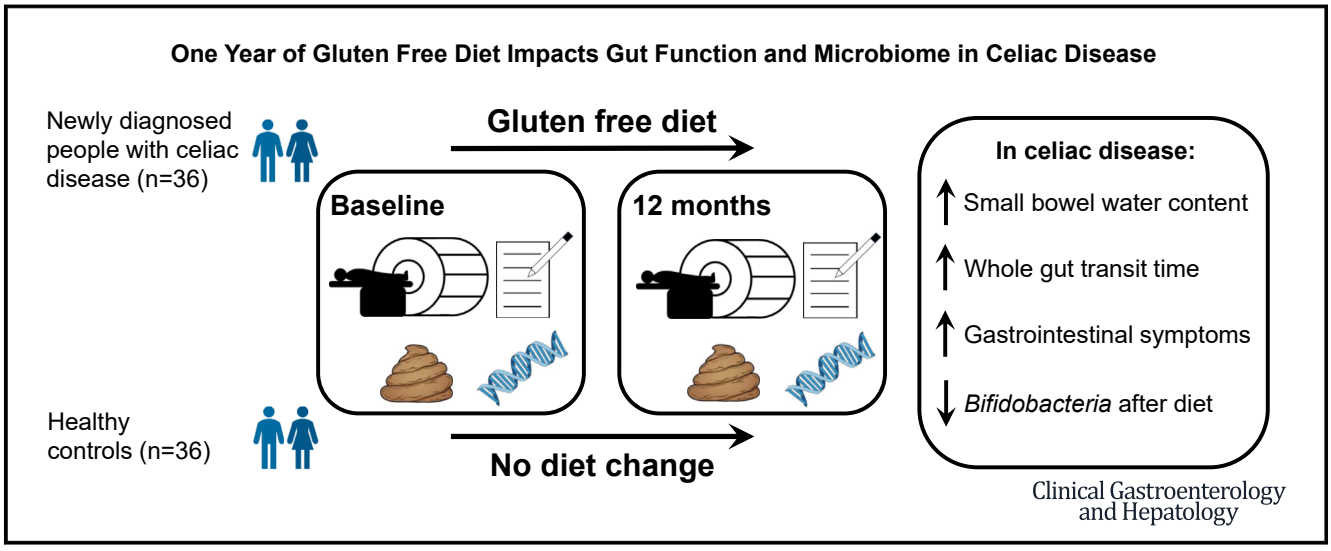
Clinical Gastroenterology and Hepatology 2025;■:■–■

One Year of Gluten-Free Diet Impacts Gut Function and Microbiome in Celiac Disease

Carolyn M. Costigan,[1,2,*] Frederick J. Warren,[3,*] Anthony P. Duncan,[3,4] Caroline L. Hoad,[5] Nina Lewis,[6] Trevor Hill,[1,2] Colin J. Crooks,[1,2] Paul S. Morgan,[7] Carolina Ciacci,[8] Paola Iovino,[8] David S. Sanders,[9] Falk Hildebrand,[3,4] Penny A. Gowland,[5] Robin C. Spiller,[1,2] and Luca Marciani[1,2]

[1]Nottingham Digestive Diseases Centre, Translational Medical Sciences, School of Medicine, University of Nottingham, Nottingham, United Kingdom; [2]NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom; [3]Quadram Institute Bioscience, Norwich, United Kingdom; [4]Digital Biology, Earlham Institute, Norwich Research Park, Norwich, Norfolk, United Kingdom; [5]Sir Peter Mansfield Imaging Centre, School of Physics and Astronomy, University of Nottingham, Nottingham, United Kingdom; [6]Gastroenterology, Nottingham University Hospitals NHS Trust, Nottingham, United Kingdom; [7]Medical Physics, School of Medicine, University of Nottingham, Nottingham, United Kingdom; [8]Department of Medicine and Surgery, University of Salerno, Salerno, Italy; and [9]Department of Clinical Medicine, School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom



BACKGROUND & AIMS: Currently, the main treatment for celiac disease (CD) is the gluten-free diet (GFD). This observational cohort study investigated the impact of CD and 1 year of GFD on gut function and microbiome.

METHODS: A total of 36 newly diagnosed patients and 36 healthy volunteers (HVs) were studied at baseline and at 12-month follow-up. Small bowel water content (SBWC), whole gut transit time (WGTT), and colon volumes were measured by magnetic resonance imaging. Stool sample DNA was subjected to shotgun metagenomic sequencing. Species-level abundances and gene functions, including CAZymes (carbohydrate active enzymes) were determined.

RESULTS: SBWC was significantly higher in people with CD (157 ± 15 mL) vs (HVs 100 ± 12 mL) ($P = .003$). WGTT was delayed in people with CD (68 ± 8 hours) vs HVs (41 ± 5 hours) ($P = .002$). The

*Authors share co-first authorship.

Abbreviations used in this paper: CD, Celiac disease; GFD, gluten-free diet; GI, gastrointestinal; HADS, Hospital Anxiety and Depression Scale; HV, healthy volunteer; MRI, magnetic resonance imaging; PHQ-15, Patient Health Questionnaire-15; SBWD, small bowel water content; VAS, visual analog scale; WGTT, whole gut transit time.

1542-3565
https://doi.org/10.1016/j.cgh.2024.11.006

differences reduced after 12 months of GFD but not significantly. Well-being in the CD group significantly improved after GFD but did not recover to control values. CD fecal microbiota showed a high abundance of proteolytic gene functions, associated with *Escherichia coli*, *Enterobacter*, and *Peptostreptococcus*. GFD significantly reduced *Bifidobacteria* and increased *Blautia wexlerae*. Microbiome composition correlated positively with WGTT, colonic volume, and *Akkermansia municphilia* but negatively with *B wexerelae*. Following GFD, the reduction in WGTT and colonic volume was significantly associated with increased abundance of *B wexlerae*. There were also significant alterations in CAZyme profiles, specifically starch- and arabinoxylan-degrading families.

CONCLUSIONS: **CD impacted gut function and microbiota. GFD ameliorated but did not reverse these effects, significantly reducing *Bifidobacteria* associated with reduced intake of resistant starch and arabinoxylan from wheat. ClinicalTrials.gov, number: NCT02551289.**

Keywords: Celiac Disease; Microbiome; Gluten-Free Diet; MRI.

Celiac disease (CD) requires a lifelong commitment to a gluten-free diet (GFD) treatment.[1,2] The aim of GFD is recovery of the small bowel mucosa and reversal of the enteropathy. Despite a long-term adherence to GFD, a significant proportion of people with CD report persistent gastrointestinal (GI) symptoms.[3]

Small bowel water content (SBWC) in untreated patients with CD, measured using magnetic resonance imaging (MRI), was increased, a feature hypothesized to be due to a combination of impaired motility and absorption.[4] Manometric, breath test, and camera pill studies described an underlying GI dysmotility with prolonged orocecal transit times.[5] Such GI motor disorders tended to resolve on GFD,[6] suggesting that the motor dysfunction may be related to mucosal inflammation.[7]

CD also impacts the gut microbiome before GFD, with reported increased abundance of Firmicutes and Proteobacteria and reductions in beneficial *Bifidobacterium*.[8] Studies in CD have identified higher levels of *Escherichia coli*, Clostridiaceae, and Enterobacteriaceae, and lower *Bifidobacterium* in the duodenal mucosa.[9,10] The GFD tends to provide a lower dietary fiber intake than habitual diets[11] and may have effects on the gut microbiome independent of CD. In healthy volunteers (HVs), GFD reduced the abundance of *Bifidobacterium* (specifically *Bifidobacterium longum*) and *Lactobacillus*, and increased the abundance of Enterobacteriaceae (specifically *E coli*).[12] There may also be a difference in the metabolic activity of intestinal microbial flora in children with CD compared with those without the disease, and this difference may not be affected by diet treatment.[13]

The reasons for alterations in the gut microbiome of CD patients and correlations with GI function and symptoms are not clearly understood. This study aimed to investigate the SBWC, gut transit time, colon volumes, symptoms, and microbiome of CD patients at diagnosis, and the impact of following 1 year of GFD using combined noninvasive MRI measurements with microbiome analysis using shotgun metagenomic approaches. The primary hypothesis was that GFD will reduce the fasting SBWC.

Materials and Methods

Study population

Thirty-six patients newly diagnosed with CD were prospectively recruited before starting treatment with a GFD. The patients commenced their GFD treatment immediately after their baseline study visit. Adherence was measured by repeating serology testing at 12 month follow-up, as well as using Biagi scores.[14] The 12-month follow-up group included patients providing a stool sample and having GFD diet adherence confirmed by tissue transglutaminase blood sample and Biagi score $\geq$3.

A parallel group of 36 HVs frequency matched for age and sex were prospectively recruited for comparison data. They were not following a GFD diet.

Both the CD and HVs groups were scanned twice, approximately 12 months apart. The 12-month follow-up group included healthy participants who did not exclude gluten from the diet, provided a stool sample, and had confirmed negative CD blood sample.

The CD group eligibility criteria also included being 18–70 years of age; being on a gluten-containing diet; having IgA-TG2 or IgA-DGP, or IgG-DGP testing, and duodenal biopsy confirming CD; being recruited within a month of their duodenal biopsy (Supplementary Table 1). The HVs were required to have a screening blood sample negative for CD and to have no comorbidities. All participants were not to have had any antibiotic or probiotic treatment in the 4 weeks preceding the study days.

Study Protocol

Participants completed Bristol Stool Form Scales for 7 days prior to each study day. On each study day they were asked to provide a stool sample. At 8 AM the day before the study day, the participants were instructed to swallow, before their breakfast, 5 inert plastic MRI transit capsule markers to measure whole gut transit time (WGTT).[15] They were also asked to fast from their

evening meal, eaten by 8 PM, until the following morning study, with only water allowed until bedtime. In the morning, fasted MRI scans were acquired to measure SBWC, colonic volumes and WGTT.

Fasting breath hydrogen (H_2) was measured using a breath H_2 meter and psychometric assessment was performed using the Hospital Anxiety and Depression Scale (HADS) and the Patient Health Questionnaire-15 (PHQ-15), also scored separately for the 3 GI symptoms (PHQ-3). GI symptom intensity on each study day was also measured using visual analog scales (VASs) (0–100 mm) for abdominal bloating, flatulence, nausea, and abdominal pain.

The patient-reported symptoms were used to divide the patients into 2 groups, 1 with persistent and 1 with resolved symptoms after GFD. Further details are available in the Supplementary Materials.

The study was registered on ClinicalTrials.gov (NCT02551289).

What You Need to Know

Background

The impact of celiac disease Celiac disease (CD) and of gluten-free diet (GFD) treatment on gut function and microbiome are not well understood.

Findings

CD increased small bowel water content, delayed transit (compared with healthy control subjects), increased microbiome proteolytic gene functions, and reduced *Bifidobacteria*. One year of GFD altered starch- and arabinoxylan-degrading families.

Implications for patient care

Assessing gut function and the metagenome can facilitate the development of targeted dietary interventions to add to a GFD to reduce the negative impact of GFD on the microbiome.

Statistical Analysis

There were no data available to estimate the size of the change in SBWC in patients after a GFD. Inference was drawn from previous MRI data[4] in 20 untreated, newly diagnosed patients with CD that showed fasting SBWC of 202 ± 115 mL. From this, it was predicted that we could detect a change of 40% (a reduction of 80 mL volume) after GFD with a power of 90% and a type I error probability of 0.05 using 24 patients in a paired study design. This would be considered a clinically significant change after GFD. A recruitment target of 36 was planned to allow for dropouts and to increase power for secondary outcomes.

All analyses, both descriptive and statistical, were carried out in Stata MP4 v18.0 (StataCorp). Normality of data was assessed via histogram with the normal distribution overlaid. For comparing outcomes between the study groups, differences between mean values at baseline and mean changes from baseline to follow-up were compared via unpaired *t* test. Where the normality of the differences was in question, the analysis was repeated using the unpaired 2-sample Wilcoxon test. Statistical tests were 2-sided. A *P* value of .05 was used to assess statistical significance. Further details of bioinformatics are in Supplementary Materials. All data are presented as mean ± SEM.

All authors had access to the study data and reviewed and approved the final manuscript.

Ethics Approval

The patients underwent clinical tests and assessments as part of their routine care. This study involved human participants and ethics approval was obtained from the United Kingdom National Research Ethics Service (approval number 14MP002). The study was conducted in accordance with the Declaration of Helsinki (6th revision, 2008). All participants gave informed written consent to participate in the study before taking part.

Data Availability

Raw read data from the metagenomic sequencing runs can be accessed through the NCBI SRA project number PRJNA1023737 at http://www.ncbi.nlm.nih.gov/bioproject/1023737. The associated metadata can be accessed through the University of Nottingham Research Repository (https://doi.org/10.17639/nott.7352). The full MATAFILER pipeline is available at https://github.com/hildebra/MATAF3. The RTK package is available at www.github.com/hildebra/Rarefaction/.

Results

Baseline Characteristics and Differences Between Groups

The study procedures and visits were well tolerated by the participants with few dropouts (Supplementary Figure 1). The 2 cohorts were matched for age, sex, and body mass index, as indicated in Table 1.

Patients with CD had higher levels of somatization at baseline compared with HVs on the PHQ-15 and its subset relating to GI problems, the PHQ-3 (Supplementary Figure 2). They also had higher scores for depression and anxiety and higher levels of GI symptoms on the VAS on the MRI study day. People with CD had no significant differences in bowel habits or short-chain fatty acids compared with HVs but had 5% lower stool water.

At baseline, SBWC of people with CD was 57% higher than that of HVs ($P = .003$) (Figure 1*A*). Baseline WGTT

Table 1. Demographic, Behavioral, and Gastrointestinal Endpoints at Baseline and 12-Month Follow-Up

	CD at Baseline (n = 36)	HVs at Baseline (n = 36)	CD vs HVs at Baseline (*P*)	CD After 12 mo of Gluten-Free Diet	HVs After 12 mo Unrestricted Diet	12-mo Change in CD vs Change in HVs (*P*)
Sex						
Female	25	24	—	12	13	—
Male	11	12	—	8	11	—
Age, y	46 ± 3	42 ± 2	NS	47 ± 3	43 ± 3	NS
Body mass index, kg/m^2	26 ±1	25.6 ±0.7	NS	26 ± 1	26.4 ± 0.8	NS
Biagi score	0.4 ± 0.1	0.0 ± 0.0	<.0001	3.0 ± 0.0	0.0 ± 0.0	<.0001
Biagi score level	All at level 1	All at level 1	NS	All at level 3	All at level 1	<.0001
PHQ-15	11 ± 1	3.0 ± 0.5	<.001	7 ± 1	4.3 ± 0.8	<.001
PHQ-3	2.9 ± 0.3	0.7 ± 0.2	<.0001	2.0 ± 0.4	1.1 ± 0.3	.0096
HADS anxiety score	8.4 ± 0.8	4.8 ± 0.6	<.001	6.2 ± 0.8	5.3 ± 0.6	NS
HADS depression score	5.6 ± 0.7	2.1 ±0.4	<.001	2.8 ± 0.5	3.4 ± 0.7	.003
HADS total score	14 ± 1	6.9 ± 0.9	<.001	9 ± 1	9 ± 1	.006
Study day symptoms VAS, mm	56 ± 8	10 ± 2	<.0001	31 ± 12	11 ± 2	NS
Stool form	3.8 ± 0.2	3.7 ± 0.2	NS	3.4 ± 0.3	3.8 ± 0.2	.0298
Stool frequency, times/day	1.8 ± 0.2	1.5 ± 0.1	NS	1.5 ± 0.1	1.7 ± 0.2	NS
Hydrogen breath test, ppm	7 ± 2	7 ± 2	NS	4 ± 1	7 ± 1	NS
Small bowel water content, mL	157 ± 15	100 ± 12	.003	150 ± 22	111 ± 19	NS
Ascending colon volume, mL	262 ± 11	251 ± 12	NS	249 ± 17	251 ± 17	NS
Transverse colon volume, mL	238 ± 17	213 ± 14	NS	277 ± 24	220 ± 14	NS
Descending colon volume, mL	181 ± 17	205 ± 19	NS	195 ± 19	178 ± 15	NS
Total colon volume, mL	691 ± 38	670 ± 34	NS	721 ± 50	648 ± 38	NS
Whole gut transit time WAPS, AU	2.2 ± 0.3	1.2 ± 0.2	.002	1.9 ± 0.3	1.2 ± 0.2	NS
Whole gut transit time, h	68 ± 8	41 ± 5	.002	60 ± 9	38 ± 5	NS
Acetic acid, μM/g dry stool	23.8 ± 0.9[a]	25.3 ± 0.6[b]	NS	23.5 ± 0.8	26.2 ± 0.7	NS
Propanoic acid, μM/g dry stool	7 ± 1[a]	7.8 ± 0.7[b]	NS	8 ± 1	9.4 ± 0.9	NS
Isobutyric acid, μM/g dry stool	1.5 ± 0.2[a]	1.6 ± 0.1[b]	NS	1.4 ± 0.2	1.6 ± 0.1	NS
Butyric acid, μM/g dry stool	8 ± 1[a]	7.1 ± 0.8[b]	NS	6.3 ± 1.1	8 ± 1	NS
Isovaleric acid, μM/g dry stool	1.9 ± 0.2[a]	1.9 ± 0.2[b]	NS	1.7 ± 0.2	1.8 ± 0.2	NS
Valeric acid, μM/g dry stool	2.0 ± 0.3[a]	2.0 ± 0.2[b]	NS	1.5 ± 0.2	2.2 ± 0.2	NS
Stool water, %	66 ± 2[a]	71 ± 2[b]	.02	68 ± 2	70 ± 2	NS

Values are n or mean ± SEM.
CD, celiac disease; HADS, Hospital Anxiety and Depression Scale; HV, healthy volunteer; PHQ, Patient Health Questionnaire; WAPS, weighted average position score; VAS, visual analog scale.
[a]n = 22 stool samples.
[b]n = 28 stool samples.

was 83% longer in people with CD compared with HVs ($P = .002$) (Figure 1*B*).

Gut transit and colonic volumes may impact gut microbiome composition.[16] Here, WGTT and colonic volumes correlated with several microbial species (Figure 1*C*). Slower transit correlated with *Akkermansia muciniphila*, several species of the genus *Alistipes* and *Bacteroides*, and the species *Ruthenibacterium lactatiformans*, while faster transit times correlated with *Faecalibacterium prausnitzii*, *Gemmiger formicilis*, and several *Agathobacter* species. These findings are consistent with recent studies which have associated these

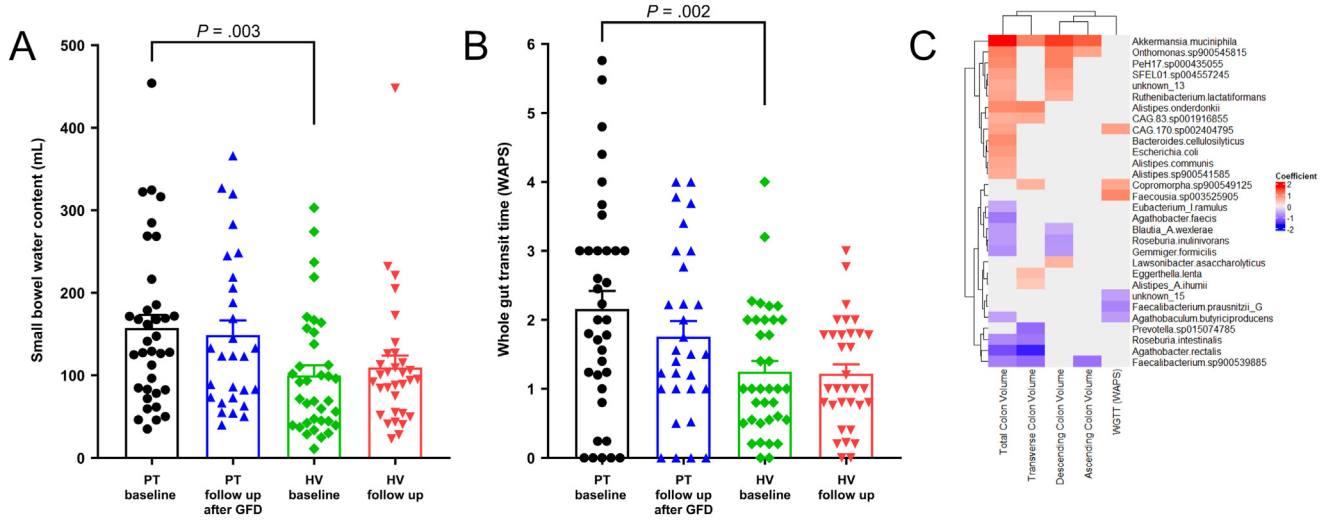


Figure 1. (*A*) SBWC. (*B*) WGTT. (*C*) Multivariate regression analysis showing bacterial species that are significantly associated with MRI markers of WGTT and colonic volumes. Red color indicates positive association and blue color indicates negative association. Data are presented as mean ± SEM. PT, patient.

microbial groups with gut transit time measured with different methods.[16,17] Species-level associations with other endpoints are shown in Supplementary Figures 3–5.

At baseline, prior to starting GFD, while there were no major community-level differences in microbiome composition ($P > 0.05$, permutational multivariate analysis of variance) (Figure 2*A*), there were significant differences in individual taxa observed between CD and HVs (Figure 2*B*). The microbiome in CD was significantly enriched in bacterial taxa including Oscillospiraceae and Peptostreptococcales compared with HVs. Conversely, the HVs had higher abundances of the genus *Ruminococcus_D* and the species *Blautia wexlerae*. It is interesting to note that *Bl wexlerae* was associated with fast WGTT and smaller colonic volumes (Figure 1), so the higher abundance of *Bl wexlerae* in HVs at baseline may reflect differences in passage time and colon volume between HVs and people with CD. At baseline, relatively few differences in CAZyme (carbohydrate active enzyme) composition (Figure 2*C*) were observed between CD and HVs, although higher levels of GH20 and GH33 were observed in CD, both enzyme families involved in host glycan metabolism, and CBM5, involved in chitin metabolism. Several metabolic pathways related to protein metabolism were elevated in CD, possibly reflecting malabsorption of protein in the upper GI tract of patients with CD.

Effects of GFD

At 12-month follow-up, all the patients included in the data had achieved a Biagi score of 3, indicating that they followed a strict GFD. After 12 months of GFD the symptoms and well-being of the group of patients with CD had significantly improved compared with baseline values against the HVs, although most of the endpoints had not recovered to HVs values. The somatization scores in people with CD had significantly reduced by 36% for the PHQ-15 and by 31% in the PHQ-3 subset of GI problems. Their total scores for depression and anxiety reduced by 36% and the VAS GI symptoms reduced by 45%. Stool form in people with CD increased on average by half a point on the Bristol stool scale and stool water increased by 2%. The higher amount of SBWC detected at baseline did not reduce after GFD. WGTT reduced by 14% after GFD, but the change was modest and not significant compared with HVs.

Following 12 months of GFD did not result in community level shifts in the gut microbiome composition of patients with CD (Figure 3*A*) but did result in changes in individual taxa and metabolic taxa, which were distinct from the baseline differences between patients with CD and HVs. While no significant differences were observed between baseline and 12-month follow-up for the HVs, several differences were observed in the patients with CD. There was an increase in abundance of *Blautia_A*, which may be related to the improvements in WGTT. A significant increase in mucin degradation pathways was also observed which may reflect changes in WGTT. There were also increases in abundance in *Bacteroides uniformis*. Notably, there was a significant reduction in Actinomycetales, the class of bacteria encompassing bifidobacterial species in patients with CD on the GFD. This was associated with a reduction in the *Bifidobacterium*–specific metabolic pathway (*Bifidobacterium* shunt), suggesting that a GFD may have specific impacts on bifidobacterial abundances.

At follow-up there was a significant difference between the HVs and the patients with CD in microbiome composition ($P = .002$, F statistic $= 2.1943$, permutational multivariate analysis of variance) (Figure 4*A*). The reduction in *Bifidobacterium* observed after GFD can clearly be observed in the CD group compared with the

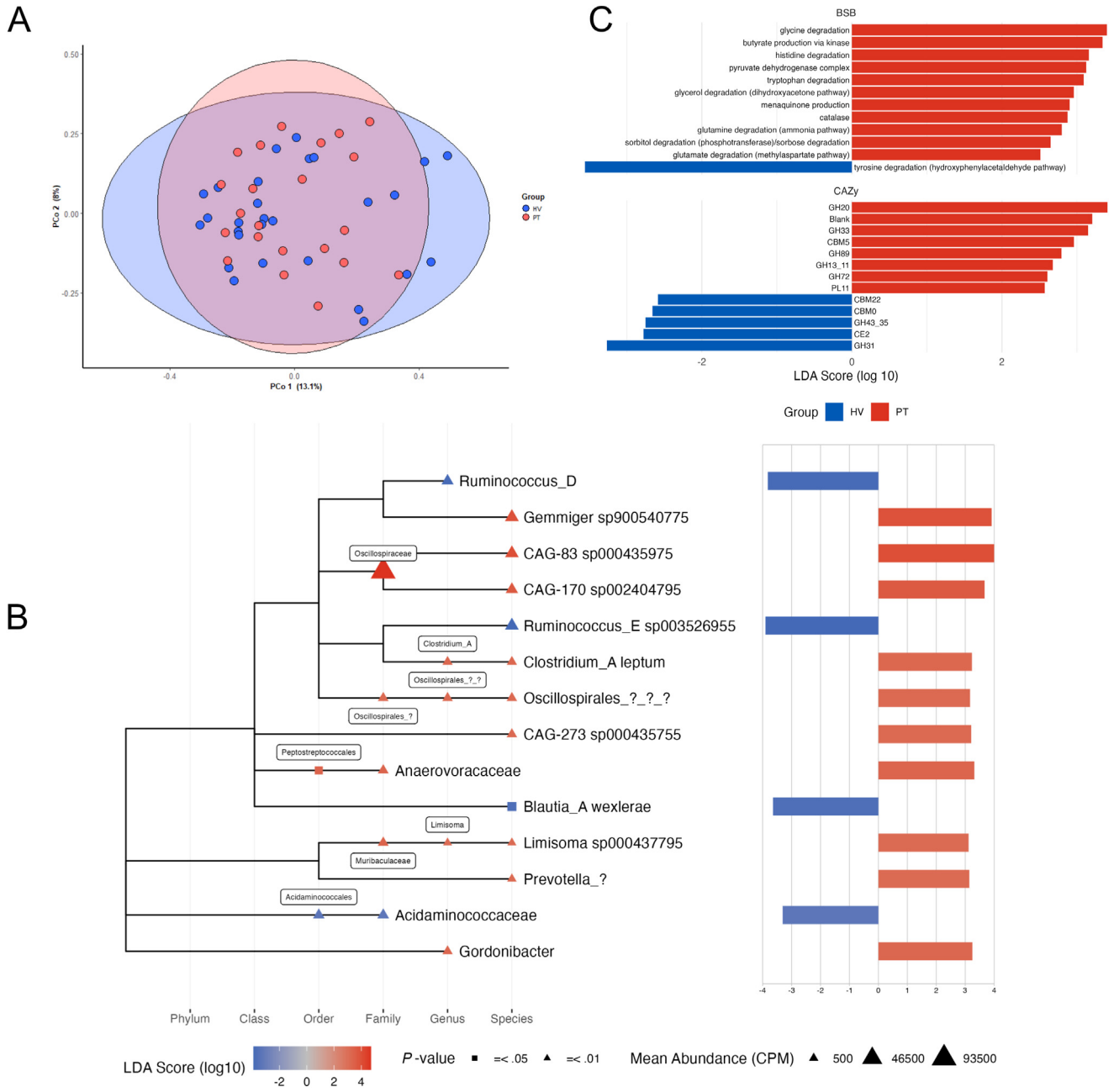


Figure 2. Baseline microbiome composition. (*A*) Principal coordinate (PCo) analysis plot of Bray-Curtis distances of species diversity colored by HV vs patient (PT) (permutational multivariate analysis of variance, $P > .05$). (*B*) Differentially abundant microbial taxa identified by the LEfSe algorithm ($P = .05$) between HVs and PTs at baseline. (*C*) Differentially abundant CAZymes and BSB gene functional pathways ($P = .05$). A blue bar indicates enrichment in the HV group and a red bar indicates enrichment in the PT group.

HVs at follow-up (Figure 4*B*). Specifically, the abundance of *Bi longum* and *Bi breve* was lower in CD patients than HVs. Additionally, *E coli*, *Bacteroides ovatus*, *Alistipes communis*, *Roseburia hominis*, and several other taxa were elevated in CD compared with HVs.

As well as differences in microbial community composition between the people with CD and HVs, alterations in microbial metabolic pathway abundances were observed (Figures 4*B* and *C*) in addition to the differences observed at baseline. These may be related to the removal of the 2 major dietary fiber constituents in wheat, resistant starch and arabinoxylan, caused by GFD. The HVs were found to have significantly higher arabinose and sucrose degradation pathways, related to arabinoxylan and starch metabolism, respectively. The differences in CAZyme profile between HVs and patients also provides evidence for this. The CAZyme profile of HVs was enriched in several GH13 subfamilies and CBM48, which are involved in starch metabolism, and in GT47 and several subfamilies of GH43, which comprise xylanase and arabinofuranosidase enzymes involved in arabinoxylan metabolism. These changes in microbiome

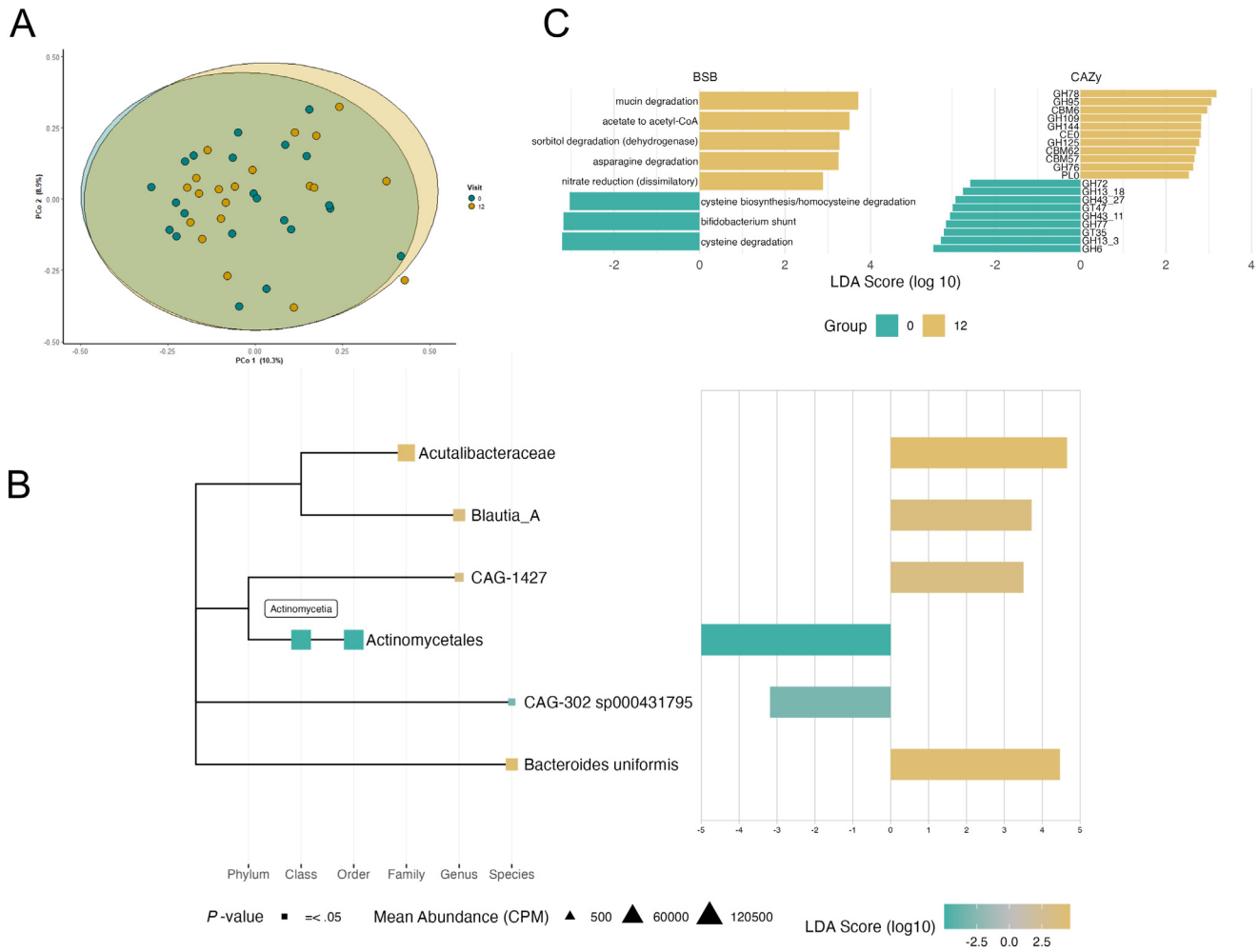


Figure 3. Gut microbiome changes from baseline to follow-up. (*A*) Principal coordinate (PCo) analysis plot of Bray-Curtis distances of species diversity colored by baseline vs follow-up (permutational multivariate analysis of variance, $P > .05$). (*B*) Differentially abundant microbial taxa identified by the LEfSe algorithm ($P = .05$) between baseline and follow-up in CD patients. (*C*) Differentially abundant CAZymes and BSB gene functional pathways ($P = .05$). A teal bar indicates enrichment at baseline and a mustard bar indicates enrichment at follow-up. CPM, counts per million; LDA, linear discriminant analysis.

composition are also associated with shifts in the enterosignature[18] groups between the HV and patient groups (Supplementary Figures 6 and 7). While at baseline there were no significant differences observed in enterosignatures between the CD and HV groups, at follow-up there was significantly higher relative abundance of the ES-Esch (comprising contributions from the genera *Escherichia*, *Citrobacter*, *Enterobacter*, *Klebsiella*, and *Staphylococcus*).[18] There was also a significant reduction in the ES_Bifi enterosignature group relative abundance (Supplementary Figure 7*B*), specifically in the CD group compared with the HV group, reflecting the changes observed in abundance of *Bi longum* and *BI breve* in Figure 4 at a community level.

Bi longum is recognized as a keystone degrader of arabinoxylan[19] and harbors a significant number of genes encoding GH43 enzymes and increases in abundance in response to dietary wheat arabinoxylan supplementation.[20] Supplementary Figure 8 shows the CAZyme profiles of *Bifidobacterium* metagenome-assembled genes. The genome of *Bi longum* contained 14 copies of GH43 genes, more than any other bifidobacterial genome identified in this study. In addition, the *BI breve* genome contained high numbers of gene copies of GH13 and CBM48, related to resistant starch degradation.

A total of 30% of the patients showed persistent or worsening symptoms following GFD (Supplementary Table 2). At follow-up, isobutyric and isovaleric branched-chain fatty acids correlated with GI symptoms ($P < .0001$); comparison of microbial community composition between patients with and without persistent symptoms showed no community-level changes (Supplementary Figure 9*A*). Species-level differences were, however, observed between the groups, particularly in species from the classes Clostridia and Actinomycetia (Supplementary Figures 9*B* and 10).

Discussion

This study provided a detailed picture of shifts in microbiome composition and function occurring in patients with CD at diagnosis and following 1 year of GFD compared with matched HVs. In common with previous

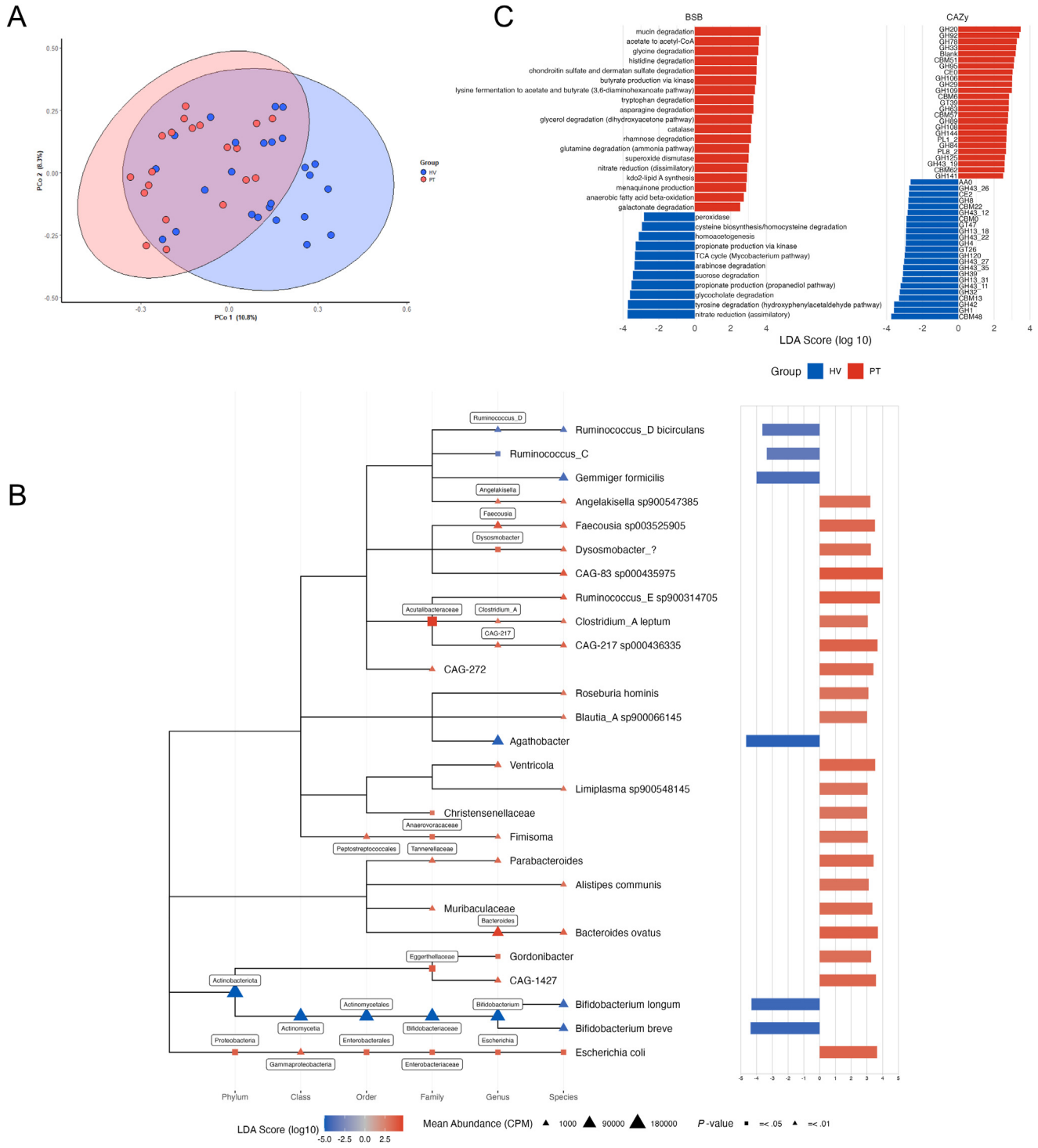


Figure 4. Twelve-month follow-up microbiome composition. (*A*) Principal coordinate (PCo) analysis plot of Bray-Curtis distances of species diversity colored by HV vs patient (PT) (permutational multivariate analysis of variance, $P = .002$, F statistic $= 2.1943$). (*B*) Differentially abundant microbial taxa identified by the LEfSe algorithm ($P = .05$) between HVs and PTs at baseline. (*C*) Results of LEfSe analysis identifying differentially abundant CAZymes and BSB gene functional pathways ($P = .05$). A blue bar indicates enrichment in HV and a red bar indicates enrichment in the PT group. CPM, counts per million; LDA, linear discriminant analysis.

studies we did not identify a CD microbiome signature.[21] We did, however, identify significant differences in individual taxa and metabolic pathways, which may be associated with metabolism of excess malabsorbed protein reaching the gut and inflammation, as has been hypothesized to occur in CD.[21]

One year of GFD resulted in further alterations to the gut microbiome and significant community-level differences between patients and HVs. Several taxa including *BI longum*, *BI breve,* and *R bicirculans* were lower in patients with CD compared with HVs, and conversely *E coli* abundance was higher.[8,22] Studies that investigated

microbiome changes in HVs following GFD have identified similar changes.[12,23] In this study, we further highlight changes in enterosignatures in patents with CD following GFD.[18]

The functional microbiome analysis indicated alterations in carbohydrate metabolism, specifically relating to the 2 major sources of dietary fiber in wheat, arabinoxylan and resistant starch.[24] *Bi longum* strains have been characterized as keystone arabinoxylan degraders[19] and are increased in abundance in response to dietary interventions with arabinoxylan.[20] Therefore, the changes observed in gut microbiome of patients with CD may result from a complex interaction of factors, including dietary as a result of following GFD and changes in gut physiology as a consequence of CD. We focused on stool microbiome, which is more responsive to diet, rather than on duodenal mucosa microbiota, which reflects more the impact of mucosal damage due to immune activation in response to ingested gluten. We did not assess the diet followed by CD group beyond being GFD, nor did we assess the diet of the HVs. Diet can have a strong effect, particularly on microbiota, and future studies could incorporate dietary assessment.

SBWC was increased in patients with CD at baseline compared with HVs, but this showed no correlation with microbiome composition. Increased SBWC in CD may reflect the net effect of impaired absorption associated with villous atrophy and increased secretion associated with crypt hyperplasia, and also impaired motility. In this study no differences in colon volumes were observed.[4] WGTT was significantly delayed in the CD cohort at baseline, and although this improved after GFD it was still slower than in HVs at follow-up. Delayed WGTT in untreated CD may be due to a number of interrelated factors such as mucosal damage and inflammation affecting gut motility, malabsorption of food constituents, and gut hormone derangement. In this study, different microbial species correlated with markers of gut environment.[21,22] Some of the gut microbiome shifts associated with GFD may therefore be related to alterations in transit.[25]

GI symptoms and quality of life reported by the newly diagnosed patients were poorer than in HVs. Although some of these measures improved after GFD, their wellbeing and symptoms were still worse overall than the HVs at follow-up. In this study, there were associations between branched-chain fatty acids and GI symptoms, and also between microbiome composition and persistence of symptoms, as have previously been suggested from studies of mucosal microbiota.[26] Several genera observed to be associated with persistent symptoms, such as *Bifidobacterium*, *Alistipes*, and *Ruminococcus*, have previously been associated with the onset of CD in an infant cohort and therefore may play a role in mediating persistent inflammation in the intestinal tract.[25]

The primary strength of this study is the combined MRI gut function and gut microbiome assessment in newly diagnosed CD before and after 12 months of GFD. Our cohort of 36 patients and 36 HVs is relatively large compared with other studies and provided an extensive dataset.

Limitations included the relatively short time of follow-up observation, as in adults 1 year of GFD does not always produce mucosal integrity and remission of symptoms. Sample size was relatively small for some of the endpoints, and in future larger numbers of patients could be studied. Due to the clinical pathways, we could not assess recovery of intestinal mucosa by histology, only by surrogate biomarkers, such as serology. Serology is an imprecise tool to assess mucosal damage compared with histology. Serology levels that normalize indicate good adherence to the diet but do not reflect occasional gluten intake. Another limitation was the stool microbiome analysis. Although strongly influenced by diet, this may miss changes associated with duodenal inflammation. Collecting duodenal aspirates is however invasive and aspirates have lower microbial load and less diversity compared with the large intestine.

In conclusion, the present study shows that gut function and microbiome are impacted by CD. One year of GFD did not reverse the abnormalities and had a negative impact on the microbiome, with reduction of bifidobacterial abundance, increase in proteolytic species, and reductions in starch-degrading and arabinoxylan-degrading CAZyme families. This potentially opens the possibility of developing interventions to reverse the negative impact of GFD with targeted prebiotic and/or symbiotic products.[27]

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at http://doi.org/10.1016/j.cgh.2024.11.006.

References

1. Rubio Tapia A. ACG clinical guidelines: diagnosis and management of celiac disease. Am J Gastroenterol 2013; 108:656–676.
2. Ciacci C, Ciclitira P, Hadjivassiliou M, et al. The gluten-free diet and its current application in coeliac disease and dermatitis herpetiformis. United European Gastroenterol J 2015;3:121–135.
3. Midhagen G, Hallert C. High rate of gastrointestinal symptoms in celiac patients living on a gluten-free diet: controlled study. Am J Gastroenterol 2003;98:2023–2026.
4. Lam C, Sanders DS, Lanyon P, et al. Increased fasting small-bowel water content in untreated coeliac disease and scleroderma as assessed by magnetic resonance imaging. United European Gastroenterol J 2019;7:1353–1360.
5. Spiller RC, Lee YC, Edge C, et al. Delayed mouth-caecum transit of a lactulose labelled liquid test meal in patients with steatorrhoea caused by partially treated coeliac disease. Gut 1987; 28:1275–1282.
6. Sadik R, Abrahamsson H, Kilander A, et al. Gut transit in celiac disease: delay of small bowel transit and acceleration after dietary treatment. Am J Gastroenterol 2004;99:2429–2436.

7. Pinto-Sanchez MI, Bercik P, Verdu EF. Motility alterations in celiac disease and non-celiac gluten sensitivity. Dig Dis 2015; 33:200–207.
8. Verdu EF, Galipeau HJ, Jabri B. Novel players in coeliac disease pathogenesis: role of the gut microbiota. Nat Rev Gastroentrol Hepatol 2015;12:497–506.
9. 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; 64:406–417.
10. de Meij TGJ, Budding AE, Grasman ME, et al. Composition and diversity of the duodenal mucosa-associated microbiome in children with untreated coeliac disease. Scand J Gastroenterol 2013;48:530–536.
11. Ohlund K, Olsson C, Hernell O, et al. Dietary shortcomings in children on a gluten-free diet. J Hum Nutr Dietet 2010; 23:294–300.
12. De Palma G, Nadal I, Collado MC, et al. Effects of a gluten-free diet on gut microbiota and immune function in healthy adult human subjects. Br J Nutr 2009;102:1154–1160.
13. Tjellstrom B, Stenhammar L, Hogberg L, et al. Gut microflora associated characteristics in children with celiac disease. Am J Gastroenterol 2005;100:2784–2788.
14. Biagi F, Andrealli A, Bianchi PI, et al. A gluten-free diet score to evaluate dietary compliance in patients with coeliac disease. Br J Nutr 2009;102:882–887.
15. Chaddock G, Lam C, Hoad CL, et al. Novel MRI tests of orocecal transit time and whole gut transit time: studies in normal subjects. Neurogastroenterol Motil 2014;26:205–214.
16. Prochazkova N, Falony G, Dragsted LO, et al. Advancing human gut microbiota research by considering gut transit time. Gut 2023;72:180–191.
17. Asnicar F, Leeming ER, Dimidi E, et al. Blue poo: impact of gut transit time on the gut microbiome using a novel marker. Gut 2021;70:1665–1674.
18. Frioux C, Ansorge R, Özkurt E, et al. Enterosignatures define common bacterial guilds in the human gut microbiome. Cell Host Microbe 2023;31:1111–1125.
19. Song A-X, Li L-Q, Yin J-Y, et al. Mechanistic insights into the structure-dependant and strain-specific utilization of wheat arabinoxylan by *Bifidobacterium longum*. Carbohydr Polym 2020;249:116886.
20. Nguyen NK, Deehan EC, Zhang Z, et al. Gut microbiota modulation with long-chain corn bran arabinoxylan in adults with overweight and obesity is linked to an individualized temporal increase in fecal propionate. Microbiome 2020;8:118.
21. Wu X, Qian L, Liu K, et al. Gastrointestinal microbiome and gluten in celiac disease. Ann Med 2021;53:1797–1805.
22. Francavilla A, Ferrero G, Pardini B, et al. Gluten-free diet affects fecal small non-coding RNA profiles and microbiome composition in celiac disease supporting a host-gut microbiota crosstalk. Gut Microbes 2023;15:2172955.
23. Caio G, Lungaro L, Segata N, et al. Effect of gluten-free diet on gut microbiota composition in patients with celiac disease and non-celiac gluten/wheat sensitivity. Nutrients 2020;12:1832.
24. Hazard B, Trafford K, Lovegrove A, et al. Strategies to improve wheat for human health. Nat Food 2020;1:475–480.
25. Leonard MM, Valitutti F, Karathia H, et al. Microbiome signatures of progression toward celiac disease onset in at-risk children in a longitudinal prospective cohort study. Proc Natl Acad Sci U S A 2021;118:e2020322118.
26. Wacklin P, Laurikka P, Lindfors K, et al. Altered duodenal microbiota composition in celiac disease patients suffering from persistent symptoms on a long-term gluten-free diet. Am J Gastroenterol 2014;109:1933–1941.
27. 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;47:139–147.

Correspondence
Address correspondence to: Luca Marciani, PhD, Nottingham Digestive Diseases Centre, Translational Medical Sciences, School of Medicine, University of Nottingham, Nottingham NG7 2UH, United Kingdom. e-mail: luca.marciani@nottingham.ac.uk.

Acknowledgments
The authors sincerely thank all participants who have contributed to this study. They gratefully acknowledge the support of the Biotechnology and Biological Sciences Research Council. The authors also thank J. Price (Nottingham University Hospitals NHS Trust), specialist celiac dietitian, for help with recruitment and clinical advice; M. Lingaya (NIHR Nottingham Biomedical Research Centre) for preparation of stool samples and DNA extraction; D. Baker and R. Evans (Quadram Institute Bioscience) for DNA normalization and sequencing library preparation; A. Wragg (NIHR Nottingham Biomedical Research Centre) for his assistance with the patient and public involvement group; and C. Lam and S. Beg for their expert gastrenterology advice. This article was posted as a preprint on bioRxiv.org: https://doi.org/10.1101/2024.06.20.599876.

CRediT Authorship Contributions
Carolyn M. Costigan (Conceptualization: Equal; Data curation: Lead; Formal analysis: Lead; Funding acquisition: Lead; Investigation: Lead; Project administration: Lead; Writing – original draft: Lead; Writing – review & editing: Equal)
Frederick J. Warren (Formal analysis: Equal; Methodology: Equal; Software: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)
Anthony P. Duncan (Formal analysis: Supporting; Methodology: Supporting)
Caroline L. Hoad (Formal analysis: Supporting; Methodology: Supporting; Software: Supporting; Writing – review & editing: Supporting)
Nina Lewis (Conceptualization: Supporting; Methodology: Supporting; Supervision: Supporting; Writing – review & editing: Supporting)
Trevor Hill (Formal analysis: Supporting)
Colin J. Crooks (Formal analysis: Supporting; Writing – review & editing: Supporting)
Paul S. Morgan (Conceptualization: Supporting; Methodology: Supporting; Supervision: Equal; Writing – review & editing: Supporting)
Carolina Ciacci (Conceptualization: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)
Paola Iovino (Conceptualization: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)
David S. Sanders (Writing – review & editing: Supporting)
Falk Hildebrand (Formal analysis: Supporting; Methodology: Supporting)
Penny A. Gowland (Conceptualization: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)
Robin C. Spiller (Conceptualization: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)
Luca Marciani, PhD (Conceptualization: Supporting; Methodology: Equal; Supervision: Lead; Writing – review & editing: Supporting)

Conflicts of Interest
The authors disclose no conflicts.

Funding
Carolyn M. Costigan was supported by a College of Radiographers Doctoral Fellowship grant. The magnetic resonance imaging scanning and other research costs were met by internal, academic research funds of the Nottingham Digestive Diseases Centre, University of Nottingham. This research was funded by the Biotechnology and Biological Sciences Research Council Institute Strategic Programme Food Microbiome and Health BB/X011054/1 and its constituent project BBS/E/F/000PR13631. This research was supported by the National Institute for Health and Care Research Nottingham Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health and Care Research, or the Department of Health and Social Care.

Supplementary Material

Patient Recruitment, Serology, Biopsies, and Stool Sample Collection

Patient participants were recruited through specialist clinics at Nottingham University Hospitals NHS Trust. The initial approach was from a member of the patient's usual care team (gastroenterologists and specialist dieticians), and information about the trial was also on display in the relevant clinical areas. All patients who expressed an interest in taking part were sent a detailed information sheet, which also contained the contact information of the study team and this gave them an opportunity to raise any questions regarding the trial before and during the first screening visit. Patients were also recruited via the specialist dietician celiac group sessions for newly diagnosed patients.

General advertisement for participants in the healthy volunteer group took place in parallel, on the campuses of the university, by poster. Healthy control subjects were also recruited using the research group's social media accounts.

All the people with celiac disease (CD) included in the study had a positive serology result from 1 of IgA-TG2 or IgA-DGP, or IgG-DGP (the particular type of serology test depended on the specific referring pathway) and duodenal biopsy confirming CD. At baseline, 21 patients had a gluten-free diet (GFD) compliance Biagi score of 0, and 15 of them had a score of 1, indicating that at baseline none of them was following a GFD.[e1]

We were unable to retrieve details of 1 biopsy result for 1 of the patients, which was reported as positive biopsy for celiac disease without detail. The other patients presented with the following Marsh gradings: n = 7 Marsh grading 1, n = 18 Marsh grading 3a, n = 9 Marsh grading 3b, n = 1 Marsh grading 3c.

Newly diagnosed celiac adult patients were referred to the specialist celiac dietician service at Nottingham University Hospitals NHS Trust and attended a group session for specialist dietary and lifestyle advice. Once following their GFD treatment regime, they had a follow-up visit with the specialist dietician, as part of their standard clinical care.

There was no formal assessment (eg, food frequency questionnaires, 24-hour dietary recall) of type of diet that the participants, both the people with CD and the healthy volunteers, followed during the 12 months in the study.

The 12 months' follow-up CD group included patients providing a stool sample and having GFD diet adherence confirmed by negative tissue transglutaminase (TTG) blood sample and Biagi score $\geq$3. TTG levels tend to fall to baseline in patients that are adherent to GFD within the first 12 months.[e2] However, this may not detect low levels of dietary transgression. Occasional gluten contamination in this period cannot, however, be ruled out, and it might have had an impact on the results. In individuals with persistent nonadherence, the TTG would be invariably still elevated at 12 months. Were an individual to switch from adherent to nonadherent, the immunological TTG response would lag by approximately 4–6 weeks.

On each study day, the participants were asked to provide a stool sample using a supplied collection kit, which included aliquot laboratory tubes and a biosafety container with freezer packs, to be kept in the home freezer (–5 °C) until needed for sample collection. The container was then brought to the study unit and the samples were then stored immediately at –80 °C until they underwent DNA extraction, sequencing to a depth of ~10 GB per sample and metagenomic processing.

Fasting breath hydrogen (H_2) was measured using a portable breath H_2 meter (Gastro+Gastrolyzer; Bedfont Scientific).

Magnetic Resonance Imaging

All participants were scanned supine on a 1.5T GE HDxt magnetic resonance imaging (MRI) scanner. The total scan duration was 15 minutes. After localizing scans, 3 different imaging sequences were acquired using short breath-holds. A coronal T2 single-shot fast spin echo sequence was used to measure the volumes of freely mobile small bowel water content.[e3] This gave high-intensity signal from areas with freely mobile fluid and little signal from body tissues. A coronal dual echo fast field echo sequence was then used to visualize the abdominal anatomy and measure colonic volumes.[e4] Last, a coronal LAVA 3-dimensional fat saturated sequence was used to identify the transit markers.[e5] To quantify whole gut transit time (WGTT), each transit capsule marker was allocated a score based on its position in the colon, and WGTT was assessed using a weighted average position score (WAPS) (in arbitrary units) of the 5 capsule markers.[e5] The WAPS can be extrapolated with some assumption to WGTT time in hours.[e5] One of the patients' MRI scans out of 20 and 3 of the control subjects' MRI scans had poor quality due to respiratory motion and were excluded from the small bowel water content dataset.

Correlations Between WGTT and Stool Form and Frequency

When all the patients' WGTT, stool form, and stool frequency data were pooled together, we observed, post hoc, a modest though significant negative correlation between stool form and WGTT (n = 59, R^2 = 0.12, P = .0079) and stool frequency (n = 59, R^2 = .15, P = .0023), whereby harder stools and less frequent stools were associated with longer WGTT. Others have reported that stool consistency correlates modestly with transit but has no relation with stool frequency.[e6] Possible differences can be explained because defecation is influenced

by many other factors including rectal sensitivity, stool consistency, and social influences, like availability of toilets. In our group of people with celiac disease, stool frequency was not a significant concern with the baseline prior to GFD being 1.8 per day.

Short-Chain Fatty Acids

The methods for fecal short-chain fatty acid (SCFA) analysis were described previously.[e7] Briefly, SCFAs were quantified by gas chromatography–mass spectrometry. Separation and detection of SCFAs of interest was achieved with splitless injection of the ethyl acetate extract using a Trace GC Ultra (Thermo Fisher Scientific) coupled with a DSQII mass spectrometer (Thermo Fisher Scientific). Compound identification was achieved by matching with database mass spectra (NIST/EPA/NIH Mass Spectral Library, Version 2.0d). Concentrations of analyte were calculated using Xcalibur software (Thermo Fisher Scientific).

DNA Extraction and Metagenomic Sequencing

DNA was extracted using an adaptation of the Qiagen QIAmp DNA mini kit. First, 0.125 g of defrosted stool sample was weighed into a 2 mL screw-cap microcentrifuge tube and was suspended in 0.5 mL of lysis buffer (500 mM NaCl, 50 mM Tris-HCl [pH 8], 50 mM EDTA, 4 % sodium dodecyl sulfate). The tubes were preloaded with 0.25 g of 0.1-mm zirconia beads and three 3-mm glass beads. The tubes were homogenized in a FastPrep (MP Biomedicinals) for 1 minute at 5.5 m/s then placed on ice for 30 seconds. This cycle was repeated 2 further times. The homogenized samples were heated to 95 °C for 15 minutes, mixing by hand every 5 minutes. They were then cooled and centrifuged at 4 °C for 5 minutes (13,000 g). The supernatant was transferred to a fresh 2 mL microcentrifuge tube. The pellet was resuspended in 150 μL of lysis buffer and the heating and centrifugation steps repeated, and the supernatants pooled. To each lysate tube, 130 μL of 10 M ammonium acetate was added, well mixed, and incubated on ice for 5 minutes, before centrifuging (4 °C, 13,000 g) for 10 minutes. The supernatant was then transferred to a fresh 2 mL microcentrifuge tube and an equal volume of isopropanol was added. This was mixed well and incubated on ice for 30 minutes. The tube was then centrifuged (4 °C, 13,000 g) for 15 minutes, the supernatant discarded, and the pellet washed with 0.5 mL of ethanol and then allowed to dry. The resulting pellet was then redissolved in 200 μL TE buffer, and 2 μL of DNase-free RNase (10 mg/mL) was added. The tube was then incubated at 37 °C for 15 minutes, before adding 15 μL of proteinase K, 200 μL of buffer AL, and heating at 70 °C for a further 10 minutes. After cooling to room temperature 200 μL of ethanol was added to the tube, and the contents were transferred to a QIAmp column. The QIAmp column was centrifuged at 6000 g for 1 minute, the flow through discarded and placed in a fresh collection tube. 500 μL of AW1 buffer was added to the column and then centrifuged at 6000 g for 1 minute, the flow through discarded and placed in a fresh collection tube. The column was then transferred to a fresh collection tube and 500 μL of AW2 buffer added. The column was then centrifuged at 20,000 g for 3 minutes, the flow through discarded, and the column then centrifuged for a further 1 minute in a fresh collection tube to dry. The column was then placed in a fresh 1.5 mL microcentrifuge tube and 200 μL AE buffer was added. This was then centrifuged at 6000 g for 1 minute to elute genomic DNA.

Genomic DNA was normalized to 5 ng/μL with elution buffer (10 mM Tris HCl). A miniaturized reaction was set up using the Nextera DNA Flex Library Prep Kit (Illumina). 0.5 μL Tagmentation Buffer 1 was mixed with 0.5 μL Bead-Linked Transposomes and 4.0 μL polymerase chain reaction (PCR)-grade water in a master mix, and 5 μL was added to each well of a chilled 96-well plate. About 2 μL of normalized DNA (10 ng total) was pipette mixed with each well of Tagmentation master mix, and the plate was heated to 55°C for 15 minutes in a PCR block. A PCR master mix was made up using 4 μL kapa2G buffer, 0.4 μL dNTPs, 0.08 μL polymerase, and 4.52 μL PCR-grade water, from the Kap2G Robust PCR kit (Sigma-Aldrich), and 9 μL was added to each well in a 96-well plate. About 2 μL each of P7 and P5 of Nextera XT Index Kit v2 index primers (catalog no. FC-131-2001 to 2004; Illumina) were also added to each well. Finally, the 7 μL of Tagmentation mix was added and mixed. The PCR was run at 72 °C for 3 minutes, 95 °C for 1 minute, 14 cycles of 95 °C for 10 seconds, 55 °C for 20 seconds, and 72 °C for 3 minutes. Following the PCR reaction, the libraries from each sample were quantified using the methods described earlier and the high sensitivity Quant-iT dsDNA Assay Kit. Libraries were pooled following quantification in equal quantities. The final pool was double-SPRI size selected between 0.5 and 0.7X bead volumes using KAPA Pure Beads (Roche). The final pool was quantified on a Qubit 3.0 instrument and run on a D5000 ScreenTape (Agilent) using the Agilent Tapestation 4200 to calculate the final library pool molarity. Quantitative PCR was done on an Applied Biosystems StepOne Plus machine. Samples quantified were diluted 1 in 10,000. A PCR master mix was prepared using 10 μL KAPA SYBR FAST qPCR Master Mix (2X) (Sigma-Aldrich), 0.4 μL ROX High, 0.4 μL 10 μM forward primer, 0.4 μL 10 μM reverse primer, 4 μL template DNA, 4.8 μL PCR-grade water. The PCR programme was 95 °C for 3 minutes, 40 cycles of 95 °C for 10 seconds, and 60 °C for 30 seconds. Standards were made from a 10 nM stock of Phix, diluted in PCR-grade water. The standard range was 20, 2, 0.2, 0.02, 0.002, and 0.0002 pmol. The pooled library was then sent to Novogene for sequencing using an Illumina NovaSeq instrument, with sample names and index combinations used. Demultiplexed FASTQs were

returned on a hard drive. A sequencing depth of ~10 GB per sample was achieved.

Metagenomic Processing and Bioinformatic Analysis

All metagenomic processing from raw reads to MGS abundance was conducted using the MATAFILER pipeline.[e8,e9] Briefly, raw shotgun metagenomes were quality filtered using sdm v1.63 with default parameters[e10] and were assembled using megahit v 1.2.9 with parameters "–k-list 25,43,67,87,101,127,"[e11] and reads were mapped onto assemblies using bowtie2 v2.3.4.1 with parameters "–end-to-end,"[e12] genes predicted with prodigal v2.6.1 with parameters "-p meta,"[e13] and a gene catalogue clustered at 95% nt identity using mmseqs2.[e14] MAGs (metagenomic assembled genomes) were binned using SemiBin2[e15] and combined in MATAFILER to MGS (metagenomic species), relying on canopy clustering.[e16] Matrix operations were carried out using rtk.[e17]

To functionally annotate genes in the gene catalogue, dbCAN3[e18] was used to annotate genes to the CAZyme database and DIAMOND[e19] to annotate Kyoto Encyclopedia of Genes and Genomes orthologues. Based on the KEGG Ortholog abundance matrices, we further calculated KEGG and Gaussian mixture model[e20] module abundances using a custom C++ implementation available on www.github.com/hildebra/Rarefaction/, similar to the protocol described in Forslund et al.[e21]

The data were total sum scaling (TSS) normalized and log transformed prior to analysis. Corrected P values were estimated using the Benjamini-Hochberg correction method. A false discovery rate–corrected P value (Q value) of .1 was selected as a cutoff for significance. Due to the longitudinal nature of the dataset, the stool donor was included in the model as a random effect. Abundance of the 5 enterosignatures[e8] were calculated using the Web server https://enterosignatures.quadram.ac.uk/. Permutational multivariate analysis of variance analysis was carried out between human volunteer and CD microbiomes at baseline and follow-up based on Bray-Curtis dissimilarity using the adonis2 function in the VEGAN library R package (v2.6.4).[e22]

LEfSe (Linear discriminant analysis Effect Size)[e23] was used to identify individual taxa and functionally annotated genes and metabolic pathways. A P-value of .05 was used as a cutoff for statistical significance, and an linear discriminant analysis value of 2.5 was used to identify discriminating features. Linear mixed-effects modeling using the MaAsLin2 package[e24] was used to identify correlations between individual taxa and functionally annotated genes and metabolic pathways. When exploring associations with MRI variables, multivariate linear regression analyses were adjusted for multiple comparisons using the Benjamini-Hochberg correction, with a significance threshold set to $P = .05$, $Q = .1$ using the MaAsLin2 software package.[e24]

Patient Groups With Persistent and Resolved Symptoms After GFD

The Patient Health Questionnaire-15 included 3 questions about GI symptoms over the previous 4 weeks, measured before and after GFD. We used this to divide the follow-up patients into 2 groups. The integral scoring of the 3 GI symptoms questions can range from 0 to 6. In keeping with the standard scoring of the Patient Health Questionnaire-15, we defined total GI symptoms scores 0–2 as low (no to mild symptoms) and scores of 3–6 as high (moderate to severe symptoms).

The first group included the patients for whom the baseline symptoms were low before and after GFD (range 0–2) or were high at baseline (range 3–6) but improved after GFD (reduced to range 0–2).

The second group included the patients for whom the baseline symptoms were high before GFD (range 3–6) and did not improve after GFD (remained in the range 3–6).

Differentially abundant taxa between patients with persistent and resolved symptoms were identified using the radEmu package.[e25]

Thirty percent of the patients showed persistent or worsening symptoms following GFD treatment, which is around the commonly quoted figure of 40%. Our study was not powered for looking at these subgroups which are small, but from the data (Supplemental Table 2) a picture emerges of the group of patients with persisting symptoms having higher anxiety and depression, higher small bowel water content, larger colon volume, and faster transit (the latter possibly related to anxiety). We also found a striking correlation of stool isobutyric acid with symptoms ($R^2 = 0.58$; $P < .0001$) and stool isovaleric acid with symptoms ($R^2 = 0.61$; $P < .0001$).

These branched-chain fatty acids (BCFAs) are consequences of protein fermentation in the colon (the result from metabolism of branched-chain amino acids) and are generally viewed negatively, as a marker of protein fermentation, and are associated with gut inflammation.[e26,e27] Elevated BCFAs could be associated with worse gut inflammation symptoms. Literature on BCFAs in CD[e28,e29] shows lower BCFAs in CD compared with healthy control subjects, although this may just be reflective of a lower overall SCFA output in people with CD compared with healthy control subjects.

Patient and Public Involvement

With the support an experienced Patient and Public Involvement and Engagement facilitator at the NIHR Nottingham Biomedical Research Centre and with the advice of charity Coeliac UK, a celiac patient focus group was set up to inform the initial trial protocol. All aspects of the trial were discussed with members of this group. Their opinion on use of language and on the patient journey during the trial informed all patient-facing

documentation and many practical aspects of the study (eg, the stool collection).

Supplementary References

e1. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol 1997;32:920–924.

e2. Elli L, Leffler D, Cellier C, et al. Guidelines for best practices in monitoring established coeliac disease in adult patients. Nat Rev Gastro Hepat 2024;21:198–215.

e3. Hoad CL, Marciani L, Foley S, et al. Non-invasive quantification of small bowel water content by MRI: a validation study. Phys Med Biol 2007;52:6909–6922.

e4. Pritchard SE, Marciani L, Garsed KC, et al. Fasting and postprandial volumes of the undisturbed colon: normal values and changes in diarrhea-predominant irritable bowel syndrome measured using serial MRI. Neurogastroenterol Motil 2014;26:124–130.

e5. Chaddock G, Lam C, Hoad CL, et al. Novel MRI tests of orocecal transit time and whole gut transit time: studies in normal subjects. Neurogastroenterol Motil 2014;26:205–214.

e6. Saad RJ, Rao SSC, Koch KL, et al. Do stool form and frequency correlate with whole-gut and colonic transit? Results from a multicenter study in constipated individuals and healthy controls. Am J Gastroenterol 2010;105:403–411.

e7. Sloan TJ, Jalanka J, Major GAD, et al. A low FODMAP diet is associated with changes in the microbiota and reduction in breath hydrogen but not colonic volume in healthy subjects. PLoS One 2018;13:e0201410.

e8. Frioux C, Ansorge R, Özkurt E, et al. Enterosignatures define common bacterial guilds in the human gut microbiome. Cell Host Microbe 2023;31:1111–1125.

e9. Hildebrand F, Gossmann TI, Frioux C, et al. Dispersal strategies shape persistence and evolution of human gut bacteria. Cell Host Microbe 2021;29:1167–1176.

e10. Özkurt E, Fritscher J, Soranzo N, et al. LotuS2: an ultrafast and highly accurate tool for amplicon sequencing analysis. Microbiome 2022;10:176.

e11. Li D, Liu C-M, Luo R, et al. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 2015;31:1674–1676.

e12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–359.

e13. Hyatt D, Chen G-L, LoCascio PF, et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 2010;11:119.

e14. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol 2017;35:1026–1028.

e15. Pan S, Zhao X-M, Coelho LP. SemiBin2: self-supervised contrastive learning leads to better MAGs for short-and long-read sequencing. Bioinformatics 2023;39:i21–i29.

e16. Nielsen HB, Almeida M, Juncker AS, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat Biotechnol 2014;32:822–828.

e17. Saary P, Forslund K, Bork P, et al. RTK: efficient rarefaction analysis of large datasets. Bioinformatics 2017;33:2594–2595.

e18. Zheng J, Ge Q, Yan Y, et al. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. Nucleic Acids Res 2023;51:W115–W121.

e19. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods 2015;12:59–60.

e20. Darzi Y, Falony G, Vieira-Silva S, et al. Towards biome-specific analysis of meta-omics data. ISME J 2016;10:1025–1028.

e21. Forslund K, Hildebrand F, Nielsen T, et al. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. Nat 2015;528:262–266.

e22. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci 2003;14:927–930.

e23. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. Genome Biol 2011;12:R60.

e24. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol 2021;17:e1009442.

e25. Clausen DS, Willis AD. Estimating fold changes from partially observed outcomes with applications in microbial metagenomics. arXiv 2024;2402:05231.

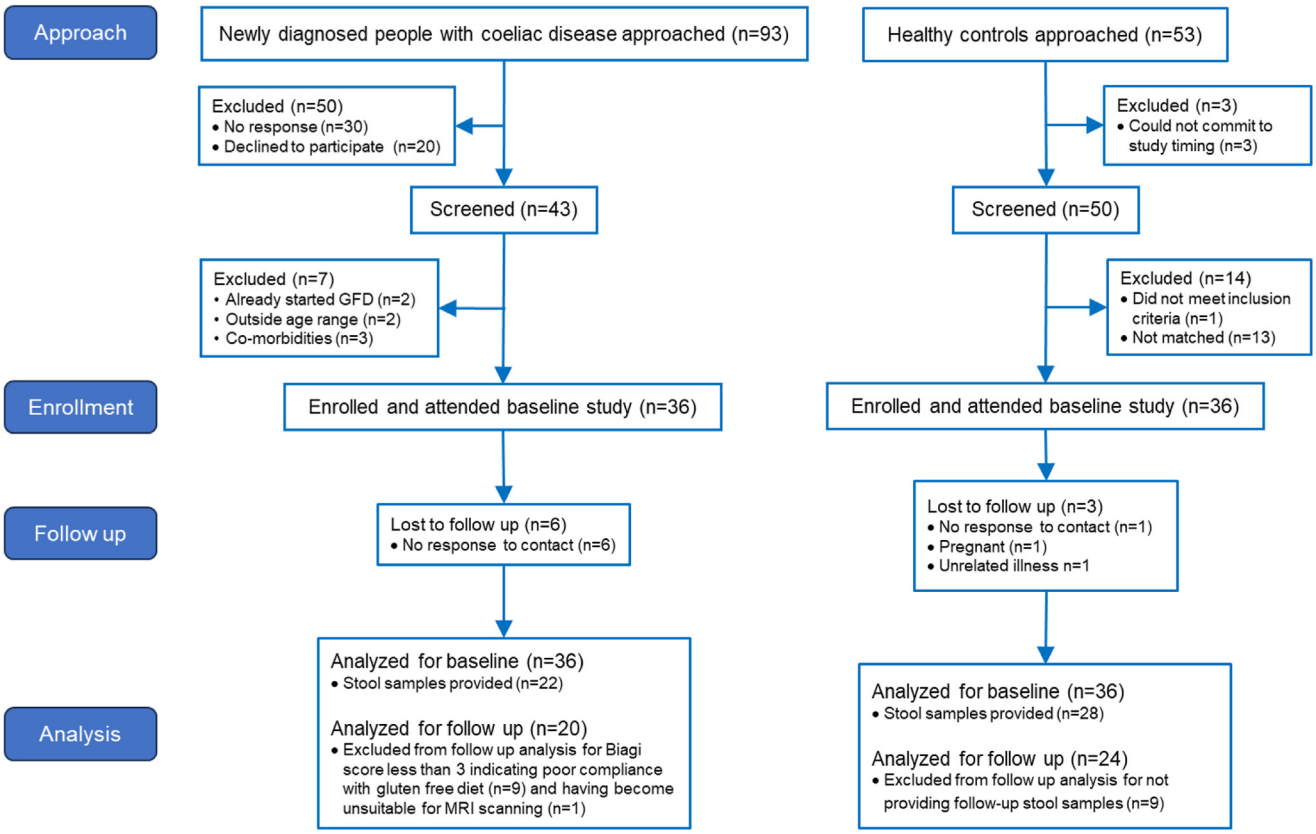
e26. Trefflich I, Dietrich S, Braune A, et al. Short- and branched-chain fatty acids as fecal markers for microbiota activity in vegans and omnivores. Nutrients 2021;13:1808.

e27. Choi BSY, Daniel N, Houde VP, et al. Feeding diversified protein sources exacerbates hepatic insulin resistance via increased gut microbial branched-chain fatty acids and mTORC1 signaling in obese mice. Nat Commun 2021;12:3377.

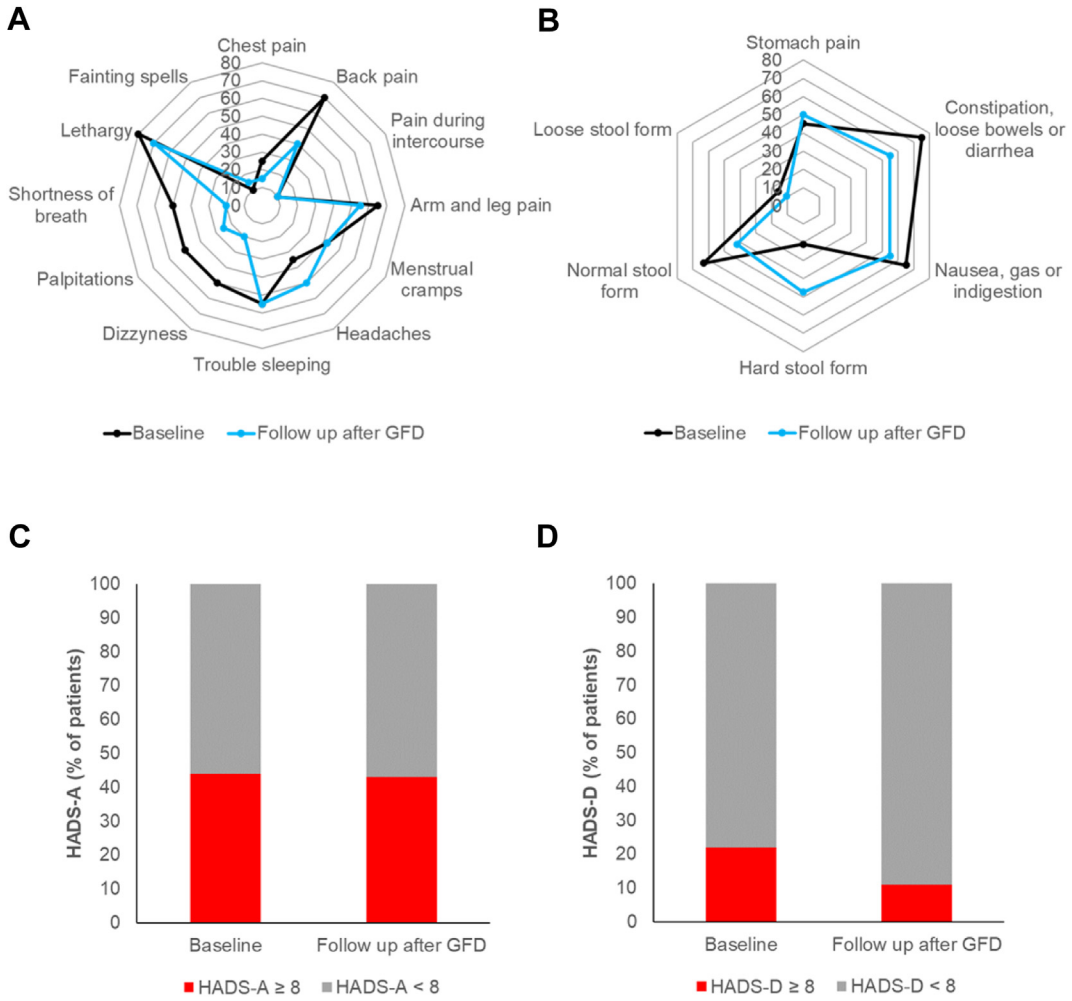
e28. Federica R, Edda R, Daniela R, et al. Characterization of the "gut microbiota-immunity axis" and microbial lipid metabolites in atrophic and potential celiac disease. Front Microbiol 2022;13:886008.

e29. Makinder M, Kassara M, Karanikolou A, et al. The metabolic activity of the gut microbiota and the impact of gluten free diet in children with coeliac disease. Proc Nutr Soc 2014;73:E5.

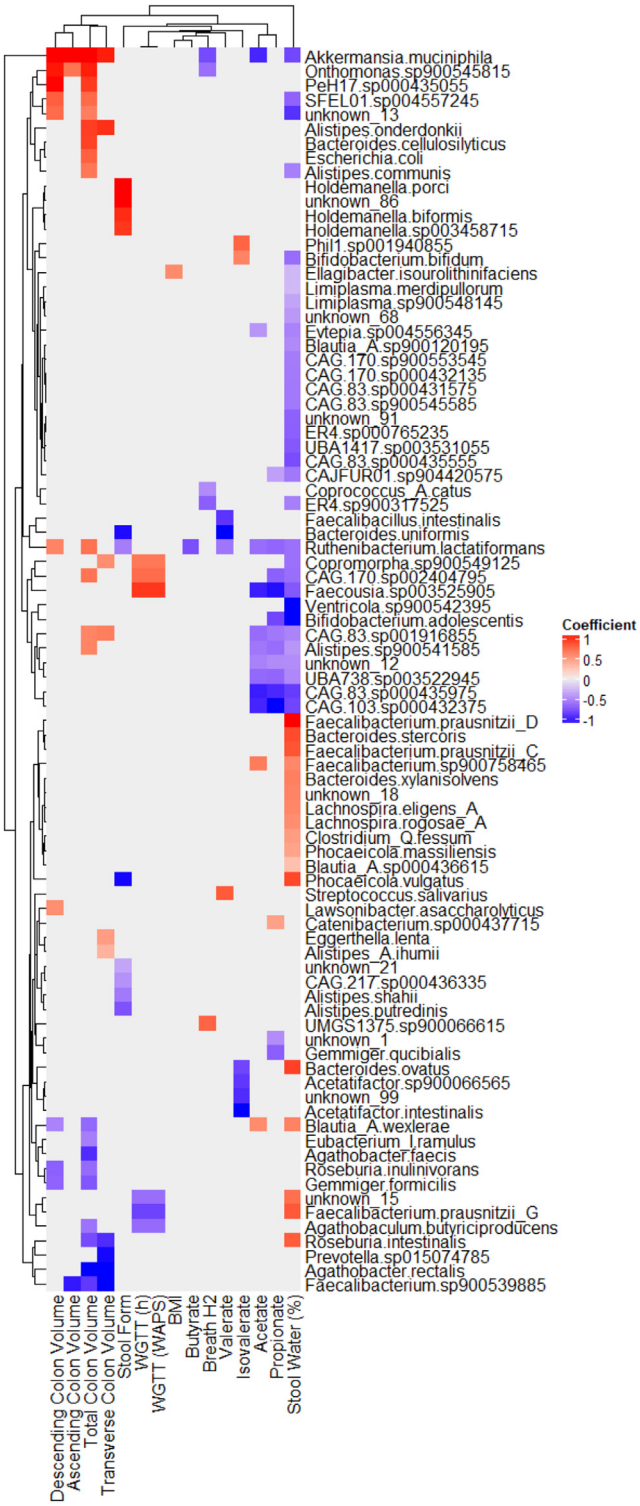
e30. Bjelland I, Dahl AA, Haug TT, et al. The validity of the Hospital Anxiety and Depression Scale - an updated literature review. J Psychosom Res 2002;52:69–77.



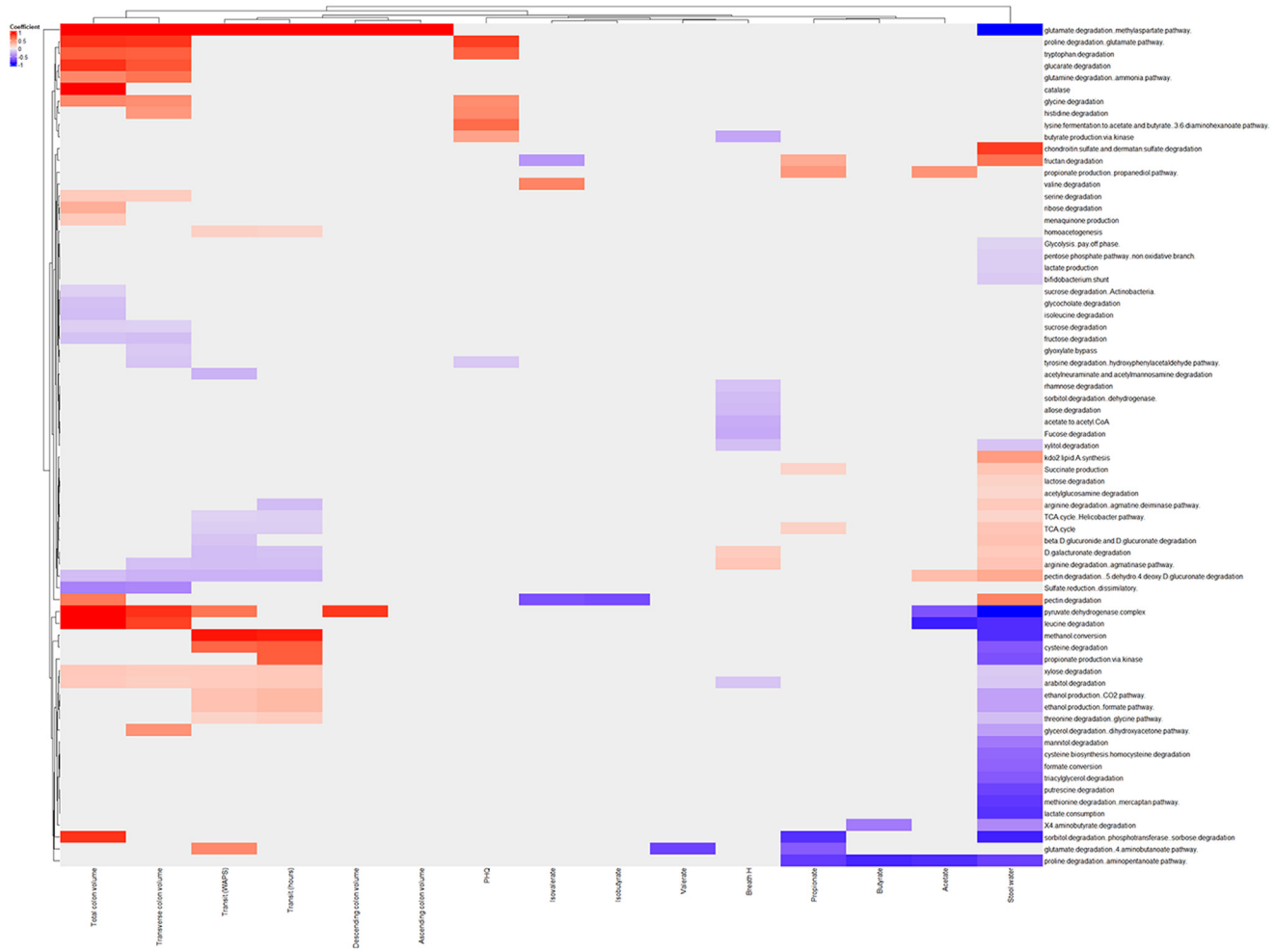
Supplementary Figure 1. CONSORT diagram for the study.



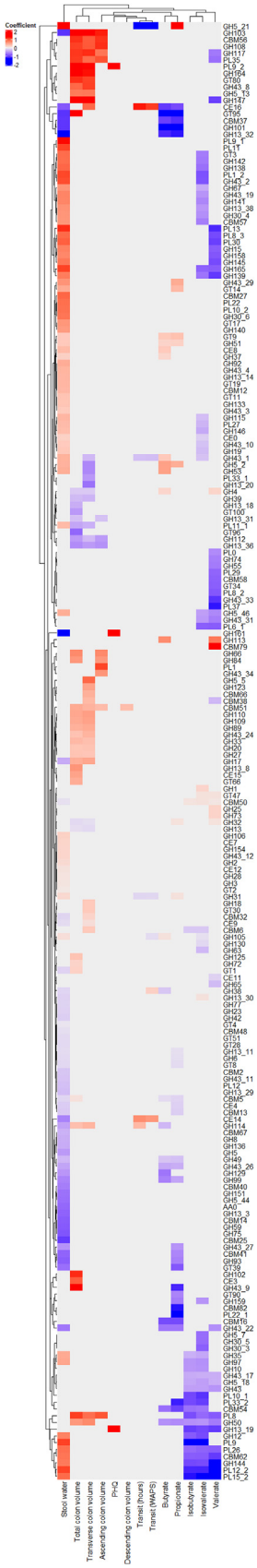
Supplementary Figure 2. Symptoms in patients with celiac disease and changes following 12 months of gluten-free diet. N = 20. People newly diagnosed with CD were studied before starting GFD treatment and 12 months after. At these 2 time points they filled in the Patient Health Questionnaire-15 (PHQ-15), the Hospital Anxiety and Depression Scale–Anxiety (HADS-A) and Hospital Anxiety and Depression Scale–Depression (HADS-D), and the Bristol Stool Form Scale. (*A*) Radar diagram of the 12 nongastrointestinal somatization symptoms of the PHQ-15, before and after following the GFD. Each data point shows the percentage of patients 'experiencing the symptom, corresponding to a score of 1 or 2 on the PHQ-15 for that symptom. (*B*) Radar diagram of the 3 gastrointestinal somatization symptoms of the PHQ-15, before and after following the GFD. The diagram also shows form of the patients' stools from the Bristol Stool Form Scale, whereby a score of 1 or 2 indicates hard stool form, a score of 3–5 indicates normal stool form, and a score of 6 or 7 indicates loose stool form. Each data point shows the percentage of patients experiencing either the symptom, corresponding to a score of 1 or 2 on the PHQ-15 for that symptom or their stool form. (*C*) Bar chart of the anxiety subscale of the Hospital Anxiety and Depression Scale. A score of 8 or greater is often used to indicate clinical levels of anxiety.[e30] The chart shows the percentage of patients with clinical levels of anxiety before and after 1 year of GFD. (*D*) Bar chart of the depression subscale of the Hospital Anxiety and Depression Scale. A score of 8 or greater is often used to indicate clinical levels of depression.[30] The chart shows the percentage of patients with clinical levels of depression before and after 1 year of GFD.



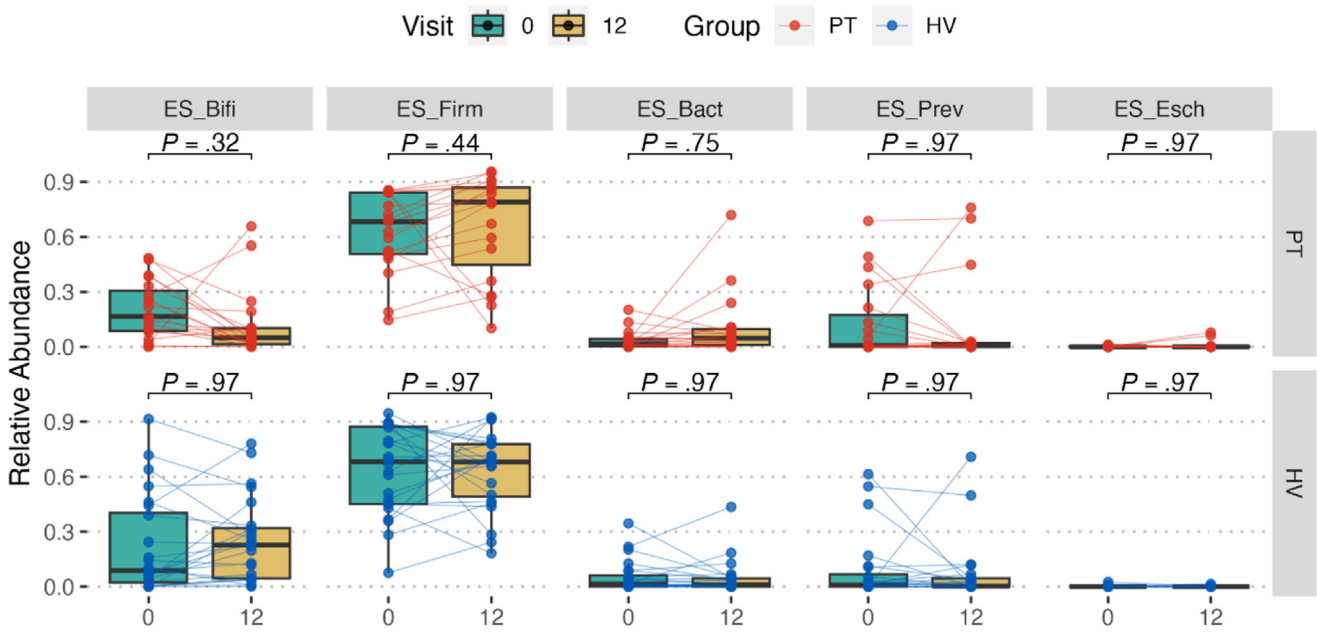
Supplementary Figure 3. Associations between MRI parameters, stool metabolites, and water content, breath hydrogen, body mass index, symptom scores (Hospital Anxiety and Depression Scale and Patient Health Questionnaire), and species identified by metagenomic sequencing. Associations were identified using the generalized linear modeling method in MaAsLin2 with TSS normalization and log transformation of data. Only significant associations (Q value $\leq$ .1) are shown, with the color indicating a positive association (red) or a negative association (blue).



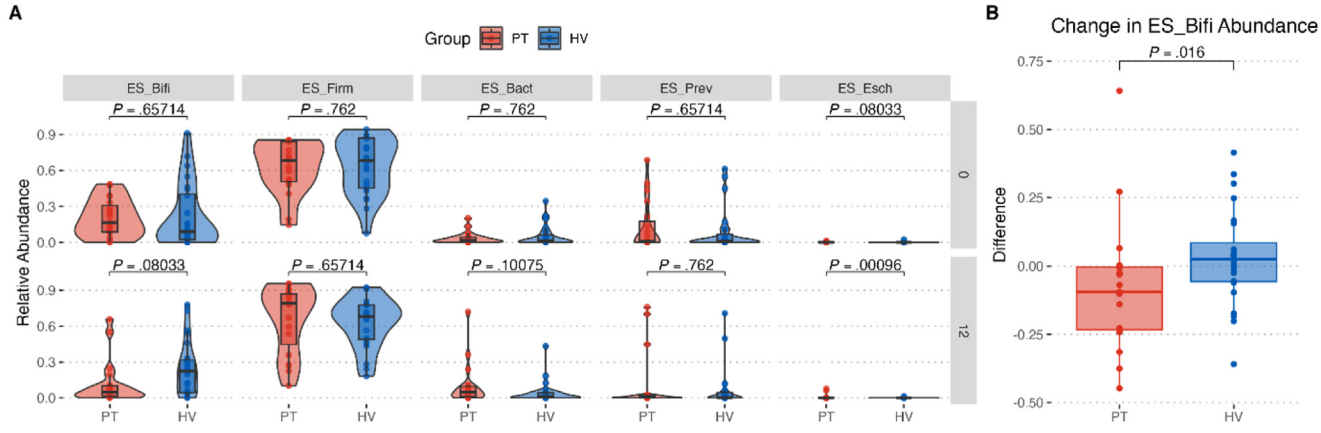
Supplementary Figure 4. Associations between MRI parameters, stool metabolites and water content, breath hydrogen, body mass index, symptom scores (Hospital Anxiety and Depression Scale and Patient Health Questionnaire), and microbial metabolic pathways. Associations were identified using the generalized linear modeling method in MaAsLin2 with TSS normalization and log transformation of data. Only significant associations (Q value $\leq .1$) are shown, with the color indicating a positive association (red) or a negative association (blue).



Supplementary Figure 5. Associations between MRI parameters, stool metabolites and water content, breath hydrogen, body mass index, symptom scores (Hospital Anxiety and Depression Scale and Patient Health Questionnaire), and CAZyme family abundances. Associations were identified using the generalized linear modeling method in MaAsLin2 with TSS normalization and log transformation of data. Only significant associations (Q value $\leq$ .1) are shown, with the color indicating a positive association (red) or a negative association (blue).

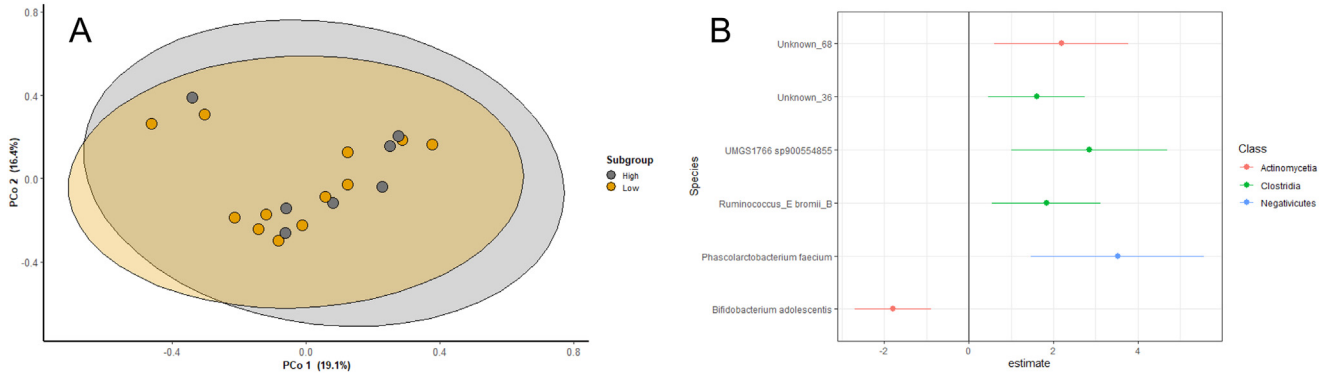


Supplementary Figure 6. Shifts in enterosignatures (ESs) from baseline to follow-up in the healthy volunteer (HV) and patient with CD (PT) groups. The statistical significance of associations between ES groups between baseline and follow-up for each group was calculated using a paired Wilcoxon test with Benjamini-Hochberg correction for multiple comparisons.

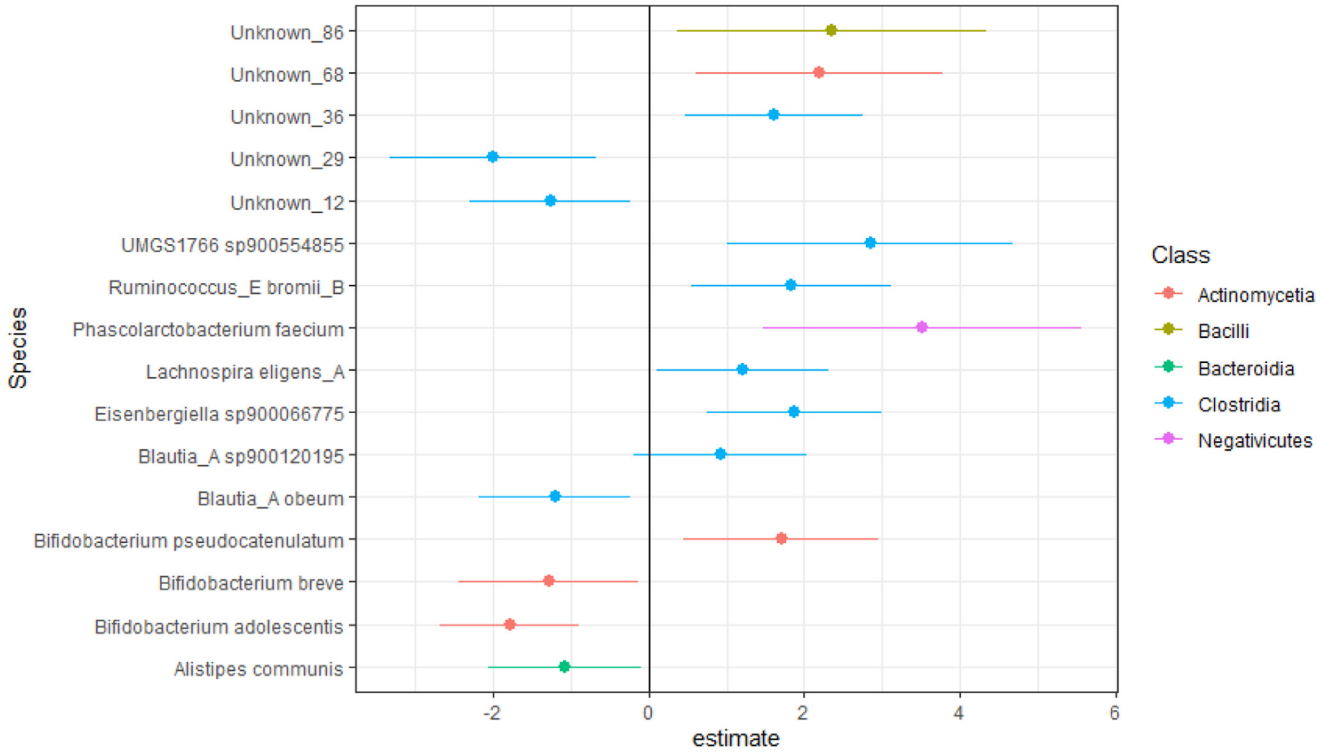


Supplementary Figure 7. Differences between enterosignatures (ESs) between healthy volunteers (HV) (n = 24) and patients with CD (PT) (n = 20) at baseline and following 12 months of follow-up. (*A*) Shifts in relative abundance of ESs. Statistical significance is indicated (*Q* values) on the basis of unpaired Wilcoxon tests with Benjamini-Hochberg correction for multiple comparisons. (*B*) Differential abundance of the ES_Bifi ES from baseline to follow-up. Statistical significance is indicated on the basis of an unpaired Wilcoxon test.

Supplementary Figure 8. CAZyme gene counts in representative metagenome assembled genomes of *Bifidobacterium* species identified in the study. CAZyme counts were determined from the DBcan algorithm based on the CAZy database. Hits were only considered significant if identified by all 3 tools in the algorithm (DIAMOND, HMMER, and dbCAN-sub). CBM, carbohydrate binding module; CE, carbohydrate esterase; GH, glycosyl hydrolase; GT, gycosyl transferase; PL, polysaccharide lyase.



Supplementary Figure 9. Microbiome compositional differences between PT with and without persistent symptoms. (*A*) PCoA plot comparing differences (Bray-Curtis dissimilarity) between gut microbiome community composition at follow-up between PT with and without persistent symptoms (PERMANOVA $P > .05$). (*B*) Differentially abundant species between PT with and without persistent symptoms at baseline ($P \leq .05$).



Supplementary Figure 10. Microbiome compositional differences between patients with CD with and without persistent symptoms. Differentially abundant species between patients with CD with and without persistent symptoms at follow-up ($P \leq .1$).

Supplementary Table 1. Serology and Biopsy Data for the People With Celiac Disease Included in the Study

Anonymized Participant ID	Sex	Age (y)	Baseline Serology	Marsh Grading	Follow-Up Biagi Score	Included in Follow-Up Analysis
MARCO_011	Female	32	Positive	3a	3	Yes
MARCO_012	Female	45	Positive	3b	3	Yes
MARCO_013	Female	20	Positive	3a	3	Yes
MARCO_014	Female	39	Positive	1	3	Yes
MARCO_015	Female	28	Positive	1	1	No
MARCO_016	Female	24	Positive	3a	N/A	No
MARCO_017	Male	67	Positive	3a	2	No
MARCO_018	Male	44	Positive	3a	3	No
MARCO_019	Female	55	Positive	1	N/A	No
MARCO_020	Female	61	Positive	1	0	No
MARCO_039	Male	59	Positive	3a	3	Yes
MARCO_040	Female	21	Positive	3b	N/A	No
MARCO_041	Female	55	Positive	3a	2	No
MARCO_042	Male	60	Positive	3a	3	Yes
MARCO_043	Male	62	Positive	1	3	Yes
MARCO_044	Male	66	Positive	3a	3	Yes
MARCO_045	Male	70	Positive	3a	3	Yes
MARCO_046	Female	38	Positive	N/A	3	Yes
MARCO_047	Female	30	Positive	3a	N/A	No
MARCO_049	Female	46	Positive	1	N/A	No
MARCO_050	Female	38	Positive	3a	3	No
MARCO_060	Female	58	Positive	3a	3	Yes
MARCO_061	Female	25	Positive	3b	3	Yes
MARCO_062	Male	70	Positive	3b	N/A	No
MARCO_063	Female	33	Positive	3b	2	No
MARCO_064	Male	40	Positive	3b	3	Yes
MARCO_065	Male	59	Positive	3a	3	Yes
MARCO_066	Male	53	Positive	3b	3	Yes
MARCO_067	Female	34	Positive	3a	3	Yes
MARCO_068	Female	31	Positive	3a	3	Yes
MARCO_069	Female	43	Positive	3b	3	Yes
MARCO_070	Female	53	Positive	3c	2	No
MARCO_071	Female	61	Positive	3a	3	Yes
MARCO_072	Female	54	Positive	1	3	Yes
MARCO_073	Female	48	Positive	3b	2	No
MARCO_074	Female	27	Positive	3a	N/A	No

N/A, not available.

Supplementary Table 2. Demographic and MRI Endpoints for People With CD With Persistent or Resolved Gastrointestinal Symptoms at Follow-Up After 12 Months of Gluten-Free Diet

	CD With Persistent Symptoms (n = 6)	CD With Resolved Symptoms (n = 14)	*P*
Sex			
Female	4	8	—
Male	2	6	—
Age, y	45 ± 4	48 ± 4	NS
Body mass index, kg/m^2	27 ± 3	26 ± 2	NS
Biagi score	3.0 ± 0.0	3.0 ± 0.0	NS
Biagi score level	All at level 3	All at level 3	—
PHQ-15	13 ± 1	4.0 ± 0.8	<.0001
PHQ-3	4.3 ± 0.5	1.0 ± 0.2	<.0001
HADS anxiety score	9 ± 2	4.9 ± 0.7	.0138
HADS depression score	4 ± 1	2.1 ± 0.3	NS
HADS total score	14 ± 3	6.9 ± 0.8	.0070
Study day symptoms VAS, mm	81 ± 32	10 ± 3	.0045
Stool form	3.1 ± 0.4	3.5 ± 0.4	NS
Stool frequency, times/day	1.3 ± 0.2	1.5 ± 0.2	NS
Hydrogen breath test, ppm	2 ± 1	5 ± 2	NS
Small bowel water content, mL	182 ± 40	139 ± 27	NS
Ascending colon volume, mL	296 ± 20	230 ± 21	NS
Transverse colon volume, mL	303 ± 41	266 ± 30	NS
Descending colon volume, mL	257 ± 23	169 ± 23	NS
Total colon volume, mL	856 ± 68	665 ± 61	NS
Whole gut transit time WAPS, AU	1.5 ± 0.2	2.1 ± 0.4	NS
Whole gut transit time, h	47 ± 7	65 ± 12	NS
Acetic acid, μM/g dry stool	24 ± 1	24 ± 1	NS
Propanoic acid, μM/g dry stool	9 ± 3	7 ± 2	NS
Isobutyric acid, μM/g dry stool	1.7 ± 0.5	1.3 ± 0.1	NS
Butyric acid, μM/g dry stool	7 ± 2	6 ± 1	NS
Isovaleric acid, μM/g dry stool	2.1 ± 0.6	1.5 ± 0.2	NS
Valeric acid, μM/g dry stool	1.6 ± 0.5	1.4 ± 0.3	NS
Stool water (%)	66 ± 5	68 ± 2	NS

Values are mean ± SEM.

CD, celiac disease; HADS = Hospital Anxiety and Depression Score; MRI, magnetic resonance imaging; PHQ, Patient Health Questionnaire; VAS, visual analog scale; WAPS, weighted average position score of transit markers.