ORIGINAL ARTICLE
ASTHMA

Differential DNA methylation in bronchial biopsies between persistent asthma and asthma in remission

Cornelis J. Vermeulen [1,2], Cheng-Jian Xu [2,3,4], Judith M. Vonk [2,5], Nick H.T. ten Hacken [1,2], Wim Timens [2,6], Irene H. Heijink [1,2,6], Martijn C. Nawijn [2,6], Jeunard Boekhoudt [6], Antoon J. van Oosterhout [7], Karen Affleck [7], Markus Weckmann [8], Gerard H. Koppelman [2,3] and Maarten van den Berge [1,2]

Affiliations: [1]Dept of Pulmonary Diseases, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [2]University Medical Center Groningen, Groningen Research Institute for Asthma and COPD (GRIAC), University of Groningen, Groningen, The Netherlands. [3]Dept of Pediatric Pulmonology and Pediatric Allergology, University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital, Groningen, The Netherlands. [4]CiiM & TWINCORE, Helmholtz-Centre for Infection Research (HZI) and the Hannover Medical School (MHH), Hannover, Germany. [5]Dept of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [6]Dept of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [7]Allergic Inflammation Discovery Performance Unit, GlaxoSmithKline, Stevenage, UK. [8]Dept of Pediatric Pneumology and Allergology, University Medical Center of Schlesswig-Holstein, Airway Research Centre North, Member of the German Centre of Lung Research, Lübeck, Germany.

Correspondence: Cornelis J. Vermeulen, University of Groningen, University Medical Center Groningen, Dept of Pulmonary Diseases, NL-9700 RB Groningen, The Netherlands. E-mail: c.j.vermeulen@umcg.nl

@ERSpublications
Former asthma patients have epigenetic modifications not present in current asthma which are associated with the activity of genes involved in the resolution of inflammation. Their epigenetic profile also shows them to be different from healthy controls. http://bit.ly/2BGmCPl

Cite this article as: Vermeulen CJ, Xu C-J, Vonk JM, *et al*. Differential DNA methylation in bronchial biopsies between persistent asthma and asthma in remission. *Eur Respir J* 2020; 55: 1901280 [https://doi.org/10.1183/13993003.01280-2019].

ABSTRACT Approximately 40% of asthmatics experience remission of asthma symptoms. A better understanding of biological pathways leading to asthma remission may provide insight into new therapeutic targets for asthma. As an important mechanism of gene regulation, investigation of DNA methylation provides a promising approach. Our objective was to identify differences in epigenome wide DNA methylation levels in bronchial biopsies between subjects with asthma remission and subjects with persistent asthma or healthy controls.

We analysed differential DNA methylation in bronchial biopsies from 26 subjects with persistent asthma, 39 remission subjects and 70 healthy controls, using the limma package. The comb-p tool was used to identify differentially methylated regions. DNA methylation of CpG-sites was associated to expression of nearby genes from the same biopsies to understand function.

Four CpG-sites and 42 regions were differentially methylated between persistent asthma and remission. DNA methylation at two sites was correlated *in cis* with gene expression at *ACKR2* and *DGKQ*. Between remission subjects and healthy controls 1163 CpG-sites and 328 regions were differentially methylated. DNA methylation was associated with expression of a set of genes expressed in ciliated epithelium.

CpGs differentially methylated between remission and persistent asthma identify genetic loci associated with resolution of inflammation and airway responsiveness. Despite the absence of symptoms, remission subjects have a DNA methylation profile that is distinct from that of healthy controls, partly due to changes in cellular composition, with a higher gene expression signal related to ciliated epithelium in remission *versus* healthy controls.

Copyright ©ERS 2020

Introduction

Asthma is a chronic inflammatory disease of the airways, characterised by variable airflow obstruction associated with symptoms of wheezing, shortness of breath, chest tightness and coughing. Chronic inflammation in asthma is associated with remodelling of the airways, *e.g.* increased basal membrane thickness, airway epithelial shedding, increase of goblet cells, increased mucus production and increase of smooth muscle mass [1]. Interestingly, in some asthma patients, symptoms disappear over time, and the need for pulmonary medication ceases. The prevalence of asthma is highest between 10 and 25 years and gradually decreases at higher age, the latter indicating that asthma remission rates have become higher than incidence rates [2]. In adulthood, the average remission rate of asthma is ∼2% per year, with a higher chance to go into remission with an earlier onset of asthma, less severe airway obstruction and cessation of smoking [3, 4]. A subset of patients with asthma in remission still have airway obstruction and/or show airway hyperresponsiveness (AHR) in provocation tests [5]. VONK *et al.* [6] introduced the terms "clinical remission" and "complete remission". Clinical remission is defined as the absence of asthma symptoms and no use of asthma medication. Complete remission of asthma is defined as the absence of asthma symptoms, no use of asthma medication, normal lung function and no AHR.

We still have minimal understanding of the cellular and molecular mechanisms that determine whether or not asthma persists or undergoes apparent spontaneous resolution. As for asthma development, genetic and environmental factors are likely to be involved in asthma remission as well [7]. The effect of both genetic and environmental variation on health outcome is often mediated by differences in gene transcription [8]. The methylation of cytosine at CpG-sites is an important epigenetic modification that regulates transcription by affecting several mechanisms [9]. Variation in DNA methylation at CpG-sites has been associated with asthma and other atopic diseases [10–15]. This warrants further investigation of this type of epigenetic variation.

In order to identify molecular mechanisms related to asthma remission in the airways, we investigated whether variation in DNA methylation in bronchial biopsies is associated with remission of asthma. Our main aim was to detect CpG-sites that are differentially methylated between well-characterised subjects in complete and clinical remission and subjects with persistent asthma or healthy controls. In addition, we investigated how differentially methylated CpG-sites associate with the expression levels of nearby genes in these bronchial biopsies, and related these to candidate molecular and cellular pathways. This furthers our ultimate aim to identify cellular mechanisms related to asthma remission in the airways.

Methods

Subjects

From previous studies [16, 17], biopsies were available from clinically well-characterised subjects with clinical remission (ClinR, n=33), complete remission (ComR, n=15) and persistent asthma (PersA, n=90). Biopsies from healthy controls (n=94) were available from another study. The study design and methods have been published previously [16, 17]. The subjects with ClinR and ComR had a documented diagnosis of asthma, confirmed by AHR testing [6, 17]. Asthma was defined as documented reversibility and/or AHR to histamine (provocative concentration causing a 20% fall in forced expiratory volume in 1 s (FEV_1) (PC_{20}) ⩽32 mg·mL^{-1}). Subjects were considered to be ClinR if they had not had an asthma attack or wheeze in the past 3 years, and did not use asthma medication (β-agonists and inhaled corticosteroids (ICS)). Asthma patients were considered to be ComR if, in addition to the previous criteria, they did not have AHR to both histamine and AMP (>32 mg·mL^{-1} in 30 s tidal breathing and >320 mg·mL^{-1} in 2 min tidal breathing, respectively) and had no signs of airflow obstruction (FEV_1 % predicted >80% pre-bronchodilator or >90% post-bronchodilator). The subjects with PersA were divided on the basis of use of ICS, since corticosteroids are known to affect gene expression and may act as a confounder on DNA methylation as well [18]. The non-asthmatic healthy controls were recruited from the NORM study (Study to Obtain Normal Values of Inflammatory Variables from Healthy Subjects; clinical trials number NCT00848406) and were all current smokers and never-smokers aged >18 years. Subjects were considered healthy if they had no respiratory symptoms, no history of respiratory disease and normal pulmonary function, defined as a post-bronchodilator FEV_1/forced vital capacity ratio higher than the lower limit of normal, absence of AHR to methacholine (PC_{20} >16 mg·mL^{-1}) and absence of FEV_1 reversibility (increase in FEV_1 after 400 µg salbutamol <10% pred). Healthy subjects were excluded if they had used inhaled or oral corticosteroids within the past 5 years, or during a total of 5 years of their lives. Biopsies from healthy subjects were collected between 2009 and 2012 and biopsies from all other subjects between 2005 and

This article has supplementary material available from erj.ersjournals.com

Received: 27 June 2019 | Accepted after revision: 18 Oct 2019

2007. All subjects were included in the same centre (University Medical Center Groningen, Groningen, the Netherlands) and bronchoscopies and processing of samples were performed by the same team. For RNA and DNA extraction and further processing, an even distribution of case–control status, age, sex and smoking status among batches was carefully controlled during all steps in the experimental design. Details on the DNA/RNA extraction, sample preparation and quality control are provided in the supplementary methods. The study protocol was approved by the local medical ethics committee. All subjects gave their written informed consent.

DNA methylation

DNA methylation levels were measured in bronchial biopsies using the Infinium HumanMethylation450 BeadChip array (450 k array) (Illumina, San Diego, CA, USA). Raw intensity data were processed using the minfi package [19]. Samples and probes failing quality control were removed and raw β values normalised using the dasen method as implemented in the wateRmelon package [20]. DNA methylation levels at each CpG-site were expressed as β-values, ranging from zero (no methylation) to one (complete methylation). A detailed description can be found in the supplementary methods. An overview of the sample dropout during quality control is shown in supplementary figure S1a.

RNA sequencing

RNA samples from airway wall biopsies were processed using the TruSeq Stranded Total RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). The cDNA fragment libraries were loaded unto an Illumina HiSeq2500 sequencer for paired-end sequencing (2×100 bp). Trimmed fastQ files where aligned to build b37 of the human reference genome using HISAT (version 0.1.5) and gene level quantification was performed by HTSeq (version 0.6.1p1) using Ensembl version 75 as gene annotation database [21]. A detailed description can be found in the supplementary methods. An overview of the sample dropout during quality control is shown in supplementary figure S1b.

Differential methylation analysis

Differential methylation between subject groups was assessed for each probe using robust linear modelling in the limma package [22, 23]. To remove heteroscedasticity, β-values were logit-transformed to M-values [24]. M-values were used as the dependent variable in all analyses. Age, sex and current smoking status were corrected for. In addition, the principal components explaining 95% of the variation in the control probes were used as covariates to correct for technical variation [25]. In addition, we analysed a model that included ex-smoker status in addition to current smoking, and found consistent results with the majority of differentially methylated regions (DMRs) remaining below the significance threshold. In remission *versus* asthma, of the 46 CpGs plus DMRs, 42 (91%) remained significant. In remission *versus* healthy controls from the 1491 CpGs plus DMRs, 1223 (82%) remained significant, but as this did not appreciably affect the results, these data are not presented. We analysed five subject categories: persistent asthma using ICS (PersA_ICS), persistent asthma not using ICS (PersA_no_ICS), ClinR, ComR and healthy controls. We pooled ClinR and ComR to increase sample size. We specified two contrasts: (ComR+ClinR) *versus* PersA_no_ICS and (ComR+ClinR) *versus* healthy controls. We refer to those contrasts as “remission *versus* asthma” and “remission *versus* healthy”, respectively. To avoid the confounding of ICS use, we focused on the group of asthmatics without ICS. We applied correction for multiple testing by controlling the false discovery rate at 5% using the Benjamini–Hochberg method [26]. The most significant hits of individual CpG-sites were annotated to the nearest gene using information from the IlluminaHumanMethylation450kanno.ilmn12.hg19 annotation package, or when no annotation was provided we searched the University of California Santa Cruz genome browser (https://genome.ucsc.edu) [27].

Differentially methylated regions

Regions of correlated CpG-sites with differential methylation were identified using the comb-p v0.48 command line tool and python library [28]. We used the pipeline with seed setting p=0.05 and distance 750 bp. The region filter settings were n=2 and a Sidak-corrected p=0.05.

Association between DNA methylation and gene expression (eQTM mapping)

We used the MatrixEQTL package to determine whether DNA methylation levels at CpG-sites were associated with gene expression levels of nearby genes within 1 Mb of the CpG-site [29]. Raw read counts from RNA sequencing were normalised for library size using the weighted trimmed mean of M-values method as implemented in the edgeR package [30]. The normalised counts were transformed to log2 (counts per million). We used methylation M-values as the explanatory variable and age, sex and current smoking status were set as covariables.

Cell type composition

Surrogate variable analysis provides estimates of latent variables, which is recommended as substitute for cell type deconvolution [31]. A detailed description can be found in the supplementary methods. We performed this analysis for all false discovery rate-significant differentially methylated CpGs from remission *versus* asthma and the 10 most differentially methylated CpGs from remission *versus* healthy.

We also took a targeted approach to the analysis of cell type composition, by extending the method proposed by Xu *et al.* [13]. For the same set of CpG-sites as used in the surrogate variable analysis, we performed genome-wide mapping of DNA methylation to gene expression. For each CpG-site, the resulting set of test statistics was used as input for mean-rank gene set enrichment analysis as implemented in limma [32]. We tested enrichment of 20 gene sets expressed by the cell type clusters taken from single cell sequencing results of airway wall biopsies as reported in Vieira Braga *et al.* [33].

Results

Patient characteristics

We obtained DNA methylation data for 179 samples, after excluding 53 samples that were lost during processing or excluded during the quality control procedure (supplementary figure S1). Of the included samples, 12 subjects had ComR, 27 subjects ClinR, 70 subjects PersA and 70 subjects were healthy controls. Of the PersA subjects, 26 were not using ICS. The patient characteristics of the five subject groups are presented in table 1.

Differential methylation in remission versus persistent asthma

Methylation levels at four individual CpG-sites and 42 regions were significantly different in remission *versus* asthma (tables 2 and 3). The most significant site was cg08364654, which showed 6% lower methylation in remission subjects than in asthmatics (figure 1a). This CpG is also part of the most significant DMR on chromosome 3 (table 3). All CpGs in the DMR are located in the *KRBOX1* gene, but we did not find these CpGs to be associated with *KRBOX1* expression. Instead, a higher methylation was associated (most significantly at cg22714811 at $p=7.9\cdot10^{-4}$) with increased expression of the atypical chemokine receptor 2 (*ACKR2*) gene (figures 2a and 3). The second CpG-site (cg23805470) also had 2% lower methylation in subjects in remission compared to persistent asthmatics. Methylation of cg23805470 was not significantly associated with expression of the gene it resides in (*tenascin XB*, *TNXB*) or any other nearby gene. The third site cg13525448 lies in an exon of *ladybird homeobox 1* (*LBX1*), but was not

TABLE 1 Patient characteristics

	PersA_ICS	PersA_no_ICS	ClinR	ComR	Healthy controls	H_0: no differences between groups. test-statistic; df; p-value
Subjects n	44	26	27	12	70	X^2=55.2; df=4; p≪0.01
PC_{20}AMP mg·mL^{-1}	64.9 (0.02–4871.5)	47.4 (0.02–3811.4)	700.2 (0.02–5090.7)	639.0 (414.8–5635.8)	2674.9 (40–5612.4)	F_{4-174}=10.9; P(emp)=0.001
FEV_1 % pred	83.1±3.02	83.3±2.39	90.8±2.70	101.8±4.05	101.0±1.34	F_{4-174}=14.0– p≪0.001
ICS dose µg·day^{-1}	800 (28– 2000)	NA	NA	NA	NA	NA
β-agonist use	40 (91)	14 (54)	0 (0)	0 (0)	0 (0)	Chi-squared=131.1; df=4; p≪0.01
Male/female	20/24	15/11	12/15	6/6	40/30	Chi-squared=2.47– df=4– p=0.650
Age years	48.8±1.85	46.3±2.44	47.5±2.33	45.3±4.78	39.5±2.03	F_{4-174}=3.37; p=0.011
FEV_1/FVC %	69.6±1.68	70.1±1.58	75.6±1.55	78.2±2.37	79.7±0.76	F_{4-174}=12.2; p≪0.001
Reversibility FEV_1 %	8.64±0.97	9.20±1.24	6.57±0.85	4.38±1.30	3.46±0.34	F_{4-173}=11.4– p≪0.001
Atopy (skin-prick)	NA	NA	NA	NA	25 (36)	NA
Atopy (phadiatop)	31 (70)	19 (73)	15 (56)	8 (67)	NA	Chi-squared=3.95– df=3– p=0.267
Smoking status						Chi-squared=43.8; df=8; p≪0.001
Current-smoker	6	10	4	4	33	
Ex-smoker	21	3	11	1	3	
Never-smoker	17	13	12	7	34	
Blood eosinophils ×10^9·L^{-1}	0.195 (0.05–0.78)	0.260 (0.00–0.90)	0.170 (0.00–0.62)	0.115 (0.00–0.28)	0.155 (0.02–0.43)	F_{4-174}=3.373; p=0.011

Data are presented as n, median (range) or mean±SEM, unless otherwise stated. PersA: persistent asthma; ICS: inhaled corticosteroids; ClinR: clinical remission; ComR: complete remission; df: degrees of freedom; FEV_1: forced expiratory volume in 1 s; PC_{20}: provocative concentration causing a 20% fall in FEV_1; FVC: forced vital capacity. Frequency data were analysed using Chi-squared tests of independence. Differences in continuous variables were tested with one-way ANOVA. Variables were transformed, if necessary, to satisfy the assumption of normality. For FEV_1 and FEV_1/FVC ratio ANOVA was performed on z-values (GLI2012_DataConversion software). When PC_{20}AMP was not reached at the highest dose, values were obtained by extrapolating from an exponential decay dose response curve. No appropriate transformation for PC_{20}AMP was found, and these data were analysed by one-way ANOVA based on a permutated distribution of F with 1000 permutations [41].

https://doi.org/10.1183/13993003.01280-2019

TABLE 2 Toptables of differentially methylated CpGs

	Genomic position	UCSC RefGene name	UCSC RefGene group	Correlated expression	logFC	Moderated t-value	p-value	Adjusted p-value	SVA corrected p-value
Remission *versus* asthma									
cg08364654	chr3:42,978,180				0.83	6.23	3.5×10^{-9}	0.0015	6.5×10^{-8}
cg23805470	chr6:32,056,820	TNXB	Body		0.60	6.07	7.9×10^{-9}	0.0017	1.1×10^{-7}
cg13525448	chr10:102,986,601				–0.48	–5.26	4.4×10^{-7}	0.0466	3.2×10^{-7}
cg00741675	chr4:967,327	DGKQ	5′UTR; 1st exon		2.96	5.28	3.8×10^{-7}	0.0466	4.5×10^{-4}
Remission *versus* healthy									
cg04886217	chr16:31,821,538				0.63	6.56	6.1×10^{-10}	0.0003	2.8×10^{-8}
cg06947286	chr5:131,596,602	PDLIM4	Body	PDLIM4; P4HA2; UQCRQ; HSPA4; KIF3A; ACSL6; ZCCHC10	–0.60	–6.20	4.2×10^{-9}	0.0006	6.8×10^{-5}
cg01716603	chr17:37,029,974	LASP1	Body	SRCIN1; PLXDC1; ARHGAP23; HNF1B; LINC00672; PIP4K2B; MLLT6; NEUROD2; STARD3; PCGF2; ENSG00000214546; CACNB1	–0.83	-6.19	4.3×10^{-9}	0.0006	1.7×10^{-5}
cg23932332	chr1:221,911,278	DUSP10	TSS1500; body; 5′UTR	ENSG00000225265; FAM177B; MARC2; DUSP10	–0.87	–6.01	1.1×10^{-8}	0.0011	0.0062
cg18763536	chr12:11,812,062	ETV6	Body	MANSC1; CREBL2; ETV6	–0.44	–5.97	1.4×10^{-8}	0.0011	1.8×10^{-5}
cg23916878	chr6:44,011,187			DLK2; LRRC73; RSPH9; DNPH1; TCTE1; SPATS1; PTK7; ENSG00000272442	–0.68	–5.94	1.6×10^{-8}	0.0011	2.9×10^{-5}
cg20001791	chr6:16,239,799	GMPR	Body	GMPR; STMND1	–0.66	–5.88	2.1×10^{-8}	0.0013	4.2×10^{-7}
cg08307963	chr1:147,245,485	GJA5	TSS200	BCL9	0.54	5.78	3.4×10^{-8}	0.0017	2.0×10^{-7}
cg10986462	chr10:135,340,539	CYP2E1	TSS1500	FUOM	–1.12	–5.77	3.6×10^{-8}	0.0017	1.2×10^{-5}
cg07224931	chr12:121,130,567	MLEC	Body	DYNLL1-AS1; CIT; MORN3; UNC119B; PXN; DYNLL1	–0.53	–5.73	4.5×10^{-8}	0.0017	7.1×10^{-5}

The table shows the probe ID, with its genomic position (hg19), the name of any nearby genes and its position relative to that gene, proximate genes with associated expression, the $\log_2$ fold change (logFC) of the β-values, the moderated t-value and its associated probability of the null hypothesis of no difference between the groups, as well as the adjusted p-value by Benjamini–Hochberg correction controlling the false discovery rate at 5%. The rightmost column lists the p-values of the analysis where surrogate variables were added as covariates (see text). Positive logFC indicates higher methylation in asthmatics than remission subjects and higher methylation in remission subjects than healthy controls. UCSC: University of California Santa Clara; SVA: surrogate variable analysis; UTR: untranslated region.

https://doi.org/10.1183/13993003.01280-2019

TABLE 3 Toptables of differentially methylated regions (DMRs)

	n_probes	z_sidak_p	t.sum	refGene_name	refGene_feature	Associated expression
Remission *versus* asthma						
3:42977776-42978248	8	9.6×10^{-20}	34.0	KRBOX1-AS1; KRBOX1	nc_exon;TSS+intron+exon+utr5	ACKR2
19:50666237-50666552	6	2.4×10^{-11}	22.5	IZUMO2	TSS+utr5+cds	RPL13A
11:124746753-124747263	6	7.7×10^{-8}	18.7	ROBO3	intron+cds	ROBO3
17:1094028-1094555	7	8.6×10^{-8}	-20.7	ABR	intron	
8:1049166-1049477	5	5.8×10^{-6}	−15.4	DLGAP2	nc_intron;intron	
15:45671000-45671347	10	3.1×10^{-5}	24.1	GATM	intron+utr5	GATM;DUOX2;SPG11;B2M;SHF
3:142666107-142666476	4	1.3×10^{-4}	12.2	LOC100507389	intergenic	GK5
15:75251490-75251733	3	2.3×10^{-4}	10.2	RPP25	intergenic	CCDC33;ISLR
1:119532043-119532352	8	2.7×10^{-4}	19.0	TBX15	TSS+exon+utr5	
11:1891931-1892307	10	3.3×10^{-4}	20.7	LSP1	TSS+intron+exon+utr5;intron+utr5;TSS +exon+utr5;intron	C11orf21;LSP1;TSPAN32;C11orf89
Remission *versus* healthy						
10:135340444-135341025	9	3.4×10^{-12}	−32.1	CYP2E1	TSS+utr5+cds	FUOM;TUBGCP2;ECHS1;SPRN
16:4714079-4714815	7	3.9×10^{-10}	−24.4	MGRN1	intron+cds;nc_intron+nc_exon	C16orf71;CORO7;SMIM22;GLYR1;CDIP1;C16orf89; ENSG00000266994;SEC14L5;GLIS2;TFAP4;ENSG00000262686
1:147245484-147245626	3	4.1×10^{-10}	14.8	NBPF19	intron	BCL9
11:69919474-69920054	5	1.8×10^{-9}	18.6	LINC02584	nc_intron	SHANK2;SHANK2-AS1;SHANK2-AS3;SHANK2-AS2;ENSG00000227726;CTTN
1:9380565-9381431	5	3.9×10^{-9}	−18.1	SPSB1	intron+utr5	ENSG00000234546;SPSB1;TMEM201;SLC45A1;H6PD;PIK3CD;CLSTN1;KIF1B;RBP7; SLC2A5
8:145018815-145019839	17	7.4×10^{-9}	−43.5	PLEC;MIR661;PLEC	intron;nc_gene;TSS+intron+utr5+cds	FAM83H-AS1;EEF1D;RHPN1-AS1;ENSG00000255224;MIR4664;TOP1MT; ARHGAP39;ZNF34;TONSL;ENSG00000254973; MAPK15;GPIHBP1;FAM83H;GPT;EXOSC4;TSTA3;C8orf31
13:114812176-114813015	6	9.5×10^{-9}	−20.1	RASA3	intron+utr5	TMCO3;CDC16;ADPRHL1;TMEM255B;R ASA3;DCUN1D2;GAS6;PCID2;TFDP1
14:105932811-105933283	4	1.1×10^{-8}	−16.1	MTA1	intron+exon+utr3;intron+cds	C14orf79;ADSSL1;INF2;IGHV3-15;IGHV3-11;IGHV3-21;IGHV3-23;IGHM;IGHA1;KIAA0125; IGHV4-39;IGHJ3P;JAG2;IGHG2;IGHA2
12:132219327-132219529	2	1.9×10^{-0}	−10.6	SFSWAP	intron	POLE;STX2;GPR133;DDX51;ULK1;MMP17
14:94392717-94392932	4	2.4×10^{-8}	−16.4	FAM181A; FAM181A-AS1	intron;nc_exon;intron+utr5	PPP4R4;UNC79;ASB2;C14orf142;ENSG00000258987;FAM181A-AS1; PRIMA1;SERPINA6;MOAP1;SERPINA10;SERPINA1;FAM181A;IFI27L2

The table shows the genomic position (hg19), the number of probes in the DMR, the one-step Sidak corrected p-value, the summed t-value, the name of any nearby genes and CpG positions relative to that gene and proximate genes with associated expression. Positive t.sum values indicate higher methylation in asthmatics than remission subjects and higher methylation in remission subjects than healthy controls.

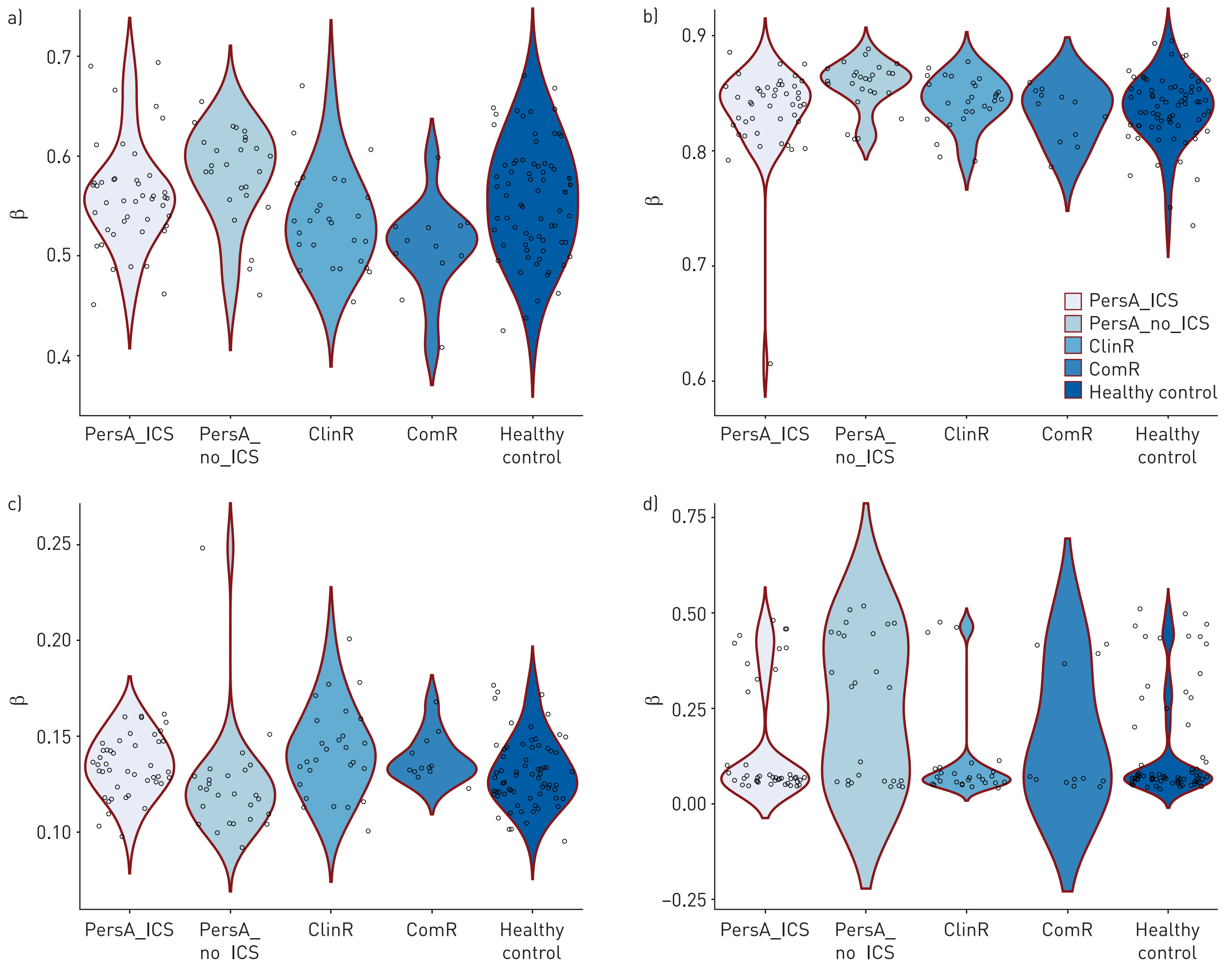

FIGURE 1 Violin plots of highly differentially methylated CpG-sites in the "asthma *versus* remission" contrast. a) cg08364654 in the *KRBOX1* gene; b) cg23805470 in the *TNXB* gene; c) cg13525448; and d) cg00741675 in the *DGKQ* gene. PersA_ICS: persistent asthma with inhaled corticosteroid (ICS) use; PersA_no_ICS: persistent asthma without ICS use; ClinR: clinical remission; ComR: complete remission.

associated with gene expression in *cis*. It had 2% higher methylation in remission than in persistent asthma (figure 1c). The fourth CpG-site (cg00741675) had 11% lower methylation in remission and is located in the 5′ untranslated region of the diacylglycerol kinase, θ 110 kDa (*DGKQ*) gene. It did not associate with gene expression at the p=0.001 cut-off, but is positively associated with gene expression of *DGKQ* at p<0.004 (figure 2b). The second significant DMR is located at chromosome 19 and associated with expression of *RPL13A*. The encoded product is a component of the interferon-γ-activated inhibitor of translation complex, which plays a role in the repression of inflammatory genes and contributes to the resolution of chronic inflammation [34]. Among the four significant individual CpG-sites from the contrast remission *versus* asthma (table 2), all remained significantly differentially methylated (at p<0.001) after correction for the first five surrogate variables, indicating that the differences in methylation were not driven by cell type composition (table 2 and supplementary figure S2a). Significance of cg00741675 dropped sharply, but for all four CpGs log fold changes remained similar after adjustment.

Differential methylation in remission versus *healthy controls*

We found 1163 individual CpG-sites and 328 regions to be differentially methylated between remission subjects and healthy controls, with one-third of the individual sites and 20% of the DMRs being higher methylated in remission (tables 2 and 3 show the top 10; complete list in supplementary tables S1B and S2B). Of the 1163 significant CpG-sites, 167 were located in a DMR. Correction for surrogate variables resulted in similar log fold changes, but did appreciably affect the significance of the 10 most significant

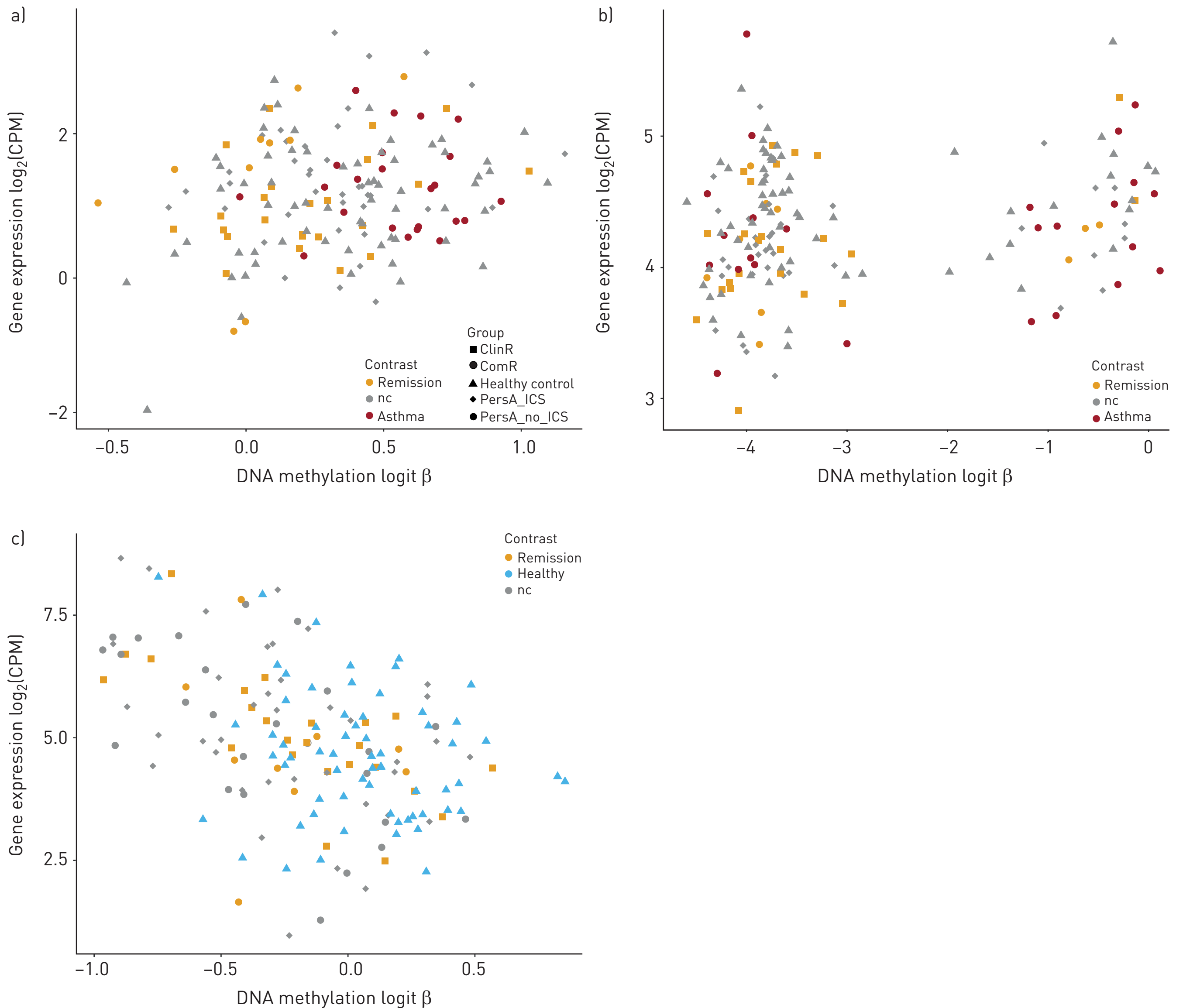

FIGURE 2 Correlation of DNA methylation at differentially methylated CpG-sites with gene expression of nearby genes. a) cg08364654 with ACKR2 expression; b) cg00741675 with DGKQ expression; c) cg18091275 with NOS2 expression. The groups used in the contrasts (either "remission *versus* asthma" or "remission *versus* healthy") are marked with non-grey colours. nc: no contrast; PersA_ICS: persistent asthma with inhaled corticosteroid (ICS) use; PersA_no_ICS: persistent asthma without ICS use; ClinR: clinical remission; ComR: complete remission.

CpG-sites, suggesting an impact of cell type composition in the contrast remission *versus* healthy (supplementary figure S2).

To further investigate effects of cell type composition we identified cell type specific gene expression profiles by associating methylation levels of the 10 most significant CpG-sites to genome-wide gene expression levels. Then we compared associated genes to gene sets characteristic for 20 major clusters of cells found in bronchial biopsies based on single cell sequencing data [33]. All 10 tested CpG-sites from the contrast remission *versus* healthy showed strong enrichment for association with genes belonging to multiple sets of cell types (supplementary table S3). The highest enriched groups throughout were ciliated epithelial cells and basal cells. However, several CpGs were most significantly enriched for other groups: smooth muscle cells (cg06947286 and cg10986462), fibroblasts (cg23916878), neutrophils (cg18763536) and T-cells (cg08307963) (figure 4a). Nine out of 10 tested CpG-sites from the contrast remission *versus* healthy controls showed strong enrichment ($p \ll 0.001$) for highly ranking genes from ciliated epithelium (figure 4). We tested whether the enrichment was due to the underlying correlation structure of our dataset, but this did not fully account for our findings. For comparison, only 49% of a randomly selected sample of CpG-sites has enrichment for this gene set at this significance level. In addition, five out of 10

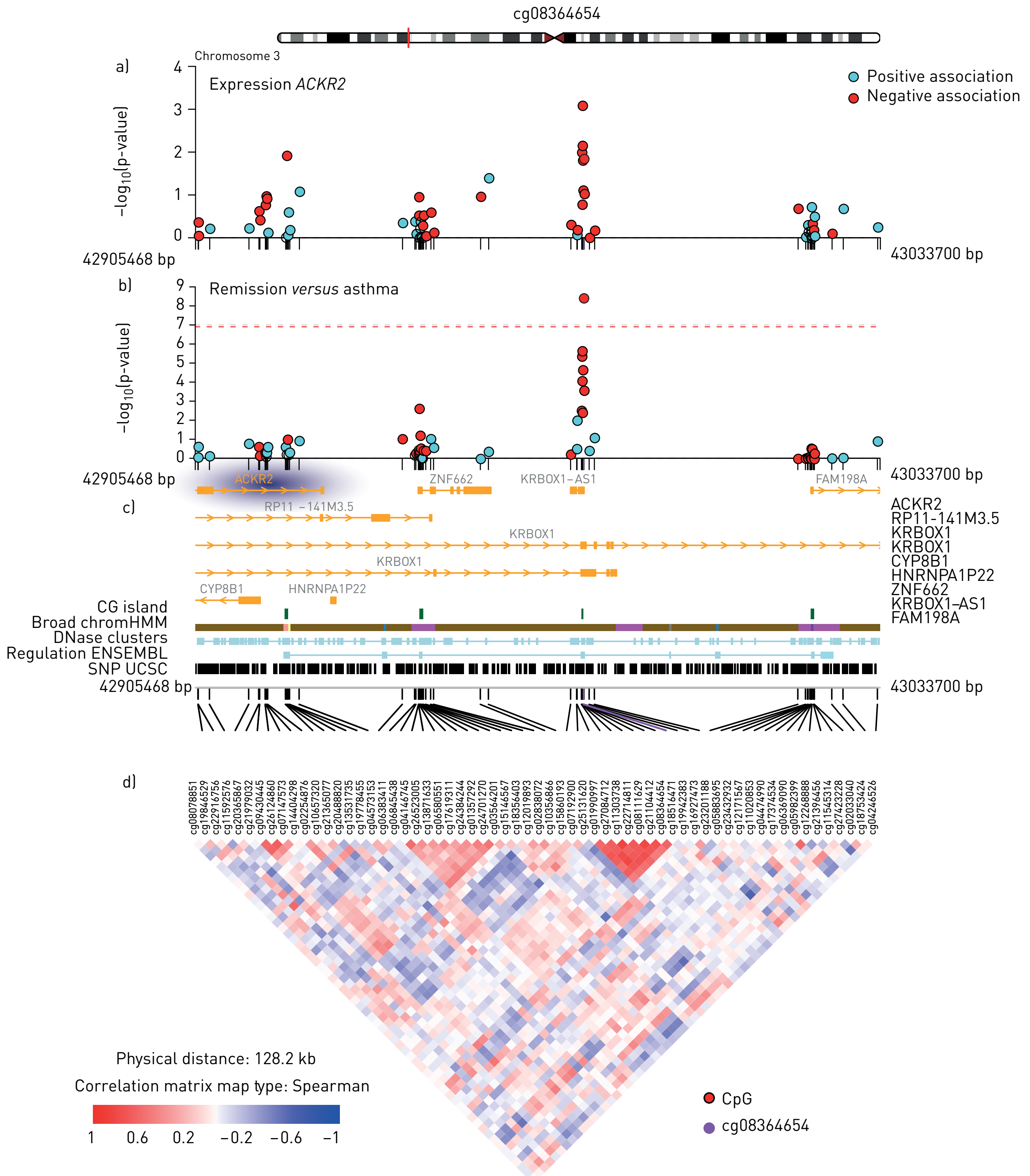

FIGURE 3 CoMET plot of the genomic region of cg08364654. a) The strength of correlation with expression of the *ACKR2* gene for all the CpG-sites in the interval; b) the strength of association in the remission *versus* asthma contrast; c) default coMET annotation tracks: ENSEMBL genes, CpG islands, chromatin-state (Broad institute), DNAse-sensitive regions and SNPs (see coMET documentation for detailed description: https://bioconductor.org/packages/release/bioc/html/coMET.html); d) correlation matrix of DNA methylation among all plotted CpG-sites.

tested CpG-sites belong to the top 1% enrichment as judged by comparison to p-values in the randomly selected sample. Figure 4b shows the ranking of correlated genes of a representative CpG-site from this set (cg07224931). Nearly all genes in the set related to ciliated epithelium have higher gene expression in remission *versus* healthy controls (99%).

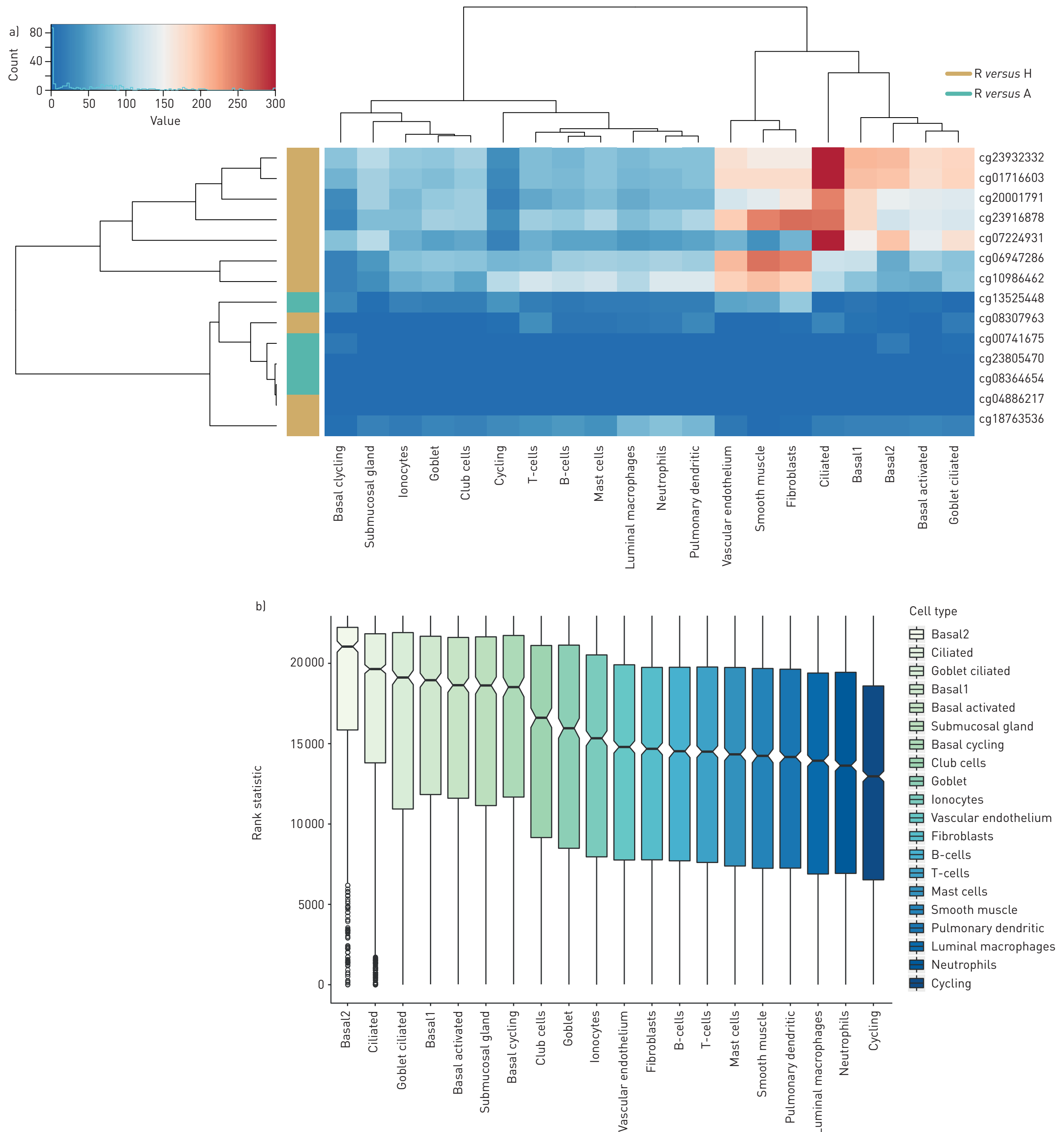

FIGURE 4 Effect of cell type composition on average methylation levels. Effect of cell type composition was gauged by genome wide correlations with gene expression. a) Heatmap and hierarchical clustering of $-\log_{10}$(p-values) of mean-rank gene set tests on the test statistics of DNA methylation–gene expression associations. The CpG-sites are the four most significant sites from the contrast remission *versus* asthma (R *versus* A) and the 10 most significant sites from the contrast remission *versus* healthy (R *versus* H). The gene sets distinguish the four main clusters of epithelial cell types in single cell sequencing. See text for details. b) Ranks of test statistic for a selected CpG-site (cg07224931) demonstrates the enrichment of highly ranking genes of the ciliated epithelium.

Discussion

We found four individual CpG-sites and 42 regions to be differentially methylated between subjects in remission and persistent asthmatics. Of the individual CpGs two associated with gene expression of *ACKR2* and *DGKQ*, respectively. These four CpGs were not strongly related to cell type composition of the airway wall biopsy. There were 1163 CpG-sites and 328 regions differentially methylated between

remission subjects and healthy controls. Many of the highly significant CpG-sites that distinguish remission subjects from healthy controls were associated with increases in the expression of genes associated with ciliated epithelial cells, consistent with cell type composition differences in remission *versus* healthy controls. We found differences in absolute methylation between groups between 2% and 11%. It should be noted that even modest differences can be relevant when they reflect changes within a specific cell type.

The most significant CpG site, as well as DMR, with lower methylation in the contrast remission *versus* asthma was associated with lower expression of *ACKR2* in subjects with remission. This gene contributes to resolution of inflammation by post-inflammatory clearance of chemokines in a mouse model [35]. Allergen-challenged mice that are deficient for *ACKR2* have more infiltrating cells in the airways, primarily dendritic cells and eosinophils. Interestingly, these *ACKR2*-deficient mice have less airway reactivity to methacholine [36]. We suggest that lower expression of *ACKR2* may play a role in the remission of asthma, by reducing AHR. Of interest, the lower methylation and *ACKR2* gene expression was paralleled by decreased airway reactivity in the *ACKR2*-deficient mice, but not by reduced inflammation as would be expected. The CpG-site in *TNXB* could not be related to expression of this gene, but tenascin genes are known to be important in airway remodelling [1]. Multiple CpG-sites in *TNXB* were less methylated in primary airway epithelium cells after exposure to interleukin-13, a type 2 cytokine and key mediator of airway inflammation and remodelling in asthma [15]. *TNXB* resides in the human leukocyte antigen (HLA) class II region, which is important for immune function [37], but we were unable to link differential methylation at cg23805470 to expression of HLA genes. CpG-site cg00741675 in *DGKQ* has lower methylation in remission subjects when compared to persistent asthmatics. *DGKQ* encodes a member of the family of diacylglycerol kinases (DGKs) that attenuate levels of the second messenger diacylglycerol in cells by converting it to phosphatidic acid [38]. The precise mechanism by which it may be involved in the remission of asthma is not clear, but members of the DGK family are expressed in immune cells, and thus may be required for proper cellular signalling in inflammatory processes [39].

Subjects in remission have a methylation profile that is distinct from that of persistent asthmatics, which may reflect differential regulation of immune functions. However, their methylation profile does not resemble the healthy situation either. Individuals with complete asthma remission retain characteristics that differ from healthy subjects without a history of asthma, for example, persistent airway remodelling as shown by increased basal membrane thickness [17]. Analogous to this situation, we find methylation differences between remission and healthy controls that do not occur between remission and asthma. We found that 21 out of 42 DMRs and three out of four differentially methylated CpGs in remission *versus* asthma overlapped with DMRs found in asthma *versus* healthy controls. For remission *versus* healthy controls this concerned 126 out of 327 DMRs and 310 out of 1163 differentially methylated CpGs. This clearly shows that differential methylation in both the remission *versus* asthma and the remission *versus* healthy controls contrasts capture aspects of methylation differences between persistent asthmatics and healthy subjects, consistent with the intermediate position of remission subjects. A possible explanation is that subjects in asthma remission still display an epigenetic fingerprint of asthma. A similar lingering difference in methylation profiles was reported for ex-smokers [40]. Alternatively, this may reflect an inherent predisposition to developing asthma, which puts remission subjects at an increased risk for relapse.

A strength of our design is the use of bronchial biopsies, which enabled us to study epigenetic profiles in the relevant tissue. In addition, we included healthy subjects and well-characterised subjects in clinical and complete remission. All remission subjects had previous objective confirmation of their asthma diagnosis with detailed follow-up. Finally, we added functional information to differentially methylated CpG-sites by their association with the expression levels of nearby genes, using the same biopsy as used for determination of methylation levels. One limitation is the absence of an independent replication cohort in which bronchial biopsies are available for remission subjects and healthy controls. To the best of our knowledge, no comparable replication cohort with a similar design is available. Another issue is that differential methylation among subject groups was partly driven by differences in cell type composition, which hampers identification of specific asthma or remission related CpG-sites. By annotating the associated CpGs with gene expression, we show that subjects in remission differ from healthy controls in methylation levels at CpG-sites associated with multiple cell types, and that there is strong enrichment of CpGs in genes associated with ciliated epithelium. A possible interpretation is that changes in the cell type composition, in particular the proportion of ciliated epithelium in remission subjects explains part of the differential methylation between remission subjects and healthy controls. This is supported by the inflation of the p-values in this particular contrast. The prominent signal from epithelium may reflect the abundance of this cell type in biopsies, and the importance of other cell types in asthma and remission should not be neglected, but it emphasises that the airways of subjects in remission still have not come to resemble the healthy profile.

Acknowledgements: We wish to thank Pieter van der Vlies, Bahram Sanjabi, Desiree Brandenburg-Weening and Harold de Bruin (University Medical Center Groningen, Groningen, the Netherlands) for their valuable work in the DNA extraction, sample preparation and hybridisation.

Author contributions: C.J. Vermeulen performed differential methylation analysis and drafted and finalised the manuscript, C-J. Xu and J.M. Vonk were involved in data management, data analysis and editing the manuscript, N.H.T. ten Hacken, W. Timens, I.H. Heijink, M.C. Nawijn, A. J. van Oosterhout, K. Affleck and G.H. Koppelman provided guidance and technical knowledge and edited the manuscript, J. Boekhoudt performed eQTM analysis, M. Weckmann contributed expression and methylation data, M. van den Berge, N.H.T. ten Hacken, W. Timens and J.M. Vonk were involved in the design of the study and M. van den Berge drafted parts of the manuscript. All authors approved the final manuscript.

Conflict of interest: C.J. Vermeulen reports grants from GSK, during the conduct of the study. C-J. Xu has nothing to disclose. J.M. Vonk has nothing to disclose. N.H.T. ten Hacken has nothing to disclose. W. Timens reports personal fees from Pfizer, GSK, Chiesi, Roche Diagnostics/Ventana, Biotest, Merck Sharp Dohme, Novartis, Lilly Oncology, Boehringer Ingelheim, AstraZeneca, Bristol-Myers-Squibb and AbbVie, and grants from the Dutch Asthma Fund, outside the submitted work. I.H. Heijink has nothing to disclose. M.C. Nawijn reports grants from GSK and the Lung Foundation of the Netherlands, outside the submitted work. J. Boekhoudt has nothing to disclose. A.J. van Oosterhout reports and holds GSK shares. K. Affleck is an employee of GSK and a member of the company share schemes. M. Weckmann has nothing to disclose. G.H. Koppelman reports grants from the Lung Foundation of the Netherlands, TEVA, GSK, Vertex and Ubbo Emmius Foundation, outside the submitted work; and has served on an international advisory board for GSK (money to institution). M. van den Berge reports grants paid to his university from AstraZeneca, TEVA, GSK and Chiesi, outside the submitted work.

Support statement: Data were generated as part of a Scientific Research Collaboration funded by GSK (COL100037293). The submitted work is co-financed by the Ministry of Economic Affairs and Climate Policy by means of the PPP. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Holgate ST, Wenzel S, Postma DS, *et al.* Asthma. *Nat Rev Dis Primers* 2015; 1; 15025.
2 Upham JW, James AL. Remission of asthma: the next therapeutic frontier? *Pharmacol Ther* 2011; 130: 38–45.
3 Holm M, Omenaas E, Gíslason T, *et al.* Remission of asthma: a prospective longitudinal study from northern Europe (RHINE study). *Eur Respir J* 2007; 30: 62–65.
4 Carpaij OA, Nieuwenhuis MAE, Koppelman GH, *et al.* Childhood factors associated with complete and clinical asthma remission at 25 and 49 years. *Eur Respir J* 2017; 49: 1601974.
5 Boulet LP, Turcotte H, Brochu A. Persistence of airway obstruction and hyperresponsiveness in subjects with asthma remission. *Chest* 1994; 105: 1024–1031.
6 Vonk JM, Postma DS, Boezen HM, *et al.* Childhood factors associated with asthma remission after 30 year follow up. *Thorax* 2004; 59: 925–929.
7 Vonk JM, Nieuwenhuis MAE, Dijk FN, *et al.* Novel genes and insights in complete asthma remission: a genome-wide association study on clinical and complete asthma remission. *Clin Exp Allergy* 2018; 48: 1286–1296.
8 Edwards SL, Beesley J, French JD, *et al.* Beyond GWASs: illuminating the dark road from association to function. *Am J Hum Genet* 2013; 93: 779–797.
9 Reddington JP, Pennings S, Meehan RR. Non-canonical functions of the DNA methylome in gene regulation. *Biochem J* 2013; 451: 13–23.
10 DeVries A, Vercelli D. Early predictors of asthma and allergy in children: the role of epigenetics. *Curr Opin Allergy Clin Immunol* 2015; 15: 435–439.
11 Sabounchi S, Bollyky J, Nadeau K. Review of environmental impact on the epigenetic regulation of atopic diseases. *Curr Allergy Asthma Rep* 2015; 15: 33.
12 Miller RL, Peden DB. Environmental effects on immune responses in patients with atopy and asthma. *J Allergy Clin Immunol* 2014; 134: 1001–1008.
13 Xu C-J, Söderhäll C, Bustamante M, *et al.* DNA methylation in childhood asthma: an epigenome-wide meta-analysis. *Lancet Respir Med* 2018; 6: 379–388.
14 Nicodemus-Johnson J, Myers RA, Sakabe NJ, *et al.* DNA methylation in lung cells is associated with asthma endotypes and genetic risk. *JCI Insight* 2016; 1: e90151.
15 Nicodemus-Johnson J, Naughton KA, Sudi J, *et al.* Genome-wide methylation study identifies an IL-13 induced epigenetic signature in asthmatic airways. *Am J Respir Crit Care Med* 2016; 193: 376–385.
16 Broekema M, Volbeda F, Timens W, *et al.* Airway eosinophilia in remission and progression of asthma: accumulation with a fast decline of FEV_1. *Respir Med* 2010; 104: 1254–1262.
17 Broekema M, Timens W, Vonk JM, *et al.* Persisting remodelling and less airway wall eosinophil activation in complete remission of asthma. *Am J Respir Crit Care Med* 2011; 183: 310–316.
18 van den Berge M, Steiling K, Timens W, *et al.* Airway gene expression in COPD is dynamic with inhaled corticosteroid treatment and reflects biological pathways associated with disease activity. *Thorax* 2014; 69: 14–23.
19 Aryee MJ, Jaffe AE, Corrada-Bravo H, *et al.* Genome analysis Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 2014; 30: 1363–1369.
20 Pidsley R, Y Wong CC, Volta M, *et al.* A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* 2013; 14: 293.
21 Anders S, Pyl PT, Huber W. HTSeq – a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015; 31: 166–169.
22 Smyth GK. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; 3: Article 3.
23 Ritchie ME, Phipson B, Wu D, *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47.

24 Du P, Zhang X, Huang C-C, *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010; 11: 587.
25 Lehne B, Drong AW, Loh M, *et al.* A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol* 2015; 16: 37.
26 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statistic Soc* 1995; 57: 289–300.
27 Kent WJ, Sugnet CW, Furey TS, *et al.* The human genome browser at UCSC. *Genome Res* 2002; 12: 996–1006.
28 Pedersen BS, Schwartz DA, Yang IV, *et al.* Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics* 2012; 28: 2986–2988.
29 Shabalin AA. Matrix eQT: ultra fast eQTL analysis *via* large matrix operations. *Bioinformatics* 2012; 28: 1353–1358.
30 McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 2012; 40: 4288–4297.
31 Teschendorff AE, Zheng SC. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. *Epigenomics* 2017; 9: 757–768.
32 Michaud J, Simpson KM, Escher R, *et al.* Integrative analysis of RUNX1 downstream pathways and target genes. *BMC Genomics* 2008; 9: 363.
33 Vieira Braga FA, Kar G, Berg M, *et al.* A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med* 2019; 25: 1153–1163.
34 Mukhopadhyay R, Jia J, Arif A, *et al.* The GAIT system: a gatekeeper of inflammatory gene expression. *Trends Biochem Sci* 2009; 34: 324–331.
35 Jamieson T, Cook DN, Nibbs RJB, *et al.* The chemokine receptor D6 limits the inflammatory response *in vivo*. *Nat Immunol* 2005; 6: 403–411.
36 Whitehead GS, Wang T, DeGraff LM, *et al.* The chemokine receptor D6 has opposing effects on allergic inflammation and airway reactivity. *Am J Respir Crit Care Med* 2007; 175: 243–249.
37 Kontakioti E, Domvri K, Papakosta D, *et al.* HLA and asthma phenotypes/endotypes: a review. *Hum Immunol* 2014; 75: 930–939.
38 Topham MK, Prescott SM. Mammalian diacylglycerol kinases, a family of lipid kinases with signaling functions. *J Biol Chem* 1999; 274: 11447–11450.
39 Yamamoto M, Tanaka T, Hozumi Y, *et al.* Expression of mRNAs for the diacylglycerol kinase family in immune cells during an inflammatory reaction. *Biomed Res* 2014; 35: 61–68.
40 Joehanes R, Just AC, Marioni RE, *et al.* Epigenetic signatures of cigarette smoking. *Circ Cardiovasc Genet* 2016; 9: 436–447.
41 Manly BFJ. Randomization, Bootstrap and Monte Carlo Methods in Biology. London, Chapman and Hall, 1997.