Published in final edited form as:

Am J Med Genet B Neuropsychiatr Genet. 2017 September ; 174(6): 608–618. doi:10.1002/ajmg.b.32544.

MTHFR Methylation Moderates the Impact of Smoking on DNA Methylation at AHRR for African American Young Adults

Steven R.H. Beach^{1,2}, Man Kit Lei², Mei Ling Ong², Gene H. Brody², Meeshanthini V. Dogan³, and Robert A. Philibert, MD PhD⁴

¹Department of Psychology, University of Georgia

²Center for Family Research, University of Georgia

³College of Engineering, University of Iowa, Iowa City, IA, USA

⁴Department of Psychiatry, University of Iowa, Iowa City, IA USA

Abstract

Smoking has been shown to have a large, reliable and rapid effect on demethylation of *AHRR*, particularly at cg05575921, suggesting that methylation may be used as an index of cigarette consumption. Because the availability of methyl donors may also influence the degree of demethylation in response to smoking, factors that affect the activity of methylene tetrahydrofolate reductase (*MTHFR*), a key regulator of methyl group availability, may be of interest. In the current investigation, we examined the extent to which individual differences in methylation of *MTHFR* moderated the association between smoking and demethylation at cg05575921 as well as at other loci on *AHRR* associated with a main effect of smoking. Using a discovery sample (AIM, $N=293$), and a confirmatory sample (SHAPE, $N=368$) of young adult African Americans, degree of methylation of loci in the first exon of *MTHFR* was associated with amplification of the association between smoking and *AHRR* demethylation at cg05575921. However, genetic variation at a commonly studied *MTHFR* variant, C677T, did not influence cg05575921 methylation. The significant interaction between *MTHFR* methylation and the smoking-induced response at cg05575921 suggests a role for individual differences in methyl cycle regulation in understanding the effects of cigarette consumption on genome wide DNA methylation.

Keywords

Smoking; Methylation; AHRR; MTHFR; Folate; African American

* Address correspondence and reprint requests to: Steven R. H. Beach, Center for Family Research, University of Georgia, 1095 College Station Road, Athens, GA 30602-4527; srhbeach@uga.edu.

Orcid:0000-0001-8053-7761

Dr. Philibert is the Chief Executive Officer and majority stockholder of Behavioral Diagnostics, Inc. (www.bdmethylation.com). He is also a co-owner of U.S. Patent 8,637,652 and 9,273,358 which cover the use of methylation at cg0557921 for the ascertainment of smoking status.

Dr. Beach has filed for a provisional patent to cover use of m*MTHFR* for quantification of smoking exposure and other potential applications.

Introduction

Smoking has enormous implications for population-level health in the US. Smoking consistently ranks as the single greatest preventable cause of morbidity and mortality in the U.S., with documented effects on increased risk for cancer, type 2 diabetes, chronic obstructive pulmonary disease, and obesity [Centers for Disease Control and Prevention 2008]. Smoking causes nearly half a million US deaths per year along with higher population rates of preventable, serious illness [Mokdad and others 2004]. This burden is not distributed equally. Among US residents, people of African descent suffer worse outcomes for smoking-related illnesses [Centers for Disease Control and Prevention 2010, November 26; Haiman and others 2006], even after adjusting for common covariates such as socioeconomic status (SES) and healthcare access. This suggests that examination of African American samples could create a better understanding of other factors which interact with smoking to alter future risk for morbidity and mortality.

The use of cg05575921 as an indicator of smoking intensity

A significant barrier to a better understanding of the impact of smoking on health outcomes has been the lack of good biomarkers of cigarette consumption. However, recently several groups have demonstrated that smoking is associated with changes in DNA methylation [Breitling and others 2011; Joubert and others 2012; Monick and others 2012; Philibert and others 2012; Philibert and others 2013; Shenker and others 2013; Zeilinger and others 2013]. These studies have shown that longer-term smoking in older adults is associated with widespread and extensive remodeling of DNA methylation, particularly at loci associated with inflammation or toxin metabolism. In contrast, the extent of genome wide remodeling is much less extensive for nascent smokers or those with less smoking history [Philibert and others 2012; Philibert and others 2013], with changes often confined to the aryl hydrocarbon receptor repressor (*AHRR*). This observation led to the proposal that measurement of a single CpG residue on *AHRR* (cg05575921) might suffice for accurate detection of nascent smoking in many cases, allowing an enhanced role for testing in pediatric care [Beach and others 2015]. In addition, this observation suggests that examination of factors influencing smoking-related demethylation may be simplified among those with recent onset because attention can be focused on *AHRR*.

The biochemical basis for sensitivity to cigarette smoking of methylation status at cg05575921 is the large amounts of dioxin and poly aromatic hydrocarbons (PAH) found in tobacco smoke [Talhout and others 2011]. When these and other toxins from cigarette smoke are inhaled and absorbed by blood passing through the lungs, there is a robust cellular response by white blood cells and pulmonary macrophages. This response includes activation of the aryl hydrocarbon receptor (AHR), a nuclear receptor which is the key regulator of the xenobiotic response, leading to activation of CYP1A1, CYP1A2 and CYP1B1. Because the unrestrained activity of these and other cytochromes can result in generation of reactive oxygen species (ROS), catabolism of key cellular components, and potentially carcinogenesis [Ramadoss and others 2005], the xenobiotic response system is under tight regulatory control. A second nuclear receptor, *AHRR*, serves as a feedback modulator of the effects of AHR activation, competing for the obligate heterodimer partner

of AHR, the aryl hydrocarbon nuclear translocator, and for DNA binding motifs for AHR. The enhancer region of *AHRR* that includes cg05575921 plays a key role in this increase in *AHRR* activity. [Zeilinger and others 2013] have shown the cg05575921 demethylation is associated with the disassociation of the repressive protein complex and the recruitment of transcriptional activators. As a result, excessive activity of the xenobiotic pathway is avoided. One direct consequence of this regulatory cycle is that methylation of *AHRR*, and specifically the region around cg05575921, is an exquisitely sensitive indicator of smoking intensity.

Do individual differences influence *AHRR* response to smoking?

Our understanding of the methylomic response to smoking is still incomplete. In particular, one question left unresolved by prior research is whether individual differences in folate-metabolism regulation may be consequential for degree of, or rapidity of, demethylation in response to smoking. Activated methyl groups are required for DNA methylation, regulation of chromatin structure, and maintenance of DNA methylation patterns when cells divide. Accordingly, individual differences affecting availability of activated methyl groups may, in turn, affect maintenance of DNA methylation patterns in replicating cells [Jones and Liang 2009]. Because smoking is associated with substantial reductions in plasma folate [Gabriel and others 2006; Okumura and Tsukamoto 2011] as well as in red blood cells [Piyathilake and others 1992; Walmsley and others 1999] and buccal cells [Piyathilake and others 1992; Piyathilake and others 1995], even after correcting for folate intake [Piyathilake and others 1995], it would be expected that most smoking effects on methylation would be in the direction of demethylation. If so, individual differences that further reduce availability of activated methyl groups could potentially amplify the long-term effect of smoking on genomewide demethylation [Jacob and others 1998], and potentially amplify as well the initial response of nascent smokers at *AHRR*. Identifying sources of individual variability in response to smoking could enhance the ability of methylation based bio-markers to even more sensitively identify smokers as well as correctly quantify level of cigarette consumption based on methylation markers.

The role of *MTHFR* in Methylation Maintenance

Because methylene tetrahydrofolate reductase (*MTHFR*) is the rate-limiting enzyme in the methyl cycle, a pathway that is critical to folate metabolism, it is a natural focus of attention as a source of potential individual differences that may be consequential for availability of activated methyl groups. Smooth functioning of this cycle maintains adequate levels of S-adenosylmethionine (SAM) in cells, providing a readily available methyl group donor for numerous methylation related reactions [Stover 2009], especially genomic methylation. Conceivably, individual differences that are associated with reduced *MTHFR* activity could lead to reductions in DNA methylation repair over time or increase the rate of loss of methylation in response to stimuli associated with demethylation such as cigarette smoking. Conversely, when *MTHFR* is more active, methylation of homocysteine should be enhanced, potentially blunting the impact of external factors like smoking that would otherwise lead to demethylation [Stover 2009].

The *MTHFR* gene is located on chromosome 1, region 1p36.3, and contains 11 exons and three alternative transcription start sites, all in the vicinity of the first exon [Frosst and others 1995; Goyette and others 1994]. Several polymorphisms that may potentially influence *MTHFR* activity have been identified, with one of the most commonly cited variants being rs1801133, a C to T substitution at position 677 (C677T) that converts a codon for alanine to one for valine (167A to 167V) [Sibani and others 2000]. The 167V isoform of *MTHFR* is less active, more thermolabile and is associated with hypohomocysteinemia [Goyette and Rozen 2000]. Not all ethnicities are affected equally; rs1801133 is in strong population disequilibrium, with the frequency of T allele in African Americans (~20%) being about half that for those of Northern European ancestry [Sherry and others 2001].

MTHFR and Folate Metabolism in Disease

Lower activity of *MTHFR* reduces the synthesis of 5-methylTHF and hence remethylation of homocysteine, an outcome associated with global hypomethylation [Castro and others 2004; Yi and others 2000]. In tissue culture, low folic acid levels lead to increased formation of micronuclei, indicating increased rates of chromosome breakage. Further, the low activity *MTHFR* TT genotype also leads to increased micronuclei formation under low folate conditions [Kimura and others 2004], and is associated with global DNA hypomethylation in circulating leukocytes, perhaps in interaction with level of folate [Friso and others 2002; Stern and others 2000]. Not surprisingly, variation in folate metabolism also has been linked to increased vulnerability for a number of disease states, including risk of cardiovascular disease, multiple cancers, and neural tube defects [Erickson 2002; Holmquist and others 2003; Lamprecht and Lipkin 2003]. In addition, the *MTHFR* TT genotype has been shown to increase risk for a wide variety of neuropsychiatric disorders including depression and schizophrenia [Gilbody and others 2007; Lewis and others 2005] as well as a broad behavioral phenotype for psychopathology [Peerbooms and others 2011]. Accordingly, there is good reason to hypothesize a role for *MTHFR*, folate metabolism, and their combination in better predicting DNA methylation in response to smoking, with potential implications for vulnerability to longer-term disease outcomes associated with smoking.

Therefore, taken as a whole, there is a substantial body of evidence that suggests that variation in methyl donor availability exists and that it could alter the epigenomic response to smoking. As a first step in the testing of that those possibilities, we examine the relationship between genetic and epigenetic differences in *MTHFR* and demethylation of cg05575921 in response to smoking.

Specific Hypotheses

H1: Early cumulative smoking will be associated with changes in cg05575921 methylation, *even after controlling* for the impact of sex, cell type variation, SES, methylation of *MTHFR*, and in the SHAPE sample genotype at *MTHFR*^{C677T}.

H2: Because it should down-regulate activity, greater methylation of the first exon of *MTHFR* will moderate (i.e., intensify) the association of smoking with demethylation of cg05575921, even after controlling for the main effect of smoking and *mMTHFR*, and the effect of cell-type variability, sex, SES-risk, and in the SHAPE sample, genotypic variation at *MTHFR*^{C677T}.

H3: Effects will replicate after removing potential false negatives, or those very recently initiating smoking, i.e., those denying smoking across all three earlier waves of self-report data collection, but who test positive for cotinine at the time of the blood draw.

H4: Patterns seen at cg05575921 will replicate at other loci on the AHRR gene previously found to be associated with smoking: cg21161138 and cg26703534.

Method

Overview

The data utilized are from two longitudinal research projects: the Strong African American Healthy Adult Project (SHAPE; $n = 322$; see [Brody and others 2013] for a full description) and the Adults in the Making project (AIM; $n = 294$; see [Brody and others 2012] for a full description). All procedures and protocols conducted in these studies were approved by the University of Georgia Institutional Review Board.

Initial Sample

For the AIM sample, 496 youth were randomly recruited from public-school lists in six rural Georgia counties; none of which overlapped with those included in SHAPE. Again, the recruitment strategy was designed to obtain a representative sample of African American youth and families. The primary caregivers worked an average of 38.5 hr per week; 42.0% of the youth lived below federal poverty standards, with a median family income of \$2,016 per month. Of the 376 youth with complete self-report data at ages 17–19, 293 (77.9%) provided blood samples at age 22. These 293 participants (107 men and 186 women) constituted our AIM sample. Comparisons of their families with the AIM families not included in the present study revealed no differences on any variables.

Replication Sample

The SHAPE participants and their families were drawn from nine rural counties in Georgia, and so resided in small towns and communities in which poverty rates are among the highest in the nation and unemployment rates are above the national average [DeNavas-Walt and Proctor 2014]. Participants were selected randomly from lists of students that schools provided. Recruitment procedures were designed to obtain a sample of African American youth that was representative of the communities from which it was drawn; no other selection criteria were used. For the 399 (79.8%) who provided blood samples, primary caregivers worked an average of 39.9 hr per week, 42% of the participants lived below federal poverty standards, and they had a median family income of \$1,732 per month. The sample for the present study comprises 368 participants (167 men and 201 women) for whom blood was drawn at age 20 and for whom data was successfully genotyped for *MTHFR* polymorphism. Comparisons between these 368 participants and the 31 participants who were excluded because of incomplete data on genotyping revealed no significant differences on any of the key study variables or confounder variables. Fifty-six participants reported no cigarette use from ages 17–19, but subsequently were found to have cotinine levels > 1 .

For both samples, all data were collected in participants' homes using standardized protocols. Interviews were conducted privately, with no other family members present or able to overhear the conversation. Participants were compensated \$100 at each wave of data collection. Informed-consent forms were completed at all data-collection points.

Self-reported smoking and potential confounder variables were assessed in SHAPE and AIM when the youth were approximately 17, 18, and 19 years of age (for AIM, average ages were Wave 1 = 17.05; Wave 2 = 18.48; Wave 3 = 19.12; for SHAPE, average ages were Wave 1 = 17.01; Wave 2 = 17.58; Wave 3 = 18.44). Subsequently, participants in each cohort provided antecubital blood samples (drawn at age 20 in SHAPE and at age 21 in AIM), and these were used to assess methylation genomewide as well as to check cotinine for secondary analyses excluding potential false negatives and recent initiators.

Measures

SES-Risk

Caregiver reports collected at the same time as youth reports of smoking were used to create our measure of concurrent Socio-economic risk. SES risk was assessed across six indicators. Each indicator was scored dichotomously (0 if absent, 1 if present). Cumulative SES risk was defined as the average number of risk factors across the three assessments, yielding an index with a theoretical range of 0 to 6 ($M = 2.33$, $SD = 1.35$). The six risk indicators were (a) family poverty, defined as being below the poverty level, taking into account both family income and number of family members; (b) primary caregiver non-completion of high school or an equivalent; (c) primary caregiver unemployment; (d) single-parent family structure; (e) family receipt of Temporary Assistance for Needy Families; and (f) income rated by the primary caregiver as not adequate to meet all needs.

Cigarette consumption

At each wave of data collection in both SHAPE and AIM participants were asked "In the past month, how much did you smoke cigarettes?" Response options included: 0 – None at all; 1 – Less than 1 cigarette a day; 2 – 1 to 5 cigarettes a day; 3 – About a half a pack a day; 4 – About a pack a day; 5 – About 1 and a half packs a day; 6 – About 2 packs a day.

Methylation

Certified phlebotomists drew antecubital blood samples of whole blood (30 ml) from each participant and shipped it to a lab in Iowa the same day for preparation. At the lab the blood tubes were inspected to ensure anticoagulation and aliquots of blood were diluted, mononuclear cell pellets were separated from the diluted blood specimen by density-gradient centrifugation, and the mononuclear cell layer was removed from the tube using a transfer pipette, resuspended, and frozen at -80 degree C until use. Genomic DNA was prepared using a QiaAmp (Qiagen, Hilden, Germany) according to manufacturer's directions. A typical DNA yield for each mononuclear cell pellet was between 10 and 15 μ g.

The Illumina (San Diego, CA, USA) HumanMethylation450 Beadchip was used to assess genome-wide DNA methylation. Participants were randomly assigned to 12 sample "slides/

chips” with groups of 8 slides being bisulfite converted in a single batch, resulting in five “batches/plates.” A replicated sample of DNA was included in each plate to aid in assessment of batch variation and to ensure correct handling of specimens. The replicated sample was examined for average correlation of beta values between plates, resulting in average correlations greater than 0.99. Prior to normalization, methylation data were filtered based on these criteria: (1) samples containing 1% of CpG sites with detection p-value > 0.05 were removed, (2) sites were removed if a beadcount of < 3 was present in 5% of samples and (3) sites with a detection p-value of > 0.05 in 1% of samples were removed. More than 99.76% of the 485,577 probes yielded statistically reliable data. Probes associated with known SNPs or close to SNPs are annotated within Illumina allowing the presence of SNPs in regions of interest to be examined.

Quantile Normalization of Methylation data

Quantile normalization methods with separate normalization of Type I and Type II assays in the Illumina array produce marked improvement in detection of relationships by correcting distributional problems inherent in the manufacturers default method for calculating β (i.e., $\beta = M/(M + U + 100)$; where M and U are methylated and unmethylated signal intensities, respectively [Pidsley and others 2013]. Accordingly, in the current investigation all loci across all plates were quantile normalized concurrently, separating methylated and unmethylated intensities, and using the wateRmelon R package [Pidsley and others 2013] to institute the “*dasen*” function. This method equalizes the backgrounds of Type I and Type II probes prior to normalization and includes between-array normalization of Type I and Type II probes separately.

Identifying and Correcting for Chip and Batch effects

As demonstrated by Sun and associates [Sun and others 2011], quantile normalization typically reduces, but may not eliminate, batch and chip effects. Accordingly, after cleaning and quantile normalizing the data, all samples were examined for batch and chip effects. The distribution of quantile normalized average β values for all samples in each chip and batch were contrasted with all others using a box and density plot to indicate both the mean and confidence intervals around the mean in each case. The results of this examination indicated that both batch and chip effects were eliminated through quantile normalization. Absence of plate effects was confirmed via direct examination of the sample replicated across plates.

Index of m*MTHFR*

Our understanding of the epigenetic regulation of gene expression is still in its infancy. However, [Brenet and others 2011] have shown that methylation of the first exon is particularly important, with hypomethylation of the first exon being associated with a relatively large effect on gene expression (LOR = -2.8) and experimentally manipulated demethylation at the first exon also producing a large effect. To create our index of methylation of *MTHFR* (m*MTHFR*), we utilized all loci annotated as being associated with the first exon of *MTHFR*. The resulting set of seven loci (cg02978542, cg08269394, cg12751404, cg14032528, cg23068701, cg23226134, cg23952195) were all also annotated as being on the CpG island for *MTHFR* and as being promoter associated. The composite index of *MTHFR* methylation was calculated by averaging the standardized scores of the

seven retained CpGs following quantile normalization. That is, $mMTHFR$ = mean (zcg02978542, zcg08269394, zcg12751404, zcg14032528, zcg23068701, zcg23226134, zcg23952195). So, all loci contributed equally to the index of $mMTHFR$. In the supplemental material we also present results using a mean index of all 24 loci annotated as being associated with $MTHFR$.

Assessing and Controlling Proportion of Cell types in Mixed Cell Populations

Mononuclear cell pellets of the sort used in the current investigation are composed of several different cell types (e.g., primarily T-helper and cytotoxic cells, monocytes, B cells, and natural killer (NK) cells [Reinius and others 2012]. Accordingly, we controlled for individual differences in cell types using an approach developed by [Houseman and others 2012] through the “EstimateCellCounts” function in the minfi Bioconductor package.

Genotyping

Youths’ participating in SHAPE were genotyped for $MTHFR^{C677T}$, a mutation associated with decreased activity of $MTHFR$ [Frosst and others 1995]. DNA was obtained at age 16 using Oragene™ DNA kits (DNA Genotek, Inc.; Ottawa, Ontario, Canada). Youths rinsed their mouths with tap water, then deposited 4 ml of saliva in the Oragene sample vial. The vial was sealed, inverted, and shipped via courier to a central laboratory in Iowa City, where samples were prepared according to the manufacturer’s specifications. The prepared DNA was robotically dispensed into 384 optimal PCR trays. Subsequently, they were genotyped for the $MTHFR^{C677T}$ (rs1801133) using standard Taqman® reagents (Applied Biosystems [ABI], Foster City, CA, USA) and an existing ABI 7900 HT Genotyping system. Of the 374 participants contributing saliva, successful genotyping for $MTHFR^{C677T}$ was achieved for 368 participants.

Plan of analysis

We utilized beta regression analyses in all cases because the dependent variable (degree of methylation) is calculated as a ratio of methylated to unmethylated loci ranging from zero to one, and so typically fails the assumption of having a normal distribution [Dolzhenko and Smith 2014]. Beta regressions, utilizing the beta distribution, were introduced by [Ferrari and Cribari-Neto 2004], and are useful when the dependent variable is bounded, or when it has a ceiling or floor effect, introducing a non-normal distribution for the dependent variable.

On step 1 of each beta regression, we entered the main effect of cumulative self-reported smoking across three years in early adulthood. Accordingly, we examined the impact of self-reported smoking in a multivariate context, controlling for the effect of sex, cell-type variation, and in the case of the SHAPE sample, genotype at $MTHFR^{C677T}$. On the second step we examined the additional main effect of $mMTHFR$, again accounting for multivariate influences, and on the third step we examined the effect of adding the interaction term, $mMTHFR \times$ cumulative smoking, and in the case of the SHAPE sample, we also added the interaction term of genotype at $MTHFR^{C677T} \times$ cumulative smoking. This basic series of steps is repeated across both the SHAPE and AIM samples as well as for the reduced samples resulting from the removal of potential false negatives, i.e., anyone reporting no

cigarette use, but nonetheless, found to have elevated cotinine at the time of the blood draw. In the supplemental materials we provide similar regression analyses for the two additional loci on *AHRR* (cg21161138 and cg26703534) that typically show main effect associations with nascent smoking. We also report parallel effects for an index comprising all 24 loci annotated as being associated with *AHRR* and an index comprising the 18 *AHRR* loci that have no annotated local SNP variation, minimizing potential cis effects.

Results

Hardy-Weinberg

Genotype at *MTHFR*^{C677T} was determined for each youth as previously described. Of the sample, 1.1% were homozygous for the low activity T allele (TT), 17.1% were heterozygous (CT), and 81.8% were homozygous for the high activity C allele (CC). The distribution of alleles did not deviate from Hardy–Weinberg equilibrium ($p = .731$, ns).

H1: Early cumulative smoking will be associated with changes in cg05575921 methylation

Table 1 presents means, SDs, and intercorrelations for all demographic and major study variables. As can be seen in Table 1, cumulative smoking across three years in late adolescence was correlated with cg05575921 in both samples. In the SHAPE sample, the *MTHFR*^{C677T} genotype was not associated with *mMTHFR* and, as expected, was found to be infrequent (18.2% had any T allele, 1.1% [4 cases only] were homozygous). Significant correlations were found between self-reported smoking and cg05575921 and sex, with marginal correlations with cell-type, and a non-significant association with SES-risk in SHAPE and a significant association in AIM.

As can be seen in Tables 2a and 2b, Column 1, even when covariates for sex, SES risk, and cell-type variation were included, cumulative smoking across three years in late adolescence continued to be strongly correlated with demethylation of cg05575921, indicating that the association of smoking with demethylation of cg05575921 is robust to covariates, including variation in cell-types comprising the PBMC pellets for each participant.

H2: Greater *mMTHFR* will moderate (i.e., intensify) the impact of smoking on demethylation of cg05575921

In Column 2 of Tables 2a and 2b, the main effect of *mMTHFR* was entered and exerted no significant main effect on methylation at cg05575921. In step 3, the interaction of *mMTHFR* and level of smoking was entered into the model, resulting in a significant interaction effect in each sample. The significant interaction effect is plotted in Figure 1a (for the AIM sample) and Figure 1b (for the SHAPE sample), showing that, as anticipated, greater methylation at *MTHFR* was associated with a **stronger** effect of smoking on demethylation of cg05575921 in response to cumulative smoking. As can be seen in table 2b, in the SHAPE sample there was no significant interaction effect of genetic variation at *MTHFR*⁶⁶⁷ with smoking, showing that variation at this locus did not account for the observed moderating effect of *mMTHFR* on the impact of smoking.

H3: Effects will replicate in the reduced sample that excludes those denying smoking despite testing positive for cotinine

To examine the effect of eliminating potential false negatives, we eliminated cotinine positive participants who denied smoking at all prior waves of data collection, thereby reducing potential error due to underreporting or recent onset of smoking. Accordingly, we replicated the findings reported above using only the cotinine verified participants in each sample (Table 3a for AIM and table 3b for SHAPE). In each case the results replicated exactly. There is a robust effect of smoking on methylation at cg05575921 controlling potential confounding variables, and the interaction of *mMTHFR* with smoking continues to be significant in the direction of an intensified demethylation response among those with greater methylation of *MTHFR*. Graphical explication of the interaction effects for tables 3a and 3b, replicating the pattern previously portrayed in Figures 1a and 1b are available in the supplemental materials (Figure S1a and Figure S1b).

H4: Effects will replicate for other AHRR loci typically found to be associated with smoking: cg21161138 and cg26703534

To examine replication of a similar interaction pattern for the two other loci on *AHRR* previously found to be associated with nascent smoking [Philibert and others 2013], we re-ran regression analyses substituting methylation at these loci as outcomes in place of cg05575921. In both samples there was a significant main effect of self-reported smoking on demethylation at the additional *AHRR* loci. Significant moderation of smoking's impact attributable to *mMTHFR* was found for both additional *AHRR* loci in the AIM sample (See supplemental tables S1a and S2a) and for one of the additional loci (cg21161138) in the SHAPE sample (See supplemental table S1b). At the second locus (cg26703534), the interaction was marginal in the SHAPE sample ($p = .067$), albeit in the same direction as the other significant effects (see supplemental table S2b). In all cases, the significant interaction of *mMTHFR* with smoking resulted in an intensification of the demethylation response among those with greater methylation of *MTHFR*. The graphical explication of the significant interaction effects are available in the supplemental materials (Supplemental Figures S2a, S2b, S2c).

Discussion

The initiation of smoking prompts a strong response by the xenobiotic pathway, leading to demethylation, particularly at cg05575921, but also secondarily at other CpG sites, particularly those on *AHRR*. This effect has been replicated in a number of samples [Elliott and others 2014; Monick and others 2012; Philibert and others 2013; Shenker and others 2013; Zeilinger and others 2013; Zhang and others 2014] and appears to be robust across ethnicities [Dogan and others 2015]. This impact of smoking on cg05575921 methylation is not surprising given the key biological role played by *AHRR* in response to the types of toxins commonly found in cigarette smoke. The current investigation adds to this finding by showing for the first time that the impact of smoking on demethylation of cg05575921 is potentially qualified by an epigenetic regulatory motif: methylation of the first exon of *MTHFR*.

Despite the success of cg05575921 as an early and sensitive indicator of incipient smoking and changes in smoking status, it is not perfectly correlated with self-reported smoking, a limitation likely due, in part, to problems inherent in self-report (e.g., [Philibert and others 2013]). When self-report errors are minimized, for example, by excluding individuals whose cotinine level suggests they are smokers even though they deny smoking, the correlation of smoking with cg05575921 methylation typically increases. However, there is sufficient variability in methylation levels among youth reporting similar, low levels of smoking, to suggest the presence of other individual difference factors moderating the impact of nascent smoking experience on demethylation of cg05575921. Such sources of variability will be particularly important when accurate quantification of prior smoking exposure is a primary goal rather than simply differentiating smokers and non-smokers. As outlined above, a potential source of such variability is individual differences in activity of *MTHFR*.

In the current investigation we found replicated evidence that for African American youth, the impact of self-reported smoking on demethylation of cg05575921 was intensified when methylation of the first exon of *MTHFR* was elevated. This is significant theoretically because it suggests that in the context of a strong stimulus favoring demethylation, like cigarette smoking, it is possible to observe direct regulatory impact of *MTHFR* methylation patterns on the epigenetic impact of smoking. The interaction was predicted on the basis of expectations that methylation of a first exon would have a strong impact on reduced expression and hence reduced activity of a gene that is central to folate metabolism. Accordingly, lower methylation of the first exon of *MTHFR* was expected to result in better maintenance of methylation at cg05575921 and so less apparent effect of smoking. It is noteworthy that the same pattern was replicated for two other *AHRR* loci known to be influenced by smoking. In addition, we examined the effect using average methylation across all loci annotated as being associated with *MTHFR* regardless of location on the gene and found that the pattern was again replicated, suggesting that observation of the pattern does not depend on restricting attention to the first exon (See supplemental Tables S3a and S3b and Figures S3a and S3b). Finally, when six loci that were on or close to known SNPs on *MTHFR* (table S4) were dropped from the index, overall *MTHFR* methylation continued to moderate the impact of smoking on methylation at cg05575921 (see supplemental Table S5a and S5b as well as supplemental Figures S4a and S4b), suggesting that the moderating effect was not due to *cis* effects of cryptic genotypic variation within *MTHFR*.

Some limitations of the current analyses should also be noted. First, given the centrality of the folate cycle to a broad range of biological activities, it is possible that the index of *mMTHFR* used in the current study is itself correlated with broader patterns of methylation, and so with variability in other gene regulatory motifs across the genome, resulting in coordinated pattern of change. If so, it may be useful in future research, using larger samples, to examine such broader networks. By necessity this will include both broader sampling of methylation of other folate cycle genes, sampling of sera to determine folate levels, consideration of dietary factors including alcohol intake, and possibly gut biome interactions. Because many of these potentially independent influences will be correlated, the study of cohorts for whom multiple levels of biological data are available should be emphasized. Likewise, no mRNA was available for the current samples and so we were not able to examine the hypothesized effect of *mMTHFR* on gene activity.

Second, the current analyses are limited to short-term effects, and it would be informative to replicate the current results in older samples with longer histories of smoking to better understand the boundaries of the observed interaction effects. Likewise, given the widely varying frequency of genetic motifs influencing *MTHFR* activity, it will be useful to examine whether the same patterns are seen in samples with a greater proportion of carriers of the *MTHFR*^{C677T} polymorphism or other polymorphisms influencing *MTHFR* gene activity, e.g. Caucasian or Asian populations. Finally, it would be useful theoretically to tease apart the relative impact of recent vs cumulative smoking on demethylation of cg05575921 to determine whether both components are equally affected by *MTHFR* activity. A related question is whether methylation of *MTHFR* may moderate the response of cg05575921 to smoking cessation (i.e. rapidity of reversion). Fortunately, because the loci utilized in the current investigation are included in the Illumina 450K array, the hypotheses advanced in the current study can be examined in any data set that includes both an assessment of smoking and Illumina 450K array data.

With regard to practical implications for assessment, the current results suggest that when demethylation of cg05575921 is used as a basis to quantify recent smoking history, it may be useful to examine m*MTHFR* as well. In the current investigation, the interaction term accounted for between 2.1 and 1.5% of the variance in demethylation of cg05575921 in response to smoking for the SHAPE and AIM cotinine verified samples respectively. Likewise, for all levels of consumption beyond zero and “less than one per day,” i.e., the two lowest levels of self-reported cigarette consumption, the level of demethylation in response to smoking differed significantly for those with higher vs. lower m*MTHFR*. For example, in the cotinine verified SHAPE sub-sample, someone with low m*MTHFR* who was found to have 80% methylation of cg0557592, would be estimated to have been a heavy smoker, over the past several years, with cigarette consumption in the range of “1 and a half packs a day”; which is a 5 on our self-report scale. Conversely, someone with high m*MTHFR* who was found to have the same level of methylation at cg0557592 would be estimated to be only a light smoker, with cigarette consumption in the range of “1 to 5 cigarettes a day”; a two on our self-report scale. Accordingly, there are potentially large differences in estimated cigarette exposure secondary to this regulatory motif, and the current results suggest the potential for improved quantification of smoking exposure given a particular value of cg0557592.

If the current findings are replicated more broadly, including replication in samples of longer-term smokers, it will be of interest to see if smoking’s broader effects on demethylation genome-wide are also moderated by regulatory motifs influencing the folate cycle. If so, such variability will be an attractive target for explication of individual differences in morbidity and mortality that are associated with smoking exposure. Specifically, it may be that cg05575921 demethylation can serve as a “canary in the coal mine,” providing an early forecast of patterns of epigenetic change likely to emerge with continued smoking, helping to identify individuals at greatest risk for smoking associated health complications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by 5R01HD030588-16A1 (Beach and Simons, PI), 1P30DA027827 from the National Institute on Drug Abuse (Brody PI) and R01DA037648 (Philibert and Gerrard, PI).

References

- Beach SRH, Gerrard M, Gibbons FX, Brody GH, Philibert RA. A role for epigenetics in broadening the scope of pediatric care in the prevention of adolescent smoking. *Epigenetic Diagnosis and Therapy*. 2015; 1(2):91–97. [PubMed: 27500081]
- Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *American Journal of Human Genetics*. 2011; 88(4): 450–457. [PubMed: 21457905]
- Brenet F, Moh M, Funk P, Feierstein E, Viale AJ, Socci ND, Scandura JM. DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS ONE*. 2011; 6(1):e14524. [PubMed: 21267076]
- Brody GH, Beach SRH, Hill KG, Howe GW, Prado G, Fullerton SM. Using genetically informed, randomized prevention trials to test etiological hypotheses about child and adolescent drug use and psychopathology. *American Journal of Public Health*. 2013; 103(S1):S19–S24. [PubMed: 23927515]
- Brody GH, Yu T, Chen Y-F, Kogan SM, Smith K. The Adults in the Making program: Long-term protective stabilizing effects on alcohol use and substance use problems for rural African American emerging adults. *Journal of Consulting and Clinical Psychology*. 2012; 80(1):17–28. [PubMed: 22182263]
- Castro R, Rivera I, Ravasco P, Camilo ME, Jakobs C, Blom HJ, De Almeida IT. 5, 10-methylenetetrahydrofolate reductase (MTHFR) 677C→T and 1298A→C mutations are associated with DNA hypomethylation. *Journal of Medical Genetics*. 2004; 41:454–458. [PubMed: 15173232]
- Centers for Disease Control and Prevention. Smoking-attributable mortality, years of potential life lost, and productivity losses—United States, 2000–2004. *MMWR Morb Mortal Wkly Rep*. 2008; 57(45): 1226–1228. [PubMed: 19008791]
- Centers for Disease Control and Prevention. Racial disparities in smoking-attributable mortality and years of potential life lost—Missouri, 2003–2007. *Morbidity and Mortality Weekly Report*. 2010 Nov 26; 59(46):1512–1522.
- DeNavas-Walt, C., Proctor, BD. Income and poverty in the United States: 2013. Washington, DC: U.S. Census Bureau; 2014. Current Population Reports 60-249
- Dogan MV, Xiang J, Beach SR, Cutrona C, Gibbons FX, Simons RL, Brody GH, Stapleton JT, Philibert RA. Ethnicity and Smoking-Associated DNA Methylation Changes at HIV Co-Receptor GPR15. *Frontiers in psychiatry*. 2015; 6
- Dolzhenko E, Smith AD. Using beta-binomial regression for high-precision differential methylation analysis in multifactor whole-genome bisulfite sequencing experiments. *BMC Bioinformatics*. 2014; 15(1):1–8. [PubMed: 24383880]
- Elliott H, Tillin T, McArdle W, Ho K, Duggirala A, Frayling T, Davey Smith G, Hughes A, Chaturvedi N, Relton C. Differences in smoking associated DNA methylation patterns in South Asians and Europeans. *Clin Epigenetics*. 2014; 6(1):4. [PubMed: 24485148]
- Erickson JD. Folic acid and prevention of spina bifida and anencephaly. 10 years after the US Public Health Service recommendation. *MMWR. Recommendations and reports: Morbidity and mortality weekly report Recommendations and reports/Centers for Disease Control*. 2002; 51:1–3.
- Ferrari SL, Cribari-Neto F. Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics*. 2004; 31(7):799–815.

- Friso S, Choi SW, Girelli D, Mason JB, Dolnikowski GG, Bagley PJ, Olivieri O, Jacques PF, Rosenberg IH, Corrocher R, Selhub J. A common mutation in the 5, 10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status. *Proceedings of the National Academy of Sciences*. 2002; 99:5606–5611.
- Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, Matthews RG, Boers GJH, den Heijer M, Kluijtmans LAJ, van den Heuvel LPWJ, Rozen R. A candidate genetic risk factor for vascular disease: A common mutation in methylenetetrahydrofolate reductase. *Nature Genetics*. 1995; 10(1):111–113. [PubMed: 7647779]
- Gabriel HE, Crott JW, Ghandour H, Dallal GE, Choi S-W, Keyes MK, Jang H, Liu Z, Nadeau M, Johnston A, Mager D, Mason JB. Chronic cigarette smoking is associated with diminished folate status, altered folate form distribution, and increased genetic damage in the buccal mucosa of healthy adults. *The American Journal of Clinical Nutrition*. 2006; 83(4):835–841. [PubMed: 16600936]
- Gilbody S, Lewis S, Lightfoot T. Methylenetetrahydrofolate Reductase (MTHFR) Genetic Polymorphisms and Psychiatric Disorders: A HuGE Review. *Am J Epidemiol*. 2007; 165(1):1–13. [PubMed: 17074966]
- Goyette P, Rozen R. The thermolabile variant 677C→T can further reduce activity when expressed in CIS with severe mutations for human methylenetetrahydrofolate reductase. *Hum Mutat*. 2000; 16(2):132–138. [PubMed: 10923034]
- Goyette P, Sumner JS, Milos R, Duncan AMV, Rosenblatt DS, Matthews RG, Rozen R. Human methylenetetrahydrofolate reductase: isolation of cDNA, mapping and mutation identification. *Nat Genet*. 1994; 7(2):195–200. [PubMed: 7920641]
- Haiman CA, Stram DO, Wilkens LR, Pike MC, Kolonel LN, Henderson BE, Le Marchand L. Ethnic and Racial Differences in the Smoking-Related Risk of Lung Cancer. *N Engl J Med*. 2006; 354(4):333–342. [PubMed: 16436765]
- Holmquist C, Larsson S, Wolk A, de Faire U. Multivitamin supplements are inversely associated with risk of myocardial infarction in men and women—Stockholm Heart Epidemiology Program (SHEEP). *The Journal of Nutrition*. 2003; 133:2650–2654. [PubMed: 12888653]
- Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012; 13 (–):Article 86.
- Jacob RA, Gretz DM, Taylor PC, James SJ, Pogribny IP, Miller BJ, Henning SM, Swendseid ME. Moderate Folate Depletion Increases Plasma Homocysteine and Decreases Lymphocyte DNA Methylation in Postmenopausal Women. *The Journal of Nutrition*. 1998; 128(7):1204–1212. [PubMed: 9649607]
- Jones PA, Liang G. Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet*. 2009; 10(11):805–811. [PubMed: 19789556]
- Joubert BR, Håberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, Huang Z, Hoyo C, Midttun Ø, Cupul-Uicab LA, Ueland PM, Wu MC, Nystad W, Bell DA, Peddada SD, London SJ. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environmental Health Perspectives*. 2012; 120(10):1425–1431. [PubMed: 22851337]
- Kimura M, Umegaki K, Higuchi M, Thomas P, Fenech M. Methylenetetrahydrofolate reductase C677T polymorphism, folic acid and riboflavin are important determinants of genome stability in cultured human lymphocytes. *The Journal of Nutrition*. 2004; 134:48–56. [PubMed: 14704292]
- Lamprecht SA, Lipkin M. Chemoprevention of colon cancer by calcium, vitamin D and folate: molecular mechanisms. *Nature Reviews Cancer*. 2003; 3:601–614. [PubMed: 12894248]
- Lewis SJ, Zammit S, Gunnell D, Smith GD. A meta-analysis of the MTHFR C677T polymorphism and schizophrenia risk. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2005; 135:2–4.
- Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA*. 2004; 291(10):1238–1245. [PubMed: 15010446]
- Monick MM, Beach SRH, Plume JM, Sears R, Gerrard M, Brody GH, Philibert RA. Coordinated changes in AHRH methylation in lymphoblasts and pulmonary macrophages from smokers.

- American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2012; 159B(2):141–151.
- Okumura K, Tsukamoto H. Folate in smokers. *Clinica Chimica Acta*. 2011; 412:521–526.
- Peerbooms OL, van Os J, Drukker M, Kenis G, Hoogveld L, De Hert M, Delespaul P, van Winkel R, Rutten BP. Meta-analysis of MTHFR gene variants in schizophrenia, bipolar disorder and unipolar depressive disorder: evidence for a common genetic vulnerability? *Brain, behavior, and immunity*. 2011; 25:1530–1543.
- Philibert RA, Beach SRH, Brody GH. Demethylation of the aryl hydrocarbon receptor repressor as a biomarker for nascent smokers. *Epigenetics*. 2012; 7(11):1331–1338. [PubMed: 23070629]
- Philibert RA, Beach SRH, Lei M-K, Brody GH. Changes in DNA methylation at the aryl hydrocarbon receptor repressor may be a new biomarker for smoking. *Clinical Epigenetics*. 2013; 5(1) Article 19.
- Pidsley R, Wong CCY, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*. 2013; 14 (–):Article 293.
- Piyathilake CJ, Hine RJ, Dasanayake AP, Richards EW, Freeberg LE, Vaughn WH, Krumdieck CL. Effect of smoking on folate levels in buccal mucosal cells. *Int J Cancer*. 1992; 52(4):566–569. [PubMed: 1399138]
- Piyathilake CJ, Macaluso M, Hine RJ, Vinter DW, Richards EW, Krumdieck CL. Cigarette smoking, intracellular vitamin deficiency, and occurrence of micronuclei in epithelial cells of the buccal mucosa. *Cancer Epidemiology Biomarkers & Prevention*. 1995; 4(7):751–758.
- Ramados P, Marcus C, Perdew GH. Role of the aryl hydrocarbon receptor in drug metabolism. *Expert Opin Drug Metab Toxicol*. 2005; 1(1):9–21. [PubMed: 16922649]
- Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlén S-E, Greco D, Söderhäll C, Scheynius A, Kere J. Differential DNA methylation in purified human blood cells: Implications for cell lineage and studies on disease susceptibility. *PLoS ONE*. 2012; 7(7):e41361. [PubMed: 22848472]
- Shenker NS, Polidoro S, van Veldhoven K, Sacerdote C, Ricceri F, Birrell MA, Belvisi MG, Brown R, Vineis P, Flanagan JM. Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Human Molecular Genetics*. 2013; 22(5):843–851. [PubMed: 23175441]
- Sherry ST, Ward M-H, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001; 29(1):308–311. [PubMed: 11125122]
- Sibani S, Christensen B, O’Ferrall E, Saadi I, Hiou-Tim F, Rosenblatt DS, Rozen R. Characterization of six novel mutations in the methylenetetrahydrofolate reductase (MTHFR) gene in patients with homocystinuria. *Hum Mutat*. 2000; 15(3):280–287. [PubMed: 10679944]
- Stern LL, Mason JB, Selhub J, Choi SW. Genomic DNA hypomethylation, a characteristic of most cancers, is present in peripheral leukocytes of individuals who are homozygous for the C677T polymorphism in the methylenetetrahydrofolate reductase gene. *Cancer Epidemiology Biomarkers & Prevention*. 2000; 9:849–853.
- Stover PJ. One-Carbon Metabolism–Genome Interactions in Folate-Associated Pathologies. *The Journal of Nutrition*. 2009; 139(12):2402–2405. [PubMed: 19812215]
- Sun Z, Chai HS, Wu Y, White WM, Donkena KV, Klein CJ, Garovic VD, Therneau TM, Kocher J-PA. Batch effect correction for genome-wide methylation data with Illumina Infinium platform. *BMC Medical Genomics*. 2011; 4 (–):Article 84.
- Talhout R, Schulz T, Florek E, Van Benthem J, Wester P, Opperhuizen A. Hazardous compounds in tobacco smoke. *Int J Environ Res Public Health*. 2011; 8(2):613–628. [PubMed: 21556207]
- Walmsley C, Bates C, Prentice A, Cole T. Relationship between cigarette smoking and nutrient intakes and blood status indices of older people living in the UK: further analysis of data from the National Diet and Nutrition Survey of people aged 65 years and over, 1994/95. *Public Health Nutr*. 1999; 2(02):199–208. [PubMed: 10447248]
- Yi P, Melnyk S, Pogribna M, Pogribny IP, Hine RJ, James SJ. Increase in plasma homocysteine associated with parallel increases in plasma S-adenosylhomocysteine and lymphocyte DNA hypomethylation. *Journal of Biological Chemistry*. 2000; 275:29318–29323. [PubMed: 10884384]

- Zeilinger S, Kühnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, Weidinger S, Lattka E, Adamski J, Peters A, Strauch K, Waldenberger M, Illig T. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One*. 2013; 8(5):e63812. [PubMed: 23691101]
- Zhang Y, Yang R, Burwinkel B, Breitling LP, Holleccek B, Schöttker B, Brenner H. F2RL3 methylation in blood DNA is a strong predictor of mortality. *Int J Epidemiol*. 2014

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1a.

Explication of significant interaction effect in Table 2a for AIM sample. Beta regression examining effects of cigarette use at ages 17–19 on cgQ5575921 moderated by *mMTHFR* ($n = 293$) using all available participants.

Figure 1b.

Explication of significant interaction effect in Table 2b for SHAPE sample. Effects of cigarette use at ages 17–19 on cg05575921 moderated by *mMTHFR* ($n = 368$) using all available participants.

Table 1

Correlation matrix for the study variables

	1	2	3	4	5	6	7	8	9	10	11	Mean _{AIM}	SD _{AIM}
1. cg05575921	—	-.493 ^{**}	-.152 ^{**}	—	-.462 ^{**}	-.095	.170 ^{**}	.236 ^{**}	-.182 ^{**}	.087	-.060	.817	.064
2. Cigarette use at ages 17–19	-.559 ^{**}	—	-.009	—	.286 ^{**}	.133 [*]	-.018	-.053	.019	.099 [†]	-.033	.309	.730
3. m <i>MTHFR</i>	-.025	.035	—	—	.176 ^{**}	.008	-.205 ^{**}	-.209 ^{**}	.098 [†]	-.090	.063	.000	.425
4. <i>MTHFR</i> (1= C:C/C:T)	.000	-.034	-.085	—	—	—	—	—	—	—	—	—	—
5. Males	-.355 ^{**}	.259 ^{**}	.060	-.048	—	.050	-.157 ^{**}	-.156 ^{**}	.387 ^{**}	-.136 [*]	.052	.365	.482
6. SES-related risk at ages 17–19	-.020	.001	.014	.031	-.007	—	-.009	-.029	-.030	-.048	-.012	2.010	1.198
7. CD8+ T cells	.063	-.044	.005	-.086 [†]	-.114 [*]	-.075	—	.308 ^{**}	-.352 ^{**}	-.008	-.351 ^{**}	.156	.065
8. CD4+ T cells	.048	.022	-.101 [†]	.004	-.005	-.047	.292 ^{**}	—	-.294 ^{**}	.051	-.309 ^{**}	.280	.084
9. Natural killer	-.014	.002	.004	-.008	.242 ^{**}	.006	-.267 ^{**}	-.208 ^{**}	—	-.190 ^{**}	.086	.073	.074
10. B-cells	.116 [*]	-.098 [†]	-.064	.014	.000	-.010	.110 [*]	.196 ^{**}	-.124 [*]	—	.049	.151	.056
11. Monocytes	-.065	.034	.046	.065	-.033	-.004	-.096 [†]	-.219 ^{**}	.053	.012	—	.163	.085
Mean _{SHAPE}	.868	.369	.002	.182	.454	2.532	.151	.299	.080	.140	.137	—	—
SD _{SHAPE}	.056	.817	.432	.386	.499	1.300	.068	.094	.079	.058	.074	—	—

Note. Correlations for SHAPE data ($N = 368$) displayed below the diagonal; correlations for AIM data ($N = 293$) displayed above the diagonal. *MTHFR* polymorphism is not available for AIM data.^{**} $p < .01$;^{*} $p < .05$;[†] $p < .10$ (two-tailed tests).

Table 2a

AIM

Beta regression models depicting moderating effects of variations at the *mMTHFR* on *cg05575921*

	Model 1		Model 2		Model 3	
	β	$e\beta$	β	$e\beta$	β	$e\beta$
Cigarette use at ages 17–19	-.193** (.023)	.824	-.196** (.023)	.822	-.206** (.023)	.814
<i>mMTHFR</i>			-.055 (.044)	.946	-.002** (.047)	.998
Cigarette use at ages 17–19 $\times$ <i>mMTHFR</i>					-.151** (.048)	.860
Males	-.244** (.042)	.783	-.235** (.042)	.791	-.230** (.042)	.795
SES-related risk at ages 17–19	-.010 (.018)	.990	-.010 (.018)	.990	-.009 (.018)	.991
CD8+ T cells	.471 (.314)	1.602	.404 (.318)	1.498	.498 (.314)	1.645
CD4+ T cells	.648** (.234)	1.912	.599* (.237)	1.820	.611** (.232)	1.842
Natural killer	.143 (.280)	1.154	.115 (.280)	1.122	.209 (.277)	1.232
B-cells	.591 [†] (.329)	1.806	.564 [†] (.330)	1.758	.645 * (.325)	1.906
Monocytes	.040 (.228)	1.041	.024 (.228)	1.024	.072 (.224)	1.074
Constant	1.296** (.120)		1.327** (.122)		1.284** (.120)	
-2LL	-984.745		-986.280		-995.889	
Pseudo R-square	.138		.139		.147	

Notes: Unstandardized coefficients (β) and odds ratio ($e\beta$) shown with standard errors in parentheses; SES-related risk are standardized by z-transformation. $n = 293$.

[†] $p < .10$,

* $p < .05$,

** $p < .01$ (two-tailed tests).

Table 2b

SHAPE

Beta regression models depicting moderating effects of variations at the *mMTHFR* on cg05575921

	Model 1		Model 2		Model 3	
	β	$e\beta$	β	$e\beta$	β	$e\beta$
Cigarette use at ages 17–19	-.230** (.019)	.795	-.229** (.019)	.795	-.216** (.021)	.806
<i>mMTHFR</i>			.040 (.043)	1.041	.132** (.046)	1.141
Cigarette use at ages 17–19 $\times$ <i>mMTHFR</i>					-.271** (.051)	.763
<i>MTHFR</i> (I = C:C/C:T)	-.013 (.047)	.987	-.009 (.047)	.991	-.019 (.051)	.981
Cigarette use at ages 17–19 $\times$ <i>MTHFR</i>					.003 (.056)	1.003
Males	-.220** (.039)	.803	-.223** (.039)	.800	-.199** (.038)	.820
SES-related risk at ages 17–19	-.011 (.018)	.989	-.011 (.018)	.989	-.015 (.018)	.985
CD8+ T cells	.070 (.287)	1.073	.059 (.286)	1.061	-.068 (.277)	.934
CD4+ T cells	.175 (.208)	1.191	.194 (.209)	1.214	.208 (.204)	1.231
Natural killer	.312 (.250)	1.366	.316 (.250)	1.372	.189 (.241)	1.208
B-cells	.546 [†] (.321)	1.726	.559 [†] (.320)	1.749	.551 [†] (.313)	1.735
Monocytes	-.256 (.243)	.774	-.260 (.243)	.771	-.234 (.236)	.791
Constant	1.960** (.093)		1.955** (.093)		1.969** (.090)	
-2LL	-1352.351		-1353.244		-1379.941	
Pseudo R-square	.120		.121		.138	

Notes: Unstandardized coefficients (β) and odds ratio ($e\beta$) shown with standard errors in parentheses; SES-related risk are standardized by z-transformation. $n = 368$.[†] $p < .10$.

. (two-tailed tests) d

**

.50' d

*

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3a

AIM

Beta regression models depicting moderating effects of variation at the *mMTHFR* on cg05575921, excluding ($n = 40$) individuals who self-reported no cigarette use across three waves from 17 to 19, but had cotinine value > 1 .

	Model 1		Model 2		Model 3	
	β	$e\beta$	β	$e\beta$	β	$e\beta$
Cigarette use at ages 17–19	-.246** (.020)	.782	-.251** (.020)	.778	-.264** (.020)	.768
<i>mMTHFR</i>			-.092* (.039)	.912	-.020 (.042)	.980
Cigarette use at ages 17–19 $\times$ <i>mMTHFR</i>					-.172** (.040)	.842
Males	-.114** (.040)	.892	-.095* (.041)	.909	-.088* (.040)	.916
SES-related risk at ages 17–19	-.023 (.017)	.977	-.024 (.016)	.976	-.024 (.016)	.976
CD8+ T cells	.028 (.283)	1.028	-.077 (.284)	.926	.038 (.275)	1.039
CD4+ T cells	.555** (.209)	1.742	.485* (.209)	1.624	.499* (.201)	1.647
Natural killer	-.101 (.258)	.904	-.163 (.257)	.850	-.008 (.251)	.992
B-cells	.422 (.293)	1.525	.387 (.291)	1.473	.481 [†] (.282)	1.618
Monocytes	-.251 (.209)	.778	-.277 (.207)	.758	-.218 (.199)	.804
Constant	1.504** (.111)		1.560** (.111)		1.507** (.107)	
-2LL	-956.706		-962.055		-980.038	
Pseudo R-square	.171		.175		.190	

Notes: Unstandardized coefficients (β) and odds ratio ($e\beta$) shown with standard errors in parentheses; SES-related risk are standardized by z-transformation. $n = 253$.

[†] $p < .10$,

* $p < .05$,

two-tailed test(s).
 d
**

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3b

SHAPE

Beta regression models depicting moderating effects of variation at the *mMTHFR* on cg05575921, excluding ($n = 56$) individuals who self-reported no cigarette use across three waves from 17 to 19, but had cotinine value > 1 .

	Model 1		Model 2		Model 3	
	β	$e\beta$	β	$e\beta$	β	$e\beta$
Cigarette use at ages 17–19	-.261** (.018)	.770	-.261** (.018)	.770	-.248** (.019)	.780
<i>mMTHFR</i>			-.007 (.040)	.993	.093* (.043)	1.097
Cigarette use at ages 17–19 $\times$ <i>mMTHFR</i>					-.269** (.045)	.764
<i>MTHFR</i> (1 = C/C:C:T)	-.017 (.045)	.983	-.018 (.045)	.982	-.022 (.049)	.978
Cigarette use at ages 17–19 $\times$ <i>MTHFR</i>					-.003 (.049)	.997
Males	-.153** (.038)	.858	-.153** (.037)	.858	-.120** (.037)	.887
SES-related risk at ages 17–19	-.016 (.017)	.984	-.016 (.017)	.984	-.021 (.017)	.979
CD8+ T cells	.007 (.272)	1.007	.008 (.273)	1.008	-.133 (.260)	.875
CD4+ T cells	.078 (.198)	1.081	.074 (.199)	1.077	.075 (.191)	1.078
Natural killer	.232 (.243)	1.261	.232 (.243)	1.261	.065 (.232)	1.067
B-cells	.759* (.311)	2.136	.759* (.311)	2.136	.734* (.301)	2.083
Monocytes	-.276 (.231)	.759	-.275 (.231)	.760	-.252 (.222)	.777
Constant	2.014** (.089)		2.015** (.089)		2.037** (.086)	
-2LL	-1269.412		-1269.447		-1302.377	
Pseudo R-square	.155		.155		.176	

Notes: Unstandardized coefficients (β) and odds ratio ($e\beta$) shown with standard errors in parentheses; SES-related risk are standardized by z-transformation. $n = 321$.

(two-tailed tests)

d_{**}

d_{*}

d_{+}

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript