OXFORD

Original Article

Serum proteomic profiling at diagnosis predicts clinical course, and need for intensification of treatment in inflammatory bowel disease

R. Kalla,[a,b,] A. T. Adams,[a,c] D. Bergemalm,[d] S. Vatn,[e] N. A. Kennedy,[a,f]
P. Ricanek,[e,o] J. Lindstrom,[g,o] A. Ocklind,[h] F. Hjelm,[h] N.T. Ventham,[a]
G. T. Ho,[b] C. Petren,[h] IBD-Character Consortium, D. Repsilber,[i]
J. Söderholm,[j] M. Pierik,[k,] M. D'Amato,[l,m] F. Gomollón,[n] C. Olbjorn,[e,o]
J. Jahnsen,[e,o] M. H. Vatn,[o] J. Halfvarson,[d] J. Satsangi[a,c]

[a]Institute of Genetics and Molecular Medicine, University of Edinburgh, UK [b]MRC Centre for Inflammation Research, Queens Medical Research Institute, University of Edinburgh, UK [c]Translational Gastroenterology Unit, Nuffield Department of Medicine, Experimental Medicine Division, University of Oxford, John Radcliffe Hospital, Oxford, UK [d]Department of Gastroenterology, Faculty of Medicine and Health, Örebro University, Örebro, Sweden [e]Department of Gastroenterology, Akershus University Hospital, Lørenskog, Norway [f]Exeter IBD and Pharmacogenetics group, University of Exeter, UK [g]Health Services Research Unit, Akershus University Hospital, Lørenskog, Norway [h]Olink Proteomics, Uppsala, Sweden [i]School of Medical Sciences, Örebro University, Örebro, Sweden [j]Department of Surgery and Clinical and Experimental Medicine, Linköping University, Linköping, Sweden [k]Maastricht University Medical Centre (MUMC), Department of Gastroenterology and Hepatology, Maastricht, Netherlands [l]BioCruces Health Research Institute and Ikerbasque, Basque Foundation for Science, Bilbao, Spain [m]School of Biological Sciences, Monash University, Victoria, Australia [n]HCU 'Lozano Blesa', IIS Aragón, Zaragoza, Spain [o]Institute of Clinical Medicine, Campus Ahus, University of Oslo, Oslo, Norway

Corresponding author: Professor J. Satsangi, Translational Gastroenterology Unit, Nuffield Department of Medicine, Experimental Medicine Division, University of Oxford, John Radcliffe Hospital, Oxford, UK. Email: jack.satsangi@ndm.ox.ac.uk.

Abstract

Background: Success in personalized medicine in complex disease is critically dependent on biomarker discovery. We profiled serum proteins using a novel proximity extension assay [PEA] to identify diagnostic and prognostic biomarkers in inflammatory bowel disease [IBD].

Methods: We conducted a prospective case-control study in an inception cohort of 552 patients [328 IBD, 224 non-IBD], profiling proteins recruited across six centres. Treatment escalation was characterized by the need for biological agents or surgery after initial disease remission. Nested leave-one-out cross-validation was used to examine the performance of diagnostic and prognostic proteins.

Results: A total of 66 serum proteins differentiated IBD from symptomatic non-IBD controls, including matrix metallopeptidase-12 [MMP-12; Holm-adjusted $p = 4.1 \times 10^{-23}$] and oncostatin-M [OSM; $p = 3.7 \times 10^{-16}$]. Nine of these proteins are associated with *cis*-germline variation [59 independent single nucleotide polymorphisms]. Fifteen proteins, all members of tumour necrosis factor-independent pathways including interleukin-1 (IL-1) and OSM, predicted escalation, over a median follow-up of 518 [interquartile range 224–756] days. Nested cross-validation of the entire data set allowed characterization of five-protein models [96% comprising five core proteins ITGAV, EpCAM, IL18, SLAMF7 and IL8], which define a high-risk subgroup in IBD [hazard ratio 3.90,

confidence interval: 2.43–6.26], or allowed distinct two- and three-protein models for ulcerative colitis and Crohn's disease respectively.

Conclusion: We have characterized a simple oligo-protein panel that has the potential to identify IBD from symptomatic controls and to predict future disease course. Further prospective work is required to validate our findings.

1. Introduction

Personalized medicine is now a major priority in healthcare research. Programmes such as the 7th framework programme for research and technological development and 100,000 Genomes Project [www.genomicsengland.co.uk] in the UK prioritize the discovery and validation of novel biomarkers in human diseases.[1] This impetus to redefine clinical practice coupled with an expanding therapeutic choice of biological agents, together with small molecules, has driven interest in risk-stratifying patients at diagnosis in inflammatory bowel disease [IBD].[2–4]

There have been recent scientific advances catalysing biomarker discovery studies. It is now apparent that genes that contribute to prognosis in Crohn's disease [CD] are distinct from those that predict disease susceptibility.[4] Studies in both adults and children have demonstrated that patients with a progressive disease display a unique transcriptional signature.[3,5–7] Critically for translation, emergent data demonstrate that early biomarker-driven therapeutic interventions can improve disease outcomes in CD.[8]

Despite significant progress in multi-omic biomarker discoveries, none has been translated into routine clinical practice. Markers such as C-reactive protein [CRP] have shown clinical utility in disease susceptibility, activity and behaviour.[1] Faecal calprotectin [FC], however, has emerged to date as the most reliable and accurate diagnostic protein biomarker in IBD.[9] Recently, randomized trial data have demonstrated that early biomarker-driven therapeutic interventions based on FC can improve disease outcomes in CD. [8] However, there are well-described limitations to faecal testing in clinical care,[2,10,11] highlighting the need for blood-based markers to maximize uptake and acceptability.

Multi-protein signatures have potentially diverse clinical applications from early detection of IBD to disease classification and behaviour, response to therapy, and monitoring disease activity. Technological limitations in multi-protein profiling have recently been overcome,[12,13] with the discovery of innovative approaches for multiplexing biological samples utilizing minimal sample volume but providing a highly sensitive and specific immunoassay. Proximity extension assays [PEAs] are antibody-based methods that utilize two or more DNA-tagged aptamers or antibodies that bind when in close proximity to the target protein or protein complex. PEAs allow multiplexing with 1 μL sample consumption, and a high sensitivity and specificity for proteins of interest.[12,13]

In this report, we explore the diagnostic and prognostic performance of circulating PEA-based proteins markers in IBD and their association with germline variations. Our study demonstrates that protein panels can predict disease and its course.

2. Materials and Methods

2.1. Study design

We conducted a prospective, multi-centre case-control study in patients with suspected or confirmed IBD, recruited at presentation as out-patients or as in-patients across six clinical centres in Europe [EU Character reference no. 305676] from May 2012 until September 2015. The diagnosis of ulcerative colitis [UC], CD and IBD-unclassified [IBDU] was based on internationally accepted criteria, following thorough clinical, microbiological, endoscopic, histological and radiological evaluation. The control group consisted of patients with gastrointestinal symptoms [symptomatic controls], who had no discernible evidence of IBD at any time during follow-up. We recorded information on demographics, clinical characteristics according to the Montreal and Paris classification, and details of drug therapies at baseline, i.e. at recruitment [Table 1].[14–16] Treatment naivety within the IBD cohort was defined as no exposure to any IBD-related medical therapies such as steroids, 5-aminosalicylic acid [5-ASA], immunomodulators and biologics [Supplementary Table 1]. Blood samples for protein profiles and genotyping and stool samples for FC were collected at baseline, i.e. at the time of recruitment. High-sensitivity CRP [hsCRP], albumin and FC were re-assayed in a single batch at the end of recruitment. Patients with IBD were followed prospectively and information on clinical outcomes was collected during follow-up. Treatment escalation was defined as the need for a biologic, ciclosporin or surgery, instituted for disease flare after initial induction therapy and aiming to induce disease remission. In UC, the definition of treatment escalation also included colectomy during index admission.

All centres were granted local ethics approval for this study and all patients gave written informed consent prior to inclusion.

2.2. Sample collection and processing

We collected blood samples [Vacuette gel tube with clot activator] and separated serum after centrifugation at 2000 *g* for 10 min and within 2 h of sampling. All serum were subsequently stored as aliquots at –80°C until further use. Whole-blood leukocyte DNA was extracted using the Nucleon BACC 3 DNA extraction kit [GE Healthcare]. We genotyped patients using the Illumina OmniExpressExome-8 Bead Chip [Illumina].

2.3. Serum protein profiling

We generated a candidate list of proteins based on the 163 IBD risk genes identified from genome-wide association studies[17] and from the existing literature relating to pathogenesis in IBD. After thorough quality control, assay analyses and validation, we built five unique multiplex protein panels, each consisting of 92 proteins. Thus, a total of 460 proteins were analysed. These proteins are involved in various IBD-related mechanisms, such as inflammation, immune regulation, metabolism and cell–cell signalling, and are listed in Supplementary Table 2.

We used the PEA technology to measure protein concentrations.[12,13] The methodology has been described in detail elsewhere.[13] Briefly, pairs of antibodies are used towards each target antigen. When both antibodies bind to the same antigen in close proximity,

Table 1. Demographics and clinical characteristic of patients with newly diagnosed inflammatory bowel disease and symptomatic controls.

	Inflammatory bowel diseases [*n* = 328]	Symptomatic controls [*n* = 224]
Mean age [range], years	34 [7–78]	34 [3–79]
Males [%]	172 [52%]	104 [46%]
Smoking status [current: never: ex: missing]	53:139:107:29	48:100:56:20
High sensitivity C-reactive protein (mg/L): median [range]	22 [0–300]	5 [0–85]
Albumin (g/L): median [range]	37 [13–50]	40[29–52]
Faecal calprotectin (µg/g): median [range]	1298 [32–6001]	78.5 [4–2647]
Subtype of IBD		
Crohn's disease	146 [45%]	
Ulcerative colitis	153 [47%]	
Inflammatory bowel disease unclassified [IBDU]	29 [8%]	
Treatment naïve	235 [72%]	
CD location at diagnosis		
L1 [terminal ileum]	46 [32%]	
L2 [colon]	43 [29%]	
L3 [ileocolon]	53 [36%]	
L4 [upper GI tract]	4 [3%]	
CD behaviour at diagnosis		
B1, B1p [non-stricturing & non-penetrating, +perianal]	111, 6 [76%, 4%]	
B2, B2p [stricturing, +perianal]	12, 0 [8%, 0%]	
B3, B3p [penetrating, +perianal]	7, 6 [5%, 4%]	
Not available	4 [3%]	
Extent of UC at diagnosis		
E1 [proctitis]	39 [25%]	
E2 [left sided]	47 [31%]	
E3 [extensive colitis]	63 [41%]	
Not available	4 [3%]	
Centre		
Edinburgh, UK	107	74
Oslo, Norway	119	60
Orebro, Sweden	57	30
Linkoping, Sweden	16	23
Zaragosa, Spain	24	37
Maastricht, Netherlands	5	0

NA, not applicable; CD, Crohn's disease; UC, ulcerative colitis; IBDU, inflammatory bowel disease unclassified.

attached oligonucleotides hybridize. The oligonucleotide templates are extended and amplified using polymerase chain reaction [PCR] [96.96, Dynamic Array IFC] on a Biomark HD Instrument. For each panel, 92 oligonucleotide-labelled antibody probe pairs are allowed to bind to their respective target present in the sample. All samples were processed at Olink Proteomics.

To minimize inter- and intraplate variation, raw data [quantitative PCR Ct values] were normalized using internal controls in each multiplex reaction, negative controls and an inter-plate control on each plate, and then transformed using a predetermined correction factor. The pre-processed data were reported as arbitrary units, i.e. normalized protein expression [NPX] on a $\log_2$ scale as described previously.[12,13,18] A high NPX represents high protein concentration and a low NPX represent low protein concentration. The limit of detection [LOD] for each protein probe was defined as the mean plus three standard deviations of the negative controls.

To reduce the effect of biologically irrelevant differences or non-informative protein features, we first excluded 147/460 proteins where > 50% of samples were below the LOD, and then excluded 33 samples in which > 20% of the remaining proteins were below the LOD. After quality control, a total of 313 proteins were analysed in 552 patients.

2.4. Statistical analysis

We used R 3.4.4 [R Foundation for Statistical Computing] and Julia 1.1.0[19] for analysis. Data were corrected for centre batch effects using ComBat. All *p*-values were adjusted for multiple testing [Holm correction].[20] Survival analysis was performed using univariate Cox proportional hazard models and including age and sex as covariates. Hazard ratios [HRs] were calculated from Cox regression coefficients. HR represents the relative risk associated with a one-unit increase in expression of the relevant protein. Diagnostic analysis including sub-analysis differentiating UC from CD was performed using binomial logistic regression. We constructed models and characterized their predictive performance using a rigorous nested cross-validation approach wherein feature selection and parameter estimation were performed in an inner leave-one-out [LOO] cross-validation loop, with the model performance assessed using the unseen outer LOO sample. Reported performance of the models is based on the combined performance in each outer LOO sample of the models derived in their respective inner loops. Models were constrained to include age and sex, with proteins added in a forward stepwise approach based on Akaike's information criterion [AIC]. The number of included proteins was based on the AIC evidence ratio assessed in the first 10% of outer loops after which models were constrained to the selected number of proteins to reduce computation. No pre-selection or filtering of the proteins by any criteria was used prior to the cross-validation. Classification was based on the optimum threshold from receiver operating characteristic [ROC] analysis of the outer cross-validation loop. Randomly permuted data [*n* = 50] were analysed with the same technique with true data outperforming every permuted dataset.

Genome Studio files were imported into R for sex mismatch removal, and further analysis. Protein quantitative trait loci [pQTLs] were found using the matrix eQTL package[21] with a distance threshold of 300 kb and a minor allele frequency [MAF] threshold of > 0.1. Age and sex were included as covariates, and Holm correction was applied to p values. Further sub-analysis was performed with treatment exposure, sex, age, body mass index, clinical centre and smoking status as covariates.

3. Results

3.1. Differentially expressed protein markers in IBD

After quality control, a total of 313 proteins were analysed in 552 patients recruited across six IBD centres in Europe [Table 1]. Linear models with age and sex as covariates identified a total of 66 protein markers that showed significant differential expression between IBD [$n = 328$] and controls [$n = 224$, Figure 1; Supplementary Table 3], including matrix metallopeptidase-12 [MMP-12, $\log_2$fold change [$\log_2$FC = 0.87, $p = 4.1 \times 10^{-23}$] and oncostatin-M [OSM, $\log_2$FC = 0.81, $p = 3.7 \times 10^{-16}$]. Over-expression in IBD was more frequent at higher significance levels [$p = 0.01$], with the top 12 proteins all being over-expressed. Of the proteins down-regulated in IBD, the most significant were growth arrest-specific-6 [GAS6] and integrin alpha-V [ITGAV].

There were 55 protein markers that were significantly differentially expressed in CD compared to controls [Supplementary Table 4]; the most significant were chemokine (C-X-C motif) ligand 9 (CXCL9) [$\log_2$FC = 1.02, $p = 5.0 \times 10^{-15}$] and OSM [$\log_2$FC = 0.82, $p = 5.8 \times 10^{-12}$]. In UC, 46 protein markers had significant expression differences

compared to controls [Supplementary Table 5], including MMP-12 [$\log_2$FC = 1.14, $p = 3.6 \times 10^{-26}$] and granzyme-B [$\log_2$FC = 1.54, $p = 7.9 \times 10^{-23}$]. Five proteins showed significant expression differences between UC and CD [Supplementary Table 6; Figure 1B], all were significantly different between CD and controls, and differed further in the same direction in UC. A clinically useful model to distinguish between CD and UC could not be established, as the accuracy of the best performing classifier [consisting of age, sex and expression of six proteins] was only 68.0%. Correlations between protein expression and inflammatory markers are shown in Supplementary Figure S1.

3.2. Diagnosis of IBD with PEAs and inflammatory markers

We next examined the diagnostic performance of PEA-based protein models using the nested cross-validation approach, independent of the differential expression analysis, using all proteins profiled in this study. Fitting logistic regression models comprising age, sex and six protein expression values in a nested cross-validation approach was 79.8% (95% confidence interval [CI]: 76.4–83.2) accurate at distinguishing IBD from controls [sensitivity 83.1%, CI: 79.1–87.2; specificity 74.8%, CI: 69.0–80.5]. The proteins selected by each inner cross-validation loop were stable, comprising Granzyme-B [selected by 100% of inner loops], MMP-12 [100%], GAS6 [99.8%], interleukin-7 [IL7, 99.6%], IL8 [99.6%] and extracellular matrix metalloproteinase inducer [EMMPRIN, 99.3%].

This model outperformed an hsCRP model with age and sex, which had a sensitivity of 77.5% [CI: 72.7–82.3], specificity of 27.8% [CI: 21.5–34.0] and accuracy of 57.2% [CI: 52.9–61.7]
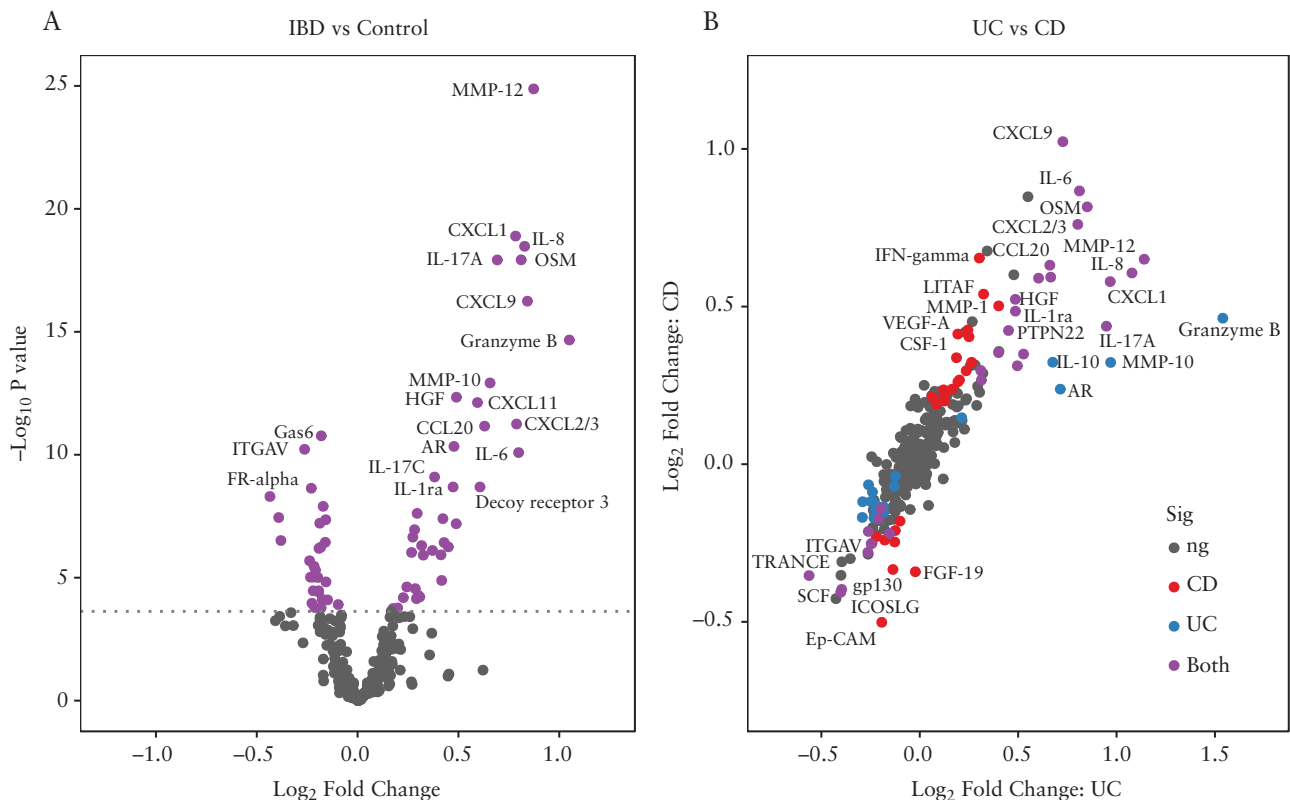

Figure 1. [A] Volcano plot displaying the $\log_2$ fold-change and significance of protein associations with IBD. Dotted line indicates the threshold for significance [Holm $p < 0.05$] after Holm correction derived from $-\log_{10}$[uncorrected $p = 1.78 \times 10^{-4}$]. [B] Fold change between ulcerative colitis [UC] and Crohn's disease [CD] respectively vs controls; points are coloured by significance after Holm correction in CD, UC, both or neither [ns].

[Supplementary Table 7]. An FC model with age and sex performed better [sensitivity 85.4%, CI: 78.1–92.7; specificity 88.4%, CI: 78.8–98.0; accuracy 86.4%, CI: 80.5–92.2%], but FC suffers from poor uptake, with only 30.4% of patients having a result between 30 days before and 7 days after inclusion.

The PEA-based models performed similarly in UC and CD [accuracy 78.4 and 77.7% respectively], and separate analysis of CD and UC did not produce more accurate models. FC was more sensitive in UC compared to CD [90.7%, CI: 83.0–98.5% vs 77.4%, CI: 62.7–92.1%; χ^2 p = 1.2 × 10^{-12}], yielding an improved accuracy of 89.7% [CI: 83.6–95.7%] vs 83.8% [CI: 75.4–92.2%] [Supplementary Table 7].

3.3. Individual proteins associated with treatment escalation

To identify proteins that associate with treatment escalation, we analysed data from 279 patients with confirmed IBD from whom follow-up data were available [Table 2; Supplementary Table 8]. Patients who required treatment escalation were younger [median age 28 vs 33 years, p = 0.02], more likely to be male [58.2 vs 51.4%, χ^2 p > 0.05] and have CD [58.2 vs 34.4%, χ^2 p = 0.004]. There was no significant association between treatment escalation and smoking status amongst patients with CD or UC.

Cox models were created to identify protein markers individually associated with treatment escalation in IBD, accounting for age and sex. Fifteen proteins [Figure 2 and Table 3] were significantly associated with treatment escalation in IBD, including ITGAV [p = 3.2 × 10^{-6}] and Epithelial Cell Adhesion Molecule (EpCAM) [p = 1.7 × 10^{-4}]. Adjusting for treatment naivety did not influence the top differentially expressed proteins among patients with IBD. In UC [n = 143], 22 proteins were significantly associated with treatment escalation [Supplementary Table 9], whereas in CD [n = 112] no individual proteins achieved significance, although the results were correlated with those obtained for UC alone [r = 0.56, p = 6.6 × 10^{-15}].

Table 2. Demographics and clinical characteristics of patients with inflammatory bowel disease who were included in the predictive analysis of disease course

Inflammatory bowel disease

	IBD escalation group [n = 67]	IBD non-escalation group [n = 212]
Males [%]	39 [58]	109 [51]
Smoking status [current: never: ex: missing]	16:34:16:1	36:98:77:1
Median FC [µg/g; range]	1631 [35–6001]	1186 [32–6001]
Median age [range], years	28 [18–67]	33 [18–77]
Edinburgh: Norway: Sweden: Spain	26:22:15:4	81:71:41:19
Disease subtype		
Crohn's disease	39	73
Ulcerative colitis	26	117
Inflammatory bowel disease unclassified [IBDU]	2	22

Ulcerative colitis

	Escalation group [n=26]	Non-escalation group [n=117]
Males [%]	19 [73]	67 [57]
Smoking status [current: never: ex: missing]	3:9:14:0	8:53:56:0
Median FC [range]	3778 [35–6001]	1367 [32–6001]
Median age [range], years	30 [18–60]	37 [18–77]
Edinburgh: Norway: Sweden: Spain	13:8:4:1	39:52:19:7
Paris extent for UC		
E1 [proctitis]	0	38 [32%]
E2 [left sided]	7 [27%]	37 [32%]
E3 [pancolitis]	19 [73%]	42 [36%]

Crohn's disease

|	Escalation group [n = 39]	Non-escalation group [n = 73]
Males [%]	19 [49]	33 [45]
Smoking status [current: never: ex: missing]	13:5:20:1	26:18:28:1
Median FC [range]	1398.5 [47–6001]	825 [70–6001]
Median age [range], years	25 [18–66]	29 [18–73]
Edinburgh: Norway: Sweden: Spain	11:14:11:3	34:17:12:10
Montreal classification for CD		
L1 [terminal ileum]	13 [33%]	25 [34%]
L2 [colonic]	9 [23%]	22 [30%]
L3 [ileocolon]	17 [44%]	25 [34%]
L4 [upper GI tract]	0	1 [1%]
Montreal behaviour for CD		
B1, B1p [non-stricturing & non-penetrating, +perianal]	29, 0 [74%, 0%]	55, 6 [75%, 8%]
B2, B2p [stricturing, +perianal]	6, 0 [15%, 0%]	4, 0 [5%, 0%]
B3, B3p [penetrating, +perianal]	2, 2 [5%, 5%]	5, 1 [7%, 1%]
Not available	0	2 [3%]
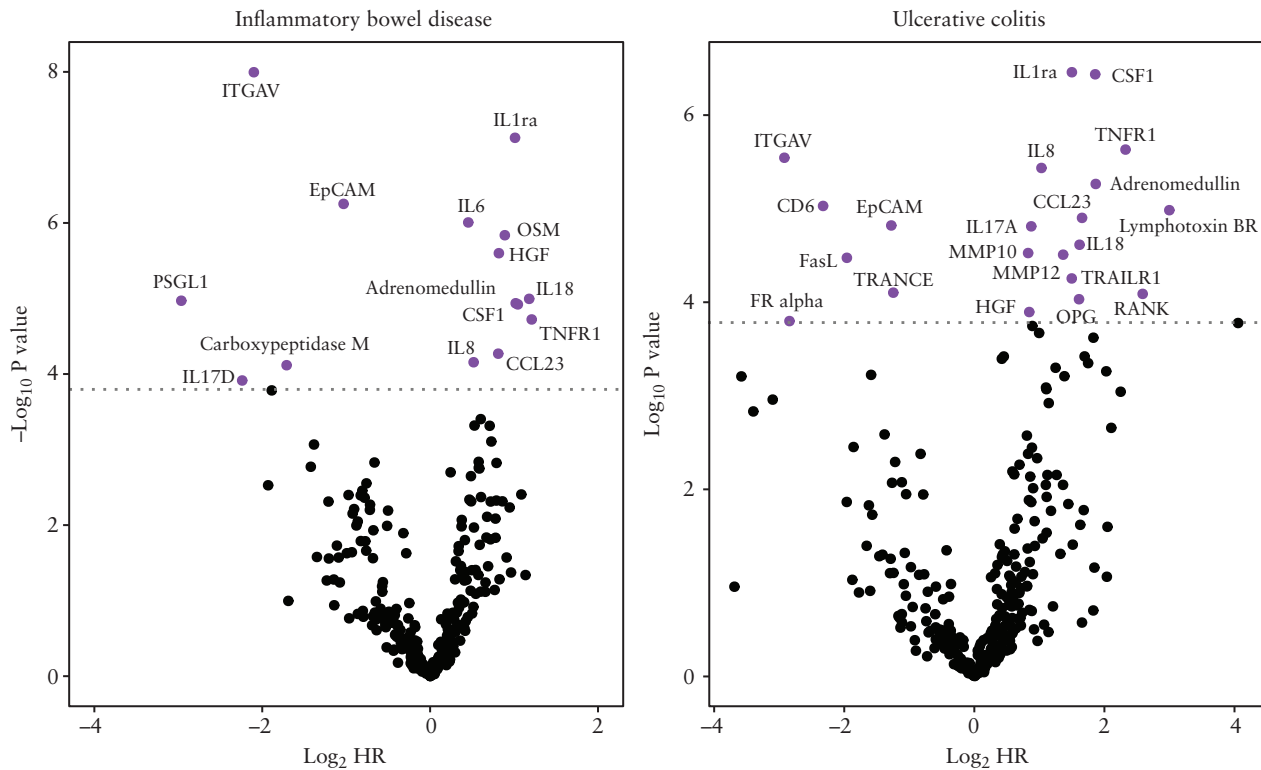
Figure 2. The significance of protein markers in predicting treatment escalation in inflammatory bowel disease and ulcerative colitis. Significance threshold after Holm correction is indicated by a dotted line.

3.4. Nested cross-validation stratifies disease subgroups that associate with treatment escalation

Models to define need for treatment escalation consisting of age, sex, IBD subtype and PEA-protein expression values were generated in each inner LOO cross-validation loop and tested in the outer loop. The models selected were highly stable. A series of five-protein models had highest predictive accuracy, with 96% of these models consisting of the same five proteins [ITGAV, EpCAM, IL18, SLAM family member 7 (SLAMF7) and IL8].

These models defined by cross-validation had 80.0% [CI: 75.3–84.7%] accuracy (sensitivity 47.6% [CI: 35.3–60.0%], specificity 89.6% [CI: 85.5–93.7], with a positive likelihood ratio [LR+] 4.59 [CI: 2.86–7.36], and negative likelihood ratio[LR−] 0.58 [CI: 0.46–0.74]). The high-risk group required treatment escalation at 3.9 [CI: 2.4–6.3] times the rate of the low-risk group. FC values were higher in patients later requiring treatment escalation [Table 2], although this finding was not significant whether analysing CD [p = 0.63] and UC [0.09] separately, or in all IBD [p = 0.14].

A simple categorization for all patients as high or low risk may not be the most useful interpretation of the protein expression panels. Subgroups can be identified at particularly high or low risk of aggressive disease tailored to an appropriate level for the intended action to be taken. Supplementary Figure 2 depicts these data in a graphical format. Each subsection represents the results from labelling a proportion of the population as low [x-axis] and high risk [y-axis]. Within each subsection the top left and bottom right numbers denote the percentage of the identified group requiring escalated treatment in the high- and low-risk groups respectively. The top right number in each subsection represents the relative risk between groups. As an example, identifying the quartiles of patients at highest and lowest risk selects a subset where 52.8% and 5.8% respectively

required treatment escalation in the first 18 months of treatment, with a relative risk ratio between groups of 9.1.

Although analysing all IBD patients [Supplementary Figure S3] in this cohort together produces models which work in both CD and UC, the accuracy achieved in UC is significantly higher than that in CD [85.1%, CI: 79.2–91.0% vs 70.9%, CI: 62.4–79.4%; χ^2 p = 0.007]. The same analytical approach applied individually to UC and CD produces simpler models [two and three proteins respectively, Supplementary Figure S4], with 79.4% [CI: 72.8–86.1%] accuracy in UC outperforming accuracy in CD [76.4%, CI: 68.4–84.3%]. As with the pan-IBD analysis, the probes selected by the inner cross-validation loops were consistent with cluster of differentiation 6 (CD6) and macrophage colony stimulating factor 1 (CSF1) in 92% of UC models and Lipopolysaccharide Induced TNF Factor (LITAF), Carboxypeptidase M (CPM) and CCL28 in 99, 97 and 88% of CD models respectively.

3.5. Performance of PEA prognostic models against conventional predictors of escalation

We compared the performance of PEA-based prognostic proteins to currently available blood and faecal biomarkers and clinical predictors in IBD and its subtypes; these are summarized in Supplementary Table 7. The performance of the PEA model is comparable to hsCRP [HR 2.74, CI:1.32–5.67 vs six-protein model HR 3.90, CI: 2.43–6.26]. However, hsCRP suffers from poor sensitivity [0.20; CI: 11.1–33.1] compared to the PEA model [sensitivity 0.48, CI: 35.3–60.0]. A Cox model trained with FC or a combined model with FC and hsCRP performed poorly at predicting treatment escalation in IBD [FC HR 1.17, CI: 0.42–3.26; FC plus hsCRP model HR 0.74, CI: 0.18–3.08 respectively]. Clinical predictors such as

Table 3. Top 15 proteins associated with escalation in treatment [anti-TNF/ciclosporin and/or surgery] and their associated biology based on the available literature

Protein	*p* value	Log$_2$ HR	Holm *p* value	Family/group	Cell of origin	Function/relevance in IBD
ITGAV	1.01×10^{-8}	−−2.12	3.16×10^{-6}	Integrin signalling	NA	Known GWAS locus[22]
IL-1RA	7.46×10^{-8}	1.01	2.33×10^{-5}	IL-1	Macrophages and monocytes	Anti-IL1 drug in phase 2 trial in UC [ISRCTN43717130]
EpCAM	5.59×10^{-7}	−1.03	1.74×10^{-4}	NA	Epithelial cells	Intercellular adhesion molecule, maintaining intestinal immune balance[23]
IL-6	9.85×10^{-7}	0.45	3.05×10^{-4}	IL-6 family	Th cells and macrophages	Pro-inflammatory response via IL1β and TNF
OSM	1.45×10^{-6}	0.89	4.49×10^{-4}	IL-6	Th cells and macrophages	Pro-inflammatory response and anti-TNF non-response[24]
HGF	2.51×10^{-6}	0.82	7.74×10^{-4}	Cytokine	Mesenchymal cells	Angiogenesis promotion and elevated levels in IBD[25]
IL-18	1.01×10^{-5}	1.18	3.10×10^{-3}	IL-1 family	Epithelial cells	IL-18 polymorphism associated with anti-TNF response[26]
PSGL1	1.07×10^{-5}	−2.94	3.28×10^{-3}	Selectin family	Leukocyte and endothelial surfaces	Anti-PSGL-1 drug in Phase 1 trial to treat CD [NIH #8307272]
ADM	1.16×10^{-5}	1.02	3.53×10^{-3}	Calcitonin peptide superfamily	Epithelial cells	Case series of mucosal healing in refractory UC with AM therapy[27]
CSF-1	1.20×10^{-5}	1.04	3.64×10^{-3}	IL-34/CSF-1 family	Various immune cells	Pro-inflammatory macrophage-induced response[28]
TNF-R1	1.89×10^{-5}	1.21	5.74×10^{-3}	TNF family	Macrophages and dendritic cells	Pro-inflammatory TNF-mediated response
CCL23	5.38×10^{-5}	0.81	0.016	CC chemokines	Epithelial and immune cells	Neutrophil activation and leukocyte migration[29]
IL-8	6.98×10^{-5}	0.52	0.021	CXC-chemokines	Epithelial cells, macrophages, monocytes	Neutrophil recruitment and pro-inflammatory response
CPM	7.64×10^{-5}	−1.69	0.023	Carboxy peptidases	Activated macrophages	Activated macrophage differentiating marker[30]
IL-17D	1.22×10^{-4}	−2.25	0.036	IL-17 family	Th-17 cells	Th-17-driven pro-inflammatory cytokine

Holm *p* represents *p* values adjusted for multiple testing. Log$_2$ HR [hazard ratio] is the relative risk associated with a one-unit increase in expression of the relevant protein.

non-B1 behaviour or perianal disease in CD, and Simple Clinical Colitis Activity Index or Harvey–Bradshaw Index scores did not significantly associate with treatment escalation, although pancolitis in UC did [uncorrected *p* = 0.002].

Compared to the overall PEA-protein model accuracy of 80.0%, the addition of FC, CRP or both did not improve model performance, yielding accuracies of 76.5% [CI: 67.5–85.5%], 77.8% [CI: 72.8–82.8%] and 72.2% [CI: 62.3–82.0%] respectively, and nor did the addition of any phenotypic characteristic such as pancolitis in UC or perianal disease in CD. We also performed correlation analyses of the top protein markers with proteins associated with IBD, hsCRP, albumin and FC and these are summarized in Supplementary Figure S5.

3.6. Circulating proteins associate with germline variation

It has been shown that expression of proteins is linked to germline variation, mainly in the *cis* regions of their encoding genes.[31] We explored the influence of germline variation on the expression of key IBD diagnostic and prognostic proteins identified in our analysis. We used linear regression models with age and sex as covariates, to analyse single nucleotide polymorphisms [SNPs; MAF > 0.1] correlated with protein expression, revealing 769 significant *cis* pQTLs affecting 51 proteins. These included 59 significant *cis* pQTLs affecting nine proteins with significant expression changes

associated with IBD [Supplementary Figure S6, Supplementary Table 10], and 35 pQTLs affecting proteins implicated in disease course [Supplementary Figure S7]. Vascular endothelial growth factor-A [VEGF-A] showed the most significant association with genotype (lead SNP rs7767396; effect [β] −0.42; MAF = 0.46; *p* = 8.7 × 10^{-18}) with a total of six significant SNP associations and 14 SNPs in linkage disequilibrium with rs7767396.

Among the proteins individually significantly associated with aggressive disease [Table 3] or frequently selected in the multi-protein models for aggressive disease, significant pQTLs were found in CD6, RANK and SLAMF7 [Supplementary Figure S7], in addition to the findings described in CCL23 above [Supplementary Figure S6].

4. Discussion

With advances in clinical care in IBD, it is widely recognized that there is a need for biomarkers that provide accurate diagnostic and prognostic testing in IBD. The key innovation in the present study is the design and evaluation of a novel multi-protein panel in newly diagnosed IBD, chosen a priori on the basis of known or suspected involvement in pathogenesis. The results substantiate the involvement of key pathways in pathogenesis, and provide targets for therapy. Importantly, we demonstrate that this strategy of biomarker discovery is feasible in diagnosis and in predicting treatment escalation in CD and UC.

A panel of six proteins had 79.8 % accuracy, 83.1% sensitivity and 74.8% specificity at differentiating IBD from controls. Whilst FC did outperform this panel [86.4% accuracy, 85.4% sensitivity, 88.4% specificity], uptake was low, with overall patient acceptability a major limiting factor. Given this widely recognized limitation of FC testing in the clinic,[10,11] we suggest that a serum protein biomarker panel could prove clinically useful as a diagnostic test in replacement of FC. Further studies are now needed to test and validate the utility of this protein panel in clinical practice.

Of the 66 differentially expressed proteins in IBD, nine demonstrated germline variation, VEGF-A being the most significant pQTL. Weaker correlations between protein expression and genetic variation were observed in four of the proteins that predicted treatment escalation, namely CCL23, RANK, CD6 and SLAM7. It has yet to be determined whether these genetic associations are causal in both disease onset and course, and our study provides a resource to investigate these associations further.

The greatest unmet need is for biomarkers that can determine disease activity, behaviour and extent, and most critically to predict response to treatment. In our dataset, we have been able to characterize and rigorously cross-validate models involving a limited number of proteins that predict disease course. The role of biomarkers in predicting the disease course has been the focus of many studies,[2–7,32,33] including our own parallel studies of glycomic and methylation profiling in the EC-funded consortia.[32,34] Lee *et al*. identified expression profiles of T cell exhaustion in CD8 T cells that predicted treatment escalation in IBD,[3] defining escalation as the need for two or more immunosuppressants and/or surgery after initial disease remission. A multi-gene signature predicting need for escalation using these original criteria has been proposed by this team in UC [HR 3.1, 95% CI: 1.25–7.72, p = 0.02] and CD [HR 2.7; CI: 1.32–5.34, p = 0.01].[7] This signature differs from the original profile of T cell exhaustion. Other studies have focused on mucosal healing, response to biological agents, and development of fistulizing or stricturing complications as endpoints—all valid in context.

In this study we decided to use more stringent criteria for escalation than those used in defining the transcriptional profile. We highlight the need for biologics or ciclosporin or surgical resection, rather than introduction of immunosuppression per se. This decision regarding the endpoint relates principally to the variable threshold for initiating immuno-modulators, which in practice have often been used as first-line therapy in CD. Our oligo-protein panels have the potential for clinical translation with significant practical benefits including the simplicity of the assay, and the ability to multiplex proteins using only 1 μL of serum.

It is noteworthy that the key prognostic proteins identified relate to pathways independent of tumour necrosis factor [TNF] signalling. OSM is a pro-inflammatory cytokine that promotes production of IL-6 to attract immune cells to the site of inflammation[24] and its intestinal expression in IBD has been shown to predict anti-TNF non-response in IBD.[24] We report that circulating levels of both IL-6 and OSM can predict treatment escalation in IBD. Similarly, we demonstrate the involvement of other pathways that predict disease course [Table 3; Supplementary Figure S8]. Of particular relevance are the proteins that show poor correlation with conventional inflammatory markers including hsCRP [Supplementary Figure S5], in particular PSGL-1. This protein is a P-selectin glycoprotein ligand that is expressed on the surface of most immune cells and facilitates immune cell trafficking across the endothelium.[35,36] A drug targeting

PSGL-1 is currently in phase 1 trial for the treatment of CD [NIH #8307272]. Future studies examining the performance of these markers in predicting response to therapy are now needed.

We recognize that clinical decisions and timing on treatment escalations may vary across centres. In this study all sites utilized a 'step-up approach' to treatment escalation, rather than a top-down approach. In this respect the clinical management is similar across centres and the consistency of our biomarker profile in predicting need for escalation across centres is especially noteworthy. Our study was not designed to detect the association between prognosis and endoscopic activity. Recently, a protein-based endoscopic healing index [EHI] has been reported that incorporates 13 proteins and performs on a par with FC in predicting endoscopic disease remission (validation cohort area under the receiver operating characteristic curve [AUROC], 0.803 for EHI vs AUROC, 0.854 for FC; p = 0.298], highlighting the translational potential of blood-based protein biomarkers in IBD.[37] The predictive capacity of our PEA model performs on a par with conventional blood tests such as hsCRP. However, it is worth noting that CRP suffers from poor sensitivity [0.20; CI: 11.1–33.1] compared to the PEA model [sensitivity 0.48, CI: 35.3–60.0]. Other markers such as FC suffer from poor uptake with only 85 FC results available for prognostic analysis in our study. Therefore, a blood-based PEA panel would be better at identifying patients likely to require treatment escalation. As the results of locally analysed CRP and albumin were known to clinicians making treatment decisions regarding escalation of therapy, it is likely these were in fact key determinants in decision-making. Because these measures are often regarded as proxies of inflammatory activity in clinical practice, these markers cannot be considered as independent predictors of disease progression. Our protein markers remain significant predictors of treatment escalation, independent of clinical confounders. We have utilized nested LOO cross-validation, which is acknowledged to produce an unbiased estimate of true error when properly nested so that the entire feature selection and parameter tuning process takes place without reference to the left out samples.[38] This methodology avoids biased estimates of performance and prevents over-fitting of the proposed models. Further validation is now needed to replicate our findings in other large multi-centre inception studies. The significance and impact of our analysis are strengthened by the pre-established evidence for these proteins in IBD or IBD-related pathways. For propriety reasons, details of the antibodies used for Olink panels were not available, but certain panels including the Olink Inflammation panel are now commercially available for further external validation. However, this is the largest inception cohort recruited in biomarker studies in adult IBD to date, allowing robust modelling and rigorous application.

With advances in IBD therapeutics, future challenges will include tailoring therapies based on individual disease biology. Our data provide insight into the importance of molecular characterization of patients with IBD at diagnosis to tailor medical therapies. These data also provide substantial progress towards the goal of developing a composite biomarker panel informing patients of their diagnosis and prognosis at their first clinic visit. In CD, the PROFILE trial is a landmark prospective biomarker-stratified study in IBD, using the Predict Immune panel, currently recruiting across the UK. [33] Our data provide an impetus for a similar protein-based biomarker trial in both CD and UC, and provide a rationale for multi-omic profiling to be integrated into trial design, and then into practice. The clear aspiration is that stratification with multi-omic biomarkers based on underlying disease mechanisms may enable personalized therapeutics.

Funding

Conflict of Interest

R.K.: financial support for research: EC IBD-Character, lecture fee[s]: Ferring; N.K.: financial support for research: Wellcome Trust, conflict with: Pharmacosmos, Takeda, Janssen, Dr Falk speaker fees, Abbvie, Janssen travel support; A.A.: none; J.S.: financial support for research: EC grant IBD-BIOM, Wellcome, CSO, MRC, conflict with: consultant for: Takeda, conflict with: MSD speaker fees. Shire travelling expenses.

Author Contributions

Study design R.K., J.H., M.D.A., M.V., J.S. Patient recruitment and sample processing N.T.V., R.K., N.A.K., D.B., S.V., A.T.A. Experimental work O.A., F.H., C.P., R.K., N.T.V., A.T.A., N.A.K. Data analysis R.K., N.A.K., A.T.A., D.R., J.L., D.B. R.K. and A.T.A. wrote the manuscript. All authors were involved in critical review, editing, revision and approval of the final manuscript.

Data Accessibility Statement

Study data can be shared on request to the corresponding author with permission from the IBD Character data access committee.

Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

References

1. Boyapati RK, Kalla R, Satsangi J, Ho GT. Biomarkers in search of precision medicine in IBD. *Am J Gastroenterol* 2016;**111**:1682–90.
2. Kalla R, Kennedy NA, Ventham NT, *et al*. Serum calprotectin: a novel diagnostic and prognostic marker in inflammatory bowel diseases. *Am J Gastroenterol* 2016;**111**:1796–805.
3. Lee JC, Lyons PA, McKinney EF, *et al*. Gene expression profiling of CD8+ T cells predicts prognosis in patients with Crohn disease and ulcerative colitis. *J Clin Invest* 2011;**121**:4170–9.
4. Lee JC, Biasci D, Roberts R, *et al*.; UK IBD Genetics Consortium. Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat Genet* 2017;**49**:262–8.
5. Kugathasan S, Denson LA, Walters TD, *et al*. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* 2017;**389**:1710–8.
6. Marigorta UM, Denson LA, Hyams JS, *et al*. Transcriptional risk scores link GWAS to eQTLs and predict complications in Crohn's disease. *Nat Genet* 2017;**49**:1517–21.
7. Biasci D, Lee JC, Noor NM, *et al*. A blood-based prognostic biomarker in IBD. *Gut* 2019;**68**:1386–95.
8. Colombel JF, Panaccione R, Bossuyt P, *et al*. Effect of tight control management on Crohn's disease (CALM): a multicentre, randomised, controlled phase 3 trial. *Lancet* 2018;**390**:2779–89.
9. van Rheenen PF, Van de Vijver E, Fidler V. Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. *BMJ* 2010;**341**:c3369.
10. Kalla R, Boyapati R, Vatn S, *et al*. Patients' perceptions of faecal calprotectin testing in inflammatory bowel disease: results from a prospective multicentre patient-based survey. *Scand J Gastroenterol* 2018;**53**:1437–42.
11. Maréchal C, Aimone-Gastin I, Baumann C, Dirrenberger B, Guéant J-L, Peyrin-Biroulet L. Compliance with the faecal calprotectin test in patients with inflammatory bowel disease. *United Eur Gastroenterol J* 2017;**5**:702–7.
12. Assarsson E, Lundberg M, Holmquist G, *et al*. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One* 2014;**9**:e95192.
13. Lundberg M, Eriksson A, Tran B, Assarsson E, Fredriksson S. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Res* 2011;**39**:e102.
14. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* 2006;**55**:749–53.
15. Levine A, Griffiths A, Markowitz J, *et al*. Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification. *Inflamm Bowel Dis* 2011;**17**:1314–21.
16. Maaser C, Sturm A, Vavricka SR, *et al*. ECCO-ESGAR Guideline for Diagnostic Assessment in IBD Part 1: Initial diagnosis, monitoring of known IBD, detection of complications. *J Crohns Colitis* 2019;**13**:144–64.
17. Jostins L, Ripke S, Weersma RK, *et al*.; International IBD Genetics Consortium (IIBDGC). Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;**491**:119–24.
18. Thorsen SB, Lundberg M, Villablanca A, *et al*. Detection of serological biomarkers by proximity extension assay for detection of colorectal neoplasias in symptomatic individuals. *J Transl Med* 2013;**11**:253.
19. Bezanson J, Edelman A, Karpinski S, Shah VB. Julia: a fresh approach to numerical computing. *SIAM Rev* 2017;**59**:65–98.
20. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat* 1979;**6**:65–70.
21. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 2012;**28**:1353–8.
22. de Lange KM, Moutsianas L, Lee JC, *et al*. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;**49**:256–61.
23. Jiang L, Shen Y, Guo D, *et al*. EpCAM-dependent extracellular vesicles from intestinal epithelial cells maintain intestinal tract immune balance. *Nat Commun* 2016;**7**:13045.
24. West NR, Hegazy AN, Owens BMJ, *et al*.; Oxford IBD Cohort Investigators. Erratum: Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease. *Nat Med* 2017;**23**:788.
25. Srivastava M, Zurakowski D, Cheifetz P, Leichtner A, Bousvaros A. Elevated serum hepatocyte growth factor in children and young adults with inflammatory bowel disease. *J Pediatr Gastroenterol Nutr* 2001;**33**:548–53.
26. Bank S, Julsgaard M, Abed OK, *et al*.; Danish IBD Genetics Working Group. Polymorphisms in the NFkB, TNF-alpha, IL-1beta, and IL-18 pathways are associated with response to anti-TNF therapy in Danish patients with inflammatory bowel disease. *Aliment Pharmacol Ther* 2019;**49**:890–903.
27. Ashizuka S, Inatsu H, Kita T, Kitamura K. Adrenomedullin therapy in patients with refractory ulcerative colitis: a case series. *Dig Dis Sci* 2016;**61**:872–80.
28. Marshall D, Cameron J, Lightwood D, Lawson AD. Blockade of colony stimulating factor-1 (CSF-I) leads to inhibition of DSS-induced colitis. *Inflamm Bowel Dis* 2007;**13**:219–24.
29. Singh UP, Singh NP, Murphy EA, *et al*. Chemokine and cytokine levels in inflammatory bowel disease patients. *Cytokine* 2016;**77**:44–9.
30. Tsakiris I, Torocsik D, Gyongyosi A, *et al*. Carboxypeptidase-M is regulated by lipids and CSFs in macrophages and dendritic cells and expressed selectively in tissue granulomas and foam cells. *Lab Invest* 2012;**92**:345–61.
31. Sun BB, Maranville JC, Peters JE, *et al*. Genomic atlas of the human plasma proteome. *Nature* 2018;**558**:73–9.
32. Clerc F, Novokmet M, Dotz V, *et al*.; IBD-BIOM Consortium. Plasma N-glycan signatures are associated with features of inflammatory bowel diseases. *Gastroenterology* 2018;**155**:829–43.
33. Parkes M, Noor NM, Dowling F, *et al*. PRedicting Outcomes For Crohn's dIsease using a moLecular biomarkEr (PROFILE): protocol

for a multicentre, randomised, biomarker-stratified trial. *BMJ Open* 2018;**8**:e026767.

34. 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 Commun* 2016;**7**:13507.

35. Guyer DA, Moore KL, Lynam EB, *et al*. P-selectin glycoprotein ligand-1 (PSGL-1) is a ligand for L-selectin in neutrophil aggregation. *Blood* 1996;**88**:2415–21.

36. Brown JB, Cheresh P, Zhang Z, Ryu H, Managlia E, Barrett TA. P-selectin glycoprotein ligand-1 is needed for sequential recruitment of T-helper 1 (Th1) and local generation of Th17 T cells in dextran sodium sulfate (DSS) colitis. *Inflamm Bowel Dis* 2012;**18**:323–32.

37. D'Haens G, Kelly O, Battat R, *et al*. Development and validation of a test to monitor endoscopic activity in patients with crohn's disease based on serum levels of proteins. *Gastroenterology* 2020;**158**:515–526.e10.

38. Lachenbruch PA, Mickey MR. Estimation of error rates in discriminant analysis. *Technometrics* 1968;**10**:1–11.

39. FANTOM Consortium and the RIKEN PMI and CLST (DGT), Forrest ARR, Kawaji H, Rehli M, *et al*. A promoter-level mammalian expression atlas. *Nature* 2014;**507**:462–70.