Long-term temporal stability of peripheral blood DNA methylation alterations in patients with inflammatory bowel disease.

Vincent Joustra MD[1], Andrew Y.F. Li Yim PhD[2,3], Ishtu Hageman MD[3], Evgeni Levin PhD[4,5], Alex Adams MBChB PhD[6], Jack Satsangi DPhil FRCP[6], Wouter J. de Jonge PhD[3], Peter Henneman PhD[2], Geert D'Haens MD PhD[1].

[1] Department of Gastroenterology and Hepatology, Amsterdam Gastroenterology Endocrinology Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands

[2] Genome Diagnostics Laboratory, Department of Human Genetics, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands.

[3] Tytgat Institute for Liver and Intestinal Research, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands.

[4] Department of Vascular Medicine, Academic Medical Center, University of Amsterdam, 1105 AZ Amsterdam, The Netherlands.

[5] Horaizon BV, 2625 GZ Delft, The Netherlands.

[6] Oxford University- Hospitals NHS Foundation Trust- John Radcliffe Hospital, Translational Gastroenterology Unit- NIHR Oxford Biomedical Research Centre, Oxford, United Kingdom.

Address for correspondence: Prof. Dr. Geert D'Haens, Amsterdam University Medical Centers,

Department of Gastroenterology and Hepatology, Room C2-208, Meibergdreef 9, 1105 AZ, Amsterdam,

The Netherlands. Email: g.dhaens@amsterdamumc.nl

Funding: This research was partially supported by the Helmsley Charitable Trust.

Conflicts of interest: VJ, ALY, IH, EL, AA, JS, PH, WJ and GD: none to declare.

Authors' contributions:

Concept design: VW, ALY, PH, GD; Methodology: VW, ALY, EL; Formal analysis: ALY, VW; Data Acquisition: VW, ALY; Writing - original draft: VW, ALY; Writing - review and editing: VW, IH, ALY, EL, AA, JS, WJ, PH, GD; Approval of final manuscript: all authors.

Abstract

Introduction

There is great current interest in the potential application of DNA methylation alterations in peripheral blood leukocytes (PBL) 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 characterised. 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 IBD patients (36 Crohn's disease (CD), 10 ulcerative colitis (UC), median age 44 (IQR: 27-56), 50% female) that received standard care without any intervention at the Amsterdam UMC. Paired PBL samples were collected at two time points with a median 7 (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%) excellent (ICC ≥0.90) stability. 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.

Conclusion

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.

Introduction

Crohn's disease (CD) and ulcerative colitis (UC) are chronic relapsing and remitting inflammatory bowel diseases (IBD) characterized by a wide variety of phenotypic manifestations[1]. While the aetiology 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, regulate gene expression thereby altering cellular behavior[7]. Over the past decade, multiple epigenome-wide association studies (EWAS) have sought to characterise, 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 (PBL) and/or mucosal tissue[9, 14-22], with only a single longitudinal study in mucosal tissue[20] and PBL[23].

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[24, 25]. Nonetheless, the influence of ageing on genome-wide DNA methylation has been well-described in monozygotic twins[26, 27] and unrelated healthy populations[28, 29], 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[30-32]. As epigenetics and thus DNA methylation, is cell-type specific, observed differences found in heterogeneous populations such as PBL or tissue might reflect differences in the cellular composition[33-35]. Nonetheless, age-related differences were found in more homogeneous populations, such as purified T-cells and monocytes[15, 36, 37]

Despite the strong effects of age on DNA methylation, a high correlation between baseline and follow-up methylation data in IBD pediatric mucosal tissue has been observed[20]. 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 normalisation of CRP[23]. 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 HumanMethylation EPIC BeadChip array. While temporal stability and intra-individual variability in PBL-derived DNA methylation has been investigated in adult healthy individuals[38-41] and patients with systemic lupus erythematosus (SLE)[42], no such study has been conducted in IBD patients.

There is widespread interest in the application of epigenetic markers in personalization of treatment[43]. 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 PBL obtained from a well-generalizable cohort of adult IBD patients with a 7-year median time between collected DNA samples.

Methods

Patient selection

We performed a single-centre, longitudinal EWAS, where we collected PBL samples from adult IBD patients at the Amsterdam University Medical Centers. The interval between the time of sampling ranged from 2 to 9 years with a median of 7 (Figure 1a). 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[44, 45]. In addition, all patients received standard care follow-up. No additional in- or exclusion criteria were used as the goal was to collect a cohort of IBD patients that reflected the overall IBD population at the Amsterdam University Medical Centers. 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 6mL EDTA tube and stored at –80 °C until further processing. Genomic DNA was isolated using the QIAsymphony 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 analysed on the Illumina HumanMethylation 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 summarised using descriptive statistics. Categorical variables are presented as percentages and continuous variables as median annotated with the interquartile range (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*[46] package (version 1.36), whereupon the raw signal intensities were normalized using functional normalization[47] and converted into methylation ratios. Differential methylation analyses was performed using *limma*[48] (version 3.46) and *eBayes*[49] regressing against time point (T2 vs T1), gender, smoking behaviour, disease, and blood cell distribution. Statistical significance was defined as a false discovery rate-adjusted *p*-value < 0.05. 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 performed using the IDOL predictor CpGs as reference[50]. Time-associated DMPs were investigated for their association with age by performing gene set enrichment analyses using the age-associated CpGs reported by Horvath[51], Hannum[52], Levine[53], and Knight[54]. Visualisations were generated using *ggplot2*[55] (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 genetic variants (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 Gaphunter tool[56] 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 normalisation nor corrected for any other potential confounders (e.g. gender, smoking behaviour, disease, and blood cell distribution) in order to identify truly stable signals. Stability analyses were conducted using intraclass correlation (ICC) analyses, where ICC estimates and their 95% confidence intervals were calculated using the *irr* package implemented in R. Specifically, a two-way mixed, single measures, consistency analysis was performed[57]. Visualisations were generated using *ggplot2*[55] (version 3.3.5) and *gghighlight* (version 0.3.2).

GO enrichment and KEGG pathway enrichment analyses

Functional enrichment analyses genes annotated to both stable and unstable methylated probes was performed using GOmeth[58] as implemented in *missMethyl*[35]. The GO terms were grouped according to biological process (BP), cellular component (CC) and molecular function (MF) and an FDR corrected p-value below 0.05 indicates a statistical significant difference.

Data availability

The raw has been made available under controlled access in the European Genome-phenome Archive (accession ID: EGAS00001006501).

Results

Patient demographics

A total of 46 adult IBD (36 CD, 10 UC) patients with a median age of 44 (IQR: 27-56) years and median disease duration of 12 (IQR: 7-21) years were included. Gender, surgical history, disease location- and behaviour were balanced within this cohort (Table 1). Notably, 32 patients (69.6%) were previously treated with an anti-TNF, prior to T1 sampling. Between T1 and T2 during regular IBD care follow-up, 10 (21.7%) patients underwent IBD-related surgery, 24 (52.2%) were treated with vedolizumab and 14 (30.4%) with ustekinumab, reflecting the tertiary referral population seen at the Amsterdam UMC. No significant differences in median CRP (*p*-value = 0.97) or leukocyte count (*p*-value = 0.85) between T1 and T2 were observed (Figure 1b and 1c).

Time-associated differential methylation expectedly associates with age-related CpGs

We first investigated the differences in methylation between both time points, identifying 194,391 (≈23%) differentially methylated positions (DMPs) when comparing T1 and T2 at a FDR-adjusted *p*-value of <0.05 (Figure 2a, b and 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 no significant differences (Figure 2d).

By contrast, the time-associated differences were, expectedly, enriched for age-associated CpGs, which have been defined as the "epigenetic clocks" from Horvath[51], Hannum[52], Levine[53], and Knight[54] (Figure 2e). Furthermore, for these specific epigenetic clock CpGs, we observed a general hypomethylated pattern at T2 relative to T1 CpG sites (Figure 2f), 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 (Supplementary Figure 1).

Time-invariant, stable methylated probes are enriched in genes involved in cell adhesion

In order 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 (≈14%) displayed a statistically significant high ICC (≥0.75), which we termed stably methylated positions (SMPs). We reasoned that probes that were associated with sites known to harbour genetic variants, both intentional and unintentional[59], 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 for quality control purposes, presented high (>0.9) ICC values (Figure 3a 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 3b). Functional enrichment analyses of the SMPs and HSMPs indicated significant enrichment of genes involved in cell-cell signaling, adhesion and neurogenesis (Supplementary Figures 2 and 3).

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[60], which included a total of 552 samples (177 CD, 132UC and 243HC) 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 4a, supplementary table 1). Nonetheless, 22 CD-associated (12.4%), 11 UC-associated (8.3%) and 24 IBD-associated loci (9.9%) show good to excellent stability over time, providing evidence that these CD-, UC-, or IBD-associated DMPs are unaffected by ageing or the exposures over time (Figure 4a, 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, 23] as well as shown found to be among the most significant IBD-associated DMPs in our meta-analysis[60]. We observed moderate consistency over time with noticeable overall hypermethylation at T2 relative to T1 for the aforementioned 3 CpGs (Figure 4b).

In addition to our meta-analysis, we also interrogated the CpGs that were CD-associated but CRP-independent as reported on by Somineni *et al*[23]. 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 0.5-0.74), and cg09171692 (*RORC*) and cg02055816 (*SMARCD3*) presented good consistency (ICC 0.75-0.89) between both time points (Figure 4c).

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*[61, 62], as well as specific IBD-associated major histocompatibility complex (MHC) 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*[61, 63-66]. 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. 5a). 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 1st exons (Supplementary figure 4 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[59], 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 hybridise with genetic variants using Gaphunter[67]. Notably, most of the high ICC values were found for CpGs that presented some type of clustering typical of genetic variants, but were not necessarily catalogued in dbSNP (v151) (Figure. 5b). 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 to *HLA* class I genes (*HLA-A*, *HLA-B,* and *HLA-C*) (Figure 5b). Besides technical artefacts, DNA methylation itself can be affected by genetic variants that occur in the vicinity[68]. As such, we cross-referenced our observations with a previous large-scale genotyping study of the *HLA* region in both CD and UC patients[64]. We indeed found multiple probes within the vicinity (<1Kb) 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 neighbouring IBD-associated GVs.

Notably, several highly stable probes (ICC ≥0.75) found to be annotated to *HLA-C*, *HLA-DPB1*, and *HLA-DPA1*, were located far away (>1Kb) from any of the IBD-associated GVs, did not associate with catalogued GVs, nor were identified as a potential GV by Gaphunter (Figure 5c, d and e, supplementary figure 5 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 IBD (36 CD and 10 UC) patients that were collected at two time points separated by a median of 7 years reflective of a real-life tertiary referral population.

Our observations indicate that the majority of all loci (≈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. While we observed no significant differences in CRP, leukocyte count, and estimated cellular composition 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 (≈14%) and 41.274 (≈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 external exposures, suggesting time-stability.

Previous studies investigating DNA methylation stability in PBL 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 – 23% of the CpG loci presented a moderate/good (ICC 0.50 - 0.79) while 8.3-12.9% of the CpG loci presented a good/excellent (ICC ≥0.8) stability over a span of 1 to 6 years[39, 41]. Focusing on IBD, when interrogating 255 CD-associated, 103 UC-associated and 221 IBD-associated DMPs identified in our own meta-analysis[60] 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. While 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[69] 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 to T1. Similar analyses performed on the CD-associated yet CRP-independent probes reported by Somineni et al.[23] showed that the majority of these CpGs did not present long-term stability in our cohort. Differences between our observations and that of Somineni 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[23, 70, 71] and therefore show promise as stable IBD-associated loci.

Given the complex, multifactorial nature of IBD[2], focussing on the interplay between genetic variation and DNA methylation rather than single gene mutations alone might prove more useful in understanding its molecular aetiology. Previous EWAS of mucosal tissue[72] and peripheral blood[73] 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, 61, 63-66, 74, 75]. 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[63] with particular impacts for the response to biological treatment[76-78] and the formation of anti-drug antibodies[66]. However, translation of the results into clinical practice has proven to be difficult due to the high number of polymorphisms of HLA $\alpha\beta$ heterodimers and strong linkage disequilibrium[63].

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 to toddlers (r=0.83) and adults (r=0.88) with type 1 diabetes[79]. While 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, stabilising effect on the epigenome.

Strengths and limitations

To our knowledge, we are the first to assess the stability of the DNA methylome obtained from PBL of IBD patients 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[39, 42, 69]. While 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) and cellular distribution were hardly different between both time points.

Conclusion

We observe considerable variability in DNA methylation measurements taken from PBL at two 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 two points. Focussing on these CpG loci during biomarker discovery might result in the identification of biologically relevant and more robust IBD-associated epigenetic biomarkers with an increased probability of independent replication.

Literature

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, et al. 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, Plainview, N.Y. 1996.
7. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev 2011;25:1010-22.
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 2020.
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 AYF, 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 2020.
16. McDermott E, Ryan EJ, Tosetto M, et al. DNA Methylation Profiling in Inflammatory Bowel Disease Provides New Insights into Disease Pathogenesis. Journal of Crohn's and Colitis 2016;10:77--86.
17. Ventham NT, Kennedy NA, Adams AT, et al. Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. Nature Communications 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. Inflammatory Bowel Diseases 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. Inflammatory Bowel Diseases 2012;18:2334--2341.
20. 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.

21. Hasler R, Feng Z, Backdahl L, et al. A functional methylome map of ulcerative colitis. Genome Research 2012;22:2130--2137.
22. Moret-Tatay I, Cerrillo E, Sáez-González E, et al. Identification of Epigenetic Methylation Signatures With Clinical Value in Crohn's Disease. Clinical and translational gastroenterology 2019:10.14309/ctg.0000000000000083.
23. Somineni HK, Venkateswaran S, Kilaru V, et al. Blood-Derived DNA Methylation Signatures of Crohn Disease and Severity of Intestinal Inflammation. Gastroenterology 2019.
24. 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.
25. Nagy C, Turecki G. Sensitive periods in epigenetics: bringing us closer to complex behavioral phenotypes. Epigenomics 2012;4:445-57.
26. 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.
27. 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;102:10604-9.
28. Salameh Y, Bejaoui Y, El Hajj N. DNA Methylation Biomarkers in Aging and Age-Related Diseases. Front Genet 2020;11:171.
29. Dobbs KR, Embury P, Koech E, et al. Age-related differences in monocyte DNA methylation and immune function in healthy Kenyan adults and children. Immun Ageing 2021;18:11.
30. 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;35:1961-73.
31. Fraga MF, Agrelo R, Esteller M. Cross-talk between aging and cancer: the epigenetic language. Ann N Y Acad Sci 2007;1100:60-74.
32. Johnson AA, Akman K, Calimport SR, et al. The role of DNA methylation in aging, rejuvenation, and age-related disease. Rejuvenation Res 2012;15:483-94.
33. 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.
34. 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.
35. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics 2016;32:286-8.
36. 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.
37. Hearps 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;11:867-75.
38. 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;24:221-9.
39. 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.
40. 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.
41. 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;24:490-7.
42. 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.
43. Leite ML, de Oliveira KBS, Cunha VA, et al. Epigenetic Therapies in the Precision Medicine Era. Advanced Therapeutics 2020;3.

44. Gionchetti P, Dignass A, Danese S, et al. 3rd European Evidence-based Consensus on the Diagnosis and Management of Crohn's Disease 2016: Part 2: Surgical Management and Special Situations. Journal of Crohns & Colitis 2017;11:135-149.

45. 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:s1-s106.

46. 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;30:1363-9.

47. 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.

48. 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.

49. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004;3:Article3.

50. Salas LA, Koestler DC, Butler RA, et al. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. Genome Biol 2018;19:64.

51. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14:R115.

52. 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.

53. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY) 2018;10:573-591.

54. 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.

55. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York, 2016.

56. 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.

57. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420-8.

58. Maksimovic J, Oshlack A, Phipson B. Gene set enrichment analysis for genome-wide DNA methylation data. Genome Biol 2021;22:173.

59. 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.

60. Joustra V. HI. Systematic Review and Meta-analysis of Peripheral Blood DNA methylation studies in Inflammatory Bowel Disease. Journal of Crohn's & Colitis (in press) 2022.

61. 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.

62. Yamazaki K, McGovern D, Ragoussis J, et al. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. Human Molecular Genetics 2005;14:3499-3506.

63. 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.

64. Goyette P, Boucher G, Mallon D, et al. 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;47:172-9.

65. Verstockt B, Smith KG, Lee JC. Genome-wide association studies in Crohn's disease: Past, present and future. Clin Transl Immunology 2018;7:e1001.

66. Sazonovs A, Kennedy NA, Moutsianas L, et al. 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.

67. 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.

68. Villicana S, Bell JT. Genetic impacts on DNA methylation: research findings and future perspectives. Genome Biology 2021;22.
69. 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.
70. 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.
71. 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.
72. Cooke J, Zhang H, Greger L, et al. Mucosal genome-wide methylation changes in inflammatory bowel disease. Inflammatory Bowel Diseases 2012;18:2128--2137.
73. 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. Inflammatory Bowel Diseases 2012;18:889--899.
74. 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.
75. Ahmad T, Marshall SE, Jewell D. Genetics of inflammatory bowel disease: the role of the HLA complex. World J Gastroenterol 2006;12:3628-35.
76. 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.
77. Dand N, Duckworth M, Baudry D, et al. HLA-C*06:02 genotype is a predictive biomarker of biologic treatment response in psoriasis. J Allergy Clin Immunol 2019;143:2120-2130.
78. Hirose W, Harigai M, Amano K, et al. 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.
79. Kindt ASD, Fuerst RW, Knoop J, et al. Allele-specific methylation of type 1 diabetes susceptibility genes. J Autoimmun 2018;89:63-74.
80. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 2016;15:155-63.

Tables

Table 1. Baseline characteristics	Crohn's disease (n=36)	Ulcerative colitis (n=10)	Total (n=46)
Female, n (%)	18 (50)	5 (50)	23 (50)
Age, years, median (IQR)	42 (27-53)	52 (36-61)	44 (27-56)
Disease duration, years, median (IQR)	11 (6-20)	13 (7-22)	12 (7-21)
Ethnic background, n (%)			
- Caucasian	30 (83.3)	9 (90)	39 (84.8)
Disease location CD, n (%)			
- 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, n (%)			
- Left-sided UC, distal to splenic flexure (E2)		3 (30)	3 (30)
- Extensive, proximal to splenic flexure (E3)		7 (70)	7 (70)
Disease behaviour, n (%)			
- 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, n (%)	19 (52.8)	1 (10)	20 (43.5)
IBD-related surgery between T1 and T2, n (%)	10 (27.8)	-	10 (21.7)
Previous medical treatment, n (%)			
- 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, n (%)			
- 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, median (IQR)	3.2 (1.4-8.0)	2 (0.4-5.8)	2.6 (1.2-6.6)
CRP T2, mg/L, median (IQR)	3 (1.6-5.3)	2.5 (0.8-7.7)	3.0 (1.4-5.3)
Leukocyte count T1, *10^9, median (IQR)	7.1 (6.3-9.1)	5.4 (4.7-6.3)	6.7 (5.4-8.3)
Leukocyte count T2, *10^9, median (IQR)	7.1 (5.7-8.9)	5.8 (5.2-7.6)	7.1 (5.3-8.5)
Smoking, n (%)			
- Active	8 (22.2)	3 (30)	11 (23.9)
- Non-smoker	28 (77.8)	7 (70)	35 (76.1)

Table 2 ICC scores

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

An overview of the ICC values classified using the system presented by Koo and Li[80]. 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.

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	0.917048378	0.960490142	LINC01307 (body)
cg06231783	chr11:1475048	0.990318642	1.53E-40	3.24E-38	-0.002016075	0.913436831	0.958698404	BRSK2 (body)
cg22113540	chr16:10125521	0.990357545	1.40E-40	2.98E-38	-0.005125811	0.726564998	0.852281861	GRIN2A (body)
cg04546413	chr19:29218101	0.992092045	1.64E-42	4.48E-40	-0.008758714	0.737190388	0.859028126	
cg05896524	chr21:47604654	0.991011374	2.90E-41	6.71E-39	-0.00862793	0.708898078	0.841027778	C21orf56 (TSS1500)
cg20540428	chr3:73045686	0.990348581	1.43E-40	3.03E-38	-0.005080004	0.778207203	0.884339751	PPP4R2 (TSS1500)
cg22243260	chr3:126946036	0.991961418	2.37E-42	6.31E-40	-0.000283585	0.989690923	0.995366819	
cg16885113	chr6:29648507	0.990173675	2.14E-40	4.44E-38	0.009565754	0.695688826	0.832438074	
cg18280909	chr6:29723301	0.990760129	5.38E-41	1.21E-38	0.009797484	0.631735087	0.788308313	
cg09271603	chrX:24482885	0.990091759	2.57E-40	5.29E-38	0.002551285	0.901311297	0.952513118	PDK3 (TSS1500)
cg08222221	chrX:139589617	0.990360759	1.39E-40	2.95E-38	-0.005243195	0.737019022	0.858927936	

CGID = Illumina CpG identifier. Coordinate = Genomic coordinate of the CpG on the human genome (build hg19). ICC = Intraclass correclation coefficient. 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. DM = Mean difference in percentage methylation. 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. Gene = Associated gene as well as the location in the gene.

Figure Legends

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

Figure 2. Time-variant methylated positions. A) A volcano plot depicting the mean difference in methylation between the two time points on the x-axis and the $-\log_{10}(p\text{-value})$ on the y-axis. B) Heatmap visualising 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-value ≤0.05). 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. Time-invariant methylated positions. A) The intraclass correlation coefficients (ICC) stratified by probe type, with QC GV representing the aforementioned quality control probes that bind genetic variants (GV), the potential GV representing probes that were annotated with a genetic variant, and predicted GV representing probes that presented a methylation signal typically found when driven by a genetic variant. Red dashed lines represent the classification boundaries introduced by Koo and Li[80], 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 visualisation of the 11 probes that present a time-invariant 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 visualisation represents the mean and standard error of the mean.

Figure 4. IBD-associated SMPs. A) The intraclass correlation coefficients (ICC) of consistent DMPs identified through meta-analyses of 4 EWAS[14, 17-19] by Joustra et al.[60] 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[80], 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 al.[18] , Ventham et al.[17], as well as the CRP-independent probes reported on by C) Somineni et al.[23] 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.

Figure 5. IBD risk genes. A) Visualisation of the intraclass correlation coefficients (ICC) for all CpGs annotated to IBD-associated GWAS genes. Boxplots show the overall median stability within each gene. Red dashed lines represent the classification boundaries introduced by Koo and Li[80], 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) 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 genetic variant, and methylation representing probes for which we have no evidence that they hybridise with any geneticm variant . Visualisations 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 (coral), predicted GV (green), or methylation (blue). Dots below represent known genetic variants as reported by Goyette[64] for CD versus healthy controls (coral) and UC versus healthy controls (turquoise).

Supplementary Figure 1. Functional enrichments analyses using GO-term and KEGG pathways for the time-associated DMPs

Supplementary Figure 2. Functional enrichments analyses using GO-term and KEGG pathways for the stable methylated positions (SMPs)

Supplementary Figure 3. Functional enrichments analyses using GO-term and KEGG pathways for the hyper-stable methylated positions (HSMPs).

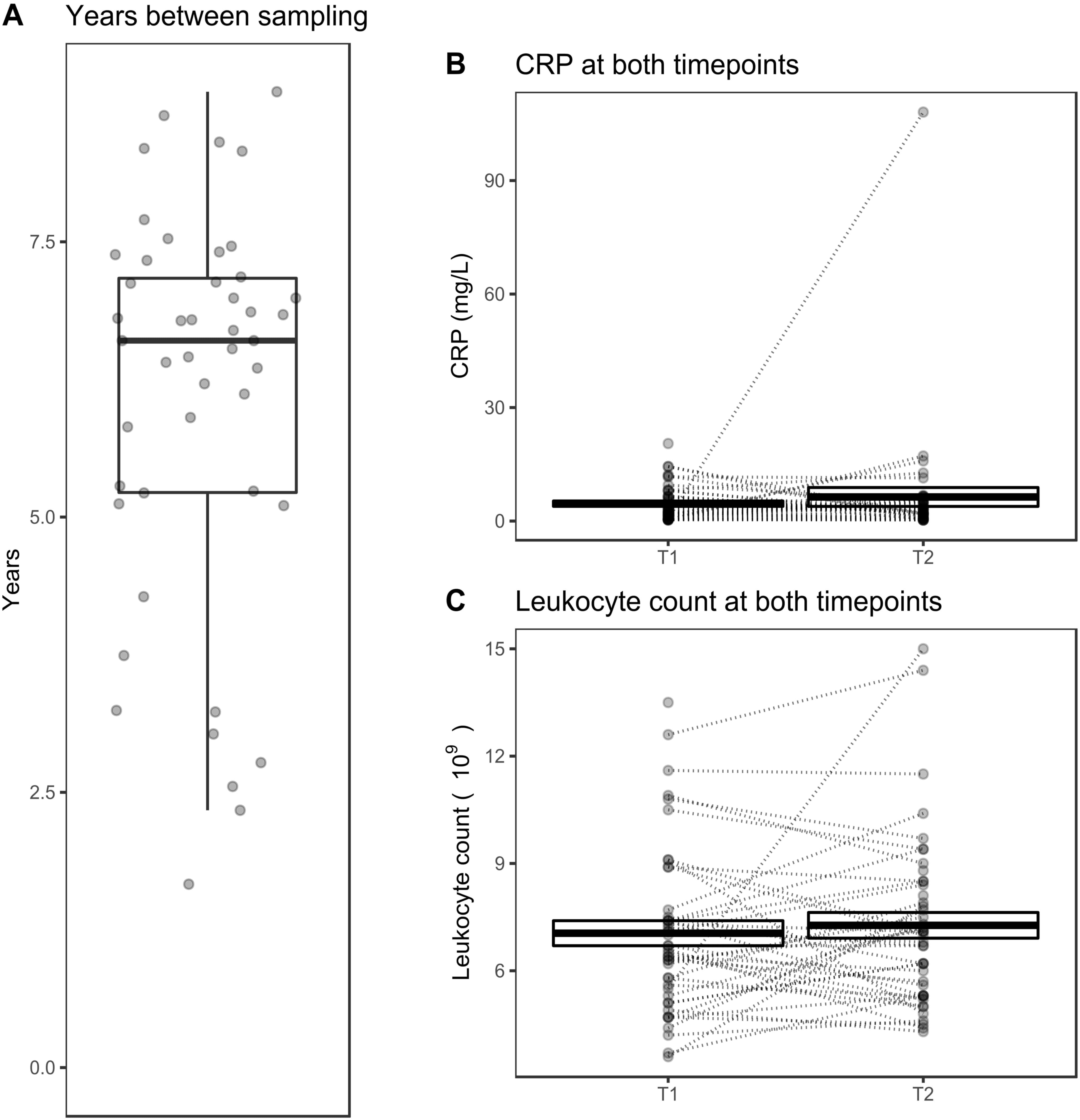
Supplementary Figure 4. Visualisation of the ICC values of all Illumina CpGs annotated to IBD GWAS risk genes, relative to their position on each gene. Dots represent individual CpG loci, dots above the red dashed line are considered SMPs (ICC≥0.75) whereas dots above the blue dashed line are HSMPs (ICC≥0.9).

Supplementary Figure 5. Visualisation 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 (coral), predicted GV (green) or methylation (blue). Dots below represent known genetic variants as reported by Goyette[64] for CD versus healthy controls (coral) and UC versus healthy controls (turquoise).

Supplementary Table 1. Summary results of the ICC results for each of the identified consistent DMPs by Joustra et al.[60], stratified by comparison (CD vs. HC, UC vs. HC and IBD vs. HC).

Supplementary Table 2. Summary of the ICC results for each of the CpGs located on the IBD GWAS risk genes.

Supplementary Table 3. Summary of the ICC results for each of the CpGs located on the HLA genes.


A
Years between sampling
Years
7.5
5.0
2.5
0.0
B
CRP at both timepoints
CRP (mg/L)
90
60
30
0
T1
T2
C
Leukocyte count at both timepoints
Leukocyte count (10^9)
15
12
9
6
T1
T2

A
T2 vs T1
-log10(p-value)
Mean difference % methylation
B
Donor
Timepoint
cg07350163
cg08064261
cg01485645
cg09281539
cg08655589
cg00820581
cg02994863
cg27344330
cg26634885
cg20095338
cg20908894
cg21168458
cg02325880
cg23465967
cg25910466
cg10460130
cg07182872
cg17218495
cg07899956
cg08429943
cg21126680
cg02963846
cg19237483
cg18318765
cg16250402
cg17958763
cg24582925
cg15911024
cg08332875
cg17978092
cg16162018
cg01828328
cg25783315
cg25288464
cg01495670
cg23101138
cg19034149
cg08064556
cg10569112
cg09642196
Donor
F006
F008
F009
F013
F014
F027
F032
F042
F047
F055
F056
F063
F066
F070
F075
F079
F082
F098
F102
F132
F135
F136
F141
F185
F188
F193
F201
F213
F216
C
-log10(p-value)
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
chrX
chrY
D
CD4T
p-value = 0.099
Neu
p-value = 0.104
NK
p-value = 0.237
Bcell
p-value = 0.264
CD8T
p-value = 0.274
Mono
p-value = 0.771
Proportion
Timepoint
T1
T2
E
Age-associated CpGs
-log10(p-value)
Mean difference % methylation
F
p-value = 1.58e-06
NES = -1.398
NES
Rank
0e+00
2e+05
4e+05
6e+05
8e+05

A
QC GV
Predicted GV
Potential GV
Methylation
-0.5
0.0
0.5
1.0
ICC
B
cg04998153
LINC01307
ICC = 0.99 [0.982, 0.995]
cg06231783
BRSK2
ICC = 0.99 [0.983, 0.995]
cg22113540
GRIN2A
ICC = 0.99 [0.983, 0.995]
cg04546413
ICC = 0.992 [0.986, 0.996]
cg05896524
C21orf56
ICC = 0.991 [0.984, 0.995]
cg20540428
PPP4R2
ICC = 0.99 [0.983, 0.995]
cg22243260
ICC = 0.992 [0.986, 0.996]
cg16885113
ICC = 0.99 [0.982, 0.995]
cg18280909
ICC = 0.991 [0.983, 0.995]
cg09271603
PDK3
ICC = 0.99 [0.982, 0.995]
cg08222221
ICC = 0.99 [0.983, 0.995]
% Methylation
Timepoint
T1
T2

A
Meta-analysis-derived DMPs
Joustra et al. 2022
IBD vs HC
UC vs HC
CD vs HC
0.00
0.25
0.50
0.75
ICC
ICC_class
Excellent
Good
Moderate
Poor
B
cg12054453
VMP1
ICC = 0.621 [0.405, 0.771]
cg16936953
VMP1
ICC = 0.59 [0.365, 0.75]
cg17501210
RPS6KA2
ICC = 0.693 [0.507, 0.818]
% Methylation
Timepoint
T1
T2
C
cg25112191
RORC
ICC = 0.408 [0.136, 0.623]
cg09171692
RORC
ICC = 0.759 [0.604, 0.859]
cg13707793
CXXC5
ICC = 0.496 [0.243, 0.686]
cg04570316
GMNN
ICC = 0.569 [0.337, 0.736]
cg21049840
GPR183
ICC = 0.492 [0.239, 0.683]
cg06460200
GPR183
ICC = 0.449 [0.185, 0.652]
cg06366627
DIDO1
ICC = 0.475 [0.217, 0.671]
cg02240291
SMARCD3
ICC = 0.745 [0.581, 0.85]
cg02055816
SMARCD3
ICC = 0.868 [0.774, 0.925]
cg00092736
ESPNL
ICC = 0.725 [0.553, 0.838]
cg09241586
ESPNL
ICC = 0.697 [0.512, 0.82]
cg15860510
EPS8L3
ICC = 0.382 [0.106, 0.603]
% Methylation
Timepoint
T1
T2

A
IBD-associated GWAS genes
ICC
1.0
0.5
0.0
HLA-DRB1
HLA-DQA1
HLA-DQB1
HLA-DPB1
HLA-DPA1
HLA-A
HLA-C
HLA-B
CARD9
NOD2
IL23R
FUT2
TYK2
TNFSF15
ATG16L1
B
ICC
HLA-A
HLA-B
HLA-C
HLA-DPA1
HLA-DPB1
HLA-DQA1
HLA-DQB1
HLA-DRB1
1.00
0.75
0.50
0.25
0.00
-0.25
Annotated GV
Gaphunter
Methylation
C
HLA-C
Gene
uc021yul.1
uc031shl.1
uc003nsy.3
uc021yuk.1
Modality
Annotated GV
Gaphunter
Methylation
$-\log_{10}$(p-value)
10
20
30
ICC value
1.0
0.5
0.0
-0.5
Comparison
CDvHC
UCvHC
IBD GVs
$-\log_{10}$(p-value)
8
6
4
2
0
31236000
31238000
31240000
D
HLA-DPB1
Gene
uc011dqq.1
uc011dqo.2
uc003ocu.2
uc011dqn.2
$-\log_{10}$(p-value)
10
20
30
40
Modality
Annotated GV
Gaphunter
Methylation
ICC value
1.0
0.5
0.0
Comparison
CDvHC
UCvHC
IBD GVs
$-\log_{10}$(p-value)
15
10
5
0
33045000
33050000
33055000
E
HLA-DPA1
Gene
uc010juk.3
uc003ocs.2
uc021ywg.1
uc021ywh.1
$-\log_{10}$(p-value)
10
20
30
40
Modality
Annotated GV
Gaphunter
Methylation
ICC value
1.0
0.5
0.0
Comparison
CDvHC
UCvHC
IBD GVs
$-\log_{10}$(p-value)
8
6
4
2
0
33035000
33040000