

OmicsNet and OmicsNetR

Web server project

Mateusz Wiktorzak

11.05.2023

What are omics?

They relate to disciplines with „-omics” suffix, such as genomics, proteomics, metabolomics, metagenomics, phenomics, transcriptomics – studies about molecules living within living organisms.

Introduction

OmicsNetR

```
# Step 1. Initiate the dataSet object  
dataSet<-Init.Data()
```

```
# Step 2. Map list of genes to the application  
dataSet<-PrepareInputList(dataSet,"#Entrez logFC  
4495 61.12  
4496 51.06  
4499 23.79  
6354 21.04  
6369 19.76", "hsa", "gene", "entrez");
```

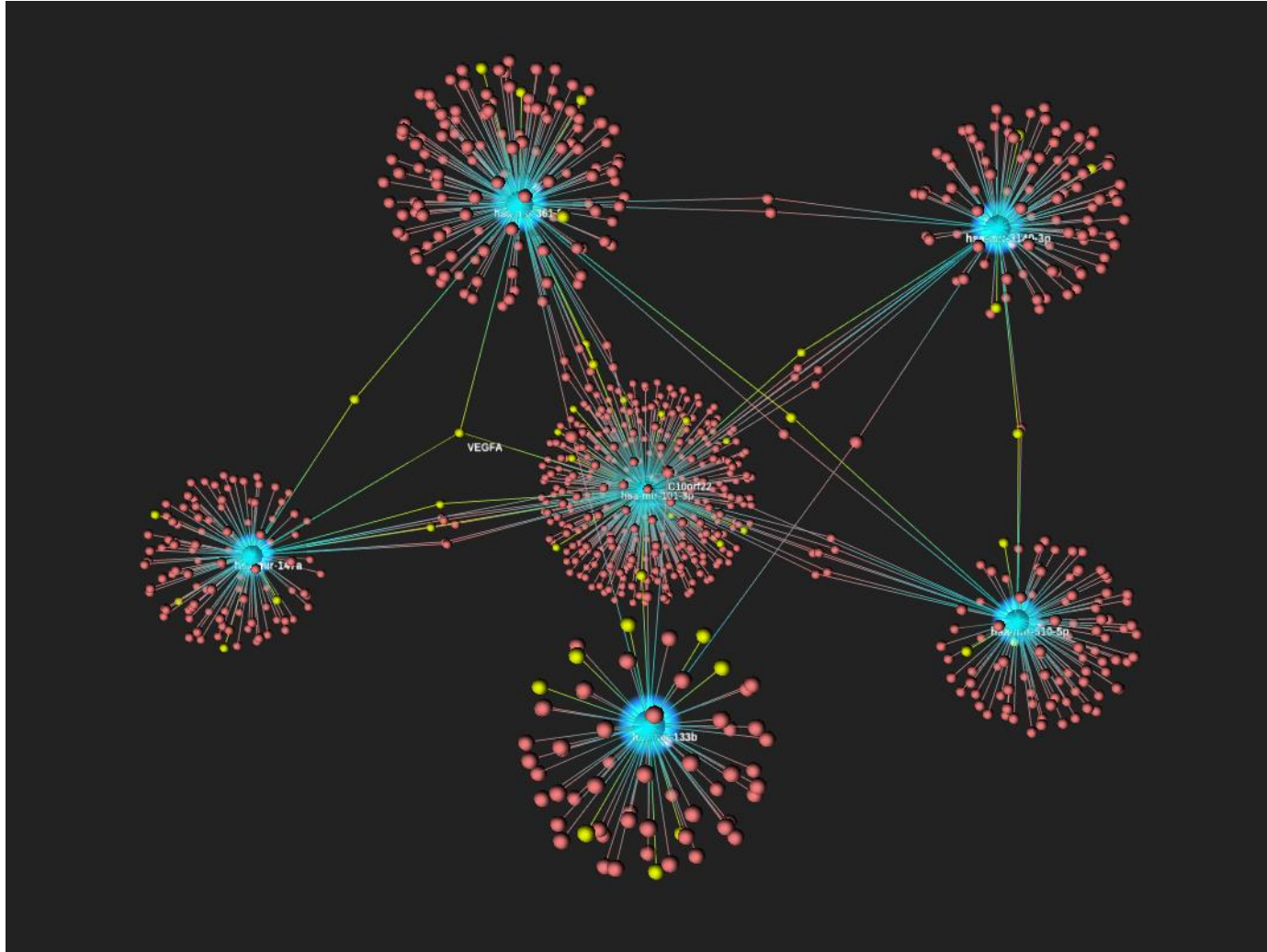
```
# Step 3. Identify interacting partners  
dataSet<-QueryNet(dataSet, "gene", "innate")
```

```
# Step 4. Build interaction subnetwork  
CreateGraph();
```

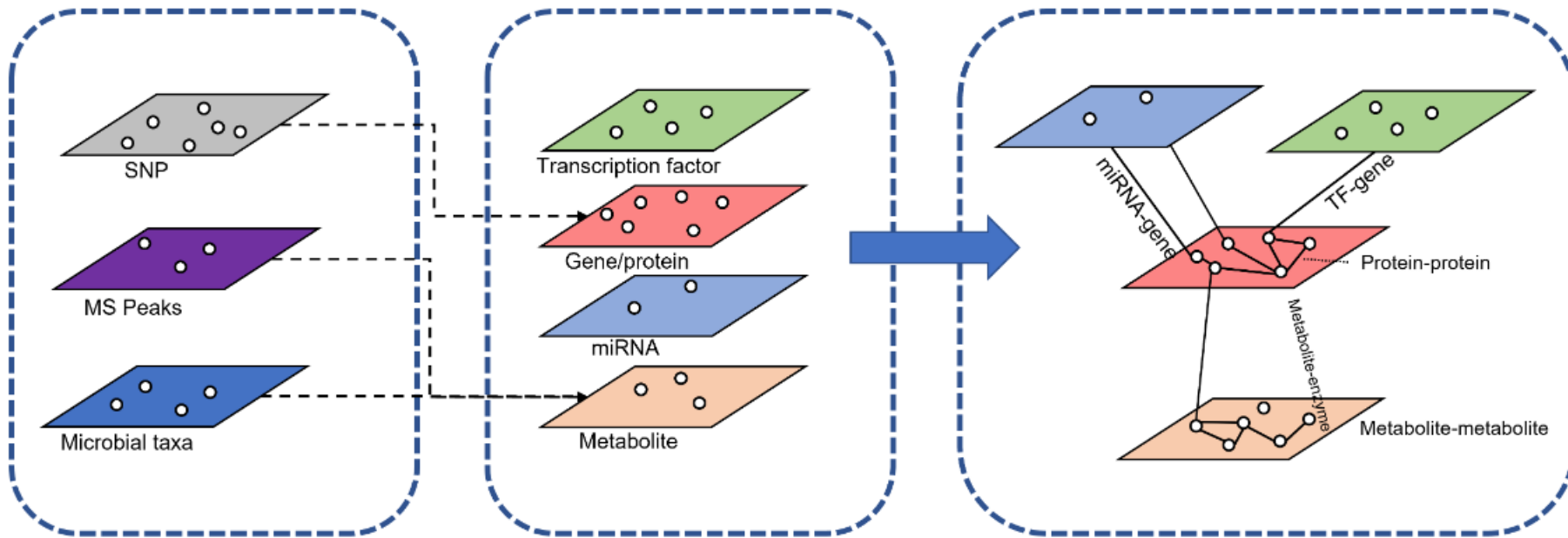
```
"id": "6354",  
"idx": 1,  
"label": "CCL7",  
"displayedLabel": "CCL7",  
"size": 32,  
"size2d": 12,  
"type": "protein",  
"molType": "protein",  
"types": "protein",  
"seedArr": "seed",  
"color": "#FF8484",  
"colorb": "#FF8484",  
"colorw": "#FF8484",  
"topocolb": "#FF8484",  
"topocolw": "#FF8484",  
"expcolb": "#BD0313",  
"expcolw": "#96000D",  
"x2d": -0.2812434,  
"y2d": -0.0641488,  
"user": {  
  "closeness": 0.05882353,  
  "eigen": 1  
},  
"attributes": {  
  "degree": 15,  
  "between": 119,  
  "expr": 21.04,  
  "closeness": 0.05882353,  
  "eigen": 1  
}  
},
```

```
"edges": [ {  
  "id": "1",  
  "source": "6354",  
  "target": "1499",  
  "color": "target"  
},  
{  
  "id": "2",  
  "source": "6354",  
  "target": "166",  
  "color": "target"  
}],
```

OmicsNet



How OmicsNet work?



Goal:
Link untargeted MS
peaks, SNPs and
microbial taxa to
knowledge framework

SNP – single nuclear polymorphism

MS peaks – mass spectrum peaks, they concern heaviness of ions

Microbial taxa – specifies how microorganisms can be grouped together

Mentioned biological molecules

- Gene: A segment of DNA, encodes a protein or RNA molecule.
- Protein: Complex molecule.
- Transcription factor: A regulatory protein that binds to DNA sequences and controls the process of transcription.
- miRNA (microRNA): Play a significant role in gene expression regulation.
- Metabolite: Products of metabolic processes in organisms.

SNP

- SNPs (Single Nucleotide Polymorphisms) are common genetic variations that occur within a population.
- In a SNP, the variation occurs when a single nucleotide (A, T, C, or G) is replaced by another nucleotide.
- They can have different effects on gene function and can be associated with various traits, diseases, and drug responses.
- Over the past two decades extensive collections of SNPs were made.
- SNPs make it possible to perform SNP-gene mapping.

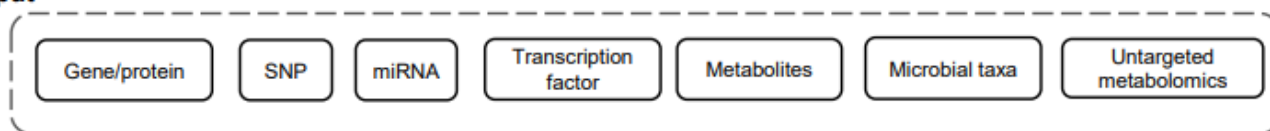
LC-MS peaks

- LC-MS peaks refer to peaks obtained in liquid chromatography-mass spectrometry analysis.
- By analyzing LC-MS peaks in OmicsNet, one can obtain information about the chemical composition of samples, identify potential biomarkers, and discover biological dependencies between chemical compounds. This tool contributes to a better understanding of complex omics data and their biological significance.

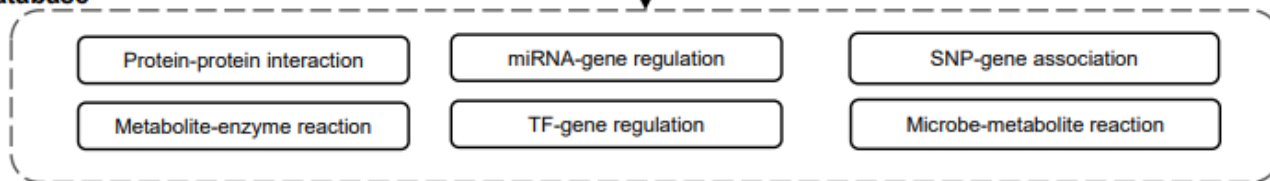
Microbial taxonomics

- Microbial taxa can contribute to network construction by representing identified microorganisms as nodes and their relationships or associations as edges.
- It can be used for the analysis and interpretation of omics data.
- Taxonomic data can be used to analyze differences in the composition of the microbiome between groups of healthy and diseased patients. It can also be used to identify microbial biomarkers, which are microorganisms or groups of microorganisms associated with specific health states or diseases.

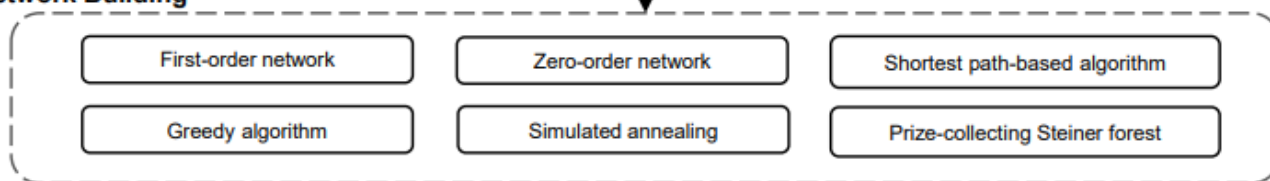
Input



Database

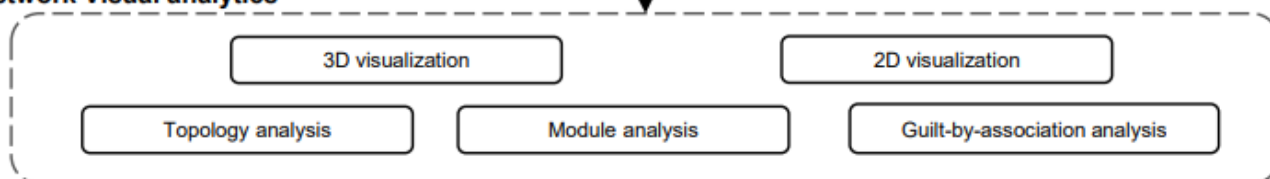


Network Building



Multi-omics network

Network Visual analytics



Export



Input

A Graph File

SNPs

Microbial Taxa

LC-MS Peaks

Genes/mRNAs

Proteins

Transcription Factors

miRNAs

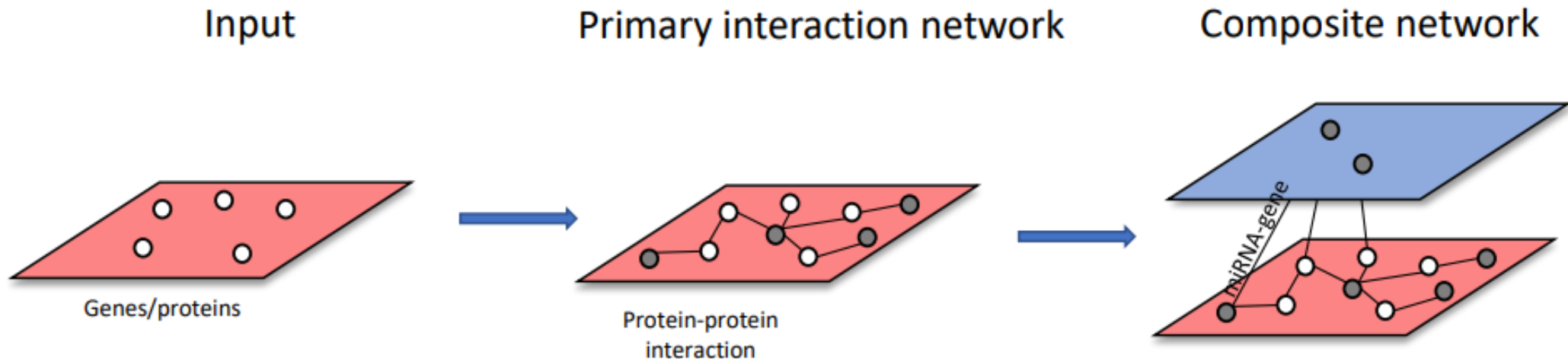
Metabolites

Key Features of OmicsNet

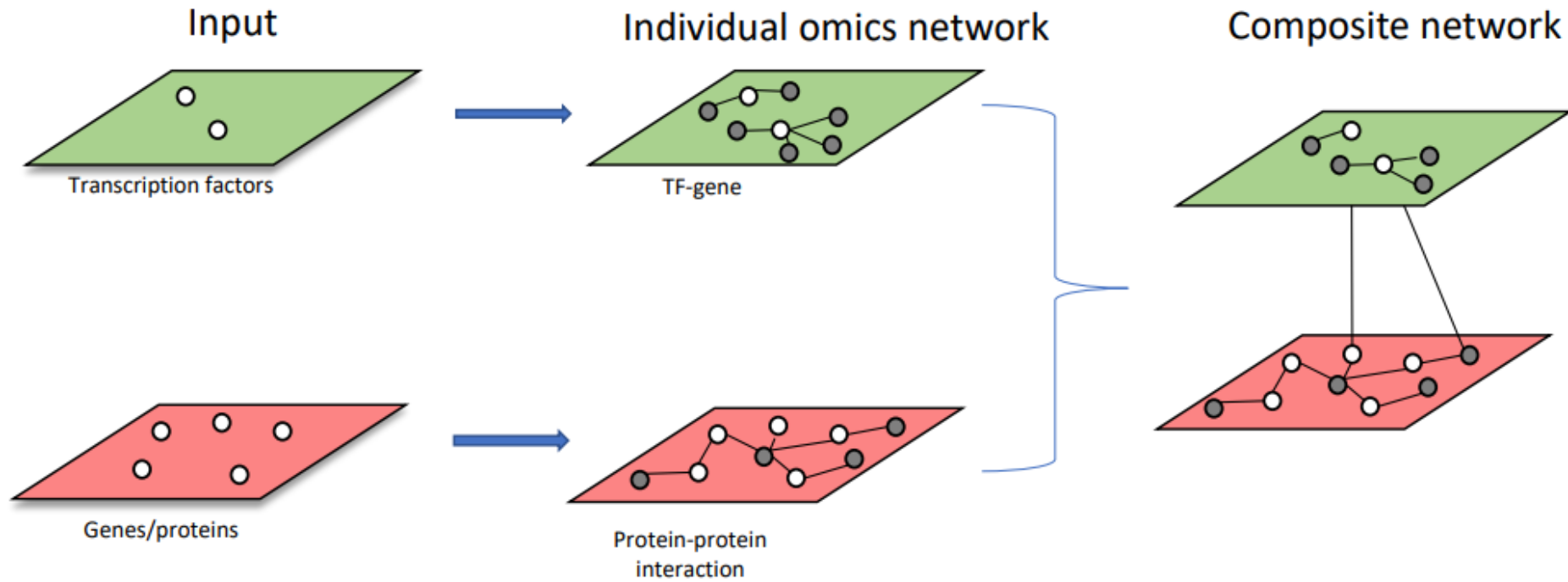
- Interactive 2D/3D network visualization in web browser.
- Enable integration of multiple types of molecular interactions.
- Built-in knowledgebase for network building:
 - Protein-protein interaction
 - Transcription factor-gene regulation
 - miRNA-gene regulation
 - Metabolite-protein interaction
- Support integration of less well-established omics data types by using computational predictions:
 - SNPs
 - Microbial taxa
 - Untargeted metabolomics MS peaks

Single list network building

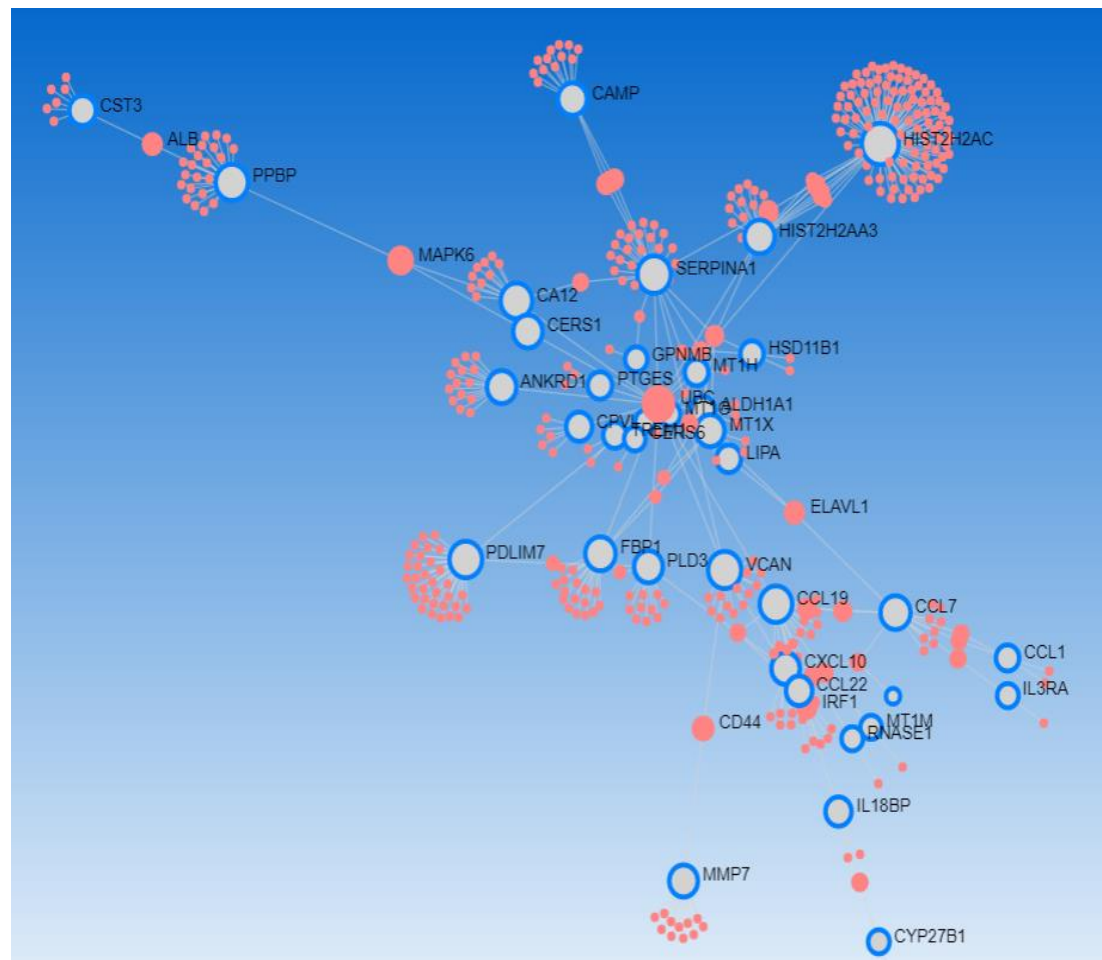
- Primary interaction network – seeds and immediate interacting partners
- Secondary interactions are searched in molecules contained in the primary network



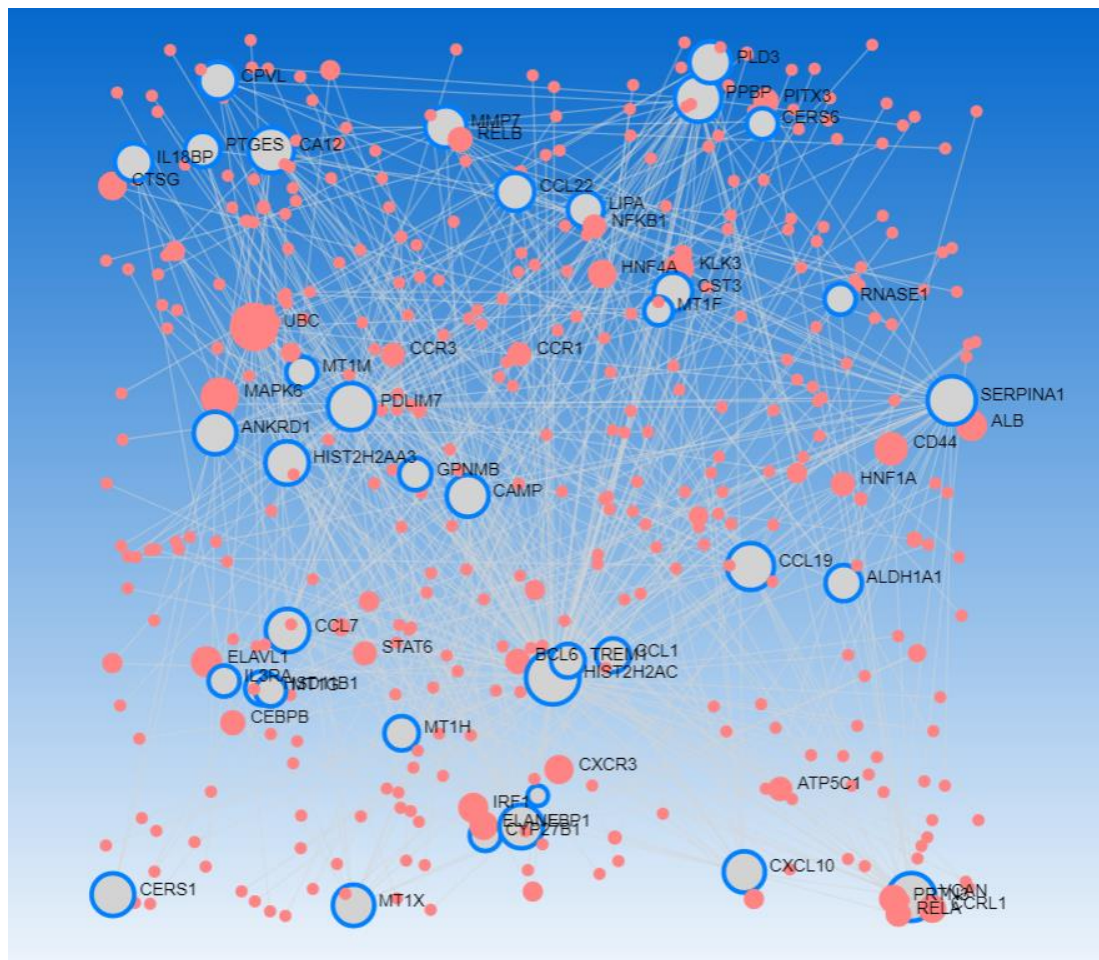
Multiple list network building



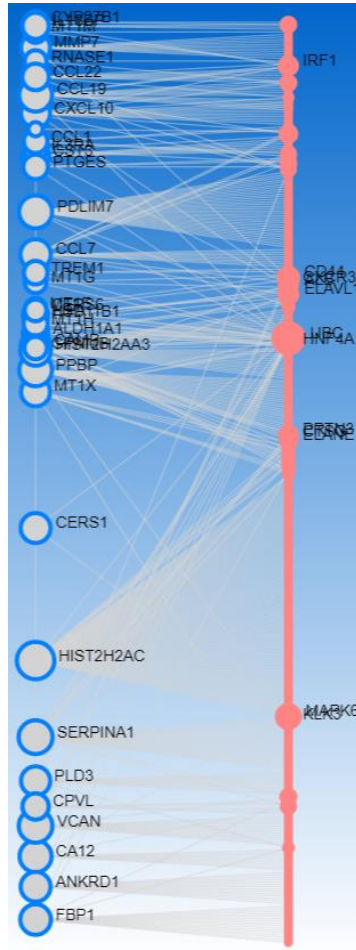
Auto Layout



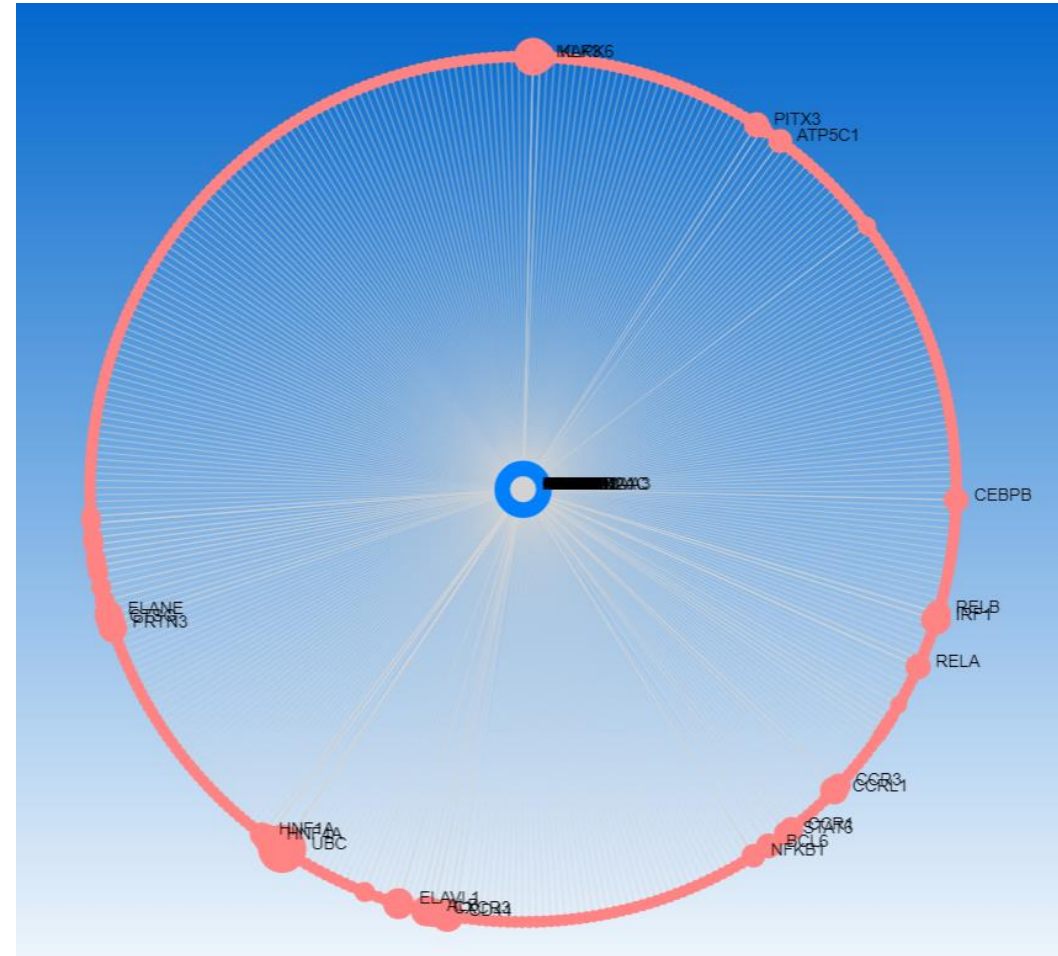
Random Layout



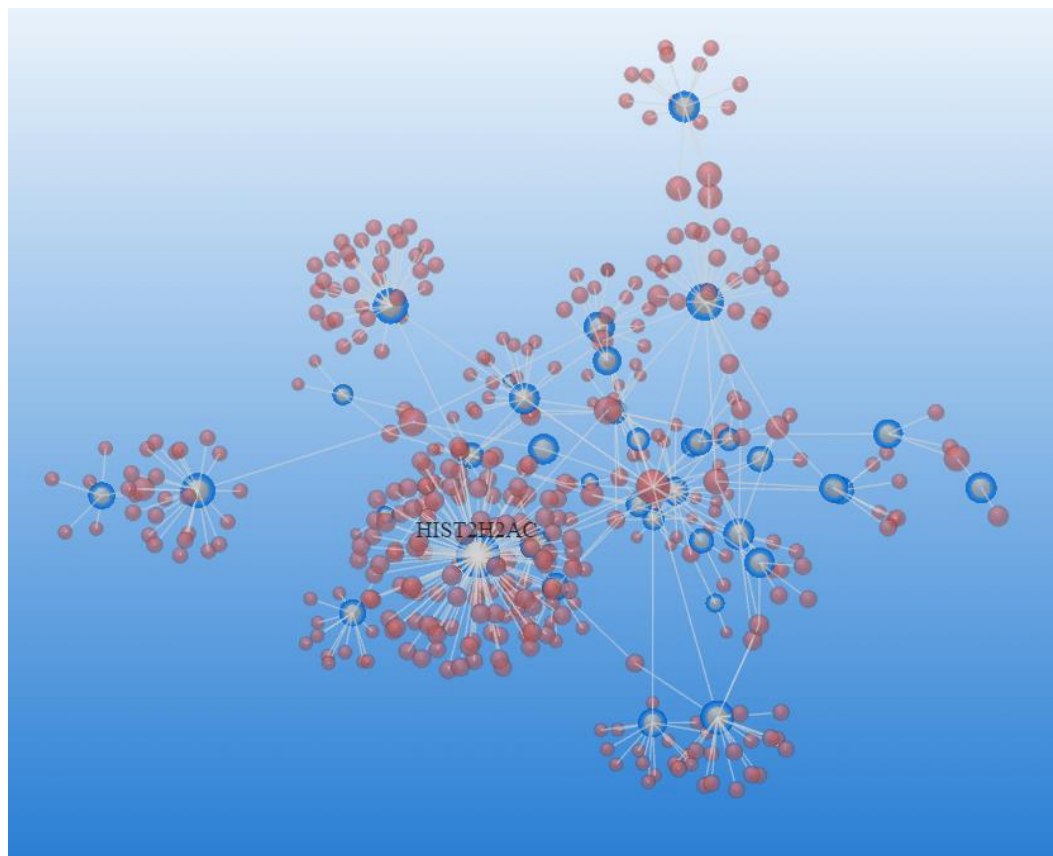
Linear Bipartite



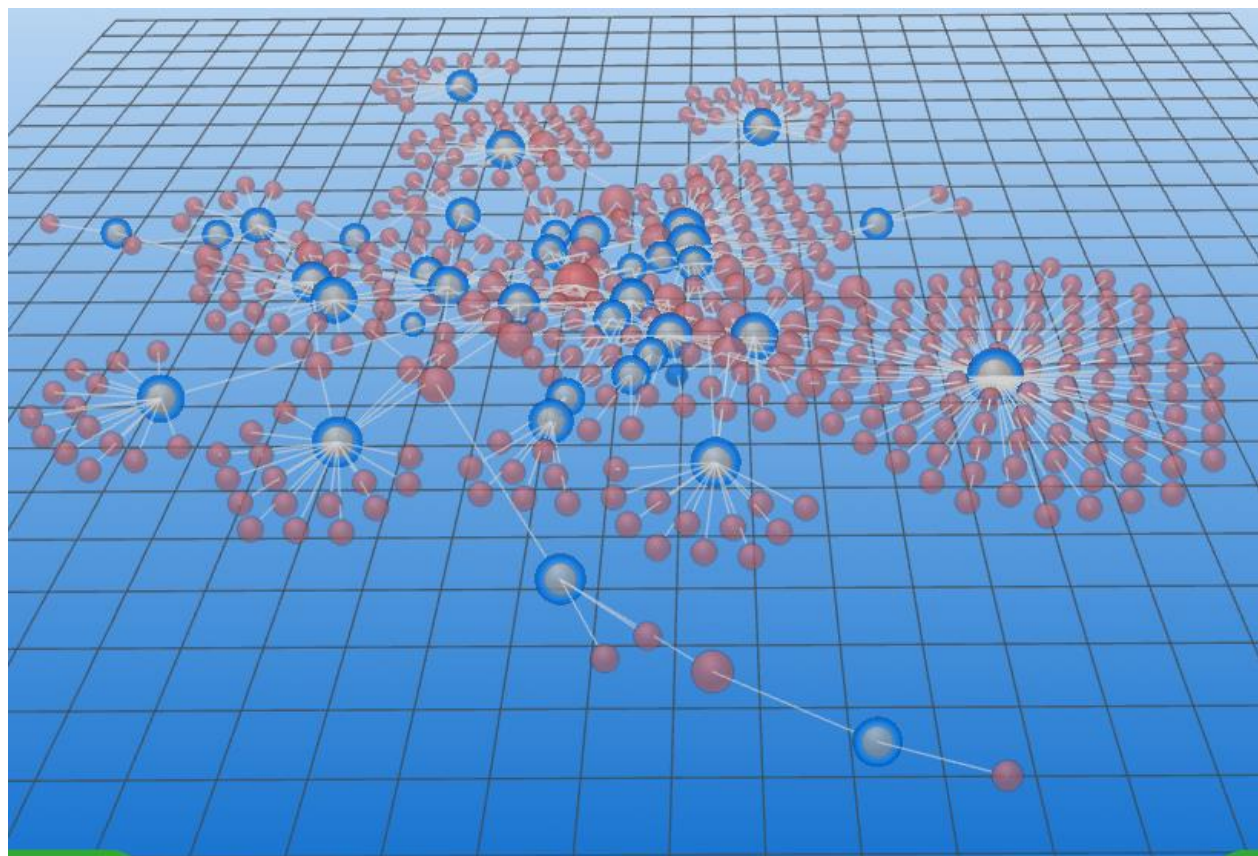
Circular Bipartite



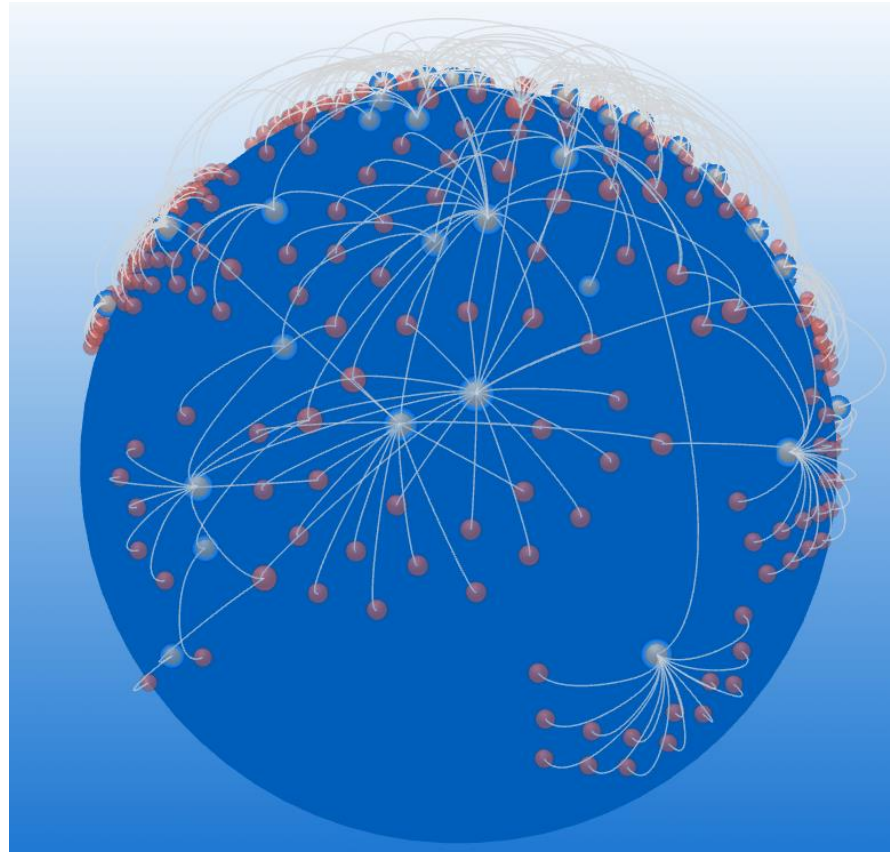
Standard



Perspective

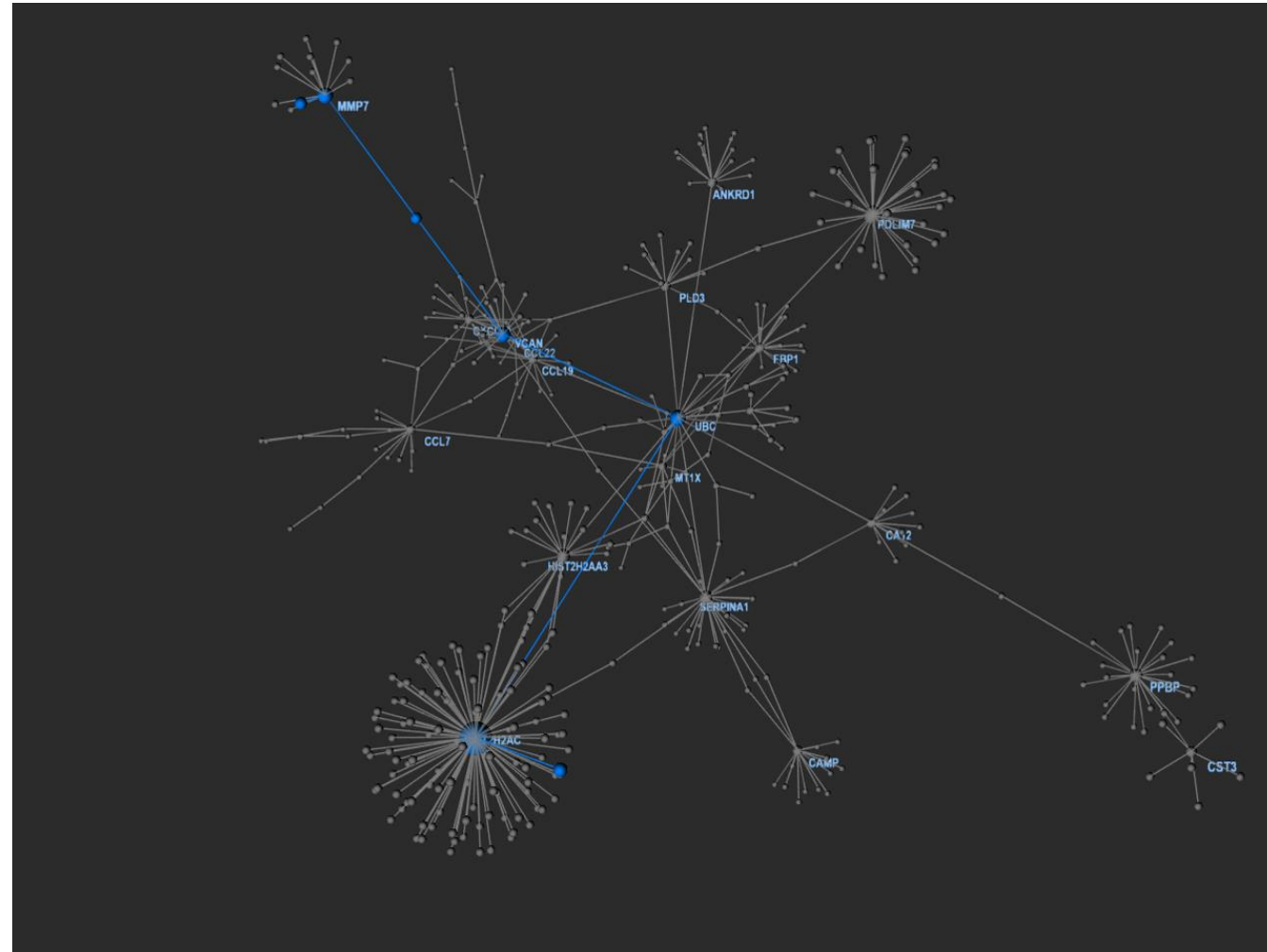


Spherical

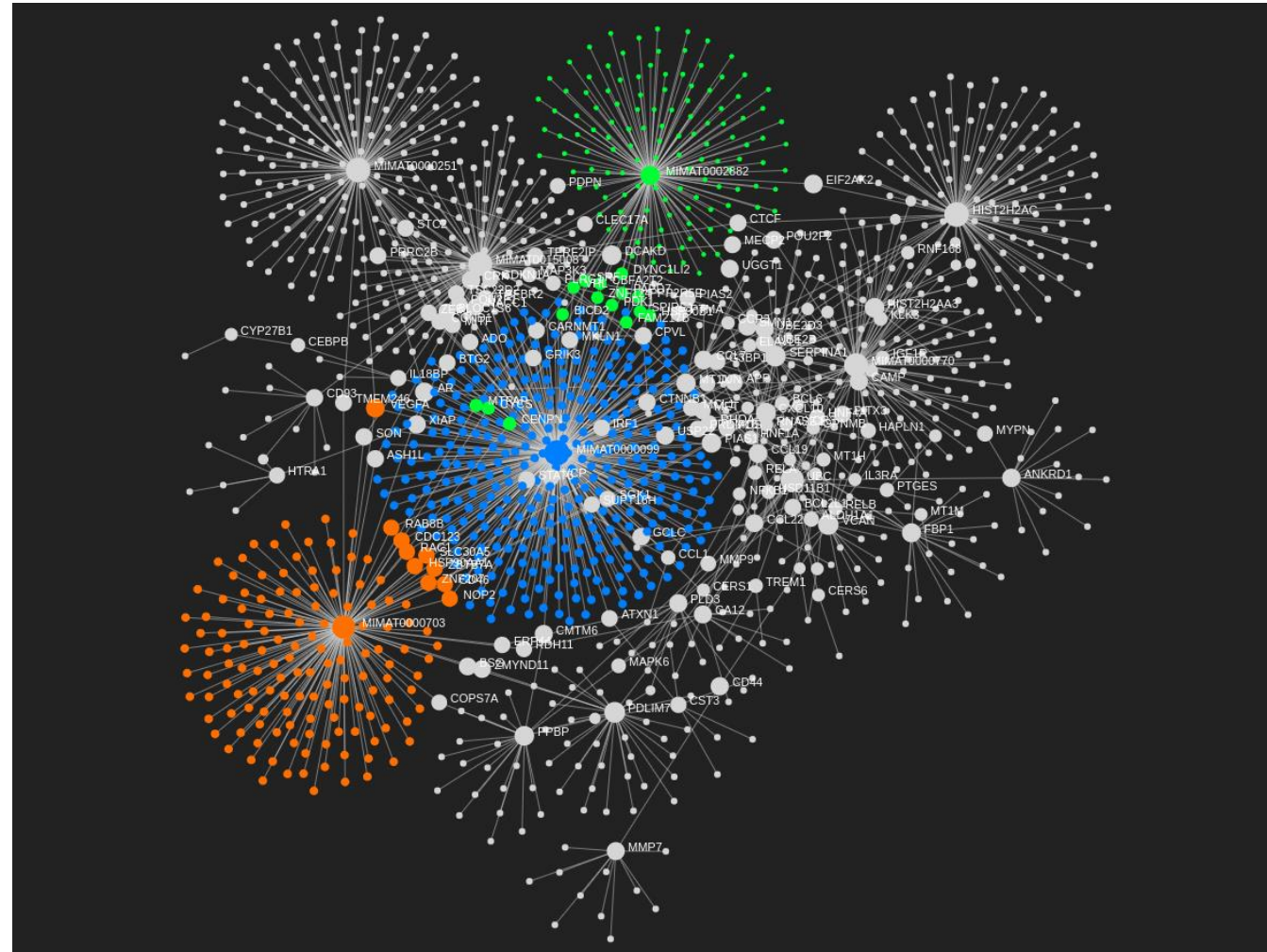


Analysis in OmicsNet

PPI network showing the shortest path between two nodes



mRNA-miRNA network with top three modules highlighted



Visualizations in OmicsNet

Transcription factor and mRNA/Protein Composite network

Global Node Styles

Type	Size	Color
Transcription factor	<input type="range"/>	<input type="color" value="#00FF00"/>
mRNA/Protein	<input type="range"/>	<input type="color" value="#808080"/>

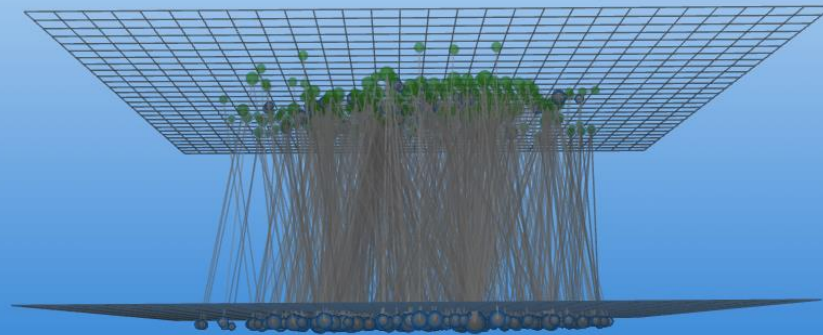
Node Table

Search

<input type="checkbox"/>	ID	Name	Degree	Expr.
<input type="checkbox"/>	1026	CDKN1A	115	
<input type="checkbox"/>	4609	MYC	91	
<input type="checkbox"/>	6667	SP1	47	
<input type="checkbox"/>	4790	NFKB1	39	
<input type="checkbox"/>	5970	RELA	38	
<input type="checkbox"/>	2099	ESR1	37	
<input type="checkbox"/>	4318	MMP9	33	
<input type="checkbox"/>	4313	MMP2	32	
<input type="checkbox"/>	355	FAS	24	
<input type="checkbox"/>	7157	TP53	23	
<input type="checkbox"/>	7852	CXCR4	23	

Page 1 of 20

Current Selections



Function Explorer

Query: Database:

Name	Hits	P-val	P-val(ad)	Color
Non-small cell lung cancer	5	0.0355	0.121	
RIG-I-like receptor signaling pathway	5	0.0355	0.121	
Citrate cycle (TCA cycle)	7	0.038	0.128	
ABC transporters	3	0.0395	0.131	
Bladder cancer	5	0.0413	0.135	
Basal cell carcinoma	5	0.0413	0.135	
Adrenergic signaling in cardiomyocytes	5	0.0433	0.14	
Leukocyte transendothelial migration	3	0.0465	0.149	
Ribosome biogenesis in eukaryotes	6	0.0484	0.153	

Module Explorer

Algorithm:

<input type="checkbox"/>	Module	Size	P-value	Color
<input type="checkbox"/>	0	528	1.28e-190	
<input type="checkbox"/>	1	14	0.000105	

[Guilt-by-Association Analysis](#)

[Regulation Explorer](#)

[Path Explorer](#)

[Batch Selection](#)

<< Previous

Proceed >>

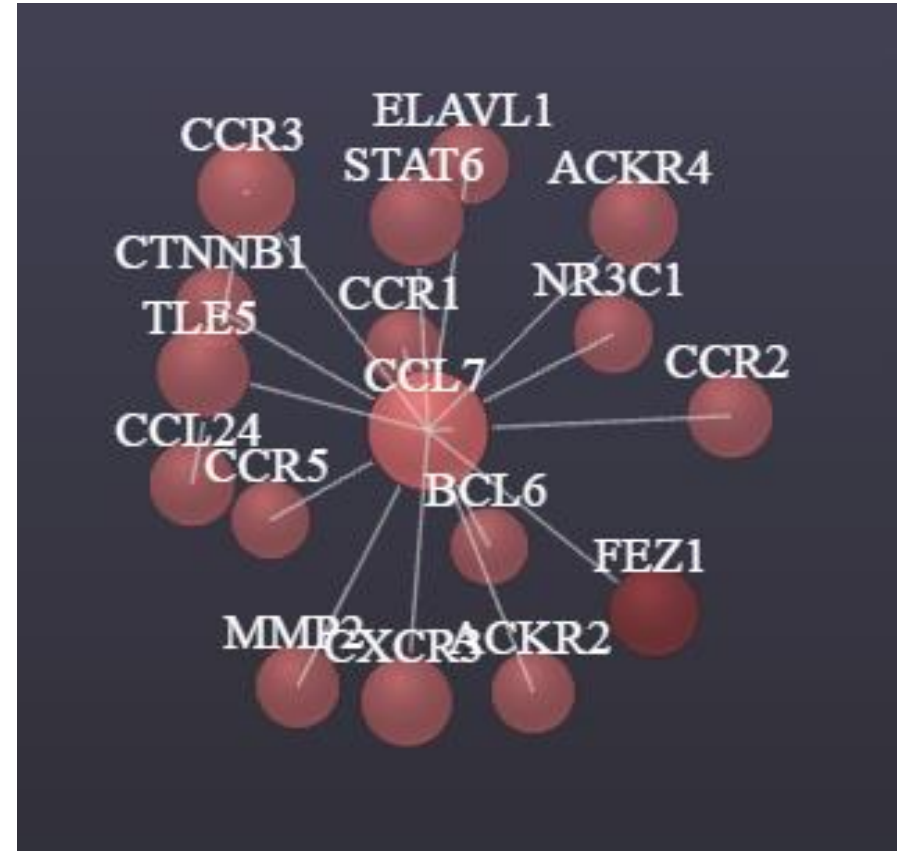
List of genes

```
# Step 1. Initiate the dataSet object
dataSet<-Init.Data()

# Step 2. Map list of genes to the application
dataSet<-PrepareInputList(dataSet,"#Entrez  logFC
4495  61.12
4496  51.06
4499  23.79
6354  21.04
6369  19.76", "hsa", "gene", "entrez");

# Step 3. Identify interacting partners
dataSet<-QueryNet(dataSet, "gene", "innate")

# Step 4. Build interaction subnetwork
CreateGraph();
```



Integration of genes and miRNA

```
# Step 1. Initiate the dataSet object
dataSet<-Init.Data();

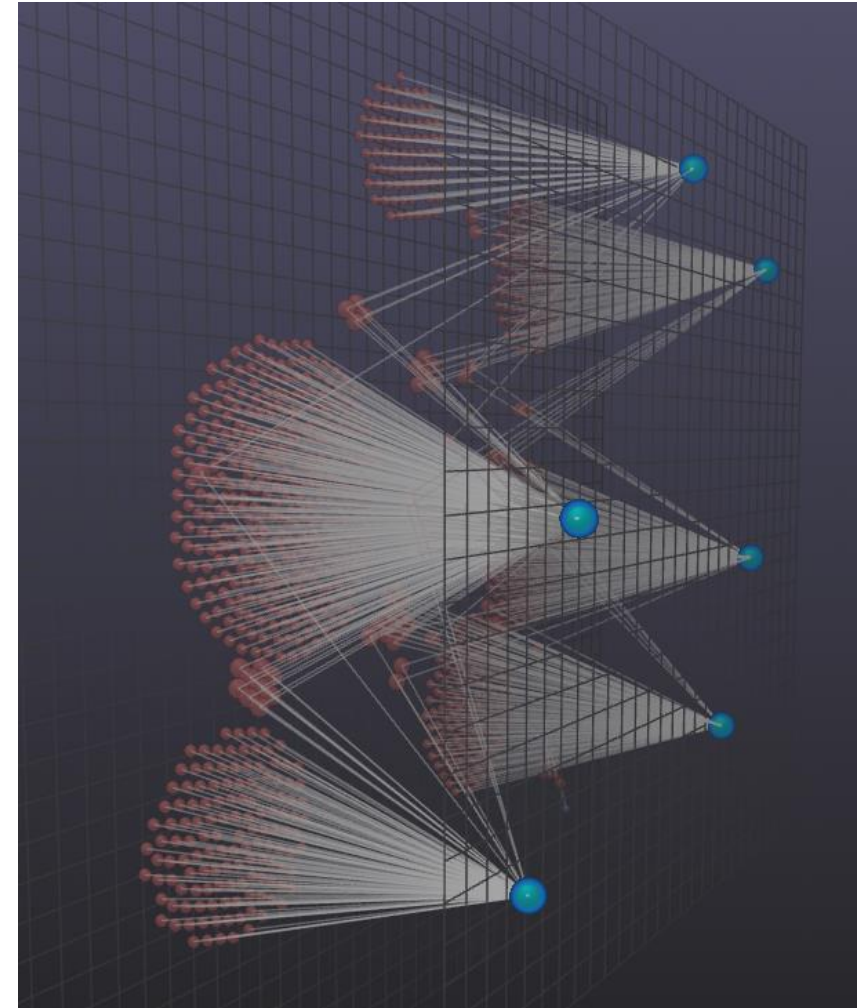
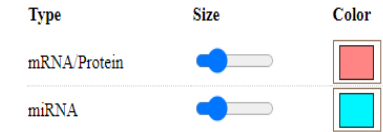
# Step 2. Map list of genes to the application
dataSet<-PrepareInputList(dataSet,
                           "#Entrez  logFC
4495  61.12
4496  51.06
4499  23.79
6354  21.04
6369  19.76", "hsa", "gene", "entrez");

# Step 2. Map list of miRNA to the application
dataSet<-PrepareInputList(dataSet,
                           "hsa-mir-101-3p
hsa-mir-133b
hsa-mir-147a
hsa-mir-3140-3p
hsa-mir-361-5p
hsa-mir-510-5p", "hsa", "mir", "mir_id");

# Step 3. Build PPI network from uploaded list of genes
dataSet<-QueryNet(dataSet, "gene", "innate", "gene");

# Step 4. Build miRNA-gene network from uploaded list of miRNA
dataSet<-QueryNet(dataSet, "mir", "mirtarbase", "mir");

# Step 5. Merge networks together through shared nodes
# and decompose into interconnected subnetworks
dataSet<-CreateGraph(dataSet);
```



Integration of SNPs and Microbial Taxa

```
# Step 1. Initiate the dataSet object
dataSet<-Init.Data();

# Step 2. Input the microbial taxa data
dataSet<-PrepareInputList(dataSet,
                           "Faecalibacterium_prausnitzii"
                           Bacteroides_uniformis
                           Eubacterium_rectale
                           Alistipes_putredinis
                           Subdoligranulum_unclassified
                           Escherichia_coli
                           Bacteroides_vulgatus
                           Clostridium_clostridioforme
                           Klebsiella_pneumoniae
                           Clostridium_hathewayi
                           Alistipes_shahii
                           Ruminococcus_obebum", "microbiome", "mic", "species");

# Step 3. Input the SNP data
dataSet<-PrepareInputList(dataSet,
                           "rs1428554"
                           rs41290504
                           rs3197999
                           rs516246
                           rs2228058", "hsa", "snp", "rsid");

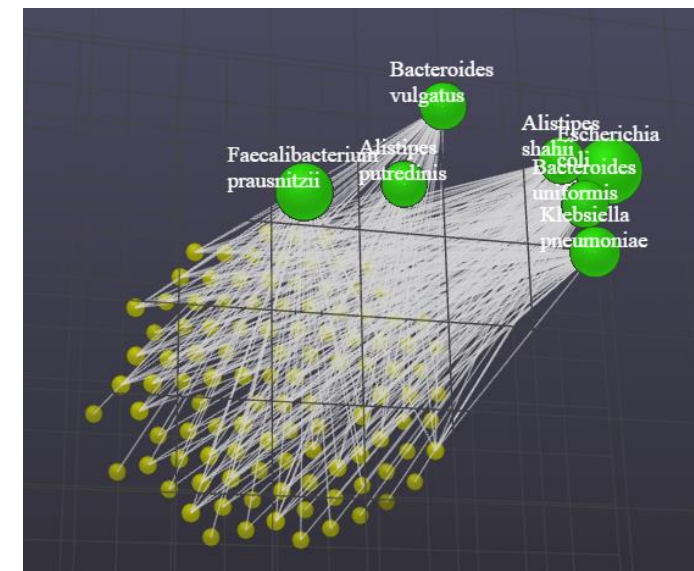
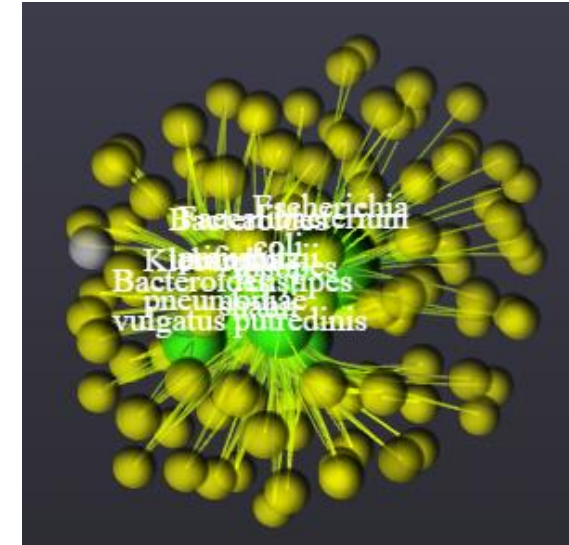
# Step 4. SNP data annotation and network construction
SetSnp2GeneOpts("vep", "", "dis", 1, 5);
dataSet<-QueryNet(dataSet, "snp", "vep", "snp");

# Step 5. Protein-protein extension
SetPpiZero(FALSE);
dataSet<-QueryNet(dataSet, "gene", "huri", "snp");

# Step 6. Protein-metabolite extension based on recon3D database
dataSet<-QueryNet(dataSet, "met", "recon3D", "snp");

# Step 7. Microbial data annotation and network construction
SetOrganism("microbiome");
SetMetPotentialOpts("0.8", "TRUE", "TRUE", "TRUE");
dataSet<-QueryNet(dataSet, "mic", "agora", "mic");

# Step 8. Merge networks together through shared nodes
# and decompose into interconnected subnetworks
dataSet<-CreateGraph(dataSet);
```



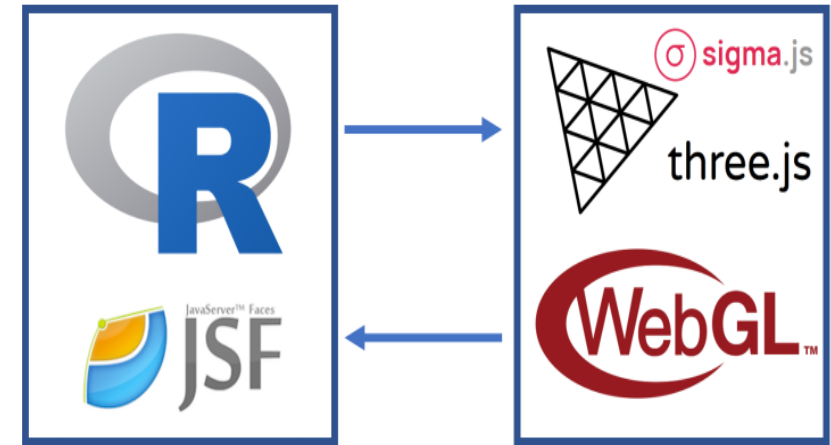
Comparison with other tools

Tools	OmicsNet	PaintOmics	MergeOmics	OmicsAnalyst	Arena3D	NeDRex	MetScape
Type	Web	Web	Web	Web	Web	Cytoscape plugin	Cytoscape plugin
Input	Lists from 8 omics types, graph files	Abundance tables from 5 omics types	Association data, gene sets, networks	Omics feature abundance tables	Graph files	Gene list, gene expression table	Gene expression and metabolomics tables
Network creation							
SNP annotation	√	-	√	-	-	-	-
Peak annotation	√	-	-	-	-	-	-
Taxon annotation	√	-	-	-	-	-	-
Network integration	√	√	√	Correlation	Multi-layer	√	√
Network visualization							
3D view	+++	-	-	√	√	-	
2D view	√	√	√	√	-	√	√
Layered layout	+++	-	-	√	√	Cytoscape	Cytoscape
Spherical layout	√	-	-	-	-	-	-
Backbone layout	√	-	-	√	-	-	-
Concentric layout	√	-	-	√	-	Cytoscape	Cytoscape
Edge bundling	√	-	-	√	-	Cytoscape	Cytoscape

Tools	OmicsNet	PaintOmics	MergeOmics	OmicsAnalyst	Arena3D	NeDRex	MetScape
Network analysis							
Enrichment Analysis	+++	+	+++	+++	-	+++	+++
Joint enrichment analysis	√	√	√	√	-	-	-
Module detection	+++	-	√	√	-	+++	-
Biomarker prioritization	++	-	√	-	-	+++	-

Implementation

- Client-server framework that uses Java and R server for network construction and some analysis functions.
- Browser-based application using WebGL interfaced by Three.js to visualize 3D network interactively.
- Sigma.js, a canvas-based graphing library is used for 2D network visualization



3D visualization technologies

- WebGL:
 - JavaScript API allowing the rendering of 3D graphics in web browser without plug-ins
 - High performance, GPU accelerated
- Three.js
 - Framework interfacing with WebGL to display 3D graphics
 - Abstract away the difficulties of WebGL.

Databases

PPI

innateDB

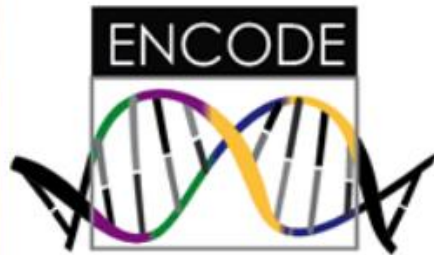
STRING

IntAct

TF-gene

TRRUST

JASPAR 2018

ENCODE

miRNA-gene

miRTarBase
<https://miRTarBase.cuhk.edu.cn>

TarBase

miRecords

Metabolic Reactions

Recon 2/3

KEGG
Kyoto Encyclopedia of
Genes and Genomes

Chemical reaction



Taxon-metabolite

Agora

EMBL



SNP-gene



Ensembl Variant
Effect Predictor

Use cases

- Protein interaction analysis: crucial for studying biological processes such as metabolic reactions.
- Metabolic pathway analysis: helps identify conditions that may lead to metabolic disorders or cancer.
- Gene expression analysis: It can assist in identifying biomarkers that can be used for disease diagnosis or predicting treatment response.
- Multi-omics data analysis: This allows for a more holistic view of biological processes and their interrelationships.
- It can aid in identifying new protein-protein or gene-gene interactions that may lead to the discovery of novel biological mechanisms.

Summary