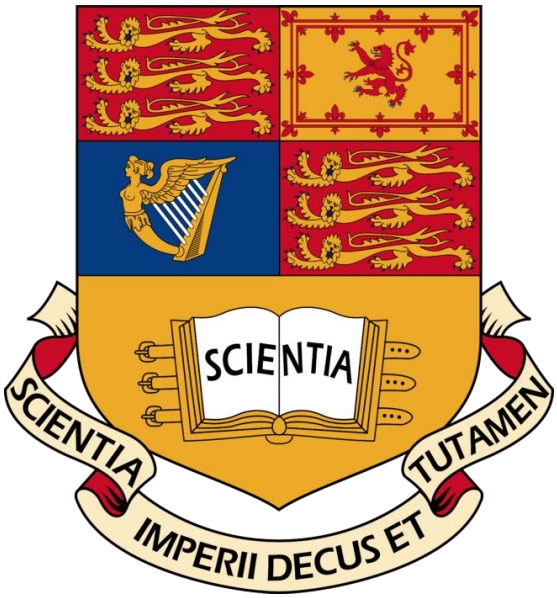
Exploring metabolite biomarkers for inflammatory bowel disease diagnosis and treatment monitoring



Full name: Ming Shin Cindy Sung

CID: 01873988

Word count: 4927

Project supervisor: Dr Jia Li, Dr Maria Valdivia-Garcia

Department of Metabolism, Digestion and Reproduction

This research report is submitted in partial fulfilment of the requirements for the degree of BSc in Medical Biosciences

Abstract

Inflammatory bowel disease (IBD) is a chronic intestinal inflammation and is mainly comprised of ulcerative colitis (UC) and Crohn's disease (CD). Diagnosis of IBD is currently based on medical history, physical examination, and laboratory tests, with endoscopic procedures being invasive and burdensome for patients. Therefore, finding a single non-invasive diagnostic tool for IBD is critical. Metabolic perturbations have been consistently reported in IBD pathophysiology, particularly in tryptophan, bile acid, and short-chain fatty acid metabolism. This study investigates serological biomarkers through utilizing untargeted and targeted metabolomics for non-invasive diagnosis and monitoring of IBD. A cross-sectional study of 101 serum samples was conducted, consisting of 37 UC patients, 23 CD patients, 20 healthy controls (HC), and 21 patients with irritable bowel syndrome (IBS) as disease control (DC) who share the same symptoms but without intestinal inflammation. Global serum fingerprints are relatively quantified using ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UHPLC/Q-TOF-MS). 44 serum metabolites of interest were absolutely quantified with liquid chromatography tandem mass spectrometry (LC-MS/MS) system. The study identified global metabolomic profile shifts in serum differentiates IBD patients from HC, specifically sphingolipid and glycerophospholipid metabolism was perturbed. Receiver operating characteristic (ROC) analysis successfully identified 5-hydroxyindole acetic acid, xanthurenic acid as potential biomarkers for diagnosing UC and CD respectively. One bile acid, glycolithocholic Acid 3-Sulfate, has potential for IBD subclassification. Butyrate was found to reflect bowel symptoms in IBD and IBS patients. The biomarker panel holds promise for clinical use in IBD diagnosis and offers novel insights into IBD pathogenesis.

Keywords

Inflammatory bowel disease, Crohn's disease, Ulcerative colitis, Metabolomics, Tryptophan, Bile acid, Short chain fatty acid, UHPLC/Q-TOF-MS, LC-MS, irritable bowel syndrome

Introduction

The increase in global prevalence of inflammatory bowel disease (IBD) poses a healthcare burden prominent in newly industrialised and Westernised countries[1]. IBD is generalized as chronic prolonged inflammation of the gastrointestinal tract[2] causing functional and structural damage[3]. Its subtypes include ulcerative colitis (UC) and Crohn's disease (CD), whereby inflammation in CD can occur as submucosal lesions in multiple parts of the gastrointestinal tract from mouth to anus but is limited to the colon in UC. Despite our increased understanding

of IBD factors involving interplays between host genetics involving more than 200 loci[4,5], environmental and gut microbiome[6], the heterogeneity of IBD poses challenges to clinicians to establish a single standard test for prognosis, diagnosis, severity monitoring and evaluating treatment response[7]. As the etiology is uncertain, no curative IBD treatments are developed[8] and patients require life-long medication. Accurate stratification and measurement of disease activity are crucial for effective treatment[9]. Under distinct clinical entities, UC and CD require different pharmaceutical treatments. Options of pharmacotherapy have been emerging which are aimed at alleviating different degrees of aberrant immune responses causing lesions for optimal care with personalized medicine[10]. There is an unmet need of biomarkers for different stages of IBD management, as defined by Sakurai and Masayuki[11].

Physicians face several challenges when diagnosing IBD. Current diagnostics include a combination of clinical symptoms, endoscopic, radiologic, and histologic examinations. Though guidelines are proposed to improve histopathological diagnosis[12,13], this could introduce subjectivity and observer variability[14,15], bearing a risk of inaccurate subclassification. For patients, endoscopic procedures are invasive which lowers quality of life and poses burden[3,16], with high cost and a risk of perforation[17]. During maintenance phase, subjective clinical symptoms cannot reliably measure disease activity, Crohn's disease activity index (CDAI) and UCDAI are widely used subjective parameters which can be incoherent with mucosal lesions[18,19]. Also, identifying changes in IBD subclassification is difficult after initial diagnosis. Studies demonstrate that a minority of patients are misdiagnosed without follow-up re-evaluation in five years and are given ineffective treatment[2,20,21], this addresses the importance of close diagnostic monitoring through systematic follow-up, which could reduce the risk of disease-related complications[22]. Laboratory biomarkers potentially help gastroenterologists and patients to non-invasively and objectively measure disease activity for convenient disease monitoring[23]. This study focuses on diagnostic serological biomarkers and the delineation of UC and CD. Blood-based biomarkers are commonly employed due to their non-invasive nature, ease of collection, low susceptibility to contamination, and widespread use[17].

A biomarker is defined as a biological measurement that predict a clinically significant intermediate or endpoint that is harder to observe[24]. An ideal IBD biomarker aims to satisfy the following criteria: being simple, rapid, cost-effective, reproducible, disease-specific[3], and related to disease pathology. To date, there is no ideal biomarker satisfying all requirements. Widely established indicator of disease activity includes erythrocyte sedimentation rate (ESR), fecal

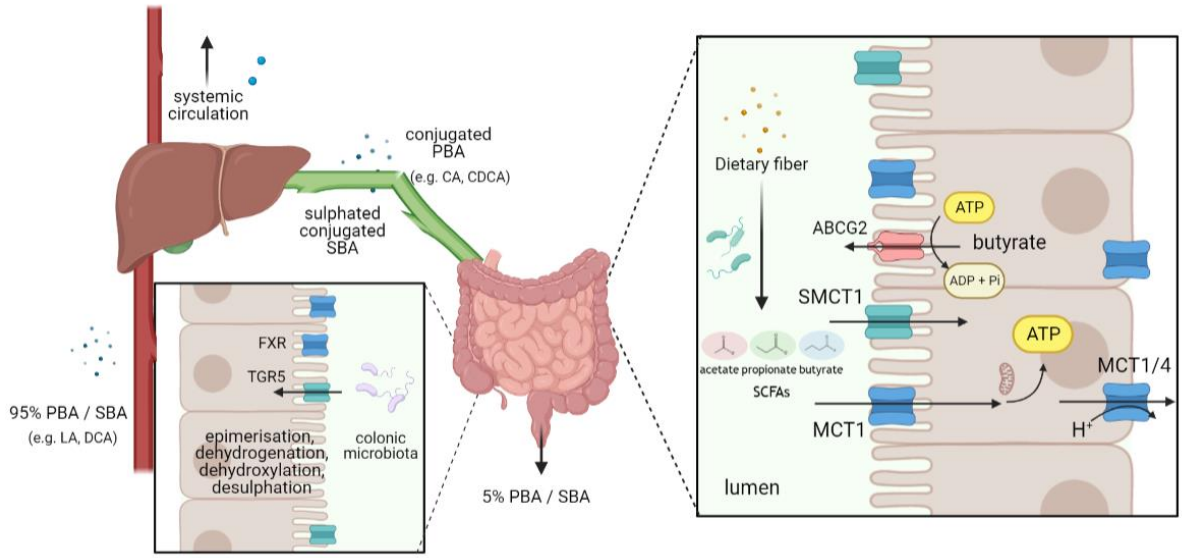
calprotectin (FCP), C-Reactive Protein (CRP) measured in stool and blood samples respectively but are often complicated by patient heterogeneity and with moderate specificity[25,26].

Metabolomics is a high-throughput quantitative analysis of small-molecule metabolites (<25 kilodaltons)[27]. In the context of IBD which has a strong metabolic basis, it provides a comprehensive snapshot that implicate physiological system and interactome from the complex etiology[28], thus emerging as a promising application[29]. This study combines targeted and untargeted metabolomics to obtain a holistic view of key variables associated with IBD pathophysiology for biomarker selection.

Untargeted serum metabolomic studies identified dysregulation in amino acid, lipid, and carbohydrate metabolism[30]. Lipidomics is a branch of metabolomics that annotates all lipid features in an analyte. It has been brought into attention due to their functionality for cell signalling as an inflammatory mediator[31]. Urszula et. al. detected an increase in lactosylceramide18:1/16:0 (LacCer), which uniquely differentiates pediatric CD from UC[32] with excellent performance for differentiating CD when combined with biomarkers ESR, CRP, FCP.

Three key metabolic alterations involved in IBD pathogenesis include tryptophan, bile acids and short chain fatty acids[33]. Here we address the involved pathways, research gaps and the significance of addressing them using targeted metabolomics.

A



B

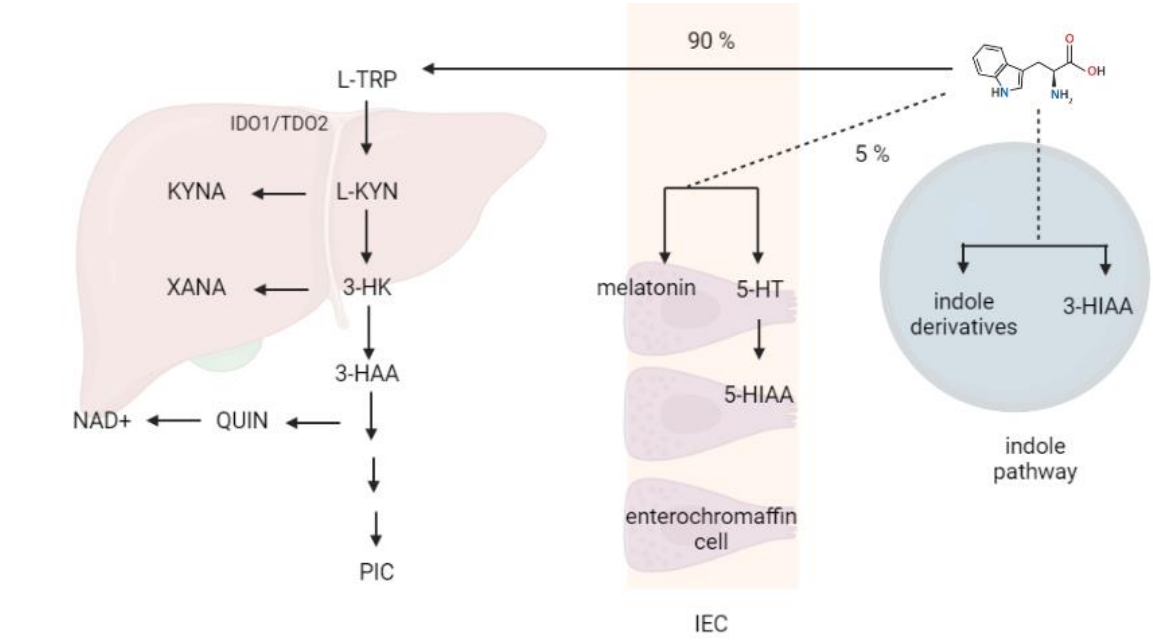


Figure 1. Major altered metabolic pathways in IBD showing (A) BA (left) and SCFA (right); (B) Tryptophan metabolism. Accumulating evidence suggests that losing commensal microbiota induces IBD. ABCG2, ATP binding cassette subfamily G member 2BA, bile acid; CDCA, chenodeoxycholic acid; DCA, deoxycholic acid; FXR, Farnesoid X receptor; IBD, inflammatory bowel disease; KYN, kynurenine; KYNA, kynurenic acid; LA, lithocholic acid; MCT1, Monocarboxylate transporter 1; NAD, Nicotinamide adenine dinucleotide; PICO, picolinic acid; PBA, primary bile acid; QUIN, quinolinic acid; SBA, secondary bile acid; SCFA, short chain fatty acid; SMCT-1, Sodium-coupled monocarboxylate transporter 1; TGR5, G-protein-coupled bile acid receptor; TRP, tryptophan; XANA, xanthurenic acid; 3-HAA, 3 hydroxy-anthranilic Acid; 3-HK, 3-hydroxykynurenine; 5-HIAA, 5-Hydroxyindoleacetic acid; 5-HT, 5-hydroxytryptamine. Illustrations are made in Biorender.

Primary bile acids (PBA), predominantly cholic acid (CA) and chenodeoxycholic acid (CDCA), are steroid molecules synthesized in the liver (Figure 1A). After taurine or glycine conjugation, they are secreted to the intestine as bile salt for lipid emulsification. The majority of BA are reabsorbed into enterohepatic circulation and the remaining 5% are metabolized by colonic microbiota to form secondary bile acids (SBA) like lithocholic acid (LCA) and deoxycholic acid (DCA)[34]. Alongside aiding digestion, both types of BA can act as host signalling molecules in the intestine like farnesoid X receptor (FXR) and G protein-coupled bile acid receptor 1 (TGR5)[35], which regulate glucose and lipid metabolism and modulate inflammation. The significance of LCA and DCA in mitigating inflammation was demonstrated in murine colitis models[36]. Hepatic sulfotransferase-mediated sulphation can act on both conjugated PBA and SBA into less toxic metabolites for removal from circulation. Shifts in BA pool shine lights on its potential role in the pathogenesis of IBD, yet it is not fully understood[34]. Research demonstrated the interrelation of dysbiosis, disrupted SBA production and pathogenesis of IBD. Increased PBA and decreased SBA[37,38] were observed in dysbiosis CD relative to non-dysbiosis stool samples due to loss in SBA-producing bacteria and BA malabsorption[39]. Therefore, serum PBA and SBA levels are worth investigating as an increase in serum SBA was observed in successful anti-TNF therapies for IBD[40,41]. Decreased sulphation of PBA or SBA was observed in IBD patients. In murine models, reduction in sulfomucin content increased intestinal permeability and susceptibility to experimental colitis[42,43,44]. However, one in vitro study showed sulphated SBA opposed the anti-inflammatory effect of SBA, hence its pathophysiological role remains to be addressed.

Short chain fatty acids (SCFAs; mainly acetate, propionate, butyrate) are secondary microbial metabolites produced by fermentation of dietary fiber in the colon, which played crucial roles in energy metabolism and anti-inflammatory actions (Figure 1). Intestinal epithelial cells (IECs) utilize 95% of SCFAs by intracellular transport mainly through transporter MCT1 and SMCT1[45]. Particularly, butyrate acts as a primary substrate for energy metabolism IECs in the TCA cycle and mediates IEC turnover and homeostasis, with its opposing effect in suppressing cancerous colonocytes[46,47], given IBD patients exhibit colorectal cancer risk five times higher than normal. Also, SCFAs are ligands that activate downstream G-protein-coupled receptor (GPCR) pathway and regulate anti-inflammatory signalling cascades and immune cell recruitment[48], cytokine, chemokine production and epigenetic regulation[49]. Therefore, it is one of the therapeutic targets through prebiotic and probiotic treatment. Several cohort studies documented a decrease in fecal SCFA levels and ratios of different SCFAs in both UC and CD. This reduction is supported by the downregulation of SCFA transporter, especially MCT1 expression[50], fewer SCFA-

producing bacteria such as F. prausnitzii and A. muciniphila[51], and impaired butyrate oxidation. However fecal SCFA only represents 5% and cannot reflect host utilization[52]. To our knowledge, serum SCFA level has not been investigated in IBD patients. Though the liver clears most of the butyrate and propionate from portal circulation[53] for gluconeogenesis, a portion still enters systemic circulation for all three SCFAs and are present in peripheral tissues[51]. Research describing their systemic roles in the context of IBD are few, and worth investigating.

Tryptophan, one of the essential amino acids, is metabolized by gut microbes, immune cells, and intestinal epithelium, where its derivatives play crucial roles in modulating intestinal inflammation, energy homeostasis and gut barrier integrity (Figure 1B). In IBD, all three major tryptophan metabolic pathways were altered: indole, serotonin /5-hydroxytryptamine (5-HT) and the kynurenine pathway (KP)[59]. 95% of absorbed tryptophan undergoes KP catalyzed by hepatic tryptophan 2,3-dioxygenase 2 (TDO2) and mucosal indoleamine 2,3-dioxygenase 1 (IDO1) to convert to downstream metabolites (Figure 1), which involves in biological processes related to immunoregulation and neurotransmission. The remaining 5% was absorbed by enterochromaffin cells, undergone 5-HT pathway to produce serotonin and is degraded into 5-hydroxyindole acetic acid (5-HIAA) as the end product before entering systemic circulation[59]. Experimental colitis models demonstrated the critical role of 5-HT in IBD pathogenesis including immune cell infiltration, enhance proinflammatory cytokine production and disease severity[54]. Unabsorbed tryptophan undergoes bacterial degradation into indole by-products; those derivatives act on aryl hydrocarbon receptor (AhR), which exerts an anti-inflammatory effect and tissue repair through the IL-22 cascade[55]; and indole-3-propionic acid receptor for barrier function and inhibit TNF production. Small-scale studies suggested serum tryptophan as a potential biomarker [56,57,58] due to its negative correlation with disease activity, measured by CRP. A large cohort study from Nikolaus et al. quantitated serum levels of downstream KP derivatives and identified increased KP activation[59] due to the upregulation of IDO. This is evidenced by a higher tryptophan to kynurenine ratio and neurotoxic quinolinic acid level, which is more pronounced in CD, whilst lower kynurenic acid and picolinic acid levels were observed only in CD when compared to healthy patients. Hence, tryptophan metabolites are potential biomarker candidates which remain to be validated. Rewiring tryptophan metabolism can influence clinical outcomes demonstrated in the murine colitis model and in post-therapeutic CD patients with ameliorated levels post-therapy, which infers its significance in IBD.

Thus, the main project objective was to explore and select serum metabolites as potential diagnostic biomarkers candidates for IBD. For sample subjects, a cross-sectional study of an

IBD cohort was recruited including 4 subtypes: UC, CD, healthy control (HC) and disease control (DC), which are patients with irritable bowel syndrome (IBS). IBS patients exhibit bowel symptoms but without inflammation. This control is included to identify whether selected biomarkers is related to the pathophysiology of IBD in regulating intestinal inflammation. Patient serum samples were collected from the IBD cohort to perform global and targeted metabolomic profiling. To accomplish the objective, multivariate statistical analysis was performed on both untargeted and targeted metabolomics datasets to select metabolites from a mathematical model. Metabolic profiles of HC are separately compared with UC or CD, and between UC and CD. For untargeted exploratory analysis, potential novel biomarkers and dysregulated metabolic pathways are revealed. In addition, the forementioned diagnostic biomarker candidates from previous studies are validated.

For targeted metabolomics approach, the defined metabolite set of interest contains 44 candidates after screening away metabolites with out-of-range measurements, which was based on alterations of tryptophan, bile acid and SCFA metabolism (Supplementary Figure 1). To evaluate the robustness of metabolite changes; diagnostic performance is measured by sensitivity and accuracy. It was hypothesized that specific metabolites or metabolic pathways in the patient's serum can serve as reliable biomarkers for disease diagnosis, activity, and response to therapy. Furthermore, the combination of these targeted metabolites with clinical parameters can enhance the accuracy and specificity of IBD diagnosis. Specifically, the metabolite concentration threshold can reveal metabolites associated with disease phenotypes and pathology.

Materials and methods

Ethics statement

The study was approved by U.K. NHS Health Research Authority West London National Research Ethics Committee, in accord with the principles in the Declaration of Helsinki. All participants provided informed consent.

Study cohort

101 patients were enrolled into the study consecutively in National Institute for Health Research (NIHR) Imperial Clinical Research Unit Facility at Hammersmith Hospital (London, UK) from April 2018 to October 2020. Patients are classified into 4 clinical subgroups, with either UC

(n=37), CD (n=23), with healthy control (n=20) and IBS patients (n=21) for comparisons. IBD patients fulfilled key eligibility criteria: female or male aged between 18 and 75 years, with diagnosis of IBD. Key exclusion criteria encompassed medications and comorbidities that potentially impede the study's safety and validity. Clinical demographics, disease subclassification, and disease activity were recorded as indicated by FCP. Whole blood samples were collected from patients during a fasting state by Dr Shiva Radhakrishn and Dr Horace Williams. Serum samples were prepared after centrifuging for metabolomic profiling.

Metabolomics Analysis

Metabolomics techniques were conducted by Dr Maria Garcia at MRC-NIHR National Phenome Centre. Untargeted metabolomic profiling of 101 serum samples were analysed with ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UHPLC/Q-TOF-MS). Analytical details including sample preparation and quality control procedures are included in previous publications[60].

UHPLC/Q-TOF-MS assays were used to capture a wide metabolite class (Supplementary Figure 2A). For global profiling, UHPLC/Q-TOF-MS assays were customized for lipophilic analyte separation utilizing reversed-phase chromatography (RPC) and lipid-phase chromatography (LPC) assays. Hydrophilic analytes were separated using hydrophilic interaction chromatography (HILIC). Columns were coupled to positive and negative ion mode electrospray ionization (ESI+/-) to collect the following datasets: RPC+, RPC-, LPC+, LPC- and HILIC+. For targeted measurement, LC-MS/MS assays were performed for bile acids, SCFA and tryptophan derivatives (Supplementary Figure 2B). Metabolite identities were confirmed using internal standards; concentrations were measured in umol / L.

Raw readings from each assay were passed into PeakPantheR package in R software for background noise removal, peak detection, alignment, and quality control (QC) processing. Each data value represents quantified area under curve (AUC) for each annotated metabolite peak.

Statistical analysis

Separate databases were built for each metabolomic assay. 6 patients with missing data were removed with resulting sample size n = 96. For each annotated metabolite, autoscaling was applied column-wise by subtracting individual values from cohort mean and dividing by standard

deviation, to account for differences between fold changes of each metabolite. Datasets were merged according to metabolomic assay.

Multivariate plots and calculations were computed with SIMCA 17 (Sartorius, Umeå, Sweden). For data overview and quality check, Principal component analysis (PCA), an unsupervised clustering method, was applied. To discriminate between samples, Orthogonal partial-least-squares-discriminant analysis (OPLS-DA) was applied and autofitted as a supervised analysis to yield group-predictive metabolite factors. Model was evaluated with cumulative regression (R2X) and performance statistic (R2Y). To ensure the validity of each OPLS-DA model, 7-fold cross-validation was employed by partitioning data into seven subsets, then iterating and testing on different combinations for model training. Permutation tests against 200 null models was performed. Only the first two components were selected for analysis to prevent overfitting. Analysis Of Variance testing of Cross-Validated predictive residuals (CV-ANOVA) were applied for significance testing. Metabolites with variable importance in projection (VIP) higher than 1.0 were considered significant for driving separation.

Univariate analysis was applied only to the targeted dataset to sort statistically significant metabolites, but not relatively quantified. Non-parametric Kruskal Wallis test was performed to compare multiple subgroups, with statistical significance set at 95% level ($p < 0.05$). The Receiver Operating Characteristic curve was plotted for selected potential biomarkers. Plots and calculations were done using Graph-Pad Prism v9.5.0. Correlations with demographics were assessed with Chi-squared test (data not shown).

Metabolic Pathway analysis

For an untargeted metabolomics approach, the selected annotated differential metabolite features were tagged with HMDB and PubChem as chemical identifiers as input data type. Compound lists were uploaded to web-based tool MetaboAnalyst 5.0[611] to test if serum metabolomic profiles shifts between groups match any metabolic pathways. Pathway analysis module was selected and visualized using a scatter plot. This module integrates enrichment analysis and pathway topology analysis which can measure the centrality of a metabolite throughout the metabolic network. Node importance was measured with relative betweenness centrality, which focuses on global network topology[622]. Homo sapiens pathway library was

[1] https://academic.oup.com/nar/article/49/W1/W388/6279832
[2] https://pubmed.ncbi.nlm.nih.gov/16880171/

selected, which contains 80 metabolic pathways from Kyoto Encyclopedia of Genes and Genomes (KEGG) database updated since Oct 2019.

Result

Study population

The study design is illustrated in Figure 2A. After excluding patients with missing datasets, cohort subjects were reduced to n=96. Clinical demographics are summarized in Table 1. To measure the landscape difference among IBD patients, OPLS-DA model was built in discriminating between two groups based on sex and ethnicity with significance, which align with previous findings[63,64]. Categorical demographics are independent with non-significant correlation based on Chi-squared test (data not shown).

T = inception	All (n = 96)	HC (n = 20)	UC (n = 36)	CD (n = 22)	DC (n = 18)
Age(median, range)	34 (18-75)	34 (26-54)	35.5 (18-75)	30.5 (18-66)	28 (18-59)
Gender (M,F)	51:45	12:8	16:20	13:9	10:8
Caucasian ratio (%)	63:33 (65.6)	12:8 (60)	25:11 (69.4)	11:11 (50)	15:3 (83.3)
DAI (median, range)	-	-	7.5 (0-15)	7 (1-11)	-
Diet (meat : veg)	73	15:5	26:10	16:6	16:2
Disease subtype	-	-	Left-sided 15 Pancolitis 14 Proctitis 7	Colonic 3 Colonic* 1 Ileocolonic 6 Ileocolonic* 3 Terminal ileitis 9	Constipation 7 Diarrhoea 4 Post-infection 2

Table 1. Baseline characteristics of the IBD patient cohort. Patient's age, gender, ethnicity, disease activity index (DAI), dietary habits and disease subdifferejntation were recorded. Data are present using bracketed parameters. Abbreviations: CD, Crohn's disease; DC, disease control; HC, healthy controls; IBD, inflammatory bowel disease; T, timepoint; UC, ulcerative colitis; veg, vegetarian; *: with perianal disease.

Untargeted metabolomics identified alterations in amino acid and lipid metabolism in IBD patients

Multivariate analysis was conducted on an untargeted metabolomic dataset, to capture the overall serum metabolic profile changes of 96 patients unbiasedly. PCA and OPLS-DA was performed on 575 annotated metabolites to examine metabolic shifts between patient groups. From the PCA scores plot, all clusters grouped by disease type overlap with no significant outliers identified nor excluded (Figure 2B). The first two principal components (PC) explain 16.4% and 11.0% of the overall variation of the model respectively. From visual inspection, data

points were dispersed without distinct clusters even when color-coded by categorical clinical observations (Supplementary 3A-C). OPLS-DA was performed on the same dataset to discriminate between two selected classes. Cross-validation plots visually showed significant separation of UC and CD patients from HC and is evidenced by CV-ANOVA. Both models have good predictive ability and overall fit to data, evidenced by Q^2 and R^2 values (Figure 2C-F). Permutation testing for both models reflect an appropriate model fit (Figure 2G-H).

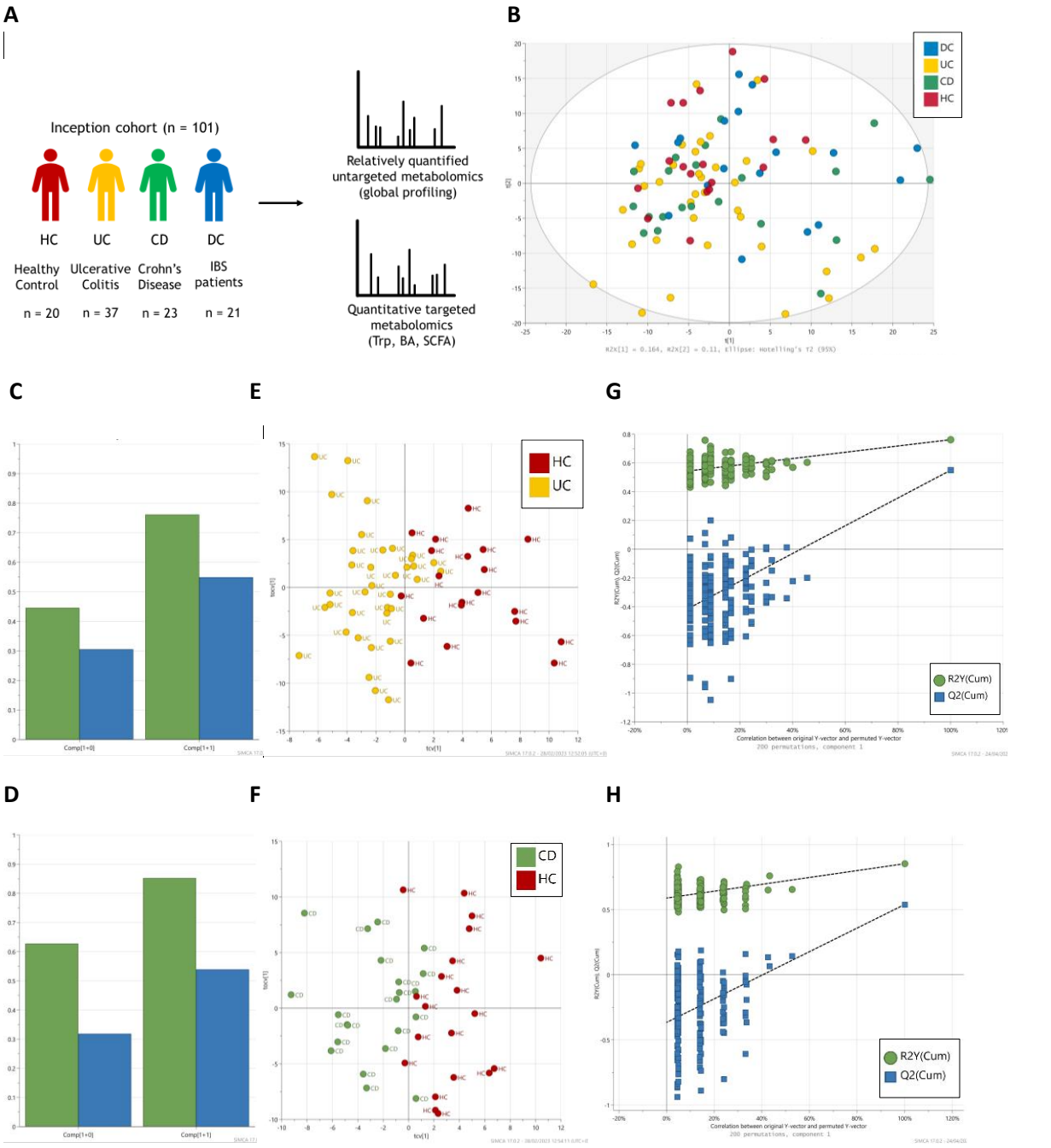


Figure 2. Differences in serum metabolomic fingerprints in UC and CD patients compared to HC. (A) Design of IBD inception cohort (n=101). Data was extracted from single time point. Multivariate statistical analysis was applied to untargeted dataset. (B) PCA scores plot of first two PC based on 575 annotated metabolites from patient serum samples (n=96). R^2Xcum= 0.275, Q^2cum=0.214. OPLS-DA models were built to UC or CD patients from HC. Summary of fit was computed based on first two PC when comparing (C) HC with UC patients (n=56).; or (D) HC with CD patients (n=42). CV scores plot differentiates HC from (E) UC (R^2Xcum= 0.198, R^2Ycum= 0.762, Q^2cum=0.549, CV-ANOVA=2.221 x 10^{-8}); (F) CD (R^2Xcum= 0.149, R^2Ycum= 0.852, Q^2cum=0.539, CV-ANOVA=6.521 x 10^{-6}). 200-time permutation plots of OPLS-DA models against null model distribution between (G) HC and UC patients, (H) HC and CD patients. Green and blue dots respectively represent R2 and Q2 values. Dashed lines represent regression lines for observed and permutated R2 and Q2. Abbreviations: ANOVA, analysis of variance; CD, Crohn's disease; Cum, cumulative; CV, cross validation; DC, disease control; HC, healthy controls; IBD, inflammatory bowel disease; OPLS-DA, Orthogonal Projections to Latent Structures Discriminant Analysis; PC, principal components; PCA, principal component analysis; UC, ulcerative colitis.

Serum fingerprints from UC and CD patients can be significantly differentiated from DC (Supplementary Figure 4B), implicating a pathophysiological effect of intestinal inflammation in IBD. No models can be generated for subdifferentiation between UC and CD patients, and between HC and DC.

To select the major variables contributing to the separation, metabolites are ranked and selected based on a threshold VIP score > 1.0 according to the literature. Among 575 annotated metabolites, 39.7% and 36.7% were significantly altered comparing UC and CD with HC respectively (Figure 2A-B). The metabolite subclasses are identified based on the LIPID MAPS database. Both comparisons between UC and CD patients to HC have similar proportions of altered metabolite classes, where glycerophospholipids take up the highest proportion, followed by sphingolipids. Comparisons between CD and HC showed higher proportions of altered fatty acyls, carboxylic acid derivatives and glycerolipids contributing to metabolic perturbation.

To identify the shifts in relative abundance, a more stringent threshold of VIP >1.5 was used to select the top metabolites (Figure 2C-D). A similar proportion of metabolites are screened out (~10%). There is a general decrease in serum glycerophospholipids and sphingolipids that is prominent in UC patients. Comparatively, there is a higher variety of metabolite classes that are altered in CD patients. While a single trend of decrease in glycerophospholipids can be observed in CD, different annotated sphingolipids and fatty acyls are shifted differently.

To test how the metabolic profile's shifts match metabolic pathways in UC and CD compared to HC, differential metabolites with VIP > 1.0 are incorporated into pathway analysis respectively for mechanistic interpretation with exclusion of non-endogenous compounds (Figure 2E-F). For both comparisons, top dysregulated metabolic pathways were relevant to sphingolipid, glycerophospholipid and amino acid metabolism including D-Glutamine and D-glutamate, alanine, aspartate and glutamate. In addition, perturbations in arginine biosynthesis and ascorbate and alderate metabolism are observed in UC patients only, while retinol and caffeine metabolic shifts are observed only in CD.

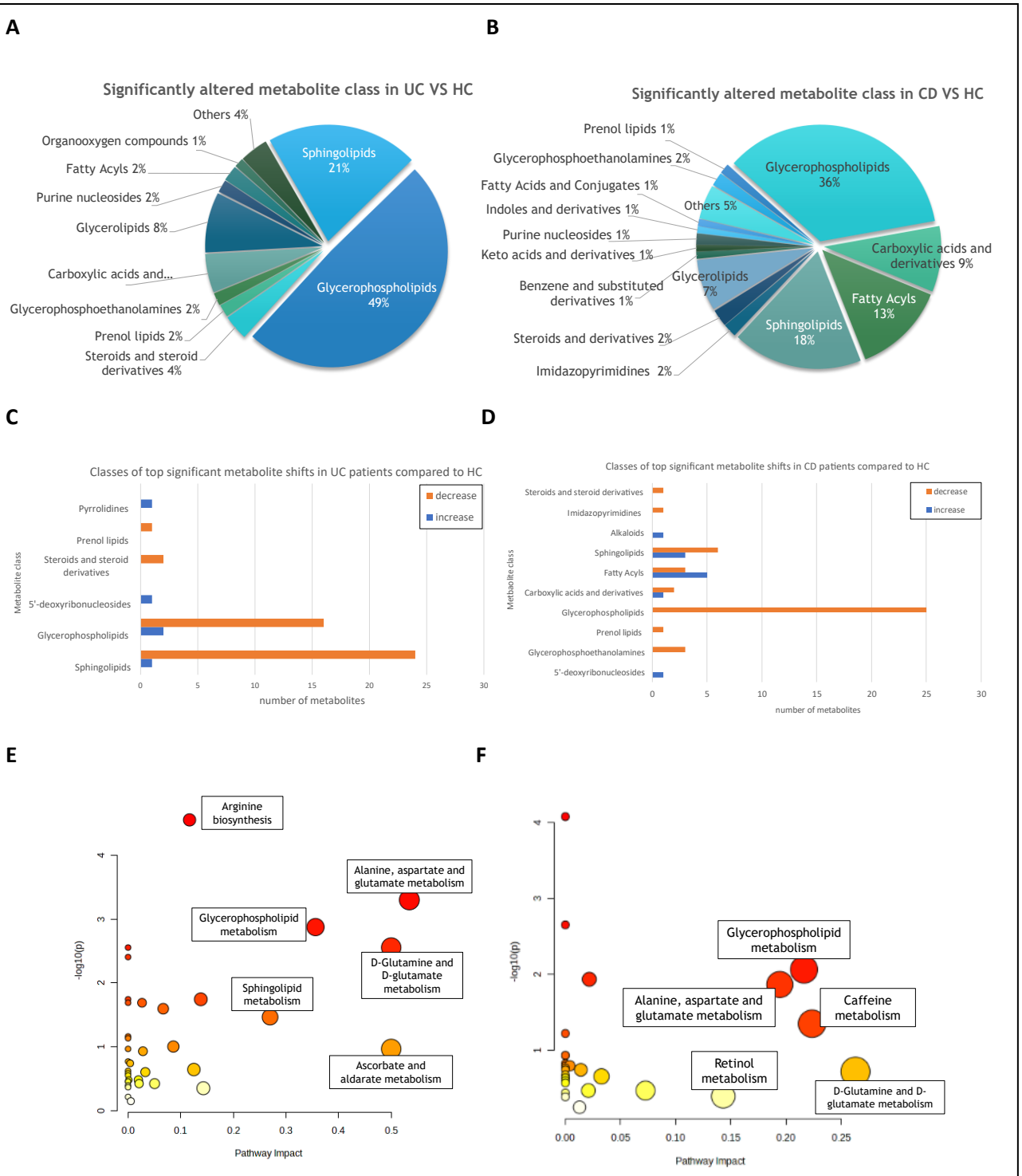


Figure 2. Altered metabolism discriminating UC and CD patients from HC. Based on VIP score > 1.0 from OPLS-DA model, major altered metabolite classes are identified for comparing (A) UC and HC (B) CD and HC, respectively selects 228 and 211 significant metabolites out of 575 annotated variables. The percentages represent the proportion among the significant metabolites. Based on VIP score >1.5, direction of shift describes IBD patients when comparing (C) UC and HC (D) CD and HC, respectively selects 52 and 60 metabolites. Metabolome view of pathway analysis generated from Metaboanalyst 5.0 comparing (E) UC with HC (F) CD with HC. Plots show matched pathways based on p values from pathway enrichment analysis and pathway impact values in topology analysis. Larger node size and colour intensity corresponds to higher degree of fold enrichment and higher significance per metabolic pathway, measured by p value. Y-axis represents fold change. Abbreviations: CD, Crohn's disease; DC, disease control; HC, healthy controls; IBD, inflammatory bowel disease; OPLS-DA, Orthogonal Projections to Latent Structures Discriminant Analysis; UC, ulcerative colitis; VIP, variable influence on projection.

Targeted metabolomics differentiates IBD patients from healthy patients

To observe and identify potential co-regulated metabolite patterns, targeted metabolomics datasets, multivariate exploratory analysis was performed investigating 41 metabolite parameters after excluding out-of-range measurements (citrulline, serotonin, valerate) (Supplementary Figure 1). 22 BAs, 8 SCFA and 11 tryptophan derivatives were included. Based on PCA scores plot, one IBS patient was excluded (Figure 3A). OPLS-DA model shows discrimination of metabolite shifts in UC and CD compared to HC (Figure 3, Supplementary Figure 3C). Significant metabolites based on VIP > 1.0 was selected; 16 and 20 metabolites were plotted in the coefficient plot. Positive coefficients represent the predictability of HC based on respective metabolite concentration, and a larger coefficient implies higher importance for classification between groups. Also, significant separation could be identified based on OPLS-DA model comparing HC and DC (Supplementary Figure 4C). Since models were built for without high performance < 0.5, they should be interpreted with caution.

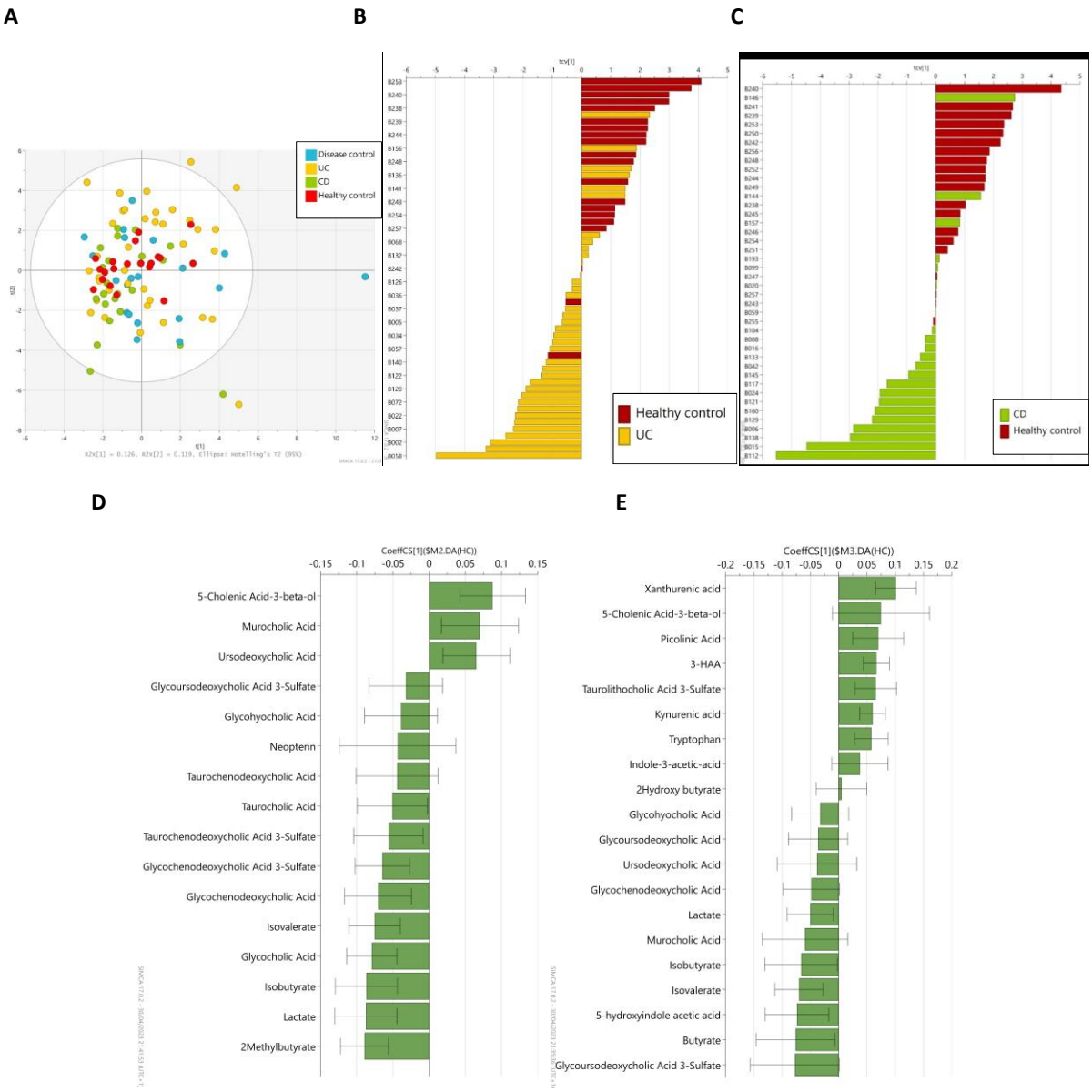


Figure 3. Significant metabolites from targeted dataset differentiating UC and CD patients from HC. (A) PCA scores plot of first two PC based on 41 targeted metabolites from patient serum samples (n=96). R^2Xcum= 0.245, Q^2cum=0.0374. OPLS-DA models were built to differentiate UC or CD patients from HC. After autofitting, one PC was fitted on CV scores plot for differentiating (E) HC from UC patients (n=56). R^2Xcum= 0.126, R^2Ycum=0.485, Q^2cum=0.39, CV-ANOVA=1.61369 x 10^{-6}; (C) HC from CD patients (n=42). R^2Xcum= 0.152, R^2Ycum=0.509, Q^2cum=0.376, CV-ANOVA=0.000103. Based on VIP score > 1.0 from OPLS-DA model, significant metabolites are arranged in coefficient plot. Positive coefficients represent prediction towards HC against (C) UC or (D) CD. Abbreviations: ANOVA, analysis of variance; CD, Crohn's disease; cum, cumulative; CV, cross validation; DC, disease control; HC, healthy controls; IBD, inflammatory bowel disease; OPLS-DA, Orthogonal Projections to Latent Structures Discriminant Analysis; PC, principal component; UC, ulcerative colitis; VIP, variable influence on projection.

Increase in serum primary bile acid level in IBD patients

We further performed univariate analysis on each targeted serum metabolite, which includes 22 bile acids, 8 SCFA and 11 tryptophan derivatives. Biomarker performance was evaluated using an ROC curve, AUC > 0.8 is considered robust with sufficient sensitivity and specificity.

When comparing to HC, 6 out of 11 PBA showed a general increase in patients with UC or DC but not in CD, except 5-Cholenic Acid-3-beta-ol which was found lower in both UC and CD patients (Supplementary 4A-C). While no SBA showed significant changes, 2 sulphated SBA, Glycolithocholic Acid 3-Sulfate and Taurolithocholic Acid 3-Sulfate, were decreased in DC whereas Glycoursodeoxycholic Acid 3-Sulfate was increased in CD. For biomarker evaluation, none showed satisfactory performance in the ROC curve, but it is worth noting that the Glycolithocholic Acid 3-Sulfate was close to the threshold (AUC = 0.78) (Figure 4A-B). Its potential and pathophysiological significance requires validation.

Increase in serum butyrate associates with bowel symptoms

Among eight SCFA metabolites, three showed potentials as IBD biomarkers. Though serum acetate and propionate level were not associated with IBD or IBS state (Supplementary), UC patients showed an increase in lactate and 2-methylbutyrate compared to HC (Figure 4C, E, G). Butyrate was increased in both IBD patients and DC and can interpret its correlation with bowel symptoms, but its significance is more pronounced in IBD. Comparison of DC with UC or CD does not infer significance. IN ROC analysis, only butyrate showed good performance when separately differentiating HC from UC and CD, and close to the threshold for DC (AUC = 0.82, 0.83, 0.78) (Figure 4D, F, H). Therefore, butyrate may be involved in common mechanisms in both IBD and IBS patients.

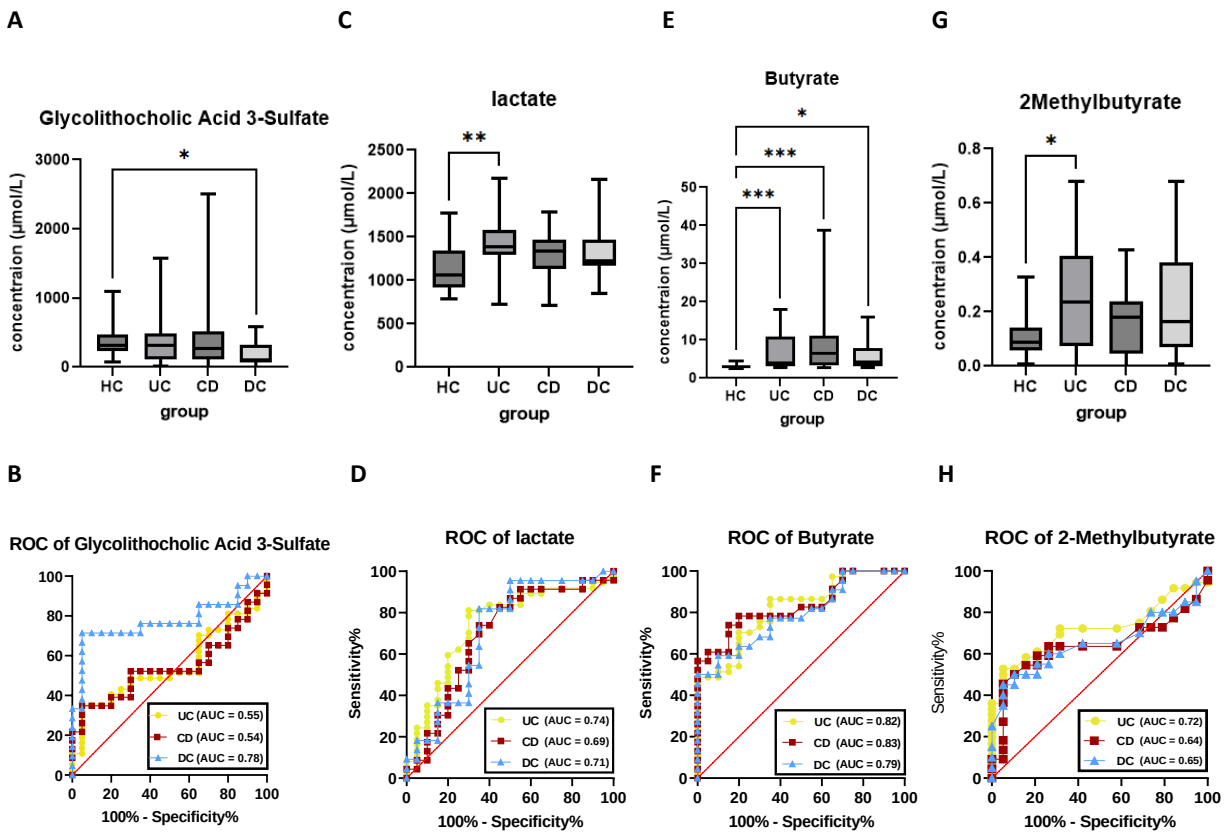


Figure 4 Biomarker performance of significantly altered BA and SCFA in targeted metabolomics. Univariate Kruskal Wallis test was performed on targeted BA and SCFA dataset. Boxplot and biomarker evaluation of ROC of (A-B) Glycolithocholic Acid 3-Sulfate, (C-D) lactate (E-F) butyrate and (G-H) 2-methylbutyrate. Threshold of biomarker performance was set at AUC = 0.8. Abbreviations: AUC, area under curve; BA, bile acid; CD, Crohn's disease; cum, cumulative; DC, disease control; HC, healthy controls; IBD, inflammatory bowel disease; ROC, Receiver operator characteristic; UC, ulcerative colitis; SCFA, short chain fatty acid. *P < 0.05; **P < 0.01; ***P < 1 x 10^{-3}.

5-HIAA and xanthurenic acid differentiates UC and CD from healthy controls

For tryptophan derivatives, serum tryptophan level showed no significance, whereas the kynurenine:tryptophan ratio was elevated in UC compared to DC only, whilst not significant compared to HC (Supplementary 4A-B). Kynurenine-related metabolites, including kynurenic acid, xanthurenic acid and picolinic acid, were reduced in CD patients in separate comparisons with and with UC (Figure 4C-D, Supplementary 4C-F), while UC patients do not show significance compared to HC. While 5-HT was out of measurement range, 5-HIAA was notably higher in IBD patients compared to HC. Though no significance was shown for Indole-3-acetic-acid, for the indole pathway, neopterin, an inflammatory biomarker of macrophage activation, was shown higher in UC relative to HC. Biomarker evaluation showed that 5-HIAA and

xanthurenic acid are robust with AUC = 0.86 and 0.83 in diagnosing UC and CD patients respectively (Figure 4-D). Picolinic acid also has potential in differentiating CD from both HC, with AUC close to the threshold (AUC = 0.78).

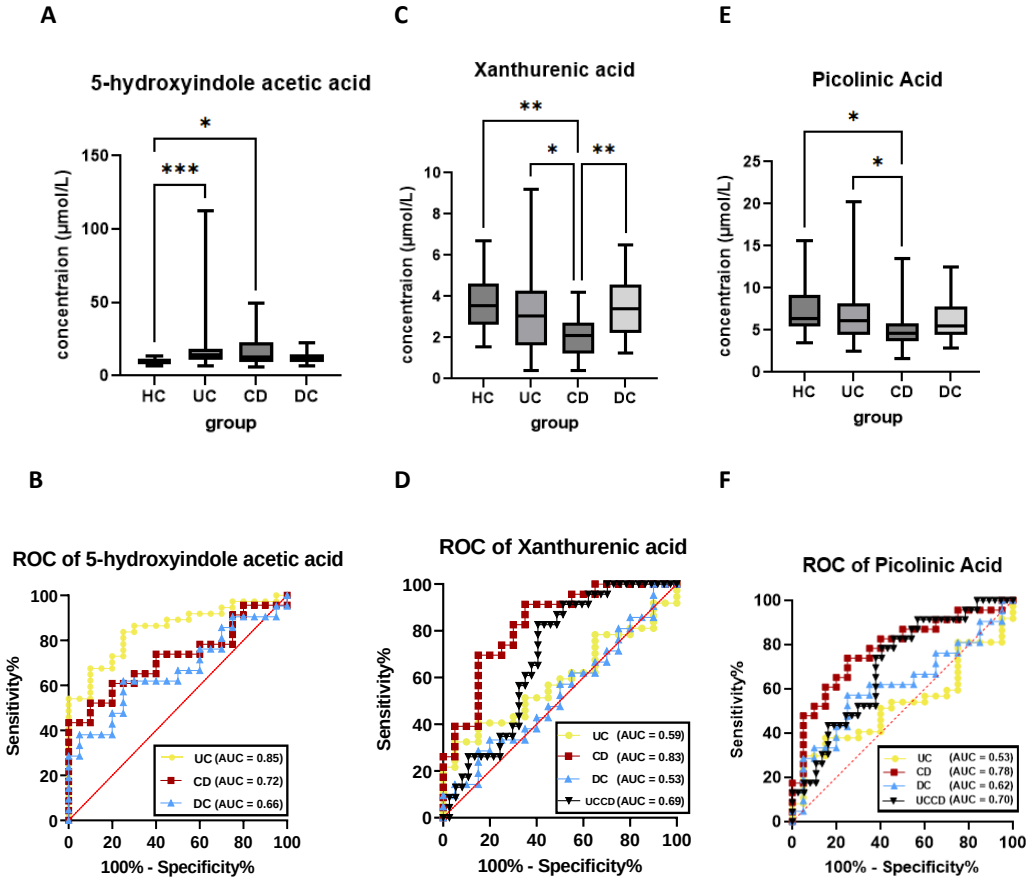


Figure 5 . Evaluation of TRP derivatives as a diagnostic biomarker for IBD. Univariate Kruskal Wallis test was performed on targeted tryptophan dataset. Boxplot and biomarker evaluation of ROC of (A-B) 5-HIAA, (C-D) xanthurenic acid, and (E-F) picolinic acid were shown. Threshold of of biomarker performance was set at AUC = 0.8. Abbreviations: AUC, area under curve; CD, Crohn's disease; DC, disease control; HC, healthy controls; IBD, inflammatory bowel disease; OPLS-DA, Orthogonal Projections to Latent Structures Discriminant Analysis; ROC, Receiver operator characteristic; UC, ulcerative colitis; VIP, variable influence on projection; 5-HIAA, 5-hydroxyindole acetic acid; *P < 0.05; **P < 0.01; ***P < 1 x 10^{-3}.

Discussion

To our knowledge, this is the first hybrid approach that combines targeted and untargeted metabolomics with MS-based superior sensitive analytical platform[65], also with the inclusion of IBS patients that helps to characterize IBD biomarkers in its relation to IBD pathophysiology. This study characterized shifts in serum metabolomic profile associated with UC and CD patients but not IBS. Contrary to Elizabeth et al. where only 5 significant metabolites were identified in serum global UC fingerprints, we identified a larger number of significant metabolites in UC with similar degree shifts of comparison between CD and HC, potentially implicating the significance of cohort size. Subdifferentiation of UC and CD cannot be determined by mathematical modelling; more work is required to identify the discriminatory factor in delineating UC and CD.

L-Arginine is a semi-essential amino acid. Its depletion in UC patients supports its critical role in the pathogenesis of IBD. With the increased enzymatic activity of NOS2 and Arg1, intraluminal L-Arginine competes for downstream effects, which include conversion into nitric oxide that triggers anti-microbial activity and stimulation of mucus production. Another pathway involves increased production of polyamines causing immune diminution. Its controversial role remains to be explored. Alteration in retinol metabolism was found in CD only. Dysregulation in vitamin A metabolism is more common in CD, which is responsible for regulating epithelial proliferation and differentiation, thus explaining the higher prevalence of hypovitaminosis in CD compared to UC[66]. Clinical trials showed the significance of vitamin supplementation in ameliorating inflammation[67]. As one of the potential targets for UC and CD subdifferentiation, the exact mechanism of vitamin A metabolism in CD remains to be elucidated. Alteration in caffeine metabolism was found in CD patients. Caffeine was shown to play opposite roles in UC and CD: it dampens inflammation in experimental colitis murine models[68], but high caffeine intake could aggravate CD symptoms[69]. This can be observed in IBS patients where high caffeine intake (400mg/day) positively associates with intestinal symptoms. Altered caffeine metabolism may serve as a potential biomarker for Crohn's disease. Identified profile shifts also revealed some signatures of IBD targets, including an increase in sphingosine-1-phsophate and 5-methyl adenosine (data not shown), implicating the significance of metabolites in the application of targeted therapy such as Sphingosine-1-Phosphate Receptor Modulators that reduced circulating lymphocytes.

This study pushes forward from existing literature in characterizing serum BA profile with increased PBA and sulfated SBA concentration in both IBD and IBS, however, the concentration difference is not enough to establish significance for clinical application. The enterohepatic circulation of BA limits its utility upon endogenous synthesis, reabsorption, and clearance, where a limited amount is detected in systemic circulation. Non-significant changes in SBA level indicates its limited diagnostic application to infer gut dysbiosis, since serum mainly reflects host metabolism only and its non-specificity in identifying IBD such as ^{75}SeHCAT test in scanning chronic diarrhea based on BA malabsorption[70]. Monitoring gut dysbiosis is better indicated from stool samples, such as using SBA for monitoring local immune cell recruitment.

This study has discovered serum butyrate as a good biomarker which can infer bowel symptoms. While studies showed reduced abundance in butyrate-producing bacteria, MCT1 and SMCT1 receptors in both IBD and IBS, we hypothesized its increase is associated with reduced expression of ATP-dependent efflux pump for exporting butyrate from colonocytes to lumen[71]. Precise mechanisms underlying the basolateral transport to systemic circulation are still an active area of research.

Contrary to previous studies, serum tryptophan reserves are not significantly depleted[56]. This implies the difference in pathway activation of tryptophan between quiescent and active IBD, where the reduction is prominent only in active CD in Nikolaus' study[59]. This further supports tryptophan as a potential indicator that reflects disease activity. We could not detect alteration in serum kynurenine to tryptophan ratio compared to HC, in which the parameter accounts for inter-variation in dietary tryptophan uptake. This can be attributed to disease state and circadian control of IDO1 expression, causing tryptophan fluctuations[72]. Decreased picolinic acid levels in CD were also observed in a larger IBD cohort. Although its role is poorly understood[73], it supports its neuroprotective role against quinolinic acid kainic acid pathway inducing depression[74]. e hypothesis as in chemokine production and macrophage activation by interacting with interferon-gamma[75]. Here we validated alteration in Kyp downstream metabolites, where results for kynurenic acid and picolinic acid are similar to previous studies, though increase in quinolinic acid cannot be detected. Here we highlight potential of 5-HIAA as diagnostic biomarker related to tryptophan-induced depression through gut-brain axis[76], although with limited specificity as an indicator for carcinoid tumors in urine, its level associates with pathogenesis of lymphocytic colitis[77]. Future studies should stratify IBS patients as one study found increased 5HIAA level in IBS-diarrhea but not IBS-constipation[78].

In conclusion, this study provided a comprehensive view of important metabolic changes and tryouts in the subdifferentiation of UC and CD. Regarding future work, identities of annotated metabolites from untargeted metabolomics could be confirmed using LC-MS/MS for absolute quantification of lipid biomarkers and evaluation. For limitations, mathematical models built in this study were tailored to identify metabolic features in Caucasians only, given that genetic predisposition from different ethnicities can influence the risk of IBD and phenotype. A larger and more diverse cohort size could compensate for sample bias. While this cross-sectional study cannot infer a causal relationship between metabolite features and pathophysiology, only the individual variation of the metabolome can be identified. A longitudinal study could provide valuable insights for identifying risk factors by tracking one's metabolic profile, amd validating previously identified biomarkers and discover their potential in predicting treatment response. Studying other biological matrices (i.e. stool, urine, colonic tissue) is desired in future IBD studies to understand metabolic changes across biological compartments, specifically colonic microbiota metabolism

Acknowledgements

I would like to thank Dr Jia Li, Dr Maria V. Garcia for their guidance and advice; Dr Li's lab group for invaluable support; patients from Hammersmith Hospital for offering samples; Dr Shiva Radhakrishn and Dr Horace Williams for study design and sample collection, and the National Phenome Centre for providing the dataset.

Reference

1. Alatab S, Sepanlou SG, Ikuta K, Vahedi H, Bisignano C, Safiri S, et al. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: A systematic analysis for the global burden of disease study 2017. The Lancet Gastroenterology & Hepatology. 2020;5(1):17–30.
2. Ananthakrishnan AN. Epidemiology and risk factors for IBD. Nature Reviews Gastroenterology & Hepatology. 2015;12(4):205–17.
3. Ford AC. Functional gastrointestinal disorders in inflammatory bowel disease: Time for a paradigm shift? World Journal of Gastroenterology. 2020;26(26):3712–9.
4. Orholm M, Munkholm P, Langholz E, Nielsen OH, Sørensen TIA, Binder V. Familial occurrence of inflammatory bowel disease. New England Journal of Medicine. 1991;324(2):84–8.
5. de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. Nature Genetics. 2017;49(2):256–61.
6. M'Koma A. The multifactorial etiopathogeneses interplay of inflammatory bowel disease: An overview. Gastrointestinal Disorders. 2018;1(1):75–105.
7. Vermeire S. Laboratory markers in IBD: Useful, magic, or unnecessary toys? Gut. 2006;55(3):426–31.
8. Elhag DA, Kumar M, Saadaoui M, Akobeng AK, Al-Mudahka F, Elawad M, et al. Inflammatory bowel disease treatments and predictive biomarkers of therapeutic response. International Journal of Molecular Sciences. 2022;23(13):6966.
9. Carter MJ. Guidelines for the management of inflammatory bowel disease in adults. Gut. 2004;53(suppl_5):v1–v16.
10. Cai Z, Wang S, Li J. Treatment of inflammatory bowel disease: A comprehensive review. Frontiers in Medicine. 2021;8.
11. Sakurai T, Saruta M. Positioning and usefulness of biomarkers in inflammatory bowel disease. Digestion. 2022;104(1):30–41.
12. Stange EF, Travis SPL, Vermeire S, Reinisch W, Geboes K, Barakauskiene A, et al. European evidence-based consensus on the diagnosis and management of ulcerative colitis: Definitions and diagnosis. Journal of Crohn's and Colitis. 2008;2(1):1–23.
13. Van Assche G, Dignass A, Panes J, Beaugerie L, Karagiannis J, Allez M, et al. The second European evidence-based consensus on the diagnosis and management of crohn's disease: Definitions and diagnosis. Journal of Crohn's and Colitis. 2010;4(1):7–27.
14. Rizzardi AE, Johnson AT, Vogel RI, Pambuccian SE, Henriksen J, Skubitz APN, et al. Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. Diagnostic Pathology. 2012;7(1).
15. Gavrielides MA, Gallas BD, Lenz P, Badano A, Hewitt SM. Observer Variability in the Interpretation of HER2/neu Immunohistochemical Expression With Unaided and Computer-Aided Digital Microscopy. Archives of Pathology & Laboratory Medicine. 2011 Feb 1;135(2):233–42.

16. Christensen KR, Ainsworth M, Skougaard M, Casper Steenholdt, Buhl S, Jørn Brynskov, et al. Identifying and understanding disease burden in patients with inflammatory bowel disease. BMJ Open Gastroenterology. 2022 Oct 1;9(1):e000994–4.

16. Chen P, Zhou G, Lin J, Li L, Zeng Z, Chen M, et al. Serum Biomarkers for Inflammatory Bowel Disease. Frontiers in Medicine [Internet]. 2020 Apr 22;7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7188783/#:~:text=C%2DReactive%20Protein
17. Laurent Peyrin-Biroulet, Reinisch W, Colombel JF, Mantzaris GJ, Kornbluth A, Diamond RH, et al. Clinical disease activity, C-reactive protein normalisation and mucosal healing in Crohn's disease in the SONIC trial. Gut. 2014 Jan 1;63(1):88–95.
18. Rodrigues BL, Mazzaro MC, Nagasako CK, Ayrizono M de LS, Fagundes JJ, Leal RF. Assessment of disease activity in inflammatory bowel diseases: Non-invasive biomarkers and endoscopic scores. World Journal of Gastrointestinal Endoscopy. 2020 Dec 16;12(12):504–20.
19. Henriksen M, Jahnsen J, Lygren I, Sauar J, Schulz T, Stray N, et al. Change of diagnosis during the first five years after onset of inflammatory bowel disease: Results of a prospective follow-up study (the IBSEN Study). Scandinavian Journal of Gastroenterology. 2006 Jan;41(9):1037–43.
20. Lee HS, Choe J, Lee HJ, Hwang SW, Park SH, Yang DH, et al. Change in the diagnosis of inflammatory bowel disease: a hospital-based cohort study from Korea. Intestinal Research [Internet]. 2016 [cited 2022 Aug 6];14(3):258. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4945530/
21. Limketkai BN, Singh S, Vipul Jairath, Sandborn WJ, Dulai PS. US Practice Patterns and Impact of Monitoring for Mucosal Inflammation After Biologic Initiation in Inflammatory Bowel Disease. Inflammatory Bowel Diseases. 2019 Oct 18;25(11):1828–37.
22. Denson LA, Curran M, McGovern DPB, Koltun WA, Duerr RH, Kim SC, et al. Challenges in IBD Research: Precision Medicine. Inflammatory Bowel Diseases [Internet]. 2019 May 16 [cited 2020 Jan 6];25(Supplement_2):S31–9. Available from: https://academic.oup.com/ibdjournal/article/25/Supplement_2/S31/5490491
23. Aronson JK, Ferner RE. Biomarkers—A General Review. Current Protocols in Pharmacology. 2017 Mar;76(1).
24. Seyedian S, Nokhostin F, Dargahi Malamir M. A review of the diagnosis, prevention, and treatment methods of inflammatory bowel disease. Journal of Medicine and Life. 2019;12(2):113–22.
25. Verstockt B, Bressler B, Martinez-Lozano H, McGovern D, Silverberg MS. Time to Revisit Disease Classification in Inflammatory Bowel Disease: Is the Current Classification of Inflammatory Bowel Disease Good Enough for Optimal Clinical Management? Gastroenterology. 2022 Apr;162(5):1370–82.
26. Yau Y, Leong RW, Zeng M, Wasinger VC. Proteomics and metabolomics in inflammatory bowel disease. Journal of Gastroenterology and Hepatology. 2013 Jun 20;28(7):1076–86.
27. Fiocchi C. Integrating Omics: The Future of IBD? Digestive Diseases. 2014;32(s1):96–102.
28. Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. Nature Reviews Drug Discovery [Internet]. 2016 Mar 11;15(7):473–84. Available from: https://www.nature.com/articles/nrd.2016.32

29. Iyer N, Corr SC. Gut Microbial Metabolite-Mediated Regulation of the Intestinal Barrier in the Pathogenesis of Inflammatory Bowel Disease. Nutrients [Internet]. 2021 Nov 26 [cited 2022 Jun 9];13(12):4259. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8704337/
30. Han X. Lipidomics for studying metabolism. Nature Reviews Endocrinology. 2016 Jul 29;12(11):668–79.
31. Daniluk U, Daniluk J, Kucharski R, Kowalczyk T, Pietrowska K, Samczuk P, et al. Untargeted Metabolomics and Inflammatory Markers Profiling in Children With Crohn's Disease and Ulcerative Colitis—A Preliminary Study. Inflammatory Bowel Diseases. 2019 Feb 17;25(7):1120–8.
32. Lavelle A, Sokol H. Gut microbiota-derived metabolites as key actors in inflammatory bowel disease. Nature Reviews Gastroenterology & Hepatology. 2020 Feb 19;17(4):223–37.
33. Thomas JP, Modos D, Rushbrook SM, Powell N, Korcsmaros T. The Emerging Role of Bile Acids in the Pathogenesis of Inflammatory Bowel Disease. Frontiers in Immunology. 2022 Feb 3;13.
34. Chiang JYL, Ferrell JM. Bile acid receptors FXR and TGR5 signaling in fatty liver diseases and therapy. American Journal of Physiology-Gastrointestinal and Liver Physiology. 2020 Mar 1;318(3):G554–73.
35. Sinha SR, Haileselassie Y, Nguyen LP, Tropini C, Wang M, Becker LS, et al. Dysbiosis-Induced Secondary Bile Acid Deficiency Promotes Intestinal Inflammation. Cell Host & Microbe. 2020 Apr;27(4):659-670.e5.
36. Duboc H, Rajca S, Rainteau D, Benarous D, Maubert MA, Quervain E, et al. Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. Gut. 2012 Sep 19;62(4):531–9.
37. Jacobs JP, Goudarzi M, Singh N, Tong M, McHardy IH, Ruegger P, et al. A Disease-Associated Microbial and Metabolomics State in Relatives of Pediatric Inflammatory Bowel Disease Patients. Cellular and Molecular Gastroenterology and Hepatology. 2016 Nov;2(6):750–66.
38. Roda, Porru, Katsanos, Skamnelos, Kyriakidi, Fiorino, et al. Serum Bile Acids Profiling in Inflammatory Bowel Disease Patients Treated with Anti-TNFs. Cells. 2019 Aug 2;8(8):817.
39. Ding NS, McDonald JAK, Perdones-Montero A, Rees DN, Adegbola SO, Misra R, et al. Metabonomics and the Gut Microbiome Associated With Primary Response to Anti-TNF Therapy in Crohn's Disease. Journal of Crohn's and Colitis [Internet]. 2020 Mar 2 [cited 2022 Dec 12];14(8):1090–102. Available from: https://academic.oup.com/ecco-jcc/article/14/8/1090/5771407
40. Adegbola SO, Sahnan K, Warusavitarne J, Hart A, Tozer P. Anti-TNF Therapy in Crohn's Disease. International Journal of Molecular Sciences [Internet]. 2018 Jul 31;19(8). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6121417/
41. Corfield AP, Myerscough N, Bradfield N, Corfield C do A, Gough M, Clamp JR, et al. Colonic mucins in ulcerative colitis: evidence for loss of sulfation. Glycoconjugate Journal [Internet]. 1996 Oct 1 [cited 2023 May 2];13(5):809–22. Available from: https://pubmed.ncbi.nlm.nih.gov/8910008/
42. Dawson PA, Huxley S, Gardiner B, Tran T, McAuley JL, Grimmond S, et al. Reduced mucin sulfonation and impaired intestinal barrier function in the hyposulfataemic NaS1 null mouse. Gut. 2009 Feb 6;58(7):910–9.

43. Raouf AH, Tsai HH, Parker N, Hoffman J, Walker RJ, Rhodes JM. Sulphation of colonic and rectal mucin in inflammatory bowel disease: reduced sulphation of rectal mucus in ulcerative colitis. Clinical Science (London, England: 1979) [Internet]. 1992 Nov 1 [cited 2023 May 2];83(5):623–6. Available from: https://pubmed.ncbi.nlm.nih.gov/1335401/
44. McNeil NI, Cummings JH, James WP. Short chain fatty acid absorption by the human large intestine. Gut. 1978 Sep 1;19(9):819–22.
45. Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunger MK, et al. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell metabolism [Internet]. 2011 [cited 2019 Dec 3];13(5):517–26. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21531334
46. Donohoe Dallas R, Collins Leonard B, Wali A, Bigler R, Sun W, Bultman Scott J. The Warburg Effect Dictates the Mechanism of Butyrate-Mediated Histone Acetylation and Cell Proliferation. Molecular Cell. 2012 Nov;48(4):612–26.
47. Park J, Kim M, Kang SG, Jannasch AH, Cooper B, Patterson J, et al. Short-chain fatty acids induce both effector and regulatory T cells by suppression of histone deacetylases and regulation of the mTOR–S6K pathway. Mucosal Immunology. 2014 Jun 11;8(1):80–93.
48. Kim MH, Kang SG, Park JH, Yanagisawa M, Kim CH. Short-chain fatty acids activate GPR41 and GPR43 on intestinal epithelial cells to promote inflammatory responses in mice. Gastroenterology [Internet]. 2013 [cited 2019 Oct 21];145(2):396-406.e110. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23665276
49. Thibault, R., Pierre de Coppet, Daly, K., Arnaud Bourreille, Cuff, M. A., Bonnet, C., Jean-François Mosnier, Galmiche, J. P., Shirazi–Beechey, S., & Jean Pierre Segain. (2007). Down-Regulation of the Monocarboxylate Transporter 1 Is Involved in Butyrate Deficiency During Intestinal Inflammation. Gastroenterology, 133(6), 1916–1927. https://doi.org/10.1053/j.gastro.2007.08.041.
50. Parada Venegas D, De la Fuente MK, Landskron G, González MJ, Quera R, Dijkstra G, et al. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. Frontiers in Immunology [Internet]. 2019 Mar 11;10(277). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6421268
51. De Preter V, Arijs I, Windey K, Vanhove W, Vermeire S, Schuit F, et al. Impaired butyrate oxidation in ulcerative colitis is due to decreased butyrate uptake and a defect in the oxidation pathway*. Inflammatory Bowel Diseases. 2012 Jun;18(6):1127–36.
52. Bloemen JG, Venema K, van de Poll MC, Olde Damink SW, Buurman WA, Dejong CH. Short chain fatty acids exchange across the gut and liver in humans measured at surgery. Clinical Nutrition. 2009 Dec;28(6):657–61.
53. Ghia J, Li N, Wang H, Collins M, Deng Y, El–Sharkawy RT, et al. Serotonin Has a Key Role in Pathogenesis of Experimental Colitis. Gastroenterology. 2009 Nov;137(5):1649–60.
54. Zenewicz LA, Yancopoulos GD, Valenzuela DM, Murphy AJ, Stevens S, Flavell RA. Innate and Adaptive Interleukin-22 Protects Mice from Inflammatory Bowel Disease. Immunity. 2008 Dec;29(6):947–57.
55. Gupta NK, Thaker AI, Kanuri N, Riehl TE, Rowley CW, Stenson WF, et al. Serum Analysis of Tryptophan Catabolism Pathway: Correlation With Crohn's Disease Activity. Inflammatory Bowel Diseases [Internet]. 2012 Jul [cited 2022 Nov 3];18(7):1214–20. Available from: https://academic.oup.com/ibdjournal/article/18/7/1214/4607697
56. Beeken WL. Serum Tryptophan in Crohn's Disease. Scandinavian Journal of Gastroenterology. 1976 Jan 1;11(7):735–40.

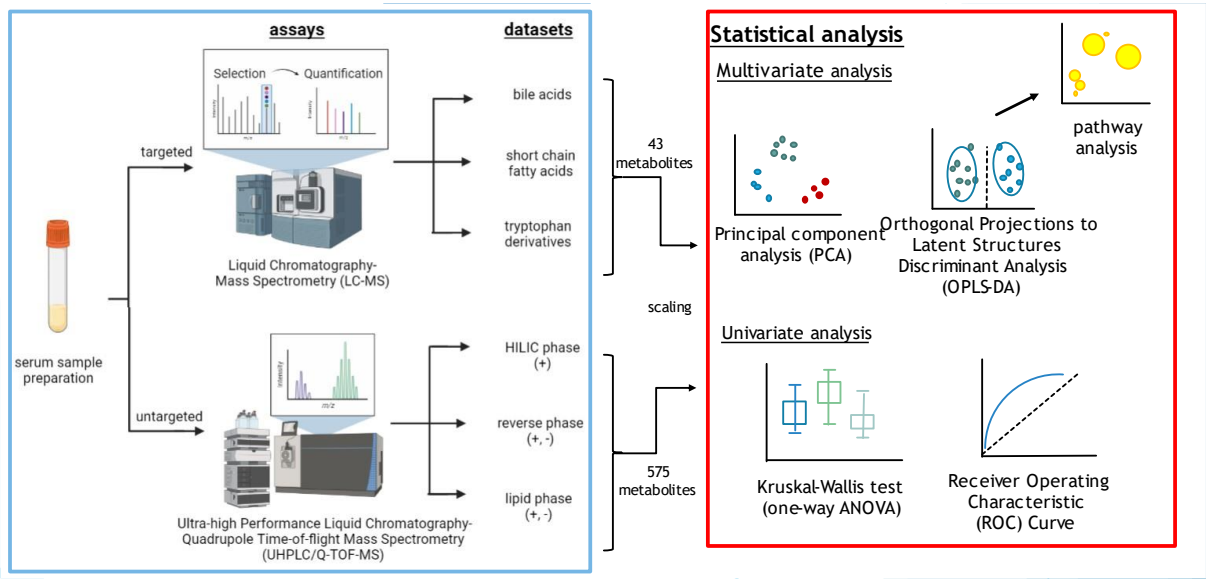
57. Clayton PE, Bridges NA, Atherton DL, Milla PJ, Malone M, Bender DA. Pellagra with colitis due to a defect in tryptophan metabolism. European Journal of Pediatrics. 1991 May 1;150(7):498–502.
58. Nikolaus S, Schulte B, Al-Massad N, Thieme F, Schulte DM, Bethge J, et al. Increased Tryptophan Metabolism Is Associated With Activity of Inflammatory Bowel Diseases. Gastroenterology [Internet]. 2017 Dec [cited 2019 Nov 8];153(6):1504-1516.e2. Available from: https://www.gastrojournal.org/article/S0016-5085(17)36061-4/fulltext
59. Jones B, Sands C, Kleopatra Alexiadou, Bloom SR, Tharakan G, Behary P, et al. The Metabolomic Effects of Tripeptide Gut Hormone Infusion Compared to Roux-en-Y Gastric Bypass and Caloric Restriction. The Journal of Clinical Endocrinology and Metabolism. 2021 Aug 30;107(2):e767–82.
60. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Research. 2021 May 21;49(W1).
61. Aittokallio T, Schwikowski B. Graph-based methods for analysing networks in cell biology. Briefings in Bioinformatics [Internet]. 2006 Sep 1 [cited 2021 Apr 25];7(3):243–55. Available from: https://academic.oup.com/bib/article/7/3/243/326999
62. Krumsiek J, Mittelstrass K, Do KT, Stückler F, Ried J, Adamski J, et al. Gender-specific pathway differences in the human serum metabolome. Metabolomics. 2015 Aug 4;11(6):1815–33.
63. Metabolomics analysis of human serum for characterisation of phenotypically ageing men in different ethnic populations [Internet]. Research Explorer The University of Manchester. [cited 2023 May 2]. Available from: https://research.manchester.ac.uk/en/studentTheses/metabolomics-analysis-of-human-serum-for-characterisation-of-phen
64. Metabolomics analysis of human serum for characterisation of phenotypically ageing men in different ethnic populations [Internet]. Research Explorer The University of Manchester. [cited 2023 May 2]. Available from: https://research.manchester.ac.uk/en/studentTheses/metabolomics-analysis-of-human-serum-for-characterisation-of-phen
65. Roma E, Krini M, Hantzi E, Sakka S, Panayiotou I, Margeli A, et al. Retinol Binding Protein 4 in children with Inflammatory Bowel Disease: a negative correlation with the disease activity. Hippokratia [Internet]. 2012 [cited 2023 May 2];16(4):360–5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3738613/
66. Nielsen OH, Hansen TI, Gubatan JM, Jensen KB, Rejnmark L. Managing vitamin D deficiency in inflammatory bowel disease. Frontline Gastroenterology [Internet]. 2019 Jan 7;10(4):394–400. Available from: https://fg.bmj.com/content/10/4/394
67. Lee IA, Low D, Kamba A, Llado V, Mizoguchi E. Oral caffeine administration ameliorates acute colitis by suppressing chitinase 3-like 1 expression in intestinal epithelial cells. Journal of Gastroenterology. 2013 Aug 8;49(8):1206–16.
68. Barthel C, Wiegand S, Scharl S, Scharl M, Frei P, Vavricka SR, et al. Patients' perceptions on the impact of coffee consumption in inflammatory bowel disease: friend or foe? – a patient survey. Nutrition Journal [Internet]. 2015 Aug 12 [cited 2020 May 15];14. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4534065/

69. Tiratterra E. Role of bile acids in inflammatory bowel disease. Annals of Gastroenterology. 2018;
70. Englund G, Jacobson A, Rorsman F, Artursson P, Kindmark A, Rönnblom A. Efflux transporters in ulcerative colitis. Inflammatory Bowel Diseases. 2007 Mar;13(3):291–7.
71. Romain Ferru-Clément, Boucher G, Forest A, Bouchard B, Bitton A, Lesage S, et al. Serum Lipidomic Screen Identifies Key Metabolites, Pathways, and Disease Classifiers in Crohn's Disease. Inflammatory Bowel Diseases. 2023 Jan 20
72. Davis I, Liu A. What is the tryptophan kynurenine pathway and why is it important to neurotherapeutics? Expert Review of Neurotherapeutics. 2015 May 24;15(7):719–21.
73. Guillemin GJ. Neurodegenerative Diseases: Tryptophan Metabolism. Springer eBooks. 2008 Nov 21;2620–3.
74. Bosco MC, Rapisarda A, Massazza S, Melillo G, Young H, Varesio L. The Tryptophan Catabolite Picolinic Acid Selectively Induces the Chemokines Macrophage Inflammatory Protein-1α and -1β in Macrophages. The Journal of Immunology. 2000 Mar 15;164(6):3283–91.
75. Chen LM, Bao CH, Wu Y, Liang SH, Wang D, Wu LY, et al. Tryptophan-kynurenine metabolism: a link between the gut and brain for depression in inflammatory bowel disease. Journal of Neuroinflammation [Internet]. 2021 Jun 14;18:135. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8204445/
76. Chojnacki C, Popławski T, Gasiorowska A, Chojnacki J, Blasiak J. Serotonin in the Pathogenesis of Lymphocytic Colitis. Journal of Clinical Medicine. 2021 Jan 14;10(2):285.
77. Chojnacki C, Błońska A, Kaczka A, Chojnacki J, Stępień A, Gąsiorowska A. Evaluation of serotonin and dopamine secretion and metabolism in patients with irritable bowel syndrome. Polish Archives of Internal Medicine. 2018 Nov 6;

metabolism	Subcategory	Metabolites	Within range
Bile acids	Primary bile acid	5-Cholenic Acid-3-beta-ol	Y
		3-alpha-Hydroxy-7 Ketolithocholic Acid	Y
		Murocholic Acid	Y
		Chenodeoxycholic Acid	Y
		Glycochenodeoxycholic Acid	Y
		Cholic Acid	Y
		Glycocholic Acid	Y
		Hyocholic Acid	Y
		Glycohyocholic Acid	Y
		Taurocholic Acid	Y
		Taurochenodeoxycholic Acid	Y
	Secondary bile acid	Deoxycholic Acid	Y
		Glycodeoxycholic Acid	Y
		Glycoursodeoxycholic Acid	Y
		Taurodeoxycholic Acid	Y
		Ursodeoxycholic Acid	Y
	Primary sulfated bile acids	Glycochenodeoxycholic Acid 3-Sulfate	Y
		Taurochenodeoxycholic Acid 3-Sulfate	Y
	Secondary sulphated bile acids	Glycolithocholic Acid 3-Sulfate	Y
		Taurolithocholic Acid 3-Sulfate	y
		Glycoursodeoxycholic Acid 3-Sulfate	Y
		Glycodeoxycholic Acid 3-Sulfate	Y
Short chain fatty acids		Lactate	Y
		Acetate	Y
		Propionate	Y
		2-Hydroxy butyrate	Y
		Isobutyrate	Y
		2-methyl butyrate	Y
		Isovalerate	Y
		Valerate	LLOQ
		Hexanoate	Y
Tryptophan	Serotonin derivatives	Serotonin / 5-HT	LLOQ/HLOQ
		5-hydroxyindole acetic acid	Y
	Kynurenine derivatives	Kynurenine	Y
		Kynurenic acid	Y
		3-HAA	Y
		Xanthurenic acid	Y
		Picolinic acid	Y
		Nicotinic acid	Y
		Quinolinic acid	Y
	Indole derivatives	Indole-3-acetic-acid	Y
	Others	Tryptophan	Y
		Citrulline	ULOQ
		Neopterin	Y

Supplementary Figure 1. Targeted metabolite list based on derivatives related to tryptophan, bile acid and short chain fatty acid metabolism. These metabolites are absolutely quantified with resepect to standards in targeted metabolomics. LLOQ, lower limit of quantification; ULOQ, upper limit of quantification; Y, yes.

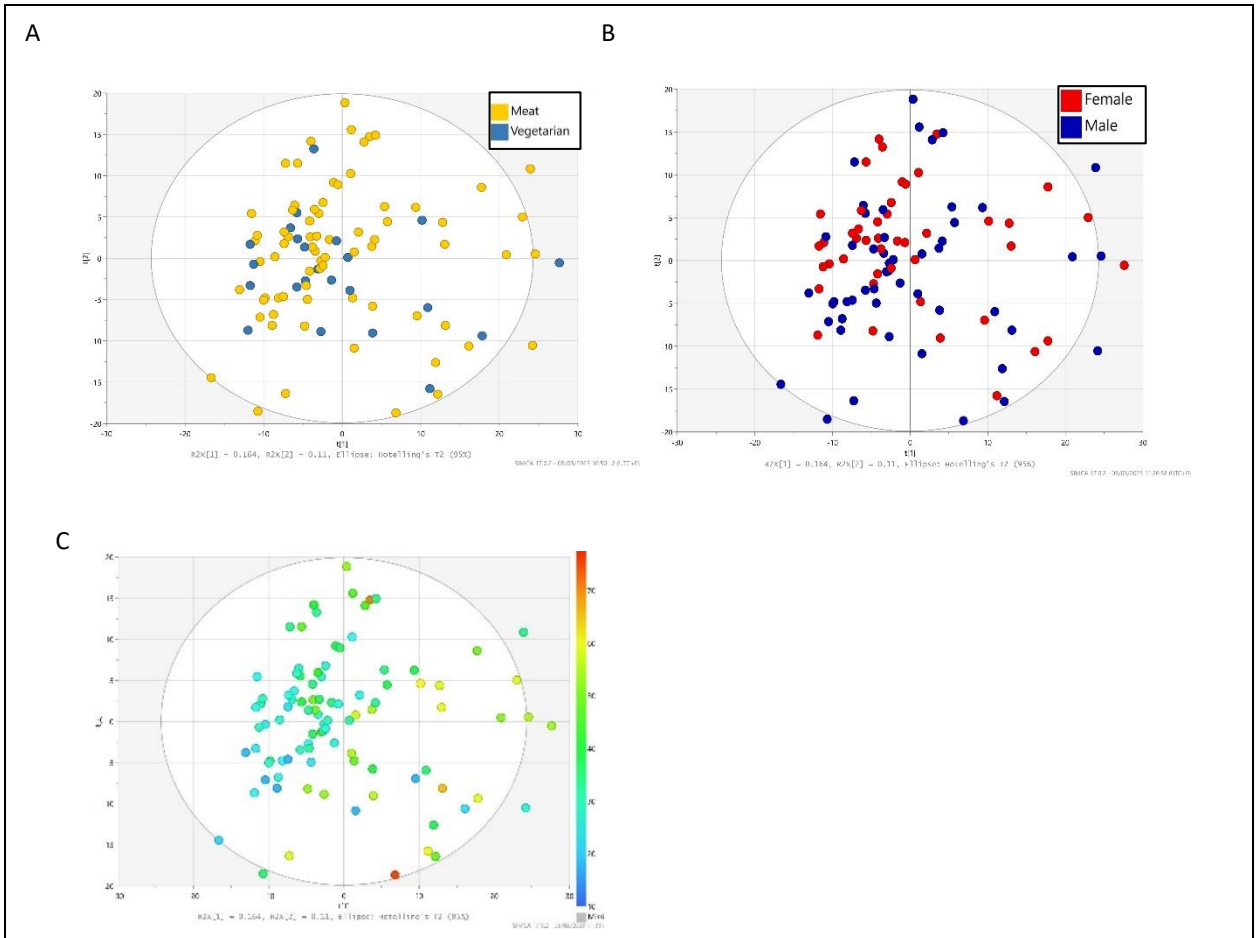
A



B

Untargeted serum dataset	Sample subject (subjects excluded)	Number of metabolites post-processing
RPC+	101 (5)	56
RPC-	101 (5)	43
LPC+	96	339
LPC-	96	97
HILIC+	101 (5)	40

Supplementary Figure 2. Summary of study workflow. (A) Overview of the from serum sample processing from sample preparation, analytical assays performed in National Phenome Centre (framed blue), and statistical analysis (framed red). (B) table summarizing datasets from 5 analytical assays performed for untargeted serum metabolomics, which includes different chromatographic phases with different ion modes of electrospray ionization. HILIC, Hydrophilic interaction liquid chromatography; LPC, lipid phase chromatography; RPC, Reverse phase chromatography; +, positive ion mode; -, negative ion mode.



Supplementary Figure 3. PCA plots of IBD cohort. (n=96). Plots are colour-coded by (A) ethnicity, (B) sex and (C) age.

A

Group comparison	n	A	R^2X (cum)	R^2Y (cum)	Q (cum)	CV-ANOVA	Statistical significance
Caucasians and non-Caucasians among IBD	58	1 + 1	0.188	0.712	0.311	0.000482137	**
Sex among IBD	58	1 + 1	0.147	0.792	0.456	1.285×10^{-6}	

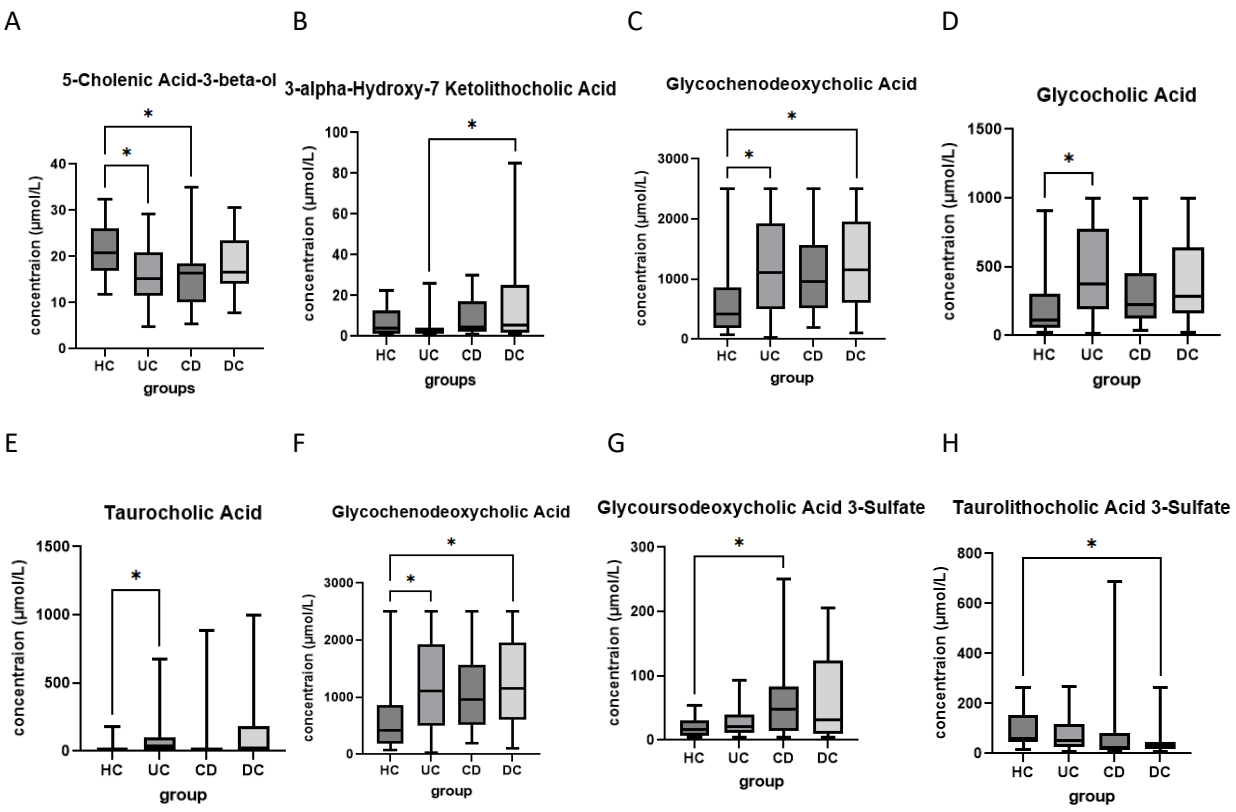
B

Group comparison	n	A	R^2X (cum)	R^2Y (cum)	Q (cum)	CV-ANOVA	Statistical significance
HC vs UC	56	1 + 1	0.198	0.762	0.549	2.221×10^{-8}	***
HC vs CD	42	1 + 1	0.149	0.852	0.539	6.521×10^{-6}	***
DC vs UC	54	1 + 1	0.187	0.716	0.367	0.000137	***
DC vs CD	40	1 + 1	0.225	0.743	0.324	0.00698	**
UC vs CD	58	0 + 0	NA	NA	NA	NA	ns

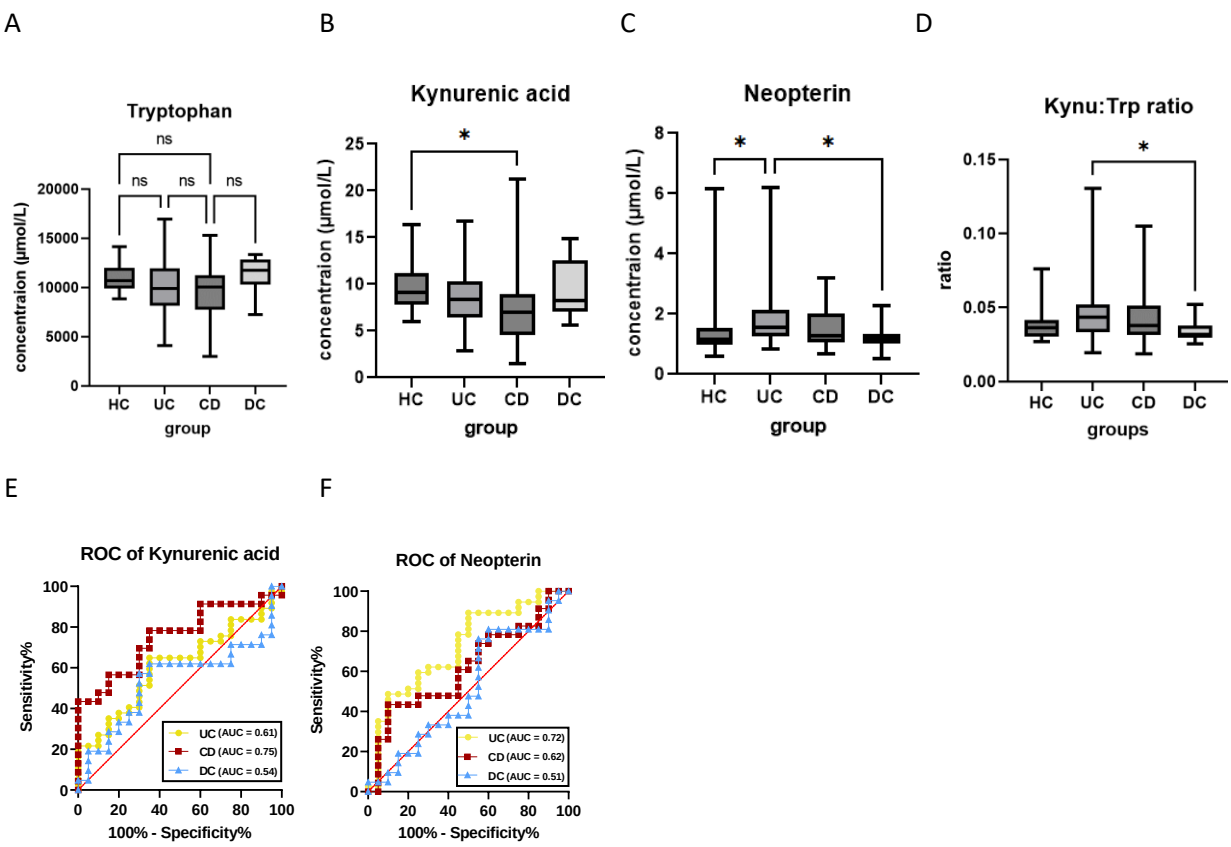
C

Group comparison	n	A	R^2X (cum)	R^2Y (cum)	Q (cum)	CV-ANOVA	Statistical significance
HC vs UC	56	1 + 0	0.126	0.485	0.39	1.61×10^{-6}	*****
HC vs CD	42	1 + 1	0.152	0.509	0.376	1.03×10^{-4}	***
HC vs DC	40	1+1	0.215	0.744	0.459	0.000186	****
DC vs UC	54	1 + 0	0.170	0.437	0.229	0.000983	****
DC vs CD	40	1 + 0	0.121	0.42	0.179	0.0219	*
UC vs CD	58	1 + 0	0.127	0.377	0.187	0.00312	***

Supplementary 4. Summary of OPLS-DA model performance based on serum metabolite profile.
Comparisons made are based on (A) patient demographics; (B) clinical subtypes. (C) Performance of OPLS-DA model on targeted metabolomics was also summarized based on clinical subtypes. All models are fitted to first two components only or less to prevent overfitting. All values are corrected to nearest 3 significant figures. ns, not significant; *P < 0.05; **P < 0.01; ***P < 1×10^{-3}; ****P < 1×10^{-4}; *****P < 1×10^{-5}.



Supplementary 5. Targeted metabolite profiles BA and conjugates, and SCFA. Univariate Kruskal Wallis tests were performed. Here shows the metabolites with significance but does not have satisfactory performance.



Supplementary 6. Targeted metabolite profiles of tryptophan derivatives. Univariate Kruskal Wallis tests were performed. Here shows the metabolites with significance but does not have satisfactory performance. ROC performance of the significant metabolites were included in (E) for kynurenic acid, and (F) neopterin.