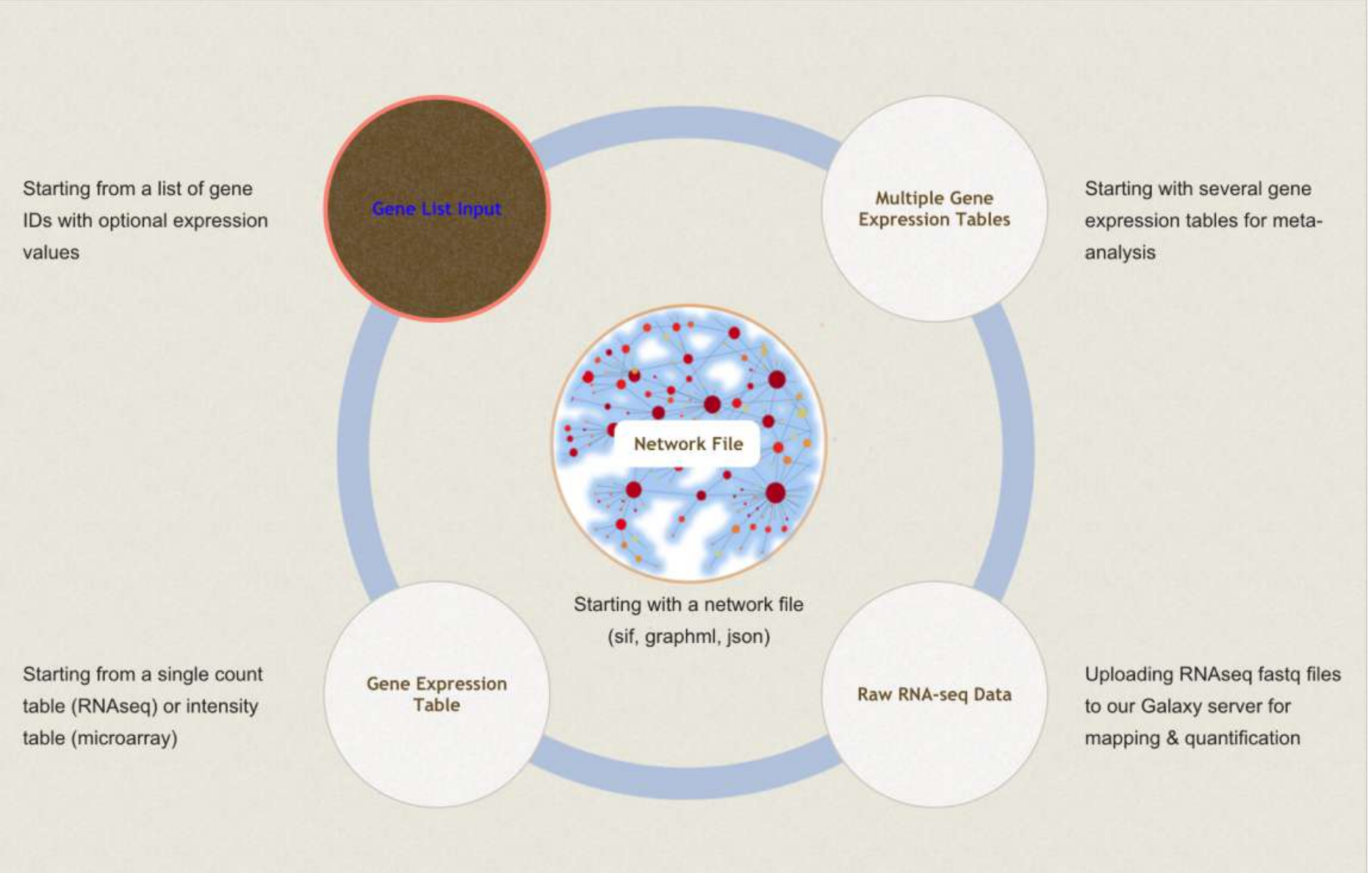
Tutorial 2b: multiple gene list input



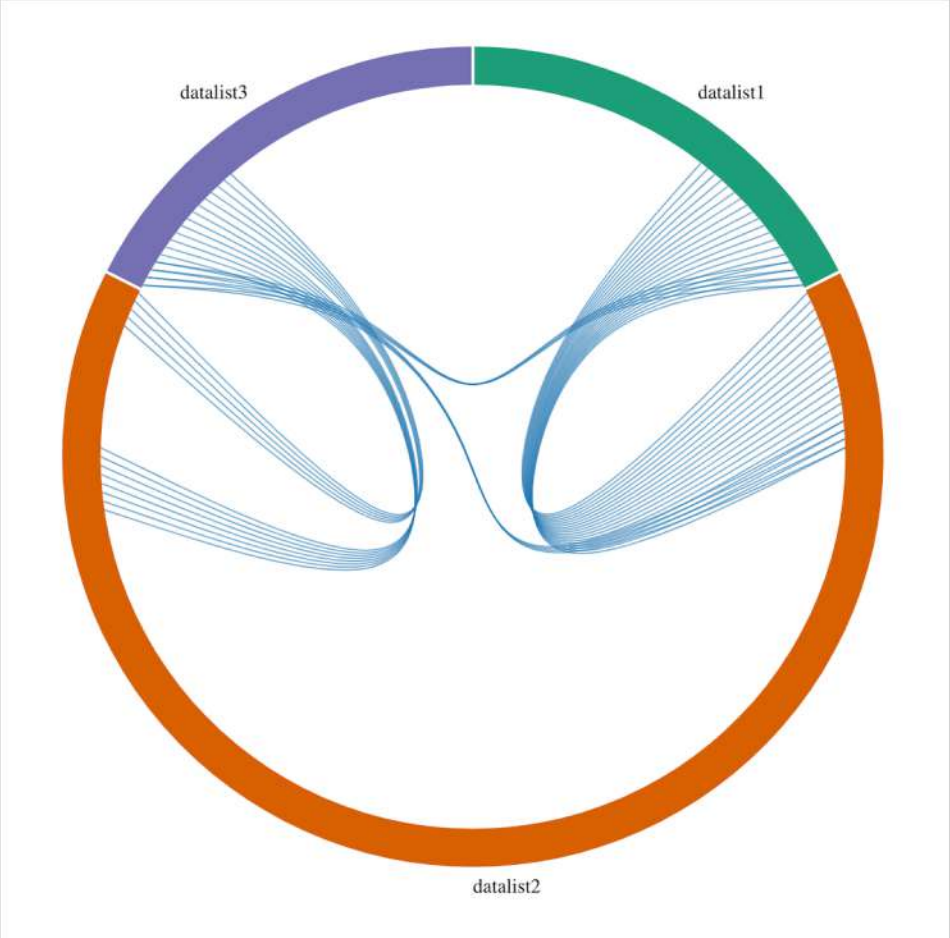
Starting from a list of gene IDs with optional expression values

Starting with several gene expression tables for meta-analysis

Starting with a network file (sif, graphml, json)

Starting from a single count table (RNAseq) or intensity table (microarray)

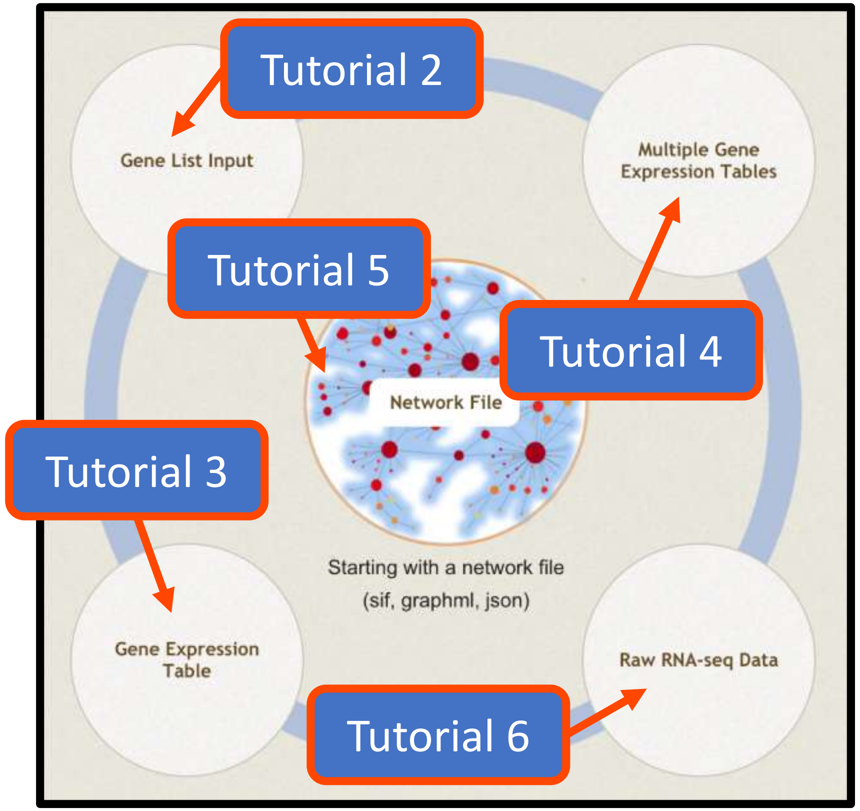
Uploading RNAseq fastq files to our Galaxy server for mapping & quantification



Intro to NetworkAnalyst

- Web application that enables complex meta-analysis and visualization
- Designed to be accessible to biologists rather than specialized bioinformaticians
- Integrates <u>advanced statistical methods</u> and <u>innovative data visualization</u> to support:
  - Efficient data comparisons
  - Biological interpretation
  - Hypothesis generation

Tutorial 1: Overview


Tutorial 2
Tutorial 5
Tutorial 4
Tutorial 3
Tutorial 6


NetworkAnalyst -- a web-based platform for gene expression profiling & biological network analysis

Computer and browser requirements

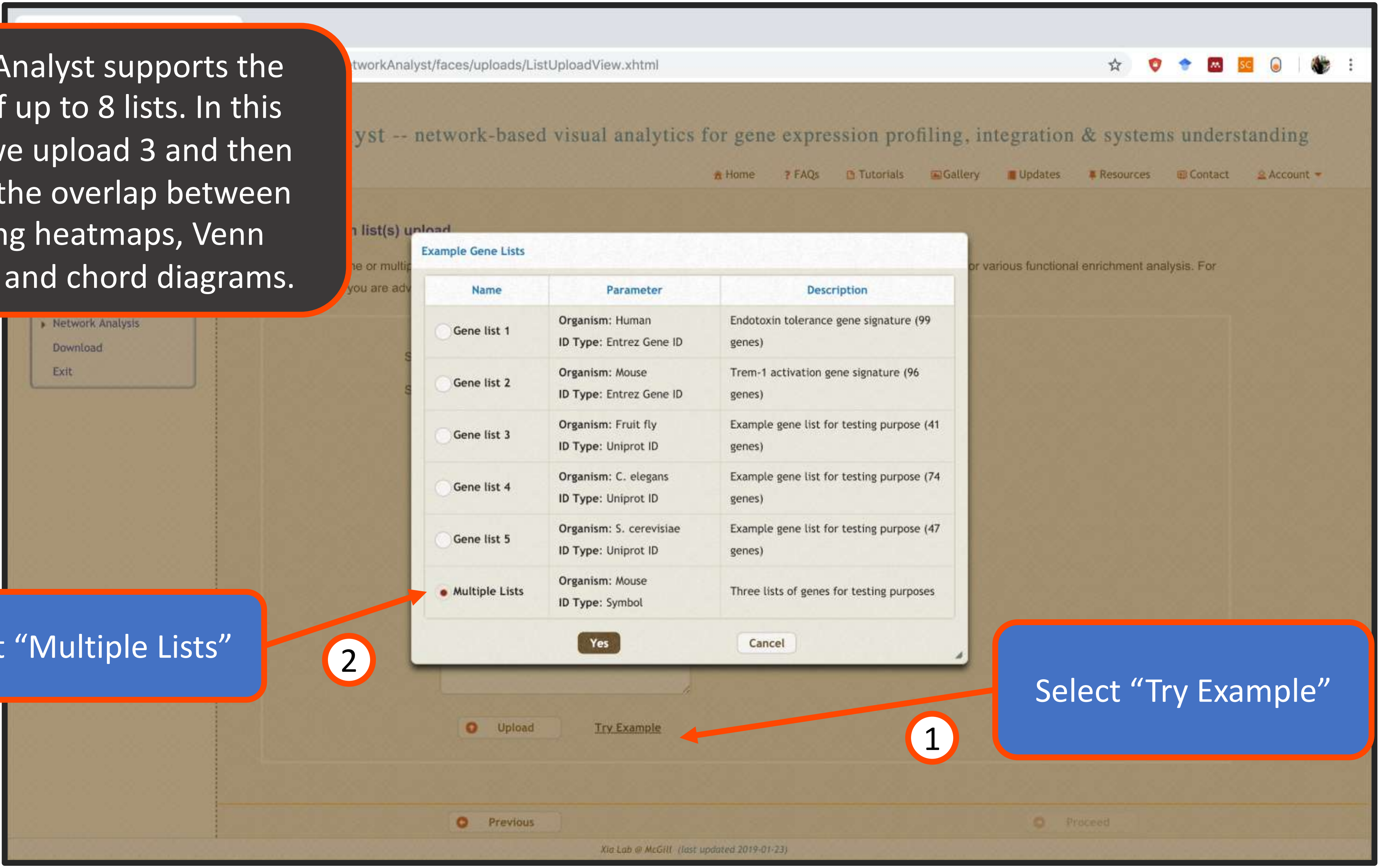
- A modern web browser with Java Script enabled
  - Supported browsers include Chrome, Safari, Firefox, and Internet Explorer 9+
- For best performance and visualization, use:
  - Latest version of Google Chrome
  - A computer with at least 4GB of physical RAM
  - A 15-inch screen or bigger (larger is better)
- Browser must be WebGL enabled for 3D network visualization
- 50MB limit for data upload
  - ~300 samples for gene expression data with 20 000 genes

Goals for this tutorial

- A challenge for ’omics analysis is to find ways to benefit from big data without being overwhelmed by it
- Visualizing key functions and relationships within large datasets is an integral part of this
- Networks capture the topology of statistical and functional relationships in graphical form – thus they are a powerful tool for ‘omics visual analytics
- The goal of this tutorial is to:
    - introduce NetworkAnalyst’s tools for visualizing 2D and 3D networks from a list of uploaded gene IDs
    - Project and visualize these genes within their biological networks to explore their relationships and collective functions

Select example data

NetworkAnalyst supports the upload of up to 8 lists. In this tutorial, we upload 3 and then compare the overlap between lists using heatmaps, Venn diagrams, and chord diagrams.



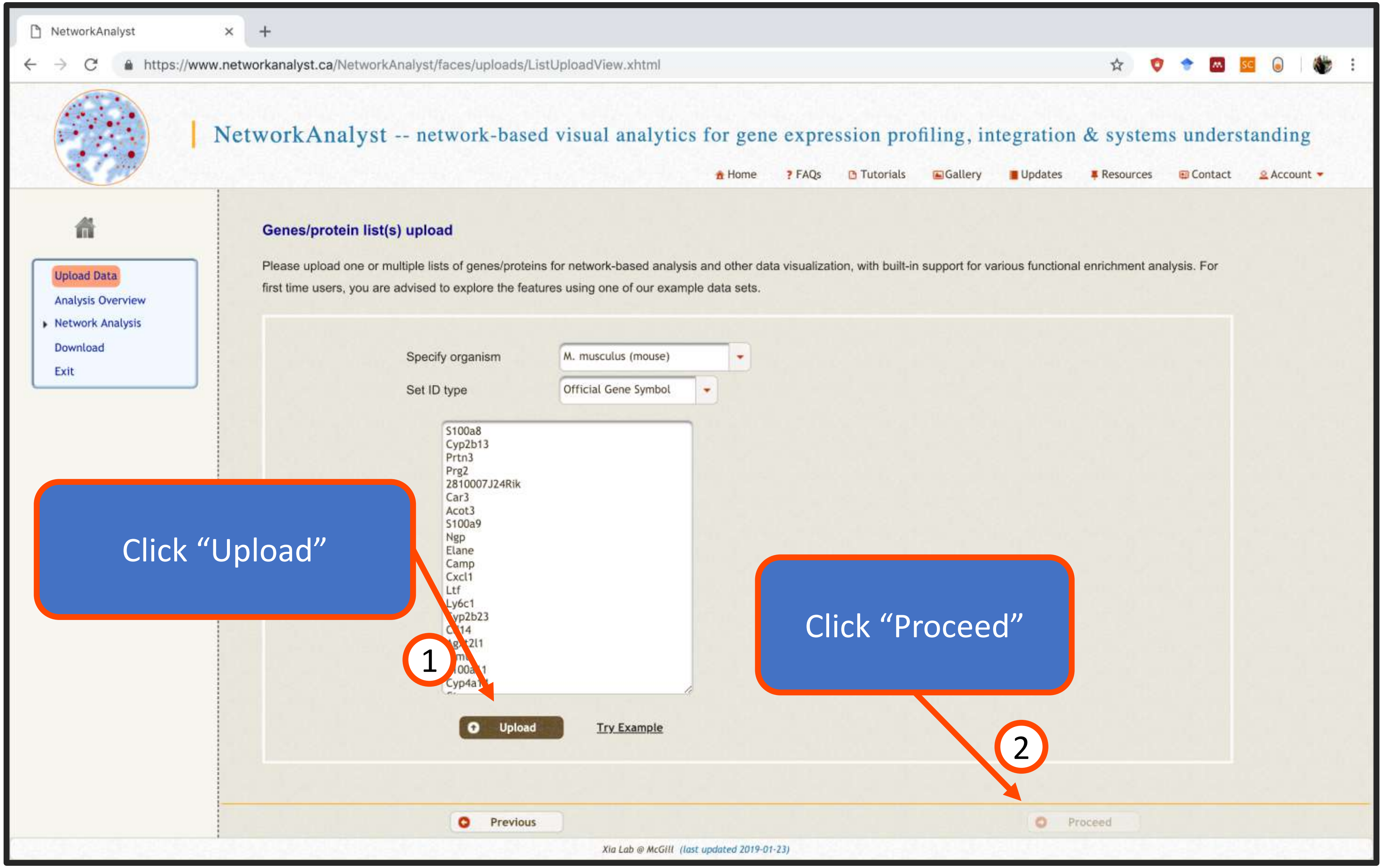
Example Gene Lists

Name	Parameter	Description
○ Gene list 1	**Organism:** Human **ID Type:** Entrez Gene ID	Endotoxin tolerance gene signature (99 genes)
○ Gene list 2	**Organism:** Mouse **ID Type:** Entrez Gene ID	Trem-1 activation gene signature (96 genes)
○ Gene list 3	**Organism:** Fruit fly **ID Type:** Uniprot ID	Example gene list for testing purpose (41 genes)
○ Gene list 4	**Organism:** C. elegans **ID Type:** Uniprot ID	Example gene list for testing purpose (74 genes)
○ Gene list 5	**Organism:** S. cerevisiae **ID Type:** Uniprot ID	Example gene list for testing purpose (47 genes)
● Multiple Lists	**Organism:** Mouse **ID Type:** Symbol	Three lists of genes for testing purposes

② Select “Multiple Lists”

① Select “Try Example”

Upload example data



Genes/protein list(s) upload

Please upload one or multiple lists of genes/proteins for network-based analysis and other data visualization, with built-in support for various functional enrichment analysis. For first time users, you are advised to explore the features using one of our example data sets.

Specify organism: M. musculus (mouse)

Set ID type: Official Gene Symbol

```
S100a8
Cyp2b13
Prtn3
Prg2
2810007J24Rik
Car3
Acot3
S100a9
Ngp
Elane
Camp
Cxcl1
Ltf
Ly6c1
yp2b23
C 14
g 2l1
m
100a 1
Cyp4a1
```

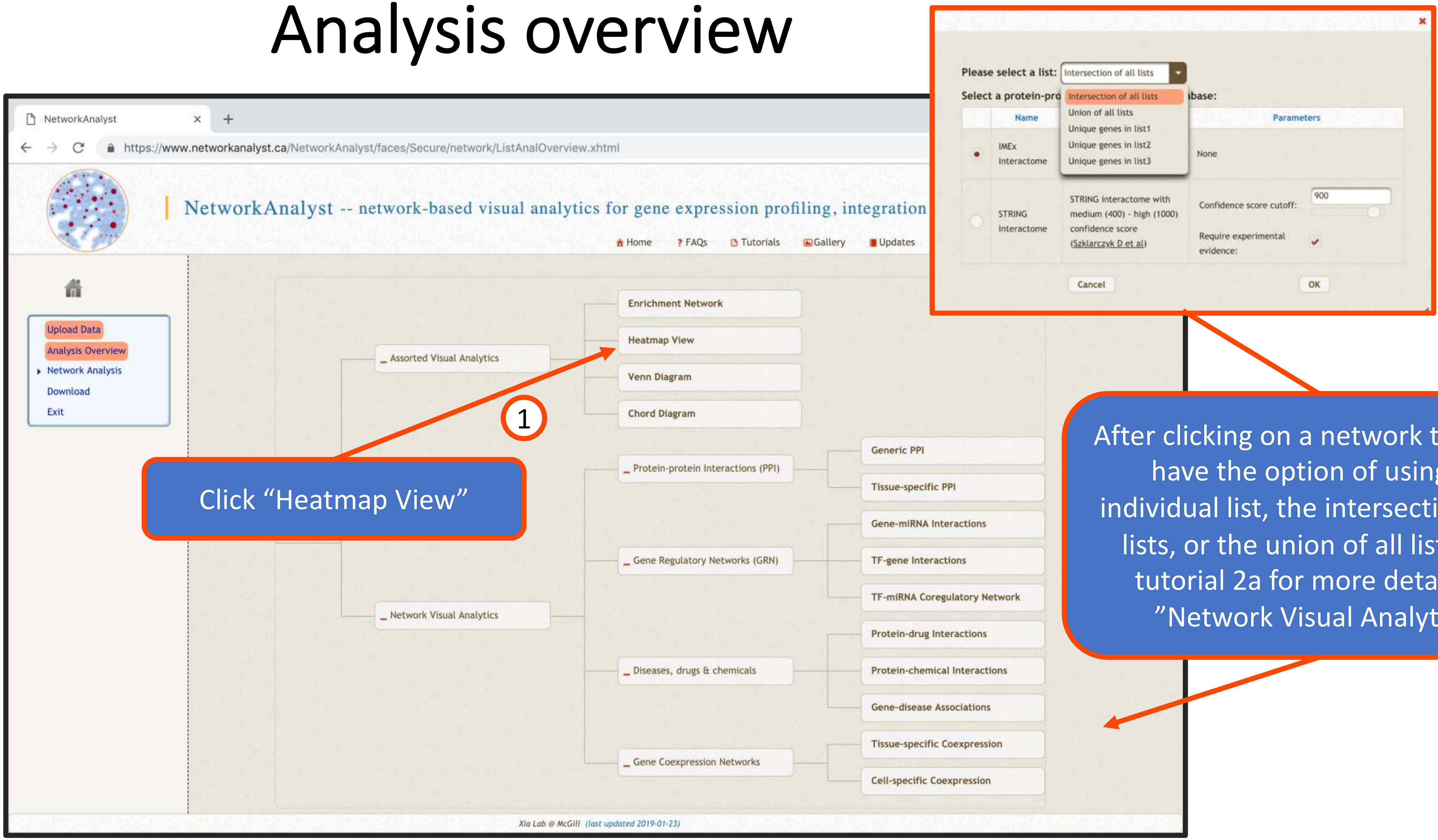
Click “Upload”

①

Click “Proceed”

②

Analysis overview



NetworkAnalyst

https://www.networkanalyst.ca/NetworkAnalyst/faces/Secure/network/ListAnalOverview.xhtml

NetworkAnalyst -- network-based visual analytics for gene expression profiling, integration

Xia Lab @ McGill (last updated 2019-01-23)

Please select a list: Intersection of all lists

Select a protein-pro… …base:

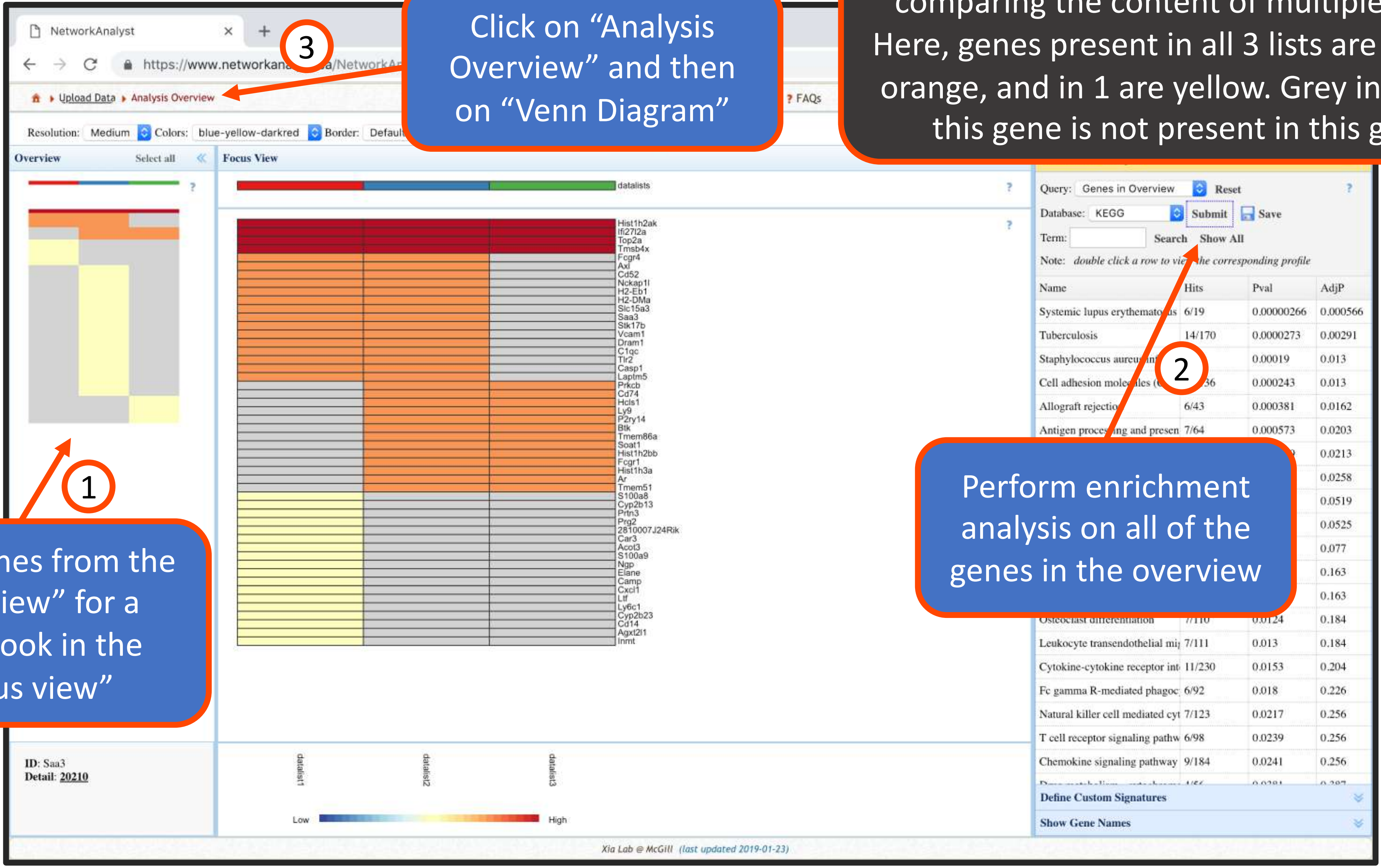
	Name		Parameters
●	IMEx Interactome		None
○	STRING Interactome	STRING interactome with medium (400) - high (1000) confidence score (Szklarczyk D et al)	Confidence score cutoff: 900 Require experimental evidence: ✓

Click “Heatmap View”

After clicking on a network type, you have the option of using an individual list, the intersection of all lists, or the union of all lists. See tutorial 2a for more details on ”Network Visual Analytics”

Heatmap view

A heatmap is a simple yet powerful way of comparing the content of multiple gene lists. Here, genes present in all 3 lists are red, in 2 are orange, and in 1 are yellow. Grey indicates that this gene is not present in this gene list.



Resolution: Medium Colors: blue-yellow-darkred Border: Defaul

ID: Saa3
Detail: 20210

Query: Genes in Overview

Database: KEGG

Term:

Note: double click a row to vie the corresponding profile

Name	Hits	Pval	AdjP
Systemic lupus erythemato us	6/19	0.00000266	0.000566
Tuberculosis	14/170	0.0000273	0.00291
Staphylococcus aureu in		0.00019	0.013
Cell adhesion mole ules (	36	0.000243	0.013
Allograft rejectio	6/43	0.000381	0.0162
Antigen proces ing and presen	7/64	0.000573	0.0203
			0.0213
			0.0258
			0.0519
			0.0525
			0.077
			0.163
			0.163
Osteoclast differentiation	7/110	0.0124	0.184
Leukocyte transendothelial mi	7/111	0.013	0.184
Cytokine-cytokine receptor int	11/230	0.0153	0.204
Fc gamma R-mediated phagoc	6/92	0.018	0.226
Natural killer cell mediated cyt	7/123	0.0217	0.256
T cell receptor signaling pathw	6/98	0.0239	0.256
Chemokine signaling pathway	9/184	0.0241	0.256

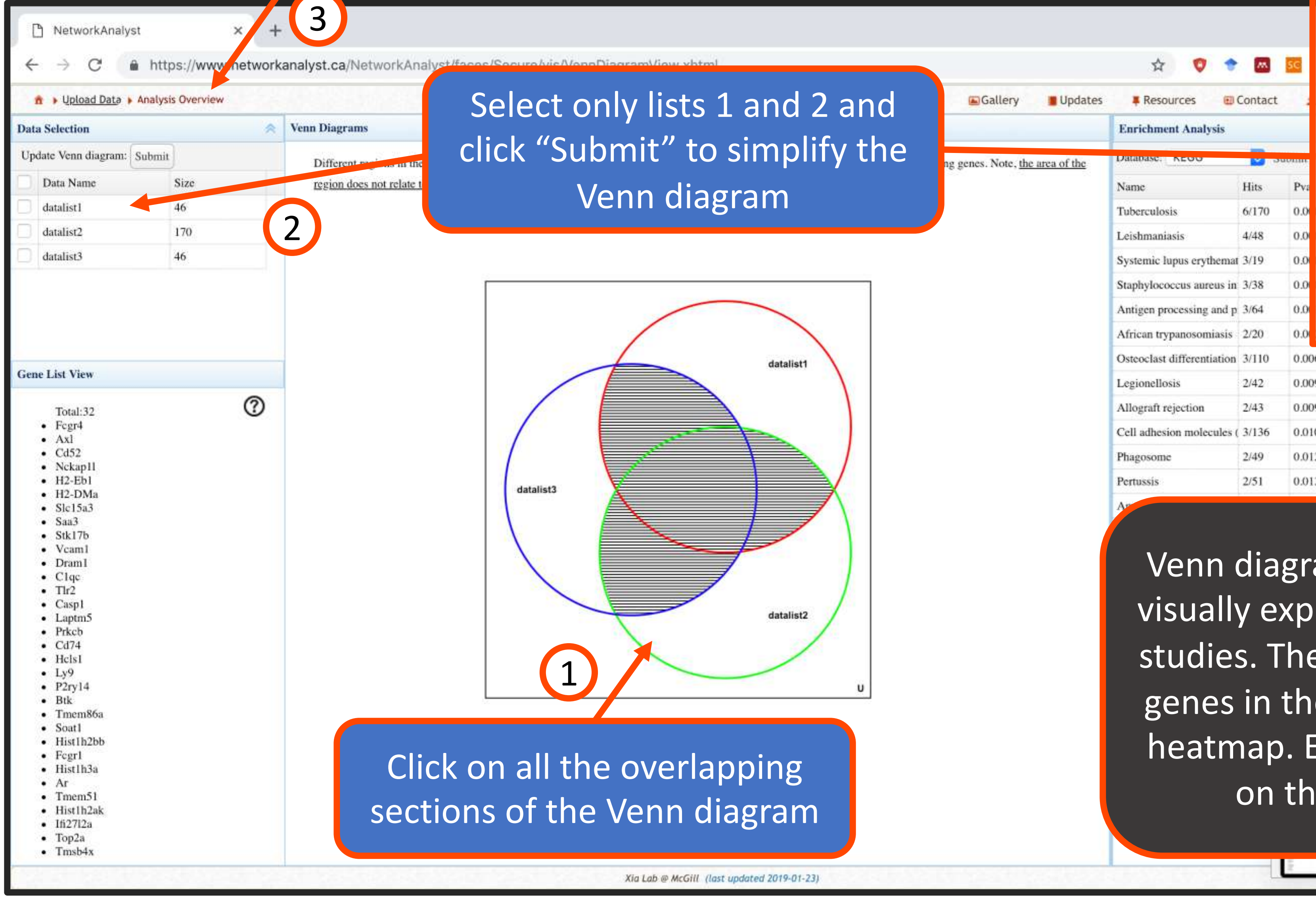
③ Click on “Analysis Overview” and then on “Venn Diagram”

① Select genes from the “Overview” for a closer look in the “Focus view”

② Perform enrichment analysis on all of the genes in the overview

Interactive Venn Diagram

When finished, click "Analysis Overview" and then "Chord Diagram"



Update Venn diagram: Submit

☐	Data Name	Size
☐	datalist1	46
☐	datalist2	170
☐	datalist3	46

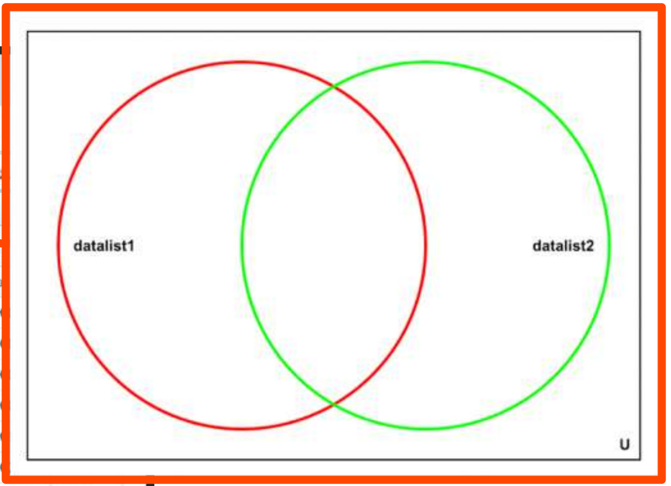
Gene List View

Total:32
- Fcgr4
- Axl
- Cd52
- Nckap1l
- H2-Eb1
- H2-DMa
- Slc15a3
- Saa3
- Stk17b
- Vcam1
- Dram1
- C1qc
- Tlr2
- Casp1
- Laptm5
- Prkcb
- Cd74
- Hcls1
- Ly9
- P2ry14
- Btk
- Tmem86a
- Soat1
- Hist1h2bb
- Fcgr1
- Hist1h3a
- Ar
- Tmem51
- Hist1h2ak
- Ifi27l2a
- Top2a
- Tmsb4x

Database: KEGG

Name	Hits	Pva	
Tuberculosis	6/170	0.0	
Leishmaniasis	4/48	0.0	
Systemic lupus erythemat	3/19	0.0	
Staphylococcus aureus in	3/38	0.0	
Antigen processing and p	3/64	0.0	
African trypanosomiasis	2/20	0.0	
Osteoclast differentiation	3/110	0.00603	0.184
Legionellosis	2/42	0.00903	0.224
Allograft rejection	2/43	0.00945	0.224
Cell adhesion molecules (	3/136	0.0108	0.231
Phagosome	2/49	0.0122	0.233
Pertussis	2/51	0.0131	0.233

Select only lists 1 and 2 and click "Submit" to simplify the Venn diagram



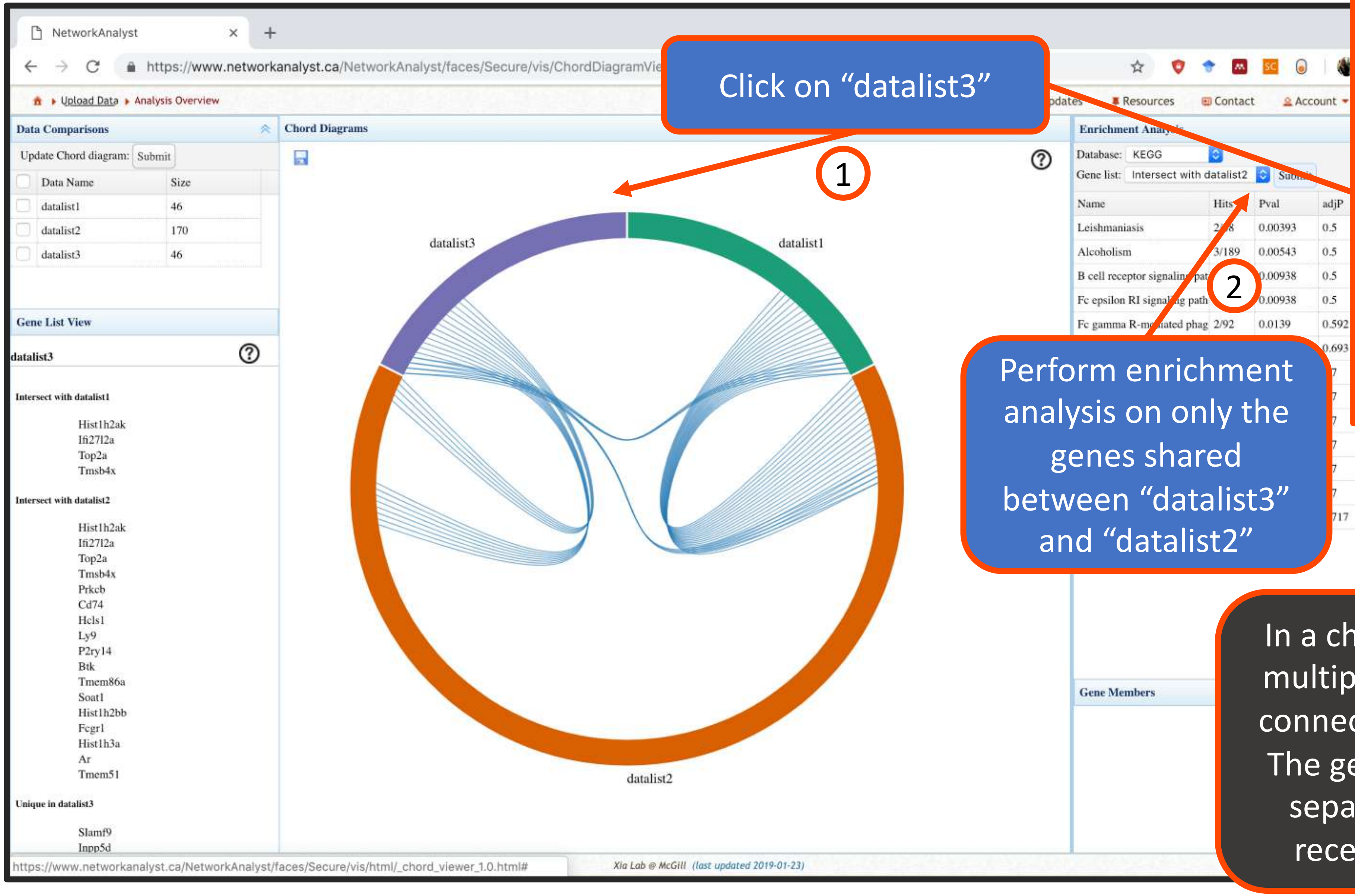
Click on all the overlapping sections of the Venn diagram

Venn diagrams are a simple but effective tool for visually exploring overlap of results from different studies. The "Gene List View" lists the union of all genes in the currently highlighted sections of the heatmap. Enrichment analysis can be performed on the genes in the "Gene List View".

Interactive chord diagram

Click on “datalist3”

①



Data Comparisons

Update Chord diagram: Submit

☐	Data Name	Size
☐	datalist1	46
☐	datalist2	170
☐	datalist3	46

Gene List View

datalist3

Intersect with datalist1

Hist1h2ak
Ifi27l2a
Top2a
Tmsb4x

Intersect with datalist2

Hist1h2ak
Ifi27l2a
Top2a
Tmsb4x
Prkcb
Cd74
Hcls1
Ly9
P2ry14
Btk
Tmem86a
Soat1
Hist1h2bb
Fcgr1
Hist1h3a
Ar
Tmem51

Unique in datalist3

Slamf9
Inpp5d

Enrichment Analysis

Database: KEGG

Gene list: Intersect with datalist2

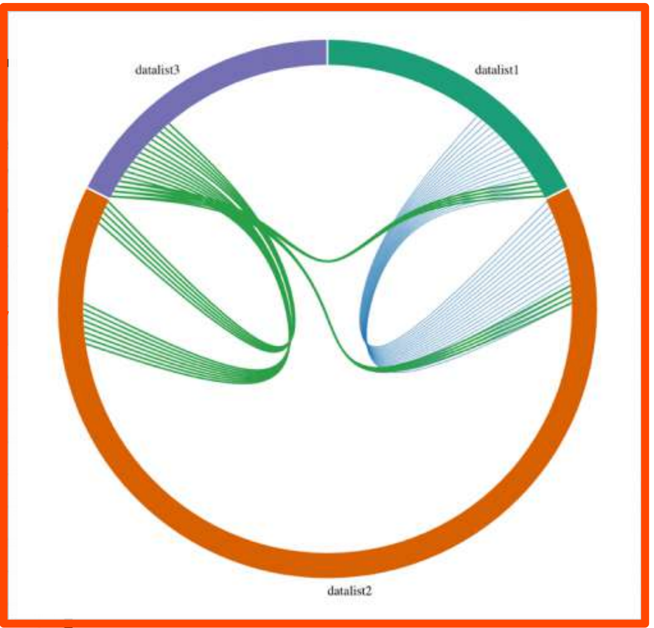
Name	Hits	Pval	adjP
Leishmaniasis	2/[illegible]8	0.00393	0.5
Alcoholism	3/189	0.00543	0.5
B cell receptor signalin[illegible] pat		0.00938	0.5
Fc epsilon RI signal[illegible]g path		0.00938	0.5
Fc gamma R-m[illegible]ated phag	2/92	0.0139	0.592

https://www.networkanalyst.ca/NetworkAnalyst/faces/Secure/vis/html/_chord_viewer_1.0.html#

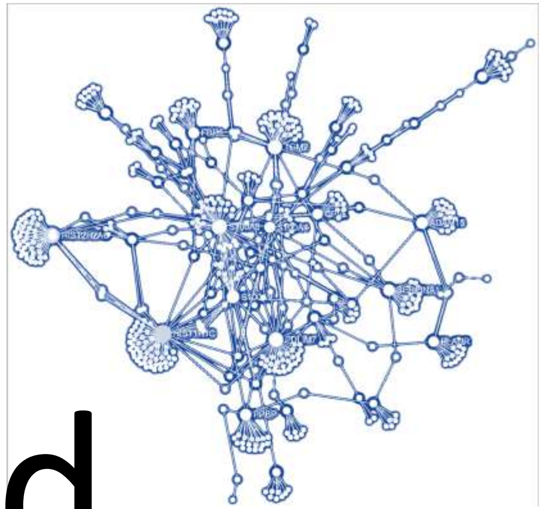
Xia Lab @ McGill (last updated 2019-01-23)

②

Perform enrichment analysis on only the genes shared between “datalist3” and “datalist2”



In a chord diagram, any gene that is in multiple lists is represented as a chord connecting the two diagram segments. The genes in the “Gene List View” are separated in reference to the most recently clicked diagram segment.



The End

For more information, visit the ***FAQs****,* ***Tutorials****,* ***Resources****,*
and ***Contact*** *pages on www.networkanalyst.ca*