

## Pipeline – Phase 3: Functional Analysis

### Preparation of files

1. Organisms – using the results obtained from the last script, create a .csv file called **organisms.csv** in which it contains the name of the microorganisms of interest separated by Genus and Species. Use the following template and fill it in:  
<https://github.com/VaneBR/Project/blob/main/Files/organisms.csv> or  
[https://drive.google.com/drive/u/0/folders/18yojrGbjKc\\_-GAFdOLppHP2UWXFu3f1](https://drive.google.com/drive/u/0/folders/18yojrGbjKc_-GAFdOLppHP2UWXFu3f1).
2. Pathways – in a .txt file write the ID number of the pathways of interest and name it **pathway\_ids.txt**. To look for the ID of the pathway go to:  
<https://www.genome.jp/kegg/pathway.html>

### Steps to use the script:

The script described below is available in the following link:

<https://colab.research.google.com/drive/1Gk60AGH4BrO4-gBCOi3PDpjUREqHk8QQ> or in github:  
[https://github.com/VaneBR/Project/blob/main/Scripts/functional\\_information.py](https://github.com/VaneBR/Project/blob/main/Scripts/functional_information.py).

1. Upload both organisms.csv and pathway\_ids.txt in GoogleColab.
2. Run the script.
3. The output **compatibility\_table\_with\_names.csv** has the enzymes that are present in both the bacteria and the pathways.
4. To help analyze your results use the script **interactive\_table.py** available in  
[https://colab.research.google.com/drive/1MsSEwkGDBfJdlfEcs\\_i8jHDhh1djcWIW?usp=drive\\_l](https://colab.research.google.com/drive/1MsSEwkGDBfJdlfEcs_i8jHDhh1djcWIW?usp=drive_l)  
ink or in github: [https://github.com/VaneBR/Project/blob/main/Scripts/interactive\\_table.py](https://github.com/VaneBR/Project/blob/main/Scripts/interactive_table.py)
5. Upload the **compatibility\_table\_with\_names.csv** file and run the script
6. Analyze your results.