Unravelling the landscape of antibiotic resistance determinants in the nasopharynx and the impact of antibiotics: a longitudinal study of preterm infants

Achal Dhariwal[1]†, Polona Rajar[1,2]†, Gabriela Salvadori[1], Heidi Aarø Åmdal[1], Dag Berild[3,4], Ola Didrik Saugstad[5], Drude Fugelseth[2,4], Gorm Greisen[6], Ulf Dahle[7], Kirsti Haaland[2], Fernanda Cristina Petersen[1]*

[1] Institute of Oral Biology, Faculty of Dentistry, University of Oslo, Oslo, Norway.
[2] Department of Neonatal Intensive Care, Division of Pediatric and Adolescent Medicine, Oslo University Hospital Ullevål, Oslo, Norway.
[3] Department of Infectious Diseases, Oslo University Hospital, Oslo, Norway.
[4] Institute of Clinical Medicine, Faculty of Medicine, Oslo University, Oslo, Norway.
[5] Department of Department of Pediatric Research, Oslo University Hospital, University of Oslo, Oslo, Norway.
[6] Department of Neonatology, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark.
[7] Centre for Antimicrobial Resistance, Norwegian Institute of Public Health, Oslo, Norway

†Shared First authorship
*Correspondence: Fernanda Cristina Petersen (f.c.petersen@odont.uio.no)

ABSTRACT

Respiratory infections caused by common pathogens that colonize the respiratory tract, including the nasopharynx, are a leading global cause of deaths associated with antimicrobial resistance. Yet, little is known about the antibiotic resistance determinants in the microbial communities of the respiratory tract. In this study, we examined the development of the resistome (collection of antibiotic resistance genes) in the nasopharynx of preterm infants and the impact of early ampicillin plus gentamicin exposure in those requiring broad-spectrum antibiotic treatment for suspected early-onset neonatal sepsis. We analyzed the resistome of 181 nasopharyngeal samples collected longitudinally from 36 preterm infants using deep shotgun metagenomic sequencing. Antibiotic resistance genes were found in most samples (~95%) of both naive and antibiotic-treated infants. We observed a significant increase in the diversity and total abundance of antibiotic resistance genes primarily immediately after early antibiotic therapy. The nasopharyngeal resistome in preterm infants was mainly driven by the microbiome ($r = 0.85$, p-value = 0.001), with inter-individual differences explaining most of the variation observed in the resistome composition ($R^2 = 30.03\%$, adjusted p-value = 0.001). We also noted a discernible impact of prenatal maternal antibiotic exposure on the overall resistome composition ($R^2 = 3.6\%$, adjusted p-value = 0.001), which was observed as a subgroup within the antibiotic-treated group. Overall, these data suggest that the nasopharynx in preterm infants is characterized by a rich resistome with a trajectory that is negatively affected by antibiotics. Although the effects appear to be transient, any adverse changes in resistome during the first days of life bear risks, particularly those selecting for drug resistant colonizers with the potential to spread and cause invasive infections.

Keywords: respiratory resistome, metagenomics, nasopharynx, microbiome, preterm infants, antibiotics, antimicrobial resistance

INTRODUCTION

Antimicrobial resistance (AMR) is one of the biggest global health threats, estimated to be associated with approximately 5 million deaths in 2019. Among these, respiratory infections are the leading AMR infection burden, accounting for more than 1.5 million deaths [1]. These are often caused by pathogens such as *Streptococcus pneumoniae, Staphylococcus aureus, Klebsiella pneumoniae, and Haemophilus influenzae*, which are part of the normal human microbiome found in the respiratory tract, including the nasopharynx. Recent studies have shown that for these pathogens, more than 90% of antibiotic exposure represents bystander events, in which they are not the infective agent [2]. This highlights the importance of the potential off-target impact of antibiotics in the selection and development of antibiotic resistance by human pathogens. These pathogens share common niches with commensals, which provide colonization resistance by preventing pathogen overgrowth. However, commensals can also serve as reservoirs of antibiotic resistance genes (ARGs) that can readily be transferred between different strains and species of bacteria in microbiomes [3-5]. Together, commensals and pathogens interact to provide additional benefits, such as immune system education and development [6, 7]. Such interactions highlight the importance of an ecological perspective when considering AMR and the impact of antibiotics on microbiomes.

Recent studies have contributed to an increased understanding of the landscape of ARGs, known as resistomes, and the impact of antibiotics on the gut microbiome and resistome [8-12]. Of critical relevance are studies focusing on vulnerable groups, in particular preterm infants, who are highly exposed to antibiotic interventions early in life [9, 13-15]. Sepsis is a serious complication for preterm infants and one of the main reasons why many receive empiric

antibiotics in the first days or weeks after birth [16, 17]. However, the risks associated with suspected early onset sepsis (sEONS) vary, and over-prescription of antibiotics in low sepsis risk preterm infants may result in unnecessary microbiome and resistome disturbances with potential short- and long-term health implications [18]. Early and prolonged antibiotic use in this group has been associated with an increased risk of various adverse outcomes, including bronchopulmonary dysplasia, late-onset sepsis, invasive fungal infections, and mortality, with conflicting results regarding necrotizing enterocolitis [19-26]. Furthermore, antibiotic use in early life may have unwanted consequences beyond the neonatal period, such as an increased risk of asthma and obesity [27-29]. In relation to antibiotic resistance, studies focusing on gut resistomes indicate that infants exhibit higher resistome burden than their adult mothers, and that antibiotic exposure amplifies this burden [9-11, 14, 15, 30, 31].

Although most of the associations between microbial disturbances and the risk of common morbidities are well-acknowledged for the gut microbiome in infants, growing evidence suggests that respiratory health is also significantly influenced by microbial communities in the upper respiratory tract. This includes the nasopharynx, which serves as a critical reservoir of respiratory pathogens and a gatekeeper of respiratory health [5-7, 32, 33]. Recent studies have highlighted the potential negative impact of antibiotic treatment on the nasopharynx microbiome and its association with an increased risk of developing asthma and respiratory infections in infants [10, 33, 34]. There is currently a significant gap in our understanding of how the infant respiratory resistome develops, including the trajectory, dynamics, and factors that can influence its establishment in preterm infants.

We recently developed an optimization strategy for DNA extraction and library preparation that helped to circumvent the challenges posed by low biomass and high host DNA content in nasopharynx samples of preterm infants [35]. In this study, we applied the optimized protocol to nasopharyngeal aspirate samples to characterize the development of the resistome in 36 preterm infants from birth up to 6 months corrected age. By analyzing ~1.85 terabytes (Tb) of metagenomic DNA data generated from deep shotgun sequencing, we obtained the most comprehensive information to date on the characteristics and dynamics of resistome development in the respiratory tract of preterm infants. Our hypothesis was that early exposure to broad-spectrum antibiotics for treatment of sEONS in preterm infants, may promote the enrichment of ARGs and compositional alterations in the respiratory resistome, which would persist until 6 months corrected age, as compared to antibiotic-naive preterm infants. Additionally, we explored the extensive metadata collected to gain insights into additional clinical covariates that may be associated with resistome development.

MATERIALS AND METHODS

Ethics statement

Premature infant metadata and all nasopharynx aspirate samples used in this study were collected at the Neonatal Intensive Care Unit (NICU) at Oslo University Hospital, Ullevål, Norway, as part of the Born in the Twilight of Antibiotic project. The study received approval from the Regional Committee for Medical and Health Research Ethics – South East, Norway (2018/1381 REKD), and followed the principles of the Declaration of Helsinki guidelines. Infants were enrolled in the study after obtaining written informed consent from their parents.

Study design and sample collection

The study is a prospective observational cohort study. Between July 2019 and January 2021, we included preterm infants born (or transferred within 48 hours) at the NICU (Ullevål) who were between 28+0 and 31+6 weeks of gestational age (GA). In total, 66 preterm infants were enrolled in the study, and followed from birth until six months corrected age (i.e., adjusted for prematurity) during their stay in the NICU and after discharge. Five of the infants were lost to follow-up. Extensive metadata and covariates from infants and mothers, including mode of delivery, gestational age (GA) at birth, postmenstrual age (PMA), gender, birth weight (BW), postnatal infant antibiotics usage (duration, type, days), maternal antibiotic usage, feeding regimen during hospitalization and after discharge, and other demographic characteristics, were collected through medical records. This data is stored in secured server at the Services for Sensitive Data (TSD), University of Oslo.

In our cohort of 66 infants, 33 infants with sEONS received broad-spectrum antibiotics, consisting of a combination of ampicillin (a beta-lactam antibiotic) and gentamicin (an aminoglycoside antibiotic), within 24 hours after birth. Due to the circumstances of premature birth and the instability of the infants, baseline samples could not be collected before antibiotic therapy initiation. Instead, the first nasopharyngeal aspirate samples (baseline) were obtained within 48 hours after the initiation of early antibiotic therapy. To investigate temporal changes in the resistome in response to antibiotics and control for potential baseline differences as confounding factors, this study included only infants who had nasopharyngeal samples collected on the same day as the initiation of early antibiotic treatment ($n = 15$) and antibiotic-naive infants

as controls (*n = 21*). The baseline characteristics of the infants stratified by antibiotic exposure are presented in Table 1.

Nasopharyngeal aspirate samples were collected from infants at six distinct time points: days of life (DOL) 0, 7, 14, 28, 56, and at 6 months corrected age. Additional samples were taken if the antibiotic therapy was initiated or terminated more than 48 hours before/after a predefined sampling time. In cases where the sample could not be collected at the scheduled time points (such as with an unstable infant), a new sample was obtained as soon as possible. Infants who were transferred to another hospital before discharged to home adhered to the study protocol. When the infants reached six months of corrected age, they were either invited to an outpatient visit at Ullevål or their local hospital or visited at home for sample collection. Nasopharynx aspirates were collected as previously described [35] and stored at -80°C until further processing.

Metagenomic DNA extraction and sequencing

Metagenomic DNA was extracted by first thawing the aspirate samples (2 ml) on ice, pelleting them by centrifugation at 10,000 g for 10 minutes at 4°C, and then using MolYsis™ for host DNA depletion (Molzym, Bremen, Germany), following the manufacturer's protocol with some modifications, as previously described [35]. After depletion of human DNA, fresh pellets were spiked with 20 µl of ZymoBIOMICS Spike-in Control II (catalog number: D6321 & D6321-10), and bacterial DNA was extracted using the MasterPure™ Gram Positive DNA Purification Kit (Epicentre, Madison, WI, USA), according to the manufacturer's recommendations. The bacterial DNA was eluted in 35 µl of TE buffer and stored at -80°C until further use. Final DNA concentration was quantified using Qubit dsDNA HS assay kits in a Qubit 4.0 Fluorometer

(Invitrogen, Thermo Fisher Scientific, USA). Metagenomic sequencing libraries were constructed using the Nextera DNA Flex library preparation kit (Illumina Inc., CA, USA). All samples were subjected to paired-end sequencing at 2 × 150 bp on an Illumina NovaSeq S4 high-output platform at the Norwegian Sequencing Centre (Oslo, Norway).

Controls for low microbial biomass

The study followed recent recommendations for low microbial biomass microbiome research [36]. This involved the following steps: (1) collecting samples from the included infants in the two groups within the same time period, (2) using standard protection measures during sample collection (such as clean suits, disposable gloves, and face masks), (3) including sampling and DNA extraction blank controls (negative controls), (4) conducting DNA extraction and library preparation while wearing laboratory coats and disposable gloves in a carefully cleaned environment , and (5) using unique barcodes in library preparation. The negative controls had no or extremely low DNA concentrations to proceed to library preparation as determined by RT-PCR using the Femto Bacterial DNA Quantification Kit from Zymo and were therefore excluded from further analysis ($n = 10$; mean = 0.00043 ng per µL). Additionally, spike-in controls ($n = 6$) were extracted and used as positive controls, and longitudinal samples from each preterm infant were extracted within the same run.

Bioinformatics processing

The raw metagenomic sequencing data was pre-processed using our in-house bioinformatics pipelines using a high-performance computing cluster inside a secured environment, i.e., TSD at

the University of Oslo. In brief, the Nextera adaptor sequences and low-quality reads were filtered using Trim Galore (v.0.6.1) with default parameters [37]. Next, the human DNA contaminant sequences were identified and removed by mapping the quality-trimmed reads against the human reference genome (GRCh38) using Bowtie2 (v.2.3.4.2) [38] (non-default parameters: *q -N 1 -k 1 --fr --end-to-end --phred33 --very-sensitive –no-discordant*) along with SAMtools (v.1.9) [39] and BEDTools (v.2.27.1) [40]. The sequence quality reports on raw and processed reads were generated using FastQC (v.0.11.9) [41] and summarized using MultiQC (v.1.7) [42].

Resistome profiling

To characterize the presence of ARGs, cleaned high-quality reads for each sample were mapped against the *nucleotide_fasta_protein_homolog_model* from the Comprehensive Antibiotic Resistance Database (CARD) (v.3.2.2) [43] using Bowtie2 under parameter *–very-sensitive-local*. The mapped reads from each sample were then filtered, sorted, and indexed using SAMtools. The number of reads mapped to each ARG was calculated using SAMtools *idxstats* and BEDTools. ARGs with a coverage of at least 80% from the alignment were retained for further downstream analyses. The mapped read counts were normalized for bacterial sequence abundances and gene lengths by calculating reads per kilobase of reference gene per million bacterial reads (RPKM) for each CARD reference sequence. The relative abundance of ARGs for each sample was computed by dividing the RPKM by the total sum of the RPKM for each sample. For keeping the acyclic hierarchical annotation structure for accurate resistome profiling at higher functional levels, ARGs were manually re-annotated based on the drug class to which

they confer resistance. ARGs belonging to macrolides, lincosamides and streptogramins were congregated into the MLS class. While ARGs belonging to carbapenem, cephamycin, cephalosporin, penem, penam and monobactam were congregated into the Beta-lactam class. Rarefaction analysis was performed using Rarefaction Analyzer [44] to determine the saturation of samples at various sequencing depths for recovery of ARGs.

Statistical analysis and data visualization

All statistical analyses were conducted in R (v.4.2.1) within RStudio (v.2022.07.2+576) [45, 46]. The ARG abundance, annotation table, and metadata file were compiled into a single data object using the phyloseq (v.1.40.0) package [47]. Figures, unless stated otherwise, were created using the ggplot2 (v.3.4.0) package [48] and further edited using Adobe Illustrator (v.16.0.0). α- and β-diversity analysis was performed using the vegan (v.2.6.2) [49] and phyloseq R packages. α-diversity was calculated using the Shannon diversity metric, and differences over time points in groups were assessed using a linear mixed-effects (LME) model with infants set as a random effect while correcting for age. The overall resistome composition of nasopharyngeal samples was visualized using principal component analysis (PCA) ordination plots. The ordinates were based on the Euclidean dissimilarity matrix of centered log-ratio (CLR) transformed relative abundance data (i.e., Aitchison distance). Associations of clinical covariates with the dispersion of sample resistome compositions were statistically evaluated with the Permutational multivariate analysis of variance (PERMANOVA) test using the *adonis2* function (vegan package) with 999 permutations. To test the homogeneity of multivariate dispersion, the permutational analysis of multivariate dispersion (PERMDISP) test was used. In case of

heterogeneity, the analysis of similarities (ANOSIM) test using the *anosim* function (vegan package) was further used to statistically re-assess the association of variables with the overall resistome composition. In the multivariable, temporal analyses, only the variables that individually showed a significant association with resistome composition were included.

To identify the resistotypes, the Dirichlet multinomial mixture (DMM) model approach was utilized using the DirichletMultinomial (v.1.38.0) R package [50]. Procrustes analysis was performed to examine the effect of underlying microbiota on the resistome. The ARG RPKM abundance matrix and the species-level count per million (CPM) abundance matrix were Hellinger-transformed and ordinated using PCoA (Principal coordinates analysis) on Bray–Curtis dissimilarities. The symmetric Procrustes correlation coefficients between the ordinations and P-values were obtained through the *protest* function from the vegan package. Spearman's correlation was applied to resistome and microbiome relative abundance profiles for paired samples. To avoid the potential bias introduced by Spearman's rank when ranking zero values, we removed ARGs and species that were present in less than half of the samples. The correlation coefficient and corresponding adjusted p-values were calculated using the associate function from the microbiome (v.1.18.0) package [51]. Associations between individual taxa and overall resistome outcomes were calculated using repeated measure correlation analysis from the rmcorr (v.0.5.4) package while controlling for inter-individual variation [52]. Heatmaps were created using the *pheatmap* function from the pheatmap R package (v.1.0.12) [53]. For simple, independent comparisons of group differences, we used a one-way analysis of variance test, Wilcoxon rank-sum test, Kruskal-Wallis test, or chi-square test as appropriate, and we

considered $p < 0.05$ to be significant. For all analyses involving multiple comparisons, we applied the Benjamini-Hochberg (BH) method to correct for multiple testing. The statistical analysis scheme used in the study is shown in Supplementary Fig. 1.

RESULTS

Study cohort and sample characteristics

Infants exposed to early antibiotics (*n = 15*) had significantly lower GA and BW, and their mothers had a more prolonged period of ruptured membranes before delivery compared to antibiotic-naive infants (*n = 21*). None of the mothers of naive infants received antibiotics during pregnancy (prenatal antibiotics) before onset of delivery (Table 1). The average early antibiotic treatment duration was 5 days (SD: 2 days). The sampling time point T2 corresponds to the first available sample collected (except in one infant) after the cessation of early antibiotic treatment. In addition to early antibiotics, one of the included infants received subsequent antibiotics initiated > 72 hours after birth and before discharge from NICU and two other infants received antibiotic therapy between discharge from NICU and 6 months corrected age. A timeline of sample collection and antibiotic administration is represented in Fig. 1.

In total, we collected and extracted metagenomic DNA from 198 nasopharyngeal samples obtained from 36 preterm infants, from day of birth until six months corrected age. The DOL at which samples were obtained for each time point varied: T1 (DOL: 0-3), T2 (DOL: 4-9), T3 (DOL: 12-19), T4 (DOL: 20-31), T5 (DOL: 32-76) and 6 months corrected age (DOL: 229-288). Four (2%) samples were excluded after DNA extraction due to negative 16S rRNA values (qRT-

PCR) after subtracting spike-in background. The remaining 194 samples proceeded to library preparation and whole metagenomic sequencing (WMS). Detailed information on sample exclusion and inclusion statistics is described in Supplementary Table 1.

Overall Sequencing results

A total of 6.14 billion raw sequencing reads were obtained with a mean value of sample reads counts of 33.97 million (M) ranging from 4.19 to 69.25 M reads. The mean Phred quality score of raw reads for samples ranged from 32 to 37. More than 99% of the total reads generated across all samples passed the sequence quality filtering and trimming. On average 46.71% (range: 1.09 - 97.69%) of total quality-filtered reads were identified as belonging to the human genome. After removal of the human-associated metagenomic reads, a total of 2.98 billion remaining high-quality clean reads across all samples with number of reads ranging from 0.76 to 59.9 M per sample, and with median of 16.4 M reads were subjected to resistome profiling. In general, 0.31% (range: 0 - 2.8%; 9.32 million in total) of the clean reads were classified as ARGs. Results from rarefaction analysis suggested that most of the samples had sufficient sequencing depth for ARG characterization (Supplementary Fig. 2). More detailed information regarding the overall sequencing results is presented in Supplementary Table 2. Ten samples were excluded from downstream analysis as no reads were assigned to ARGs. In two infants there was more than one sample available for the defined time points (collected due to initiation or discontinuation of antibiotics more than 48 hours before/after scheduled time point). Therefore, to conduct time point-wise analysis, only one sample was considered for each of the

time points (Supplementary Table 1 & Fig. 1). Collectively, we assessed the resistome of the nasopharyngeal microbiome in a total of 181 samples collected from 36 preterm infants.

Distribution and characterization of ARGs in the preterm infant nasopharyngeal microbiome

In total, we found 373 ARGs belonging to 15 ARG classes conferring resistance via 5 distinct mechanisms. Multidrug efflux mediated ARGs comprised the majority of the resistome with an average of 27% at T1 to T5 and 17.8% at T6 (six months corrected age) of the total relative abundance. There was a median of 18 (mean: 24) unique ARGs ranging from 1 to 157 per analyzed sample. On average, maximum number of ARGs were detected after the discontinuation of treatment (T2) (mean: 40) in most infants receiving early antibiotics treatment. In the non-treated group, the corresponding mean value at T2 was 19. Only 13 ARGs were both most prevalent (detected in > 40% of all samples) and highly abundant with mean relative abundance of 64.8% (SD: 32.06%) across all preterm infant nasopharyngeal resistome. These core set of resistance genes included multidrug resistance genes (*acrB, oqxB, mexI),* macrolide-lincosamide-streptogramin (MLS) resistance genes (*mel, pmrA, rlmA(II)*), two beta-lactam resistance genes (*blaZ, SST-1*), fluoroquinolone resistance genes (*patA, patB*), tetracycline resistance genes (*tet(41), tetM*), and the aminoglycoside resistance gene *AAC(6')-Ic*.

In both antibiotic-treated and naive group, the most abundant class of ARG was multidrug (mean: 27.85%; SD: 21.90%), followed by fluoroquinolone (mean: 17.41%; SD: 20.18%), beta-lactam (mean: 15.27%; SD 16.99%), tetracycline (mean: 14.70%; SD: 10.75%), MLS (mean:

13.06%; SD: 15.29%) and aminoglycoside (mean: 8.77%; SD: 12.48%) (Fig. 2A). In terms of prevalence, aminoglycoside (76/82) and beta-lactam (75/82) were the most common class present in the samples from antibiotic-treated group, while multidrug (92/99) and beta-lactam (91/99) were the most common class in the naive infants. When classified based on resistance mechanism, antibiotic efflux (mean: 59.56%; SD: 20.97%) was the most abundant category in both the groups across all time points (Supplementary Fig. 3). At baseline (T1), the ARG composition of antibiotic-treated group was found to be more heterogenous (higher inter-individual variability) and diverse compared to naive infants (Fig. 2B).

Homology to ARGs posing major current threat for public health, including those identified by recent omics-based framework [54], were identified in nasopharyngeal samples of our cohort. Most of these potential high-risk ARGs (rank I) were detected at multiple time points across infants in both antibiotic-treated and naive group (Supplementary Fig. 4). We also observed a few high risk and clinically relevant ARGs that appeared directly after the early antibiotic treatment at T2. For instance, high-risk ARGs specific to the antibiotics encoding for aminoglycoside-modifying enzymes (*AAC(3)-II*) along with extended-spectrum TEM beta-lactamase (bla_{TEM-1}) were detected in two infants. These ARGs co-occurred with other non-targeted high-risk ARGs such as *qnrS*, *floR* and *aadA* (rank II). Previous studies have shown the co-carriage of these ARGs on multidrug resistance plasmids in Enterobacter species [55]. Also, other clinically relevant Extended Spectrum Beta-Lactamase (ESBL) genes belonging to SHV beta-lactamase (rank II) family were detected in two other infants following antibiotic exposure. In one patient, the plasmid-mediated AmpC-type β-lactamase (ACT-beta lactamase), known to

confer resistance against all classes of beta-lactam antibiotics such as penicillins, cephalosporins and carbapenems was detected only after the treatment cessation (T2). These high-risk ARGs were observed mainly in early antibiotics exposed infants whose mothers had also received antibiotics during pregnancy (prenatal). Nonetheless, they did not persist in nasopharyngeal samples collected after discharge at the corrected age of 6 months. Amongst ARGs that are found in Gram-positive pathogens, the *Staphylococcus aureus* methicillin-resistance gene i.e., *mecA* was detected in 44 samples from almost all (14/15) antibiotic-treated infants and 10 of the 21 naive infants.

Impact of early antibiotics on preterm infant nasopharyngeal resistome

Firstly, we investigated the total ARG abundance (quantified as the sum of RPKM) and ARG α-diversity (quantified by Shannon index) across sampling time points. In both the antibiotic-naive and treated groups, we found that total ARG abundance and diversity increased from T1 to T3-T4, before descending at T5 until 6 month corrected age (T6). In treated infants, total ARG abundance and diversity increased more rapidly upon antibiotic treatment, whereas the changes were generally more gradual over time points in naive infants (Fig. 3A, B). To identify the direct effects of early antibiotic treatment on the total ARG abundance and diversity, we used a generalized linear mixed model with individual set as random effect while correcting for age (DOL). Early antibiotics were significantly associated with increase in both total ARG abundance (LME model: $p.adj = 0.008$) and Shannon diversity ($p.adj = 0.008$) after treatment (T2) as compared to baseline (T1) samples in preterm infant nasopharyngeal resistome. Moreover, the significant effect of antibiotics on increase in total ARG abundance was observed

until T3 (*p.adj* = 0.001). Nonetheless, such effects were short-lived, as we did not identify significant differences at any of the following timepoints compared to baseline (T1) in antibiotic-treated group. In naive group of infants, no significant differences appeared in total ARG abundance and α-diversity between any timepoints during the first 6 months of life while correcting for age.

β-diversity analysis revealed that the overall resistome composition did significantly differ between the two groups (PERMANOVA: R^2 = 8.2%, *p.adj* = 0.02) at baseline (T1). Infants that received early antibiotic treatment had more dissimilar (dispersed) resistome composition than the naive infants (Fig. 4A). Further, we observed that nasopharyngeal samples were clustering more effectively according to birth weight group (R^2 = 16.4%, *p.adj* = 0.012; Supplementary Fig. 5A). No significant difference in baseline resistome composition between antibiotic-treated and naive infants (R^2 = 0.03, *p.adj* = 0.319) was found while correcting for birth weight group. However, a minor yet significant compositional difference in resistome was still observed between the two groups (R^2 = 5.9%, *p.adj* = 0.03) directly after the cessation of early antibiotic treatment (T2) (Fig. 4B). Moreover, the compositional changes due to early antibiotics at T2 were more discernible in infants whose mothers had received prenatal antibiotics during pregnancy (prenatal + early (*n* = *8*) vs naive (*n* = *20*): R^2 = 10.09%, *p.adj* = 0.006; Supplementary Fig. 5B). Following treatment, resistome normalized rapidly as no significant differences in composition between naive and treated group were detected from T3 until 6 months of corrected age (T6) (Fig. 4C).

Clinical covariates associated with preterm infant nasal resistome composition development

The associations between resistome composition and clinical factors were analyzed using PERMANOVA-tests (all *p.adj* ≤ 0.001) and principal component analysis (PCA) across all nasopharyngeal samples from preterm birth until 6 months corrected age, using centered log-ratio (CLR) transformed relative abundance data (Supplementary Fig. 1). The PCA was explained by two main principal components, describing 18.5% and 17.7% of the variation, respectively. Besides inter-individual variation (PERMANOVA (univariate): R^2 = 37.7%, *p.adj* = 0.001), age was the other significant covariate found to be associated with overall resistome composition in the overarching cohort. A moderate yet statistically significant effect was observed with GA (R^2 4%, *p.adj* = 0.001) and BW group (R^2 4.1%, *p.adj* = 0.001), which are highly correlated variables (as indicated by chi-squared test: p < 2.2e-16). Mode of delivery, gender, and intrapartum antibiotic treatment did not appear to shape the overall nasopharyngeal resistome composition (Supplementary Table 3). Covariates that differed between the antibiotic-treated and naive infants at baseline and that also had a significant association with the overall resistome composition (GA and prenatal antibiotics (R^2 4.5%, *p.adj* = 0.001)) were used for correction in the subsequent multivariate analysis.

Multivariate PERMANOVA analysis showed that the largest effect on overall resistome composition remained attributed to inter-individual variation (R^2 30.03%), followed by PMA (R^2 9.7%) and chronological age group (sampling time points) (R^2 8.4%, all *p.adj* = 0.001). Although we observed a high degree of variation between samples with no obvious discrete

spatial clustering in relation to these variables upon visual inspection (Fig. 5A-C). We found no significant influence of early antibiotics (R^2 = 0.6%, *p.adj* = 0.31) and its duration (R^2 = 1.2%, *p.adj* = 0.009) after adjusting for confounding variables. However, we observed a significant effect of prenatal antibiotics (R^2 = 3.6%, *p.adj* = 0.001) and its duration (categorized) (R^2 = 4.4%, *p.adj* = 0.001) in shaping the overall nasopharyngeal resistome composition in preterm birth. To account for the likelihood of prenatal antibiotics with early antibiotics exposure in our analyses (Chi-squared test, $p < 0.0003$), we further categorized our early antibiotic-treated group based on whether the mothers had also received antibiotics during pregnancy (prenatal + early, $n = 9$) or not (only early, $n = 6$). Post-hoc analysis revealed significant differences in the overall resistome composition of infants having exposure to both prenatal and early life antibiotics (prenatal + early) compared to infants receiving either only early or naive (no prenatal + no early) (*p.adj* = 0.001; Fig. 5D). No difference in overall resistome composition was observed between the only early antibiotics and naive group (*p.adj* = 0.14), underlining the relevance of prenatal antibiotic exposure in shaping the overall resistome composition of preterm infants during the first six months of life (corrected).

Unsupervised clustering classifies nasopharyngeal samples into distinct clusters based on composition differences

Next, we determined whether the nasopharyngeal samples were organized into clusters (i.e., resistotypes) according to their ARG abundance profiles using an unsupervised method, i.e., DMM models. This way, we identified three distinct clusters (or resistotypes) that were driven by specific ARGs. Majority of our samples (~90%) were classified into two main resistotypes,

R1 and R2. No significant difference in the sequencing depth, was detected between the two Resistotypes (pairwise Wilcox test: $p = 0.14$), suggesting that sequencing depth was not driving their separation. The first resistotype (R1) was characterized by a predominance of fluoroquinolone resistance genes (*patA*, *patB*) and *rlmA(II)* (Supplementary Fig. 6). On the other hand, resistotype R2 was driven mainly by relatively higher abundance of the beta-lactam resistance gene *SST-1* and the tetracycline resistance gene *tet(41)*. While resistotype R3 was enriched with the blaZ β-lactamase resistance gene. The identified resistotypes did vary in their overall ARG composition (β-diversity), as shown in a PCA plot (Fig. 6A). Further, PERMANOVA analysis showed that resistotypes were more strongly (larger effect size) and significantly associated with the resistome than other clinical covariates or early-life factors in preterm infants during the first 6 months of life ($R^2 = 19.05\%$, $p.adj = 0.001$).

Nasopharyngeal resistome composition of preterm infants is linked to microbial composition

To investigate the impact of microbial composition on resistome composition, we taxonomically profiled the infant nasopharyngeal samples using MetaPhlAn3 (manuscript under preparation, Rajar et al. 2023). We noticed that microbial community types, as determined by DMM models, were highly linked with the resistotypes (chi-square test: $p < 0.0001$). Resistotype R2 was very specific to the *Serratia* community type. Resistotype R1 was associated with the community types mainly driven by *Streptococcus* and *Gemella* and Resistotype R3 was enriched in community types predominated by *Staphylococcus* and *Cutibacterium* (Fig. 6B). These community types were also significantly associated to the overall composition of the resistome

(PERMANOVA: R^2 = 13.01%, *p.adj* = 0.001; Fig. 6C). Additionally, result from Procrustes analysis showed a strong and significant association between the microbial taxa abundance and ARG abundance profiles (correlation coefficient (r) = 0.85, p = 0.001, m^2 = 0.26), underlining that bacterial community composition structured the resistome composition in the nasopharyngeal microbiome of preterm infants. However, the correlation was more robust in the antibiotic-naive infants (r = 0.94, p = 0.001, m^2 = 0.11) as compared to antibiotic-treated infants (r = 0.85, p = 0.001, m^2 = 0.27; Fig. 7A, B).

Next, we performed pairwise Spearman's correlation analysis between ARG and microbial species abundances to predict the origin or potential microbial host of ARGs in nasopharyngeal samples. As expected, the strongest positive correlation ($r \geq 0.8$, *p.adj* < 0.05) was found between the Resistotype 2 enriched ARGs i.e., *SST-1*, *AAC(6')-Ic*, *tet(41)*, *MexI* and *Serratia marcescens/nematodiphila*. The highly prevalent and Resistotype R1 associated ARGs like *patA, patB and rlmA(II)* were found to have strong co-occurrence with *Streptococcus mitis/oralis* and *Gemella haemolysans/sanguinis* (Fig. 7C). While *Staphylococcus aureus* and *epidermidis* correlated only with *blaZ* beta-lactam ARG. Interestingly, these associations are in accordance with information present in the literature and AMR gene databases on the known microbial host of these ARGs, highlighting the potential of such analysis approach to predict the origin of ARG in metagenomes. However, this analysis is only applicable on the features (ARGs or taxa) that are present in high prevalence across samples to reduce the bias due to false ranking of features with many zero values.

Lastly, we also examined whether some of these strongly associated bacterial species with ARGs would also directly correlate with our overall resistome outcomes. We plotted the relative abundance of each species versus the total ARG abundance (sum of RPKM) and the number of unique ARGs from all samples. Our analysis revealed that only the relative abundance of *Streptococcus mitis* (rmcorr correlation coefficient (r^{rm} = 0.29, *p.adj* = 0.01) showed a moderately linear association with the total ARG abundance while adjusting for inter-individual variability (Supplementary Fig. 7). However, we did not observe a significant correlation between the relative abundance of *S. mitis* and the number of unique ARGs (r^{rm} = 0.19, *p.adj* = 0.12). To statistically assess whether *S. mitits* relative abundance explains the dispersion of resistomes across samples, we conducted PERMANOVA test and found that *Streptococcus mitis* significantly explains 11.69% (*p.adj* = 0.001) of the variation in the overall resistome composition. Only *Serratia marcescens* (R^2 = 11.8%, *p.adj* = 0.001) can describe rather similar variation than *Streptococcus mitis* in the resistome, though it was not found to be positively associated with any overall resistome outcomes in our data.

DISCUSSION

In this study, we (1) examined the dynamics of nasopharynx resistome development in preterm infants, (2) assessed the impact of early antibiotics on its trajectory, and (3) explored the potential associations between clinical variables and the overall resistome composition in preterm infants. We detected ARGs in nearly all (~95%) samples from both the antibiotic-treated and naive groups, including some potentially high-risk ARGs. Our analysis revealed that early antibiotic exposure was transiently associated with an increase in both the diversity and total abundance of ARGs, as well as changes in the overall resistome composition. The resistome

composition was strongly correlated with the microbiome composition, with inter-individual variation and age explaining most of the compositional variation in the resistomes. Among the clinical variables, prenatal maternal antibiotic use, which was a subgroup found only in the antibiotic group of preterm infants, appeared to be one of the most significant factors influencing the overall nasopharyngeal resistome composition in preterm infants.

The nasopharynx represents a highly accessible microbial community and also serves as a crucial diagnostic window in combating respiratory infections and AMR [56]. It is also the primary ecological niche for common respiratory pathogens [6]. To date, one of the closest attempts for an in-depth investigation of ARGs in the respiratory tract using a culture-free approach identified ARGs in 86% of the samples. However, this was based on the microbiomes of the anterior nares and from an adult population (Human Microbiome Project) [57], which are different from the groups and collection sample sites used in our study. In a recent study focusing on infants, ARGs were detected in 64% of nasopharynx samples [58]. Samples in this study were enriched for streptococci, thus precluding general conclusions on the extension of ARG presence in the natural microbiome. In our study, ARGs were present in 95% of the preterm infant samples from the nasopharynx, regardless of whether they were exposed to antibiotics or not. Although our study differs in several aspects from the ones above, the combined findings indicate that the respiratory tract is likely an important reservoir of ARGs and an integral component of human respiratory microbiomes. For preterm infants, the detection of ARGs in the nasopharynx during the first few hours of life might occur due to microbiome colonization from multiple sources such as transgenerational transfer from mothers and the environment [31, 59]. In addition,

preterm infants usually have longer stay in the hospital than term infants after birth. This most likely contributes to an increased risk for respiratory tract colonization with drug resistant microorganisms, as recently reported in relation to the gut resistome [9].

We found that ARGs conferring multidrug resistance via efflux pumps encompassed a large proportion of the nasopharyngeal resistome in the preterm infants. Similar findings have been previously reported in the gut and oral resistome of preterm and term infants by other metagenomic studies [10, 59-61]. In addition, we detected many of the same aminoglycoside-modifying enzymes, tetracycline protection proteins, and beta-lactamases ARGs identified previously in the nasopharynx of South African infants, despite the fact that the samples in the South African study were culture-enriched for streptococci and were from a country with higher AMR challenges than the origin of samples used in the present study. Such similarities in resistome composition may, at least in part, be explained by the large proportion of streptococci found in the natural microbial communities of the infant nasopharynx. Together, these findings suggest the existence of a core infant nasopharyngeal resistome.

Staphylococcus aureus is one of the leading pathogens causing neonatal sepsis, and one of the priority drug resistant pathogens by the WHO [62]. We found a strong correlation between the abundance of *Staphylococcus aureus/epidermidis* and the PC1-beta-lactamase (*blaZ*) gene. This is a gene found in the chromosome and mobile plasmids of staphylococci, thus representing high risk ARGs. Another strong correlation was found between *Serratia* species and specific resistance genes such as *SST-1*, which confers resistance against cephalosporins. Other resistance genes correlated with this group included aminoglycoside and tetracycline resistance genes. This is of relevance, since *Serratia spp.* are associated with hospital infection outbreaks, especially in

NICUs [63]. However, it is uncertain whether the identified *Serratia spp.* in the study were acquired from the hospital environment or not, as this group of bacteria has been also found as part of the nasal microbiome in healthy adults [64].

Our results showed a significant effect of early broad-spectrum antibiotic regimen (ampicillin + gentamicin) on the total abundance, diversity, and composition of ARGs in the nasopharynx of preterm infants. These effects on the resistome were short-lived and only observed directly after the cessation of antibiotic treatment (T2). In a recent study investigating the impact of antibiotics used in the first week of life on intestinal resistome, Reyman et al. reported that three other antibiotic combinations (penicillin + gentamicin, co-amoxiclav + gentamicin, and amoxicillin + cefotaxime) also had a transient impact on the infant resistome [11]. As the effects of antibiotics on microbiome and resistome composition are unique to each individual and dependent on the antibiotics used and the body site, caution should be taken when extending the interpretation of such general findings. Interestingly, the transient effects of ampicillin + gentamicin on the nasopharyngeal resistome were more discernible in preterm infants whose mothers were exposed to antibiotics during pregnancy before delivery (prenatal antibiotics). Moreover, we also observed a more profound effect of prenatal antibiotic exposure on the nasopharyngeal resistome of preterm infants at T2 (after cessation of early antibiotics), which persisted until the last sampling time point before discharge from the NICU (T4-T5), but not at 6 months corrected age. Although very few studies have focused on the association between prenatal antibiotic exposure and the resistome, a recent study based on the gut indicates a possible correlation of prenatal antibiotics with increased abundance of ARGs, at least when used 5 to 6 weeks before birth [10]. Additionally, in concordance with other studies on the infant gut [15, 59], our data showed no

significant impact of intrapartum antibiotic therapy on the overall nasopharyngeal resistome composition. The nasopharyngeal samples of preterm infants at 6 months corrected age were still clustered according to the birth weight group (similar to T1, Supplementary Fig. 5C), suggesting a strong but transient influence of other environmental factors, such as antibiotic exposures. Nonetheless, there is a paucity of relevant studies on nasopharynx that can corroborate our findings, as they are mainly restricted to the infant gut resistome [9, 11, 14, 15, 31, 59].

One of the strengths of this study is that it presents the most comprehensive information to date on the characteristics and dynamics of resistome development in the nasopharynx, including results for the first six months of life (corrected) – a critical period of airway development in infants. In addition, the naive group was enrolled at the same time as our treated group from the same hospital, thus removing potential confounding differences in enrollment years or sampling duration. On the other hand, the small number of samples at each time point and inter-individual response to antibiotic exposure may be insufficient to unravel significant associations between antibiotic regimen and resistome development. Moreover, the baseline samples for the antibiotic-treated group ideally would represent samples taken before treatment. However, antibiotic treatment initiation for sEONS and clinical stabilization of the preterm infants were a high priority that needed to precede nasopharynx sample collection. To partially address this issue, we included only preterm infants with samples collected no more than 24 hours after antibiotic initiation. Even so, this choice may have resulted in the selection of more clinically stable infants, which could possibly limit the generalizability of our results. Despite this approach, the differences in resistome changes between baseline and after treatment found in our study may underestimate the true impact of antibiotics on the resistome. We rigorously adjusted for such

confounding factors and inter-individual variation using multivariable methods or matched analysis to ensure the reliability of our observations as a consequence of early antibiotics. Nevertheless, we cannot isolate early antibiotic effects from other adverse pre-life events that coincide with premature birth, such as prenatal and antepartum maternally administered antibiotics. Furthermore, for low biomass samples such as those used in our study, sensitive approaches such as deep sequencing are necessary to obtain meaningful microbiome information. This increases the risk of detecting potential DNA contaminants. Although it is not possible to completely eliminate the potential identification of contaminants, we adopted various control procedures, starting from sample collection to analysis, to ensure that the data presented in our study is a reliable representation of the ARGs found in the nasopharynx.

In conclusion, our findings suggest that early-life treatment with broad-spectrum antibiotics for sEONS can cause acute perturbations in the nasopharyngeal resistome of preterm infants. However, these perturbations were transient, as we did not observe a significant alteration in diversity, abundance, composition, or carriage persistence of high-risk ARGs in treated preterm infants compared to controls at 6 months of corrected age. The individual variability in the uneven distribution of high-risk ARGs and the changes in the resistome following antibiotic treatment highlight the potential benefits of developing resistome profiling tools for point-of-care use, which could aid in informing and tailoring antibiotic treatment decisions [65]. Future multicenter studies with long-term follow-up are needed to assess the ecological side-effects of different antibiotic regimens, including alternatives to broad-spectrum antimicrobial therapy on infant respiratory resistome development. This study also emphasizes the need to further

investigate the underappreciated effect of different antibiotic exposures (pre, peri, post-birth) during the critical period of the nasopharyngeal microbiome and resistome development in infants, and the subsequent development of respiratory infections later in life. Understanding the development of the respiratory resistome and the impact of antibiotic therapies is relevant for striking a balance between the conflicting goals of minimizing the risk of serious infectious complications to preterm birth on one side, and minimizing the selection of organisms with ARGs on the other. Such knowledge can also help to strengthen antibiotic stewardship programs. In general, the disproportionate contribution of respiratory pathogens to drug-resistant deaths at all ages highlights the importance of future studies focusing on this yet rather unexplored reservoir of pathogens and ARGs.

REFERENCES

1. Murray, C., et al., *& Naghavi, M.(2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis.* The Lancet. **399**(10325): p. 629-655.
2. Tedijanto, C., et al., *Estimating the proportion of bystander selection for antibiotic resistance among potentially pathogenic bacterial flora.* Proceedings of the National Academy of Sciences, 2018. **115**(51): p. E11988-E11995.
3. Von Wintersdorff, C.J., et al., *Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer.* Frontiers in microbiology, 2016. **7**: p. 173.
4. Kilian, M., et al., *Parallel evolution of Streptococcus pneumoniae and Streptococcus mitis to pathogenic and mutualistic lifestyles.* MBio, 2014. **5**(4): p. e01490-14.
5. Pailhoriès, H., et al., *Antibiotic resistance in chronic respiratory diseases: from susceptibility testing to the resistome.* European Respiratory Review, 2022. **31**(164).
6. Man, W.H., W.A. de Steenhuijsen Piters, and D. Bogaert, *The microbiota of the respiratory tract: gatekeeper to respiratory health.* Nature Reviews Microbiology, 2017. **15**(5): p. 259-270.
7. Huang, Y.J., *Nasopharyngeal microbiota: gatekeepers or fortune tellers of susceptibility to respiratory tract infections?* 2017, American Thoracic Society. p. 1504-1505.
8. Dhariwal, A., et al., *Differential response to prolonged amoxicillin treatment: long-term resilience of the microbiome versus long-lasting perturbations in the gut resistome.* Gut Microbes, 2023. **15**(1): p. 2157200.
9. Gasparrini, A.J., et al., *Persistent metagenomic signatures of early-life hospitalization and antibiotic treatment in the infant gut microbiota and resistome.* Nature microbiology, 2019. **4**(12): p. 2285-2297.
10. Li, X., et al., *The infant gut resistome associates with E. coli, environmental exposures, gut microbiome maturity, and asthma-associated bacterial composition.* Cell Host & Microbe, 2021. **29**(6): p. 975-987. e4.
11. Reyman, M., et al., *Effects of early-life antibiotics on the developing infant gut microbiome and resistome: a randomized trial.* Nat Commun, 2022. **13**(1): p. 893.
12. Thänert, R., et al., *The resistance within: Antibiotic disruption of the gut microbiome and resistome dynamics in infancy.* Cell Host & Microbe, 2022. **30**(5): p. 675-683.

668 13. Gasparrini, A.J., et al., *Antibiotic perturbation of the preterm infant gut microbiome and*
669 *resistome.* Gut microbes, 2016. **7**(5): p. 443-449.
670 14. Gibson, M.K., et al., *Developmental dynamics of the preterm infant gut microbiota and*
671 *antibiotic resistome.* Nat Microbiol, 2016. **1**: p. 16024.
672 15. Lebeaux, R.M., et al., *The infant gut resistome is associated with E. coli and early-life*
673 *exposures.* BMC Microbiol, 2021. **21**(1): p. 201.
674 16. Tripathi, N., C.M. Cotten, and P.B. Smith, *Antibiotic use and misuse in the neonatal*
675 *intensive care unit.* Clinics in perinatology, 2012. **39**(1): p. 61-68.
676 17. Mukhopadhyay, S., E.C. Eichenwald, and K.M. Puopolo, *Neonatal early-onset sepsis*
677 *evaluations among well-appearing infants: projected impact of changes in CDC GBS*
678 *guidelines.* Journal of perinatology, 2013. **33**(3): p. 198-205.
679 18. Rajar, P., et al., *Antibiotic stewardship in premature infants: a systematic review.*
680 Neonatology, 2020. **117**(6): p. 673-686.
681 19. Kuppala, V.S., et al., *Prolonged initial empirical antibiotic treatment is associated with*
682 *adverse outcomes in premature infants.* The Journal of pediatrics, 2011. **159**(5): p. 720-
683 725.
684 20. Novitsky, A., et al., *Prolonged early antibiotic use and bronchopulmonary dysplasia in*
685 *very low birth weight infants.* American journal of perinatology, 2015. **32**(01): p. 043-
686 048.
687 21. Ting, J.Y., et al., *Association between antibiotic use and neonatal mortality and*
688 *morbidities in very low-birth-weight infants without culture-proven sepsis or necrotizing*
689 *enterocolitis.* JAMA pediatrics, 2016. **170**(12): p. 1181-1187.
690 22. Ting, J.Y., et al., *Duration of initial empirical antibiotic therapy and outcomes in very low*
691 *birth weight infants.* Pediatrics, 2019. **143**(3).
692 23. Cotten, C.M., et al., *Prolonged duration of initial empirical antibiotic treatment is*
693 *associated with increased rates of necrotizing enterocolitis and death for extremely low*
694 *birth weight infants.* Pediatrics, 2009. **123**(1): p. 58-66.
695 24. Esaiassen, E., et al., *Antibiotic exposure in neonates and early adverse outcomes: a*
696 *systematic review and meta-analysis.* Journal of antimicrobial chemotherapy, 2017.
697 **72**(7): p. 1858-1870.

25. Li, Y., et al., *Early use of antibiotics is associated with a lower incidence of necrotizing enterocolitis in preterm, very low birth weight infants: the NEOMUNE-NeoNutriNet cohort study.* The Journal of pediatrics, 2020. **227**: p. 128-134. e2.
26. Vatne, A., et al., *Early empirical antibiotics and adverse clinical outcomes in infants born very preterm: a population-based cohort.* The Journal of Pediatrics, 2023. **253**: p. 107-114. e5.
27. Tamburini, S., et al., *The microbiome in early life: implications for health outcomes.* Nature medicine, 2016. **22**(7): p. 713-722.
28. Bailey, L.C., et al., *Association of antibiotics in infancy with early childhood obesity.* JAMA pediatrics, 2014. **168**(11): p. 1063-1069.
29. Arrieta, M.-C., et al., *Early infancy microbial and metabolic alterations affect risk of childhood asthma.* Science translational medicine, 2015. **7**(307): p. 307ra152-307ra152.
30. Pärnänen, K., et al., *Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements.* Nature Communications, 2018. **9**(1): p. 3891.
31. Yassour, M., et al., *Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability.* Science translational medicine, 2016. **8**(343): p. 343ra81-343ra81.
32. Kang, H.M. and J.H. Kang, *Effects of nasopharyngeal microbiota in respiratory infections and allergies.* Clinical and Experimental Pediatrics, 2021. **64**(11): p. 543.
33. Teo, S.M., et al., *The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development.* Cell host & microbe, 2015. **17**(5): p. 704-715.
34. Toivonen, L., et al., *Antibiotic treatments during infancy, changes in nasal microbiota, and asthma development: population-based cohort study.* Clinical Infectious Diseases, 2021. **72**(9): p. 1546-1554.
35. Rajar, P., et al., *Microbial DNA extraction of high-host content and low biomass samples: Optimized protocol for nasopharynx metagenomic studies.* Frontiers in Microbiology, 2022. **13**.
36. Eisenhofer, R., et al., *Contamination in low microbial biomass microbiome studies: issues and recommendations.* Trends in microbiology, 2019. **27**(2): p. 105-117.

37. Krueger, F., *Trim Galore: a wrapper tool around Cutadapt and FastQC.* Trim Galore, 2012.
38. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nature methods, 2012. **9**(4): p. 357-359.
39. Li, H., et al., *The sequence alignment/map format and SAMtools.* bioinformatics, 2009. **25**(16): p. 2078-2079.
40. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-842.
41. Andrews, S., *FastQC: a quality control tool for high throughput sequence data.* 2010, Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
42. Ewels, P., et al., *MultiQC: summarize analysis results for multiple tools and samples in a single report.* Bioinformatics, 2016. **32**(19): p. 3047-3048.
43. Alcock, B.P., et al., *CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database.* Nucleic acids research, 2020. **48**(D1): p. D517-D525.
44. Doster, E., et al., *MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data.* Nucleic acids research, 2020. **48**(D1): p. D561-D569.
45. Allaire, J., *RStudio: integrated development environment for R.* Boston, MA, 2012. **770**(394): p. 165-171.
46. R Core Team, R. and R.C. Team, *R: a language and environment for statistical computing. R Foundation for Statistical Computing; 2020.* 2021.
47. McMurdie, P.J. and S. Holmes, *phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data.* PloS one, 2013. **8**(4): p. e61217.
48. Wickham, H., *ggplot2.* Wiley interdisciplinary reviews: computational statistics, 2011. **3**(2): p. 180-185.
49. Oksanen, J., et al., *Vegan: Community Ecology Package (R Package Version 2.6-2). 2022.*
50. Holmes, I., K. Harris, and C. Quince, *Dirichlet multinomial mixtures: generative models for microbial metagenomics.* PloS one, 2012. **7**(2): p. e30126.
51. Lahti, L. and S. Shetty, *Introduction to the microbiome R package.* 2018.
52. Bakdash, J.Z. and L.R. Marusich, *Repeated measures correlation.* Frontiers in psychology, 2017. **8**: p. 456.

53. Kolde, R. and M.R. Kolde, *Package 'pheatmap'.* R package, 2018. **1**(10).
54. Zhang, A.-N., et al., *An omics-based framework for assessing the health risk of antimicrobial resistance genes.* Nature communications, 2021. **12**(1): p. 4765.
55. Long, Y., et al., *High Carriage Rate of the Multiple Resistant Plasmids Harboring Quinolone Resistance Genes in Enterobacter spp. Isolated from Healthy Individuals.* Antibiotics, 2022. **11**(1): p. 15.
56. Flynn, M. and J. Dooley, *The microbiome of the nasopharynx.* Journal of Medical Microbiology, 2021. **70**(6).
57. Consortium, H.M.P., *Structure, function and diversity of the healthy human microbiome.* Nature, 2012. **486**(7402): p. 207-214.
58. Manenzhe, R.I., et al., *Longitudinal changes in the nasopharyngeal resistome of South African infants using shotgun metagenomic sequencing.* PloS one, 2020. **15**(4): p. e0231887.
59. Parnanen, K., et al., *Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements.* Nat Commun, 2018. **9**(1): p. 3891.
60. Rahman, S.F., et al., *Machine learning leveraging genomes from metagenomes identifies influential antibiotic resistance genes in the infant gut microbiome.* MSystems, 2018. **3**(1): p. e00123-17.
61. Sukumar, S., et al., *The development of the oral resistome during the first decade of life.* 2022.
62. Tacconelli, E., *Global Priority List of Antibiotic-Resistant Bacteria to Guide Research, Discovery, and Development.* 2017.
63. Acar, J.F., *Serratia marcescens infections.* Infection Control & Hospital Epidemiology, 1986. **7**(5): p. 273-280.
64. Liu, C.M., et al., *Staphylococcus aureus and the ecology of the nasal microbiome.* Sci Adv, 2015. **1**(5): p. e1400216.
65. Leggett, R.M., et al., *Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens.* Nature Microbiology, 2020. **5**(3): p. 430-442.

ACKNOWLEDGEMENTS

We would like to thank the parents of participating infants, and the clinical staff at the Neonatal Intensive Care units at Ullevål hospital for their assistance with sample collection. The sequencing service for this work was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no) and the computations were performed on resources provided by Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway.

FUNDING

This work was financed by the Norwegian Research Council (NFR) (project number: 273833), the Olav Thon Foundation (project number: 421258), and the Faculty of Dentistry at the University of Oslo.

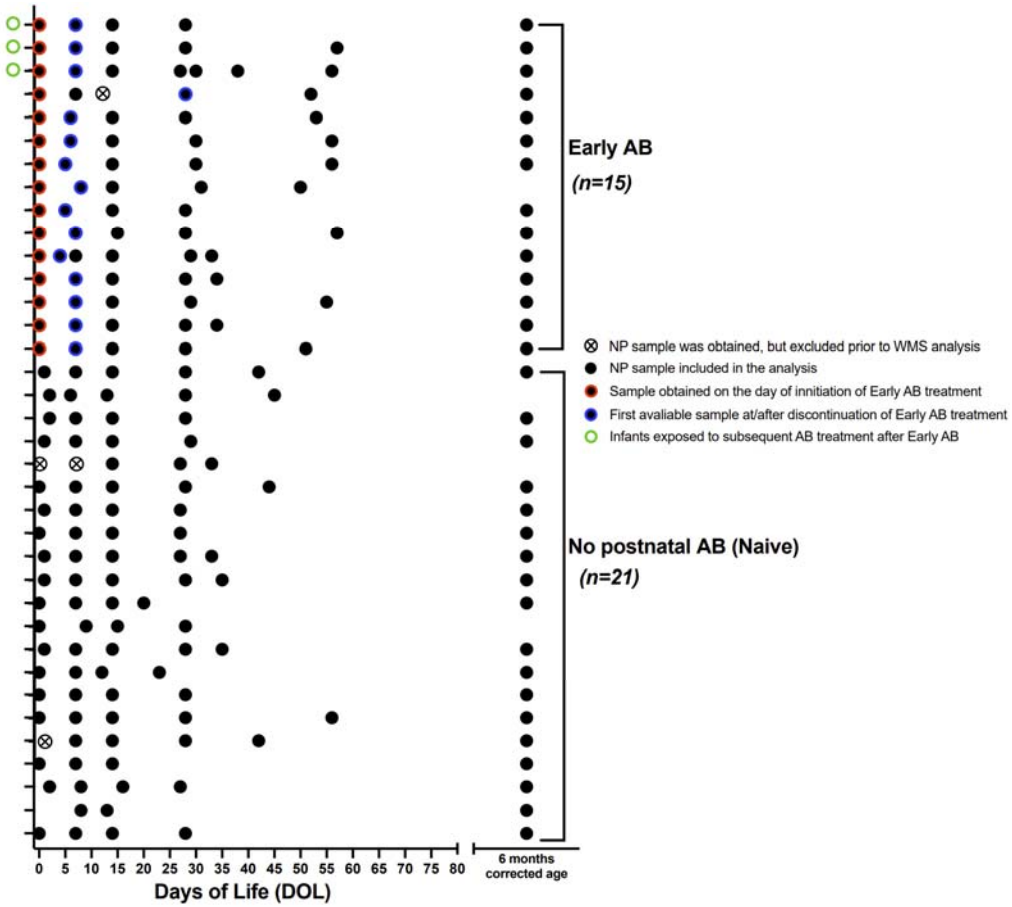
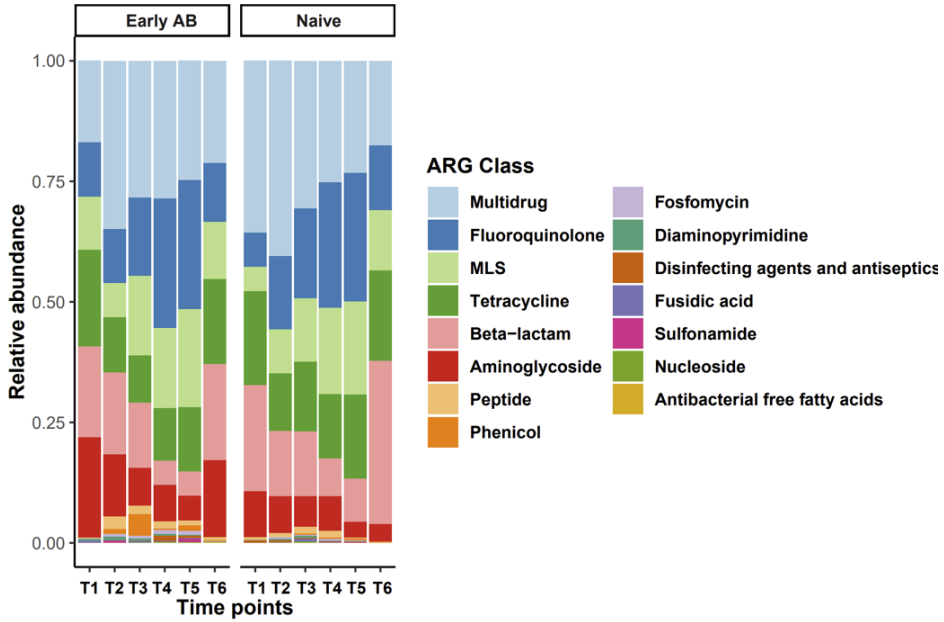


Fig. 1: Sample collection and antibiotic exposure for preterm infants in this study. Sampling timeline of 198 nasopharyngeal samples collected from 36 preterm infants over first six months (corrected age), together with information regarding the time of early antibiotic exposure (red and blue colored circles), exposure to subsequent antibiotic treatment (green colored circles) and included/excluded samples. The x-axis represents the days of life (i.e., age) and the y-axis represents the infants. NP = nasopharyngeal; AB = antibiotics.

(A)


Early AB
Naive
Relative abundance
1.00
0.75
0.50
0.25
0.00
T1 T2 T3 T4 T5 T6
T1 T2 T3 T4 T5 T6
Time points
ARG Class
Multidrug
Fluoroquinolone
MLS
Tetracycline
Beta−lactam
Aminoglycoside
Peptide
Phenicol
Fosfomycin
Diaminopyrimidine
Disinfecting agents and antiseptics
Fusidic acid
Sulfonamide
Nucleoside
Antibacterial free fatty acids

(B)

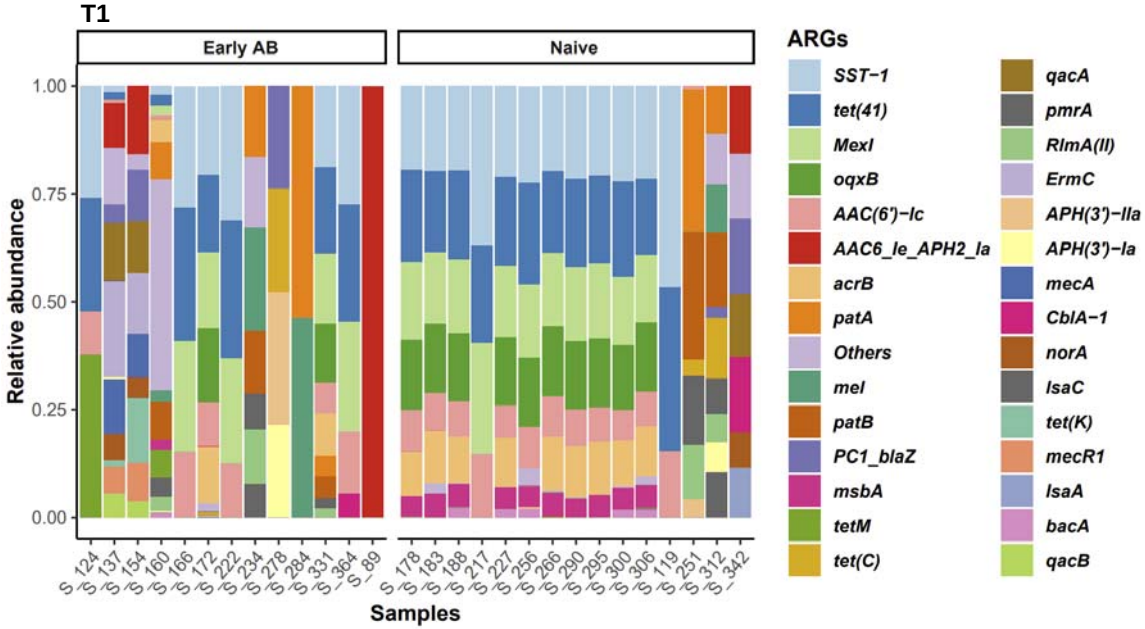
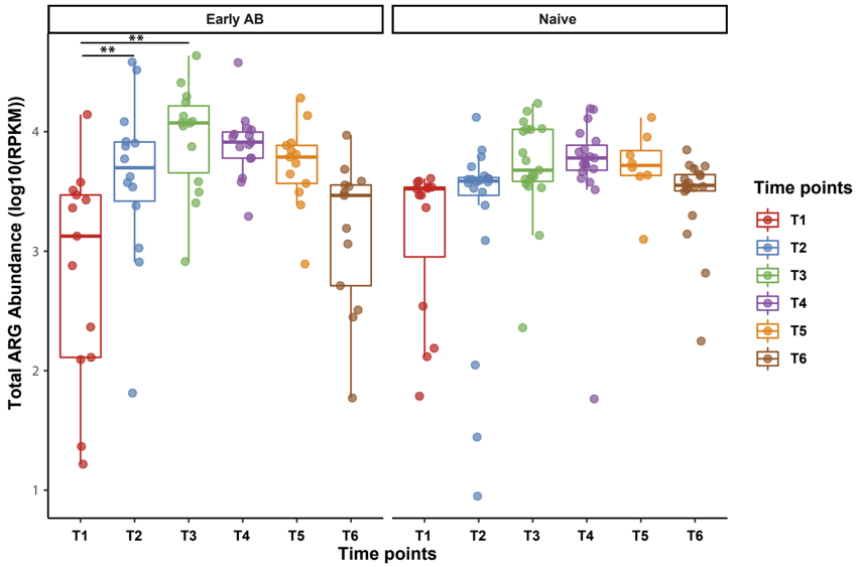

T1
Early AB
Naive
Relative abundance
1.00
0.75
0.50
0.25
0.00
S_124
S_137
S_154
S_160
S_166
S_172
S_222
S_234
S_278
S_284
S_331
S_364
S_89
S_178
S_183
S_188
S_217
S_227
S_256
S_266
S_290
S_295
S_300
S_306
S_119
S_251
S_312
S_342
Samples
ARGs
SST−1
tet(41)
MexI
oqxB
AAC(6')−Ic
AAC6_Ie_APH2_Ia
acrB
patA
Others
mel
patB
PC1_blaZ
msbA
tetM
tet(C)
qacA
pmrA
RlmA(II)
ErmC
APH(3')−IIa
APH(3')−Ia
mecA
CblA−1
norA
lsaC
tet(K)
mecR1
lsaA
bacA
qacB

Fig. 2: Resistome composition in the nasopharynx microbiome of preterm infants. Stacked bar plot showing the relative abundance of (A) antibiotic resistance genes (ARGs) clustered at the Class level for the naive and antibiotic-treated groups over time; (B) top 30 most abundant ARGs identified across samples at baseline (T1) in early antibiotics treated (left) and naive (right) groups. All other ARGs are grouped into the "Other" category. MLS = macrolide-lincosamide-streptogramin.

(A)


Early AB
Naive
**
**
Total ARG Abundance (log10(RPKM))
4
3
2
1
T1
T2
T3
T4
T5
T6
T1
T2
T3
T4
T5
T6
Time points
Time points
T1
T2
T3
T4
T5
T6

(B)

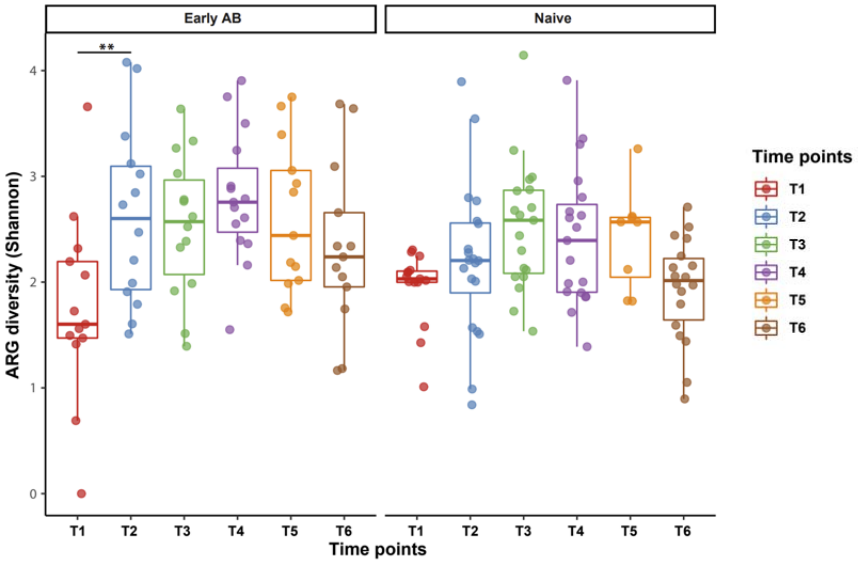
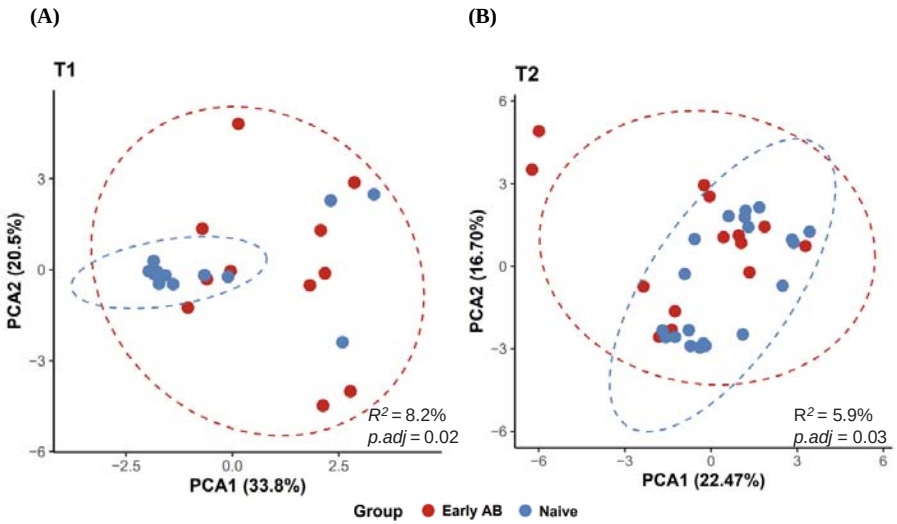

Early AB
Naive
**
ARG diversity (Shannon)
4
3
2
1
0
T1
T2
T3
T4
T5
T6
T1
T2
T3
T4
T5
T6
Time points
Time points
T1
T2
T3
T4
T5
T6

Fig. 3: Impact of early antibiotic treatment on nasopharynx total ARG abundance and ARG α-diversity Boxplots showing the **(**A) total ARG abundance (sum of RPKM) and (B) ARG Shannon diversity per sample, stratified by antibiotic treatment and time points. The lower and upper horizontal box lines correspond to the first and third quartiles, with the line in the middle of the boxes representing the median. Whiskers extending from the boxes represent the range of data points within 1.5 times the interquartile range below the first quartile and within 1.5 times the interquartile range above the third quartile. (Statistical significance: LME mixed effect model. Adjusted p-values (p): *p < .05, **p < .01 and ***p < .001). RPKM = reads per kilobase of reference gene per million bacterial reads.



(C)

Time points	R^2 (%)	p.adj
T3	3.4%	0.326
T4	4.4%	0.129
T5	4.99%	0.387
T6	3.58%	0.349

Fig. 4: β-diversity PCA plots of overall nasopharyngeal resistome composition stratified per time point. Principal component analysis (PCA) based on Euclidean distances of CLR-transformed ARG abundances (*i.e.*, Aitchison dissimilarity) between samples, visualizing the overall nasopharyngeal resistome composition stratified for antibiotic-treated infants and naive at time point (A) T1 (baseline) and (B) T2 (immediately after cessation of early antibiotic therapy). Each data point represents the resistome composition of one sample. The ellipses represent the 95% confidence interval around the centroid of resistome composition for each group. Statistical significance (adjusted p-values) and effect size (R^2) of the differences in beta diversity were assessed using PERMANOVA. In panel (C), the results of the PERMANOVA-tests for the later

time points are presented. Data points of early antibiotic-treated infants are colored red and those of naive infants are blue. PERMANOVA = permutational multivariate analysis of variance; p.adj = adjusted p-values.

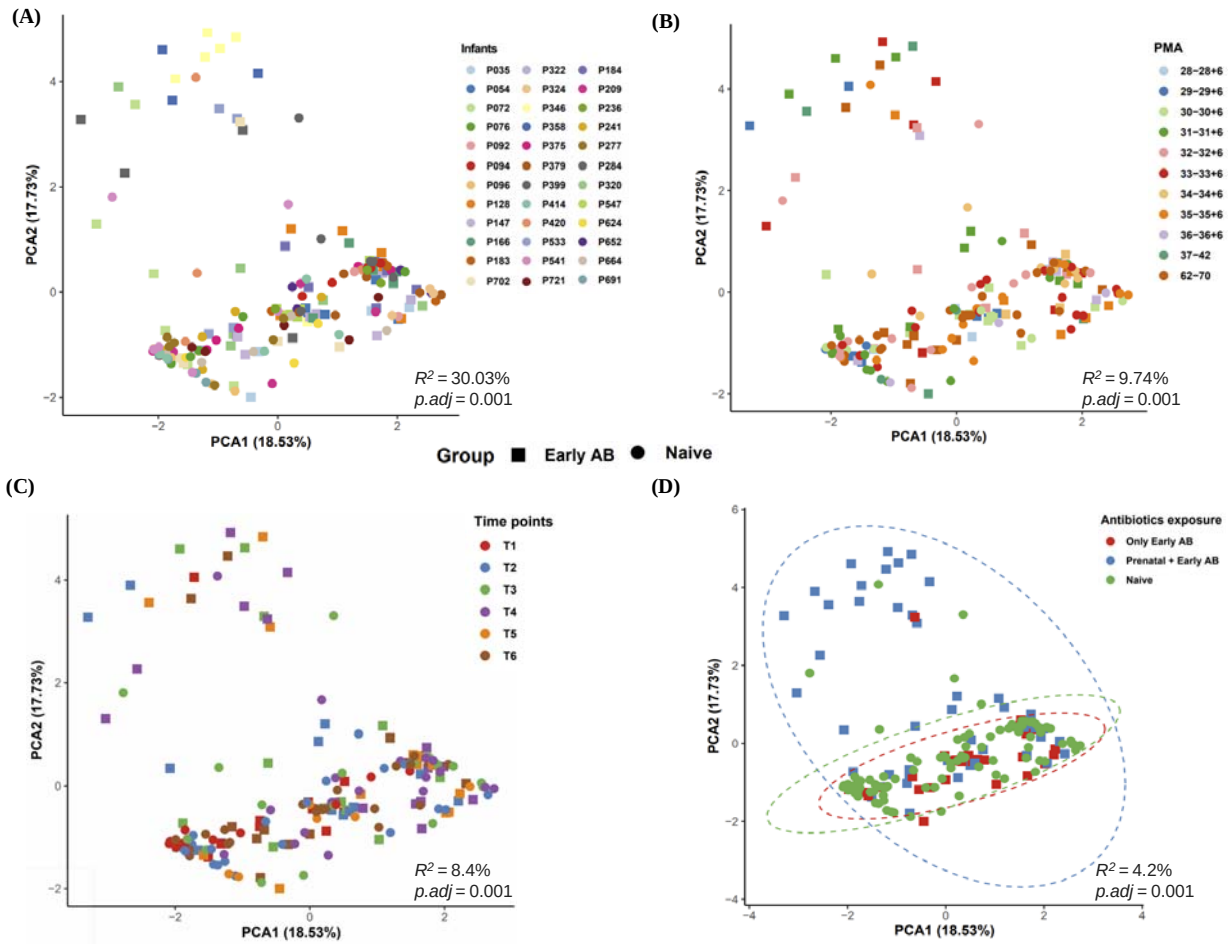


Fig. 5: Factors influencing the overall nasopharyngeal resistome composition in the preterm infants. PCA visualizations of Aitchison dissimilarities between all the nasopharyngeal samples from birth until 6 months corrected age at ARG level. Each data point corresponds to a sample, colored by (A) Infant, (B) PMA, (C) Time points and (D) Prenatal and post-natal antibiotic exposure and shaped according to early antibiotic exposure or naive groups. Effect

sizes (R^2) and corresponding adjusted p-values (p.adj, corrected using the Benjamini-Hochberg method) are calculated using PERMANOVA-tests (as shown in the bottom right of each plot). Percentage refers to the percentage of variance explained by the principal component. PMA = postmenstrual age; PC1 = principal component 1; PC2 = principal component 2; PERMANOVA = permutational multivariate analysis of variance; p.adj = adjusted p-values.

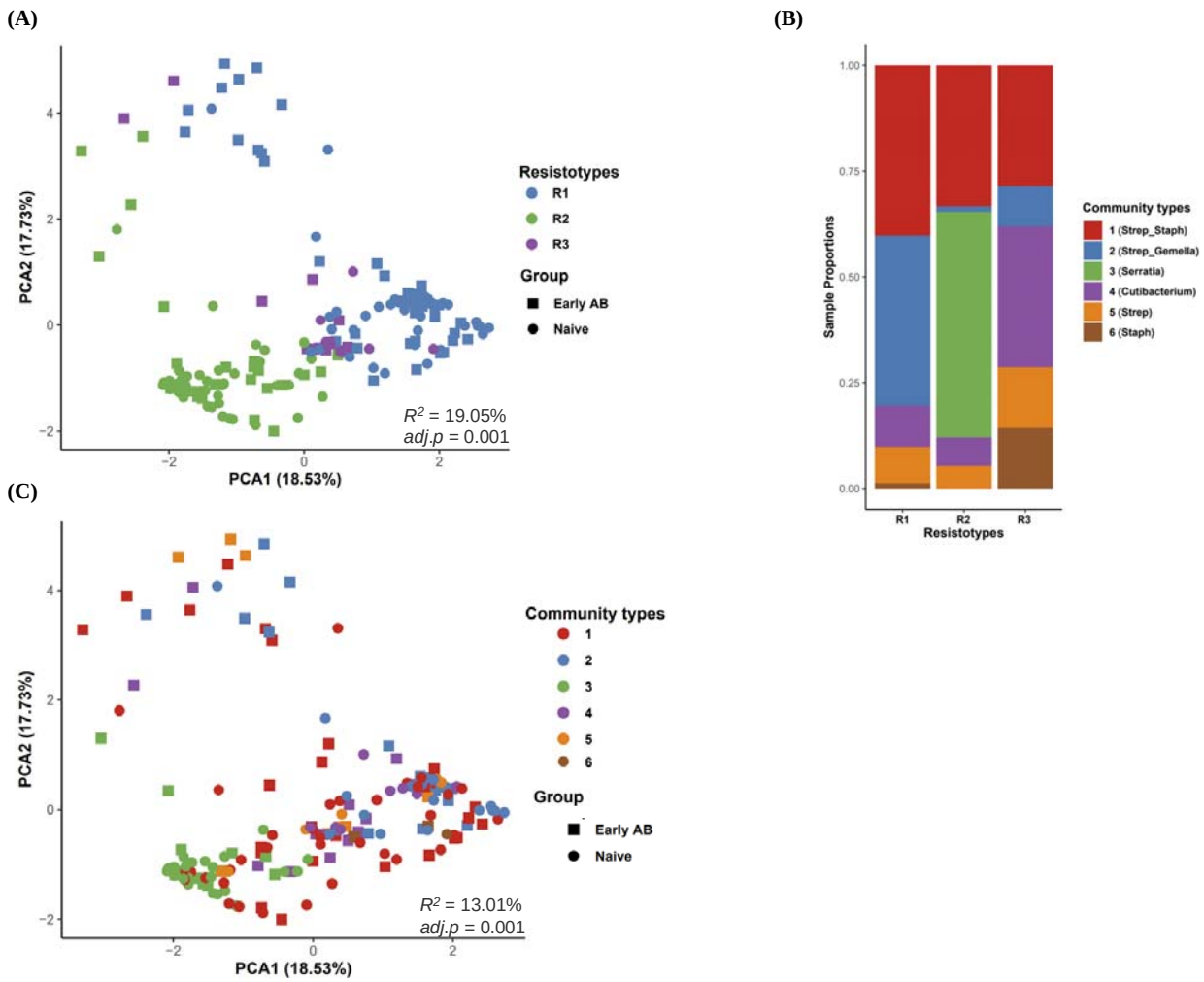


Fig. 6: Association between resistotypes, community types, and ARG abundance profiles. The PCA plot showing the association between overall ARG abundance profile with respect to their (A) resistotypes and (C) community types identified among all the nasopharyngeal samples

using the DMM models. R^2 was obtained using the PERMANOVA test. (B) Samples proportions for each resistotype shown as a function of community type. DMM = dirichlet multinomial mixture; Strep = *Streptococcus* ; Staph = *Staphylococcus*.

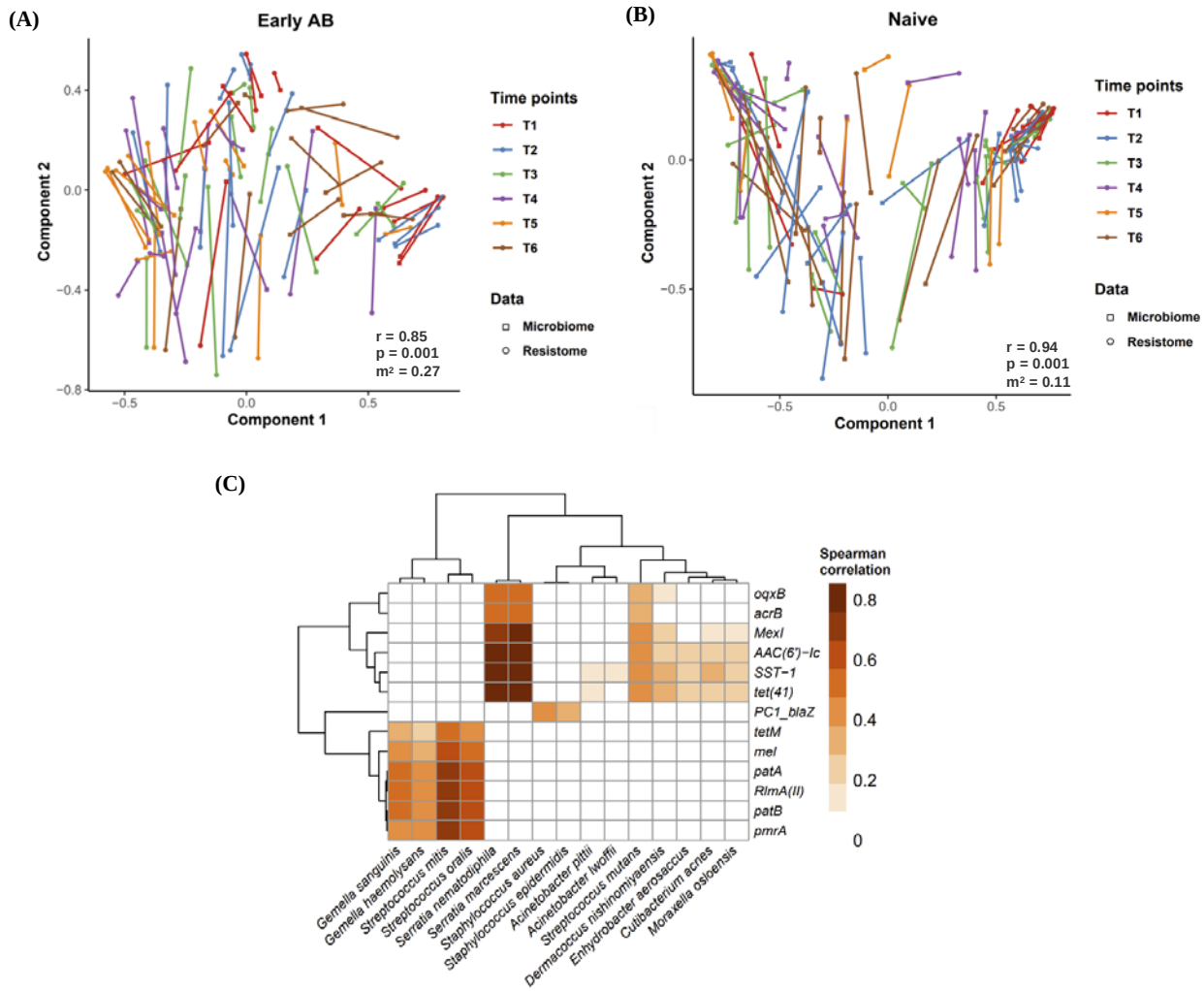
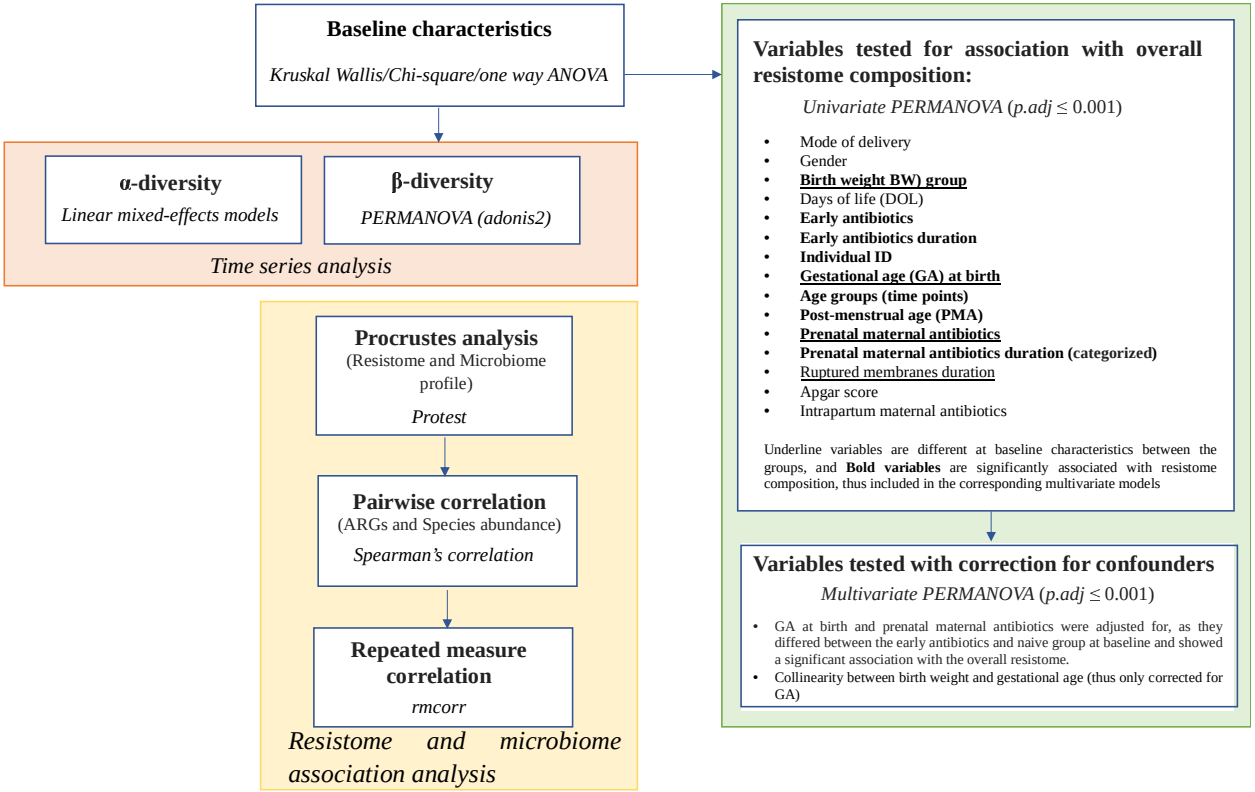
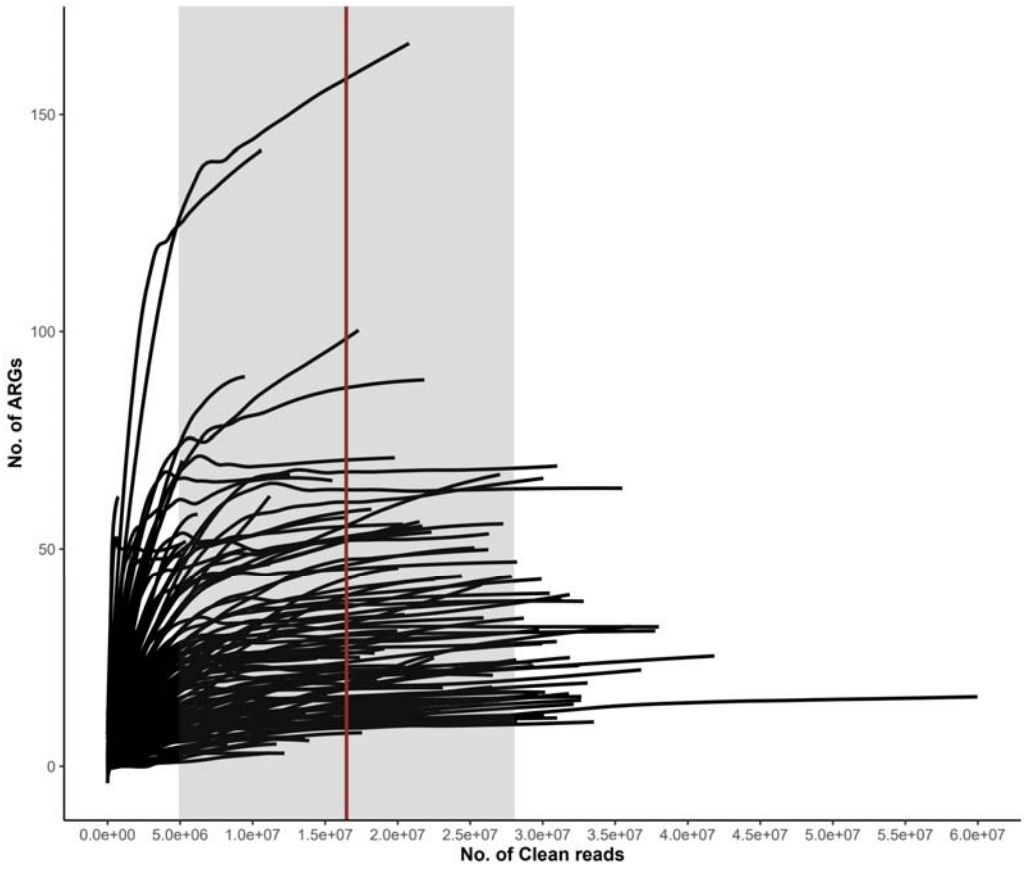


Fig. 7: The resistome variation in nasopharynx is explained by microbiome variation. The ARG and microbial abundance profiles were correlated with each other for antibiotic treated (A) and naive (B) groups using Procrustes analysis. The length of line connecting two points denotes

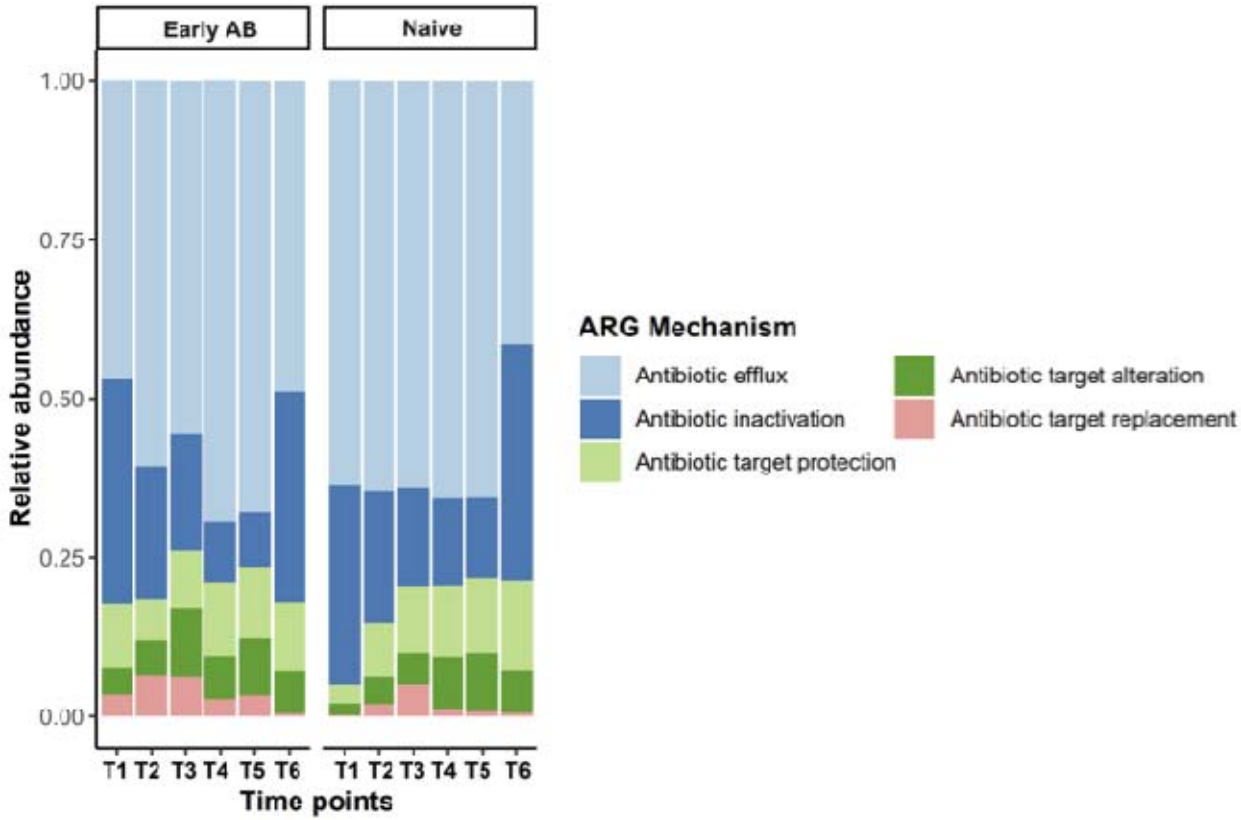
the degree of dissimilarity or distance between the microbiome (filled square) and resistome (filled circle) composition of the same sample. The line and points are colored based on time points in both groups. The correlation coefficients and significance were calculated using the *protest* function in vegan package. (C) Heatmap depicts the Spearman's pairwise correlation coefficients (*r*) between individual ARGs and microbial species from nasopharyngeal samples. Hierarchical clustering using Euclidean distance was performed on both rows and columns. Correlation coefficients (r) was shown only where adjusted p-value < 0.05. ARGs = antibiotic resistance genes.



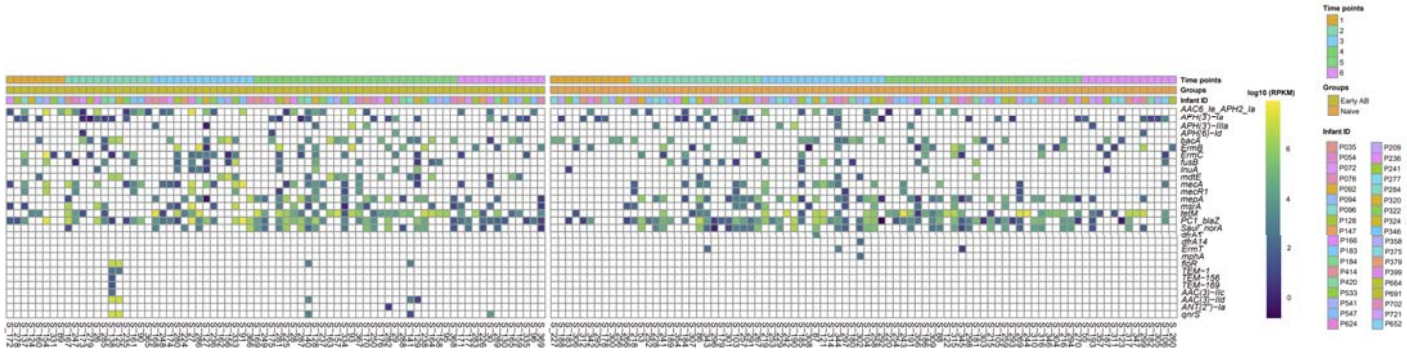
Supplementary Fig. 1: Statistical analysis design used in the study. ARGs = antibiotic resistance genes; PERMANOVA = permutational multivariate analysis of variance.



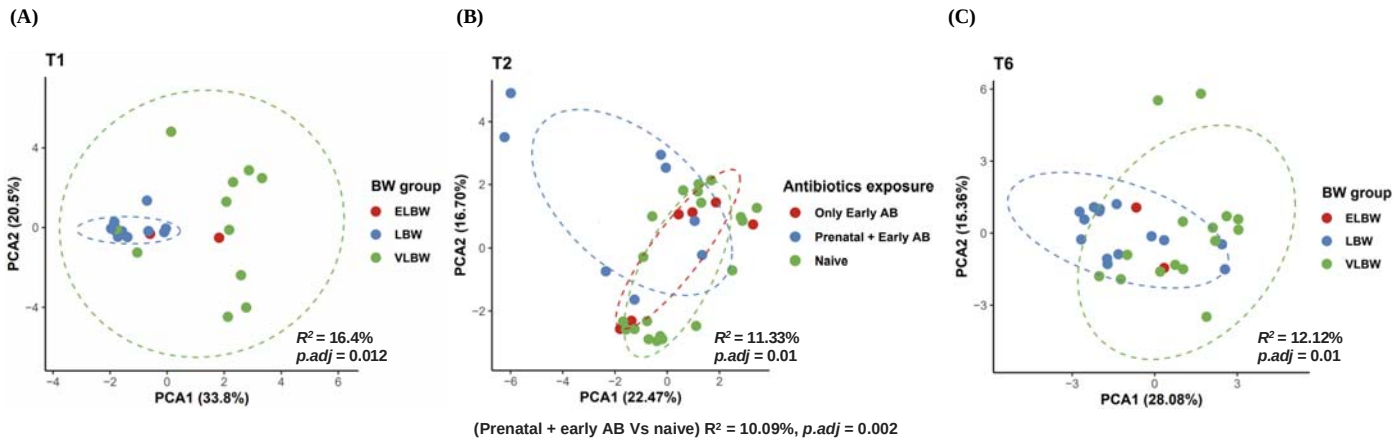
Supplementary Fig. 2: Rarefaction curve for the number of unique ARGs identified at subsampled sequencing depth. The red line indicates the median sequencing depth for all samples, while the gray shading represents one standard deviation.



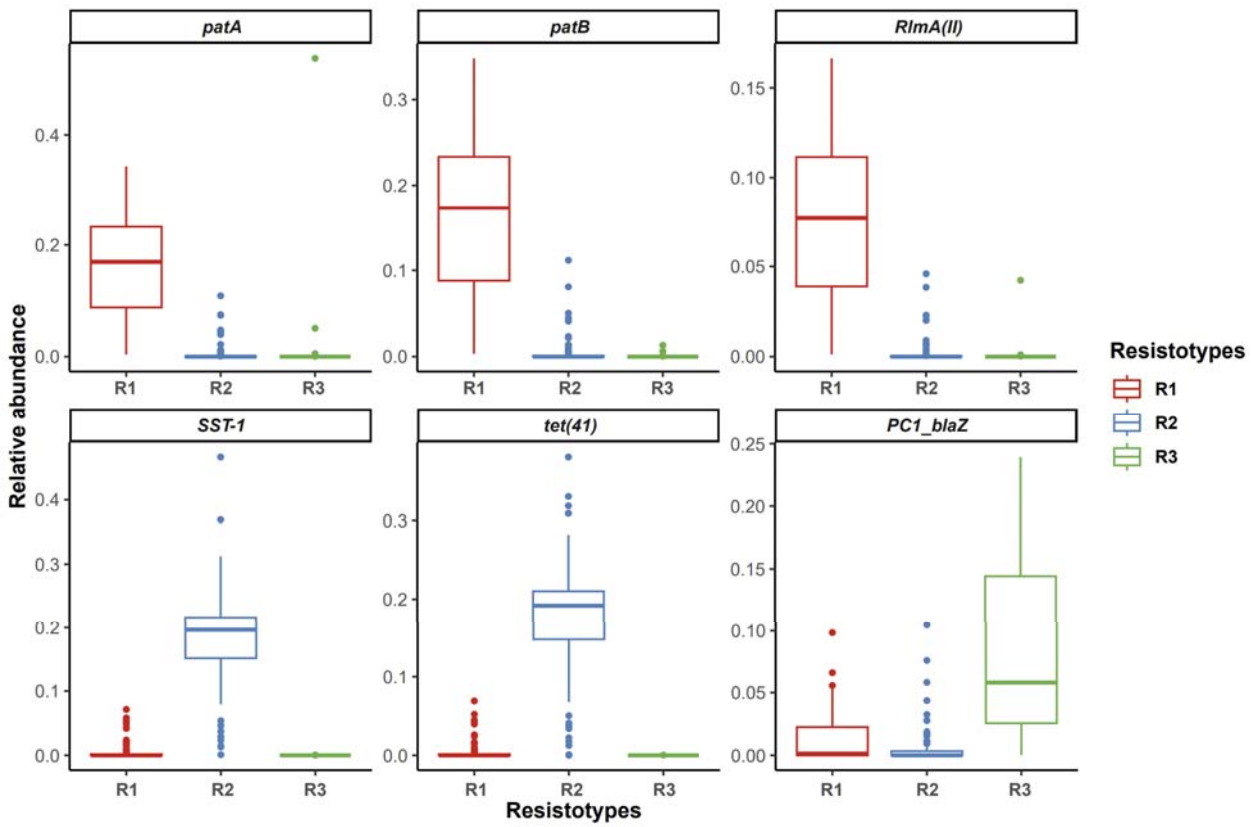
Supplementary Fig. 3: Mean normalized relative abundances of different ARG mechanisms found in the nasopharynx of early antibiotic treated and naive preterm infants at different time points (T1-T6).



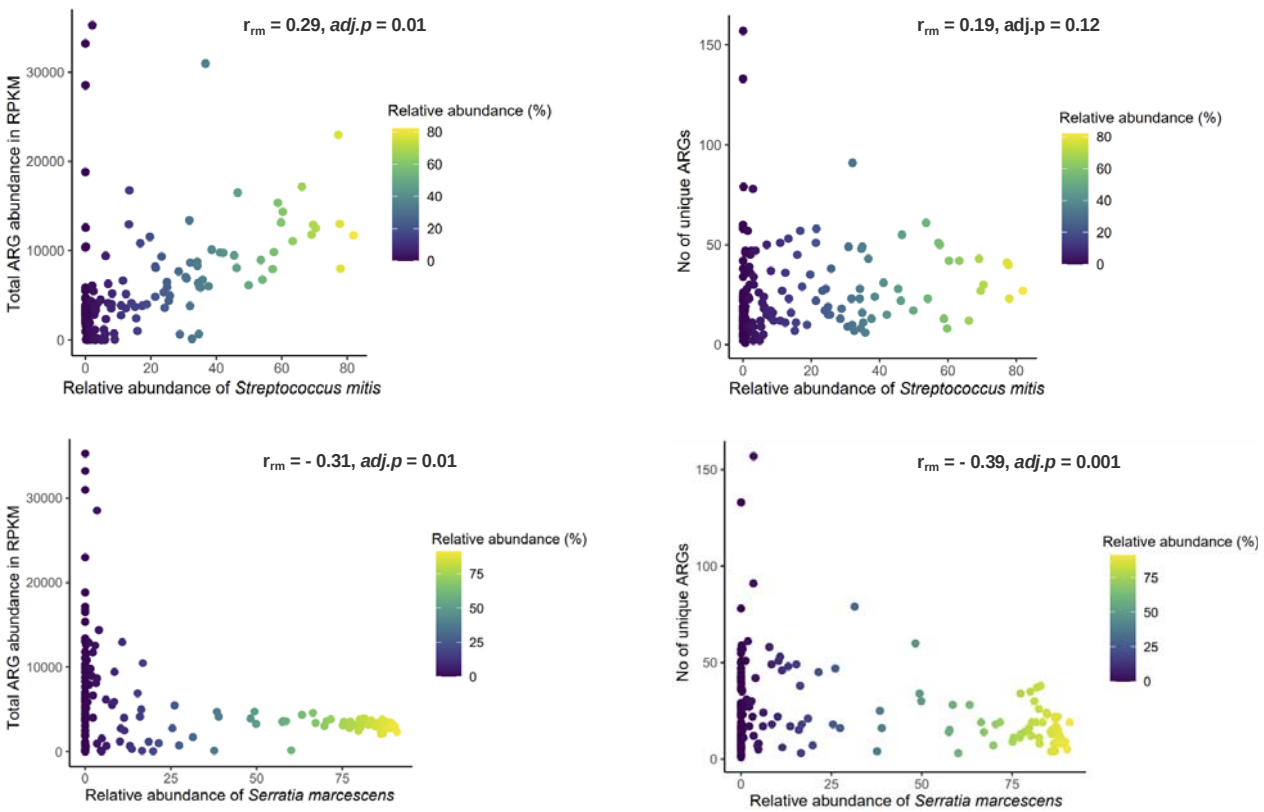
Supplementary Fig. 4: Heatmap showing the log10-transformed RPKM abundance of potential high-risk antibiotic resistance genes (ARGs) identified in the nasopharynx of all preterm infants across samples. The x-axis represents the names of samples, which are arranged by treatment group and time points. RPKM = reads per kilobase of reference gene per million bacterial reads.



Supplementary Fig. 5: PCA visualizations of beta diversity analysis using the Aitchison distances at time points: (A) T1, (B) T2 and (C) T6 (6 month corrected age). Each point corresponds to a sample colored based on birth weight group (A, C) and prenatal and postnatal antibiotic exposure (B). Ellipses indicate 95% confidence intervals (CI). Adjusted p-values and effect size (R^2) are based on PERMANOVA test. CLR = centered-log ratio; BW= Birth Weight.



Supplementary Fig. 6: Box plots of the relative abundances of the main ARG contributors are shown for each Resistotype cluster identified across all the preterm infants' nasopharyngeal resistomes using DMM models. The X-axis displays the three identified clusters, and the Y-axis represents the relative abundance of the ARGs. DMM = dirichlet multinomial mixture.



Supplementary Fig. 7: Scatterplots showing the association between correlated species and overall resistome outcomes across all infant samples from birth to 6 months corrected age. The total ARG abundance in RPKM was positively correlated with the relative abundance of *Streptococcus mitis* (A) and non-linearly negatively associated with the relative abundance of *Serratia marcescens* (C). Plots (B) and (D) depict the relationship between the number of unique ARGs and the relative abundance of *Streptococcus mitis* and *Serratia marcescens*, respectively. Plot (A) and (C) are colored by the relative abundance of *Streptococcus mitis*, while plots (B)

and (D) are colored by the relative abundance of *Serratia marcescens*. RPKM = reads per kilobase of reference gene per million bacterial reads.