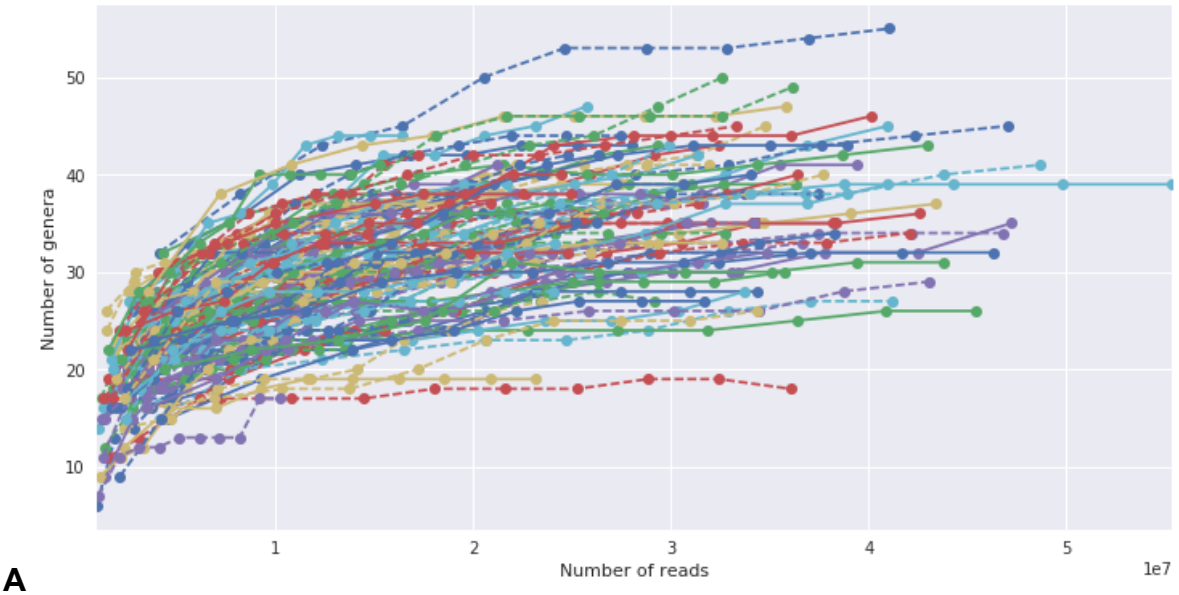
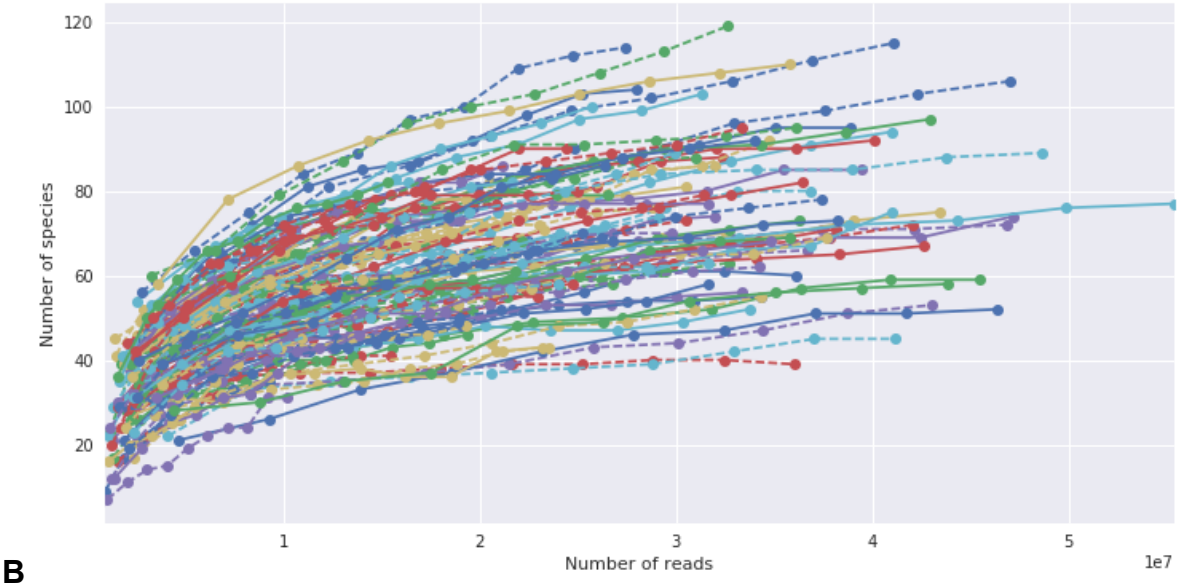
Supplementary data



A



B

Figure S1. The results of metagenomic rarefaction analysis. The number of genera (**A**) / species (**B**) (X axis) classifications using MetaPhlAn2 depending on reads coverage (Y axis). The ID of samples is shown different colors.

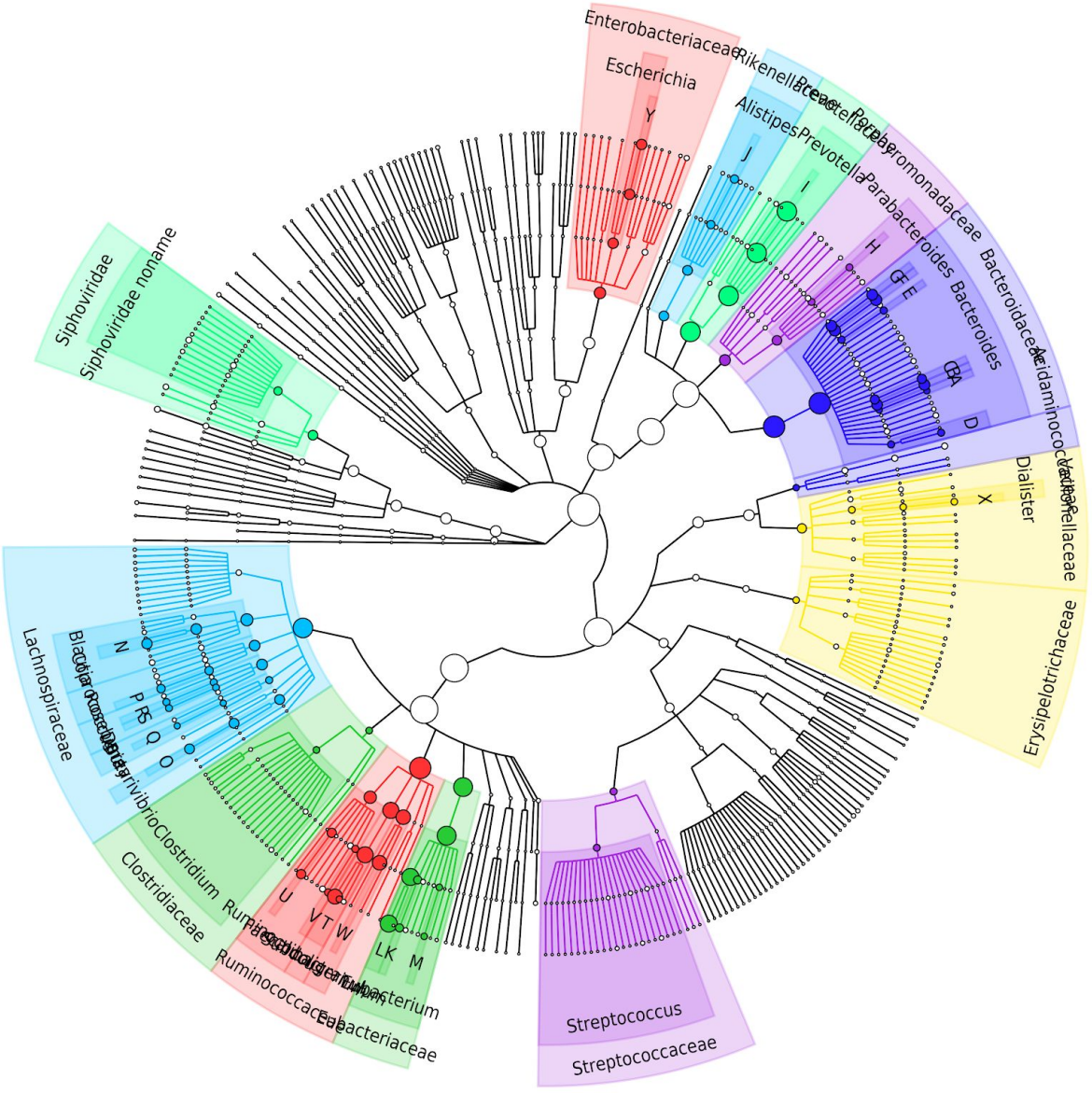


Figure S2. GraPhlAn cladogram based on MetaPlAn2 output show general taxonomic composition of patients microbiota (1st and 2nd time points). Different color denote different taxonomic levels combined by family level. Each dot represents bacterial taxon. The dot size show relative abundance value. The various shades of colors highlighting the background indicates that the changes consistent at several taxonomic levels.

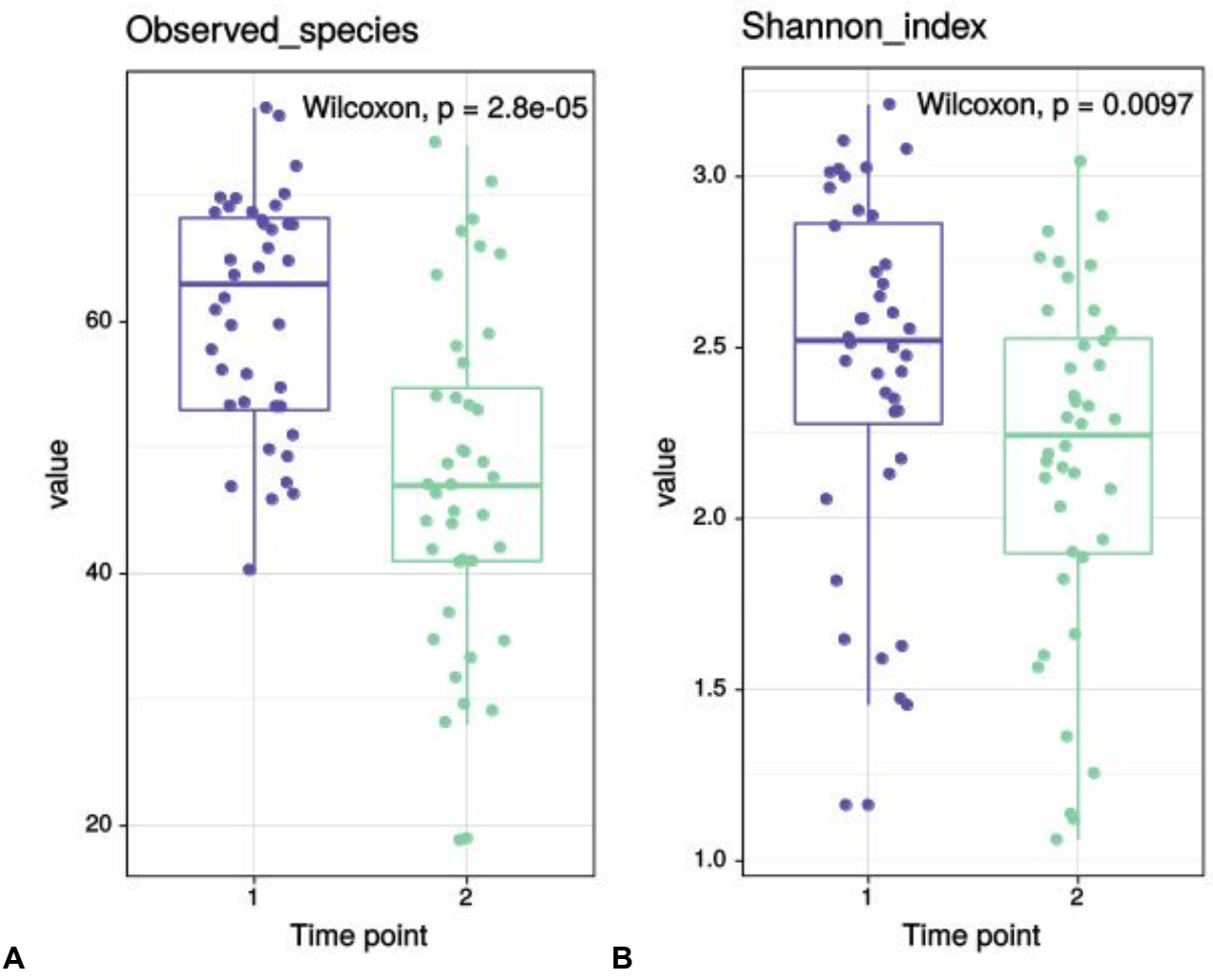


Figure S3. Boxplots represent the 25th to 75th percentile range of observed species (**A**) and Shannon index (**B**) among sample for 1st and 2nd time points, while whiskers indicate the full range of sample values.

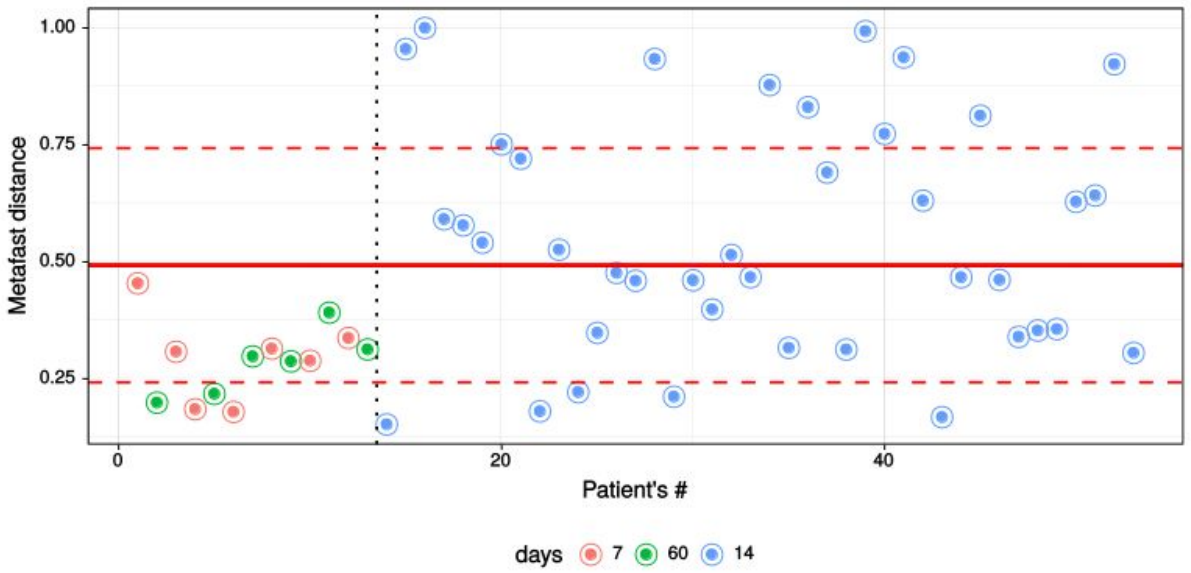
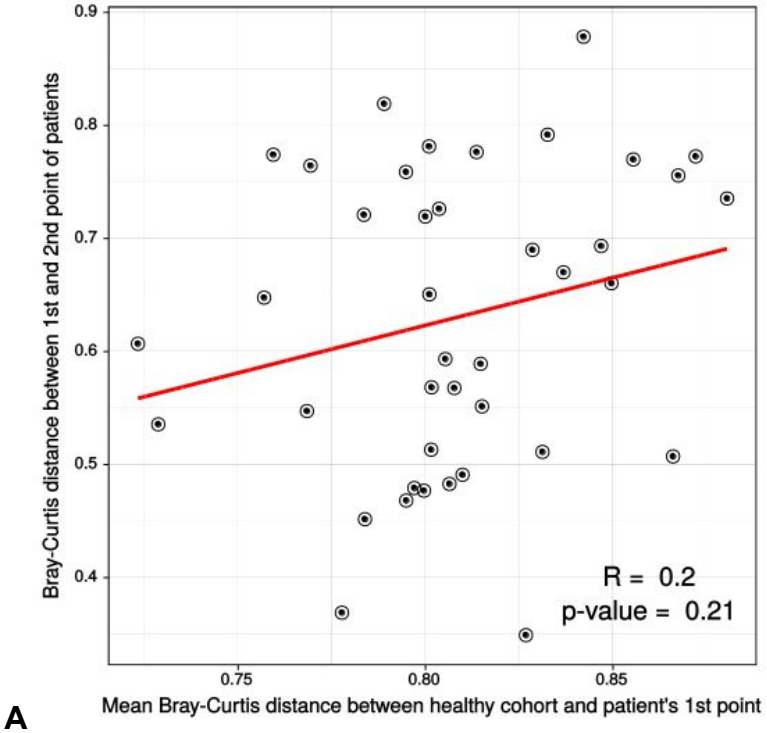
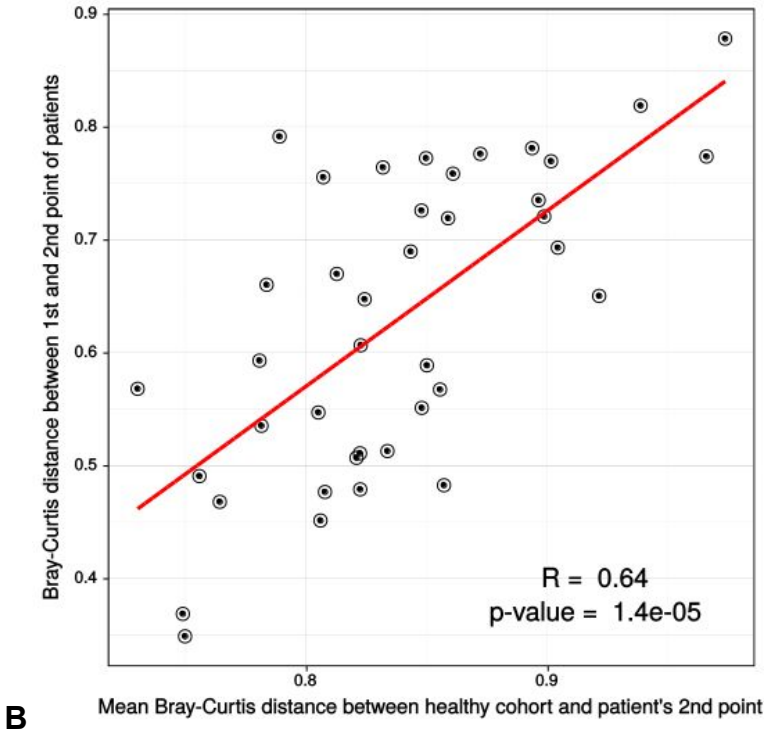


Figure S4. The changes scale of the MetaFast distance caused by *H. pylori* eradication therapy. The X axis denote the patient numbers, the Y axis - MetaFast distance value between different time points samples. Time points is shown different colors. The dashed vertical line separates the control sample (on the left - red and green dots) from the experiment group (on the right - blue dots). The red solid horizontal line shown mean value MetaFast distance. Dashed horizontal red lines show the values mean±SD of MetaFast distance.



A



B

Figure S5. Dependence of the Brey-Curtis distance between the 1st and 2nd time points metagenomes of experimental group from A) the range of the first time point from a healthy cohort (Spearman R=0.20, P=0.21). **B)** the range of the second time point from a healthy cohort (Spearman R=0.64, P = 1.4×10^{-5}).

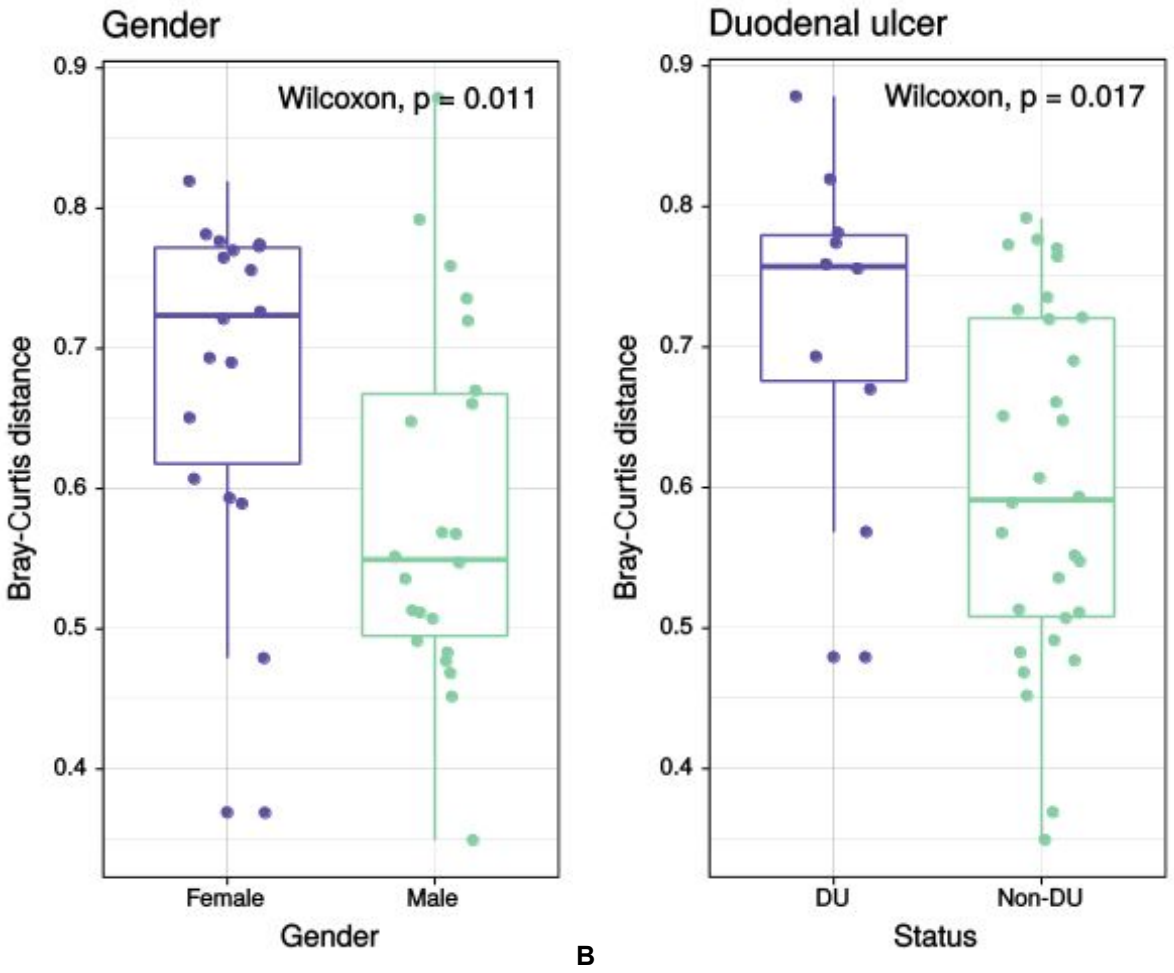


Figure S6. Boxplots represent the 25th to 75th percentile range of difference Bray-Curtis dissimilarity (between 1st and 2nd time points) among sample for (**A**) female and male, (**B**) duodenal ulcer and non-duodenal ulcer, while whiskers indicate the full range of sample values.

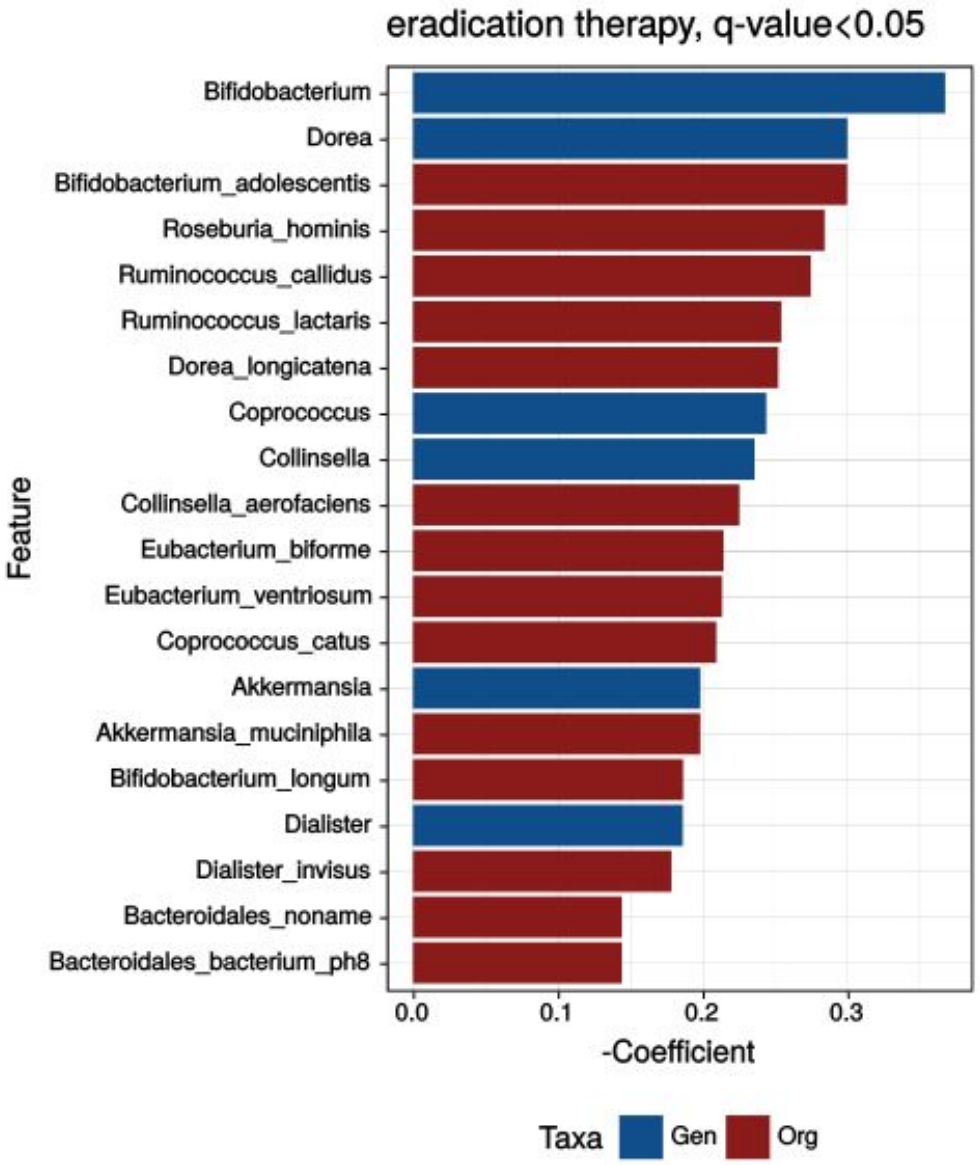
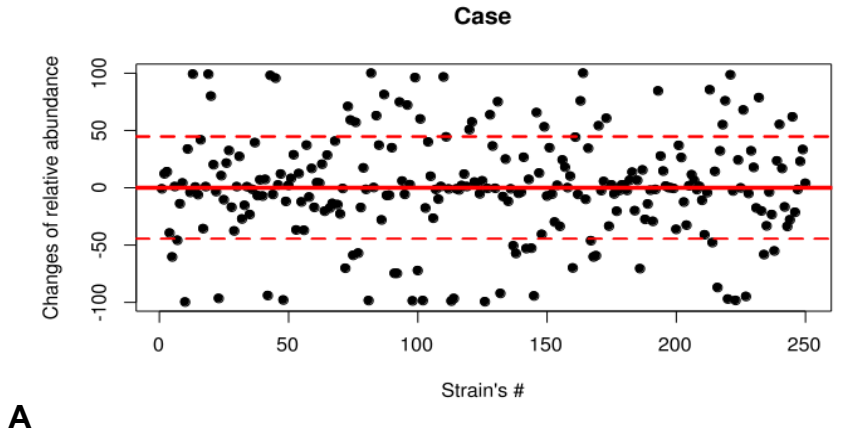
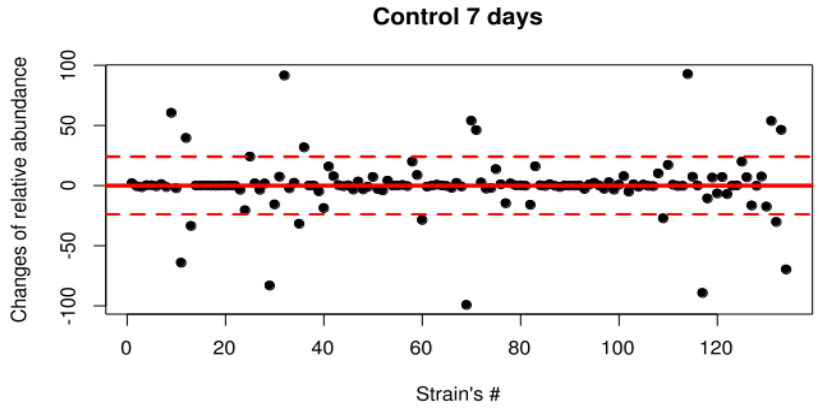


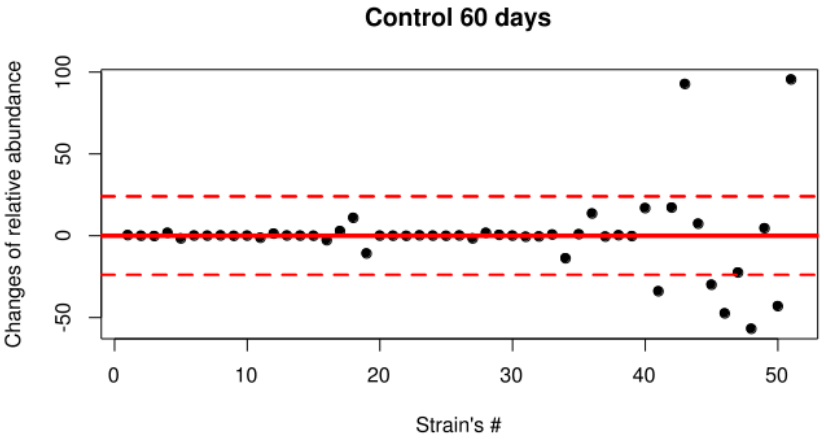
Figure S7. Microbial species significantly associated with *H. pylori* eradication therapy. The figure is shows coefficients × (-1) of the linear model obtained by applying MaAsLin method to reference-mapping based taxonomic composition vectors (q-value < 0.05). The values denoted scale reverse correlation between relative abundance of different taxa with eradication therapy.



A



B



C

Figure S8. The scale of changes in relative abundance for each individual strain between the experimental and control groups. A - experimental group, **B** - control 7 days, **C** - control 60 days. The red solid line indicates mean value of relative abundance changes. Dotted lines show mean ± SD values of relative abundance changes. Intra-specific structures with non-zero relative representation at two time points were selected (for the control group, separately analyzed metagenomes selected at 7 and 60 days after the first sampling). When comparing the dispersion, significant differences were revealed between the change in the relative abundance of strains in the experimental and control groups Siegel-Tukey test A-B $P=6.493 \times 10^{-12}$, A-C $P=1.383 \times 10^{-8}$, B-C $P=0.051$.

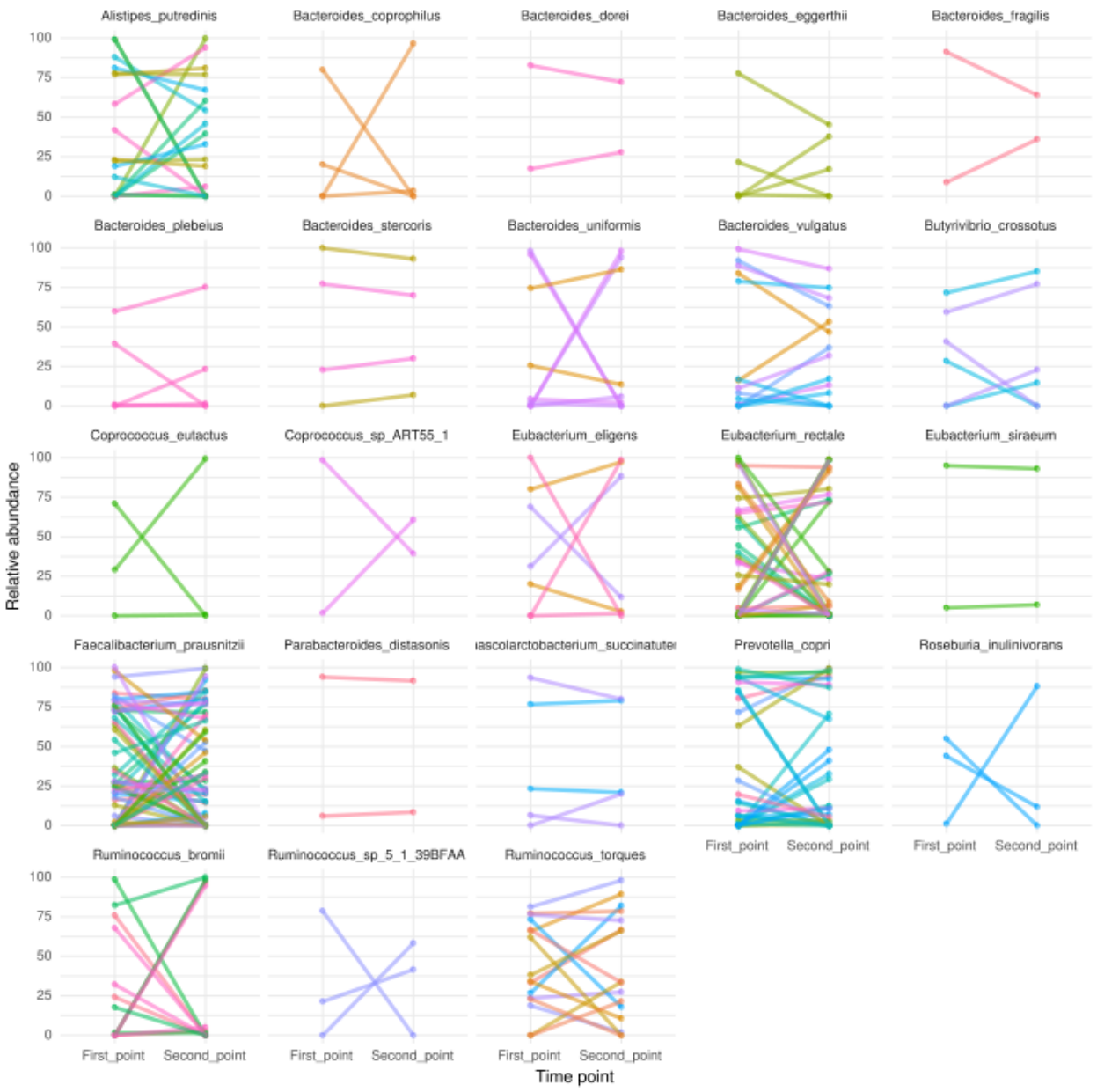


Figure S9. The changes in intraspecies relative abundances of intraspecies structures ("strains") between 1st and 2nd time points. The colors denoted the patients. The relative abundances were received by ConStrains.

Note 1: *Enterococcus faecium* isolation from stool samples of patients study.

From faecal specimens of two patients (HP_003, HP_010) *E. faecium* strains were isolated. For each strains, resistance level to six antibiotics (azithromycin (AZM), ciprofloxacin (CIP), tetracycline (TE), penicillin (P), vancomycin (VAN) and cefixime (CFM)) were determined (see Supplementary Note S2). The results of testing the obtained isolates are presented in Table S14.

Table S16. Sensitivity to different classes of antibiotics of *E. faecium* strains isolated from stool samples of patients HP003 and HP010 obtained by disk-diffusion method. The diameter of the inhibition zone of bacterial growth with considering the diameter of antibiotic disk is indicated (mm). The dose of antibiotics is shown by the number beside to the antibiotic name. The red color correspond higher resistance level to antibiotics.

Patient	Time point	AZM 15	CIP 5	TE 30	P 10	VAN 30	CFM 5
HP_003	1	22	16	24	25	20	0
HP_003	1	9	15	26	14	19	0
HP_010	1	11	0	25	13	13	0
HP_010	1	10	16	27	22	15	0
HP_003	2	0	13	0	0	16	0
HP_010	2	8	12	7	0	13	0

The presented data shown the trend of increasing resistance to antibiotics at the second time point. In the case of the patient HP003, *E. faecium* at the first point were susceptible to 5 antibiotics and resistant to cefixime, while at the second point the sensitivity decreased (decreased inhibition zone) to azithromycin, tetracycline and penicillin, while maintaining resistance to cefixime. In the case of patient HP010, sensitivity to ciprofloxacin increased and decreased to penicillin at the second time point. Thus, the resistance of isolates to the classes of antibiotics used in the therapy regimen (macrolides and beta-lactams) at the second time point.

The search for ARGs in the genomic assemblies was performed (see Supplementary Note S2). The results of data analysis for isolate genomes presented on Figure S10. After eradication therapy, the number of ARGs to tetracyclines and macrolides in the genomes of *E. faecium* strains was increased (Supplementary Figure S10A). Also, these strains are phylogenetically close to each other (Supplementary Figure S10B). The isolates genomes of the 1st and 2d time points were phylogenetically distant.

For study the relative abundance of these isolates in patient metagenomas at different time points, we mapped the metagenomic reads to the contigs obtained by assembling the genomic data. The results presented on Supplementary Figure S10C and Supplementary Figure S10D. Based on the obtained results, all strains (from 1st and 2nd time points) were present in the metagenomic data at all time points. After therapy, isolates with an increased number of ARGs were more highly abundance. The obtained results may indicate a change in the strain distribution within the intestinal microbiota.

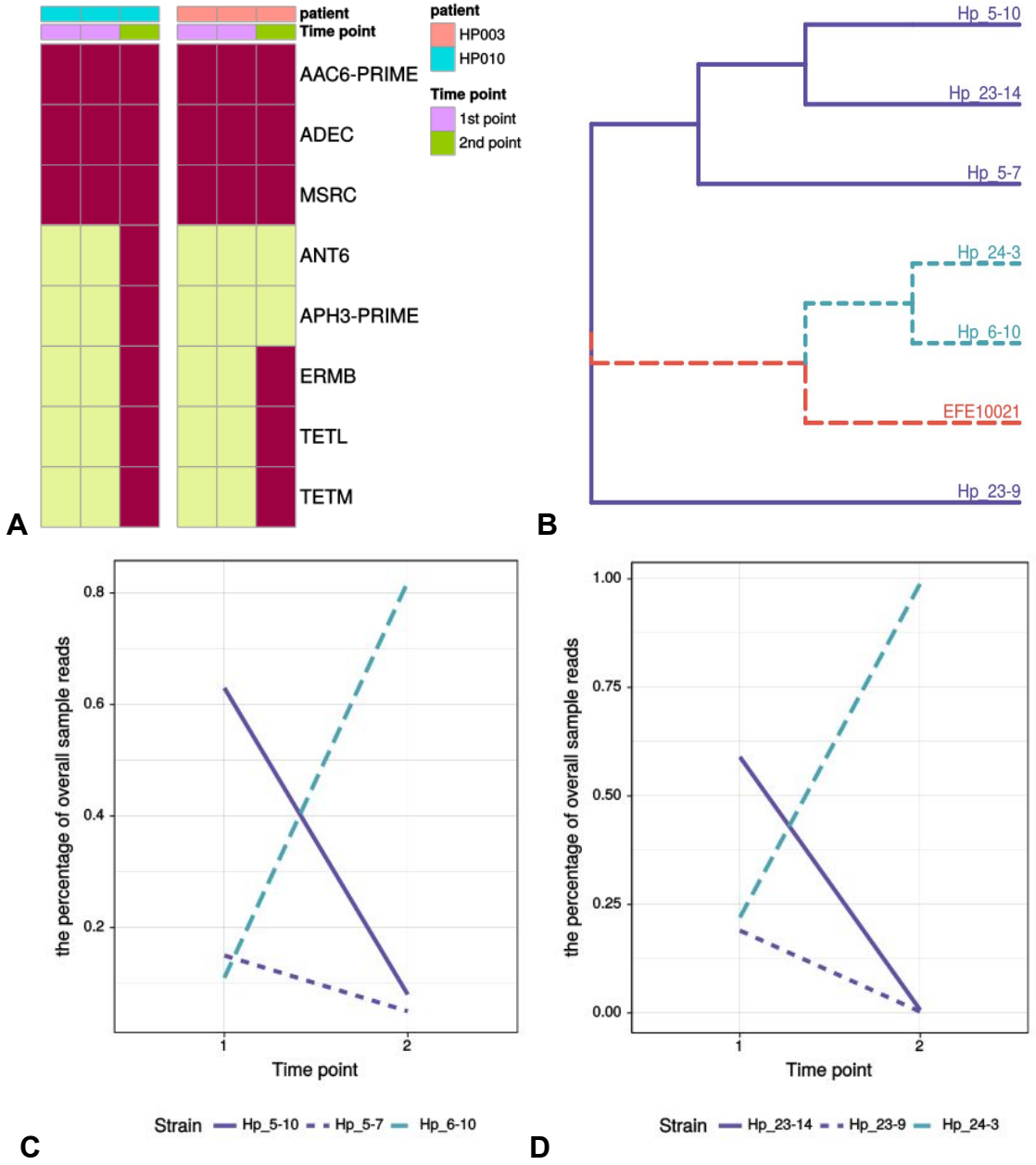


Figure S10. Results on the study of *Enterococcus faecium* isolates from stool samples of patients before and after eradication therapy. A. The ARGs presented in the genomes of *E. faecium* isolate from stool samples of patients HP003 and HP010 (1st and 2nd time points). The columns correspond to the samples/patients. The patient ID and time points is denoted with a top color bar. Presence ARGs is shown a red color. **B.** The phylogenetic tree built using PARSNP with default parameters for genomic assemblies of *E. faecium* strains. Color branches correspond to time points (Solid line denoted 1st point, dot line - 2nd point). The reference genome is shown by orange color. Shifts of relative abundance (of total abundance = 100%) of *E.faecium* strains isolated from HP003 (**C**) and HP010 (**D**) patient stool among samples for each time points.

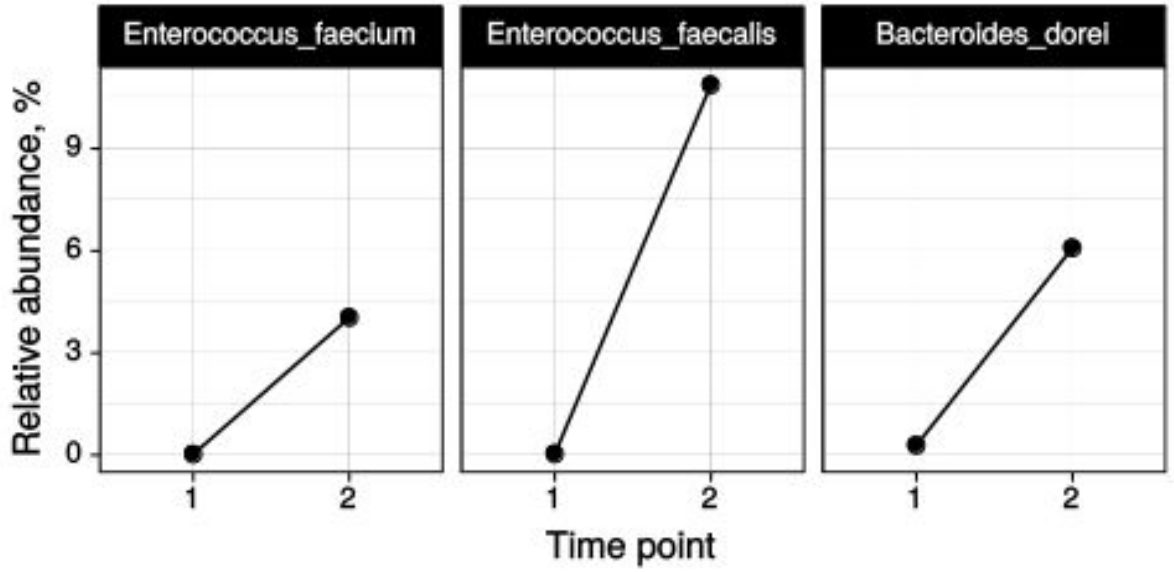


Figure S11. Line-point plots of *Bacteroides dorei*, *Enterococcus faecium* and *Enterococcus faecalis* relative abundance change associated with eradication therapy. The metagenomic data of HP003 patient (2nd time point) was used for figure constructed. Relative abundance values was obtained by MetaPhlAn2.

Note 2: Additional materials and methods

Enterococcus faecium isolation scheme

Patient stool samples for isolating *E. faecium* were selected based on taxonomic profiling by MetaPhlAn2. Metagenomic samples with non-zero relative abundance *E. faecium* at two time points were selected for isolating. The *E. faecium* strains that could be isolated at two time points were sequenced. The sampling scheme is shown in Table S13. The Description of *E. faecium* strains isolation, whole-genome sequencing (WGS) and genome assembly methods are presented in manuscript Prianichnikov et al. (2018).

Determine of *Enterococcus faecium* strains susceptibility to antibiotics

The resistance levels for azithromycin, ciprofloxacin, tetracycline, penicillin, vancomycin and cefixime for the isolates were determined by disk diffusion method. The plates containing Mueller-Hinton agar were inoculated by pre-grown McFarland-standard bacterial culture, then disks with antibiotics were placed and tightly pressed onto the surfaces of the inoculated plates to ensure constant diffusion of antibiotic. The plates were incubated at 37°C in a thermostat for 24 h. Inhibition zone diameter was measured to within 1 mm taking into account the diameter of the disk itself. Only the complete inhibition of visible growth was measured; the smallest colonies that were only detected under specific lighting conditions and that had a barely visible plume at the edge of the inhibition zone were not counted. Formation of the large colonies within a clear zone of growth inhibition was considered as an evidence of the presence of extraneous microflora or multiple resistance of the population. In this cases, re-identification and re-testing for antibiotic resistance from both zones were conducted. Strains were considered sensitive or resistant according to the CLSI criteria [CLSI, 2009]. *E. faecium* genomes isolated from patients were sequenced on the Illumina HiSeq 2500 platform and deposited in NCBI (project ID PRJNA41282).

Determine *Enterococcus faecium* strains relative abundances in metagenomic data

To determine strains relative abundances in metagenomes of patients we mapped metagenomic reads (Illumina) to genome sequences. For each patient bowtie2 database was built from combined assemblies of *E. faecium* colonies isolated from this patient. From closely related genomes only one representative strain was selected (i.e. because genomes of Hp_6-10_S9 and Hp_6-9_S18 are very similar

differing in a XXX SNPs only the former was included in database). Reads were mapped with bowtie2 with default parameters. Reads having mapping quality (MAPQ) less than 10 which means they are not uniquely mapped were disregarded.

ARGs search in *Enterococcus faecium* genome assemblies

The search for ARGs in the received contigs was effected using BLAST (similarity criterion: e-value < 10^{-5}, percent identity >80% for >80% of length) and MEGARes database. The phylogenetic tree was built using PARSNP [Treangen et al., 2014] by default parameters, as a reference, using the genome *E. faecium* isolate EFE1002 (https://www.ncbi.nlm.nih.gov/bioproject/PRJEB12395) from NCBI database.

Table S21. Scheme of *Enterococcus faecium* strains isolation from patients stool samples. Strains that have been sequenced shown yellow color.

Patient ID / Time point	1st	2nd
HP003	+	+
HP010	+	+
HP024	-	+
HP059	-	-
HP063	-	+

References

1. Prianichnikov NA, Malakhova MV, Babenko VV, Larin AK, Olekhnovich EI, Starikova EV, Chuvelev DI, Glushchenko OE, Samoilov AE, Manolov AI, Kovarsky BA. Draft genomes of *Enterococcus faecium* strains isolated from human feces before and after eradication therapy against *Helicobacter pylori*. Data in brief. 2018 Feb 28;16:511-4.

2. CLSI. 2009. Performance standards for antimicrobial susceptibility testing; 19th informational supplement. CLSI document M100-S19. Clinical and Laboratory Standards Institute, Wayne, PA.

3. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome biology. 2014 Nov;15(11):524.

4. Voigt AY, Costea PI, Kultima JR, Li SS, Zeller G, Sunagawa S, Bork P. Temporal and technical variability of human gut metagenomes. Genome biology. 2015 Dec;16(1):73.