ORIGINAL RESEARCH

Current Smoking Alters Gene Expression and DNA Methylation in the Nasal Epithelium of Patients with Asthma

Karosham D. Reddy[1,2], Andy Lan[3,4], Ilse M. Boudewijn[3,4], Senani N. H. Rathnayake[1], Gerard H. Koppelman[4,5], Hananeh Aliee[6], Fabian Theis[6], Brian G. Oliver[1,2], Maarten van den Berge[4*], and Alen Faiz[1,2,3,4*]

[1]School of Life Sciences, University of Technology Sydney, Sydney, Australia; [2]Respiratory Cellular and Molecular Biology, Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia; [3]Department of Pulmonary Diseases, [4]Groningen Research Institute for Asthma and COPD, and [5]Beatrix Children's Hospital Department of Pediatric Pulmonology and Pediatric Allergology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; and [6]German Research Center for Health and Environment (GmbH), Helmholtz Centre, Munich, Germany

ORCID IDs: 0000-0001-9002-6930 (K.D.R.); 0000-0002-5631-7264 (S.N.H.R.).

Abstract

Current smoking contributes to worsened asthma prognosis and more severe symptoms and limits the beneficial effects of corticosteroids. As the nasal epithelium can reflect smoking-induced changes in the lower airways, it is a relevant source to investigate changes in gene expression and DNA methylation. This study explores gene expression and DNA methylation changes in current and ex-smokers with asthma. Matched gene expression and epigenome-wide DNA methylation samples collected from nasal brushings of 55 patients enrolled in a clinical trial investigation of current and ex-smoker patients with asthma were analyzed. Differential gene expression and DNA methylation analyses were conducted comparing current smokers with ex-smokers. Expression quantitative trait methylation (eQTM) analysis was completed to explore smoking-relevant genes by CpG sites that differ between current and ex-smokers. To investigate the relevance of the smoking-associated DNA methylation changes for the lower airways, significant CpG sites were explored in bronchial biopsies from patients who had stopped smoking. A total of 809 genes and 18,814 CpG sites were differentially associated with current smoking in the nose. The *cis*-eQTM analysis uncovered 171 CpG sites with a methylation status associated with smoking-related gene expression, including *AHRR*, *ALDH3A1*, *CYP1A1*, *and CYP1B1*. The methylation status of CpG sites altered by current smoking reversed with 1 year of smoking cessation. We confirm that current smoking alters epigenetic patterns and affects gene expression in the nasal epithelium of patients with asthma, which is partially reversible in bronchial biopsies after smoking cessation. We demonstrate the ability to discern molecular changes in the nasal epithelium, presenting this as a tool in future investigations into disease-relevant effects of tobacco smoke.

Keywords: asthma; nasal epithelium; gene expression; DNA methylation; smoking

The long-term adverse health effects of cigarette smoking are well established. Short-term tobacco smoke exposure negatively affects lung function, whereas long-term exposure is linked with chronic respiratory disease (1). Cigarette smoking increases the risk of asthma development and severity (2). Asthma causes 450,000 deaths worldwide annually, and 350 million people have it (3). The smoking rate among patients with asthma is similar to that of the general population, and they are less likely to quit (2). Smoking patients

(Received in original form December 2, 2020; accepted in final form April 19, 2021)

*These authors contributed equally to this work.

Supported by Australian Government Research Training Program Scholarships (K.D.R. and S.N.H.R.).

Author Contributions: M.v.d.B. and A.F. conceived the idea. K.D.R., A.L., I.M.B., S.N.H.R., M.v.d.B., and A.F. contributed to, performed, and discussed data analysis and interpretation. H.A. and F.T. contributed to the data analysis. The National Heart, Lung, and Blood Institute LungMAP Consortium contributed to the data analysis. K.D.R., I.M.B., G.H.K., B.G.O., M.v.d.B., and A.F. discussed and contributed to the drafted manuscript composition.

Correspondence and requests for reprints should be addressed to Karosham D. Reddy, B.Sc. (Hons.), Woolcock Institute of Medical Research, Cells Group, Level 3, 431 Glebe Point Road, Glebe NSW 2037, Australia. E-mail: Karosham.Reddy@sydney.edu.au.

This article has a related editorial.

This article has a data supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

Am J Respir Cell Mol Biol Vol 65, Iss 4, pp 366–377, October 2021

Originally Published in Press as DOI: 10.1165/rcmb.2020-0553OC on May 14, 2021
Internet address: www.atsjournals.org

with asthma are a complex and vulnerable population because they are more susceptible to the development of respiratory diseases such as chronic obstructive pulmonary disease (COPD), in addition to a poor asthma prognosis (4). Smoking patients with asthma experience more life-threatening exacerbations, making disease management difficult. The beneficial effect of inhaled corticosteroids on lung function outcomes is negated by smoking (5). As such, an altered physiological response occurs in the airways because of smoke exposure, causing worsened asthma disease outcomes (2).

The respiratory airway epithelium is the body's first line of defense against inhaled toxic agents and exhibits smoking-associated changes (6). Although bronchial samples are considered the most representative (7), the nasal epithelium can reflect smoking-induced changes in the lower airways. Functionally similar genes in paired nasal and bronchial epithelium samples can differentiate between healthy never-smokers and current smokers (7–11). Although nasal brushings do not perfectly mirror the lower airways, the response to tobacco smoking is shown to be reflective of transcriptional changes, and they are an attractive alternative to invasive bronchial biopsies.

Epigenetics refers to chemical changes to the genome that do not alter the nucleotide sequence (12). DNA methylation is the most widely studied epigenetic mark and is known to influence gene expression. Modifications in DNA methylation can be caused by environmental stimuli, such as cigarette smoke (13, 14). These changes can persist and contribute to increased disease susceptibility (15). As smoking increases asthma severity and reduces treatment effectiveness, greater understanding of gene expression and DNA methylation changes due to smoke exposure is needed. Methylation commonly occurs at positions in the DNA where a cytosine nucleotide is followed by a guanine nucleotide, these locations are referred to as CpG sites. Differentially methylated CpG sites exist between current, ex-, and never-smokers in samples taken from the lung and blood tissues. These sites are associated with the genes *ALDH3A1*, *CYP1A1*, *CYP1B1*, and *AHRR* (16, 17). Most studies investigate DNA methylation with blood samples; however, examination of structural cells is vital because they are exposed to smoke. Some studies investigate changes in gene expression and DNA methylation using patient-matched lung samples (11, 18, 19) but are limited by size. These analyses enable a direct investigation between gene expression and DNA methylation, increasing our understanding of molecular processes that influence the tobacco smoke response.

The current study investigates nasal differential gene expression and DNA methylation profiles between current and ex-smokers with asthma. We examined differential gene expression due to current smoking in two *in vitro* datasets and in two *in vivo* smoking cessation datasets. We tested these gene expression patterns for an association with relevant clinical measures in asthma. Analysis of DNA methylation and gene expression using expression quantitative trait methylation (eQTM) analysis offered insight into the relationship between DNA methylation status and gene expression relative to smoking status. To address the dynamics of smoking-associated changes in the airways, significant nasal CpG sites were explored in bronchial biopsies from patients who had stopped smoking. This study contributes to the current understanding of nasal gene expression as an alternative to invasive bronchial biopsies. It elucidates mechanisms and identifies appropriate biomarkers affecting processes in smoking patients with asthma.

Methods

Participants and Sample Collection

Study subjects were selected from participants of the OLiVIA (Effects of Extra-Fine Particle HFA-Beclomethasone Versus Coarse Particle Treatment in Smokers and Ex-Smokers with Asthma) study (ClinicalTrials.gov #NCT01741285) (20). This study was approved by the local medical ethics committee, with written and informed consent received before collection. Patients were included if they were either current or ex-smokers (smoking cessation $\geq$6 mo), aged 18–65 years old, and had doctor-diagnosed asthma with a smoking history of 5 pack-years or more. Exclusion criteria comprised treatment with oral steroids, forced expiratory volume in 1 second (FEV_1) of $\leq$1.2 L, and an upper respiratory tract infection $\leq$4 weeks before inclusion, had used inhaled corticosteroids in the previous 4–6 weeks, and an asthma exacerbation $\leq$6 weeks before inclusion. A schematic representation of the study design and analysis is included in the supplementary file (Figure E1 in the data supplement).

Sample Collection, Processing, and Quality Assessment

Detailed methods are presented in the data supplement. Nasal brushings were taken from the inferior turbinate of patients. Two separate brushings were taken for RNA and DNA collections. RNA was isolated from samples using an RNeasy kit (#74106, Qiagen) and quality tested before single-end sequencing. Quality control of gene expression was performed using R, with concordance between reported sex and sex-associated genes (*XIST* and Y-chromosome genes) confirmed. Genes with mean expression of $\leq$1 fragments/million across all samples were removed. The influence of technical variation was assessed via principal component analysis using the R package EDASeq. The variation in library sizes was explored for gene expression among samples by making a relative log expression plots. The library size cut-off ensured that sequences would be larger than one-quarter of the median of all library sizes. Data were $\log_2$ transformed and normalized.

DNA extraction was completed on a second nasal brushing using a DNA extraction kit (#51306, Qiagen), according to the manufacturer's instructions. DNA methylation processing was completed using the Infinium MethylationEPIC BeadChip array. DNA methylation data quality and normalization were also assessed using the R packages *MethylAid* and *watermelon* with five diagnostic filter variables with the following thresholds: MU = 10, OP = 12, BS = 11.50, HC = 12.75, and DP = 0.95 (the definitions of these variables are provided in the data supplement). Discordant samples were discarded as "swapped" or "contaminated" samples. Probes with high intensities indistinguishable from the background, cross-reactive probes, and type I probes with high-intensity signals were discarded. Probes were excluded if intensities *1*) were indistinguishable from background levels, *2*) were cross-reactive, or *3*) were type I probes displaying high-intensity signals. The final number of CpG sites tested after filtration was 797,128 sites. Normalization was completed using the *dasen* method in the *watermelon* package, in which type 1 and type 2 background intensities are equalized separately using quantile normalization before the generation of β values. Detailed sample processing methodology is included in the data supplement.

Smoker versus Ex-Smoker: Differential Gene Expression Analysis

Differential expression was assessed using least-squares linear modeling in the R package limma. We corrected for age and sex by including each factor as a covariate in the linear model, whereas smoking status was the tested

variable. Correction for multiple testing was applied through controlling the false discovery rate (FDR) using the Benjamin-Hochberg (BH) procedure. Significant differentially expressed genes were determined with an FDR of less than 0.05 and a fold-change (FC) greater than |1.0|.

Cellular Deconvolution Analysis

Cellular deconvolution was completed using the Cibersort method on single-cell RNA sequencing signatures for basal, ciliated, and mucus-secretory cells (club and goblet cells) to investigate the composition of nasal brushings. Cell composition was then compared between current and ex-smokers. The complete methodology is included in the data supplement.

Gene Expression Profiling and Mapping with *In Vitro* Datasets

Independent publicly available *in vitro* datasets on the Gene Expression Omnibus were used to conduct gene set expression analysis (GSEA) and gene set variation analysis (GSVA). GSEA used datasets analyzing airway epithelial cells from patients exposed to cigarette smoke for 30 minutes four times (GSE30660) before gene expression analysis and bronchial epithelium air–liquid interface (ALI) cultures exposed cigarette smoke for 24 hours before gene expression analysis (GSE82137) (21). We conducted GSVA on a dataset that analyzed nasal epithelium gene expression from smokers who had quit smoking, over a 6-month period (GSE83364) (22). We also conducted GSVA on datasets that analyzed changes gene expression in A549 cells in which NRF2 (GSE113519) and AHR (GSE109576) (23) are inhibited. The methodology and patient demographics (where appropriate) are located in the data supplement.

Identification of Biological Pathways Altered by Smoke Exposure in the Nose

Pathway analysis was completed to identify the biological pathways that involve significant differentially expressed genes in nasal epithelium. The g:Profiler web-based tool was used to conduct the analysis, using a subset of the top 25 significant (FDR $<$ 0.05 and FC $>$ |1.0|) genes.

Association between Gene Expression Variation and Clinical Parameters

The top 25 significantly increased genes from the differential expression analysis were analyzed by GSVA to generate gene variation enrichment scores for each patient. This "smoking score" was analyzed for correlation with multiple clinical parameters, including pack-years, age, lung function (FEV_1/forced vital capacity), and peripheral immune cell counts. Data were analyzed using a nonparametric Spearman's correlation test, with statistical significance determined at $P < 0.05$.

Smoker versus Ex-Smoker: Differential DNA Methylation Analysis

Differential DNA methylation analysis used robust linear modeling in the R package limma. To remove the dependence of the mean (heteroscedasticity), β values were logit-transformed to *M* values. Smoking status, age, and sex were covariates in the linear model. Major principal components, explaining 95% of the variation in technical control probes, were included as covariates to account for variation, as recommended by Lehne and colleagues (24). The Δβ values were calculated by subtracting the average β value per CpG site for the current smoker group from the average β value for the ex-smoker group. Correction for multiple testing was controlled at FDR $<$ 0.05 using the BH procedure. The BH correction method was preferred to the Bonferroni method to ensure consistency across analyses; because BH is the standard method for differential gene expression, it was used for DNA methylation analysis. As such, DNA methylation analysis was less restrictive to enable an appropriate comparison with gene expression. Bonferroni significance was still indicated in the figures for reference.

Expression Quantitative Trait Methylation Analysis

cis-eQTM analysis was performed using the R package MatrixEQTL (version 2.2). Age and sex were entered as covariates in the model. For the analysis, a 1-Mb region on either side of the gene was selected. The analysis was completed for each significant gene from the differential gene expression analysis (FDR $<$ 0.05 and FC $>$ |1.0|) individually against significant CpG sites from the differential DNA methylation analysis (FDR $<$ 0.05). Correction for multiple testing was completed using the BH procedure.

Mapping of DNA Methylation Sites with Transcription Factor Binding Sites

The relevance of significant differential DNA methylation sites was profiled by comparison with independent online chromatin IP sequencing (ChIP-Seq) datasets. GSE90550 investigates the binding locations of smoking-related transcription factors AHR and AHRR in MCF-7 cells treated with DMSO or 2,3,7,8-Tetrachlorodibenzodioxin (TCDD) for 24 hours (25), whereas GSE75812 investigates transcription factor NRF2 binding locations in BEAS-2B cells treated with sulforaphane for 5 hours (26). The location of transcription binding sites was aligned with gene transcript locations and significant differentially methylated CpG sites identified in the eQTM analysis.

Longitudinal Analysis of Changes to DNA Methylation Sites with Smoking Cessation

The longitudinal effects of smoking cessation were evaluated using an independent 1-year smoking cessation model (27). The independent study completed differential DNA methylation analysis on paired bronchial biopsies taken from smokers with COPD before and after they had stopped smoking for 1 year. From a total of 63 participants, only 19 patients had successfully ceased smoking by the end of the study. Significant *cis*-eQTM CpG sites identified in the nasal epithelium as part of the current study were extracted from the differential DNA methylation results of the smoking cessation analysis. We investigated how the DNA methylation sites affected by current smoking the nasal epithelium are altered in bronchial samples upon 1 year of smoking cessation.

Statistical Analysis

Differential gene expression, differential DNA methylation, and eQTM analyses were all conducted using R software (version 4.0.1). All other analyses were performed using the GraphPad Prism software (version 8), with statistical significance determined at $P < 0.05$.

Results

Patient Demographics and Sample Processing

Ten samples were removed after mRNA quality checking processes, leaving a final number of 55 high-quality samples for analysis. Gene expression data were collected from these brushings. After processing, 31 samples had sufficient quality DNA to undergo DNA methylation analysis. Finally, 29 samples remained in which both gene expression and DNA methylation data were

available for eQTM analysis. Patient information and demographics are summarized in Table 1.

Differential Gene Expression in Nasal Epithelium between Current and Ex-Smokers

Differential gene expression analysis identified 809 genes being upregulated or downregulated in the nasal epithelium by current smoking. In total, 153 genes showed increased expression and 656 showed decreased expression (FDR $<$ 0.05 and log-fold change $>$ |1.0|; Figure 1A). Using hierarchy clustering, the majority of smokers clustered separately from the ex-smokers. Table E1 lists the top significantly differentially expressed genes. Among the significantly upregulated genes in active smoker nasal epithelium were *CYP1A1*, *CYP1B1*, *ALDH3A1*, *NQO1*, and *AHRR*. The most significant downregulated genes included *SAA1*, *SAA2*, *CYP4X1*, *COL9A2*, and *CHST4* (Figure 1B). Of particular note are genes *CYP1A1* and *CYP1B1*, which were greatly upregulated in current smokers. with a log-fold change of 7.44 (FDR = 5.80×10^{-19}) and 4.31 (FDR = 8.19×10^{-14}), respectively (Figure 1B).

Cellular Composition of Nasal Brushings Are Similar between Current and Ex-Smoker Patients

To assess the uniformity of nasal samples and whether epithelium from current and ex-smokers differs, we conducted cellular deconvolution using the support vector regression (SVR) and non-negative least squares (NNLS) analytical methods. The following four cell types were most abundant across the samples: goblet cells, basal epithelium, dendritic cells, and ciliated epithelium. Both techniques reported minimal variability in cell composition between samples taken from current and ex-smokers (Figure E2). Therefore, we observed no difference in sample cell composition from nasal brushings between current and ex-smokers.

Smoking Affected Genes in the Nose Are Similarly Altered in Independent *In Vitro* Analysis of Airway Cells

To investigate whether differences in gene expression were caused by changes in the nasal epithelium or to inflammatory cells, we analyzed *in vitro* datasets of airway epithelial cells exposed to gaseous tobacco smoke. Using GSEA, we investigated whether the 809 genes altered by current smoking in nasal epithelium are similarly affected by cigarette smoke in independent ALI cell culture datasets. The genes *ALDH3A1*, *AHRR*, *CYP1A1*, *CYP1B1*, *CYP1B1-AS1*, and *SLC7A11* were core enriched among the genes that had increased expression by smoke exposure in the human epithelial ALI dataset (GSE30660; Figure 2A) and the bronchial epithelial ALI dataset (GSE82137; Figure 2B). The genes *CYP4X1*, *COL9A2*, *SAA1*, and SAA2 were core enriched among genes downregulated by smoke exposure in the ALI datasets (Figures 2A and 2B). In Figures 2A and 2B, enriched genes are represented by black vertical lines indicating the amount these differentially expressed genes from the dataset are overrepresented in nasal epithelium gene expression. Core enriched genes, represented by red vertical lines, are genes that are highly represented in a gene set, indicating that they are highly likely to be associated with the stimulus. This shows that gene expression changes caused by smoke exposure in the nasal epithelium align with independent findings, indicating that gene expression occurring in the upper airways in response to cigarette smoke is reflected in airway epithelial cells in an *in vitro* setting.

AHR and NRF2 Biological Signaling Pathways Are Involved in the Nasal Response to Smoking

Next, to determine which pathways are enriched in our analysis, we conducted a pathway analysis using g:Profiler. Here, we identified five pathways, including general pathways such as response to external stimulus and metabolism of xenobiotics together with well-known smoking-related pathways such as the AHR pathway and the NRF2 pathway (Table 2).

To investigate these pathways, we analyzed two publicly available datasets that

Table 1. Clinical Characteristics of the Study Population

Characteristic	Differential Gene Expression Analysis		Differential Methylation Analysis		eQTM Analysis	
	Current Smoker	Ex-Smoker	Current Smoker	Ex-Smoker	Current Smoker	Ex-Smoker
N	26	29	14	17	13	16
Female, %	53.85	55.17	53.85	55.17	46.15	56.25
Age, yr	40 ± 11	50 ± 10*	38 ± 11	50 ± 11†	38 ± 12	49 ± 11*
Pack-years	22 ± 13	21 ± 19	18 ± 10	17 ± 10	19 ± 9	17 ± 10
Daily cigarettes	17.68 ± 10.54	—	17.68 ± 10.54	—	18.15 ± 7.18	—
FEV_1% predicted	86.08 ± 13.75	89.34 ± 18.17	82.14 ± 12.45	85.06 ± 17.97	82.38 ± 10.03	87.69 ± 16.81
FEV_1/FVC, %	83.87 ± 11.73	82.64 ± 13.82	82.08 ± 10.21	80.86 ± 12.29	82.88 ± 11.16	81.88 ± 10.80
Peripheral blood cell absolute count, $\times 10^3$ cell/μl						
Leukocytes	9.14 ± 2.15	6.64 ± 1.50‡	9.14 ± 2.15	6.64 ± 1.50†	9.10 ± 2.19	6.39 ± 1.72†
Neutrophils	5.21 ± 1.23	3.74 ± 1.57†	5.21 ± 1.23	3.74 ± 1.57†	5.36 ± 1.65	3.65 ± 1.47†
Lymphocytes	2.80 ± 0.81	2.04 ± 0.46‡	2.80 ± 0.81	2.04 ± 0.46†	2.68 ± 0.81	1.91 ± 0.33†
Monocytes	0.70 ± 0.17	0.53 ± 0.16†	0.70 ± 0.17	0.53 ± 0.16†	0.68 ± 0.15	0.51 ± 0.16†
Eosinophils	0.36 ± 0.20	0.27 ± 0.18	0.36 ± 0.20	0.27 ± 0.18	0.32 ± 0.18	0.26 ± 0.19

Definition of abbreviations: eQTM = expression quantitative trait methylation; FEV_1 = forced expiratory volume in 1 second; FVC = forced vital capacity.
Data are presented as arithmetic mean ± SD and analyzed using a Mann-Whitney test.
*$P < 0.05$ (current smoker vs. ex-smoker).
†$P < 0.001$ (current smoker vs. ex-smoker).
‡$P < 0.0001$ (current smoker vs. ex-smoker).

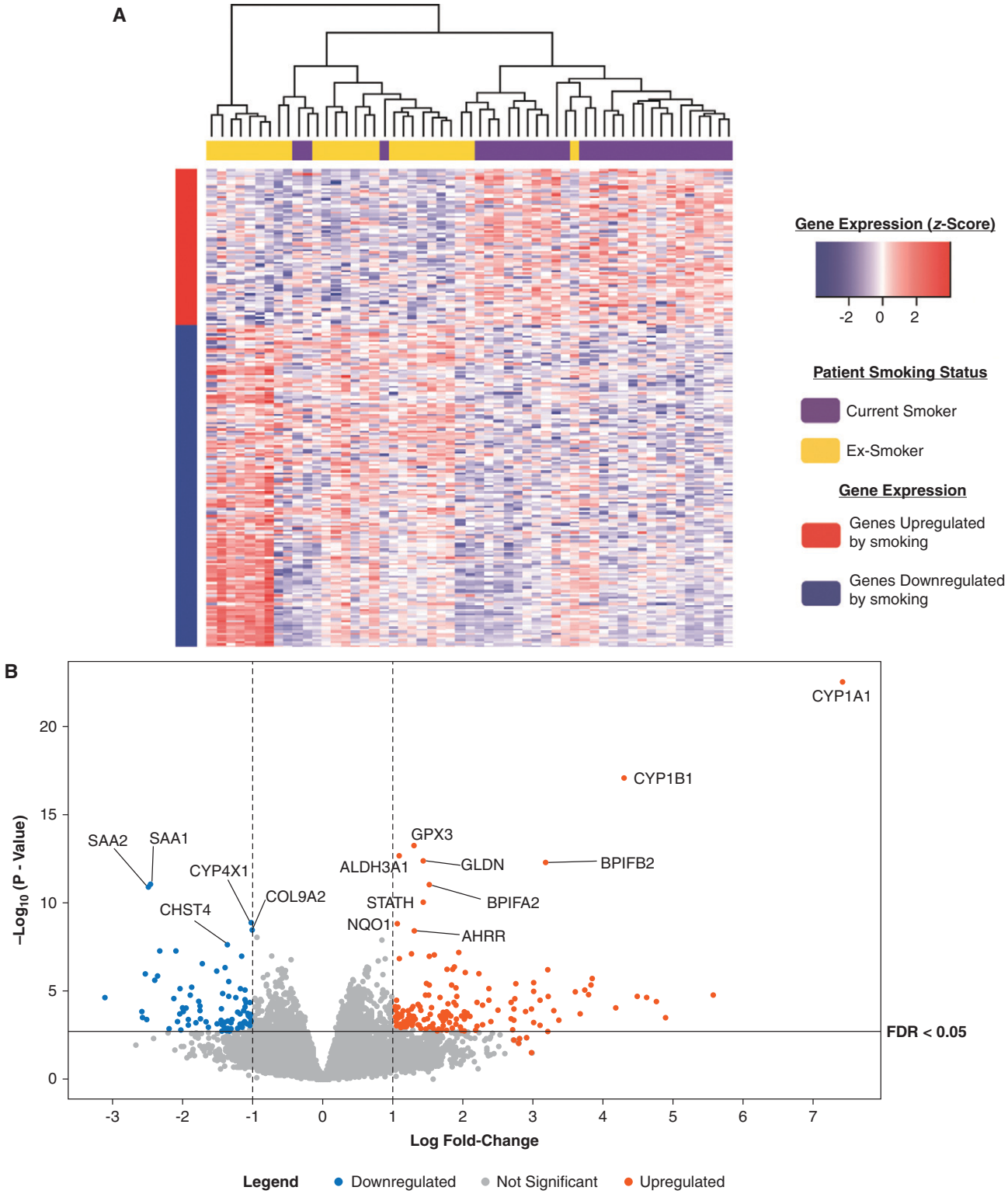

Figure 1. Heatmap and volcano plot visualization of differential gene expression between current smokers and ex-smokers. (*A*) An unsupervised heatmap of the 809 significant genes differentially expressed between smokers (purple) and ex-smokers (gold). Increased genes are colored red, and decreased genes are colored blue. (*B*) Volcano plot of $-\log_{10}$ (*P* value) against log-fold change in gene expression in smokers compared with ex-smokers. Significantly increased and decreased expressed genes in smokers are indicated by red and blue, respectively. Dotted vertical lines indicate fold-change of |1.0|. Statistical significance was determined at FDR-adjusted *P* value of less than 0.05, determined using the Benjamini-Hochberg method. $n = 26$ (current smokers) and 29 (ex-smokers). AHRR = aryl-hydrocarbon receptor repressor; ALDH3A1 = aldehyde dehydrogenase 1 family, member A3; CYP1A1 = cytochrome P450, family 1, subfamily A, polypeptide 1; CYP1B1 = cytochrome P450, family 1, subfamily B, polypeptide 1; FDR = false discovery rate.

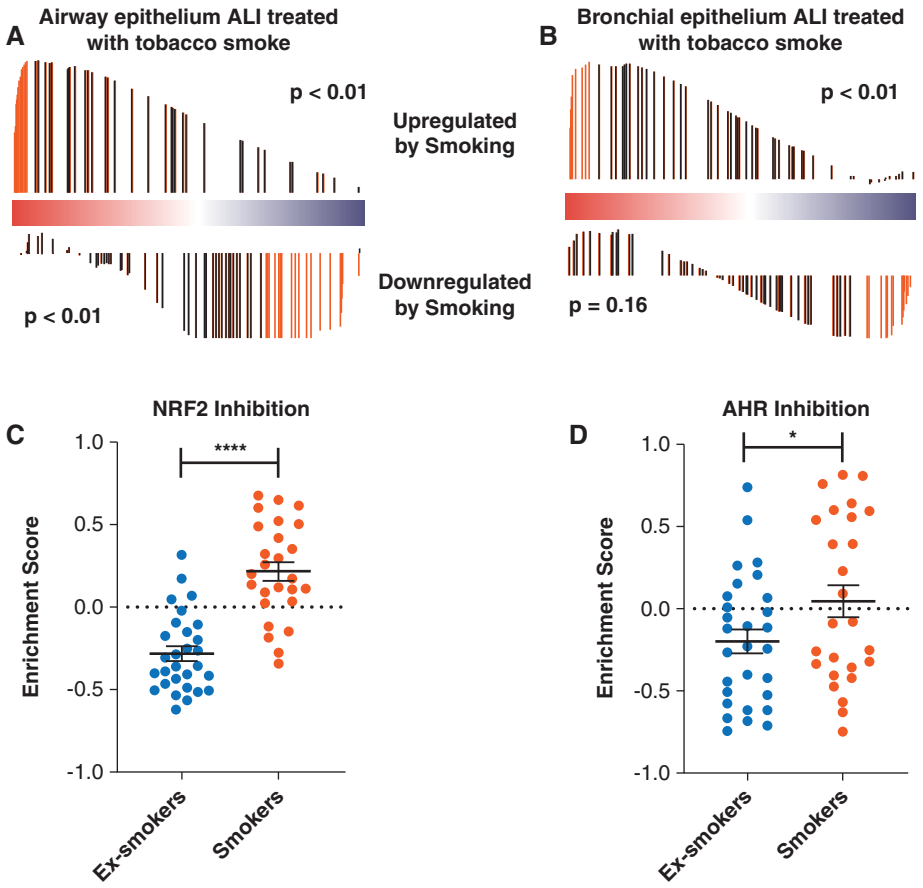

Figure 2. Analysis of smoking affected genes expressed in the nasal epithelium within independent *in vitro* datasets. (*A*) Gene set expression analysis (GSEA) enrichment of genes upregulated and downregulated in air–liquid interface (ALI) cultures of primary airway epithelial cells exposed cigarette smoke compared with nasal epithelium genes from patients with asthma. (*B*) GSEA enrichment of genes increased or decreased in bronchial epithelial ALI cultures upon cigarette smoke exposure compared with nasal epithelium genes from patients with asthma. Colored horizontal bars represent genes identified in each independent dataset ranked by *t* value; red indicates a positive association with gene expression, whereas blue indicates a negative association. The black vertical lines represent differentially expressed genes between smokers and ex-smokers in the nasal epithelium. The height of the black lines indicates the running enrichment scores of each gene, signifying the amount that the genes are overrepresented in the nasal epithelium. Red bars indicate core enriched genes; these are genes highly represented in the gene set. (*C* and *D*) Gene set variation analysis analyzing genes that indicate decreased expression in A549 cells upon NRF2 (*C*) or AHR inhibition (*D*) in A549 cells have increased expression in current smoker (red) nasal epithelial compared with ex-smokers (blue). Dots represent individual patients from nasal differential gene expression analysis. Data are presented as mean ± SEM and analyzed using Student's parametric *t* test. Statistical significance between ex-smoker and current smoker groups is indicated by $^{*}P < 0.05$ and $^{****}P < 0.0001$. $n = 26$ (current smokers) and 29 (ex-smokers) for all analyses. AHR = aryl-hydrocarbon receptor.

specifically inhibited NRF2 (GSE113519) or AHRR (GSE109576) in airway epithelial cells. To understand the function of differentially expressed genes from these datasets in nasal epithelium from smokers with asthma, GSVA analysis was conducted. Because of limited genes indicating increased expression (FDR < 0.05) in these datasets, we focused on gene expression decreased by NRF2 and AHR inhibition. Genes that decreased in expression upon inhibition of either protein demonstrate an increased amount of expression in current smoker nasal epithelium (Figures 2C and 2D). These results indicate that both NRF2 and AHR have a significant role in the regulation of genes associated with current smoking.

Gene Expression Levels Reverse with Smoking Cessation

To explore the effect of smoking cessation on genes that change with smoke exposure, we performed a GSVA using an independent *in vivo* nasal smoking cessation dataset (22). A subset of the 25 top significantly increased and a second subset of the 25 top significantly decreased genes in the nasal epithelium were analyzed for changes associated with smoking cessation. Gene expression was analyzed 4, 8, 16, and 24 weeks after cessation. After 4 weeks, genes upregulated by current smoking in the nasal epithelium are downregulated, which is sustained over 24 weeks (Figure 3A). Comparatively, genes downregulated by current smoking in the nasal epithelium are upregulated 8 weeks after smoking cessation, which is sustained after 16 weeks (Figure 3B); at 24 weeks, this was no longer significant. It is unclear what might be the physiological effect of this change.

GSVA Smoking Score Correlates with Clinical Parameters

After this, we investigated whether transcriptional signatures were associated with clinical parameters (measures of lung function and peripheral inflammation). The subset of the top 25 significantly increased genes (FDR < 0.05 and FC > 1) in the nasal epithelium was used to generate a GSVA enrichment "smoking score." Figure 4 shows a positive correlation with our gene expression signature and pack-years ($P = 0.0077$). A significant correlation was also observed for peripheral blood cells such as leukocytes ($P < 0.0001$), neutrophils ($P = 0.0012$), monocytes ($P = 0.0015$), and lymphocytes ($P = 0.0010$); these have been included in the data supplement (Figure E3). However, when the correlation analysis is completed within each group, no significant correlation is observed for either current or ex-smokers ($P > 0.05$). No correlation was observed with age, FEV_1/forced vital capacity, number of cigarettes per day, or absolute eosinophil counts. No correlation is reported within the current smoker group for the clinical parameters against the GSVA smoking enrichment score, despite a correlation observed for pack-years. This might be explained by a lack of a broad range of smoking levels within the current smoker group. Therefore, the inclusion of a greater range of smoking levels will offer more nuanced insight between the ability to measure the effects of smoking on clinical parameters using this methodology.

Table 2. Summary of g: Profiler Pathway Analysis Investigating the Top 20 Significantly Upregulated Genes by Current Smoking

Pathway Name	Term ID	FDR
Cellular response to xenobiotic stimulus	GO:0071466	7.98×10^{-04}
Metabolism of xenobiotics by cytochrome P450	KEGG:00980	8.53×10^{-03}
Aryl Hydrocarbon receptor pathway	WP:WP2873	5.20×10^{-07}
Nuclear receptors meta-pathway	WP:WP2882	1.43×10^{-05}
NRF2 pathway	WP:WP2884	5.21×10^{-03}

Definition of abbreviations: FDR = false discovery rate; GO = Gene Ontology; ID = identification; KEGG = Kyoto Encyclopedia of Genes and Genomes; WP = WikiPathways.

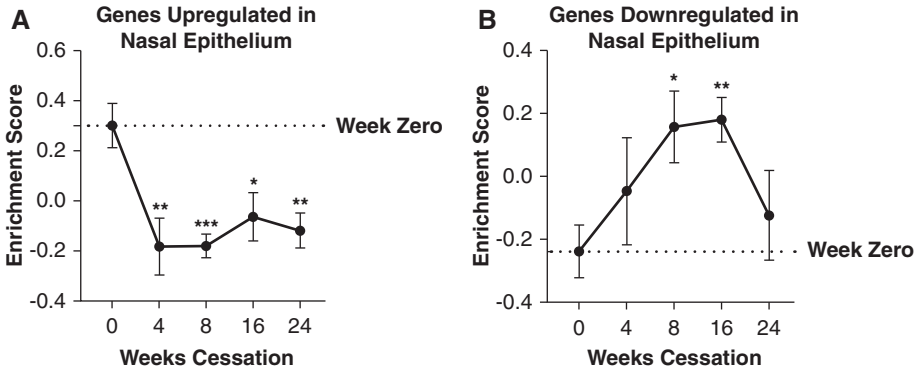

Figure 3. Gene set variation analysis demonstrating gene expression reverses upon smoking cessation. (*A*) Upon smoking cessation, genes upregulated by current smoking in the nasal epithelium are downregulated for 24 weeks. (*B*) Smoking cessation causes genes that are downregulated by current smoking in the nasal epithelium to increase in expression up to 16 weeks. Data are presented as mean ± SEM and analyzed using Student's parametric *t* test. Statistical significance to baseline (Week 0) is indicated by $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. $n = 8$.

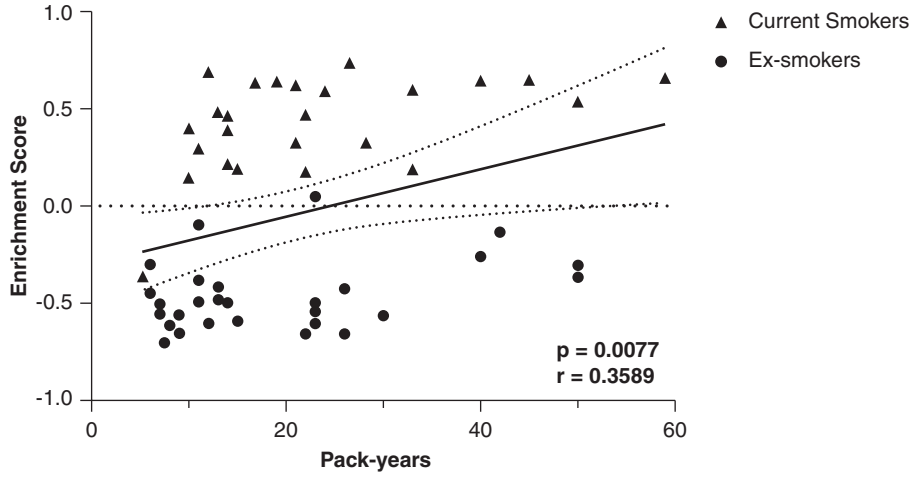

Figure 4. Correlation between the gene variation score of smoking-related upregulated genes and pack-years. A gene enrichment score of 0 is indicated by a dotted line, a solid regression line illustrates the direction of correlation between the *X* and *Y* variables, with 95% confidence intervals indicated by a dotted line. Triangles represent current smokers. Circles indicate ex-smokers. Data were analyzed by nonparametric Spearman's correlation (*r* = correlation coefficient). Statistical significance was determined at $P < 0.05$. $n = 26$ (current smokers) and 29 (ex-smokers) for all analyses.

Differential DNA Methylation Patterns Exist between Current and Ex-Smokers

To investigate the effect of current smoking on epigenetics (DNA methylation), we conducted a differential DNA methylation analysis. We identified 18,814 CpG sites altered by current smoking, with 10,697 sites with higher methylation and 8,117 sites with lower methylation (FDR < 0.05) in current smokers relative to ex-smokers (Figure 5A). Table E3 lists the top significantly differentially methylated CpG sites. Figure 5B highlights the most significant CpG sites that are associated with either increased (red) or decreased (blue) methylation (Δβ value) in current smokers compared with ex-smokers. Figure E5 illustrates this relationship with respect to log-fold change.

The cellular composition of samples can affect results from differential DNA methylation analysis (28). We completed a sensitivity analysis by including the cellular deconvolution data as covariates in the linear model. These results reported that 95% of significantly affected methylation sites were retained when cell composition was adjusted for, compared with the original analysis.

Changes in DNA Methylation Are Associated with Changes in the Expression of Smoking-related Genes

The *cis*-eQTM analysis aimed to investigate whether an association exists between significantly differentially expressed genes and methylated CpG sites in response to current smoking. A total of 171 *cis*-eQTM relationships were identified with key smoking-related genes (*ALDH3A1*, *AHRR*, *BPIFA2*, *CYP1A1*, *CYP1B1*, *CYP1B1-AS1*, and *SLC7A11*); the most significant *cis*-eQTM relationships per gene are listed in Table E4. Figures 6A–6G illustrate the relationship between gene expression (counts per million) and DNA methylation status for the most significant CpG sites identified by eQTM analysis. Figure 6A indicates that increased methylation of cg19949948 correlates with decreased expression of *ALDH3A1* (FDR = 5.74×10^{-04}). Conversely, Figure 6G shows that increased methylation of the CpG site cg13779050 correlates with increased transcription of *SLC7A11* (FDR = 3.81×10^{-02}).

Mapping of Smoking Affected Genes in an Independent *In Vitro* Dataset

To explore a possible explanation for how DNA methylation sites may influence the expression of the identified genes, we

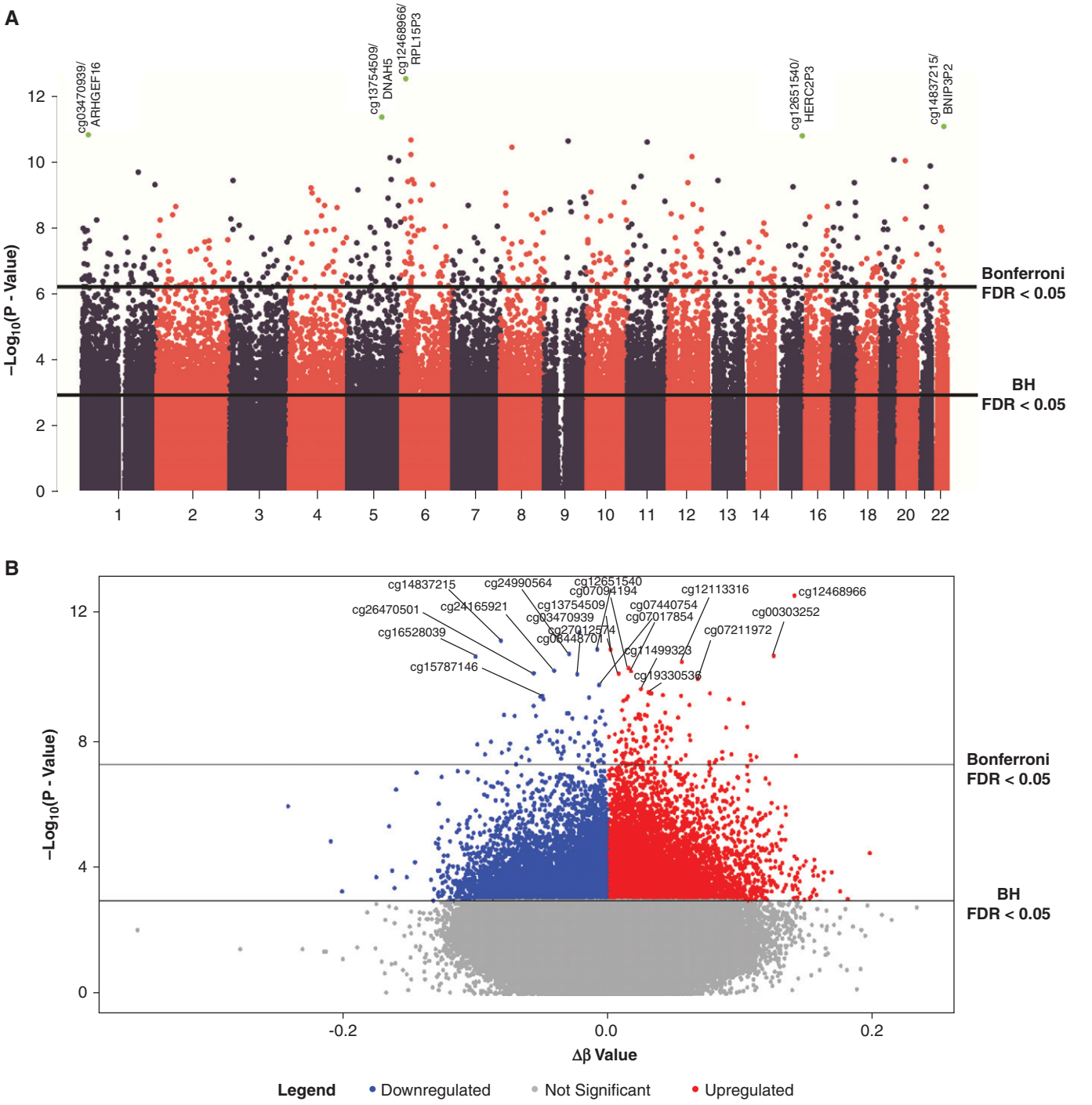

Figure 5. Manhattan and volcano plot visualization of lower and higher methylation of CpG sites in current smokers versus ex-smokers. (*A*) Manhattan plot showing the amount of differentially methylated sites on individual chromosomes, with the most significant sites highlighted and nearby genes indicated. (*B*) Volcano plot illustrating the CpG sites that are associated with either lower (blue) or higher (red) methylation and the $-\log_{10}$ (*P* value) plotted against the Δβ values for each CpG site. Genome-wide statistical significance is denoted in both figures by a solid black line at FDR-adjusted *P* value < 0.05 using the Bonferroni and BH method. *n* = 14 (current smokers) and 17 (ex-smokers) for all analyses. BH = Benjamin-Hochberg.

investigated publicly available ChIP-Seq datasets of well-known smoking-related transcription factors AHR, AHRR, and NRF2. With eQTM analysis, we identified three genes with a total of 17 CpG sites to be located in proximity to AHR and AHRR binding sites. No results were found for NRF2. Supplementary Figure E6 maps the transcript location for *CYP1B1*, *CYP1B1-AS1*, and *AHRR*, with AHR and AHRR transcription factor binding sites indicated in dark and light blue upstream of the gene. Multiple significant

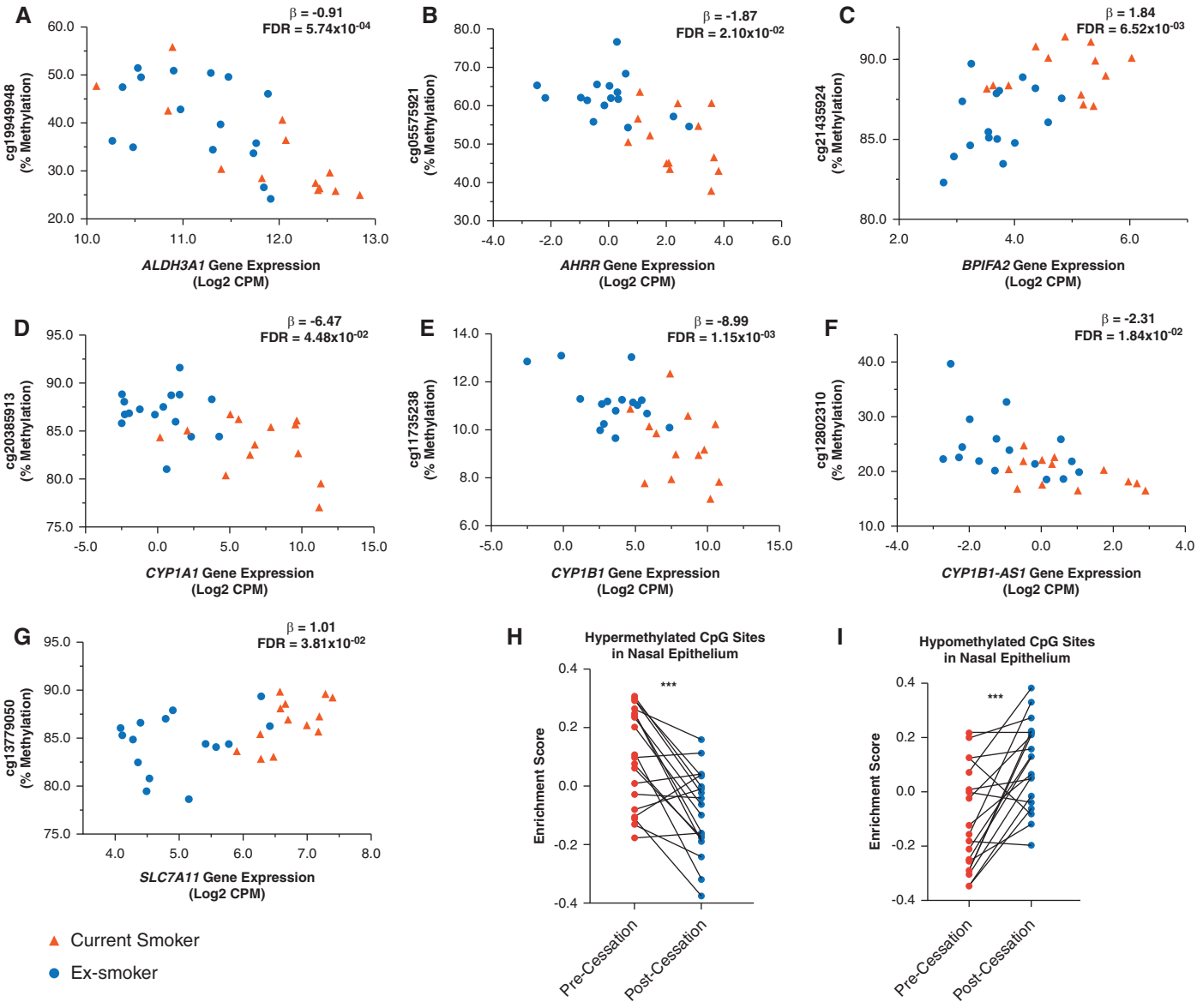

Figure 6. Representation of the relationship between DNA methylation status and gene expression in ex-smoker (blue) and current smoker (red) patients. (*A–G*) Percentage methylation values are reported on the *y*-axis and normalized gene expression is Log2 count per million (CPM) reported on the *x*-axis. The β-value indicates the slope of the correlation. All results returned an FDR <0.05. $n = 13$ (current smokers) and 16 (ex-smokers) for all analyses. (*H* and *I*) Gene set variation analysis data are represented as mean ± SEM and analyzed using a Student's paired parametric *t* test. Each dot represents an individual patient. Statistical significance is indicated by ***$P < 0.001$. $n = 19$. BPIFA2 = BPI fold-containing family A member 2; CYP1B1-AS1 = CYP1B1 antisense RNA 1; SLC7A11 = solute carrier family 7 member 11.

CpG sites identified by eQTM analysis were shown to be located within these sites, potentially influencing the binding of the transcription factors and therefore gene transcription.

We conducted a bootstrap analysis to increase the confidence that the identified list of CpG sites did not occur by chance. Across 1,000 permutations, an average of 1.23 CpG sites were identified with a 95% confidence interval of 1.15–1.31 CpG sites. Therefore, the total number of 17 CpG sites found in the original analysis was significant, with a *P* value of less than 0.0001. A frequency histogram has been included in the supplementary material to illustrate the results of the bootstrap analysis (Figure E7).

DNA Methylation of Specific CpG Sites Reverses with Smoking Cessation

Next, to investigate the long-term effect of smoking cessation on DNA methylation profiles and address the relevance for the lower airways, we used a bronchial biopsy dataset that followed current smokers for 1 year after smoking cessation (27). A total of 25 *cis*-eQTM CpG sites in the nasal epithelium were significantly altered in the reverse direction in bronchial biopsies after 1 year of smoking cessation (Table 3). GSVA of all differentially methylated CpG sites in the nasal epithelium indicated an overall reversal of DNA methylation status in the biopsies. Sites more

Table 3. List of *cis*-eQTM CpG Sites in Nasal Epithelium Altered after 1 Year of Smoking Cessation

		Nasal Epithelium Current Smoking			Bronchial Biopsy 1 Year after Smoke Cessation		
CpG Site	Gene	Log FC	FDR	Change in DNA Methylation	Log FC	FDR	Change in DNA Methylation
cg11183875	AKR1B10	−0.951	3.58×10^{-04}	Decreased	0.341	2.55×10^{-02}	Increased
cg04530175	AKR1B10	−1.112	4.52×10^{-02}	Decreased	0.458	2.23×10^{-02}	Increased
cg27638168	ALDH3A1	−0.694	3.15×10^{-05}	Decreased	0.337	1.25×10^{-02}	Increased
cg01375106	ALDH3A1	−0.375	4.15×10^{-02}	Decreased	0.425	1.20×10^{-02}	Increased
cg12828741	ALDH3A1	−0.518	2.30×10^{-02}	Decreased	0.421	1.19×10^{-02}	Increased
cg24752487	BPIFA2	−0.732	1.32×10^{-02}	Decreased	0.424	4.16×10^{-02}	Increased
cg02790305	BPIFA2	−0.720	4.72×10^{-02}	Decreased	0.264	3.73×10^{-02}	Increased
cg02162897	COL9A2	−1.701	1.36×10^{-03}	Decreased	0.489	3.06×10^{-02}	Increased
cg08761909	COL9A2	−1.086	8.77×10^{-03}	Decreased	0.500	6.81×10^{-03}	Increased
cg04968473	CYP1B1	0.565	1.99×10^{-02}	Increased	−0.269	2.89×10^{-02}	Decreased
cg01838114	CYP1B1	−0.887	8.92×10^{-05}	Decreased	0.274	4.81×10^{-02}	Increased
cg20408276	CYP1B1	−1.390	5.47×10^{-03}	Decreased	0.432	2.65×10^{-02}	Increased
cg26144569	CYP1B1	−0.663	4.73×10^{-03}	Decreased	0.384	4.93×10^{-02}	Increased
cg24577031	CYP1B1	0.435	3.97×10^{-02}	Increased	−0.308	1.17×10^{-02}	Decreased
cg12099669	CYP1B1-AS1	1.095	1.54×10^{-02}	Increased	−0.543	2.20×10^{-02}	Decreased
cg00305453	LHX6	−0.657	1.85×10^{-02}	Decreased	0.274	4.62×10^{-02}	Increased
cg13921444	MTNR1A	−0.734	2.09×10^{-02}	Decreased	0.333	2.19×10^{-02}	Increased
cg20385913	MTNR1A	−0.507	1.15×10^{-03}	Decreased	0.442	2.96×10^{-02}	Increased
cg13127741	SAA2	−0.507	2.70×10^{-03}	Decreased	0.246	4.87×10^{-02}	Increased
cg03296424	SAA4	−0.672	1.78×10^{-04}	Decreased	0.319	2.12×10^{-02}	Increased
cg06589222	SNTG2	−0.648	1.88×10^{-02}	Decreased	0.243	1.98×10^{-02}	Increased
cg07921828	STC1	−1.105	3.56×10^{-02}	Decreased	0.524	2.61×10^{-02}	Increased
cg26267490	STC1	0.812	2.72×10^{-02}	Increased	−0.281	3.75×10^{-02}	Decreased
cg03041510	SUSD2	0.398	1.05×10^{-02}	Increased	−0.463	1.56×10^{-02}	Decreased
cg05575921	TRHDE	−0.719	2.29×10^{-02}	Decreased	0.341	1.90×10^{-02}	Increased

Definition of abbreviations: *cis*-eQTM = *cis*-expression quantitative trait methylation; FDR = false discovery rate (Benjamini-Hochberg method); LogFC = log-fold change.

highly methylated in current smokers become more lowly methylated with 1 year of smoking cessation, and the opposite effect occurs for lowly methylated sites (Figures 6H and 6I).

Discussion

The genes that most consistently associated with current smoking in patients with asthma include *ALDH3A1*, *CYP1A1*, *CYP1B1*, *CYP1B1-AS1*, and *AHRR*. These genes were identified in the nasal epithelium and validated in controlled *in vitro* smoke exposure using ALI and *in vivo* airway cell models, demonstrating immediate changes in gene expression upon smoke exposure. We observe that genes that are increased by smoke exposure are downregulated with smoking cessation in upper and lower airways. Gene expression increased by current smoking correlate with peripheral blood immune cell counts, which may suggest a relationship between these genes and the innate immune response. Differentially methylated CpG sites were located in the proximity of key transcription factor binding sites, possibly affecting the expression of the nearby genes. We also demonstrate changes in DNA methylation in the nasal epithelium are reflected in bronchi, highlighting a "field of injury" throughout the airways.

Genes significantly upregulated by current smoking in nasal epithelium of patients with asthma include *ALDH3A1*, *CYP1A1*, *CYP1B1,* and *AHRR*. *CYP1A1/B1* catalyze the breakdown of drug components producing reactive metabolites that can cause DNA damage (11, 19). Hence, prolonged expression of these genes increases the susceptibility of cancer development (29). *ALDH3A1* increases expression in response to aldehydes in smoke and other teratogenic components (30). Prolonged exposure to tobacco smoke products causes increased expression of these genes. This is reinforced by changes to DNA methylation, possibly at transcription factor binding sites, potentially influencing gene expression. The majority of the CpG sites indicated lower methylation levels in the nasal epithelium of current smokers, thus indicating that smoking may demethylate transcription factor binding sites, facilitating gene expression. Transcription factor binding is known to be impacted by DNA methylation; however, the findings presented in our study require experimental validation to confirm this interaction. DNA methylation at AHRR and AHR binding sites need to be investigated to completely understand the dynamics between methylation and the ability of transcription factors to initiate gene transcription. AHRR has been linked to cigarette smoke and immune response in various tissues (1, 31–33). AHRR inhibits the transcription factor AHR, which regulates the expression of *CYP1A1* and *CYP1B1*. Dysregulated gene expression may prolong inflammation, leading to exacerbations in diseases such as asthma and COPD (34). *CYP1B1-AS1* is an antisense RNA that showed increased expression because of current smoking, indicating an effect of tobacco smoke on the regulation of protein translation.

Concurrent investigation of DNA methylation status and gene expression demonstrated a correlation between smoking status and methylation of CpG sites in proximity to smoking-related genes. Tobacco smoke has a heterogeneous effect on DNA

methylation status, potentially increasing or decreasing gene expression depending on the location of the site within a gene body (35). A similar investigation in small airway epithelium found changes in the methylation status of CpG sites associated with *ALDH3A1*, *CYP1A1*, and *CYP1B1* (18). The lower methylation of CpG site cg05575921 is observed to downregulate expression of *AHRR* in various tissues and cell types (19, 31, 36–38). Analysis of differential DNA methylation data in conjunction with independent ChIP-Seq data of AHR and AHRR illustrates the proximity of smoking-related DNA methylation sites to gene promoter regions. This reinforces the significance of these transcription factors as key regulators of the physiological response to tobacco smoke.

A correlation between the GSVA smoking score in the nose and the immune response is observed through increased peripheral blood neutrophil, lymphocyte, and monocyte absolute counts, whereas no association with peripheral eosinophil numbers was observed. Our findings agree with those of previous studies showing that current smoking increases peripheral blood neutrophils and that the nasal epithelium can reflect smoking-induced neutrophilic inflammation (39). Therefore, altered nasal gene expression and DNA methylation due to current smoking might relate to increased immune cell count and worsening asthma symptoms.

We identify a correlation between smoking-induced changes in DNA methylation and gene expression in nasal brushed cells. Investigating an independent *in vivo* smoking dataset demonstrated a reversal of smoking-associated upregulated genes within 4 weeks of cessation. A reversal in the expression of smoking affected genes is sustained for 6 months with continued smoking cessation. In bronchial samples, we observed a reversal of DNA methylation status after 1 year of smoking cessation. We observed no strong cell-type contribution through cellular deconvolution. Surprisingly, a lack of goblet cell hyperplasia, which has been associated with smoking, was reported. This is possibly due to the maximal collection of these cells, reducing the sensitivity of the comparison. A consistent pattern and alteration of DNA methylation from the nose to bronchus reflects the concept of a "field of injury" throughout the respiratory tract caused by tobacco smoke exposure. Studies in adipose tissue support this, with cessation causing some reversal in methylation of CpG sites (19). Our analysis involved comparing nasal with bronchial epithelium CpG sites in different patient cohorts. Therefore, further investigation into the dynamics of changes in DNA methylation and the effect on gene expression within the same tissue in response to smoke is necessary.

A strength of our investigation is the ability to analyze both gene expression and DNA methylation in nasal epithelial samples from the same patients. This enables a direct study of the correlation between expression and DNA methylation status. However, there were a number of limitations in the current study. Our investigation compares data collected from varying cell types, limiting the ability to interpret relationships and correlations and to determine conclusions. However, these analyses, in combination with the primary analysis of paired gene and DNA methylation samples, offer valuable insight into potential mechanisms driving the biological response to tobacco smoke. We acknowledge the relatively low number of patients in our study. Therefore, follow-up studies are required using a larger cohort, and the inclusion of a third never-smoker group in such a study would also be advisable, as it offers a more holistic understanding of the effect of tobacco smoke on the studied parameters.

This study confirms that current cigarette smoking alters epigenetic patterns and affects gene expression in nasal brushed cells of individuals with asthma. Current smoking affects the genes *ALDH3A1*, *CYP1A1*, *CYP1B1*, and *AHRR*, which are involved in physiological responses in detoxification and oxidative stress. Moreover, differentially methylated CpG sites due to current smoking demonstrate a relationship with the expression of these identified genes. Smoking cessation for 6 months indicates a reversal in gene expression pattern caused by smoking, with a similar pattern seen for DNA methylation levels after 1 year of cessation. Hence, a complex and dynamic relationship between DNA methylation and gene expression exists. Our findings contribute to the application and understanding of nasal gene expressions, such as the identification of appropriate biomarkers for disease prognosis and the elucidation of mechanisms contributing to pathological processes in smoking patients with asthma. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank the patients who participated in the study. The authors also thank H. Aliee, F. J. Theis, and M. C. Nawjin of the National Heart, Lung, and Blood Institute LungMAP Consortium for their contribution to the cellular deconvolution analysis. The authors also thank the anonymous reviewers of this manuscript, as their suggestions and comments elevated the quality of the work.

References

1. Kodal JB, Kobylecki CJ, Vedel-Krogh S, Nordestgaard BG, Bojesen SE. *AHRR* hypomethylation, lung function, lung function decline and respiratory symptoms. *Eur Respir J* 2018;51:1701512.
2. Jiménez-Ruiz CA, Andreas S, Lewis KE, Tonnesen P, van Schayck CP, Hajek P, *et al.* Statement on smoking cessation in COPD and other pulmonary diseases and in smokers with comorbidities who find it difficult to quit. *Eur Respir J* 2015;46:61–79.
3. Global Initiative for Asthma. Global strategy for asthma management and prevention online appendix 2020. 2020 [accessed 2021 Aug 9]. Available from: https://webmed.irkutsk.ru/doc/pdf/ginareport.pdf.
4. Bakakos P, Kostikas K, Loukides S. Smoking asthma phenotype: diagnostic and management challenges. *Curr Opin Pulm Med* 2016;22:53–58.
5. Dijkstra A, Vonk JM, Jongepier H, Koppelman GH, Schouten JP, ten Hacken NH, *et al.* Lung function decline in asthma: association with inhaled corticosteroids, smoking and sex. *Thorax* 2006;61:105–110.
6. Billatos E, Faiz A, Gesthalter Y, LeClerc A, Alekseyev YO, Xiao X, *et al.* Impact of acute exposure to cigarette smoke on airway gene expression. *Physiol Genomics* 2018;50:705–713.
7. Imkamp K, Berg M, Vermeulen CJ, Heijink IH, Guryev V, Kerstjens HA, *et al.* Nasal epithelium as a proxy for bronchial epithelium for smoking-induced gene expression and expression quantitative trait loci. *J Allergy Clin Immunol* 2018;142:314–317.e315.
8. Zhang X, Sebastiani P, Liu G, Schembri F, Zhang X, Dumas YM, *et al.* Similarities and differences between smoking-related gene expression in nasal and bronchial epithelium. *Physiol Genomics* 2010;41:1–8.
9. Sridhar S, Schembri F, Zeskind J, Shah V, Gustafson AM, Steiling K, *et al.* Smoking-induced gene expression changes in the bronchial airway are reflected in nasal and buccal epithelium. *BMC Genomics* 2008;9:259.
10. Poole A, Urbanek C, Eng C, Schageman J, Jacobson S, O'Connor BP, *et al.* Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. *J Allergy Clin Immunol* 2014;133:670–678.e612.

11. Tekpli X, Zienolddiny S, Skaug V, Stangeland L, Haugen A, Mollerup S. DNA methylation of the CYP1A1 enhancer is associated with smoking-induced genetic alterations in human lung. *Int J Cancer* 2012;131:1509–1516.
12. Zakarya R, Adcock I, Oliver BG. Epigenetic impacts of maternal tobacco and e-vapour exposure on the offspring lung. *Clin Epigenetics* 2019;11:32.
13. Lee KW, Pausova Z. Cigarette smoking and DNA methylation. *Front Genet* 2013;4:132.
14. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 2011;12:529–541.
15. Qi C, Xu CJ, Koppelman GH. The role of epigenetics in the development of childhood asthma. *Expert Rev Clin Immunol* 2019;15:1287–1302.
16. Guida F, Sandanger TM, Castagné R, Campanella G, Polidoro S, Palli D, *et al.* Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. *Hum Mol Genet* 2015;24:2349–2359.
17. Ringh MV, Hagemann-Jensen M, Needhamsen M, Kular L, Breeze CE, Sjöholm LK, *et al.* Tobacco smoking induces changes in true DNA methylation, hydroxymethylation and gene expression in bronchoalveolar lavage cells. *EBioMedicine* 2019;46:290–304.
18. Buro-Auriemma LJ, Salit J, Hackett NR, Walters MS, Strulovici-Barel Y, Staudt MR, *et al.* Cigarette smoking induces small airway epithelial epigenetic changes with corresponding modulation of gene expression. *Hum Mol Genet* 2013;22:4726–4738.
19. Tsai PC, Glastonbury CA, Eliot MN, Bollepalli S, Yet I, Castillo-Fernandez JE, *et al.* Smoking induces coordinated DNA methylation and gene expression changes in adipose tissue with consequences for metabolic health. *Clin Epigenetics* 2018;10:126.
20. Cox CA, Boudewijn IM, Vroegop SJ, Schokker S, Lexmond AJ, Frijlink HW, *et al.* Extrafine compared to non-extrafine particle inhaled corticosteroids in smokers and ex-smokers with asthma. *Respir Med* 2017;130:35–42.
21. Moses E, Wang T, Corbett S, Jackson GR, Drizik E, Perdomo C, *et al.* Molecular impact of electronic cigarette aerosol exposure in human bronchial epithelium. *Toxicol Sci* 2017;155:248–257.
22. Hijazi K, Malyszko B, Steiling K, Xiao X, Liu G, Alekseyev YO, *et al.* Tobacco-related alterations in airway gene expression are rapidly reversed within weeks following smoking-cessation. *Sci Rep* 2019;9:6978.
23. Procházková J, Strapáčová S, Svržková L, Andrysík Z, Hýžd'alová M, Hrubá E, *et al.* Adaptive changes in global gene expression profile of lung carcinoma A549 cells acutely exposed to distinct types of AhR ligands. *Toxicol Lett* 2018;292:162–174.
24. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, *et al.* A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol* 2015;16:37.
25. Yang SY, Ahmed S, Satheesh SV, Matthews J. Genome-wide mapping and analysis of aryl hydrocarbon receptor (AHR)- and aryl hydrocarbon receptor repressor (AHRR)-binding sites in human breast cancer cells. *Arch Toxicol* 2018;92:225–240.
26. Wang X, Campbell MR, Lacher SE, Cho H-Y, Wan M, Crowl CL, *et al.* A polymorphic antioxidant response element links nrf2/smaf binding to enhanced mapt expression and reduced risk of parkinsonian disorders. *Cell Rep* 2016;15:830–842.
27. Willemse BW, ten Hacken NH, Rutgers B, Lesman-Leegte IG, Postma DS, Timens W. Effect of 1-year smoking cessation on airway inflammation in COPD and asymptomatic smokers. *Eur Respir J* 2005;26:835–845.
28. Houseman EA, Kim S, Kelsey KT, Wiencke JK. DNA methylation in whole blood: Uses and challenges. *Curr Environ Health Rep* 2015;2:145–154.
29. Port JL, Yamaguchi K, Du B, De Lorenzo M, Chang M, Heerdt PM, *et al.* Tobacco smoke induces CYP1B1 in the aerodigestive tract. *Carcinogenesis* 2004;25:2275–2281.
30. Jang JH, Bruse S, Liu Y, Duffy V, Zhang C, Oyamada N, *et al.* Aldehyde dehydrogenase 3A1 protects airway epithelial cells from cigarette smoke-induced DNA damage and cytotoxicity. *Free Radic Biol Med* 2014;68: 80–86.
31. Zeilinger S, Kühnel 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.
32. Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenetics* 2015;7:113.
33. Arimilli S, Madahian B, Chen P, Marano K, Prasad GL. Gene expression profiles associated with cigarette smoking and moist snuff consumption. *BMC Genomics* 2017;18:156.
34. Stringer KA, Freed BM, Dunn JS, Sayers S, Gustafson DL, Flores SC. Particulate phase cigarette smoke increases MnSOD, NQO1, and CINC-1 in rat lungs. *Free Radic Biol Med* 2004;37:1527–1533.
35. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012;13:484–492.
36. 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.
37. Bakulski KM, Dou J, Lin N, London SJ, Colacino JA. DNA methylation signature of smoking in lung cancer is enriched for exposure signatures in newborn and adult blood. *Sci Rep* 2019;9:4576.
38. Bojesen SE, Timpson N, Relton C, Davey Smith G, Nordestgaard BG. *AHRR* (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax* 2017;72:646–653.
39. Vachier I, Vignola AM, Chiappara G, Bruno A, Meziane H, Godard P, *et al.* Inflammatory features of nasal mucosa in smokers with and without COPD. *Thorax* 2004;59:303–307.