Similar Gut Bacterial Composition Between Patients With Ulcerative Colitis and Healthy Controls in a High Incidence Population: A Cross-sectional Study of the Faroe Islands IBD Cohort

Marjun á Fríðirksmørk Berbisá, Cand.Scient,∗¶, Kári Rubek Nielsen, MD,† Anna Cäcilia Ingham, PhD,‡ Jóngerð Midjord, MSc,† Turíð Hammer, PhD,§ Poula Patursson, Cand.Scient,† Nicolina M. O. Vest, Cand.Scient,§ Noomi O. Gregersen, PhD,‖ Johan Burisch, MD, PhD,∗∗#, and Amanda Vang, PhD¶,††#, Address correspondence to: Amanda Vang, PhD, Hoyviksvegur 51, 100 Tórshavn, Faroe Islands (amanda@fiskaaling.fo).

Background: The Faroe Islands has the world’s highest incidence of inflammatory bowel disease (IBD). Epidemiological studies have characterized this unique cohort and a decreased risk of developing IBD with emigration. Therefore, this well-characterized Faroese IBD cohort gives the opportunity to better understand this complex disease. This study represents the first investigation of the gut microbiota for the cohort.

Methods: This cross-sectional study consisted of 41 patients with established ulcerative colitis and 144 age- and sex-matched healthy controls recruited through the Faroe Genome project. Participants donated a 1-time fecal sample and completed questionnaires on food frequency, background health, and lifestyle. 16S rRNA amplicon sequencing of the V3-V4 region was performed followed by bioinformatic analysis of taxonomy and diversity metrics.

Results: The overall bacterial composition in both groups was dominated by Firmicutes and Bacteroidetes. No significant differences were found based on metrics of alpha or beta diversity. However, discriminatory analysis identified differential abundance of several indicator taxa in healthy controls and ulcerative colitis participants, whereas Akkermansia was completely absent from 27% of all study participants. Food frequency questionnaires revealed similar dietary patterns between the 2 groups.

Conclusion: The similarity in bacterial community composition and absence of the beneficial Akkermansia genus in both groups raise further questions concerning the underlying susceptibility toward inflammatory disorders within this high-risk population. Results vary widely by study design and geographic location, which speaks to the need for regionally tuned reference groups and disease-based studies on the Faroe Islands.

Lay Summary
The Faroe Islands has the world’s highest incidence of IBD. This is the first investigation characterizing gut microbiota for this unique cohort. No significant differences were found between ulcerative colitis and healthy controls based on alpha or beta diversity.

Key Words: ulcerative colitis, Faroe Islands, isolated population, gut microbiota, gut microbiome

Introduction
Inflammatory bowel disease (IBD) is generally accepted to result from a complex relationship between genetic susceptibility, environmental factors, and alterations in the intestinal microbiota. The global incidence and prevalence of noncommunicable inflammatory diseases, including IBD, has steadily increased since the 1970s and is considered to be linked to societal industrialization, including changes in lifestyle and diet.1,3

The Faroe Islands constitute a unique genetically and geographically isolated population located in the North Atlantic Ocean. Previous epidemiological studies have found the worldwide highest incidence rate of IBD in the Faroe Islands at 80 per 100 000 person-years.4,5 Ulcerative colitis (UC) is the predominant phenotype with an incidence estimate at 39 per 100 000 person-years compared with the incidence rate of IBD unclassified at 32 per 100 000

© 2022 Crohn’s & Colitis Foundation. Published by Oxford University Press on behalf of Crohn’s & Colitis Foundation. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
person-year and Crohn’s Disease at 10 per 100 000 person-years. \(^4\) The foundation for IBD research within the Faroese population is a nationwide IBD patient cohort that includes all IBD patients diagnosed with UC, Crohn’s Disease and IBD unclassified since 1960. \(^7\) Despite high familial aggregation of IBD, \(^8\) the Faroese population has not always been at high risk of IBD. A study covering 60 years of the Faroese IBD cohort reported an approximately 14-fold increase in the IBD incidence during the study period. \(^4\) Furthermore, Faroese immigrants to Denmark show a decreased risk of developing UC over time, which highlights the potential influence of environmental factors on disease risk in this unique population. \(^9\)

The complexity of understanding the molecular underpinning of IBD is further highlighted by the variability seen with respect to the intestinal microbiota. Shifts in the balance between commensal beneficial bacteria and opportunistic pathogenic bacteria, referred to as dysbiosis, are well documented by extensively reviewed IBD studies. \(^1\) Decrease in alpha diversity and separate clustering of the intestinal microbiota are common findings in fecal samples from IBD compared with healthy controls (HCs). \(^1\) However, conclusions based on dysbiosis status are difficult to generalize because study results vary widely, \(^11\) especially regarding potential indicator taxa. \(^12\), \(^13\) Geographic location and ethnic background have been suggested to have a stronger influence on the microbial community structure than disease status. \(^14\) Therefore, studying isolated populations with a high occurrence of IBD is a unique opportunity to understand the interaction between genetics and environment, which is an a priori condition for disease manifestation.

**Aim**

The aim of this study was to generate an initial dataset describing the bacterial composition of the gut microbiota among a subset of patients from the Faroese IBD cohort and from individuals without IBD from the background population. To date, no studies have characterized the microbiota of the Faroese population. This dataset thus provides a first look that can direct further microbiota-related investigations within this population and adds a deeper insight into this epidemiologically well-characterized IBD cohort.

**Materials and Methods**

**Study Design**

This cross-sectional observational study compared patients with established UC (n = 41) with HCs (n = 144). Only UC patients were included in order to focus on the predominate IBD phenotype within the Faroese cohort.

**Recruitment**

HC participants were matched and recruited by the Faroese Genome project under the authority of the Genetic Biobank of the Faroe Islands. UC patients were identified from the PROGENY-based IBD cohort database maintained by the responsible clinician at the gastroenterology outpatient clinic at the National Hospital of the Faroe Islands. Extent of disease for the included UC participants was registered by the Montreal classifications as E1 (rectum), E2 (distal to splenic flexure), or E3 (proximal to splenic flexure). \(^15\)

Both groups were recruited by an invitation letter sent to their home address. Some additional UC patients were recruited at routine clinical consultations. Following written consent, participants received a collection kit and corresponding questionnaires to return by mail to the Genetic Biobank of the Faroe Islands. Recruitment and sample collection occurred from October 2018 to June 2019.

**Inclusion Criteria**

All recruited patients met the Copenhagen Diagnostic Criteria for UC with disease diagnosed by endoscopy and verified by histology. \(^16\) Genealogical information, clinical information concerning disease history, medications, and extraintestinal disease manifestations, and epidemiological information were curated in the PROGENY database at Genetic Biobank of the Faroe Islands (Table 1).

All participants resided on the Faroe Islands at the time of recruitment and had at least 1 parent of Faroese origin. Age was restricted to older than 18 years at recruitment and 20-40 years at diagnosis. No participant had consumed antibiotics within 3 months prior to sample collection. Participants who had consumed probiotics, prebiotics, or other supplements were not excluded, but consumption was registered by questionnaire as described in the following section. HCs had no diagnosed gastrointestinal, rheumatological, or dermatological diseases. No UC patients had undergone colectomy. To create 2 cohorts with similar age and sex distributions, UC patients were matched with HCs based on current age and sex.

**Cohort Questionnaires**

All study participants completed questionnaires regarding family history of immune-mediated diseases (asthma, psoriasis, heart disease, etc.), details about their childhood (breastfed, exposure to passive smoking, family pets, childhood residency, etc.), employment history, time spent living abroad, physical activities, body mass index (BMI), and smoking status.

In addition, all participants were asked to complete a food frequency questionnaire (FFQ), developed in consultation with a clinical nutritionist, regarding their diet during the week prior to fecal sample collection. The FFQ consisted of a finite list of foods and beverages, including consumption of traditional Faroese foods, such as fermented lamb, dried fish, pilot whale meat, and blubber. Questions regarding use of antibiotics, probiotics, vitamins, or other supplements during the past 3 months were also included. Responses to the FFQ were recorded as 0 (have not consumed in the past week), 1 (consumed 4-7 days ago), 2 (consumed 2-3 days ago), 3 (consumed once yesterday), or 4 (consumed more than once yesterday/daily).

UC patients also completed the Simple Clinical Colitis Activity Index (SCCAI) survey at the time of sample collection to estimate disease activity.

**Sample Collection and Storage**

The OMNIgene-Gut Stool Microbiome Kit from DNA Genotek was used to collect and stabilize the fecal sample. Fecal microbiota samples collected with the OMNIgene-Gut kit have shown to closely resemble that of immediately frozen samples and allow for sample return without a cold chain. \(^17\) Sample aliquots were stored at −20 °C until DNA
Table 1. Demographic and clinical characteristics of UC patients and HCs

|                     | UC (%) | HCs (%) | Total (%) |
|---------------------|--------|---------|-----------|
| Total               | 41 (22)| 144 (78)| 185 (100)|
| Sex                 |        |         |           |
| Male                | 18 (44)| 57 (40)| 75 (40)  |
| Female              | 23 (56)| 87 (60)| 110 (60)|
| Age                 |        |         |           |
| Mean age at inclusion (SD; y) | 48 | 55 | 53 |
| Mean BMI (SD; kg/m²) | 26.4 | 25.4 | 25.7 |
| Smoking             |        |         |           |
| Current             | 7 (17) | 23 (16)| 30 (16)  |
| Former              | 10 (24)| 42 (29)| 52 (28)  |
| Never               | 24 (59)| 79 (55)| 103 (56)|
| Participant reported co-morbidities |       |         |           |
| None                | 25 (61)| 121 (84)| 146 (79)|
| Asthma              | 0 (0)  | 0 (0)  | 0 (0)    |
| Heart disease       | 4 (10) | 0 (0)  | 4 (2)    |
| Peripheral joint pain/ arthritis | 2 (5) | 3 (2) | 5 (3) |
| Axial joint pain/ arthritis | 3 (7) | 0 (0) | 3 (2) |
| Psoriasis           | 5 (12) | 4 (3)  | 9 (5)    |
| Diabetes            | 2 (5)  | 0 (0)  | 2 (1)    |
| Immune-related diseases in family members |       |         |           |
| Asthma              | 4 (10) | 20 (14)| 24 (13)  |
| Heart disease       | 16 (39)| 38 (26)| 54 (29)  |
| Peripheral joint pain/ arthritis | 11 (27)| 27 (19)| 38 (21)|
| Axial joint pain/ arthritis | 3 (7) | 5 (4) | 8 (4) |
| Psoriasis           | 10 (24)| 22 (15)| 32 (17)  |
| Diabetes            | 11 (27)| 34 (24)| 45 (24)|
| Mean age at UC diagnosis (SD; y) | 27 | – | – |
| UC disease duration (IQR; y) | 19.5 (10–30.5) | – | – |
| UC extent           |        |         |           |
| E1, proctitis       | 22 (54)|        |          |
| E2, left-sided      | 9 (22) |        |          |
| E3, extensive       | 10 (24)|        |          |
| UC disease activity |        |         |           |
| Active              | 17 (42)|        |          |
| Remission           | 24 (59)|        |          |
| Previous abdominal surgery | 0 | – | – |
| Medical treatment   |        |         |           |
| None                | 3 (7)  |        |          |
| 5-ASA               | 38 (90)|        |          |
| Immunosuppressants  | 22 (54)|        |          |
| Biological          | 9 (22) |        |          |
| Combination therapy | 22 (54)|        |          |

Abbreviations: BMI, body mass index; E1, rectum; E2, distal to splenic flexure; E3, proximal to splenic flexure; HCs, healthy controls; IQR, interquartile range; UC, ulcerative colitis; 5-ASA, 5-Amino-salicylic acid.

bMean (SD).

Median (IQR).

DNA was extracted using the DNeasy PowerSoil Pro kit from Qiagen according to the manufacturer’s recommendations, which include a bead beading step to ensure bacterial cell lysis and a polymerase chain reaction (PCR) inhibitor removal step. DNA was eluted in sterile 10 mM Tris-Cl, pH 8.5, and quantified by Qubit. Sterile stabilizing liquid from an unused OMNLgene-Gut kit was routinely processed as a negative control.

Sequencing Library Preparation

DNA extraction. As part of processing, stabilization media from an unused OMNLgene-Gut kit was routinely stored for parallel processing as a negative control.

Sequencing Library Preparation

DNA was extracted using the DNeasy PowerSoil Pro kit from Qiagen according to the manufacturer’s recommendations, which include a bead beading step to ensure bacterial cell lysis and a polymerase chain reaction (PCR) inhibitor removal step. DNA was eluted in sterile 10 mM Tris-Cl, pH 8.5, and quantified by Qubit. Sterile stabilizing liquid from an unused OMNLgene-Gut kit was routinely processed as a negative control along with processing of returned fecal samples.

The Zymo Quick 16S library prep kit was used to first amplify the V3-V4 region using primers based on the Klindworth V3-V4 341F and 785R primer set followed by magnetic bead cleanup and dual index barcoding. Library preparation was performed using quantitative polymerase chain reaction (qPCR), which allowed each individual amplicon and indexing PCR reaction to be visually verified for successful amplification and index attachment and quantification prior to pooling. The integrity of select libraries was verified by Agilent Bioanalyzer. Libraries for negative controls as well as Zymo mock community positive controls were prepared along with experimental samples.

Sequencing

DNA sequencing resulted in a total number of 8694870 reads, with 4158271 reads generated after quality filtering from 192 samples, including the negative and positive controls. An average of 48% of raw reads per sample remained after filter trimming and removal of chimeras.

The 2 negative controls had 0 and 2 reads, respectively, after quality filtering, indicating no systematic contamination in the data set. To identify putative contaminants amplified in samples with low DNA concentration, we applied the frequency method (threshold 0.1) of package decontam. After quality filtering, indicating no systematic contamination in the data set. To identify putative contaminants amplified in samples with low DNA concentration, we applied the frequency method (threshold 0.1) of package decontam.

Rarefaction curves verify that all included samples contained more than 10 000 reads after quality filtering and contamination removal and that no further ASVs would be detected with additional reads (Supplementary Figure 1). The final data set comprised 13 386 ASVs in 185 samples.

Two mock community samples were sequenced as positive controls. These samples contained bacterial species as well as yeast species in defined abundances. The results reflected the expected taxa and relative abundance, and no yeast taxa were identified, highlighting the quality of the data (Supplementary...
Figure 2). Taken together, the mock community controls and rarefaction plots show that the data can be considered biologically representative.

Statistical Analyses
Statistical analyses were performed in R (version 4.0.1). The package phyloseq (version 1.32.0) was used to integrate the ASV count table, taxonomy, and metadata. The results were visualized with the packages ggplot2 (version 3.3.1) and metacoder (version 0.3.4). Bacterial alpha diversity was calculated from raw counts by the Shannon index. To compare alpha diversity between groups (UC vs HC, sex, smoking status, medication, extent of disease), Kruskal-Wallis tests and pairwise Wilcoxon rank sum tests with Benjamini-Hochberg correction for multiple testing were performed. Pearson’s product moment-correlation with Benjamini-Hochberg correction for multiple testing was subsequently used to determine significant differences in beta diversity. Taxonomic heat trees based on relative abundance up to genus level were generated with the package metacoder.

To identify ASVs that are differentially abundant between UC patients and HCs, we applied DESeq2 (version 1.28.1) on the 100 most abundant ASVs in the data set. In addition, we used a compositional method to assess whether groups of ASVs characterize the microbiota of UC patients and HCs (package selbal, version 0.1.0). Moreover, linear discriminant analysis effect size (LefSe) was estimated on genus count data (top 50) normalized by cumulative sum scaling (package microbiomeMarker) to assess differentially abundant taxa on several taxonomic levels (here, kingdom to genus).

Ethical Considerations
This study was approved by the Research Ethics Council of the Faroe Islands on June 14, 2018 and by the Faroese Data Protection Authority (18/00201#3) on August 31, 2018.

Results
Study Cohort
The study cohort included 185 participants, with 41 patients (22%) diagnosed with UC and 144 HCs (78%) (Table 1). The sex distribution in the cohort was 40% males and 60% females, and age distribution in the UC group and HC group was equivalent. Mean BMI was 25.7. A total 16% of the participants were current smokers, 28% former smokers, and 56% nonsmokers. Extent of disease for UC patients was 54% diagnosed with E1, 22% with E2, and 24% with E3. For medication, 85% of the cohort was treated with 5-ASA, approximately 50% were receiving immunosuppressants or combination therapy, 25% were receiving biologics, and 4% were unmedicated. According to the SCCAI scores, 41% of the UC patients had active disease at the point of sample collection while 59% were in remission.

Background questionnaires regarding immune-mediated diseases for cohort participants showed joint pain (9%-13% UC, 0%-3.5% HC) and/or psoriasis comorbidity (18.2% UC and 2.7% HC) for the cohort participants. Heart disease (11.4% UC) and diabetes (4.5% UC) were reported in the UC group but not the HC group. In first- and second-degree relatives, heart disease and diabetes were the most common reported diseases, followed by psoriasis, peripheral joint pain, and asthma. Psoriasis and peripheral joint pain were reported more frequently for family members of IBD participants, whereas more asthma was reported for family members of HCs. Statistics were not calculated for immune-mediated diseases because these were self-reported data and not clinically verified.

Food Frequency
On the day of fecal collection, participants were required to report their eating habits over the previous week. Median and average frequency of consumption for each food item and supplement type were calculated for UC and HC groups as well as the total cohort (Supplementary Figure 3). The results from the FFQ resembled a typical western diet, with high consumption of dairy, starch, meat, vegetables, fruits, coffee or tea, and sweets. No apparent differences were revealed in the pattern of eating habits between the study groups. Traditional Faroese foods were in the least frequently consumed food groups during the study period, whereas processed foods such as cold meat cuts, snacks, sweets, and high-starch foods were more common. Approximately 40% of the study participants consumed nutritional supplements, such as vitamin D, multivitamins, fish oil, magnesium, and other supplements daily; however, the UC patients consumed more supplements than the HC group. Four percent of the participants reported consumption of probiotics but were not identified as outliers during data analysis.

Alpha Diversity
There was no significant difference in observed ASV richness or Shannon diversity between UC and HC groups (Figure 1). Further analysis related to the demographic characteristics, sex, smoking, or BMI did not reveal any additional separation regarding alpha-diversity between the HC and UC group. Shannon diversity was significantly positively correlated with age (at sampling) in the UC group (P-value = .02), but not in the HC group (Supplementary Figure 4). Comparing within the UC group based on disease activity, extent of disease or age of onset also resulted in similar alpha diversity with no significant differences between groups. However, patients with standard treatment had significantly lower Shannon diversity compared with patients with no treatment (P-value = .04) (Supplementary Figure 4).

Beta Diversity
Principal coordinates analysis plots of the Bray-Curtis dissimilarity index were generated to address beta diversity. Following analysis of similarities testing, there were no significant differences between UC and HC groups (Supplementary Figure 5).
Taxonomical Classification of the Cohort
Relative abundance of the microbiota at family level is presented as a percent of total identified ASVs for each participant to show the variation between individuals in the entire cohort (Figure 2). To visualize the pattern associated with the entire cohort across multiple taxonomical levels, heat trees were constructed to show the relative abundance of taxa (down to genus level) by utilizing color and branch-size for each taxonomical unit (Supplementary Figure 6).

Because the diversity measurements showed no significant difference between the UC and HC group, the following characterization of the taxonomical levels is representative of the combined UC and HC cohort. Firmicutes accounted for 64.5% of the total ASVs at phylum level (Supplementary Figure 7), with the highest abundance attributed to the Clostridia (57.9%) class. Lachnospiraceae (26.3%) and Ruminococcaceae (26.4%) represented the most abundant taxa within the family level.

At the genus level within these 2 families, Blautia (7.1%) was the most abundant within the family of Lachnospiraceae, followed by Agathobacter (3.7%), and Faecalibacterium (7%) was the most abundant genus within the family of Ruminococcaceae, followed by Ruminococcaceae_UCG (5.9%) and Ruminococcus (5.4%).

The second-most abundant phylum was Bacteroidetes (20.95%; Supplementary Figure 7). At the genus level, Bacteroides (10.2%), Prevotella_9 (3.9%), and Alistipes (1.7%) accounted for the majority of ASVs.

Actinobacteria accounted for 12% of the total ASVs, with Bifidobacterium (4%) and Collinsella (4%) as the majority at the genus level. Proteobacteria accounted for 2.8% of the total ASVs, with Sutterella (0.75%) at the genus level. Verrucomicrobia represented 1.2% of the total ASVs, exclusively including the Akkermansia genus.

Presence/Absence of Taxa
Because there was no significant difference based on alpha and beta diversity between the HC and UC groups, we decided to sort the taxonomical tables based on presence/absence to explore if any taxa routinely associated with intestinal health were absent from both UC and HC groups as well as to detect the presence of pathogens, which might suggest intestinal infection in both groups.

We found that 17 out of 41 (42%) UC samples and 33 out of 144 (23%) HC samples had 0 sequence reads assigned to the Akkermansia genus (Supplementary Figure 8), which is 27% of all study participants.

ASVs assigned to the Escherichia/Shigella genera were identified in 21 out of 41 (51%) UC samples and 64 out of 144 (43%) HC samples. ASVs assigned to the Klebsiella genus were identified in 2 UC samples (0.5%) and 11 HC samples (0.1%).

Indicator Taxa
We also performed 3 different discriminatory analyses to identify potential indicator taxa.

Mean relative abundances for significant taxa for the entire cohort were calculated by LefSe, determining differentially abundant taxa between HC and UC. Statistics are shown in the supplemental data. The most noteworthy results showed that the whole Verrucomicrobia phylum, which exclusively consist of the Akkarmansia genus, was significantly increased in HC compared with UC (Figure 3). Additionally, the LefSe plot and cladogram showed increased abundance of Coprococcus, Rikenellaceae RC9 gut group, Dialister, Veillonellaceae, and Lachnospiraceae ND3007 group in HC, and Ruminiclostridium was increased in UC.

DESeq2 differential abundance analysis was performed on the top 100 most abundant ASVs. One specific ASV affiliated with Prevotella_9 was significantly higher in HC compared with UC (P < .05) (Figure 4).

To capture indicator taxa that rely on a network, we next used a selbal compositional analysis approach, which identifies groups of taxa whose abundance and impact depend on each other, rather than comparing taxa abundance individually. Selbal analysis supports the differences highlighted by the LefSe plot and cladogram (Figure 3), with genus Coprococcus of the Lachnospiraceae family characterizing HCs and genus Ruminiclostridium of the Ruminococcaceae family characterizing UC patients (Supplementary Figure 9).

Discussion
In this descriptive cross-sectional study, we report the first fecal microbiota characterization of the Faroese IBD cohort.
To our knowledge, the composition of the gut microbiota and its impact on the overall health of the Faroese population has not previously been reported. This study was performed to gain appreciation of the architecture in gut microbiota within the Faroese background population and to investigate if the UC group, the predominate IBD phenotype on the Faroe Islands, was characterized by a clear dysbiosis.

The average UC patient in our study was 41 years old with a mean disease duration of 17 years. The results from the FFQ revealed that UC and HC participants consumed a typical westernized diet with starches, sugars, and processed meats consumed daily. The similarity between the groups may be due to well-established disease with patients not experiencing food as connected to their disease management. On the phylum level, high-starch, low-fiber diets have been associated with higher abundance of Firmicutes than Bacteroidetes. Our results are in line with those findings, and the overall characterization of the Faroese microbiota is similar to other Western European populations. Traditional Faroese foods such as whale and fermented meats were rarely consumed during the week prior to sample collection. Although it might be enlightening to follow the relationship between traditional Faroese foods and fecal microbiota over a longer timeframe, that is outside the scope of this study.

In this study, we did not find any significant differences in alpha or beta diversity between the UC and HC group. Reports of a decrease in alpha diversity in IBD patients and separate beta-diversity clustering has led to the assumption that this is the overall trend of HC and IBD. However, there is no consensus on what characterizes a healthy vs a UC microbiota profile. A systematic review determined that only

Figure 2. Stacked bar-chart showing relative abundance of the 10 overall most abundant taxonomical families in each participant at the family level. Data are presented as percent of total abundance.

Figure 3. A, Linear discriminant analysis effect size (LefSe) of significantly differentially abundant taxa markers in healthy controls (HCs) and ulcerative colitis patients (UC). The linear discriminant analysis score represents the influencing degree of the bacterial marker. B, Cladogram of significantly differentially abundant microbial taxa obtained using LefSe. Beginning in the middle of the cladogram, the taxa levels are shown for phylum, class, order, family, and genus of bacteria. White nodes mark taxa that were not significantly different between the 2 groups. Black nodes mark taxa with significantly different abundance between groups, HCs (grey), and UC (black).
Differential abundance and compositional analysis of indicator taxa revealed a decreased abundance of the genera *Prevotella* and *Coprococcus* in UC compared with HC, and an increased abundance of the genus *Ruminococcus*. This is consistent with indicator taxa for IBD found in previous studies. The identified indicator taxa are primarily beneficial commensal bacteria. *Coprococcus* and *Ruminococcus* belong to the Firmicutes phylum, which consists of anaerobic gram-positive bacteria with antiinflammatory properties. *Prevotella* is a genus in the Bacteroidetes phylum, which mainly includes commensal bacterium with antiinflammatory and health beneficial qualities. However, the role of *Prevotella* is debatable. Some studies find it to be beneficial, whereas others find it to be associated with disease, exacerbating intestinal inflammation.

Sorting the taxonomical tables based on presence/absence identified 42% UC samples and 23% of HC as having 0 reads of *Akkermansia* (Supplementary Figure 8), a species associated with mucosal renewal and intestinal barrier integrity that accounts for 1%-4% of the fecal microbiota in healthy cohorts. The lack of *Akkermansia* may be an indication of a shared predisposition or tendency toward metabolic and intestinal dysfunction in the general Faroese population. Specifically, the background population may have reduced mucus available to act as an energy source for *Akkermansia* and thus have a deficit in the ability to maintain a protective mucus barrier in the intestines. As shown by our LefSe analysis, the further reduction in *Akkermansia* associated with the UC group may be both partly causative and a consequence of the ongoing inflammatory process associated with UC.

In addition to *Akkermansia*, approximately 50% of the UC and HC samples contained ASVs assigned to potential pathobionts *Escherichia/Shigella*, commensal microbes that may act as pathogens in the context of environmental triggers in genetically susceptible hosts.

Taken together, the similarity in our dataset and the lack of *Akkermansia* raises the question whether intestinal inflammation is a hallmark that connects various immune-mediated and metabolic diseases within the general Faroese population through a genetic predisposition.

In addition to IBD, the Faroese population also holds the world’s highest incidence rates for several metabolic diseases: carnitine transport deficiency, glycogen storage disease type IIIA, and SUCLA2 deficiency. Exome sequencing has been performed by the Faroese Genome Project for all UC and HC participants included in this study and will be published as a companion study that includes analysis of disease phenotype and carrier frequency for these metabolic diseases.

In addition to the documented frequency of metabolic diseases, the self-reporting of immune-related diseases in families of both UC and HC groups underscores the need for further investigations into inflammatory, rheumatic, and dermatologic diseases, such as spondyloarthritis and psoriasis, for the Faroese population. Gut, skin, and joint inflammation is known to be interrelated with common genetic loci conferring disease susceptibility. Therefore, we expect the companion genetics study of our cohort to provide the first insights concerning if there is a shared genetic predisposition to metabolic and inflammatory diseases for the Faroese population.

Recent international studies have shown geographic variation, latitude, and ethnic group to be significant contributing factors to the composition of the gut microbiota. Because the Faroese population originates from a restricted and isolated genetic background in addition to a distinct geographic location, it is plausible that the influence of shared genetics and/or environment is a stronger determinant of bacterial composition than having a UC diagnosis.

Because this was a study with an exploratory nature, time of diagnosis was not the main priority and led to a study cohort where most of the included UC patients were not newly diagnosed. Following the successful establishment of the workflow and this initial cross-sectional study, a longitudinal study was initiated in 2021. The longitudinal study collects fecal samples and metadata from suspected IBD patients prior to first colonoscopy and follow them with repeated sampling if diagnosed with IBD. The compositions of the gut mycobiome, virome, and helminths may be investigated as well. These colonies exist in symbiosis with the bacteriome and affect the critical microbial functions in the gastrointestinal tract. Studies have shown that the compositions of these biomes have an impact on the development of IBD.

Additionally, variations in study design, lack of standardized methodology, and geographic location influence the results, which speaks to the need for reference groups limited to the Faroese community and disease-based experiments for high-risk populations, such as the Faroe Islands.

**Conclusion**

This study represents the first 16S rRNA amplicon sequencing data from the Faroese IBD and background population. The similarity in the gut microbiota from UC and HC samples within this study, as well as the absence of the beneficial taxon...
Akkermansia in both groups, is especially interesting. These results raise further questions concerning the underlying susceptibility toward inflammatory and metabolic disorders within this population.

To ensure compatibility between similar studies worldwide, a standardized methodology is necessary, because variations in study design, ethnicity, and geographic location impact final results. The need for limited Faroese reference groups and disease-based studies for high-risk populations, such as the Faroe Islands, is highlighted.

Supplementary Data
Supplementary data is available at Inflammatory Bowel Diseases online.

Acknowledgments
First and foremost, we extend our gratitude to the study participants. Thank you for your trust, contribution, and involvement in this project.

We would like to thank Iľegusavnið and the FarGen project. In particular, Katrin D. Apol for her coordination of the HCs and returned samples, Ólavur Mortensen, and Leivur Nattestad Lydersen for sharing broad expertise and helpful troubleshooting. Finally, Guðrið Andorsdóttir for her guidance and support.

Lastly, we would like to thank our colleagues at research Park iNOVA for helpful discussions and generosity.

Graphical abstract was created with BioRender.com (License number: ZC234F31GL).

Author Contributions
M.F.B.: Performed all lab investigations, formally analyzed and interpreted data, visualized data, and drafted the manuscript.

K.R.N.: Responsible clinician, conceptualized and designed the project, curated clinical metadata.

A.C.I.: Performed the bioinformatic analysis and interpreted the data, created data visualizations, drafted, and revised the manuscript.

J.M.: Clinical coordinator, curated clinical metadata and patient recruitment.

T.H.: Contributed to the analysis and interpretation of the metadata and critically revised the manuscript.

P.P.: Designed the food frequency questionnaire and contributed to the interpretation of the data.

N.M.O.V.: Designed the background questionnaires, visualized sampling protocol and acquired data.

N.O.G.: Responsible for the recruitment and metadata for the healthy control cohort, designed the background questionnaires.

J.B.: Conceptualized and designed the project, interpreted data, and critically revised the manuscript, supervised the project, and secured funding.

A.V.: Designed the project, acquired, analyzed, and interpreted data, and drafted and revised the manuscript, supervised the project, and secured funding.

Funding
This work was supported by the Novo Nordisk Fund (NNF-16OC0022278 awarded to J.B.) and the Faroese Research Council (0347 awarded to A.V.).

Conflict of Interest
Dr Burisch reports personal fees from AbbVie, personal fees from Janssen-Cilag, personal fees from Celgene, grants and personal fees from MSD, personal fees from Pfizer, grants and personal fees from Takeda, grants and personal fees from Tillots Pharma, personal fees from Samsung Bioepis, grants from Bristol Myers Squibb, grants from Novo Nordisk, and personal fees from Pharmacosmos outside the submitted work.

References
1. Clemente JC, Manasson J, Scher JU. The role of the gut microbiome in systemic inflammatory disease. BMJ. 2018;360:j5145.
2. Gordon H, Trier Moller F, Andersen V, Harbord M. Heritability in inflammatory bowel disease: from the first twin study to genomewide association studies. Inflamm Bowel Dis. 2015;21:1428-1434.
3. Hammer T, Langholz E. The epidemiology of inflammatory bowel disease: balance between East and West? A narrative review. Dig Med Res. 2020;3:48.
4. Hammer T, Nielsen KR, Munkholm P, Burisch J, Lynge E. The Faroese IBD Study: incidence of inflammatory bowel diseases across 5 years of population-based data. J Crohns Colitis. 2016;10:934-942.
5. Nielsen KR, Midjord J, Hammer T, Lophaven S, Burisch J. P607 The Faroese IBD Study: update on incidence from 2015-2020 and prevalence from 1960-2020. J Crohn’s Colitis. 2021;15(Supplement_1):S552-S553.
6. Burisch J, Pedersen N, Čuković-Čavka S, et al.; EpiCom-group. East-West gradient in the incidence of inflammatory bowel disease in Europe: the ECCO-EpiCom inception cohort. Gut. 2014;63:588-597.
7. Röin F, Röin J. Inflammatory bowel disease of the Faroe Islands, 1981-1988: a prospective epidemiologic study: primary report. Scand J Gastroenterol. 1989;24(sup170):44-46.
8. Nielsen KR, Olsen J, Andorsdóttir G, Gislason E, Burisch J. P665 Familial aggregation of inflammatory bowel disease on the Faroe Islands: a Faroese inflammatory bowel disease cohort study. Eur J Crohn’s Colitis Organ. 2016;10:S440-S441.
9. Hammer T, Lophaven SN, Nielsen KR, et al. Inflammatory bowel diseases in Faroese-born Danish residents and their offspring: further evidence of the dominant role of environmental factors in IBD development. Aliment Pharmacol Ther. 2017;45:1107-1114.
10. Ni J, Wu GD, Albenberg L, Tomov VT. Gut microbiota and IBD: causation or correlation? Nat Rev Gastroenterol Hepatol. 2017;14:573-584.
11. Pittayanon R, Lau JT, Leontiadis GI, et al. Differences in gut microbiota in patients with vs without inflammatory bowel diseases: a systematic review. Gastroenterology. 2020;158:930-946.e1.
12. Gorovitovskaia A, Holmes SP, Huse SM. Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. Microbiome. 2016;4:15.
13. Almonacid DE, Kraal L, Ossandon FJ, et al. Correction: 16S rRNA gene sequencing and healthy reference ranges for 28 clinically relevant microbial taxa from the human gut microbiome. PLoS One. 2019;14:e0212474.
14. He Y, Wu W, Zheng HM, et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nat Med. 2018;24:1532-1535.
15. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut. 2006;55:749-753.
16. Langholz E. Ulcerative colitis. An epidemiological study based on a regional inception cohort, with special reference to disease course and prognosis. Dan Med Bull. 1999;46:400-415.
17. Panek M, Ćipčić Paljetak H, Barešić A, et al. Methodology challenges in studying human gut microbiota - effects of collection, storage, DNA extraction and next generation sequencing technologies. *Sci Rep.* 2018;8:5143.

18. Klindworth A, Pruesse E, Schweer T, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 2013;41:e1.

19. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome.* 2018;6:226.

20. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2020. Scientific Research Publishing. ProMED-mail. Accessed May 30, 2021. https://www.r-project.org/

21. McMurdie PJ, Holmes S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One.* 2013;8:e61217.

22. Wickham H. *Ggplot2.* 2nd ed. New York: Springer International Publishing; 2016. doi:10.1007/978-3-319-24277-4

23. Foster ZS, Sharpton TJ, Grünwald NJ. Metacoder: an R package for visualization and manipulation of community taxonomic diversity data. *PLoS Comput Biol.* 2017;13:e1005404.

24. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.

25. Deschasaux M, Bouter KE, Prodan A, et al. Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nat Med.* 2018;24:1526-1531.

26. Zuo T, Kamm MA, Colombel JF, Ng SC. Urbanization and the gut microbiota in health and inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol.* 2018;15:440-452.