31 10 2024

Hayat HAGE[1], Jérémie BECKER[1], Marion DARNAUD[1], Juan Pablo Rodriguez Ruiz[2], Leen Timbermont[2], Philippe LEISSNER[1]

[1]BIOASTER, Lyon, France. [2]University of Antwerp, Belgium.

D2.3 Report on national pneumococcal phylogenetic and antimicrobial profiles

VALUE-Dx

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 820755. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and bioMérieux SA, Janssen Pharmaceutica NV, Abbott, Bio-Rad Laboratories, BD Switzerland Sàrl, and The Wellcome Trust Limited.


University
of Antwerp


BIOMÉRIEUX


wellcome


UNIVERSITY OF
OXFORD


UMC Utrecht


Penta
Child Health Research


UNIVERSITÀ
di VERONA


THE UNIVERSITY
of EDINBURGH


University Medical Center Groningen

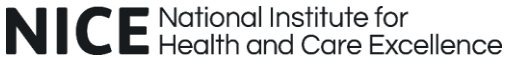

RAMBAM
Health Care Campus


UNIVERSIDAD
DE LA RIOJA


BOSTON
UNIVERSITY


BIOASTER
MICROBIOLOGY TECHNOLOGY INSTITUTE


FIND


NICE National Institute for
Health and Care Excellence


Gesundheit Österreich
GmbH


ESCMID
EUROPEAN SOCIETY
OF CLINICAL MICROBIOLOGY
AND INFECTIOUS DISEASES


ERS
EUROPEAN
RESPIRATORY
SOCIETY
every breath counts


Berry Consultants
Statistical Innovation


ZonMw


janssen


Abbott


BIO-RAD


BD

Contents

1 Introduction

Within VALUE-Dx, Task 2.1.3 aimed to quantifying antimicrobial resistance and the associated microbiome in community-acquired acute respiratory tract infections (CA-ARTI) patients.

This task utilized samples from the clinical studies PRUDENCE and ADEQUATE performed in work package 4 (WP4). Upper respiratory (PRUDENCE and ADEQUATE) and stool (only PRUDENCE) samples were collected at the time of the primary consultation and before the start of antibiotic, respectively, and then at several timepoints, to assess the impact of antibiotics on the respiratory and intestinal microbiota. DNA extraction and sequencing were carried out at the University of Antwerp, as well as the 16S sequencing analysis, while the bioinformatic and statistical analyses post shotgun metagenomic sequencing were performed at BIOASTER.

Shotgun metagenomic sequencing from oropharyngeal swabs (ADEQUATE) and stools (PRUDENCE), and 16S sequencing from nasopharyngeal swabs (PRUDENCE), allowed identification and quantification of the bacterial aetiologies in CA-ARTI patients. However, these approaches did not enable the identification of viral pathogens and therefore distinction between bacterial and viral (co)aetiologies.

Microbiota composition and differences between patients treated with antibiotics and untreated controls were studied, as well as microbiota variation across timepoints. Differential analysis of antimicrobial resistant genes was further investigated, and statistical models were applied to highlight the difference in AMR selection by antibiotics.

2 Material & Method

2.1 Covariate selection for randomization

2.1.1 Introduction

This section presents a strategy to prioritize clinical covariates used to randomize samples in the extraction step. The goal of the randomization is to ensure that confounding variables (clinical or technical) are uniformly distributed across batches, making their correction possible. Before this short analysis, discussions within the project team led to the identification of four commonly selected covariates, sex, age, symptom, treatment. Since both cohorts are concerned with community acquired acute respiratory tract infection (CA-ARTI), the sampling site also seems an important information to consider (due to variation in care). Given that the targets for inclusion are relatively low (49 and 114 patients for ADEQUATE and PRUDENCE), the number of covariates should be restricted to a handful to ensure a robust estimation of the confounding effects. The content of the clinical forms is detailed section 2.1.2, the strategy underlying the variable selection is laid out in section 2.1.3.

2.1.2 Form content

Despite the differences in metadata files between ADEQUATE and PRUDENCE, their format remains similar, enabling a shared data processing. Table 1 displays the count of samples and variables categorized by type for each file. The way metadata are structured is as follows, a first categorical variable (often binary) indicates whether the patient presents a particular symptom, has undergone a specific treatment or a clinical examination. A second variable (mainly categorical) then specifies the gravity of the symptom, the type of treatment or the examination result, from a pre-specified list (most of the time). In the case of cough for example, the first variable indicates the presence of the symptom, the second relates to the severity (mild, moderate or severe). This two-level structure was not leveraged in this work, instead we decided to focus on all categorical variables because they are easier to handle and make up the majority of covariates.

Cohort	Form	N samples	N checkbox radiobutton	N text	N integer decimal	N date
ADEQUATE	02aP - Signs and Symptoms at ER and Management Plan at Baseline	49	63	30	35	2
	02b - Vaccination	49	17	4	6	3
	02bP - Participant Background	49	13	1	3	0
	02cP - Co-morbidities and Chronic Medication	49	16	9	0	0
	03aP - Clinical Decision after Randomisation and Initial Results	49	4	1	4	2
	03b - Study Samples (only for intervention group)	24	59	2	6	2
	(Serious) Adverse Events (Paediatric)	0	27	14	2	13
PRUDENCE	Demographics_20240307	115	1	0	1	0
	FWdiag_20240307	115	1	1	0	0
	Med_20240307	115	7	1	0	0
	Medhis_20240307	115	17	4	0	0
	Prescribing_20240307	115	11	2	0	0
	SAE_20240307	3	18	16	9	8
	SAE_med_20240307	13	1	4	1	2
	SARSCOV_20240307	115	3	1	0	1

Table 1: Metadata files used for the study.

2.1.3 Strategy used for covariate selection

In this analysis, we employed a two-step strategy, first we pre-filtered covariates based on their frequencies, then, we refined them manually. The initial step involves selecting covariates that have sufficient observations across all levels to be randomized in multiple

batches. The second, prioritizes manually the covariates that are most likely to have the highest impact on the microbiota composition.

2.1.3.1 Pre-filtering step

Each run will contain 24 samples extracted manually. While all runs will include negative controls, only one will be spiked with a positive control. For simplicity we assumed that 22 clinical samples will be processed in each run. The pre-filtering step involved retaining a categorical covariate only if at least two combinations of levels can be included in triplicate in each run. The idea is that, despite some covariates may have under-populated levels, when combined together, the resulting meta-levels become sufficiently large for randomization. Practically, for a given variable, the sample size associated with each level is compared to a threshold defined as (total samples – nReplicates x nBatches) samples, with nReplicates=3 and nBatches the total number of batches needed. If all levels are smaller than this threshold, it means that the samples are not assigned to a single level. Therefore, in the most extreme case, the smallest levels can be merged in a way that at least 3 replicates of this meta-level are present in each batch. In the case of cough severity (ADEQUATE), out of a total of 49 patients, 45 were positive and 10, 24, 10 were found with mild, moderate and severe cough. The total number of patients being 49, 3 runs will be needed, leading to a threshold = 49 – (3x3) = 40. All levels are below this threshold, which implies that this covariate successfully passes the pre-filtering step. If this variable was to be used for randomization, the 4 negatives samples will have to be merged, probably with the mild patients, to ensure that 3 replicates of each level are present in each batch. 44 and 22 covariates were pre-filtered using this criterion and are presented in Table 2.

ADEQUATE	**Site name** **Fever, Severity** **Cough, Severity** **Cough with sputum production, Is symptom present?** **Cough with sputum production, Severity** **Sore throat, Is symptom present?** **Sore throat, Severity** **Nasal congestion or runny nose, Severity** **Low energy and/or tired, Severity** **Abdominal pain, Is symptom present?** **Abdominal pain, Severity** **Nausea and/or vomiting, Is symptom present?** **Nausea and/or vomiting, Severity** **Diarrhoea, Is symptom present?** **Diarrhoea, Severity** **Not sleeping well, Is symptom present?** **Not sleeping well, Severity** **Feeling generally unwell, Is symptom present?** **Feeling generally unwell, Severity** **Respiratory rate available?** **Has the patient stayed away from usual day care?** **Has the patient stayed away from usual school?"** **Has the patient stayed away from usual day care / school / work? Not applicable** **For how many days has at least one parent been absent from work for childcare in this episode of participant's illness?** **Has the patient used other health services in this episode of illness?** **Service used: Other medical specialist (e.g. paediatrician for children)**

	Service used: Pharmacist **Did the patient receive any other medication previously to this consultation?** **Medication: Pain or fever medication** **Medication: Other** **What is the suspected diagnosis?** **What is the suspected etiology?** **Were additional blood tests performed?** **Were additional tests performed?** **COVID-19 specific vaccine information collected?** **Pertussis, most recent approximate year the patient received the vaccine?** **Diphtheria, most recent approximate year the patient received the vaccine?** **Status other vaccinations?,** **Sex at birth** **Standard daytime childcare arrangement: Parent at home** **Standard daytime childcare arrangement: Creche/kindergarten** **Standard daytime childcare arrangement: School** **Prior confirmed COVID-19?** **Time test results were generated?** **Time test results were received by care team?**
PRUDENCE	Site name Sex Indicate any other prescribed medication for this illness: Paracetamol and/or NSAIDS Indicate any other prescribed medication for this illness: Other Indicate any other prescribed medication for this illness: None of the above Cardiovascular disease Diabetes Chronic respiratory condition (e.g asthma, COPD) Hepatic, hematologic, neurologic or neurodevelopmental condition Flu vaccination in the last year Does the participant have a POSITIVE SARS-CoV-2 test result for this illness episode? If yes, 1st dose If yes, 2nd dose If yes, 3rd dose Pneumococcal vaccination in last five years Medical history of antibiotic allergy/intolerance Smoking Any Comorbidities? (completed automatically) Yes Antibiotic prescribing at this stage: Why was this decision made? Probable bacterial infection Why was this decision made? Diagnostic test result (if randomized to test) Does the participant have a POSITIVE SARS-CoV-2 test result for this illness episode?

Table 2: Pre-filtered covariates for randomization.

2.1.3.2 Manual refining of the list

One can first note that the majority of the 5 covariates identified prior to the analysis (sex, age, symptom, treatment, location) are included in one or two lists, only the age, excluded because continuous, cannot be found.

In ADEQUATE, we suggest using site name, sex and age in the randomization. Although many symptoms and vaccines are returned, the antibiotic intake and smoking variables are absent from the shortlist. Examining the clinical data, we find that only 4 out of 49 patients were given antibiotics and 1 patient was reported smoking, fractions too small for effective randomization. Fever, on the other hand is reported in all patients with different level of severity (8 mild, 13 moderate and 28 severe). It could be used as a general indicator of the infection gravity.

In PRUDENCE, we also suggest using site name, sex and age to randomize samples. This time, 26 out of 115 patients were prescribed antibiotics and 13 were reported smoking, making these two variables easily amenable to randomization. The only symptoms available in PRUDENCE are adverse events that only affect 3 patients.

Despite being present in a small fraction of samples, both antibiotic intake (ADEQUATE) and adverse events (PRUDENCE) should be included in the randomization (manually) because their effect is too large to be disregarded. The selected covariates are summarized in the Table 3.

ADEQUATE	**site name, sex, age, antibiotic intake**
PRUDENCE	site name, sex, age, diabetes, obesity, COVID19, antibiotic intake

Table 3: Selected covariates for randomization.

2.2 DNA extraction & sequencing

2.2.1 NP swabs (PRUDENCE)

Nasopharyngeal (NP) swabs were collected using the DNA/RNA Shield Collection Tube w/Swab (Zymo Research). Before DNA extraction, 20µl of ZymoBIOMICS Spike-in Control II (Low Microbial Load) (Zymo research) were added to 400µl of sample. DNA extraction was performed with the ZymoBIOMICS DNA Miniprep kit (Zymo Research) according to manufacturer's instructions, and DNA concentration was measured with Qubit® dsDNA HS Assay Kit (Invitrogen). 16S amplification was performed with the KAPA HiFi ReadyMix kit (KAPA Biosystems) targeting the V3-V4. Barcoded libraries were sequenced in a MiSeq instrument with the MiSeq Reagent Kit v2 (500 cycles) (Illumina).

2.2.2 Stool samples (PRUDENCE)

Stool samples were collected using the DNA/RNA Shield-Fecal Collection Tube (Zymo Research). Tube contents were mixed using a Bag Mixer MiniMix 100 (Interscience) in Stomacher 80 sterile bags (MLS) for 3 minutes at speed 9. Before DNA extraction, 20µl of ZymoBIOMICS Spike-in Control I (High Microbial Load) (Zymo Research) were added to 400 µl of sample. DNA extraction was performed with the Fast DNA Spin Kit for Feces (MP Biomedicals) according to manufacturer's instructions, and DNA concentration was measured with Qubit® dsDNA HS Assay Kit (Invitrogen). Finally, libraries were prepared with the Nextera® XT DNA Sample Preparation Kit (Illumina) and sequenced in a NextSeq instrument with the NextSeq 500/550 High Output kit v2.5 (300 cycles) (Illumina).

2.2.3 OP swabs (ADEQUATE)

Oropharyngeal (OP) swabs were collected using the DNA/RNA Shield Collection Tube w/Swab (Zymo Research). Before DNA extraction, 20µl of ZymoBIOMICS Spike-in Control II (Low Microbial Load) (Zymo research) were added to 400µl of sample. DNA extraction was performed with the Fast DNA Spin Kit for Feces (MP Biomedicals) according to

manufacturer's instructions, and DNA concentration was measured with Qubit® dsDNA HS Assay Kit (Invitrogen). Finally, libraries were prepared with the Nextera® XT DNA Sample Preparation Kit (Illumina) and sequenced in a NextSeq instrument with the NextSeq 500/550 High Output kit v2.5 (300 cycles) (Illumina).

2.3 Bioinformatic analyses post 16S sequencing of NP (PRUDENCE) samples

Quality of the reads was assessed using FastQC v0.11.9 and quality trimming was performed with TrimGalore v0.6.4 using a quality threshold of 30. Trimmed reads were aligned to the SILVA database v.132 and clustered into operational taxonomical units (OTUs) using mothur v1.44.1. Sequencing error rate was assessed with mothur by incorporating a well-characterized mock community (D6300, Zymo Research), resulting in an average error rate of 0.67% (0.57-0.79%). Negative controls during extraction and library preparation were included and used to remove contaminants from the OTU table with decontam v.1.14.0. Rare taxa (<6 classified reads across entire dataset) were removed prior to further analysis. Shannon alpha diversity and Bray-Curtis dissimilarity matrix were calculated using Vegan 2.6.4. Differences in alpha-diversity between groups were analysed using ANOVA, and Adonis was used to evaluate variables influencing the community dissimilarity. DESeq2 v1.34.0 was used to investigate differentially abundant OTUs between timepoints and antibiotic treatment.

2.4 Bioinformatic analyses post shotgun metagenomic sequencing of OP (ADEQUATE) and stool (PRUDENCE) samples

2.4.1 Reads preprocessing & quality control

The quality control workflow is presented in Figure 1. Briefly, the quality of the generated sequencing reads was assessed using FastQC (v0.11.9) (Andrews. 2010) with default parameters. Fastp (v0.23.4) (Chen et al. 2018) was then used to filter out low-quality reads, discarding reads with more than 5 N bases or those shorter than 15 bases after trimming.

High-quality reads that passed filtering were then aligned to the host genome using Bowtie2 (v2.5.0) (Langmead et Salzberg. 2012). The human reference genome (T2T5-CHM13v2.0, 2022) was used for this alignment. Reads aligning to the human genome, considered contamination, were discarded.

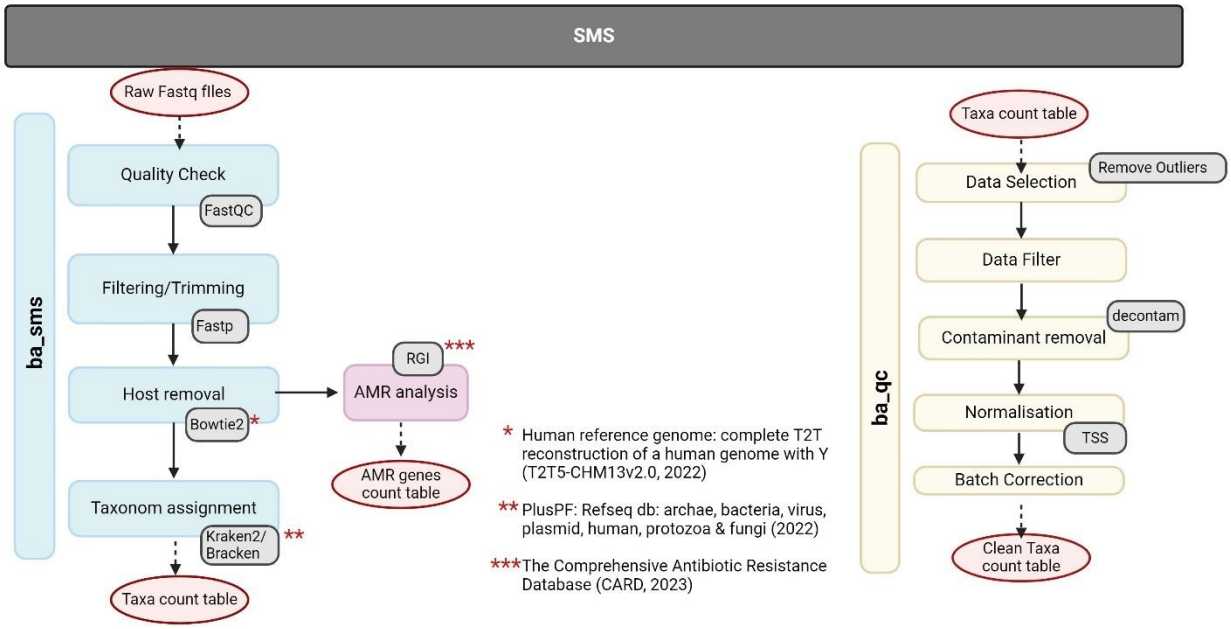


Figure 1: Data quality control, taxonomy analysis and resistance gene analysis workflow.

2.4.2 Mapping and quantification

Non-host reads were then aligned to the RefSeq PlusPF database for taxonomic classification using Kraken2 (v2.1.1) (Wood, Lu, et Langmead. 2019), followed by refinement with Bracken (v2.7) (Lu et al. 2017) to improve the accuracy of species-level abundance estimates. RefSeq PlusPF database contains archaea, bacteria, viral, plasmid, fungi, protozoa and Human. It is important to have the host genome included in the database in order to catch any human read that did not align in the previous step with bowtie2.

The R decontam package was used to identify and remove contaminant species that were not truly present in the sampled community. This contaminating DNA can come from several sources and can be identified from negative control samples in which sequencing was performed on blanks without any biological sample added. We used the prevalence method which assumes that contaminants will be more prevalent in negative control samples compared to true biological samples.

2.4.3 Resistance gene analysis

Resistance gene identifier (RGI) v.6.0.3 was used to map reads against Comprehensive Antibiotic Resistance Database (CARD) database v.3.2.9 (McArthur et al. 2013). The command rgi bwt was used with the aligner bowtie2. AMR genes identified with less than 100 mapped reads were filtered out.

2.4.4 Normalization and variance partitioning

The taxa table was normalized using TMM (Robinson and Oshlack. 2010, *EdgeR* v4.2.0 package), the resulting factors were then applied onto the resistance gene table. An alternative method, additive log-ratio (*Compositions* v2.0-8), which leverages spike-ins, was also applied and compared to TMM. Confounding effects were corrected using removeBatchEffect (*Limma* v3.60.3).

Variance partitioning analysis was performed to evaluate the proportion of variance explained by each technical and clinical variable (*variancePartition* v.1.22.0). This step was carried out before and after correction of confounding effects to ensure the latter were properly adjusted for.

2.4.5 Alpha & beta diversity

Alpha and beta diversity coefficients are metrics derived from ecology. The former describes the composition of each sample independently, while the latter quantifies the (dis)similarities between groups of samples.

Applied to metagenomics, alpha diversity provides information on richness (number of species) and evenness of a given sample. Richness is generally estimated as the number of species, with an additional corrective term to account for species that are rare (missed during sampling). Evenness indicates whether species are present in similar abundances or if, conversely, a small number of them dominate the biomass. Beta diversity measures the difference in bacterial composition between two groups of samples by first calculating pairwise distances between samples and then projecting these distances into a low-dimensional space.

In this project, we used two standard measures of alpha diversity: the Chao estimator for richness and the Shannon index for evenness. The Bray-Curtis distance, combined with the Adonis test (also known as permutational MANOVA), were further chosen to calculate beta diversity and assess the significance of between group differences (*Vegan* R package, v2.6-6.1).

2.4.6 Differential analysis

Differential analysis was performed with limma-voom (*Limma,* v3.60.3), an approach well suited to longitudinal analysis due to its ability to account for intra-individual correlations. Unlike other methods, Limma effectively incorporates a random effect for patients, similar to a mixed-effects model. Limma-Voom combines the generalized linear models of Limma (moderated t-test based on pooling variance across features) with an estimation of the mean-variance relationship through a weighting of observations via Voom.

Only features (taxa or resistant genes) passing background filters were retained, reducing the multiple testing correction (Benjamini and Hochberg) on the p-values. In the user manual, the authors recommend setting the threshold such that the mean-variance trend decreases monotonically. The values obtained here were log2(cpm) = 2.5 and 7.5 on the taxa and resistance gene datasets.

Limma-Voom requires the input data to be raw counts (from the alignment steps), which implies that confounding effects must be included in the model to be corrected. As mentioned earlier, the individual effect was modelled as a random effect, while the others were modelled as fixed effects.

In this work, we conducted two types of comparisons: intra-group and inter-group. The intra-group comparison assesses the same individuals at two different time points, while the inter-group comparison contrasts two groups at a single time point, after adjusting for baselines (Figure 2). For instance, the inter-group comparisons between antibiotic and untreated patients at D7 is expressed as: (ATB_D7-ATB_D1)-(Untr_D7-Untr_D1). This "double contrast" approach ensures that the inter-group difference ATB_D7-Untr_D7 is unaffected by baseline heterogeneity between antibiotic and untreated groups.

A feature (taxon or resistance gene) was deemed significant if the adjusted p-value was < 5% (Benjamini-Hochberg's procedure) and $|\log_2(\text{fold-change})| > 0.5$. The relatively permissive threshold on $\log_2(\text{fold-change})$ was chosen due to the limited number of findings in the study. Differential analysis results are provided in supplementary materials, enabling consortium members to refine the results using alternative thresholds.

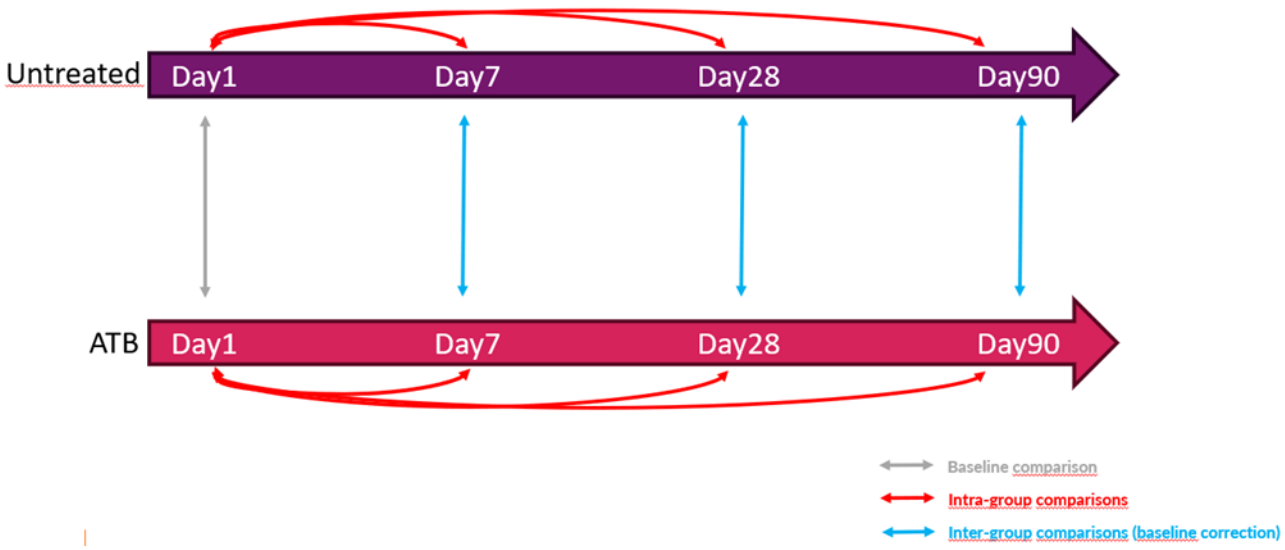


Figure 2: Illustration of intra and inter group comparisons used in Prudence.

2.4.7 Longitudinal analysis

Two approaches were considered to identify features significantly affected by the time. (i) A likelihood ratio test comparing two nested models, a full model and a reduced model, to assess whether the treatment induces changes in taxa abundance at any time point after baseline. The full model accounts for confounding effects and includes treatment, time factors, and their interaction, while the reduced model excludes the interaction. (ii) All significant taxa from intra-group differential analyses were pooled. The first approach identifies taxa with trajectories that differ between groups, while the second captures taxa that differ from baseline at one or more time points. Using Gaussian Mixture Model

(MClust package v6.1.1), an additional step of clustering was applied to group taxa according to their trajectories.

3 Results

3.1 PRUDENCE: Respiratory microbiota

3.2 Sample collection

Participants of the PRUDENCE trial (adults attending primary care or patients residing in long-term care facilities) were randomized to undergo standard care or rapid diagnostic testing for Influenza, Group A Streptococcus or C-Reactive Protein. Participants were approached for participation in the microbiology study, and those consenting provided nasopharyngeal (NP) samples at randomization, Day 7 and Day 28. Upon sample collection, they were frozen at -80°C until shipment to the UA biobank. In total, 309 samples obtained from 113 patients were processed.

3.2.1 Sequencing quality control

Overall, 36171 OTUs were obtained, of which 100 were deemed as contaminants by decontam based on comparison to the negative controls used during extraction and library preparation. After rare taxa filtering, 2733 OTUs were retained for further analysis, and rarefaction curves showed that they microbial communities of most samples are well represented (Figure 3).

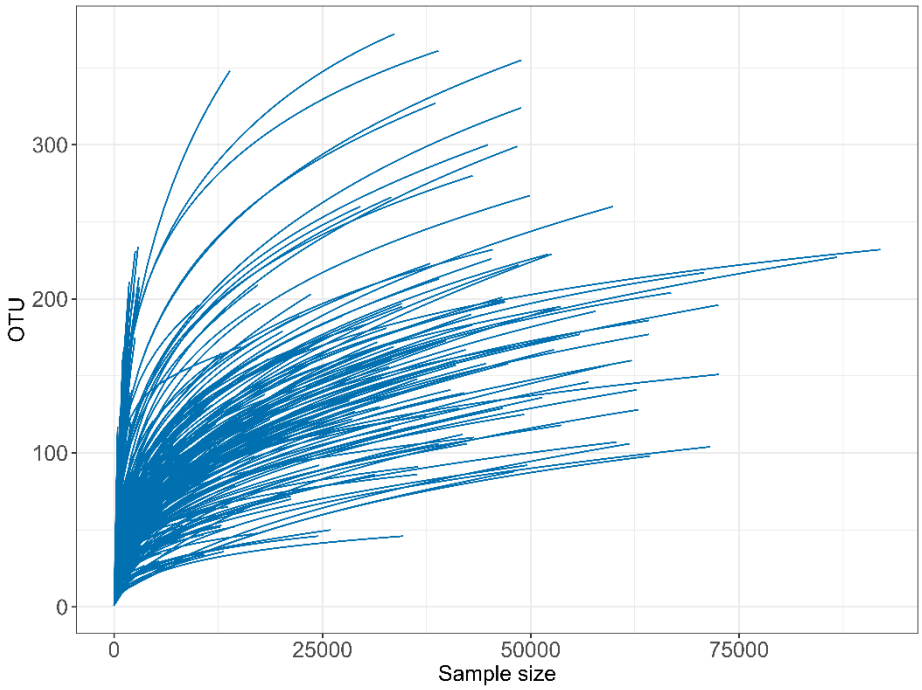


Figure 3: Rarefaction curves showing the number of detected species at various sequencing depth, each curve representing a sample.

3.2.2 Diversity analysis

3.2.2.1 Alpha diversity

Alpha-diversity did not show significant differences between any of the studied groups, including timepoint, prior antibiotic treatment, antibiotic treatment at randomization visit, antibiotic treatment after first visit, country, site and gender (Figure 4).

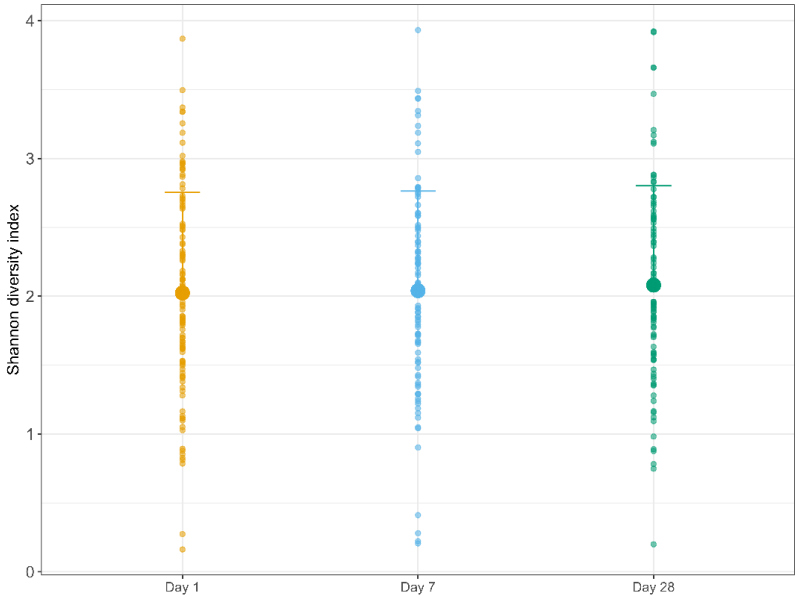


Figure 4: Alpha diversity.

3.2.2.2 Beta diversity

On the other hand, several variables represented a significant influence in community dissimilarity, including site (highest weight), country, gender, extraction batch and sequencing run (Figure 5).

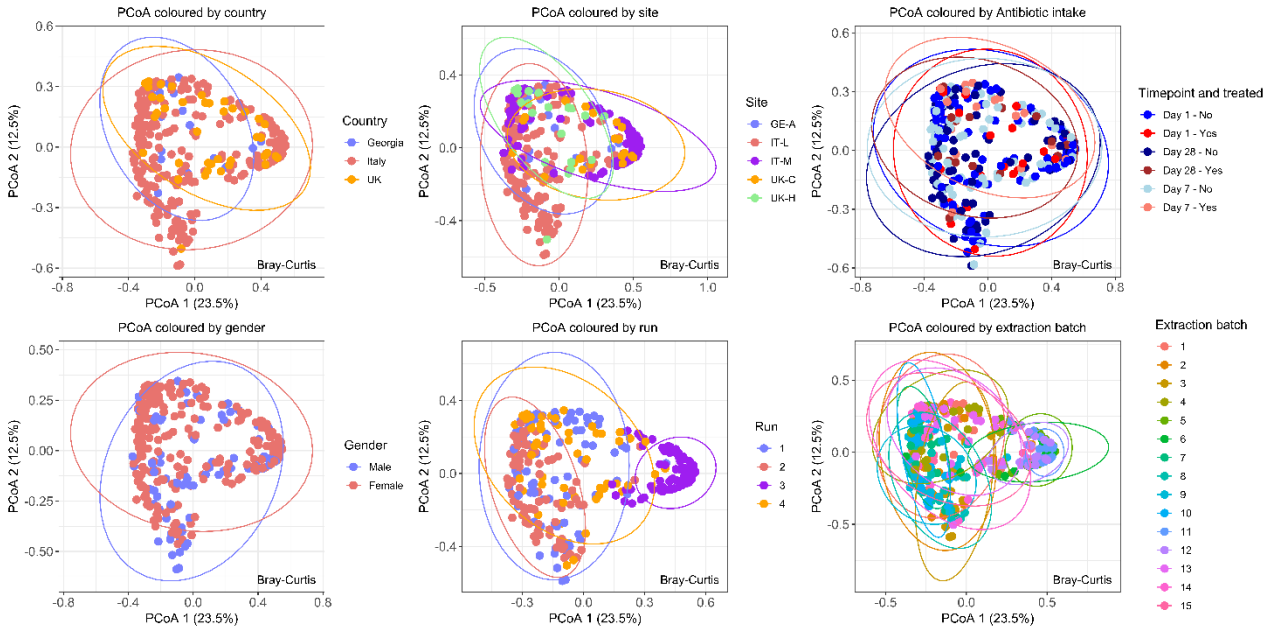


Figure 5: Beta diversity.

3.2.3 Differential analysis

A small amount of OTUs was differentially abundant between patients treated with antibiotics and those untreated at each timepoint (Figure 6). At Day 1, 13 OTUs showed differential abundance, of which only OTU00003 (*Staphylococcus*) had more than 100 average reads across the dataset. At Day 7, 15 OTUs showed differential abundance, of which only OTU00005 (*Corynebacterium*) and OTU00006 (*Dolosigranulum*) had more than 100 average reads across the dataset. At Day 28, 8 OTUs showed differential abundance, of which only OTU00002 (*Corynebacterium*) had more than 100 average reads across the dataset. All these differentially abundant OTUs were more abundant in the untreated group. Especially relevant is the depletion of *Corynebacterium* and *Dolosigranulum* at Day 7, after antibiotic treatment, as these are keystone taxa associated to health and lower recurrence of upper respiratory tract infections.

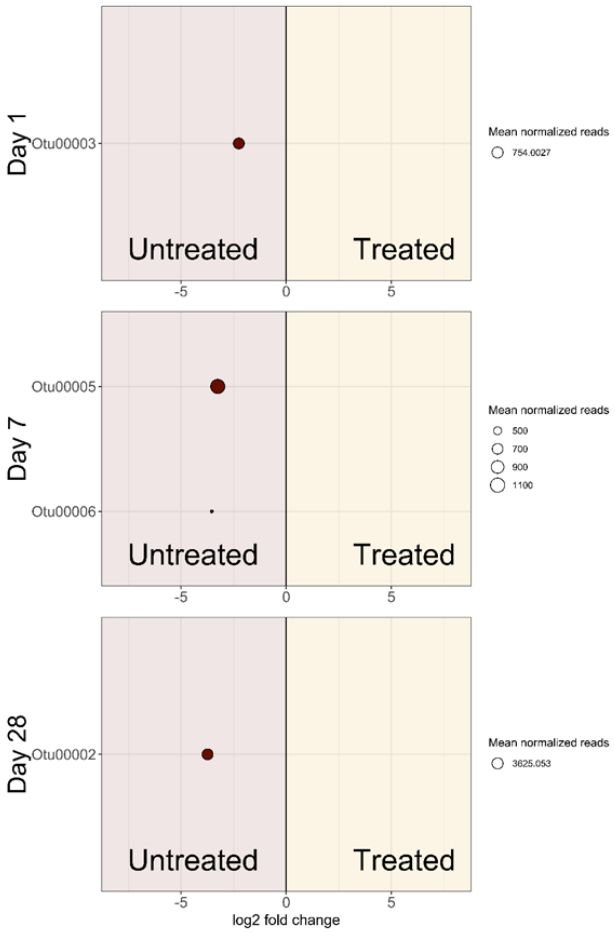


Figure 6: Differential analysis.

When inspecting the abundance of genera, it can be noticed that abundance of potential pathogens like *Streptococcus*, *Haemophilus* and *Moraxella* is overall low, and the nasopharyngeal microbial community is composed mainly of *Corynebacterium*, *Dolosigranulum* and *Staphylococcus*, with differences between countries and sites (Figure 7).

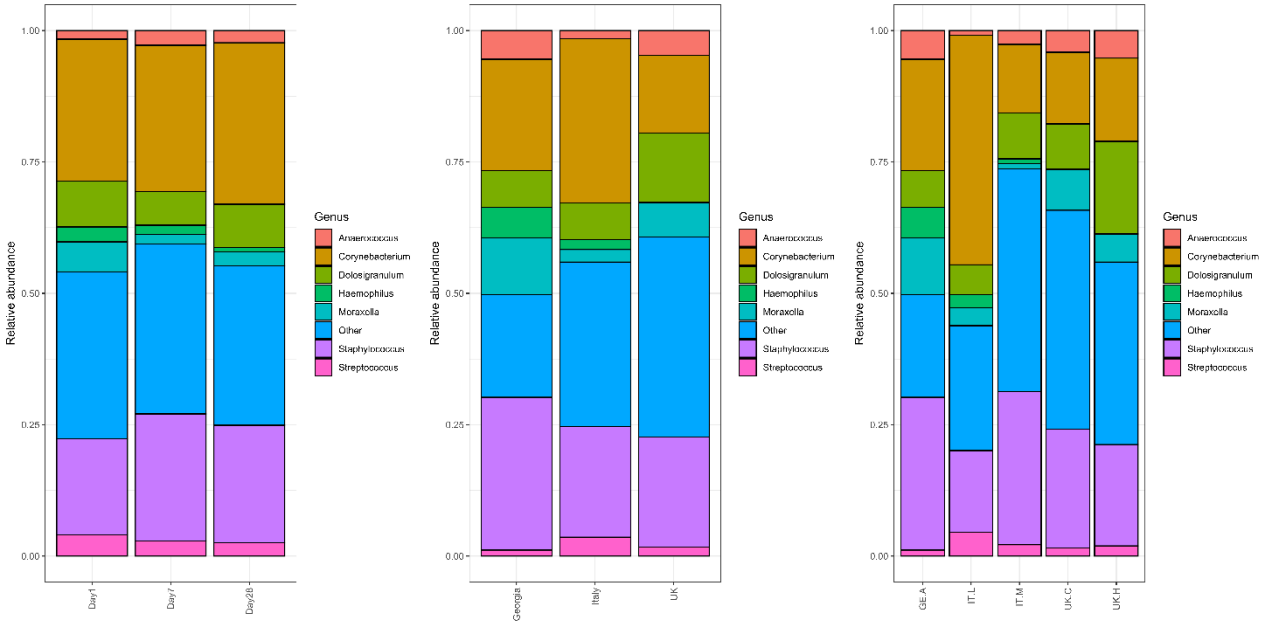


Figure 7: Relative abundance of the nasopharyngeal microbial community at the genus level.

The genera distribution of the mock communities shows a large proportion of OTUs misclassified as Other, indicating a large presence of spurious results (Figure 8). Additionally, *Listeria* could only be classified correctly in 4/17 positive controls, corresponding to decreased reads classified as Enterobacteriaceae_unclassified, which are associated to *Salmonella*. Additionally, one positive control showed a low number of reads, mostly classified as Other, indicating issues during extraction or library preparation that need to be further investigated. Overall, positive control taxonomy distribution is distinct from the rest of samples, clustered in sample Total.

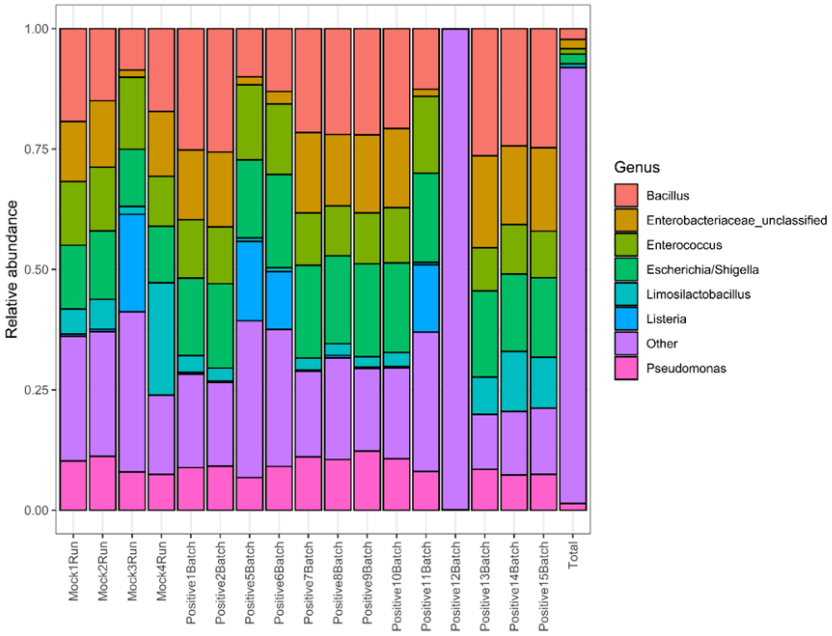


Figure 8: Genera distribution of the mock communities.

Regarding negative controls, after contaminant removal, taxonomy composition was distinct form the actual samples. Besides other genera with low number of counts, usual members of the negative controls included *Acinetobacter*, Enterobacteriaceae, *Escherichia*, *Herbaspirillum*, *Pelomonas*, *Pseudomonas* and *Ralstonia*, which are not expected in nasopharyngeal environment (Figure 9).

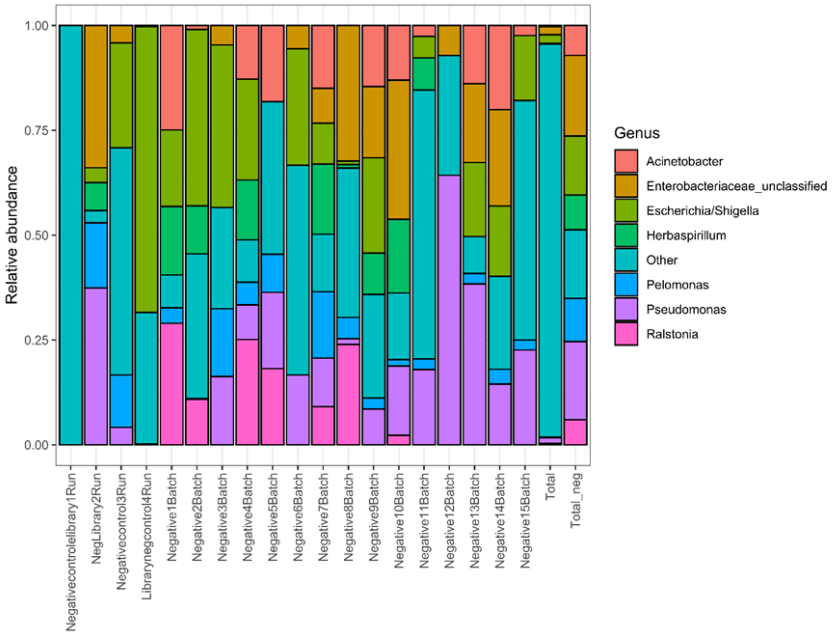


Figure 9: Genera distribution of the negative controls.

3.3 PRUDENCE: Intestinal microbiota

3.3.1 Evaluation of data quality

3.3.1.1 Sample distribution

A total of 330 stool samples were collected from 102 unique patients across 5 sites (GE-A, IT-L, IT-M, UK-C, UK-H) at 4 time points: Day 1, Day 7, Day 28, and Day 90 (denoted as D1, D7, D28, and D90 hereafter). On average, 3.2 time points were sampled per patient, with 13 patients missing the baseline (D1) sample. A maximum of 22 samples could be extracted simultaneously, resulting in 15 extraction batches, which were then pooled into 8 sequencing runs. To detect and remove potential contaminants, a negative control was included in each extraction batch.

The 330 samples were randomized in a way that confounding variables were uniformly distributed across extraction batches, making thus their correction possible in downstream analyses. The workflow used to select clinical covariates for randomization led to the selection of site, sex, age, diabetes, obesity, COVID-19 test, visit (time point) and antibiotic intake. The last two variables are the main variable of interest, the other are confounders that were adjusted for. Figure 10 illustrates that the distribution of these 8 clinical variables is comparable to that of the full dataset, ensuring thus that both technical and clinical effects can be accurately estimated.

Table 4 summarises the sample distribution across time points and antibiotic classes. Two classes, co-amoxiclav and “Other”, contain one sample per time point and were therefore removed from downstream analyses. Similarly, quinolone presents a single sample at baseline making this class not amenable to differential analysis.

Note that some patients showed conflicting meta data, including:

- An antibiotic was prescribed but intake was not recorded at any visit (patients IT-M-004, IT-M-021 and IT-M-023). Although IT-M-004 might just be a case of delayed therapy, so prescribed antibiotic would be “No” and duration “None taken”
- No antibiotic was prescribed nor recorded, but every visit shows antibiotic intake (patient IT-L043)

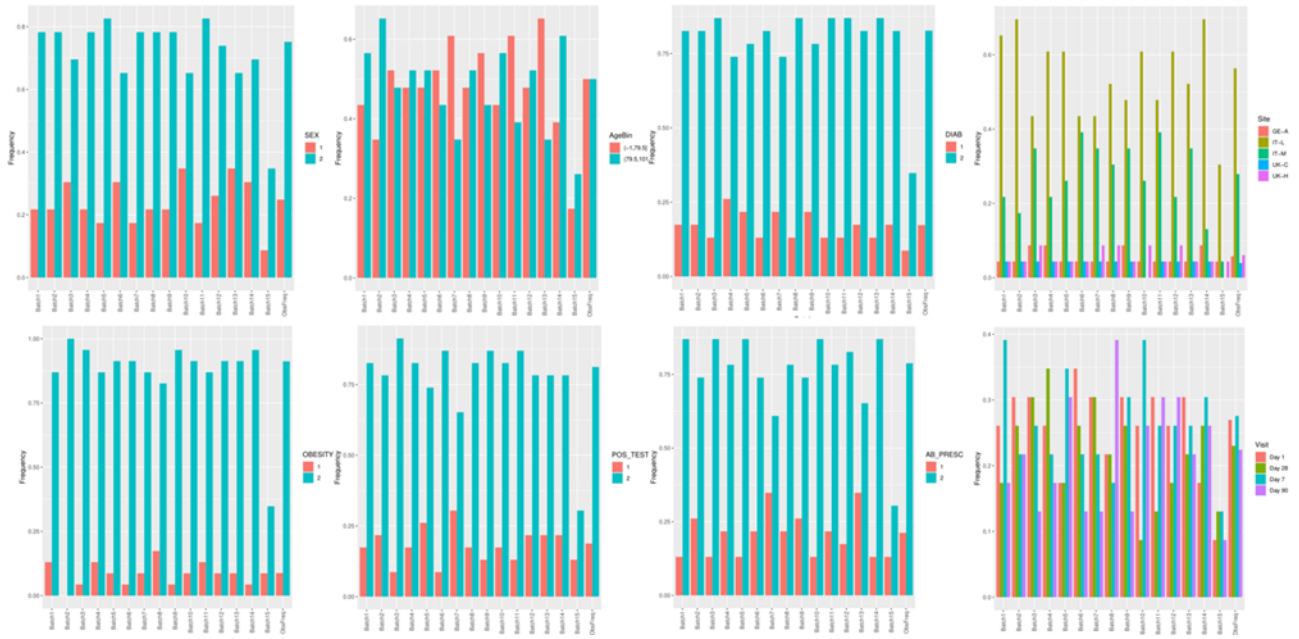


Figure 10: Sample randomization into extraction batches. Each barplot shows, for a given variable, the sample distribution in the 15 batches and the full dataset (observed frequencies, the rightmost category).

	Day1	Day7	Day28	Day90
Broad_spectrum_penicillin	**4**	**5**	**4**	**5**
Co_amoxiclav	1	1	1	1
Macrolide	**5**	**5**	**3**	**4**
Narrow_spectrum_penicillin	**4**	**2**	**4**	**5**
Other	1	1	1	1
Quinolone	1	4	3	2
Untreated	71	71	59	56

Table 4: Sample distribution across time points and antibiotic classes.

3.3.1.2 Sequencing quality control

Sequencing produced high-quality reads in each sample, with a Phred score of Q30, indicating excellent base call accuracy. A Q30 score corresponds to 99.9% accuracy, meaning there is only a 0.1% chance of an incorrect base call.

Around 10 million sequencing reads were generated per samples (Figure 11, top). No human DNA contamination was detected in the samples. Only two samples—IT-M-020_Day1 and IT_L_045_Day7—had fewer than 1 million microbial reads and were considered outliers due to their significantly lower read counts.

Thanks to the high sequencing depth and absence of host contamination in all the samples, the rarefaction curves reach a plateau (Figure 11, middle). This indicates that, with the current sequencing depth, all species potentially present in the samples could be detected.

A further indicator was used to investigate the number of detected species relative to the sequencing depth (Figure 11, bottom). Iterations between differential analysis and this scatter plot allowed to flag 5 additional outliers that, when included, led to significant results at log2(Fold-Change)=0 (IT-L-030_Day28, IT_L_019_Day1, IT-L-032_Day7, IT_L_032_Day1, IT_L_015_Day7).

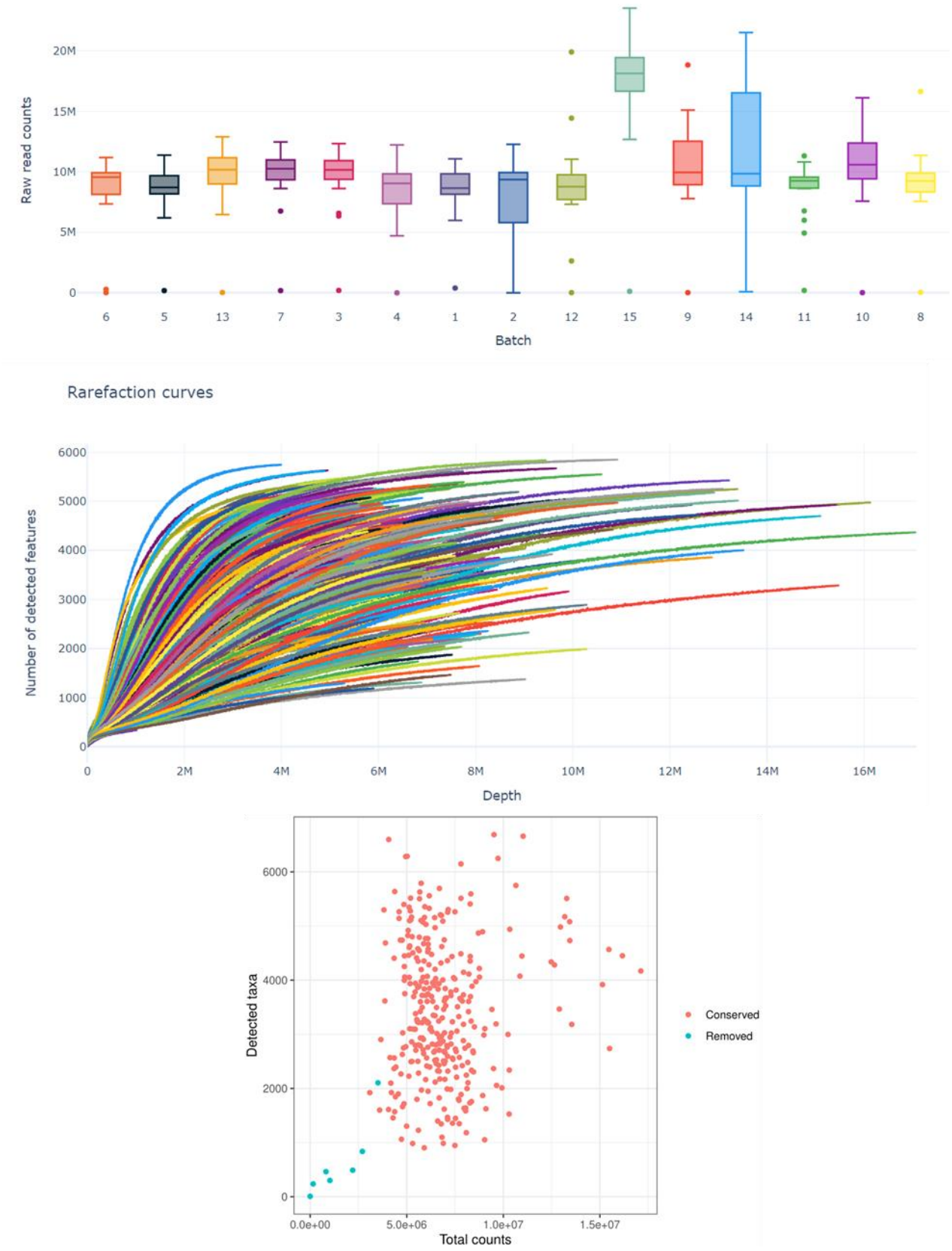


Figure 11: (top) Number of generated reads per sample. Samples are grouped by extraction batch. (middle) Rarefaction curves showing the number of detected species at various sequencing depth, each curve representing a sample. (bottom) Detected taxa versus total counts (sequencing depth).

3.3.1.3 Spike-ins and contaminating species

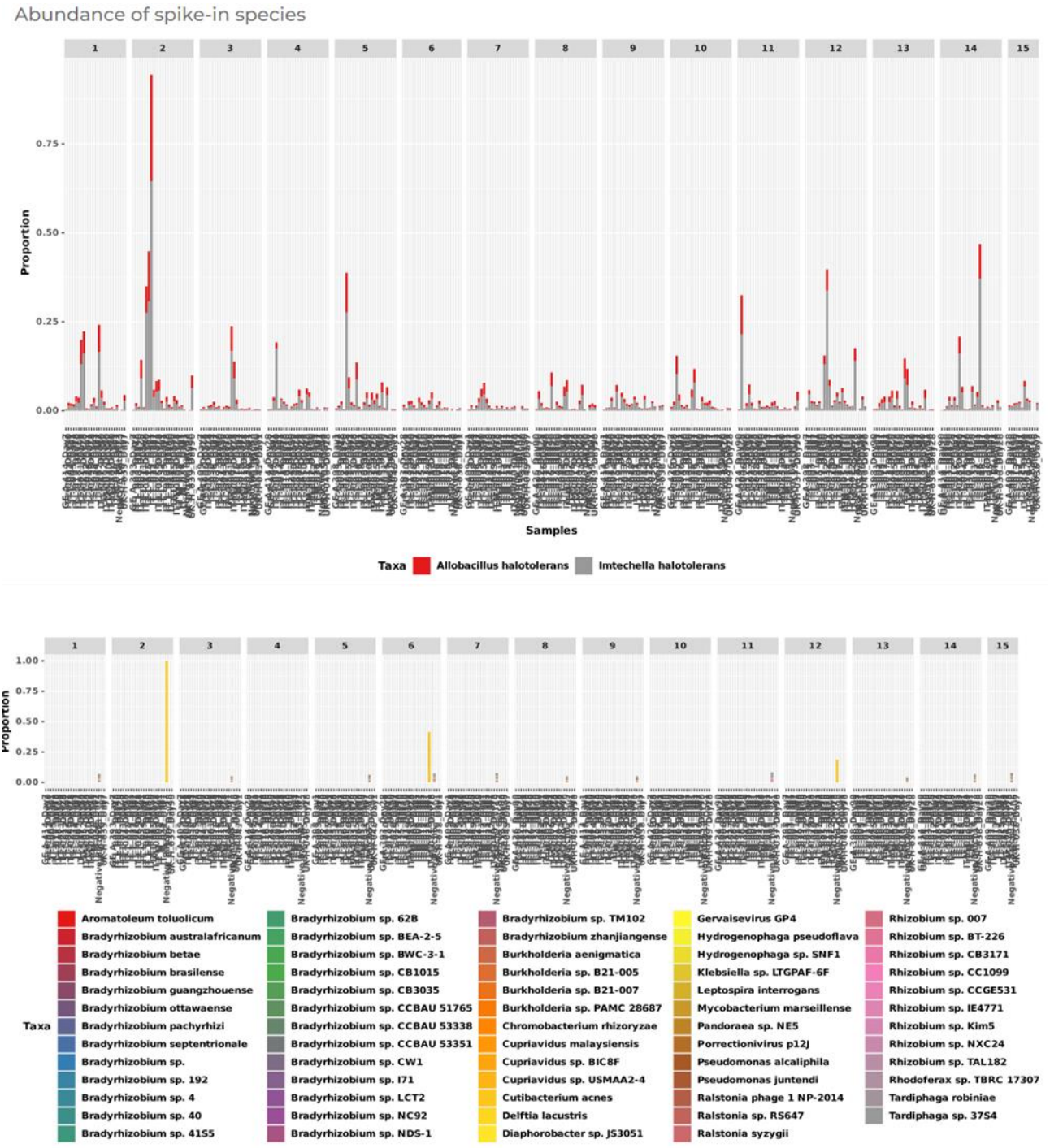


Figure 12: (top) Barplot of abundance of three spike-in species in the samples. Samples are grouped by batch. (bottom) Barplot of abundance of contaminant species in the samples. Samples are grouped by batch.

Two spike-in species, *Imtechella halotolerans* and *Allobacillus halotolerans*, were added to each sample to monitor potential biases during DNA extraction or library preparation. The expected ratio between these species was 10:1, with *Imtechella halotolerans* being 10 times more abundant than *Allobacillus halotolerans*. Overall, the total abundance of these spike-in species was relatively consistent and low across samples, except for a few

cases where their abundance exceeded 25% (Figure 12, top). Notably, in sample IT_L_019_Day1, spike-in species accounted for 94% of the total reads. These spike-in species were removed prior to downstream analysis to avoid skewing the results.

A high number of species have been detected in the negative control samples even with more stringent threshold of 0.1. Those contaminant species have been removed from the following analysis (Figure 12, bottom).

3.3.1.4 Data normalization

Two normalization methods were evaluated in this study: TMM, a widely used approach for NGS data and default in Limma-Voom, and additive log-ratio (ALR), which leverages spike-in controls. A moderate correlation was found ($R^2 = 0.6$, Figure 13), suggesting that normalization choice may significantly impact downstream analyses. In the absence of orthogonal data to favour one method over the other, differential analyses were conducted using both methods.

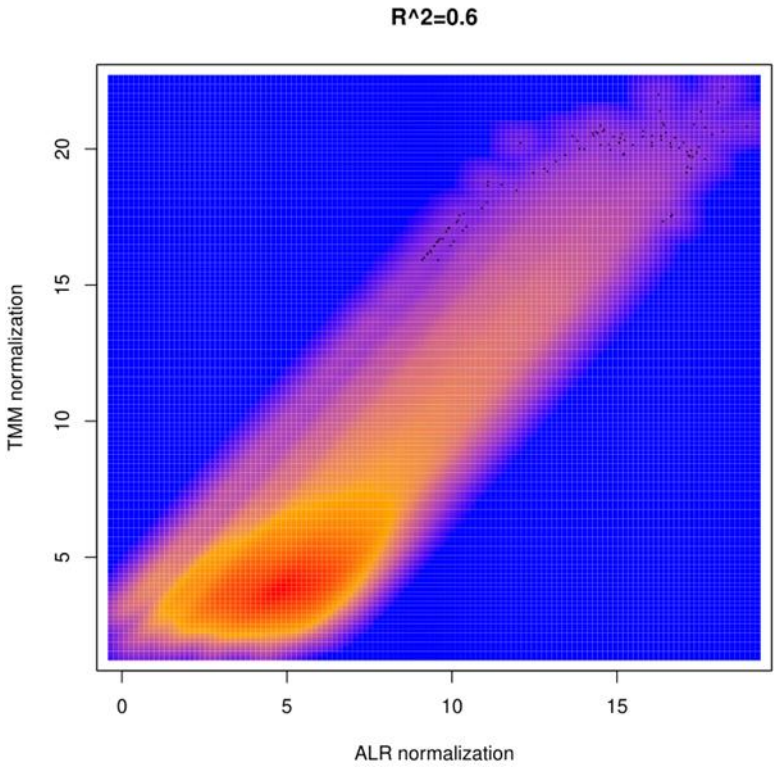


Figure 13: TMM versus ALR normalization scatterplot.

3.3.1.5 Variance partitioning

The variance partitioning step provides a visual way to examine the relative contribution of confounders and variables of interest to the overall variance. It can first be noted that the residual variance is high, indicating that the majority of the variance remains unexplained by the clinical and technical covariates. Before correction, site, age and extraction factors have the largest effect. After adjusting for confounders, their

contribution becomes negligible, leaving only the effects of antibiotics and visit remain (Figure 14).

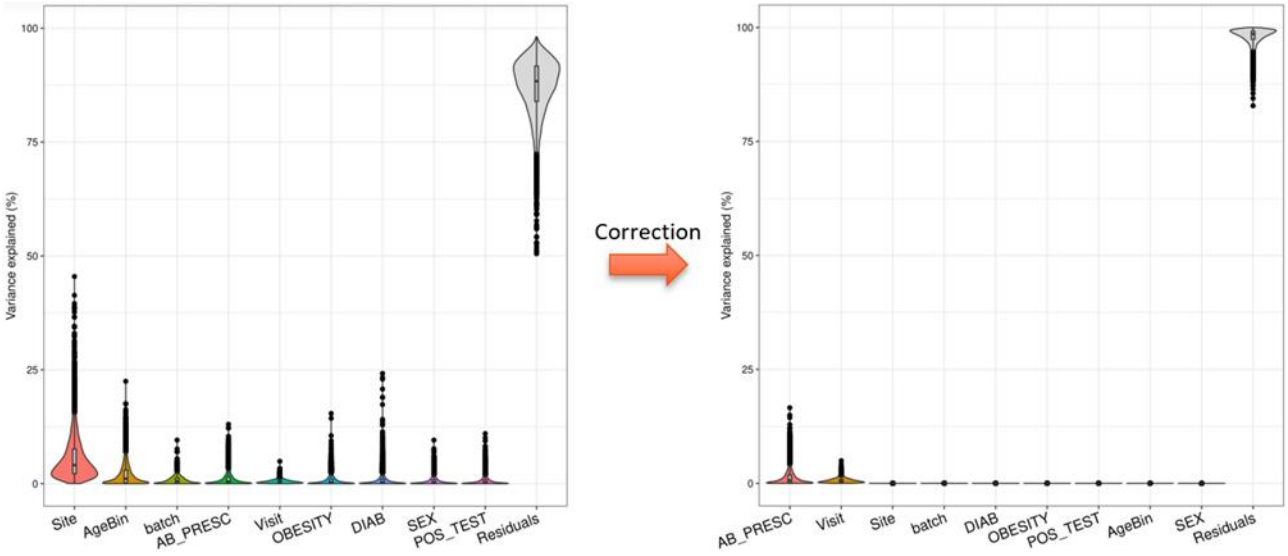


Figure 14: Variance partitioning before and after correction of confounding effects.

3.3.2 Diversity analysis

3.3.2.1 Alpha diversity

Richness and evenness were computed with the Chao estimator and Shannon index for each sample. Both metrics, like the read counts, were adjusted for confounding effects. A Student's t-test was performed at each time point to compare antibiotic-treated groups (analysed both by individual classes and combined) with the untreated group. The Co-amoxiclav and 'Other' antibiotic classes were excluded due to limited sample size. It is expected that antibiotic treatments induce a microbiota dysbiosis, impacting thus the number of species and their relative abundance. Figure 15 reveals that no difference is significant when considering all antibiotic classes together (p-values > 5%). It can nevertheless be noticed that while the median richness is higher in antibiotic-treated patients at day 1 and 90, an opposite trend is observed for evenness at the same time points. At the class level, only one difference is found significant at D7 between macrolide and untreated patients (p-value=2.5%). Although quinolone displays a large effect at baseline and D28, it is not significant due to small sample size.

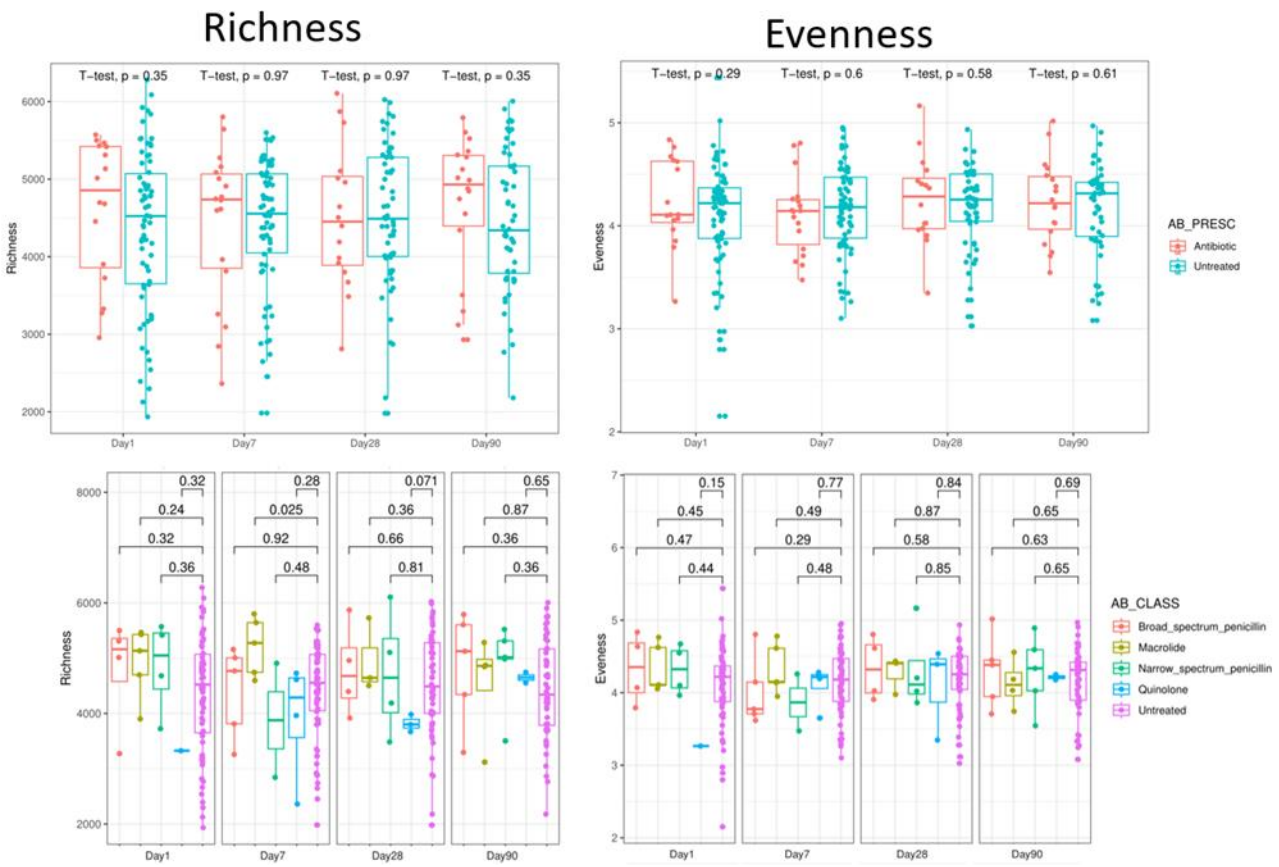


Figure 15: Richness and evenness boxplots grouped (top) or stratified by antibiotic class (bottom). A Student's t-test was performed at each time point to compare antibiotic-treated and untreated patients.

3.3.2.2 Beta diversity

To compute the beta diversity, the Bray-Curtis distance matrix was built on the normalized, adjusted counts. The Adonis test was run to test whether there is a difference in the centroid and dispersion of the groups in the multivariate space. Although significant differences were detected at the antibiotic level across all time points, the effect sizes are modest, consistent with the findings with alpha diversity. By contrast, all p-values exceed 5% at the class level, despite substantial different ellipse shapes and directions. Overall, beside the Quinolone class whose centroid appears slightly shifted from the other groups at D7, no clear trend can be deduced from the beta diversity (Figure 16).

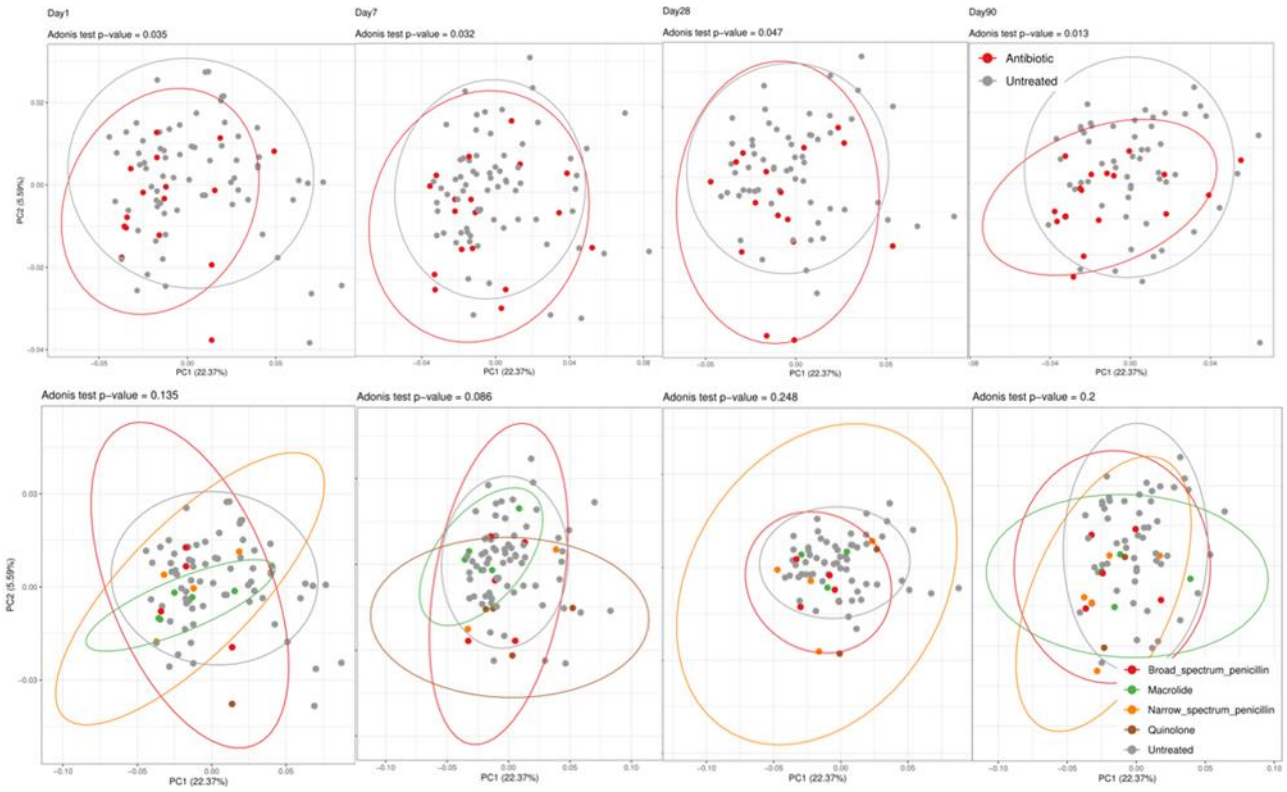


Figure 16: Beta diversity calculated using Bray-Curtis distance, shown at the antibiotic (top) and class (bottom) levels. An Adonis (PERMANOVA) was performed at each time point to compare antibiotic-treated and untreated patients.

3.3.3 Differential and longitudinal analyses in taxa

This section presents intra- and inter-group comparisons at the taxon level, considering 5 different response variables:

- Antibiotic versus untreated
- 3 Individual antibiotic (narrow and broad spectrum penicillin, macrolide) classes versus untreated
- COVID-19 positive versus negative test

For each comparison, a volcano plot and a result table with average expression in $\log_2$(cpm), $\log_2$(fold-change), p-values, were generated and are available in supplementary materials (Supplementary Tables 1-6, Figures 1-6).

3.3.3.1 Most abundant taxa

Before getting into differential analyses, this section describes the most abundant taxa. Out of a total of 7291 taxa detected in all samples, a handful of species (25 shown in Figure 17) make up approximately half of the biomass. Out of the 17 genera detected in these most abundant bacteria, 5 belong to genus previously described as dominant in the human gut microbiota (***Bacteroides***, *Prevotella*, ***Alistipes***, ***Akkermansia***, *Oscillibacter*, *Clostridium*, ***Faecalibacterium***, *Eubacterium*, *Ruminococcus*, *Roseburia*, *and* ***Bifidobacterium***) (Segata et al. 2012; Tremaroli et Bäckhed. 2012). Of note, none of these

top 25 most abundant bacteria are found differentially abundant in the following analyses.

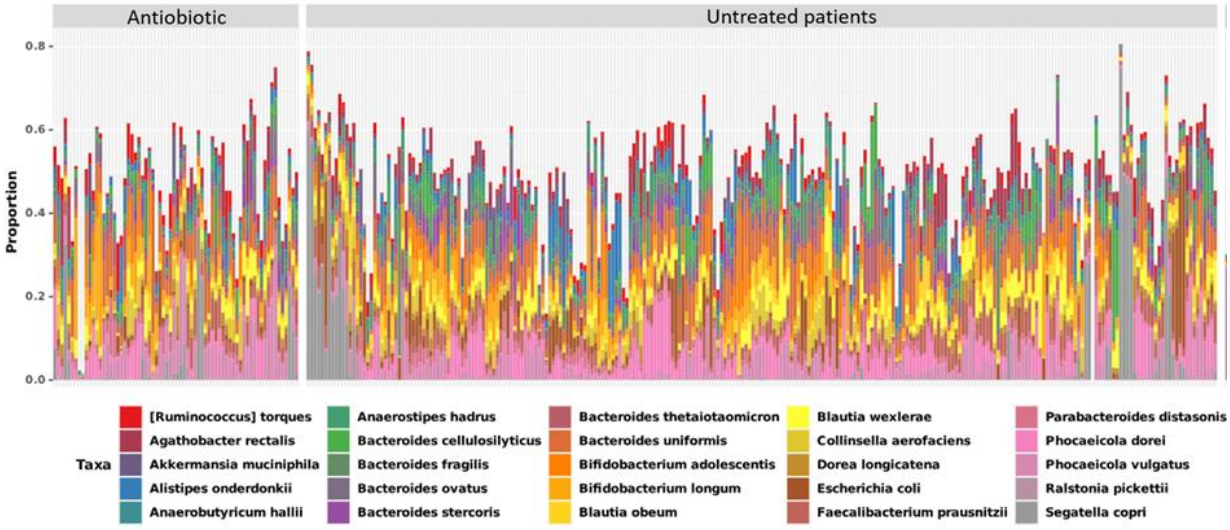


Figure 17: Relative abundance of the top 25 taxa stratified grouped by treatment.

3.3.3.2 Background noise determination

The authors of Limma-Voom recommend adjusting the background noise threshold so that the variance strictly decreases with the mean abundance. In case of bell-shaped curve, the left arm of the curve indicates a reduction in variance at low abundance levels, primarily due to the high number of zeros. Figure 18 reveals that the mean-variance trend resembles a bell-shaped curve, with many taxa at low abundance levels. Applying a background noise threshold of $\log_2(\text{cpm}) > 2.5$ resulted in the removal of 3953 taxa out of 7291.

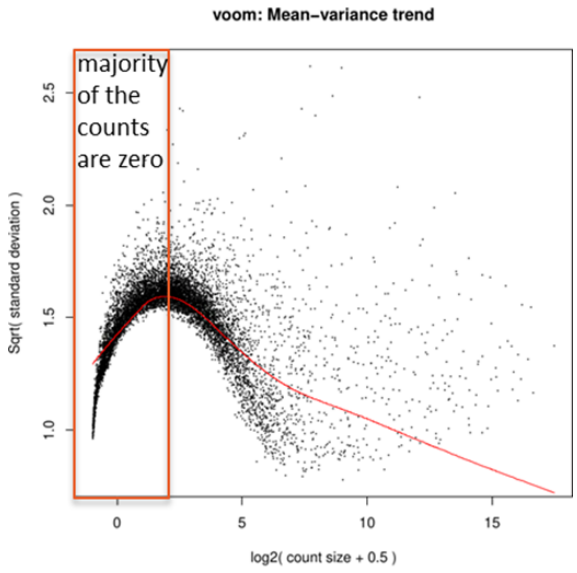


Figure 18: Mean-variance scatterplot. Each dot is the estimated mean and sqrt (standard deviation) of a given taxon. The red curve is a regression spline that estimate the overall mean-variance relationships. All taxa with mean $\log_2(\text{cpm})$ <2.5 were excluded from differential analyses.

3.3.3.3 Comparison at baseline

An initial comparison was conducted at baseline on the 5 variables of interest. Although many taxa showed |log2(fold-change)| values up to 5 and were significant before multiple testing correction (Figure 19), only 2 remained significant after correction in antibiotic and macrolide inter-group comparisons (*P. massiliensis* and *A. propionicum* respectively, see Table 5). In contrast to the antibiotic and the 3 classes comparisons that feature a large fraction of taxa with |log2(fold-change)|>1, the COVID-19 comparison includes a handful of taxa that exceed this threshold, in line with the weak COVID-19 effect estimated in variance partitioning. Of note, *Ruminococcus* sp. FMB CY1 presents a large log2(fold-change) of 9.2 in broad spectrum penicillin, without being significant though.

On top of the top10 most significant taxa, Table 5 also includes 5 taxa commonly studied in respiratory related pathologies (*Pseudomonas aeruginosa, Streptococcus pneumoniae, Klebsiella pneumoniae, Haemophilus influenzae, Staphylococcus aureus*). None of those 5 taxa reached the adjusted p-values = 5% threshold (full list available in Supplementary Figure 1-3).

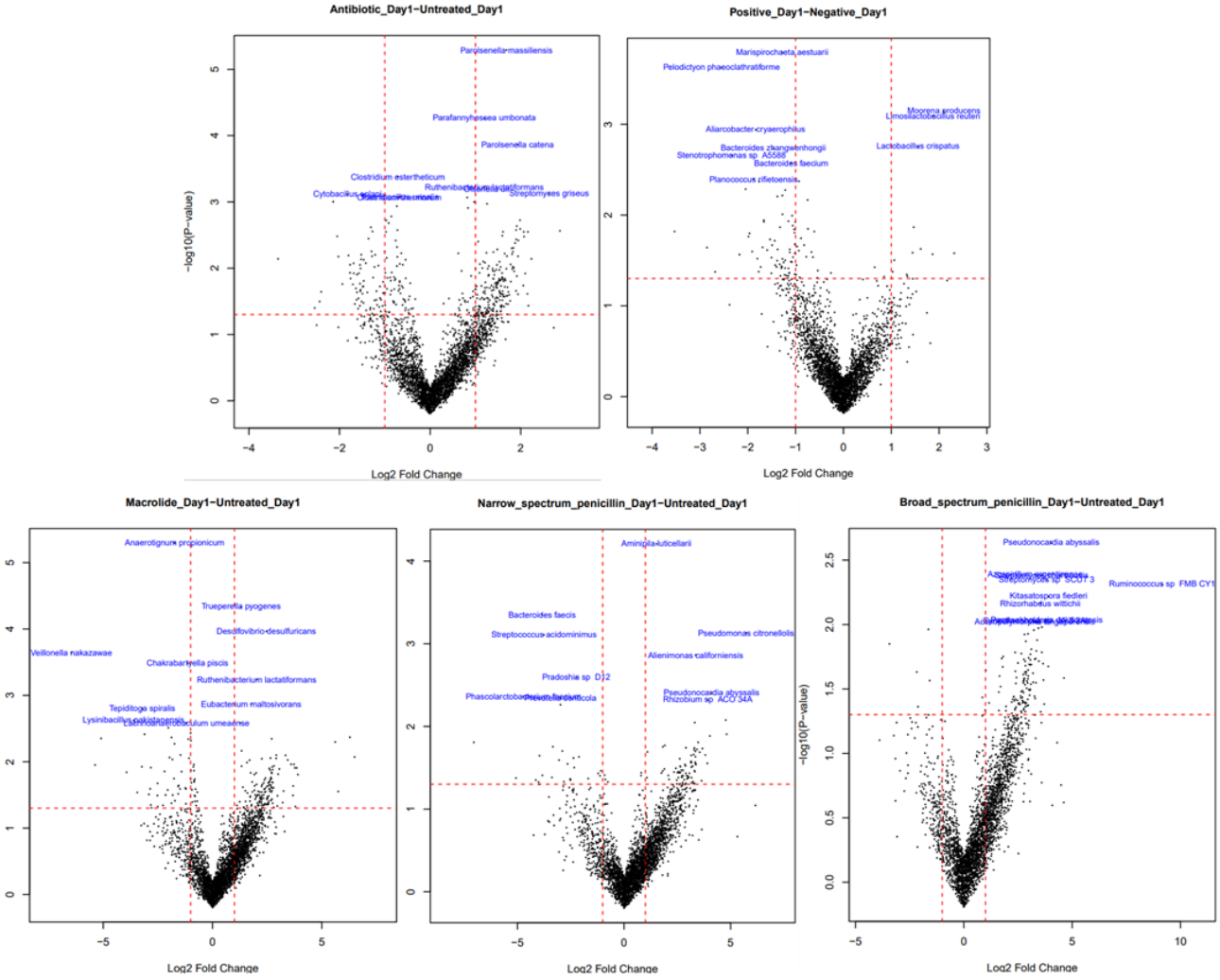


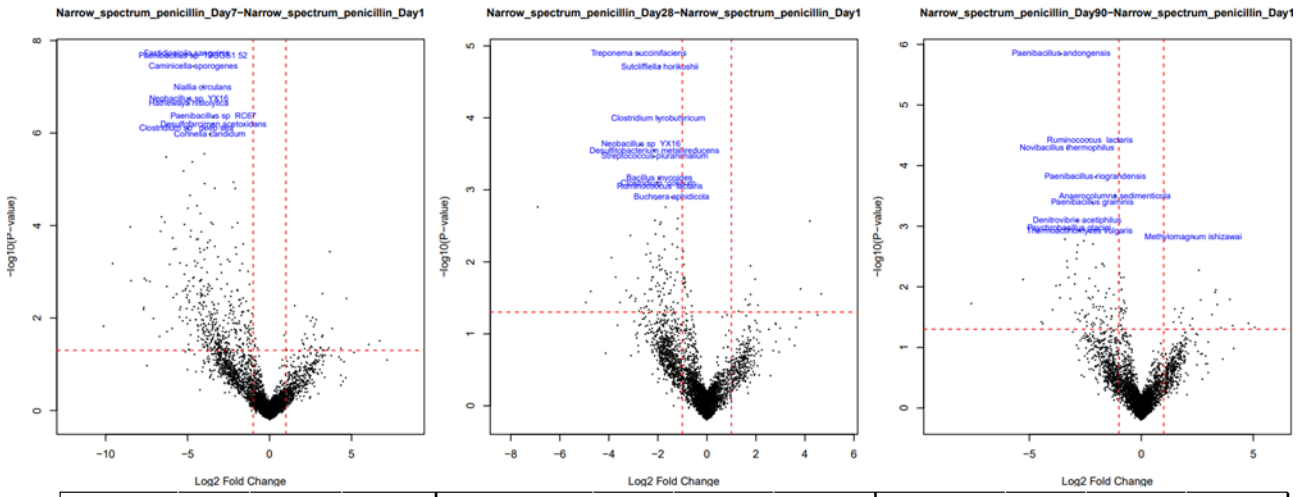
Figure 19: Volcano plot showing taxa with largest differences in the 5 comparisons of interest at baseline. Each point represents a taxon, with log_2(fold-change) on the x-axis and -log_{10}(p-value) on the y-axis. The horizontal and vertical dotted lines represents the (unadjusted) p-value=5% and |log2(fold-change)|<1 thresholds respectively.

	Antibiotic_Day1-Untreated_Day1				Macrolide_Day1-Untreated_Day1			
Category	Taxa name	logFC	AveExpr	adj.P.Val	Taxa name2	logFC3	AveExpr4	adj.P.Val5
Taxa of interest in respiratory related pathologies	Pseudomonas aeruginosa	1,024	7,432	0,411	Pseudomonas aeruginosa	1,454	7,432	0,703
	Streptococcus pneumoniae	-0,67	5,013	0,525	Streptococcus pneumoniae	-0,928	5,013	0,703
	Klebsiella pneumoniae	-1,168	8,411	0,551	Klebsiella pneumoniae	-1,737	8,411	0,703
	Haemophilus influenzae	-0,705	1,089	0,738	Staphylococcus aureus	-0,075	5,475	0,973
	Staphylococcus aureus	-0,099	5,475	0,914	Haemophilus influenzae	-0,07	1,089	0,991
Top 10 differentially abundant taxa	Parolsenella massiliensis	1,684	6,102	0,025	Anaerotignum propionicum	-1,734	5,533	0,019
	Parafannyhessea umbonata	1,197	6,007	0,114	Trueperella pyogenes	1,306	4,35	0,111
	Parolsenella catena	1,93	6,107	0,114	Desulfovibrio desulfuricans	2,46	5,666	0,163
	Streptomyces griseus	2,628	0,63	0,249	Veillonella nakazawae	-6,453	2,771	0,214
	Olsenella uli	1,242	5,124	0,249	Chakrabartyella piscis	-1,142	4,875	0,336
	Paenibacillus urinalis	-0,661	3,798	0,249	Ruthenibacterium lactatiforma	2,021	11,745	0,377
	Clostrm estertheticum	-0,713	5,389	0,249	Lysinibacillus pakistanensis	-3,61	0,843	0,669
	Bactero caccae	-2,139	11,073	0,249	Eubacterium maltosivorans	1,77	7,795	0,669
	Ruthenibacterium lactatiformans	1,195	11,745	0,249	Lachnoanaerobaculum umeae	-1,188	5,757	0,669
	Denitrobacterium detoxificans	0,973	5,007	0,249	Tepoga spiralis	-3,213	0,782	0,674

Table 5: Differential analysis results at baseline in antibiotic and macrolide inter-group comparisons. Shown are the log_2(fold-change), average abundance (log_2CPM), raw and adjusted p-values associated with 5 taxa of interest in respiratory pathologies and the top10 taxa.

3.3.3.4 Intra-group comparisons in patients treated with antibiotics

Unlike inter-group comparisons at baseline that resulted in only 2 significant taxa, intra-group comparisons led to 67, 2, and 3 species showing significant changes at D7, D28, and D90 respectively in the narrow-spectrum penicillin comparison—all exhibiting decreased abundance relative to baseline, as illustrated by the leftward shift in the volcano plots (Figure 20). These significant results should however be interpreted with caution given that D7, the time point with the largest number of significant taxa, consists of 2 samples only (see Table 4). Finally, no significant changes were detected in the 2 other classes or when taking the classes altogether.



D7-D1				D28-D1				D90-D1			
Taxa name	logFC	AveExpr	adj.P.Val	Taxa name	logFC	AveExpr	adj.P.Val	Taxa name	logFC	AveExpr	adj.P.Val
Haemophilus influenzae	-5,473	1,089	0,127	Pseudomonas aeruginosa	-1,112	7,432	1	Staphylococcus aureus	-0,189	5,475	1
Staphylococcus aureus	-0,614	5,475	0,955	Klebsiella pneumoniae	0,702	8,411	1	Pseudomonas aeruginosa	-0,203	7,432	1
Pseudomonas aeruginosa	-0,511	7,432	0,999	Streptococcus pneumoniae	-0,176	5,013	1	Haemophilus influenzae	0,254	1,089	1
Streptococcus pneumoniae	-0,262	5,013	0,999	Haemophilus influenzae	0,178	1,089	1	Streptococcus pneumoniae	-0,095	5,013	1
Klebsiella pneumoniae	0,355	8,411	0,999	Staphylococcus aureus	0,021	5,475	1	Klebsiella pneumoniae	-0,174	8,411	1
Paenibacillus sp 19GGS152	-4,666	2,567	0	Treponema succinifaciens	-2,779	3,406	0,042	Paenibacillus andongensis	-3,576	2,25	0,004
Fastsipila sanguinis	-5,048	3,385	0	Sutcliffiella horikoshii	-1,919	3,207	0,042	Ruminococcus lactaris	-2,347	9,931	0,039
Caminicella sporogenes	-4,655	3,25	0	Clostrm tyrobutyricum	-1,983	3,544	0,152	Novibacillus thermophilus	-3,318	1,674	0,039
Niallia circulans	-4,081	3,245	0	Neobacillus sp YX16	-2,693	2,56	0,152	Paenibacillus riograndensis	-2,055	3,173	0,198
Neobacillus sp YX16	-4,924	2,56	0	Desulfitobacterium metallireducens	-2,142	2,994	0,152	Paenibacillus graminis	-2,185	2,787	0,279
Hathewaya histolytica	-4,929	2,958	0	Streptococcus pluranimalium	-2,13	3,201	0,152	Anaerocolumna sedimenticola	-1,185	6,231	0,279
Desulfofarcimen acetoxs	-3,421	2,954	0	Bacillus myco	-1,94	3,237	0,412	Denitrovibrio acetiphilus	-2,867	1,984	0,281
Paenibacillus sp RC67	-3,429	2,823	0	Ruminococcus lactaris	-1,964	9,931	0,412	Methylomagnum ishizawai	2,319	1,918	0,469
Clostrm sp deep sea	-5,007	2,436	0	Clostrm colinum	-2,036	3,87	0,412	Staphylococcus epmidis	-2,568	3,086	0,469
Cohnella cand	-3,649	3,345	0	Streptococcus sp FDAARGOS_192	-6,904	4,105	0,542	Thermoactinomyces vulgaris	-2,757	1,872	0,469

Figure 20: Narrow spectrum ampicillin intra-group comparisons. Volcano plots (top); result table associated with taxa of interest in respiratory related pathologies and top10 differentially abundant taxa (bottom).

3.3.3.5 Intra-group comparisons in untreated patients

We next examined whether differences exist between baseline and subsequent time points (D7, D28, and D90) in untreated patients. Only a small fraction of taxa had an effect size greater than $|\log_2(\text{fold-change})| = 1$, and the lowest adjusted p-value was 0.52, indicating that time had very little impact on the microbiota of untreated patients (Supplementary Table 1, Figure 1). Consequently, inter-group comparisons with untreated patients effectively reduce to intra-group comparisons. For instance, the inter-group comparisons between antibiotic and untreated patients at D7 is expressed as: (ATB_D7-ATB_D1)-(Untr_D7-Untr_D1). Since the second term is nearly zero, this inter-group comparison is reduced to the antibiotic intra-group comparison. This point is supported

by the scatterplot of intra versus inter-group log2(fold-change) estimates which shows a high correlation in the antibiotic versus untreated comparison (Figure 21). For this reason, inter-group comparisons based on untreated patients were excluded from further analysis.

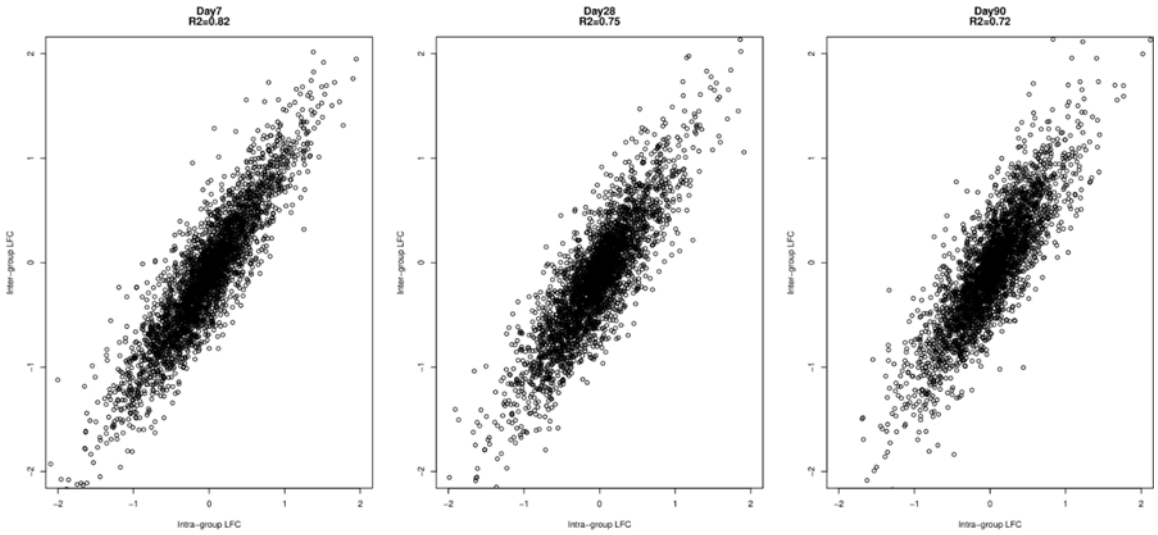


Figure 21: Scatterplot of log2(fold-change) estimated in antibiotic intra (x-axis) versus inter (y-axis) group comparisons. The obtained corrections motivated us to exclude inter-group comparisons.

3.3.3.6 Intra and inter-group comparisons in COVID-19 patients

So far, COVID-19 infection has been regarded as a confounding factor that could potentially hinder the estimation of antibiotic treatment effects. In this paragraph, we shift the focus from antibiotic to COVID-19 infection, treating COVID-19 as the primary clinical variable of interest, while considering antibiotic intake as confounding factor. Among the 102 patients enrolled in Prudence, 18 tested positive for COVID-19. Both intra- and inter-group comparisons revealed no taxa that were significantly differentially abundant over time or between COVID-19 status groups (Supplementary Figure 3 and Table 3). Again, this result is in line with the variance partitioning analysis where COVID-19 (POS_TEST) has the smallest median explained variance (Figure 14).

3.3.3.7 Longitudinal analysis

Two approaches were employed to identify bacteria showing significant changes over time within clinical groups, thereby offering insights into bacterial dynamics. The first approach (DESeq2 LRT) targets taxa with trajectories that differed between groups (e.g., antibiotic-treated versus untreated), while the second (Limma pairwise TPs) focuses on taxa with significant changes from baseline at one or more time points.

In line with the intra-group comparisons, Table 6 shows that most of the significant taxa are found in narrow spectrum penicillin using Limma pairwise TPs. On the other hand, only a handful of taxa were identified with DESeq2 LRT.

Clinical comparison		Limma Pairwise TPs	DESeq2 LRT
Antibiotic vs untreated		1	0
ATB Classes	Narrow spectrum penicillin	78	1
	Broad spectrum penicillin	0	1
	Macrolide	1	3
COVID19 positive vs negative		0	7

Table 6: Number of significant taxa found in longitudinal analysis using limma pairwise TPs and DESeq2 LRT in the 5 clinical comparisons of interest.

For a given comparison, no overlap was observed between the two approaches, indicating that each method identified distinct patterns. Gaussian mixture clustering was applied to the 78 taxa identified in narrow-spectrum penicillin using Limma, resulting in a single cluster. This cluster shows a decreased abundance at D1 followed by a return to baseline by D28 and D90. Examining the first 10 taxa alphabetically (Figure 22) reveals minimal variation in other groups, suggesting these taxa are specifically sensitive to narrow-spectrum penicillin (full list available in Supplementary Figure 10). As noted in section 3.3.3.4, these findings should however be interpreted with caution and further validated given that they are based primarily on two samples at D7.

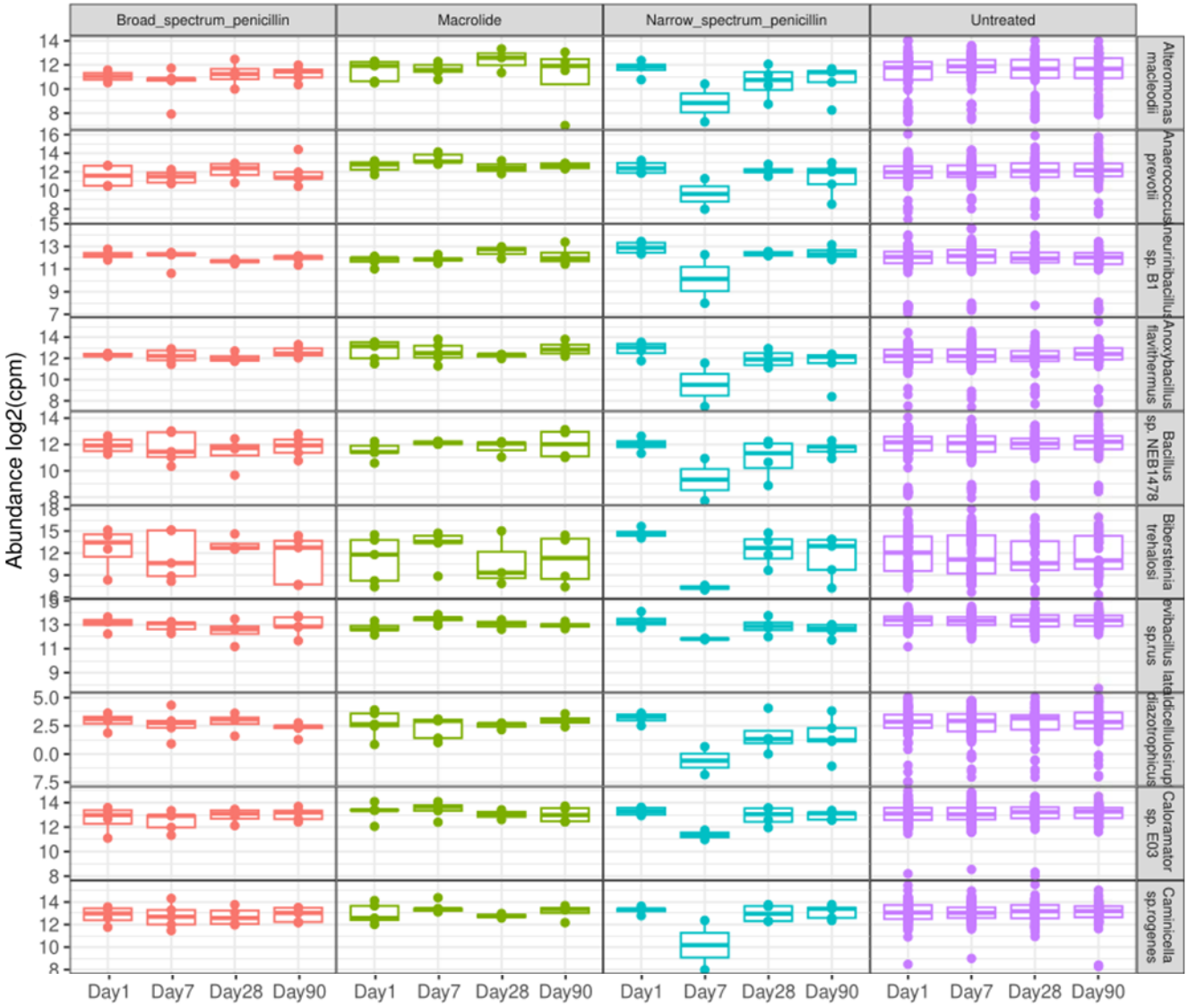


Figure 22: Trajectories of the first 10 bacteria (in alphabetical order) identified as significant in narrow spectrum penicillin with Limma Pairwise TP.

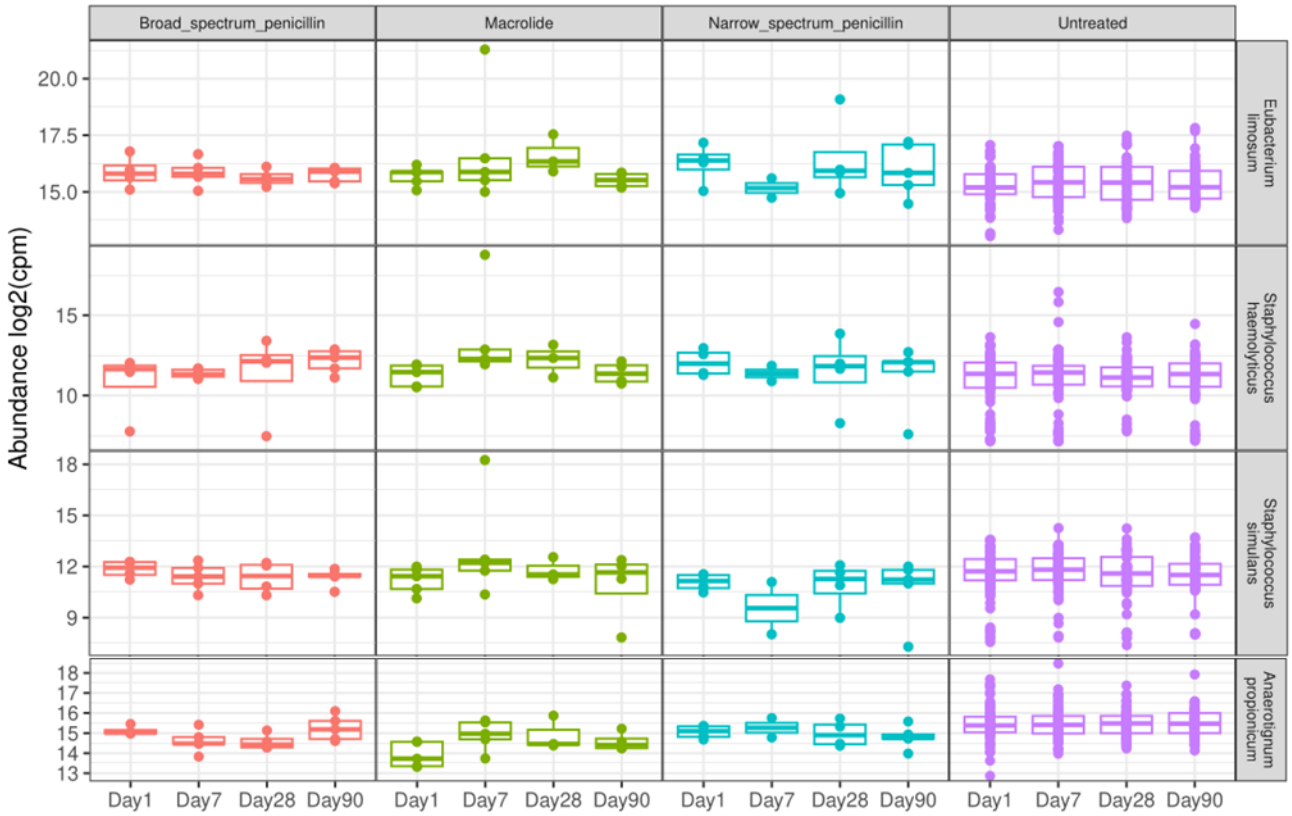


Figure 23: Trajectories of the 4 bacteria identified as significant in macrolide DESeq2 LRT (top 3 taxa) and Limma Pairwise TP (bottom taxon).

Looking at the 4 taxa found in macrolide group (Figure 23), an opposite trend can be observed where a peak occurs at D1-D28 and a return to baseline at D90. This pattern was found both by DESeq2 LRT (top 3 taxa) and Limma (bottom taxon).

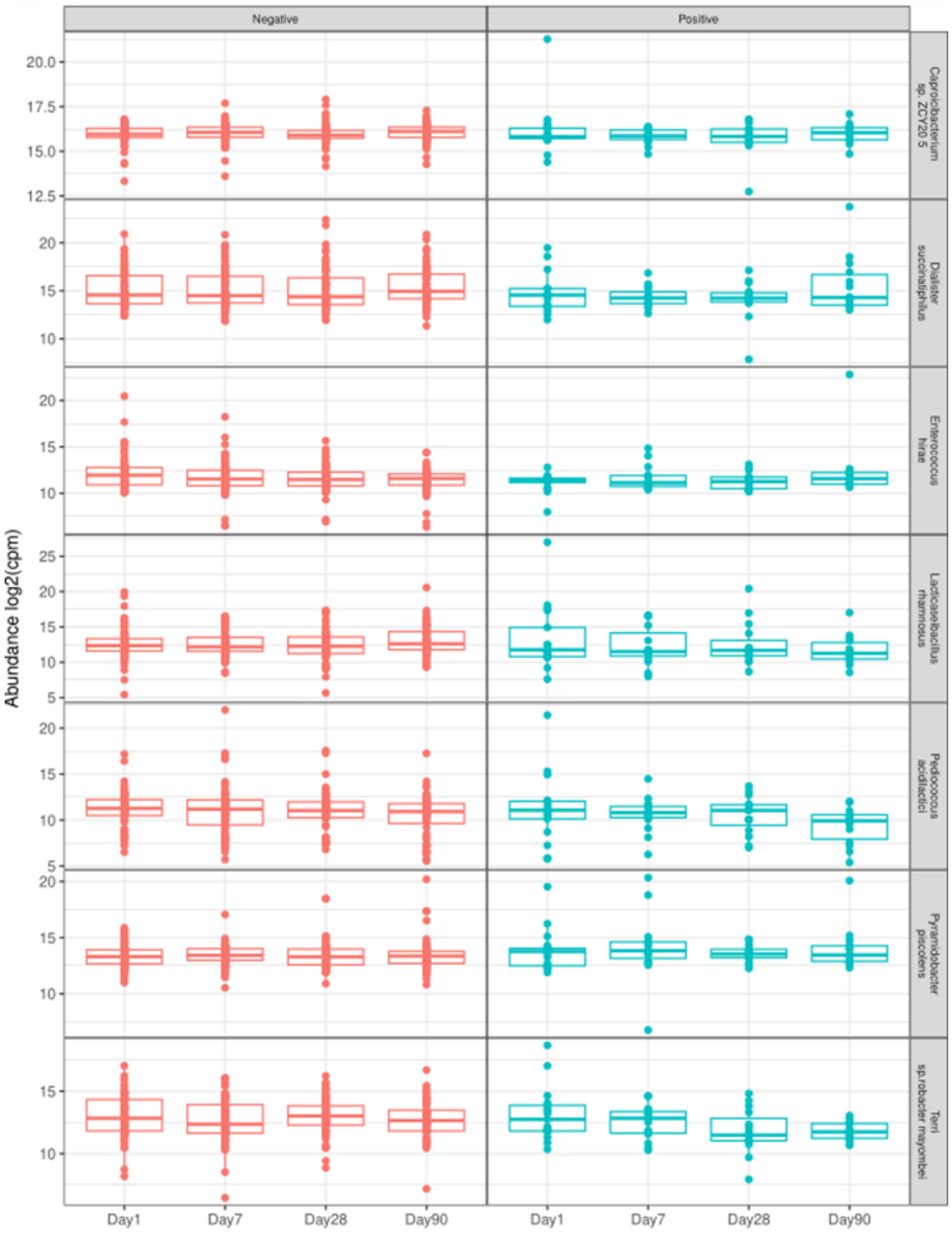


Figure 24: Trajectories of the 7 bacteria identified as significant in COVID-19 comparison using DESeq2 LRT.

Finally, in the COVID-19 comparison, beside a downward trend observed in 3 taxa, little differences can be noticed between positive and negative patients (Figure 24). This result is not unexpected since COVID-19 (POS_TEST) explained little variation in Figure 14 and no taxa were found differentially abundant in the COVID-19 comparisons.

3.3.3.8 Additive log-ratio normalization

In this section we evaluate the impact of the normalization method (TMM versus ALR) on differential analysis results. Table 7 shows the number of significant taxa identified in the clinical intra-group comparisons of interest. Despite a moderate correlation found between the two methods (R^2=0.6, Figure 13), the number of significant taxa is highly similar. Few, if any, significant taxa were found in the antibiotic vs untreated and COVID-19 comparisons, while most differentially abundant taxa were identified in the narrow spectrum penicillin group.

Class	Days	Taxa (TMM)	Taxa (ALR)
Antibiotic	Day7-Day1	0	0
	Day28-Day1	0	0
	Day90-Day1	0	0
	Day28-Day7	0	1
	Day90-Day7	0	0
	Day90-Day28	0	1
Untreated	Day7-Day1	0	0
	Day28-Day1	0	0
	Day90-Day1	0	0
	Day28-Day7	0	0
	Day90-Day7	0	0
	Day90-Day28	0	0

Class	Days	Taxa (TMM)	Taxa (ALR)
Covid19 Positive	Day7-Day1	0	0
	Day28-Day1	0	0
	Day90-Day1	0	0
	Day28-Day7	0	0
	Day90-Day7	0	0
	Day90-Day28	0	0
Covid19 Negative	Day7-Day1	0	0
	Day28-Day1	0	0
	Day90-Day1	0	0
	Day28-Day7	0	0
	Day90-Day7	0	0
	Day90-Day28	0	0

Class	Days	Taxa (TMM)	Taxa (ALR)
Narrow spectrum penicillin	Day7-Day1	67	34
	Day28-Day1	2	0
	Day90-Day1	3	1
	Day28-Day7	13	24
	Day90-Day7	32	26
	Day90-Day28	0	1
Broad spectrum penicillin	Day7-Day1	0	0
	Day28-Day1	0	0
	Day90-Day1	0	0
	Day28-Day7	0	0
	Day90-Day7	0	1
	Day90-Day28	0	0
Macrolide	Day7-Day1	0	0
	Day28-Day1	0	0
	Day90-Day1	0	0
	Day28-Day7	0	0
	Day90-Day7	0	0
	Day90-Day28	0	0

Table 7: Number of significant taxa found with TMM and ALR normalizations in the clinical comparisons of interest.

We then sought to determine whether the lists of significant taxa were similar across methods. To do so, Venn Diagrams were generated for intra-group comparisons, where significant taxa were identified in both methods. Figure 25 shows that across all comparisons, the intersection is large, with 3 out of 4 cases showing that most, if not all ALR results are encompassed within the TMM taxa. Depending on the reader's interest into TMM or ALR, differential analysis results are available for both methods in supplementary materials.

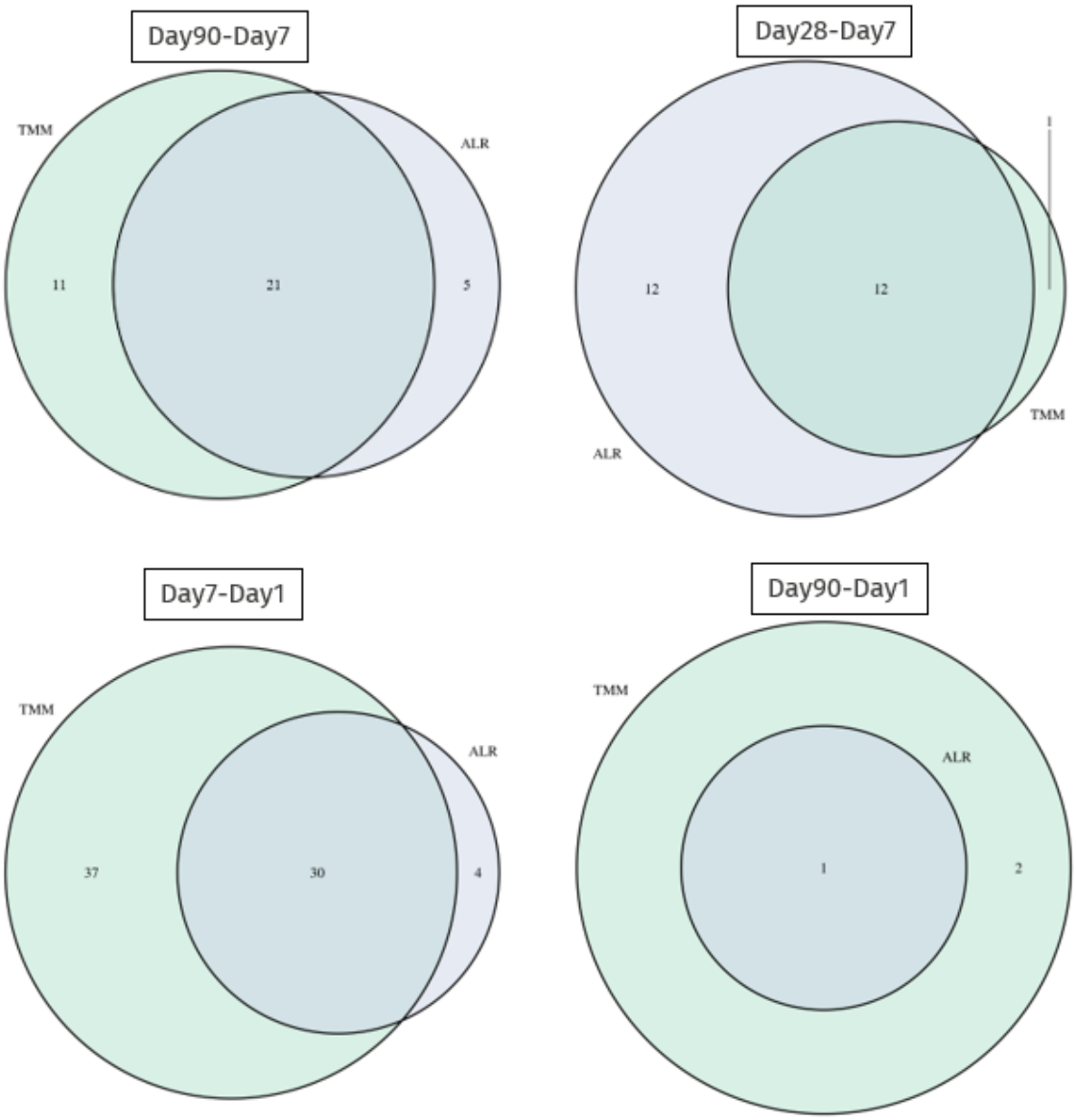


Figure 25: Intersections between taxa identified as significant in narrow spectrum penicillin intra-group comparisons.

3.3.4 Differential and longitudinal analyses in resistance genes

In this section, we repeat the previous differential and longitudinal analyses at the resistance gene level. The full tables and associated Volcano plots are available in Supplementary Tables 7-9 and Figures 7-9.

3.3.4.1 Background noise determination

The background noise threshold was adjusted to ensure a monotonically decreasing mean-variance trend. As previously observed at the taxa level, Figure 26 shows that the mean-variance trend resembles a logarithmic curve, with many genes at low abundance. Setting a background noise threshold of log2(cpm) > 2.5 resulted in the removal of 1,205

out of 1,331 genes. This stringent filtering reflects the matrix's high sparsity, with approximately 90% of entries being zeros, as illustrated in the histogram below.

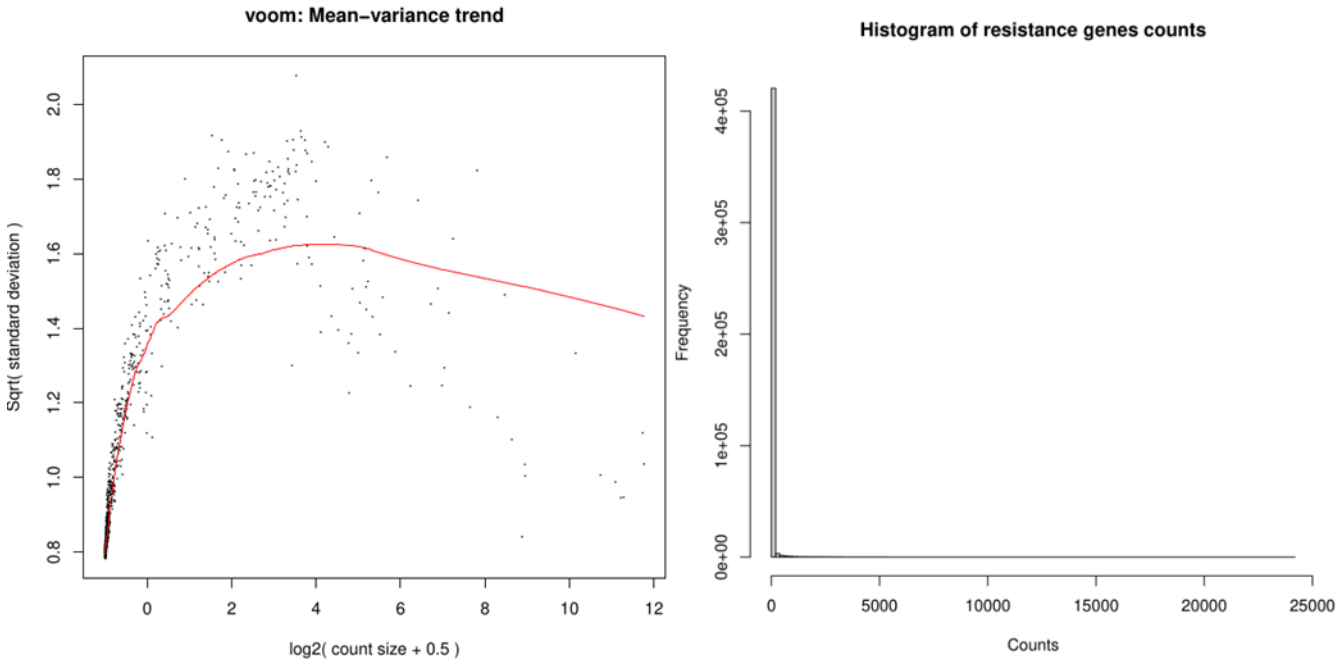


Figure 26: Mean-variance scatterplot (left). All taxa with mean log2(cpm) <2.5 were excluded from differential analyses. Histogram of raw counts in the resistance gene matrix (right).

3.3.4.2 Comparison at baseline

Consistent with the results at the taxa level, differential analyses at baseline across the five comparisons (Figure 27) identified only one significant resistance gene, *vanY* in *vanF*, in macrolide (top 10 genes are shown in Table 8). Despite weak effects at baseline in all comparisons, it can further be noted that the range of $\log_2$(fold-change) is notably smaller in the COVID-19 than the other 4 comparisons.

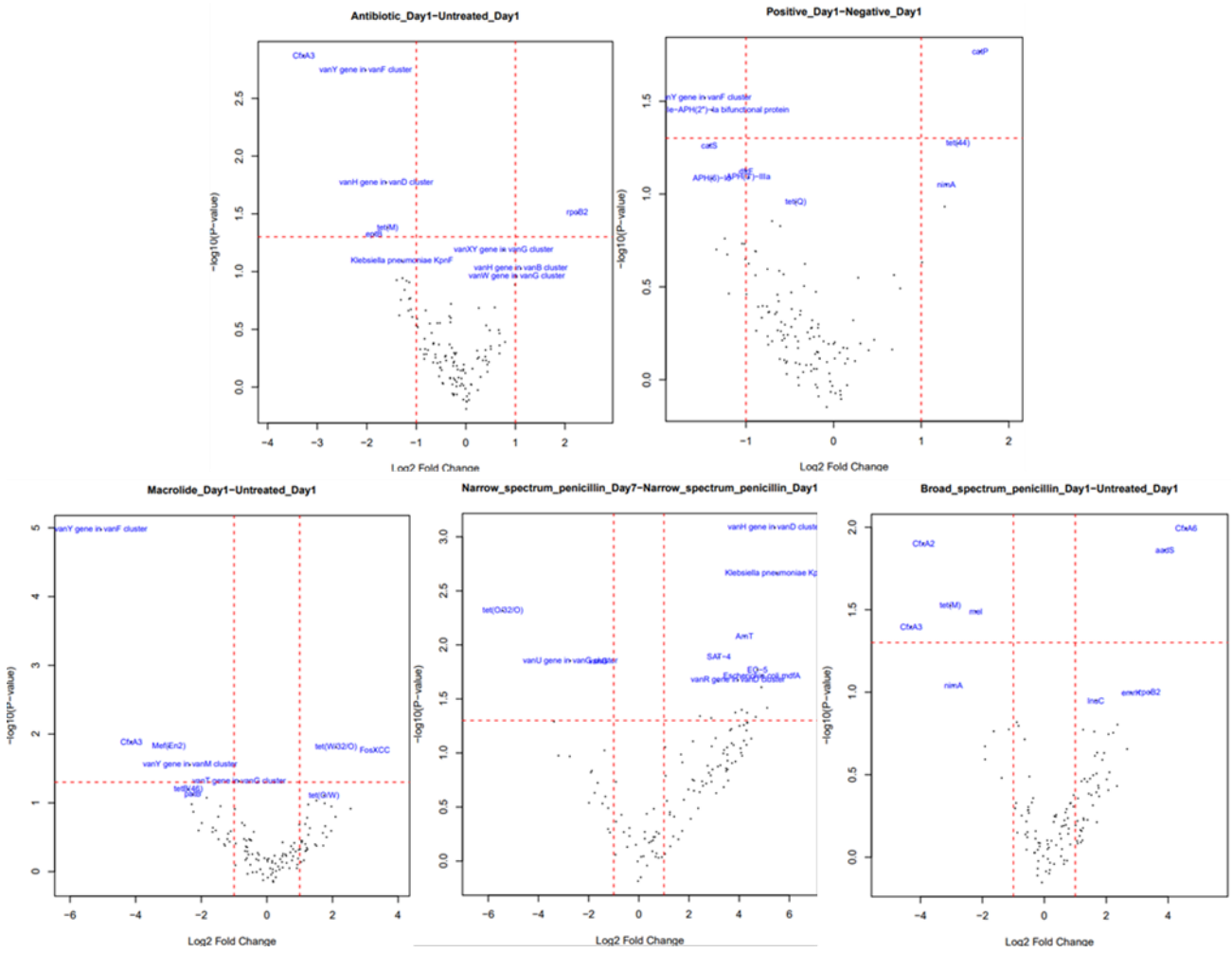


Figure 27: Volcano plot showing taxa with largest differences in the 5 comparisons of interest at baseline. Each point represents a resistant gene, with log_2(fold-change) on the x-axis and -log_{10}(p-value) on the y-axis. The horizontal and vertical dotted lines represents the (unadjusted) p-value=5% and |log2(fold-change)|<1 thresholds respectively.

Macrolide_Day1-Untreated_Day1			
Gene name	logFC	AveExpr	adj.P.Val
vanY gene in vanF	-5,055	9,834	0,001
Mef(En2)	-2,99	11,211	0,512
FosXCC	3,292	10,106	0,512
CfxA3	-4,129	8,326	0,512
tet(W/32/O)	2,108	9,566	0,514
vanY gene in vanM	-2,341	9,163	0,514
tetA(46)	-2,293	7,433	0,848
SAT-4	1,761	9,988	0,848
vanT gene in vanG	-0,845	15,884	0,848
vanY gene in vanB	-0,979	13,74	0,848

Table 8: Top resistant genes found at baseline in macrolide versus untreated comparison. Shown are the log_2(fold-change), average abundance (log_2CPM) and adjusted p-values.

3.3.4.3 Intra-group comparisons

In contrast to the numerous significant taxa found in intra-group comparisons, only 5 resistant genes reached significance (Table 9, Table 10) in narrow spectrum penicillin (vanG in D90-D1, vanY gene in vanG cluster, vanR gene in vanD cluster in D28-D7) and macrolide (vanY gene in vanF cluster in D90-D1 and D90-D7, APH(6)-Id in D90-D7) intra-group comparisons.

Class	Days	Resistant Genes (TMM)
Antibiotic	Day7-Day1	0
	Day28-Day1	0
	Day90-Day1	0
	Day28-Day7	0
	Day90-Day7	0
	Day90-Day28	0
Untreated	Day7-Day1	0
	Day28-Day1	0
	Day90-Day1	0
	Day28-Day7	0
	Day90-Day7	0
	Day90-Day28	0

Class	Days	Resistant Genes (TMM)
Narrow spectrum penicillin	Day7-Day1	0
	Day28-Day1	0
	Day90-Day1	1
	Day28-Day7	2
	Day90-Day7	0
	Day90-Day28	0
Broad spectrum penicillin	Day7-Day1	0
	Day28-Day1	0
	Day90-Day1	0
	Day28-Day7	0
	Day90-Day7	0
	Day90-Day28	0
Macrolide	Day7-Day1	0
	Day28-Day1	0
	Day90-Day1	1
	Day28-Day7	0
	Day90-Day7	2
	Day90-Day28	0

Class	Days	Resistant Genes (TMM)
Covid19 Positive	Day7-Day1	0
	Day28-Day1	0
	Day90-Day1	0
	Day28-Day7	0
	Day90-Day7	0
	Day90-Day28	0
Covid19 Negative	Day7-Day1	0
	Day28-Day1	0
	Day90-Day1	0
	Day28-Day7	0
	Day90-Day7	0
	Day90-Day28	0

Table 9: Number of genes found significant in intra-group comparison across the 5 clinical comparisons of interest.

Narrow spectrum penicillin versus untreated							
Day90-Day1				Day28-Day7			
Colonne1	logFC	AveExpr	adj.P.Val	Colonne1	logFC	AveExpr	adj.P.Val
vanG	-1,923	13,750	0,009	vanY gene in vanG cluster	-4,580	9,633	0,032
vanY gene in vanB cluster	-1,525	13,740	0,175	vanR gene in vanD cluster	-5,808	8,585	0,032
tet(O/32/O)	-3,973	7,846	0,175	vanH gene in vanD cluster	-4,596	8,354	0,112
SAT-4	2,431	9,988	0,219	AmT	-4,148	10,027	0,212
aad(6)	2,120	10,313	0,313	Klebsiella pneumoniae Kpr	-4,343	9,230	0,212
vanW gene in vanI cluster	-1,392	13,432	0,313	mdtO	-5,085	7,692	0,212
eptA	3,631	9,085	0,313	poxtA	5,150	8,120	0,212
ArnT	2,479	10,027	0,313	AcrF	-5,292	8,536	0,212
poxtA	-3,426	8,120	0,313	mdtF	-5,203	8,252	0,212
EC-5	3,041	7,810	0,313	APH(6)-Id	-4,078	6,984	0,212
Macrolide versus untreated							
Day90-Day1				Day90-Day7			
Colonne1	logFC	AveExpr	adj.P.Val	Colonne1	logFC	AveExpr	adj.P.Val
vanY gene in vanF cluster	4,533	9,834	0,002	vanY gene in vanF cluster	3,695	9,834	0,042
ErmX	4,494	6,479	0,101	APH(6)-Id	4,556	6,984	0,042
eptA	-3,878	9,085	0,499	lnuC	-2,601	10,377	0,190
mdtF	-3,856	8,252	0,499	ErmX	3,836	6,479	0,218
lnuC	-2,008	10,377	0,499	catP	2,767	7,706	0,330
vanY gene in vanM cluster	2,076	9,163	0,499	vanH gene in vanB cluster	-2,644	9,916	0,330
APH(6)-Id	2,870	6,984	0,499	Mef(En2)	-2,580	11,211	0,330
gadW	-3,135	7,206	0,499	APH(3")-Ib	3,453	7,712	0,330
vanY gene in vanG cluster	-1,979	9,633	0,499	vanY gene in vanG cluster	-1,947	9,633	0,543
evgA	-2,717	6,954	0,499	FosXCC	-2,469	10,106	0,543

Table 10: Differential analysis results in narrow spectrum penicillin and macrolide intra-group comparisons. Shown are the log_2(fold-change), average abundance (log_2CPM), raw and adjusted p-values associated with the top10 genes.

3.3.4.4 Intra versus inter group comparisons

Intra- and inter-group log2(fold-change) values were then compared in the antibiotic versus untreated comparison. As noted earlier on taxa, the two types of comparisons are highly similar at the gene level (Figure 28) due to little variation in untreated patients along the time. For this reason, inter-group comparisons based on untreated patients were again excluded from further analysis.

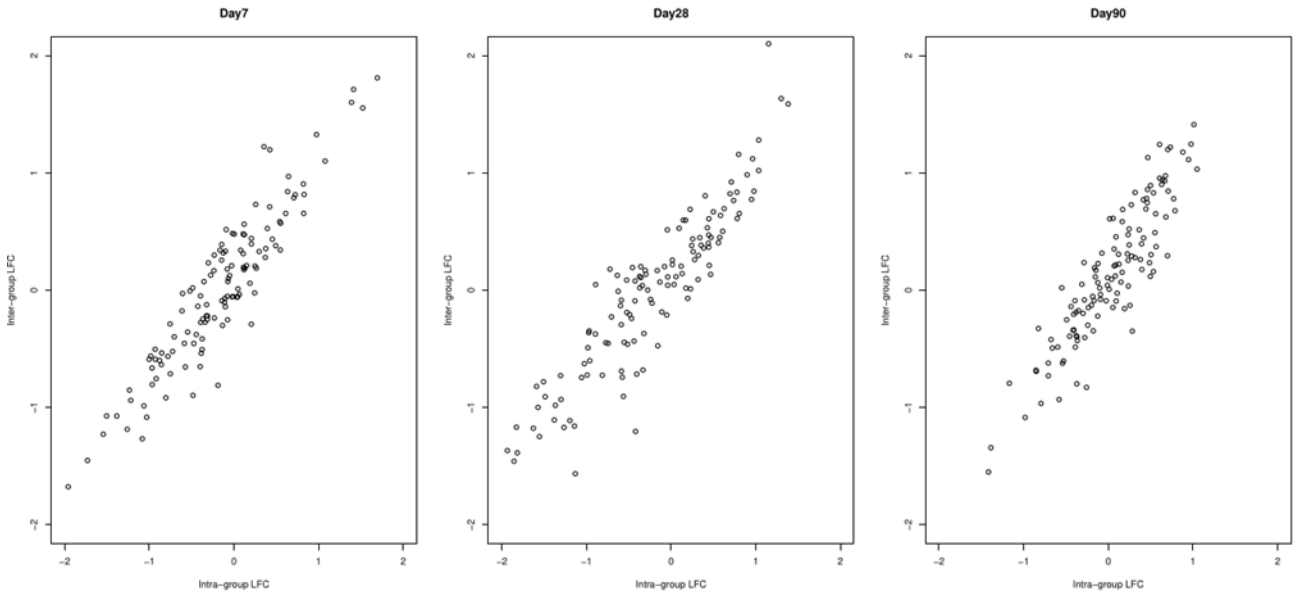


Figure 28: Scatterplot of log2(fold-change) estimated in antibiotic intra (x-axis) versus inter (y-axis) group comparisons. The obtained corrections motivated us to exclude inter-group comparisons.

3.3.4.5 Longitudinal analysis

Reflecting the intra-group comparisons, Table 11shows that the 5 significant genes are found in narrow spectrum penicillin and macrolide using Limma pairwise TPs. No additional genes are uncovered by the DESeq2 LRT approach.

Clinical comparison		Limma Pairwise TPs	DESeq2 LRT
Antibiotic vs untreated		0	0
ATB Classes	Narrow spectrum penicillin	3	0
	Broad spectrum penicillin	0	0
	Macrolide	2	0
COVID19 positive vs negative		0	0

Table 11: Number of significant genes found in longitudinal analysis using limma pairwise TPs and DESeq2 LRT in the 5 clinical comparisons of interest.

Unlike what was observed at the taxa level, these significant genes show large variations along the time in all 3 antibiotics (Figure 29). Only untreated patients display little variation. The 3 significant genes identified in narrow-spectrum penicillin exhibit distinct patterns: *vanG* decreases steadily over time, while *vanY* and *vanR* peak at D7. In contrast, *vanY* and *APH(6)-Id* in the macrolide treatment show a consistent upward trend over time. This pattern could be consistent with the development of a resistant strain.

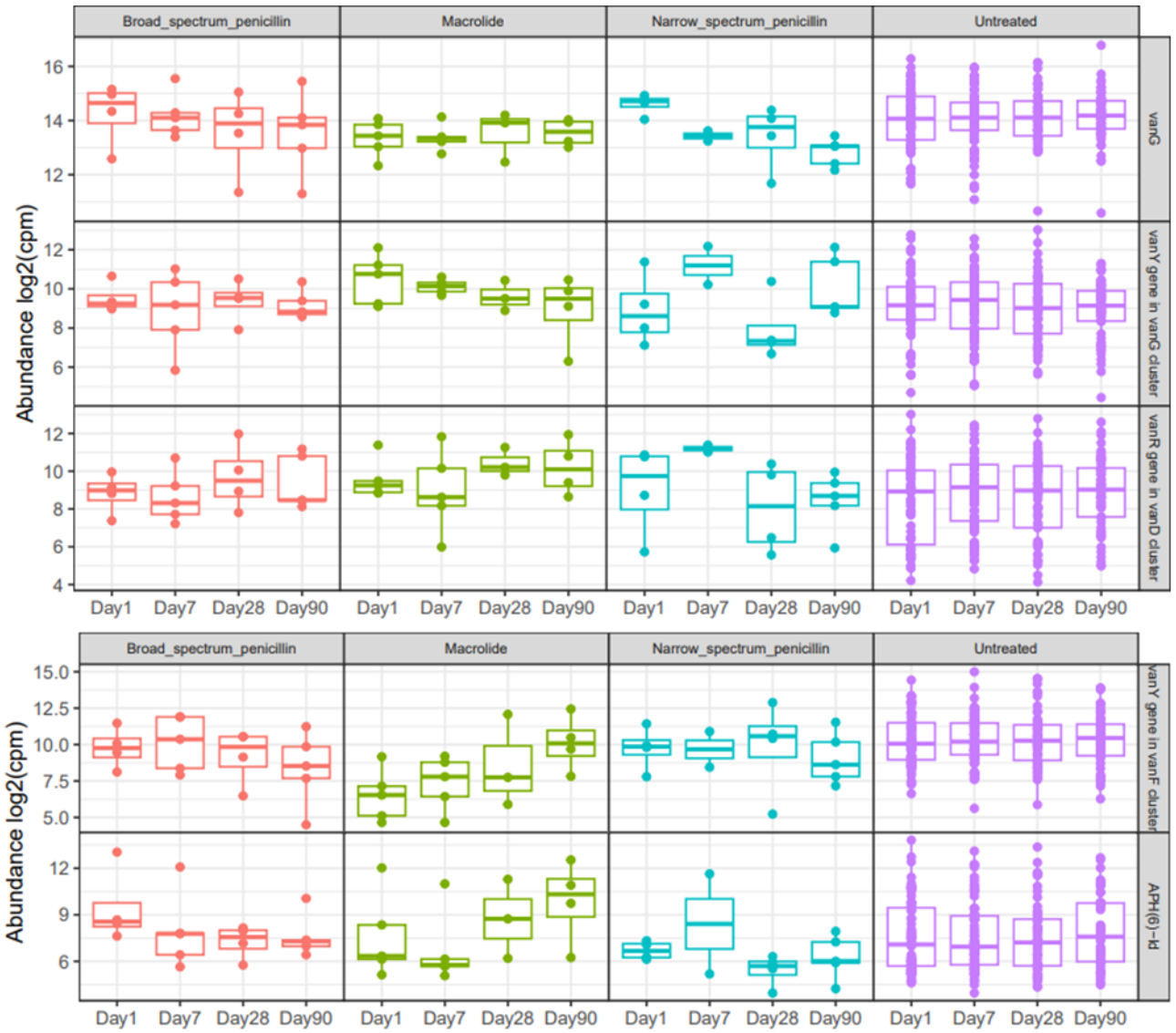


Figure 29: Trajectories of the 5 genes significant identified in narrow spectrum penicillin (top) and macrolide (bottom) with Limma Pairwise TP.

3.4 ADEQUATE: Respiratory microbiota

3.4.1 Evaluation of data quality

3.4.1.1 Sample distribution

A total of 86 oropharyngeal samples were collected from 49 unique patients across 5 sites (SP-P-01, GR-P-01, GE-P-03, SW-P-01, UK-P-01) at 2 time points: Day 0 and Day 30 (denoted as D0 and D30 hereafter). On average, 1.8 time points were sampled per patient, with 12 patients missing the Day 30 sample. A maximum of 22 samples could be extracted simultaneously, resulting in 4 extraction batches, which were then pooled into 2 sequencing runs. To detect and remove potential contaminants, a negative control was included in each extraction batch.

The 86 samples were randomized in a way that confounding variables were uniformly distributed across extraction batches, making thus their correction possible in downstream analyses. The workflow used to select clinical covariates for randomization led to the selection of site, sex, age, fever, obesity (BMI), visit (time point) and antibiotic intake. The last two variables are the main variable of interest, the other are confounders that were adjusted for. Figure 30 illustrates that the distribution of these 7 clinical variables is comparable to that of the full dataset, ensuring thus that both technical and clinical effects can be accurately estimated. Table 12summarises the sample distribution across time points and treatments.

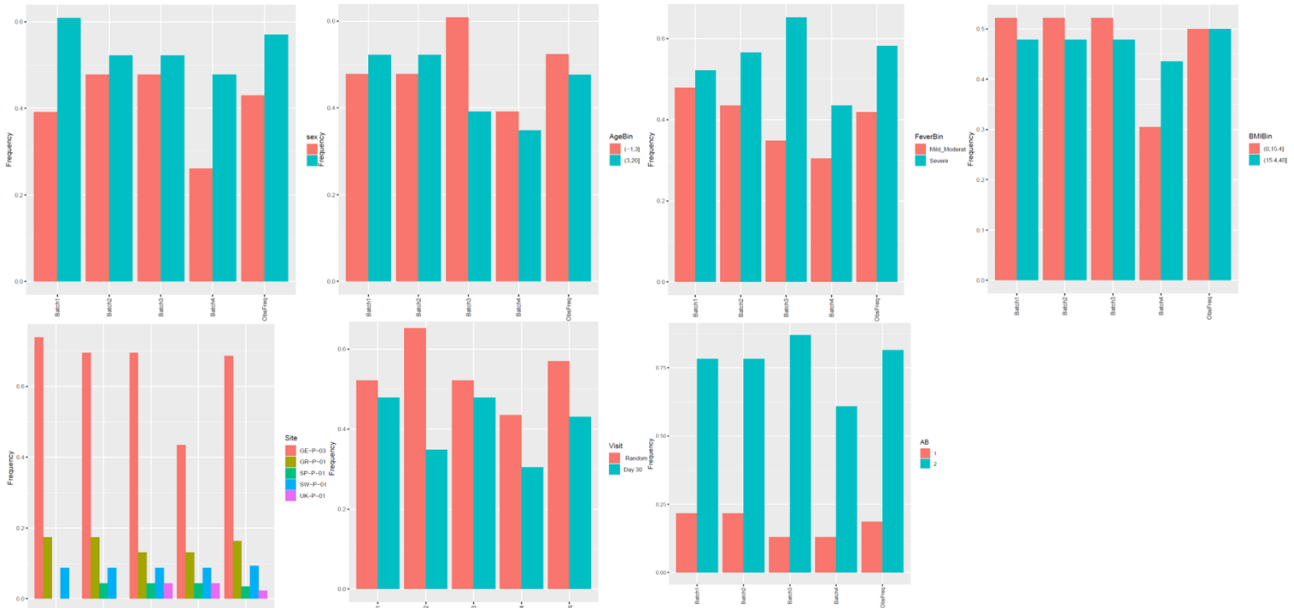


Figure 30: Sample randomization into extraction batches. Each barplot shows, for a given variable, the sample distribution in the 4 batches and the full dataset (observed frequencies, the rightmost category).

	Antibiotic	Untreated
D0	14	35
D30	12	25

Table 12: Sample distribution across time points and treatments.

3.4.1.2 Sequencing quality control

Sequencing produced high-quality reads in each sample, with a Phred score of Q30, indicating excellent base call accuracy. A Q30 score corresponds to 99.9% accuracy, meaning there is only a 0.1% chance of an incorrect base call.

However, the number of sequencing reads generated varied across samples, with some containing fewer than 1 million reads. Additionally, the proportion of human reads was highly variable, with some samples containing up to 98% human reads, while others were composed entirely of microbial reads (Figure 31). The presence of human reads is expected in swab samples, but 36 samples had fewer than 1 million microbial reads (Supplementary Table 10), limiting the amount of microbial information available for analysis and comparison with other samples. These samples were flagged as cautionary due to their limited data.

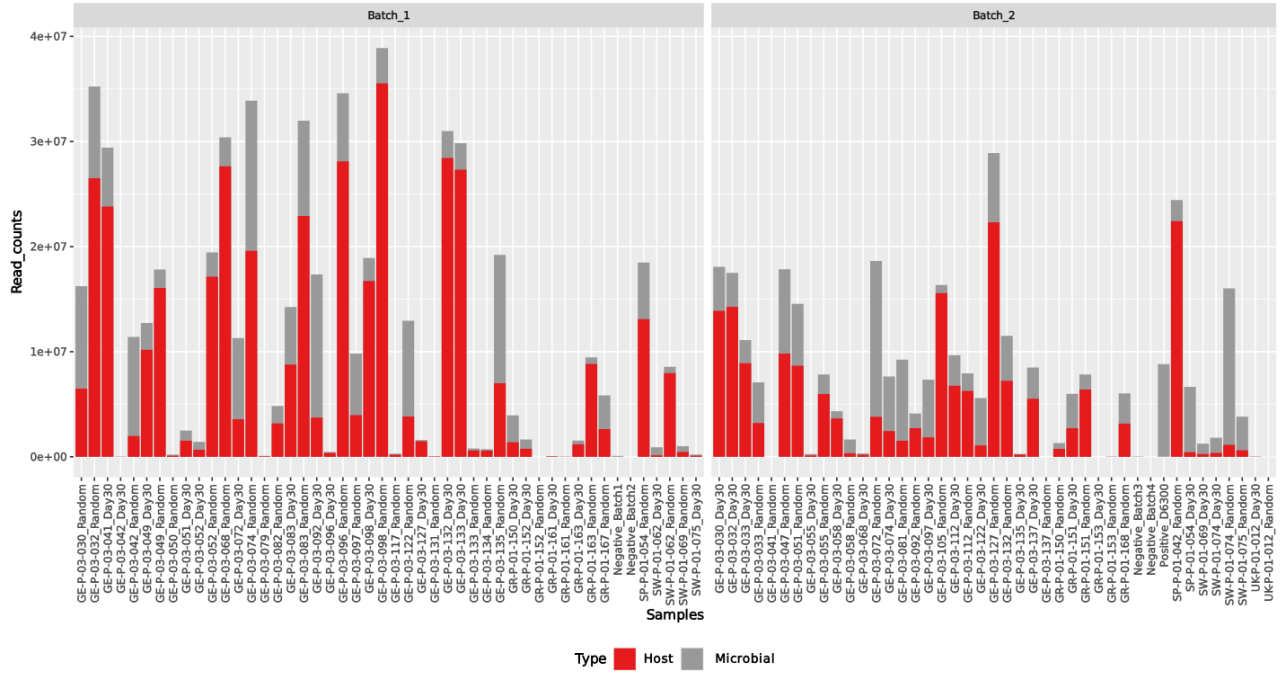


Figure 31: Number of generated Host and microbial reads per sample. Samples are grouped by batch.

Due to the low sequencing depth and/or high levels of host contamination in many samples, the rarefaction curves did not reach a plateau (Figure 32). This indicates that,

with the current sequencing effort, not all species present in the samples—particularly those of low abundance—could be detected. However, it is still feasible to focus on the most abundant bacteria, which are well represented in the data, to draw meaningful insights.

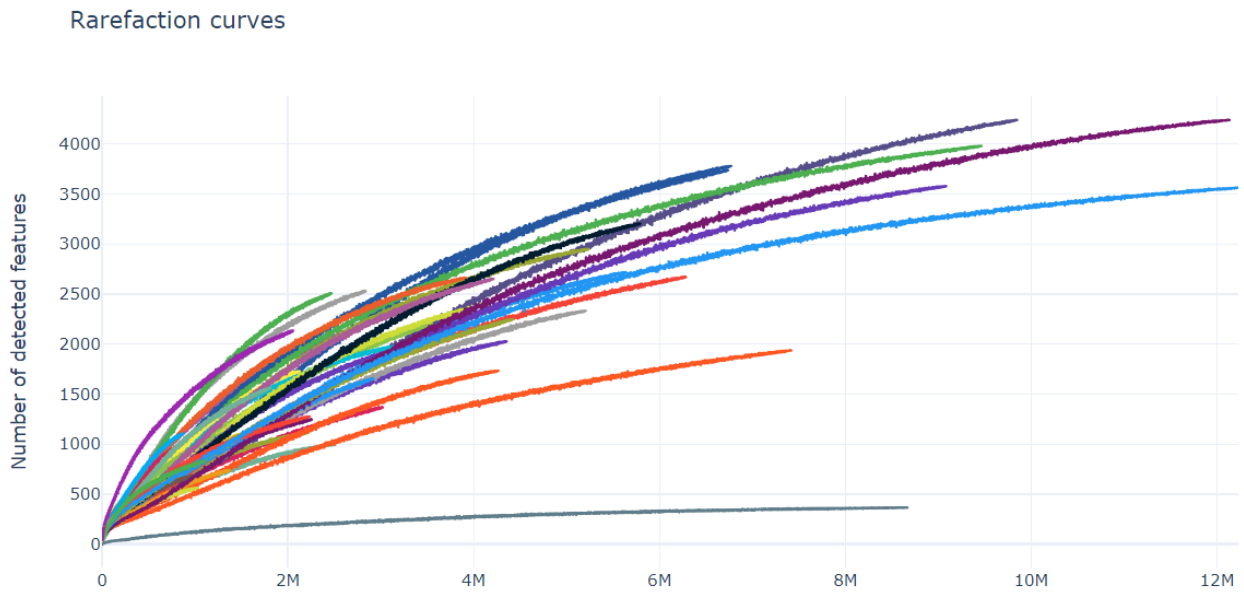


Figure 32: Rarefaction curves showing the number of detected species with various sequencing depth. Each curve represents a sample.

3.4.1.3 Spike-ins and contaminating species

Three spike-in species were added to each sample: *Truepera radiovictrix*, *Imtechella halotolerans*, and *Allobacillus halotolerans*. By comparing the relative abundance of these organisms in the sequencing data, potential biases during DNA extraction or library preparation can be detected. The expected ratio is 10:1:1, with *Truepera radiovictrix* being 10 times more abundant than *Imtechella halotolerans*, and *Imtechella halotolerans* being 10 times more abundant than *Allobacillus halotolerans*. However, *Allobacillus* was largely absent from the samples, and the abundances of *Truepera radiovictrix* and *Imtechella halotolerans* were nearly identical (Figure 33). Nevertheless, the total abundance of those spike-in species was relatively homogenous between the samples. The spike-in species were removed before the downstream analysis.

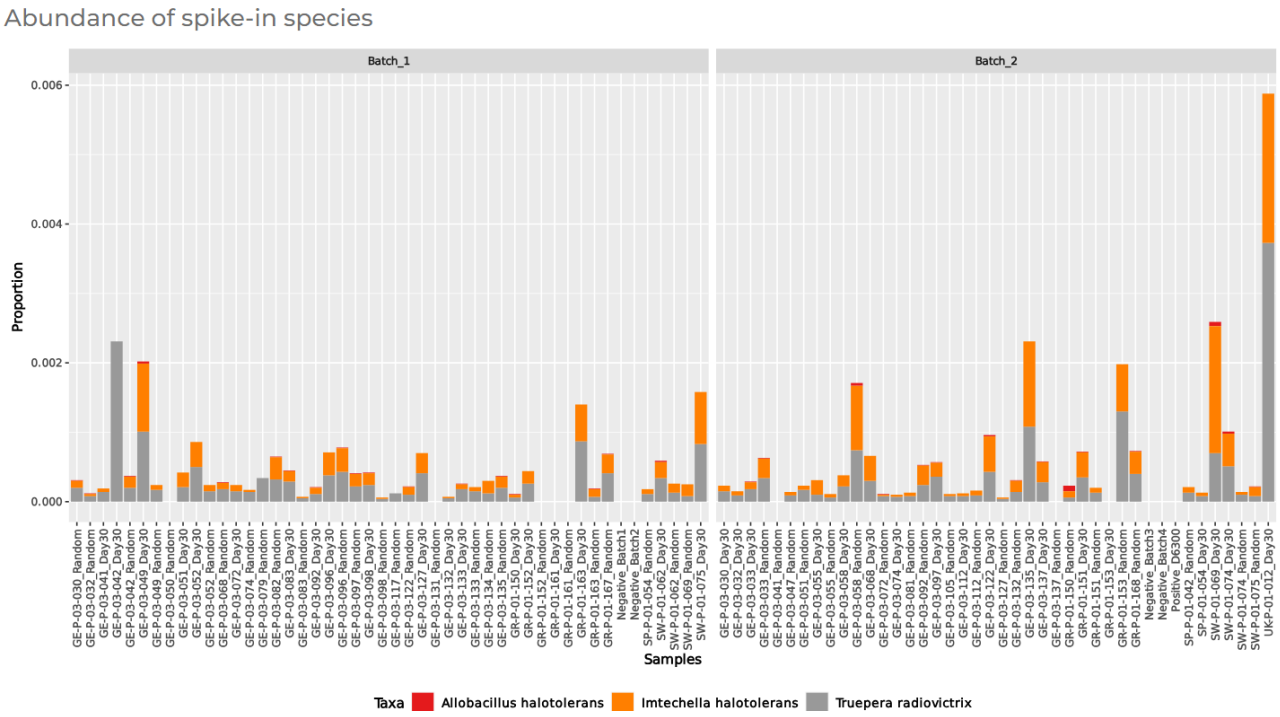


Figure 33: Barplot of abundance of three spike-in species in the samples. Samples are grouped by batch.

18 species have been detected in the negative control samples and therefore have been removed from the following analysis (Figure 34).

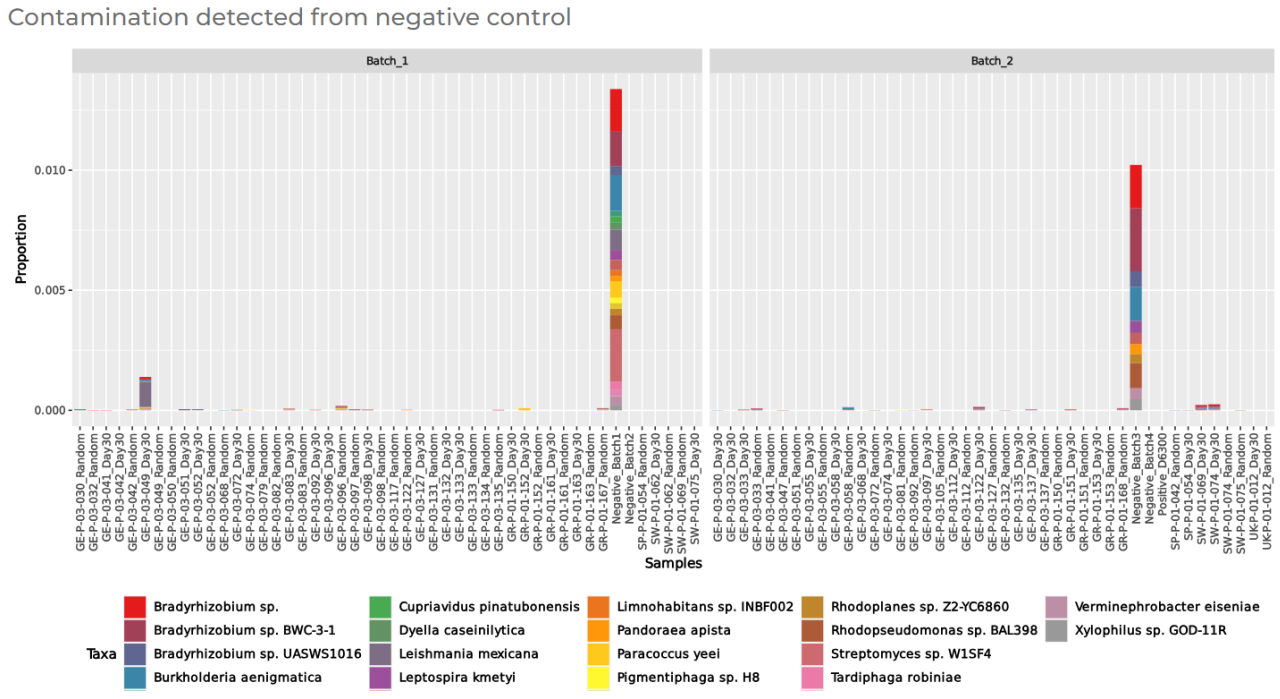


Figure 34: Barplot of abundance of contaminant species in the samples. Samples are grouped by batch.

3.4.1.4 Workflow validation with a mock sample

In order to validate the wet lab as well as the taxonomic classification workflow, we used a mock sample with well-defined mock community of microbes with known abundances. The mock samples comprised of 10 different species to mimic the human gut microbiome. All bacteria were detected with a similar abundance to the expected ones (Figure 35). The observed abundance of some gram-negative species seemed to slightly exceed the expected abundance while the abundance of gram-positive species seemed to be slightly lower than expected. These discrepancies are known and are due to the tough-to-lyse gram-positive bacteria leading to an under-representation of the latter and an over-representation of gram-negative bacteria. Therefore, the wet-lab and dry-lab workflow are validated.

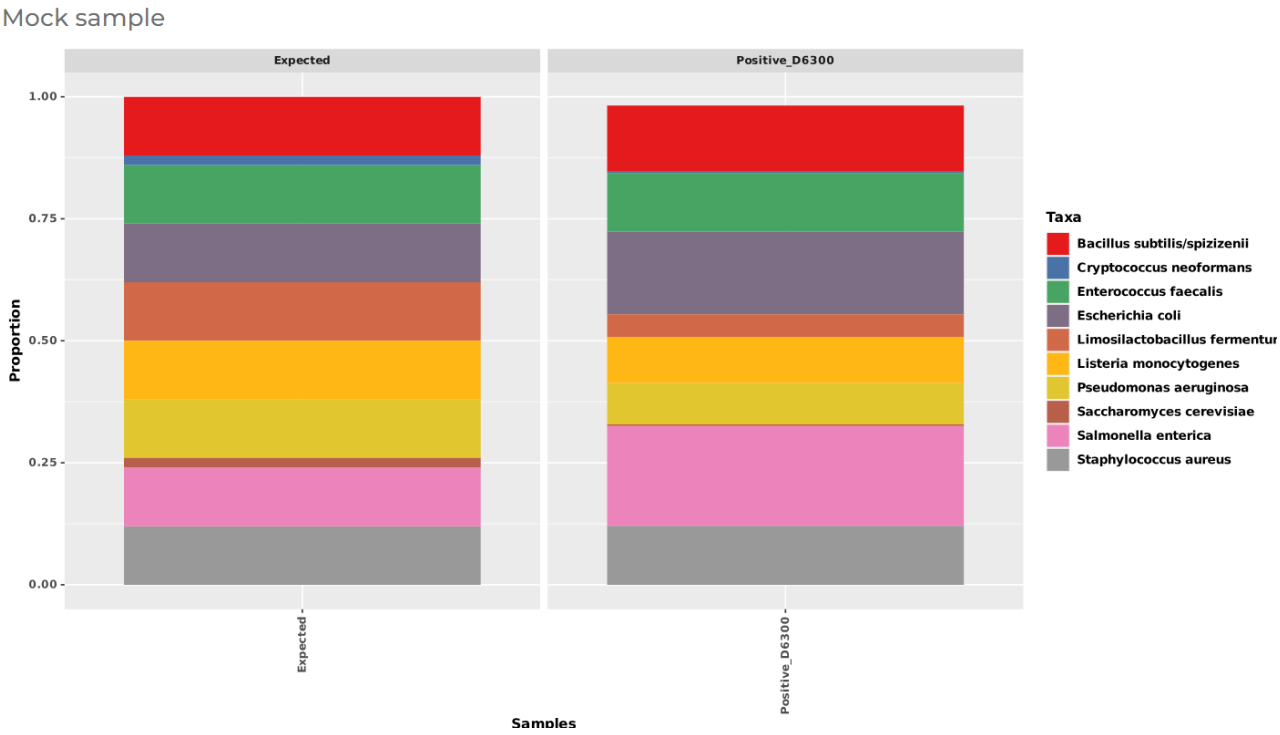


Figure 35: Barplot of the expected and observed abundance of the 10 bacteria present in the mock sample.

3.4.1.5 Data Normalization and sample filtering

TMM normalization, a widely used approach for NGS data and the default in *limma-voom*, was applied in this analysis. This method trims extreme values before calculating normalization factors, so it requires each sample to contain a substantial fraction of non-zero counts to be effective. Rather than excluding the 36 samples with fewer than 1 million reads, samples with fewer than 10% non-zero features (taxa or resistance genes) were instead excluded to ensure reliable normalization. Overall, 71 samples were retained in the following analyses (Table 13).

	Taxa		Resistance genes	
Day	**Antibiotic**	**Untreated**	**Antibiotic**	**Untreated**
D0	10	28	11	28
D30	9	24	10	22

Table 13: Sample distribution after filtering.

3.4.1.6 Variance partitioning

The variance partitioning step provides a visual way to examine the relative contribution of confounders and variables of interest to the overall variance. It can first be noted that the residual variance is high, indicating that the majority of the variance remains unexplained by the clinical and technical covariates (Figure 36). Before correction, site and extraction batch factors have the largest effect. After adjusting for confounders, their contribution becomes negligible, leaving only the effects of antibiotics and visit remain. Of note, after correction, a handful of taxa still exhibit explained variance associated with the confounding variable. These taxa have extremely low abundance; removing them results in an explained variance of zero.

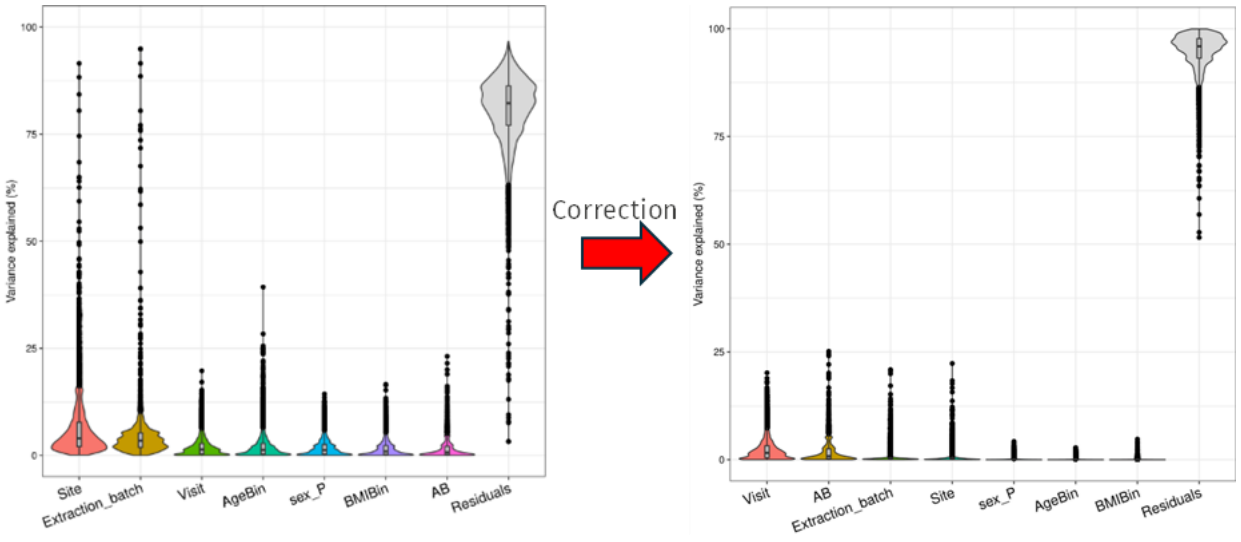


Figure 36: variance partitioning before and after correction of confounding effects.

3.4.2 Diversity analysis

3.4.2.1 Alpha diversity

Richness and evenness were computed with the Chao estimator and Shannon index for each sample. Both metrics, like the read counts, were adjusted for confounding effects.

A Student's t-test was performed at each time point to compare antibiotic versus untreated patients. It is expected that antibiotic treatments induce a microbiota dysbiosis, impacting thus the number of species and their relative abundance. Figure 37 actually reveals that no difference is significant for both richness and evenness indicators (p-values > 5%). It can nevertheless be noticed that while the median richness and evenness are higher at D0 in antibiotic-treated patients, an opposite trend is observed at D30. This latter observation is in line with the idea that antibiotics decrease diversity and evenness by suppressing or eliminating certain species. This creates an imbalance, with some resistant or unaffected species becoming more dominant.

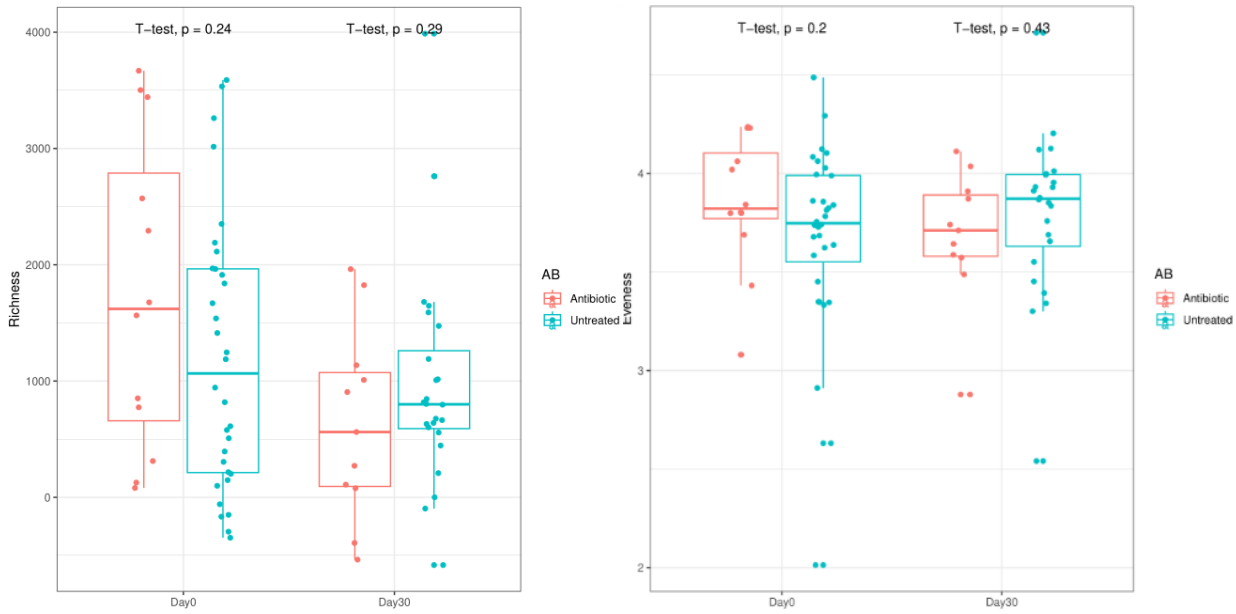


Figure 37: Richness and evenness boxplots. A Student's t-test was performed at each time point to compare antibiotic-treated and untreated patients.

3.4.2.2 Beta diversity

To compute the beta diversity, the Bray-Curtis distance matrix was built on the normalized, adjusted counts. The Adonis test was run to test whether there is a difference in the centroid and dispersion of the groups in the multivariate space. Although no significant difference was detected at either time point, a modest effect size can be observed at D30, suggesting a potential effect of antibiotics (Figure 38).

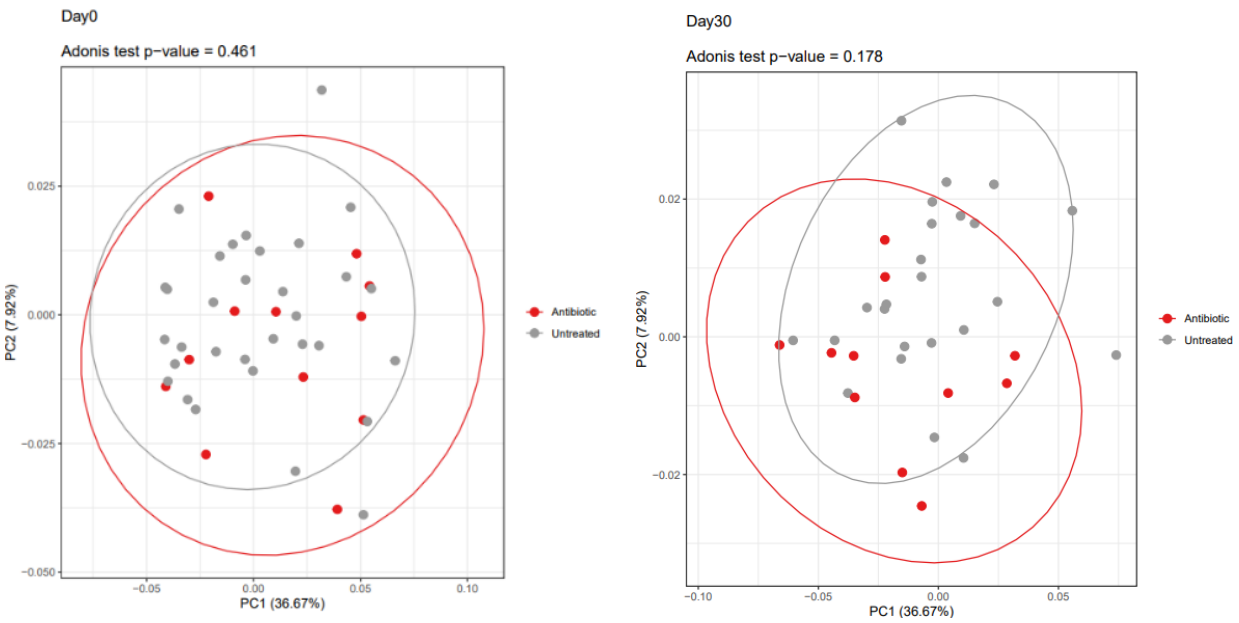


Figure 38: Beta diversity calculated using Bray-Curtis distance. An Adonis (PERMANOVA) was performed at each time point to compare antibiotic-treated versus untreated patients.

3.4.3 Differential analyses in taxa

This section presents antibiotic versus untreated intra- and inter-group comparisons at the taxon level. For each comparison, a volcano plot and a result table with average expression in $\log_2$(cpm), $\log_2$(fold-change), p-values, were generated and are available in supplementary materials (Supplementary Table 11 and Figure 11).

3.4.3.1 Most abundant taxa

Before getting into differential analyses, this section describes the most abundant taxa. Out of a total of 7919 taxa detected in all samples, a handful of species (25 shown in Figure 39) make up approximately half of the biomass. Out of the 14 generea detected in these most abundant bacteria, 6 belong to 7 genus previously described as dominant in the human oropharyngeal microbiota (***Streptococcus***, ***Haemophilus***, ***Neisseria***, ***Veillonella***, ***Prevotella***, ***Rothia***, *Actinomyces*, (Bogaert et al., 2011; de Steenhuijsen Piters. 2016). Of note, none of these top 25 most abundant bacteria are found differentially abundant in the following analyses, which is consistent with the observed high inter-individual variability.

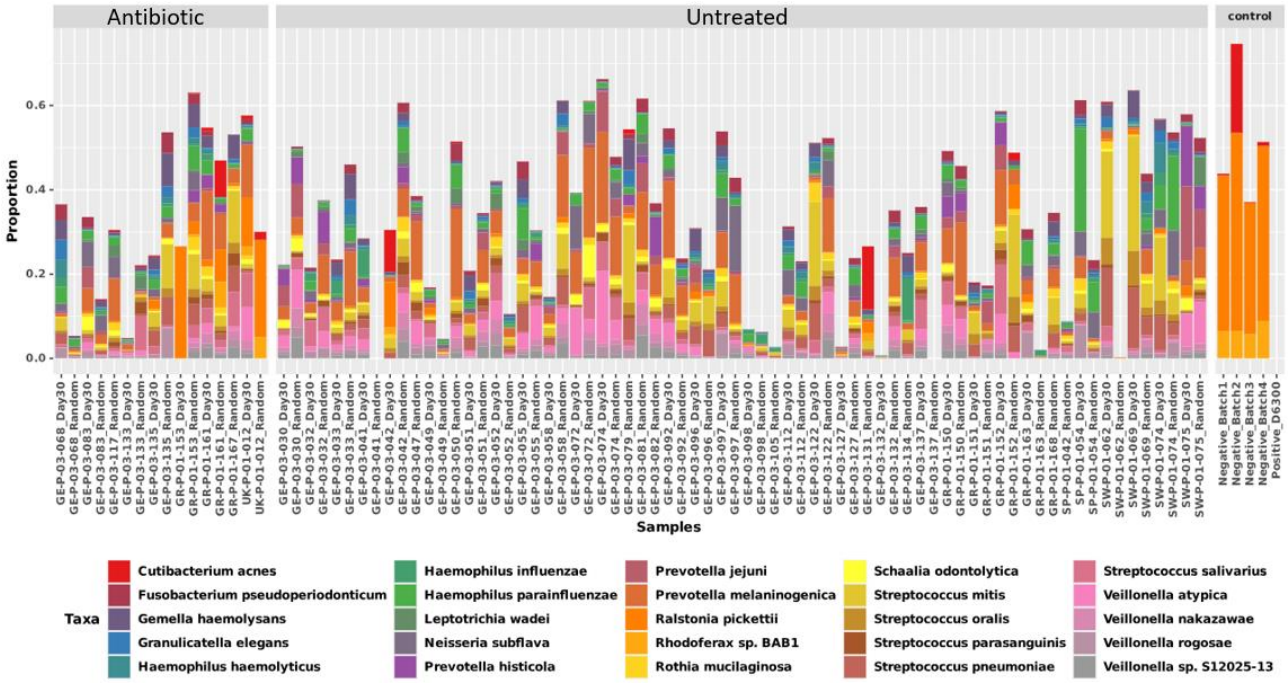


Figure 39: Relative abundance of the top 25 taxa stratified grouped by treatment.

3.4.3.2 Background noise determination

The authors of Limma-Voom recommend adjusting the background noise threshold so that the variance strictly decreases with the mean abundance. In case of bell-shaped curve, the left arm of the curve indicates a reduction in variance at low abundance levels, primarily due to the high number of zeros. Figure 40 reveals that the mean-variance trend resembles a log curve, with many taxa at low abundance levels. Applying a background noise threshold of $\log_2(\text{cpm}) > 2.5$ resulted in the removal of 6275 taxa out of 7919.

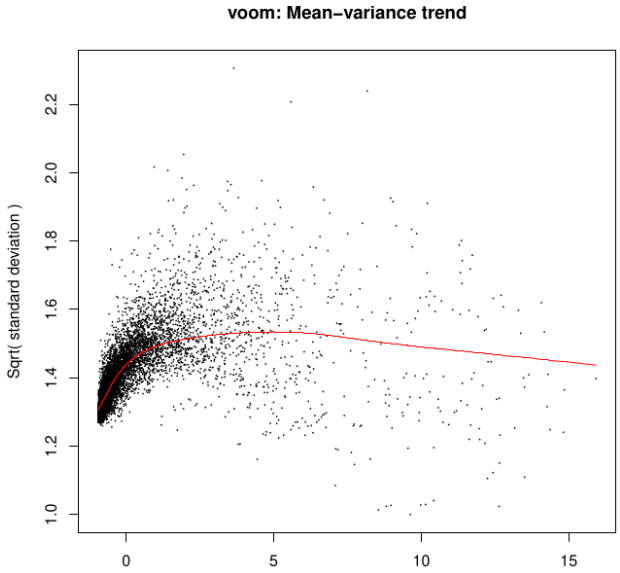
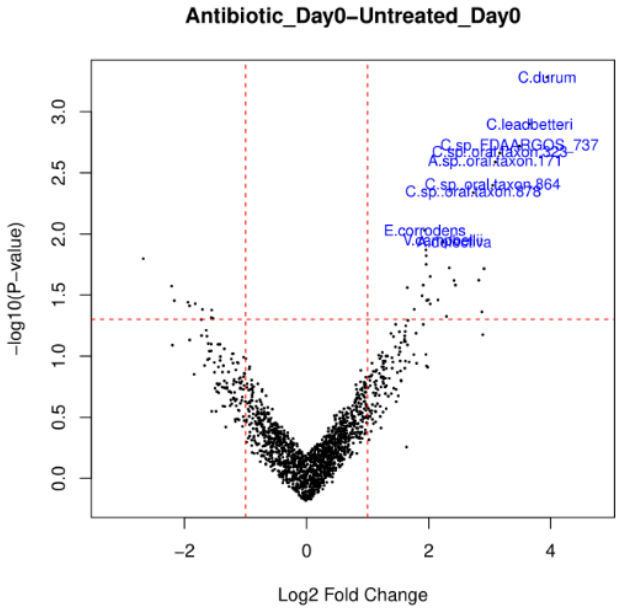


Figure 40: Mean-variance scatterplot. Each dot is the estimated mean and sqrt(standard deviation) of a given taxon. The red curve is a regression spline that estimate the overall mean-variance relationships. All taxa with mean log2(cpm) <2.5 were excluded from differential analyses.

3.4.3.3 Comparisons at baseline

An initial comparison was conducted at baseline. Although many taxa showed large |log2(fold-change)| values (up to 4) and were significant before multiple testing correction, none remained significant after correction (Figure 41). This result is in line with the absence of differences observed in beta diversity at D0. Although the results are not statistically significant, taxa with the largest log2(fold-change) tend to have higher abundance in the antibiotic group.



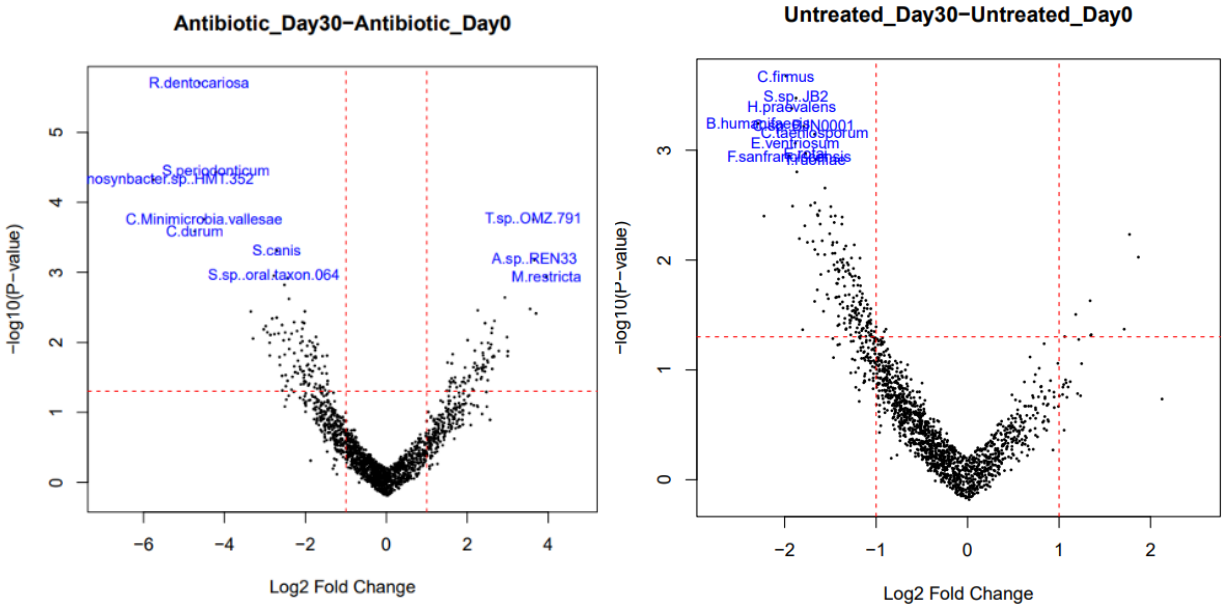
Taxa name	logFC	AveExpr	adj.P.Val
Corynebacterium.durum	3,941	4,782	0,867
Capnocytophaga.leadbetteri	3,652	8,715	0,939
Capnocytophaga.sp..oral.taxon.323	3,166	4,236	0,939
Capnocytophaga.sp..FDAARGOS_737	3,488	8,797	0,939
Capnocytophaga.sp..oral.taxon.864	3,051	4,517	0,939
Actinomyces.sp..oral.taxon.171	3,092	4,942	0,939
Capnocytophaga.sp..oral.taxon.878	2,731	7,055	0,994
Abiotrophia.defectiva	2,417	7,673	0,997
Vibrio.campbellii	2,232	0,171	0,997
Treponema.sp..OMZ.838	2,907	1,391	0,997

Figure 41: Volcano plot showing taxa with largest differences between antibiotic and untreated patients at baseline (left). The horizontal and vertical dotted lines represents the (unadjusted) p-value=5% and |log2(fold-change)|<1 thresholds respectively. Differential analysis results at baseline (right). Shown are the log_2(fold-change), average abundance (log_2CPM), raw and adjusted p-values in the top10 taxa.

3.4.3.4 Intra-group comparisons

We next investigated whether differences between D30 and baseline exist in both antibiotic-treated and untreated patients. Figure 42 reveals that the range of log2(fold-change) and p-values is substantially larger in the antibiotic group, indicating a stronger effect. While no taxa reach significance in the untreated group, a trend toward negative log2(fold-change) suggests a decline in the abundance of certain species. In the antibiotic

group, three bacteria - *Rothia dentocariosa*, *Candidatus Nanosynbacter*, and *Streptococcus periodonticum* - show significant negative log2(fold-change), reflecting a decrease in abundance at D30. Considering that a limited number of patients received antibiotics (Table 13), it is plausible that more significant bacterial changes could emerge in a larger study.

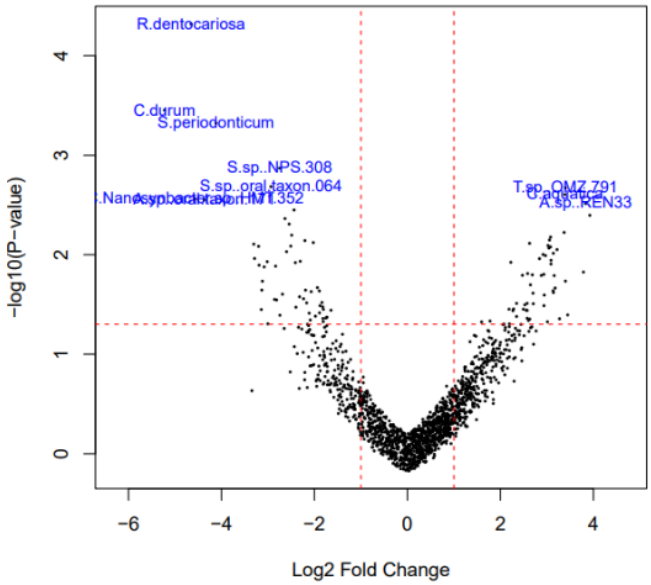


Taxa name	logFC	AveExpr	adj.P.Val	Taxa name	logFC	AveExpr	adj.P.Val
Rothia.dentocariosa	-4,624	6,556	0,004	Cytobacillus.firmus	-1,986	1,517	0,213
Candidatus.Nanosynbacter.sp..HMT.352	-5,759	9,961	0,026	Sarcina.sp..JB2	-1,876	2,369	0,213
Streptococcus.periodonticum	-4,213	4,882	0,026	Halanaerobium.praevalens	-1,922	0,988	0,213
Corynebacterium.durum	-4,736	4,782	0,075	Eubacterium.ventriosum	-1,885	3,014	0,213
Treponema.sp..OMZ.791	3,632	-1,091	0,075	Bacteroides.humanifaecis	-2,290	0,544	0,213
Candidatus.Minimicrobia.vallesae	-4,502	6,321	0,075	Fructilactobacillus.sanfranciscensis	-1,949	1,295	0,213
Agrococcus.sp..REN33	3,658	0,320	0,150	Clostridium.sp..BJN0001	-1,792	2,379	0,213
Streptococcus.canis	-2,711	5,174	0,150	Clostridium.taeniosporum	-1,672	2,279	0,216
Streptococcus.sp..oral.taxon.064	-2,780	7,861	0,177	Enterococcus.rotai	-1,772	1,616	0,278
Malassezia.restricta	3,958	1,172	0,177	Ignavigranum.ruoffiae	-1,656	1,031	0,278

Figure 42: Volcano plots showing taxa with largest differences in antibiotic (left) and untreated (right) intra-group comparisons (top). Top 10 taxa from the differential analyses associated with antibiotic and untreated comparisons (bottom).

3.4.3.5 Inter-group comparison at Day 30

The inter-group comparison between antibiotic-treated and untreated patients is represented as (ATB_D30 - ATB_D0) - (Untr_D30 - Untr_D0). Given that the intra-group difference in the antibiotic group is greater than that in the untreated group, the inter-group comparison primarily reflects changes seen in the antibiotic-treated group. The examination of the most differentially abundant taxa (Figure 43) reveals that five of the top ten species are common to both the antibiotic intra-group and inter-group comparisons.



Taxa name	logFC	AveExpr	adj.P.Val
Rothia.dentocariosa	-4,654	6,556	0,061
Corynebacterium.durum	-5,227	4,782	0,316
Streptococcus.periodonticum	-4,119	4,882	0,316
Streptococcus.sp..NPS.308	-2,749	9,296	0,515
Actinomyces.sp..oral.taxon.171	-4,371	4,942	0,515
Agrococcus.sp..REN33	3,833	0,320	0,515
Comamonas.aquatica	3,386	3,651	0,515
Streptococcus.sp..oral.taxon.064	-2,937	7,861	0,527
Treponema.sp..OMZ.791	3,398	-1,091	0,527
Streptococcus.toyakuensis	-2,633	11,313	0,527

Figure 43: Volcano plot showing taxa with largest differences in inter-group comparison at D30 (top). Top 10 taxa from the differential analysis (bottom).

3.4.4 Differential analyses in resistance genes

In this section, we repeat the previous differential and longitudinal analyses at the resistance gene level. All figures and full tables are available in supplementary materials (Supplementary Table 12 and Figure 12).

3.4.4.1 Background noise determination

The background noise threshold was adjusted to ensure a monotonically decreasing mean-variance trend. As previously observed at the taxa level, Figure 44 shows that the mean-variance trend resembles a noisy logarithmic curve, with many genes at low abundance. Setting a background noise threshold of log2(cpm) > 2 resulted in the removal of 500 out of 638 genes. This stringent filtering reflects the matrix's high sparsity, with approximately 90% of entries being zeros, as illustrated in the histogram below.

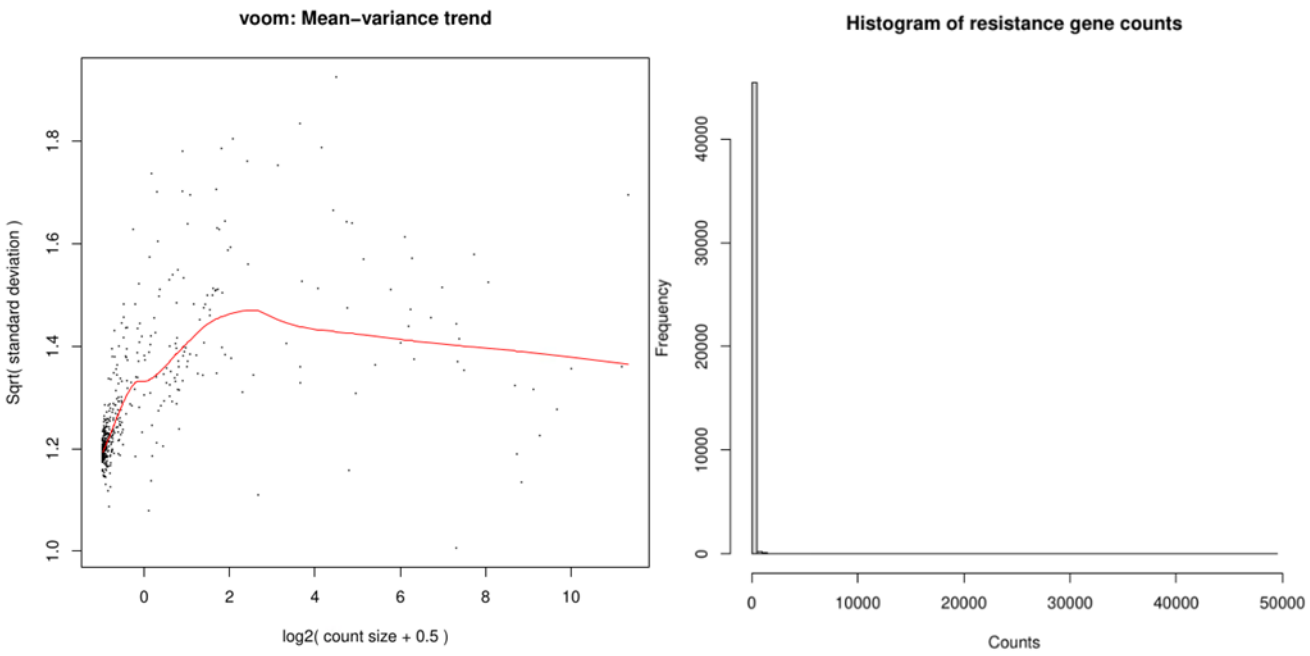
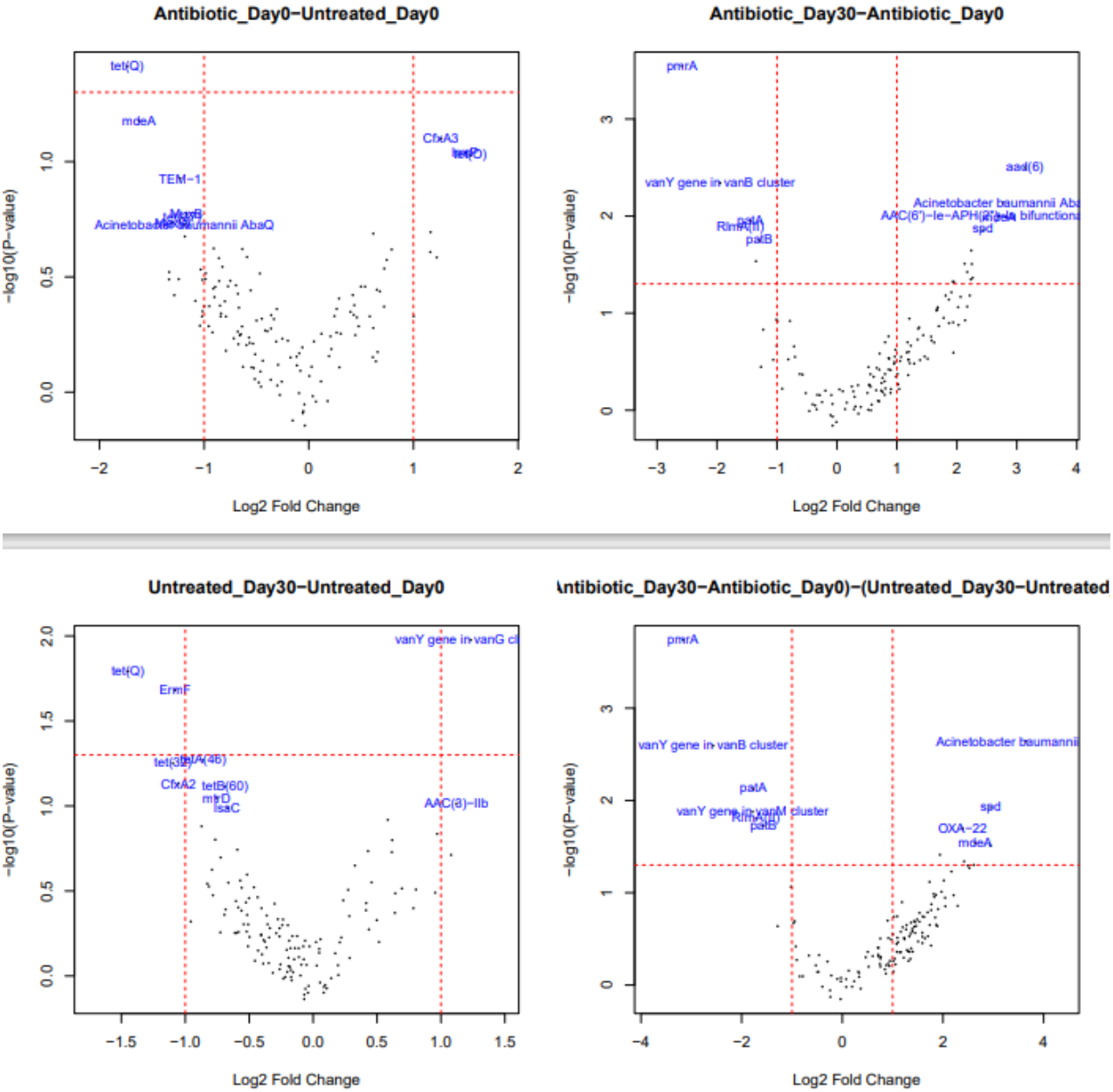


Figure 44: Mean-variance scatterplot (left). All taxa with mean log2(cpm) <2 were excluded from differential analyses. Histogram of raw counts in the resistance gene matrix (right).

3.4.4.2 Intra- and inter-group comparisons

Of the four comparisons, only the antibiotic intra-group (D30 - D0) and inter-group comparisons at D30 yielded significant results: *pmrA* showed a log2(fold-change) of -2.6 (adjusted p-value = 6.1%) and -3.2 (adjusted p-value = 3%) in these comparisons, respectively (Figure 45). No genes reached significance at baseline or in the untreated intra-group comparison. Similar to the taxa results, the top genes in the D30 inter-group comparison largely overlap with those in the antibiotic intra-group (D30 - D0) comparison. However, unlike taxa, their magnitude and significance are greater in the inter-group comparison. In other words, these genes exhibit opposite trends over time between the antibiotic and untreated groups. As noted in the taxa section, the small number of patients receiving antibiotics limits the power to detect significant changes, thereby hindering the identification of resistant genes.



Antibiotic_Day0-Untreated_Day0				Untreated_Day30-Untreated_Day0			
Gene name	logFC	AveExpr	adj.P.Val	Gene name	logFC	AveExpr	adj.P.Val
tet(Q)	-1,735	11,984	0,905	tet(Q)	-1,448	11,984	0,965
CfxA3	1,263	13,913	0,905	vanY gene in vanG cluster	1,234	9,528	0,965
mdeA	-1,622	8,060	0,905	ErmF	-1,079	11,380	0,967
tet(O)	1,541	7,947	0,905	tet(32)	-1,096	9,484	0,967
lmrP	1,496	6,396	0,905	tetB(60)	-0,688	10,403	0,967
tet(M)	1,161	11,860	0,905	tetA(46)	-0,860	11,728	0,967
Acinetobacter baumannii AbaQ	-1,188	6,661	0,905	CfxA2	-1,053	10,152	0,967
vanY gene in vanF cluster	1,164	7,182	0,905	mtrD	-0,755	13,137	0,967
TEM-1	-1,229	7,335	0,905	lsaC	-0,672	11,580	0,967
APH(3')-IIIa	1,223	7,270	0,905	AAC(3)-IIb	1,121	8,483	0,967
Antibiotic_Day30-Antibiotic_Day0				**(Antibiotic_Day30-Antibiotic_Day0)-(Untreated_Day30-Untreated_Day0)**			
Gene name	logFC	AveExpr	adj.P.Val	Gene name	logFC	AveExpr	adj.P.Val
pmrA	-2,582	13,073	0,061	pmrA	-3,167	13,073	0,030
aad(6)	3,135	6,551	0,230	vanY gene in vanB cluster	-2,565	15,546	0,104
vanY gene in vanB cluster	-1,947	15,546	0,230	Acinetobacter baumannii AbaQ	3,647	6,661	0,104
Acinetobacter baumannii AbaQ	2,775	6,661	0,230	patA	-1,768	14,742	0,264
AAC(6')-Ie-APH(2'')-Ia bifunction	2,812	7,320	0,230	spd	2,949	6,985	0,352
patA	-1,441	14,742	0,230	patB	-1,566	15,282	0,367
mdeA	2,703	8,060	0,230	vanY gene in vanM cluster	-1,780	16,817	0,367
RlmA(II)	-1,610	13,105	0,230	RlmA(II)	-1,718	13,105	0,367
spd	2,434	6,985	0,258	tet(37)	2,938	8,013	0,367
patB	-1,292	15,282	0,258	OXA-22	2,391	8,517	0,367

Figure 45: Volcano plot showing taxa with largest differences in intra and inter-group comparisons (top). Top 10 taxa from the differential analysis (bottom).

4 Conclusion

Due to the COVID-19 pandemic, clinical studies (WP4) suffered a delayed, which in turn negatively impacted patient enrolments. This resulted in smaller cohorts and reduced statistical power, especially at the antibiotic class level, where few patients received the same class of antibiotics.

In the ADEQUATE study, out of 49 patients recruited, 15 received antibiotics but 3 samples were missing, reducing the treated group to 12 patients. Among these, 7 were treated with broad spectrum penicillins, 2 with macrolide, 1 with both, 1 with narrow spectrum penicillin and 1 with an antibiotic of the lincosamide class. Similarly, in the PRUDENCE study, only 3 antibiotic classes (broad and narrow spectrum penicillins, and macrolides) had sufficient sample sizes for biostatistical analysis.

In addition to the limited sample size, start and end dates of antibiotic therapy were not systematically available, especially for delayed antibiotherapy, resulting in hypothetical treatment between day 1 and day 7.

Across the three analyses, three main trends emerged: (1) microbiota variation was primarily influenced by sampling site, necessitating adjustment for this and other confounders; (2) untreated patients' microbial communities showed no change over time; and (3) antibiotics had a modest overall impact on both nasopharyngeal and stool microbiota.

More specifically, in the PRUDENCE nasopharyngeal analysis, the differential analysis identified only two taxa, *Corynebacterium* and *Dolosigranulum*, both associated with healthy microbiota. In stool samples, two species (*P. massiliensis* and *A. propionicum*) were found significant at baseline in the antibiotic and macrolide inter-group comparisons. In macrolide intra-group comparisons, 67 taxa showed significant U-shape trajectories, reaching minimum abundance at day 7. These results should be interpreted cautiously, as they rely on two samples at this time point. Additionally, three and two resistant genes were identified in narrow-spectrum penicillin and macrolide respectively, the latter (*vanY* and *APH(6)-Id*) showing an upward trend indicative of potential resistance development. Overall, these findings suggest that antibiotics had a limited impact on the intestinal microbiota, but that narrow-spectrum penicillin and macrolides may pose a higher risk for developing AMR in CA-ARTI patients.

In the ADEQUATE oropharyngeal analysis, fewer significant results were observed. Three species (*Rothia dentocariosa, Candidatus Nanosynbacter*, and *Streptococcus periodonticum*) and one resistance gene (*pmrA*) were found with significantly decreased abundance by day 30. Unexpectedly, *S. pneumoniae*, *P. aeruginosa* and *K. pneumoniae*, the most common Gram-positive or Gram-negative pathogens expected in CA-ARTI patients, did not vary significantly across time points or between treatment groups. Unfortunately, the small sample size and low sequencing depth limited conclusive interpretations in the ADEQUATE study.

Finally, enrichment analysis was initially considered to uncover common antibiotic classes associated with resistant genes, but this analysis was deferred due to the limited number of significant findings.

5 References

Andrews, S. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. 2010 [cited 2022 Feb 18]. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bogaert, D., Keijser, B., Huse, S., Rossen, J., Veenhoven, R., Van Gils, E., ... & Sanders, E. (2011). Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. *PloS one*, *6*(2), e17035.

Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.

de Steenhuijsen Piters, W. A., Heinonen, S., Hasrat, R., Bunsow, E., Smith, B., Suarez-Arrabal, M. C., ... & Mejias, A. (2016). Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. American journal of respiratory and critical care medicine, 194(9), 1104-1115.

Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Comput Sci. 2017;3:e104.

McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The Comprehensive Antibiotic Resistance Database. Antimicrob Agents Chemother. 2013;57:3348–57.

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology*, *11*, 1-9.

Segata, N., Haake, S. K., Mannon, P., Lemon, K. P., Waldron, L., Gevers, D., ... & Izard, J. (2012). Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. *Genome biology*, 13, 1-18.

Tremaroli, V., & Bäckhed, F. (2012). Functional interactions between the gut microbiota and host metabolism. *Nature*, *489*(7415), 242-249.

Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biology. 2019;20:257.