Omics Data Science Training Course

Winter 2024

XiaLab Team

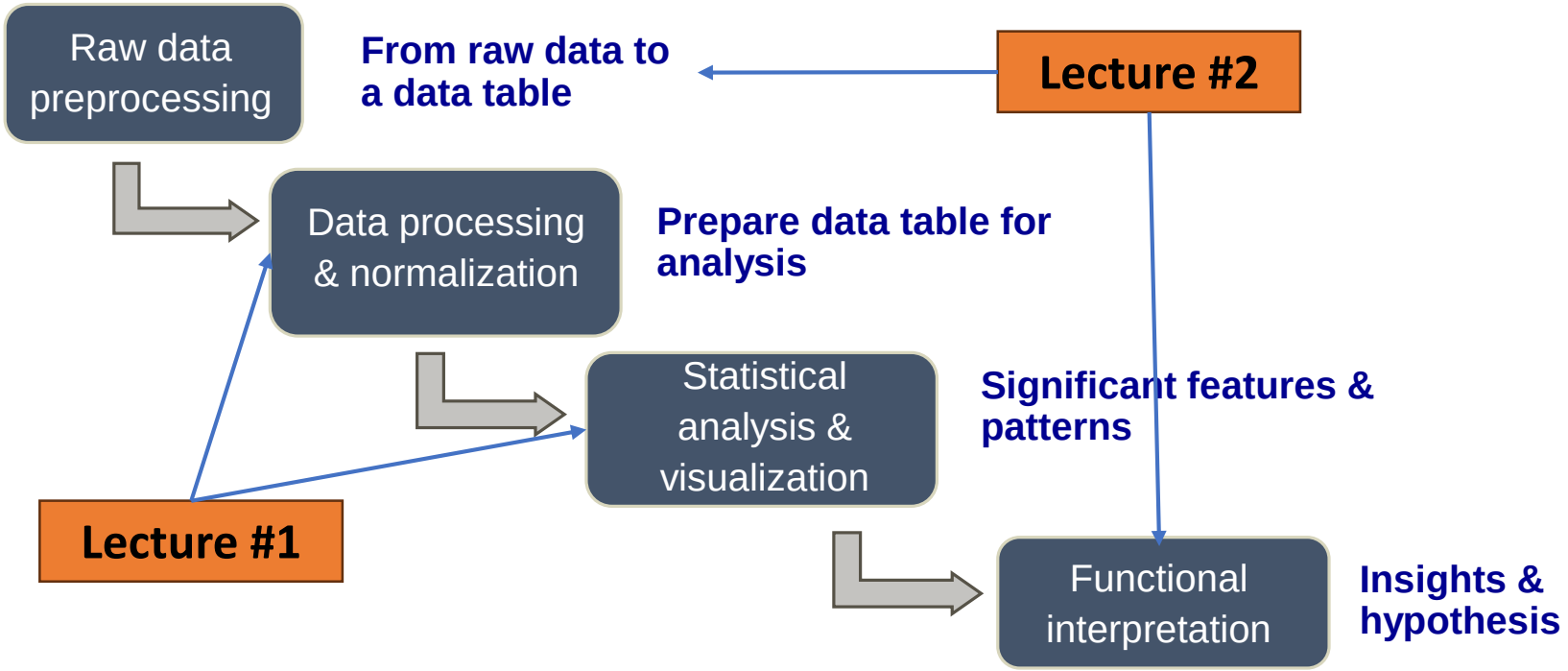
Jianguo (Jeff) Xia, PhD MD
Canada Research Chair in
Bioinformatics & big data analytics
jeff.xia@xialab.ca

Guangyan (Joe) Zhou, PhD
Transcriptomics, visual analytics
guangyan.zhou@xialab.ca

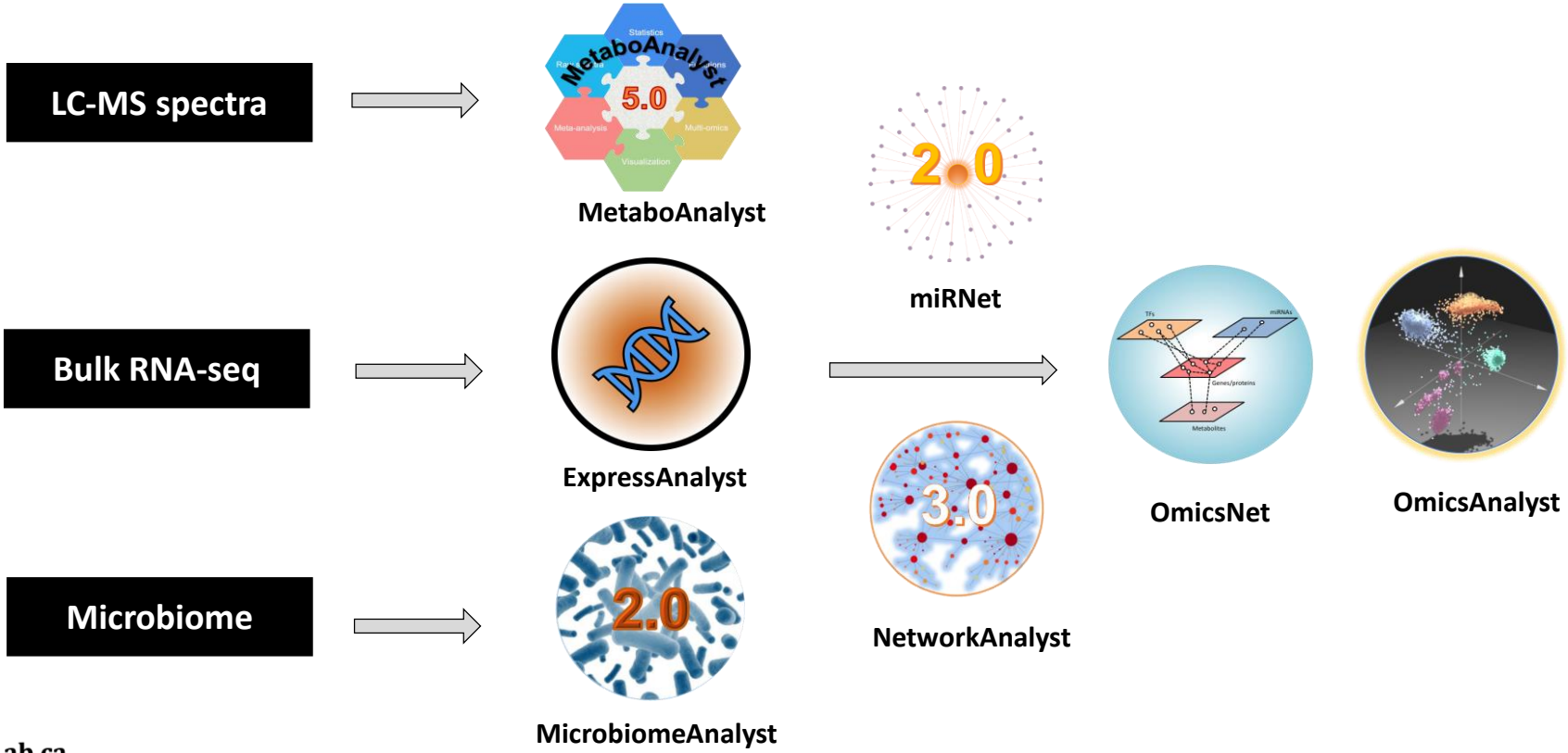
Zhiqiang (Qiang) Pang, PhD
Metabolomics, big data analytics
zhiqiang.pang@xialab.ca

Yao Lu, PhD Candidate
Microbiomics, multi-omics
yao.lu@xialab.ca

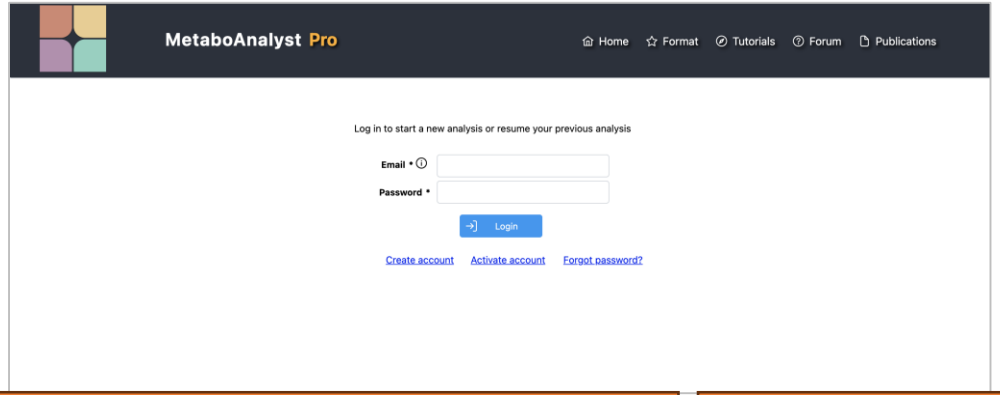
Two foundation lectures (prerequisite)



Raw data ➔ statistics ➔ networks ➔ functions



More on the “pro” tools



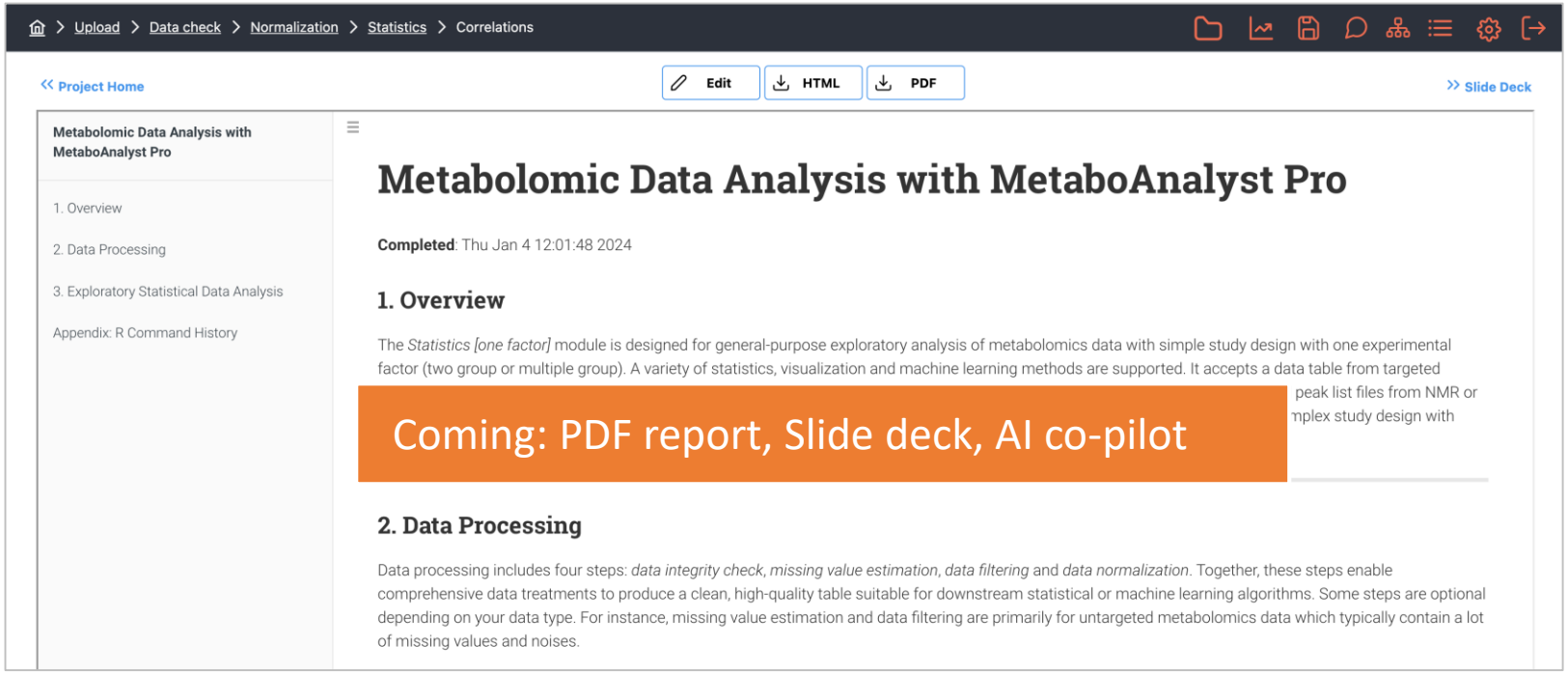
Community version
~ 5000 users / day
~ 100s concurrent users

A total of six tools have their “pro” versions

- https://pro.metaboanalyst.ca
- https://pro.microbiomeanalyst.ca
- https://pro.expressanalyst.ca
- https://pro.omicsnet.ca
- https://pro.omicsanalyst.ca
- https://pro.mirnet.ca

1. Dedicated computing
2. Live report & project management
3. More stable
4. Prioritized support

Live Report & Project Management



Our Resources

Recordings & slides

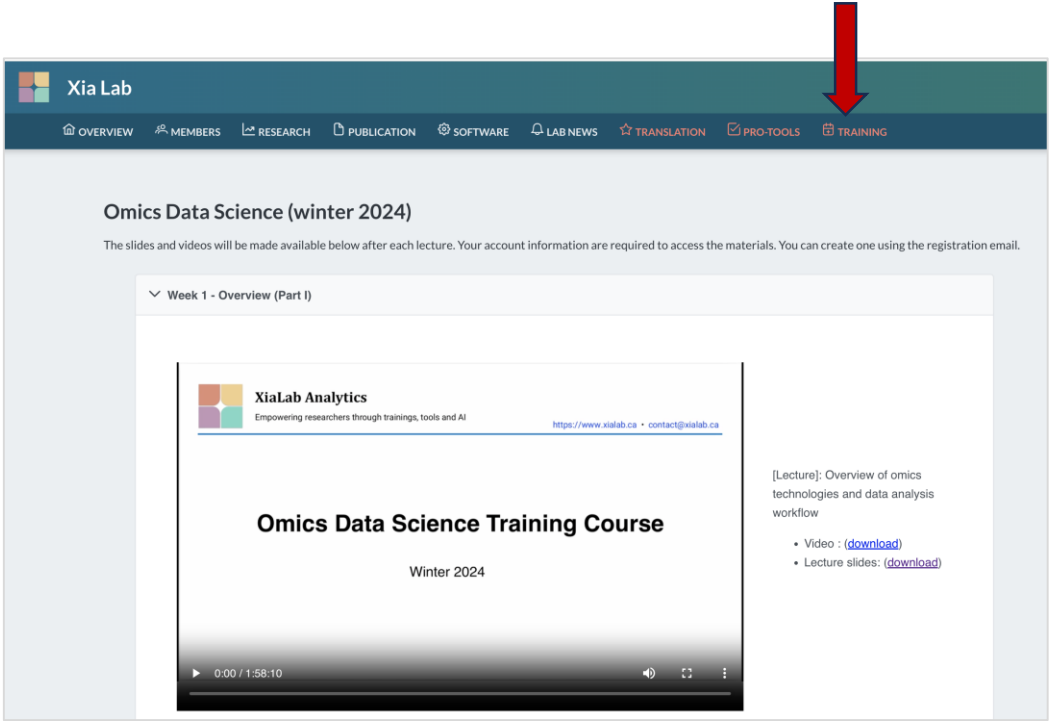
- https://www.xialab.ca under the "**Training**" tab within 3 days after lecture
- Faster with Firefox

Community tool:

- https://www.####.ca

Pro Tools:

- https://pro.####.ca
  - You will be assigned to one of the cloud nodes

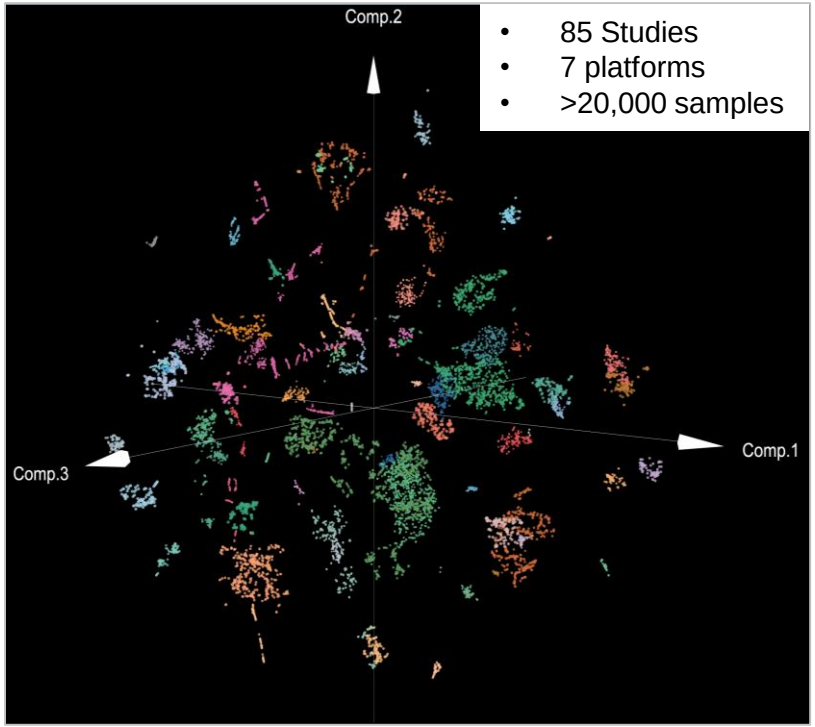


Our Syllabus

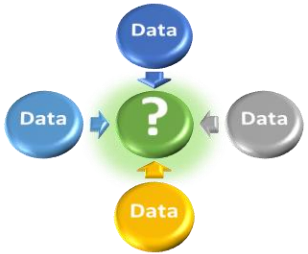
Topic	Date	Lecture	Lab
Omics Data Science Foundations	Jan. 6	Omics data processing, statistics and visualization	--
	Jan. 13	From raw data to functional insights	--
Transcriptomics	Jan. 20	Gene expression data analysis (part I)	ExpressAnalyst & NetworkAnalyst
	Jan. 27	Gene expression data analysis (part II)	ExpressAnalyst & Seq2Fun
Biological Networks	Feb. 3	Biological network analysis & gene regulatory networks	NetworkAnalyst & miRNet
Proteomics & Biomarker	Feb. 10	Proteomics & biomarker analysis	ExpressAnalyst & MetaboAnalyst
Metabolomics	Feb. 17	Targeted metabolomics data analysis	MetaboAnalyst
	Feb. 24	LC-MS untargeted metabolomics data analysis	MetaboAnalyst
Microbiomics	Mar. 2	Marker gene data analysis	MicrobiomeAnalyst
	Mar. 9	Shotgun metagenomics data analysis	MicrobiomeAnalyst
Multi-omics	Mar. 16	Knowledge-driven multi-omics integration	OmicsNet
	Mar. 23	Data-driven multi-omics integration	OmicsAnalyst

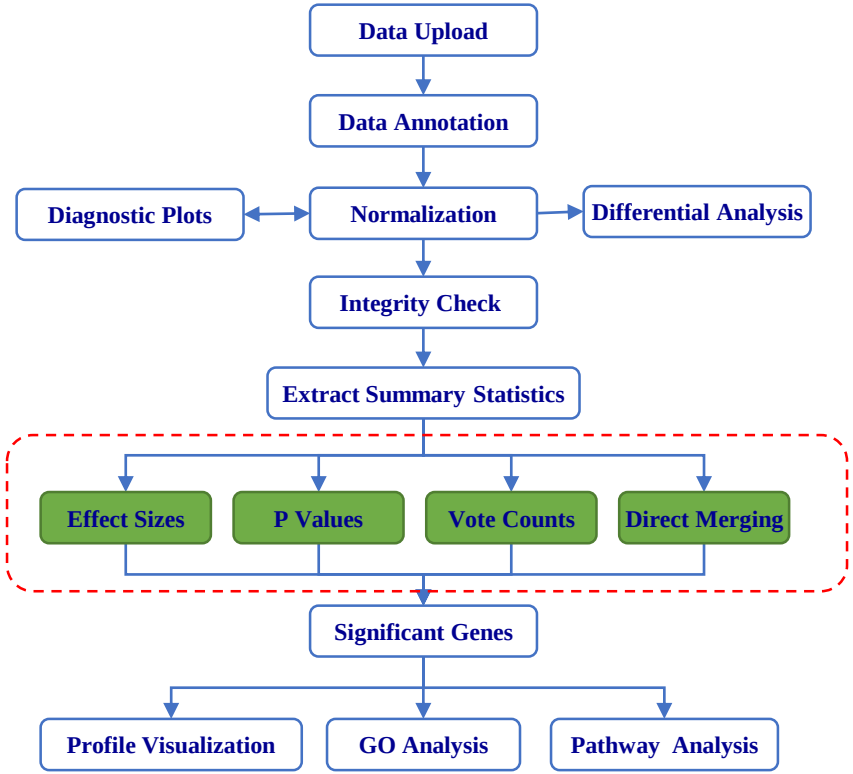
Recap: meta-analysis

- Integrating results from multiple independent studies
- Based on summary statistics that are more “integrable”
  - P-values
  - Effect sizes
- Apply “batch effect” adjustment to reduce study specific effects

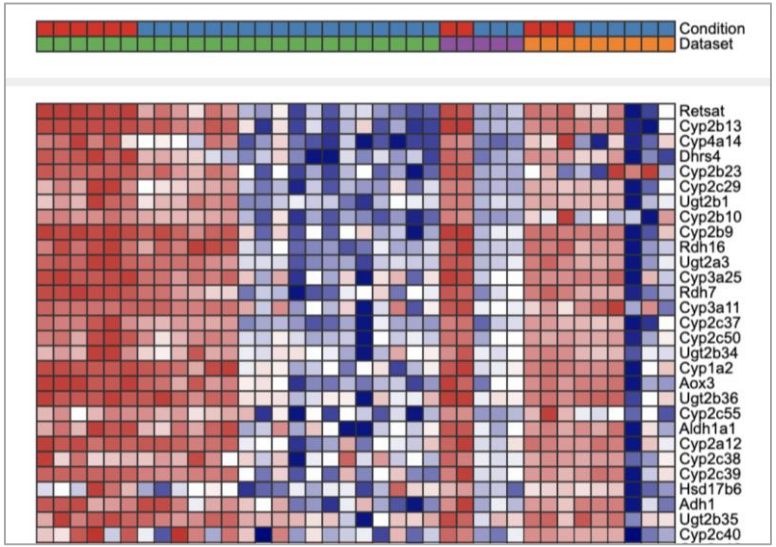


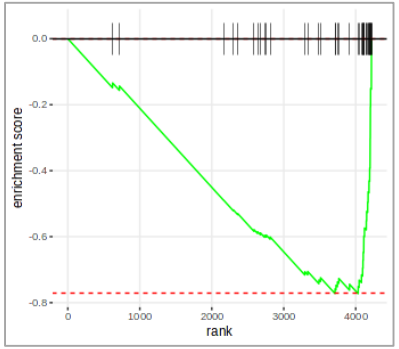
Feature-level meta-analysis

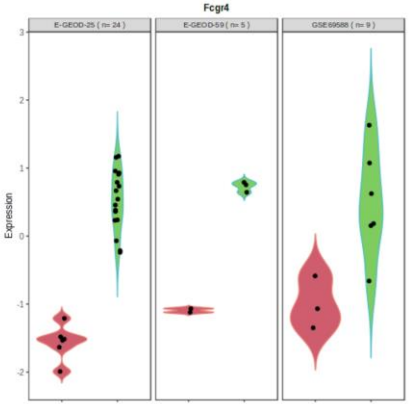




Visualization across different studies

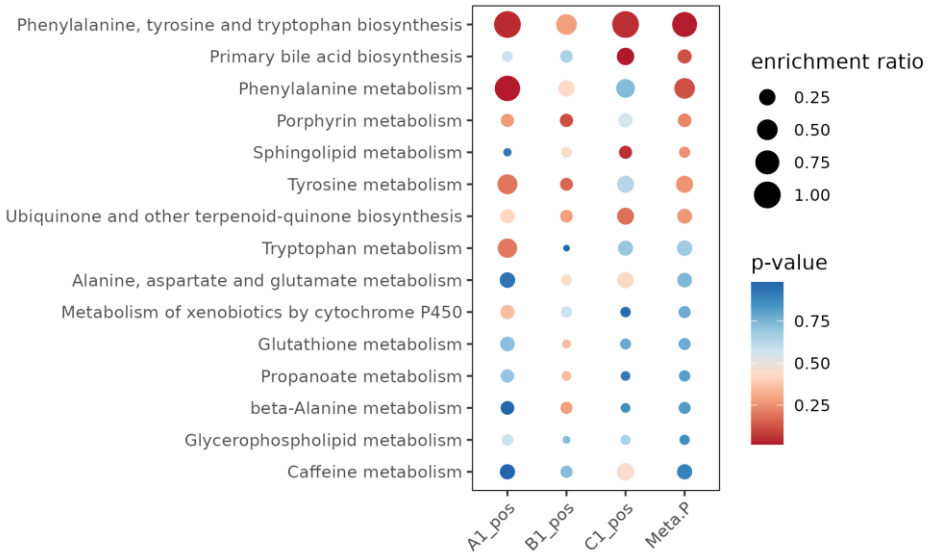






Meta-analysis at pathway level

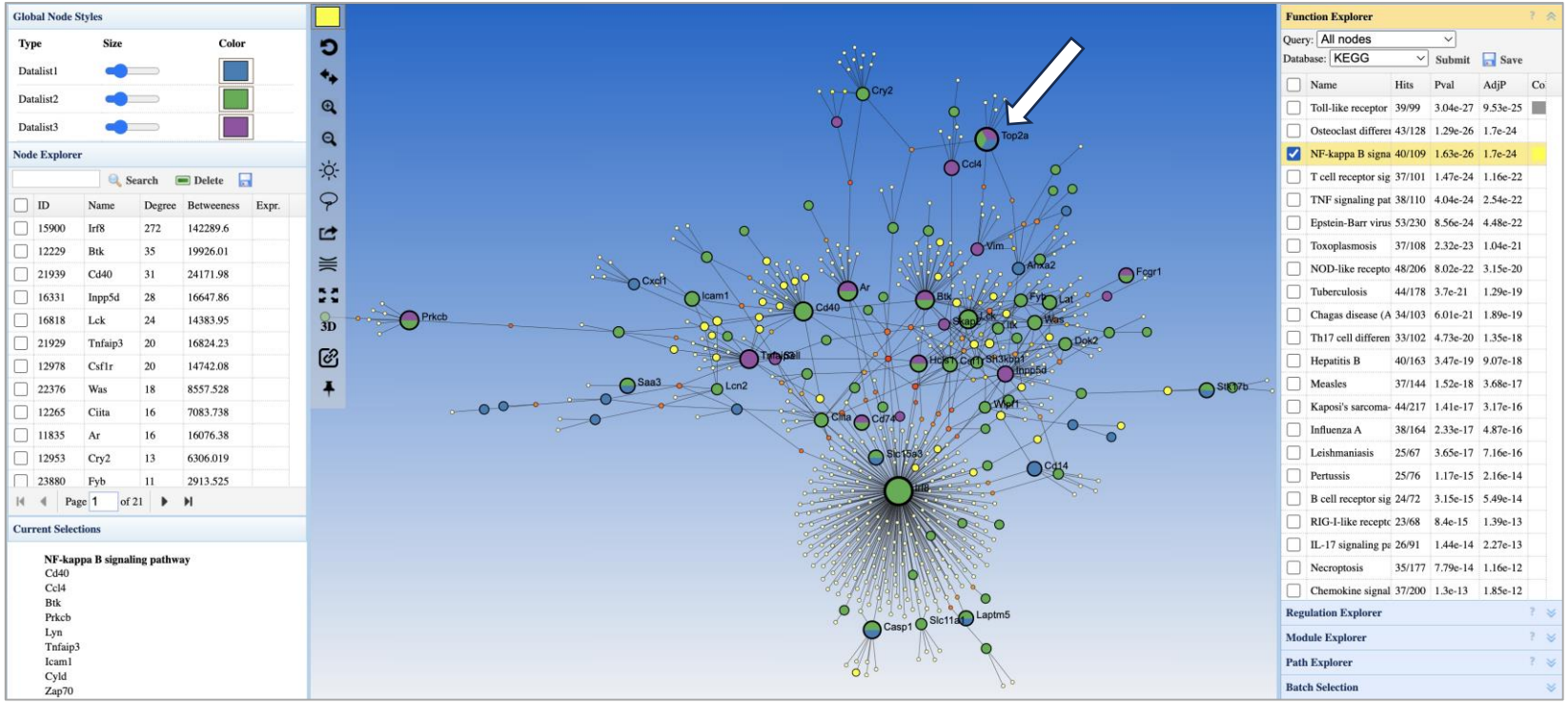
- Summary statistics can be computed at different levels such as pathway enrichment p-values
- Treat pathways as operational units (instead of features)
  - Useful when features are similar but not identical. Such as in multi-omics, untargeted metabolomics based on MS peaks



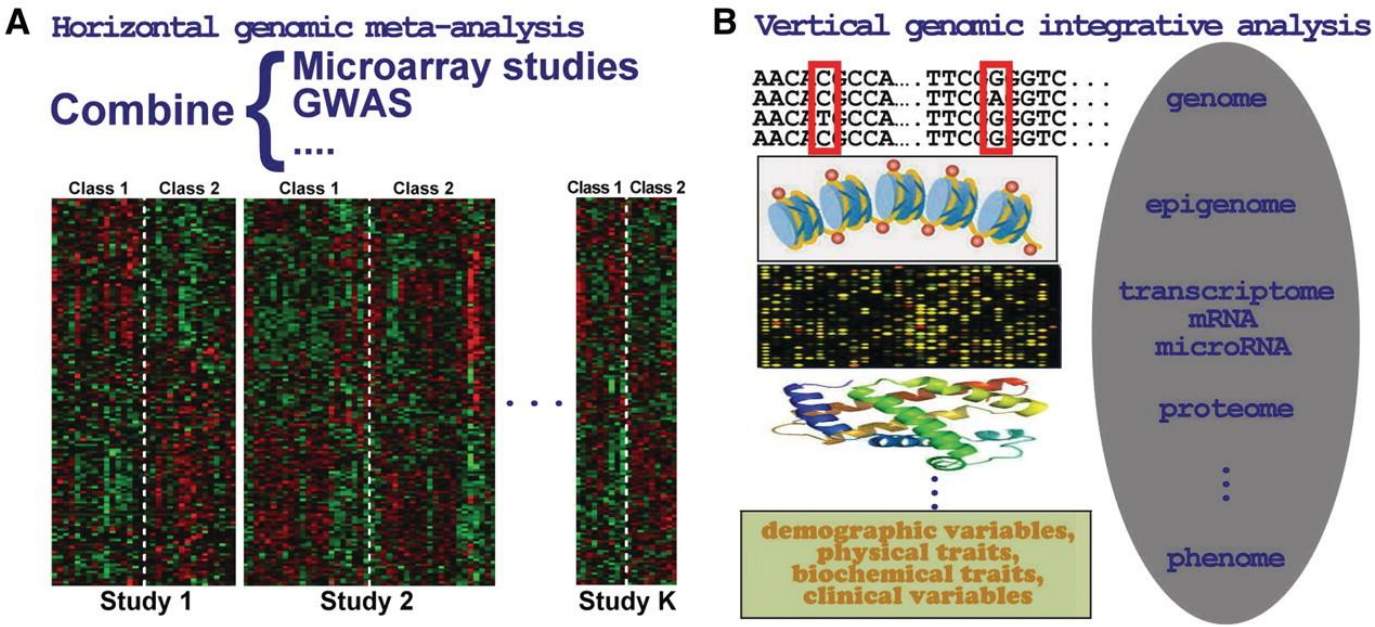
Network is “naturally” integrative

1. Obtain gene signatures from different studies
2. Enter them into NetworkAnalyst
   - Signatures separated by “***//***”
3. View their direct relationships & shared context in networks
   - Shared genes will be shown as pie chart indicating source
   - Similar genes will be close to each other
     - They will share the same context (i.e. interaction partners)
4. Detect enriched functions or common themes

Using networks to integrate gene signatures



Horizontal & vertical integration



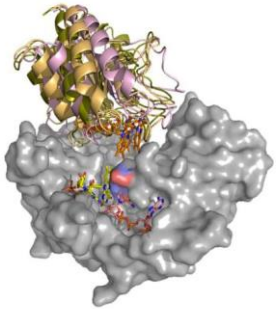
https://doi.org/10.1093/nar/gkr1265

Schedule for today

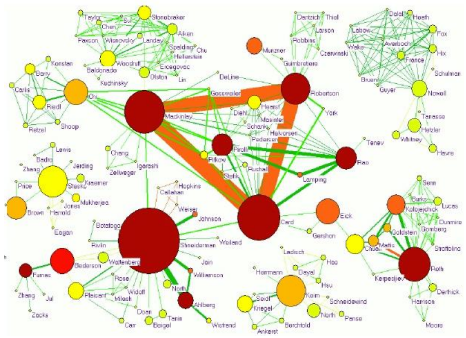
Time	Topics
9:00 – 9:15	General introduction & recap
9:15 – 9:40	Network analysis and visualization
9:45 – 10.15	NetworkAnalyst live demo & hands on
10:20 – 10:50	Gene regulatory networks
10:55 – 11:25	miRNet live demo & hands on
Summary and discussion	

Goals of omics data analytics

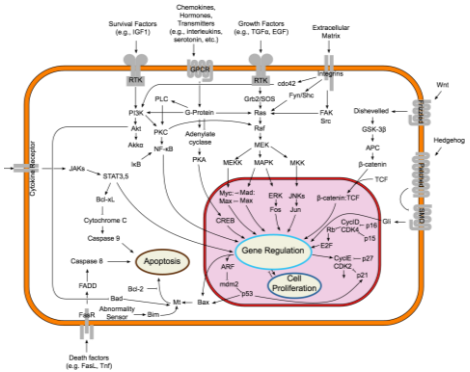
Individual Molecules



Omics & Multi-omics

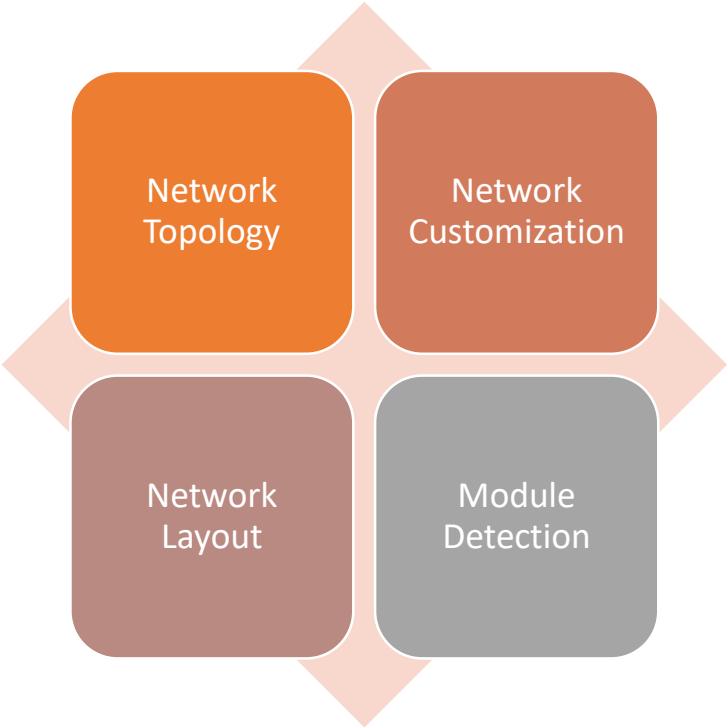


Systems Biology



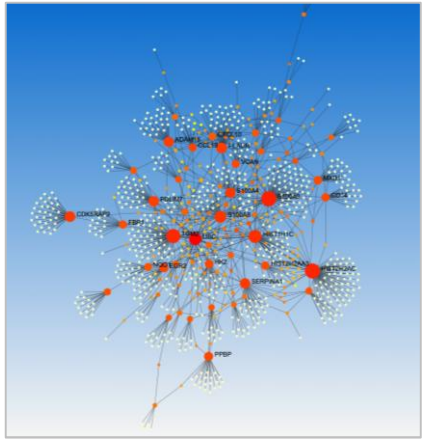
Mechanism ← **Patterns, functions, hypothesis** → Knowledge

Network Analysis

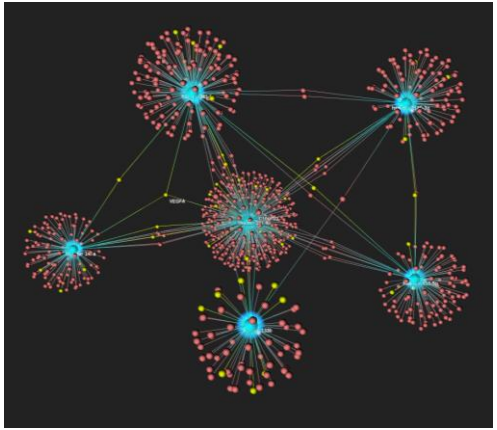


Network Topology

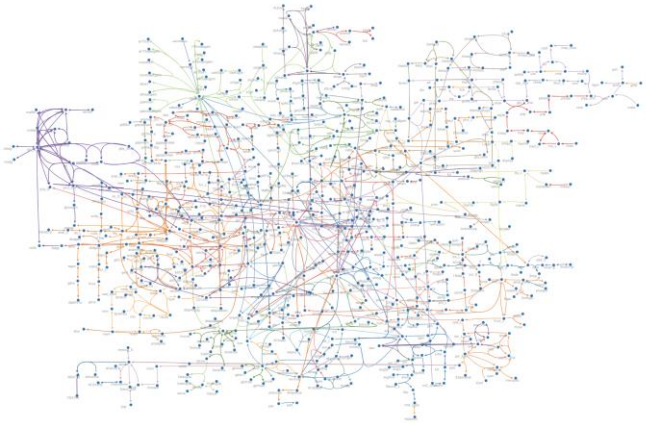
Common types of networks



PPI network



miRNA-gene target

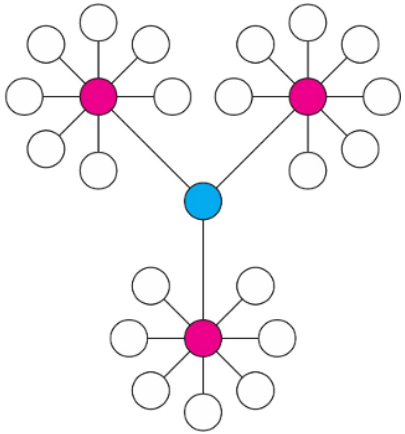


Metabolic networks

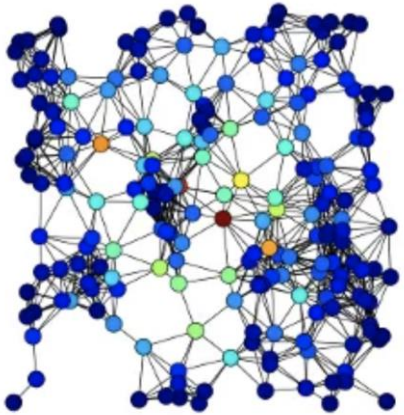
Network topology – centrality measures

- Degree centrality
  - Number of links to other nodes
  - Popular or densely connected nodes
- Betweenness centrality
  - How many times shortest paths cross a node
  - Nodes influence the information flow around a system
- Closeness centrality
  - Average farness (inverse of shortest distance) to all other nodes. Nodes with a high closeness score have the shortest distances to all other nodes
  - Nodes that can influence the entire network most quickly

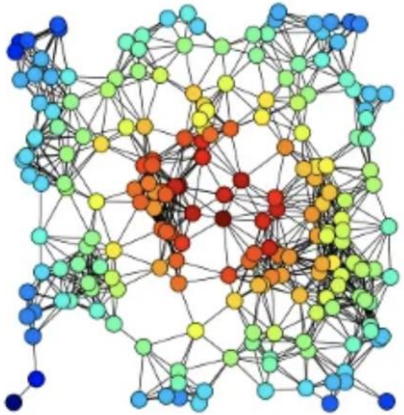
Illustrations



High degree centrality



High betweenness centrality

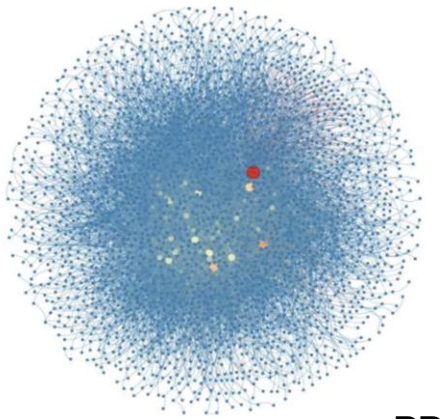


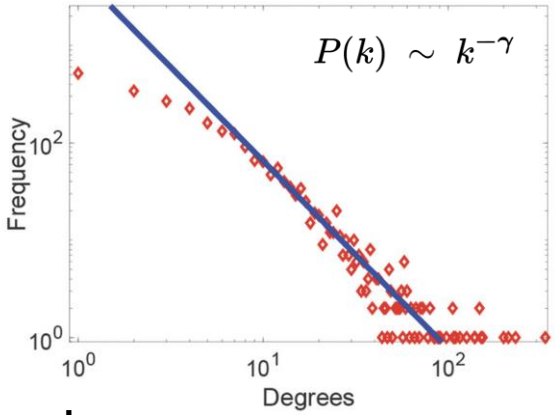
High closeness centrality

Scale-free network

A network whose degree distribution follows a power law

- Scale-free networks are more robust against failure (random node failure).
- Scale-free networks are more vulnerable against non-random attacks.
- Scale-free networks have short average path lengths.





PPI network

Network customization

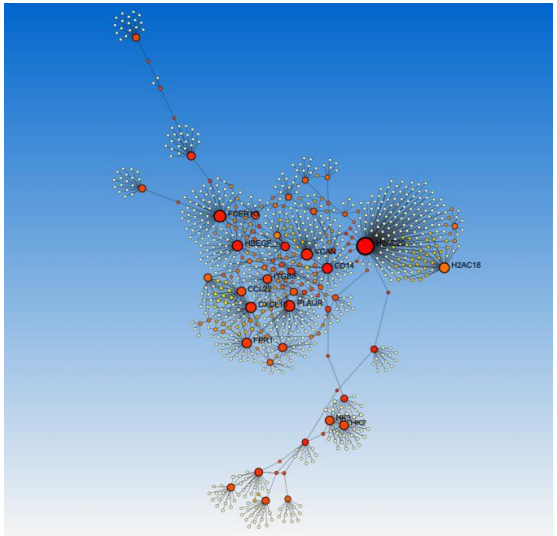
Large networks are difficult to view and interpret. How can we reduce the network size while keep its main structure and main features?

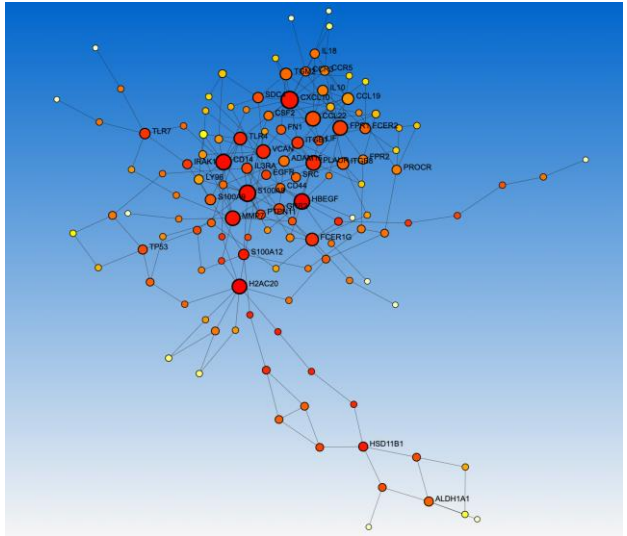
- **Zero-order Network**: keep only seed proteins that directly interact with each other - suitable when there are a large number of seeds;
- **Minimum Network**: compute pair-wise shortest paths between all seed nodes and remove the nodes that are not on the shortest paths.
- **Steiner Forest Network**: this is similar to minimum network, but based on a fast heuristic Prize-collecting Steiner Forest (PCSF) algorithm
- **Batch Exclusion**: Enter a list of genes that you want to exclude from the resulting network.
- **Tissue Filter**: filter genes according to gene expression measured in specific tissues.

Aiming for 200 – 2000 nodes

Minimum network

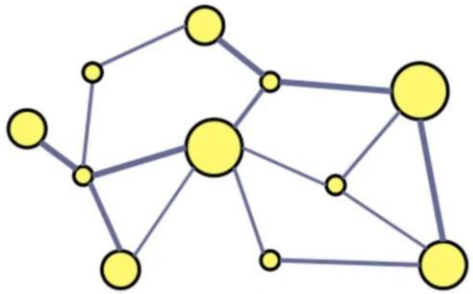
- Based on pair-wise shortest paths to compute the minimum nodes to be introduced in order to connect all seeds (queries)?

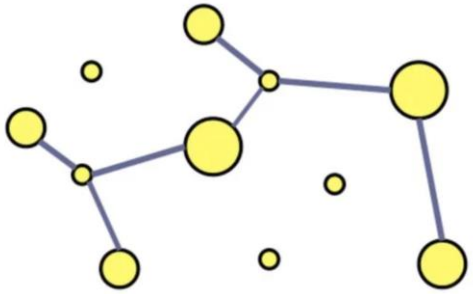




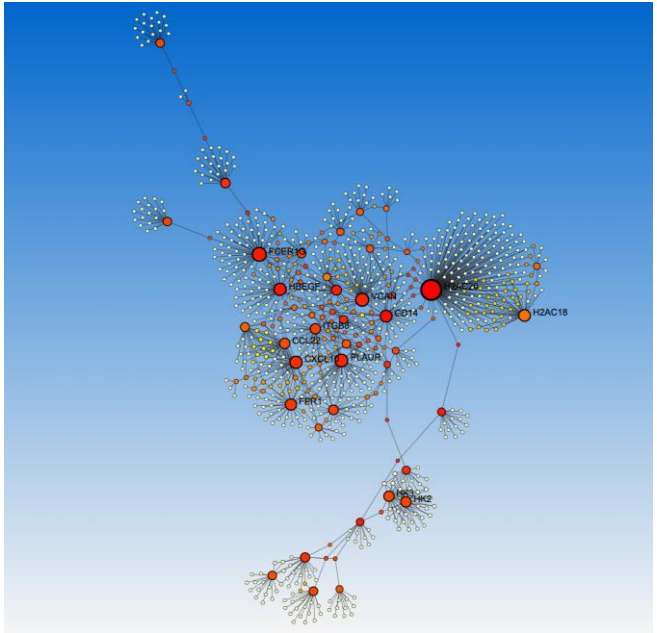
Prize-collecting Steiner Forest (PCSF) method

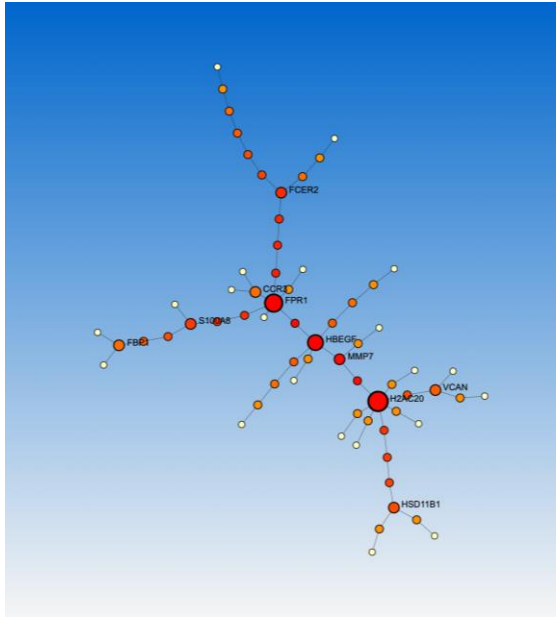
- The goal is to establish connections between pairs of nodes in a given network at minimum **cost**.
- The cost is customizable based on different context
  - ➔ We can adjust weights to encourage traveling through “seeds”
- A “global” optimization problem





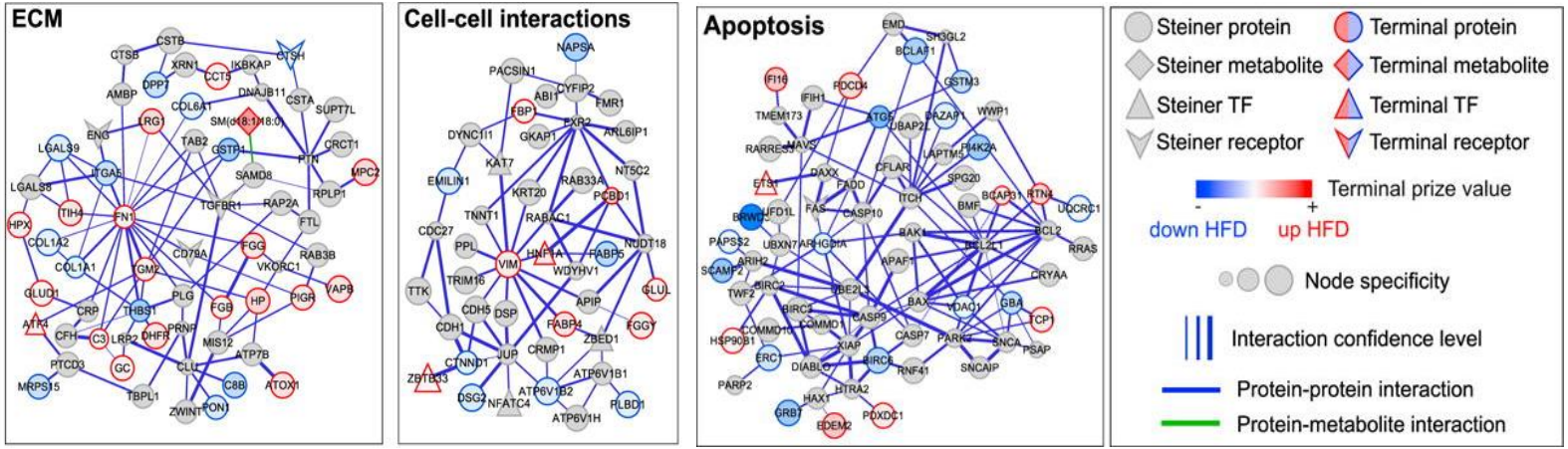
PCSF network





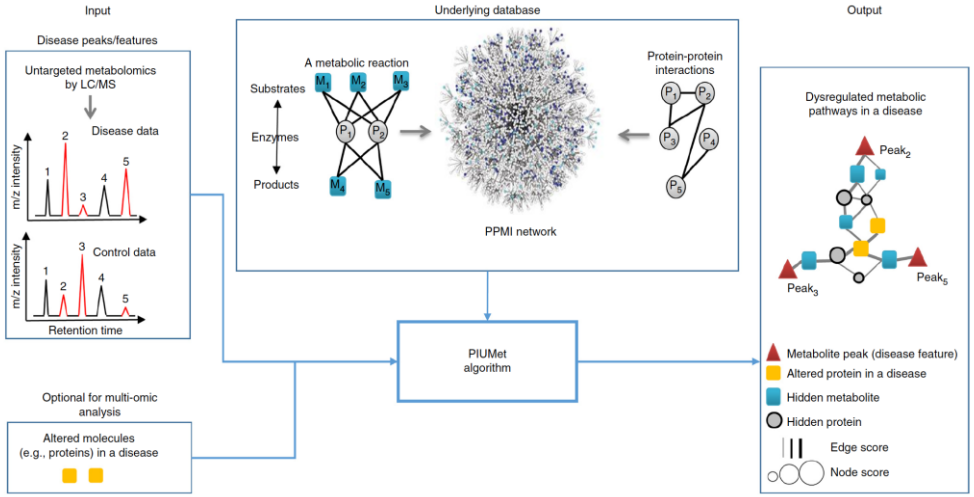
PCSF applications in biology

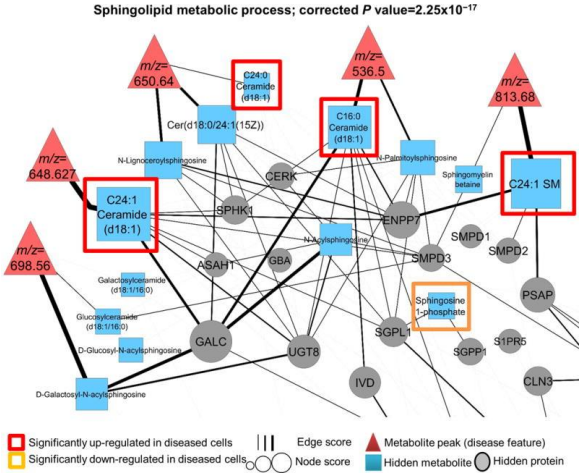
Computing PCSF sub-networks to select meaningful biological processes



Cell Rep. 2017;21(11):3317-3328.

Revealing disease-associated pathways by PCSF





Integrating metabolic network with PPI

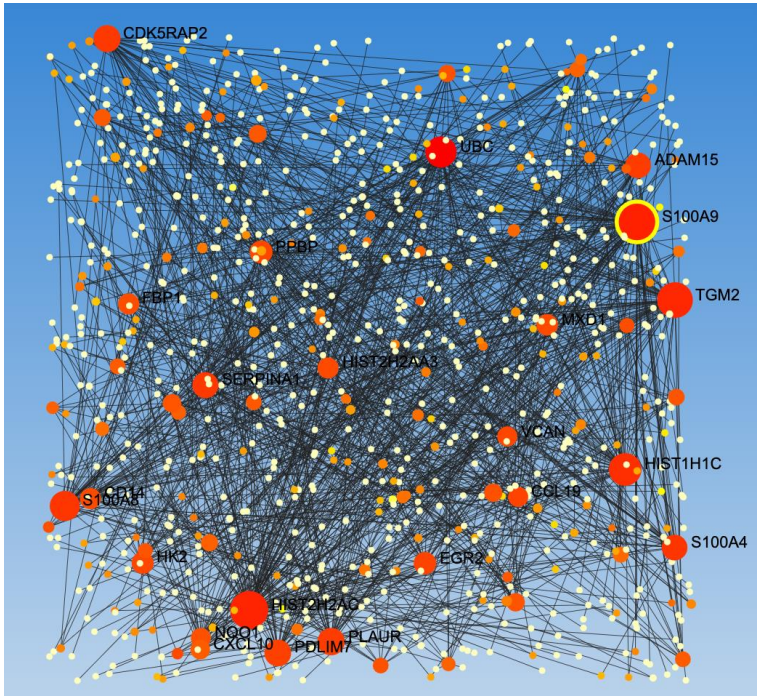
Pirhaji, et al 2017. Nature Methods **13**, pages770–776

Network Layout

Mix art with science

A layout algorithm takes as input a set of nodes and edges along with a set of options and return their locations

- Heuristic
- Subjective
- Often used in combination with
  - Color
  - Size
  - …

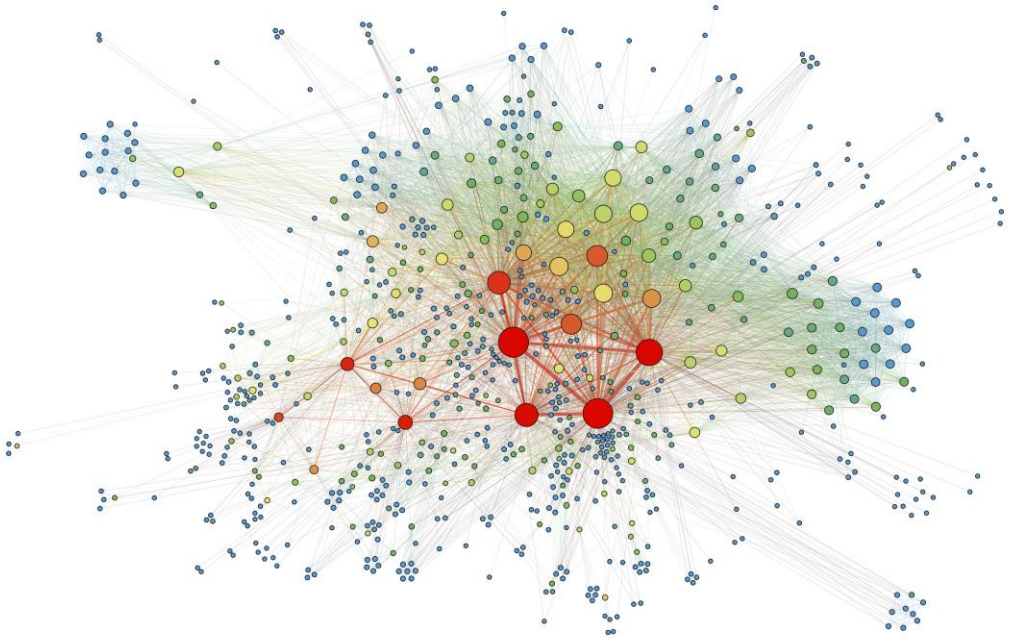


Random Layout

Network layout: force-directed layout

- The connections between nodes are treated like metal springs attached to the pair of nodes.
- These springs repel or attract their end points according to a force function.

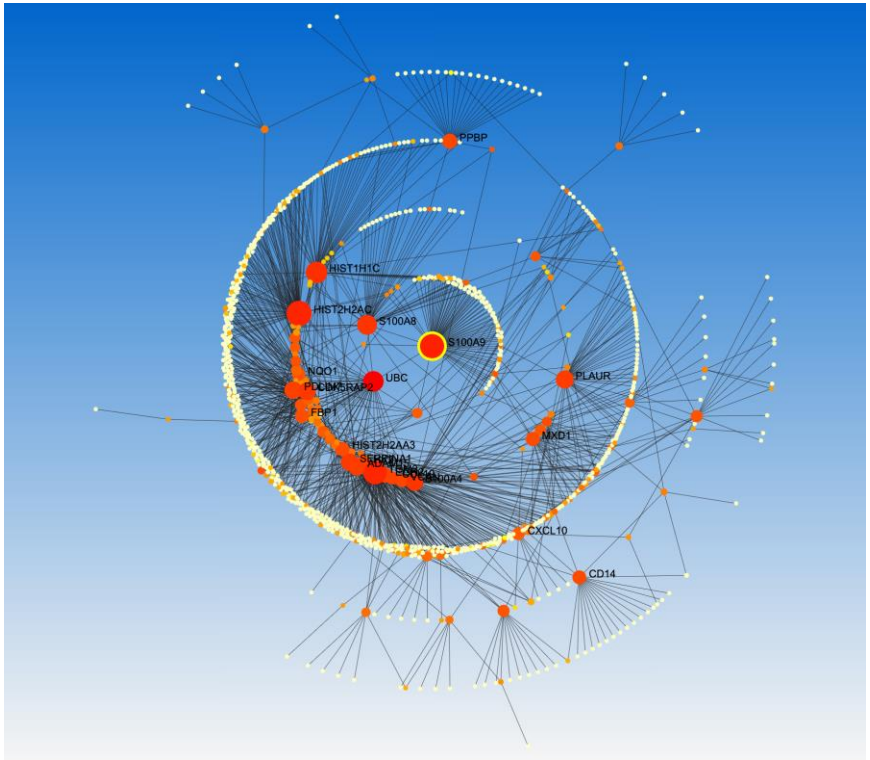
Highlights the underlying topology of the graph. We can visually identify clusters, cliques, and bridges.



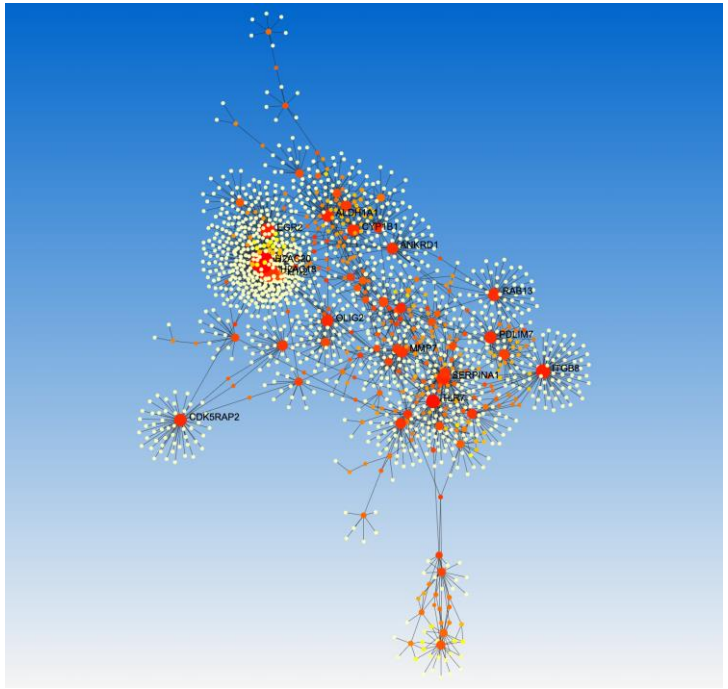
Network layout: concentric circle

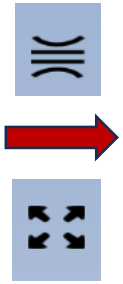
- Organize the nodes into concentric circles, based on the specified metric.
- The nodes with the highest metric values are placed in the innermost circle, and the metric values of the nodes descend for each outward circle.
- Each circle has nodes with metric values between a specified range, with nodes within a circle sorted accordingly.

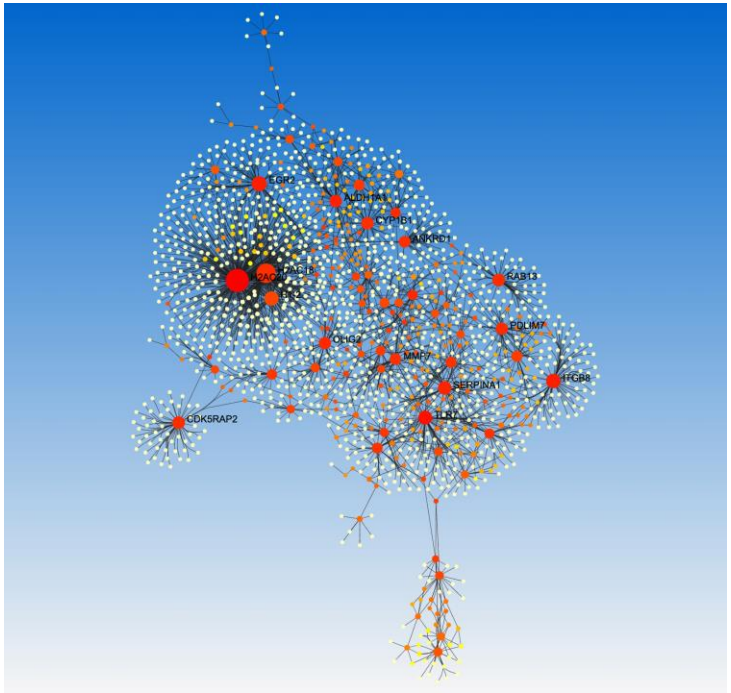
Useful for highlighting relative importance of the nodes



Tips: Edge Bundling & Reduce Overlap

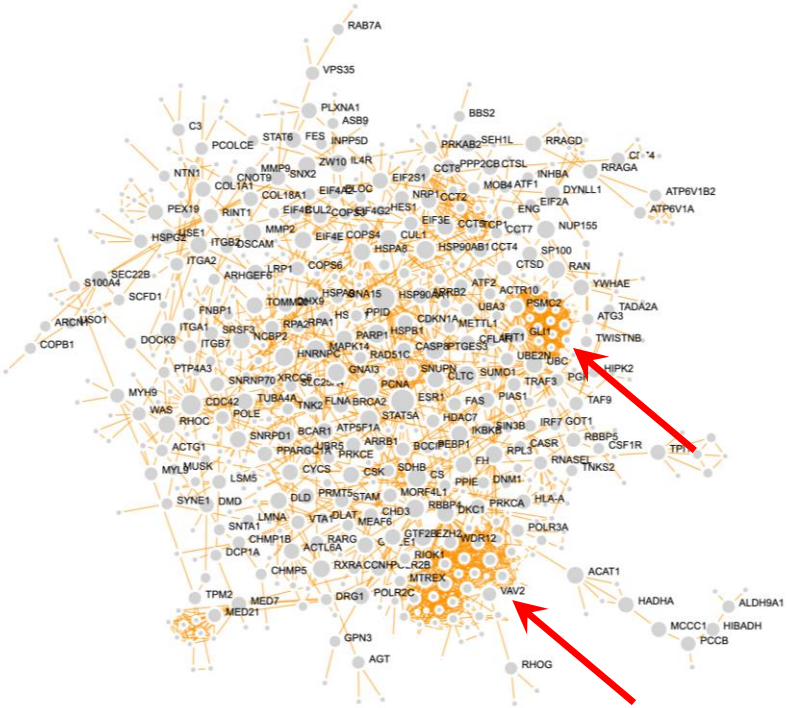






Network modules (communities)

- Community: having dense connections between the nodes within modules but sparse connections between nodes in different modules
- Could have biological or functional implications
- Structural module ≈ Functional module



Common module detection methods

Label propagation

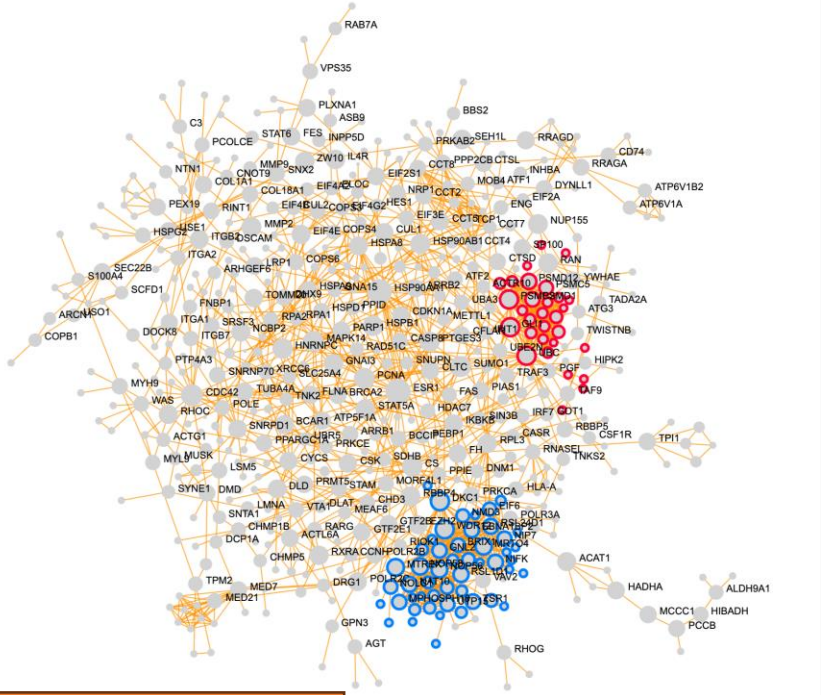
- Each node in the network is assigned to the same community as the majority of its neighbours (through iterative process)

InfoMap

- Based on information flow through a network using random walk

WalkTrap

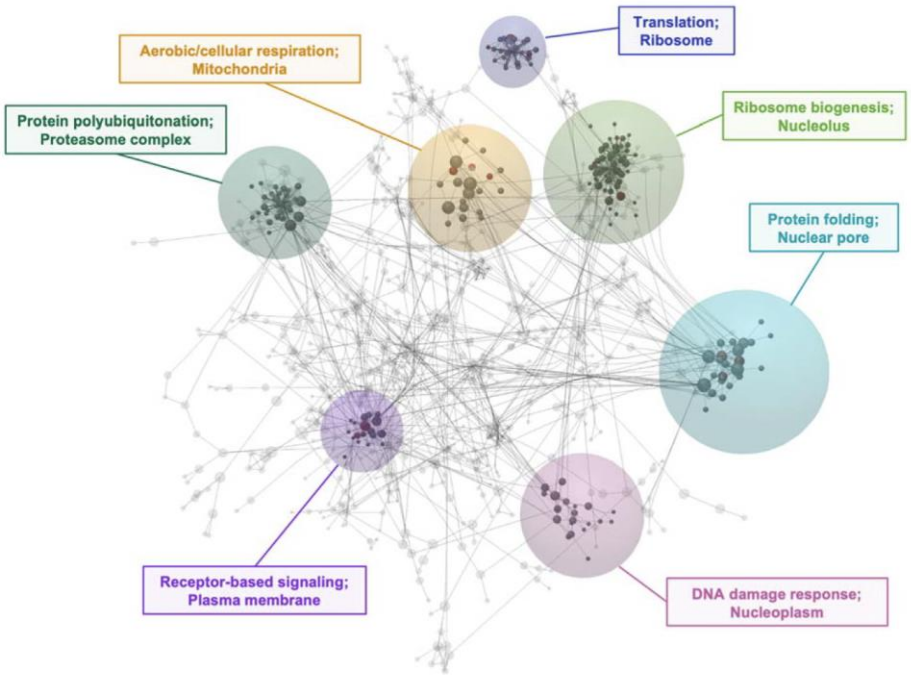
- Short distance random walks tend to stay in the same community



All Heuristic Methods

Understand & view modules

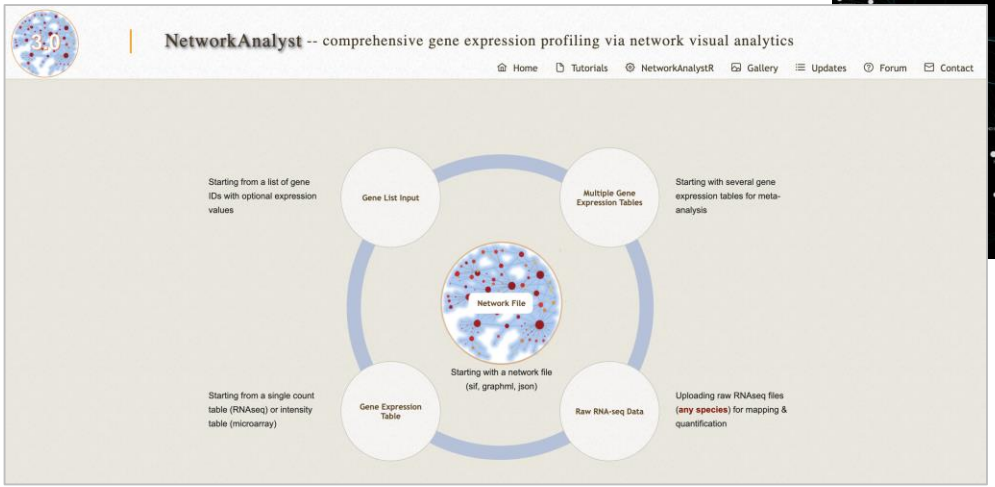
- We can use Wilcoxon rank-sum test to compare the number of connections of each node with that of other nodes within the module and with that for nodes outside the module.
- We can test whether certain biological functions are enriched in these modules.
- We can also extract the modules for more focused visualization



Cytoscape vs. NetworkAnalyst

Cytoscape:

- Standalone platform for network visualization
- Need to install plug-ins for other analysis



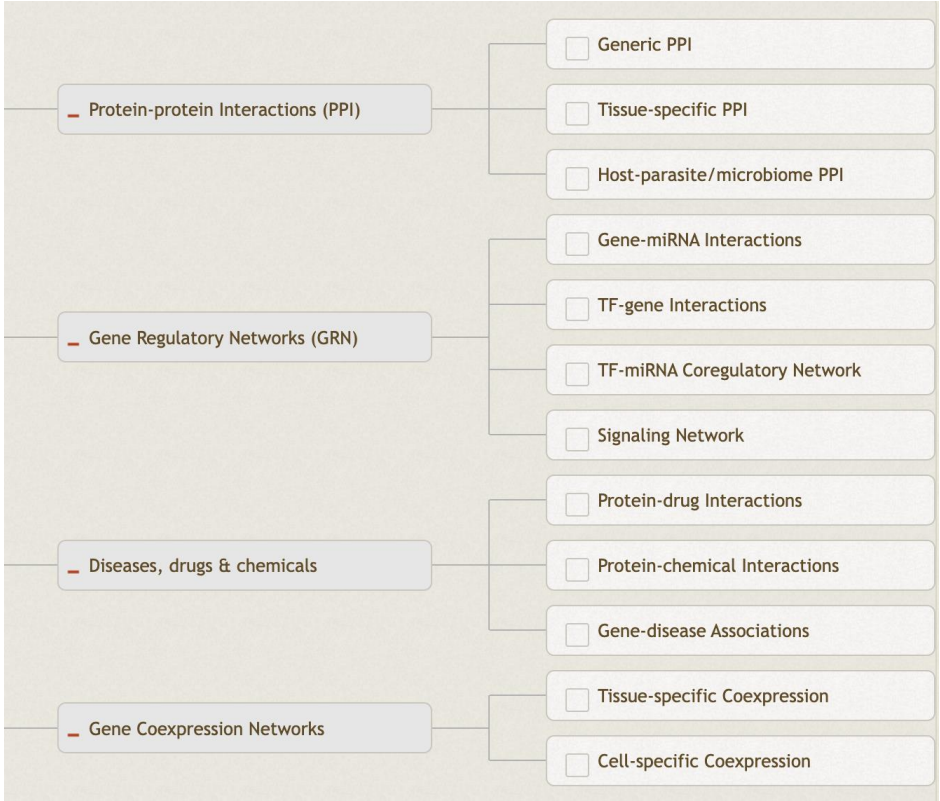
NetworkAnalyst:

- Web-based platform for network visualization
- Built-in databases for network creation & functional analysis

NetworkAnalyst – network creation

Built-in databases:

- PPI network
- miRNA-gene
- TF-gene
- Signaling network
- Protein-drugs
- …..



NetworkAnalyst – Network Refinement

Comprehensive support for automatic filtering (topology and biology), as well as manual network filtering

Mapping Overview

The genes of interest (seeds) from previous analysis are mapped to the corresponding molecular interaction database. The procedure typically produces one big subnetwork ("continent") with several smaller ones ("islands"). Please note:

- Subnetworks with **at least 3 nodes** are listed below. You can visually explore them in the next step.
- The common **network topology measures** are also computed based on well-established the igraph R package
- These subnetworks can be downloaded as **SIF (simple interaction format)** files to be explored in other tools (i.e. Cytoscape).

Networks	Nodes	Edges	Seeds	Topology	Interactions
subnetwork1	916	1118	85	View	SIF
subnetwork2	3	2	1	View	SIF
subnetwork3	3	2	1	View	SIF

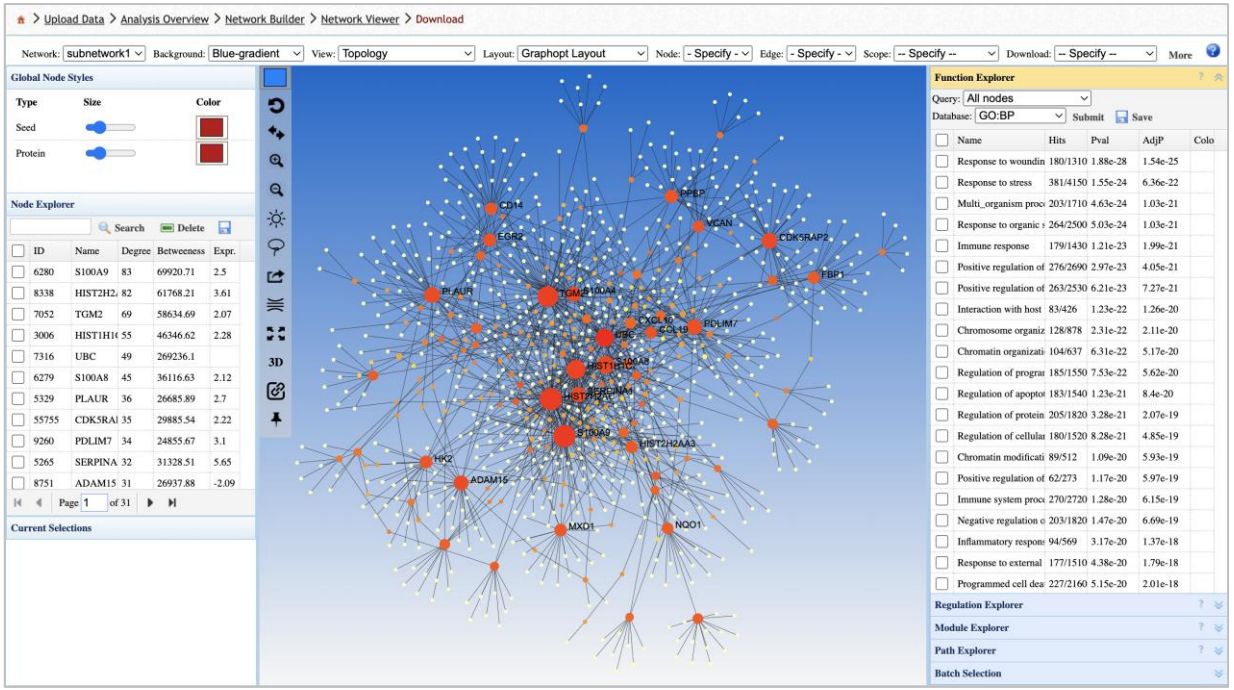
Network Tools:

- First Order Network
- Zero-order Network
- Second-order Network
- Minimum Network
- Steiner Forest Network
- Degree Filter
- Betweenness Filter
- Tissue Filter
- Batch Exclusion

NetworkAnalyst – Network Visualization

Build-in network visual analytics system

- ❖ Multiple algorithms for layout & module detection
- ❖ Enrichment analysis
- ❖ 2D and 3D support!



Schedule for today

Time	Topics
9:00 – 9:15	General introduction & recap
9:15 – 9:40	Network analysis and visualization
9:45 – 10.15	NetworkAnalyst live demo & hands on
10:20 – 10:50	Gene regulatory networks
10:55 – 11:25	miRNet live demo & hands on
Summary and discussion	

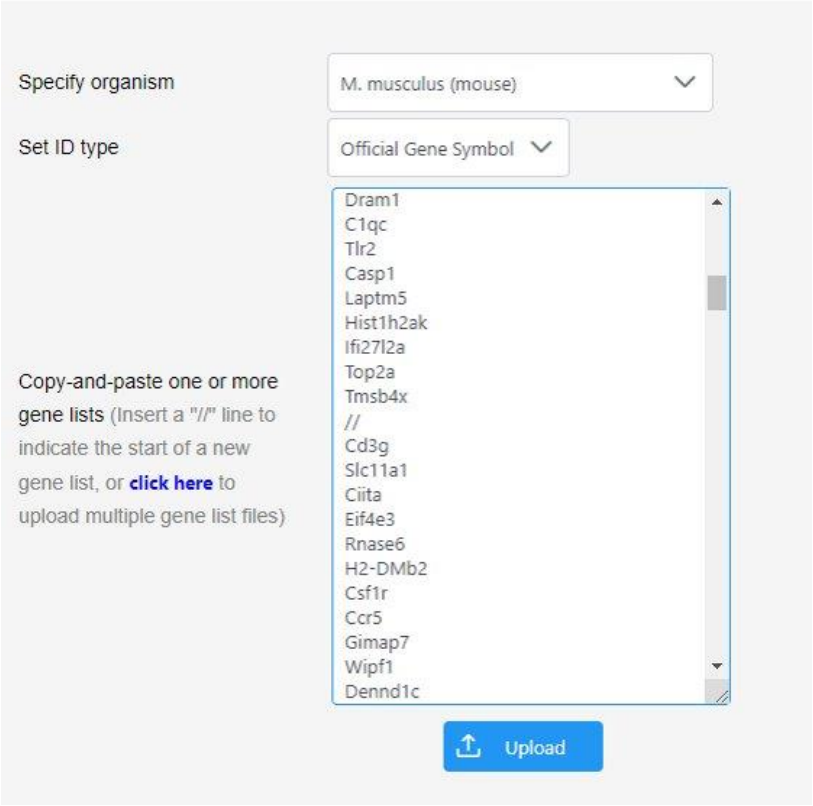
Live Demo & Hands On

Tasks

1. Multi-list upload and analysis.
2. Targeted functional analysis - extract module/subnetworks
3. Save a network state

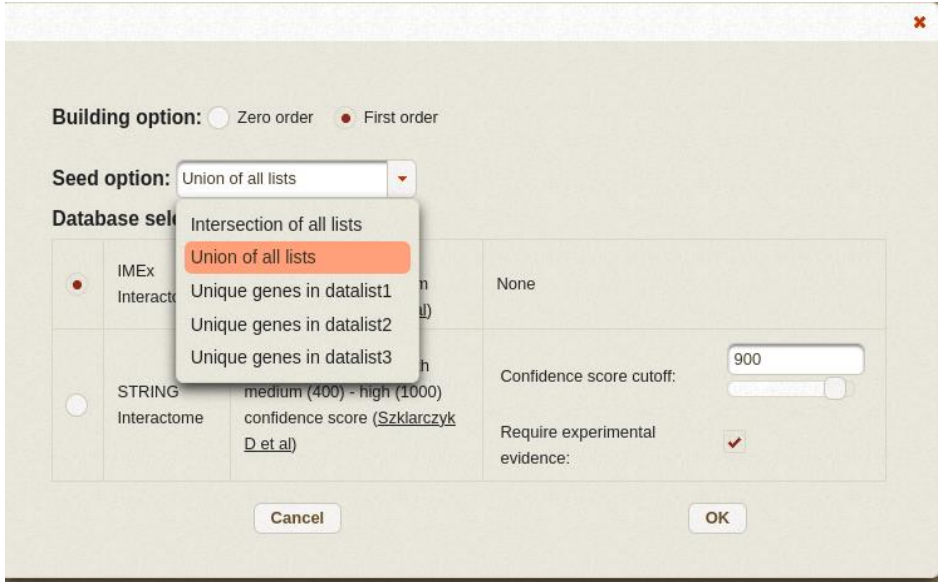
Task 1: Multi-list

- For multiple lists, insert **"//"** in a line to indicate the start of new list.



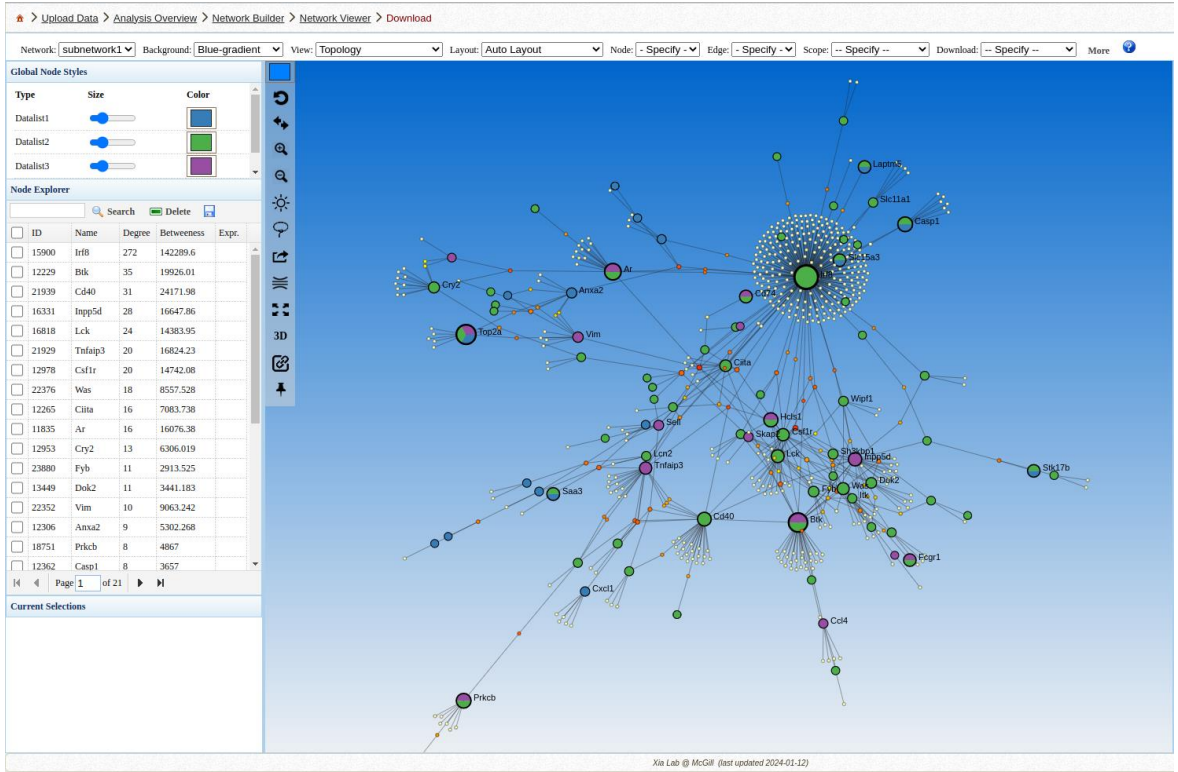
Task 1: Multi-list

- Allow you to explore relationships between different lists:



Task 1: Multi-list

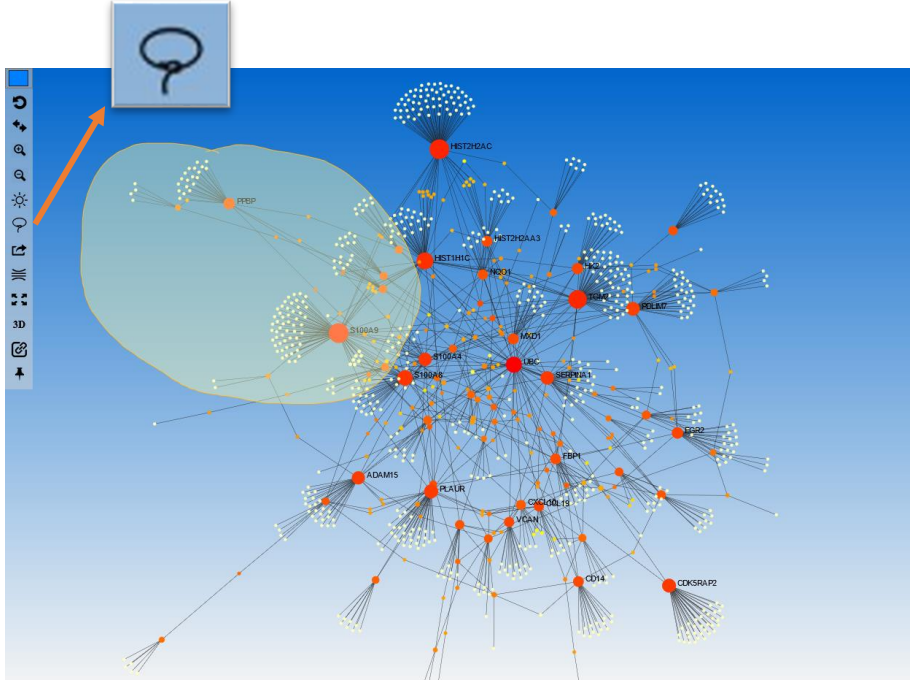
- Visualize intersections in network view --> **Pie nodes**
- Color legend on top right corner
- Genes coming from any of the list(s) are larger in size



Task 2: Targeted analysis

- Focus on a region of the network that you are interested in.

- Select nodes based on:
  - Functional category/pathway
  - Predefined list of genes
  - Community detection
  - Manual selection

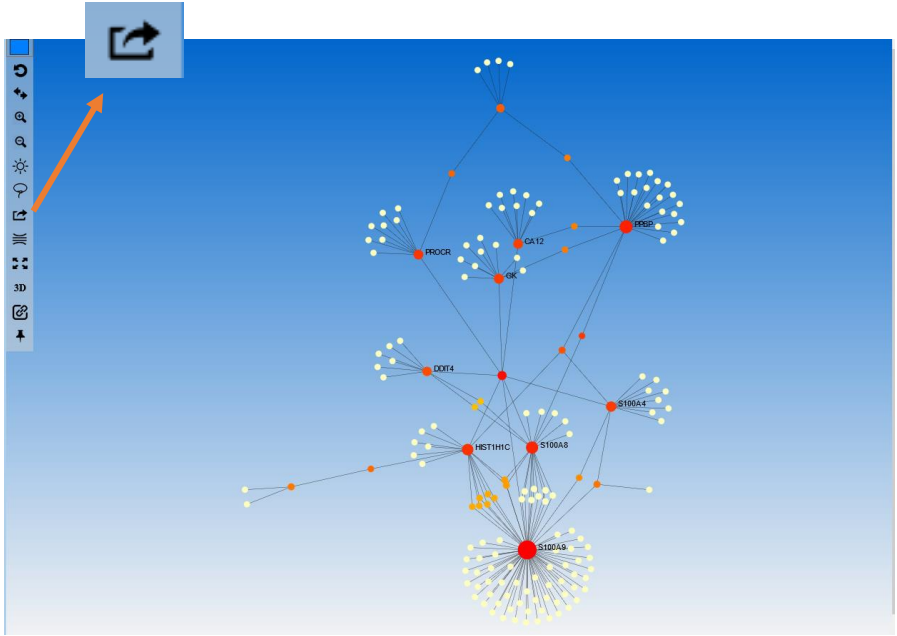


Example of drag-and-select manual selection

Task 2: Targeted analysis

Analyze the nodes of interest:

- Enrichment analysis on **highlighted nodes** in the current network
- **Extract** the nodes as a connected subnetwork
  - Use "**Minimum Network**" approach --> pairwise shortest paths and remove those that are not traversed

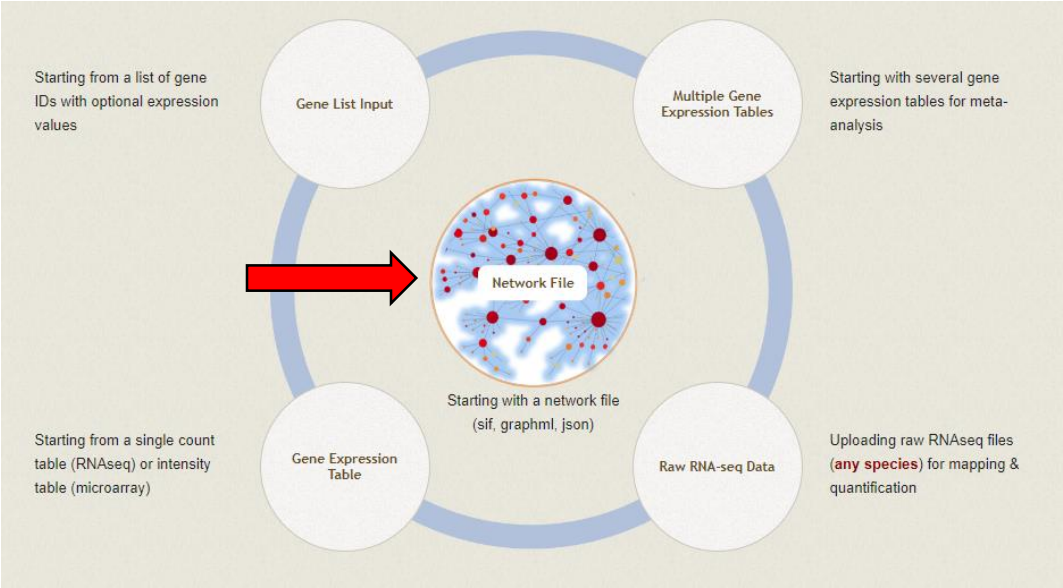


Extract a subnetwork

Task 3: Save and load subnetwork

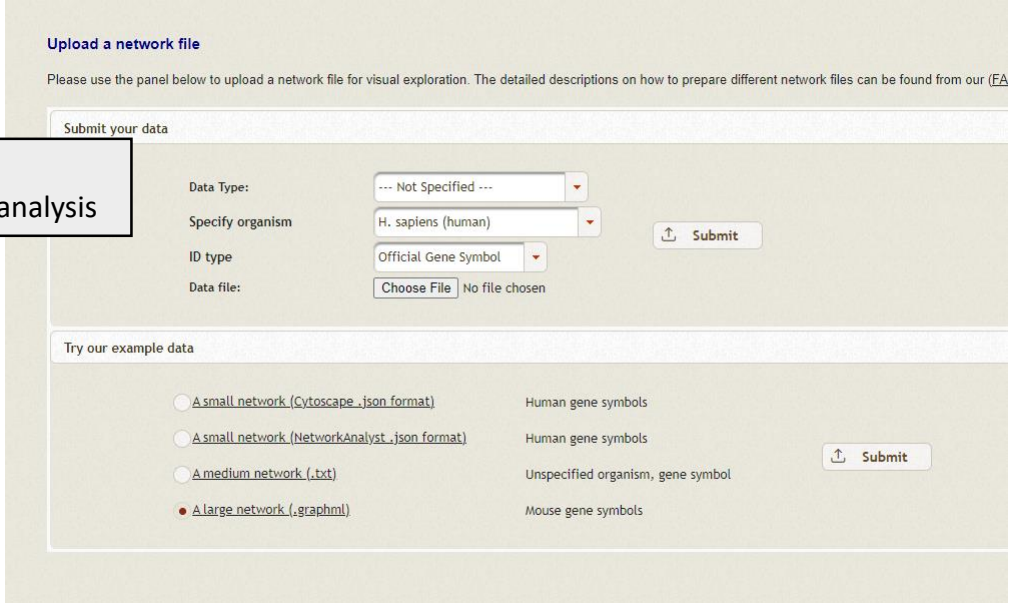
- You can save your network as a .json file that you can reload later.

- NetworkAnalyst can take **graph file** as input.



Task 3: Save and load subnetwork

ID mapping (optional)
- allows functional enrichment analysis



Schedule for today

Time	Topics
9:00 – 9:10	General introduction & recap
9:15 – 9:40	Network analysis and visualization
9:45 – 10.15	NetworkAnalyst live demo & hands on
10:20 – 10:50	Gene regulatory networks
10:55 – 11:25	miRNet live demo & hands on
Summary and discussion	

Biological Networks

- **Protein-protein interaction network:** proteins that are connected in physical interactions
  - Proteins and their interaction partners
- **Gene regulatory network**: a collection of molecular regulators that interact with each other and with other substances in the cell to govern the gene expression levels of mRNA and protein
  - miRNA, transcription factors, non-coding genes, etc
- **Metabolic network:** metabolic products and substrates that participate in biochemical reactions
  - Metabolites, enzymes, co-factors

Protein-protein interaction (PPI) network

Experimental

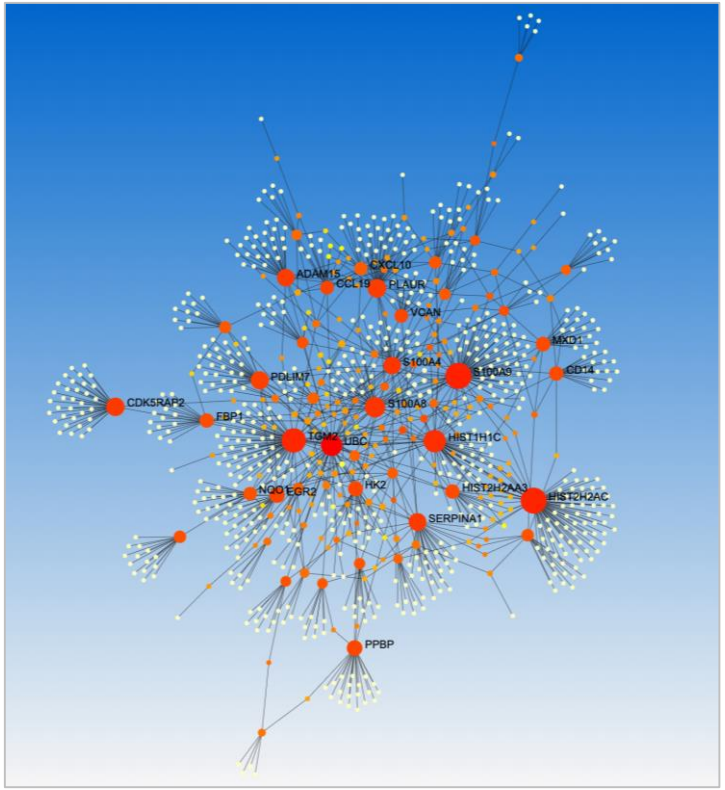
- InnataDB/IMEx Interactome
- Rolland Interactome (2014)
- HuRi (human reference interactome, 2020): Pairwise combinations of human protein-coding genes are tested systematically using high throughput yeast two-hybrid screens to detect protein-protein interactions.

Predicted

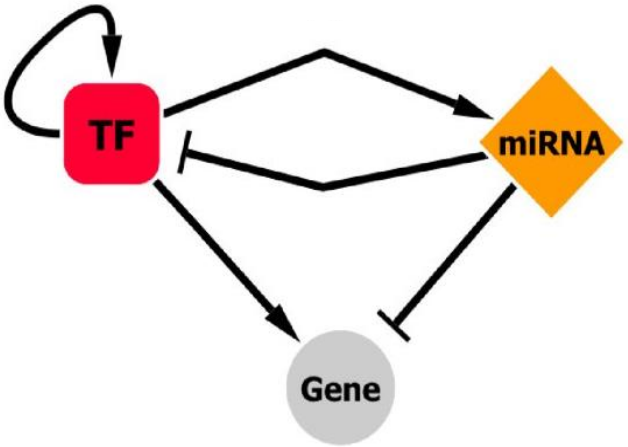
- Binding motifs, evolution, etc

Both:

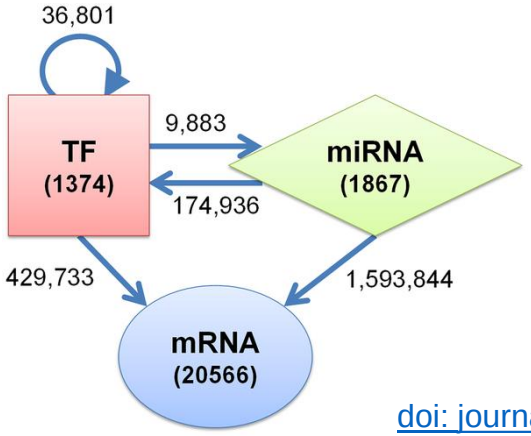
- STRING:
  - Experimentally derived interactions through literature curation;
  - Computationally predicted interactions through text mining and inference from other organisms
  - Supporting > 10,000 organisms



Core Gene Regulatory Networks



GRN in human (2015)



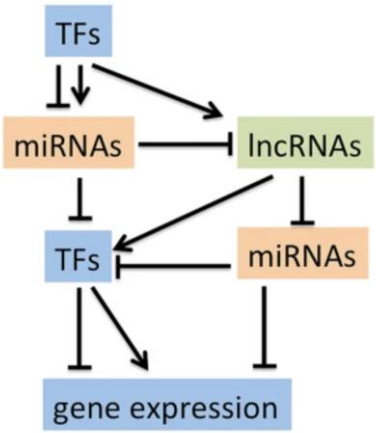
doi: journal.pcbi.1004504

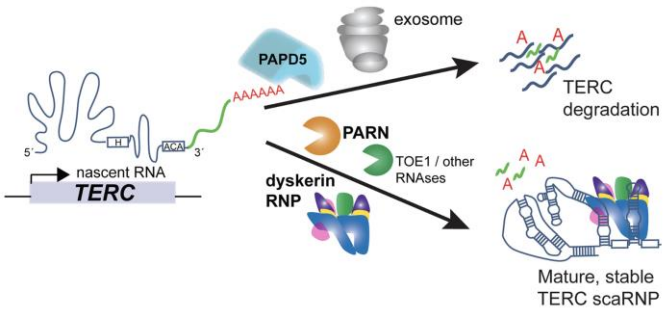
TF: a protein that controls the rate of transcription from DNA to mRNA, by binding to a specific DNA sequence.

miRNA: small, single-stranded, non-coding RNA molecules (21-23 nt), involved in post-transcriptional regulation of gene expression

Long noncoding RNAs (lncRNAs)

- Transcripts that are over 500 nucleotides and don't encode proteins
- Over 20,000 lncRNAs have reported, most of which have no known functions
- Known functions
  - Gene transcriptional regulation
    - Interact with miRNAs
    - Interact with TFs
  - Function as scaffolds
    - Assembles telomerase





How to obtain GRN?

Infer network from gene expression data

- Data driven, suitable for non-model species without good database
  - Weight Gene Correlation Network Analysis (WGCNA)
  - GEne Network Inference with Ensemble of trees (GENIE3)
- **Issues:**

  require a lot of data, computing intensive, high false positives

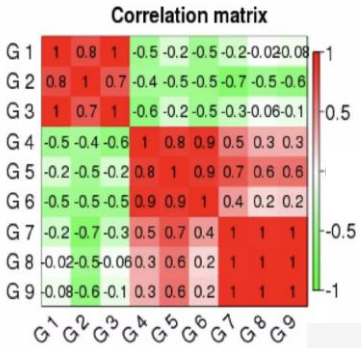
Search from the databases

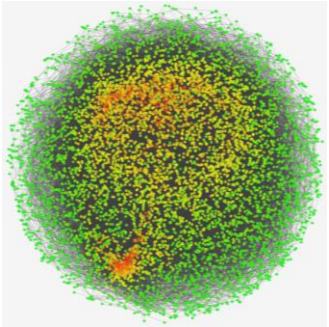
- Suitable for model species with high-quality knowledge & coverage
  - miRNA-target gene database
  - TF-target databases
- **Issues:**

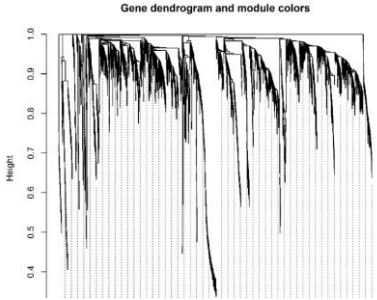
  tend to be generic, need to add more context specific information

WGCNA

1. Compute the gene **co-expression matrix**
2. Create a **weighted co-expression network** by selecting the threshold for constructing the network to reflect <u>small world, power law distribution</u>
3. Compute topological overlap measure (TOM) adjacency matrix based on overlap of network neighbors for **module detections.**







DREAM Challenges

DREAM (**D**ialogue for **R**everse **E**ngineering **A**ssessment and **M**ethods)

DREAM5: network inference challenge on genome-scale transcriptional regulatory network from gene expression data

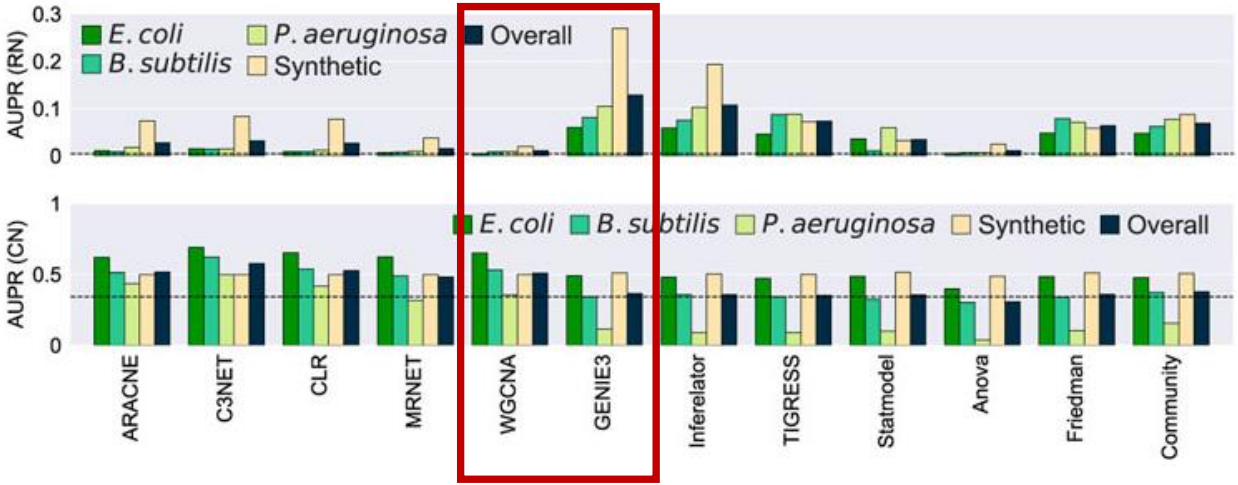
- Community driven;
- Synthetic data (ground truth known);
- Time series experimental data

Method	Network type	Directed network	Main References
ARACNE	COEX	FALSE	Margolin et al. (2006)
C3NET	COEX	FALSE	Altay and Emmert-Streib (2010)
CLR	COEX	FALSE	Faith et al. (2007)
MRNET	COEX	FALSE	Meyer et al. (2007)
LSTrAP	COEX	FALSE	Proost et al. (2017)
RNA-seqNet	COEX	FALSE	Proost et al. (2017)
WGCNA	COEX	FALSE	Zhang and Horvath (2005)
GENIE3	CAUS	TRUE	Huynh-Thu et al. (2010)
INFERELATOR	CAUS	TRUE	Bonneau et al. (2006)
TIGRESS	CAUS	TRUE	Haury et al. (2012)
StatModel	CAUS	TRUE	Zorro-Aranda et al. (2022)
iBMA	CAUS	TRUE	Annest et al. (2009)
ScanBMA	CAUS	TRUE	Young et al. (2014)
ANOVA	HYBR	TRUE	Zorro-Aranda et al. (2022)
FRIEDMAN	HYBR	TRUE	Zorro-Aranda et al. (2022)

https://doi.org/10.3389/fgene.2023.1143382

DREAM5 result

- ❖ GENIE3 is the best tool for bulk RNA-seq network inference
- ❖ WGCNA is the most widely used tool for co-expression network inference



TF - target gene databases

- **ENCODE**
  - ✓ Transcription factor and gene target data derived from the ENCODE ChIP-seq data
  - ✓ Integrates ChIP-seq of TFs or chromatin regulators with differential gene expression data to infer direct target genes
  - ✓ 379 TFs, ~400k interactions total
- **JASPAR**
  - ✓ Manually curated high-quality and non-redundant DNA-binding profiles for transcription factors (TFs) across taxa.
  - ✓ https://jaspar.elixir.no/
  - ✓ 111 TFs, ~150k interactions total
- **CHEA**
  - ✓ TF–gene co-expression from RNA-seq studies, TF–target associations from ChIP-seq experiments, and TF–gene co-occurrence
  - ✓ https://maayanlab.cloud/chea3/
  - ✓ 199 TFs, ~386k interactions

TRRUST TF-target gene database

- Transcriptional Regulatory Relationships Unraveled by Sentence-based Text mining
- Text mining, followed by manual curation

Table 1.
Summary of databases of transcription factor (TF)–target interactions

	Human					Mouse				
	TF	TG	Interaction	MoR	PMID	TF	TG	Interaction	MoR	PMID
TRRUST v2	800	2521	8444	4866	6562	828	2187	6552	4106	5069
TRRUST v1	748	2375	8015	4861	6175					
PAZAR	173	2747	4918		489	162	1726	2329		328
TFactS	275	1873	4209	1468	776	238	1080	2149	1176	472
TRED	119	1621	3275		1542	75	236	508		745
TFe	76	623	1058	134	510	147	388	673	90	353

TF, the number of TF; TG, the number of target genes; Interaction, the number of TF–target interactions;

https://doi.org/10.1093/nar/gkx1013

miRBase

miRbase: the microRNA database
the archive for microRNA sequences
and annotations

https://mirbase.org/

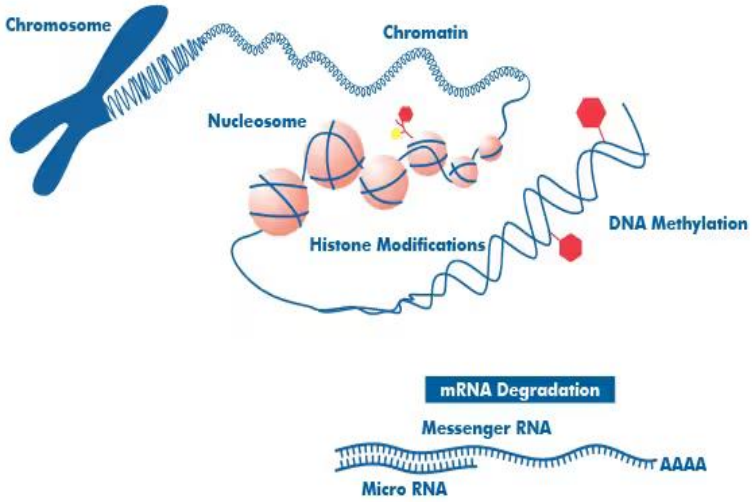
miRNA - target gene databases

Database	Number of miRs	Number of target genes	Number of MTIs	Number of citations (WoS)	Experimental validation methods
MiRTarBase (10, 16–20)	2599	15 064	380 639	3217	CLIP-Seq, Luciferase assay, microarray, NGS, pSILAC, western blot
starBase/ENCORI (11, 21, 38)	155	687	1286	2605	CLIP-Seq
DIANA-TarBase (3, 22, 23–25)	1084	20 820	422 662	1894	AGO-IP, biotin miRNA tagging, CLASH, CLEAR-CLIP, CLIP-Seq, ELISA, IMPACT-seq, microarrays, PAR-CLIP, proteomics, qPCR, Reporter genes, RIP-seq, RNA-seq, RPF-seq, TRAP, western blot, other
miRWalk (26–28)	2656	19 128	31 235 408	1737	CLIP-Seq, Luciferase assay, microarray, NGS, pSILAC, western blot
miRecords (29)	303	1112	1748	986	ELISA, immunocytochemistry, northern blot, qRT-PCR, reporter assay, western blot, other
miRGator (30–32)	N/A	N/A	N/A	266	qRT-PCR, reporter assay, western blot

https://doi.org/10.1093/database/baad014

miRNA-Epigenetics modifier

- Regulation between epigenetics and miRNA:
  - Methylation of CpG island --> reduced transcription of corresponding miRNAs
  - Histone modifications --> up or down regulation of miRNAs expression
  - miRNA --> methylation and histone modifications by targeting genes encoding enzymes
- EpimiR --> database of experimentally validated regulations collected from 2000 papers
  - 19 types of epigenetic modifications
  - 617 miRNAs across 7 species
  - 1974 regulatory relationships



http://www.jianglab.cn/EpimiR/index.jsp

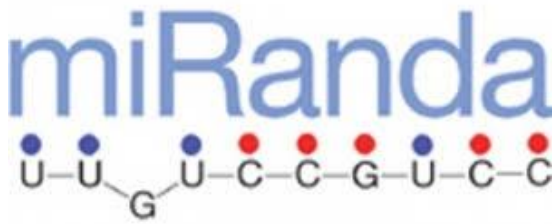
Xeno-miRNA regulation

- Exogeneous miRNAs have been detected to interact with host – Cross-species interactions
- miRNet contains both experimentally detected ~~and computational predicted~~ interactions

Hosts	Tissue/ sources	Xeno-species	Xeno-miRNAs (detected/ predicted)	Potential targets
Human	18	40	296/625	20,791
Mouse	18	27	83/418	19,430
Pig	4	14	20/116	12,537
Chicken	2	15	23/10	16,459
Fruit fly	6	6	16/44	12,445
C. elegans	8	6	15/36	16,391
Total	49	54 (unique)	1702	98,053

miRNA target prediction

- Algorithms to predict interactions between miRNAs and target mRNA sequences
- i.e. miRanda
  - **Complementarity** of nucleotides, especially at seed region
  - **Stability** of interactions by measuring free energy of miRNA-mRNA
  - **Conservation** of target site across species (More conserved --> more likely to be functionally relevant)




```
Query:    3' cgucuugaAUCGGUGACACUu 5'
                        |:|  |||||||
Ref:      5' cactcttcTGGAAACTGTGAg 3'

Energy:  -14.890000 kCal/Mol
```

miRNet & miRNA-centric knowledgebase

miRBase

TAM 2.0

starBase For RNA Interactome

EpimiR

miRanda

PhenomiR

SNP2TFBS

SM2miR

TransmiR

ADmiRE

PolymiRTS

miR2Disease

miRNet 2.0

HMDD

MOTA

Pharmacomir

miRTarBase

TarBase

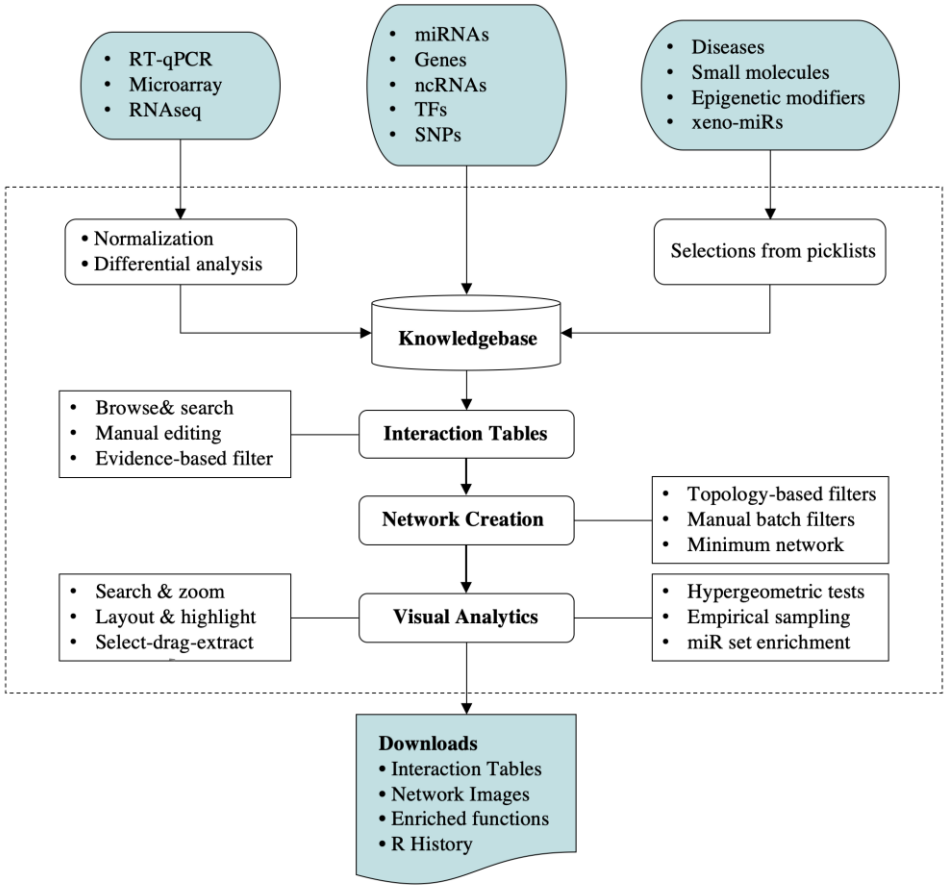
miRecords

Comparison with other tools

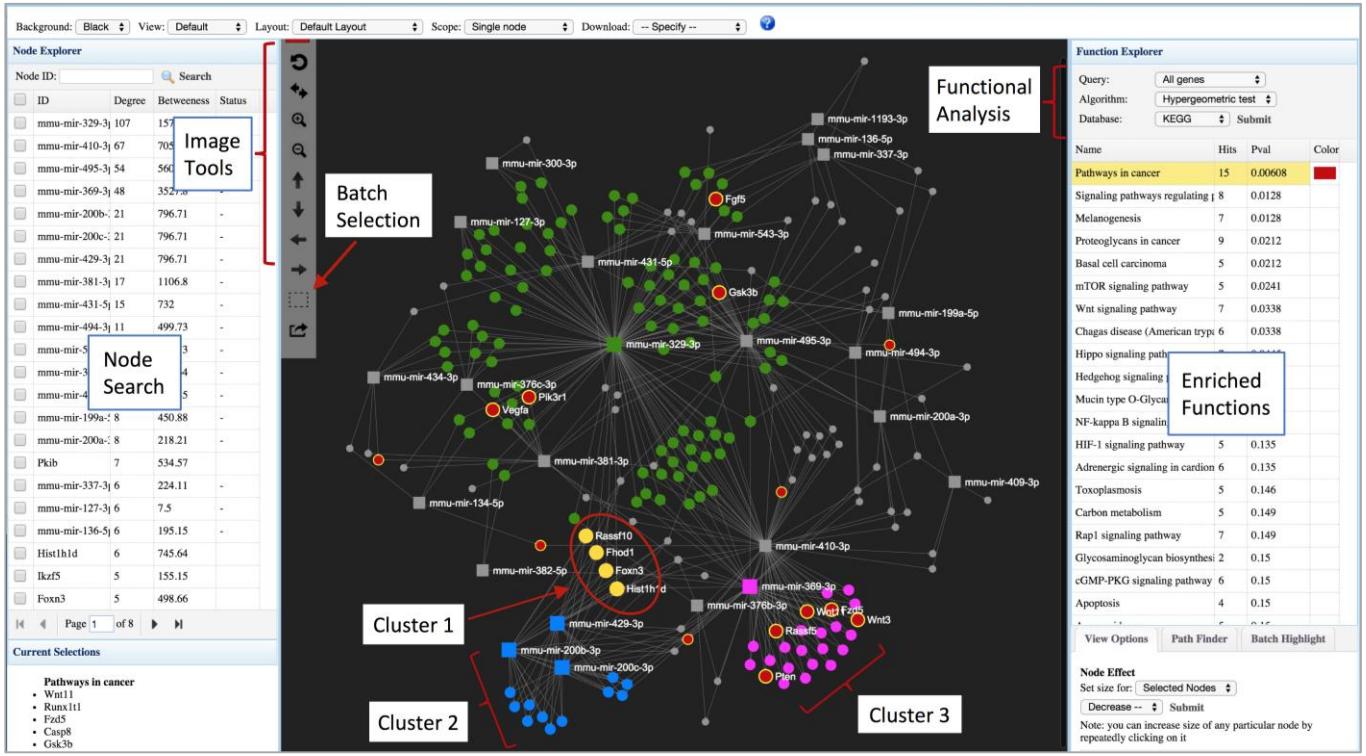
Tool name	miRNet 2.0	miRNet 1.0	miRTargetLink	MIENTURNET	Arena-Idb	starBase
Data input and processing						
Species #	11	8	1	6	1	23
Target genes						
Experimental	+++	++	++	+	++	+++
Predicted	√	√	√	√	√	√
Other targets & associations						
miR-SNP	√	-	-	-	-	-
TF	√	-	-	-	-	-
ncRNA	+++	+	-	-	+	++++
xeno-miRNA	√	-	-	-	-	-
Disease	+++	++	-	-	++	++
Epigenetic modifier	√	√	-	-	-	-
Small molecule	√	√	-	-	-	-
Expression profiling	√	√	-	-	-	-
Enrichment analysis						
Hypergeometric tests	√	√	√	√	-	√
Empirical sampling	√	√	-	-	-	-
miR-set enrichment	√	-	-	-	-	-
Network visual analytics						
Multiple query types	√	-	-	-	-	-
Integration with PPI network	√	-	-	-	-	-
Multipartite network visualization	√	-	-	-	-	-
Subnetwork extraction	√	-	-	-	-	-

- miRTargetLink: https://ccb-web.cs.uni-saarland.de/mirtargetlink/
- MIENTURNET: http://userver.bio.uniroma1.it/apps/mienturnet/
- Arena-Idb: http://ncrnadb.scienze.univr.it/sites/arenaidb/
- starBase: http://starbase.sysu.edu.cn/index.php

miRNet 2.0 Workflow



Comprehensive miRNA functional analysis

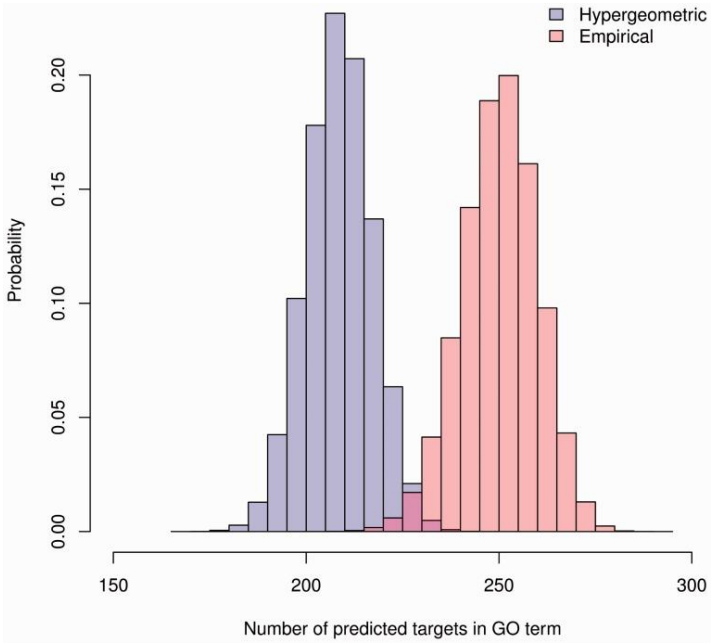


How to understand the functions of miRNAs?

Standard approach: apply hypergeometric test to identify enriched functions based on the target genes.

Potential issue: hypergeometric distribution describes the situation where samples are picked <u>uniformly at random</u> from a finite population which contains a labelled subset.

But miRNA target gene set is not random



https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4426843/

Using empirical sampling to address bias

The method is used to estimate the null distribution of the target genes as selected based on the input miRNAs. The procedures can be divided into the following steps:

1. A list of miRNAs of the same size are randomly selected from all the miRNAs with known targets in the database;
2. The functional annotations (i.e. KEGG) are then performed for the list;
3. The process is repeated 1000 times (default);
4. Compare the KEGG pathway hits. The empirical p values are calculated as the proportion of overlaps from 1000 random process that equal or larger than the original. A p value < 0.001 will be reported if no results from random process are better than the original.

Function Explorer

Query: All genes
Algorithm: Emprical sampling
Database: KEGG
Submit

Name	Hits	Pval	adj.Pval
Pentose and glucuronate interconvers	3	0.0595	0.401
Carbohydrate digestion and absorptic	5	0.0612	0.401
Rheumatoid arthritis	4	0.0748	0.401
Type I diabetes mellitus	1	0.0775	0.401
Drug metabolism - cytochrome P450	3	0.089	0.401
Steroid hormone biosynthesis	4	0.092	0.401
Autoimmune thyroid disease	1	0.0992	0.401
Drug metabolism - other enzymes	4	0.106	0.401
Taste transduction	3	0.125	0.401
Folate biosynthesis	1	0.128	0.401
Metabolism of xenobiotics by cytoch	4	0.128	0.401
Pancreatic cancer	19	0.145	0.401
Phototransduction	3	0.146	0.401

Schedule for today

Time	Topics
9:00 – 9:10	General introduction & recap
9:15 – 9:40	Network analysis and visualization
9:45 – 10.15	NetworkAnalyst live demo & hands on
10:20 – 10:50	Gene regulatory networks
10:55 – 11:25	miRNet live demo & hands on
Summary and discussion	

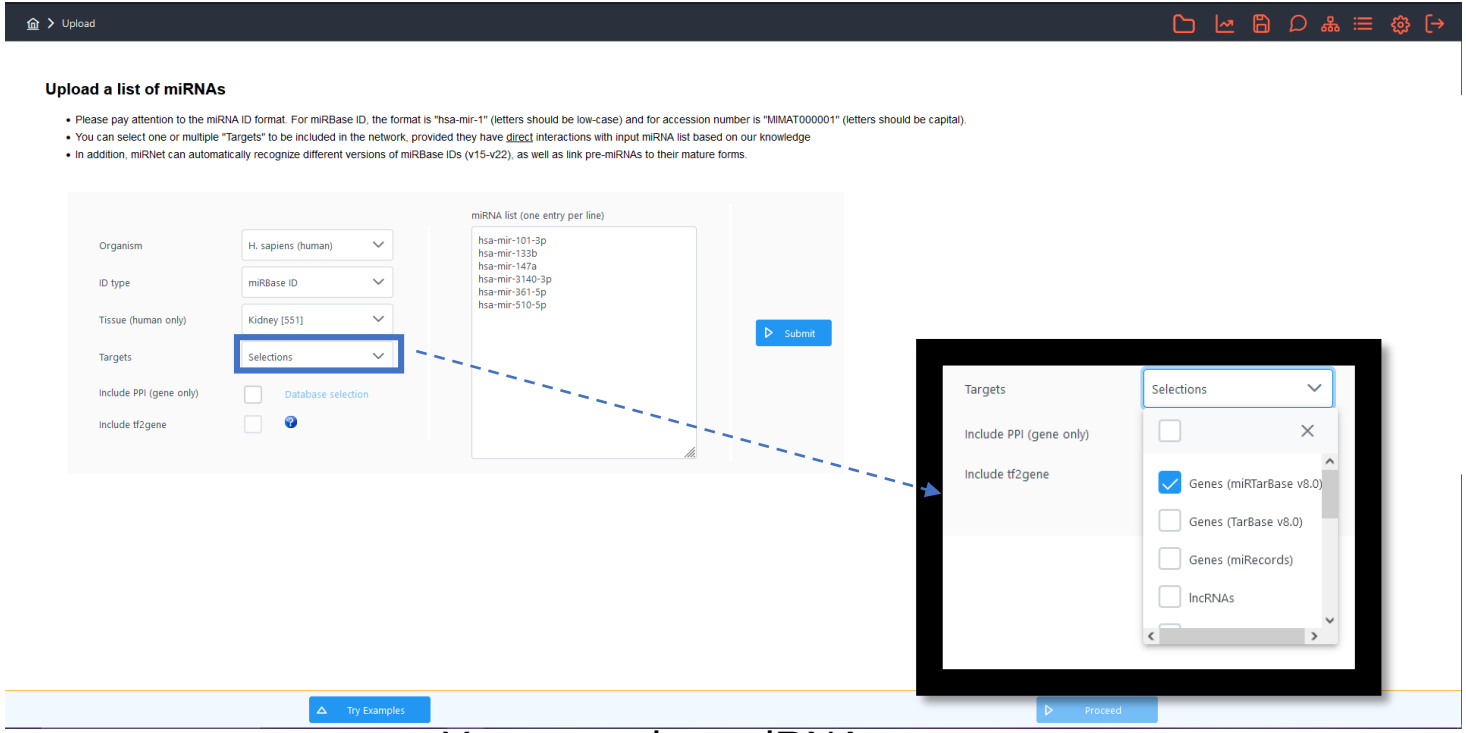
Live Demo & Hands On

Tasks

- Understand the network building process in miRNet
  - Single input upload
  - Multi-input upload

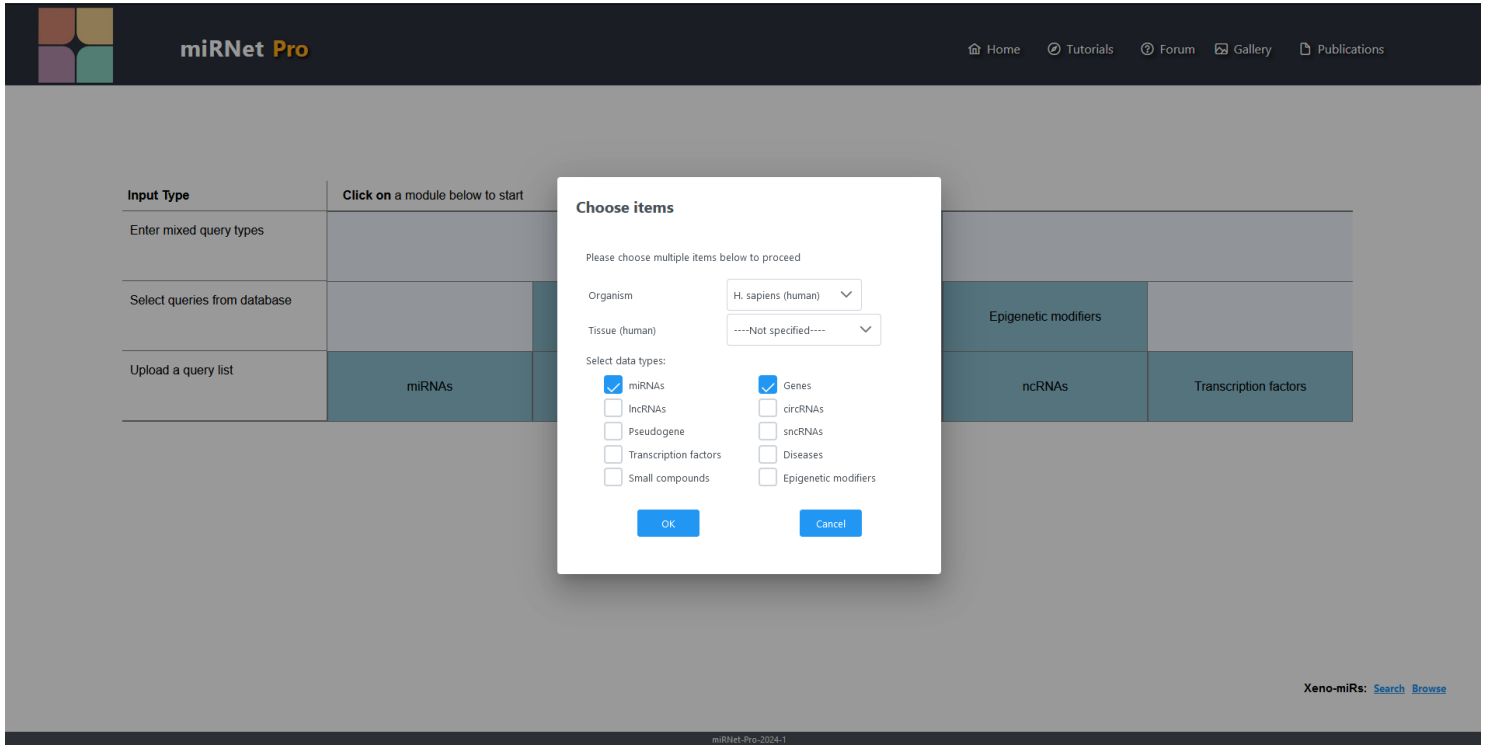
- Explore different layouts for network visualization

Single input



You can select miRNA targets

Multi-input upload



Multi-input upload

Input tabs

Upload progress

Upload

Upload multiple lists

Please use the tabs below to prepare your input for: **H. sapiens (human)**. Make sure to click **Submit** button before switching to the next tab.

miRNAs Genes

ID type: miRBase ID

Target type: Genes (miRTarBase v8.0)

miRNA list (one entry per line):

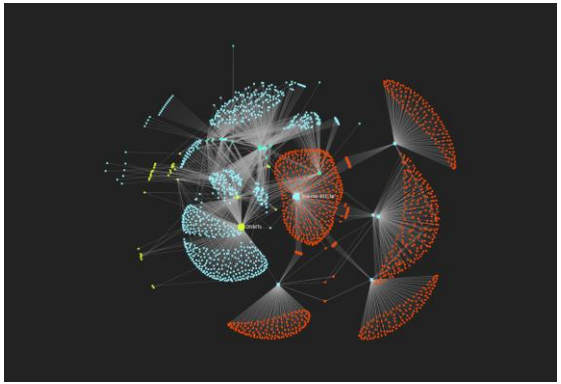
hsa-mir-101-3p
hsa-mir-133b
hsa-mir-147a
hsa-mir-3140-3p
hsa-mir-361-5p
hsa-mir-510-5p

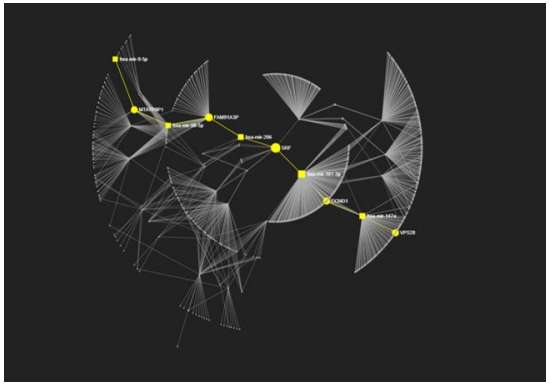
Submit

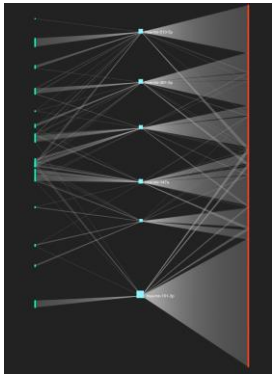
Try example

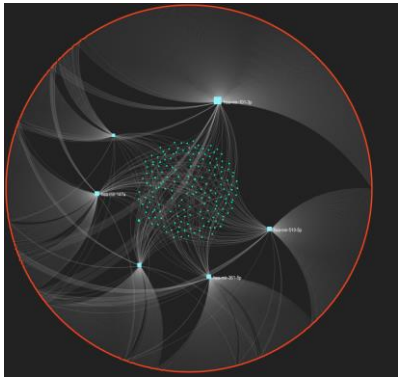
Data uploaded: mirna gene **Include PPI:** Database selection Proceed

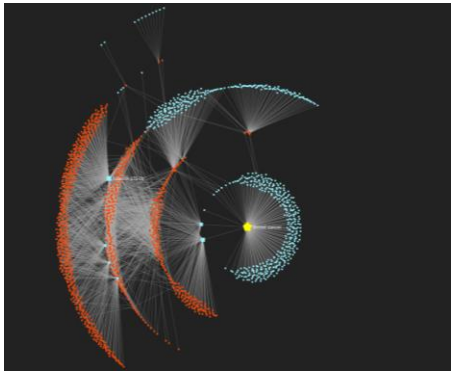
Network Layout









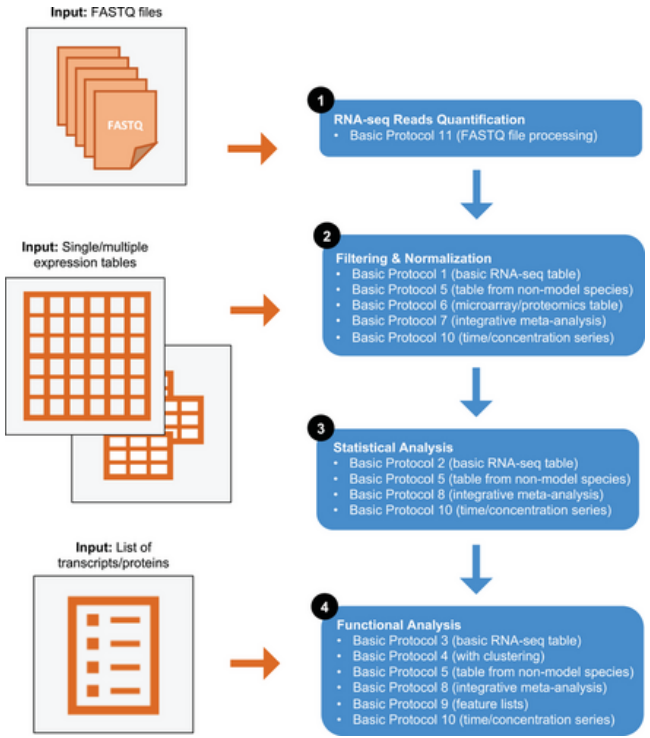


Next Lecture

Topic	Date	Lecture	Lab
Omics Data Science Foundations	Jan. 6	Omics data processing, statistics and visualization	--
	Jan. 13	From raw data to functional insights	--
Transcriptomics	Jan. 20	Gene expression data analysis (part I)	ExpressAnalyst
	Jan. 27	Gene expression data analysis (part II)	ExpressAnalyst
miRNAs & non-coding RNAs	Feb. 3	MicroRNAs, noncoding RNAs and biological networks	miRNet & NetworkAnalyst
Proteomics	Feb. 10	Proteomics data analysis and integration	ExpressAnalyst & NetworkAnalyst
Metabolomics	Feb. 17	Targeted metabolomics data analysis	MetaboAnalyst
	Feb. 24	LC-MS untargeted metabolomics data analysis	MetaboAnalyst
Microbiomics	Mar. 2	Marker gene data analysis	MicrobiomeAnalyst
	Mar. 9	Shotgun metagenomics data analysis	MicrobiomeAnalyst
Multi-omics	Mar. 16	Knowledge-driven multi-omics integration	OmicsNet
	Mar. 23	Data-driven multi-omics integration	OmicsAnalyst

Comprehensive protocols (76 pages)

Basic Protocol 1: RNA-seq count table uploading, processing, and normalization
Basic Protocol 2: Differential expression analysis with linear models
Basic Protocol 3: Functional analysis with volcano plot, enrichment network, and ridgeline visualization
Basic Protocol 4: Hierarchical clustering analysis of transcriptomics data using interactive heatmaps
Basic Protocol 5: Cross-species gene expression analysis based on ortholog mapping results
Basic Protocol 6: Proteomics and microarray data processing and normalization
Basic Protocol 7: Preparing multiple gene expression tables for meta-analysis
Basic Protocol 8: Statistical and functional meta-analysis of gene expression data
Basic Protocol 9: Functional analysis of transcriptomics signatures
Basic Protocol 10: Dose-response and time-series data analysis
Basic Protocol 11: RNA-seq reads processing and quantification with and without reference transcriptomes



Tutorials

CURRENT PROTOCOLS

PROTOCOL | Open Access

Using ExpressAnalyst for Comprehensive Gene Expression Analysis in Model and Non-Model Organisms

Jessica Ewald, Guangyan Zhou, Yao Lu, Jianguo Xia

First published: 06 Novem

Basic Protocol 6: Proteomics data processing and normalization

Metabolomics (2013) 9:280–299
DOI 10.1007/s11306-012-0482-9

REVIEW ARTICLE

biomarker analysis

Translational biomarker discovery in clinical metabolomics: an introductory tutorial

Jianguo Xia · David I. Broadhurst · Michael Wilson · David S. Wishart

We would like to hear your comment & feedback

contact@xialab.ca

See (most of) you next week!