



G.L.A.S.S

Gene List Analysis using Shiny Server

LONG PROJECT

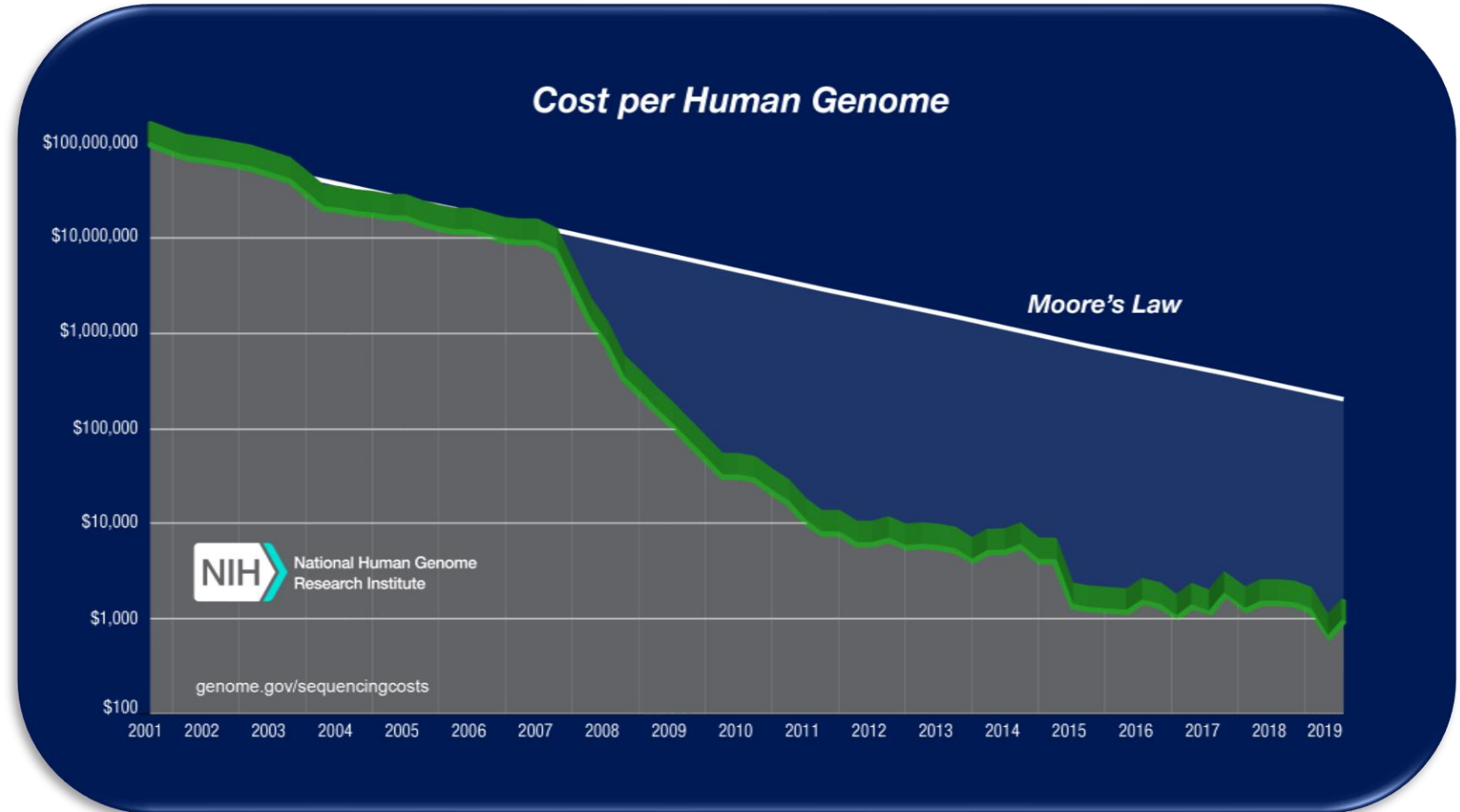
Coralie CAPRON

2019/2020

Tutor : Costas BOUYIOUKOS

Introduction

- Next Generation Sequencing (NGS)
- Result :
Gene List of interest



Introduction

- How to analyze a gene list to get functional information?

- First option : Analyze each gene manually ..



- Second option : Use a tool designed to automate this task

2003



Overlaps between user-supplied
gene lists and the curated databases

2005



Uses every datapoint in
its statistical algorithm

2012

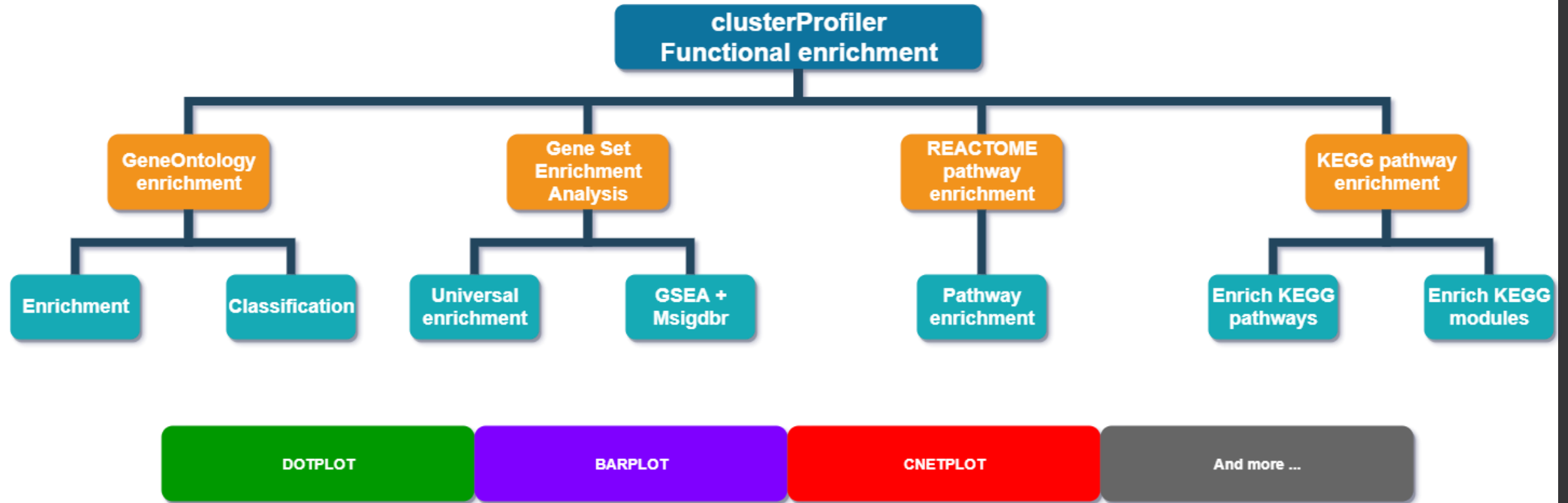


Analysis module and visualization
module were combined

Problematic

Develop an R-Shiny application for the downstream treatment and analysis of a list of “genes of interest”

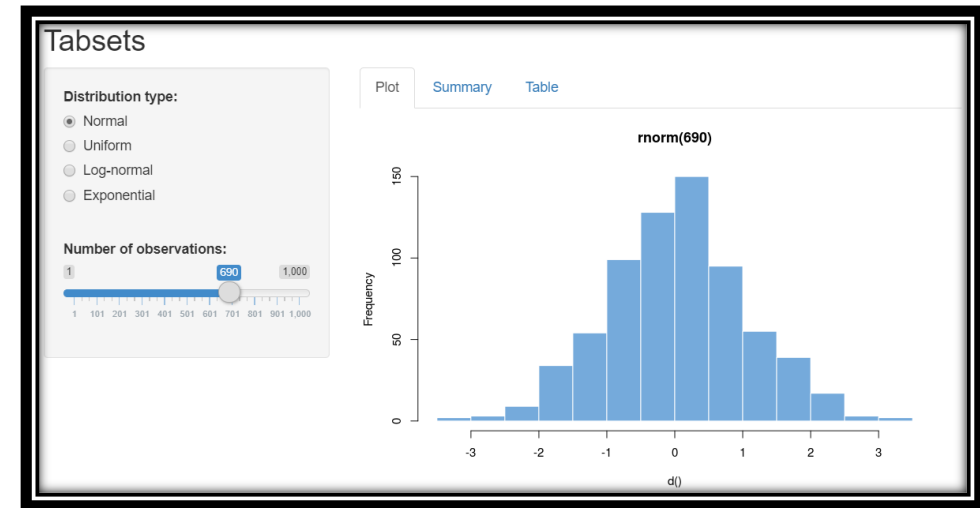
Materials and methods



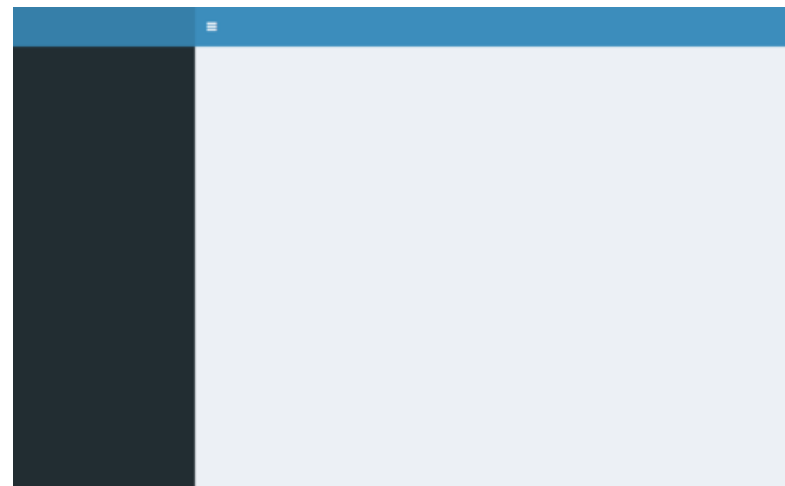
Msigdb : annotated gene sets for use with GSEA

Materials and methods

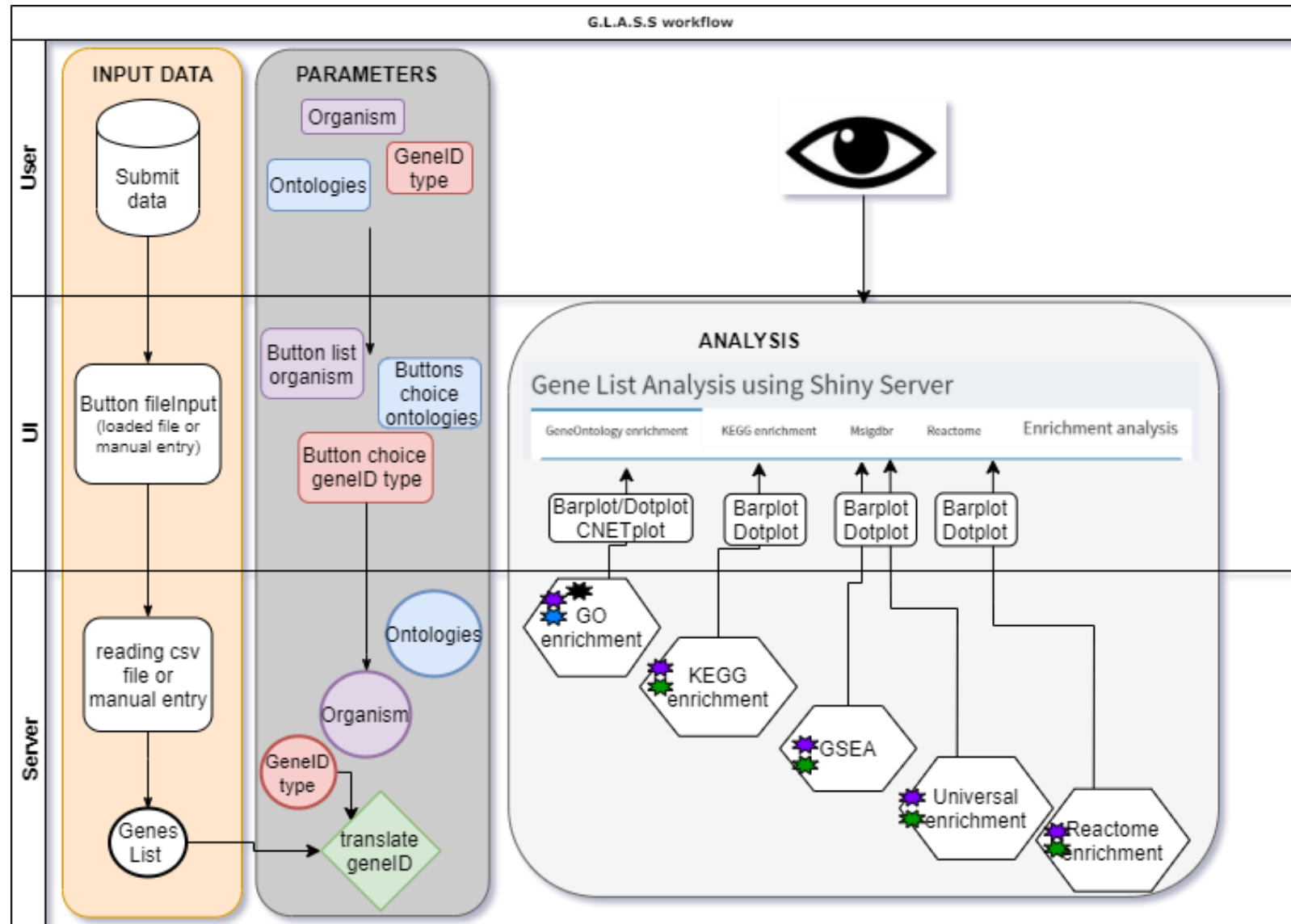
- Shiny



- Shiny Dashboard



Materials and methods



Results

- Example file : Genes impacted by the chemical compound Perfluorooctane sulfonate (PFOS)
 - PFOS is persistent, bioaccumulative and toxic to mammalian species and was used to waterproof materials
- First step : Import your data



G.L.A.S.S

Import data

Enrichments parameters

Analysis

About

Gene List Analysis using Shiny Server

Choose a file

Browse...

PFOS_interactome.csv

Upload complete

...Or write your list of genes

geneID1

geneID2

Write an ordered or unordered list of genes lines by lines with the same separator

Results

- Second step : Choose your enrichments parameters and global parameters

G.L.A.S.S

Import data

Enrichments parameters

Analysis

About

Gene List Analysis using Shiny Server

Select organism

Human (Homo sapiens)

Gene identifier type

☐ ENTREZ

☐ ENSEMBL

☒ SYMBOL

Ontologies (Select at least one of them)

☒ Molecular Function

☐ Cellular Component

☒ Biological Process

Do you want to upload your own Universe ? (GSEA analysis)

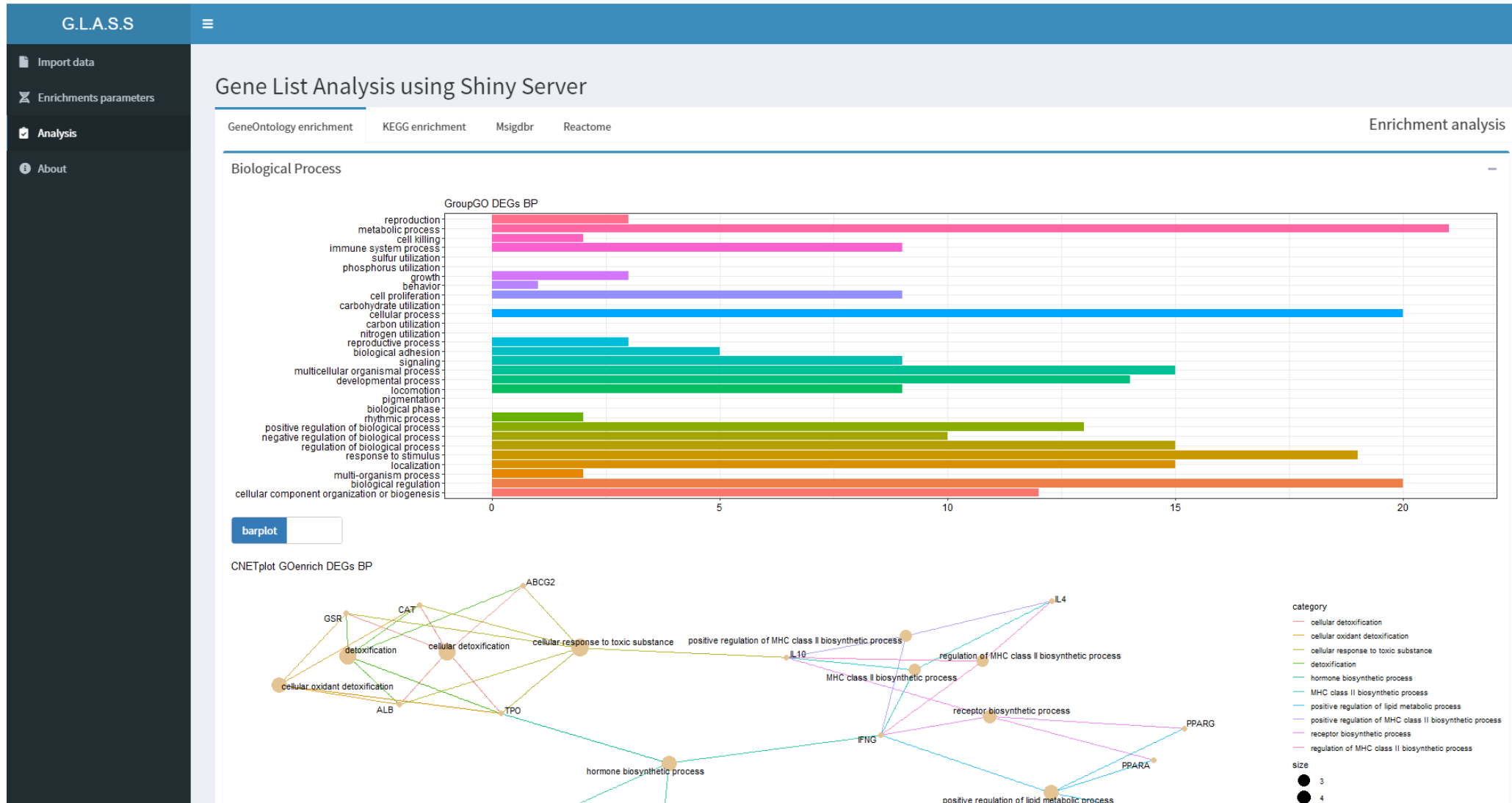
Yes

No, use default (all human genes)

Submit

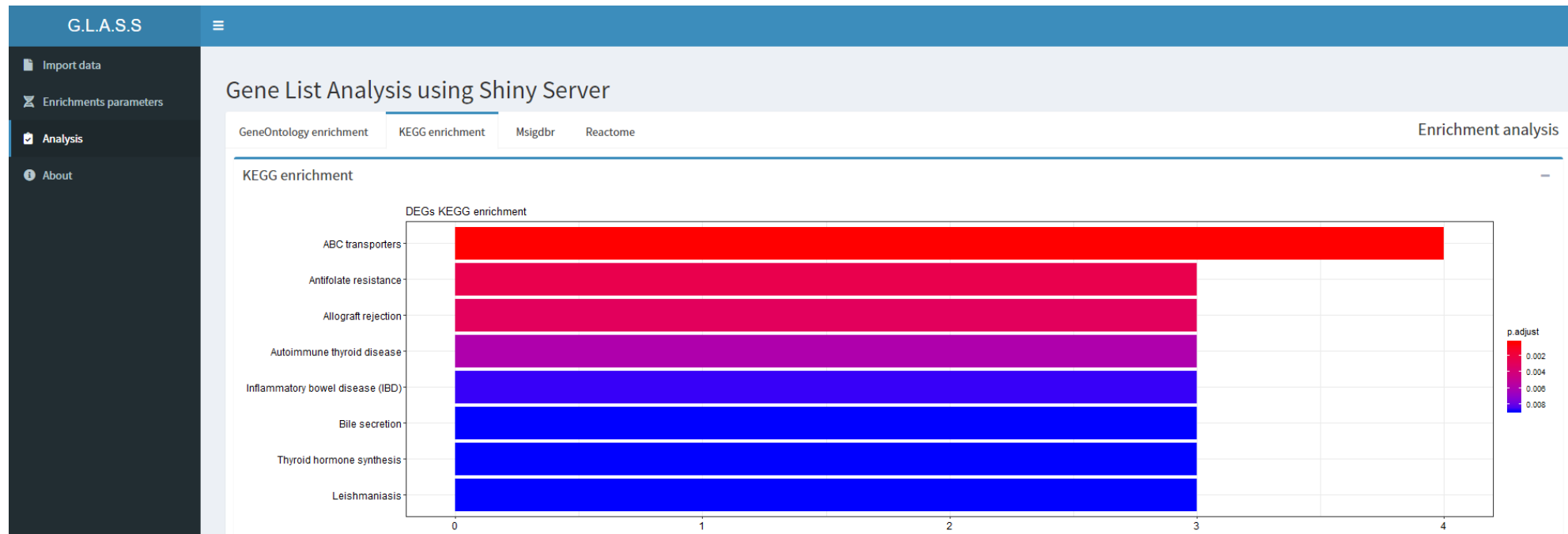
Results

- Third step : Time to analyze !
 - GeneOntology enrichment : terms enriched related to metabolic process, cellular process



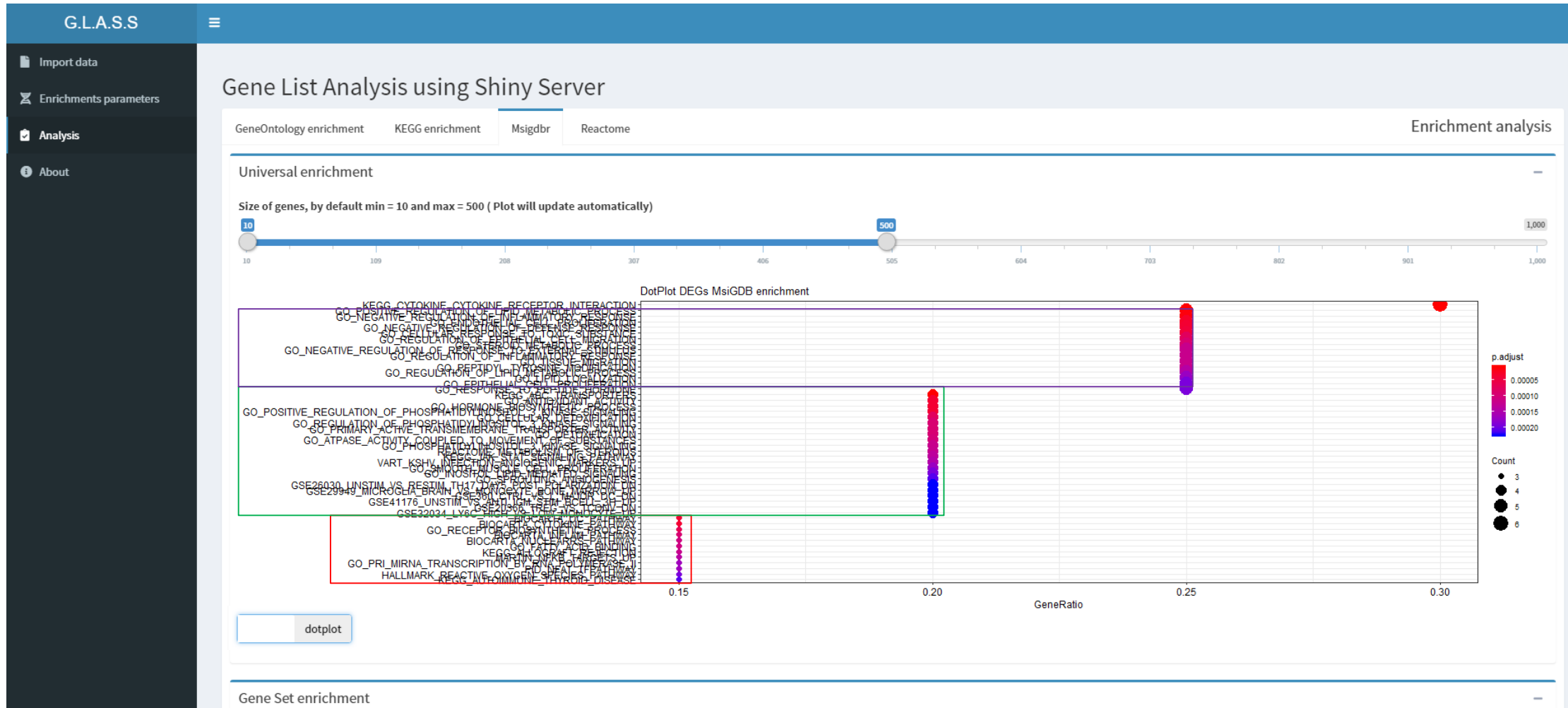
Results

- KEGG enrichment
 - Biological pathways enriched : transporters, diseases, etc.



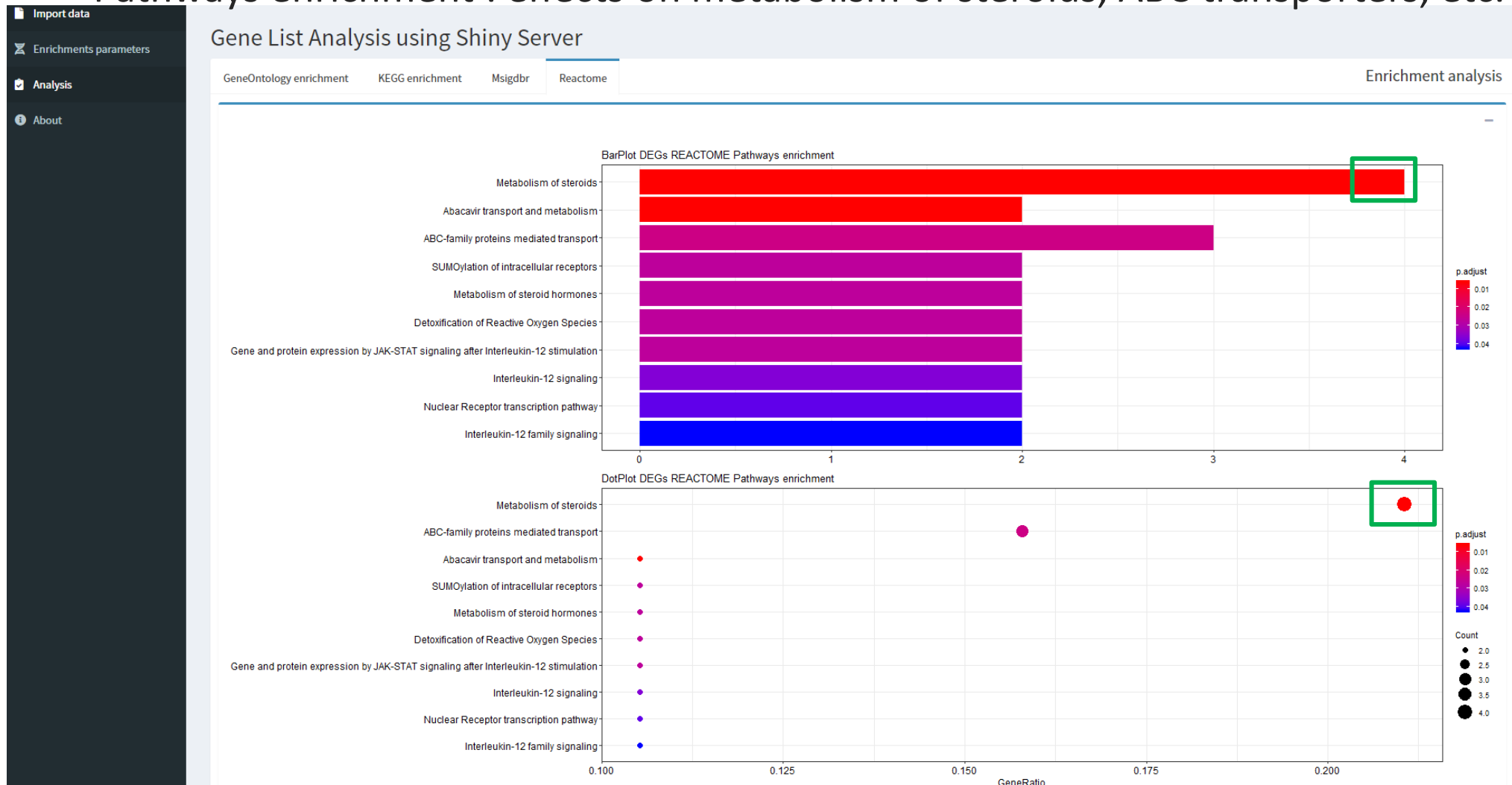
Results

- Universal enrichment and GSEA :
- Observe 4 different clusters



Results

- Reactome enrichment:
 - Pathways enrichment : effects on metabolism of steroids, ABC transporters, etc.



Conclusion and prospects

G.L.A.S.S application :

- Local Shiny server used to analyze a gene list of interest :
 - Gene Ontologies enrichment
 - Gene Set Enrichments Analysis (in case an ordered list is provided)
 - REACTOME pathway enrichment
 - KEGG pathway enrichment
- Some improvements can be done :
 - Decrease the processing time by modifying the library clusterProfiler itself
 - Add some features like interactive view in dotplots, add fold change values, etc.
 - Put the server online

References

- Ashburner, Michael, Catherine A. Ball, Judith A. Blake, David Botstein, Heather Butler, J. Michael Cherry, Allan P. Davis, et al. 2000. « Gene Ontology: tool for the unification of biology ». *Nature genetics* 25 (1): 25-29. <https://doi.org/10.1038/75556>.
- Huang, Da Wei, Brad T. Sherman, et Richard A. Lempicki. 2009. « Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources ». *Nature Protocols* 4 (1): 44-57. <https://doi.org/10.1038/nprot.2008.211>.
- Huber, W., V.J. Carey, R. Gentleman. «Orchestrating high-throughput genomic analysis with Bioconductor». *Morgan Nature Methods*, 2015:12, 115.
- Kanehisa, Minoru, et Susumu Goto. 2000. « KEGG: Kyoto Encyclopedia of Genes and Genomes ». *Nucleic Acids Research* 28 (1): 27-30.
- Mardis, Elaine R. 2011. « A Decade's Perspective on DNA Sequencing Technology ». *Nature* 470 (7333): 198-203. <https://doi.org/10.1038/nature09796>.
- Metzker, Michael L. 2010. « Sequencing Technologies - the next Generation ». *Nature Reviews. Genetics* 11 (1): 31-46. <https://doi.org/10.1038/nrg2626>.
- Subramanian, Aravind, Pablo Tamayo, Vamsi K. Mootha, Sayan Mukherjee, Benjamin L. Ebert, Michael A. Gillette, Amanda Paulovich, et al. 2005. « Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles ». *Proceedings of the National Academy of Sciences of the United States of America* 102 (43): 15545-50. <https://doi.org/10.1073/pnas.0506580102>.
- « Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. - PubMed - NCBI ». s. d. Consulté le 16 janvier 2020. <https://www.ncbi.nlm.nih.gov/pubmed/19131956?dopt=Abstract>.
- Wetterstrand, KA. s. d. « DNA Sequencing Costs: Data ». Genome.Gov. Consulté le 11 janvier 2020. <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>.
- Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson (2019). shiny: Web Application Framework for R
- Yu, Guangchuang, Li-Gen Wang, Yanyan Han, et Qing-Yu He. 2012. « clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters ». *OMICS: a Journal of Integrative Biology* 16 (5): 284-87. <https://doi.org/10.1089/omi.2011.0118>.

Thank you for your attention

Annex

- GO enrichment window, « bloc organization »

The screenshot displays the G.L.A.S.S. (Gene List Analysis using Shiny Server) interface. On the left is a dark sidebar with navigation links: 'Import data', 'Enrichments parameters', 'Analysis' (highlighted), and 'About'. The main content area has a blue header with the title 'Gene List Analysis using Shiny Server'. Below the header, there are four tabs: 'GeneOntology enrichment' (selected), 'KEGG enrichment', 'Msigdb', and 'Reactome'. The 'GeneOntology enrichment' tab shows a table with three rows, each representing a GO term category: 'Biological Process', 'Molecular Function', and 'Cellular Component'. Each row has a '+' icon on the right side, indicating expandable content. The title 'Enrichment analysis' is positioned at the top right of the table area.

Enrichment analysis	
Biological Process	+
Molecular Function	+
Cellular Component	+

Annex

- GO enrichment Biological Process dotplot



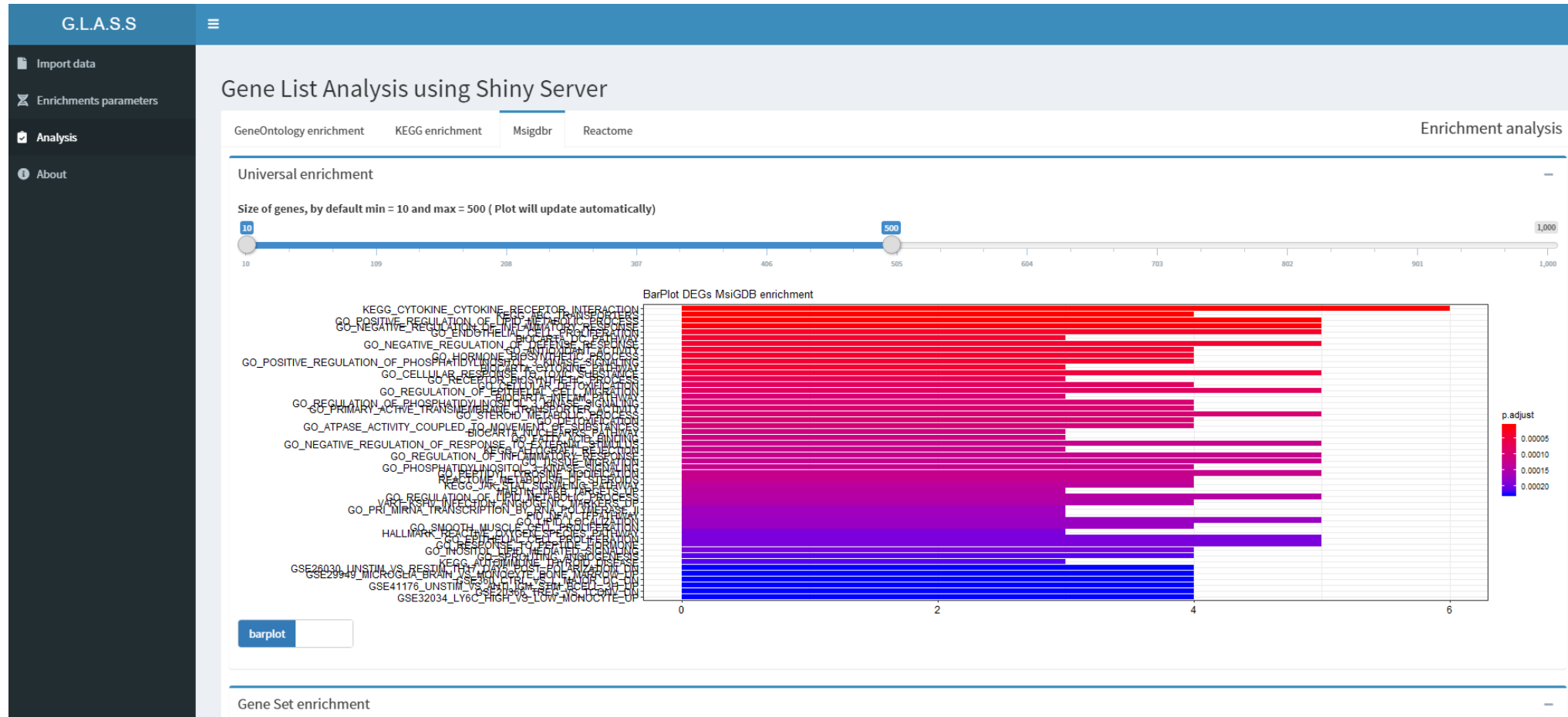
Annex

- GSEA message error

The screenshot displays the G.L.A.S.S. (Gene List Analysis using Shiny) web application interface. A dark blue header bar at the top contains the text "G.L.A.S.S." and a hamburger menu icon. A left sidebar lists navigation options: "Import data", "Enrichments parameters", "Analysis", and "About". The main content area is titled "Gene List Analysis using Shiny S" and features three tabs: "GeneOntology enrichment", "KEGG enrichment", and "Msigdb". The "KEGG enrichment" tab is currently selected. A modal dialog box is overlaid on the interface, titled "About Gene Set Enrichment". It contains the text "GSEA cannot work with an unordered list, please change the input genes list" and a "Dismiss" button. The background of the application is dimmed, and a loading spinner is visible in the center of the main content area. The text "Enrichment analysis" is visible in the top right corner of the main content area.

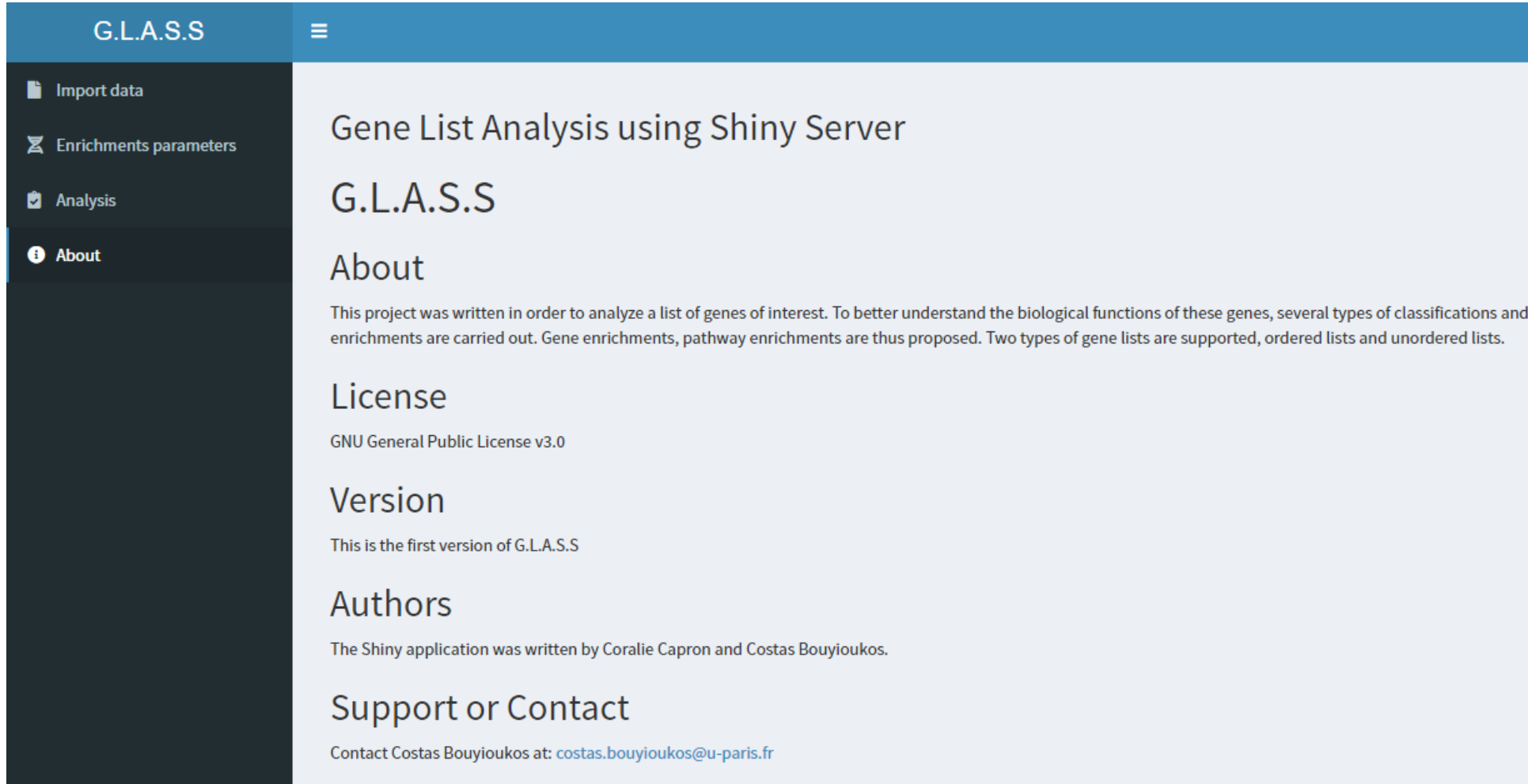
Annex

- Universal enrichment (barplot)



Annex

- About section



The screenshot displays the G.L.A.S.S. web application interface. On the left is a dark sidebar with a menu containing four items: 'Import data' (with a document icon), 'Enrichments parameters' (with a funnel icon), 'Analysis' (with a checkmark icon), and 'About' (with an information icon). The 'About' section is currently selected and highlighted. The main content area on the right has a light blue background and contains the following sections:

- Gene List Analysis using Shiny Server**
- G.L.A.S.S**
- About**

This project was written in order to analyze a list of genes of interest. To better understand the biological functions of these genes, several types of classifications and enrichments are carried out. Gene enrichments, pathway enrichments are thus proposed. Two types of gene lists are supported, ordered lists and unordered lists.
- License**

GNU General Public License v3.0
- Version**

This is the first version of G.L.A.S.S
- Authors**

The Shiny application was written by Coralie Capron and Costas Bouyioukos.
- Support or Contact**

Contact Costas Bouyioukos at: costas.bouyioukos@u-paris.fr

Annex

- Hidden sidebar

G.L.A.S.S.

Gene List Analysis using Shiny Server

Choose a file

Browse...

No file selected

...Or write your list of genes

geneID1
geneID2

Write an ordered or unordered list of genes lines by lines with the same separator

Annex

- First implementation of G.L.A.S.S without using shiny dashboard

