Published in final edited form as:

Circ Cardiovasc Genet. 2015 October ; 8(5): 707–716. doi:10.1161/CIRCGENETICS.115.001097.

DNA Methylation of the Aryl Hydrocarbon Receptor Repressor Associations with Cigarette Smoking and Subclinical Atherosclerosis

Lindsay M. Reynolds, PhD[1], **Ma Wan, MD, PhD**[2], **Jingzhong Ding, MD, PhD**[1], **Jackson R. Taylor, PhD**[1], **Kurt Lohman Mstat**[1], **Dan Su, MD, PhD**[2], **Brian D. Bennett, PhD**[2], **Devin K. Porter, BS**[2], **Ryan Gimple, BS**[2], **Gary S. Pittman, MSPH**[2], **Xuting Wang, PhD**[2], **Timothy D. Howard, PhD**[1], **David Siscovick, MD, MPH**[3], **Bruce M. Psaty, MD, PhD, MPH**[4], **Steven Shea, MD, MS**[5], **Gregory L. Burke, MD, MS**[1], **David R. Jacobs Jr., PhD**[6], **Stephen S. Rich, PhD**[7], **James E. Hixson, PhD**[8], **James H. Stein, MD**[9], **Hendrik Stunnenberg, PhD**[10], **R. Graham Barr, MD, DrPH**[5], **Joel D. Kaufman, MD, MPH**[4], **Wendy S. Post, MD, MS**[11], **Ina Hoeschele, PhD**[12], **David M. Herrington, MD, MHS**[1], **Douglas A. Bell, PhD**[2,*], and **Yongmei Liu, MD, PhD**[1,*]

[1]Dept of Epidemiology & Prevention, Division of Public Health Sciences, Ctr for Human Genomics, Depts of Internal Medicine & Gerontology and Geriatric Medicine, J. Paul Sticht Ctr on Aging, Wake Forest School of Medicine, Winston-Salem [2]National Inst of Environmental Health Sciences, NIH, Research Triangle Park, NC [3]New York Academy of Medicine, New York, NY [4]Depts of Environmental & Occupational Health Sciences, Medicine & Epidemiology, Univ of Washington, Seattle, WA [5]Depts of Medicine & Epidemiology, Columbia Univ Medical Center, New York, NY [6]Division of Epidemiology & Community Health, School of Public Health, University of Minnesota, Minneapolis, MN [7]Center for Public Health Genomics, Univ of Virginia, Charlottesville, VA [8]Human Genetics Ctr, School of Public Health, Univ of Texas Health Science Center at Houston, Houston, TX [9]Univ of Wisconsin School of Medicine & Public Health, Madison, WI [10]Molecular Biology, Nijmegen Centre for Molecular Life Sciences (NCMLS), Nijmegen, Netherlands [11]Depts of Pathology & Cardiology, Johns Hopkins Univ, Baltimore, MD [12]Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA

Abstract

Background—Tobacco smoke contains numerous agonists of the aryl-hydrocarbon receptor (AhR) pathway, and activation of the AhR pathway was shown to promote atherosclerosis in mice. Intriguingly, cigarette smoking is most strongly and robustly associated with DNA modifications to an AhR pathway gene, the aryl-hydrocarbon receptor repressor (*AHRR).* We hypothesized that altered *AHRR* methylation in monocytes, a cell type sensitive to cigarette

Correspondence: Yongmei Liu, MD, PhD, Department of Epidemiology and Prevention, Division of Public Health Sciences, Medical Center Blvd., Wake Forest School of Medicine, Winston-Salem, NC 27157, Tel: 336-716-9503, Fax: 336-716-6427, yoliu@wakehealth.edu.

* contributed equally

Conflict of Interest Disclosures: None.

Journal Subject Terms: Epidemiology; Functional Genomics; Gene Expression and Regulation

smoking and involved in atherogenesis, may be a part of the biological link between cigarette smoking and atherosclerosis.

Methods and Results—DNA methylation profiles of *AHRR* in monocytes (542 CpG sites ± 150kb of *AHRR*, using Illumina 450K array) were integrated with smoking habits and ultrasound-measured carotid plaque scores from 1,256 participants of the Multi-Ethnic Study of Atherosclerosis (MESA). Methylation of cg05575921 significantly associated ($p = 6.1\times10^{-134}$) with smoking status (current vs. never). Novel associations between cg05575921 methylation and carotid plaque scores ($p = 3.1\times10^{-10}$) were identified, which remained significant in current and former smokers even after adjusting for self-reported smoking habits, urinary cotinine, and well-known CVD risk factors. This association replicated in an independent cohort using hepatic DNA (n = 141). Functionally, cg05575921 was located in a predicted gene expression regulatory element (enhancer), and had methylation correlated with *AHRR* mRNA profiles ($p = 1.4\times10^{-17}$) obtained from RNA sequencing conducted on a subset (n = 373) of the samples.

Conclusions—These findings suggest *AHRR* methylation may be functionally related to *AHRR* expression in monocytes, and represents a potential biomarker of subclinical atherosclerosis in smokers.

Keywords

smoking; atherosclerosis; gene expression/regulation; epidemiology; epigenetics; DNA methylation

Background

Smoking is a major preventable risk factor for many human diseases including cardiovascular diseases[1, 2] (CVD), which cause a substantial health and economic burden worldwide. Decoding the molecular mechanisms by which smoking drives the development of CVD may provide both preclinical diagnostic biomarkers and drug targets for therapeutics of CVD.

Epigenetic modifications such as DNA methylation of cytosine-(phosphate)-guanine (CpG) dinucleotides are functional biochemical alterations to DNA that are implicated in regulation of gene expression and disease development[2, 3]. Cigarette smoke contains thousands of chemicals[4], including many known carcinogens such as polycyclic aromatic hydrocarbons (PAHs)[5], which may lead to both mutagenic and epigenetic alterations to the genomes of smokers. Recent genome-wide association studies have shown that smoking is most significantly and robustly associated with DNA methylation alterations of the Aryl Hydrocarbon Receptor Repressor (*AHRR*)[6–9], which appear to be at least partially reversible upon smoking cessation[9], and associated with higher mRNA expression levels of *AHRR* in human lung[10], and lymphoblasts[11].

Previous studies indicate that AHRR, a basic helix–loop–helix/Per-ARNT-Sim (bHLH/PAS) family transcription factor, is transcriptionally regulated through activation of the Ah Receptor (AhR) pathway[12, 13]. Tobacco smoke contains numerous agonists of AhR signaling[14] and persistent activation of the AhR signaling pathway has been hypothesized to contribute to atherogenesis[15–18]. While there is currently no known role for AHRR in

atherogenesis, recent studies implicate up-regulation of *AHRR* expression in monocyte-to-macrophage differentiation[19] and suppression of anti-inflammation[13]. We hypothesized that altered *AHRR* methylation levels in CD14+ monocytes, a cell type sensitive to cigarette smoking[20] and involved in atherogenesis[21, 22], might provide a biological link between smoking and atherosclerotic plaque formation. Therefore, we integrated methylation profiles of CpG sites within 150kb of *AHRR* (measured using the Illumina 450K array) in primary CD14+ monocytes with smoking status and ultrasound-measured carotid plaque scores from 1,256 participants of the Multi-Ethnic Study of Atherosclerosis (MESA)[23], and sought replication in the Pathobiological Determinants of Atherosclerosis in Youth Study (PDAY, n = 141)[24]. To look for potential functional effects of smoking-associated *AHRR* methylation, we incorporated histone modification data in monocytes from the ENCODE[25] and BLUEPRINT[26] projects, and gene expression measured in monocytes from a subset of the population (n = 373) using RNA-sequencing. Methylation and gene expression results were validated in independent monocyte samples using reduced representation bisulfite sequencing (RRBS, n = 4) and real-time PCR (n = 10).

Methods

Study Population

The Multi-Ethnic Study of Atherosclerosis (MESA) was designed to investigate the prevalence, correlates, and progression of subclinical cardiovascular disease in a population cohort of 6,814 participants. Since its inception in 2000, five clinic visits collected extensive clinical, socio-demographic, lifestyle, behavior, laboratory, nutrition, and medication data[23]. The present analysis is primarily based on analyses of purified monocyte samples from the April 2010 – February 2012 examination (exam 5) of 1,264 randomly selected MESA participants from four MESA field centers (Baltimore, MD; Forsyth County, NC; New York, NY; and St. Paul, MN). The study protocol was approved by the Institutional Review Board at each site. All participants signed informed consent.

Data collection

Smoking status, average number of cigarettes smoked per day, age of smoking cessation, and pack-years of cigarette smoking were assessed via standard American Thoracic Society questionnaire items[27]. Participants who reported smoking fewer than 100 cigarettes in their lifetimes were classified as never smokers. Among participants who reported smoking greater than 100 cigarettes in their lifetime, those who reported smoking during the last 30 days were classified as current smokers and those who did not were classified as former smokers. Pack-years of cigarette smoking were calculated from age of starting to quitting (or current age among current smokers) × (cigarettes per day/20). Urinary cotinine levels were measured via immunoassay (Immulite 2000 Nicotine Metabolite Assay; Diagnostic Products Corp., Los Angeles, CA)[28].

To obtain carotid artery plaque scores, readers adjudicated carotid plaque presence or absence, defined as a focal abnormal wall thickness (carotid IMT >1.5 mm) or a focal thickening of >50% of the surrounding IMT, as reported previously[29, 30]. Presence or absence of plaque acoustic shadowing was recorded. A total plaque score (range 0–12) was

calculated to describe carotid plaque burden. One point per plaque was allocated for the near and far walls of each segment (CCA, bulb, ICA) of each carotid artery that was interrogated. For carotid plaque presence, intra-reader reproducibility was excellent (per reader Kappa = 0.82–1.0, overall Kappa = 0.83, 95% confidence interval [CI] 0.70–0.96), as was inter-reader reproducibility (Kappa = 0.89; 95% CI 0.72 – 1.00)[31].

Weight was measured with a Detecto Platform Balance Scale to the nearest 0.5 kg. Height was measured with a stadiometer (Accu-Hite Measure Device with level bubble) to the nearest 0.1 cm. BMI was defined as weight in kilograms divided by square of height in meters. Type 2 Diabetes (T2D) was defined as fasting glucose ≥ 6.99 mmol/L (≥ 126 mg/dL) or use of hypoglycemic medication. Fasting serum glucose was measured by rate reflectance spectrophotometry using thin-film adaptation of the glucose oxidase method on the Vitros analyzer (Johnson & Johnson Clinical Diagnostics, Rochester, NY). To achieve consistency of the serum glucose assay over examinations, 200 samples were reanalyzed and then recalibrated the original observations. LDL-cholesterol was calculated using the Friedewald equation. Resting blood pressure was measured three times in the seated position using a Dinamap model Pro 100 automated oscillometric sphygmomanometer (Critikon, Tampa, FL). The average of the last two measurements was used in analysis. Hypertension was assessed using the JNC VI (1997) criteria (systolic pressure greater than or equal to 140 mm Hg, diastolic pressure greater than or equal to 90 mm Hg or current use of anti-hypertensive medication).

Purification of Monocytes

Monocytes were isolated with anti-CD14 monoclonal antibody coated magnetic beads, using AutoMACs automated magnetic separation unit (Miltenyi Biotec, Bergisch Gladbach, Germany). Initially flow cytometry analysis of 18 monocyte samples collected from all four MESA field centers was performed to assess the cell purification quality across the labs and technicians. The purity was > 90% for all samples (see Data Supplement for details).

DNA/RNA Extraction

DNA and RNA were isolated from samples simultaneously using the AllPrep DNA/RNA Mini Kit (Qiagen, Inc., Hilden, Germany; see Data Supplement for details).

Methylation Quantification

Illumina HumanMethylation450 BeadChips and HiScan reader were used to perform the epigenome-wide methylation analysis (see Data Supplement for details). This methylation data has been deposited in the NCBI Gene Expression Omnibus and is accessible through GEO Series accession number (GSE56046).

Quality Control and Pre-Processing of Microarray Data

Data pre-processing and quality control (QC) analyses were performed in *R* (http://www.r-project.org/) using *Bioconductor* (http://www.bioconductor.org/) packages (see Data Supplement for detailed methods). The Illumina HumanMethylation450 BeadChip included probes for 485K CpG sites. Of these CpG sites, 484,817 passed the QC elimination criteria

including: 'detected' methylation levels <90% of MESA samples using a detection p-value cut-off of 0.05 or overlap with a repetitive element or region.

To estimate residual sample contamination for data analysis, we generated separate enrichment scores for neutrophils, B cells, T cells, monocytes, and natural killer cells. We implemented a Gene Set Enrichment Analysis[32] as previously described[33] to calculate the enrichment scores using the gene signature of each blood cell type from previously defined lists[34].

To remove technical error in methylation levels associated with batch effects across the multiple chips, positional effects of the sample on the chip, and residual sample contamination with non-monocyte cell types, we adjusted methylation M-values for chip, sample position on the chip, and estimated residual sample contamination with neutrophils, B cells, T cells, monocytes, and natural killer cells.

mRNA quantification using RNA seq

A subset of 374 samples was randomly selected from the 1,264 MESA monocyte samples for RNA sequencing of mRNA (see Data Supplement for details).

Association Analyses

The overall goal of the association analysis was to characterize associations between smoking and *AHRR* methylation, and *AHRR* methylation and carotid plaque scores. Analyses were performed using the linear model (*lm*) function of the *Stats* package and the *stepAIC* function of the *MASS* package in *R*. Reported correlations (r) represent the partial Pearson product-moment correlation coefficient, after adjusting for covariates.

To identify methylation sites associated with smoking, we fit separate linear regression models with smoking status as a predictor of the M-value for each CpG site genome-wide that passed QC. Covariates were age, sex, race/ethnicity, and study site. P-values were adjusted for multiple testing using the q-value FDR method[35]. The FDR was calculated at the genome-wide level, including all 484,817 CpGs passing QC, although only results from CpG sites within 150 kb of *AHRR* (542 CpGs) are presented.

To identify potential bimodal distributions, we calculated the kurtosis and skewness of the M-value distributions of the 542 probes investigated. Six probes (1.1%) were found to have strong negative kurtosis (kurtosis < −.75) and some skewness (skewness < −.3 or skewness > .3), suggesting bimodal distributions. Although linear regression estimates are robust to the normality assumption in large samples, we next performed permutation testing with the 542 CpGs (H_0: beta = 0, i.e. no difference in current smoking vs. never smoking) to ensure the validity of the linear regression results. The number of permutations was set to be 1,000 for all CpGs, and for CpGs with p-value < .10, we also ran 100,000 permutations. We found very consistent permutation p-values compared to the normal p-values from the linear regression analysis for all probes, including the six probes with bimodal distributions. The correlation between the normal and permutation p-values is 0.999, the correlation between the ranks based on the normal and permutation p-values is also 0.999, and the sum of the

absolute differences between the normal and permutation p-values divided by the sum of the normal p-values is 2% (0.02).

To characterize the association between *AHRR* methylation and carotid plaque levels, we fit separate linear regression models with the M-value for the 542 CpGs within 150 kb of *AHRR* as a predictor of carotid plaque scores [natural log (carotid plaque score + 1)]. Initial analysis covariates were age, sex, race/ethnicity, and study site. Secondary analyses adjusted for smoking status, the number of cigarettes per day, pack-years, traditional CVD risk factors (BMI, LDL cholesterol, hypertension, diabetes mellitus) and statin use.

Methylation of the most significantly differentially methylated CpG associated with smoking status and with carotid plaques scores (i.e. cg05575921) was also investigated for association with nearby gene expression levels and genetic variants. We arbitrarily chose to investigate a large window (± 1Mb) surrounding the CpG to avoid missing any potential effects. For gene expression analysis, we fit separate linear regression models with the M-value as a predictor of mRNA expression for any autosomal gene within 1 Mb of the CpG. Covariates were age, sex, and race/ethnicity, and study site. For analyses of nearby genetic variants, we fit separate linear regression models with single nucleotide polymorphisms (SNPs) located within 1 Mb as a predictor of the M-value in the MESA samples from Caucasian participants, including SNPs with a minor allele frequency > 0.05 in the MESA Caucasian population. P-values were adjusted for multiple testing using the q-value FDR method[35].

Mediation analysis

We performed mediation analysis to investigate the hypothesis that smoking may have an effect on atherosclerosis mediated through methylation alteration. We used Structural Equation Modeling (SEM) with bootstrapping as implemented in the R package lavaan[36] to estimate the indirect effects (mediated through methylation) of current smoking (cigarettes per day) and cumulative exposure (pack-years) on carotid plaque score in current and former smokers for the most significantly differentially methylated CpG associated with smoking status and with carotid plaques score.

Replication study

Replication analysis for the most significantly differentially methylated CpG associated with smoking status and with carotid plaques score included samples from 141 males from the Pathobiological Determinants of Atherosclerosis in Youth Study (PDAY)[24] (see Data Supplement for details). Tests for association between methylation measured in hepatic samples and the extent of fatty streaks observed in the right coronary arteries and the left halves of the aortas were performed using linear regression, including the following risk factors: age, race, lipids (HDL, non-HDL), BMI, and glucose levels.

Reduced Representation Bisulfite Sequencing (RRBS) methods and DMR Identification

To investigate the relationship between smoking and CpG methylation not captured by microarray, RRBS libraries of CD14+ monocytes from two smokers and two non-smokers were generated as previously described[37] with minor modifications (see Data Supplement).

To identify genomic regions that are differentially methylated regions (DMRs) between smokers and non-smokers, we used the method described in Ziller *et al.*[38]. Briefly, this method first uses the methylated and unmethylated counts to identify a set of dynamic CpGs, which are CpGs that are significantly differentially methylated between groups, using a model based on the beta difference distribution. Next, it merges any dynamic CpGs whose genomic locations are within 500 bp into a set of CpG clusters. Finally, all clusters that are significantly differentially methylated between groups are retained as the final set of DMRs. We defined the DMR location as the region between the outermost CpGs in the cluster. We excluded any DMRs that had fewer than five CpGs with nonzero methylation counts, and we also excluded any DMRs that had an overall methylation difference of less than 10%.

Functional characterization

Histone modifications in monocytes and monocyte-derived macrophages from BLUEPRINT[19] were downloaded from GEO (Series GSE58310) including H3K27ac and H3K4me1. Other histone modifications in monocytes from ENCODE[39] were downloaded through the UCSC browser[25], including H3K4me1 (GEO accession GSM1003535), H3K4me3 (GEO accession GSM1003536), and H3K27me3 (GEO accession GSM1003564), which were produced by the Bernstein lab at the BROAD Institute.

RT-qPCR

Reverse Transcription Quantitative Polymerase Chain Reaction (RTqPCR) was used for replication of gene expression in CD14+ monocytes from nonsmokers and smokers collected at the NIEHS Clinical Research Unit (see Data Supplement).

Results

The study population (n = 1,256; Table 1) was composed of 47% Caucasians, 32% Hispanics, and 21% African Americans that were recruited from four different study sites across the US, as previously described[33].

AHRR Methylation and Smoking

We identified 33 of 542 CpGs within 150 kb of *AHRR* whose degree of methylation was significantly associated with cigarette smoking at a genome-wide FDR threshold of 5% (p ranging 6.1×10^{-134} – 2.5×10^{-4}, Figure 1) in monocytes from 114 current smokers and 502 never smokers. Methylation of cg05575921 (chr5:373,378; hg19) within the *AHRR* gene body represented the most significantly ($p = 6.07\times10^{-134}$, FDR = 2.67×10^{-128}) differentially methylated CpG.

To explore the smoking-associated cg05575921 methylation in more detail, we examined the relationship between smoking history/status, the dosage of exposure, the level of urinary cotinine, and cg05575921 methylation. In agreement with the proposed reversibility of smoking-responsive cg05575921 methylation[9], the average methylation measured in former smokers (mean beta = 0.79) was more similar to never smokers (mean beta = 0.83, Figure 2a) than current smokers (mean beta = 0.59). Furthermore, we found that cg05575921 methylation was positively correlated with time since quitting in former smokers (partial

correlation: r = 0.35, $p < 2\times10^{-16}$, Figure 2b), and was negatively correlated with recent exposure measurements in current smokers such as the average number of cigarettes smoked per day (r = −0.39, $p = 6.1\times10^{-5}$, Figure 2c) and urinary cotinine levels (r = −0.43, $p < 9.09\times10^{-4}$, Figure 2d), after adjusting for age, sex, race, and study site. Consistent with these findings, cg05575921 methylation was also significantly negatively correlated with cumulative exposure measured by pack-years in current smokers (r = −0.34, $p = 3.97\times10^{-4}$) and former smokers (r = −0.35, $p < 2\times10^{-16}$; Data Supplement Figure 1).

AHRR Methylation and Carotid Plaque Scores

Among all investigated CpGs within 150 kb of *AHRR*, methylation profiles of two CpGs were associated with higher carotid plaque scores (cg05575921, $p = 3.08\times10^{-10}$; cg21161138, $p = 7.73\times10^{-8}$; Figure 3a), both of which were also associated with smoking. Including both CpG methylation profiles in the same model did not find independent effects for these CpGs on carotid plaque score (methylation correlation = 0.37).

Importantly, even after adjusting for statin use and traditional CVD risk factors including body mass index (BMI), low-density lipoprotein (LDL) cholesterol, hypertension, and diabetes mellitus, the CpG with methylation most strongly associated with carotid plaque score remained significant (cg05575921, r = −0.18, $p = 1.14\times10^{-9}$, Figure 3b). After additionally adjusting for smoking status, the effect of cg05575921 methylation on carotid plaque score remained significant (r = −0.13, $p = 2.2\times10^{-5}$). We then stratified by smoking status and found the estimated effect of cg05575921 methylation was strongest among current smokers (r = −0.32, $p = 2.93\times10^{-3}$) and was attenuated in former smokers (r = −0.13, $p = 1.64\times10^{-3}$). No significant relationship was detected between *AHRR* methylation and carotid plaque score in never smokers (r = −0.03; p = 0.46), likely due to the predominantly hypermethylated state of cg05575921 in never smokers, and the associated reduced variability of methylation in never smokers compared with current and former smokers.

In current and former smokers, the effect of cg05575921 methylation on carotid plaque score was further attenuated, but remained significant ($p < 0.05$) after including self-reported pack-years and urinary cotinine levels in the model (Data Supplement Table 1). Standardized effect estimates of cg05575921 methylation (beta = −0.13, p = 0.01) and pack-years (beta = 0.13, p = 0.007) on carotid plaque score were similar in magnitude, each uniquely explaining ~1% of the variance of carotid plaque score. The standardized effect size of cg05575921 methylation on carotid plaque score was similar in magnitude to the effect of LDL cholesterol on carotid plaque score (beta = 0.16, p = 0.001), as shown in the full model results presented in the Data Supplement Table 1.

Using mediation analysis to explore if methylation was intermediate in the association between smoking (pack-years and cigarettes per day) and higher carotid plaque score, we found cg05575921 methylation significantly mediated 37% of the total effect of self-reported pack-years on carotid plaque score (indirect effect: β ± SE: 0.25 ± 0.06; $p = 3.1\times10^{-4}$) in current and former smokers, and 60% of the total effect of cigarettes per day on carotid plaque score (indirect effect: β ± SE: 0.011 ± 0.004; $p = 4.3\times10^{-3}$) in current smokers.

We replicated both the associations of cg05575921 methylation (in hepatic cells) with smoking ($p = 0.002$), and with subclinical atherosclerosis (extent of fatty streaks in the right coronary arteries and left halves of the aortas; $p = 0.002$) measured at autopsy in biopsies from 141 young males (< 35 years) from the PDAY study.

Functional genomic characterization of the cg05575921 locus

Using reduced representation bisulfite sequencing (RRBS) to measure CpG methylation in non-MESA monocyte samples from two smokers and two non-smokers, we recapitulated the association between smoking and methylation of cg05575921. Additionally, RRBS revealed seven adjacent (within 180 bp) CpGs not included on the Illumina 450K microarray which displayed similar directional changes in methylation with smoking status, composing a smoking-associated differentially methylated region (DMR) (Figure 4 and Data Supplement Figure 2).

The *AHRR* DMR (chr5:373,378–373,556 hg19) is located in a potential poised enhancer region, based on the presence of both activating (H3K27ac/H3K4me1) and repressive (H3K27me3) histone modifications in a monocyte sample (from ENCODE[25]). This enhancer region appears to be strongly activated (based on the enrichment of histone marks H3K27ac/H3K4me1) upon monocyte differentiation to macrophage (from BLUEPRINT[19, 26]; Figure 4).

DNA methylation of cg05575921 and *cis*-gene expression

In a subset of the MESA monocyte samples with both RNA-sequencing and methylation profiles (n=373, including 29 current smokers and 152 never smokers), cg05575921 methylation significantly (FDR < 0.05) associated with mRNA expression of only one gene within 1 MB, *AHRR* ($r = -0.42$, $p = 1.4\times10^{-17}$, FDR = $4.32\times10^{-15;}$ Figure 5a). Analysis stratified by smoking status revealed that *AHRR* expression was strongly up-regulated in current smokers ($p = 3.29\times10^{-21}$, Data Supplement Figure 3a), which was replicated using real-time PCR in a set of independent monocyte samples ($p = 0.003$; Data Supplement Figure 3b).

Next, we tested for associations between *AHRR* expression and carotid plaque score and found suggestive evidence, even in our small sample size of current smokers with RNA-sequencing measurements, for a positive association between *AHRR* expression and carotid plaque score ($r = 0.53$, $p = 0.07$, $n = 21$; Figure 5b). Carotid plaque score was also positively associated with *AHRR* expression in former smokers ($r = 0.17$, $p = 0.03$, $n = 175$), but not in never smokers ($p > 0.05$; $n = 147$).

Nearby genetic variants were also investigated for association with cg05575921 methylation; however, no significant (FDR < 0.05) associations were observed with common genetic variants.

Discussion

Extending previous findings in other tissues, we found decreased methylation levels in human monocytes at the *AHRR* CpG cg05575921 is a significant biomarker of smoking.

This finding is in line with several previous reports in other cell types[6, 7, 11]. Cg05575921 methylation appeared to be dosage-dependent and at least partially reversible upon smoking cessation. These results, along with other published reports[7], indicate that methylation levels at cg05575921 represent a biomarker of "effect" that integrates both smoking dose and duration.

Importantly, we uncovered novel associations between smoking-responsive methylation of *AHRR* in monocytes and subclinical atherosclerosis in a large, multi-ethnic human cohort, and replicated the association between subclinical atherosclerosis and *AHRR* methylation in hepatic cells from an independent cohort. These results suggest that smoking-altered methylation of *AHRR* in multiple cell types may represent a biomarker for the risk of atherosclerosis associated with smoking.

Our results also indicate that *AHRR* methylation could be a useful biomarker of atherosclerosis in addition to self-reported smoking exposure and other known CVD risk factors. For instance, in current and former smokers, *AHRR* methylation remained significantly associated with subclinical atherosclerosis even after adjustment for self-reported smoking exposure, urinary cotinine, and many other known CVD risk factors. However, longitudinal studies are necessary to confirm *AHRR* methylation as a predictor of atherosclerosis.

Genome-wide association studies have suggested that differential DNA methylation is often associated with expression changes of nearby genes[33]. Functionally, we found reduced methylation levels at cg05575921, measured by microarray, were associated with up-regulated *AHRR* mRNA expression in monocytes from 373 participants, measured using RNA-sequencing, and in an independent sample using RT-PCR. Combined with a recent report in lung tissue from five smokers and five non-smokers that found smoking-associated methylation of *AHRR* was linked with increased *AHRR* expression ($p=0.0047$)[10], these results provide evidence that altered *AHRR* methylation profiles may mediate increased transcription of *AHRR* in smokers.

Previous studies indicate that *AHRR* is transcriptionally up-regulated by activation of the Ah Receptor pathway[12, 13]. Interestingly, RRBS showed that the *AHRR* cg05575921 locus is located in a smoking-associated DMR, which appears to overlap an intragenic enhancer that is activated, accompanied by up-regulation of *AHRR* expression, when monocytes differentiate into macrophages[19]. Taken together with the ability of DNA methylation profiles to predict enhancer activity[40, 41], we hypothesize that smoking-induced Ah Receptor pathway activation drives methylation changes in the *AHRR* gene accompanied by histone modifications and enhancer activation, resulting in up-regulation of *AHRR* expression in monocytes. Residual levels of altered methylation and enhancer activity at the cg05575921 DMR in former smokers could partially explain why *AHRR* gene expression was previously reported to be up-regulated in former smokers compared with never smokers even decades after smoking cessation[42].

The consequences of *AHRR* up-regulation in monocytes is unclear, but given the potential role of AHRR in macrophage biology[19] and suppression of anti-inflammation[13], up-

regulation of *AHRR* could plausibly alter the function of monocytes and macrophages in smokers. While there is currently no known role for *AHRR* in atherogenesis, a number of studies have shown that activation of the Ah Receptor pathway promotes atherosclerosis. For instance, using *in vitro* and *in vivo* models in mice, Wu *et al*[18] demonstrated that Ah Receptor pathway activation with treatments such as cigarette smoke extract or 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD), increased expression of inflammatory markers in macrophages, and were associated with the accumulation of lipids in macrophages and plaque formation. Vogel *et al*.[43] showed that Ah Receptor pathway activation increased cytokine production and foam cell formation. A limitation of our study is it remains unclear if smoking-associated *AHRR* expression is causally related to atherogenesis, or occurs in parallel to the Ah Receptor activation-related effects on plaque formation noted above. Mediation analysis results do not distinguish confounding from mediating effects; therefore, although we observe cg05575921 methylation significantly mediating the effect of pack-years on carotid plaque score, these results do not confirm biological mediation and should be interpreted with caution.

In conclusion, we identified smoking-responsive methylation profiles at the cg05575921 locus which are associated with up-regulated *AHRR* expression in monocytes, and with subclinical atherosclerosis in a human population. Future functional studies are necessary to understand the potential role of AHRR in response to cigarette smoke and in atherogenesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org. The authors would also like to thank Drs. Michael Ziller and Alexander Meissner of the Broad Institute of MIT and Harvard for their generosity in providing the R code to identify differentially methylated regions.

Funding Sources: This research was supported by contracts N01-HC-from the National Heart, Lung and Blood Institute, by grants UL1-RR-024156 and UL1-RR-025005 from the National Center for Research Resources, and support from the Intramural Research Program (projects: ZO1-ES100475 and Z01 ES046008) of the National Institute of Environmental Health Sciences. The MESA Epigenomics & Transcriptomics Study was funded by NHLBI grant R01HL101250 to Wake Forest University Health Sciences. Urinary Cotinine measurements are available to the MESA study courtesy of MESA Lung contract R01-HL077612. The research described in this publication was funded in part by the U.S Environmental Protection Agency through RD831697 to the University of Washington (MESA Air); it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

References

1. Csiszar A, Podlutsky A, Wolin MS, Losonczy G, Pacher P, Ungvari Z. Oxidative stress and accelerated vascular aging: implications for cigarette smoking. Front Biosci (Landmark Ed). 2009; 14:3128–3144. [PubMed: 19273262]
2. Breitling LP. Current genetics and epigenetics of smoking/tobacco-related cardiovascular disease. Arterioscler Thromb Vasc Biol. 2013; 33:1468–1472. [PubMed: 23640490]
3. Petronis A. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. Nature. 2010; 465:721–727. [PubMed: 20535201]

4. Talhout R, Schulz T, Florek E, van BJ, Wester P, Opperhuizen A. Hazardous compounds in tobacco smoke. Int J Environ Res Public Health. 2011; 8:613–628. [PubMed: 21556207]
5. Muto H, Takizawa Y. Dioxins in cigarette smoke. Arch Environ Health. 1989; 44:171–174. [PubMed: 2751353]
6. Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. Environ Health Perspect. 2012; 120:1425–1431. [PubMed: 22851337]
7. Philibert RA, Beach SR, Lei MK, Brody GH. Changes in DNA methylation at the aryl hydrocarbon receptor repressor may be a new biomarker for smoking. Clin Epigenetics. 2013; 5:19. [PubMed: 24120260]
8. Dogan MV, Shields B, Cutrona C, Gao L, Gibbons FX, Simons R, et al. The effect of smoking on DNA methylation of peripheral blood mononuclear cells from African American women. BMC Genomics. 2014; 15:151. [PubMed: 24559495]
9. Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. PLoS One. 2013; 8:e63812. [PubMed: 23691101]
10. Shenker NS, Polidoro S, van Veldhoven K, Sacerdote C, Ricceri F, Birrell MA, et al. Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. Hum Mol Genet. 2013; 22:843–851. [PubMed: 23175441]
11. Monick MM, Beach SR, Plume J, Sears R, Gerrard M, Brody GH, et al. Coordinated changes in AHRR methylation in lymphoblasts and pulmonary macrophages from smokers. Am J Med Genet B Neuropsychiatr Genet. 2012; 159B:141–151. [PubMed: 22232023]
12. Kazantseva MG, Highton J, Stamp LK, Hessian PA. Dendritic cells provide a potential link between smoking and inflammation in rheumatoid arthritis. Arthritis Res Ther. 2012; 14:R208. [PubMed: 23036591]
13. Awji EG, Chand H, Bruse S, Smith KR, Colby JK, Mebratu Y, et al. Wood Smoke Enhances Cigarette Smoke-Induced Inflammation by Inducing the Aryl Hydrocarbon Receptor Repressor in Airway Epithelial Cells. Am J Respir Cell Mol Biol. 2015; 52:377–386. [PubMed: 25137396]
14. Dertinger SD, Silverstone AE, Gasiewicz TA. Influence of aromatic hydrocarbon receptor-mediated events on the genotoxicity of cigarette smoke condensate. Carcinogenesis. 1998; 19:2037–2042. [PubMed: 9855021]
15. Kerley-Hamilton JS, Trask HW, Ridley CJ, Dufour E, Lesseur C, Ringelberg CS, et al. Inherent and benzo[a]pyrene-induced differential aryl hydrocarbon receptor signaling greatly affects life span, atherosclerosis, cardiac gene expression, and body and heart growth in mice. Toxicol Sci. 2012; 126:391–404. [PubMed: 22228805]
16. Humblet O, Birnbaum L, Rimm E, Mittleman MA, Hauser R. Dioxins and cardiovascular disease mortality. Environ Health Perspect. 2008; 116:1443–1448. [PubMed: 19057694]
17. Savouret JF, Berdeaux A, Casper RF. The aryl hydrocarbon receptor and its xenobiotic ligands: a fundamental trigger for cardiovascular diseases. Nutr Metab Cardiovasc Dis. 2003; 13:104–113. [PubMed: 12929624]
18. Wu D, Nishimura N, Kuo V, Fiehn O, Shahbaz S, Van WL, et al. Activation of aryl hydrocarbon receptor induces vascular inflammation and promotes atherosclerosis in apolipoprotein E−/− mice. Arterioscler Thromb Vasc Biol. 2011; 31:1260–1267. [PubMed: 21441140]
19. Saeed S, Quintin J, Kerstens HH, Rao NA, Aghajanirefah A, Matarese F, et al. Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science. 2014; 345:1251086. [PubMed: 25258085]
20. Bergmann S, Siekmeier R, Mix C, Jaross W. Even moderate cigarette smoking influences the pattern of circulating monocytes and the concentration of sICAM-1. Respir Physiol. 1998; 114:269–275. [PubMed: 9926990]
21. Ley K, Miller YI, Hedrick CC. Monocyte and macrophage dynamics during atherogenesis. Arterioscler Thromb Vasc Biol. 2011; 31:1506–1516. [PubMed: 21677293]

22. Lessner SM, Prado HL, Waller EK, Galis ZS. Atherosclerotic lesions grow through recruitment and proliferation of circulating monocytes in a murine model. Am J Pathol. 2002; 160:2145–2155. [PubMed: 12057918]
23. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, et al. Multi-ethnic study of atherosclerosis: objectives and design. Am J Epidemiol. 2002; 156:871–881. [PubMed: 12397006]
24. Wissler RW. USA Multicenter Study of the pathobiology of atherosclerosis in youth. Ann N Y Acad Sci. 1991; 623:26–39. [PubMed: 2042833]
25. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, et al. ENCODE data in the UCSC Genome Browser: year 5 update. Nucleic Acids Res. 2013; 41:D56–D63. [PubMed: 23193274]
26. Adams D, Altucci L, Antonarakis SE, Ballesteros J, Beck S, Bird A, et al. BLUEPRINT to decode the epigenetic signature written in blood. Nat Biotechnol. 2012; 30:224–226. [PubMed: 22398613]
27. Ferris BG. Epidemiology Standardization Project (American Thoracic Society). Am Rev Respir Dis. 1978; 118:1–120. [PubMed: 742764]
28. Rodriguez J, Jiang R, Johnson WC, MacKenzie BA, Smith LJ, Barr RG. The association of pipe and cigar use with cotinine levels, lung function, and airflow obstruction: a cross-sectional study. Ann Intern Med. 2010; 152:201–210. [PubMed: 20157134]
29. Stein JH, Korcarz CE, Hurst RT, Lonn E, Kendall CB, Mohler ER, et al. Use of carotid ultrasound to identify subclinical vascular disease and evaluate cardiovascular disease risk: a consensus statement from the American Society of Echocardiography Carotid Intima-Media Thickness Task Force. Endorsed by the Society for Vascular Medicine. J Am Soc Echocardiogr. 2008; 21:93–111. [PubMed: 18261694]
30. Touboul PJ, Hennerici MG, Meairs S, Adams H, Amarenco P, Bornstein N, et al. Mannheim carotid intima-media thickness consensus (2004–2006). An update on behalf of the Advisory Board of the 3rd and 4th Watching the Risk Symposium, 13th and 15th European Stroke Conferences, Mannheim, Germany, 2004, and Brussels, Belgium, 2006. Cerebrovasc Dis. 2007; 23:75–80. [PubMed: 17108679]
31. Tattersall MC, Gassett A, Korcarz CE, Gepner AD, Kaufman JD, Liu KJ, et al. Predictors of carotid thickness and plaque progression during a decade: the multi-ethnic study of atherosclerosis. Stroke. 2014; 45:3257–3262. [PubMed: 25213342]
32. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–15550. [PubMed: 16199517]
33. Liu Y, Ding J, Reynolds LM, Lohman K, Register TC, de la Fuente A, et al. Methylomics of gene expression in human monocytes. Hum Mol Genet. 2013; 22:5065–5074. [PubMed: 23900078]
34. Abbas AR, Baldwin D, Ma Y, Ouyang W, Gurney A, Martin F, et al. Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. Genes Immun. 2005; 6:319–331. [PubMed: 15789058]
35. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A. 2003; 100:9440–9445. [PubMed: 12883005]
36. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. J Stat Softw. 2012; 48:1–36.
37. Boyle P, Clement K, Gu H, Smith ZD, Ziller M, Fostel JL, et al. Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome Biol. 2012; 13:R92. [PubMed: 23034176]
38. Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, Kohlbacher O, et al. Charting a dynamic DNA methylation landscape of the human genome. Nature. 2013; 500:477–481. [PubMed: 23925113]
39. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]
40. Schmidl C, Klug M, Boeld TJ, Andreesen R, Hoffmann P, Edinger M, et al. Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. Genome Res. 2009; 19:1165–1174. [PubMed: 19494038]

41. Wiench M, John S, Baek S, Johnson TA, Sung MH, Escobar T, et al. DNA methylation status predicts cell type-specific enhancer activity. EMBO J. 2011; 30:3028–3039. [PubMed: 21701563]

42. Bosse Y, Postma DS, Sin DD, Lamontagne M, Couture C, Gaudreault N, et al. Molecular signature of smoking in human lung tissues. Cancer Res. 2012; 72:3753–3763. [PubMed: 22659451]

43. Vogel CF, Sciullo E, Matsumura F. Activation of inflammatory mediators and potential role of ah-receptor ligands in foam cell formation. Cardiovasc Toxicol. 2004; 4:363–373. [PubMed: 15531779]

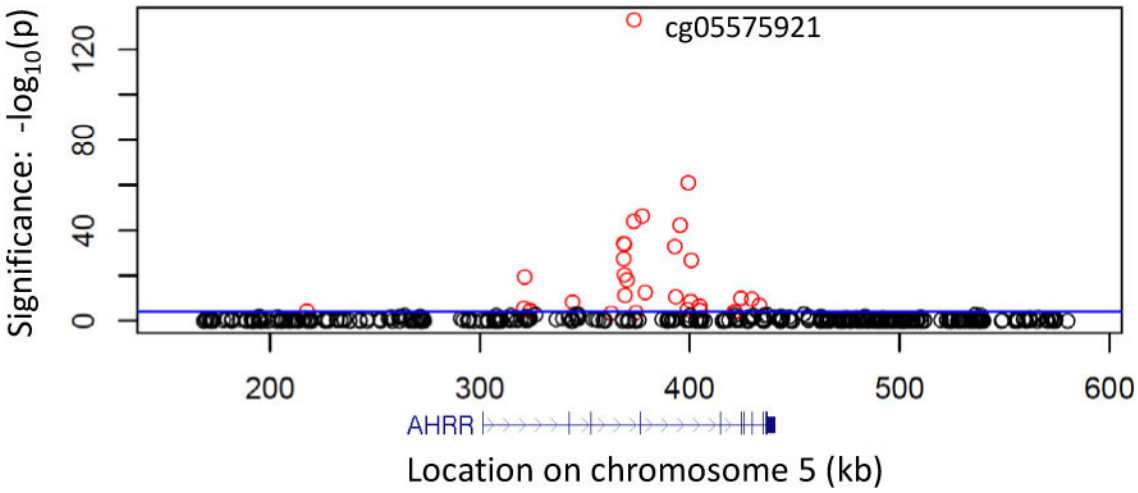

Figure 1.

Cigarette smoking status is strongly associated with DNA methylation profiles of *AHRR* in monocytes from current smokers and never smokers. Regional association plot shows the significance (y-axis: $-\log_{10}$ p-value; $FDR_{genome\text{-}wide} < 0.05$ colored red; $FDR_{genome\text{-}wide} \geq 0.05$ colored black) of the association between smoking status (114 current vs. 502 never smokers) and methylation in CD14+ monocytes (542 CpGs within 150 kb of *AHRR*, measured using the Illumina 450K microarray). The blue horizontal line represents a Bonferroni corrected significance threshold ($\alpha = 0.05$) for the 542 CpGs examined. The genomic position (hg19) of *AHRR* relative to the CpGs is shown below the association plot (in blue). Regression covariates included age, sex, race, and study site.

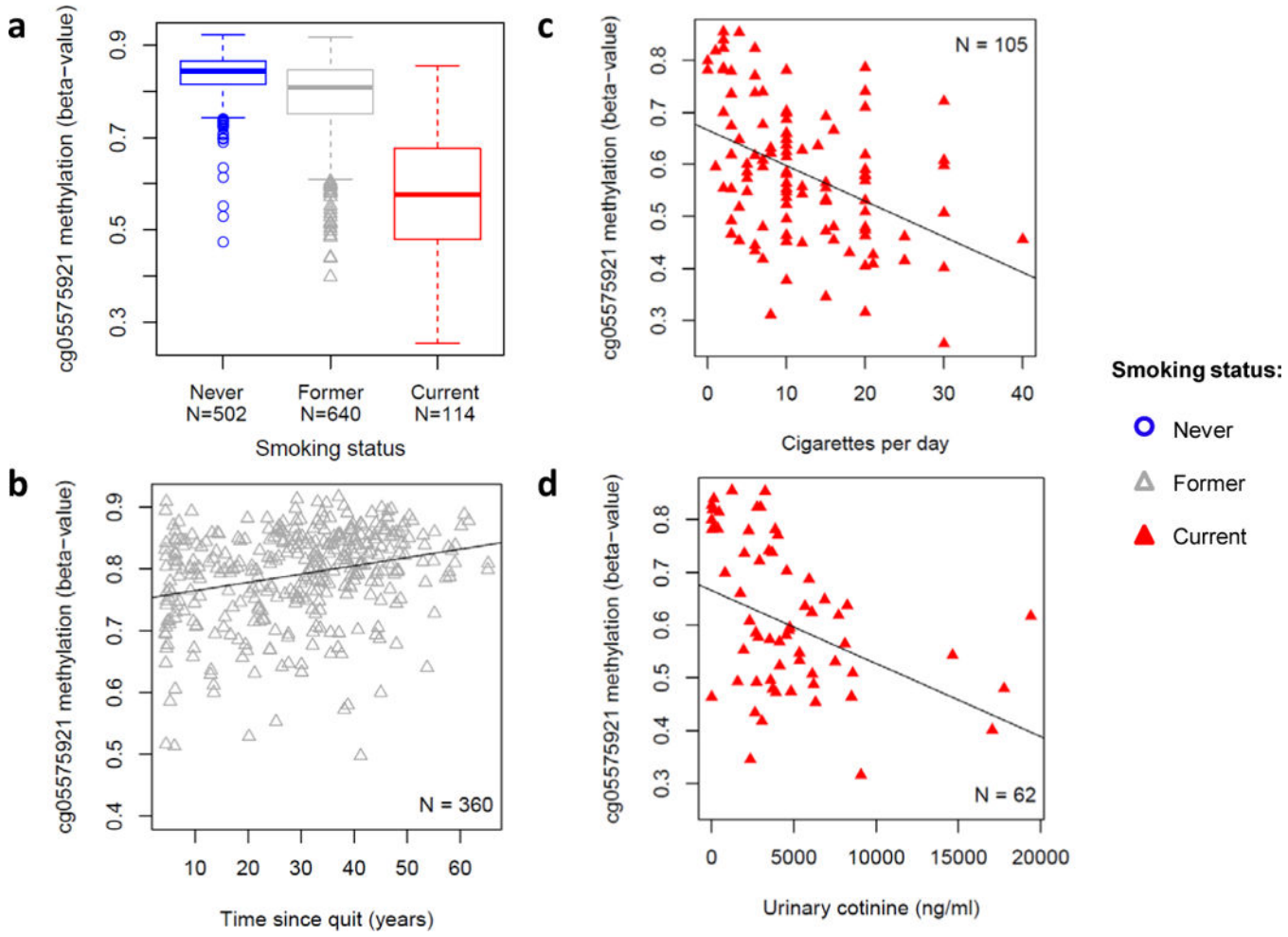

Figure 2.

Most significant smoking-associated methylation profile in monocytes strongly reflects smoking exposure. **(a)** Box plots showing distributions of *AHRR* cg05575921 methylation (y-axis: beta-value) by smoking status **(b–d)** cg05575921 methylation (y axis: beta-value) was **(b)** positively correlated with years since quitting smoking in former smokers (r = 0.35, 95% CI: 0.26 – 0.44, $p < 2\times10^{-16}$) and negatively correlated with **(c)** the number of cigarettes smoked per day (r = −0.39, 95% CI: −0.54 – −0.22, $p = 6.1\times10^{-5}$) and **(d)** urinary cotinine levels (ng/ml) (r = −0.43, 95% CI: −0.62 – −0.21, $p = 9.09\times10^{-4}$) in current smokers. Regression covariates included age, sex, race, and study site.

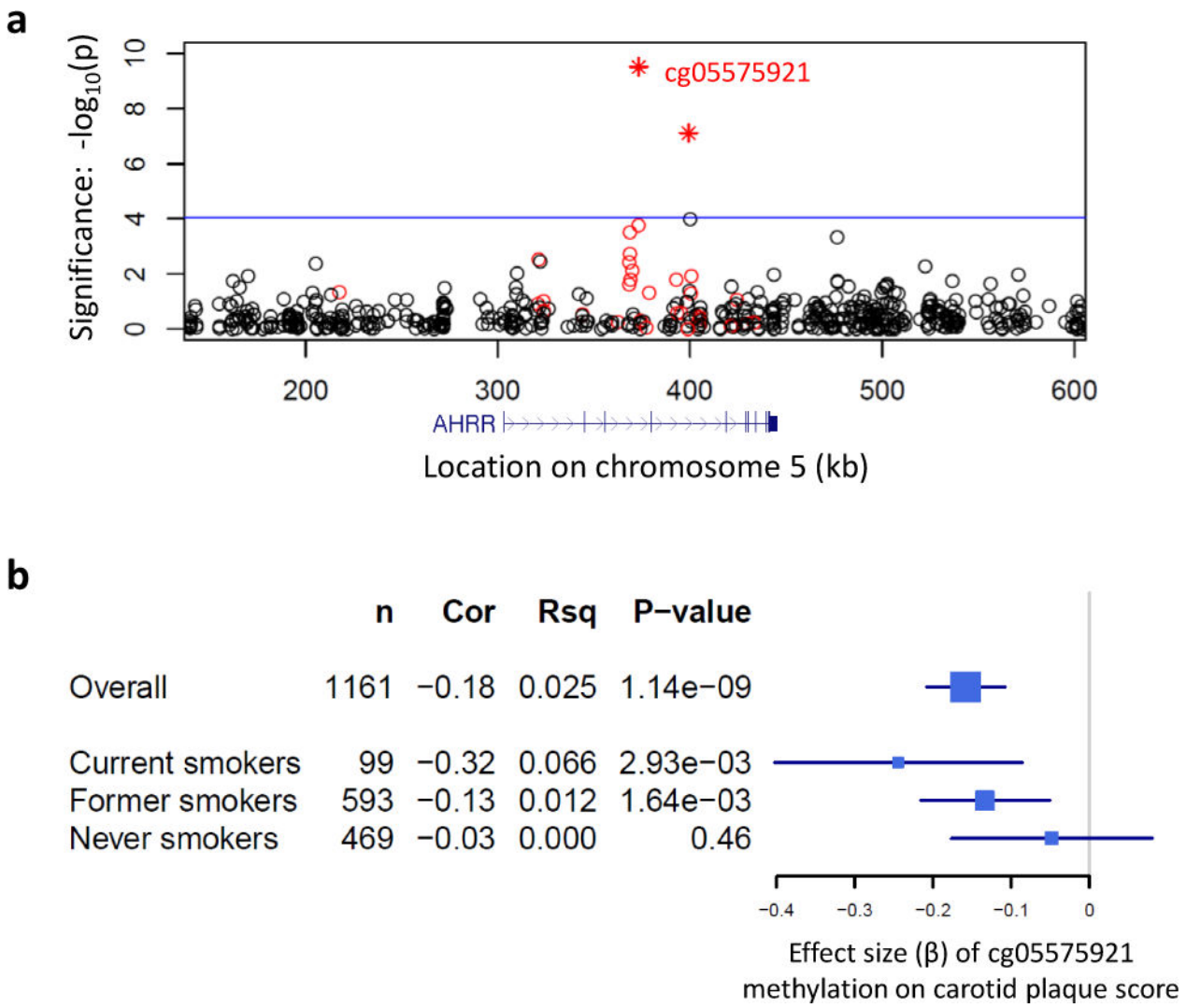

Figure 3.

AHRR methylation is associated with carotid plaque score. **(a)** Regional association plot shows the significance (y-axis: $\log_{10}$ P-value) of the association between methylation (542 CpGs within 150 kb of *AHRR* measured in 1,176 CD14+ monocyte samples) and ultrasound-measured carotid plaque score [ln(plaque score + 1)]. Two methylation profiles (cg05575921 and cg21161138, r = 0.37) significantly associated with carotid plaque score (indicated as stars), based on a Bonferroni adjusted significance threshold for 542 CpGs (α = 0.05; indicated by the blue horizontal line). CpGs with methylation associated with smoking ($FDR_{genome\text{-}wide}<0.05$) are colored red. The genomic position (hg19) of *AHRR* relative to the CpGs is shown below the association plot (in blue). Regression covariates included age, sex, race, and study site. **(b)** Forest plot shows the estimated effect size (x-axis: beta and 95% CI) of cg05575921 methylation on carotid plaque score (overall and stratified by smoking status). Also shown are the sample size (n), the partial Pearson's correlation coefficient (Cor) between methylation and carotid plaque score, the unique variance (Rsq) of carotid plaque score explained by cg05575921 methylation, and the significance (P-value).

Regression covariates included age, sex, race, study site, BMI, LDL cholesterol, diabetes, hypertension, and statin use.

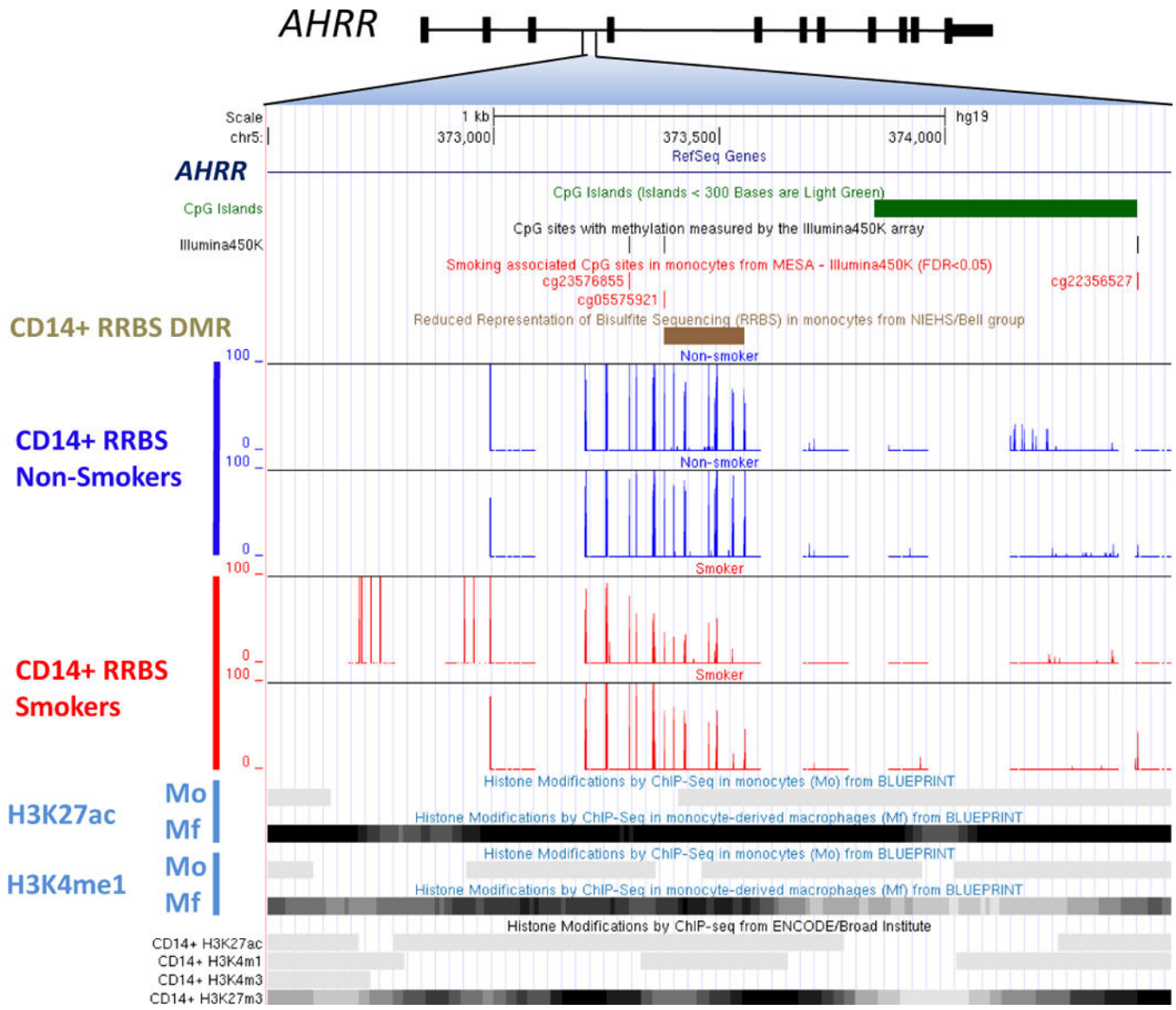

Figure 4.

Functional genomic characterization of the smoking-and carotid plaque score-associated locus within *AHRR*. Cg05575921 is located on chromosome 5p15.33 in an *AHRR* intron and a CpG island 'shore'. Within 500 bp of cg05575921, the Illumina450 microarray captured two additional CpGs, both of which were associated with smoking in MESA ($FDR_{genome\text{-}wide} < 0.05$). Fine mapping using reduced representation of bisulfite sequencing (RRBS) revealed a statistically significant ($p = 2.39\times10^{-5}$) differentially methylation region (DMR, brown bar, chr5:373,378 – 373,556 hg19) by smoking status, including cg05575921 and seven other CpGs within 200 bp that were not included on the Illumina 450K microarray. RRBS was performed in non-MESA monocyte samples from two non-smokers (shown in dark blue) and two current smokers (shown in red), with the resulting percent methylation detected for nearby CpGs shown on the y-axis. This DMR overlaps potential functional features including histone modifications indicative of an active enhancer

(H3K27ac and H3K4me1) in monocyte-derived macrophages (Mf; from BLUEPRINT) and a poised enhancer (active: H3K4m1; repressed: H3K27m3) in a CD14+ monocyte sample (from ENCODE).

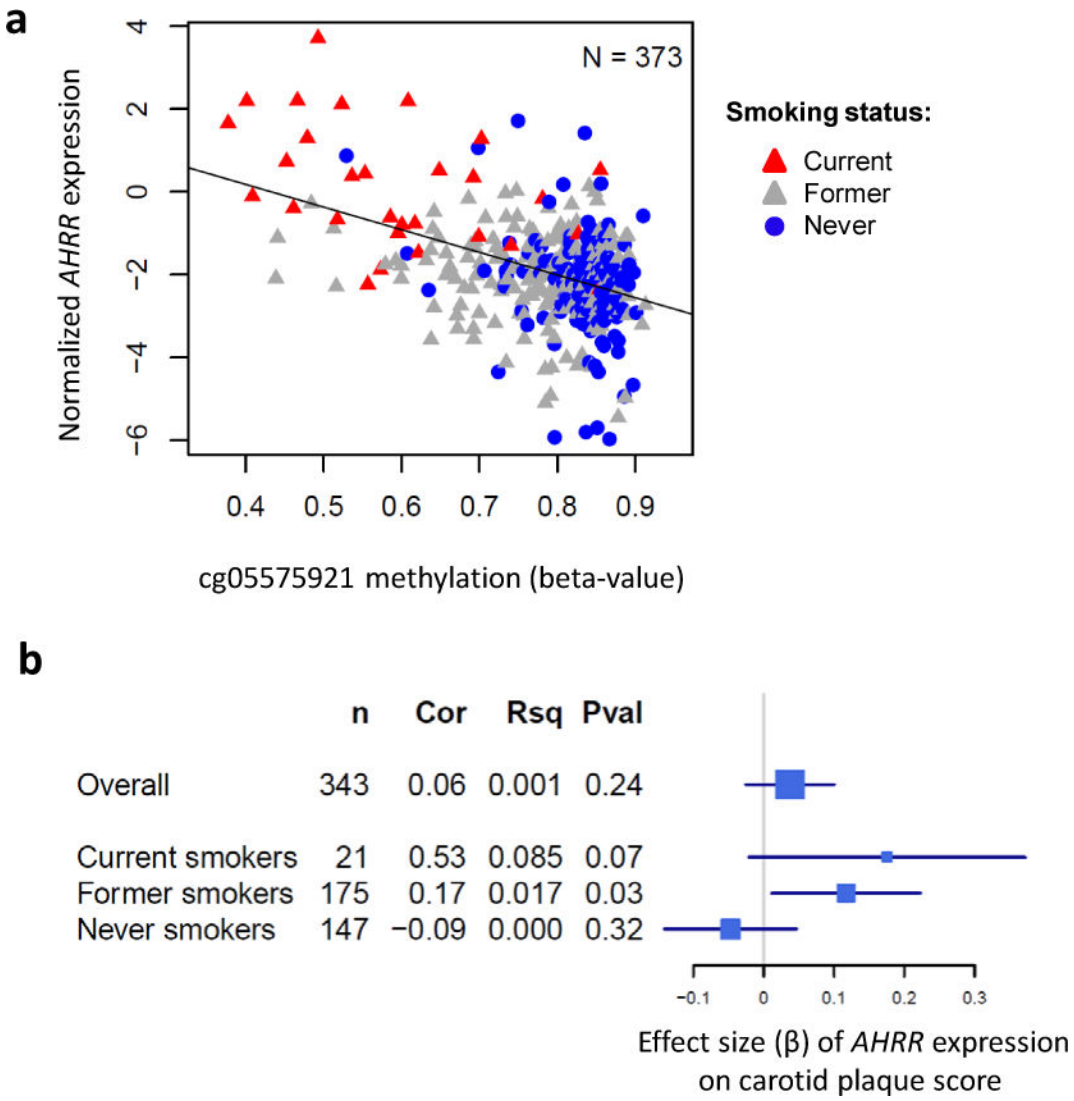

Figure 5.

AHRR expression associations with *AHRR* methylation and carotid plaque score. **(a)** mRNA expression profiles of *AHRR* measured using RNA sequencing (y-axis, normalized values) negatively correlate ($r = -0.42$, $p = 1.36\times10^{-17}$) with methylation of *AHRR* cg05575921 (x-axis, beta-value) measured using the Illumina 450K microarray. Regression covariates included age, sex, race, and study site. **(b)** Forest plot shows the estimated effect size (x-axis: beta and 95% CI) of *AHRR* mRNA expression on carotid plaque score (overall and stratified by smoking status). Also shown are the sample size (n), the partial Pearson's correlation coefficient (Cor) between *AHRR* expression and carotid plaque score, the unique variance of carotid plaque score explained by *AHRR* expression (Rsq), and the significance (Pval). Regression covariates included age, sex, race, BMI, LDL cholesterol, diabetes, and hypertension.

Table 1

Characteristics of the MESA participants

	Overall (n = 1,256)	Cigarette Smoking Status				
		Never (n = 502)	Former (n = 640)		Current (n = 114)	
	Proportion, Mean (SD), or Median (IQR)	Proportion, Mean (SD), or Median (IQR)	Proportion, Mean (SD), or Median (IQR)	P-val*	Proportion, Mean (SD), or Median (IQR)	P-val*
Demographic						
Mean age (years)	70 (9)	70 (10)	70 (9)	ns	65 (8)	<0.001
Gender (% female)	51	62	44	<0.001	45	<0.001
Race						
Caucasian (%)	47	43	50	0.03	45	ns
Hispanic (%)	32	34	31	ns	25	ns
African American (%)	21	23	19	ns	31	ns
Smoking exposure						
Mean pack-years			16 (20)		36 (27)	
Mean cigarettes per day					12 (8)	
Median urine cotinine (ng/ml)†	7 (0)	7 (0)	7 (0)	<0.001	3,708 (3,772)	<0.001
Clinical and laboratory						
Mean BMI	30 (6)	30 (6)	30 (5)	ns	28 (6)	0.05
Mean LDL cholesterol (mg/dl)	105 (32)	108 (32)	102 (33)	0.005	105 (32)	ns
Diabetes (%)	20	19	21	ns	26	ns
Hypertension (%)	61	60	63	ns	54	ns
Median carotid plaque score†	2 (4)	1 (3)	2 (4)	<0.001	2 (4)	0.003

SD, standard deviation; IQR, interquartile range (IQR); ns, P-val > 0.05

*compared to never smokers; t-test to compare means, Wilcoxon rank sum test to compare medians, and chi-square test to compare proportions

†measured in a subset of samples: carotid plaque (n = 1,176) urine cotinine (n = 928)