Comparison of Small Gut and Whole Gut Microbiota of First-Degree Relatives with Adult Patients with Celiac Disease and Controls

Sudarshan A. Shetty[1,4#], Rahul Bodkhe[1#], Dhiraj P. Dhotre[1], Anil K. Verma[2,5], Khushbo Bhatia[2], Asha Mishra[2], Gurvinder Kaur[3], Pranav Pande[1], Dhinoth K. Bangarusamy[3], Beena P. Santosh[3], Rajadurai C. Perumal[3], Vineet Ahuja[2], Yogesh S. Shouche[1*], Govind K. Makharia[2*]

[1]National Centre for Microbial Resource, National Centre for Cell Science, Pune-411007, India.

[2]Department of Gastroenterology and Human Nutrition, All India Institute of Medical Sciences, New Delhi, India.

Transplant Immunology & Immunogenetics, All India Institute of Medical Sciences, New Delhi, India

[3]AgriGenome Labs Pvt Ltd. Kerala, India.

[4] Current address: Laboratory of Microbiology, Wageningen University and Research, Building 124, Stippeneng 4, 6708 WE Wageningen, The Netherlands.

[5] Current address: Celiac Disease Research Laboratory, Department of Pediatrics, Università Politecnica delle Marche, Ancona, Italy.

*Co-corresponding authors: Yogesh S. Shouche (yogesh@nccs.res.in) and Govind K. Makharia (govindmakharia@gmail.com)

# Authors with equal contribution.

Abstract

Objectives: Gut microbiota gets altered in patients with celiac disease (CeD) and whether these microbiota changes are the cause or effect of the disease is not well understood to date. The first degree relatives (FDRs) of CeD patients are genetically susceptible and may represent a pre-diseased state. Therefore, understanding differences in duodenal and faecal microbiota composition between the FDR and CeD subjects is of interest. To investigate this, we characterised the microbiota in duodenal biopsies and faeces of CeD patients (n = 23), FDRs (n = 15) and control subjects (DC, n= 24) by 16S rRNA gene sequencing.

Results: Duodenal biopsies showed more diverse pattern in microbial community composition and structure than faecal samples. In duodenal biopsies, 52 OTUs and 41 OTUs were differentially abundant between the FDR and DC group, and between the FDR and CeD group respectively ($P < 0.01$). In faecal samples, 30 OTUs were differentially abundant between FDR and DC, and 81 between FDR and CeD ($P < 0.01$). Predicted metagenomes from duodenal microbiomes of FDR and CeD showed a lower genetic potential for metabolizing gluten as compared to controls.

Conclusions: The microbial communities of FDR and CeD groups are more similar to each other than to the control groups. Significant differences at OTU level suggest that specific bacterial taxa may be important for pathogenesis of CeD. Moreover, the predicted differences in gluten metabolism potential by the FDR and CeD microbiota point towards the need for investigating functional capabilities of specific bacterial taxa in healthy FDR and CeD patients.

Key words

Gluten, Microbiota, Malabsorption, *Acinetobacter*, *Lactobacillus.*

Background

Celiac disease (CeD) is a common, chronic immune mediated enteropathy of the small intestine which affects approximately 0.7% of the global population (1). Once thought to be uncommon in Asia, CeD is now prevalent in many Asian countries including India (2). CeD is caused by the consumption of gluten proteins present in cereals such as wheat, barley and rye in genetically susceptible individuals (3). While many genes are involved in the development of CeD, thus far only the presence of HLA-DQ2 or DQ8 haplotype is considered to be essential (4). Additional factors that contribute to pathogenesis include other co-genetic factors (genome wide association studies have identified several markers), wheat-related factors (age of ingestion, type and quantity of wheat) and the way gluten is metabolized in the intestine (5,6). About 30-40% of the gluten protein consists of glutamine and proline. Since humans are unable to enzymatically break the molecular bonds between these two amino-acids, many immunogenic peptides are produced (5). There remains a possibility that enzymes secreted by the small intestinal microbiota convert some of these immunogenic peptides to non-immunogenic peptides.

While 20-30% of individuals in many countries including India are genetic susceptibility to develop CeD and the majority of them are exposed to wheat, only 1% of them develop CeD. This brings forth the role of other factors such as the gut microbiota in the pathogenesis of CeD (7). Recently, numerous studies have highlighted the potential role of gut microbiota in inflammatory gastro-intestinal diseases (4,8–12,12–29).

Whether these changes in the microbial community structure and function in patients with CeD are cause or effect of the disease state however remains unclear to date. In order to answer this question, one has to examine the status of the gut microbiota in the pre-disease state. Such

observations have not yet been possible due to the lack of well-established animal models for CeD. While 70-80% percent of first-degree relatives (FDRs) of patients with CeD have HLADQ2/DQ8 haplotype (compared to 30% in the general population); only approximately 8.5% of FDRs develop CeD (30). Thus, the question arises; Why do only few FDRs develop CeD and what is the role of the gut microbiome in disease protection? Indirect evidence of altered microbiota in relatives of patients with CeD is suggested by significantly lower levels of acetic and total short chain fatty acids, and higher fecal tryptic activity (31). Nevertheless, to date there is no information on the gut microbial composition and function in FDRs of patients with CeD, especially using the latest sequencing approaches. Additionally, it is important to explore the status of the microbiota in both the small intestine, the site of the disease, and faeces, as representative of whole gut microbiome.

To test the hypothesis that gut microbiome of FDR is different from CeD and could potentially play an important role in the pathogenesis of CeD, we explored the composition of both small intestinal and the whole gut microbiome using Illumina MiSeq in a subset of patients with CeD, first degree relatives and controls. We further investigated the potential microbial functions that are characteristic of FDR and CeD microbiota.

Patients and Methods

Human subjects, duodenal biopsies and faecal sample collection

A total of 62 subjects participated in this study including 23 treatment naïve patients with CeD [all HLA-DQ2/DQ8+, having high titre of anti-tissue transglutaminase antibodies (tTG Ab) and having villous abnormalities of modified Marsh grade 2 or more], 15 healthy first-degree relatives of patients with CeD [having normal titre of anti-tTG Ab and having no villous

abnormalities of modified Marsh grade 0 or 1], and 24 controls (patients with Hepatitis B Virus carriers or those having functional dyspepsia; having normal titre of anti-tTG Ab and having no villous abnormalities) (Table 1). Duodenal biopsies and faecal samples were collected from each of the above mentioned subjects at All India Institute of Medical Sciences, New Delhi, and sent to National Centre for Cell Sciences, Pune for microbiome analysis. The ethics committees of All India Institute of Medical Sciences, New Delhi, and National Centre for Cell Sciences, Pune, India approved the study. Informed and written consent was obtained from all the participants. Further details of patients and controls have been provided in the (Additional file 1: Table S1.

DNA extraction and 16S rRNA gene sequencing

Total DNA was extracted from duodenal biopsies using QIAGEN DNeasy Blood and Tissue kit (QIAGEN, Germany) and faecal samples using the QIAamp fast DNA stool Mini Kit (QIAGEN, Germany) according to the manufacturer's instructions. We used Illumina MiSeq sequencing to determine the bacterial composition of the duodenal biopsies and faecal samples. PCR was set up in 50 µl reaction using AmpliTaq Gold PCR Master Mix (Life Technologies, USA) and with 16S rRNA V4 variable region specific bacterial primers 515F (5´-GTGCCAGCMGCCGCGGTAA- 3´) and 806R (5´- GGACTACHVGGGTWTCTAAT-3´)

Bacterial community analysis

QIIME (v1.8.0) was used to process the raw data files from the sequencer and construct the OTU table (32). Closed reference based OTU picking was used to cluster reads into OTUs at 97% sequence similarity using the UCLUST algorithm and a representative sequence from each OTU was selected for downstream analysis. All OTUs were assigned to the lowest possible taxonomic rank by utilizing RDP Classifier 2.2 and Greengenes database 13.8 with a confidence score of at least 80% (33). The OTU table was filtered to remove OTUs belonging to

Mitochondria, Chloroplasts and Archaea. Alpha diversity and beta diversity analysis and composition analysis was done using the R-package phyloseq (v1.21) (34) and microbiome R package (v1.1.10008) (35). The alpha diversity measures were tested for significance using the Wilcoxon test for pairwise comparisons. OTU counts were transformed to relative abundance and Analysis of Similarities (ANOSIM) on Bray-Curtis distances was used to test for similarities in bacterial communities between sample types and diagnosis groups. Canonical correspondence analysis (CCA) was done separately on $\log10_{(1+X)}$ transformed microbial count data for faecal and duodenal biopsy samples to test for differences between the different diagnosis groups. PREMANOVA was carried out using the adonis function in vegan package (v 2.4-4) to calculate significance of PCoA clustering based on the Bray-Curtis distances (36). We used DESeq2 (v1.18.0) for identifying differentially abundant OTUs in pairwise comparisons between diagnosis groups (37). All OTUs that were significantly (alpha = 0.01) different in abundance between the diagnosis groups were reported and were adjusted for multiple comparisons using the Benjamini-Hochberg, false discovery rate procedure. Data was visualized using ggplot2 (v 2.2.1) in R (34). All scripts and files to reanalyse microbiota profiling data from this study are available at https://github.com/microsud/Gut-microbiota-Celiac-disease.

Predicted metagenome analysis: We used PICRUSt to infer community gene content from 16S rRNA gene amplicon sequencing data, followed by STAMP to get significantly different predicted genes in diagnosis groups. For comparison of taxa contributions to the inferred metagenomes we used a recently devloped tool, BURRITO (Browser Utility for Relating micRobiome Information on Taxonomy and functiOn) available from (https://elbo-spice.gs.washington.edu/shiny/burrito/).

Results

Characteristics of the study cohort:

The characteristics of the study subjects have been summarized in the Table 1.

Comparison of faecal and duodenal microbial community in the study cohort

Having both duodenal biopsies and faecal samples provided the opportunity to investigate differences in both site-specific and whole gut bacterial diversity and community structure in patients with CeD, FDRs and controls. The microbial community was significantly different between the faecal and duodenal biopsies irrespective of whether they were from CeD, FDR or DC groups (Fig. 1a, Analysis of similarities; Anosim test*;* R-statistic = 0.45, P−value = 0.001). Hence, further analyses were carried out separately for faecal and duodenal samples in different groups.

Analysis of alpha diversity using Shannon index between the sampling sites suggested significant differences between the sampling sites (Wilcoxon test $P < 0.05$). The duodenal biopsies showed larger inter-individual variation and consisted of samples with high and low diversity communities in all groups (Fig. 1b).

Site specific bacterial community structure in FDRs, CeD and controls

Duodenal and faecal microbiota composition and structure is distinct in FDRs, CeD and control groups:

To investigate if patients with CeD, FDRs or DC had site specific dissimilarities in microbiota composition, we tested duodenal and faecal samples separately. Bray-Curtis distance for microbiota in duodenal biopsy was not different between CeD, FDRs or DC (Analysis of similarities; Anosim test*;* R-statistic = 0.02192, $p = 0.202$). To analyse whether CeD and FDR have differences in microbial community, we performed PREMANOVA on ordination excluding

DC. Although there was no significant difference in total community (R^2 = 0.0608, Pr(>F)= 0.225), we identified the top ten variable OTUs between CeD and FDR microbiota (Additional file 2: Figure S1). These included four OTUs classified as *Acinetobacter*, of which two (OTU-562618, OTU-543942) were associated with CeD, while the other two (OTU-1009894, OTU-988314) were associated with FDR. In addition, *Pseudomonas* (two OTUs), and *Stenotrophomonas* (one OTU) were associated with FDR. *Corynebacterium*, *Commamonas* and *Novosphingobium* were differentiating genera in CeD. Canonical correspondence analysis (CCA) constrained for diagnosis status revealed that the clustering of subjects was marginally different in the duodenal bacterial community from FDRs, CeD and controls (F = 1.185, $Pr(>F)$ = 0.031 , Fig. 2a). Bray-Curtis distances for microbiota in faeces were not different between CeD, FDRs or DC (Analysis of similarities; Anosim test; R-statistic = 0.032, p = 0.125). In addition, the CCA analysis suggested no significant differences between faecal microbiota of FDRs, CeD and controls (F = 1.0704, $Pr(>F)$ = 0.154, Fig. 2b).

Alpha diversity measures were independently calculated for duodenal biopsies and faecal samples and compared between FDRs, CeD and controls (Fig. 3a and 3b). Pairwise comparisons of alpha diversity in duodenal biopsies between FDRs, CeD and controls suggested no significant differences (Fig. 3a). In faecal samples, pairwise comparison of alpha diversity demonstrated significant differences between FDR and CeD subjects (Wilcoxon test, FDR adjusted p = 0.029) (Fig. 3b).

Taxonomic differences in microbiota from duodenal biopsies of FDRs CeD, and controls:

At phylum level, there were no statistically significant differences in the major phyla (Proteobacteria, Firmicutes, Bacteroidetes and Verrucomicrobia) in the duodenal biopsies of FDRs, CeD and controls (Additional file 3: Figure S2), however, inter-individual variation

within the diagnosis groups was detected. At OTU level, after filtering out for low abundance OTUs (less than 5 counts in 50% of the samples in each group), CeD and FDR shared 124 OTUs, CeD and DC shared 25 OTUs, FDR and DC shared 23 OTUs and a total of 447 were shared by all three groups (Additional file 4: Figure S3).

To further investigate OTUs that are different between the diagnosis groups, FDR vs DC, CeD vs DC and FDR vs CeD, we used the DESeq2 default parameters. The enriched OTUs described in the following section were significantly different between the groups (BH adjusted, $P < 0.01$).

Difference in OTUs between FDRs and DC in the duodenal biopsies

In total, 52 OTUs belonging to 15 genera and one unclassified genera were identified as differentially abundant in between FDR and DC. Of these, four (*Streptococcus*, *Pseudomonas*, *Parvimonas* and *Acinetobacter*) were differentially abundant in FDR and 12 (*Sporosarcina*, *Planomicrobium*, *Planococcus*, *Lysinibacillus*, unclassified genus from Planococcaceae, *Enhydrobacter*, *Lactobacillus*, *Comamonas*, *Desemzia*, *Bacillus*, *Anoxybacillus*, and an unclassified genus were differentially abundant in DC (Fig. 4a).

Difference in OTUs between CeD and DC in the duodenal biopsies

Comparison of duodenal microbiota of CeD with that of the controls identified 106 differentially abundant OTUs associated with 22 genera. Amongst these, several OTUs from the genera *Acinetobacter*, Lactobacillaceae, *Corynebacter*, and one OTU each for *Prevoella* and *Pseudomoans* were abundant in CeD (Fig.4b).

Difference in OTUs between FDRs and CeD in the duodenal biopsies

Comparison of FDR with CeD identified 41 OTUs belonging to 12 genera that were significantly different (Fig.4c). These included OTUs classified as genus *Streptococcus, Stenotrophomonas, Acinetobacter, Mogibacterium,* Enterococcaceae, *Atopobium,* unclassified Coriobacteriaceae, *Brevundimonas, Bacillus, Actinomyces and Parvimonas,*were abundant in FDR, other OTUs within *Acinetobacter* were abundant in CeD group. In addition, CeD had higher abundance of *Corynebacterium* and OTUs members of *Lactobacillaceae*.

Taxonomic differences in the faecal microbiota in patients with CeD, FDRs and controls

Similar to duodenal biopsies, at phylum level, there was not statistically significant difference in the major phyla (Proteobacteria, Firmicutes, Bacteroidetes and Verrucomicrobia) in fecal samples of patients with CeD, FDRs and controls. (Additional file 5: Figure S4). At OTU level, after filtering for low abundance OTUs (counts less than 5 in 50% of the samples in each group), CeD and FDR shared 1 OTU, CeD and DC shared 7 OTUs, FDR and DC shared 101 OTUs and a total of 110 were shared by all three groups (Additional file 6: Figure S5).

Differences in OTUs between FDRs and DC in faeces

A total of 30 OTUs from 12 genera were differentially abundant between FDRs and DC. Among these, *Streptococcus, Prevotella and Acinetobacter* were abundant in DC, while *Ruminococcus*, unclassified Ruminococcaceae, *Pseudomonas*, *Lysinibacillus*, unclassified Planococcaceae, *Lactobacillus*, *Enterococcus*, unclassified Enterobacteriaceae, unclassified Coriobacteriaceae, *Comamonas*, *Bacteroides*, *Bacillus* and an unclassified genus were abundant in FDR (Fig. 5a).

Differences in OTUs between CeD and DC in faeces

Comparison of CeD and DC groups revealed a total of 86 OTUs from 12 genera that were differentially abundant. The OTUs classified to genus unclassified Planococcaceae *Weissella*, unclassified Coriobacteriaceae, unclassified Christensenellaceae, *Bacteroides* and *Bacillus* were abundant only in CeD. On the contrary, unclassified Clostridiaceae, unclassified Enterobacteriaceae *Coprococcus*, *Acinetobacter*, unclassified Peptostreptococcaceae, *Trabulsiella*, *Turicibacter*, unclassified Ruminococcaceae and unclassified genus were abundant in in DC (Fig. 5b).

Difference in OTUs between FDRs and CeD in faeces

A total of 81 OTUs belonging to 13 genera including one unclassified genus were differentially abundant between FDR and CeD. Genus *Streptococcus*, *Prevotella*, unclassified Leuconostoceaceae, *Lactobacillus*, , unclassified Lachnospiraceae, *Eubacterium* (Erysipelotrichaceae), *Enterococcus*, unclassified Enterococcaceae, unclassified Enterobacteriaceae, *Collinsella*, *Bacteroides* and were abundant in CeD, while *Comamonas*, *Lysinibacillus*, *Serratia*, *SMB53*, *Trabulsiella*, unclassified Peptostreptococcaceae and unclassified Ruminococcaceae were abundant in FDR (Fig. 5c).

Imputed metagenome of FDR and CeD duodenal microbiome shows reduced proportion of genes involved in gluten metabolism in comparison to that of the controls

In addition to differentially abundant microbial taxa, different study groups might have altered microbial community functions by enriching or depleting taxa that encode specific metabolic modules. Of specific interest were the enzymes related to peptidases as they play a role in the breakdown of gliadin residues. For an overview of taxon contributions to this class of enzymes, we used a recently developed tool called BURRITO (Browser Utility for Relating micRobiome Information on Taxonomy and functiOn) (https://elbo-

spice.gs.washington.edu/shiny/burrito/). This tool employs PICRUst to infer metagenome predictions using 16S rRNA gene data (38). The differential abundances of *Acinetobacter* and *Pseudomonas* were also correlated to the enrichment of peptidases in the total community (Fig. 6).

The proportion of pyroglutamyl peptidase [3.4.19.3], subtilisin [3.4.21.62] and x-pro dipeptidase [3.4.13.9] genes which are involved in gluten degradation were reduced in FDRs and CeD as compared to controls (Fig. 7a,b,c). Moreover, aminopeptidase [3.4.11] was also reduced in FDRs and CeD (Fig. 7d).

Genes for beta-N-acetylhexosaminidase [3.2.1.52] and lysozyme [3.2.1.17] were predicted to occur in higher abundance in CeD than that in FDRs and DC (Additional file 7: Table S2). Higher proportions of beta-N-acetylhexosaminidase enzymes might cause or be the reflection of a weak mucosal barrier and indicate an enriched, mucolytic bacterial population.

Discussion

The aim of the present study was to investigate differences in the duodenal and faecal microbiota of FDRs compared to CeD and controls. The FDR group was included for two main reasons: 1) They represent a population which is genetically-susceptible to develop CeD; 2) They provide a unique opportunity to identify features of the host as well as of the associated microbiota that may be involved in the protection against developing CeD. We collected both duodenal biopsies and fecal samples to investigate both local and overall changes in the microbiota in FDR, patients with CeD and controls.

The differences in microbial diversity and community structure between the small intestinal microbiome represented by biopsy samples and whole gut microbiome represented by faecal samples were significant. To the best of our knowledge, reports on site specific microbiota patterns in patients with CeD remain scarce, and no results on both site specific and whole gut microbiome on FDRs have been reported to date. Present study provides an overall view on differences of both site-specific changes as well as changes in the faecal microbiota of FDRs, CeD and DC.

The duodenal microbial community structure of FDR is more similar to CeD than that of DC. In addition, both FDR and CeD individuals have high abundances of *Acinetobacter* and *Pseudomonas*. In addition, inferred genus-function relationship demonstrated these two genera to contribute to increased peptidases in the duodenal microbial community. Both *Acinetobacter* and *Pseudomonas* include opportunistic pathogenic species that are linked to inflammation (39,40). *Pseudomonas* is reported to be higher in children with CeD (41). On the other hand, the duodenal biopsies of FDR showed lower numbers of OTUs of *Lactobacillus* compared to DC and CeD. In a previous study, higher abundance of *Lactobacilus* was observed with higher glutenase activitiy in oral microbiome of patients with CeD (14). A lower number of *Lactobacillus* in FDRs in the present study may indicate their reduced ability to breakdown gluten into pro-inflammatory peptides in their small intestine. This is also supported to some extent by the observation of lower abundance of pyro-glutamyl peptidase [EC 3.4.19.3] in the predicted metagenome of the duodenal biopsies of FDR in the present study.

Interestingly, *Corynebacterium* was abundant in patients with CeD compared to both the FDR and DC (Fig. 4b and c). However, in the comparison between DC and FDR, abundance of *Corynebacterium* was not detected to be differentially abundant (Fig. 4a). Previously, the genus

Corynebacterium was reported to be present in high abundance in infants with higher risks for developing CeD (42). This suggests the need to further investigate the role of bacteria from this genus in CeD. Another observation is the higher abundance of OTUs of *Streptococcus* in FDR when compared to both the DC and CeD. Both *Lactobacillus* and *Streptococcus* are lactic acid bacteria and known to have peptidase activity to breakdown gliadin peptides (43). Further, mechanistic investigations will be necessary to ascertain, if there is a trade-off for abundance of specific strains of *Lactobacillus* or *Streptococcus* in subjects genetically susceptible to developing CeD. Another important difference between FDR and CeD is the higher abundance of *Actinomyces* in FDR. Specific strains of *Actinomyces* are shown to breakdown the highly immunogenic α-gliadin 33-mer peptide (44). These observations suggest that the FDR and CeD duodenal microbiota differs in the bacterial composition and that loss or gain of specific bacteria capable of creating immunogenic or non-immunogenic gliadin peptides may be crucial.

Recently, a serine endopeptidase, subtilisin a novel class of gluten-degrading enzyme belonging to the S8 family of peptidases has been described. This enzyme is able to cleave and abolish gluten immunogenic epitopes (45). In this study, the gene proportion for this enzyme was found to be depleted in FDR as compared to the control. Moreover, genes coding for Xaa-Pro dipeptidase (EC:3.4.13.9) (prolidase) which assist in gluten degradation by splitting dipeptides with a prolyl residue in the carboxyl terminal position was reduced in FDR in comparison to that in controls. Also, gene proportion for aminopeptidase which may help in degradation of gluten was found to be decreased proportion in FDRs. In summary, results from differential taxa abundance analysis and predicted metagenome suggest that the microbiome in the small intestine of FDR might be different from that of CeD and DC in the way gluten is metabolized.

In the faecal samples, the alpha diversity indices of FDR and DC were more similar to each other compared to that of CeD. Higher diversity was observed in FDR compared to CeD ($p = 0.034$). In patients with active CeD, there is high inflammation in the intestine and such an environment is known to be strongly associated with reduced richness of microbial community (46,47). Higher diversity is often linked to higher resilience of the intestinal microbiome and a barrier to invasion by pathogenic species (47,48). The comparable and higher diversity of faecal microbiota in FDR and controls may suggest that the intestinal microbiota may act as a barrier to invasion of pro-inflammatory bacteria in FDRs.

In contrast to observations of duodenal biopsies, both FDR and CeD had lower abundance of *Acinetobacter* compared to DC in the faeces. However, *Pseudomonas* was abundant in FDR compared to DC. In addition, OTUs for *Bacillus* and *Bacteroides* were abundant in FDR and CeD. However, one specific OTU (OTU-535375) classified as *Bacteroides fragilis* was significantly abundant in CeD compared to FDR. Previously in *in-vitro* condition nine different strains of *B. fragilis* showed gliadin-hydrolysing activity and some of them generated immunogenic peptides that increased inflammatory cytokine production and showed increased ability to permeate through Caco-2 cell cultures (18). These observations suggest the need for investigating the role higher abundance of specific *Bacteroides fragilis* strains in CeD.

The FDR and DC do not show differences in abundance of many known and predicted butyrate, acetate and propionate producing bacteria from family Ruminococcaceae and Lachnospiraceae. However, CeD fecal microbiota has significantly low abundance of these compared to DC group. Reduction in butyrate producing bacteria was previously reported in inflammatory bowel disease (49).

This present study was conducted to investigate if the duodenal and faecal microbiota of FDR of CeD patients are different from that of controls and CeD. At microbial community level, the structure of FDR was more similar to CeD than that of DC. In fecal samples the FDR microbiota was characterised by higher species diversity. Several OTUs could differentiate microbiota of FDRs from that of CeD and DC. These OTUs belong to genera that are known to have different abilities to breakdown gluten as well as some with ability to produce butyrate.

We used 16S rRNA gene sequencing which gives sufficient power to obtain the microbial community profile (50). To investigate potential functions (specifically those related to gluten metabolism), we use predictive metagenomics (38). Our observations from beta diversity (Additional file 1: Figure S1) and differential abundance (Figure 4 and 5) suggests variation at taxonomic levels lower than genus. These potential strain level variations and functional aspects using metagenomics and functional omics need to be investigated in follow-up studies. However, metagenomics studies of biopsy samples remain a challenge because of high proportion of host DNA. Thus, predictive metagenomics using 16S rRNA gene as a practical solution was employed for biopsies. In this initial exploratory study, we investigated the gut microbiome with respect to the disease status only and future studies considering other confounding factors such as diet, body mass index age, sex, frequency and quantity of gluten intake among others will be required for a better understanding the gut microbiome in CeD and FDRs. Nonetheless, the study provides support to the emerging view that the gut microbiome and its function could possibly have a pivotal role in the etiopathogenesis of CeD.

In summary, present study highlights the specific differences in the microbiota of FDR compared to that in patients with CeD and controls. Difference in FDR microbiota in both the faecal and duodenal biopsy samples compared to CeD and DC suggests microbiota of FDR have

unique features. These unique features should be addressed in future mechanistic studies to understand etiopathogenesis of CeD.

List of abbreviations

CeD: Celiac disease, **DC**: Diseased controls (dyspeptic), FDR: First degree relatives. **OTU**: Operational taxonomic unit, **PERMANOVA**: Permutational multivariate analysis of variance, **rRNA**: Ribosomal Ribonucleic acid, **PCoA**: Principal coordinates analysis, **CCA**: Canonical correspondence analysis.

Declarations

Acknowledgement

We acknowledge the help of Aishwairya Sharma, a dietician, who counselled the patients for gluten free diet. The authors are grateful to Dr. Nachiket Marathe for fruitful discussion related to the project. The authors thank Johanna Gutleben for proofreading the manuscript. We sincerely acknowledge the support of each participant of this study who accepted to participate in this study.

Funding

We acknowledge the support of the Food and Nutrition Division of Department of Biotechnology, Government of India who provided the research grant for this study (Sanction no BT/PR-14956/FNS/20/497/2010).

Availability of data and materials

Sequence data generated in this study is available from the NCBI Sequence Read Archive within the Bioproject ID accession PRJNA385740. (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA385740) and to reproduce the analysis

done in R, the R Markdown file and required data are available at https://github.com/microsud/Gut-microbiota-Celiac-disease.

Conflicts of interest: All the authors disclose no conflict of interest

Authors Contributions:

The research study was conceptualized, designed and supervised by GKM, YSS and VA. Patient recruitment, diagnosis and endoscopic examination was done by GKM; HLA testing was done by GK; biological sample collection (duodenal biopsy/stool) storage and maintenance was done by AKV, KB and AM. The extraction of genomic DNA was done by RB and PP. DKB, BPS and RCP were involved in amplicon sequencing. Bioinformatics analysis for amplicon data was done by SAS, DPD and RB. Data acquisition, data interpretation and drafting of the manuscript was done by SAS and GKM. YSS, DPD and VA critically reviewed the manuscript. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

The Ethics Committees of All India Institute of Medical Sciences, New Delhi, and National Centre for Cell Sciences, Pune, India approved the study. Informed and written consent was obtained from all the participants.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References:

1. Singh P, Arora A, Strand TA *et al.* Global prevalence of celiac disease: Systematic review and Meta-analysis. Clin Gastroenterol Hepatol (In press).

2. Makharia GK, Verma AK, Amarchand R *et al.* Prevalence of celiac disease in the northern part of India: A community based study. J Gastroenterol Hepatol. 2011;26:894–900.

3. Caminero A, Nistal E, Herrán AR *et al.*Differences in gluten metabolism among healthy volunteers, coeliac disease patients and first-degree relatives. Br J Nutr. 2015;114:1157–67.

4. Sanz Y, Pama G De. Unraveling the Ties between Celiac Disease and Intestinal Microbiota. Int Rev Immunol. 2011;30:207–18.

5. Kagnoff MF, Kagnoff MF, Green PH *et al.* Celiac disease: pathogenesis of a model immunogenetic disease. J Clin Invest. 2007;117:41–9.

6. Jabri B, Sollid LM. Mechanisms of Disease: immunopathogenesis of celiac disease. Nat Clin Pract Gastroenterol Hepatol. 2006;3:516–25.

7. Sánchez E, Donat E, Ribes-Koninckx C *et al.* Intestinal Bacteroides species associated with coeliac disease. J Clin Pathol. 2010;63:1105–11.

8. Mohan M, Chow CET, Ryan CN *et al.*Dietary Gluten-Induced Gut Dysbiosis Is Accompanied by Selective Upregulation of microRNAs with Intestinal Tight Junction and Bacteria-Binding Motifs in Rhesus Macaque Model of Celiac Disease. Nutrients. 2016;8:684.

9. Sánchez E, Donat E, Ribes-Koninckx C *et al.* Duodenal-mucosal bacteria associated with celiac disease in children. Appl Environ Microbiol. 2013;79:5472–9.

10. Tjellström B, Stenhammar L, Högberg L *et al.* Gut microflora associated characteristics in children with celiac disease. Am J Gastroenterol. 2005;100:2784–8.

11. de Sousa Moraes LF, Grzeskowiak LM, de Sales Teixeira TF *et al.* Intestinal microbiota and probiotics in celiac disease. Clin Microbiol Rev. 2014;27:482–9.

12. Nistal E, Caminero A, Vivas S *et al.* Differences in faecal bacteria populations and faecal bacteria metabolism in healthy adults and celiac disease patients. Biochimie. 2012;94:1724–9.

13. Collado MC, Calabuig M, Sanz Y. Differences between the fecal microbiota of coeliac infants and healthy controls. Curr Issues Intest Microbiol. 2007;8:9–14.

14. Tian N, Faller L, Leffler DA *et al.* Salivary gluten degradation and oral microbial profiles in healthy individuals and celiac disease patients. Appl Environ Microbiol. 2017;83: e03330-16.

15. Francavilla R, Ercolini D, Piccolo M *et al.*Salivary microbiota and metabolome associated with celiac disease. Appl Environ Microbiol. 2014;80:3416–25.

16. Golfetto L, Senna FD de, Hermes J *et al.*. Lower Bifidobacteria Counts in Adult Patients With Celiac Disease on a Gluten-Free Diet. Arq Gastroenterol. 2014;51:139–43.

17. Cheng J, Kalliomäki M, Heilig HG *et al.* Duodenal microbiota composition and mucosal homeostasis in pediatric celiac disease. BMC Gastroenterol. 2013;13:113.

439 18. Sánchez E, Laparra JM, Sanz Y. Discerning the role of bacteroides fragilis in celiac
440 disease pathogenesis. Appl Environ Microbiol. 2012;78:6507–15.

441 19. Béres NJ, Sziksz E, Vannay Á *et al.* Role of the Microbiome in Celiac Disease. Int J
442 Celiac Dis Attila J Szabó, Gábor Veres Int J Celiac Dis. 2014;2:150–3.

443 20. Kopečný J, Mrázek J, Fliegerová K *et al.* The intestinal microflora of childhood patients
444 with indicated celiac disease. Folia Microbiol.2008;53:214–6.

445 21. Pisarello MLJ, Vintiñi EO, González SN *et al.* Decrease in lactobacilli in the intestinal
446 microbiota of celiac children with a gluten-free diet, And selection of potentially probiotic
447 strains. Can J Microbiol. 2015;61:32–7.

448 22. Caminero A, Herrán AR, Nistal E *et al.* Diversity of the cultivable human gut microbiome
449 involved in gluten metabolism: isolation of microorganisms with potential interest for
450 coeliac disease. FEMS Microbiol Ecol. 2014;88:309–19.

451 23. Sánchez E, De Palma G, Capilla A *et al.* Influence of environmental and genetic factors
452 linked to celiac disease risk on infant gut colonization by Bacteroides species. Appl
453 Environ Microbiol. 2011;77:5316–23.

454 24. Nylund L, Kaukinen K, Lindfors K. The microbiota as a component of the celiac disease
455 and non-celiac gluten sensitivity. Clin Nutr Exp. 2016;6:17–24.

456 25. Schippa S, Iebba V, Barbato M *et al.* A distinctive “microbial signature” in celiac
457 pediatric patients. BMC Microbiol. 2010;10:175.

458 26. D’Argenio V, Casaburi G, Precone V *et al.* Metagenomics Reveals Dysbiosis and a
459 Potentially Pathogenic N. flavescens Strain in Duodenum of Adult Celiac Patients. Am J

Gastroenterol. 2016;111:879-90.

27. Anania C, Olivero F, Olivero E *et al.*Probiotics in Childhood Celiac Disease. In: World â€™s largest Science , Technology & Medicine Open Access book publisher Probiotics in Childhood Celiac Disease. 2016;95:223–39.

28. Nadal I, Donant E, Ribes-Koninckx C *et al.* Imbalance in the composition of the duodenal microbiota of children with coeliac disease. J Med Microbiol. 2007;56:1669–74.

29. Hedberg ME, Israelsson A, Moore ERB *et al.* Prevotella jejuni sp. nov., isolated from the small intestine of a child with coeliac disease. Int J Syst Evol Microbiol. 2013;63:4218–23.

30. Singh P, Arora S, Lal S *et al.*Risk of Celiac Disease in the First- and Second-Degree Relatives of Patients With Celiac Disease: A Systematic Review and Meta-Analysis. Am J Gastroenterol. 2015;110:1539–48.

31. Tjellström B, Stenhammar L, Högberg L *et al.* Gut microflora associated characteristics in first-degree relatives of children with celiac disease. Scand J Gastroenterol. 2007;42:1204–8.

32. Caporaso JG, Kuczynski J, Stombaugh J *et al.*QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7:335–6.

33. Wang Q, Garrity GM, Tiedje JM *et al.* Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl Environ Microbiol. 2007;73:5261–7.

34. McMurdie PJ, Holmes S. Phyloseq: An R Package for Reproducible Interactive Analysis

and Graphics of Microbiome Census Data. Watson M, editor. PLoS One. 2013 Apr 22;8:e61217.

35. Leo L, Shetty S. microbiome R package. Bioconductor. 2017.

36. Dixon P. VEGAN, a package of R functions for community ecology. Vol. 14, Journal of Vegetation Science. 2003;14:927–30.

37. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

38. Langille MGI, Zaneveld J, Caporaso JG *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol. 2013;31:814–21.

39. Thierry P, Mulet C, Catherine D. A Crypt-Specific Core Microbiota Resides in the Mouse Colon Thierry. MBio. 2012;3:e00116-12.

40. Hakansson A, Molin G. Gut microbiota and inflammation. Nutrients. 2011;3:637–82.

41. Cheng J, Kalliomäki M, Heilig HG *et al*. Duodenal microbiota composition and mucosal homeostasis in pediatric celiac disease. BMC Gastroenterol. 2013;13:113.

42. Olivares M, Neef A, Castillejo G *et al.* The HLA-DQ2 genotype selects for early intestinal microbiota composition in infants at high risk of developing coeliac disease. Gut. 2015;64:406–17.

43. Montserrat CA, Andreu CM, Ramon VD *et al.* Lactic acid bacteria for coeliac disease.2011; US Patents:US20130121976 A1.

44. Fernandez-Feo M, Wei G, Blumenkranz G *et al.* The cultivable human oral gluten-

degrading microbiome and its potential implications in coeliac disease and gluten sensitivity. Clin Microbiol Infect. 2013;19:E386-94.

45. Wei G, Tian N, Siezen R *et al.* Identification of food-grade subtilisins as gluten-degrading enzymes to treat celiac disease. Am J Physiol - Gastrointest Liver Physiol. 2016;311:G571–80.

46. Gevers D, Kugathasan S, Denson LA *et al.* The treatment-naive microbiome in new-onset Crohn's disease. Cell Host Microbe. 2014;15:382–92.

47. Shetty SA, Hugenholtz F, Lahti L *et al.* Intestinal microbiome landscaping: Insight in community assemblage and implications for microbial modulation strategies. FEMS Microbiol Rev. 2017;41:182–99.

48. Levine JM, D'Antonio CM. Elton Revisited: A Review of Evidence Linking Diversity and Invasibility. Oikos. 1999;87:15.

49. Wang W, Chen L, Zhou R *et al.* Increased proportions of Bifidobacterium and the Lactobacillus group and loss of butyrate-producing bacteria in inflammatory bowel disease. J Clin Microbiol. 2014;52:398–406.

50. Walters W, Hyde ER, Berg-lyons D *et al.* Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems. 2015;1:e0009-15.

522 **Table 1: Demographic characteristics on study subjects**

Features		**Groups**		
		CeD	**FDR**	**DC**
Number of subjects		23	15	24
Age (mean±S.D.)		23.4±9.5	31.6±10.8	30.6±12.3
Gender	Male	10	6	22
	Female	13	9	2
Sampling site	Faecal samples	21	15	23
	Biopsy samples	16	13	14
Villous abnormalities (as per Modified Marsh criteria)	0	-	15	22
	1	-	-	2
	3a	2	-	-
	3b	7	-	-
	3c	14	-	-
HLA DQ2/DQ8+		23	13	6
tTG Titre (mean±S.D.)		199.9±72.1	4.36±2.6	4.09±2.8

523

524 **Abbreviations:** CeD: Celiac disease patients; FDR: First degree relatives of CeD; DC: control

525 subject; tTG: tissue transglutaminase.

526

527 **Figure Legends**

Figure 1: Community wide differences in beta and alpha diversity. a. Principal coordinates analysis (PCoA) of bacterial diversity based on Bray–Curtis distance. **b.** Comparison of alpha diversity measures between sampling sites (Wilcoxon test was used for pairwise comparisons).

Figure 2: Constrained ordination of individuals using CCA with diagnosis as constraint a. Duodenal biopsies. **b.** Faecal samples.

Figure 3: Comparison of alpha diversity between diagnosis groups. a. Duodenal biopsy. **b.** Faecal samples. Wilcoxon test was used for statistical comparison.

Figure 4: **Comparison of differential abundance of bacterial taxa between the diagnosis groups in biopsy samples. a.** Differential abundance DC vs FDR **b.** Differential abundance CeD vs DC **c**. Differential abundance CeD vs FDR. Only genera with significant differences ($P < 0.01$) in log2 fold change are depicted.

Figure 5: Comparison of differential abundance of bacterial taxa between the diagnosis groups in faecal samples. a. Differential abundance DC vs FDR. **b.** Differential abundance CeD vs DC. **c.** Differential abundance CeD vs FDR. Only genera with significant differences ($P < 0.01$) in average log2 fold change are depicted.

Figure 6: Inferred genus-function relationships across duodenal biopsies demonstrating higher contribution of *Acinetobacter* and *Pseudomonas* to peptidase abundances in the total community.

Figure 7: Comparison of proportion of sequence percentage of specific genes related to gluten metabolism by bacteria inferred from predicted metagenome for duodenal biopsy. a. pyroglutamyl peptidase [EC 3.4.19.3] **b.** Subtilisin [3.4.21.62] **c.** X-Pro dipeptidase [3.4.13.9] **d.** Aminopeptidase [EC 3.4.11].

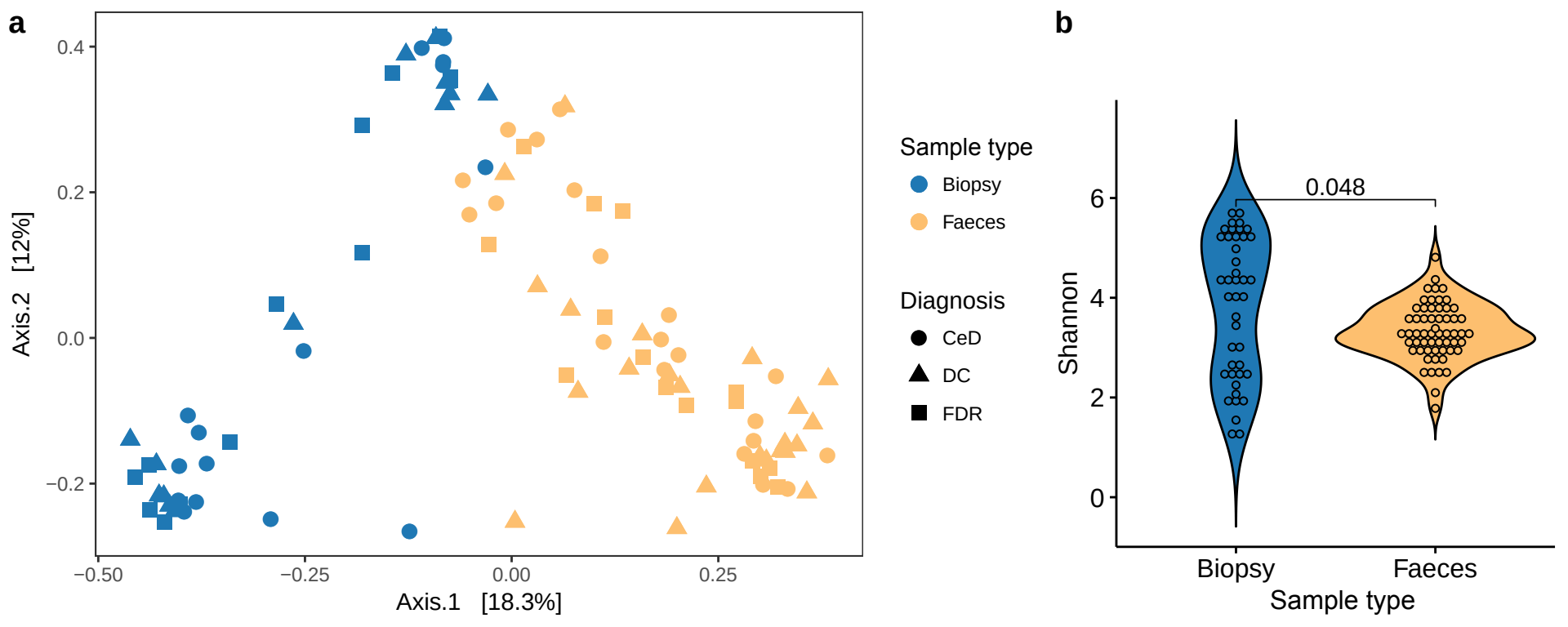
Additional file, Figure S1: Coefficients for the top OTUs separating CeD and FDR microbiota (PERMANOVA, R^2 = 0.03063, Pr(>F)= 0.45). Top half (Red) are OTUs characteristics of CeD group while bottom half (lavender) are OTUs characteristics of FDR group.

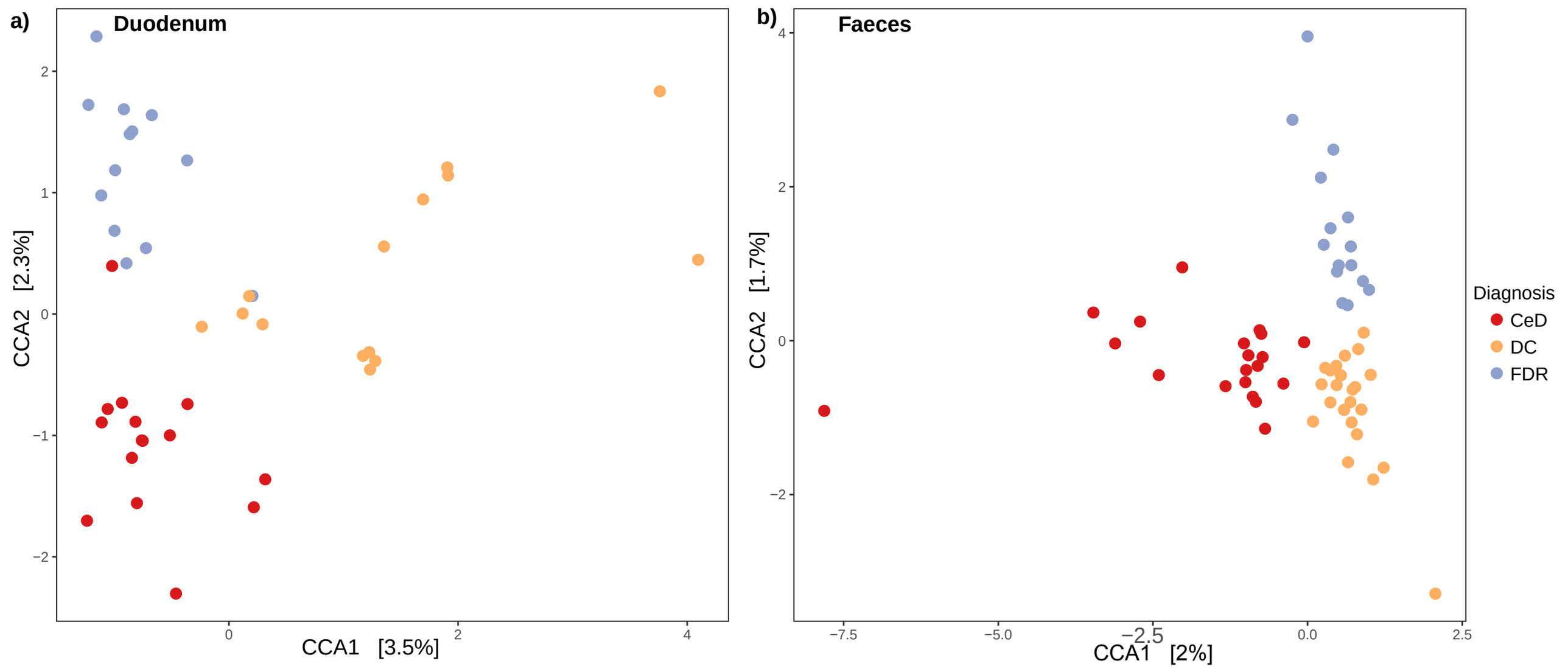
Additional file, Figure S2: Comparison of top four bacterial phyla in duodenal biopsy. Wilcoxon test was used for statistical comparison.

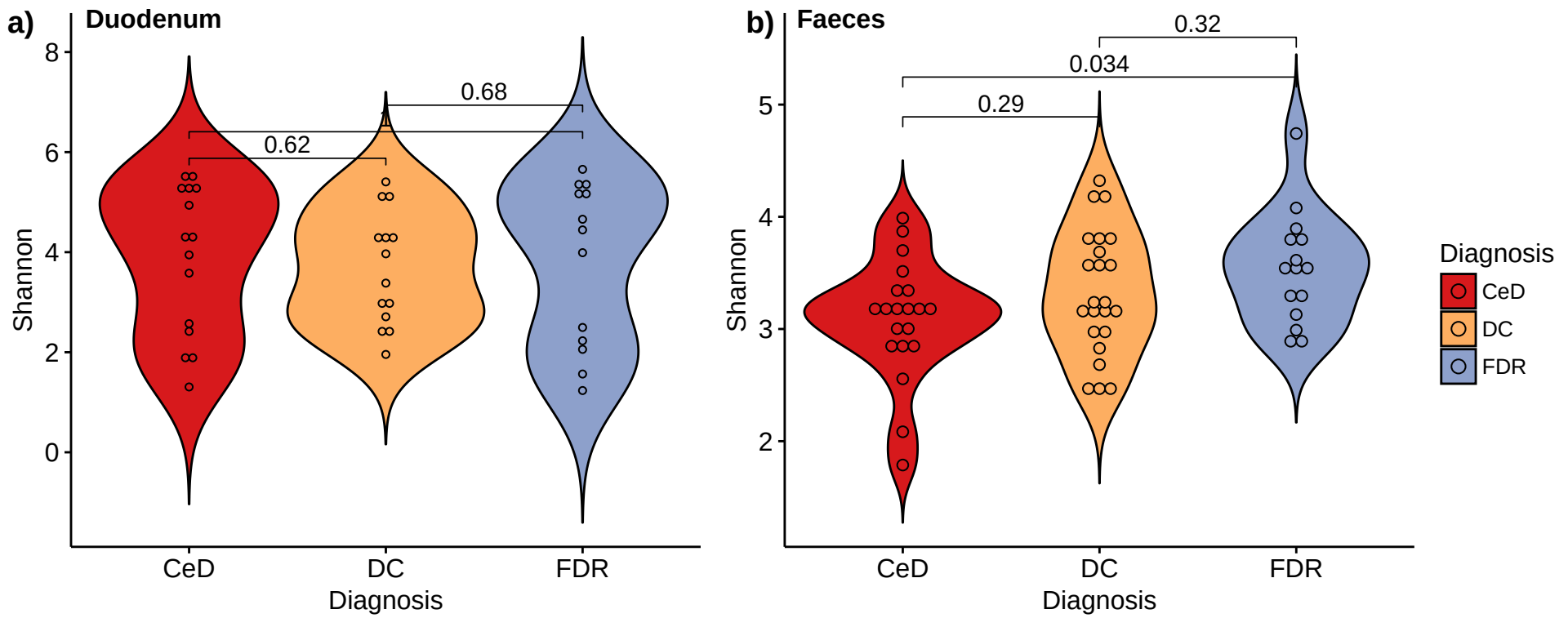
Additional file, Figure S3: Venn Diagram depicting shared and unique OTUs in duodenal biopsies between FDR, CeD and DC.

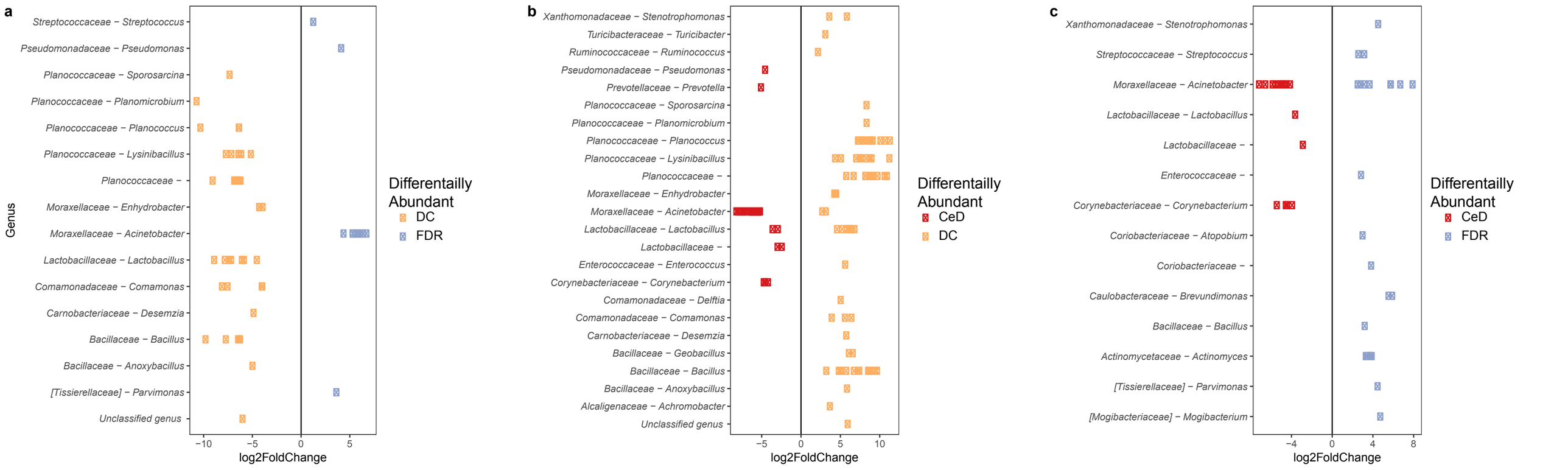
Additional file, Figure S4: Comparison of top four bacterial phyla in faeces. Wilcoxon test was used for statistical comparison.

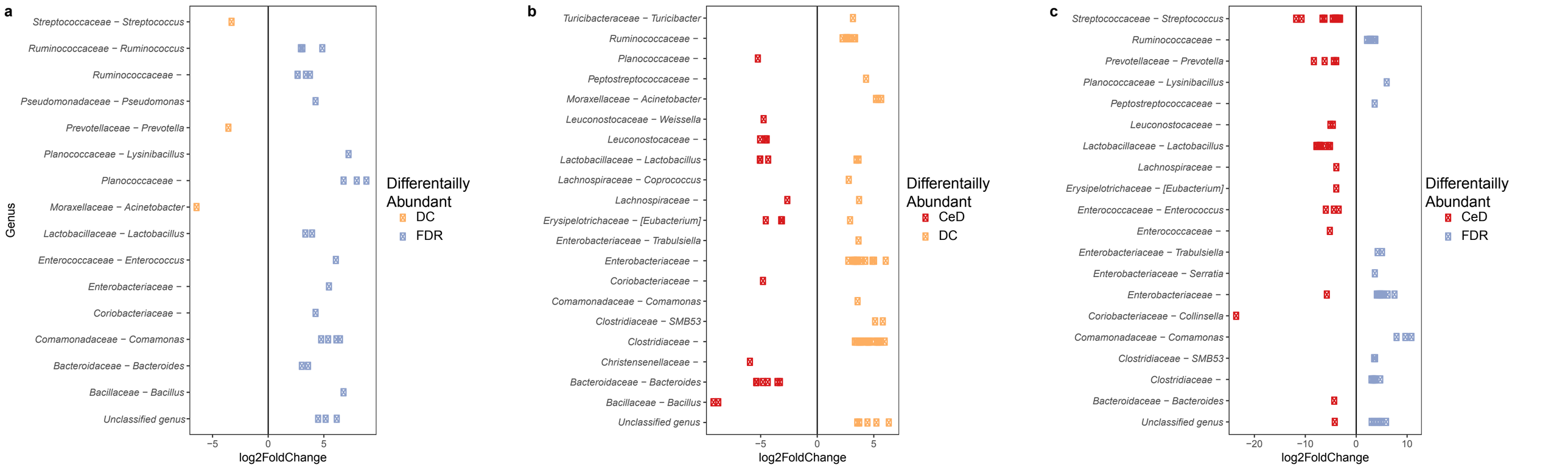
Additional file, Figure S5: Venn Diagram depicting shared and unique OTUs in faeces between FDR, CeD and DC.

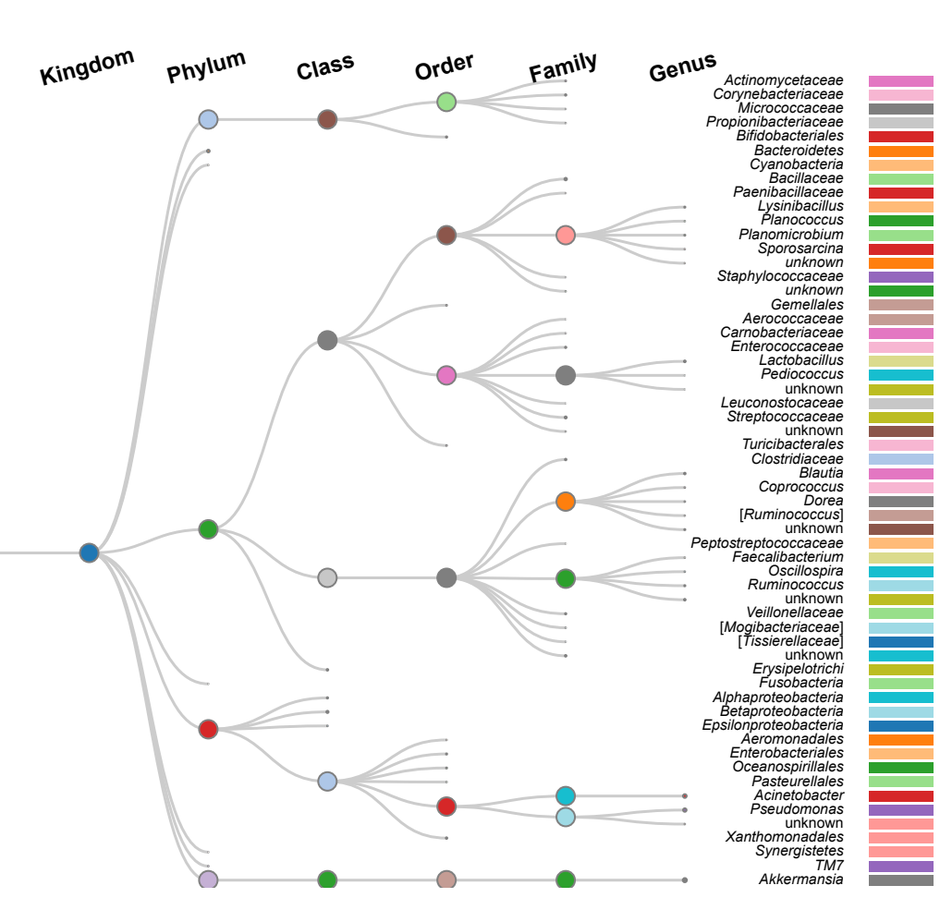

a
0.4
0.2
0.0
−0.2
Axis.2 [12%]
−0.50
−0.25
0.00
0.25
Axis.1 [18.3%]
Sample type
Biopsy
Faeces
Diagnosis
CeD
DC
FDR
b
Shannon
6
4
2
0
0.048
Biopsy
Faeces
Sample type

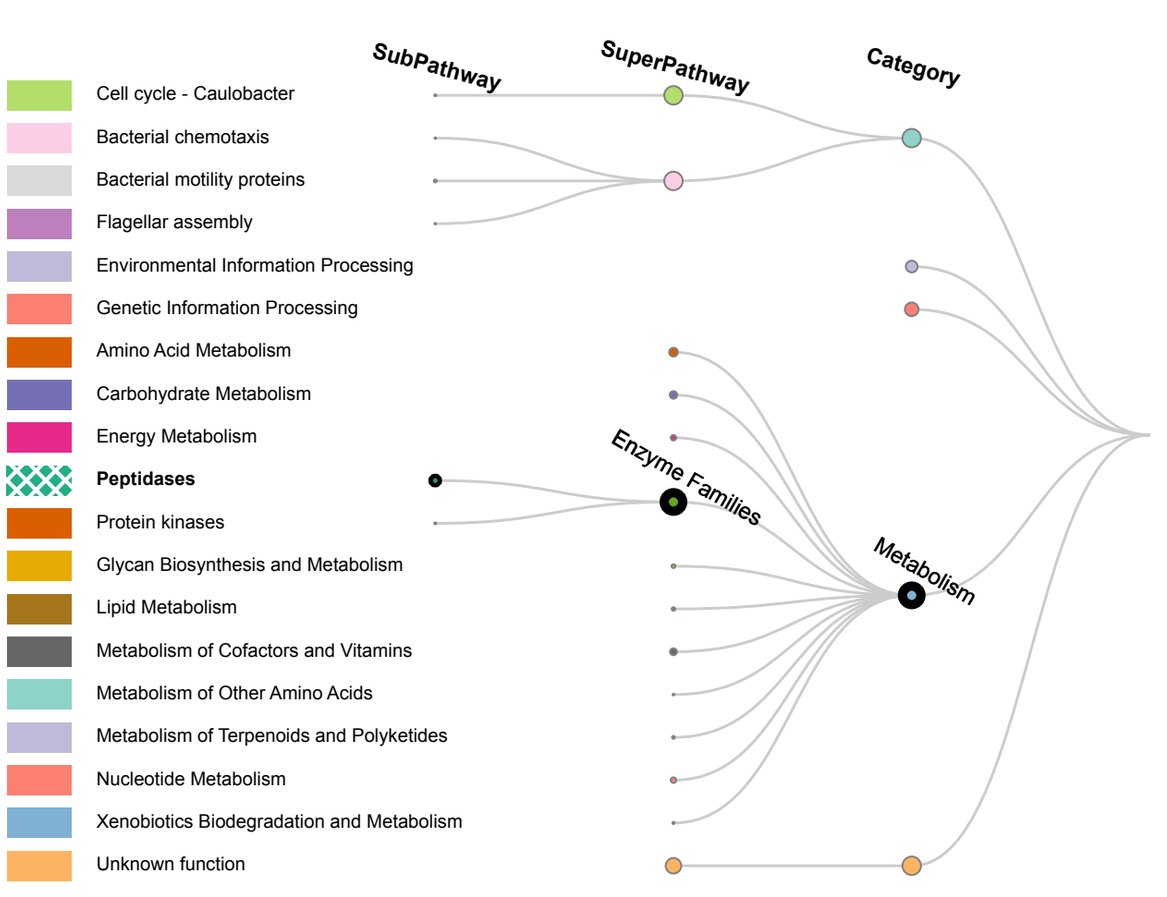

a)
Duodenum
CCA2 [2.3%]
CCA1 [3.5%]
b)
Faeces
CCA2 [1.7%]
CCA1 [2%]
Diagnosis
CeD
DC
FDR

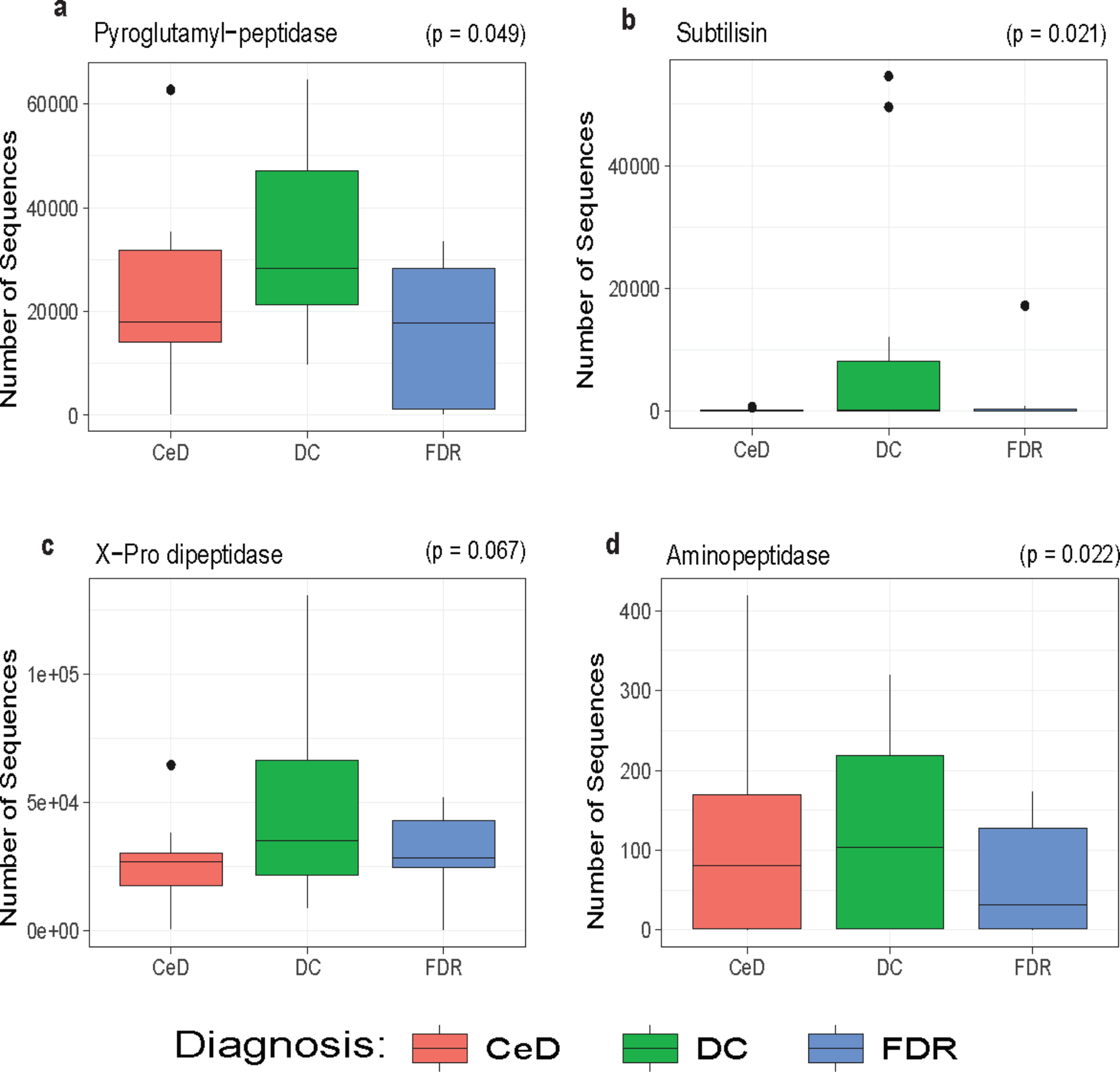

a)
Duodenum
Shannon
8
6
4
2
0
0.62
0.68
CeD
DC
FDR
Diagnosis
b)
Faeces
Shannon
5
4
3
2
0.29
0.034
0.32
CeD
DC
FDR
Diagnosis
Diagnosis
CeD
DC
FDR

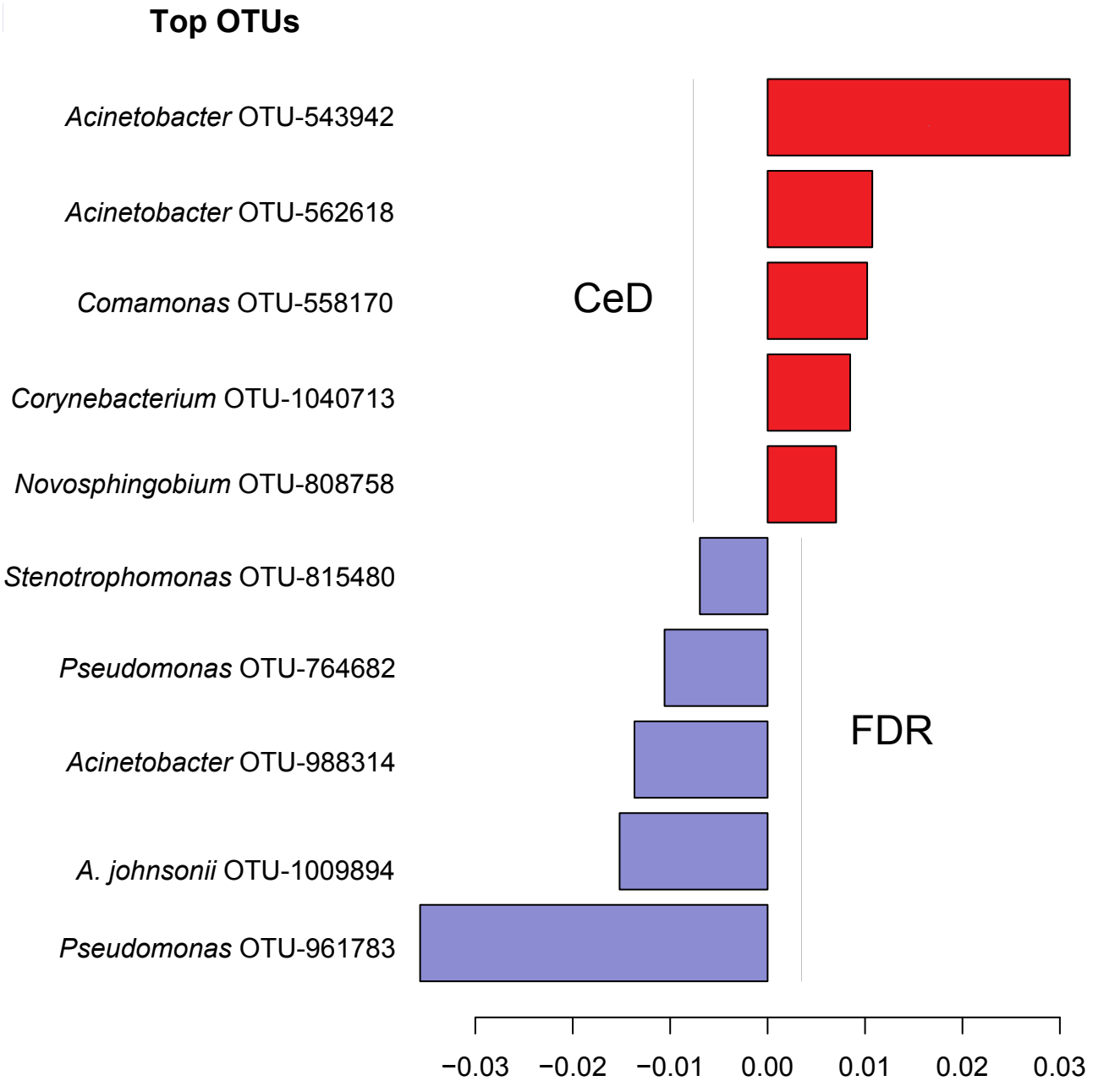

a
Streptococcaceae − Streptococcus
Pseudomonadaceae − Pseudomonas
Planococcaceae − Sporosarcina
Planococcaceae − Planomicrobium
Planococcaceae − Planococcus
Planococcaceae − Lysinibacillus
Planococcaceae −
Moraxellaceae − Enhydrobacter
Moraxellaceae − Acinetobacter
Lactobacillaceae − Lactobacillus
Comamonadaceae − Comamonas
Carnobacteriaceae − Desemzia
Bacillaceae − Bacillus
Bacillaceae − Anoxybacillus
[Tissierellaceae] − Parvimonas
Unclassified genus
Genus
−10
−5
0
5
log2FoldChange
Differentailly Abundant
DC
FDR
b
Xanthomonadaceae − Stenotrophomonas
Turicibacteraceae − Turicibacter
Ruminococcaceae − Ruminococcus
Pseudomonadaceae − Pseudomonas
Prevotellaceae − Prevotella
Planococcaceae − Sporosarcina
Planococcaceae − Planomicrobium
Planococcaceae − Planococcus
Planococcaceae − Lysinibacillus
Planococcaceae −
Moraxellaceae − Enhydrobacter
Moraxellaceae − Acinetobacter
Lactobacillaceae − Lactobacillus
Lactobacillaceae −
Enterococcaceae − Enterococcus
Corynebacteriaceae − Corynebacterium
Comamonadaceae − Delftia
Comamonadaceae − Comamonas
Carnobacteriaceae − Desemzia
Bacillaceae − Geobacillus
Bacillaceae − Bacillus
Bacillaceae − Anoxybacillus
Alcaligenaceae − Achromobacter
Unclassified genus
−5
0
5
10
log2FoldChange
Differentailly Abundant
CeD
DC
c
Xanthomonadaceae − Stenotrophomonas
Streptococcaceae − Streptococcus
Moraxellaceae − Acinetobacter
Lactobacillaceae − Lactobacillus
Lactobacillaceae −
Enterococcaceae −
Corynebacteriaceae − Corynebacterium
Coriobacteriaceae − Atopobium
Coriobacteriaceae −
Caulobacteraceae − Brevundimonas
Bacillaceae − Bacillus
Actinomycetaceae − Actinomyces
[Tissierellaceae] − Parvimonas
[Mogibacteriaceae] − Mogibacterium
−4
0
4
8
log2FoldChange
Differentailly Abundant
CeD
FDR

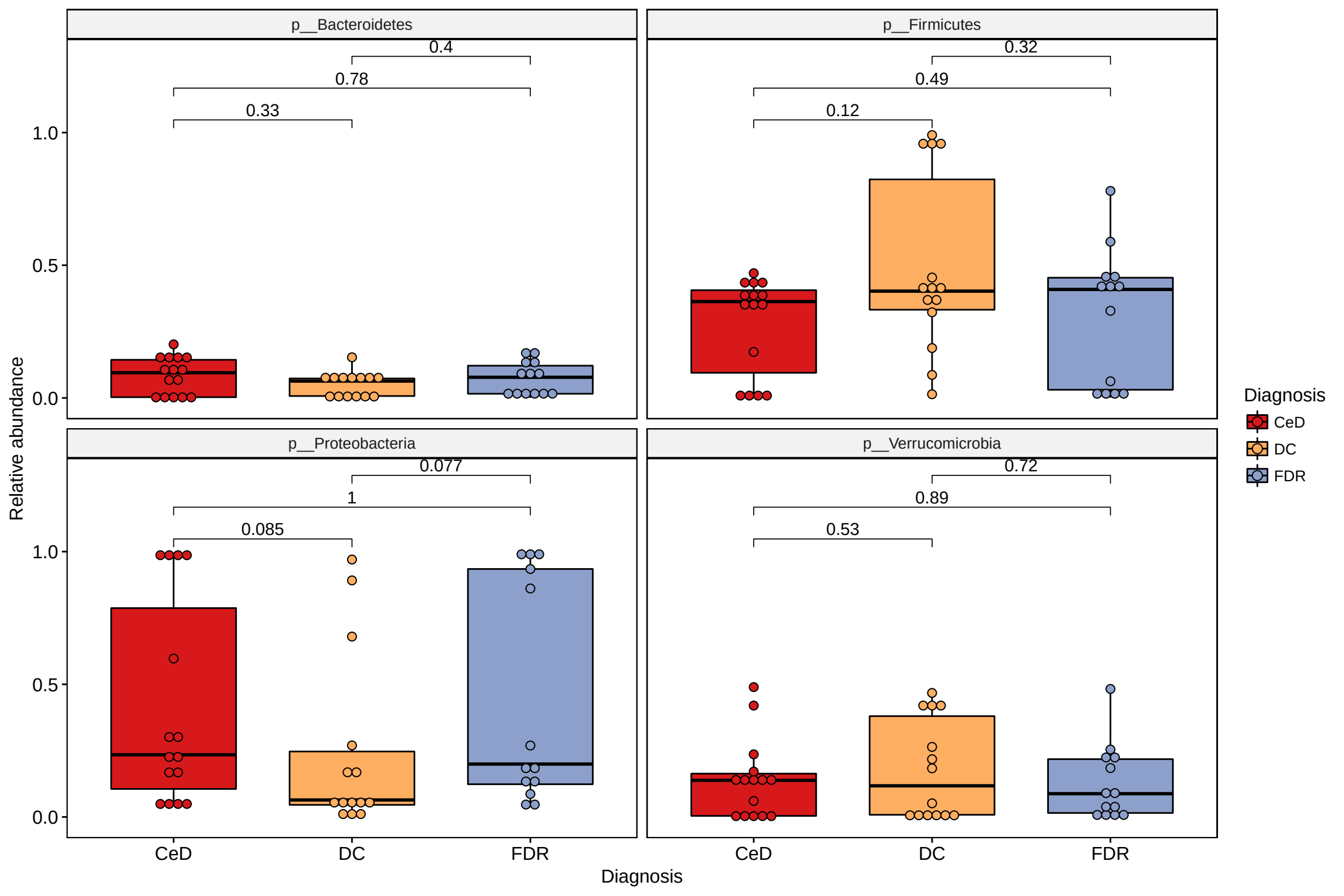

a
Streptococcaceae − Streptococcus
Ruminococcaceae − Ruminococcus
Ruminococcaceae −
Pseudomonadaceae − Pseudomonas
Prevotellaceae − Prevotella
Planococcaceae − Lysinibacillus
Planococcaceae −
Moraxellaceae − Acinetobacter
Lactobacillaceae − Lactobacillus
Enterococcaceae − Enterococcus
Enterobacteriaceae −
Coriobacteriaceae −
Comamonadaceae − Comamonas
Bacteroidaceae − Bacteroides
Bacillaceae − Bacillus
Unclassified genus
Genus
−5
0
5
log2FoldChange
Differentailly Abundant
DC
FDR
b
Turicibacteraceae − Turicibacter
Ruminococcaceae −
Planococcaceae −
Peptostreptococcaceae −
Moraxellaceae − Acinetobacter
Leuconostocaceae − Weissella
Leuconostocaceae −
Lactobacillaceae − Lactobacillus
Lachnospiraceae − Coprococcus
Lachnospiraceae −
Erysipelotrichaceae − [Eubacterium]
Enterobacteriaceae − Trabulsiella
Enterobacteriaceae −
Coriobacteriaceae −
Comamonadaceae − Comamonas
Clostridiaceae − SMB53
Clostridiaceae −
Christensenellaceae −
Bacteroidaceae − Bacteroides
Bacillaceae − Bacillus
Unclassified genus
−5
0
5
log2FoldChange
Differentailly Abundant
CeD
DC
c
Streptococcaceae − Streptococcus
Ruminococcaceae −
Prevotellaceae − Prevotella
Planococcaceae − Lysinibacillus
Peptostreptococcaceae −
Leuconostocaceae −
Lactobacillaceae − Lactobacillus
Lachnospiraceae −
Erysipelotrichaceae − [Eubacterium]
Enterococcaceae − Enterococcus
Enterococcaceae −
Enterobacteriaceae − Trabulsiella
Enterobacteriaceae − Serratia
Enterobacteriaceae −
Coriobacteriaceae − Collinsella
Comamonadaceae − Comamonas
Clostridiaceae − SMB53
Clostridiaceae −
Bacteroidaceae − Bacteroides
Unclassified genus
−20
−10
0
10
log2FoldChange
Differentailly Abundant
CeD
FDR

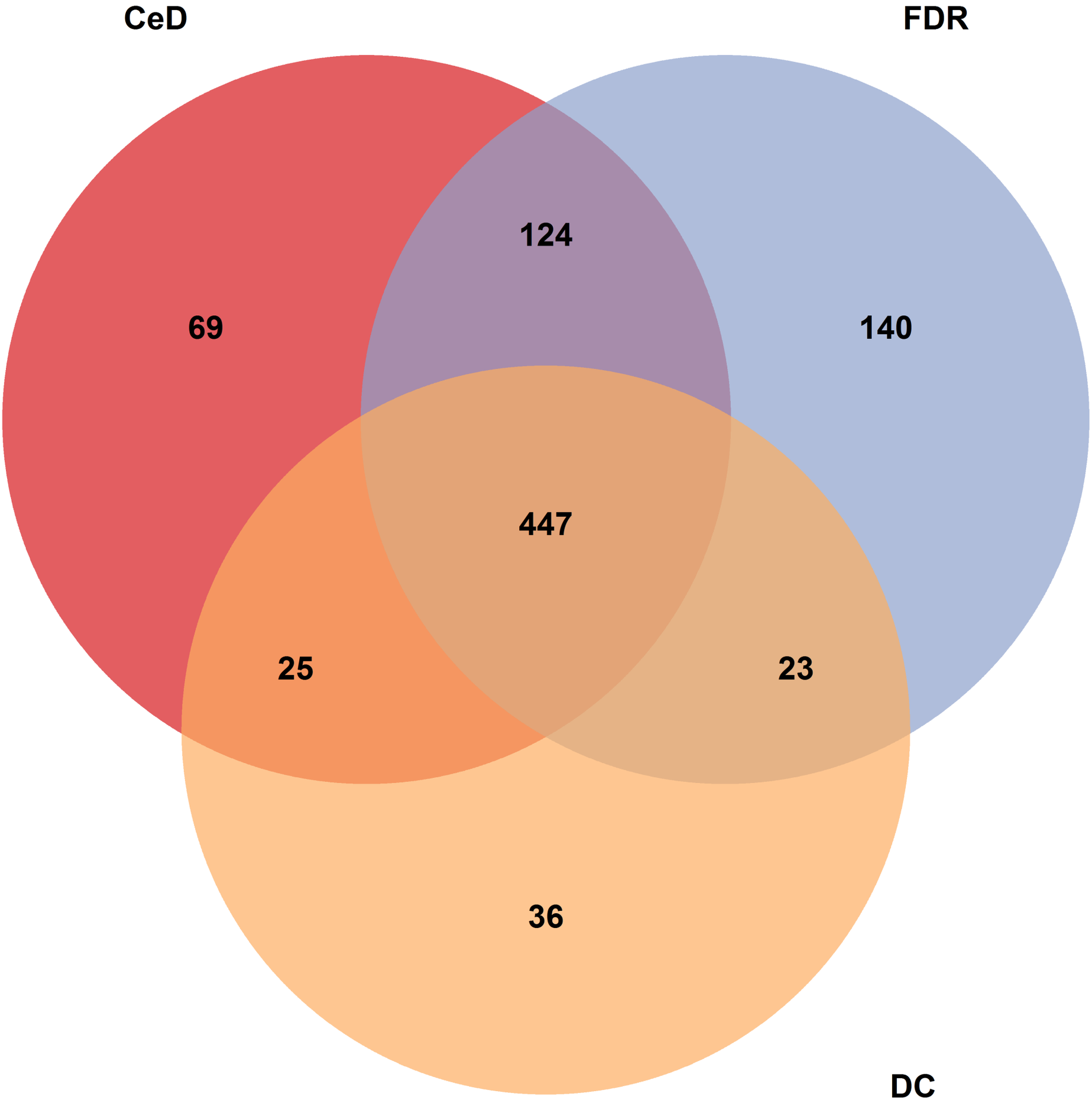

Kingdom
Phylum
Class
Order
Family
Genus
Actinomycetaceae
Corynebacteriaceae
Micrococcaceae
Propionibacteriaceae
Bifidobacteriales
Bacteroidetes
Cyanobacteria
Bacillaceae
Paenibacillaceae
Lysinibacillus
Planococcus
Planomicrobium
Sporosarcina
unknown
Staphylococcaceae
unknown
Gemellales
Aerococcaceae
Carnobacteriaceae
Enterococcaceae
Lactobacillus
Pediococcus
unknown
Leuconostocaceae
Streptococcaceae
unknown
Turicibacterales
Clostridiaceae
Blautia
Coprococcus
Dorea
[Ruminococcus]
unknown
Peptostreptococcaceae
Faecalibacterium
Oscillospira
Ruminococcus
unknown
Veillonellaceae
[Mogibacteriaceae]
[Tissierellaceae]
unknown
Erysipelotrichi
Fusobacteria
Alphaproteobacteria
Betaproteobacteria
Epsilonproteobacteria
Aeromonadales
Enterobacteriales
Oceanospirillales
Pasteurellales
Acinetobacter
Pseudomonas
unknown
Xanthomonadales
Synergistetes
TM7
Akkermansia

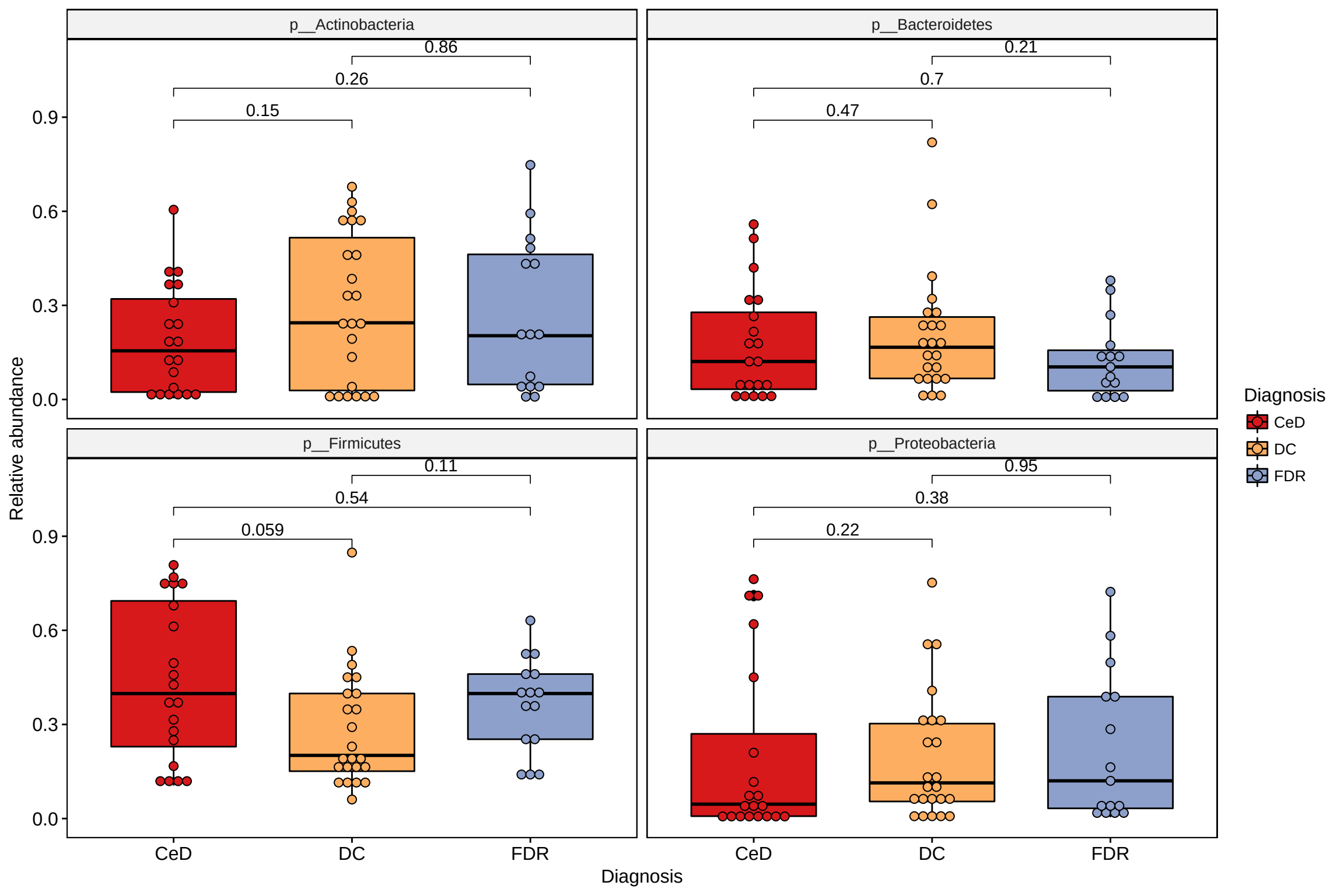

SubPathway
SuperPathway
Category
Cell cycle - Caulobacter
Bacterial chemotaxis
Bacterial motility proteins
Flagellar assembly
Environmental Information Processing
Genetic Information Processing
Amino Acid Metabolism
Carbohydrate Metabolism
Energy Metabolism
Peptidases
Protein kinases
Glycan Biosynthesis and Metabolism
Lipid Metabolism
Metabolism of Cofactors and Vitamins
Metabolism of Other Amino Acids
Metabolism of Terpenoids and Polyketides
Nucleotide Metabolism
Xenobiotics Biodegradation and Metabolism
Unknown function
Enzyme Families
Metabolism

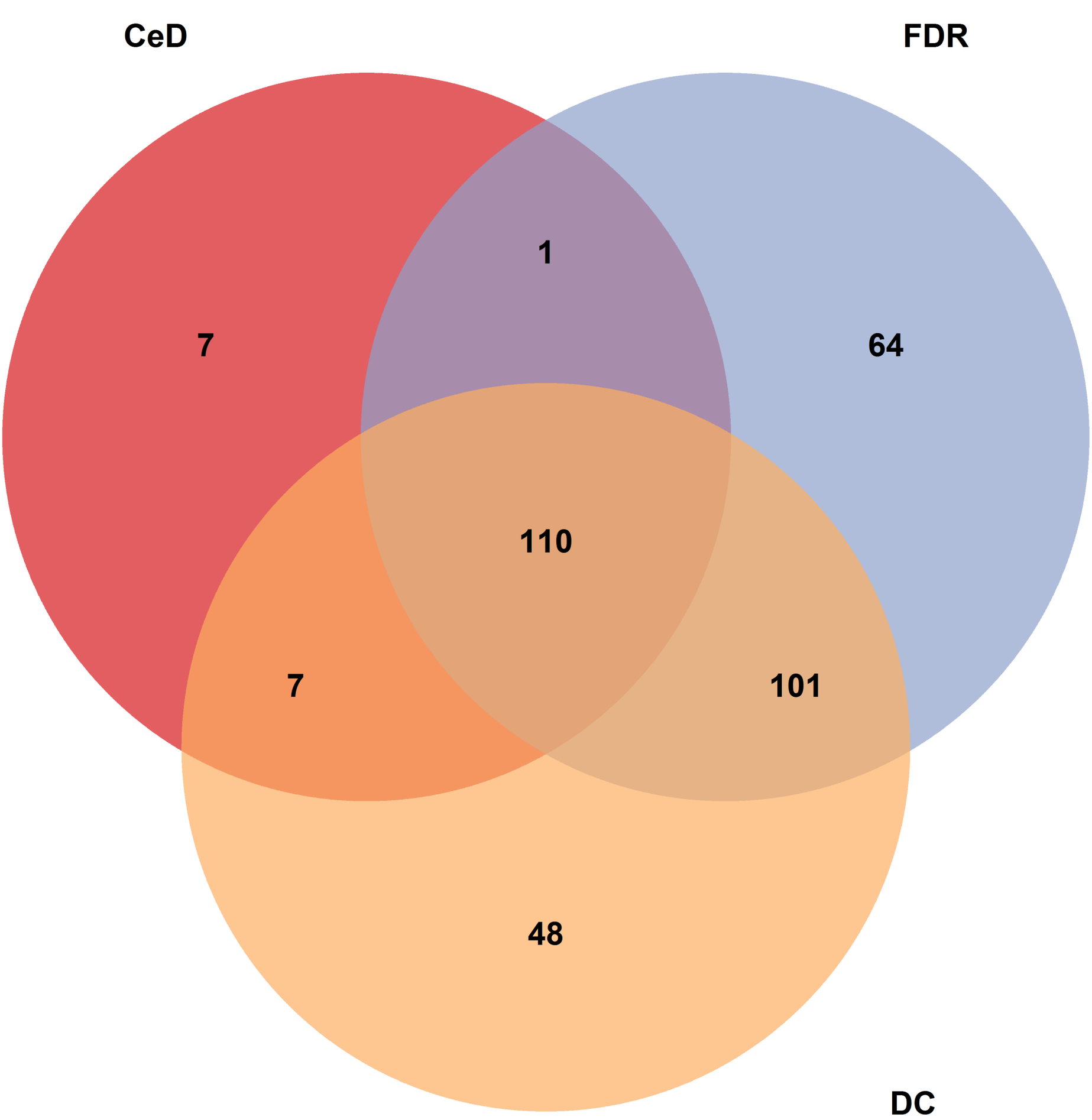

a
Pyroglutamyl−peptidase
(p = 0.049)
Number of Sequences
60000
40000
20000
0
CeD
DC
FDR
b
Subtilisin
(p = 0.021)
Number of Sequences
40000
20000
0
CeD
DC
FDR
c
X−Pro dipeptidase
(p = 0.067)
Number of Sequences
1e+05
5e+04
0e+00
CeD
DC
FDR
d
Aminopeptidase
(p = 0.022)
Number of Sequences
400
300
200
100
0
CeD
DC
FDR
Diagnosis:
CeD
DC
FDR


Top OTUs
Acinetobacter OTU-543942
Acinetobacter OTU-562618
Comamonas OTU-558170
Corynebacterium OTU-1040713
Novosphingobium OTU-808758
Stenotrophomonas OTU-815480
Pseudomonas OTU-764682
Acinetobacter OTU-988314
A. johnsonii OTU-1009894
Pseudomonas OTU-961783
CeD
FDR
−0.03
−0.02
−0.01
0.00
0.01
0.02
0.03


p__Bacteroidetes
p__Firmicutes
p__Proteobacteria
p__Verrucomicrobia
0.4
0.78
0.33
0.32
0.49
0.12
0.077
1
0.085
0.72
0.89
0.53
1.0
0.5
0.0
Relative abundance
CeD
DC
FDR
Diagnosis
Diagnosis
CeD
DC
FDR


CeD
FDR
124
69
140
447
25
23
36
DC


p__Actinobacteria
p__Bacteroidetes
p__Firmicutes
p__Proteobacteria
0.86
0.26
0.15
0.21
0.7
0.47
0.11
0.54
0.059
0.95
0.38
0.22
0.9
0.6
0.3
0.0
Relative abundance
CeD
DC
FDR
Diagnosis
Diagnosis
CeD
DC
FDR


CeD
FDR
DC
7
1
64
110
7
101
48