Morteza Kouhsar

AmirKabir University of Technology

Computational Biology Research Center

MSC: Computer Science, Tarbiat Modares University.
PhD: Bioinformatics, Laboratory of Systems Biology and Bioinformatics (LBB), Institute of Biochemistry and Biophysics (IBB), University of Tehran.

Kouhsar, Morteza, Fatemeh Zare-Mirakabad, and Yousef Jamali. "WCOACH: Protein complex prediction in weighted PPI networks." Genes & genetic systems 90, no. 5 (2015): 317-324.

Motieghader, Habib, **Morteza Kouhsar**, Ali Najafi, Balal Sadeghi, and Ali Masoudi-Nejad. "mRNA–miRNA bipartite network reconstruction to predict prognostic module biomarkers in colorectal cancer stage differentiation." Molecular BioSystems 13, no. 10 (2017): 2168-2180.

Mehranfar, Adele, Nasser Ghadiri, **Morteza Kouhsar**, and Ashkan Golshani. "A Type-2 fuzzy data fusion approach for building reliable weighted protein interaction networks with application in protein complex detection." Computers in Biology and Medicine 88 (2017): 18-31.

AmirKabir University of Technology

Computational Biology Research Center

Computational Methods on a System Level: PPI Networks

L.B.B Laboratory of Systems Biology and Bioinformatics

Morteza Kouhsar

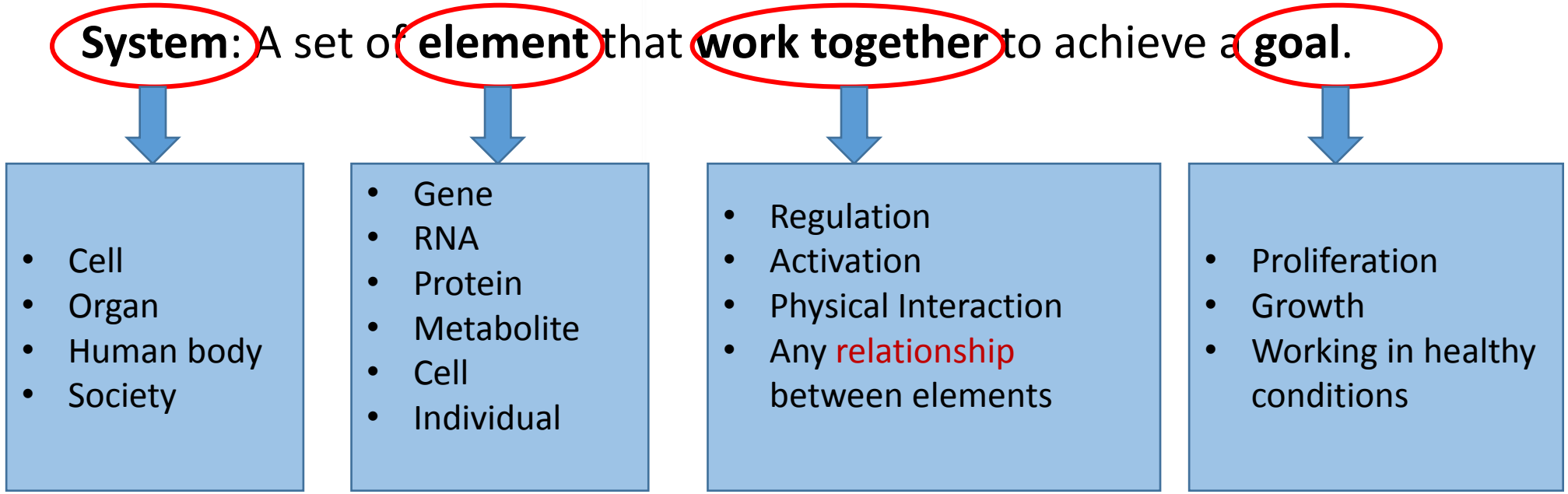
Laboratory of Systems Biology and Bioinformatics (LBB)

Outline

- Systems Biology and Biological Networks
- PPI Networks
  - Protein Physical Interactions
  - Resources and Databases
- PPI Network Analysis
  - Module Analysis
    - Module Detection
    - Gene Ontology and Enrichment Analysis
  - Centrality Measures and Node Analysis
- Cytoscape: A Powerful Framework for Systems Biology Analysis
  - Import Data
  - Installing Apps
  - Module detection and Enrichment

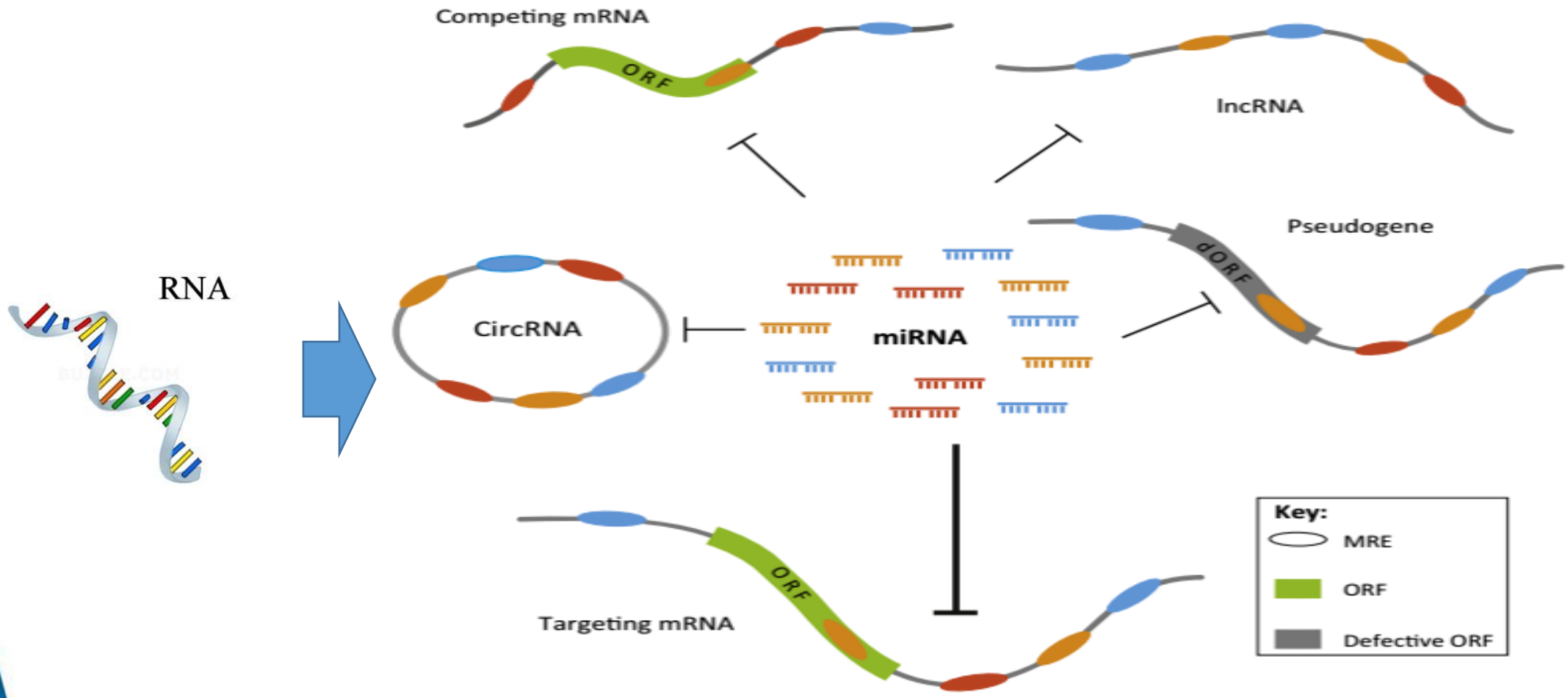
Systems Biology

System: A set of **element** that **work together** to achieve a **goal**.

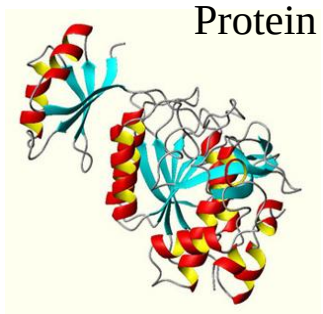


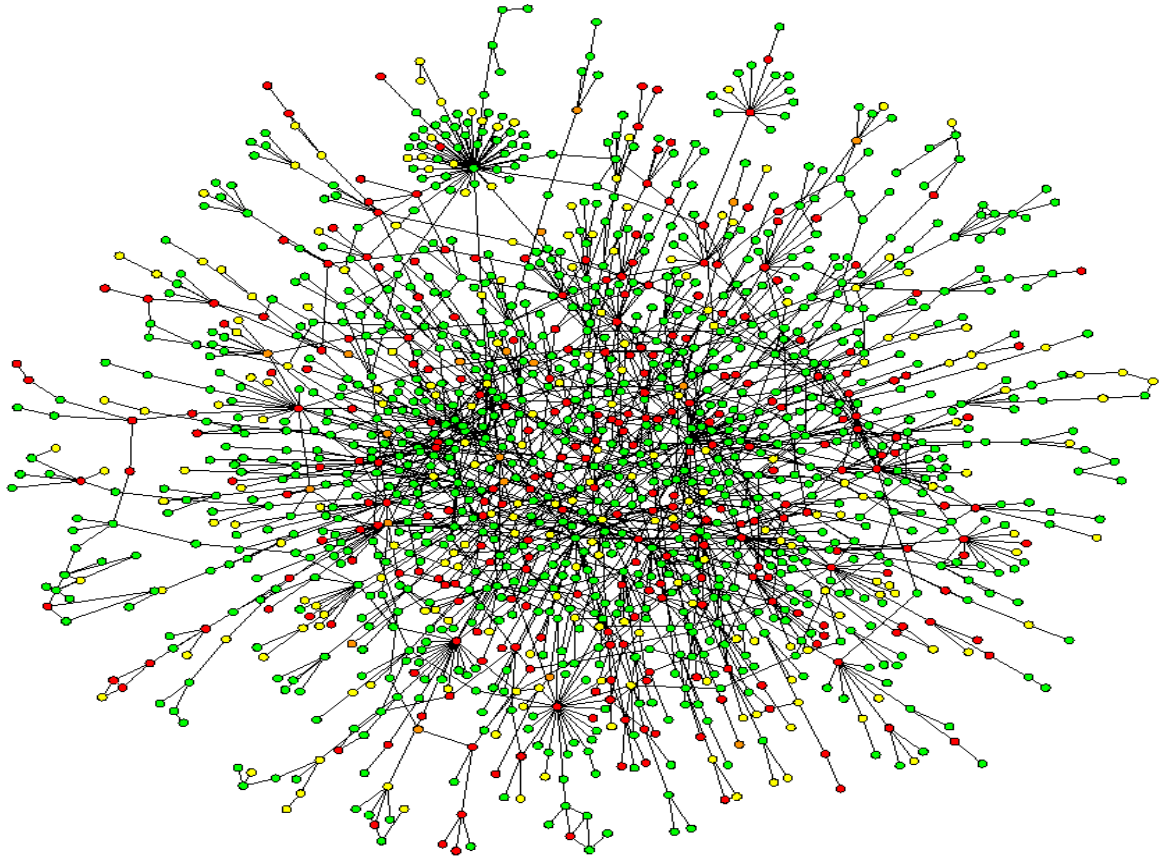
Systems Biology = Systematic view to the Biology

Systems Biology



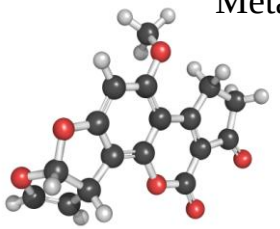
Systems Biology

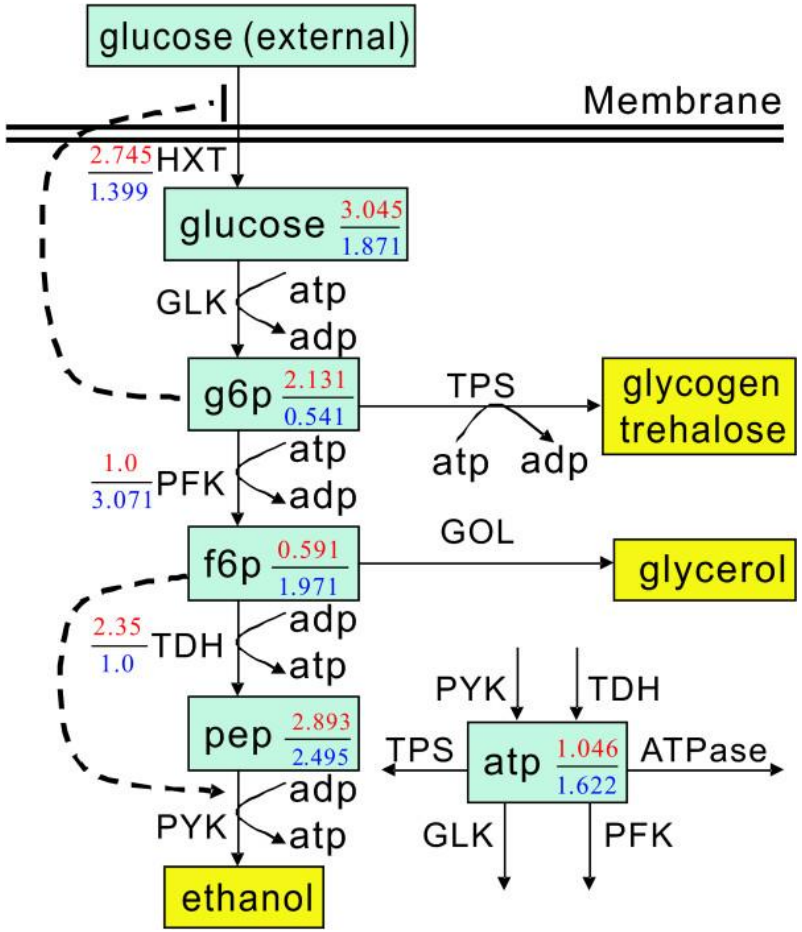




Systems Biology

Metabolite

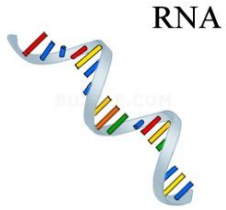




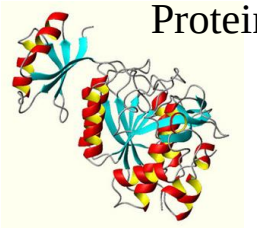
Systems Biology

Biological Elements

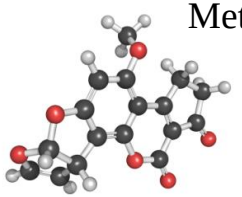
RNA

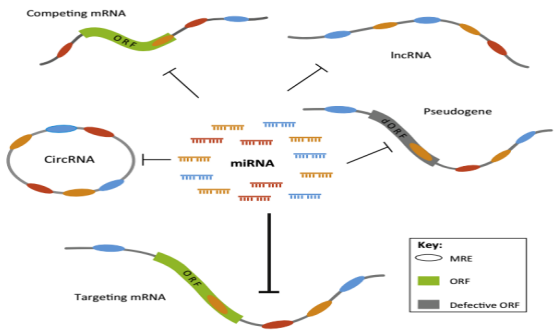


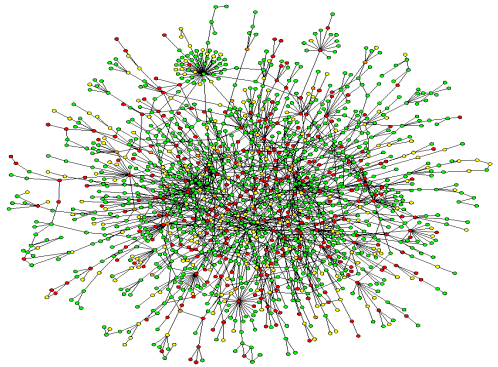
Protein

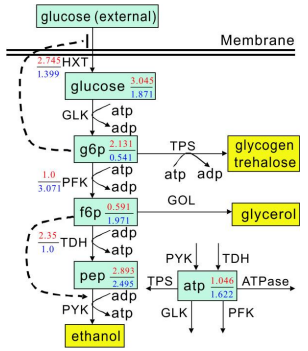


Metabolite









Biological Systems or Networks

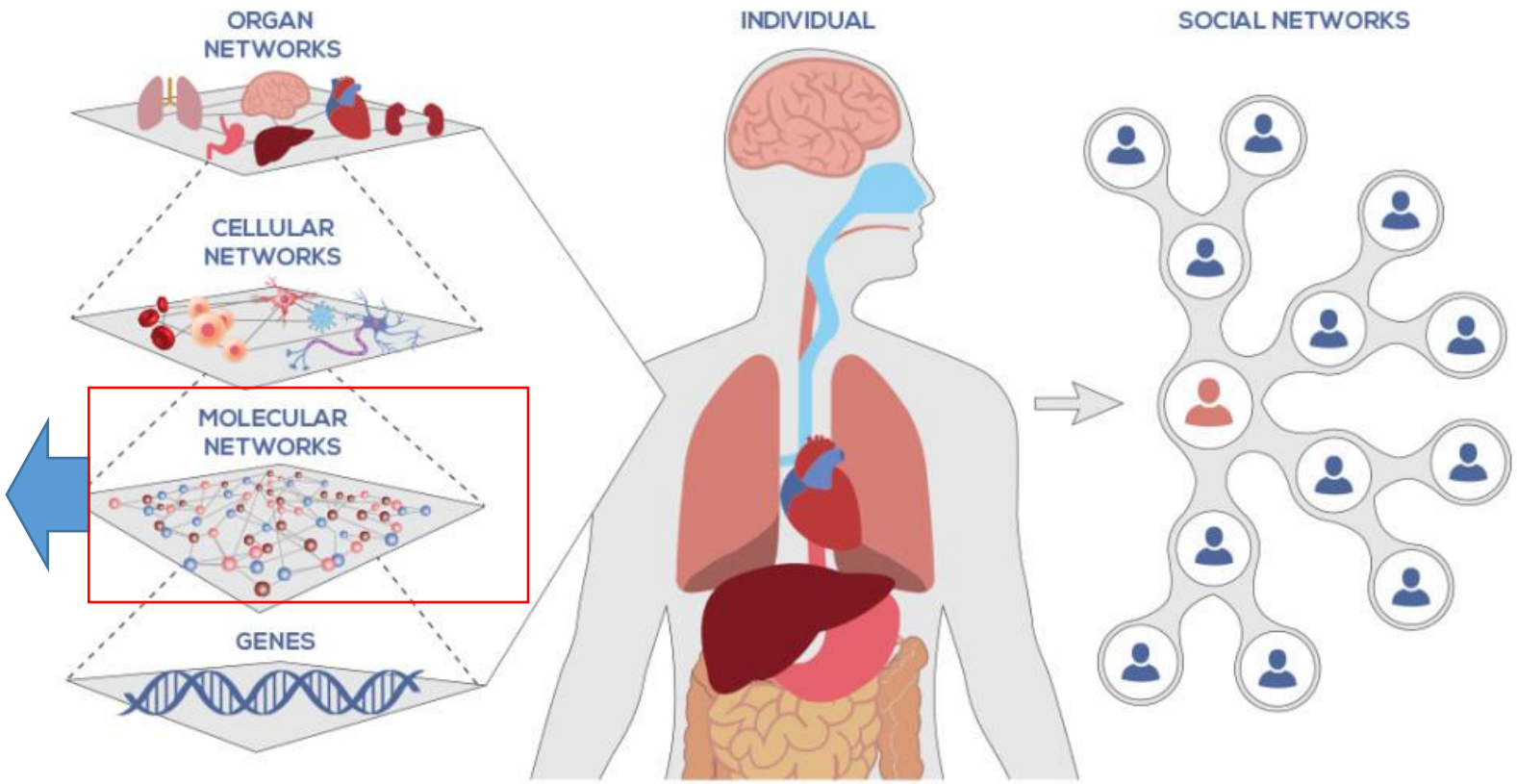
1st International Computational Biology workshop

Systems Biology

Analysis of **Biological Systems**
instead of
Biological elements individually
to understand
System behavior

Systems Biology Resources

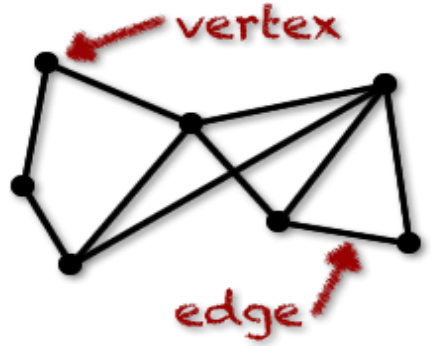
- Biological Networks:

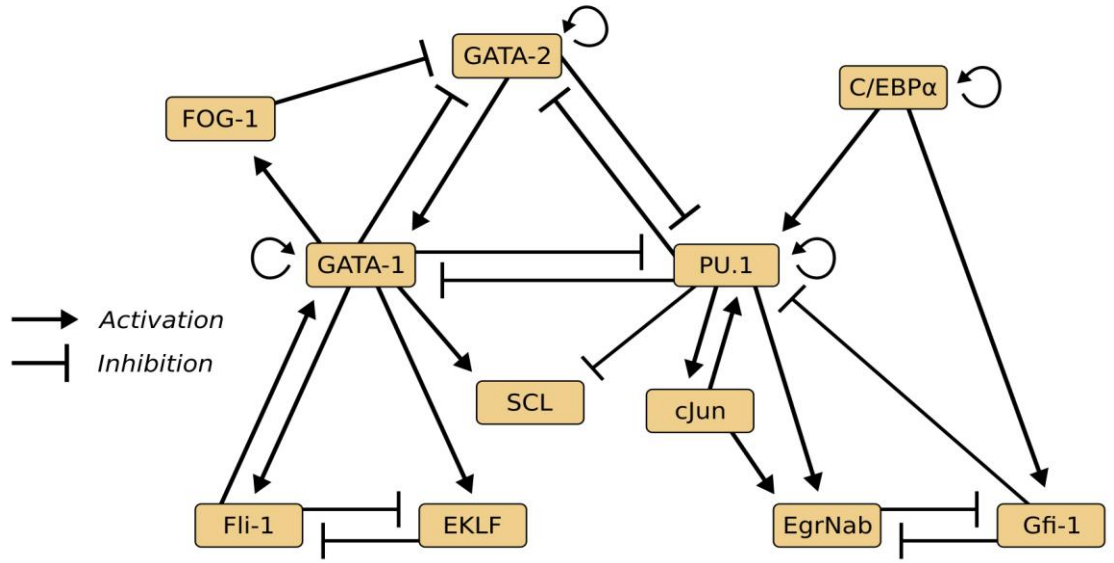


- PPI
- Gene Regulatory
- Metabolic
- DNA-Protein
- Co-expression

Systems Biology Resources

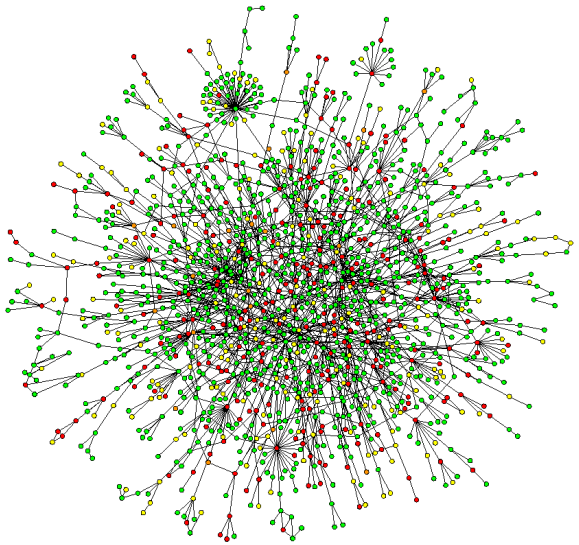
- Biological Networks:
  - Elements (node or vertex) + Relationships (Edges)
  - Modeled by Mathematical Graphs
- Gene regulatory Networks
- Metabolic Networks
- Protein Interaction Networks
- Co-expression Networks

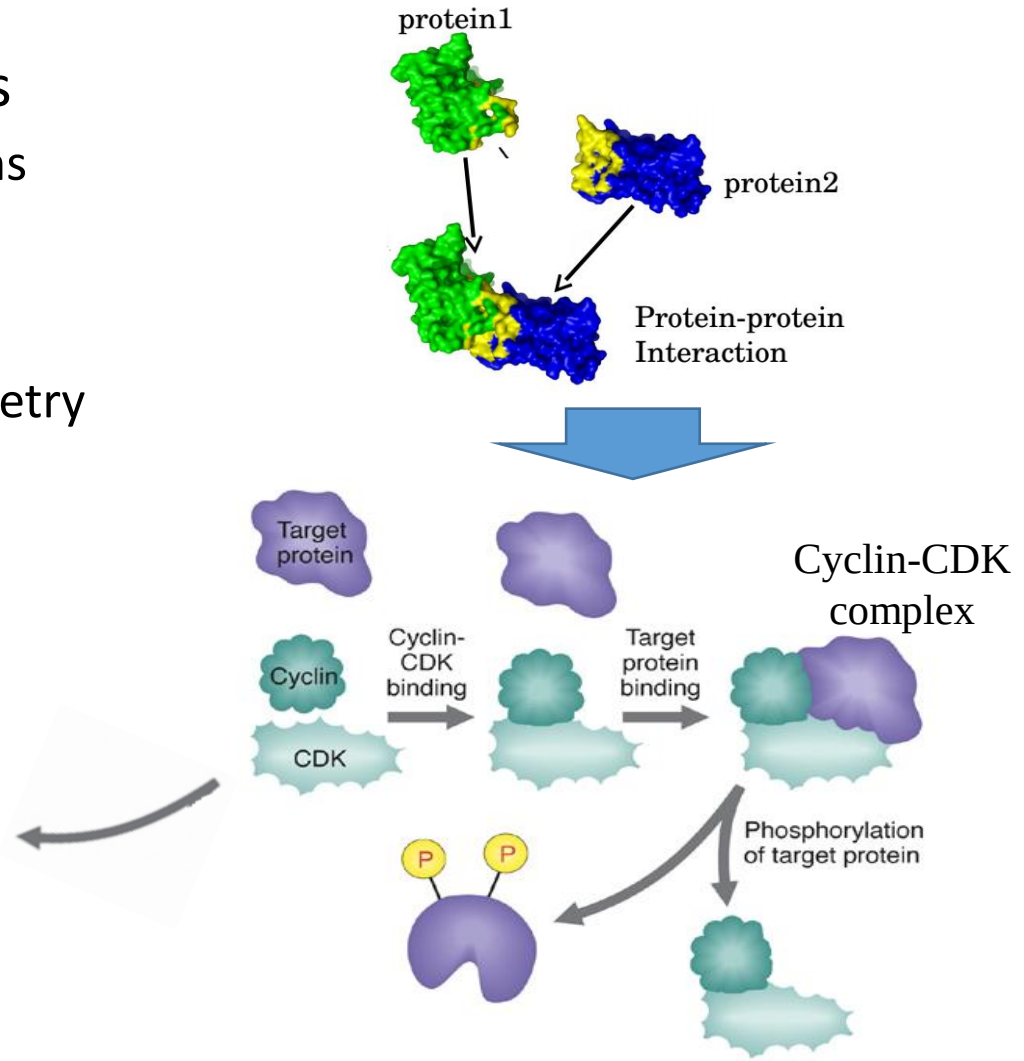




Protein-Protein Interaction (PPI) Networks

- Physical Interaction between Proteins
  - Physical contacts of two or more proteins
- Detection Methods:
  - Yeast-two hybrid
  - Affinity purification with mass spectrometry
  - Text mining



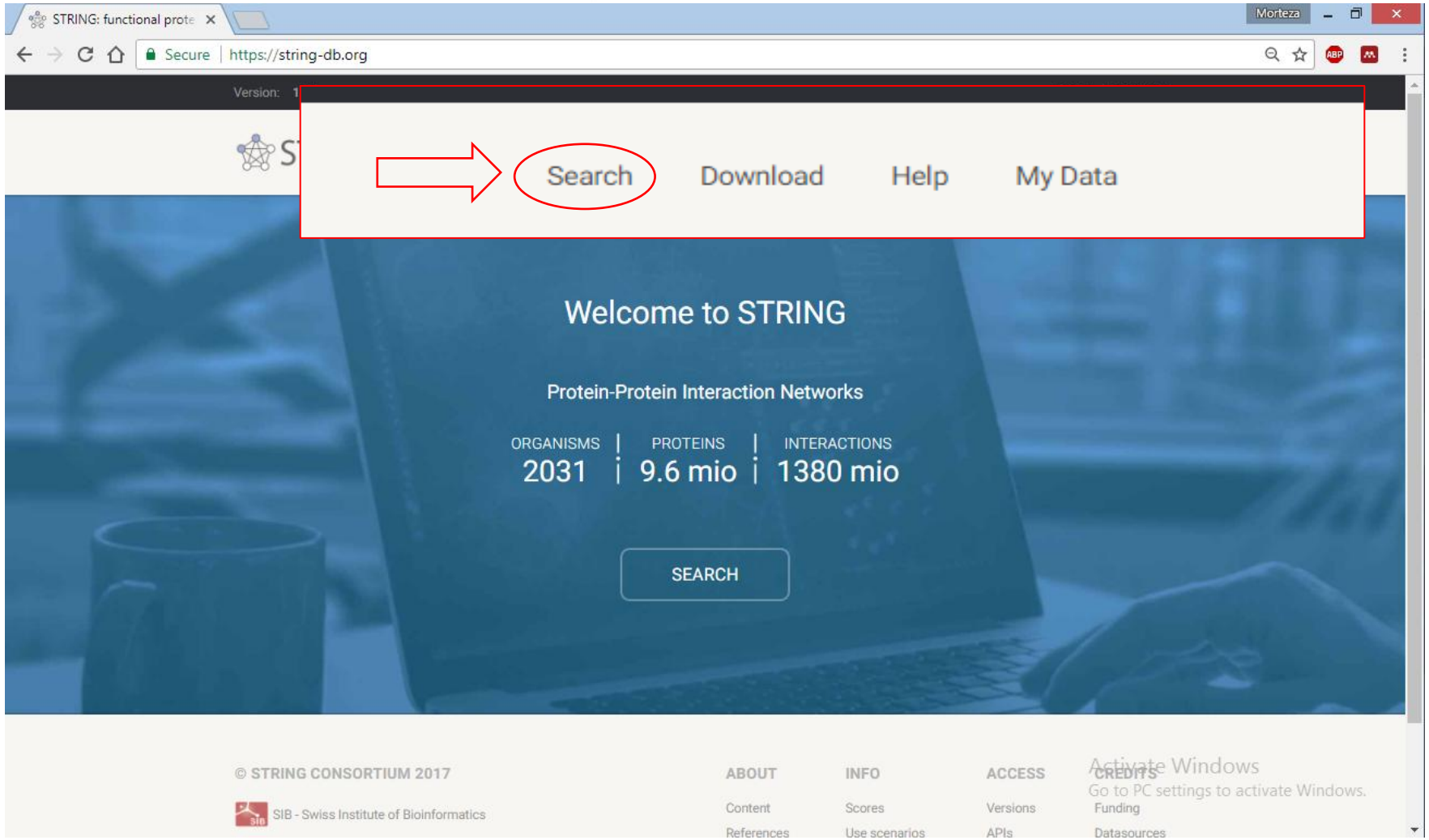


PPI Networks

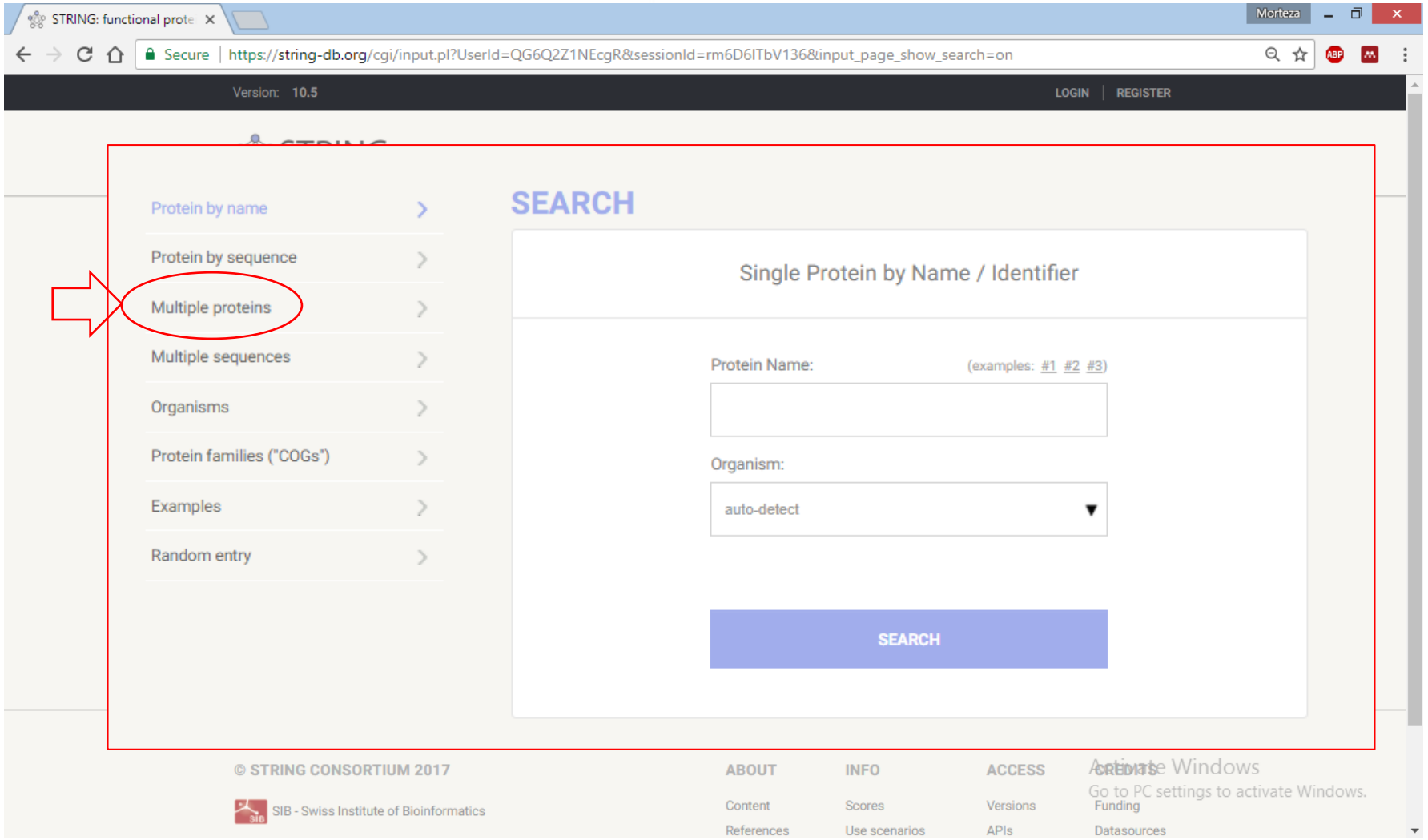
- Databases:

Database name	URL	Latest update
DIP [26]	http://dip.doe-mbi.ucla.edu/dip/	**2016/07/31**
BioGRID [27]	https://thebiogrid.org/	**2016/12/25**
STRING [28]	http://string-db.org/	**2016/04/16**
IntAct [29]	http://www.ebi.ac.uk/intact/	**2016/12/13**
HIPPIE [30]	http://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/index.php	**2016/06/24**

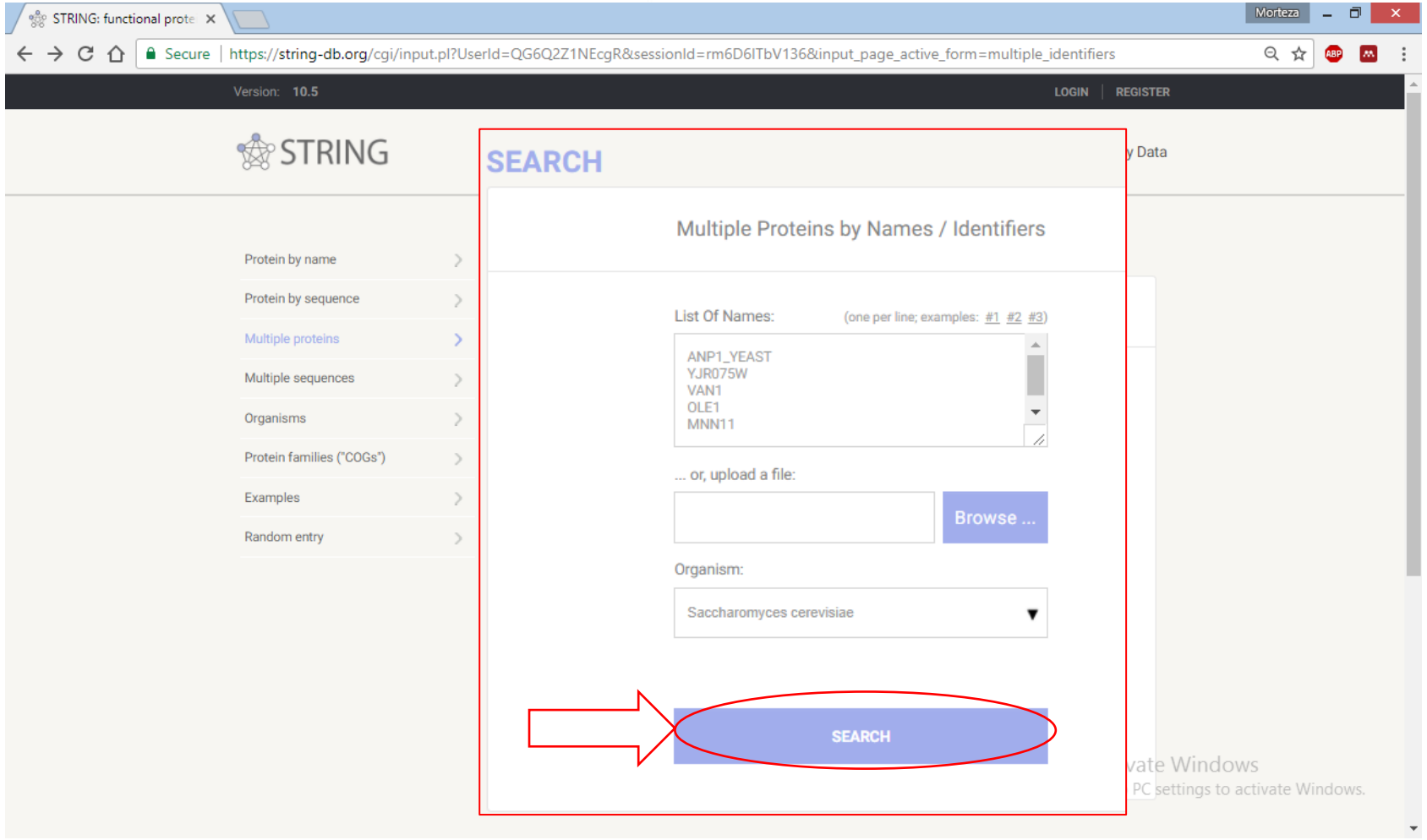
String Database



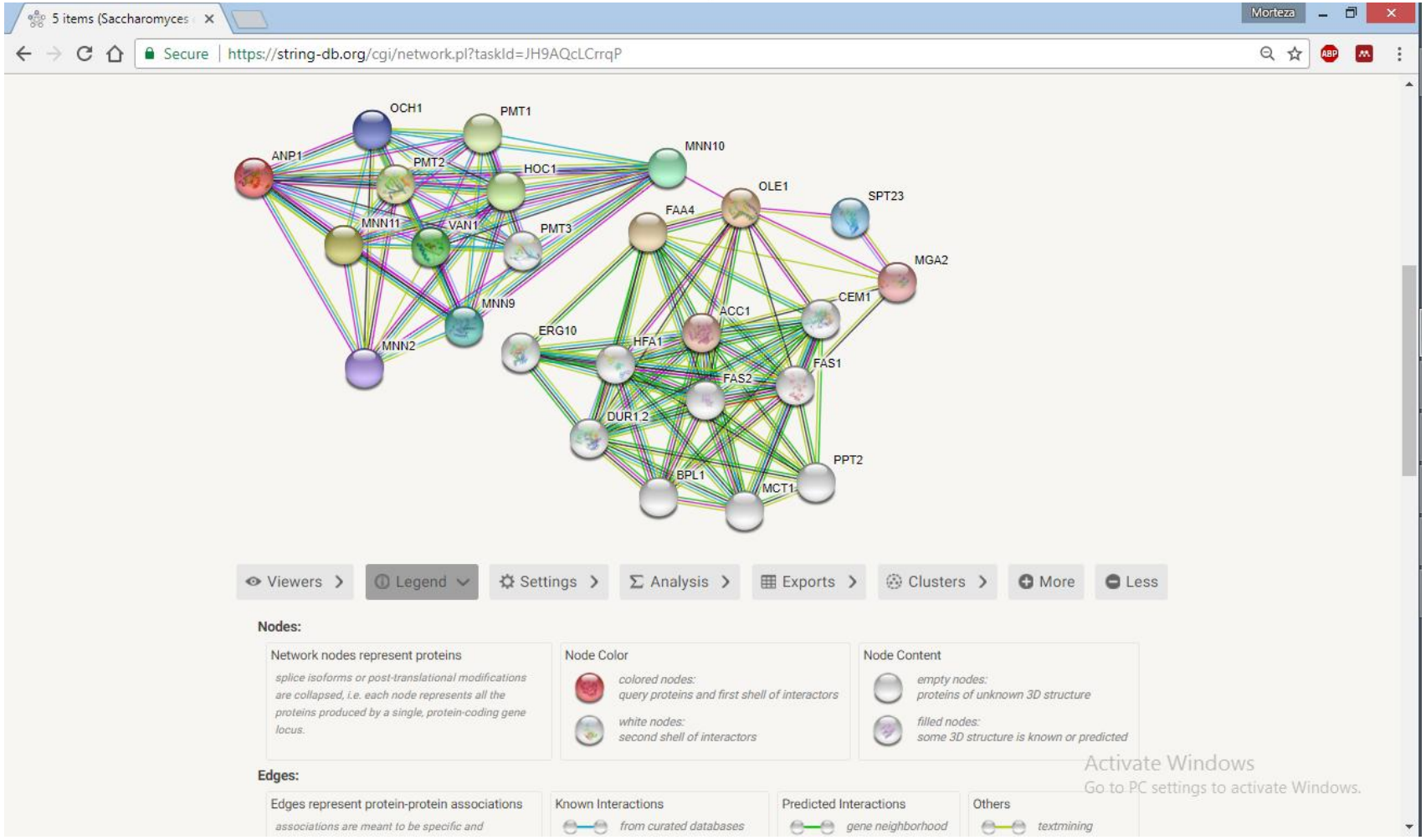
String Database



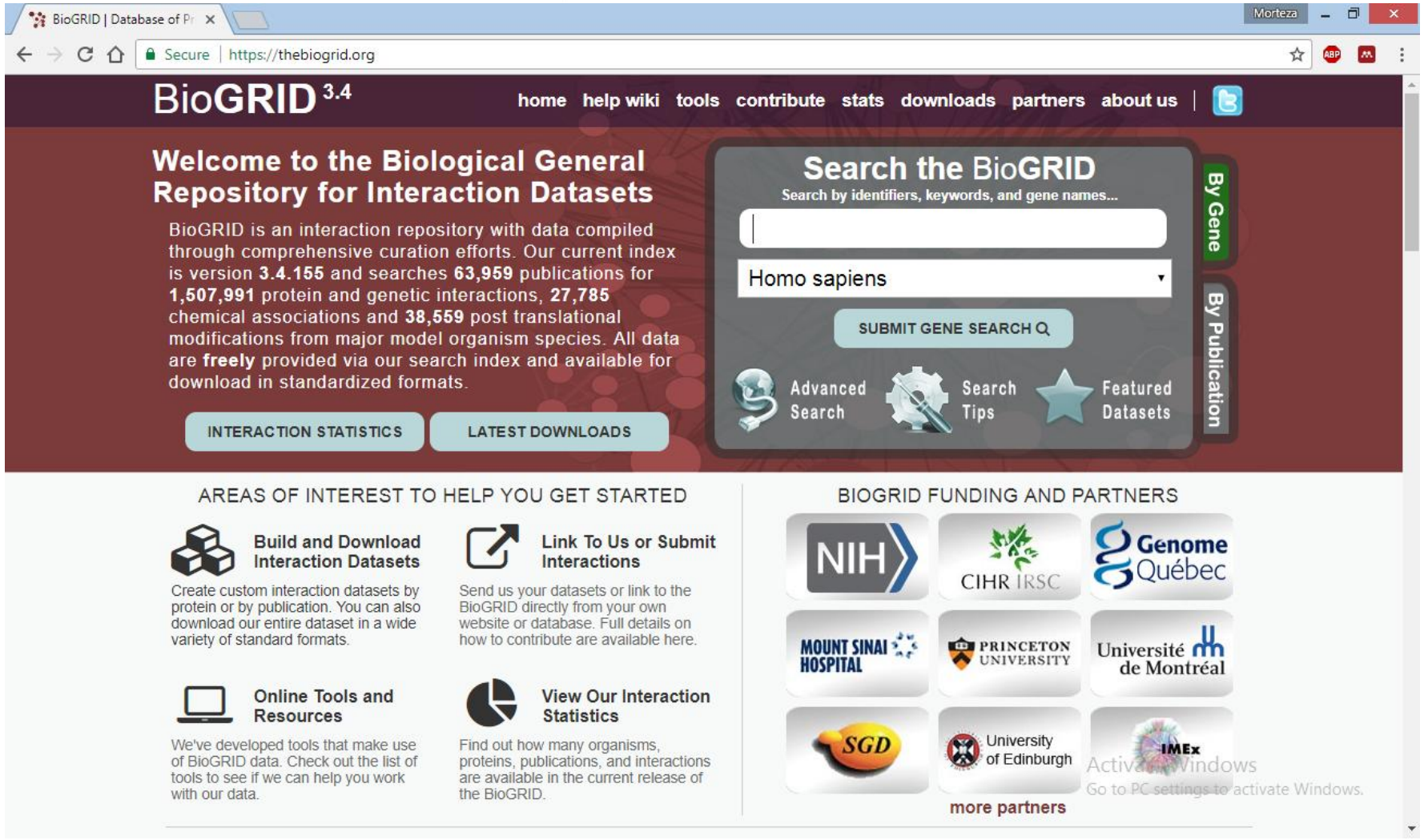
String Database



String Database

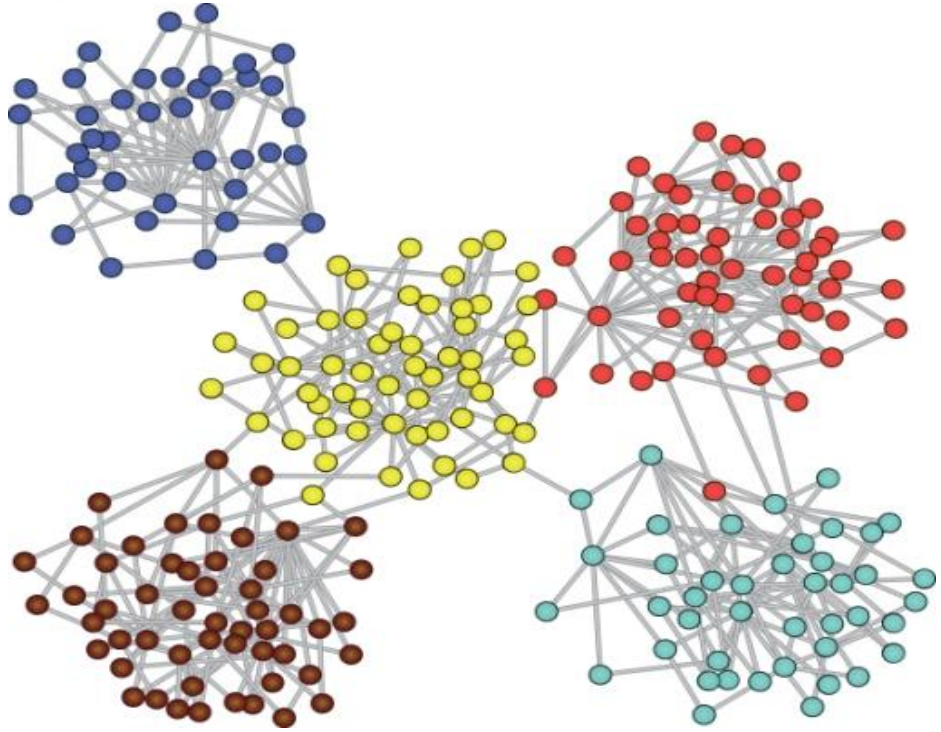


BioGRID Database



PPI Networks Analysis

- Modules or Clusters in PPI:
  - Basic definition: **Dens subnetworks (highly connected)**
  - Biological Definition: **A group of proteins that shared a biological process (Protein Complexes)**



Modules Analysis Workflow

Finding Candidate Disease genes

- Disease gene database

Extract PPI Network

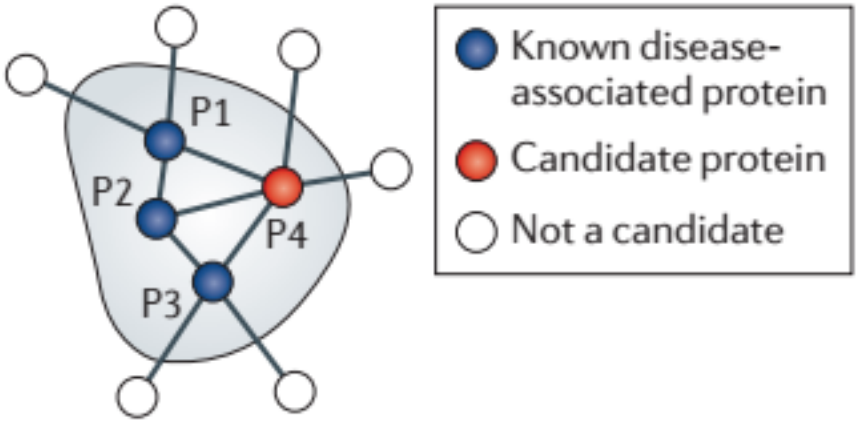
- STRING
- BioGRID
- PPI Database

Module Detection

- Detection Methods
- MCODE
- ClusterOne
- MCL

Module Analysis

- Enrichment Analysis
- New Disease genes
- Disease Pathways
- Drug Targets



Module Detection Methods

- Using Topological Properties of the Network:
  - Finding **Cliques** (Fully Connected Subnetworks): **Cfinder**
  - Finding **Dense Subnetworks**: **MCODE**, **ClusterOne**
  - Simulating **Random Walks** on the Network: **MCL**
- Integrating Biological Information:
  - Integrate **Gene Ontology** Information: **WCOACH**
  - Finding **Active Modules** by Integrating Expression Data: **JactiveModules**

Gene Set Enrichment Analysis

Is a set of genes (or gene products) significantly involved in a biological process or pathway?

Search these genes in a biological process or pathway database and calculate a P-Value

Gene Set Enrichment

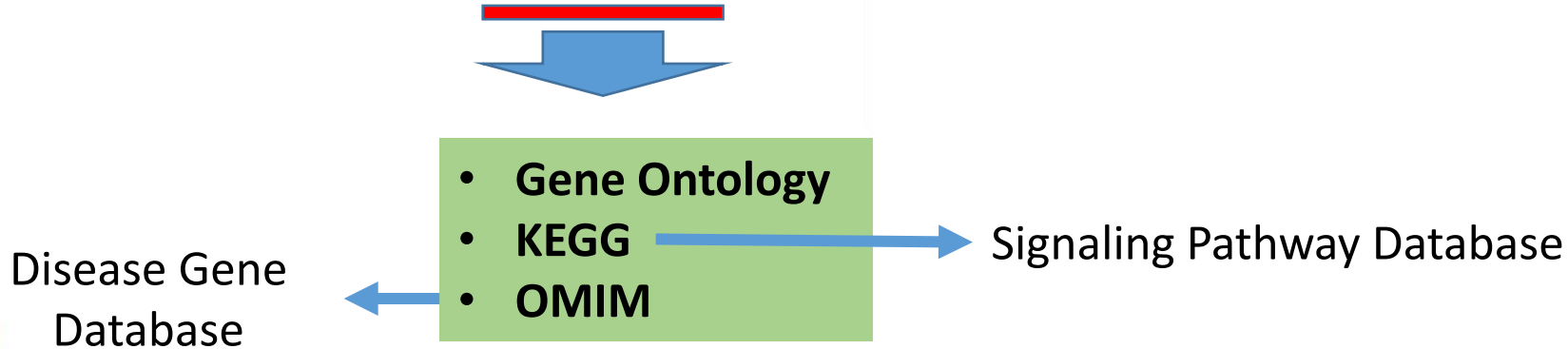
Gene Set Enrichment Analysis

- Input:

  A **set of** genes or **proteins** (A **Module**)

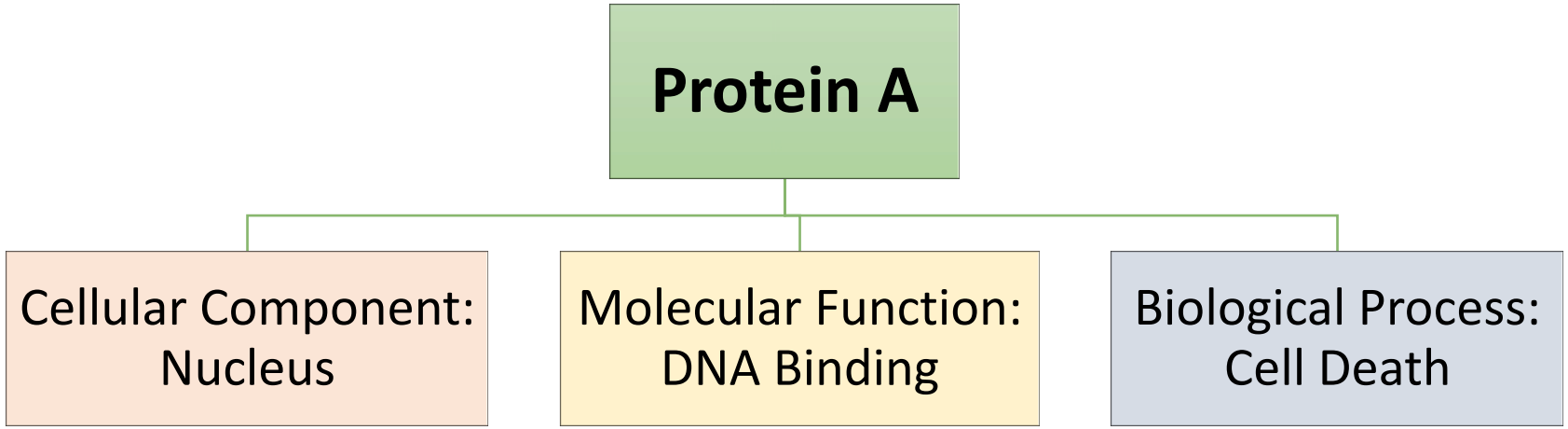
- Output:

  Common Biological **Processes** or **Pathways** or **Diseases** based on a **Database** with P-Value



Gene Ontology Database

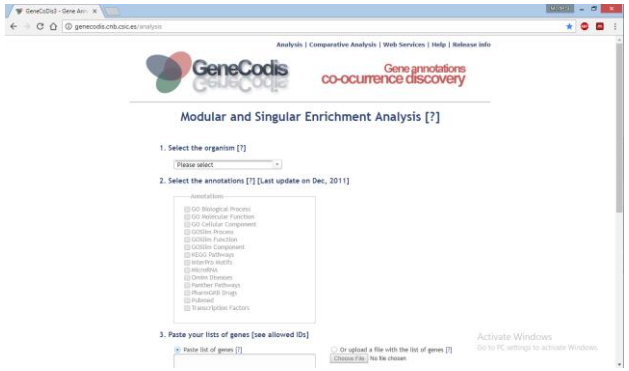
- Answer to three question about a Protein (or Gene):
  - What is the **location** of a protein in the cell? **Cellular Component (CC)** of a protein
  - What is the **Function** of a protein? **Molecular Function (MF)** of a Protein
  - What **Process** does a Protein involve? **Biological Process (BP)** of a Protein

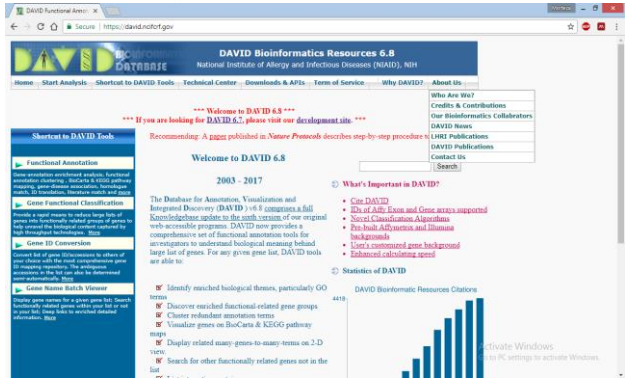


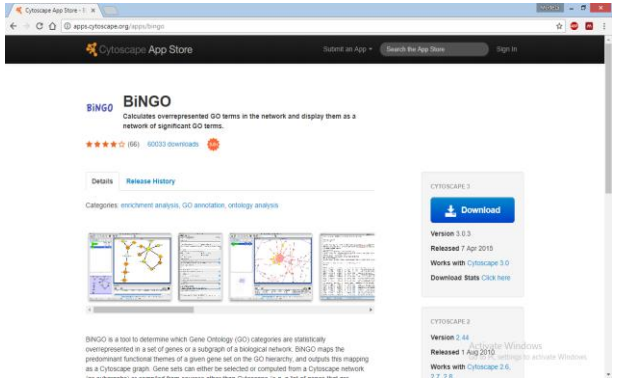
Enrichment Analysis Tools

- Online Tools
  - DAVID: https://david.ncifcrf.gov/
  - GeneCodis: http://genecodis.cnb.csic.es/analysis

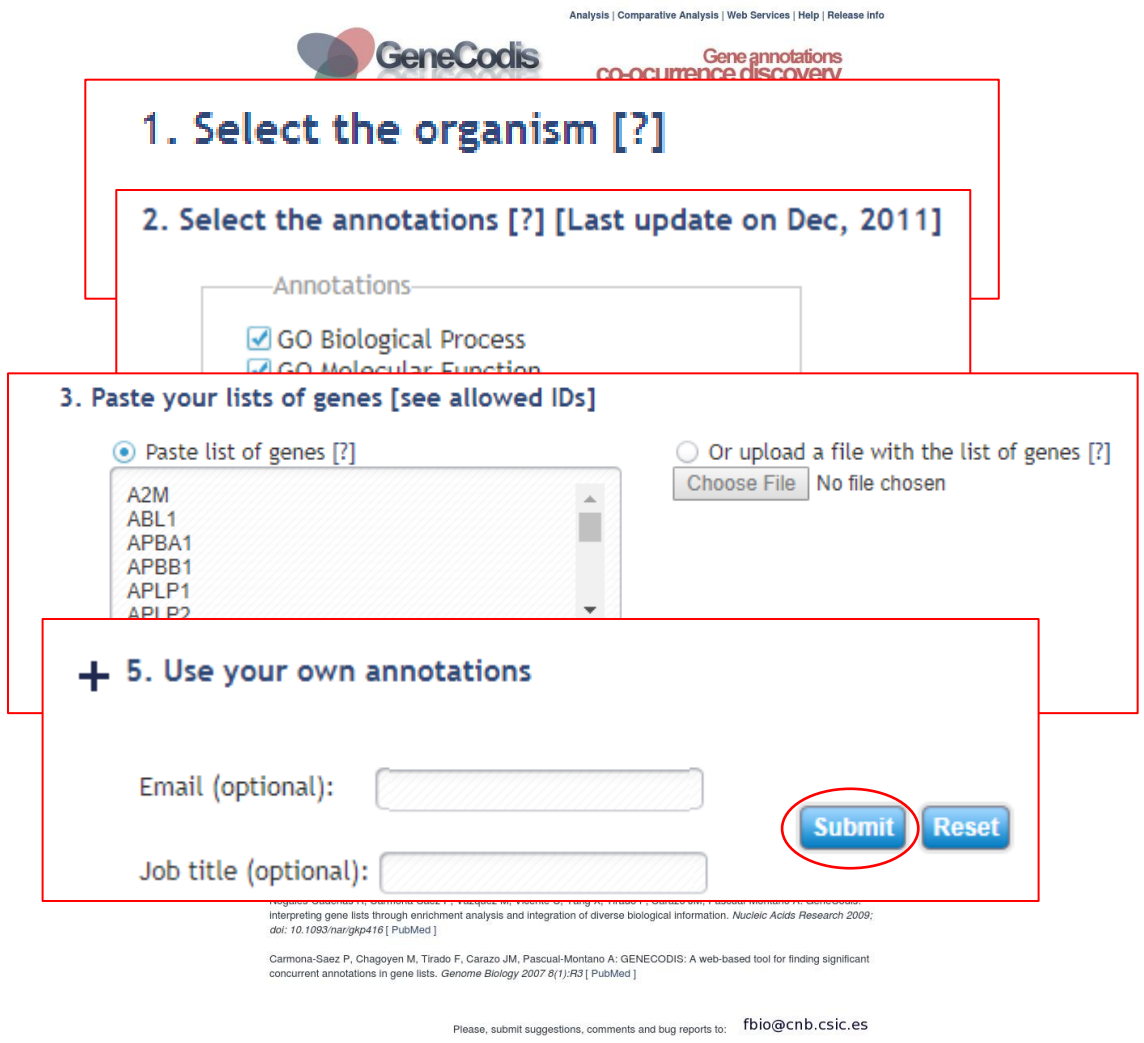
- Offline Tools:
  - BiNGO: http://apps.cytoscape.org/apps/bingo



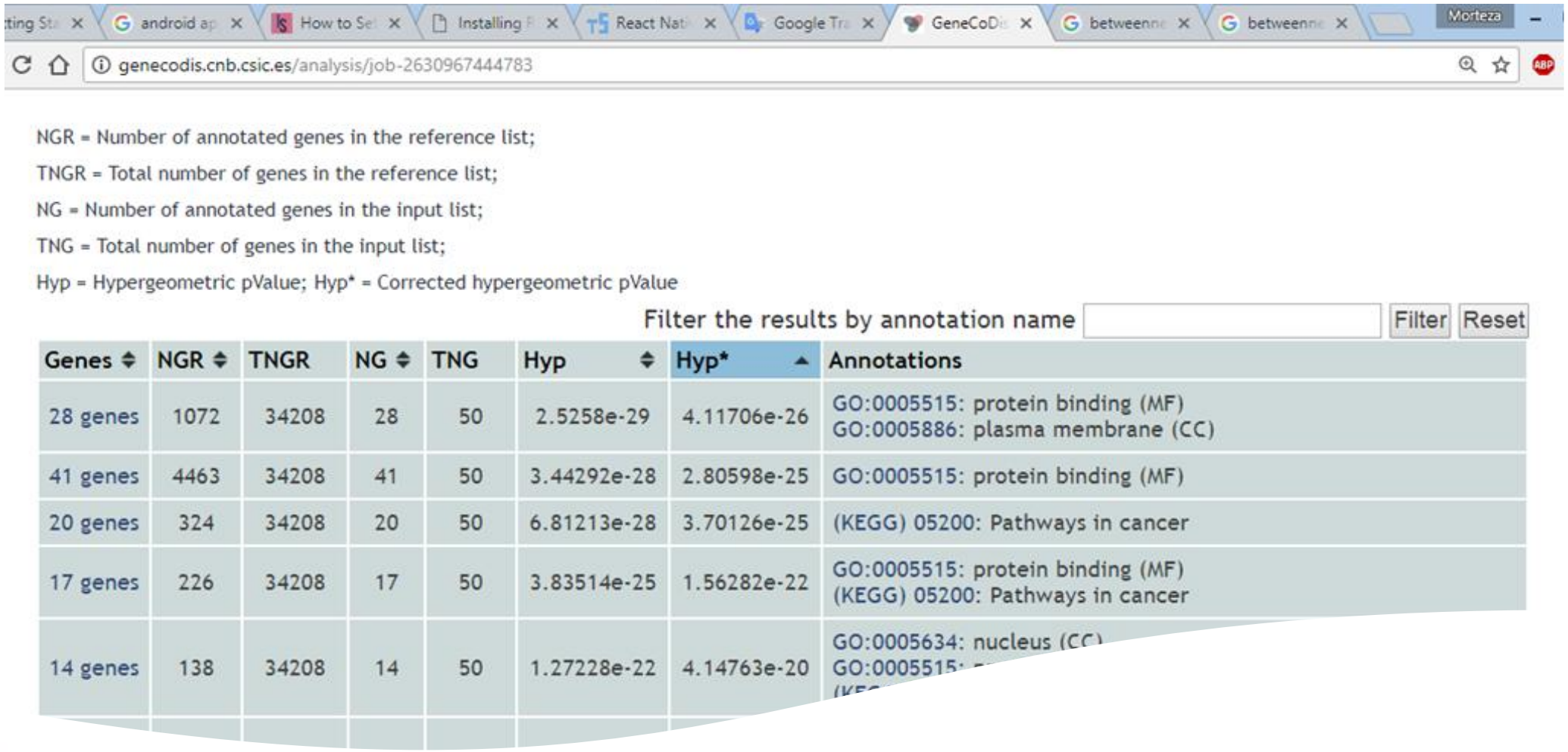




Enrichment Analysis Tools

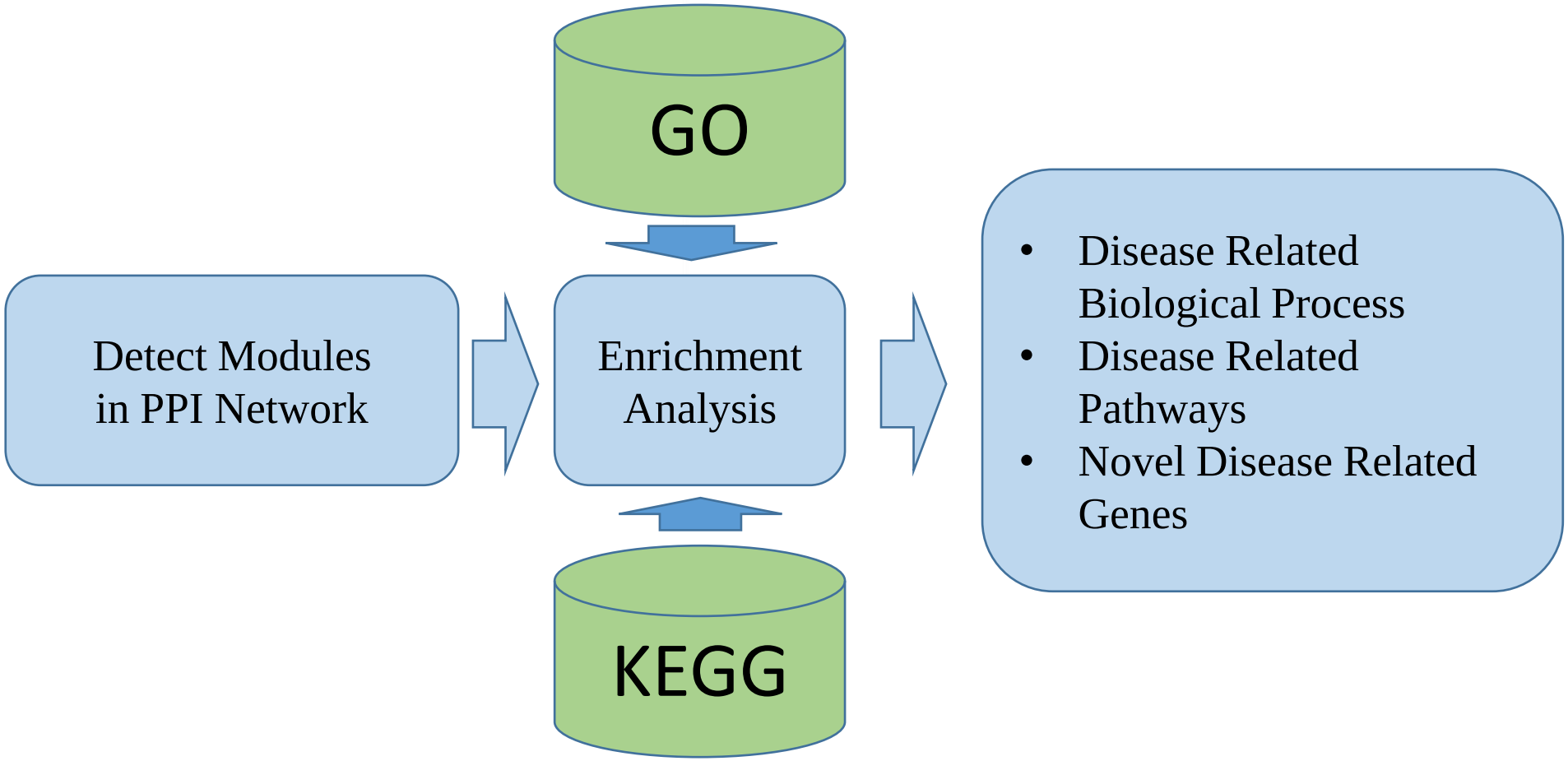


Enrichment Analysis Tools



Genes	NGR	TNGR	NG	TNG	Hyp	Hyp*	Annotations
28 genes	1072	34208	28	50	2.5258e-29	4.11706e-26	GO:0005515: protein binding (MF) GO:0005886: plasma membrane (CC)
41 genes	4463	34208	41	50	3.44292e-28	2.80598e-25	GO:0005515: protein binding (MF)
20 genes	324	34208	20	50	6.81213e-28	3.70126e-25	(KEGG) 05200: Pathways in cancer
17 genes	226	34208	17	50	3.83514e-25	1.56282e-22	GO:0005515: protein binding (MF) (KEGG) 05200: Pathways in cancer
14 genes	138	34208	14	50	1.27228e-22	4.14763e-20	GO:0005634: nucleus (CC) GO:0005515:

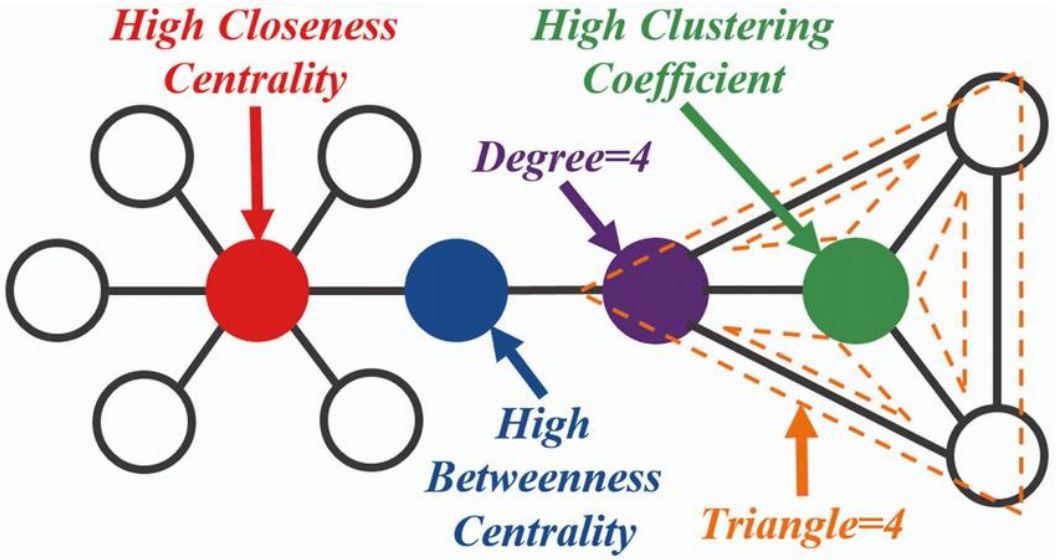
PPI Network Module Analysis



Centrality Measures and Finding Key Proteins

- Centrality Measure: A Metric to **rank nodes** based on specific topological property in a network
  - Degree Centrality: Finding Hub Proteins
  - Betweenness Centrality: Finding Between Cluster Proteins
  - Clustering Coefficient

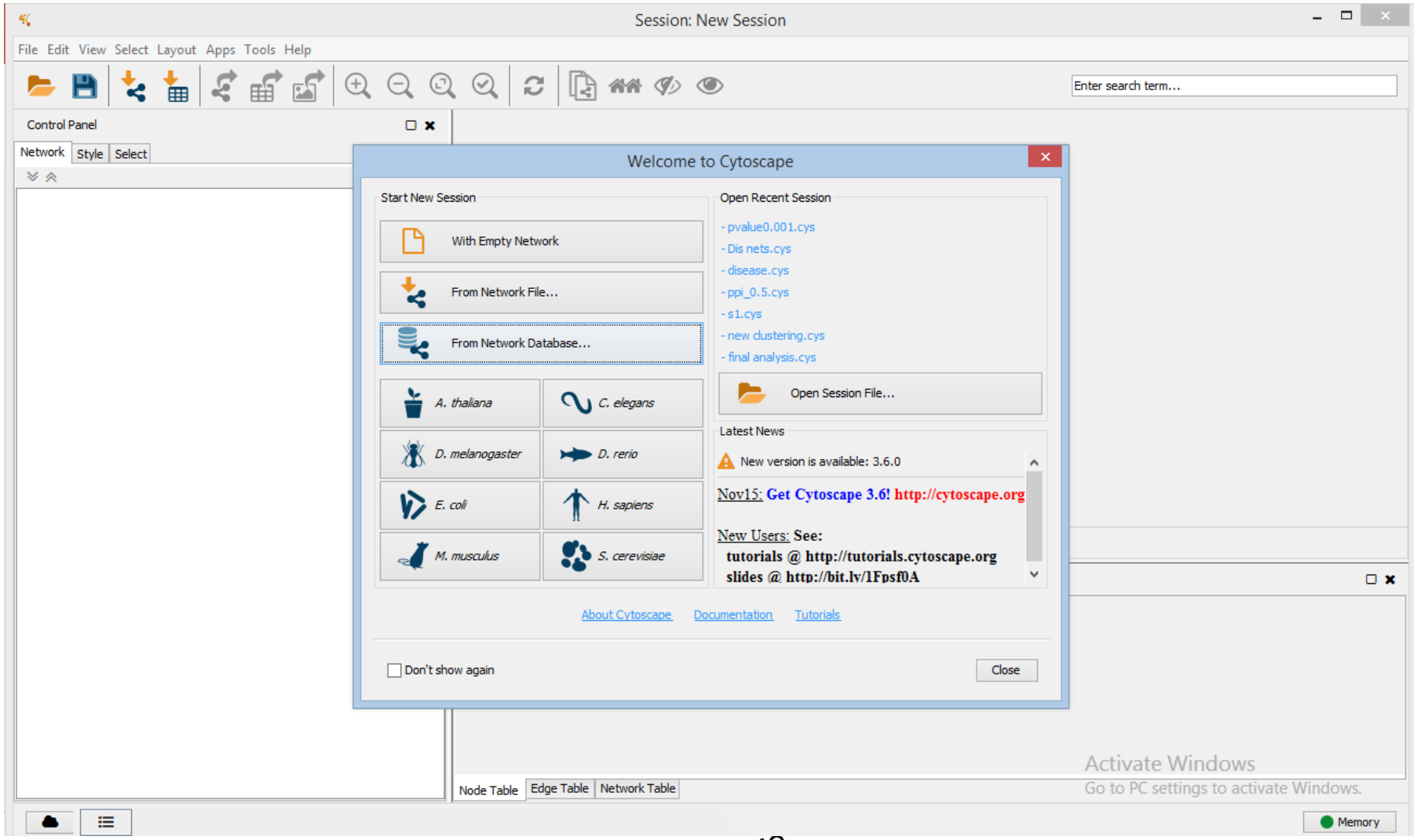
Key Proteins can be introduced as **Drug Targets**



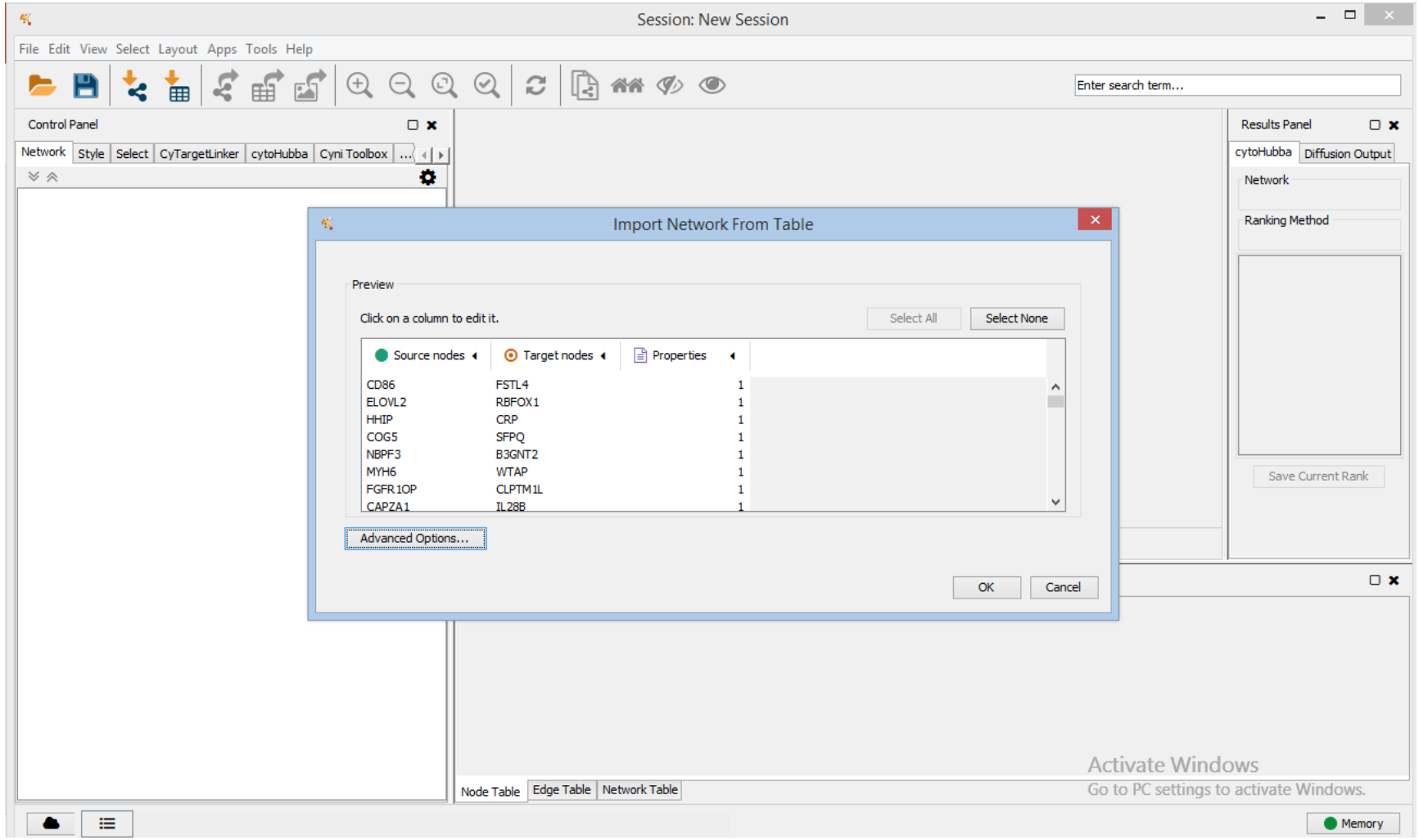
Cytoscape: A Powerful Framework for Systems Biology Analysis

- Installing Requirement:
  - Java
- Easy Installation on any OS
- Find and add many Apps to Analysis Biological Networks

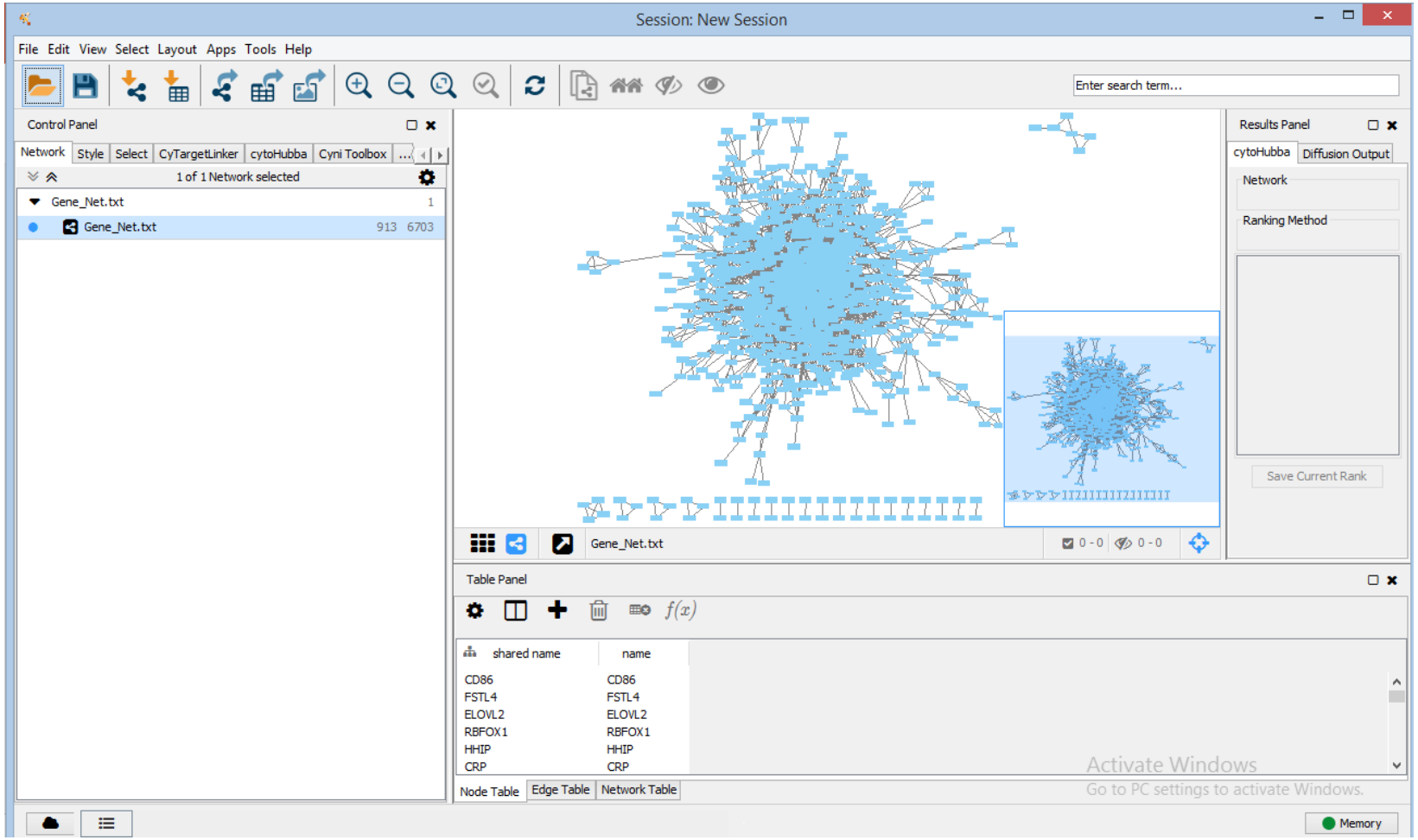
Cytoscape: Welcome Screen



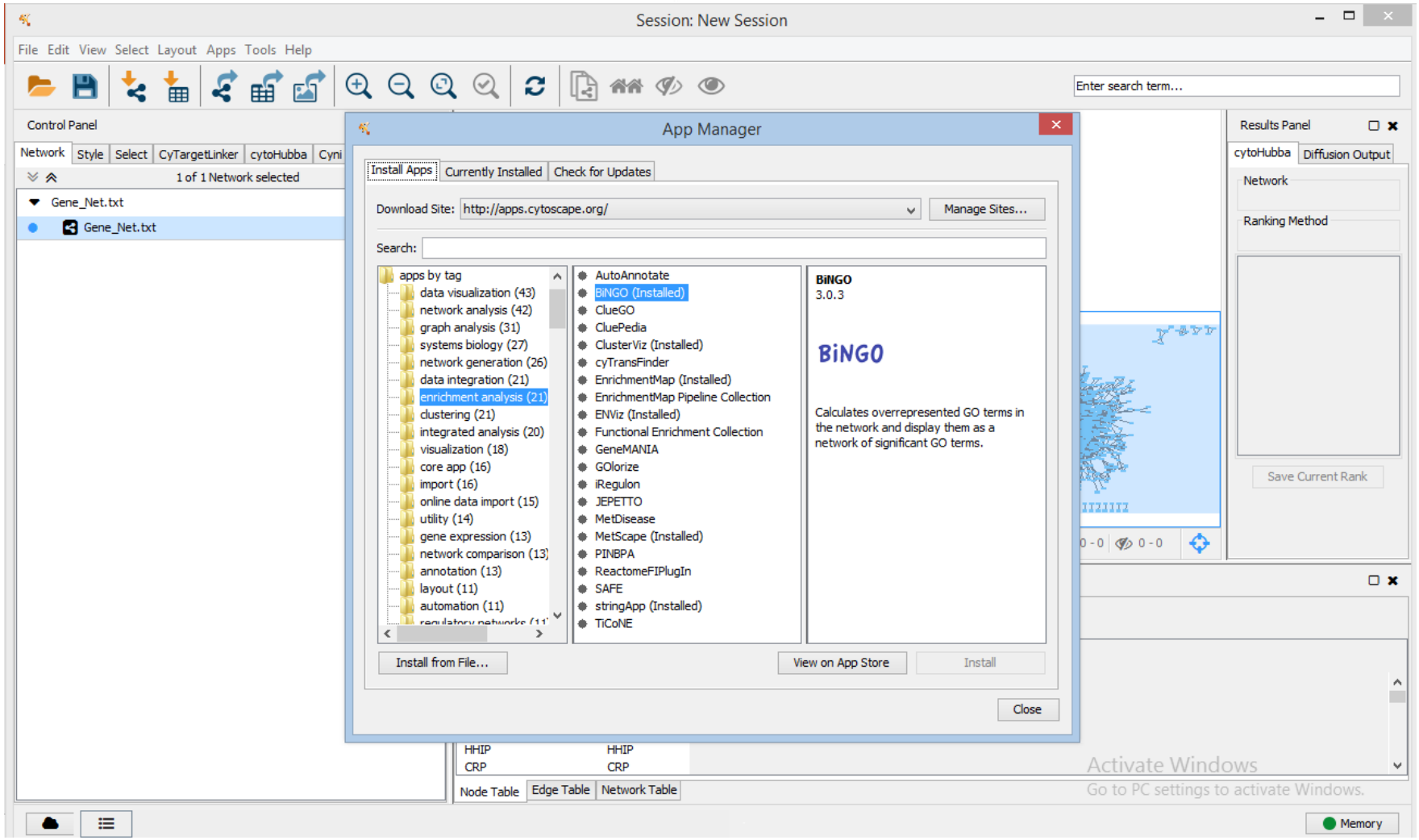
Cytoscape: Import Network from File



Cytoscape: View Network



Cytoscape: App Manager



Thank you for your attention