

# Colonic Mucosa-associated Mycobiota in Individuals With Normal Colons

**Themistoklis Kourkoumpetis**

Baylor College of Medicine (BCM)

**Kristi L. Hoffman**

Department of Molecular Virology and Microbiology, BCM

**Yuna Kim**

Dornsife School of Public Health, Drexel University

**David Y. Graham**

Michael E. DeBakey Veterans Affairs Medical Center (MEDVAMC)

**Donna L. White**

Effectiveness and Safety, MEDVAMC and Dan L. Duncan Comprehensive Cancer Center, BCM

**R. Sue Day**

University of Texas School of Public Health

**Eric L. Brown**

University of Texas School of Public Health

**Gail Rosen**

College of Engineering, Drexel University

**Hashem B. El-Serag**

Section of Gastroenterology and Hepatology, Michael E. DeBakey Veterans Affairs Medical Center (MEDVAMC)

**Joseph F Petrosino**

Department of Molecular Virology and Microbiology, BCM

**Li Jiao** (✉ [ljiao@bcm.edu](mailto:ljiao@bcm.edu))

Baylor College of Medicine <https://orcid.org/0000-0002-0237-0262>

---

## Research Article

**Keywords:** fungi, microbiome, distribution, biogeography, ITS2, Malassezia

**Posted Date:** December 21st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1147234/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

To characterize the spatial variation of the mucosa-associated adherent mycobiota along the large intestine in individuals with a normal-colon, we performed eukaryotic rRNA operon's internal transcribed spacer-2 sequencing to profile fungal community composition and structure in 70 mucosal biopsies taken from the cecum, ascending, transverse, descending colon, and rectum of 14 polyp-free individuals. The bacteriome of these samples was previously characterized by sequencing the V4 region of the 16S rRNA gene. We identified 64 amplicon sequence variants (ASVs) with the relative abundance no less than 0.05% from these colonic mucosa samples. Each individual has a unique community composition of the gut mycobiome ( $P = 0.001$  for beta diversity). Alpha-diversity and beta-diversity did not differ significantly across the colon segments. The most common phyla (relative abundance) were Ascomycota (45.4%) and Basidiomycota (45.3%). The most common genera were *Malassezia* (28.2%) and *Candida* (13.4%). *Malassezia* was found in 13 of 14 individuals. Other fungi genera were sporadically found in the large intestine. The most common species were *Malassezia restricta* (22.7%), *Candida albicans* (11.9%), *Malasseziales sp.* (8.80%), unclassified fungi (7.80%), and *Penicillium paneum* (5.70%). *Malasseziaceae* was co-abundant with Enterobacteriaceae and co-exclusive with Barnesiellaceae, Rikenellaceae, and Acidaminococcaceae. *Malassezia* was widely colonized whereas other fungal genera were sporadically colonized in the large intestine. The physiologic and pathogenic functions of fungi in human gastrointestinal tract including *Malasseziaceae* that may interact with several bacterial families remain to be fully elucidated.

## Introduction

The colon harbors the most diverse and abundant population of microorganisms in the human body [1]. The community composition and structure of colonic intestinal microorganisms may reflect the differences in nutrient abundance, temperature, gastrointestinal motility, and oxygen content, which are known to change gradually from the cecum to the rectum [2].

The longitudinal distribution of the gut bacteriome along the colon has been investigated. Some studies have reported significant differences in bacterial diversity, density, and depth of invasion along the intestinal tract [3, 4]. Other studies have suggested the mucosa-associated bacteriome is homogeneous along the colon axis [5–7]. Fewer studies have investigated the distribution of other normal colonic cohabitants, such as fungi, viruses, and protozoa in the large intestine in humans.

Gut fungi – also known as the "mycobiome" – comprise a smaller proportion of the human intestinal microbiome ( $\sim 10^8$  organisms per gram of wet stool) compared to bacteria ( $\sim 10^{11}$  organisms per gram of wet stool)[8]. Shotgun metagenomic sequencing has shown that fungi account for  $\sim 0.1\%$  of the gut microbiome in feces [9]. Emerging evidence has implicated fungal dysbiosis in inflammatory bowel disease (IBD) [10], colorectal cancer (CRC) [11], cirrhosis [12], alcoholic liver disease [13], and pancreatic cancer [14]. Nevertheless, the feces-based studies by far suggested that there are no or only a small set of species that colonize the large intestine [15, 16].

A few studies have analyzed the mucosal mycobiota in humans with adenoma or IBD [10, 17–20]. The community composition and structure of the fungi along the large intestine in the normal colon have not been purposefully characterized in humans. We addressed this knowledge gap by leveraging next generation sequencing (NGS) technology using the colonic biopsies taken from adults with macroscopically normal colons. We hypothesized that the community composition and structure of the mucosa-associated mycobiota would vary along the colonic

longitudinal axis and among participants. In addition, because the 16S rRNA gene sequencing data were available to all the samples, we examined interactions between gut bacteria and fungi.

## Method

### Study participants and data collection

The study design, participant eligibility, and data collection have been reported previously [6]. Briefly, 402 participants were prospectively recruited and underwent routine colonoscopy at the Michael E. DeBakey Veterans Affairs Medical Center (MEDVAMC) in Houston, Texas, between July 2013 and April 2016.

All participants attended a regular educational session one to two weeks before the colonoscopy. During the session, a trained research coordinator obtained the informed consent from all participants and administered a questionnaire to collect information on lifestyle, social history, and medical history.

### Colonoscopy And Biopsy Acquisition

To ensure adequate endoscopic visualization of the colorectal mucosa, each participant took 3.78 liters of polyethylene glycol (Golytely®) bowel cleanser the night before their colonoscopy [21]. Patients were advised to stop taking aspirin, anti-inflammatory drugs, blood thinners, iron or vitamins with iron seven days before the procedure, and stop taking antidiabetic medications two days before the colonoscopy. The procedure was performed under conscious sedation and involved the visualization of the entire colon up to the cecum using standard Olympus® endoscopes.

Among 402 enrolled participants, a total of 120 were found to be polyp-free and have no other colonic pathological conditions. We collected the mucosal biopsies using cold-biopsy forceps introduced through the scope-operating-channel from multiple colonic segments (cecum, ascending colon, transverse colon, descending colon, and rectum) when feasible. Each biopsy sample was immediately placed in a sterile tube (RNase and DNase-free) on dry ice and transferred to a -80°C freezer within 15 minutes.

We sent multiple biopsies from the first enrolled 27 participants for gut microbiota profiling using the 16S rRNA gene amplicon sequencing. To perform the mycobiota profiling, we excluded nine participants with relatively lower sequencing count in the 16S rRNA gene survey because of the concern over poor DNA quality. We exclude four participants with a history of diabetes. Thus, our current mycobiota analyses were performed on 70 colonic biopsies taken from 14 participants. The study protocol was approved by the Institutional Review Board of Baylor College of Medicine (BCM) and Michael E. DeBakey VA Medical center (MEDVAMC) (protocol number: #H30941).

### Library Construction And Sequencing Analysis

#### DNA extraction and Internal Transcribed Spacer (ITS)-2 sequencing

The Alkek Center for Metagenomics and Microbiome Research (CMMR) at BCM performed microbial DNA extraction, library construction, and mycobiota sequencing using the methods previously described by Nash *et al*

[22]. In brief, DNA was extracted using the MO BIO PowerLyzer UltraClean Tissue & Cell Isolation Kit (Qiagen, Germantown, MD). The eukaryotic rRNA operon's second internal transcribed spacer 2 (ITS2) was amplified by PCR using modified versions of primers of ITS3 and ITS4 [23] and sequenced on the MiSeq platform (Illumina, San Diego, CA) using the 2x300bp paired-end protocol. The PCR primer modifications included the addition of a PrimerProspector [24] Illumina adapter and linker sequence. The reverse primer (ITS4) also included a unique 12-base pair Golay barcode for each sample [25]. A rarefaction curve was constructed using the sequence data for each sample to ensure we sampled the majority of its fungal diversity with the cutoff threshold of 70 per sample. The rarefaction resulted in excluding 8 out of the 70 samples from the analysis.

## Bioinformatics And Taxonomic Assignment

We used the CMMR pipeline incorporating phylogenetic and alignment-based approaches to maximize data resolution. This analysis provided summary statistics and quality control measurements for each sequencing run, as well as multi-run reports and data-merging capabilities for validating built-in controls and characterizing microbial communities across large numbers of samples.

The ITS2 read pairs were demultiplexed based on their unique molecular barcodes and merged using USEARCH (v7.0.1001) [26], requiring a minimum overlap of 50 base pairs. Merged reads were trimmed at the first base with quality score Q5 and filtered to remove those containing a > 0.5% expected error rate. The primers, adaptors, and linkers were removed. We used the DADA2 pipeline (version 1.8) to generate the amplicon sequence variant (ASV) table [27]. DADA2 implemented the naïve Bayesian classifier method for a taxonomic assignment using the UNITE database and NCBI databases (2018 version). The ASV table records the number of times each exact amplicon sequence variant was observed in each sample. ASVs can be resolved exactly down to the level of single-nucleotide differences over the sequenced gene region. Hence, it has a higher-resolution than operational taxonomic unit (OTU) tables, which roughly cluster sequences using sequence identity thresholds. A total of 64 ASVs were identified from our samples.

## 16S rRNA sequencing for gut microbiota profiling

We amplified and sequenced the 16S rRNA gene V4 region to determine the microbiota profile as described previously [6].

## Statistical analysis

We performed data analysis using the Agile Toolkit for Incisive Microbial Analyses (ATIMA), which is part of the CMMR pipeline. We used the Kruskal-Wallis test to test the differences in Shannon index for alpha diversity, as well as the difference in relative abundance of ASVs across the colonic segments. The difference in the beta diversity of the mycobiota by each individual or by the colon segment was visualized using the Principal Coordinate Analysis (PCoA) plots based on the weighted Bray-Curtis as the distance matrix and the difference was tested using PERMANOVA. *P*-values were adjusted for multiple comparisons with the Benjamini-Hochberg false discovery rate (FDR) algorithm [28]. FDR adjusted *P*-value < 0.05 indicated statistical significance.

## Fungal-bacterial Interaction

We used the SparCC [29] implementation in the R package SpiecEasi (version 1.1.0) [30] to analyze the interaction between fungi and bacteria. We evaluated the interaction between 46 families of bacteria and eight families of fungi based on the ASV data.

## Results

### Description of study population

Table 1 presents the basic demographic and selected lifestyle characteristics of 14 participants. Most of the participants were men and obese.

Table 1  
Selected characteristics of 14 study participants with normal colon

Characteristics	Mean (range, median)
Age (years)	59.8 (51.0-71.0, 60.4)
Body mass index (kg/m <sup>2</sup> )	30.4 (21.0-41.7, 29.3)
	<b>N (%)</b>
Sex	
Men	12 (85.7)
Women	2 (14.3)
Race/Ethnicity	
African-American	8 (57.1)
White/Caucasian	6 (42.9)
Body mass index (kg/m <sup>2</sup> )	
<30	8 (57.1)
≥30	6 (42.9)
Alcohol usage	
Never/rarely	5 (35.7)
Past use	2 (14.3)
Active use	7 (50.0)
Cigarette smoke	
Never	5 (35.7)
Past use	6 (42.9)
Active use	3 (21.4)
Hypertension, yes	6 (42.8)

## **Fungal Diversity: Inter-individually And Across The Colon Segments**

We identified a total of 64 unique ASVs in our samples. There was a significant difference in the beta-diversity of gut mycobiota among 14 participants (Fig. 1). However, there was no compositional dissimilarity of fungi across all colonic segments (Fig. 2). There was no significant difference in the Shannon index across all colon segments (Fig. 3).

## **The relative abundance and the prevalence of fungal taxa in the colon**

From the 64 fungal ASVs with relative abundance of no less than 0.05%, we identified ten phyla, 23 classes, 31 orders, 45 families, 53 genera, and 60 species.

Table 2 shows the most abundant fungal taxa and their relative abundance in all samples. Ascomycota and Basidiomycota were the most common phyla, with an overall mean relative abundance of 45.4% and 45.3%, respectively. These two phyla could be found in all the samples (Fig. 4A).

Table 2  
Mean relative abundance (%) of the top fungal taxa in colonic biopsies of 14 individuals

Phylum	Order	Family	Genus	Species
Ascomycota (45.4)	Saccharomycetales (22.1)	Saccharomycetales (13.4)	Candida (13.4)	Candida_albicans (11.9)
		Debaryomycetaceae (3.30)	Yamadazyma (2.20)	Yamadazma_mexicana (2.20)
	Capnodiales (6.95)	Cladosporiaceae (3.90)	UNC (3.52)	UNC (3.52)
		Mycosphaerellaceae (3.04)	Ramularia (3.04)	UNC (3.04)
		Eurotiales (7.04)	Aspergillaceae (6.84)	Penicillium (5.70)
Basidiomycota (45.3%)	Malasseziales (37.0)	Malasseziaceae (28.2)	Malassezia (28.2)	Malassezia_restricta (22.7)
		UNC_17133 (8.80)	UNC_14318 (8.80)	Malassezia_sp (8.80)
				UNC (4.52)
		Sporidiobolales (2.59)		
UNC fungi (7.70)	UNC (7.70)		UNC (7.68)	UNC (7.68)
UNC: unclassified.				

The most common genera were *Malassezia* (28.2%), *Candida* (13.4%), *Malasseziales\_unclassified* (8.8%), *Penicillium* (5.70%), *Ramularia* (3.04%), *Cladosporiaceae\_unclassified* (3.52%), and *Yamadazyma* (2.20%). *Malassezia* was detected in 92.8% (13 of 14) individuals. *Candida* and *Penicillium* were detected in less than five individuals. The *Saccharomyces* genus was only found in the cecum of one individual with a relative abundance of 0.18%. The fungi were generally present in the entire colon if they could be detected in an individual. Fig. 4B and Fig. 4C shows the top 10 genera and species respectively in all the samples.

The most common species were *Malassezia restricta* (22.7%), *Candida albicans* (11.9%), *Malasseziales sp.* (8.8%), *Penicillium paneum* (5.7%), *Yamadazyma mexicana* (2.20%), and unclassified fungi, *Cladosporium spp.*, and *Malasseziales* (Table 3). *S. cerevisiae* was not detected in our samples.

Table 3  
Mean relative abundance (%) of top 20 most abundant fungal species across colon segments

Species	Relative abundance (%)						
	Colon segment						
	Overall	Cecum	Ascending	Transverse	Descending	Rectum	<i>q</i> value
<i>Malassezia_restricta</i>	23.7	9.48	26.36	28.41	34.03	20.48	0.900
<i>Candida_albicans</i>	11.34	5.84	5.58	22.06	15.19	9.68	0.770
<i>Malasseziales_sp</i>	8.88	12.73	9.48	5.4	1.04	16.51	0.691
<i>Fungi (ASV09)</i>	5.41	3.38	7.27	2.7	7.01	6.35	1.000
<i>Penicillium_paneum</i>	5.27	9.09	8.57	0	6.75	0	0.816
<i>Malassezia (ASV13)</i>	4.73	9.74	0.78	10	0.13	3.81	0.816
<i>Cladosporiaceae (ASV16)</i>	3.56	8.31	0.39	0	3.77	4.92	0.691
<i>Yamadazyma_mexicana</i>	2.32	0	8.96	0	0	2.22	0.671
<i>Phaeococcomyces_sp</i>	1.96	0	9.09	0	0	0	0.653
<i>Alternaria (ASV14)</i>	1.82	0	0	9.05	0.13	1.11	0.653
<i>Cyberlindnera_jadinii</i>	1.79	6.1	2.21	0	0	0	0.653
<i>Fungi (ASV06)</i>	1.71	0	7.92	0	0	0	0.653
<i>Sporobolomyces_johnsonii</i>	1.6	5.58	0	0.48	0	1.75	0.653
<i>Candida_tropicalis</i>	1.57	1.69	5.58	0	0	0	0.653
<i>Ascomycota (ASV01)</i>	1.46	1.56	0	0	5.19	0	0.691
<i>Ramularia (ASV35)</i>	1.46	6.75	0	0	0	0	0.653
<i>Trichosporon (ASV34)</i>	1.43	0	0	0	0	8.1	0.653
<i>Bolbitius_demangei</i>	1.12	0	0	6.35	0	0	0.653
<i>Ramularia (ASV10)</i>	1.01	0	0	0	4.68	0	0.653
<i>Saccharomycetales (ASV38)</i>	1.01	0	0	5.71	0	0	0.653

## The relative abundance of major fungal phyla and genera across the colon segments

Table 3 shows no statistically significant difference in the relative abundance of fungal species across five colon segments. However, some observations were worth noting. The relative abundance of *Malassezia restricta* in the

cecum was much lower than in other segments. The relative abundance of *Candida albicans* was also lower in the cecum and ascending colon than in other segments. The variability in relative abundance of *Penicillium paneum* across the colon segments was also noted.

## Fungal-bacterial interaction

SparCC network plot of correlations between bacterial families and fungal families is shown in Fig. 5. Only significant correlations (two-sided  $P$  value  $< 0.001$  based on bootstrapping of 1000 repetitions) with an absolute correlation coefficient  $\geq 0.3$  are presented. We found that Malasseziaceae was the only fungal family significantly correlated with bacteria. It was inversely correlated with Barnesiaceae (correlation coefficient  $r = -0.48$ ), Rikenellaceae ( $r = -0.49$ ), Acidaminococcaceae ( $r = -0.49$ ), and positively correlated with Enterobacteriaceae ( $r = 0.42$ ), respectively.

## Discussion

We surveyed the mucosa-associated mycobiota across the longitudinal axis of normal-appearing adult human colon using the ITS2 amplicon sequencing. Ascomycota and Basidiomycota were the two major phyla. *Malassezia* and *Candida* were the two major genera. *M. restricta*, *C. albicans*, *P. paneum*, and *M. spp* were the major species. We found that almost all individuals were colonized with *Malassezia*, with sporadic colonization with other fungal genera. Notably, although we observed significant inter-individual variability in fungal composition across individuals, the distribution of fungi longitudinally along the colon was not statistically heterogeneous within each individual. The prevalence of fungi with other co-occurring bacteria was observed only for Malasseziaceae and a few bacterial families including Barnesiaceae, Rikenellaceae, Acidaminococcaceae, and Enterobacteriaceae.

Ascomycota, Basidiomycota, and unclassified fungi accounted for 97% of the fungi phyla. A previous study using the adjacent normal colonic mucosa of 27 patients with colorectal adenoma found Ascomycota was the dominant phylum (80.5%), followed by Glomeromycota (3.1%) and Basidiomycota (2.5%) [18]. The other study using the mucosal sample of 14 patients with UC and 14 healthy controls found Ascomycota and Basidiomycota were the two dominant phyla [19]. In these two deep sequencing studies conducted in China, the relative abundance of Ascomycota was much higher than Basidiomycota. However, in our study, these two common phyla had equal abundance as detected in the normal mucosa of polyp-free individuals.

*Malassezia*, which belongs to Basidiomycota, was the most common genus with *M. restricta* being the most common species in our study. *Malassezia* is abundant in human breast milk and colonizes the gut during the neonatal period [31] and can survive in the gut [32]. *Malassezia* has previously been recognized as the dominant fungal genus in the oral mycobiome [33]. *Malassezia* is also a skin commensal fungus frequently associated with mild skin infections such as tinea versicolor, and occasionally results in fungemia in immunosuppressed individuals [34, 35]. Using the fecal sample, *Malassezia* was identified as an opportunistic fungus in patients with colorectal tumor [36]. *Malassezia restricta* was shown to be a key player in the pathogenesis of Crohn's disease in a mouse model because this fungus can elicit innate inflammatory responses through caspase recruitment domain family member 9 (CARD9) [37]. While our study identified that *Malassezia* was the most abundant fungal genus, another study reported that *Candida* was the most abundant genus using 54 mucosal samples taken from cecum or rectum of 28 healthy individuals, followed by *Cyberlindnera*, *Fusarium*, *Galactomyces*, and *Malassezia*. The differences in these two studies included geography (California vs. Texas), race/ethnicity (Asian vs. African-

American predominantly), and the taxonomic assignment tool (ASV vs. OTU-based) used [37]. Our finding suggested that the role of *Malassezia* in colonic physiology and pathology deserves more investigations.

*Candida* was the second most abundant genus in our study with a mean relative abundance of 11.8%, which is similar to the abundance reported by Nash using 317 fecal samples in the Human Microbiome Project [22]. Fechney et al reported that *Candida* is the most abundant oral colonizer in 17 children with or without dental caries [33]. *C. albicans*, and *C. tropicalis* were the major *Candida spp* in our study. *C. albicans* is not only a commensal in healthy individuals but also a pathogen of the gastrointestinal (GI) tract, at mucosal surfaces, and the blood under opportunistic conditions including disease- and drug-related immunosuppression or injury. *Candida* has recently been implicated in the pathogenesis of IBD [20], *Clostridium difficile* colitis [38], and graft-versus-host disease [39]. *C. tropicalis* is a common opportunistic infectious agent, especially in patients hospitalized in intensive care units [40]. *C. tropicalis* has also been associated with Familial Crohn's disease [41]. In our study, *Candida* was not present in the colon of all participants. However, its occurrence in 43% of our colonically normal participants coupled with its known potential pathogenic significance suggested its presence and role should be further investigated.

*Penicillium*, commonly found in soil and belongs to Ascomycetes phylum and Aspergillaceae family [42], also colonizes the colon. *P. paneum* was the only species identified in our samples. *P. paneum*, isolated from the baled grass silage, has been shown to produce mycotoxins such as Roquefortine C, marcfortine A, and andrastin A and diverse secondary metabolites [43]. However, the implication of *P. paneum* in human health is essentially unknown.

*Yamadazyma mexicana*, unclassified fungi, and unclassified members of *Cladosporium* were also detected in the colonic mucosa. *Yamadazyma mexicana* is one of the six yeast strains that can degrade environmental hydrocarbons [44] and its role in the gut is completely unknown. *Cladosporium spp* is a ubiquitous fungus, which can cause skin infections and has been linked to allergic rhinitis, with its cell wall being classified as a potential allergen [45]. Among the 20 most abundant species, 10 were unclassified using our sequencing and bioinformatics pipeline. Additional research will be necessary to characterize these unclassified fungi in order to understand their potential function and significance in colon health and disease pathogenesis.

The fecal mycobiome has been examined using culture-based and non-culture-based methods. The following core fungal genera were proposed: *Candida* (especially *C. albicans*), *Saccharomyces*, *Penicillium*, *Cryptococcus*, *Malassezia* (especially *M. restricta*), *Cladosporium*, *Galactomyces*, *Aspergillus*, *Debaryomyces*, and *Trichosporon* [46]. However, Nash et al were unable to find the latter four genera using NGS in 317 stool samples [22]. Similarly, we did not identify *Cryptococcus* and *Galactomyces* in our mucosal samples. *Saccharomyces* was only detected in the cecum of one individual with a relative abundance of 0.18%. *Saccharomyces cerevisiae* was not detected in our samples. Because *Saccharomyces* is a yeast that is most abundant in foods, our observation supported the notion that foodborne *Saccharomyces cerevisiae* is likely not a mucosa-associated fungus and is only transiently present in the human GI tract [47]. A *Saccharomyces*-free diet has been shown to eliminate its presence in the stool [15]. In our study, participants had a full colonic preparation as well as fasted for 12 hours before the procedure. The absence of food-related fungi may be the indicator of proper colonic preparation. On the other hand, the colonization of fungi in the colon mucosa should be further investigated. Overall, the structure of the fungi in the colonic mucosa differed from that of feces shown by other studies.

Our findings did not concur with a previous research using the mucosal sample. Li et al used the 18S rRNA region to survey the mycobiome in normal terminal ileum tissues (non-colonic) samples of seven healthy volunteers [20]. They found *Saccharomyces cerevisiae*, *Saccharomyces castellii*, *Candida albicans*, *Candida tropicalis*, *Gibberella*

*moniliformis*, and *Sclerotinia sclerotiorum* to be the most abundant species. The difference between our study and Li's study included anatomic locations surveyed, target regions sequenced, and the genome assemblies for fungal identification. ITS2 has improved resolution of the mycobiome membership compared to metagenomics and 18S rRNA gene sequencing [48]. More research is needed to characterize the fungi entity in humans using the standardized approach.

In an earlier study, we found the bacterial distribution was largely homogeneous along the colon axis [6]. In the present study, we found that the distribution of the common fungal genus, *Malassezia*, was homogeneous along the colon. Most of other fungal species were only sporadically detected in the colon. Whether the rare or sporadic microorganism has any role to play in patchy colonic lesions should be elucidated further.

We found the bacterial-bacterial interaction in colonic mucosa of these normal individuals detected by SparCC. However, because fungi were sporadic, we only observed the interaction between the most common fungi, Malasseziaceae family, with Acidaminococcaceae and Barnesiellaceae, and possibly with Enterobacteriaceae. We also observed co-abundance of Aspergillaceae and unclassified member of Lactobacillales. Taken together, our data suggested that in the normal colon, there was no extensive interaction between fungi and bacteria because of the rarity of fungi.

Our study had several strengths. It was among the first to characterize fungi in the normal colon in humans. Although we provided a snapshot of the fungal profile, the survey of the mucosal mycobiota likely reflects the long-term colonic indigenous mycobiota that are less likely to be disturbed by bowel preparation. In addition, we used the identical procedure to collect biopsies where the contamination was minimized in a clinical setting. Lastly, we used the ASVs for taxonomic assignment, which is thought to be more precise than the OTU-based assignment [49].

Our study also had several limitations. First, the ITS2 ribosomal domain for fungal identification is different from that of ITS1. The taxonomic characteristics may be biased towards Ascomycota and Basidiomycota [50]. Second, the generalizability of this study was limited as it included predominantly overweight and obese middle-aged men using the VA healthcare system. Third, our sample size was limited in testing the distribution of less abundant fungi across the colon segment. Lastly, the research provided a snap-shot of the fungal distribution in the gut after bowel preparation. Additional study is needed to describe the trajectory of fungal colonization over time and with aging.

In summary, using the ITS2 targeted-amplicon sequencing and the present fungal genome assemblies, we found fungi that are commonly present in indoor and outdoor environments also inhabit the lower GI tract in humans. Our study findings argue against the absence of fungi in "healthy" adult colons. Nevertheless, *Malassezia* was the only prevalent fungi and others were sporadic. The potential co-occurring and co-exclusive correlation between Malasseziaceae and bacteria and its physiological implication deserve further research. The characterization of the fungal origin, richness, composition, and structure in the human colon is essential for informing future research examining the role of fungi in both maintenances of colonic health and the pathogenesis of colonic diseases. Appropriate use of antifungal therapy will be dependent on understanding the role of fungi in health and disease.

## Declarations

## FUNDING

This study was supported by the Cancer Prevention and Research Institute of Texas (RP#140767, Principal Investigators [PIs]: Jiao, Li and Petrosino, Joseph), Gillson Longenbaugh Foundation and Golfers Against Cancer organization (PI: Jiao, Li), and Houston Veterans Affairs Health Services Research Center of Innovations for Innovations in Quality, Effectiveness, and Safety (CIN13-413). The Texas Medical Center Digestive Disease Center (P30 DK56338, PI: Hashem B. El-Serag). Drs. White and Jiao receive research salary support from the U.S. Department of Veterans Affairs (CX001430). Dr. Kourkoumpetis was supported by the T32-National Institutes of Health Training grant (PI: Dr. Hashem B. El-Serag). The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government. The funders played no role in the design or conduct of the study or interpretation and presentation of study findings.

## **Acknowledgment**

We thank Drs. David Ramsey, PhD, Kathryn Royse, PhD, Liang Chen, MS, and Nadim Ajami, PhD at Baylor College of Medicine for assistance with data and project management. We also thank Ashely Johnson, BS; Sarah Plew, BS, Preksha Shah, MPH, Ava Smith, PA, Mahmoud Al-Saadi, MD, and Jocelyn Uriostegui, PA for collecting and processing biopsy samples and conducting patient interviews. We wish to thank the veterans who generously participated in this study and the nurses and doctors at the endoscopy unit of the Michael E. DeBakey VA Medical Center for their contributions.

## **Conflict of interest**

The authors have no relevant financial or non-financial interests to disclose.

## **Ethics approval**

The study protocol (#H30941) was approved by the Institutional Review Board of Baylor College of Medicine and Michael E. DeBakey VA Medical center.

## **Consent to participate**

A trained research coordinator obtained informed consent from each participant.

## **Consent for publication**

Participants provided informed consent to publish the research data in aggregate.

## **Availability of data and material**

Available upon request.

## **Authors' contributions**

Themistoklis Kourkoumpetis drafted the manuscript. Kristi Hoffman managed the project and performed the bioinformatics analysis. Yuna Kim and Gail Rosen performed the bioinformatics analysis. David Y. Graham and Hashem B. El-Serag managed the clinic and edited the manuscript. Donna L. White participated in patient enrollment and edited the manuscript. R. Sue Day and Eric L. Brown edited the manuscript. All authors approved the content of the manuscript. Li Jiao was responsible for the design, data analysis, and writing. Joseph F. Petrosino was responsible for microbiota profiling and bioinformatics.

## References

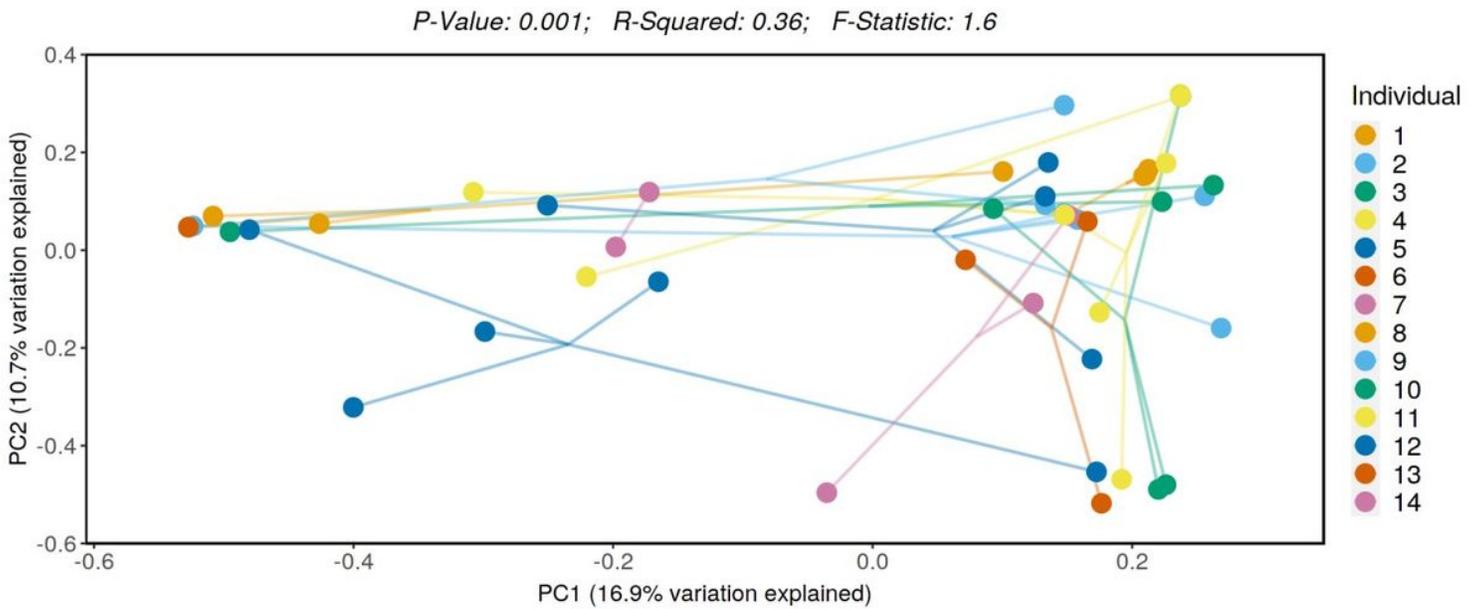
1. Lozupone CA (2016) Getting to know the microbiome. *Nat Microbiol* 1:16030. doi:10.1038/nmicrobiol.2016.30
2. Blau S, Rubinstein A, Bass P, Singaram C, Kohen R (1999) Differences in the reducing power along the rat GI tract: lower antioxidant capacity of the colon. *Mol Cell Biochem* 194 (1-2):185-191
3. Zhang Z, Geng J, Tang X, Fan H, Xu J, Wen X, Ma ZS, Shi P (2014) Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. *The ISME journal* 8 (4):881-893. doi:10.1038/ismej.2013.185
4. Nava GM, Carbonero F, Croix JA, Greenberg E, Gaskins HR (2012) Abundance and diversity of mucosa-associated hydrogenotrophic microbes in the healthy human colon. *The ISME journal* 6 (1):57-70. doi:10.1038/ismej.2011.90
5. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA (2005) Diversity of the human intestinal microbial flora. *Science* 308 (5728):1635-1638. doi:10.1126/science.1110591
6. Jiao L, Kourkoumpetis T, Hutchinson D, Ajami NJ, Hoffman K, White DL, Graham DY, Hair C, Shah R, Kanwal F, Jarbrink-Sehgal M, Husain N, Hernaez R, Hou J, Cole R, Velez M, Ketwaroo G, Kramer J, El-Serag HB, Petrosino JF (2021) Spatial Characteristics of Colonic Mucosa-Associated Gut Microbiota in Humans. *Microbial ecology*. doi:10.1007/s00248-021-01789-6
7. Stearns JC, Lynch MD, Senadheera DB, Tenenbaum HC, Goldberg MB, Cvitkovitch DG, Croitoru K, Moreno-Hagelsieb G, Neufeld JD (2011) Bacterial biogeography of the human digestive tract. *Scientific reports* 1:170. doi:10.1038/srep00170
8. Huseyin CE, O'Toole PW, Cotter PD, Scanlan PD (2017) Forgotten fungi-the gut mycobiome in human health and disease. *FEMS microbiology reviews* 41 (4):479-511. doi:10.1093/femsre/fuw047
9. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Meta HITC, Bork P, Ehrlich SD, Wang J (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464 (7285):59-65. doi:10.1038/nature08821
10. El Mouzan MI, Korolev KS, Al Mofarreh MA, Menon R, Winter HS, Al Sarkhy AA, Dowd SE, Al Barrag AM, Assiri AA (2018) Fungal dysbiosis predicts the diagnosis of pediatric Crohn's disease. *World journal of gastroenterology* 24 (39):4510-4516. doi:10.3748/wjg.v24.i39.4510
11. Coker OO, Nakatsu G, Dai RZ, Wu WKK, Wong SH, Ng SC, Chan FKL, Sung JJY, Yu J (2019) Enteric fungal microbiota dysbiosis and ecological alterations in colorectal cancer. *Gut* 68 (4):654-662. doi:10.1136/gutjnl-2018-317178

12. Bajaj JS, Liu EJ, Kheradman R, Fagan A, Heuman DM, White M, Gavis EA, Hylemon P, Sikaroodi M, Gillevet PM (2018) Fungal dysbiosis in cirrhosis. *Gut* 67 (6):1146-1154. doi:10.1136/gutjnl-2016-313170
13. Yang AM, Inamine T, Hochrath K, Chen P, Wang L, Llorente C, Bluemel S, Hartmann P, Xu J, Koyama Y, Kisseleva T, Torralba MG, Moncera K, Beeri K, Chen CS, Freese K, Hellerbrand C, Lee SM, Hoffman HM, Mehal WZ, Garcia-Tsao G, Mutlu EA, Keshavarzian A, Brown GD, Ho SB, Bataller R, Starkel P, Fouts DE, Schnabl B (2017) Intestinal fungi contribute to development of alcoholic liver disease. *J Clin Invest* 127 (7):2829-2841. doi:10.1172/JCI90562
14. Kazmierczak-Siedlecka K, Dvorak A, Folwarski M, Daca A, Przewlocka K, Makarewicz W (2020) Fungal Gut Microbiota Dysbiosis and Its Role in Colorectal, Oral, and Pancreatic Carcinogenesis. *Cancers* 12 (5). doi:10.3390/cancers12051326
15. Auchtung TA, Fofanova TY, Stewart CJ, Nash AK, Wong MC, Gesell JR, Auchtung JM, Ajami NJ, Petrosino JF (2018) Investigating Colonization of the Healthy Adult Gastrointestinal Tract by Fungi. *mSphere* 3 (2). doi:10.1128/mSphere.00092-18
16. Suhr MJ, Banjara N, Hallen-Adams HE (2016) Sequence-based methods for detecting and evaluating the human gut mycobiome. *Letters in applied microbiology* 62 (3):209-215. doi:10.1111/lam.12539
17. Liguori G, Lamas B, Richard ML, Brandi G, da Costa G, Hoffmann TW, Di Simone MP, Calabrese C, Poggioli G, Langella P, Campieri M, Sokol H (2016) Fungal Dysbiosis in Mucosa-associated Microbiota of Crohn's Disease Patients. *J Crohns Colitis* 10 (3):296-305. doi:10.1093/ecco-jcc/jjv209
18. Luan C, Xie L, Yang X, Miao H, Lv N, Zhang R, Xiao X, Hu Y, Liu Y, Wu N, Zhu Y, Zhu B (2015) Dysbiosis of fungal microbiota in the intestinal mucosa of patients with colorectal adenomas. *Scientific reports* 5:7980. doi:10.1038/srep07980
19. Qiu X, Ma J, Jiao C, Mao X, Zhao X, Lu M, Wang K, Zhang H (2017) Alterations in the mucosa-associated fungal microbiota in patients with ulcerative colitis. *Oncotarget* 8 (64):107577-107588. doi:10.18632/oncotarget.22534
20. Li Q, Wang C, Tang C, He Q, Li N, Li J (2014) Dysbiosis of gut fungal microbiota is associated with mucosal inflammation in Crohn's disease. *J Clin Gastroenterol* 48 (6):513-523. doi:10.1097/MCG.0000000000000035
21. Committee ASoP, Saltzman JR, Cash BD, Pasha SF, Early DS, Muthusamy VR, Khashab MA, Chathadi KV, Fanelli RD, Chandrasekhara V, Lightdale JR, Fonkalsrud L, Shergill AK, Hwang JH, Decker GA, Jue TL, Sharaf R, Fisher DA, Evans JA, Foley K, Shaukat A, Eloubeidi MA, Faulx AL, Wang A, Acosta RD (2015) Bowel preparation before colonoscopy. *Gastrointest Endosc* 81 (4):781-794. doi:10.1016/j.gie.2014.09.048
22. Nash AK, Auchtung TA, Wong MC, Smith DP, Gesell JR, Ross MC, Stewart CJ, Metcalf GA, Muzny DM, Gibbs RA, Ajami NJ, Petrosino JF (2017) The gut mycobiome of the Human Microbiome Project healthy cohort. *Microbiome* 5 (1):153. doi:10.1186/s40168-017-0373-4
23. White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications* 18 (1):315-322
24. Walters WA, Caporaso JG, Lauber CL, Berg-Lyons D, Fierer N, Knight R (2011) PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics* 27 (8):1159-1161. doi:10.1093/bioinformatics/btr087
25. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME journal* 6 (8):1621-1624. doi:10.1038/ismej.2012.8

26. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26 (19):2460-2461
27. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods* 13 (7):581-583. doi:10.1038/nmeth.3869
28. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. *Behavioural brain research* 125 (1-2):279-284. doi:10.1016/s0166-4328(01)00297-2
29. Friedman J, Alm EJ (2012) Inferring correlation networks from genomic survey data. *PLoS computational biology* 8 (9):e1002687. doi:10.1371/journal.pcbi.1002687
30. Kurtz ZD, Muller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA (2015) Sparse and compositionally robust inference of microbial ecological networks. *PLoS computational biology* 11 (5):e1004226. doi:10.1371/journal.pcbi.1004226
31. Boix-Amoros A, Puente-Sanchez F, du Toit E, Linderborg KM, Zhang Y, Yang B, Salminen S, Isolauri E, Tamames J, Mira A, Collado MC (2019) Mycobiome profiles in breast milk from healthy women depend on mode of delivery, geographic location and interaction with bacteria. *Applied and environmental microbiology*. doi:10.1128/AEM.02994-18
32. Spatz M, Richard ML (2020) Overview of the Potential Role of *Malassezia* in Gut Health and Disease. *Frontiers in cellular and infection microbiology* 10:201. doi:10.3389/fcimb.2020.00201
33. Fechney JM, Browne GV, Prabhu N, Irinyi L, Meyer W, Hughes T, Bockmann M, Townsend G, Salehi H, Adler CJ (2019) Preliminary study of the oral mycobiome of children with and without dental caries. *J Oral Microbiol* 11 (1):1536182. doi:10.1080/20002297.2018.1536182
34. Crespo-Erchiga V, Florencio VD (2006) *Malassezia* yeasts and pityriasis versicolor. *Curr Opin Infect Dis* 19 (2):139-147. doi:10.1097/01.qco.0000216624.21069.61
35. Wheeler ML, Limon JJ, Underhill DM (2017) Immunity to Commensal Fungi: Detente and Disease. *Annu Rev Pathol* 12:359-385. doi:10.1146/annurev-pathol-052016-100342
36. Gao R, Kong C, Li H, Huang L, Qu X, Qin N, Qin H (2017) Dysbiosis signature of mycobiota in colon polyp and colorectal cancer. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 36 (12):2457-2468. doi:10.1007/s10096-017-3085-6
37. Limon JJ, Tang J, Li D, Wolf AJ, Michelsen KS, Funari V, Gargus M, Nguyen C, Sharma P, Maymi VI, Iliev ID, Skalski JH, Brown J, Landers C, Borneman J, Braun J, Targan SR, McGovern DPB, Underhill DM (2019) *Malassezia* Is Associated with Crohn's Disease and Exacerbates Colitis in Mouse Models. *Cell host & microbe* 25 (3):377-388 e376. doi:10.1016/j.chom.2019.01.007
38. Panpetch W, Somboonna N, Palasuk M, Hiengrach P, Finkelman M, Tumwasorn S, Leelahavanichkul A (2019) Oral *Candida* administration in a *Clostridium difficile* mouse model worsens disease severity but is attenuated by *Bifidobacterium*. *PloS one* 14 (1):e0210798. doi:10.1371/journal.pone.0210798
39. Kumari R, Palaniyandi S, Hildebrandt GC (2019) Microbiome: An Emerging New Frontier in Graft-Versus-Host Disease. *Dig Dis Sci* 64 (3):669-677. doi:10.1007/s10620-018-5369-9
40. Xie J, Du H, Guan G, Tong Y, Kourkoumpetis TK, Zhang L, Bai FY, Huang G (2012) N-acetylglucosamine induces white-to-opaque switching and mating in *Candida tropicalis*, providing new insights into adaptation and fungal sexual evolution. *Eukaryot Cell* 11 (6):773-782. doi:10.1128/EC.00047-12
41. Hoarau G, Mukherjee PK, Gower-Rousseau C, Hager C, Chandra J, Retuerto MA, Neut C, Vermeire S, Clemente J, Colombel JF, Fujioka H, Poulain D, Sendid B, Ghannoum MA (2016) Bacteriome and Mycobiome Interactions

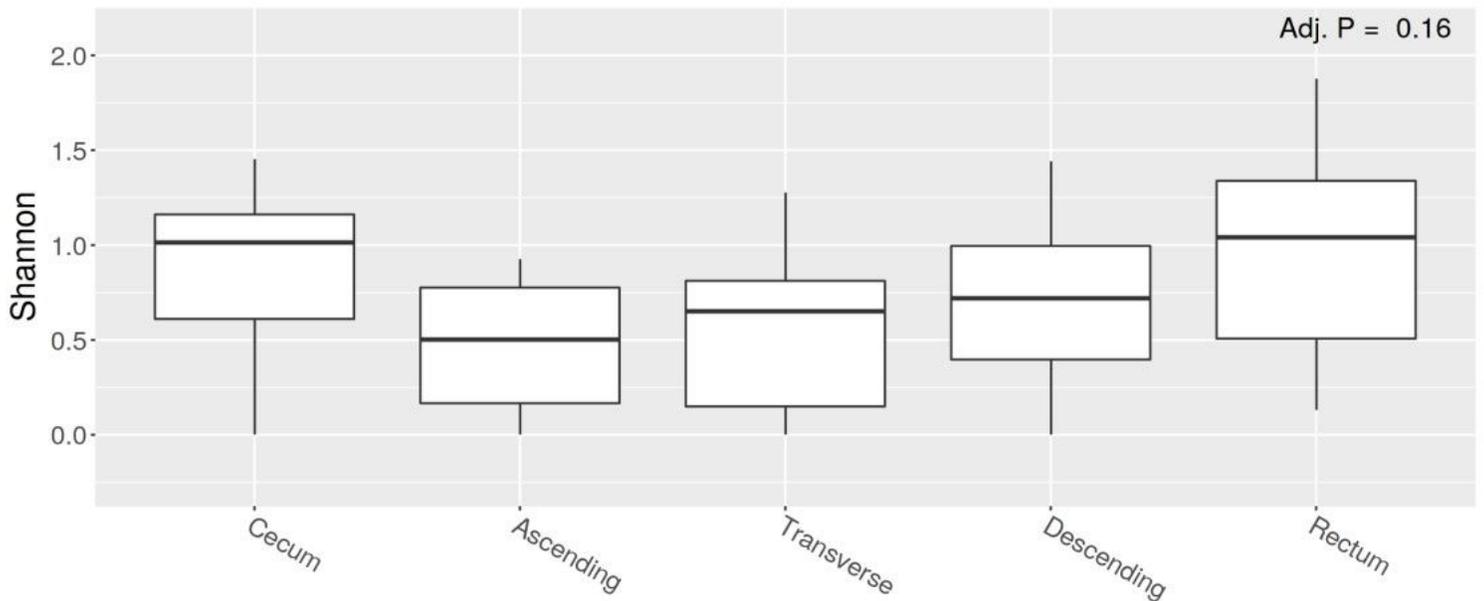
- Underscore Microbial Dysbiosis in Familial Crohn's Disease. *mBio* 7 (5). doi:10.1128/mBio.01250-16
42. Blesa J, Soriano JM, Molto JC, Manes J (2006) Factors affecting the presence of ochratoxin A in wines. *Crit Rev Food Sci Nutr* 46 (6):473-478. doi:10.1080/10408390500215803
  43. O'Brien M, Egan D, O'Kiely P, Forristal PD, Doohan FM, Fuller HT (2008) Morphological and molecular characterisation of *Penicillium roqueforti* and *P. paneum* isolated from baled grass silage. *Mycol Res* 112 (Pt 8):921-932. doi:10.1016/j.mycres.2008.01.023
  44. Abdel-Azeem AM, Held BW, Richards JE, Davis SL, Blanchette RA (2019) Assessment of biodegradation in ancient archaeological wood from the Middle Cemetery at Abydos, Egypt. *PloS one* 14 (3):e0213753. doi:10.1371/journal.pone.0213753
  45. Weichel M, Schmid-Grendelmeier P, Rhyner C, Achatz G, Blaser K, Cramer R (2003) Immunoglobulin E-binding and skin test reactivity to hydrophobin HCh-1 from *Cladosporium herbarum*, the first allergenic cell wall component of fungi. *Clin Exp Allergy* 33 (1):72-77
  46. Richard ML, Sokol H (2019) The gut mycobiota: insights into analysis, environmental interactions and role in gastrointestinal diseases. *Nature reviews Gastroenterology & hepatology*. doi:10.1038/s41575-019-0121-2
  47. Garcia-Mazcorro JF, Ishaq SL, Rodriguez-Herrera MV, Garcia-Hernandez CA, Kawas JR, Nagaraja TG (2019) Review: Are there indigenous *Saccharomyces* in the digestive tract of livestock animal species? Implications for health, nutrition and productivity traits. *Animal : an international journal of animal bioscience*:1-9. doi:10.1017/S1751731119001599
  48. Cervantes JL, Weinerman B, Basole C, Salazar JC (2012) TLR8: the forgotten relative revindicated. *Cellular & molecular immunology* 9 (6):434-438. doi:10.1038/cmi.2012.38
  49. Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal* 11 (12):2639-2643. doi:10.1038/ismej.2017.119
  50. Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, Kauserud H (2010) ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol* 10:189. doi:10.1186/1471-2180-10-189

## Figures



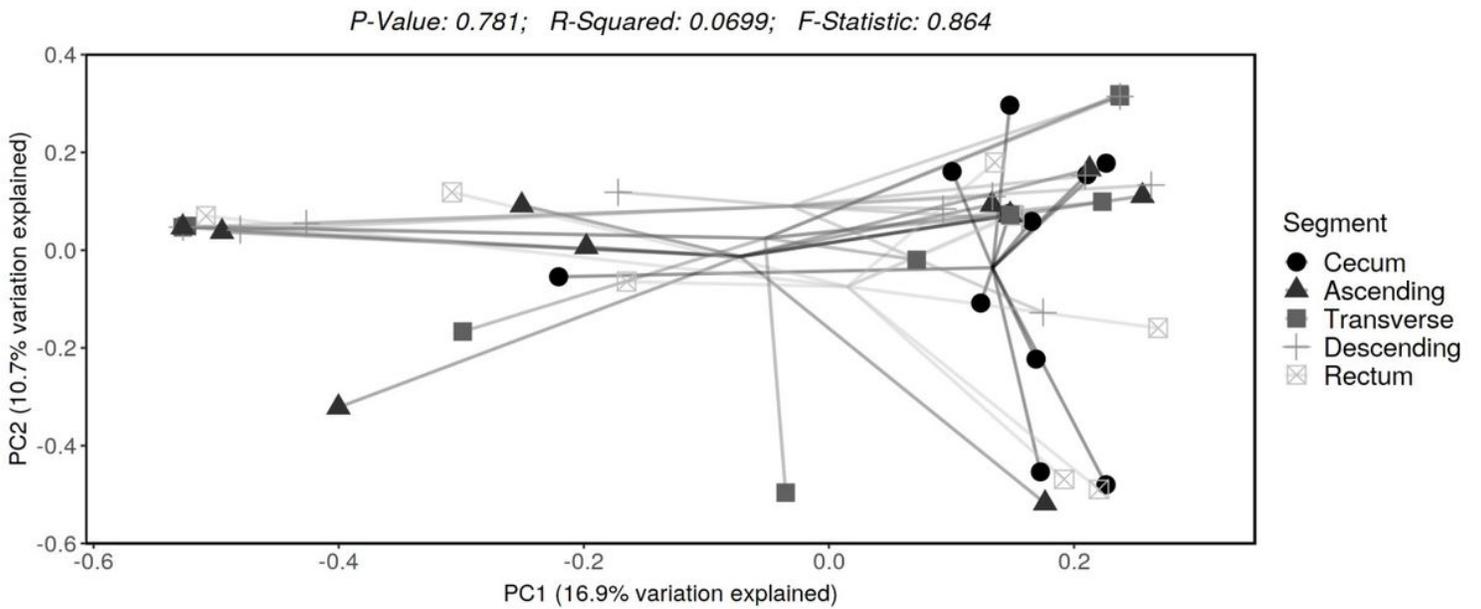
**Figure 1**

Mycobiota beta diversity inter-individually. Principal-coordinate analysis (PCoA) with weighted Bray-Curtis dissimilarity shows that there was a significant inter-individual difference in fungal beta diversity in 14 participants (*P* value = 0.001,  $R^2=0.36$ , PERMANOVA test). PC1 and PC2 represented the top two principal coordinates that capture most of the diversity. The fraction of diversity captured by the coordinate was shown as a percentage in the corresponding axis. Each centroid represented each individual with five colon segments.



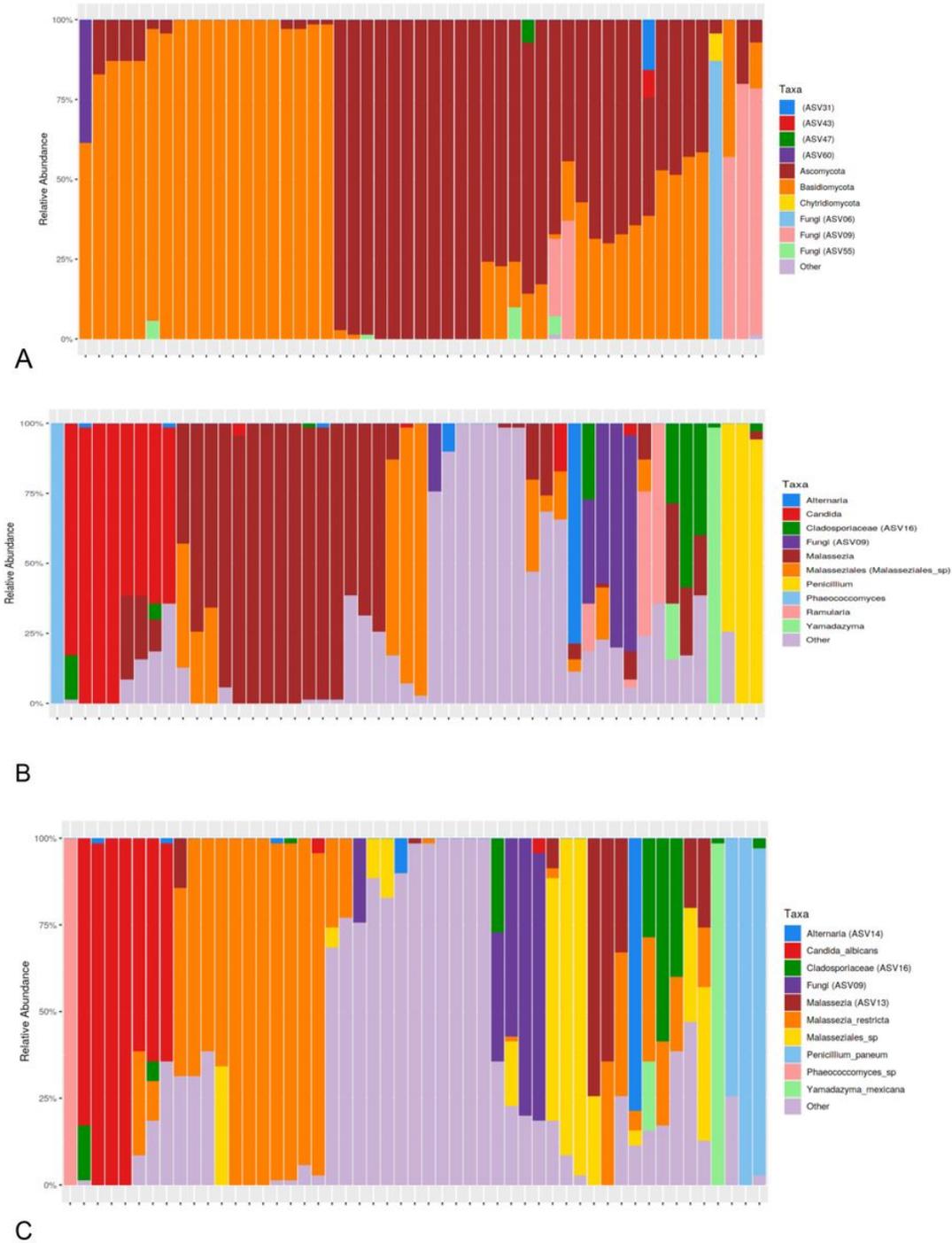
**Figure 2**

Box plot of alpha diversity of fungi across the colon segments. There was no significant difference in fungi richness and evenness (the Shannon index) across the colon segments (*q* value = 0.16, Kruskal-Wallis test).



**Figure 3**

Mycobiota beta diversity across the colon segments. Principal-coordinate analysis (PCoA) with weighted Bray-Curtis dissimilarity shows there was no significant difference in the fungal beta diversity across the colon segments in 14 participants ( $P$  value = 0.78,  $R^2=0.07$ , PERMANOVA test). PC1 and PC2 represented the top two principal coordinates that capture most of the diversity. The fraction of diversity captured by the coordinate was shown as a percentage in the corresponding axis. Each centroid represented each colon segment.



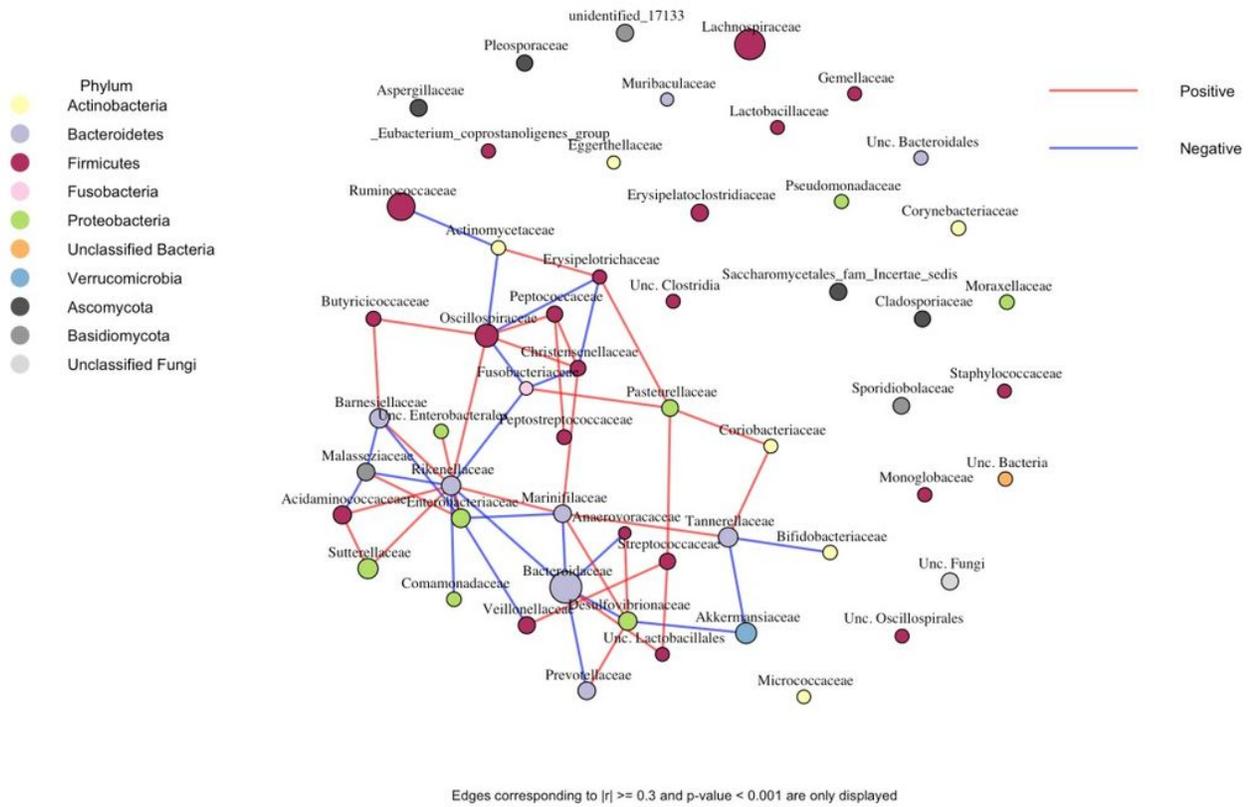
**Figure 4**

**A** Stacked bar chart of the relative abundance (% ,Y axis) of fungal phyla in all study samples.

**B** Stacked bar chart of the relative abundance (% ,Y axis) of fungal genera in all study samples.

**C** Stacked bar chart of the relative abundance (% ,Y axis) of fungal species in all study samples.

### SparCC on family-level data of all samples (N=50)



**Figure 5**

SparCC-fungal and bacterial families network analysis. The node size represents the relative abundance of the families. Node color corresponds to phylum taxonomic classification. Edges between the nodes represent correlation between the nodes they connect. Red is the positive correlation and blue is the negative correlation respectively. Only significant correlations (two-sided  $P$  value  $< 0.001$  based on bootstrapping of 1000 repetitions) with an absolute correlation magnitude  $\geq 0.3$  are presented.