Food and Chemical Toxicology 123 (2019) 501–510

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

Tobacco Heating System 2.2 has a limited impact on DNA methylation of candidate enhancers in mouse lung compared with cigarette smoke

Mohamed-Amin Choukrallah, Nicolas Sierro, Florian Martin, Karine Baumer, Jerome Thomas, Sonia Ouadi, Julia Hoeng*, Manuel C. Peitsch, Nikolai V. Ivanov

PMI R&D, Philip Morris Products S.A., Quai Jeanrenaud 5, CH-2000, Neuchâtel, Switzerland

ARTICLE INFO

Keywords:
Cigarette smoke
DNA methylation
Epigenetics
Modified risk tobacco products
Enhancers
Gene regulation

ABSTRACT

Cigarette smoke (CS) exposure has been shown to correlate with changes in DNA methylation levels, however, the impact of CS on DNA methylation at genome-wide scale is missing. Here, we used whole-genome bisulfite sequencing to assess the effects of CS extract and aerosol from the Tobacco Heating System (THS) 2.2, a candidate modified risk tobacco product, on DNA methylation in lung and liver tissues from apolipoprotein E-deficient mice during an eight-month period of exposure. We found that in lung tissue, CS mainly induced hypermethylation of candidate enhancers at late time points, while promoters were less affected. This effect was strongly reduced upon cessation or switching to THS 2.2. By contrast, chronic exposure to THS 2.2 had a limited effect on DNA methylation at both promoters and enhancers. We also identified members of the Ets and Fox families of transcription factors as potential players in the epigenetic response to CS exposure in lung tissue. In contrast to the lung, DNA methylation in the liver was largely insensitive to all investigated exposures.

In summary, our investigations indicate that CS-related DNA methylation alterations are tissue-specific, occur mainly at enhancers and are strongly reduced upon smoking cessation or switching to THS2.2.

1. Introduction

Cigarette smoke (CS) is a major long-term risk factor for various human disorders, including disorders of the respiratory (Doll and Hill, 1950) and cardiovascular (Roth, 1954) systems (Chu et al., 2013). The effects of CS are thought to be partly mediated by genetic and epigenetic changes (Zhang et al., 2016; Rogers et al., 2017). DNA methylation, 5-methylcytosine (5 mC), is the most studied epigenetic mechanism in relation to CS exposure. 5 mC occurs almost exclusively in the context of CpG dinucleotides and is considered as a repressive mechanism incompatible with active transcription. 5 mC is catalyzed by three DNA methyltransferases: DNMT1, DNMT3a, and DNMT3b (Bestor, 2000). DNMT1 is a maintenance enzyme that ensures the inheritance of 5 mC patterns during DNA replication, while DNMT3a and DNMT3b are de novo DNA methyltransferases. Cytosine methylation is dynamic and reversible: 5 mC can be converted to 5-hydroxymethylcytosine during oxidative reaction catalyzed by ten-eleven translocation (Tet) enzymes (He et al., 2011).

Previous studies have reported that CS induced modest changes of DNA methylation mainly in blood cells (Shenker et al., 2013; Joubert et al., 2012; Breitling et al., 2011; Joehanes et al., 2016). However, the majority of these studies assessed DNA methylation using array-based technologies that are limited mainly to annotated loci and poorly cover the complex network of distal regulatory elements, such as enhancers, which can be located hundreds of kilobases (kb) away from their target genes.

DNA methylation alteration is often reported at the single CpG level without considering the genomic context. This approach may lead to a misinterpretation of DNA methylation readout. Although DNA methylation was historically associated with gene silencing, recent genome-scale investigations demonstrated clearly that DNA methylation readout varies depending on the genomic context. High levels of DNA methylation at promoters and enhancers are associated with transcriptional silencing, while within the gene body, it correlates positively with transcriptional activity (Lister et al., 2009; Laurent et al., 2010). Importantly, advanced quantitative analysis of whole-genome bisulfite sequencing (WGBS) has shown that the methylome can be segmented into three distinct classes: fully methylated regions (FMR), unmethylated regions (UMR) and low-methylated regions (LMR). FMRs represent 90% of the genome, and UMRs correspond to the majority of CpG islands and active promoters, while LMRs exhibit enhancer features, such as specific histone marks and binding of transcription factors (TFs) (Stadler et al., 2011a). Furthermore, the WGBS signal itself has

* Corresponding author.
E-mail address: Julia.Hoeng@pmi.com (J. Hoeng).

https://doi.org/10.1016/j.fct.2018.11.020
Received 10 August 2018; Received in revised form 19 October 2018; Accepted 7 November 2018
Available online 09 November 2018

Abbreviations

5 mC	5-methylcytosine
ApoE	Apolipoprotein E
bp	Base pair
ChIP-Seq	Chromatin immunoprecipitation sequencing
CRE	Cis-regulatory element
CS	Cigarette smoke
dsDNA	DNMT DNA Methyltransferase
dsDNA	Double-strand DNA
ETS	E26 transformation-specific family of transcription factors
FDR	False discovery rate
FMR	Fully methylated regions
FOX	Forkhead box family of transcription factors
H3K27ac	Acetylated lysine 27 on histone H3
H3K4me1	Monomethylated lysine 4 on histone H3
H3K4me3	Trimethylated lysine 4 on histone H3
kb	Kilobase
LMR	Low methylated region
MRTP	Modified Risk Tobacco Product
Tet	Ten-eleven translocation
TF	Transcription factor
THS 2.2	Tobacco Heating System 2.2
UMR	Unmethylated region
WGBS	Whole genome bisulfite sequencing

been used to identify candidate regulatory elements by detecting LMRs and UMRs (Burger et al., 2013).

Gene expression patterns are controlled directly by DNA cis-regulatory elements (CRE). Promoters and enhancers are the two major types of CREs in eukaryotes. While promoter features are well characterized, distal regulatory elements, such as enhancers, are cell type-specific, and their activity states are highly dynamic and depend strongly on the physiological conditions (Heintzman et al., 2007, 2009; Creyghton et al., 2010; Bulger and Groudine, 2011; Shen et al., 2012). Therefore, building a universal set of enhancers for all cell types and environments remains challenging. It is also known that enhancers are much more dynamic than promoters and may reflect genomic changes that are not observed in annotated loci or at the level of gene expression (Choukrallah et al., 2015).

We have previously shown that in the apolipoprotein E-deficient ($ApoE^{-/-}$) mouse model cessation or switching to the aerosol from a modified risk tobacco product (MRTP), the Tobacco Heating System (THS) 2.2, retarded the progression of atherosclerotic and emphysematous changes associated with CS exposure and that exposure to THS 2.2 aerosol alone had no obvious adverse effects (Phillips et al., 2016). Furthermore, in contrast to CS, THS 2.2 had no significant effect on gene expression profiles in the lung (Phillips et al., 2016). In contrast to lung tissue, both CS and THS 2.2 aerosols had a weak effect on gene expression in liver tissue (Lo Sasso et al., 2016). Although these investigations provided important insights about CS effects on various biological processes, the impact of CS and THS 2.2 on the epigenome remained to be elucidated.

Here, we assessed the effect of the conventional CS (from a 3R4F reference cigarette) and the aerosol from THS 2.2 on DNA methylation in the lung and liver using WGBS. The samples used in the current report are from the same $ApoE^{-/-}$ mice that were used in our previous reports (Phillips et al., 2016; Lo Sasso et al., 2016) to investigate other biological endpoints.

DNA methylation was quantified at different genomic targets, including annotated promoters, gene bodies, and candidate enhancers identified as LMRs. Our results show that CS induced mainly hypermethylation at hundreds of candidate lung enhancers. The number of affected loci increased with prolonged exposure times, while hypermethylation was reduced by cessation or switching to THS 2.2. In contrast to CS, chronic exposure to THS 2.2 altered DNA methylation at very few loci. Surprisingly, we found that both CS and THS 2.2 exposure had a limited effect on promoter methylation. In contrast to lung tissue, the effect of CS and the aerosol from THS 2.2 affected the methylation of a limited number of enhancers and promoters in the liver.

2. Material and methods

2.1. Study design and mice experimental conditions

The study design and experimental conditions were identical to those in our previously published study that reported on other biological endpoints for the same exposures and from the same animals (Phillips et al., 2016). Female $ApoE^{-/-}$ mice were randomized into five groups (Fig. 1A): (1) Sham (exposed to air), (2) 3R4F (exposed to CS from the 3R4F reference cigarette), (3) THS 2.2 (exposed to mainstream aerosol from THS 2.2 at nicotine levels matched to those of 3R4F CS), (4) smoking cessation, and (5) switching to THS 2.2 aerosol. To model the effects of smoking cessation and switching to THS 2.2 aerosol, mice from the cessation and switch groups were first exposed to 3R4F for two months and then switched to air or THS 2.2 aerosol, respectively, for up to six additional months (Fig. 1A).

The mice were whole body-exposed to diluted mainstream smoke from 3R4F cigarettes (600 mg TPM/m^3, equivalent to 29.9 mg nicotine/m^3), THS 2.2 aerosol (nicotine-matched to 3R4F, 29.9 mg/m^3), or filtered air for 3 h per day, five days per week, for up to eight months. Intermittent daily exposure to fresh filtered air for 30 min after the first hour of smoke exposure and for 60 min after the second hour of exposure (Fig. 1B) was provided to avoid a buildup of excessive carboxyhemoglobin concentrations in the 3R4F group. For the Sham group, mice were exposed to air.

3R4F reference cigarettes were purchased from the University of Kentucky (http://www.ca.uky.edu/refcig). The candidate MRTP, THS 2.2, heats a specially designed tobacco product with an electronically controlled heating blade that does not exceed the maximum operating

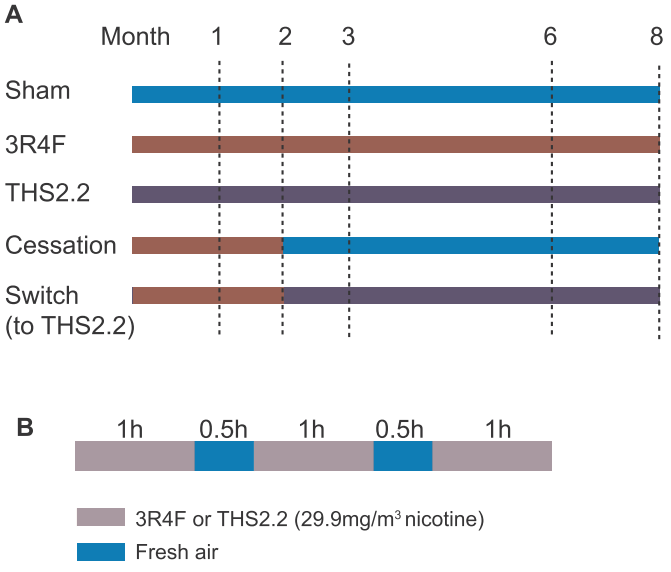

Fig. 1. Study design.
A. Experimental groups and time points are indicated. Mice were exposed to CS (3R4F), aerosols from THS 2.2 (THS 2.2), or fresh air (Sham = control). Cessation or switching to THS 2.2 were also assessed. B. Daily exposure schedule.

temperature of 350 °C.

2.2. DNA extraction, bisulfite conversion, and sequencing

Genomic DNA was extracted manually from lung and liver slices without grinding using a QiaAMP tissue mini kit (QIAGEN, Hilden, Germany) and resuspended in 50 µl buffer. The concentration of the extracted DNA was checked using a Qubit (Thermo Fisher Scientific, Ecublens, Switzerland). After isolation, 100 ng of genomic DNA and 0.5 ng of unmethylated lambda DNA (Promega, Madison, USA) were resuspended in 52.5 µl TE buffer. The DNA was sheared using a Covaris E220 sonicator (Covaris Inc., Brighton, UK) to obtain double-strand DNA (dsDNA) fragments of about 200 base pairs (bp). After ligation of indexed sequencing adapters from the Ovation Ultralow Methyl-Seq Library System (NuGEN Technologies, Inc., San Carlos, CA, USA), bisulfite conversion of the dsDNA was carried out using the EpiTect Bisulfite Kit (QIAGEN). The converted dsDNA was recovered using the MinElute PCR purification Kit (QIAGEN), amplified by PCR, and cleaned on Agencourt RNAClean XP beads (Beckman Coulter, Nyon, Switzerland). The concentrations and sizes of the sequencing libraries were verified on a Bioanalyzer (Agilent, Santa Clara, CA, USA). Normalized libraries were pooled together in multiplexes of six to 14 libraries and clustered on Illumina PE high-output flow cell v3 using an Illumina cBot (Illumina Inc., San Diego, California, USA) and Illumina TruSeq PE Cluster Kit v3 – cBot – HS kits. Sequencing of the flow cells was performed on an Illumina HiSeq 2500 (Illumina Inc.) using Illumina TruSeq SBS v3 – HS kits 200 cycles and Illumina TruSeq PE Cluster Kit v3 – cBot – HS in a 2 × 101 paired-end mode.

2.3. Data processing

Sequencing reads were aligned to the mouse genome (version mm10) using the qAlign function in the QuasR package (v 1.10.0) (Gaidatzis et al., 2015), which internally uses bowtie (v 1.10.0) with alignment parameters fitting directional bisulfite-converted libraries. Methylation was quantified using the qMeth function from the QuasR package. Only cytosines in a CpG context were considered, and the counts were strand-combined. Counts were summed per region, and methylation levels corresponding to the ratio between methylated and total events were presented as a 0 to 1 range, with 0 being a fully unmethylated state (methylated counts = 0) and 1 being a fully methylated state (methylated counts = total counts). The significance of

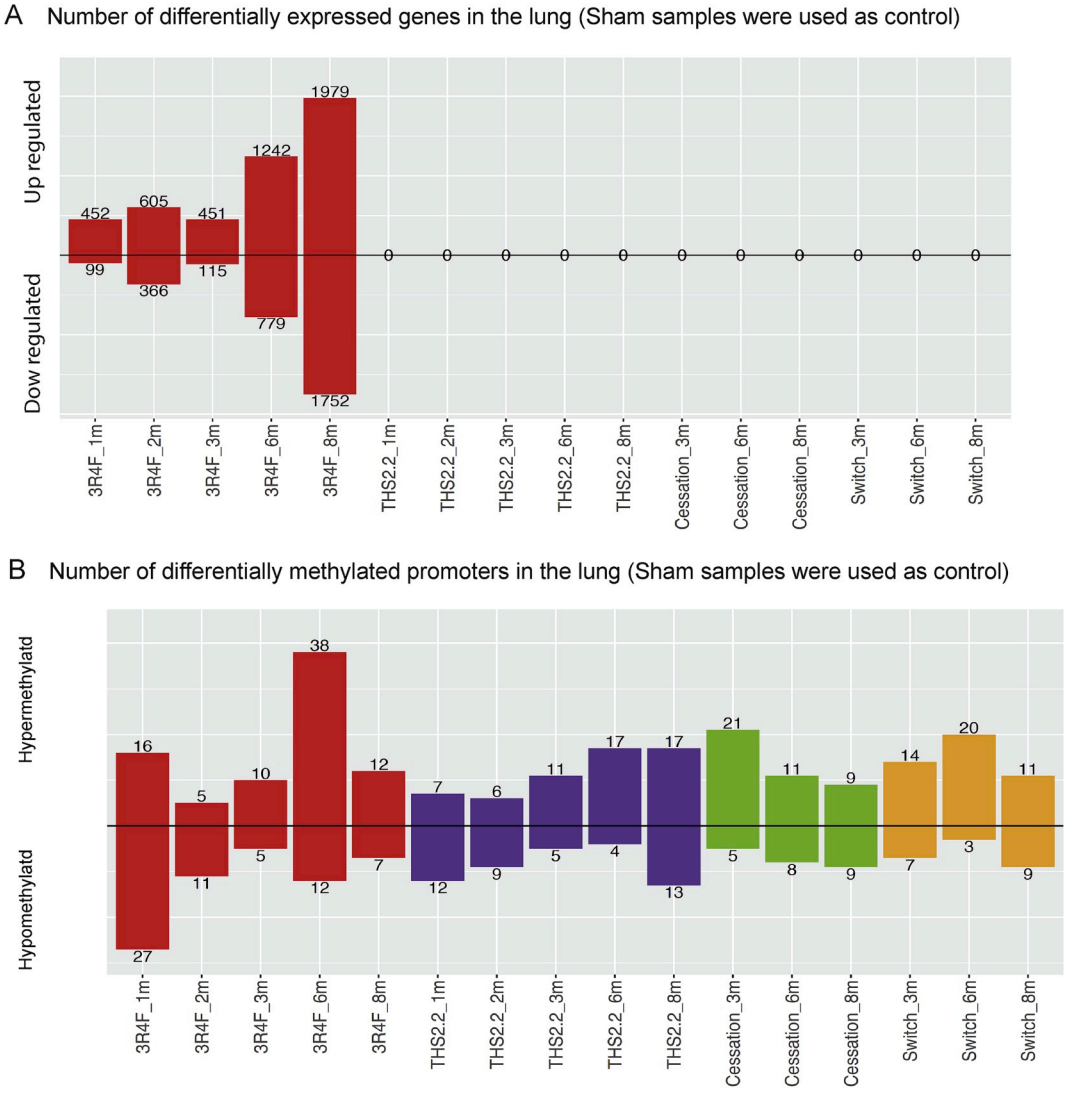

Fig. 2. Assessment of DNA methylation at promoters and comparison with gene expression profiles in the lung.
A. Barplot representing the number of differentially expressed genes based on the FDR cut-off (0.05) between the indicated treatments (*x*-axis) and the corresponding Sham samples from the respective time points (data from (Phillips et al., 2016)). B. Barplot representing the number of differentially methylated promoters (transcription start site [TSS] −/+ 500 bp) based on the FDR cut-off (0.05). Sham samples were used as controls.

differential methylation between the treatment groups and the respective Sham samples was assessed using the false discovery rate (FDR)-adjusted *p*-value from the betabinomial model implemented in the *betabin* function from the aod package (https://cran.r-project.org/web/packages/aod/index.html). The coupled values (number of methylated reads and number of total reads) per replicate were used as input as follows: betabin(formula = cbind(y, n - y) ~ 1 + group, random = ~1, data = data.frame (y, n, group), link = "logit"). The variables y and n respectively indicate the number of methylated reads and the total number of reads and the variable "group" indicates the experimental group. For every contrast (treatment versus control), only loci covered by at least 15 reads in all of the computed samples were considered. Promoters were defined as non-overlapping −500-bp to +500-bp intervals around transcription start sites from refseq data base

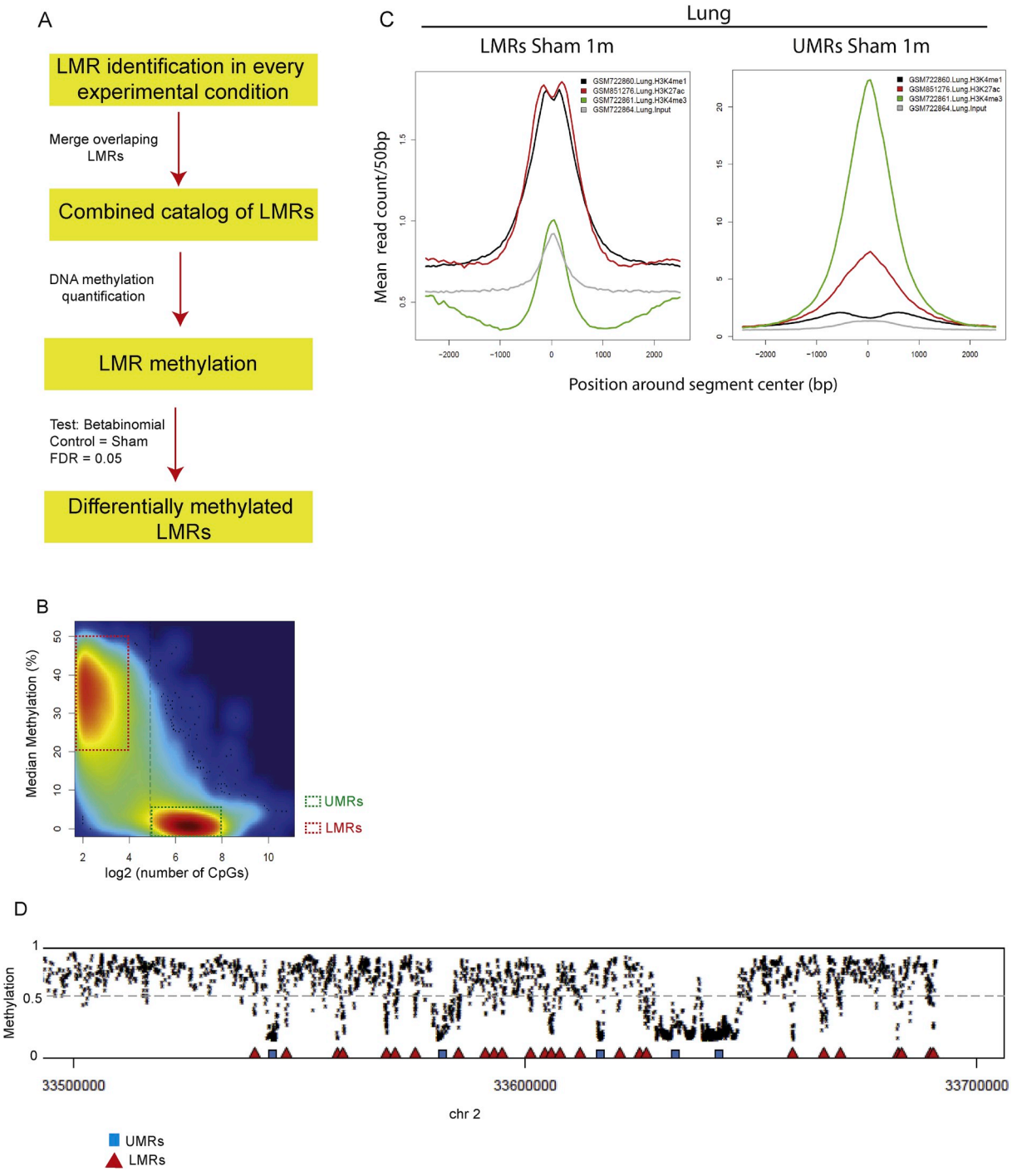

Fig. 3. Workflow for identification of LMRs and assessment of differential DNA methylation at these elements.
A. Workflow to detect differentially methylated LMRs between the treatments and the corresponding Sham samples. B. Smooth scatterplots representing the number of CpGs
(log_2) per region and the methylation median. Two populations of loci can be distinguished: unmethylated CpG-rich loci (mainly CpG islands) and low methylated CpG-poor loci (candidate enhancers). C. Chromatin profiles (GEO: GSE29184, lung samples from eight-week-old mice) at LMRs and UMRs detected in Sham samples at the one-month time point, which corresponds to 12- to 14-week-old mice. Enhancer marks are enriched at LMRs, and promoter marks are enriched at UMRs. D. Genome browser snapshot illustrating UMR and LMR profiles.

(https://www.ncbi.nlm.nih.gov/refseq/). Gene bodies were defined as transcripts with unique start and end coordinates from the same database.

2.3.1. LMR and UMR identification

LMRs and UMRs were identified using methylSeekR package with default parameters and without partially methylated domain filtering. LMRs were detected for every experimental group by merging the counts for the respective replicates. Then, the LMRs from all of the experimental groups were combined in one unified catalog for each tissue, whereby elements with more than 60% overlap were merged to reduce the redundancy. Methylation levels for all of the elements in the unified sets of LMRs were quantified separately for every replicate and used to compute differential methylation between the treatment samples and the respective Sham samples.

2.3.2. TF motif analysis

TF motif discovery was performed using the HOMER software (Version 4.8.3) (Heinz et al., 2010) with the following command: findMotifsGenome.pl input mm10 output_directory -nomotif, with the input being the different sets of LMRs.

2.3.3. ChIP-seq data analysis

Sra files were downloaded from the GEO DataSets website (https://www.ncbi.nlm.nih.gov/gds) and converted to fastq format using the fastq-dump function from the SRA-tools kit (https://www.ncbi.nlm.nih.gov/books/NBK158900/). Sequencing reads were mapped to the mouse genome (mm10) using the qAlign function in the QuasR package with default settings. Quantification of chromatin immunoprecipitation sequencing (ChIP-seq) signals at LMRs and UMRs was performed using QuasR qCount function. The following ChIP-Seq samples were used: Lung H3K4me1 (GSM722860), Lung H3K27ac(GSM851276), Lung H3K4me3 (GSM722861), Lung Input(GSM722864), Liver H3K4me1(GSM722760), Liver H3K27ac(GSM851275), Liver H3K4me3(GSM722761) and Liver Input(GSM722764).

3. Results

3.1. Generation of high-resolution methylomes

ApoE$^{-/-}$ transgenic mice were exposed to CS or the aerosol from THS 2.2 (nicotine concentration matched to CS: 29.9 μg/L) over an eight-month period. DNA methylation was assessed after one, two, three, six, and eight months of exposure to CS or aerosol from THS 2.2. Cessation or switching from CS to THS 2.2 were also assessed: after two months of CS exposure, one group of animals was switched to THS 2.2, and a second group underwent cessation (Fig. 1). Fresh air (Sham) was used as control. Every experimental group contained eight biological replicates corresponding to eight different animals, leading to a total of 168 methylomes for each tissue.

3.2. CS-dependent transcriptomic alteration in the lung is largely unrelated to DNA methylation at promoters and gene bodies

We have shown previously that CS alters the expression of thousands of genes in murine lung tissue (Fig. 2A) (Phillips et al., 2016). To evaluate the effect of the different treatments on DNA methylation and a possible link to the previously observed changes in gene expression, we first quantified DNA methylation at promoter regions, defined as 1 kb window flanking transcription start sites. DNA methylation levels correspond to the ratio between the number of reads overlapping with methylated events and the total number of reads. This ratio ranges from 0 to 1, where 0 indicates an unmethylated state, and 1 indicates a fully methylated state.

Regardless of the experimental groups, DNA methylation levels at promoter regions showed a bimodal distribution where highly methylated promoters were either not expressed or weakly expressed (Fig. S1A). Methylation levels at promoters in samples belonging to the 3R4F, THS 2.2, cessation, and switching groups were compared to those of the Sham group from the respective time points. Differentially methylated promoters were identified based on FDR-adjusted *p*-values derived from the betabinomial model and using a cut-off of 0.05 (see Section 2). With these criteria, we found that all of the treatments had a minor effect on DNA methylation at promoters in both lung (Fig. 2B) and liver (Fig. S2B) tissues. Only a few promoters underwent significant changes in methylation upon treatment, with the maximum number of 50 differentially methylated promoters observed for the 3R4F-treated group at six months in the lung (Fig. 2B). This observation is in striking contrast to the changes observed in gene expression in the lung, where more than 3000 genes were found to be differentially expressed in the 3R4F-treated samples.

DNA methylation within the gene body positively correlates with expression level in some cell types and is unrelated to expression status in others (Lister et al., 2009). Gene bodies are often found to be highly methylated, except at intragenic enhancer elements. In both lung and liver samples, we found that the majority of gene bodies were highly methylated regardless of gene expression level (Fig. S1B).

Taken together, our analysis indicated that the changes observed in gene expression in the lung due to CS exposure were largely unrelated to DNA methylation changes at annotated promoters and gene bodies.

3.3. CS, but not THS 2.2, alters DNA methylation at enhancer elements in the lung but not in the liver

Enhancers are distal regulatory elements that control gene expression in a temporal and cell type-specific manner. They are characterized by specific histone marks, such as monomethylated lysine 4 (H3K4me1) and acetylated lysine 27 (H3K27ac) on histone 3 (Heintzman et al., 2007; Creyghton et al., 2010). Enhancers are also characterized by reduced local methylation (Stadler et al., 2011a). Therefore, identifying LMRs in high-resolution methylomes directly predicts TF binding sites and enhancer activity (Burger et al., 2013; Boller et al., 2016; Xu et al., 2016).

To assess DNA methylation at enhancers, we first identified LMRs in every experimental condition and tissue (Fig. 3) using MethylSeekR algorithm (Burger et al., 2013). LMRs from all of the experimental conditions were combined into one unified set for each tissue. To reduce loci redundancy, LMRs with more than 60% overlap were merged. These unified sets were used in all subsequent analyses.

To ensure the accuracy of LMR detection and the role of these regions as potential enhancers, we took advantage of the public ChIP-Seq data performed from lung and liver chromatin of eight-week-old mice (GEO: GSE29184) (Shen et al., 2012). We quantified H3K4me1, trimethylated lysine 4 on histone 3 (H3K4me3), and H3K27ac histone marks and the input control (DNA extracted from the chromatin prior to the immunoprecipitation) at LMRs and UMRs detected in Sham-exposed mice at the one-month time point, corresponding to 12- to 14-week-old mice. As expected, LMRs harbored an enhancer signature that consists of high H3K4me1 and H3K27ac and low H3K4me3 signals (Figure 3C, S2A). By contrast, UMRs were enriched in H3K4me3 and H3K27ac signals, which are hallmarks of active promoters, and showed a low level of H3K4me1 signal (Fig. 3C, Fig. S2A). Therefore, this comparison of our data with public data from independent samples and with distinct methods consolidated the approach used in our study to identify regulatory elements.

Similarly to promoters, DNA methylation levels at LMRs from the different samples were compared to those in Sham samples from the respective time points. We found that CS induced a change in DNA methylation level at hundreds of LMRs in the lung when compared with Sham control. At the early time points (one, two, and three months), CS altered the DNA methylation levels of approximately 200 LMRs, and this number increased over time, reaching a maximum of

approximately 1000 LMRs at eight months. The majority of affected LMRs underwent hypermethylation, while only a small number were hypomethylated (Fig. 4). In contrast to CS, THS 2.2 altered DNA methylation of a small set of LMRs, with the maximum number observed at one month (75 hypermethylated and 74 hypomethylated LMRs) and deceasing over time. At eight months, only 53 LMRs were hypermethylated and 37 hypomethylated in the THS 2.2-treated samples (Figs. 4 and 5). Interestingly, at the later time points (six and eight months) the number of affected LMRs in the cessation and switching samples was higher than in the samples from the continuous exposure to THS 2.2 group. Surprisingly, we observed a limited overlap between differentially methylated LMRs at different time points from the same exposure group, suggesting heterogeneous and stochastic events. In contrast to the lung, DNA methylation levels at LMRs in liver methylomes were largely unchanged in all experimental conditions (Fig. S2C).

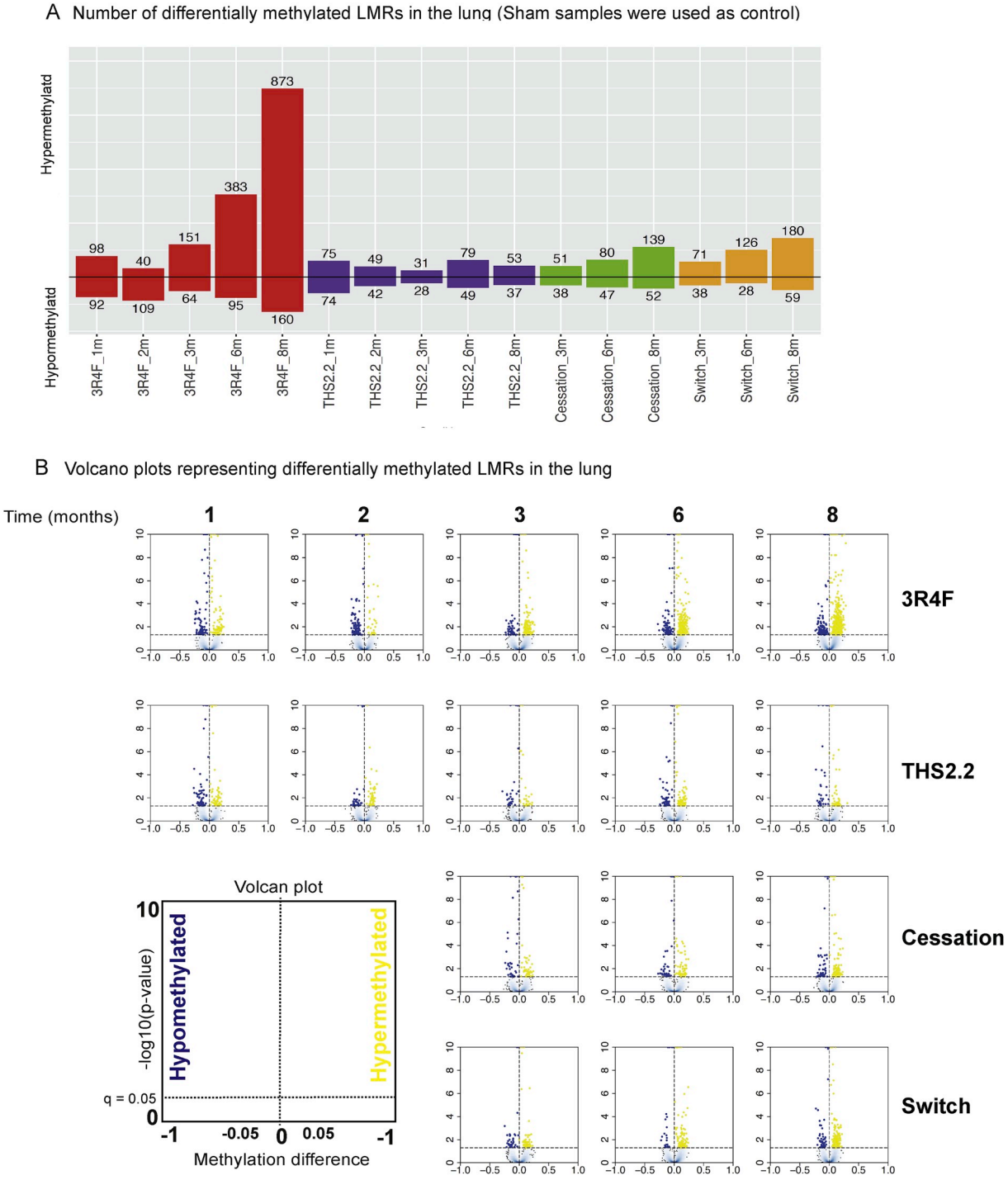

Fig. 4. Assessment of DNA methylation changes at LMRs in the lung.
A. Barplot representing the number of differentially methylated LMRs based on the FDR cut-off (0.05). The respective Sham samples were used as controls. B. Volcano plots representing the amplitude and significance of methylation changes between the indicated treatment and the respective Sham samples. Methylation difference (treatment vs. Sham) is plotted on the *x*-axis, and the statistical significance, calculated as $-\log_{10}$ (FDR-adjusted *p*-value) is plotted on the *y*-axis. Yellow and blue dots indicate hypermethylated and hypomethylated LMRs, respectively, relative to the Sham samples.

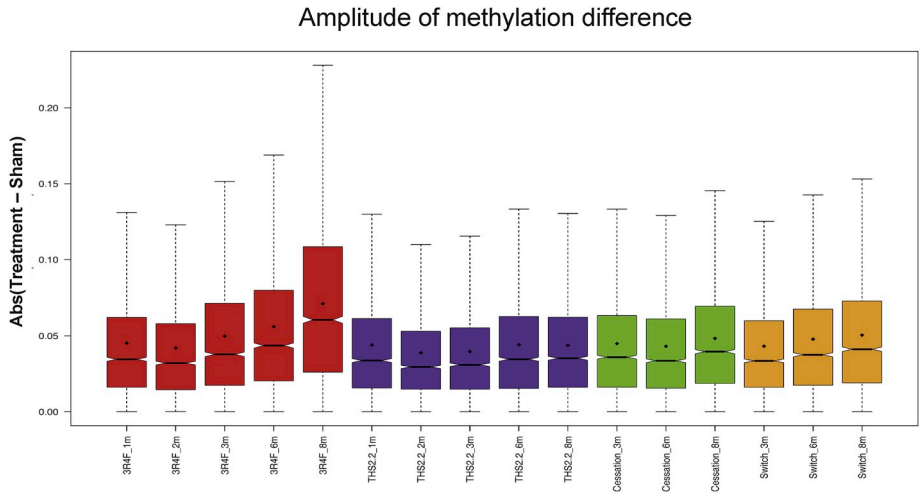

Fig. 5. Amplitude of methylation difference at differentially methylated LMRs.
Boxplots representing absolute values of methylation difference between treatment and Sham groups for all differentially methylated LMRs. Mean values are indicated by black dots.

3.4. Ets and Fox TF motifs are enriched in lung enhancers and member of these families are downregulated in CS-treated samples

TF motif analysis of the unified catalogue of LMRs from lung methylomes showed that these loci were enriched mainly for motifs corresponding to the E-twenty-six (Ets) and forkhead (Fox) families of TFs and the insulator factor CTCF, with the ERG, ETV2, ETS1, and FLI1 motifs having the lowest *p*-values (Supplementary Table 1). By contrast, LMRs detected from liver methylomes were enriched in motifs corresponding to TFs known to be involved in liver function, including Foxa1 (Moya et al., 2012), Foxa2 (Wolfrum et al., 2004), and HNF4a as well as CTCF (Costa et al., 1989; Drewes et al., 1996) (Supplementary Table 2).

It has been shown that TF binding shapes DNA methylation profiles by creating LMRs (Stadler et al., 2011b). We hypothesized that the 3R4F-induced hypermethylation of lung LMRs may be due to the downregulation of some of the TFs that bind to these elements. In line with this hypothesis, we found that *Erg, Ets1*, and *Fli1* genes were downregulated in 3R4F-treated groups, mainly at the eight-month time point, but not in the other experimental groups (Fig. 6A). Similarly, many genes coding for FOX TFs were downregulated in 3R4F-treated samples (Fig. 6B). Interestingly, the CTCF motif, which is highly enriched in the common LMR catalogue (p-value $< 1e-50$) is not enriched in the 3R4F hypermethylated LMRs in the lung ((Supplementary Tables 1 and 3). This suggests that CS may specifically target members of the Ets and Fox families of TFs, leading to reduced binding to their target loci and therefore to reduced protection from the DNA methylation machinery that ultimately causes the observed hypermethylation of these loci.

4. Discussion

Previous studies reported that CS can alter DNA methylation in various human tissues. However, the majority of these studies interrogated DNA methylation at genomic loci covered by methylation arrays, which consist mainly of annotated regulatory elements, such as promoters and a small subset of enhancers. This strategy is convenient for large-scale clinical studies but has major limitations given the complexity of mammalian genomes and the central role of distal regulatory elements (enhancers), which are tissue- and cell type-specific and therefore conceptually impossible to capture with arrays. Genome-wide investigations of genomic and epigenetic alterations related to CS are crucial to better understand the subsequent phenotypes linked to this exposure and to develop candidate genomic biomarkers to assess the harm of cigarettes or alternative tobacco products.

Here we investigated DNA methylation alterations related to CS or aerosol from THS 2.2 exposure in the ApoE$^{-/-}$ mouse model in lung and liver tissues. Considering the genomic context readout of CpG methylation, we first investigated DNA methylation at promoters and gene bodies. While DNA methylation at promoters showed a clear bimodal distribution and anticorrelated with gene expression levels, DNA methylation inside gene bodies was equally high, with no correlation with gene expression, which is similar to what has been observed in human embryonic stem cells (Lister et al., 2009). DNA methylation levels at promoters were largely unchanged in all experimental conditions in both lung and liver tissues. These results strongly suggest that the massive changes observed in gene expression levels for 3R4F-exposed samples in the lung (Phillips et al., 2016) were not related to DNA methylation at promoters or at gene bodies. The observed transcriptional changes may be a result of other epigenetic mechanisms or the activity of TFs. Theoretically, almost every promoter in mammalian genomes is regulated by at least one enhancer, and the transcriptomic dynamics are often associated with changes in enhancer activity. To investigate the effect of CS and THS 2.2 aerosol exposure on enhancer activity, we used methylome data to identify candidate enhancers as regions with low methylation levels in every experimental group. We found that CS altered DNA methylation at almost 1000 candidate lung enhancers. This effect was largely reduced in THS 2.2 aerosol-exposed lung samples and in the cessation and switching groups. Interestingly, CS exposure had a limited effect on DNA methylation in liver tissue at both enhancers and promoters, arguing for a tissue-specific impact of CS on the epigenome.

Importantly, we observed that the amplitudes of DNA methylation changes observed in lung samples upon CS exposure were modest (median = 15%) and did not increase with exposure time. Rather, we observed an increase in the number of affected loci; those loci were different between time points, suggesting highly dynamic and potentially stochastic epigenetic events induced by CS exposure.

CS-induced DNA methylation changes observed in the lung were largely oriented toward hypermethylation. To determine whether this directional change is linked to dysregulation of DNA methylation-associated enzymes, we verified the expression levels of genes coding for DNMTs and Tet proteins and found that none of these genes were deregulated by all of the exposures tested. This strongly suggests that the observed alterations in DNA methylation are not linked to a general perturbation of DNA methylation machinery but rather result from locus-specific events, probably involving sequence-specific DNA

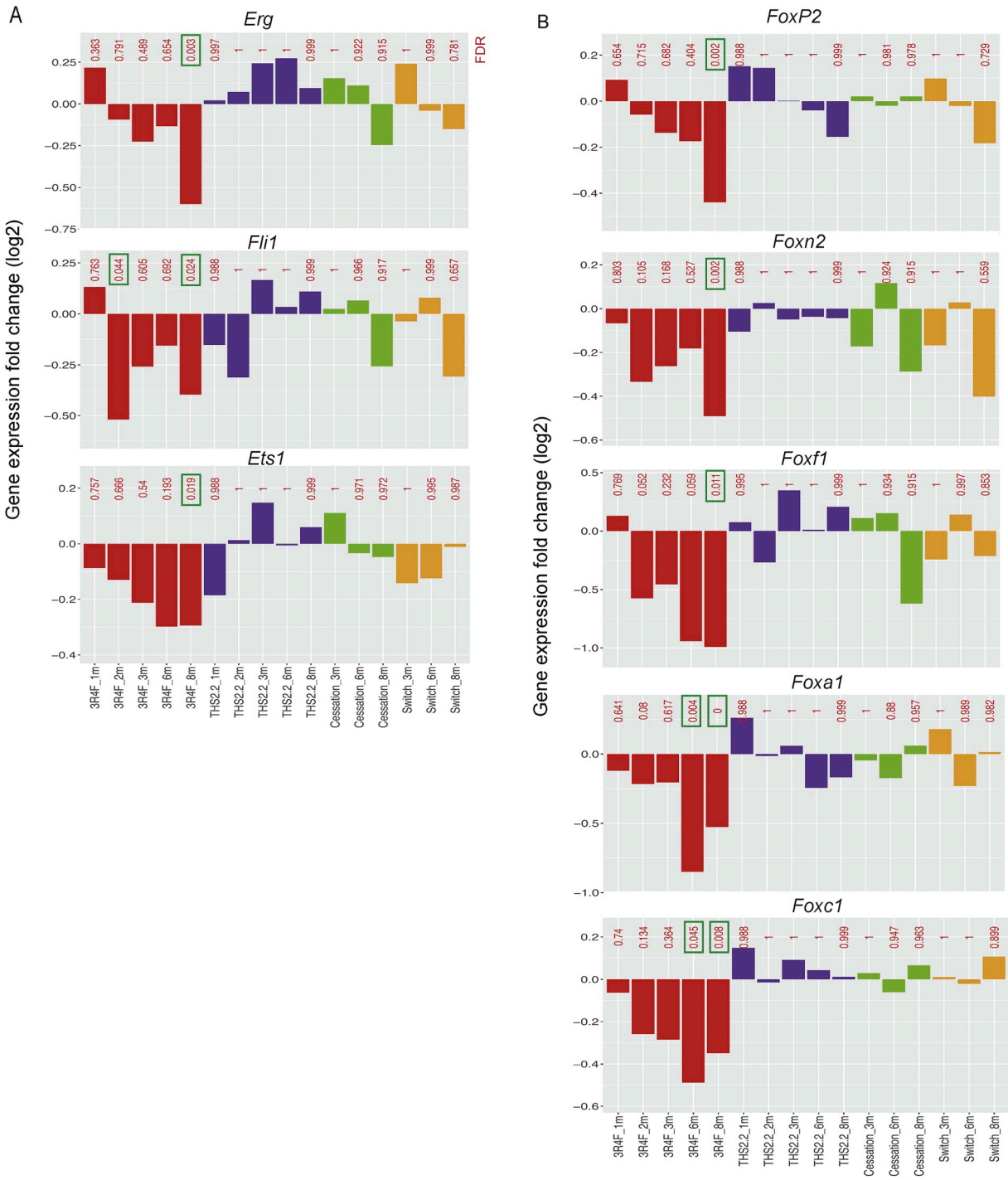

Fig. 6. Expression levels of TFs whose motifs are enriched in lung LMRs.
Expression levels of genes coding for some Ets (A) and Fox (B) TFs that were found to be deregulated in at least one experimental group. *Y*-axis indicates fold change in $\log_2$ scale. FDR-adjusted *p*-values are indicated in red, and significant changes are indicated by green rectangles.

binding proteins, such as TFs.

TF motif analysis of lung data suggests that members of the Ets and Fox families may play roles in the epigenetic and transcriptional response to CS exposure. The Ets family of TFs are important regulators of a variety of genes that control endothelial homeostasis (Oh et al., 2015; Yuan et al., 2009). For example, ERG and FLI1 play important roles in lung homeostasis by controlling the expression of many inflammatory genes (Grzegorzewska et al., 2017). Members of the Fox family have also been reported to be involved in lung development (Mahlapuu et al., 2001; Yang et al., 2010), function (Clevidence et al., 1994), and related diseases (Kim et al., 2006). Many genes coding for Ets and Fox TFs were found to be downregulated in CS-exposed samples, in agreement with the enrichment of their respective motifs in hypermethylated LMRs. This observation suggests that CS-related hypermethylation of regulatory elements may occur through the perturbation of TF binding profiles. It is known that TF binding protects DNA from methylation (Stadler et al., 2011a; Boller et al., 2016). Downregulation of some TFs may lead to aberrant hypermethylation of their genomic targets. Further investigations of the binding profiles of these TFs using ChIP-Seq approach are necessary to better understand the relationship between

CS exposure and TF activity. It is also important to mention that cell type heterogeneity in lung tissue may dilute cell type-specific alterations. Investigating homogenous cell populations may lead to the identification of more obvious differences in the activity of these TFs and their target loci.

Linking changes in enhancer activity to the expression of target promoters is a challenging task without information about the physical interaction between enhancers and promoters. This type of information can be achieved by chromosome capture strategies (Lieberman-Aiden et al., 2009; van Berkum et al., 2010), but to date, the implementation of such an approach in a large-scale systems toxicology study is challenging because of the high sequencing depth that is required to generate high-resolution maps. Future improvements of sequencing strategies may allow to circumvent these limitations.

5. Conclusions

This study reveals that the effect of CS exposure on DNA methylation mainly occurs at distal regulatory elements. This effect is tissue-specific, as it is observed in the lung but not in the liver. The majority of affected loci underwent hypermethylation at the late time points, after six months of exposure. The observed DNA methylation alterations are strongly reduced after the cessation of CS exposure or after switching to THS 2.2. The affected loci were enriched for motifs corresponding to the ETS and FOX families of TFs, suggesting a potential role of these proteins in the epigenetic response to CS exposure.

Declarations

Ethics approval

All procedures involving animals were performed in an Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC)-accredited, Agri-Food & Veterinary Authority of Singapore-licensed facility with approval from an Institutional Animal Care and Use Committee (IACUC protocol 15015), in compliance with the National Advisory Committee for Laboratory Animal Research Guidelines on the Care and Use of Animals for Scientific Purposes (NACLAR, 2004).

Competing interests

All authors are employees of Philip Morris International. Philip Morris International is the sole source of funding and sponsor of this project.

Public data

The publicly available datasets used in this study are indicated in Fig. 3 and Supplementary Fig. 2.

Data availability

Raw read counts per CpG are available on Intervals platform at the following URL: https://www.intervals.science/studies/#/apoe-ths22-sw. Lung and liver methylome datasets are labeled “Lung_DNA_Methylation_CpG_count” and “Liver_DNA_Methylation_CpG_count” respectively.

Funding

Philip Morris International is the sole source of funding and sponsor of this project.

CRediT authorship contribution statement

Mohamed-Amin Choukrallah: Data curation, Formal analysis, Writing – original draft. **Nicolas Sierro:** Supervision. **Florian Martin:** Formal analysis. **Karine Baumer:** Methodology. **Jerome Thomas:** Methodology. **Sonia Ouadi:** Methodology. **Julia Hoeng:** Conceptualization. **Manuel C. Peitsch:** Conceptualization. **Nikolai V. Ivanov:** Supervision.

Acknowledgements

We are grateful to the members of the *in vivo* experimentation teams at Philip Morris International Research Laboratories, Singapore, for their technical contributions to this project. We thank the members of the bioinformatic team at Philip Morris International R&D for their computational support and help with data management.

Transparency document

Transparency document related to this article can be found online at https://doi.org/10.1016/j.fct.2018.11.020.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fct.2018.11.020.

References

Bestor, T.H., 2000. The DNA methyltransferases of mammals. Hum. Mol. Genet. 9, 2395–2402.
Boller, S., Ramamoorthy, S., Akbas, D., Nechanitzky, R., Burger, L., Murr, R., Schubeler, D., Grosschedl, R., 2016. Pioneering activity of the C-terminal domain of EBF1 shapes the chromatin landscape for B cell programming. Immunity 44, 527–541.
Breitling, L.P., Yang, R., Korn, B., Burwinkel, B., Brenner, H., 2011. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. Am. J. Hum. Genet. 88, 450–457.
Bulger, M., Groudine, M., 2011. Functional and mechanistic diversity of distal transcription enhancers. Cell 144, 327–339.
Burger, L., Gaidatzis, D., Schubeler, D., Stadler, M.B., 2013. Identification of active regulatory regions from DNA methylation data. Nucleic Acids Res. 41, e155.
Choukrallah, M.-A., Song, S., Rolink, A.G., Burger, L., Matthias, P., 2015. Enhancer repertoires are reshaped independently of early priming and heterochromatin dynamics during B cell differentiation. Nat. Commun. 6.
Chu, K.M., Cho, C.H., Shin, V.Y., 2013. Nicotine and gastrointestinal disorders: its role in ulceration and cancer development. Curr. Pharmaceut. Des. 19, 5–10.
Clevidence, D.E., Overdier, D.G., Peterson, R.S., Porcella, A., Ye, H., Paulson, K.E., Costa, R.H., 1994. Members of the HNF-3/forkhead family of transcription factors exhibit distinct cellular expression patterns in lung and regulate the surfactant protein B promoter. Dev. Biol. 166, 195–209.
Costa, R.H., Grayson, D.R., Darnell Jr., J.E., 1989. Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and alpha 1-antitrypsin genes. Mol. Cell Biol. 9, 1415–1425.
Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U. S. A. 107, 21931–21936.
Doll, R., Hill, A.B., 1950. Smoking and carcinoma of the lung; preliminary report. Br. Med. J. 2, 739–748.
Drewes, T., Senkel, S., Holewa, B., Ryffel, G.U., 1996. Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. Mol. Cell Biol. 16, 925–931.
Gaidatzis, D., Lerch, A., Hahne, F., Stadler, M.B., 2015. QuasR: quantification and annotation of short reads in R. Bioinformatics 31, 1130–1132.
Grzegorzewska, A.P., Han, R., Stawski, L., Marden, G., Iwamoto, M., Trojanowska, M., 2017 Jul. Synergistic role of endothelial ERG and FLI1 in mediating pulmonary vascular homeostasis. Am. J. Respir. Cell Mol. Biol. 57 (1). https://doi.org/10.1165/rcmb.2016-0200OC.
He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al., 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science (New York, NY) 333, 1303–1307.
Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al., 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. 39, 311–318.
Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al., 2009. Histone modifications at human

enhancers reflect global cell-type-specific gene expression. Nature 459, 108–112.
Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., Glass, C.K., 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell. 38, 576–589.
Joehanes, R., Just, A.C., Marioni, R.E., Pilling, L.C., Reynolds, L.M., Mandaviya, P.R., Guan, W., Xu, T., Elks, C.E., Aslibekyan, S., et al., 2016. Epigenetic signatures of cigarette smoking. Circ. Cardiovasc. Genet. 9, 436–447.
Joubert, B.R., Haberg, S.E., Nilsen, R.M., Wang, X., Vollset, S.E., Murphy, S.K., Huang, Z., Hoyo, C., Midttun, O., Cupul-Uicab, L.A., et al., 2012. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. Environ. Health Perspect. 120, 1425–1431.
Kim, I.M., Ackerson, T., Ramakrishna, S., Tretiakova, M., Wang, I.C., Kalin, T.V., Major, M.L., Gusarova, G.A., Yoder, H.M., Costa, R.H., Kalinichenko, V.V., 2006. The Forkhead Box m1 transcription factor stimulates the proliferation of tumor cells during development of lung cancer. Cancer Res. 66, 2153–2161.
Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Kin Sung, K.W., Rigoutsos, I., Loring, J., Wei, C.L., 2010. Dynamic changes in the human methylome during differentiation. Genome Res. 20, 320–331.
Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293.
Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., et al., 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462, 315–322.
Lo Sasso, G., Titz, B., Nury, C., Boue, S., Phillips, B., Belcastro, V., Schneider, T., Dijon, S., Baumer, K., Peric, D., et al., 2016. Effects of cigarette smoke, cessation and switching to a candidate modified risk tobacco product on the liver in ApoE -/- mice–a systems toxicology analysis. Inhal. Toxicol. 28, 226–240.
Mahlapuu, M., Enerback, S., Carlsson, P., 2001. Haploinsufficiency of the forkhead gene Foxf1, a target for sonic hedgehog signaling, causes lung and foregut malformations. Development 128, 2397–2406.
Moya, M., Benet, M., Guzman, C., Tolosa, L., Garcia-Monzon, C., Pareja, E., Castell, J.V., Jover, R., 2012. Foxa1 reduces lipid accumulation in human hepatocytes and is down-regulated in nonalcoholic fatty liver. PloS One 7, e30014.
Oh, S.Y., Kim, J.Y., Park, C., 2015. The ETS factor, ETV2: a master regulator for vascular endothelial cell development. Mol. Cell. 38, 1029–1036.
Phillips, B., Veljkovic, E., Boue, S., Schlage, W.K., Vuillaume, G., Martin, F., Titz, B., Leroy, P., Buettner, A., Elamin, A., et al., 2016. An 8-month systems toxicology inhalation/cessation study in ApoE-/- mice to investigate cardiovascular and respiratory exposure effects of a candidate modified risk tobacco product, THS 2.2, compared with conventional cigarettes. Toxicol. Sci. 149, 411–432.
Rogers, S., de Souza, A.R., Zago, M., Iu, M., Guerrina, N., Gomez, A., Matthews, J., Baglole, C.J., 2017. Aryl hydrocarbon receptor (AhR)-dependent regulation of pulmonary miRNA by chronic cigarette smoke exposure. Sci. Rep. 7, 40539.
Roth, G.M., 1954. Effects of smoking of tobacco on the cardiovascular system of normal persons and patients with hypertension. J. Am. Geriatr. Soc. 2, 270–273.
Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., Ren, B., 2012. A map of the cis-regulatory sequences in the mouse genome. Nature 488, 116–120.
Shenker, N.S., Polidoro, S., van Veldhoven, K., Sacerdote, C., Ricceri, F., Birrell, M.A., Belvisi, M.G., Brown, R., Vineis, P., Flanagan, J.M., 2013. 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. 22, 843–851.
Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Scholer, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., et al., 2011a. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 480, 490–495.
Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Scholer, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., et al., 2011b. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 480, 490–495.
van Berkum, N.L., Lieberman-Aiden, E., Williams, L., Imakaev, M., Gnirke, A., Mirny, L.A., Dekker, J., Lander, E.S., 2010 May. Hi-C: a method to study the three-dimensional architecture of genomes. JoVE 6 (39). https://doi.org/10.3791/1869. pii: 1869.
Wolfrum, C., Asilmaz, E., Luca, E., Friedman, J.M., Stoffel, M., 2004. Foxa2 regulates lipid metabolism and ketogenesis in the liver during fasting and in diabetes. Nature 432, 1027–1032.
Xu, J., Hu, H., Dai, Y., 2016. LMethyR-svm: predict human enhancers using low methylated regions based on weighted support vector machines. PloS One 11, e0163491.
Yang, Z., Hikosaka, K., Sharkar, M.T., Tamakoshi, T., Chandra, A., Wang, B., Itakura, T., Xue, X., Uezato, T., Kimura, W., Miura, N., 2010. The mouse forkhead gene Foxp2 modulates expression of the lung genes. Life Sci. 87, 17–25.
Yuan, L., Nikolova-Krstevski, V., Zhan, Y., Kondo, M., Bhasin, M., Varghese, L., Yano, K., Carman, C.V., Aird, W.C., Oettgen, P., 2009. Antiinflammatory effects of the ETS factor ERG in endothelial cells are mediated through transcriptional repression of the interleukin-8 gene. Circ. Res. 104, 1049–1057.
Zhang, Y., Elgizouli, M., Schottker, B., Holleczek, B., Nieters, A., Brenner, H., 2016. Smoking-associated DNA methylation markers predict lung cancer incidence. Clin. Epigenet. 8, 127.