

1

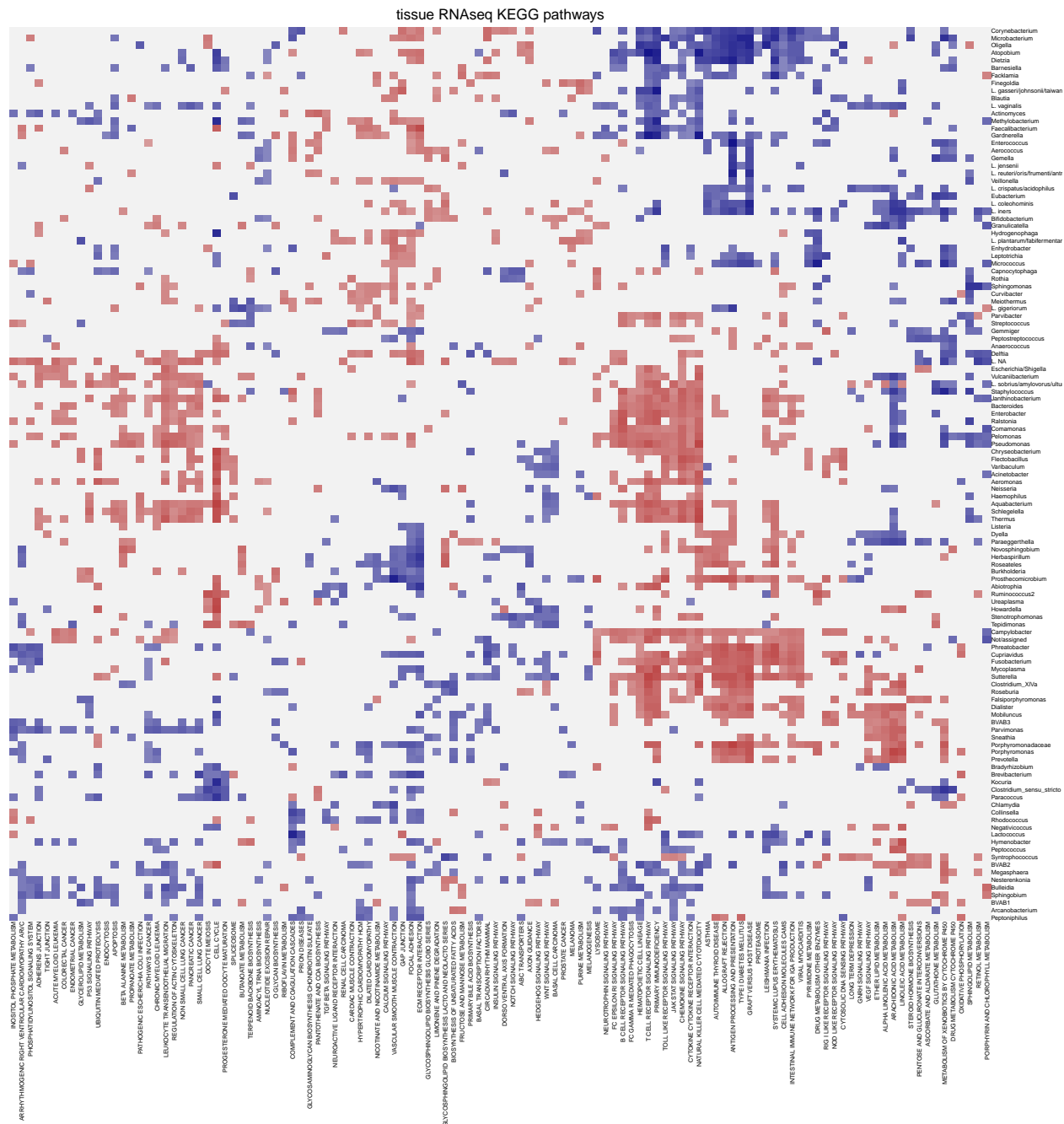


Figure 9. Heatmaps showing the functional association of microbiome datasets with the expression profiles in the RNAseq dataset. Briefly, bacterial abundances from each dataset were correlated with the gene expression of the top 5000 highly variable genes from the RNAseq dataset, generating a correlation matrix between bacteria and genes. Then for each bacteria, we rank genes based on their correlation to that bacteria and perform gene set enrichment analysis (GSEA) using the KEGG gene annotation database. This, in turn, will result in a matrix associating every bacteria with every KEGG process in the tissue. The heatmap shows the normalized enrichment score (NES). Only enrichments with pvalue below 0.05 are shown. Only bacterium and pathways with at least 10 significant NES scores were included in the heatmap.