OPEN ACCESS

Original research

DNA methylation is associated with airflow obstruction in patients living with HIV

Ana I Hernandez Cordero,[1] Chen Xi Yang,[1] Maen Obeidat,[1] Julia Yang,[1] Julie MacIsaac,[2] Lisa McEwen,[2] David Lin,[2] Michael Kobor,[2] Richard Novak,[3] Fleur Hudson,[4] Hartwig Klinker,[5] Nila Dharan,[6] SF Paul Man,[1] Don D Sin,[1] Ken Kunisaki ,[7] Janice Leung,[1] on behalf of the INSIGHT START Pulmonary and Genomic Substudy Groups

[1]Centre for Heart Lung Innovation, The University of British Columbia, Vancouver, British Columbia, Canada
[2]Center for Molecular Medicine and Therapeutics, The University of British Columbia, Vancouver, British Columbia, Canada
[3]Section of Infectious Diseases, University of Illinois at Chicago, Chicago, Illinois, USA
[4]MRC Clinical Trials Unit, University College London, London, UK
[5]Medizinische Klinik und Poliklinik II, Universitatsklinikum Wurzburg, Wurzburg, Germany
[6]The Kirby Institute, UNSW Sydney, Sydney, New South Wales, Australia
[7]Minneapolis Veterans Affairs Health Care System, Section of Pulmonary, Critical Care and Sleep Medicine and the Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Department of Medicine, University of Minnesota, Minneapolis, Minnesota, USA

Correspondence to
Dr Janice Leung, Centre for Heart Lung Innovation, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada;
Janice.Leung@hli.ubc.ca

Received 27 July 2020
Revised 20 November 2020
Accepted 23 November 2020
Published Online First 18 December 2020

To cite: Hernandez Cordero AI, Yang CX, Obeidat M, *et al*. *Thorax* 2021;**76**:448–455.

ABSTRACT

Introduction People living with HIV (PLWH) suffer from age-related comorbidities such as COPD. The processes responsible for reduced lung function in PLWH are largely unknown. We performed an epigenome-wide association study to investigate whether blood DNA methylation is associated with impaired lung function in PLWH.

Methods Using blood DNA methylation profiles from 161 PLWH, we tested the effect of methylation on FEV_1, FEV_1/FVC ratio and FEV_1 decline over a median of 5 years. We evaluated the global methylation of PLWH with airflow obstruction by testing the differential methylation of transposable elements Alu and LINE-1, a well-described marker of epigenetic ageing.

Results Airflow obstruction as defined by a FEV_1/FVC<0.70 was associated with 1393 differentially methylated positions (DMPs), while 4676 were associated with airflow obstruction based on the FEV_1/FVC<lower limit of normal. These DMPs were enriched for biological pathways associated with chronic viral infections. The airflow obstruction group was globally hypomethylated compared with those without airflow obstruction. 103 and 7112 DMPs were associated with FEV_1 and FEV_1/FVC, respectively. No positions were associated with FEV_1 decline.

Conclusion A large number of DMPs were associated with airflow obstruction and lung function in a unique cohort of PLWH. Airflow obstruction in even relatively young PLWH is associated with global hypomethylation, suggesting advanced epigenetic ageing compared with those with normal lung function. The disturbance of the epigenetic regulation of key genes not previously identified in non-HIV COPD cohorts could explain the unique risk of COPD in PLWH.

Key messages

What is the key question?
- What explains the increased risk of COPD in patients living with HIV?

What is the bottom line?
- Peripheral blood methylation disruptions are numerous in patients with both airflow obstruction and HIV infection.

Why read on?
- Epigenetic disturbance related to chronic viral infections may hold clues to early disease-causing mechanisms in COPD.

INTRODUCTION

The progress in the treatment of HIV has led to an increase in life expectancy and a decrease in immunodeficiency syndrome-related conditions among people living with HIV (PLWH).[1] Age-related comorbidities, though, have become common, including COPD,[2] which is associated with increased mortality[3] and significant respiratory symptom burden.[4] Whether tobacco exposure, illicit drug use, repeated infections, or chronic inflammation are the key causes of this increased risk for COPD in PLWH is still unclear.

Lung function decline in PLWH was recently reported in a large, multinational cohort demonstrating that the timing of antiretroviral therapy initiation alone has no influence on the rate of decline.[5] As in any population, the variability of lung function is likely the consequence of complex environmental and genetic factors, as well as their interaction[6]; however, the underlying molecular processes that explain these relationships in PLWH remain elusive. In this study, we explore the possibility that epigenetic alterations such as DNA methylation may influence lung function variability in PLWH. DNA methylation involves the addition of a methyl group to a cytosine base located next to guanine base (CpG site). The methylation of CpG sites in regulatory elements (ie, promoter regions) often results in decreased gene expression[7] and can potentially affect other traits. Environmental factors such as tobacco use and chronic diseases or infections as well as the natural ageing process can all influence DNA methylation. Age-related diseases, for instance, are characterised by genome-wide hypomethylation.[8] The methylation of ubiquitous transposable elements like Alu and LINE-1 is used as markers for global methylation and is thought to play key roles in age-related genomic instability,[9] which may lead to tumorigenesis and senescence.

Previous efforts to understand the effect of DNA methylation on lung function have focused mainly on non-HIV cohorts.[10 11] Evidence from these studies suggests, however, that methylation

may play an important role in lung function and the aetiology of COPD.[12] For this study, we conducted an epigenome-wide association analysis to investigate the relationship of blood DNA methylation with lung function of PLWH. We hypothesised that PLWH with airflow obstruction have a distinct methylation pattern when compared with those with normal lung function. We also hypothesised that differential DNA methylation is associated with lung function decline in PLWH.

METHODS

Study cohort

The study cohort consisted of 161 PLWH over the age of 40 years who were enroled in the genomic and the pulmonary substudies of the Strategic Timing of Antiretroviral Therapy (START, Clinicaltrials.gov NCT00867048) trial, which has been previously described.[13 14] Briefly, this was a multicentre, international, randomised controlled trial designed to compare immediate versus deferred initiation of antiretroviral therapy. The START cohort included adult PLWH with CD4 T cell counts >500 cells/mm^3 and who had not yet been exposed to antiretroviral therapy.[5]

Lung function and filtering criteria

All participants underwent spirometry testing yearly for up to 6 years, according to methods previously described.[5] Participants with three or more spirometric tests were retained for the analysis on FEV_1 decline. In total, 152 participants were retained for FEV_1 decline analysis, while all 161 subjects were included in cross-sectional analyses (online supplemental figure S1). Participants were characterised as having airflow obstruction if the FEV_1/FVC ratio was <0.70. In addition, airflow obstruction was also assessed based on FEV_1/FVC ratio <lower limit of normal (LLN), according to the Global Lung Function Initiative 2012 normative equations.[15]

DNA methylation profiling and quality control

Participants had a whole blood sample drawn at study entry. The DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) was used to extract DNA from this sample. Unmethylated cytosine residues present in the DNA extract were converted to uracil using the EZ DNA Methylation Kit (Zymo, Irvine, California, USA). DNA methylation profiles were obtained using the Illumina Infinium MethylationEPIC BeadChip microarray which interrogates 863 904 CpG sites and covers 95% of all genes and 95% of CpG islands.[16] The ratio of the methylated probe intensity to the overall intensity (β value) was calculated for each CpG and ranged from 0 (all unmethylated) to 1 (all methylated) and then transformed to M-values (log2 ratio of the intensity of the methylated probe to unmethylated CpG probe). CpG probes were filtered based on the detection quality and probes with a detection p>1e−10 were excluded from downstream analyses. In addition, non-CpG, XY-linked, single nucleotide polymorphism (SNP) and cross-hybridisation probes were also removed (839 418 CpGs were retained). Lastly, background correction, normalisation and batch correction were performed using the normal–exponential out-of-band,[17] mixture quantile normalisation[18] and ComBat[19] methods, respectively.

Alu and LINE-1 imputation

Global methylation status can be inferred from the methylation of repetitive and transposable elements along the genome, of which Alu and LINE-1 are among the most abundant.[20] Hypomethylation of these sites is associated with ageing as well as with worse lung function in non-PLWH study cohorts.[21] Alu and LINE-1 sites were imputed using the machine learning tool repetitive element methylation prediction.[22]

Statistical analysis

The cell type proportion of each sample was estimated using the deconvolution method of Houseman *et al*[23] which provides the proportion of CD4 T cells, CD8 T cells, natural killer cells, monocytes, granulocytes (neutrophils+eosinophils) and B cells. The EPISTRUCTURE software[24] was used to infer the ancestry. This software calculates the principal components using CpGs that are highly correlated with SNPs, capturing the genetic variation within a population. Additional covariates were chosen based on the algorithm outlined by Lee *et al*.[25] To identify differentially methylated positions (DMPs) between PLWH with airflow obstruction and with normal lung function, we performed an epigenome-wide association study using a robust linear model (rlm) implemented in the MASS R package (M-estimation),[26] and adjusted as follows:

$$Methylation\ (M\ value) \sim Airflow\ obstruction\ status + Age_{baseline} + Sex + CD8\ T\ cells + CD4\ T\ cells + NK\ cells + B\ cells + Monocytes + Granulocytes + Epistructure\ PC1\ to\ PC5$$

The model presented above was also used to interrogate Alu and LINE-1 methylation sites.

To investigate the effect of methylation over cross-sectional FEV_1, and FEV_1/FVC (at baseline visit) in PLWH we also used rlm adjusted for the following covariates:

$$Lung\ function\ trait \sim Methylation\ (beta\ value) + Sex + Age + Age^2 + Height + Height^2 + Smoking\ status + Smoking\ Pack - Years + CD8\ T\ cells + CD4\ T\ cells + NK\ cells + B\ cells + Monocytes + Granulocytes + Epistructure\ PC1\ to\ PC10$$

The effect of methylation over FEV_1 decline over the course of 6 years was studied using a random coefficient model (lme) implemented in the nlme R package[27]; both a random intercept term and a random slope term were included and the model was adjusted for the following covariates:

$$FEV_{1_{Year1-6}} \sim (Year \times CpGs\ (beta\ value)) + (Year \times Current\ smoking\ status) + (Year \times Former\ smoking\ status) + (Year \times Sex) + (Year \times Age_{year0}) + (Year \times Age^2) + (Year \times height) + (Year \times height^2) + (Year \times Smoking\ Pack - Years) + (Year \times CD4\ T\ cells) + (Year \times CD8\ T\ cells) + (Year \times B\ cells) + (Year \times NK\ cells) + (Year \times Monocytes) + (Year \times Granulocytes) + (Year \times Epistructure\ PC1\ to\ PC10)$$

The DMPs for each model were defined at a false discovery rate (FDR)<0.10. The R package DMRcate[28] was used to identify differentially methylated regions (DMRs), defined with at least three CpGs. The R package clusterProfiler was used to identify Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways that were significantly (FDR<0.10) enriched by genes that corresponded to DMPs. The FEV_1 decline analysis was executed first for all PLWH, then additionally separately by each ethnic group.

RESULTS

Description of the study cohort

The baseline demographic characteristics of the study cohort grouped by airflow obstruction status based on the fixed ratio

Table 1 Summary of the study cohort by airflow obstruction group (FEV_1/FVC<0.70 criteria)*

Status (FEV1/FVC<0.70 criteria)	Normal lung function (n=141)	Airflow obstruction (n=20)	P value†
Age (years±SD)	46 (43–51)	49.5 (43.75–52.75)	0.21
Female, %	9.99	5	0.70
Smoking status			
Current, %	36.88	70	0.01
Former, %	23.4	10	0.25
Never, %	39.72	20	0.14
Pack years	2 (0–15)	23.75 (2.81–32.44)	3.00×10^{-03}
Race			0.11
African, %	22.69	10	
Asian, %	0.71	0	
Caucasian, %	59.57	90	
Hispanic, %	14.18	0	
Other, %	2.84	0	
Baseline characteristics			
BMI	25.10 (23.03–28.04)	22.61 (21.55–24.54)	3.00×10^{-03}
Height, cm	175.26 (170-181)	178 (166.75–182.22)	0.68
Weight, kg	77.7 (69.9–87.5)	70 (64.23–74.55)	8.00×10^{-03}
FEV1, mL	3560 (3050–3950)	2895 (2307.5–3407.5)	1.51×10^{-03}
FEV1 %	95.79 (88.25–103.38)	75.18 (66.88–88.97)	4.99×10^{-06}
FVC, mL	4470 (3880–5060)	4300 (3740–5127.5)	0.97
FVC %	94.53 (87.87–102.73)	89.34 (85.17–108.18)	0.66
FEV1/FVC ratio	0.79 (0.77–0.84)	0.67 (0.61–0.68)	5.05×10^{-13}
CD4 T cells/mm^3	630 (577–742)	660 (589–753)	0.67
HIV RNA viral load, copies/mm^3	20 250 (3851–60 798)	26 100 (7710–92 300)	0.33

*Median and IQRs.
†P values correspond to Mann-Whitney U test (continuous variables) and Fisher's exact test (discrete variables).
BMI, body mass index.

criteria are shown in table 1. Demographic data grouped by airflow obstruction based on the LLN criteria are shown in online supplemental table S1. Fifteen individuals met both airflow obstruction criteria, whereas five met criteria only for FEV_1/FVC<0.70 and one met criteria only for FEV_1/FVC<LLN. Based on Mann-Whitney U tests, the airflow obstruction group had a larger proportion of smokers and lower body mass index ($p<0.05$) than the normal lung function group, irrespective of the criteria used for airflow obstruction characterisation. There was no significant difference in CD4 T cell counts and HIV RNA viral load between the groups.

Airflow obstruction in PLWH is associated with methylation

We identified 1392 DMPs (online supplemental table S2) and 2 DMRs (online supplemental table S3) that were associated with airflow obstruction (based on the FEV_1/FVC ratio <0.70 criteria). Of these, 846 DMPs were hypermethylated in individuals with airflow obstruction, while 546 were hypomethylated. Twenty-eight per cent of DMPs were located in CpG Islands and 46% in genomic region of low CpG density (Open Sea), in addition 21% of DMPs were within promoter regions (online supplemental table S4).

Based on the criteria of FEV_1/FVC ratio <LLN, 4675 DMPs and 9 DMRs were identified (online supplemental tables S2 and S3). We found 2843 hypomethylated DMPs in PLWH experiencing airflow obstruction, while 1832 were hypermethylated. The majority of DMPs were in the Open Sea region (61%), while 14% are in CpG Islands and 11% in promoter regions (online supplemental table S4). Moreover, there were 745 DMPs that overlapped between the analyses by FEV_1/FVC ratio <LLN and FEV_1/FVC ratio <0.70 criteria.

Figure 1 shows the range of statistical significance and methylation beta difference for DMPs distinguishing those with and without airflow obstruction. The most significant DMPs correspond to the genes *HK2*, *HBEGF*, *TAPBP*, *MAD1L1*, *GPR153*, *VGLL4* and *ADCY7* (table 2). The DMPs for airflow obstruction (criteria: FEV_1/FVC<0.70) enriched multiple KEGG pathways including 'Small cell lung cancer', 'Hepatitis B', 'Epstein-Bar virus infection' and 'Human Papillomavirus infection'. The top 10 pathways are shown in figure 1C. Only one biological pathway, cAMP signalling pathway, was enriched by DMPs for the FEV_1/FVC<LLN criteria.

Global methylation: Alu and LINE-1 sites

We investigated the overall differences in methylation between those with and without airflow obstruction using the Alu and LINE-1 CpGs as markers of global methylation. Results show that of 122 differentially methylated Alu sites and 13 differentially methylated LINE-1 sites, 117 and 12, respectively, were hypomethylated in those with a FEV_1/FVC ratio <70% (online supplemental table S5). Moreover, of 781 differentially methylated Alu sites and 105 LINE-1 sites, 768 and 101, respectively, were hypomethylated in those with a FEV_1/FVC ratio <LLN (online supplemental table S5 and figure 2).

Cross-sectional lung function and methylation

We assessed blood methylation and lung function as a continuous measure in PLWH and identified 103 DMPs and 9 DMRs associated with FEV_1 (online supplemental tables S2 and S3). The absolute effect of a 1% change in methylation of the DMPs for FEV_1 was on average 110.69 mL and ranged from 1.71 mL to 850.63 mL (online supplemental figure S2). In total, 7112 DMPs and 888 DMRs were identified as having an association with FEV_1/FVC (online supplemental tables S2 and S3). The average effect size of a 1% change in methylation of DMPs on FEV_1/FVC was 1.67%, and the minimum and maximum were 0.10% and 27.03%, respectively. While the DMPs for FEV_1 did not enrich any KEGG pathways, online supplemental figure S3 shows that DMPs for FEV_1/FVC enriched multiple KEGG pathways, the most significant of which was inflammatory mediator regulation of transient receptor potential channels.

The majority of the DMPs were located in regions with low CpG density (Open Sea and Shelf) (online supplemental table S4). Approximately 45% of FEV_1 DMPs were located within regions with high density (CpG Island) or close to CpG Islands (Shores), while 13.97% and 8% of FEV_1/FVC ratio DMPs were located in CpG Islands and Shores, respectively. In addition, 4.8% and 11.14% of DMPs influencing FEV_1 and FEV_1/FVC, respectively, are located within known promoter regions.

Table 3 shows the seven genes corresponding to the top DMPs associated with lung function, which include *MACROD1*, *DPF3*, *CACNA1G*, *PADI4*, *DLEC1*, *FERMT3* and *ELANE*. Overall hypermethylation of the top DMPs was associated with decreased FEV_1. Moreover, 78% of all DMPs for FEV_1 also have negative effects. Similar results were found for the FEV_1/FVC

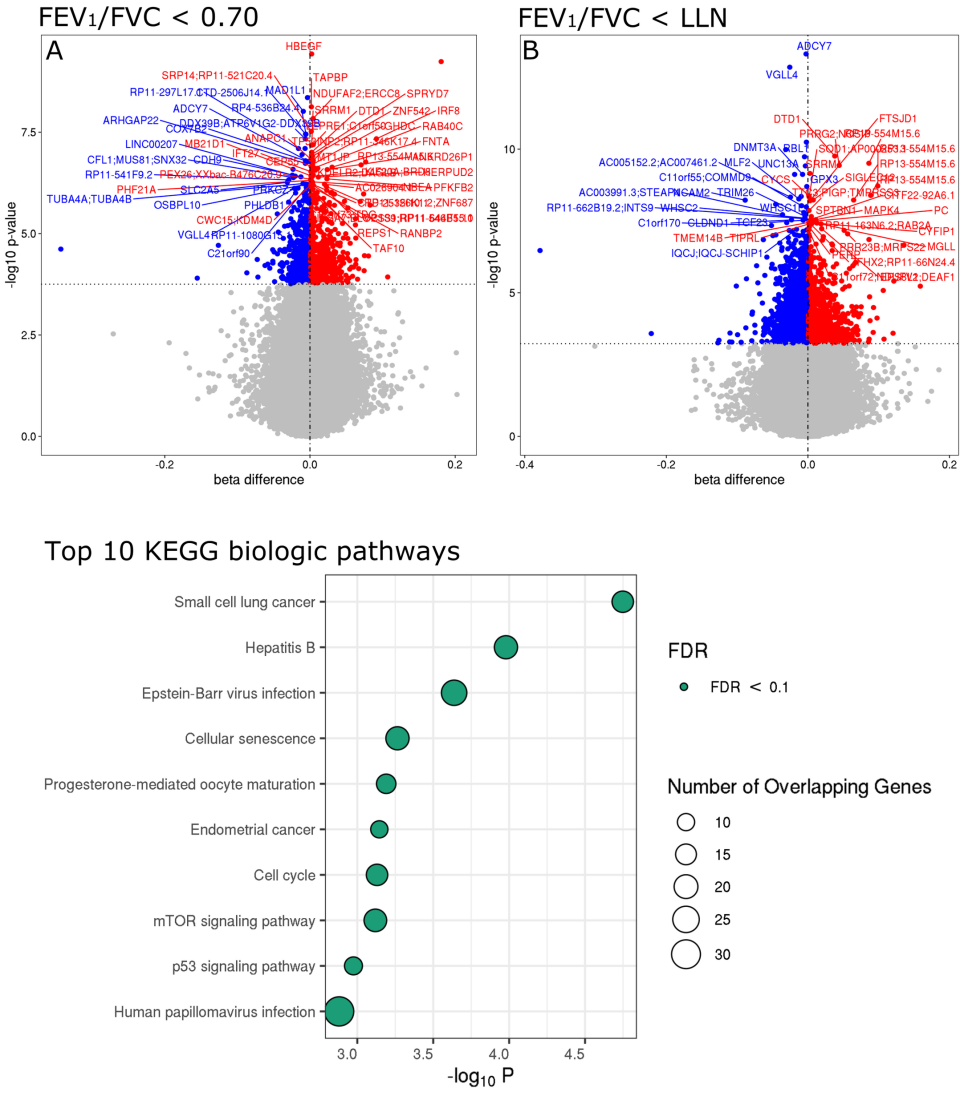

Figure 1 Differentially methylated points (DMPs) for airflow obstruction and top 10 biologic-enriched pathways for airflow obstruction (C). The x-axis on A and B represents the effect size difference of the DMPs between subjects with and without airflow obstruction (reference group: no airflow obstruction). The y-axis on figures A and B represents the level of statistical significance for each DMPs. Airflow obstruction was defined as (A) FEV_1/FVC ratio was <0.70 and (B) FEV_1/FVC <lower limit of normal. The red and blue colours represent hypomethylation and hypermethylation, respectively. The dashed horizontal line in A and B represents the −log10 p value that corresponds to the false discovery rate (FDR)<0.10. The axis in C represents the enrichment level of significance (×) for each biological pathway (y). The size of the circles inside the figure represents the number of overlapping genes in the pathways and the genes characterised by DMPs. The colour green represents significant enrichment based on the FDR<0.1. KEGG, Kyoto Encyclopedia of Genes and Genomes.

ratio, where 86% of the DMPs had effect sizes with a negative direction. The top DMR associated with FEV_1 included 7 CpGs and corresponding to *CTHRC1*, while the top DMR for FEV_1/FVC included 27 CpGs and correspond to *CTSZ* (online supplemental tables S3).

Lung function decline

Over a median of 5 years, PLWH had a small decline in FEV_1 (online supplemental figure S4). Overall FEV_1 declined on average by 20 mL/year; however, the decline was not statistically significant (p=0.43, 95% CI −187.04 to 430.19). Caucasian PLWH showed an overall FEV_1 decline, but this was not statistically significant either (32 mL/year, p=0.33, 95% CI −205.06 to 586.28). There were no DMPs associated with FEV_1 decline in the overall cohort. In order to remove the noise in FEV_1 decline that was due to the variability between ethnic groups, we analysed Africans (n=31), Caucasians (n=97) and Hispanics (n=20) separately; however, only the Caucasians yielded significant results. We found 53 DMPs and 4 DMRs (online supplemental tables S2 and S3) that were significantly associated with FEV_1 decline in Caucasian PLWH. The top five DMPs are shown in table 4; two of them, cg09595479 and cg08625260, were located within a CpG Island and corresponded to *PRPH* and *IRS2*, respectively.

DISCUSSION

This study is the first epigenome-wide association analysis on airflow obstruction and lung function in a multiethnic cohort of PLWH. Previous research on DNA methylation and lung function has focused on general or COPD-specific populations.[12 29] Because DNA methylation can be altered by environmental factors including chronic infections, these past results may

Table 2 Most significant differentially methylated positions (DMPs) for airway obstruction in people living with HIV

DMP								
Trait	**Probe**	**Chr****	**Gene**	**Relation to Island**	**Beta**	**SE**	**P value**	**FDR**
FEV_1/FVC ratio <70%	cg01175605	2	*HK2*	Open Sea	−0.00345	0.032	4.42×10^{-09}	1.16×10^{-03}
	cg20868410	5	*HBEGF*	South Shore	0.00231	0.037	3.71×10^{-10}	2.26×10^{-04}
	cg27385940	6	*TAPBP*	Island	0.13592	0.023	7.47×10^{-09}	1.47×10^{-03}
	cg13209990	7	*MAD1L1*	Open Sea	−0.22058	0.038	9.65×10^{-09}	1.52×10^{-03}
	cg13632595	19	Unknown	Open Sea	0.18069	0.320	5.74×10^{-10}	2.26×10^{-04}
FEV_1/FVC ratio <LLN	cg13071306	1	*GPR153*	Island	−0.00215	0.029	5.72×10^{-11}	1.50×10^{-05}
	cg23455629	2	Unknown	Open Sea	−0.40923	0.063	1.03×10^{-10}	2.04×10^{-05}
	cg15288310	3	*VGLL4*	South Shelf	−0.02561	0.047	1.40×10^{-13}	5.53×10^{-08}
	cg02910002	16	*ADCY7*	South Shelf	−0.00262	0.048	4.85×10^{-14}	3.82×10^{-08}
	cg08050464	20	Intergenic	South Shore	−0.27178	0.043	1.84×10^{-10}	2.42×10^{-05}

*Chromosome.
FDR, false discovery rate.

not reflect the relationship between methylation and lung function in the HIV-specific context. Our study revealed that PLWH with airflow obstruction have a distinct blood DNA methylation profile compared with PLWH with normal lung function, and that airflow obstruction is linked with global hypomethylation in HIV. Furthermore, our results indicate that although DNA methylation is associated with cross-sectional lung function, there was minimal influence on lung function decline.

Most methylated CpGs in the genome are located in CpG-rich sequences of the transposable elements Alu and LINE-1;

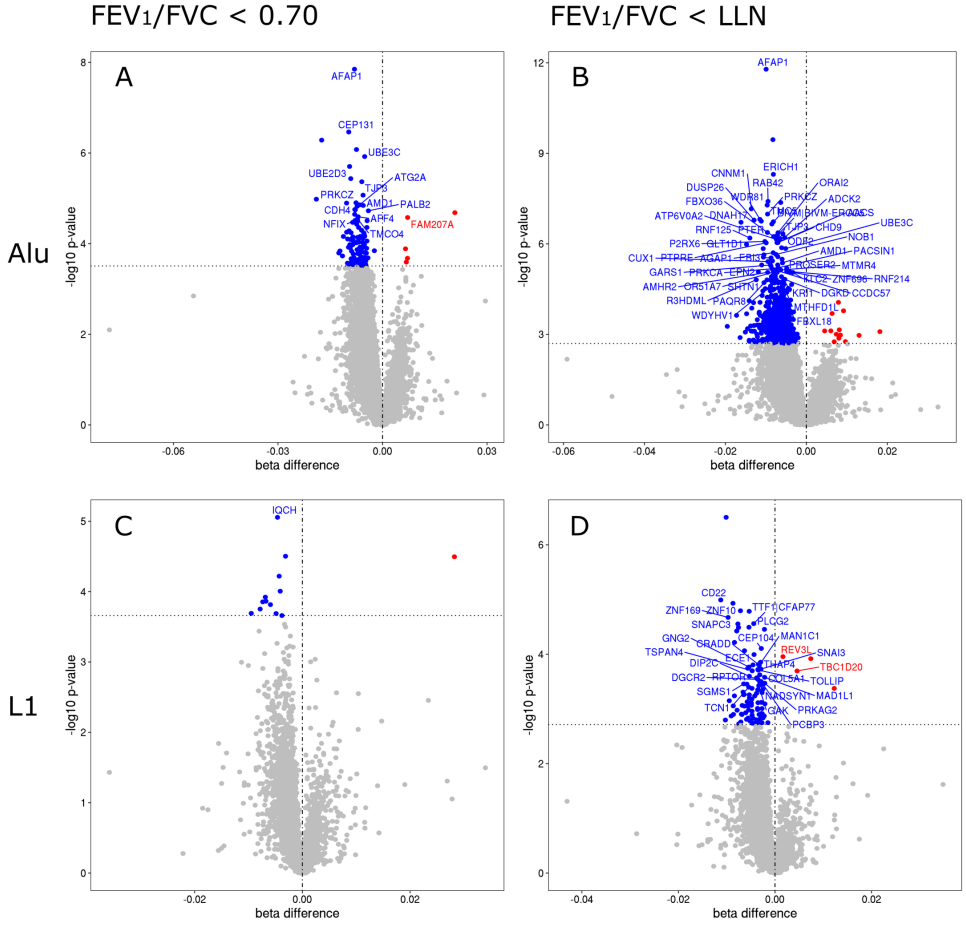

Figure 2 Alu and LINE-1 differentially methylated sites in people living with HIV with airflow obstruction. (A) and (B) Differentially methylated points (DMPs) in transposable element Alu between subjects with and without airflow obstruction (reference group: without airflow obstruction). (C) and (D) DMPs in transposable element LINE-1 between subjects with and without airflow obstruction (reference group: without airflow obstruction). The x-axis on the plots represents the effect size difference of the DMPs between subjects with and without airflow obstruction. The y-axis on the plots represents the level of significance for each DMPs. Airflow obstruction was defined as FEV_1/FVC ratio was <0.70 (A and C) and FEV_1/FVC <lower limit of normal (B and D). The dashed horizontal line inside the plots represents the −log10 p value that corresponds to the false discovery rate<0.10.

Table 3 Most significant differentially methylated positions for baseline FEV_1 and FEV_1/FVC

Trait	Probe	Chr*	Gene	Relation to Island	Beta	SE	P value	FDR
FEV_1	cg09006039	4	Intergenic	Open Sea	94.41	17.42	6.05×10^{-08}	9.54×10^{-03}
	cg22040274	5	Unknown	Open Sea	−127.22	22.38	1.32×10^{-08}	2.74×10^{-03}
	cg01557460	11	*MACROD1*	North Shore	−53.31	9.31	1.06×10^{-08}	2.74×10^{-03}
	cg17903071	14	*DPF3*	Island	−850.63	147.35	7.79×10^{-09}	2.74×10^{-03}
	cg23599559	17	*CACNA1G*	Island	−302.09	53.24	1.39×10^{-08}	2.74×10^{-03}
FEV_1/FVC ratio	cg16091981	1	*PADI4*	Open Sea	−0.0223	0.003	7.38×10^{-12}	1.94×10^{-06}
	cg02703606	3	*DLEC1*	Open Sea	−0.0134	0.002	1.19×10^{-10}	2.35×10^{-05}
	cg03045038	11	*FERMT3*	North Shore	−0.0094	0.001	7.38×10^{-12}	1.94×10^{-06}
	cg26615186	16	Unknown	Open Sea	−0.0088	0.001	1.70×10^{-10}	2.68×10^{-05}
	cg06100973	19	*ELANE*	North Shore	−0.0085	0.001	1.49×10^{-12}	1.17×10^{-06}

*Chromosome.
FDR, false discovery rate.

therefore, CpG sites within these elements are often used as markers for global methylation. Hypomethylation along these elements has also been used as a marker for ageing[9] and has also been associated with lower lung function in healthy older men.[21] Our results on Alu and LINE-1 show that those PLWH with airflow obstruction have greater hypomethylation when compared with those with normal lung function. This is despite the fact that subjects with airflow obstruction in our cohort still had relatively mild decrements in FEV_1 and were relatively young. These results suggest that the process of airflow obstruction in PLWH may reflect an advanced ageing process, concurrent with the observations of accelerated ageing and comorbid age-related conditions in HIV.[30] The biological implications of global hypomethylation and lung obstruction need further investigation, however given that global hypomethylation can potentially lead to overexpression of genes and activation of transposable elements and thus promote tumourigenesis in the lung,[31] the interplay between smoking, airflow obstruction and lung cancer may be mediated by this specific methylation process. As PLWH are at higher risk of developing lung cancer,[32] this hypothesis should be explored further.

The top pathways enriched by DMPs for airflow obstruction included the Hepatitis B, Epstein-Barr virus and Human Papillomavirus pathways, which raises the intriguing possibility that concurrent viral infections in PLWH may be drivers of airflow obstruction. PLWH are known to be at higher risk of coinfection with these particular viruses,[33–35] although to our knowledge the association of these chronic infections with COPD from an epidemiologic standpoint has not yet been reported. Another top pathway enriched by DMPs, the cAMP signalling pathway, is related to small airway remodelling in COPD, and therapeutic compounds that target proteins in this pathway such as roflumilast have been used to treat COPD.[36] Furthermore, differentially methylated genes in the small airways of patients with COPD have been also found to be enriched in the cAMP signalling pathway.[29] While more research is needed to validate that the differential methylated genes could alter enriched biologic pathways, no previous research has linked these pathways at the DNA methylation level with airflow obstruction in PLWH.

Because of the uniqueness of our study cohort, a multiethnic group of PLWH who were ART naïve at study entry, we could not replicate this analysis. However, there is a modest overlap between genes identified in our study with previous work looking at methylation in COPD populations (online supplemental figure S5 and table S2). In accordance with previous findings on the small airway methylation profiles of patients with COPD,[29] we identified a large number of DMPs associated with airflow obstruction in PLWH . One of our discovered DMPs (cg13071306) corresponds to a gene previously described in airways diseases, *GPR153*. The function of GPR153 is poorly understood; however, this gene belongs to a rhodopsin family of G protein-coupled receptors (GPCRs), which are mediators of airway smooth muscle contraction and increased airway resistance. GPCRs, for example, are frequently dysregulated in asthma.[37] In addition, one of our top hits for airflow obstruction (cg01175605) is located in an exon of *HK2*, which has previously been linked to COPD and lung cancer. *HK2* is a hexokinase predominantly localised to the mitochondrial membrane as part of the glucose metabolism pathway, but has also been reported to be expressed in the lung.[38] Specifically, the CpG site cg18638581 in the promoter region of *HK2* was associated with COPD, FEV_1 and FEV_1/FVC in a previously reported COPD cohort.[39] This effect was independent of tobacco use. Previous

Table 4 Most significant differentially methylated positions (DMPs) for FEV_1 decline in Caucasian people living with HIV

DMP								
Trait	**Probe**	**Chr***	**Gene**	**Relation to Island**	**Beta**	**SE**	**P value**	**FDR**
FEV_1 decline	cg13911697	11	Intergenic	Open Sea	−11.36	2.11	1.30×10^{-07}	2.78×10^{-02}
	cg15056794	11	*BLID*	Open Sea	−12.97	2.32	4.44×10^{-08}	2.78×10^{-02}
	cg09595479	12	*PRPH*	Island	19.50	3.59	1.02×10^{-07}	2.78×10^{-02}
	cg05300248	18	*CHST9*	Open Sea	10.85	2.02	1.41×10^{-07}	2.78×10^{-02}
	cg08625260	13	*IRS2*	Island	19.59	3.71	2.19×10^{-07}	3.25×10^{-02}

*Chromosome.
FDR, false discovery rate.

work has demonstrated as well that *HK2* is upregulated in non-small cell lung cancer.[40] Possible regulation of *HK2* expression may occur through epigenetic changes to influence the development of COPD and lung cancer.

While this study provides novel findings, it also has several limitations. First, our study cohort was restricted to PLWH over 40 years of age, with a detectable viral load and CD4 T cell count >500 cells/mm^3, and who were not at the time of study entry on antiretroviral therapy. Whether these results apply to PLWH who have been on antiretroviral therapy for many years and have achieved viral suppression cannot be ascertained here. Second, the FEV_1 decline analysis suggested that decline is not likely to be affected by methylation changes; however it is also possible that our analysis of FEV_1 decline was simply underpowered. The proportion of our cohort meeting criteria for airflow obstruction was small and analyses of methylation in cohorts with a greater fraction of patients with COPD should be performed in the future. It is possible that some of the effects identified by our study also apply to non-HIV cohort; however this was outside the scope of our study. Third, the direction of effect, whether DNA methylation disruptions influence the progression of airflow obstruction or conversely whether airflow obstruction alters DNA methylation profiles cannot be ascertained by these data. Further study in cohorts with longitudinal DNA methylation profiling would be essential to solving this problem. Finally, because of the pressing need to extrapolate findings to diverse populations of PLWH, we included multiple ethnic groups in our analysis while controlling for population structure to the best of our abilities. However, since some methylation sites are specific to certain ethnicities, and would only be identified in homogenous populations, future efforts should focus on increasing the sample size of underrepresented minority groups. Despite these limitations, we have identified for the first time linkages between lung function, airflow obstruction and methylation in a unique cohort of PLWH. Epigenetic disruptions at key genes may hold clues to the increased risk of chronic lung diseases in this population.

Acknowledgements The authors thank the study participants who were involved in the START trial (see reference [13] for the complete list of START investigators, and [14] for the complete list of START Pulmonary Sub study investigators). They used computational resources from Compute Canada and WestGrid to conduct this research.

Contributors AIHC, CXY, MO, SFPM, DDS, KK, JL: study design and conception, interpretation of data, manuscript drafting and editing. JY: data acquisition, manuscript drafting and editing. JM, DL, RN, FH, HK, ND: data acquisition, manuscript editing. LM: data acquisition, interpretation of data, manuscript editing. MK: study design, interpretation of data, manuscript editing.

Funding AIHC is supported by MITACS accelerate. This study was supported by the Canadian Institutes of Health Research (F15-00958). JML is supported by the Michael Smith Foundation for Health Research and the Canadian Institutes of Health Research. The START Pulmonary Substudy was supported by the National Heart Lung and Blood Institute [R01 HL096453]; the parent START trial was primarily supported by the National Institute of Allergy and Infectious Diseases Division of AIDS [UM1 AI068641 and UM AI120197] with additional support from the German Ministry of Education and Research, the European AIDS Treatment Network (NEAT), the Australian National Health and Medical Research Council, and the UK Medical Research Council and National Institute for Health Research. The Veterans Health Administration Office of Research and Development also provided protected research time in support of this study. The University of Minnesota served as sponsor of the study.

Disclaimer None of the funders nor sponsor had any input regarding the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study was conducted under the University of British Columbia Ethics Board Approval Number H15-02166.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Deidentified data are available upon reasonable request. The full methylation data will be deposited into a public repository in 2022 upon conclusion of the START study. Requests can be directed to Dr Janice Leung at janice.leung@hli.ubc.ca.

ORCID iD

Ken Kunisaki http://orcid.org/0000-0001-8644-2827

REFERENCES

1 Palella FJ, Baker RK, Moorman AC, *et al*. Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study. *J Acquir Immune Defic Syndr* 2006;43:27–34.
2 Drummond MB, Kirk GD, Astemborski J, *et al*. Association between obstructive lung disease and markers of HIV infection in a high-risk cohort. *Thorax* 2012;67:309–14.
3 Gingo MR, Nouraie M, Kessinger CJ, *et al*. Decreased lung function and all-cause mortality in HIV-infected individuals. *Ann Am Thorac Soc* 2018;15:192–9.
4 Diaz PT, Wewers MD, Pacht E, *et al*. Respiratory symptoms among HIV-seropositive individuals. *Chest* 2003;123:1977–82.
5 Kunisaki KM, Niewoehner DE, Collins G, *et al*. Pulmonary function in an international sample of HIV-positive, treatment-naïve adults with CD4 counts >500 cells/µL: a substudy of the INSIGHT Strategic Timing of AntiRetroviral Treatment (START) trial. *HIV Med* 2015;16 Suppl 1:119–28.
6 Shrine N, Guyatt AL, Erzurumluoglu AM, *et al*. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet* 2019;51:481–93.
7 Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012;13:484–92.
8 Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. *Biochim Biophys Acta* 1775;2007:138–62.
9 Erichsen L, Beermann A, Arauzo-Bravo MJ, *et al*. Genome-Wide hypomethylation of LINE-1 and Alu retroelements in cell-free DNA of blood is an epigenetic biomarker of human aging. *Saudi J Biol Sci* 2018;25:1220–6.
10 Bell JT, Tsai P-C, Yang T-P, *et al*. Epigenome-Wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet* 2012;8:e1002629.
11 Bolund ACS, Starnawska A, Miller MR, *et al*. Lung function discordance in monozygotic twins and associated differences in blood DNA methylation. *Clin Epigenetics* 2017;9:132.
12 de Vries M, Nedeljkovic I, van der Plaat DA, *et al*. DNA methylation is associated with lung function in never smokers. *Respir Res* 2019;20:268.
13 INSIGHT START Study Group, Lundgren JD, Babiker AG, *et al*. Initiation of antiretroviral therapy in early asymptomatic HIV infection. *N Engl J Med* 2015;373:795–807.
14 Kunisaki KM, Niewoehner DE, Collins G, *et al*. Pulmonary effects of immediate versus deferred antiretroviral therapy in HIV-positive individuals: a nested substudy within the multicentre, international, randomised, controlled strategic timing of antiretroviral treatment (start) trial. *Lancet Respir Med* 2016;4:980–9.
15 Quanjer PH, Stanojevic S, Cole TJ, *et al*. Multi-Ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012;40:1324–43.
16 Logue MW, Smith AK, Wolf EJ, *et al*. The correlation of methylation levels measured using illumina 450K and EPIC BeadChips in blood samples. *Epigenomics* 2017;9:1363–71.
17 Triche TJ, Weisenberger DJ, Van Den Berg D, *et al*. Low-Level processing of illumina Infinium DNA methylation BeadArrays. *Nucleic Acids Res* 2013;41:e90.
18 Teschendorff AE, Marabita F, Lechner M, *et al*. A beta-mixture quantile normalization method for correcting probe design bias in illumina Infinium 450 K DNA methylation data. *Bioinformatics* 2013;29:189–96.
19 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;8:118–27.
20 Yang AS, Estécio MRH, Doshi K, *et al*. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res* 2004;32:38e–38.
21 Lange NE, Sordillo J, Tarantini L, *et al*. Alu and LINE-1 methylation and lung function in the normative ageing study. *BMJ Open* 2012;2:e001231.
22 Zheng Y, Joyce BT, Liu L, *et al*. Prediction of genome-wide DNA methylation in repetitive elements. *Nucleic Acids Res* 2017;45:8697–711.
23 Houseman EA, Accomando WP, Koestler DC, *et al*. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012;13:86.

24 Rahmani E, Shenhav L, Schweiger R, *et al*. Genome-Wide methylation data mirror ancestry information. *Epigenetics Chromatin* 2017;10:1.
25 Lee MK, Hong Y, Kim S-Y, *et al*. Epigenome-Wide association study of chronic obstructive pulmonary disease and lung function in Koreans. *Epigenomics* 2017;9:971–84.
26 Venables WN, Ripley BD. *Modern applied statistics with S*. 4th ed. New York: Springer, 2002. https://www.springer.com/gp/book/9780387954578
27 Pinheiro J, Bates D, DebRoy S, *et al*. nlme: linear and nonlinear mixed effects models, 2015. Available: https://cran.r-project.org/web/packages/nlme/index.html [Accessed 21 Jul 2020].
28 Peters TJ, Buckley MJ, Statham AL, *et al*. De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin* 2015;8:6.
29 Vucic EA, Chari R, Thu KL, *et al*. DNA methylation is globally disrupted and associated with expression changes in chronic obstructive pulmonary disease small airways. *Am J Respir Cell Mol Biol* 2014;50:912–22.
30 Van Epps P, Kalayjian RC. Human immunodeficiency virus and aging in the era of effective antiretroviral therapy. *Infect Dis Clin North Am* 2017;31:791–810.
31 Pfeifer GP, Rauch TA. DNA methylation patterns in lung carcinomas. *Semin Cancer Biol* 2009;19:181–7.
32 Sigel K, Wisnivesky J, Gordon K, *et al*. HIV as an independent risk factor for incident lung cancer. *AIDS* 2012;26:1017–25.
33 Greer AE, Ou S-S, Wilson E, *et al*. Comparison of hepatitis B virus infection in HIV-infected and HIV-uninfected participants enrolled in a multinational clinical trial: HPTN 052. *J Acquir Immune Defic Syndr* 2017;76:388–93.
34 Ferenczy A, Coutlée F, Franco E, *et al*. Human papillomavirus and HIV coinfection and the risk of neoplasias of the lower genital tract: a review of recent developments. *CMAJ* 2003;169:431–4.
35 Ling PD, Vilchez RA, Keitel WA, *et al*. Epstein-Barr virus DNA loads in adult human immunodeficiency virus type 1-infected patients receiving highly active antiretroviral therapy. *Clin Infect Dis* 2003;37:1244–9.
36 Oldenburger A, Maarsingh H, Schmidt M. Multiple facets of cAMP signalling and physiological impact: cAMP compartmentalization in the lung. *Pharmaceuticals* 2012;5:1291–331.
37 Penn RB, Bond RA, Walker JKL. GPCRs and arrestins in airways: implications for asthma. *Handb Exp Pharmacol* 2014;219:387–403.
38 Heikkinen S, Suppola S, Malkki M, *et al*. Mouse hexokinase II gene: structure, cDNA, promoter analysis, and expression pattern. *Mamm Genome* 2000;11:91–6.
39 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.
40 Li W, Gao F, Ma X, *et al*. Deguelin inhibits non-small cell lung cancer via down-regulating hexokinases II-mediated glycolysis. *Oncotarget* 2017;8:32586–99.

DNA Methylation is Associated with Airflow Obstruction in Patients Living with Human Immunodeficiency Virus

Supplementary Figures

Ana I. Hernández Cordero,[1] Chen Xi Yang,[1] Ma'en Obeidat,[1] Julia Yang,[1] Julie MacIsaac,[2] Lisa McEwen,[2] David Lin,[2] Michael Kobor,[2] Richard M. Novak,[3] Fleur Hudson,[4] Hartwig Klinker,[5] Nila Dharan,[6] S.F. Paul Man,[1,7] Don D. Sin,[1,7] Ken M. Kunisaki,[8] Janice M. Leung[1,7] for the INSIGHT START Pulmonary and Genomic Substudy Groups

1. Centre for Heart Lung Innovation, University of British Columbia, Vancouver, BC
2. Centre for Molecular Medicine and Therapeutics, University of British Columbia, Vancouver, BC, Canada
3. Section of Infectious Diseases, University of Illinois at Chicago, Chicago, IL, United States of America
4. MRC Clinical Trials Unit, University College London, London, United Kingdom
5. Medizinische Klinik und Poliklinik II, Universitätsklinikum Würzburg, Würzburg, Germany
6. Kirby Institute, Sydney, Australia
7. Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, BC, Canada
8. Minneapolis Veterans Affairs Health Care System, Section of Pulmonary, Critical Care and Sleep Medicine and the Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Department of Medicine, University of Minnesota, Minneapolis, MN, United States of America

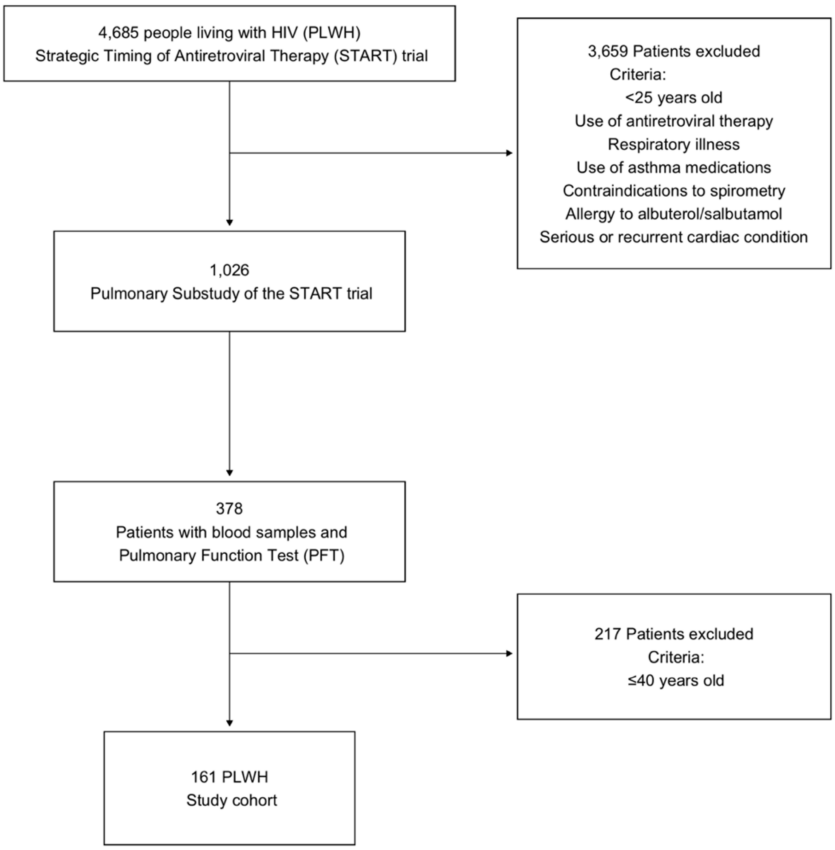

Figure S1. Diagram of the study cohort. The diagram shows the exclusion criteria used to obtain the study cohort.

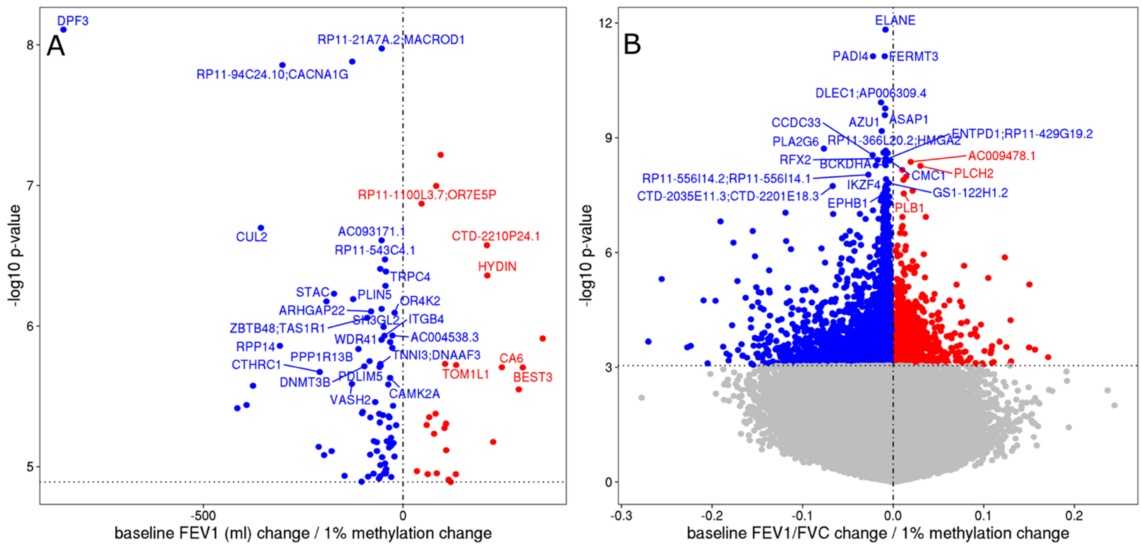

Figure S2. DMPs Associated with FEV_1 and FEV_1/FVC. The figure shows the absolute effect of a 1% change in methylation of the DMPs for FEV_1 (A) and FEV_1/FVC (B) (x-axis). The y-axis on the plots represents the level of statistical significance for each DMP. The dashed horizontal line in A and B represents the -log10 p-value that correspond to the FDR<0.1.

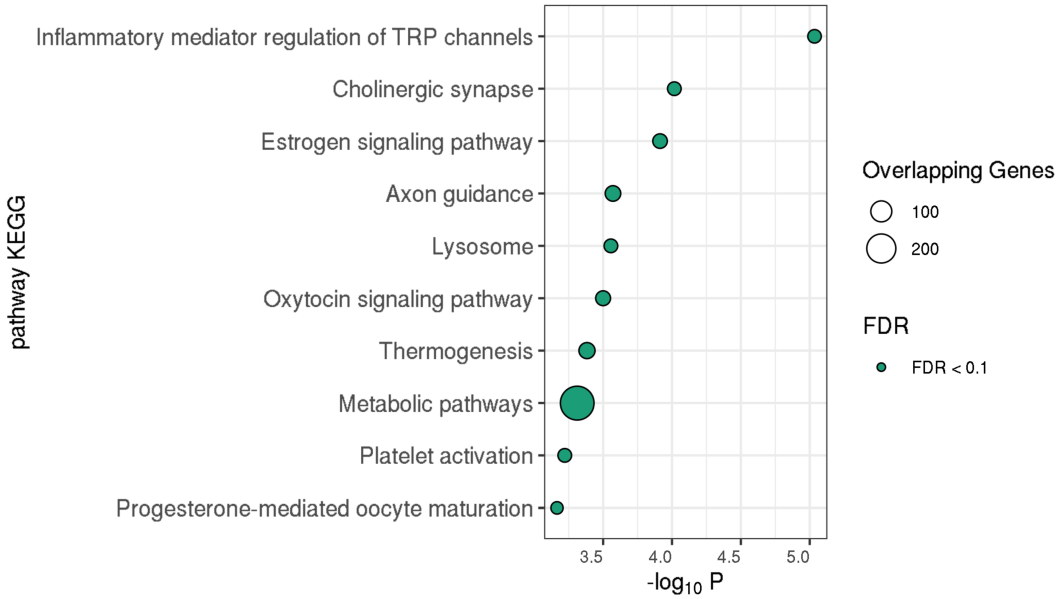

Figure S3. Top 10 KEGG biologic pathways enriched by genes identified for the FEV_1/FVC ratio. The axis on the figure represents the pathway's enrichment level of significance (x-axis) for each biologic pathway (y-axis). The size of the circles inside the figure area represents the number of overlapping genes between the genes in the pathways and the genes characterized by DMPs. The color green represents statistically significant enrichment based on the FDR<0.1.

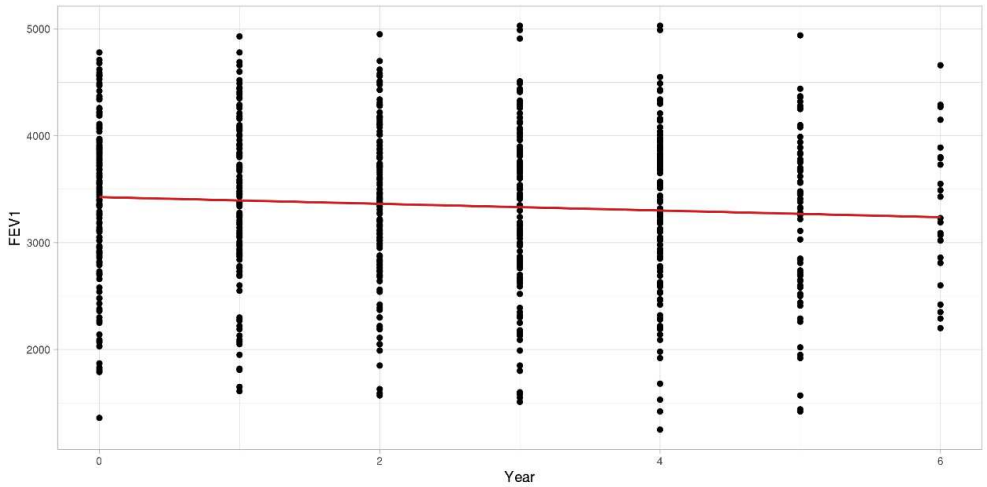

Figure S4. FEV_1 decline in PLWH over the course of a median of 5 years. The plot's x- and y-axis represent the number of years and forced expiratory volume in 1 second (FEV_1) records, respectively. Dots inside the plot represent records for the subjects in the study cohort. The red line represents the linear decline of FEV_1 over the course of the Strategic Timing of Antiretroviral Therapy (START) trial.

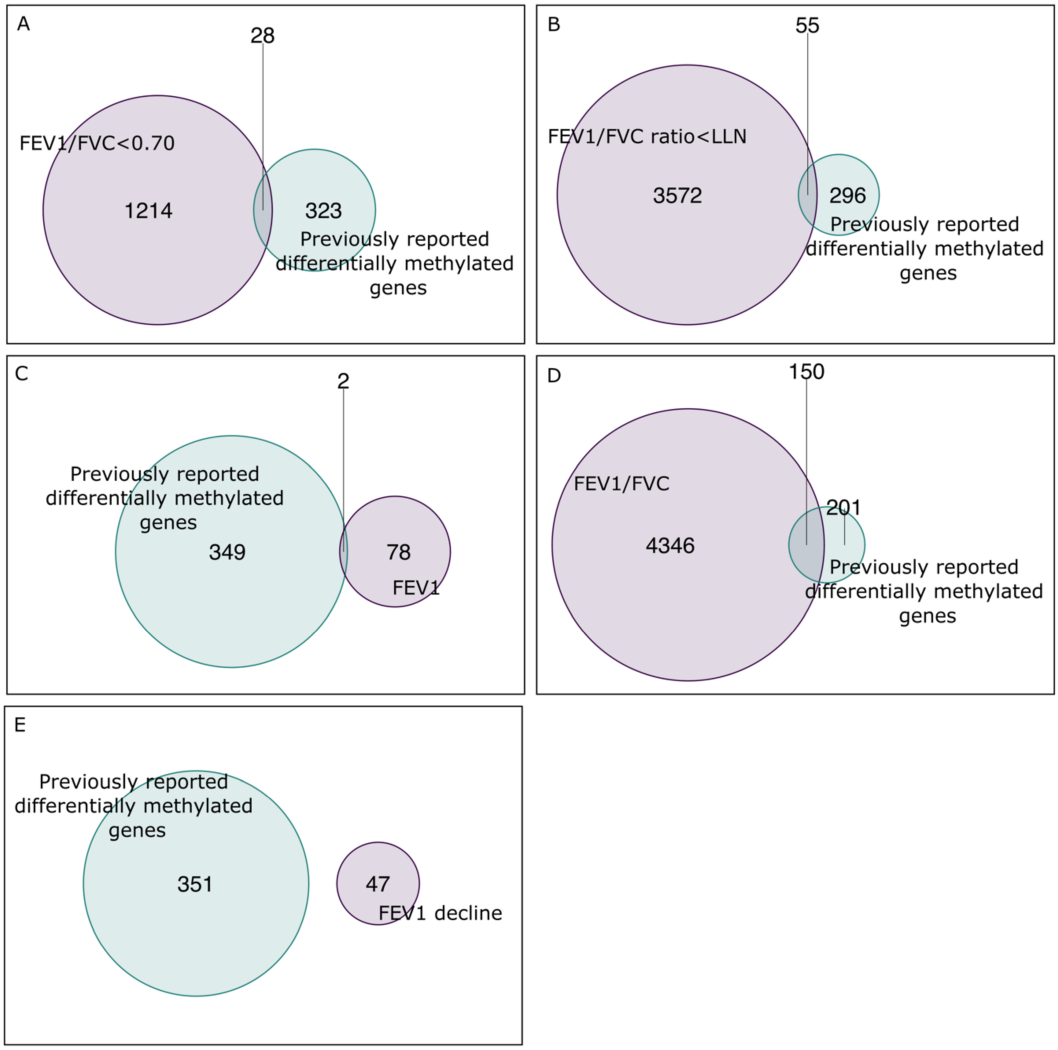

Figure S5. Overlapping of differentially methylated genes identified in previously published COPD and population-based cohorts shown in green [1,2] and in PLWH shown in purple. Panels represent the overlap with differentially methylated genes for (A) FEV1/FVC<0.70; (B) FEV1/FVC<LLN; (C) FEV1 (D) FEV1/FVC; (E) FEV1 decline.

References

1 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. doi:10.1164/rccm.201108-1382OC

2 de Vries M, Nedeljkovic I, van der Plaat DA, *et al.* DNA methylation is associated with lung function in never smokers. *Respiratory Research* 2019;**20**:268. doi:10.1186/s12931-019-1222-8

DNA Methylation is Associated with Airflow Obstruction in Patients Living with Human Immunodeficiency Virus

Supplementary Figures

Ana I. Hernández Cordero,[1] Chen Xi Yang,[1] Ma'en Obeidat,[1] Julia Yang,[1] Julie MacIsaac,[2] Lisa McEwen,[2] David Lin,[2] Michael Kobor,[2] Richard M. Novak,[3] Fleur Hudson,[4] Hartwig Klinker,[5] Nila Dharan,[6] S.F. Paul Man,[1,7] Don D. Sin,[1,7] Ken M. Kunisaki,[8] Janice M. Leung[1,7] for the INSIGHT START Pulmonary and Genomic Substudy Groups

1. Centre for Heart Lung Innovation, University of British Columbia, Vancouver, BC
2. Centre for Molecular Medicine and Therapeutics, University of British Columbia, Vancouver, BC, Canada
3. Section of Infectious Diseases, University of Illinois at Chicago, Chicago, IL, United States of America
4. MRC Clinical Trials Unit, University College London, London, United Kingdom
5. Medizinische Klinik und Poliklinik II, Universitätsklinikum Würzburg, Würzburg, Germany
6. Kirby Institute, Sydney, Australia
7. Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, BC, Canada
8. Minneapolis Veterans Affairs Health Care System, Section of Pulmonary, Critical Care and Sleep Medicine and the Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Department of Medicine, University of Minnesota, Minneapolis, MN, United States of America

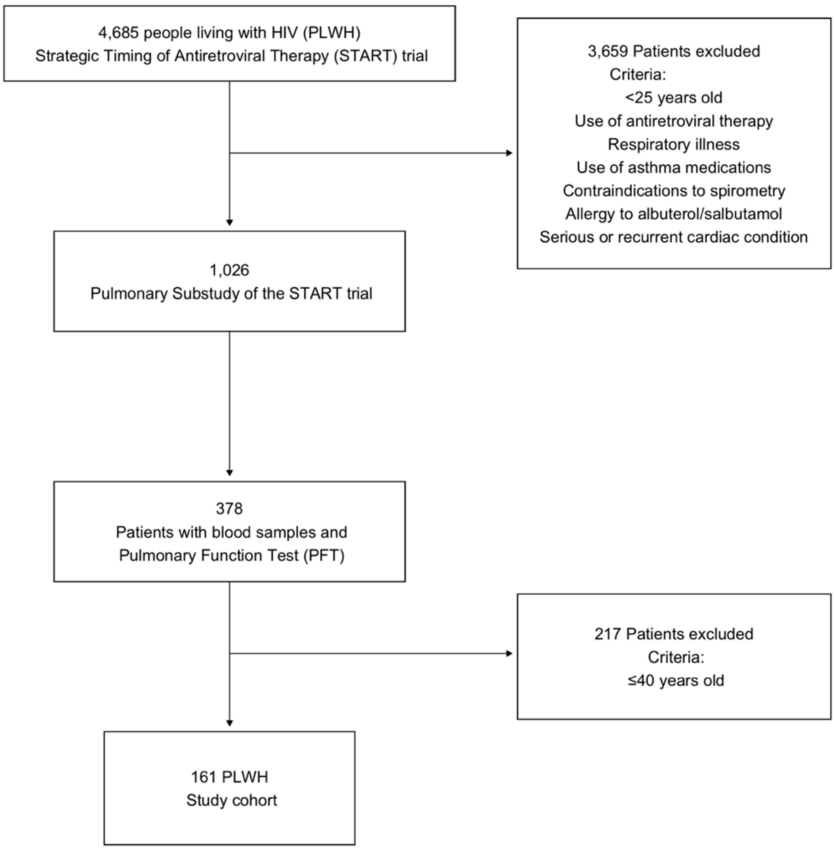

Figure S1. Diagram of the study cohort. The diagram shows the exclusion criteria used to obtain the study cohort.

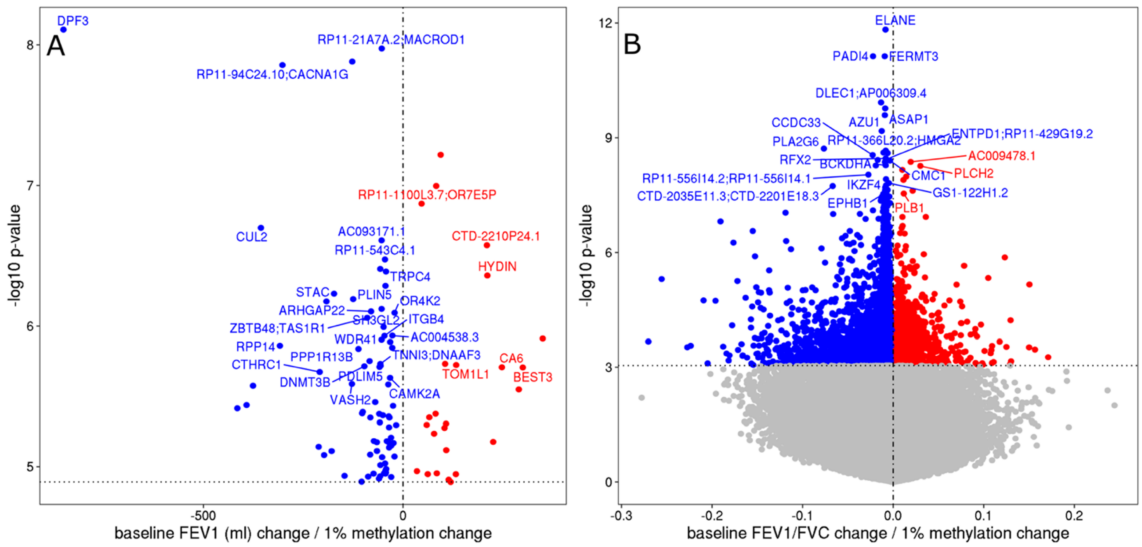

Figure S2. DMPs Associated with FEV_1 and FEV_1/FVC. The figure shows the absolute effect of a 1% change in methylation of the DMPs for FEV_1 (A) and FEV_1/FVC (B) (x-axis). The y-axis on the plots represents the level of statistical significance for each DMP. The dashed horizontal line in A and B represents the -log10 p-value that correspond to the FDR<0.1.

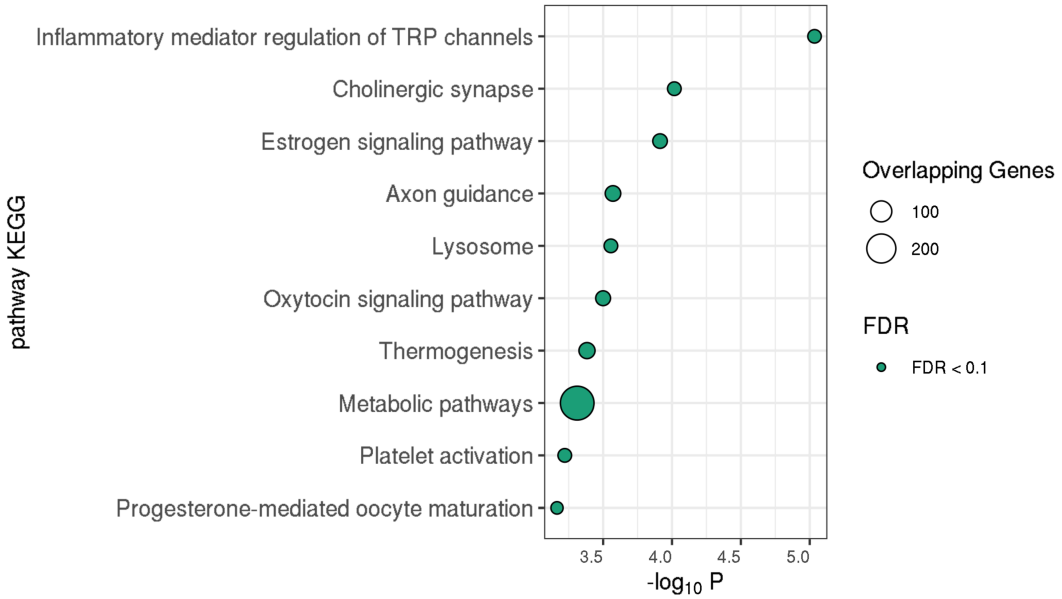

Figure S3. Top 10 KEGG biologic pathways enriched by genes identified for the FEV_1/FVC ratio. The axis on the figure represents the pathway's enrichment level of significance (x-axis) for each biologic pathway (y-axis). The size of the circles inside the figure area represents the number of overlapping genes between the genes in the pathways and the genes characterized by DMPs. The color green represents statistically significant enrichment based on the FDR<0.1.

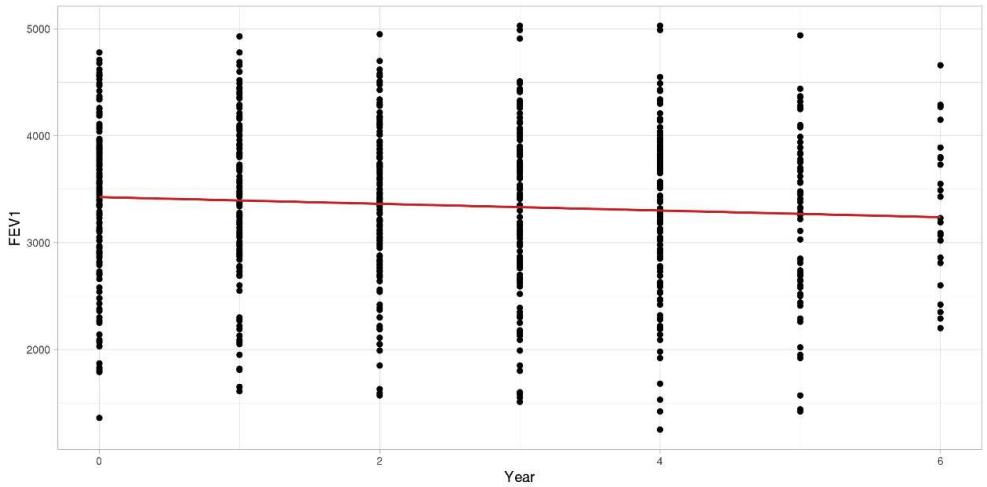

Figure S4. FEV_1 decline in PLWH over the course of a median of 5 years. The plot's x- and y-axis represent the number of years and forced expiratory volume in 1 second (FEV_1) records, respectively. Dots inside the plot represent records for the subjects in the study cohort. The red line represents the linear decline of FEV_1 over the course of the Strategic Timing of Antiretroviral Therapy (START) trial.

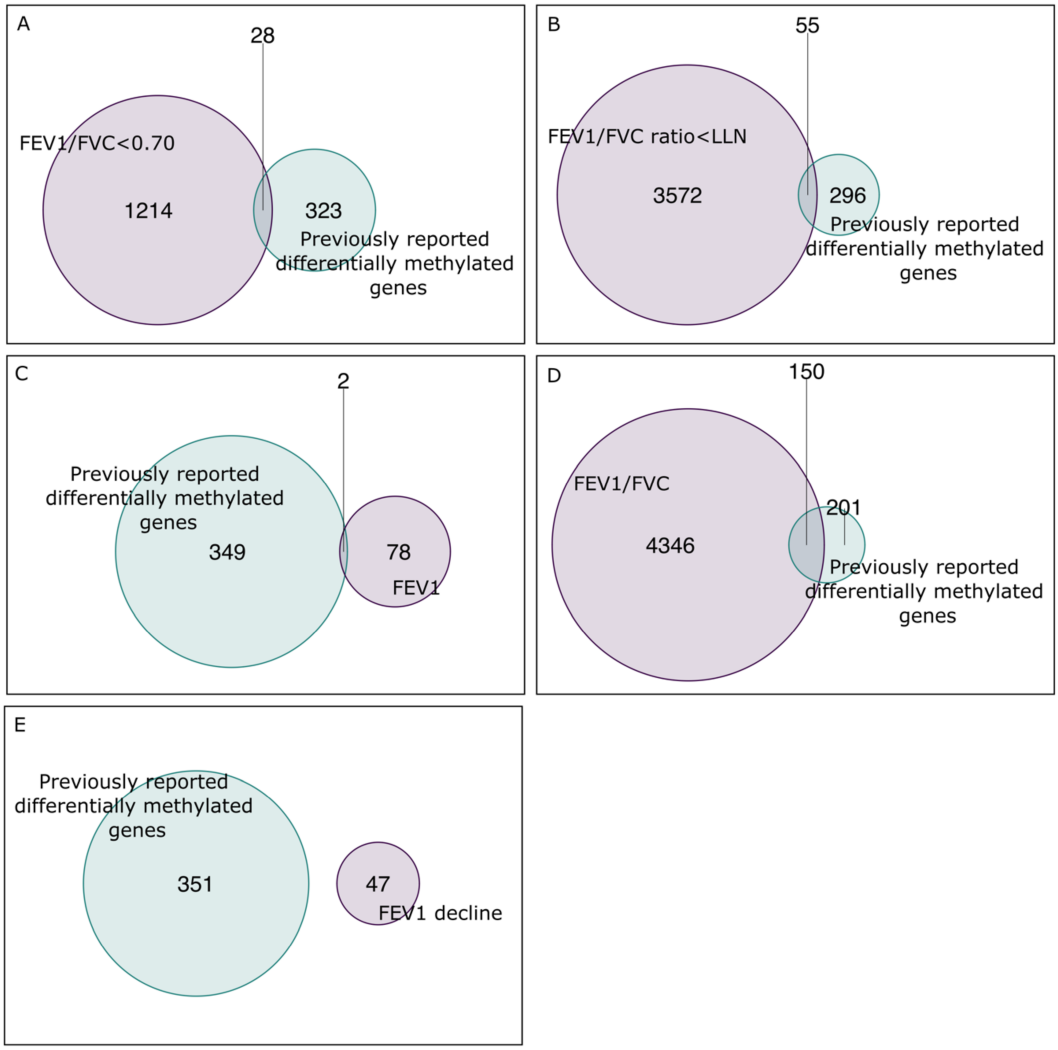

Figure S5. Overlapping of differentially methylated genes identified in previously published COPD and population-based cohorts shown in green [1,2] and in PLWH shown in purple. Panels represent the overlap with differentially methylated genes for (A) FEV1/FVC<0.70; (B) FEV1/FVC<LLN; (C) FEV1 (D) FEV1/FVC; (E) FEV1 decline.

References

1 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. doi:10.1164/rccm.201108-1382OC

2 de Vries M, Nedeljkovic I, van der Plaat DA, *et al.* DNA methylation is associated with lung function in never smokers. *Respiratory Research* 2019;**20**:268. doi:10.1186/s12931-019-1222-8