genes were associated globally with mutagen sensitivity (Table III). Principal components of *DNMT1* and *DNMT3B* were globally associated with the number of breaks per cell at the variance threshold of 80% ($P < 0.03$ and $P < 0.04$, respectively). No global association between *DNMT3A* and the number of breaks per cell was seen ($P = 0.09$).

The association between sequence variations of *DNMT1* and *DNMT3B* and mutagen sensitivity was further evaluated by a haplotype-based approach because these two genes were associated with mutagen sensitivity globally. Because of the similar haplotype block structure between non-Hispanic whites and Hispanics and highly correlated LD statistics for all pairs of SNPs in each *DNMT* for both ethnic groups, haplotype block boundaries were defined based on all subjects: *DNMT1*, block1 (rs11085587–rs8112895, 70 kb) and block2 (rs16999714–rs8106605, 7 kb); *DNMT3B*, block1 (rs6119279–rs8118663, 68 kb) and block2 (rs853858–rs242544, 1 kb). The entire coding area of *DNMT1* and *DNMT3B* is covered within a large block (block1 for either *DNMT1* or *DNMT3B*, supplementary Figures 1 and 2, available at *Carcinogenesis* Online), and the value of multiallelic D' between the common haplotypes in block1 and block2 is high for both *DNMT1* and *DNMT3B* (multiallelic $D' > 0.85$); therefore, only the results in block1 are presented for each gene. The haplotypes in block1 of *DNMT1* and *DNMT3B* with frequencies >5% in either non-Hispanic whites, Hispanics or subjects with other ethnicities are listed in Table IV. Three and six common haplotypes were observed that accounted for at least 95 and 75% of all chromosomes in each ethnic group in block1 of *DNMT1* and *DNMT3B*, respectively (Table IV). The association between common haplotypes in block1 of *DNMT1* and *DNMT3B* and mutagen sensitivity is shown in Table V. The likelihood ratio global tests of association indicate that haplotypes of *DNMT1* and *DNMT3B* are associated with mutagen sensitivity ($P = 0.037$ and $P = 0.023$, respectively). The haplotype H284 in block1 of *DNMT1* was associated with a 16% increase in breaks per cell with all other haplotypes as the reference ($P < 0.01$). The haplotype H200 in block1 of *DNMT1* was associated with a 13% reduction in breaks per cell with all other haplotypes as the reference ($P < 0.01$). Haplotypes H284 and H200 make a Yin–Yang haplotype pair within block1 of *DNMT1* according to Zhang's criteria (30). The haplotype H70 in block1 of *DNMT3B* was associated with a 24% increase in breaks per cell with all other haplotypes as the reference ($P < 0.01$).

Table II. BPDE-induced chromatid breaks per cell by age, sex, race, smoking history, cryopreservation duration and seeding count of cells in study subjects

Variables	*n*	Number of chromatid breaks per cell	*P*-value
		Geometric mean (95% CI)[a]	
Age (years)[b]			0.546
<57.5	139	0.23 (0.21–0.26)	
≥57.5	139	0.22 (0.20–0.25)	
Gender			0.026
Female	147	0.25 (0.22–0.28)	
Male	131	0.21 (0.19–0.23)	
Race			0.824
Non-Hispanic white	207	0.23 (0.21–0.25)	
Hispanic	53	0.24 (0.20–0.28)	
Others	18	0.18 (0.14–0.24)	
Current smoking status			0.012
Current	127	0.25 (0.23–0.28)	
Former	151	0.21 (0.19–0.23)	
Pack-years (pack × year)[b]			0.400
<35	139	0.22 (0.20–0.25)	
≥35	139	0.24 (0.21–0.26)	
Cryopreservation duration (days)[b]			0.258
<383	139	0.24 (0.22–0.27)	
≥383	139	0.22 (0.20–0.24)	
Seeding count ($\times 10^6$ per culture)[b]			<0.001
<1.6	136	0.27 (0.24–0.30)	
≥1.6	142	0.19 (0.17–0.22)	

CI, confidence interval.
[a]Least squares means (95% CI) were calculated based on natural log-transformed number of chromatid breaks per cell with adjustment for gender, current smoking status and seeding count and then were converted to their exponential form, i.e. geometric mean (95% CI).
[b]Categorized by median of selected variables among 278 study subjects.

Table III. Association between major principal components of the *DNMTs* and BPDE-induced chromatid breaks per cell

Statistics	*DNMT1*	*DNMT3A*	*DNMT3B*
Number of PCs	2	8	2
Cumulative variance explained (%)	87.7	82.4	87.3
P-value of LRT[a]	0.028	0.092	0.033

PC, principal components.
[a]Adjustment for sex, seeding count and current smoking status were included in the multiple linear regression model.

Table IV. Common (>5%) haplotypes and frequencies in block1[a] of *DNMT1* or *DNMT3B* by ethnicity

Gene	Haplotype	Allele	Frequency (%)[b]			
			Non-Hispanic whites (n = 207)	Hispanics (n = 53)	Others[c] (n = 18)	All (n = 278)
DNMT1	H284	CAGGGGTTTGCTGGCAAGATA	51	56	37	51
	H200	GGGATATCCGCTAACGTAATA	39	28	19	36
	H48	GACATGCCCTGCGGTGTGGGG	6	13	28	9
DNMT3B	H145	AAGTTTTCGGCTCTATTCCTGGTAGCAGACATCAGTTCTCGAAA	23	19	19	22
	H149	AAGTTTTCGGCTCTATTCCTGGTAGCAGACATCAGTTCTCGGAA	21	14	23	20
	H91	GGTCACTTGATCGCGCCTTCAGCATTGAGATGTGTCCGCTAGCG	12	19	11	14
	H70	GAGCATCCGATCGCGCCTTCAACGTTGAGATGTGTCCGCTAGCA	12	8	8	11
	H49	GGGCACTTGATCGCGCCTTCAGCATTGAGATGTGTCCGCTAGCG	5	18	8	8
	H11	GGTCACTTAATCGCGCCTTCAACGTTGAGATGTGTCCGCTAGCA	0	8	5	2

[a]Haplotype block boundaries were defined based on all subjects: *DNMT1*, block1 (rs11085587–rs8112895, 70 kb) and *DNMT3B*, block1 (rs6119279–rs8118663, 68 kb).
[b]Population haplotype frequencies were estimated from the genotype data in non-Hispanic whites, Hispanics and others separately using Bayesian statistical method implemented in the program PHASE (version 2.1).
[c]Others included Native American (n = 3), Black (n = 1), Asian (n = 4) and mixed ethnicity (n = 10).

Table V. Association between haplotypes in block1 of *DNMT1* or *DNMT3B* and a number of BPDE-induced chromatid breaks per cell

Haplotype	Frequency (%)	Effect per haplotype allele		*P*-value[b]	Permuted *P*-value[b,c]	Likelihood ratio *P*-value[d]
		Estimate[a]	95% CI			
DNMT1					0.015	0.037
H284	51	1.16	1.05–1.28	0.003		
H200	36	0.87	0.79–0.97	0.010		
H48	9	1.03	0.86–1.23	0.760		
Hother	4	0.87	0.70–1.08	0.203		
DNMT3B					0.056	0.023
H145	22	0.94	0.82–1.07	0.321		
H149	20	1.02	0.89–1.16	0.785		
H91	14	0.99	0.84–1.16	0.893		
H70	11	1.24	1.05–1.45	0.009		
H49	8	0.93	0.77–1.12	0.445		
Hother	25	0.93	0.82–1.04	0.205		

CI, confidence interval.
[a]Estimate is the ratio of breaks per cell between any common haplotype and all other haplotypes in each gene.
[b]*P*-value is from multiple linear regression model comparing haplotype to all other haplotypes and including adjustment for sex, seeding count and current smoking status.
[c]The mutagen sensitivity data were permuted for the 278 subjects 10 000 times at random. Permuted *P*-values were calculated based on Chen *et al.* (28).
[d]*P*-value is from multiple linear regression model with all haplotypes considered simultaneously and adjustment for sex, seeding count and current smoking status.

The joint effect between *DNMT1* and *DNMT3B* was evaluated by combining the two putative adverse haplotypes (H284 and H70). A 50% elevation in the mean number of breaks per cell was seen in lymphocytes from subjects with three or four adverse haplotypes compared with subjects without adverse alleles for *DNMT1* and *DNMT3B* ($P < 0.01$) (Table VI).

Minimal *P*-value permutation tests were conducted to adjust for multiple testing for each gene. The *P*-values were 0.015 and 0.056 for *DNMT1* and *DNMT3B*, respectively (Table V). This result indicates that, if a similar study was repeated under a null distribution (i.e. no haplotype in *DNMT1* or *DNMT3B* associated with mutagen sensitivity induced by BPDE), an association similar to that observed in *DNMT1* and *DNMT3B* would occur by chance only 1.5 and 5.6% of the time, respectively.

Table VI. Association between number of adverse haplotypes in *DNMT1* and *DNMT3B* and number of BPDE-induced chromatid breaks per cell

No. of adverse haplotypes[a]	n	Chromatid breaks per cell	*P*-value
		Geometric mean (95% CI)[b]	
0	58	0.19 (0.16–0.22)	Reference
1	119	0.22 (0.19–0.24)	0.152
2	79	0.27 (0.23–0.30)	0.001
≥3	22	0.29 (0.23–0.38)	0.004

CI, confidence interval.
[a]*DNMT1* H284 and *DNMT3B* H70 were the haplotypes evaluated for joint effect on mutagen sensitivity.
[b]Least squares means (95% CI) were calculated based on natural log-transformed number of chromatid breaks per cell with adjustment for gender, current smoking status and seeding count and then were converted to their exponential form, i.e. geometric mean (95% CI).

Discussion

This study evaluated a comprehensive panel of common and coding SNPs in three *DNMTs* in relation to mutagen sensitivity induced by BPDE among 278 cancer-free smokers. Both *DNMT1* and *DNMT3B* were associated with mutagen sensitivity. Two haplotypes (*DNMT1* H284 and *DNMT3B* H70) were significantly associated with a 16–24% increase in chromatid breaks per cell, suggesting a role for

DNMT1 and *3B* in protecting the cell from BPDE-induced DNA damage. The 18 subjects with other ethnicities did not bias the association of these haplotypes with DNA damage since exclusion of these subjects had no influence on the identified associations.

Several strategies were taken to control the false-positive rate that may result from the multiple comparisons in this candidate gene-based genetic association study. Association testing was based solely on the hypothesis that each *DNMT* as a genetic locus may be associated with mutagen sensitivity induced by BPDE. First, a test based on principal component analysis that captured the LD from multiple SNPs within a candidate region was used to assess whether a candidate locus was associated with a trait (25). Second, global assessment of haplotypes within each gene was conducted using the LRT. Finally, the statistical significance of the association between the haplotypes in *DNMT1* and *DNMT3B* and mutagen sensitivity was assessed by a minimal *P*-value permutation test within each gene that helps to minimize family-wise type I error rate when performing multiple testing (28). The permutation test compares each haplotype with all other haplotypes and the LRT assesses all haplotypes simultaneously. As indicated by Setiawan *et al* (31), if the contribution of a gene is primarily through one haplotype, then the permutation test may be more sensitive than the LRT. On the other hand, if multiple haplotypes contribute to the significance of the gene, then the LRT may be more appropriate.

Three putatively functional SNPs identified in *DNMT1* and *3B* were not associated with mutagen sensitivity (data not shown). In addition, rs8111085, the only non-synonymous SNP (Ile311Val) in *DNMT1*, has a very low LD with haplotype H284 and H200 in block1 of *DNMT1* ($r^2 < 0.09$). Furthermore, two SNPs (rs6087990 and rs2424913) that affect basal promoter activity of *DNMT3B* or Δ*DNMT3B* (a newly discovered *DNMT3B* alternative transcript in tumor tissue) also have a very low LD with the haplotype H70 in block1 of *DNMT3B* ($r^2 < 0.14$). The genetic variation for *DNMT1* and *DNMT3B* is contained in large haplotype blocks (70 kb for *DNMT1* and 68 kb for *DNMT3B*). Therefore, it is plausible that the SNPs or haplotypes significantly associated with the number of breaks per cell may be in complete or perfect LD with SNPs in transcriptional regulatory regions or in microRNA-binding sites, which in turn could influence the messenger RNA levels of DNMT1 or DNMT3B following exposure to BPDE. Testing this hypothesis will require the identification of the minimal promoters, critical transcription factors and possible microRNA-binding sites for these genes in conjunction with resequencing of their regulatory regions.

The involvement of *DNMT1* and *3B* in protecting the cell from BPDE-induced DNA damage is biologically plausible. Bulky DNA adducts caused by BPDE can block DNA replication fork progression (32). Persistent blockage of the replication fork by bulky lesions leads to DNA double-strand breaks (DSBs) that elicit histone H2AX phosphorylation and ataxia telangiectasia-mutated kinase/CHK2-mediated events (32). The link between *DNMT1* and BPDE-induced chromatid breaks could partially be explained by the involvement of *DNMT1* in DNA damage response to DSBs. One of the earliest responses to DSB damage is phosphorylation by ataxia telangiectasia-mutated kinase of the histone H2AX, which then facilitates accumulation of repair and signaling proteins. Palii *et al.* (14) found that expression of a variant DNMT1 causing an 80% reduction in protein in HCT116 cells completely abrogates the phosphorylation of H2AX by ataxia telangiectasia-mutated kinase, suggesting that *DNMT1* is an essential DNA damage signaling molecule in DSB repair. Furthermore, doxorubicin treatment (which induces DSBs) enhanced the coimmunoprecipitation of DNMT1 and CHK1 in HeLa nuclear extract, confirming a direct link between DNMT1 and the downstream events in DSB repair (14). Several studies showed that in cells treated with doxorubicin, DNMT1 is recruited by activated p53 and binds to functional Sp1 sites within promoters of the survivin, cdc2 and cdc25 genes (15–17). Therefore, it is plausible that activated p53 also could recruit DNMT1 to DNA damaged sites to facilitate the repair of DSBs because phosphorylated p53 rapidly binds to sites of DNA breaks after irradiation (33). Together, these studies strongly support a direct role of *DNMT1* in signaling and repairing DSBs induced by BPDE. The sequence variants in *DNMT1* may reduce the level of DNMT1 protein and impair the cells' capacity to protect the integrity of the genome upon environmental insult.

The association between *DNMT3B* and chromatid breaks induced by BPDE is not unexpected because several studies support a role for *DNMT3B* in modulating the DNA damage response. Palii *et al.* (14) found that *DNMT3B*-deficient cells showed a reduced DNA damage response after 5-azadC treatment. The coimmunoprecipitation and/or cofractionation (34,35) between DNMT1 and 3B further substantiate a direct interaction between these two DNMTs in cells that support the joint effect of the two putative adverse haplotypes in *DNMT1* and 3B found in this study. Although these two methyltransferases are best studied for their role in maintaining and modulating the distribution of 5-methylcytosines in the genome, our work supports a new and potentially equally important role of protecting the cell against DNA damage-induced chromosomal aberrations.

Supplementary material

Supplementary Table I and Figures 1–3 can be found at http://carcin.oxfordjournals.org/

Funding

National Institute of Health/National Cancer Institute (U01 CA097356); State of New Mexico as a direct appropriation from the Tobacco Settlement Fund.

Acknowledgements

Conflict of Interest Statement: None declared.

References

1. Hsu,T.C. *et al*. (1991) Mutagen sensitivity: a biological marker of cancer susceptibility. *Cancer Epidemiol. Biomarkers Prev.*, **1**, 83–89.
2. Wu,X. *et al*. (2007) Mutagen sensitivity: a genetic predisposition factor for cancer. *Cancer Res.*, **67**, 3493–3495.
3. Au,W.W. (2003) Mutagen sensitivity assays in population studies. *Mutat. Res.*, **544**, 273–277.
4. Wei,Q. *et al*. (1996) Benzo(*a*)pyrene diol epoxide-induced chromosomal aberrations and risk of lung cancer. *Cancer Res.*, **56**, 3975–3979.
5. Bondy,M.L. *et al*. (2001) Gamma-radiation sensitivity and risk of glioma. *J. Natl Cancer Inst.*, **93**, 1553–1557.
6. Wang,L.E. *et al*. (2005) *In vitro* sensitivity to ultraviolet B light and skin cancer risk: a case-control analysis. *J. Natl Cancer Inst.*, **97**, 1822–1831.
7. Wu,X. *et al*. (2006) Mutagen sensitivity has high heritability: evidence from a twin study. *Cancer Res.*, **66**, 5993–5996.
8. Lin,J. *et al*. (2007) Mutagen sensitivity and genetic variants in nucleotide excision repair pathway: genotype-phenotype correlation. *Cancer Epidemiol. Biomarkers Prev.*, **16**, 2065–2071.
9. Zhu,Y. *et al*. (2008) Modulation of DNA damage/DNA repair capacity by XPC polymorphisms. *DNA Repair (Amst.)*, **7**, 141–148.
10. Herman,J.G. *et al*. (2003) Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.*, **349**, 2042–2054.
11. Schermelleh,L. *et al*. (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res.*, **35**, 4301–4312.
12. Spada,F. *et al*. (2007) DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. *J. Cell Biol.*, **176**, 565–571.
13. Mortusewicz,O. *et al*. (2005) Recruitment of DNA methyltransferase I to DNA repair sites. *Proc. Natl Acad. Sci. USA*, **102**, 8905–8909.
14. Palii,S.S. *et al*. (2008) DNA methylation inhibitor 5-aza-2′-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B. *Mol. Cell Biol.*, **28**, 752–771.
15. Estève,P.O. *et al*. (2007) Molecular mechanisms of transactivation and doxorubicin-mediated repression of survivin gene in cancer cells. *J. Biol. Chem.*, **282**, 2615–2625.
16. Le Gac,G. *et al*. (2006) DNA damage-induced down-regulation of human Cdc25C and Cdc2 is mediated by cooperation between p53 and maintenance DNA (cytosine-5) methyltransferase 1. *J. Biol. Chem.*, **281**, 24161–24170.

17. Estève,P.O. *et al*. (2005) Human maintenance DNA (cytosine-5)-methyltransferase and p53 modulate expression of p53-repressed promoters. *Proc. Natl Acad. Sci. USA*, **102**, 1000–1005.
18. Wang,Y.A. *et al*. (2005) DNA methyltransferase-3a interacts with p53 and represses p53-mediated gene expression. *Cancer Biol. Ther.*, **4**, 1138–1143.
19. Tesfaigzi,Y. *et al*. (2006) Genotypes in matrix metalloproteinase 9 are a risk factor for COPD. *Int. J. Chron. Obstruct. Pulmon. Dis.*, **1**, 267–278.
20. Leng,S. *et al*. (2008) Double-strand break damage and associated DNA repair genes predispose smokers to gene methylation. *Cancer Res.*, **68**, 3049–3056.
21. Hsu,T.C. *et al*. (1989) Sensitivity to genotoxic effects of bleomycin in humans: possible relationship to environmental carcinogenesis. *Int. J. Cancer*, **43**, 403–409.
22. Lee,S.J. *et al*. (2005) DNMT3B polymorphisms and risk of primary lung cancer. *Carcinogenesis*, **26**, 403–409.
23. Wang,L. *et al*. (2006) A novel DNMT3B subfamily, DeltaDNMT3B, is the predominant form of DNMT3B in non-small cell lung cancer. *Int. J. Oncol.*, **29**, 201–207.
24. Barrett,J.C. *et al*. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
25. Gauderman,W.J. *et al*. (2007) Testing association between disease and multiple SNPs in a candidate gene. *Genet. Epidemiol.*, **31**, 383–395.
26. Stephens,M. *et al*. (2001) A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.*, **68**, 978–989.
27. Stephens,M. *et al*. (2003) A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.*, **73**, 1162–1169.
28. Chen,B.E. *et al*. (2006) Resampling-based multiple hypothesis testing procedures for genetic case-control association studies. *Genet. Epidemiol.*, **30**, 495–507.
29. Wu,X. *et al*. (1998) A parallel study of *in vitro* sensitivity to benzo[*a*]pyrene diol epoxide and bleomycin in lung carcinoma cases and controls. *Cancer*, **83**, 1118–1127.
30. Zhang,J. *et al*. (2003) Genomewide distribution of high-frequency, completely mismatching SNP haplotype pairs observed to be common across human populations. *Am. J. Hum. Genet.*, **73**, 1073–1081.
31. Setiawan,V.W. *et al*. (2006) A systematic assessment of common genetic variation in CYP11A and risk of breast cancer. *Cancer Res.*, **66**, 12019–12025.
32. Bi,X. *et al*. (2005) DNA polymerase kappa is specifically required for recovery from the benzo[*a*]pyrene-dihydrodiol epoxide (BPDE)-induced S-phase checkpoint. *J. Biol. Chem.*, **280**, 22343–22355.
33. Al Rashid,S.T. *et al*. (2005) Evidence for the direct binding of phosphorylated p53 to sites of DNA breaks *in vivo*. *Cancer Res.*, **65**, 10810–10821.
34. Kim,G.D. *et al*. (2002) Co-operation and communication between the human maintenance and *de novo* DNA (cytosine-5) methyltransferases. *EMBO J.*, **21**, 4183–4195.
35. Datta,J. *et al*. (2003) Biochemical fractionation reveals association of DNA methyltransferase (Dnmt) 3b with Dnmt1 and that of Dnmt 3a with a histone H3 methyltransferase and Hdac1. *J. Cell. Biochem.*, **88**, 855–864.

Received March 12, 2008; revised May 8, 2008; accepted May 8, 2008