Effects of a low FODMAP diet on gut microbiota in individuals with treated celiac disease having persistent gastrointestinal symptoms – a randomized controlled trial

Anne Mari Herfindal¹, Frida van Megen²³⁴, Mari K. O. Gilde¹, Jørgen Valeur⁵, Knut Rudi¹, Gry Skodje⁶, Knut E. A. Lundin²⁷, Christine Henriksen⁴, Siv K. Bøhn¹* ¹Department of Chemistry, Biotechnology and Food Sciences, Norwegian University of Life Sciences, Ås, Norway ²KG Jebsen Coeliac Disease Research Centre, University of Oslo, Oslo, Norway ³Unit for Clinical Nutrition, Division of Cancer Medicine, Oslo University Hospital Rikshospitalet, Oslo, Norway ⁴Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway ⁵Unger-Vetlesen Institute, Lovisenberg Diaconal Hospital, Oslo, Norway ⁶Healthy Life Centre, Municipality of Nes, Norway ⁷Department of Gastroenterology, Oslo University Hospital Rikshospitalet, Oslo, Norway  *Corresponding author: Siv K. Bøhn, Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P. O. Box 5003, N-1432, Ås, Norway, email address: sivb@nmbu.no, telephone: +47 67 23 25 02,

Short version of the title: FODMAPs, microbiota, and celiac disease.

Keywords: Treated celiac disease, low FODMAP diet, gut microbiota, short-chain fatty acids (SCFAs), neutrophil gelatinase-associated lipocalin (NGAL).

This peer-reviewed article has been accepted for publication but not yet copyedited or typeset, and so may be subject to change during the production process. The article is considered published and may be cited using its DOI 10.1017/S0007114523001253

The British Journal of Nutrition is published by Cambridge University Press on behalf of The Nutrition Society
Abstract

Individuals with celiac disease (CeD) often experience gastrointestinal symptoms despite adherence to a gluten-free diet (GFD). While we recently showed that a diet low in fermentable oligo-, di-, monosaccharides, and polyols (FODMAPs) successfully provided symptom relief in GFD-treated CeD patients, there have been concerns that the low FODMAP diet (LFD) could adversely affect the gut microbiota. Our main objective was therefore to investigate whether the LFD affects the faecal microbiota and related variables of gut health. In a randomized controlled trial GFD-treated CeD adults, having persistent gastrointestinal symptoms, were randomized to either consume a combined LFD and GFD (n = 39) for four weeks or continue with GFD (controls, n = 36). Compared to the control group, the LFD group displayed greater changes in the overall faecal microbiota profile (16S rRNA gene sequencing) from baseline to follow-up (within-subject β-diversity, \( P < 0.001 \)), characterized by lower and higher follow-up abundances (%) of genus Anaerostipes (\( P_{\text{group}} < 0.001 \)) and class Erysipelotrichia (\( P_{\text{group}} = 0.02 \)), respectively. Compared to the control group, the LFD led to lower follow-up concentrations of faecal propionic and valeric acid (GC-FID) in participants with high concentrations at baseline (\( P_{\text{interaction}} \leq 0.009 \)). No differences were found in faecal bacterial α-diversity (\( P_{\text{group}} \geq 0.20 \)) or in faecal neutrophil gelatinase-associated lipocalin (ELISA), a biomarker of gut integrity and inflammation (\( P_{\text{group}} = 0.74 \)), between the groups at follow-up. The modest effects of the LFD on the gut microbiota and related variables in the CeD patients of the present study are encouraging given the beneficial effects of the LFD strategy to treat functional GI symptoms. (Registered at clinicaltrials.gov as NCT03678935).
1 Introduction

Celiac disease (CeD) is a common (1) immune-mediated disease (2) where intolerance to gluten can lead to severe health problems with a wide range of gastrointestinal (GI) and extra-intestinal symptoms (3). Although the majority of people with CeD recover when adhering to a gluten-free diet (GFD) (3), it is estimated that 29% experience ongoing GI symptoms (4). Persistent GI symptoms in CeD, which overlaps with the typical features of irritable bowel syndrome (IBS) (5), have been associated with reduced quality of life (6; 7). A diet low in fermentable oligo-, di-, monosaccharides, and polyols (FODMAPs), developed by researchers at Monash University in Australia (8; 9), has shown to be successful in reducing GI symptoms in IBS (10) and GFD-treated CeD (11; 12; 13).

FODMAPs naturally occur in many foods (14) and include fructose (in excess of glucose), lactose, oligosaccharides (fructan, galacto-oligosaccharides [GOS]), and polyols (i.e., mannitol and sorbitol). When reaching the small intestine, they may increase the osmotic pressure attracting water into the gut lumen. Also, the microbes harbouring our intestine (i.e., the gut microbiota (15)) produce gases such as CH₄, H₂ and CO₂ when fermenting food components (e.g., undigested FODMAPs) that reach the colon, possibly resulting in luminal distention of the intestine. Thus, FODMAP-rich foods may trigger GI problems such as bloating and abdominal pain in susceptible individuals, although the direct mechanisms remain unclear (10).

While we have previously shown that a low FODMAP diet (LFD) successfully provided symptom relief in GFD-treated CeD patients (11), there have been concerns that a diet that limits the nutrient source for the gut bacteria could adversely affect colonic health (16) by inducing changes in the composition of the gut bacteria and the metabolites they produce. Importantly, short-chain fatty acids (SCFAs) produced by colonic bacteria supports mucosal homeostasis locally (e.g., epithelial function, intestinal barrier function, and immunoregulation) and can also affect other parts of the body when taken up into the circulation (17). The most abundant SCFAs produced in the colon are acetic, butyric, and propionic acid. An LFD has in some studies (18; 19; 20) reduced the concentrations of faecal SCFAs in patients with IBS and inflammatory bowel disease (IBD) but did not have this effect in other studies (21; 22; 23). Interestingly, although the evidence is limited, dietary FODMAP restriction may also reduce immune activation (19; 24) and visceral sensitivity by lowering the concentration of faecal lipopolysaccharides (LPS) (25).
To our knowledge, effects of an LFD on gut microbiota and related markers of intestinal health have not been reported in GFD-treated CeD patients. The main aims of the present study were therefore to investigate whether an LFD in this patient group affects (i) the faecal microbiota, (ii) the concentrations of faecal SCFAs, and (iii) the concentrations of faecal human neutrophil gelatinase-associated lipocalin (NGAL, also known as lipocalin 2/LCN2), a biomarker of gut inflammation (26). In addition, we investigated whether LFD-induced changes in faecal microbiota and SCFAs could explain changes in GI symptoms, and the importance of baseline faecal characteristics for the symptom response to the LFD.

2 Materials and methods

2.1 Participants and study design

The present study is part of a clinical trial which followed a nonblinded, parallel randomized design, conducted from October 2018 to August 2019, at Oslo University Hospital, Rikshospitalet, Norway. The design and study population have been described previously (11). In short, participants with verified CeD having persistent GI symptoms despite adherence to a GFD were recruited using a web-based national survey. For inclusion, the participants had to be well-treated on a GFD for at least 12 months with serological and mucosal recovery prior to inclusion (anti-tissue transglutaminase 2 IgA <4 U/mL, anti-deaminated gliadin peptides IgG <20 U, Marsh score 0-1). The criterion for persistent GI symptoms was a Gastrointestinal Symptom Rating Scale IBS version (27) (GSRS-IBS) total score ≥30 at screening. Exclusion criteria are described in detail by van Megen et al. (11).

The participants were randomized to either an LFD group or a control group. The random allocation sequence was created by an investigator with no clinical involvement in the trial using STATA 15 (StatCorp LP) with random block sizes of 2, 4, and 6. The block sizes were varied to avoid the possibility of guessing the pattern in the randomization list. The allocation sequence was concealed for participants and the involved researchers until after completion of baseline assessments. Participants randomized to the LFD group were instructed to consume a GFD low in FODMAPs, while participants randomized to the control group were instructed to continue their regular GFD. Both diet groups were followed up for four weeks as out-patients and all participants were requested to refrain from introducing other dietary changes during the study period. Biological sampling including stool samples were performed at baseline and at 4-week follow-up.
2.2 Outcomes

The present study reports on the secondary outcomes of the randomized controlled trial (RCT) designed to investigate whether reduction of dietary FODMAPs can relieve GI symptoms in CeD patients with persistent IBS-like symptoms despite GFD-adherence. The primary outcome of the RCT was change in GI symptoms (GSRS-IBS scores), and the results have been reported previously (11). Briefly, GSRS-IBS is a questionnaire developed for self-evaluation of GI symptoms in IBS patients (27). It consists of 13 items which together covers symptoms related to satiety (two items), abdominal pain (two items), diarrhoea (four items), constipation (two items), and bloating (three items). All the items have a response scale from 1 (no discomfort at all) to 7 (very severe discomfort).

The primary outcomes of the present study were differences in faecal microbiota, SCFAs, and NGAL between the control and LFD group at 4-week follow-up. Additionally, we investigated associations between changes in GI symptoms (GSRS-IBS scores) and changes in gut microbiota and SCFAs, and associations between faecal characteristic at baseline and symptom response.

2.3 Collection and preparation of stool samples

The participants were instructed to collect stool samples in 13 mL tubes at home the same weeks as the sampling days (baseline, 4-week follow-up) and store the samples in their own freezers until the visits at the clinic. The samples were transported to Oslo University Hospital by the participants at the day of their visits, keeping the sample under cold conditions (~4°C) enabled by a specially made freezing element in Styrofoam boxes. The samples were temporary stored at -20°C for about four weeks followed by long-term storage at -80°C and transferred between collaborating institutions on dry ice.

Approximately 24 hours prior to aliquotation, the samples were moved from -80 to -20°C. The samples were slightly thawed on ice before preparing smaller aliquots in 1.5 mL Eppendorf tubes using sterile scalpels in a laminar air flow cabinet to avoid contamination of the environment. The samples were kept on ice during the aliquotation procedure to maintain a semi-frozen condition. To avoid contamination between samples, disposable scalpels and working tools were used when possible, and non-disposable tools were washed between each sample with warm detergent-containing water followed by disinfection with 70% ethanol and DNA decontamination reagent (43944, Sigma-Aldrich). Faecal aliquots were weighted and logged, and the aliquots designated for DNA extraction were added 600 μL room temperature
Stool transport and Recovery Buffer (cat. no. 03335208001, Roche) and vortexed. All aliquots were stored at -80°C until further processing.

2.4 DNA extraction from faecal material

Faecal aliquots for DNA extraction (~0.2 g in 600 μL Stool transport and Recovery Buffer) were crudely homogenized by vortexing and 300 μL of the resulting suspensions were transferred to tubes containing acid-washed glass beads (~0.2g <106 μm, ~0.2g 425-600 μm, and 1 2.5-3.5 mm, Sigma-Aldrich). All samples were processed twice on FastPrep 96 to obtain cell lysis (1,800 rpm, 40 s, 5 min cooling step at room temperature in-between, MP BioMedicals). Processed samples were centrifuged (~13,226 × g, 10 min) and 50 μL supernatants were transferred to 96-well plates for protease treatment followed by DNA extraction using Mag Midi LGC kit (cat. no. NAP40420, LGC Genomics) according to the manufacturer’s protocol on a KingFisher Flex DNA extraction robot (Thermo Fisher Scientific).

2.5 Library preparation and gene sequencing of 16S rRNA

Gene sequencing of 16S rRNA was performed according to the workflow reported by Avershina et al. (28). After DNA extraction from faecal material, 16S rRNA genes were amplified by polymerase chain reaction (PCR) using prokaryote-targeting primers (Supplemental Table 1&2). The resulting PCR products (~466 bp) were purified using AMPure XP (Beckman-Coulter) and 10 further PCR cycles with index primers modified with Illumina adapters were performed (Supplemental Table 3-5), resulting in PCR products of ~594 bp. All PCR products were qualitatively confirmed by gel electrophoresis. Quantification, normalization, and pooling of individual libraries were followed by purification by AMPure XP and quantification of the pooled library. The pooled library was diluted to 6 pM and sequenced with the MiSeq Reagent Kit V3 (Illumina, cat. nr. MS-102-3003, Illumina) on the Illumina MiSeq following Illumina’s protocol (16S Metagenomic Sequencing Library Preparation Part# 15044223 Rev. B), except we used nuclease free-water instead of Tris for PhiX library dilution. PhiX (20%, Illumina, cat. nr. FC-110-3001) served as internal control.

2.6 Processing of 16S rRNA sequencing data

The resulting 300 bp forward and reverse paired-end reads from gene sequencing of 16S rRNA were assembled, split into their respective samples, and quality-filtered using QIIME (29). The resulting dataset included in total 4,058,950 high-quality and chimera-checked
sequences from the 138 samples (10,367 to 114,847 sequences/sample), with mean (SD) sequencing depth of 29,413 (14,540) sequences/sample. Sequences were clustered into taxonomically assigned operational taxonomic units (OTUs) with ≥97% identity using the closed-reference USEARCH algorithm (version 8) (30; 31) against the SILVA database (release 128) (32). The OTU counts for each sample were normalized by down-sampling (rarefying) to an even library size of 10,000 sequences/sample in QIIME (core_diversity_analyses.py). No relationship was found between number of observed OTUs in the normalized OTU table and the original sequencing depth (Supplemental Figure 1). In the normalized dataset, 986 OTUs were identified in total (rarefaction curves in Supplemental Figure 2A). The OTUs were taxonomically binned into nine phylum (p_), 16 class (c_), 19 order (o_), 34 family (f_) and 119 genus (g_-) level taxa when excluding OTUs with no taxonomic assignment below kingdom level (OTUs classified as “unassigned” or “unknown bacteria”) and those taxa detected in less than 30% of the samples at both time points in both diet groups. For analyses where only baseline samples were included, six genera were excluded since not detected in more than 30% of the samples in either diet group. This OTU filtering was chosen to achieve a state of representativeness and robustness. Taxa abundances are presented as relative abundance (%) where the lowest detectable abundance was 0.01%.

2.7 Bacterial α- and β-diversity

Bacterial between-sample diversities (β-diversity; Jaccard, Bray-Curtis, weighted UniFrac distances (33)) and within-sample diversities (α-diversity; Chao1 and Shannon-Wiener index) were calculated based on the normalized OTU table(s) using QIIME default scripts (core_diversity_analyses.py). The Chao1 and Shannon-Wiener index aim to describe bacterial richness (estimated number of species (OTUs)) and total diversity (richness and evenness combined) of a bacterial community, respectively. Higher α-diversity indicates a more diverse community. Rarefaction curves for the Chao1 and Shannon-Wiener index are provided in Supplemental Figure 2B&C. β-diversity aim to quantify the overall dissimilarity (“distance”) between two bacterial communities, where weighted Unifrac distances are phylogenetic distances, Bray-Curtis abundance-based distances, and Jaccard incidence-based distances. All β-diversity indices used are expressed as dissimilarities ranging from 0 to 1, where higher numbers indicate less similar communities. For each participant, the β-diversity of the baseline and 4-week follow-up sample was used as a measure of change in the overall bacterial community structure (referred to as “within-subject β-diversity”). Ordination of β-
Accepted manuscript

diversities was performed using non-metric multidimensional scaling (NMDS). Details are provided in Supplemental method description.

2.8 SCFA measurements by GC-FID

Faecal homogenates for SCFA measurements by gas chromatography (GC) with flame ionization detection (FID) were prepared by homogenizing ~0.5 g faeces in 2 mL distilled water containing 3 mmol/L 2-ethylbutyric acid (internal standard) and 0.5 mmol/L sulfuric acid. A fraction of 2.5 mL of the homogenate was vacuum distilled, according to the method of Zijlstra et al. (34), as modified by Hoverstad et al. (35) and added 12% formic acid (30 µL per 500 µL distillate). The distillate was analysed with GC (Agilent 6850, Agilent) using a HP-FFAP WAX capillary column (part number 19091F-112E, serial number USE400345H, Agilent J&W GC columns; Agilent). The SCFAs acetic, propionic, butyric, iso-butyric, valeric, iso-valeric, caproic, and iso-caproic acid were quantified using internal standardization employing FID. The results for all SFCAs were adjusted for wet weight and are expressed as mmol/kg faeces and relative amount (% of total).

2.9 NGAL measurements by ELISA

Preparation of faecal aliquots for NGAL measurements by enzyme linked immunosorbent assay (ELISA) was based on Chassaing et al. (36) with some modifications. Aliquots (~100 mg) were mixed with 1 mL room temperature Dulbecco’s phosphate-buffered saline (DPBS, L0615, Biowest) with 0.1% Tween20 (P1379, Sigma-Aldrich). Homogenous faecal suspensions were obtained by vigorously mixing using a MM2 mixer mill (Retsch, speed 60, 20 min) with adapters for Eppendorf tubes (22.008.0005, Retsch). The suspensions were centrifuged for 10 min (13,523 × g, 4°C) and the resulting supernatants were diluted 1:100 in DPBS prior to ELISA. The 1:100 dilution was considered the most optimal for avoiding inhibitory effects based on trial experiments testing various dilutions. ELISA procedure was performed using Human lipocalin-2/NGAL Duoset ELISA (DY1757, R&D Systems) according to the manufacturer’s instructions with some modification. Modifications included using DPBS without Tween20 for the first wash and blocking with Reagent Diluent at 4°C overnight during plate preparations, using a 11-point standard curve, and after adding samples, incubating plates with shaking (300 rpm, horizontal movement). Standard curves were generated using 4-parameter logistic curve fit in Softmax Pro (version 6.5, 2015, Molecular Devices). Samples were analysed in duplicates. The NGAL results were adjusted for wet weight and are expressed as ng/g faeces.
2.10 Statistics

All statistical analyses were performed using R\textsuperscript{(37)}. Details are provided in \textbf{Supplemental method description}, including evaluation of test assumptions and specific use of packages and functions in R. All statistical analyses were performed per protocol, where only participants with available faecal outcomes at baseline and 4-week follow-up were included. Alternatives to exclusion, i.e., imputation of missing values, were not considered suitable since stool samples were only collected at two time points.

2.10.1 Sample size

As previously reported\textsuperscript{(11)}, the sample size estimation was based on a clinically relevant change in the GSRS-IBS total score, defined as a reduction of 7 points (SD, 10 points) from baseline to 4-week follow-up. With a power of 80\%, and a 2-sided 5\% significance level, it was estimated that 34 participants were needed in each group. The total number was increased to 74 participants to account for 10\% drop-out.

2.10.2 Between-group comparisons

Multiple linear regression (ANCOVA)\textsuperscript{(38)} was performed to estimate the effect of diet group (difference between the control and LFD group) at 4-week follow-up while adjusting for baseline values of the outcome variable in question (follow-up values \~ baseline values + diet group + (baseline values * diet group)). The interaction term (baseline values * diet group) was excluded when $P_{\text{interaction}} > 0.05$, leaving the reduced model with only baseline values and diet group as explanatory variables (follow-up values \~ baseline values + diet group). The estimated interaction or diet group effects are reported together with corresponding $P$ values ($P_{\text{interaction}}$ or $P_{\text{group}}$). Transformation of baseline and follow-up values (log10 or square root) was used when model assumptions were not reasonably met. ANCOVA was performed for the outcome variables Chao1 index, Shannon-Wiener index, taxa abundances, SCFAs, and NGAL.

Independent two-sample t-test or Wilcoxon rank-sum exact test (in cases where residuals were not reasonably normally distributed) was performed to compare within-subject $\beta$-diversity (single-value responses) between diet groups.

2.10.3 Within-group comparisons

Prior to between-group comparisons of taxa abundances, within-group comparisons (baseline versus 4-week follow-up within diet groups) using Wilcoxon signed-rank test (exact or with continuity correction) were performed. This within-group analysis approach was included as
a measure for reducing the number of taxa that would be subject to between-group comparison. Only taxa which had changed or tended to change from baseline to follow-up within the LFD group were included for between-group comparisons.

2.10.4 Correlations
Spearman’s correlation was run to assess associations between change in GI symptoms and change in faecal outcomes variables (taxa abundances, propionic acid, valeric acid) and between change in GI symptoms and faecal baseline characteristics (α-diversity, taxa abundances, SCFAs, NGAL). Spearman's rank correlation coefficients (Rho) and $P$ values are reported. Rho between 0 and (-) 0.3 is described as negligible, between (-) 0.3 and (-) 0.5 as weak, between (-) 0.5 and (-) 0.7 as moderate, between (-) 0.7 and (-) 0.9 as high, and between (-) 0.9 and (-) 1 as strong.

2.10.5 Statistical significance
For most statistical testing, unadjusted $P$ values below 0.05 were considered statistically significant. However, for within-group comparison of taxa abundances and Spearman’s correlation for GI symptoms against faecal baseline characteristics, an unadjusted $P$ value below 0.05 was only considered statistically significant if the Benjamini-Hochberg (BH) false discovery rate-adjusted $P$ value was also below 0.05, otherwise the result was referred to as a tendency.

2.11 Ethics and approvals
This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all procedures involving human participants were approved by the Regional Committee for Medical and Health Research Ethics, South East A on August 24, 2018 (identification 25809). Written informed consent was obtained from all participants. The study was also registered at clinicaltrials.gov with the registration number NCT03678935. The manuscript was prepared according to the CONSORT statement (http://www.consort-statement.org).

3 Results
3.1 Participants, GI symptoms, and FODMAP intake
Participant flow chart is provided in Figure 1. In total, 894 persons were invited by email to participate in the study. Of the 253 that responded, 165 met the inclusion criterion of GSRS-IBS total score of 30 or higher. Of these, 52 were excluded for not meeting the other inclusion criteria or for meeting at least one exclusion criterion, 27 declined to participate,
and 11 were excluded due to other reasons. The remaining 75 participants were randomized and allocated to the control ($n = 36$) or LFD group ($n = 39$). Two participants in the LFD group discontinued the intervention due to concomitant illness requiring antibiotic treatment, two participants withdrew, and one participant was excluded since CeD was not confirmed. GSRS-IBS score data, estimated FODMAP intake, and stool samples were collected at baseline and 4-week follow-up from the 70 participants that completed the study (control group: $n = 36$; LFD group: $n = 34$). NGAL analysis was performed on faecal samples from all participants while one participant in the LFD group was excluded from the microbiota analysis due to insufficient faecal material. Insufficient faecal material also led to exclusion of four participants from the SCFA analysis (two from each diet group).

Baseline characteristics of the participants stratified by diet groups are presented in Table 1. Briefly, the participants were mostly females (84%) around 40 years of age in which approximately 50% had BMI categorized as overweight or obese. Furthermore, most participants had a Marsh score equal to 0, with maximum score of 1, and several participants reported food allergies or intolerances (43%). Additional baseline characteristics have been previously reported.

As described by van Megen et al., the participants in the LFD group reported a lower intake of total FODMAPs at follow-up than in the control group. The reduction was mainly achieved by a reduced intake of excess fructose (monosaccharide), lactose (disaccharide), sorbitol, and mannitol (polyols) (Figure 2A&B). No difference between the LFD and control group was found for fructan or GOS (oligosaccharides) at follow-up. For the primary outcome of the RCT, it was found that the LFD significantly reduced GI symptoms in the LFD group as measured by GSRS-IBS total score and dimension scores abdominal pain, bloating, diarrhoea, and satiety, but not for constipation.

3.2 Effects of the LFD on the faecal microbiota

To evaluate effects of the LFD intervention on faecal microbiota, we performed a 16S rRNA gene-based sequencing approach, comparing the diet groups with regards to β-diversity, α-diversity, and relative abundance of bacterial taxa (Figure 2C-I).

3.2.1 Change in the overall faecal microbiota (Bray-Curtis β-diversity) was greater for the LFD group

Three different measures of between-sample diversity (β-diversity) were used to assess the overall change (“distance”) in bacterial community structure between the baseline and
follow-up sample for each participant (“within-subject β-diversity”): 1) Jaccard, 2) Bray-Curtis, and 3) weighted Unifrac distances (Figure 2C). Higher within-subject β-diversity indicates more change from baseline to follow-up and the three distance measures utilize different community factors as illustrated by correlation plots provided in **Supplemental Figure 3**.

The within-subject Bray-Curtis distances were higher in the LFD group than in the control group (t-test, estimate: 0.08, P < 0.001, Figure 2C) implying that the participants switching to an LFD had on average more changes in bacterial abundances from baseline to follow-up than those in the control group. No difference in within-subject β-diversities was found between the control and LFD group for within-subject Jaccard (t-test, estimate: 0.01, P = 0.21, Figure 2C) or weighted Unifrac distances (Wilcoxon rank-sum test, P = 0.12, Figure 2C).

To further evaluate whether the participants in the LFD group obtained a more uniform microbiota, distinguishable from the control group, we constructed NMDS ordination plots of the Bray-Curtis distances for the total sample population as well as separately for baseline, follow-up, and for each of the diet groups (**Supplemental Figure 4**). No distinct clustering of the samples according to diet group or timepoint was observed. Most importantly, the follow-up samples from the control and LFD group did not cluster according to group (Supplemental Figure 4C) and the baseline and follow-up samples from the LFD group did not cluster according to timepoint (Supplemental Figure 4E). These results indicate that the LFD did not induce large systematic changes in the faecal microbiota.

### 3.2.2 No effect of the LFD on faecal bacterial richness or evenness; α-diversity

Two different indices α-diversity were calculated and compared between the diet groups at follow-up with baseline adjustment (ANCOVA): 1) the Chao1 index (richness; estimated number of species) and 2) the Shannon-Wiener index (richness and evenness combined). No difference was found between the diet groups for Chao1 (estimate: -7.7, P_{group} = 0.20) or Shannon-Wiener (estimate: 0.01, P_{group} = 0.86) (Figure 2D). ANCOVA model plots for the Chao1 and Shannon-Wiener index are provided in **Supplemental Figure 5**.

### 3.2.3 Lower faecal abundance of Anaerostipes in the LFD group

To obtain a more detailed understanding of the changes in faecal microbiota induced by the LFD, we evaluated the effect of the LFD on bacterial relative abundances (%) at different taxonomic levels (phylum, class, order, family, and genus). First, the number of comparisons
between the diet groups was reduced by excluding the taxa where no change had occurred from baseline to follow-up within the LFD group (Supplemental Table 6). When only focusing on the taxa that had changed or tended to change within the LFD group (ten in total), between-group comparisons by ANCOVA with baseline adjustment (Table 2, ANCOVA model plots in Supplemental Figure 6) showed that the LFD group had significantly lower abundance of *Anaerostipes* (*P* < 0.001, Figure 2E) and higher abundance of *Erysipelotrichia* (*P* = 0.02, Figure 2F) at follow-up than the control group. The low abundant genus *Lachnospiraceae UCG-010* was also found to be higher in the LFD group than in the control group at follow-up (*P* = 0.03, Figure 2G). *Anaerostipes* and *Erysipelotrichia* were detected in faecal samples from all participants, while *Lachnospiraceae UCG-010* was detected in 60% of the samples.

Between-group comparisons by ANCOVA could not be performed for the abundant class Actinobacteria and the low abundant genera *Lachnospiraceae UCG-008*, *Tyzzerella3*, and *Candidatus Soleaferrea* due to strong deviations from parametric model assumptions, low abundance and/or low detectability (Table 2). The within-group comparisons (Supplemental Table 6) showed that Actinobacteria (mainly genus *Bifidobacterium*, Figure 2H) and *Tyzzerella3* tended to decrease, while *Lachnospiraceae UCG-008* and *Candidatus Soleaferrea* tended to increase within the LFD group. For these four taxa, the within-group analysis showed no tendency for change from baseline to follow-up within the control group.

Generally, the faecal microbiota of the participants was dominated by the phylum (median relative abundance) Firmicutes (65%), followed by Bacteroidetes (30%), Proteobacteria (2.3%) and Actinobacteria (1%). Only these four phyla were detected in faecal samples from all participants (Figure 2I). We also observed that the variation in relative abundances were high between and within participants. To exemplify, the relative abundance of Firmicutes ranged from 27 to 91%, and Actinobacteria ranged from 0.08 to 11%.

### 3.3 Effects of the LFD on faecal SCFAs

To investigate the impact of the LFD on microbial metabolism we compared the control and LFD group with regards to faecal concentrations of SCFAs which are produced in the colon through bacterial fermentation of carbohydrates and proteins. Average faecal concentrations (mmol/kg faeces) and relative amounts (% of total) of SCFAs at baseline and follow-up stratified by diet groups are presented in Table 3.
Using ANCOVA with baseline adjustment we found an interaction between baseline concentrations and diet group for both propionic ($P_{interaction} = 0.009$) and valeric acid ($P_{interaction} = 0.003$) (Table 3, Figure 2J&K). These interactions indicate that, compared to the control group, the LFD led to lower follow-up concentrations of propionic and valeric acid in participants with high concentrations at baseline. No difference between the control and LFD group at follow-up was found for total SCFAs, acetic, butyric, iso-butyric, iso-valeric, or caproic acid concentrations, and no differences between the diet groups were found for SCFA relative amounts (Table 3). Iso-caproic acid was measured but excluded from statistical analysis since only detected in 9% of the samples. Boxplot with response for each participant and ANCOVA model plot for all SCFAs can be found in Supplemental Figure 7 (concentrations) and Supplemental Figure 8 (relative amounts).

### 3.4 Effect of the LFD on faecal NGAL

To evaluate the potential effect of the LFD on intestinal inflammation, we measured the faecal concentrations of the sensitive inflammation marker NGAL (Figure 2L). At baseline, the control and LFD group had mean (SD) NGAL concentrations of 250 (130) and 261 (129) ng/g faeces respectively, and at follow-up, the concentrations were 250 (133) and 264 (120) ng/g faeces, respectively. By applying ANCOVA with adjustment for baseline values, the LFD group was estimated to have NGAL concentration of 8.5 ng/g faeces higher than the control group at follow-up, which was considered a nonsignificant difference ($P_{group} = 0.74$). ANCOVA model plot can be found in Supplemental Figure 9. One participant in the LFD group was excluded from the statistical analysis due to technical issues with sample handling in the laboratory.

### 3.5 Correlations between GI symptoms and faecal microbiota

Since the LFD was shown to improve GI symptoms in the study population (11), we assessed the potential association between changes in faecal microbiota and changes in GSRS-IBS symptom scores.

By performing Spearman’s correlation tests between changes in GSRS-IBS scores and within-subject Bray-Curtis distances independent of diet group, no significant correlations were found (Supplemental Table 7, $P \geq 0.11$). Within the LFD group, smaller within-subject Bray-Curtis distances (i.e., less change in overall microbiota structure) were weakly associated with larger reduction in diarrhoea (Rho = 0.44, $P = 0.010$, Supplemental Figure 10A, Supplemental Table 7). Within the control group, smaller within-subject Bray-Curtis
distances were weakly associated with larger reduction in constipation (Rho = 0.40, P = 0.01, Supplemental Figure 10B, Supplemental Table 7).

Spearman’s correlation tests between changes in bacterial abundances and changes in GSRS-IBS scores were also performed. A correlogram in Figure 3 summarizes the results while details (Rho and P values) are provided in Supplemental Table 8-10. Only taxa that changed or tended to change in LFD were included in the analysis (see Table 2 for taxa selection) and tests were first performed independent of diet group. Reduced relative abundances of both Anaerostipes (Rho = 0.26, P = 0.03) and Actinobacteria (Rho = 0.25, P = 0.04) were associated with reduced GSRS-IBS total scores. Regarding GSRS-IBS dimensions, reduced Anaerostipes abundances were associated with reduced bloating (Rho = 0.31, P = 0.01) and satiety scores (Rho = 0.30, P = 0.01) while reduced Actinobacteria abundances were associated with reduced bloating (Rho = 0.25, P = 0.04) and diarrhoea scores (Rho = 0.29, P = 0.02). Furthermore, increased abundances of Lachnospiraceae UCG-010 were associated with reduced GSRS-IBS total scores (Rho = -0.25, P = 0.04) and bloating scores (Rho = -0.36, P = 0.003). Finally, increased abundances of Candidatus Soleaferrea were associated with reduced bloating scores (Rho = -0.27, P = 0.02). All mentioned correlations were considered weak or negligible. When repeating the correlation analyses separately for each diet group, only Lachnospiraceae UCG-010 in the LFD group (Rho = -0.49, P = 0.004) and Candidatus Soleaferrea in the control group (Rho = 0.34, P = 0.04) displayed significant correlations.

3.6 Correlations between GI symptoms and faecal SCFAs

The LFD intervention seemed to decrease the faecal SCFA concentrations (mmol/kg faeces) of propionic and valeric acid in participants with high baseline concentrations. We therefore assessed whether changes in propionic and valeric acid concentrations were associated with changes in GSRS-IBS symptom scores. Spearman’s correlation tests were first performed independent of diet group (Supplemental Table 11). Increased concentrations of propionic acid were negligibly associated with reduced satiety scores (Rho = -0.25, P = 0.04) while changes in valeric acid concentrations were not associated with changes in GSRS-IBS total scores or dimensions (P ≥ 0.28). When performing correlation analyses separately for each diet group, no correlations were considered significant (Supplemental Table 11, P ≥ 0.23).
3.7 Correlations between GI symptoms and faecal baseline characteristics

The primary analyses of faecal microbiota, SCFAs, and NGAL revealed only minor differences between the control and LFD group at follow-up. Additionally, few significant correlations were found between changes in faecal microbiota/SCFAs and changes in GI symptoms. We therefore performed exploratory analyses to investigate whether the different faecal characteristics at baseline could be associated with the GSRS-IBS response to the LFD. Spearman’s correlation was used to assess associations between the faecal baseline characteristics (a-diversity, relative abundance of taxa, SCFAs, and NGAL) and changes in GSRS-IBS total scores. Analyses were performed independent of diet group and within each diet group.

No associations were found between changes in GSRS-IBS total scores and baseline α-diversity (Rho ≤ 0.11, P ≥ 0.53, Supplemental Figure 11A&B). Of the bacterial taxa present in at least 30% of the samples in at least one of the diet groups at baseline, abundance of 11 taxa tended to be associated with changes in GSRS-IBS total scores independent of diet group and/or within the LFD group (unadjusted P < 0.05, BH-adjusted P > 0.05, Supplemental Table 12). These included the abundant family Bacteroidaceae (completely dominated by genus Bacteroides), family Ruminococcaceae, and class Negativicutes (mainly the family Veillonellaceae). Lower Bacteroidaceae abundances at baseline were weakly associated with reduced GSRS-IBS total scores (Rho ≥ 0.25, P ≤ 0.04, independent of diet group and within the LFD group, Supplemental Figure 11C), higher baseline abundances of Ruminococcaceae were weakly associated with reduced GSRS-IBS total scores (Rho = -0.36, P = 0.04, within the LFD group, Supplemental Figure 11D), and lower abundances of Negativicutes at baseline were weakly associated with reduced GSRS-IBS total scores (Rho ≥ 0.26, P ≤ 0.04, independent of diet group and within the LFD group, Supplemental Figure 11E). In the control group, only two low abundant taxa (family Clostridiaceae1, genus Clostridiumsensustricto1) were positively associated with changes in GSRS-IBS total scores (Rho = 0.33, P = 0.05, Supplemental Table 12).

For faecal SCFAs, low concentrations and relative amounts of caproic acid were weakly associated with reduced GSRS-IBS total scores (Rho ≥ 0.27, P ≤ 0.03, Supplemental Figure 11F&G). Results from all SCFA variables are summarized in Supplemental Table 13.

Finally, no associations were found between changes in GSRS-IBS total scores and baseline NGAL concentrations (Rho ≤ 0.13, P ≥ 0.29, Supplemental Figure 11H).
4 Discussion

To our knowledge, we are the first to report on the impact of an LFD on the faecal microbiota composition and related variables of gut health in GFD-treated CeD patients with persistent GI symptoms. After four weeks, certain differences in gut microbiota were detected between the control and LFD group, and the SCFA results indicated that the LFD resulted in lower concentrations of propionic and valeric acid in participants with initially high concentrations. Biomarker of gut inflammation was however unaffected by the LFD. Importantly, we show that the participants in the LFD group had greater shifts in their overall microbiota profiles (Bray-Curtis within-subject β-diversity) from baseline to 4-week follow-up than the control participants. Furthermore, since a difference between the diet groups was only observed for the Bray-Curtis measure of within-subject β-diversity, and not for Jaccard and Weighted Unifrac, this suggests that the greater microbiota shifts in the LFD group were mainly characterized by abundance changes, and not by the bloom or disappearance of species or by phylogenetic changes. Our result contrast the result from a study of LFD in IBS, where no difference in β-diversity was found between the LFD and control group (39).

Furthermore, although the faecal microbiota had changed more in the LFD group than in the control group, these shifts did not result in any difference in bacterial α-diversity between the groups at follow-up. The finding that the LFD did not affect the bacterial α-diversity is consistent with that of other studies (20; 24; 40; 41).

In previous studies of LFDs, a decline in Bifidobacterium abundance has been the most consistent finding (19; 20; 21; 22; 39; 41; 42; 43). Because Bifidobacterium spp. are beneficial for human health (45), a reduction in these spp. due to FODMAP restriction has risen a concern for a potential unbeneficial effect for gut health. In the present study, we did not see a strong “anti-bifidogenic” effect by the LFD as reported by others. We only observed a tendency for reduced abundance of the class Actinobacteria (mainly genus Bifidobacterium) within the LFD group. This discrepancy might be related to the lack of reduction in FODMAP fibres (fructan/GOS) in the present study, as compared to other LFD studies that report on a significant reduction in these FODMAPs (20; 21; 22; 23; 39; 40; 43). Another explanation for the discrepancy might be a possible presence of initially low abundances of Bifidobacterium, which has been reported in other trials of GFD-treated CeD patients (46; 47; 48; 49; 50; 51). A low-gluten diet (52) and a GFD (53) has been shown to reduce Bifidobacterium abundance in healthy adult.
High abundance of the genus *Anaerostipes* has been identified as a potential biomarker of a healthy core microbiota in humans (54) and mice (55), possibly mediated through the production of the SCFA butyric acid (17). Thus, our distinct finding of lower follow-up abundances of *Anaerostipes* in the LFD group could indicate a potential undesired effect of the LFD. Also, the LFD group had modestly higher follow-up abundances of bacteria belonging to the class Erysipelotrichia and genus *Lachnospiraceae UCG-010*. Although literature on the implication of these bacteria in health and disease is unresolved (56) and scarce, a few papers point in the direction of association between high bacterial abundances and adverse gut-related conditions (54; 56; 57; 58).

While an LFD has previously been shown to affect bacterial metabolism with reduced faecal concentrations of the SCFAs acetic (18; 20) and butyric (18; 19) acid, no effect of the LFD was observed for these acids in the present study, possibly due to the lack of systematic reduction in fructan and GOS intake. Instead, our results indicate that the LFD reduced the concentrations of propionic and valeric acid in those participants with initially high concentrations. To our knowledge, a reduction in these SCFAs due to an LFD has not been reported before. While SCFAs are generally reported to be positively associated with gut health, certain evidence suggests that a reduction in SCFAs following the LFD could be regarded as favourable in the patient group of present study; Firstly, high concentrations of SCFAs (including propionic and valeric acid) have been measured in adults with long-term GFD-treated CeD compared with healthy controls (59; 60), possibly indicative of a bacterial dysbiosis in the intestine. Secondly, high concentrations of SCFAs (acetic and propionic acid) have been associated with more abdominal pain in IBS (61). Thus, if “unusual” high SCFA concentrations indicate a state of bacterial dysbiosis that potentially triggers GI symptoms, reduced concentrations by the LFD may be a favourable effect. However, there are many potential mechanisms that can explain the reduced faecal concentrations of valeric and propionic acid, including reduced production, increased absorption by colonocyte, and increased bacterial utilization (17). Interpretations of the reduced SCFA concentrations in the present study therefore remain speculative without further assessments. Also, since the present study did not include healthy controls, we do not know if the faecal SCFA concentrations were abnormal.

Low-grade inflammation has been implicated in IBS pathology (62; 63) and could also be involved in triggering IBS-like concomitant symptoms in GFD-treated CeD (64; 65). Furthermore, since an LFD has been shown to reduce the blood concentrations of pro-
inflammatory cytokines in IBS patients (19), we wanted to investigate whether an LFD could induce changes in a faecal marker of intestinal inflammation. We considered NGAL a relevant marker for this purpose since NGAL has been characterized as a highly sensitive marker of intestinal inflammation in IBD, with tendency for elevated concentrations in IBS (26). However, in our study population, no difference in faecal NGAL concentrations was found between the control and LFD group at follow-up. It is possible that we could have measured an effect of the LFD if a more sensitive faecal marker had been investigated. Beta-defensin 2 is one such candidate that could be worth investigating in symptomatic GFD-treated CeD since it was measured in higher concentrations in IBS patients compared to healthy controls (66; 67).

Since the LFD to some extent affected the gut microbiota and SCFAs, we investigated whether these changes were associated with changes in GI symptoms. Interestingly, within the LFD group, lesser change in the overall microbiota profiles (lower within-subject β-diversity) was weakly associated with larger reduction in diarrhoea. This finding could indicate that an optimal symptom response to the LFD is obtained when it is reducing the amount of water in the small intestinal and/or gas production while not affecting the bacterial composition. We also explored whether faecal characteristics at baseline were determinants for symptom response to the LFD since not all participants in the LFD group experienced the same level of symptom relief. We found that low baseline abundance of Bacteroides was weakly associated with higher symptom response to LFD, which is in line with some studies (43; 68). However, other studies have observed the opposite association (69; 70). Furthermore, in agreement with previous reports (18; 19; 22), the symptom response to the LFD in the present study was not associated with baseline concentrations of any of the major SCFAs produced in the colon. However, we did observe that higher symptom response was weakly associated with low baseline concentrations of the minor abundant SCFA caproic acid (also known as hexanoic acid), but the potential mechanism behind this association is unknown.

As already mentioned, one explanation for the lack of uniform changes in gut microbiota and SCFAs by the LFD in the present study could be the lack of reduction in FODMAP fibres (fructan/GOS), although the intake of total FODMAPs was reduced on average from 21 to 8 g/day (11). For fructan/GOS, the mean intake at baseline and follow-up in the LFD group was 2.8/0.7 and 3.8/0.5 g/day, respectively, with no difference in follow-up intake between the control and LFD. For comparison, the mean amount of total FODMAPs and FODMAP fibres in a typical Australian diet was ~24 and ~5.5 g/day (22). No systematic reduction in
fructan/GOS during the LFD in the present trial may be related to the fact that gluten-containing cereals and grains already excluded from the study participant’s diet, are also a considerable dietary source of FODMAP oligosaccharides (71). The FODMAP restriction in the presents study was instead mainly characterized by reduced intake of lactose, excess fructose, and polyols. In this regard, it is important to recognize that the effect of reduced intake of lactose and excess fructose on the gut microbiota highly depends on how much (or how little) of the consumed FODMAPs that escapes digestion in the small intestine and becomes available for bacterial fermentation. For example, in a study of people with or without lactose malabsorption, *Bifidobacterium* abundance was only increased by lactose intake in the group with lactose malabsorption (72). While lactose malabsorption is a common issue in CeD under active intestinal inflammation (73) less is known about the prevalence of fructose and lactose malabsorption/intolerance in long-term GFD-treated CeD, but it has been reported (74; 75; 76; 77). In the present study, participants were not assessed for malabsorption, so we are not able to link the variable bacterial responses directly to variations between participants in fructose/lactose malabsorption status. However, several participants reported having GI symptoms as response to the intake of lactose-containing foods which could indicate that at least lactose malabsorption was present in our study population (78). Thus, variations in lactose and/or fructose malabsorption could potentially explain why few uniform changes were found in specific bacterial taxa in response to the LFD. It should also be noted that the baseline intake of lactose and excess fructose in the LFD group in the present study was considerably higher (13.9 versus 1.35 g/day) and lower (2.1 versus 12.7 g/day), respectively, compared to the previously mentioned Australian study (22). Such differences in baseline diets highlights the difficulties regarding between-study comparisons.

This study has several limitations that should be acknowledged. Importantly, the sample size estimation for the RCT was based on a clinically relevant change in total GSRS-IBS score reported by van Megen et al. (11). It is therefore possible that the sample size of 32-36 participants in each diet group provided insufficient power to detect difference between the control an LFD regarding the faecal outcomes assessed in this study (i.e., microbiota, SCFAs, and NGAL). The assumed effect size for the total GSRS-IBS score was relatively high (Cohen’s d of ~0.7, calculated from data reported in van Megen et al. (11)), which might not be realistic for the faecal outcomes in this study. The duration of the intervention period should also be noted. In this study, we employed a 4-week FODMAP restriction period, which is commonly used in clinical practice. However, since longer restriction periods are
also possible, it is essential to investigate the effects of an LFD on gut microbiota and related variables for more extended periods. However, it is important to note that the LFD is not a long-term diet. Instead, it serves as method for determining the FODMAP subtypes that a patient can tolerate through a restriction, reintroduction, and personalization phase \(^{(79)}\). Therefore, examining the effect of LFD on gut microbiota in all three phases, rather than solely after the restriction phase, may be equally important. Furthermore, since the RCT was not blinded, we cannot rule out the possibility that the LFD group, while anticipating improvements, would overestimate symptom-improvements. Thus, our analyses for assessing associations between faecal outcomes and GI symptoms should be interpreted with caution. Lastly, it must be emphasized that the microbiota characterization in the present study was performed by the 16S rRNA sequencing approach and thus limited to measure the effects of the LFD on bacteria and archaea. Future studies should therefore investigate the impact of LFD on other microbes that harbour the human intestine, such as fungi. Other aspects related to the gut bacteria could be of relevance, including total bacterial load (as opposed to relative abundances), secondary bile acids, and gas production.

5 Conclusion

In conclusion, the present study demonstrates that a 4-week LFD impacts the gut microbiota to some degree in adult GFD-treated CeD patients having persistent GI symptoms. Specifically, the LFD led to changes in overall community structure of the faecal microbiota, with a possible unfavourable low faecal abundance of *Anaerostipes*, and low concentrations of the faecal SCFA propionic and valeric acid in participants with high concentrations of these acids at baseline. However, few uniform abundance changes of specific bacterial taxa within the LFD group were observed, indicating a highly individualized microbial response to LFD. Nevertheless, the relatively modest effects of the LFD on the gut bacterial composition and related variables in the CeD patients of the present study are encouraging given the beneficial effects of the LFD strategy to treat functional GI symptoms.
Acknowledgments
We acknowledge Gunn Helen Malmstrøm and Jennifer T. Fiennes (Lovisenberg Diaconal Hospital) for performing the analysis of the SCFAs, and Maria Fossli and Marianne Stendahl Deocareza (master students, University of Oslo) for contributing to the collection of stool samples. We are also thankful to Trude Elise Aspholm (Norwegian University of Life Sciences) for supervision during the NGAL analysis. We thank Morten Nilsen, Inga Leena Angell, and Ida Ormaasen for technical support. Finally, we would like to thank all the participants in the study.

Financial support
This work was supported by the Norwegian research fund for celiac disease (S.K.B); and South-Eastern Norway Regional Health Authority (C.H.). The funding sources had no role in the design, analysis, or writing of this article.

Conflict of interest
None.

Authorship
The authors’ responsibilities were as follows – A.M.H.: conducted research (microbiota laboratory analysis), analysed data and performed statistical analyses, wrote paper; F.V.M.: designed research (clinical trial administration); M.K.O.G.: conducted research (NGAL analysis); J.V.: conducted research (responsible for SCFA analysis); K.R.: provided essential materials (facilitating the microbiota analysis); G.S.: designed research (conceptualization); C.H.: designed research (conceptualization); K.E.A.L.: designed research (conceptualization, funding acquisition); S.K.B.: designed research (conceptualization, funding acquisition), conducted research (supervision), analysed data, wrote paper, had primary responsibility for final content; and all authors: read, commented, and approved the final manuscript.
References


Figure 1. Flow chart of the population selection process with inclusion and exclusion of participants and overview of analysed samples/data. FODMAP, fermentable oligosaccharides, disaccharides, monosaccharides, and polyols; GSRS-IBS, gastrointestinal symptom rating scale IBS version; IBS, irritable bowel syndrome; LFD, low FODMAP diet; NGAL, neutrophile gelatinase-associated lipocalin; SCFAs, short-chain fatty acids.
Figure 2. FODMAP intake, faecal microbiota, faecal SCFAs, and faecal NGAL for the control (CON) and low FODMAP diet (LFD) group. (A) FODMAP intake (g/day) at baseline (b) and 4-week follow-up (f) for the CON (n = 36) and LFD (n = 34) group. Colours represents subgroups of FODMAPs (lactose, fructan, excess fructose, sorbitol, GOS, and mannitol). (B) Change in total FODMAP intake (g/day) from baseline to follow-up for the
CON \((n = 36)\) and LFD \((n = 34)\) group. Each circle/triangle represents the change for one participant. Change is defined as follow-up minus baseline value. (C) Faecal within-subject β-diversity for the CON \((n = 36)\) and LFD \((n = 33)\) group. From left to right: binary Bray-Curtis, Jaccard, and weighted Unifrac. Each circle/triangle represents the β-diversity between the baseline and follow-up sample for one participant. Between-group comparison was performed using independent two-sample t-test (Jaccard, Bray-Curtis) or Wilcoxon rank-sum exact test (weighted Unifrac). (D) Faecal α-diversity at baseline \((b)\) and follow-up \((f)\) for the CON \((n = 36)\) and LFD \((n = 33)\) group. From left to right: Chao1 and Shannon-Wiener index. Between-group comparison at follow-up was performed using ANCOVA with adjustment for baseline values (no interaction). (E-H) Faecal relative abundances (% or square root transformed relative abundances at baseline \((b)\) and follow-up \((f)\) for the CON \((n = 36)\) and LFD \((n = 33)\) group of the bacterial (E) class Actinobacteria, (F) genus Anaerostipes, (G) class Erysipelotrichia (identical to family Erysipelotrichaceae) and (H) genus Lachnospiraceae UCG-010. Each circle represents one sample from one participant, and samples from the same participant are connected with a line. Between-group comparison at follow-up was performed using ANCOVA with adjustment for baseline values (no interaction). (I) Cumulative faecal relative abundances (%) of all detected phyla at baseline \((b)\) and follow-up \((f)\) for the CON \((n = 36)\) and LFD \((n = 33)\) group. One bar represents one participant at one timepoint. Only Firmicutes, Bacteroidetes, Proteobacteria, and Actinobacteria were present in all samples (bold font). (J,K) Faecal concentrations of \((J)\) propionic and \((K)\) valeric acid concentrations \((\text{mmol/kg faeces})\) at baseline and follow-up for the CON \((n = 34)\) and LFD \((n = 32)\) group. Each circle/triangle represents one participant. Red and blue line represents the fitted ANCOVA model with follow-up values as response variable and with baseline values, diet group, and diet group baseline interaction as explanatory variables. (L) Faecal concentrations \((\text{ng/g faeces})\) of NGAL at baseline \((b)\) and follow-up \((f)\) for the CON \((n = 36)\) and LFD \((n = 33)\) group. Each circle represents one sample from one participant, and samples from the same participant are connected with a line. Between-group comparison at follow-up was performed using ANCOVA with adjustment for baseline values (no interaction). FODMAP; fermentable oligosaccharides, disaccharides, monosaccharides, and polyols; GOS, galacto-oligosaccharides; NGAL, neutrophile gelatinase-associated lipocalin; SCFA, short-chain fatty acid. ns: \(P > 0.05\); *: \(P < 0.05\); **: \(P < 0.0001\).
Figure 3. Correlogram illustrating Spearman’s rank correlation coefficients (Rho) between change in bacterial abundances and change in GSRS-IBS scores (total score and dimensions pain, bloating, constipation, diarrhoea, and satiety). Analysis was performed for the total population (left panel, n = 69) and separately for the control (CON, middle panel, n = 36) and low FODMAP diet (LFD, right panel, n = 33) group. Change is defined as 4-week follow-up minus baseline value. Only the taxa which tended to change in abundance from baseline to follow-up in the LFD group are included (see Table 2). Details (Rho and P values) are provided in Supplemental Table 8-10. Black arrows next to taxon names indicate the direction of change that was observed in the LFD group for that specific taxon (↑: increased abundance; ↓: decreased abundance). Colour scale (red to blue) indicate strength and direction of the correlations (Rho). Significant correlations are marked with stars. Positive correlation indicates that reduced bacterial abundances are associated with reduced GSRS-IBS scores, while negative correlation indicates that increased abundances are associated with reduced GSRS-IBS scores. FODMAP, fermentable oligosaccharides, disaccharides, monosaccharides, and polyols; GSRS-IBS, gastrointestinal symptom rating scale IBS version; IBS, irritable bowel syndrome. *: P < 0.05; **: P < 0.01.
Table 1. Participant characteristics at baseline. *

<table>
<thead>
<tr>
<th></th>
<th>Control group (n = 36)</th>
<th>LFD group (n = 34)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female/male</strong></td>
<td>29/7 (81/19)</td>
<td>30/4 (88/12)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td>45 ± 15</td>
<td>46 ± 13</td>
</tr>
<tr>
<td><strong>European</strong></td>
<td>36 (100)</td>
<td>34 (100)</td>
</tr>
<tr>
<td><strong>BMI (kg/m^2)</strong></td>
<td>25 [23,28] †</td>
<td>25 [22,27]</td>
</tr>
<tr>
<td><strong>BMI category</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight (&lt;18.5)</td>
<td>0 (0) †</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Normal weight (18.5-24.9)</td>
<td>17 (49) †</td>
<td>16 (47)</td>
</tr>
<tr>
<td>Overweight (25.0-29.9)</td>
<td>13 (37) †</td>
<td>13 (38)</td>
</tr>
<tr>
<td>Obese (&gt;30.0)</td>
<td>5 (14) †</td>
<td>4 (12)</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school</td>
<td>1 (3)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>High school</td>
<td>13 (36)</td>
<td>5 (15)</td>
</tr>
<tr>
<td>University/college, lower degree</td>
<td>12 (33)</td>
<td>14 (41)</td>
</tr>
<tr>
<td>University/college, higher degree</td>
<td>10 (28)</td>
<td>14 (41)</td>
</tr>
<tr>
<td><strong>Allergy/intolerance</strong></td>
<td>14 (39)</td>
<td>17 (50)</td>
</tr>
<tr>
<td><strong>CeD diagnostics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of CeD (years)</td>
<td>7 [4,15]</td>
<td>9 [5,14]</td>
</tr>
<tr>
<td>anti-tTG2 IgA (10^3 U/L)</td>
<td>0.5 [0.5,1.4]</td>
<td>0.5 [0.5,1.5]</td>
</tr>
<tr>
<td>anti-DGP IgG (U)</td>
<td>2.5 [2.5,4.4]</td>
<td>2.5 [2.5,2.5]</td>
</tr>
<tr>
<td>HLA DQ 2</td>
<td>27 (82) ‡</td>
<td>28 (88) ‡</td>
</tr>
<tr>
<td>HLA DQ 8</td>
<td>6 (18) ‡</td>
<td>4 (13) ‡</td>
</tr>
<tr>
<td>Marsh score 0</td>
<td>26 (72)</td>
<td>25 (74)</td>
</tr>
<tr>
<td>Marsh score 1</td>
<td>10 (28)</td>
<td>9 (27)</td>
</tr>
<tr>
<td><strong>GSRS-IBS total score ‖</strong></td>
<td>37 ± 13</td>
<td>39 ± 11</td>
</tr>
</tbody>
</table>

* Values are presented as means ± SDs for normally distributed variables, medians [25th,75th percentiles] for non-normally distributed variables, or n (%) for categorical variables. Median is used for variables where data were not normally distributed in at least one group. LFD, low FODMAP diet; FODMAP, fermentable oligosaccharides, disaccharides, monosaccharides, and polyols; BMI, body mass index; CeD, celiac disease; tTG, tissue transglutaminase; DGP, deaminated gliadin peptide; HLA, human leukocyte antigen; GSRS, gastrointestinal symptom rating scale; IBS, irritable bowel syndrome.

† n = 35 (one missing).
‡ n = 33 (three missing).
§ n = 32 (two missing).
‖ GSRS-IBS total scores are presented as mean ± SD but are not checked for normality because the method is based on the mean independently of normality.
Table 2. Between-group comparisons of faecal bacterial abundances at baseline and 4-week follow-up in the control and LFD group. †

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Control group (n = 36)</th>
<th>LFD group (n = 33)</th>
<th>ANCOVA (n = 69)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Follow-up</td>
<td>Baseline</td>
</tr>
<tr>
<td>c_Actinobacteria</td>
<td>0.52 [0.02,1.07]</td>
<td>0.41 [0.04,0.97]</td>
<td>0.29 [0.13,1.05]</td>
</tr>
<tr>
<td>f_Rikenellaceae §</td>
<td>3.36 [2.58,5.87]</td>
<td>3.70 [2.93,6.05]</td>
<td>3.62 [2.49,6.04]</td>
</tr>
<tr>
<td>g_Alistipes §</td>
<td>3.25 [2.41,5.87]</td>
<td>3.70 [2.93,5.70]</td>
<td>3.62 [2.40,6.04]</td>
</tr>
<tr>
<td>g_Anaerostipes §</td>
<td>0.86 [0.45,2.00]</td>
<td>1.14 [0.65,1.43]</td>
<td>1.01 [0.41,1.90]</td>
</tr>
<tr>
<td>g_Lachnospiraceae UCG-008</td>
<td>0.01 [0.00,0.03]</td>
<td>0.03 [0.01,0.04]</td>
<td>0.02 [0.01,0.04]</td>
</tr>
<tr>
<td>g_Lachnospiraceae UCG-010 §</td>
<td>0.01 [0.00,0.02]</td>
<td>0.01 [0.00,0.02]</td>
<td>0.01 [0.00,0.02]</td>
</tr>
<tr>
<td>g_Tyzerella3</td>
<td>0.01 [0.00,0.06]</td>
<td>0.00 [0.00,0.07]</td>
<td>0.01 [0.00,0.12]</td>
</tr>
<tr>
<td>g_CandidatusSoleaferrea</td>
<td>0.00 [0.00,0.01]</td>
<td>0.00 [0.00,0.01]</td>
<td>0.00 [0.00,0.01]</td>
</tr>
<tr>
<td>g_Flavonifractor §</td>
<td>0.03 [0.02,0.08]</td>
<td>0.03 [0.01,0.08]</td>
<td>0.03 [0.01,0.06]</td>
</tr>
<tr>
<td>c_Erysipelotrichia §, ‡</td>
<td>0.39 [0.24,0.58]</td>
<td>0.40 [0.24,0.69]</td>
<td>0.48 [0.30,0.82]</td>
</tr>
</tbody>
</table>

† Bacterial abundances (%) are presented as medians [25th,75th percentiles]. LFD, low FODMAP diet; FODMAP, fermentable oligosaccharides, disaccharides, monosaccharides, and polyols.

‡ Between-group comparison: ANCOVA models were fitted with follow-up values as response variable and with baseline values and diet group as explanatory variables. Estimated difference between the diet groups at follow-up (control as reference) with corresponding \( P \) group value from ANCOVA. Only taxa which had shown a change or tendency for change within the LFD group (within-group comparison using Wilcoxon signed-rank test, see Supplemental Table 6) were subject to between-group comparison. ANCOVA could not be performed for four out of 10 taxa due to strong deviations from parametric model assumptions, low abundance, and/or low detectability.

§ ANCOVA performed on square root transformed baseline and follow-up values.

‡ Completely dominated by family Erysipelotrichaceae.

*: \( P < 0.05 \).
Table 3. Between-group comparisons of faecal SCFAs at baseline and 4-week follow-up in the control and LFD group. †

<table>
<thead>
<tr>
<th>SCFA</th>
<th>Control group (n = 34)</th>
<th>LFD group (n = 32)</th>
<th>ANCOVA (n = 66)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Follow-up</td>
<td>Baseline</td>
</tr>
<tr>
<td>Total SCFAs, conc</td>
<td>57 ± 21</td>
<td>55 ± 19</td>
<td>54 ± 20</td>
</tr>
<tr>
<td>Acetic acid, conc</td>
<td>32 ± 12</td>
<td>31 ± 11</td>
<td>31 ± 12</td>
</tr>
<tr>
<td>Butyric acid, conc</td>
<td>8.6 [7.2, 14]</td>
<td>9.2 [5.8, 14]</td>
<td>9.8 ± 4.9</td>
</tr>
<tr>
<td>Propionic acid, conc</td>
<td>8.9 ± 3.9</td>
<td>8.9 ± 4.7</td>
<td>8.7 ± 3.6</td>
</tr>
<tr>
<td>Iso-valeric acid, conc</td>
<td>1.9 ± 0.82</td>
<td>1.8 ± 0.61</td>
<td>1.8 ± 0.94</td>
</tr>
<tr>
<td>Valeric acid, conc</td>
<td>1.3 ± 0.47</td>
<td>1.4 ± 0.67</td>
<td>1.3 [0.91, 1.5]</td>
</tr>
<tr>
<td>Iso-butyric acid, conc</td>
<td>1.3 ± 0.51</td>
<td>1.3 ± 0.37</td>
<td>1.2 ± 0.60</td>
</tr>
<tr>
<td>Caproic acid, conc</td>
<td>0.57 ± 0.67</td>
<td>0.58 ± 0.66</td>
<td>0.54 ± 0.64</td>
</tr>
<tr>
<td>Acetic acid, %</td>
<td>57 ± 6.0</td>
<td>57 ± 7.3</td>
<td>57 ± 6.3</td>
</tr>
<tr>
<td>Butyric acid, %</td>
<td>18 ± 5.7</td>
<td>17 ± 4.8</td>
<td>18 ± 5.3</td>
</tr>
<tr>
<td>Propionic acid, %</td>
<td>16 ± 3.8</td>
<td>16 ± 4.5</td>
<td>16 ± 3.5</td>
</tr>
<tr>
<td>Iso-valeric acid, %</td>
<td>3.8 ± 1.8</td>
<td>3.7 ± 1.7</td>
<td>3.5 ± 1.6</td>
</tr>
<tr>
<td>Valeric acid, %</td>
<td>2.4 ± 0.93</td>
<td>2.6 ± 1.0</td>
<td>2.5 ± 0.86</td>
</tr>
<tr>
<td>Iso-butyric acid, %</td>
<td>2.6 ± 1.1</td>
<td>2.6 ± 0.96</td>
<td>2.4 ± 0.97</td>
</tr>
<tr>
<td>Caproic acid, %</td>
<td>0.72 [0.12, 1.7]</td>
<td>0.79 [0.1, 1.6]</td>
<td>0.98 ± 1.1</td>
</tr>
</tbody>
</table>

† SCFA concentrations (“conc”, mmol/kg feces) and percentages (“%”, of total SCFAs) are presented as means ± SDs or medians [25th,75th percentiles]. SCFA, short-chain fatty acid; LFD, low FODMAP diet; FODMAP, fermentable oligosaccharides, disaccharides, monosaccharides, and polyols.
‡ Between-group comparison: ANCOVA models were fitted with follow-up values as response variable and with baseline values and diet group as explanatory variables. Estimated difference between the diet groups at follow-up (control as reference) with corresponding P value from ANCOVA, except where specified.
§ Estimated diet group * baseline interaction (control as reference) with corresponding P interaction value from ANCOVA.
|| ANCOVA performed on square root transformed baseline and follow-up values.
*: P < 0.05.