Bacterial and Fungal Profiles as Markers of Infliximab Drug Response in Inflammatory Bowel Disease

Ventin-Holmberg, Rebecka

2021-06-22

Ventin-Holmberg , R , Eberl , A , Saqib , S , Korpela , K , Virtanen , S , Salonen , A , Sipponen , T M , Saavalainen , P & Nissilä , E 2021 , ' Bacterial and Fungal Profiles as Markers of Infliximab Drug Response in Inflammatory Bowel Disease ' , Journal of Crohn's and Colitis , vol. 15 , no. 6 , pp. 1019-1031 . https://doi.org/10.1093/ecco-jcc/jjaa252

http://hdl.handle.net/10138/341370
https://doi.org/10.1093/ecco-jcc/jjaa252

unspecified
acceptedVersion

Bacterial and fungal profiles as markers of infliximab drug response in inflammatory bowel disease

Rebecka Ventin-Holmberg[1], Anja Eberl[2], Schahzad Saqib[3], Katri Korpela[3], Seppo Virtanen[4], Taina Sipponen[2], Anne Salonen[3], Päivi Saavalainen[1] and Eija Nissilä[1*]

[1] Translational Immunology Research Program, University of Helsinki, Finland

[2] Department of Gastroenterology, Helsinki University Hospital, University of Helsinki, Finland

[3] Human Microbiome Research Program, University of Helsinki, Finland

[4] Obstetrics and Gynaecology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Corresponding author: Eija Nissilä, PhD, Translational Immunology Research Program, Biomedicum 1, 00014 Helsinki, Finland. Tel.: +358 294126731; email: eija.nissila@helsinki.fi

Short title: Faecal microbiota markers of IFX response in IBD (50 characters)

Conference: UEG, Barcelona, 2019: Eija Nissilä, Rebecka Ventin-Holmberg, Anja Eberl, Taina Sipponen, Anne Salonen, Päivi Saavalainen. Gut microbiota composition as a predictor of infliximab treatment response in inflammatory bowel disease patients.

Abstract

Background: Inflammatory bowel diseases (IBDs), Crohn's disease (CD) and ulcerative colitis (UC), are globally increasing chronic gastro-intestinal inflammatory disorders associating with altered gut microbiota. Infliximab (IFX), a TNF-alpha blocker, is used to treat IBD patients successfully though one third of the patients do not respond to therapy. No reliable biomarkers are available for prediction of IFX response. **Aims**: Our aim was to investigate the faecal bacterial and fungal communities during IFX therapy and find predictors for IFX treatment response in IBD patients. **Methods**: 72 IBD patients (25 CD and 47 UC) started IFX therapy and were followed for one year or until IFX treatment was discontinued. Amplicon sequencing approach targeting the bacterial 16S rRNA gene and fungal ITS 1 region separately was used to determine the microbiota profiles in faecal samples collected before IFX therapy, two, six, twelve weeks and one year after initiation of therapy. The response to IFX was evaluated by colonoscopy and clinically at twelve weeks after initiation. **Results**: Both the faecal bacterial and fungal profiles differed significantly between response groups before start of IFX treatment. Non-responders had lower abundances of short chain fatty acid producers, particularly of the class Clostridia and higher abundances of pro-inflammatory bacteria and fungi, such as the genus *Candida*, compared to responders. This was further indicated by bacterial taxa predicting the response in both CD and UC patients (area under curve > 0.8). **Conclusions**: Faecal bacterial and fungal microbiota composition could provide a predictive tool to estimate IFX response in IBD patients.

Keywords: Microbiota, IBD, infliximab

1. Introduction

Inflammatory bowel diseases (IBDs) are chronic inflammatory conditions of the intestine, of which the most common subtypes are Crohn's disease (CD) and ulcerative colitis (UC) [1]. These are differentiated by an inflammation that is limited to the colonic mucosa in UC and an inflammation that is transmural and can manifest anywhere in the gastrointestinal tract in CD [2]. No defined pathogenesis has been established for IBD, but the disease derives from several environmental factors and particularly an imbalanced gut microbiota in a host, who is genetically prone to IBD [3, 4, 5].

An emerging number of studies over the past 20 years have shown that gut microbiota with altered composition plays the pivotal role in the pathogenesis of IBD as reviewed by Nishida *et al*. [6]. The reduction of bacteria with anti-inflammatory capacities, such as short chain fatty acid (SCFA)-producing bacteria belonging to the Clostridia class and increases in bacteria with inflammatory capacities are detected in faecal samples of patients with IBD when compared to healthy subjects. IBD patients also show a decreased diversity of gut microbiota and a lower abundance of Firmicutes compared to healthy subjects [6]. However, most of the previous studies on intestinal microbiota have concentrated on the bacterial community (bacteriome), and research into gut fungal communities (mycobiota) in IBD has started only recently. This has enabled more comprehensive research in the role of intestinal bacteria and fungi in the pathobiology of IBD. The fungi present in the human gut are known to elicit immunomodulatory effects especially via polysaccharides in their cell wall and contribute to disease progression as reviewed by Galloway-Peña and Kontoyiannis [7].

There is no cure for IBD, but inflammation can be treated with different medications, such as steroids, thiopurines or biologicals [8]. Faecal microbiota transplant has been used as a treatment in IBD patients successfully by restoring microbial diversity [9]. Of the IBD patients receiving

conventional treatment with thiopurines, approximately 10–15% require further treatment to achieve and maintain remission [10]. Tumour necrosis factor alpha (TNF-α) is a cytokine that is part of the proinflammatory cascade, which is activated in an autoimmune reaction and is associated with both CD and UC [11]. Infliximab (IFX) is a chimeric monoclonal antibody that blocks the activity of TNF-α by binding to it and is used to induce and maintain remission in numerous autoimmune diseases, including IBD [12]. However, up to 40% of the CD patients and approximately 50% of the UC patients treated with IFX do not respond to the medication. Additionally, up to 40% of primary responders lose their response over time [13]. Early identification of the patients who do not respond to IFX therapy allows for prompt modification of the medical treatment which might reduce potential harmful side effects [14] and the cost of therapy [15]. Previous studies have investigated the gut bacteriome for the prediction of response against TNF-α blocker in IBD and suggested that the gut microbiota may provide potential biomarkers for monitoring and predicting IBD treatment outcomes although a number of these studies have been constrained by limited power [16, 17, 18, 19, 20]. To our knowledge, similar studies have not been performed on the gut mycobiota in adult IBD patients.

The potential of fungal and bacterial profiles individually and in combination as markers for predicting IFX response during treatment has not been fully explored. In the present study we investigated first, the influence of IFX therapy on both gut bacterial and fungal communities in a prospective IBD patient cohort consisting of both CD and UC patients and second, characterized potential markers of gut fungal and bacterial genera for predicting IFX drug response in IBD.

2. Materials and methods

2.1. Study design and subjects

The cohort consisted of adult patients with a diagnosis of CD or UC, for whom IFX therapy was initiated at the Department of Gastroenterology at Helsinki University Hospital between February 2017 and 2019. The reason for starting IFX treatment was active inflammation without response or with intolerance to prior conventional or biological IBD medications. IFX treatment response was assessed after induction at week twelve after treatment initiation. Stool and serum samples were collected before initiation of IFX therapy (two days to a few hours before the first infusion of IFX) and at two, six and twelve weeks, and at one year. 16S and ITS sequencing were performed at different timepoints, therefore a different number of samples is included in the analyses.

2.2. Evaluation of response against IFX

Clinicians performed endoscopy at week twelve after IFX initiation to assess endoscopic activity. Treatment outcome was evaluated using endoscopic and clinical indices. For UC and IBD unclassified (IBDU) patients, the Mayo Score (MS), combining clinical and endoscopic assessment, was applied. Remission (R), or responder to IFX, was defined as an MS of ≤ 2 points combined with an endoscopic sub score of 0 or 1 points. Partial remission (PR), or partial responder to IFX, was defined as an MS of 3 or 4 point with and endoscopic sub score of 1 or 2. Patients with an MS of ≥ 5 points and an endoscopic sub score of ≥ 2 were defined as non-responders (NR) [10, 21].

For CD patients The Simple Endoscopic Score for Crohn's Disease (SES-CD) was used to determine endoscopic activity at week twelve [22]. Endoscopic remission was defined as a SES-CD score of 0-2 points and endoscopically mildly active disease as a SES-CD score of 3-6 points

and was regarded as a PR. SES-CD 7–15 suggested moderately active and SES-CD ≥ 16 severely active disease and were regarded as NR to IFX treatment [23, 24].

Some recruited patients did not undergo endoscopy at twelve weeks after initiation of IFX treatment. Disease activity of these patients was evaluated with clinical scores and faecal calprotectin (fCal) measured at week twelve after IFX initiation. For fCal measurement a quantitative enzyme immunoassay (PhiCal Test, Calpro AS, Oslo, Norway) was applied and fCal values under 100 µg/g were considered as normal [25]. In UC patients the Partial Mayo Score (PMS) and in CD patients the modified Harvey-Bradshaw -index (mHBI, no abdominal palpation) was applied. In UC patients a PMS of ≤ 2 combined with a normal fCal value was defined as R. PR was determined as a decrease of PMS ≥ 3 points from baseline [26]. R of CD patients was defined as an mHBI ≤ 4 points and a normal fCal, and PR as an mHBI reduction of ≥ 3 points from baseline [27, 28]. If treatment was discontinued due to surgery other than colectomy, or other reason, treatment outcome at week twelve was not assessed and the patient's data not included in the bacteriome and mycobiota analysis.

2.3. ASCA IgG/IgA ELISA assay

The methods are described in Supplementary Methods 1.

2.4. Faecal DNA extraction

A total of 297 faecal samples were collected from 72 IBD patients and transported to the research facilities within approximately 8 hours. The samples were stored at -80 ˚C until the DNA was extracted from the samples using the repeated bead beating (RBB) method optimized for faecal DNA extraction as previously described [29].

2.5. Library preparation

The bacterial composition was analysed by Illumina MiSeq sequencing of the hypervariable V3-V4 region of the 16S rRNA gene with primers 341FWD 5′-CCTACGGGNGGCWGCAG-3′ and 785REV 5′-GACTACHVGGGTATCTAATCC-3′ in two separate runs for a total of 297 samples from 72 IBD patients. The library was prepared as previously described [30]. **For the fungal composition analysis,** DNA was amplified in a separate reaction with the PCR primer pair ITS1F (FWD, CTTGGTCATTTAGAGGAAGTAA) and ITS2 (REV, GCTGCGTTCTTCATCGATGC), which target the conserved ITS1 region of fungal DNA [31]. The libraries were prepared in a three-step PCR, described in Supplementary Methods 2, for 285 samples of 71 IBD patients. Illumina MiSeq paired-end sequencing was performed for both 16S and ITS in Functional Genomics Unit, University of Helsinki, Helsinki, Finland, in separate runs. The 16S and ITS rDNA amplicon sequences are deposited in the European Nucleotide Archive (ENA). Other data that support the findings of this study are available from the corresponding authors upon reasonable request.

2.6. Analysis of sequencing data

For bacterial composition, the 16S rRNA amplicon sequencing reads were analysed using the R package mare [32]. The median number reads obtained per sample was 34499 (range 185 - 87774) for the first MiSeq run and 55753 (range 4325 - 124080) for the second MiSeq run. Data from both separate runs were combined and analysed together. The processing was done using the ProcessReads function in the mare package using the default parameters. Only forward reads truncated to 150 bases were used and reads below the abundance of 0.000015% were discarded. After pre-processing the median number of reads obtained per sample was 29361 (range 105 - 65089) in combined data. Taxonomic annotation was conducted using USEARCH [33] by mapping the reads to the SILVA 16S rRNA reference database version 115 [34] including only gut-associated taxa. Diversity was measured as the inverse Simpson diversity index and richness

as the number of operational taxonomic units (OTUs; reads clustered at 97% similarity). Samples with <2000 reads (N = 5) were not included in the analyses. There was an association between the read count and microbiota richness (r = 0.377, Pearson, p < *0.0001*) and hence all statistical analyses were performed using the number of reads per sample as the offset. **For fungal composition**, the MiSeq sequencing data was pre-processed using DADA2 version 1.12.1 [35]. Both forward and reverse reads were included in the analysis. The total number of ITS reads per sample before pre-processing was at an average of 54 700 reads. The processing was done according to the DADA2 pipeline for ITS processing with a few exceptions: due to low quality of reverse reads, Ns were not removed from the reads prior to removing the primers by using Cutadapt version 2.10 [36]. Additionally, parameters in the filterAndTrim function were set to maxEE = 4, maxN = 0 and minLen = 100. After pre-processing ITS reads were received for 284 out of the 285 samples from the study. The amplicon sequence variants (ASVs) were annotated with BLAST [37], using the nt database, with "Fungi" as the organism specific filter. Out of these samples, 195 samples (68%) were successfully annotated with the read count being at an average of 1900 reads for successfully annotated taxa. Finally, the R package mare [32] was used for analysis and visualization. There was an association between the read count and mycobiota richness (r = 0.188, Pearson, p = *0.008*). Species level annotations, included in brackets, were the most likely hits based on BLAST [37] or SILVA [34] annotations.

2.7. Statistical analysis

Statistical analysis for the relative abundance data of bacteria and fungi on all taxonomical levels was performed in R with package mare [23] with tools from packages vegan [38], MASS [39] and nlme [40]. P-values for taxon-specific differences were corrected for false discovery rate (FDR; Benjamini–Hochberg [41]). Associations between the overall microbiota composition and background variables were studied using principal coordinates analysis (PCoA), with R package

vegan [38], and multivariate permutational analysis of variance (PERMANOVA). The data-based selection of confounders was based on PERMANOVA and factors that were statistically significantly associated with the variation in the bacterial and fungal taxa were used as confounders in respective analyses. For both bacterial and fungal data, generalised linear models with negative binomial distribution (glm.nb) from the MASS package [39] and Generalized Least Squares (gls) from the nlme package [40] were used for analysing differences in response groups and IBD subtypes. In all 16S analyses age, sex, smoking and use of medication were used as confounders and 0.1 was used as min prevalence and 0.01 as min abundance. In all ITS analyses age, sex, smoking, IBD subtype and use of the drugs mesalazine, mercaptopurine and azathioprine were used as confounders and 500 was used as read count cut-off. In ITS analysis 0.1 was used as min prevalence and 0.01 as min abundance. When assessing the effect of medication on the bacteriome, the use of medicines was categorized into 10 categories based on the medication usage in patients (Supplementary Methods 3). A set of negative control samples containing only PCR reagents were included and sequenced together with the samples as described in the Supplementary Methods 4.

Groupwise comparisons of univariate data (e.g. richness, diversity and read counts) were conducted with GraphPad Prism 7.0 (GraphPad Software, Inc., San Diego, CA). Parametric tests (Anova, parametric t-test) were used for normally distributed values and non-parametric tests (Mann-Whitney U-test) for values not normally distributed or low in numbers.

PathModel function in mare [23] was used to identify the bacterial genera that differed significantly between response groups and find the ideal glm model [39] to fit the data. Age, sex, smoking, IBD subtype and use of the drugs mesalazine, mercaptopurine and azathioprine were used as confounding variables in the PathModel function. For CD patients the predictive power was further investigated by constructing an initial model using the bacterial taxa identified as significantly different between response groups in this article and reducing it stepwise using

Akaike Information Criterion (AIC) [42]. The performance of the final models was tested and visualized with receiver operating characteristic (ROC) analysis using the pROC package [43]. The bacterial genera used within this model are listed in Supplementary Methods 5.

Finally, the correlation between bacterial and fungal genera was studied by calculating Spearman correlations and p values. Phyloseq objects were created from both the 16S and the ITS data separately [44]. Filtering was done by first summarizing the unique genera based on OTUs and including only the taxa with an abundance of over 100 reads.

2.8. Ethics statement

The study was approved by the Ethics Committee of the Helsinki University Central Hospital (147/13/03/01/16) and was registered in European Union Drug Regulating Authorities Clinical Trials Database (EUDRA-CT-Number: 2016-001278-13). The patients signed an informed consent form to participate in the present study before the start of IFX therapy.

3. Results

3.1. Patient characteristics

Altogether, 75 patients (26 CD and 49 UC) were recruited for the study, but three patients withdrew consent before data collection was initiated and were excluded. Data of 72 patients were included in baseline characteristics, presented in Table 1. For data analysis, IBDU patients (n=2) were combined with the UC group.

Two patients terminated treatment before week twelve due to colectomy. These patients were regarded as NRs and their data were included in the bacteriome and mycobiota analysis. One patient underwent planned upper gastrointestinal tract surgery, and one patient developed an allergic reaction at the second infusion due to immunization and IFX was discontinued. Since the response for these patients was not determined, both patients' data were excluded from analyses predicting IFX response. Data of 68 patients were available for evaluation of treatment outcome at week twelve. Endoscopy was performed in 59 patients (17 CD patients, 42 UC/IBDU). In nine patients (6 CD, 3 UC) treatment response was evaluated by clinical scores and fCal, if available (see Figure 1).

3.1.2. Response to IFX therapy

IFX response was evaluated in 70 patients after twelve weeks of IFX treatment. Out of the 70 patients, 44 (62.9%) were Rs at week twelve. Twelve patients (17.1%) were PRs to IFX-treatment and 14 patients (20.0%) were considered NRs as presented in Figure 1. The individual responses of each patient to IFX treatment at week twelve are shown in Supplementary Table 1 (ST1). During the follow-up time of one year 15 patients discontinued IFX due to loss of response or other reasons like pregnancy.

3.2. Composition, diversity and richness of gut microbiota in IBD patients

3.2.1. Overview of faecal bacterial composition

Faecal DNA sequencing was used to determine the gut bacterial community composition among patients in the IFX treatment cohort. Firmicutes (68.0%), Bacteroides (15.5%), Actinobacteria (9.9%) and Proteobacteria (5.8%) were the main phyla making up 99.1% of the bacteriome. We detected 105 bacterial genera in IBD patients. The most abundant bacterial genera were an unknown genus of *Lachnospiracea*, *Faecalibacterium*, *Bacteroides*, *Bifidobacterium*, *Subdoligranulum*, *Blautia*, *Roseburia* and *Enterobacter*.

3.2.2. Overview of faecal fungal composition

DNA sequencing was also used to determine the composition of the gut fungal communities among patients in the IFX treatment cohort. The samples that did not annotate successfully to any gut or environment-specific taxa had lower read counts before pre-processing compared to the ones that did annotate successfully. Additionally, there was an association between the samples that did not annotate successfully and bacteriome samples with lower read counts ($p = 0.01$). The gut mycobiota in the IBD patients was composed of phyla Basidiomycota and Ascomycota, of which Ascomycota was more abundant (73.0% of the mycobiota, Basidiomycota 18.4% and the rest were uncultured Eukaryota). In the samples, 48 different genera were observed, of which *Candida* was the most abundant (25.6% of the mycobiota), followed by *Clavispora* (10.3%) and uncultured *Galactomyces* (7.7%). The prevalence of *Candida* was 40%, *Clavispora* 20% and uncultured *Galactomyces* 12%. The prevalence of all genera is presented in ST2. For reliable statistical analyses, only the genera with a prevalence above 10% were included. The fungal faecal composition is presented at genus level for each sample at different timepoints in Supplementary Figures S1-5.

3.2.3. Diversity and Richness

3.2.3.1. Diversity and richness in gut bacteriome

The bacterial diversity ranged from 1.0 to 28.2 (median 9.0) and the richness varied between 33 and 192 (median 106) among all samples. There was a trend with the diversity being higher in UC compared to CD at baseline and two, twelve and 52 weeks of treatment although no statistical significance was reached. Bacteriome richness did not differ at any timepoint between CD and UC. The richness was lower in samples among patient who underwent surgery ($p = 0.0004$) compared to patients who completed the study. When the patients were divided into groups stratified by IFX response, the diversity was higher in Rs compared to NRs at six weeks ($p = 0.01$, Mann-Whitney, Supplementary Figure S6). In addition, a similar trend was seen at baseline ($p = 0.06$) and at other timepoints as well, although not significant. Richness did not differ significantly between response groups during the study at any timepoint (Supplementary Figure S7).

3.2.3.2. Diversity and richness in gut mycobiota

The fungal diversity ranged from 1.0 to 4.4 (median 1.0) and the richness varied between 1 and 10 (median 2.0) among all samples. No significant differences were observed in fungal diversity ($p = 0.2$, MW) or richness ($p = 0.2$, MW) between IBD subtypes at baseline. A trend with diversity being lower in NRs compared to Rs was observed at all timepoints except for the sixth week timepoint, although not significant (Supplementary Figure S8). The same trend between response groups was also present in richness (Supplementary Figure S9) although not statistically significant.

3.3. IBD subtypes and gut microbiota composition

Principal coordinates analysis (PCoA) based on Bray-Curtis dissimilarity of all samples revealed that the gut bacterial communities clustered according to the diagnosis (3% of variation explained by diagnoses, p = *0.001*, Supplementary Figure S10). Therefore, IBD subtypes were included as a confounding variable in later analyses. The composition of the main bacterial genera in all IBD patients, in CD and in UC is shown in Figure 2 (A-C, respectively). Before initiation of IFX therapy it was observed that the abundance was significantly higher in CD compared to UC for the class *Bacilli* and the order Lactobacillales by eight-fold (FDR-corrected p (p FDR) = *0.005*, glm.nb), the family *Streptococcaceae* by five-fold (p FDR = *0.03*, glm.nb) and the genus *Desulfovibrio* by thirteen-fold (p = *0.005*, glm.nb), and lower for the class Clostridia and the order Clostridiales by one-fold (p < *0.0001*, glm.nb) and the family *Ruminococcaceae* by one-fold (p < *0.0001*, glm.nb).

The gut fungal composition in the IBD subtypes was analysed before the initiation of IFX therapy and it was observed that the abundance of Ascomycata was significantly higher by one-fold (p FDR < *0.0001*, glm.nb) in CD compared to UC. Since significant differences were observed between CD and UC, it was considered in further analysis.

3.4. Influence of IFX treatment on bacterial composition

We investigated the influence of the response to IFX therapy on composition of gut bacteria in IBD subtypes. Principal coordinates analysis (PCoA) based on Bray-Curtis dissimilarity of all samples revealed that the gut bacterial communities clustered according to the IFX response in all patients (2% of variation explained by IFX response, p = *0.001*), in CD (7%, p = *0.001*) and in UC (2%, p = *0.02*, Supplementary Figure S11) suggesting that the microbiota composition varies more between response groups in CD than in UC. Our main focus was on comparing the differences between NRs and Rs.

Gut bacterial composition showed differences at family, order and genus levels at different time points during the study between the IFX response groups in CD, UC and all IBD patients combined. Differences at family level in all IBD patients are shown in Figure 3 as well as differences at family level in CD and UC (in Supplementary Figure S12). CD patients had more consistent and numerous differences in taxa between the response groups. The relative abundances of several genera such as *Odoribacter*, *Alistipes*, *Butyricimonas* and *Anaerofilum* were lower in NRs, while *Parasutterella*, *Haemophilus* and *Veillonella* were higher in NRs compared to Rs at several timepoints in all IBD patients. There were fewer changes in several taxa between timepoints during the study in Rs compared to PRs and NRs suggesting more stable bacterial composition in Rs. Similarly, in CD the most consistent findings at genus level were lower relative abundances of *Alistipes* and *Butyricimonas* and higher relative abundances of *Veillonella* in NRs compared to Rs. In addition, the CD patients had lower relative abundances of *Bifidobacterium*, *Barnesiella*, *Enterobacter* and *Phascolarctobacteria* as well as higher relative abundances of *Streptococcus* in NRs compared to Rs. In UC, clear and uniform differences between response groups were not as numerous as in CD but were seen at genus level in lower relative abundances of *Alistipes*, *Anaerococcus* and *Odoribacter* and higher abundances of uncultured *Prevotella* and *Sutterella*.

We compared the changes in bacterial composition between response groups over time. No significant (FDR-corrected) differences in changes of taxa were seen between NRs and Rs in all IBD patients or in UC. In CD, Enterobacteriales were increased in Rs and decreased in NRs and this change was significantly different (p FDR < *0.0001*) between the baseline and two weeks timepoint. No other changes after FDR-correction were seen but based on raw p-values some taxa were changed significantly during the study in both CD and UC. PR had more differences compared to NR than R in general among all patients.

3.5. Prediction of IFX response using microbial profiles at baseline

We explored whether the bacterial and fungal gut microbiota differed between response groups at baseline. We observed clear differences between IFX response groups among all patients and in CD and UC (Figure 4). In our analysis, we wanted to focus on comparing NRs to Rs.

NRs had less Clostridia (by one-fold, p FDR < *0.0001*), of which *Ruminococcaceae* (by one-fold, p FDR < *0.0001*) and an unknown genus of *Ruminococcaceae* (by two-fold, p FDR < *0.0001*) were significantly decreased at family and genus level, respectively (ST3). Differences were seen at family level in *Carnobacteriaceae, Peptostreptococcaceae, Ruminococcaceae* and *Enterobacteriaceae* between NRs and Rs. In addition, *Odoribacter* and unknown *Ruminococcacea* were significantly increased in Rs compared to NRs and *Granulicatella, Enterobacter* and an unknown genus of *Peptostreptococcae* were increased in NRs compared to Rs. In CD patients Bacteroidetes were elevated in Rs (by nine-fold, p FDR = *0.03*), while Firmicutes were elevated in NRs (by two-fold, p FDR = *0.02*) (ST4). In UC patients, Bacteroidetes was elevated in NRs (by one-fold, p FDR < *0.0001*) at baseline (ST5).

In CD patients, the orders Bifidobacteriales Micrococcales, Lactobacillales, Burkholderiales and Pseudomonales were significantly more abundant in NRs, while Bacteroidales and Desulfovibrionales were significantly elevated in Rs. Thirteen genera and eight families differed significantly (by FDR-corrected p-values) between response groups (Figure 4A-D). In UC patients, the order Bacteroidales, families *Enterococcae*, *Clostridiaceae*, *Peptostreptococcaceae* and *Ruminococcaceae* and seven genera differed significantly (by FDR-corrected p-values) between groups NR and R (Figure 4A-D).

The genus *Candida* (*C. albicans*) was significantly more abundant in NRs compared to Rs by two-fold (p FDR < *0.0001*, ST6) before the initiation of IFX therapy and remained more abundant in NRs among all IBD patients at week two, six and 52. *Candida* (*C. albicans*) was more abundant also in both CD and UC patients in the group of NRs at baseline. Similarly, Ascomycota was elevated in both CD and UC at baseline by one-fold (p FDR < *0.0001*, ST7, ST8). The ratio between relative abundances of Ascomycota and Basidiomycota was calculated at all timepoints, but there were no significant differences between the response groups.

The response to IFX therapy was predictable with high AUC values (Figure 5A-C) before start of treatment. When ROC analyses were done using PathModel function to select predictive genera the AUC value was 0.797 for all IBD patients (Figure 5A), 0.842 in CD patients and 0.791 in UC patients (Supplementary Figure S13). The AUC value rose to 0.933 in CD patients (Figure 5B) and to 0.818 in UC patients (Figure 5C) when using the genera we found to differ between the response groups (see Supplementary Methods 5). In UC ***Candida*** was included in the ROC analysis in addition to the bacterial genera.

3.6. Correlation between bacteria and fungi

We performed Spearman correlation analyses between fungal and bacterial genera at baseline to investigate the cross-kingdom relationships in the microbiota in relation to IFX therapy response. Multiple significant correlations were observed between fungal and bacterial genera and the correlations differed between response groups (ST9-11).

In Rs at baseline, *Clavispora* correlated negatively with four bacterial genera (ST9). Uncultured *Galactomyces* correlated negatively with *Bilophila* (r = -0.48, p = *0.01*, ST9) and positively with *Citrobacter* (r = 0.42, p = *0.03*, ST9). Uncultured *Saccharomyces* correlated negatively with three bacterial genera (ST9). In PRs at baseline *Candida* correlated negatively with *Bacteroides* (r = -

0.76, p = *0.02*, ST10), *Clostridium* (r = -0.78, p = *0.01*, ST10), *Coprococcus* (r = -0.73, p = *0.02*, ST10), *Dorea* (r = -0.87, p = *0.002*, ST10), *Pseudomonas* (r = -0.70, p = *0.04*, ST10) and *Roseburia* (r = -0.83, p = *0.006*, ST10). Additionally, *Gelatoporia* correlated positively (r = 1.0, p < *0.001*) with *Christensenella* in PRs at baseline. In NRs, *Candida* correlated positively with *Klebsiella* (r = 0.83, p = *0.02*, ST11) and *Lactococcus* (r = 0.94, p = *0.002*, ST11). Additionally, uncultured *Saccharomyces* correlated with eleven bacterial genera (see ST11).

3.7. ASCA IgG/IgA ELISA assay results

IgG and IgA values were more elevated in CD patients compared to UC (p < *0.0001*). No correlation was observed between ASCA values and response to IFX therapy. Neither IgG nor IgA values correlated with the % relative abundance of *Saccharomyces* or *Candida* at baseline.

4. Discussion

We have investigated the faecal bacterial and fungal compositions in a prospective IBD cohort treated with the TNF-α blocker IFX and followed for one year. The gut bacterial compositions in IBD patients have previously been investigated, but only a few studies have explored the fungal composition thus far. We identified significant differences in the gut bacteriome and mycobiota in patients depending on their response to IFX treatment already before the initiation of therapy, also indicated by high predictive power. During the therapy, bacterial composition was found to be more stable in responders compared to partial or non-responders. Additionally, responders and non-responders showed differences in correlations between bacterial and fungal genera before start of IFX treatment. These findings support the suggestion that bacterial and fungal microbiota profiles could provide predictive markers for response to IFX therapy.

The bacteriome differed between CD and UC patients and bacterial diversity was higher in UC compared to CD, as previously found [45, 46]. The composition at phylum level also agreed with previous studies with Firmicutes, Bacteroides, Actinobacteria and Proteobacteria being the four main phyla representing 99.1% of all phyla [47, 48]. The detected fungi of the gut mycobiota in IBD patients belonged to the phyla Ascomycota and Basidiomycota, as has previously been observed [49, 50]. The most common genera in the human gut mycobiota are *Saccharomyces* and *Candida* [49, 52, 53], while *Candida* was the most abundant in this study. The prevalence of genera was consistently low in our study, and therefore only genera with a prevalence above 10% were included. Others have reported similar issues [53] and this should be considered when interpreting the results. It has been reported that *Candida* is more abundant in IBD patients compared to healthy controls [54, 55] while *Saccharomyces* species are decreased in IBD patients [46]. The mycobiota is less stable than the bacteriome and varies over time [56].

Our results demonstrate significant differences in bacterial taxa between IFX treatment response groups before initiation of IFX treatment. Several studies have previously explored the use of microbial markers in predicting (TNF-α) treatment relapse in patients with CD or UC, both in paediatric patients [17, 55, 57, 58] and in adults [16, 59, 60]. As reviewed by Ananthakrishnan [61], the patient number has been limited in some studies, which might hamper significant findings. The evaluation of response, microbiota analysis and whether the study is performed with control subjects vary and introduce difficulties when comparing results [61].

Zhou *et al*. demonstrated an increase in relative abundance of Clostridiales in responders compared to those who relapsed in a study with 16 CD patients and that Clostridiales could predict treatment effectiveness in combination with calprotectin [16]. Similarly, Clostridiales was more abundant in Rs compared to NRs in our study. The Clostridiales order and its families *Lachnospiraceae* and *Ruminococcaceae,* several of which are butyrate producing organisms, represent the most dominant group of bacteria in the healthy human gut. Butyrate and butyrate-producing bacteria have been shown to be elevated in IBD patients in response to TNF-α treatment, and even presented as possible markers for response [59]. We observed a lower abundance of *Odoribacter* and unknown *Ruminococcaceae* in NRs compared to Rs before start of IFX therapy. These genera are SCFA producers [62, 63] and have been shown to produce butyrate [62, 64] and have also previously been reported to be decreased in NRs to IFX [17, 19, 65]. Wang *et al*. reported that *Lachnospira, Ruminococcus, Sutterella, Roseburia* and *Bilophila* are decreased in paediatric CD non-responders to IFX [17]. Additionally, *Granulicatella* [19], *Bifidobacterium* [58], *Bacteroides* [16, 65], *Enterobacter* [58], *Sutterella* [19, 66], *Clostridium* [19], unknown *Lachnospiraceae* [19, 60], *Alistipes* [17], *Enterococcus* [17] and *Faecalibacterium* [17, 18, 19] have been reported to associate with IFX response, as in our study. Further, the genus *Rothia*, which was increased in CD NRs, has previously been shown to be associated with disease progression [67] and *Dialister*,

which was increased in UC Rs, has been shown to predict reduction in dysbiosis [55]. *Clostridium* has been observed to be elevated in responders to IFX [19], but in our study it was more abundant in non-responders compared to responders.

The possibility to predict IFX therapy response from the gut microbiota was further analysed by ROC curves, which indicated the performance of the models created from a list of predictive genera. The ROC analyses were performed using bacterial genera that significantly explained the difference between response groups. Out of these *Bifidobacterium, Sutterella, Enterococcus, Enterobacter, Alistipes, Clostridium* and *Faecalibacterium* have previously been found to associate with IFX response [16, 17, 19, 58, 60, 66]. Additionally, *Candida* was included in UC ROC, indicating its predictive value. These analyses demonstrated high predictive power for IFX response.

Only few studies have investigated the possible markers of response to IFX therapy in the gut mycobiota and, to our knowledge, none have studied the response in gut microbiota of adult CD and UC patients. We found that *Candida* (*C. albicans*) was more abundant in NRs at baseline in both CD and UC patients, in line with its possible dysbiosis-driving properties. Although *Candida* colonizes the healthy gut, it has been observed that the host regulates its growth [68, 69]. In a recent study done in type 1 diabetes patients, *Candida* and *Saccharomyces* were observed as elevated in patients in whom the disease progressed [53]. In addition, it has been reported that IBD patients are characterized by elevated *C. albicans* levels and reduced *Saccharomyces cerevisiae* levels, further highlighting the dysbiotic properties of *C. albicans* [49]. In a reduction of SCFA producers caused by antibiotics, there was an increased abundance of *C. albicans* [70] which is consistent with our results that NRs have lower abundance of SCFA (namely butyrate) producers and increased abundance of *C. albicans* compared to Rs. Interestingly, some fungal species, primarily *Saccharomyces boulardii,* is used for probiotic purposes in treating diarrhoea [71, 72,

73]. *Saccharomyces* has been observed to be decreased during IBD flare, indicating that it protects against disease [49]. These contradicting observations reveal the difficulty in determining whether a fungal organism has a probiotic or dysbiotic property and highlight their species-specific properties and effects on the host. We also observed a clear trend with lower diversity in NRs compared to Rs in both the gut bacteriome and mycobiota, although not significant at all timepoints. Consistent with earlier findings, a lower diversity signals a more inflamed gut [74].

IBD has successfully been treated by faecal microbiota transplantation [75, 76]. Both anti-fungal treatments and fungal probiotics have been studied in treating active IBD with successful results highlighting the potential in such treatments, as reviewed by Lam *et al*. [77]. There is also supportive evidence that diet affects the gut mycobiota, with carbohydrate-rich diets causing elevation in fungal species [52, 78] and protein-rich diets the contrary [52]. These findings support the suggestion that the microbiota can affect treatment outcome.

The bacterial and fungal communities in the gut microbiota interact [49, 69, 79]. Following antibiotic treatment in mice, there was a significant increase in the growth of fungi [80] suggesting that the decrease of bacteria causes an increase in fungal growth. We investigated this in our study by calculating Spearman correlations between the kingdoms. Interestingly, we observed differences in correlations between fungal and bacterial genera in Rs, PRs and NRs. In Rs, there was no significant correlation between *Candida* and bacterial genera. We observed correlations between bacterial genera and *Clavispora*, uncultured *Galactomyces* and uncultured *Saccharomyces*, although the low r values indicated only weak to moderate correlation. In NRs, *Candida* correlated positively with *Lactococcus,* a bacterial genus that was elevated in responders to IFX [61] and is considered to have probiotic properties [81], and *Klebsiella*, a genus previously associated with non-response to IFX [17, 55, 61]. In PRs, *Candida* correlated negatively with *Pseudomonas,* which is considered a driver of dysbiosis [82, 83], and *Clostridium* was elevated in

CD and UC NRs. All the other bacterial genera that correlated negatively with *Candida* in PRs at baseline, namely *Bacteroides*, *Coprococcus*, *Dorea* and *Roseburia*, have previously been associated with low abundance in non-responders to IFX therapy [17, 18, 55]. The correlations reveal the possibility that not only bacterial and fungal organisms alone, but also their interplay can be part of predicting IFX response.

Other potential predictive biomarkers for IFX response, besides the microbiota, have also been investigated. Recently some genetic and protein-based biomarkers, e.g., TREM1, ZNF133 and Oncostatin M have been identified to associate with IFX response in IBD patients [84, 85, 86]. Biomarkers predicting treatment response are urgently required to select optimal medical treatment for IBD patients as about one third are NRs to TNF-α blockers [87]. Interestingly, our prediction of response calculated by ROC curve analyses was stronger than that of some other biomarkers, particularly for CD patients [84, 85, 88], although our results need to be validated in a larger cohort.

The anti-*Saccharomyces cerevisiae* antibody (ASCA) is an antibody against oligomannosides in the cell wall of bakers' yeast [89] *S. cerevisiae*, which is present in the human gut mycobiota. It is used as a biomarker for CD and has been observed to correlate with both *S. cerevisiae* and species of the genus *Candida* [79, 90]. Here we studied ASCA values as a possible marker for response. ASCA values measured from baseline serum samples did not correlate with response groups. The proportion of primary NRs to IFX induction in our study was similar to the 10-20% of NRs reported in earlier studies, and the overall response rate of 80% is comparable to the results of other studies [10, 91, 92]. The higher remission rate of 63% in our study can be explained with the combination of endoscopically and clinically assessed treatment groups as endoscopic remission rates after TNF-α blocker induction are reported to be around 30% in CD and 45% in UC patients [93].

The strengths of our study are the longitudinal collection of faecal samples and high endoscopy rate of 87% at week twelve that provide an objective assessment of disease activity and thus a reliable classification of response groups. In previously published studies investigating gut microbiota as a predictor for IFX response, disease activity was mainly evaluated clinically or with surrogate parameters like CRP or fCal [18, 20]. We acknowledge that the lack of baseline endoscopic data and incomplete endoscopy at week twelve are limitations, as endoscopic disease activity and clinical symptoms often poorly correlate [94, 95]. The limited number of patients, especially CD patients, and the lack of mucosal-wash sampling at the time of endoscopy were additional limitations of the study, which might give rise to an issue with validation. For the ROC curve analysis, our strength is that around half of the predictive genera used were previously found to predict IFX therapy, while the limitation is that not all the genera were previously published to predict IFX response. Additionally, a validation cohort would be required for investigating whether the results can be used for prediction of IFX response in IBD patients.

In conclusion, we found significant differences between response groups to IFX therapy in IBD patients. Non-responders to IFX therapy had lower abundance of butyrate-producing bacteria, particularly Clostridiales, and a higher abundance of *Candida*, also indicated by high predictive power. These results further strengthen previously published results that the gut microbiota could provide promising biomarkers for IFX therapy response prediction in the future.

Acknowledgements

We would like to acknowledge the excellent assistance by study nurses Pirkko Tuukkala ja Virpi Pelkonen. Additionally, we thank Hanne Ahola for handling the samples, Tinja Kanerva for assisting in DNA extraction and library preparation and Pinja Elomaa for assisting in DNA extraction. Finally, we will thank Heli Pessa for thoroughly reading through and revising the manuscript.

Funding

This work was supported by Finnish Funding Agency for Innovation (Tekes) through the SalWe GET IT DONE Personalized Diagnostics and Care program (PS), Sigrid Juselius Foundation (PS) and Foundation of Mary and Georg C. Ehrnrooth (TS).

Conflicts of interest

TS has received speaker or consultant fees from Abbvie, Ferring, Janssen-Cilag, Pfizer, Takeda, and Tillotts Pharma in addition to research grants from Janssen-Cilag and Takeda. AE has received speaker or consultant fees from Janssen-Cilag, Pfizer and Takeda. The authors (RVH, EN, SS, SV, KK, AS, PS) declare that there is no conflict of interest.

Author Contributions

RVH, EN: designing the study, DNA extraction, library preparation, data analysis and interpretation and drafting the manuscript. AE: designing the study, patient recruitment, evaluating IFX response and drafting the manuscript. SS, SV, KK, AS: data analysis and revising the article. KK, AS: assistance in study design and interpretation of results. PS, TS: study design, revising the article and reviewing the manuscript. All authors approved the final version of the manuscript. We did not use any writing assistance.

Supplementary Data

Supplementary are available at ECCO-JCC online.

5. References

1. Baumgart D C, Carding S R. Inflammatory bowel disease: cause and immunobiology. *Lancet* 2007;369:1627–1640. DOI:10.1016/S0140-6736(07)60750-8

2. Yu Y R, Rodriguez J R. Clinical presentation of Crohn's, ulcerative colitis, and indeterminate colitis: Symptoms, extraintestinal manifestations, and disease phenotypes. *Semin Pediatr Surg* 2017;26:349–355. DOI:10.1053/j.sempedsurg.2017.10.003

3. Kostic A D, Xavier R J, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 2014;146:1489–1499. DOI:10.1053/j.gastro.2014.02.009

4. Ni J, Wu G D, Albenberg L, Tomov V T. Gut microbiota and IBD: causation or correlation?. *Nat Rev Gastroenterol hepatol* 2017;14:573–584. DOI:10.1038/nrgastro.2017.88

5. Jostins L, Ripke S, Weersma R K, *et al*. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119–124. DOI:10.1038/nature11582

6. Nishida A, Inoue R, Inatomi O, *et al*. Gut microbiota in the pathogenesis of inflammatory bowel disease. *Clin J Gastroenterol* 2018;11:1–10. DOI:10.1007/s12328-017-0813-5

7. Galloway-Peña J R, Kontoyiannis D P. The gut mycobiome: The overlooked constituent of clinical outcomes and treatment complications in patients with cancer and other immunosuppressive conditions. *PLoS pathogens* 2020;16:e1008353. DOI:10.1371/journal.ppat.1008353

8. Ardizzone S, Cassinotti A, Manes G, Porro G B. Immunomodulators for all patients with inflammatory bowel disease? *Therap Adv Gastroenterol* 2010;3:31–42. DOI:10.1177/1756283X09354136

9. Borody T J, Eslick G D, Clancy R L. Fecal microbiota transplantation as a new therapy: from Clostridioides difficile infection to inflammatory bowel disease, irritable bowel syndrome, and colon cancer. *Curr Opin Pharmacol* 2019;49:43–51. DOI:10.1016/j.coph.2019.04.017.

10. Rutgeerts P, Sandborn W J, Feagan B G, *et al*. Infliximab for induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2005;353:2462–2476. DOI:10.1056/NEJMoa050516

11. Murch S H, Lamkin V A, Savage M O, *et al*. Serum concentrations of tumour necrosis factor in childhood chronic inflammatory bowel disease. *Gut* 1991;32:913–917. DOI:10.1136/gut.32.8.913

12. Knight D M, Trinh H, Le J, *et al*. Construction and initial characterization of a mouse-human chimeric anti-TNF antibody. *Mol Immunol* 1993;30:1443–1453. DOI:10.1016/0161-5890(93)90106-l

13. Nielsen O H, Ainsworth M A. Tumor necrosis factor inhibitors for inflammatory bowel disease. *N Engl J Med* 2013;369:754–762. DOI:10.1056/NEJMct1209614

14. Rosenblum H, Amital H. Anti-TNF therapy: Safety aspects of taking the risk. *Autoimmun. Rev* 2011;10:563–568. DOI:10.1016/j.autrev.2011.04.010

15. Gaujoux R, Starosvetsky E, Maimon N, *et al*. Cell-centred meta-analysis reveals baseline predictors of anti-TNFα non-response in biopsy and blood of patients with IBD. *Gut* 2019;68:604–614. DOI:10.1136/gutjnl-2017-315494

16. Zhou Y, Xu Z Z, He Y, *et al*. Gut microbiota offers universal biomarkers across ethnicity in inflammatory bowel disease diagnosis and infliximab response prediction. *mSystems* 2018;30:3e00188-17. DOI:10.1128/mSystems.00188-17

17. Wang Y, Gao X, Ghozlane A, *et al*. Characteristics of faecal microbiota in paediatric Crohn's disease and their dynamic changes during infliximab therapy. *J Crohns Colitis* 2018;12337–346. DOI:10.1093/ecco-jcc/jjx153

18. Rajca S, Grondin V, Louis E, *et al*. Alterations in the intestinal microbiome (dysbiosis) as a predictor of relapse after infliximab withdrawal in Crohn's disease. *Inflamm Bowel Dis* 2014;20:978–986. DOI:10.1097/MIB.0000000000000036

19. Dovrolis N, Michalopoulos G, Theodoropoulos G E, *et al*. The interplay between mucosal microbiota composition and host gene-expression is linked with infliximab response in inflammatory bowel diseases. *Microorganisms* 2020;8:438. DOI:10.3390/microorganisms8030438.

20. Ding N S, McDonald J A K, Perdones-Montero A, *et al*. Metabonomics and the gut microbiome associated with primary response to anti-TNF therapy in Crohn's disease. *J Crohns Colitis* 2020;jjaa039. DOI:10.1093/ecco-jcc/jjaa039

21. Schroeder K W, Tremaine W J, Ilstrup D M. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N Engl J Med* 1987;317:1625–1629. DOI:10.1056/NEJM198712243172603

22. Daperno M, D'Haens G, Van Assche G, *et al*. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. *Gastrointest Endosc* 2004;60:505–512. DOI:10.1016/s0016-5107(04)01878-4

23. Sipponen T, Nuutinen H, Turunen U, Färkkilä M. Endoscopic evaluation of Crohn's disease activity: comparison of the CDEIS and the SES-CD. *Inflamm Bowel Dis* 2010;16:2131–2136. DOI:10.1002/ibd.21300

24. Moskovitz D N, Daperno M, Van Assche G. Defining and validating cut-offs for the Simple Endoscopic Score for Crohn's Disease. *Saudi J Gastroenterol* 2007;132:1097. DOI:10.4103/1319-3767.182455

25. Sipponen T, Kolho KL. Fecal calprotectin in diagnosis and clinical assessment of inflammatory bowel disease. *Scand J Gastronterol* 2015;50:74–80. DOI:10.3109/00365521.2014.987809

26. Lewis J D, Chuai S, Nessel L, *et al*. Use of the noninvasive components of the Mayo score to assess clinical response in ulcerative colitis. *Inflamm Bowel Dis* 2008;14:1660–1666. DOI:10.1002/ibd.20520

27. Best W R. Predicting the Crohn's disease activity index from the Harvey-Bradshaw Index. *Inflamm Bowel Dis* 2006;12:304–310. DOI:10.1097/01.MIB.0000215091.77492.2a

28. af Bjorkesten C G, Nieminen U, Turunen U, *et al*. Surrogate markers and clinical indices, alone or combined, as indicators for endoscopic remission in anti-TNF-treated luminal Crohn's disease. *Scand J Gastroenterol* 2012;47:528-537. DOI:10.3109/00365521.2012.660542

29. Salonen A, Nikkilä J, Jalanka-Tuovinen J, *et al*. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J Microbiol Methods* 2010;81:127–134. DOI:10.1016/j.mimet.2010.02.007

30. Korpela K, Salonen A, Vepsäläinen O, *et al*. Probiotic supplementation restores normal microbiota composition and function in antibiotic-treated and in caesarean-born infants. *Microbiome* 2018;6:1–11. DOI:10.1186/s40168-018-0567-4

31. White T J, Bruns T, Lee S, Taylor J W. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: A Guide to Methods and Applications* 1990;18:315–322.

32. Korpela K. *Mare: microbiota analysis in R easily. R package version 1.0. 2016.* Accessed October 26, 2020, at https://github.com/katrikorpela/mare

33. Edgar R C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;26:2460–2461. DOI:10.1093/bioinformatics/btq461

34. Quast C, Pruesse E, Yilmaz P, *et al*. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012;4:D590–D596. DOI:10.1093/nar/gks1219

35. Callahan B J, McMurdie P J, Rosen M J, *et al*. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;13:581–583. DOI:10.1038/nmeth.3869

36. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* 2011:ISSN2226–6089. DOI:10.14806/ej.17.1.200

37. Altschul S F, Gish W, Miller W, *et al*. *BLAST* PROGRAMS. Basic local alignment search tool. *J Mol Biol* 1990;215:403-410.

38. Oksanen J, Blanchet F G, Kindt R, *et al*. Package 'vegan'. *Community ecology package, version, 2*. 2013;9:1–295. Accessed October 26, 2020, at cran.r-project.org.

39. Venables W, Ripley B. Modern Applied Statistics with S. *Springer* 2002.

40. Pinheiro J, Bates D, DebRoy S, Sarkar D R. *nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-14.* Accessed October 26, 2020, at cran.r-project.org.

41. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 1995;57:289–300. DOI:10.2307/2346101

43. Burnham KP , Anderson DR . Multimodel inference - understanding AIC and BIC in model selection . *Sociol Methods Res* 2014;33:261–304.

42. Robin X, Turck N, Hainard A, *et al*. pROC: anopen-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12:77. DOI:10.1186/1471-2105-12-77

44. McMurdie P J, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;8:e61217. DOI:10.1371/journal.pone.0061217

45. Alam M T, Amos G C, Murphy A R, *et al*. Microbial imbalance in inflammatory bowel disease patients at different taxonomic levels. *Gut pathog* 2020;12:1–8. DOI:10.1186/s13099-019-0341-6

46. Pascal V, Pozuelo M, Borruel N, *et al*. A microbial signature for Crohn's disease. *Gut* 2017;66:813–822. DOI:10.1136/gutjnl-2016-313235

47. Eckburg P B, Bik E M, Bernstein C N, *et al*. Diversity of the human intestinal microbial flora. *Science* 2005;308:1635–1638. DOI:10.1126/science.1110591

48. Ley R E, Hamady M, Lozupone C, *et al*. Evolution of mammals and their gut microbes. *Science* 2008;320:1647–1651. DOI:10.1126/science.1155725

49. Sokol H, Leducq V, Aschard H, *et al*. Fungal microbiota dysbiosis in IBD. *Gut* 2017;66:1039–1048. DOI:10.1136/gutjnl-2015-310746

50. Ott S J, Kühbacher T, Musfeldt M, *et al*. Fungi and inflammatory bowel diseases: Alterations of composition and diversity. *Scand J Gastroenterol* 2008;43:831–841. DOI:10.1080/00365520801935434

51. Huseyin C E, Rubio R C, O'Sullivan O, *et al*. The fungal frontier: a comparative analysis of methods used in the study of the human gut mycobiome. *Fron Microbiol* 2017;8:1–15. DOI:10.3389/fmicb.2017.01432

52. 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. DOI:10.1371/journal.pone.0066019

53. Honkanen J, Vuorela A, Muthas D, *et al*. Fungal dysbiosis and intestinal inflammation in children with beta-cell autoimmunity. *Front Immunol* 2020;11:468. DOI:10.3389/fimmu.2020.00468

54. Chehoud C, Albenberg L G, Judge C, *et al*. Fungal signature in the gut microbiota of pediatric patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2015;21:1948–1956. DOI:10.1097/MIB.0000000000000454

55. Lewis J D, Chen E Z, Baldassano R N, *et al*. Inflammation, antibiotics, and diet as environmental stressors of the gut microbiome in pediatric Crohn's disease. *Cell Host Microbe* 2015;18:489–500. DOI:10.1016/j.chom.2015.09.008

56. Hallen-Adams H E, Kachman S D, Kim J, *et al*. Fungi inhabiting the healthy human gastrointestinal tract: A diverse and dynamic community. *Fungal Ecol* 2015;15:9–17. DOI:10.1080/21505594.2016.1247140

57. Michail S, Durbin M, Turner D, *et al*. Alterations in the gut microbiome of children with severe ulcerative colitis. *Inflamm Bowel Dis* 2012;18:1799–1808. DOI:10.1002/ibd.22860

58. Kolho K L, Korpela K, Jaakkola T, *et al*. Fecal microbiota in pediatric inflammatory bowel disease and its relation to inflammation. *Am J Gastroenterol* 2015;110:921–930. DOI:10.1038/ajg.2015.149

59. Aden K, Rehman A, Waschina S, *et al*. Metabolic functions of gut microbes associate with efficacy of tumor necrosis factor antagonists in patients with inflammatory bowel diseases. *Gastroenterology* 2019;157:1279–1292.e11. DOI:10.1053/j.gastro.2019.07.025

60. Ribaldone D G, Caviglia G P, Abdulle A, *et al*. Adalimumab therapy improves intestinal dysbiosis in Crohn's disease. *J Clin Med* 2019;8:1646. DOI:10.3390/jcm8101646

61. Ananthakrishnan A N. Microbiome-Based Biomarkers for IBD. *Inflamm Bowel Dis* 2020;izaa071. DOI:10.1093/ibd/izaa071

62. Vital M, Karch A, Pieper D H. Colonic butyrate-producing communities in humans: an overview using omics data. *mSystems* 2017;2:e00130-17. DOI:10.1128/mSystems.00130-17

63. LeBlanc J G, Chain F, Martín R, *et al*. Beneficial effects on host energy metabolism of short-chain fatty acids and vitamins produced by commensal and probiotic bacteria. *Microb Cell Fact* 2017;16:79. DOI:10.1186/s12934-017-0691-z

64. Qin P, Zou Y, Dai Y, *et al*. Characterization a novel butyric acid-producing bacterium Collinsella aerofaciens subsp. shenzhenensis subsp. nov. *Microorganisms* 2019;7:78. DOI:10.3390/microorganisms7030078

65. Doherty M K, Ding T, Koumpouras C, *et al*. Fecal microbiota signatures are associated with response to ustekinumab therapy among Crohn's disease patients. *mBio* 2018;9:e02120–17. DOI:10.1128/mBio.02120-17

66. Hyams J S, Thomas S D, Gotman N, *et al*. Clinical and biological predictors of response to standardised paediatric colitis therapy (PROTECT): a multicentre inception cohort study. *The Lancet* 2019;393:1708–1720. DOI:10.1016/S0140-6736(18)32592-3

67. Kugathasan S, Denson L A, Walters T D, *et al*. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* 2017;389:1710–1718. DOI:10.1016/S0140-6736(17)30317-3

68. Fan D, Coughlin L A, Neubauer M M, *et al*. Activation of HIF-1α and LL-37 by commensal bacteria inhibits Candida albicans colonization. *Nat Med* 2015;21:808–814. DOI:10.1038/nm.3871

69. De Luca A, Zelante T, D'angelo C, *et al*. IL-22 defines a novel immune pathway of antifungal resistance. *Mucosal Immunol* 2010;3:361–373. DOI:10.1038/mi.2010.22

70. Guinan J, Wang S, Hazbun T R, *et al*. Antibiotic-induced decreases in the levels of microbial-derived short-chain fatty acids correlate with increased gastrointestinal colonization of Candida albicans. *Scientific reports* 2019;9:1–11. DOI:10.1038/s41598-019-45467-7

71. McFarland L V. Systematic review and meta-analysis of Saccharomyces boulardii in adult patients. *World J Gastroenterol* 2010;16:2202–2222. DOI:10.3748/wjg.v16.i18.2202

72. McFarland L V, Surawicz C M, Greenberg R N, *et al*. A randomized placebo-controlled trial of Saccharomyces boulardii in combination with standard antibiotics for Clostridium difficile disease. *JAMA* 1994;271:1913–1918.

73. Madoff S E, Urquiaga M, Alonso C D, Kelly C P. Prevention of recurrent Clostridioides difficile infection: A systematic review of randomized controlled trials. *Anaerobe* 2020;61:102098. DOI:10.1016/j.anaerobe.2019.102098

74. Knox N C, Forbes J D, Van Domselaar G, Bernstein C N. The gut microbiome as a target for IBD treatment: are we there yet?. *Curr Treat Options Gastroenterol* 2019;17:115–126. DOI:10.1007/s11938-019-00221-w

75. Seekatz A M, Aas J, Gessert C E, *et al*. Recovery of the gut microbiome following fecal microbiota transplantation. *mBio* 2014;5:e00893–00814. DOI:10.1128/mBio.00893-14

76. Paramsothy S, Kamm M A, Kaakoush N O, *et al*. Multidonor intensive faecal microbiota transplantation for active ulcerative colitis: a randomised placebo-controlled trial. *Lancet* 2017;389:1218–1228. DOI:10.1016/S0140-6736(17)30182-4

77. Lam S, Zuo T, Ho M, *et al*. Review article: fungal alterations in inflammatory bowel diseases. *Aliment Pharmacol Ther* 2019;50:1159–1171. DOI:10.1111/apt.15523

78. Zamani A, Jeihanipour A, Edebo L, *et al*. Determination of glucosamine and N-acetyl glucosamine in fungal cell walls. *Agric Food Chem* 2008;56:8314–8318. DOI:10.1021/jf801478j

79. Hoarau G, Mukherjee P K, Gower-Rousseau C, *et al*. Bacteriome and mycobiome interactions underscore microbial dysbiosis in familial Crohn's disease. *mBio* 2016;7:e01250–16. DOI:10.1128/mBio.01250-16

80. Franzosa E A, Sirota-Madi A, Avila-Pacheco J, *et al*. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol* 2019;4:293–305. DOI:10.1038/s41564-018-0306-4

81. Song A A, In L L A, Lim S H E, Rahim R A A review on Lactococcus lactis: from food to factory. *Microb Cell Fact* 2017;16:55. DOI:10.1186/s12934-017-0669-x

82. Shimizu K, Ogura H, Goto M, *et al*. Altered gut flora and environment in patients with severe SIRS. *J Trauma* 2006;60:126–133. DOI:10.1097/01.ta.0000197374.99755.fe

83. Hayakawa M, Asahara T, Henzan N, *et al*. Dramatic changes of the gut flora immediately after severe and sudden insults. *Dig Dis Sci* 211;56:2361–2365. DOI:10.1007/s10620-011-1649-3

84. Verstockt B, Verstockt S, Dehairs J, *et al*. Low TREM1 expression in whole blood predicts anti-TNF response in inflammatory bowel disease. *EBioMedicine* 2019;40:733–742. DOI:10.1016/j.ebiom.2019.01.027

85. Jung E S, Choi K W, Kim S W, *et al*. ZNF133 is associated with infliximab responsiveness in patients with inflammatory bowel diseases. *J Gastroenterol Hepatol* 2019;34:1727–1735. DOI:10.1111/jgh.14652

86. West N, Hegazy A, Owens B, *et al*. Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor–neutralizing therapy in patients with inflammatory bowel disease. *Nat Med* 2017;23:579–589. DOI:10.1038/nm.4307

87. Papamichael K, Lin S, Moore M, *et al*. Infliximab in inflammatory bowel disease. *Ther Adv Chronic Dis* 2019;10:2040622319838443. DOI:10.1177/2040622319838443

88. Minar P, Lehn C, Tsai Y T, *et al*. Elevated pretreatment plasma oncostatin M is associated with poor biochemical response to infliximab. *Crohn's & colitis* 2019;360:otz026. DOI:10.1093/crocol/otz026

89. Main J, McKenzie H, Yeaman G R, *et al*. Antibody to Saccharomyces cerevisiae (baker's yeast) in Crohn's disease. *BMJ* 1988;297:1105–1106. DOI:10.1136/bmj.297.6656.1105

90. Standaert–Vitse A, Jouault T, Vandewalle P, *et al*. Candida albicans is an immunogen for anti–Saccharomyces cerevisiae antibody markers of Crohn's disease. *Gastroenterology* 2006;130:1764–1775. DOI:10.1053/j.gastro.2006.02.009

91. Ben-Horin S, Kopylov U, Chowers Y. Optimizing anti-TNF treatments in inflammatory bowel disease. *Autoimmun Rev* 2014;13:24–30. DOI:10.1016/j.autrev.2013.06.002

92. Schnitzler F, Fidder H, Ferrante M, *et al*. Long-term outcome of treatment with infliximab in 614 patients with Crohn's disease: results from a single-centre cohort. *Gut* 2009;58:492–500. DOI:10.1136/gut.2008.155812

93. Cholapranee A, Hazlewood G S, Kaplan G G, *et al*. Systematic review with meta-analysis: Comparative efficacy of biologics for induction and maintenance of mucosal healing in Crohn's disease and ulcerative colitis controlled trials. *Aliment Pharm Ther* 2017;45:1291–1302. DOI:10.1111/apt.14030

94. Peyrin-Biroulet L, Reinisch W, Colombel J F, *et al*. Clinical disease activity, C-reactive protein normalisation and mucosal healing in Crohn's disease in the SONIC trial. *Gut* 2014;63:88–95. DOI:10.1136/gutjnl-2013-304984

95. Colombel J F, Keir M E, Scherl A, *et al*. Discrepancies between patient-reported outcomes, and endoscopic and histological appearance in UC. *Gut* 2017;66:2063–2068. DOI:10.1136/gutjnl-2016-312307

Figure Legends

Figure 1. Patients flowchart. R equals remission (responder), NR non-responders and PR partial remission (partial responders) to infliximab therapy.

Figure 2. Gut bacterial composition at genus level during the study at different time points in all IBD patients, in CD and in UC (including 2 IBDU subjects). The plots present the most abundant taxa (mean abundance = 0.01, prevalence = 0.1) showing the relative abundance of the taxa in the response groups. The taxa are colour-coded and shown on the right side of the panel. Numbers of faecal samples available at each timepoint (weeks) for analysis are presented below each bar (n). IBD equals inflammatory bowel disease, CD Crohn's disease, UC ulcerative colitis and IBDU unclassified inflammatory bowel disease.

Figure 3. Relative abundances of bacteria that showed Significant differences in generalized linear models from the MASS package at family level before and during IFX treatment in all IBD patients stratified by IFX treatment response groups. The numbering at x-axis indicates time points as follows: 1 = baseline, 2 = week 2, 3 = week 6, 4 = week 12 and 5 = week 52 from initiation of IFX therapy. The groups non-responder and partial-responder at each time point are compared to group responders. The group means and standard errors of relative abundance are shown. Significant differences are indicated with asterisks (* $p < 0.05$, ** $p < 0.001$, *** $p < 0.0001$). IFX equals infliximab and IBD inflammatory bowel disease.

Figure 4. Gut bacterial (A) composition of main taxa and differences at bacterial genus level in all IBD patients (B), CD (Crohn's disease, C) and UC (ulcerative colitis, D) at baseline stratified by response to infliximab treatment after twelve weeks of treatment. A) The plot presents the most abundant bacterial genera (mean abundance = 0.01), showing the relative abundance of the taxa in the response groups. The taxa are color-coded and shown on the right side of the panel. B-D) Effect of IFX response on taxa at baseline in CD and UC. The fold changes represent the difference in the relative abundance of the taxon between non-responders and partial responders compared to responders. The asterisks indicate the significance of difference (based on generalized linear model or generalized least squares, see supplementary tables ST3-5): * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0001$. IBD equals inflammatory bowel disease and R equals remission (responder), NR non-responders and PR partial remission (partial-responders) to infliximab therapy.

Figure 5. Receiver operating characteristic (ROC) curve of the relation between treatment response (to distinguish NR from R) and bacterial genera as predictive markers for the response to infliximab (IFX) therapy in all IBD patients (A), and CD (B) and UC (C) patients separately. The genera used for ROC analyses were those found to differ significantly in response groups by using PathModel function for all IBD and the genera that we found to differ significantly in this study for CD and UC patients (see Supplementary Methods 5). The area under curve (AUC) is indicated. IBD equals inflammatory bowel disease, CD Crohn's disease and UC ulcerative colitis.

Tables

Table 1. Basic clinical characteristic of IBD patients.

Characteristics	n (%) or Median (IQR)
No. of patients	72
Female	30 (41.7)
Crohn's disease (CD)	25 (34.7)
Ulcerative colitis (UC)	47 (65.3)
Montreal classification, UC	47
E1/E2/E3	1 (2.1) / 11 (23.4) / 35 (74.5)
Montreal classification, CD	25
A1/A2/A3	3 (12) / 18 (72) / 4 (16)
L1/L2/L3	3 (12) / 8 (32) / 14 (56)
B1/B2/B3	14 (56) / 5 (20) / 6 (24)
Perianal disease	12 (48)
Age at diagnosis, years	25 (19-25)
Age at IFX initiation, years	31 (24-45)
Disease duration, years	2 (0-7)
Smoking	11 (15.3)
Extraintestinal manifestations	
Arthritis/sacroiliitis	8 (11)
Iritis/uveitis	1 (1.4)
Primary sclerosing cholangitis	2 (2.8)
Prior surgery	9 (12.5)
Concomitant medication at IFX initiation	
Metronidazole	6 (8.3)
Ciprofloxacin	7 (9.7)
Steroid	35 (48.6)
Thiopurine	50 (69.5)
Methotrexate	5 (6.9)
5-Aminosalicylic acid	39 (54.2)

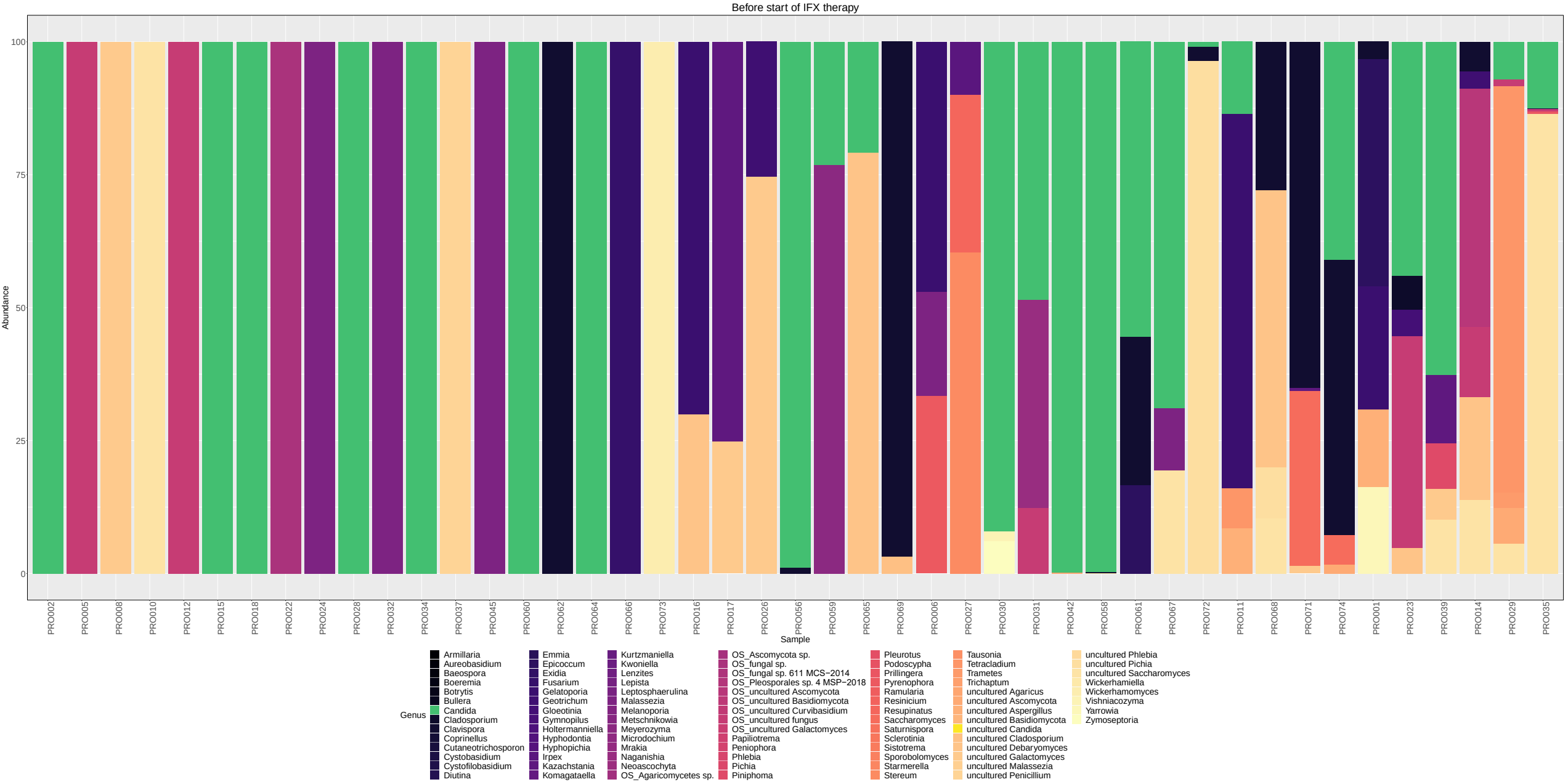
Supplementary Figures

Supplementary Figure S1-S13

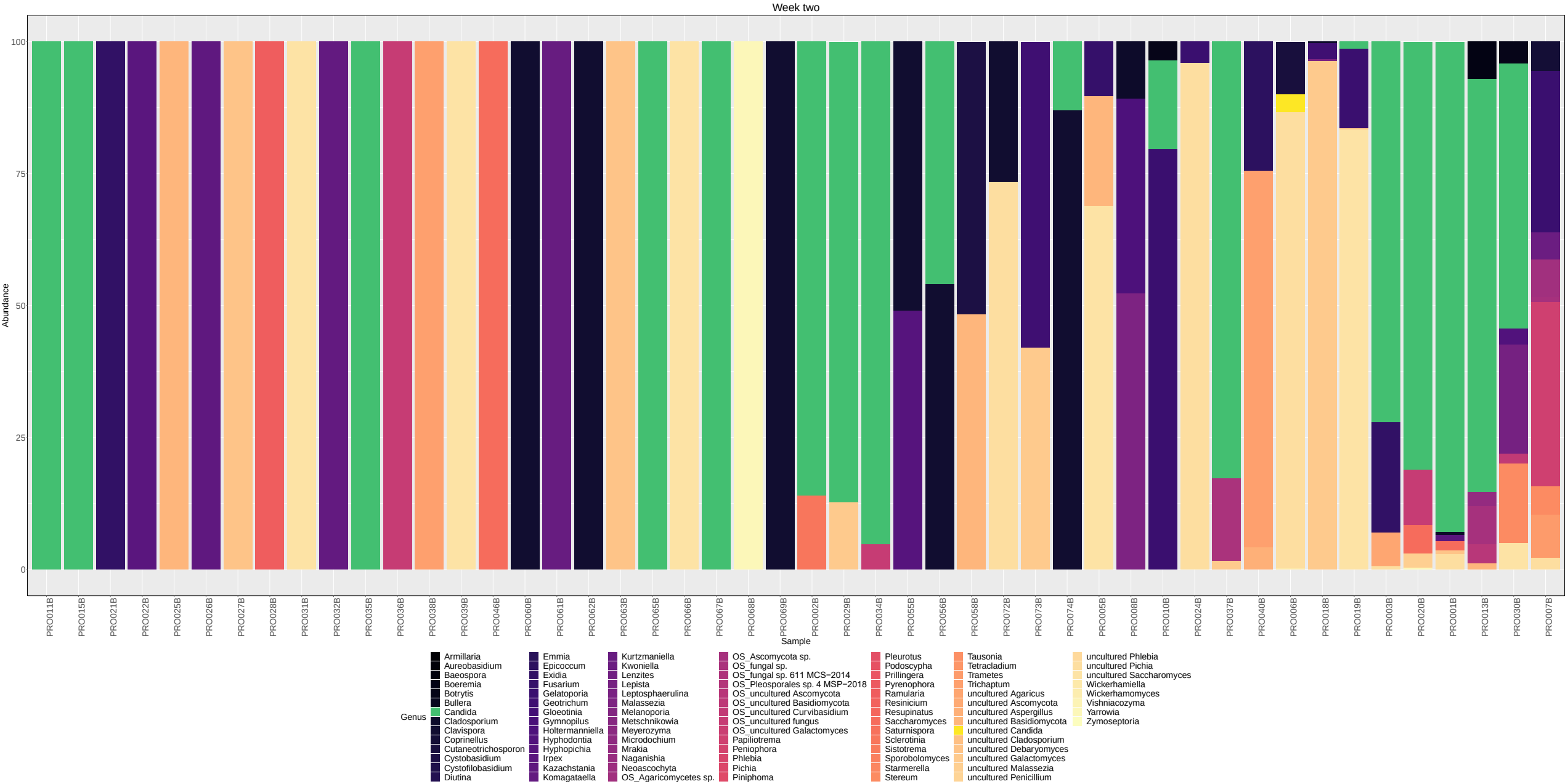
Bacterial and fungal profiles as markers of infliximab drug response in inflammatory bowel disease

Rebecka Ventin-Holmberg, Anja Eberl, Schahzad Saqib, Katri Korpela, Seppo Virtanen, Taina Sipponen,

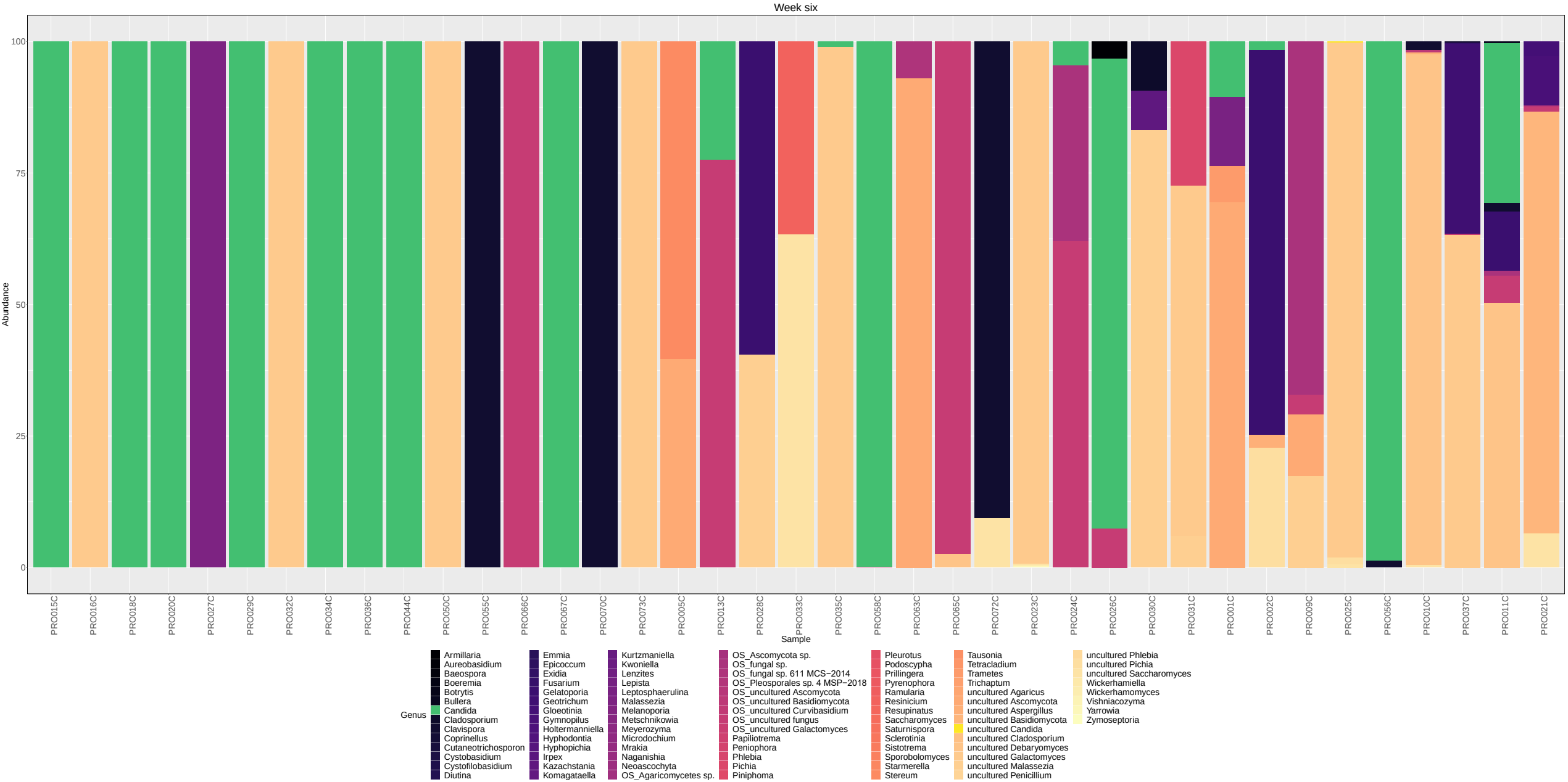
Anne Salonen, Päivi Saavalainen and Eija Nissilä



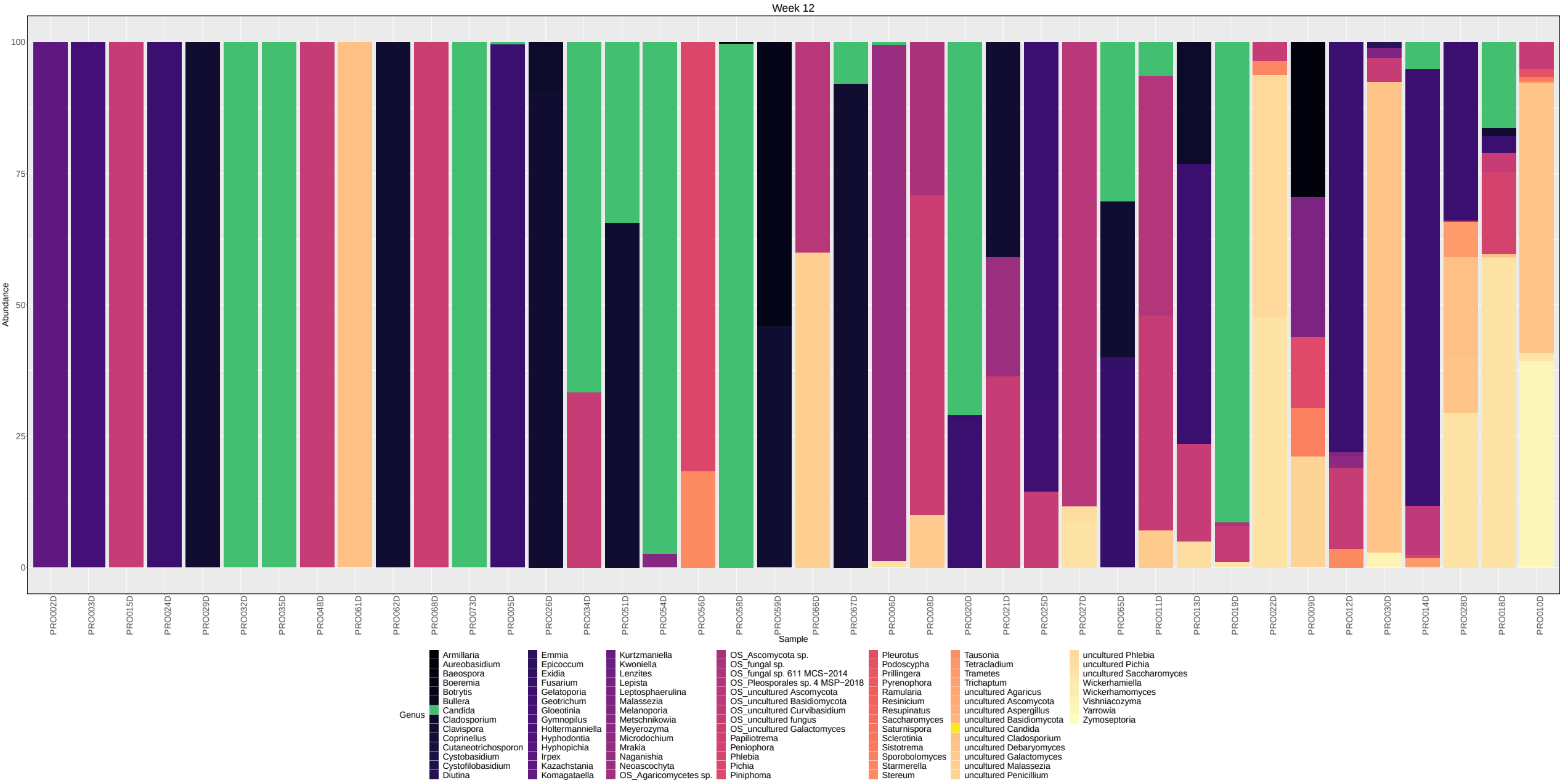
Supplementary Figure S1



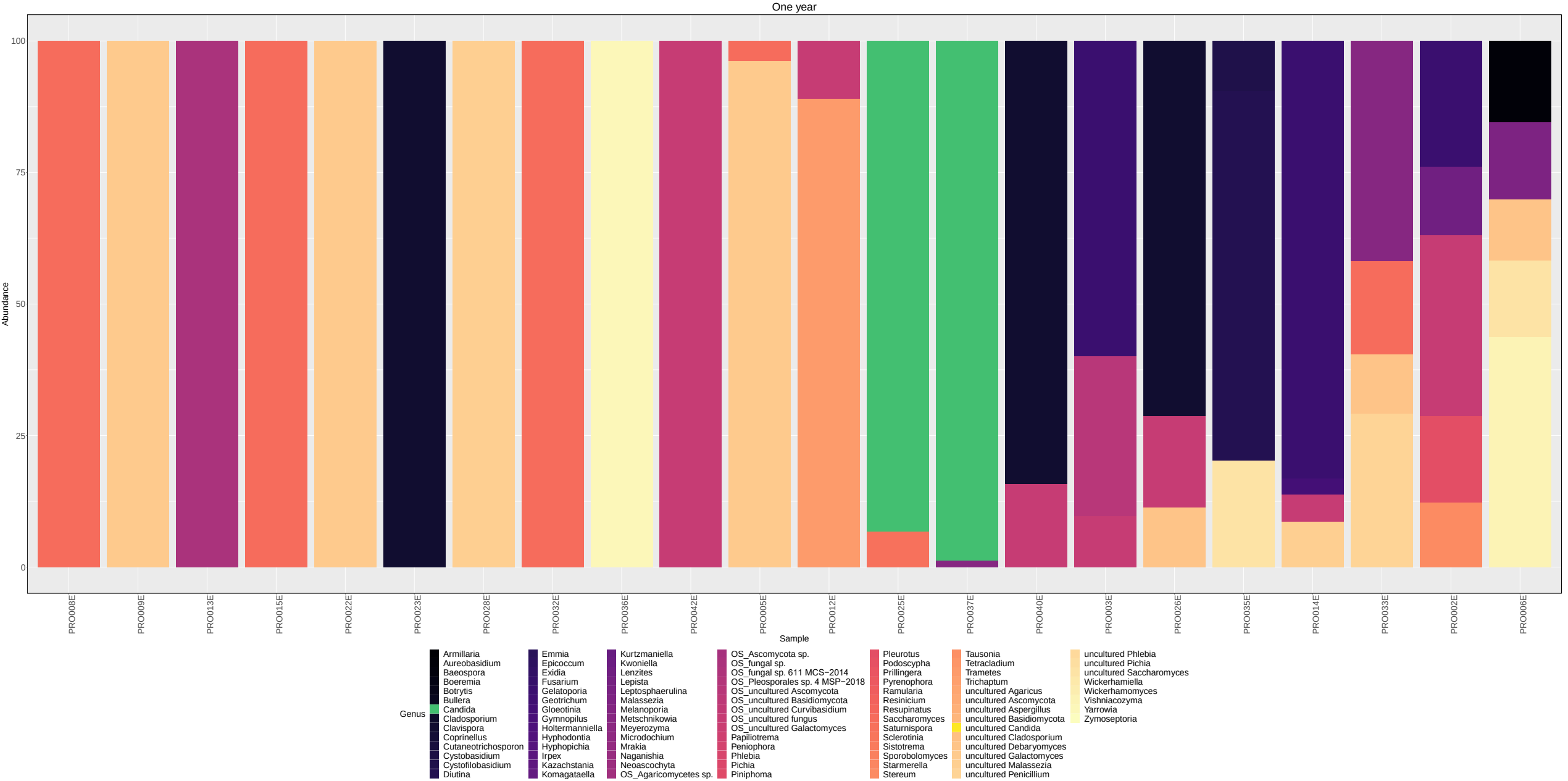
Supplementary Figure S2



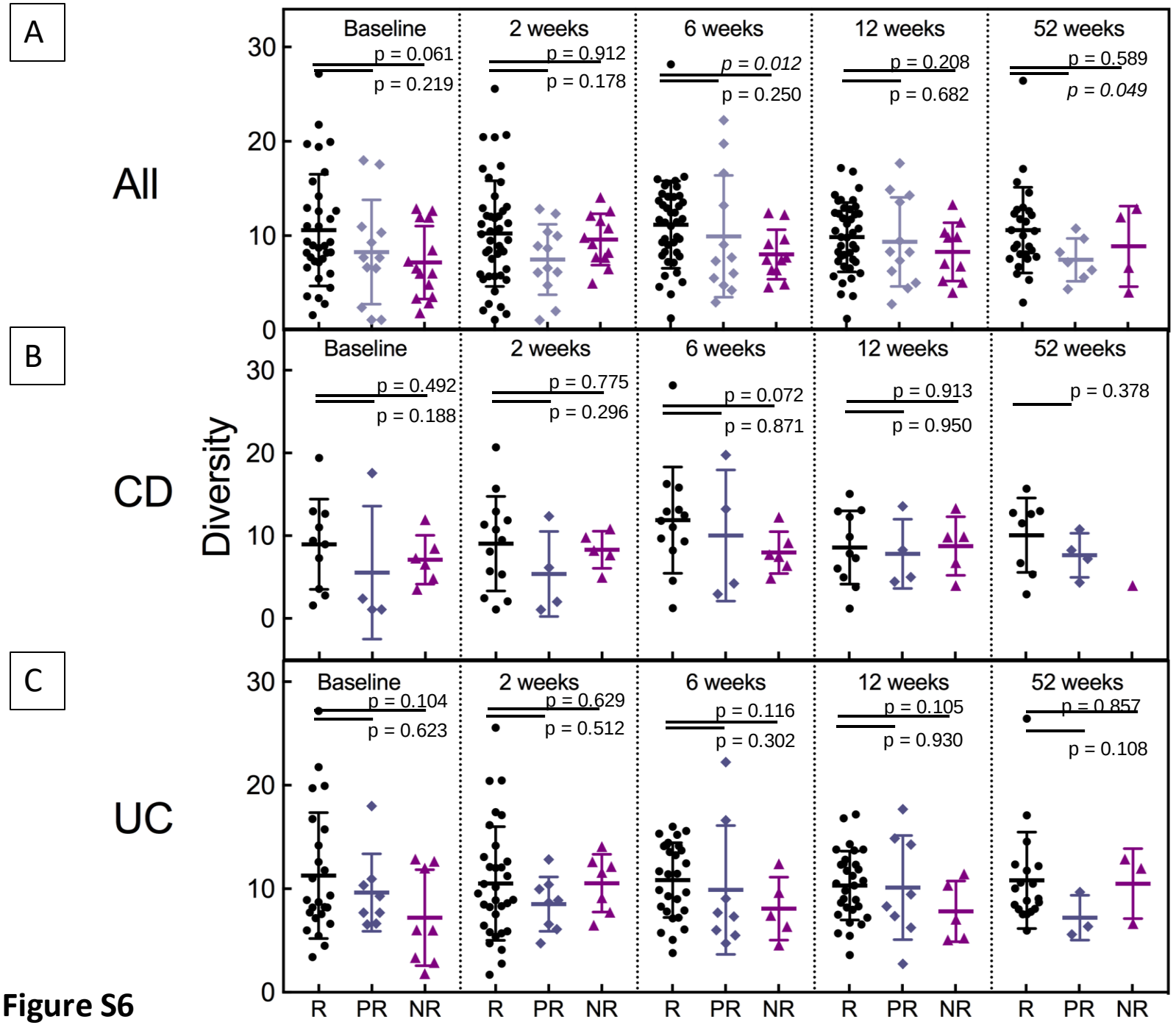
Supplementary Figure S3



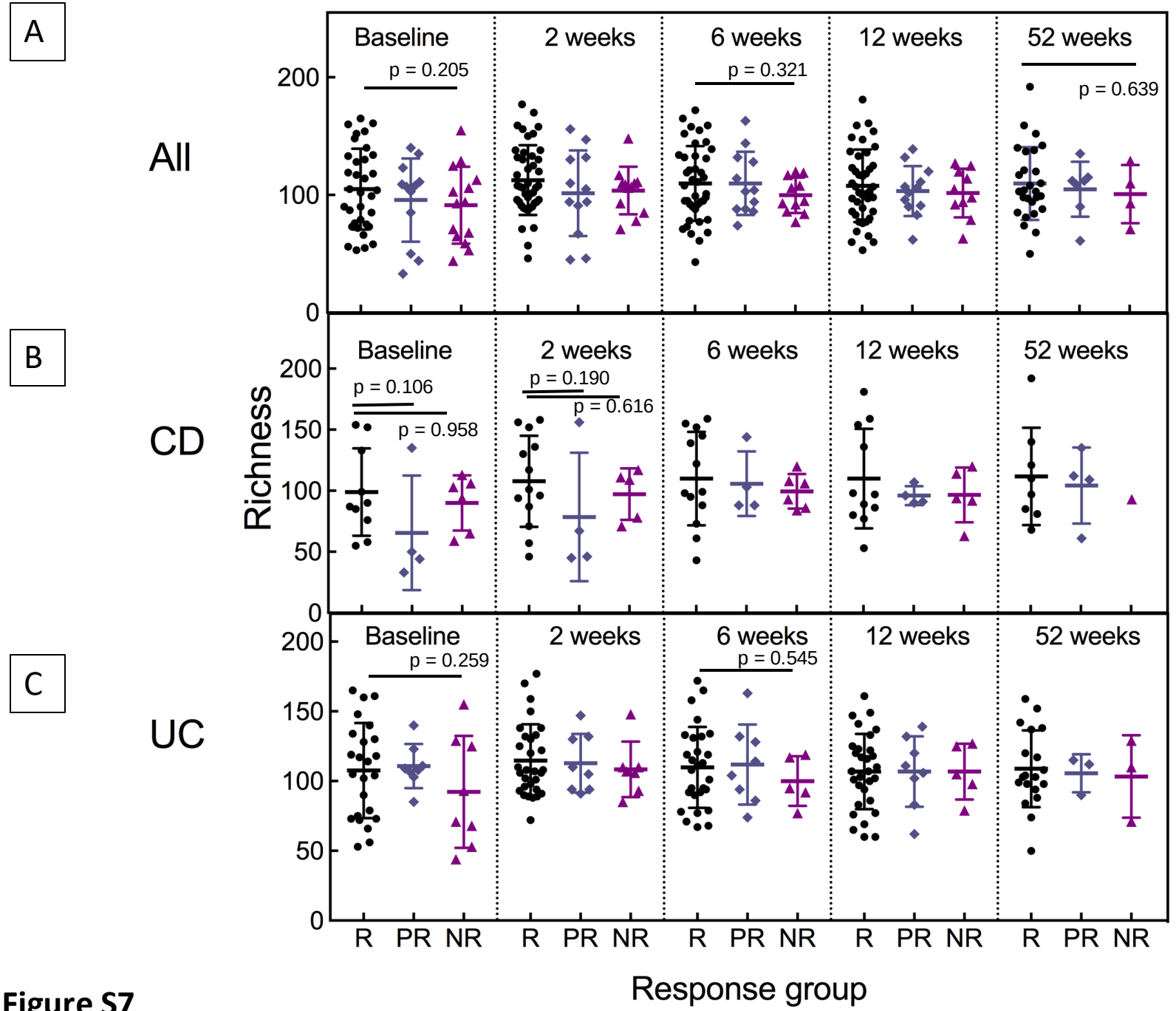
Supplementary Figure S4



Supplementary Figure S5



Supplementary Figure S6

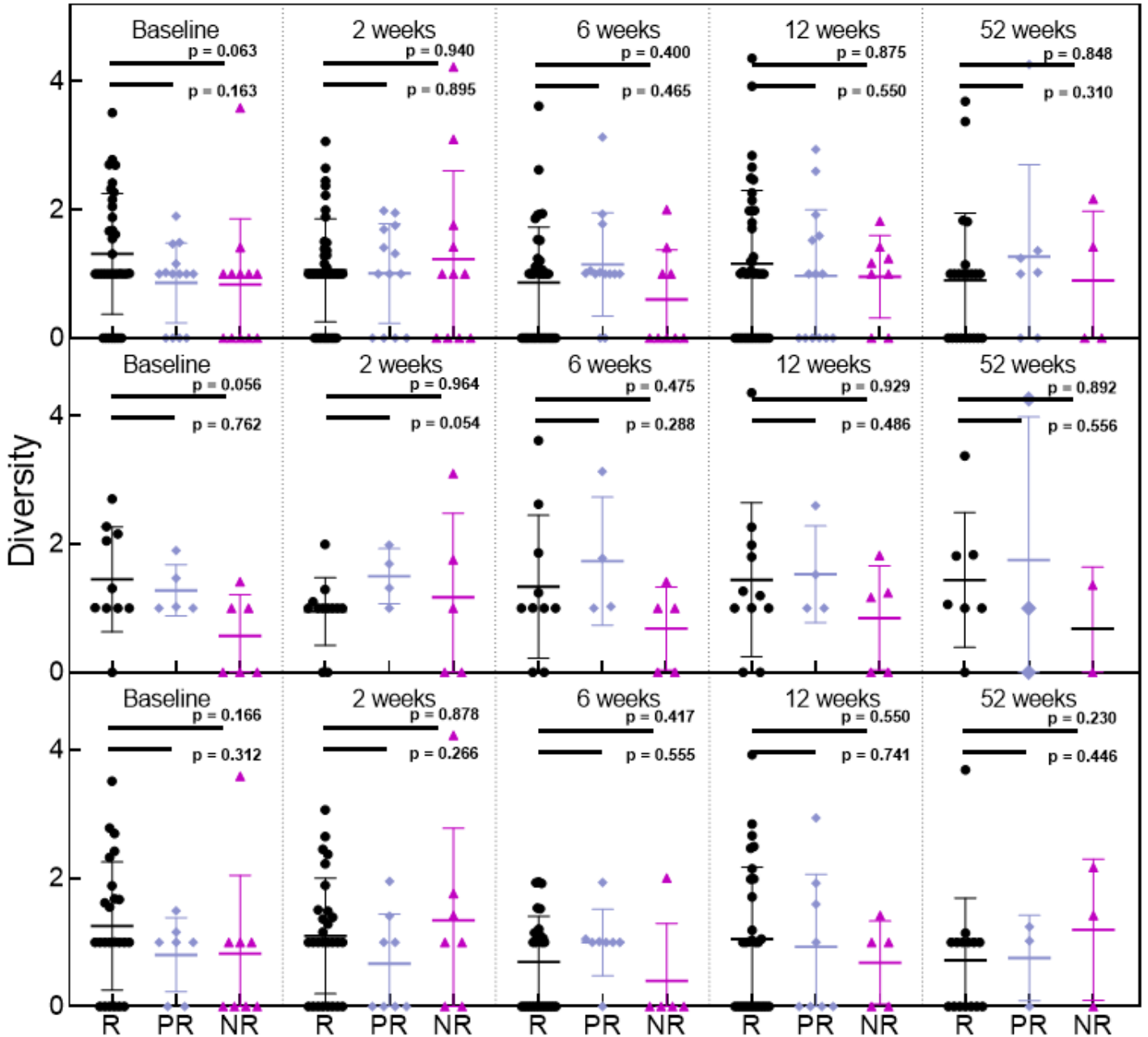


Supplementary Figure S7

A All

B CD

C UC

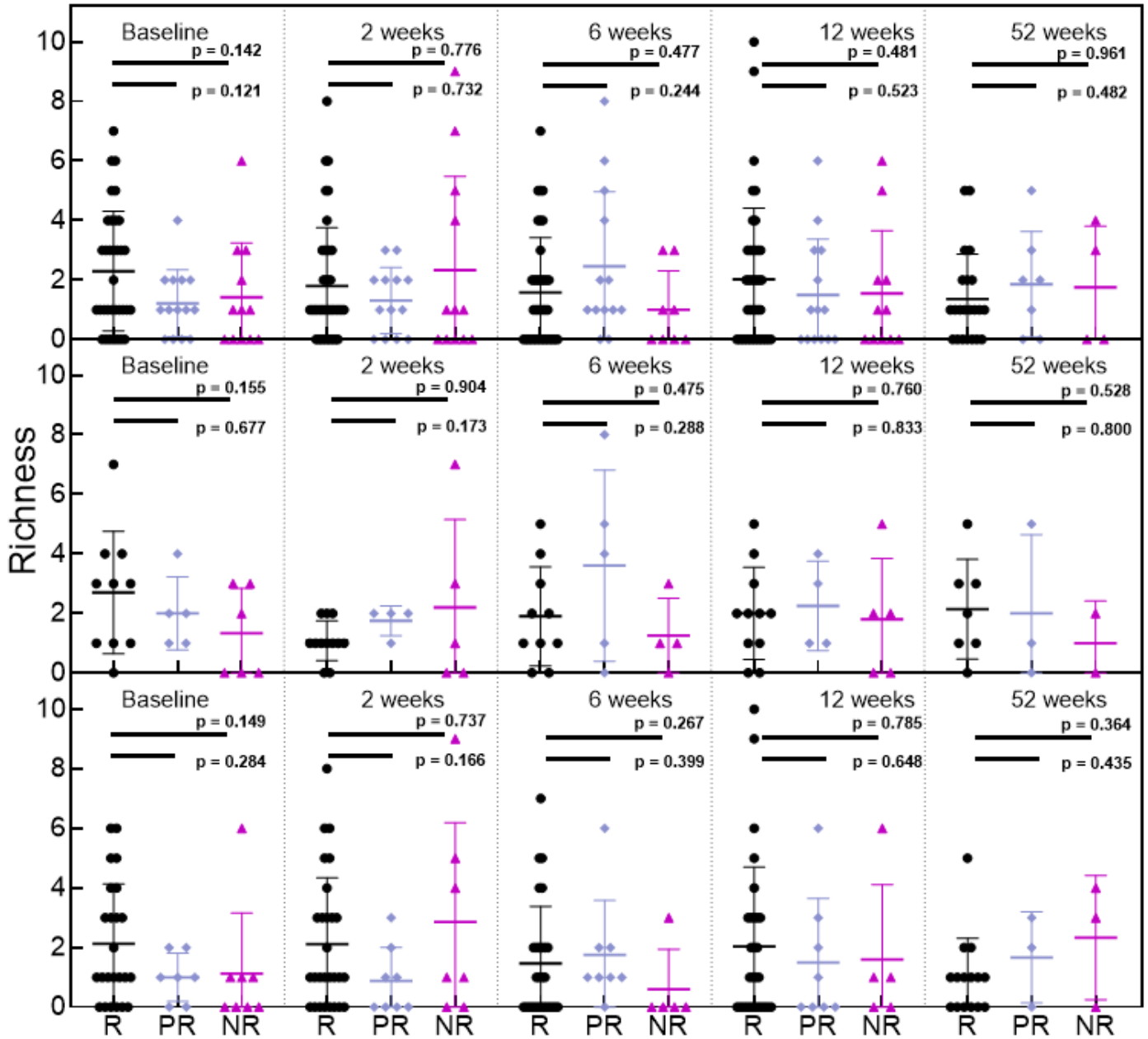


Supplementary Figure S8

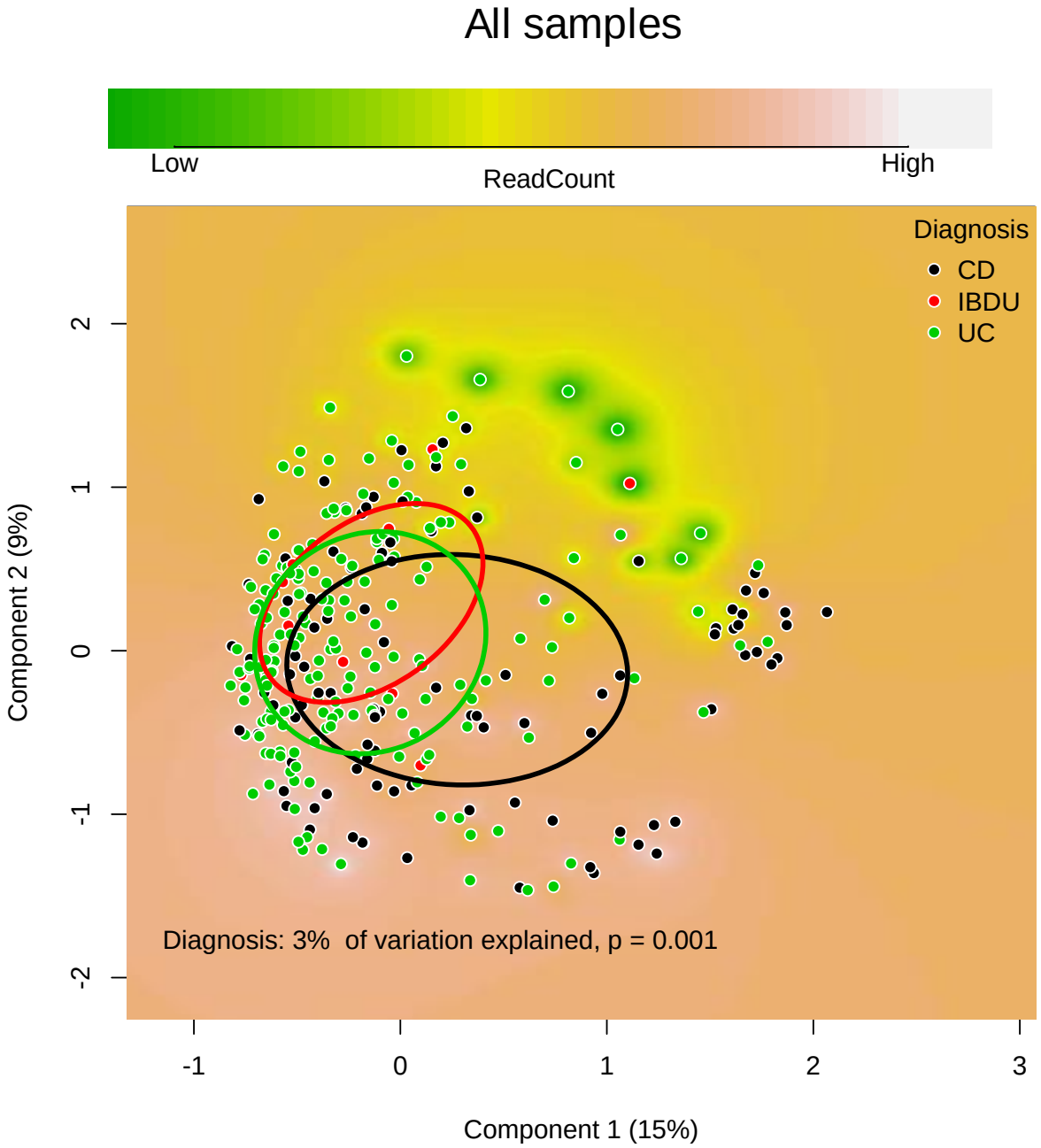
A All

B CD

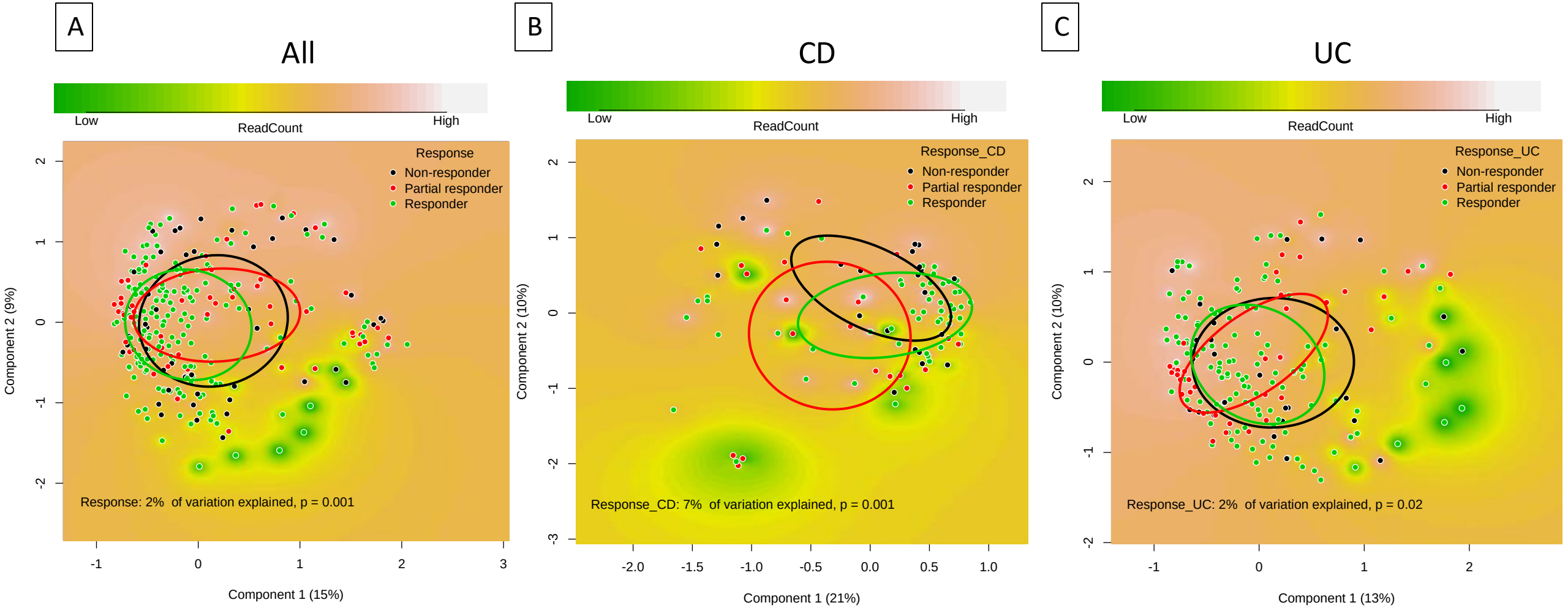
C UC



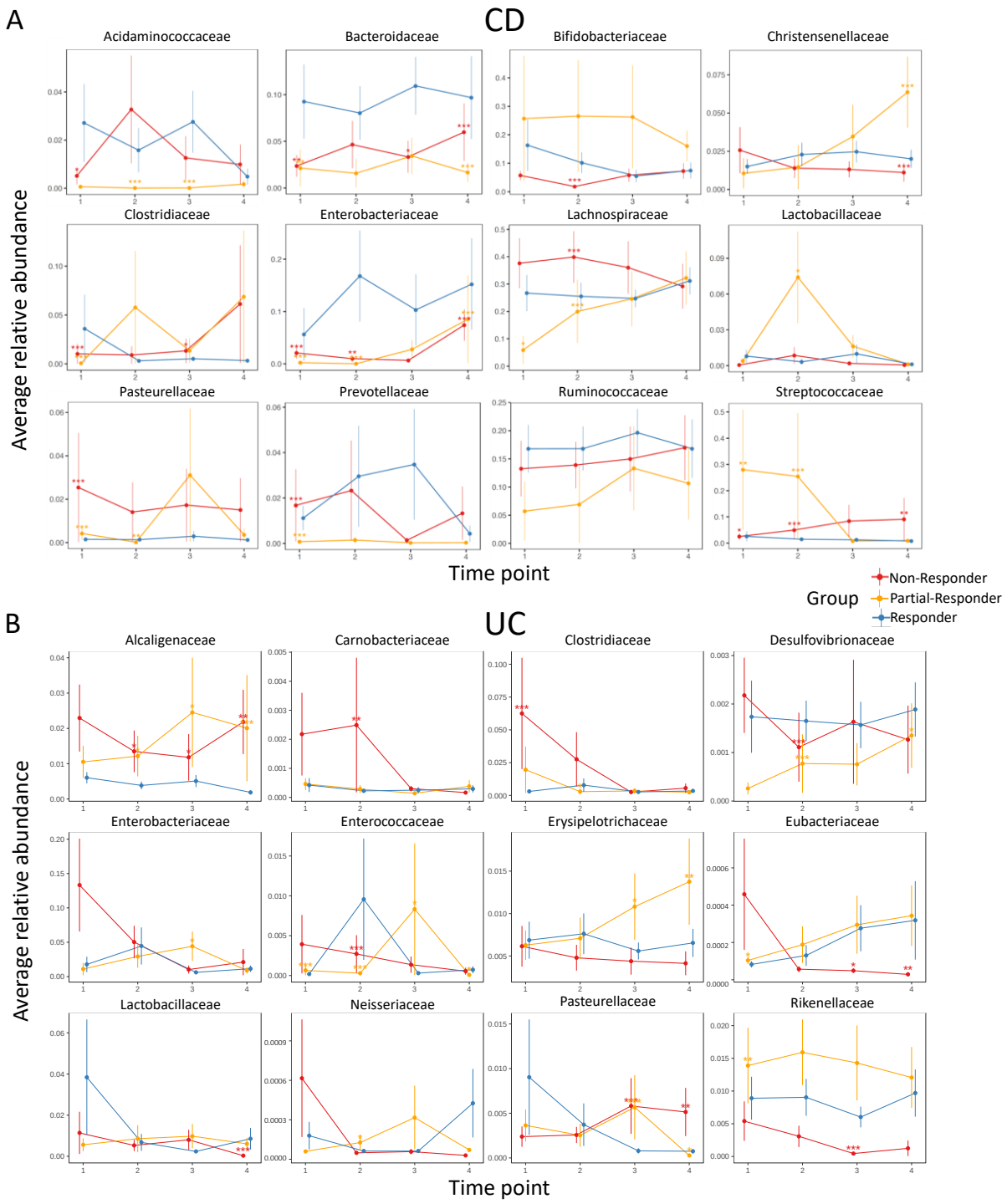
Supplementary Figure S9



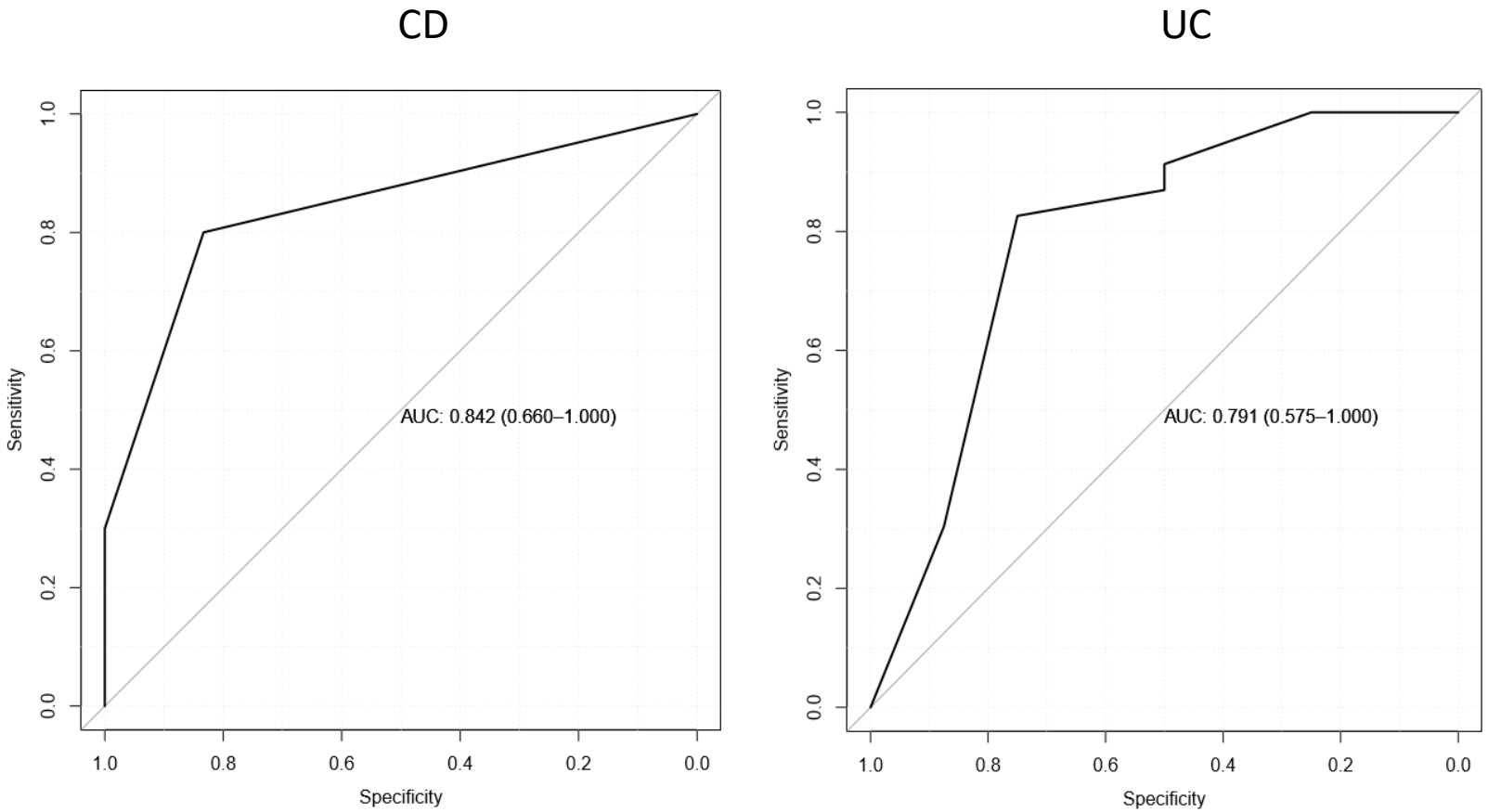
Supplementary Figure S10



Supplementary Figure S11



Supplementary Figure S12



Supplementary Figure S13

Supplementary materials - Bacterial and fungal profiles as markers of infliximab drug response in inflammatory bowel disease

Rebecka Ventin-Holmberg, Anja Eberl, Schahzad Saqib, Katri Korpela, Seppo Virtanen, Taina Sipponen, Anne Salonen, Päivi Saavalainen and Eija Nissilä

1. Supplementary methods:

1. ASCA IgG/IgA ELISA assay

IgG and IgA anti-*Saccharomyces cerevisiae* antibody (ASCA) concentrations were measured from IBD serum samples using a commercialized ELISA assay (ASCA IgG/IgA ELISA, Demeditec Diagnostics GmbH). The serum samples were diluted to 1:100, as recommended by the manufacturer, and were analysed as singlicates for both IgG and IgA ELISA assays. Plates were read at 405 nm by the Hidex Sense Microplate Reader. The OD 405 nm values were analysed by comparison to the standards included in the kit. Values exceeding or equal to 10 U ml^{-1} were considered as positive outcomes, as recommended by the manufacturer.

2. Mycobiota library preparation

The ITS1 region was first amplified with the ITS1F and ITS2 primers mentioned. In the second PCR, the PCR product from the initial PCR reaction was amplified using ITS1F (FWD, ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGGTCATTTAGAGGAAGTAA) and ITS2 (REV, AGACGTGTGCTCTTCCGATCTGCTGCGTTCTTCATCGATGC) primers with Illumina adapters for MiSeq sequencing. The PCRs were completed in a BioRad T100™ Thermal Cycler using the following conditions: denaturation at 98 ˚C for 60 s, 44 cycles at 98 ˚C for 10 s, 58 ˚C for 40 s, and 72 ˚C for 40 s, followed by a final extension time of 10 min at 72 ˚C. In the reaction mixture, 2x Phusion High-Fidelity PCR Master Mix with HF Buffer (ThermoFisher Scientific), 2 µl of each primer (5 µM), and 0.6 µl DMSO was used. A final volume of 20 µl for each reaction was reached by adding water. For the initial PCR reaction, 5 ng DNA was used and in the following PCR reaction, 4.4 µl of the PCR product was used as the template. After the second PCR, the product was

purified using AMPure XP beads (Beckman Coulter, Copenhagen, Denmark) according to the 16S metagenomic library prep guide using 0.8x concentration of beads [1]. The purified products were analysed using LabChip® GX Touch™ nucleic acid analyser. The DNA concentrations of the samples were measured with the Quant-It PicoGreen dsDNA Assay Kit and the FLUOstar plate reader was used to measure the concentration. In the third PCR reaction, barcodes for sequencing were added to the samples by index PCR, primers previously published [2]. As previously, the PCR products were purified with AMPure XP beads and analysed with LabChip and the concentration was measured. Finally, the samples were pooled for MiSeq sequencing by adding 10 nM of each sample to the pool.

3. Medication categorization

The categorization used in the 16S analyses are presented in Table 1.

Table 1.

Category number	Medication
1	Steroid
2	Steroid and azathiopurine
3	Steroid and merkaptopurine
7	Steroid and methotrexate
5	Steroid and mesalazine
6	Steroid and mesalazine and merkaptopurine
7	Steroid and mesalazine and azathiopurine
8	Mesalazine
9	Mesalazine and azathiopurine
10	Azathiopurine
11	Merkaptopurine
12	Methotrexate
13	No medication

4. Negative controls

The number of reads in the negative controls was small (median 403) compared to the real samples (median 40720 (combined data of two run)). This suggests that contaminants caused only a few hundred reads per sample, which would not affect the overall observed composition. The most abundant taxa in the negative controls were *Bifidobacterium* and an uncultured *Collinsella* bacterium, which made a very small abundance in the real samples, being 0.3% and 0.24% of total reads on average. In fungal negative controls, no annotations against gut-specific fungi were observed.

5. Taxa used for ROC analyses

In CD and UC patients we used the genera (see Figure 4) that we found to differ significantly between the response groups. In CD seven out of the twelve predictive genera were used in the model, based on AIC selection: *Bifidobacterium, Rothia, Atopobium, Gemella, Pseudoflavonifractor, Sutterella* and *Pseudomonas.* In UC the following genera were used: *Enterococcus, Clostridium, Peptostreptococcus, Faecalibacterium* and *Candida*.

For the ROC analysis the *Enterobacter* and *Alistipes* were used for all IBD patients. selected by using the PathModel function from mare. The genera that were used for the ROC analyses done based on PathModel function for CD and UC are found in Supplementary Figure S13.

References:

1. Illumina. 16S metagenomic sequencing library preparation guide. 2014.

2. Kozich J J, Westcott S L, Baxter N T, *et al*. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 2013;79:5112–5120.

2. Supplementary Table legends (ST):

All Supplementary Tables are found in separate excel file “Revised_Supplementary_Tables”.

Content:

ST1_Responses_to_IFX – Supplementary Table 1. Responses to infliximab (IFX) treatment evaluated after 12 weeks in IBD subtypes.

ST2_Prevalence_ITS – Supplementary Table 2. Faecal fungal prevalence in IBD patients.

ST3_ResponseGroups_Baseline – Statistical test on faecal bacterial microbiota – IFX response at baseline in all IBD patients

ST4_ResponseGroups_CD_Baseline – Statistical test on faecal bacterial microbiota – IFX response at baseline in CD patients

ST5_ResponseGroups_UC_Baseline – Statistical test on faecal bacterial microbiota – IFX response at baseline in UC patients (including IBDU patients)

ST6_ResponseGroups_Baseline – Statistical test on faecal fungal microbiota – IFX response at baseline in all IBD patients

ST7_ResponseGroups_CD_Baseline – Statistical test on faecal fungal microbiota – IFX response at baseline in CD patients

ST8_ResponseGroups_UC_Baseline – Statistical test on faecal fungal microbiota – IFX response at baseline in UC patients (including IBDU patients)

ST9_correlation_R – Spearman correlations between faecal bacterial and fungal genera at baseline in responders to IFX therapy

ST10_correlation_PR – Spearman correlations between faecal bacterial and fungal genera at baseline in partial responders to IFX therapy

ST11_correlation_NR – Spearman correlations between faecal bacterial and fungal genera at baseline in non-responders to IFX therapy

3. Supplementary Figure legends:

All Supplementary Tables are found in separate excel file "Revised_Supplementary_Figures".

Supplementary Figure S1. **Faecal fungal microbiota at genus level in IBD patients before start of IFX therapy.** The plot shows the mycobiota composition for each patient at baseline. The fungal genera are colour-coded and showed under the plot. *Candida* is highlighted with a green colour. IBD equals inflammatory bowel disease and IFX infliximab.

Supplementary Figure S2. **Faecal fungal microbiota at genus level in IBD patients at two weeks after start of IFX therapy.** The plot shows the mycobiota composition for each patient at baseline. The fungal genera are colour-coded and showed under the plot. *Candida* is highlighted with a green colour. IBD equals inflammatory bowel disease and IFX infliximab.

Supplementary Figure S3. **Faecal fungal microbiota at genus level in IBD patients at six weeks after start of IFX therapy.** The plot shows the mycobiota composition for each patient at baseline. The fungal genera are colour-coded and showed under the plot. *Candida* is highlighted with a green colour. IBD equals inflammatory bowel disease and IFX infliximab.

Supplementary Figure S4. **Faecal fungal microbiota at genus level in IBD patients at twelve weeks after start of IFX therapy.** The plot shows the mycobiota composition for each patient at baseline. The fungal genera are colour-coded and showed under the plot. *Candida* is highlighted with a green colour. IBD equals inflammatory bowel disease and IFX infliximab.

Supplementary Figure S5. **Faecal fungal microbiota at genus level in IBD patients at one year after start of IFX therapy.** The plot shows the mycobiota composition for each patient at baseline. The fungal genera are colour-coded and showed under the plot. *Candida* is highlighted with a green colour. IBD equals inflammatory bowel disease and IFX infliximab.

Supplementary Figure S6. **Bacterial diversity during study in IBD patients stratified in IFX treatment response groups. Diversity is presented at different timepoints during study in all IBD patient, in CD and UC.** Infliximab Response was evaluated after 12 weeks from initiation of IFX treatment. Diversity is presented at baseline, 2 weeks, 6 weeks, 12 weeks and one year after initiation of treatment in all IBD patients (A), CD patients (B) and UC patients(C). R equals remission, PR partial remission and NR no remission achieved. One dot represents an individual sample. Differences between response groups were analysed using Mann-Whitney test. IBD equals inflammatory bowel disease, CD Crohn's disease, UC ulcerative colitis and IFX infliximab.

Supplementary Figure S7. **Bacterial richness during study in IBD patients stratified in IFX treatment response groups. Richness is presented at different timepoints during study in all IBD patient, in CD and UC.** Response was evaluated after 12 weeks from initiation of IFX treatment. Richness is presented at baseline, 2 weeks, 6 weeks, 12 weeks and one year after initiation of treatment in all IBD patients (A), CD patients (B) and UC patients(C). R equals responder, PR partial responder and NR non-responder. Differences between response groups were analysed using Mann-Whitney test. Each dot presents one faecal sample. Mean and standard deviation values of

each group are shown as lines. IBD equals inflammatory bowel disease, CD Crohn's disease, UC ulcerative colitis and IFX infliximab.

Supplementary Figure S8. Fungal diversity during study in IBD patients stratified in IFX treatment response groups. **Diversity is presented at different timepoints during study in all IBD patient, in CD and UC.** Infliximab Response was evaluated after 12 weeks from initiation of IFX treatment. Diversity is presented at baseline, 2 weeks, 6 weeks, 12 weeks and one year after initiation of treatment in all IBD patients (A), CD patients (B) and UC patients (C). Group R equals remission (responder), PR partial remission (partial responder) and NR no remission achieved (non-responder). IBD equals inflammatory bowel disease, CD Crohn's disease, UC ulcerative colitis and IFX infliximab. One dot represents one faecal sample. Differences between response groups were analysed using Mann-Whitney test.

Supplementary Figure S9. **Fungal richness during study in IBD patients stratified in IFX treatment response groups. Richness is presented at different timepoints during study in all IBD patient, in CD and UC.** Response was evaluated after 12 weeks from initiation of IFX treatment. Richness is presented at baseline, 2 weeks, 6 weeks, 12 weeks and one year after initiation of treatment in all IBD patients (A), CD patients (B) and UC patients (C). R equals responder, PR partial responder, NR non-responder, IBD equals inflammatory bowel disease, CD Crohn's disease, UC ulcerative colitis and IFX infliximab. Differences between response groups were analysed using Mann-Whitney test. Each dot presents one faecal sample. Mean and standard deviation values of each group are shown as lines.

Supplementary Figure S10. **Principal coordinates analysis (PCoA) of faecal bacteria in Crohn's disease (black, n = 103)), ulcerative colitis (green, n = 184) and IBDU (red, n = 10) subjects (A).** The background colour indicates interpolated values of read count. IBDU equals unclassified inflammatory bowel disease, CD Crohn's disease and UC ulcerative colitis.

Supplementary Figure S11. **Principal coordinates analysis (PCoA) of faecal bacteria in all patients (A), in Crohn's disease (B) and in ulcerative colitis (C) stratified in IFX treatment response groups.** Each dot stands for one sample, and the ellipses group stands for the three groups responding to IFX treatment at 12 weeks. IFX equal infliximab. The background colour indicates interpolated values of read count.

Supplementary Figure S12. **Average relative abundances of 12 most abundant bacteria that showed significant differences in generalized linear models from the MASS package at family level before and during IFX treatment in CD (A) and UC (B) patients stratified by IFX treatment response groups.** The numbering at x-axis indicates time points as follows: 1 = baseline, 2 = week 2, 3 = week 6, 4 = week 12 and 5 = week 52 from initiation of IFX therapy. The groups non-responders and partial responders at each time point are compared to group responders. The group means and standard errors of relative abundance are shown. Significant differences are indicated with asterisks (* $p < 0.05$, ** $p < 0.001$, *** $p < 0.0001$).CD equals Crohn's disease, UC ulcerative colitis and IFX infliximab.

Supplementary Figure S13. **Receiver operating characteristic (ROC) curve of the relation between treatment response (to distinguish NR from R) and bacterial genera as predictive markers for the response to infliximab (IFX) therapy in CD and UC patients.** The genus used for ROC analyses was *Anaerofilum* for CD patients and unknown *Lachnospiraceae* based on PathModel function from mare package in R. The area under curve (AUC) is indicated. R equals responder and NR non-responder.

[InternetShortcut] URL=https://helsinkifi-my.sharepoint.com/:b:/g/personal/ekelloko_ad_helsinki_fi/EWPC93Qv24VGmNuX3Qh-KvwBHPzdWbIMAlwpPXLcEWJhXg IDList=
HotKey=0 IconFile=C:\Users\lhuovine\AppData\Local\Mozilla\Firefox\Profiles\kxdztn42.default\shortcutCache\y1vh+YaiDbXlN_WISnG_zQ==.ico
IconIndex=0

[InternetShortcut] URL=https://helsinkifi-my.sharepoint.com/:b:/g/personal/ekelloko_ad_helsinki_fi/EfuMztT90ztCjiMiUC9otS0BKfeBA9MIIvSaPtY7u0ndmQ IDList=
HotKey=0 IconFile=C:\Users\lhuovine\AppData\Local\Mozilla\Firefox\Profiles\kxdztn42.default\shortcutCache\Kszr3f92QLsDkfB_v2dKrw==.ico
IconIndex=0

[InternetShortcut] URL=https://helsinkifi-my.sharepoint.com/:b:/g/personal/ekelloko_ad_helsinki_fi/EbgTY_AGhdtIsgKsoRKx5dMBnH3PxA2-82ycWOIntKubnA IDList=
HotKey=0 IconFile=C:\Users\lhuovine\AppData\Local\Mozilla\Firefox\Profiles\kxdztn42.default\shortcutCache\o4UmZ7BLiR4eJP5bz9F_3g==.ico
IconIndex=0

[InternetShortcut] URL=https://helsinkifi-my.sharepoint.com/:b:/g/personal/ekelloko_ad_helsinki_fi/ER7Y1SpaKmNGiH_NhsI9beQBaTKPvRwfYl5fhwxeoSBjvA IDList=
HotKey=0 IconFile=C:\Users\lhuovine\AppData\Local\Mozilla\Firefox\Profiles\kxdztn42.default\shortcutCache\HsjOUFXOXVJSGCmlhdNRmw==.ico
IconIndex=0

[InternetShortcut] URL=https://helsinkifi-my.sharepoint.com/:b:/g/personal/ekelloko_ad_helsinki_fi/EXMKNBn09-NKsC0b5Fo8lO4BR_E1tvF8x8uDO7lICGSZpw IDList=
HotKey=0 IconFile=C:\Users\lhuovine\AppData\Local\Mozilla\Firefox\Profiles\kxdztn42.default\shortcutCache\iZTqKZ0rNiJF1QlN67Vi6g==.ico
IconIndex=0