Gut microbiota differs between children with inflammatory bowel disease and healthy siblings in taxonomic and functional composition: a metagenomic analysis

Knoll, R.L., Forslund, K., Kultima, J.R., Meyer, C.U., Kullmer, U., Sunagawa, S., Bork, P., Gehring, S.

American Journal of Physiology Gastrointestinal and Liver Physiology
2017 APR 01 ; 312(4): G327-G339
2017 APR 01 (first published online)
doi: 10.1152/ajpgi.00293.2016

Publisher: American Physiological Society

Articles in PresS. Am J Physiol Gastrointest Liver Physiol (December 30, 2016). doi:10.1152/ajpgi.00293.2016

Gut microbiota differs between children with Inflammatory Bowel Disease and healthy siblings in taxonomic and functional composition - a metagenomic analysis

Running Title: Metagenomic analysis of the pediatric IBD microbiome

Rebecca L. Knoll*[1], Kristoffer Forslund *[2], Jens Roat Kultima [2], Claudius U. Meyer [1], Ulrike Kullmer [1], Shinichi Sunagawa [2,3], Peer Bork [2,4,5,6]+, Stephan Gehring [1]+

*: These authors contributed equally

+: Corresponding authors

Affiliations:

[1] Children's Hospital, University Medical Center, Johannes Gutenberg University, Mainz 55131, Germany

[2] European Molecular Biology Laboratory, Structural and Computational Biology Unit, 69117 Heidelberg, Germany

[3] Institute of Microbiology, ETH Zurich, 8092 Zurich, Switzerland.

[4]Molecular Medicine Partnership Unit, University of Heidelberg and European Molecular Biology Laboratory, 69120 Heidelberg, Germany.

[5]Max Delbrück Centre for Molecular Medicine, 13125 Berlin, Germany.

[6]Department of Bioinformatics, Biocenter, Am Hubland, 97074 Würzburg, Germany.

Contact information corresponding author:

Stephan Gehring
Children's Hospital, University Medical Center, Johannes Gutenberg University
Langenbeckstraße 1
55131 Mainz, Germany
Electronic address: stephan.gehring@uni-mainz.de
Telephone: +49 (0) 6131 17-3560; Fax number: +49 (0) 6131 17-6624

Abstract

Current treatment for pediatric IBD patients is often ineffective, with serious side effects. Manipulating the gut microbiota via fecal microbiota transplantation (FMT) is an emerging treatment approach but remains controversial. We aimed to assess the composition of the fecal microbiome through a comparison of pediatric IBD patients to their healthy siblings, evaluating risks and prospects for FMT in this setting. A Case–Control (Sibling) Study was conducted analyzing fecal samples of six children with Crohn's Disease (CD), six children with Ulcerative Colitis (UC) and 12 healthy siblings by metagenomic sequencing. In addition, lifetime antibiotic intake was retrospectively determined. Species richness and diversity were significantly reduced in UC patients compared to control (MWU FDR = 0.011). In UC, bacteria positively influencing gut homeostasis e.g. *Eubacterium rectale* and *Faecalibacterium prausnitzii* were significantly reduced in abundance (MWU FDR = 0.05). Known pathobionts like *Escherichia coli* were enriched in UC patients (MWU FDR = 0.084). Moreover, *E. coli* abundance correlated positively with that of several virulence genes (SCC > 0.65, FDR < 0.1). A shift towards antibiotic resistant taxa in both IBD groups distinguished them from controls (MWU BY FDR = 0.062 in UC, MWU BY FDR = 0.019 in CD). The collected results confirm a microbial dysbiosis in pediatric UC, and to a lesser extent in CD patients, replicating associations found previously using different methods. Taken together, these observations suggest microbiotal remodeling therapy from family donors, at least for children with UC, as a viable option.

New and Noteworthy

In this sibling study, prior reports of microbial dysbiosis in IBD patients from 16S rRNA sequencing was verified using deep shotgun sequencing and augmented with insights into the abundance of bacterial virulence genes and bacterial antibiotic resistance determinants, seen against the background of data on the specific antibiotic intake of each of the study participants. The observed dysbiosis, which distinguishes patients from siblings, highlights such siblings as potential donors for microbiotal remodeling therapy in IBD.

Keywords: Metagenomics, Microbiome, pediatric gastroenterology, Inflammatory Bowel Diseases, Fecal Microbiota Transplantation (FMT)

Abbreviations:

BY FDR	Benjamini-Hochberg-Yekutieli procedure False Discovery rate (adjustment)
CD	Crohn' s Disease
CTRL	Control
(BH) FDR	Benjamini-Hochberg procedure False Discovery Rate (adjustment)
FMT	Fecal Microbiota Transplantation
GABA	gamma-Aminobutyric acid
IBD	Inflammatory Bowel Disease
KW	Kruskal-Wallis test
LRT	likelihood ratio test
mOTUs	metagenomic Operational Taxonomic Units
MWU	Mann-Whitney-U test
PCDAI	Pediatric Crohn's Disease Activity Index
PUCAI	Pediatric Ulcerative Colitis Activity Index
rRNA	ribosomal RNA
SCC	Spearman rank Correlation Coefficient
UC	Ulcerative Colitis

Introduction

The gut microbiota plays a crucial role in human physiology and host development (41), and maintaining equilibrium between the commensal microbiota and the host immune system is required for a healthy gut homeostasis.

There is increasing evidence for a contribution of the gut microbiome to the etiology of IBD (18). In the past, genome–wide association studies have revealed multiple host gene loci in both UC and CD, with alleles associated with functional aberrations of the intestinal immune system (18). Recent studies based on novel DNA sequencing methods have revealed major differences in bacterial taxonomic composition between IBD patients and healthy individuals (see Appendix Table 1). Still, it remains unclear whether these observed alterations are the cause or result of inflammation.

Since most studies on the IBD microbiome to date have relied on 16S rRNA sequencing, the functional impact of the reported dysbiosis is not well understood (23), and reported associations are largely limited to identifying more general taxa (ranging from bacterial phyla to genera) as being associated with disease, given the limitations of 16S for reliable species identification Whole genome sequencing allows higher resolution and sensitivity than the more common and less expensive 16S rRNA sequencing and offers new insights into the functional context of the IBD microbiome (e.g. abundance of metabolic pathways and the distribution of genes determining virulence or resistance to antibiotics).

Current treatment strategies for pediatric IBD patients often come with serious side effects or provide insufficient treatment responses. Therefore, there is a need for novel treatment approaches.

Manipulating the gut microbiota via fecal microbiota transplantation (FMT) appears as an intriguingly facile and harmless therapy option for children with IBD. Efficacy in the treatment of pediatric UC via fecal enemas has been suggested (22), and the administration of FMT via nasogastric tube has led to improved well-being in pediatric CD patients (44). Still, its outcome remains controversial, since results from the first two randomized placebo-controlled trials in adults

are in contradiction (29, 36). While no serious adverse events for FMT in children have been reported, data on feasibility and safety in a long-term perspective (e.g. the risk of transferring a pathogenic disease state via FMT) are missing (12). Microbiome-based validation of suitable donors might help better predict treatment outcome (11). The present pilot study therefore aims to elucidate microbiome correlates of juvenile IBD so as to help in designing criteria for when FMT might be employed, using a pediatric cohort of CD and IBD patients as well as their healthy siblings.

Methods

Cohort recruitment

Patients were selected from the IBD patient collective at the Zentrum für Kinder- und Jugendmedizin (University Children's Hospital) Mainz in Mainz, Germany. In all patients IBD diagnosis was histologically confirmed by prior endoscopy. Disease activity was assessed by the Pediatric Ulcerative Colitis Activity Index (PUCAI) (45) and the Pediatric Crohn's Disease Activity Index (PCDAI) (14). In total, fecal samples from six Ulcerative Colitis (UC) patients and six Crohn's Disease (CD) patients and from 12 controls were collected from August 2013 to July 2014. Controls were healthy siblings of the IBD patients, sharing the same parents and living in the same households, thereby sharing genetic background, environment and diet. In addition, study participants completed a questionnaire about health/disease status, living conditions, and alimentary habits. Children who had taken antibiotics during the last 2 months prior to sampling were excluded from the study. For additional cohort characteristics see Table 1. Since all study participants are older than 8 years, and thus should have relatively stable microbiomes, we do not anticipate age differences between patients and healthy siblings to have a strong impact on results.

The study design was approved by the ethics committee of the State Chamber of Physicians of Rhineland-Palatinate (Ethikkomission der Landesärztekammer Rheinland-Pfalz, reference number 837.292.13 (8977-F)).

Written informed consent was given by both parents and the study participants prior to sample collection and medical assessment.

Assessment of lifetime antibiotic intake

Antibiotic intake of the study population was retrospectively recorded, starting from birth. Therefore, all physicians/hospitals who have treated the study participants were contacted and asked to transfer their antibiotic prescription data for each child. Out of 306 person years of total participant lifetime to date, only 16 (5 %) could not be covered by this assessment.

Sample collection

Fecal samples were collected at home with help of a *stool sampling kit*. The *stool sampling kit* consisted of a plastic lining to cover the toilet, two stool sample tubes with spoons, two plastic bags and a clipping system for safe closure of the outer bag. Collected samples were stored at home between 4°C and 8°C and transferred to the laboratory within 24h. In the laboratory of the Zentrum für Kinder- und Jugendmedizin Mainz, Germany, the samples were frozen at – 80°C. All samples were gathered there and then shipped on dry ice (- 78.5°C) to the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany.

Probe processing

All probe processing was conducted by the Genomics Core Facility, EMBL, Heidelberg. DNA-Extraction and Library preparation were performed according to the protocol from Zeller et al. (49). Whole genome shotgun sequencing was executed on the Illumina HiSeq 2000/2500 (Illumina, San Diego, USA) platform. All samples were paired-end sequenced with a read length of 100 bp and a targeted sequencing depth of 5 Gbp (49).

Data analysis

Data analysis was performed at the Structural and Computational Biology Unit, EMBL, Heidelberg.

Taxonomic profiling

Using the MOCAT pipeline (21), gene sequences were annotated to their bacterial taxonomy. Since many gut bacteria are known to belong to species for which no genome yet exists in public databases, two different alignment procedures were used. The first procedure was based on metagenomic operational taxonomic units (mOTUs) (43) which also encompass uncharacterized bacteria identified in metagenomic dataset. The second procedure used species clusters defined from bacterial genomes for which publicly deposited genomes do exist (27). In this manner as complete coverage of the taxonomic composition of the samples as possible was achieved.

Analysis of bacterial diversity, species richness and evenness

Based on the mOTU taxonomic composition, Shannon diversity index, species richness and evenness were calculated using the vegan R package (http://cran.r-project.org/web/packages/vegan/index.html), for details see Zeller et al. 2014 (49).

Functional profiling

To gain insight into metabolic functions of the microbiome, the metagenomic catalogue was aligned to the Kyoto Encyclopedia of Genes and Genomes (KEGG) Database (16). For further details see Zeller et al., 2014 (49).

Identifying virulence factors

For virulence factor profiling of the metagenome samples, reads were mapped to virulence factor gene families via an annotated gene catalog as described in Kultima et al., 2016 (20). For each gene catalog entry annotated to a virulence factor gene family, we traced which sequenced taxa have genes similar enough to the reference that reads originating from them could map to it. For the virulence genes found significantly different under IBD in this study, from sequence alone their specific origin cannot be determined; on average each virulence gene family is found in 87 taxa (See Supplementary Table 2).

Identifying antibiotic resistance genes

To estimate the abundance and prevalence antibiotic resistance genes in the studied microbiomes, the reference gene catalogue was aligned to the Antibiotic Resistance Genes Database (ARDB)(25), as described in Forslund et al. (9). Metronidazole resistance genes (nim, nimA, nimB, nimC, nimD, nimF) were further annotated through bidirectional best hits between the reference gene catalog and all annotated such sequences from UniProt, otherwise as previously described. Only antibiotic classes which had been taken by at least one individual in the study cohort were considered.

For each sample, three measures of antibiotic resistance gene carriage were determined: the raw relative abundance of antibiotic resistance genes, the relative abundance of potentially resistance gene-carrying bacterial species, and the antibiotic resistance potential (abundance of antibiotic resistance genes relative to the abundance of potentially resistance gene-carrying species). The abundance of potentially resistant species was calculated as the fraction of genetic material from each sample that map to species belonging to genera with known examples of species carrying resistance genes of the appropriate type, in the reference genome database used (see Forslund et al. (9)).

Statistical analysis

In subsequent analyses, all read/base counts were transformed into relative abundances (by division by the total number of reads/bases sequenced per sample) as described previously by Zeller et al., 2014 (49).

In order to identify significant differences between sample categories for each metagenomic feature, a nonparametric Kruskal-Wallis-test (KW) was performed whenever the UC-, CD- and control-group were compared. This was followed by pairwise Mann-Whitney-U-tests (MWU) as a post-hoc procedure.

For data sets not normally distributed, the Wilcoxon signed-rank test was employed to compare sib pairs directly. Correlation analyses were performed using Spearman rank's correlation coefficient (SCC), in order to fit potentially non-normal data.

In cases of multiple testing p-value correction via the Benjamini-Hochberg false discovery rate (FDR)(2) or the Benjamini-Hochberg-Yekutieli procedure (BY) (3) was performed. For corrected p-values significance threshold was set at < 0.5; corrected p-values ranging $0.05 < FDR < 0.1$ we refer to as approaching significance.

To reduce the number of tests for taxonomic abundances and thereby increase statistical power, a pre-selection of hypotheses was applied by collecting microbiome associations described in the literature, including taxa directly associated with IBD or associations projected from such reports by phylogenetic relationships. This collection consisted of all directly reported IBD-associated taxa in the referenced works and their containing taxonomic superclades (see Appendix Table 1). In addition, thus far uncharacterized taxa sorting immediately under each clade previously reported associated with IBD were added, to account for incompleteness of sequenced genomes (e.g. unknown Firmicutes for class analysis, as Firmicutes were reported to be different on the Phylum level, or unknown Clostridia for analysis on the order level, as Clostridia were previously reported to be associated with IBD on the class level); the mOTU technology used here was explicitly designed to allow detection of unculturable human gut taxa.

A likelihood ratio test (LRT) was conducted to test whether the abundance of potentially resistant species can be explained by antibiotic selection pressure (as represented through lifetime antibiotic intake) or not. To calculate these likelihood ratios, mixed-effect linear models taking into account 1) overall lifetime antibiotic intake and disease status, 2) lifetime antibiotic intake alone, 3) disease status alone and 4) solely a background constant model were compared to each other. Prior to this analysis in each case the general assumptions for linear models were tested (*independence, absence of co-linearity, homoscedasticity, linearity and normality of the residuals and absence of influential data points)* using graphpad (http://graphpad.com/quickcalcs/PValue1.cfm) with p-values calculated for a χ^2-distribution with one degree of freedom.

Results

Pediatric UC patients have reduced microbial biodiversity compared to their healthy siblings

Tests were conducted on the Shannon diversity index to characterize overall gut biodiversity. Only in UC patients reduction of biodiversity reached significance compared to controls (MWU, n_{UC} = 6/$n_{Control}$ = 12, FDR = .011) (Figure 1). Analyzing the data on aggregate, children with UC and CD both displayed significantly reduced species richness (MWU, n_{UC} = 6/ n_{CD} = 6/ $n_{Control}$ = 12, FDR = .011, FDR= .045 respectively) (Figure 1) compared to the set of control children. This reduction was significant also when matching each IBD patient to their control sibling in a pairwise test (WT, n_{UC} = 6/ n_{CD} = 5/ $n_{Control}$ = 11, FDR= .048). Similar observations could be made for microbial community evenness (see Figure 1).

Differences in taxonomic abundances from former 16S rRNA analyses are validated by WGS

Considering those bacterial taxa previously noted as being altered in prevalence or abundance in IBD (cf Appendix Table 1), many of these findings could be recaptured as significant in the present cohort even with its limited size, comparing the set of all controls to the set of each subset of IBD subjects (Figure 2). Pairwise testing of siblings (Wilcoxon rank test) showed the same trends as the aggregated case-control comparison, though due to limited sample numbers, approached significance only for associations of Clostridium (Clostridiacae) and *Eubacterium rectale* to IBD (Supplementary Table 1).

If we do not restrict analysis to validation of 16S results only, power is reduced due to the increased number of tested hypotheses. However, 55% (17 out of 31) of the findings described above also show associations with IBD approaching significance without such preselection and with subsequently more stringent FDR correction (Supplementary Table 3). The previously described associations constitute the majority of significant associations found in the unrestricted search, suggesting most obvious associations already were reported at least once from 16S data. However, another four taxa approached significance in the unrestricted test, which had not previously been linked to IBD. On the taxonomic order level, a group of uncharacterized Firmicutes were

diminished in the UC population compared to the healthy cohort. At the level of bacterial genera, *Anaerococcus* ($MR_{(Control)}$ 10.83, $MR_{(UC)}$ 18.33, $MR_{(CD)}$ 10.00, KW FDR = .093) and a thus far uncharacterized genus belonging to Clostridiales ($MR_{(Control)}$ 17.08, $MR_{(UC)}$ 5.83, $MR_{(CD)}$ 10.00, KW FDR = .093), differed in abundance between IBD subjects and controls. Most importantly, *Clostridium ramosum* was significantly enriched in UC patients compared to the set of controls (FDR <.001).

Enrichment of pathobiont species and depletion of commensals

Reduction of gut homeostasis positively influencing *E. rectale* and *Faecalibacterium prausnitzii* were observed in UC and CD samples compared to the set of controls samples, but only the comparison between control and UC samples approached significance (Figure 2). Similarly, the pathobiont *Escherichia coli* was increased in abundance in both IBD conditions, but approached significance only for UC samples (Figure 2). An increase in *Ruminoccus gnavus* approached significance (at the species cluster level, not significant at mOTU level). *Bilophila wadsworthia* was decreased in UC, also approaching significance. The pathogen *Fusobacterium nucleatum* (resolved at the species cluster level, not significant at mOTU level) was found in IBD patients but not in the healthy siblings, with this difference approaching significance for UC patients (Figure 2).

The abundance of specific virulence factors correlates with the abundance of certain bacterial species

Figure 3 gives an overview of the specific virulence genes, which were significantly increased in the UC population (CD and UC samples compared to the full set of control samples) of the present study. To identify correlations with species present in the study cohort, a Spearman correlation analysis was conducted. Most virulence factors were correlated (Spearman test FDR < 0.1) with abundance of *E. coli*. All of these genes, except hslT, were previously described and identified as occurring in strains of *E. coli* (see Supplementary Table 2).

Abundance of certain families of genes facilitating the survival of bacteria within macrophages and the evasion of immune response correlated with abundance of particular gut microbial species (*Bacteroides vulgatus*/*Gemella morbillorum*). Virulence genes adhD, aslA, sitA, Ndk, fur, gcvT and fepA had no significant correlation with any species found significantly different in abundance between controls and IBD patients. These observations suggest a broader involvement of those genes in UC pathogenesis, such that they may play a role which is not reducible to a simple taxonomic difference.

Abundance of GABA shunt genes is elevated in the IBD microbiome

A Kruskal-Wallis test was performed to analyze KEGG pathways and KEGG modules for significant differences in gene abundance comparing all IBD cases to all controls. Only the KEGG module representing the GABA (gamma-Aminobutyric acid) shunt, considered a potential virulence factor by some investigators (8), showed such a significant difference with a mean rank of 6.67 for the Control group, 18.67 for UC and 18.00 for CD; BY <.000.

Little difference in antibiotic intake

Neither the cumulative lifetime antibiotic intake per person nor the mean antibiotic intake per life year was significantly different comparing the full groups of control, UC and CD.

In all three groups cephalosporin was the most commonly prescribed antibiotic, followed by penicillin, with highest intake for both in the CD group (Figure 4). Metronidazole was only taken by IBD patients, and a tetracycline antibiotic was only taken once by one subject in the control group.

The resistome of IBD patients and their healthy siblings

In accordance with previous studies (9), tetracycline was the antibiotic for which the most resistance capacity was found in the microbiomes of the study cohort (Figure 5), followed by cephalosporin. No significant difference in the abundance of resistance genes against any antibiotic

class was observed comparing the full groups of control, UC and CD samples on aggregate (See Appendix Figure 1). While we augment the ARDB antibiotic resistance gene database with known resistance genes for metronidazole, the antibiotic most commonly prescribed to IBD patients, such genes are still largely uncharacterized. As a result, it is possible we fail to observe an association between such resistance and metronidazole exposure history.

The high relative fraction of potentially resistant bacterial species does distinguish the gut microbiome of IBD patients from the microbiomes of controls. Higher abundances of bacterial species where strains with resistance genes have been sequenced were observed for all antibiotic classes (BY FDR = 0.062 for UC; BY FDR = 0.019 for CD) (Fig. 5, see also Appendix Figure 1).

To evaluate whether this high abundance of potentially resistant species was due to antibiotic selection pressure, a likelihood ratio test was performed testing fits of the data to nested mixed-effects linear models. The total abundance of potentially resistant species was modeled from overall lifetime antibiotic intake of each individual together with disease status (Control, UC, CD). A model including both antibiotic intake and disease status did not fit significantly better than a model including group affiliation alone (LRT test, $P > 0.4$). We therefore at this point cannot refer the greater relative prevalence of potentially resistant microbial taxa in the IBD patients than in their healthy siblings' microbiomes to any higher lifetime antibiotic consumption in these children.

Discussion

Our results from comparatively analyzing the microbiome of pediatric IBD patients constitute a deep sequencing-based consolidated validation of findings from previous research which relied mostly on 16S rRNA sequencing (see Appendix Table 1 and Supplementary Table 1). To our knowledge, this is the first study which contrasts metagenomic analysis of antibiotic resistance genes in relation to lifetime antibiotic intake of study participants in an IBD cohort. We also provide the first direct metagenomic functional profiling of pediatric IBD samples, including analysis of virulence factor differential abundance in cases versus controls. By a study design using as controls

healthy siblings of IBD participants sharing similar diet and living circumstances, we are able to minimize potential confounding factors. In addition, although we expected higher antibiotic consumption within the patient population, which could have provided an alternate interpretation of the observed dysbiosis, antibiotic intake was not significantly higher among pediatric IBD patients than other children.

Dysbiosis of the intestinal microbiome may involve reduced microbial diversity, enrichment of potentially pathogenic taxa and/or depletion of beneficial microbiota (31). Microbial dysbiosis has been observed in numerous disease conditions, e.g. colorectal carcinoma (40), hypertension (48), psoriatic arthritis (37), diabetes mellitus type 1 (42) and type 2 (17) and especially in pediatric IBD (see Appendix Table 1).

Several research groups observed how microbial diversity and the amount of beneficial commensals is reduced in the microbiome of both pediatric IBD conditions (10, 28, 39). At the same time, potential pathobionts like *E. coli* are more prevalent in the IBD microbiome. These findings are quantifiable using the present setup and we were able to validate them. While in the present study, we find more microbiome associations of UC than of CD, previous studies found stronger such associations to CD (see Appendix Table 1 and Supplementary Table 1). As we observe similar trends in both IBD subtypes for most comparison, we interpret this as that our small sample size may make us underpowered to detect many of the associations to CD, which larger studies using similar methodology may be able to recover.

Decrease of gut homeostasis-promoting species in the IBD microbiome

The gut homeostasis-promoting bacterial species *F. prausnitzii* and *E. rectale* were depleted in both IBD conditions in the present cohort. This observation is in consensus with former microbiome analysis of UC and CD patients relying on 16S rRNA sequencing (6, 10, 26). Both species are important producers of short-chain fatty acids (SCFA), including propionate and butyrate (34).

SCFAs are among the most central metabolites produced by microbes, influencing both colon physiology and the intestinal immune system (30).

Increase of E. coli and F. nucleatum in pediatric IBD

In contrast, pathobionts such as *E. coli* and *F. nucleatum* were enriched in the IBD microbiome. This is in agreement with former 16S rRNA analyses, where *E. coli* was increased in the IBD microbiome (6, 28, 39). *E. coli*, especially its subspecies *adherent-invasive E. coli (AIEC)*, has been directly implicated in IBD development (23). A strong correlation was observed between *E. coli* abundance and that of genes implicated in the expression of fimbriae or leading to the adhesion of bacteria to the intestinal mucosa. Those genes are characteristic of pathogenic *E. coli.* subspecies (e.g. *enterohaemorrhagic, uropathogenic E. coli*) (1, 5). However, most of these genes are found also in many other bacterial taxa, and as such it cannot be ruled out that other bacterial species also may contribute to the observed increase of virulence capacity in IBD patients (see Supplementary Table 2).

The strong correlation of the abundance of virulence genes enriched in the UC microbiome with *E. coli* abundance suggests that this species is a central driver of a functional shift towards virulence in UC patients of the present cohort, underscoring its potential role as an IBD partial cause.

Survival strategies of bacteria – relating virulence capacity to pathology

The virulence genes enriched in the UC microbiome included TonB, sitA, IroN, fepA and Fur. They are all genes implicated in metal acquisition of Enterobacteria, where the last is the key regulator of iron transport systems (33). The enriched iron uptake/transport systems contribute to the virulence of their host bacteria by counteracting the so called nutritional immunity which deprives bacteria of iron. The homeostasis of intracellular iron concentrations is maintained by Fur (the ferric uptake regulator), since high intracellular accumulation also is toxic for the bacteria (32). In its role as a sensor for iron availability, Fur is also able to regulate directly or indirectly (through iron concentration) the expression of several other virulence factors in pathogenic bacteria.

Adaptation of microbes in IBD to a lower environmental pH, as a feature of inflammation, is reflected by higher abundance of the genes encoding the GABA shunt module in the IBD microbiomes studied here (Figure 4B). The GABA shunt supplies bacteria with nitrogen and promotes their survival in acidic conditions and other environmental stresses (8). Feehily and Karatzas therefore suggested that this module could facilitate microbial pathogenicity.

Exploration of the IBD resistome in correlation with antibiotic consumption

Several studies have displayed a significant relation of antibiotic intake in childhood with the later development of IBD, particularly Crohn's Disease (13, 19, 47). Analogously, antibiotic consumption often exacerbates dysbiosis, seen as a reduced species richness and the bloom of pathogens (e.g. *Clostridium difficile*) in the gut microbiome (46). In addition, administration of antibiotics not only leads to shifts in taxonomic composition, but also induces the lateral transfer and spread of antibiotic resistance genes in a microbial ecosystem (38).

Notably, the quantification of antibiotic resistance genes in the microbiomes of the present cohort revealed no significant difference in the relative abundance of these genes between IBD patients and their healthy siblings, and similar retrospective assessments of antibiotic intake in both groups. In other words, while it seems unlikely the IBD discordance within the sib pairs can be explained by differences between siblings in antibiotic exposure, the possibility remains that the families themselves differ from other families in both antibiotic exposure and IBD prevalence, highlighting the inter-sibling differences in microbiome composition as rather reflecting IBD risk keeping antibiotic exposure constant. It is further conceivable that resistant bacteria acquired from intra-host adaptation to antibiotic treatment or from healthcare setting exposure may propagate between family members living in close contact. In contrast, a significant shift in both IBD conditions towards a higher abundance of species potentially carrying antibiotic resistance genes was revealed. This measure could not be reliably predicted from individual antibiotics use history within the present dataset, however.

Besides antibiotic treatment, intake of IBD medications has recently been lined out as a covariate of microbiome variation in a large-scale study (7). Due to cohort size, we had not the power to detect such effects, although they need to be assumed (Table 1).

Clinical implementation of study results – are siblings suitable donors in FMT?

It has been argued whether or not siblings or near relatives are suitable choices for donors in FMT, because they share genetic and environmental risk factors with the patients (4). A recent study on the IBD microbiome in pediatric patients and their first-grade family members suggested a disease-related microbial and metabolomic state in some relatives, suggesting this risk may be relevant (15). Moreover, Li et al. (24) recently outlined the importance of the compatibility of the donor and recipient microbiomes to receive persistent responses from FMT treatment. They concluded that each recipient will need an idiosyncratic donor to attain successful colonization by allogenic strains. Donor strains will colonize more successfully when the species is already present in the recipient microbiome. In the present study, gut microbial diversity (Shannon diversity index) and particularly species richness was significantly diminished in IBD patients when comparing each patient with their sibling. This observation of overall dysbiosis in pediatric IBD suggests that healthy siblings, who have more analogy than total strangers, but still provide a healthier microbiome composition, may indeed be suitable as donors from an efficacy point of view.

Future outlook

This study must be considered a pilot effort, by design and given the limited number of subjects included. A future study with more pediatric IBD patients and healthy siblings would yield greater statistical power and will hereby provide both further validation (esp. intra-sibling disparity) and the potential discovery of novel associations, which was very limited at the present time for gut microbial taxonomic correlates of IBD, as our research was mainly based on previous 16S rRNA findings (Appendix Table 1). Likewise, it would be very interesting to further employ direct sib pair comparisons, e.g. regarding the analyses of virulence and resistance genes. However, in the present

study the efficacy of direct sib pair comparison was limited for reasons of statistical power, and most analysis therefore relied on aggregated analysis (comparing the sets of all control samples to the samples in the UC/CD-cohorts), obscuring intra-sibling disparity. The additional analysis of tissue samples should be considered in the future, as recent research has demonstrated a discrepancy between luminal and mucosal samples in microbial composition (35) and as deep sequencing of bioptic samples is becoming more feasible. Today, a crucial limitation inherent to all functional metagenomic studies, excluding those using direct selection, is that only known gene families can be quantified for generation of functional profiles. Consequently, it is possible that results are reflecting research biases e.g. virulence factors to be better known in some taxa than others. With databases of curated resistance determinants growing in scope, and with technologies for large-scale functional screen metagenomics gradually maturing, these difficulties should decrease in coming years. We anticipate further studies building on top of what is reported here.

Conclusion

Our observations of dysbiosis, higher abundance of virulence factors and a shift towards gut bacterial taxa with known resistance gene carrying strains in the IBD population provide evidence for siblings as appropriate potential stool donors for FMT treatment of afflicted children. Pediatric IBD patients differ from their healthy siblings in several regards that can be linked to the disease. Employing such siblings as donors would further facilitate the possibility of “home-done” transplantation. Furthermore, the ethical, esthetic and clinical barriers against stool transplants likely are easier to overcome with respect to a close relative donor than in the case of an outside / adult stool donor. The same holds regarding considerations of the potential transmission of infectious diseases and the adaptations of the microflora to dietary habits, where siblings dwelling in the same home already should be similar.

Finally, our observations suggest utility of taking metagenomic measurements of microbial dysbiosis in both donor and recipient pre- and post-procedure in order to correlate FMT outcome with microbiome characteristics.

Appendix

Appendix Table 1: **Bacteria found associated with inflammatory bowel diseases by previous work groups = literature selection.**

↑/↓ = increase/decrease in patients with IBD in comparison with healthy controls; ↑CD /↓CD = increase/decrease in patients with Crohn's Disease; ↑UC /↓UC = increase/decrease in patients with Ulcerative Colitis; Superscripted numbers are referring to the authors, see reference table below.

Phylum	Class	Order	Family	Genus	Species
Bacteroidetes ↓[2] ↑[4,8] ↑CD[5]	Bacteroidetes	Bacteroidales ↓CD[6]	Porphyromonadaceae	Odoribacter ↓[1] Parabacteroides ↓CD[6]	
			Bacteroidaceae	Bacteroides ↑[8] ↓[17] ↓CD[6]	*B. thetaiotaomicron* ↓[2] *B. vulgatus* ↓CD[6] *B. caccae* ↓CD[6]
			Prevotellaceae	Prevotella ↓UC[3]	
			Rikenellaceae ↓CD[6]	Alistipes	
Firmicutes ↓[2,4] ↓UC[14]	Clostridia ↓UC[14] ↓CD[3,17]	Clostridiales ↓CD[3,6]	Lachnospiraceae ↓[1, 2] ↓CD[6]	Roseburia ↓[1] ↓CD[3,6]	*R. hominis* ↓UC [7] *R. faecis* ↓CD[5] *R. intestinalis* ↓CD[6]s
				Coprococcus ↓CD[6]	*C. eutactus* ↓CD[5] *C. comes* ↓CD[6]s
			Clostridiaceae	Clostridium ↓[8]	*C. leptum* ↓[9] *C. nexile* ↓CD[6]s *C. bolteae* ↓CD[6]s
				Blautia ↓CD[6]s	*B. coccoides* ↓[9] *B. hanseni* ↓CD[6]
				Dorea ↓CD[6]	
				Butyricicoccus	*B. pullicaecorum* [1] ↓[16]
			Ruminococcaceae ↓CD[1,3,5,6]	Acetivibrio [1]	
				Ruminococcus ↓CD[1] ↓CD[6]	*R. gnavus* ↑CD[3] ↓CD[6] *R. torques* ↓CD[6]
				Faecalibacterium ↓[1] ↑CD [10] ↓CD[3,6,11,17,19] ↓UC[12]	*F. prausnitzii* ↓UC [7] ↓CD[6] ↓CD[9]
				Oscillospira ↓CD[5,6]	
				Subdoligranulum ↓CD[5]	
			Peptococcaceae ↓CD[3]	Peptococcus ↓CD[3]	
				Phascolactobacterium ↓[1]	

Phylum	Class	Order	Family	Genus	Species
			Eubacteriaceae	Eubacterium	*E. rectale* ↓CD[6]
	Bacilli ↑CD[3]	Lactobacillales↓[17]↑CD[3]	Leuconostocaceae ↓UC[1] ↑CD[3]		
			Lactobacillaceae	Lactobacillus ↑[13]↑ CD[3,6]	
			Streptococcaceae↓UC[3]	Streptococcus↑CD[6]↓UC[3]	
			Enterococcaceae	Enterococcus ↑ CD[6]	
		Gemellales	Gemellaceae ↑ CD[6]	Gemella	*G. morbillorum* ↑ CD[6]
	Erysipelotrichi ↑UC[14]	Erysipelotrichiales	Erysipelotrichiaceae ↓CD[6] ↑ CD[6]s	Catenibacterium ↓UC[3]	
	Negativicutes	Selenomonadales	Acidaminococcaceae	Acidaminococcus ↑CD[3]	
			Veillonellaceae↑ CD[6]	Veillonella ↑CD[3]	*V. parvula* ↑ *CD[6]*
				Dialister ↓CD[6]	
Actinobacteria ↑[2]	Actinobacteridae	Bifidobacteriales	Bifidobacteriaceae↓CD[6] ↑CD[6]	Bifidobacterium ↓[1,11]↑[13]	*B. bifidum* ↓CD[6] *B. longum* ↓CD[6] *B. adolescentis*↓CD[6] *B. dentum* ↓CD[6] *B. infantis*
	Coriobacteridae	Coribacteriales	Coriobacteriaceae [1]		
Proteobacteria ↑[2]↑UC14 ↑CD[3,5]	Gammaproteobacteria ↑UC[14] ↑CD[3]	Enterobacteriales ↑[18]↑CD[3]	Enterobacteriaceae ↑CD[3,4,6]	Escherichia ↑[1]↑CD[19]	*E. coli* ↑CD[6]s ↑[9,11]↑UC[14] AIEC↑ [1]↑CD[15]
				Shigella ↑[1] ↑CD[3,19]	
		Aermonadales	Aeromonadaceae ↑CD[3]	Aeromonas ↑CD[3]	
		Pasteurellales	Pasteurellaceae↑ CD[6]	Haemophilus (spp.) ↑ CD[6]	*H. parainfluenzae*↑ CD[6]
	Betaproteobacteria	Neisseriales	Neisseriaceae ↑CD[6]	Eikenella	*E. corrodens* ↑ CD[6]
		Burkholderiales	Sutterellaceae	Sutterella ↓CD[6]	
	Deltaproteobacteria	Desulfovibrionales	Desulfovibrionaceae	Bilophila ↓CD[6]	*B. wadsworthia* ↓[20]
				Desulfovibrio	
	Alphaproteobacteria	Rhizobiales ↑CD[5]	Bradyrhizobiaceae↑CD[5]		
Fusobacteria ↑CD[3]↑UC[14]	Fusobacteria ↑CD[3]	Fusobacteriales ↑CD[3]	Fusobacteriaceae ↑CD[3,6] ↓CD[6]	Fusobacterium ↑CD[3]	*F. nucleatum* ↑ CD[6]
Spirochaetae	Spirochaetes ↑UC[14]	Spirochaetales			
Verrucomicrobia↓UC[14,18]	Verrucomicrobiae ↓UC[18]	Verrucomicrobiales ↓UC[18]	Verrucomicrobiaceae ↓UC[18]		
Lentisphaerae ↓UC[14]					
Tenericutes ↓[3]	Mollicutes ↓[3]	Anaeroplasmatales ↓[3]	Anaeroplasmataceae ↓[3]	Asteroleplasma ↓[3]	

Ref-Nr.	Author	Methods	Study cohort
1	Morgan et al., 2012	• 16S rRNA-sequencing • **WGS** of 11 fecal samples	• 121 CD • 75 UC • 27 Controls • Age: 13-45 y
2	Frank et al., 2007	• Q-PCR -rRNA sequencing	• 190 CD, UC, Controls
3	Willing et al., 2010	• 16S rRNA-sequencing	• 29 CD • 16 UC • 35 Controls • Age: 40-67 y
4	Walker et al., 2011	• 16S rRNA- sequencing	• 6 CD • 6 UC • 5 Controls • Age: 24-73 y
5	Kaakoush et al., 2012	• 16S rRNA- sequencing	• 19 CD (n.o.) • 21 Controls • Age: 5-15 y
6	Gevers et al., 2014	• 16S rRNA-sequencing • **WGS** of 43 Fecal samples	• 447 CD (n.o.) • 221 Controls • Age: 3-17 y
7	Machiels et al., 2013	• 16S rRNA sequencing	• 127 UC • 87 Controls
8	Andoh et al., 2011	• 16S rRNA sequencing	• 31CD • 31 UC • 30 Controls
9	Duboc et al., 2013	• 16S rRNA-sequencing	• 12 CD • 30 UC • 29 Controls • Age: 20-60 y
10	Hansen et al., 2012	• 16S rRNA-sequencing	• 13 CD • 12 UC • 12 Controls • Age: Children
11	Schwiertz et al., 2010	• 16S rRNA-sequencing	• 69 IBD • 25 Controls • Age: 1-20
12	Varela et al., 2013	• qRT-PCR	• 116 UC patients • 29 first degree relatives • 31 Controls • Age 18–75 y
13	Wang et al., 2014	• 16S rRNA-sequencing	• 36 CD • 63 UC • 21 Controls
14	Michail et al., 2012	• 16S rRNA-sequencing	• 27 UC • 26 Controls • Mean Age: 13,5 y
15	Martinez-Medina et al., 2009	• Colony dependent Rep-PCR • Pulsed field gel electrophoresis • Adhesion and invasion assays	• 20 CD • 28 Controls
16	Eeckhaut et al., 2013	• 16S rRNA-sequencing	• 51 CD • 91 UC • 88 Controls • Mean Age: 40 y
17	Aomatsu et al., 2012	• 16S rRNA gene sequencing • T-RFLP- analysis	• 10 CD • 14 UC • 27 Controls

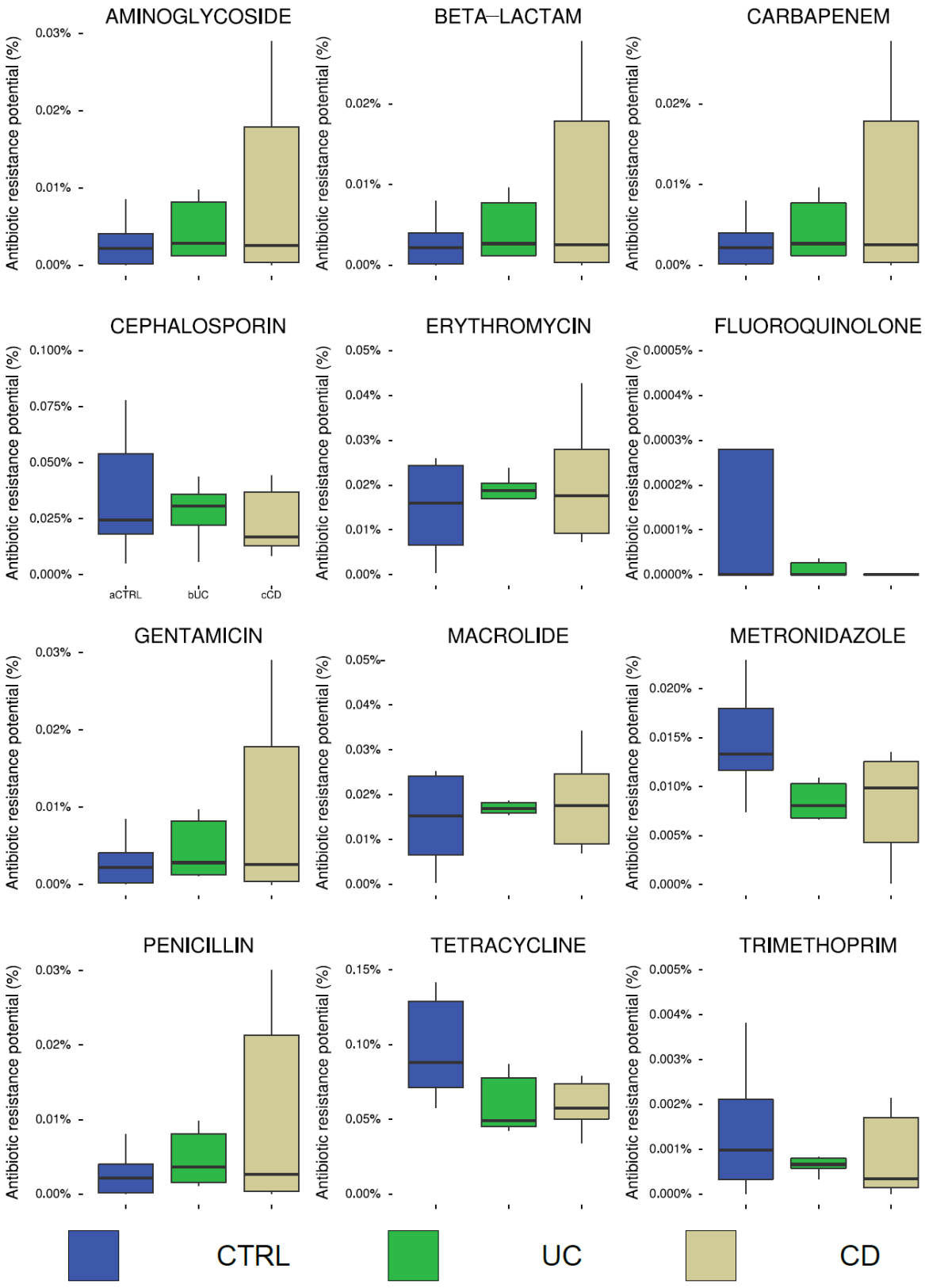
Ref-Nr.	Author	Methods	Study cohort
			• Age: 1-18 y
18	Papa et al., 2012	• 16S rRNA-sequencing	• 23 CD • 43 UC • 24 Controls • Mean Age: 13 y
19	Thorkildsen et al., 2013	• 16S rRNA-sequencing	• 30 CD • 33 UC • 33 Controls • Mean Age: 33 y
20	Jia W et al., 2012	• PCR of the dsrAB gene	• 20 CD • 14 UC • 18 Controls

References for Appendix Table 1

1. Morgan XC, Tickle TL, Sokol H, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome biology.* 2012;13(9):R79.
2. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proceedings of the National Academy of Sciences of the United States of America.* 2007;104(34):13780-13785.
3. Willing BP, Dicksved J, Halfvarson J, et al. A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology.* 2010;139(6):1844-1854.e1841.
4. Walker AW, Sanderson JD, Churcher C, et al. High-throughput clone library analysis of the mucosa-associated microbiota reveals dysbiosis and differences between inflamed and non-inflamed regions of the intestine in inflammatory bowel disease. *BMC microbiology.* 2011;11:7.
5. Kaakoush NO, Day AS, Huinao KD, et al. Microbial dysbiosis in pediatric patients with Crohn's disease. *Journal of clinical microbiology.* 2012;50(10):3258-3266.
6. Gevers D, Kugathasan S, Denson LA, et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell host & microbe.* 2014;15(3):382-392.
7. 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.* 2013.
8. Andoh A, Imaeda H, Aomatsu T, et al. Comparison of the fecal microbiota profiles between ulcerative colitis and Crohn's disease using terminal restriction fragment length polymorphism analysis. *Journal of gastroenterology.* 2011;46(4):479-486.
9. Duboc H, Rajca S, Rainteau D, et al. Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. *Gut.* 2013;62(4):531-539.
10. Hansen R, Russell RK, Reiff C, et al. Microbiota of de-novo pediatric IBD: increased Faecalibacterium prausnitzii and reduced bacterial diversity in Crohn's but not in ulcerative colitis. *The American journal of gastroenterology.* 2012;107(12):1913-1922.
11. Schwiertz A, Jacobi M, Frick JS, Richter M, Rusch K, Kohler H. Microbiota in pediatric inflammatory bowel disease. *The Journal of pediatrics.* 2010;157(2):240-244.e241.
12. Varela E, Manichanh C, Gallart M, et al. Colonisation by Faecalibacterium prausnitzii and maintenance of clinical remission in patients with ulcerative colitis. *Alimentary pharmacology & therapeutics.* 2013;38(2):151-161.
13. Wang W, Chen L, Zhou R, et al. Increased proportions of Bifidobacterium and the Lactobacillus group and loss of butyrate-producing bacteria in inflammatory bowel disease. *Journal of clinical microbiology.* 2014;52(2):398-406.
14. Michail S, Durbin M, Turner D, et al. Alterations in the gut microbiome of children with severe ulcerative colitis. *Inflammatory bowel diseases.* 2012;18(10):1799-1808.
15. Martinez-Medina M, Aldeguer X, Lopez-Siles M, et al. Molecular diversity of Escherichia coli in the human gut: new ecological evidence supporting the role of adherent-invasive E. coli (AIEC) in Crohn's disease. *Inflammatory bowel diseases.* 2009;15(6):872-882.

16. Eeckhaut V, Machiels K, Perrier C, et al. Butyricicoccus pullicaecorum in inflammatory bowel disease. *Gut.* 2013;62(12):1745-1752.
17. Aomatsu T, Imaeda H, Fujimoto T, et al. Terminal restriction fragment length polymorphism analysis of the gut microbiota profiles of pediatric patients with inflammatory bowel disease. *Digestion.* 2012;86(2):129-135.
18. Papa E, Docktor M, Smillie C, et al. Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. *PloS one.* 2012;7(6):e39242.
19. Thorkildsen LT, Nwosu FC, Avershina E, et al. Dominant fecal microbiota in newly diagnosed untreated inflammatory bowel disease patients. *Gastroenterology research and practice.* 2013;2013:636785.
20. Jia W, Whitehead RN, Griffiths L, Dawson C, Bai H, Waring RH, Ramsden DB, Hunter JO, Cauchi M, Bessant C, Fowler DP, Walton C, Turner C, and Cole JA. Diversity and distribution of sulphate-reducing bacteria in human faeces from healthy subjects and patients with inflammatory bowel disease. *FEMS Immunol Med Microbiol* 65: 55-68, 2012.

492 Appendix Figure 1



493

Appendix Figure 1: **Antibiotic resistance potential for the different classes of antibiotics which have been used by the study population.** The antibiotic resistance potential is shown as box plots. No significant difference in antibiotic resistance potential was observed.

Acknowledgements

This publication comprises results from R. L. Knoll's medical doctor thesis.

Grants

This work was supported by the "Naturwissenschaftlich-medizinsches Forschungszentrum (NMFZ) ", University Medical Centre of the Johannes Gutenberg University, Mainz, Germany; and FP7: MetaCardis, grant agreement [HEALTH-2012-305312]. The funding bodies had no direct role in the design or conduct of the study, collection, management, analysis or interpretation of the data, or preparation, review or approval of the manuscripts.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

Author Contributions

R.L.K.: Study concept and design; collection, analysis and interpretation of data; drafting of the manuscript. K.F.: Study concept and design, analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content. J.R.K. and S.S.: analysis of data; critical revision of the manuscript for important intellectual content. U.K.: Study design; acquisition of data, critical revision of the manuscript for important intellectual content. C.U.M.: contributed to the design of the study, critical revision of the manuscript for important intellectual content. P.B.: Study concept and design; critical revision of the manuscript; obtained funding; S.G.: Study concept and design; acquisition of data; critical revision of the manuscript for important intellectual content; obtained funding. All authors approved the final version of the manuscript.

References

1. **Bahrani-Mougeot FK, Buckles EL, Lockatell CV, Hebel JR, Johnson DE, Tang CM, and Donnenberg MS**. Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic Escherichia coli virulence determinants in the murine urinary tract. *Molecular microbiology* 45: 1079-1093, 2002.
2. **Benjamini Y, and Hochberg Y**. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* 57: 289-300, 1995.
3. **Benjamini Y, and Yekutieli D**. The control of the false discovery rate in multiple testing under dependency. 1165-1188, 2001.
4. **Borody TJ, Paramsothy S, and Agrawal G**. Fecal microbiota transplantation: indications, methods, evidence, and future directions. *Current gastroenterology reports* 15: 337, 2013.
5. **Donnenberg MS, Tacket CO, James SP, Losonsky G, Nataro JP, Wasserman SS, Kaper JB, and Levine MM**. Role of the eaeA gene in experimental enteropathogenic Escherichia coli infection. *The Journal of clinical investigation* 92: 1412-1417, 1993.
6. **Duboc H, Rajca S, Rainteau D, Benarous D, Maubert MA, Quervain E, Thomas G, Barbu V, Humbert L, Despras G, Bridonneau C, Dumetz F, Grill JP, Masliah J, Beaugerie L, Cosnes J, Chazouilleres O, Poupon R, Wolf C, Mallet JM, Langella P, Trugnan G, Sokol H, and Seksik P**. Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. *Gut* 62: 531-539, 2013.
7. **Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, Kurilshikov A, Bonder MJ, Valles-Colomer M, Vandeputte D, Tito RY, Chaffron S, Rymenans L, Verspecht C, De Sutter L, Lima-Mendez G, D'Hoe K, Jonckheere K, Homola D, Garcia R, Tigchelaar EF, Eeckhaudt L, Fu J, Henckaerts L, Zhernakova A, Wijmenga C, and Raes J**. Population-level analysis of gut microbiome variation. *Science (New York, NY)* 352: 560-564, 2016.
8. **Feehily C, and Karatzas KA**. Role of glutamate metabolism in bacterial responses towards acid and other stresses. *J Appl Microbiol* 114: 11-24, 2013.
9. **Forslund K, Sunagawa S, Kultima JR, Mende DR, Arumugam M, Typas A, and Bork P**. Country-specific antibiotic use practices impact the human gut resistome. *Genome research* 23: 1163-1169, 2013.
10. **Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M, Morgan XC, Kostic AD, Luo C, Gonzalez A, McDonald D, Haberman Y, Walters T, Baker S, Rosh J, Stephens M, Heyman M, Markowitz J, Baldassano R, Griffiths A, Sylvester F, Mack D, Kim S, Crandall W, Hyams J, Huttenhower C, Knight R, and Xavier RJ**. The treatment-naive microbiome in new-onset Crohn's disease. *Cell host & microbe* 15: 382-392, 2014.
11. **Grinspan AM, and Kelly CR**. Fecal Microbiota Transplantation for Ulcerative Colitis: Not Just Yet. *Gastroenterology* 149: 15-18, 2015.
12. **Hourigan SK, and Oliva-Hemker M**. Fecal microbiota transplantation in children: a brief review. *Pediatric research* 80: 2-6, 2016.
13. **Hviid A, Svanstrom H, and Frisch M**. Antibiotic use and inflammatory bowel diseases in childhood. *Gut* 60: 49-54, 2011.
14. **Hyams JS, Ferry GD, Mandel FS, Gryboski JD, Kibort PM, Kirschner BS, Griffiths AM, Katz AJ, Grand RJ, Boyle JT, and et al.** Development and validation of a pediatric Crohn's disease activity index. *Journal of pediatric gastroenterology and nutrition* 12: 439-447, 1991.
15. **Jacobs JP, Goudarzi M, Singh N, Tong M, McHardy IH, Ruegger P, Asadourian M, Moon B-H, Ayson A, Borneman J, McGovern DPB, Fornace AJ, Jr., Braun J, and Dubinsky M**. A Disease-Associated Microbial and Metabolomics State in Relatives of Pediatric Inflammatory Bowel Disease Patients. *Cellular and Molecular Gastroenterology and Hepatology*.
16. **Kanehisa M, and Goto S**. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* 28: 27-30, 2000.
17. **Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, Nielsen J, and Backhed F**. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498: 99-103, 2013.
18. **Kostic AD, Xavier RJ, and Gevers D**. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 146: 1489-1499, 2014.

19. **Kronman MP, Zaoutis TE, Haynes K, Feng R, and Coffin SE**. Antibiotic exposure and IBD development among children: a population-based cohort study. *Pediatrics* 130: e794-803, 2012.
20. **Kultima JR, Coelho LP, Forslund K, Huerta-Cepas J, Li SS, Driessen M, Voigt AY, Zeller G, Sunagawa S, and Bork P**. MOCAT2: a metagenomic assembly, annotation and profiling framework. *Bioinformatics (Oxford, England)* 2016.
21. **Kultima JR, Sunagawa S, Li J, Chen W, Chen H, Mende DR, Arumugam M, Pan Q, Liu B, Qin J, Wang J, and Bork P**. MOCAT: a metagenomics assembly and gene prediction toolkit. *PloS one* 7: e47656, 2012.
22. **Kunde S, Pham A, Bonczyk S, Crumb T, Duba M, Conrad H, Jr., Cloney D, and Kugathasan S**. Safety, tolerability, and clinical response after fecal transplantation in children and young adults with ulcerative colitis. *Journal of pediatric gastroenterology and nutrition* 56: 597-601, 2013.
23. **Li J, Butcher J, Mack D, and Stintzi A**. Functional Impacts of the Intestinal Microbiome in the Pathogenesis of Inflammatory Bowel Disease. *Inflammatory bowel diseases* 2014.
24. **Li SS, Zhu A, Benes V, Costea PI, Hercog R, Hildebrand F, Huerta-Cepas J, Nieuwdorp M, Salojarvi J, Voigt AY, Zeller G, Sunagawa S, de Vos WM, and Bork P**. Durable coexistence of donor and recipient strains after fecal microbiota transplantation. *Science (New York, NY)* 352: 586-589, 2016.
25. **Liu B, and Pop M**. ARDB--Antibiotic Resistance Genes Database. *Nucleic acids research* 37: D443-447, 2009.
26. **Machiels K, Joossens M, Sabino J, De Preter V, Arijs I, Eeckhaut V, Ballet V, Claes K, Van Immerseel F, Verbeke K, Ferrante M, Verhaegen J, Rutgeerts P, and Vermeire S**. A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. *Gut* 2013.
27. **Mende DR, Sunagawa S, Zeller G, and Bork P**. Accurate and universal delineation of prokaryotic species. *Nature methods* 10: 881-884, 2013.
28. **Michail S, Durbin M, Turner D, Griffiths AM, Mack DR, Hyams J, Leleiko N, Kenche H, Stolfi A, and Wine E**. Alterations in the gut microbiome of children with severe ulcerative colitis. *Inflammatory bowel diseases* 18: 1799-1808, 2012.
29. **Moayyedi P, Surette MG, Kim PT, Libertucci J, Wolfe M, Onischi C, Armstrong D, Marshall JK, Kassam Z, Reinisch W, and Lee CH**. Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology* 149: 102-109 e106, 2015.
30. **Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, and Pettersson S**. Host-gut microbiota metabolic interactions. *Science (New York, NY)* 336: 1262-1267, 2012.
31. **Petersen C, and Round JL**. Defining dysbiosis and its influence on host immunity and disease. *Cell Microbiol* 16: 1024-1033, 2014.
32. **Porcheron G, and Dozois CM**. Interplay between iron homeostasis and virulence: Fur and RyhB as major regulators of bacterial pathogenicity. *Veterinary microbiology* 179: 2-14, 2015.
33. **Porcheron G, Garenaux A, Proulx J, Sabri M, and Dozois CM**. Iron, copper, zinc, and manganese transport and regulation in pathogenic Enterobacteria: correlations between strains, site of infection and the relative importance of the different metal transport systems for virulence. *Frontiers in cellular and infection microbiology* 3: 90, 2013.
34. **Pryde SE, Duncan SH, Hold GL, Stewart CS, and Flint HJ**. The microbiology of butyrate formation in the human colon. *FEMS Microbiol Lett* 217: 133-139, 2002.
35. **Ringel Y, Maharshak N, Ringel-Kulka T, Wolber EA, Sartor RB, and Carroll IM**. High throughput sequencing reveals distinct microbial populations within the mucosal and luminal niches in healthy individuals. *Gut microbes* 6: 173-181, 2015.
36. **Rossen NG, Fuentes S, van der Spek MJ, Tijssen JG, Hartman JH, Duflou A, Lowenberg M, van den Brink GR, Mathus-Vliegen EM, de Vos WM, Zoetendal EG, D'Haens GR, and Ponsioen CY**. Findings From a Randomized Controlled Trial of Fecal Transplantation for Patients With Ulcerative Colitis. *Gastroenterology* 149: 110-118 e114, 2015.
37. **Scher JU, Ubeda C, Artacho A, Attur M, Isaac S, Reddy SM, Marmon S, Neimann A, Brusca S, Patel T, Manasson J, Pamer EG, Littman DR, and Abramson SB**. Decreased bacterial diversity characterizes the altered gut microbiota in patients with psoriatic arthritis, resembling dysbiosis in inflammatory bowel disease. *Arthritis & rheumatology (Hoboken, NJ)* 67: 128-139, 2015.
38. **Schmieder R, and Edwards R**. Insights into antibiotic resistance through metagenomic approaches. *Future microbiology* 7: 73-89, 2012.

39. **Schwiertz A, Jacobi M, Frick JS, Richter M, Rusch K, and Kohler H**. Microbiota in pediatric inflammatory bowel disease. *The Journal of pediatrics* 157: 240-244.e241, 2010.
40. **Sobhani I, Tap J, Roudot-Thoraval F, Roperch JP, Letulle S, Langella P, Corthier G, Tran Van Nhieu J, and Furet JP**. Microbial dysbiosis in colorectal cancer (CRC) patients. *PloS one* 6: e16393, 2011.
41. **Sommer F, and Backhed F**. The gut microbiota--masters of host development and physiology. *Nature reviews Microbiology* 11: 227-238, 2013.
42. **Soyucen E, Gulcan A, Aktuglu-Zeybek AC, Onal H, Kiykim E, and Aydin A**. Differences in the gut microbiota of healthy children and those with type 1 diabetes. *Pediatrics international : official journal of the Japan Pediatric Society* 56: 336-343, 2014.
43. **Sunagawa S, Mende DR, Zeller G, Izquierdo-Carrasco F, Berger SA, Kultima JR, Coelho LP, Arumugam M, Tap J, Nielsen HB, Rasmussen S, Brunak S, Pedersen O, Guarner F, de Vos WM, Wang J, Li J, Dore J, Ehrlich SD, Stamatakis A, and Bork P**. Metagenomic species profiling using universal phylogenetic marker genes. *Nature methods* 10: 1196-1199, 2013.
44. **Suskind DL, Brittnacher MJ, Wahbeh G, Shaffer ML, Hayden HS, Qin X, Singh N, Damman CJ, Hager KR, Nielson H, and Miller SI**. Fecal microbial transplant effect on clinical outcomes and fecal microbiome in active Crohn's disease. *Inflammatory bowel diseases* 21: 556-563, 2015.
45. **Turner D, Otley AR, Mack D, Hyams J, de Bruijne J, Uusoue K, Walters TD, Zachos M, Mamula P, Beaton DE, Steinhart AH, and Griffiths AM**. Development, validation, and evaluation of a pediatric ulcerative colitis activity index: a prospective multicenter study. *Gastroenterology* 133: 423-432, 2007.
46. **Vangay P, Ward T, Gerber JS, and Knights D**. Antibiotics, pediatric dysbiosis, and disease. *Cell host & microbe* 17: 553-564, 2015.
47. **Virta L, Auvinen A, Helenius H, Huovinen P, and Kolho KL**. Association of repeated exposure to antibiotics with the development of pediatric Crohn's disease--a nationwide, register-based finnish case-control study. *American journal of epidemiology* 175: 775-784, 2012.
48. **Yang T, Santisteban MM, Rodriguez V, Li E, Ahmari N, Carvajal JM, Zadeh M, Gong M, Qi Y, Zubcevic J, Sahay B, Pepine CJ, Raizada MK, and Mohamadzadeh M**. Gut dysbiosis is linked to hypertension. *Hypertension* 65: 1331-1340, 2015.
49. **Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, Amiot A, Böhm J, Brunetti F, Habermann N, Hercog R, Koch M, Luciani A, Mende DR, Schneider MA, Schrotz-King P, Tournigand C, Tran Van Nhieu J, Yamada T, Zimmermann J, Benes V, Kloor M, Ulrich CM, von Knebel Doeberitz M, Sobhani I, and Bork P**. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Molecular Systems Biology* 10: n/a-n/a, 2014.

Figure Captions

***Figure 1:* Shannon diversity index, species richness and evenness in comparison between groups (CTRL, UC, CD).** *Boxplots show all samples contrasted with MWU tests, showing significances as +: FDR < 0.1, *: FDR < 0.05. Scatterplots show each sib pair, revealing significantly higher diversity of each type in controls than in siblings with IBD.*

Figure 2:

A) Taxonomic distribution of bacterial taxa in relation to their relative abundance compared between the control cohort and both IBD populations (UC = Ulcerative colitis; CD = Crohn's disease). *Nodes represent taxa. Node color indicates direction of associations, with white nodes marking cases where an association from Appendix Table 1 could not be validated in the novel dataset (see bottom of figure for colour key). Coloured nodes with a narrow edge represent associations from the literature set (Appendix Table 1) which approached significance (MWU FDR < 0.1) when testing only hypotheses from this set. Coloured nodes with a bold edge represent such associations which also approached significance (MWU or KW FDR < 0.1) without restricting the hypothesis space. Coloured nodes with a bold, dashed outer edge represent IBD associations approaching significance (MWU or KW FDR < 0.1) when testing all possible associations, which are novel relative to the set of previous literature-derived findings. Detailed FDR scores and results from sibpair testing are given in Supplementary Table 1. Abbreviations: E. rectale = Eubacterium rectale, R. gnavus = Ruminococcus gnavus, F. prausnitzii = Faecalibacterium prausnitzii, C. ramosum = Clostridium ramosum, F. nucleatum = Fusobacterium nucleatum, E. coli = Escherichia coli, B. wadsworthia = Bilophila wadsworthia.*

B) Box plots of relative abundance of species in % (square root scale for visibility) for which differences in abundance approached significance (* indicates MWU FDR < 0.1) in this cohort, comparing the sets of UC or CD samples to the set of Controls. *Detailed FDR scores and results from sibpair testing are given in Supplementary Table 1. The observed species-level increase of C. ramosum in UC compared to controls is novel relative to the set of literature-derived findings.*

***Figure 3:* Correlation of species abundance with the abundance of virulence genes significantly increased in the microbiomes of UC patients.** *Dark colored = positive correlation of species abundance and the abundance of the corresponding virulence gene (SCC > 0.65, FDR < 0.1), red = increase of the species*

abundance in UC patients (MWU, FDR < 0.1), blue = no significant difference of abundance of species between UC and Control.

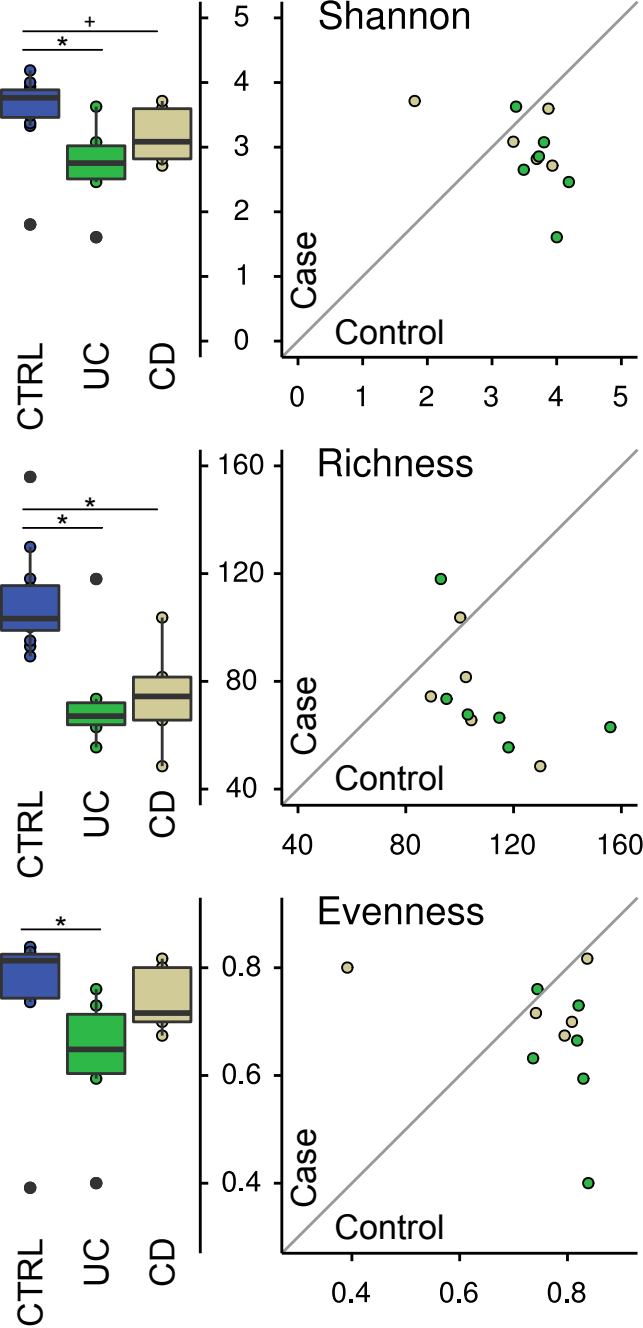
Figure 4: ***Mean lifetime number of antibiotic courses in comparison between Control, Ulcerative colitis (UC) and Crohn's disease (CD).*** *Bar charts are divided into segments representing the different antibiotic classes covered. For the most frequently used antibiotics, mean lifetime number of therapies are indicated.*

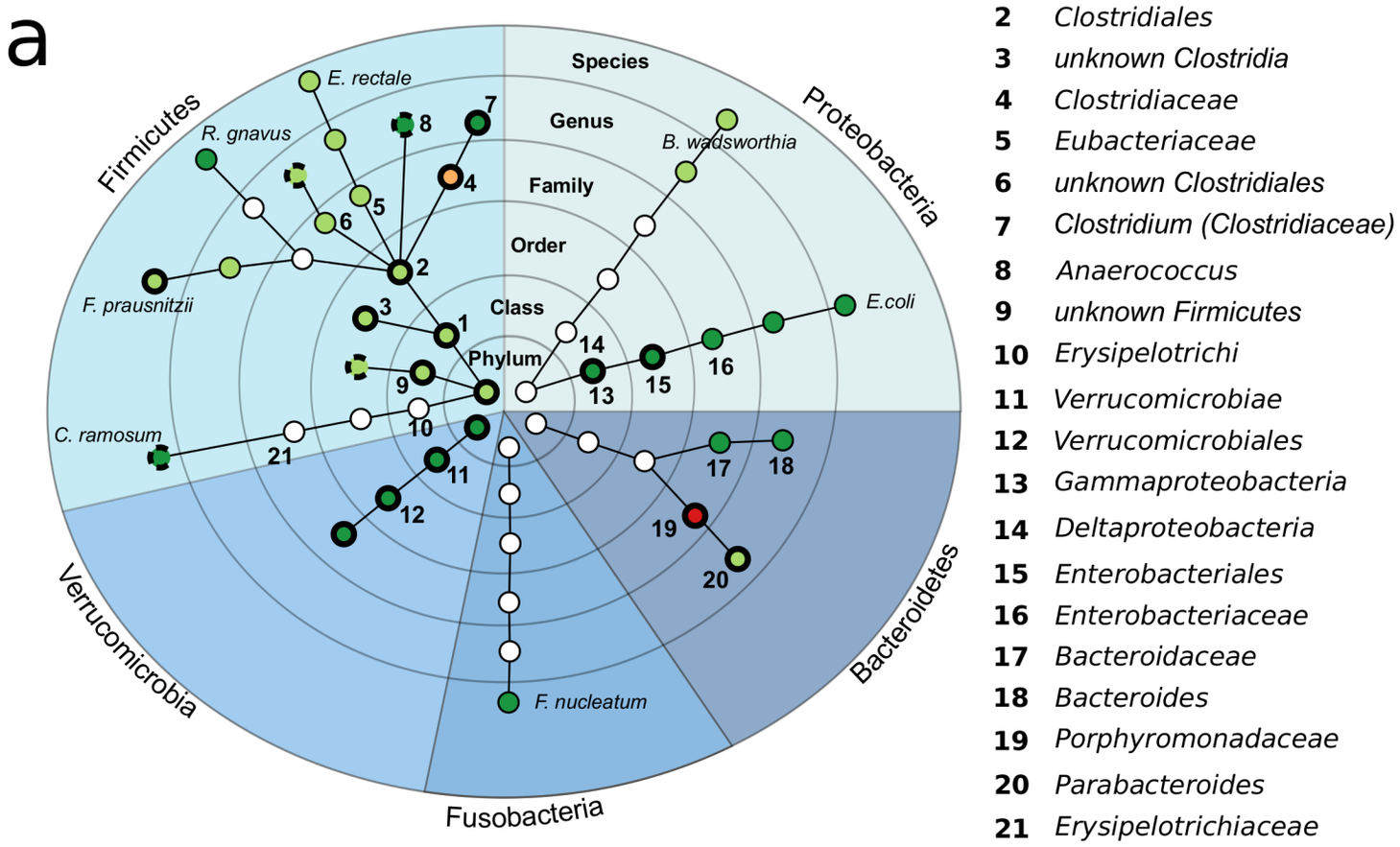
Figure 5: ***Fraction of potential carrier species:*** *For relative abundance of potentially resistant species, significant difference was observed for each antibiotic class between control and Ulcerative colitis (UC) (BY FDR = 0.062) and control and Crohn's disease (BY FDR = 0.019).*

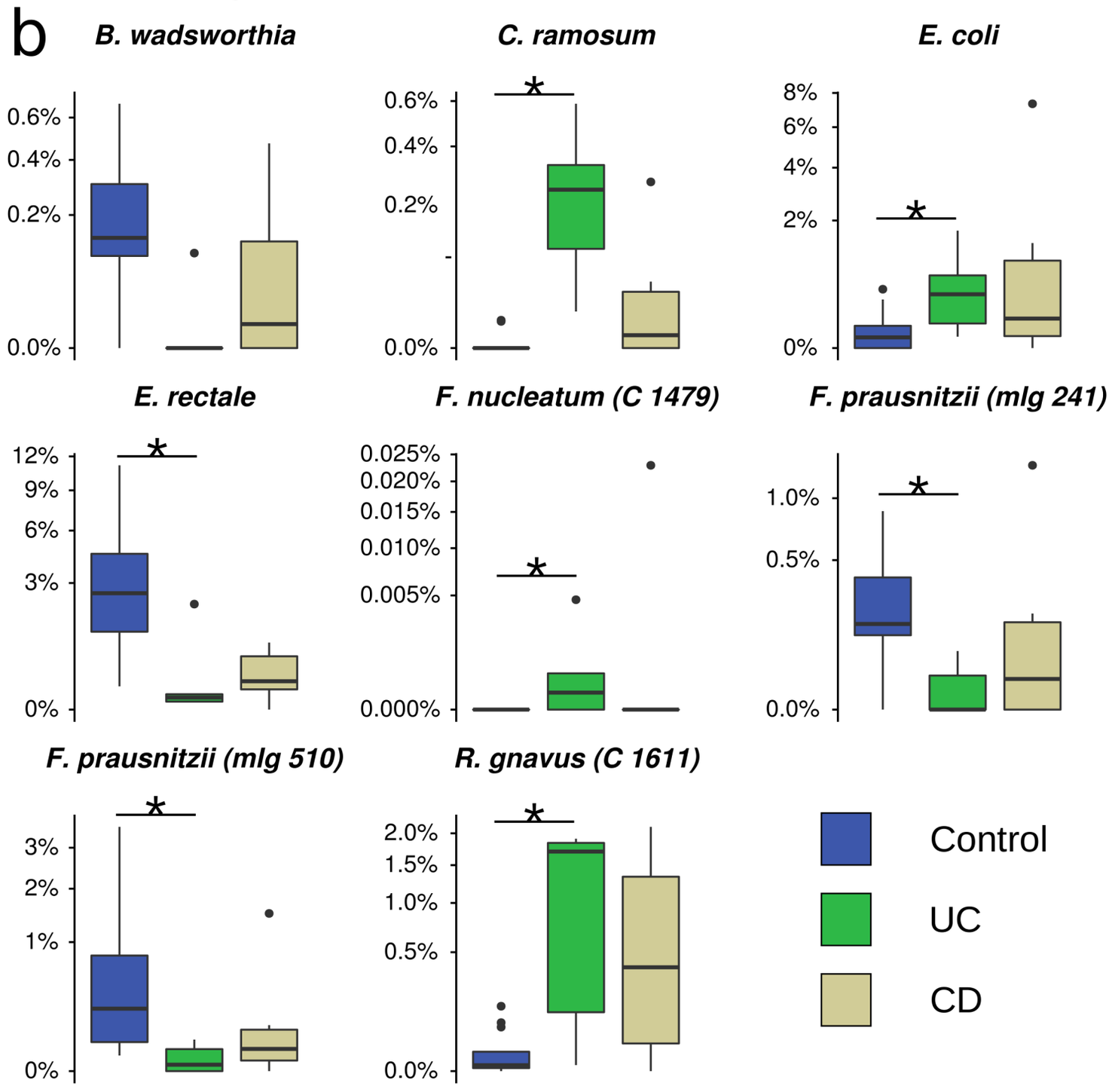
Table 1: **Cohort characteristics**

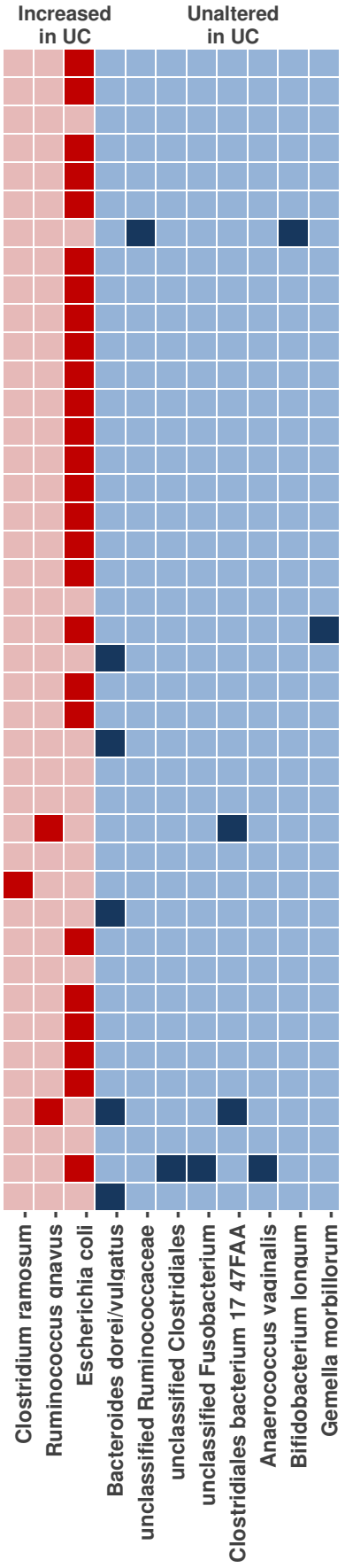
Characteristics	UC	CD	Control
n =	6	6	12
Demographics			
male: female	2:4	3:3	6:6
Age (years)			
Median ± SD	13 ± 2.7	14 ± 2.0	12.5 ± 3.6
Range	10 - 17	11 - 16	8 – 20
BMI (z-score)			
Median ± SD	-0.09 ± 0.73	-0.10 ± 1.09	0.26 ± 1.10
< -1 (%)	0	34	25
> +1 (%)	17	0	34
Age at onset (years)			
Median ± SD	9.5 ± 2.8	10 ± 1.8	
Range	4 - 12	9 - 14	
Disease duration (months)			
Median ± SD	22.5 ± 54.4	29 ± 31.6	
Range	13 - 144	0 - 82	
Disease activity			
inactive	3	1	
mild	1	2	
moderate-severe	2	3	
Medications (%)			
Steroids	67	50	
Anti-TNF	17	50	
Azathioprine	50	34	
Mesalazine	50	17	
TGF-β2	0	34	
UDCA	34	0	
Colchizine	17	0	

UC = Ulcerative colitis, CD = Crohn's disease; % = per group (UC, CD, Control); Disease activity based on PUCAI, PCDAI, respectively; Anti-TNF = infliximab/adalimumab; TGF-β2 = transforming growth factor beta 2 nutritional support formula; UDCA = ursodeoxycholic acid

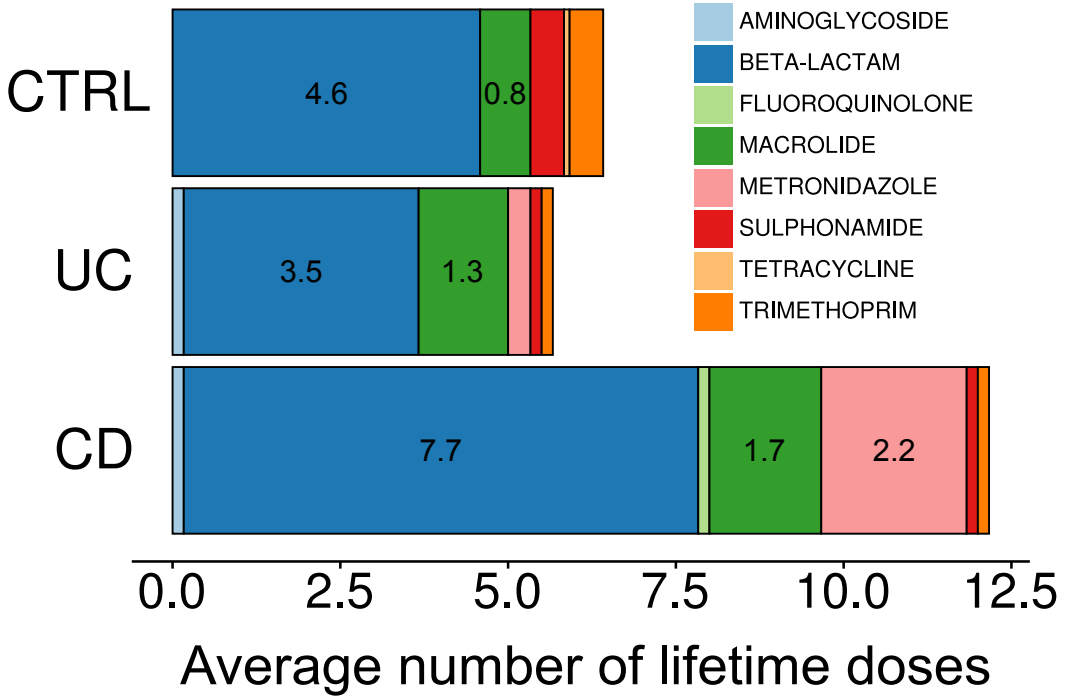

Shannon
Richness
Evenness
CTRL
UC
CD
Case
Control
+
*
0
1
2
3
4
5
40
80
120
160
0.4
0.6
0.8

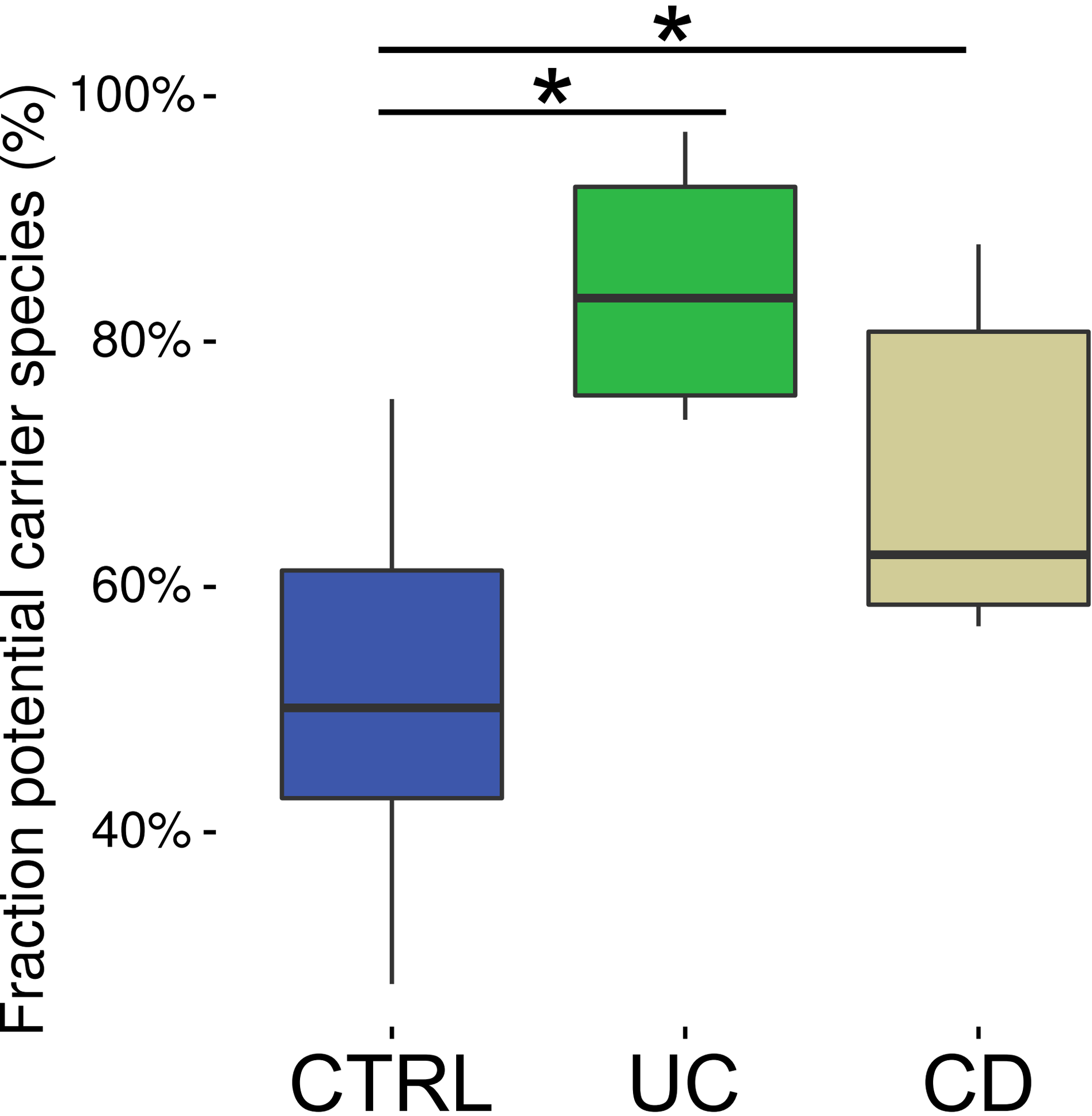

a
Species
Genus
Family
Order
Class
Phylum
Firmicutes
Proteobacteria
Verrucomicrobia
Bacteroidetes
Fusobacteria
E. rectale
R. gnavus
B. wadsworthia
F. prausnitzii
E.coli
C. ramosum
F. nucleatum
1 Clostridia
2 Clostridiales
3 unknown Clostridia
4 Clostridiaceae
5 Eubacteriaceae
6 unknown Clostridiales
7 Clostridium (Clostridiaceae)
8 Anaerococcus
9 unknown Firmicutes
10 Erysipelotrichi
11 Verrucomicrobiae
12 Verrucomicrobiales
13 Gammaproteobacteria
14 Deltaproteobacteria
15 Enterobacteriales
16 Enterobacteriaceae
17 Bacteroidaceae
18 Bacteroides
19 Porphyromonadaceae
20 Parabacteroides
21 Erysipelotrichiaceae
Literature set finding, validated
Literature set finding, not validated
Independent rediscovery
Novel association
Enriched in CD
Enriched in CD & UC
Enriched in UC
Depleted in UC


b
B. wadsworthia
C. ramosum
E. coli
E. rectale
F. nucleatum (C 1479)
F. prausnitzii (mlg 241)
F. prausnitzii (mlg 510)
R. gnavus (C 1611)
Control
UC
CD



Gene	Product	Category
gfcB	putative outer membrane lipoprotein	Capsule / Cell wall
ymcC	putative regulator	
adhD	zinc-type alcohol dehydrogenase AdhD	
yeeI	putative inner membrane protein	Colonisation
yhiJ	hypothetical protein	
yjeJ	hypothetical protein	
nrdF1	ribonucleotide-diphosphate reductase subunit β	DNA synthesis
eaeA	intimin EaeA	Fimbriae / Adherence
yagZ	hypothetical protein	
ydiV	hypothetical protein	
fimA	major type 1 subunit fimbrin (pilin)	
fimI	fimbrin-like protein fimI precursor	
csgF	curli assembly protein CsgF	
fimD	outer membrane usher protein fimD precursor	
fimE	tyrosine recombinase	
fimG	FimG protein precursor	
stcC	putative outer membrane usher protein	
fimC	fimbrial chaperone protein	
fimF	FimF protein precursor; putative fimbrial protein	
ndk	nucleoside diphosphate kinase	Intracellular survival, evading immune response
tonB	outer membrane receptor-mediated transport energizer	
tpx	thiol peroxidase	
slyA	transcriptional regulator SlyA	
soxS	DNA-binding transcriptional regulator	
ibeB	invasin of brain endothelial cells, IbeB	Invasion
aslA	arylsulfatase	
sitA	iron transport protein	Iron uptake, transport
iroN	outer membrane receptor FepA	
fur	ferric uptake regulator	
aroE	3-phosphoshikimate 1-carboxyvinyltransferase	Metabolic pathways
bioF1	8-amino-7-oxononanoate synthase	
solA	N-methyl-L-tryptophan oxidase	
gcvT	glycine cleavage system aminomethyltransferase T	
fliQ	flagellar biosynthesis protein FliQ	Motility
hslT	heat shock protein A	Resilience
celA	PTS system, cellobiose-specific IIB component	Transport
cyoB	cytochrome o ubiquinol oxidase subunit I	
pitB	partial putative transport protein	
fepA	outer membrane receptor FepA	
marT	putative transcriptional regulator (MarT)	Virulence
serB2	phosphoserine phosphatase	


CTRL
UC
CD
4.6
0.8
3.5
1.3
7.7
1.7
2.2
AMINOGLYCOSIDE
BETA-LACTAM
FLUOROQUINOLONE
MACROLIDE
METRONIDAZOLE
SULPHONAMIDE
TETRACYCLINE
TRIMETHOPRIM
0.0
2.5
5.0
7.5
10.0
12.5
Average number of lifetime doses


*
*
100%
80%
60%
40%
Fraction potential carrier species (%)
CTRL
UC
CD