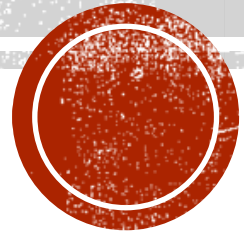
DSEA

http://dsea.tigem.it/



Diego Carrella

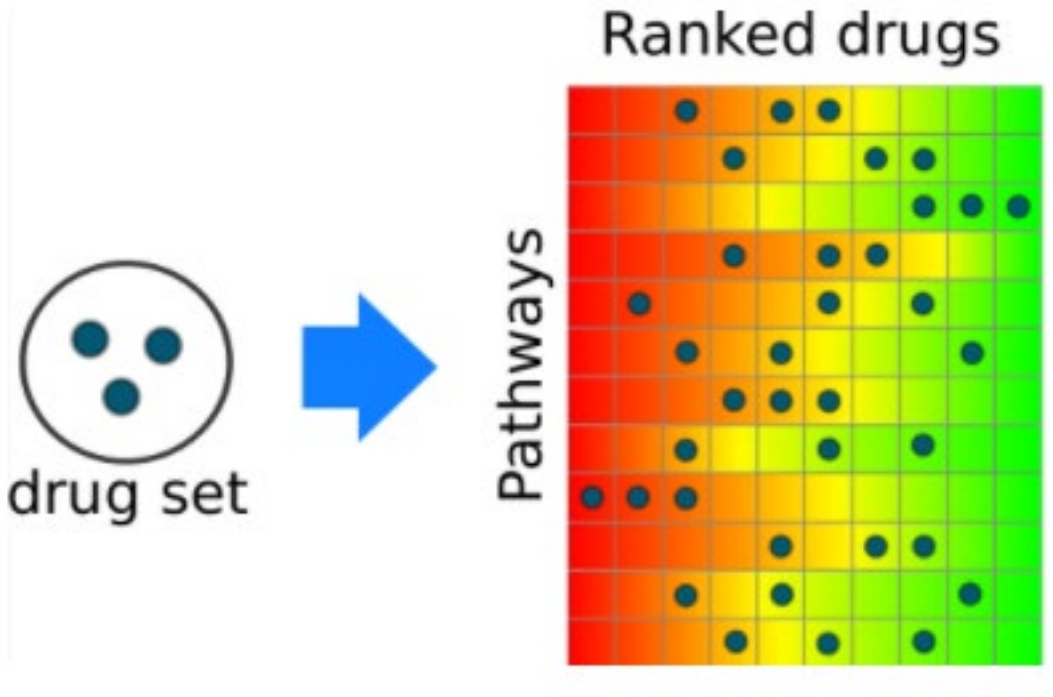
DSEA

- **Drug Set Enrichment Analysis (DSEA)** detects molecular pathways that are consistently up- or down- regulated by a set of drugs.

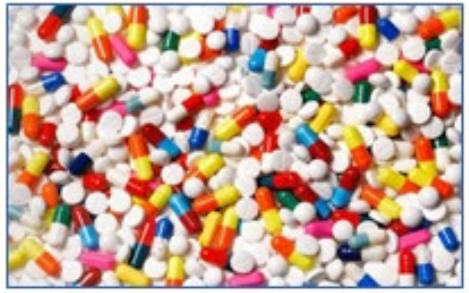
- When you do automatic drug screening you get a set of molecules that have nothing to do with each other.

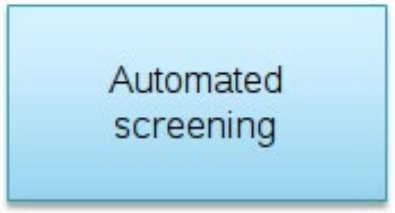
- With the DSEA we try to understand what they have in common by looking at the regulated paths by all of them (or most of them).

METHOD



- The Drug Set Enrichment Analysis (DSEA) works on the same principles as GSEA.
- A set of drugs of interest is tested against a database of pathways.
- Each pathway in the database is stored as a ranked list of drugs, sorted from the one most up-regulating the pathway to the one most down-regulating it.





METHOD

- DSEA tend to highlight pathways that are most regulated by a set of drugs of interest compared with the other drugs in the database.
- This could be the key behind the efficacy of the screening hits and the inefficacy of the other drugs.

CURRENTLY SUPPORTED GENE SET DATABASES

- Pathways analyzed by the DSEA are defined as gene sets collected by various sources, as summarized in the following table:

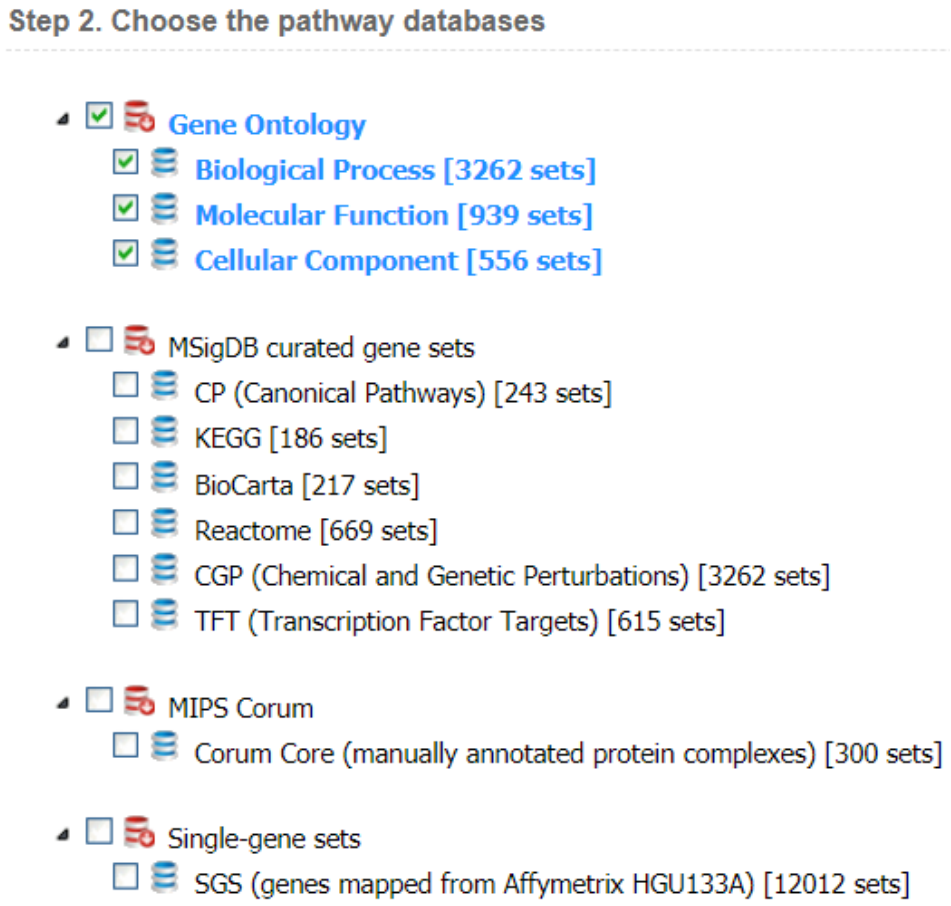
Source	Name	Description	#
BioMart	GO BP	Gene Ontology - Biological Processes	3262
BioMart	GO MF	Gene Ontology - Molecular Function	939
BioMart	GO CC	Gene Ontology - Cellular Component	556
MSigDB	CP	Expert-defined Canonical Pathways	243
MSigDB	KEGG	Kyoto Encyclopedia of Genes and Genomes	186
MSigDB	Biocarta	Community-fed molecular relationships	217
MSigDB	Reactome	Open-source, open access, manually curated and peer-reviewed pathway database	674
MSigDB	CGP	Genetic and Chemical Perturbations	2427
Mips	CORUM	Comprehensive Resource of Mammalian protein complexes	1343
-	SGS	Sets containing single genes mapped from Affymetrix chip U133A	12012

RUNNING A DSEA ANALYSIS

- **Step 1. Define the drug set**
  Create a drug set by selecting drugs from the *Find drug* box and/or directly pasting a list into the *Drug set* box (one per line).

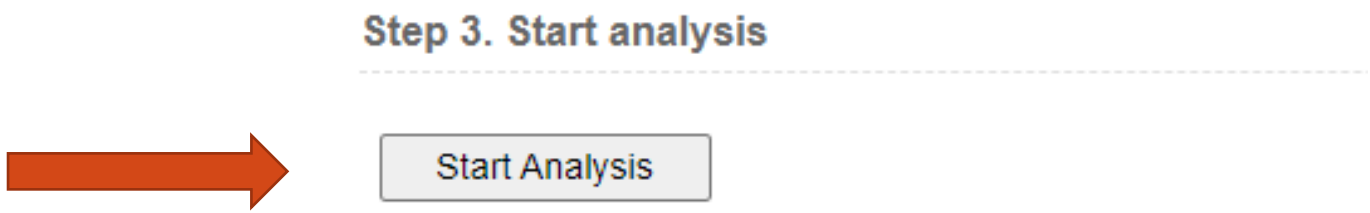
RUNNING A DSEA ANALYSIS

- **Step 2. Choose the pathway databases**
  - DSEA will be performed for all the pathways included in the chosen databases.
  - Analyses on the different databases are independent.



RUNNING A DSEA ANALYSIS

- **Step 3. Start analysis**
  Click the *Start analysis* button to run the analysis and wait for processing. A results page will be showed when finished.



RESULTS

Gene Ontology

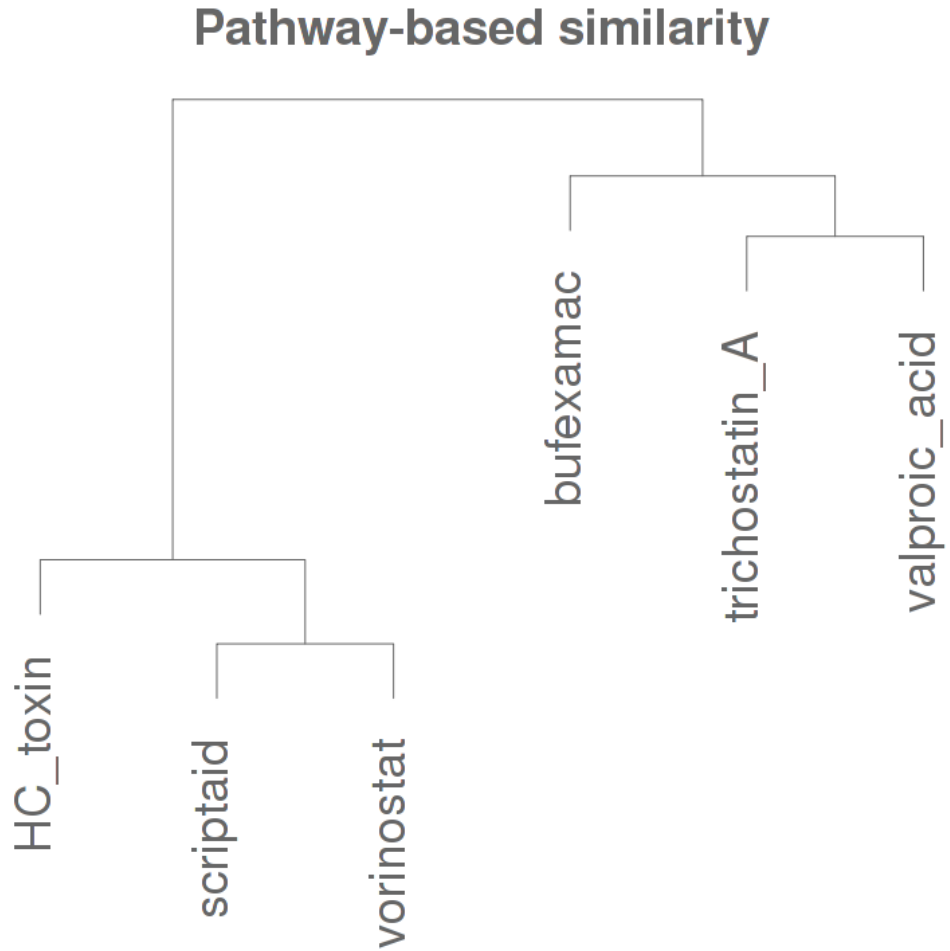
Biological Process

Rank	Pathway Name	EScore	Pvalue
1	cellular protein localization	0.94	1.36e-7
2	nerve growth factor signaling pathway	0.93	2.52e-7
3	mRNA cis splicing, via spliceosome	-0.91	1.06e-6
4	mRNA export from nucleus	-0.91	1.50e-6
5	nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay	-0.89	3.23e-6
6	synaptic vesicle exocytosis	0.89	3.51e-6
7	transcription from RNA polymerase II promoter	-0.88	6.37e-6
8	cell-matrix adhesion	0.88	6.37e-6
9	negative regulation of mRNA splicing, via spliceosome	-0.87	1.06e-5
10	cellular response to insulin stimulus	0.87	1.26e-5
11	tRNA modification	-0.86	1.34e-5

- Results for each database are shown as separate tables, in which pathways are sorted according to relevance, together with the corresponding Enrichment Scores (ESs) and nominal p-values.
- Top 10% pathways are shown for each database.
- Positive ESs correspond to up-regulated pathways, negative ESs correspond to down-regulated pathways.
- P-values indicate how much the ranks of the chosen drugs are consistently up-regulated or down-regulated for each pathway.

RESULTS

- The figure on the right of each table shows a hierarchical clustering of the drugs according to their similarities, based on pathway rankings for each database. Similar drugs are those that up-regulate and down-regulate the same pathways.

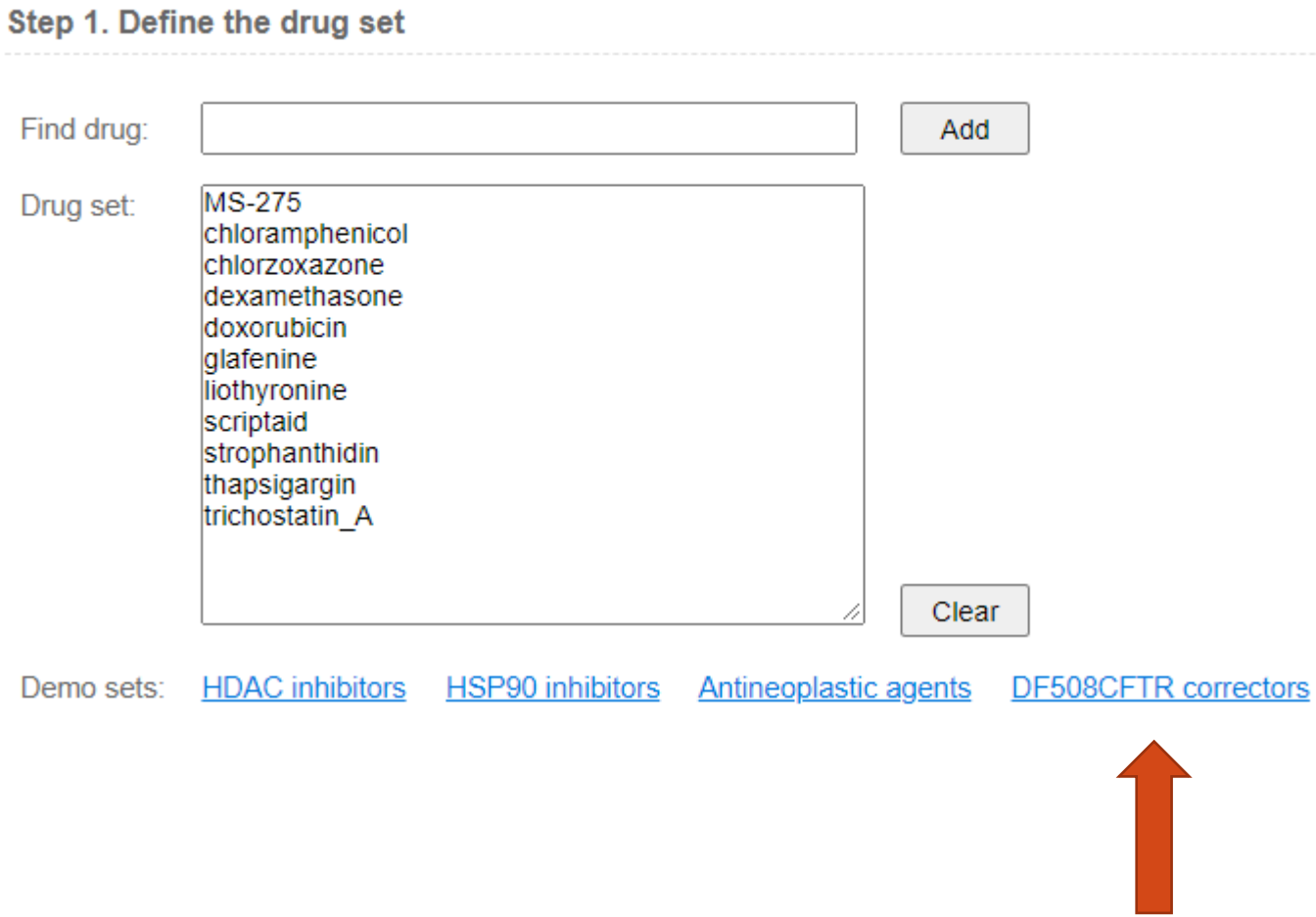


RESULTS

- As soon as the Excel report of the results is ready, the "Export" button on top of the page will become active.

- Excel files contain all the database results as different sheets, together with additional information including individual drug ranks.

EXAMPLE: CYSTIC FIBROSIS



- Cystic fibrosis is caused by mutations in the gene coding for the CFTR (CF transmembrane conductance regulator) protein. The most frequent mutation is the deletion of phenylalanine 508 (DF508).
- We applied DSEA to a drug-set consisting of 11 drugs reported to act as DF508-CFTR correctors in Cystic Fibrosis (CF).

EXAMPLE: CYSTIC FIBROSIS

Cellular Component

Rank	Pathway Name	EScore	Pvalue
1	chloride channel complex	0.62	4.73e-4
2	dendrite membrane	0.6	8.73e-4
3	mRNA cleavage factor complex	-0.58	1.40e-3
4	signal recognition particle	0.57	1.60e-3
5	nuclear membrane	-0.56	2.13e-3
6	axonemal dynein complex	0.56	2.13e-3
7	transcription factor TFIIC complex	0.56	2.21e-3

- DSEA predicts that one mode of action shared in common by the 11 drugs is the upregulation of **chloride channel genes' expression**. ES score associated to chloride channel complex for the drug-set is positive.

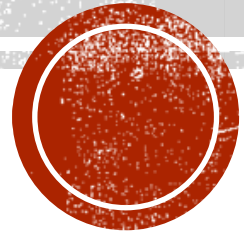
- The 11 drugs belong to very different pharmacological classes, the effect on the **chloride channel** gene expression is detected by DSEA only because it is a common 'side-effect' shared by most of them.

PUBLICATIONS

- Napolitano, F., Carrella, D., Mandriani, B., Pisonero-Vaquero, S., Sirci, F., Medina, D.L., Brunetti-Pierri, N., and di Bernardo, D. *Gene2drug: a computational tool for pathway-based rational drug repositioning.* Bioinformatics (2017).

GENE2DRUG

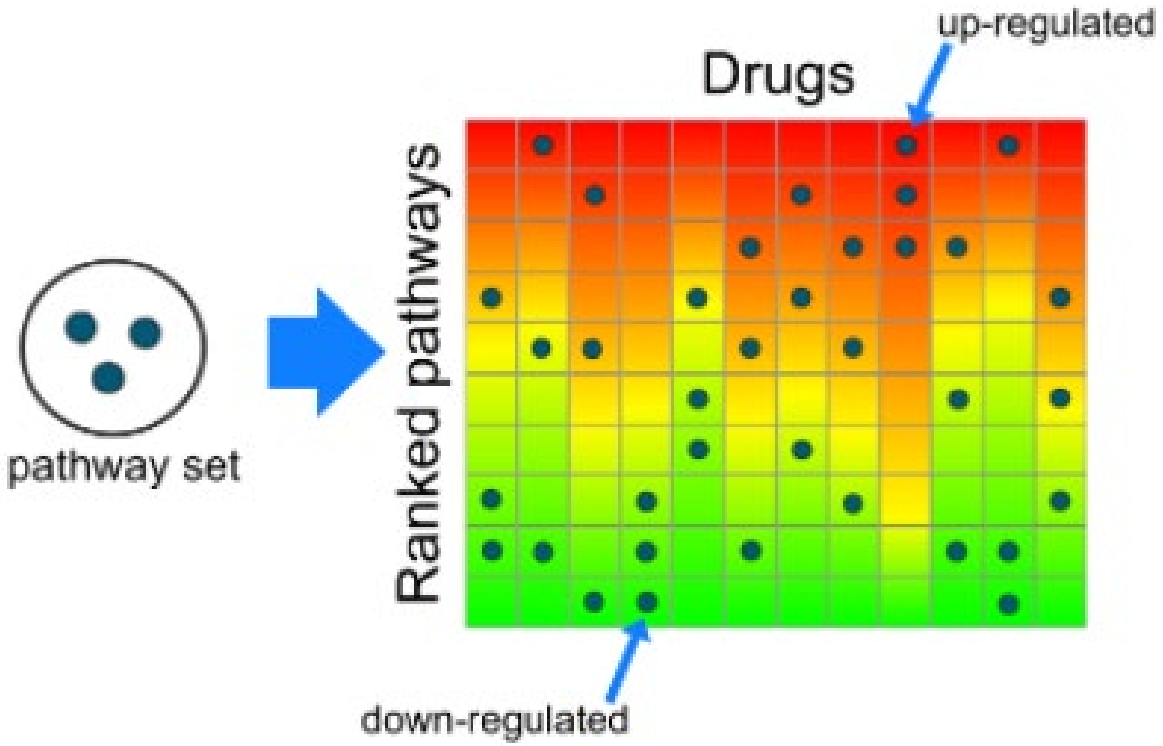
http://gene2drug.tigem.it/



Diego Carrella

- **Gene2drug** ranks small molecules according to their ability to regulate an input set of pathways. Sets of pathways can be defined starting from a gene and exploiting its pathway annotations from a number of publicly available databases.
- The effects of direct drug targets are usually not detectable at the mRNA level and a more systematic approach is required to link transcriptional data to therapeutic effects.
- Therefore, Gene2drug searches for sets of pathways that are regulated by a drug.
- One way to define a set of pathways is to start from a target gene and collect the pathways that it is annotated to, thus assessing the effects of perturbing the target gene indirectly through cellular mechanisms that are expected to be involved.

METHOD



- Gene2drug uses a method we called “Pathway-set Enrichment Analysis” (PSEA), analogous to the Gene-set Enrichment Analysis (GSEA).
- Gene expression profiles from the Connectivity Map are converted to “pathway expression profiles” and ranked according to the p-value (Kolmogorov-Smirnov statistic).
- Given a set of pathways il tool search for drugs that consistently up-regulate or down-regulate most pathways in the set.

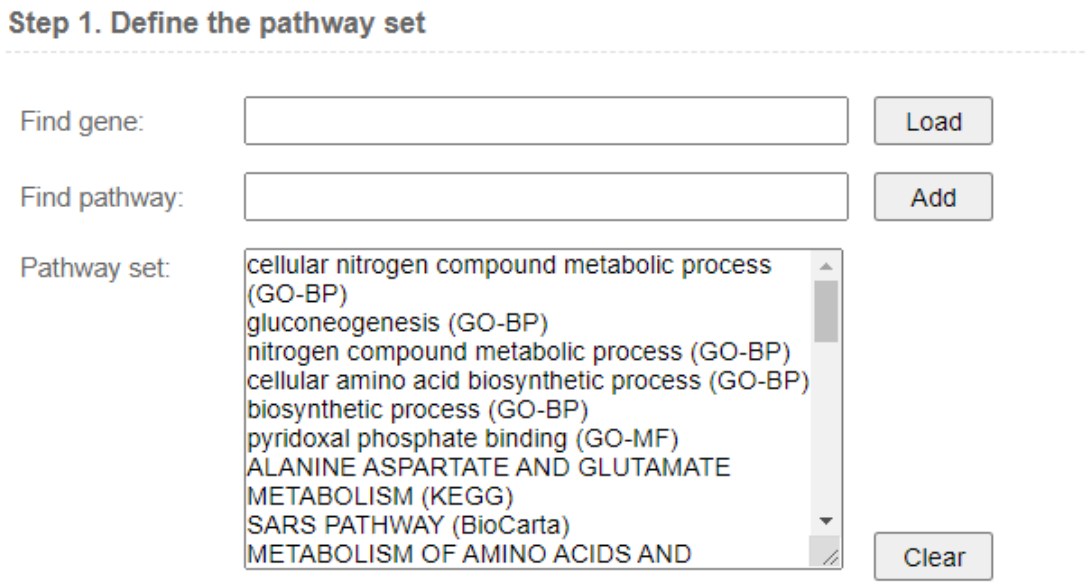
RUNNING A DSEA ANALYSIS

- Gene2drug needs a set of pathways as input to run the PSEA analysis. The set of pathways can be defined in three different ways:

1. By typing the name of a gene in the “Find gene” box and choosing a corresponding result from the autocompletion popup window. The “Load” button will add to the “Pathway set” box all the sets that the chosen gene is annotated to.

2. By typing the name of a pathway in the “Find pathway” box and choosing a corresponding result from the autocompletion popup window. The “Add” button will add the chosen pathway to the “Pathway set” box.

3. By pasting a list of pathways, one per line, directly to the “Pathway set” box. However, pathway names must match our list of pathways, including “([database])”.



RESULTS

Reactome

Rank	Compound Name	EScore	Pvalue
1	fulvestrant	0.98	1.07e-3
2	citalopram	-0.98	1.22e-3
3	tomatidine	0.97	1.88e-3
4	clonidine	-0.97	2.07e-3
5	nifuroxazide	0.97	2.47e-3
6	terguride	0.96	3.14e-3
7	ricinine	-0.96	3.63e-3
8	monastrol	0.96	3.89e-3
9	Prestwick-1080	-0.96	4.16e-3
10	triflusal	0.96	4.16e-3
11	5707885	0.96	4.44e-3

- The results page shows a list of drugs (top 10%).
- Top drugs are those most regulating (up or down) the pathways in the input set.
- Enrichment score sign indicates if the regulation is “up” or “down”.

RESULTS

- The “Export Data” button becomes active as soon as a full report of the analysis is ready.

EXAMPLE: HYPEROXALURIA

- The **GPT** (liver-specific enzyme GPT) plays a key role in the intermediary metabolism of glucose and amino acids.
- GPT overexpression reduces oxalate in mouse models with a rare genetic disorder (**hyperoxaluria**).
- We applied Gene2Drug to find drugs effective at increasing the expression of GPT to reduce the hyperoxaluria.

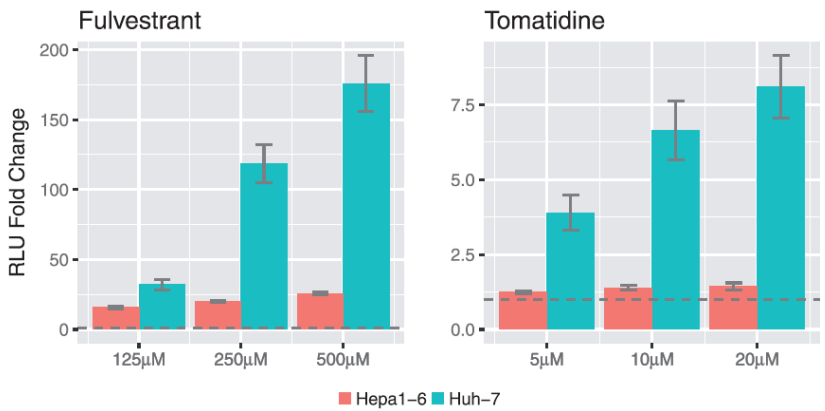
EXAMPLE: HYPEROXALURIA

Reactome

Rank	Compound Name	EScore	Pvalue
1	fulvestrant	0.98	1.07e-3
2	citalopram	-0.98	1.22e-3
3	tomatidine	0.97	1.88e-3
4	clonidine	-0.97	2.07e-3

- In **Reactome** DB the first 2 compounds ranked as those ones most upregulating the pathways involving GPT are: **Fulvestrant** and **Tomatidine**.

- We experimentally tested the efficacy of these drugs to upregulate GPT in two different cell lines:
  - Fulvestrant resulted in significant upregulation of luciferase only at concentrations above 125 µM.
  - Tomatidine resulted in increase of luciferase expression at low concentrations.



PUBLICATIONS

- Napolitano, F., Sirci, F., Carrella, D. & di Bernardo, D. *Drug-set enrichment analysis: a novel tool to investigate drug mode of action.* **Bioinformatics** (2015). doi:10.1093/bioinformatics/btv536.