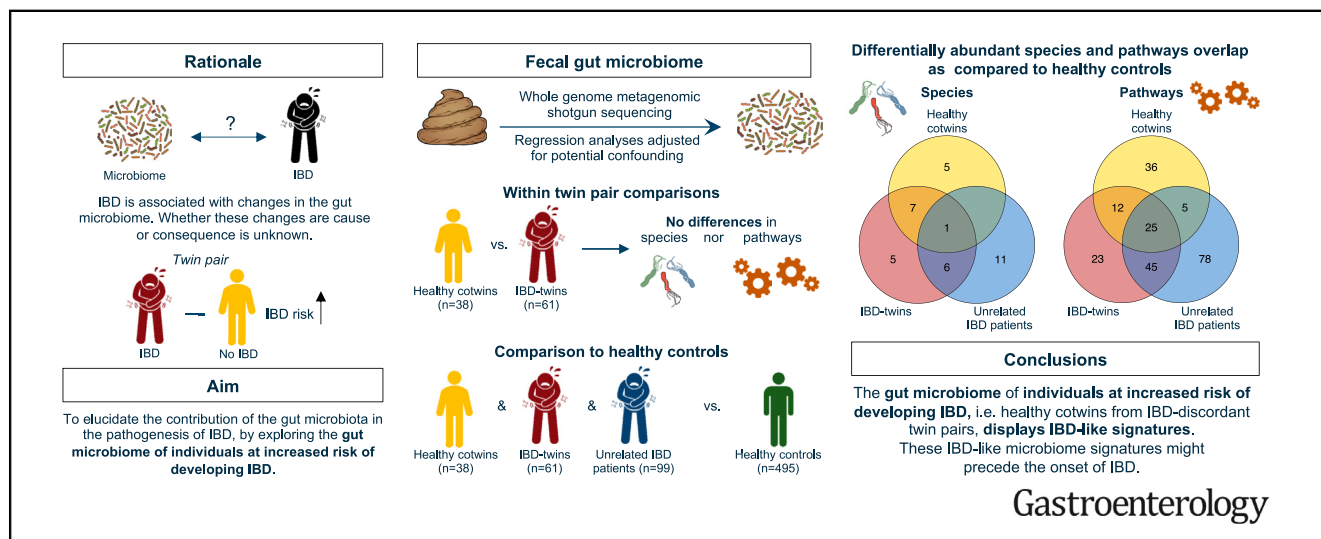


# Healthy Cotwins Share Gut Microbiome Signatures With Their Inflammatory Bowel Disease Twins and Unrelated Patients



Eelco C. Brand,<sup>1,2,\*</sup> Marjolein A. Y. Klaassen,<sup>3,4,\*</sup> Ranko Gacesa,<sup>3,4</sup> Arnau Vich Vila,<sup>3,4</sup> Hiren Ghosh,<sup>3,4</sup> Marcel R. de Zoete,<sup>5</sup> Dorret I. Boomsma,<sup>6,7</sup> Frank Hoentjen,<sup>8</sup> Carmen S. Horjus Talabur Horje,<sup>9</sup> Paul C. van de Meeberg,<sup>10</sup> Gonneke Willemsen,<sup>6,7</sup> Jingyuan Fu,<sup>4,11</sup> Cisca Wijmenga,<sup>4</sup> Femke van Wijk,<sup>2</sup> Alexandra Zhernakova,<sup>4</sup> Bas Oldenburg,<sup>1,§</sup> and Rinse K. Weersma,<sup>3,4,§</sup> on behalf of the Dutch TWIN-IBD consortium and the Dutch Initiative on Crohn and Colitis

<sup>1</sup>Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; <sup>2</sup>Center for Translational Immunology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; <sup>3</sup>Department of Gastroenterology and Hepatology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; <sup>4</sup>Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands; <sup>5</sup>Department of Medical Microbiology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; <sup>6</sup>Department of Biological Psychology, Behavioral and Movement Sciences, Vrije Universiteit, Amsterdam, the Netherlands; <sup>7</sup>Amsterdam Public Health Research Institute, Vrije Universiteit Medical Center, Amsterdam, the Netherlands; <sup>8</sup>Department of Gastroenterology and Hepatology, Radboud University Medical Center, Nijmegen, the Netherlands; <sup>9</sup>Crohn & Colitis Center Rijnstate, Department of Gastroenterology and Hepatology, Rijnstate Hospital, Arnhem, the Netherlands; <sup>10</sup>Department of Gastroenterology & Hepatology, Slingeland Hospital, Doetinchem, the Netherlands; and <sup>11</sup>Department of Pediatrics, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands



## See Covering the Cover synopsis on page 1910.

**BACKGROUND & AIMS:** It is currently unclear whether reported changes in the gut microbiome are *cause* or *consequence* of inflammatory bowel disease (IBD). Therefore, we studied the gut microbiome of IBD-discordant and -concordant twin pairs, which offers the unique opportunity to assess individuals at increased risk of developing IBD, namely healthy cotwins from IBD-discordant twin pairs. **METHODS:** Fecal samples were obtained from 99 twins (belonging to 51 twin pairs), 495 healthy age-, sex-, and body mass index-matched controls, and 99 unrelated patients with IBD. Whole-genome metagenomic shotgun sequencing was performed. Taxonomic and functional (pathways) composition was compared

among healthy cotwins, IBD-twins, unrelated patients with IBD, and healthy controls with multivariable (ie, adjusted for potential confounding) generalized linear models. **RESULTS:** No significant differences were observed in the relative abundance of species and pathways between healthy cotwins and their IBD-twins (false discovery rate <0.10). Compared with healthy controls, 13, 19, and 18 species, and 78, 105, and 153 pathways were found to be differentially abundant in healthy cotwins, IBD-twins, and unrelated patients with IBD, respectively (false discovery rate <0.10). Of these, 8 (42.1%) of 19 and 1 (5.6%) of 18 species, and 37 (35.2%) of 105 and 30 (19.6%) of 153 pathways overlapped between healthy cotwins and IBD-twins, and healthy cotwins and unrelated patients with IBD, respectively. Many of the shared species and pathways have previously been associated with IBD.

The shared pathways include potentially inflammation-related pathways, for example, an increase in propionate degradation and L-arginine degradation pathways. **CONCLUSIONS:** The gut microbiome of healthy cotwins from IBD-discordant twin pairs displays IBD-like signatures. These IBD-like microbiome signatures might precede the onset of IBD. However, longitudinal follow-up studies are needed to infer a causal relationship.

**Keywords:** Preclinical; Prediagnostic; Crohn's Disease; Ulcerative Colitis; Family Studies; Microbiota; Prediction; Discordant Twin Design.

Over the past decade, a consistent body of evidence has accumulated that supports the association between the gut microbiota and inflammatory bowel disease (IBD).<sup>1,2</sup> Differences have been observed in the microbiome of Crohn's disease (CD) and ulcerative colitis (UC) patients compared with individuals not affected by IBD in a large number of studies, most of which were based on 16S ribosomal RNA gene (16S rRNA) sequencing.<sup>2,3</sup> It is, however, currently unclear whether these changes are a cause or a consequence of intestinal inflammation, because longitudinal microbiome studies are scarce and data from the pre-diagnostic phase of IBD are lacking. Moreover, the large interindividual heterogeneity induced by the effect of (childhood) environmental factors<sup>4-7</sup> and host genetics<sup>6,8,9</sup> on the gut microbiome hampers the elucidation of the exact contribution of the gut microbiome to the onset of IBD.

In the context of IBD, healthy cotwins with an IBD-affected twin are at increased risk of developing IBD, with reported concordance rates among monozygotic twin pairs as high as 64% for CD and 28% for UC.<sup>10</sup> Studying the gut microbiome in IBD-discordant and concordant twin pairs can thus provide important insights in its role in the pathogenesis of IBD, also because of the shared (childhood) environmental and genetic factors between twins from the same twin pair.

Five previous studies, aiming to explore the gut microbiome in IBD-affected twin pairs, using fecal<sup>11-13</sup> and mucosal biopsy<sup>11,14,15</sup> samples, reported differences in the gut microbiome composition in IBD-affected twins compared with their healthy cotwin. These studies were, however, performed in small numbers of IBD-discordant or -concordant twin pairs (10 or fewer),<sup>12-15</sup> did not include an unrelated matched healthy control group,<sup>11,15</sup> or only a small nonmatched healthy control group,<sup>12-14</sup> and were based on 16S rRNA sequencing,<sup>11,13-15</sup> which does not allow for prediction of microbial functional pathways.

The goal of the present study was to elucidate the contribution of the gut microbiota in the risk of developing IBD, by exploring the gut microbiome of healthy cotwins - at increased risk of developing IBD - in comparison with the gut microbiome of their IBD-twins, unrelated patients with IBD, and those of unrelated healthy controls.

## Material and Methods

### Study Population

Participants from 2 cohorts were included in the present cross-sectional study.

### WHAT YOU NEED TO KNOW

#### BACKGROUND AND CONTEXT

Changes in the composition of the gut microbiome are associated with inflammatory bowel disease (IBD), but it is presently unclear whether these changes are cause or consequence of IBD.

#### NEW FINDINGS

The gut microbiome composition of individuals at increased risk of developing IBD (i.e. healthy cotwins from IBD-discordant twin pairs) displays IBD-like signatures on a species and pathway level.

#### LIMITATIONS

This is a cross-sectional study. Future follow-up studies will help to identify those cotwins who will develop IBD in the long run, thereby allowing confirmation and further in-depth analyses of our findings.

#### IMPACT


The overlap in gut microbial features between healthy cotwins at increased risk of developing IBD and related and unrelated IBD patients suggests that these IBD-like microbiome signatures might precede the onset of IBD. This potentially opens new avenues for diagnosis and therapy in individuals with pre-symptomatic IBD.

**(1) IBD-twins and healthy cotwins: "TWIN-study".** Twin pairs were included from the ongoing prospective longitudinal Dutch "Twin cohort for the study of (pre) clinical inflammatory bowel disease in the Netherlands" (TWIN) study (Netherlands Trial Register: NTR6681). Twin pairs,  $\geq 16$  years of age, either IBD-discordant or -concordant, from the Netherlands, were recruited via their treating physician, through awareness for the study on (social) media, or via the Netherlands Twin Register (NTR).<sup>16</sup> Potential candidates in the NTR were identified through previously administered questionnaires or those who gave consent for record linkage were identified by linking the NTR-database with an IBD-directed search in the nationwide network and registry of histo- and cytopathology in the Netherlands (PALGA).<sup>17</sup>

Participants were followed longitudinally and evaluated at intervals of 6 months at the University Medical Center (UMC) Utrecht. Samples of blood, urine, feces, oropharyngeal swabs, and (depending on consent) rectal biopsies were collected

\* Authors share co-first authorship; § Authors share co-senior authorship.

**Abbreviations used in this paper:** 16S rRNA, 16S ribosomal RNA; BMI, body mass index; CD, Crohn's disease; FDR, false discovery rate; IBD, inflammatory bowel disease; IBDU, inflammatory bowel disease unclassified; NTR, Netherlands Twin Register; PALGA, Nationwide network and registry of histo- and cytopathology in the Netherlands; PCoA, principal coordinate analyses; PERMANOVA, permutational multivariate analysis of variance; PPI, proton pump inhibitor; SCFA, short chain fatty acid; TWIN-study, twin cohort for the study of (pre)clinical inflammatory bowel disease in the Netherlands study; UC, ulcerative colitis; UMC, University Medical Center.

 Most current article

© 2021 by the AGA Institute. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

0016-5085

<https://doi.org/10.1053/j.gastro.2021.01.030>

during study visits. In addition, questionnaires were used to obtain information on demographics, family composition, diet, general health, IBD-specific characteristics, medication use, quality of life, and environmental factors.

For the present cross-sectional study, we analyzed 1 fecal sample per participant, collected between October 2017 and August 2019. Fecal samples of a twin pair were not included for whole-metagenome sequencing if only 1 twin from a pair had provided a fecal sample or when a twin had an ostomy or pouch. If the whole-metagenome shotgun sequencing results of a twin were excluded after sequencing, we kept the results of their cotwin in the data analyses. All fecal samples were collected within the same week per twin pair, except for 1 twin pair (6 months between feces collection). Both IBD-discordant (ie, one of the twins of a pair is affected by IBD) and IBD-concordant (ie, both twins of the twin pair are affected by IBD regardless of the IBD subtype) twin pairs are included in the present study. No multiples consisting of more than 2 individuals were included. Individuals from IBD-discordant twin pairs who are not diagnosed with IBD are referred to as “healthy cotwins” and individuals from IBD-discordant or IBD-concordant twin pairs who are diagnosed with IBD are referred to as “IBD-twins.”

**(2) Healthy controls and unrelated patients with IBD: “Dutch Microbiome Project”.** The Dutch Microbiome Project is part of *LifeLines*,<sup>18</sup> a large-scale population-based cohort that prospectively included 167,729 individuals between 8 and 84 years of age and their families, who all have been resident in the 3 Northern provinces of the Netherlands. In 2015, 10,000 *LifeLines* participants were asked to collect a fecal sample as part of the Dutch Microbiome Project. Of these, ~9700 individuals collected 1 fecal sample between 2015 and 2018. Demographics, medical information, and medication use were collected via surveys. For the present cross-sectional study, we included age-, sex- and BMI-matched healthy participants from the Dutch Microbiome Project. Healthy individuals were matched to each twin individual (ie, healthy cotwins and IBD-twins) with a healthy control:twin ratio of 5:1. Matching was performed with the MatchIt package in R using the method “optimal”.

Within the Dutch Microbiome Project, 99 participants had a self-reported diagnosis of IBD. These individuals were included in this study as “unrelated patients with IBD.”

### Ethical Considerations

The TWIN-study and Dutch Microbiome Project study were conducted in accordance with the declaration of Helsinki and the Dutch Medical Research Involving Human Subjects Act. All participants provided informed consent. The TWIN-study was approved by the medical ethics committee of the UMC Utrecht. The Dutch Microbiome Project was approved by the medical ethics committee of the UMC Groningen. The PALGA search was approved by the PALGA Privacy Commission and Scientific Council. Linkage took place only for NTR participants who had provided approval for linkage to external registers.

### Participant Characteristics

**IBD characteristics.** For the TWIN cohort, diagnoses and phenotypes were verified by review of medical records. For participants in the Dutch Microbiome Project, diagnoses and phenotypes were self-reported. IBD unclassified (IBDU) patients were grouped with UC patients in the analyses. In the

TWIN-study signs of active disease was defined as a patient Harvey Bradshaw Index<sup>19</sup> >4 in case of CD, and a patient Short Clinical Colitis Activity Index<sup>20</sup> >4 in case of UC or IBDU, or endoscopic signs of inflammation as noted during proctoscopy in the TWIN-study or endoscopy for clinical care. If no signs of inflammation were found on proctoscopy in patients with UC, patients were classified as having quiescent disease. In the Dutch Microbiome Project, the Montreal classification, date of diagnosis, symptoms, and endoscopic data were not available.

**Twin-specific characteristics.** The zygosity of twins was based on self-reported zygosity or, in case of doubt, on a zygosity questionnaire, based on childhood similarity between twins. This questionnaire was developed by the NTR and has been shown to have a 95.9% accuracy in predicting DNA zygosity.<sup>16</sup> Cohabitation was based on survey data.

**Demographics.** In both the TWIN-study and the Dutch Microbiome Project, data on sex, age, body mass index (BMI), smoking behavior, history of appendectomy, and history of bowel resections were collected by surveys at the moment of feces collection.

**Medication use.** Medication use, including antibiotics (in the past 3 months), current proton pump inhibitors (PPIs), and IBD medication, was determined by targeted medication surveys in the TWIN-study and the Dutch Microbiome Project.

### Stool Sampling and Analysis of the Gut Microbiome

**Fecal sample collection.** In the TWIN-study, fecal samples were kept at room temperature and transported to the research facility by the participants within 31 hours of collection, and subsequently stored at  $-80^{\circ}\text{C}$  (median time at room temperature until  $-80^{\circ}\text{C}$ : 9.7 hours; Q1-Q3: 4.8–19.8). In the Dutch Microbiome Project, fresh fecal samples were frozen at the participants' homes in a standardized manner at  $-20^{\circ}\text{C}$ . Subsequently, frozen fecal samples were shipped on dry ice and stored at  $-80^{\circ}\text{C}$  on arrival. Fecal samples from both cohorts remained frozen at  $-80^{\circ}\text{C}$  until DNA extraction.

**Microbial DNA extraction and sequencing.** Microbial genomic DNA was isolated via the QIAamp Fast DNA Stool Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions by the same research technician in the same time period for both cohorts. The QIAcube (Qiagen) automated sample preparation system was used for DNA isolation. Whole-genome shotgun metagenomic sequencing was performed at Novogene (HK) Company Limited (Wan Chai, Hong Kong) on the Illumina (San Diego, CA) HiSeq 2000 platform to generate approximately 8 Gb of 150 base pairs, paired-end, reads per sample (mean 7.9 gigabytes, standard deviation 1.2 gigabytes, median 14.4 million reads).

**Quality control and determination of microbiome parameters.** Samples with a read depth below 10 million reads were removed. From the raw metagenomic reads, the Illumina adapters were removed using kneadData (v0.5.1) toolkit, and reads were trimmed to PHRED quality 30. Trimmed reads that aligned to the human genome (GRCh37/hg19) were removed using kneadData integrated Bowtie2 tool (version 2.3.4.1), and after this, the quality of the metagenomes was tested using the FastQC toolkit (version 0.11.7). Taxonomic composition of the metagenomes was profiled by the MetaPhlAn2 tool (version 2.7.2) using the MetaPhlAn database of marker genes (version mpa v20 m200). In addition, profiling of



genes encoding microbial biochemical pathways (ie, the functional potential of the gut microbiome) was performed using the HUMAnN2 pipeline (version 0.11.1) integrated with the DIAMOND alignment tool (version 0.8.22), uniref90 protein database (version 0.1.1), and the ChocoPhlAn pangenome database (version 0.1.1).

Non-bacteria were filtered out and the taxonomic species level was maintained. After these steps, 586 species and 576 pathways were identified across samples. For the diversity and dissimilarity analyses, no further filtering steps were used. For the regression analyses, only species and pathways that were prevalent in  $\geq 10\%$  of samples were included.

### Design of Data Analysis

To answer our question of to which extent the gut microbiome of healthy cotwins displays a signature of an IBD-like microbiome, we performed our analyses in 2 steps. First, we compared the healthy cotwins with their IBD-twins (within-twin pairs comparison). Second, we compared the gut microbiome of healthy cotwins, IBD-twins, and unrelated patients with IBD to the microbiome of healthy controls. By means of these comparisons, we aimed to identify whether the microbiomes of healthy cotwins display a healthy signature (ie, is similar to the gut microbiome of unrelated healthy controls) or already displays an IBD signature (ie, is more similar to the microbiome of their cotwins and unrelated patients with IBD).

We analyzed the gut microbiome at the following levels: (1)  $\alpha$ - and  $\beta$ -diversity; (2) (dis)similarity in gut microbiome composition among individuals based on IBD concordance, IBD phenotype, zygosity, and cohabitation; and (3) differential relative abundances of individual species and pathways including assessment of overlap in identified differentially (compared with healthy controls) abundant species and pathways among the healthy cotwins, IBD-twins, and unrelated patients with IBD.

### Statistical Analyses

All statistical analyses were performed in R version 4.0.0 for macOS<sup>21</sup> (analyses scripts can be found at: [https://github.com/WeersmaLabIBD/Microbiome/blob/master/IBD\\_Twins\\_Microbiome\\_Utrecht\\_Groningen.md](https://github.com/WeersmaLabIBD/Microbiome/blob/master/IBD_Twins_Microbiome_Utrecht_Groningen.md)). Baseline characteristics are shown as numbers and proportions for categorical variables, and as means and standard deviations for continuous variables.

**Microbiome diversity measures.** We estimated diversity measures for both for the taxonomic (ie, species) and functional (ie, pathways) composition. The  $\alpha$ -diversity is expressed as the Shannon Index for species and gene richness for pathways. Differences between groups were tested with the Mann-Whitney *U* test. The  $\beta$ -diversity was calculated as Bray-Curtis distance and is visually shown using principal coordinate analyses (PCoA) plots, showing PCoA1 and PCoA2 per plot. The proportion of explained variance ( $R^2$ ) in the  $\beta$ -diversity was assessed based on permutational multivariate analysis of variance (PERMANOVA) using distance matrices implemented as *adonis* function in *vegan* package for R.

For the within-twin pair comparisons, we estimated the proportion of explained variance ( $R^2$ ) in the  $\beta$ -diversity adjusting for twin pairs, thereby taking the clustering of twins within their twin pair into account, and stratified for zygosity for the

following variables separately: IBD phenotype (no IBD, CD, or UC), IBD-activity, belonging to the same twin pair, disease location, history of bowel resection, age, sex, BMI, antibiotics use, PPI use, and read depth. For the comparison of healthy cotwins, IBD-twins, and unrelated patients with IBD with healthy controls, we estimated the explained variance in  $\beta$ -diversity comparing IBD-twins, healthy cotwins, healthy controls, and unrelated patients with IBD, and for the variables IBD phenotype, sex, age, BMI, antibiotics use, PPI use, and read depth.

**Similarity in gut microbiome composition.** To test whether the gut microbiome composition was similar between pairs of individuals on a taxonomic or functional level, the pairwise Bray-Curtis dissimilarity was calculated, ranging from 0 to 1 with lower scores indicating more similarity in the microbiome composition. We assessed the associations between dissimilarity and 4 variables: IBD concordance (ie, being discordant or concordant for IBD), IBD phenotype, zygosity, and cohabitation. To this end, dissimilarities between twins from the same twin pair were calculated and compared with the dissimilarities of 1000 random pairs of unrelated healthy controls. IBD phenotype pairs were logically formed between unrelated IBD-twins. For zygosity, we also compared the dissimilarity with random pairs of unrelated twins (ie, pairs of twins coming from 2 different twin pairs).

**Differentially abundant taxa and pathways.** We assessed whether taxa and pathways were differentially abundant by using multivariable general linear regression models (GLM) of the Gaussian family, implemented in the GLM function in R, using the *MaAsLin2* package.<sup>22</sup> The microbiome taxa and pathway relative abundances that serve as outcomes in the GLMs were arcsine square-root transformed. Species and pathways were included in the analyses only if they were present in both cohorts and were prevalent in at least 10% of the samples.

For the comparison between IBD-twins and healthy cotwins, we used twin pair as a random effect to take within-twin pair effects (ie, clustering) into account. Furthermore, the regression analyses with 119 species and 343 pathways as outcomes were adjusted for IBD phenotype (ie, CD, UC, or no IBD), disease location, age, sex, BMI, zygosity, antibiotics use in the past 3 months, current PPI use, and sequence read depth.

Second, we assessed differential abundance of taxa and pathways comparing healthy cotwins, IBD-twins, and unrelated patients with IBD with healthy controls in one model. These regression analyses with 116 individual species and 330 individual pathways as outcomes were adjusted for potentially confounding factors: IBD phenotype (ie, CD, UC, or no IBD), age, sex, BMI, antibiotics use in the past 3 months, current PPI use, and sequence read depth. Next, we assessed if overlap existed in the identified differentially abundant taxa and pathways as present in the microbiomes of healthy cotwins, IBD-twins, and unrelated patients with IBD compared with healthy controls. Last, we compared the relative abundances of 112 species and 339 pathways between IBD-twins and unrelated patients with IBD directly adjusting for the same potentially confounding factors as in the comparison with healthy controls.

**Multiple testing correction.** To correct for multiple testing, the Benjamini-Hochberg correction was applied to calculate the false discovery rate (FDR). An FDR  $< 0.10$  was regarded as statistically significant. This significance threshold

Table 1. Baseline and Sample Characteristics

	The TWIN-study		The Dutch Microbiome Project	
	Healthy cotwins (n = 38)	IBD-twins (n = 61)	Healthy controls (n = 495)	Unrelated IBD patients (n = 99)
Demographics and clinical characteristics				
Female sex	25 (65.8)	39 (63.9)	312 (63.0)	59 (59.6)
Age (y)	44.7 (15.4)	40.1 (15.4)	42.3 (15.0)	52.8 (10.0)
BMI (kg/m <sup>2</sup> )	24.9 (3.8)	24.2 (3.3)	24.4 (4.0)	26.0 (3.9)
Current smoker	6 (15.8)	10 (16.4)	49 (9.9)	14 (14.1)
History of appendectomy	4 (10.5)	6 (9.8)	29 (5.9)	11 (11.1)
History of bowel resection	0	10 (16.4)	NA <sup>e</sup>	NA <sup>e</sup>
Current PPI use	6 (15.8)	9 (14.8)	24 (4.8)	12 (12.1)
Antibiotic use in past 3 mo	3 (7.9)	8 (13.1)	23 (4.6)	10 (10.1)
Twin pair characteristics <sup>a</sup>				
Zygoty				
Monozygotic	16 (42.1)	37 (60.7)	NA	NA
Dizygotic	22 (57.9)	24 (39.3)	NA	NA
Concordance for IBD				
Concordant for IBD	0	24 (39.3)	NA	NA
Discordant for IBD	38 (100)	37 (60.7)	NA	NA
Cohabitation with cotwin during sampling	4 (10.5)	10 (16.4)	NA	NA
IBD characteristics				
IBD phenotype				
CD	—	33 (54.1)	—	28 (28.3)
UC	—	26 (42.6)	—	72 (72.7)
IBD unclassified	—	2 (3.3)	—	0
No IBD	38 (100)	—	495 (100)	—
IBD duration (mo)	NA	139 (112)	NA	NA <sup>e</sup>
Signs of active disease <sup>b</sup>	NA	20 (32.8)	NA	NA <sup>e</sup>
Age of diagnosis CD (Montreal classification) <sup>c</sup>				
A1 (≤16 y)	NA	3 (9.1)	NA	NA <sup>e</sup>
A2 (17–39 y)	NA	26 (78.8)	NA	NA <sup>e</sup>
A3 (≥40 y)	NA	4 (12.1)	NA	NA <sup>e</sup>
Location CD (Montreal classification) <sup>c</sup>				
L1 (ileum only)	NA	13 (39.4)	NA	NA <sup>e</sup>
L2 (colon only)	NA	7 (21.2)	NA	NA <sup>e</sup>
L3 (ileocolonic)	NA	13 (39.4)	NA	NA <sup>e</sup>
L4 (proximal of ileum)	NA	4 (12.1)	NA	NA <sup>e</sup>
Behavior CD (Montreal classification) <sup>c</sup>				
B1 (nonstricturing, nonpenetrating)	NA	18 (54.5)	NA	NA <sup>e</sup>
B2 (stricturing)	NA	11 (33.3)	NA	NA <sup>e</sup>
B3 (penetrating)	NA	4 (12.1)	NA	NA <sup>e</sup>
P (perianal modifier)	NA	3 (9.1)	NA	NA <sup>e</sup>
Location UC (Montreal classification) <sup>c</sup>				
E1 (ulcerative proctitis)	NA	6 (23.1)	NA	NA <sup>e</sup>
E2 (left-sided UC)	NA	7 (26.9)	NA	NA <sup>e</sup>
E3 (extensive UC)	NA	11 (42.3)	NA	NA <sup>e</sup>
Missing	NA	2 (7.7)	NA	NA <sup>e</sup>
Current IBD-medication use				
No IBD-medication	38 (100)	10 (16.4)	495 (100)	NA <sup>e</sup>
5-aminosalicylic acid	—	20 (32.8)	—	NA <sup>e</sup>

Table 1. Continued

	The TWIN-study		The Dutch Microbiome Project	
Corticosteroids	NA	3 (4.9)	NA	NA <sup>e</sup>
Methotrexate	NA	1 (1.6)	NA	NA <sup>e</sup>
Ciclosporin	NA	0	NA	NA <sup>e</sup>
Thiopurine	NA	24 (39.3)	NA	NA <sup>e</sup>
Anti-tumor necrosis factor- $\alpha$	NA	12 (19.7)	NA	NA <sup>e</sup>
Vedolizumab (anti-integrin $\alpha 4\beta 7$ )	NA	2 (3.3)	NA	NA <sup>e</sup>
Ustekinumab (anti-interleukin12/23)	NA	0	NA	NA <sup>e</sup>
Tofacitinib (pan-janus kinase inhibitor)	NA	0	NA	NA <sup>e</sup>
<b>Fecal sample and sequence characteristics</b>				
Bristol stool scale <sup>d</sup>				
Type 1 or 2	7 (18.4)	10 (16.4)	47 (9.5)	6 (6.1)
Type 3 or 4	26 (68.4)	23 (37.7)	349 (70.5)	53 (53.5)
Type 5, 6, or 7	5 (13.2)	28 (45.9)	68 (13.7)	34 (34.3)
Missing	0	0	31 (6.3)	6 (6.1)
Number of bowel movements per wk	7.4 (2.71)	12.5 (11.5)	8.3 (4.17) Missing: n = 31	11.0 (7.4) Missing: n = 6
Number of sequence reads	16,108,523 (4,134,472)	15,271,177 (3,989,233)	24,057,674 (3,552,760)	25,111,866 (4,072,961)

NOTE. Continuous variables are depicted as mean (standard deviation), and categorical variables as number (proportion) unless indicated otherwise.

n, number of participants; NA, not applicable.

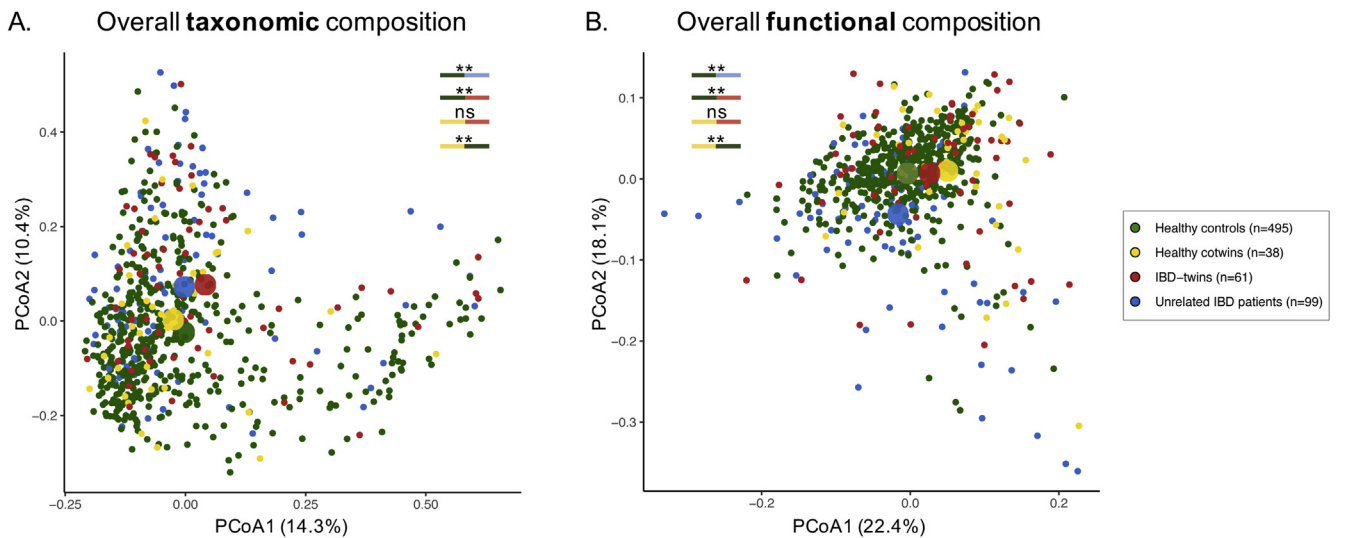
<sup>a</sup>All twin-related characteristics are here depicted per individual, but logically apply to both of the twin pair.

<sup>b</sup>Signs of active disease was defined as Harvey Bradshaw Index<sup>19</sup> >4 in case of CD, and a Short Clinical Colitis Activity Index<sup>20</sup> >4 in case of UC or IBD unclassified, or endoscopic signs of inflammation as noted during either proctoscopy performed as part of the TWIN-study protocol or during a colonoscopy which was performed as part of clinical care.

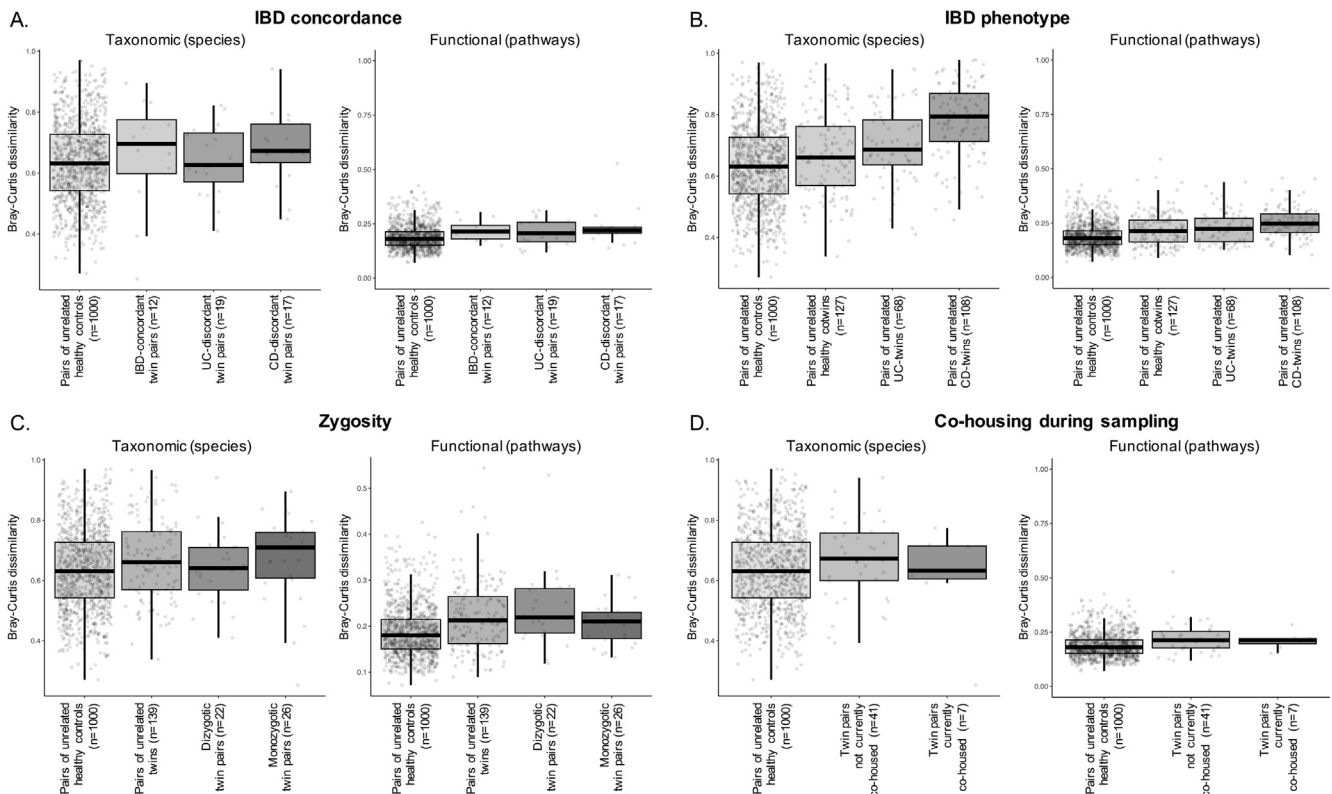
<sup>c</sup>All proportions for Montreal classification characteristics are only calculated for those participants to whom it applies.

<sup>d</sup>The Bristol stool scale ranges from 1 to 9, with 1 being separate hard lumps, and 9 entirely liquid consistency without solid pieces. In the healthy controls and unrelated patients with IBD this reflects the mean over a week.

<sup>e</sup>These variables are not assessed in the Dutch Microbiome Project.



**Figure 1.** No differences were detected in the gut microbiome composition (beta-diversity) between healthy cotwins and IBD-twins, but both gut microbiomes differ from healthy controls. PCoA plots for the (A) taxonomic (species) and (B) functional (pathways) composition of the gut microbiome. Each small dot represents 1 fecal sample. The larger centroids depict the center per group (ie, healthy cotwins, IBD-twins, healthy controls, and unrelated patients with IBD). On the x- and y-axes the first and second principal coordinate and proportion of explained variance are displayed. Based on the FDR from univariable PERMANOVA analyses performed on the whole Bray-Curtis distance matrix (Supplementary Material 4), the healthy cotwins and IBD-twins are not significantly different from each other on a taxonomic and functional level, whereas the gut microbiome composition was statistically significantly different comparing healthy cotwins, IBD-twins, and unrelated patients with IBD with healthy controls. \*\*, FDR < 0.05; ns, FDR > 0.10.



**Figure 2.** The gut microbiome composition of random pairs of healthy individuals is not more dissimilar than of twin pairs stratified for IBD concordance, zygosity or co-housing. Box plots show Bray-Curtis dissimilarity (lower values reflect more similarity) for random pairs of unrelated healthy controls and/or unrelated patients with IBD, random unrelated pairs of twins, and true twin pairs. Each dot represents 1 pair. The left panel per subfigure shows the dissimilarity for the relative abundance of species and the right panel for pathways. (A) The dissimilarity of IBD-concordant and IBD-discordant twin pairs is comparable to that of random pairs of healthy controls. (B) Dissimilarity increases from random pairs of healthy controls, random unrelated pairs of healthy cotwins, random unrelated pairs of UC-twins to random unrelated pairs of CD-twins, probably reflecting the heterogeneity of microbiome changes associated with IBD. (C) and (D) show that the gut microbiomes of true twin pairs, stratified for zygosity and co-housing during sampling, are not more similar than the gut microbiomes of random pairs of unrelated healthy controls or random pairs of unrelated twins.

was chosen because of the explorative setup of our study, to capture both common associations and those with low effect sizes.

## Results

### Participant and Sample Characteristics

In total, 99 twins from 51 twin pairs (61/99 IBD-twins, and 38/99 healthy cotwins) from the TWIN-study, and 495 age-, sex-, and BMI-matched unrelated healthy controls and 99 unrelated patients with IBD from the Dutch Microbiome Project were included in the present study (Table 1 and Supplementary Materials 1 and 2). Fifty-three of the 99 twins were part of a monozygotic twin pair and 46 were part of a dizygotic twin pair. Twelve (23.5%) of the 51 twin pairs were IBD-concordant and 39 (76.5%) IBD-discordant. Seven (13.7%) of the 51 twin pairs were living together at the time of fecal sampling (Table 1 and Supplementary Material 2).

Healthy controls from the Dutch Microbiome Project had a comparable age, sex, and BMI, but smoked less frequently

as compared with all twins, whereas the unrelated patients with IBD were older and had a higher BMI compared with the TWIN cohort. A history of appendectomy and antibiotic or PPI use was encountered more frequently in twins and unrelated patients with IBD than in healthy controls. IBD-twins were compared with the unrelated patients with IBD more often diagnosed with CD (54.1% vs 28.3%, Table 1). The sequence read depth was higher in healthy controls and unrelated patients with IBD as compared with the TWIN participants (Table 1), the linear regression analyses were therefore adjusted for sequence read depth.

### Healthy Cotwins and IBD-Twins Are Alike in Microbiome Diversity and Differ From Healthy Controls

There were no significant differences in  $\alpha$ -diversity, that is, the per-participant diversity, on taxonomic and functional level between the healthy cotwins and the IBD-twins (FDR = 0.14, FDR = 0.62, respectively, Supplementary Material 3). Furthermore, no statistically significant differences in the Shannon Index were observed between healthy



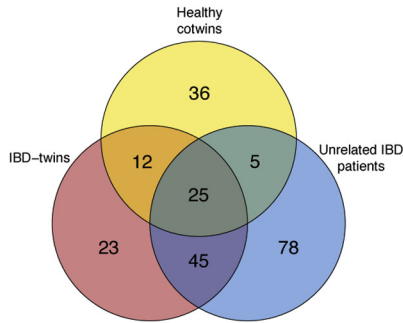
cotwins, IBD-twins, unrelated patients with IBD, and healthy controls. The gene richness was statistically significantly higher in healthy cotwins compared with healthy controls,

and IBD-twins compared with healthy controls (FDR = 0.007, FDR = 0.007, respectively, [Supplementary Material 3](#)).

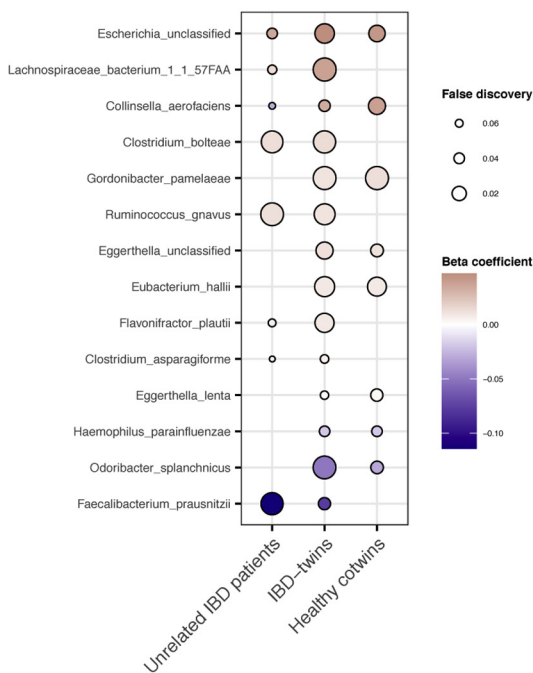
**A. Differentially abundant species compared to healthy controls (FDR < 0.10)**



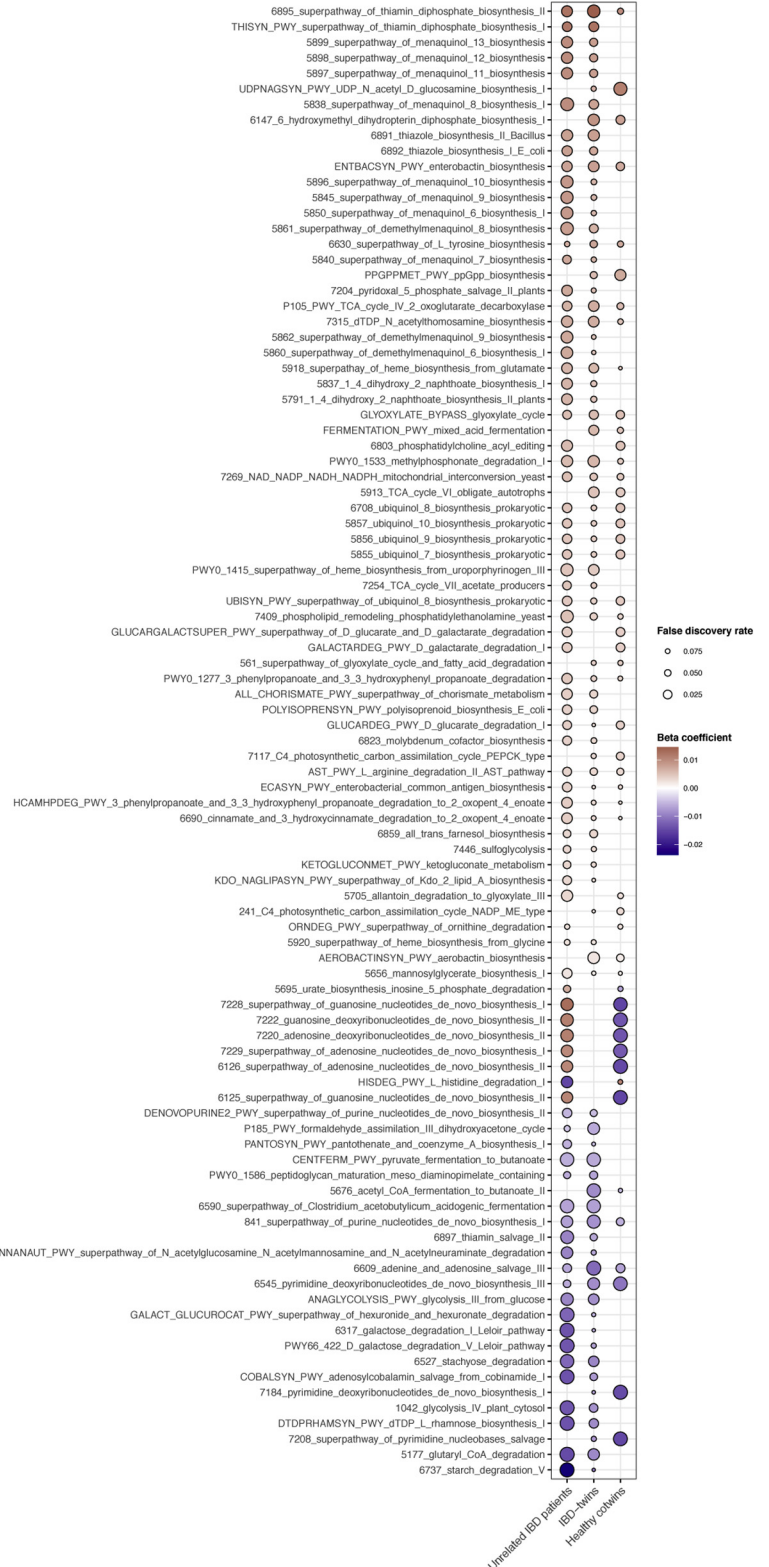
**Differentially abundant pathways compared to healthy controls (FDR < 0.10)**



**B. Shared differentially abundant species compared to healthy controls (FDR < 0.10)**



**C. Shared differentially abundant pathways compared to healthy controls (FDR < 0.10)**



CLINICAL AT



No statistically significant differences were found in the gut microbiome of the healthy cotwins compared with the IBD-twins for the overall taxonomic (PERMANOVA of Bray-Curtis  $\beta$ -diversities:  $R^2 = 0.015$ , FDR = 0.12) and functional (PERMANOVA:  $R^2 = 0.011$ , FDR = 0.36) composition (Figure 1, Supplementary Material 4). The overall taxonomic and functional composition was different for healthy cotwins, IBD-twins and unrelated patients with IBD compared with healthy controls (PERMANOVA: taxonomic:  $R^2 = 0.004$ , 0.012, 0.010; FDR = 0.012, 0.0002, 0.0002; functional:  $R^2 = 0.013$ , 0.012, 0.020; FDR = 0.0002, 0.0002, 0.0002 respectively, Figure 1, Supplementary Material 4). Overall, no differences between the gut microbiome composition of healthy cotwins and IBD-twins were detected, whereas both groups' microbiome composition differed from healthy controls.

### Association Between IBD Concordance, IBD Phenotype, Zygosity, and Co-housing and the Composition of Gut Microbiome

The associations of IBD concordance, IBD phenotype, zygosity, and co-housing during sampling with the interindividual heterogeneity in gut microbiome composition were assessed by using the Bray-Curtis dissimilarity by comparing 2 individuals. In line with our observations for the similarity in microbiome diversity between healthy cotwins and IBD-twins, we found no differences in the dissimilarity between IBD-discordant and IBD-concordant twin pairs, both at a taxonomic and functional levels (Figure 2A). When looking at unrelated random pairs of twins, we observed an increase in dissimilarity from healthy, healthy cotwins, to UC, to CD, underscoring the heterogeneity in changes in the gut microbiome composition in IBD (Figure 2B). The gut microbiomes of monozygotic or dizygotic twin pairs or twin pairs co-housing during sampling was not more similar than those of random unrelated healthy controls (Figure 2C and D).

### Healthy Cotwins, IBD-Twins, and Unrelated IBD Patients Differ in Relative Abundance of Species and Pathways From Healthy Controls

To identify whether the gut microbial features of the healthy cotwins are more alike to those displayed in an IBD microbiome (ie, IBD-twins and unrelated patients with IBD) or more alike to those displayed in healthy controls, we performed multivariable linear regression analyses adjusted

for IBD subtype (CD, UC, and no IBD), age, sex, BMI, antibiotics and PPI use, and sequencing depth. In the within-twin pairs analyses, additionally adjusted for zygosity and disease location, and with twin pair as random effects, none of the 119 species and 343 predicted bacterial pathways tested were differentially abundant between the healthy cotwins and the IBD-twins (FDR >0.1, Supplementary Material 5). However, compared with the healthy controls, 19 species and 105 pathways were differentially abundant within the gut microbiome of IBD-twins, and 18 species and 153 pathways in unrelated patients with IBD (Figure 3, and Supplementary Materials 6 and 7, FDR <0.1). The relative abundance of 13 species and 78 pathways differed significantly between microbiomes of healthy cotwins and healthy controls (Figure 3, Supplementary Material 8).

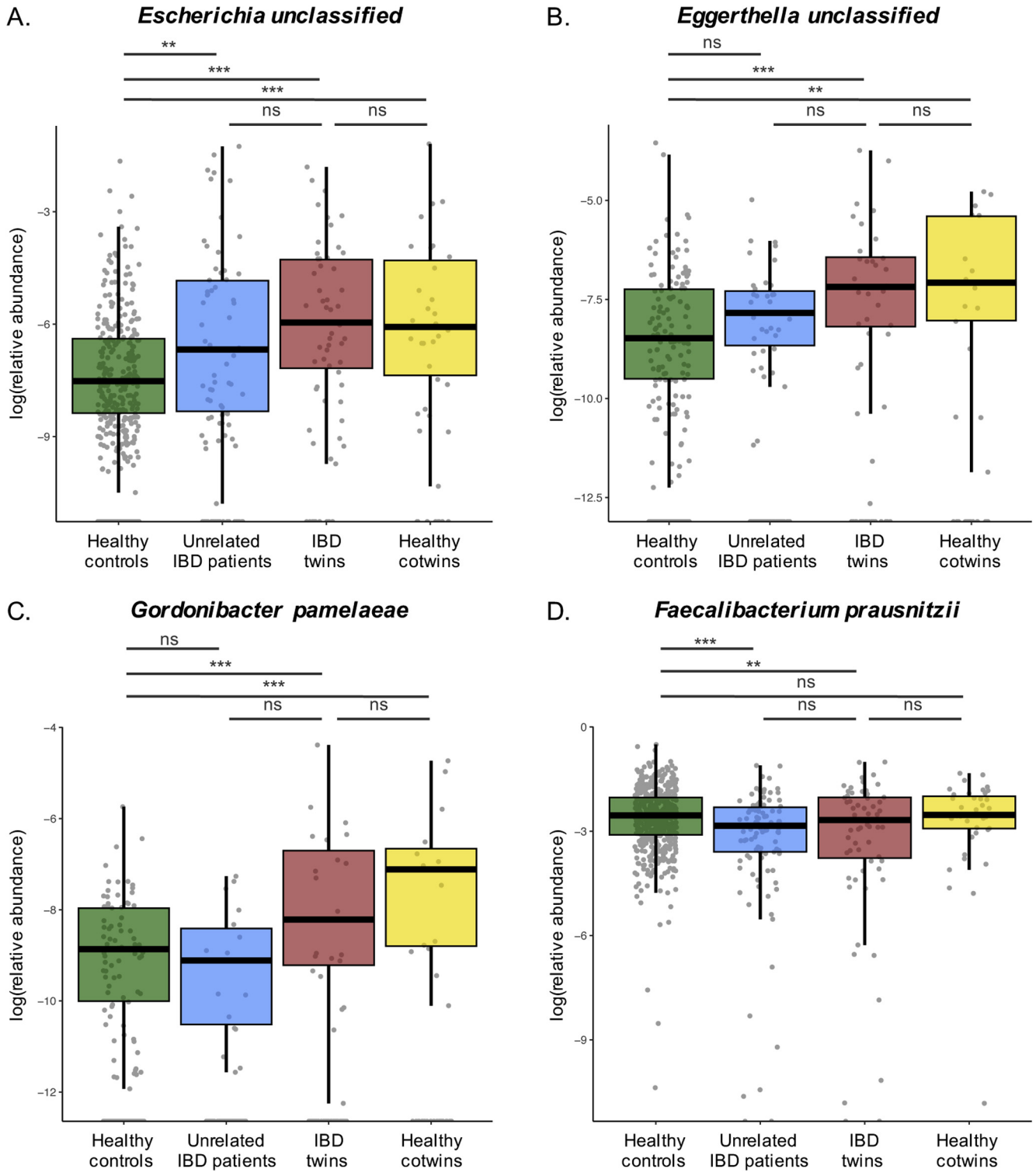
The number of species and pathways that were differentially abundant when we compared IBD-twins and unrelated patients with IBD directly, 7 species and 34 pathways (Supplementary Material 9), was less than when we compared both groups with healthy controls. In combination with the overlap of 7 species and 70 pathways that were differentially abundant compared with healthy controls, this indicates that the gut microbiome composition of IBD-twins and unrelated patients with IBD was comparable.

### Healthy Cotwins Share Differentially Abundant Taxa and Pathways With IBD-Twins and Unrelated IBD Patients Compared With Healthy Controls

Of the 19 species and 105 pathways that were differentially abundant between IBD-twins compared with healthy controls, 8 of these species (42.1%) and 37 pathways (35.2%) were also differentially abundant between the healthy cotwins compared with healthy controls. Moreover, of the 18 species and 153 pathways that were differentially abundant between the unrelated patients with IBD and the healthy controls, 1 of these species (5.6%) and 30 pathways (19.6%) were also differentially abundant between the healthy cotwins compared with the healthy controls (Figure 3 and Supplementary Material 10).

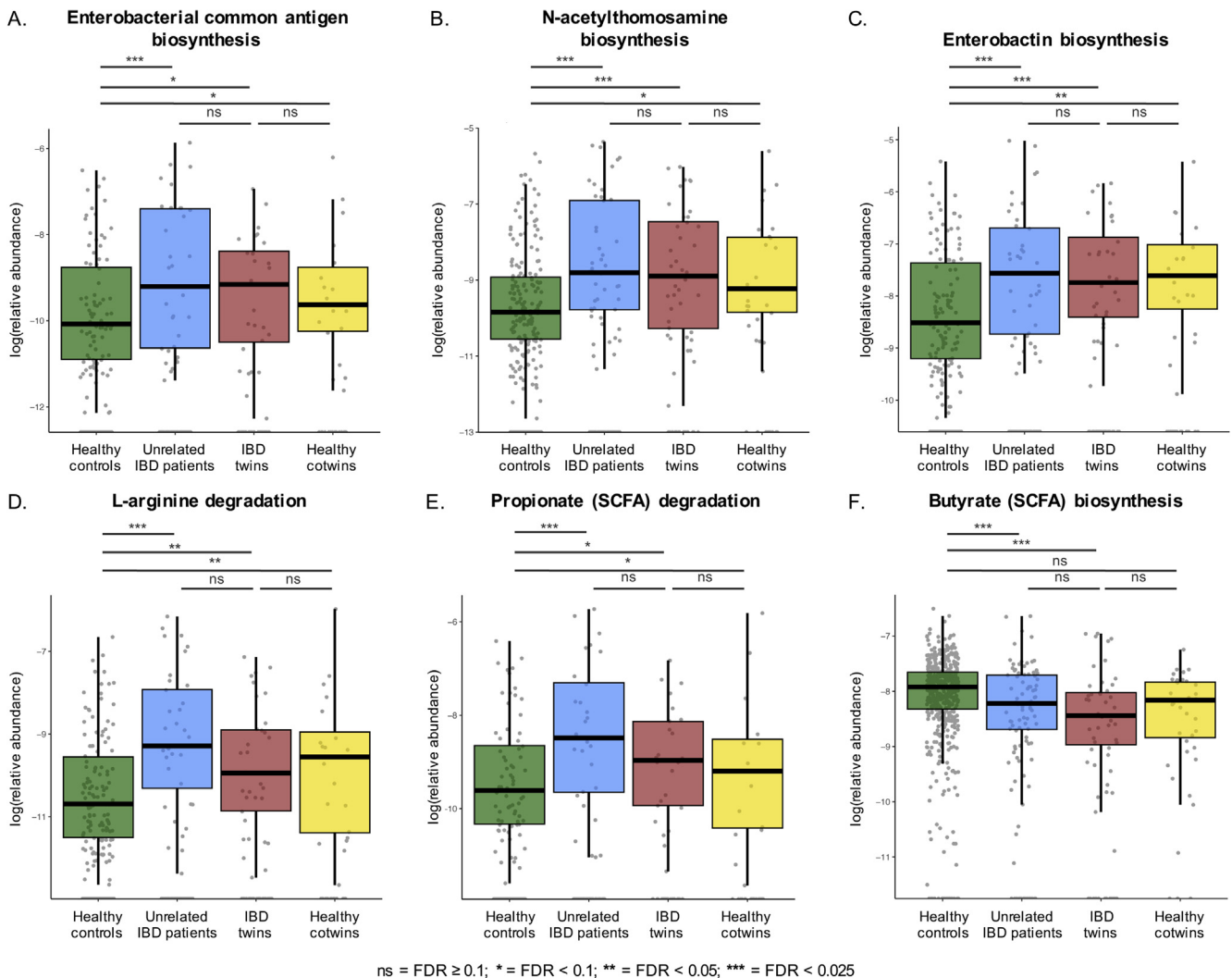
Among these, an increase in the relative abundance of potentially pathogenic species *Escherichia unclassified*, *Gordonibacter pamelaee*, and *Eggerthella unclassified*, among others, was observed in the gut microbiomes of healthy cotwins and IBD-twins or unrelated patients with IBD, as compared with the healthy controls (Figures 3B and 4A-C). *Faecalibacterium prausnitzii* was only statistically

**Figure 3.** The relative abundance of species and pathways in the gut microbiomes of healthy cotwins, IBD-twins, and unrelated patients with IBD differs from healthy controls and these partly overlap. Results from multivariable general linear model analyses adjusted for age, sex, BMI, antibiotics use, PPI use, IBD phenotype, and sequence read depth. (A) Venn diagrams depicting the absolute number of differentially abundant (FDR <0.10) species (top) and pathways (bottom) compared with healthy controls and the overlap between healthy cotwins, IBD-twins, and unrelated patients with IBD. Balloon plots show the individual (B) species and (C) pathways that are differentially abundant compared with healthy controls among at least 2 groups. Larger symbols depict lower FDR values, the color depicts the size of the effect. A beta coefficient <0 implicates a decrease in relative abundance, and >0 an increase in relative abundance compared with healthy controls. Not only are species and pathways shared differentially abundant among the groups, the effect size and direction are mostly similar in the shared species and pathways. Pathway names are based on the HUMAnN 2.0 pipeline.



CLINICAL AT

**Figure 4.** Relative abundance of a selection of species. The relative abundance of IBD-associated species, (A) *Escherichia unclassified*, (B) *Eggerthella unclassified*, (C) *Gordonibacter pamelaeae* is increased among healthy cotwins and shared differentially between IBD-twins and/or unrelated patients with IBD compared with healthy controls. The relative abundance of (D) *Faecalibacterium prausnitzii*, often inversely associated with IBD, is decreased in unrelated patients with IBD and IBD-twins, but not in healthy cotwins compared with healthy controls. The FDR values are based on multivariable general linear model analyses adjusted for age, sex, BMI, antibiotics use, PPI use, IBD phenotype, and sequence read depth.



ns = FDR  $\geq$  0.1; \* = FDR < 0.1; \*\* = FDR < 0.05; \*\*\* = FDR < 0.025

**Figure 5.** Relative abundance of a selection of pathways. The relative abundance of pathways of (A) enterobacterial common antigen biosynthesis, (B) N-acetylthomosamine (an enterobacterial common antigen component) biosynthesis, (C) enterobactin (a siderophore and virulence factor) biosynthesis, (D) L-arginine degradation, and (E) propionate (an SCFA) degradation are increased in healthy cotwins and shared differentially abundant in the gut microbiomes of IBD-twins and unrelated patients with IBD compared with healthy controls. (F) Butyrate biosynthesis is decreased in unrelated patients with IBD and IBD-twins, but not in healthy cotwins compared with healthy controls. The FDR values are based on multivariable general linear model analyses adjusted for age, sex, BMI, antibiotics use, PPI use, IBD phenotype, and sequence read depth.

significantly decreased compared with healthy controls in IBD-twins and unrelated patients with IBD, but not in healthy cotwins (Figures 3B and 4D).

Two pathways involved in the biosynthesis of family-specific surface enterobacterial common antigen, which is specific to the family Enterobacteriaceae, among which the genus *Escherichia*<sup>23</sup> (ECASYN\_PWY\_enterobacterial\_common\_antigen\_biosynthesis, 7315\_dTDP\_N\_acetylthomosamine\_biosynthesis) were increased in the healthy cotwins, IBD-twins, and unrelated patients with IBD, as compared with the healthy individuals (FDR < 0.1) (Figures 3C and 5A and 5B). The relative abundance of pathways for the biosynthesis of siderophores, iron-chelating molecules that are known potential virulence factors,<sup>24</sup> were increased in healthy cotwins, IBD-twins, and unrelated patients with IBD, namely enterobactin,

(ENTBACSYN\_PWY\_enterobactin\_biosynthesis, Figure 5C), and in healthy cotwins and IBD-twins, namely aerobactin (AEROBACTINSYN\_PWY\_aerobactin\_biosynthesis, Figure 3C). Moreover, the relative abundance of genes encoding the degradation of amino acid L-arginine, which is known to promote gut integrity via tight junctions between enterocytes,<sup>25,26</sup> (AST\_PWY\_L\_arginine\_degradation\_II\_AST\_pathways) was increased in healthy cotwins, IBD-twins, and unrelated patients with IBD (Figure 5D) and the degradation of short-chain fatty acid (SCFA) propionate (PWY0\_1277\_3\_phenylpropanoate\_and\_3\_3\_hydroxyphenylpropanoate\_degradation, HCAMHPDEG\_PWY\_3\_phenylpropanoate\_and\_3\_3\_hydroxyphenylpropanoate\_degradation\_to\_2\_oxopent\_4\_enoate) was increased in the gut microbiomes of healthy cotwins and IBD-twins, as compared with the healthy individuals (Figure 5E, FDR < 0.1). In line with the decrease in

the relative abundance of *F prausnitzii*, the relative abundance of 2 butyrate synthesis pathways (CENTFERM\_PWY\_pyr-uvate\_fermentation\_to\_butanoate and 6590\_superpathway\_of\_Clostridium\_acetobutylicum\_acidogenic\_fermentation) were only statistically significantly decreased in IBD-twins and unrelated patients with IBD, and not in healthy cotwins (Figure 5F), although 1 butyrate biosynthesis pathway was decreased in healthy cotwins and IBD-twins compared with healthy controls (5676\_acetyl\_CoA\_fermentation\_to\_butanoate\_II).

Taken all together, an increase in potentially pathogenic species and proinflammatory pathways was noted in the gut microbiome of healthy cotwins of IBD-discordant twin pairs compared with healthy controls. A substantial proportion of these differences are shared with IBD-twins and unrelated patients with IBD.

## Discussion

In this unique cross-sectional study analyzing the gut microbiome in IBD twin pairs, using metagenomic shotgun sequencing, we observed no differences in  $\alpha$  and  $\beta$  diversity and the relative abundance of individual species and pathways between healthy cotwins and their IBD-twins. Compared with age-, sex-, and BMI-matched healthy controls, healthy cotwins displayed a large overlap in differentially abundant species and pathways, not only with their IBD-twins, but also with unrelated patients with IBD. Among these overlapping species and pathways, species associated with IBD and potentially inflammation-related pathways were present. This implies that either the diagnosis of IBD might be preceded by microbial compositional changes or an increased risk of IBD is associated with an altered microbial composition or a combination of both.

In the present study, the relative abundances of taxa and pathways were not statistically significantly different between IBD-twins and their healthy cotwins. This is in contrast with previous, smaller-sized, 16S rRNA sequencing based microbiome IBD twin studies,<sup>11,13-15</sup> in which gut microbiome differences between IBD-affected twins, especially individuals with ileal CD, and their healthy cotwins were reported.<sup>11,13</sup> In these previous twin studies, furthermore, a microbial profile linked to ileal CD that differed from healthy cotwins and patients with colonic CD,<sup>11-13</sup> and a less diverse microbial composition at the intestinal mucosa in UC-affected twins compared with their healthy cotwins was observed.<sup>14</sup> This last phenomenon was found to be less pronounced in healthy cotwins, but was still decreased compared with healthy non-IBD related twins.<sup>14</sup> A Belgian study in siblings and parents of CD-affected patients, using denaturing gradient gel electrophoresis, found a fecal microbiome dysbiosis in the unaffected relatives as compared with healthy controls characterized by mucin degrading microbiota.<sup>27</sup> An English study among siblings of patients with CD also noted a dysbiosis in the fecal microbiome of the IBD-unaffected siblings.<sup>28</sup> The results of these smaller family (ie, non-twin) studies based on techniques with lower resolution are in line with our findings with a

changed composition of the fecal microbiome in healthy cotwins with IBD-affected twins.

The question arises whether shared genetics and (childhood) environment, rather than IBD phenotype or disease activity, might have resulted in a more similar composition of the gut microbiome between healthy cotwins and their IBD-twins. Previous work underscored the large impact of environmental factors during childhood on the adult gut microbiome composition.<sup>5</sup> To further unravel the processes shaping the gut microbiome and explore the putative microbial drivers of IBD, we included unrelated patients with IBD, in addition to the IBD-twins. Unrelated patients with IBD showed only minor differences in individual taxa and pathways with IBD-twins. Interestingly, healthy cotwins did not only have overlap in differentially abundant species and pathways with their IBD-twins, but also with the unrelated patients with IBD compared with healthy controls. This renders shared environment and genetics as sole explanation for the overlap in microbiome signatures between IBD-twins and their healthy cotwins less probable, and suggests a mechanistic role for the gut microbiome in the development of IBD in the prediagnostic state whether or not related to subclinical inflammation, which we cannot rule out. Although prediagnostic microbiome data are lacking for IBD, in type 1 diabetes mellitus gut microbiome changes, albeit moderate, have been shown to occur before diagnosis,<sup>29</sup> hinting toward the possibility of gut microbial changes occurring in the mechanistic chain of disease development.

One of the IBD signatures detected in healthy cotwins was an increase in the relative abundance of *Escherichia unclassified* (Figure 4A). *Escherichia unclassified* belongs to the genus of gram-negative, facultative anaerobic taxa that display lipopolysaccharide, a potent stimulator of innate immune responses, on their outer surface. *Escherichia unclassified* has previously been associated with UC and CD, and is considered a key pathogenic driver in IBD.<sup>30,31</sup> Moreover, 2 pathways involved in the biosynthesis of family-specific surface enterobacterial common antigen, which is shared by all members of, and restricted to, the family Enterobacteriaceae,<sup>23</sup> and a pathway for the biosynthesis of enterobactin, a siderophore that chelates Fe<sup>3+</sup> and is known to be a virulence factor,<sup>24</sup> were increased in the healthy cotwins, IBD-twins, and unrelated patients with IBD, as compared with the healthy individuals. The relative abundance of the genes encoding L-arginine degradation was increased in healthy cotwins, IBD-twins, and unrelated patients with IBD, as compared with the healthy controls, as well. L-Arginine is a precursor for the synthesis of polyamines. Polyamines contribute to the integrity of the gut and reduced expression of proinflammatory cytokines by monocytes and macrophages.<sup>32</sup> L-arginine has been associated with a protective effect of colitis in mice models,<sup>33,34</sup> and decreased levels of L-arginine have been observed in the intestinal epithelium in active UC possibly caused by decreased cellular uptake and increased consumption by nitric oxide synthase 2.<sup>32</sup> A decreased L-arginine biosynthesis or increased L-arginine degradation has been reported previously in metagenomic shotgun sequencing-



based microbiome studies in IBD.<sup>3,35</sup> Likewise, an increase in the relative abundance of the gene encoding the degradation of propionate, an SCFA, was observed in healthy cotwins, IBD-twins and unrelated patients with IBD, as compared with healthy controls. SCFAs are an important energy source for enterocytes, and induce tolerogenic and anti-inflammatory enterocyte and T-cell phenotypes by multiple mechanisms.<sup>36,37</sup> SCFAs constitute an increasingly convincing link between the gut microbiome and the IBD phenotype, and SCFA supplementation has been found to attenuate colonic inflammation.<sup>38</sup> Interestingly, the relative abundance of *F prausnitzii* (often inversely associated with IBD<sup>2</sup>), a well-known butyrate producer, and 2 butyrate biosynthesis pathways was found to be decreased in IBD-twins and unrelated patients with IBD, but not in healthy cotwins. We could thus not replicate the data from 1 Spanish<sup>39</sup> and 1 English<sup>28</sup> smaller-sized 16S rRNA quantitative polymerase chain reaction-based study, in which a decrease of *F prausnitzii* in relatives of patients with UC<sup>39</sup> and relatives of patients with CD<sup>28</sup> was found.

To date, insights in the preclinical phase of IBD are scarce.<sup>40</sup> Previous epidemiological studies have shown an increased risk of developing IBD in healthy cotwins from IBD-discordant twin pairs.<sup>10</sup> It is therefore tempting to speculate that changes in the gut microbiome of healthy cotwins precede IBD development and are involved in the pathogenesis. An alternative explanation might be that these microbiome alterations reflect the impact of shared genetic makeup or environmental factors in these individuals, but do not necessarily lead to IBD development. Longitudinal studies, with sampling at multiple timepoints in high-risk individuals before the onset of IBD (for example the GEM-project<sup>41</sup>), are therefore needed to get a true insight in the preclinical phase of IBD.

Our study is, to the best of our knowledge, the largest twin study in the field of IBD and the microbiome up to date with high-resolution assessment of the taxonomic and functional composition of the gut microbiome. An important strength of our study included the comparison of carefully phenotyped IBD-discordant and -concordant twin pairs, and a large cohort of age-, sex-, and BMI-matched healthy controls and unrelated patients with IBD, allowing us to study several aspects of the overlap of the microbiome between healthy cotwins and patients with IBD. Although the participants came from 2 cohorts, the DNA isolation, library preparation, and sequencing were for all samples performed in the same way, at the same location and time by the same technician, minimizing the risk for batch effects. Furthermore, we adjusted our analyses for potential confounding factors, thereby reducing the chance of identifying spurious associations between the microbiome composition and IBD-status.

Our study does, however, have its limitations. Further increasing the sample size could have increased the power to detect more subtle differences in microbiome composition. Furthermore, multicollinearity prevented us from correcting for the use of IBD medication (which has been shown, however, to be only mildly associated with the gut microbiome composition<sup>42</sup>), and stool consistency. Patients

with CD and patients with UC were grouped together in our analyses. However, by including IBD phenotype as a covariate in our regression models we corrected for its potential confounding effects. Although we sought, as described, to minimize batch effects as much as possible, this could not completely be avoided, given the fact that the fecal samples were collected from individuals included in 2 different cohorts. Last, the unrelated patients with IBD were self-reported and, therefore, detailed information on Montreal classification, IBD-medication use, and disease activity was not available. Nonetheless, the gut microbiome composition of IBD-twins and unrelated patients with IBD considerably overlapped, as would have been expected when comparing 2 cohorts of patients with IBD.

In conclusion, we found that the gut microbiome of healthy cotwins from IBD-discordant twin pairs displays IBD-like signatures, both at a taxonomic and functional level. The gut microbiome of these individuals at increased risk of developing IBD displays similarities to the gut microbiome of their IBD-affected twins and unrelated patients with IBD, and is different from healthy controls. These IBD-like microbiome signatures could be a reflection of a shared genetic background and environment and might precede IBD development. However, longitudinal studies are needed to infer a causal relationship.

## Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online, version of *Gastroenterology* at [www.gastrojournal.org](http://www.gastrojournal.org), and at <https://doi.org/10.1053/j.gastro.2021.01.030>.

## References

1. Balfour Sartor R, Wu GD. Roles for intestinal bacteria, viruses, and fungi in pathogenesis of inflammatory bowel diseases and therapeutic approaches. *Gastroenterology* 2017;152:327–339.e4.
2. 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.
3. Vich Vila A, Imhann F, Collij V, et al. Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci Transl Med* 2018;10:eaap3914.
4. Turnbaugh PJ, Hamady M, Yatsunencko T, et al. A core gut microbiome in obese and lean twins. *Nature* 2009;457:480–484.
5. Rothschild D, Weissbrod O, Barkan E, et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature* 2018;555:210–215.
6. Xie H, Guo R, Zhong H, et al. Shotgun metagenomics of 250 adult twins reveals genetic and environmental impacts on the gut microbiome. *Cell Syst* 2016;3:572–584.e3.
7. Yatsunencko T, Rey FE, Manary MJ, et al. Human gut microbiome viewed across age and geography. *Nature* 2012;486:222–227.

8. Goodrich JK, Waters JL, Poole AC, et al. Human genetics shape the gut microbiome. *Cell* 2014;159:789–799.
9. **Imhann F, Vich Vila A**, Bonder MJ, et al. Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. *Gut* 2018;67:108–119.
10. Ek WE, D'Amato M, Halfvarson J. The history of genetics in inflammatory bowel disease. *Ann Gastroenterol* 2014;27:294–303.
11. Willing BP, Dicksved J, Halfvarson J, et al. A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology* 2010;139:1844–1854.e1.
12. **Erickson AR, Cantarel BL, Lamendella R**, et al. Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS One* 2012;7:e49138.
13. **Dicksved J, Halfvarson J**, Rosenquist M, et al. Molecular analysis of the gut microbiota of identical twins with Crohn's disease. *ISME J* 2008;2:716–727.
14. **Lepage P, Häslér R, Spehlmann ME**, et al. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* 2011;141:227–236.
15. **Willing B, Halfvarson J**, Dicksved J, et al. Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflamm Bowel Dis* 2009;15:653–660.
16. Ligthart L, van Beijsterveldt CEM, Kevenaar ST, et al. The Netherlands Twin Register: longitudinal research based on twin and twin-family designs. *Twin Res Hum Genet* 2019;22:623–636.
17. Casparie M, Tiebosch ATMG, Burger G, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* 2007;29:19–24.
18. Zhernakova A, Kurilshikov A, Bonder MJ, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* 2016;352:565–569.
19. Bennebroek Evertsz' F, Hoeks CCMQ, Nieuwkerk PT, et al. Development of the patient Harvey Bradshaw index and a comparison with a clinician-based Harvey Bradshaw index assessment of Crohn's disease activity. *J Clin Gastroenterol* 2013;47:850–856.
20. Bennebroek Evertsz' F, Nieuwkerk PT, Stokkers PCF, et al. The patient simple clinical colitis activity index (P-SCCAI) can detect ulcerative colitis (UC) disease activity in remission: a comparison of the P-SCCAI with clinician-based SCCAI and biological markers. *J Crohns Colitis* 2013;7:890–900.
21. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2018. Available at: <https://www.R-project.org/>.
22. Mallick H, Tickle T, McIver L, et al. Multivariable association in population-scale meta'omic surveys. Available at: <http://huttenhower.sph.harvard.edu/maaslin2>.
23. Kuhn HM, Meier-Dieter U, Mayer H. ECA, the enterobacterial common antigen. *FEMS Microbiol Rev* 1988;4:195–222.
24. Holden VI, Bachman MA. Diverging roles of bacterial siderophores during infection. *Metallomics* 2015;7:986–995.
25. Liu L, Guo X, Rao JN, et al. Polyamines regulate E-cadherin transcription through c-Myc modulating intestinal epithelial barrier function. *Am J Physiol Cell Physiol* 2009;296:C801–C810.
26. Chen J, Rao JN, Zou T, et al. Polyamines are required for expression of Toll-like receptor 2 modulating intestinal epithelial barrier integrity. *Am J Physiol Gastrointest Liver Physiol* 2007;293:G568–G576.
27. Joossens M, Huys G, Cnockaert M, et al. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* 2011;60:631–637.
28. Hedin CR, McCarthy NE, Louis P, et al. Altered intestinal microbiota and blood T cell phenotype are shared by patients with Crohn's disease and their unaffected siblings. *Gut* 2014;63:1578–1586.
29. Vatanen T, Franzosa EA, Schwager R, et al. The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. *Nature* 2018;562:589–594.
30. Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 2019;569:655–662.
31. Mirsepasi-Lauridsen HC, Vallance BA, Krogfelt KA, et al. *Escherichia coli* pathobionts associated with inflammatory bowel disease. *Clin Microbiol Rev* 2019;32. e00060-18.
32. Coburn LA, Horst SN, Allaman MM, et al. L-Arginine availability and metabolism is altered in ulcerative colitis. *Inflamm Bowel Dis* 2016;22:1847–1858.
33. **Ren W, Yin J**, Wu M, et al. Serum amino acids profile and the beneficial effects of L-arginine or L-glutamine supplementation in dextran sulfate sodium colitis. *PLoS One* 2014;9:e88335.
34. Coburn LA, Gong X, Singh K, et al. L-arginine supplementation improves responses to injury and inflammation in dextran sulfate sodium colitis. *PLoS One* 2012;7:e33546.
35. **Klaassen MAY, Imhann F**, Collij V, et al. Anti-inflammatory gut microbial pathways are decreased during crohn's disease exacerbations. *J Crohns Colitis* 2019;13:1439–1449.
36. Lavelle A, Sokol H. Gut microbiota-derived metabolites as key actors in inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol* 2020;17:223–237.
37. Koh A, Vadder F De, Kovatcheva-Datchary P, et al. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* 2016;165:1332–1345.
38. Maslowski KM, Vieira AT, Ng A, et al. Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* 2009;461:1282–1286.
39. Varela E, Manichanh C, Gallart M, et al. Colonisation by *Faecalibacterium prausnitzii* and maintenance of clinical

remission in patients with ulcerative colitis. *Aliment Pharmacol Ther* 2013;38:151–161.

40. Torres J, Burisch J, Riddle M, et al. Preclinical disease and preventive strategies in IBD: perspectives, challenges and opportunities. *Gut* 2016;65:1061–1069.
41. Turpin W, Lee S-H, Raygoza Garay JA, et al. Increased intestinal permeability is associated with later development of Crohn's disease. *Gastroenterology* 2020; 159:2092–2100.e5.
42. Vich Vila A, Collij V, Sanna S, et al. Impact of commonly used drugs on the composition and metabolic function of the gut microbiota. *Nat Commun* 2020;11:362.

Author names in bold designate shared co-first authorship.

Received August 7, 2020. Accepted January 7, 2021.

#### Correspondence

Address correspondence to: Rinse K. Weersma, MD, PhD, Department of Gastroenterology and Hepatology & Department of Genetics, University of Groningen and University Medical Center Groningen, PO Box 30.001, 9700RB Groningen, the Netherlands; e-mail: [r.k.weersma@umcg.nl](mailto:r.k.weersma@umcg.nl); fax: +31 50 361 9306. or Bas Oldenburg, MD, PhD, Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands; e-mail: [b.oldenburg@umcutrecht.nl](mailto:b.oldenburg@umcutrecht.nl); fax: +31 88 75 55081.

#### Acknowledgments

The authors thank all participants. They thank Dianne Jansen for her work on isolating microbial DNA from the fecal samples, and Bart Müskens for his help with study visits in the TWIN-study. The authors thank all funding partners for their support in this study.

#### CRedit Authorship Contributions

Eelco C. Brand, MD (Conceptualization: Equal; Data curation: Equal; Formal analysis:

Equal; Investigation: Equal; Methodology: Equal; Visualization: Equal; Writing – original

draft: Lead; Writing – review & editing: Lead; Generation and acquisition of data: Equal).

Marjolein A.Y. Klaassen, BSc (Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal; Visualization: Equal; Writing – original draft: Lead; Writing – review & editing: Lead).

Ranko Gacesa, PhD (Data curation: Equal; Formal analysis: Supporting; Investigation: Supporting; Methodology: Equal; Writing – review & editing: Equal).

Anau Vich Vila, PhD (Formal analysis: Supporting; Investigation: Supporting; Methodology: Supporting; Writing – review & editing: Equal).

Hiren Ghosh, PhD (Data curation: Supporting; Formal analysis: Supporting; Writing – review & editing: Equal).

Marcel R. de Zoete, PhD (Writing – review & editing: Equal).

Dorret I. Boomsma, PhD (Writing – review & editing: Equal).

Frank Hoentjen, MD, PhD (Writing – review & editing: Equal).

Carmen S. Horjus Talabur Horje, MD, PhD (Writing – review & editing: Equal).

Paul C. van de Meeberg, MD, PhD (Writing – review & editing: Equal).

Gonneke Willemsen, PhD (Writing – review & editing: Equal).

Jingyuan Fu, PhD (Writing – review & editing: Equal; Generation and acquisition of data: Equal).

Cisca Wijmenga, PhD (Writing – review & editing: Equal; Generation and acquisition of data: Equal).

Femke van Wijk, PhD (Writing – review & editing: Equal; Generation and acquisition of data: Equal).

Alexandra Zernakova, MD, PhD (Writing – review & editing: Equal; Generation and acquisition of data: Equal).

Bas Oldenburg, MD, PhD (Conceptualization: Equal; Formal analysis: Supporting; Investigation: Equal; Methodology: Supporting; Supervision: Lead; Writing – original draft: Supporting; Writing – review & editing: Equal; Generation and acquisition of data: Equal).

Rinse K. Weersma, MD, PhD (Conceptualization: Equal; Formal analysis: Supporting; Investigation: Equal; Methodology: Supporting; Supervision: Lead; Writing – original draft: Supporting; Writing – review & editing: Equal).

#### Conflict of interest

These authors disclose the following: Eelco C. Brand is supported by the Alexandre Suerman program for MD and PhD candidates of the University Medical Center Utrecht, Netherlands, and is co-applicant on an Investigator-Initiated research grant of Pfizer. Marcel R. de Zoete is supported by a VIDI

grant from the Netherlands Organization for Scientific Research (NWO, grant 91715377) and the Utrecht Exposome Hub of Utrecht Life Sciences, and a co-founder and consultant of Artizan Biosciences. Frank Hoentjen has received unrestricted grants from Janssen-Cilag, AbbVie, and Takeda; and consulting fees from Celgene and Janssen-Cilag. Jingyuan Fu is funded by the Netherlands Heart Foundation (IN CONTROL, CVON2018-27) and NWO Gravitation Netherlands Organ-on-Chip Initiative (024.003.001). Jingyuan Fu and Alexandra Zernakova are further supported by a CardioVasculair Onderzoek Nederland grant (CVON 2012-03). Cisca Wijmenga is supported by a Spinoza award (NWO SPI 92-266), an ERC advanced grant (ERC-671274), a grant from the Netherlands' Top Institute Food and Nutrition (GH001), the NWO Gravitation Netherlands Organ-on-Chip Initiative (024.003.001), the Stiftelsen Kristian Gerhard Jebsen foundation (Norway), and the RuG investment agenda grant Personalized Health. Femke van Wijk is supported by a VIDI career development grant from The Netherlands Organization for Health Research and Development (ZonMw), unrestricted grants from Pfizer and Regeneron, and has been a consultant for Sanofi. Alexandra Zernakova is supported by the ERC Starting Grant 715772, Netherlands Organization for Scientific Research NWO-VIDI grant 016.178.056, the Netherlands Heart Foundation CVON grant 2018-27, and the NWO Gravitation grant ExposomeNL 024.004.017. Bas Oldenburg has received unrestricted grants from AbbVie, Janssen, Pfizer, Ferring, Dr. Falk, Takeda, and MSD; and served on the advisory board of Janssen, Pfizer, Takeda, and Cablon. Rinse K. Weersma has received unrestricted grants from Takeda, Johnson & Johnson, Tramedico, and Ferring; and served as a consultant for Takeda. The remaining authors disclose no conflicts.

#### Funding

Eelco Brand is supported by the Alexandre Suerman program for MD and PhD candidates of the University Medical Center Utrecht, the Netherlands.

#### Members of the Dutch TWIN-IBD consortium:

In alphabetical order of the affiliation. Bas Oldenburg, MD, PhD, Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, The Netherlands, Femke van Wijk, PhD, Center for Translational Immunology, University Medical Center Utrecht, Utrecht, The Netherlands, Eelco C. Brand, MD, Department of Gastroenterology and Hepatology & Center for Translational Immunology, University Medical Center Utrecht, Utrecht, The Netherlands, Pieter Honkoop, MD, PhD, Department of Gastroenterology and Hepatology, Albert Schweitzer Hospital, Dordrecht, Zwiindrecht, Slidrecht, The Netherlands Rutger J. Jacobs, MD, PhD, Department of Gastroenterology and Hepatology, Alrijne Hospital, Leiden, Leiderdorp, Alphen aan den Rijn, The Netherlands, Cyriel Y. Ponsioen, MD, PhD, Department of Gastroenterology and Hepatology, Amsterdam University Medical Centers, location AMC, Amsterdam, The Netherlands, Nanne K.H. de Boer, MD, PhD, Department of Gastroenterology and Hepatology, Amsterdam UMC, Vrije Universiteit Amsterdam, AGEM institute, Amsterdam, The Netherlands, Yasser A. Alderlieste, MD, Department of Gastroenterology and Hepatology, Beatrix Hospital, Gorinchem, The Netherlands, Margot A. van Herwaarden, MD, PhD, Department of Gastroenterology and Hepatology, Deventer Hospital, Deventer, The Netherlands, Sebastiaan A.C. van Tuyl, MD, PhD, Department of Gastroenterology and Hepatology, Diakonessen Hospital, Utrecht, The Netherlands, Maurice W. Lutgens, MD, PhD, Department of Gastroenterology and Hepatology, Elisabeth-TweeSteden Hospital, Tilburg, The Netherlands, C. Janneke van der Woude, MD, PhD, Department of Gastroenterology and Hepatology, Erasmus Medical Center, Rotterdam, The Netherlands, Wout G.M. Mares, MD, Department of Gastroenterology and Hepatology, Gelderse Vallei Hospital, Ede, The Netherlands, Daan B. de Koning, MD, Department of Gastroenterology and Hepatology, Gelre Hospitals, Apeldoorn, Zutphen, The Netherlands, Joukje H. Bosman, MD, Department of Gastroenterology and Hepatology, Groene Hart Hospital, Gouda, The Netherlands, Juda Vecht, MD, PhD, Department of Gastroenterology and Hepatology, Isala, Zwolle, The Netherlands, Anneke M.P. de Schryver, MD, PhD, Department of Gastroenterology and Hepatology, Jeroen Bosch Hospital, Den Bosch, The Netherlands, Andrea E. van der Meulen-de Jong, MD, PhD, Department of Gastroenterology and Hepatology, Leiden University Medical Center, Leiden, The Netherlands, Marieke J. Pierik, MD, PhD, Department of Gastroenterology and Hepatology, Maastricht University Medical Center, Maastricht, The Netherlands, Paul J. Boekema, MD, PhD, Department of Gastroenterology and Hepatology, Maxima Medical Center, Veldhoven, Eindhoven, The Netherlands, Robert J. Verburg, MD, PhD, Department of Gastroenterology and Hepatology, Medical Center Haaglanden, Den Haag, The Netherlands, Bindia Jharap, MD, PhD, Department of Gastroenterology and Hepatology, Meander Medical Center, Amersfoort, The Netherlands, Gonneke Willemsen, PhD, Department of Biological Psychology, Behavioral and Movement Sciences, Vrije Universiteit, & 7. Amsterdam Public Health Research Institute, Vrije Universiteit Medical Center, Amsterdam, The Netherlands, Dorret I. Boomsma, PhD, Department of Biological Psychology, Behavioral and Movement Sciences, Vrije Universiteit, & 7. Amsterdam Public Health Research Institute, Vrije Universiteit Medical Center, Amsterdam, The Netherlands, Jeroen M. Jansen, MD, Department of Gastroenterology and Hepatology, OLVG Oost, Amsterdam, The Netherlands, Pieter C.F. Stokkers, MD, PhD, Department of Gastroenterology

and Hepatology, OLVG West, Amsterdam, The Netherlands, Frank Hoentjen, MD, PhD, Department of Gastroenterology and Hepatology, Radboud University Medical Center, Nijmegen, The Netherlands, Rutger Quispel, MD, Department of Gastroenterology and Hepatology, Reinier de Graaf Gasthuis, Delft, The Netherlands, Carmen S. Horjus Talabur Horje, MD, PhD, Department of Gastroenterology and Hepatology, Rijnstate Hospital, Arnhem, The Netherlands, Paul C. van de Meeberg, MD, PhD, Department of Gastroenterology and Hepatology, Slingeland Hospital, Doetinchem, The Netherlands, Nofel Mahmmod, MD, Department of Gastroenterology and Hepatology, St. Antonius Hospital, Nieuwegein, Utrecht, The Netherlands, Rachel L. West, MD, PhD, Department of Gastroenterology and Hepatology, St. Franciscus Gasthuis, Rotterdam, The Netherlands, Marleen Willems, MD, Department of Gastroenterology and Hepatology, St. Jansdal Hospital, Harderwijk, The Netherlands, Itta M. Minderhoud, MD, PhD, Department of Gastroenterology and Hepatology, Tergooi Hospital, Blaricum, Hilversum,

The Netherlands, Herma H. Fidder, MD, PhD, Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, The Netherlands, Fiona D.M. van Schaik, MD, PhD, Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, The Netherlands, Meike M.C. Hirdes, MD, PhD, Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, The Netherlands, Nynke A. Boontje, MSc, Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, The Netherlands, Bart L.M. Müskens, BSc, Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, The Netherlands, Rinse K. Weersma, MD, PhD, Department of Gastroenterology and Hepatology, University Medical Center Groningen, Groningen, The Netherlands, Marielle J.L. Romberg-Camps, MD, PhD, Department of Gastroenterology and Hepatology, Zuyderland Hospital, Sittard-Geleen, Heerlen, The Netherlands