Alterations in Intestinal Microbiota of Children With Celiac Disease at Time of Diagnosis and on a Gluten-Free Diet

Konstantina Zafeiropoulou, Ben Nichols, Mary Mackinder, Olga Biskou, Eleni Rizou, Antonia Karanikolou, Clare Clark, Elaine Buchanan, Tracey Cardigan, Hazel Duncan, David Wands, Julie Russell, Richard Hansen, Richard K. Russell, Paraic McGrogan, Christine A. Edwards, Umer Z. Ijaz, Konstantinos Gerasimidis

PII: S0016-5085(20)35023-X
DOI: https://doi.org/10.1053/j.gastro.2020.08.007
Reference: YGAST 63676

To appear in: Gastroenterology
Accepted Date: 2 August 2020

Please cite this article as: Zafeiropoulou K, Nichols B, Mackinder M, Biskou O, Rizou E, Karanikolou A, Clark C, Buchanan E, Cardigan T, Duncan H, Wands D, Russell J, Hansen R, Russell RK, McGrogan P, Edwards CA, Ijaz UZ, Gerasimidis K, Alterations in Intestinal Microbiota of Children With Celiac Disease at Time of Diagnosis and on a Gluten-Free Diet, Gastroenterology (2020), doi: https://doi.org/10.1053/j.gastro.2020.08.007.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 by the AGA Institute
Title: Alterations in Intestinal Microbiota of Children With Celiac Disease at Time of Diagnosis and on a Gluten-Free Diet

Konstantina Zafeiropoulou1, Ben Nichols1, Mary Mackinder1, Olga Biskou1, Eleni Rizou1, Antonia Karanikolou1, Clare Clark1, Elaine Buchanan2, Tracey Cardigan2, Hazel Duncan2, David Wands2, Julie Russell3, Richard Hansen2, Richard K Russell2, Paraic McGrogan2, Christine A Edwards1, Umer Z Ijaz3, Konstantinos Gerasimidis1*

1 Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life Sciences, University of Glasgow, New Lister Building, Glasgow Royal Infirmary, G31 2ER, Glasgow
2 Department of Paediatric Gastroenterology, Royal Hospital for Children, Glasgow
3 Civil Engineering, School of Engineering, Oakfield Avenue, University of Glasgow, Glasgow

*Corresponding author

Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life Sciences, University of Glasgow, New Lister Building, Glasgow Royal Infirmary, G31 2ER, Glasgow;

Tel: 0044 141 201 8689

Email: Konstantinos.gerasimidis@glasgow.ac.uk

Conflicts of interest

R Hansen has received speaker’s fees, conference support or consultancy fees from Nutricia and 4D Pharma. RK Russell has received speaker’s fees, travel support, and participated in medical board meetings with Abbvie, Janssen, Takeda, Celltrion, Pharmacosmos and Nestle. K Gerasimidis reports personal fees from Nutricia, research grants and personal fees from Nestle, personal fees from Dr Falk. C Edwards is chair of working group for ILSI Europe. The rest of the authors have no conflicts of interest to disclose.

Author contributions

KZ: performed measurements of short chain fatty acids, part of bioinformatics and statistical analysis and produced the first draft of the paper; BN performed bioinformatics and statistical analysis and
produced the paper illustrations; MM designed the study, applied for ethics, recruited participants collected some of the samples; OB, AK, CC collected samples and performed part of the laboratory analysis for measurements of microbiota, lactic acid, sulfides, ER performed the dietary assessment; EB, TC, HD contributed to participants’ recruitment and clinical follow-up; JR performed the 16S rRNA sequencing; RH, RKR co-ordinated the clinical activities and collected clinical data, PMc, CAE supervised the student, designed the study, applied for ethical permission, co-ordinated activities; UZI trained and supervised KZ in bioinformatics; DW collected clinical information for patients; KG conceived and designed the study, applied for ethical application and funding award, co-ordinated the field and laboratory activities, trained and supervised the field researchers and edited the first draft for publication. All authors reviewed the final version of the manuscript and agreed to its content prior submission. 5 indicates shared first authorship

FUNDING

This study was funded by Nutricia Research Foundation; Dr Ben Nichols was funded by a grant from the Biotechnology and Biological Sciences Research Council (BB/R006539/1) and The Catherine McEwan Foundation.
ABSTRACT

Background & Aims: It is not clear whether alterations in the intestinal microbiota of children with celiac disease cause the disease or are a result of disease and/or its treatment with gluten-free diet (GFD).

Methods: We obtained 167 fecal samples from 141 children (20 with new-onset celiac disease, 45 treated with a GFD, 57 healthy children, and 19 unaffected siblings of children with celiac disease) in Glasgow, Scotland. Samples were analyzed by 16S rRNA sequencing and diet-related metabolites were measured by gas chromatography. We obtained fecal samples from 13 of the children with new-onset CD after 6 and 12 months on GFD. Relationships between microbiota with diet composition, gastrointestinal function, and biomarkers of GFD compliance were explored.

Results: Microbiota α diversity did not differ among groups. Microbial dysbiosis was not observed in children with new-onset celiac disease. In contrast, 2.8% (Bray-Curtis dissimilarity index, P=.025) and 2.5% (UniFrac distances, P=.027) of the variation in microbiota composition could be accounted for by the GFD. Between 3% to 5% of all taxa differed among all group comparisons. Eleven distinctive operational taxonomic units composed a microbe signature specific to celiac disease with high diagnostic probability. Most of the operational taxonomic units that differed between patients on GFD with new-onset celiac disease vs healthy children were associated with nutrient and food group intake (from 75% to 94%), and with biomarkers of gluten ingestion. Fecal levels of butyrate and ammonia decreased during the GFD.

Conclusions: Although several alterations in the intestinal microbiota of children with established celiac disease appear to be effects of a GFD, there are specific bacteria that are distinct biomarkers of celiac disease. Studies are needed to determine whether these bacteria contribute to pathogenesis of celiac disease.

Keywords: OTU, pediatric, microbiome, short chain fatty acids
INTRODUCTION

Celiac disease (CD) is an autoimmune destruction of small intestinal villi triggered by ingestion of gluten in genetically susceptible individuals. It causes nutrient malabsorption, leading to intestinal and extra-intestinal symptoms. Lifelong adherence to a gluten-free diet (GFD) is the only treatment. The underlying pathogenesis of CD is multifactorial and while genetic predisposition occurs in 30-40% of the general population, only a small proportion of these individuals will develop CD, suggesting that environmental factors are at play. This premise is supported by the rise in the incidence of CD over past decades, suggesting that changes in population genetics cannot explain this increase, nor can changes in gluten consumption. Weaning practices, antibiotic exposure, and viral gastrointestinal (GI) infections have been implicated as risk factors for CD onset. Recent evidence points to the involvement of the gut microbiota; driven by the rapid expansion of microbiota research.

Several non-communicable diseases, like inflammatory bowel disease, show increasing incidence similar to CD, and have been associated with distinct features of microbiota structure and function, also termed ‘dysbiosis’. There has been increasing interest in research exploring the role of the microbiota in CD. However, results remain inconclusive. Previous research was mostly of cross-sectional design providing a snapshot of the gut microbiota in CD and most importantly, did not determine whether a disturbed microbiota in feces of new-onset, untreated patients (UCD) is implicated in disease pathogenesis or is predominantly an epiphenomenon of the underlying disease, altered gastrointestinal (GI) motility and excessive substrate availability, owing to nutrient malabsorption and intestinal inflammation. Research has also described an altered microbiota in treated patients with CD (TCD) but did not explore the extent to which these signals were attributed to dietary modification and exclusions imposed during treatment with GFD. Studies describing the human fecal microbiota of CD using next-generation sequencing do not exist and previous research relied on characterization of selective microbial groups.

This study characterized the gut microbiota of children with CD. We combined cross-sectional and prospective patient cohorts and collected data on determinants of the microbiota likely to change in children with CD during GFD, or likely to differ when comparing CD with healthy controls (HC). For the first time in the literature, we profiled the fecal microbiota using 16S rRNA gene sequencing and measured metabolites likely to be influenced by adherence to GFD. Additionally, we explored associations with dietary intake, GI symptoms and novel biomarkers of GFD compliance.
We hypothesized that the gut microbiota of CD patients differs from HC; several of these microbial signals are secondary effects of dietary modifications during GFD, but others may be implicated in the underlying disease pathogenesis.
MATERIAL & METHODS

Subjects

Fecal samples were collected from children with CD attending the Royal Hospital for Children in Glasgow. New-onset patients with CD were referred from primary healthcare services whereas previously diagnosed children were recruited from clinics which patients attend annually. Siblings of CD children with no clinical symptoms and negative tissue transglutaminase IgA antibodies (tTG-IgA), and healthy volunteers recruited via advertisement were used for comparative analysis. In these two groups, healthy status was defined as children who did not visit their general practitioner for a medical condition regularly or who did not receive regular medication, and who had no past history of chronic GI disorders. All consecutive children were invited to participate unless they met one of the exclusion criteria (i.e. antibiotics use or regular use of probiotics/prebiotics in the preceding 3 months. Patients with other comorbidities were also excluded.

Celiac disease was confirmed by small bowel biopsy using UK guidelines in place at the time of recruitment. Dietary intake (macronutrient and food groups) was evaluated using a Scottish food frequency questionnaire (FFQ). GI symptoms were evaluated using the PedsQL-GS questionnaire (version 1). The higher the PedsQL-GS score, the lower GI symptoms are. Recent gluten ingestion, a proxy marker of GFD adherence, was evaluated by measuring the fecal gluten immunogenic peptide (GIP) levels (iVYLISA, Biomedal, Spain).

Fecal sample collection

The entire bowel movement was collected, stored under anaerobic (Anaerocult® A, Merck, Germany) cold conditions and processed within 2 hours of defecation. The entire sample was homogenized with a hand-blender and stored appropriately for downstream analysis. Except for the UCD patients a single sample was collected. For the UCD group we aimed to collect 3 samples; one baseline sample prior to diagnostic endoscopy while the patients were on a gluten-containing diet, and again at 6 and 12 months on GFD.

Fecal water content, pH and ammonia

Fecal pH was measured in aqueous slurries and fecal ammonia with an analyzer (Hanna HI 93715, Bedfordshire, United Kingdom). Fecal water content was calculated after lyophilization.
Microbiota profiling

Genomic DNA was isolated using the chaotropic method within 3 months of sample collection\textsuperscript{24}. Sequencing of the V4 region of the 16S rRNA gene was performed with MiSeq (Illumina) using the Golay barcodes on the reverse strand\textsuperscript{18, 25}. Barcoded amplicons were purified using the Zymoclean Gel DNA Recovery Kit (D4001).

Fecal SCFA, lactate and sulfide

The short chain fatty acids (SCFA) acetic, propionic, butyric and valeric acids, and the branch chain fatty acids (BCFA) (isobutyric and isovaleric acids) were measured using gas chromatography (Agilent 7820A) with a DB–WAX UI column on diethyl ether extracts\textsuperscript{23, 25}. Nitrogen was the carrier gas. D- and L-lactate were measured using a commercial kit (D, L Lactic Acid, Roche, Cat No; 11112821035) scaled-down for use in a 96-well plate. Free and total sulfide were measured colorimetrically\textsuperscript{23}.

Bioinformatics analysis

Quality trimming was done using Sickle, applying a sliding window approach and trimming regions where average base quality (PHRED score) dropped below 20\textsuperscript{26}. Assembly of paired reads was done using PANDAseq\textsuperscript{27}. USEARCH was used for dereplication and clustering of sequences into operational taxonomic units (OTUs) of 97% similarity as described in https://docs.google.com/document/d/1BcZAk28k7Uycr7iKKAVSiZ0MB9jDs9bODpdPZtYFH3Y/pub#h.agz7wlf8m6. Chimera detection involved de novo and reference-based steps, using the ChimeraSlayer gold database (http://drive5.com/uchime/uchime_download.html) derived from the ChimeraSlayer reference database (http://microbiomeutil.sourceforge.net/). OTUs were taxonomically classified using QIIME with SILVA reference database (version 123). An approximately-maximum-likelihood phylogenetic tree was produced using the ‘ginsi’ alignment algorithm in MAFFT\textsuperscript{28}, followed by FastTree\textsuperscript{29}.

Statistical analysis

General linear models on Box-Cox transformed data were used to compare groups, accounting for the paired design of the prospective cohort, and adjusted using Bonferroni correction. We first compared the fecal metabolite concentrations and microbiota between HC and siblings of the CD patients. In the absence of significant difference, we removed the sibling group from further analysis. Multivariate statistical analysis was performed in R (version 3.4.0) using the packages
vegan, phyloseq and DESeq2. Samples were rarefied to 5,000 reads before calculating α diversity, measured as species richness and Shannon index. Microbial composition was assessed using non-metric multidimensional scaling plots (NMDS) at genus and OTU level based on Bray-Curtis dissimilarity indices and unweighted UniFrac distances. The former considers bacterial taxon abundance, while the latter considers phylogenetic distances between bacterial taxa through presence/absence, regardless of proportional representation. Permutation ANOVA was applied using the vegan Adonis function on distance matrices (Bray-Curtis/ unweighted UniFrac) with data stratified by subject to allow for repeated sampling from UCD participants during follow-up. Local contribution to β-diversity (LCBD) analysis was performed to measure the contribution of each sample to the total OTU β diversity; samples with high LCBD represent samples that are markedly different from the average β diversity of all study samples. Differences in OTU abundances between groups were found using the DESeq2 method, with participant ID included as a variable in the input formula for paired data. For correlations Kendall rank correlation was used. Benjamini-Hochberg correction was applied to cases of multiple testing. Analysis using the Bioenv function in vegan produced subsets of OTUs whose Euclidean distance matrices correlate maximally with the Bray-Curtis dissimilarity matrices derived from complete OTU tables, thus indicating major determinants of community structure.

Random forest analysis used OTUs that significantly differed in abundance between HC versus both UCD and TCD patients inclusive. Models were generated using Log-proportional abundances via the randomForest R package, with 10,000 decision trees used per model. Model performance was assessed using the rf.significance function in the rfUtilities R package and ROC analysis using the ROCR R package.

**Ethical considerations**

The study was approved by the West of Scotland Research Ethics Committee (Ref:11/WS/0006). All authors had access to the study data and reviewed and approved the final manuscript.
RESULTS

Participants

141 children participated including, 45 TCD children on GFD, 20 UCD children on gluten containing diet, 19 siblings of 18 TCD children and 57 HC (Table 1). Thirty-three eligible participants (12 females) declined participation; 10 other (6 females) CD patients did not meet the inclusion criteria (i.e. 3 unable to comprehend English, 1 had developmental delay, 1 child was in foster care, 4 had Type 1 diabetes, 1 had congenital hypothyroidism). There was no difference in age (p=0.11) or gender (p=0.87) between participants and those who declined. All healthy children who expressed an interest participated in the study.

All UCD children were recommended to follow a GFD. From the 20 UCD patients with baseline fecal samples, 13 (65%) provided follow-up samples at 6 and 12 months after GFD initiation (prospective cohort); 4 (20%) patients were lost at follow-up and 3 (15%) others provided paired samples at 12 months only but were subsequently removed to avoid bias in statistical analysis. There was no difference in age (p=0.27), gender (p=0.998) or BMI z-scores (p=0.63) as well as in α and β diversity of the baseline microbiota between the 13 patients with follow-up samples and the 7 others who did not provide all samples. In total, 167 fecal samples were collected across all groups.

After commencement of a GFD, tTG-IgA titer decreased in the UCD children (Table 1). The UCD group experienced more GI problems than TCD and HC. The mean PedsQL-GS score was also lower in TCD than HC, suggesting that TCD children had more GI symptoms than HC. In the UCD group, GI symptoms improved only 12 months post-diagnosis.

Thirty-eight of 45 TCD children (84%) had undetectable GIP, indicating at least recent compliance with GFD; the remaining 7 (16%) TCD children had detectable levels indicating either transgression from GFD recommendations or accidental exposure to gluten (Table 1). Compared to baseline, GIP concentration was almost 13 times lower at 6 months and 6 times lower at 12 months on GFD. At 6 and 12 months after recommendation to adhere to a GFD, 2 (15%) and 3 out of 13 (23%) UCD patients had detectable GIP.

Fecal microbiota profiling

There was no difference in α-diversity between TCD, UCD and HC groups, neither at OTU nor genus level (Figure 1a & Supplementary Table 1). In contrast, 2.8% (p=0.025) and 2.5% (p=0.027) of the variation in OTU community structure (β-diversity) for the Bray-Curtis dissimilarity index and unweighted UniFrac distance analyses, respectively, were explained by participant grouping. TCD clustered separately to HC and tended to do so with respect to UCD (Supplementary Table 1),
suggesting a significant effect of GFD on microbiota structure. Similar findings were observed at genus level. Gender did not influence this effect. No separation in community structure was seen between the HC and UCD groups, suggesting an absence of profound dysbiosis at disease onset. LCBD analysis confirmed that the microbiota structure of TCD individuals differed from that of UCD and HC, with no difference seen between the latter 2 groups (Figure 1a).

Using the Bioenv workflow, a subset of 13 OTUs maintained the group clustering, explaining 92.3% of the variance described by the complete OTU dataset (Supplementary Table 2). When the Bioenv analysis was repeated for pairings of UCD vs TCD and HC vs TCD, separately, subsets of 14 and 12 OTUs were retrieved, explaining 92.0% and 91.4% of the variance described by the complete datasets, respectively. Of note, 9 OTUs featured in both pairwise comparisons, suggesting their strong influence on the fecal microbiota structure of children with CD on a GFD. Compared to disease diagnosis, there was no difference in β-diversity of the 13 UCD children at 6 and 12 months after initiation of GFD (Figure 1b).

**Differential analysis in OTU abundance between new-onset, untreated CD and HC**

The UCD and HC groups were characterized by a total of 1,033 distinct OTUs. Thirty-one OTUs (3%) differed significantly between the 2 groups; all of which had a significantly lower abundance in UCD than HC (Figure 2). Of these 31 discriminatory OTUs, only the abundance of OTU_1054 Alistipes correlated positively with PedsQL-GS score; thus suggesting that the remaining 30 discriminatory OTUs were less likely to be explained by differences in GI symptoms between the 2 groups.

**Differential analysis in OTU abundance between patients on recommendation to a GFD with untreated CD or HC**

Next, we explored the effect GFD might have on the CD microbiota. First, we looked for differences in OTU abundances between UCD and TCD. Fifty-one of 1,082 OTUs (5%) differed significantly in abundance between the 2 groups (Figure 2 & Supplementary Table 3). Forty-eight OTUs (94%) had significantly higher abundance in TCD than UCD apart from OTU_31 Megamonas, OTU_143 Ruminococcus 1, and OTU_135 Holdemanella.

Likewise, 29 of 1,082 OTUs (3%) had different abundance between HC and TCD. Almost half (n=13, 45%) were increased in TCD compared with HC (Figure 2). Of the 13 OTUs more abundant in TCD than HC, 10 (77%) were significantly higher in TCD than UCD too, suggesting that treatment with GFD influences these taxa independently of disease status. Of the 16 OTUs with lower relative
abundance in TCD than HC, OTU_31 *Megamonas*, OTU_143 *Ruminococcus* 1, and OTU_135 *Holdemanella* had significantly lower abundance in TCD than UCD (Figure 2). Of note, these 3 OTUs were the only ones with lower abundance in TCD than UCD, strongly suggesting that their modulation is the consequence of treatment with GFD.

**A celiac disease-specific microbiota signature**

Irrespective of treatment with GFD, the relative abundance of 11 OTUs were consistently lower in children with CD than HC (Figure 2, Supplementary Table 3); hence composing a microbial signature specific to CD. This was visualized through NMDS analysis including only these 11 discriminant OTUs (Supplementary Figure 1). None of these 11 OTUs associated with disease duration [median, 3.1; IQR, 1.5: 7.3 yrs] in the TCD group. Using these 11 discriminatory OTUs, random forest classifier distinguished between HC and CD patients with an ‘out-of-bag’ error rate of 21.5%. This was significantly more effective than random classification (permutation ANOVA p<0.001 and AUC=0.789) (Figure 3). The 2 most influential OTUs were OTU_53 *Clostridium sensu stricto 1*, followed by OTU_143 *Ruminococcus*.

**Discriminant analysis in OTU abundance in new-onset CD following recommendation to GFD**

In the prospective cohort of UCD patients followed-up at 6 and 12 months on GFD, fecal samples were characterized by a total of 835 OTUs; 31 (3.7%) and 12 (1.4%) of which differed significantly 6 and 12 months after initiation of GFD, respectively (Figure 4). Compared to CD diagnosis, at both 6 and 12 months on GFD, the relative abundance of 7 and 3 OTUs significantly decreased and increased, respectively. It is noteworthy that in this prospective cohort the mean effect size of GFD on OTU abundance (Figure 4) was more pronounced than the magnitude of OTU abundance difference between TCD and UCD (Figure 2).

**Effect of dietary nutrients and food groups on the gut microbiota of CD**

The effect dietary modifications, during treatment with GFD, might have on microbiota was explored using stepwise data analysis. First, we correlated the intake of macronutrients (e.g. carbohydrates) and food groups (e.g. dairy portions/day) with the abundance of all OTUs characterizing the microbiota of HC (Supplementary Figure 2 & Supplementary Figure 3). Significantly related OTUs were subsequently cross-referenced with the discriminant OTUs for comparisons between UCD or HC with TCD, as well as in the subset of UCD patients with paired data at 6 and 12 months. We
applied this analysis workflow as, despite assessing the dietary intake of our CD patients with the FFQ, complete nutritional composition of gluten-free products is currently unavailable and therefore the outcome of dietary assessment, in this group, would have been incomplete and findings misleading.

Of the 200 OTUs which associated with either the macronutrient or food group intake of HC, 39 OTUs were differentially abundant between UCD and TCD children (Figure 2) suggesting that differences in the abundance of 39 of the 51 (76%) discriminatory OTUs between UCD and TCD are likely to be explained by changes in dietary nutrient intake after GFD recommendation. Likewise, 23 of the 200 OTUs were significantly differentially abundant between HC and TCD children (Figure 2). Therefore, differences in abundance of 23 of the 29 (79%) OTUs that discriminated between HC and TCD are likely to be explained by changes in dietary intake after initiation of GFD.

In the prospective cohort, from the 31 and 12 OTUs whose relative abundance changed at 6 and 12 months of treatment with GFD, 29 (94%) and 11 (82%) correlated with macronutrient or food group intake in HC (Figure 4), further supporting the hypothesis that the gut microbiota of treated CD patients differs to HC predominantly as the result of dietary modification during GFD.

Differential analysis in OTU abundance between patients with and without recent consumption of gluten

In pooled analysis (cross-sectional and prospective cohorts together), 12 (17%) children with recommended adherence to a GFD had detectable and 59 (83%) undetectable GIP. When we looked for differences in OTU abundance between these 2 groups, 89 OTUs differed (Supplementary Figure 4). Among these, all but 2 (OTU_99 Senegalimassilia and OTU_239 Clostridiales vadinBB60 group) OTUs were higher in children with undetectable fecal GIP.

For a few discriminatory OTUs between UCD and/or HC with TCD the direction of their change in abundance differed to that of those OTUs which discriminated between patients with and without recent gluten consumption. 19 out of the 87 OTUs (22%) with higher abundance in children without recent gluten ingestion were significantly increased in TCD compared with UCD, with 8 of them (9%) also significantly increased in TCD over HC (Supplementary Figure 4). Similarly, 1 of the 2 OTUs (50%) with lower abundance in children without recent gluten ingestion was also significantly lower in TCD than in HC.

Comparison in microbiota between treated patients with CD and their unaffected siblings
There was no difference in the microbiota structure (β-diversity) of the TCD patients and their unaffected siblings (Supplementary Table 1). The microbiota structure of the unaffected siblings did not differ to HC either. Fifty-six of 964 OTUs (6%) were differentially abundant between TCD and their unaffected siblings, with 36 (64%) significantly decreased in the TCD group (Supplementary Table 5).

**Diet-related microbiota metabolites**

There was no difference between the unaffected siblings of TCD children and HC in all bacterial metabolites assayed, water content and pH in feces (Table 2). No difference was found also in fecal water content and the absolute concentration of SCFA/BCFA among the UCD, TCD and HC. However, the relative abundance (%) of acetic acid was higher, and that of butyric and valeric acids were lower in TCD than HC (Figure 5a).

In the prospective group of UCD patients the absolute concentrations of butyric acid (p=0.053) and the 2 BCFA (isovaleric acid, p=0.052, isobutyric acid, p=0.063) non-significantly decreased after GFD initiation (Table 2). The effect of GFD on SCFA production was reflected also in their proportional profile (Figure 5b). Compared with disease diagnosis, the relative abundance of acetic acid increased and the relative abundances of propionic, butyric and valeric acids decreased at 6 and/or 12 months on GFD, mirroring the observations in the cross-sectional cohort and between the TCD children and HC or UCD.

Samples from TCD children had lower ammonia concentration than HC or UCD, and patients with UCD had significantly less free sulfide than HC (Table 2). Mean fecal L-lactic acid concentration was significantly lower in UCD than TCD and HC, but its D-isomer was higher in UCD than TCD. During the follow-up of the 13 UCD children, a non-significant (p=0.067) decrease in ammonia levels and a corresponding increase in free sulfide (p=0.074) and L-lactate (p=0.087) concentrations were observed.

There was no difference in bacterial metabolites, except for fecal ammonia which was significantly lower in patients who had undetectable GIP than those who had consumed gluten (Supplementary Table 6).
DISCUSSION

It is still unclear the extent to which an altered microbiota observed in previous research\(^8\text{-}17\) is involved in CD pathogenesis or if these are secondary effects of disease pathology, including increased epithelial cell turnover and nutrient malabsorption. Following diagnosis, adherence to GFD may associate with a decreased intake of non-digestible carbohydrates from cereals, thus affecting fiber fermenting species and colonic production of SCFA\(^32\). This is important as CD patients may be unable to compensate for decreased fiber intake from gluten-containing foods by increasing its intake from other sources, including fruits and vegetables. Here, by performing a data-rich study we tried to discern which microbial signals in patients with new-onset and treated CD are potentially involved in disease pathogenesis and which are secondary disease effects, including from treatment.

Even though we identified differences in the abundance of a few species between treatment-naïve UCD patients and HC, the profound microbial dysbiosis noted in Crohn’s disease was not observed, at least using crude diversity indices\(^18\). Instead, significant effects were observed in TCD patients after recommendation to a GFD, confirming our \textit{a priori} hypothesis. More importantly, we identified three major groups of bacterial taxa (Figure 2); one group which is CD-specific and non-responsive to treatment with GFD; a second group which is associated with new-onset CD but which is also treatment responsive; and a third one which is treatment dependent but does not differentiate between disease and the health state. Of these, the first cluster represents the microbial signature of CD which can distinguish children with from those without CD with a reasonably high likelihood, as demonstrated using machine learning algorithms. The magnitude of microbial alterations observed here are similar to other non-communicable diseases, in which the microbiota has been implicated in their underlying pathogenesis, such as type 1 diabetes\(^33\).

Although in the second group several other bacteria were different between UCD and the HC, these discriminatory microbial signals vanished following treatment with GFD. These represent bacteria which responded to recovery of gut pathology following treatment with GFD, or bacteria whose underlying role in CD cause and treatment might be important. The role of these taxa, the majority of which belong to Bacteroidetes warrants further research. Using denaturing gradient gel electrophoresis with group-specific primers Sánchez showed also that \textit{Bacteroides} diversity was higher in duodenal biopsies from controls than in samples from patients with active and treated CD\(^15\).

The third and largest group should almost certainly represent microbial noise attributed to dietary modification during treatment with GFD and amelioration of disease activity. This speculation is supported by the fact that several differential OTUs between the TCD with the UCD
children or HC overlapped; most of these discriminatory species were associated with participants’ diet, as well as by the observation that in patients on GFD almost 100 OTUs differed between patients with positive and negative gluten contents in their feces. As a prime example of these effects, avoidance of wheat products likely explains the decrease of Megamonas in TCD patients and the same may apply for the fiber fermenters Coprococcus, Ruminococcus and Anaerostipes, and Bifidobacterium. The reduction in Bifidobacterium in CD is in accordance with previous research using molecular fingerprinting techniques and quantitative PCR. Of note, 41% of the genera which were influenced by GFD in the current study were also influenced in healthy adults who adhered to other dietary interventions which were gluten free and low in fiber (Supplementary Figure S) in previous research. This further corroborates our conclusions that to a large extent OTUs which are altered in TCD are the result of GFD and low fiber intake. Changes in the abundance of butyric acid producers paralleled a decrease in butyric acid levels and its proportional abundance in patients on GFD. This finding confirms previous observations, but here we provide evidence that this is a secondary effect and not primary disease defect. The concentration of ammonia and BCFA was lower in patients on GFD; both in the cross-sectional and prospective cohorts. This likely indicates reduction in protein intake or lower epithelial cell shedding, with amelioration of intestinal inflammation, both resulting in less protein reaching the colon and being fermented. The reason behind the low concentration of free sulfide at disease diagnosis is unclear but in conjunction with the significantly reduced abundance of Methanobrevibacter, a methane producer, suggests an altered hydrogen metabolism in newly diagnosed CD. Hydrogen sulfide has been implicated in various processes of gut function including motility, epithelial secretion and protection from inflammation.

Findings from the prospective cohort corroborated the results of the cross-sectional group analysis although often different species were affected by GFD between the two cohorts, highlighting substantial variation in inter-individual responses. The observation that almost three times fewer species were affected after 12 compared with 6 months diagnosis indicates either better adaptation of CD patients to GFD and broader food choices to compensate for gluten-containing food with time or, most likely and supported by the change in GIP levels between these periods, loss of strict compliance to GFD in some patients.

There are significant implications for future research and clinical practice arising from the findings of this study. The role of CD-associated microbiota in disease pathogenesis, including those organisms which respond to GFD, needs to be unraveled in mechanistic research. The fact that the CD microbial signature we observed persisted in patients on GFD and was independent of disease duration suggests that the effects on these 11 OTUs are unrelated to bacterial fermentation of the
luminal glycocalyx and other gastrointestinal secretions or mediators of the innate immune system. It is therefore possible that these bacterial species are important modifiers of risk of CD onset, particularly in individuals who are genetically susceptible to developing the illness. In previous research, healthy exclusively breastfed infants who were carriers of the HLA-DQ2 haplotype and also had family history of CD, had less *Bifidobacterium*⁴⁶. Provided that the majority of patients with CD are carriers of the HLA-DQ2 haplotype the observation of a lower *Bifidobacterium* abundance reported in the current and previous study⁴⁶ suggests that genetic factors which impede the early colonization of the gut with species beneficial for human health, may potentiate the risk of development of CD; this effect extends beyond disease diagnosis and remains independent of GFD treatment.

The role of *Bifidobacterium* in the underlying microbial origins of CD pathogenesis has received extensive attention within mechanistic research. Inoculation of peripheral blood mononuclear cells with feces from active and asymptomatic CD patients increased TNF-α production and CD86 expression, while decreased IL-10 cytokine production and CD4 expression compared with samples from HC but specific *Bifidobacterium* strains suppressed this Th1 pro-inflammatory milieu, characteristic of CD ³⁷. In a subsequent study of the same group, addition of *Bifidobacterium* strains changed the gliadin-derived peptide pattern and attenuated production of TNF-α and IL-1β and expression of NF-κB and chemokine CXCR3 receptor from Caco-2 cells exposed to gliadin digestions ³⁸. These in-vitro data were replicated in gliadin-induced enteropathy murine models sensitized with interferon-γ where *Bifidobacterium longum* CECT 7347 attenuated the production of TNF-α and the CD4 mediated immune response and increased the tissue mRNA levels of NFκB and IL-10 ³⁹.

The exact mechanism by which Bifidobacteria may exhibit immune-modulating properties is not yet clear but it has been demonstrated that *Bifidobacterium longum* NCC2705 produces a serine protease inhibitor which attenuates gliadin-induced immunopathology and impacts on intestinal microbial composition in the NOD/DQ8 mouse model of gluten sensitivity⁴⁰. In one of the few clinical trials available, administration of *Bifidobacterium infantis* decreased Paneth cells and expression of α-defensin-5 in duodenal biopsy of patients with active CD⁴¹; an effect which was associated with symptom improvement but which did not modify abnormal intestinal permeability⁴².

Very few other species identified here as disease specific biomarkers have been studied in the context of CD pathogenesis. Commensal Clostridia belonging to clusters IV and XIVa are important inducers of Tregs in the colon⁴³. It is therefore possible that the highly discriminant OTU_53 belonging to an unknown *Clostridium* is less abundant in CD thus influencing the induction of Tregs required for maintaining immune homeostasis. It has also been shown that bacteria could potentially reduce gluten immunogenicity by producing enzymes that effectively cleave proteolytic-
resistant sequences in gluten peptides which activate Th1 response. Pseudomonas aeruginosa, isolated from the duodenum of CD patients, produces, through its elastase activity, a multitude of peptides that activate gluten-specific T cells in HLA-DQ2 CD patients but conversely, Lactobacillus spp from healthy subjects, degrade such modified peptides and decrease their immunogenic potential.

Future research should explore the role of the disease-specific species identified here in disease pathogenesis. Such studies may include in-vitro experiments with candidate species and immune cell co-cultures triggered by gliadin epitopes and dietary interventions aiming to change their abundance in the gut coupled with measurements of disease outcomes. The ability of microbiota signatures of unaffected siblings to predict risk of CD onset alongside other environmental factors is important to study and in a similar way to ongoing large cohort studies in Crohn’s disease.

Irrespective of their primary role in CD pathogenesis, the disease-specific microbial signature identified here might be used as another adjutant, non-invasive biomarker to screen for CD. The observation that the abundance of fiber fermenters or cross-feeders and production of butyric acid diminish in patients on GFD has implications for the dietary management of this population. Dietary fiber intake in the westernized diet is low and adherence to GFD with low consumption of cereals may decrease patient intake even further. It is therefore important for colonic health and gut motility to promote intake of non-gluten containing sources of fiber in this population and routinely fortify gluten-free products with a broad variety of fibers (e.g. pectin, ispaghula).

Limitations of the current study include the modest sample size of the prospective group. Although the mean effect size of microbial changes were more pronounced than the cross-sectional group, we may have been underpowered to identify smaller size differences. Some patients from the UCD group were lost at follow-up or their measurements were excluded. However, this group of patients did not differ in characteristics and microbiota features from patients who were retained in the analysis. Also, CD is a condition of the small bowel; hence the role of the fecal microbiota may be considered less relevant to its pathogenesis. While this is a fair argument to propose, it is possible that events in the large bowel influence disease pathogenesis upstream along the GI tract, as is perhaps the case in Crohn’s disease where colonic microbiota changes can be seen in patients with disease affecting their small intestine. It is also possible that several fecal microbes are markers of the small bowel resident community. Collado et al previously showed that similar bacteria were related to CD in both fecal and duodenal biopsies; but further research is required to clarify the role of each of these gut niches and in the mucosal adherent microbiota as suggested. In HC we did not have ethical permission to screen for CD. The fact though that none of the siblings of the CD
patients screened positive for CD infers that the proportion of HC with undiagnosed CD would have been small and unlikely to have influenced the main results presented here.

In conclusion, we identified a set of bacteria which may comprise another important environmental factor in the pathogenesis of CD and which warrant further research, but also demonstrated that several alterations in the microbiota of patients with established CD are likely to be secondary effects of disease treatment. The suppression of butyric acid production and fiber fermenters is likely a biomarker of diminished consumption of fermentable carbohydrate and may suggest a need for the development of fiber-enriched gluten-free products and interventions with prebiotics.

FIGURE & TABLE LEGENDS

Figure 1: α and β-diversity for the cross-sectional (a) and prospective (b) cohort
HC: Healthy controls, UCD: newly-diagnosed celiac disease, TCD: patients with celiac disease on gluten-free diet; NMDS: non-metric multidimensional scaling

Figure 2: Statistically significant differences (log2-fold change) in relative abundance of OTUs between groups and correlations between these discriminatory OTUs with dietary nutrients and food groups
HC: Healthy controls, UCD: newly-diagnosed celiac disease, TCD: patients with celiac disease on gluten-free diet; A negative log2-fold change represents a lower abundance in the second of the two-group comparison

Figure 3: Most influential OTUs, among the 11 disease-specific, in predicting disease or health status with associated receivers operating curves and area under the curve
**Figure 4:** Statistically significant differences (log2-fold change) in relative abundance of OTUs between follow-up timepoints and correlations between these discriminatory OTUs with dietary nutrients and food groups

UCD: newly-diagnosed celiac disease; GFD 6 and 12 mos: UCD patients on gluten-free diet for 6 and 12 months, respectively.

**Figure 5:** Relative proportion (%) of short-chain fatty acids

C2: acetic, C3: propionic, C4: butyric, C5: valeric, iC4: isobutyric, iC5: isovaleric acids; HC: Healthy controls, UCD: newly-diagnosed celiac disease, TCD: patients with celiac disease on gluten-free diet; 6 and 12 mos: UCD patients on gluten-free diet for 6 and 12 months, respectively.

**Table 1:** Participants characteristics of the cross-sectional study and prospective cohort

**Table 2:** Fecal characteristics and microbiota metabolites in the cross-sectional study and prospective cohort

**REFERENCES**

1. Husby S, Koletzko S, Korponay-Szabo IR, et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J Pediatr Gastroenterol Nutr 2012;54:136-60.
2. Leffler DA, Green PH, Fasano A. Extraintestinal manifestations of coeliac disease. Nat Rev Gastroenterol Hepatol 2015;12:561-71.
3. D’Argenio V, Casaburi G, Precone V, et al. Metagenomics Reveals Dysbiosis and a Potentially Pathogenic N. flavescens Strain in Duodenum of Adult Celiac Patients. Am J Gastroenterol 2016;111:879-90.
4. Dydensborg Sander S, Nybo Andersen AM, Murray JA, et al. Association Between Antibiotics in the First Year of Life and Celiac Disease. Gastroenterology 2019;156:2217-2229.
5. Jiang HY, Zhang X, Zhou YY, et al. Infection, antibiotic exposure, and risk of celiac disease: A systematic review and meta-analysis. J Gastroenterol Hepatol 2019.
6. Kahrs CR, Chuda K, Tapia G, et al. Enterovirus as trigger of coeliac disease: nested case-control study within prospective birth cohort. BMJ 2019;364:l231.
7. Sarno M, Discepolo V, Troncone R, et al. Risk factors for celiac disease. Ital J Pediatr 2015;41:57.
8. Collado MC, Calabuig M, Sanz Y. Differences between the fecal microbiota of coeliac infants and healthy controls. Curr Issues Intest Microbiol 2007;8:9-14.
9. Collado MC, Donat E, Ribes-Koninckx C, et al. Imbalances in faecal and duodenal Bifidobacterium species composition in active and non-active coeliac disease. BMC Microbiol 2008;8:232.
10. De Palma G, Nadal I, Medina M, et al. Intestinal dysbiosis and reduced immunoglobulin-coated bacteria associated with coeliac disease in children. BMC Microbiol 2010;10:63.
11. Di Cagno R, De Angelis M, De Pasquale I, et al. Duodenal and faecal microbiota of celiac children: molecular, phenotype and metabolome characterization. BMC Microbiol 2011;11:219.
12. Di Cagno R, Rizzello CG, Gagliardi F, et al. Different fecal microbiotas and volatile organic compounds in treated and untreated children with celiac disease. Applied and environmental microbiology 2009;75:3963-3971.
13. Olivares M, Castillejo G, Varea V, et al. Double-blind, randomised, placebo-controlled intervention trial to evaluate the effects of Bifidobacterium longum CECT 7347 in children with newly diagnosed coeliac disease. Br J Nutr 2014;112:30-40.
14. Reddel S, Putignani L, Del Chierico F. The Impact of Low-FODMAPs, Gluten-Free, and Ketogenic Diets on Gut Microbiota Modulation in Pathological Conditions. Nutrients 2019;11.
15. Sanchez E, Donat E, Ribes-Koninckx C, et al. Intestinal Bacteroides species associated with coeliac disease. J Clin Pathol 2010;63:1105-11.
16. Sánchez E, Nadal I, Donat E, et al. Reduced diversity and increased virulence-gene carriage in intestinal enterobacteria of coeliac children. BMC gastroenterology 2008;8:50-50.
17. Sanz Y, Sanchez E, Marzotto M, et al. Differences in faecal bacterial communities in coeliac and healthy children as detected by PCR and denaturing gradient gel electrophoresis, FEMS Immunol Med Microbiol 2007;51:562-8.
18. Quince C, Ijaz UZ, Loman N, et al. Extensive Modulation of the Fecal Metagenome in Children With Crohn's Disease During Exclusive Enteral Nutrition. Am J Gastroenterol 2015;110:1718-29; quiz 1730.
19. Murch S, Jenkins H, Auth M, et al. Joint BSPGHAN and Coeliac UK guidelines for the diagnosis and management of coeliac disease in children. Arch Dis Child 2013;98:806-11.
20. Tsiountsioura M, Wong JE, Upton J, et al. Detailed assessment of nutritional status and eating patterns in children with gastrointestinal diseases attending an outpatients clinic and contemporary healthy controls. Eur J Clin Nutr 2014;68:700-6.
21. Varni JW, Bendo CB, Denham J, et al. PedsQL Gastrointestinal Symptoms Module: Feasibility, Reliability, and Validity. Journal of Pediatric Gastroenterology and Nutrition 2014;59:347-355.
22. Gerasimidis K, Zafeiropoulou K, Mackinder M, et al. Comparison of Clinical Methods With the Faecal Gluten Immunogenic Peptide to Assess Gluten Intake in Coeliac Disease. J Pediatr Gastroenterol Nutr 2018;67:356-360.
23. Gerasimidis K, Bertz M, Hanske L, et al. Decline in presumptively protective gut bacterial species and metabolites are paradoxically associated with disease...
improvement in pediatric Crohn's disease during enteral nutrition. Inflamm Bowel Dis 2014;20:861-71.
24. Gerasimidis K, Bertz M, Quince C, et al. The effect of DNA extraction methodology on gut microbiota research applications. BMC research notes 2016;9:365-365.
25. Svolos V, Hansen R, Nichols B, et al. Treatment of Active Crohn's Disease With an Ordinary Food-based Diet That Replicates Exclusive Enteral Nutrition. Gastroenterology 2019;156:1354-1367.e6.
26. Joshi NAF, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). 1.33 ed, 2011.
27. Masella AP, Bartram AK, Truszkowski JM, et al. PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics 2012;13:31.
28. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013;30:772-80.
29. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One 2010;5:e9490.
30. Evans JSM, M.A, Holden, Z.A., Cushman, S.A. . Modeling species distribution and change using Random Forests. in Predictive species and habitat modeling in landscape ecology: concepts and applications: (ed. C. A. Drew, Y. F. W. and F. H.), 2011:139-159.
31. Sing T, Sander O, Beerenwinkel N, et al. ROCviewer: visualizing classifier performance in R. Bioinformatics 2005;21:3940-1.
32. Nylund L, Kaukinen K, Lindfors K. The microbiota as a component of the celiac disease and non-celiac gluten sensitivity. Clinical Nutrition Experimental 2016;6:17-24.
33. Zheng P, Li Z, Zhou Z. Gut microbiome in type 1 diabetes: A comprehensive review. Diabetes Metab Res Rev 2018;34:e3043.
34. De Paepe K, Kerckhof FM, Verspreekt J, et al. Inter-individual differences determine the outcome of wheat bran colonization by the human gut microbiome. Environ Microbiol 2017;19:3251-3267.
35. Kalantar-Zadeh K, Berean KJ, Burgell RE, et al. Intestinal gases: influence on gut disorders and the role of dietary manipulations. Nat Rev Gastroenterol Hepatol 2019;16:733-747.
36. Olivares M, Neef A, Castillejo G, et al. The HLA-DQ2 genotype selects for early intestinal microbiota composition in infants at high risk of developing coeliac disease. Gut 2015;64:406-17.
37. Medina M, De Palma G, Ribes-Koninckx C, et al. Bifidobacterium strains suppress in vitro the pro-inflammatory milieu triggered by the large intestinal microbiota of coeliac patients. Journal of Inflammation 2008;5:19.
38. Laparra JM, Sanz Y. Bifidobacteria inhibit the inflammatory response induced by gliadins in intestinal epithelial cells via modifications of toxic peptide generation during digestion. J Cell Biochem 2010;109:801-7.
39. Laparra JM, Olivares M, Gallina O, et al. Bifidobacterium longum CECT 7347 modulates immune responses in a gliadin-induced enteropathy animal model. PLoS One 2012;7:e30744.
40. McCarville JL, Dong J, Caminero A, et al. A Commensal Bifidobacterium longum Strain Prevents Gluten-Related Immunopathology in Mice through Expression of a Serine Protease Inhibitor. Appl Environ Microbiol 2017;83.
41. Pinto-Sánchez MI, Smecuol EC, Temprano MP, et al. Bifidobacterium infantis NLS Super Strain Reduces the Expression of α-Defensin-5, a Marker of Innate Immunity,
in the Mucosa of Active Celiac Disease Patients. J Clin Gastroenterol 2017;51:814-817.

42. Smecuol E, Hwang HJ, Sugai E, et al. Exploratory, randomized, double-blind, placebo-controlled study on the effects of Bifidobacterium infantis natren life start strain super strain in active celiac disease. J Clin Gastroenterol 2013;47:139-47.

43. Atarashi K, Tanoue T, Shima T, et al. Induction of colonic regulatory T cells by indigenous Clostridium species. Science 2011;331:337-41.

44. Caminero A, Herrán AR, Nistal E, et al. Diversity of the cultivable human gut microbiome involved in gluten metabolism: isolation of microorganisms with potential interest for coeliac disease. FEMS Microbiology Ecology 2014;88:309-319.

45. Caminero A, Galipeau HJ, McCarville JL, et al. Duodenal Bacteria From Patients With Celiac Disease and Healthy Subjects Distinctly Affect Gluten Breakdown and Immunogenicity. Gastroenterology 2016;151:670-683.

46. Kevans D, Turpin W, Madsen K, et al. Determinants of intestinal permeability in healthy first-degree relatives of individuals with Crohn's disease. Inflamm Bowel Dis 2015;21:879-87.

47. Collado MC, Donat E, Ribes-Koninckx C, et al. Specific duodenal and faecal bacterial groups associated with paediatric coeliac disease. J Clin Pathol 2009;62:264-9.

48. Cheng J, Kalliomäki M, Heilig HGHJ, et al. Duodenal microbiota composition and mucosal homeostasis in pediatric celiac disease. BMC Gastroenterology 2013;13:113.
Table 1: Participants characteristics of the cross-sectional study and prospective cohort

|                  | Cross-sectional study | Prospective study |
|------------------|-----------------------|-------------------|
|                  | HC (57)               | Siblings (19)     | UCD (20) | TCD (45) | Diagnosis (13) | 6 months GFD (13) | 12 months GFD (13) |
| Age (y)          | 7.8 (0.41)            | 9.1 (0.76)        | 10.1 (0.70)\(^a\) | 9.3 (0.47) | 9.5 (0.87) | 10.1 (0.87) | 10.6 (0.86)          |
| Gender (M/F)     | 27/30                 | 8/11              | 10/10    | 20/25    | 6/7         | 6/7         | 6/7                   |
| Weight (Kg)      | 29.1 (1.6)            | 33.7 (3.8)        | 33.7 (2.9) | 32.3 (1.7) | 30.5 (3.1)\(^b\) | 32.5 (3.3)\(^c\) | 34.8 (3.5)          |
| Height z-score   | 0.29 (0.15)           | 0.43 (0.28)       | -0.16 (0.22) | 0.06 (0.16) | -0.19 (0.30) | -0.15 (0.28) | -0.20 (0.27)         |
| <-2SD [n(%)]     | 1 (1.8)               | 0 (0)             | 2 (10.0) | 0 (0)    | 2 (15.4)    | 1 (7.7)     | 1 (7.7)               |
| BMI (Kg/m\(^2\)) | 16.8 (0.34)           | 17.5 (0.62)       | 17.1 (0.58) | 17.4 (0.38) | 16.5 (0.67)\(^d\) | 16.6 (0.71) | 17.1 (0.79)          |
| BMI z-score      | 0.06 (0.15)           | 0.24 (0.23)       | -0.23 (0.25) | 0.18 (0.17) | -0.40 (0.34) | -0.45 (0.33) | -0.31 (0.33)         |
| <-2SD [n(%)]     | 1 (1.8)               | 0 (0)             | 2 (10.0) | 0 (0)    | 2 (15.4)    | 2 (15.4)    | 2 (15.4)             |
| >2SD [n(%)]      | 4 (7.0)               | 2 (10.5)          | 0 (0)    | 3 (6.7)  | 0 (0)       | 0 (0)       | 0 (0)                |
| tTG (U/mL)       | -                     | -                 | 64.8 (13.3) [9] | 7.9 (3.0)\(^e\) [2] | 68.5 (19.6)\(^f\) [7] | 9.8 (3.3)\(^g\) [4] | 7.7 (2.0)\(^h\) [4] |
| <-7 [n(%)]       | -                     | -                 | 1 (9.1)  | 34 (79.1) | 1 (16.7)    | 3 (33.3)    | 6 (66.7)             |
| ≥7 [n(%)]        | -                     | -                 | 10 (90.9) | 9 (20.9)  | 5 (83.3)    | 6 (66.7)    | 3 (33.3)             |
|                          | <0.156 [n(%)] | ≥0.156 [n(%)] |
|--------------------------|---------------|---------------|
| GIP (μg/g)               | 3.5 (0.6) [1] | 2.95 (0.76) [1] |
|                          | 0.25 (0.06)*  | 0.22 (0.06)   |
|                          | 0.49 (0.23)   |               |
| <0.156 [n(%)]           | 1 (5.3)       | 18 (94.7)     |
|                          | 38 (84.4)     | 7 (15.6)      |
|                          | 0 (0)         | 12 (100)      |
|                          |               | 2 (15.4)      |
| ≥0.156 [n(%)]           |               | 3 (23.1)      |
| PedsQL-GS score         | 91.4 (1.7) [1]| 57.1 (4.8)*   |
|                          | 88.2 (2.9)    | 77.5 (2.7)    |
|                          | 58.3 (6.2)    | 67.1 (5.3)    |
|                          | 73.6 (6.4)    |               |

Values expressed as mean (SEM); GLM for UCD, TCD & HC in the cross-sectional study and GLM accounted for paired design in the prospective study; Box-Cox transformation with optimal λ in all but BMI z-score; pairwise comparison with Bonferroni correction; the number of missing data is shown in brackets; "*: p-value 0.017 compared to HC; "#: p<0.0001 compared to 6 months GFD & 12 months GFD; "*: p<0.0001 compared to 12 months GFD; "#: p<0.0001 compared to HC; "#: p<0.0001 compared to UCD and HC; "#: p<0.0001 compared to UCD; "#: p<0.0001 compared to 12 months GFD; "#: p<0.0001 compared to HC; "#: p<0.0001 compared to UCD and HC; "#: p=0.009 compared to 12 months GFD; "*: for all-group (except for siblings) comparison
Table 2: Fecal characteristics and microbiota metabolites in the cross-sectional study and prospective cohort

|                      | Cross-sectional study |                      |                      |                      |                      |                      | Prospective study |
|----------------------|-----------------------|----------------------|----------------------|----------------------|----------------------|-------------------|-------------------|
|                      | HC (57)               | Siblings (19)        | UCD (20)             | TCD (45)             | Diagnosis (13)       | 6 months GFD (13) | 12 months GFD (13) |
| Faecal pH            | 6.9 (0.08)            | 6.9 (0.15) [1]       | 6.7 (0.3)            | 7.1 (0.1) [2]        | 6.4 (0.46)           | 7.2 (0.18) [1]    | 6.8 (0.16) [1]    |
| Faecal water content (%) | 67.8 (0.7)        | 65.5 (1.6)           | 66.3 (1.7)           | 69.2 (1.2) [2]       | 67.6 (0.46)          | 65.8 (1.1)        | 69.1 (2.0)        |
| Ammonia (*10⁻⁴ mg/g) | 11.5 (0.8)            | 11.4 (1.4)           | 19.6 (8.2)           | 7.8 (0.8) [2]        | 11.2 (0.96)          | 8.0 (1.2)         | 11.1 (1.6) [2]    |
| Free sulphide (μmol/g) | 0.13 (0.02) [2]   | 0.13 (0.03)          | 0.06 (0.01) [2]      | 0.10 (0.01) [4]      | 0.03 (0.01) [1]      | 0.05 (0.01)       | 0.09 (0.02)       |
| Total sulphide (μmol/g) | 0.83 (0.10)         | 1.15 (0.17)          | 0.83 (0.13) [2]      | 1.03 (0.11) [3]      | 0.87 (0.17) [1]      | 0.72 (0.13)       | 0.57 (0.11)       |
| L-lactic acid (μg/g) | 126.8 (41.3)          | 114.0 (11.1) [2]     | 60.1 (14.0) [c,d,1]  | 100.8 (11.6) [7]     | 55.3 (16.6)          | 67.6 (10.9)       | 81.5 (11.2)       |
| D-lactic acid (μg/g) | 92.5 (18.4)           | 60.4 (6.4) [2]       | 119.1 (23.6) [1]     | 62.5 (4.1) [7]       | 130.2 (33.9)         | 116.9 (16.9)      | 123.2 (15.6)      |
| Total lactic acid (μg/g) | 219.3 (59.4)        | 174.4 (11.6) [2]     | 179.2 (29.0) [1]     | 163.3 (13.6) [7]     | 185.5 (38.6)         | 184.4 (17.5)      | 204.7 (18.0)      |
| Acetic acid (μmol/g) | 128.2 (5.2)           | 117.9 (8.6)          | 119.9 (10.0) [1]     | 124.6 (6.9) [2]      | 121.9 (13.8)         | 124.0 (12.6)      | 116.6 (11.1)      |
| Propionic acid (μmol/g) | 25.8 (1.6)           | 23.1 (3.0)           | 23.2 (2.7) [1]       | 23.1 (2.1) [2]       | 26.7 (3.5)           | 22.0 (2.9)        | 23.6 (3.0)        |
| Butyric acid (μmol/g) | 24.4 (1.5)           | 21.9 (2.8)           | 23.2 (3.3) [1]       | 21.0 (2.5) [2]       | 26.7 (4.4)           | 19.8 (5.4)        | 20.5 (2.5)        |
| Valeric acid (μmol/g) | 3.1 (0.18)            | 2.9 (0.31)           | 3.0 (0.37) [1]       | 2.5 (0.24) [2]       | 3.2 (0.51)           | 2.3 (0.37)        | 2.4 (0.33)        |
| Metabolite                  | UCD Mean (SEM) | TCD Mean (Error) | HC Mean (SEM) | \[1\] Mean (SEM) | \[2\] Mean (SEM) | HC Mean (SEM) |
|-----------------------------|----------------|------------------|---------------|------------------|------------------|---------------|
| Isobutyric acid (μmol/g)    | 3.6 (0.22)     | 3.5 (0.35)       | 3.8 (0.46)    | 3.0 (0.25)       | 4.0 (0.64)       | 3.1 (0.37)    |
| Isovaleric acid (μmol/g)    | 3.6 (0.23)     | 3.5 (0.32)       | 3.9 (0.48)    | 3.1 (0.28)       | 4.1 (0.66)       | 3.1 (0.37)    |
| Total SCFA (μmol/g)         | 188.6 (7.9)    | 172.8 (14.6)     | 176.9 (15.9)  | 177.3 (11.4)     | 186.6 (21.7)     | 174.4 (21.3)  |

Values expressed as mean (SEM); GLM for UCD, TCD & HC in the cross-sectional study and GLM accounted for paired data in the prospective study; Box-Cox transformation with optimal λ; pairwise comparison, Bonferroni method; the number of missing data is shown in brackets; metabolites are measured per wet matter; \(^a\): p=0.001 compared to HC & UCD; \(^b\): p=0.009 compared to HC; \(^c\): p=0.012 compared to HC; \(^d\): p=0.003 compared to TCD; \(^e\): p=0.037 compared to UCD; \(^f\): p=0.009 compared to HC; \(^g\): p=0.003 compared to diagnosis; \(^h\): p=0.017 compared to HC; \(^i\): p=0.013 compared to diagnosis; \(^j\): p=0.045 compared to HC; \(^k\): p=0.008 compared to diagnosis; \(^l\): p=0.014 compared to diagnosis; \(*\): for all groups comparison
What you need to know:

Background and Context: It is not clear whether alterations in the intestinal microbiota of children with celiac disease cause the disease or are a result of disease and/or its treatment with gluten-free diet (GFD).

New Findings: Although several alterations in the intestinal microbiota of children with established celiac disease appear to be effects of a GFD, there are specific bacteria that are distinct biomarkers of celiac disease.

Limitations: It is not clear whether the microbes identified contribute to pathogenesis of celiac disease or are the result of it.

Impact: The GFD alters the intestinal microbiota, but in patients with celiac disease, there are additional differences, compared with healthy children.

Lay Summary: Children with celiac disease have differences in composition of intestinal microbes compared to healthy children. Some of these differences are caused by a gluten-free diet, but studies are needed to determine whether the other changes are a cause or a result of celiac disease.
**Supplementary Figure 1:** Non-metric multidimensional scaling of the microbial community formed by only the 11 discriminant OTUs between children with celiac disease and healthy controls.

**Supplementary Figure 2:** All significant correlations between the abundances of OTUs from the microbiota of the healthy controls with their dietary nutrient intake.

%TDEI: %Total dietary energy intake; MUFA: Mono-unsaturated fatty acids; NMES: Non-milk extrinsic sugars; NSP: Non-starch polysaccharides; DRV: Dietary Reference Value; RNI: Reference Nutrient Intake; PUFA: Polyunsaturated fatty acids; SFA: Saturated fatty acids.

**Supplementary Figure 3:** All significant correlations between the abundances of OTUs from the microbiota of the healthy controls with the intake of food groups.

**Supplementary Figure 4:** Differential OTUs between patients, on gluten-free diet recommendation with positive and negative fecal gluten immunogenic peptide. Bar color indicates significant OTUs with significant difference between 2 groups.

**Supplementary Figure 5:** Changes in genera during gluten free diet (GFD) in children with celiac disease of the current study and their overlap with genera which changed following treatment with exclusive enteral nutrition (EEN) or a new dietary therapy for Crohn’s disease (CD-TREAT) in healthy adults in previous research.

Red color indicates increase and blue color decrease in relative abundance.
Supplementary Table 1: PERMANOVA analysis using Bray-Curtis and unweighted UniFrac metrics at OTU and genus level

Supplementary Table 2: Bioenv selected OTU that explain part [(a) 92.3% in UCD, TCD and HC, (b) 92% in UCD and TCD, (c) 91.4% in HC and TCD children] of the variance in microbiota structure described by the full OTU dataset

Supplementary Table 3: OTUs with significantly different relative abundance in fecal samples of UCD, TCD and HC

Supplementary Table 4: OTUs with significantly different relative abundance in fecal samples of (a) paired data from 13 CD children at diagnosis and at six and 12 months after the initiation of GFD, and (b) from cross-sectional data from UCD patients (n=20) and TCD patients (n=45)

Supplementary Table 5: OTUs with significantly different relative abundance in fecal samples of 18 TCD children and their unaffected 19 siblings

Supplementary Table 6: Fecal metabolites in celiac disease children with or without recent ingestion of gluten
Bray-Curtis NMDS - Disease Specific OTUs

Condition
- CD
- HC
| CD-TREAT | EEN | GFD |
|----------|-----|-----|
| [Eubacterium] hallii group |  |  |
| Clostridium sensu stricto 1 |  |  |
| Cronobacter |  |  |
| Dialister |  |  |
| Dorea |  |  |
| Eggerthella |  |  |
| Erysipelotrichis |  |  |
| Flavonifractor |  |  |
| Intestinimonas |  |  |
| [Eubacterium] ventriosum group |  |  |
| Howardella |  |  |
| Blautia |  |  |
| Decrease |  |  |

| Phascolarctobacterium |  |  |
| Ruminoclostridium 9 |  |  |
| Ruminococcaceae NK4A214 group |  |  |
| Ruminococcaceae UCG-002 |  |  |
| Ruminococcaceae UCG-005 |  |  |
| Ruminococcus 2 |  |  |
| Subdoligranulum |  |  |
| Terrisporobacter |  |  |
| Tyzzerella 4 |  |  |
| Veillonella |  |  |
| Ruminococcus 1 |  |  |
| Eisenbergiella |  |  |
| Family_XIII_AD3011_group |  |  |
| Anaerostipes |  |  |
| Anaerotruncus |  |  |
| Bifidobacterium |  |  |
| Lachnospiraceae FCS020 group |  |  |
| Escherichia/Shigella |  |  |
| Ruminoclostridium |  |  |
| Pseudobutyrivibrio |  |  |
| Lachnoclostridium |  |  |
| Hungatella |  |  |
| Bilophila |  |  |
| Actinomyces |  |  |
| Anaerofilum |  |  |
| Oscillibacter |  |  |
| Faecalibacterium |  |  |
| Ruminococcaceae UCG-013 |  |  |
| Lachnospiraceae UCG-004 |  |  |
| Candidatus Soleaeferrea |  |  |
| Ruminoclostridium 5 |  |  |
| Lactococcus |  |  |
| Butyricimonas |  |  |
| Tyzzerella 4 |  |  |
| Senegalimassilia |  |  |
| Ruminoclostridium 6 |  |  |
| Ruminococcaceae UCG-002 |  |  |
| Ruminococcaceae UCG-005 |  |  |
| Gordonibacter |  |  |
| Alistipes |  |  |
| Ruminococcaceae UCG-010 |  |  |
| Prevotella |  |  |
| Lactobacillus |  |  |
| Desulfovibrio |  |  |
| Ruminococcaceae UCG-004 |  |  |
| Lachnospiraceae UCG-010 |  |  |
| Coprococcus 2 |  |  |
| Fusobacterium |  |  |
| Prevotella 9 |  |  |
| Ruminococcus 2 |  |  |
| Faecalitalea |  |  |
| Holdemanella |  |  |
| Catenibacterium |  |  |
| Granulicatella |  |  |
| Lachnospiraceae ND3007 group |  |  |
| Tyzzerella 3 |  |  |
| Incertae_Sedis |  |  |
| Peptococcus |  |  |
| Lachnospira |  |  |
| Lachnospiraceae UCG-005 |  |  |
| Cloacibacillus |  |  |
| Lachnospiraceae UCG-001 |  |  |
| Megamonas |  |  |
| Haemophilus |  |  |
| Ruminococcaceae UCG-003 |  |  |
| [Eubacterium] ventriosum group |  |  |
| Howardella |  |  |
| Blautia |  |  |
| Increase |  |  |
Supplementary Table 1: PERMANOVA analysis using Bray-Curtis and unweighted UniFrac distance metrics at OTU and genus level

|                           | Bray-Curtis | Unweighted UniFrac |
|---------------------------|-------------|--------------------|
| **OTU level**             |             |                    |
| Overall                   | p=0.025, R²=2.76% | p=0.027, R²=2.49%  |
| **Within group comparison**|             |                    |
| UCD (20)                  | HC (57)     | p=0.506, R²=1.21%  | p=0.125, R²=1.77% |
| TCD (45)                  | HC (57)     | p=0.017, R²=2.32%  | p=0.045, R²=1.57% |
|                           | UCD (20)    | p=0.106, R²=2.35%  | p=0.056, R²=2.46% |
| UCD (13)                  | 6mos GFD (13)| p=0.951, R²=1.53%  | p=0.377, R²=2.54% |
|                           | 12mos GFD (13)| p=0.762, R²=2.42%  | p=0.691, R²=2.44% |
| Siblings (19)             | TCD (18)    | p=0.384, R²=2.89%  | p=0.336, R²=2.96% |
|                           | HC (57)     | p=0.745, R²=0.97%  | p=0.183, R²=1.67% |
| **Genus level**           |             |                    |
| Overall                   | p=0.026, R²=3.0%  | n/a                |
| **Within group comparison**|             |                    |
| UCD (20)                  | HC (57)     | p=0.515, R²=1.15%  | n/a                |
| TCD (45)                  | HC (57)     | p=0.013, R²=2.56%  | n/a                |
|                           | UCD (20)    | p=0.069, R²=2.77%  | n/a                |
| UCD (13)                  | 6mos GFD (13)| p=0.998, R²=1.28%  | n/a                |
|                           | 12mos GFD (13)| p=0.904, R²=2.21%  | n/a                |
| Siblings (19)             | TCD (18)    | p=0.296, R²=3.21%  | n/a                |
|                           | HC (57)     | p=0.696, R²=0.94%  | n/a                |

OTU: operational taxonomic unit; n/a: not applicable
**Supplementary Table 2**: Bioenv selected OTU that explain part [(a) 92.3% in UCD, TCD and HC, (b) 92% in UCD and TCD, (c) 91.4% in HC and TCD children] of the variance in microbiota structure described by the full OTU dataset:

### a)

```
| OTU       | Phylum     | Subdivision           | Class     | Order          | Family      | Genus          |
|-----------|------------|-----------------------|-----------|----------------|-------------|----------------|
| OTU_6     | Bacteria   | Firmicutes            | Selenomonadales | Veillonellaceae | Dialister;  |                |
| OTU_4     | Bacteroidetes | Bacteroidia         | Bacteroidales | Bacteroidaceae | Bacteroides; |                |
| OTU_2     | Actinobacteria | Actinobacteria    | Bifidobacteriales | Bifidobacteriaceae | Bifidobacterium; |                |
| OTU_881  | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Ruminococcaceae | Faecalibacterium; |
| OTU_14    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Ruminococcaceae | Subdoligranulum; |
| OTU_33    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Roseburia;     |                |
| OTU_129   | Bacteria   | Actinobacteria        | Coriobacteriales | Coriobacteriaceae | Eggerthella; |                |
| OTU_568   | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Lachnospiraceae | Lachnospiraceae UCG-008; |
| OTU_72    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Lachnospiraceae | Lachnocostridium; |
| OTU_35    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Blautia;       |                |
| OTU_20    | Bacteria   | Bacteroidetes         | Bacteroidia | Bacteroidales  | Bacteroidaceae | Bacteroides;    |
| OTU_458   | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Lachnospiraceae | Blautia;       |
| OTU_10    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Faecalibacterium; |                |
```

### b)

```
| OTU*      | Phylum     | Subdivision           | Class     | Order          | Family      | Genus          |
|-----------|------------|-----------------------|-----------|----------------|-------------|----------------|
| OTU_6     | Bacteria   | Firmicutes            | Selenomonadales | Veillonellaceae | Dialister;  |                |
| OTU_4     | Bacteroidetes | Bacteroidia         | Bacteroidales | Bacteroidaceae | Bacteroides; |                |
| OTU_2     | Actinobacteria | Actinobacteria    | Bifidobacteriales | Bifidobacteriaceae | Bifidobacterium; |                |
| OTU_881  | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Ruminococcaceae | Faecalibacterium; |
| OTU_14    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Ruminococcaceae | Subdoligranulum; |
| OTU_33    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Roseburia;     |                |
| OTU_129   | Bacteria   | Actinobacteria        | Coriobacteriales | Coriobacteriaceae | Eggerthella; |                |
| OTU_568   | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Lachnospiraceae | Lachnospiraceae UCG-008; |
| OTU_72    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Lachnospiraceae | Lachnocostridium; |
| OTU_35    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Blautia;       |                |
| OTU_20    | Bacteria   | Bacteroidetes         | Bacteroidia | Bacteroidales  | Bacteroidaceae | Bacteroides;    |
| OTU_10    | Bacteria   | Firmicutes            | Clostridia  | Clostridiales  | Faecalibacterium; |                |
```
**c)**

| OTU    | Phylum          | Class          | Order           | Family           | Genus       | Species                   |
|--------|-----------------|----------------|-----------------|------------------|-------------|---------------------------|
| OTU_57 | Bacteria        | Firmicutes     | Clostridiae     | Lachnospiraceae  | Lachnospira |                           |
| OTU_6  | Bacteria        | Firmicutes     | Negativicutes   | Selenomonadales  | Veillonellaceae | Dialister                |
| OTU_4  | Bacteria        | Firmicutes     | Bacteroidia     | Bacteroidales    | Bacteroides  |                           |
| OTU_60 | Bacteria        | Firmicutes     | Bacteroidia     | Bacteroidales    | Bacteroides  |                           |
| OTU_14 | Bacteria        | Firmicutes     | Clostridiae     | Clostridiales    | Subdoligranulum |                          |
| OTU_33 | Bacteria        | Firmicutes     | Clostridiae     | Lachnospiraceae  | Roseburia    |                           |
| OTU_129| Bacteria        | Actinobacteria | Actinobacteria  | Coriobacteriales | Eggerthella  |                           |
| OTU_568| Bacteria        | Actinobacteria | Actinobacteria  | Bifidobacteriales | Bifidobacterium | Bifidobacterium bifidum NCIMB 41171 |
| OTU_72 | Bacteria        | Firmicutes     | Clostridiae     | Lachnospiraceae  | Lachnoclostridium |                         |
| OTU_655| Bacteria        | Bacteroidetes  | Bacteroidales   | Bacteroidaceae   | Bacteroides  |                           |
| OTU_86 | Bacteria        | Bacteroidetes  | Bacteroidales   | Rikenellaceae    | Alistipes    |                           |
| OTU_41 | Bacteria        | Firmicutes     | Clostridiae     | Clostridiales    | Peptostreptococaceae | Intestinibacter   |

*Asterisk indicates Bioenv-selected OTUs that partially explain the variance in community structure of UCD vs TCD as well as HC vs TCD; OTU: operational taxonomic unit*
**Supplementary Table 3:** OTUs with significantly different relative abundance in faecal samples of UCD, TCD and HC

| Group comparison | BaseMean | log2Fold Change | p-value | p-value (adjusted) |
|------------------|----------|-----------------|---------|--------------------|
| **OTU_60 Bacteroides** | | | | |
| HC Vs UCD | 7.06 | -0.23 | 6.3E-01 | 7.83E-01 |
| HC Vs TCD | 122.96 | 5.17 | 3.03E-23 | 2.17E-20 |
| UCD Vs TCD | 80.26 | 4.15 | 1.43E-08 | 8.46E-07 |
| **OTU_143 Ruminococcus 1** | | | | |
| HC Vs UCD | 25.1 | -2.08 | 8.85E-04 | 1.16E-02 |
| HC Vs TCD | 18.11 | -4.37 | 8.31E-23 | 2.98E-20 |
| UCD Vs TCD | 3.30 | -2.28 | 5.05E-08 | 2.44E-06 |
| **OTU_31 Megamonas** | | | | |
| HC Vs UCD | 39.6 | -2.09 | 4.56E-03 | 3.55E-02 |
| HC Vs TCD | 28.46 | -4.72 | 3.43E-21 | 8.21E-19 |
| UCD Vs TCD | 4.94 | -2.64 | 3.99E-08 | 2.12E-06 |
| **OTU_244 Ruminococcus 1** | | | | |
| HC Vs UCD | 24.72 | -1.37 | 3.2E-02 | 1.06E-01 |
| HC Vs TCD | 17.06 | -4.26 | 5.95E-21 | 1.07E-18 |
| UCD Vs TCD | 2.15 | -1.19 | 1.28E-03 | 9.45E-03 |
| **OTU_78 Ruminiclostridium 5** | | | | |
| HC Vs UCD | 2.71 | -0.69 | 9.89E-02 | 2.34E-01 |
| HC Vs TCD | 30.16 | 4.40 | 1.54E-20 | 2.21E-18 |
| UCD Vs TCD | 44.97 | 5.09 | 1.97E-12 | 5.22E-10 |
| **OTU_42 Methanobrevibacter** | | | | |
| HC Vs UCD | 38.78 | -3.84 | 1.33E-07 | 1.47E-05 |
| HC Vs TCD | 29.62 | -4.36 | 1.88E-17 | 2.24E-15 |
| UCD Vs TCD | 2.84 | -0.51 | 2.78E-01 | 4.25E-01 |
| **OTU_114 Caprococcus 2** | | | | |
| HC Vs UCD | 32.92 | -0.25 | 7.13E-01 | 8.41E-01 |
| HC Vs TCD | 38.73 | -4.27 | 2.92E-17 | 2.99E-15 |
| UCD Vs TCD | 4.24 | -0.79 | 1.22E-01 | 2.32E-01 |
| **OTU_70 Ruminicoccaceae** | | | | |
| HC Vs UCD | 11.29 | -3.43 | 1.28E-07 | 1.47E-05 |
| **UCG-014** | | | | |
| HC Vs TCD | 8.89 | -3.36 | 5.94E-14 | 5.32E-12 |
| UCD Vs TCD | 1.42 | 0.07 | 8.52E-01 | n/a |
| **OTU_120 Lachnospiraceae** | | | | |
| HC Vs UCD | 4.48 | 0.63 | 1.34E-01 | 2.97E-01 |
| **UCG-005** | | | | |
| HC Vs TCD | 69.03 | 3.59 | 7.21E-13 | 5.74E-11 |
| OTU | Taxon | UCD Vs TCD | HC Vs UCD | HC Vs TCD | UCD Vs TCD |
|-----|-------|------------|-----------|-----------|-----------|
| OTU_135 | Holdemanella | UCD Vs TCD | 99.05 | 4.43 | 3.74E-11 | 3.43E-09 |
| OTU_197 | Ruminococcaceae | HC Vs UCD | 21.19 | -1.04 | 1.10E-01 | 2.55E-01 |
| OTU_908 | Anaerostipes | HC Vs UCD | 12.70 | -0.83 | 1.60E-01 | 3.32E-01 |
| OTU_979 | Bifidobacterium | HC Vs UCD | 45.70 | -0.18 | 7.36E-01 | 8.59E-01 |
| OTU_53 | Clostridium sensu stricto 1 | HC Vs UCD | 200.98 | -2.58 | 1.1E-05 | 4.04E-04 |
| OTU_186 | Gastranaerophilales | HC Vs UCD | 1.35 | -0.11 | 7.66E-01 | n/a |
| OTU_84 | Enterorhabdus | HC Vs UCD | 7.93 | -2.34 | 1.20E-04 | 2.54E-03 |
| OTU_174 | Catenibacterium | HC Vs UCD | 1.56 | 0.70 | 2.74E-02 | n/a |
| OTU_336 | Holdemanella | HC Vs UCD | 7.99 | -2.38 | 6.68E-05 | 1.76E-03 |
| OTU_776 | Bifidobacterium | HC Vs UCD | 21.81 | -2.25 | 5.29E-04 | 8.59E-03 |
| UCG – 005 | Pseudocatenulatum | HC Vs TCD | 22.96 | 2.76 | 1.12E-11 | 7.29E-10 |
| OTU_98 | UCG – 005 | HC Vs TCD | 23.17 | 0.85 | 1.42E-01 | 2.56E-01 |
| OTU_186 | Catenibacterium | HC Vs UCD | 1.35 | -0.11 | 7.66E-01 | n/a |
| OTU_186 | Catenibacterium | HC Vs UCD | 7.93 | -2.34 | 1.20E-04 | 2.54E-03 |
| OTU_186 | Catenibacterium | HC Vs UCD | 1.35 | -0.11 | 7.66E-01 | n/a |
| OTU   | Species          | HC Vs UCD | HC Vs TCD | UCD Vs TCD |
|-------|------------------|-----------|-----------|------------|
| OTU_88 | Bacteroides     | 281.31    | 0.02      | 9.79E-01   | 9.82E-01   |
|       |                  | 863.13    | 2.49      | 1.02E-07   | 3.48E-06   |
|       |                  | 1,187.7   | 2.48      | 2.66E-04   | 2.77E-03   |
| OTU_1002 | Lachnospira   | 8.97      | -1.75     | 8.59E-04   | 1.16E-02   |
|        |                  | 33.92     | 2.49      | 1.08E-07   | 3.51E-06   |
|        |                  | 45.22     | 3.23      | 1.54E-06   | 3.89E-05   |
| OTU_28 | Bacteroides     | 151.07    | -1.06     | 6.32E-02   | 1.7E-01    |
|        |                  | 728.21    | 2.57      | 2.80E-07   | 8.72E-06   |
|        |                  | 552.60    | 3.13      | 1.08E-06   | 2.85E-05   |
| OTU_1049 | Lachnoclostridium | 53.35    | 0.60      | 3.14E-01   | 5.04E-01   |
|        |                  | 128.17    | 2.28      | 8.26E-07   | 2.28E-05   |
|        |                  | 180.41    | 1.68      | 1.09E-02   | 5.23E-02   |
| OTU_125 | Parabacteroides | 34.72     | -1.85     | 9.65E-04   | 1.22E-02   |
|        |                  | 116.40    | 2.28      | 9.30E-07   | 2.47E-05   |
|        |                  | 148.10    | 4.02      | 3.88E-11   | 3.43E-09   |
| OTU_278 | Akkermansia     | 4.46      | 0.34      | 4.50E-01   | 6.47E-01   |
|        |                  | 15.59     | 2.17      | 1.13E-06   | 2.88E-05   |
|        |                  | 20.63     | 2.37      | 1.87E-04   | 2.2E-03    |
| OTU_537 | Dialister      | 18.0      | -2.46     | 1.10E-04   | 2.43E-03   |
|        |                  | 15.10     | -2.20     | 1.37E-06   | 3.40E-05   |
|        |                  | 4.68      | 0.23      | 6.11E-01   | 7.29E-01   |
| OTU_62 | Ruminococcaceae | 23.89     | -3.94     | 1.84E-08   | 1.01E-05   |
| UCG-014 |                | 11.97     | -2.27     | 1.69E-06   | 3.90E-05   |
|        |                  | 3.30      | 0.88      | 7.0E-02    | 1.79E-01   |
| OTU_752 | Subdoligranulum| 43.81     | -1.87     | 1.49E-03   | 1.68E-02   |
|        |                  | 164.0     | 2.47      | 2.72E-06   | 5.57E-05   |
|        |                  | 122.66    | 3.48      | 5.24E-07   | 1.74E-05   |
| OTU_99 | Senegalimassilia| 73.24     | -3.88     | 1.05E-07   | 1.47E-05   |
|        |                  | 61.70     | -2.47     | 9.10E-06   | 1.52E-04   |
|        |                  | 14.06     | 1.48      | 2.37E-02   | 8.5E-02    |
| OTU_112 | Bacteroidetes   | 22.60     | -3.07     | 1.16E-07   | 1.47E-05   |
| OTU | Species          | HC Vs UCD  | - | p-value  | FDR  |
|-----|------------------|------------|---|----------|------|
| OTU_1054 | Alistipes     | 31.79 | -1.19 | 1.0E-02 | 4.2E-02 |
|      |                 | 13.92 | 2.42  | 1.43E-05 | 3.15E-04 |
| OTU_146 | Slackia        | 17.29 | -2.95 | 7.37E-07 | 6.78E-05 |
|      |                 | 30.12 | 0.81  | 8.83E-02 | 2.18E-01 |
|      |                 | 28.41 | 3.74  | 1.82E-09 | 1.38E-07 |
| OTU_226 | [Eubacterium] oxideroducens group | 7.95 | -3.09 | 1.02E-06 | 8.05E-05 |
|      |                 | 7.16 | -1.73 | 1.27E-04 | 1.26E-03 |
|      |                 | 2.57  | 1.36  | 9.44E-03 | 4.91E-02 |
| OTU_45 | Parabacteroides | 56.41 | -3.08 | 3.60E-06 | 2.21E-04 |
|      |                 | 73.21 | 0.84  | 1.04E-01 | 2.47E-01 |
|      |                 | 69.9   | 3.44  | 3.06E-07 | 1.25E-05 |
| OTU_895 | Barnesiella    | 58.0   | -3.11 | 5.04E-06 | 2.77E-04 |
|      |                 | 54.13 | -1.48 | 5.16E-03 | 2.4E-02 |
|      |                 | 21.56 | 1.65  | 1.02E-02 | 5.03E-02 |
| OTU_101 | Anaerotruncus  | 15.61 | -3.11 | 5.53E-06 | 2.77E-04 |
|      |                 | 17.95 | -0.46 | 3.83E-01 | 5.67E-01 |
|      |                 | 11.03 | 2.64  | 4.55E-05 | 8.05E-04 |
| OTU_572 | Family XVIII    | 9.54   | -2.43 | 7.14E-06 | 3.16E-04 |
| AD3011 group |                | 11.51 | -0.19 | 6.55E-01 | 7.82E-01 |
|      |                 | 7.94  | 2.20  | 6.21E-05 | 9.70E-04 |
| OTU_230 | Bacteroides     | 6.20   | -2.79 | 7.43E-06 | 3.16E-04 |
|      |                 | 6.79  | -0.60 | 1.79E-01 | 3.62E-01 |
|      |                 | 4.02  | 2.18  | 1.13E-04 | 1.58E-03 |
| OTU_65 | Bacteroides      | 9.79   | -2.44 | 8.18E-06 | 3.23E-04 |
|      |                 | 25.82 | 1.73  | 7.37E-05 | 1.28E-03 |
|      |                 | 29.9  | 4.19  | 1.17E-12 | 5.22E-10 |
| OTU_49 | Alistipes        | 56.57 | -2.40 | 1.86E-05 | 6.42E-04 |
|      |                 | 174.05 | 1.50  | 3.31E-03 | 1.72E-02 |
| OTU   | Species          | HC Vs UCD | HC Vs TCD | UCD Vs TCD | HC Vs UCD | HC Vs TCD | UCD Vs TCD | HC Vs UCD | HC Vs TCD | UCD Vs TCD | HC Vs UCD | HC Vs TCD | UCD Vs TCD |
|-------|------------------|-----------|-----------|------------|-----------|-----------|------------|-----------|-----------|------------|-----------|-----------|------------|
| OTU_109 | Ruminococcus 2   | 6.0       | 5.40      | 2.22       | -2.54     | -1.57     | 0.96       | 2.39E-05  | 2.12E-04  | 3.83E-02  | 7.75E-04  | 1.88E-03  | 1.17E-01  |
| OTU_51  | Ruminococcaceae  | 13.48     | 16.15     | 11.22      | -2.49     | -0.22     | 2.25       | 4.88E-05  | 6.38E-01  | 1.20E-04  | 1.50E-03  | 7.66E-01  | 1.60E-03  |
| OTU_111 | Blautia          | 23.54     | 28.78     | 20.82      | -2.33     | -0.13     | 2.19       | 6.14E-05  | 7.87E-01  | 4.69E-04  | 1.7E-03   | 8.72E-01  | 4.61E-03  |
| OTU_868 | Bacteroides      | 136.18    | 293.71    | 310.64     | -2.64     | 1.33      | 4.83       | 7.33E-05  | 1.42E-02  | 8.47E-12  | 1.76E-03  | 5.4E-02   | 1.12E-09  |
| OTU_923 | Mavinbryantia    | 10.23     | 13.06     | 10.29      | -2.00     | 0.07      | 2.06       | 7.03E-05  | 8.63E-01  | 2.49E-04  | 1.76E-03  | 9.2E-01   | 2.7E-03   |
| OTU_190 | Alistipes        | 6.26      | 11.02     | 13.20      | -2.16     | 0.07      | 3.35       | 1.09E-04  | 8.74E-01  | 5.57E-08  | 2.43E-03  | 9.26E-01  | 2.46E-06  |
| OTU_168 | Alistipes        | 4.43      | 3.92      | 1.75       | -2.01     | -1.54     | 0.45       | 4.11E-04  | 8.38E-05  | 2.5E-02   | 7.31E-03  | 9.4E-04   | 3.9E-01   |
| OTU_3   | Akkermansia      | 2,041.4   | 2,863.9   | 2,466.0    | -2.17     | 0.33      | 1.45       | 1.41E-03  | 5.55E-01  | 5.12E-02  | 1.63E-02  | 7.09E-01  | 1.43E-01  |
| OTU_1   | Prevotella 9     | 6,760.4   | 5,953.8   | 193.2      | -2.68     | -1.79     | 0.85       | 3.77E-03  | 1.22E-02  | 2.36E-01  | 3.15E-02  | 4.84E-03  | 3.75E-01  |
| OTU_302 | Eisenbergiella   | 26.32     | 55.97     | 97.89      | -1.38     | 1.44      | 3.49       | 1.01E-02  | 7.21E-04  | 2.47E-09  | 4.74E-02  | 5.23E-03  | 1.64E-07  |
| OTU     | Species                  | Comparison | p-value | q-value |
|---------|--------------------------|------------|---------|---------|
| OTU_259 | *Erysipelatoclostridium* | HC Vs UCD  | 4.86    | 1.07E-02 |
|         |                          | HC Vs TCD  | 8.45    | 5.0E-02  |
|         |                          | UCD Vs TCD | 8.6     | 1.64E-05 |
| OTU_448 | *Ruminiclostridium*      | HC Vs UCD  | 9.97    | 1.56E-03 |
|         |                          | HC Vs TCD  | 15.06   | 1.57E-01 |
|         |                          | UCD Vs TCD | 13.72   | 1.7E-05  |
| OTU_270 | *Butyricimonas*          | HC Vs UCD  | 5.25    | 7.63E-03 |
|         |                          | HC Vs TCD  | 10.0    | 3.52E-02 |
|         |                          | UCD Vs TCD | 10.47   | 2.34E-05 |
| OTU_248 | *Clostridium sensu stricto* 1 | HC Vs UCD  | 2.07    | 1.87E-01 |
|         |                          | HC Vs TCD  | 4.37    | 3.5E-04  |
|         |                          | UCD Vs TCD | 7.89    | 2.69E-05 |
| OTU_317 | *Lachnospiraceae*        | HC Vs UCD  | 3.81    | 2.89E-03 |
|         | *NK4A136 group*          | HC Vs TCD  | 7.29    | 1.99E-02 |
|         |                          | UCD Vs TCD | 7.67    | 2.79E-05 |
| OTU_34  | *Bacteroides*            | HC Vs UCD  | 243.37  | 3.29E-02 |
|         |                          | HC Vs TCD  | 425.59  | 1.19E-01 |
|         |                          | UCD Vs TCD | 431.67  | 1.44E-04 |
| OTU_235 | *Lachnospiraceae*        | HC Vs UCD  | 6.69    | 1.60E-02 |
|         |                          | HC Vs TCD  | 10.74   | 1.21E-01 |
|         |                          | UCD Vs TCD | 10.62   | 4.15E-04 |
| OTU_1036| *Bacteroidetes*          | HC Vs UCD  | 59.31   | 1.14E-01 |
|         |                          | HC Vs TCD  | 76.45   | 7.8E-01  |
|         |                          | UCD Vs TCD | 61.94   | 4.22E-04 |
| OTU_73  | *Bacteroides*            | HC Vs UCD  | 6.31    | 4.38E-02 |
|         |                          | HC Vs TCD  | 11.31   | 3.24E-02 |
|         |                          | UCD Vs TCD | 12.07   | 5.89E-04 |
| OTU_154 | *Odoribacter*            | HC Vs UCD  | 30.05   | 2.48E-01 |
|         |                          | HC Vs TCD  | 61.85   | 7.18E-03 |
|         |                          | UCD Vs TCD | 72.08   | 8.68E-04 |
| OTU_139 | *Bilophila*              | HC Vs UCD  | 22.05   | 3.99E-01 |
| OTU     | Taxonomy                  | HC Vs UCD | HC Vs TCD | UCD Vs TCD |
|---------|---------------------------|-----------|-----------|------------|
| OTU_77  | Lachnoclostridium         | 47.57     | 120.45    | 128.1      |
|         |                           | -0.75     | 1.11      | 2.40       |
|         |                           | 1.61E-01  | 1.82E-02  | 7.94E-05   |
|         |                           | 3.32E-01  | 6.59E-02  | 1.14E-03   |
| OTU_343 | Christensenellaceae       | 1.84      | 4.18      | 4.98       |
|         |                           | -0.42     | 1.82      | 2.14       |
|         |                           | 2.68E-01  | 4.55E-07  | 7.97E-05   |
|         |                           | 4.55E-01  | 1.36E-05  | 1.14E-03   |
| OTU_355 | Lachnoclostridium         | 24.25     | 43.33     | 46.44      |
|         |                           | -0.88     | 1.19      | 2.01       |
|         |                           | 9.18E-02  | 3.82E-03  | 1.19E-04   |
|         |                           | 2.20E-01  | 1.9E-02   | 1.60E-03   |
| OTU_98  | Ruminococcaceae           | 58.58     | 118.5     | 139.21     |
|         |                           | -0.55     | 1.54      | 2.06       |
|         |                           | 3.0E-01   | 1.46E-04  | 1.39E-04   |
|         |                           | 4.9E-01   | 1.36E-03  | 1.80E-03   |
| OTU_282 | Christensenellaceae       | 4.57      | 11.23     | 12.46      |
|         |                           | -0.49     | 1.37      | 2.21       |
|         |                           | 3.34E-01  | 1.8E-03   | 1.59E-04   |
|         |                           | 5.24E-01  | 1.1E-02   | 1.92E-03   |
| OTU_1104| Akkermansia               | 8.04      | 17.76     | 21.41      |
|         |                           | -0.82     | 1.65      | 2.46       |
|         |                           | 9.76E-02  | 2.25E-04  | 1.56E-04   |
|         |                           | 2.32E-01  | 1.97E-03  | 1.92E-03   |
| OTU_46  | Incertae Sedis            | 192.02    | 480.32    | 512.68     |
|         |                           | -0.67     | 1.09      | 2.34       |
|         |                           | 1.88E-01  | 1.38E-02  | 1.94E-04   |
|         |                           | 3.69E-01  | 6.17E-02  | 2.24E-03   |
| OTU_63  | Blautia                   | 154.0     | 341.42    | 415.46     |
|         |                           | -0.70     | 1.69      | 2.33       |
|         |                           | 2.24E-01  | 4.03E-04  | 2.2E-04    |
|         |                           | 4.02E-01  | 3.14E-03  | 2.43E-03   |
| OTU_176 | Lachnoclostridium         | 10.20     | 19.53     | 21.77      |
|         |                           | -0.71     | 1.37      | 2.03       |
|         |                           | 1.39E-01  | 8.38E-04  | 2.17E-04   |
|         |                           | 3.04E-01  | 5.89E-03  | 2.43E-03   |
| OTU_134 | Ruminiclostridium         | 30.41     | 46.6      |           |
|         |                           | -1.31     | 0.72      |           |
|         |                           | 9.4E-03   | 7.65E-02  | 1.94E-01   |
| OTU       | Group 1     | Group 2     | log2Fold Change | p-value 1   | p-value 2   |
|-----------|-------------|-------------|----------------|------------|------------|
| OTU_136   | HC vs UCD   | 117.66      | -1.32          | 5.59E-02   | 1.57E-01   |
| Eubacterium coprostanoligenes group | HC vs TCD   | 158.56      | 0.38           | 4.9E-01    | 6.54E-01   |
|           | UCD vs TCD  | 218.81      | 2.38           | 1.57E-03   | 1.13E-02   |
| OTU_245   | HC vs UCD   | 11.55       | -0.64          | 2.71E-01   | 4.57E-01   |
| Christensenellaceae R-7 group | HC vs TCD   | 23.45       | 1.52           | 1.55E-03   | 9.45E-03   |
|           | UCD vs TCD  | 25.55       | 2.00           | 2.23E-03   | 1.54E-02   |

n/a: not applicable; log2Fold change is negative when the second group in the pairwise comparison has lower relative abundance than the first one; OTU: operational taxonomic unit
Supplementary Table 4: OTUs with significantly different relative abundance in faecal samples of (a) paired data from 13 CD children at diagnosis and at six and 12 months after the initiation of GFD, and (b) from cross-sectional data from UCD patients (n=20) and TCD patients (n=45)

| OTU | Group comparison | BaseMean | log2Fold Change | p-value | p-value (adjusted) |
|-----|------------------|----------|-----------------|---------|--------------------|
| OTU_576 Akkermansia | UCD Vs GFD 6 mos | 212.2 | -8.53 | 1.79E-05 | 1.27E-03 |
| | UCD Vs GFD 12 mos | 109.1 | -7.00 | 3.20E-06 | 1.31E-03 |
| | UCD Vs TCD | 124.9 | 1.19 | 1.27E-01 | 2.38E-01 |
| OTU_6 Dialister | UCD Vs GFD 6 mos | 1,428.3 | 6.82 | 1.99E-06 | 2.83E-04 |
| | UCD Vs GFD 12 mos | 1,283.0 | 6.88 | 1.48E-05 | 1.52E-03 |
| | UCD Vs TCD | 1,540.8 | -0.21 | 7.67E-01 | 8.48E-01 |
| OTU_5 | UCD Vs GFD 6 mos | 1,522.3 | -9.80 | 5.98E-11 | 1.70E-08 |
| Phascolarctobacterium | UCD Vs GFD 12 mos | 1,256.2 | -7.66 | 8.71E-06 | 1.52E-03 |
| | UCD Vs TCD | 2,385.3 | 0.83 | 3.26E-01 | 4.69E-01 |
| OTU_537 Dialister | UCD Vs GFD 6 mos | 25.8 | 6.82 | 2.51E-05 | 1.42E-03 |
| | UCD Vs GFD 12 mos | 25.9 | 7.69 | 1.15E-05 | 1.52E-03 |
| | UCD Vs TCD | 4.68 | 0.23 | 3.11E-01 | 7.29E-01 |
| OTU_8 Ruminococcus 2 | UCD Vs GFD 6 mos | 2,208.9 | 4.79 | 6.18E-03 | 4.46E-02 |
| | UCD Vs GFD 12 mos | 3,458.3 | 6.24 | 1.71E-04 | 1.40E-02 |
| | UCD Vs TCD | 2,692.5 | 0.21 | 7.48E-01 | 8.33E-01 |
| OTU_43 Veillonella | UCD Vs GFD 6 mos | 343.6 | 4.80 | 6.14E-04 | 1.43E-02 |
| | UCD Vs GFD 12 mos | 1,222.8 | 5.26 | 3.85E-04 | 2.26E-02 |
| | UCD Vs TCD | 367.3 | -0.42 | 4.66E-01 | 6.00E-01 |
| OTU_55 [Eubacterium] coprostanoligenes group | UCD Vs GFD 6 mos | 119.2 | 6.46 | 1.39E-03 | 2.28E-02 |
| | UCD Vs GFD 12 mos | 129.2 | 6.76 | 3.69E-04 | 2.26E-02 |
| | UCD Vs TCD | 86.6 | -0.38 | 5.89E-01 | 7.09E-01 |
| OTU_1045 Roseburia | UCD Vs GFD 6 mos | 98.6 | 0.85 | 5.22E-01 | 8.06E-01 |
| | UCD Vs GFD 12 mos | 140.7 | 5.49 | 5.53E-04 | 2.52E-02 |
| | UCD Vs TCD | 276.2 | -0.27 | 6.85E-01 | 7.87E-01 |
| OTU_50 Tyzzerella 4 | UCD Vs GFD 6 mos | 59.9 | -6.20 | 1.22E-04 | 3.84E-03 |
| | UCD Vs GFD 12 mos | 62.6 | -5.39 | 5.08E-04 | 2.52E-02 |
| OTU  | Taxon  | Treatment 1 | Treatment 2 | p-value 1 | p-value 2 |
|------|--------|-------------|-------------|-----------|-----------|
| OTU_191 | Christensenellaceae | UCD Vs TCD | 55.6 | 1.74 | 5.06E-03 | 3.02E-02 |
| R-7 group | GFD 6 mos | 16.9 | 6.09 | 1.01E-04 | 3.57E-03 |
| OTU_22 | Ruminococcaceae | UCD Vs TCD | 17.4 | 6.00 | 6.65E-04 | 2.73E-02 |
| UCG-002 | GFD 12 mos | 14.4 | -0.74 | 2.07E-01 | 3.39E-01 |
| OTU_18 | Phascolarctobacterium | UCD Vs TCD | 748.0 | 3.92 | 1.29E-03 | 2.28E-02 |
| OTU_166 | Cronobacter | UCD Vs TCD | 772.1 | 6.01 | 8.30E-04 | 2.84E-02 |
| OTU_113 | Clostridium sensu stricto 1 | UCD Vs TCD | 785.6 | -0.22 | 7.35E-01 | 8.24E-01 |
| OTU_259 | Erysipelatoclostridium | UCD Vs TCD | 732.0 | -0.48 | 7.88E-01 | 9.85E-01 |
| OTU_24 | Lachnoclostridium | UCD Vs TCD | 109.8 | -5.38 | 7.91E-04 | 2.84E-02 |
| OTU_14 | Subdoligranulum | UCD Vs TCD | 9.4 | 0.96 | 1.36E-01 | 2.46E-01 |
| OTU_83 | Alistipes | UCD Vs TCD | 25.3 | -7.90 | 5.84E-06 | 5.53E-04 |
| OTU_66 | Ruminococcaceae | UCD Vs TCD | 268.6 | 5.73 | 7.91E-05 | 3.21E-03 |
| UCG-005 | GFD 12 mos | 297.8 | 4.92 | 2.03E-03 | 6.43E-02 |
| UCD Vs GFD 6 mos | 26.8 | -6.95 | 7.37E-05 | 3.21E-03 |
| UCD Vs TCD | 3.57 | -0.02 | 9.91E-01 | 9.98E-01 |
| UCD Vs TCD | 3.4 | 0.65 | 1.54E-01 | 2.73E-01 |
| UCD Vs TCD | 26.8 | -6.95 | 7.37E-05 | 3.21E-03 |
| UCD Vs TCD | 10.5 | -1.98 | 2.78E-01 | 9.98E-01 |
| UCD Vs TCD | 8.6 | 2.61 | 4.31E-07 | 1.64E-05 |
| UCD Vs TCD | 1,315.9 | 5.22 | 2.79E-04 | 7.91E-03 |
| UCD Vs TCD | 1,471.1 | 2.51 | 9.34E-02 | 6.32E-01 |
| UCD Vs TCD | 939.0 | -0.70 | 2.28E-01 | 2.83E-02 |
| UCD Vs TCD | 1,006.0 | 4.16 | 6.55E-04 | 1.43E-02 |
| UCD Vs TCD | 1,258.2 | 2.21 | 1.30E-01 | 7.30E-01 |
| UCD Vs TCD | 1,446.3 | 1.24 | 2.59E-02 | 9.09E-02 |
| UCD Vs TCD | 93.9 | 5.52 | 6.50E-04 | 1.43E-02 |
| UCD Vs TCD | 113.2 | 5.40 | 9.00E-03 | 1.83E-01 |
| UCD Vs TCD | 132.2 | 0.65 | 3.06E-01 | 4.49E-01 |
| UCD Vs TCD | 558.9 | 4.49 | 7.20E-04 | 1.46E-02 |
| UCD Vs TCD | 505.6 | 2.38 | 1.64E-01 | 7.98E-01 |
| UCD Vs TCD | 304.6 | 0.77 | 1.55E-01 | 2.73E-01 |
| OTU      | Family                     | UCD Vs GFD 6 mos | UCD Vs GFD 12 mos | UCD Vs TCD |
|----------|----------------------------|------------------|-------------------|-----------|
| OTU_98   | Ruminococcaceae            | 57.0             | 53.1              | 139.2     |
|          |                            | -3.66            | -1.81             | 2.06      |
|          |                            | 8.64E-04         | 2.57E-01          | 1.39E-04  |
| OTU_3    | Akkermansia                | 1,033.2          | 1,582.8           | 2,466.0   |
|          |                            | 5.40             | 2.19              | 1.45      |
|          |                            | 1.45E-03         | 2.35E-01          | 5.12E-02  |
| OTU_129  | Eggerthella                | 73.1             | 72.6              | 70.16     |
|          |                            | -3.76            | -1.69             | 0.53      |
|          |                            | 2.99E-03         | 2.12E-01          | 2.89E-01  |
| OTU_893  | [Eubacterium] hallii group | 82.7             | 85.3              | 50.2      |
|          |                            | -2.95            | -1.93             | 0.48      |
|          |                            | 2.60E-03         | 8.97E-02          | 3.50E-01  |
| OTU_303  | Lachnospiraceae            | 46.0             | 62.5              | 114.8     |
|          | UCG-008                    | -3.69            | -0.42             | 1.47      |
|          |                            | 2.90E-03         | 7.67E-01          | 3.77E-03  |
| OTU_115  | Parasutterella             | 56.8             | 73.4              | 36.5      |
|          |                            | 4.42             | 3.26              | 1.62      |
|          |                            | 2.33E-03         | 4.43E-02          | 5.14E-03  |
| OTU_1005 | Fusicatenibacteri          | 9.71             | 8.7               | 4.25      |
|          |                            | -4.70            | -2.47             | 0.11      |
|          |                            | 2.31E-03         | 1.28E-01          | 8.07E-01  |
| OTU_133  | Flavonifractor             | 57.0             | 59.1              | 99.5      |
|          |                            | -3.64            | -2.48             | 1.02      |
|          |                            | 4.45E-03         | 1.07E-01          | 5.73E-02  |
| OTU_1044 | [Eubacterium] hallii group | 150.3            | 149.9             | 152.3     |
|          |                            | -3.14            | -2.11             | 0.04      |
|          |                            | 5.18E-03         | 1.40E-01          | 9.38E-01  |
| OTU_185  | Dorea                      | 19.4             | 20.0              | 6.5       |
|          |                            | 4.18             | 3.10              | -0.05     |
|          |                            | 5.41E-03         | 7.03E-02          | 9.29E-01  |
| OTU_90   | Ruminococcaceae            | 103.6            | 82.7              | 50.2      |
|          |                            | 4.71             | 2.99E-03          | 0.48      |
|          |                            | 5.05E-03         | 3.14E-02          | 4.91E-01  |
| OTU   | Genus      | UCD Vs GFD 6 mos | UCD Vs GFD 12 mos | UCD Vs TCD |
|-------|------------|------------------|-------------------|------------|
| UCG-002 |            | 92.9             | 83.8             | 83.8       |
| OTU_122 | Ruminiclostridium 9 | 25.3             | 25.3             | 16.9       |
| OTU_15 | Dorea      | 275.7            | 446.8            | 538.8      |
| OTU_68 | Ruminococcaceae | 152.6            | 77.5             | 150.1      |
| UCG-005 |            |                  |                  |            |

n/a: not applicable; log2Fold change is negative when the second group in the pairwise comparison has lower relative abundance than the first one; OTU: operational taxonomic unit
**Supplementary Table 5:** OTUs with significantly different relative abundance in faecal samples of 18 TCD children and their unaffected 19 siblings

| OTU/Group Description                                      | BaseMean | log2Fold Change | p-value | p-value (adjusted) | Higher in |
|-----------------------------------------------------------|----------|-----------------|---------|--------------------|-----------|
| OTU_38 [Eubacterium] coprostanoligenes group              | 75.9     | 5.27            | 3.09E-12| 1.88E-09           | SIBLINGS  |
| OTU_107 Prevotella 7                                      | 42.2     | 5.05            | 7.62E-10| 2.32E-07           | SIBLINGS  |
| OTU_111 Blautia                                           | 34.0     | 4.17            | 4.98E-09| 1.01E-06           | SIBLINGS  |
| OTU_62 Ruminococcaceae UCG-014                           | 19.1     | 4.01            | 3.05E-08| 4.65E-06           | SIBLINGS  |
| OTU_114 Coprococcus 2                                     | 12.9     | 3.85            | 9.00E-08| 1.10E-05           | SIBLINGS  |
| OTU_568 Bifidobacterium bifidum NCIMB 41171              | 232.0    | 4.26            | 2.10E-07| 1.72E-05           | SIBLINGS  |
| OTU_135 Holdemanella                                      | 10.6     | 3.39            | 2.26E-07| 1.72E-05           | SIBLINGS  |
| OTU_143 Ruminococcus 1                                    | 9.9      | 3.51            | 2.18E-07| 1.72E-05           | SIBLINGS  |
| OTU_295 Christensenellaceae R-7 group                     | 6.8      | 3.43            | 4.18E-07| 2.83E-05           | SIBLINGS  |
| OTU_1059 [Eubacterium] coprostanoligenes group           | 8.5      | 3.61            | 2.93E-06| 1.79E-04           | SIBLINGS  |
| OTU_228 Faecalibacterium                                  | 6.5      | 3.36            | 3.61E-06| 2.00E-04           | SIBLINGS  |
| OTU_212 Ruminiclostridium 9                              | 7.5      | 2.73            | 6.28E-06| 3.06E-04           | SIBLINGS  |
| OTU_74 [Eubacterium] ruminantium group                    | 5.7      | 3.18            | 6.53E-06| 3.06E-04           | SIBLINGS  |
| OTU_78 Ruminiclostridium 5                               | 11.2     | -3.08           | 1.11E-05| 4.60E-04           | TCD       |
| OTU_140 Collinsella                                       | 6.9      | 3.13            | 1.13E-05| 4.60E-04           | SIBLINGS  |
| OTU_230 Bacteroides                                       | 6.8      | -2.84           | 1.21E-05| 4.62E-04           | TCD       |
| OTU_299 [Ruminococcus] gauvreauii group                   | 9.2      | -2.72           | 1.71E-05| 6.07E-04           | TCD       |
| OTU_324 Peptococcaceae                                    | 4.7      | 2.60            | 1.79E-05| 6.07E-04           | SIBLINGS  |
| OTU_88 Bacteroides                                        | 954.3    | -2.89           | 1.95E-05| 6.23E-04           | TCD       |
| OTU    | Family               | Genus          | Fraction | Abundance | p-value  | T-test  | Group     |
|--------|----------------------|----------------|----------|-----------|----------|---------|-----------|
| OTU_665 | Ruminococaceae      | UCG-014        | 9.1      | 3.26      | 2.20E-05 | 6.68E-04| SIBLINGS  |
| OTU_130 | Christensenellaceae | R-7 group      | 10.1     | 2.72      | 2.67E-05 | 7.39E-04| SIBLINGS  |
| OTU_576 | Akkermansia         |                | 28.4     | -3.03     | 3.18E-05 | 8.43E-04| TCD       |
| OTU_80  | Ruminococaceae      | UCG-014        | 5.7      | 2.77      | 3.39E-05 | 8.60E-04| SIBLINGS  |
| OTU_70  | Ruminococaceae      | UCG-014        | 5.2      | 2.75      | 4.15E-05 | 1.01E-03| SIBLINGS  |
| OTU_178 | Ruminococaceae      |                | 8.4      | 2.42      | 4.96E-05 | 1.16E-03| SIBLINGS  |
| OTU_42  | Methanobrevibacter  |                | 17.2     | 3.21      | 5.55E-05 | 1.25E-03| SIBLINGS  |
| OTU_908 | Anaerostipes        |                | 4.3      | 2.50      | 7.12E-05 | 1.46E-03| SIBLINGS  |
| OTU_85  | Bacteroidales       | S24-7 group    | 6.5      | -2.68     | 7.21E-05 | 1.46E-03| TCD       |
| OTU_272 | [Eubacterium]       | oxidoreducens group | 3.4 | 2.35 | 7.14E-05 | 1.46E-03| SIBLINGS  |
| OTU_180 | Ruminiclostridium  | 5              | 6.2      | 2.71      | 8.01E-05 | 1.57E-03| SIBLINGS  |
| OTU_442 | Christensenellaceae |                | 3.7      | 2.44      | 9.44E-05 | 1.80E-03| SIBLINGS  |
| OTU_163 | Ruminiclostridium  | 5              | 26.9     | -2.34     | 1.12E-04 | 2.07E-03| TCD       |
| OTU_33  | Roseburia           |                | 427.7    | -2.68     | 1.17E-04 | 2.09E-03| TCD       |
| OTU_622 | Lachnospiraceae     | UCG-001        | 52.9     | -2.73     | 1.26E-04 | 2.20E-03| TCD       |
| OTU_783 | Bacteroides         |                | 4.4      | -2.05     | 2.31E-04 | 3.71E-03| TCD       |
| OTU_315 | Christensenellaceae | R-7 group      | 8.9      | 2.40      | 2.25E-04 | 3.71E-03| SIBLINGS  |
| OTU_337 | Coriobacteriaceae   |                | 7.0      | 2.45      | 2.31E-04 | 3.71E-03| SIBLINGS  |
| OTU_67  | Christensenellaceae | R-7 group      | 53.2     | 2.71      | 2.56E-04 | 4.00E-03| SIBLINGS  |
| OTU_434 | Christensenellaceae |                | 3.7      | 2.19      | 3.06E-04 | 4.55E-03| SIBLINGS  |
| OTU_155 | Lachnoclostridium   |                | 23.8     | -2.56     | 3.25E-04 | 4.71E-03| TCD       |
| OTU_179 | Howardella          |                | 3.6      | 2.28      | 4.45E-04 | 6.30E-03| SIBLINGS  |
| OTU          | Taxonomy                          | Relative Abundance | % Exp. | p-Value | TCD     | SIBLINGS |
|--------------|-----------------------------------|--------------------|--------|---------|---------|----------|
| OTU_441 Thalassospira | 3.8 2.13 4.58E-04 6.34E-03 | SIBLINGS |
| OTU_895 Barnesiella  | 21.4 -2.62 5.37E-04 7.26E-03 | TCD     |
| OTU_868 Bacteroides   | 112.1 -2.68 6.00E-04 7.77E-03 | TCD     |
| OTU_704 Collinsella   | 91.7 -2.16 6.34E-04 7.88E-03 | TCD     |
| OTU_133 Flavonifractor| 82.0 -2.07 9.31E-04 1.09E-02 | TCD     |
| OTU_225 Butyrivibrio   | 3.1 2.03 1.16E-03 1.26E-02 | SIBLINGS |
| OTU_149 Lachnospiraceae NK4A136 group | 25.9 -2.37 1.55E-03 1.58E-02 | TCD     |
| OTU_195 Bacteroides plebeius DSM 17135 | 12.2 -2.48 1.84E-03 1.81E-02 | TCD     |
| OTU_116 Ruminococcaceae UCG-014 | 9.2 2.07 2.64E-03 2.40E-02 | SIBLINGS |
| OTU_1105 Prevotella 9 | 192.5 3.03 3.33E-03 2.98E-02 | SIBLINGS |
| OTU_137 Ruminococcaceae UCG-014 | 6.6 2.08 3.48E-03 3.05E-02 | SIBLINGS |
| OTU_1049 Lachnoclostridium | 181.3 -2.06 4.73E-03 3.60E-02 | TCD     |
| OTU_63 Blautia        | 388.5 -2.06 4.87E-03 3.66E-02 | TCD     |
| OTU_156 Christensenellaceae R-7 group | 58.7 2.25 4.94E-03 3.76E-02 | SIBLINGS |
| OTU_1054 Alistipes    | 45.1 -2.01 5.82E-03 4.07E-02 | TCD     |

OTU: operational taxonomic unit
### Supplementary Table 6: Faecal metabolites in coeliac disease children with or without recent ingestion of gluten, based on GIP results

| Metabolite                        | GIP negative (59) | GIP positive (12) | p*  |
|-----------------------------------|-------------------|-------------------|-----|
| Ammonia \(10^{-4}\) mg/g          | 11.2 (1.8)        | 7.7 (6.1)         | 0.043 |
| Free sulphide \(\mu\)mol/g        | 0.11 (0.03)       | 0.08 (0.01)       | 0.228 |
| Total sulphide \(\mu\)mol/g       | 1.2 (0.25)        | 0.8 (0.08)        | 0.121 |
| L-lactic acid \(\mu\)g/g          | 106.4 (30.6)      | 86.4 (6.5)        | 0.750 |
| D-lactic acid \(\mu\)g/g          | 80.5 (8.4)        | 87.1 (7.5)        | 0.925 |
| Total lactic acid \(\mu\)g/g      | 186.9 (33.9)      | 173.5 (9.1)       | 0.414 |
| Acetic acid \(\mu\)mol/g          | 110.2 (9.5)       | 125.7 (6.1)       | 0.338 |
| Propionic acid \(\mu\)mol/g       | 22.8 (3.1)        | 23.0 (1.7)        | 0.956 |
| Butyric acid \(\mu\)mol/g         | 19.3 (3.1)        | 21.0 (2.2)        | 0.675 |
| Valeric acid \(\mu\)mol/g         | 2.5 (0.51)        | 2.5 (0.19)        | 0.856 |
| Isobutyric acid \(\mu\)mol/g      | 3.1 (0.42)        | 3.0 (0.20)        | 0.794 |
| Isovaleric acid \(\mu\)mol/g      | 3.2 (0.49)        | 2.95 (0.22)       | 0.640 |
| Total SCFA \(\mu\)mol/g           | 161.0 (13.3)      | 178.0 (9.9)       | 0.722 |

Values expressed as mean (SEM); * Mann-Whitney non-parametric test; gluten ingestion indicated by a faecal GIP concentration \(\geq 0.156\)μg/g