



FIG S3 Heatmap of the 50 most abundant intestinal bacterial taxa relative abundance in samples collected from 113 preterm infants enrolled in the study. The fecal microbiota was characterized by high-throughput sequencing of the V3-V4 variable regions of 16S rRNA genes. Ward linkage clustering was used to cluster samples based on their Jensen-Shannon distance calculated in *vegan* package in R (1). The number of clusters was validated using gap statistics implemented in the *cluster* package in R (2) by calculating the goodness of clustering measure.

1. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. 2011. *vegan: Community Ecology Package. R package version 2.0-2.*
2. Maechler M. 2016. *cluster: "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al.*