

Computational analysis of pathways and interaction networks

Gary Bader

Terrence Donnelly Center for Cellular and Biomolecular Research
(DCCBR)

BBDMR, SLRI

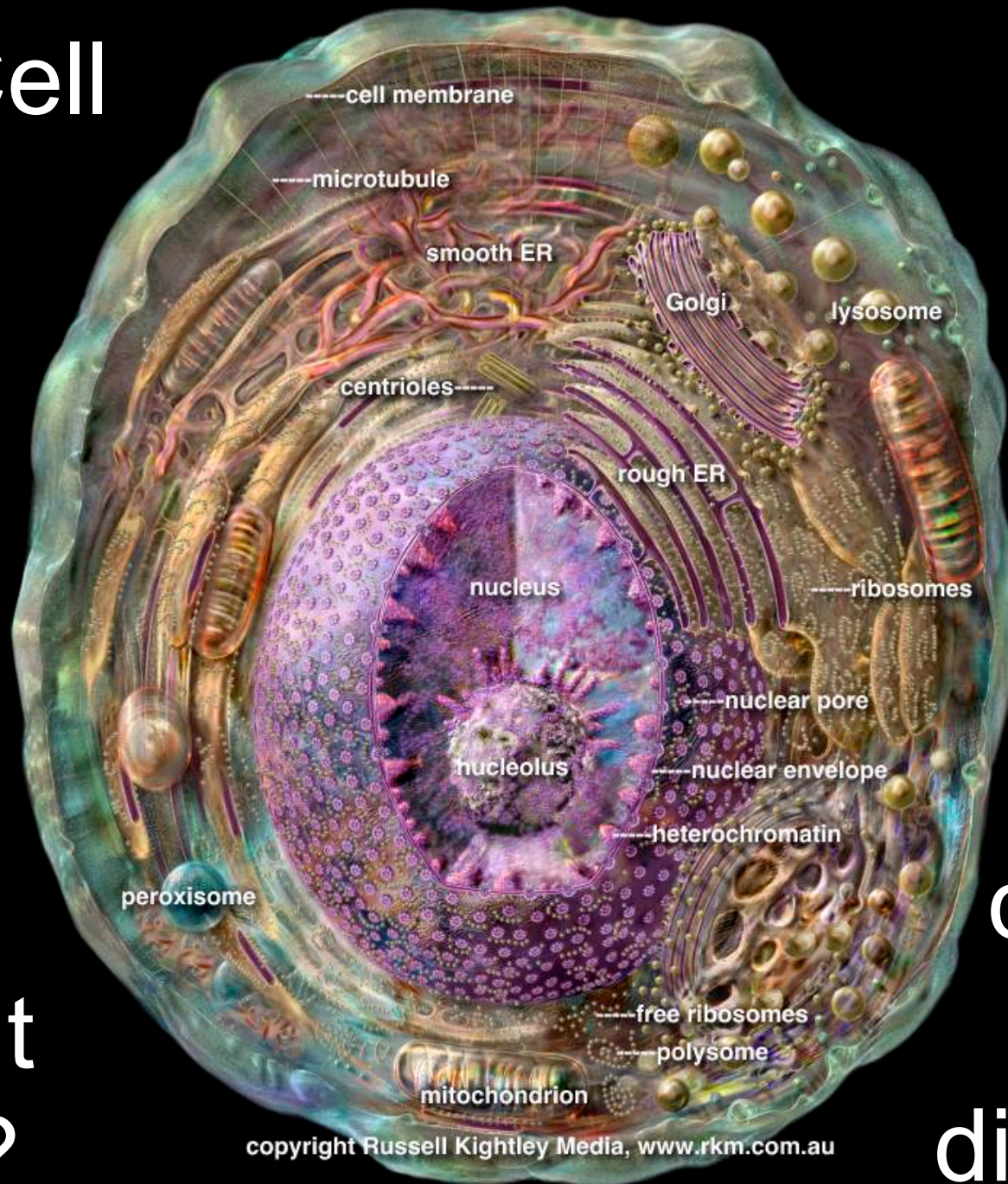
University of Toronto

JTB2010H

<http://baderlab.org>



The Cell



How
does it
work?

How
does it
fail in
disease?

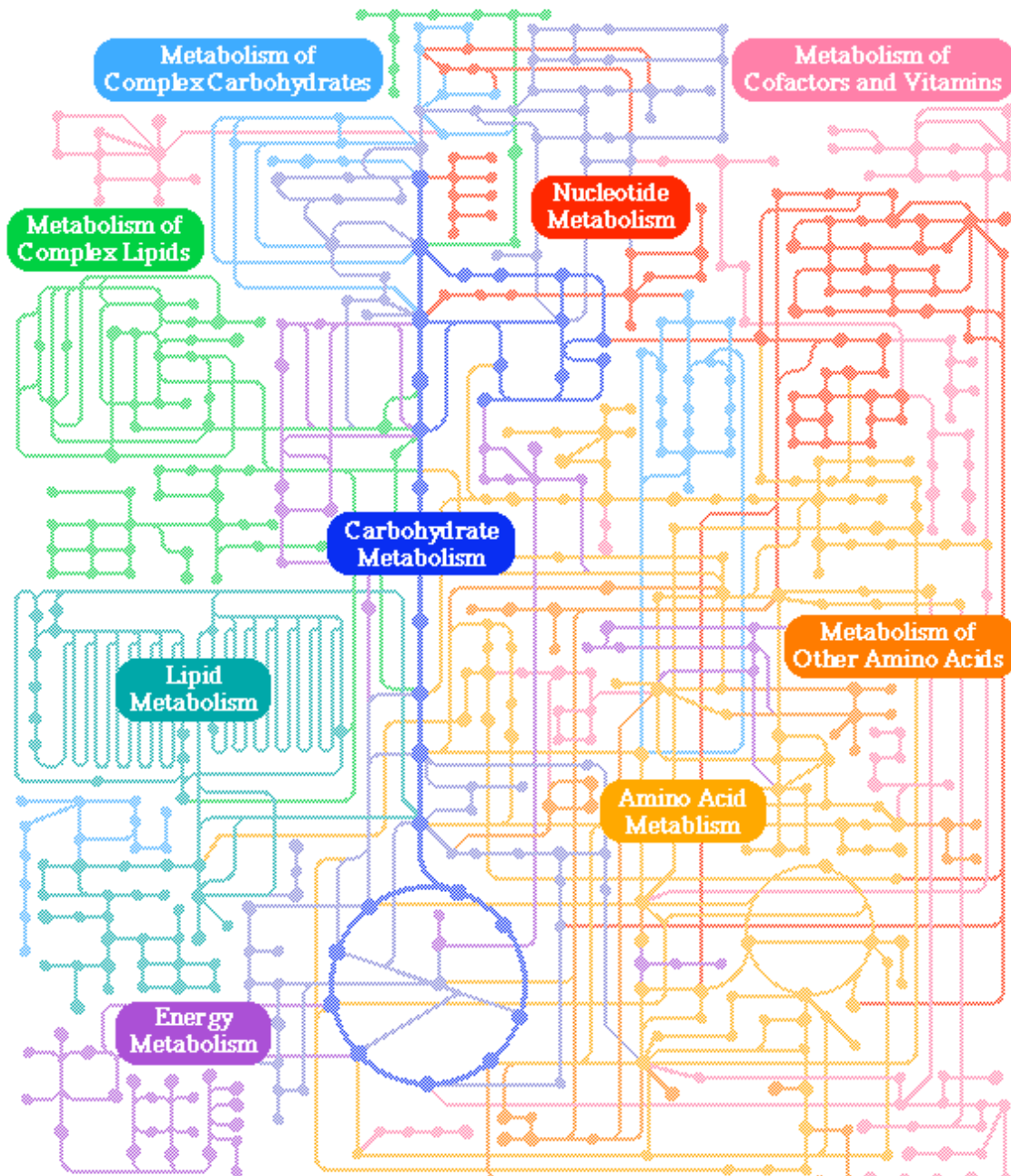
Pathways?

- Pathways are biological processes
- But, not really pathways → networks
- Metabolic, signaling, regulatory and genetic
- Define gene function at many different levels
- Human constructs to organize the cell

Figure S3: View of the entire HMS-PCI dataset.
Thick blue lines represent literature-derived interactions from PsBIND+MIPS in the HMS-PCI dataset.
Thin orange lines represent experimental novel interactions from the HMS-PCI dataset.

Note: generally out of cell context

METABOLIC PATHWAYS



Signaling Pathway

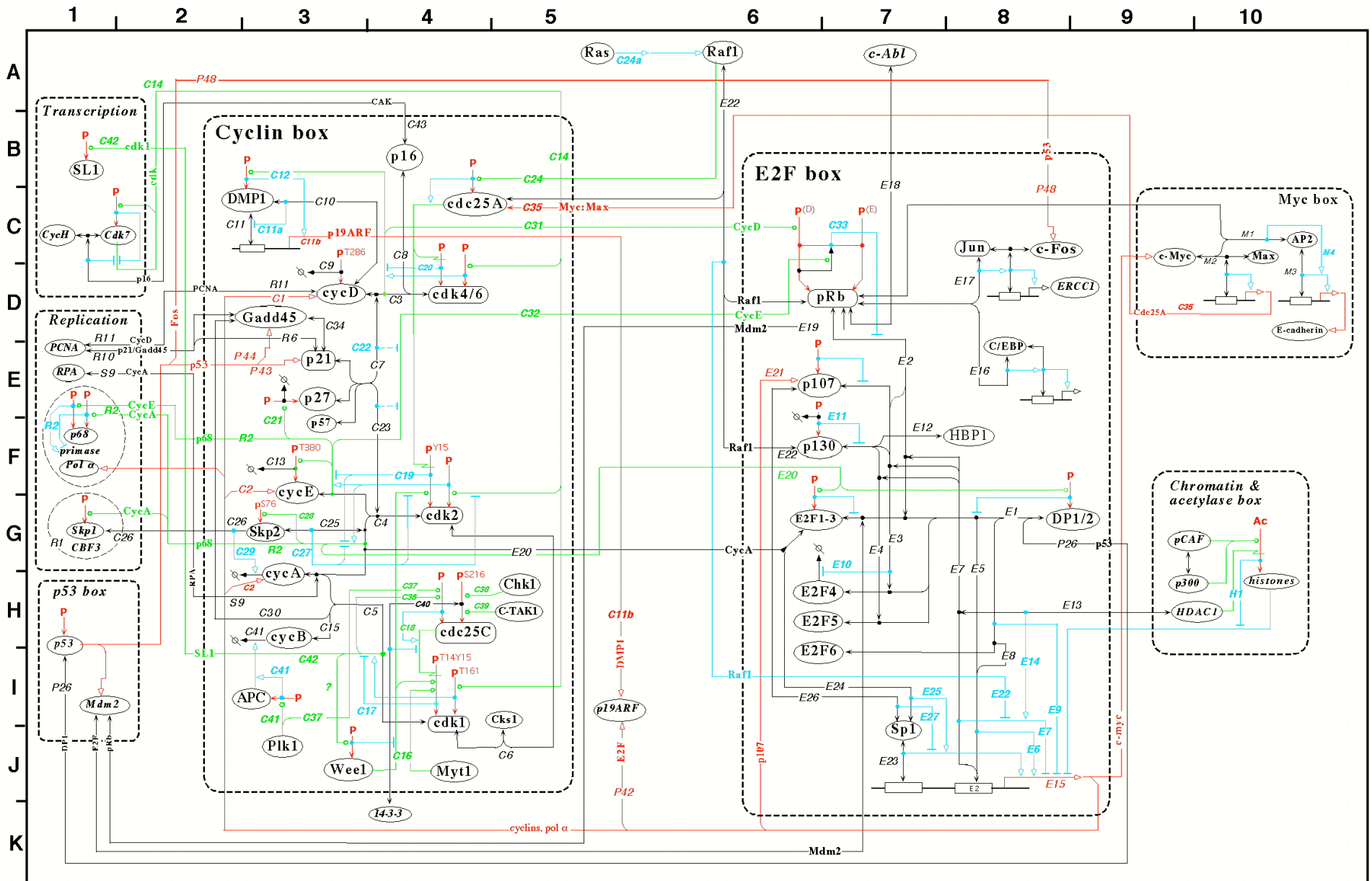
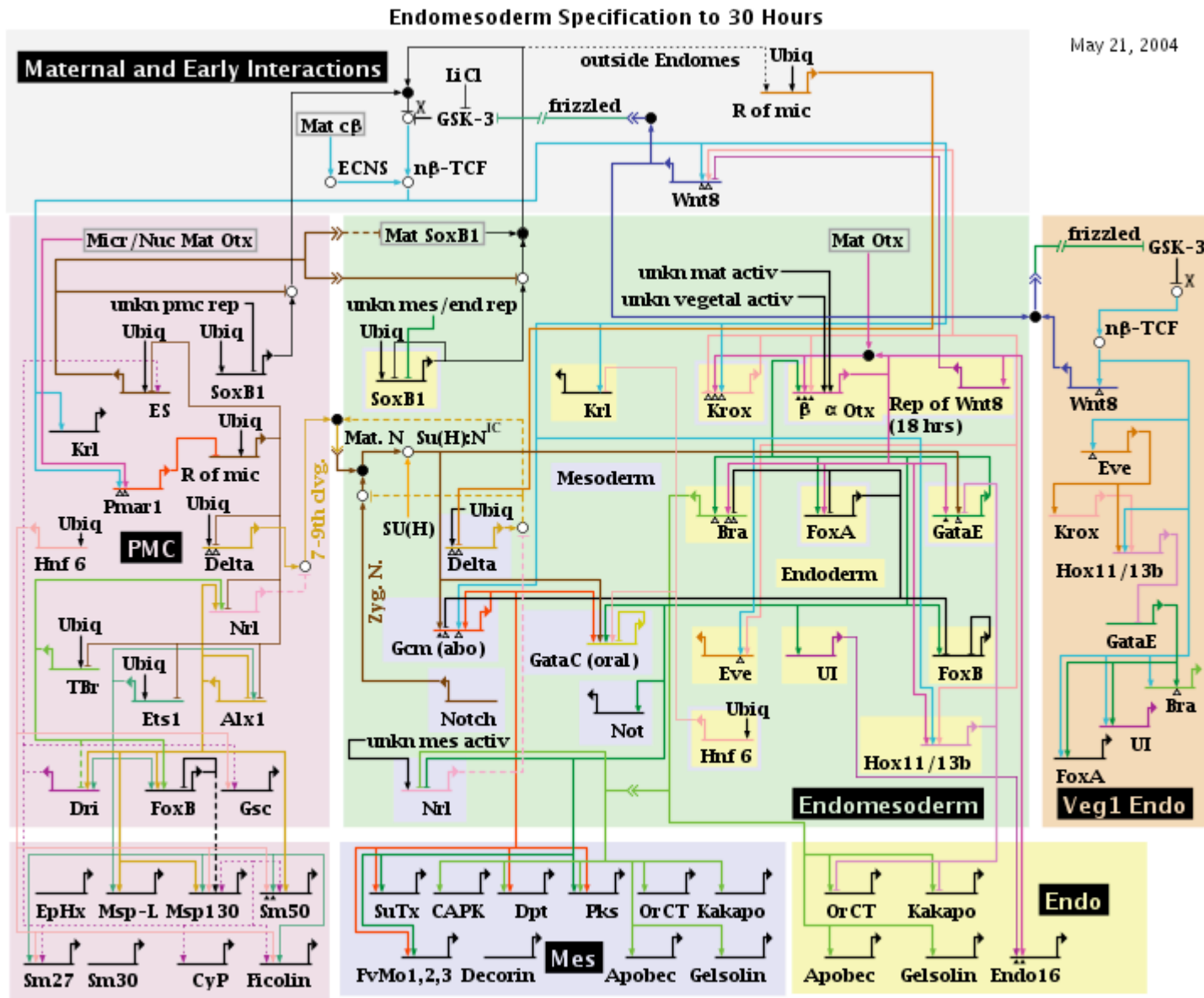
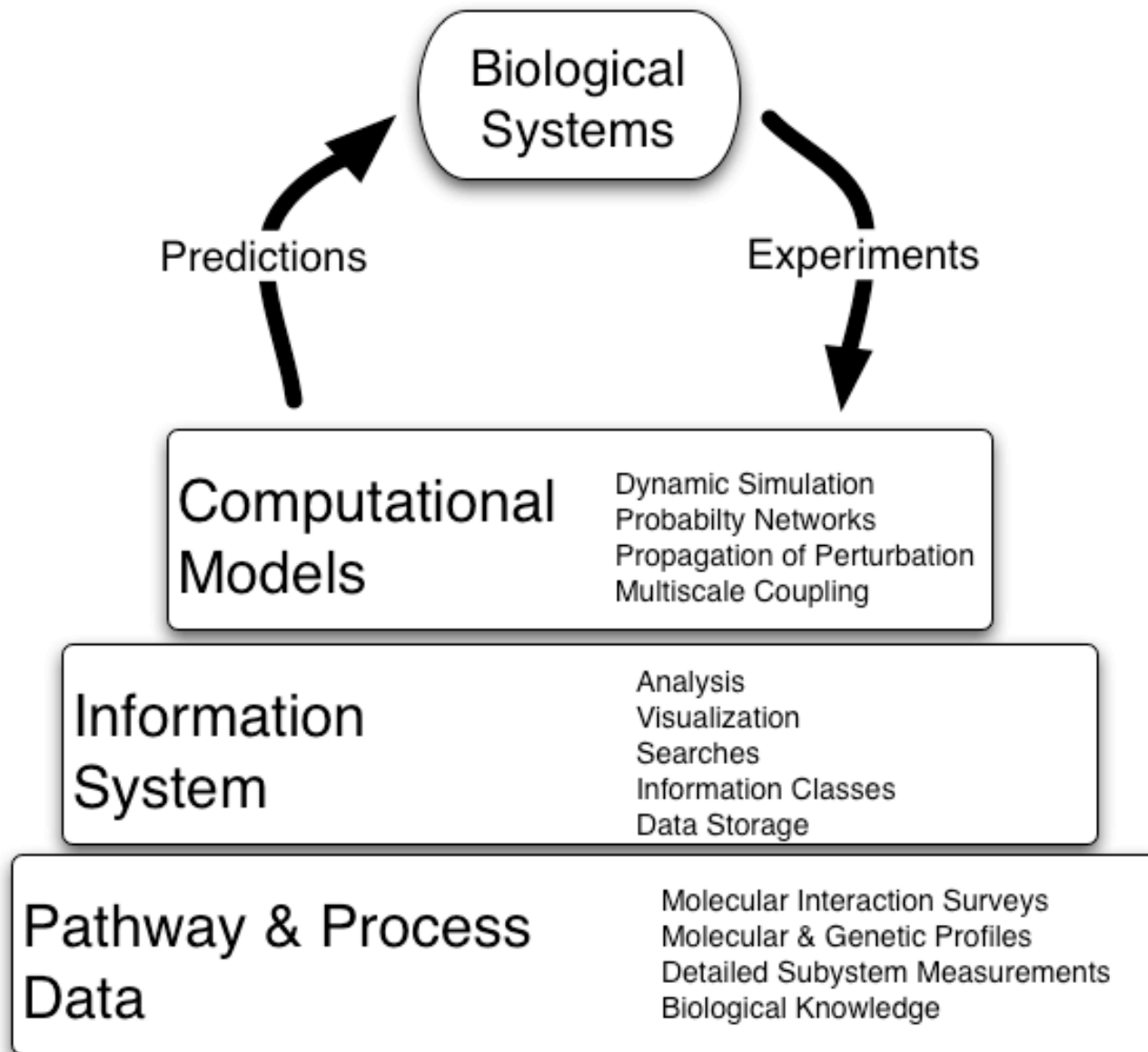


Figure 6A: The Cyclin - E2F cell cycle control system (version 3a - June 8, 1999) http://discover.nci.nih.gov/kohnk/interaction_maps.html

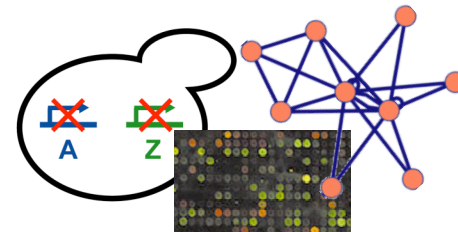
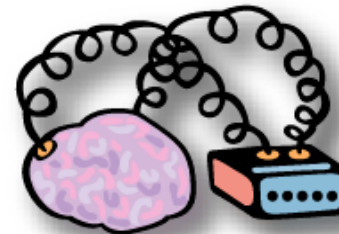
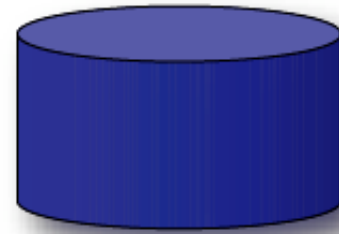
Gene Regulatory Pathway





Pathway Information

- Databases
 - Fully electronic
 - Easily computer readable
- Literature
 - Increasingly electronic
 - Human readable
- Biologist's brains
 - Richest data source
 - Limited bandwidth access
- Experiments
 - Basis for models



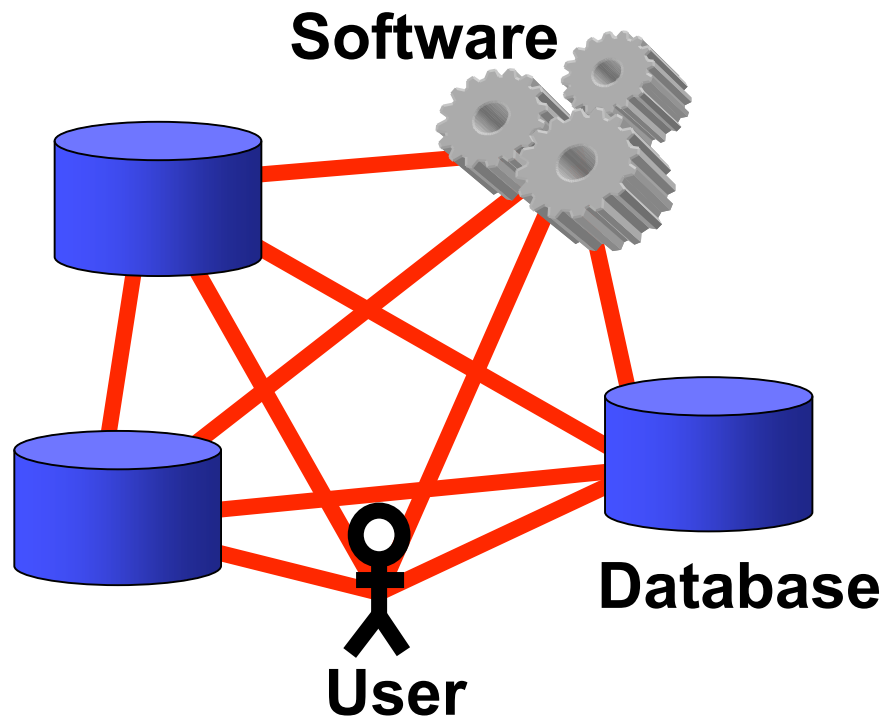
Pathway Databases

<p>Metabolic Pathways aMAZE - Protein Function and Biochemical Pathways Project BioCyc - BioCyc Knowledge Library BioModels - BioModels Database BioSilico - BioSilico BRENDA - Comprehensive Enzyme Information System CellML Repository - CellML Model Repository EcoCyc - Encyclopedia of E. coli Genes and Metabolism EMP - Enzymes and Metabolic Pathways Database ENZYME - Enzyme nomenclature database GeneNet - Genetic Networks GenMAPP - Gene MicroArray Pathway Profiler GOLDdb - Genomics of Lipid-associated Disorders Indigo - Gene Neighborhoods and Codon Usage IntEnz - Integrated relational Enzyme database KEGG - Kyoto Encyclopedia of Genes and Genomes LIGAND - Database of Chemical Compounds and Reactions in Biological Pathways Malaria - Malaria Parasite Metabolic Pathways MetaCore - MetaCore pathway database MetaCyc - Metabolic Pathway Database MIPS CYGD - MIPS Comprehensive Yeast Genome Database MorMod - Modeler's Notebook and Datasore MPB - Metabolic Pathways of Biochemistry MRAD - Metabolic Reaction Analysis Database NetBiochem - Medical Biochemistry Resource Nicholson Minimaps - IJLMB-Nicholson Minimaps PathArt - Pathway Articulor PathDB - Pathways Database PATKA - Pathway Analysis Tool for Integration and Knowledge Acquisition PDS - Pathways Database System PharmGKB - The Pharmacogenetics and Pharmacogenomics Knowledge Base ProcessDB - ProcessDB Reactome - Reactome KnowledgeBase SBML Model Repository - SBML Model Repository SoyBase - USDA-ARS Soybean Genetics and Genome Database TECR - Thermodynamics of Enzyme-Catalyzed Reactions UM-BBD - Biocatalysis/Biodegradation Database WIT - What Is There?</p>	<p>BRITe - Biomolecular Relations in Information Transmission and Expression CTDB - Calmodulin Target Database DDIB - Database of Domain Interactions and Bindings DIP - Database of Interacting Proteins DIP - Database of oligomerization domains from lambdaDB experiments DopaNet - DopaNet DRAC - Database of Ribosomal Crosslinks DSM - Dynamic Signaling Maps FIMM - Functional Molecular Immunology FlyNets - FlyNets FusionDB - Prokaryote Gene Fusion Events GPCR-PD - G-protein-coupled receptors protein database GRID - General Repository for Interaction Datasets GroEL_PPI - Proteins that interact with GroEL and factors that affect their release HIV-1 at NCBI - HIV-1, Human Protein Interaction Database at NCBI HIVIND - HIV Molecular Immunology Database HP-DPI - Helicobacter pylori Database of Protein Interactions HPID - Human Protein Interaction Database HPRD - Human Protein Reference Database HSP1 PPI - Protein-Protein Interactions Table for Human herpesvirus HumanPSD - Human Proteome Survey Database ICBS - Inter-Chain Beta-Sheets IntAct - IntAct INTERACT - INTERACT Protein-protein interaction database InterDom - Database of Interacting Domains Interfaces - DATASET OF PROTEIN-PROTEIN INTERFACES Interolog - Interolog/Regulog Database JenPeP - JenPeP Peptide Binding database KIDBI - Kinetic Data of Bio-molecular Interactions database KinaseDB - Kinase Pathway Database MHCPEP - Database of MHC binding peptides MINT - Molecular Interactions Database MIPS CYGD - MIPS Comprehensive Yeast Genome Database MPID - MHC-Peptide Interaction Database MycoPathPD - Human Fungal Pathogens Proteome Database NetPro - Molecular Connections NetPro ooTFD - Object Oriented Transcription Factors Database OPHD - The Online Predicted Human Interaction Database Pathcall - Curagen Pathcall PathDB - Pathways Database PDZBase - PDZ domain protein-protein interaction database PhosphoELM - Post-translational phosphorylation database PhosphoSite - Database of phosphorylation sites PhosphoSite - Cell Signaling Technology's PhosphoSite Database PBASE - PBASE PIMab - Drosophila Protein Interaction Map Database PIMRider - Protein Interaction Map - Hybrigenics PINDb - Proteins Interacting in the Nucleus database POINT - Prediction of Intercome PombePD - Schizosaccharomyces pombe Proteome Database</p>	<p>PPID - Protein-Protein Interaction Database PPV (FANTOM) - RIKEN FANTOM Protein Protein Interaction Viewer POS - Protein Quaternary Structure database Predictome - Predictome ProChart - ProChart database of signal transduction pathway information Prolinks - Prolinks ProMesh - ProMesh Protein-Protein Interaction Database ProNet - Protein-protein Interaction Database</p> <p>CSNDB - Cell Signaling Networks Database DOQCS - Database of Quantitative Cellular Signaling DSM - Dynamic Signaling Maps Predictome - Predictome GeneNet - Genetic Networks GenMAPP - Gene MicroArray Pathway Profiler GOLDdb - Genomics of Lipid-associated Disorders GON - Genomic Object Net INOH - Integrating Network Objects with Hierarchies</p>	<p>STCDB - Signal Transduction Classification Database STKE - Signal Transduction Knowledge Environment TRANSPATH - Signal Transduction Browser TRMP - Therapeutically Relevant Multiple Pathways Database TRRD - Transcription Regulatory Regions Database XPD - Phosphorylation Site Database</p> <p>Pathway Diagrams BBID - Biological Biochemical Image Database BioCarta - BioCarta Pathway Diagrams BMPH - Boehringer Mannheim Biochemical Pathways Wall Chart DSM - Dynamic Signaling Maps eMIM - Electronic Molecular Interaction Map HPRD - Human Protein Reference Database INOH - Integrating Network Objects with Hierarchies KEGG - Kyoto Encyclopedia of Genes and Genomes KIMB - Kohn Molecular Interaction Maps Malaria - Malaria Parasite Metabolic Pathways MIPS CYGD - MIPS Comprehensive Yeast Genome Database MPB - Metabolic Pathways of Biochemistry NetBiochem - Medical Biochemistry Resource Nicholson Minimaps - IJLMB-Nicholson Minimaps PharmGKB - The Pharmacogenetics and Pharmacogenomics Knowledge Base PID - CMAP Pathway Interaction Database SPAD - Signaling Pathway Database TRMP - Therapeutically Relevant Multiple Pathways Database</p> <p>Transcription Factors / Genetic Regulatory Networks COMPEL - Database on Composite Regulatory Elements DBTBS - Database of Bacillus subtilis Promoters and Transcription Factors DPIInteract - DNA-Protein Interactions Database GeNet - Gene Networks Database HoxPro - HOX Pro Interolog - Interolog/Regulog Database JASPAR - JASPAR Transcription Factor Binding Profile Database MAPPER - MAPPER ooTFD - Object Oriented Transcription Factors Database PRODORIC - Prokaryotic database of gene regulation RegulonDB - Database on Transcriptional Regulation and Genome Organization SCPD - The Promoter Database of Saccharomyces cerevisiae TRACTOR_DB - TRACTOR_DB TRANSFAC - Transcription Factor Database TRED - Transcriptional Regulatory Element Database TRRD - Transcription Regulatory Regions Database</p>	<p>CTD - Comparative Toxicogenomics Database Het-PDB_Navi - Het-PDB Navi KIBank - KIBank Ligand Depot - Ligand Depot MDB - Metalloprotein Database NRR - Nuclear Receptor Resource ORDB - Olfactory Receptor Database PDB-Ligand - PDB-Ligand PDSP - Psychoactive Drug Screening Program KI Database PLD - Protein Ligand Database RelBase - Protein-ligand database query tool ResNet - ResNet TTD - Therapeutic Target Database</p> <p>Genetic Interaction Networks BIND - Biomolecular Interaction Network Database GeneNet - Genetic Networks GenePath - GenePath GRID - General Repository for Interaction Datasets KNIFE - Drosophila pattern formation knowledge base MIPS CYGD - MIPS Comprehensive Yeast Genome Database</p> <p>Protein Sequence Focused ASC - Active Sequences Collection InBase - The Intra Database MEROPS - MEROPS Peptidase Database NRR - Nuclear Receptor Resource ORDB - Olfactory Receptor Database PANTHER - PANTHER PhosphoBase - Database of phosphorylation sites REBASE - Restriction Enzyme Database SCOPE - Database of protein catalytic domains Senta - Senta Signal Transduction Database TGD - Tumor Gene Database XPD - Phosphorylation Site Database</p>
<p>Protein-Protein Interactions 3CID - 3D interacting domains ABCdb - Archaea and Bacteria ABC transporter database AICS - Alliance for Cellular Signaling Molecule Pages Database AIFuse - Functional Associations of Proteins in Complete Genomes ASEdb - Alanine Scanning Energetics Database ASD - Artificial Selected Proteins/Peptides Database BID - Binding Interface Database BIND - Biomolecular Interaction Network Database BindvDB - The Binding Database</p>	<p>Pubase - Protein Structural Interactome Database PubGene - PubGene ResNet - ResNet S/MARt_DB - The S/MARt transaction DataBase Scansite - Scansite SPID - Subtilis Protein interaction Database Scanite - Scanite SPIN-PP - Surface Properties of Interfaces - Protein-Protein Interfaces STRING - Search Tool for the Retrieval of Interacting Genes/Proteins SYFPEITHI - Database of MHC Ligands and Peptide Motifs WormPD - Caenorhabditis elegans Proteome Database YPD - Yeast Proteome Database</p>	<p>Signaling Pathways ACS - Alliance for Cellular Signaling Molecule Pages Database aMAZE - Protein Function and Biochemical Pathways Project BioModels - BioModels Database CellML Repository - CellML Model Repository COPE - Cytokines Online Pathfinder Encyclopedia</p>	<p>MetaCore - MetaCore pathway database Monod - Modeler's Notebook and Datasore PANTHER - PANTHER PathArt - Pathway Articulor PathDB - Pathways Database Pathways Knowledge Base - Ingenuity Pathways Knowledge Base PATKA - Pathway Analysis Tool for Integration and Knowledge Acquisition PDS - Pathways Database System PhosphoSite - Cell Signaling Technology's PhosphoSite Database PD - CMAP Pathway Interaction Database ProcessDB - ProcessDB Reactome - Reactome KnowledgeBase ResNet - ResNet ROSPath - Reactive Oxygen Species related Signaling Pathway SBML Model Repository - SBML Model Repository Senta - Senta Signal Transduction Database SigPath - Signaling Pathway Information System SPAD - Signaling Pathway Database</p>	<p>Protein-Compound Interactions CLIBE - Computerized Lipid Binding Energy</p>

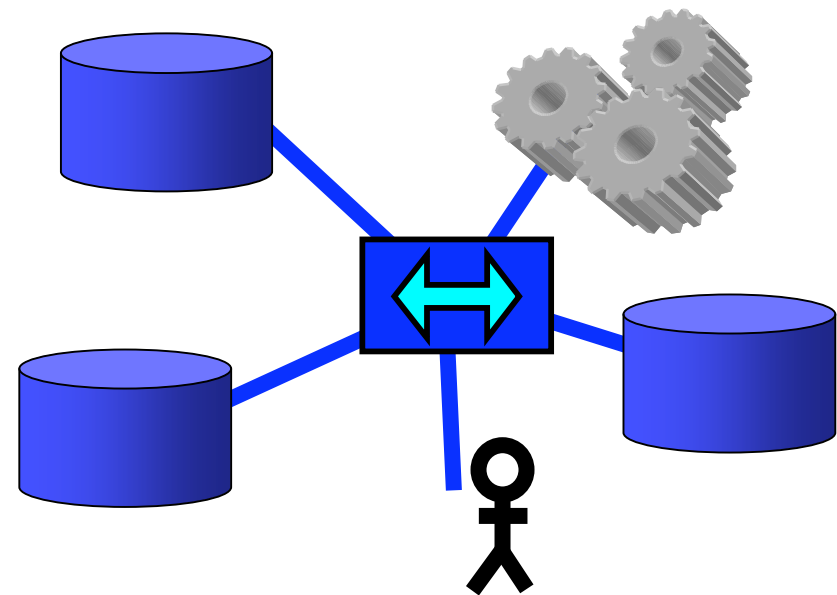
220 Pathway Databases!

- Arguably the most accessible data source, but...
 - Varied formats, representation, coverage
 - Pathway data extremely difficult to combine and use
- Pathguide Pathway Resource List (<http://www.pathguide.org>)**

Biological Pathway Exchange (BioPAX)



Before BioPAX
>100 DBs and tools
Tower of Babel



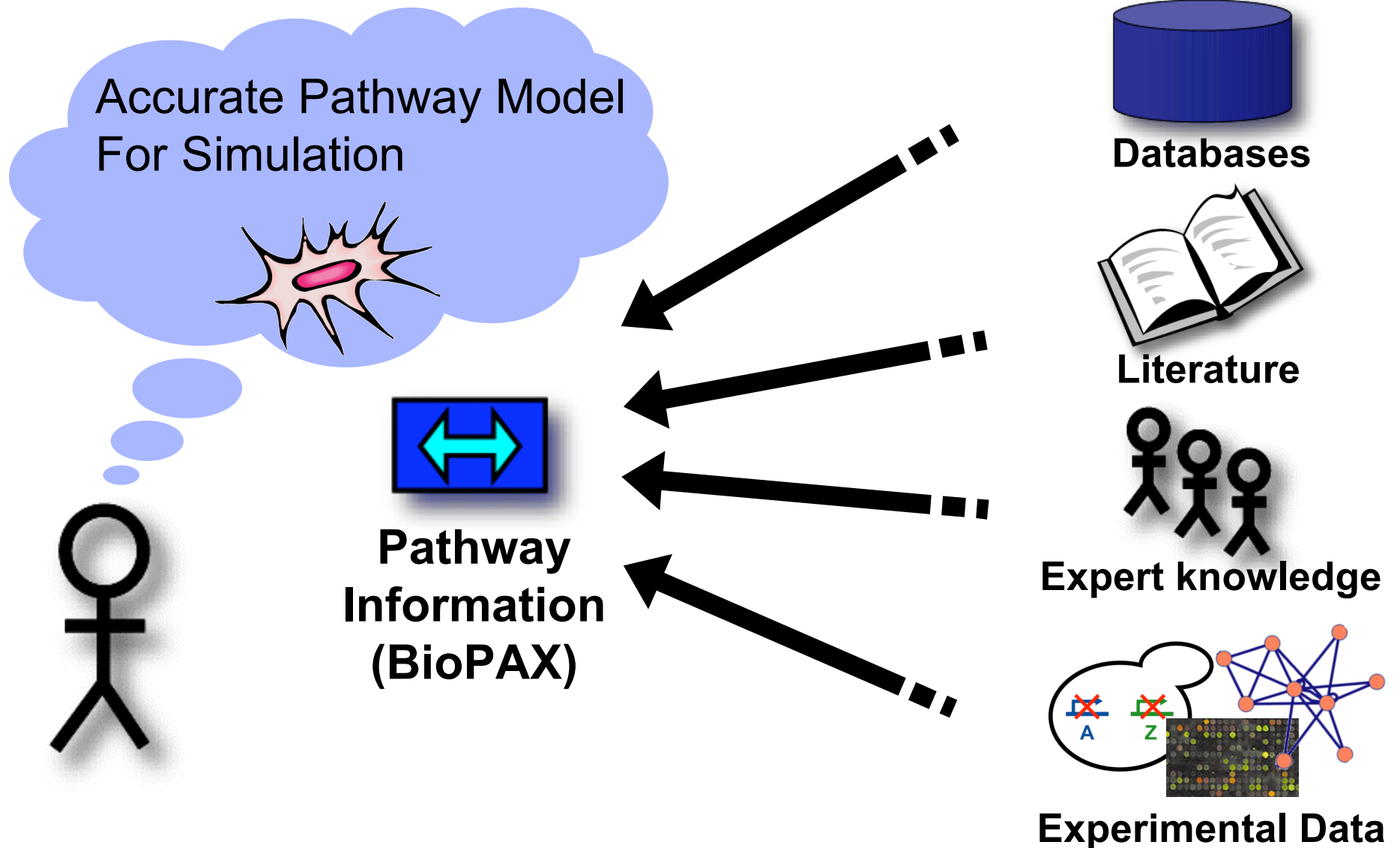
After BioPAX
Unifying language

Reduces work, promotes collaboration, increases accessibility

BioPAX Pathway Language

- Represent:
 - Metabolic pathways
 - Signaling pathways
 - Protein-protein, molecular interactions
 - Gene regulatory pathways
 - Genetic interactions
- Community effort: pathway databases distribute pathway information in standard format

Using Pathway Information

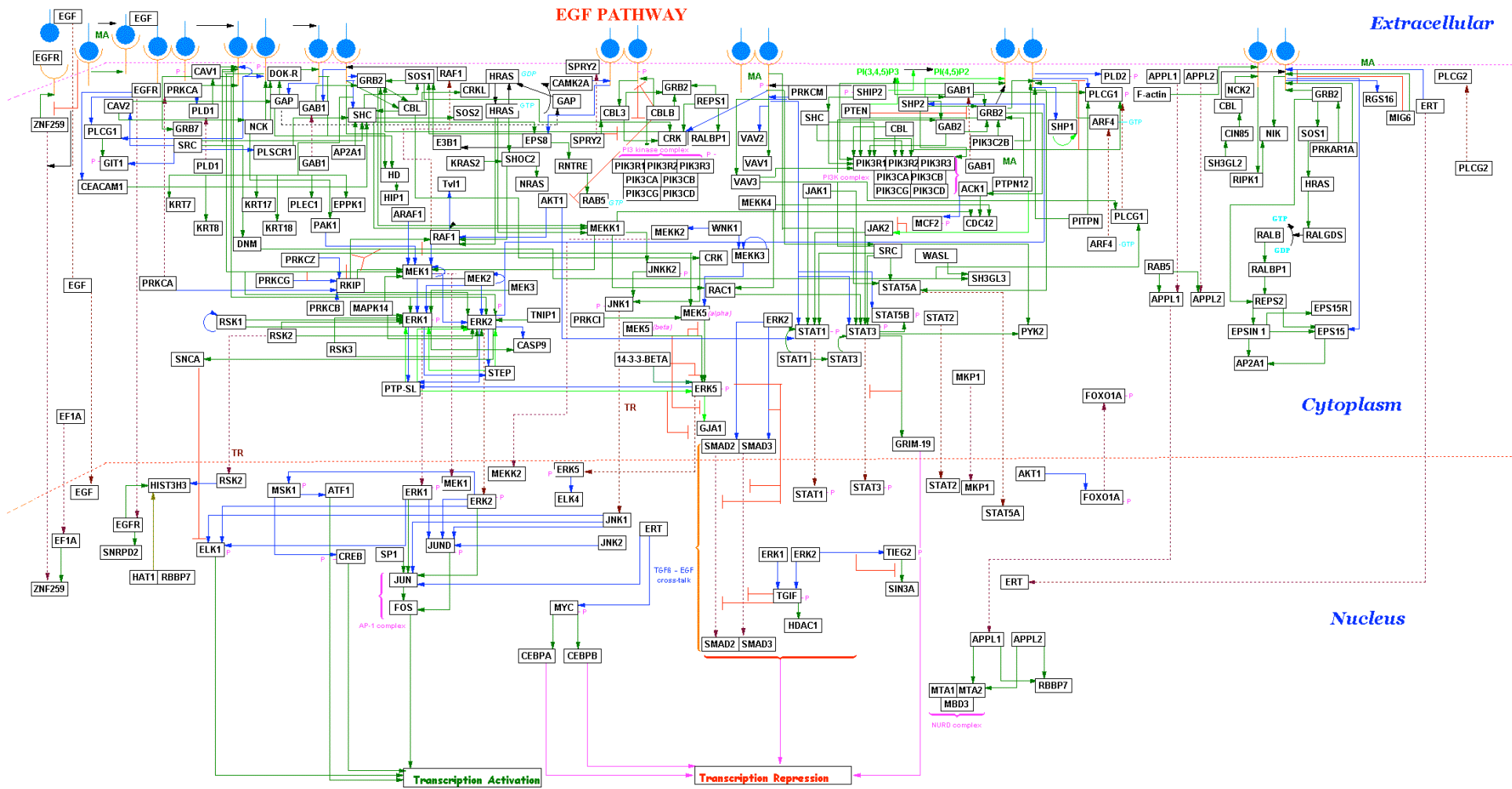


The Cancer Cell Map

- EGF, TGFR, AR, Delta-notch, A6B4 Integrin, Id, Kit, TNF-alpha, Wnt, Hedgehog (10 pathways)

<http://cancer.cellmap.org>

- Details on interaction, reactions, post-translational modifications from membrane to nucleus
- Derived from original articles
- Reviewed by MSKCC experts in Massague, Benezra, Besmer, Gerald, Giancotti labs



EGF Pathway

>170 Proteins

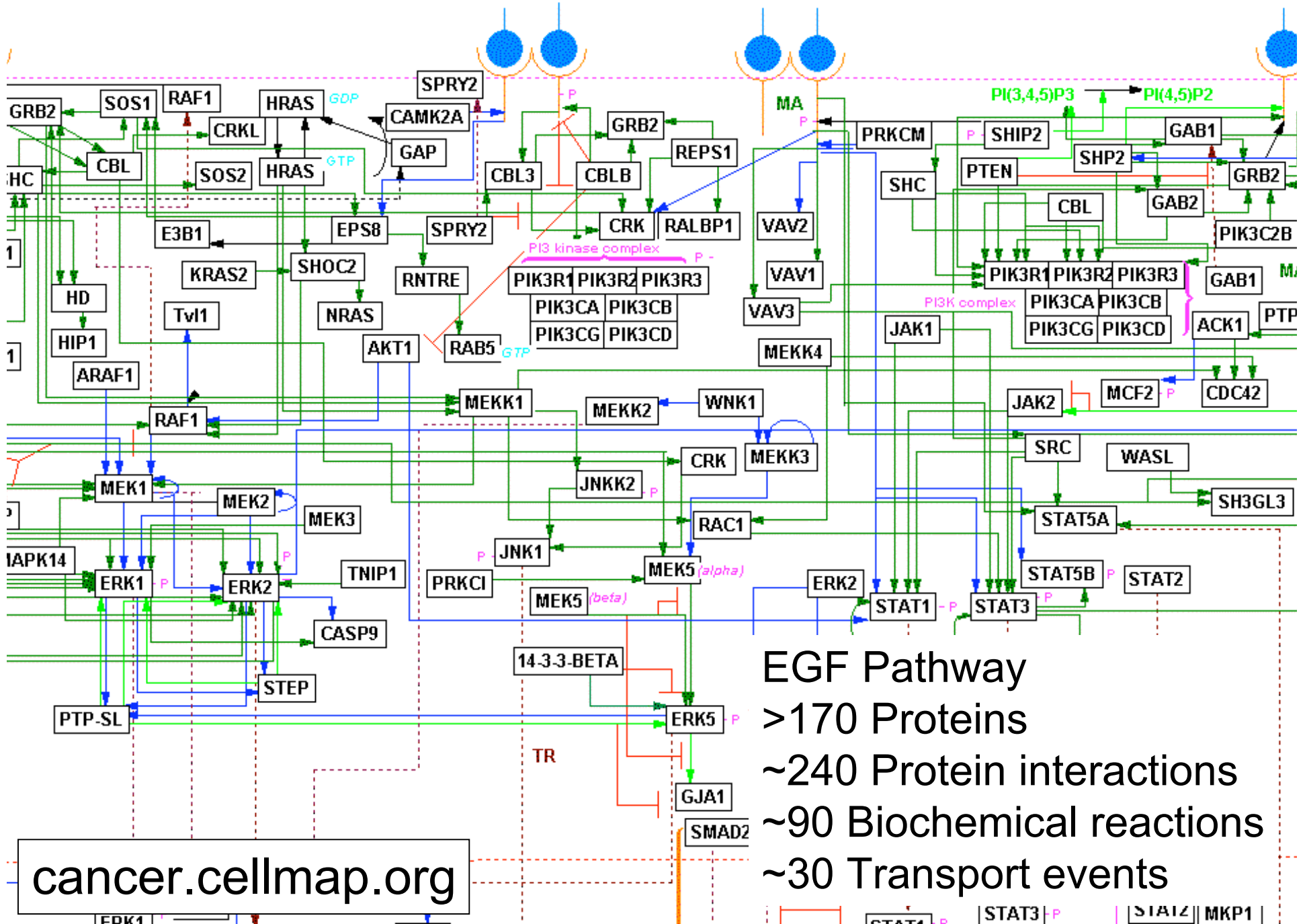
~240 Protein interactions

~90 Biochemical reactions

~30 Transport events

cancer.cellmap.org

EGF PATHWAY



EGF Pathway
 >170 Proteins
 ~240 Protein interactions
 ~90 Biochemical reactions
 ~30 Transport events

cancer.cellmap.org

Legend: [STAT1]-P, [STAT3]-P, [STAT2]-P, [STAT12]-MKP1

The Cancer Cell Map

Memorial Sloan-Kettering Cancer Center

SEARCH EXAMPLES

[Home](#) [Organisms](#) [FAQ](#) [Web Service](#) [Cytoscape](#) [About](#)

EGFR1 (Pathway) from Homo sapiens

Data The Cancer Cell Map at Memorial Sloan-Kettering Cancer Center (<http://cancer.cellmap.org>)

Source:

Availability: Freely available from The Cancer Cell Map (created by Memorial Sloan-Kettering Cancer Center and the Institute of Bioinformatics) under the Creative Commons license (<http://creativecommons.org/licenses/by/2.5/>)

Comment: The Epidermal growth factor (EGF) signals through the epidermal growth factor (EGF) receptor (EGFR) family composed of four receptor tyrosine kinases (RTKs) designated EGFR, ErbB2 (also known as HER2 or neu), ErbB3 (HER3), and ErbB4 (HER4). In EGFR1 pathway, it includes interactions of EGF-EGFR only. Binding of EGF to the extracellular domain of these RTKs leads to receptor dimerization, activation of the intrinsic protein tyrosine kinase (PTK) activity, tyrosine autophosphorylation, and recruitment of various signaling proteins to the receptor. Tyrosine phosphorylation of the EGFR leads to the recruitment of various signaling proteins, including the adaptor proteins Grb2, Nck, phospholipase C gamma, Shc and other molecules. The EGF receptor activates RAS and the MAP kinase pathways, ultimately causing phosphorylation of transcription factors such as c-Fos, c-Jun, ELK-1. The STAT transcription factors are also activated by JAK kinases in response to EGF. This pathway is involved in the regulation of a wide spectrum of biological processes including cell proliferation, apoptosis, cell cycle etc.

[View Expression Data on this Pathway](#)

Contains the Following Molecules (Showing 1 - 20 of 177) [\[display all\]](#)

[AKT1](#) [AP2A1](#) [APPL](#) [APPL2](#) [ARAF1](#) [ARF4](#) [ATF1](#) [BCAR1](#) [CAMK2A](#) [CASP9](#) [CAV1](#) [CAV2](#) [CBL](#) [CBLB](#)
[CBL3](#) [CDC42](#) [CEACAM1](#) [CEBPA](#) [CEBPB](#) [CREB](#)

Contains the Following Interactions / Pathways (Showing 1 - 20 of 475) [\[display all\]](#)

Biochemical Reaction(s)

[View Details](#) [EGFR](#) → [EGFR](#) (Phosphorylated)

[View Details](#) [ELK4](#) → [ELK4](#) (Phosphorylated)

[View Details](#) [GAP](#) → [GAP](#) (Phosphorylated)

[View Details](#) [ERK5](#) (Phosphorylated) → [ERK5](#)

[View Details](#) [MEKK3](#) → [MEKK3](#) (Phosphorylated)

[View Details](#) [EPS15](#) → [EPS15](#) (Phosphorylated)

[View Details](#) [JUN](#) → [JUN](#) (Phosphorylated)

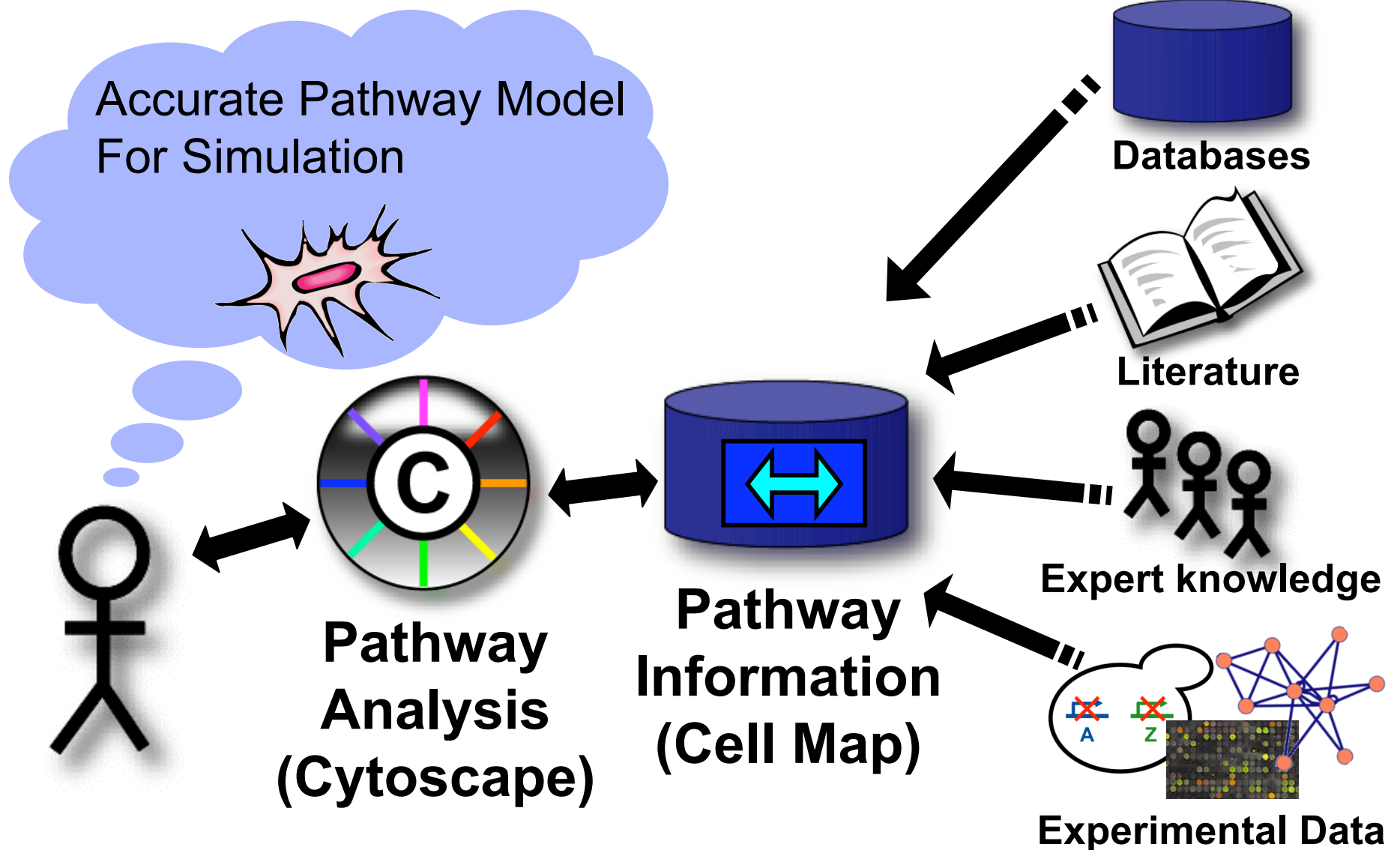
[View Details](#) [VAV2](#) → [VAV2](#) (Phosphorylated)

[View Details](#) [EGFR](#) → [EGFR](#) (Ubiquitinated)

[View Details](#) [EGFR](#) → [EGFR](#) (Ubiquitinated)

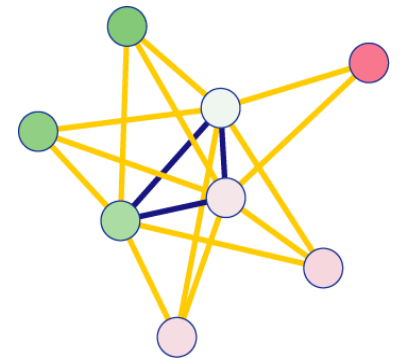
cancer.cellmap.org

Using Pathway Information



Network visualization and analysis tool: Cytoscape

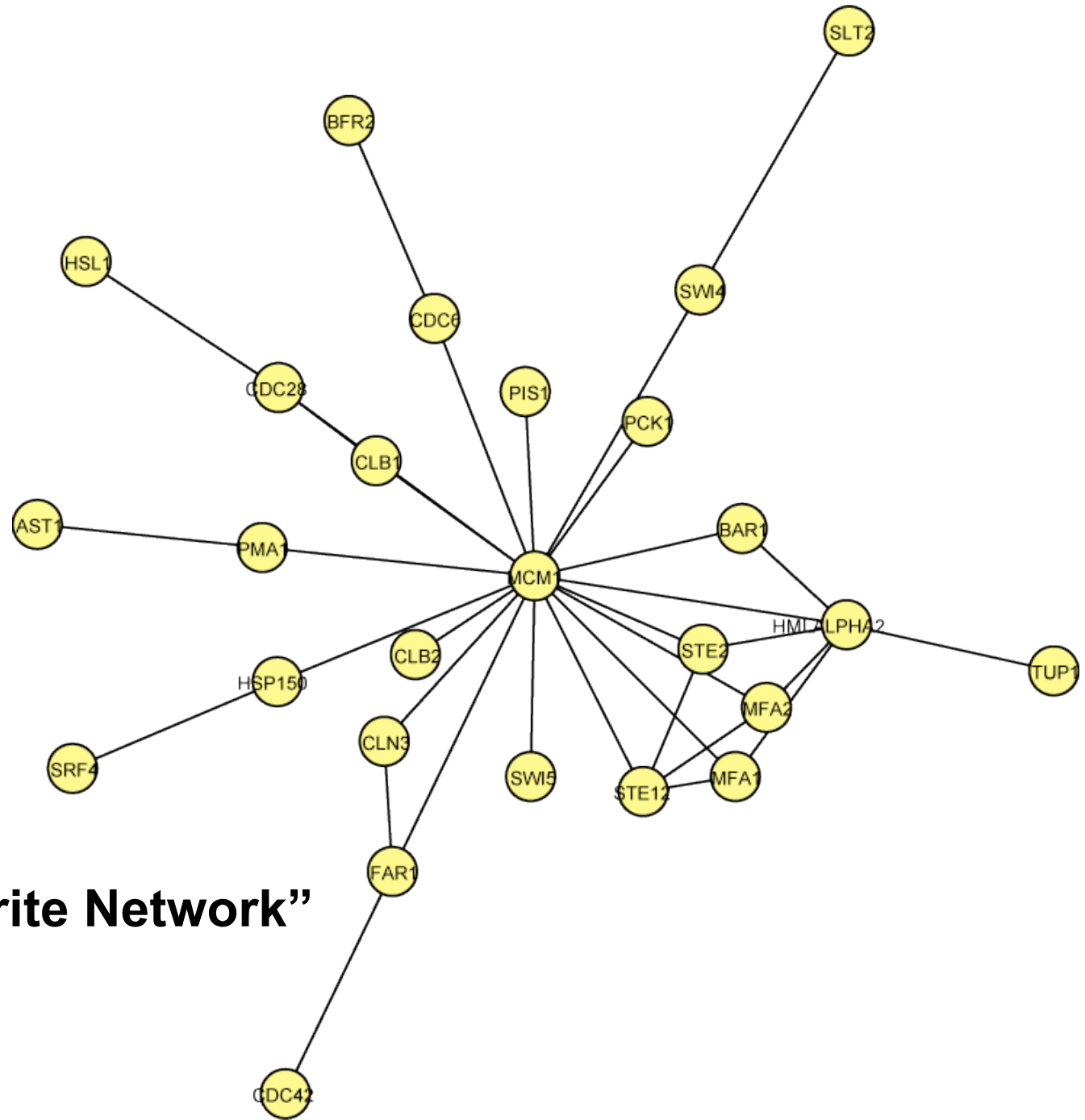
- Network-based molecular profiling analysis
 - Transcriptionally active network modules
- Network comparison
 - PathBLAST
- PubMed search (literature mining)



Visual Style

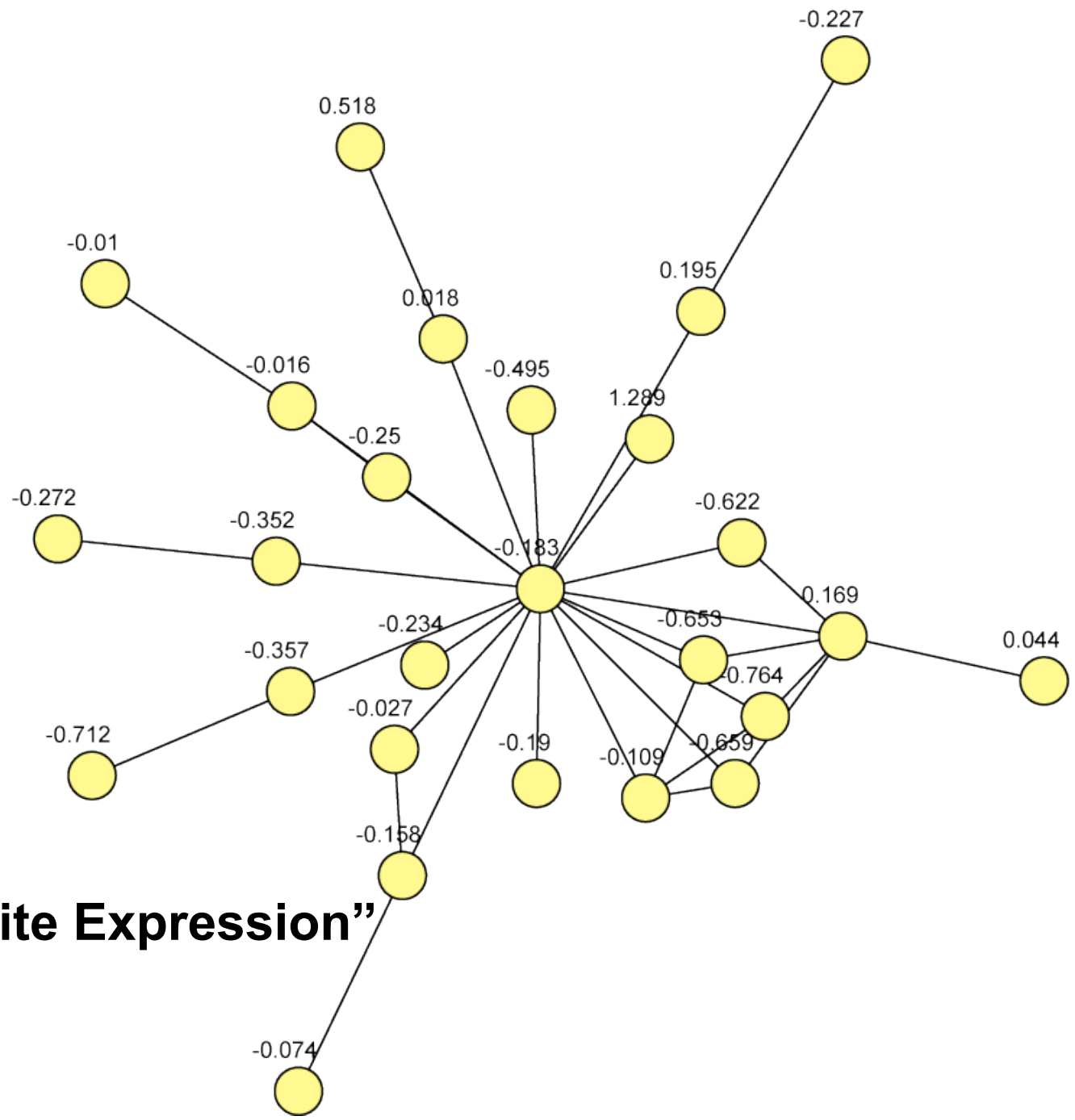
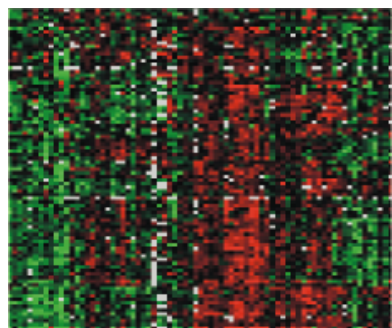
- Customized views of experimental data in a network context
- Network has node and edge attributes
 - E.g. expression data, interaction type, GO function
- Mapped to visual attributes
 - E.g. node/edge size, shape, colour...
- E.g. Visualize gene expression data as node colour gradient on the network

Visual Style



Load “Your Favorite Network”

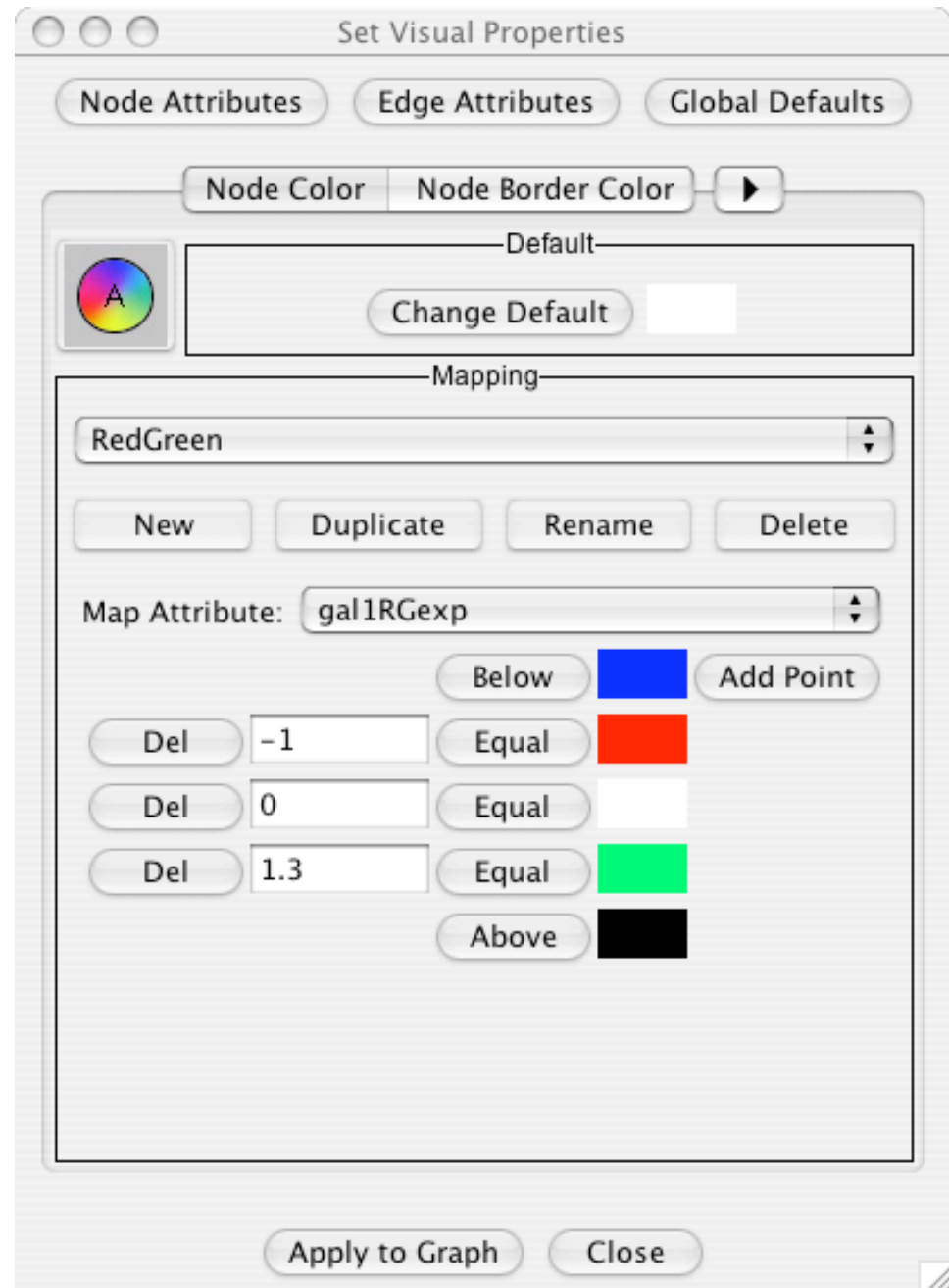
Visual Style



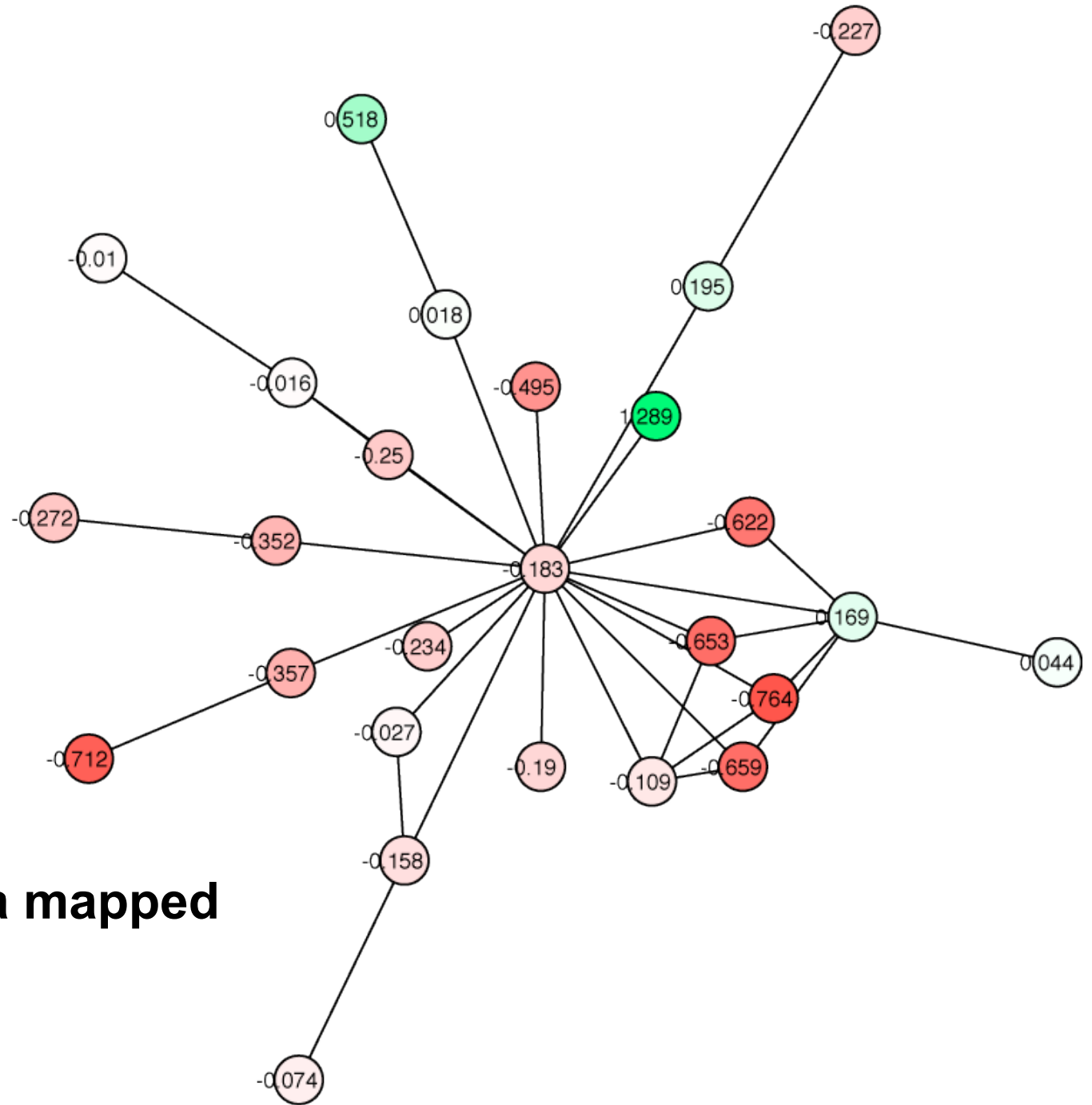
**Load “Your Favorite Expression”
Dataset**

Visual Style

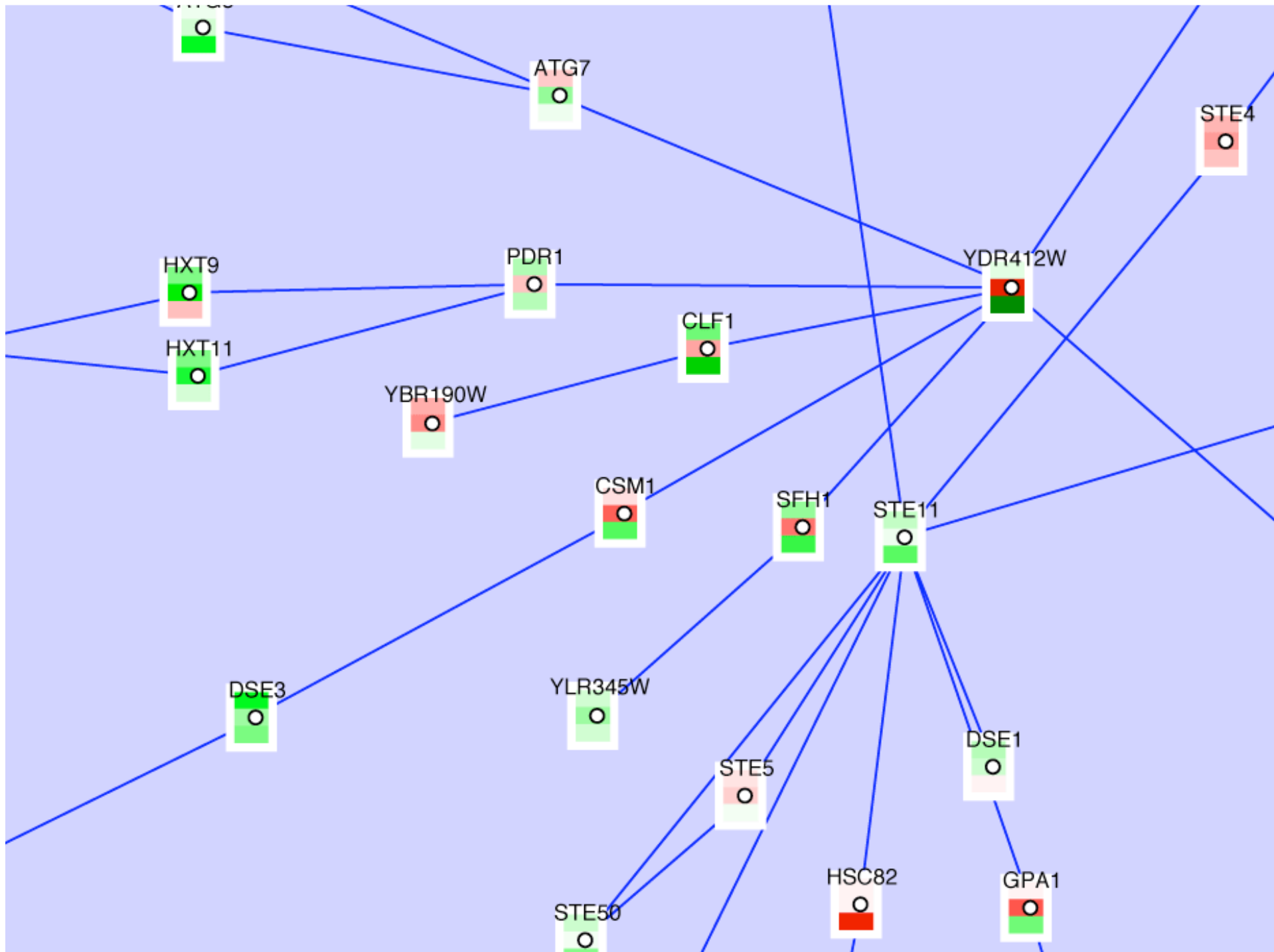
Map expression values to node colours using a continuous mapper



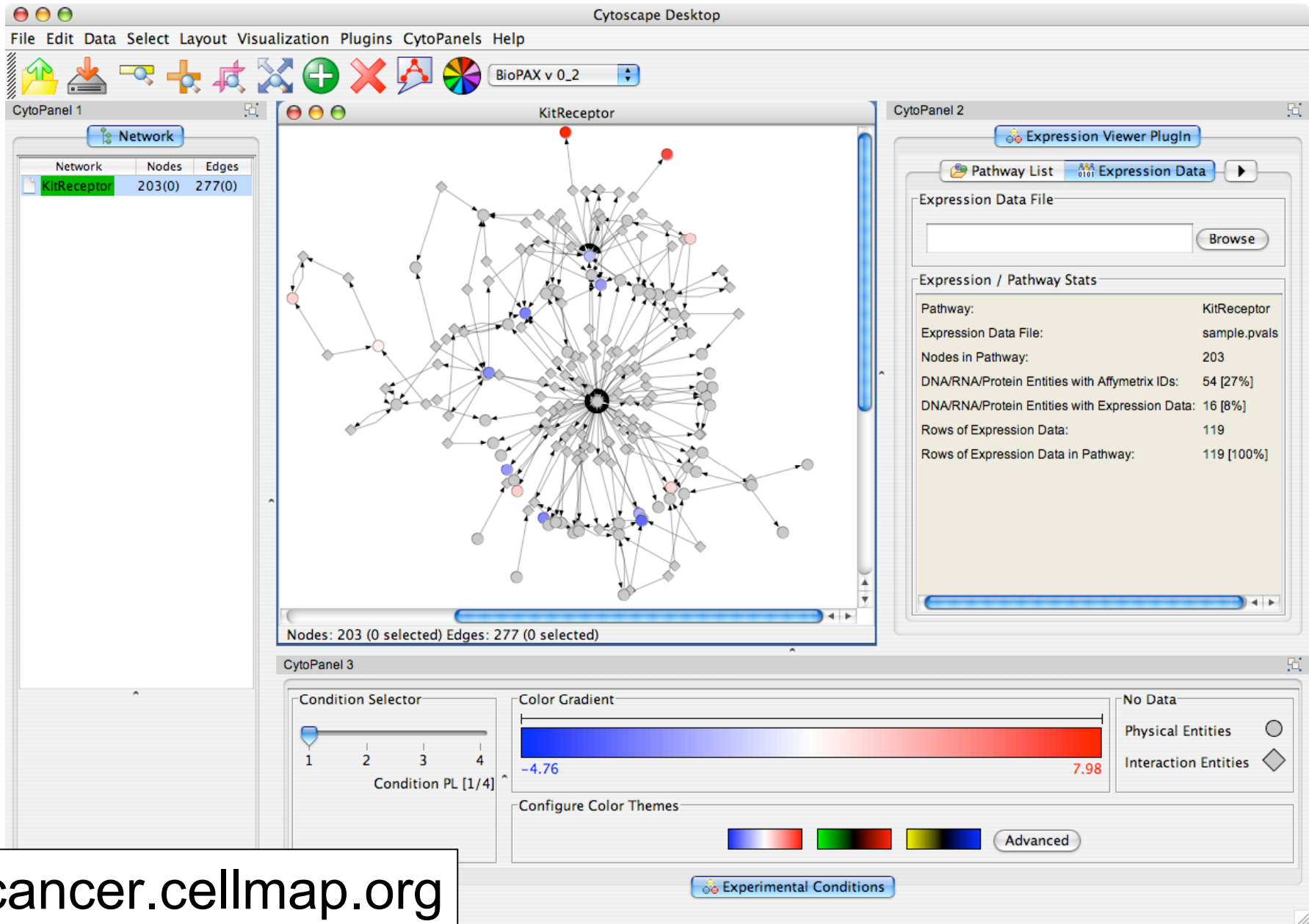
Visual Style



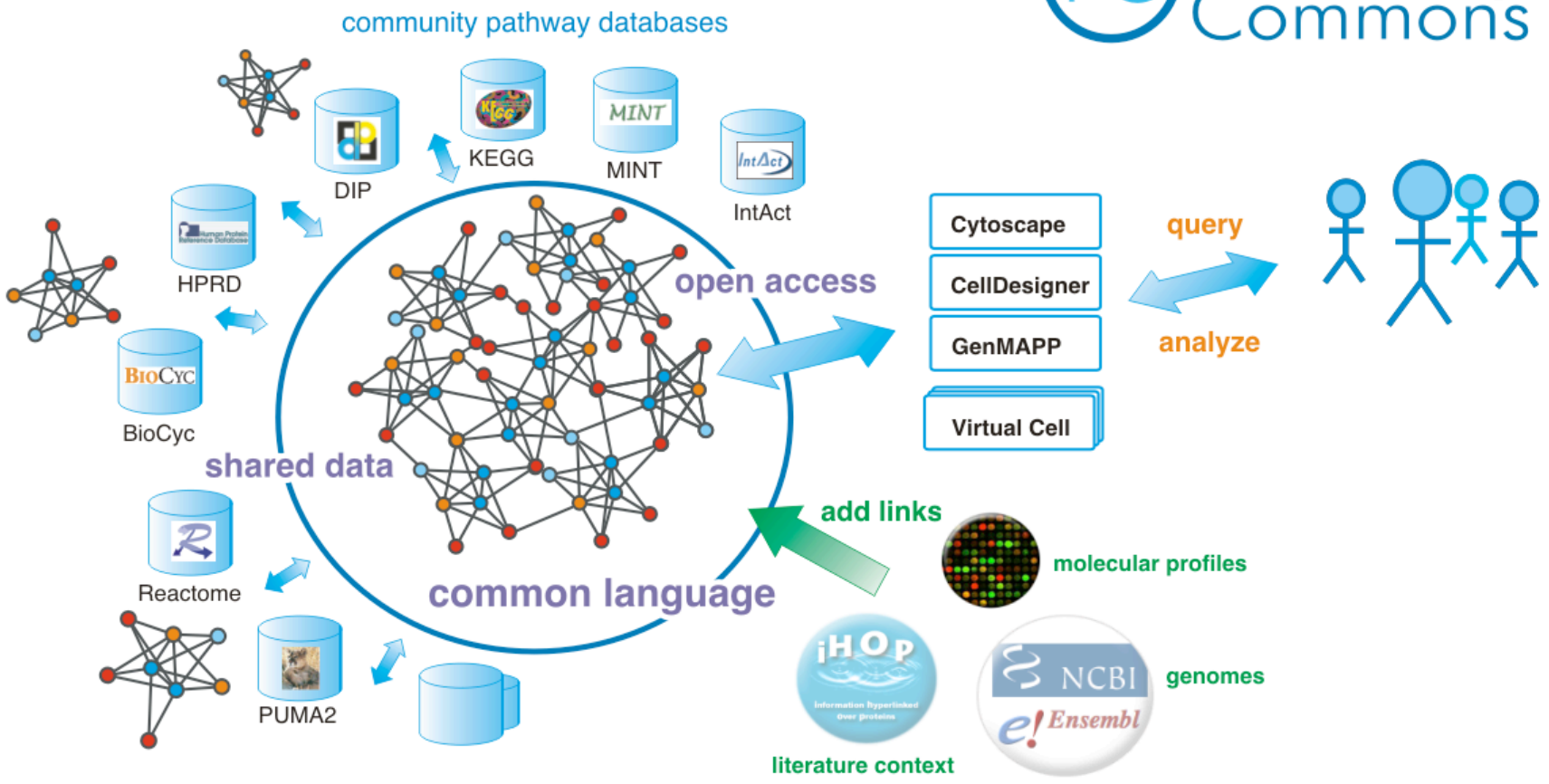
**Expression data mapped
to node colours**



The Cancer Cell Map



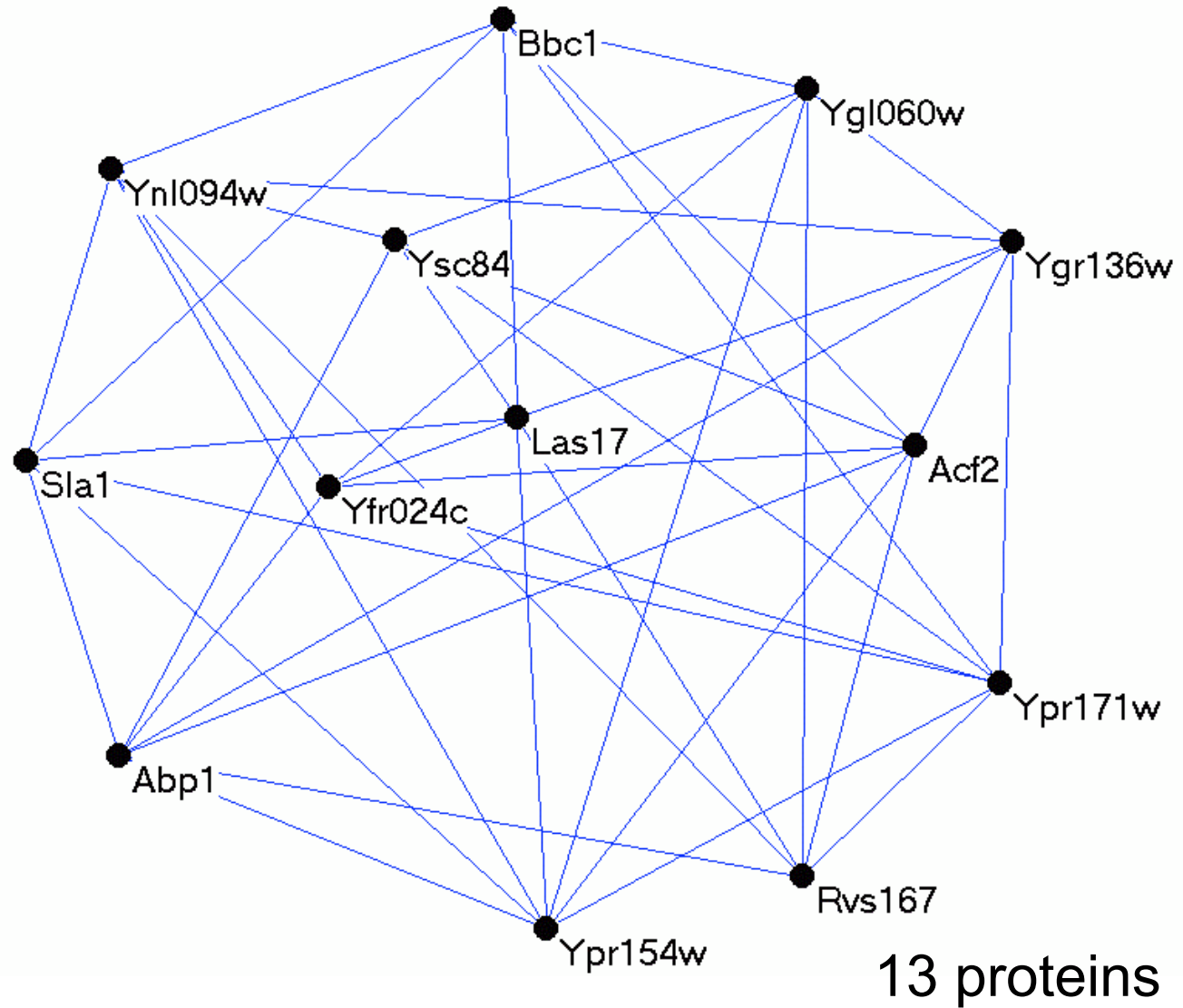
PC Pathway Commons



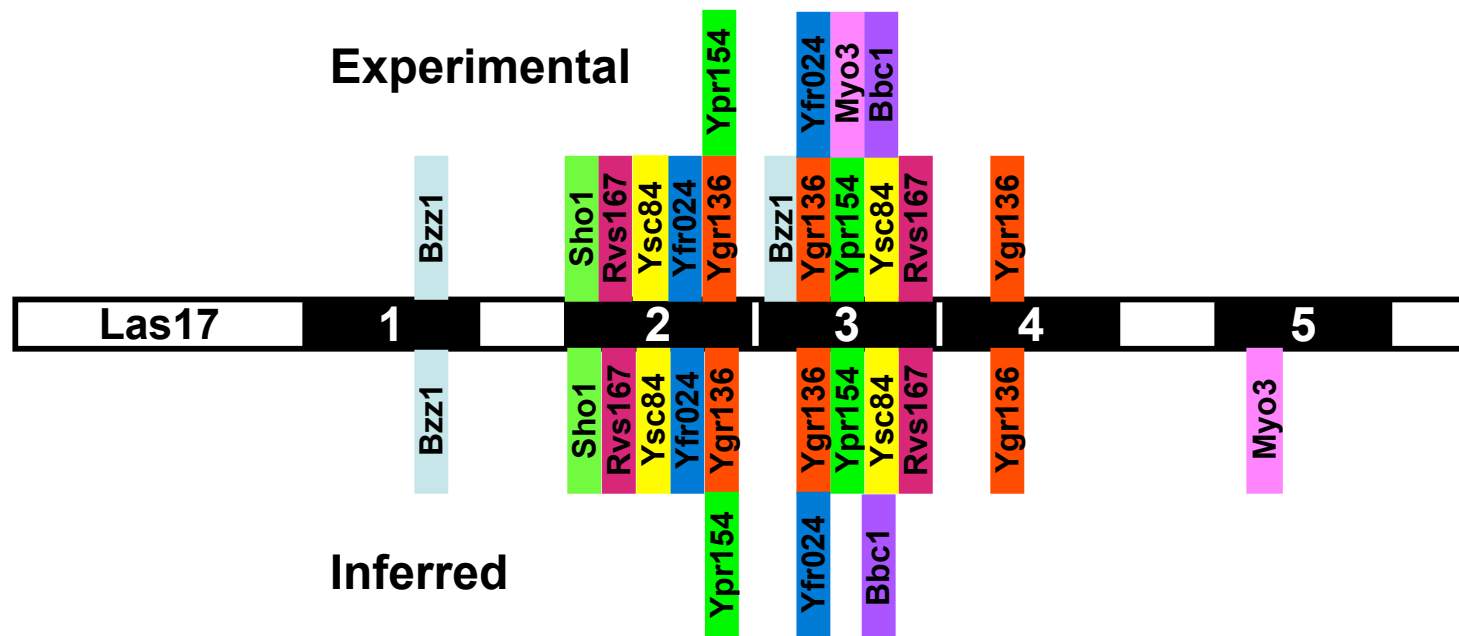
Practical uses of pathway information

- Use as a reference
 - Summary of what is known about a cellular process or gene
 - Visual communication
- Analyze molecular profiles
 - After ranking and clustering
 - Summary of ranked data or clusters (1000 genes, now what?)
- Visualize molecular profiles in pathway context
- Find active pathways in sets of genes defined by ranking or clustering
- Find protein complexes in protein interaction networks

Highly Connected 6-Core = Las17 Actin Assembly Complex?



Experimental Validation of Las17 Complex

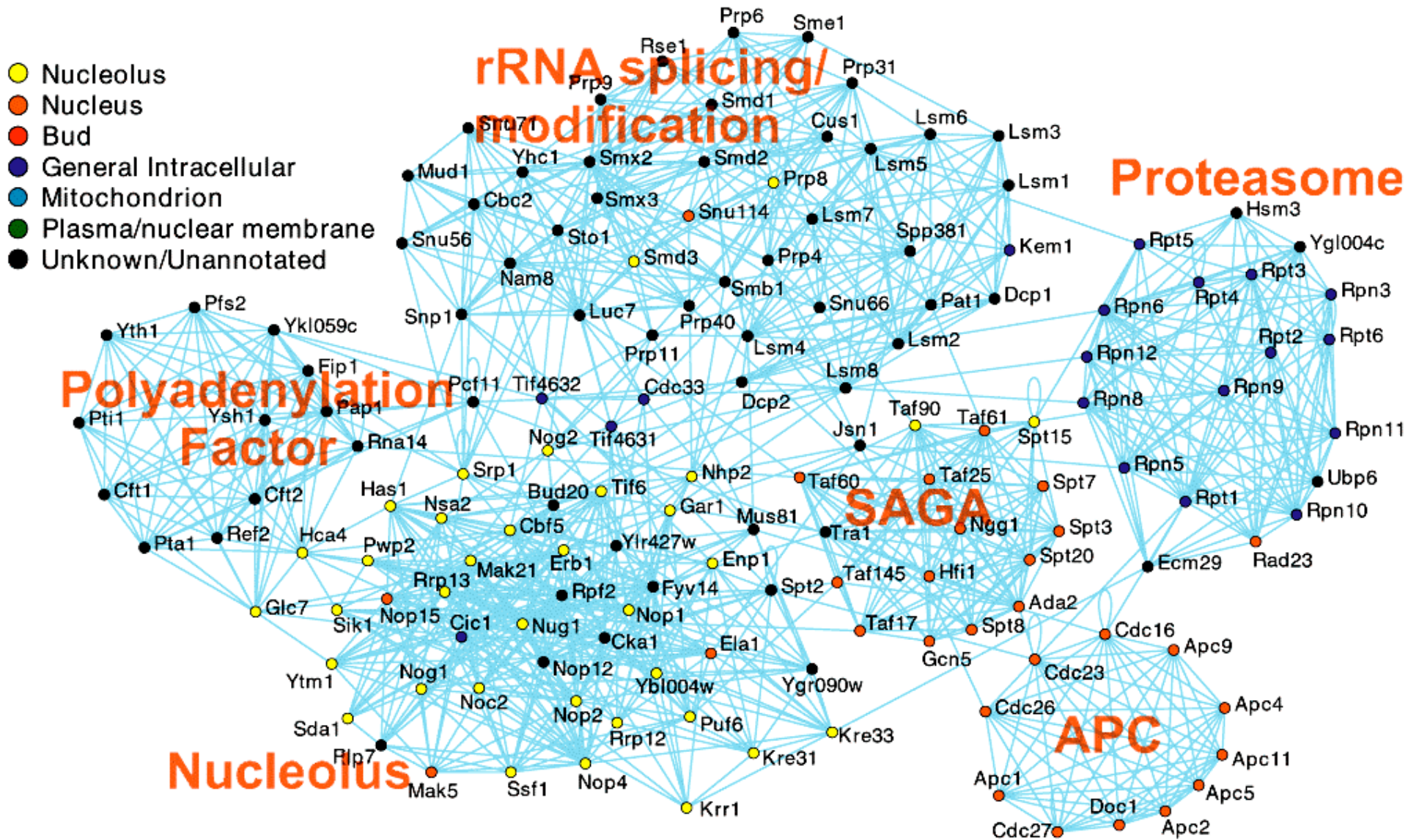


ELISA; Bbc1, Bzz1, Ygr136w, Ypr154w, Yfr024c, Ysc84 CoIP; Colocalized

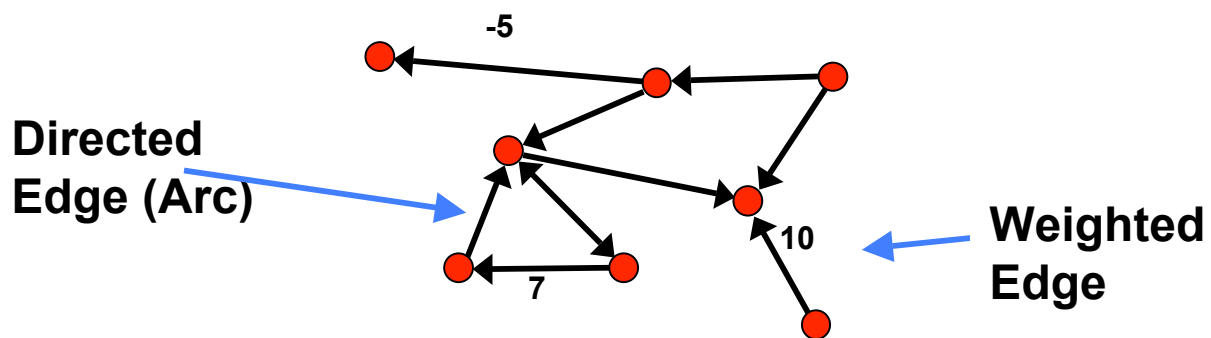
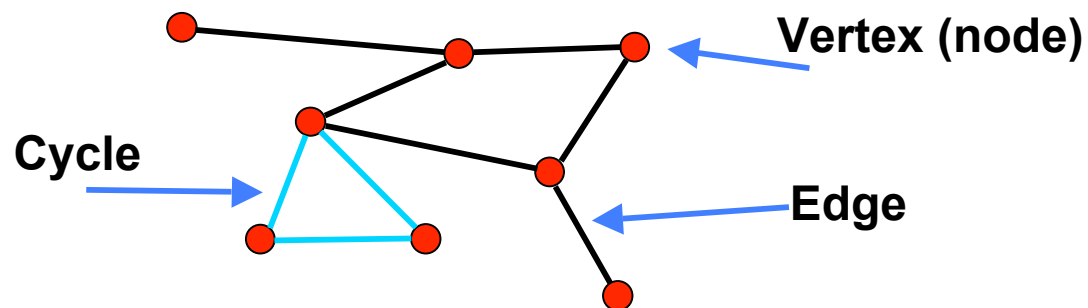
So...

- Based on observations, densely interconnected regions of an interaction network may represent molecular complexes
- Complexes can provide another level of annotation above classic 'guilt by association'
- ∴ Methods that find dense network regions can help us understand biological systems (using only qualitative connectivity information)

Nuclear Complexes



Graph Theory



We map molecular interaction networks to graphs

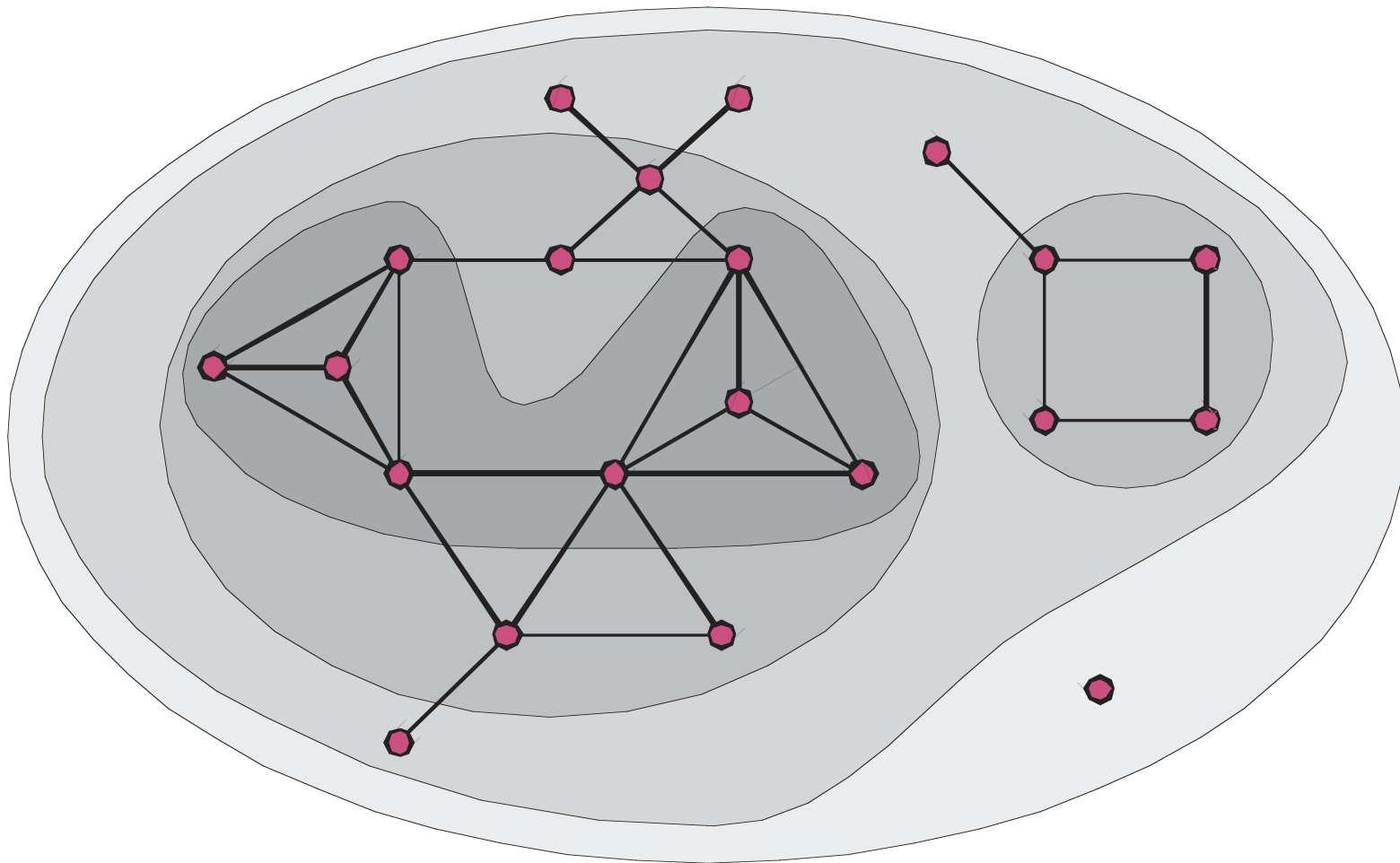
Mapping Biology to a Graph

- A simple mapping
 - one compound/node, one interaction/edge
- A more realistic mapping
 - Cell localization, cell cycle, cell type, taxonomy
 - Only represent physiologically relevant interaction networks
- Graph can be relaxed by representing similar molecules as one node (COGs)

k-core

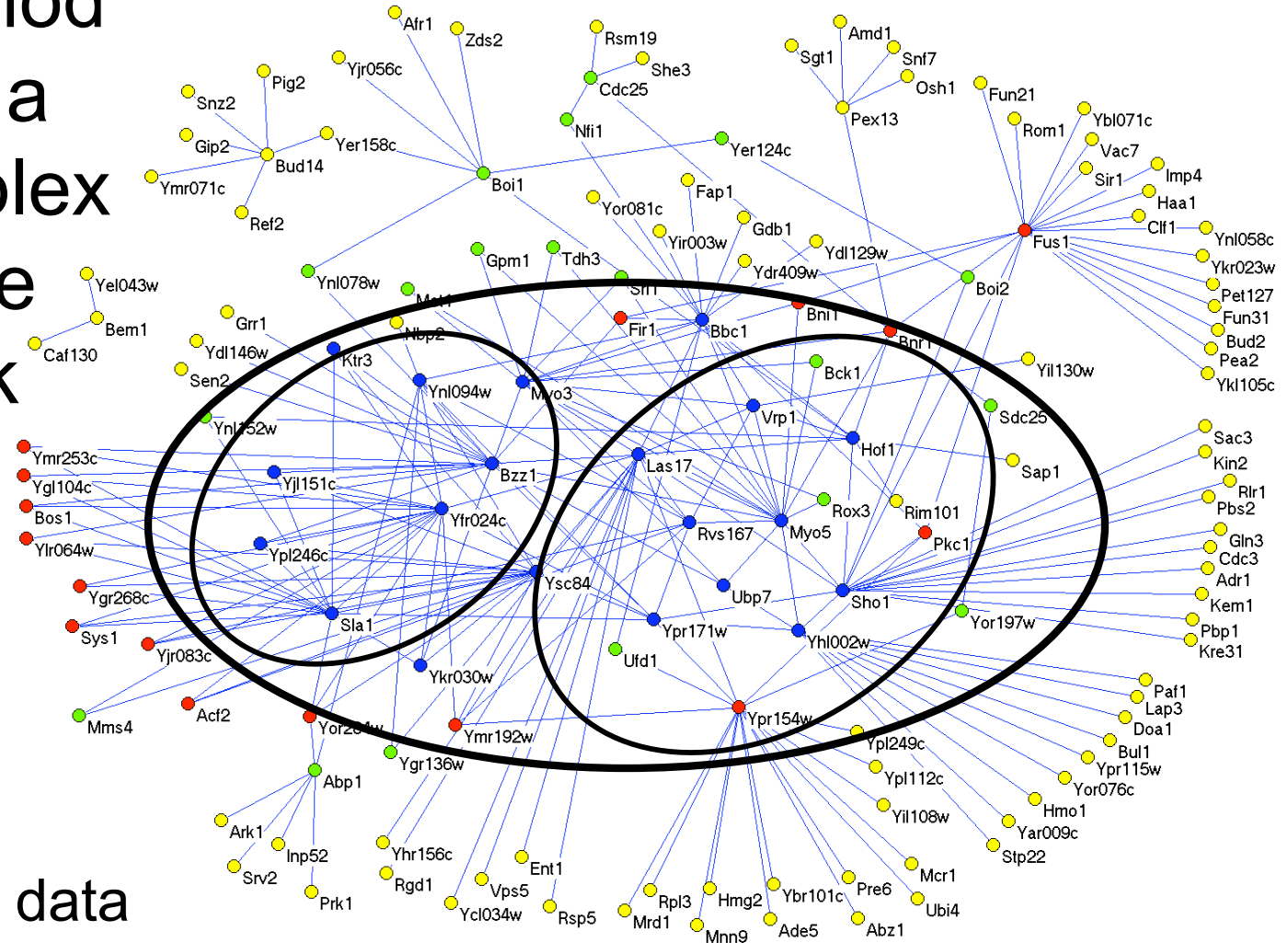
- A part of a graph where every node is connected to other nodes with at least k edges ($k=0,1,2,3\dots$)
- Highest k -core is a central most densely connected region of a graph
- Regions of dense connectivity may represent molecular complexes
- Therefore, high k -cores may be molecular complexes

k-core



A Better Complex Finder

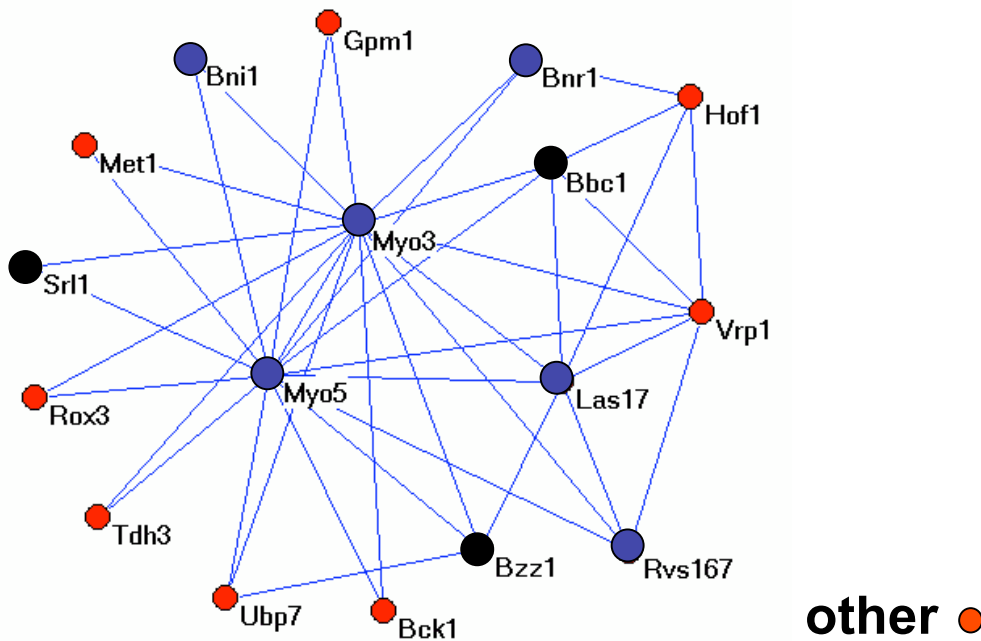
K-core method
is limited to a
single complex
in the middle
of a network



2 Complexes

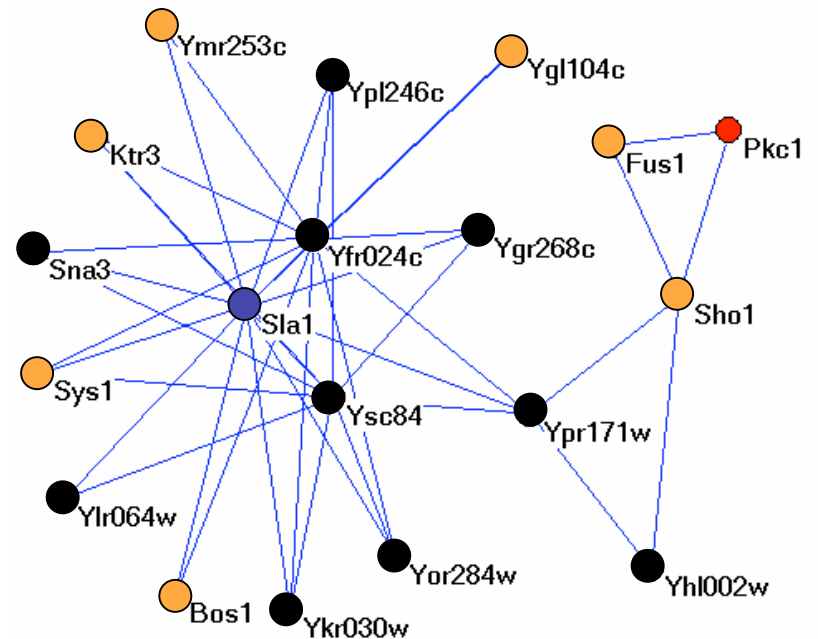
6/17 cell polarity ●
3/17 unknown role ●

actin cytoskeleton rearrangement

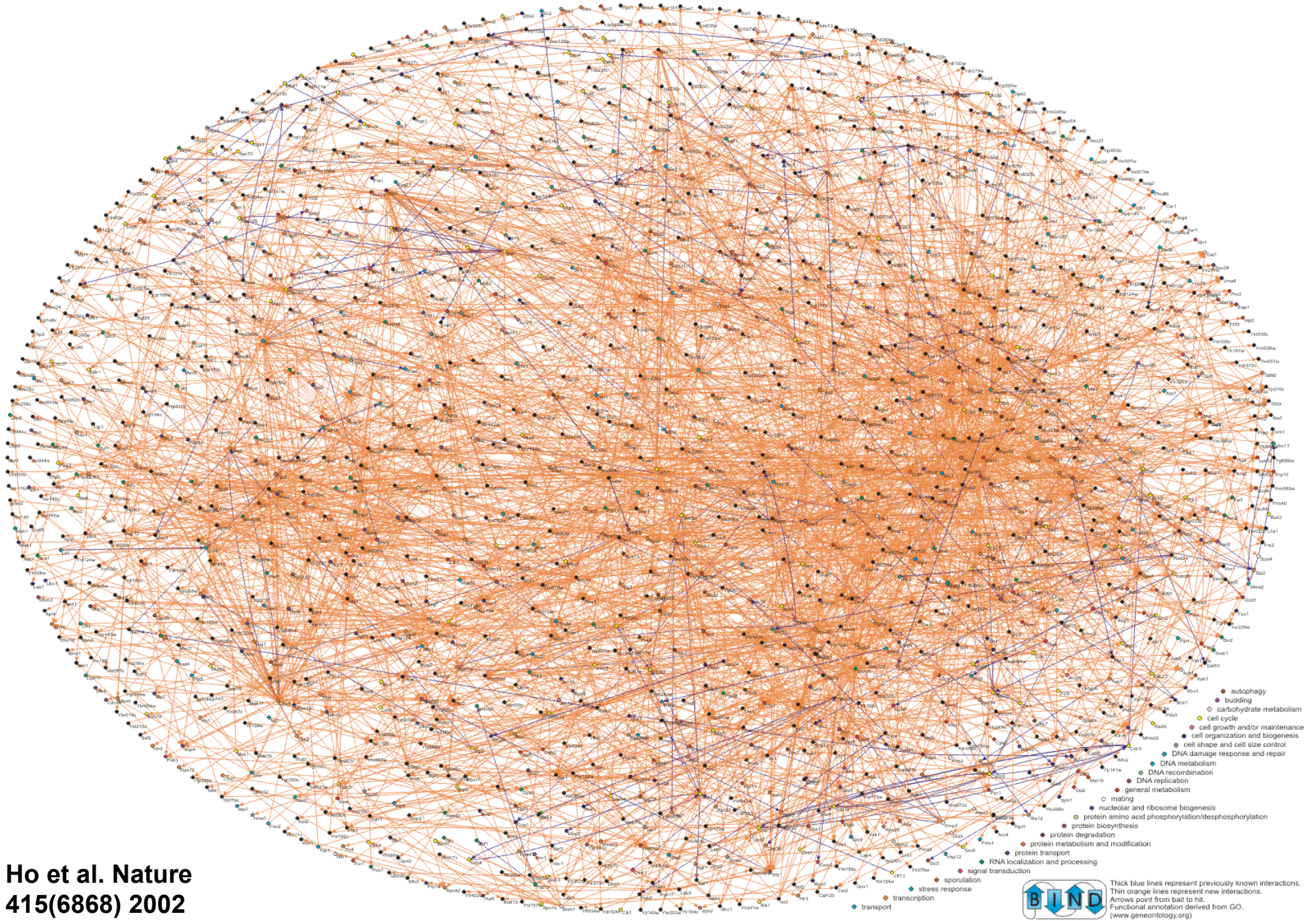


7/19 membrane ●
10/19 unknown ●
1/19 cytoskeletal ●

Signal transduction



Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry



Ho et al. Nature
415(6868) 2002

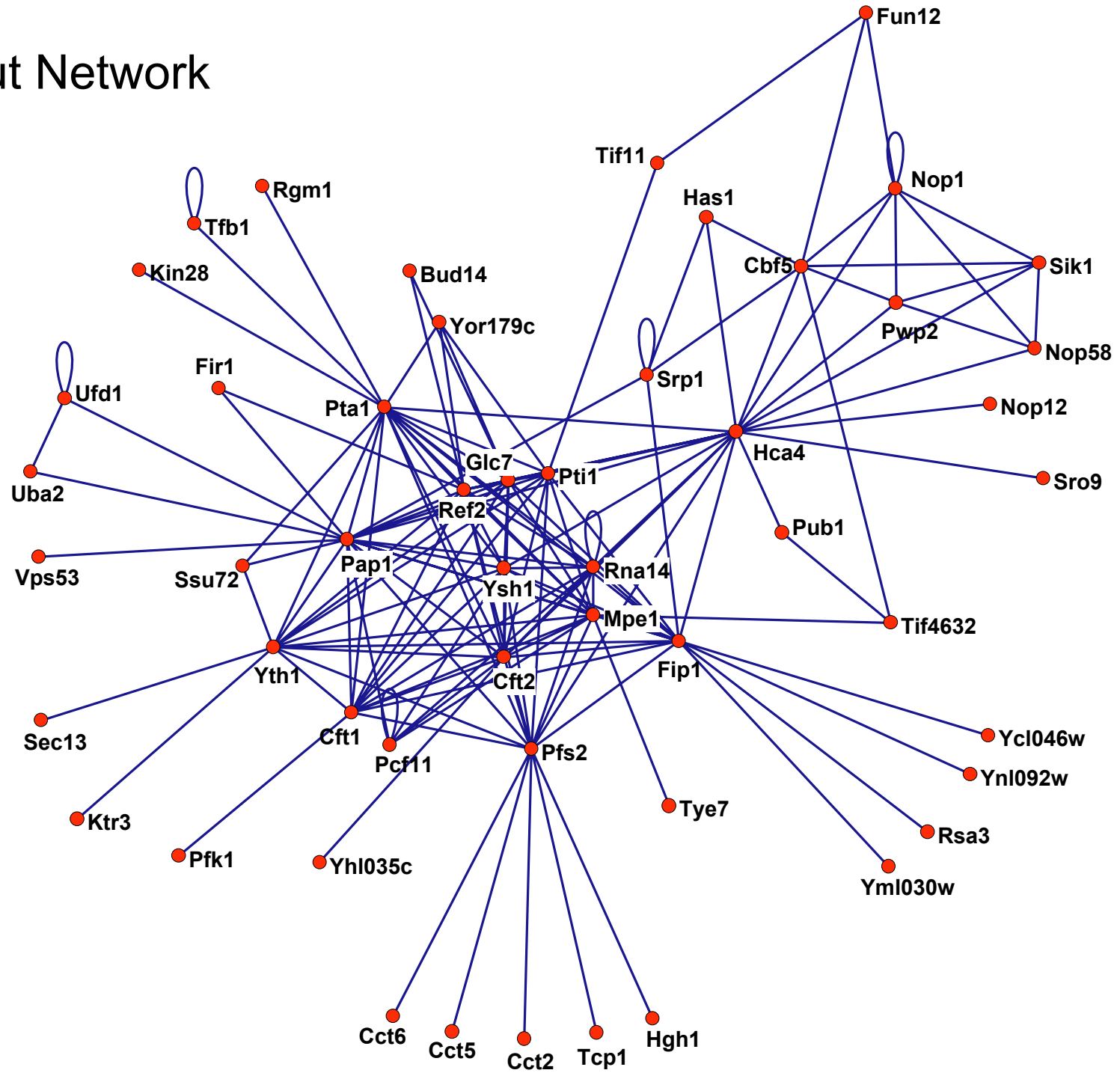


Molecular Complex Detection

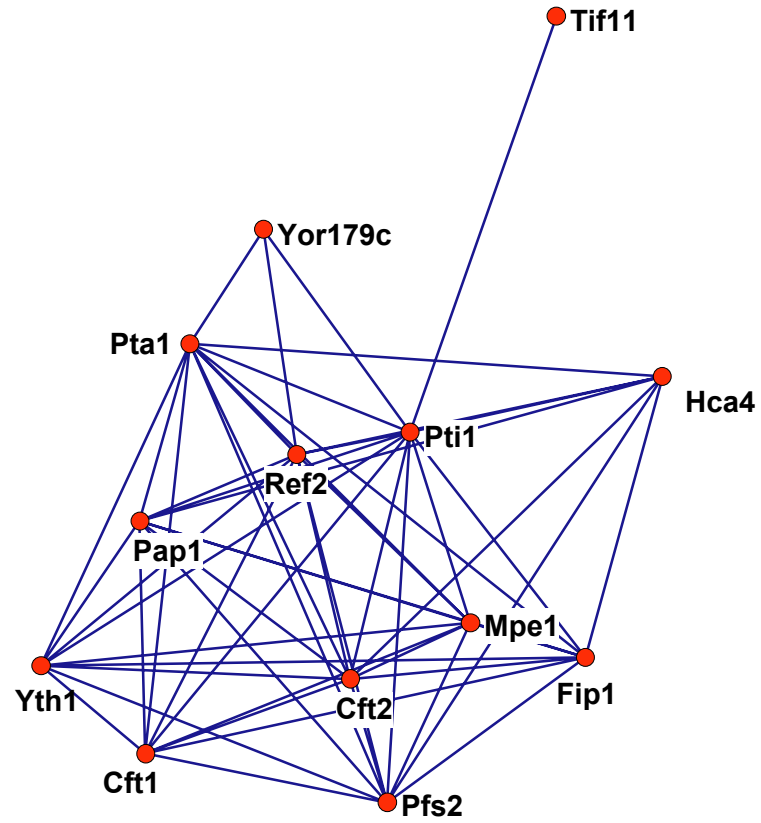
MCODE

- MCODE finds densely connected regions of a network
- Graph theoretic based clustering algorithm
 - Three stages:
 - Network Weighting
 - Complex Detection
 - Optional Post-processing

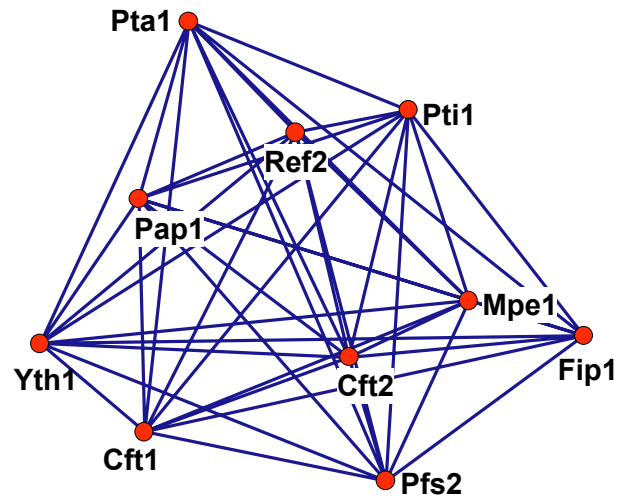
Input Network



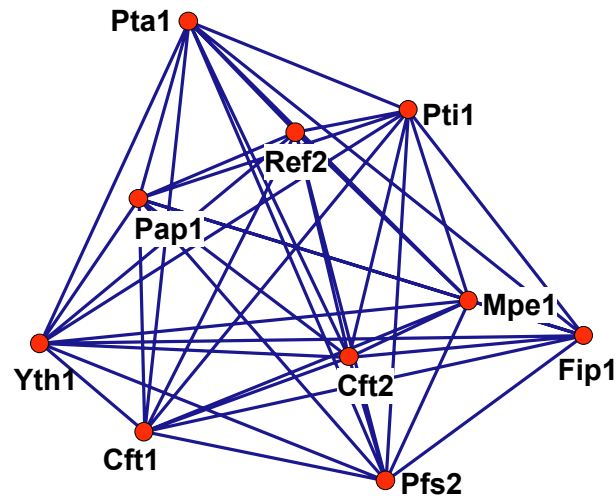
Find neighbors of Pti1



Find highest k-core (8-core)

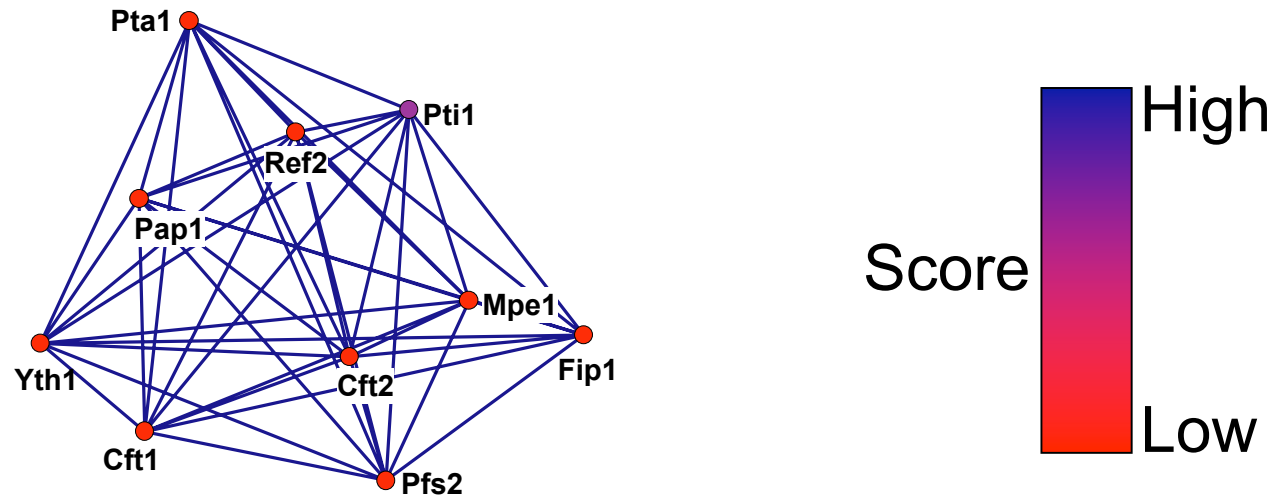


Find graph density



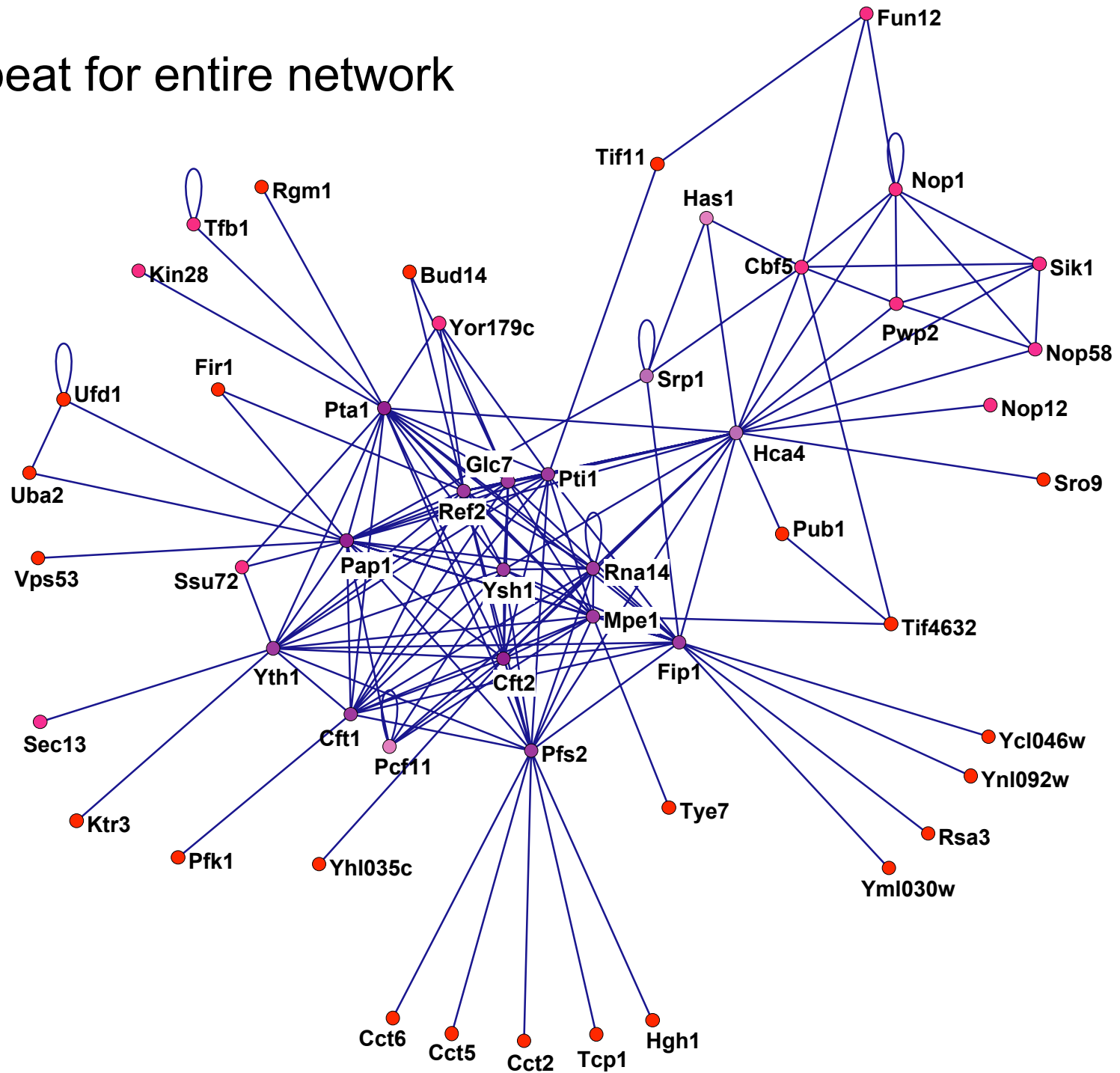
$$\text{Density} = \frac{\text{Number edges}}{\text{Number possible edges}} = 44/55 = 0.8$$

Calculate score for Pti1



$$\text{Score} = \text{highest k-core} * \text{density} = 8 * 0.8 = 6.4 = \text{●}$$

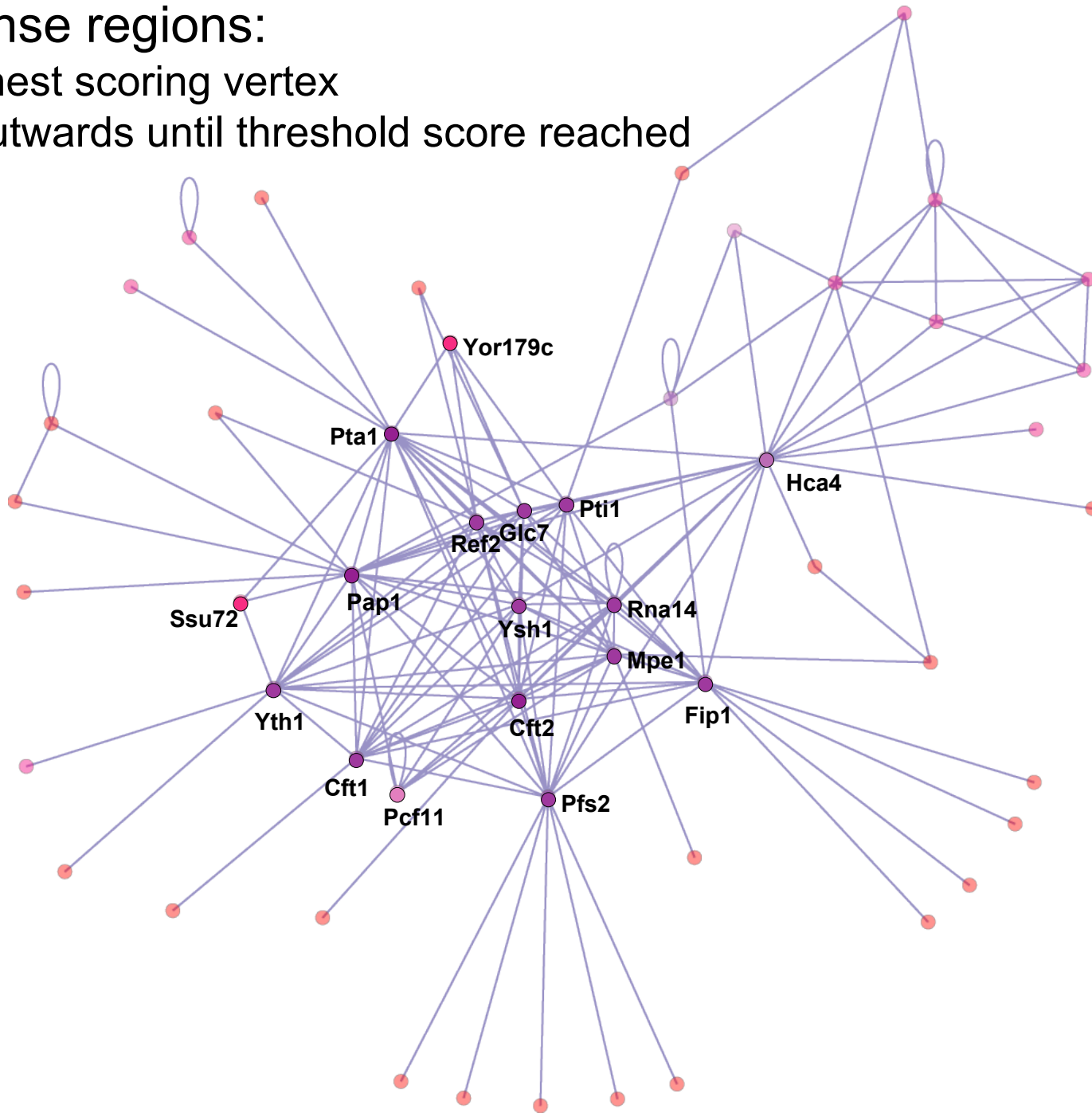
Repeat for entire network



Find dense regions:

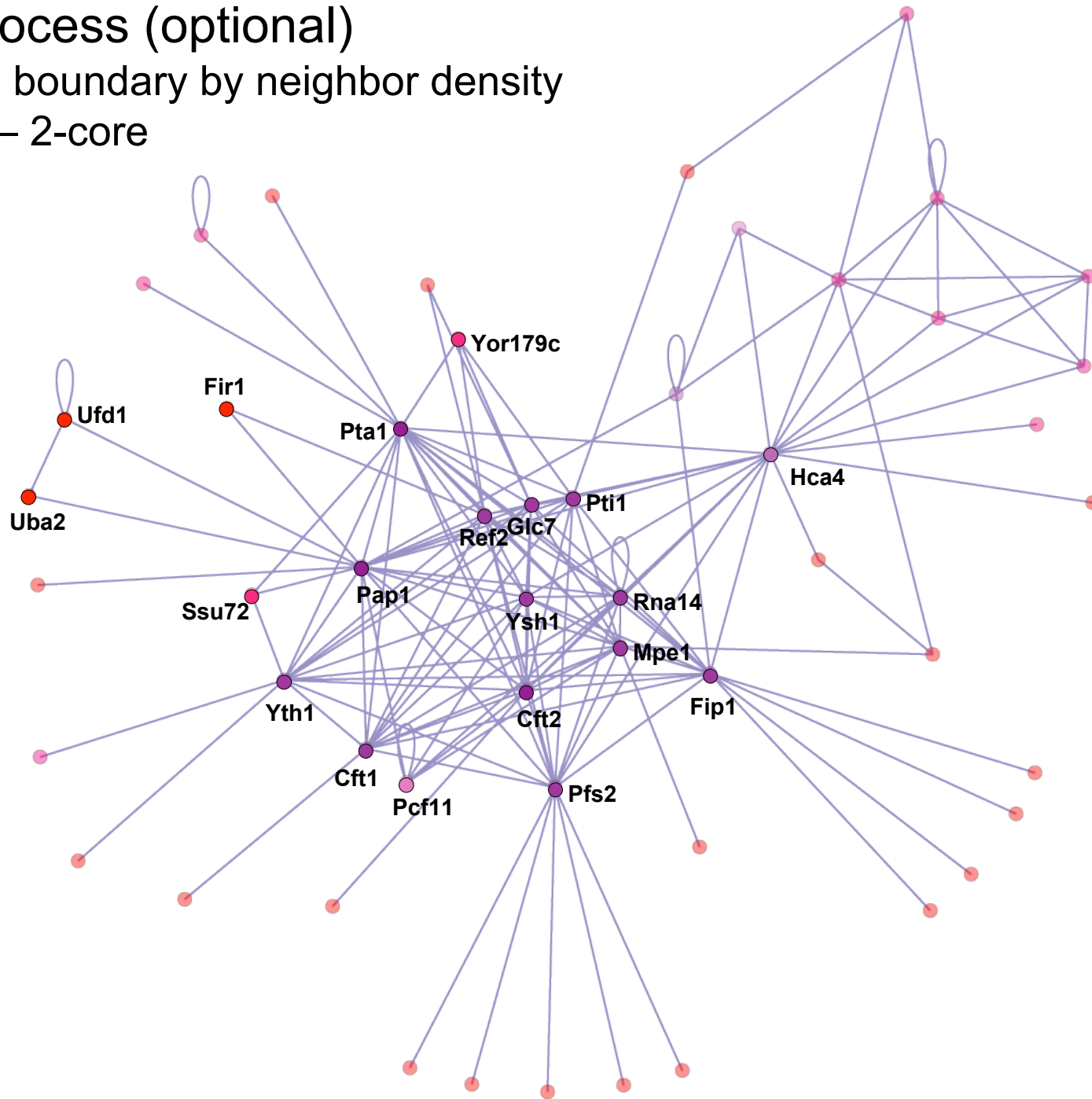
-Pick highest scoring vertex

-'Paint' outwards until threshold score reached

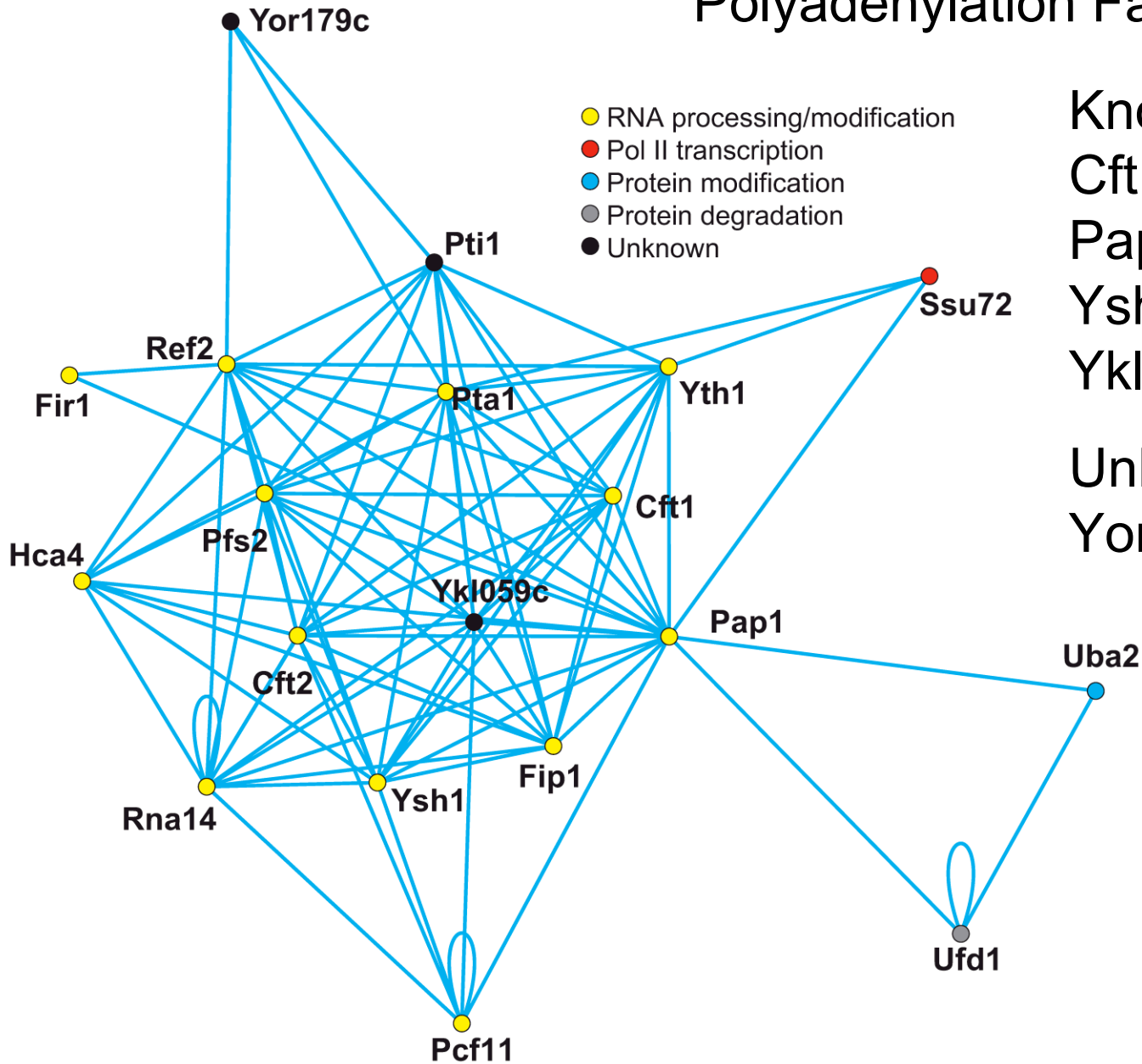


Post-process (optional)

- Fluff the boundary by neighbor density
- Haircut – 2-core



Polyadenylation Factor I Complex



Known:

Cft1, Cft2, Fip1,
Pap1, Pfs2, Pta1,
Ysh1, Yth1 and
Ykl059c

Unknown:

Yor179c and Pti1

Evaluation

- Yeast
- Requires a list of known complexes for comparison:
Gavin *et al.* (221), MIPS (208)

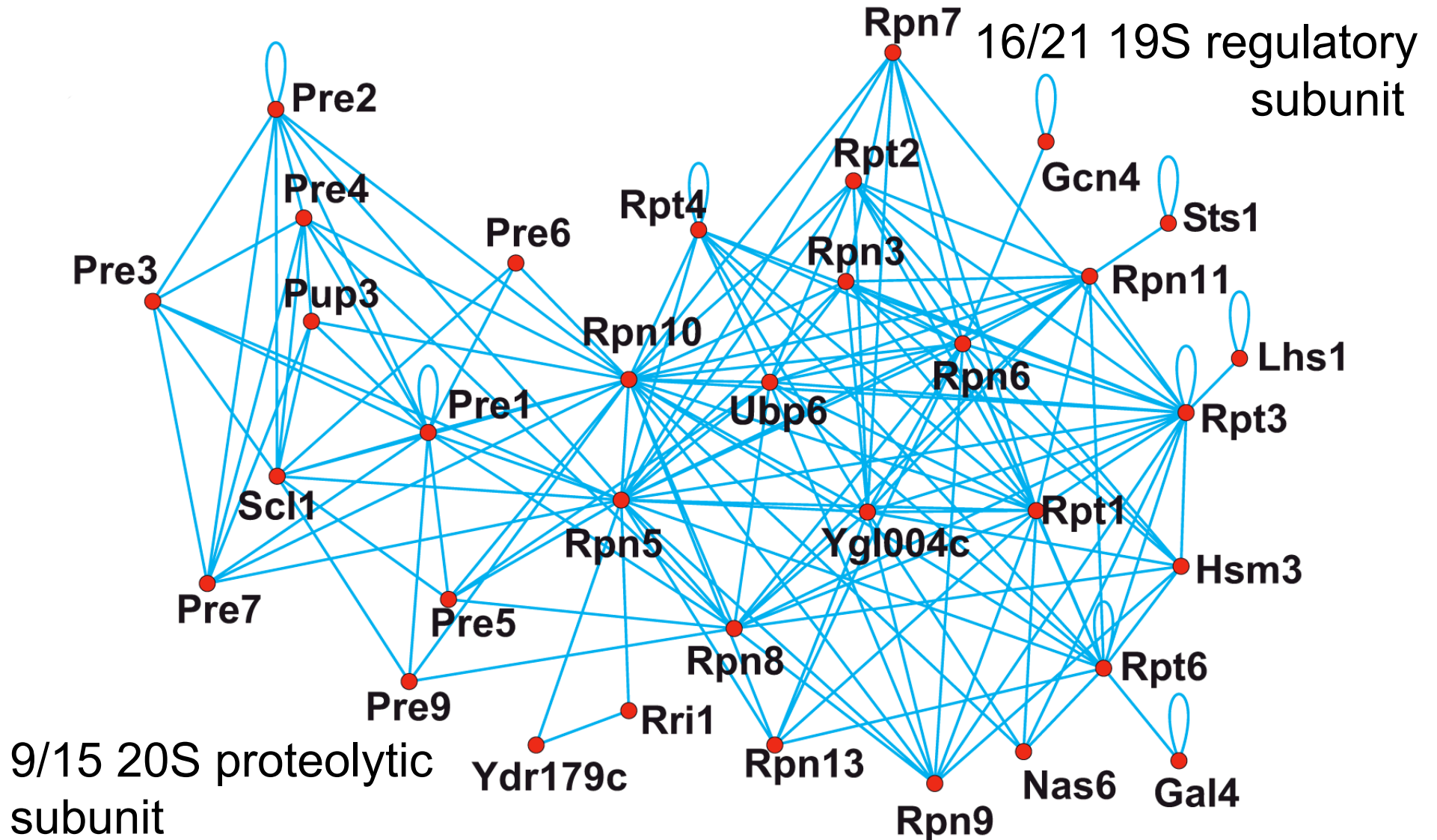
Data Set	Number of Proteins	Number of Interactions	Number of Predicted Complexes	MCODE Complexes Predicted Above $w=0.2$	Matched Benchmark Complexes	Complex Benchmark	Best MCODE Parameters
Gavin Spoke	1363	3225	77	63	88	Gavin	$hFfT\backslash 0.05\backslash 0.05$
Pre HTMS	4379	9088	166	52	64	MIPS	$hTfT\backslash 0.1\backslash 0.2$

- Not perfect, but neither is the data

Application to Yeast Network

- From a list of 15,143 known yeast intx among 4,825 proteins 209 complexes predicted
- 100 random network permutations:
 - Average of 27.4 complexes (SD=4.4)
 - Random complexes ~5x larger
 - Did not match any known complexes
 - Large annotation spread
- Thus, number, size, functional composition unlikely to occur by chance
- Not affected by high number of false positives in high-throughput data sets

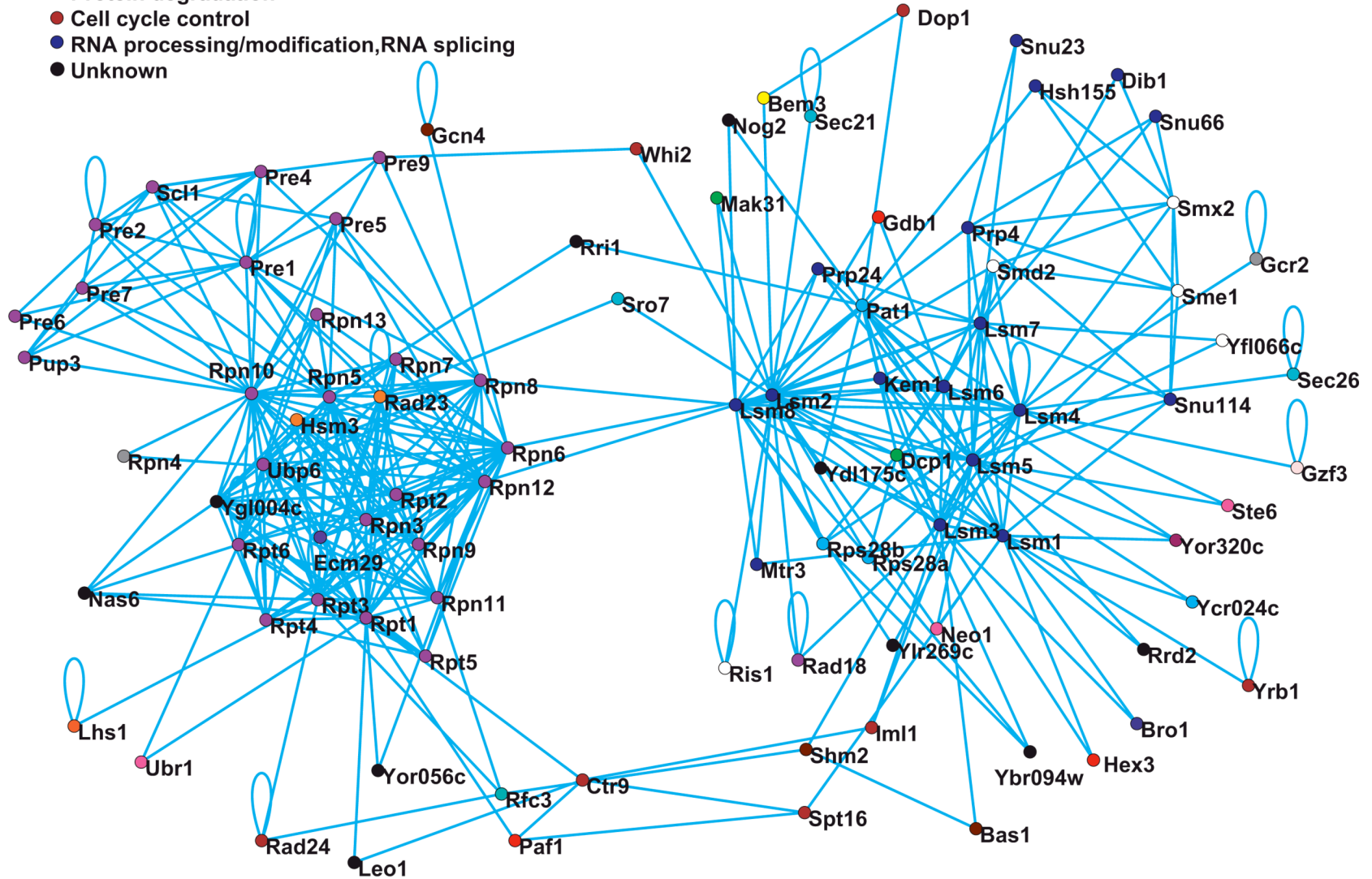
The Yeast 26S Proteasome



Basic structure is evident

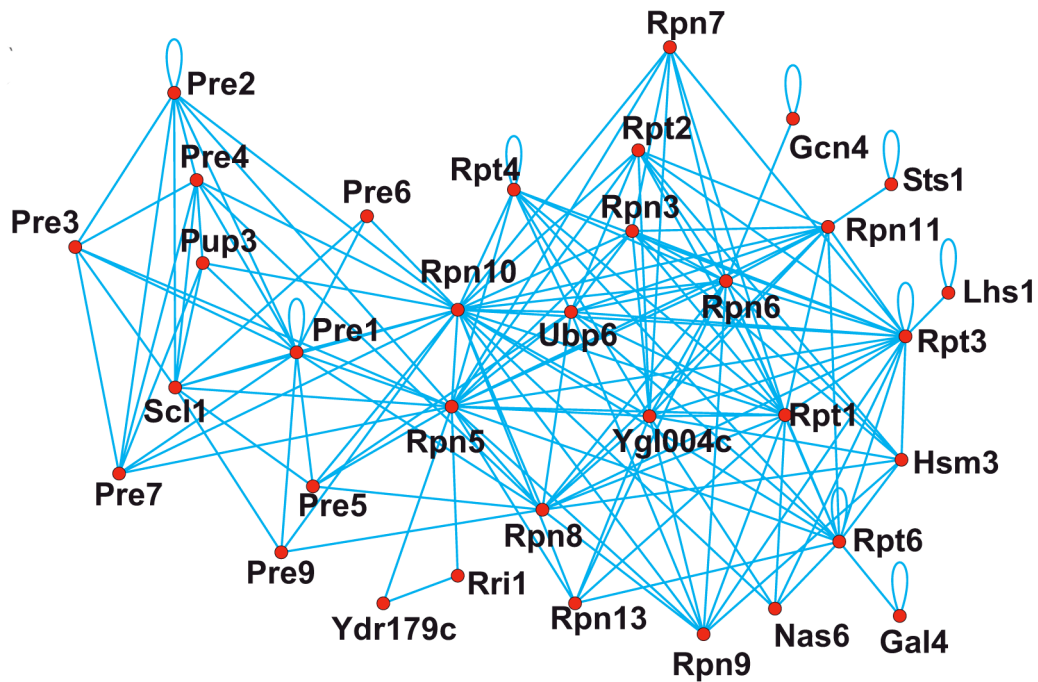
Directed Mode

- Chromatin/chromosome structure
- Protein degradation
- Cell cycle control
- RNA processing/modification, RNA splicing
- Unknown

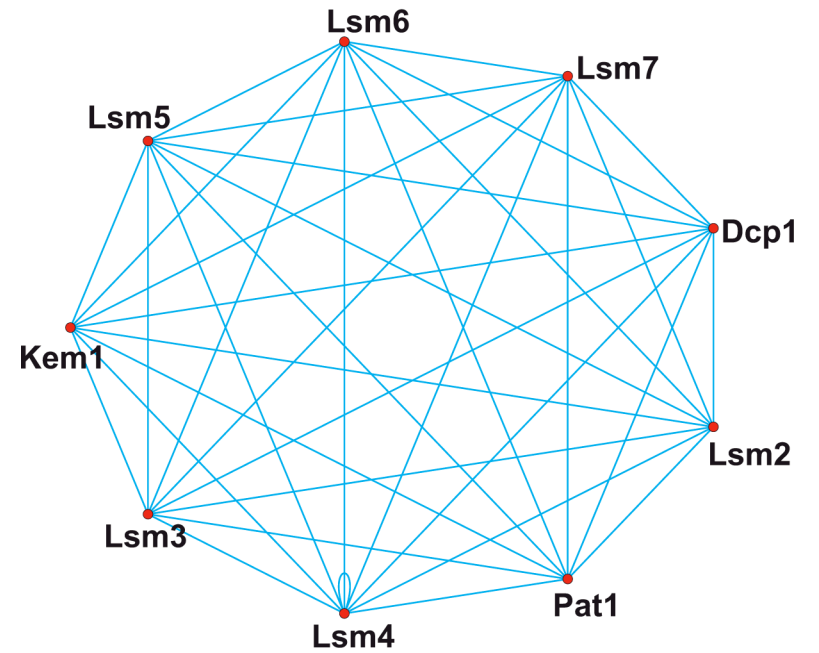


Directed Mode - Split

26S proteasome



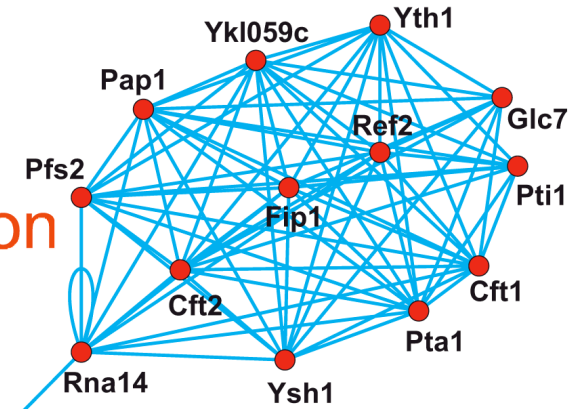
Lsm mRNA Splicing



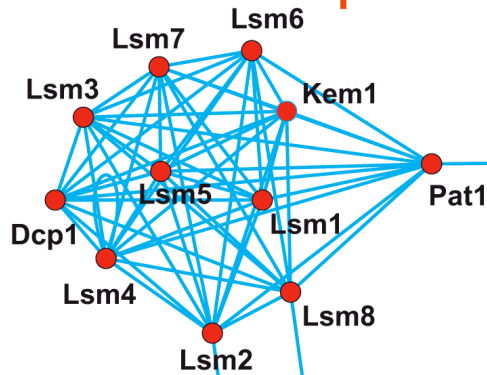
Shared Subunits

Functional Connections Between Complexes

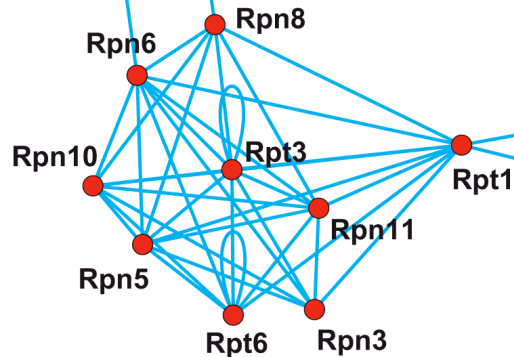
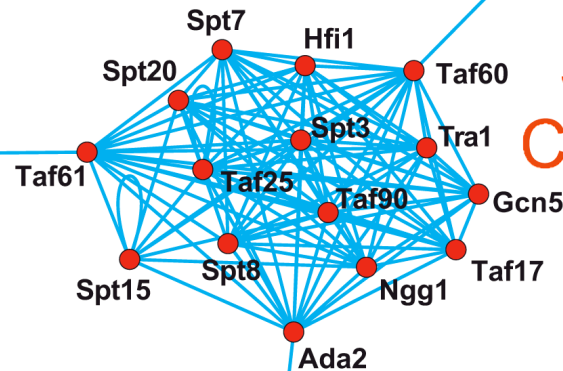
mRNA
Cleavage/
Polyadenylation



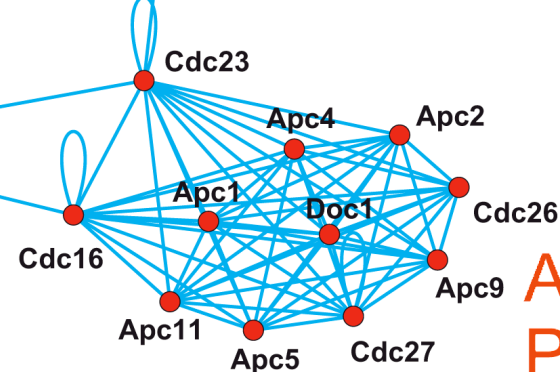
Lsm Complex



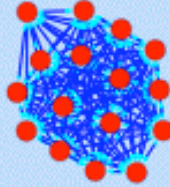
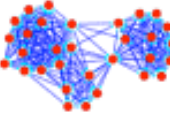
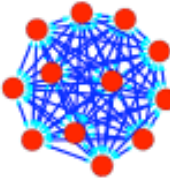
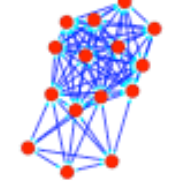
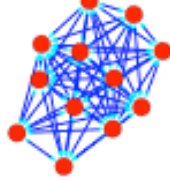
SAGA
Complex



19S Proteasome



Anaphase
Promoting
Complex

MCODE Results Summary				
Rank	Score	Size	Names	Complex
1	7.25	16,116	YGR232W, YDL007W, YKL145W, YFR052W, YFR004W, YLR421C, YOR261C, YDL147W, YDR427W, YHR200W, YER021W, YOR117W, YDL097C, YOR259C, YPR108W, YDR394W	
2	6.387	31,198	YPL093W, YBL004W, YOR272W, YNL110C, YKL009W, YFL002C, YOL077C, YPL126W, YIL035C, YLR409C, YLR129W, YOR061W, YKR060W, YCR057C, YDR449C, YOR039W, YJL109C, YPL012W, YGR103W, YLR449W, YOR206W, YKL014C, YLL008W, YKL172W, YNL002C, YLR002C, YGL111W, YOL041C, YGL019W, YOR145C, YPR016C	
3	5.417	12,65	YGL011C, YOL038W, YPR103W, YMR314W, YBL041W, YOR362C, YER012W, YJL001W, YML092C, YGR253C, YER094C, YGR135W	
4	5	15,75	YPL043W, YMR290C, YER006W, YKR081C, YDR496C, YDL031W, YNL061W, YNL132W, YLR222C, YLR197W, YMR049C, YHR052W, YJL069C, YKL099C, YDL014W	
5	5	12,60	YPR187W, YPR010C, YPR110C, YNL248C, YOR341W, YNR003C, YKL144C, YOR207C, YPR190C, YNL113W, YOR116C, YBR154C	

Create a new child network.

Advantages

- Compared to other clustering algorithms
- Directed mode
- Complex connectivity mode
- Does not force all data points into clusters
- Fast: ~ a few seconds for 15,000 protein network on laptop
- Makes visualization of large networks more manageable

Conclusions

- Initial step in taking advantage of current purely qualitative connectivity information
- Future networks need to have more information about time, space and quality (from e.g. interaction databases) → p-value weights on edges
- <http://cbio.mskcc.org/~bader/software/mcode/>

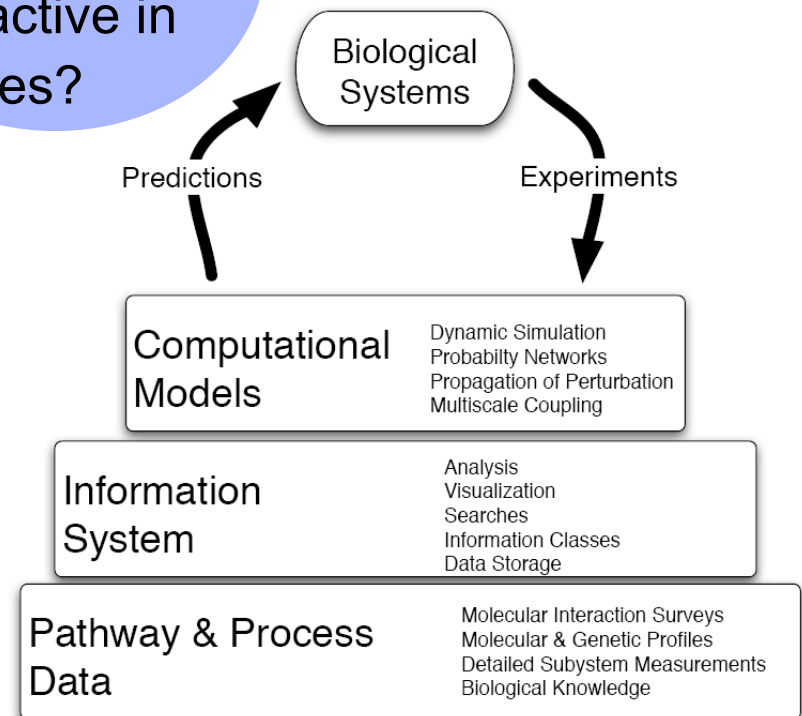
How do cellular biological systems composed of interacting parts work?
Can we accurately predict biologically relevant interactions?

How do genome sequence changes underlying disease or through evolution affect the molecular network in the cell?

What biological processes are active in different cell types and stages?



Cary, Bader, Sander, FEBS Letters 579 (2005) 1815-20

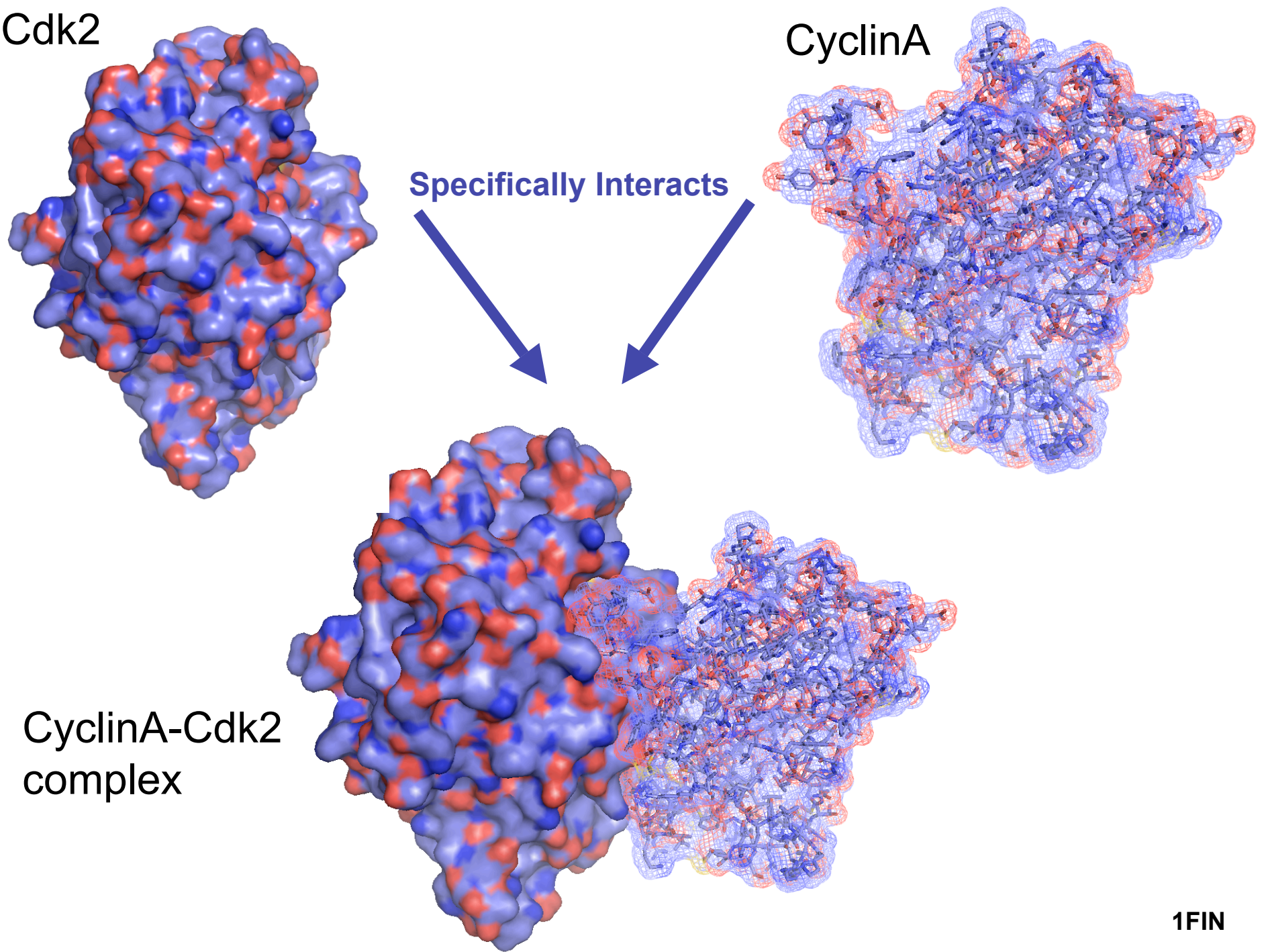


Cdk2

CyclinA

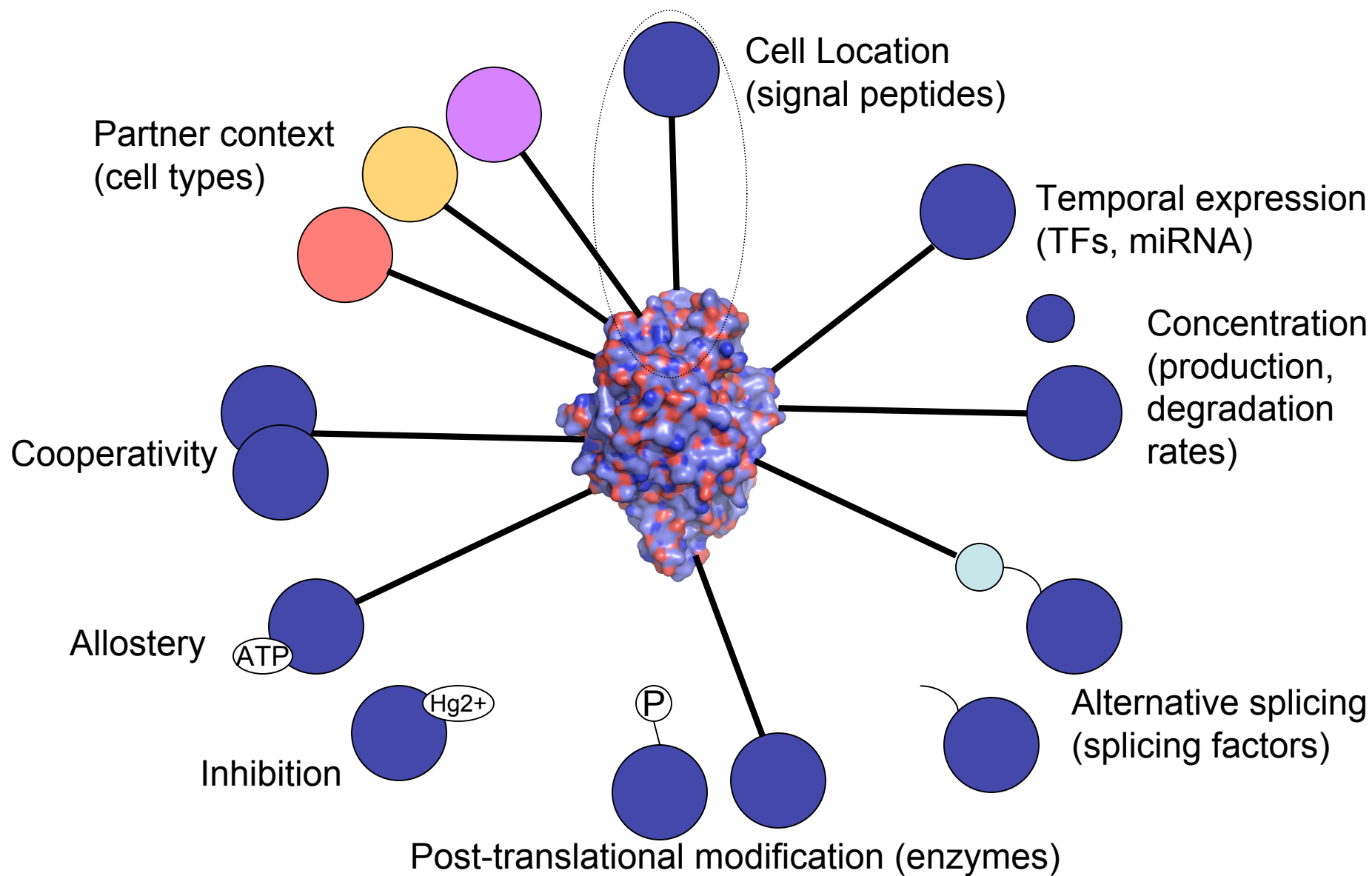
Specifically Interacts

CyclinA-Cdk2
complex



1FIN

Interaction Regulation Modes



Acknowledgements

MCODE

Chris Hogue

Data

Boone lab, Tyers lab,
MDSP, SLRI, Fields
lab, Cesareni lab

Bader Lab

Anastasija Baryshnikova
Kevin Ha
Vuk Pavlovic
Chris Tan
Graham Cromar

Cytoscape

Trey Ideker (UCSD)

Ryan Kelley, Kei Ono, Mike Smoot,
Peng Liang Wang
(Nerius Landys, Chris Workman, Mark
Anderson, Nada Amin, Owen Ozier,
Jonathan Wang)

Lee Hood (ISB)

Sarah Killcoyne
(Iliana Avila-Campillo, Rowan
Christmas, Andrew Markiel, Larissa
Kamenkovich, Paul Shannon)

Benno Schwikowski (Pasteur)

Melissa Cline, Tero Aittokallio

Chris Sander (MSKCC)

Ethan Cerami, Ben Gross (Robert
Sheridan)

Annette Adler (Agilent)

Allan Kuchinsky, Aditya Vailaya, Mike
Creech

Bruce Conklin (UCSF)

Alex Pico, Kristina Hanspers

