

MATrIX : a Shiny application to explore microarray data

Franck Soubès¹, Yannick Lippi¹

¹GeT-TRiX - Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-Purpan, UPS, Toulouse, France

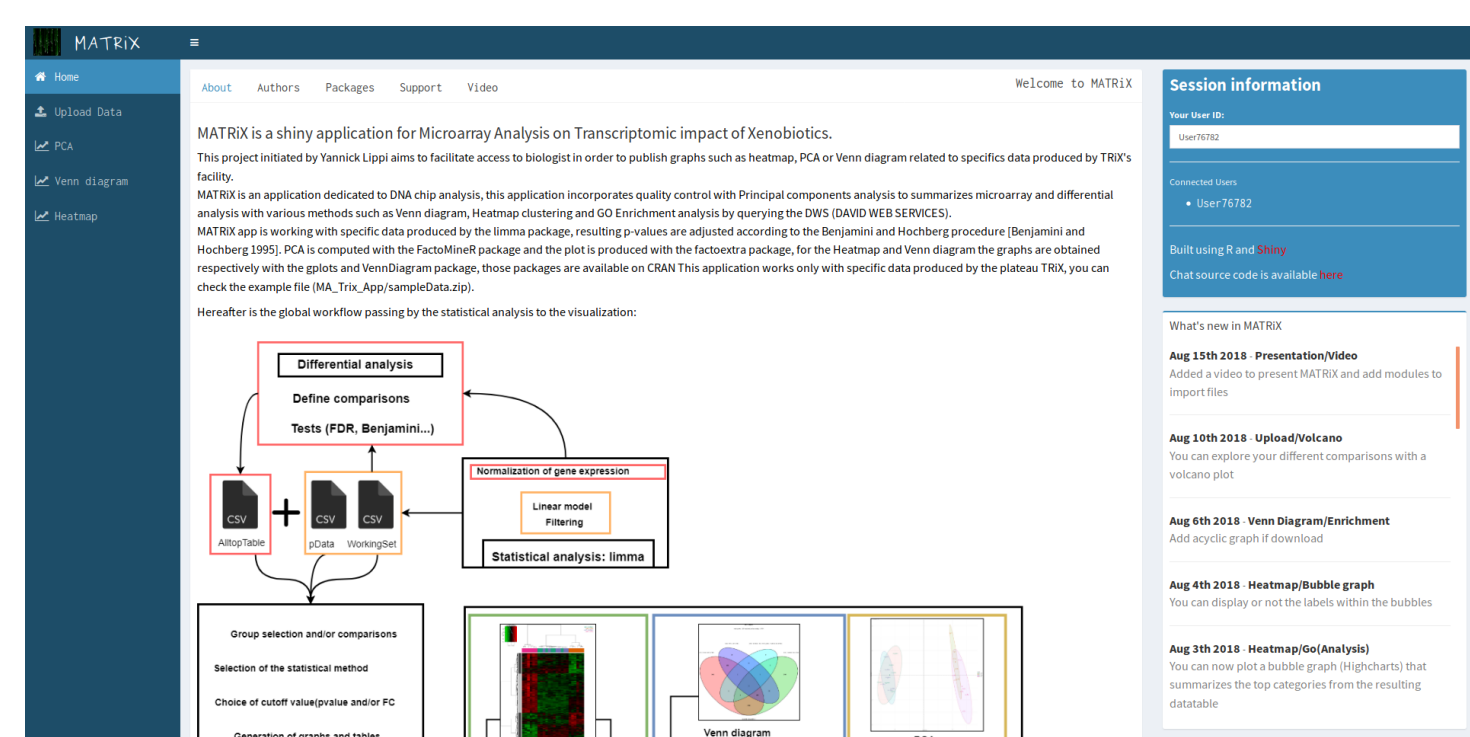
Context

Transcriptomic is a common tool to assess **differential gene expression** in order to compare them with multiple conditions.

The GeT-TRiX facility offers a service for acquisition and statistical analysis of transcriptomic data on **DNA chips**. The data analyses are performed using **R scripts** from raw data pretreatments to statistical analysis (Bioconductor [1] packages).

It is noted that statistical analyses **define arbitrary filtering criteria** to select relevant genes. Hence, biologists are led to reconsider these criteria according to their own scientific question. Thus, the data generated are **difficult to exploit** and for people not familiar with bioinformatics.

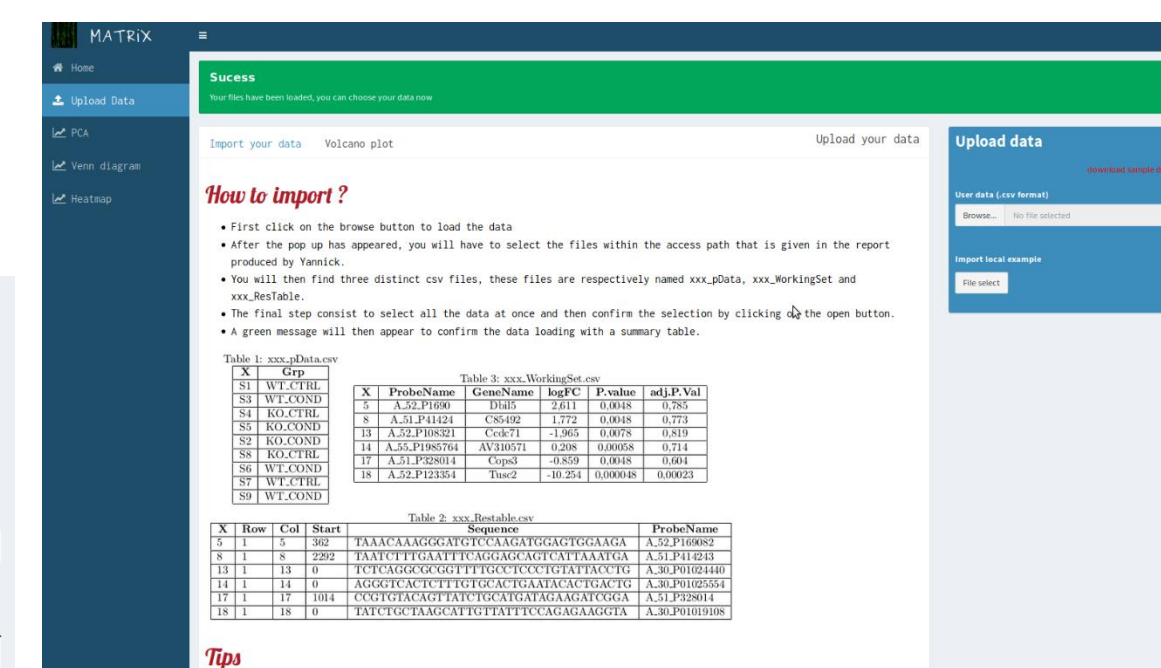
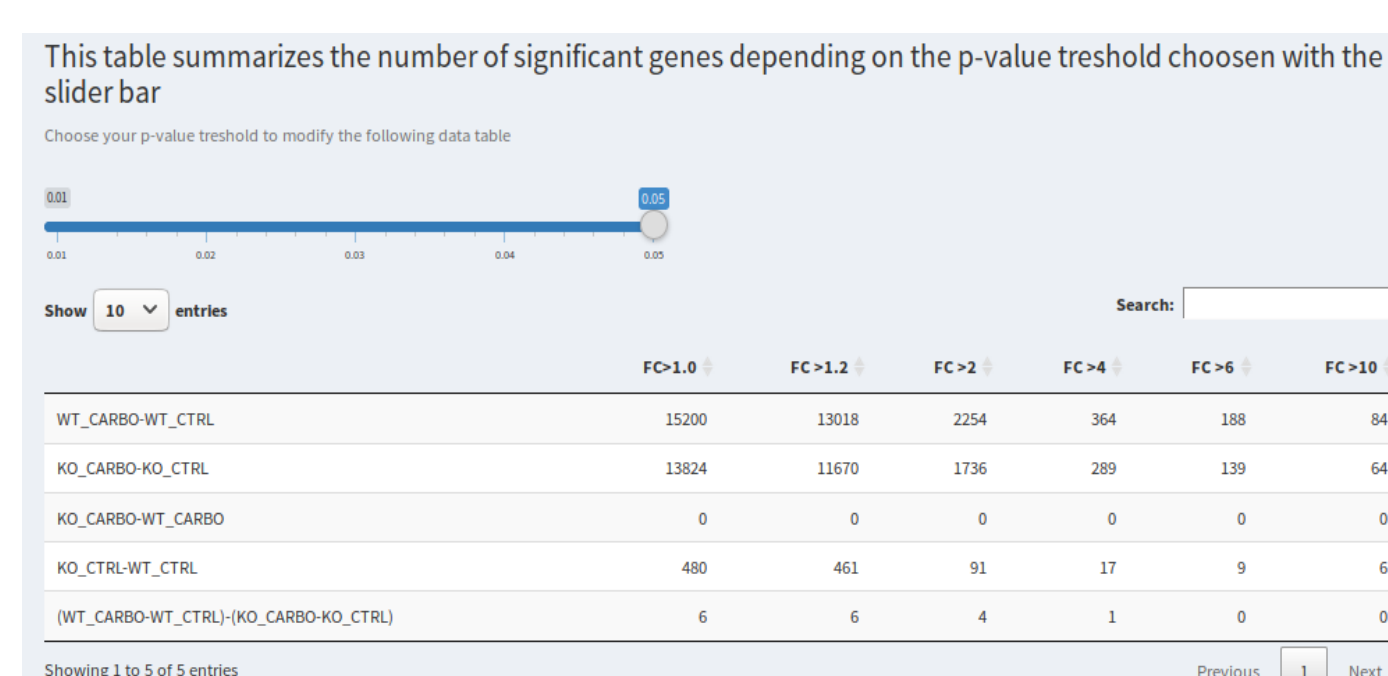
In order to allow biologists to **explore** their datasets in **autonomy** and generate graphical representations, we have deployed a **turnkey** solution built with the Shiny framework [2]



This open source project is available at:
https://forgemia.inra.fr/GET_TRiX/matrix

Import data

- MATrIX needs 3 distinct csv files as input: a design table, an expression table and a statistical table.
- These data are supplied through GeT-TRiX's analysis workflow.
 - A validation message appears to validate the import.
 - Then the HTML procedure is hidden.

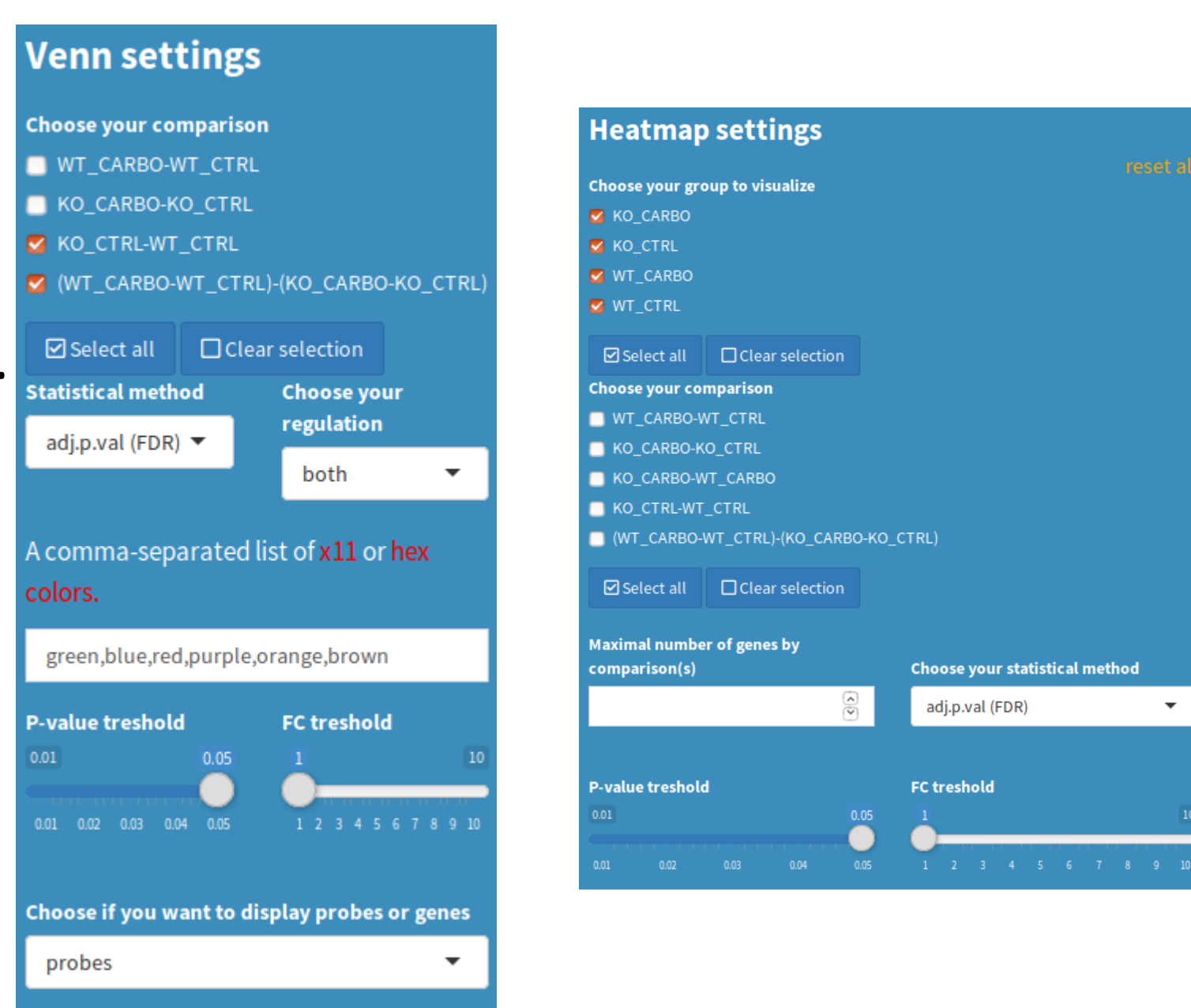


- A summarize table is displayed showing the number of significant genes depending on the thresholds (pvalue, logFC)

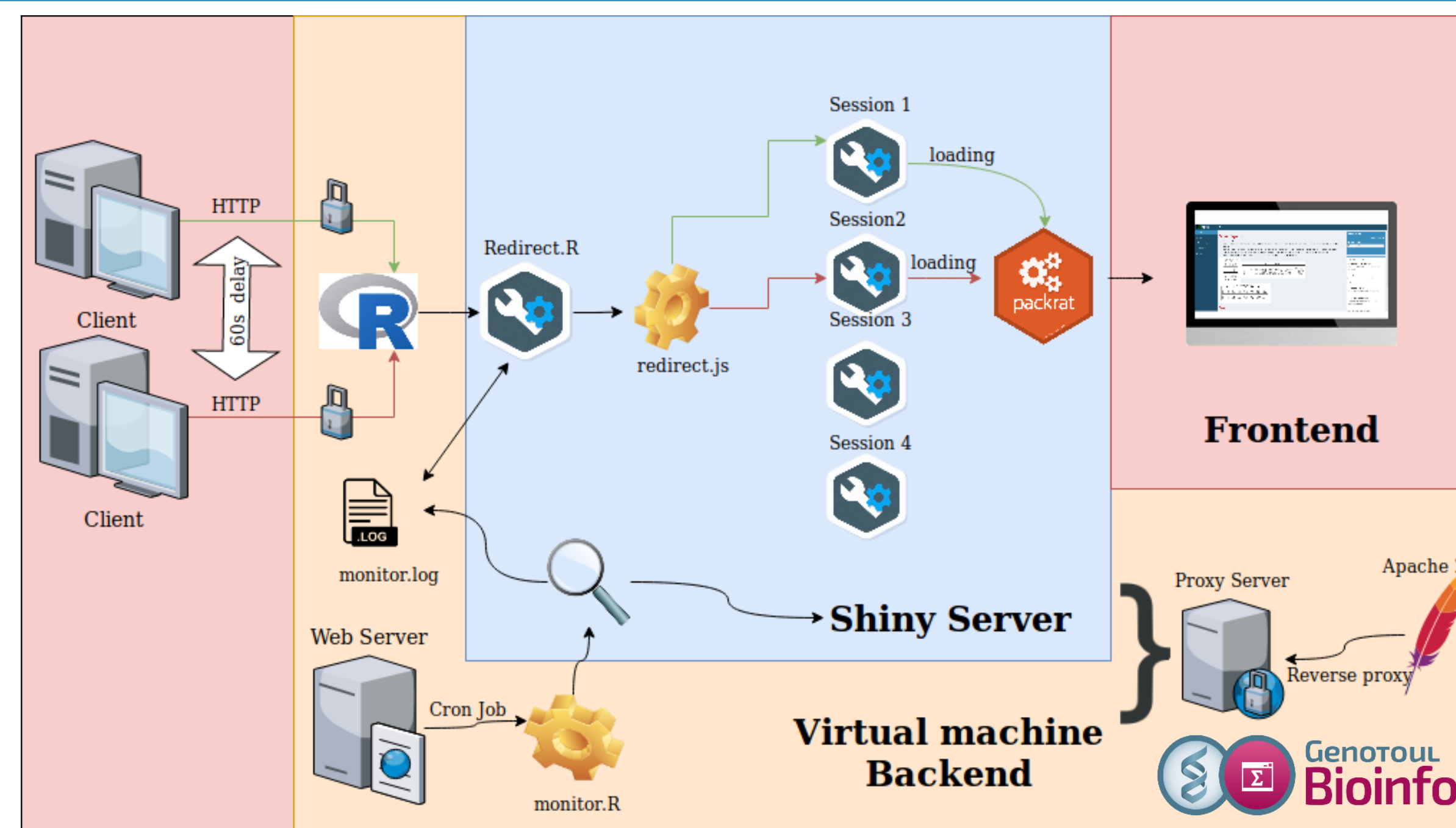
User interface

MATrIX is responsive, user friendly and offers multiple functionalities:

- Group selection and/or comparisons.
- Selection of the statistical method (pval, adj pval).
- Choice of cutoff value (p-value and/or FC)
- Generation of graphs and tables.
- Queries the DAVID database.
- Export of data (csv,xlsx,png,eps,svg,pdf).
- Graphics customization with widgets.
- Graphics interaction (js libraries).
- Filtering data in tables.
- ...

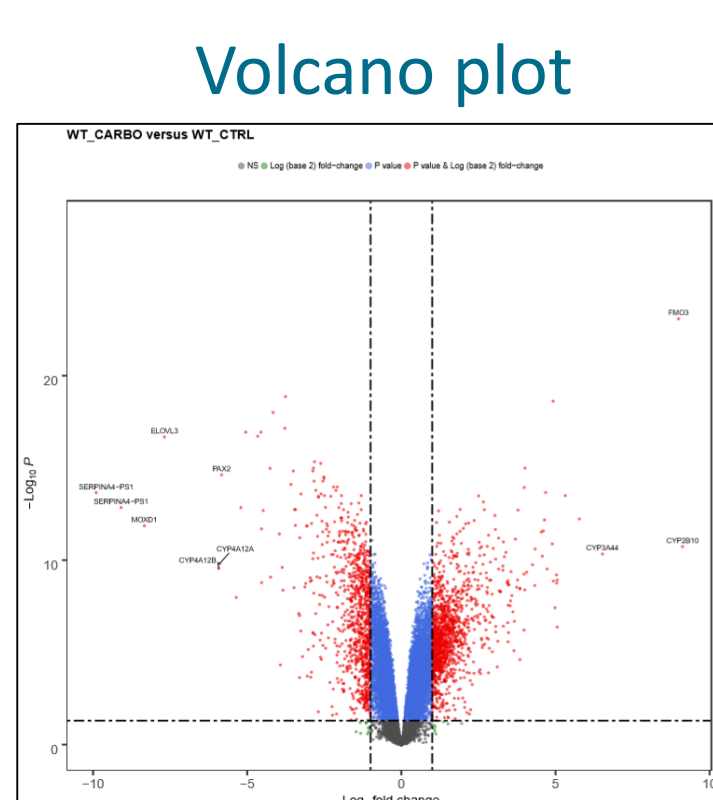
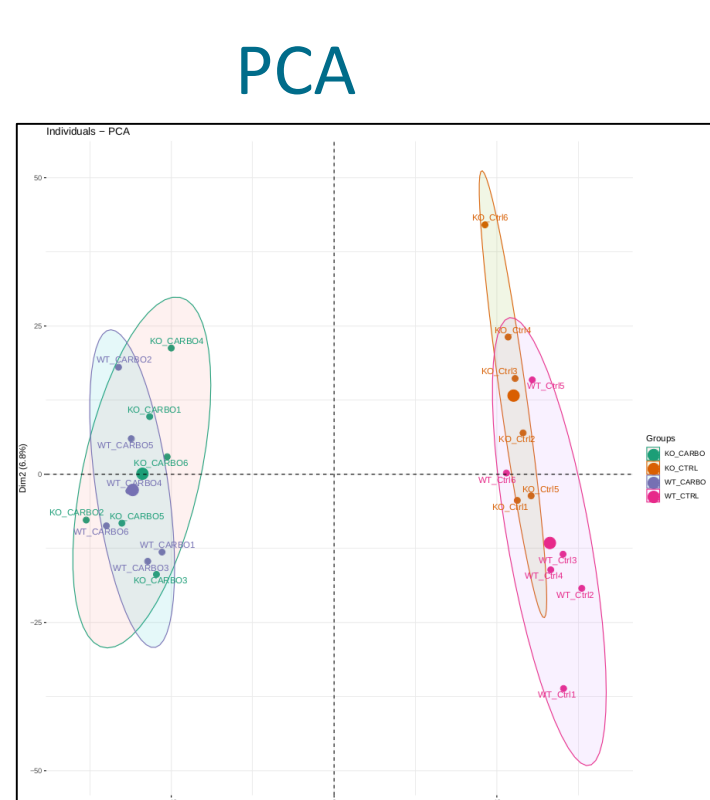


Deployment of the web service

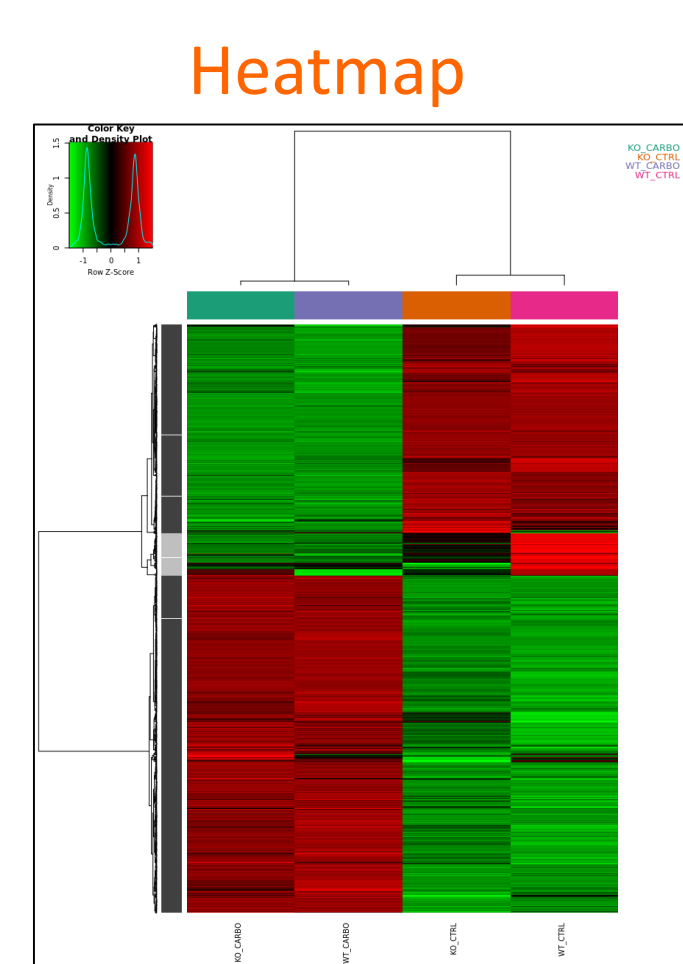
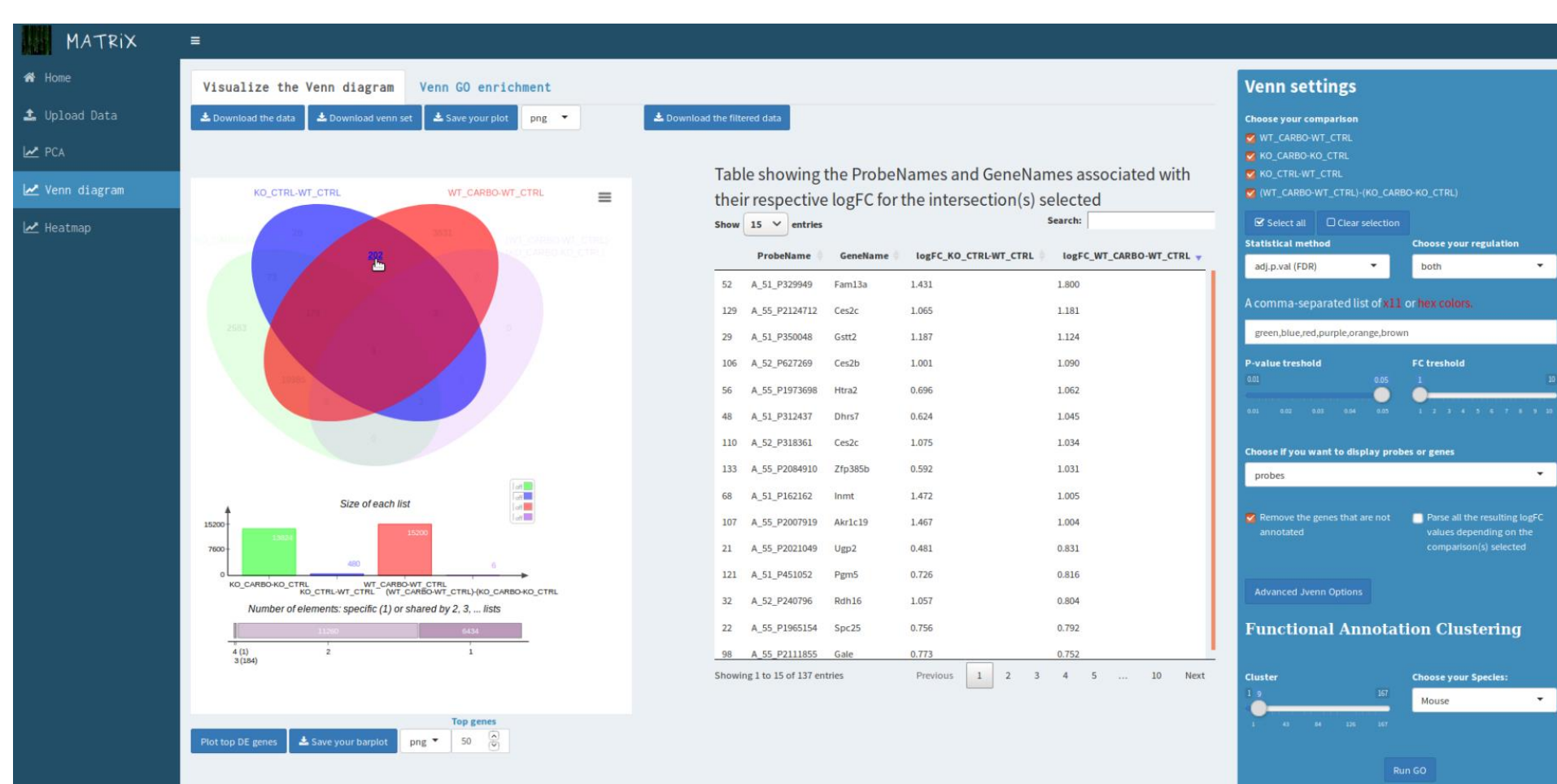
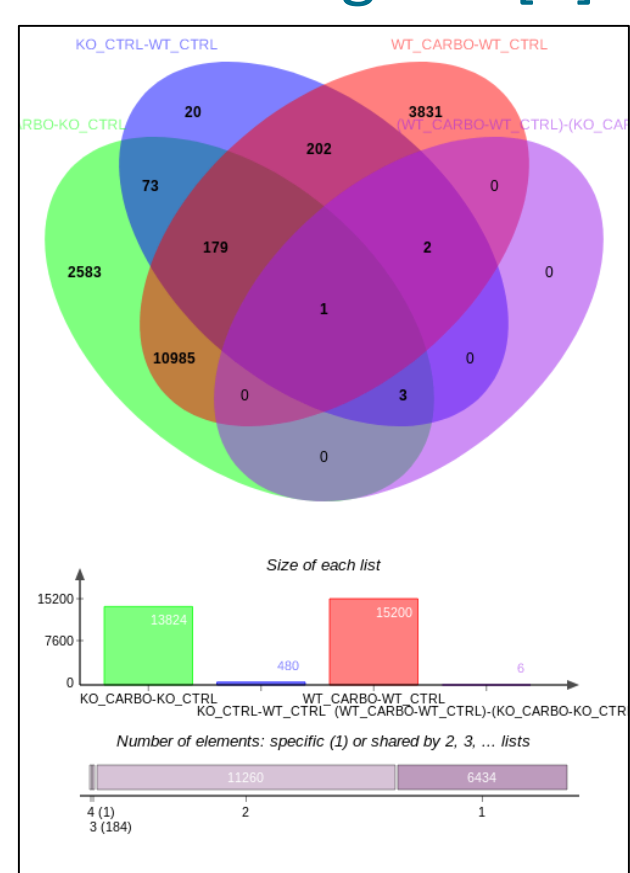


Deployed with Shiny Open Source.

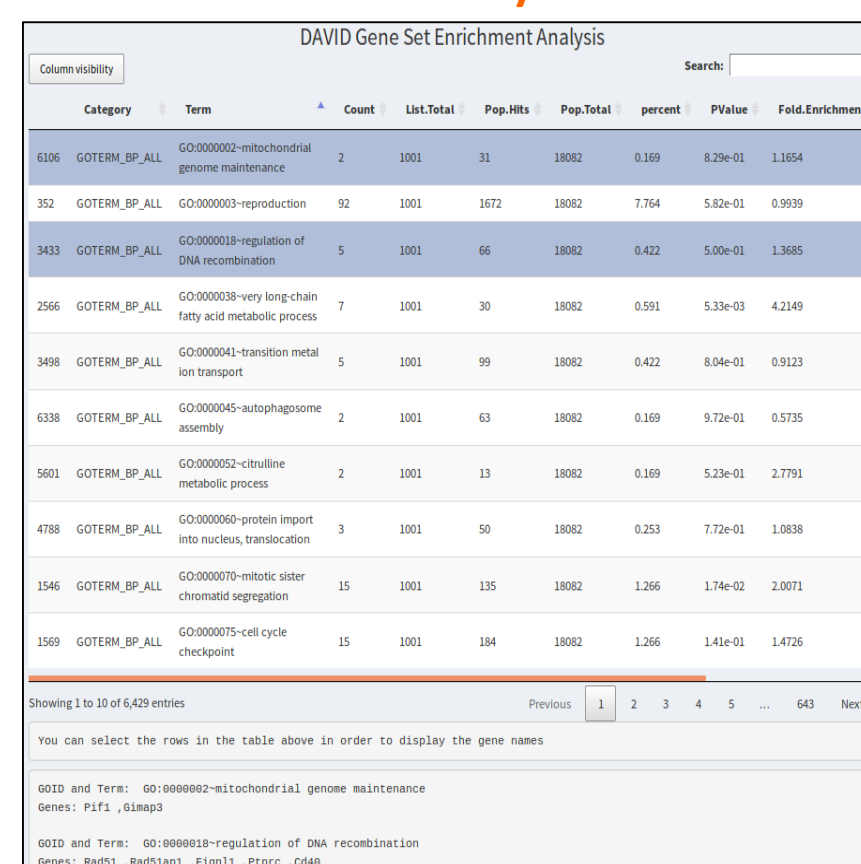
Visualisations



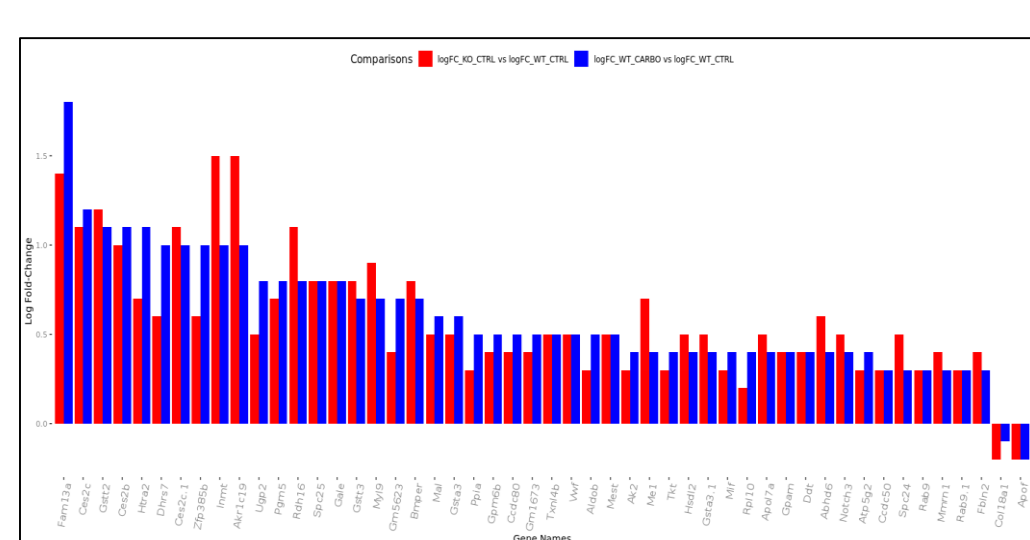
Venn diagram [3]



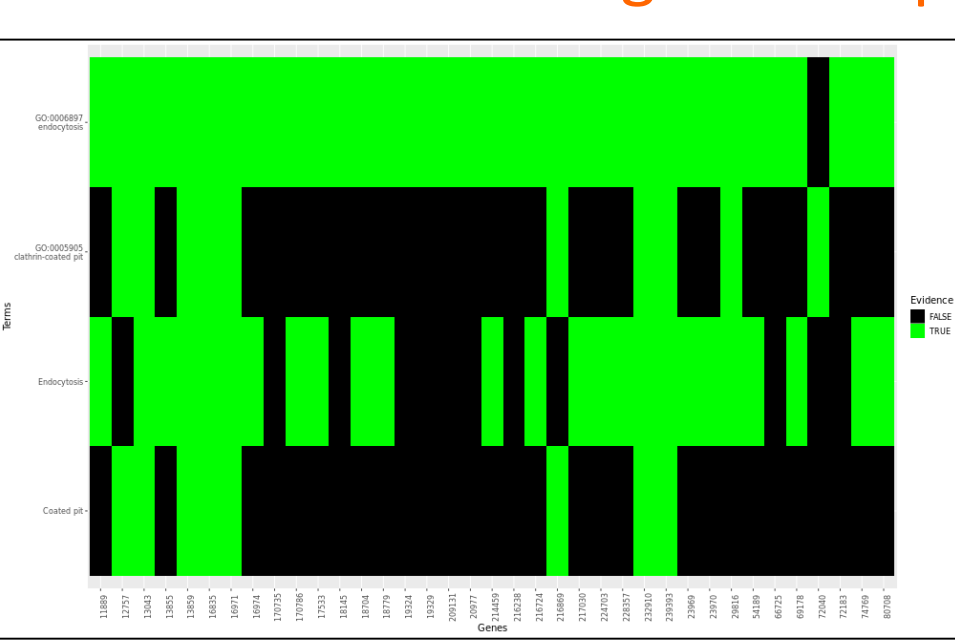
Functional analysis result



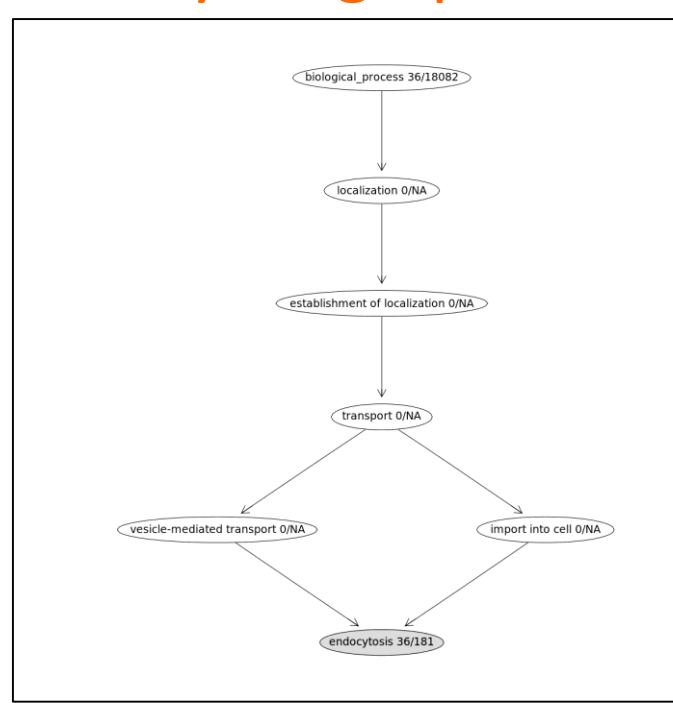
Bar chart



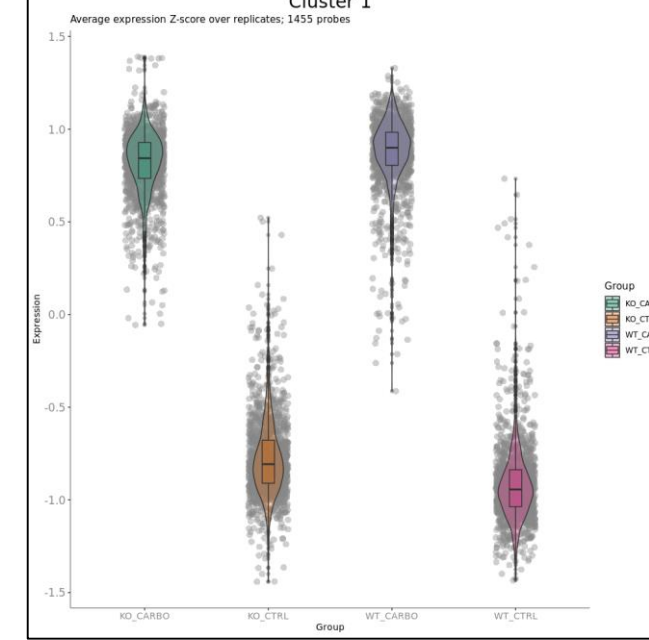
Functional Clustering Heatmap



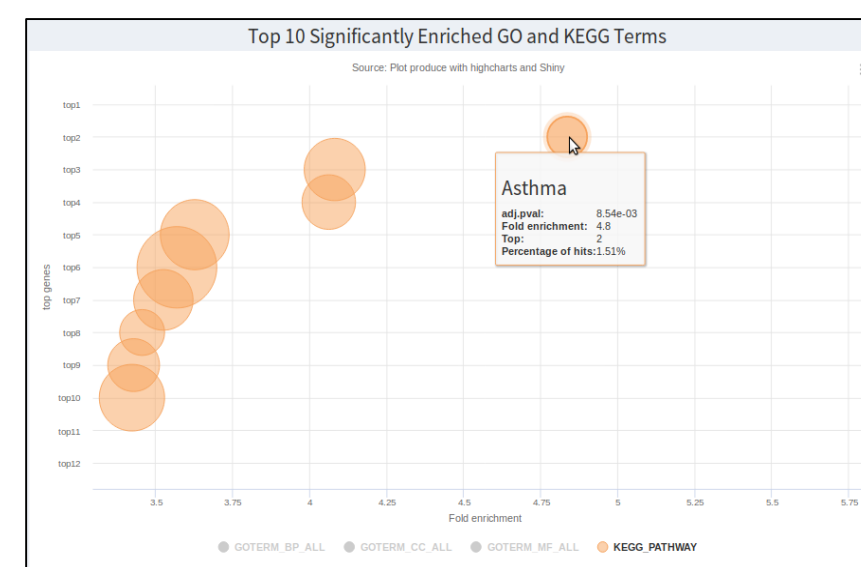
Acyclic graph



Boxplot



Bubble graph



Exploratory analysis
Functional analysis

Conclusion & Perspective

Main features:

- ✓ A turnkey functional application
- ✓ MATrIX leverages the cumulative functionalities of 38 R and 3 js libraries
- ✓ Integrates most of the outputs (graphics, analyses) required for publication
- ✓ Solution for the lack of scalability

Possible improvements:

- Dockerize MATrIX 
- Generating downloadable reports 
- Directly load the data from the server

References

- [1] Gentleman, R.C., Carey, V.J. et al. 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 5, R80. doi:10.1186/gb-2004-5-10-r80
- [2] RStudio, Inc. shiny : Easy web applications in R, 2014.
- [3] Philippe Bardou, Jérôme Mariette, Frédéric Escudié, Christophe Djemiel and Christophe Klopp. jvenn: an interactive Venn diagram viewer. BMC Bioinformatics 2014, 15:293 doi:10.1186/1471-2105-15-293