

RESEARCH

Open Access



From blood to lung tissue: effect of cigarette smoke on DNA methylation and lung function

Maaïke de Vries^{1,2*}, Diana A van der Plaat^{1,2}, Ivana Nedeljkovic³, Rikst Nynke Verkaik-Schakel⁴, Wierd Kooistra^{2,5}, Najaf Amin³, Cornelia M van Duijn³, Corry-Anke Brandsma^{2,5}, Cleo C van Diemen⁶, Judith M Vonk^{1,2} and H Marïke Boezen^{1,2}

Abstract

Background: Genetic and environmental factors play a role in the development of COPD. The epigenome, and more specifically DNA methylation, is recognized as important link between these factors. We postulate that DNA methylation is one of the routes by which cigarette smoke influences the development of COPD. In this study, we aim to identify CpG-sites that are associated with cigarette smoke exposure and lung function levels in whole blood and validate these CpG-sites in lung tissue.

Methods: The association between pack years and DNA methylation was studied genome-wide in 658 current smokers with >5 pack years using robust linear regression analysis. Using mediation analysis, we subsequently selected the CpG-sites that were also associated with lung function levels. Significant CpG-sites were validated in lung tissue with pyrosequencing and expression quantitative trait methylation (eQTM) analysis was performed to investigate the association between DNA methylation and gene expression.

Results: 15 CpG-sites were significantly associated with pack years and 10 of these were additionally associated with lung function levels. We validated 5 CpG-sites in lung tissue and found several associations between DNA methylation and gene expression.

Conclusion: This study is the first to validate a panel of CpG-sites that are associated with cigarette smoking and lung function levels in whole blood in the tissue of interest: lung tissue.

Keywords: Cigarette smoking, Lung function, DNA methylation, EWAS, Lung tissue

Background

Chronic Obstructive Pulmonary Disease (COPD) is a chronic and progressive inflammatory lung disease with cigarette smoking as the main risk factor. COPD is characterized by persistent airflow limitation and COPD patients suffer from severe respiratory symptoms, overall resulting in a poor quality of life [1]. The global prevalence of COPD is 10.7% [2], resulting in a high economic and societal burden. In 2015, 5% of all deaths globally were caused by COPD and

it is expected that COPD will be the third leading cause of death worldwide in 2030 (WHO 2016).

The development of COPD is known to be associated with both genetic [3–5] and environmental factors [6] and their interactions [7]. However, the genome-wide association studies (GWAS) and genome-wide interaction studies (GWIS) that have been performed so far identified variants in COPD susceptibility genes that only explain a very small part of the variation in the onset of COPD [8]. As a consequence, the epigenome is increasingly recognized as an important link between changes to the inherited genome and environmental exposures such as cigarette smoke [9]. The epigenome comprises several epigenetic mechanisms that affect gene expression without modifying

* Correspondence: m.de.vries04@umcg.nl

¹University of Groningen, University Medical Center Groningen, Department of Epidemiology, Hanzplein 1, Groningen 9713, GZ, The Netherlands

²University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD (GRIAC), Groningen, The Netherlands

Full list of author information is available at the end of the article



the DNA sequence. These epigenetic mechanisms are highly dynamic and changes can be induced by environment exposures, diseases and ageing [10].

One well-defined epigenetic mechanism is DNA methylation, which is tissue-specific and involves the binding of a methyl group to a cytosine base adjacent to a guanine base, a so called CpG-site. CpG rich sites are found in the regulatory regions of the DNA and methylation of CpG-sites in these regulatory regions leads to a decrease in gene expression [11]. It has been shown that DNA methylation is highly affected by environmental exposures such as air pollution and cigarette smoke [12–14]. Next to the fact that exposure to cigarette smoke is an important risk factor for COPD, it is also strongly associated with lower lung function levels [15, 16]. Hence, we postulate that DNA methylation plays an important role in the etiology of COPD by mediating the effect that cigarette smoking has on lung function levels. In this study, we aim to identify these CpG-sites by performing an epigenome-wide association study (EWAS) in whole blood in current smokers and validate these smoking-related differences in DNA methylation in lung tissue.

Methods

Study design

As indicated in the flowchart in Fig. 1, our complete study can be divided into two consecutive studies.

Study I: Epigenome-wide association study (EWAS) in whole blood

We performed an epigenome-wide association study (EWAS) in whole blood to identify differential DNA methylation related to pack years in a selection of current smokers from the LifeLines population-based cohort study. Subsequently, we studied whether the identified CpG-sites were also associated with lung function levels using mediation analysis.

Study II: Validation study in lung tissue

We validated the significant mediating CpG-sites in actual lung tissue. Subsequently, we studied the association between these CpG-sites and gene expression in lung tissue.

A detailed description of the study population and measurements can be found in the online supplement (Additional file 1).

Study I: Epigenome-wide association study (EWAS) in whole blood

Study population

The association between pack years as a cumulative measurement of cigarette smoking and DNA methylation in whole blood was studied in 658 subjects from the Dutch general population-based cohort study LifeLines at baseline [17]. All selected subjects were current smokers with a smoking history of at least 5 pack years.

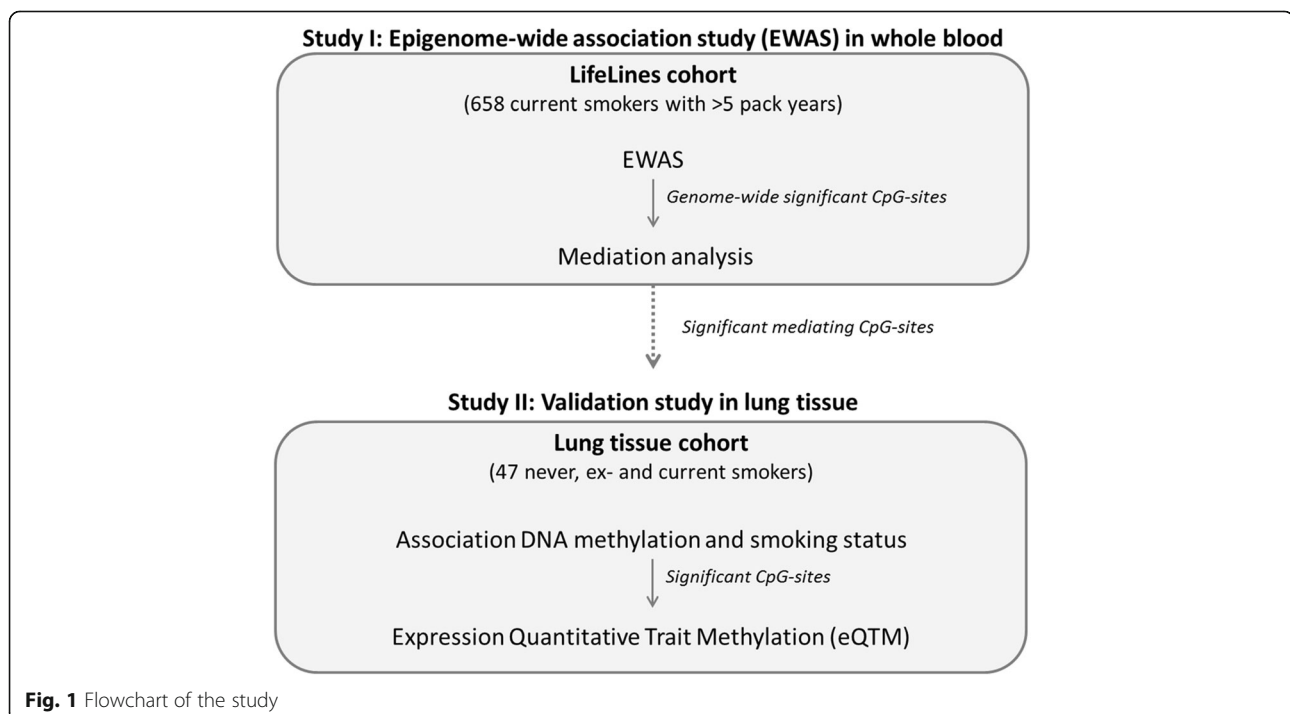


Fig. 1 Flowchart of the study

Measurements

DNA methylation levels in whole blood were determined with the Illumina Infinium Human Methylation 450 K array. After quality control, the final dataset contained DNA methylation data presented as beta values for 420,938 CpG probes [18] (See supplementary methods for sample processing and quality control).

Statistical analyses

Epigenome-wide association analysis (EWAS) in whole blood We assessed the association between pack years in current smokers and epigenome-wide DNA methylation in whole blood. We performed robust linear regression using the *MASS* package in R and adjusted the model for the potential confounders age and gender. To adjust for technical variation, principal components were calculated using the 220 control probes incorporated in the Illumina 450 K chip. The principal components that explain >1% of the technical variation ($N = 7$) were included in the analysis. To adjust for the cellular heterogeneity of the whole blood samples, we included proportional white blood cell counts of mononuclear cells, lymphocytes, neutrophils and eosinophils obtained by standard laboratory techniques.

Mediation analysis To assess if the CpG-sites identified in the EWA study were also associated with lung function levels, we performed mediation analysis using the R-package *Mediation* [19]. For every individual CpG-site, the Average Causal Mediation Effect (ACME), Average Direct Effect (ADE), Total Effect and Proportion Mediated was calculated. When the p -value of ACME was below 0.05 for a particular CpG-site (tested 2-sided), this CpG-site was considered to be significantly associated with lung function levels.

Study II: Validation study in lung tissue

Study population

DNA methylation in lung tissue was studied in patients undergoing lung resection, lung volume reduction or transplant surgery in the University Medical Center Groningen as previously described [20]. If resection surgery was conducted for tumor removal, macroscopically normal lung tissue was taken far distant from the tumor. All samples were histologically checked for abnormalities using standard haematoxylin and eosin staining. To reach a power of 90% with an alpha of 5%, we selected a total of 47 subjects, being 16 never smokers, 15 ex-smokers and 16 current smokers. The power analysis can be found in the online supplement. All ex- and current smokers had a smoking history of >5 pack years and all ex-smokers quit smoking for at least 1 year.

We studied gene expression in 36 of the 47 selected subjects included in our lung tissue validation study [20].

Measurements

DNA methylation in lung tissue DNA methylation in lung tissue was determined with the PyroMark Q48 Autoprep System. Quality of the pyrosequencing was assessed for every CpG-site in all the subjects individually and the subject was excluded for a particular CpG-site if the default software quality control was not passed. Primers for cg01940273 did not match with the original sequence and cg01940273 was excluded from the validation study. Primer sequences are shown in the online supplement.

Gene expression in lung tissue Total RNA was extracted from lung tissue at Rosetta Inpharmatics Gene Expression Laboratory (Seattle, WA, USA). mRNA profiling was performed using a custom-made Affymetrix HU133 array (GPL 10379) containing 751 controls probe sets and 51,627 non-control probe sets. Gene expression was normalized with the Robust Multichip Average method implemented in the Affymetrix Power Tools software [20].

Statistical analysis

DNA methylation in lung tissue The association between smoking status and DNA methylation in lung tissue was studied using linear regression analysis using R. The analyses were adjusted for the potential confounders age and gender.

Expression quantitative trait methylation (eQTM) analysis in lung tissue For the CpG-sites that were differentially methylated in lung tissue, we tested the association between DNA methylation levels and gene expression, i.e. expression quantitative trait methylation (eQTM) analysis. Per CpG-site, all the gene expression probe sets within 1 MB on either side were selected. An association was considered significant when the p -value was lower than the Bonferroni adjusted p -value ($=0.05 /$ the number of probe sets within a 2 MB window of the indicated CpG-site).

Results

Study I: Epigenome-wide association study (EWAS) in whole blood

Study population

Population characteristics are shown in Table 1. Notably, the study population was non-random, i.e. selected based on the presence or absence of airway obstruction.

Epigenome-wide association analysis (EWAS) in whole blood

The EWAS in whole blood identified 15 CpG-sites that were significantly (Bonferroni corrected p -value $< 0.05/420938 = 1.19 \times 10^{-7}$) associated with pack years (Fig. 2). Higher number of pack years was associated with lower levels of DNA methylation in 13 of the 15 identified CpG-sites (see Table 2). We identified two novel CpG-sites, cg22994830

Table 1 Population characteristics

	LifeLines Cohort
Number	658
Males, N (%)	375 (57,0)
Age, years (range)	46 (22–79)
Pack year, years, mean (range)	20,6 (5–100)
FEV ₁ /FVC (%), mean (range)	71,7 (40,1 – 92,4)
FEV ₁ (%pred)*, mean (range)	94,4 (47,7 – 138,4)
COPD cases, N (%)	279 (42,4)
GOLD COPD Stage ≥ 2†, N (%)	102 (15,6)

* Calculated with GLI-2012 if possible

† COPD GOLD stage > 2 (FEV₁/FVC < 70% and FEV₁ between 50 and 80% of predicted)

COPD, Chronic Obstructive Pulmonary Disease; FEV₁, Forced Expiratory Volume in 1 s; FVC, Forced Vital Capacity

annotated to Protein Kinase CAMP-Dependent Type 1 Regulatory Subunit Beta (*PRKAR1B*) and cg20451986 located in the intergenic region of *11q25* between the genes *loc100126239* and Junctional Adhesion Molecule 3 (*JAM3*). As indicated in Table 2, we identified several CpG-sites to be related to pack years, that were described in earlier studies in association with smoking status [12, 21–24]. An overview of all significant CpG sites at nominal *p*-value < 0.05 can be obtained upon request by the corresponding author.

Mediation analysis

The mediation analysis revealed that 10 CpG-sites were significantly associated with lung function levels (Additional file 2).

Study II: Validation study in lung tissue

Study population

We validated DNA methylation levels in lung tissue of 47 subjects, divided into 3 equal groups based on smoking status (see Additional file 3 for subject characteristics).

DNA methylation in lung tissue

Figure 3 shows that the current smokers had significant lower DNA methylation levels at 4 out of the 9 CpG sites than never smokers (Bonferroni corrected *p*-value of 0.05/9=0.00556). For cg21566642, the DNA methylation levels were significantly different between current and never smokers at a nominal *p*-value of 0.05. An overview of the complete analysis can be found in the online supplement (Additional file 4).

Expression quantitative trait methylation (eQTM) analysis in lung tissue

DNA methylation at cg05575921 and cg21161138 (the two CpG-sites annotated to *AHRR*) was significantly associated with gene expression levels of *AHRR* in lung tissue (Fig. 4a and c). As indicated in Fig. 4b and d, current smokers showed lower DNA methylation levels together with higher expression levels compared to ex- and never smokers. For the other 3 differentially methylated CpG-sites in lung tissue, we found novel significant associations with gene expression levels. DNA methylation levels at cg05951221 and cg21566642, located in the intergenic region on chromosome 2, were associated with gene expression levels of the genes Autophagy Related 16 Like1 (*ATG16L1*) and DIS3 Like 3'-5' Exoribonuclease (*DIS3L2*) (Fig. 5a-d). DNA methylation levels at cg06126421, located in the intergenic region on chromosome 6, were associated with gene expression levels of the genes Mucin 21 (*MUC21*) and Tubulin Beta Class I (*TUBB*) (Fig. 5e-f). A complete overview of the eQTM analysis can be found in Additional file 5.

Discussion

In this study, we showed that DNA methylation at 15 CpG-sites was significantly associated with pack years. Next to previously described CpG-sites, we identified 2 novel CpG-sites: cg22994830, annotated to *PRKAR1B*,

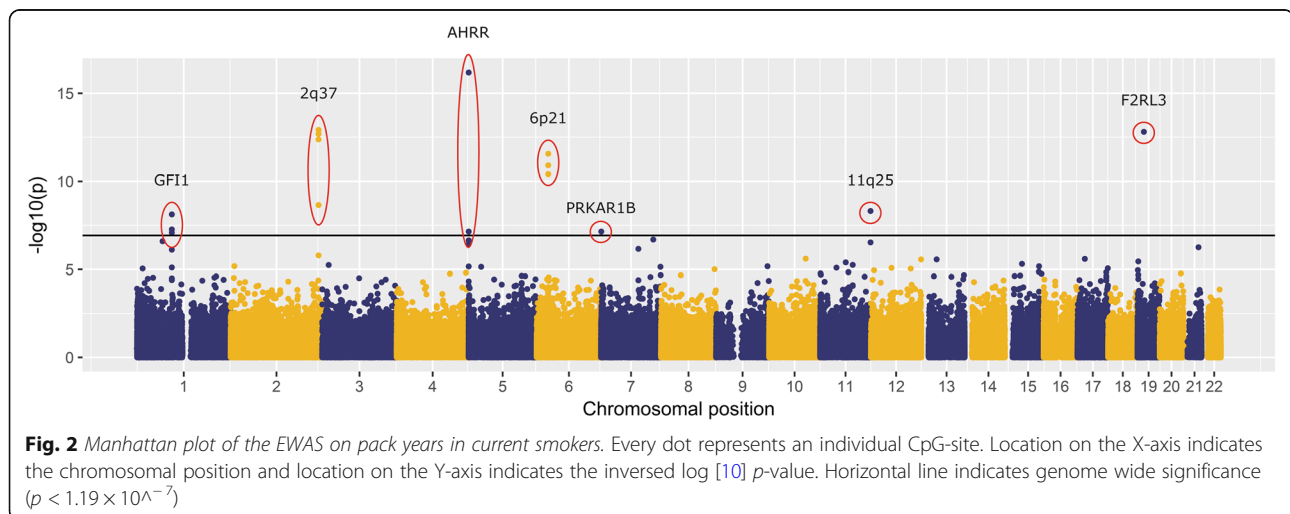
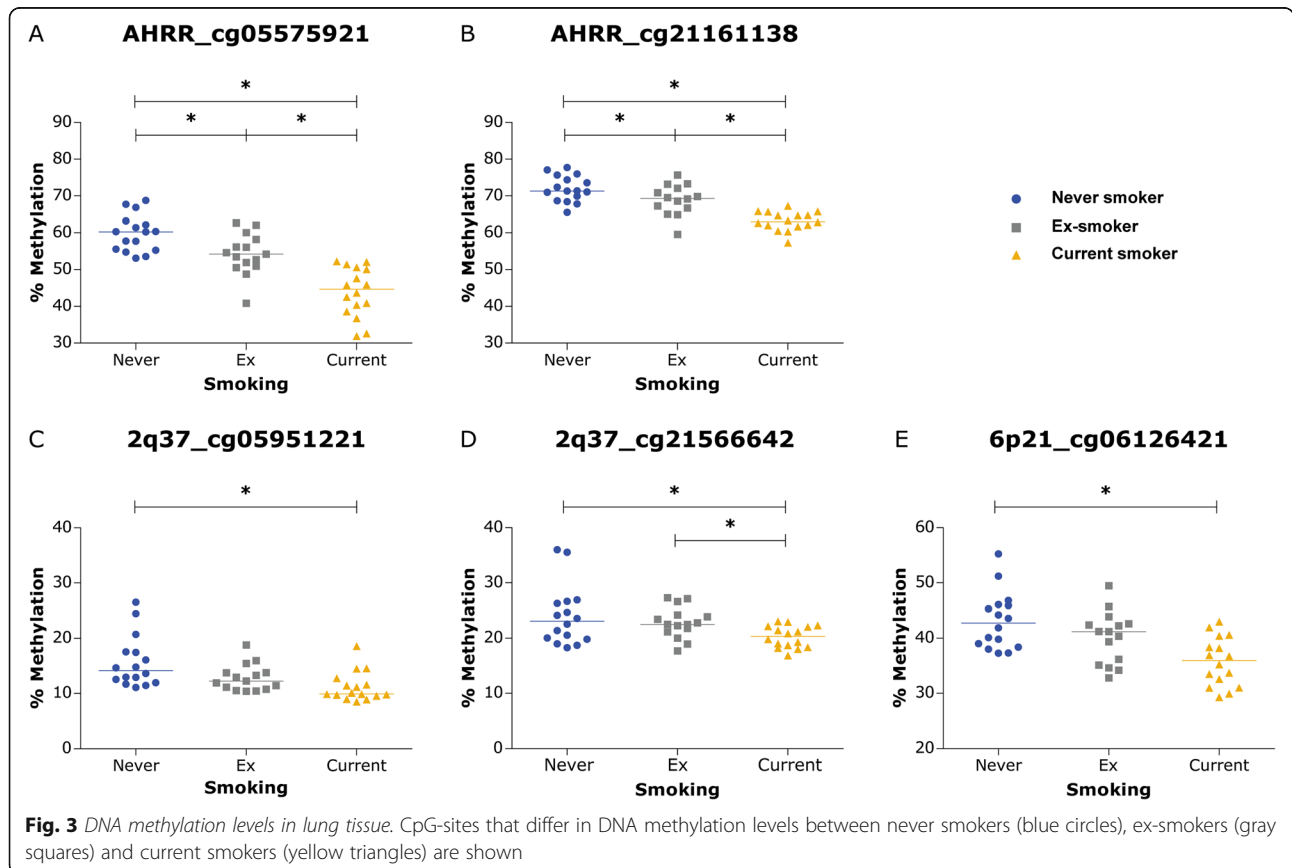


Table 2 Genome-wide significant results of EWAS in whole blood

CpG-site	B	SE	P-value	Adj. P-value	Gene annotation	Gene function [#]
cg05575921*	-0.00360	0.00043	6.80E-17	2.86E-11	AHRR	Repressor of AHR, important in dioxin toxicity and involved in regulation of cell growth and differentiation
cg21566642*	-0.00176	0.00024	1.21E-13	5.09E-08	NA	NA
cg03636183*	-0.00211	0.00029	1.55E-13	6.53E-08	F2RL3	Plays a role in platelet activation and hypomethylation may be associated with human lung cancer
cg05951221*	-0.00130	0.00018	2.01E-13	8.48E-08	NA	NA
cg01940273*	-0.00143	0.00020	4.20E-13	1.77E-07	NA	NA
cg06126421*	-0.00230	0.00033	2.75E-12	1.16E-06	NA	NA
cg15342087*	-0.00092	0.00014	1.23E-11	5.19E-06	NA	NA
cg24859433*	-0.00089	0.00014	3.91E-11	1.65E-05	NA	NA
cg27241845*	-0.00110	0.00018	2.21E-09	9.30E-04	NA	NA
cg20451986	0.00104	0.00018	4.85E-09	2.04E-03	NA	NA
cg12876356*	-0.00244	0.00042	7.55E-09	3.18E-03	GFI1	Transcription repressor essential for hematopoiesis
cg09935388*	-0.00218	0.00040	5.42E-08	2.28E-02	GFI1	Transcription repressor essential for hematopoiesis
cg21161138*	-0.00106	0.00020	7.10E-08	2.99E-02	AHRR	Repressor of AHR, important in dioxin toxicity and involved in regulation of cell growth and differentiation
cg22994830	0.00102	0.00019	7.17E-08	3.02E-02	PRKAR1B	Involved in cAMP signaling in cells
cg18146737*	-0.00256	0.00048	8.19E-08	3.45E-02	GFI1	Transcription repressor essential for hematopoiesis

*Associated with smoking status in previous studies

[#]Gene function obtained by www.genecards.org



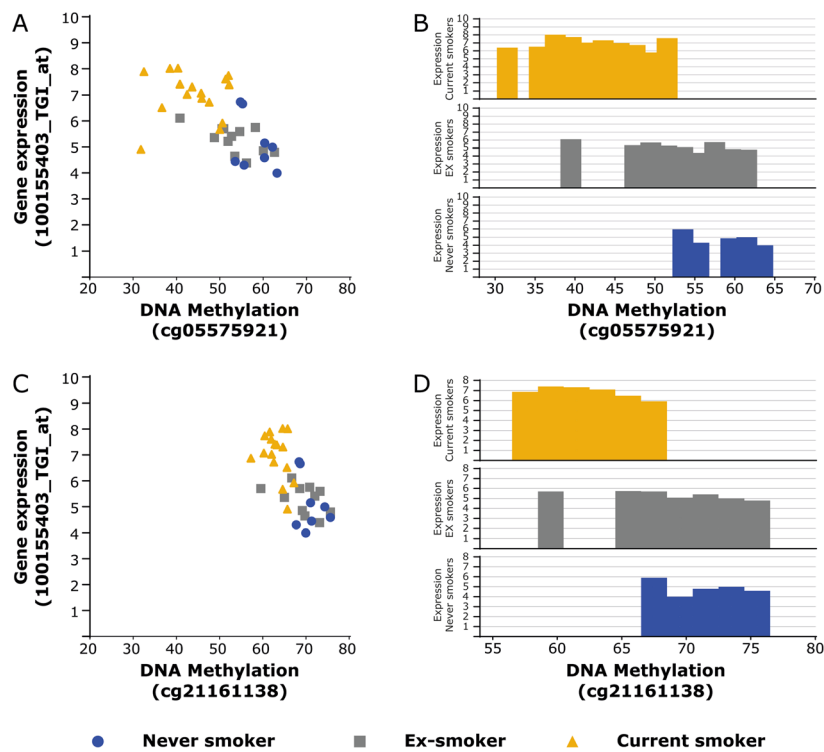


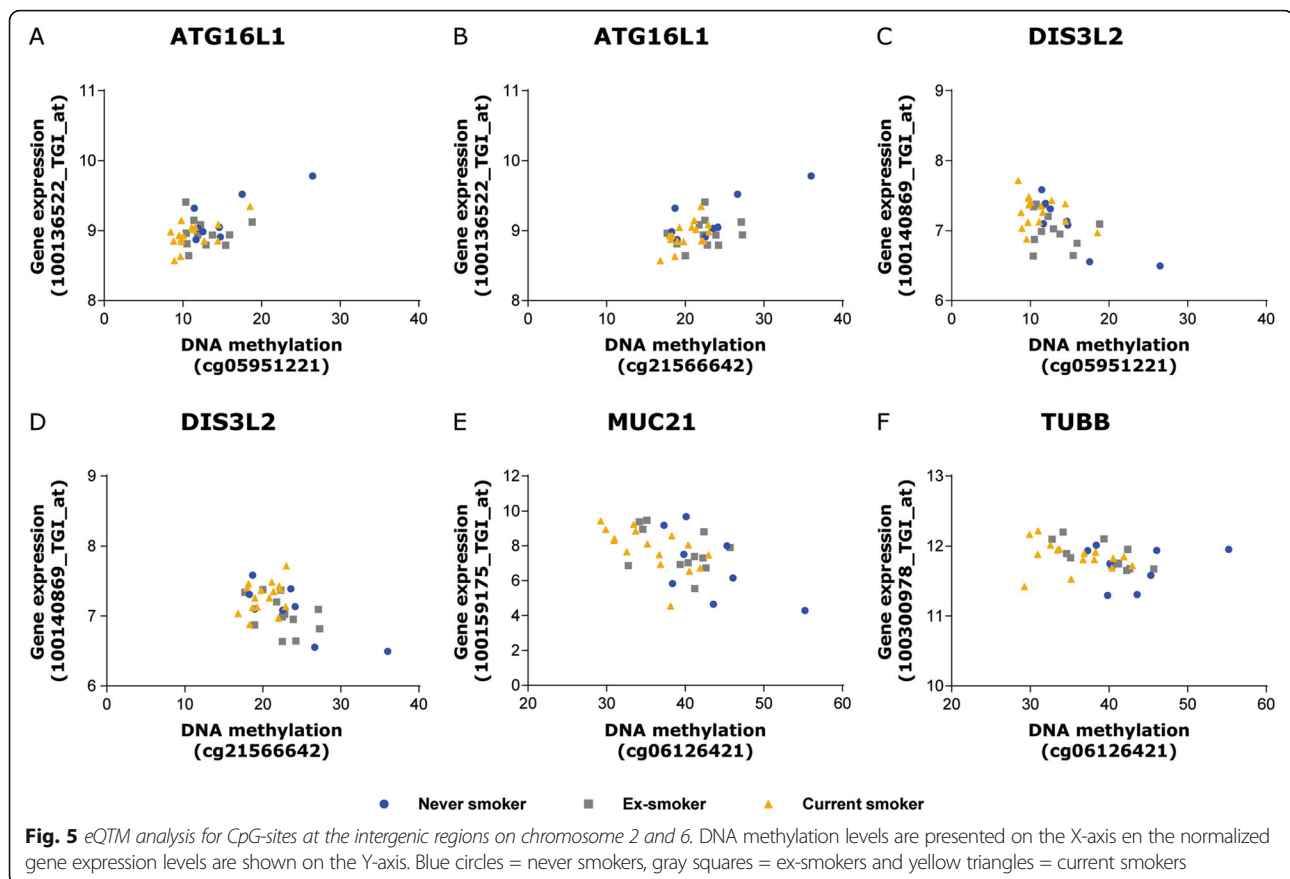
Fig. 4 Association between DNA methylation levels and gene expression of *AHRR* in lung tissue. **a** and **c**) DNA methylation levels are presented on the X-axis and the normalized gene expression levels are shown on the Y-axis. Blue circles = never smokers, gray squares = ex-smokers and yellow triangles = current smoker. **b** and **d**) DNA methylation levels are presented on the X-axis and the average of the normalized gene expression levels distributed per 2% DNA methylation are shown on the Y-axis. The upper part with yellow bars represents current smokers, the middle part with gray bars represents ex-smokers and the lower part with blue bars represent never smokers

and cg20451986, located in the intergenic region of chromosome 11. 10 CpG-sites were additionally associated with lung function levels and we validated 5 of these CpG-sites in lung tissue. We found several significant associations between DNA methylation and gene expression in lung tissue. Moreover, we found novel associations for the CpG-sites located in the intergenic regions of chromosome 2 and 6 and biological plausible genes in lung tissue.

In our EWA study, we identified 2 novel CpG-sites associated with pack years. One of them, cg22994830, is located in the body of *PRKAR1B*. As a regulatory subunit of cyclic AMP-dependent protein kinase A, this gene is involved in several cellular events including ion transport, metabolism and transcription. *PRKAR1B* has shown to be implicated in neurodegenerative disorders, however, a role for *PRKAR1B* in pulmonary diseases is currently unknown [25]. The other novel CpG-site cg20451986 has not been annotated to a gene yet. The gene that is closest to this CpG-site, at approximately 10,000 base pairs, is *JAM3*, a gene involved in cell-cell adhesion. While we performed an EWA study on the association between DNA methylation and pack years as a

cumulative measurement of exposure to cigarette smoke, as reviewed by Gao et al, earlier studies investigated differences in DNA methylation levels between never and current smokers [22]. Nevertheless, we identified several CpG-sites that have been described before. Our third most significant CpG-site cg03636183 located in the gene *F2RL3*, was one of the first sites that was found to be associated with smoking status [23]. In addition, a study by Zeilinger et al showed associations of CpG-sites located in the genes *AHRR*, *GFI1*, *F2RL3* and the unknown intergenic regions at chromosome 2 and 6 with smoking status [12]. Nine of our identified CpG-sites were identical to the CpG-sites discovered by Zeilinger's study.

A major strength of our study is that we validated our findings in lung tissue, the actual tissue of interest. This is in contrast to previous studies in which the most significant CpG-sites from the EWAS were validated by replication in whole blood of other populations [12, 13, 23]. In our study, DNA methylation at one of the CpG-sites located in the body of *AHRR*, cg21161138, was significantly lower in current smokers compared to never smokers in lung tissue, in line with previous findings by Shenker et al [24]. In addition, the association



between DNA methylation of cg21161138 and lower gene expression of *AHRR* in lung tissue also confirms findings of that earlier study. However, we were able to validate 4 additional CpG-sites in lung tissue. The fact that we could confirm a total of 5 CpG-sites, identified in whole blood, in lung tissue suggests that changes in DNA methylation in response to cigarette smoking directly occur at the local level in the lung. For cg03636183 located in the body of *F2RL3*, we neither found differences upon smoking in lung tissue nor an association between DNA methylation levels and gene expression. Since the intensities of the nucleotide peaks in the pyrosequence run of cg03636183 were low compared to other assays, despite the use of several different primer sets, we cannot exclude that actual differences in DNA methylation in lung tissue are masked by technical issues.

Several cross-sectional studies strongly suggest that DNA methylation in whole blood can be partially normalized upon smoking cessation [21, 26, 27]. In our study, we found that DNA methylation levels of ex-smokers are between the levels of never and current smokers for the 5 CpG-sites that are differentially methylated upon smoking. Moreover, with the eQTM analysis, we showed that for

AHRR both DNA methylation and gene expression levels of ex-smokers tends more towards the levels observed for never smokers compared to current smokers. Even though this is cross-sectional data, it implies a reversible character of DNA methylation.

After identification and validation of CpG-sites that are associated with exposure to cigarette smoke in lung tissue, an important next step is to assess the functional relevance of the CpG-sites. For *AHRR*, the function is well studied. Briefly, *AHRR* inhibits the aryl hydrocarbon receptor pathway, which is involved in the removal of harmful environmental chemicals, including cigarette smoke. Cigarette smoking results in decreased DNA methylation and this decreased DNA methylation subsequently increases the expression of *AHRR*. This leads to reduced removal of harmful compounds and thereby thus potentially increasing the damage caused by these compounds [28]. In contrast, the other 3 CpG-sites that are different with smoking status in lung tissue have not yet been annotated to a gene, making it impossible to assess the potential function of DNA methylation at these sites. By performing an eQTM analysis, we tried to identify the genes that are regulated by these CpG-sites. For the 2 CpG-sites at chromosome 2, we found associations

with gene expression of *ATG16L1* and *DIS3L2*. *ATG16L1* is an essential component of the autophagy pathway and mutations in this gene have been associated with inflammatory bowel disease. Since proper function of *ATG16L1* is necessary for host-defense responses against micro-organisms and inflammatory responses in Crohn's disease, this gene might be relevant in the respiratory system as well [29]. *DIS3L2* belongs to the family of exo-ribonucleases, key enzymes involved in the control of messenger RNA stability. *DIS3L2* has been associated with human diseases such as Perlman syndrome and Wilm's tumor, but a role in pulmonary diseases has not been described [30]. For cg06126421 located at the intergenic region of chromosome 6, we found an association between DNA methylation levels and the expression of the genes *TUBB* and *MUC21*. *TUBB* encodes a beta tubulin protein of the microtubule cytoskeleton and has shown to be involved in microcephaly with structural brain abnormalities in humans. A role for *TUBB* in lung-related pathologies, however, is currently unknown [31]. *MUC21* belongs to group of 22 mucin proteins, the major glycoprotein components of mucus, which forms the protective layer of the epithelial surface. Since overproduction of mucins is associated with common respiratory diseases including COPD, asthma and cystic fibrosis [32], *MUC21* is a biological plausible gene and further investigations into the inverse association between DNA methylation at cg06126421 and the expression of *MUC21* are warranted.

While our study is one of the first studies to validate a large panel of CpG-sites in actual lung tissue, we have to consider some limitations of our study. First, the results of the mediation analysis should be interpreted with caution. In our study, we used mediation analysis to select CpG-sites that were associated with lung function levels rather than implying biological mediation. We used pack years as cumulative measurement of the exposure to cigarette smoke to assess the association between cigarette smoking and DNA methylation. However, the number of pack years is estimated from self-reported information obtained from questionnaires of the LifeLines population-based cohort study. It has been suggested that self-reported estimations of smoking may underestimate the true smoking prevalence, since cigarette smoking is often interpreted as socially undesirable behavior [33]. Moreover, it has been stated that DNA methylation levels at specific CpG-sites in the genes *AHRR* and *F2RL3* are a better estimate for the exposure to cigarette smoke than pack years derived from self-report [34]. Within the mediation analysis, this potential misclassification of exposure to cigarette smoke may lead to overestimation of the mediation effect [35]. Furthermore, with the cross-sectional design of our study, we cannot infer causality from the mediation analysis. A second potential limitation of our study is that most of the lung tissue was obtained from

tumor resection surgery. Although all the tissue has been histologically checked for abnormalities, DNA methylation might be affected by the tumor and potential metastases. However, since this holds for all groups under study, we assume that it will not lead to differences in DNA methylation between the groups.

Conclusions

We identified 15 CpG-sites that were associated with pack years, of which 2 CpG-sites were novel. From the 10 CpG-sites that were also associated with lung function levels, 5 CpG-sites could be validated in lung tissue. Significant associations with expression levels of biological plausible genes suggest a functional role for DNA methylation in lung tissue. Overall, our study showed for the first time that a panel of CpG-sites associated with cigarette smoking and lung function levels identified in whole blood is also associated with exposure to cigarette smoke in the tissue of interest, i.e. the lung. Further research should reveal the functional relevance of these CpG-sites in the lung and their role in the etiology of COPD.

Additional files

- Additional file 1:** Online supplement. (DOCX 158 kb)
- Additional file 2:** Overview of the results of the mediation analysis. (DOCX 23 kb)
- Additional file 3:** Subject characteristics lung tissue cohort. (DOCX 17 kb)
- Additional file 4:** Overview results of the association between DNA methylation and smoking status in lung tissue. (DOCX 19 kb)
- Additional file 5:** Overview results of the eQTM analysis. (PDF 553 kb)

Acknowledgements

Part of the data presented in this manuscript is from the Merck-Laval-UBC-Groningen Lung eQTL consortium. We thank the principal investigators of this consortium including Peter Paré, Wim Timens, Yohan Bossé, Maarten van den Berge, Don D. Sin, Philippe Joubert, Ke Hao, and David Nickle.

Funding

This work was supported by consortium grant number 4.1.13.007 of the Lung foundation Netherlands and a research grant of Stichting Astmabestrijding (2017/015). The LifeLines initiative has been made possible by funds from FES (Fonds Economische Structuurversterking), SNN (Samenwerkingsverband Noord Nederland) and REP (Ruimtelijk Economisch Programma).

Availability of data and materials

The datasets and/or analyzed during the current study are available from the corresponding authors on reasonable request.

Authors contributions

M.d.V., D.A.v.d.P., I.N., N.A., C.M.v.D., D.S.P., C.A.B., C.C.v.D., J.M.V. and H.M.B. were involved in the conception and design of the research; M.d.V., D.A.v.d.P., R.N.V.S. and W.K. performed the analyses and experiments; M.d.V. and H.M.B. interpreted the results. M.d.V. prepared figures and drafted the manuscript. H.M.B. critically reviewed and revised the manuscript; all authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

LifeLines Groningen.

Written informed consents was provided by all included subjects and the study was approved by the Medical Ethics Committee of the University Medical Center Groningen (2007/152).

Lung tissue cohort.

Lung tissue was collected from patients undergoing lung resection, lung volume reduction or transplant surgery in Groningen (The Netherlands) in accordance with the Research Code of the University Medical Center Groningen and Dutch national ethical and professional guidelines.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹University of Groningen, University Medical Center Groningen, Department of Epidemiology, Hanzplein 1, Groningen 9713, GZ, The Netherlands.

²University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD (GRIAC), Groningen, The Netherlands. ³Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands. ⁴University of Groningen, University Medical Center Groningen, Department of Obstetrics and Gynaecology, Groningen, The Netherlands. ⁵University of Groningen, University Medical Center Groningen, Department of Pathology and Medical Biology, Groningen, The Netherlands. ⁶University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands.

Received: 29 August 2018 Accepted: 28 September 2018

Published online: 03 November 2018

References

- From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2015. Available from: <http://www.goldcopd.org/>.
- Adeloye D, Chua S, Lee C, Basquill C, Papan A, Theodoratou E, et al. Global and regional estimates of COPD prevalence: systematic review and meta-analysis. *J Glob Health*. 2015;5(2):020415.
- van der Plaats DA, de Jong K, Lahousse L, Faiz A, Vonk JM, van Diemen CC, et al. Genome-wide association study on the FEV1/FVC ratio in never-smokers identifies HHIP and FAM13A. *J Allergy Clin Immunol*. 2016;6.
- Bosse Y. Updates on the COPD gene list. *Int J Chron Obstruct Pulmon Dis*. 2012;7:607–31.
- Hobbs BD, de Jong K, Lamontagne M, Bosse Y, Shrine N, Artigas MS, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet*. 2017;49(3):426–32.
- de Jong K, Boezen HM, Kromhout H, Vermeulen R, Postma DS, Vonk JM. Association of occupational pesticide exposure with accelerated longitudinal decline in lung function. *Am J Epidemiol*. 2014;179(11):1323–30.
- de Jong K, Vonk JM, Timens W, Bosse Y, Sin DD, Hao K, et al. Genome-wide interaction study of gene-by-occupational exposure and effects on FEV1 levels. *J Allergy Clin Immunol*. 2015;136(6):1664–72.e1–14.
- Kheirallah AK, Miller S, Hall IP, Sayers I. Translating lung function genome-wide association study (GWAS) findings. *New Insights for Lung Biology Adv Genet*. 2016;93:57–145.
- Madrigano J, Baccarelli A, Mittleman MA, Wright RO, Sparrow D, Vokonas PS, et al. Prolonged exposure to particulate pollution, genes associated with glutathione pathways, and DNA methylation in a cohort of older men. *Environ Health Perspect*. 2011;119(7):977–82.
- Yang IV, Lozupone CA, Schwartz DA. The environment, epigenome, and asthma. *J Allergy Clin Immunol*. 2017;140(1):14–23.
- Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484–92.
- Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One*. 2013;8(5):e63812.
- Qiu W, Baccarelli A, Carey VJ, Boutaoui N, Bacherman H, Klanderman B, et al. Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function. *Am J Respir Crit Care Med*. 2012;185(4):373–81.
- Panni T, Mehta AJ, Schwartz JD, Baccarelli AA, Just AC, Wolf K, et al. Genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. *Environ Health Perspect*. 2016;124(7):983–90.
- Kohansal R, Martinez-Cambor P, Agusti A, Buist AS, Mannino DM, Soriano JB. The natural history of chronic airflow obstruction revisited: an analysis of the Framingham offspring cohort. *Am J Respir Crit Care Med*. 2009;180(1):3–10.
- Mohamed Hoesein FA, Zanen P, Boezen HM, Groen HJ, van Ginneken B, de Jong PA, et al. Lung function decline in male heavy smokers relates to baseline airflow obstruction severity. *Chest*. 2012;142(6):1530–8.
- Scholten S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol*. 2015;44(4):1172–80.
- van der Plaats DA, de Jong K, de Vries M, van Diemen CC, Nedeljkovic I, Amin N, et al. Occupational exposure to pesticides is associated with differential DNA methylation. *Occup Environ Med*. 2018;75(6):427–35.
- Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal mediation analysis. *J Stat Softw*. 2014;59(5).
- Hao K, Bosse Y, Nickle DC, Pare PD, Postma DS, Laviolette M, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet*. 2012;8(11):e1003029.
- Wan ES, Qiu W, Baccarelli A, Carey VJ, Bacherman H, Rennard SI, et al. Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum Mol Genet*. 2012;21(13):3073–82.
- 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;7:113–015–0148-3 eCollection 2015.
- Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet*. 2011;88(4):450–7.
- Shenker NS, Polidoro S, van Veldhoven K, Sacerdote C, Ricceri F, Birrell MA, et al. Epigenome-wide association study in the European prospective investigation into Cancer and nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Hum Mol Genet*. 2013;22(5):843–51.
- Wong TH, Chiu WZ, Breedveld GJ, Li KW, Verkerk AJ, Hondius D, et al. PRKAR1B mutation associated with a new neurodegenerative disorder with unique pathology. *Brain*. 2014;137(Pt 5):1361–73.
- Guida F, Sandanger TM, Castagne 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(8):2349–59.
- Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J, et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics*. 2014;9(10):1382–96.
- Lee KW, Pausova Z. Cigarette smoking and DNA methylation. *Front Genet*. 2013;4:132.
- Salem M, Ammitzboell M, Nys K, Seidelin JB, Nielsen OH. ATG16L1: a multifunctional susceptibility factor in Crohn disease. *Autophagy*. 2015;11(4):585–94.
- Pashler AL, Towler BP, Jones CI, Newbury SF. The roles of the exoribonucleases DIS3L2 and XRN1 in human disease. *Biochem Soc Trans*. 2016;44(5):1377–84.
- Ngo L, Haas M, Qu Z, Li SS, Zenker J, Teng KS, et al. TUBB5 and its disease-associated mutations influence the terminal differentiation and dendritic spine densities of cerebral cortical neurons. *Hum Mol Genet*. 2014;23(19):5147–58.
- Evans CM, Koo JS. Airway mucus: the good, the bad, the sticky. *Pharmacol Ther*. 2009;121(3):332–48.
- Connor Gorber S, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. *Nicotine Tob Res*. 2009;11(1):12–24.
- Lee DH, Hwang SH, Lim MK, Oh JK, Song DY, Yun EH, et al. Performance of urine cotinine and hypomethylation of AHRR and F2RL3 as biomarkers for smoking exposure in a population-based cohort. *PLoS One*. 2017;12(4):e0176783.
- Valeri L, Reese SL, Zhao S, Page CM, Nystad W, Coull BA, et al. Misclassified exposure in epigenetic mediation analyses. Does DNA methylation mediate effects of smoking on birthweight? *Epigenomics*. 2017;9(3):253–65.