

Title: Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome

Authors: Molly K. Gibson¹, Bin Wang^{1,2}, Sara Ahmadi^{1,2}, Carey-Ann D. Burnham^{2,3}, Phillip I. Tarr^{3,4}, Barbara B. Warner³, Gautam Dantas^{1,2,4,5,*}

Affiliations:

¹Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St Louis, MO, USA

²Department of Pathology and Immunology, Washington University School of Medicine, St Louis, MO, USA

³Department of Pediatrics, Washington University School of Medicine, St Louis, MO, USA

⁴Department of Molecular Microbiology, Washington University School of Medicine, St Louis, MO, USA

⁵Department of Biomedical Engineering, Washington University, St Louis, MO, USA

*To whom correspondence should be addressed: dantas@wustl.edu

Supplementary Tables

Supplementary Table 1. Metadata for 401 stool samples and 84 individuals.

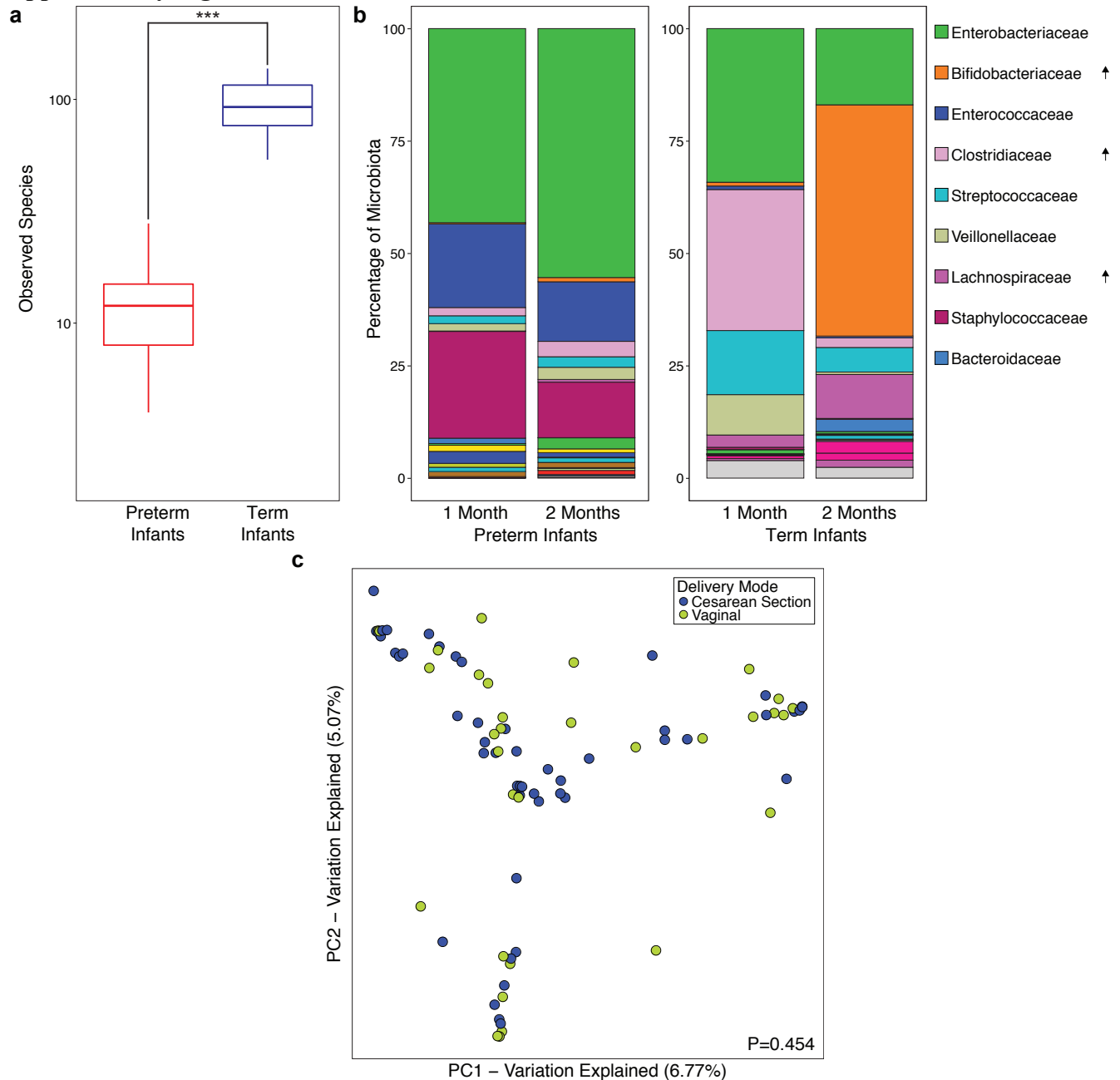
- (a) Sample metadata for 401 shotgun sequenced stool samples.
- (b) Individual metadata for 84 preterm infants included in study.
- (c) Full antibiotic history of 84 preterm infants included in study.
- (d) Summary of antibiotic exposure for all preterm infants who received a particular antibiotic during hospitalization.
- (e) Stool sample and individual identification number for all samples with 16S rRNA sequencing used for comparison to term infant gut microbiota composition.

Supplementary Table 2. Functional metagenomic selections of 21 fecal metagenomic expression libraries to 16 antibiotics.

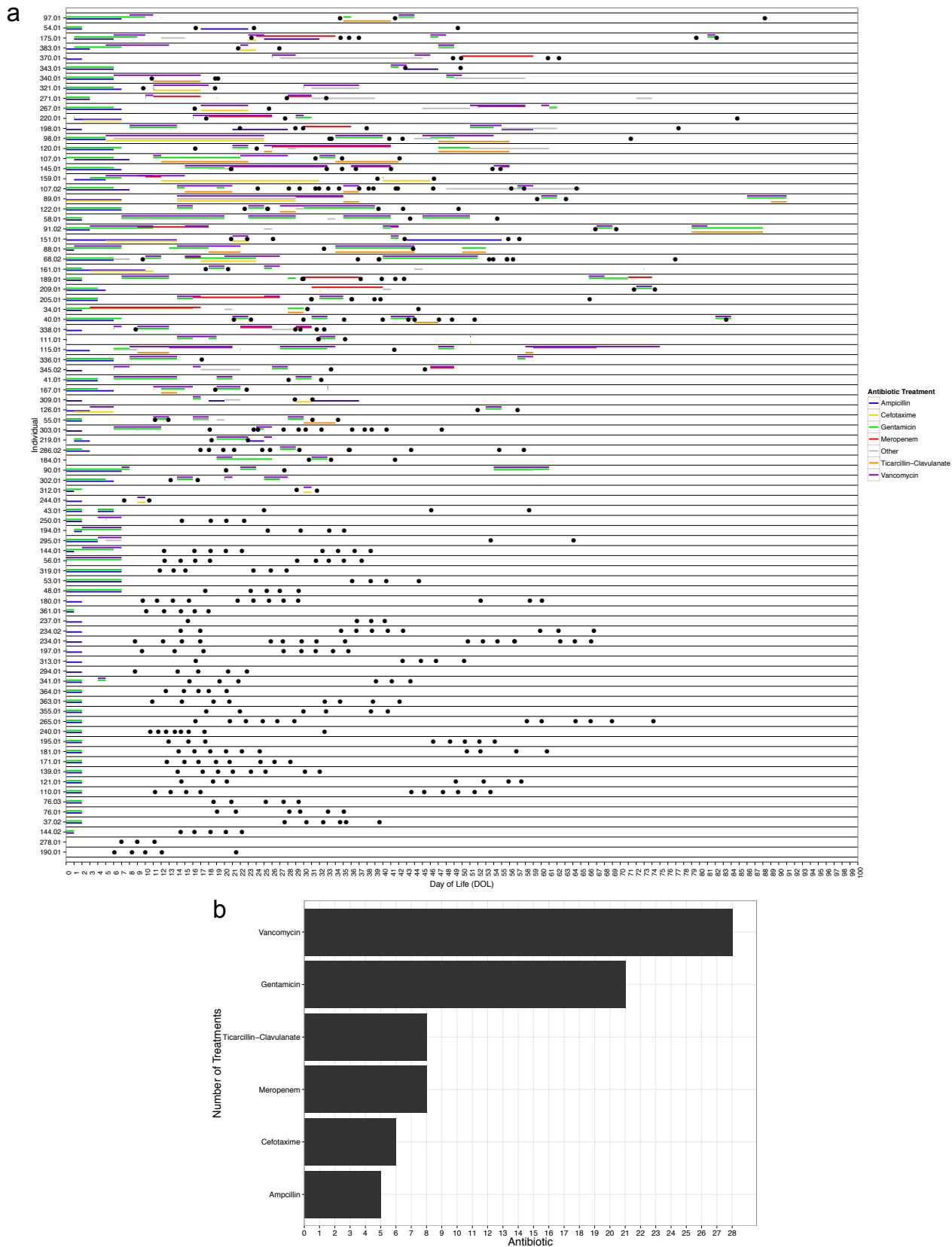
- (a) Metagenomic expression libraries constructed from 21 preterm infant stool samples.
- (b) Antibiotics and minimum inhibitory concentration used for functional selections.

Supplementary Table 3. VelvetOptimiser assembly stats for 312 assembled preterm infant metagenomes.

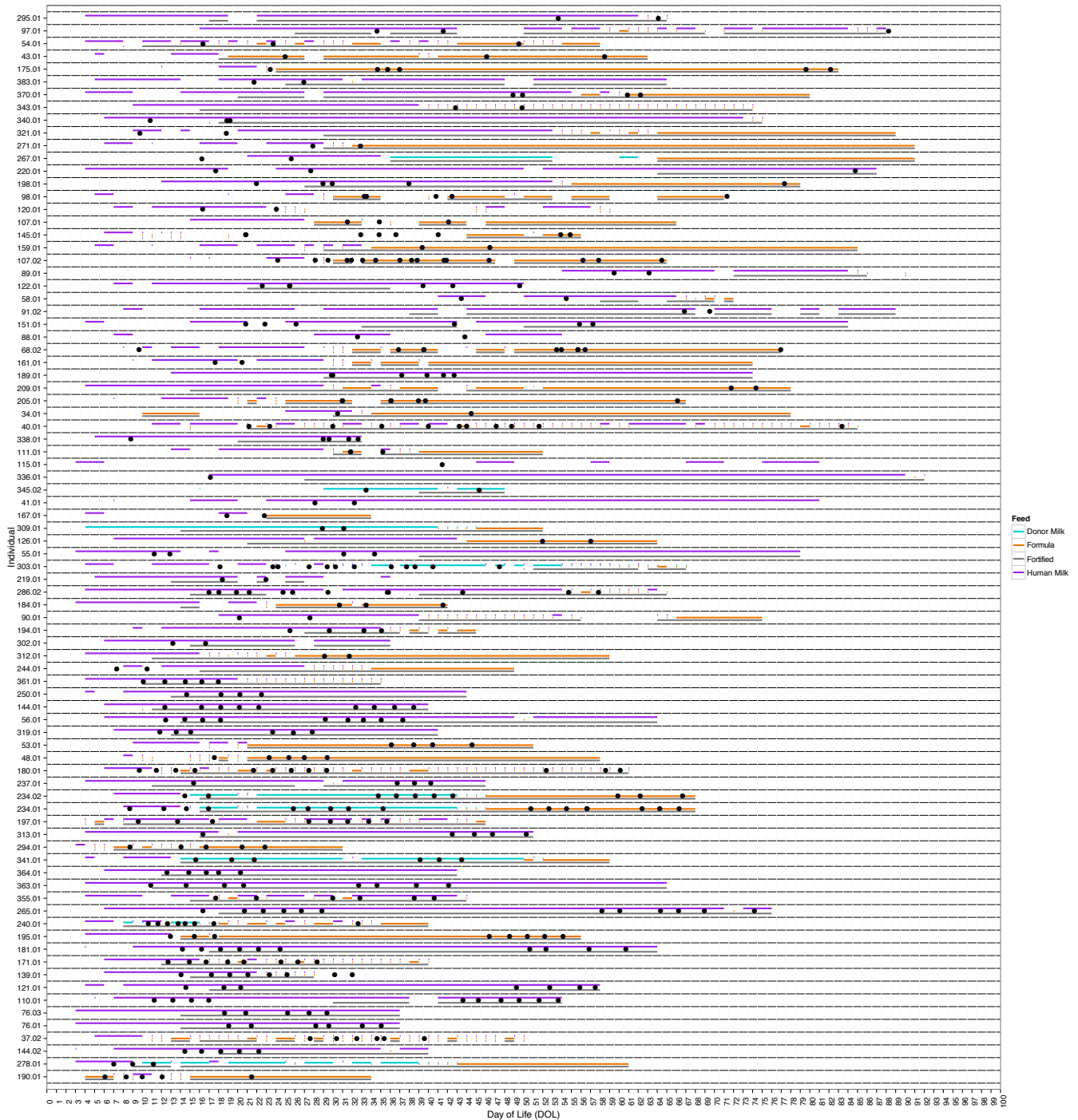
Supplementary Figures



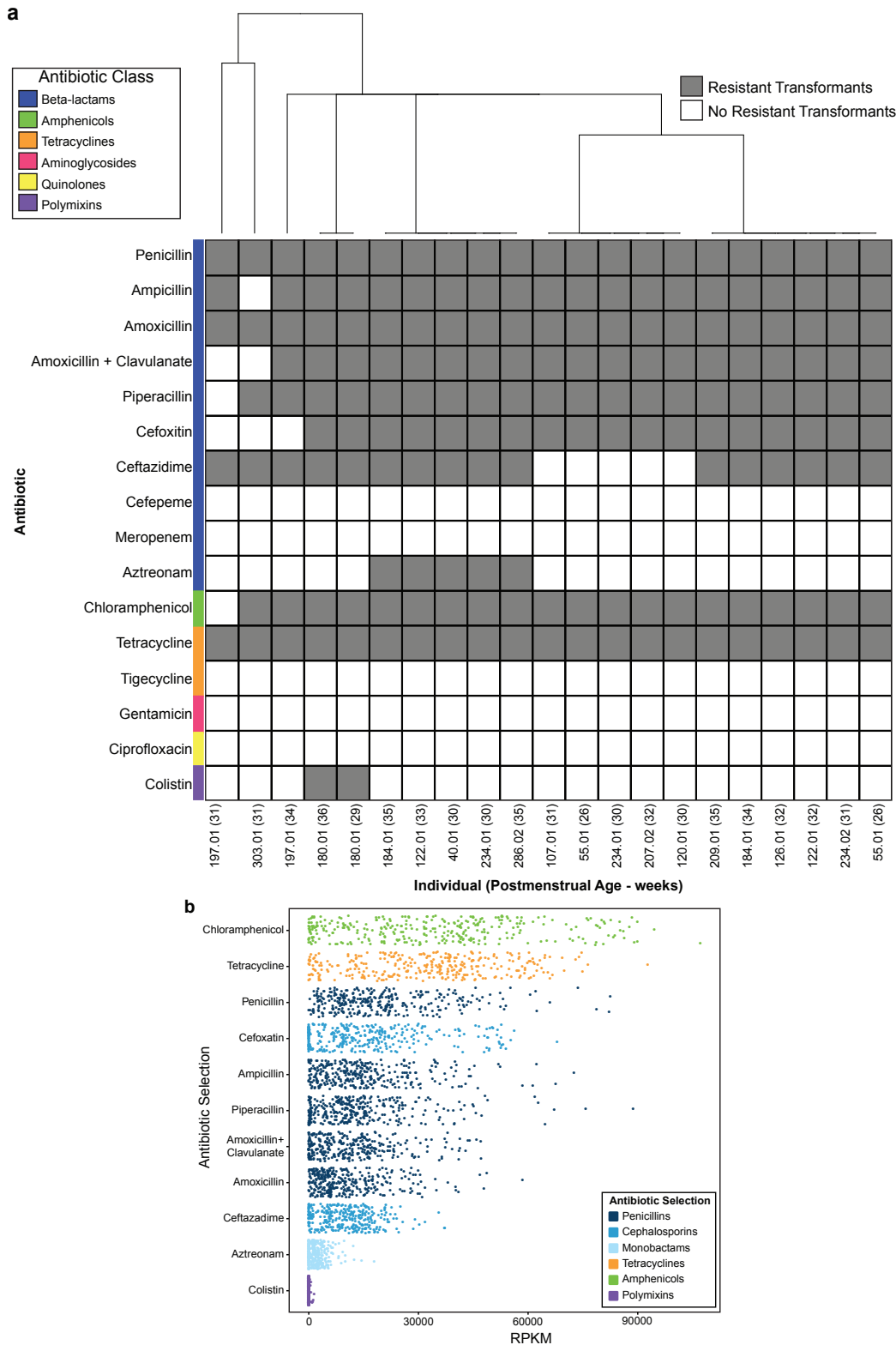
Supplementary Figure 1. The preterm infant gut microbiota is lower in species richness and compositionally distinct from age-matched term infants. (a) Preterm infant gut microbiota in the first two months of life are significantly lower in species richness than age-matched term infants ($***P < 0.001$; Wilcoxon) as determined by 16S rRNA marker gene sequencing. (b) Gut microbiota composition of preterm infants ($N=124$ 16S rRNA sequenced samples) and age-matched term infants ($N=10$ previously sequenced 16S rRNA sequenced samples) colored by bacterial family. Arrows represent the bacterial families that are significantly enriched in term infant gut microbiota ($P < 0.05$, non-parametric Student's T-test with 10,000 permutations). (c) Preterm infant gut microbiota composition is not influenced by delivery mode ($P=0.452$; PERMANOVA) as determined by species profiling of shotgun metagenomic sequencing data. Only the first sample from each individual is included ($N=84$ shotgun sequenced samples). PCoA depicts bray-curtis distance.



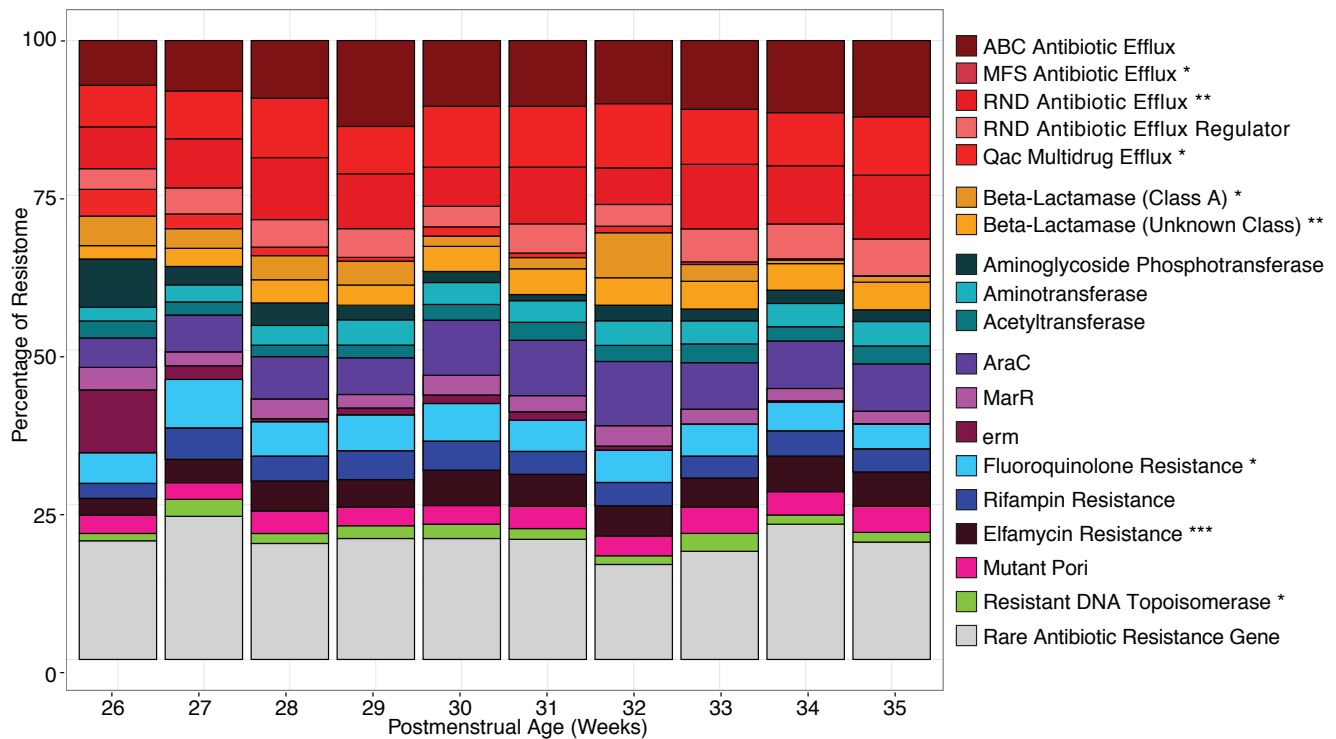
Supplementary Figure 2. 84 preterm infants and 401 samples included in this study with antibiotic treatments. (a) Day of life depicted along the x-axis for each preterm infant along the y-axis represented by Individual ID. Colored lines represent duration of a specific antibiotic treatment. Circles represent samples that were analyzed using shotgun sequencing. **(b)** Number of specific discrete antibiotic treatments where a fecal sample was collected and analyzed directly before initiation and directly after termination of antibiotic treatment (within 48 hours).



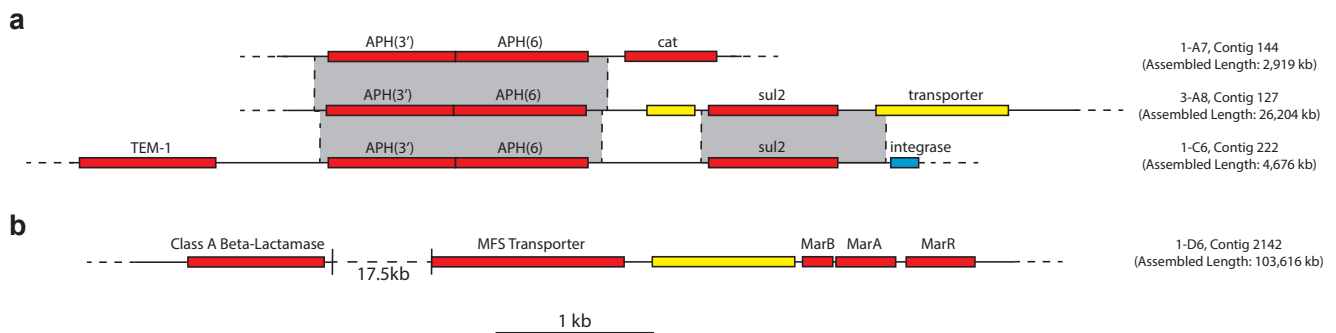
Supplementary Figure 3. 84 preterm infants and 401 samples included in this study with feeding. Day of life depicted along the x-axis for each preterm infant along the y-axis represented by Individual ID. Colored lines represent duration of a specific feed type. Circles represent samples that were analyzed using shotgun sequencing.



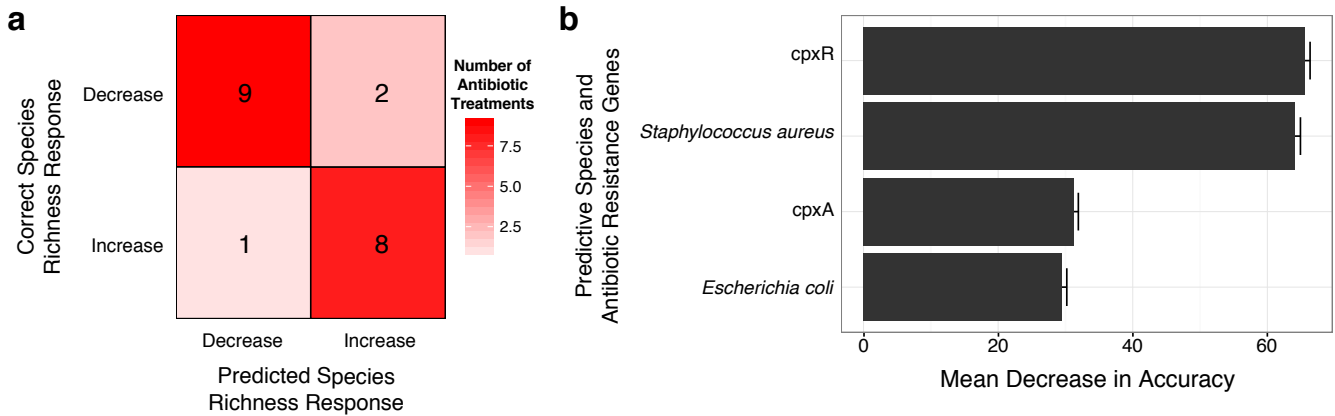
Supplementary Figure 4. Functional selections of 21 preterm infant gut metagenomes for resistance against 16 antibiotics. (a) Phenotypic results of selections. A dark grey cell means that a resistance phenotype was observed whereas white cells indicate the absence of any drug-tolerant transformants. **(b)** Relative abundance (RPKM) of functionally selected antimicrobial resistance genes with observed resistance to 11 antimicrobials across all 401 preterm infant gut metagenomes.



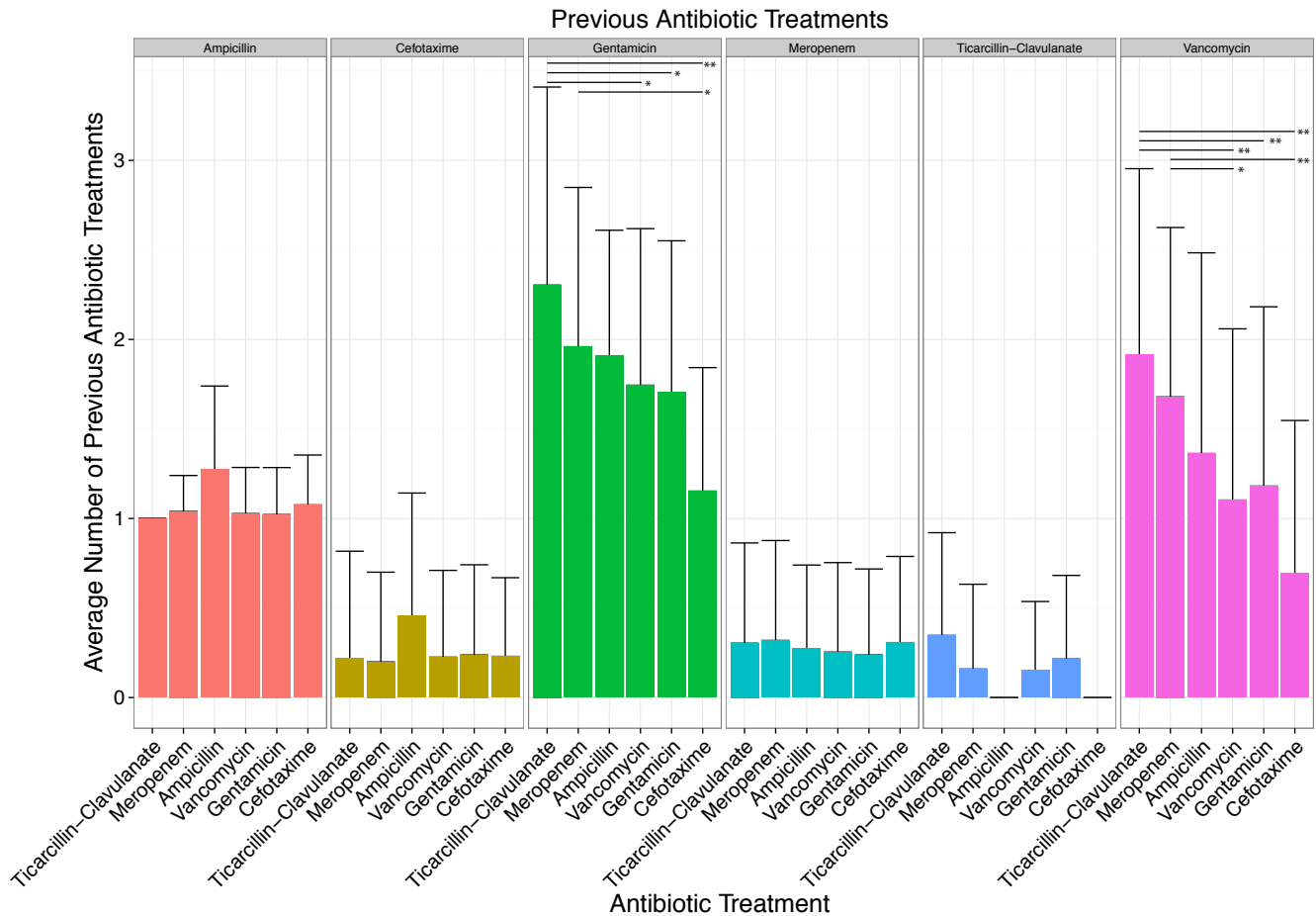
Supplementary Figure 5. Resistome development in preterm infants early in life. Total antibiotic resistance gene relative abundance binned by postmenstrual age. Rare antibiotic resistance genes are defined as resistance genes that comprise <3% of all postmenstrual age bins. Four classes significantly change over time (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; ANOVA).



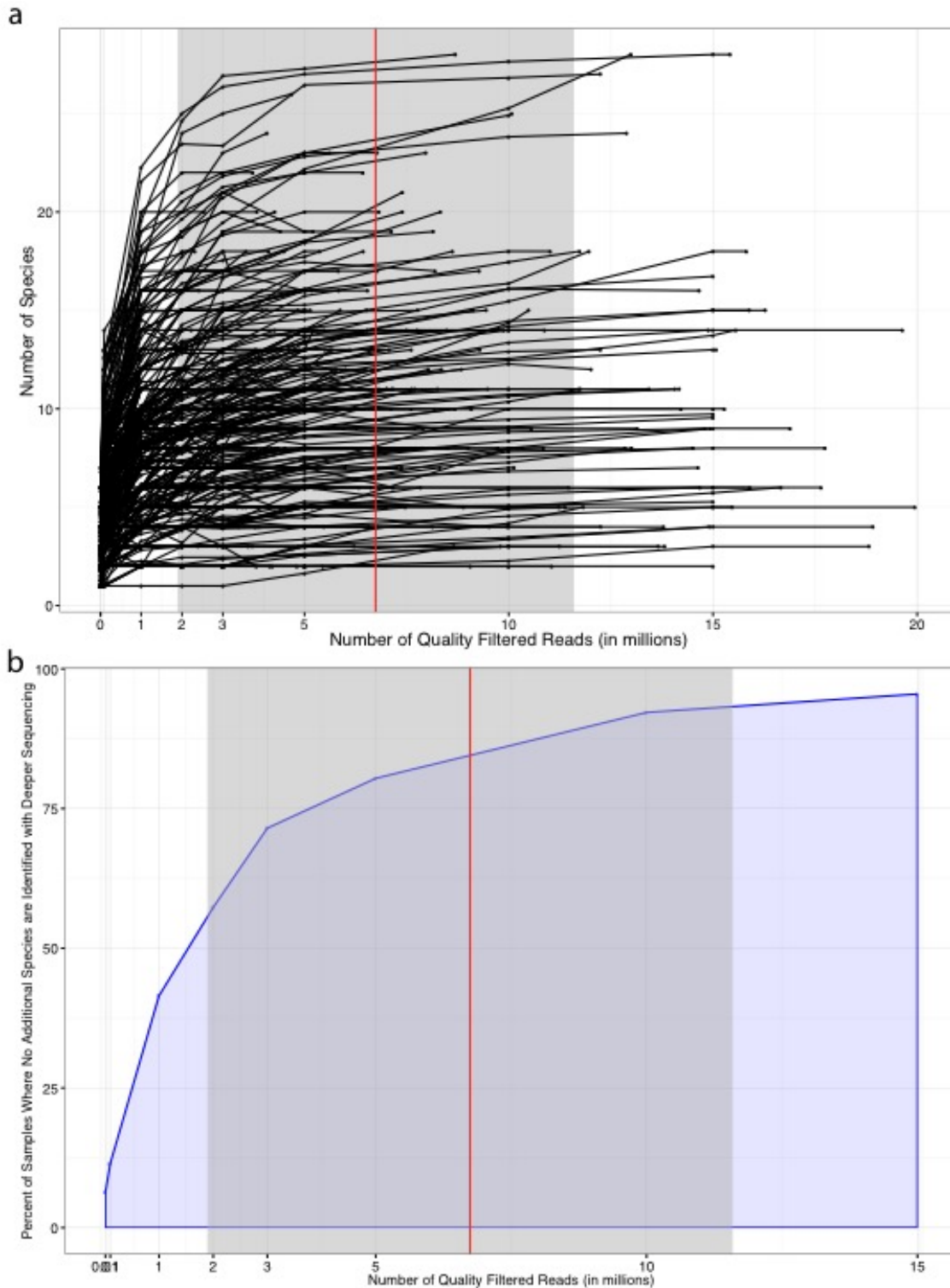
Supplementary Figure 6. Representative multidrug resistance clusters encoded in the preterm infant gut microbiota. (a) Cluster of beta-lactam, aminoglycoside, and sulfamethoxazole antibiotic resistance genes identified in 8% of samples and 12% of preterm infants. Grey shading represent >99% identity in assembled contigs. **(b)** Multidrug resistant cluster connected to beta-lactam antibiotic resistance gene identified in 8% of sample and 14% of preterm infants. (red=antibiotic resistance; blue=integration element; yellow=hypothetical or other function).



Supplementary Figure 7. Predictable response of preterm infant gut microbiota species richness to gentamicin and vancomycin combination treatment. (a) Random forest classification confusion matrix for prediction of species richness response direction using relative abundance of 2 species and 2 antibiotic resistance (AR) genes before treatment. Model only includes treatments where a response was observed (i.e. species richness change not equal to zero) and both gentamicin and vancomycin were administered in combination. Random Forests classification resulted in 15% out-of-box error rate. (b) Two predictive species and two predictive AR genes identified by applying Random Forests classification and minimizing the out-of-box error with the fewest number of predictors. Inclusion of predictors was based on mean decrease in classification accuracy when the relative abundance of the species or AR genes were randomly permuted (mean decrease +/- s.d., n=1000 replicates).



Supplementary Figure 8. Previous antibiotic exposure by antibiotic treatment. For each specific antibiotic specific, the average total number of independent prior antibiotic exposures to that specific antibiotic is shown. Specific antibiotic treatments are displayed on the x-axis and bars are grouped by prior antibiotic exposure. A prior antibiotic exposure was defined as a discrete antibiotic treatment that initiated prior to the start of the current antibiotic treatment. Specific antibiotic treatments differed in the total number of previous antibiotic exposure to both gentamicin and vancomycin. * $P < 0.05$, ** $P < 0.01$, Dunn's Test with Benjamini-Hochberg correction for multiple comparisons.



Supplementary Figure 9. Rarefaction curves for species identified by sequencing depth (after quality-filtering and removal of human-sequence reads). (a) Rarefaction curve for the number of unique species identified by rarefied sequencing depth. Red bar represents mean sequencing depth of quality filtered sequences with grey shading representing +/- s.d. **(b)** The percentage of samples where no additional species are identified with deeper sequencing depth than the indicated sequencing depth on the x-axis. Red bar represents mean sequencing depth of quality-filtered sequences with grey shading representing +/- s.d.