**Assessing the colonic microbiota in children: Effects of sample site and bowel preparation**

Alexander G. **Shaw**\*†1, Ph.D., Naomi **Black**\*2, BSc, Azelea **Rushd**2, BSc, Kathleen **Sim**1, M.R.C.P.C.H., Paul **Randell**1, F.R.C.Path., J. Simon **Kroll**1, F.Med.Sci., Jenny **Epstein**2, Ph.D.

\* Contributed equally to the manuscript

1 Department of Medicine, Section of Paediatrics, Imperial College London, London, United Kingdom

2 Department of Paediatric Gastroenterology, Chelsea and Westminster Hospital, London, United Kingdom

†To whom correspondence should be addressed: [a.shaw@imperial.ac.uk](mailto:a.shaw@imperial.ac.uk), Section of Paediatrics, Wright-Fleming Institute, Norfolk Place, London, W2 1PG, Telephone: 020 7594 3717, Fax: 020 7954 3984

**Funding Statement**

This work was supported by funding from Meningitis Now, the Winnicott Foundation and the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial Healthcare NHS Trust and Imperial College London.  KS has been the recipient of an NIHR Doctoral Research Fellowship [NIHR-DRF-2011-04-128].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Word Count:** 3677

**Number of Figures:** 6

**Number of Tables:** 1

**Conflicts of Interest and Source of Funding:** The authors have no conflicts of interest to declare.

**Author Contributions:**

AGS drafted the manuscript, performed sample preparation and next generation sequencing, conducted the bioinformatics processing and statistical analysis of the data, and submitted read data to an online repository.

NB drafted the manuscript, collected patient samples and clinical data, and performed sample preparation and next generation sequencing.

AR drafted the manuscript, collected patient samples and clinical data, and performed sample preparation and next generation sequencing.

KS performed sample preparation and next generation sequencing and revised the manuscript.

PR performed sample preparation and next generation sequencing and revised the manuscript.

JSK jointly led the study and revised the manuscript.

JE jointly led the study, conducted the ethical application, collected patient samples and clinical data, and revised the manuscript.

**Abstract**

**Objectives**

Inflammatory bowel disease (IBD) states are associated with gastrointestinal dysbiosis. Mucosal biopsy sampling, retrieving the bacterial community that most directly interacts with the host, is an invasive procedure, and we hypothesize may be sufficiently approximated by other sampling methods. We investigate the relatedness of samples obtained by different methods and the effects of bowel preparation on the gastrointestinal community in a paediatric population.

**Methods**

We recruited a cohort of patients undergoing colonoscopy, collecting serial samples via differing methods (rectal swabs, biopsies and faecal matter/luminal contents) pre-bowel preparation, during colonoscopy and post-colonoscopy. Next generation sequencing was used to determine the structure of the microbial community.

**Results**

The microbial community in luminal contents collected during colonoscopy was found to be more similar to that of mucosal biopsies than rectal swabs. Community traits of the mucosal biopsies could be used to segregate IBD patients from other patients, and the similarity of the communities in the luminal contents was sufficient for the segregation to be reproduced. Microbial communities sampled by rectal swabs and pre-bowel preparation faeces were less similar to mucosal biopsies. Bowel preparation was found to have no significant long term effects on the microbial community, despite the transient effects evident during colonoscopy.

**Conclusions**

A clinically relevant description of the mucosal microbial community can be obtained via the non-invasive collection of luminal contents after bowel cleansing. Bowel preparation in a paediatric population results in no consistent sustained alterations to the gastrointestinal microbiota.

**Key words:** paediatric, colonoscopy, biopsy, bacteria

**What is known:**

Altered gastrointestinal microbiota states have been associated with inflammatory bowel disease (IBD).

Bowel preparation in non-paediatric populations has little long-term effects on the gastrointestinal microbiota.

Rectal swabs taken from un-prepared individuals provide notably different observations of the gastrointestinal microbiota when compared to mucosal biopsies.

**What is new:**

Luminal contents (post-bowel preparation) provide the closest approximation of the mucosal microbiota; rectal swabs and faecal samples taken pre-preparation are more disparate.

This similarity is sufficient to segregate IBD patients from other patients requiring endoscopies.

Bowel preparation in paediatric populations produces significant shifts in community structure and diversity evident during colonoscopy, and these effects are transient.

**Introduction**

Humans carry approximately 1011 bacterial organisms per gram of faecal material in their gastrointestinal tract alone (1), collectively known as the gastrointestinal microbiota. The coevolution of the human body and its gut microbiota has led to a taxonomically diverse ecosystem which not only differs greatly between individuals but also varies within individuals throughout their lifetime. Analytical advances due to the Human Microbiome Project have helped increase our knowledge of specific crucial species in the microbiota. Research into to the colonic microbiota suggests changes and imbalances in this relationship may result in disease states. There is increasing evidence to suggest an association between dysbiosis of the gut microbiota and disease states, most notably inflammatory bowel disease (IBD), colorectal cancer and obesity (2-11).

The gut microbiota can be sampled by a rectal mucosal biopsy, faecal matter/luminal contents or rectal swab, but which is most representative? All three methods have distinct advantages and disadvantages when practiced in paediatric populations. Currently, the gold standard for colonic analysis is a rectal biopsy of the mucosal layer which, as well as providing tissue for histological analysis, identifies the mucosal adherent bacteria. These bacteria interact directly with the host immune system and are thus potentially able to modulate the manifestation of diseases such as IBD. However, a biopsy requires a colonoscopy or sigmoidoscopy, often performed under general anaesthetic for children, which itself carries risks. Where the interest is in the microbiota, bacterial populations and diversity can also be assessed in faecal matter or rectal swabs. Faecal matter sampling is non-invasive and easy, although patients dislike handling faeces and especially where sampling is requested at home, compliance is often poor. Rectal swabbing is non-invasive, less expensive, less time consuming, and more comfortable and acceptable for both adult and paediatric patients, so this might be an ideal sampling method. However, the extent to which either approach provides information comparable to mucosal biopsy is unknown.

We hypothesize that rectal swabbing and/or faecal sampling may be suitable, non-invasive alternatives to a mucosal biopsy for sampling the gut microbiota. To test this we assessed concordance in composition and diversity of the microbiota retrieved by these methods from a paediatric cohort of patients compared to results obtained from a rectal mucosal biopsy sample. We investigated the impact of laxative bowel preparation on the composition and diversity of the gut microbiota, as studies have shown this to be perturbed by bowel cleansing.

**Methods and Materials**

**Study populations**

The study population comprised paediatric patients undergoing sigmoidoscopy or colonoscopy in our unit between February and June 2014 for any clinical indication. There were no exclusion criteria. 31 patients were recruited, with samples obtained from eighteen of these, aged between four and seventeen years. Sixteen patients provided a full set of samples (swab, biopsy and faecal sample) at colonoscopy. Clinical metadata and sample availability for the patients who provided samples are shown in Table 1.

Colonoscopy patients received standard bowel preparation of sodium picosulfate with magnesium citrate and senna before their procedure as per local protocol (Supplemental Digital Content 1). Sigmoidoscopy was performed on an unprepared colon. All patients underwent the procedures under general anaesthetic.

**Ethics declaration**

Written consent was obtained from parents of all study participants prior to any sample collection and ethical approval was granted by Chelsea Research Ethics Service, United Kingdom (IRAS REC number 12/LO/0240).

**Sample collection**

In order to achieve a longitudinal representation of the gut microbiota a total of six samples were collected from each participant at three different time points (see Figure 1); Pre-colonoscopy, colonoscopy and post-colonoscopy. The pre-colonoscopy faecal sample was obtained prior to bowel preparation. A rectal swab, biopsy and luminal contents were taken at colonoscopy. The post-colonoscopy faecal sample was obtained at a subsequent outpatient visit two weeks or more after the procedure. Samples collected for each patient are indicated in Table 1.

**Faecal sample collection**

Pre-colonoscopy faecal samples were collected by the patient prior to beginning bowel preparation. Luminal contents were obtained at colonoscopy, either directly or via colonoscopic suction. In all cases the faecal liquid present at the most distal colonic location was sampled.The samples were stored in a DNAase, RNAase free, sterile microcentrifuge tube and stored at -80oC.

**Mucosal biopsy collection**

Rectal mucosal biopsies were obtained during colonoscopy using disposable flexible biopsy forceps and were stored in a DNAase, RNAase free, sterile microcentrifuge tube at -80oC.

**Rectal swab collection**

Rectal swabs were taken at colonoscopy and post-colonoscopy. The swab specimen was collected by inserting a sterile cotton-tipped swab just beyond the anal canal and rotating once. Swabs were then removed and stored at -80oC.

**Bacterial DNA extraction**

Faecal samples (200 mg) were processed using the FastDNA SPIN Kit for Soil (MP Biomedicals, Solon, US), incorporating a bead-beating step for mechanical disruption of cells. Extractions were performed following the manufacturer’s protocol except that the final elution step was into TRIS (10 mM) low-ethylenediaminetetraacetic acid (EDTA) (0.1 mM) buffer. Swabs underwent the same extraction process, with swab heads being removed with sterilised scissors. Biopsy samples (a maximum of 500 mg) was placed in the FastDNA SPIN Kit for Soil LME tubes and homogenised twice using FastPrep-24 (MP Biomedicals) for 40 seconds at maximum setting. Samples were incubated at 37oC for 10 minutes on a rotator then incubated for 18 hours at 55oC with 100 µl of 20 mg/ml Proteinase K. After the incubation, samples were heated to 90oC for 1 hour and then entered the faecal sample extraction process as above.

**Polymerase Chain Reaction amplification and pyrosequencing of variable regions of the bacterial 16S rRNA gene**

The V3-V5 region of the bacterial 16S rRNA genes was amplified from each DNA sample using a primer pair tagged with individually unique 12-bp error-correcting Golay barcodes (12-14). Polymerase Chain Reaction (PCR) was performed as previously described (15). Replicate amplicons were pooled and purified, and two pyrosequencing runs were performed on a 454 Life Sciences GS Junior (Roche, Basel, Switzerland) following the Roche Amplicon Lib-L protocol. Replicate samples spread over all sequencing runs acted as internal controls.

Negative controls were included in all sequencing runs to identify potential contamination.

**Bioinformatics**

454 Shotgun processed data were denoised using AmpliconNoise (16) as part of the ‘Quantitative Insights Into Microbial Ecology’ (QIIME) (17) package followed by chimera-removal with Perseus (16) and demultiplexing. Sequences were aligned using the SILVA rRNA database 119 (18) release for reference and clustered at 97% sequence identity using the UCLUST algorithm (19) into operational taxonomic units (OTUs). The OTU table was filtered to remove singletons (sequences present only once in the dataset) and OTUs only present in one sample. Rarefaction to 556 reads was performed, removing heterogeneity of sequencing reads per sample. The OTU table was summarised to genus and phylum level using the QIIME script summarize\_taxa.py, with the resulting taxonomic descriptors being referred to as “taxonomic groups”. In order to calculate generalised UniFrac Distances (20), a taxonomic tree was created based on consensus sequences for each genus level classification. The consensus sequences created using the EMBOSS cons tool (21), combining the representative sequences for the OTUs that were compressed into each classification.

**Data availability**

16S rRNA amplicon data have been deposited at the European Nucleotide Archive under accession number PRJEB11439.

**Statistics**

Statistical analysis was performed using the R software package (version 3.1.3).

Alpha diversity has been calculated using the Shannon-Weaver index and has been analysed for the filtered OTU table. Alpha diversity was analysed using either the Wilcoxon signed-rank test, Mann-Whitney U test or the Friedman test.

Where taxonomic group count data is presented, it is drawn from either the genus or phylum level summarised tables. The stacked barcharts presented are those derived from the genus level summarised table. General linear models, implementing a negative binomial regression, and including sample donor ID and age as additional variables where applicable, were used to detect associations across the datasets.

Generalised UniFrac distances were calculated using the summarised OTU table and were compared using the Wilcoxon signed-rank test, Mann-Whitney U test or a Kruskal Wallis ANOVA.

Where post hoc analysis or multiple tests have been performed, p-values have been adjusted with the Bonferroni correction.

**Results**

**Sequencing Results**

The microbial communities of 88 samples were successfully sequenced using the Roche GS Junior platform, with total of 181,486 sequencing reads. The thirteen samples with below 556 reads were discarded, with rarefaction curves showing that this number of reads was sufficient to accurately capture the full microbial diversity. After the removal of duplicates, seventeen biopsy samples, sixteen luminal content samples and eighteen swabs taken at colonoscopy, eleven pre-bowel preparation faecal samples and seven post-bowel preparation faecal samples were retained, inclusive of fourteen complete sets of sample taken at colonoscopy from patients who had undergone bowel preparation.

**Can luminal contents or swabs taken during colonoscopy act as a surrogate for biopsy samples?**

Given the invasive nature of obtaining mucosal biopsy samples, a less invasive surrogate sample would prove valuable if observations of the microbial community are sufficiently similar. The fourteen complete sets (a biopsy, a swab and luminal contents taken at colonoscopy) from patients who had undergone a bowel preparation were compared, with the biopsy being considered the ‘gold standard’ representation of the mucosa-associated gut microbiota.

Generalised UniFrac distances, summarising the similarity of samples according to both microbial diversity and abundance, were measured, with the distances between biopsies and paired luminal contents being significantly lower (p = 0.002) than the distance between biopsies and paired swabs (paired t-tests, mean distances 0.32 and 0.57 respectively). Biopsies were also found to be more similar to their paired luminal contents than to other biopsies (p < 0.001, mean distance between paired luminal contents and biopsies = 0.32 faeces (n = 14), mean distance between biopsies = 0.57 (n = 91 comparisons made between fourteen biopsies). These comparisons are shown in Figure 1.

The diversity of the microbial communities was compared between biopsies and the surrogate samples. A Friedman test indicated significant differences between sample types (p = 0.008), and although the community diversity was lower in luminal contents than paired biopsies, post-hoc analysis found this to be non-significant (median diversities: luminal contents = 3.23, rectal swabs = 4.08, biopsies = 4.04).

General linear models (GLMs) were then used to identify associations between sample types and taxonomic groups that comprised >1% of the dataset reads. At the genus level, compared to paired biopsies, luminal contents were found to have a significantly lower abundance of Ruminococcaceae (p = 0.032) and a significantly higher abundance of Anaerostipes (p = 0.026). Rectal swabs had a greater number of taxonomic groups that differed from paired biopsies, with significant decreases in three Lachnospiraceae taxonomic groups (p < 0.001, p = 0.027, p = 0.038), Bacteroides (p = 0.015), Blautia (p < 0.001) and Ruminococcaceae (p < 0.001) and significant increases in Peptoniphilus and Finegoldia (p-values < 0.001). Comparisons of the taxonomic groups for the fourteen sets of samples are shown in Figure 3.

Differential taxonomic group abundance was also investigated at the level of phylum. GLMs indicated that rectal swabs had significantly increased abundances of Actinobacteria and Fusobacteria (p-values < 0.001) compared to paired biopsies.

**Similarity of biopsies to faecal samples taken before bowel preparation**

Given that luminal contents, despite the lower diversity at the time of colonoscopy, are more similar than rectal swabs to paired biopsies, we sought to determine the similarity of faecal samples taken prior to bowel preparation versus paired biopsy samples. Eight sets of pre-bowel preparation faecal samples, luminal contents and biopsies were available. Generalised UniFrac distances between the paired biopsies and pre-bowel preparation faecal samples were found to be significantly lower than distances between biopsies (p = 0.03, median distance = 0.53 between biopsies (n = 28) and 0.44 between paired biopsies and pre-bowel preparation faecal samples (n = 8)). Biopsies were however significantly more similar to paired luminal contents than they were to paired pre-colonoscopy faecal samples (p = 0.04, mean distance between paired biopsies and colonoscopy faecal samples = 0.53, Figure 4).

Pre-bowel preparation faecal samples and paired biopsies did not have significantly different microbial community diversities (median diversities of 4.48 and 4.35 respectively); however multiple taxonomic groups were found to be differentially abundant. Compared to paired biopsies, the pre-bowel preparation faecal samples were found to have significantly higher abundances of Ruminococcus, Bifidobacterium, Blautia, Subdoligranulum and Peptostreptococcaceae (p-values < 0.001) and significantly decreased abundances of Faecalibacterium (p = 0.002) and a Lachnospiraceae taxonomic group (p < 0.001). At the phylum level, Erysipelotrichia and Actinobacteria were significantly more abundant in pre-bowel preparation faecal samples compared to paired biopsies (p = 0.010 and p < 0.001 respectively) whilst Bacteroidia and Betaproteobacteria were significantly lower in abundance (p = 0.033 and p = 0.006).

**What are the effects of bowel preparation on the faecal microbiota?**

To determine the effects of bowel preparation on the faecal microbiota six sets of patient samples were considered, each including faecal samples taken pre-bowel preparation, luminal contents and post-colonoscopy faecal samples (a median of 55 days after colonoscopy).

The diversity of the microbial communities for each sample set was found to differ significantly (Friedman test, p = 0.04), with lower diversity communities in luminal content samples (median pre-bowel preparation diversity = 4.82, luminal contents diversity = 4.33 and post-colonoscopy diversity = 4.85), although post hoc analysis found the differences to be non-significant after correction. No significant change in diversity was found comparing pre-bowel preparation and post-colonoscopy.

GLMs were used to identify shifts in taxonomic group abundance. A significant increase in Faecalibacterium (p < 0.001) and significant decreases in Ruminococcus, Escherichia, Pseudobutyrivibrio and Subdoligranulum (p-values < 0.001) were found in luminal contents when compared to pre-bowel preparation faecal samples. However, when pre-bowel preparation samples were compared to paired post colonoscopy samples, only minor changes in taxonomic group abundances were found (a significant increase in Christensenellaceae (p < 0.001) in the post colonoscopy samples). At the phylum level, despite some temporary changes as a result of the bowel preparation (a significant reduction in the abundance of Gammaproteobacteria (p < 0.001) and an increase in Bacteroidia (p = 0.014) in luminal contents), no significant alterations were found between pre- and post-bowel preparation samples. Overall, it would appear that there are relatively minor *concordant*changes as a result of the bowel preparation. Conversely, when considering the percentage of reads that were assigned to different genus-level taxonomic groups when comparing pre and post-treatment samples, considerable alteration of the microbial community was observed (a median of 35% of taxonomic group reads being reassigned). Even when performed at phylum level, a median of 20% of taxonomic group reads were reassigned to different phyla. The taxonomic groups observed at genus level found in the six sets of patient data are shown in Figure 5.

**Are there any associations between microbial community and disease state?**

Available samples were split into two groups according to the patient diagnoses made by attending clinicians; either inflammatory bowel disease (IBD, inclusive of Crohn’s Diseases (CD) and Ulcerative Colitis (UC)) or non-IBD diagnoses (see Table 1). GLMs were performed including the age of the donor patients as a confounding factor, and only patients having undergone bowel preparation were included in the analyses.

Microbial data for biopsies were available for five CD patients, five UC and five non-IBD patients. Samples from IBD patients were found to have a significantly less diverse microbial communities (p = 0.005, median diversity: IBD = 3.56, non-IBD = 4.77). Biopsies from IBD patients had a significantly higher abundance of Escherichia (p < 0.001) and lower abundances of Lachnospiraceae and Ruminococcus (non-significant after multiple testing corrections). At the phylum level, IBD biopsies were associated with a significantly higher abundance of Gammaproteobacteria (p < 0.001).

Microbial data from luminal contents were available for five UC patients, four CD patients and five non-IBD patients. Similar to the biopsy results, microbial community diversity was found to be significantly lower in samples from IBD patients compared to non-IBD (p = 0.029), median diversity: IBD = 2.92, non-IBD = 4.33). IBD luminal contents also showed significant shifts in the same taxonomic groups as the biopsies: significantly increased Escherichia (p = 0.002) and lower abundances of Lachnospiraceae and Ruminococcus (non-significant after multiple testing corrections). Similarly, at the phylum level, IBD luminal contents were again associated with a significantly higher abundance of Gammaproteobacteria (p < 0.001).

Whilst a non-significant reduction in community diversity was evident in pre-bowel cleansed faecal samples (median diversity: IBD = 3.61, non-IBD = 4.63), no matching taxonomic group differences were evident (IBD n = 5, non-IBD n = 6).

The three significant factors found to differ between the biopsy sample groups (the Shannon-Weaver diversity index, and Escherichia and Gammaproteobacteria sequencing reads) were used to calculate a matrix of distances between samples, with the aim of discriminating between IBD and non-IBD samples. Improved separation was achieved through incorporating Lachnospiraceae and Ruminococcus sequencing reads into the calculations. The distances derived from these five factors were used to cluster the biopsy samples (Figure 6a) and the faecal samples taken at colonoscopy (Figure 6b) by similarity. The same factors were used to cluster pre-bowel preparation samples but resulted in poor separation.

**Discussion**

We conclude that microbial community composition from rectal swabs was unrelated to rectal biopsies and luminal contents, and thus, whilst a non-invasive and more rapid sampling method, they do not serve as a useful surrogate for sampling the mucosa-associated microbiota. These observations were supported by Araujo-Perez et al*.* who also documented the difference in bacterial community composition between swabs and biopsies (22). However, our findings suggest a concordance between the bacterial diversity and composition from rectal biopsies and luminal contents from a prepared bowel, raising the possibility of luminal contents being a viable alternative to invasive biopsies for the purpose of determining the microbial community.

This finding is given greater importance by our demonstration that measurements of the microbial community drawn from biopsy samples can be used to segregate CD and UC patients from non-IBD patients, and that these factors can also be measured in luminal contents and result in similar clustering. This shows that not only are the samples similar, but they share enough homology to be of clinical use.

Zoetendal et al. found significant differences between the microbial communities of biopsies and faecal samples in adults (23). This is comparable to our finding of dissimilarity between biopsy sample microbiota and the pre-treatment samples analysed. In our work, bowel preparation appears sufficiently to cleanse the lumen to ensure that the luminal contents sampled at colonoscopy – being derived from surface/mucus-adherent microbes rather than the bulk faecal stream – contain a microbial community more similar to that of biopsy samples.

Our analysis comparing patient’s faecal microbiota pre- and post-colonoscopy indicates that bowel preparation does lead to changes in the faecal microbial community, but that these do not appear to occur in a co-ordinated manner. Alterations occur both at the genus and phylum level. Such differences may therefore be either due to stochastic recovery of the community after treatment or simply background noise from sampling and sequencing.

Studies by O’Brien et al. and Jalanka et al. confirm that colonoscopy preparation does not significantly affect the gut microbiota in the long term (24, 25). The short term effects of bowel preparation may be due to the majority of the bacteria being washed out in a non-discriminatory manner, which could reduce low-abundance taxonomic groups to levels below detection; on average, bowel preparation with polyethylene glycol can reduce bacterial load by 34.7-fold (25).

Many studies into the aetiology of CD and UC focus on the important role played by the intestinal microbiota. Our findings of the association between IBD and the gut microbiota are in accordance with previous studies. Increased abundance of Escherichia (or related Enterobacteriaceae, given to difficulty of resolving this family with 16S rRNA-based methodologies (26)) has also been found in adult UC patients compared to healthy individuals in biopsies (27) and faecal samples (28). Maukonen et al.found an association between reduced Lachnospiraceae and paediatric IBD, particularly in the case of UC (2). In a discordant twin study, Lepage et al. found reduced levels of Lachnospiraceae in the UC twins (3). Some studies have suggested more precise microbial associations, with reductions in *Roseburia* (a genus of the Lachnospiraceae family) in both UC and CD (4), and in UC alone (5), compared to healthy controls. An array of both mucosal and stool based studies have observed a reduced diversity of the microbial community in association with UC (3, 6, 29) and with both UC and CD (27). Our observation of these factors is therefore both supported and further reinforces the utility of luminal contents from a prepared bowel as an accurate surrogate for the mucosal microbial community. These signals are not evident in samples prior to bowel preparation however, suggesting that faecal material during colonoscopy (and post bowel preparation) is more representative of the mucosal layer.

The study population was comparable in size to similar studies, and provides new paediatric data which builds upon what is known from adult populations.

Whilst larger patient cohorts could further explore our conclusions, we provide evidence that the paediatric colonic microbiota is subject to systematic variation depending upon sample method and bowel cleansing effects, and that in pursuit of the mucosa-associated microbiota an invasive gut biopsy may not be required. Our findings support the ongoing development of the human gut microbiota as a practical tool; able to segregate patients based on disease state, and of potential utility in pathophysiological study and clinical practice.

**References**

1 Thiel R, Blaut M An improved method for the automated enumeration of fluorescently labelled bacteria in human faeces. J.Microbiol.Methods 2005;61(3):369-79.

2 Maukonen J, Kolho KL, Paasela M, et al. Altered Fecal Microbiota in Paediatric Inflammatory Bowel Disease. J Crohns Colitis 2015;9(12):1088-95.

3 Lepage P, Hasler R, Spehlmann ME, et al. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. Gastroenterology 2011;141(1):227-36.

4 Morgan XC, Tickle TL, Sokol H, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol 2012;13(9):R79.

5 Machiels K, Joossens M, Sabino J, et al. A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. Gut 2014;63(8):1275-83.

6 Nishikawa J, Kudo T, Sakata S, et al. Diversity of mucosa-associated microbiota in active and inactive ulcerative colitis. Scand J Gastroenterol 2009;44(2):180-6.

7 Sobhani I, Tap J, Roudot-Thoraval F, et al. Microbial dysbiosis in colorectal cancer (CRC) patients. PLoS One 2011;6(1):e16393.

8 Wang T, Cai G, Qiu Y, et al. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. ISME J 2012;6(2):320-9.

9 Ahn J, Sinha R, Pei Z, et al. Human gut microbiome and risk for colorectal cancer. J Natl Cancer Inst 2013;105(24):1907-11.

10 Ley RE, Turnbaugh PJ, Klein S, et al. Microbial ecology: human gut microbes associated with obesity. Nature 2006;444(7122):1022-3.

11 Wu GD, Chen J, Hoffmann C, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science 2011;334(6052):105-8.

12 Fierer N, Hamady M, Lauber CL, et al. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. Proc.Natl.Acad.Sci.U.S.A 2008;105(46):17994-99.

13 Muyzer G, de Waal EC, Uitterlinden AG Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl.Environ.Microbiol. 1993;59(3):695-700.

14 Sim K, Cox MJ, Wopereis H, et al. Improved detection of bifidobacteria with optimised 16S rRNA-gene based pyrosequencing. PLoS One 2012;7(3):e32543.

15 Sim K, Shaw AG, Randell P, et al. Dysbiosis anticipating necrotizing enterocolitis in very premature infants. Clin Infect Dis 2015;60(3):389-97.

16 Quince C, Lanzen A, Davenport RJ, et al. Removing noise from pyrosequenced amplicons. BMC Bioinformatics 2011;12(38.

17 Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. Nat.Methods 2010;7(5):335-36.

18 Pruesse E, Quast C, Knittel K, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 2007;35(21):7188-96.

19 Edgar RC Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;26(19):2460-1.

20 Chen J, Bittinger K, Charlson ES, et al. Associating microbiome composition with environmental covariates using generalized UniFrac distances. Bioinformatics 2012;28(16):2106-13.

21 Rice P, Longden I, Bleasby A EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 2000;16(6):276-7.

22 Araujo-Perez F, McCoy AN, Okechukwu C, et al. Differences in microbial signatures between rectal mucosal biopsies and rectal swabs. Gut Microbes 2012;3(6):530-5.

23 Zoetendal EG, von WA, Vilpponen-Salmela T, et al. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl.Environ.Microbiol. 2002;68(7):3401-07.

24 O'Brien CL, Allison GE, Grimpen F, et al. Impact of colonoscopy bowel preparation on intestinal microbiota. PLoS One 2013;8(5):e62815.

25 Jalanka J, Salonen A, Salojarvi J, et al. Effects of bowel cleansing on the intestinal microbiota. Gut 2015;64(10):1562-8.

26 Sproer C, Mendrock U, Swiderski J, et al. The phylogenetic position of Serratia, Buttiauxella and some other genera of the family Enterobacteriaceae. Int J Syst Bacteriol 1999;49 Pt 4(1433-8.

27 Rehman A, Lepage P, Nolte A, et al. Transcriptional activity of the dominant gut mucosal microbiota in chronic inflammatory bowel disease patients. J Med Microbiol 2010;59(Pt 9):1114-22.

28 Sokol H, Lepage P, Seksik P, et al. Temperature gradient gel electrophoresis of fecal 16S rRNA reveals active Escherichia coli in the microbiota of patients with ulcerative colitis. J Clin Microbiol 2006;44(9):3172-7.

29 Martinez C, Antolin M, Santos J, et al. Unstable composition of the fecal microbiota in ulcerative colitis during clinical remission. Am J Gastroenterol 2008;103(3):643-8.

**Figure Legends**

*Figure 1- Timeline representation of the six samples collected during the study period. Median days of sample collection are indicated above the bar.*

*Figure 2 - Generalised UniFrac distances between biopsies and: Left, paired luminal contents; Centre, paired swabs; Right, other biopsies. \*\* indicates highly significant differences (p < 0.001). Bars indicate data median.*

*Figure 3 – A comparison of taxonomic groups found in each set of patient samples. Samples labelled as: B = Biopsy, L = Luminal contents, R = Rectal swab. The most abundant taxonomic groups, making up 90% of the dataset’s reads, are displayed.*

*Figure 4 - Generalised UniFrac distances between biopsies and: Left, paired luminal contents; Centre, paired pre-bowel preparation faecal samples; Right, other biopsies. \* indicates significant differences (p < 0.05). Bars indicate data median.*

*Figure 5 – A Comparison of taxonomic groups found in each set of patient faecal samples. Samples labelled as: Pre = Pre colonoscopy faecal sample, Lum. = Luminal Contents, Post = Faecal sample taken an average of 55 days after colonoscopy.*

*Figure 6- Clustering of samples according to four measures of the microbial community. a) The clustering of biopsy samples. b) The clustering of luminal contents. Circles indicate samples from UC patients, triangles indicate samples from CD patients and crosses indicate samples from non-IBD patients.*

**Table Legends**

*Table 1 - Patient clinical metadata and sample availability. Asterisks indicate samples that were collected but not successfully sequenced.*