📊 **Alpha Diversity Analysis: Aerobic vs Anaerobic Microbiome Samples**

This R script performs **alpha diversity** analysis using the **Shannon index** to compare microbial diversity across aerobic, anaerobic, and uncultured caecal microbiome samples from D7 HMA mice. The analysis includes **non-parametric statistical testing** and visualizes the results with annotated **boxplots**.

📁 **Directory Structure**

Ensure your working directory includes the following input files:

* merged\_abundance\_table.txt — OTU abundance table
  + exported from the tool MetaPhlAn
  + taxonomic labels follow GTDB-style formatting (e.g., d\_\_Bacteria;p\_\_Firmicutes;...)
  + taxonomic lineages should be annotated in clade\_name
  + tab-delimited
* AerobevsAnaerobe\_scraping\_metadata.txt — Sample metadata file,
  + create it based on Supplementary Table 17
  + include grouping variables such as annot1: experimental group (WT/ D2 HMA/ D7 HMA) and annot2: condition (uncultured/ aerobic/ anaerobic)
* Update the setwd() path in the script as needed.

**🔧 Requirements**

Install the following R packages if not already installed:

install.packages(c(

"vegan", "ape", "ggplot2", "RColorBrewer", "stringr", "gridExtra",

"dunn.test", "ggprism", "ggsignif", "multcompView", "tidyverse"

))

🚀 **How to Run**

* Set your working directory inside the R script to the folder where the data files are located: setwd("/path/to/your/project\_directory")
* Run the script AervsAnaer\_alpha\_div.R using your preferred R environment.
* An alpha diversity boxplot (alpha.plot) will be generated and displayed, with statistical annotations.

**🧬 Overview of Analysis**

**Step 1: Load Libraries and Data**

* Loads necessary R packages for data manipulation, statistical analysis, and plotting.
* Loads the OTU table and metadata file.
* Filters OTU data to include only species-level taxonomic entries (lines matching "s\_\_" and excluding "t\_\_").
* Filters metadata to exclude other experimental groups (e.g., "WT" and "D2 HMA").

**Step 2: Alpha Diversity Analysis**

* Computes Shannon diversity index for each sample.
* Performs statistical testing:
  + **Kruskal-Wallis test** to detect differences between groups.
  + **Dunn's test** for post-hoc pairwise comparisons (with **Benjamin-Hochberg correction**).

**Step 3: Visualisation**

* Creates a customized ggplot2 boxplot with jittered points, significance bars, and stylized themes using ggprism.

**📈 Output**

* A **box plot** of Figure 6A, Supplementary Figure 2C or Supplementary Figure 6C is generated, showing the alpha diversity of samples.

**⚠️ Notes**

* The final analysis includes only **D7 HMA** samples.
  + To visualize other experimental groups (e.g., **WT or D2 HMA**),adjust **lines 52–53** in the script.