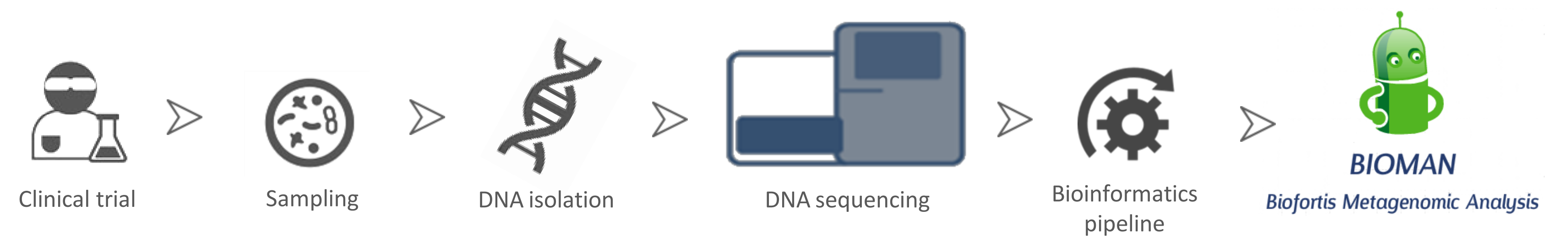
BioMAn™: A user friendly interface for targeted metagenomic data visualization and analysis

Pauline Vaissié, **Christophe Camus, Yao Amouzou, Thomas Carton, Sophie Le Fresne-Languille, Françoise Le Vacon, Murielle Cazaubiel and Sébastien Leuillet.**
Biofortis Mérieux NutriSciences, Saint-Herblain, France

Start with BioMAn

With the recent advances in the field of **next-generation sequencing** (NGS), metagenomics allow to explore the biodiversity of microbiota. Dedicated bioinformatic pipeline focusing on **targeted metagenomics** (such as 16S rRNA) provides to biologists the bacterial composition of the samples. Faced with this large amount of data, their **visualization and appropriate statistical analysis** are essential for scientists to adequately explore and interpret their experiments.



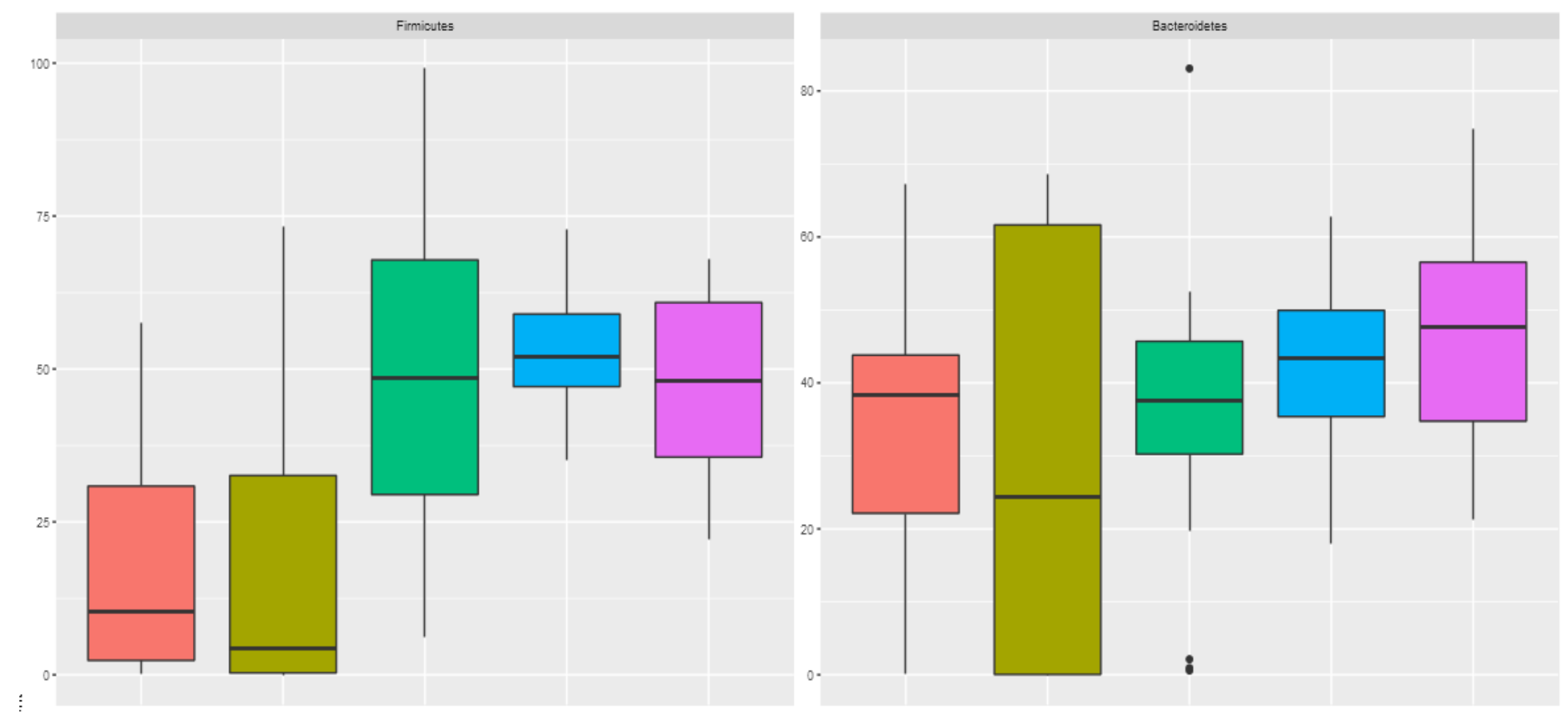
In this context, Biofortis has developed an R [1] and Shiny [2] web based platform called **BioMAn^{TM}** (Biofortis Metagenomics Analysis) which mixes the statistical power of dedicated R packages (metagenomeSeq, mixOmics…) and a user friendly web design.

This interface allows users to **interactively** look into their project by **manipulating**, **filtering** or **gathering** information for further interpretation or communications purposes.

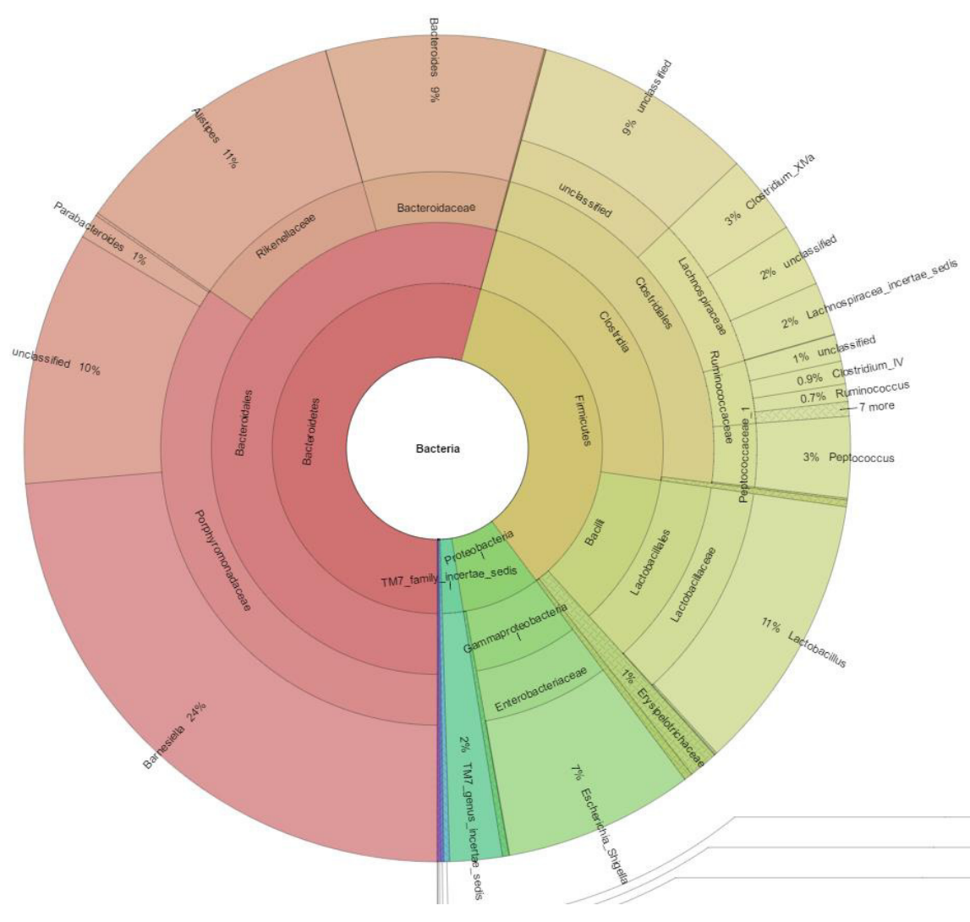
BioMAn^{TM} is deployed on a **Shiny Server Pro**, implemented by a **secure health data hosting** provider according to the French regulatory requirements, to protect the **confidentiality**, **integrity** and **availability** of patient and user data.

Data visualization

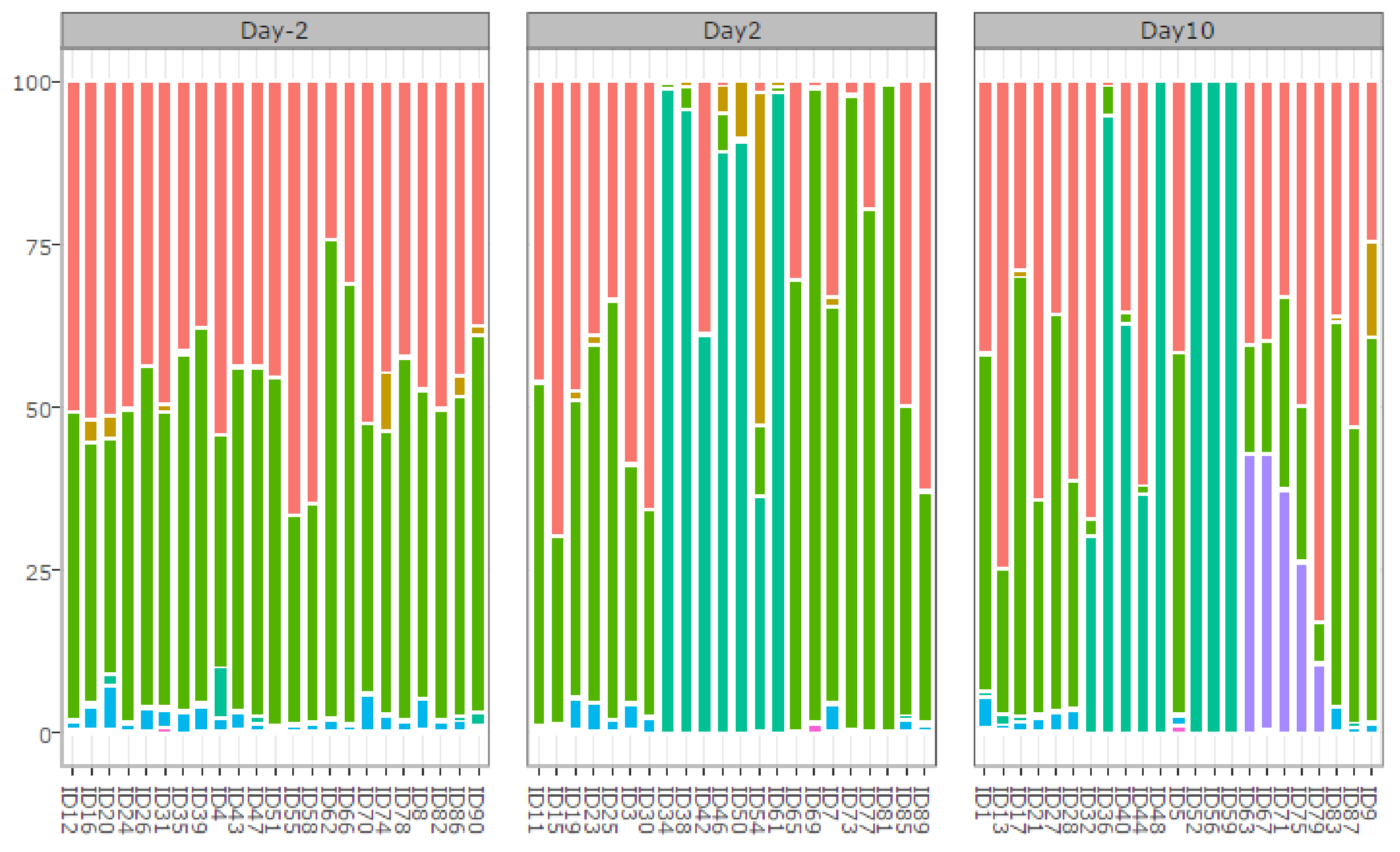
Seven tools are available **to interactively visualize taxa composition** of samples (or grouped samples) at different ranks: barplot, heatmap, boxplot, pie chart, sunburst, Krona chart [3] and interactive hierarchical tree. Most of them are created using ggplot2/plotly or javascript and are **user customizable** through several graphical parameters such as colors, fonts, …



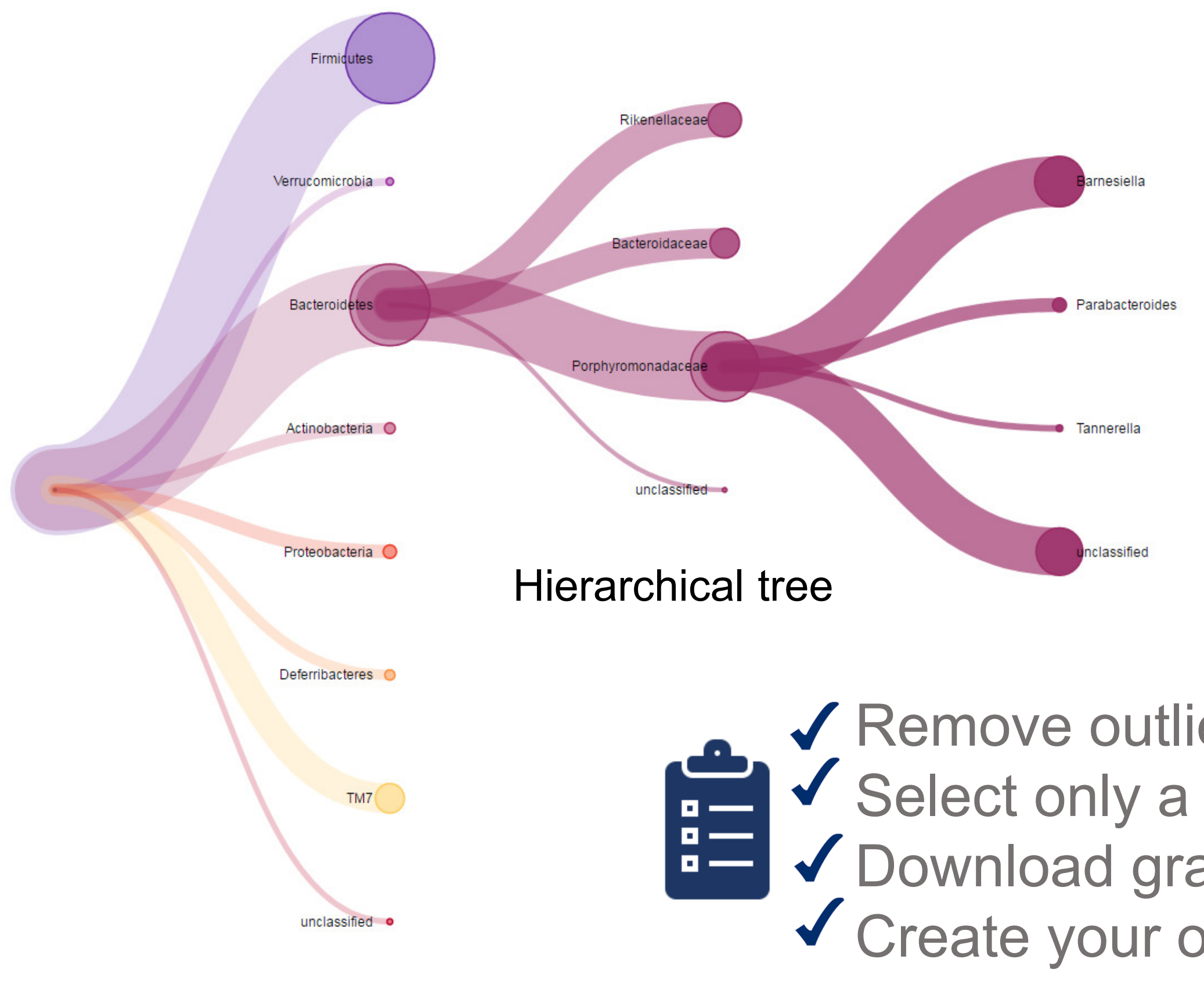
Boxplot



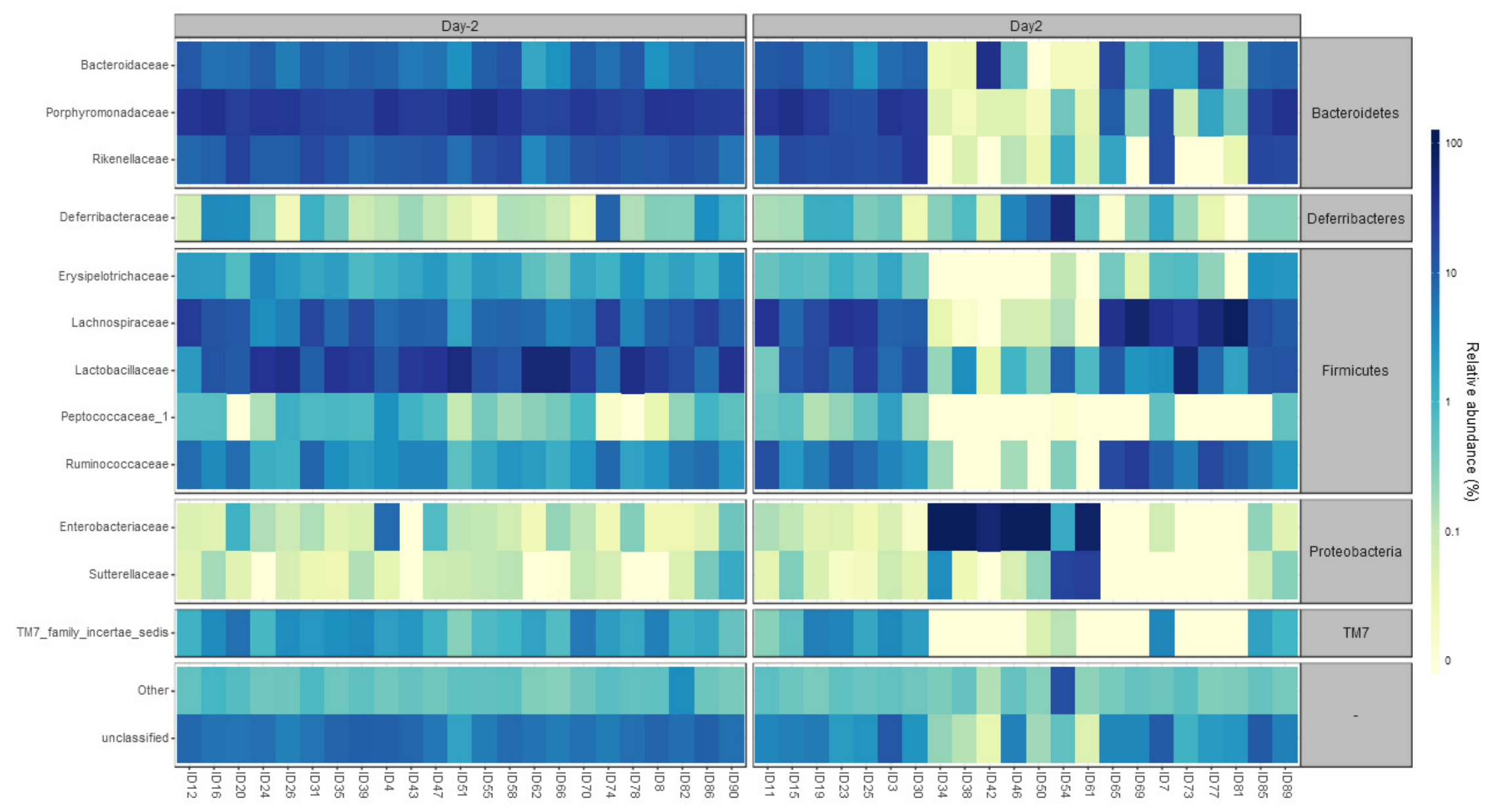
Krona plot



Barplot



Hierarchical tree



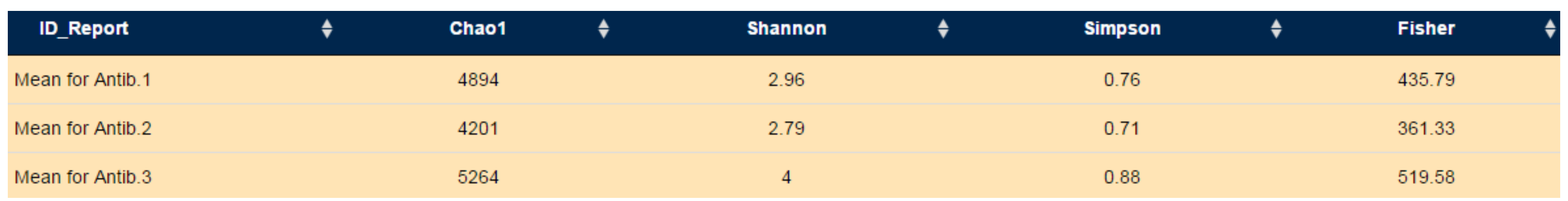
Heatmap

- ✓ Remove outliers
- ✓ Select only a sub-samples according to metadata properties
- ✓ Download graphics in high quality / tables in Excel format
- ✓ Create your own report by adding graphics to a custom PPT template

Diversity and statistical analysis

α and β diversity

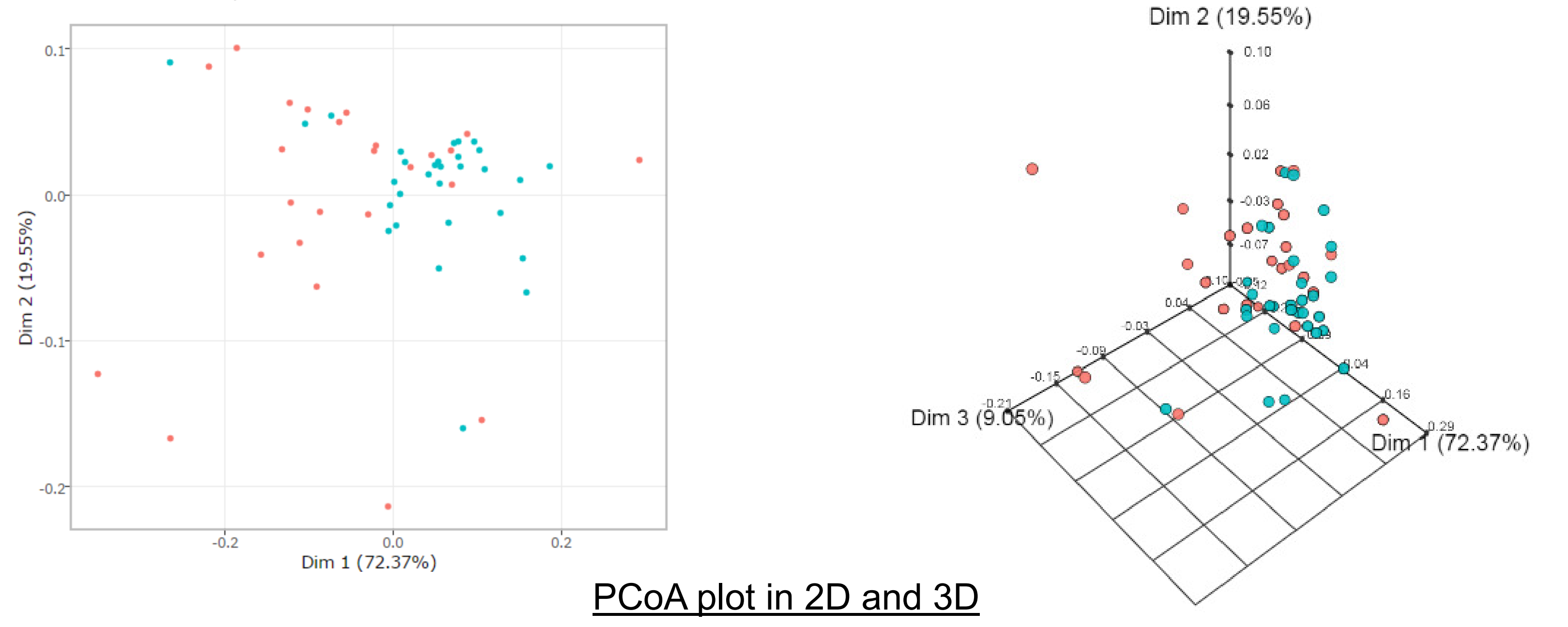
Diversity of each sample (or α diversity) can be expressed using **several indices** including Shannon, Chao, Simpson and Fisher indices.



ID_Report	Chao1	Shannon	Simpson	Fisher
Mean for Antib.1	4894	2.96	0.76	435.79
Mean for Antib.2	4201	2.79	0.71	361.33
Mean for Antib.3	5264	4	0.88	519.58

Graphical representation of the α diversity index of interest can be generated by a simple click on the table.

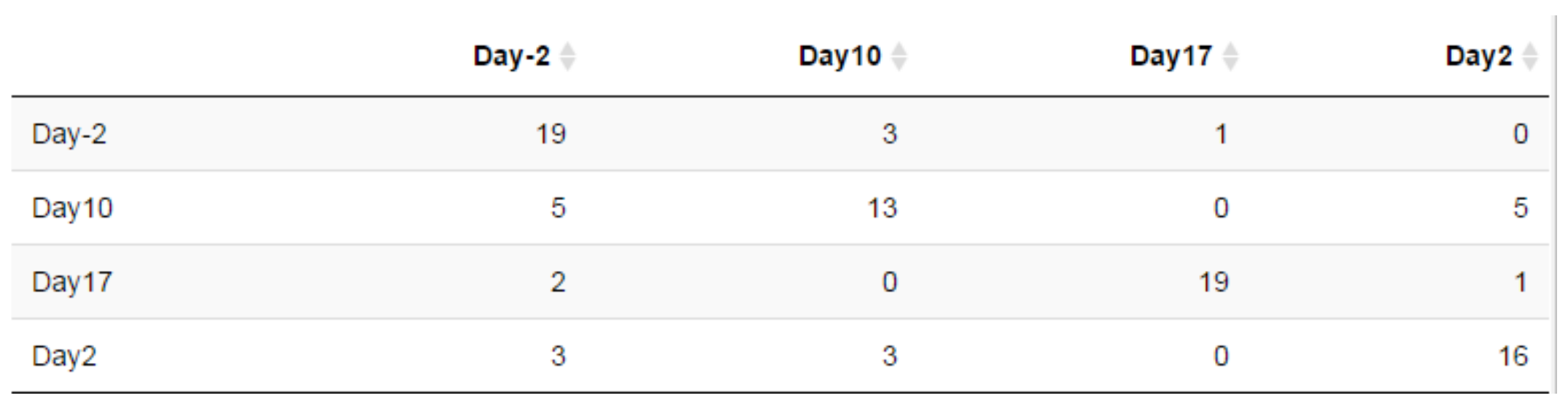
Visualisation of **β diversity** (which represents **distances/dissimilarities in taxa composition** between samples or group of samples) is available through **PCoA** map, using 2D/3D representations. All usual distance methods are available: Bray-Curtis, Jaccard,…



PCoA plot in 2D and 3D

Statistical Analysis

A first approach was implemented into BioMan in order to **predict** the sample belonging to metadata groups by taxa information.
First, a **PLS-DA** (Partial Least Squares – Discriminant Analysis) coupled with an analysis of VIP (Variable Importance in Projection) is conducted to select taxa of interest for discrimination. This step is particularly important to reduce the total number of predictive variables for further steps. A calculation of VIF (Variance Influence Factor) is also conducted to identify taxa involved in strong multicolinearity into the data.
With the retained variables, a **discriminant analysis** is then conducted, in predictive and exploratory purposes.



	Day-2	Day10	Day17	Day2
Day-2	19	3	1	0
Day10	5	13	0	5
Day17	2	0	19	1
Day2	3	3	0	16

Confusion matrix

In the **next version of BioMAn**:
- ✓ Implementation of PERMANOVA
- ✓ Differential analysis: DESeq / edgeR
- ✓ LEfSe analysis

…

[1] R Core Team. R:A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2013.
[2] Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson (2017). shiny: Web Application Framework for R. R package version 1.0.0. https://CRAN.R-project.org/package=shiny
[3] Ondov, Brian D., Nicholas H. Bergman, and Adam M. Phillippy. "Interactive metagenomic visualization in a Web browser." BMC bioinformatics 12.1 (2011): 385.

For more information, please contact biofortis-contact@mxns.com or visit us at www.biofortisinnovation.fr