



Epigenetic Signatures Discriminate Patients With Primary Sclerosing Cholangitis and Ulcerative Colitis From Patients With Ulcerative Colitis

OPEN ACCESS

Edited by:

Cumali Efe,
Harran University, Turkey

Reviewed by:

Swapna Mahurkar-Joshi,
University of California, Los Angeles,
United States
Federica Facciotti,
European Institute of Oncology (IEO),
Italy

***Correspondence:**

Cyriel Y. Ponsioen
c.y.ponsioen@amsterdamumc.nl

[†]These authors have contributed
equally to this work and share
first authorship

[‡]These authors share last authorship

Specialty section:

This article was submitted to
Autoimmune and
Autoinflammatory Disorders,
a section of the journal
Frontiers in Immunology

Received: 21 December 2021

Accepted: 14 February 2022

Published: 16 March 2022

Citation:

de Krijger M, Hageman IL, Li Yim AYF,
Verhoeff J, Garcia Vallejo JJ,
van Hamersveld PHP, Levin E,
Hakvoort TBM, Wildenberg ME,
Henneman P, Ponsioen CY and
de Jonge WJ (2022) Epigenetic
Signatures Discriminate Patients
With Primary Sclerosing Cholangitis
and Ulcerative Colitis From Patients
With Ulcerative Colitis.
Front. Immunol. 13:840935.
doi: 10.3389/fimmu.2022.840935

Manon de Krijger^{1,2†}, Ishtu L. Hageman^{1,2†}, Andrew Y. F. Li Yim^{3†}, Jan Verhoeff^{1,4},
Juan J. Garcia Vallejo⁴, Patricia H. P. van Hamersveld¹, Evgeni Levin^{5,6},
Theodorus B. M. Hakvoort¹, Manon E. Wildenberg^{1,2}, Peter Henneman³,
Cyriel Y. Ponsioen^{2*‡} and Wouter J. de Jonge^{1,7‡}

¹ Tytgat Institute for Liver and Intestinal Research, Amsterdam Gastroenterology Endocrinology Metabolism, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands, ² Department of Gastroenterology and Hepatology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands, ³ Department of Clinical Genetics, Genome Diagnostics Laboratory, Amsterdam Reproduction and Development, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands, ⁴ Department of Molecular Cell Biology and Immunology, Amsterdam Infection & Immunity and Cancer Center Amsterdam, Amsterdam University Medical Centers, Free University of Amsterdam, Amsterdam, Netherlands, ⁵ Department of Vascular Medicine, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands, ⁶ Horaizon BV, Delft, Netherlands, ⁷ Department of Surgery, University Clinic of Bonn, Bonn, Germany

Background: Primary sclerosing cholangitis (PSC) is a chronic inflammatory liver disease affecting the intra- and extrahepatic bile ducts, and is strongly associated with ulcerative colitis (UC). In this study, we explored the peripheral blood DNA methylome and its immune cell composition in patients with PSC-UC, UC, and healthy controls (HC) with the aim to develop a predictive assay in distinguishing patients with PSC-UC from those with UC alone.

Methods: The peripheral blood DNA methylome of male patients with PSC and concomitant UC, UC and HCs was profiled using the Illumina HumanMethylation Infinium EPIC BeadChip (850K) array. Differentially methylated CpG position (DMP) and region (DMR) analyses were performed alongside gradient boosting classification analyses to discern PSC-UC from UC patients. As observed differences in the DNA methylome could be the result of differences in cellular populations, we additionally employed mass cytometry (CyTOF) to characterize the immune cell compositions.

Results: Genome wide methylation analysis did not reveal large differences between PSC-UC and UC patients nor HCs. Nonetheless, using gradient boosting we were capable of discerning PSC-UC from UC with an area under the receiver operator curve (AUROC) of 0.80. Four CpG sites annotated to the *MINJ2* gene were found to strongly contribute to the predictive performance. While CyTOF analyses corroborated the largely

similar blood cell composition among patients with PSC-UC, UC and HC, a higher abundance of myeloid cells was observed in UC compared to PSC-UC patients.

Conclusion: DNA methylation enables discerning PSC-UC from UC patients, with a potential for biomarker development.

Keywords: primary sclerosing cholangitis, ulcerative colitis, DNA methylation/methylome, peripheral blood, 850k methylation array, mass cytometry

INTRODUCTION

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease, characterized by inflammation and fibrosis of the intra- and extrahepatic bile ducts. The etiology of PSC is largely unknown and there is still no medical treatment with a proven benefit on disease progression (1). The male to female ratio is 2:1, and interestingly, up to 80% of PSC patients has concomitant inflammatory bowel disease (IBD), of which the majority has ulcerative colitis (UC) (2–4). Comparing patients with UC and patients with PSC-UC could give more insights in pathophysiological processes underlying PSC.

The co-occurrence of PSC and IBD has led to various hypotheses linking these two disease entities. One of these hypotheses is the aberrant gut-homing lymphocyte paradigm, which hypothesizes that circulating T-lymphocytes that are primed in the gut and express gut-homing molecules integrin $\alpha 4\beta 7$ and C-Chemokine Receptor 9 (CCR9), can infiltrate in the liver *via* aberrantly expressed adhesion molecules in PSC-affected liver (5, 6). This mechanism may be due to genetic predisposition of the host and/or epigenetic changes in circulating lymphocytes.

Over the past few years, several genome-wide association studies (GWAS) in PSC have identified multiple susceptibility loci for PSC-IBD, with none of the loci being associated with this gut-lymphocyte homing paradigm. Interestingly, PSC-IBD patients and IBD patients share only a limited number of risk loci, suggesting that the combination of PSC-IBD comprises a distinct disease manifestation (7–9). Nevertheless, genetic variation alone cannot account for total disease liability in PSC-IBD, emphasizing the role of internal and external exposures (the “exposome”) including epigenetic factors that may contribute to the disease (10).

Epigenetics comprises heritable processes that involve transcriptional regulation without changing the nucleotide sequence. One of the most studied epigenetic marks is DNA methylation, which represents the addition of a methyl to a base, typically a cytosine that is directly followed by a guanine (CpG). DNA methylation in the promotor region of a gene has been inversely associated with gene expression and is thought to prevent binding of transcription factors, thereby silencing gene expression (11, 12). Differences in DNA methylation of blood cells has been described previously in the context of IBD, where patients with Crohn’s disease (CD) differed from patients with UC (13) and healthy controls (14). Similar observations have been made in colonic tissue of IBD patients (13). Furthermore, patients with PSC with and without IBD share common methylation differences compared to controls (15). Recently, it

was reported that in patients with PSC the DNA methylation age, as estimated using the so-called Horvath clock, was higher than the chronological age (16, 17). This age acceleration has proven to have predictive properties regarding disease activity in different diseases, both based on peripheral blood as well as, amongst others, in liver tissue (16–18).

In this study, we hypothesized that the peripheral blood DNA methylome of patients with concomitant PSC-UC is distinct from that of patients with UC without PSC and healthy controls (HC). Accordingly, we investigated the DNA methylome and performed supervised classification analyses to see whether we could accurately discern PSC-UC from UC and identify CpG loci that contributed to this classification. To ascertain that discriminating differences in DNA methylation between PSC-UC and UC were not solely the result of underlying differences in cellular composition in these different disease states, we also performed mass cytometry to characterize the cellular heterogeneity.

MATERIALS AND METHODS

Patients

Patients included in the current study were selected from a well-characterized cohort of the ‘Epi PSC PBC project’, a large population-based cohort to study patients with cholestatic liver diseases [PSC and primary biliary cholangitis (PBC)] as well as IBD patients in the Netherlands (2). Patients with prior liver transplantation, colorectal carcinoma, cholangiocarcinoma or prior bowel surgeries were excluded. Out of 1183 cases, 18 patients with PSC-UC, 17 patients with UC, as well as 12 healthy controls (HC) were selected for DNA methylation analysis. Only male patients were included, and groups were matched for age, UC and medication use (**Table 1**).

For mass cytometry analysis, peripheral blood was collected from 10 patients with PSC-UC, 10 patients with UC and 9 healthy volunteers (**Table 1**). Again, only male patients were included, and biological use was an exclusion criterion to minimize influences on immunological cell distribution. Both protocols were approved by the Medical Ethics Committee at the Amsterdam UMC, University of Amsterdam (METC 06–267/E and METC 2018-050). All samples were collected with written informed consent.

DNA Isolation and *In Vitro* DNA Methylation Analysis

Genomic DNA was extracted from whole peripheral blood samples using an Autopure LS system (Qiagen, according to the

TABLE 1 | Baseline characteristics.

	Cohort 1				Cohort 2			
	PSC-UC (n=17)	UC (n=17)	Healthy controls (n=12)	p-value	PSC-UC (n=10)	UC (n=10)	Healthy controls (n=10)	p-value
Male [n (%)]	17 (100)	17 (100)	12 (100)	-	10 (100)	10 (100)	10 (100)	-
Age [median (IQR)]	41 (30-55)	34 (31-54)	39 (29-50)	0.911	40 (35-58)	58 (49-62)	30 (28-53)	0.022
Age at diagnosis PSC [median (IQR)]	33 (24-46)	-	-	-	33 (25-37)	-	-	-
Duration PSC in years [median (IQR)]	5 (2-12)	-	-	-	7 (6-15)	-	-	-
Age at diagnosis UC [median (IQR)]	27 (20-37)	29 (19-40)	-	0.865	35 (22-39)	30 (19-46)	-	0.796
Duration UC in years [median (IQR)]	16 (7-21)	11 (5-15)	-	0.218	9 (5-19)	22 (14-34)	-	0.019
Medication use [n (%)]								
UDCA	17 (100)	0 (0)	-	0.000	8 (80)	0 (0)	-	0.001
Anti-TNF- α	0 (0)	0 (0)	-	-	0 (0)	0 (0)	-	-
Mesalazine	17 (100)	17 (100)	-	-	9 (90)	9 (90)	-	1.000
Thiopurines	8 (47)	9 (53)	-	1.000	0 (0)	1 (10)	-	1.000
Montreal classification [n (%)]				0.144				0.164
Pancolitis	15 (88)	10 (59)	-	-	7 (70)	6 (60)	-	-
Left sided	2 (12)	5 (29)	-	-	1 (10)	4 (40)	-	-
Proctitis	0 (0)	1 (6)	-	-	2 (20)	0 (0)	-	-

For continuous variables the Kruskal Wallis test was used for comparing 3 groups, whereas the Mann Whitney-U test was used for comparing 2 groups. For dichotomous variables, Fisher's Exact test was used. P-value < 0.05 was considered statistically significant. PSC, primary sclerosing cholangitis; UC, ulcerative colitis; IQR, interquartile range; UDCA, Ursodeoxycholic acid.

manufacturer protocol) and stored at 4°C. Genomic DNA was sent to GenomeScan (Leiden) for bisulfite conversion and DNA methylation profiling. In short, bisulfite converts unmethylated cytosines into uracil, whereas the methylated cytosines remain unchanged. DNA methylation profiling was performed using the Illumina HumanMethylation Infinium EPIC BeadChip (850K) array, yielding the methylation status of approximately 850.000 CpG sites (19). Samples were randomized across the different chip slides to reduce possible batch effects.

In Silico DNA Methylation Analysis

Raw microarray data was imported in R statistical programming environment (v4.0.2) using the Bioconductor (v3.11) package *minfi* (v1.32.0) (20). Quality control was performed using *MethylAid* (v1.22.0) (21) and *shinyMethyl* (v1.24.0) (22), suggesting no apparent technical issues, such as slide-related batch effects. The raw data was then normalized either through functional normalization for differential methylation analyses using *limma* (v3.44.3) (23), or through *noob* for classification analyses (24). Preprocessing of the data consisted of removing probes that were associated to known genetic variants and cross-reactive probes (25). Allosome-associated probes were not specifically filtered out in the preprocessing steps as the cohort consisted of only males. Repetitive element-binding probes were not excluded as their effect on the percentage methylation was found to be minimal (26). Quality control based on principal component analysis resulted in the removal of one sample with PSC-UC due to its outlier status. Differential methylation analyses using *limma* were performed using the following design matrix:

$$methylation \sim age + steroids + UC : PSC,$$

where we compared PSC-UC with UC and PSC-UC with HC. The resultant p-values were adjusted for multiple testing using the Benjamini-Hochberg method. For the hypothesis-driven approach we extracted all probes of length N_{gene} associated to a particular gene of interest and calculated the aggregated p-value using the Fisher method. Next, we constructed a null distribution of p-values by aggregating p-values calculated from 5000 randomly selected stretches of N_{gene} consecutive of probes thereby capturing the correlated nature of DNA methylation that occurs in a particular region. By comparing the observed aggregated p-value with that of the null distribution, we obtained the final p-value. Visualizations were generated using *ggplot2* (v3.3.2) (27) and *ggbio* (v1.36.0) (28).

DNA Methylation Acceleration Analysis

The DNA methylation age in years was calculated using the Horvath clock as implemented in *wateRmelon* (v2.0.0) (17, 29). The difference between the DNA methylation age and the chronological age called the age acceleration.

DNA Methylation Blood Cell Estimation

The blood cell distribution as estimated from the DNA methylation data using the *estimateCellCounts2* function as implemented in the *FlowSorted.Blood.EPIC* (v1.12.1) package (30). In short, this package estimates the cellular composition per sample using a quadratic programming approach. Resultant

estimates were subsequently compared between groups using a two-way ANOVA test as implemented in R.

Gradient Boosting Analysis

Gradient boosting analysis was used to classify patients with PSC-UC from patients with UC. To identify the CpGs that contributed the most to the predictive performance, covered information disentanglement (CID) was implemented (31–34). In short, data was split up in a train (2/3) and test (1/3) set, whereupon the classifier was trained through repeated cross-validation on the training set. The performance of the resulting model was subsequently evaluated on the withheld test set. The area under the receiver operating characteristic (AUROC) scores were computed within each repetition of cross-validation and averaged for the final test AUROC. We reported the CID-derived feature importance scores for the CpG sites that contributed the most to the prediction model (31, 32).

DNA Sequencing NINJ2

DNA sequencing of *NINJ2*-associated loci was performed through Sanger sequencing using the BigDye Terminator v1.1 Cycle Sequencing kit. In short, primers were designed against the region encompassing the CpG loci of interest (**Supplementary Figure 5** and **Supplementary Table 4**). Genomic DNA of all patients included in the DNA methylation data was used as input for specific PCR amplification of the region of interest. A comprehensive overview of the PCR amplification protocol can be found in the supplementary methods. The resultant PCR products were subsequently sequenced at the Core Facility Genomics, Amsterdam UMC. Reads were then aligned using BioEdit and the CpGs of interest were analyzed for variants that might have introduced bases other than cytosine. Seven samples (n=3 PSC-UC and n=4 UC) were excluded after quality control.

mQTL Database Analysis

The mQTL database was interrogated for the *NINJ2*-associated predictor CpGs to identify the potential relationship with catalogued genetic variants (mQTL; <http://www.mqtl.db.org/>).

Bisulfite Conversion, PCR and Illumina MiSeq Sequencing

Technical validation of the *NINJ2*-associated loci of interest (cg26654770 and cg14911689) annotated to the *NINJ2* gene was performed through targeted amplicon sequence analysis using the Illumina MiSeq platform. Primers were designed in MethPrimer (35) (**Supplementary Table 4**). DNA samples from five PSC-UC, two UC and two HC patients (cohort 1, **Table 1**) were bisulfite converted according to standard protocol using the EZ DNA methylation kit (Zymo Research) (36). Amplicons were made from bisulfite converted DNA PCR and further purified with the Agencourt AMPure PCR purification kit (Beckman Coulter). During a second PCR, amplicons were elongated using TruSeq indices and Illumina sequence adapters, whereupon they were purified and pooled in stoichiometric amounts. Quality control of the amplicon length within the pools was performed using Agilent 2100 BioAnalyzer. DNA concentrations were measured using Qubit 2.0 Fluorometer (ThermoFisher) and

equalized to equimolar concentrations for all subject pools. MiSeq amplicon sequencing was then performed according to the standard protocol. Raw sequence data was mapped, aligned, and analyzed using Bismark (37) and visualized using Integrative Genomics viewer (v 2.3.57) against the bisulfite-converted human genome hg19. A minimum of 120 reads per sample-derived amplicon was deemed successful.

Quantitative Real-Time Polymerase Chain Reaction

Messenger RNA was extracted from frozen PBMCs using the Bioline ISOLATE II RNA mini kit (GC biotech B.V. Alphen a/d Rijn, the Netherlands) according to manufacturer's instructions. RNA concentration was measured using the Nanodrop 1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA). cDNA was synthesized using the Revertaid first strand cDNA synthesis kit (Fermentas, St. Leon-Rot, Germany). A quantitative polymerase chain reaction (qPCR) was performed using SensiFAST SYBR No-ROX (GC Biotech B.V.) on a BioRad (CFX96 real-time qPCR thermocycler). Resultant gene expression levels were calculated using LinRegPCR (38). After stability analysis in geNorm, two human reference genes were selected for normalization; Glyceraldehyde 3-Phosphate Dehydrogenase (*GAPDH*) and Hypoxanthine-guanine-fosforibosyl-transferase (*HPRT*) (39). Primers were either obtained by Qiagen or synthesized by Sigma (**Supplementary Table 4**).

Mass Cytometry

Cryopreserved PBMCs were thawed and washed with medium (RPMI+20% fetal bovine serum (FBS)) whereupon they were resuspended in PBS. For cellular viability assessment, single-cell suspensions were incubated with Cisplatin (5 μ M, Fluidigm) for 5 minutes and washed with cell staining buffer (CSB, Fluidigm). Cells were incubated with Human TruStain FcX Fc receptor blocking solution (Biolegend), after which cells were stained with a mix of metal-conjugated antibodies against cell surface markers (**Supplementary Table 5**, 'pre-fixation'), washed with CSB and fixed with 1.6% PFA. Cells were permeabilized by Maxpar Barcode Perm Buffer (Fluidigm), incubated with mass tag barcodes and stained with the remaining metal-conjugated antibodies (**Supplementary Table 5**, Fluidigm). For intracellular staining, cells were washed with Perm-S buffer (Fluidigm) and incubated with antibodies against CTLA-4 and CES-1 (**Supplementary Table 5**, 'nuclear staining'), washed and incubated with the corresponding secondary antibodies. Antibodies were fixated with 1.6% PFA/PBS, washed and incubated overnight with $^{191/193}$ Ir DNA intercalator (1:4000) diluted in Fix-and-Perm Buffer (Fluidigm). Acquisition was performed on the Cytometry by time of flight (CyTOF)3-Helios. Sample was diluted in H₂O and supplemented with 10% v/v of EQ Four Element Calibration beads (Fluidigm). After acquisition data was normalized and individual files were deconvoluted using the CyTOF software v6.7 functions.

Normalized.fcs files were uploaded in to Cytobank (40) for analysis and quality control. Viable CD45⁺ singlets were selected according to gating strategy previously described (41). Both batches of samples included a technical replicate. Potential batch effects were investigated through manual inspection of

marker distributions and overlap in clusterings. Different lineages (B-cells, CD4⁺ T-cells, CD8⁺ T-cells, myeloid cells and NK-cells) were clustered and color-coded using FlowSOM and subsequent manual annotation (42). Data is visualized using viSNE, a visualization tool for high-dimensional single-cell data based on the t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm (43).

Patient Characteristic Statistical Analyses

Patient characteristics are presented as median and interquartile range (IQR; 25th- 75th percentile). Dichotomous variables are presented as percentage (%) of the cohort. Differences were calculated with the chi-square test or Fisher's exact test for categorical variables. Numerical data were compared using a Mann-Whitney U test or One-way ANOVA, or a Kruskal-Wallis test with Dunn's correction for multiple testing. Statistical analysis was performed in SPSS statistical software for Windows version 26.0 (SPSS, Chicago, USA) or GraphPad Prism 8. A p-value <0.05 was considered statistically significant.

RESULTS

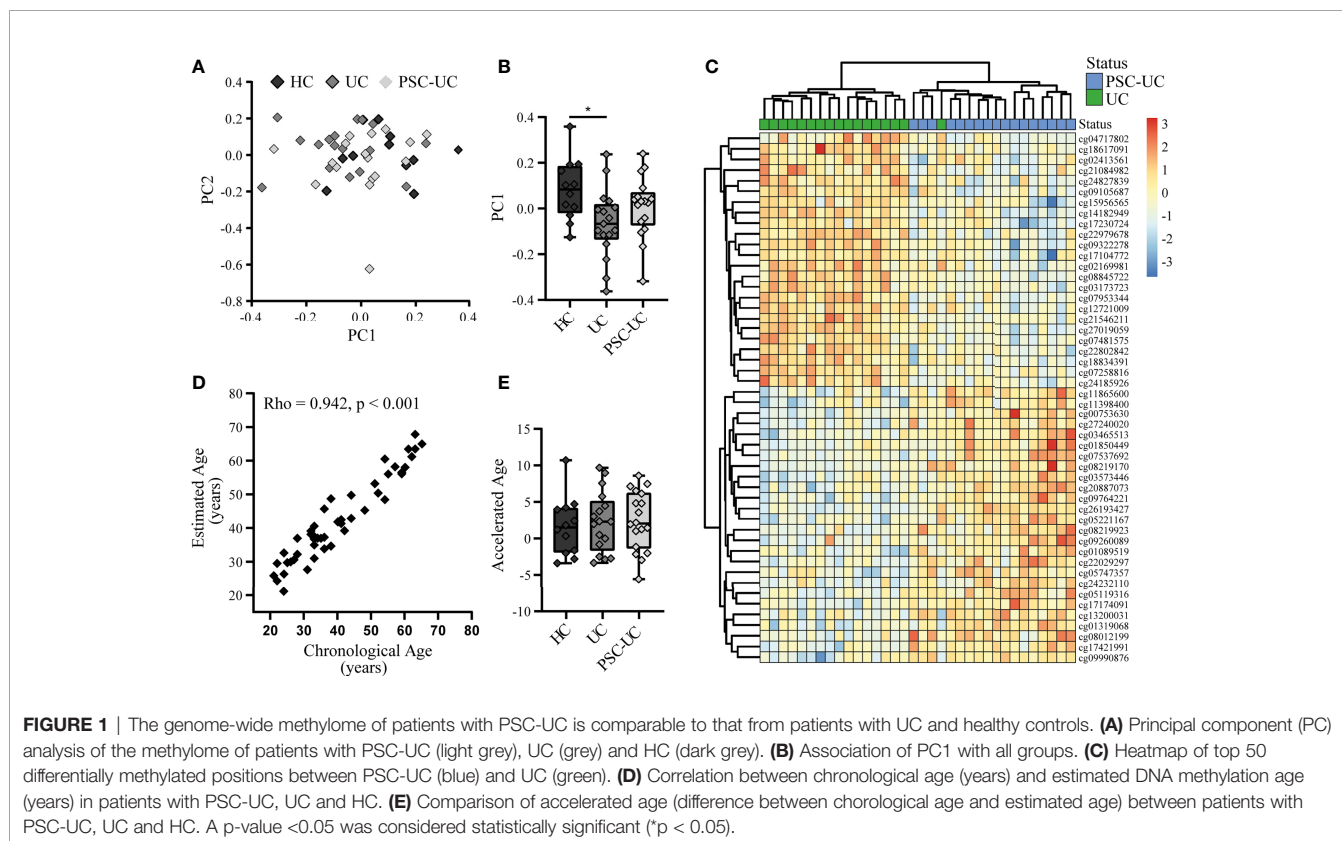
The Genome-Wide Methylome of Patients With PSC-UC Is Comparable to That From Patients With UC and Healthy Controls

DNA methylation was investigated in peripheral blood samples from 17 patients with PSC-UC, 17 patients with UC and 12 HCs.

The characteristics of the included patients are shown in **Table 1** (cohort 1). To limit the number of variables, we included only male patients, with all groups being matched for age (median age 41, 34 and 39 years for PSC-UC, UC and HC, respectively), UC duration (16 and 11 years for PSC-UC and UC, respectively) and medication use (all patients with PSC-UC and UC used mesalazine, 50% used thiopurins and none used biologicals). Median PSC duration at time of inclusion was 5 years. None of the patients had undergone liver transplantation at the time of sampling.

Global methylation analysis through principal component (PC) analysis did not present a clear separation of the samples according to disease status (**Figure 1A**). Notably, PC1 presented a larger separation between HC and UC (p-value = 0.021), than HC and PSC-UC (p-value = 0.347) or PSC-UC and UC (p-value = 0.292) (**Figure 1B**).

Next, we specifically investigated whether any of the probes were differentially methylated when comparing PSC-UC with UC. While we were able to observe differentially methylated positions (DMPs) that were visibly differentially methylated (**Figure 1C**), they were not statistically significant after correction for multiple testing (**Table 2**). The most differentially methylated probe was cg02169981 (annotated to *WNT11*), which displayed hypermethylation among patients with UC compared to patients with PSC-UC (**Supplementary Figure 1A**). Expanding our search to regions of contiguous differential methylation (DMRs) when comparing either PSC-UC with UC, or PSC-UC with HC yielded no statistically significant differences (**Supplementary Table 1**).



Having identified no clear differences in methylation at a genome-wide level, we adopted a hypothesis-driven approach where we interrogated the methylation status of previously reported loci/genes of interest, obtained from previous genome- (GWAS) or transcriptome- wide association studies (TWAS) (7–9, 44–49). When comparing PSC-UC with UC, only two genes were significantly enriched for nominally differentially methylated probes: *BACH2* and *ASAP2* (**Supplementary Table 2** and **Supplementary Figure 2**). Comparing PSC-UC with HC indicated that probes associated with eight genes showed a statistically significant difference (**Supplementary Table 2**). *UBASH3A*, the most significant SNP in a large GWAS study and associated with a lower risk of PSC (8), showed differential methylation centering around the promoter region when comparing PSC-UC with HC (**Supplementary Figure 2C**).

In a recent study, it was reported that patients with PSC presented an increased DNA methylation age relative to their chronological age, which was especially apparent in samples obtained from patients at an advanced disease state (16). Using the same method we estimated the DNA methylation age of all samples and compared the estimated and chronological age (17). Overall, we found that the DNA methylation age correlated well with the chronological age (**Figure 1D**; Spearman's Rho 0.942, p -value < 0.001). Calculating the difference between the DNA methylation age and the chronological age revealed a median difference of 2.1 years in patients with PSC-UC (range -5.6-8.6), which did not significantly differ from either patients with UC or HCs (median 2.3 (range -3.3-9.7) and 1.5 (range -3.4-10.7), respectively (p -value = 0.874, **Figure 1E**). We note however that the use of different normalization methods resulted in larger or smaller differences between the predicted age and the chronological age, with a more notable deviation when using quantile normalization. However, even with the larger differences from the quantile normalization, we did not identify any significant differences between phenotypes (**Supplementary Figure 3**).

Taken together, we observed no big differences between the three groups at a DNA methylome-wide level, nor did we find any statistically significant differences between DMPs or DMRs of patients with PSC-UC and either patients with UC or healthy controls.

Classification Analysis Distinguishes PSC-UC From UC

Next, we investigated whether DNA methylation could be of use for distinguishing patients with PSC-UC from patients with UC without PSC. To this end, we performed a classification analysis using repeated cross-validation with gradient boosting, where we sought to capitalize on potential non-linear relationships between the CpG loci of interest and the presence of PSC-UC (34). Altogether, the classification analysis yielded a predictive model with an area under the receiver-operator characteristic curve (AUROC) of 0.80. Subsequent permutation analyses indicated that 18 CpGs contributed significantly to the model (**Table 3** and **Figure 2A**). Of these 18 CpGs, PSC-UC-associated hypermethylation was observed for *NINJ2*, cg12219587, *SERPINB9*, *DNAJC17*, cg19079513, *OR51A7*, cg12313868 and *SOX6* whereas hypomethylation was found for cg00980980, *TTC15*, *THUMPD1*, *TRAPPC12* and *CYP4F22* (**Figure 2B**).

While most predictive CpGs were embedded in a gene as single CpG, four were found to be all localized within the gene *NINJ2*. All four of the *NINJ2*-associated predictor CpGs were adjacent to one another and were located in an intronic region (**Figure 3A** and **Supplementary Figure 5**). Technical validation through bisulfite sequencing of two of the four predictor CpGs (cg26654770 and cg14911689) confirmed the methylation pattern (**Figure 3B**). Notably, the observed methylation signal of the *NINJ2*-associated DMPs presented itself as a clustered pattern around 0%, 50% and 100% methylation (**Figure 3C**), a pattern characteristic of underlying genetic variants (50). However, sequencing the underlying region of interest revealed no single nucleotide polymorphism (SNP) or other local genetic variants at the loci of interest (**Figure 3D**). As more distal genetic variants might affect DNA methylation, which have been termed methylation quantitative trait loci (mQTL), we interrogated the mQTLs database (<http://www.mqtl.org>) (51) to identify potential catalogued genetic variants that have been associated with the DNA methylation of our sites of interest. Altogether, we identified 1330, 1374, 1329 and 1329 mQTLs where genetic variation was strongly associated with the loci cg01201512, cg26371957, cg14911689 and cg26654770, respectively (**Additional File_mQTLs**).

TABLE 2 | Differentially methylated positions (DMPs) when comparing PSC-UC with UC.

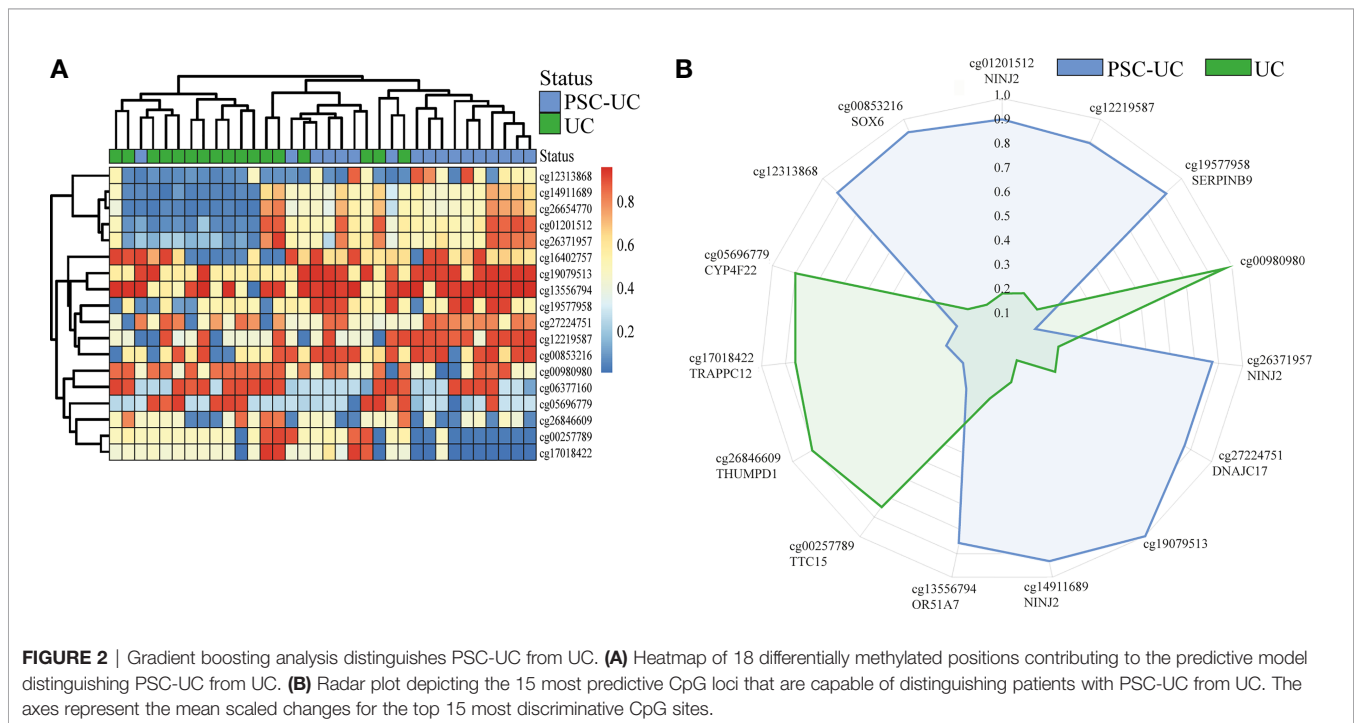
Location	Associated gene	Chromosome	Difference in methylation in PSC-UC vs UC	p-value	Adj. p-value	Name
cg02169981	<i>WNT11</i>	Chr 11	Hypomethylation	2.20E-06	0.770179	Wnt Family Member 11
cg26193427	<i>INPP5A</i>	Chr 10	Hypermethylation	5.44E-06	0.770179	Inositol Polyphosphate-5-Phosphatase A
cg07953344	<i>NRP2</i>	Chr 2	Hypomethylation	5.59E-06	0.770179	Neuropilin 2
cg17421991	<i>MXN1</i>	Chr 7	Hypermethylation	5.80E-06	0.770179	Motor Neuron And Pancreas Homeobox 1
cg08012199	<i>ARHGAP20*</i>	Chr 11	Hypermethylation	6.23E-06	0.770179	Rho GTPase Activating Protein 20
cg27019059	<i>RYK*</i>	Chr 3	Hypomethylation	8.73E-06	0.770179	Receptor Like Tyrosine Kinase
cg13200031	<i>FAM163A</i>	Chr 1	Hypermethylation	9.66E-06	0.770179	Family With Sequence Similarity 163 Member A
cg09990876		Chr 1	Hypermethylation	1.04E-05	0.770179	
cg00753630	<i>ATP6VOE2-AS1</i>	Chr 7	Hypermethylation	1.18E-05	0.770179	Vacuolar Proton Pump Subunit E 2
cg18617091	<i>TIMPRSS6</i>	Chr 22	Hypomethylation	1.36E-05	0.770179	Transmembrane Serine Protease 6

The 10 most differentially methylated positions. *Neighbour gene. P-values were calculated through linear regression using *limma* and adjusted for multiple testing using the Benjamini-Hochberg method.

TABLE 3 | Predictor CpG loci capable of distinguishing PSC-UC from UC.

Location	Associated gene	Chromosome	Difference in methylation in PSC-UC vs UC	Feature Importance	Name
cg01201512	<i>NINJ2</i>	Chr 12	Hypermethylation	1.442312	Nerve Injury-Induced Protein 2, Ninjurin-2
cg12219587			Hypermethylation	1.295918	
cg19577958	<i>SERPINB9</i>	Chr 6	Hypermethylation	1.265306	Serpin Family B Member 9, Serpin B9, Granzyme B inhibitor
cg00980980			Hypomethylation	0.858523	
cg26371957	<i>NINJ2</i>	Chr 12	Hypermethylation	0.639473	Nerve Injury-Induced Protein 2, Ninjurin-2
cg27224751	<i>DNAJC17</i>	Chr 15	Hypermethylation	0.632653	DnaJ Heat Shock Protein Family (Hsp40) Member C17
cg19079513			Hypermethylation	0.612245	
cg14911689	<i>NINJ2</i>	Chr 12	Hypermethylation	0.583521	Nerve Injury-Induced Protein 2, Ninjurin-2
cg13556794	<i>OR51A7</i>	Chr 11	Hypermethylation	0.520408	Olfactory Receptor Family 51 Subfamily A Member 7, Odorant receptor
cg00257789	<i>TRAPPC12/TTC15</i>	Chr 2	Hypomethylation	0.448323	Trafficking Protein Particle Complex 12
cg26846609	<i>THUMP1</i>	Chr 16	Hypomethylation	0.397421	THUMP domain-containing protein 1
cg17018422	<i>TRAPPC12/TTC15</i>	Chr 2	Hypomethylation	0.355172	Trafficking Protein Particle Complex 12
cg05696779	<i>CYP4F22</i>	Chr 19	Hypomethylation	0.306122	Cytochrome P450 family 4 subfamily F member 22
cg12313868			Hypermethylation	0.244898	
cg00853216	<i>SOX6</i>	Chr 11	Hypermethylation	0.204082	Transcription factor SOX-6
cg26654770	<i>NINJ2</i>	Chr 12	Hypermethylation	0.196723	Nerve Injury-Induced Protein 2, Ninjurin-2
cg06377160	<i>SNX19</i>	Chr 11	Hypomethylated	0.173469	Sorting nexin-19
cg16402757	<i>CUL2</i>	Chr 10	Hypermethylated	0.134404	Cullin-2

The 18 CpG loci returned by gradient boosting that were capable of distinguishing PSC-UC from UC.



Next, we sought to investigate whether the relation between CpG methylation that we found associated with PSC-UC was reflected in differential gene expression of this particular gene, through measuring *NINJ2* transcripts using quantitative PCR. However, we found no differential expression of *NINJ2* between PSC-UC and UC patients (**Figure 3E**), suggesting that the observed difference in methylation does not affect the expression of *NINJ2*.

Additional classification analysis on PSC-UC and HC samples yielded a predictive model with an area under the

receiver-operator characteristic curve (AUROC) of 0.83. The CpG positions observed in this classification did not overlap with the CpGs as compared to the predictive model on PSC-UC and UC (**Supplementary Table 3**).

Peripheral Blood Cell Distribution Is Comparable Between Patients With PSC-UC and UC

Due to the epigenetic and hence cell-specific nature of DNA methylation, observed differences could be the result of

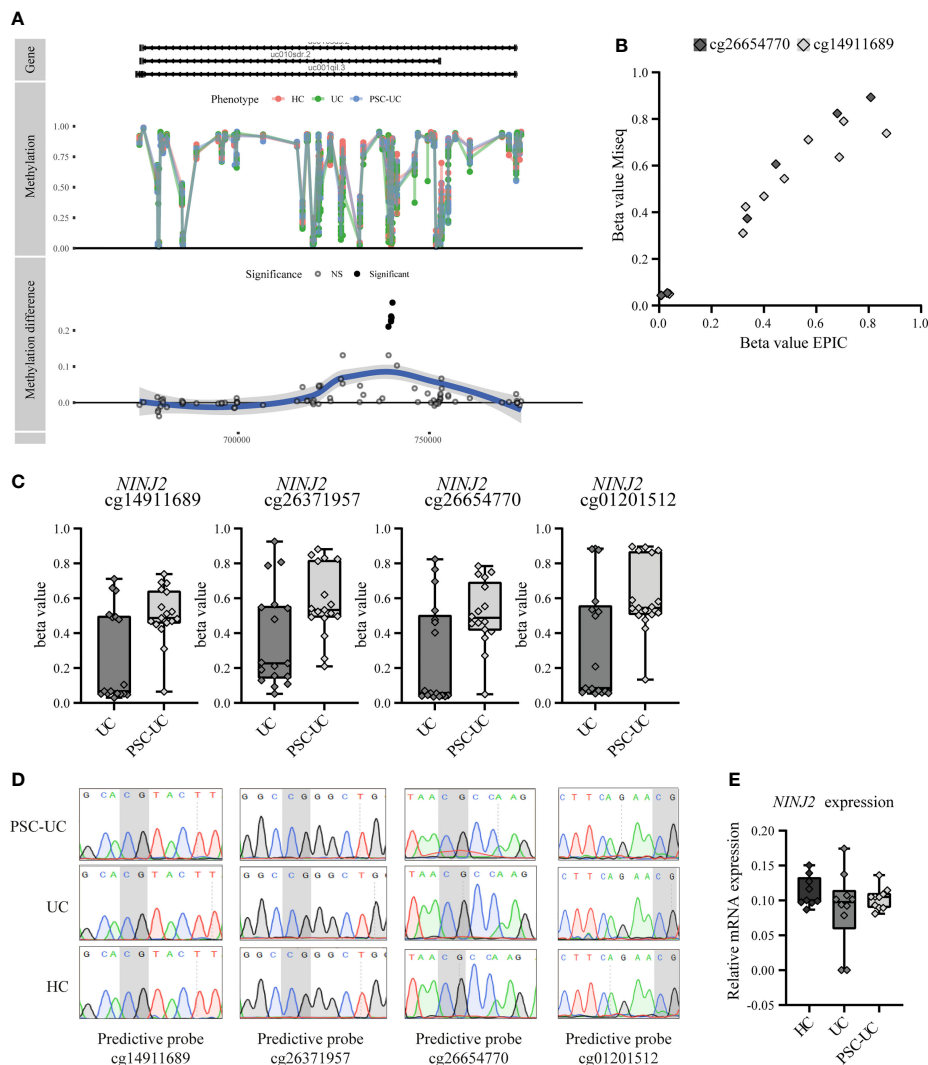


FIGURE 3 | Predictive CpGs annotated to *NINJ2* were all hypermethylated in PSC-UC compared to UC patients. **(A)** Visualization of *NINJ2* by plotting the actual percentage methylation for PSC-UC, UC and HC (“Methylation”) as well as the difference between PSC-UC and UC in percentage methylation (“Methylation difference”) relative to the position on the genome. **(B)** Visual correlation of the percentage methylation observed for cg26654770 and cg14911689 as calculated using the 850k DNA methylation array and through MiSeq sequencing for 5 PSC-UC, 2 UC and 2 HC patients. **(C)** Differences in percentage methylation (beta value) in patients with UC ($n = 17$) and PSC-UC ($n = 17$). **(D)** Representative images of Sanger sequencing traces surrounding the CpG loci of interest (marked in grey) that are annotated to *NINJ2* in PSC-UC ($n = 14$), UC ($n = 13$) and HCs ($n = 12$). **(E)** Relative mRNA expression of *NINJ2* normalized to the household genes *GAPDH* and *HPRT* in peripheral blood mononuclear cells of PSC-UC ($n = 10$), UC ($n = 10$) and HCs ($n = 9$).

differences in the cellular composition of peripheral blood between PSC-UC and UC patients (52, 53). We therefore estimated the cellular composition using the algorithm described in Houseman et al. (Figure 4A) (54). We found no evidence that any of the estimated lineages ($CD4^+$ T-cells, $CD8^+$ T-cells, B-cells, natural killer (NK)-cells, monocytes and neutrophils) were different when comparing PSC-UC with UC. By contrast, we observed a significantly lower abundance of NK cells in the PSC-UC and UC samples as compared to HC (p -value = $5.33E-04$).

As cellular estimates based on DNA methylation are currently limited in their resolution, we explored the peripheral blood

mononuclear cellular distribution in more detail using mass cytometry. Blood samples were collected from a new cohort of PSC-UC ($n = 10$), UC ($n = 10$) and HC ($n = 9$). The median age at inclusion was 40, 58 and 30 years ($p=0.022$) (Table 1, cohort 2), for PSC-UC, UC and HC respectively. UC duration was significantly longer in patients with UC without PSC (median 22 years (IQR 14-34) compared to 9 years (IQR 5-19) in patients with PSC-UC ($p=0.019$). The majority of PSC-UC and UC patients used Mesalazine (90% in both groups) with only one patient with UC (without PSC) using a thiopurine. Different lineages were identified based on cellular phenotype and visualized in tSNE plots (Figure 5, Figure 4B). Akin to our

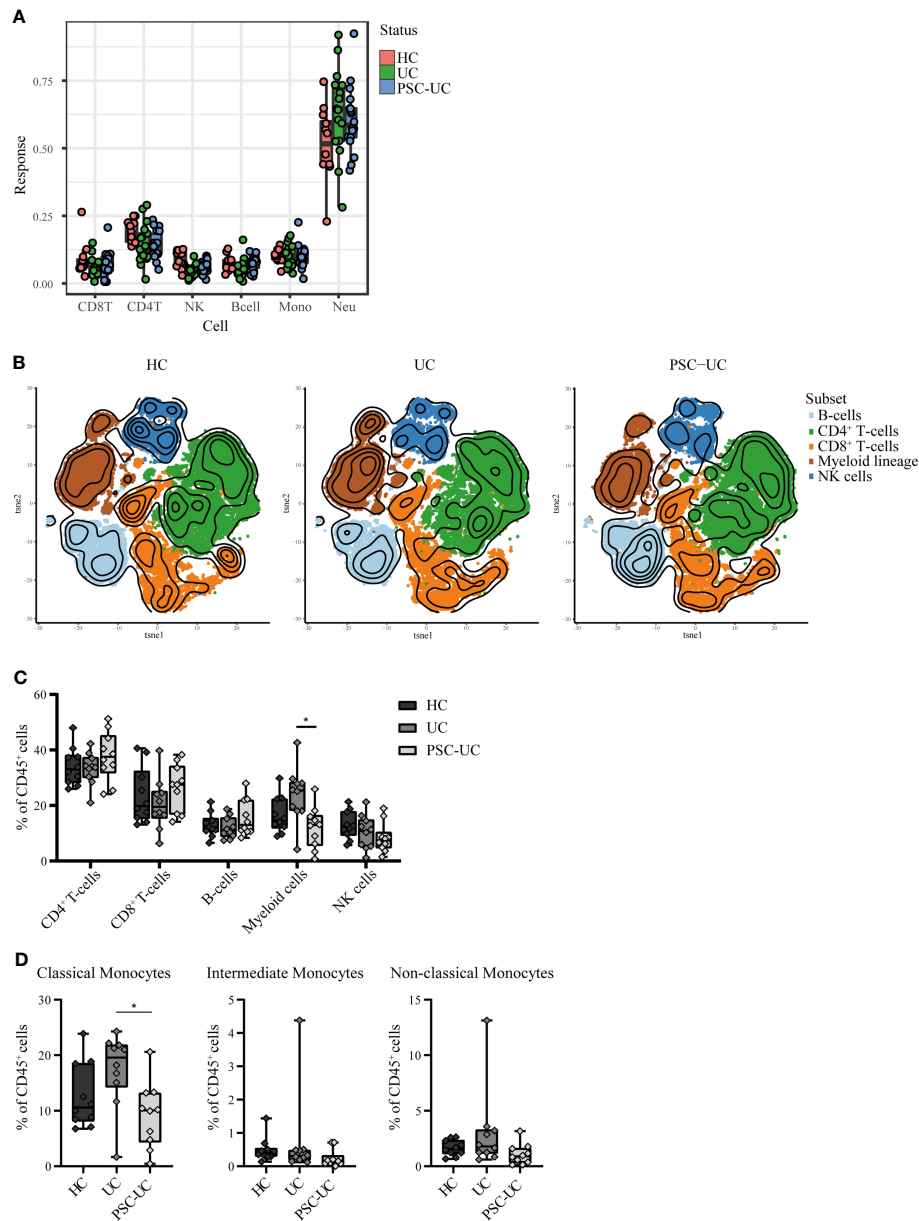
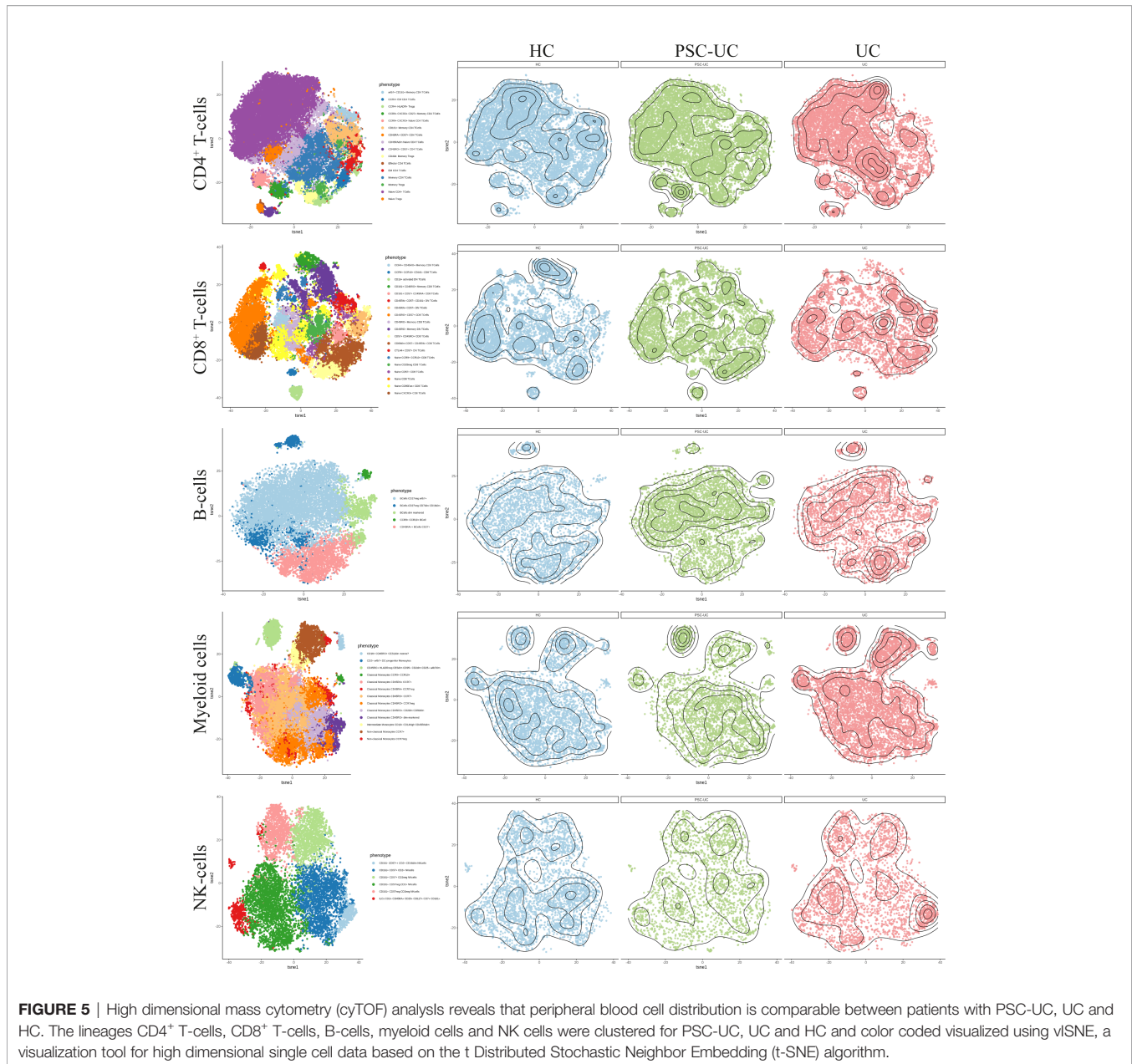


FIGURE 4 | High-dimensional mass cytometry (cyTOF) analysis reveals that peripheral blood cell distribution is comparable between patients with PSC-UC, UC and HC. **(A)** Estimated cell proportions as derived from Houseman algorithm cell mixture deconvolution from DNA methylation data of PSC-UC (blue), UC (green) and HC (red). **(B)** The lineages CD4⁺T-cells, CD8⁺T-cells, B-cells, myeloid cells and NK-cells were clustered for PSC-UC, UC and HC and color-coded visualized using tSNE. **(C)** Differences in frequencies of CD4⁺T-cells, CD8⁺T-cells, B-cells, Myeloid lineage and NK-cells as percentage of total CD45⁺ cells in HC (n = 9), UC (n = 10) and PSC-UC (n = 10). **(D)** Differences in frequencies of classical monocytes, intermediate monocytes and non-classical monocytes percentage of total CD45⁺ cells in HC (n = 9), UC (n = 10) and PSC-UC (n = 10). Statistical testing was performed using Kruskal Wallis with Dunn's correction for multiple testing. A p-value < 0.05 was considered statistically significant (*p < 0.05).

observations made at the level of DNA methylation, the abundance of CD4⁺ T-cells, CD8⁺ T-cells and B-cells was not statistically different between PSC-UC, UC and HCs (**Figure 4C** and **Supplementary Figure 4**). Notably, we observed a similar trend in NK cell distribution as seen in the Houseman algorithm estimation. The myeloid lineage was significantly more abundant in patients with UC compared to patients with PSC-UC (p-value = 0.013,

Figure 4C). Disentangling the myeloid lineage into the constituent monocyte subtypes revealed that this difference was predominantly observable in the classical monocyte (CD14⁺CD16⁺) population (**Figure 4D**). Within the classical monocyte populations, we particularly saw a decrease in CD45RA⁺CCR7⁺ classical monocytes and CD45RO⁺ CD2^{DIM} CD69^{DIM} classical monocytes among PSC patients compared to UC patients. Although the overall



T-cell population did not differ between PSC-UC and UC, we observed distinct proportions of memory T-cells expressing CD161, a c-type lectin like-receptor expressed by NK cells, T helper 17 cells and mucosal invariant T (MAIT) cells (55) between patients with PSC-UC and UC (**Supplementary Figure 4**).

DISCUSSION

Distinguishing patients with UC and concomitant PSC from patients with UC could give more insights in disease pathophysiology of PSC and would be of clinical value for early diagnosis. In this explorative study, we performed genome-wide DNA methylation and mass cytometry analysis

on whole peripheral blood from patients with PSC-UC, UC and HCs. We show that minor differences exist in the peripheral blood methylome when comparing male patients with PSC-UC to male patients with solely UC. Notwithstanding, classification analysis yielded a predictive model capable of distinguishing patients with PSC-UC and UC.

The overall lack of large-scale differences in methylation between the various groups corroborates the observations made by Moore et al., where the authors did not observe large global methylation changes in the peripheral blood of patients with PSC compared to healthy controls and between patients with PSC with and without IBD (15). While peripheral blood is a practical tissue for biomarker use due to its ease of access, it contains a heterogeneous population of cells, which might be less

representative for disease features, such as the PSC-UC-associated phenotypes that manifest primarily in the liver and gut tissue. Subtle differences between PSC-UC, UC and HC may therefore remain hidden in the analysis performed.

Indeed mass cytometry analyses revealed subtle differences in blood cell composition across patient groups. While we demonstrated that the peripheral blood cell distribution was largely comparable between patients with PSC-UC and UC, the myeloid lineage, the monocyte CD14/CD16 subsets in particular, were more abundant in patients with UC compared to PSC-UC and HCs (56). As DNA methylation signal can vary between different cell types, observed differences might be reflective of changes in the underlying population (53).

Importantly, we could not confirm an increased DNA methylation age in PSC patients as was reported by Trauner et al. (16). The discrepancy between our observations might be related to their phenotype of interest, namely progression of fibrosis, for which we have no information of at the time of sampling. The fact that the median PSC duration was low in our cohort (5 years), might have influenced our accelerated age differences. Notwithstanding, we believe that the observed differences were likely influenced by the choice of normalization as Trauner et al. utilized quantile normalization, a method that presented one of the largest offset as compared to other normalization methods (57).

Classification analysis yielded a predictive model, which enables a distinction between peripheral blood of patients with PSC-UC and UC with an AUROC of 0.80, indicating that these two disease entities do have a distinct epigenetic architecture. The discrepancy between our DMP analysis and the classification analysis may be due to the non-linear relationships identified through gradient boosting. Whereas none of the predictive CpG-associated genes had been described within the context of PSC previously, two genes are associated to hepatocellular carcinoma, namely *CUL2* and *SOX6* (58–60). Moreover, both *CUL2* and *SOX6* are associated with colitis as well, while *SERPINB9* and *NINJ2* have been associated with colorectal cancer (61–64).

The four predictive CpG associated to *NINJ2* were found to be hypermethylated in PSC-UC compared to UC patients, which we validated through bisulfite sequencing. While the four CpGs presented a methylation pattern reminiscent of an underlying genetic variant (65), we did not identify any genetic variants through Sanger sequencing at the sites of interest. This does not eliminate the possibility that other genetic variants might have conferred the observed methylation signal, such as copy number variations (CNV), indels and more distal SNPs associated with differences in DNA methylation (51, 66, 67), which we were not able to pick up with the techniques utilized in the current study. With over a thousand catalogued mQTLs associated to the four *NINJ2*-associated predictive CpGs, the possibility is present that distal genetic variants may influence the methylation status of *NINJ2* (51).

The biological consequence of the observed difference in methylation of *NINJ2* remains to be established. The encoded Ninjurin 2 or nerve injury induced protein 2 (*NINJ2*) is an

adhesion molecule expressed in neurons and glial cells and is involved in nerve regeneration (68). Notably, Ninjurin 1, a homologue of Ninjurin 2, as well as *NINJ2* are highly expressed in myeloid cells and peripheral leukocytes, suggesting a role in immune-mediated diseases as well (44, 69, 70). In vascular endothelial cells it was found that *NINJ2* regulates monocyte-adhesion as well as endothelial inflammation through the expression of pro-inflammatory cytokines such as IL-1 β , TNF- α , IL-8, IL-6, ICAM-1 and E-selectin (70). Through mass cytometry we observed a diminished abundance of the myeloid population among the PSC-UC patients relative to the UC patients. The observed differences in myeloid cell population abundance may therefore be related to the observed difference in *NINJ2* methylation. However, while *NINJ2* expression was reportedly associated with DNA methylation in CD4⁺ T-cells (71), we observed no difference in *NINJ2* gene expression in the current cohort. This might be attributable to the differences in methylation being observed in the first intron rather than the promoter region. In summary, the role of *NINJ2* in relation to PSC-UC remains unclear.

Our study is not without its limitations: we only included male subjects to reduce variables and limit confounding in DNA methylation analysis. While having a same-sex cohort decreases variance, it makes the observations less translatable to the general PSC-IBD population. While previous EWAS have indicated that DNA methylation profiles differ between males and females (72), limited to no sex-associated difference were reported in PSC-associated EWAS (15, 16). One study did show an enrichment of differentially methylated CpG sites located on chromosome X in patients with PSC-IBD and PSC without IBD compared to controls but did not make a comparison with the methylomes of patients with PSC alone (15). While the majority of the patients with PSC are male, we acknowledge that further validation experiments in a larger and more heterogeneous population with both male and female subjects are necessary to properly ascertain the robustness of the predictive CpGs and the potential utility in the clinic in defining PSC-UC.

CONCLUSIONS

Our study provides a novel approach of exploring DNA methylation analysis to differentiate patients with PSC-UC from patients with UC. Further validation in a different cohort has to confirm the biomarker potential of these methylation differences for early detection of PSC.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and

accession number(s) can be found below: <https://ega-archive.org/>, EGAS00001005832.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethics Committee at the Amsterdam UMC, University of Amsterdam, Ref. nrs: METC 06–267/E and METC 2018-050. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Concept and design of the study: MK, CP, and WJ. MK assembled the patient cohorts, prepared samples for the DNA methylation array and performed the mass cytometry. PvH performed RNA isolation. JV, JGV, and MW analyzed the mass cytometry data. IH and TH performed Sanger sequencing and next generation sequencing. ALY and EL analyzed and visualized the results acquired from the HumanMethylation EPIC BeadChip array. EL performed machine learning. MK, IH, and ALY drafted the manuscript. PH, TH, CP, and WJ supervised the study. All authors read and approved the final manuscript.

REFERENCES

- Williamson KD, Chapman RW. New Therapeutic Strategies for Primary Sclerosing Cholangitis. *Semin Liver Dis* (2016) 36(1):5–14. doi: 10.1055/s-0035-1571274
- Boonstra K, Weersma RK, van Erpecum KJ, Rauws EA, Spanier BW, Poen AC, et al. Population-Based Epidemiology, Malignancy Risk, and Outcome of Primary Sclerosing Cholangitis. *Hepatology* (2013) 58(6):2045–55. doi: 10.1002/hep.26565
- Sorensen JO, Nielsen OH, Andersson M, Ainsworth MA, Ytting H, Belard E, et al. Inflammatory Bowel Disease With Primary Sclerosing Cholangitis: A Danish Population-Based Cohort Study 1977–2011. *Liver Int* (2018) 38(3):532–41. doi: 10.1111/liv.13548
- Weismuller TJ, Trivedi PJ, Bergquist A, Imam M, Lenzen H, Ponsioen CY, et al. Patient Age, Sex, and Inflammatory Bowel Disease Phenotype Associate With Course of Primary Sclerosing Cholangitis. *Gastroenterology* (2017) 152(8):1975–84 e8. doi: 10.1053/j.gastro.2017.02.038
- Adams DH, Eksteen B. Aberrant Homing of Mucosal T Cells and Extra-Intestinal Manifestations of Inflammatory Bowel Disease. *Nat Rev Immunol* (2006) 6(3):244–51. doi: 10.1038/nri1784
- de Krijger M, Wildenberg ME, de Jonge WJ, Ponsioen CY. Return to Sender: Lymphocyte Trafficking Mechanisms as Contributors to Primary Sclerosing Cholangitis. *J Hepatol* (2019) 71(3):603–15. doi: 10.1016/j.jhep.2019.05.006
- Liu JZ, Hov JR, Folseraas T, Ellinghaus E, Rushbrook SM, Doncheva NT, et al. Dense Genotyping of Immune-Related Disease Regions Identifies Nine New Risk Loci for Primary Sclerosing Cholangitis. *Nat Genet* (2013) 45(6):670–5. doi: 10.1038/ng.2616
- Ji SG, Juran BD, Mucha S, Folseraas T, Jostins L, Melum E, et al. Genome-Wide Association Study of Primary Sclerosing Cholangitis Identifies New Risk Loci and Quantifies the Genetic Relationship With Inflammatory Bowel Disease. *Nat Genet* (2017) 49(2):269–73. doi: 10.1038/ng.3745
- Ellinghaus D, Jostins L, Spain SL, Cortes A, Bethune J, Han B, et al. Analysis of Five Chronic Inflammatory Diseases Identifies 27 New Associations and Highlights Disease-Specific Patterns at Shared Loci. *Nat Genet* (2016) 48(5):510–8. doi: 10.1038/ng.3528
- Cheung AC, LaRusso NF, Gores GJ, Lazaridis KN. Epigenetics in the Primary Biliary Cholangitis and Primary Sclerosing Cholangitis. *Semin Liver Dis* (2017) 37(2):159–74. doi: 10.1055/s-0037-1603324

FUNDING

ALY was funded by the European Union's Horizon 2020 research and innovation program under Grant Agreement No. ITN-2014-EID-641665. ALY and IH were funded by the Helmsley charitable trust.

ACKNOWLEDGMENTS

M.S. Tulek, K. Boonstra: collection peripheral blood samples of PSC-UC patients, UC patients and healthy controls. We thank the Microscopy and Cytometry Core facility at the VU Medical Center Amsterdam for use of the Helios and support in sample preparation and acquisition. We thank the company Genome Diagnostics B.V (GENDX) for use of the Illumina Miseq machine to perform next generation sequencing

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.840935/full#supplementary-material>

- Razin A, Cedar H. DNA Methylation and Gene Expression. *Microbiol Rev* (1991) 55(3):451–8. doi: 10.1128/mr.55.3.451-458.1991
- Deaton AM, Bird A. CpG Islands and the Regulation of Transcription. *Genes Dev* (2011) 25(10):1010–22. doi: 10.1101/gad.2037511
- McDermott E, Ryan EJ, Tosetto M, Gibson D, Burrage J, Keegan D, et al. DNA Methylation Profiling in Inflammatory Bowel Disease Provides New Insights Into Disease Pathogenesis. *J Crohns Colitis* (2016) 10(1):77–86. doi: 10.1093/ecco-jcc/jjv176
- Li Yim AYE, Duijvis NW, Zhao J, de Jonge WJ, D'Haens G, Mannens M, et al. Peripheral Blood Methylation Profiling of Female Crohn's Disease Patients. *Clin Epigenet* (2016) 8:65. doi: 10.1186/s13148-016-0230-5
- Moore RM, Sun Z, Juran BD, Lazaridis KN. Genome-Wide Resolution Peripheral Blood Methylation Profiling Reveals Signatures for Cholestatic Liver Disease. *Epigenomics* (2020) 12(16):1363–75. doi: 10.2217/epi-2020-0048
- Trauner M, Gindin Y, Jiang Z, Chung C, Subramanian GM, Myers RP, et al. Methylation Signatures in Peripheral Blood Are Associated With Marked Age Acceleration and Disease Progression in Patients With Primary Sclerosing Cholangitis. *JHEP Rep* (2020) 2(1):100060. doi: 10.1016/j.jhep.2019.11.004
- Horvath S. DNA Methylation Age of Human Tissues and Cell Types. *Genome Biol* (2013) 14(10):R115. doi: 10.1186/gb-2013-14-10-r115
- Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA Methylation-Based Measures of Biological Age: Meta-Analysis Predicting Time to Death. *Aging (Albany NY)* (2016) 8(9):1844–65. doi: 10.18632/aging.101020
- Bibikova M, Le J, Barnes B, Saedinia-Melnyk S, Zhou L, Shen R, et al. Genome-Wide DNA Methylation Profiling Using Infinium(R) Assay. *Epigenomics* (2009) 1(1):177–200. doi: 10.2217/epi.09.14
- Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: A Flexible and Comprehensive Bioconductor Package for the Analysis of Infinium DNA Methylation Microarrays. *Bioinformatics* (2014) 30(10):1363–9. doi: 10.1093/bioinformatics/btu049
- van Iterson M, Tobi EW, Sliker RC, den Hollander W, Luijk R, Slagboom PE, et al. MethyLaid: Visual and Interactive Quality Control of Large Illumina 450k Datasets. *Bioinformatics* (2014) 30(23):3435–7. doi: 10.1093/bioinformatics/btu566

22. Fortin JP, Fertig E, Hansen K. Shinymethyl: Interactive Quality Control of Illumina 450k DNA Methylation Arrays in R. *F1000Res* (2014) 3:175. doi: 10.12688/f1000research.4680.2
23. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007
24. Triche TJ Jr., Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-Level Processing of Illumina Infinium DNA Methylation Beadarrays. *Nucleic Acids Res* (2013) 41(7):e90. doi: 10.1093/nar/gkt090
25. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of Polymorphic and Off-Target Probe Binding Sites on the Illumina Infinium Methylation Beadchip. *Genom Data* (2016) 9:22–4. doi: 10.1016/j.gdata.2016.05.012
26. Zhou W, Laird PW, Shen H. Comprehensive Characterization, Annotation and Innovative Use of Infinium DNA Methylation Beadchip Probes. *Nucleic Acids Res* (2017) 45(4):e22. doi: 10.1093/nar/gkw967
27. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York: ISBN 978-3-319-24277-4 (2016). Available at: <https://ggplot2.tidyverse.org>.
28. Yin T, Cook D, Lawrence M. Ggbio: An R Package for Extending the Grammar of Graphics for Genomic Data. *Genome Biol* (2012) 13(8):R77. doi: 10.1186/gb-2012-13-8-r77
29. Pidsley R, YW CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A Data-Driven Approach to Preprocessing Illumina 450K Methylation Array Data. *BMC Genomics* (2013) 14:293. doi: 10.1186/1471-2164-14-293
30. Salas LA, Koestler DC, Butler RA, Hansen HM, Wiencke JK, Kelsey KT, et al. An Optimized Library for Reference-Based Deconvolution of Whole-Blood Biospecimens Assayed Using the Illumina Humanmethylation Beadarray. *Genome Biol* (2018) 19(1):64. doi: 10.1186/s13059-018-1448-7
31. Pereira J, Stroes ESG, Zwinderman AH, Levin E. *Covered Information Disentanglement: Model Transparency via Unbiased Permutation Importance*. Available at: <https://arxiv.org/pdf/2111.09744.pdf>2021.
32. Reeskamp LF, Venema A, Pereira JPB, Levin E, Nieuwdorp M, Groen AK, et al. Differential DNA Methylation in Familial Hypercholesterolemia. *EBioMedicine* (2020) 61:103079. doi: 10.1016/j.ebiom.2020.103079
33. de Groot P, Nikolic T, Pellegrini S, Sordi V, Imangaliyev S, Rampanelli E, et al. Faecal Microbiota Transplantation Halts Progression of Human New-Onset Type 1 Diabetes in a Randomised Controlled Trial. *Gut* (2021) 70(1):92–105. doi: 10.1136/gutjnl-2020-322630
34. van der Vossen EWJ, Bastos D, Stols-Goncalves D, de Goffau MC, Davids M, Pereira JPB, et al. Effects of Fecal Microbiota Transplant on DNA Methylation in Subjects With Metabolic Syndrome. *Gut Microbes* (2021) 13(1):1993513. doi: 10.1080/19490976.2021.1993513
35. Li LC, Dahiya R. Methprimer: Designing Primers for Methylation Pcrs. *Bioinformatics* (2002) 18(11):1427–31. doi: 10.1093/bioinformatics/18.11.1427
36. Holmes EE, Jung M, Meller S, Leisse A, Sailer V, Zech J, et al. Performance evaluation of kits for bisulfite-conversion of DNA from tissues, cell lines, FFPE tissues, Aspirates, Lavages, Effusions, Plasma, Serum, and Urine. *PLoS One* (2014) 9(4):e93933. doi: 10.1371/journal.pone.0093933
37. Krueger F, Andrews SR. Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications. *Bioinformatics* (2011) 27(11):1571–2. doi: 10.1093/bioinformatics/btr167
38. Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, et al. Amplification Efficiency: Linking Baseline and Bias in the Analysis of Quantitative PCR Data. *Nucleic Acids Res* (2009) 37(6):e45. doi: 10.1093/nar/gkp045
39. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paep A, et al. Accurate Normalization of Real-Time Quantitative RT-PCR Data by Geometric Averaging of Multiple Internal Control Genes. *Genome Biol* (2002) 3(7):RESEARCH0034. doi: 10.1186/gb-2002-3-7-research0034
40. Kotzcha N, Krutzik PO, Irish JM. Web-Based Analysis and Publication of Flow Cytometry Experiments. *Curr Protoc Cytom* (2010) Chapter 10:Unit10.7. doi: 10.1002/0471142956.cy1017s53
41. Bagwell CB, Inokuma M, Hunsberger B, Herbert D, Bray C, Hill B, et al. Automated Data Cleanup for Mass Cytometry. *Cytometry A* (2020) 97(2):184–98. doi: 10.1002/cyto.a.23926
42. Van Gassen S, Callebaut B, Van Helden MJ, Lambrecht BN, Demeester P, Dhaene T, et al. Flowsom: Using Self-Organizing Maps for Visualization and Interpretation of Cytometry Data. *Cytometry A* (2015) 87(7):636–45. doi: 10.1002/cyto.a.22625
43. Amir el AD, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, et al. Visne Enables Visualization of High Dimensional Single-Cell Data and Reveals Phenotypic Heterogeneity of Leukemia. *Nat Biotechnol* (2013) 31(6):545–52. doi: 10.1038/nbt.2594
44. Aoki CA, Dawson K, Kenny TP, Gershwin ME, Bowls CL. Gene Expression by PBMC in Primary Sclerosing Cholangitis: Evidence for Dysregulation of Immune Mediated Genes. *Clin Dev Immunol* (2006) 13(2-4):265–71. doi: 10.1080/17402520600800085
45. Alberts R, de Vries EMG, Goode EC, Jiang X, Sampaziotis F, Rombouts K, et al. Genetic Association Analysis Identifies Variants Associated With Disease Progression in Primary Sclerosing Cholangitis. *Gut* (2018) 67(8):1517–24. doi: 10.1136/gutjnl-2016-313598
46. Ellinghaus D, Folseraas T, Holm K, Ellinghaus E, Melum E, Balschun T, et al. Genome-Wide Association Analysis in Primary Sclerosing Cholangitis and Ulcerative Colitis Identifies Risk Loci at GPR35 and TCF4. *Hepatology* (2013) 58(3):1074–83. doi: 10.1002/hep.25977
47. Srivastava B, Mells GF, Cordell HJ, Muriithi A, Brown M, Ellinghaus E, et al. Fine Mapping and Replication of Genetic Risk Loci in Primary Sclerosing Cholangitis. *Scand J Gastroenterol* (2012) 47(7):820–6. doi: 10.3109/00365521.2012.682090
48. Folseraas T, Melum E, Rausch P, Juran BD, Ellinghaus E, Shiryaev A, et al. Extended Analysis of a Genome-Wide Association Study in Primary Sclerosing Cholangitis Detects Multiple Novel Risk Loci. *J Hepatol* (2012) 57(2):366–75. doi: 10.1016/j.jhep.2012.03.031
49. Melum E, Franke A, Schramm C, Weismuller TJ, Gotthardt DN, Offner FA, et al. Genome-Wide Association Analysis in Primary Sclerosing Cholangitis Identifies Two Non-HLA Susceptibility Loci. *Nat Genet* (2011) 43(1):17–9. doi: 10.1038/ng.728
50. Daca-Roszak P, Pfeifer A, Zebracka-Gala J, Jarzab B, Witt M, Zietkiewicz E. Euresa_Gplex—a New Snapshot Assay for Continental Population Discrimination and Gender Identification. *Forensic Sci Int Genet* (2016) 20:89–100. doi: 10.1016/j.fsigen.2015.10.004
51. Gaunt TR, Shihab HA, Hemani G, Min JL, Woodward G, Lyttleton O, et al. Systematic Identification of Genetic Influences on Methylation Across the Human Life Course. *Genome Biol* (2016) 17:61. doi: 10.1186/s13059-016-0926-z
52. Murphy TM, Mill J. Epigenetics in Health and Disease: Herald of the EWAS Era. *Lancet* (2014) 383(9933):1952–4. doi: 10.1016/S0140-6736(14)60269-5
53. Jaffe AE, Irizarry RA. Accounting for Cellular Heterogeneity is Critical in Epigenome-Wide Association Studies. *Genome Biol* (2014) 15(2):R31. doi: 10.1186/gb-2014-15-2-r31
54. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA Methylation Arrays as Surrogate Measures of Cell Mixture Distribution. *BMC Bioinf* (2012) 13:86. doi: 10.1186/1471-2105-13-86
55. Walker LJ, Kang YH, Smith MO, Tharmalingham H, Ramamurthy N, Fleming VM, et al. Human MAIT and CD8 α Cells Develop From a Pool of Type-17 Precommitted CD8 $^{+}$ T Cells. *Blood* (2012) 119(2):422–33. doi: 10.1182/blood-2011-05-353789
56. Kunzmann LK, Schoknecht T, Poch T, Henze L, Stein S, Kriz M, et al. Monocytes as Potential Mediators of Pathogen-Induced T-Helper 17 Differentiation in Patients With Primary Sclerosing Cholangitis (PSC). *Hepatology* (2020) 72(4):1310–26. doi: 10.1002/hep.31140
57. McEwen LM, Jones MJ, Lin DTS, Edgar RD, Husquin LT, MacIsaac JL, et al. Systematic Evaluation of DNA Methylation Age Estimation With Common Preprocessing Methods and the Infinium Methylation Beadchip Array. *Clin Epigenet* (2018) 10(1):123. doi: 10.1186/s13148-018-0556-2
58. Meng J, Chen S, Han JX, Qian B, Wang XR, Zhong WL, et al. Twist1 Regulates Vimentin Through Cul2 Circular RNA to Promote EMT in Hepatocellular Carcinoma. *Cancer Res* (2018) 78(15):4150–62. doi: 10.1158/0008-5472.CAN-17-3009
59. Guo X, Yang M, Gu H, Zhao J, Zou L. Decreased Expression of SOX6 Confers a Poor Prognosis in Hepatocellular Carcinoma. *Cancer Epidemiol* (2013) 37(5):732–6. doi: 10.1016/j.canep.2013.05.002

60. Cao X, Zhang J, Apaer S, Yao G, Li T. MicroRNA-19a-3p and MicroRNA-376c-3p Promote Hepatocellular Carcinoma Progression Through SOX6-Mediated Wnt/Beta-Catenin Signaling Pathway. *Int J Gen Med* (2021) 14:89–102. doi: 10.2147/IJGM.S278538
61. Rivas MA, Beaudoin M, Gardet A, Stevens C, Sharma Y, Zhang CK, et al. Deep Resequencing of GWAS Loci Identifies Independent Rare Variants Associated With Inflammatory Bowel Disease. *Nat Genet* (2011) 43(11):1066–73. doi: 10.1038/ng.952
62. Kinchen J, Chen HH, Parikh K, Antanaviciute A, Jagielowicz M, Fawcner-Corbett D, et al. Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. *Cell* (2018) 175(2):372–86.e17. doi: 10.1016/j.cell.2018.08.067
63. Vycital O, Pitule P, Hosek P, Kriz T, Treska V, Liska V. Expression of Serpin B9 as a Prognostic Factor of Colorectal Cancer. *Anticancer Res* (2019) 39(11):6063–6. doi: 10.21873/anticancerres.13813
64. Li G, Zhou LN, Yang H, He X, Duan Y, Wu F. Ninjurin 2 Overexpression Promotes Human Colorectal Cancer Cell Growth *In Vitro* and *In Vivo*. *Aging (Albany NY)* (2019) 11(19):8526–41. doi: 10.18632/aging.102336
65. Daca-Roszak P, Pfeifer A, Zembracka-Gala J, Rusinek D, Szybinska A, Jarzab B, et al. Impact of Snps on Methylation Readouts by Illumina Infinium HumanMethylation450 Beadchip Array: Implications for Comparative Population Studies. *BMC Genomics* (2015) 16:1003. doi: 10.1186/s12864-015-2202-0
66. Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, et al. Genetic Control of Individual Differences in Gene-Specific Methylation in Human Brain. *Am J Hum Genet* (2010) 86(3):411–9. doi: 10.1016/j.ajhg.2010.02.005
67. Shi X, Radhakrishnan S, Wen J, Chen JY, Chen J, Lam BA, et al. Association of Cnvs With Methylation Variation. *NPJ Genom Med* (2020) 5(1):41. doi: 10.1038/s41525-020-00145-w
68. Araki T, Milbrandt J. Ninjurin2, a Novel Homophilic Adhesion Molecule, is Expressed in Mature Sensory and Enteric Neurons and Promotes Neurite Outgrowth. *J Neurosci* (2000) 20(1):187–95. doi: 10.1523/JNEUROSCI.20-01-00187.2000
69. Ahn BJ, Lee HJ, Shin MW, Choi JH, Jeong JW, Kim KW. Ninjurin1 is Expressed in Myeloid Cells and Mediates Endothelium Adhesion in the Brains of EAE Rats. *Biochem Biophys Res Commun* (2009) 387(2):321–5. doi: 10.1016/j.bbrc.2009.07.019
70. Wang J, Fa J, Wang P, Jia X, Peng H, Chen J, et al. NINJ2- a Novel Regulator of Endothelial Inflammation and Activation. *Cell Signal* (2017) 35:231–41. doi: 10.1016/j.celsig.2017.04.011
71. Rhead B, Brorson IS, Berge T, Adams C, Quach H, Moen SM, et al. Increased DNA Methylation of SLFN12 in CD4+ and CD8+ T Cells From Multiple Sclerosis Patients. *PLoS One* (2018) 13(10):e0206511. doi: 10.1371/journal.pone.0206511
72. Yousefi P, Huen K, Dave V, Barcellos L, Eskenazi B, Holland N. Sex Differences in DNA Methylation Assessed by 450 K Beadchip in Newborns. *BMC Genomics* (2015) 16:911. doi: 10.1186/s12864-015-2034-y

Conflict of Interest: Author EL was employed by Horaizon BV.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 de Krijger, Hageman, Li Yim, Verhoeff, Garcia Vallejo, van Hamersveld, Levin, Hakvoort, Wildenberg, Henneman, Ponsioen and de Jonge. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.