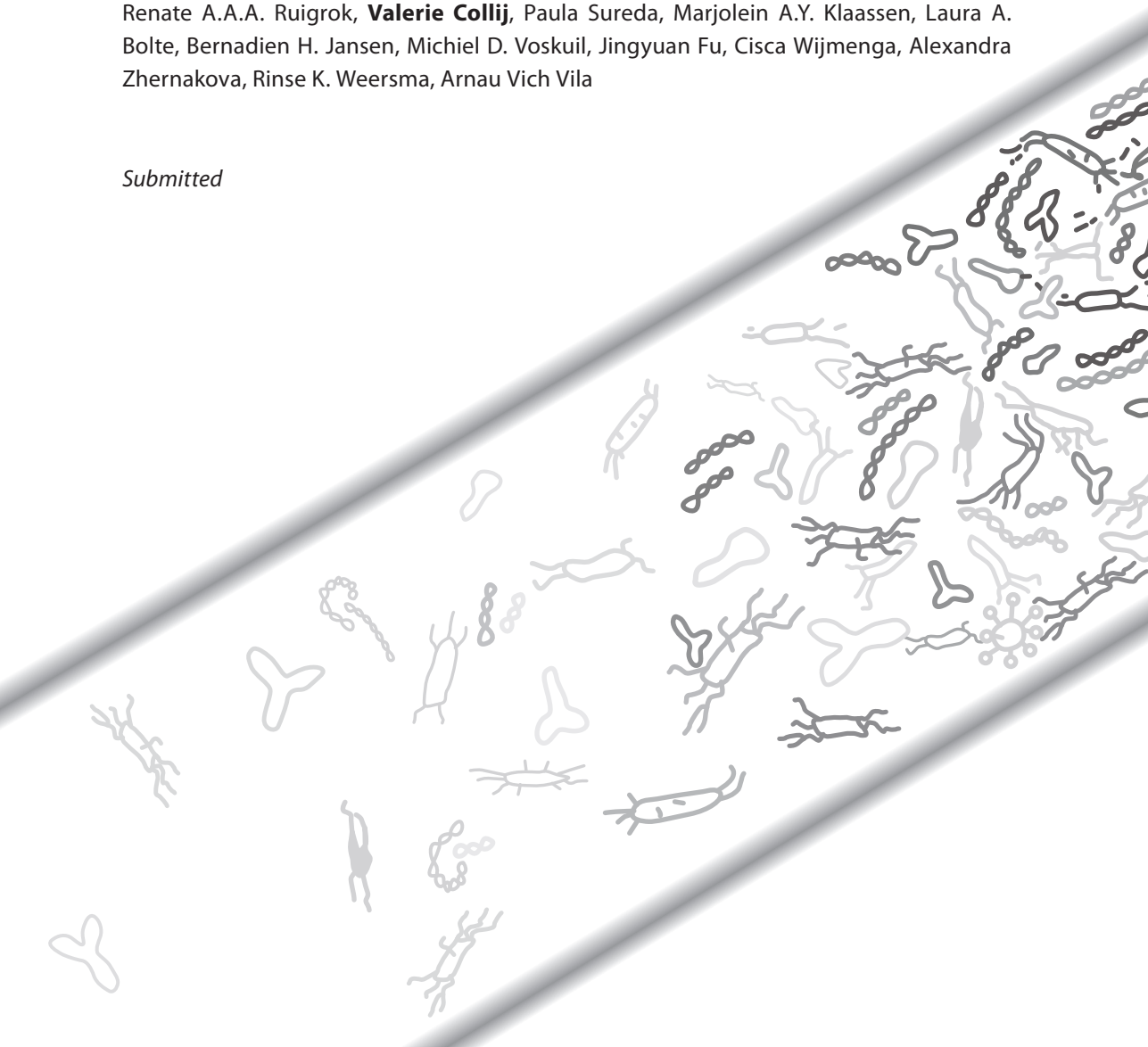
4

The composition and metabolic potential of the human small intestinal microbiota within the context of inflammatory bowel disease

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

Submitted



Abstract

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). 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 *MetaPhlAn2* and *HUMAnN2,* and association analyses were performed using multivariate linear regression. Small intestinal samples had a significantly lower bacterial diversity, compared with the general population and, to a lesser extent, IBD samples. *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 proinflammatory potential. We conclude that the colonic microbiome of IBD patients, particularly with intestinal resections, showed resemblance to that of the small intestine. These results highlight the importance of studying the small intestinal microbiota to add new insight into disease pathogenesis.

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.[1,2] 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.[2-4] However, the use of faecal samples in the majority of these studies has meant that most findings are largely specific to the colonic content.[5] 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.[6,7] Dysbiosis of the duodenal microbiota has been associated with certain GI-related disorders and complaints, such as functional dyspepsia, bloating and diarrhea.[8,9] Additionally, the ectopic colonization of microbes typical of the oral cavity has been hypothesised to play a role in the pathogenesis of several disorders; a phenomenon termed "oralization".[10,11] 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 colonization in the colon.[12] However, oral-considered bacteria 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. Majority of studies to date have relied on using mucosal samples collected either during routine endoscopies, following intestinal resections or from sudden death individuals.[13] However, they 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.

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 faecal microbiota of the general population (n = 1178) and patients with IBD (n = 478).

Materials and methods

Cohort description

To study the small intestinal microbiota and its potential implications in IBD, two independent Dutch cohorts were used: 1) 1000IBD cohort, established at the IBD center at the University Medical Center Groningen (UMCG), the Netherlands;[14] 2) Lifelines DEEP, a general population cohort from the northern provinces of the Netherlands.[15] Metagenomic in combination with phenotypic data was 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, was 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 was also available for 1178 Lifelines DEEP participants. Phenotypic data was collected through participant questionnaires which included questions concerning, (GI-related) medical history, medication use and diet. All participants signed a form of informed consent prior to 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).

Group stratification and description

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

1) General population (n=1178): Lifelines DEEP participants for whom both phenotypic and microbiome data was 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 (i.e. small intestinal, ileocecal valve or colonic).
4) IBD small intestine (SI; n=57): 1000IBD participants who had either an ileostomy (n=48) or ileoanal pouch (n=9).

Faecal sample collection and metagenomic sequencing

All faecal samples were collected in the same manner, which has been previously described.[14,15] In short, all participants were asked to collect and freeze (at -20°C) their faecal samples at home, within 15 minutes of faeces production. Samples were subsequently collected from the participant's house, transported on dry ice and stored in the lab at -80°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, Massachusetts, USA) for metagenomic shotgun sequencing (MGS) using the Illumina HiSeq platform. Low-quality reads were filtered out at the sequencing facility.

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)[16] and *HUMAnN2* (v 0.10.0), were applied to the resulting reads to generate taxonomic and microbial pathway abundance profiles, respectively.[17] The taxonomic profiles were subsequently processed as follows: 1) redundant taxa and; 2) taxa present in less than 15% of the samples were removed; 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. 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.

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 with colonic involvement) disease location.

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: 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 univariate correlation analyses, between a total of 120 host phenotypes and each of the species or microbial pathway (i.e. microbial feature) abundances, using the total samples in this study, to identify potential phenotypic confounders. Wilcoxon test was used for categorical phenotypes and Spearman correlation for numerical. Phenotypes with most associations were selected for subsequent multivariate analyses

(Table S1; 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 univariable tests.

Bacterial composition & 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:

Tests
i) SI vs general population
ii) SI vs IBD-NoRes
iii) SI vs IBD-Res

The multivariate analyses were performed using generalized linear models as implemented in the R *MaAsLin* (v 0.0.5) package, allowing the boosted feature selection step.[18] The processed taxonomic or pathway data generated from the metagenomes, plus the selected phenotypes, were used as input (see previous sections; Table S1). 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. Descriptions of the multivariate models can be found in Tables S15-17 and 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

Exploration of low prevalence bacteria

As part of the quality control in the previous analyses, species with a prevalence of less than 15% of the total samples (n=1713) were filtered out. Due to an underrepresentation of SI samples in the total cohort (n=48), any bacteria prevalent in the small intestine samples, but rare in the samples from the general population and patients with IBD, would have been filtered. To explore this, we performed a logistic regression on the species filtered out between the SI group and the other groups combined (i.e. general population, IBD-NoRes & IBD-Res). Bacterial species relative abundances were coded as 0 for absence and 1 for presence. Age and sex were included in the model as covariates and corrected for multiple testing (FDR < 0.05).

Table 1. Clinical characteristics of the tested groups.

	General Population		IBD*		SI		Wilcoxon test p-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			
Sex [f/m]	689/489 [58/42%]	0	285/193 [60/40%]	0	42/15 [74/26%]	0	0.023	0.040
Age at faecal sampling [SD]	45 [13.6]	0	42.9 [12.8]	0	45.2 [10.9]	0	0.968	0.132
BMI [SD]	25.3 [4.2]	0	25.4 [5.0]	6	26.4 [6.85]	1	0.650	0.631
FecalCalprotectinOver200 [n/y] [%]	1124/48 [96/4%]	6	238/185 [56/44%]	55	32/10 [76/24%]	15	0.000	0.013
CRP Divided by 5 [SD]	NA	1178	1.7 [1.9]	2	1.75 [2.13]	0	NA	0.695
Current diagnosis [CD/IBDU/UC] [%]	NA	0	274/29/175 [57/6/37%]	0	23/1/33 [40/2/58%]	0	NA	0.005
IBD Disease duration in years [SD]	NA	1178	11.8 [8.8]	8	16.6 [10.4]	0	NA	0.000
IBD Active Disease [active/not active] [%]	NA	1178	114/358 [24/76%]	6	5/52 [9/91%]	0	NA	0.009
IBD Disease location [both/colon/ileum] [%]	NA	1178	111/221/97 [26/51/23%]	49	14/36/2 [27/69/4%]	5	NA	0.056
Any intestinal resections [n/y] [%]	NA	1178	309/169 [65/35%]	0	0/57 [0/100%]	0	NA	0.0
Number of intestinal resections [SD]	NA	1178	0.82 [1.6]	0	2.53 [2.28]	2	NA	0.0
IleocecalValveInSitu [n/y]	NA	1178	123/351 [26/74%]	4	52/3 [95/5%]	2	NA	0.0
Steroids [n/y] [%]	NA	1178	379/82 [82/18%]	17	46/6 [88/12%]	5	NA	0.258
Immunosuppressants [n/y] [%]	NA	1178	256/205 [56/44%]	17	37/15 [71/29%]	5	NA	0.031
AntiTNF [n/y] [%]	NA	1178	336/125 [73/27%]	17	43/9 [83/17%]	5	NA	0.127
Thiopurines [n/y] [%]	NA	1178	319/159 [67/33%]	0	51/6 [89/11%]	0	NA	0.000
ACE inhibitor [n/y] [%]	1091/44 [96/4%]	43	455/23 [95/5%]	0	56/1 [98/2%]	0	0.412	0.292
Angiotensin II receptor antagonist [n/y] [%]	1101/34 [97/3%]	43	467/11 [98/2%]	0	55/2 [96/4%]	0	0.825	0.576
Antibiotics [n/y] [%]	1165/13 [99/1%]	0	467/11 [98/2%]	0	54/3 [95/5%]	0	0.007	0.186
Beta Blockers [n/y] [%]	1115/63 [95/5%]	0	443/35 [93/7%]	0	49/8 [86/14%]	0	0.006	0.078
Laxatives [n/y] [%]	1114/21 [99/1%]	43	447/31 [94/6%]	0	53/4 [93/7%]	0	0.008	0.878
Mesalazines [n/y] [%]	1129/6 [99/1%]	43	310/168 [65/35%]	0	52/5 [91/9%]	0	0.000	0.000
Metformin [n/y] [%]	1162/16 [99/1%]	0	472/6 [99/1%]	0	55/2 [96/4%]	0	0.186	0.186
NSAID [n/y] [%]	1093/42 [96/4%]	43	448/30 [94/6%]	0	55/2 [96/4%]	0	0.940	0.405
Opiates [n/y] [%]	1122/13 [99/1%]	43	473/5 [99/1%]	0	55/2 [96/4%]	0	0.118	0.122
Oral Steroid [n/y] [%]	1173/5 [99/1%]	0	384/94 [80/20%]	0	47/10 [82/18%]	0	0.0	0.702
PPI [n/y] [%]	1079/99 [92/8%]	0	366/112 [77/23%]	0	37/20 [65/35%]	0	0.0	0.054
Antidepressant (SSRI) [n/y] [%]	1106/29 [97/3%]	43	469/9 [98/2%]	0	57/0 [100/0%]	0	0.222	0.297
Statin [n/y] [%]	1079/56 [95/5%]	43	447/31 [94/6%]	0	54/3 [95/5%]	0	0.911	0.721

Results

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 small intestine (SI). Average age and BMI were comparable between the groups ($p > 0.05$) (Table 1, Table S2). The SI group had a significantly larger proportion of females compared with both the general population and IBD groups (proportion females = 74%, 58% & 60%, respectively; $p < 0.05$) and a higher use of proton pump inhibitors and antibiotics when compared with the general population group (% PPI users = 35% & 8%, respectively; % antibiotic users = 5% & 1%, respectively; $p < 0.05$). Compared with the IBD groups, the SI group had a significantly larger proportion of individuals with UC and a lower mesalazine use (% UC = 37% & 58%, respectively; % mesalazine users = 35% & 9%; $p < 0.05$). Within the SI group, 5 individuals (9%) had active ileal disease at the time of faecal sampling.

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 were associated with gut microbial alterations in the SI group, we conducted association analyses between the respective phenotypes and species abundances. We identified no significant associations for all these analyses (FDR > 0.05, Table S3-S5).

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

On average, samples belonging to the SI group had a lower microbial richness when compared with the other groups (Shannon $Index_{mean}$ SI = 1.71; Shannon $Index_{mean}$ IBD-Res = 2.44, $p = 5.10 \times 10^{-14}$; Shannon $Index_{mean}$ IBD-NoRes = 2.77, $p = 2.22 \times 10^{-16}$; Shannon $Index_{mean}$ General population = 2.84, $p = 2.22 \times 10^{-16}$) (Figure 1, Table S6). To get an overview of the bacterial compositions between the groups, we measured the beta diversity using Bray-Curtis dissimilarity (Figure 2, Table S7). Samples from the SI group on average clustered furthest away from general population samples. IBD-Res and IBD-NoRes samples formed a gradient between SI and general population samples, with the IBD-Res samples positioning slightly more towards SI samples. Among all samples, SI samples explained 7.2%, and among IBD samples the presence of intestinal resections explained 5.6%, of the compositional dissimilarities ($p = 0.001$) (Table S8).

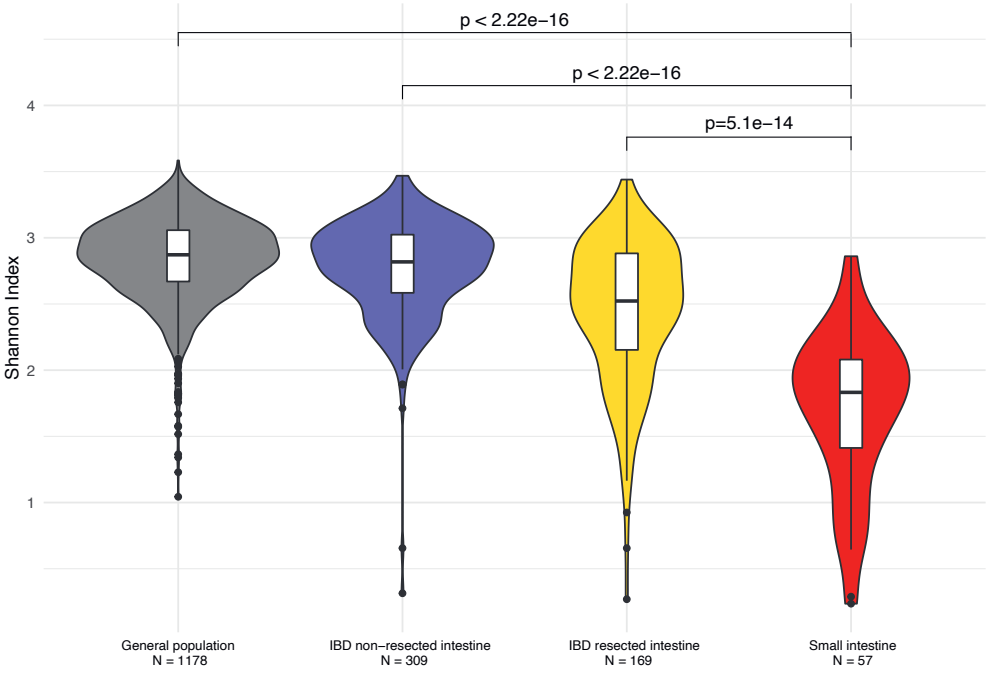


Figure 1. Microbial community in the small intestine shows a lower bacterial diversity compared with colonic samples. Violin plots displaying the distribution of Shannon index values per study groups. Small intestinal samples have a lower average bacterial diversity score (mean= 1.71) when compared with both the general population (mean= 2.84) and samples from patients with IBD, without or with intestinal resections (mean= 2.77 & 2.44, respectively).

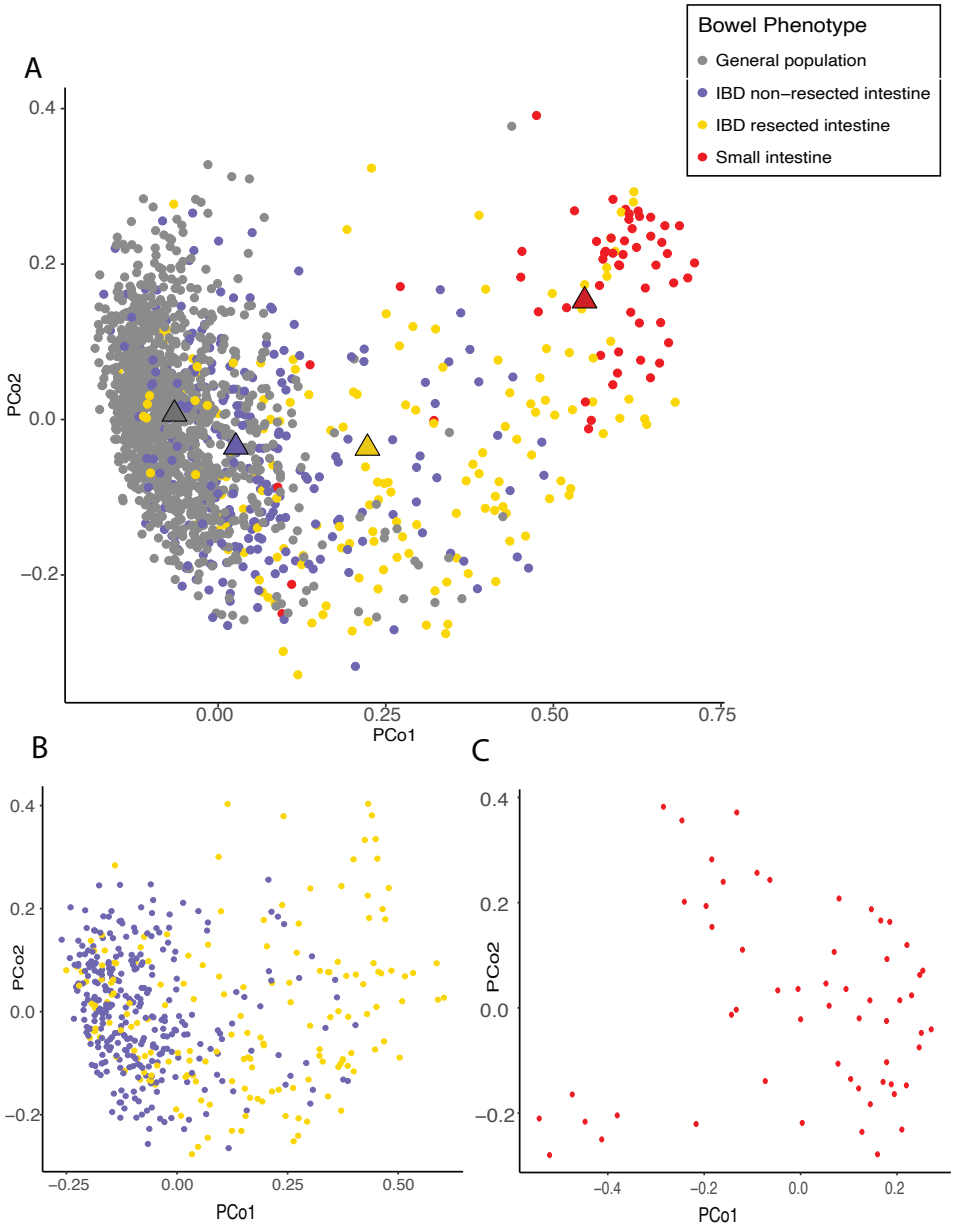


Figure 2. Principal coordinate analysis. Scatter plots showing Bray-Curtis distances between samples based on the bacterial composition. In grey, general population samples (n=1178); purple, patients with IBD without intestinal resections (n=309); yellow, patients with IBD with intestinal resections (n=169); red, the small intestinal content (n=57)). Panel A shows 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. Panels B and C highlight the heterogeneity between IBD samples representing the small intestinal content, respectively.

The overall genus composition in the small intestine

To characterise the differences observed in the beta-diversity analysis, we compared the top 10 most abundant genera in the SI and general population samples (Table S9-10). The most abundant genera in the 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% & 4.8%, respectively). Except for *Blautia*, all abundances were significantly higher when compared with the other groups (*Streptococcus*: SI vs general population, FDR = 2.73×10^{-23}; SI vs IBD-NoRes, FDR = 6.39×10^{-17}; SI vs IBD-Res, FDR = 3.73×10^{-14}; see supplementary table S9 for a complete table of values) (Figure 3). Notably, IBD-Res group had the second highest total mean abundance of the genera and the general population 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% & 5.2%, respectively); the total relative abundance increased in the order: SI, IBD-Res, IBD-NoRes and general population (Figure 3, Table S10).

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 S11). A total of 3617 associations were identified, involving 106 phenotypes and 134 species (FDR < 0.05). The phenotype representing 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 SI group, however, we did not identify any significant associations (FDR > 0.05) (Table S12). Lastly, 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 associations were identified for either of the variables (FDR > 0.05, Table S13-S14, Figure S1).

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

In total, 89 species were differentially abundant in the SI group when compared to the general population individuals, 82 compared with IBD-NoRes and 49 in the comparison between SI samples and IBD-Res (FDR < 0.05; Figure 4a; Table S15-S17). Of the 89 species differentially abundant in the SI compared with the general population samples, 22 were enriched in the SI. This included 9 belonging to the genera *Streptococcus*, 3 to *Veillonella* and 3 to *Actinomyces* (FDR < 0.05; Figure 4a; Table S15). 67 species were therefore underrepresented in the SI group, of which 6 belonged to the genera *Ruminococcus*, 8

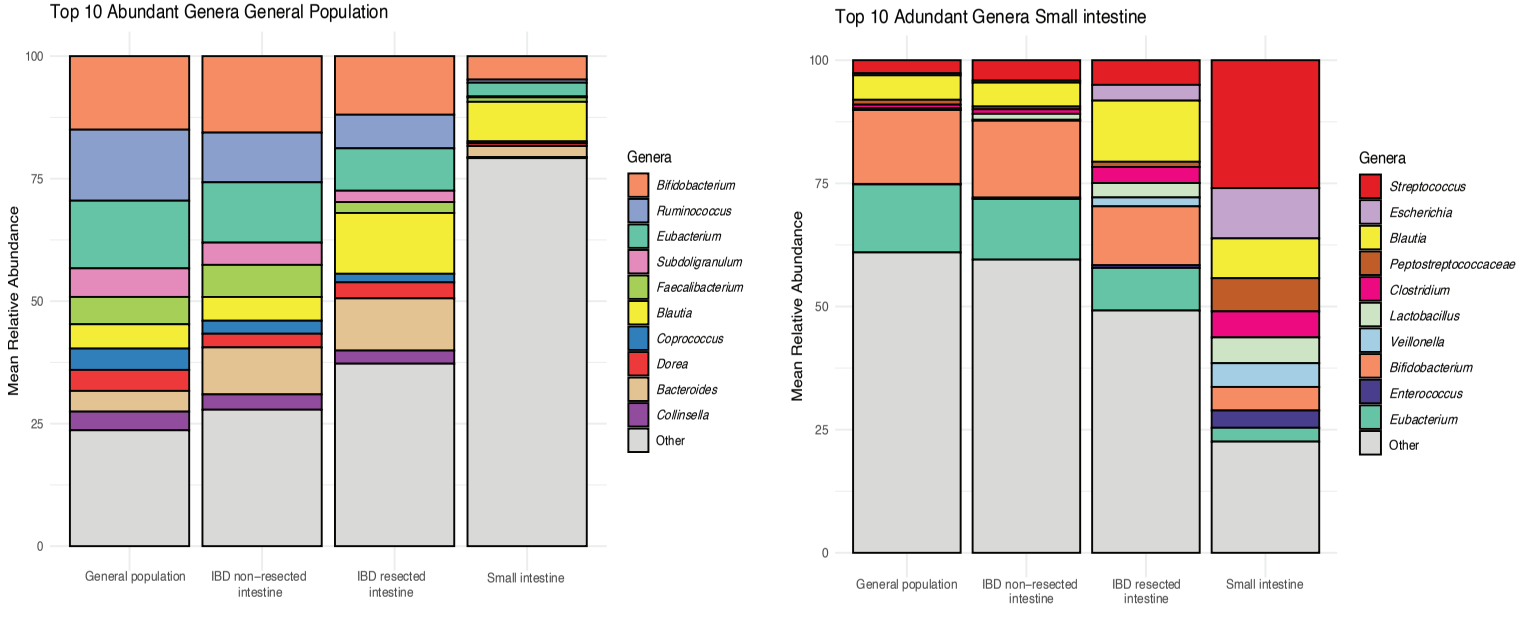


Figure 3. Bacterial genus profile is markedly different in samples from the small intestine when compared to colonic samples. Bar plots represent the top 10 most abundant genera among the general population samples and among the small intestine samples, respectively. In both the general population and IBD samples from patients without intestinal resections, *Bifidobacterium* and *Eubacterium* are the most abundant bacterial genera. In contrast, samples representing the small intestine microbial composition are predominated by bacteria belonging to the genera *Streptococcus* and *Escherichia* (Tables S9 & S10).

to *Eubacterium*, 10 to *Bacteroides* and 5 to *Alistipes*. Moreover, *Bifidobacterium dentium, Actinomyces odontolyticus, Streptococcus mutans* and *Streptococcus salivarius* were exclusively associated with this comparison (FDR < 0.05; Figure 4a, Table S15). Of the associations between SI and IBD-NoRes samples, 11 were unique, including a lower relative abundance of the butyrate producer *Pseudoflavonifractor capillosus* in SI samples (FDR = 1.95×10^{-5}; Figure 4a; Table S16). A lower relative abundance of a *Parabacteroides* species in SI individuals was only observed when comparing SI with the IBD-Res group (FDR = 0.047; Figure 4a; Table S17). *Veillonella atypica, Streptococcus mitis oralis pneumoniae, Streptococcus infantis, Streptococcus sanguinis, Actinomyces graevenitzii* and *Haemophilus parainfluenzae,* which are typically found in the oral cavity, were consistently found to be enriched in the SI compared with the other groups (FDR < 0.05; Figure 4a; Table S15-S17).

Rare colonic bacteria are prevalent in the small intestine

When comparing the prevalence of bacteria that were present in less that 15% of the cohort between the SI and the other three groups combined, we found that 110 of these species were significantly more prevalent in the SI group (FDR < 0.05, Table S18). Among the most prevalent species, 6 belonged to the genera *Streptococcus*, 3 to *Clostridium*, 4, to *Actinomyces*, 3 to *Klebsiella*, 5 to *Lactobacillus,* 3 to *Gemella,* 2 to *Atopobium* and 3 to *Enterococcus* (prevalence range in SI = 15-75%; prevalence range in other groups combined = 0.1-13%). Specific species that were enriched included *Veillonella dispar* (FDR = 9.39×10^{-23}), *Klebsiella pneumoniae* (FDR = 1.39×10^{-16}), *Enterococcus faecalis* (FDR = 2.39×10^{-23}), *Enterococcus faecium* (FDR = 1.32×10^{-12}) and *Lactobacillus fermentum* (FDR = 8.37×10^{-12}) (Table S18).

The small intestinal microbiota is largely characterised by pathways involved in sugar metabolism and quinone, heme, 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. 252 (74%) of the pathways were associated with at least one of the test comparisons: 243 pathways in the comparison SI vs general population, 147 in the comparison SI vs IBD-NoRes and 65 in the comparison SI vs IBD-Res (FDR < 0.05; Figure 4b; Table S19-21). Of these identified pathways, 52 were associated with all three tests. Examples included an increase in pathways related to sugar degradation, fermentation to lactate and quinone, heme, fatty acid and lipid biosynthesis, and an underrepresentation of pathways involved in degradation of complex carbohydrates and pyruvate fermentation to propanoate and butanoate (FDR < 0.05; Figure 4b; Table S19-21). Pathways that were exclusively enriched in the 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;

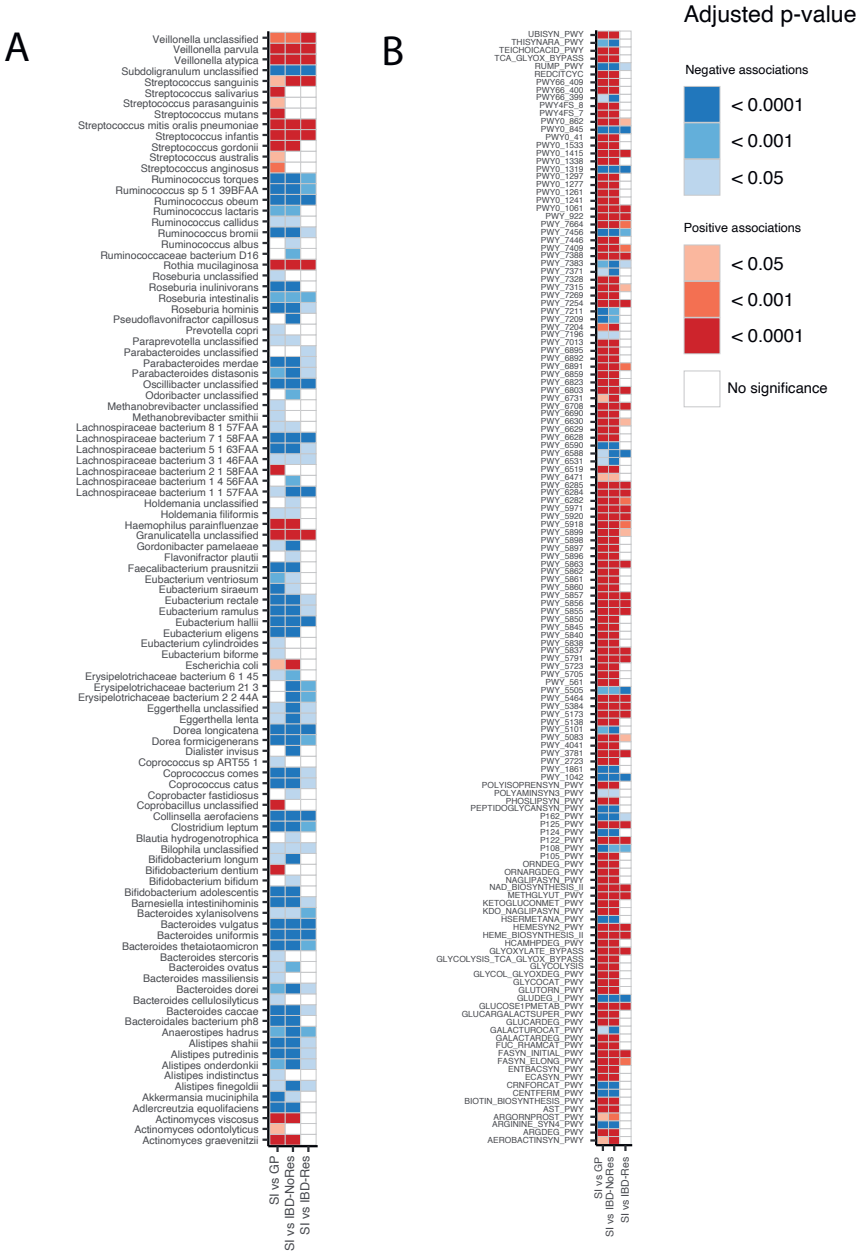


Figure 4. Microbial composition and metabolic potential of the small intestine. Heat maps represent the significant enrichment (in red) or underrepresentation (in blue) of bacterial species (panel A) and microbial pathways (panel B) in the samples representing the small intestinal content, compared with the respective study groups (GP, samples from general population cohort; IBD-NoRes, patients with IBD without intestinal resections; IBD-Res, patients with IBD and intestinal resections). Metabolic pathways are coded based on the accession ID in the MetaCyc database. P-values were calculated using multivariate linear regression models (see Methods) and adjusted for multiple testing (FDR<0.05). (Tables S15-S17 & S19-S21)

Figure 4b; Table S19). Conversely, pathways exclusively underrepresented in the SI were related to methanogenesis and pantothenate biosynthesis and amino acid biosynthesis pathways both increased and decreased (FDR < 0.05; Figure 4b; Table S19). Pathways that were associated with the comparison between SI and general population or IBD-NoRes group (FDR < 0.05), but similar in abundance between SI and IBD-Res samples (FDR > 0.05), are 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 4b; Table S19-21). These pathways, except for butanoate production were enriched in the SI. Of note, pathways such as quinone, heme, fatty acid and endotoxin biosynthesis suggest a proinflammatory potential.

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. Whilst correcting for potential confounders, we analysed metagenomes derived from faecal samples of 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 distances. 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.[13,19-21]

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 graevenitzii*. Additionally, species belonging to genera such as *Clostridium, Lactobacillus, Klebsiella, Gemella a*nd *Enterococcus*, which were rarely observed in the general population faecal samples, had a significantly higher prevalence between 15%-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.[6,13,19,22,23]

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 to the general population, including biotin biosynthesis pathways.[19] Biotin, also called vitamin B_7 or B_8, is an important cofactor for several carboxylases that are essential for glucose, amino acid and fatty acid metabolism.[24] Biotin is also thought to have anti-inflammatory effects by inhibiting NF-kB, a pro-inflammatory signaling molecule, expression. 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, relative to general population, samples.[25] We noted at least 4 pathways related to fatty acid and lipid metabolism that 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.[26] 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.[27]

When comparing the SI group with IBD-NoRes and IBD-Res groups, fewer species (n = 82 & 49, respectively) were associated, as when compared with the general population group (n = 89), suggesting increased colonization 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.[17,28-31] In fact, many bacteria enriched in the small intestine compared with the other groups have been associated with IBD. Examples of which include *Veillonella spp.*, *Streptococcus spp., Enterococcus faecalis*, *Enterococcus faecium* and *Klebsiella pneumoniae*.[12,17,28,32,33] On a functional level, fewer pathways were associated with the comparison between the SI and the two IBD groups (n = 147 [SI vs IBD-NoRes], 65 [SI vs IBD-Res] & 243 [SI vs general population]). Pathways involved in lactate and acetate production and degradation of arginine, which were enriched in SI samples compared with the general population, were no longer associated with the SI vs IBD-Res comparison. This is in line with reports of elevated abundances of lactate, as well as lactate-producing bacteria (e.g. *Lactobacilli, Enterococci, Streptococci* and *Pediococci*) in faecal samples of patients with IBD.[17] Similarly, pantothenate (vitamin B_5) biosynthesis and methanogenesis pathways were underrepresented in the SI compared with the general population group, but not compared with IBD samples. Pantothenate metabolites have been previously found to be decreased in IBD faecal samples.[17] Vitamin B_5 is absorbed in the colon and its deficiency has been associated with

the production of pro-inflammatory molecules.[24] Methanogenesis is the formation of methane from hydrogen and carbon dioxide. Methane has been reported to slow intestinal transit, thus reduced methanogenesis is consistent with the shorter transit times observed in the small intestine.[34] 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.[34] Taken together, these results support a role for small intestinal, rather than per se oral, pathobionts in IBD disease pathogenesis.

Whilst 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 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 study the small intestinal gut microbiota non-invasively. Lastly, one might argue that the individuals within our small intestine group are heterogeneous due to, for example, the inclusion of patients with pouches. 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.

References

1. Brestoff JR, Artis D. Commensal bacteria at the interface of host metabolism and the immune system. Nat Immunol. 2013;14:676-684
2. Clemente JC, Ursell LK, Parfrey LW, et al. The impact of the gut microbiota on human health: An integrative view. Cell. 2012;148:1258-1270
3. Ni J, Wu GD, Albenberg L, et al. Gut microbiota and IBD: Causation or correlation? Nat Rev Gastroenterol Hepatol. 2017;14:573-584
4. Collins SM, Surette M, Bercik P. The interplay between the intestinal microbiota and the brain. Nat Rev Microbiol. 2012;10:735-742
5. 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;174:1388-1405
6. Aidy SE, van den Bogert B, Kleerebezem M. The small intestine microbiota, nutritional modulation and relevance for health. Curr Opin Biotechnol. 2015;32:14-20
7. Ermund A, Schütte A, Johansson MEV, et al. Studies of mucus in mouse stomach, small intestine, and colon. Gastrointestinal mucus layers have different properties depending on location as well as over the Peyer's patches. Am J Physiol - Gastrointest Liver Physiol. 2013;305:341-347
8. Zhong L, Shanahan ER, Raj A, et al. Dyspepsia and the microbiome: Time to focus on the small intestine. Gut. 2017;66:1168-1169
9. Saffouri GB, Shields-Cutler RR, Chen J, et al. Small intestinal microbial dysbiosis underlies symptoms associated with functional gastrointestinal disorders. Nat Commun. 2019;10:2012
10. Schmidt TSB, Hayward MR, Coelho LP, et al. Extensive transmission of microbes along the gastrointestinal tract. Elife. 2019;8
11. 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
12. Atarashi K, Suda W, Luo C, et al. Ectopic colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation. Science. 2017;358:359-365
13. Kastl AJ, Terry NA, Wu GD, et al. The structure and function of the human small intestinal microbiota: Current understanding and future directions. Cmgh. 2020;9:33-45
14. Imhann F, Van Der Velde KJ, Barbieri R, et al. The 1000IBD project: Multi-omics data of 1000 inflammatory bowel disease patients; data release 1. BMC Gastroenterol. 2019;19:5
15. Tigchelaar EF, Zhernakova A, Dekens JAM, et al. Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: Study design and baseline characteristics. BMJ Open. 2015;5:1-9
16. Truong DT, Franzosa EA, Tickle TL, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods. 2015;12:902-903
17. 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. MaAsLiN. [http://huttenhower.sph.harvard.edu/maaslin].
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;6:1415-1426
20. Friedman ES, Bittinger K, Esipova T V., et al. Microbes vs. chemistry in the origin of the anaerobic gut lumen. Proc Natl Acad Sci U S A. 2018;115:4170-4175
21. Nugent S, Kumar D, Rampton D, et al. Intestinal luminal pH in inflammatory bowel disease : possible determinants and implications for therapy with aminosalicylates and other drugs. Gut. 2001;48:571-577
22. Booijink CCGM, El-Aidy S, Rajilić-Stojanović M, et al. High temporal and inter-individual variation detected in the human ileal microbiota. Environ Microbiol. 2010;12:3213-3227
23. Villmones HC, Haug ES, Ulvestad E, et al. Species level description of the human ileal bacterial microbiota. Sci Rep. 2018;8:1-9
24. Yoshii K, Hosomi K, Sawane K, et al. Metabolism of dietary and microbial vitamin b family in the regulation of host immunity. Front Nutr. 2019;6:1-12

25. Magnúsdóttir S, Ravcheev D, De Crécy-Lagard V, et al. Systematic genome assessment of B-vitamin biosynthesis suggests cooperation among gut microbes. Front Genet. 2015;6:148
26. 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;23:458-469
27. Araújo JR, Tazi A, Burlen-Defranoux O, et al. Fermentation products of commensal bacteria alter enterocyte Llpid metabolism. Cell Host Microbe. 2020;27:358-375
28. Gevers D, Kugathasan S, Denson LA, et al. The treatment-naïve microbiome in new-onset Crohn's disease. Cell Host Microbe. 2014;15:382-392
29. 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
30. 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;2:1-11
31. Nahum A, Filice G, Malhotra A. A complicated thread: abdominal actinomycosis in a young woman with Crohn disease. Case Rep Gastroenterol. 2017;11:377-381
32. 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;20:1-18
33. Balish E, Warner T. Enterococcus faecalis induces inflammatory bowel disease in interleukin-10 knockout mice. Am J Pathol. 2002;160:2253-2257
34. Jahng J, Jung IS, Choi EJ, et al. The effects of methane and hydrogen gases produced by enteric bacteria on ileal motility and colonic transit time. Neurogastroenterol Motil. 2012;24:185-191