Gene Set Analysis –Methods and Tools
Exercise 1

Antonio Mora, Ph.D.
Lv Xuanyi
(Xie Chengshu)

21.09.2020

Mora Lab
FUNCTIONAL BIOINFORMATICS

creative commons

Attribution-Share Alike 2.5 Canada

You are free:

to Share — to copy, distribute and transmit the work

to Remix — to adapt the work

APPROVED FOR Free Cultural Works

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar licence to this one.

- For any reuse or distribution, you must make clear to others the licence terms of this work.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- The author's moral rights are retained in this licence.

Disclaimer

Learn how to distribute your work using this licence

Exercise. Pathway databases.

1. Go to REACTOME: reactome.org
2. Find the search box and type: "VEGFB" (we will look for pathways related to this protein).
3. You will get options for a single protein and for a complex (because VEGFB forms a dimer). Choose "VEGFB dimer". What is the location of this dimer? What is its Reactome ID?
4. Go to 'Signal transduction (Homo sapiens)' (click on the "plus" icon to the left of the names and observe all the structure of pathways).
5. Now let's choose one specific pathway. Click on "Homodimerization of VEGF proteins (Homo sapiens)". The 'pathway browser' should appear and show you the pathway. What is the Reactome ID of this pathway? Take a look to the entire pathway (zoom in and out using your mouse).
6. Check the "Description" and "Molecules" tabs: What are the UniProt IDs of the proteins involved in this pathway? Click on each protein ID and open their UniProt pages. What are their names and functions? In which tissues are they expressed?
7. Find the "Downloads" tab and download the pathway in pdf format (how many pages long is the pdf report?)

Exercise. Pathway databases.

8. Go to KEGG: www.kegg.jp
9. Find the search box and, again, type: "VEGFB"
10. Under "KEGG GENES", you will find that the gene is present in several species. How many?
11. Choose the one for human ("hsa:7423"). How many KEGG pathways are related to this protein? How related are they to the ones in Reactome?
12. Choose "Pathways in cancer" ("hsa05200") and see the typical KEGG pathway display. In the search box, type "VEGFB": what happened?

13. Go to PATHWAY COMMONS: www.pathwaycommons.org. Again, search for "VEGFB". How many pathways do you get? From how many different databases?
14. Finally, compare a pathway in two different databases. Go to both REACTOME and KEGG, and type "Glycolysis". What are the KEGG and REACTOME IDs for Glycolysis? How different are their pathway diagrams? Use these databases to find out what is glycolysis, which proteins are involved, and which diseases are associated. Which database is your favorite?

Mora Lab
FUNCTIONAL BIOINFORMATICS