

ORIGINAL RESEARCH

Long-term Temporal Stability of Peripheral Blood DNA Methylation Profiles in Patients With Inflammatory Bowel Disease

Vincent Joustra,¹ Andrew Y. F. Li Yim,^{2,3} Ishtu Hageman,³ Evgeni Levin,^{4,5} Alex Adams,⁶ Jack Satsangi,⁶ Wouter J. de Jonge,³ Peter Henneman,² and Geert D'Haens¹

¹Department of Gastroenterology and Hepatology, Amsterdam Gastroenterology Endocrinology Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands; ²Genome Diagnostics Laboratory, Department of Human Genetics, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands; ³Tytgat Institute for Liver and Intestinal Research, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands; ⁴Department of Vascular Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands; ⁵Horaizon BV, Delft, the Netherlands; and ⁶Oxford University- Hospitals NHS Foundation Trust- John Radcliffe Hospital, Translational Gastroenterology Unit- NIHR Oxford Biomedical Research Centre, Oxford, United Kingdom

SUMMARY

This work provides insight into the long-term intra-individual stability of the peripheral blood DNA methylome in patients with inflammatory bowel disease, a key aspect of predictive biomarker development. The data could serve to pre-select stable biomarkers to increase the probability of independent validation. In addition, the marked stability HLA-associated cytosine-phosphate-guanines have potential implications in understanding disease pathogenesis.

BACKGROUND & AIMS: There is great current interest in the potential application of DNA methylation alterations in peripheral blood leukocytes (PBLs) as biomarkers of susceptibility, progression, and treatment response in inflammatory bowel disease (IBD). However, the intra-individual stability of PBL methylation in IBD has not been characterized. Here, we studied the long-term stability of all probes located on the Illumina HumanMethylation EPIC BeadChip array.

METHODS: We followed a cohort of 46 adult patients with IBD (36 Crohn's disease [CD], 10 ulcerative colitis [UC]; median age, 44 years; interquartile range [IQR] 27–56 years; 50% female) that received standard care follow-up at the Amsterdam University Medical Centers. Paired PBL samples were collected at 2 time points with a median of 7 years (range, 2–9 years) in between. Differential methylation and intra-class correlation (ICC) analyses were used to identify time-associated differences and temporally stable CpGs, respectively.

RESULTS: Around 60% of all EPIC array loci presented poor intra-individual stability (ICC <0.50); 78.114 (≈9%) showed good (ICC, 0.75–0.89), and 41.274 (≈5%) showed excellent (ICC ≥0.90) stability, between both measured time points. Focusing on previously identified consistently differentially methylated positions indicated that 22 CD-, 11 UC-, and 24 IBD-associated loci demonstrated high stability (ICC ≥0.75) over time; of these, we observed a marked stability of CpG loci associated to the *HLA* genes.

CONCLUSIONS: Our data provide insight into the long-term stability of the PBL DNA methylome within an IBD context,

facilitating the selection of biologically relevant and robust IBD-associated epigenetic biomarkers with increased potential for independent validation. These data also have potential implications in understanding disease pathogenesis. (*Cell Mol Gastroenterol Hepatol* 2023;15:869–885; <https://doi.org/10.1016/j.jcmgh.2022.12.011>)

Keywords: Biomarkers; Epigenetics; Personalized Medicine.

Crohn's disease (CD) and ulcerative colitis (UC) are chronic relapsing and remitting inflammatory bowel diseases (IBDs) characterized by a wide variety of phenotypic manifestations.¹ Although the etiology of IBD remains unknown, it is thought to arise as a result of a complex interplay between the host and microbial composition, triggered by environmental factors, such as tobacco smoking or diet.^{2–4}

Accordingly, much effort has been invested in understanding the interaction between host and environment, which is thought to be mediated by the epigenome.⁵ The epigenome represents the set of mitotically heritable modifications that can affect gene transcriptions without altering the primary DNA sequence.⁶ DNA methylation, one of the most studied epigenetic mechanisms, involves the attachment of methyl groups to cytosine-phosphate-guanine (CpG) nucleotide sequences on the DNA. This covalent attachment is mitotically heritable and can, under certain conditions,

Abbreviations used in this paper: CD, Crohn's disease; CpG, cytosine-phosphate-guanine; CRP, C-reactive protein; DMP, differentially methylated position; EWAS, epigenome-wide association study; FDR, false discovery rate; GVs, genetic variants; HSMP, hyper-stable methylated position; IBD, inflammatory bowel disease; ICC, intra-class correlation; IQR, interquartile range; PBLs, peripheral blood leukocytes; SMP, stably methylated position; UC, ulcerative colitis.

Most current article

© 2023 The Authors. Published by Elsevier Inc. on behalf of the AGA Institute. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2352-345X

<https://doi.org/10.1016/j.jcmgh.2022.12.011>

Table 1. Baseline Characteristics

Baseline characteristics	CD (n = 36)	UC (n = 10)	Total (N = 46)
Female	18 (50)	5 (50)	23 (50)
Age, y	42 (27–53)	52 (36–61)	44 (27–56)
Disease duration, y	11 (6–20)	13 (7–22)	12 (7–21)
Ethnic background			
Caucasian	30 (83.3)	9 (90)	39 (84.8)
Disease location CD			
Ileal disease (L1)	14 (38.9)	–	14 (38.9)
Colonic disease (L2)	4 (11.1)	–	4 (11.1)
Ileocolonic disease (L3)	18 (50)	–	18 (50)
Disease location UC			
Left-sided UC, distal to splenic flexure (E2)		3 (30)	3 (30)
Extensive, proximal to splenic flexure (E3)		7 (70)	7 (70)
Disease behavior			
Non-stricturing/penetrating (B1)	9 (25)	–	9 (25)
Stricturing (B2)	13 (36.1)	–	13 (36.1)
Penetrating (B3)	14 (38.9)	–	14 (38.9)
Perianal disease (p)	12 (36.1)	–	12 (36.1)
Previous IBD-related surgery	19 (52.8)	1 (10)	20 (43.5)
IBD-related surgery between T1 and T2	10 (27.8)	–	10 (21.7)
Previous medical treatment			
Immunomodulator (AZA, 6MP, 6TG, MTX)	29 (80.6)	7 (70)	36 (78.3)
Anti-TNF (IFX and/or ADA)	28 (77.8)	4 (40)	32 (69.6)
Ustekinumab	1 (2.8)	–	1 (2.2)
Treatment between T1 and T2			
Immunomodulator (AZA, 6MP, 6TG, MTX)	14 (38.9)	3 (30)	17 (37)
Anti-TNF (IFX and/or ADA)	16 (44.4)	7 (70)	23 (50)
Vedolizumab	17 (47.2)	7 (70)	24 (52.2)
Ustekinumab	13 (36.1)	1 (10)	14 (30.4)
CRP T1, mg/L	3.2 (1.4–8.0)	2 (0.4–5.8)	2.6 (1.2–6.6)
CRP T2, mg/L	3 (1.6–5.3)	2.5 (0.8–7.7)	3.0 (1.4–5.3)
Leukocyte count T1, *10 ⁹	7.1 (6.3–9.1)	5.4 (4.7–6.3)	6.7 (5.4–8.3)
Leukocyte count T2, *10 ⁹	7.1 (5.7–8.9)	5.8 (5.2–7.6)	7.1 (5.3–8.5)
Smoking			
Active	8 (22.2)	3 (30)	11 (23.9)
Non-smoker	28 (77.8)	7 (70)	35 (76.1)

Note: Data are presented as number (%) or median (interquartile range). CD, Crohn's disease; CRP, C-reactive protein; IBD, inflammatory bowel disease; UC, ulcerative colitis.

regulate gene expression, thereby altering cellular behavior.⁷ Over the past decade, multiple epigenome-wide association studies (EWAS) have sought to characterize, classify, and predict IBD and its various phenotypes using DNA methylation.^{8–13} However, most EWAS in IBD to date have been cross-sectional in design, reporting aberrant DNA methylation signatures in peripheral blood leukocytes (PBLs) and/or mucosal tissue,^{9,11,12,14–20} with only a single longitudinal study in mucosal tissue¹² and PBLs.²¹

Previous literature has shown that the intra-individual variability of DNA methylation is most prominent during the early stages of life, which gradually diminishes and presents a more stable phenotype after 5 years of life.^{22,23} Nonetheless, the influence of aging on genome-wide DNA methylation has been well-described in monozygotic twins^{24,25} and unrelated healthy populations,^{26,27} demonstrating a global decrease in methylation as individuals age, as well as site-specific increases in methylation in CpG-rich areas, both of which are thought

to result from dynamic external and internal environmental changes.^{28–30} As epigenetics, and thus DNA methylation, is cell-type specific, observed differences found in heterogeneous populations such as PBLs or tissue might reflect differences in the cellular composition.^{31–33} Nonetheless, age-related differences were found in more homogeneous populations, such as purified T-cells and monocytes.^{15,34,35}

Despite the strong effects of age on DNA methylation, a high correlation between baseline and follow-up methylation data in pediatric IBD mucosal tissue has been observed.¹² In contrast, IBD-associated differences in blood have shown to largely revert back to patterns observed in non-IBD controls during follow-up as the result of treatment and normalization of C-reactive protein (CRP).²¹ It is noteworthy that these studies focused only on a subset of IBD-associated CpGs, and did not report on the long-term stability of all CpG probes located on the Illumina Human-Methylation EPIC BeadChip array. Although temporal

stability and intra-individual variability in PBL-derived DNA methylation has been investigated in adult healthy individuals³⁶⁻³⁹ and patients with systemic lupus erythematosus,⁴⁰ no such study has been conducted in patients with IBD.

There is widespread interest in the application of epigenetic markers in personalization of treatment.⁴¹

If epigenetic biomarkers are to be used as pathognomonic for IBD or its (sub)phenotypes, the features of interest would need to remain stable throughout the duration of the disease and, hence, over time, without being affected by various internal/external exposures. In addition, for biomarker development, loci that are time-stable reduce the number of false-positive findings, thereby increasing the probability of independent replication. Furthermore, selection of time-stable epigenetic biomarkers would help overcome current practical barriers in sample collection at specific time frames, thereby facilitating the use of larger samples sizes with similar phenotypes needed to enhance predictive power. We therefore sought to identify CpG positions that present stable DNA methylation in PBLs obtained from a well-generalizable cohort of adult patients with IBD with a 7-year median time between collected DNA samples.

Results

Patient Demographics

A total of 46 adult patients with IBD (36 CD, 10 UC) with a median age of 44 years (interquartile range [IQR], 27–56 years) and median disease duration of 12 years (IQR, 7–21 years) were included. Gender, surgical history, disease location, and disease behavior were balanced within this cohort (Table 1). Notably, 32 patients (69.6%) were previously treated with an anti-tumor necrosis factor, prior to T1 sampling. Between T1 and T2 during regular IBD care follow-up, 10 patients (21.7%) underwent IBD-related surgery, 24 patients (52.2%) were treated with vedolizumab, and 14 patients (30.4%) were treated with ustekinumab, reflecting the tertiary referral population seen at the Amsterdam University Medical Centers (UMC). No significant differences in median CRP ($P = .97$) or leukocyte count ($P = .85$) between T1 and T2 were observed (Figure 1, B–C).

Time-associated Differential Methylation

Expectedly Associates With Age-related CpGs

We first investigated the differences in methylation between both time points, identifying 194,391 ($\approx 23\%$) differentially methylated positions (DMPs) when comparing T1 and T2 at a false-discovery rate (FDR)-adjusted P -value of $< .05$ (Figure 2, A–C), which we termed time-associated DMPs. As our sample of interest was derived from peripheral blood, we investigated whether differences in the cellular composition were observable. Comparing the predicted blood cell composition yielded significant increase of the B- ($P = .017$) and CD4+ T-cells ($P = .013$) over time, whereas the neutrophils present a significant decrease ($P = .02$) (Figure 2, D).

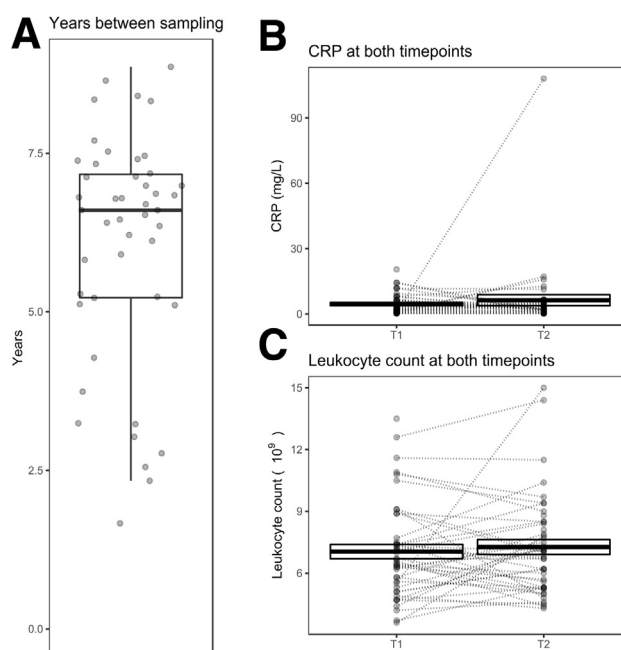


Figure 1. Patient characteristics over time. (A), Visualization of the number of years between both samplings per patient. Visualization of the CRP (mg/L) (B) and leukocyte count (10^9) (C) between both time points, where connected samples were obtained from the same patient annotated with the mean difference and P -value.

Expectedly, the time-associated differences were enriched for age-associated CpGs, which have been defined as the “epigenetic clocks” from Horvath,⁴² Hannum,⁴³ Levine,⁴⁴ and Knight⁴⁵ (Figure 2, E). Furthermore, for these specific epigenetic clock CpGs, we observed a general hypomethylated pattern at T2 relative to T1 CpG sites (Figure 2, F), suggesting that the observed differences in DNA methylation are enriched for age-related differences. Functional enrichment analyses of the time-associated DMPs displayed several cancer-associated pathways (Figure 3).

Time-invariant, Stable Methylated Probes are Enriched in Genes Involved in Cell Adhesion

To identify CpGs that were consistently methylated at both time points, we performed intra-class correlation (ICC) analysis, which indicated that the majority of the CpGs (517,576 probes or around 60%) present poor intra-individual stability over time ($ICC < 0.50$) (Table 2). Conversely, 119,388 CpGs ($\approx 14\%$) displayed a statistically significant high ICC (≥ 0.75), which we termed stably methylated positions (SMPs). Expectedly, CpGs with high ICC values typically presented less difference in mean methylation (Figure 4). We reasoned that probes that were associated with sites known to harbor genetic variants, both intentional and unintentional,⁴⁶ should present the highest stability, as the genome of an individual typically does not change over the course of 7 years. Indeed, splitting the data by modality suggested that the CpG sites associated with known germline variants, namely those that were included

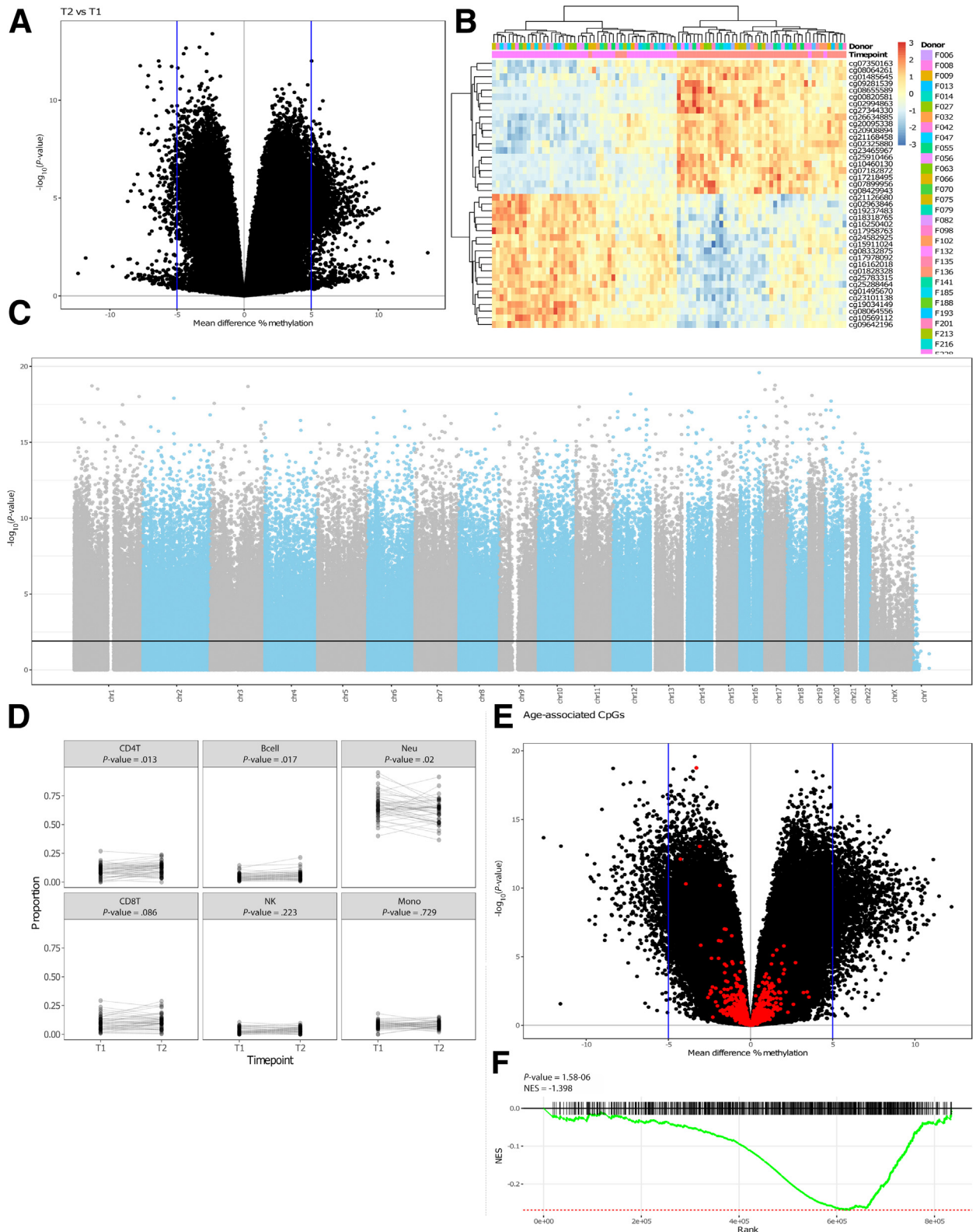


Figure 2. Time-variant methylated positions. (A), A volcano plot depicting the mean difference in methylation between the 2 time points on the x-axis and the $-\log_{10}(P\text{-value})$ on the y-axis. (B), Heatmap visualizing the percentage methylation for the 25 most hyper- and 25 most hypo methylated DMPs. (C), Manhattan plot showing the chromosomal distribution of all Illumina HumanMethylation EPIC array probes. Each dot represent a single CpG locus; dots above the black line are statistically significantly different between T1 and T2 (FDR-adjusted $P\text{-value} \leq .05$). (D), Estimated blood cell distribution stratified by time. Dashed lines connect samples obtained from the same donor. Statistical significance was calculated using a Mann-Whitney U test. (E), Volcano plot colored for age-associated CpGs. (F), Gene set enrichment analysis barcode plot representing the overrepresentation of the age-related CpGs among the time-associated DMPs.

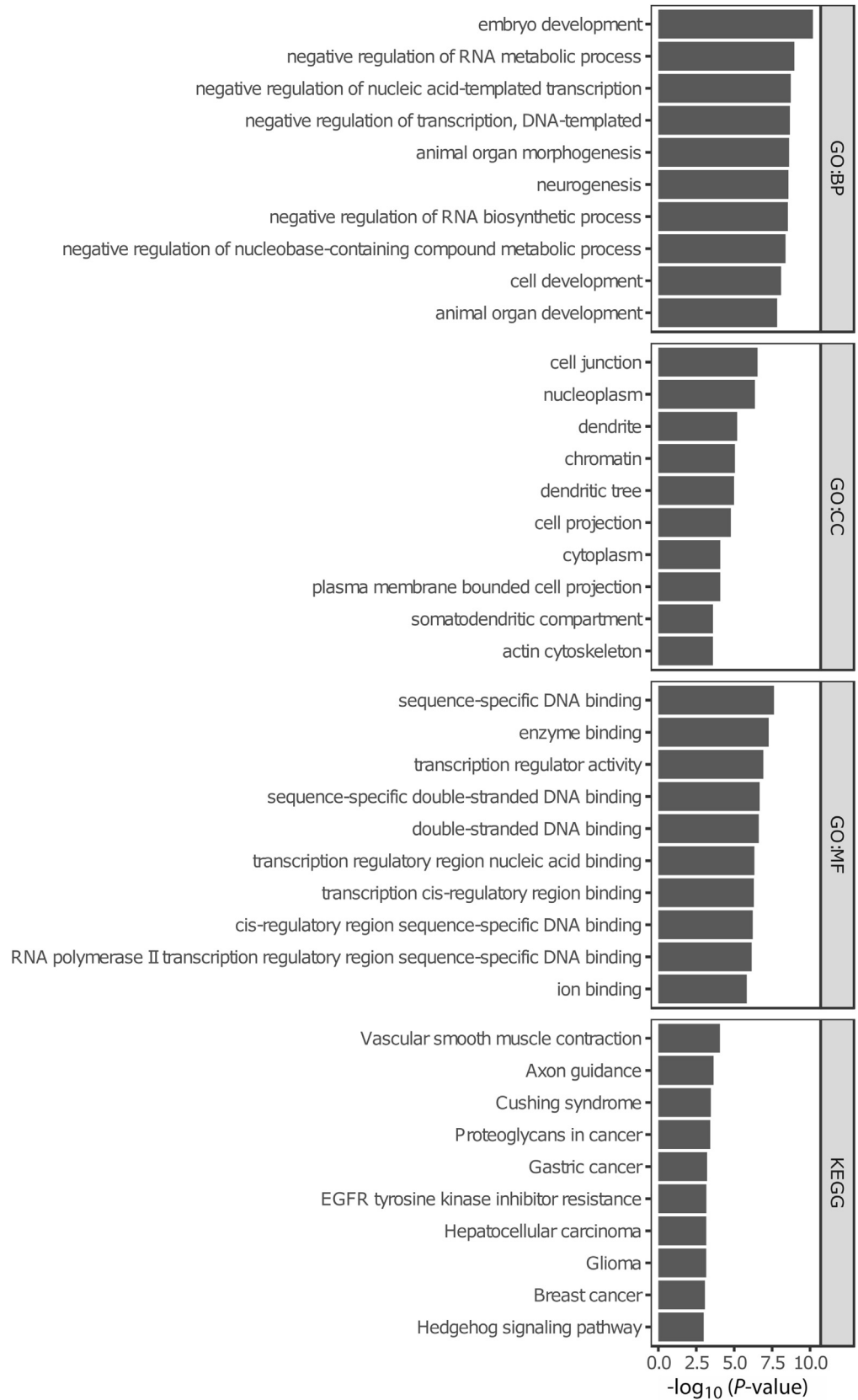


Figure 3. Functional enrichments analyses using Gene Ontology-term and Kyoto Encyclopedia of Genes and Genomes pathways for the time-associated DMPs.

Table 2. ICC Scores

	Poor (ICC <0.5)	Moderate (ICC, 0.5–0.74)	Good (ICC, 0.75–0.89)	Excellent (0.9 \geq ICC)
QC GV	0	0	0	59
Annotated GV	10462	4141	1539	576
Predicted GV	18583	25029	22199	24873
Methylation	488531	168790	54376	15766

Note: An overview of the ICC values classified using the system presented by Koo and Li.⁴⁷ QC GV = Quality control probes that bind genetic variant only. Annotated GV = Methylation probes that are annotated to bind genetic variants at the CpG of interest. Predicted GV = Methylation probes that are annotated to bind genetic variants that were annotated by Gaphunter to be caused by genetic variants. Methylation = Methylation probes for which we have no evidence that they bind genetic variants.

GV, Genetic variants; ICC, interclass correlation.

for quality control purposes, presented high (>0.9) ICC values (Figure 5, A and Table 2). Of the SMPs with ICC values over 0.9, 15,766 SMPs (around 2%) presented no indication that they bind predicted or potential genetic variants, which we classified as hyper-stable methylated positions (HSMPs) (Table 3 and Figure 5, B). Functional

enrichment analyses of the SMPs and HSMPs indicated significant enrichment of genes involved in cell-cell signaling, adhesion and neurogenesis (Figures 6 and 7).

Stability Analysis of Previous IBD-associated DMPs, HLA, and IBD-susceptibility genes

We next investigated whether previously reported IBD-associated DMPs were found to be invariant over time. To do so, we evaluated ICC values of 255 CD-associated, 103 UC-associated, and 221 IBD-associated consistent DMPs identified in our systematic review and meta-analysis on CD-, UC-, and IBD-associated differential methylation,⁴⁸ which included a total of 552 samples (177 CD, 132 UC, and 243 HC) from 4 different EWAS.^{14,17-19} Focusing on the stability of these DMPs in this cohort, we show that the majority (151 or 59.2% CD-associated, 73 or 70.9% UC-associated, and 156 or 70.6% IBD-associated) present poor to moderate stability, indicating that the methylation status of these DMPs are affected by age or other exposures over time (Figure 8, A; Supplementary Table 1). Nonetheless, 22 CD-associated (12.4%), 11 UC-associated (8.3%), and 24 IBD-associated (9.9%) loci show good to excellent stability over time, providing evidence that these CD-, UC-, or IBD-associated DMPs are unaffected by aging or the exposures over time (Figure 8, A; Supplementary Table 1).

Among the many IBD-associated DMPs, we specifically zoomed in on *VMP1* (cg12054453 and cg16936953) as well as *RPS6KA2* (cg17501210), as they were identified in multiple IBD-EWAS,^{17,18,21} as well as shown found to be among the most significant IBD-associated DMPs in our meta-analysis.⁴⁸ We observed moderate consistency over time, with noticeable overall hypermethylation at T2 relative to T1 for the aforementioned 3 CpGs (Figure 8, B).

In addition to our meta-analysis, we also interrogated the CpGs that were CD-associated but CRP-independent, as reported on by Somnineni et al.²¹ Interrogation thereof using our cohort revealed that cg25112191 (*RORC*), cg13707793 (*CXXC5*), cg21049840 (*GPR183*), cg06460200 (*GPR183*), cg06366627 (*DIDO1*), and cg15860510 (*ESP8L3*) presented poor consistency (ICC <0.5); cg04570316 (*GMNN*), cg02240291 (*SMARCD3*), cg00092736 (*ESPNL*), and cg00092736 (*ESPNL*) presented moderate consistency (ICC,

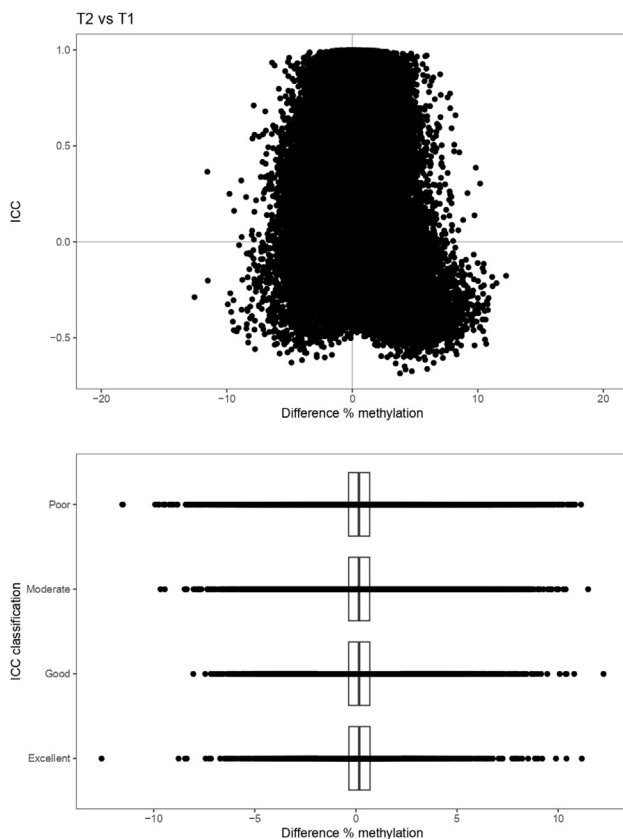


Figure 4. Difference in methylation according to ICC values. The x-axis represents the mean difference in methylation relative to the ICC on the y-axis. Note that the shape is slight conical, with the mean difference in methylation decreasing as the ICC increases. This becomes slightly more visible when binning the ICC values by the poor (ICC < 0.5), moderate (0.5 \leq ICC < 0.75), good (0.75 \leq ICC < 0.9), and excellent (0.9 \leq ICC) ICC classification, shown below.

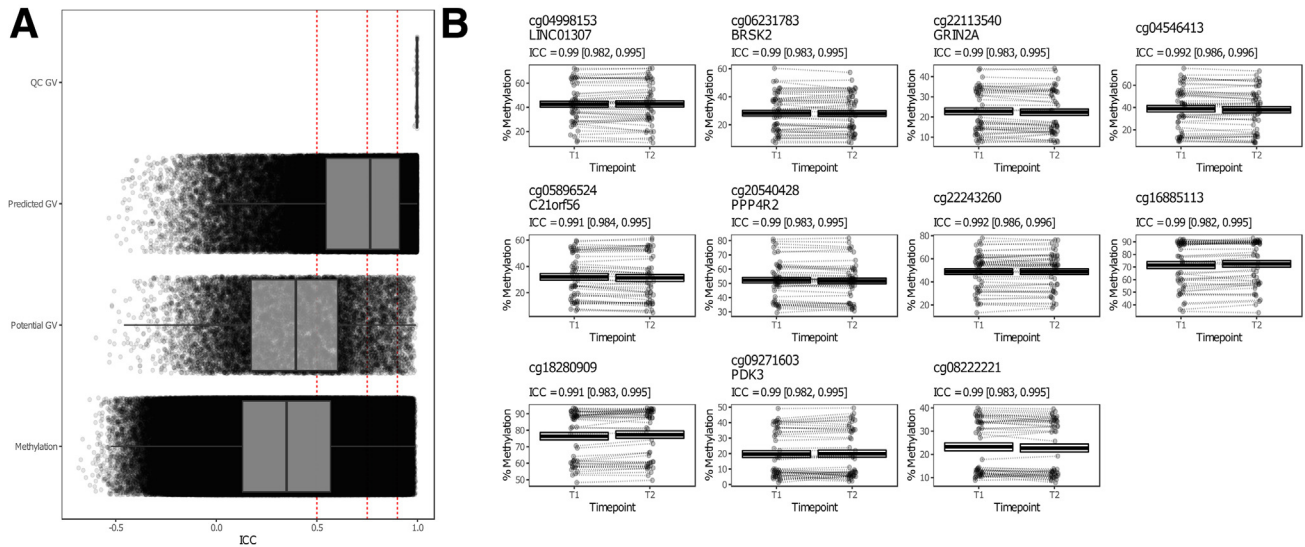


Figure 5. Time-invariant methylated positions. (A), The ICCs stratified by probe type, with QC GV representing the aforementioned quality control probes that bind GVs, the potential GV representing probes that were annotated with a genetic variant, and predicted GV representing probes with a methylation signal typically found when driven by a genetic variant. *Red dashed lines* represent the classification boundaries introduced by Koo and Li,⁴⁷ with blocks representing poor ($ICC < 0.5$), moderate ($0.5 \leq ICC < 0.75$), good ($0.75 \leq ICC < 0.9$), and excellent ($0.9 \geq ICC$). (B), Jittered visualization of the 11 probes that present a time in variant difference that is as stable as the aforementioned quality control probes with the percentage methylation on the y-axis and the time point on the x-axis. *Dashed lines* connect samples obtained from the same donor. The cross-bar visualization represents the mean and standard error of the mean.

0.5–0.74); and cg09171692 (*RORC*) and cg02055816 (*SMARCD3*) presented good consistency (ICC, 0.75–0.89) between both time points (Figure 8, C).

We next interrogated the stability of all CpG loci associated with several well-known GWAS-identified IBD risk genes involved in IBD pathogenesis, namely *ATG16L1*, *NOD2*, *IL23R*, *CARD9*, *FUT2*, *TYK2*, and *TNFSF15*,^{49,50} as well as specific IBD-associated major histocompatibility complex encoding *HLA* genes previously reported on in GWAS studies, namely *HLA-DRB1*, *HLA-DQB1*, *HLA-DQA1*, *HLA-DPA1*, *HLA-DPB1*, *HLA-A*, *HLA-B*, and *HLA-C*.^{49,51–54} Comparing all IBD risk genes, we noticed that the *HLA* genes presented the highest stability, all of which had a median ICC score over 0.5, whereas the majority of CpGs that annotate to non-*HLA* IBD risk genes had poor ICC values (< 0.5) (Figure 9, A). Nonetheless, for each of these non-*HLA* IBD risk genes, we identified highly stable methylated positions, several of which located to transcription start sites or first exons (Figure 10 and Supplementary Table 2), implicating potential regulatory function.

As DNA methylation measured using deamination technologies cannot distinguish DNA methylation from genetic variants located at the CpG of interest,⁴⁶ we investigated whether such technical artefacts were found among the *HLA* SMPs by interrogating the dbSNP (v151) database for catalogued variants, as well as by investigating for a typical clustered methylation signal when probes hybridize with genetic variants (GVs) using Gaphunter.⁵⁵ Notably, most of the high ICC values were found for CpGs that presented some type of clustering typical of GV but were not necessarily catalogued in dbSNP (v151) (Figure 9, B). In addition, *HLA*

class II genes (*HLA-DPA1*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, and *HLA-DRB1*) appeared to have a larger proportion of highly stable probes compared with *HLA* class I genes (*HLA-A*, *HLA-B*, and *HLA-C*) (Figure 9, B). Besides technical artefacts, DNA methylation itself can be affected by GVs that occur in the vicinity.⁵⁶ As such, we cross-referenced our observations with a previous large-scale genotyping study of the *HLA* region in both patients with CD and UC.⁵² We indeed found multiple probes within the vicinity (< 1 Kb) of CD- or UC-associated *HLA* alleles, many of which did not appear to have an annotated dbSNP identifier, indicating that the observed differences were either unannotated GVs, or CpGs whose methylation status is strictly controlled by neighboring IBD-associated GVs. Notably, several highly stable probes ($ICC \geq 0.75$) found to be annotated to *HLA-C*, *HLA-DPB1*, and *HLA-DPA1* were located far away (> 1 Kb) from any of the IBD-associated GVs, did not associate with catalogued GVs, nor were identified as a potential GV by Gaphunter (Figure 9, C–E; Figure 11; and Supplementary Table 3), suggesting by default strong methylation stability over time.

Discussion

Biomarker research often involves samples taken prior to or within a strictly pre-defined timeframe to the outcome of interest to mitigate the number of additional variables. In this study, we performed long-term longitudinal stability analyses of the PBL DNA methylome obtained from a cohort of adult patients with IBD (36 CD and 10 UC) that were collected at 2 time points separated by a median of 7 years reflective of a real-life tertiary referral population.

Table 3. Hyper-stable Methylated Positions (ICC ≥ 0.9)

CGID	Coordinate	ICC	P -value _{ICC}	P -adjusted _{ICC}	DM	P -value _{DM}	P -adjusted _{DM}	Gene
cg04998153	chr1:101823331	0.990153046	2.24E-40	4.64E-38	0.002350766	.917048378	.960490142	LINC01307 (body)
cg06231783	chr11:1475048	0.990318642	1.53E-40	3.24E-38	-0.002016075	.913436831	.958698404	BRSK2 (body)
cg22113540	chr16:10125521	0.990357545	1.40E-40	2.98E-38	-0.005125811	.726564998	.852281861	GRIN2A (body)
cg04546413	chr19:29218101	0.992092045	1.64E-42	4.48E-40	-0.008758714	.737190388	.859028126	
cg05896524	chr21:47604654	0.991011374	2.90E-41	6.71E-39	-0.00862793	.708898078	.841027778	C21orf56 (TSS1500)
cg20540428	chr3:73045686	0.990348581	1.43E-40	3.03E-38	-0.005080004	.778207203	.884339751	PPP4R2 (TSS1500)
cg22243260	chr3:126946036	0.991961418	2.37E-42	6.31E-40	-0.000283585	.989690923	.995366819	
cg16885113	chr6:29648507	0.990173675	2.14E-40	4.44E-38	0.009565754	.695688826	.832438074	
cg18280909	chr6:29723301	0.990760129	5.38E-41	1.21E-38	0.009797484	.631735087	.788308313	
cg09271603	chrX:24482885	0.990091759	2.57E-40	5.29E-38	0.002551285	.901311297	.952513118	PDK3 (TSS1500)
cg08222221	chrX:139589617	0.990360759	1.39E-40	2.95E-38	-0.005243195	.737019022	.858927936	

CGID, Illumina CpG identifier; Coordinate, Genomic coordinate of the CpG on the human genome (build hg19); DM, mean difference in percentage methylation; ICC, intraclass correlation coefficient; Gene, associated gene as well as the location in the gene; P -value_{ICC}, P -value associated with the intraclass correlation coefficient; P -adjusted_{ICC}, Benjamini-Hochberg-adjusted P -value associated with the intraclass correlation coefficient; P -value_{DM}, P -value associated with the mean difference in percentage methylation; P -adjusted_{DM}, Benjamini-Hochberg-adjusted P -value associated with the mean difference in percentage methylation.

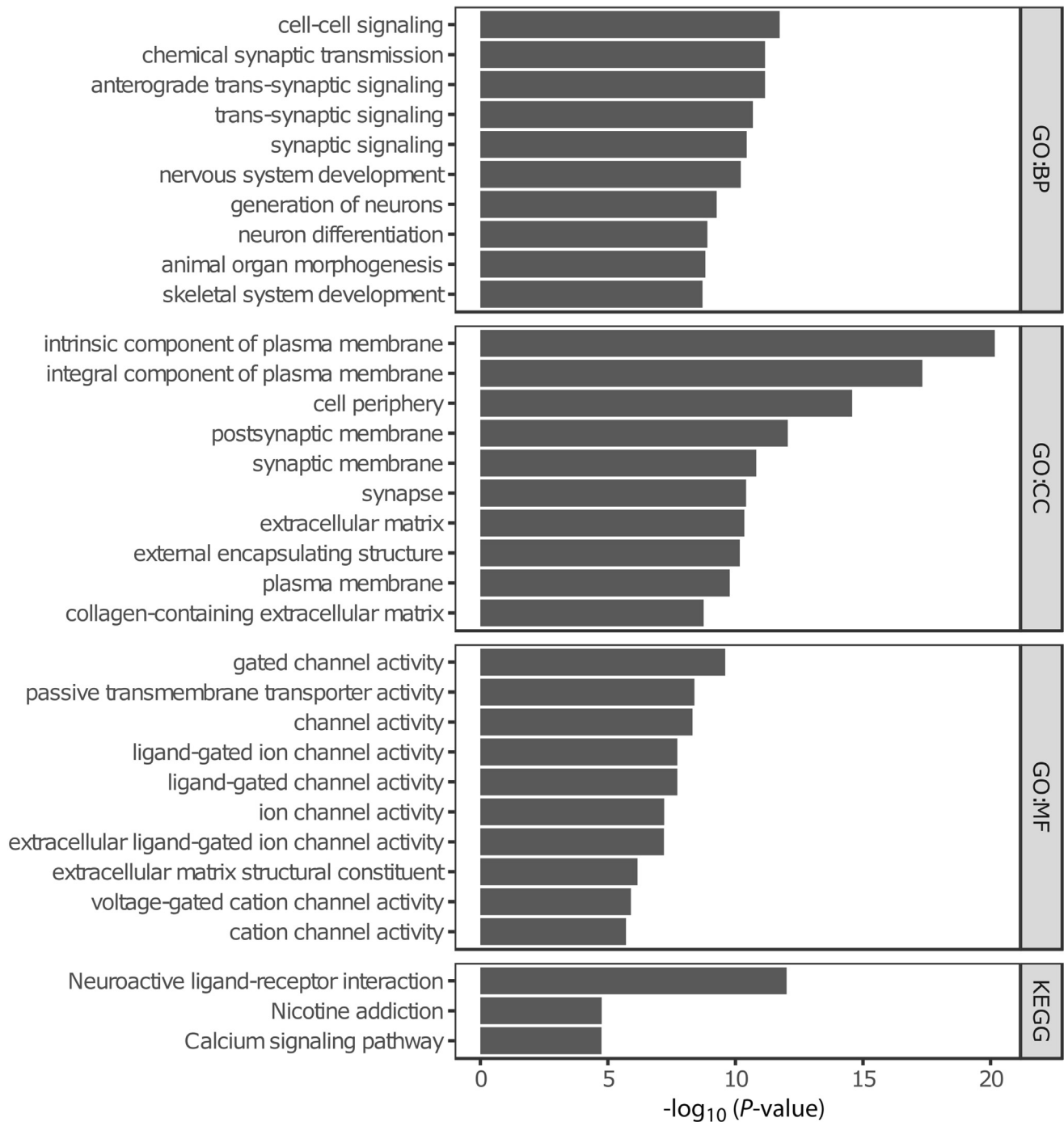


Figure 6. Functional enrichments analyses using Gene Ontology term and Kyoto Encyclopedia of Genes and Genomes pathways for the SMPs.

Our observations indicate that the majority of all loci ($\approx 60\%$) measured on the Illumina HumanMethylation EPIC BeadChip array present notable intra-individual variation in methylation over time ($ICC < 0.5$), which is enriched for age-associated CpGs. Nonetheless, not all time-associated DMPs in our cohort were previously reported as age-associated CpGs. Although we observed no significant differences in CRP and leukocyte count between both time points, other external or environmental

exposures, such as smoking, dietary alterations, therapy failure or switch, IBD-related surgery, or disease progression might have altered methylation status contributing to the observed time-associated differences. By contrast, 119,388 ($\approx 14\%$) and 41,274 ($\approx 5\%$) loci presented a highly stable pattern across both time points, with ICC values ≥ 0.75 and ≥ 0.90 , respectively. Such loci retained their degree of methylation even after the aforementioned known IBD-associated and unknown

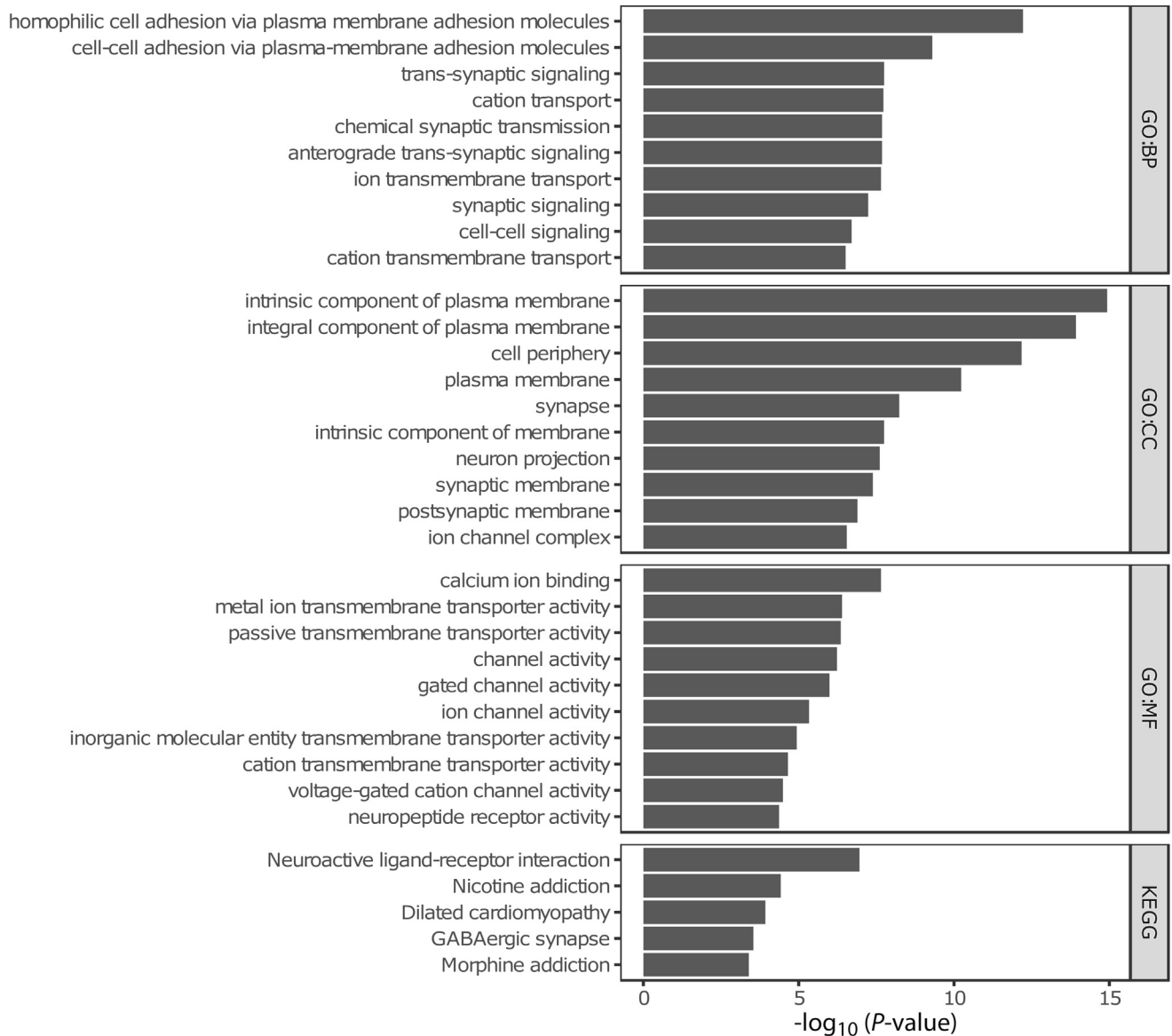


Figure 7. Functional enrichments analyses using Gene Ontology term and Kyoto Encyclopedia of Genes and Genomes pathways for the HSMPs.

external exposures as well as differences in cellular composition, suggesting time stability.

Previous studies investigating DNA methylation stability in PBLs of healthy adults using both the Illumina HumanMethylation EPIC BeadChip array, as well as its predecessor the Illumina HumanMethylation 450k BeadChip array, presented similar observations. In these studies, 16.9% to 23% of the CpG loci presented a moderate/good (ICC, 0.50–0.79), whereas 8.3% to 12.9% of the CpG loci presented a good/excellent (ICC ≥ 0.8) stability over a span of 1 to 6 years.^{37,39} Focusing on IBD, when interrogating 255 CD-associated, 103 UC-associated, and 221 IBD-associated DMPs identified in our own meta-analysis⁴⁸ of 4 IBD EWAS,^{14,17-19} we observed that the majority presented poor to moderate stability, suggesting that the aforementioned IBD-associated loci might also be affected by exposures over time that might or might not be

related with IBD. Although interesting, such time-variant probes should be interpreted with care when used as predictive biomarkers, given their association with exposures that occurred during both time points. To that end, our data could be used as a resource to preselect time-invariant CpG loci before independent validation when performing IBD-associated EWAS,⁵⁷ thereby increasing the potential to identify replicable predictive biomarkers better reflecting the underlying biology of IBD. In addition, such an approach would enable a larger pool of samples to be used as samples need not to be obtained within the same age range when performing DNA methylation studies on IBD and its phenotypes.

When specifically interrogating the IBD-associated probes cg12054453 (*VMP1*), cg16936953 (*VMP1*), and cg17501210 (*RPS6KA2*), we find moderate consistency over time with a noticeable hypermethylation at T2 compared with T1. Similar

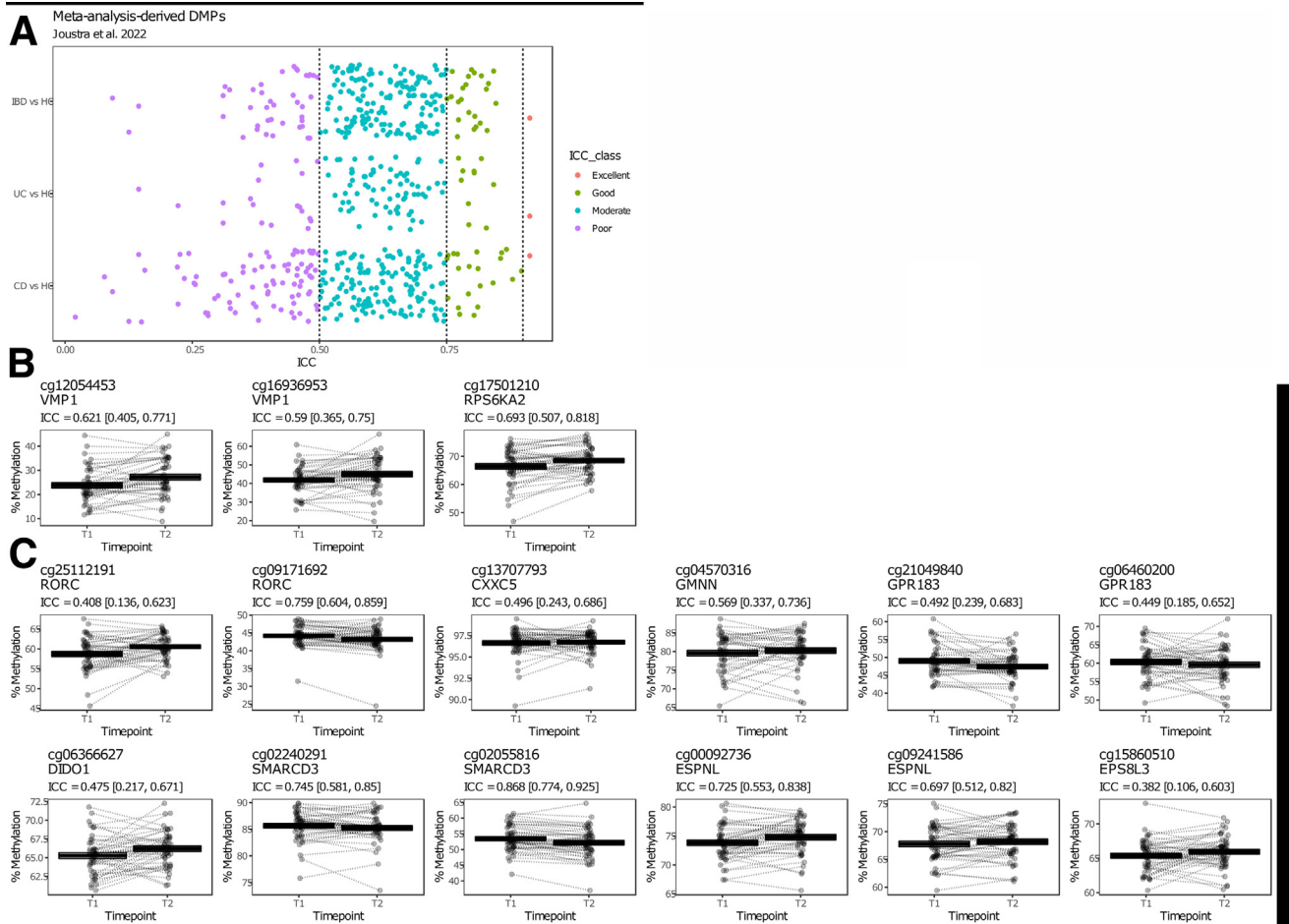


Figure 8. IBD-associated SMPs. (A), The ICCs of consistent DMPs identified through meta-analyses of 4 EWAS^{14,17-19} by Joustra et al⁴⁸ stratified by comparison (CD vs HC; UC vs HC; and IBD vs HC). *Black dashed lines* represent the classification boundaries introduced by Koo and Li,⁴⁷ with blocks representing poor (ICC < 0.5), moderate (0.5 ≤ ICC < 0.75), good (0.75 ≤ ICC < 0.9), and excellent (0.9 ≥ ICC). (B), Jittered visualization of the IBD-associated DMPs as reported on by Adams et al¹⁸ and Venham et al,¹⁷ as well as the CRP-independent probes reported on by Sominen et al (C).²¹ The percentage methylation is plotted on the y-axis and the time point on the x-axis. *Dashed lines* connect samples obtained from the same donor. The cross-bar visualization represents the mean and standard error of the mean.

analyses performed on the CD-associated yet CRP-independent probes reported by Sominen et al²¹ showed that the majority of these CpGs did not present long-term stability in our cohort. Differences between our observations and that of Sominen might be attributable to differences between adult and paediatric cohorts or might simply reflect non-inflammatory changes in methylation that occur over time. By contrast, cg09171692 (*RORC*) and cg02055816 (*SMARCD3*) presented high ICC values (0.76 and 0.87, respectively), indicating good stability in our cohort, irrespective of CRP or non-inflammatory exposures. Notably, both genes have previously been associated with IBD in multiple studies^{21,58,59} and therefore, show promise as stable IBD-associated loci.

Given the complex, multifactorial nature of IBD,² focusing on the interplay between genetic variation and DNA methylation rather than single gene mutations alone might prove more useful in understanding its molecular etiology. Previous EWAS of mucosal tissue⁶⁰ and peripheral blood⁵¹ both demonstrated differential methylation between IBD and controls for well-known GWAS-identified

IBD-associated risk genes, suggesting differential methylation of key risk genes to affect disease susceptibility. Our observations corroborate this hypothesis, showing highly stable methylation for particular CpG loci within these risk genes. Interestingly, several of these HSMPs were located in or near to the transcription start sites, potentially regulating gene transcription by maintaining the aberrant phenotype.¹⁸

There has been extensive interest in (epi)genetic alterations of the highly polymorphic *HLA* region related to IBD pathogenesis, most consistently reported for *HLA* class II genes involved in the presentation of bacterial antigens to CD4+ T-cells.^{14,18,49,51-54,62,63} Specifically, genetic variation of classical *HLA* genes has been suggested to play a role in the aberrant response to the dysbiotic microbiome observed in IBD,⁵¹ with particular impacts for the response to biological treatment⁶⁴⁻⁶⁶ and the formation of anti-drug antibodies.⁵⁴ However, translation of the results into clinical practice has proven to be difficult due to the high number of polymorphisms of *HLA* αβ heterodimers and strong linkage disequilibrium.⁵¹

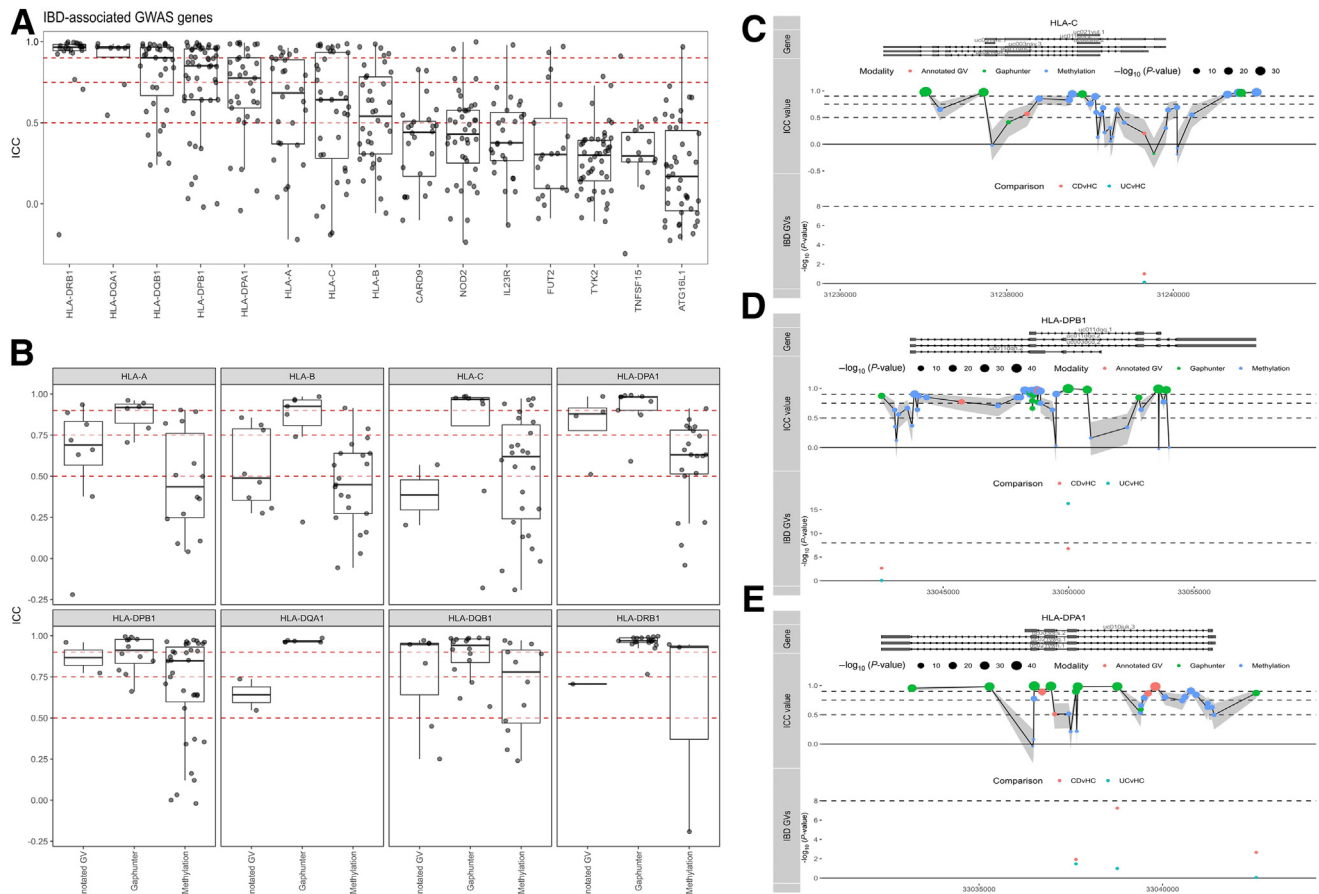


Figure 9. IBD risk genes. (A), Visualisation of the ICCs for all CpGs annotated to IBD-associated GWAS genes. Box plots show the overall median stability within each gene. Red dashed lines represent the classification boundaries introduced by Koo and Li,⁴⁷ with blocks representing poor ($ICC < 0.5$), moderate ($0.5 \leq ICC < 0.75$), good ($0.75 \leq ICC < 0.9$), and excellent ($0.9 \geq ICC$). (B), ICC values of all CpG loci of class I and II HLA genes. The potential GV representing probes that were annotated with a genetic variant, predicted GV representing probes that presented a methylation signal typically found when driven by a GV, and methylation representing probes for which we have no evidence that they hybridize with any GV. Visualizations of the ICC values of all Illumina CpGs annotated to HLA-C (C), HLA-DPB1 (D), and HLA-DPA1 (E), relative to their position on each gene and grouped as potential GV (pink), predicted GV (green), or methylation (blue). Dots below represent known genetic variants as reported by Goyette53 for CD vs healthy controls (pink) and UC vs healthy controls (turquoise).

In our study, we observed multiple HSMPs in HLA genes, suggesting that the DNA methylation profile of these genes is very stable over time. Our results corroborate with previous array data showing highly significant correlations between CpG loci on several *HLA* class II genes of neonates compared with toddlers ($r = 0.83$) and adults ($r = 0.88$) with type 1 diabetes.⁶⁷ Although further interrogation of these HSMPs indicates that multiple CpGs might be actual genetic variants, we also find multiple HSMPs that are not genetic variants. Nonetheless, several of such epigenetic HSMPs do occur within the vicinity of known IBD-associated HLA-variants, providing evidence that particular HLA-alleles might impart a strong, stabilizing effect on the epigenome.

Strengths and Limitations

To our knowledge, we are the first to assess the stability of the DNA methylome obtained from PBLs of patients with IBD with a median 7-year follow-up period in a real-life disease

exposure setting. This study is explorative in nature, using a moderate sample size without prior power calculation. Nonetheless, we note that studies of similar design have been conducted with a similar sample size.^{37,40,57} Although we can be reasonably confident in identifying the time-invariant aspect of the SMPs, we cannot fully eliminate the possibility that the SMPs would remain stable in a more diverse IBD cohort, as the typical markers of inflammation (CRP and leukocyte count) were hardly different between both time points.

Conclusion

We observe considerable variability in DNA methylation measurements taken from PBL at 2 different time points separated by a median of 7 years. By contrast, around 14% of all CpG loci could be considered highly stable even after IBD-specific exposures during the 2 points. Focusing on these CpG loci during biomarker discovery might result in the identification of biologically relevant and more robust

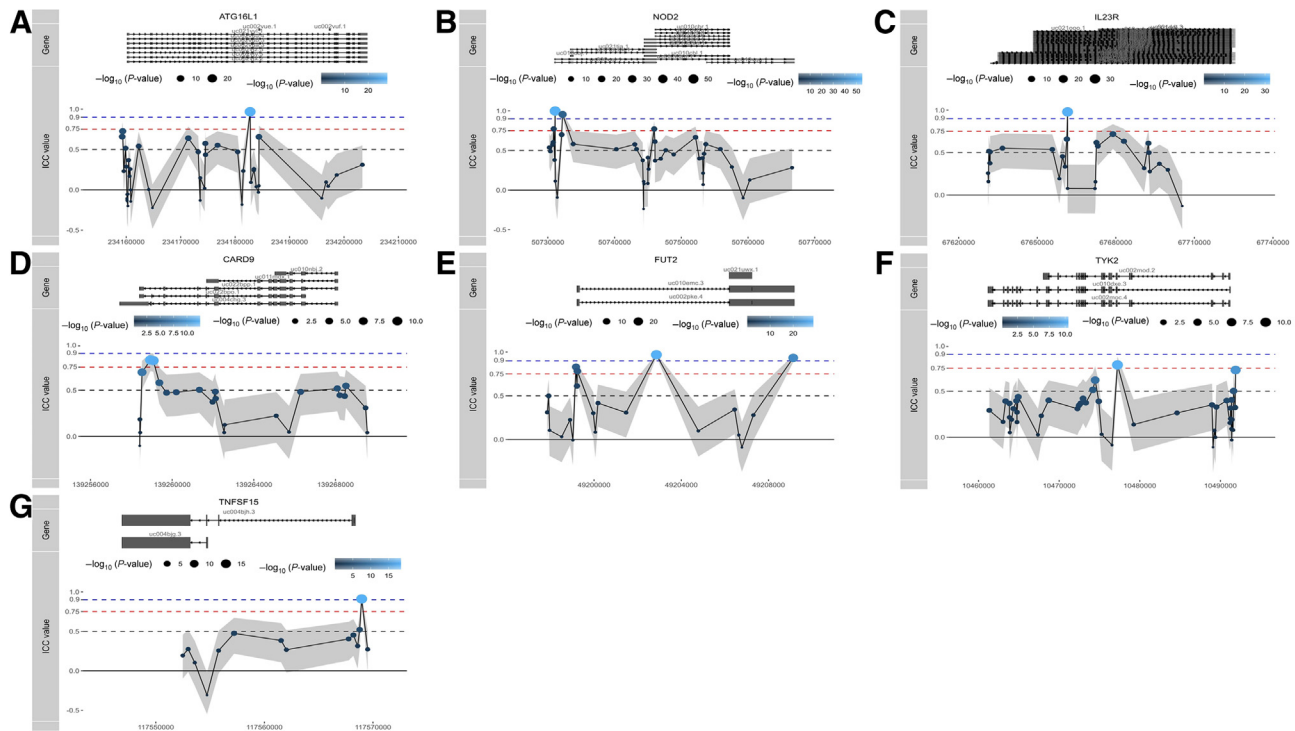


Figure 10. Visualization of the ICC values of all Illumina CpGs annotated to IBD GWAS risk genes, relative to their position on each gene. Dots represent individual CpG loci, dots above the red dashed line are considered SMPs ($ICC \geq 0.75$), whereas dots above the blue dashed line are HSMPs ($ICC \geq 0.9$).

IBD-associated epigenetic biomarkers with an increased probability of independent replication.

Methods

Patient Selection

We performed a single-center, longitudinal EWAS, where we collected PBL samples from adult patients with IBD at the Amsterdam UMC. The interval between the time of sampling ranged from 2 to 9 years with a median of 7 (Figure 1, A). All included patients were historically diagnosed with either CD or UC on the basis of a combination of clinical symptoms and endoscopic inflammation as confirmed by histology per the current guidelines.^{68,69} In addition, all patients received standard care follow-up. No additional inclusion or exclusion criteria were used as the goal was to collect a cohort of patients with IBD that reflected the overall IBD population at the Amsterdam UMC. This study was approved by the medical ethics committee of the academic medical hospital (METC NL24572.018.08 and NL53989.018.15), and written informed consent was obtained from all subjects prior to sampling.

Sample Collection and DNA Methylation Analysis

Whole peripheral blood samples were collected in a 6-mL EDTA tube and stored at -80°C until further processing. Genomic DNA was isolated using the QIAasympomy, whereupon the quantity of the DNA was assessed using the FLUOstar OMEGA and quality of the high-molecular weight DNA on a

0.8% agarose gel. Genomic DNA was bisulfite converted using the Zymo EZ DNA Methylation kit, randomized per plate to limit batch effects, and analyzed on the Illumina Human-Methylation EPIC BeadChip array at the Core Facility Genomics, Amsterdam UMC, Amsterdam, the Netherlands.

Statistical Analysis of Clinical Data

Baseline characteristics of all included patients were summarized using descriptive statistics. Categorical variables are presented as percentages and continuous variables as median and IQR. Differences in CRP and leukocyte count levels between T1 and T2 were calculated using the Wilcoxon signed ranks test. Analyses of clinical data were performed in IBM SPSS statistics version 26 and methylation analyses in the R statistical environment version 4.2.1.

Time-dependent DNA Methylation Data Analyses

For differential methylation analyses, raw DNA methylation data were imported into the R statistical environment using the Bioconductor *minfi*⁷⁰ package (version 1.36), whereupon the raw signal intensities were normalized using functional normalization⁷¹ and converted into methylation ratios. Differential methylation analyses was performed using *limma*⁷² (version 3.46) and *eBayes*⁷³ regressing against time point (T2 vs T1), gender, smoking behavior, disease, and blood cell distribution. Statistical significance was defined as an FDR-adjusted P -value $< .05$. In addition to identifying time-associated differences in methylation, we also investigated differences in methylation associated with CRP and leukocyte count. Blood cell estimations were

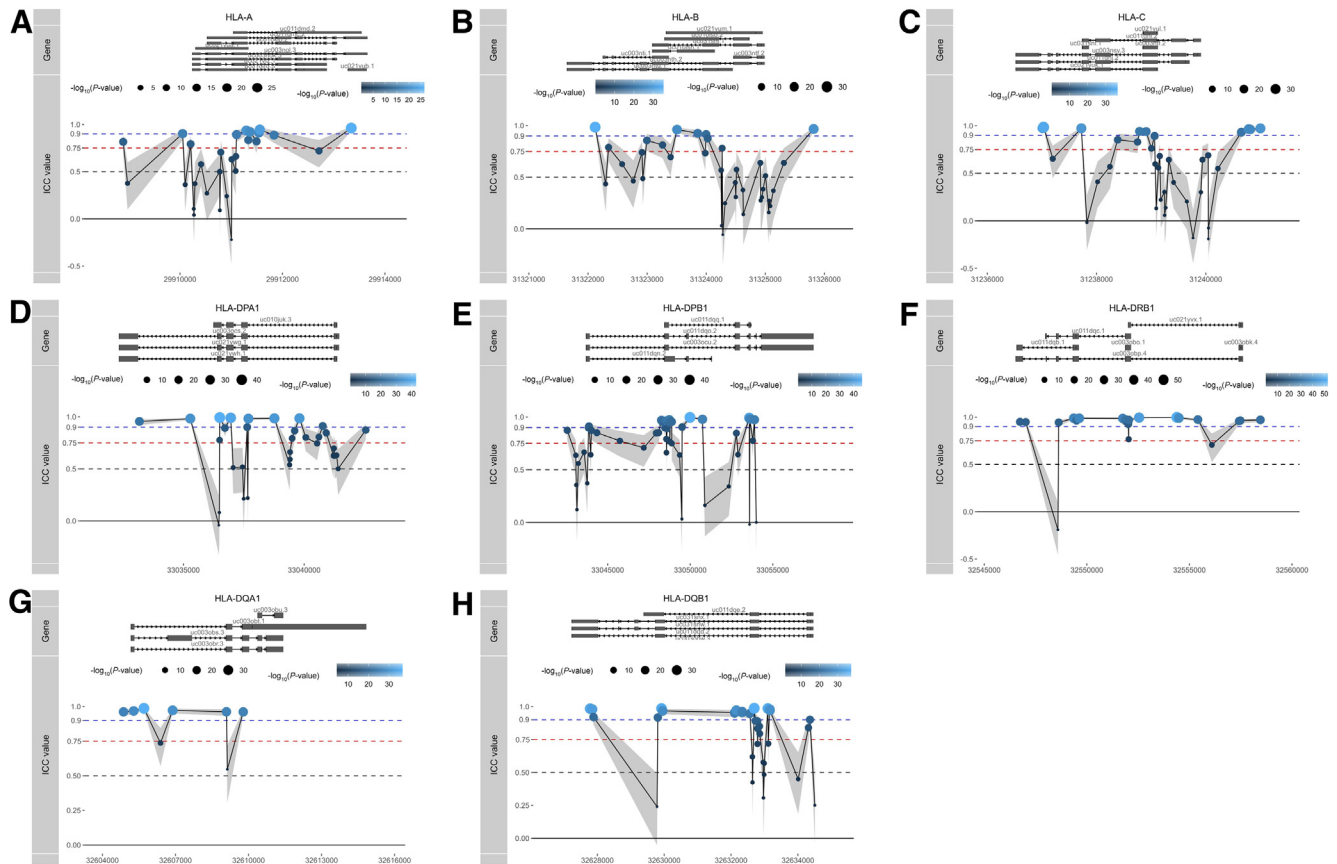


Figure 11. Visualization of the ICC values of all Illumina CpGs annotated to HLA-A (A), HLA-B (B), HLA-DQA1 (C), HLA-DQB1 (D), and HLA-DRB1 (E) relative to their position on each gene and grouped as potential GV (pink), predicted GV (green), or methylation (blue). Dots below represent known genetic variants as reported by Goyette53 for CD vs healthy controls (pink) and UC vs healthy controls (turquoise).

performed using the IDOL predictor CpGs as reference.⁷⁴ Time-associated DMPs were investigated for their association with age by performing gene set enrichment analyses using the age-associated CpGs reported by Horvath,⁴² Hanum,⁴³ Levine,⁴⁴ and Knight.⁴⁵ Visualizations were generated using *ggplot2*⁷⁵ (version 3.3.5) and *gghighlight* (version 0.3.2).

Time-independent DNA Methylation Data Analyses

For the time stability analyses, raw DNA methylation data were imported using *ewastools* to retain the 89 quality control probes that bind GVs. Methylation probes that might bind GVs were identified on the basis of the *minfi*-provided annotation files, which we termed as potential GVs. Additional GV-binding probes were estimated using the Gap-hunter tool⁵⁵ as implemented in *minfi*, which we termed the predicted GVs. Moreover, as opposed to the differential methylation analysis, for the time stability analyses, we did not perform normalization nor did we correct for any other potential confounders (eg, gender, smoking behavior, disease, and blood cell distribution) to identify truly stable signals. Stability analyses were conducted using ICC analyses, where ICC estimates and their 95% confidence

intervals were calculated using the *irr* package implemented in R. Specifically, a 2-way mixed, single measures, consistency analysis was performed.⁷⁶

Visualizations were generated using *ggplot2*⁷⁵ (version 3.3.5) and *gghighlight* (version 0.3.2).

Gene Ontology Enrichment and Kyoto Encyclopedia of Genes and Genomes Pathway Enrichment Analyses

Functional enrichment analyses genes annotated to both stable and unstable methylated probes was performed using GOMeth⁷⁷ as implemented in *missMethyl*.³³ The gene ontology terms were grouped according to biological process, cellular component, and molecular function, and an FDR corrected *P*-value below .05 indicates a statistical significant difference.

References

1. Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* 2017;390:2769–2778.

2. de Souza HSP, Fiocchi C, Iliopoulos D. The IBD interactome: an integrated view of aetiology, pathogenesis and therapy. *Nat Rev Gastroenterol Hepatol* 2017; 14:739–749.
3. de Souza HS, Fiocchi C. Immunopathogenesis of IBD: current state of the art. *Nat Rev Gastroenterol Hepatol* 2016;13:13–27.
4. Carreras-Torres R, Ibanez-Sanz G, Obon-Santacana M, et al. Identifying environmental risk factors for inflammatory bowel diseases: a Mendelian randomization study. *Sci Rep* 2020;10:19273.
5. Ventham NT, Kennedy NA, Nimmo ER, Satsangi J. Beyond gene discovery in inflammatory bowel disease: the emerging role of epigenetics. *Gastroenterology* 2013; 145:293–308.
6. Riggs AD, Russo VEA, Martienssen RA. *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor Laboratory Press; 1996.
7. Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes Dev* 2011;25:1010–1022.
8. Jung G, Hernandez-Illan E, Moreira L, et al. Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat Rev Gastroenterol Hepatol* 2020;17:111–130.
9. Samarani S, Dupont-Lucas C, Marcil V, et al. CpG methylation in TGFbeta1 and IL-6 genes as surrogate biomarkers for diagnosis of IBD in children. *Inflamm Bowel Dis* 2020;26:1572–1578.
10. Tao W, Concepcion AN, Vianen M, et al. Multiomics and machine learning accurately predict clinical response to adalimumab and etanercept therapy in patients with rheumatoid arthritis. *Arthritis Rheumatol* 2021; 73:212–222.
11. Moret-Tatay I, Cerrillo E, Saez-Gonzalez E, et al. Identification of epigenetic methylation signatures with clinical value in Crohn's disease. *Clin Transl Gastroenterol* 2019; 10:e00083.
12. Howell KJ, Kraiczy J, Nayak KM, et al. DNA methylation and transcription patterns in intestinal epithelial cells from pediatric patients with inflammatory bowel diseases differentiate disease subtypes and associate with outcome. *Gastroenterology* 2018;154:585–598.
13. Kalla R, Adams AT, Vatn S, et al. Epigenetic alterations at diagnosis predict susceptibility, prognosis and treatment escalation in inflammatory bowel disease-IBD character. *Gut* 2017;66:A24–A25.
14. Li Yim AYW, Duijvis NW, Zhao J, et al. Peripheral blood methylation profiling of female Crohn's disease patients. *Clin Epigenetics* 2016;8:65.
15. Gasparetto M, Payne F, Nayak K, et al. Transcription and DNA methylation patterns of blood-derived CD8(+) T cells are associated with age and inflammatory bowel disease but do not predict prognosis. *Gastroenterology* 2021;160:232–244.e7.
16. McDermott E, Ryan EJ, Tosetto M, et al. DNA methylation profiling in inflammatory bowel disease provides new insights into disease pathogenesis. *J Crohns Colitis* 2016;10:77–86.
17. Ventham NT, Kennedy NA, Adams AT, et al. IBD BIOM consortium; IBD CHARACTER consortium. Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat Comm* 2016;7:13507.
18. Adams AT, Kennedy NA, Hansen R, et al. Two-stage genome-wide methylation profiling in childhood-onset Crohn's disease implicates epigenetic alterations at the VMP1/MIR21 and HLA Loci. *Inflamm Bowel Dis* 2014; 20:1784–1793.
19. Harris RA, Nagy-Szakal D, Pedersen N, et al. Genome-wide peripheral blood leukocyte DNA methylation microarrays identified a single association with inflammatory bowel diseases. *Inflamm Bowel Dis* 2012; 18:2334–2341.
20. Hasler R, Feng Z, Backdahl L, et al. A functional methyome map of ulcerative colitis. *Genome Res* 2012; 22:2130–2137.
21. Sominen HK, Venkateswaran S, Kilaru V, et al. Blood-derived DNA methylation signatures of Crohn disease and severity of intestinal inflammation. *Gastroenterology* 2019;156:2254–2265.e3.
22. Perez RF, Santamarina P, Tejedor JR, et al. Longitudinal genome-wide DNA methylation analysis uncovers persistent early-life DNA methylation changes. *J Transl Med* 2019;17:15.
23. Nagy C, Turecki G. Sensitive periods in epigenetics: bringing us closer to complex behavioral phenotypes. *Epigenomics* 2012;4:445–457.
24. Planterose Jimenez B, Liu F, Caliebe A, et al. Equivalent DNA methylation variation between monozygotic co-twins and unrelated individuals reveals universal epigenetic inter-individual dissimilarity. *Genome Biol* 2021; 22:18.
25. Fraga MF, Ballestar E, Paz MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 2005;102:10604–10609.
26. Salameh Y, Bejaoui Y, El Hajj N. DNA methylation biomarkers in aging and age-related diseases. *Front Genet* 2020;11:171.
27. Dobbs KR, Embury P, Koech E, et al. Age-related differences in monocyte DNA methylation and immune function in healthy Kenyan adults and children. *Immune Ageing* 2021;18:11.
28. Gentilini D, Mari D, Castaldi D, et al. Role of epigenetics in human aging and longevity: genome-wide DNA methylation profile in centenarians and centenarians' offspring. *Age (Dordr)* 2013;35:1961–1973.
29. Fraga MF, Agrelo R, Esteller M. Cross-talk between aging and cancer: the epigenetic language. *Ann N Y Acad Sci* 2007;1100:60–74.
30. Johnson AA, Akman K, Calimport SR, et al. The role of DNA methylation in aging, rejuvenation, and age-related disease. *Rejuvenation Res* 2012;15:483–494.
31. Houseman EA, Kelsey KT, Wiencke JK, et al. Cell-composition effects in the analysis of DNA methylation array data: a mathematical perspective. *BMC Bioinformatics* 2015;16:95.
32. Bartlett AH, Liang JW, Sandoval-Sierra JV, et al. Longitudinal study of leukocyte DNA methylation and biomarkers for cancer risk in older adults. *Biomark Res* 2019;7:10.

33. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics* 2016; 32:286–288.
34. Reynolds LM, Taylor JR, Ding J, et al. Age-related variations in the methylome associated with gene expression in human monocytes and T cells. *Nat Commun* 2014;5:5366.
35. Hears AC, Martin GE, Angelovich TA, et al. Aging is associated with chronic innate immune activation and dysregulation of monocyte phenotype and function. *Aging Cell* 2012;11:867–875.
36. Flanagan JM, Brook MN, Orr N, et al. Temporal stability and determinants of white blood cell DNA methylation in the breakthrough generations study. *Cancer Epidemiol Biomarkers Prev* 2015;24:221–229.
37. Zaimi I, Pei D, Koestler DC, et al. Variation in DNA methylation of human blood over a 1-year period using the Illumina MethylationEPIC array. *Epigenetics* 2018;13:1056–1071.
38. Forest M, O'Donnell KJ, Voisin G, et al. Agreement in DNA methylation levels from the Illumina 450K array across batches, tissues, and time. *Epigenetics* 2018;13:19–32.
39. Shvetsov YB, Song MA, Cai Q, et al. Intraindividual variation and short-term temporal trend in DNA methylation of human blood. *Cancer Epidemiol Biomarkers Prev* 2015;24:490–497.
40. Coit P, Ortiz-Fernandez L, Lewis EE, et al. A longitudinal and transancestral analysis of DNA methylation patterns and disease activity in lupus patients. *JCI Insight* 2020;5: e143654.
41. Leite ML, de Oliveira KBS, Cunha VA, et al. Epigenetic therapies in the precision medicine era. *Advanced Therapeutics* 2020;3:1900184.
42. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol* 2013;14:R115.
43. Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell* 2013;49:359–367.
44. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* 2018;10:573–591.
45. Knight AK, Craig JM, Theda C, et al. An epigenetic clock for gestational age at birth based on blood methylation data. *Genome Biol* 2016;17:206.
46. Daca-Roszak P, Pfeifer A, Zebracka-Gala J, et al. Impact of SNPs on methylation readouts by Illumina Infinium HumanMethylation450 BeadChip Array: implications for comparative population studies. *BMC Genomics* 2015; 16:1003.
47. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–163.
48. Joustra V, Hageman IL, Satsangi J, et al. Systematic review and meta-analysis of peripheral blood DNA methylation studies in inflammatory bowel disease. *J Crohns Colitis* 2022.
49. Momozawa Y, Dmitrieva J, Theatre E, et al. IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. *Nat Commun* 2018; 9:2427.
50. Yamazaki K, McGovern D, Ragoussis J, et al. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum Mol Genet* 2005; 14:3499–3506.
51. Ashton JJ, Latham K, Beattie RM, et al. Review article: the genetics of the human leucocyte antigen region in inflammatory bowel disease. *Aliment Pharmacol Ther* 2019;50:885–900.
52. Goyette P, Boucher G, Mallon D, et al. International Inflammatory Bowel Disease Genetics Consortium; Australia and New Zealand IBDGC; Belgium IBD Genetics Consortium; Italian Group for IBD Genetic Consortium; NIDDK Inflammatory Bowel Disease Genetics Consortium; United Kingdom IBDGC; Wellcome Trust Case Control Consortium; Quebec IBD Genetics Consortium. High-density mapping of the MHC identifies a shared role for HLA-DRB1*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. *Nat Genet* 2015;47:172–179.
53. Verstockt B, Smith KG, Lee JC. Genome-wide association studies in Crohn's disease: past, present and future. *Clin Transl Immunology* 2018;7:e1001.
54. Sazonovs A, Kennedy NA, Moutsianas L, et al. PANTS Consortium. HLA-DQA1*05 carriage associated with development of anti-drug antibodies to infliximab and adalimumab in patients with Crohn's disease. *Gastroenterology* 2020;158:189–199.
55. Andrews SV, Ladd-Acosta C, Feinberg AP, et al. Gap hunting" to characterize clustered probe signals in Illumina methylation array data. *Epigenetics & Chromatin* 2016;9:56.
56. Villicana S, Bell JT. Genetic impacts on DNA methylation: research findings and future perspectives. *Genome Biology* 2021;22:127.
57. Sugden K, Hannon EJ, Arseneault L, et al. Patterns of reliability: assessing the reproducibility and integrity of DNA methylation measurement. *Patterns (N Y)* 2020;1: 100014.
58. Bogaert S, Laukens D, Peeters H, et al. Differential mucosal expression of Th17-related genes between the inflamed colon and ileum of patients with inflammatory bowel disease. *BMC Immunol* 2010;11:61.
59. Sipos F, Galamb O, Wichmann B, et al. Peripheral blood based discrimination of ulcerative colitis and Crohn's disease from non-IBD colitis by genome-wide gene expression profiling. *Dis Markers* 2011;30:1–17.
60. Cooke J, Zhang H, Greger L, et al. Mucosal genome-wide methylation changes in inflammatory bowel disease. *Inflamm Bowel Dis* 2012;18:2128–2137.
61. Nimmo ER, Prendergast JG, Aldhous MC, et al. Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. *Inflamm Bowel Dis* 2012;18:889–899.
62. Agliata I, Fernandez-Jimenez N, Goldsmith C, et al. The DNA methylome of inflammatory bowel disease (IBD) reflects intrinsic and extrinsic factors in intestinal mucosal cells. *Epigenetics* 2020;15:1068–1082.
63. Ahmad T, Marshall SE, Jewell D. Genetics of inflammatory bowel disease: the role of the HLA complex. *World J Gastroenterol* 2006;12:3628–3635.

64. van Vugt LJ, van den Reek J, Hannink G, et al. Association of HLA-C*06:02 status with differential response to ustekinumab in patients with psoriasis: a systematic review and meta-analysis. *JAMA Dermatol* 2019;155:708–715.
65. Dand N, Duckworth M, Baudry D, et al, BADBIR Study Group; BSTOP Study Group; PSORT Consortium. HLA-C*06:02 genotype is a predictive biomarker of biologic treatment response in psoriasis. *J Allergy Clin Immunol* 2019;143:2120–2130.
66. Hirose W, Harigai M, Amano K, et al, TOF-ABT Study Group. Impact of the HLA-DRB1 shared epitope on responses to treatment with tofacitinib or abatacept in patients with rheumatoid arthritis. *Arthritis Res Ther* 2021;23:228.
67. Kindt ASD, Fuerst RW, Knoop J, et al. Allele-specific methylation of type 1 diabetes susceptibility genes. *J Autoimmun* 2018;89:63–74.
68. Gionchetti P, Dignass A, Danese S, et al. European evidence-based consensus on the diagnosis and management of Crohn's disease 2016: part 2: surgical management and special situations. *J Crohns Colitis* 2017;11:135–149.
69. Lamb CA, Kennedy NA, Raine T, et al. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. *Gut* 2019;68(Suppl 3):s1–s106.
70. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 2014;30:1363–1369.
71. Fortin JP, Labbe A, Lemire M, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol* 2014;15:503.
72. 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.
73. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3:Article3.
74. Salas LA, Koestler DC, Butler RA, et al. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina Human-MethylationEPIC BeadArray. *Genome Biol* 2018;19:64.
75. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag; 2016.
76. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–428.
77. Maksimovic J, Oshlack A, Phipson B. Gene set enrichment analysis for genome-wide DNA methylation data. *Genome Biol* 2021;22:173.

Received August 31, 2022. Accepted December 20, 2022.

Correspondence

Address correspondence to: Prof Dr Geert D'Haens, Amsterdam University Medical Centers, Department of Gastroenterology and Hepatology, Room C2-208, Meibergdreef 9, 1105 AZ, Amsterdam, the Netherlands. e-mail: g.dhaens@amsterdamumc.nl.

CRedit Authorship Contributions

Vincent Wilhelmus Joustra, MD (Conceptualization: Equal; Data curation: Supporting; Formal analysis: Supporting; Methodology: Lead; Project administration: Lead; Visualization: Supporting; Writing – original draft: Lead)

Andrew Y.F. Li Yim, PhD (Conceptualization: Equal; Formal analysis: Lead; Methodology: Equal; Supervision: Supporting; Visualization: Lead; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Ishtu Hageman, MD (Data curation: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)

Evgeni Levin, PhD (Conceptualization: Supporting; Formal analysis: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)

Alex Adams, MBChB, PhD (Conceptualization: Supporting; Formal analysis: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)

Jack Satsangi, DPhil, FRCP (Conceptualization: Supporting; Methodology: Supporting; Supervision: Equal; Writing – review & editing: Supporting)

Wouter J. de Jonge, PhD (Conceptualization: Supporting; Data curation: Supporting; Funding acquisition: Equal; Methodology: Supporting; Resources: Equal; Supervision: Supporting; Writing – review & editing: Supporting)

Peter Henneman, PhD (Conceptualization: Supporting; Formal analysis: Supporting; Methodology: Supporting; Supervision: Supporting; Writing – review & editing: Supporting)

Geert D'Haens, MD, PhD (Conceptualization: Equal; Data curation: Supporting; Funding acquisition: Lead; Methodology: Supporting; Supervision: Equal; Writing – review & editing: Supporting)

Conflicts of interest

The authors disclose no conflicts.

Funding

This research was partially supported by the Helmsley Charitable Trust.

Data Availability

The raw data has been made available under controlled access in the European Genome-phenome Archive (accession ID: EGAS00001006501). All authors had access to the study data and have reviewed and approved the final manuscript.