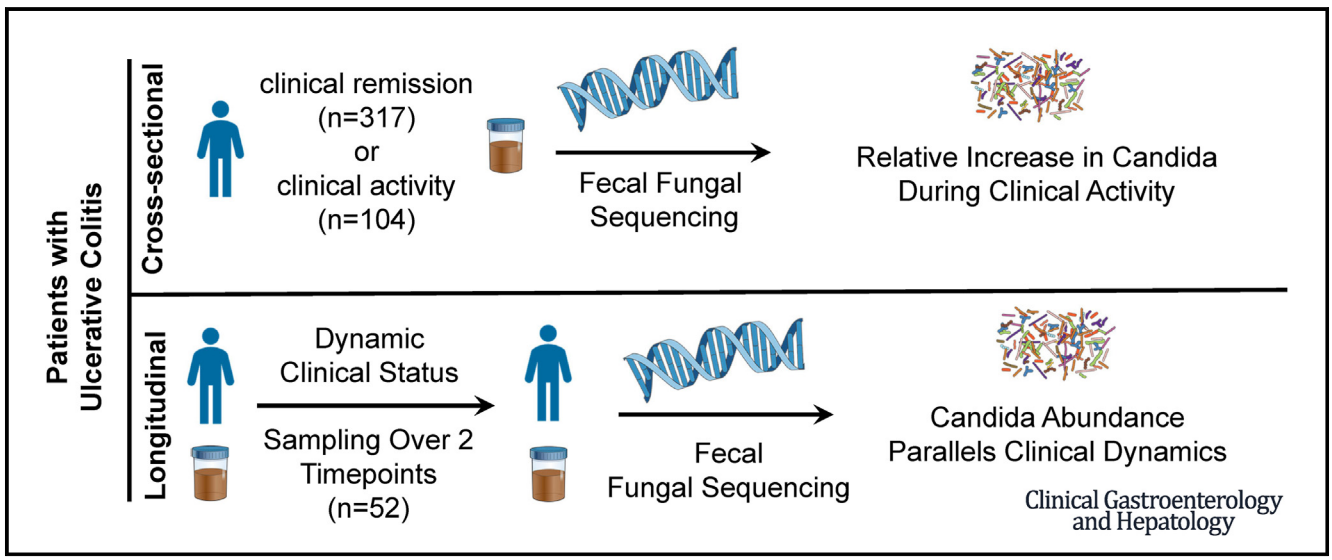
Dynamics of the Gut Mycobiome in Patients With Ulcerative Colitis

Sushrut Jangi,[1] Katie Hsia,[1] Naisi Zhao,[2] Carol A. Kumamoto,[3] Sonia Friedman,[1] Siddharth Singh,[4] and Dominique S. Michaud[2]

[1]*Department of Medicine, Tufts Medical Center, Boston, Massachusetts;* [2]*Public Health and Community Medicine, Tufts University School of Medicine, Boston, Massachusetts;* [3]*Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts; and* [4]*Division of Gastroenterology, Department of Medicine, University of California, San Diego, California*



BACKGROUND & AIMS: **Intestinal fungi have been implicated in the pathogenesis of ulcerative colitis (UC). However, it remains unclear if fungal composition is altered during active versus quiescent disease.**

METHODS: **We analyzed clinical and metagenomic data from the Study of a Prospective Adult Research Cohort with Inflammatory Bowel Disease (SPARC IBD), available via the IBD Plexus Program of the Crohn's & Colitis Foundation. We evaluated the fungal composition of fecal samples from 421 patients with UC during clinical activity and remission. Within a longitudinal subcohort (n = 52), we assessed for dynamic taxonomic changes across alterations in clinical activity over time. We examined if fungal amplicon sequence variants and fungal-bacterial relationships were altered during activity versus remission. Finally, we classified activity in UC using a supervised machine learning random forest model trained on fungal abundance data.**

RESULTS: **During clinical activity, the relative abundance of genus *Candida* was increased 3.5-fold (P-adj < 1×10^{-4}) compared with during remission. Patients with longitudinal reductions in clinical activity demonstrated parallel reductions in *Candida* relative abundance ($P < .05$). *Candida* relative abundance correlated with *Parabacteroides diastonis*, *Faecalibacterium prausnitzii*, and *Bacteroides dorei* relative abundance ($P < .05$) during remission; however, these correlations were disrupted during activity. Fungal abundance data successfully classified patients with**

Abbreviations used in this paper: ASV, amplicon sequence variant; FMT, fecal microbial transplantation; IBD, inflammatory bowel disease; IQR, interquartile range; ITS2, Internal Transcribed Spacer 2; PRO-2, 2-item patient-reported outcome; UC, ulcerative colitis.

Most current article

1542-3565/$36.00
https://doi.org/10.1016/j.cgh.2023.09.023

active or quiescent UC (area under the curve ~0.80), with *Candida* relative abundance critical to the success of the model.

CONCLUSIONS: **Clinical activity in UC is associated with an increased relative abundance of *Candida*, cross-sectionally and dynamically over time. The role of fecal *Candida* as a target for therapeutics in UC should be evaluated.**

Keywords: Fungome; Gut Microbiome; Fecal Microbial Transplantation.

Ulcerative colitis (UC) is a form of inflammatory bowel disease (IBD) that primarily affects the colon, leading to frequent relapses, hospitalizations, surgeries, and increased lifetime morbidity.[1] The pathogenesis of UC is likely driven by disruptions in host-microbiome homeostasis, with alterations in specific bacterial taxa and microbial products linked to colonic inflammation.[2] However, therapies directed at modulating the gut microbiome, including probiotics, antibiotics, or fecal microbial transplantation (FMT), have had modest and inconsistent effects in treating UC, presumably because of interindividual differences in an incompletely characterized microbiome.[3,4] Specifically, the role of nonbacterial microbial kingdoms in UC pathogenesis, notably fungi, and the effect of manipulating their populations on disease course is poorly understood.

Gut fungi, and the fungal genus *Candida*, have been previously implicated in the pathogenesis of UC. Dysbiosis of the fungal microbiome (mycobiome) has been observed in patients with IBD compared with healthy individuals, with patients with UC showing relative increases in *Candida*.[5] Furthermore, genetic polymorphisms in fungal antigen-sensing genes, including *Dectin-1*, have been linked to severe forms of UC.[6] Oral gavage of *Candida* in mice exacerbates Th17-mediated inflammation,[7] and filamentous forms of *Candida* can activate the inflammasome and induce colonic Th17 responses.[8] Interestingly, patients with UC undergoing FMT who have higher fecal *Candida* before transplant, demonstrate favorable responses to the microbial therapy.[9] We also previously showed that *Candida* is enriched during endoscopic activity versus remission in fecal samples among a small cohort of patients with UC ($n = 53$).[10]

Despite these findings, it remains unclear whether the gut mycobiome changes during active UC compared with quiescent disease. In this study, we investigated potential associations between the mycobiome and UC by performing an expanded, longitudinal secondary analysis of clinical and metagenomic metadata obtained from a prospective cohort (Study of a Prospective Adult Research Cohort with Inflammatory Bowel Disease [SPARC]).[11] We hypothesized that *Candida* would be enriched during clinical activity. By identifying potential culprit fungal taxa associated with inflammation, we hoped to characterize fungal-UC phenotypes that could support personalized approaches to therapies, including FMT, probiotics, or antifungal treatments.

Methods

Cohort Description

The cohort of patients included in this study was derived from SPARC IBD, a geographically diverse longitudinal research cohort of patients with IBD using standardized data and biosample collection methods and processing techniques.[11] Demographics and collection methods of this cohort have been previously described.[11]

Inclusion/Exclusion Criteria for Study Cohort

Cross-sectional. For this study, we included 421 patients from the SPARC cohort who had a history of UC, underwent colonoscopy, and had available Internal Transcribed Spacer 2 (ITS2) fungal mycobiome sequencing data (Supplementary Figure 1). Patients were either clinically active or in clinical remission.

Longitudinal. Among the 421 patients in the cross-sectional cohort, we examined the subset of patients ($n = 52$) who contributed 2 serial fecal samples over time. Patients either had clinical activity or clinical remission at baseline, then had repeat clinical assessment at follow-up.

Definitions

A 2-item patient reported outcome (PRO-2) is a validated practical index of disease activity in UC that includes a rectal bleeding score and stool frequency score.[12] Clinical activity was defined as PRO-2 $\geq$ 2. Clinical remission was defined by PRO-2 $\leq$ 1. See the Supplementary Methods for additional definitions.

Microbial Preparation and Sequencing Analysis

For details regarding microbial isolation, library preparation, and sequencing analysis, see the Supplementary Methods.

Analyses

Relative abundance, fractional prevalence, diversity, and differential abundance. To evaluate the relative abundance, we excluded rare organisms, defined as taxa that constituted less than 1% of the total abundance.

Proportions were calculated by dividing the number of reads assigned to each taxon by the total number of reads. Fractional prevalence was calculated by dividing the number of patients containing a specific taxon by the total number of patient samples. Alpha diversity of the mycobiome was assessed using observed amplicon sequence variants (ASVs) and Shannon diversity indices. For beta-diversity, weighted UniFrac was used to generate a distance matrix, which was then ordinated using nonmetric multidimensional scaling. A negative binomial model (DESeq2, version 3.15)[13] was used to calculate the differential abundance of specific taxa across comparison groups in patients with UC, adjusted for potential confounding effects of age, gender, immunosuppressive therapy, steroid use, antibiotic use, and probiotic use. Dynamic changes in microbial abundance were assessed over longitudinal samples obtained at 2 timepoints across clinical activity categories. For heterogeneity and machine learning analysis, see the Supplementary Methods.

Statistical Analysis

We conducted all analyses and figure preparation using R studio version 2023.03.0[14] or Quantitative Insights Into Microbial Ecology 2.[15] We performed a Wilcoxon rank sum test ($P < .05$) to evaluate whether alpha diversity indices differed significantly between comparators. To evaluate significance for beta-diversity, permutational analysis of variance analysis was used to calculate significance between groups ($P < .05$). For differential abundance analysis, P values were corrected for multiple testing using the Benjamini and Hochberg method.[16] See the Supplementary Methods for additional statistical analyses.

Ethical Considerations

Ethical approval for the study was obtained from the Tufts Medical Center Institutional Review Board.

Results

Clinical Cohort

Of 467 patients with ITS2 sequencing of the mycobiome, 421 patients had PRO-2 scores obtained concurrently with stool specimen collection (median interval, 0 months). One hundred and four patients (25%) had clinically active disease, with either mild-moderate clinically active disease (n = 73; PRO-2 of 2–3) or severe clinically active disease (n = 31; PRO-2 of 4–6), whereas 317 patients (75%) were in clinical remission (PRO-2 $\leq$ 1). Study cohort characteristics are given in Table 1. Longitudinal fecal sampling was available among 52 patients, who each contributed 2 fecal samples (total of 104 samples) over a median of 4 months (interquartile range [IQR], 2–8) (Supplementary Table 1). Within this subcohort (n = 52), at baseline, 37 patients were in clinical remission (71%) and 15 patients had clinical activity (29%).

What You Need to Know

Background

The fungal gut microbiome has been previously implicated in the pathogenesis of inflammatory bowel disease. *Candida*, in particular, can drive gut inflammatory responses and has been shown to be increased in IBD patients compared to healthy people.

Findings

This study found that stool *Candida* was increased during inflammation in ulcerative colitis compared to remission. Furthermore, *Candida* decreased over time as patients moved from active disease towards quiescence.

Implications for patient care

Elevated *Candida* is linked to active ulcerative colitis, and future studies might evaluate if treating *Candida* may contribute to improved outcomes.

Relative Abundance and Prevalence of Fungal Phylum and Genera in Ulcerative Colitis

We assessed the relative abundance of fungal phylum across all samples (Figure 1*A*). We found that most fungal ASVs belonged to the phylum Ascomycota (86% of ASVs), with a minority classified into the phylum Basidiomycota (3%) (Figure 1*B*). Using Nucleotide Basic Local Alignment Search Tool, unidentifiable ASVs (11%) represented low prevalence fungi or dietary contaminants (Supplementary Table 2). We assessed the relative abundance of fungal genera across all samples (Figure 1*C*). Forty percent of ASVs belonged to genus *Saccharomyces*, 30% to *Candida*, 7% to *Penicillium*, 3% to *Rhodotorula*, 3% to *Agaricus*, with the remaining (17%) belonging to unidentified genera (Figure 1*D*). Using Nucleotide Basic Local Alignment Search Tool, 2 fungal ASVs were resolvable to *Geotrichum candidum* and *Candida glabrata* (Supplementary Table 2).

Diversity of the Mycobiome in Clinically Active and Remission Ulcerative Colitis

Among 421 patients with UC, the alpha and beta diversity of the mycobiome was similar during clinical activity compared with remission (mean Shannon diversity index = 0.94 vs 0.93, $P = .42$; beta diversity $P = .92$, permutational multivariate analysis of variance) (Supplementary Figure 2*A*, *B*).

Table 1. Study Cohort

		Clinical status		
	Total cohort	Clinical activity	Clinical remission	*P* value
n	421	104	317	
Age, mean (SD)	46.4 (15.2)	46.4 (15.7)	46.4 (15.0)	.97
Gender, female, frequency, n (%)	227 (53.9)	57 (54.8)	170 (53.5)	.83
Disease extent, n (%)				
E1	30 (7.1)	11 (10.6)	19 (6)	.11
E2	89 (21.1)	18 (17.3)	71 (22.4)	.27
E3	226 (53.7)	53 (51)	173 (54.6)	.52
Unknown	76 (18.1)	22 (21.2)	54 (17)	.34
Disease duration, *y*, median (IQR)	11 (5–20)	10 (4–20)	12 (6–20)	.81
PRO-2 score, median (IQR)	0 (0–1)	3 (2–4)	0 (0–0)	**< .005**
Rectal bleeding score, median (IQR)	0 (0–0)	1 (0–3)	0 (0–0)	**< .005**
Stool frequency score, median (IQR)	0 (0–1)	2 (1–3)	0 (0–0)	**< .005**
Physician's Global Assessment Score, median (IQR)	0 (0–1) (n = 288)	1 (1–2) (n = 72)	0 (0–0) (n = 216)	**< .005**
MES, median (IQR)	1 (0–2) (n = 117)	2 (1–3) (n = 30)	0 (0–1) (n = 87)	**< .005**
Interval, fecal sample collection and PRO-2 scoring, *mo*, median (IQR)	0 (0–0)	0 (0–0)	0 (0–0)	.71
Interval, fecal sample collection and PGA scoring, *mo*, median (IQR)	0 (0–0)	0 (0–0)	0 (0–0)	.87
Interval, fecal sample collection and MES scoring, *mo*, median (IQR)	0 (0–0)	0 (0–0)	0 (0–1)	< .05
Longitudinal sampling available	104	25	79	—
Immunosuppressive exposure, n (%)	300 (71.3)	72 (69.2)	228 (71.9)	.59
Vedolizumab, n (%)	102 (24.2)	33 (31.7)	69 (21.7)	**< .05**
Anti-TNF, n (%)	156 (37.1)	35 (33.7)	121 (38.1)	.41
Thiopurine/methotrexate, n (%)	136 (32.3)	29 (27.9)	107 (33.6)	.27
Ustekinumab, n (%)	19 (4.5)	12 (11.5)	7 (2.2)	**< .005**
Tofacitinib, n (%)	28 (6.7)	13 (12.5)	15 (4.7)	**< .05**
Tacrolimus/cyclosporine, n (%)	20 (4.8)	3 (2.9)	17 (5.3)	.42
Concurrent steroid use (n = 321), n (%)	**29 (9.0)**	**17 (22.4)**	**12 (4.8)**	**< .005**
Concurrent antibiotic use (n = 255), n (%)	**42 (16.5)**	**14 (36.8)**	**28 (12.9)**	**< .005**
Concurrent probiotic use (n = 281), n (%)	**64 (22.8)**	**16 (27.1)**	**48 (21.6)**	.37

Bolded values are statistically significant.
IQR, interquartile range; MES, Mayo endoscopic subscore; PGA, Physician's Global Assessment; PRO-2, 2-item patient-reported outcome; SD, standard deviation; TNF, tumor necrosis factor.

Candida is Increased in Relative Abundance During Clinical Activity in Ulcerative Colitis

Among 421 patients with UC, during clinical activity, genus *Candida* showed 2.8-fold higher relative abundance compared with during clinical remission (*P*-adj < .001) (Figure 2*A*). Genera *Agaricus* and *Rhodotorula* showed lower relative abundances during clinical activity compared with during remission (*P*-adj < .001) (Figure 2*A, B*). After adjusting for the effects of age, gender, and immunosuppressive exposure, patients with UC with clinical activity showed a 3.5-fold higher relative abundance of *Candida* compared with remission (*P*-adj < 1×10^{-4}), with a 2.6-fold lower relative abundance of *Agaricus* during clinical activity (*P*-adj < 1×10^{-4}) (Figure 2*C*). After adjusting for concurrent steroid use, patients with UC with clinical activity showed a 2.7-fold higher relative abundance of *Candida* compared with remission (*P*-adj < .001). Similarly, adjusting for concurrent antibiotic use, patients with UC with clinical activity showed a 2.3-fold higher relative abundance of *Candida* compared with remission (*P*-adj < .001).

Although adjusting for probiotic use still resulted in patients with UC with clinical activity showing a 1.8-fold increased relative abundance of *Candida* compared with remission, this result was less significant (*P*-adj < .07). In a stratified analysis, there was no difference in *Candida* relative abundance among patients in remission or activity who did or did not take probiotics.

Candida Differential Relative Abundance Increases With Worsening Disease Severity Index in Ulcerative Colitis

The relative abundance of *Candida* linearly increased with PRO-2 score (ranging from 0–6; *P*-adj < .05) (Supplementary Figure 3*A*). Additionally, we found a trend toward increased *Candida* relative abundance with endoscopic severity (*P*-adj = .15) (Supplementary Figure 3*B*). *Candida* differential relative abundance was not significantly affected by immunosuppressive exposure (*P*-adj = .23) (Supplementary Figure 3*C*). Furthermore, *Candida* relative abundance was not significantly affected by concurrent steroid (*P*-adj = .78), antibiotic (*P*-adj = .33), or probiotic use (*P*-adj = .56) (Supplementary Figure 4).

Candida Longitudinal Dynamics Are Perturbed by Changes in Clinical Activity in Ulcerative Colitis

Candida relative abundance was significantly altered across 2 serial timepoints (median of 4 months; IQR, 2–8) among all patients with longitudinal fecal sampling (n = 52 patients) (*P* < .05) (Figure 3*A*). Patients who had no change in clinical activity (n = 39; median time interval of 4 months; IQR, 2–8) had stable *Candida* relative abundance across time (*P* = .32) (Figure 3*B*). Patients who experienced a change in clinical activity (n = 13, either from clinical remission to activity or vice versa; median time interval of 5 months; IQR, 3–9) also experienced a change in *Candida* relative abundance over time (*P* < .05) (Figure 3*C*).

Among patients persistently in clinical remission (n = 33), there was no difference in the overall low *Candida* mean relative abundance between timepoints (0.02 vs 0.06; *P* = .72) (Figure 3*D*). Among patients with persistent clinical activity (n = 6), there was no difference in the overall high *Candida* mean relative abundance between timepoints (0.28 vs 0.40; *P* = .29) (Figure 3*F*). Among patients with clinical activity who then developed

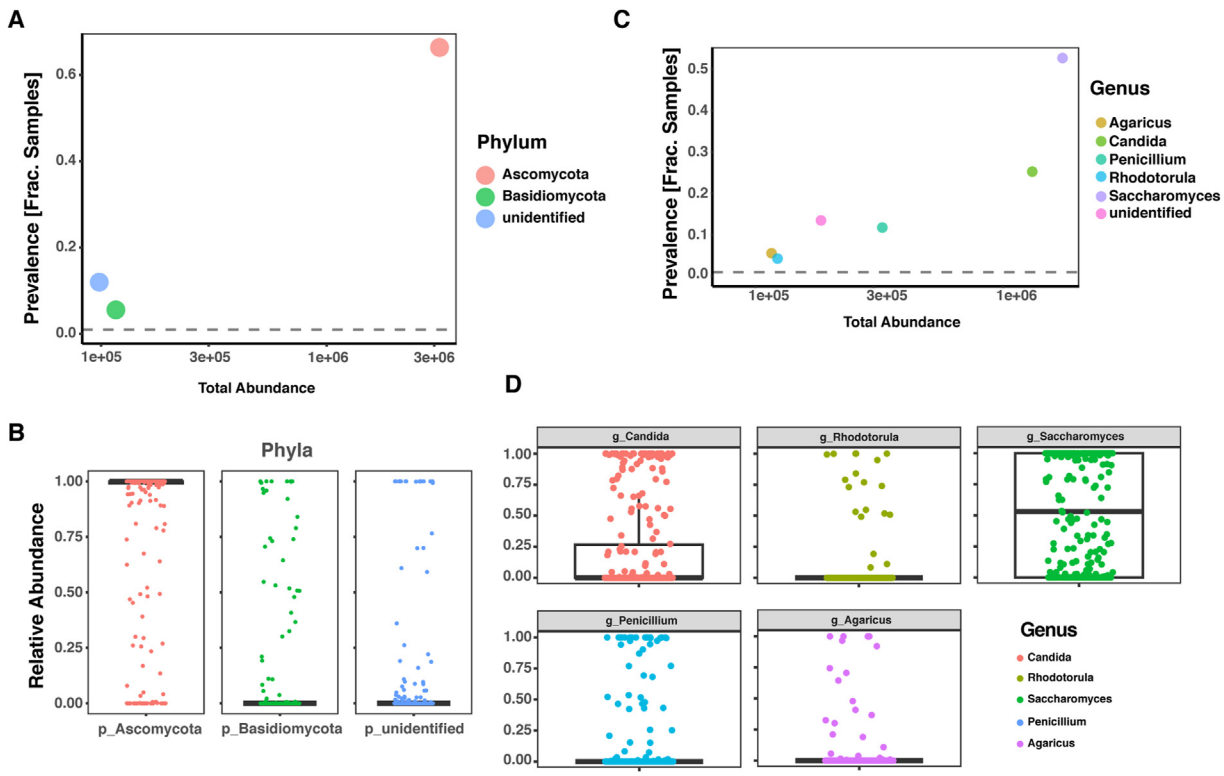


Figure 1. Prevalence and abundance of fungal taxa among patients with UC. (*A*) Fractional prevalence and total abundance of fungal phyla. X-axis shows the proportion of total samples with detected phylum, y-axis shows the phylum microbial counts. (*B*) Relative abundance of fungal phyla. X-axis shows relative abundance, y-axis shows fungal phylum name. (*C*) Fractional prevalence and total abundance of fungal genera. X-axis shows the proportion of samples with detected genus, y-axis shows the total genus microbial counts. (*D*) Relative abundance of fungal genus. X-axis shows relative abundance, y-axis shows fungal genus name.

clinical remission (n = 9), *Candida* mean relative abundance concomitantly decreased from high to low (0.43 vs 0.02; $P < .05$) (Figure 3*E*). Among patients in clinical remission who then developed clinical activity (n = 4), *Candida* was not detectable in any samples.

Among clinically active patients who maintained activity (n = 6) or progressed to remission (n = 9), there was no significant difference in baseline steroid use (2/5 vs 0/6 patients; $P = .18$), antibiotic use (2/2 vs 3/7 patients; $P = 1.0$), or probiotic use (1/1 vs 3/6 patients; $P = 1.0$). Among patients in remission who maintained remission (n = 33) or progressed to activity (n = 4), there was no significant baseline difference in steroid use (0/33 vs 1/4; $P = .11$), antibiotic use (5/22 vs 0/2; $P = 1.0$), or probiotic use (3/22 vs 1/2 patients; $P = .31$).

Heterogeneity Among Candida Amplicon Sequence Variants

Within the entire patient cohort, we identified 5 unique *Candida* ASVs (Figure 4*A*). Three of these ASVs represent different strains of *Candida albicans* (ASV 1-3), whereas 1 ASV represents *Candida tropicalis* (ASV 4) or *C glabrata* (ASV 5). *C albicans* was detected in 108 patients (28.1%; ASV 1–3), whereas *C tropicalis* was detected in 2 patients (0.5%) and *C glabrata* was detected in 9 patients (2%) (Figure 4*B*). On differential abundance analysis, *C albicans* (combined aggregate of ASV 1–3) demonstrated a 3-fold increase during clinical activity versus remission (P-adj $< 1.1 \times 10^{-3}$), whereas *C glabrata* and *C tropicalis* were of low prevalence and did not exhibit significant differences (Figure 4*C*).

Classification of Ulcerative Colitis Using a Supervised Machine Learning Model

A supervised random forest machine learning model was trained and tested on the fungal sequencing data from the patient cohort achieved a testing area under the curve of ~0.80, with overall accuracy of 81% in classifying clinically quiescent versus active UC (Figure 5*A*, *B*). *Saccharomyces* and *Candida* demonstrated the highest feature importance to the model (0.22 and 0.20, respectively) (Figure 5*C*, Supplementary Table 3), with individual performance characteristics for each predictor provided in (Supplementary Figure 5).

Candida-Bacterial Interkingdom Relationships

Within the entire cohort (n = 421), *Candida* relative abundance was positively correlated with the relative abundances of bacterial genera *Parabacteroides diastonis* (+0.15; $P < .005$), and negatively correlated with *Eubacterium hallii* (−0.10; $P < .05$) and *Bifidobacterium adolescentis* (−0.11; $P < .05$) (Supplementary Figure 6*A*). During clinically active disease (n = 93), *Candida* relative abundance did not significantly correlate with any bacterial taxa (Supplementary Figure 6*B*). However, during

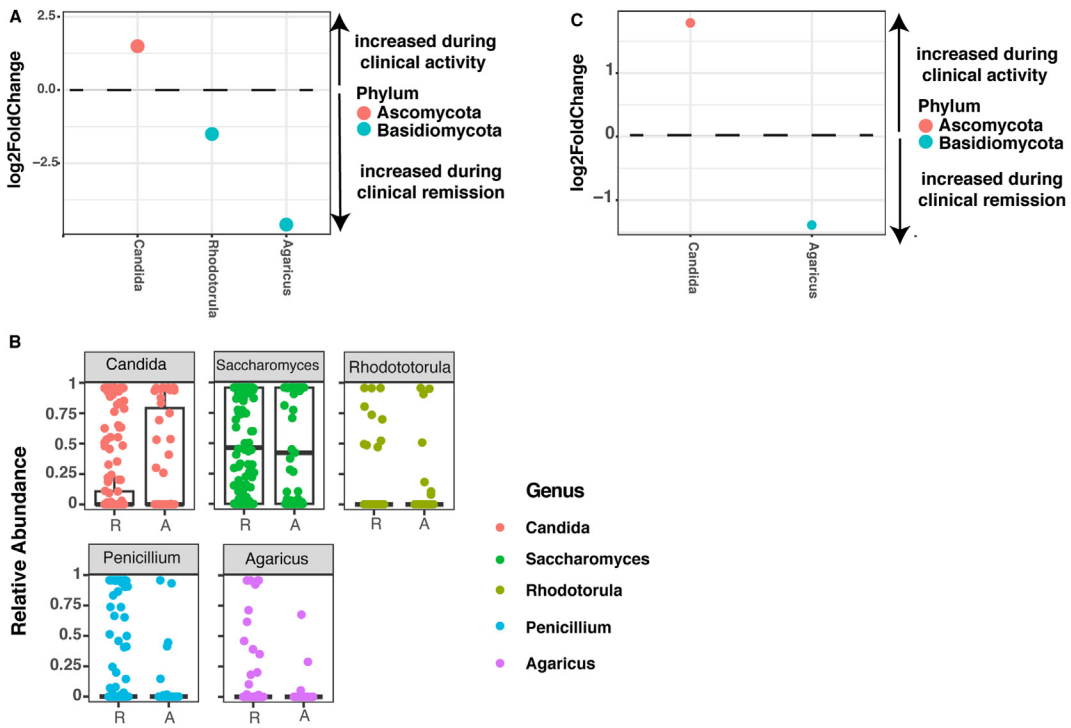


Figure 2. Differential abundance of fungal genera in clinical activity versus remission among patients with UC. (*A*) Differential log 2-fold changes in fungal genera among patients in clinical activity versus clinical remission. (*B*) Relative abundance of highly abundant fungal genera in patients with remission (R) versus activity (A). (*C*) Differentially abundant fungal genera, after adjusting for age, sex, and immunosuppressive exposure.

remission (n = 284), *Candida* relative abundance was positively correlated with *P diastonis* (+0.16; $P < .05$), *Faecalibacterium prausnitzii* (+0.14; $P < .05$), and *Bacteroides dorei* (+0.13; $P < .05$) (Supplementary Figure 6*C*).

Discussion

Previous deep-sequencing profiles of several IBD cohorts have consistently shown a high relative abundance of *Candida* in patients with IBD compared with healthy control subjects. However, how *Candida* and other fungi are altered during inflammation in patients with UC has not been thoroughly examined. In this large, secondary analysis of a prospective adult cohort of 421 patients with UC, we found that the relative abundance of *Candida* increased 3-fold in patients with clinical activity. This association is strengthened by the large sample size, adjustment for medication use, the relative high abundance and prevalence of *Candida* across patient samples, the use of a validated definition of clinical activity, and dynamic assessment over longitudinally sampled patients. Furthermore, fecal samples and clinical indices were obtained concurrently in time. Overall, these findings substantiate a link between *Candida* and inflammatory activity in UC, lending support to the hypothesis that the mycobiome may prove useful as a target for microbial manipulation to improve disease outcomes in UC.

Mouse models have previously shown that oral gavage of *C albicans* worsens inflammation in dextran sodium sulfate colitis,[7] mediated through Dectin-1 receptors[17] and downstream fine tuning of Th1/Th17 balance in the colon.[18] Complementary to these mechanistic studies, our study provides epidemiologic data linking *Candida* to colonic inflammation in UC.

The current study is further bolstered by assessment of the mycobiome dynamically over time. Longitudinal changes in the mycobiome in humans have been few, with none performed to date in IBD or UC. *Candida* has been previously reported to be more temporally stable on repeated measurements within the same person,[19] suggesting that it is more likely to be a persistent resident commensal in the human gut. Our finding that the relative abundance of *Candida* decreases in patients who evolve from clinical activity to clinical remission supports that *Candida* populations may dynamically parallel the degree of colonic inflammation.

The current study also examined the ASV composition of genus *Candida*. ASVs representing *C albicans* predominated and were significantly linked to inflammatory status. A prior study found that gut *C albicans* isolates from human colonic samples demonstrated high genetic variability coinciding with altered pathogenic transcriptional programs.[20] Further assessment of ASV heterogeneity and strain-level analyses in future studies may provide additional insight into *Candida* abundance, genetic variation, and UC pathogenesis. In our analysis of fungal-bacterial correlations, during conditions of quiescence, *Candida* abundance positively correlated with *P diastonis*, *F prausnitzii*, and *B dorei*, genera known to induce regulatory responses in the gut.[21–23] *Candida* can influence gut bacterial assembly through metabolic competition and collaboration.[24–26] Given these findings, it remains possible that during activity, the expansion of *Candida* is associated with bacterial dysbiosis and a loss of anti-inflammatory bacteria, such as *Faecalibacterium*.

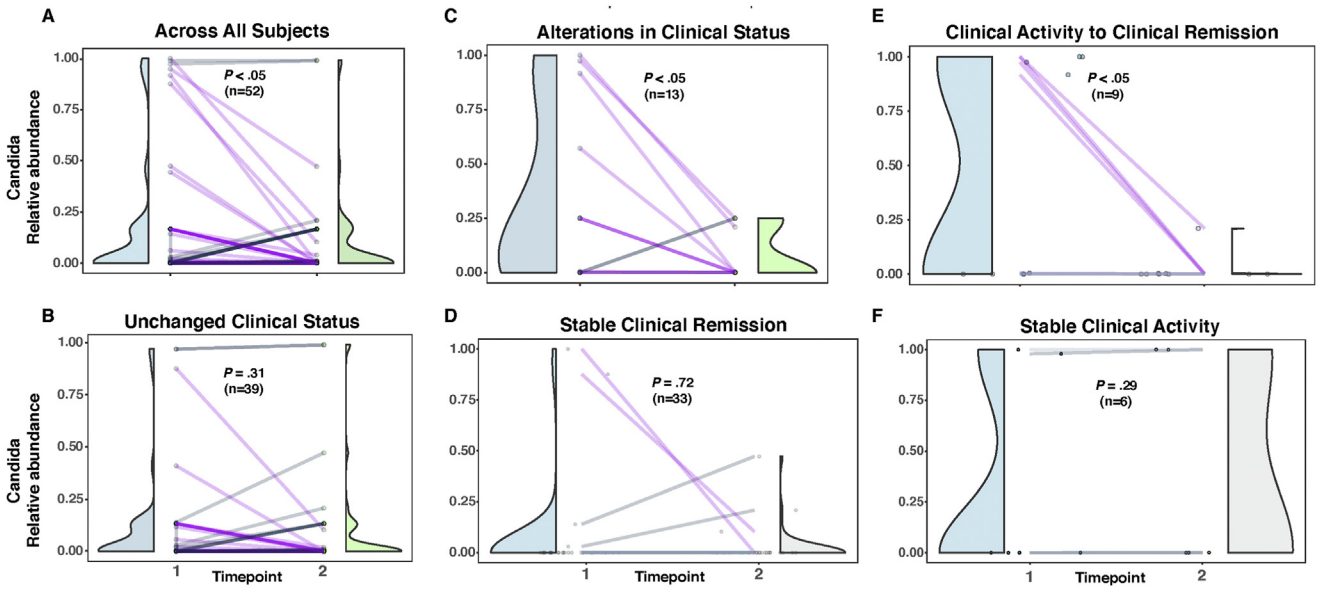


Figure 3. Longitudinal relative abundance of *Candida* across disease activity. Relative abundance of *Candida* at timepoint 1 and 2 in patients with UC (*A*) in entire longitudinal cohort (n = 52), (*B*) with no change in clinical status (n = 39), (*C*) with shift in clinical status (n = 13), (*D*) with stable clinical remission (n = 33), (*E*) with clinical activity evolving to clinical remission (n = 9), and (*F*) with persistent clinical activity (n = 6).

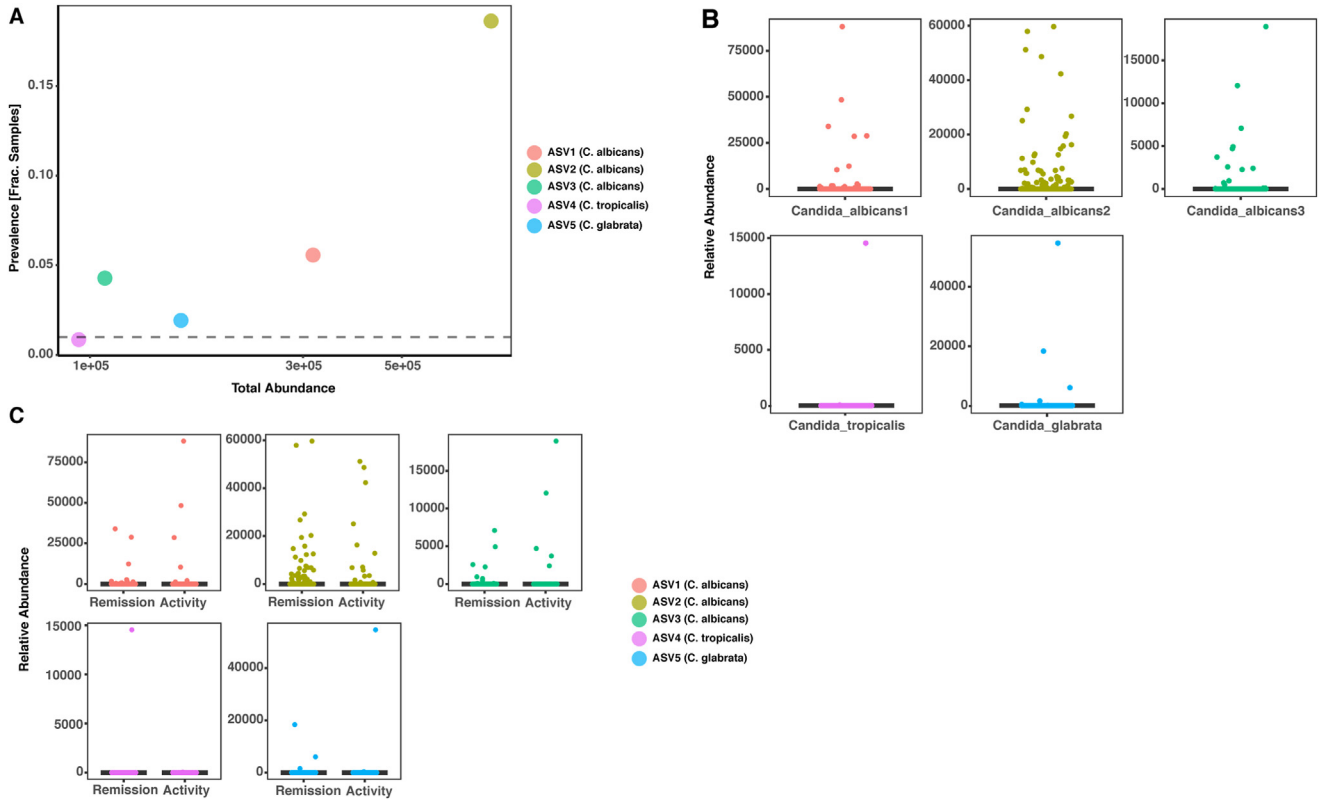


Figure 4. Heterogeneity in *Candida* ASVs across patients with UC. (*A*) Fractional prevalence (x-axis) and total counts (y-axis) of ASVs attributed to *Candida* genus. (*B*) Relative abundance of each *Candida* ASV. (*C*) Relative abundance of each *Candida* ASV during clinical remission versus activity.

Our study also used fungal abundance data to train a supervised machine learning random forest model, to classify patients into active or quiescent UC. The model achieved an area under the curve of ∼0.80 using cross-sectional microbiome data from highly abundant fungal taxa. Future studies could combine fungal-bacterial features to improve the operating characteristics of an externally validated model.

Given our observed association of *Candida* relative abundance with clinical activity in UC, the implication of our study is that treating patients with antifungal therapy may contribute to inducing remission. In a proof-of-concept, *C albicans* reduced FMT efficacy in a mouse model of *Clostridium difficile* infection, with antifungal therapy restoring FMT response.[25] A randomized controlled trial of fluconazole in *Candida* colonized patients with UC led to clinical and biochemical improvements over placebo.[27] Selecting patients for antifungal therapy alongside standard medications is a promising consideration in future trials, especially in refractory or severe UC.

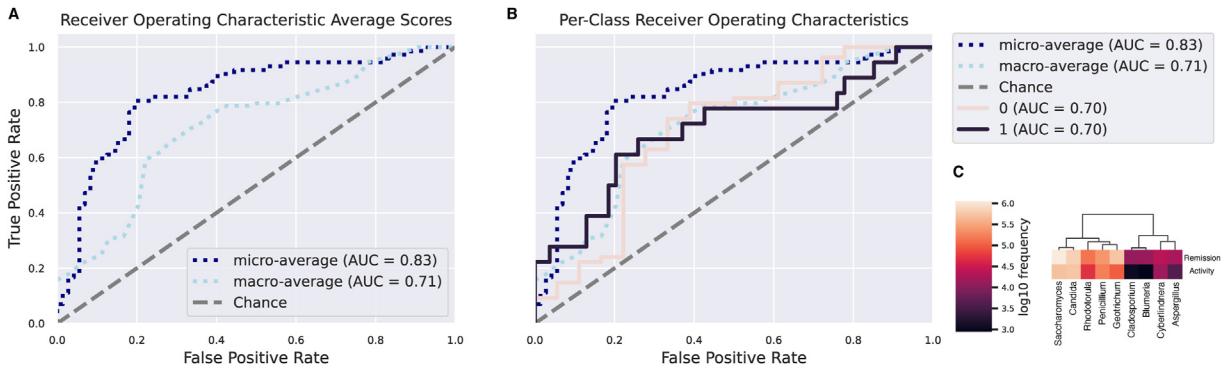


Figure 5. Classification of clinical activity in ulcerative colitis using a supervised machine learning model trained on fungal taxa abundance data. Y-axis represents the true-positive rate and x-axis represents false-positive rate, with (*A*) overall micro-averaging (*dark blue squares*) and macroaveraging (*light blue squares*) (*B*) per class, with prediction of quiescence (*pink line*) and prediction of activity (*black line*). (*C*) Heat map demonstrates the feature importance of each predictor taxa contributing to the model, with *light squares* demonstrating high sequence counts and *dark squares* demonstrating low sequence counts. AUC, area under the curve.

Our study is limited by lack of dietary intake history, because nutritional components can influence fungal colonization.[28] For example, although we found enriched *Agaricus* (mushrooms) during remission, this may reflect ingestion of edible mushrooms, rather than increases in a true gut colonizer.[29] Another shortcoming is that the study primarily rested on ITS2-based sequencing. Clinical data were also limited by the smaller subcohort size among longitudinal samples, the absence of histologic data, and lack of reporting on antifungal use. We were additionally limited by the availability of concurrent endoscopic scoring; however, our use of the validated PRO-2 allowed for longitudinal comparisons, and reflects a practical and clinically meaningful association between PRO-2 and *Candida* that can be tested in subsequent investigations. Future studies that combine large-scale gut compositional and functional assessments will better translate host-microbe associations toward generating mechanistic insights that may power clinically meaningful microbial interventions.

Overall, we report that among a large, prospective, well-characterized cohort of 421 patients, the abundance of genus *Candida* is significantly associated with clinical activity. The strength of this association is underlined even after adjustment for immunosuppressive exposure, antibiotic use, steroid use, and probiotic use. ASVs belonging to *C albicans* strongly linked to inflammatory activity. Interkingdom relationships between *Candida* and anti-inflammatory bacterial taxa evident during clinical quiescence were disrupted during clinical activity. Supervised machine learning, resting primarily on *Candida* and *Saccharomyces* abundance features, achieved a promising ability to classify patients with active versus quiescent UC. In the light of prior observations supporting a mechanistic role for *Candida* in colonic inflammation, this report substantiates a relationship between *Candida* and clinical activity in UC, opening the door for trials exploring antifungal, probiotic, or microbial transplantation therapeutics alongside conventional treatments to optimize outcomes in UC.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at http://doi.org/10.1016/j.cgh.2023.09.023.

References

1. Fumery M, Singh S, Dulai PS, et al. Natural history of adult ulcerative colitis in population-based cohorts: a systematic review. Clin Gastroenterol Hepatol 2018;16:343–356.
2. Sartor RB, Wu GD. Roles for intestinal bacteria, viruses, and fungi in pathogenesis of inflammatory bowel diseases and therapeutic approaches. Gastroenterology 2017;152:327–339.
3. Dang X, Xu M, Liu D, et al. Assessing the efficacy and safety of fecal microbiota transplantation and probiotic VSL#3 for active ulcerative colitis: a systematic review and meta-analysis. PLoS One 2020;15:e0228846.
4. Narula N, Kassam Z, Yuan Y, et al. Systematic review and meta-analysis: fecal microbiota transplantation for treatment of active ulcerative colitis. Inflamm Bowel Dis 2017;23:1702–1709.
5. Sokol H, Leducq V, Aschard H, et al. Fungal microbiota dysbiosis in IBD. Gut 2017;66:1039–1048.
6. Gantner BN, Simmons RM, Underhill DM. Dectin-1 mediates macrophage recognition of *Candida albicans* yeast but not filaments. EMBO J 2005;24:1277–1286.
7. Jawhara S, Thuru X, Standaert-Vitse A, et al. Colonization of mice by *Candida albicans* is promoted by chemically induced colitis and augments inflammatory responses through galectin-3. J Infect Dis 2008;197:972–980.
8. Joly S, Ma N, Sadler JJ, et al. Cutting edge: *Candida albicans* hyphae formation triggers activation of the Nlrp3 inflammasome. J Immunol 2009;183:3578–3581.
9. Leonardi I, Paramsothy S, Doron I, et al. Fungal trans-kingdom dynamics linked to responsiveness to fecal microbiota transplantation (FMT) therapy in ulcerative colitis. Cell Host Microbe 2020;27:823–829.
10. Hsia K, Zhao N, Chung M, et al. Alterations in the fungal microbiome in ulcerative colitis. Inflamm Bowel Dis 2023; 29:1613–1621.
11. Raffals LE, Saha S, Bewtra M, et al. The development and initial findings of a Study of a Prospective Adult Research Cohort with Inflammatory Bowel Disease (SPARC IBD). Inflamm Bowel Dis 2022;28:192–199.
12. Dragasevic S, Sokic-Milutinovic A, Stojkovic Lalosevic M, et al. Correlation of Patient-Reported Outcome (PRO-2) with endoscopic and histological features in ulcerative colitis and Crohn's disease patients. Gastroenterol Res Pract 2020;2020:2065383.
13. Weiss S, Xu ZZ, Peddada S, et al. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 2017;5:27.
14. Posit team. Posit Software P. Boston, MA: RStudio: Integrated Development Environment for R. Posit Software, 2023.
15. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 2019;37:852–857.
16. Benjamini YH. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995;57:289–300.
17. Iliev ID, Funari VA, Taylor KD, et al. Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. Science 2012;336:1314–1317.
18. Bonifazi P, Zelante T, D'Angelo C, et al. Balancing inflammation and tolerance in vivo through dendritic cells by the commensal *Candida albicans*. Mucosal Immunol 2009;2:362–374.
19. Kondori N, Nowrouzian F, Ajdari M, et al. *Candida* species as commensal gut colonizers: a study of 133 longitudinally followed Swedish infants. Med Mycol 2020;58:485–492.
20. Li XV, Leonardi I, Putzel GG, et al. Immune regulation by fungal strain diversity in inflammatory bowel disease. Nature 2022; 603:672–678.
21. Kverka M, Zakostelska Z, Klimesova K, et al. Oral administration of *Parabacteroides distasonis* antigens attenuates experimental murine colitis through modulation of immunity and microbiota composition. Clin Exp Immunol 2011;163:250–259.
22. Qiu X, Zhang M, Yang X, et al. *Faecalibacterium prausnitzii* upregulates regulatory T cells and anti-inflammatory cytokines

in treating TNBS-induced colitis. J Crohns Colitis 2013; 7:e558–e568.
23. Zhao H, Xu H, Chen S, et al. Systematic review and meta-analysis of the role of *Faecalibacterium prausnitzii* alteration in inflammatory bowel disease. J Gastroenterol Hepatol 2021; 36:320–328.
24. Erb Downward JR, Falkowski NR, Mason KL, et al. Modulation of post-antibiotic bacterial community reassembly and host response by *Candida albicans*. Sci Rep 2013;3:2191.
25. Zuo T, Wong SH, Cheung CP, et al. Gut fungal dysbiosis correlates with reduced efficacy of fecal microbiota transplantation in *Clostridium difficile* infection. Nat Commun 2018; 9:3663.
26. Garcia-Gamboa R, Kirchmayr MR, Gradilla-Hernandez MS, et al. The intestinal mycobiota and its relationship with overweight, obesity and nutritional aspects. J Hum Nutr Diet 2021; 34:645–655.
27. Jena A, Dutta U, Shah J, et al. Oral fluconazole therapy in patients with active ulcerative colitis who have detectable *Candida* in the stool: a double-blind randomized placebo-controlled trial. J Clin Gastroenterol 2022;56:705–711.
28. Hoffmann C, Dollive S, Grunberg S, et al. Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. PLoS One 2013;8:e66019.
29. Hess J, Wang Q, Gould T, et al. Impact of *Agaricus bisporus* mushroom consumption on gut health markers in healthy adults. Nutrients 2018;10:1402.

Correspondence
Address correspondence to: Sushrut Jangi, MD, Division of Gastroenterology, Tufts Medical Center, 800 Washington Street, Boston, Massachusetts 02111. e-mail: sushrut.jangi@tuftsmedicine.org.

Acknowledgments
The authors acknowledge the Tufts University High Performance Computing Cluster (https://it.tufts.edu/high-performance-computing), which was used for the research reported in this article. The results published here are in whole or part based on data obtained from the IBD Plexus program of the Crohn's & Colitis Foundation.

CRediT Authorship Contributions
Sushrut Jangi (Conceptualization: Lead; Data curation: Lead; Formal analysis: Lead; Funding acquisition: Lead; Investigation: Lead; Methodology: Lead; Supervision: Lead; Visualization: Lead; Writing – original draft: Lead; Writing – review & editing: Lead)
Katie Hsia (Investigation: Supporting; Software: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)
Naisi Zhao (Conceptualization: Supporting; Formal analysis: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)
Carol A. Kumamoto (Conceptualization: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)
Sonia Friedman (Writing – review & editing: Supporting)
Siddharth Singh (Writing – review & editing: Supporting)
Dominique S. Michaud (Conceptualization: Supporting; Writing – review & editing: Supporting)

Conflicts of interest
These authors disclose the following: Sushrut Jangi received consulting fees from Bristol Myers Squibb. Siddharth Singh has received institutional research grants from AbbVie and Pfizer; and received personal fees from Pfizer. The remaining authors disclose no conflicts.

Funding
The project described was supported by the National Center for Advancing Translational Sciences, National Institutes of Health (NIH), Award Number KL2TR002545. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Additionally, this work was supported by the Charlton Research Award and Natalie V. Zucker Research Center for Women Scholars Award, both from the Tufts University School of Medicine.

Supplementary Methods

Additional Definitions

The Mayo endoscopic subscore was graded endoscopically as either 0 (normal colon), 1 (erythema, blurring of vascular pattern), 2 (friability, absence of vascular pattern, or erosions), or 3 (spontaneous bleeding or ulcers). Ascertainment of clinical indices and stool sample collections occurred simultaneously. Longitudinal patients underwent an additional timepoint of clinical assessment and stool sample collection. Immunosuppressive exposure was defined as use of any thiopurines, anti–tumor necrosis factor inhibitors, vedolizumab, Janus kinase inhibitors, anti-interleukin-12 or interleukin-23 inhibitors, or cyclosporine or tacrolimus up until the time of index clinical assessment and stool collection for microbiome sequencing. Immunosuppressive-naive patients were untreated or exclusively on mesalamine therapy up until the time of index clinical assessment and stool collection for microbiome sequencing. Steroid use, antibiotic use, and probiotic use was reported within 2–3 months of fecal sampling for most patients (Table 1).

Fungal Isolation and Library Preparation

ITS-based deep sequencing of fungal rDNA from fecal samples from SPARC IBD was performed by Diversigen. DNA from 100 mg of fecal sample per subject was extracted using MagAttract PowerSoil DNA EP kit (Qiagen) using mechanical-based lysis with garnet beads. Samples were prepared using AccuPrime High Fidelity kit (Invitrogen) to amplify ITS2 (ITS3/ITS4). Sequencing was performed using the 2 × 300 bp paired-end protocol. The primers used for amplification contained adaptors for MiSeq sequencing (Illumina) and single-end barcodes allowing pooling and direct sequencing of polymerase chain reaction products. Target-specific primer sequences were: ITS3F: 5'GCATCGATGAAGAACGCAGC 3' and ITS4: 5' TCCTCCGCTTATTGATATGC 3'. Libraries were quantified with Quant-iT PicoGreen dsDNA assay (Invitrogen) and confirmed using gel electrophoresis.

Bacterial Library Preparation and Annotation

DNA from 100 mg of fecal sample per subject was extracted using MagAttract PowerSoil DNA EP kit (Qiagen) using mechanical-based lysis with garnet beads. Samples were quantified using the Quant-iT PicoGreen dsDNA assay (Invitrogen) and libraries were prepared with a procedure adapted from the Illumina DNA prep kit. Libraries were quantified with Quant-iT PicoGreen dsDNA assay (Invitrogen) and confirmed using the Fragment Analyzer (Agilent). Libraries were sequenced on an Illumina NovaSeq 6000 using a paired-end 2 × 150 bp flow cell. DNA sequences were taxonomically classified using the MetaPhlAn2 analysis tool.[21] MetaPhlAn2 maps reads to clade-specific marker genes identified from ∼17,000 reference genomes and estimates clade abundance within a sample from these mappings.

Internal Transcribed Spacer Sequencing Analysis

Sequences were quality filtered and denoised using the Quantitative Insights Into Microbial Ecology (QIIME2; version 2021.11).[14,15] Sequences were checked for quality and trimmed for barcodes and polymerase chain reaction primers, with low-quality reads lower than a Q-score of 30 dropped using DADA2 (version 2022.8.0).[15] Reads were binned for a minimal sequence length of 280 forward and 261 reverse. We built fungal community matrices from the resulting unique ASVs. Sequence alignment was performed with MAFFT[16] and phylogenic trees were built with FastTree.[17] Taxonomy was assigned with a fitted classifier using the UNITE (version 8.3) fungal reference data.[18,19] Normalization of fungal sequencing abundances was performed by standardizing abundances to the median scaling depth. Median normalization was applied before differential abundance analysis.[20]

Heterogeneity Analysis of Amplicon Sequence Variants

Nucleotide Basic Local Alignment Search Tool (BLASTN)[15] was used to identify fungal sequences that could not be identified by the UNITE fungal classifier. Prevalence and differential abundance of these ASVs were compared across the cohort, during clinical activity and remission.

Supervised Machine Learning

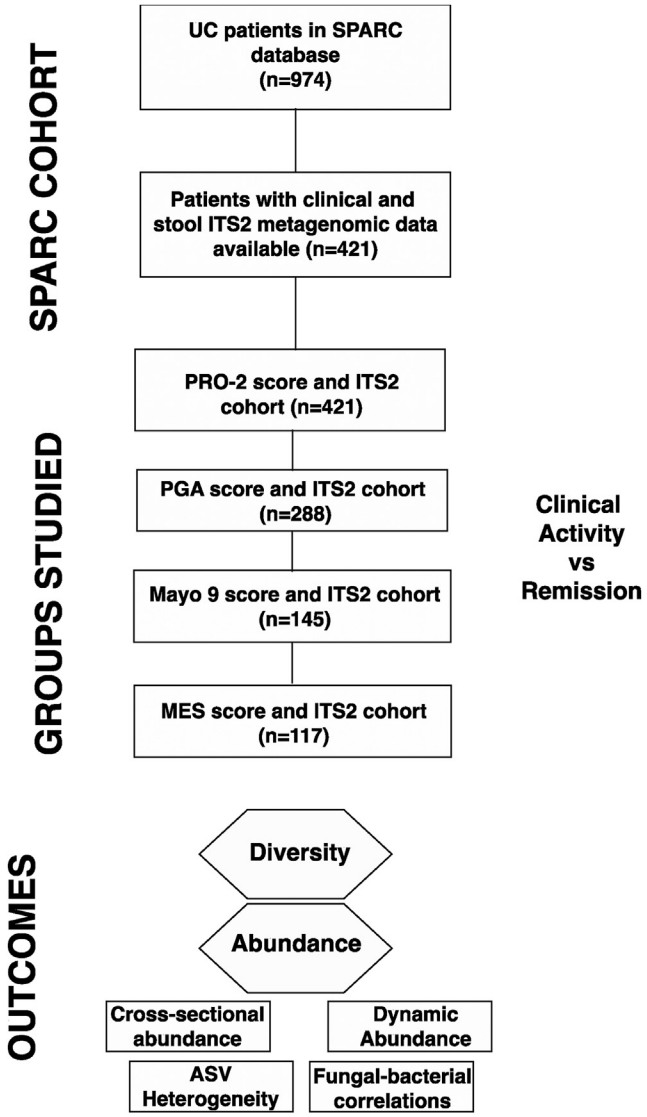
To classify clinical activity versus remission in patients with ulcerative colitis, a random forest model was trained using the features of fungal taxa, with assessment of the model's performance characteristics. The supervised machine learning random forest model was developed using the QIIME2 plugin q2-sample-classifier.[20] To reduce the dimensionality of the feature space, the model was trained using only abundant (n = 9) fungal taxa, sorted to the genus level, with low prevalence (<1%) organisms filtered out. Data were split into a training set (75%) and testing set (25%). Training was performed using 5-fold cross-validation. Assessment of the performance characteristics on the tested samples included area under the receiver operating characteristic curves, accuracy, and feature importance of each fungal taxa to model performance. A separate receiver operating characteristic curve for each predictor (fungal taxa)

in the model was also constructed, by testing the performance of each predictor in classifying disease activity or remission.

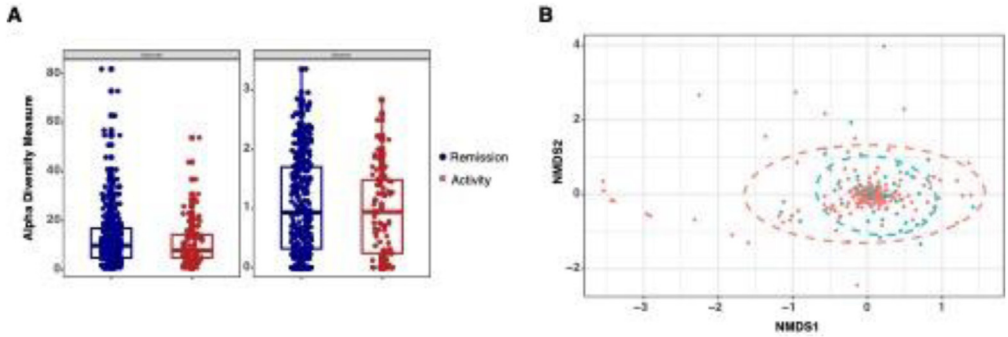
Additional Statistical Analyses

For longitudinal assessments, transformed abundances were assessed with a linear-mixed effects model. Correlations between the bacteriome and mycobiome were determined by drafting a matrix of Spearman correlations between the relative abundances of bacterial genera and fungal genera. All QIIME2 and R code used in this analysis are available on github at: https://github.com/katiehsia/dynamic_fungome.

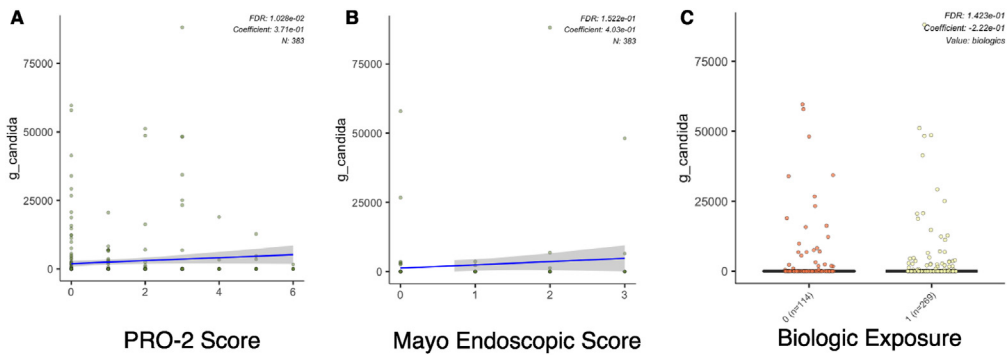
To assess for univariate associations between clinical metadata on the differential abundance of fungal groups, fitting was performed using the Maaslin2 package in R.[20] Subjects were included as random effects to account for correlations in repeated measures, whereas disease category and time were modeled as fixed effects. Only the 30 most abundant bacterial taxa and only significantly altered fungal taxa were included in the analysis. Correlation analysis was performed using the CorrPlot package in R.[20]



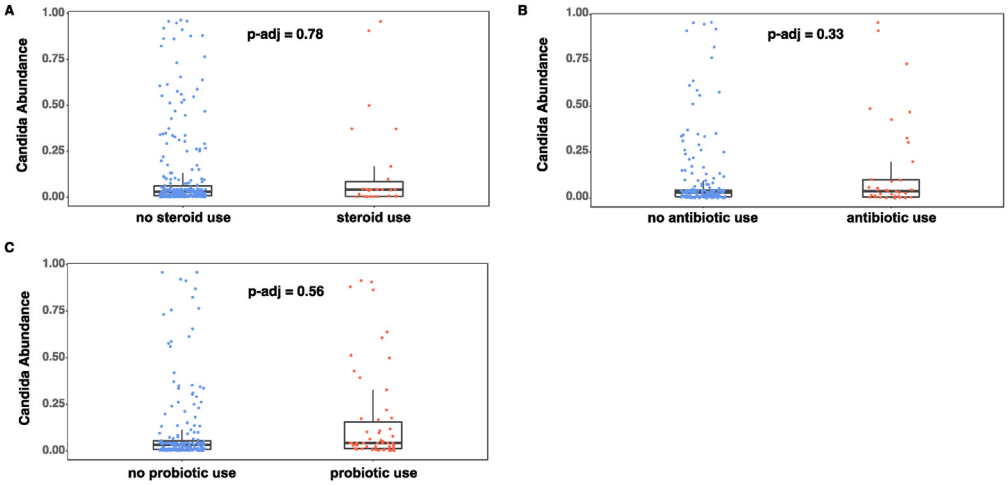
Supplementary Figure 1. Study scheme. MES, Mayo endoscopic subscore; PGA, Physician's Global Assessment.



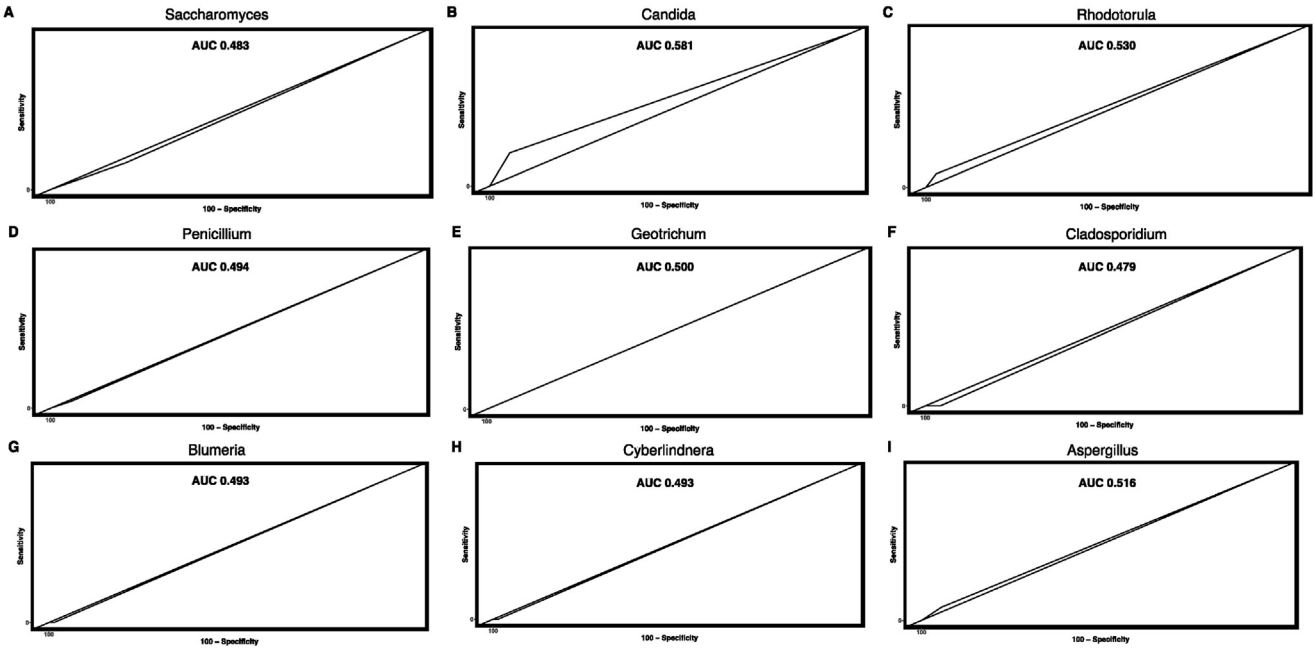
Supplementary Figure 2. Fungal diversity during clinical activity and remission among patients with ulcerative colitis. (*A*) Alpha diversity as measured by observed and Shannon diversity index, with significance tests performed using Wilcoxon signed rank sum tests. (*B*) Beta diversity, using weighted Unifrac for distance metrics and nonmetric multidimensional scaling for ordination, with significance tested using permutational analysis of variance analysis.



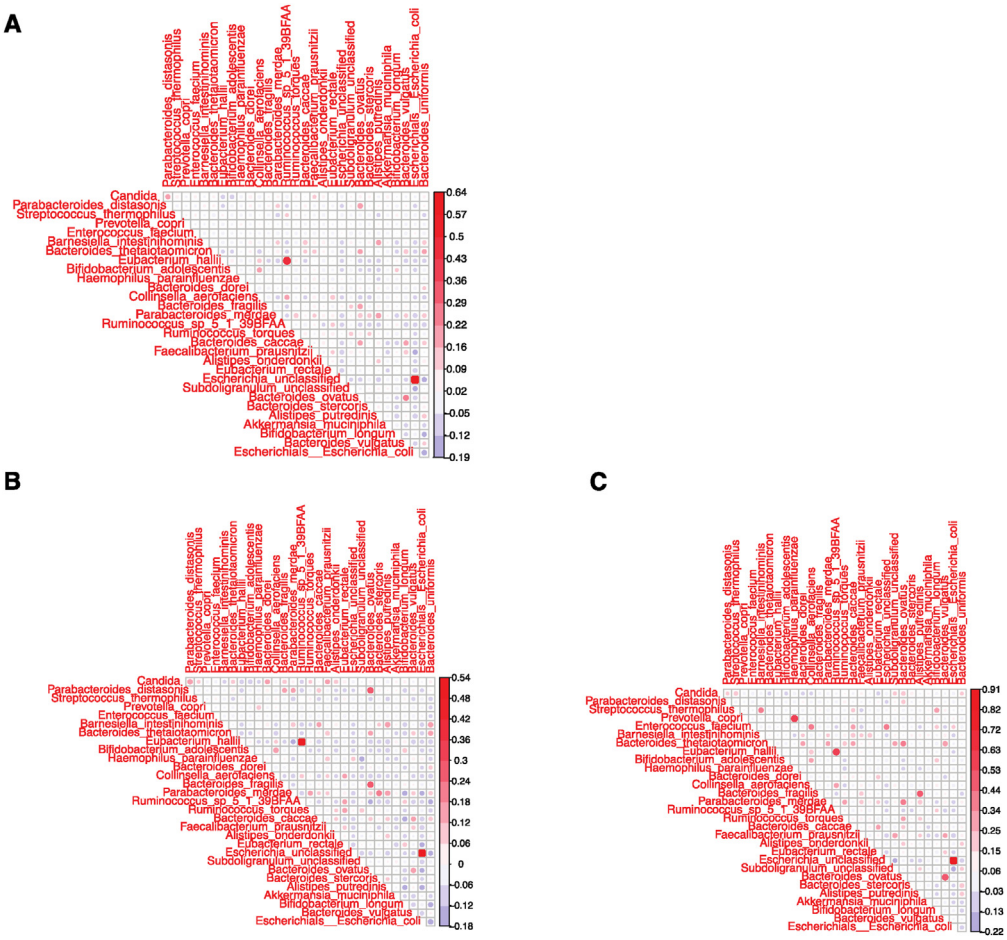
Supplementary Figure 3. Relative abundance of *Candida* genus across disease severity scores with x-axis representing index score and y-axis representing *Candida* relative abundance including (*A*) the 2-item PRO-2 score and (*B*) the Mayo endoscopic score (*C*) in patients with immunosuppressive exposure versus immunosuppressive naive.



Supplementary Figure 4. Relative abundance of *Candida* genus across patients (*A*) with no steroid use versus steroid use, (*B*) no antibiotic use versus antibiotic use, and (*C*) no probiotic use versus probiotic use.



Supplementary Figure 5. Receiver operating characteristic curves for fungal taxa in classifying activity versus remission in patients with ulcerative colitis, including (*A*) *Saccharomyces*, (*B*) *Candida*, (*C*) *Rhodotorula*, (*D*) *Penicillium*, (*E*) *Geotrichum*, (*F*) *Cladosporium*, (*G*) *Blumeria*, (*H*) *Cyberlindnera*, and (*I*) *Aspergillus*, with AUC (area under the curve) indicated for each plot.



Supplementary Figure 6. Correlogram displaying Spearman correlations between *Candida* genus and bacterial relative abundance. Area of *dots* is proportional to correlation coefficient, with *dark red* representing strongly positive correlations and *dark blue* representing strongly negative correlations.

Supplementary Table 1. Longitudinal Cohort of Patients with UC

	Baseline clinical status		
	Clinical remission	Clinical activity	*P* value
N	37	15	
Age, mean (SD)	48.6 (15.8)	46.2 (15.6)	.80
Pancolitis, n (%)	19 (51)	11 (73)	.22
Disease duration, *y*, median (IQR)	13 (8–26)	14 (9–21)	.79
PRO-2 score, median (IQR)	0 (0–0)	3 (2–3)	< .05
Mayo endoscopic score, median (IQR)	1 (0–2)	3 (2–3)	<.05
Remission at second timepoint, n (%) frequency	33 (89)	9 (60)	—
Clinical activity at second timepoint, n (%) frequency	4 (11)	6 (40)	—
Concurrent steroids, n (%) (n = 43)	1 (3.1)	2 (18)	.16
Concurrent antibiotics, n (%) (n = 35)	5 (20.8)	5 (45.5)	.23
Concurrent probiotics, n (%) (n = 32)	4 (16.7)	5 (62.5)	< .05

IQR, interquartile range; PRO-2, 2-item patient-reported outcome; SD, standard deviation; UC, ulcerative colitis.

Supplementary Table 2. Classification of Unidentified Fungal ASVs Using the Nucleotide Basic Local Alignment Search Tool

Unidentified ASVs	Scientific name (genus/species)	Common name	Query coverage, %	E value	Number of patients with ASV	Mean reads within cohort
ASV1	*Geotrichum candidum*	Fungus	72	$<1 \times 10^{-10}$	4	236.8
ASV2	*Daucus carota*	Carrot	100	0	56	210.0
ASV3	*Chenopodium ficifolium*	Goosefoot	99.7	0	7	183.9
ASV4	*Candida glabrata*	Fungus	100	0	9	342.3

ASV, amplicon sequence variant.

Supplementary Table 3. Feature Importance of Each Taxa Predictor to the Random Forest Classifier Model

Genus	Feature importance
Saccharomyces	0.22
Candida	0.20
Penicillium	0.14
Blumeria	0.13
Aspergillus	0.09
Cyberlindnera	0.07
Geotrichium	0.06
Rhodotorula	0.05
Cladosporium	0.04