

# No association between DNA methylation and COPD in never and current smokers

Maaïke de Vries,<sup>1,2</sup> Diana A van der Plaats,<sup>1,2</sup> Judith M Vonk,<sup>1,2</sup> H Marïke Boezen<sup>1,2</sup>

**To cite:** de Vries M, van der Plaats DA, Vonk JM, *et al*. No association between DNA methylation and COPD in never and current smokers. *BMJ Open Resp Res* 2018;**5**:e000282. doi:10.1136/bmjresp-2018-000282

Received 6 February 2018  
Revised 7 June 2018  
Accepted 9 June 2018

## ABSTRACT

**Introduction** Chronic obstructive pulmonary disease (COPD) is a progressive inflammatory lung disease with cigarette smoke as the main risk factor for its development. Since not every smoker develops COPD, other factors likely underlie differences in susceptibility to develop COPD. Here, we tested if DNA methylation may be such a factor by assessing the association between DNA methylation levels and COPD in never and current smokers from the general population.

**Methods** For the current study, 1561 subjects were non-randomly selected from the LifeLines cohort study. We included 903 never smokers and 658 current smokers with and without COPD, defined as pre-bronchodilator forced expiratory volume in 1 s/forced vital capacity (FEV<sub>1</sub>/FVC) <70%. Subsequently, we performed robust regression analysis on whole blood DNA methylation levels of 420 938 CpG sites with COPD as outcome.

**Results** None of the CpG sites in both the never and the current smokers were genome-wide significantly associated with COPD. CpG site cg14972228 annotated to *SIPAL3* was most significant ( $p=5.66 \times 10^{-6}$ ) in the never smokers, while CpG site cg08282037 annotated to *EPS8L1* was most significant ( $p=1.45 \times 10^{-5}$ ) in the current smokers.

**Conclusion** In contrast to a previous, smaller study, we did not observe any significant association between DNA methylation levels and the presence of COPD, independent of smoking status. Apparently, DNA methylation studies are highly variable.

## INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a progressive inflammatory lung disease, characterised by persistent airflow limitation. Patients with COPD suffer from severe respiratory symptoms, resulting in a worse quality of life. The development of COPD is associated with both genetic and environmental factors and their interactions. Although exposure to cigarette smoke is the main risk factor for the development of COPD, not every smoker will develop COPD.

The results of the genome-wide association studies<sup>1,2</sup> and genome-wide interaction studies<sup>3</sup> that have been performed so far indicate that only a small part (approximately

## Key messages

- ▶ We did not identify any genome-wide significant association between COPD and DNA methylation in never and current smokers in a large, non-random sample from the general population.
- ▶ The results of our study are in contrast to a previous, smaller study, indicating that studies on DNA methylation might be highly variable.
- ▶ The high variability in DNA methylation studies should be taken into account when interpreting these types of studies.

5.5%) of the so called ‘susceptibility to develop COPD’ can be explained by genetic variation.<sup>4</sup> As a consequence, the epigenome is increasingly recognised as an important link between the inherited genome and environmental exposures. One well-defined epigenetic modification is DNA methylation, which is tissue-specific and involves binding of a methyl group to a cytosine base adjacent to a guanine base, a so-called CpG site, leading to changes in gene expression.<sup>5</sup> Therefore, differences in DNA methylation might partly explain the variance in the susceptibility to develop COPD. In the current study, we assessed whether differences in DNA methylation levels are associated with COPD in a non-randomly selected subset of the general population of never and current smokers.

## METHODS

### Study population

For the current study, 1561 subjects were non-randomly selected from the LifeLines general population-based cohort study.<sup>6,7</sup> The LifeLines cohort study is approved by the medial ethical committee of the University Medical Centrum Groningen, The Netherlands (METc 2007/152) and all participants have signed informed consent. Subjects were non-randomly selected based on smoking history (never smoker or current smoker with >5 pack years), COPD (defined by



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, The Netherlands

<sup>2</sup>University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, Groningen, The Netherlands

### Correspondence to

Dr Maaïke de Vries;  
m.de.vries04@umcg.nl

pre-bronchodilator forced expiratory volume in 1 s/forced vital capacity ( $FEV_1/FVC$ ) <70%) and job-related exposures. Pre-bronchodilator spirometry was performed with a Welch Allyn Version 1.6.0.489, PC-based Spiroperfect with CA Workstation software according to the ATS/ERS guidelines. Technical quality and results were evaluated by well-trained assistants and abnormal results were re-evaluated by a lung physician.

### Measurements

DNA methylation levels in whole blood were determined using the Illumina Infinium Human Methylation 450K Array. After quality control, DNA methylation data presented as beta values were available for 420 938 CpG sites.<sup>8</sup>

### Epigenome-wide association study

To analyse the association between genome-wide DNA methylation and COPD, we performed robust logistic regression analysis with COPD as outcome and DNA methylation as predictor, stratified for smoking status (never vs current). We adjusted for the potential confounders age, gender, pack years, batch effects and white blood cell composition and applied false discovery rate (FDR) correction for multiple testing.

## RESULTS

### Subject characteristics

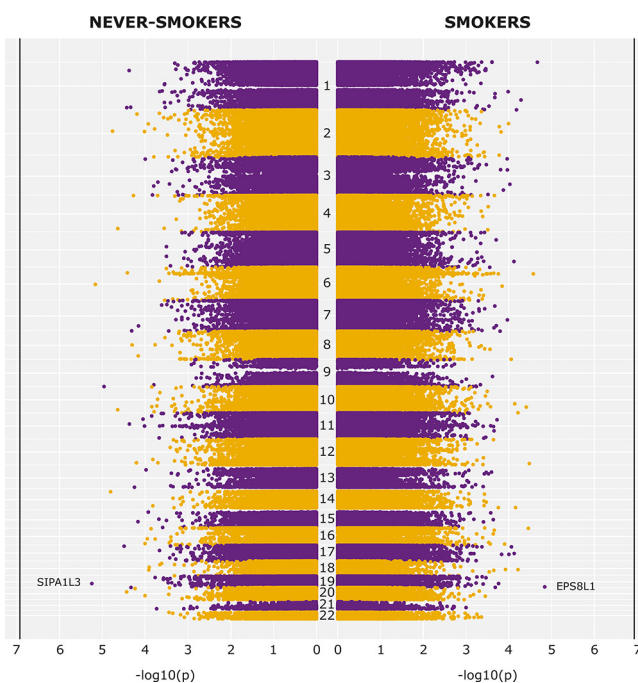
A total of 903 never smokers and 658 current smokers were included in this study. An overview of the subject characteristics is shown in table 1.

Table 1 Subject characteristics	Never smokers	Current smokers
	Number of subjects, N (%)	903 (57.8)
Male, N (%)	508 (56.3)	375 (57.0)
Age (years), median (min–max)	46 (18–80)	46 (22–79)
Pack years (years), mean (min–max)	–	20.6 (5.1–95.0)
COPD ( $FEV_1/FVC$ <70%), N (%)	316 (35.0)	279 (42.4)
Lung function, mean (SD)		
$FEV_1$ (L)	3.5 (0.9)	3.4 (0.9)
$FEV_1$ %predicted (%)*	100.5 (14.5)	94.4 (14.9)
$FEV_1/FVC$ (%)	84.5 (8.2)	71.7 (8.8)
Moderate COPD, N (%) †	56 (6.2)	100 (15.2)

\*Calculated with GLI-2012 if possible.

†COPD GOLD stage >2 ( $FEV_1/FVC$  <70% and  $FEV_1$  between 50% and 80% of predicted).

COPD, chronic obstructive pulmonary disease;  $FEV_1$ , forced expiratory volume in 1 s; FVC, forced vital capacity.



**Figure 1** Manhattan plot for the association between COPD and DNA methylation in never smokers (left) and current smokers (right). Robust logistic regression analysis was performed, adjusted for age, gender, pack years, batch effects and white blood cell composition. Vertical line indicates genome wide significance ( $p < 1.19 \times 10^{-7}$ ). COPD, chronic obstructive pulmonary disease.

### Epigenome-wide association study

In both never and current smokers, none of the CpG sites were genome-wide significantly associated with COPD (see Manhattan plot in figure 1). In never smokers, CpG site cg14972228 on chromosome 19, annotated to *Signal-Induced Proliferation-Associated 1-Like Protein 3 (SIPA1L3)* and involved in the GTP-ase activating cascade, was most significant ( $p = 5.66 \times 10^{-6}$ ). While *SIPA1L3* has been associated with epithelial cell morphogenesis, polarity and cytoskeletal organisation in eye abnormalities, a similar role for *SIPA1L3* in lung-related tissues has not yet been described.<sup>9</sup> In current smokers, CpG site cg08282037 on chromosome 19, annotated to *Epidermal Growth Factor Receptor Pathway Substrate 8-Related Protein 1 (EPS8L1)*, was most significant ( $p = 1.45 \times 10^{-5}$ ). It has been postulated that *EPS8L1* can link growth factor stimulation to actin reorganisation, thereby playing a role in the regulation of actin cytoskeletal remodelling.<sup>10</sup> However, the function of *EPS8L1* in the respiratory system is currently unknown.

## DISCUSSION

To our knowledge, this is the first genome-wide methylation study with the main focus on COPD in a general population-based sample of never and current smokers.<sup>11</sup> The only other study exploring the association between genome-wide DNA methylation and COPD so far<sup>12</sup> focused on a much more severe COPD phenotype than the common COPD

phenotype we have studied. Our results are in contrast with those from *Qui et al*, which described a strong association between DNA methylation and COPD. Several reasons can be put forward as a potential explanation why the outcome of these two, on first sight, comparable studies is so different.

First, *Qiu et al* used the Illumina 27K array to determine the DNA methylation levels, while we used the Illumina 450K array. Of the CpG sites on the 27K array, 23 787 CpG sites were also present among the 420 938 CpG sites that passed quality control in our analysis, but all failed to reach significance (after applying FDR correction for multiple testing for 23 787 CpG sites as used by *Qui et al*). *Qui et al* identified 349 CpG sites that were significant for their three selected phenotypes COPD, FEV<sub>1</sub> and FEV<sub>1</sub>/FVC and replicated at  $p < 10^{-3}$  in their replication cohort. All these 349 CpG sites were available in our study too. Hence, different Illumina platforms cannot explain the differences in the results of both studies.

Second, we adjusted for age, gender, pack years, batch effects and white blood cell composition while *Qui et al* decided to avoid potential over adjustment for age and sex, arguing that these confounders are already included in the definition of COPD. Nevertheless, they show that in their covariate-adjusted model, the association between COPD and DNA methylation remains highly significant. However, the p-values are considerably larger than in the unadjusted models and none of the CpG sites were significantly replicated in their replication cohort. Interestingly, an unadjusted analysis in our cohort revealed several genome-wide significant CpG sites (data not shown), but did not replicate any of their top hits. Since *Qiu et al* did not adjust for white blood cell composition, we could not determine if differences in white blood composition might explain the differences between both studies. Overall, adjustments for confounders definitively affect the results of the studies, but it is not clear if the adjustments can be held completely responsible for the observed differences between the studies.

Third, our cohort is a selected sample from the general population including subjects with mild to moderate severe COPD and controls. In contrast, *Qiu et al* included only patients with severe COPD based on a strict cut-off of FEV<sub>1</sub>% predicted of 70% together with a FEV<sub>1</sub>/FVC <70%. Furthermore, in our cohort only never smokers and current smokers are included, while ex-smokers are not present. These differences in smoking habits and phenotype between both cohorts might explain the differences between the results of both analyses to some extent.

Finally, both studies have a cross-sectional design and we cannot rule out the fact that COPD may also cause differential DNA methylation itself. Together with the different phenotypes of COPD in both studies, this might contribute to the observed differences between the studies. With respect to our study, some additional cohort specific characteristics have to be emphasised. Our DNA methylation cohort is a non-random selection of the LifeLines Cohort Study. Since COPD was one of the selection criteria, the percentages of COPD cases should not be interpreted as prevalence.

Furthermore, COPD was defined as a fixed pre-bronchodilator FEV<sub>1</sub>/FVC ratio below 70%. While this definition of COPD is commonly used, it is known that this definition can overestimate the number of COPD cases, especially in older subjects.<sup>13</sup>

In conclusion, we did not observe any genome-wide significant association between DNA methylation levels and the presence of COPD, independent of smoking status. Results of DNA methylation studies appear to be highly variable, which should be taken into account for the interpretation of future studies.

**Contributors** MdV and HMB: conception and design of the research. MdV and DAvdP: performed the analyses; MdV and HMB: interpreted the results. MdV: prepared the figure and drafted the manuscript. HMB and JMV: critically reviewed and revised the manuscript. All authors read and approved the final version of the manuscript.

**Funding** This work was supported by consortium grant 4.1.13.007 of the Lung Foundation Netherlands.

**Competing interests** None declared.

**Patient consent** Not required.

**Ethics approval** METc University Medical Center Groningen.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## REFERENCES

- Hobbs BD, de Jong K, Lamontagne M, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet* 2017;49:426–32.
- van der Plaats DA, de Jong K, Lahousse L, et al. Genome-wide association study on the FEV<sub>1</sub>/FVC ratio in never-smokers identifies HHIP and FAM13A. *J Allergy Clin Immunol* 2017;139.
- de Jong K, Vonk JM, Timens W, et al. Genome-wide interaction study of gene-by-occupational exposure and effects on FEV1 levels. *J Allergy Clin Immunol* 2015;136:1664–72.
- Kheirallah AK, Miller S, Hall IP, et al. Translating Lung Function Genome-Wide Association Study (GWAS) Findings: New Insights for Lung Biology. *Adv Genet* 2016;93:57–145.
- Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012;13:484–92.
- Stolk RP, Rosmalen JG, Postma DS, et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *Eur J Epidemiol* 2008;23:67–74.
- Scholtens S, Smidt N, Swertz MA, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int J Epidemiol* 2015;44:1172–80.
- van der Plaats DA, de Jong K, de Vries M, et al. Occupational exposure to pesticides is associated with differential DNA methylation. *Occup Environ Med* 2018;75:427–35.
- Greenlees R, Mihelec M, Yousoof S, et al. Mutations in SIPA1L3 cause eye defects through disruption of cell polarity and cytoskeleton organization. *Hum Mol Genet* 2015;24:5789–804.
- Offenhäuser N, Borgonovo A, Disanza A, et al. The eps8 family of proteins links growth factor stimulation to actin reorganization generating functional redundancy in the Ras/Rac pathway. *Mol Biol Cell* 2004;15:91–8.
- Machin M, Amaral AF, Wielscher M, et al. Systematic review of lung function and COPD with peripheral blood DNA methylation in population based studies. *BMC Pulm Med* 2017;17:54,017–397.
- Qiu W, Baccarelli A, Carey VJ, et al. Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function. *Am J Respir Crit Care Med* 2012;185:373–81.
- Celli BR, Halbert RJ, Isonaka S, et al. Population impact of different definitions of airway obstruction. *Eur Respir J* 2003;22:268–73.