

Patients and samples

Exclusion criteria were as follows: age under 1 or over 18 years; confirmed acute or chronic infections; gastrointestinal inflammatory diseases (especially inflammatory bowel disease); diabetes; thyroid disorders; obesity; congenital or acquired immunodeficiencies (particularly selective IgA deficiency); active neoplastic diseases (particularly gastrointestinal cancers); use of antibiotics and/or probiotics within 30 days before stool sample collection; and lack of consent to participate.

All participants also underwent routine blood tests, and transglutaminase IgA antibodies (anti-tTG) were measured by enzyme-linked immunosorbent assay (ELISA); titers above 20 RU/mL were considered abnormal. In addition, HLA-DQ2/DQ8 allele typing was performed using the CeliacStrip test (Operon, Cuarte de Huerva [Zaragoza], Spain).

All laboratory analyses were conducted at the University Children's Hospital laboratory in Krakow, following International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) guidelines and using IVD/FDA-certified diagnostic kits included in the hospital's clinical protocol.

Genomic library preparation

The reaction mixture consisted of 5 µL of DNA, 12.5 µL of Kapa Biosystems (Roche, Basel, Switzerland), 0.5 µL of each primer (10 mmol/L, Genomed, Warsaw, Poland), and 6.5 µL of water (A&A Biotechnology, Gdansk, Poland). The thermal cycling conditions included an initial denaturation at 95 °C for 5 minutes, followed by 50 cycles (for fungi) or 30 cycles (for bacteria) of 95 °C for 30 seconds, 55 °C for 30 seconds, and 72 °C for 30 seconds. A 5 µL volume of each amplicon was analyzed via electrophoresis on a 1.5% agarose gel (Prona, Basica Le, Burgos, Spain) diluted tenfold with TBE buffer (Sigma-Aldrich, Saint Louis, United States). The PCR products were visualized under UV light using the FastGene FAS-Digi Pro system (Nippon Genetics, Duren, Germany) to assess library quality.

Taxonomic profiling and microbial community analyses

Statistical comparisons of microbial abundance between study groups were conducted at each taxonomic level using non-parametric Wilcoxon rank-sum tests, with multiple testing corrections by Benjamini–Hochberg. Linear discriminant analysis effect size (LEfSe) and Kruskal–Wallis tests (p -value < 0.05 , LDA effect size threshold = 2) were applied to identify taxa with statistically significant and biologically relevant differential abundance. The median abundance differences between groups were assessed using Wilcoxon rank-sum tests.

Biomarker discovery using LefSe

The results from differential abundance testing were integrated to discover common microbial signatures associated with phenotypes of interest. To identify microbial taxa and features that significantly discriminated between study groups, linear discriminant analysis effect size (LEfSe) was applied. The analysis was conducted on relative abundance data at multiple taxonomic levels using the LEfSe pipeline with the following parameters: p -value for the Kruskal–Wallis test set at < 0.05 for detecting features with significantly different abundances, and a logarithmic Linear Discriminant Analysis (LDA) score threshold of 2. LEfSe combines non-parametric statistical testing with effect size estimation to uncover biologically meaningful biomarkers. Identified biomarkers were further validated by examining normalized relative abundance matrices and confirmed using Wilcoxon rank-sum tests adjusted for multiple comparisons.

Diversity analyses

The PERMANOVA and ANOSIM tests were applied to evaluate the significance of the effect of phenotype on community composition.

Alpha diversity indices including observed species richness, ACE, Chao1, Shannon, Simpson and Fisher diversity were calculated from the downsized and normalized species abundance tables at rarefaction thresholds ranging from 10,000 to 100,000 counts per sample using the `estimate_richness` function.

Beta diversity was evaluated to assess differences in microbial community composition between samples. Three complementary dissimilarity indices were

computed: Bray–Curtis (abundance-based), Jaccard (presence/absence-based), and Jensen–Shannon divergence (distribution-based). Beta diversity was calculated for every pair of samples to generate a distance of dissimilarity.

The significance of clustering by clinical group and other covariates was tested using permutational multivariate analysis of variance (PERMANOVA).

Statistical analyses

Post-hoc power analysis was conducted for the comparison of three groups (CeD, siblings, controls) using one-way ANOVA, assuming small ($f = 0.10$), medium ($f = 0.25$), and large ($f = 0.40$) effect sizes.

Post-hoc power analysis was conducted for alpha diversity using ANOVA and for beta diversity using PERMANOVA assumptions. Post-hoc power analysis was also calculated for categorical variables using chi-square tests of independence on contingency tables for three study groups. For each variable, we computed the chi-square statistic, Cramér's V as a measure of effect size, and then estimated power using the observed effect size, total sample size, and $\alpha = 0.05$. Post-hoc power analysis was also performed for all quantitative biochemical variables using a one-way ANOVA model. Effect size was calculated as Cohen's f based on η^2 derived from between-group and within-group variance. Power was computed using the observed effect size, total sample size, and $\alpha = 0.05$.

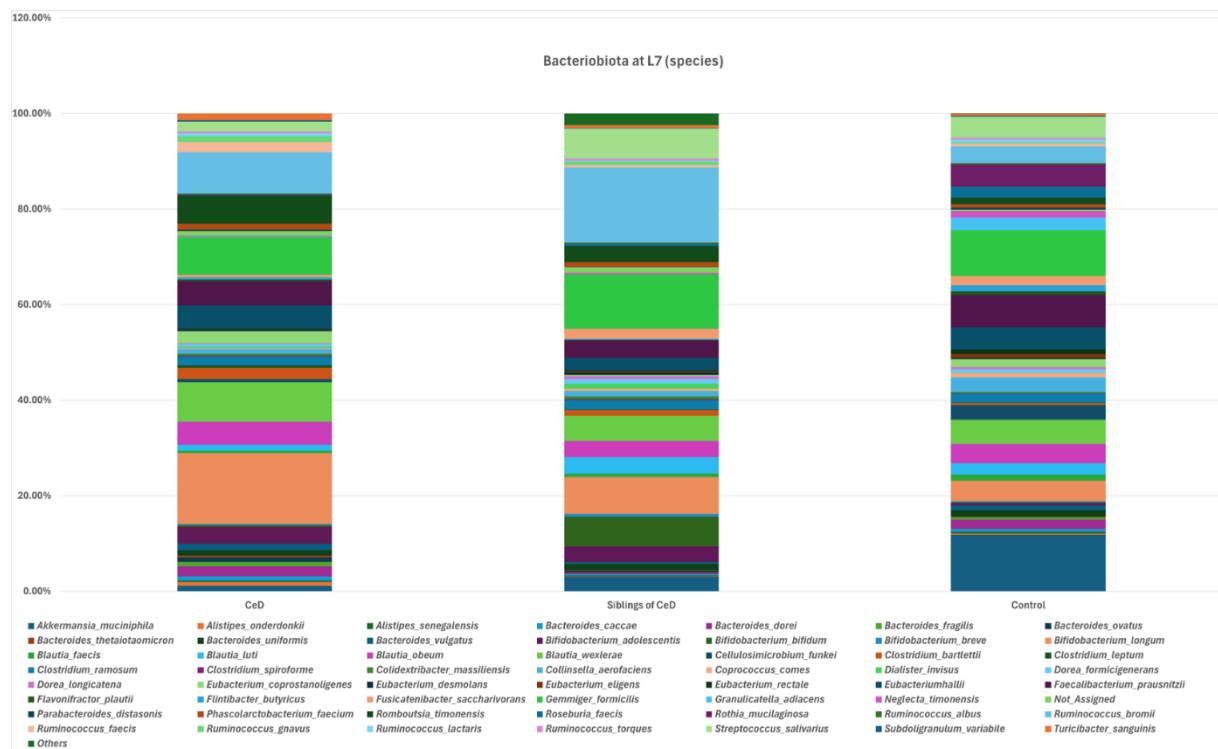
All post-hoc power analyses were conducted in Python (version 3.11) using the statsmodels package (version 0.14.0; <https://www.statsmodels.org>) within a Windows 11 environment.

The descriptive statistics was carried out with IBM SPSS Statistics 28 and were calculated for all variables. Normality of data distribution was assessed using the Shapiro–Wilk test for each variable within the three study groups (CeD, siblings, controls). Variables with normal distribution are presented as mean \pm standard deviation (SD), while non-normally distributed variables are presented as median and quartile 1 (Q1) and quartile 3 (Q3). Multivariate analysis was performed to determine the statistically significant differences of the means when comparing three groups' results. The Chi-square and Fisher's tests were used to analyze the variables on the

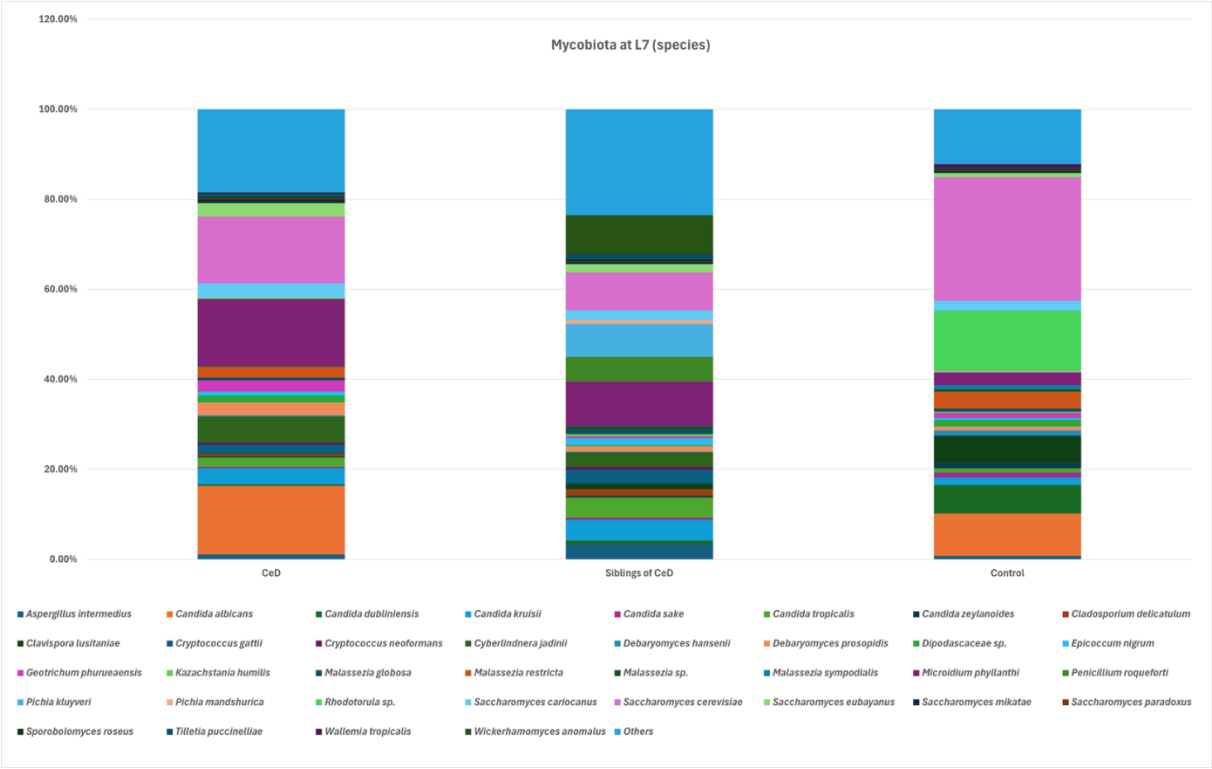
nominal scale. The Kruskal–Wallis test was used for quantitative variables, and a post-hoc test (Mann–Whitney U test) was used for each pair of groups with Bonferroni correction to account for multiple comparisons. After Bonferroni correction, the significance level was set at $\alpha = 0.017$ for three pairwise comparisons

Pairwise comparisons were made between the three groups: CeD vs siblings, CeD vs controls, and siblings vs controls. Microbiota profiles at L6 and L7 were correlated with anti-tTG antibody levels and HLA DQ2/DQ8 genotypes. Linear regression models were applied to assess the impact of clinical variables, sequencing covariates, and disease status on alpha diversity indices. Wilcoxon rank-sum tests evaluated differences between groups on both raw and residual diversity measures (after adjustment for covariates). PERMANOVA assessed the marginal effects of covariates on beta diversity. Differences in microbial composition were tested with Chi-square tests.

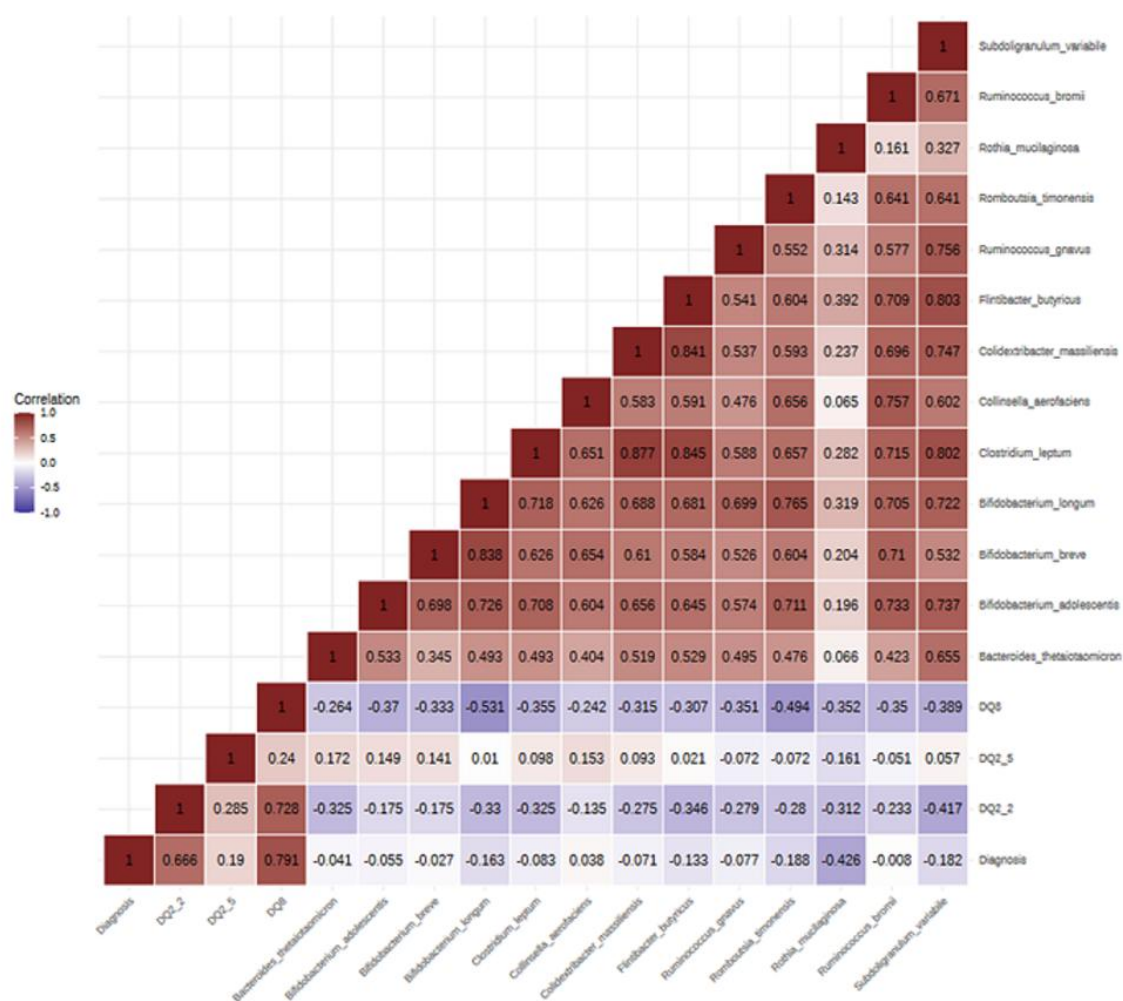
Associations between microbial features and disease status were investigated via linear regression model adjusting. The dependence between the numbers of specific bacteria/fungi and HLA DQ2/DQ8 alleles, anti- tTG titre and diagnosis of the celiac disease was evaluated using correlation analysis, in which traits deviating from the normal distribution were analyzed using Spearman's rank correlation coefficient. Because multiple correlations were tested, p-values were adjusted for multiple comparisons using the Benjamini-Hochberg procedure to control the false discovery rate. Statistical significance was set at $p < 0.05$ after correction. All analyses were performed using the R software package (v. 4.5.1 [June, 2025] v1.30.0).



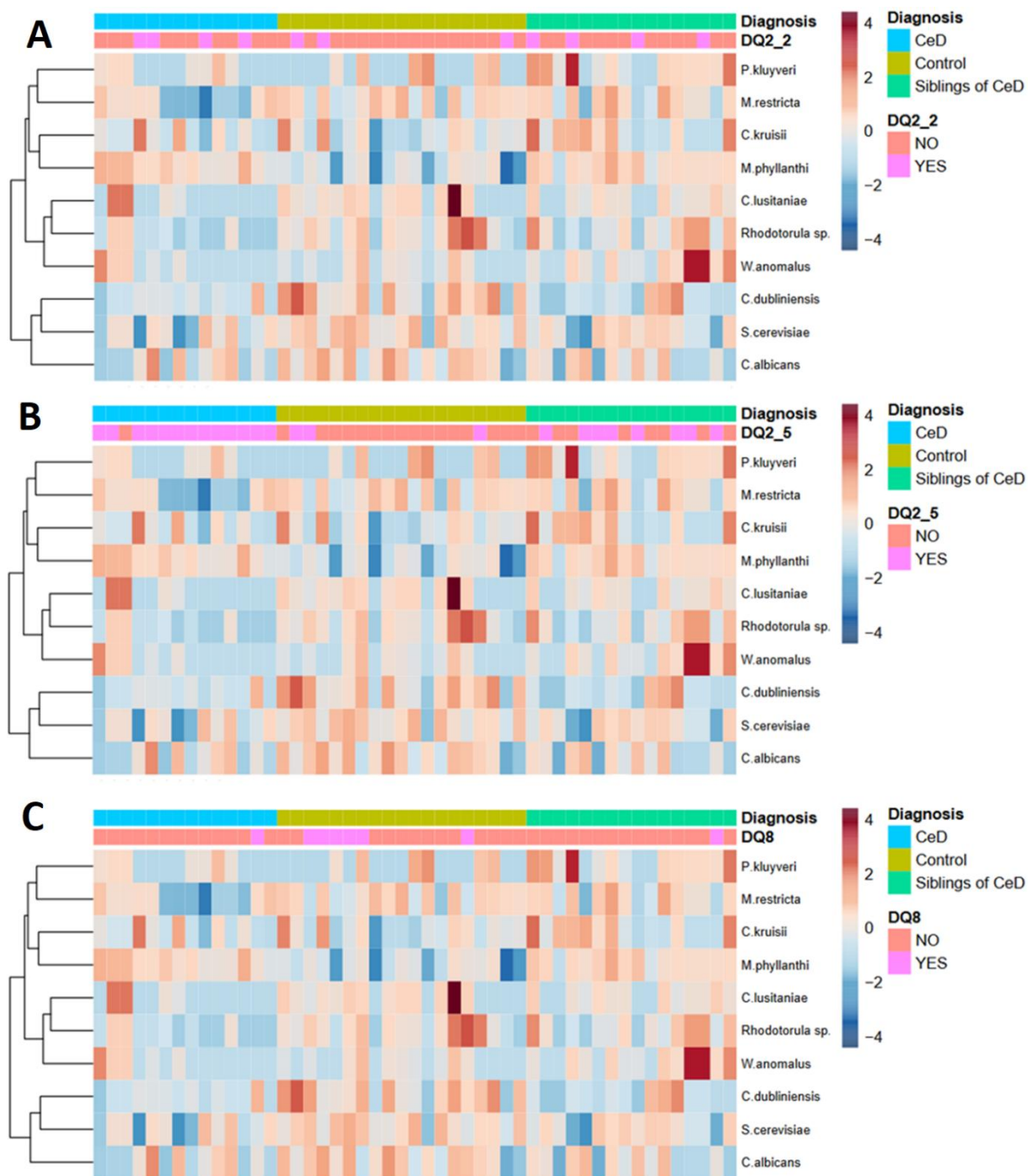
Supplementary Figure 1



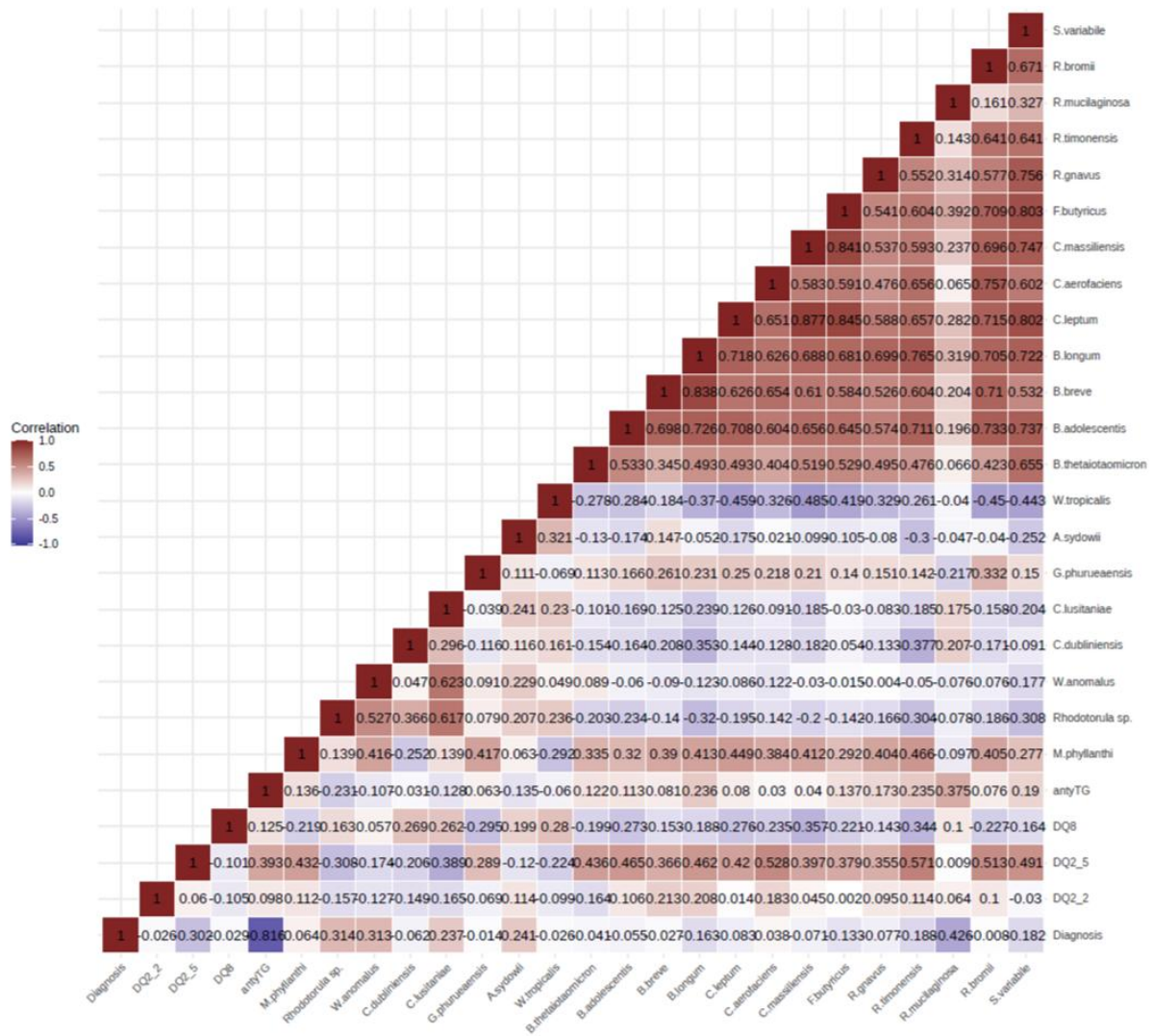
Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4



Supplementary Figure 5

Supplementary Table 1 The primer sequences (Genomed, Warsaw, Poland) targeting the 16S and ITS-1 regions. Adapter sequences for the MiSeq sequencer (Illumina, San Diego, United States) were shown in bold

Primer Sequence 5'→3'	Bacteria (16S region)
F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCCTACGGGNGGCWGCAG
R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGTGCGGGCCCCCGTCAATT
	Fungi (ITS-1 region)
F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTAAAAGTCGTAACAAGGTTTC
R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTTCAAAGAYTCGATGATTCAC