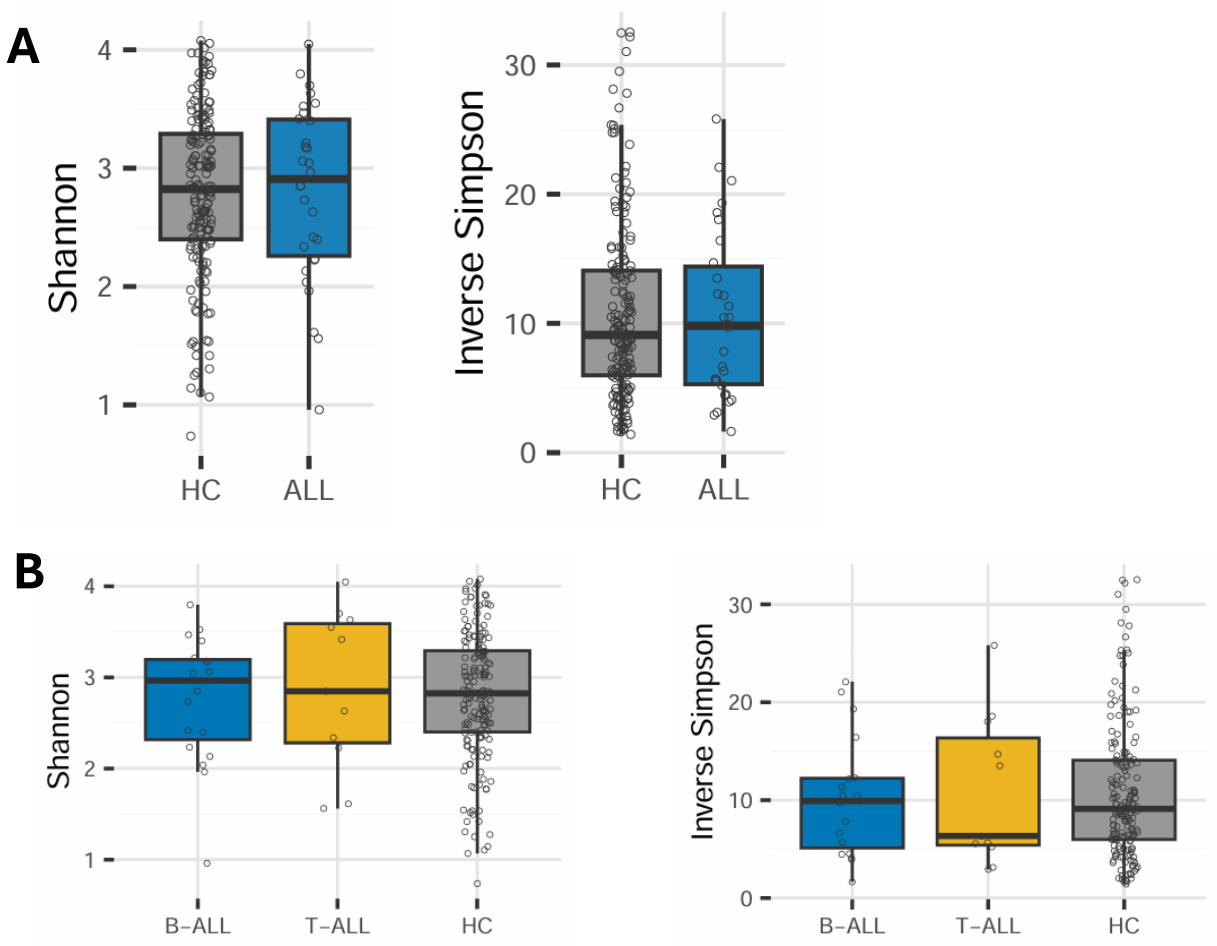
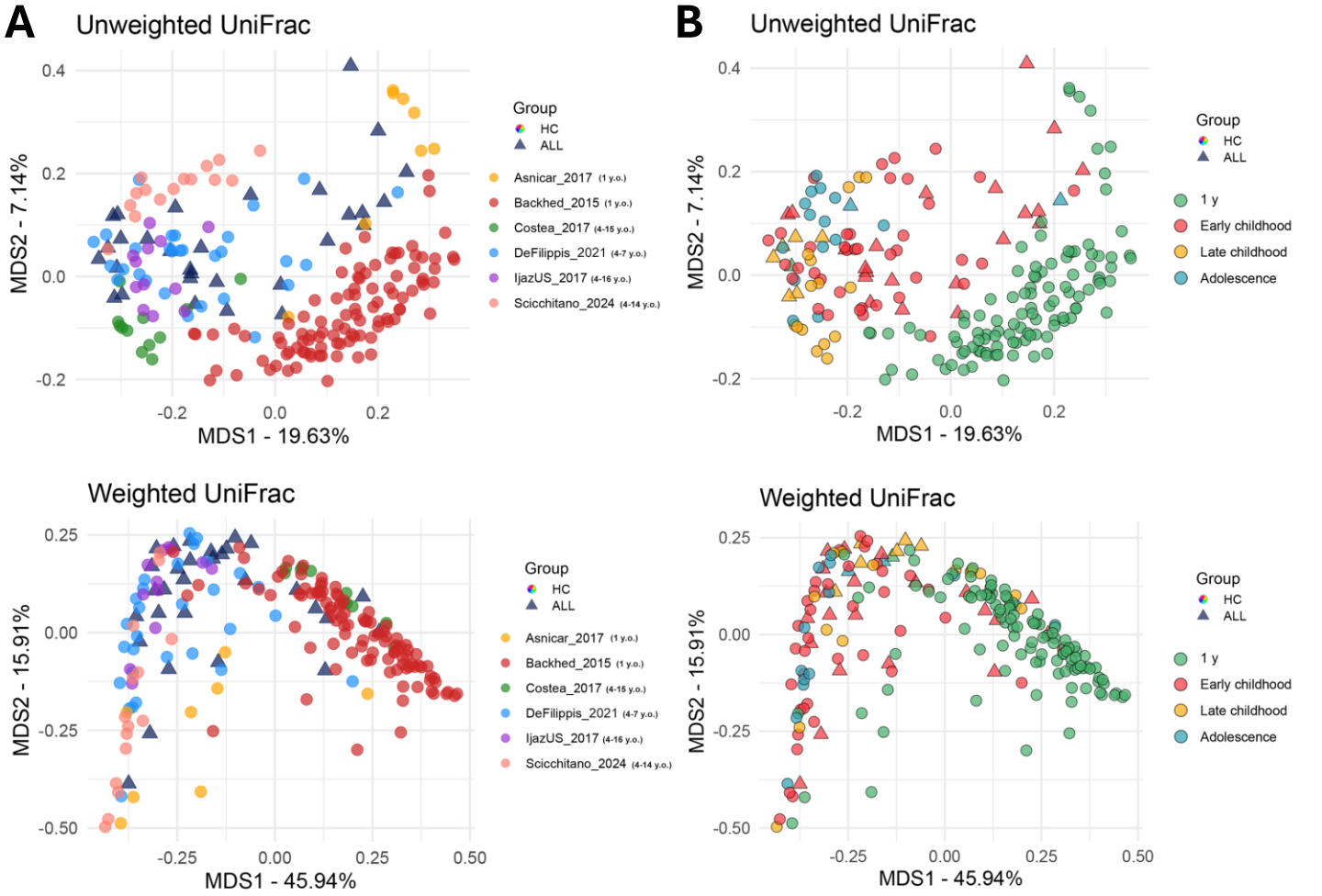


Supplementary Figure 1. Alpha diversity of the gut microbiome in healthy controls (HC) and pediatric acute lymphoblastic leukemia (ALL) patients

(**A**) Boxplots showing Shannon and Inverse Simpson alpha diversity indices in HC and ALL patients at diagnosis. (**B**) Alpha diversity metrics in HC and ALL patients stratified by immunophenotype (B-ALL and T-ALL). Differences between groups were assessed using Wilcoxon rank-sum tests with false discovery rate (FDR) correction for multiple comparisons. No significant differences were observed, indicating that gut microbiota diversity at ALL onset does not differ from that of healthy controls.

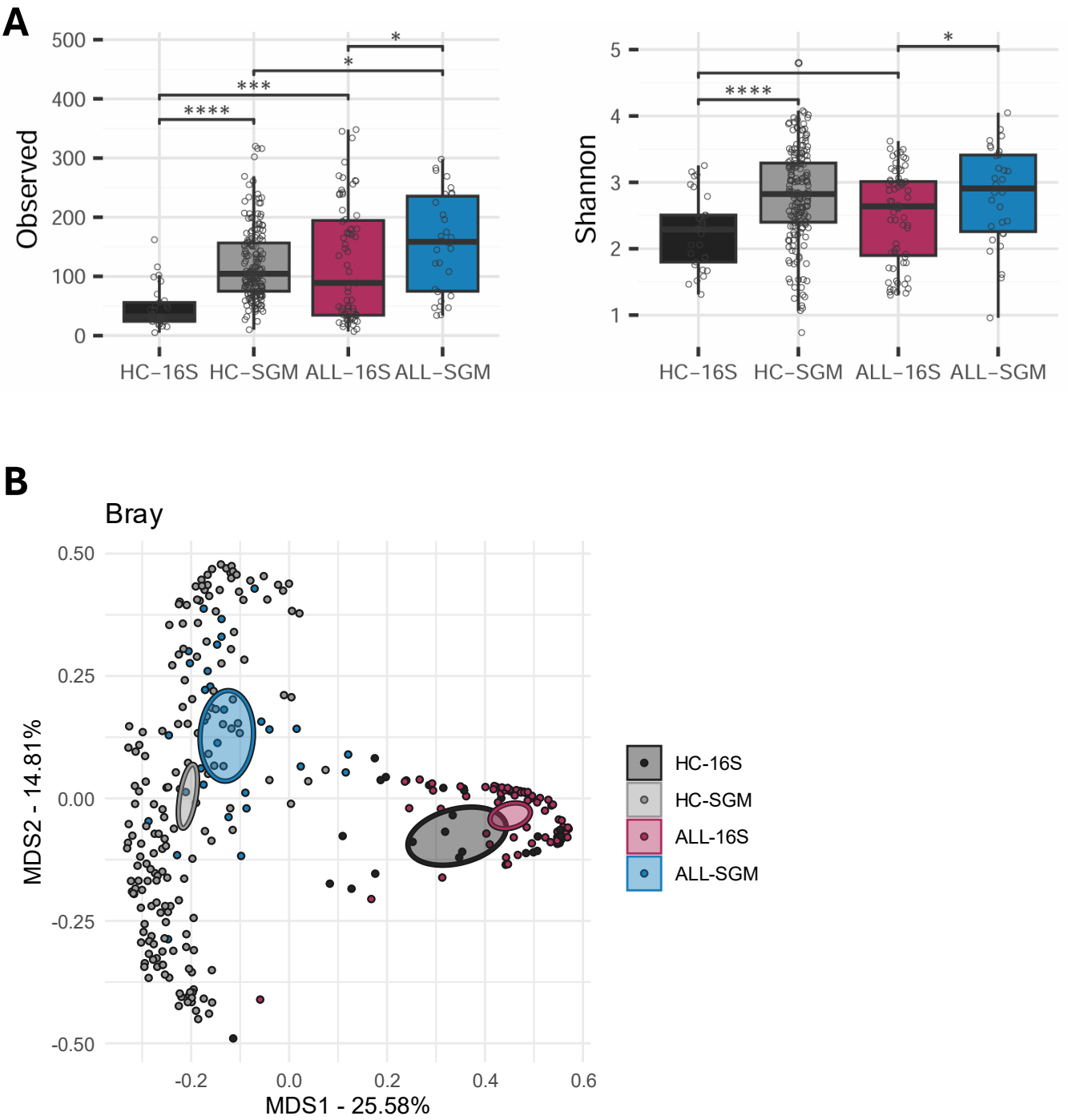


Supplementary Figure 2. Beta diversity principal coordinates analysis (PCoA) stratifying healthy controls (HCs) by cohort and age group

Circles denote HCs, colored according to their respective cohort (**A**) and age (**B**), while triangles indicate acute lymphoblastic leukemia (ALL) patients. The first principal coordinate (MDS1) primarily separates samples by age, especially with the Unweighted UniFrac measurements, demonstrating that clustering is the study and control cohort is mostly driven by age. Age groups have been defined by interpretable bins reflecting developmental stages: 0-12 months: "*1 y*" group; 4-7 yo: "*Early childhood*" group; 8-12 yo: "*Late childhood*" group; >13 yo: "*Adolescence*"

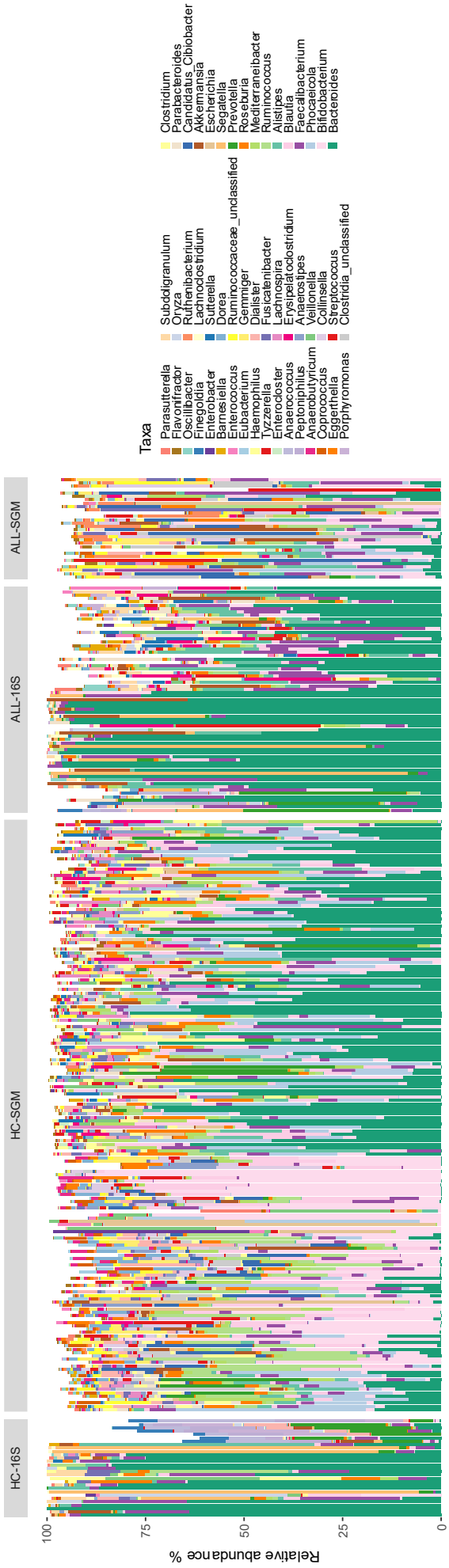
Supplementary Figure 3. Taxonomic contributors to functional pathways in the gut metagenomes of acute lymphoblastic leukemia (ALL) patients.

Stacked barplots show the relative contribution of microbial taxa to HUMAnN-derived pathway abundances (expressed as percentage of total pathway counts, normalized in copies per million reads, CPM) in B-ALL, T-ALL, and healthy controls (HCs). For each pathway, only the top five contributing taxa are displayed.



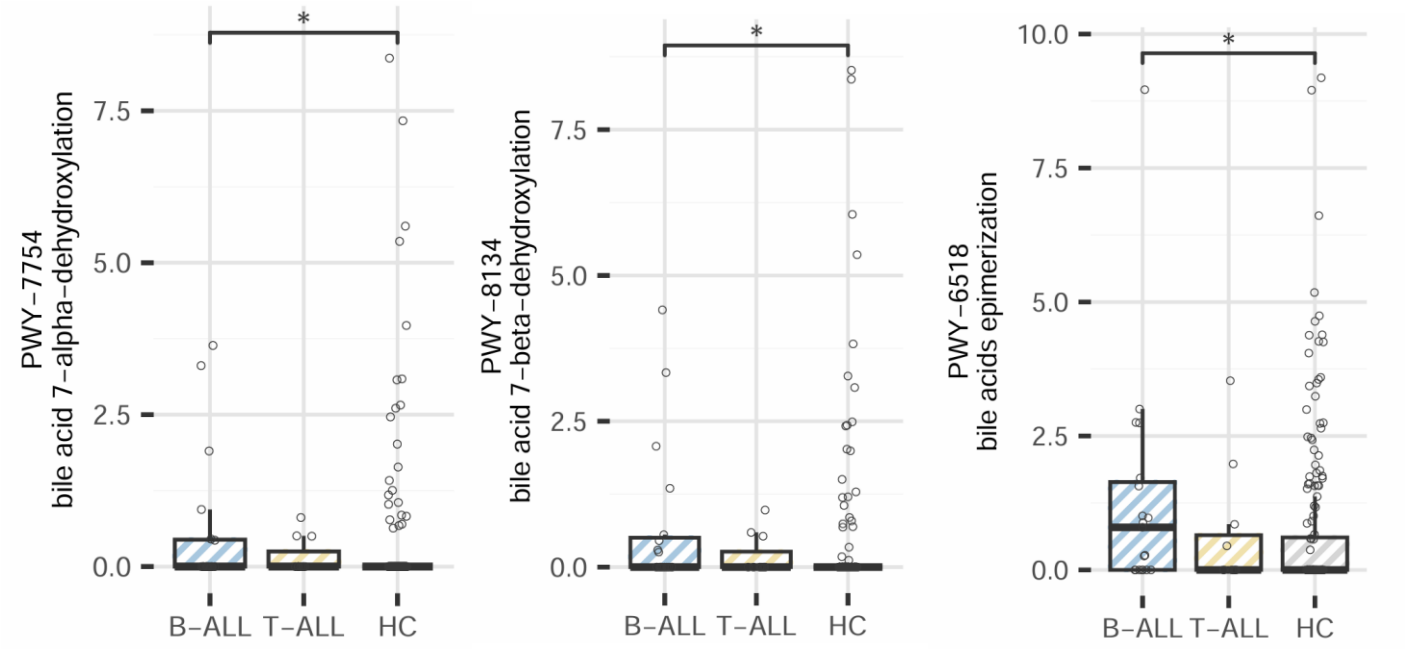
Supplementary Figure 4. Alpha and beta diversity comparison between 16S rRNA and shotgun metagenomics datasets

(**A**) Boxplots showing alpha diversity metrics (Observed features and Shannon indices) across healthy controls (HCs) and acute lymphoblastic leukaemia (ALL) samples, stratified by sequencing technique (16S rRNA vs shotgun metagenomics [SGM]).(**B**) Principal Coordinates Analysis (PCoA) plots based on Weighted UniFrac distances, illustrating clustering patterns by sequencing method and sample type. Ellipses represent 95% confidence intervals for each group's distribution. Together, these plots demonstrate higher alpha diversity and distinct beta diversity clustering in SGM datasets compared to 16S datasets, reflecting the technique-driven nature of diversity estimates.



Supplementary Figure 5. Comparison of genus-level composition between shotgun metagenomics (SGM) and 16S rRNA data in ALL and HCs

Barplot illustrating the relative abundances of the top 50 genera shared by both SGM and 16S rRNA datasets analyzed in this study, facilitating direct comparison of microbial composition across patient grouping and sequencing techniques.



Supplementary Figure 6. Pathway counts for bile acid metabolism in the gut metagenomes of acute lymphoblastic leukemia (ALL) patients and healthy controls (HCs)

Boxplots depict pathway counts (copies per million reads, CPM) for B-ALL and T-ALL subtypes compared to HCs. Pairwise Wilcoxon tests (FDR-corrected) identified significant differences ($p < 0.05$).