SUPPLEMENTARY FIGURES

Figure S1

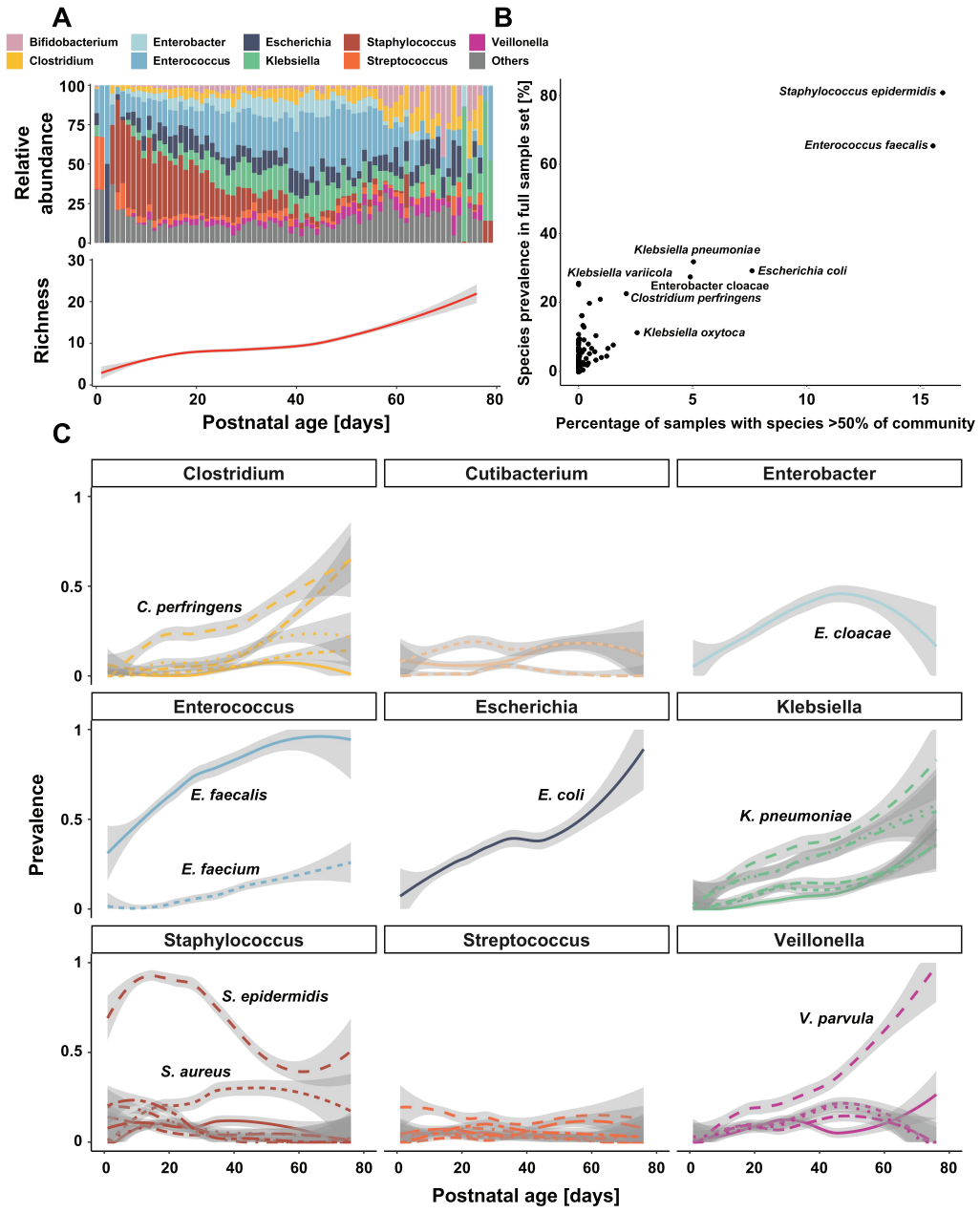

Figure S1. Microbiota assembly in the preterm infant gut microbiome during hospitalization, Related to Figure 1. A) Average relative microbiota abundance (top) and richness (bottom) over postnatal days. **B)** Species prevalence and relative abundance greater than 50% in preterm infant gut microbiomes. Selected species are labeled. **C)** Species prevalence over postnatal day of life for selected genera. Species within genera are differentiated by line type. Species of interest are labelled.

Figure S2

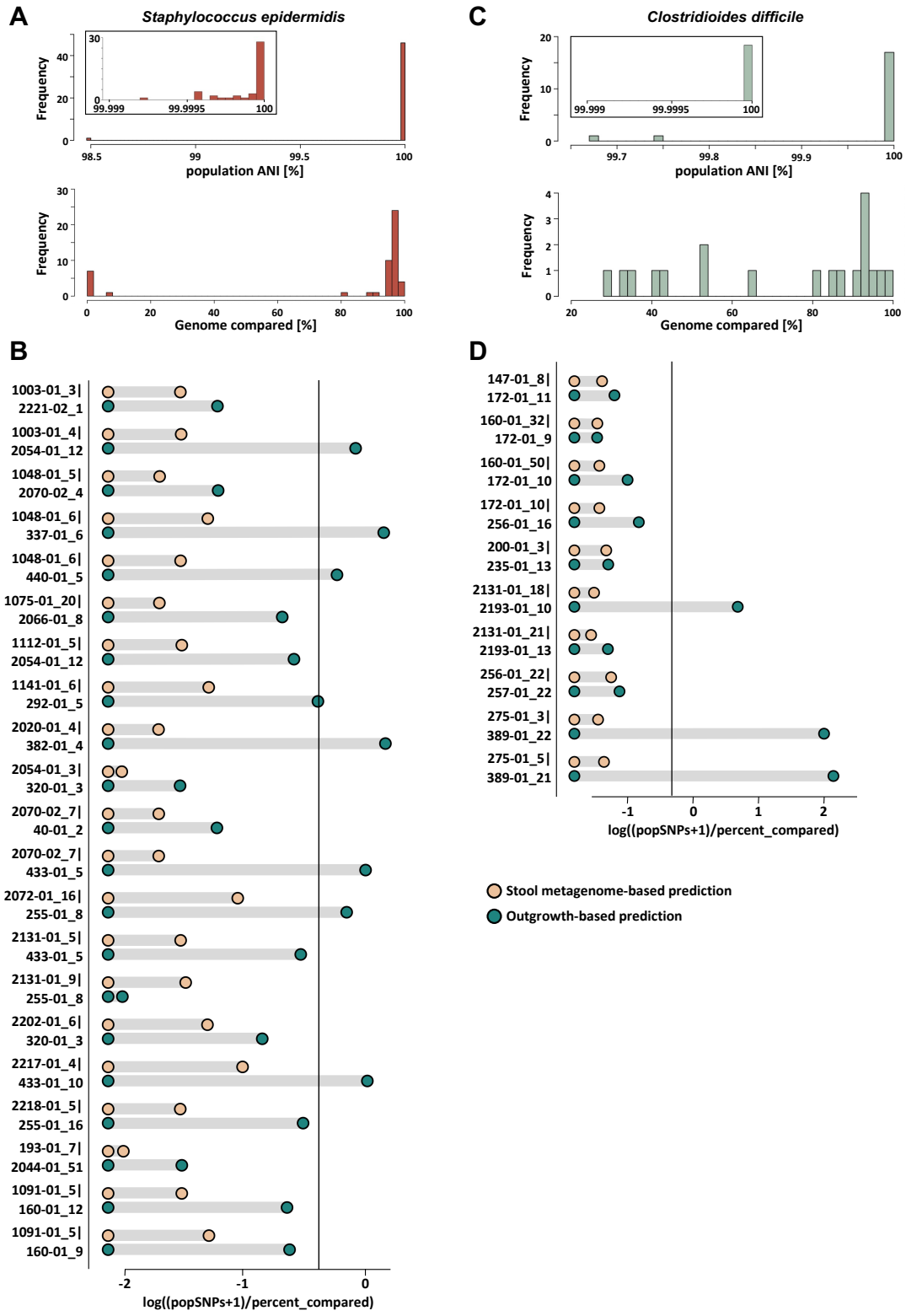

Figure S2. Confirmation of strain-sharing events across infants using whole-genome sequencing, Related to Figure 1. A) Pairwise population average nucleotide identity (popANI, top) and percent of

genome compared (bottom) values based on whole-genome sequencing of selectively outgrown *S. epidermidis* MAGs for infant sample pairs predicted to share the same strain. **B)** Pairwise population single nucleotide polymorphism to percentage of genome compared ratio for *S. epidermidis* strains determined from stool shotgun metagenomes (beige) or selectively outgrown genomes (green) as determined via inStrain. **C)** Pairwise popANI (top) and percent of genome compared (bottom) values based on whole-genome sequencing of selectively outgrown *C. difficile* MAGs for infant sample pairs predicted to share the same strain. **D)** Pairwise population single nucleotide polymorphism to percentage of genome compared ratio for *C. difficile* strains determined from stool shotgun metagenomes (beige) or selectively outgrown genomes (green) as determined via inStrain. Vertical line represents SNPs equivalent to a population ANI of 99.999%.

Figure S3

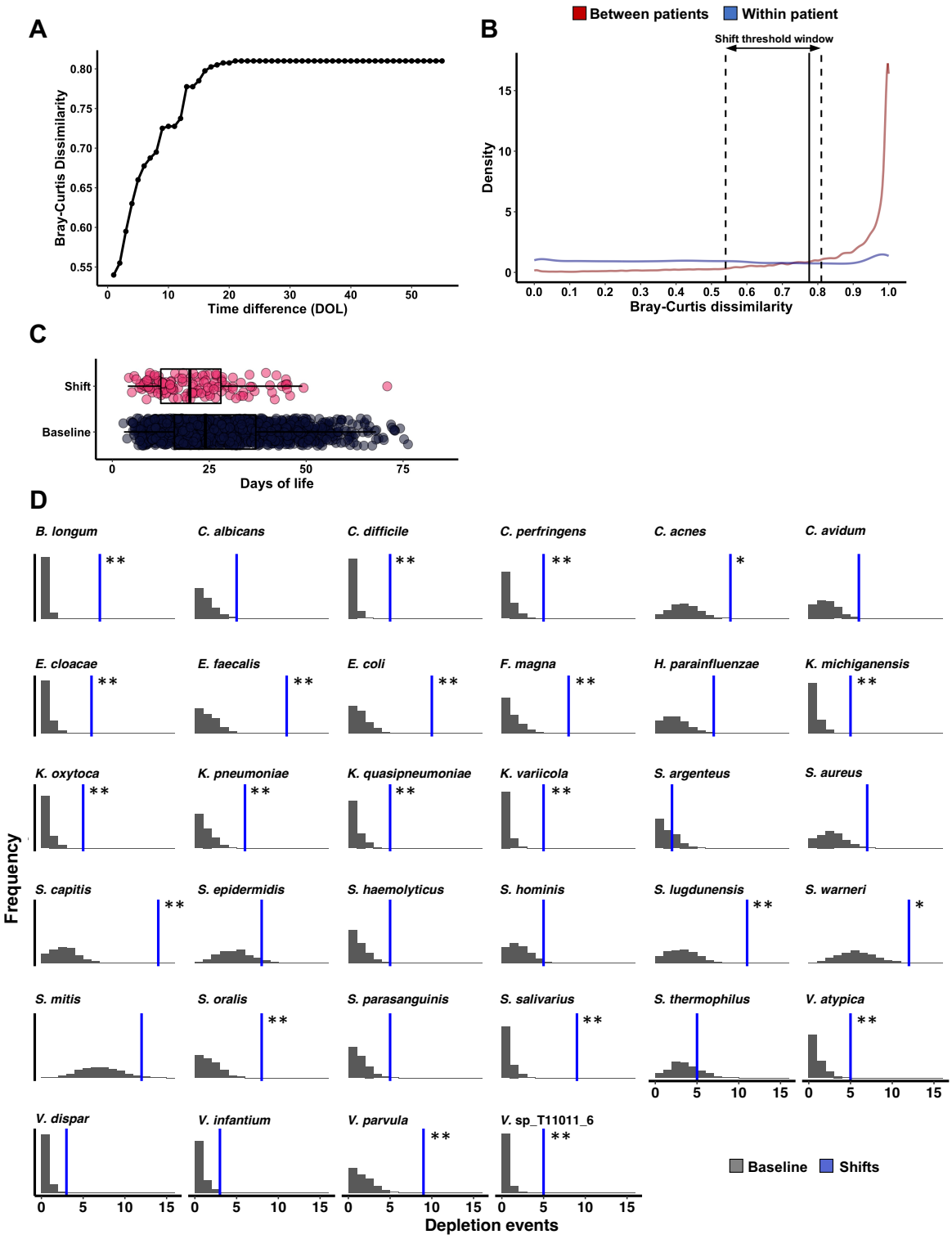

Figure S3. Microbiome shifts occur during NICU hospitalization, Related to Figure 2. A) Bray-Curtis dissimilarity between gut microbiota profiles of subsequent samples collected from the same infant by day of life difference. **B)** Microbiome shifts are defined as Bray-Curtis dissimilarities between consecutive

samples collected from the same patient (blue) that are $\geq$ Bray-Curtis dissimilarities between samples from unrelated individuals (red) collected with the same time interval. Microbiome shift definition window is determined as the point when same-patient samples equal unrelated patient distance for all possible time differences between sample collection. **C)** Microbiome shifts plotted by their occurrence in the postnatal day of life timeline. Time of sample collection for non-shift samples is plotted as a comparison. **D)** Observed microbiota depletion in shift events compared to expected frequency of depletions of the same species occurring in 999 permutations of 131 non-shift samples (Permutation test, $*p<0.05$, $**p<0.01$). Full taxa and statistics appear in data file S19.

Figure S4

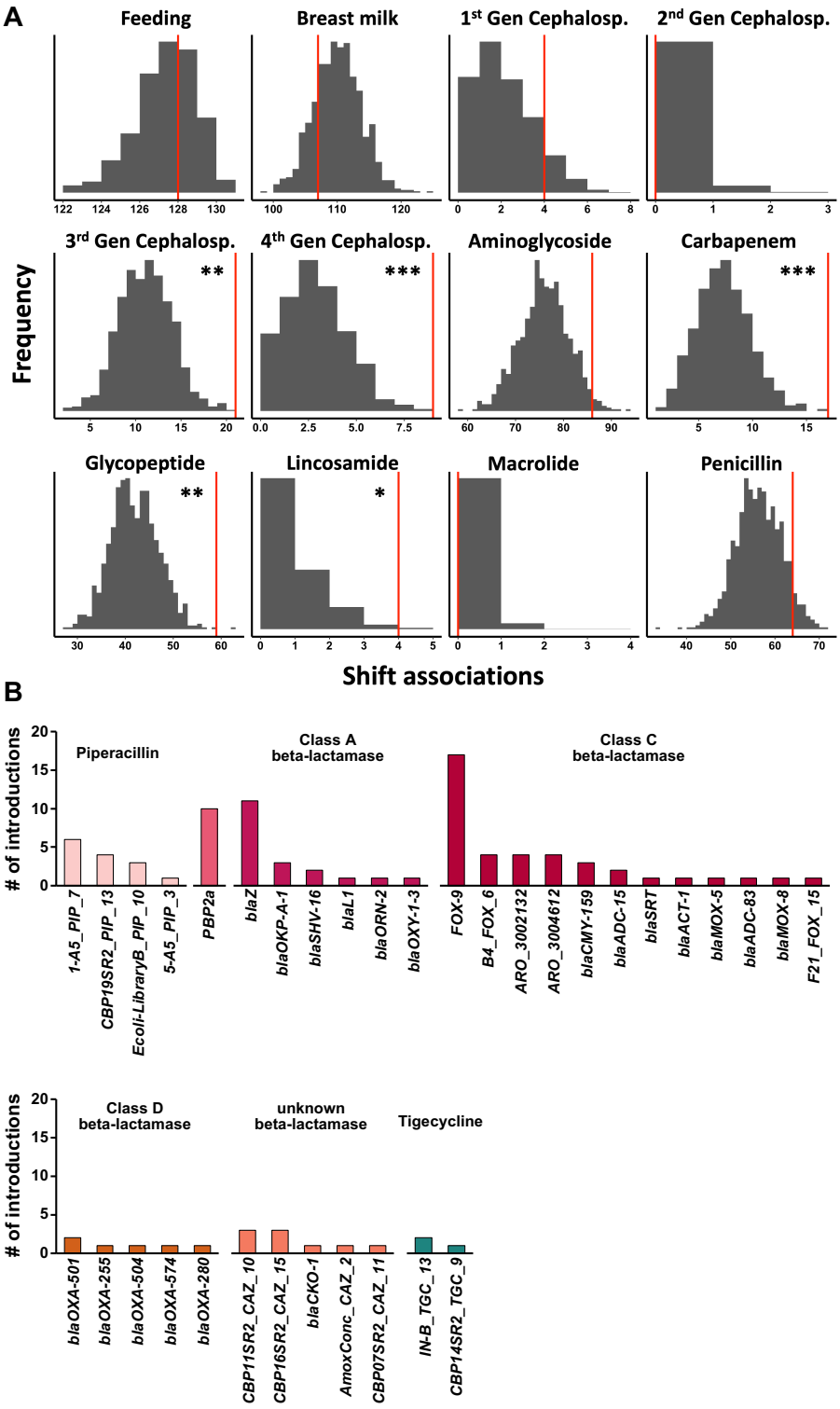

Figure S4. Microbiome shifts are driven by antibiotic exposures and introduce antibiotic resistance genes into the preterm gut microbiome, Related to Figure 2. A) Exposures associated with microbiome shift events (Permutation test, *p<0.05, **p<0.01). **B)** ARGs introduced at microbiome shift events by antibiotic class.

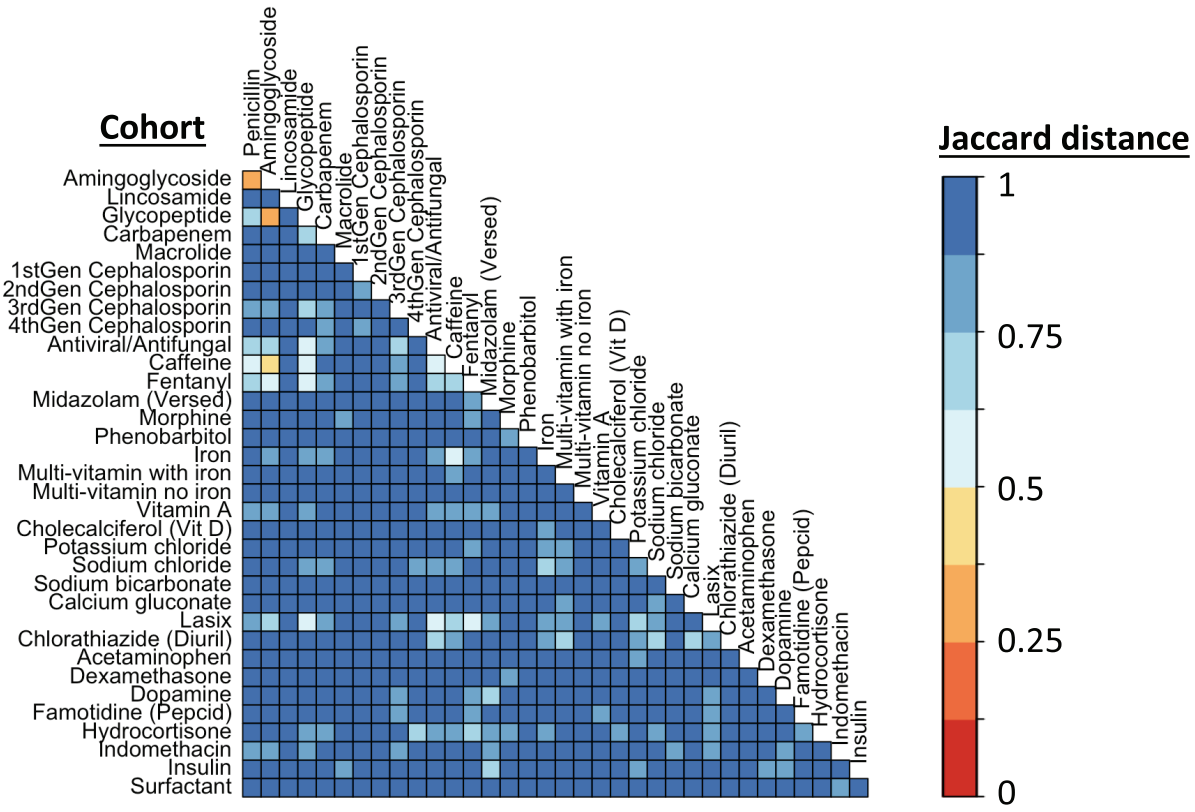

Figure S5. Co-exposure of NICU medications, Related to Figure 3. Jaccard distance between samples' exposure profiles determined based on a matrix of current treatments for the complete time of hospitalization.

Figure S6

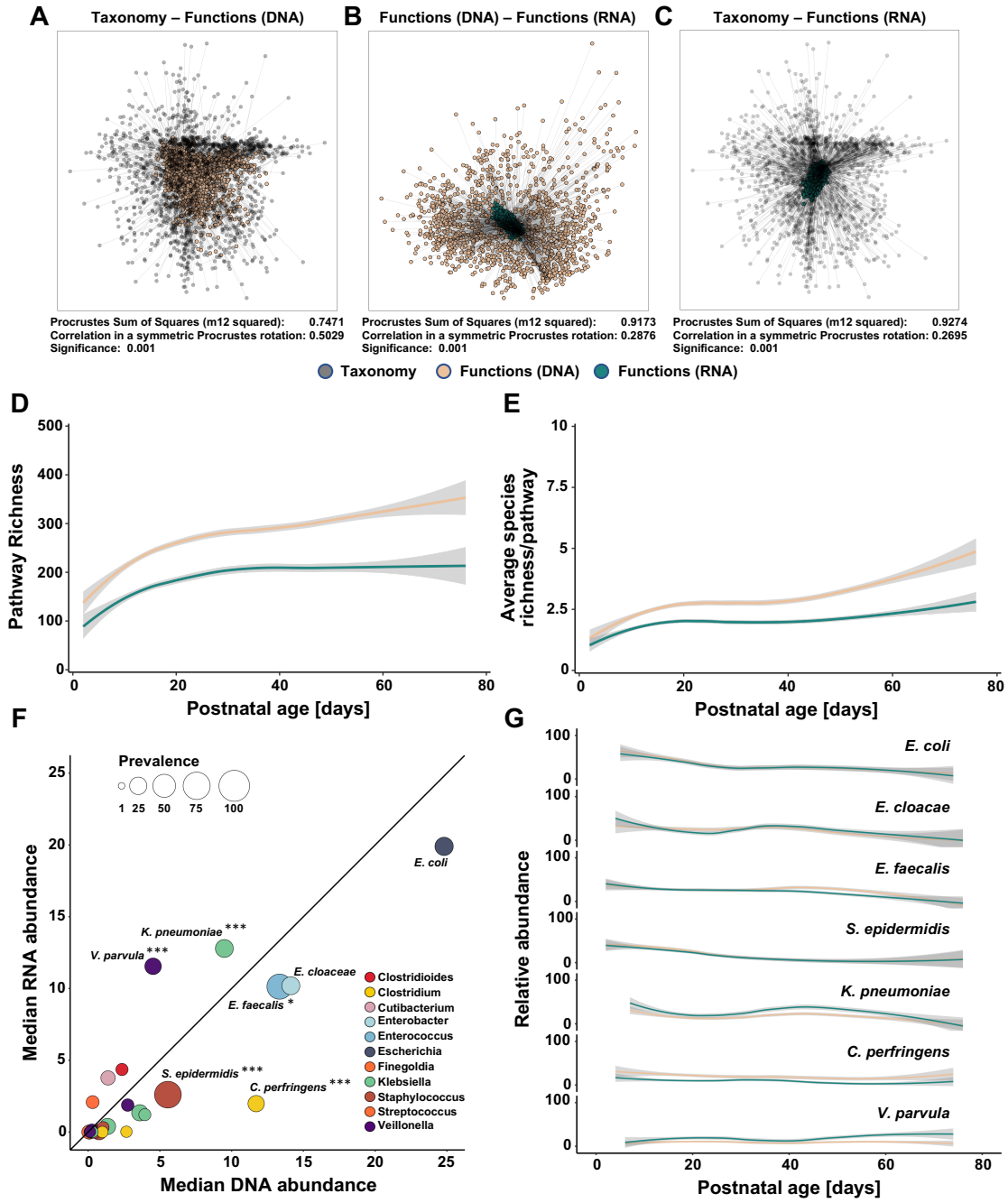

Figure S6. Transcriptional microbiota profiles systemically mirror DNA-based microbiome profiles but are distinct on a per-taxa basis, Related to Figure 4. A) Procrustes analysis of taxonomic and functional microbiota DNA profiles. **B)** Procrustes analysis of functional microbiota DNA and RNA profiles. **C)** Procrustes analysis of taxonomic DNA and functional expression profiles. **D)** Pathway richness on the DNA (beige) and RNA (green) level. **E)** Average species richness per pathway on the DNA (beige) and RNA (green) level. **F)** Median abundance of species on the RNA and DNA level. Prevalence is depicted as bubble size and species are colored by genus. Transcriptional over-/under-representation is determined using zero-inflated gaussian models (*p<0.05, **p<0.01, ***I<0.001). **G)** Species abundance on the DNA (beige) and RNA (green) level of selected samples over postnatal days.

Figure S7

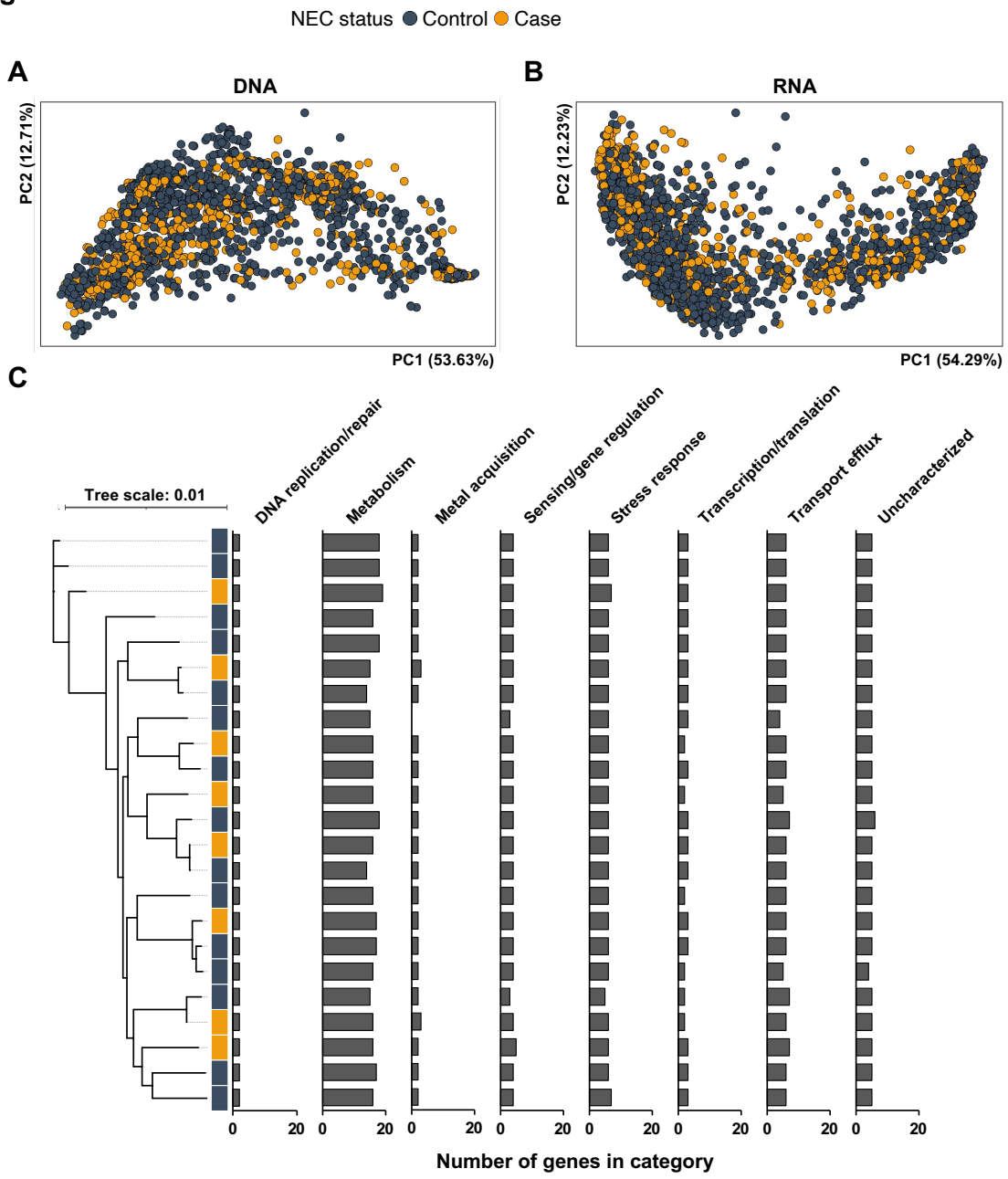

Figure S7. Bacterial virulence profiles do not differ between gut microbiomes of NEC cases and matched controls, Related to Figure 5. A-B) PCoA based on virulence genes (**A**) and transcription (**B**) of the microbiota in NEC cases (yellow) and matched cases (blue). **C)** Virulence factors annotated into *K. pneumoniae* metagenome-assembled genomes from NEC cases (yellow) and controls (blue). *K. pneumoniae* virulence profiles are clustered based on a virulence factor presence absence/matrix. Virulence factors presence in genomes is summarized by functional category.

Figure S8

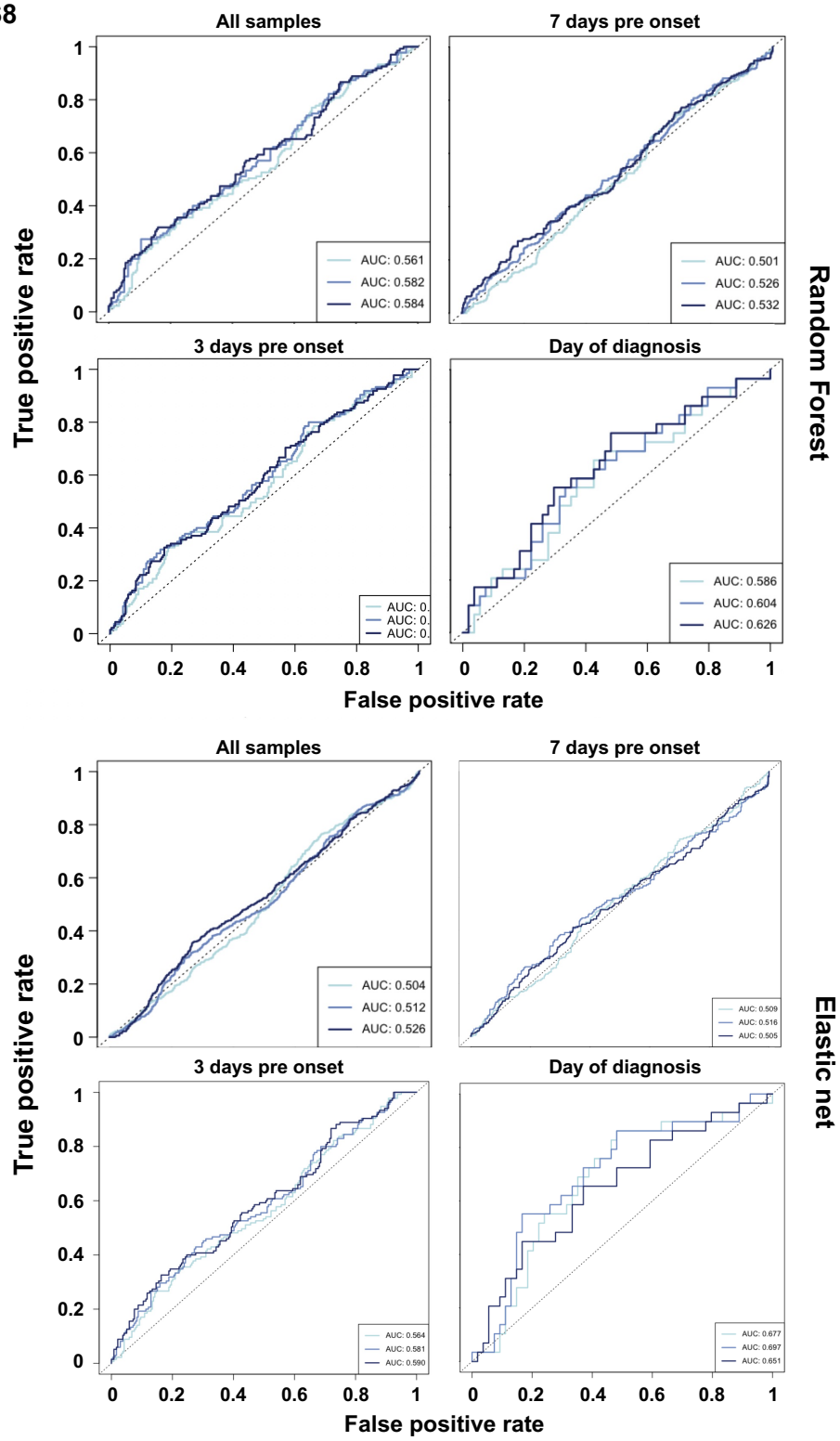

Figure S8. Microbiome profiles do not reliably predict NEC onset until day of diagnosis, Related to Figure 5. Receiver operating characteristic curves of random forest models (top) and elastic nets (bottom) accounting for repeat measures utilizing all microbiome data collected from all samples, or samples taken in the week before, 3 days before, or at NEC onset.

Figure S9

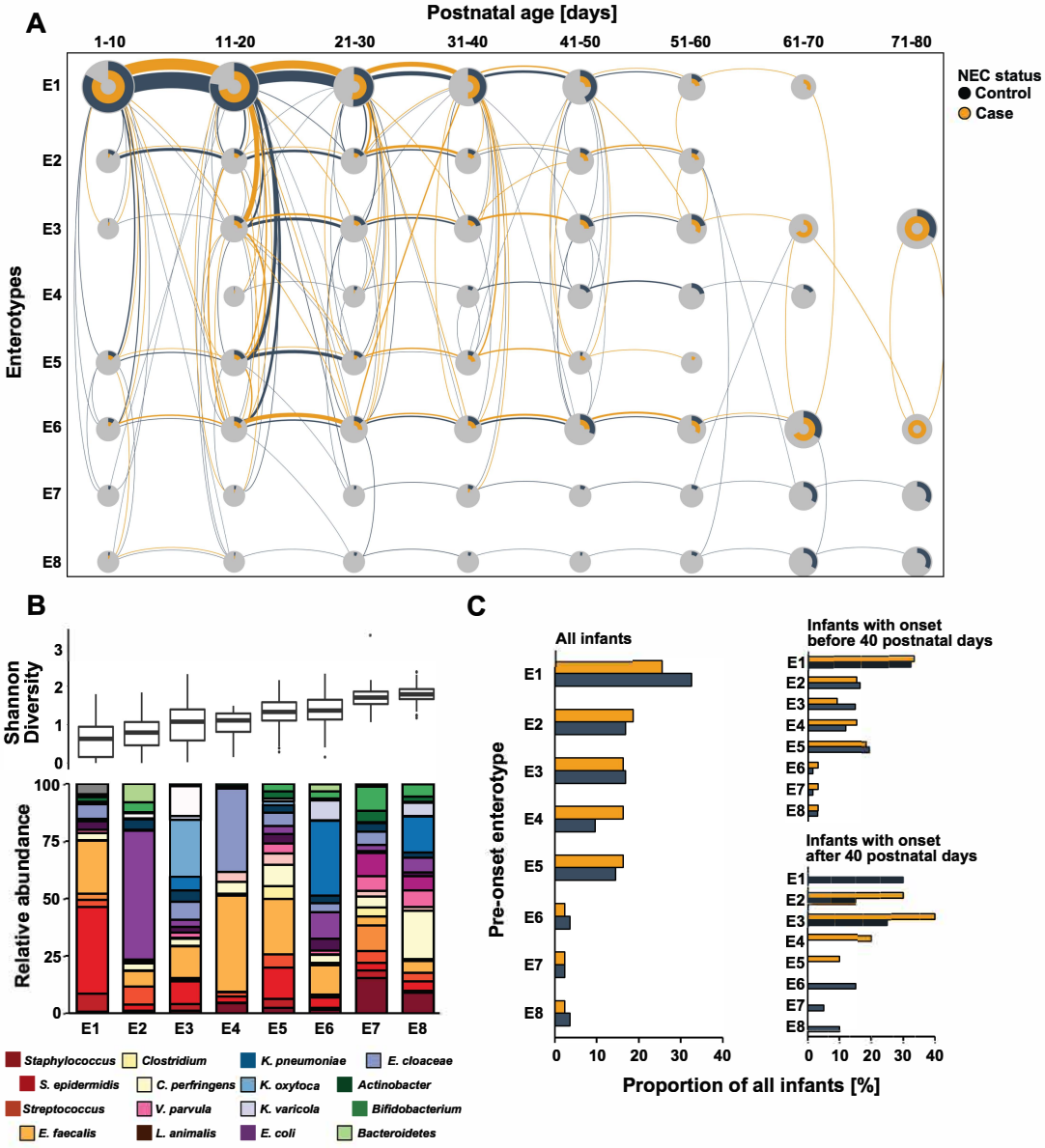

Figure S9. Enterotype trajectories do not predict NEC onset but differentiate cases and controls with onset after postnatal day 40, Related to Figure 5. A) Enterotype trajectories of NEC cases (yellow) and matched controls (blue) over postnatal days of life. Bubble size corresponds to the number of infants in a respective enterotype during a given temporal window. Lines illustrate transitions of infants between enterotypes. Line thickness corresponds to the number of infants transitioning between two given enterotypes. **B)** Bacterial composition and diversity of enterotypes. **C)** Summary of enterotype identity for all NEC cases (yellow) and matched controls (blue) for all infants (left), early-onset cases and matched controls (right top), and late-onset cases and matched controls (right bottom).