

Supplementary Information

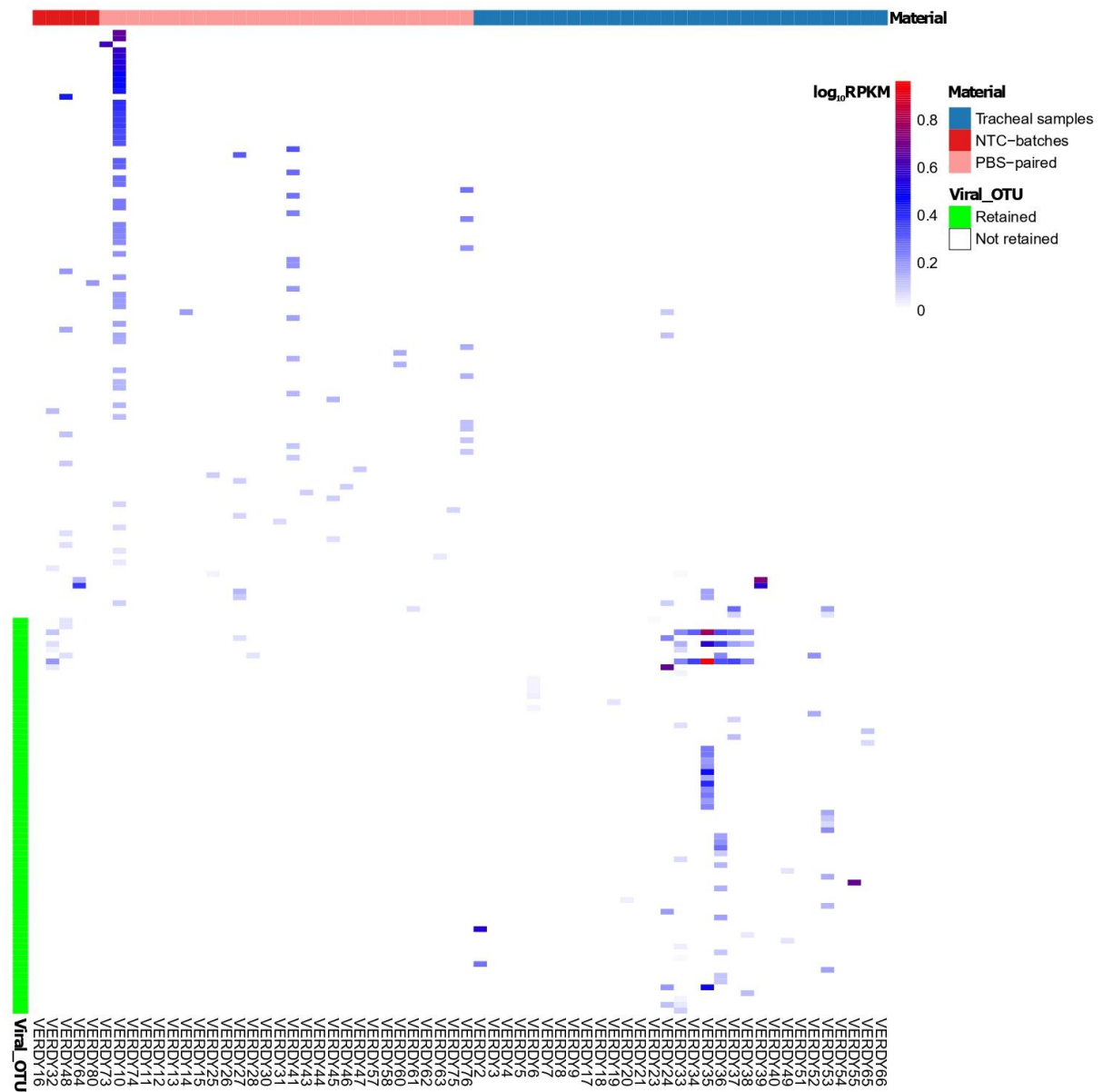
Endotracheal virome and metatranscriptome of preterm neonates at birth in relation with progression to bronchopulmonary dysplasia

Destras G^{1,2}, Anani H^{1,2}, Godbert I³, Burfin G¹, Pastor-Diez B³, Oddoux E¹, Semanas Q¹, Lina B^{1,2}, Picaud J.C³, Josset L^{1,2}

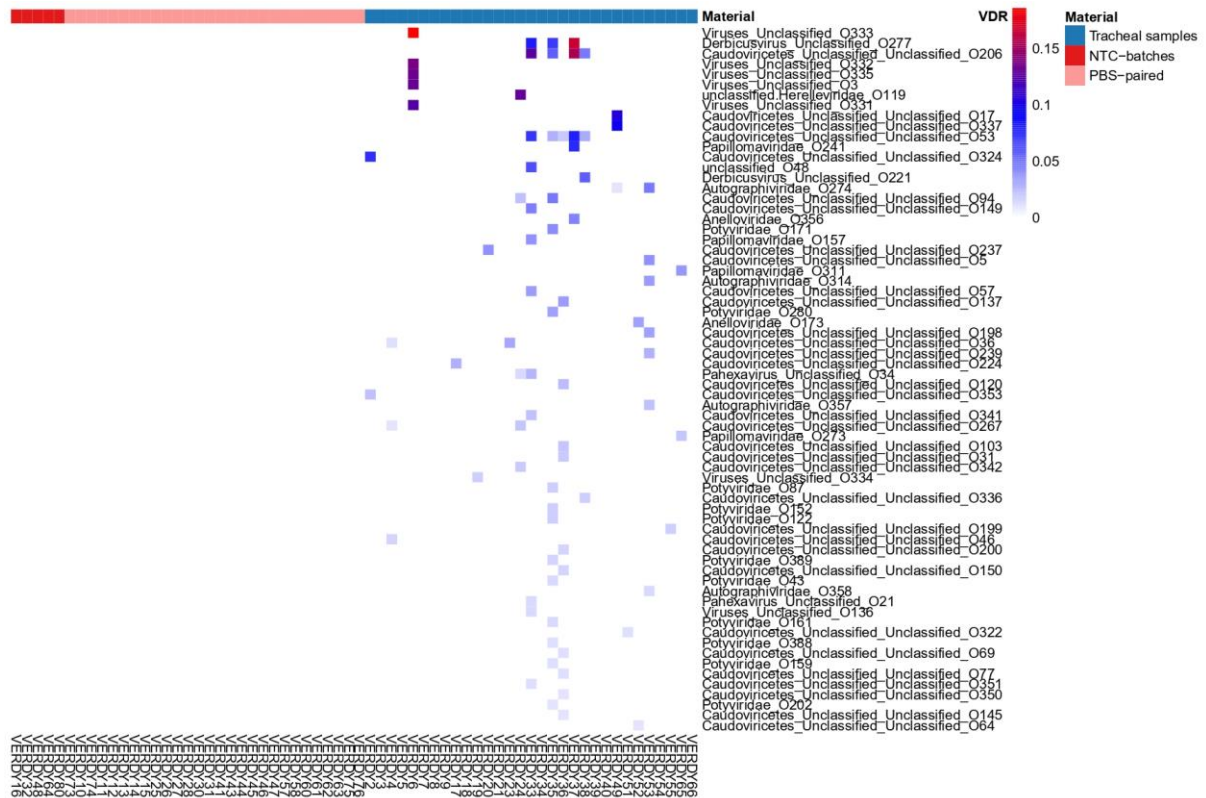
¹Laboratoire de Virologie, Institut des Agents Infectieux, Groupement Hospitalier Nord, Hospices Civils de Lyon, Lyon, France

²Université Claude Bernard Lyon 1, CIRI, Inserm U1111 CNRS UMR5308, Virpath, Lyon, France

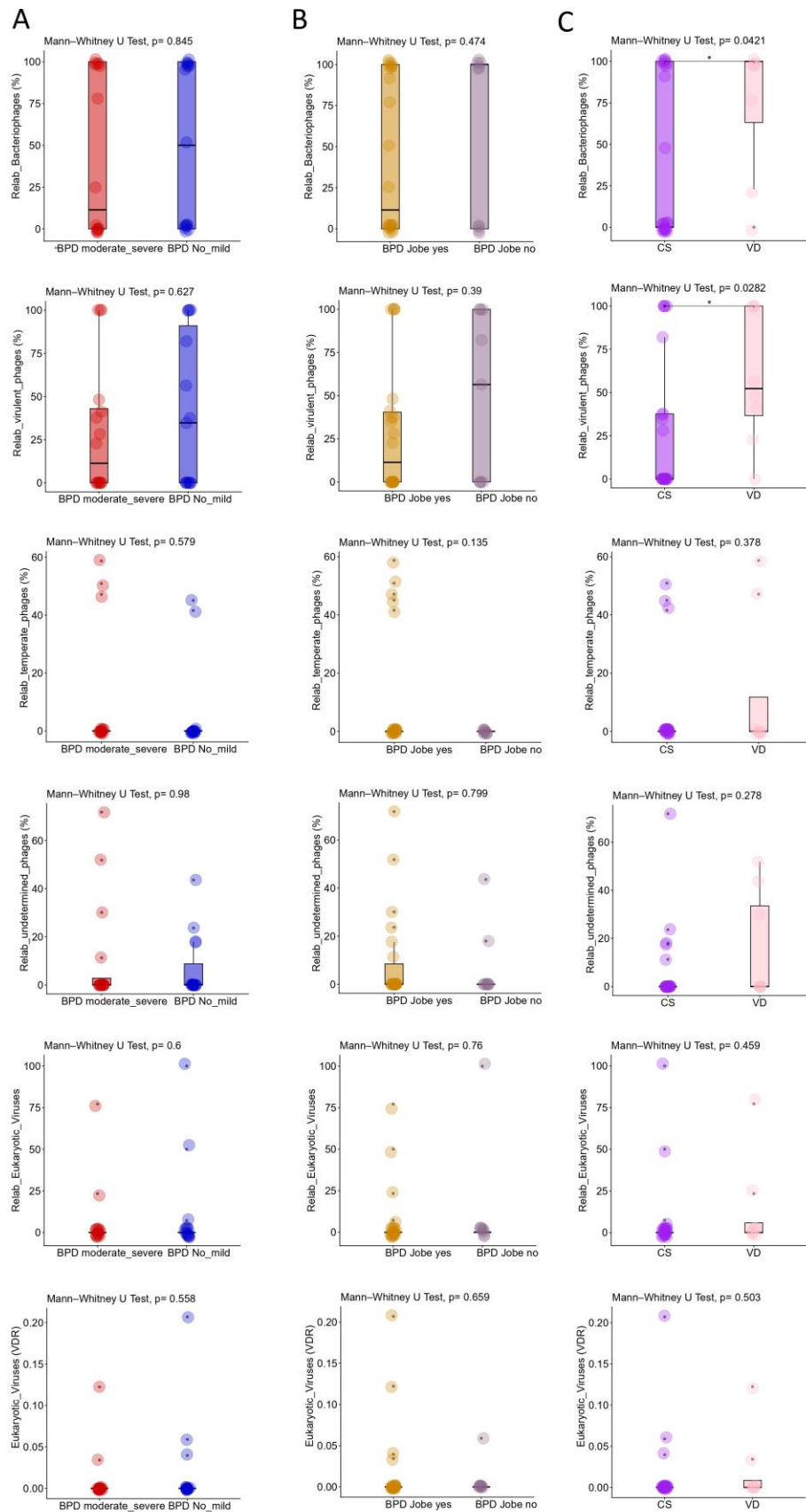
³Service de réanimation néonatale, Groupement Hospitalier Nord, Hospices Civils de Lyon, Lyon, France



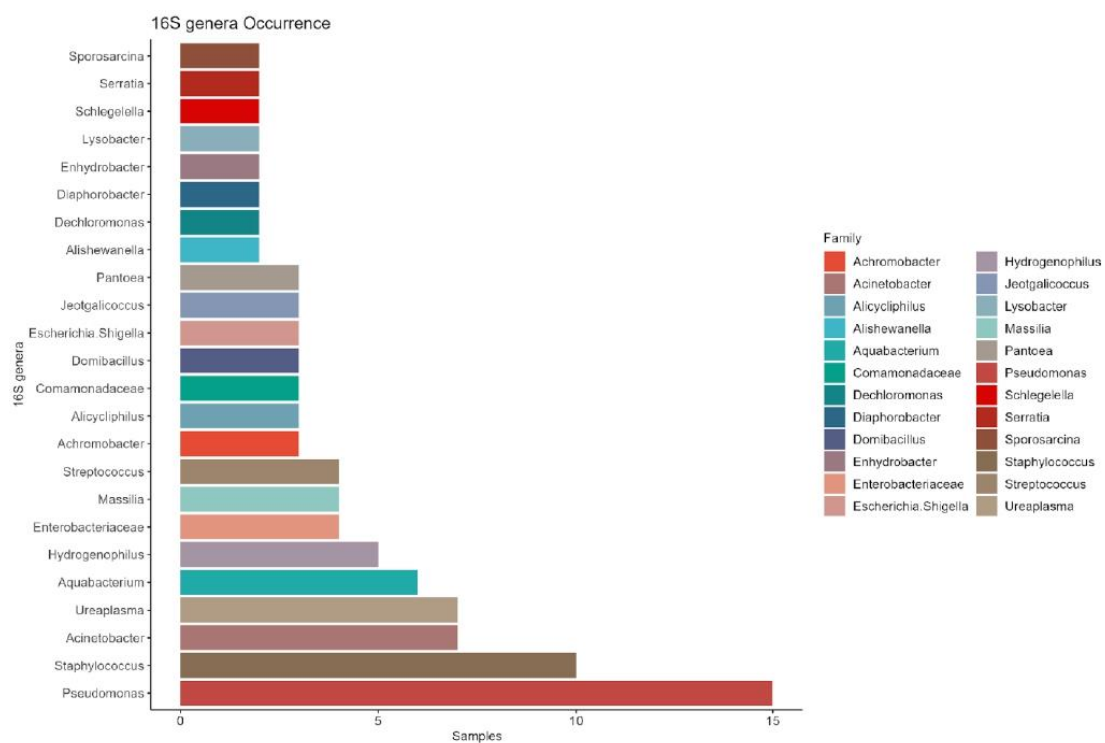
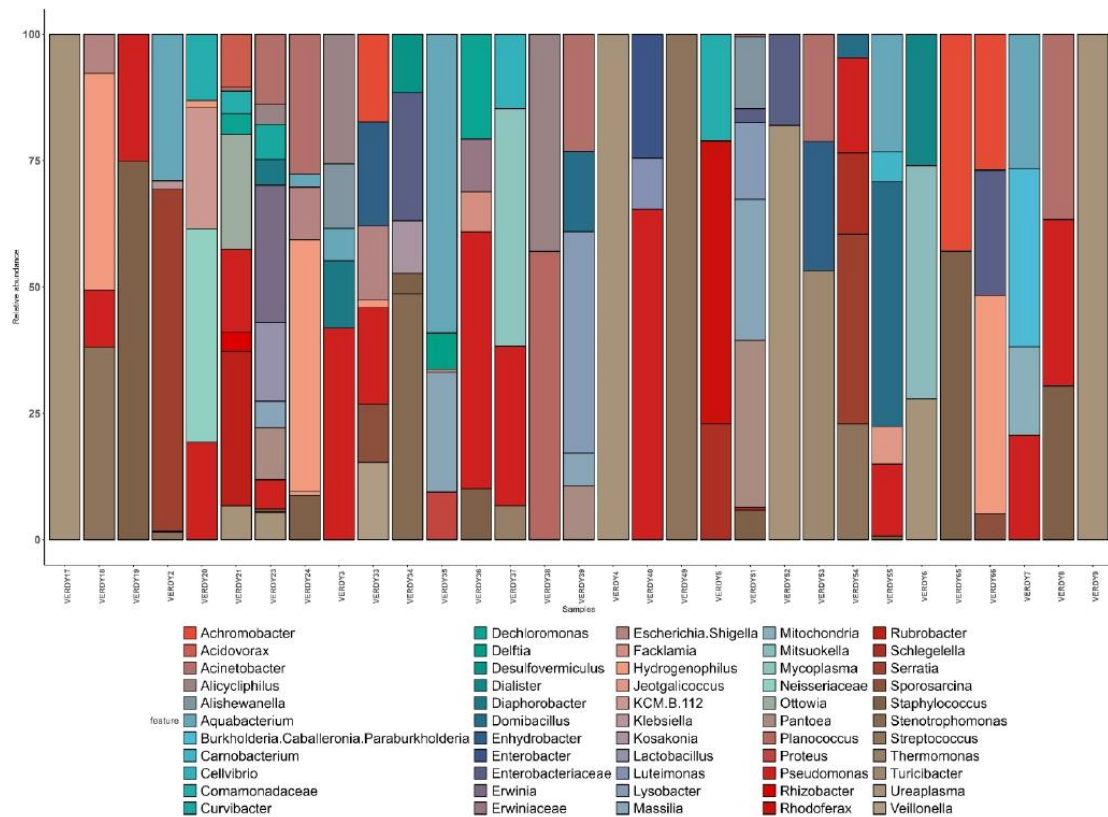
Supplementary Figure 1. Heatmap representing vOTUs identified in the different materials: no-template controls (red), paired PBS controls (pink) and endotracheal samples (blue) after the first decontamination step. The vOTUs retained for analysis are shown in green. The intensity of vOTUs, expressed as RPKM is represented according to a colour gradient ranging from blue (low level) to red (high level).



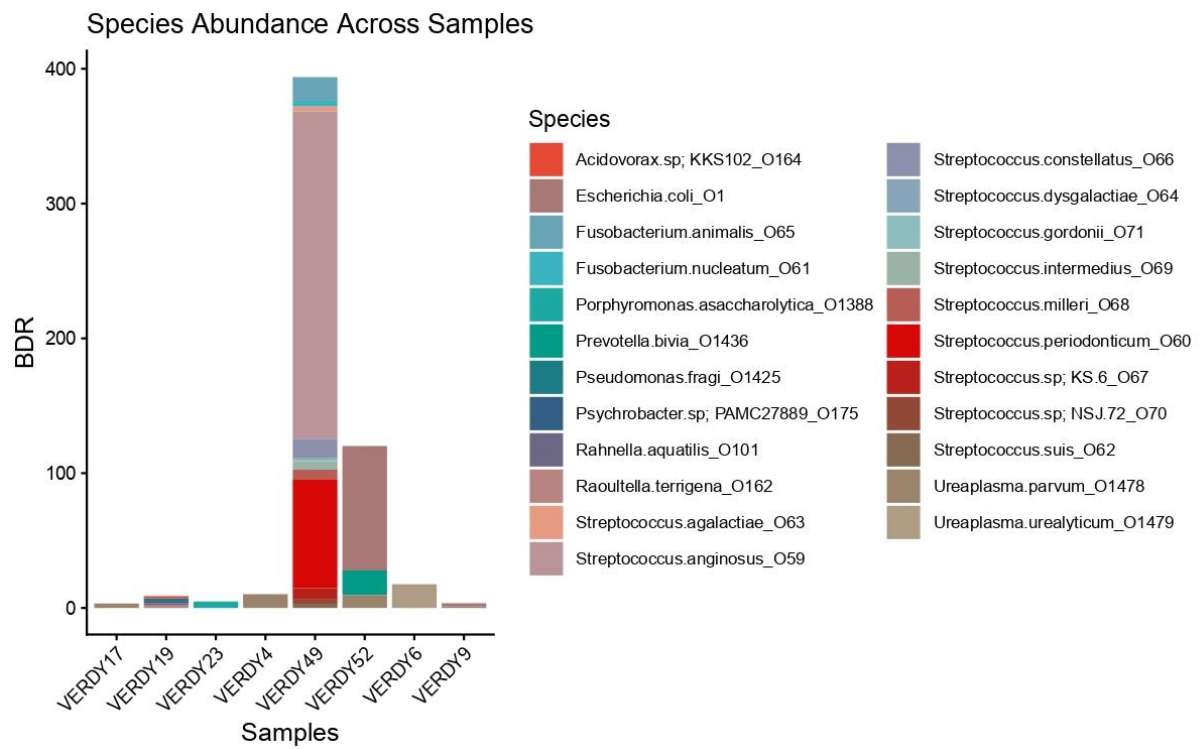
Supplementary Figure 2. Heatmap representing vOTUs identified in the different materials: no-template controls (red), paired PBS controls (pink) and endotracheal samples (blue) after the final decontamination step consisting of subtracting the maximum control signal from each sample. The intensity of vOTUs, expressed as Viral Decontaminated Ratio (VDR), is represented according to a colour gradient ranging from blue (low level) to red (high level).



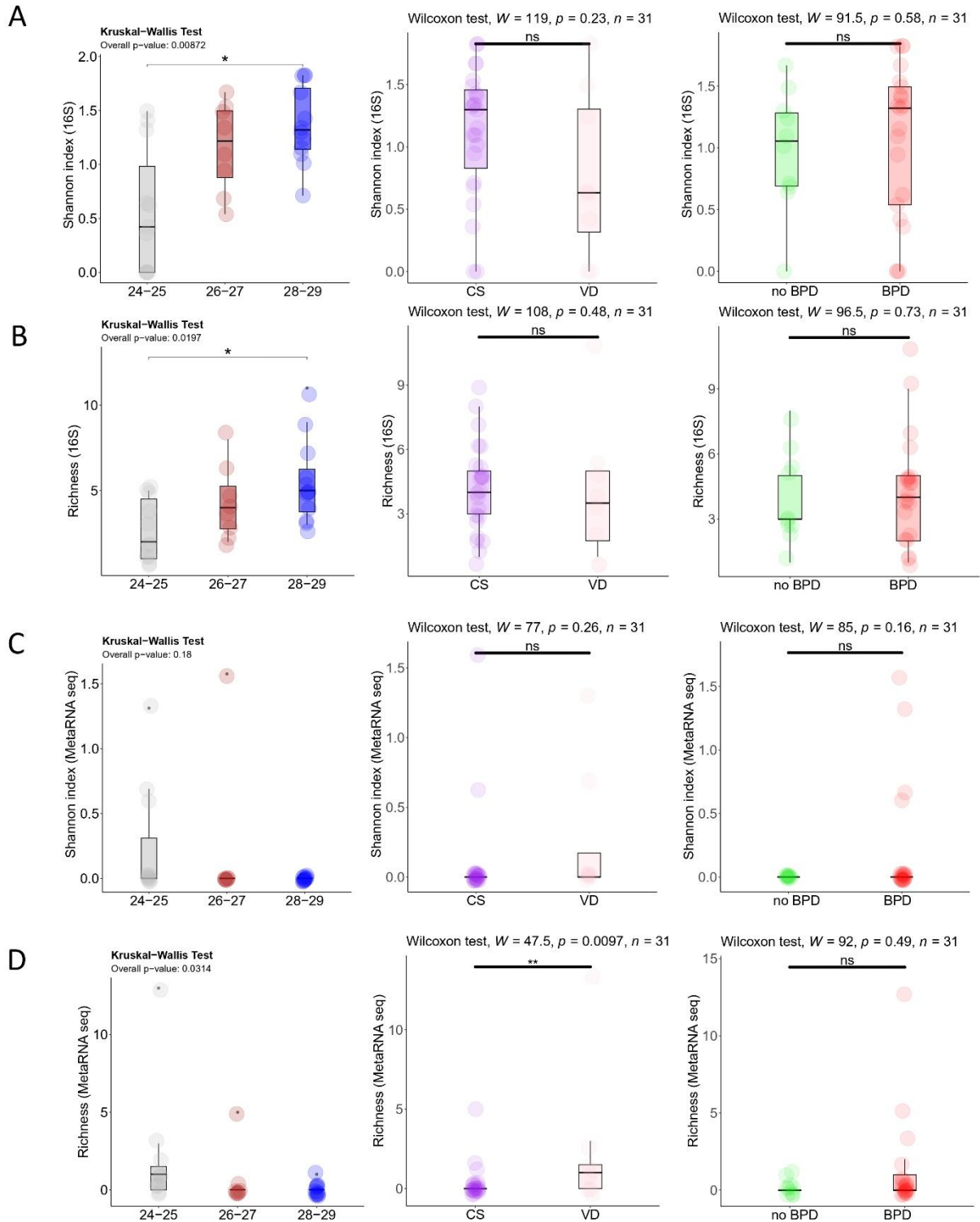
Supplementary Figure 3. Boxplots representing relative abundances of bacteriophages and eukaryotic viruses stratified by BPD status (left), oxygen requirement (middle) and mode of delivery (right). (Mann-Whitney U test: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)



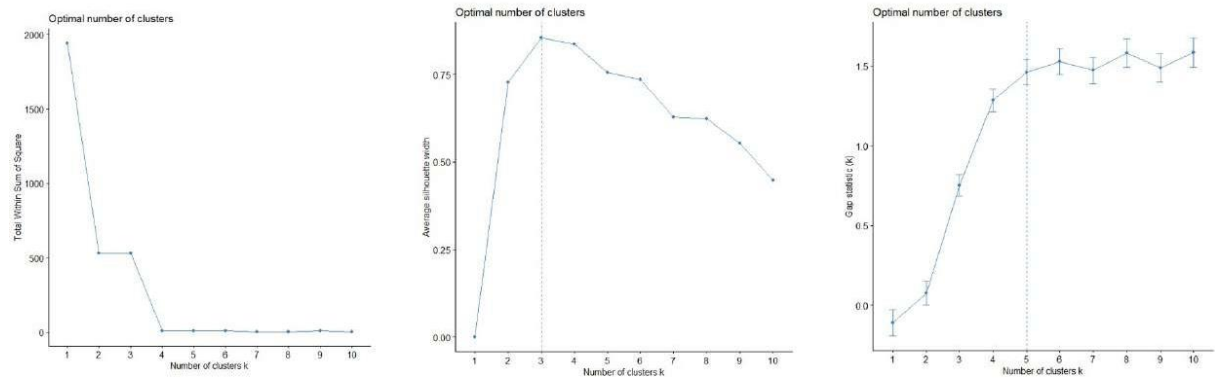
Supplementary Figure 4. Relative abundances (top) and occurrences (bottom) of endotracheal bacterial genera identified via 16S rRNA-seq.



Supplementary Figure 5. Abundance of transcriptionally active bacterial species (Bacterial Decontaminated Ratio) in endotracheal samples.



Supplementary Figure 6. Boxplots of Shannon index and richness according to bacterial genera identified by 16S rRNA-seq (A-B) and RNA bacterial species identified by metaRNA-seq (C-D), and stratified by gestational age (left), delivery mode (middle) and BPD status (right). (Mann-Whitney U test: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)



Supplementary Figure 7. Optimal number of clusters as suggested by Total Within Sum of Square, Average Silhouette width and gap statistic methods.