

Original Article

# The Composition and Metabolic Potential of the Human Small Intestinal Microbiota Within the Context of Inflammatory Bowel Disease

Renate A. A. Ruigrok,<sup>a,b</sup> Valerie Collij,<sup>a,b</sup> Paula Sureda,<sup>a,b</sup>  
Marjolein A. Y. Klaassen,<sup>a,b</sup> Laura A. Bolte,<sup>a,b</sup> Bernadien H. Jansen,<sup>a</sup>  
Michiel D. Voskuil,<sup>a,b</sup> Jingyuan Fu,<sup>b,c</sup> Cisca Wijmenga,<sup>b</sup>  
Alexandra Zhernakova,<sup>b</sup> Rinse K. Weersma,<sup>a</sup> Arnau Vich Vila<sup>a,b</sup>

<sup>a</sup>Department of Gastroenterology and Hepatology, University Medical Center Groningen, Groningen, The Netherlands

<sup>b</sup>Department of Genetics, University Medical Center Groningen, Groningen, The Netherlands <sup>c</sup>Department of Pediatrics, University Medical Center Groningen, Groningen, The Netherlands

Corresponding author: A. Vich Vila, PO Box 30.001, 9700RB Groningen, The Netherlands. Tel: +31 50 361 2620; Fax: +31 50 361 9306; Email: [a.vich.vila@umcg.nl](mailto:a.vich.vila@umcg.nl)

## Abstract

**Background and Aims:** The human gastrointestinal tract harbours distinct microbial communities essential for health. Little is known about small intestinal communities, despite the small intestine playing a fundamental role in nutrient absorption and host-microbe immune homeostasis. We aimed to explore the small intestine microbial composition and metabolic potential, in the context of inflammatory bowel disease [IBD].

**Methods:** Metagenomes derived from faecal samples and extensive phenotypes were collected from 57 individuals with an ileostomy or ileoanal pouch, and compared with 1178 general population and 478 IBD faecal metagenomes. Microbiome features were identified using *MetaPhAn2* and *HUMAnN2*, and association analyses were performed using multivariate linear regression.

**Results:** Small intestinal samples had a significantly lower bacterial diversity, compared with the general population and, to a lesser extent, IBD samples. Comparing bacterial composition, small intestinal samples clustered furthest from general population samples and closest to IBD samples with intestinal resections. *Veillonella atypica*, *Streptococcus salivarius*, and *Actinomyces graevenitzii* were among the species significantly enriched in the small intestine. Predicted metabolic pathways in the small intestine are predominantly involved in simple carbohydrate and energy metabolism, but also suggest a higher pro-inflammatory potential.

**Conclusions:** We described the bacterial composition and metabolic potential of the small intestinal microbiota. The colonic microbiome of IBD patients, particularly with intestinal resections, showed resemblance to that of the small intestine. Moreover, several features characterising the small intestinal microbiome have been previously associated with IBD. These results highlight the importance of studying the small intestinal microbiota to gain new insight into disease pathogenesis.

**Key Words:** Inflammatory bowel disease; shotgun sequencing; small intestinal microbiota

## 1. Introduction

The human gut microbiota, which refers to the trillions of bacteria, viruses, fungi, and archaea that inhabit the gastrointestinal [GI] tract, plays an important role in maintaining health.<sup>1,2</sup> Alteration to the composition of the gut microbiota has already been widely described for several disorders ranging from GI, including inflammatory bowel disease [IBD], to neurological.<sup>2-4</sup> The use of faecal samples in the majority of these studies, however, has meant that most findings are largely specific to the colonic content<sup>5</sup>; that is, the faecal microbiome does not capture all the microbial communities inhabiting other parts of the GI tract, such as the small intestine, which remain considerably understudied.

The small intestine is responsible for approximately 90% of the body's total nutrient absorption from the diet and plays a central role in the maintenance of host-microbe immune homeostasis.<sup>6,7</sup> Dysbiosis of the duodenal microbiota has been associated with certain GI-related disorders and complaints, such as functional dyspepsia, bloating, and diarrhoea.<sup>8,9</sup> Additionally, the ectopic colonisation of microbes typical of the oral cavity has been hypothesised to play a role in the pathogenesis of several disorders: a phenomenon termed 'oralization'.<sup>10,11</sup> Specific strains of *Klebsiella pneumoniae* isolated from the salivary microbiota of patients with IBD, for example, were shown to cause aberrant activation of the immune system in colitis-prone mice, following their colonisation in the colon.<sup>12</sup> Bacteria considered oral have also been identified in the small intestine.

Studying the small intestinal content, especially within a healthy context, is challenging due to its poor accessibility. Most studies to date have relied on using mucosal samples collected during routine endoscopies, following intestinal resections or from sudden death individuals.<sup>13</sup> Such sampling methods, however, are prone to contamination and may be hampered by the lavage treatment that precedes some of these procedures. Moreover, they do not represent the luminal content of the small intestine and are limited by the lower taxonomic and functional resolution of 16s rRNA sequencing. Individuals with an ileostomy or ileoanal pouch, following treatment for an intestinal-related complication or disease such as IBD, present a unique group in which to non-invasively sample the small intestine. Faecal samples from these individuals currently provide the closest representation of the luminal content in the small intestine, although the disease context should be kept in mind.

Here, we aimed to characterise the composition and metabolic potential of the small intestinal microbiota, with a specific focus on its possible implications in IBD. We analysed shotgun metagenomes derived from faecal samples collected from 1713 participants, including 57 samples from individuals with an ileostomy or ileoanal pouch due to IBD, which represented the small intestinal microbiota. The small intestinal metagenomes were compared with the remaining metagenomes representing the colonic microbiota of the general population [ $n = 1178$ ] and of patients with IBD [ $n = 478$ ].

## 2. Materials and Methods

### 2.1. Cohort description

To study the small intestinal microbiota and its potential implications in IBD, two independent Dutch cohorts were used: 1] the 1000IBD cohort<sup>14</sup> established in the IBD centre at the University

Medical Center Groningen [UMCG], The Netherlands; and 2] Lifelines DEEP,<sup>15</sup> a general population cohort from the northern provinces of The Netherlands.

Metagenomic, in combination with phenotypic, data were available for 535 of the 1000IBD cohort participants. All participants were diagnosed previously with IBD by means of standard radiological, endoscopic, and histopathological investigation, in addition to evaluation by the respective treating physician. Phenotypic data, which included information about physical characteristics, medical history [including surgery within the GI tract], and medication use, were gathered using medical records, and food questionnaires were used to obtain additional information on dietary intake; 57 of the subjects had an ileostomy or ileoanal pouch, forming the small intestinal group. Metagenomic and phenotypic data were also available for 1178 Lifelines DEEP participants. Phenotypic data were collected through participant questionnaires which included questions concerning [GI-related] medical history, medication use, and diet. From the dietary data, only information on daily macronutrient intake [ie, percentage total energy intake from animal protein, plant protein, fat, carbohydrates, etc.] was included in this study, to identify potential confounding effects of specific dietary groups on the microbiota structure and thus correct for the effects of interindividual differences in dietary intake on species abundance.

All participants signed a form of informed consent before sample collection. Institutional ethics review board [IRB] approval was obtained for both cohorts from the UMCG IRB; Lifelines DEEP [ref. M12.113965] and 1000 IBD [IRB number 2008.338].

### 2.2. Group stratification and description

Participants were stratified into four groups according to their intestinal physiology and respective cohort at the time of faecal sampling:

- 1] general population [ $n = 1178$ ]: Lifelines DEEP participants for whom both phenotypic and microbiome data were available;
- 2] IBD non-resected intestine [IBD-NoRes;  $n = 309$ ]: 1000IBD participants without any form of intestinal resection;
- 3] IBD resected intestine [IBD-Res;  $n = 169$ ]: 1000IBD participants who had at least one segmental intestinal resection [ie, small intestinal, ileocaecal valve, or colonic];
- 4] IBD small intestine [IBD-SI;  $n = 57$ ]: 1000IBD participants who had either an ileostomy [ $n = 48$ ] or ileoanal pouch [ $n = 9$ ].

### 2.3. Faecal sample collection and metagenomic sequencing

All faecal samples were collected according to the same protocol, which has been previously described.<sup>14,15</sup> Furthermore, all samples were collected during the same time period, handled by the same group of technicians, and processed using the same protocols and machines. In short, all participants were asked to collect and freeze [at  $-20^{\circ}\text{C}$ ] their faecal samples at home, within 15 min of faeces production. Samples were subsequently collected from the participant's house, transported on dry ice, and stored in the laboratory at  $-80^{\circ}\text{C}$  to minimise any technical confounders. Microbial DNA was isolated from the samples using Qiagen AllPrep DNA/RNA Mini kit [Qiagen; cat. #80204] in combination with mechanical lysis. Isolated DNA was sent to the Broad Institute [Boston, MA, USA] for metagenomic shotgun sequencing [MGS] using the Illumina HiSeq platform. Low-quality reads were filtered out at the sequencing facility.

#### 2.4. Microbiome characterisation

Metagenomic sequencing reads that mapped to the human genome or aligned to Illumina adapters were identified and removed using *KneadData* [v 0.4.6.1]. Biobakery pipeline tools, *MetaPhlan2* [v 2.2]<sup>16</sup> and *HUMAnN2* [v 0.10.0],<sup>17</sup> were applied to the resulting reads to generate taxonomic and microbial pathway abundance profiles, respectively. The taxonomic profiles were subsequently processed as follows: 1) taxa below the level of species [ie, strain] were removed due to inaccurate strain profiles; 2) taxa present in less than 15% of the samples were removed and analysed separately in a logistic regression analysis whereby the abundance values were converted to a binary trait, namely 1 for non-zero values [or ‘presence’] and 0 for zero values [or ‘absence’]; 3) relative abundance values were normalised using arcsine square root transformation. Microbial pathway abundance values were converted to relative abundance and  $\log_{10}$  transformed. Pathways present in fewer than 15% of samples were filtered out.

#### 2.5. Microbial diversity and community description: diversity and composition

Alpha diversity was determined per group by calculating the Shannon index for each sample using the *diversity* function [*index = ‘shannon’*] and Bray-Curtis dissimilarities were calculated using the *vegdist* function [*method = ‘bray’*], also from the R *vegan* [v 2.5–6] package.

#### 2.6. The gut microbiota within small intestinal samples

Differences in the colonic microbiome of UC and CD patients have been reported, as well as dysbiosis in the pouch microbiome of individuals with an ileoanal pouch due to UC. To explore these host-related factors within the small intestinal group, we carried out association analyses using the Wilcoxon test, comparing species relative abundance between: 1) CD vs UC samples; 2) ileostomy vs ileoanal pouch samples; and 3) samples with a colon-only disease location vs ileal [with or without colonic involvement] disease location.

#### 2.7. Phenotypic influences on microbial communities in the small intestine vs colon

To evaluate the relationship between host phenotypes and microbial interindividual variation [represented as Bray-Curtis dissimilarities] within the different groups we performed three PERMANOVA analyses: IBD-SI samples only, IBD-NoRes and IBD-Res samples combined, and general population samples only. Each test was performed using the *adonis* function from the R *vegan* package [*permutations = 1000, method = ‘bray’*].

Next, we performed a univariate correlation analysis between a total of 120 host-related phenotypes and species abundance, using the total samples in this study, to identify potential phenotypic confounders. The Wilcoxon test was used for categorical phenotypes and Spearman correlation for numerical. Phenotypes with most associations were selected for subsequent multivariate analyses [Table S1, available as Supplementary data at ECCOC-JCC online; see following section]. The relationship between number of intestinal resections, as well as resection location [ileal vs colonic] and species abundance, was additionally analysed within the IBD-Res group using the same univariate tests.

#### 2.8. Bacterial composition and metabolic potential in the small intestine

To characterise the microbial composition and metabolic potential in the small intestine, we performed multivariate linear model analyses for the following comparisons:

- i) IBD-SI vs general population;
- ii) IBD-SI vs IBD-NoRes;
- iii) IBD-SI vs IBD-Res;
- iv) IBD-Res vs IBD-Res [taxa only].

The multivariate analyses were performed using generalised linear models as implemented in the R *MaAsLin* [v 0.0.5] package,<sup>18</sup> allowing the boosted feature selection step. The processed taxonomic or pathway data generated from the metagenomes, plus the selected phenotypes, were used as input [see previous sections; Table S1]. Selected phenotypes, such as sequencing read depth [Figure S1, available as Supplementary data at ECCOC-JCC online], were included in the model to correct for the potential effects of interindividual variation in these phenotypes on the microbiota structure and function. All default arguments were used with the exception of two filtering parameters [*dMinAbd = 0* and *dMinSamp = 0*]. Multiple testing corrections were applied using the false discovery rate [FDR] < 0.05. Codes used for the analyses can be accessed via the following link: [https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/tree/master/Projects/Small\_Intestine].

#### 2.9. Exploration of low prevalence bacteria

As part of the quality control in the previous analyses, species present in less than 15% of the total samples [*n = 1713*] were filtered out to better deal with zero inflation. A separate logistic regression analysis was performed on the filtered out [ie, low prevalence] species, between the IBD-SI group and the remaining three groups combined [ie, general population, IBD-NoRes, and IBD-Res]. Relative abundances were coded as 0 for absence [zero values] and 1 for presence [non-zero values]. Age and sex were included in the model as covariates and *p*-values were corrected for multiple testing [FDR < 0.05].

### 3. Results

#### 3.1. Study cohort clinical characteristics

The study cohort consisted of four groups: general population, IBD patients without resections [IBD-NoRes], IBD patients with resections [IBD-Res], and IBD small intestine [IBD-SI]. Average age and body mass index [BMI] were comparable between the groups [*p* > 0.05] [Table 1; Table S2, available as Supplementary data at ECCOC-JCC online]. The IBD-SI group had a significantly larger proportion of females compared with both the general population and the IBD groups [proportion females = 74%, 58%, and 60%, respectively; *p* < 0.05] and a higher use of proton pump inhibitors [PPI] and antibiotics when compared with the general population group [PPI users = 35% and 8%, respectively; antibiotic users = 5% and 1%, respectively; *p* < 0.05]. Compared with the IBD groups, the IBD-SI group had a significantly larger proportion of individuals with UC and a lower mesalazine use [UC = 37% and 58%, respectively; mesalazine users = 35% and 9%; *p* < 0.05]. Within the IBD-SI group, five individuals [9%] had active ileal disease at the time of faecal sampling.

Table 1. Cohort clinical characteristics.

	General Population		IBD*		SI		Wilcoxon test <i>p</i> -value	
	Average [%] or Count [SD]	NA [n]	Average [%] or Count [SD]	NA [n]	Average [%] or Count [SD]	NA [n]	General Population vs SI	IBD* vs SI
Number of samples [n]	1178	-	478	-	57	-	-	-
Sequencing read depth [SD]	32919455 [12276237]	0	25031114 [10203888]	3	22828376 [10140543]	0	1.33E-09	0.138995323
Sex [female/male]	689/489 [58/42%]	0	285/193 [60/40%]	0	42/15 [74/26%]	0	0.022680447	0.039755647
Age at faecal sampling [SD]	45 [13.6]	0	42.9 [12.8]	0	45.2 [10.9]	0	0.968143509	0.131811149
Body mass index [SD]	25.3 [4.2]	0	25.4 [5.0]	6	26.4 [6.85]	1	0.650352121	0.630972528
Faecal calprotectin level over 200 mg/kg [n/y]	1124/48 [96/4%]	6	238/185 [56/44%]	55	32/10 [76/24%]	1.5	4.03E-09	0.012657476
C-reactive protein levels in mg/L divided by 5 [SD]	NA	1178	1.7 [1.9]	2	1.75 [2.13]	0	NA	0.695116431
Current IBD diagnosis [CD/IBDU/UC]	NA	0	274/29/175 [57/6/37%]	0	23/1/33 [40/2/58%]	0	NA	0.005232666
Disease duration of IBD in years [SD]	NA	1178	11.8 [8.8]	8	16.6 [10.4]	0	NA	0.000410007
IBD disease activity [active/not active]	NA	1178	114/358 [24/76%]	6	5/52 [9/91%]	0	NA	0.008679626
Disease location of IBD [both/colon/ileum]	NA	1178	111/221/97 [26/51/23%]	49	14/36/2 [27/69/4%]	5	NA	0.05559574
Ever had a stoma or ileoanal pouch [n/y]	NA	1178	446/32 [93/7%]	0	0/57 [0/100%]	0	NA	2.27E-71
Previous intestinal resections in IBD [n/y]	NA	1178	309/169 [65/35%]	0	0/57 [0/100%]	0	NA	1.05E-20
Number of intestinal resections in IBD [SD]	NA	1178	0.82 [1.6]	0	2.53 [2.28]	2	NA	1.1E-18
Ileocecal valve in situ [n/y]	NA	1178	123/351 [26/74%]	4	52/3 [95/5%]	2	NA	1.53E-24
Total stools formed per 24hrs [SD]	1.4 [0.7]	47	2.65 [2.34]	0	2.56 [3.77]	0	0.001870973	0.000157164
Steroid use [n/y] [%]	NA	1178	379/82 [82/18%]	17	46/6 [88/12%]	5	NA	0.257620812
Immunosuppressants use [n/y] [%]	NA	1178	256/205 [56/44%]	17	37/15 [71/29%]	5	NA	0.031106634
Anti-TNF $\alpha$ use [n/y]	NA	1178	357/121 [75/25%]	0	50/7 [88/12%]	0	NA	0.029402356
Thiopurines use [n/y]	NA	1178	319/159 [67/33%]	0	51/6 [89/11%]	0	NA	0.000447962
Antidiarrhoea use [n/y]	NA	1178	412/49 [89/11%]	17	49/3 [94/6%]	5	NA	0.271482266
Bile acids use [n/y]	NA	1178	437/24 [95/5%]	17	49/3 [94/6%]	5	NA	0.863254283
ACE-inhibitor use [n/y]	1091/44 [96/4%]	43	455/23 [95/5%]	0	56/1 [98/2%]	0	0.412227688	0.292318977
Angiotensin II receptor antagonist use [n/y]	1101/34 [97/3%]	43	467/11 [98/2%]	0	55/2 [96/4%]	0	0.825236536	0.576077019
Antihistamine use [n/y]	1066/69 [94/6%]	43	461/17 [96/4%]	0	53/4 [93/7%]	0	0.773224751	0.203838823
Antibiotics use [n/y]	1165/13 [99/1%]	0	467/11 [98/2%]	0	54/3 [95/5%]	0	0.006704643	0.185885969
Benzodiazepine derivatives use [n/y]	1107/28 [98/2%]	43	459/19 [96/4%]	0	55/2 [96/4%]	0	0.62427707	0.864119807
Beta-blockers use [n/y]	1115/63 [95/5%]	0	443/35 [93/7%]	0	49/8 [86/14%]	0	0.00594789	0.0783367
Beta-sympathomimetic inhaler use [n/y]	1070/65 [94/6%]	43	463/15 [97/3%]	0	54/3 [95/5%]	0	0.882948381	0.400759801

Table 1. Continued

	General Population			IBD*			SI			Wilcoxon test <i>p</i> -value	
	Average [%] or Count [SD]	NA [n]	Average [%] or Count [SD]	Average [%] or Count [SD]	NA [n]	Average [%] or Count [SD]	Average [%] or Count [SD]	NA [n]	General Population vs SI	IBD* vs SI	
Bisphosphonates use [n/y]	1125/10 [99/1%]	43	464/14 [97/3%]	0	0	53/4 [93/7%]	0	0	0.0000274	0.105948295	
Calcium channel blocker use [n/y]	1114/21 [98/2%]	43	470/8 [98/2%]	0	0	55/2 [96/4%]	0	0	0.374611026	0.334008032	
Calcium use [n/y]	1164/14 [99/1%]	0	395/83 [83/17%]	0	0	47/10 [82/18%]	0	0	2.49E-18	0.973008941	
Iron preparations use [n/y]	1171/7 [99/1%]	0	460/18 [96/4%]	0	0	51/6 [89/11%]	0	0	7.34E-13	0.019880859	
Folic acid use [n/y]	1171/7 [99/1%]	0	444/34 [93/7%]	0	0	53/4 [93/7%]	0	0	0.000000468	0.978869397	
Laxatives use [n/y]	1114/21 [99/1%]	43	447/31 [94/6%]	0	0	53/4 [93/7%]	0	0	0.007918352	0.878043045	
Mesalazines use [n/y]	1129/6 [99/1%]	43	310/168 [65/35%]	0	0	52/5 [91/9%]	0	0	2.17E-10	0.0000582	
Metformin use [n/y]	1162/16 [99/1%]	0	472/6 [99/1%]	0	0	55/2 [96/4%]	0	0	0.185963733	0.185554342	
NSAID use [n/y]	1093/42 [96/4%]	43	448/30 [94/6%]	0	0	55/2 [96/4%]	0	0	0.940326774	0.40539998	
Opiates use [n/y]	1122/13 [99/1%]	43	473/5 [99/1%]	0	0	55/2 [96/4%]	0	0	0.118446382	0.122305559	
Oral contraceptive use [n/y]	1019/116 [90/10%]	43	420/58 [88/12%]	0	0	56/1 [98/2%]	0	0	0.036135809	0.018152845	
Oral steroid use [n/y]	1173/5 [99/1%]	0	384/94 [80/20%]	0	0	47/10 [82/18%]	0	0	1.05E-30	0.702312302	
Antidepressants use [n/y]	1125/10 [99/1%]	43	457/21 [96/4%]	0	0	53/4 [93/7%]	0	0	0.0000274	0.375357542	
Paracetamol use [n/y]	1166/12 [99/1%]	0	434/44 [91/9%]	0	0	55/2 [96/4%]	0	0	0.082979844	0.14742216	
Platelet aggregation inhibitor use [n/y]	1101/34 [97/3%]	43	451/27 [94/6%]	0	0	53/4 [93/7%]	0	0	0.091814089	0.676119002	
Proton pump inhibitor use [n/y]	1079/99 [92/8%]	0	366/112 [77/23%]	0	0	37/20 [65/35%]	0	0	2.64E-11	0.053880957	
SSRI-antidepressant use [n/y]	1106/29 [97/3%]	43	469/9 [98/2%]	0	0	57/0 [100/0%]	0	0	0.221991439	0.29657215	
Statin use [n/y]	1079/56 [95/5%]	43	447/31 [94/6%]	0	0	54/3 [95/5%]	0	0	0.910998854	0.720945956	
Steroid inhaler use [n/y]	1078/57 [95/5%]	43	459/19 [96/4%]	0	0	55/2 [96/4%]	0	0	0.607418172	0.864119807	
Steroid nose spray [n/y]	1079/56 [95/5%]	43	473/5 [99/1%]	0	0	55/2 [96/4%]	0	0	0.625700166	0.122305559	
Thiazide diuretic use [n/y]	1092/43 [96/4%]	43	466/12 [97/3%]	0	0	56/1 [98/2%]	0	0	0.426920135	0.726271253	
Levothyroxine use [n/y]	1109/26 [98/2%]	43	468/10 [98/2%]	0	0	55/2 [96/4%]	0	0	0.553695825	0.495158799	
Tricyclic antidepressant use [n/y]	1124/11 [99/1%]	43	467/11 [98/2%]	0	0	57/0 [100/0%]	0	0	0.455430863	0.247608241	
Triptans use [n/y]	1115/20 [98/2%]	43	473/5 [99/1%]	0	0	57/0 [100/0%]	0	0	0.312355796	0.438298036	
Vitamin B12 use [n/y]	1168/10 [99/1%]	0	387/91 [81/19%]	0	0	44/13 [77/23%]	0	0	5.01E-33	0.497073848	
Vitamin D use [n/y]	1164/14 [99/1%]	0	403/75 [84/16%]	0	0	44/13 [77/23%]	0	0	1.2E-27	0.171099283	

Bold values indicate *p*-value < 0.05.

\*IBD = IBD-NoRes + IBD-Res.



### 3.2. Bacterial species profiles are similar within small intestine group

To test whether IBD subtype [CD vs UC], inflammation location [ileal with or without colonic vs colonic only], or the presence of an ileoanal pouch versus an ileostomy were associated with gut microbial alterations in the IBD-SI group, we conducted association analyses between the respective phenotypes and species abundances. We did not identify significant associations in any of the three analyses [FDR > 0.05; Tables S3-S5, available as Supplementary data at ECCOC-JCC online].

### 3.3. The small intestinal microbiota is characterised by lower microbial richness and a distinct bacterial composition

On average, samples belonging to the IBD-SI group had a lower microbial richness when compared with the other groups [Shannon Index<sub>mean</sub> IBD-SI = 1.71; Shannon Index<sub>mean</sub> IBD-Res = 2.44,  $p = 5.10 \times 10^{-14}$ ; Shannon Index<sub>mean</sub> IBD-NoRes = 2.77,  $p = 2.22 \times 10^{-16}$ ; Shannon Index<sub>mean</sub> General population = 2.84,  $p = 2.22 \times 10^{-16}$ ] [Figure 1a]. To get an overview of the bacterial compositions between the groups, we measured the beta diversity using Bray-Curtis dissimilarity [Figure 1b and c; Table S6, available as Supplementary data at ECCOC-JCC online]. Samples from the IBD-SI group on average clustered furthest away from general population samples. IBD-Res and IBD-NoRes samples formed a gradient between IBD-SI and general population samples, with the IBD-Res samples positioning slightly more towards SI samples. Among all samples, IBD-SI samples explained 7.2%, and among IBD samples the presence of intestinal resections explained 5.6%, of the compositional dissimilarities [ $p = 0.001$ ]. The differences in microbial richness and overall bacterial composition between the groups remained significant after correcting for potential confounders [FDR < 0.05; Tables S7 and S8, available as Supplementary data at ECCOC-JCC online].

### 3.4. The overall genus composition in the small intestine

To characterise the differences observed in the beta-diversity analysis, we compared the top 12 most abundant genera in the general population and IBD-SI samples [Tables S12 and S13, available as Supplementary data at ECCOC-JCC online]. The most abundant genera in the IBD-SI group were *Streptococcus*, *Escherichia*, *Blautia*, *Peptostreptococcaceae noname*, *Clostridium*, *Lactobacillus*, and *Veillonella* [mean relative abundance = 26%, 10%, 8.1%, 6.7%, 5.3%, 5.2%, and 4.8%, respectively]. Except for *Blautia*, all abundances were significantly higher when compared with the other groups [*Streptococcus*: IBD-SI vs general population, FDR =  $2.73 \times 10^{-23}$ ; IBD-SI vs IBD-NoRes, FDR =  $6.39 \times 10^{-17}$ ; IBD-SI vs IBD-Res, FDR =  $3.73 \times 10^{-14}$ ; see Supplementary Table S13 for a complete table of values] [Figure 2b and d]. Notably, the IBD-Res group had the second highest total mean abundance of the genera and the general population had the lowest. The reverse trend was seen for the most abundant genera in the general population, which included *Bifidobacterium*, *Ruminococcus*, *Eubacterium*, *Subdoligranulum*, and *Faecalibacterium* [mean relative abundance = 15%, 15%, 14%, 5.8%, 5.6%, and 5.2%, respectively]; the total relative abundance increased in the order: IBD-SI, IBD-Res, IBD-NoRes, and general population [Figure 2a and c; Table S12]. For a more global overview of the microbiota structure in each group, these analyses were also performed at the bacterial phylum, family, and species level [Figure

S3; Tables S9-S11 and S14 and S15, available as Supplementary data at ECCOC-JCC online]. Similar patterns were observed between the four study groups for each taxonomic level.

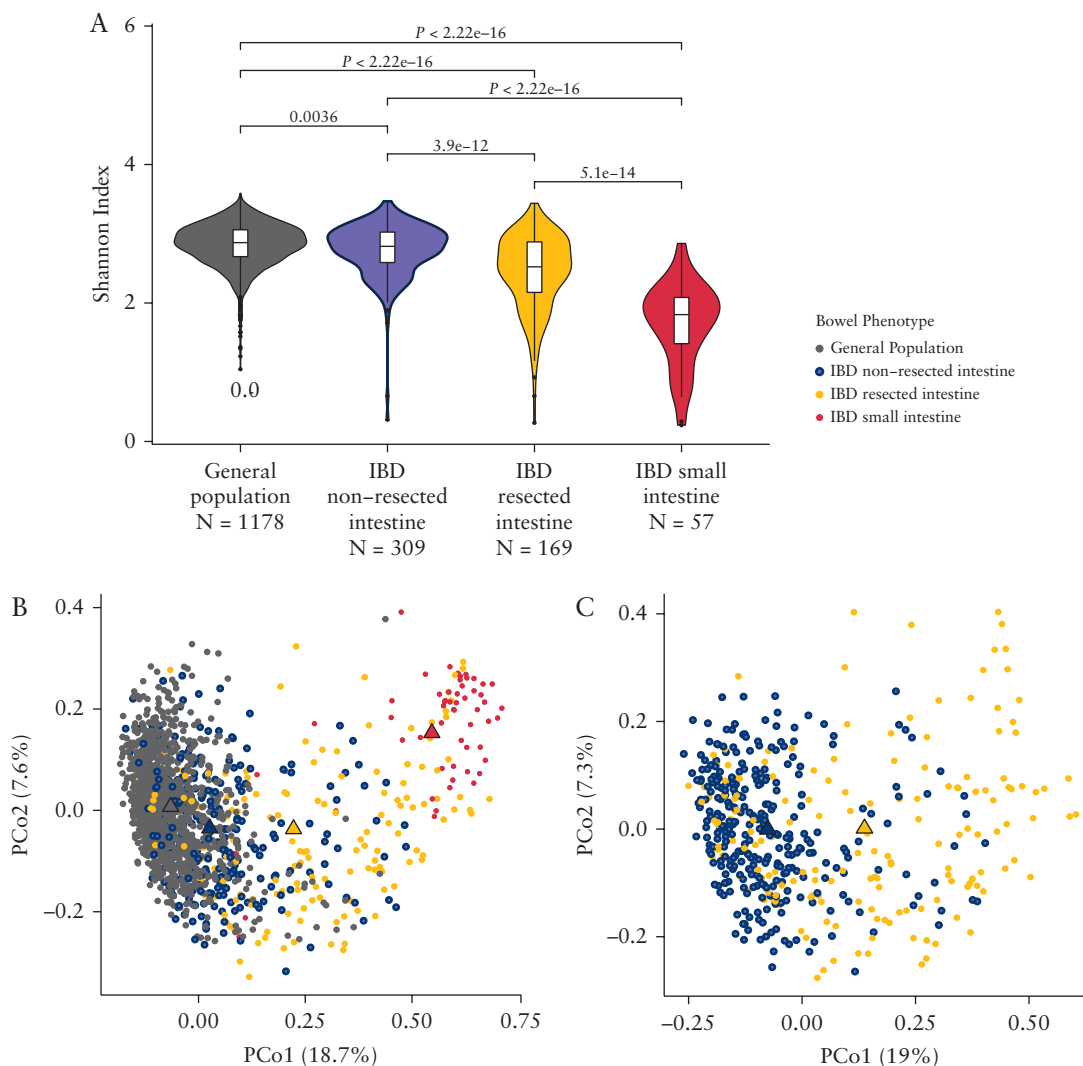
### 3.5. Host-related characteristics associated with the gut bacterial composition

To evaluate potential phenotypes driving differences in the bacterial composition between the groups, we performed correlation analyses between a total of 120 phenotypes and species abundance [Table S16, available as Supplementary data at ECCOC-JCC online]. A total of 3617 associations were identified, involving 106 phenotypes and 134 species [FDR < 0.05]. The phenotype representing current IBD diagnosis had the most associations at 240, involving 108 different species, including *Ruminococcus gnavus* and *Escherichia coli*. Vitamin B12 intake [ $n = 62$ ], sequencing depth [ $n = 62$ ], and PPI use [ $n = 40$ ] were also among the top phenotypes.

Next, we tested if certain phenotypes were specifically associated with the microbial interindividual variation within the IBD-SI group; however, we did not identify any significant associations [FDR > 0.05] [Table S17, available as Supplementary data at ECCOC-JCC online]. Last, given the differences in bacterial composition observed between IBD-NoRes and IBD-Res samples, we asked if the number of intestinal resections, or the location of the resection, is associated with bacterial species abundance within the IBD-Res group. No significant differences were identified for either of the variables [FDR > 0.05] [Tables S18 and S19, available as Supplementary data at ECCOC-JCC online].

### 3.6. *Veillonella*, *Streptococcus*, and *Actinomyces* species are enriched in the small intestine

In total, 89 species were differentially abundant in the IBD-SI group when compared with the general population individuals, 82 compared with IBD-NoRes, and 49 in the comparison between IBD-SI samples and IBD-Res [FDR < 0.05] [Figure 3a; Tables S20-S22, available as Supplementary data at ECCOC-JCC online]. Of the 89 species differentially abundant in the IBD-SI compared with the general population samples, 22 were enriched in the IBD-SI. This included nine belonging to the genus *Streptococcus*, three to *Veillonella*, and three to *Actinomyces* [FDR < 0.05] [Figure 3a; Table S20, available as Supplementary data at ECCOC-JCC online]. Sixty seven species were under-represented in the IBD-SI group, of which six belonged to the genus *Ruminococcus*, eight to *Eubacterium*, 10 to *Bacteroides*, and five to *Alistipes*. Moreover, *Bifidobacterium dentium*, *Actinomyces odontolyticus*, *Streptococcus mutans*, and *Streptococcus salivarius* were exclusively associated with this comparison [FDR < 0.05] [Figure 3a; Table S20]. Of the associations between IBD-SI and IBD-NoRes samples, 11 were unique, including a lower relative abundance of the butyrate producer *Pseudoflavonifractor capillosus* in IBD-SI samples [FDR =  $1.95 \times 10^{-5}$ ] [Figure 3a; Table S21, available as Supplementary data at ECCOC-JCC online]. A lower relative abundance of a *Parabacteroides* species in IBD-SI individuals was only observed when comparing IBD-SI with the IBD-Res group [FDR = 0.047] [Figure 3a; Table S22, available as Supplementary data at ECCOC-JCC online]. *Veillonella atypica*, *Streptococcus mitis oralis pneumoniae*, *Streptococcus infantis*, *Streptococcus sanguinis*, *Actinomyces graevenitzi*, and *Haemophilus parainfluenzae*, which are typically found in the oral cavity, were consistently found to be enriched in the IBD-SI compared with the other groups [FDR < 0.05] [Figure 3a; Table S20-S22]. In the comparison between IBD-Res and IBD-NoRes samples, 19 species were significantly associated



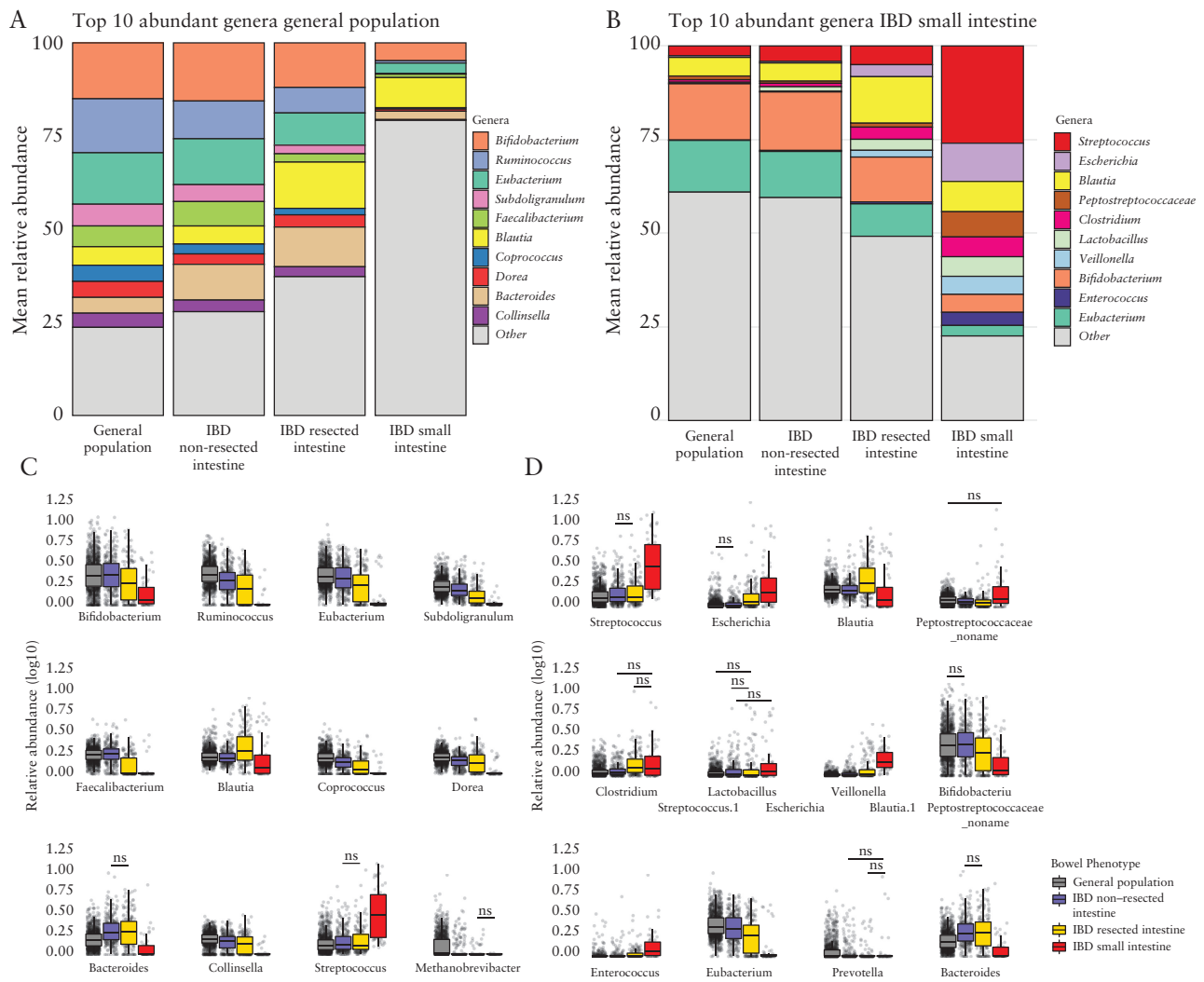
**Figure 1.** Lower bacterial diversity and a distinct composition in the small intestinal, compared with colonic, samples. [a] Violin plots representing the distribution of Shannon index values per study group. Small intestinal samples have, on average, a lower bacterial diversity [mean Shannon index = 1.71] when compared with general population samples, IBD non-resected intestine samples and IBD resected intestine samples [mean Shannon index = 2.84, 2.77, and 2.44, respectively];  $p$ -values were calculated using the two-sided Wilcoxon test. Boxplots show the median and interquartile range [25th and 75th]. Whiskers show the  $1.5 \times IQR$  range. b, c] Scatter plots show the Bray-Curtis dissimilarity between the samples, as a measure of the differences in overall bacterial composition [principal coordinate analysis]. Samples are coloured according to the group classification used throughout this study (grey, general population samples [ $n = 1178$ ]; purple, samples from patients with IBD without intestinal resections [ $n = 309$ ]; yellow, samples from patients with IBD with intestinal resections [ $n = 169$ ]; red, samples representing the small intestine [ $n = 57$ ]). Percentages on the x and y axes represent the total variance explained by each coordinate. Triangles indicate the mean coordinate value per group. Panel [b] highlights the dissimilarities between all the samples used in this study. Small intestinal samples form a defined cluster with little overlap with general population samples. Samples from patients with IBD [purple and yellow] form a gradient between the small intestine and general population clusters. Panel [c] highlights the heterogeneity between IBD samples only. IBD, inflammatory bowel disease; IQR, interquartile range.

[FDR < 0.05] [Figure S4; Table S23, available as Supplementary data at ECCOC-JCC online], three of which were significantly enriched in IBD-Res samples, namely *Ruminococcus gnavus*, *Escherichia coli*, and *Granulicatella unclassified*. The remaining species associated with the comparison were significantly under-represented in IBD-Res samples, including *Faecalibacterium prausnitzii*, *Coprococcus catus*, *Barnesiella intestinihominis*, and *Ruminococcus bromii*.

### 3.7. Rare colonic bacteria are prevalent in the small intestine

When comparing the prevalence of bacterial species that were present in less than 15% of the cohort between the IBD-SI and the other

three study groups combined, we found that 110 of these species were significantly more prevalent in the IBD-SI group [FDR < 0.05; Table S24, available as Supplementary data at ECCOC-JCC online]. Among the most prevalent species, six belonged to the genus *Streptococcus*, three to *Clostridium*, four to *Actinomyces*, three to *Klebsiella*, five to *Lactobacillus*, three to *Gemella*, two to *Atopobium*, and three to *Enterococcus* [prevalence range in IBD-SI = 15–75%; prevalence range in other groups combined = 0.1–13%]. Specific species that were enriched included *Veillonella dispar* [FDR =  $9.39 \times 10^{-23}$ ], *Klebsiella pneumoniae* [FDR =  $1.39 \times 10^{-16}$ ], *Enterococcus faecalis* [FDR =  $2.39 \times 10^{-23}$ ], *Enterococcus faecium* [FDR =  $1.32 \times 10^{-12}$ ], and *Lactobacillus fermentum* [FDR =  $8.37 \times 10^{-12}$ ] [Figure 3b; Table S24].



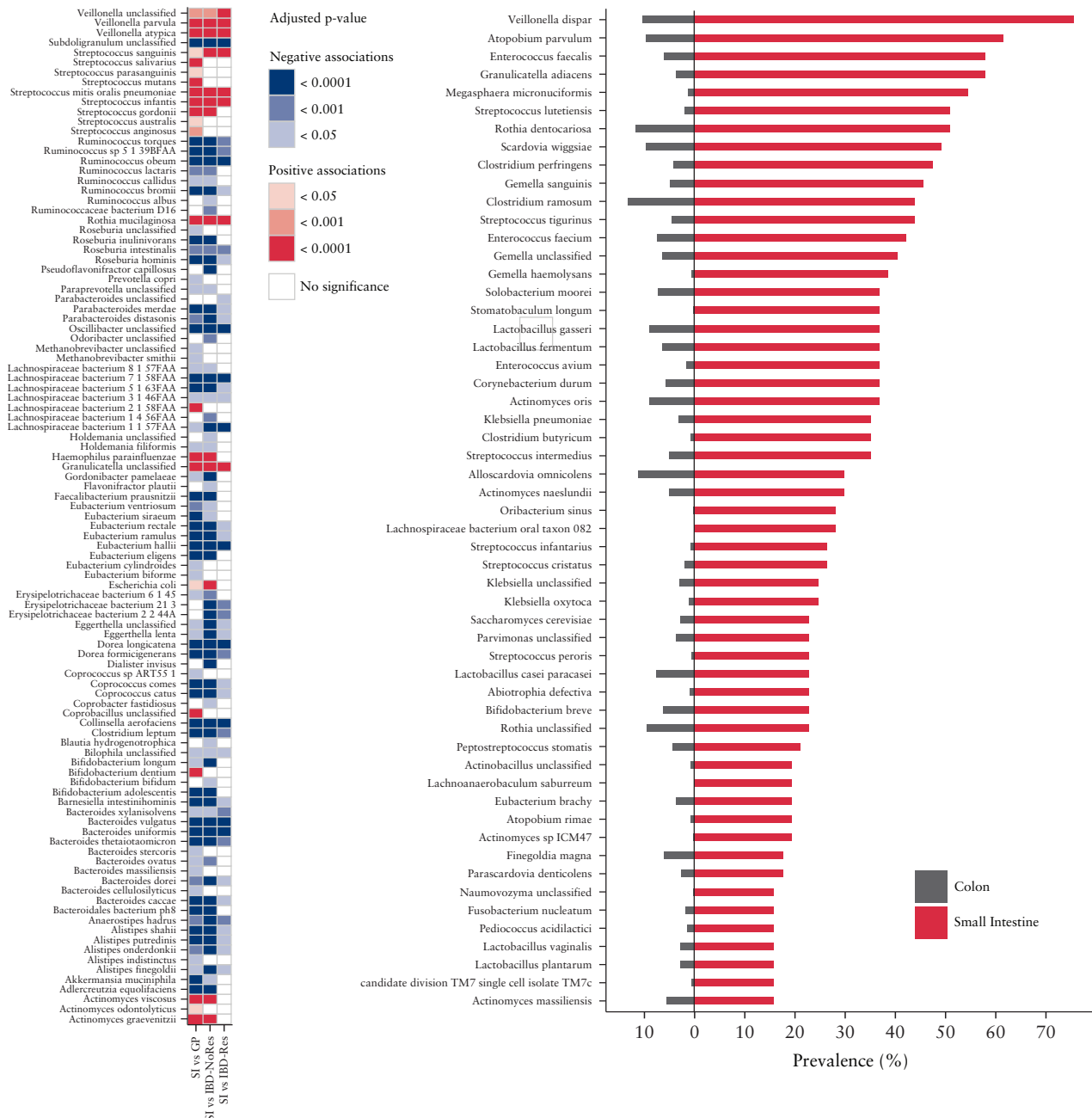
**Figure 2.** Bacterial genus profile is markedly different in samples representing the small intestinal microbiota compared with samples representative of the colon. [a,b] Bar plots representing the relative abundance per study group of the top 10 most abundant genera among the general population samples and among the small intestinal samples, respectively. *Bifidobacterium* and *Eubacterium* are the most abundant bacterial genera in both the general population samples and IBD samples from patients without intestinal resections. In contrast, genera *Streptococcus* and *Escherichia* are most abundant in samples representing the small intestinal content. [c,d] Boxplots showing the average relative abundance per test group of the top 12 abundant bacteria in the general population samples [c] and small intestinal samples [d]. Comparisons marked with 'ns' are not significant. The two-sided Wilcoxon test was used to test significance. Boxplots show the median and interquartile range [25th and 75th]. Whiskers show the  $1.5 \times IQR$  range. Each grey dot represents one sample [Supplementary tables S12 and S13]. IBD, inflammatory bowel disease; IQR, interquartile range; ns, non-significant.

### 3.8. The small intestinal microbiota is largely characterised by pathways involved in sugar metabolism and quinone, haeme, fatty acid and lipid biosynthesis

To investigate the functional potential of the small intestinal microbial community and its possible role in IBD, we analysed the relative abundance of 341 predicted metabolic pathways that were present in at least 15% of the total samples. Of these, 252 [74%] of the pathways were associated with at least one of the test comparisons: 243 pathways in the comparison IBD-SI vs general population; 147 in the comparison IBD-SI vs IBD-NoRes; and 65 in the comparison IBD-SI vs IBD-Res [FDR < 0.05] [Figure 4; Tables S25-27, available as Supplementary data at ECCOC-JCC online]. Of all the pathways associated, 52 were associated with all three tests. Examples included an increase in pathways related to sugar degradation, fermentation

to lactate and quinone, haeme, fatty acid and lipid biosynthesis, and an under-representation of pathways involved in degradation of complex carbohydrates and pyruvate fermentation to propanoate and butanoate [FDR < 0.05] [Figure 4; Tables S25-S27]. As expected, we observed that pathways which clustered together by Euclidean distance showed, in general, similar associations with the respective test comparison. Thus pathways enriched in the small intestinal samples tended to cluster together and those under-represented tended to cluster together. Pathways that were exclusively enriched in the IBD-SI compared with general population samples were also related to sugar [derivatives] degradation and energy metabolism, as well as nucleotide, nucleoside, and biotin biosynthesis [FDR < 0.05] [Figure 4; Table S25]. Conversely, pathways exclusively under-represented in the IBD-SI were related to methanogenesis and pantothenate biosynthesis and amino acid biosynthesis pathways, both increased and decreased [FDR < 0.05; Figure 4, Table S25]. Pathways that were





**Figure 3.** Bacterial composition in the small intestine. [a] Heatmap showing the bacterial species significantly enriched [red] or under-represented [blue] in the small intestinal group compared with at least one of the three other study groups [general population, GP; IBD non-resected intestine, IBD-NoRes; IBD resected intestine, IBD-Res];  $p$ -values were calculated using multivariate linear regression models [see Methods] and adjusted for multiple testing [FDR < 0.05] [Supplementary tables S20-S22]. [b] Barplot showing species prevalence in colonic samples [grey] versus small intestinal samples [red], of the species present in less than 15% of the total samples. Only the species differentially prevalent between the two groups, with a prevalence of 15% or more in at least one of the groups, are plotted. Logistic regression was used to test significance [see Methods] [Supplementary table S24]. IBD, inflammatory bowel disease; FDR, false-discovery rate.

associated with the comparison between IBD-SI and general population or IBD-NoRes group [FDR < 0.05], but similar in abundance between IBD-SI and IBD-Res samples [FDR > 0.05], were involved in methylglyoxal and arginine degradation, biotin and quinone biosynthesis, sugar metabolism, butanoate production, and endotoxin biosynthesis, such as enterobacterial common antigen and lipopolysaccharides [Figure 4; Tables S25-S27]. These pathways, except for butanoate production, were enriched in the IBD-SI. Of note,

pathways such as quinone, haeme, fatty acid, and endotoxin biosynthesis suggest a pro-inflammatory potential.

#### 4. Discussion

In this study, we explored the bacterial composition and metabolic potential of the human small intestinal microbiota and have highlighted its potential implications in IBD. While correcting for



potential phenotypic confounders, we analysed metagenomes derived from faecal samples from 57 individuals with an ileostomy or ileoanal pouch, following colonic resection due to IBD, in comparison with metagenomes from general population individuals and patients with IBD with or without a history of intestinal resections.

We found that samples belonging to the small intestine group had a significantly lower bacterial diversity as compared with the other groups. Small intestinal samples were also visibly distinct from samples representing the general population, in terms of overall bacterial composition expressed as Bray-Curtis dissimilarities. These findings highlight the known physiological differences observed between the small intestine and colon which can drive bacterial selection. The small intestine, for example, is known to be a harsh environment for microbial existence due to its acidic environment, higher oxygen concentrations, short transit times, and regular inflow of digestive enzymes and bile.<sup>13,19–21</sup>

Bacterial species that were markedly enriched in the small intestine as compared with the faecal microbiota of the general population included *Veillonella atypica*, *Streptococcus mitis oralis pneumoniae*, *Streptococcus salivarius*, *Bifidobacterium dentium*, *Haemophilus parainfluenzae*, and *Actinomyces graevenitzi*. Additionally, species belonging to genera such as *Clostridium*, *Lactobacillus*, *Klebsiella*, *Gemella*, and *Enterococcus*, which were by rarely observed in the general population faecal samples, had a significantly higher prevalence of between 15% and 75% in the small intestinal samples. These results suggest a specific small intestinal niche formed by these bacteria. Consistent with our results, *Veillonella*, *Streptococcus*, *Actinomyces*, *Gemella*, *Clostridium*, and *Lactobacillus* species have also been identified by other small intestinal microbiome studies.<sup>6,13,19,22,23</sup> We also observed that the bacterial richness was significantly lower in patients with IBD and intestinal segmental resections compared with those without resections. Composition differences were also observed between the two groups, such as an enrichment of *Escherichia coli* and *Ruminococcus gnavus* and reduced *Faecalibacterium prausnitzii* and *Ruminococcus bromii*. These findings are consistent with the study of Yilmaz *et al.*, in which reduced species richness, increased Enterobacteriaceae, and reduced Ruminococcaceae were reported in the microbiota of operated, compared with non-operated, patients with CD.<sup>24</sup>

The use of metagenomes also allowed us to study the predicted metabolic potential of the small intestinal microbiota. In line with the findings of Zoetendal *et al.*, we identified an enrichment of microbial pathways related to simple carbohydrate degradation and fermentation and energy metabolism in the SI compared with the general population, including biotin biosynthesis pathways.<sup>19</sup> Biotin, also called vitamin B<sub>7</sub> or B<sub>8</sub>, is an important cofactor for several carboxylases that are essential for glucose, amino acid, and fatty acid metabolism.<sup>25</sup> Biotin is also thought to have anti-inflammatory effects by inhibiting expression of NF-κB, a pro-inflammatory signalling molecule. Although gut bacteria-derived biotin is mostly absorbed in the colon, our results indicate that biotin biosynthesis is performed to a larger extent in the small intestine. Moreover, bacteria belonging to the phyla Proteobacteria, Fusobacteria, and Bacteroidetes are reported to possess a biotin biosynthesis pathway, which is consistent with our observation that Bacteroidetes and Proteobacteria were overall more abundant in small intestinal samples relative to general population samples.<sup>26</sup> We noted at least four pathways related to fatty acid and lipid metabolism which were more abundant in the small intestine. This is in accordance with studies demonstrating the importance of small intestinal bacteria in intestinal lipid digestion and absorption.<sup>27</sup> Moreover, we also observed an enrichment in

small intestinal samples of *E. coli* and *Lactobacillus casei paracasei*, which have been shown to alter enterocyte lipid metabolism via their secretion of acetate and L-lactate, respectively.<sup>28</sup>

When comparing the IBD-SI group with IBD-NoRes and IBD-Res groups, fewer species [ $n = 82$  and  $49$ , respectively] were associated, as when compared with the general population group [ $n = 89$ ], suggesting increased colonisation of certain small intestinal bacteria in the IBD colon. Examples include *Bifidobacterium dentium*, *Actinomyces odontolyticus*, *Streptococcus mutans*, *Streptococcus salivarius*, and *Haemophilus parainfluenzae*, which, with the exception of *B. dentium*, have been previously associated with IBD and/or intestinal complications.<sup>17,29–32</sup> In fact, many other bacteria enriched in the small intestine compared with the other groups have been associated with IBD. Examples include *Veillonella spp.*, *Streptococcus spp.*, *Enterococcus faecalis*, *Enterococcus faecium*, and *Klebsiella pneumoniae*.<sup>12,17,29,33,34</sup> On a functional level, fewer pathways were associated with the comparison between the IBD-SI and the two IBD groups ( $n = 147$  [IBD-SI vs IBD-NoRes],  $65$  [IBD-SI vs IBD-Res], and  $243$  [IBD-SI vs general population]). Pathways involved in lactate and acetate production and degradation of arginine, which were enriched in IBD-SI samples compared with the general population, were no longer associated with the IBD-SI vs IBD-Res comparison. This is in line with reports of elevated abundances of lactate, as well as lactate-producing bacteria [eg, *Lactobacilli*, *Enterococci*, *Streptococci*, and *Pediococci*] in faecal samples of patients with IBD.<sup>17</sup> Similarly, pantothenate [vitamin B<sub>5</sub>] biosynthesis and methanogenesis pathways were under-represented in the small intestine group when compared with the general population, but not when compared with the IBD groups. Pantothenate metabolites have been previously found to be decreased in IBD faecal samples.<sup>17</sup> Vitamin B<sub>5</sub> is absorbed in the colon and its deficiency has been associated with the production of pro-inflammatory molecules.<sup>25</sup> Methanogenesis is the formation of methane from hydrogen and carbon dioxide. Methane has been reported to slow intestinal transit and thus reduced methanogenesis is consistent with the shorter transit times observed in the small intestine.<sup>35</sup> Reduced methanogenesis in the colon may, however, contribute to the development of diarrhoea, which is a common symptom of IBD. We also observed an enrichment in the small intestine and IBD colon of a lactose/galactose degradation pathway whereby hydrogen is produced. Hydrogen has been demonstrated to shorten colonic transit times, predominantly in the proximal colon.<sup>35</sup> Taken together, these results support a role for small intestinal, rather than per se oral, pathobionts in IBD disease pathogenesis.

Whereas the results of this study offer a detailed insight into the small intestinal microbiota and its possible implications in IBD, there are some limitations that need to be addressed. Due to the cross-sectional nature of this study, we were not able to take temporal variation of the gut microbiota into account. Functional experiments such as culturomics and animal models are therefore still required to provide causal validation and a mechanistic understanding of the implications of these bacteria in the pathogenesis of IBD. Additionally, untargeted metabolomics data integration will help to better understand the significance of the microbial pathway results presented in this study.

Furthermore, our entire IBD-SI group consisted of individuals with an IBD context. Although 'healthy' individuals with an ileostomy or ileoanal pouch do not exist, replicating the findings in non-IBD patients with an ileostomy would be beneficial to studying the small intestinal gut microbiota non-invasively. Last, one might argue that the individuals within our small intestine group are

heterogeneous due to, for example, the inclusion of patients with pouches.<sup>36–38</sup> We compared the bacterial communities between ileostomy and pouch-derived faecal samples and found no significant differences in the relative bacterial abundances between the two groups. We also did not identify any associations between IBD subtypes or the location of inflammation and the abundance of bacterial species.

Overall, we have provided a high-resolution description of the bacterial composition and potential metabolic functions characteristic of the small intestinal microbiota. Moreover, we have shown that the colonic content in a subset of patients with IBD resembles the distinct small intestinal microbiome, suggesting the translocation of small intestinal pathobionts to the colon. Further supporting this, we observed that the small intestinal microbiome harbours potentially pathogenic features that could be relevant for IBD pathogenesis, and ultimately future targets for therapeutic intervention. Instead of focusing on the faecal microbiome and the role of oral bacteria, it is worth turning our attention and efforts towards elucidating the mechanisms that define the small intestinal microbiota and its interaction with the host, to better understand health maintenance and disease development.

The raw sequences of the faecal metagenomes are available at the European Genome-phenome Archive data repository upon request. The 1000IBD cohort data are available via the following link: [<https://www.ebi.ac.uk/ega/datasets/EGAD00001004194>]. LifeLines DEEP data are available via this link: [<https://www.ebi.ac.uk/ega/datasets/EGAD00001001991>]. The codes used for microbial profiling are available via the following link: [[https://github.com/WeersmaLabIBD/Microbiome/blob/master/Protocol\\_metagenomic\\_pipeline.md](https://github.com/WeersmaLabIBD/Microbiome/blob/master/Protocol_metagenomic_pipeline.md)]. The codes used for the statistical analyses performed in this study are available via this link: [[https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/tree/master/Projects/Small\\_Intestine](https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/tree/master/Projects/Small_Intestine)].

## Funding

RW and AZ are supported by VIDI grants [016.136.308 and 016.178.056] from the Netherlands Organization for Scientific Research [NWO]. RKW is further supported by a Diagnostics Grant from the Dutch Digestive Foundation [D16-14] and received unrestricted grants from Takeda, Johnson and Johnson, Tramedico, Ferring and acted as consultant for Takeda. AZ holds a Rosalind Franklin fellowship from the University of Groningen. AZ is further supported by the Dutch Heart Foundation IN-CONTROL [CVON2018-27], an ERC starting grant [ERC-715772] and the NWO Gravitation grant Exposome-NL [024.004.017]. JF is supported by the Dutch Heart Foundation IN-CONTROL [CVON2018-27], the ERC Consolidator grant [grant agreement No. 101001678] and the Netherlands Organ-on-Chip Initiative, an NWO Gravitation project [024.003.001] funded by the Ministry of Education, Culture and Science of the government of The Netherlands. Sequencing of the Lifelines DEEP cohort was funded by a grant awarded to CW from the Netherlands' Top Institute Food and Nutrition GH001. CW is further supported by an ERC advanced grant [ERC-671274] and a Spinoza award [NWO SPI 92-266].

## Conflict of Interest

Nothing to declare.

## Acknowledgments

We would like to thank all patients with IBD from the 1000IBD cohort and participants from the general population cohort, LifeLines DEEP, for faecal sample collection and providing additional phenotypic data. Furthermore, we would like to thank the IBD doctors and nurses at the UMCG IBD clinic for patient inclusion.

## Author Contributions

AVV and RKW designed the study. BHJ provided logistic and laboratory support. RAAAR, AVV, VC, and MAY gathered and prepared the data. RAAAR and AVV analysed the data. RAAAR, AVV, VC, and LB wrote the manuscript. MAY, PS, BHJ, MDV, JF, CW, AZ, and RKW critically reviewed the manuscript.

## Supplementary Data

Supplementary data are available at ECCO-JCC online. Supplementary tables can be accessed via the following link: [[https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/blob/master/Projects/Small\\_Intestine/Supplementary\\_Tables\\_JCC\\_Final\\_REVISED.xlsx.zip](https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/blob/master/Projects/Small_Intestine/Supplementary_Tables_JCC_Final_REVISED.xlsx.zip)].

## References

- Brestoff JR, Artis D. Commensal bacteria at the interface of host metabolism and the immune system. *Nat Immunol* 2013;14:676–84.
- Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. *Cell* 2012;148:1258–70.
- Ni J, Wu GD, Albenberg L, Tomov VT. Gut microbiota and IBD: causation or correlation? *Nat Rev Gastroenterol Hepatol* 2017;14:573–84.
- Collins SM, Surette M, Bercik P. The interplay between the intestinal microbiota and the brain. *Nat Rev Microbiol* 2012;10:735–42.
- Zmora N, Zilberman-Schapira G, Suez J, *et al*. Personalized gut mucosal colonization resistance to empiric probiotics is associated with unique host and microbiome features. *Cell* 2018;174:1388–405.e21.
- El Aidy S, van den Bogert B, Kleerebezem M. The small intestine microbiota, nutritional modulation and relevance for health. *Curr Opin Biotechnol* 2015;32:14–20.
- Ermund A, Schütte A, Johansson ME, Gustafsson JK, Hansson GC. Studies of mucus in mouse stomach, small intestine, and colon. I. Gastrointestinal mucus layers have different properties depending on location as well as over the Peyer's patches. *Am J Physiol Gastrointest Liver Physiol* 2013;305:G341–7.
- Zhong L, Shanahan ER, Raj A, *et al*. Dyspepsia and the microbiome: time to focus on the small intestine. *Gut* 2017;66:1168–9.
- Saffouri GB, Shields-Cutler RR, Chen J, *et al*. Small intestinal microbial dysbiosis underlies symptoms associated with functional gastrointestinal disorders. *Nat Commun* 2019;10. doi:10.1038/s41467-019-09964-7.
- Schmidt TSB, Hayward MR, Coelho LP, *et al*. Extensive transmission of microbes along the gastrointestinal tract. *Elife* 2019;8. doi:10.7554/eLife.42693.
- Zhang X, Zhang D, Jia H, *et al*. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat Med* 2015;21:895–905.
- Atarashi K, Suda W, Luo C, *et al*. Ectopic colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation. *Science* 2017;358:359–65.
- Kastl AJ Jr, Terry NA, Wu GD, Albenberg LG. The structure and function of the human small intestinal microbiota: current understanding and future directions. *Cell Mol Gastroenterol Hepatol* 2020;9:33–45.
- Imhann F, Van der Velde KJ, Barbieri R, *et al*. Correction to: the 1000IBD project: multi-omics data of 1000 inflammatory bowel disease patients; data release 1. *BMC Gastroenterol* 2019;19:44.
- Tigchelaar EF, Zhernakova A, Dekens JA, *et al*. Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: study design and baseline characteristics. *BMJ Open* 2015;5:e006772.
- Truong DT, Franzosa EA, Tickle TL, *et al*. MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat Methods* 2015;12:902–3.
- Franzosa EA, Sirota-Madi A, Avila-Pacheco J, *et al*. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol* 2019;4:293–305.

18. Morgan XC, Tickle TL, Sokol H, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* 2012;13:R79. doi:10.1186/gb-2012-13-9-r79.
19. Zoetendal EG, Raes J, van den Bogert B, et al. The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J* 2012;6:1415–26.
20. Friedman ES, Bittinger K, Esipova TV, et al. Microbes vs. chemistry in the origin of the anaerobic gut lumen. *Proc Natl Acad Sci U S A* 2018;115:4170–5.
21. Nugent SG, Kumar D, Rampton DS, Evans DF. Intestinal luminal pH in inflammatory bowel disease: possible determinants and implications for therapy with aminosalicylates and other drugs. *Gut* 2001;48:571–7.
22. Booijink CC, El-Aidy S, Rajilić-Stojanović M, et al. High temporal and inter-individual variation detected in the human ileal microbiota. *Environ Microbiol* 2010;12:3213–27.
23. Villmones HC, Haug ES, Ulvestad E, et al. Species level description of the human ileal bacterial microbiota. *Sci Rep* 2018;8:4736.
24. Yilmaz B, Juillerat P, Öyäs O, et al. Microbial network disturbances in relapsing refractory Crohn's disease. *Nat Med* 2019;25:701.
25. Yoshii K, Hosomi K, Sawane K, Kunisawa J. Metabolism of dietary and microbial vitamin B family in the regulation of host immunity. *Front Nutr* 2019;6:48.
26. Magnúsdóttir S, Ravcheev D, de Crécy-Lagard V, Thiele I. Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Front Genet* 2015;6:148.
27. Martinez-Guryn K, Hubert N, Frazier K, et al. Small intestine microbiota regulate host digestive and absorptive adaptive responses to dietary lipids. *Cell Host Microbe* 2018;23:458–69.e5.
28. Araújo JR, Tazi A, Burlen-Defranoux O, et al. Fermentation products of commensal bacteria alter enterocyte lipid metabolism. *Cell Host Microbe* 2020;27:358–75.e7.
29. Gevers D, Kugathasan S, Denson LA, et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 2014;15:382–92.
30. 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:eaap8914.
31. Kojima A, Nakano K, Wada K, et al. Infection of specific strains of *Streptococcus mutans*, oral bacteria, confers a risk of ulcerative colitis. *Sci Rep* 2012;2:332.
32. Nahum A, Filice G, Malhotra A. A complicated thread: abdominal actinomycosis in a young woman with Crohn disease. *Case Rep Gastroenterol* 2017;11:377–81.
33. Seishima J, Iida N, Kitamura K, et al. Gut-derived *Enterococcus faecium* from ulcerative colitis patients promotes colitis in a genetically susceptible mouse host. *Genome Biol* 2019;20:252.
34. Balish E, Warner T. *Enterococcus faecalis* induces inflammatory bowel disease in interleukin-10 knockout mice. *Am J Pathol* 2002;160:2253–7.
35. Jahng J, Jung IS, Choi EJ, Conklin JL, Park H. The effects of methane and hydrogen gases produced by enteric bacteria on ileal motility and colonic transit time. *Neurogastroenterol Motil* 2012;24:185–90, e92.
36. Sinha SR, Haileselassie Y, Nguyen LP, et al. Dysbiosis-induced secondary bile acid deficiency promotes intestinal inflammation. *Cell Host Microbe* 2020;27:659–70.e5.
37. Turpin W, Kelly O, Borowski K, et al. Mucosa-associated microbiota in ileoanal pouches may contribute to clinical symptoms, particularly stool frequency, independent of endoscopic disease activity. *Clin Transl Gastroenterol* 2019;10:1–7.
38. Morgan XC, Kabakchiev B, Waldron L, et al. Associations between host gene expression, the mucosal microbiome, and clinical outcome in the pelvic pouch of patients with inflammatory bowel disease. *Genome Biol* 2015;16:67.