

Gluten free diet does not fully restore the faecal metabolome in paediatric coeliac disease

Nicholas Rattray

nicholas.rattray@strath.ac.uk

University of Strathclyde <https://orcid.org/0000-0002-3528-6905>

Patricia Kelly

Gillian farrell

Richard Russell

Richard Hansen

Christine Edwards

Peter Gillett

Konstantinos Gerasimidis

Article

Keywords: coeliac disease, faecal metabolomics, gluten-free diet, bile acids, paediatric gastroenterology

Posted Date: July 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10187795/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Abstract

Background and Aims

Coeliac disease (CD) is an autoimmune enteropathy triggered by gluten ingestion, with a strict gluten-free diet (GFD) as the only currently effective treatment. Despite clinical and serological improvements with dietary management, the extent to which the gut metabolic environment normalises remains poorly understood. This study aimed to characterise the faecal metabolome of children with CD across disease states and following GFD intervention.

Methods

Faecal samples were collected from 143 children aged 2–15 years, comprising untreated CD (UCD; n = 27), treated CD (TCD; n = 40), healthy controls (HC; n = 56), and unaffected siblings (n = 20). A prospective cohort of UCD patients (n = 12) was sampled at diagnosis and at 6 and 12 months after GFD initiation. Untargeted liquid chromatography–mass spectrometry (LC-MS) metabolomics was performed, with multivariate (OPLS-DA, PC-CVA) and univariate statistical analyses applied, alongside ROC-based biomarker evaluation.

Results

Three distinct metabolite patterns were identified: (1) CD-specific metabolites persistently altered in both UCD and TCD relative to HC, including bile acid derivatives, sterols, acylcarnitines, and amino acid dipeptides (combined AUC = 0.868); (2) treatment-responsive metabolites normalised by GFD, predominantly amino acid oligopeptides and microbial-derived compounds such as indole-3-lactic acid; and (3) diet-driven metabolites altered in TCD independently of disease status, reflecting GFD-related shifts in tryptophan and purine metabolism. Longitudinal analysis confirmed dynamic, time-dependent metabolic changes over 12 months of treatment.

Conclusion

Paediatric CD is characterised by distinct faecal metabolic signatures that only partially resolve with GFD. Persistent alterations in bile acid, sterol, and microbial metabolite pathways may reflect ongoing gut dysfunction and represent promising non-invasive biomarkers for disease monitoring.

Introduction

Coeliac disease (CD) is a multifactorial autoimmune disorder of the small intestine caused by ingestion of gluten in genetically predisposed individuals. Disease incidence has increased in recent years, with 1.4% of the global population currently estimated to be living with the condition [1]. Beyond the growing prevalence, the economic and personal burden of CD is substantial. In the United States, the average cost associated with diagnosis is estimated at approximately \$6,000 per individual per year [2].

Adherence to a gluten-free diet (GFD) imposes additional financial strain, with gluten-free products costing an estimated 159% more than their conventional counterparts [3].

CD results in small intestinal villous atrophy and inflammation causing abdominal pain, nutrient malabsorption, diarrhoea, bloating, and weight loss. The diagnosis of CD can often be delayed, both due to symptom overlap with other conditions, and, in some patients with a silent disease phenotype, an absence of overt gastrointestinal symptoms. A combination of diagnostic tools and biomarkers in blood are therefore required alongside evidence of clinical manifestations. Serological tests are used to measure levels of IgA anti-tissue transglutaminase antibodies (tTG) and anti-endomysial antibodies (EMA), which are produced in response to gluten-activated immune pathways.

The only effective treatment of CD at present is strict and lifelong adherence to a GFD. While effective in alleviating symptoms and promoting mucosal healing, a GFD can be difficult to maintain, as gluten is widely present in food products and is frequently encountered through cross-contamination during food processing. The currently limited therapeutic options highlight the need for a more comprehensive understanding of disease pathogenesis. Identification of novel disease biomarkers and mechanistic insights will help to improve patient outcomes through earlier diagnosis, improved disease and treatment monitoring, and prediction of complications.

It is now recognised that the gut microbiota plays a crucial role in CD pathophysiology [4, 5]. A recent study observed distinct changes in the gut microbiota in children with CD compared to healthy children [6]. Most notably, a microbial signature specific to CD was identified comprising eleven operational taxonomic units, with a distinct altered microbial composition found following a GFD. Additionally, the predictive ability of the gut microbiota in the diagnosis of CD has been explored, with recent studies suggesting the possibility of non-invasive diagnostic or early predictive microbiome-associated biomarkers for CD, achieving up to 82% accuracy [7] and identifying microbial and metabolite shifts 18 months before disease onset in at-risk infants [8, 9]. In line with evidence supporting an altered microbiome as part of CD pathophysiology, complementary study of the metabolome may facilitate further understanding of the disease onset and progression.

While there is clear evidence for disease-associated metabolic alterations, the determination of a metabolic fingerprint for CD is still in the early phases. Liquid chromatography mass spectrometry (LC-MS) based metabolomics approaches are beneficial due to their high sensitivity and selectivity for detecting molecules in a range of applications. Importantly, stool samples are an advantageous non-invasive biotransform for studying gastrointestinal disease, providing a window into gut microbial and metabolic function. We hypothesise that the application of faecal metabolomics to investigate coeliac disease has the potential to unravel relationships between microbial metabolism, the immune response, and the diet. This will help not only in the understanding of disease mechanisms and dietary impact but also holds promise in complementary diagnostic and prognostic tools.

The aim of the present study is to characterise the faecal metabolome of children with CD. The utilisation of both cross-sectional and prospective cohorts allowed an in-depth analysis of the metabolic

impact of disease and treatment. Comparisons were performed between treatment-naïve CD patients at the point of disease diagnosis (UCD), patients with established CD and on a GFD (TCD), healthy siblings of TCD patients, and healthy controls (HC). UCD patients were examined at the time of diagnosis and at 6 and 12-month follow-up time points after treatment with a GFD.

Materials and Methods

Subjects

Faecal samples were collected from children with CD receiving care at the Royal Hospital for Children in Glasgow. Children with a CD diagnosis were recruited from annual clinic appointments, while newly diagnosed cases were referred by primary healthcare services. Unaffected siblings of children diagnosed with CD were screened negative to tTG antibodies. Exclusion criteria included antibiotic use in the past three months, regular prebiotic and probiotic supplementation, or presence of other comorbidities.

Confirmation of a CD diagnosis was by small bowel biopsy, following the UK guidelines that were effective at the time of recruitment [10]. Assessment of GI symptoms was carried out using the PedsQL-GS questionnaire (version 1) [11]. Compliance to GFD was assessed by measuring gluten immunogenic peptides (GIP) in faeces plus dietetic assessment and repeat measurements of tTG (iVYLISA; Biomedal, Seville, Spain) [12].

Sample Collection and Storage

Complete bowel movements were collected and stored under cold anaerobic conditions (Anaerocult A; Merck, Darmstadt, Germany) and processed within 2 hours. Samples were homogenised and stored at -80°C until further processing. A single faecal sample was obtained from all groups except the UCD cohort, from whom three samples were collected: a baseline sample prior to diagnostic endoscopy while patients were consuming a gluten-containing diet, and follow-up samples at 6 and 12 months after commencing a GFD.

Chemicals and Reagents

LC-MS grade methanol, acetonitrile, and water were purchased from Fisher Scientific (Geel, Belgium), and formic acid was purchased from Thermo Scientific (Czech Republic).

Faecal Metabolite Extraction and LC-MS Analysis

Metabolite extraction was performed using a previously optimised method for LC-MS analysis of faecal samples [13]. Untargeted metabolomics analysis was performed on a binary Vanquish ultra-high performance liquid chromatography (UHPLC) system (ThermoFisher Scientific, Bremen, Germany)

coupled to an Orbitrap Exploris 240 (ThermoFisher Scientific, Bremen, Germany) mass spectrometer. Full LC-MS parameters are provided in **Supplementary Information Table S1**.

Quality Control and Analytical Design

Quality control (QC) samples were prepared by pooling all samples and analysed throughout each batch. Samples were analysed as two batches and batch correction of peak areas was performed using QC-MXP software [14]. Software configuration is shown in **Supplementary Information Table S2**.

Mass Spectrometry Data Processing

Thermo Scientific Xcalibur format raw data files (.RAW) were processed using Compound Discoverer 3.2. Full parameters are shown in **Supplementary Information Table S3**.

Data and Statistical Analysis

Statistical analyses were performed using R. Univariate statistical analyses were performed using unpaired t-tests and one-way ANOVA. p-values were adjusted using the Benjamini-Hochberg false discovery rate (FDR) procedure, with $p < 0.05$ considered statistically significant. Principal Component-Canonical Variate Analysis (PC-CVA), a supervised multivariate method that combines PCA-based dimensionality with discriminant analysis, was performed using the *vegan* package [15]. Orthogonal partial least squares discriminant analysis (OPLS-DA) was performed using the *ropls* package [16] and evaluated using R^2 and Q^2 values, with model quality defined by parameter values > 0.5 [17].

Receiver Operating Characteristic (ROC) analyses were performed using the *pROC* R package [18]. Logistic regression was applied with cross-validation using the *glmnet* package [19].

Results

Participant Demographics and Clinical Parameters

This study investigated a total of 143 participants aged between two and 15 years old, including 56 healthy controls (HCs), 27 UCD children, 40 TCD children on a GFD, and 20 unaffected siblings of TCD children (Table 1). There were no significant differences in gender, or mean age, weight, or body mass index (BMI) among the study groups ($p < 0.05$). UCD children had a higher incidence of gastrointestinal symptoms (PedsQL-GS) compared to both the TCD group and HCs. The TCD group additionally presented a lower PedsQL-GS score than HCs. In the prospective cohort of UCD, tTG levels decreased with GFD treatment, paralleled by an increase in PedsQL-GS score.

Table 1
Patient demographics of cross-sectional and prospective cohorts.

Cross-Sectional Study					Prospective Study		
Variable	HC n = 56	Siblings n = 20	UCD group n = 27	TCD group n = 40	At diagnosis n = 12	GFD 6 months n = 12	GFD 12 months n = 12
Age (y)	7.8 (0.41)	9.3 (0.62)	9.6 (0.60)	9.2 (0.48)	9.3 (0.92)	9.9 (0.92)	10.5 (0.91)
Gender (F/M)	30/26	10/10	14/13	23/17	6/6	6/6	6/6
Weight z- score (SD)	0.14 (0.15)	0.34 (0.26)	-0.19 (0.27)	0.15 (0.18)	-0.30 (0.36)	-0.35 (0.33)	-0.29 (0.32)
Height z- score (SD)	0.27 (0.15)	0.55 (0.28)	-0.18 (0.20)	-0.05 (0.16)	-0.02 (0.27)	-0.02 (0.27)	-0.06 (0.26)
BMI (kg/m ²)	16.8 (0.46)	17.3 (0.68)	17.2 (0.75)	17.5 (0.41)	16.3 (0.69)	16.4 (0.72)	16.8 (0.79)
BMI z- score (SD)	0.05 (0.14)	0.08 (0.24)	-0.15 (0.25)	0.26 (0.17)	-0.46 (0.36)	-0.53 (0.35)	-0.41 (0.35)
tTG (U/mL)			75.3 (11.21)	11.7 (4.41)	105.7 (14.37)	23.9 (11.97)	7.7 (2.08)
PedsQL-GS score	100.0 (2.32)		56.3 (7.28)	77.7 (3.95)	51.2 (6.60)	65.97 (5.66)	71.2 (8.73)

Values are mean (SEM) unless otherwise stated. HC, healthy control; UCD, Untreated coeliac disease, TCD, Treated coeliac disease.

Faecal Metabolome Profiling

To explore overall metabolic variation across study groups, we performed principal components-canonical variate analysis (PC-CVA) on the faecal metabolomics dataset. In the cross-sectional cohort, this approach revealed clear group-level clustering between HC, UCD, and TCD groups (Fig. 1A). In the prospective UCD cohort, metabolite trajectories across 0, 6, and 12 months of GFD treatment were observed (Fig. 1B).

Differential Analysis in Metabolome between Untreated CD and HCs

Multivariate analysis using orthogonal partial least squares discriminant analysis (OPLS-DA) revealed partial separation of the global metabolome between UCD patients and HCs ($R^2Y = 0.555$, $Q^2 = 0.267$;

Fig. 2A). Although the Q^2 value indicates moderate predictive ability and falls below the conventional 0.5 threshold for strong model performance, this is consistent with the subtle metabolic perturbations expected in UCD. Nonetheless, key metabolites driving class discrimination based on their variable importance in projection (VIP) scores were identified (**Supplementary Information Figure S2**). Subsequent univariate analysis identified 38 metabolites to be significantly altered between the two groups, of which 17 were elevated and 21 were reduced relative to HCs (Fig. 2B). Among the most discriminant features were acetylcarnitine, hippuric acid, glycine, leucyl-prolyl-tyrosine, nutriacholic acid, undecanoylcarnitine, 9(11)-dehydroxysterol, and S-allyl-L-cysteine (Fig. 2C). All significantly altered metabolites are displayed in **Supplementary Information Table S4**.

Differential Analysis in Metabolome Between Patients on Recommendation to a GFD with Untreated CD or HCs

To investigate the impact that GFD had on the faecal metabolome, we compared metabolite profiles between TCD patients with both UCD patients and HC groups. We first explored the effects of a GFD on the metabolome of CD patients, with VIP scores from the OPLS-DA model highlighting the metabolites that most strongly contributed to distinguishing TCD from UCD patients (**Supplementary Figure S3**). This was complemented by pairwise analysis identifying 33 significantly altered metabolites (11 decreased and 22 increased), relative to UCD. The metabolites which are altered only between TCD and UCD groups, reflecting the metabolic changes upon a GFD in individuals with CD, are shown in Fig. 3. Metabolites increased in TCD patients compared to UCD patients includes a panel of amino acid dipeptides and tripeptides such as Leu-Ala-Ser, Val-Leu, Trileucine, Ile-Leu, Leu-Leu, Val-Tyr, and L-Valyl-L-tyrosine, alongside additional changes in phenolic and purine metabolites (ferulic acid, 9-methyluric acid, 1,7-dimethyluric acid). In contrast, metabolites decreased in TCD patients included microbial-derived and nucleotide-related compounds such as indole-3-lactic acid, 1,7-dimethyluric acid, and other purine derivatives.

When comparing metabolite profiles between TCD patients and HCs, OPLS-DA analysis identified discriminant metabolites, with univariate testing confirming 24 metabolites significantly altered between the groups (**Supplementary Figure S4**). Notably, almost half of these metabolites (11 out of the 24, 46%) overlapped with those also significantly altered in the comparison between TCD and UCD patients. Within this panel, several metabolites were reduced in TCD relative to both UCD and HC, including 2-aminobutyric acid, phenylalanine, N-(indole-3-acetyl) leucine, isovaline, C12-carnitine, galacosylhydroxylysine, and xanthurenic acid. Additionally, xanthine, N-acetylmuramic acid, indole-3-lactic acid, and 1,7-dimethyluric acid were elevated in TCD patients. This indicates a subset of metabolites that are modulated by a GFD independently of disease status rather than by disease activity, with the metabolite panel illustrated in (**Supplementary Figure S5**).

Coeliac Disease-specific Metabolome Signature

To identify metabolite features specific to CD, we identified the metabolites that were significantly altered between both HCs and UCD as well as between HCs and TCD (Fig. 4A). Receiver operating characteristic (ROC) analysis (Fig. 4B/4C) was performed to characterise the predictive value of the most influential disease metabolites. The combined metabolite model demonstrated strong classification performance, showing an area under the curve (AUC) value of 0.868, based on five-fold cross-validated predictions. The optimal probability threshold, determined using Youden's index, was 0.716, corresponding to sensitivity of 81.6% and a specificity of 79.6%.

Discriminant Analysis in Metabolome Profile in New-onset CD After Recommendation to a GFD

Analysis of the prospective cohort was performed at three timepoints: (1) at disease diagnosis, (2) 6 months and (3) 12 months following initiation of a GFD. At follow-up, faecal GIP was detectable in 15% of patients at 6 months and 23% of patients at 12 months, with GIP positive patients excluded from the analysis. Metabolite profiling revealed 10 significantly altered metabolites at six months and nine metabolites altered at 12 months relative to baseline (Fig. 5). Of these, 7 were uniquely increased at 6 months, while 6 were only increased at 12-months follow-up, indicating dynamic and time-dependent metabolic changes to a GFD. Importantly, monolinolenin, N-indole-3-acetyl-leucine, and 9(11)-dehydroxysterol were consistently altered at both follow-up points, representing persistent treatment-responsive metabolites. These longitudinal findings align with trends observed in the cross-sectional cohort, where 10 out of the 16 metabolites (63%) were similarly altered (**Supplementary Table S5**).

Comparison in Metabolome Between Treated Patients with CD and Their Unaffected Siblings

Last, the faecal metabolome between TCD patients and their unaffected siblings was compared to distinguish metabolic alterations specifically associated with CD from those related to shared environment and genetic risk. A total of 41 metabolites differed significantly between TCD and their unaffected siblings (19 decreased and 22 increased in the TCD group), as shown in **Supplementary Information Figure S6**. Comparison of siblings with HCs revealed 29 differential metabolites, including 18 decreased and 11 increased relative to HC. When these metabolites were compared with the CD-specific biomarker panel, there was no overlap, except hippuric acid. This indicates that the remaining 11 candidate biomarkers identified as CD-specific are independent of shared familial influences.

Discussion

While previous studies have characterised microbial alterations in CD, the corresponding functional consequences on the gut metabolome remain largely unexplored. In this study, we profiled the faecal metabolome of children with new-onset and treated CD, alongside HC and first-degree relatives using untargeted LC-MS metabolomics, to identify metabolites that may reflect disease processes or

treatment effects. The majority of CD metabolomics work to date has focussed on serum or plasma samples [21–25], however, there is a growing need to investigate intestinal disorders using faecal samples, given their non-invasive nature and ability to provide direct insights into gut environment. Faecal samples reflect both the composition and functional activity of the gut microbiota, while also capturing diet-derived substrates and host-derived metabolites that interact within the intestinal environment. This is particularly pertinent in CD, where microbial dysbiosis has been increasingly implicated in disease pathogenesis [6].

Our analysis revealed distinct clustering between HCs, UCD patients, and TCD patients, highlighting the presence of disease- and treatment- associated metabolic signatures. Consistent with the absence of pronounced microbial dysbiosis reported previously [6, 26], UCD patients showed largely unchanged global faecal metabolome profiles compared with HCs. However, subtle changes in specific metabolites were observed between the groups. UCD patients exhibited alterations in amino acid, bile acid, and lipid metabolites compared with HCs, suggesting shifts in nutrient absorption and microbial metabolism. Notably, UCD patients showed increased levels of acylcarnitines (acetylcarnitine, undecanoylcarnitine), amino acids and dipeptides (glycine, leucyl-prolyl-tyrosine), and the bile acid derivative nutriacholic acid. Acetylcarnitines are important in CD as markers of disrupted energy metabolism and mitochondrial function [27, 28]. Their altered levels reflect malabsorption, inflammation, and gut microbiota changes, making them potential non-invasive biomarkers for diagnosis and disease monitoring.

Importantly, three major patterns of metabolite profiles were identified, paralleling microbial groups previously described in this cohort [6]. The first group comprised CD-specific, non-treatment responsive metabolites that remain altered despite adherence to a GFD, reflecting core disease-associated metabolic signatures that are not fully corrected by dietary treatment. Among these, a consistent change was observed in 3-oxocholic acid, which was elevated in both UCD and TCD. This suggests that aspects of bile acid handling may remain altered in CD, potentially reflecting disease biology, impaired intestinal absorption, or microbial modulation of bile acids. Given the regulatory roles of bile acids in nutrient uptake, epithelial integrity, and immune signalling [29–31], such alterations may contribute to subtle but long-term disturbances in gut homeostasis. Also notable were widespread changes in amino acid and peptide metabolism. There persistent alterations observed in leu-pro-tyr, leucylproline, valylproline, 4-hydroxyprolylleucine, S-allyl-L-cysteine, and nicotinic acid point to long-term imbalances in protein turnover and amino acid metabolism. Previous studies have reported systemic alterations in amino acid profiles in CD patients [32, 33].

Changes in sterol profiles, highlighted by the alteration of 9(11)-dehydroaxinysterol, extend these findings to lipid-related metabolites. Although the specific biological role of 9(11)-dehydroaxinysterol is not well understood, sterols are key components of cell membranes and precursors for a wide range of signalling molecules [34, 35]. Sterol metabolism may intersect with other lipid-handling pathways, particularly involving carnitine and β -oxidation [36]. This connection likely arises through shared intermediates such as acyl-CoA, which serve as central nodes linking sterol modification, fatty acid oxidation, and energy metabolism. Consequently, disruptions in sterol homeostasis could contribute to

altered acyl-CoA flux, impaired carnitine shuttling, and mitochondrial dysfunction, further emphasising the broader metabolic consequences of sterol dysregulation. These findings suggest that changes in 9(11)-dehydroaxinysterol and related sterols may not only reflect shifts in membrane lipid composition but also signal deeper perturbations in mitochondrial lipid handling and energy balance. Altered sterol levels have previously been reported in newly diagnosed CD patients [37], suggesting that disturbances in sterol metabolism may be a recurring feature of CD.

Receiver operating characteristic analysis confirmed the predictive value of this metabolite panel, with the top 12 features achieving an AUC of 0.885, highlighting their potential utility as a screening diagnostic tool. Collectively, these results suggest that CD-specific metabolites represent fundamental perturbations across multiple metabolic classes that are only partially resolved by dietary treatment and may serve as robust biomarkers of disease presence independent of gluten exposure.

Metabolic dysregulation was also observed in individuals with treatment-responsive, new-onset CD. These metabolites were altered in UCD patients relative to HCs but normalise following treatment with a GFD, reflecting dietary modulation of disease-associated pathways. Comparison of TCD and UCD patients identified 27 metabolites significantly altered by dietary intervention. Most of these changes involved amino acid dipeptides and tripeptides, including Leu-Ala-Ser, Val-Leu, and Trileucine, which were consistently increased in TCD relative to UCD. The restoration of these small peptides may indicate remodelling of amino acid metabolism during dietary recovery. This could reflect improved intestinal absorption and mucosal repair but also changes in protein turnover, both from host digestive processes and microbial proteolysis, which together shape the availability of amino acids and peptides in the gut. These shifts may also relate to autophagy-mediated protein recycling, a process known to contribute to amino acid and peptide pools under conditions of cellular stress and recovery [38]. In parallel, treatment also altered microbial and purine metabolites, for example, indole-3-lactic acid, a tryptophan-derived microbial product, was reduced in TCD patients, suggesting shifts in microbial tryptophan catabolism with gluten exclusion. Indole-3-lactic acid is increasingly recognised as a key immunoregulatory and barrier-protective metabolite [39], capable of activating the aryl hydrocarbon receptor (AhR) to promote anti-inflammatory signalling, enhance epithelial repair, and regulate mucosal immune responses [40]. Similarly, reductions in the metabolite 1,7-dimethyluric acid, a purine derivative that can arise from both exogenous methylxanthines and endogenous purine metabolism, may reflect changes in nucleotide turnover or xanthine oxidase activity following treatment. Together, these findings highlight that dietary treatment not only restores aspects of amino acid metabolism but also reshapes microbial and nucleotide-related pathways, underscoring the complex metabolic adaptation that occurs following gluten withdrawal.

The third group comprised treatment-dependent, non-disease-specific metabolites that are primarily influenced by dietary intervention rather than disease status. Comparison of TCD patients with HCs revealed that 24 metabolites were significantly altered following GFD adherence, of which nearly half (46%) overlapped with metabolites also altered between TCD and UCD patients. This suggests that many of these changes occur independently of CD itself and are likely driven by the GFD. TCD patients

exhibited widespread alterations in amino acid metabolism, with reductions in tryptophan-derived metabolites highlighting perturbations in microbial and host pathways. The decrease in N-(indole-3-acetyl) leucine and xanthurenic acid indicates a shift in tryptophan metabolism, with microbial indole production and host kynurenine pathway activity both diminished. Given that CD is an immune-mediated disease, these changes are particularly notable. Tryptophan metabolism plays a central role in regulating mucosal immunity, largely through the indoleamine 2,3-dioxygenase (IDO) axis and the production of kynurenine-derived immunomodulatory metabolites. Reductions in these associated metabolites observed in TCD patients likely reflect a downregulation of this inflammatory response and a return toward immune homeostasis. These findings may point to remodelling of tryptophan utilisation under dietary treatment, consistent with dietary modulation of host-microbiota interactions [41]. Beyond tryptophan metabolism, several additional metabolites were altered between the groups. Reductions in 2-aminobutanoic acid, phenylalanine, isovaline, and lysopine point to broader perturbations in amino acid pathways. Decreased galactosylhydroxylysine, a collagen-derived metabolite, may reflect altered connective tissue turnover [42], while reduced N-acetylmuramic acid, a component of bacterial cell walls, may indicate shifts in host-microbial metabolic exchange [43]. The metabolic shifts observed with a GFD underscore the importance of considering diet-driven metabolic effects when interpreting faecal metabolomic profiles, particularly in TCD patients, where residual changes may not reflect disease activity but rather secondary dietary consequences. Persistent disturbances despite adherence may signal ongoing inflammation or incomplete remission, underscoring the need to map key metabolic changes to inflammatory markers and clinical outcomes. This is particularly important in non-responsive or refractory CD, where patients face higher risks of complications, including intestinal lymphoma and adenocarcinoma [44] and may require additional drug treatments, emphasising the value of early identification of this patient subgroup.

To gain further insights into CD, investigation was carried out into the metabolome of unaffected siblings by comparing them to the TCD group. Unaffected siblings are an informative reference group, as they share genetic susceptibility and often similar environmental exposures, but do not currently have disease. In this study, all siblings were confirmed to be tTG-negative through screening following the diagnosis of their affected relatives, supporting their classification as non-coeliac controls. By comparing their metabolomic profiles to those of treated CD patients, we can identify metabolic alterations that are associated with disease, rather than genetic susceptibility or lifestyle factors.

The outcomes observed from analysis of the prospective cohort build upon the findings obtained from the cross-sectional study. Looking into the effect of a GFD across 6-months, it was shown that aspartic acid, malic acid, and serine metabolites increased, while only ornithine was shown at higher levels after 12-months of treatment. This co-occurred with a decrease in tTG levels and an increase in PedsQL-GS score, indicating improvement of GI symptoms upon initiation of a GFD. While it is interesting to note that metabolic alteration was more profound at 6 months than at 12 months on a GFD, the same consensus was revealed from the microbiome analysis of the same CD study cohort [6]. The authors assumed that this can likely be explained by loss of adherence to GFD with time, as supported by the change in GIP levels, which further emphasises the difficulties associated with sustaining complete

exclusion of gluten in the diet. These findings do, however, contrast one study suggesting the metabolome fingerprint returns to normal after 12 months [22] as we did not observe complete metabolic restoration after this time.

This study has various limitations; firstly, there was a relatively modest sample size included in the prospective cohort. This is particularly relevant for analysis of faecal samples which are inherently complex and interindividual metabolite matrices, highly influenced by factors such as geographical location and dietary patterns. Although OPLS-DA models were used to explore multivariate separation between the groups (UCD vs HC and UCD vs TCD), both models demonstrated low predictive power. As such, these multivariate findings should be interpreted as exploratory and supported primarily by univariate analyses. Future studies with larger cohorts will be essential for improving the predictive stability and strengthening biological interpretations. Building the current findings through multi-omics approaches will provide a deeper understanding of the intricate interplay between the microbiome, metabolome, and disease to maximise clinical relevance.

Interpretation of the findings should also consider the duration of GFD exposure in the treated cohort. Although 12 months on a GFD is generally associated with substantial clinical and serological improvements, this time frame may not reflect complete mucosal or metabolic normalisation. Persistent inflammation, incomplete villous atrophy, or ongoing microbiome remodelling may still be present at this stage, potentially influencing the faecal metabolomic profile. Consequently, the TCD group may represent an early recovery or transitional metabolic state rather than a fully healed phenotype. Future studies incorporating longer-term follow-up and histological assessment of mucosal recovery would be valuable to determine whether the metabolomic signature observed here reflects long-term disease remission or ongoing biological adaptation.

Conclusion

In conclusion, we provide a comprehensive characterisation of the faecal metabolome in children with CD. While adherence to a GFD improved clinical parameters and partially restored the metabolome, a subset of metabolites remained persistently altered, highlighting a panel of CD-specific metabolites. By distinguishing CD-specific metabolites, treatment-responsive metabolites, and diet-driven metabolites, we reveal distinct metabolic signatures that reflect different aspects of disease biology and recovery. These results highlight persistent metabolic disturbances in treated CD that may contribute to ongoing mucosal injury, altered host-microbiota interactions, underscoring the need for strategies that address both disease pathophysiology and the effects of dietary restriction.

Declarations

Ethics Statement

The study was approved by the West of Scotland Research Ethics Committee (reference no.11/WS/0006). All participants and their carers provided written consent.

Financial support

Nutricia Research Foundation

Potential competing interests

The authors have no conflicts of interest to declare.

References

1. Singh, P., et al., *Global Prevalence of Celiac Disease: Systematic Review and Meta-analysis*. Clin Gastroenterol Hepatol, 2018. 16(6): p. 823–836.e2.
2. Mearns, E.S., et al., *Systematic Literature Review of the Economic Burden of Celiac Disease*. Pharmacoeconomics, 2019. 37(1): p. 45–61.
3. Bozorg, S.R., et al., *The Economic Iceberg of Celiac Disease: More Than the Cost of Gluten-Free Food*. Gastroenterology, 2024. 167(1): p. 172–182.
4. Di Biase, A.R., et al., *Gut microbiota signatures and clinical manifestations in celiac disease children at onset: a pilot study*. J Gastroenterol Hepatol, 2021. 36(2): p. 446–454.
5. Olivares, M., et al., *Gut microbiota trajectory in early life may predict development of celiac disease*. Microbiome, 2018. 6(1): p. 36.
6. Zafeiropoulou, K., et al., *Alterations in Intestinal Microbiota of Children With Celiac Disease at the Time of Diagnosis and on a Gluten-free Diet*. Gastroenterology, 2020. 159(6): p. 2039–2051.e20.
7. El Mouzan, M., A. Assiri, and A. Al Sarkhy, *Gut microbiota predicts the diagnosis of celiac disease in Saudi children*. World J Gastroenterol, 2023. 29(13): p. 1994–2000.
8. Leonard, M.M., et al., *Microbiome signatures of progression toward celiac disease onset in at-risk children in a longitudinal prospective cohort study*. Proc Natl Acad Sci U S A, 2021. 118(29).
9. Girdhar, K., et al., *Dynamics of the gut microbiome, IgA response, and plasma metabolome in the development of pediatric celiac disease*. Microbiome, 2023. 11(1): p. 9.
10. Murch, S., et al., *Joint BSPGHAN and Coeliac UK guidelines for the diagnosis and management of coeliac disease in children*. Archives of Disease in Childhood, 2013. 98(10): p. 806–811.
11. Varni, J.W., et al., *PedsQL gastrointestinal symptoms module: feasibility, reliability, and validity*. J Pediatr Gastroenterol Nutr, 2014. 59(3): p. 347–55.
12. Gerasimidis, K., et al., *Comparison of Clinical Methods With the Faecal Gluten Immunogenic Peptide to Assess Gluten Intake in Coeliac Disease*. J Pediatr Gastroenterol Nutr, 2018. 67(3): p. 356–360.
13. Kelly, P.E., et al., *An Optimised Monophasic Faecal Extraction Method for LC-MS Analysis and Its Application in Gastrointestinal Disease*. Metabolites, 2022. 12(11).

14. Broadhurst, D.I., *QC:MXP Repeat Injection based Quality Control, Batch Correction, Exploration & Data Cleaning (version 1)* Zendono. <https://zenodo.org/doi/10.5281/zenodo.11101541>. Retrieved from <https://github.com/broadhurstdavid/QC-MXP>. 2024.
15. Oksanen J, S.G., Blanchet F, et al., *vegan: Community Ecology Package*. R package version 2.6–10. 2025.
16. Thévenot, E.A., et al., *Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses*. J Proteome Res, 2015. 14(8): p. 3322–35.
17. Westerhuis, J.A., et al., *Multivariate paired data analysis: multilevel PLSDA versus OPLSDA*. Metabolomics, 2010. 6(1): p. 119–128.
18. Robin, X., et al., *pROC: an open-source package for R and S + to analyze and compare ROC curves*. BMC Bioinformatics, 2011. 12: p. 77.
19. Friedman, J., T. Hastie, and R. Tibshirani, *Regularization Paths for Generalized Linear Models via Coordinate Descent*. J Stat Softw, 2010. 33(1): p. 1–22.
20. Sumner, L.W., et al., *Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI)*. Metabolomics, 2007. 3(3): p. 211–221.
21. Solakivi, T., et al., *Serum fatty acid profile in celiac disease patients before and after a gluten-free diet*. Scand J Gastroenterol, 2009. 44(7): p. 826–30.
22. Bertini, I., et al., *The metabonomic signature of celiac disease*. J Proteome Res, 2009. 8(1): p. 170–7.
23. Lewis, N.R., et al., *Cholesterol profile in people with newly diagnosed coeliac disease: a comparison with the general population and changes following treatment*. Br J Nutr, 2009. 102(4): p. 509–13.
24. Upadhyay, D., et al., *Abnormalities in metabolic pathways in celiac disease investigated by the metabolic profiling of small intestinal mucosa, blood plasma and urine by NMR spectroscopy*. NMR Biomed, 2020. 33(8): p. e4305.
25. Jamnik, J., D.J. Jenkins, and A. El-Sohemy, *Biomarkers of cardiometabolic health and nutritional status in individuals with positive celiac disease serology*. Nutr Health, 2018. 24(1): p. 37–45.
26. Aboulaghras, S., et al., *Meta-Analysis and Systematic Review of HLA DQ2/DQ8 in Adults with Celiac Disease*. Int J Mol Sci, 2023. 24(2).
27. Dambrova, M., et al., *Acylcarnitines: Nomenclature, Biomarkers, Therapeutic Potential, Drug Targets, and Clinical Trials*. Pharmacological Reviews, 2022. 74(3): p. 506–551.
28. Bene, J., et al., *Plasma carnitine ester profile in adult celiac disease patients maintained on long-term gluten free diet*. World Journal of Gastroenterology, 2005. 11(42): p. 6671–6675.
29. Chiang, J.Y.L. and J.M. Ferrell, *Bile Acids as Metabolic Regulators and Nutrient Sensors*. Annu Rev Nutr, 2019. 39: p. 175–200.
30. Sun, R., et al., *Critical roles of bile acids in regulating intestinal mucosal immune responses*. Therap Adv Gastroenterol, 2021. 14: p. 17562848211018098.

31. Shi, L., L. Jin, and W. Huang, *Bile Acids, Intestinal Barrier Dysfunction, and Related Diseases*. Cells, 2023. 12(14).
32. Khalkhal, E., et al., *Screening of Altered Metabolites and Metabolic Pathways in Celiac Disease Using NMR Spectroscopy*. Biomed Res Int, 2021. 2021: p. 1798783.
33. Drabińska, N., et al., *Plasma profile and urine excretion of amino acids in children with celiac disease on gluten-free diet after oligofructose-enriched inulin intervention: results of a randomised placebo-controlled pilot study*. Amino Acids, 2018. 50(10): p. 1451–1460.
34. Dufourc, E.J., *Sterols and membrane dynamics*. J Chem Biol, 2008. 1(1–4): p. 63–77.
35. Wollam, J. and A. Antebi, *Sterol regulation of metabolism, homeostasis, and development*. Annu Rev Biochem, 2011. 80: p. 885–916.
36. Chandel, N.S., *Lipid Metabolism*. Cold Spring Harb Perspect Biol, 2021. 13(9).
37. Vuoristo, M. and T.A. Miettinen, *Serum cholesterol precursor sterols in coeliac disease: effects of gluten free diet and cholestyramine*. Gut, 1986. 27(11): p. 1312–9.
38. Maiuri, L., et al., *Defective proteostasis in celiac disease as a new therapeutic target*. Cell Death & Disease, 2019. 10(2): p. 114.
39. Meng, D., et al., *Indole-3-lactic acid, a metabolite of tryptophan, secreted by Bifidobacterium longum subspecies infantis is anti-inflammatory in the immature intestine*. Pediatric Research, 2020. 88(2): p. 209–217.
40. Rocha, C.S., et al., *Microbial remodeling of gut tryptophan metabolism and indole-3-lactate production regulate epithelial barrier repair and viral suppression in human and simian immunodeficiency virus infections*. Mucosal Immunology, 2025. 18(3): p. 583–595.
41. Sinha, A.K., et al., *Dietary fibre directs microbial tryptophan metabolism via metabolic interactions in the gut microbiota*. Nature Microbiology, 2024. 9(8): p. 1964–1978.
42. Moro, L., et al., *The glycosides of hydroxylysine are final products of collagen degradation in humans*. Biochim Biophys Acta, 1993. 1156(3): p. 288–90.
43. Li, C., et al., *Gut microbiota-derived GlcNAc-MurNAc is a TLR4 agonist that protects the host gut*. Nature Communications, 2025. 16(1): p. 5577.
44. Demiroren, K., *Possible relationship between refractory celiac disease and malignancies*. World J Clin Oncol, 2022. 13(3): p. 200–208.

Figures

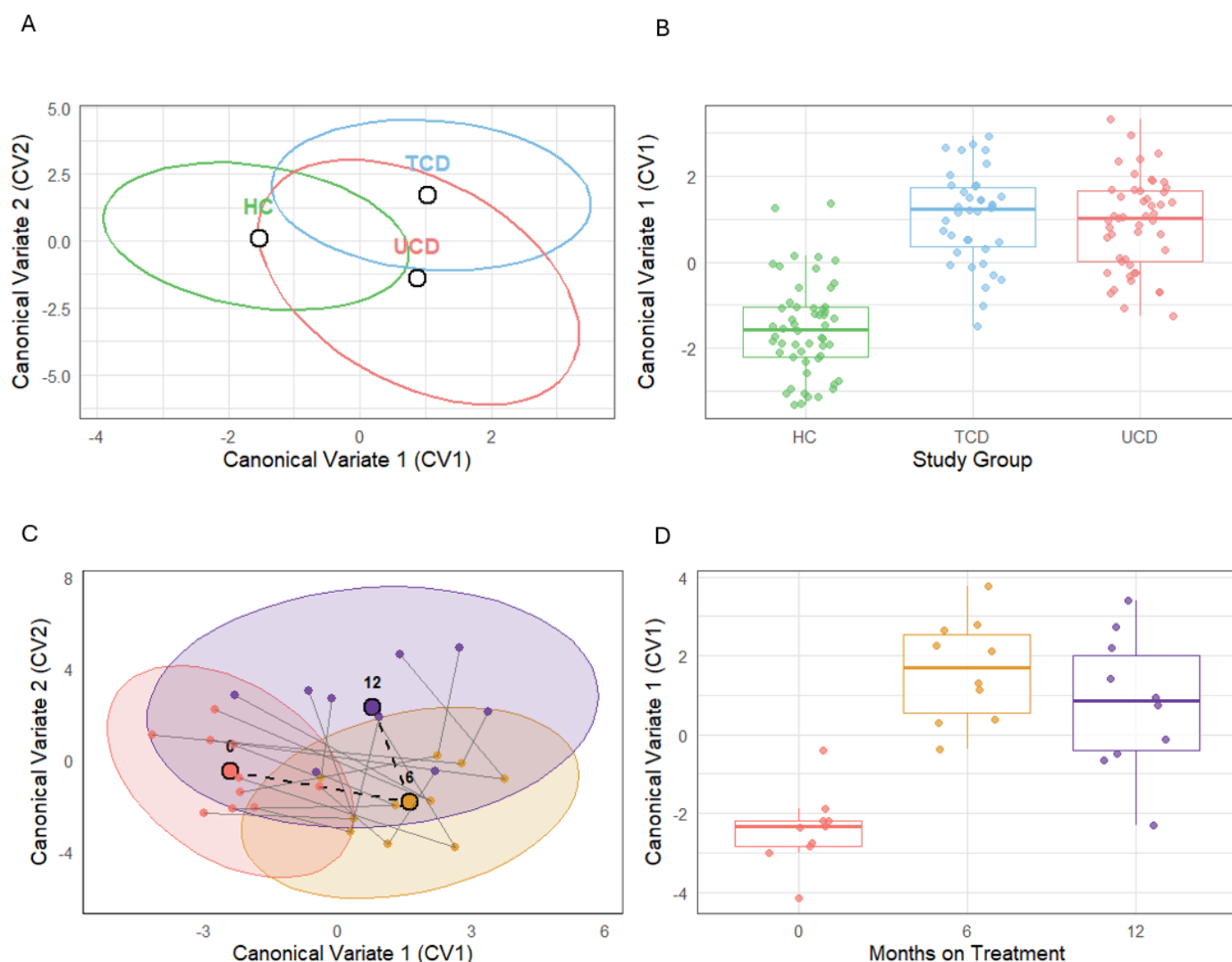


Figure 1

Faecal metabolome profiles in the cross-sectional and prospective cohorts. (A) Principal Coordinates–Canonical Variate Analysis (PC-CVA) of the faecal metabolome across cross-sectional study groups (TCD, UCD, and HC). Group centroids and 95% confidence ellipses are shown. (B) Boxplots of canonical variate 1 (CV1) values corresponding to panel A. (C) PC-CVA of the faecal metabolome in the prospective cohort at baseline and after 6 and 12 months on a gluten-free diet (GFD). Grey lines represent individual patient trajectories across timepoints. (D) Boxplot of CV1 values corresponding to panel C.

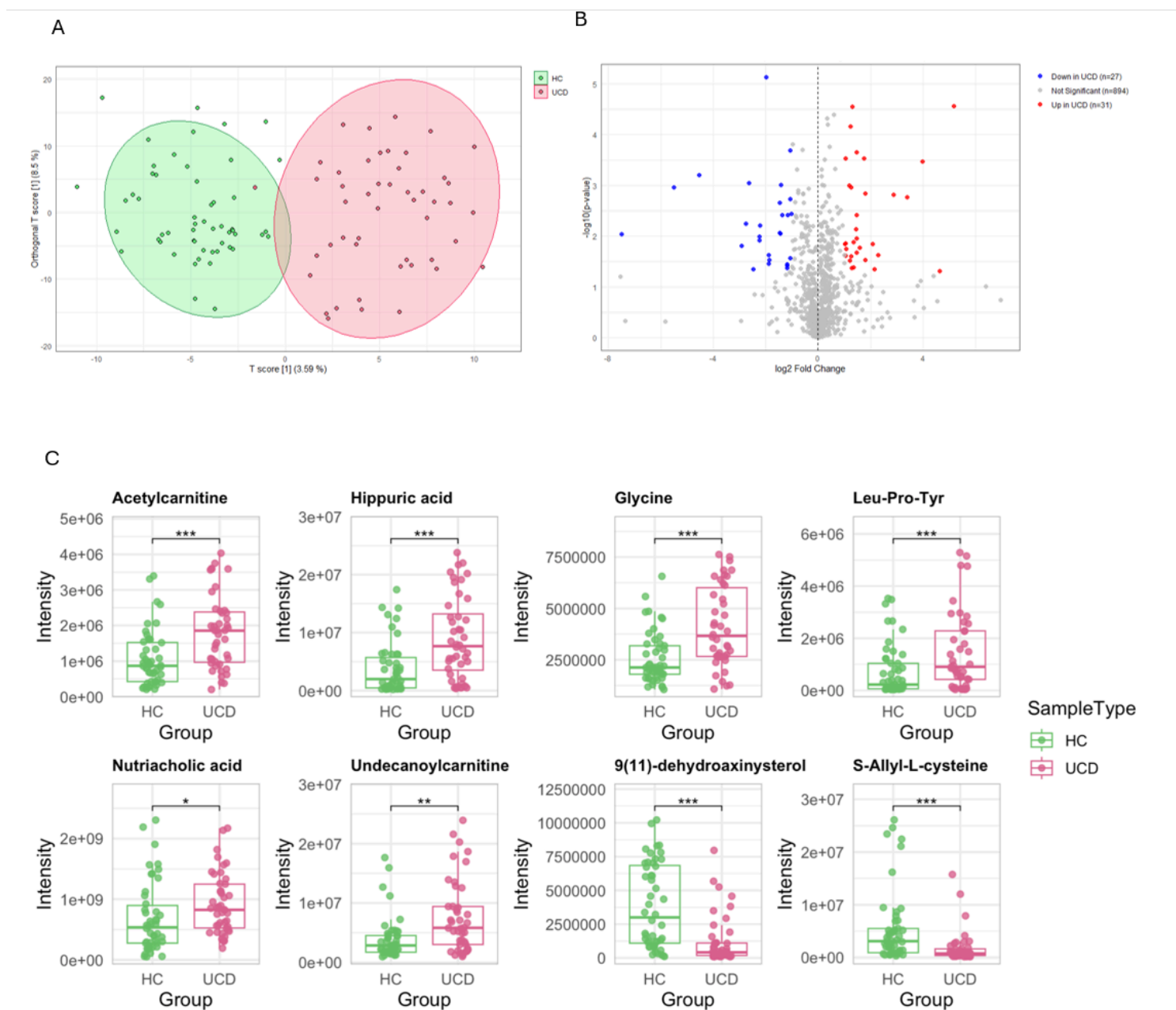


Figure 2

Metabolome profile of children with newly diagnosed coeliac disease compared to healthy controls (n = 82). (A) Scores plot of the orthogonal partial least square discriminant analysis (OPLS-DA) model, $R^2Y = 0.555$, $Q^2 = 0.267$. (B) Volcano plot of significantly differential faecal metabolites, $p < 0.05$, fold change = 2. (C) Box and whisker plots of the top significantly differential metabolites, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

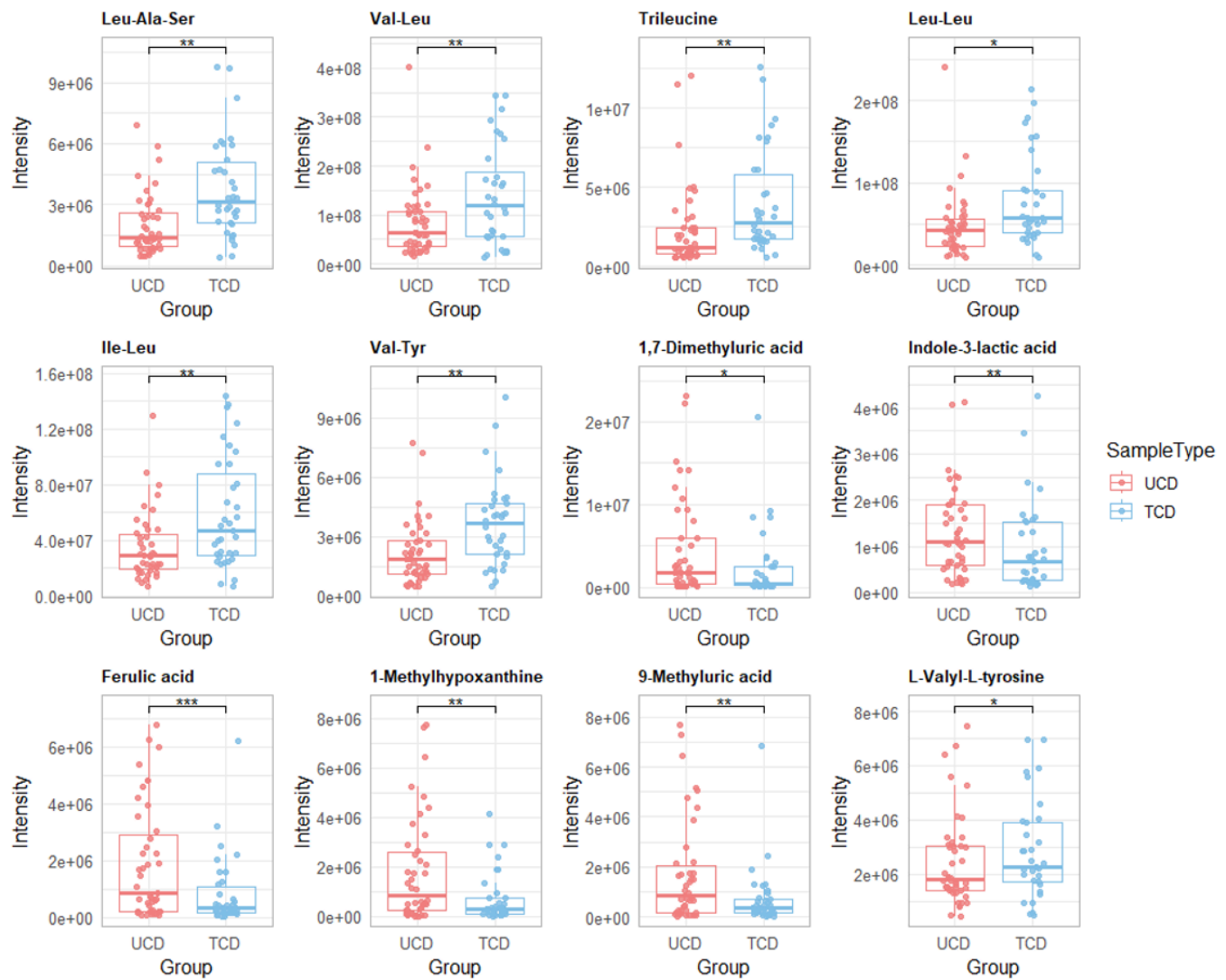
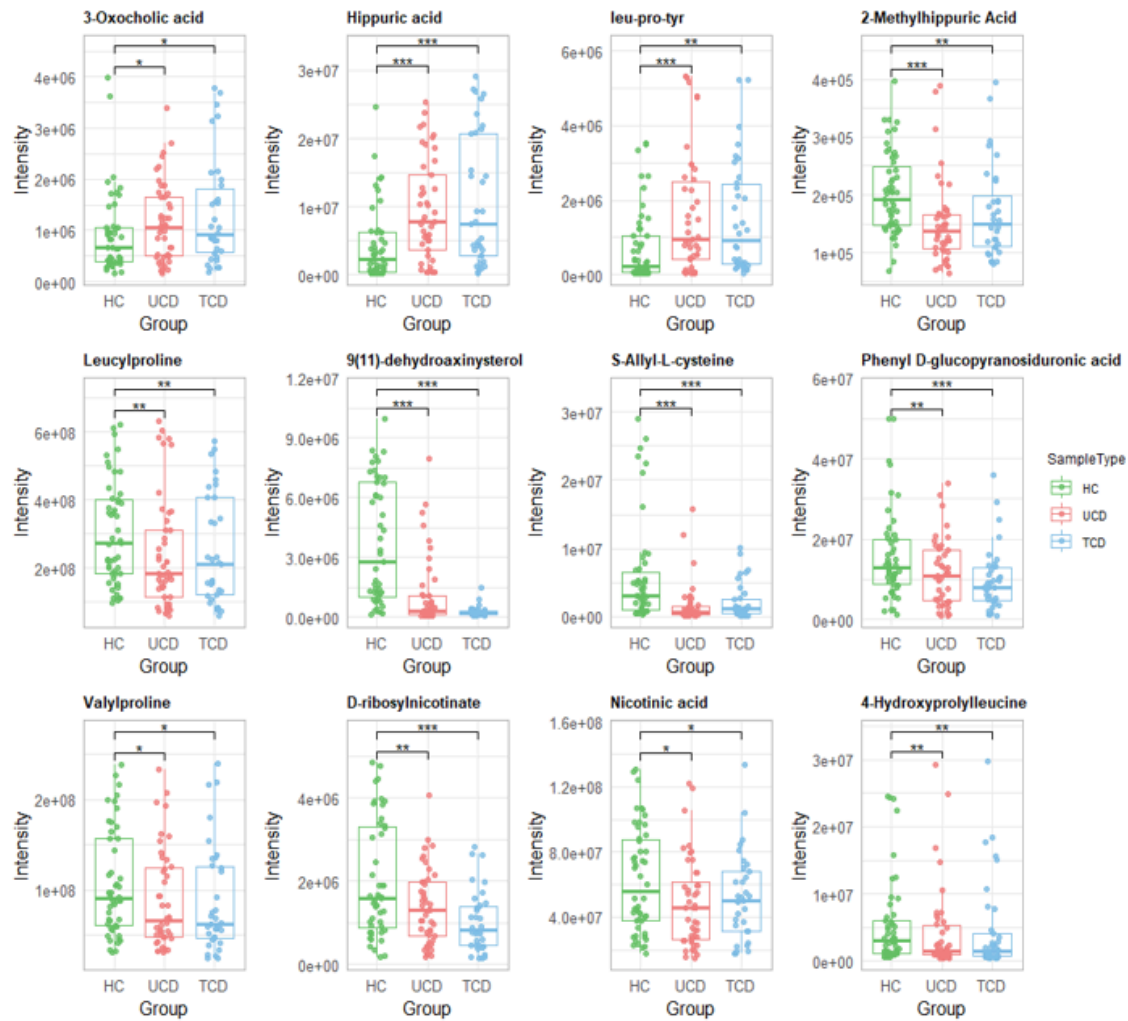


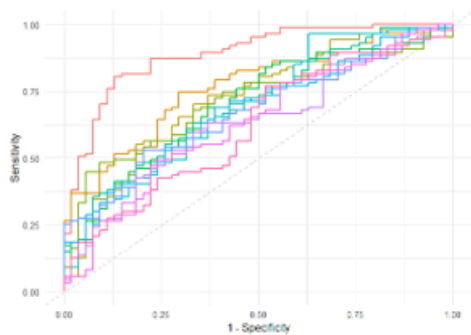
Figure 3

Treatment-responsive faecal metabolites in patients with CD. Boxplots showing the relative intensities of metabolites significantly altered between untreated coeliac disease (UCD) and treated coeliac disease (TCD). Statistical significance is indicated by *p < 0.05, **p < 0.01, ***p < 0.001.

A



B



C

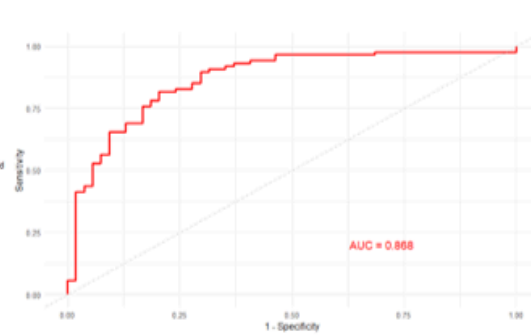


Figure 4

Coeliac disease-specific metabolites. (A) Representative boxplots showing relative intensities of the top metabolites significantly altered in both untreated (UCD) and treated (TCD) patients compared with HCs. (B) Receiver operating characteristic (ROC) curves for the top 12 discriminant metabolites. (C) Combined ROC curve generated from a multivariate logistic regression model incorporating all 12 metabolites, AUC = 0.868.

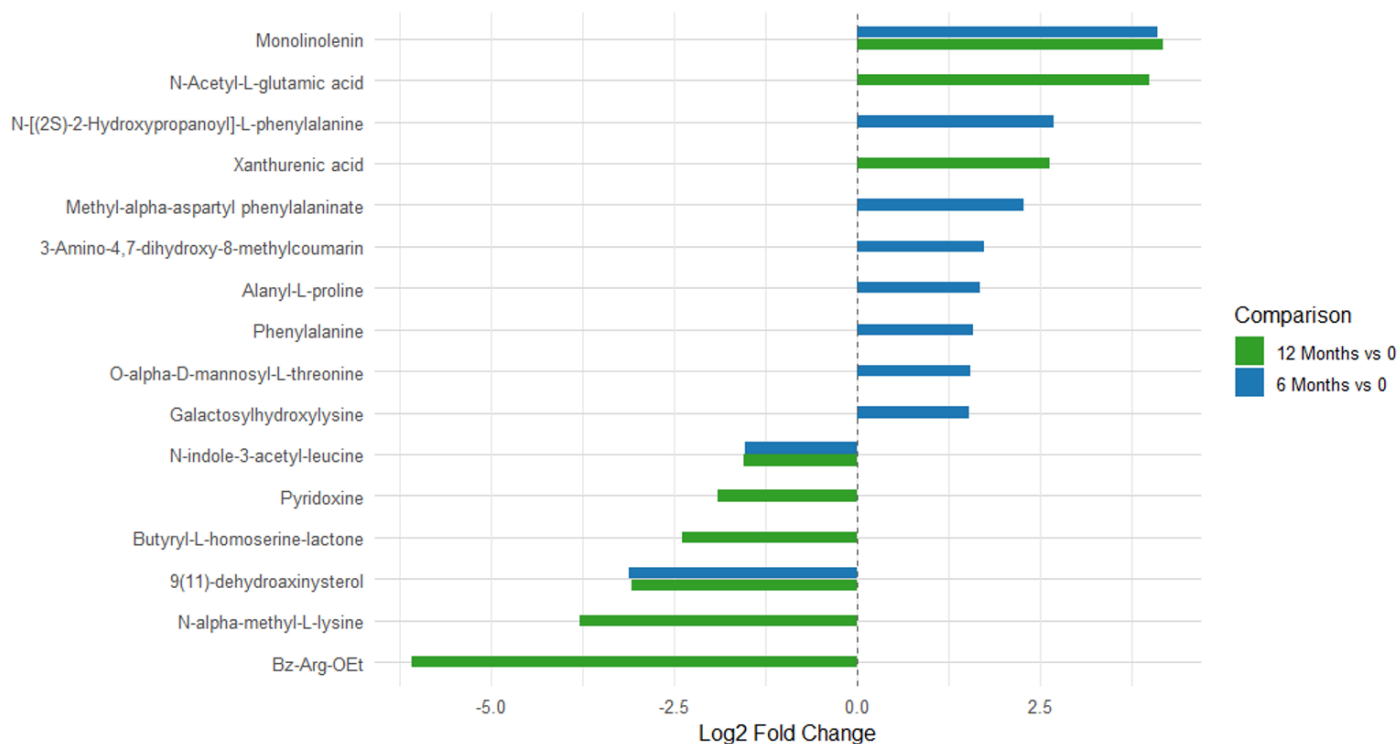


Figure 5

Statistically significant differences (log2 fold change) in metabolite levels between coeliac disease diagnosis and at 6- and 12-months follow-up after starting a gluten-free diet.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.docx](#)