http://www.pathwaycommons.org

Pathway Commons

A public library of biological pathways on the semantic web

June.13 2007 - NETTAB, Pisa

Gary Bader
University of Toronto

http://baderlab.org

Chris Sander
MSKCC, New York

http://cbio.mskcc.org

C
C
B
R

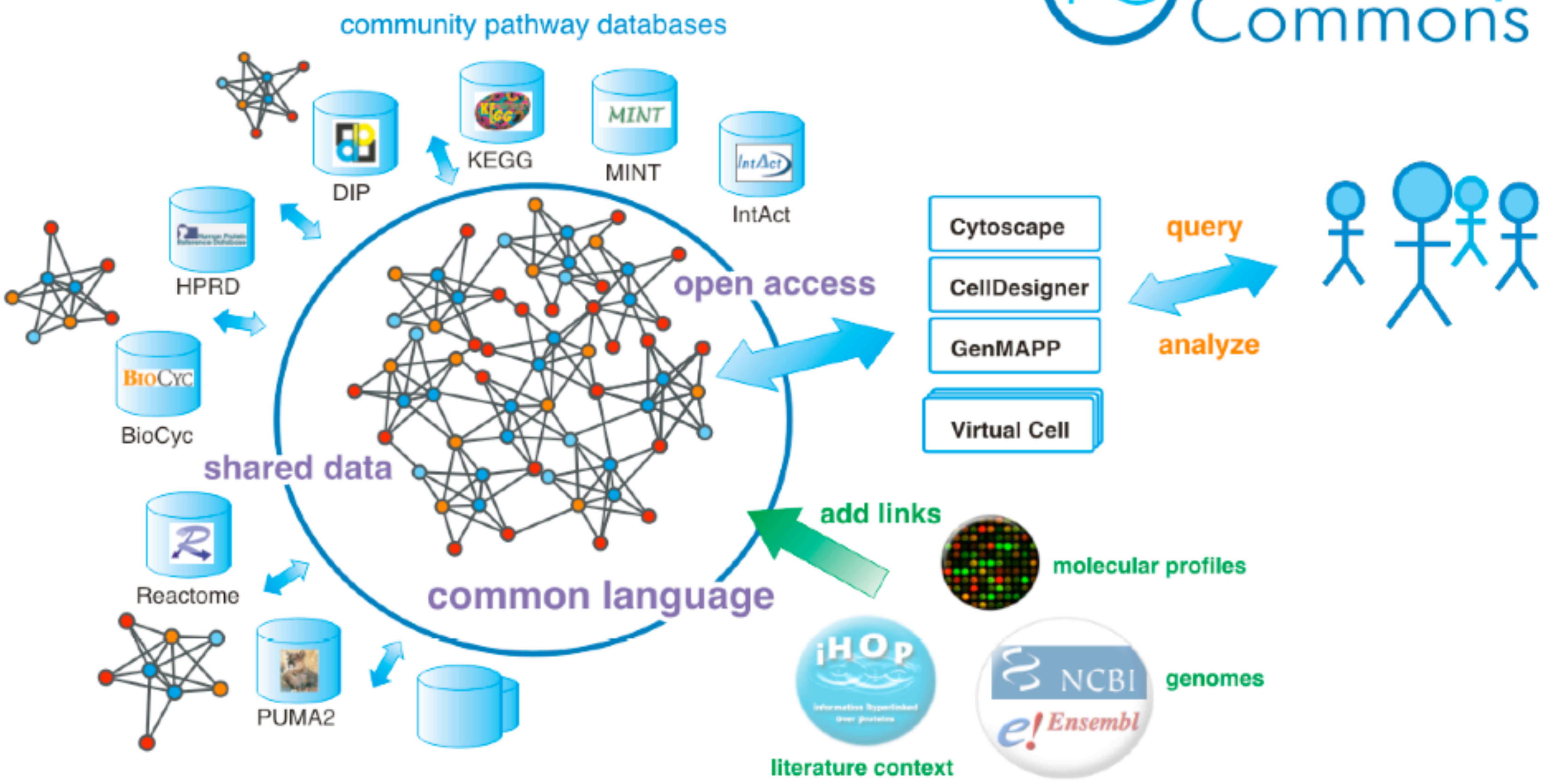
MEMORIAL SLOAN-KETTERING
CANCER CENTER

computational biology center

Aim: Convenient Access to Pathway Information

http://www.pathwaycommons.org

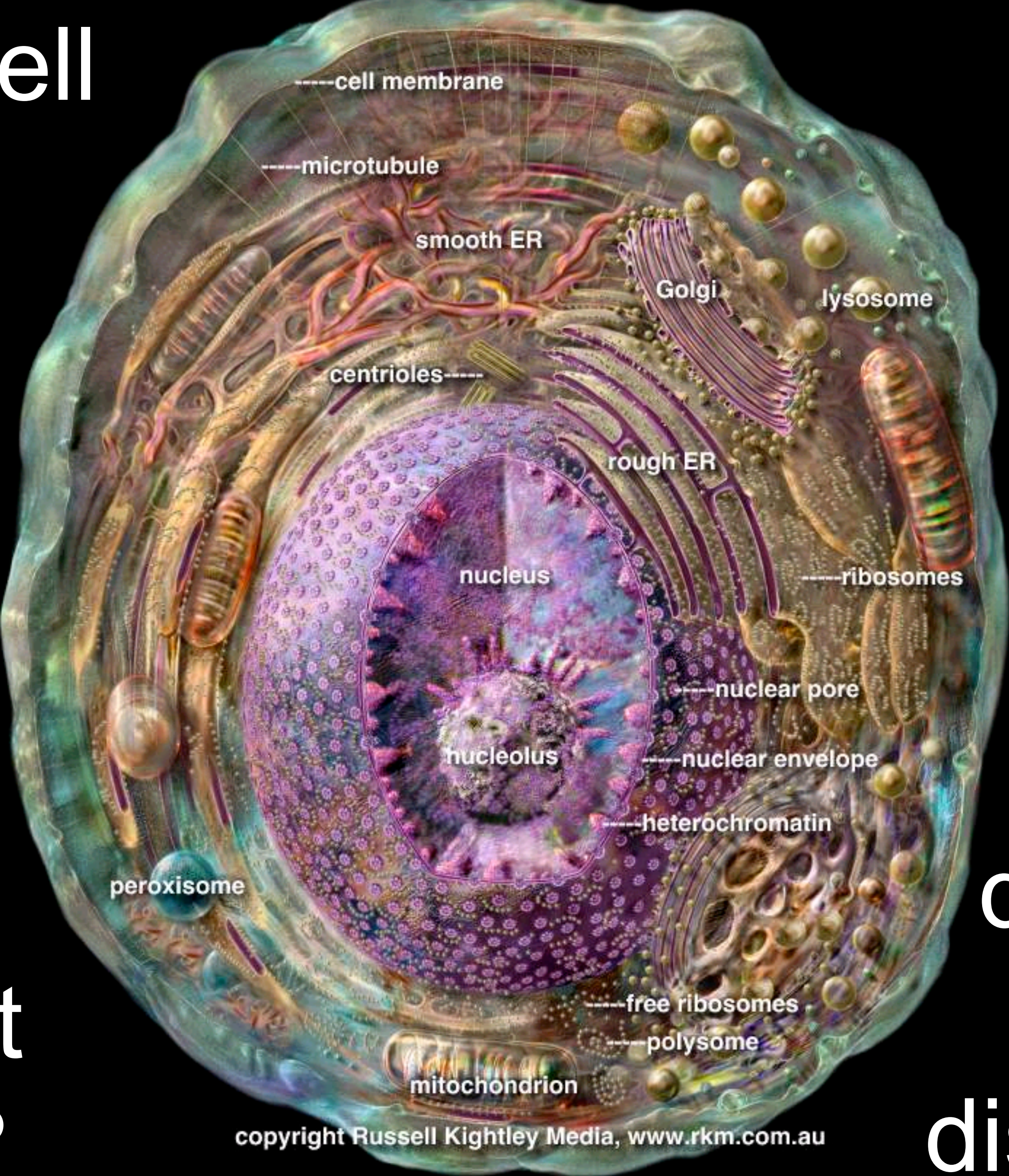
PC Pathway Commons



Facilitate creation and communication of pathway data
Aggregate pathway data in the public domain
Provide easy access for pathway analysis

Long term: Converge to integrated cell map

The Cell



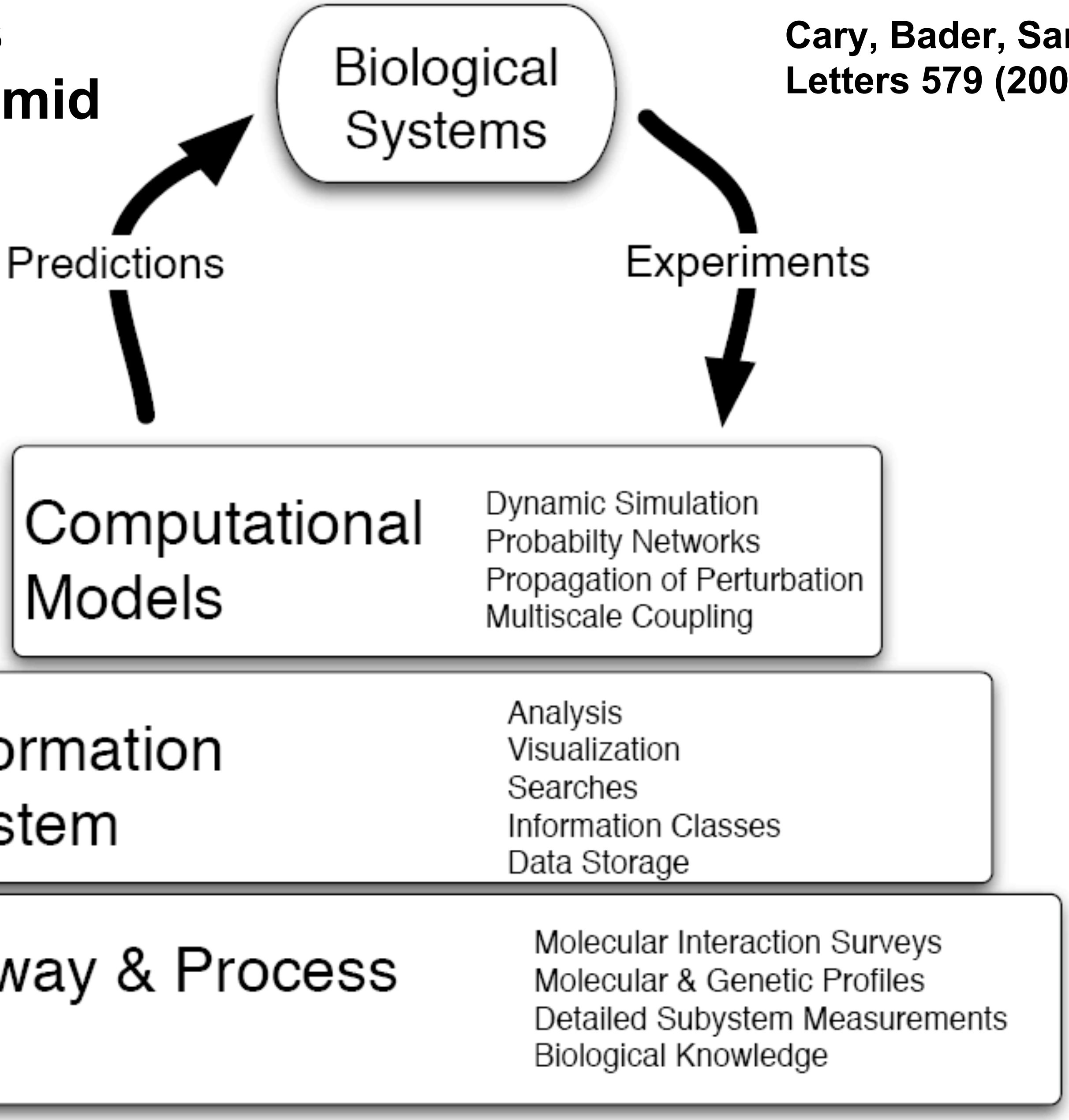
copyright Russell Kightley Media, www.rkm.com.au

How does it work?

How does it fail in disease?

The Systems Biology Pyramid

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



Biological Systems

Predictions

Experiments

Computational Models

- Dynamic Simulation
- Probabilty Networks
- Propagation of Perturbation
- Multiscale Coupling

Information System

- Analysis
- Visualization
- Searches
- Information Classes
- Data Storage

Pathway & Process Data

- Molecular Interaction Surveys
- Molecular & Genetic Profiles
- Detailed Subystem Measurements
- Biological Knowledge

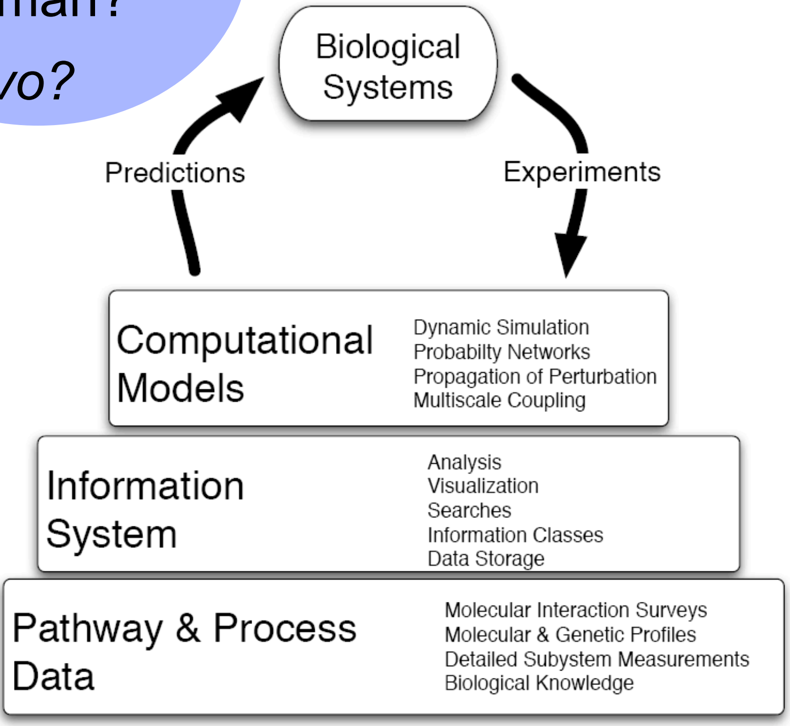
How are biological networks in the cell encoded in the genome?

Can we accurately predict biologically relevant interactions from a genome?

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

Can we predict how well model pathways or phenotypes will translate to human?

Can we design networks *de novo?*



Biological Systems

Experiments

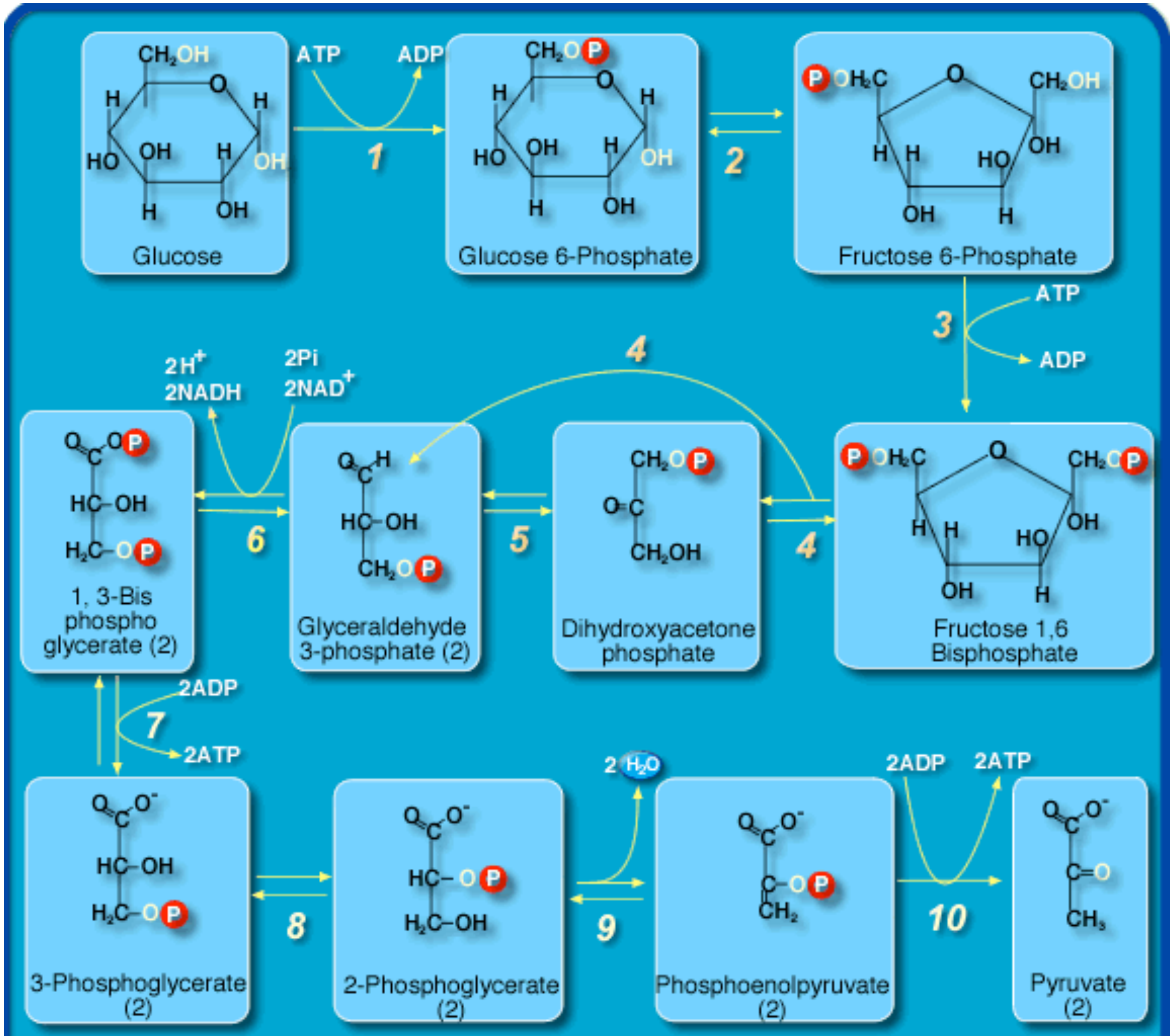
Predictions

Computational Models
- Dynamic Simulation
- Probabilty Networks
- Propagation of Perturbation
- Multiscale Coupling

Information System
- Analysis
- Visualization
- Searches
- Information Classes
- Data Storage

Pathway & Process Data
- Molecular Interaction Surveys
- Molecular & Genetic Profiles
- Detailed Subystem Measurements
- Biological Knowledge

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



ENZYMES ● **Preparatory phase** ● **Payoff phase**

1 **Hexokinase**
2 **Glucose Phosphate Isomerase**
3 **Phosphofructokinase**
4 **Fructose diphosphate aldolase**
5 **Triose phosphate Isomerase**
6 **Glyceraldehyde Phosphate Dehydrogenase**
7 **Phosphoglycerate Kinase**
8 **Phosphoglyceromutase**
9 **Enolase**
10 **Pyruvate Kinase**

Signaling Pathway

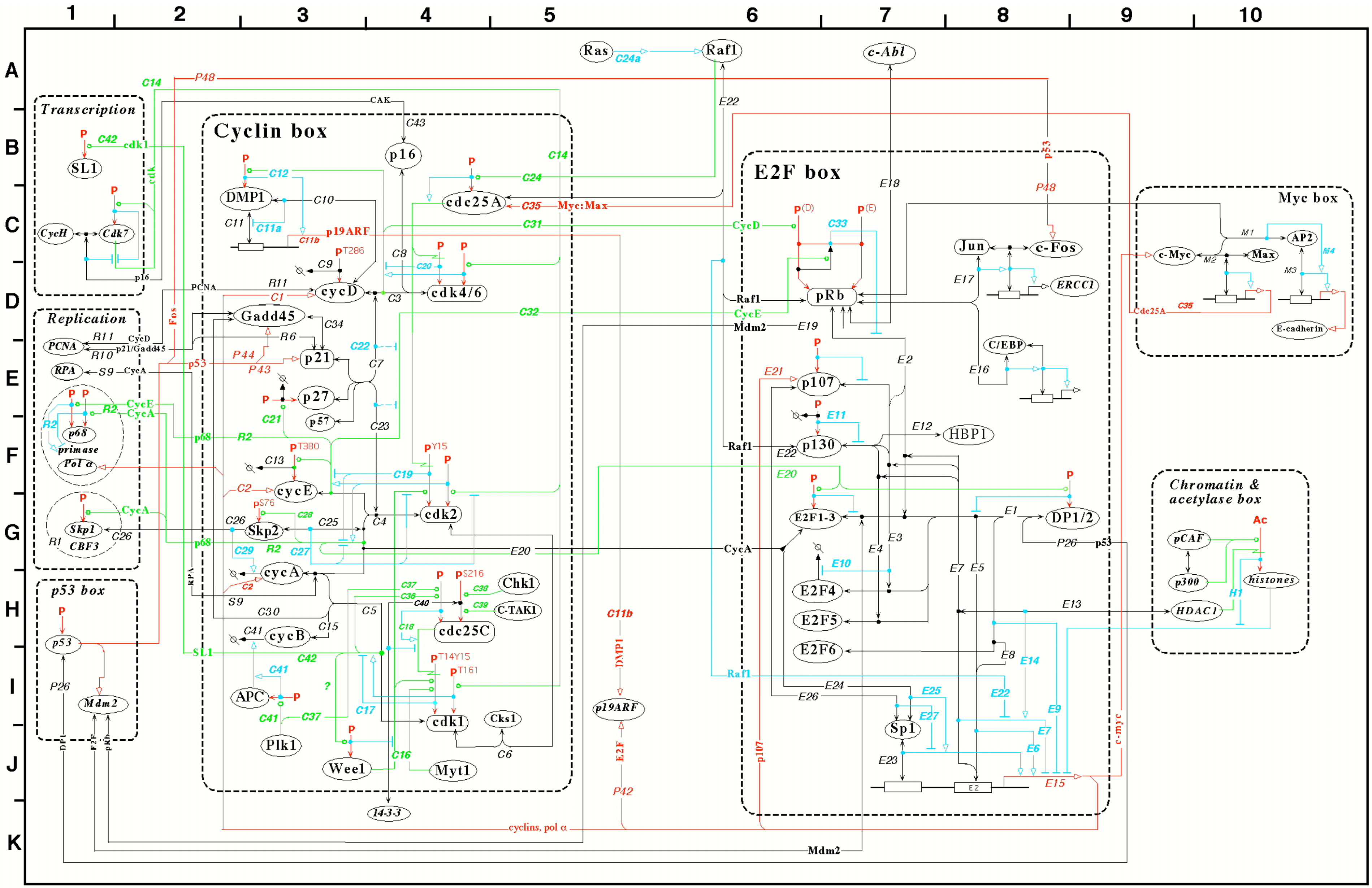


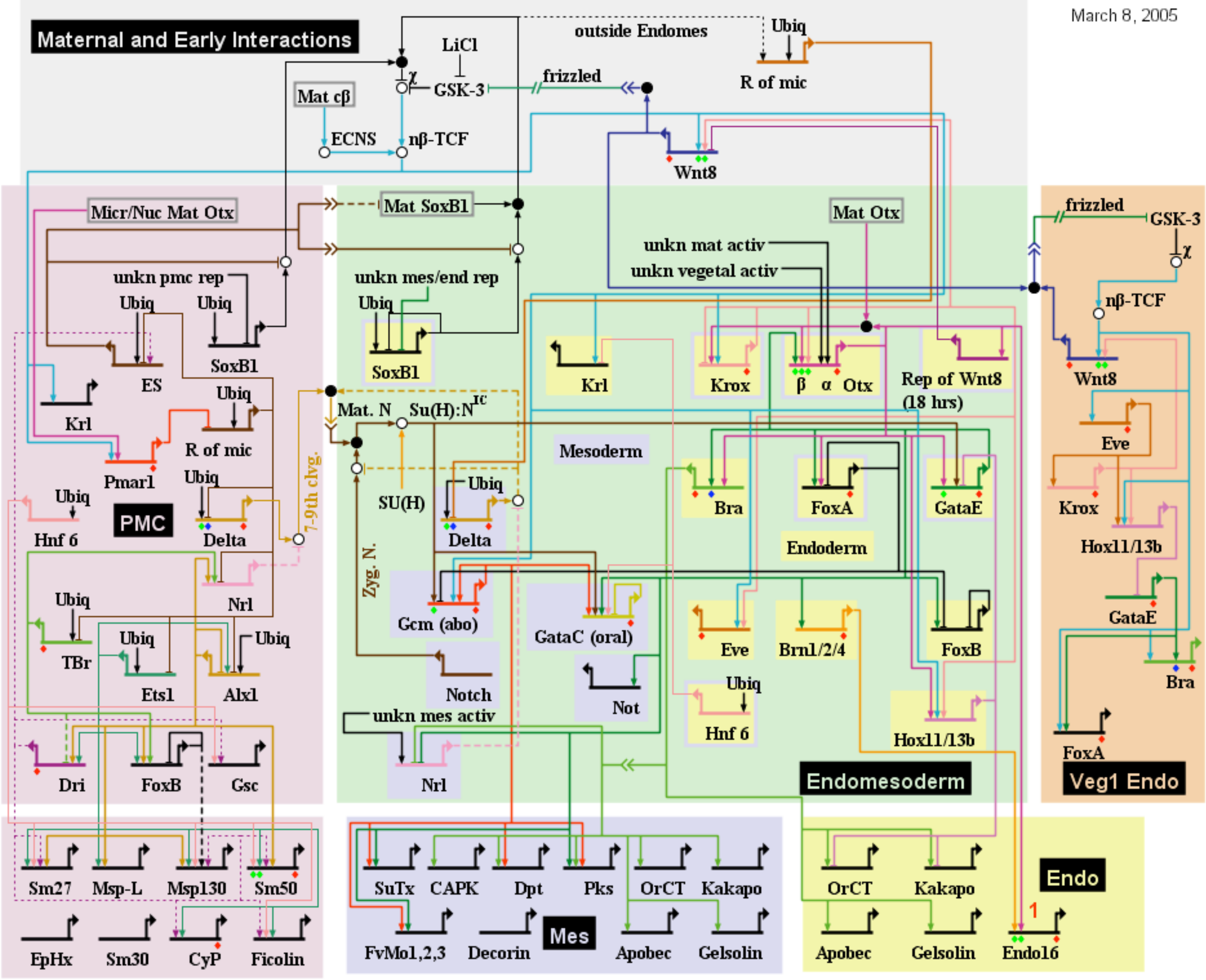
Figure 6A: The Cyclin - E2F cell cycle control system (version 3a - June 8, 1999)

http://discover.nci.nih.gov/kohnk/interaction_maps.html

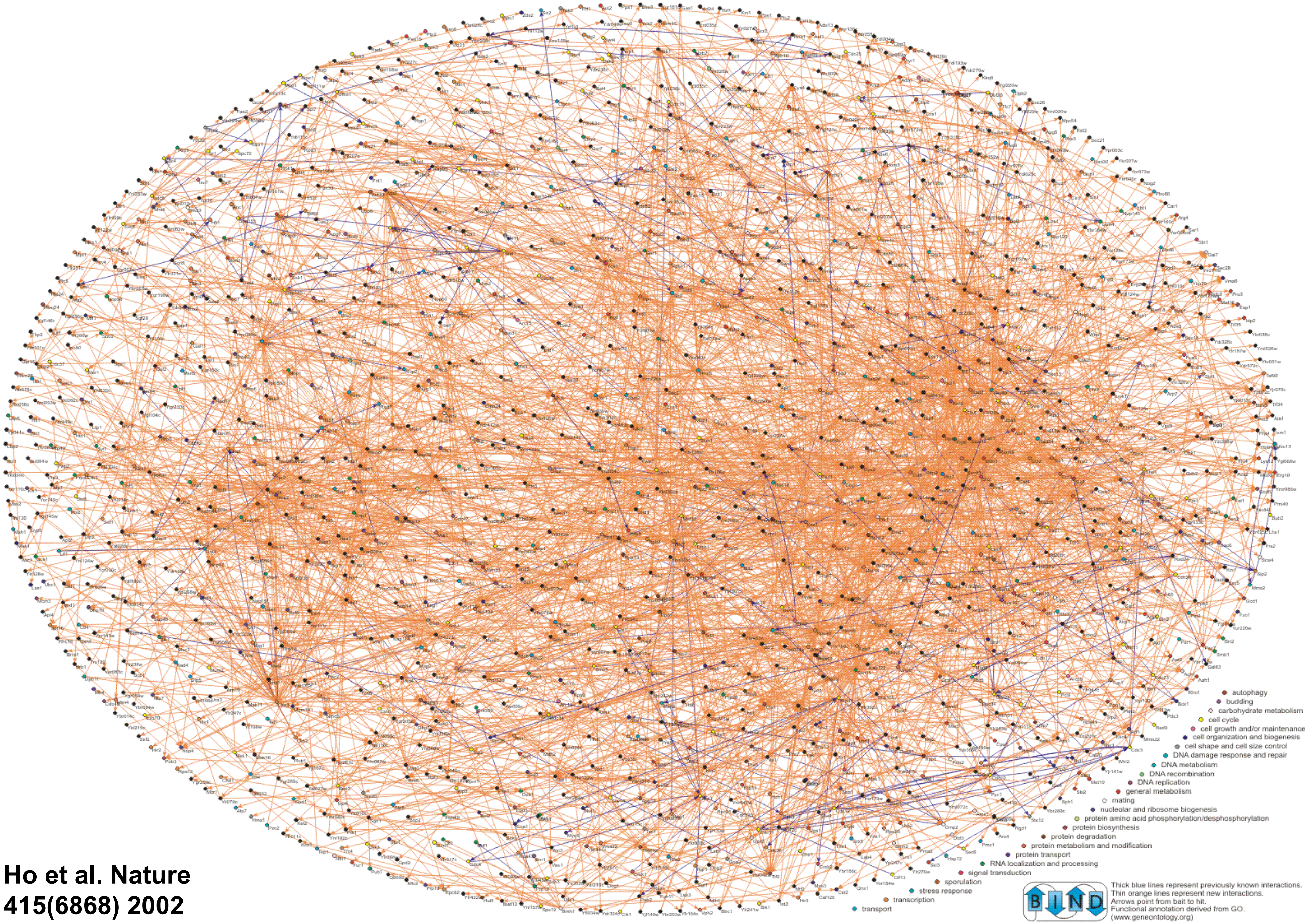
MIMC&E 2k

Endomesoderm Specification to 30 Hours

March 8, 2005



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

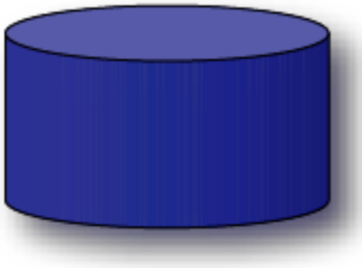


Thick blue lines represent previously known interactions.
Thin orange lines represent new interactions.
Arrows point from bait to hit.
Functional annotation derived from GO.
(www.geneontology.org)

Ho et al. Nature 415(6868) 2002

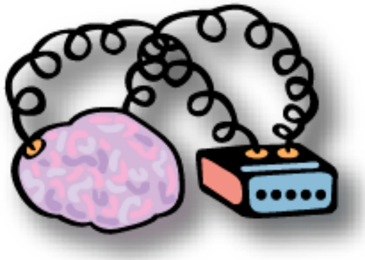
Pathway Information

- Databases
  - Fully electronic
  - Easily computer readable

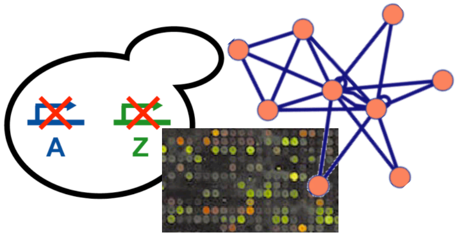


- Literature
  - Increasingly electronic
  - Human readable

- Biologist’s brains
  - Richest data source
  - Limited bandwidth access



- Experiments
  - Basis for models



Pathway Databases

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
GOLD.db - 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
Monod - Modeler's Notebook and Datastore
MPB - Metabolic Pathways of Biochemistry
MRAD - Metabolic Reaction Analysis Database
NetBiochem - Medical Biochemistry Resource
Nicholson Minimaps - IUBMB-Nicholson Minimaps
PathArt - Pathway Articulator
PathDB - Pathways Database
PATIKA - 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?

Protein-Protein Interactions
3DID - 3D interacting domains
ABCdb - Archaea and Bacteria ABC transporter database
AfCS - Alliance for Cellular Signaling Molecule Pages Database
AllFuse - Functional Associations of Proteins in Complete Genomes
ASEdb - Alanine Scanning Energetics Database
ASPD - Artificial Selected Proteins/Peptides Database
BID - Binding Interface Database
BIND - Biomolecular Interaction Network Database
BindingDB - The Binding Database
BRITE - Biomolecular Relations in Information Transmission and Expression
CTDB - Calmodulin Target Database
DDIB - Database of Domain Interactions and Bindings
DIP - Database of Interacting Proteins
Doodle - Database of oligomerization domains from lambda experiments
DopaNet - DopaNet
DRC - 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
HIVMID - HIV Molecular Immunology Database
HP-DPI - Helicobacter pylori Database of Protein Interactomes
HPID - Human Protein Interaction Database
HPRD - Human Protein Reference Database
HSV1 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
KDBI - 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
OPHID - The Online Predicted Human Interaction Database
Pathcalling - Curagen Pathcalling
PathDB - Pathways Database
PDZBase - PDZ domain protein-protein interaction database
Phospho.ELM - Post-translational phosphorylation database
PhosphoBase - Database of phosphorylation sites
PhosphoSite - Cell Signaling Technology's PhosphoSite Database
PIBASE - PIBASE
PIMdb - Drosophila Protein Interaction Map Database
PIMRider - Protein Interaction Map - Hybrigenics
PINdb - Proteins Interacting in the Nucleus database
POINT - Prediction of Interactome
PombePD - Schizosaccharomyces pombe Proteome Database
PPID - Protein-Protein Interaction Database
PPIV (FANTOM) - RIKEN FANTOM Protein Protein Interaction Viewer
PQS - 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
PSIbase - Protein Structural Interactome Database
PubGene - PubGene
ResNet - ResNet
S/MARt_DB - The S/MAR transaction DataBase
Scansite - Scansite
SPiD - Subtilis Protein interaction Database
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

Signaling Pathways
AfCS - 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
CSNDB - Cell Signaling Networks Database
DOQCS - Database of Quantitative Cellular Signaling
DSM - Dynamic Signaling Maps
eMIM - Electronic Molecular Interaction Map
GeneNet - Genetic Networks
GenMAPP - Gene MicroArray Pathway Profiler
GOLD.db - Genomics of Lipid-associated Disorders
GON - Genomic Object Net
INOH - Integrating Network Objects with Hierarchies
MetaCore - MetaCore pathway database
Monod - Modeler's Notebook and Datastore
PANTHER - PANTHER
PathArt - Pathway Articulator
PathDB - Pathways Database
Pathways Knowledge Base - Ingenuity Pathways Knowledge Base
PATIKA - Pathway Analysis Tool for Integration and Knowledge Acquisition
PDS - Pathways Database System
PhosphoSite - Cell Signaling Technology's PhosphoSite Database
PID - CMAP Pathway Interaction Database
ProcessDB - ProcessDB
Reactome - Reactome KnowledgeBase
ResNet - ResNet
ROSPath - Reactive Oxygen Species related Signaling Pathway
SBML Model Repository - SBML Model Repository
Sentra - Sentra Signal Transduction Database
SigPath - Signaling Pathway Information System
SPAD - Signaling Pathway Database
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

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
KMIM - 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 - IUBMB-Nicholson Minimaps
PharmGKB - The Pharmacogenetics and Pharmacogenomics Knowledge Base
PID - CMAP Pathway Interaction Database
SPAD - Signaling Pathway Database
TRMP - Therapeutically Relevant Multiple Pathways Database

Transcription Factors / Genetic Regulatory Networks
COMPEL - Database on Composite Regulatory Elements
DBTBS - Database of Bacillus subtilis Promoters and Transcription Factors
DPInteract - 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

Protein-Compound Interactions
CLiBE - Computed Ligand Binding Energy
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
Relibase - Protein-ligand database query tool
ResNet - ResNet
TTD - Therapeutic Target Database

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

Protein Sequence Focused
ASC - Active Sequences Collection
InBase - The Intein Database
MEROPS - MEROPS Peptidase Database
NRR - Nuclear Receptor Resource
ORDB - Olfactory Receptor Database
PANTHER - PANTHER
PhosphoBase - Database of phosphorylation sites
REBASE - Restriction Enzyme Database
SCOPEC - Database of protein catalytic domains
Sentra - Sentra Signal Transduction Database
TGDB - Tumor Gene Database
XPD - Phosphorylation Site Database

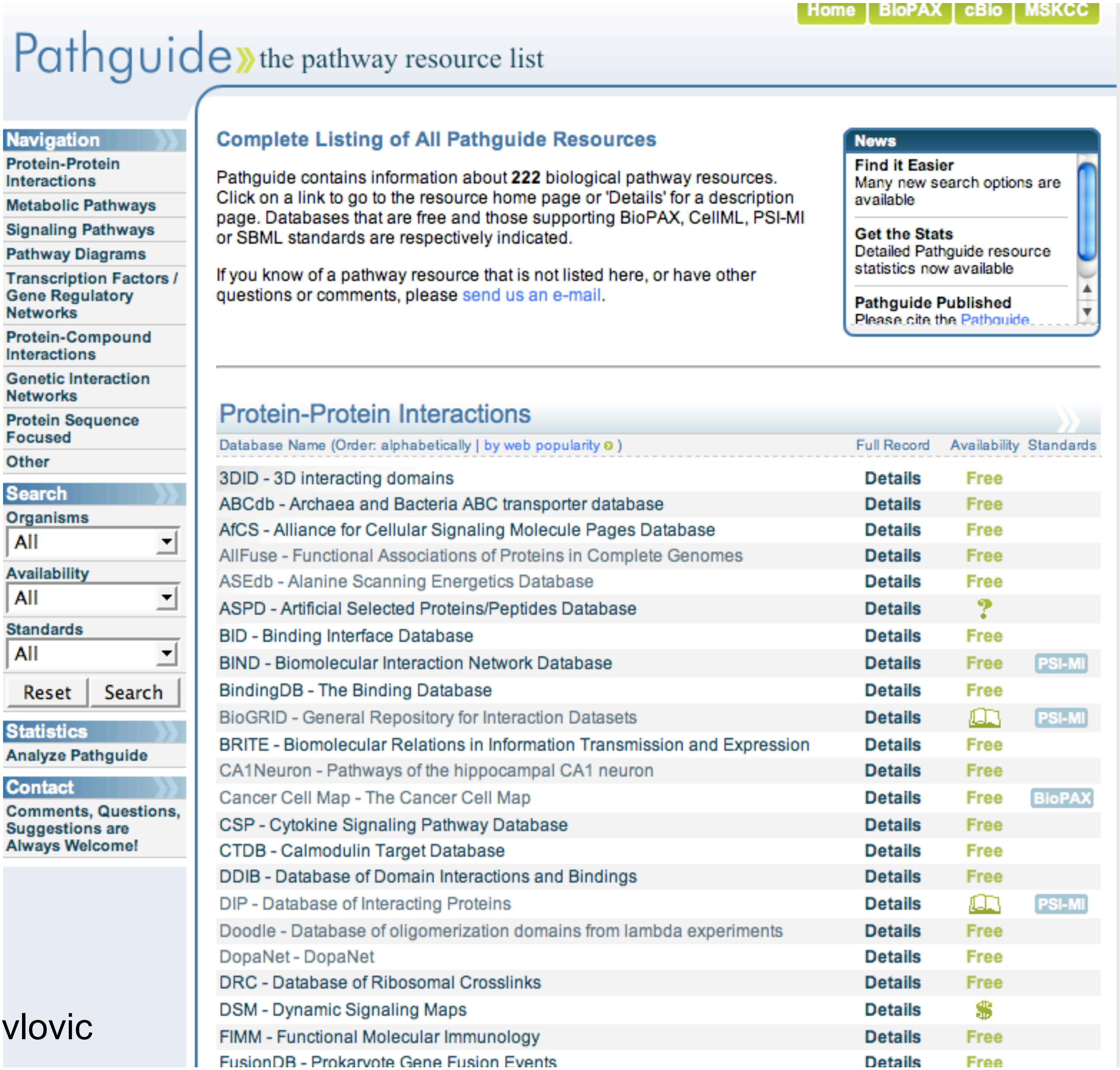
Other
SELEX_DB - Randomized DNA/RNA sequence database
AARSDB - Aminoacyl-tRNA Synthetase Database
MedGene - MedGene
AANT - Amino Acid-Nucleotide Interaction Database
ProNIT - Thermodynamic Database for Protein-Nucleic Acid Interactions
TCDB - Transport Classification Database
TransportDB - TransportDB
DPIDB - DNA-Protein Interaction Database
iHOP - Information Hyperlinked Over Proteins
PRID - Protein-RNA Interaction Database
Precise - Predicted and Consensus Interaction Sites in Enzymes

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)

http://pathguide.org



Pathguide » the pathway resource list

Organisms: All

Availability: All

Standards: All

Complete Listing of All Pathguide Resources

Pathguide contains information about **222** biological pathway resources. Click on a link to go to the resource home page or 'Details' for a description page. Databases that are free and those supporting BioPAX, CellML, PSI-MI or SBML standards are respectively indicated.

If you know of a pathway resource that is not listed here, or have other questions or comments, please send us an e-mail.

News

Find it Easier
Many new search options are available

Get the Stats
Detailed Pathguide resource statistics now available

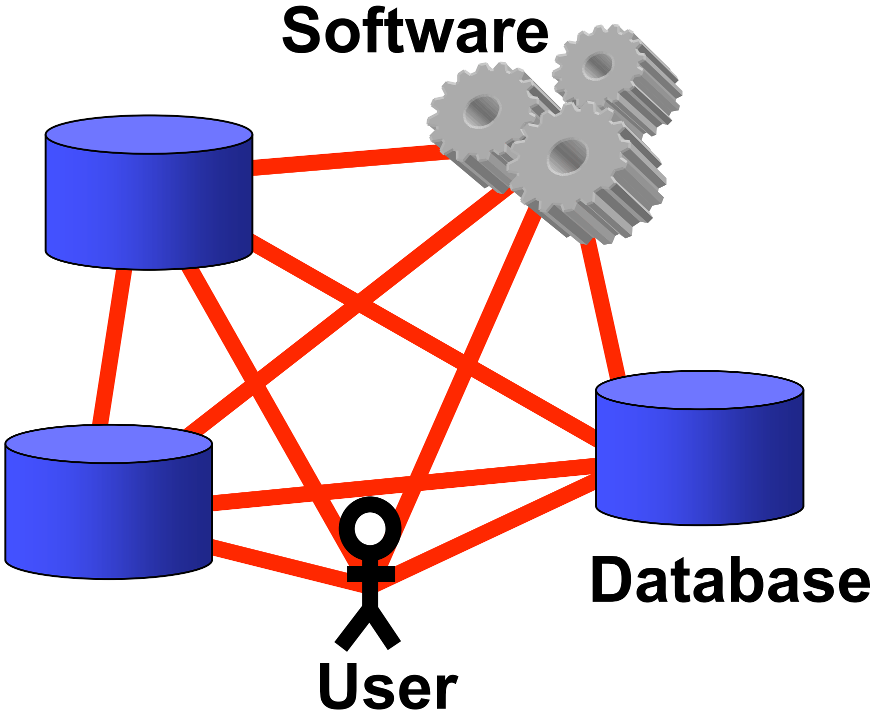
Pathguide Published
Please cite the Pathguide

Protein-Protein Interactions

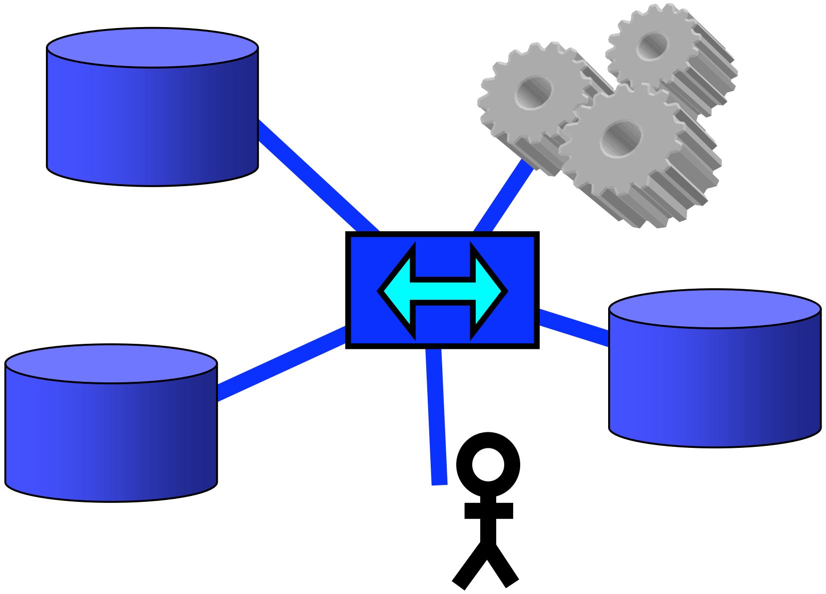
Database Name (Order: alphabetically | by web popularity)	Full Record	Availability	Standards
3DID - 3D interacting domains	Details	Free	
ABCdb - Archaea and Bacteria ABC transporter database	Details	Free	
AfCS - Alliance for Cellular Signaling Molecule Pages Database	Details	Free	
AllFuse - Functional Associations of Proteins in Complete Genomes	Details	Free	
ASEdb - Alanine Scanning Energetics Database	Details	Free	
ASPD - Artificial Selected Proteins/Peptides Database	Details	?	
BID - Binding Interface Database	Details	Free	
BIND - Biomolecular Interaction Network Database	Details	Free	PSI-MI
BindingDB - The Binding Database	Details	Free	
BioGRID - General Repository for Interaction Datasets	Details	📖	PSI-MI
BRITE - Biomolecular Relations in Information Transmission and Expression	Details	Free	
CA1Neuron - Pathways of the hippocampal CA1 neuron	Details	Free	
Cancer Cell Map - The Cancer Cell Map	Details	Free	BioPAX
CSP - Cytokine Signaling Pathway Database	Details	Free	
CTDB - Calmodulin Target Database	Details	Free	
DDIB - Database of Domain Interactions and Bindings	Details	Free	
DIP - Database of Interacting Proteins	Details	📖	PSI-MI
Doodle - Database of oligomerization domains from lambda experiments	Details	Free	
DopaNet - DopaNet	Details	Free	
DRC - Database of Ribosomal Crosslinks	Details	Free	
DSM - Dynamic Signaling Maps	Details	$	
FIMM - Functional Molecular Immunology	Details	Free	
FusionDB - Prokaryote Gene Fusion Events	Details	Free	

Vuk Pavlovic

Biological Pathway Exchange (BioPAX)



Before BioPAX
>100 DBs and tools
Tower of Babel



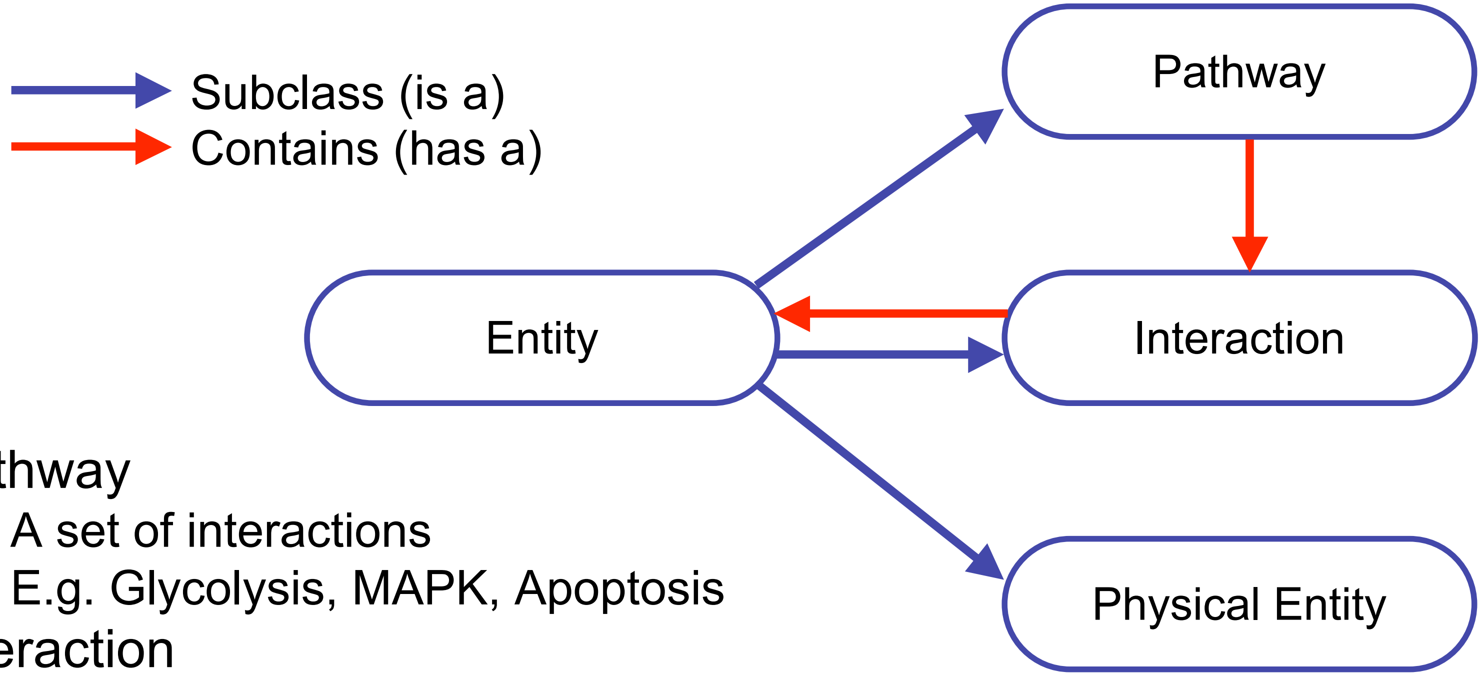
After BioPAX
Unifying language

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

BioPAX Structure

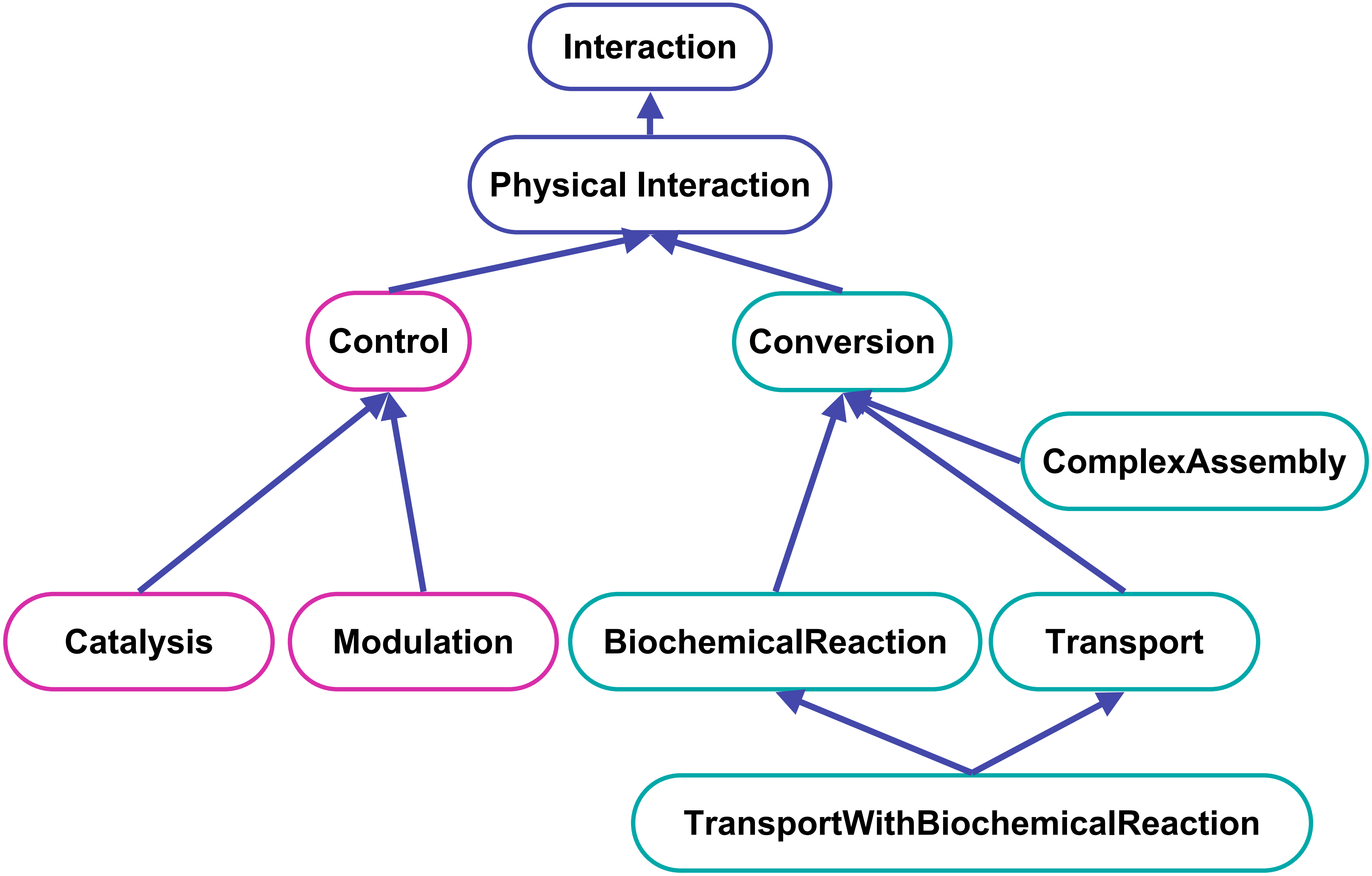


Subclass (is a)
Contains (has a)

Pathway
Entity
Interaction
Physical Entity

- Pathway
  - A set of interactions
  - E.g. Glycolysis, MAPK, Apoptosis
- Interaction
  - A basic relationship between a set of entities
  - E.g. Reaction, Molecular Association, Catalysis
- Physical Entity
  - A building block of simple interactions
  - E.g. Small molecule, Protein, DNA, RNA

BioPAX: Interactions



Interaction

Physical Interaction

Control

Conversion

ComplexAssembly

Catalysis

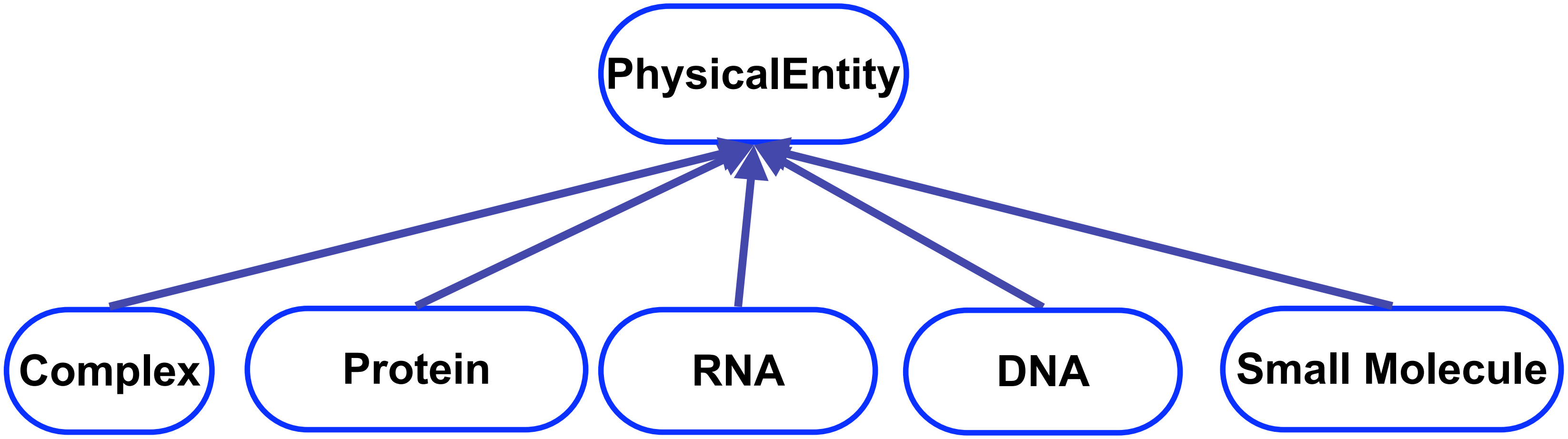
Modulation

BiochemicalReaction

Transport

TransportWithBiochemicalReaction

BioPAX: Physical Entities



PhysicalEntity

Complex

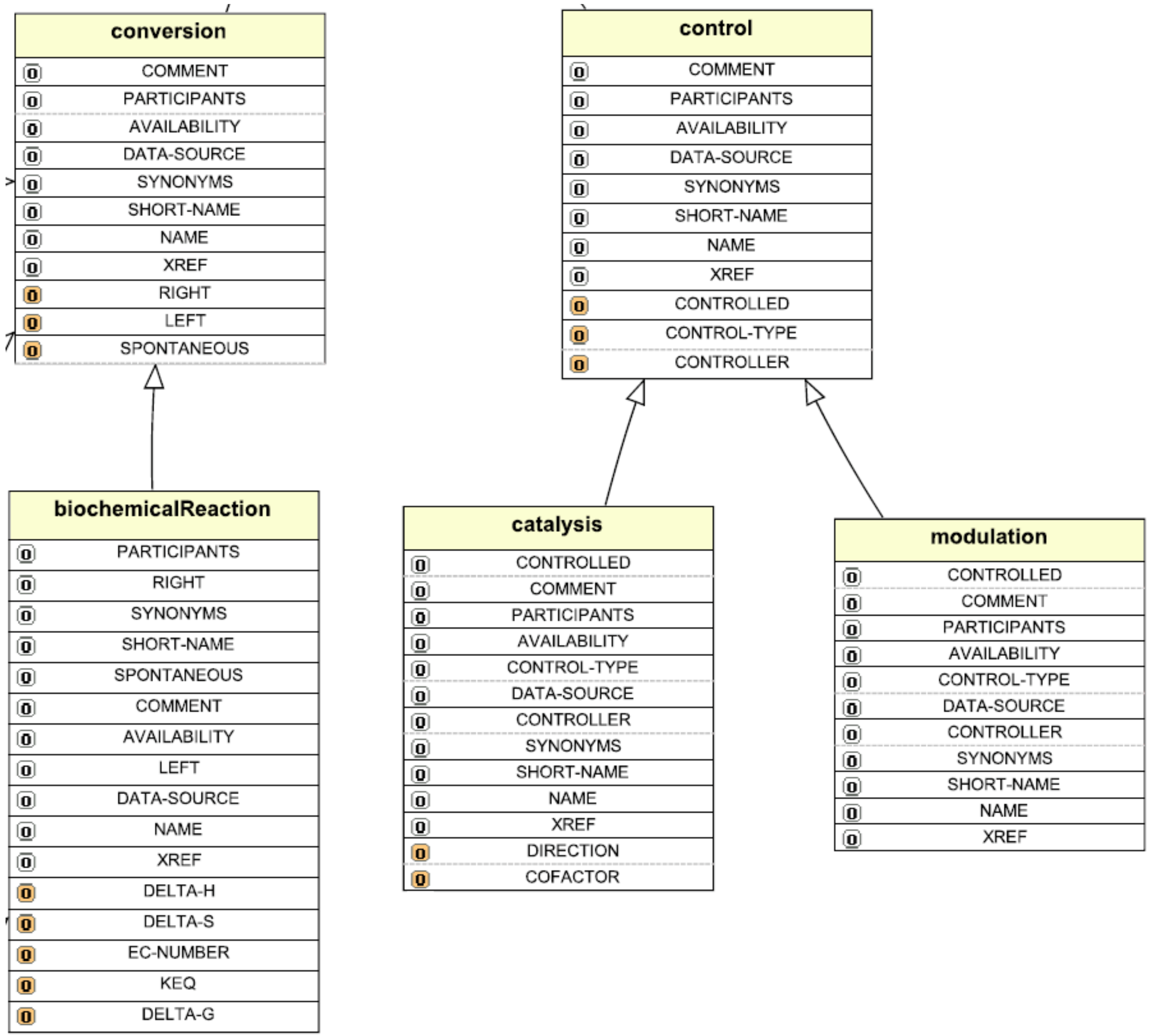
Protein

RNA

DNA

Small Molecule

BioPAX Ontology



conversion

- COMMENT
- PARTICIPANTS
- AVAILABILITY
- DATA-SOURCE
- SYNONYMS
- SHORT-NAME
- NAME
- XREF
- RIGHT
- LEFT
- SPONTANEOUS

control

- COMMENT
- PARTICIPANTS
- AVAILABILITY
- DATA-SOURCE
- SYNONYMS
- SHORT-NAME
- NAME
- XREF
- CONTROLLED
- CONTROL-TYPE
- CONTROLLER

biochemicalReaction

- PARTICIPANTS
- RIGHT
- SYNONYMS
- SHORT-NAME
- SPONTANEOUS
- COMMENT
- AVAILABILITY
- LEFT
- DATA-SOURCE
- NAME
- XREF
- DELTA-H
- DELTA-S
- EC-NUMBER
- KEQ
- DELTA-G

catalysis

- CONTROLLED
- COMMENT
- PARTICIPANTS
- AVAILABILITY
- CONTROL-TYPE
- DATA-SOURCE
- CONTROLLER
- SYNONYMS
- SHORT-NAME
- NAME
- XREF
- DIRECTION
- COFACTOR

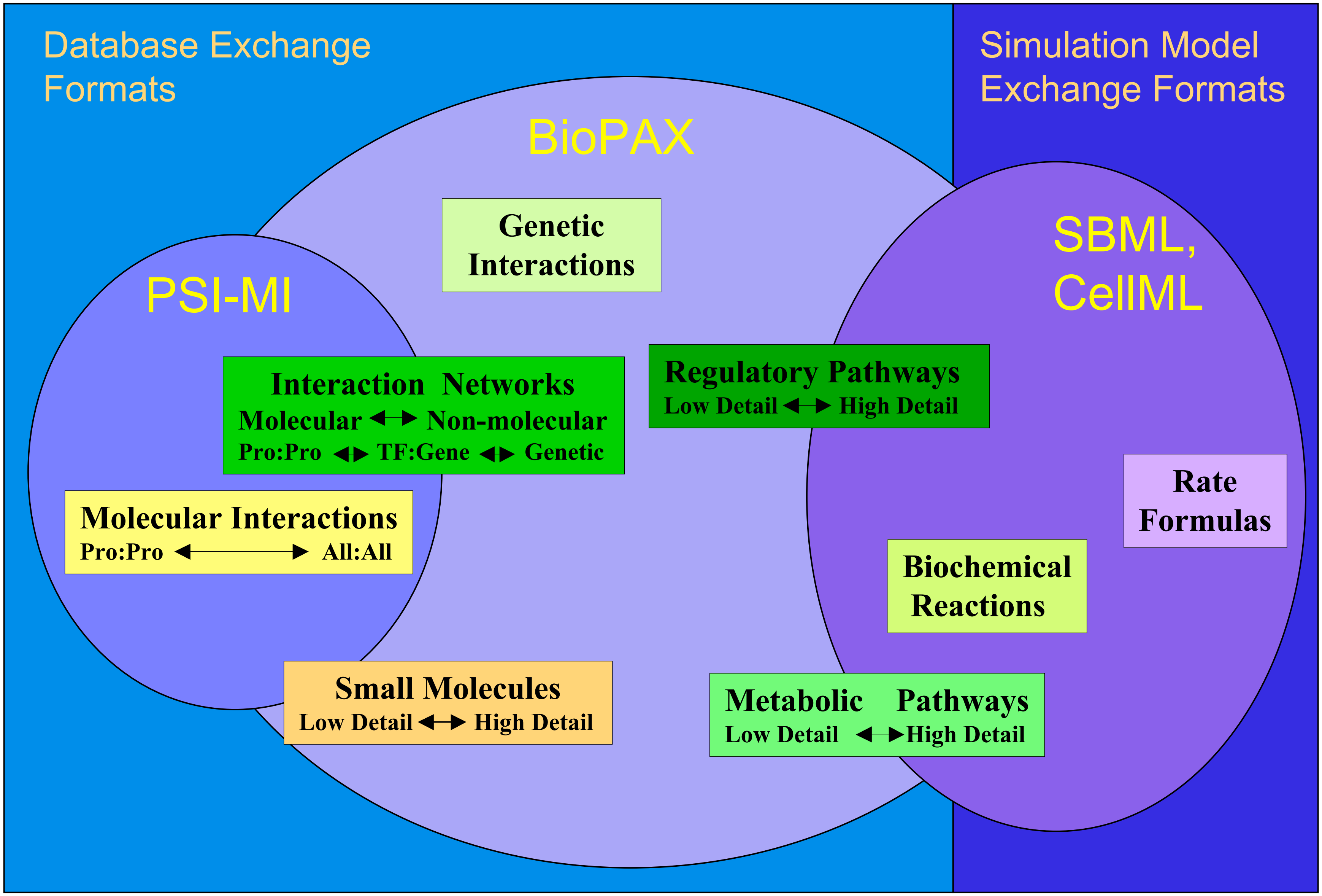
modulation

- CONTROLLED
- COMMENT
- PARTICIPANTS
- AVAILABILITY
- CONTROL-TYPE
- DATA-SOURCE
- CONTROLLER
- SYNONYMS
- SHORT-NAME
- NAME
- XREF

XML Snippet (OWL)

```
<bp:biochemicalReaction rdf:ID="biochemicalReaction37">
  <bp:DATA-SOURCE rdf:resource="#dataSource14"/>
  <bp:LEFT>
    <bp:physicalEntityParticipant rdf:ID="physicalEntityParticipant26">
      <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>
      <bp:PHYSICAL-ENTITY>
        <bp:smallMolecule rdf:ID="smallMolecule27">
          <bp:SHORT-NAME rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
          >a-D-glu-6-p</bp:SHORT-NAME>
          <bp:CHEMICAL-FORMULA rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
          >C6H13O9P</bp:CHEMICAL-FORMULA>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
          >&lt;FONT FACE="Symbol">a&lt;/FONT>-D-glucose-6-phoshate</bp:SYNONYMS>
          <bp:XREF>
            <bp:unificationXref rdf:ID="unificationXref30">
              <bp:ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
              >C00668</bp:ID>
              <bp:DB rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
              >KEGG</bp:DB>
            </bp:unificationXref>
          </bp:XREF>
          <bp:XREF rdf:resource="#unificationXref29"/>
          <bp:MOLECULAR-WEIGHT>260.14</bp:MOLECULAR-WEIGHT>
          <bp:AVAILABILITY rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
          >see http://www.amaze.ulb.ac.be/</bp:AVAILABILITY>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
          >glucose-6-P</bp:SYNONYMS>
          <bp:DATA-SOURCE rdf:resource="#dataSource14"/>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
          >alpha-D-glucose-6-p</bp:SYNONYMS>
          <bp:STRUCTURE>
            <bp:chemicalStructure rdf:ID="chemicalStructure28">
              <bp:STRUCTURE-FORMAT>SMILES</bp:STRUCTURE-FORMAT>
              <bp:STRUCTURE-DATA>C(OP(=O)(O)O)[CH]1([CH](O)[CH](O)[CH](O)[CH](O)O1)</bp:STRUCTURE-DATA>
            </bp:chemicalStructure>
          </bp:STRUCTURE>
          <bp:NAME>alpha-D-glucose 6-phosphate</bp:NAME>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
          >alpha-D-glucose-6-phosphate</bp:SYNONYMS>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
          > D-glucose-6-P</bp:SYNONYMS>
          <bp:DATA-SOURCE rdf:resource="#KB_439584_Individual_47"/>
        </bp:smallMolecule>
      </bp:PHYSICAL-ENTITY>
      <bp:CELLULAR-LOCATION rdf:resource="#openControlledVocabulary15"/>
    </bp:physicalEntityParticipant>
  </bp:LEFT>
  <bp:DELTA-G rdf:datatype="http://www.w3.org/2001/XMLSchema#double"
  >0.4</bp:DELTA-G>
  <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
  >alpha-D-Glucose 6-phosphate &lt;=> beta-D-Fructose 6-phosphate </bp:SYNONYMS>
  <bp:RIGHT>
    <bp:physicalEntityParticipant rdf:ID="physicalEntityParticipant38">
      <bp:CELLULAR-LOCATION rdf:resource="#openControlledVocabulary15"/>
      <bp:PHYSICAL-ENTITY>
        <bp:smallMolecule rdf:ID="smallMolecule39">
```

Exchange Formats in the Pathway Data Space



Database Exchange Formats

Simulation Model Exchange Formats

BioPAX

PSI-MI

SBML, CellML

Genetic Interactions

Interaction Networks
Molecular ↔ Non-molecular
Pro:Pro ↔ TF:Gene ↔ Genetic

Regulatory Pathways
Low Detail ↔ High Detail

Molecular Interactions
Pro:Pro ↔ All:All

Rate Formulas

Biochemical Reactions

Small Molecules
Low Detail ↔ High Detail

Metabolic Pathways
Low Detail ↔ High Detail

How to participate?

- Visit biopax.org and join the discussion mailing list
  - biopax-discuss@biopax.org
- Make pathway data available in BioPAX
- Build software that supports BioPAX
- Contribute BioPAX worked examples, documentation and specification reviews
- Spread the word about BioPAX
- Review documentation and specifications

BioPAX Supporting Groups

Current Participants

- Memorial Sloan-Kettering Cancer Center: E.Demir, M. Cary, C. Sander
- University of Toronto: G. Bader
- SRI Bioinformatics Research Group: P. Karp, S. Paley, J. Pick
- Bilkent University: U. Dogrusoz
- Université Libre de Bruxelles: C. Lemer
- CBRC Japan: K. Fukuda
- Dana Farber Cancer Institute: J. Zucker
- Millennium: J. Rees, A. Ruttenberg
- Cold Spring Harbor/EBI: G. Wu, M. Gillespie, P. D'Eustachio, I. Vastrik, L. Stein
- BioPathways Consortium: J. Luciano, E. Neumann, A. Regev, V. Schachter
- Argonne National Laboratory: N. Maltsev, E. Marland, M.Syed
- Harvard: F. Gibbons
- AstraZeneca: E. Pichler
- BIOBASE: E. Wingender, F. Schacherer
- NCI: M. Aladjem, C. Schaefer
- Università di Milano Bicocca, Pasteur, Rennes: A. Splendiani
- Vassar College: K. Dahlquist
- Columbia: A. Rzhetsky

Collaborating Organizations

- Proteomics Standards Initiative (PSI)
- Systems Biology Markup Language (SBML)
- CellML
- Chemical Markup Language (CML)

Databases

- BioCyc, WIT, KEGG, BIND, PharmGKB, aMAZE, INOH, Transpath, Reactome, PATIKA, eMIM, NCI PID, CellMap

Wouldn’t be possible without

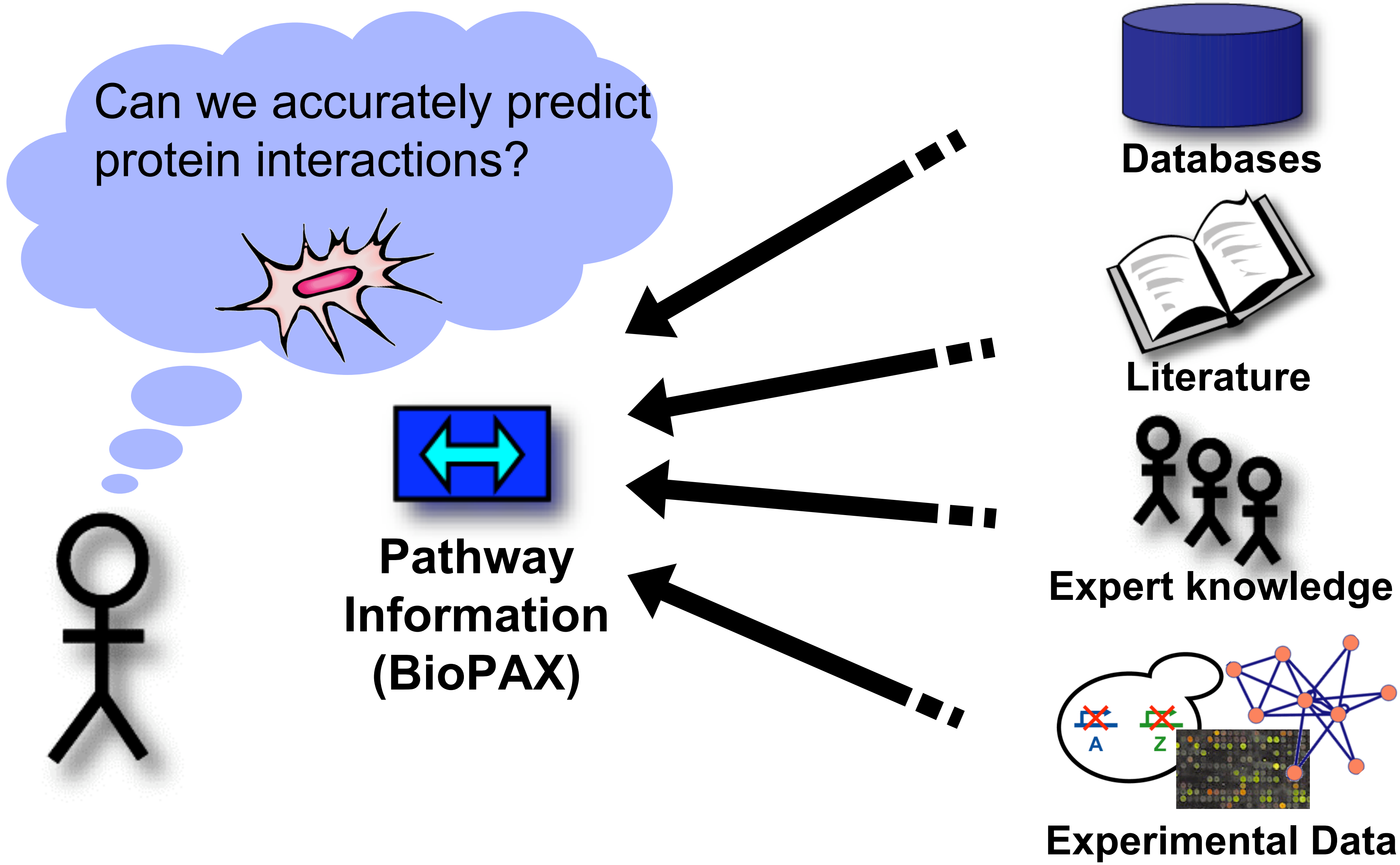
Gene Ontology

Protégé, U.Manchester, Stanford

Grants/Support

- Department of Energy (Workshop)
- caBIG

Using Pathway Information



Can we accurately predict protein interactions?

Pathway Information (BioPAX)

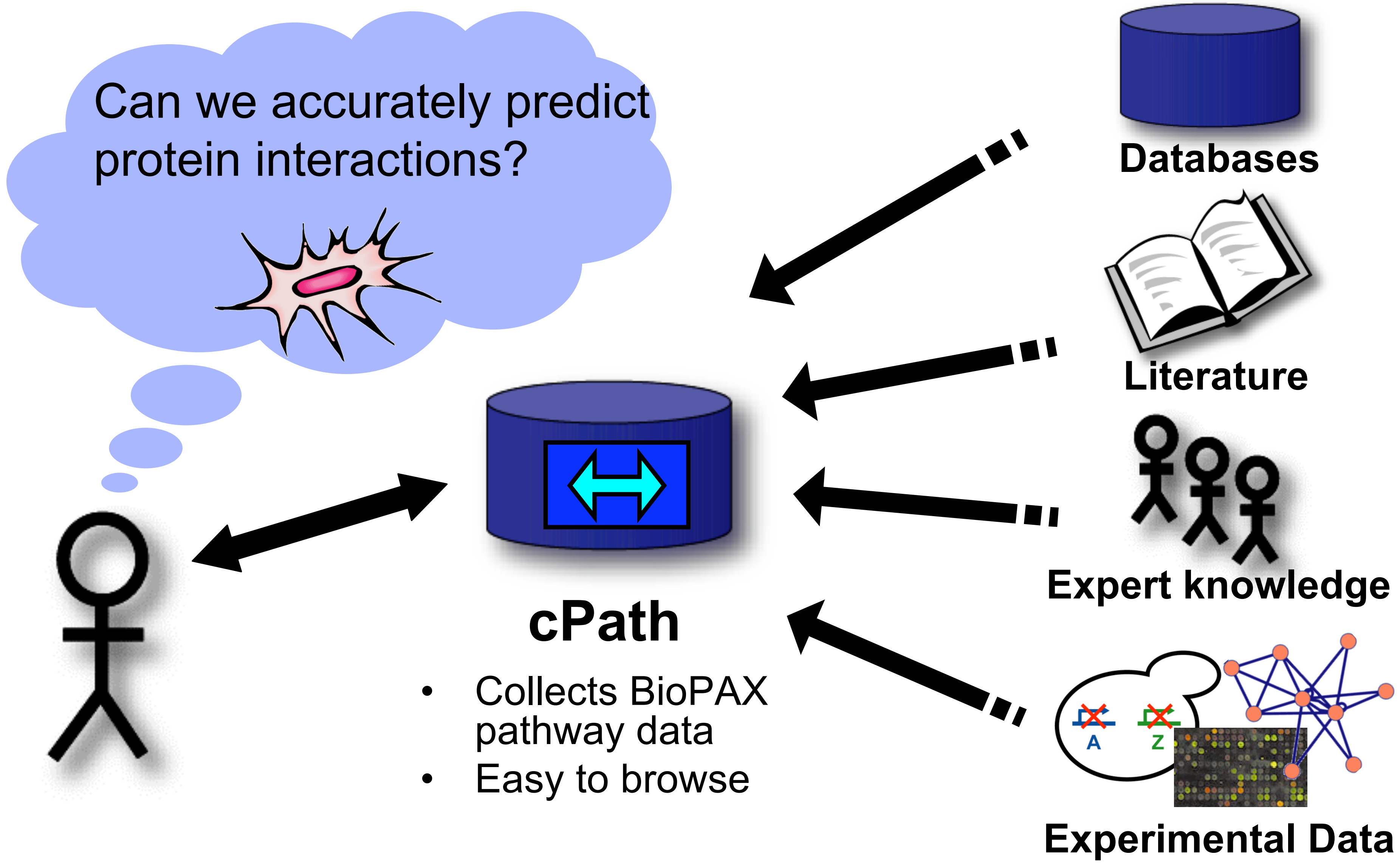
Databases

Literature

Expert knowledge

Experimental Data

Using Pathway Information



Can we accurately predict protein interactions?

Databases

Literature

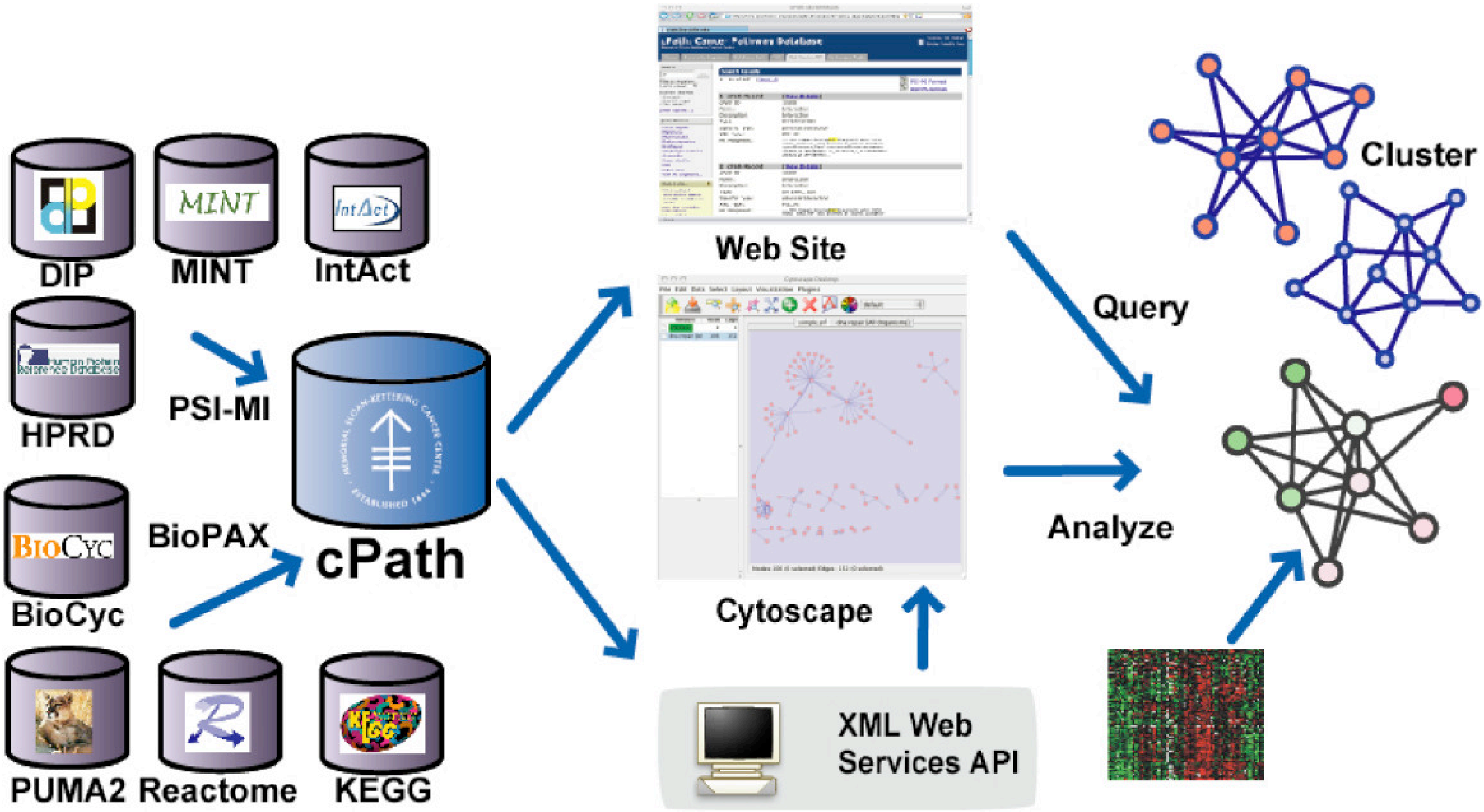
Expert knowledge

Experimental Data

cPath

- Collects BioPAX pathway data
- Easy to browse

cPath Pathway Database Software



DIP MINT IntAct

HPRD

PSI-MI

BioPAX

BioCyc

PUMA2 Reactome KEGG

cPath

Web Site

Cytoscape

XML Web Services API

Query

Analyze

Cluster

Collect → Browse / Query → Analyze

cPath Key Features

- Identifier mapping system e.g. proteins
- Scalable pathway data aggregation
- Simple web interface for browse and query
- Standard web service API for application communication
- 100% open source
  - Java, Tomcat, MySQL, Lucene, Struts, YUI
- Local installation and customization

http://cbio.mskcc.org/cpath

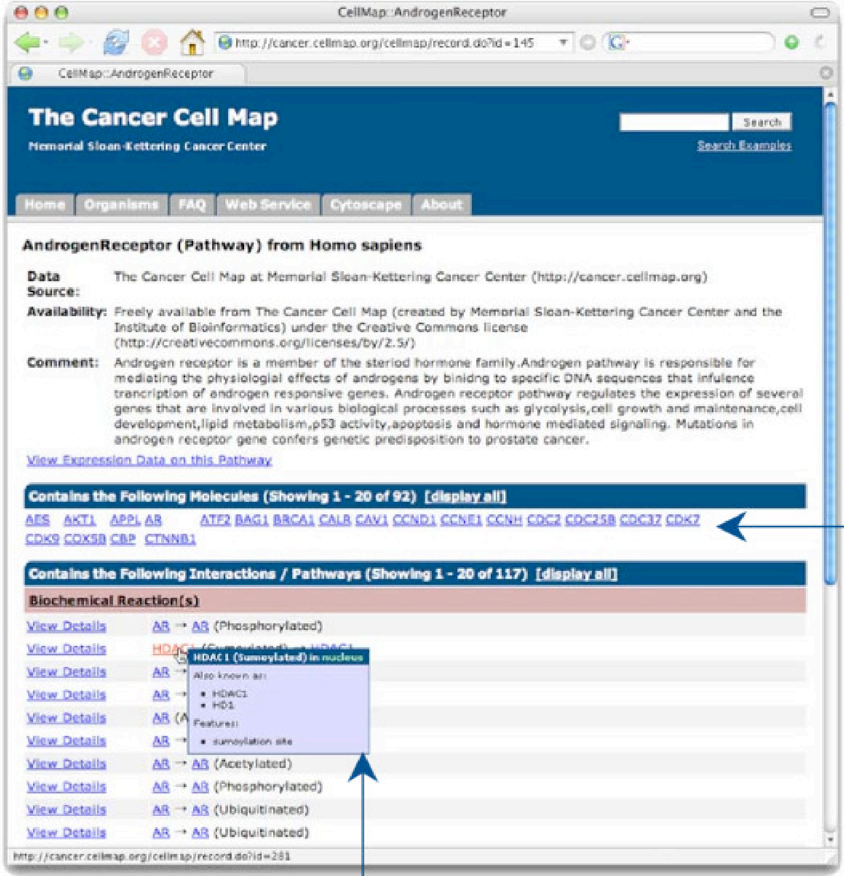
Cerami EG, Bader GD, Gross BE, Sander C.
BMC Bioinformatics. 2006 Nov 13;7:497

Customizable Content

Customizable headers, footers, and home page content.

Search Box

Enables users to search for pathways or proteins of interest.



The Cancer Cell Map

Memorial Sloan-Kettering Cancer Center

AndrogenReceptor (Pathway) from Homo sapiens

Data Source:	The Cancer Cell Map at Memorial Sloan-Kettering Cancer Center (http://cancer.cellmap.org)
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:	Androgen receptor is a member of the sterIod hormone family.Androgen pathway is responsible for mediating the physiologial effects of androgens by binidng to specific DNA sequences that infulence trancription of androgen responsive genes. Androgen receptor pathway regulates the expression of several genes that are involved in various biological processes such as glycolysis,cell growth and maintenance,cell development,lipid metabolism,p53 activity,apoptosis and hormone mediated signaling. Mutations in androgen receptor gene confers genetic predisposition to prostate cancer.

View Expression Data on this Pathway

Contains the Following Molecules (Showing 1 - 20 of 92) [display all]

AES AKT1 APPL AR ATF2 BAG1 BRCA1 CALR CAV1 CCND1 CCNE1 CCNH CDC2 CDC25B CDC37 CDK7 CDK9 COX5B CBP CTNNB1

Contains the Following Interactions / Pathways (Showing 1 - 20 of 117) [display all]

Biochemical Reaction(s)

View Details AR → AR (Phosphorylated)
View Details HDAC1 (Sumoylated) → HDAC1
View Details AR →
View Details AR →
View Details AR (A
View Details AR →
View Details AR → AR (Acetylated)
View Details AR → AR (Phosphorylated)
View Details AR → AR (Ubiquitinated)
View Details AR → AR (Ubiquitinated)

HDAC1 (Sumoylated) in nucleus

Also known as:
- HDAC1
- HD1

Features:
- sumoylation site

http://cancer.cellmap.org/cellmap/record.do?id=281

Basic Pathway Data

Includes pathway name, organism, source of data, availability restrictions, and comments.

Interaction List

Lists all interactions that participate In the pathway.

Sample Pop-Up Box

Rolling over members of an interaction opens a pop-up box with cellular location a and synonyms.

Molecule List

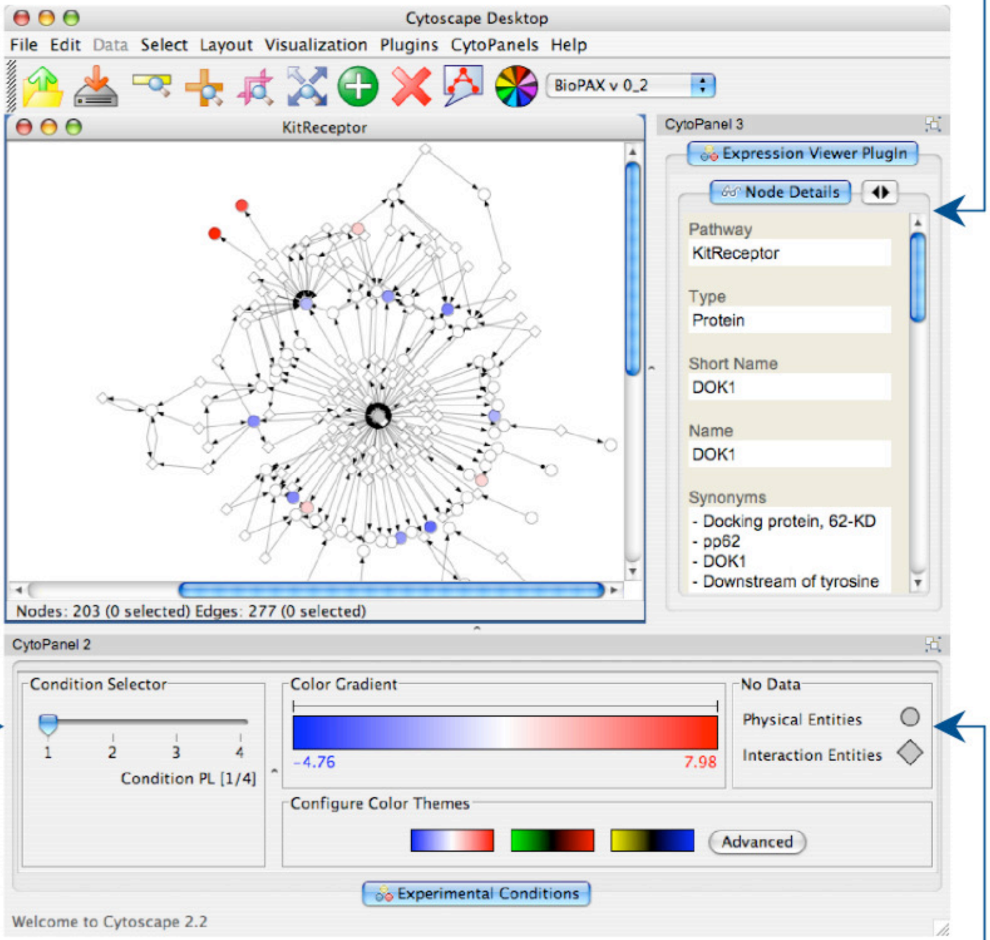
Lists all molecules that participate In the pathway.

Node / Edge Details

Window for viewing interactor or interaction details.

Network Visualization

Graphical rendering of the Kit Receptor pathway.



Pathway
KitReceptor

Type
Protein

Short Name
DOK1

Name
DOK1

Synonyms
- Docking protein, 62-KD
- pp62
- DOK1
- Downstream of tyrosine

Nodes: 203 (0 selected) Edges: 277 (0 selected)

Condition PL [1/4]

-4.76 7.98

Welcome to Cytoscape 2.2

Overlay Expression Data

Node color indicates expression level. Red indicates high expression level. Blue indicates low expression level.

Condition Selector

Slider for selecting current experimental condition.

Visual Legend

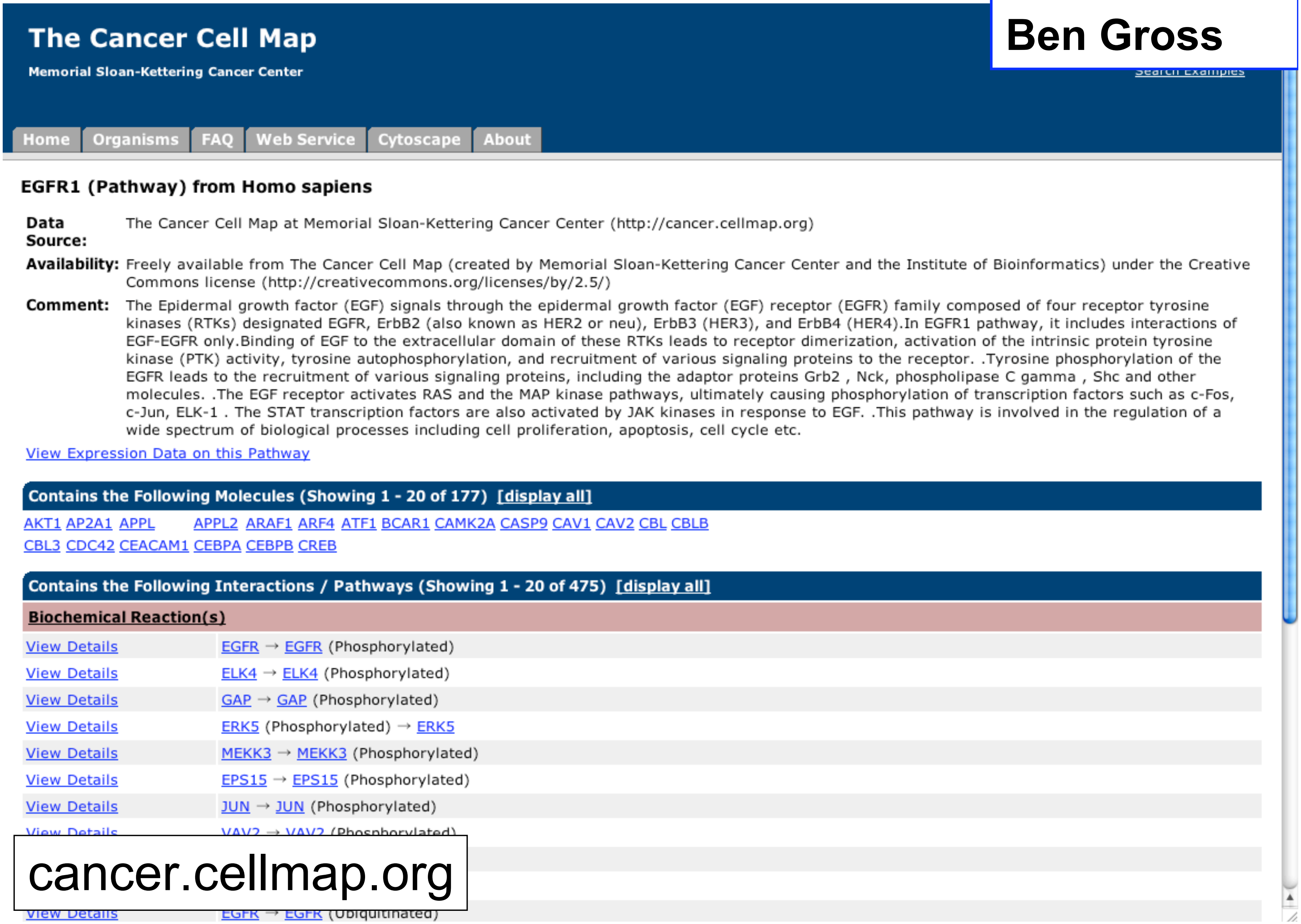
Visual legend for viewing and configuring the color gradient.

cPath web service API

- Queried by URL (RESTful architecture)
- getPathway, getNeighbors, getPathwayList, search
- webservice.do?cmd=get_pathway_list&version=2.0&q=O14763&input_id_type=UNIPROT

```
Database:ID       Pathway_Name                            Pathway_Database_Name   Internal_ID
UNIPROT:O14763    Apoptosis                               REACTOME                        579
UNIPROT:O14763    Extrinsic Pathway for Apoptosis         REACTOME                        580
UNIPROT:O14763    Death Receptor Signalling               REACTOME                        581
UNIPROT:O14763    FasL/ CD95L signaling                   REACTOME                        582
UNIPROT:O14763    TRAIL signaling                         REACTOME                        584
UNIPROT:O14763    Caspase-8 is formed from procaspase-8   REACTOME                        585
UNIPROT:O14763    Activation of Pro-Caspase 8             REACTOME                        586
```

Ethan Cerami
Ben Gross



The Cancer Cell Map

Memorial Sloan-Kettering Cancer Center

EGFR1 (Pathway) from Homo sapiens

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

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)

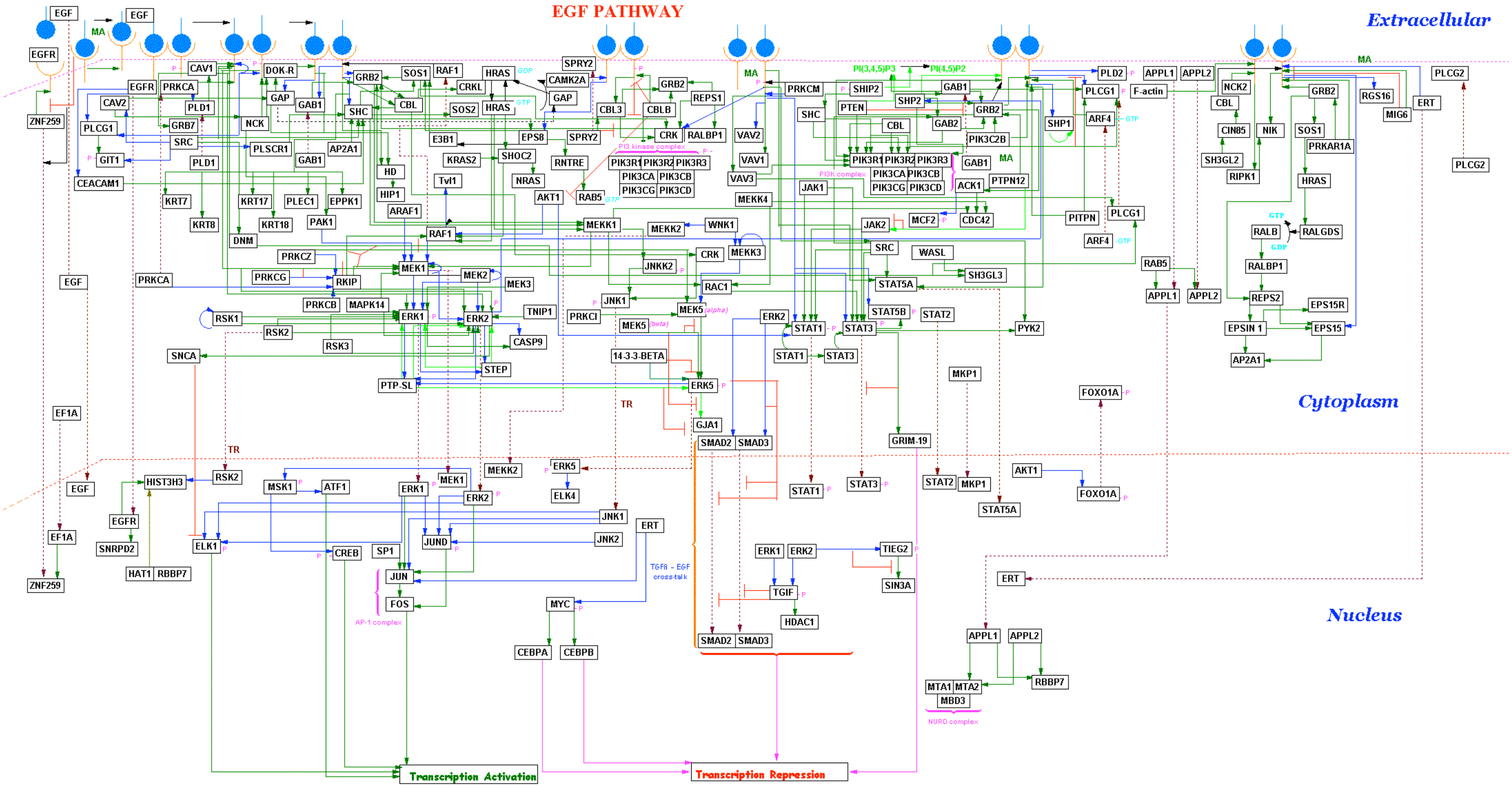
cancer.cellmap.org

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 + Wiley lab (PNNL)
- Institute of Bioinformatics, Bangalore
- Free under Creative Commons, BioPAX, easy to share

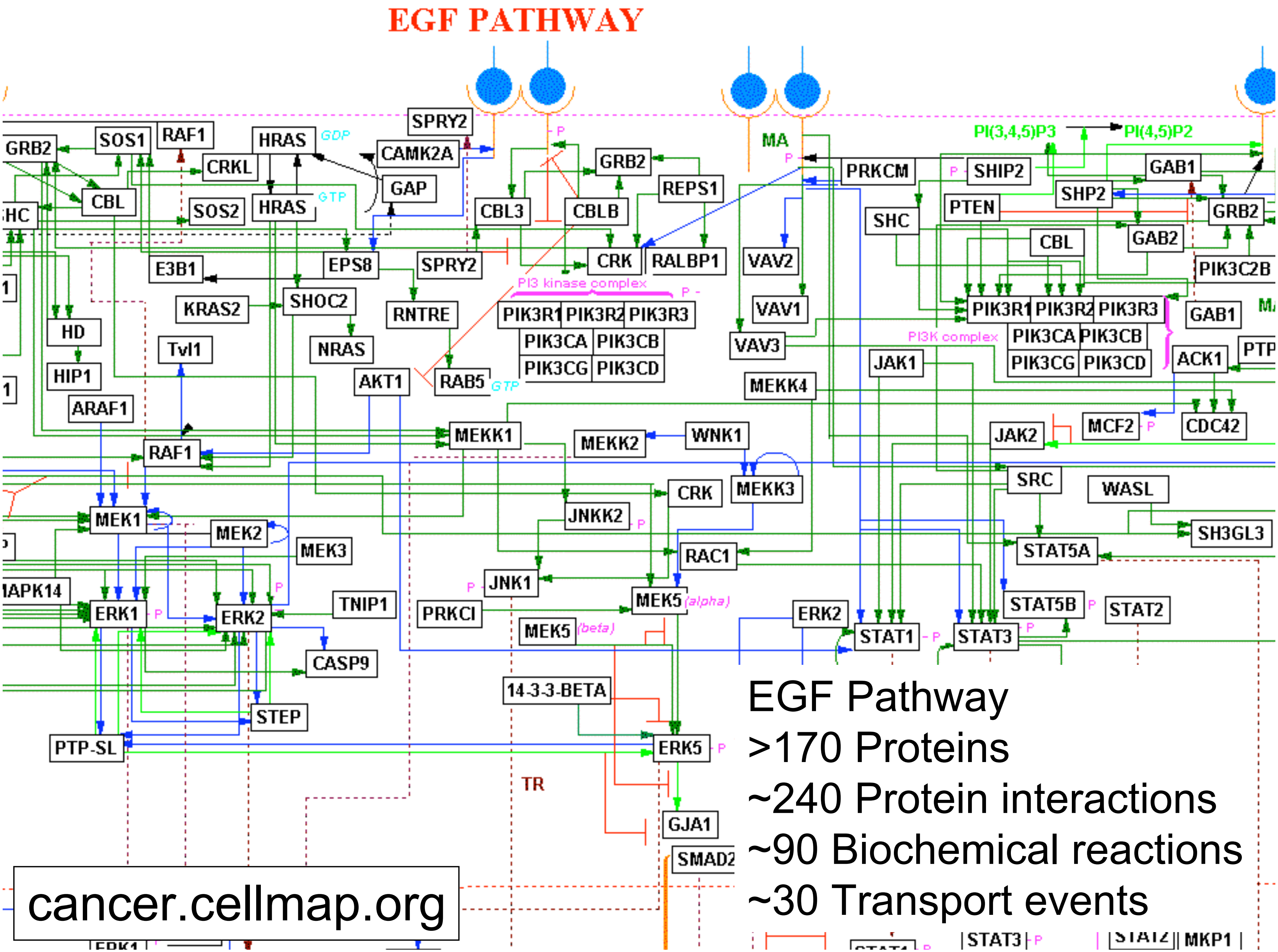


MA ⟶ Molecular Association
CONTROL
EC — Enzyme Catalysis
⟶ Phosphorylation
⟶ Dephosphorylation
⟶ Proteolytic Cleavage
⟶ Ubiquitination
⟶ Acetylation
TC — Transport Catalysis
TR --→ Transport
—| — Inhibition

Made with GenMAPP

cancer.cellmap.org

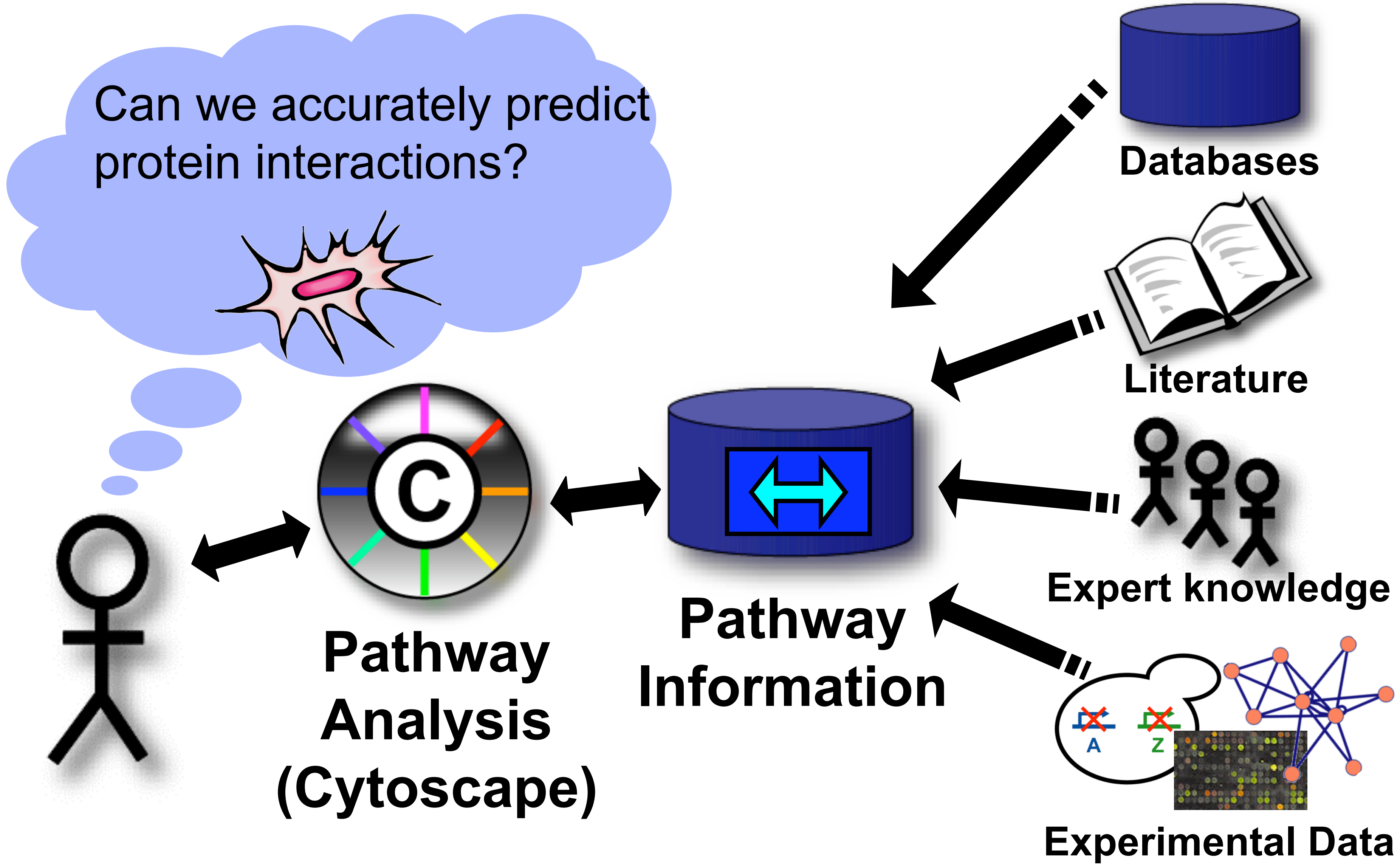
EGF Pathway
>170 Proteins
~240 Protein interactions
~90 Biochemical reactions
~30 Transport events



cancer.cellmap.org

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

Using Pathway Information



Can we accurately predict protein interactions?

Pathway Analysis (Cytoscape)

Pathway Information

Databases

Literature

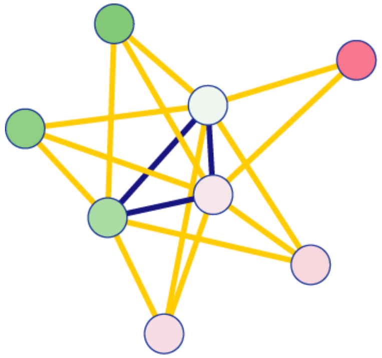
Expert knowledge

Experimental Data

Network visualization and analysis tool: Cytoscape

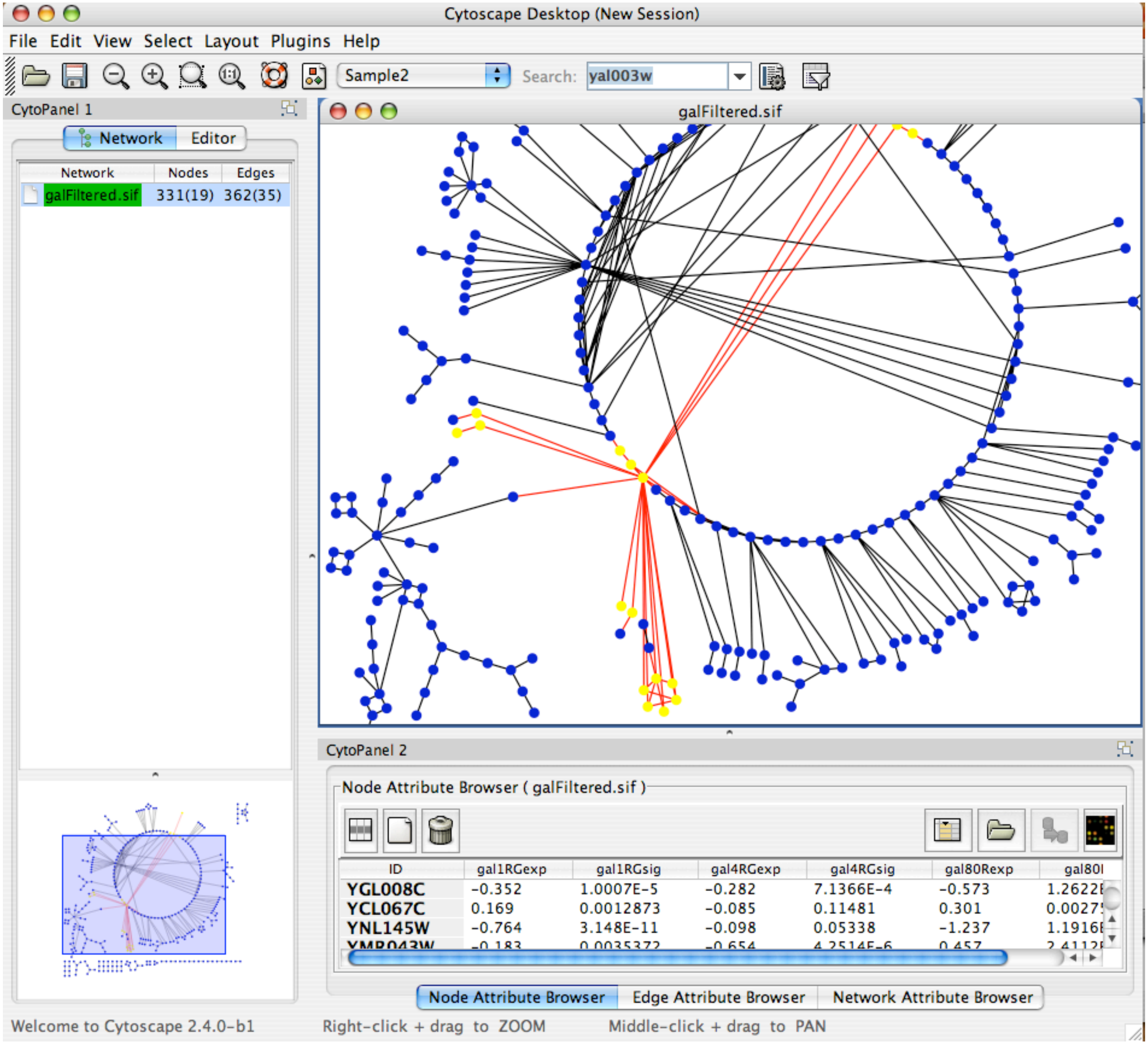
http://cytoscape.org

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



UCSD, ISB, Agilent, MSKCC, Pasteur, UCSF

Other software: Osprey, BioLayout, VisANT, Navigator, PIMWalker, ProViz



Network	Nodes	Edges
galFiltered.sif	331(19)	362(35)

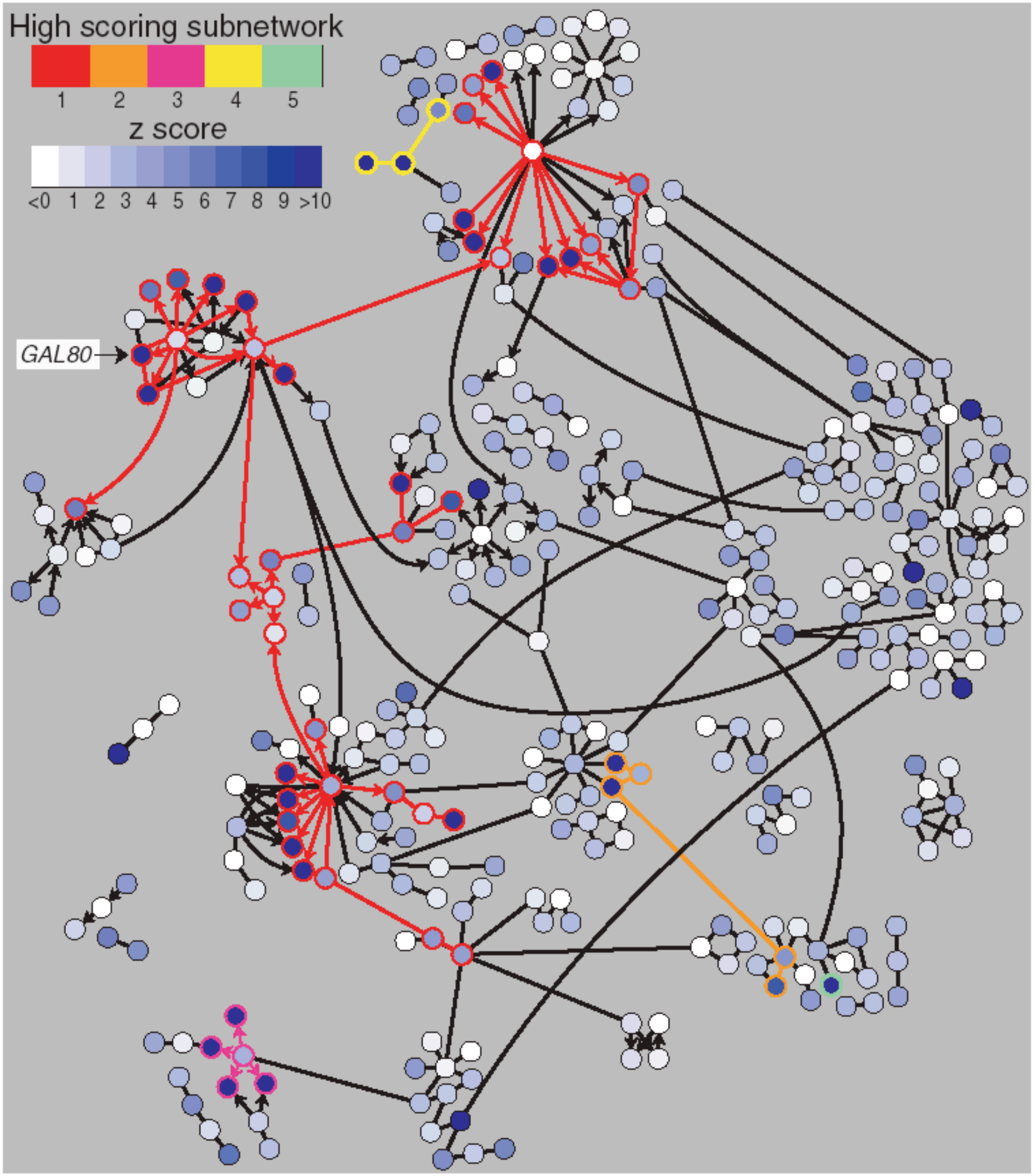
ID	gal1RGexp	gal1RGsig	gal4RGexp	gal4RGsig	gal80Rexp	gal80I
YGL008C	-0.352	1.0007E-5	-0.282	7.1366E-4	-0.573	1.2622
YCL067C	0.169	0.0012873	-0.085	0.11481	0.301	0.0027
YNL145W	-0.764	3.148E-11	-0.098	0.05338	-1.237	1.1916
YMR043W	-0.183	0.0035372	-0.654	4.2514E-6	0.457	2.4112

Welcome to Cytoscape 2.4.0-b1

Right-click + drag to ZOOM

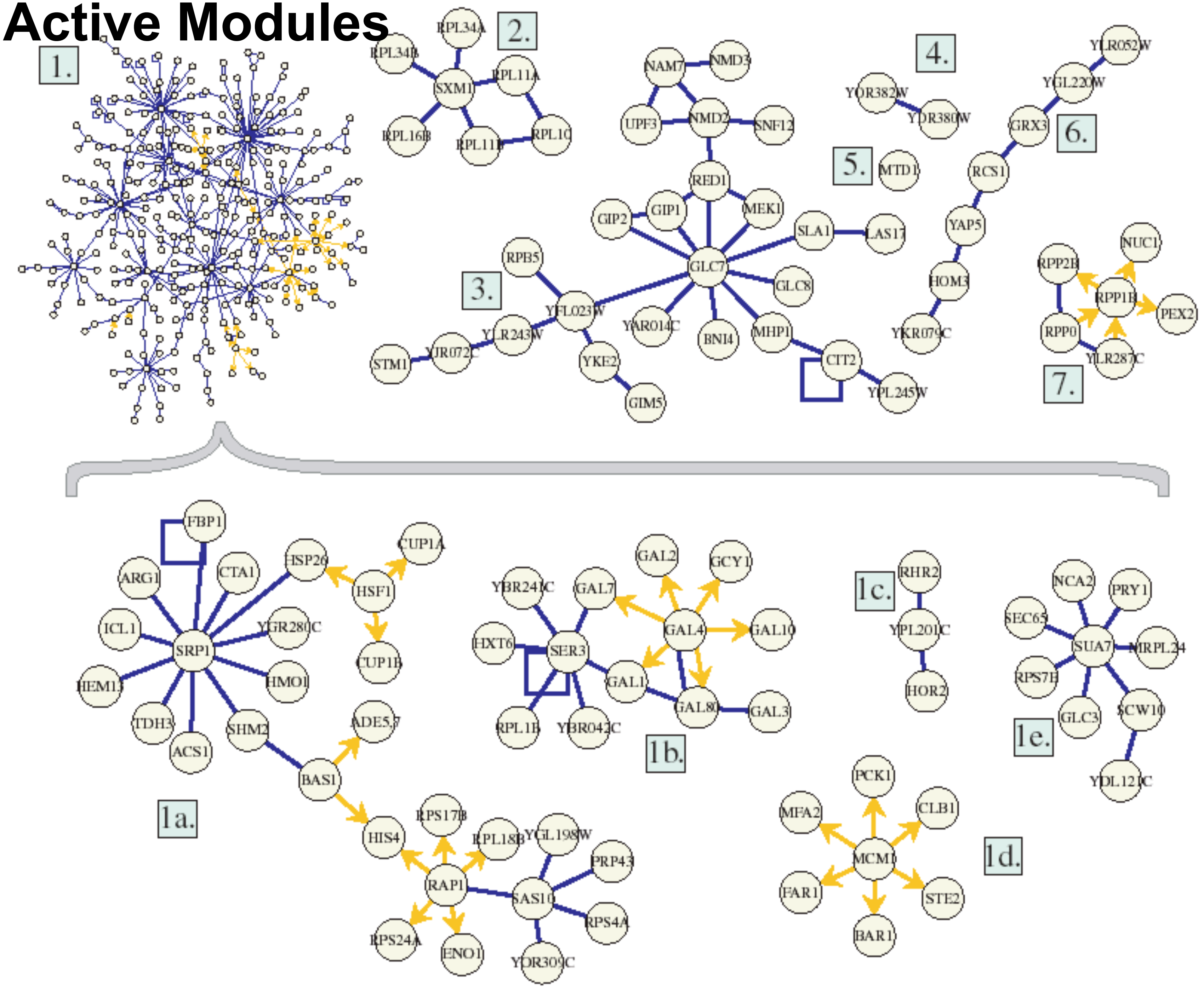
Middle-click + drag to PAN

Active Modules (UCSD)

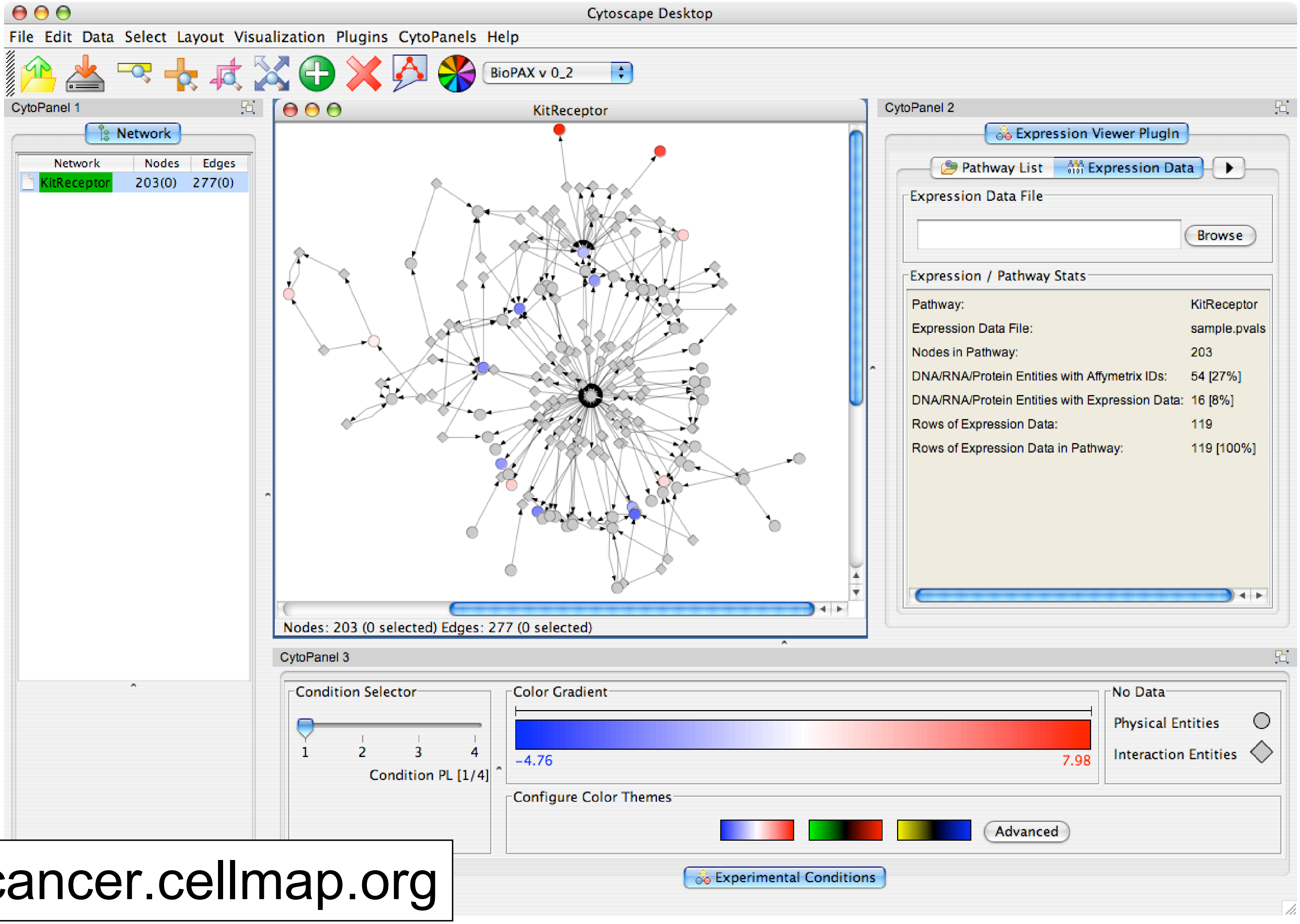


Ideker T, Ozier O, Schwikowski B, Siegel AF
Bioinformatics. 2002;18 Suppl 1:S233-40

Active Modules



The Cancer Cell Map



Network	Nodes	Edges
KitReceptor	203(0)	277(0)

Nodes: 203 (0 selected) Edges: 277 (0 selected)

Pathway:	KitReceptor
Expression Data File:	sample.pvals
Nodes in Pathway:	203
DNA/RNA/Protein Entities with Affymetrix IDs:	54 [27%]
DNA/RNA/Protein Entities with Expression Data:	16 [8%]
Rows of Expression Data:	119
Rows of Expression Data in Pathway:	119 [100%]

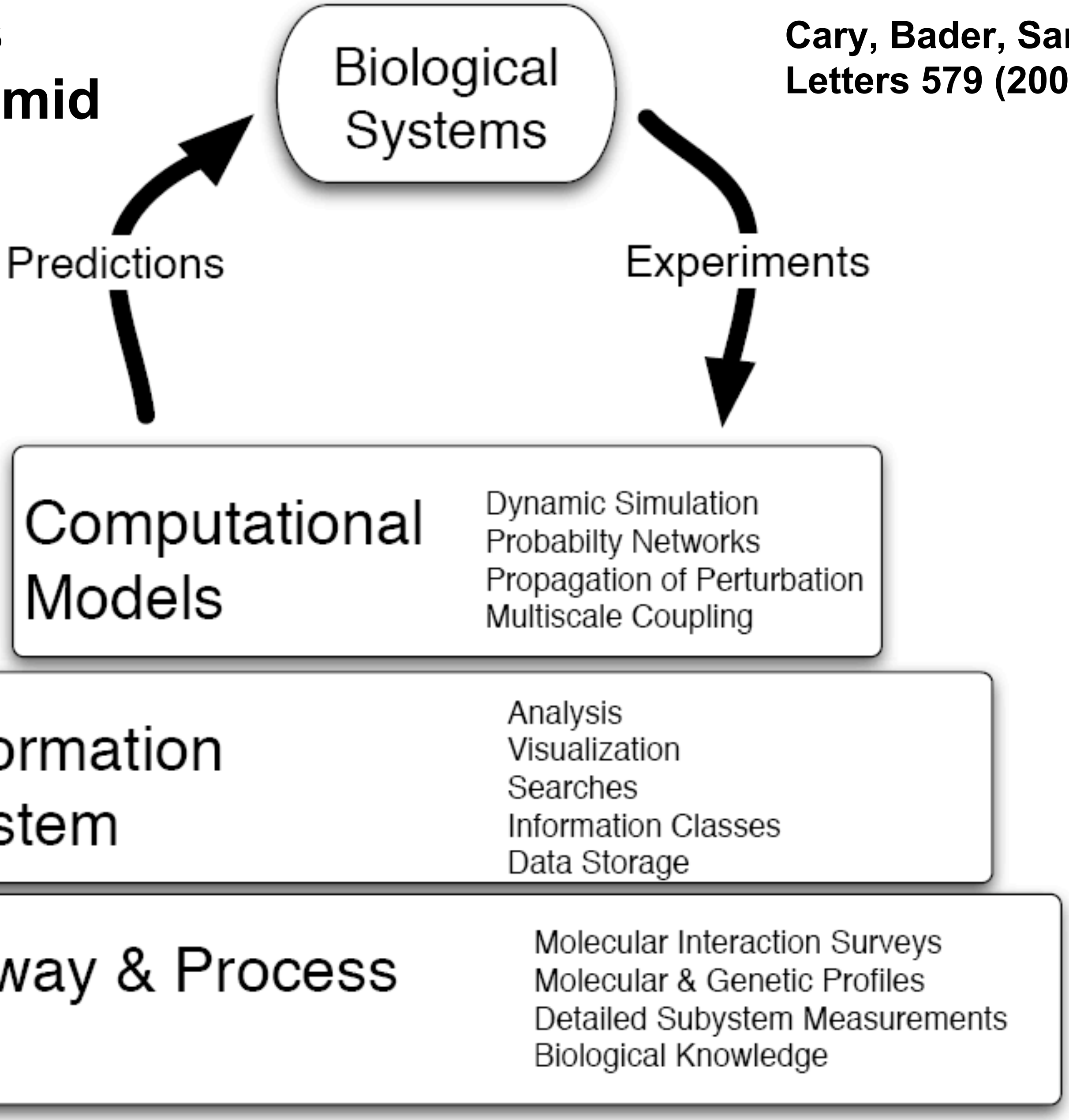
Condition PL [1/4]

-4.76 7.98

cancer.cellmap.org

The Systems Biology Pyramid

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



Biological Systems

Predictions

Experiments

Computational Models

- Dynamic Simulation
- Probabilty Networks
- Propagation of Perturbation
- Multiscale Coupling

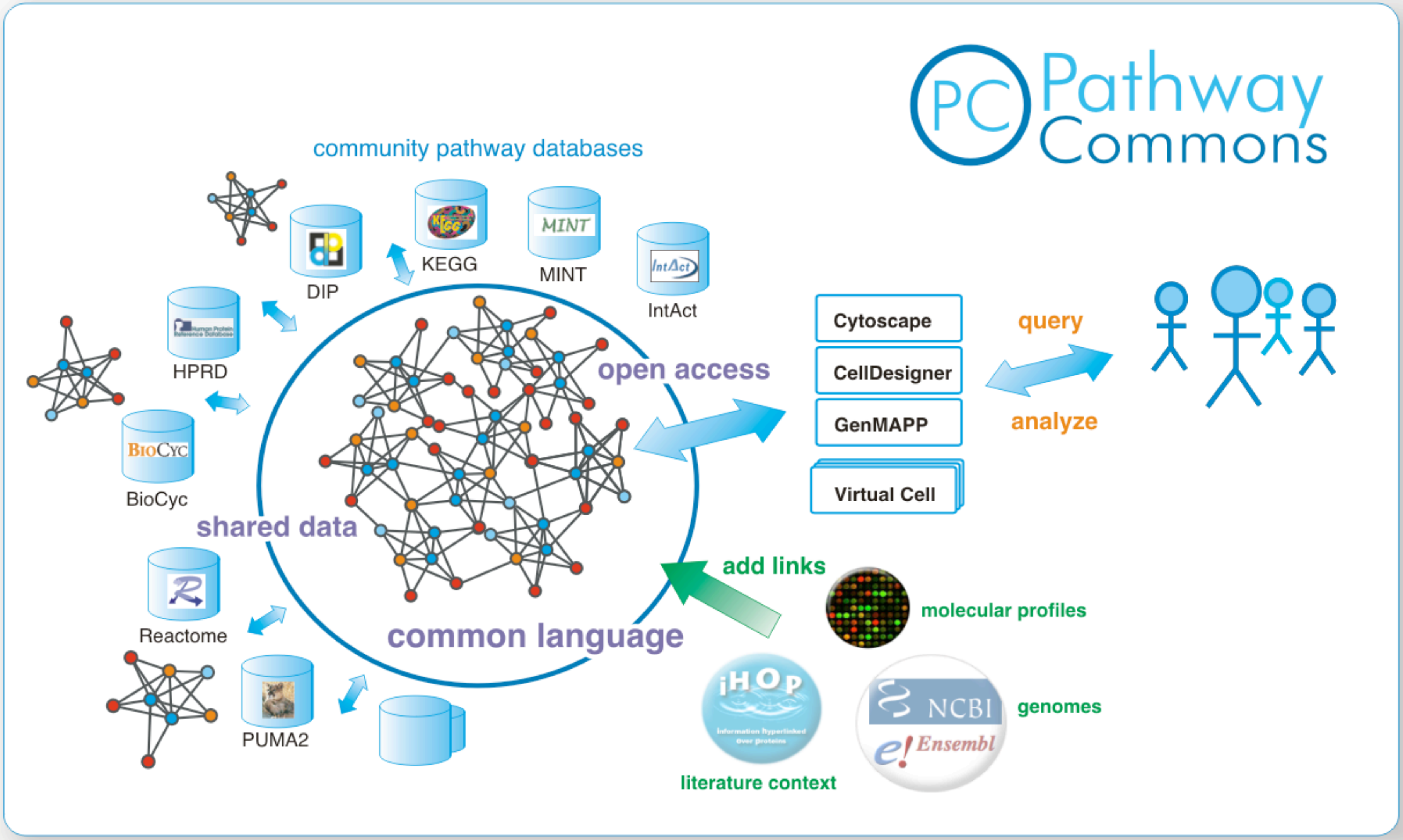
Information System

- Analysis
- Visualization
- Searches
- Information Classes
- Data Storage

Pathway & Process Data

- Molecular Interaction Surveys
- Molecular & Genetic Profiles
- Detailed Subystem Measurements
- Biological Knowledge

Pathway Commons: A Public Library

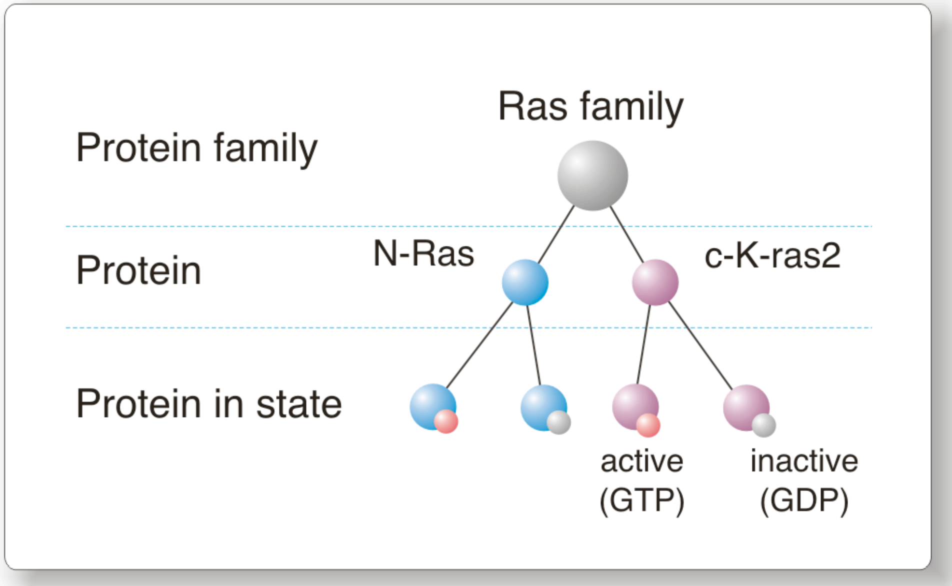


- Books: Pathways
- Lingua Franca: BioPAX OWL
- Index: cPath pathway database software
- Translators: translators to BioPAX

- Open access, free software
- No competition: Author attribution
- Aggregate ~ 20 databases in BioPAX format

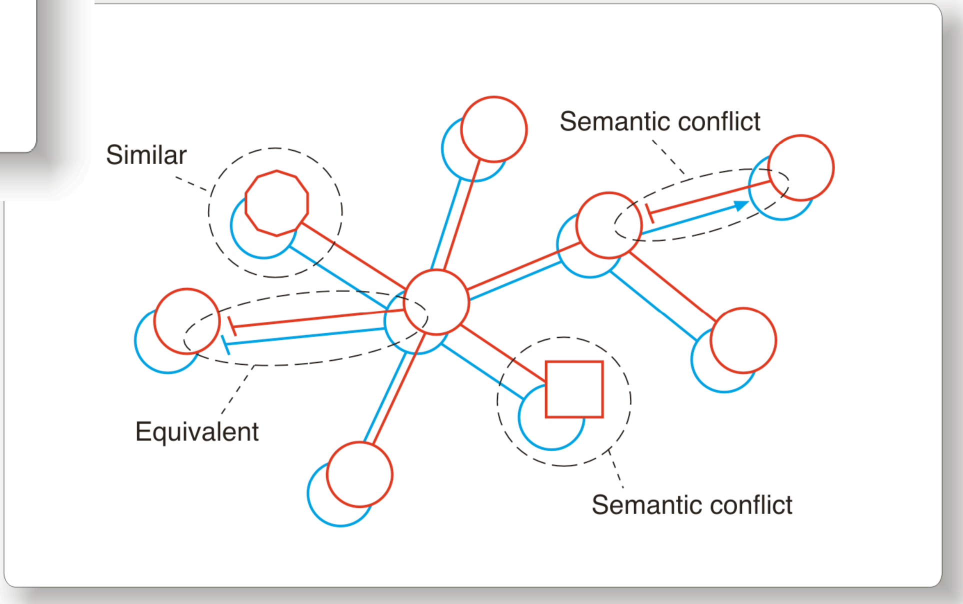
Towards an Integrated Cell Map

- Semantic pathway integration is very hard



Physical entities

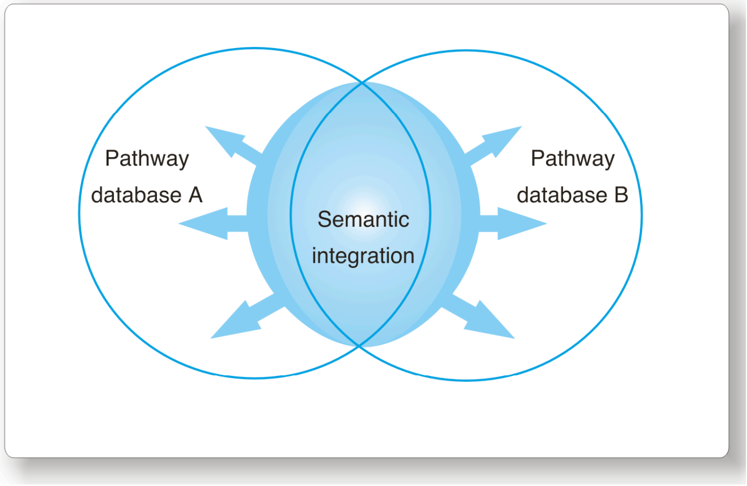
Relationships



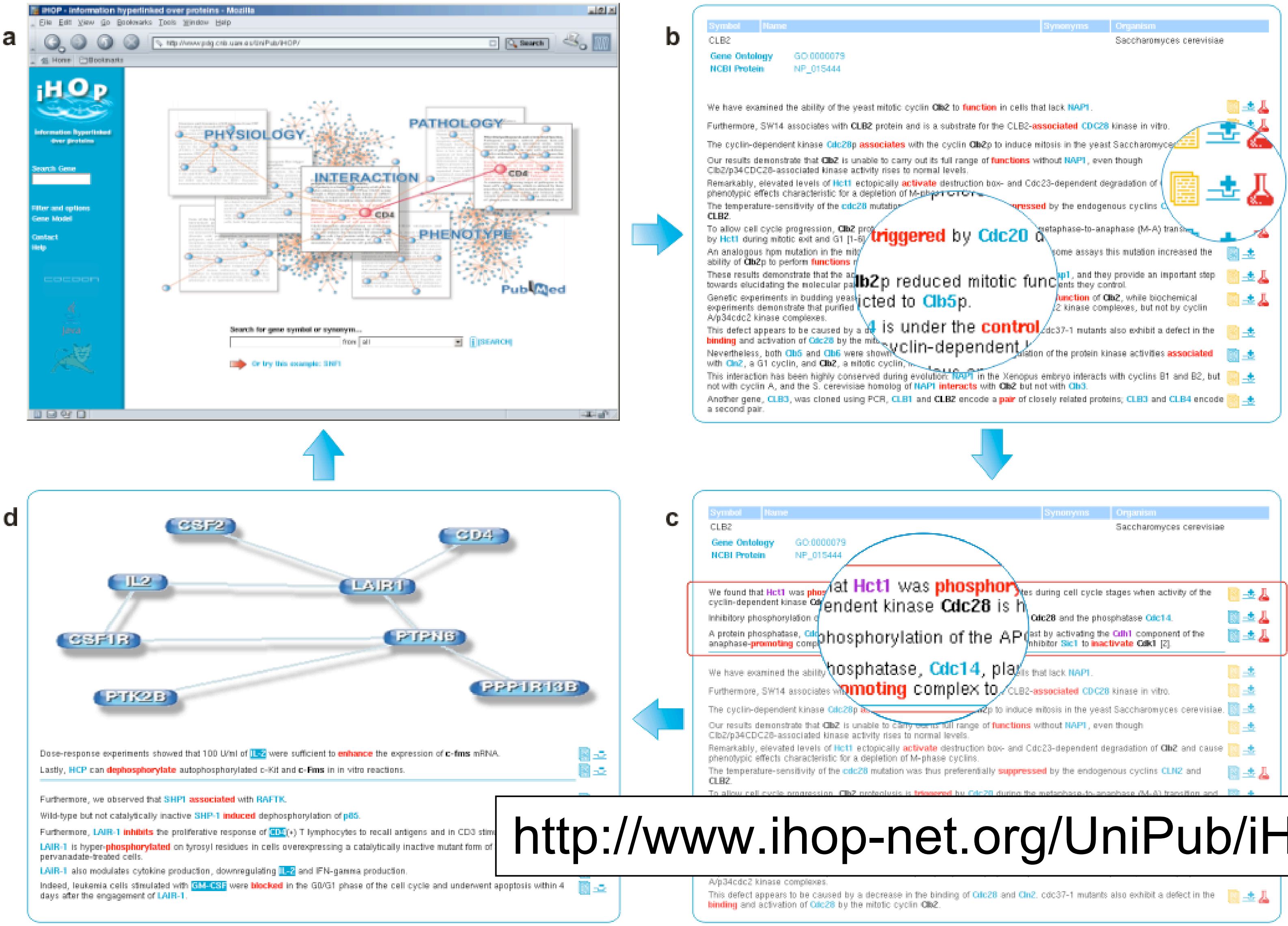
Practical Semantic Integration

- Minimize errors
  - Integrate only where possible with high accuracy
  - Detect and flag conflicts, errors for users, no revision
  - Promote best-practices to minimize future errors
  - Interaction confidence algorithms
  - Validation software
  - Allow users to filter and select trusted sources
- Converge to standard representation
  - Community process

Doable: hundreds of curators globally in >200 databases (GDP) - make it more efficient



Add Value via Text Mining



Symbol	Name	Synonyms	Organism
CLB2			Saccharomyces cerevisiae

Gene Ontology GO:0000079
NCBI Protein NP_015444

We have examined the ability of the yeast mitotic cyclin Clb2 to function in cells that lack NAP1.

Furthermore, SW14 associates with CLB2 protein and is a substrate for the CLB2-associated CDC28 kinase in vitro.

The cyclin-dependent kinase Cdc28p associates with the cyclin Clb2p to induce mitosis in the yeast Saccharomyces cerevisiae.

Our results demonstrate that Clb2 is unable to carry out its full range of functions without NAP1, even though Clb2/p34CDC28-associated kinase activity rises to normal levels.

Remarkably, elevated levels of Hct1 ectopically activate destruction box- and Cdc23-dependent degradation of Clb2 and cause phenotypic effects characteristic for a depletion of M-phase cyclins.

The temperature-sensitivity of the cdc28 mutation was thus preferentially suppressed by the endogenous cyclins CLN2 and CLB2.

To allow cell cycle progression, Clb2 proteolysis is triggered by Cdc20 during the metaphase-to-anaphase (M-A) transition and by Hct1 during mitotic exit and G1 [1-6].

An analogous hpm mutation in the mit… some assays this mutation increased the ability of Clb2p to perform functions r…

These results demonstrate that the ac…p1, and they provide an important step towards elucidating the molecular pa…ents they control.

Genetic experiments in budding yeas…function of Clb2, while biochemical experiments demonstrate that purified …2 kinase complexes, but not by cyclin A/p34cdc2 kinase complexes.

This defect appears to be caused by a decrease in the binding of Cdc28 and Cln2. cdc37-1 mutants also exhibit a defect in the binding and activation of Cdc28 by the mitotic cyclin Clb2.

Nevertheless, both Clb5 and Clb6 were shown… …lation of the protein kinase activities associated with Cln2, a G1 cyclin, and Clb2, a mitotic cyclin, …

This interaction has been highly conserved during evolution: NAP1 in the Xenopus embryo interacts with cyclins B1 and B2, but not with cyclin A, and the S. cerevisiae homolog of NAP1 interacts with Clb2 but not with Clb3.

Another gene, CLB3, was cloned using PCR, CLB1 and CLB2 encode a pair of closely related proteins; CLB3 and CLB4 encode a second pair.

triggered by Cdc20 d…
…lb2p reduced mitotic func…
…icted to Clb5p.
…4 is under the control…
…cyclin-dependent k…

We found that Hct1 was phos…tes during cell cycle stages when activity of the cyclin-dependent kinase Cdc28 is h…

Inhibitory phosphorylation o… Cdc28 and the phosphatase Cdc14.

A protein phosphatase, Cdc14, pla…ast by activating the Cdh1 component of the anaphase-promoting comp… …nhibitor Sic1 to inactivate Cdk1 [2].

…hat Hct1 was phosphory…
…endent kinase Cdc28 is h…
…phosphorylation of the AP…
…hosphatase, Cdc14, pla…
…omoting complex to…

Dose-response experiments showed that 100 U/ml of IL-2 were sufficient to enhance the expression of c-fms mRNA.

Lastly, HCP can dephosphorylate autophosphorylated c-Kit and c-Fms in in vitro reactions.

Furthermore, we observed that SHP1 associated with RAFTK.

Wild-type but not catalytically inactive SHP-1 induced dephosphorylation of p85.

Furthermore, LAIR-1 inhibits the proliferative response of CD4(+) T lymphocytes to recall antigens and in CD3 stim…

LAIR-1 is hyper-phosphorylated on tyrosyl residues in cells overexpressing a catalytically inactive mutant form of … pervanadate-treated cells.

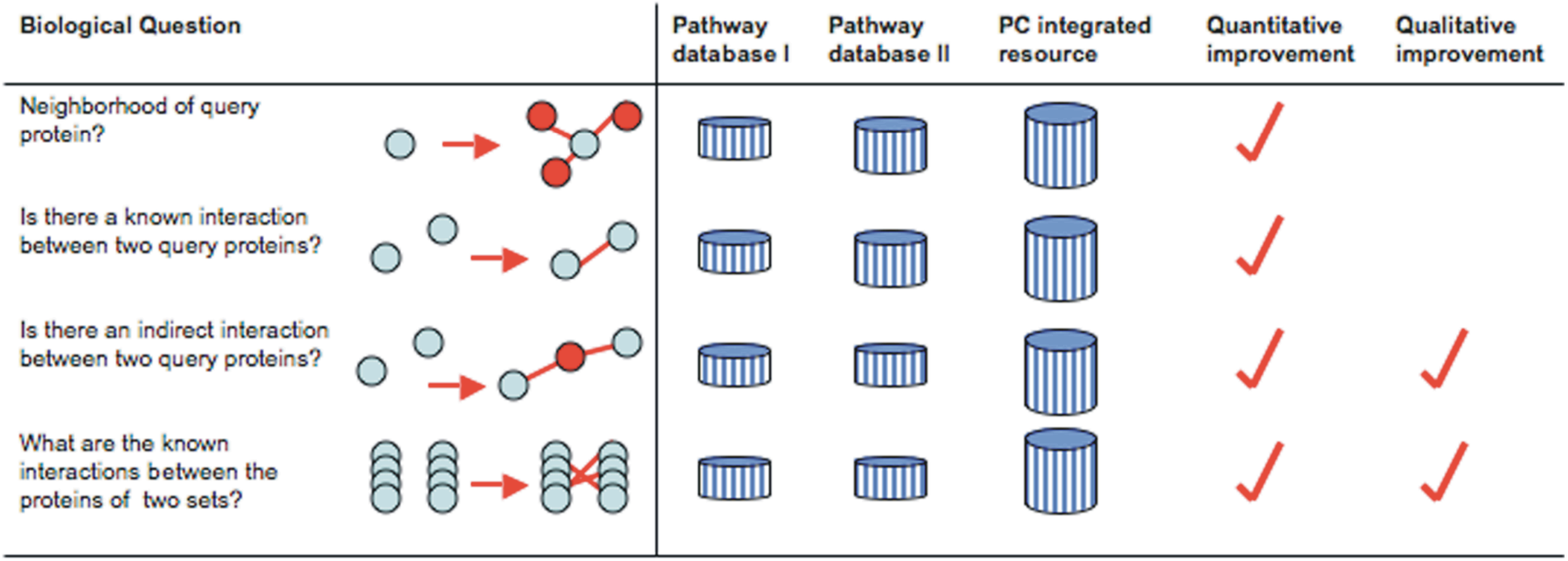
LAIR-1 also modulates cytokine production, downregulating IL-2 and IFN-gamma production.

Indeed, leukemia cells stimulated with GM-CSF were blocked in the G0/G1 phase of the cell cycle and underwent apoptosis within 4 days after the engagement of LAIR-1.

http://www.ihop-net.org/UniPub/iHOP/

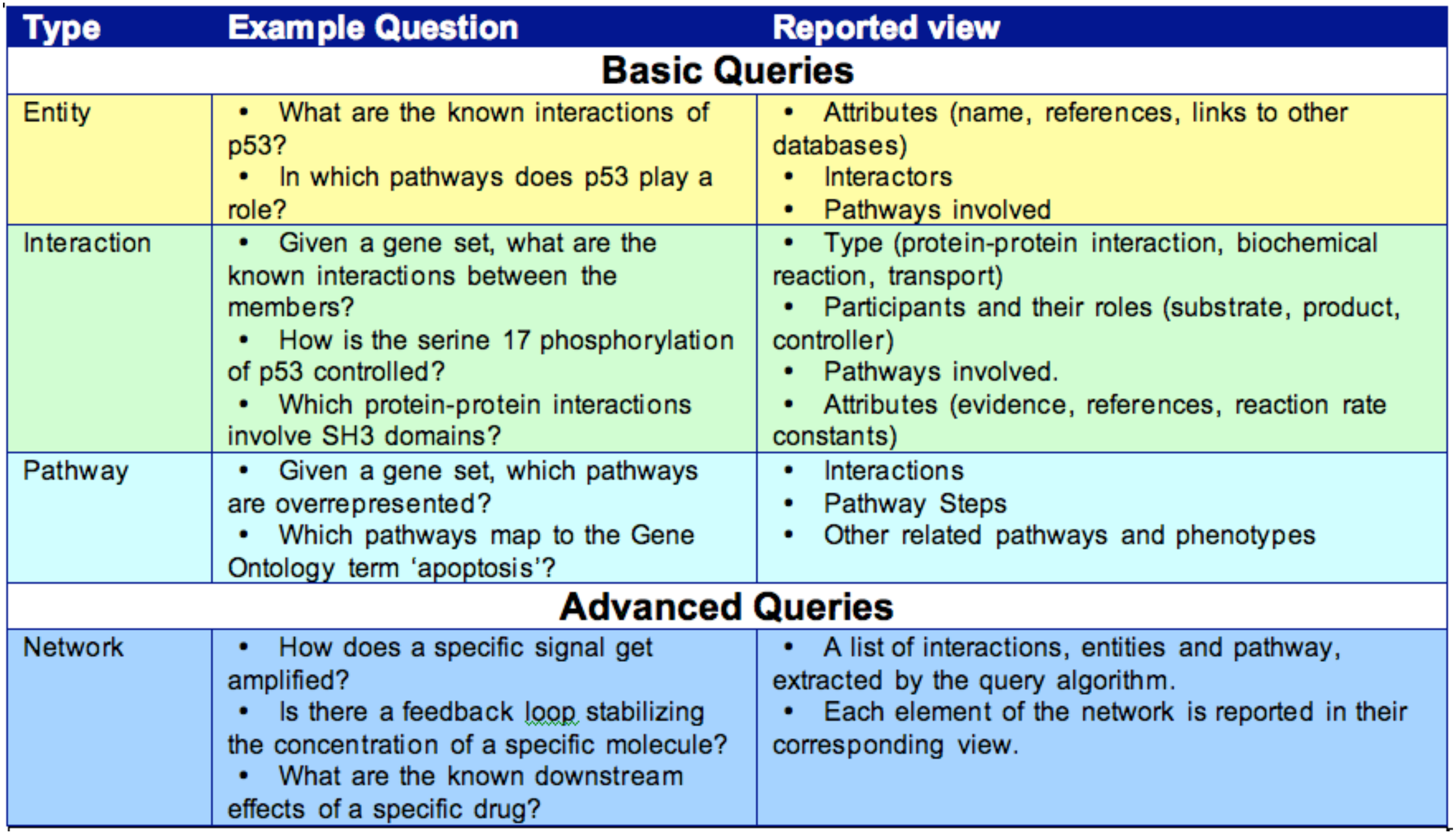
Robert Hoffmann, Alfonso Valencia, Chris Sander

Improved Queries



Biological Question	Pathway database I	Pathway database II	PC integrated resource	Quantitative improvement	Qualitative improvement
Neighborhood of query protein?				✓	
Is there a known interaction between two query proteins?				✓	
Is there an indirect interaction between two query proteins?				✓	✓
What are the known interactions between the proteins of two sets?				✓	✓

Pathway Commons Queries



Type	Example Question	Reported view
	Basic Queries	
Entity	• What are the known interactions of p53? • In which pathways does p53 play a role?	• Attributes (name, references, links to other databases) • Interactors • Pathways involved
Interaction	• Given a gene set, what are the known interactions between the members? • How is the serine 17 phosphorylation of p53 controlled? • Which protein-protein interactions involve SH3 domains?	• Type (protein-protein interaction, biochemical reaction, transport) • Participants and their roles (substrate, product, controller) • Pathways involved. • Attributes (evidence, references, reaction rate constants)
Pathway	• Given a gene set, which pathways are overrepresented? • Which pathways map to the Gene Ontology term 'apoptosis'?	• Interactions • Pathway Steps • Other related pathways and phenotypes
	Advanced Queries	
Network	• How does a specific signal get amplified? • Is there a feedback loop stabilizing the concentration of a specific molecule? • What are the known downstream effects of a specific drug?	• A list of interactions, entities and pathway, extracted by the query algorithm. • Each element of the network is reported in their corresponding view.

Platform for research for more advanced queries

http://pathwaycommons.org

Pathway Commons is a project for the scientific community. We welcome your feedback at: pc-info@pathwaycommons.org.

Pathway Commons is a convenient point of access to biological pathway information collected from public pathway databases, which you can browse or search. Pathways include biochemical reactions, complex assembly, transport and catalysis events, and physical interactions involving proteins, DNA, RNA, small molecules and complexes. more...

Search Pathway Commons:

To get started, enter a gene name, gene identifier or pathway name in the text box above. For example: p53, P38398 or mTOR.

To restrict your search to specific data sources or specific organisms, update your global filter settings.

Biologists: Browse and search pathways across multiple valuable public pathway databases.

Computational biologists: Download an integrated set of pathways in BioPAX format for global analysis.

Software developers: Build software on top of Pathway Commons using our soon-to-be released web service API. Download and install the cPath software to create a local mirror.

Pathway Commons currently contains the following data sources:

Cancer Cell Map, Release: 1.0 [19-May-06]
Browse

HumanCyc, Release: 10.5 [18-Sep-06]
Browse

NCI / Nature Pathway Interaction Database [01-Jan-07]
Browse

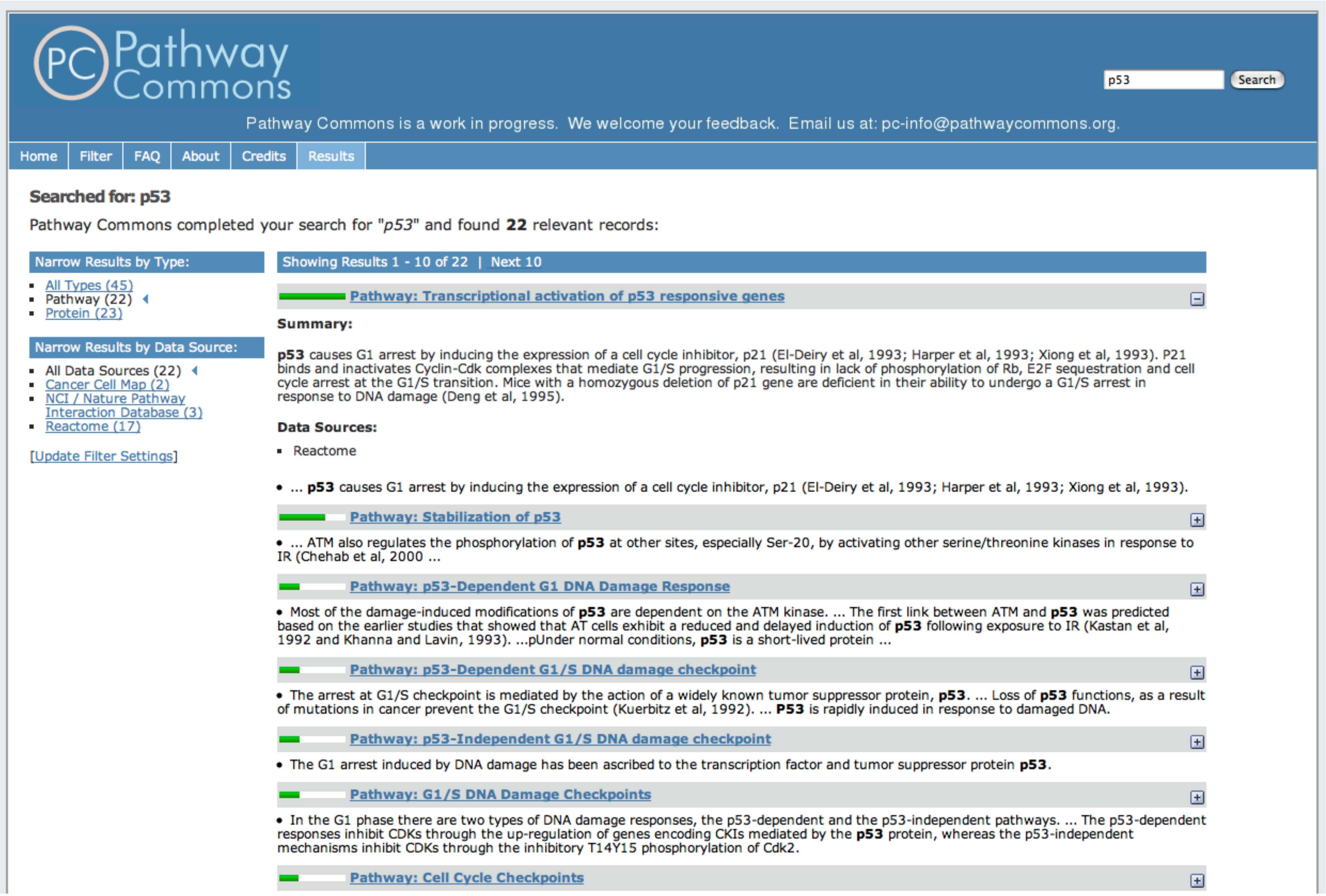
Reactome, Release: 19 [16-Nov-06]
Browse

Pathway Commons Quick Stats:

Number of Pathways:	921
Number of Interactions:	9,924
Number of Physical Entities:	15,515
Number of Organisms:	10

Integration of additional data sources is planned in the near future. For a comprehensive directory of interaction and pathway databases, please refer to Pathguide.

Pathway Commons is built and maintained by the **Computational Biology Center** at **Memorial Sloan-Kettering Cancer Center** and the **University of Toronto**.



Pathway Commons is a work in progress. We welcome your feedback. Email us at: pc-info@pathwaycommons.org.

Searched for: p53

Pathway Commons completed your search for "*p53*" and found **22** relevant records:

Narrow Results by Type:

- All Types (45)
- Pathway (22)
- Protein (23)

Narrow Results by Data Source:

- All Data Sources (22)
- Cancer Cell Map (2)
- NCI / Nature Pathway Interaction Database (3)
- Reactome (17)

[Update Filter Settings]

Showing Results 1 - 10 of 22 | Next 10

Pathway: Transcriptional activation of p53 responsive genes

Summary:

p53 causes G1 arrest by inducing the expression of a cell cycle inhibitor, p21 (El-Deiry et al, 1993; Harper et al, 1993; Xiong et al, 1993). P21 binds and inactivates Cyclin-Cdk complexes that mediate G1/S progression, resulting in lack of phosphorylation of Rb, E2F sequestration and cell cycle arrest at the G1/S transition. Mice with a homozygous deletion of p21 gene are deficient in their ability to undergo a G1/S arrest in response to DNA damage (Deng et al, 1995).

Data Sources:

- Reactome

- ... **p53** causes G1 arrest by inducing the expression of a cell cycle inhibitor, p21 (El-Deiry et al, 1993; Harper et al, 1993; Xiong et al, 1993).

Pathway: Stabilization of p53

- ... ATM also regulates the phosphorylation of **p53** at other sites, especially Ser-20, by activating other serine/threonine kinases in response to IR (Chehab et al, 2000 ...

Pathway: p53-Dependent G1 DNA Damage Response

- Most of the damage-induced modifications of **p53** are dependent on the ATM kinase. ... The first link between ATM and **p53** was predicted based on the earlier studies that showed that AT cells exhibit a reduced and delayed induction of **p53** following exposure to IR (Kastan et al, 1992 and Khanna and Lavin, 1993). ...pUnder normal conditions, **p53** is a short-lived protein ...

Pathway: p53-Dependent G1/S DNA damage checkpoint

- The arrest at G1/S checkpoint is mediated by the action of a widely known tumor suppressor protein, **p53**. ... Loss of **p53** functions, as a result of mutations in cancer prevent the G1/S checkpoint (Kuerbitz et al, 1992). ... **P53** is rapidly induced in response to damaged DNA.

Pathway: p53-Independent G1/S DNA damage checkpoint

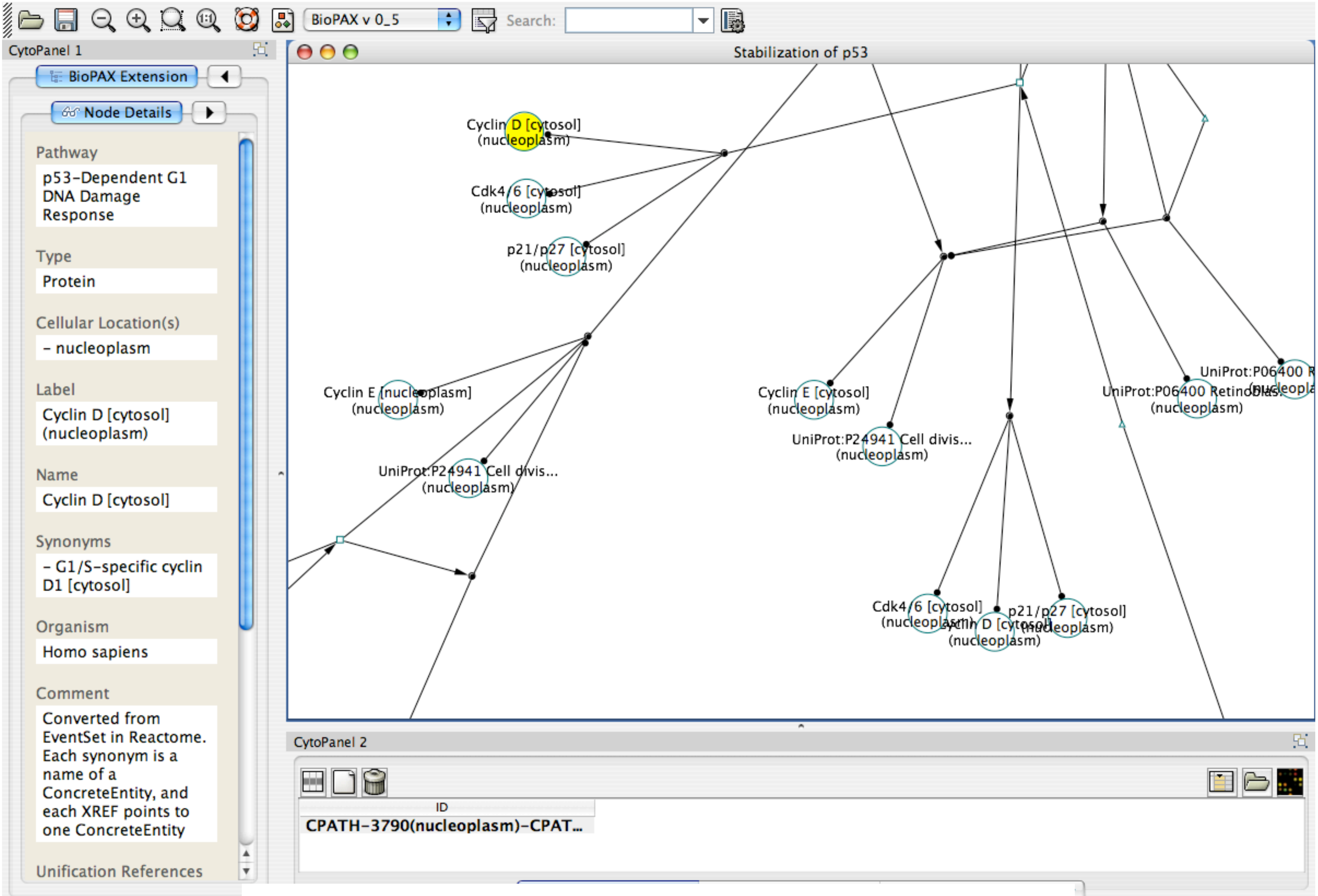
- The G1 arrest induced by DNA damage has been ascribed to the transcription factor and tumor suppressor protein **p53**.

Pathway: G1/S DNA Damage Checkpoints

- In the G1 phase there are two types of DNA damage responses, the p53-dependent and the p53-independent pathways. ... The p53-dependent responses inhibit CDKs through the up-regulation of genes encoding CKIs mediated by the **p53** protein, whereas the p53-independent mechanisms inhibit CDKs through the inhibitory T14Y15 phosphorylation of Cdk2.

Pathway: Cell Cycle Checkpoints

http://pathwaycommons.org



Pathway
p53-Dependent G1 DNA Damage Response

Type
Protein

Cellular Location(s)
- nucleoplasm

Label
Cyclin D [cytosol] (nucleoplasm)

Name
Cyclin D [cytosol]

Synonyms
- G1/S-specific cyclin D1 [cytosol]

Organism
Homo sapiens

Comment
Converted from EventSet in Reactome. Each synonym is a name of a ConcreteEntity, and each XREF points to one ConcreteEntity

Unification References

ID
CPATH-3790(nucleoplasm)-CPAT...

http://pathwaycommons.org

Pathway Commons Status

Cancer Cell Map, Release: 1.0 [19-May-06]
Browse

HumanCyc, Release: 10.5 [18-Sep-06]
Browse

NCI / Nature Pathway Interaction Database
[30-Apr-07]
Browse

Reactome, Release: 21 [15-May-07]
Browse

Pathway Commons Quick Stats:

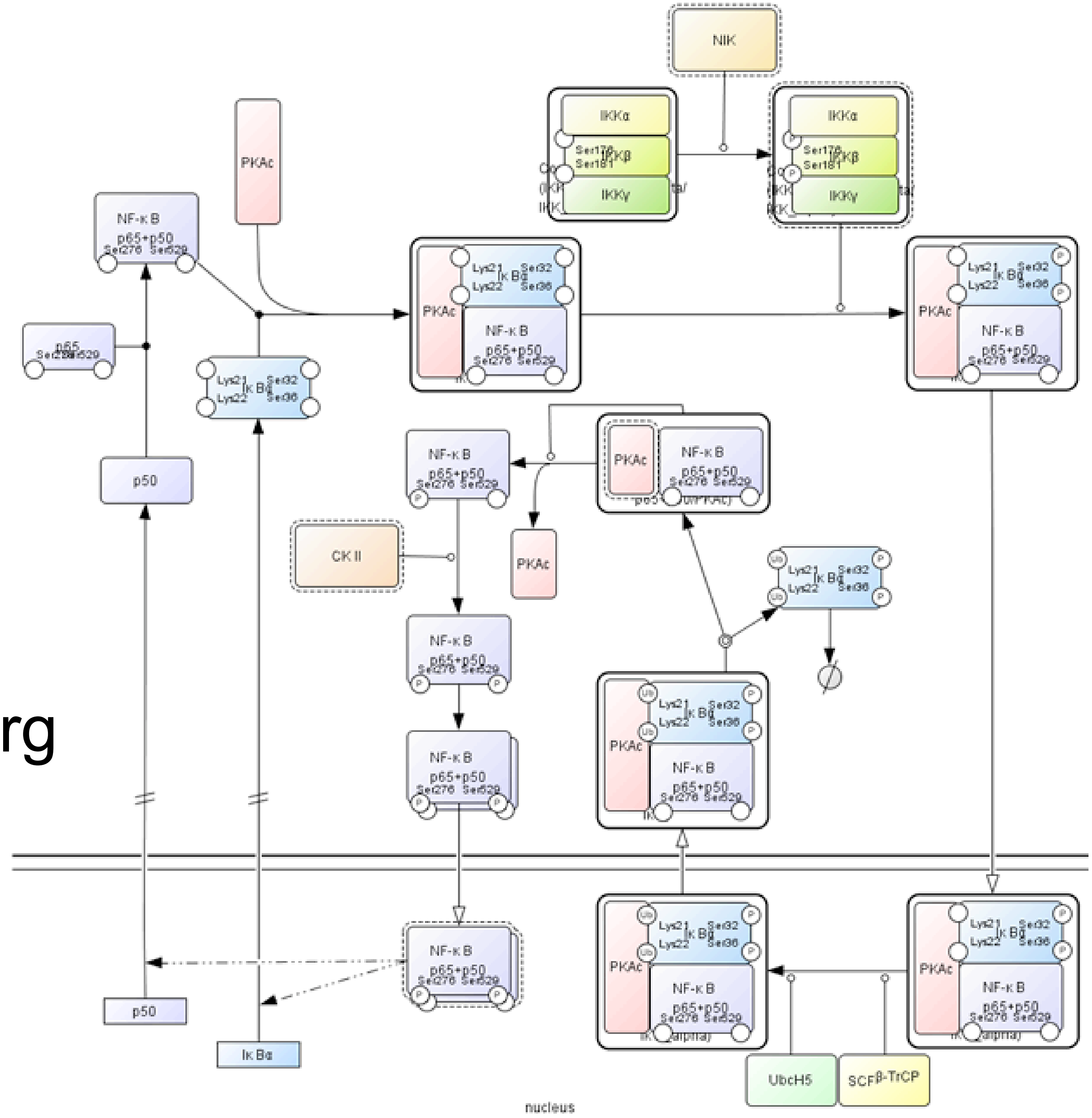
Number of Pathways:	994
Number of Interactions:	12,550
Number of Physical Entities:	16,933
Number of Organisms:	9

- Next Databases (human)
  - BioCarta, KEGG
  - Protein-protein interactions (IntAct, MINT)
  - iHOP annotation
- Web service API under development
  - getNeighbors
  - getPathwayList
- Neighborhood visualization
- Cytoscape integration

Open Challenges

- Data: Author entry systems
  - From individual publications (The Cashew Prize)
  - For pathways (review)
  - Curator tools (advanced)
- Semantic integration (ID resolution)
- Visualization
  - Pathway diagrams (SBGN)
  - Automated layout
- Algorithms for compound graphs
- Linking discrete and dynamic representations
  - Including use by modelers

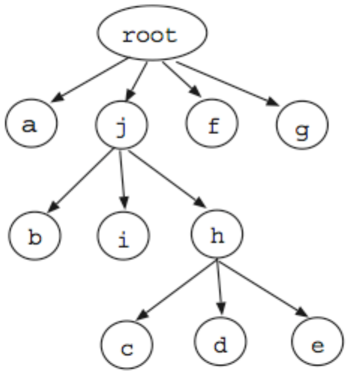
Systems Biology Graphical Notation



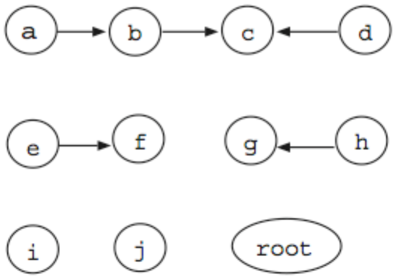
http://sbgn.org

In progress

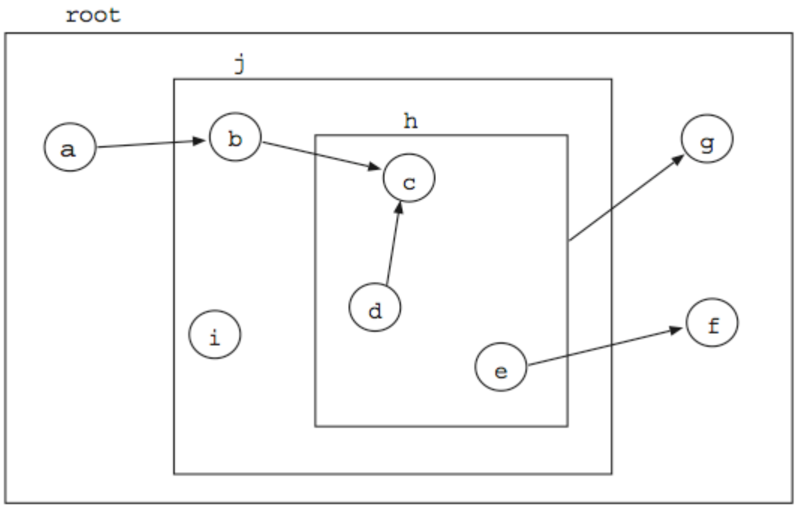
Compound Graph Algorithms



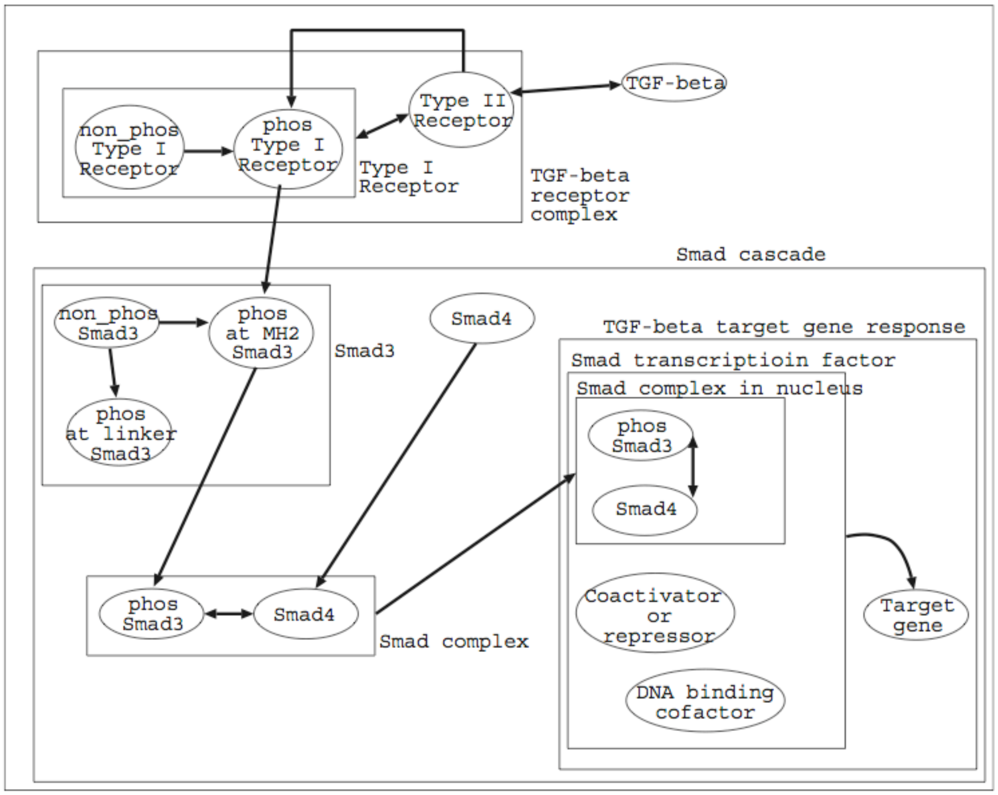
decomposition tree T



interaction graph G

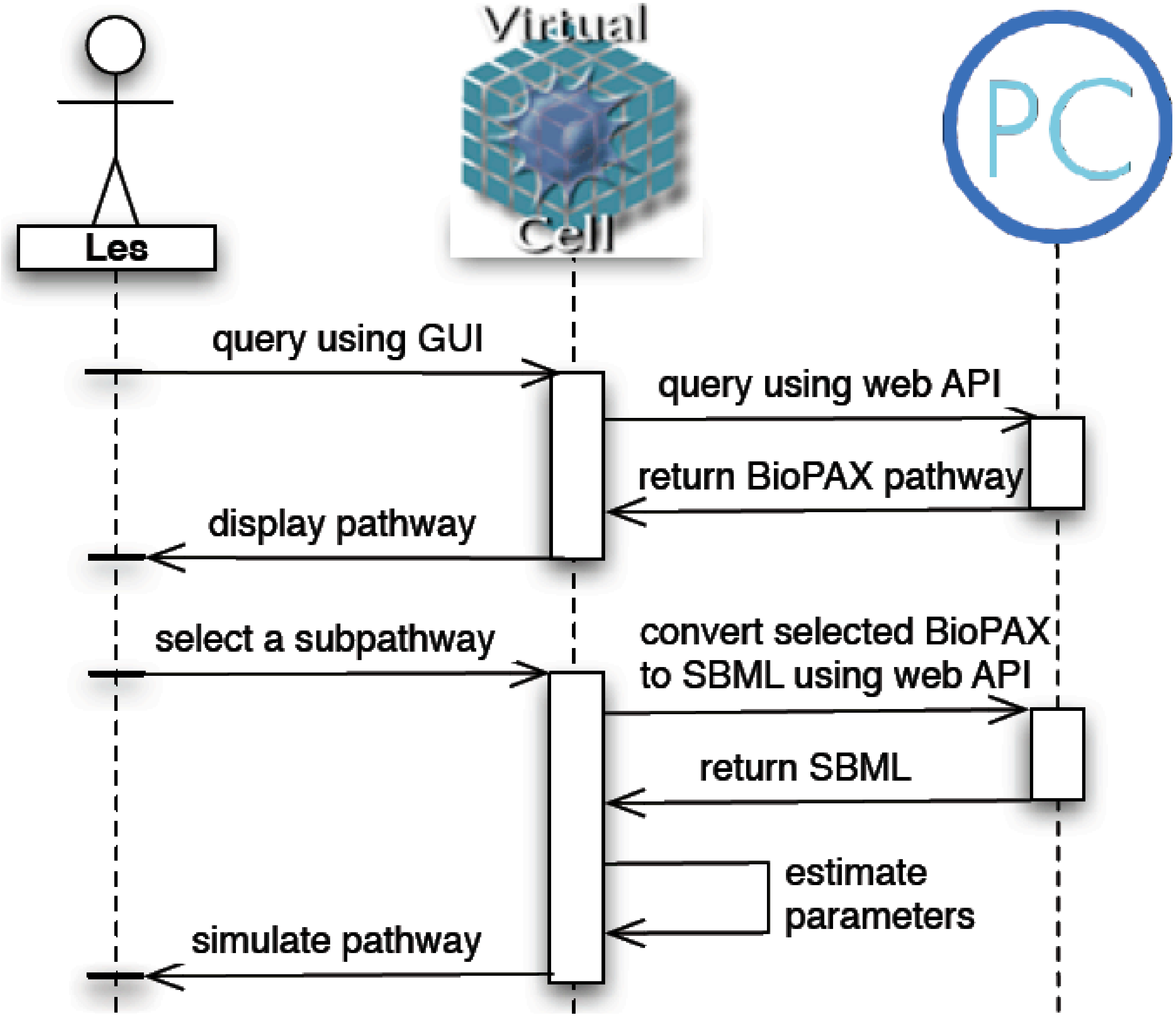


compound graph CG



TGF-beta/Smad signal transduction pathway(root)

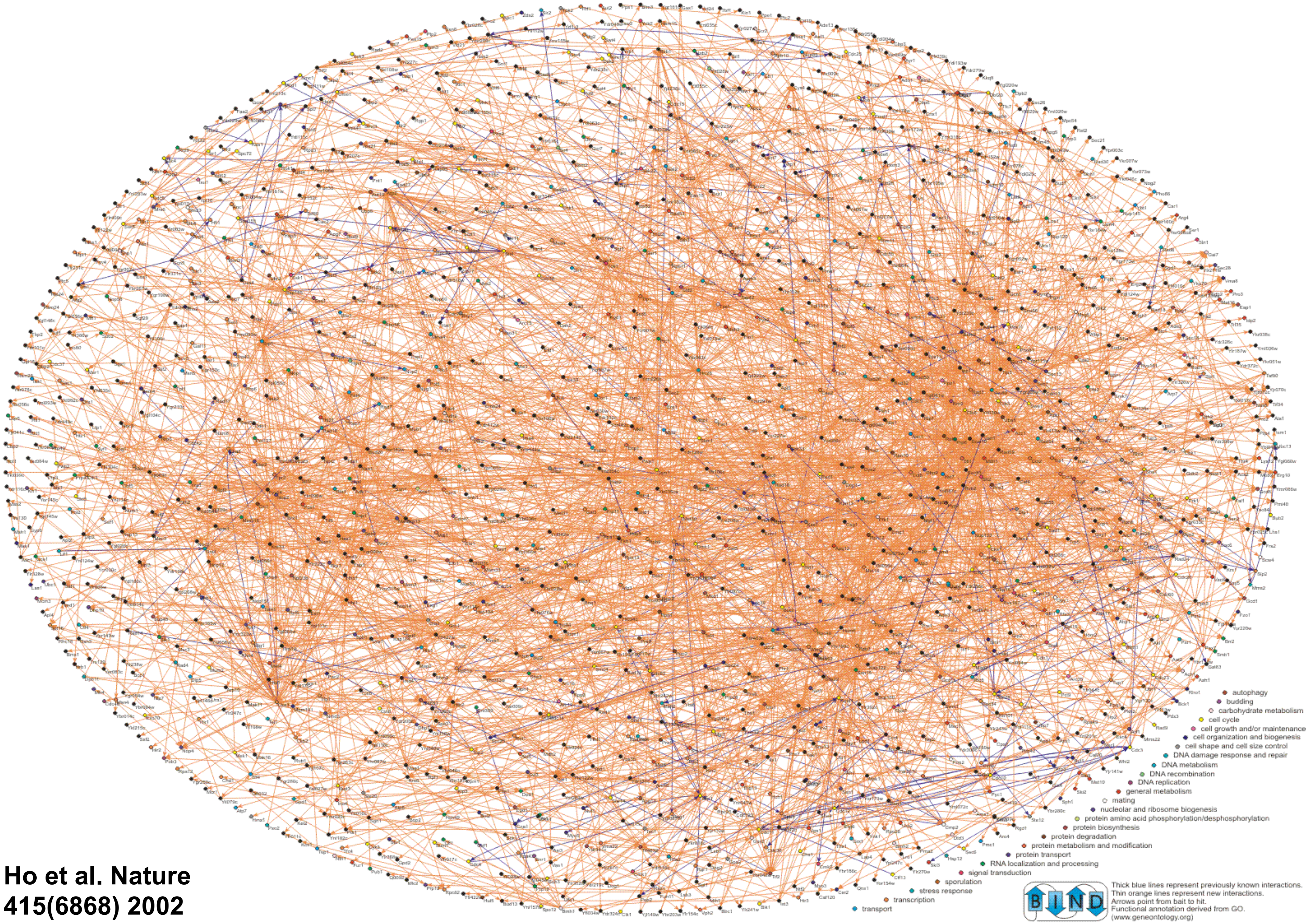
Fukuda K, Takagi T (Bioinformatics 2001)


Les
Virtual
Cell
PC
query using GUI
query using web API
return BioPAX pathway
display pathway
select a subpathway
convert selected BioPAX
to SBML using web API
return SBML
estimate
parameters
simulate pathway

Future Pathway Analyses

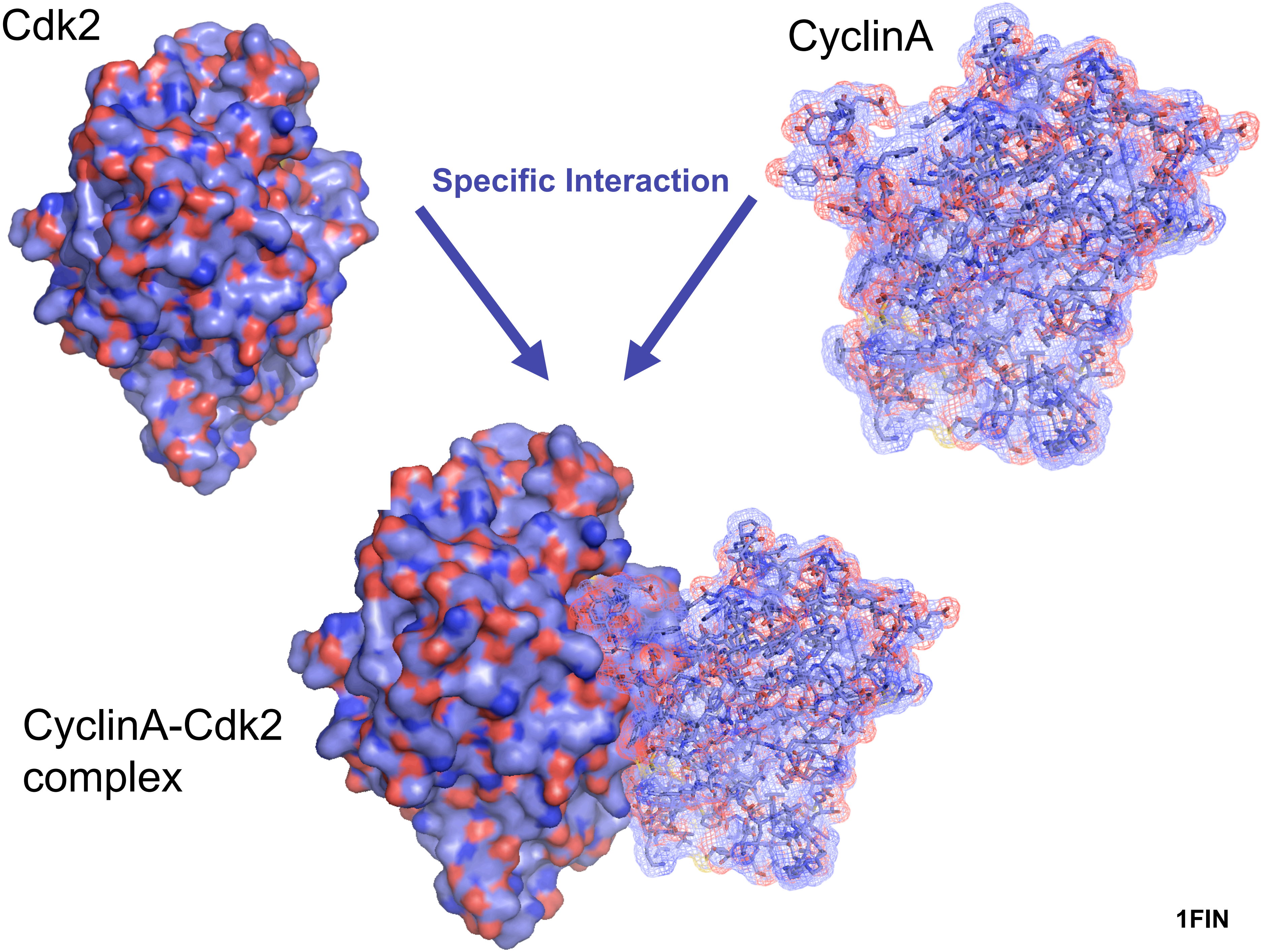
- Find known pathways in new species using Pathblast (www.pathblast.org)
- Find active pathways from molecular profiles (e.g. active modules, activity centers, GOALIE)
- Molecular interaction and pathway prediction from genome sequence
- Pathway simulation to predict drugs and drug combinations that activate or inhibit specific biological processes

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



Thick blue lines represent previously known interactions.
Thin orange lines represent new interactions.
Arrows point from bait to hit.
Functional annotation derived from GO.
(www.geneontology.org)

Ho et al. Nature 415(6868) 2002



Cdk2

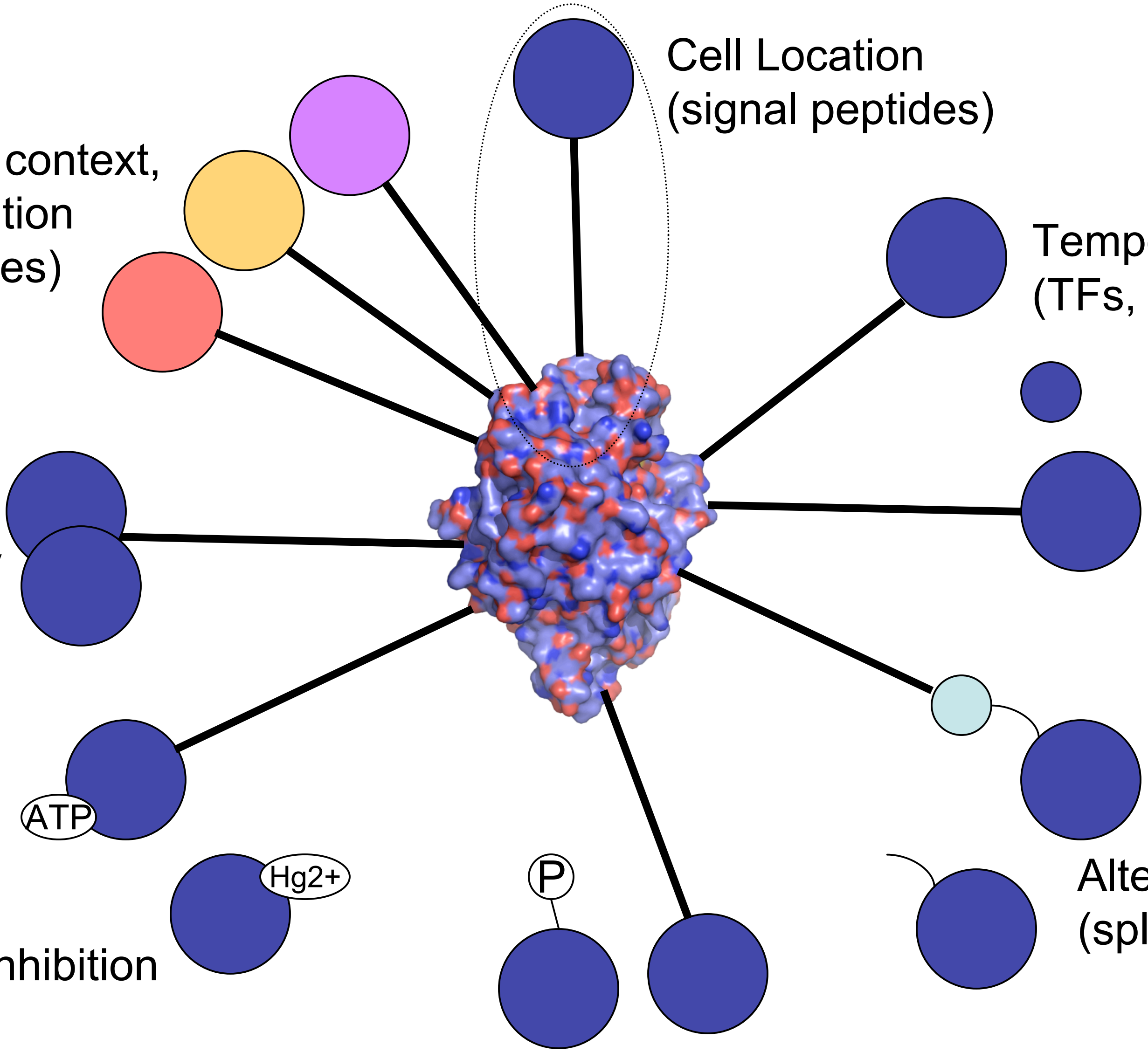
Specific Interaction

CyclinA

CyclinA-Cdk2
complex

1FIN

Towards More Biologically Relevant Networks



Cell Location
(signal peptides)

Partner context,
competition
(cell types)

Temporal expression
(TFs, miRNA)

Concentration
(production,
degradation
rates)

Cooperativity

Allostery

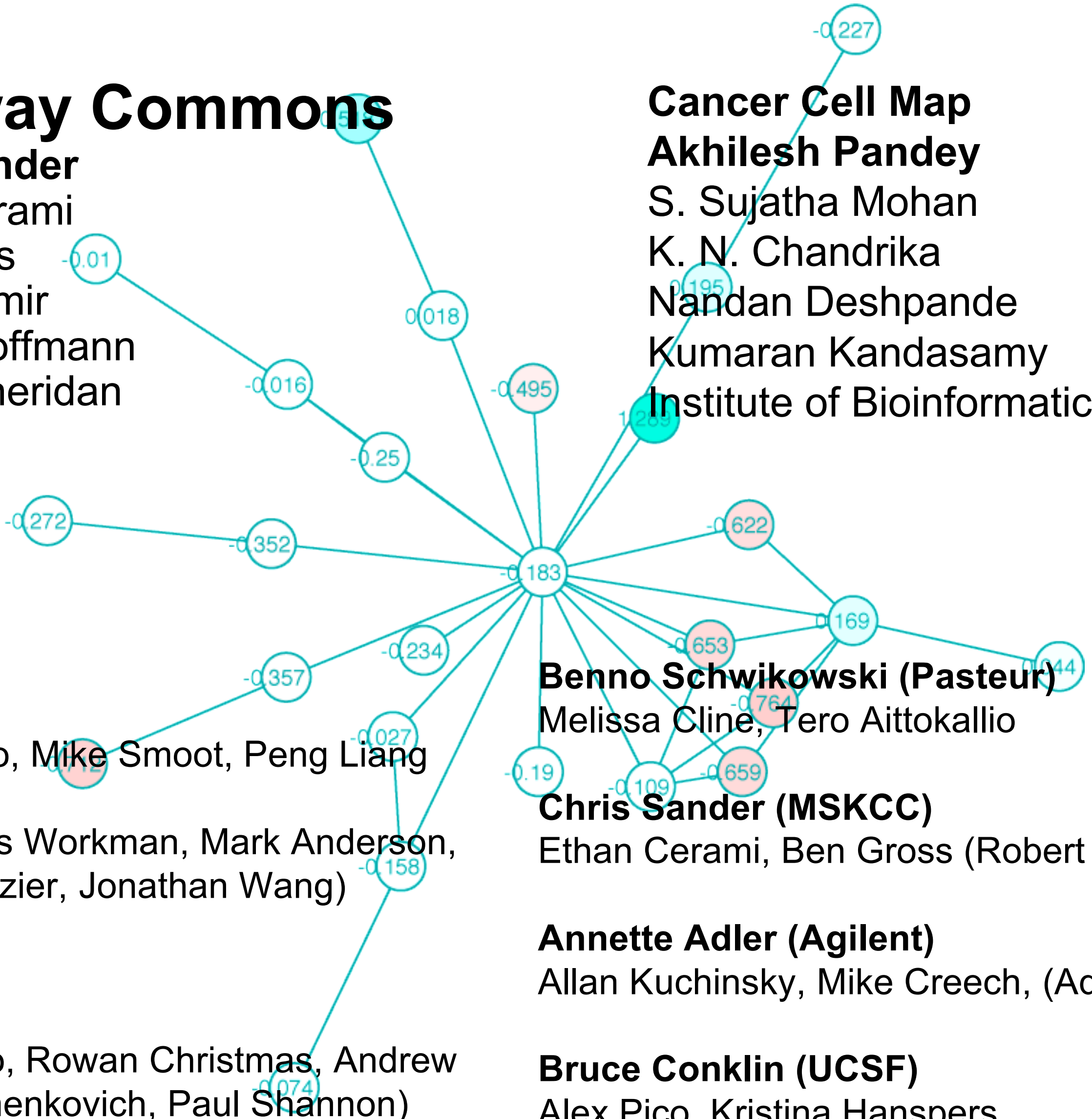
Inhibition

Alternative splicing
(splicing factors)

Post-translational modification (enzymes)

Where we want to be with cellular data integration and visualization…

Acknowledgements



Pathway Commons

Chris Sander
Ethan Cerami
Ben Gross
Emek Demir
Robert Hoffmann
Robert Sheridan

Cancer Cell Map
Akhilesh Pandey
S. Sujatha Mohan
K. N. Chandrika
Nandan Deshpande
Kumaran Kandasamy
Institute of Bioinformatics

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)

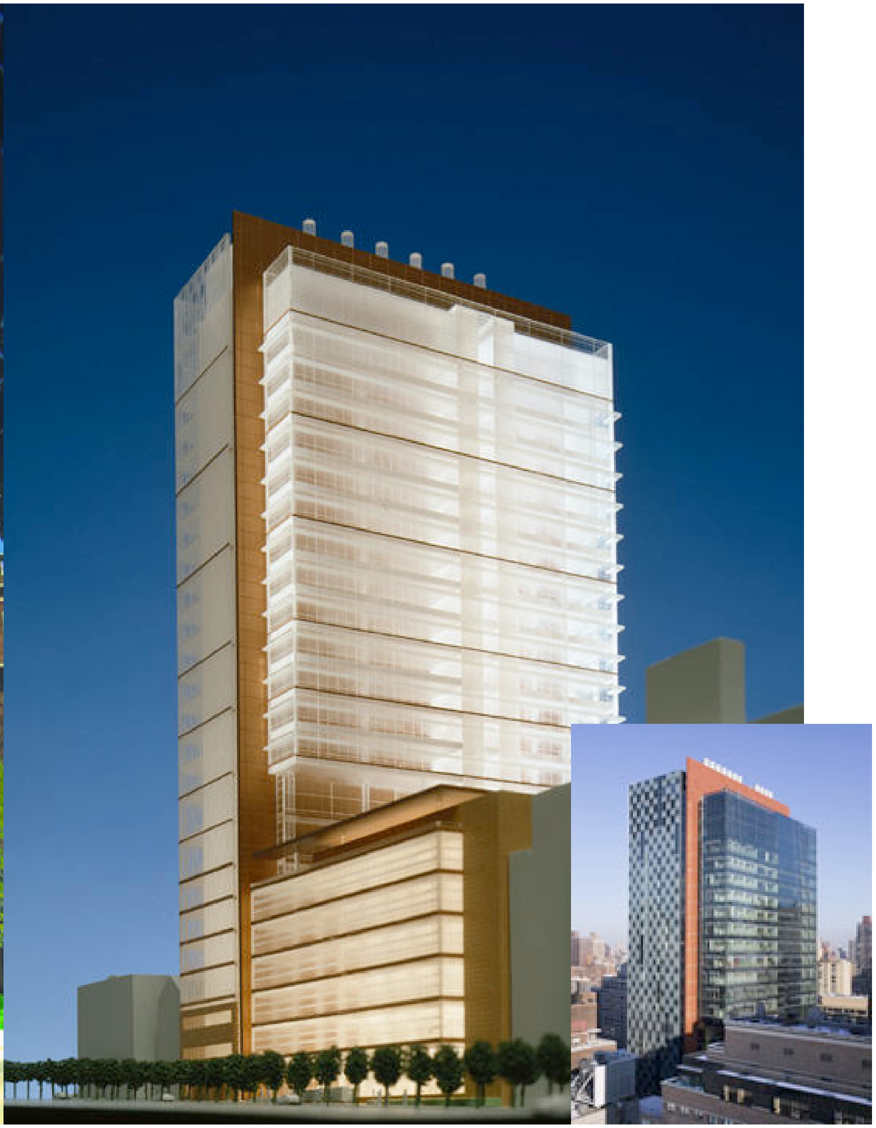
Benno Schwikowski (Pasteur)
Melissa Cline, Tero Aittokallio

Chris Sander (MSKCC)
Ethan Cerami, Ben Gross (Robert Sheridan)

Annette Adler (Agilent)
Allan Kuchinsky, Mike Creech, (Aditya Vailaya),

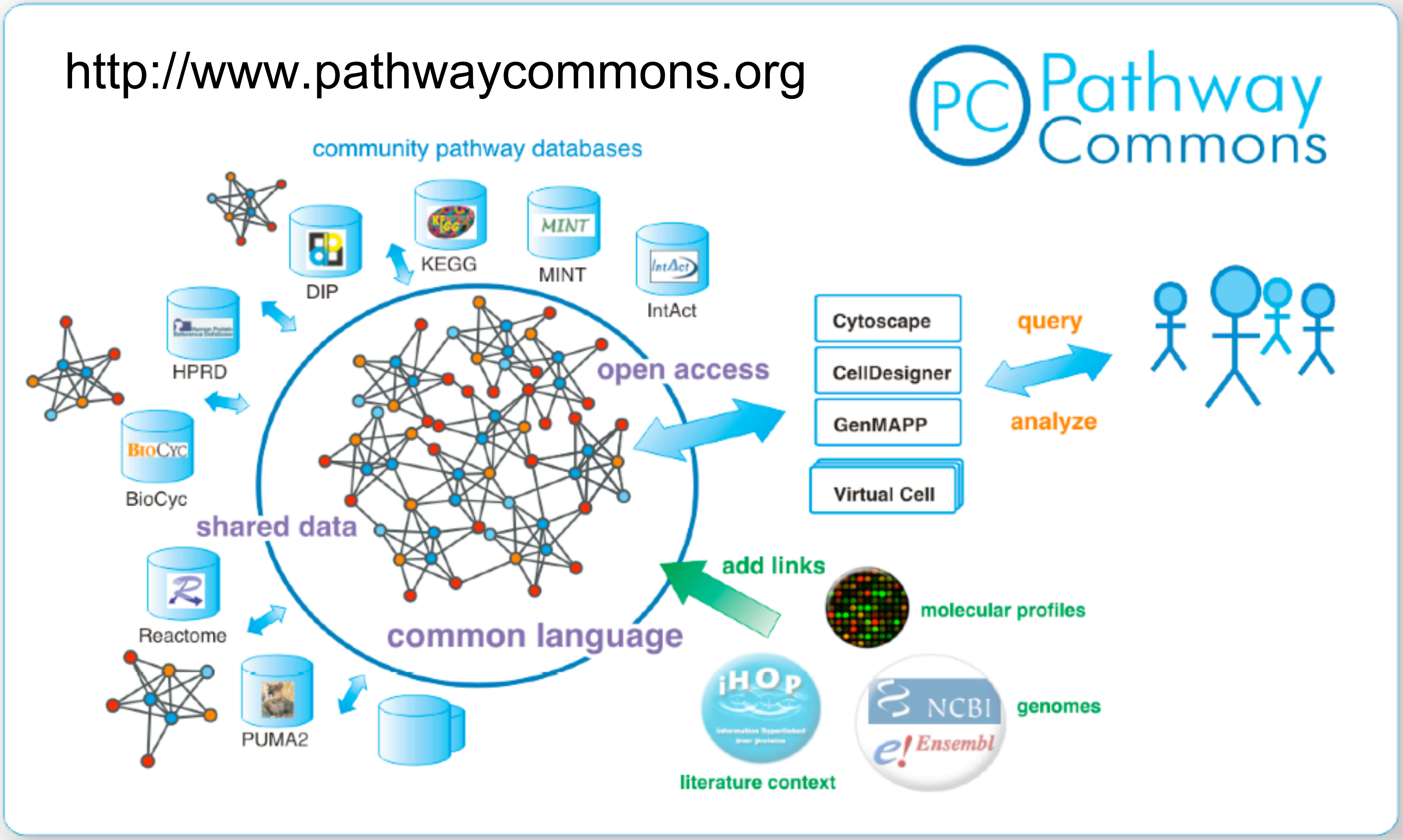
Bruce Conklin (UCSF)
Alex Pico, Kristina Hanspers
John ‘Scooter’ Morris (Ferrin lab, UCSF)

Donnelly Center for Cellular
and Biomolecular Research
University of Toronto
5 open faculty positions



Computational Biology Center
Memorial Sloan Kettering Cancer Center
New York
3 open faculty positions

Aim: Convenient Access to Pathway Information



Facilitate creation and communication of pathway data
Aggregate pathway data in the public domain
Provide easy access for pathway analysis

Long term: Converge to integrated cell map