

<http://www.pathwaycommons.org>

Pathway Commons

A public library of biological pathways

July.14 2008 - Bioinformatics to Systems Biology 2008

Gary Bader
University of Toronto

<http://baderlab.org>

Chris Sander
MSKCC, New York

<http://cbio.mskcc.org>



MEMORIAL SLOAN-KETTERING
CANCER CENTER

computational biology center
.....

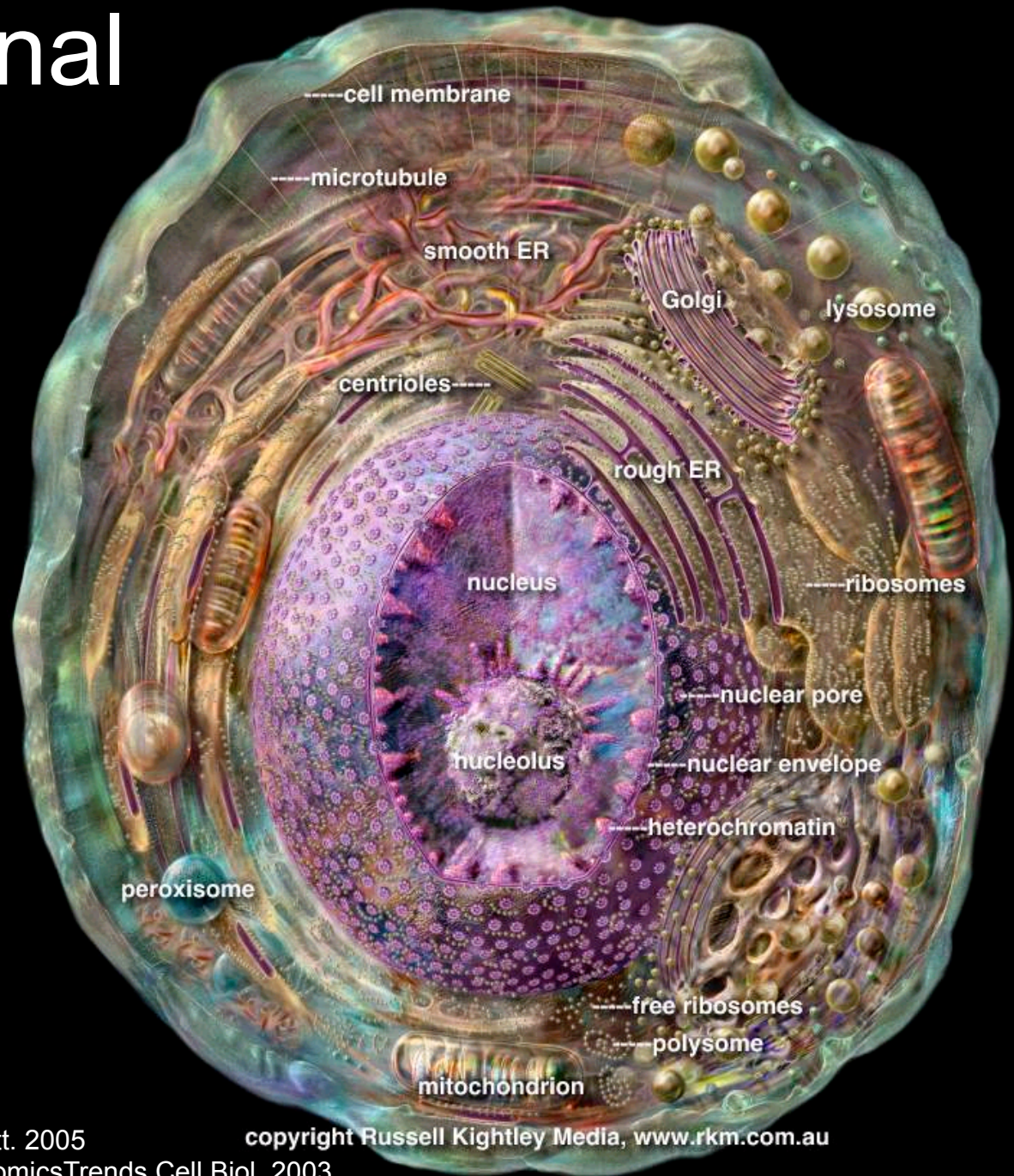
Computational Cell Map

Map the cell

- Predict map from genome
- Multiple perturbation mapping
- Active cell map
- Map visualization and analysis software

Read map to understand

- Cell processes
- Gene function
- Disease effects
- Map evolution



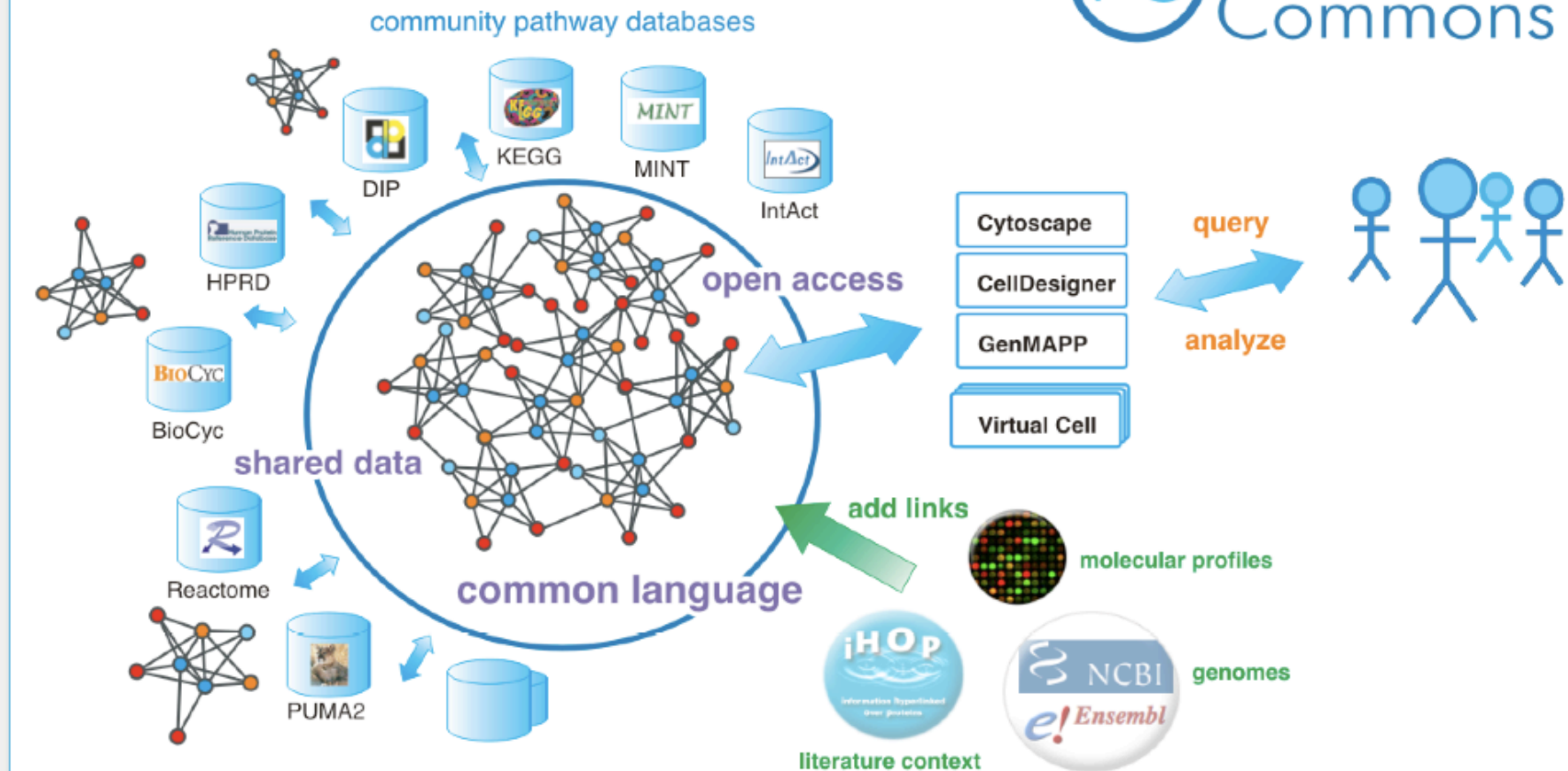
Cary MP et al. Pathway information... FEBS Lett. 2005

Bader GD et al. Functional genomics and proteomics Trends Cell Biol. 2003

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

Aim: Convenient Access to Pathway Information

<http://www.pathwaycommons.org>

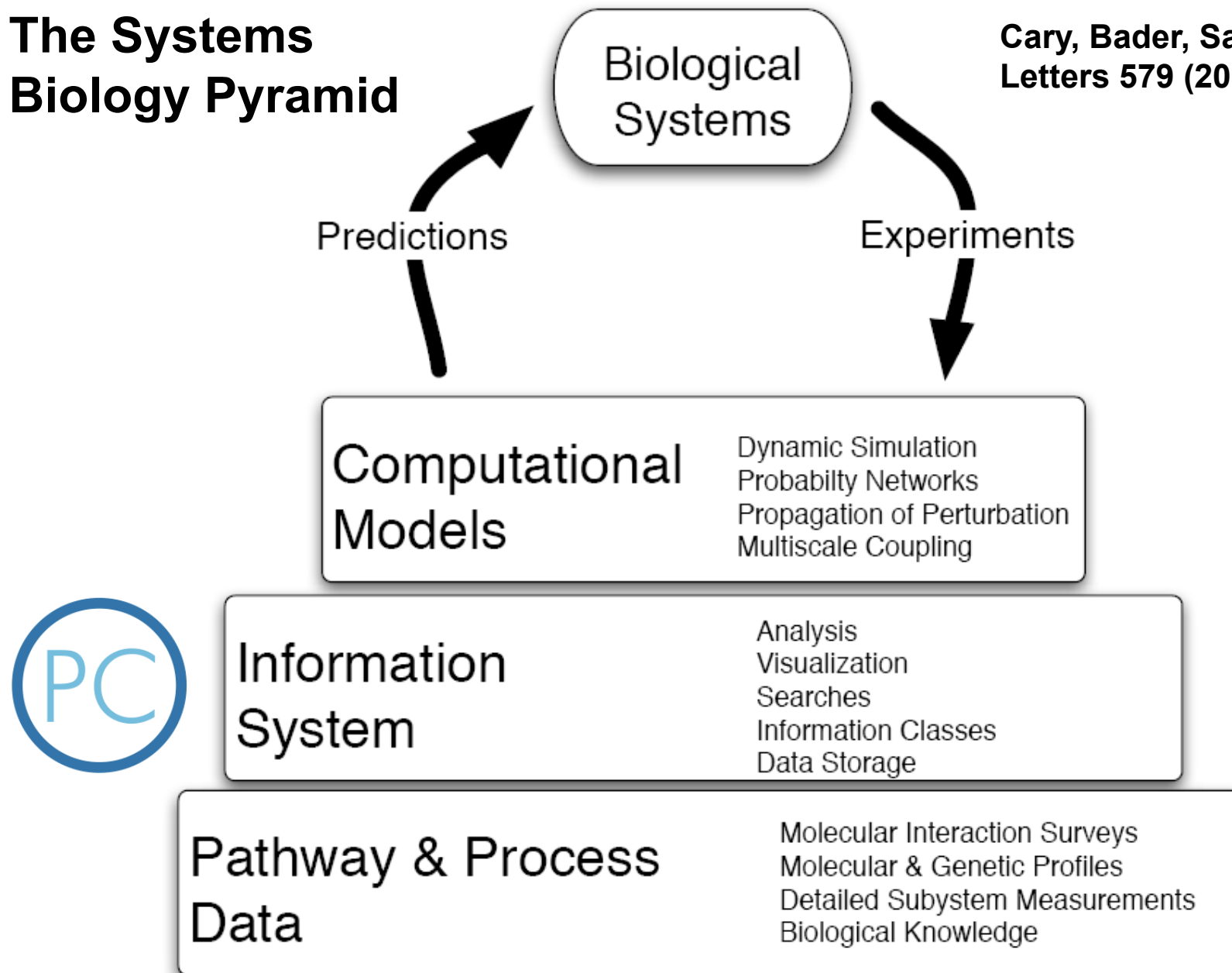


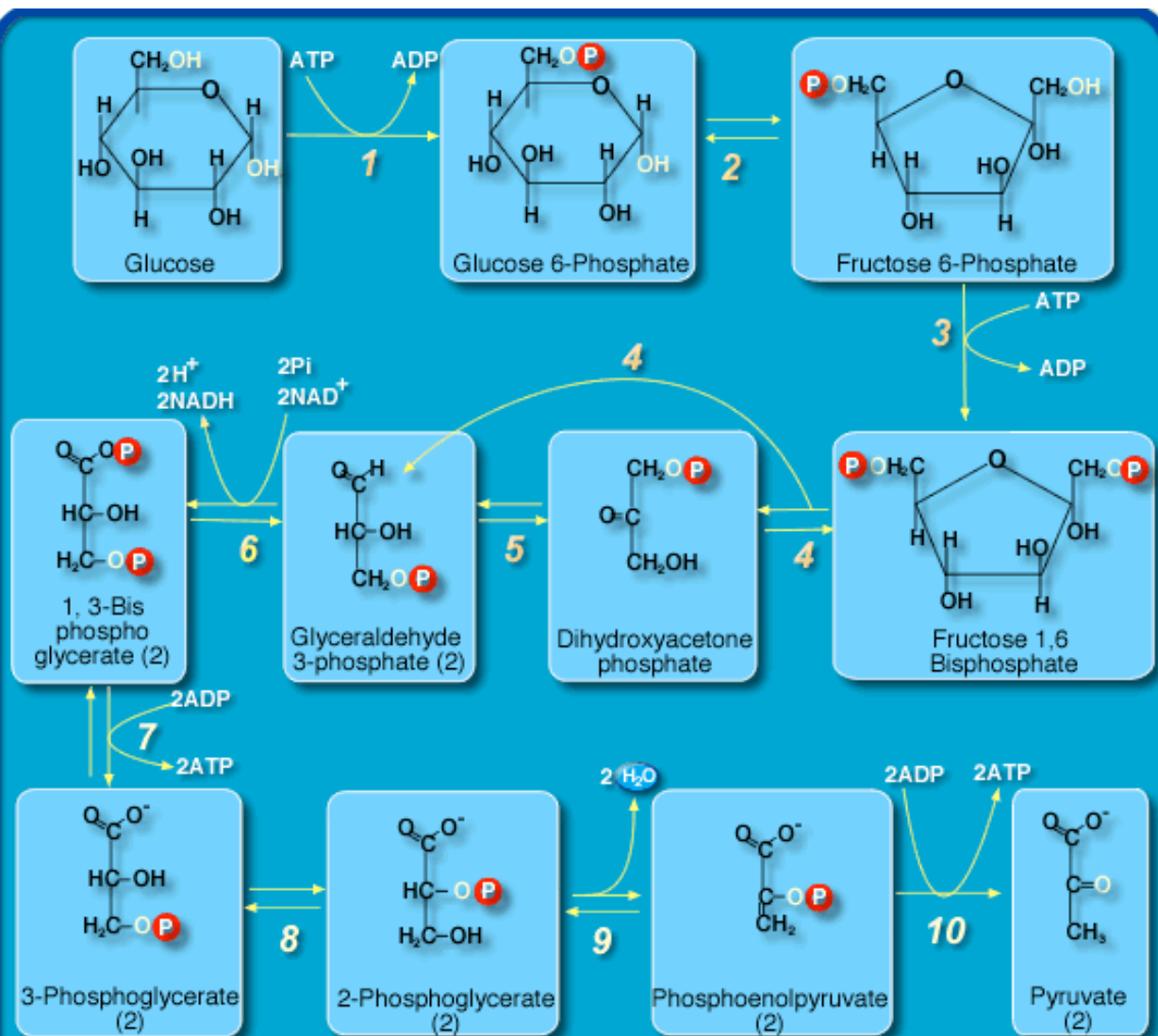
Facilitate creation and communication of pathway data
Aggregate pathway data in the public domain
Provide easy access for pathway analysis

Long term: Converge
to integrated cell map

The Systems Biology Pyramid

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





ENZYMES

- 1 Hexokinase
- 2 Glucose Phosphate Isomerase
- 3 Phosphofructokinase
- 4 Fructose diphosphate aldolase

● Preparatory phase

- 5 Triose phosphate Isomerase
- 6 Glyceraldehyde Phosphate Dehydrogenase

● Payoff phase

- 7 Phosphoglycerate Kinase
- 8 Phosphoglyceromutase
- 9 Enolase
- 10 Pyruvate Kinase



Signaling Pathway

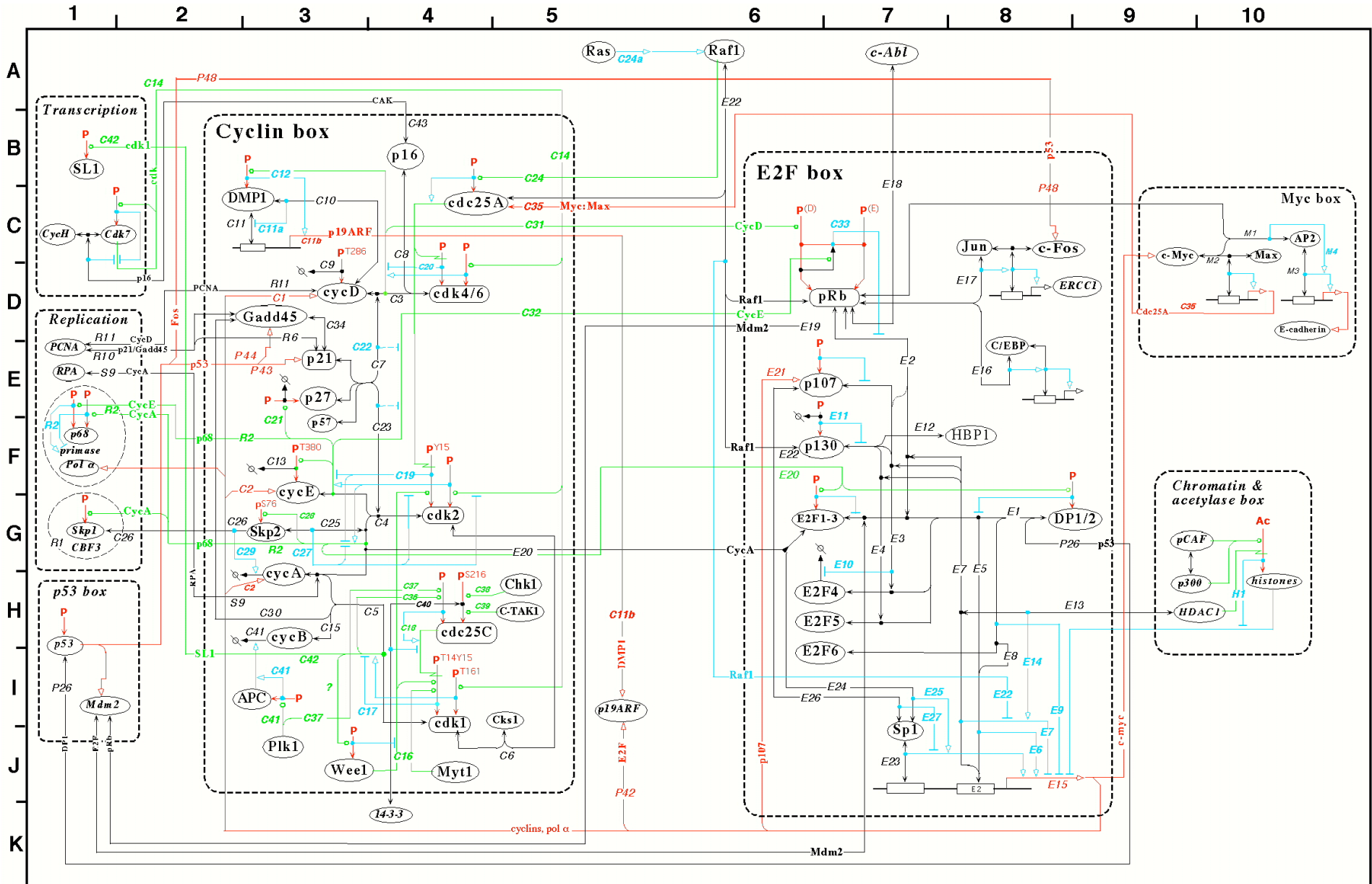
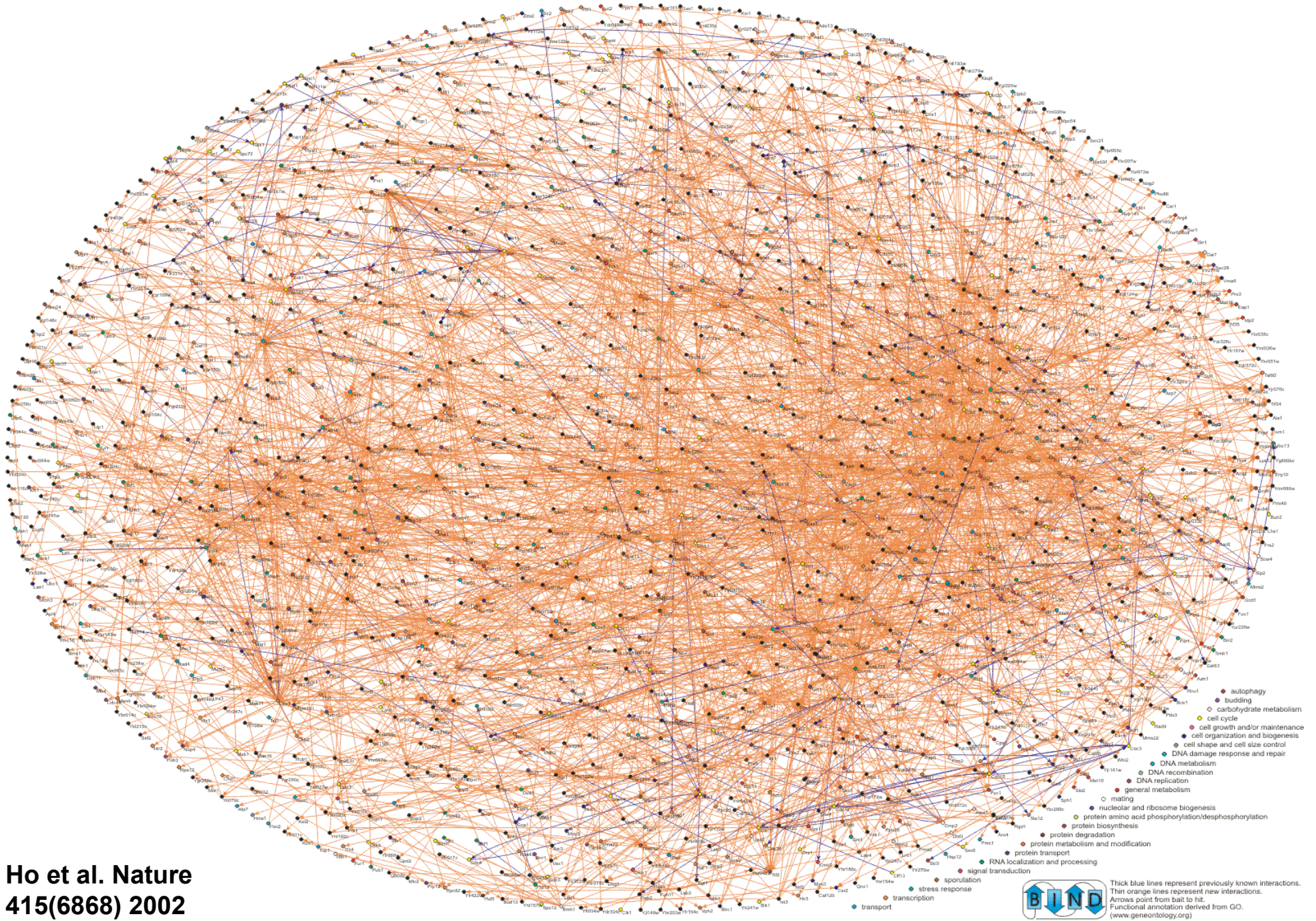


Figure 6A: The Cyclin - E2F cell cycle control system (version 3a - June 8, 1999) http://discover.nci.nih.gov/kohnk/interaction_maps.html

March 8, 2005



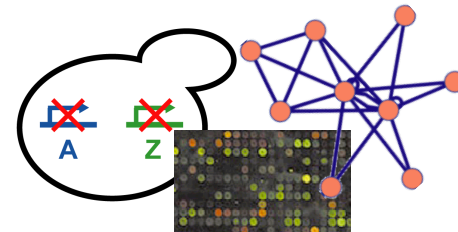
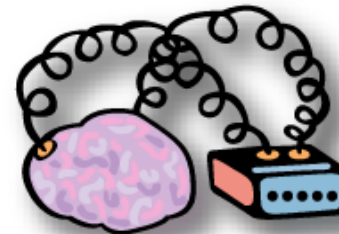
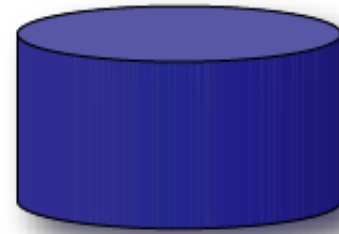
Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry



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



http://pathguide.org

Vuk Pavlovic

Pathguide» the pathway resource list

Home BioPAX cBio MSKCC

>240 Pathway Databases!

Navigation

- Protein-Protein Interactions
- Metabolic Pathways
- Signaling Pathways
- Pathway Diagrams
- Transcription Factors / Gene Regulatory Networks
- Protein-Compound Interactions
- Genetic Interaction Networks
- Protein Sequence Focused
- Other

Search

Organisms:

Availability:

Standards:

Statistics

Analyze Pathguide

Contact

Comments, Questions, Suggestions are Always Welcome!

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](#).

Get the Stats
Detailed Pathguide resource statistics now available

Pathguide Published
Please cite the [Pathguide](#).

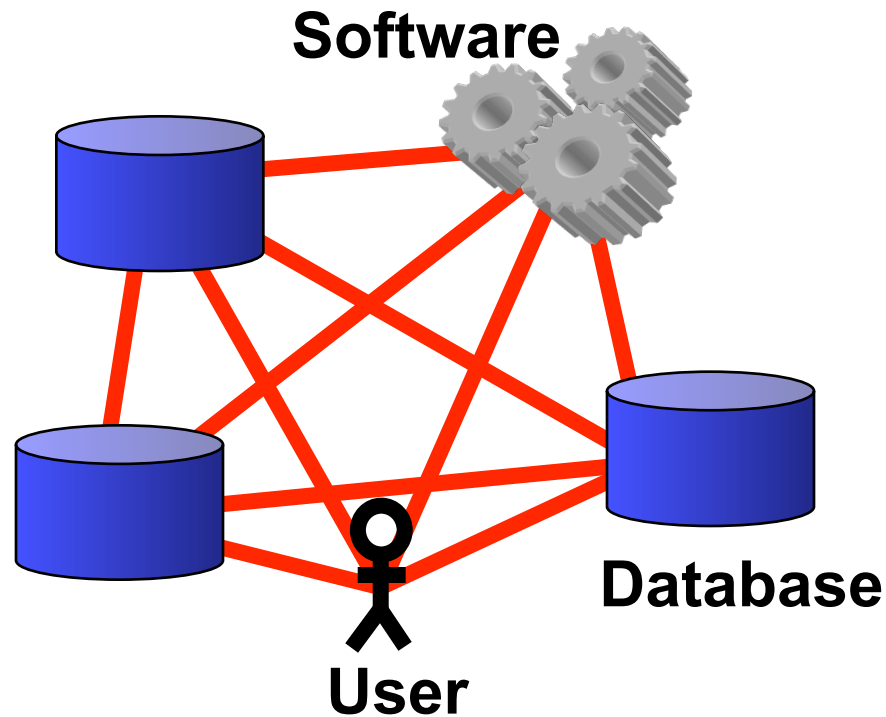
Protein-Protein Interactions

Database Name (Order: alphabetically | [by web popularity](#))

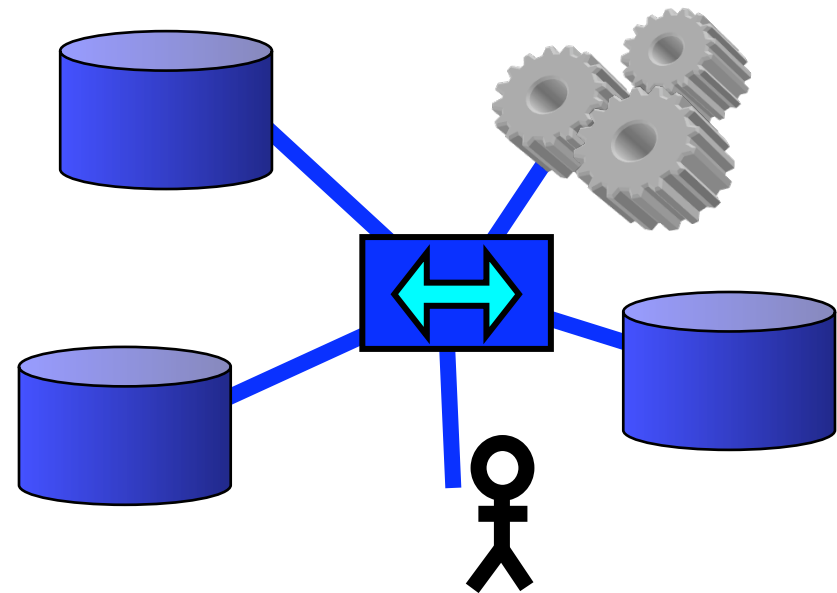
Database Name	Full Record	Availability	Standards
3DID - 3D interacting domains	Details	Free	
ABCdb - Archaea and Bacteria ABC transporter database	Details	Free	
AfCS - Alliance for Cellular Signaling Molecule Pages Database	Details	Free	
AllFuse - Functional Associations of Proteins in Complete Genomes	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 oligomeri	Details		
DopaNet - DopaNet	Details		
DRC - Database of Ribosomal	Details		
DSM - Dynamic Signaling Maps	Details		
FIMM - Functional Molecular Im	Details		
FusionDB - Prokaryote Gene Fu	Details		

- Varied formats, representation, coverage
- Pathway data extremely difficult to combine and use

Biological Pathway Exchange (BioPAX)



Before BioPAX
>100 DBs and tools
Tower of Babel



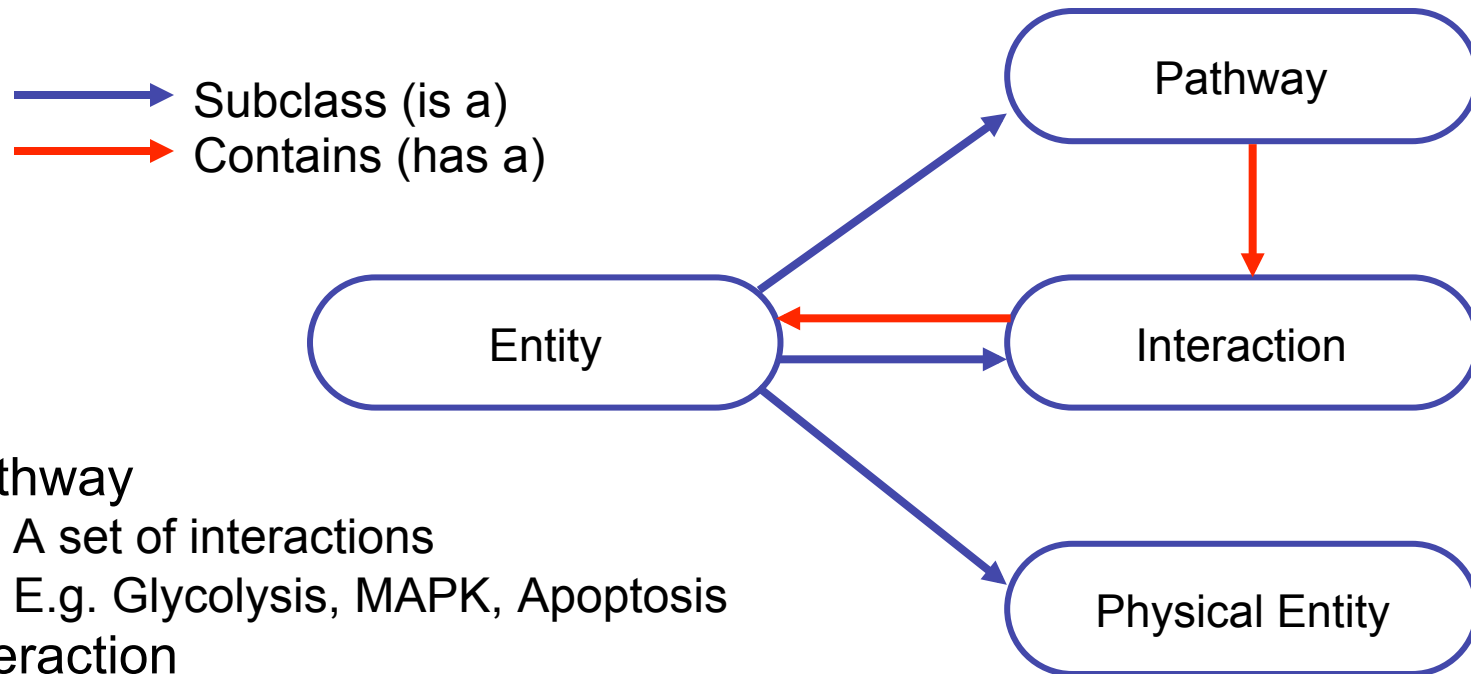
After BioPAX
Unifying language

Reduces work, promotes collaboration, increases accessibility

BioPAX Pathway Language

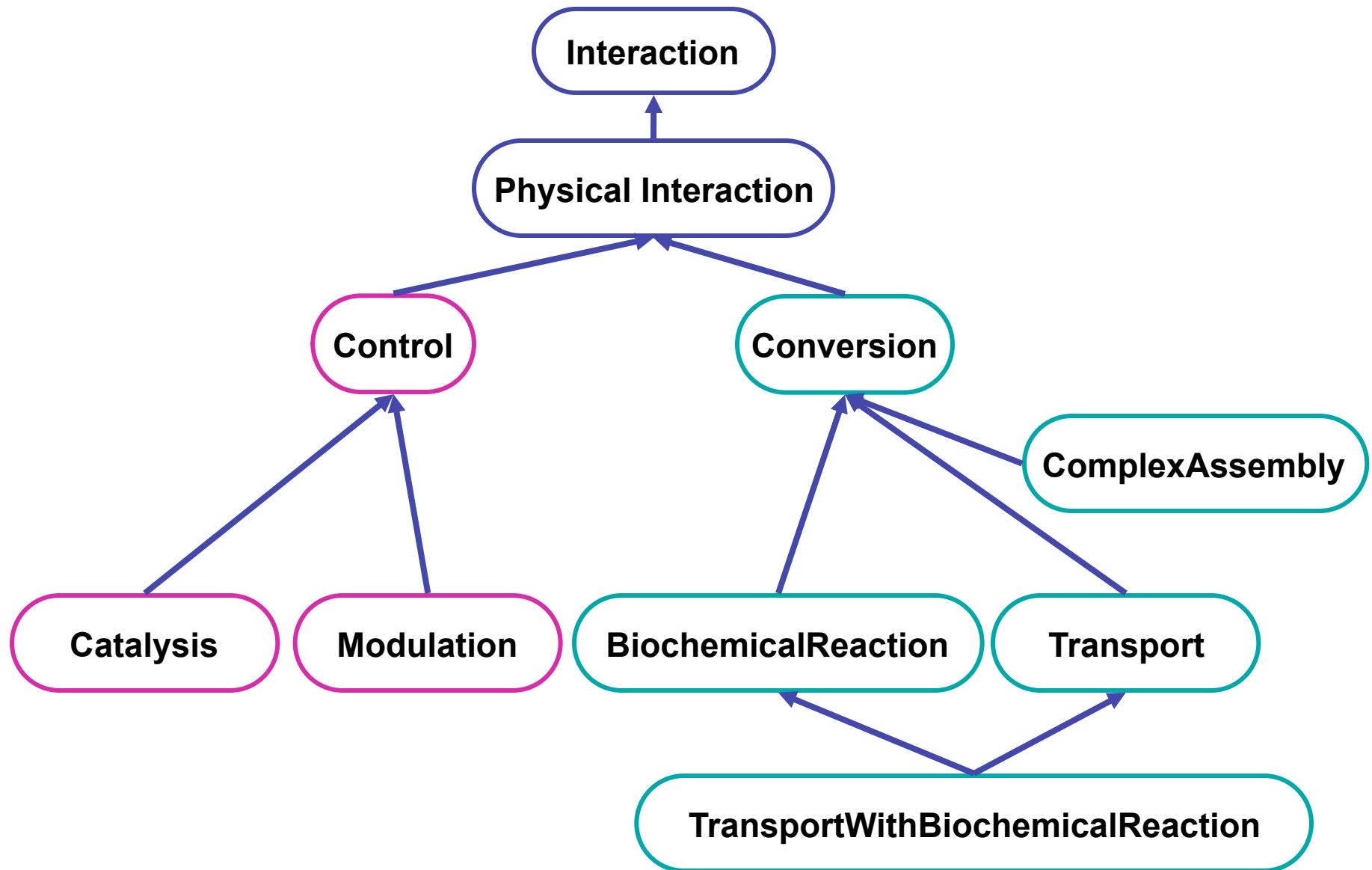
- Represent:
 - Metabolic pathways
 - Signaling pathways
 - Protein-protein, molecular interactions
 - Gene regulatory pathways
 - Genetic interactions
- Community effort: pathway databases distribute pathway information in standard format

BioPAX Structure



- 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



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

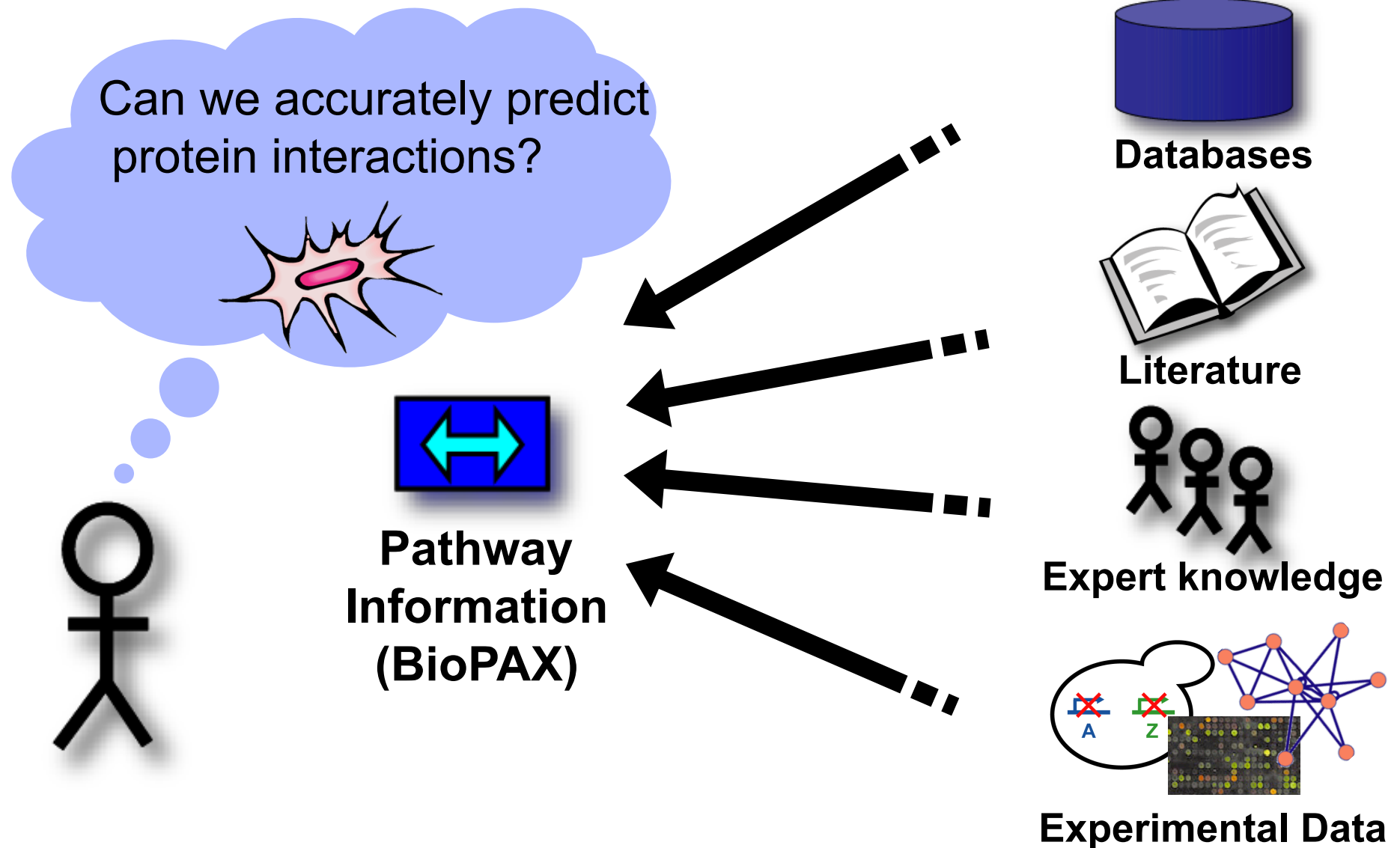
Protégé, U.Manchester, Stanford

Grants/Support

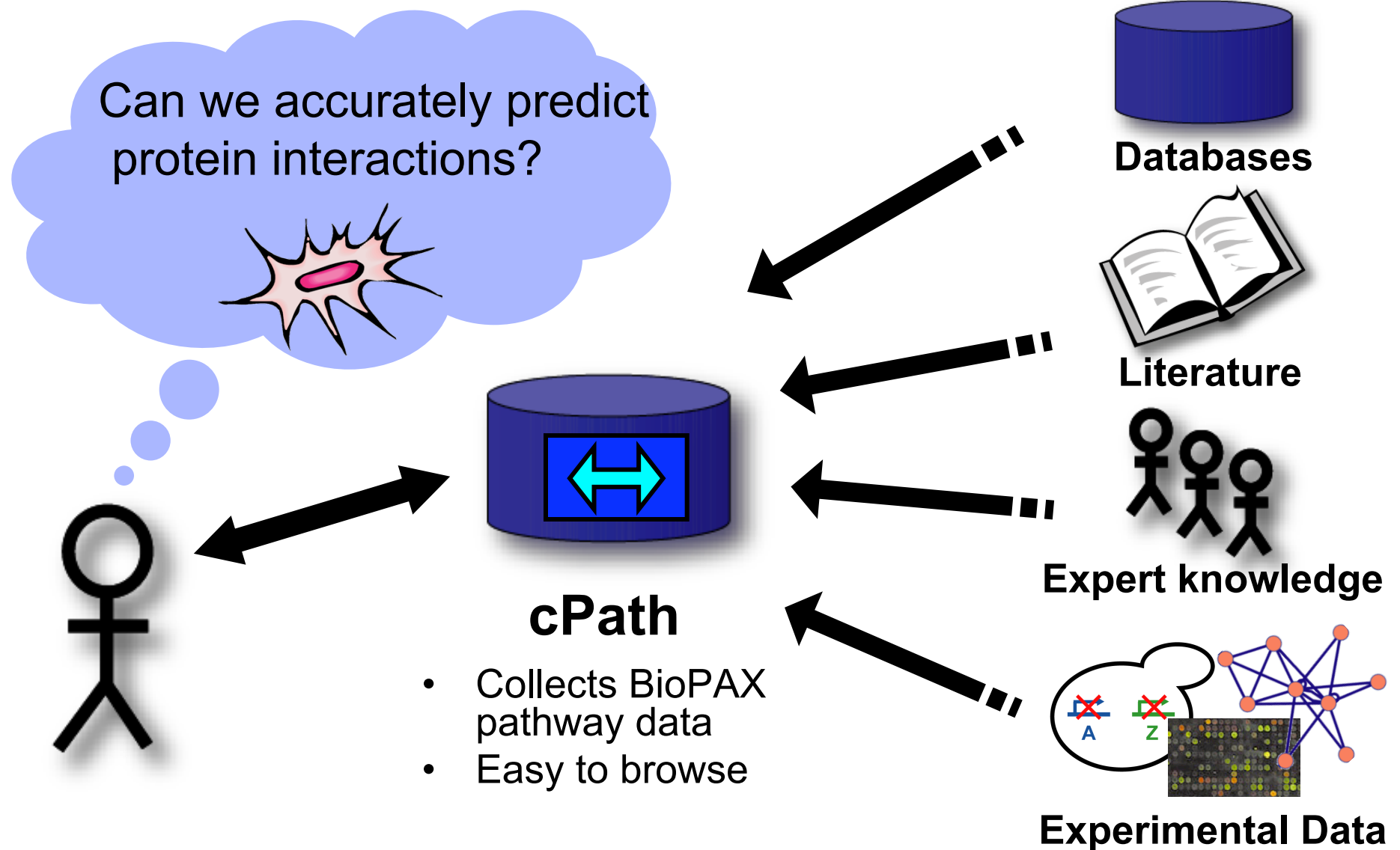
- Department of Energy (Workshop)
- caBIG



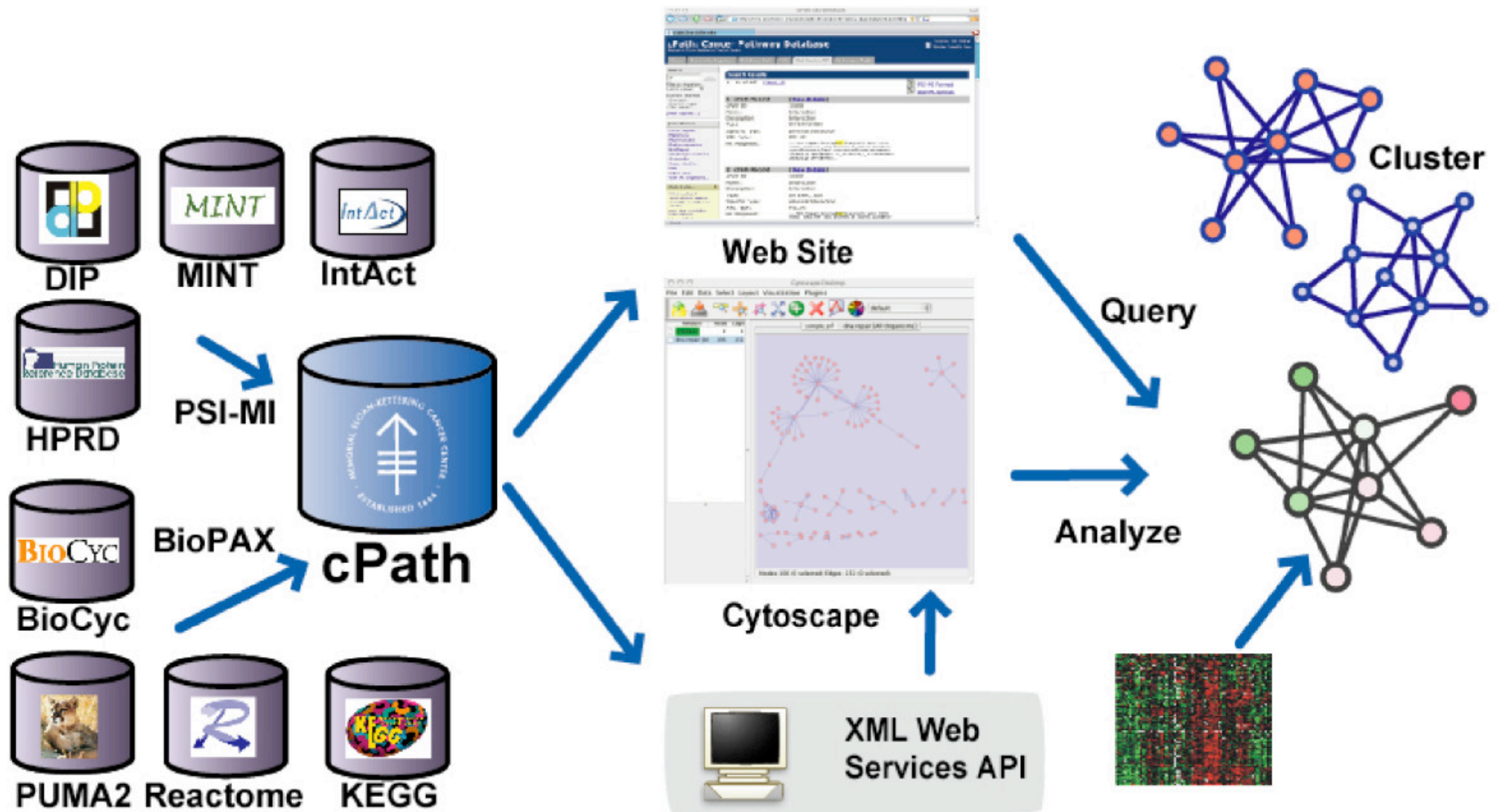
Using Pathway Information



Using Pathway Information



cPath Pathway Database Software



Collect → **Browse / Query** → **Analyze**

cPath Key Features

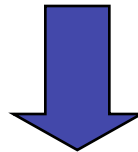
- Identifier mapping system e.g. proteins
- Scalable pathway data aggregation
- Simple web interface for browse and query
- Standard web service API for application communication
- 100% open source
 - Java, Tomcat, MySQL, Lucene, Struts, YUI
- Local installation and customization

<http://cbio.mskcc.org/cpath>

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

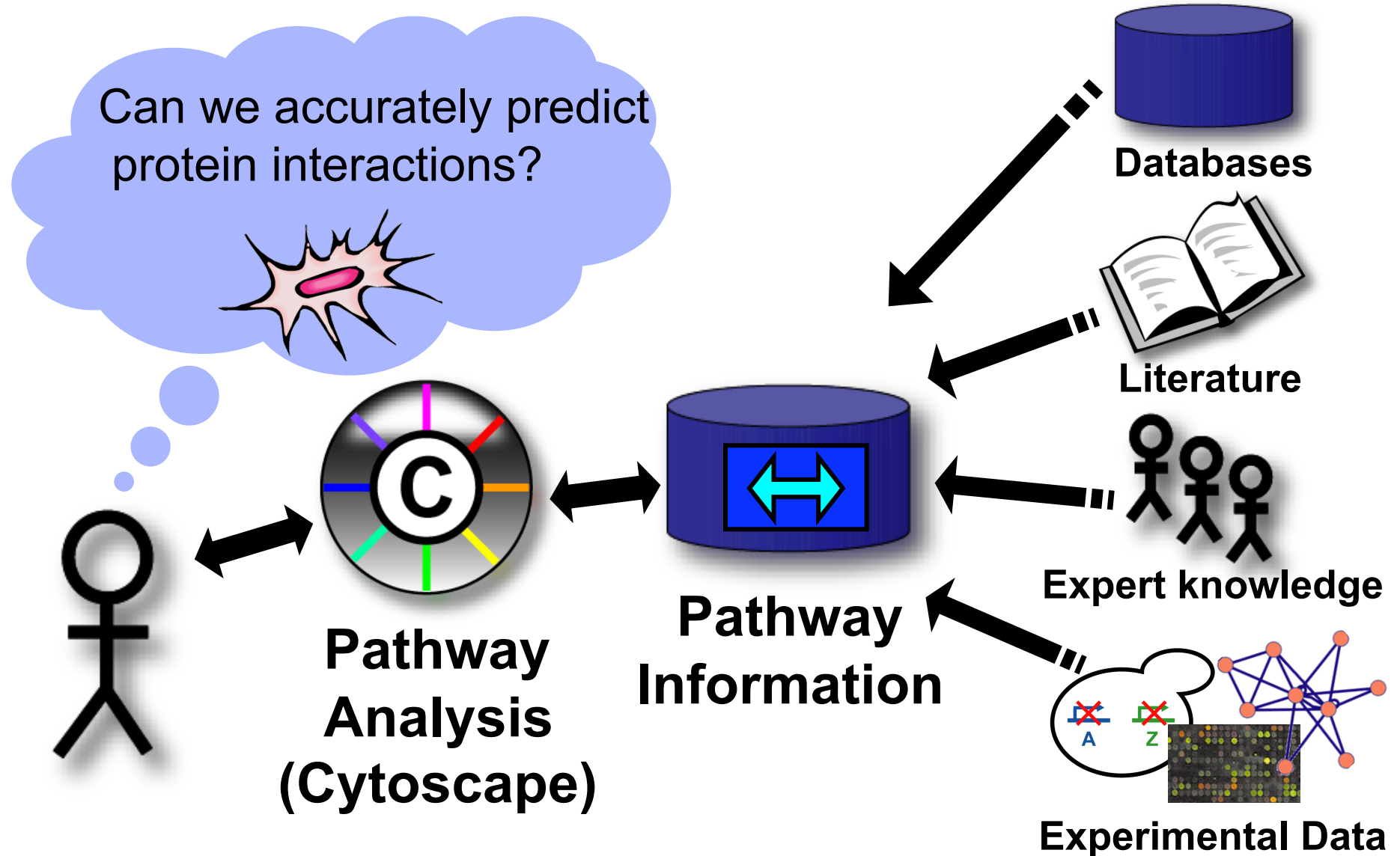
cPath web service API

- 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

Using Pathway Information



<http://cytoscape.org>

Network
visualization
and analysis

Pathway comparison

Literature mining

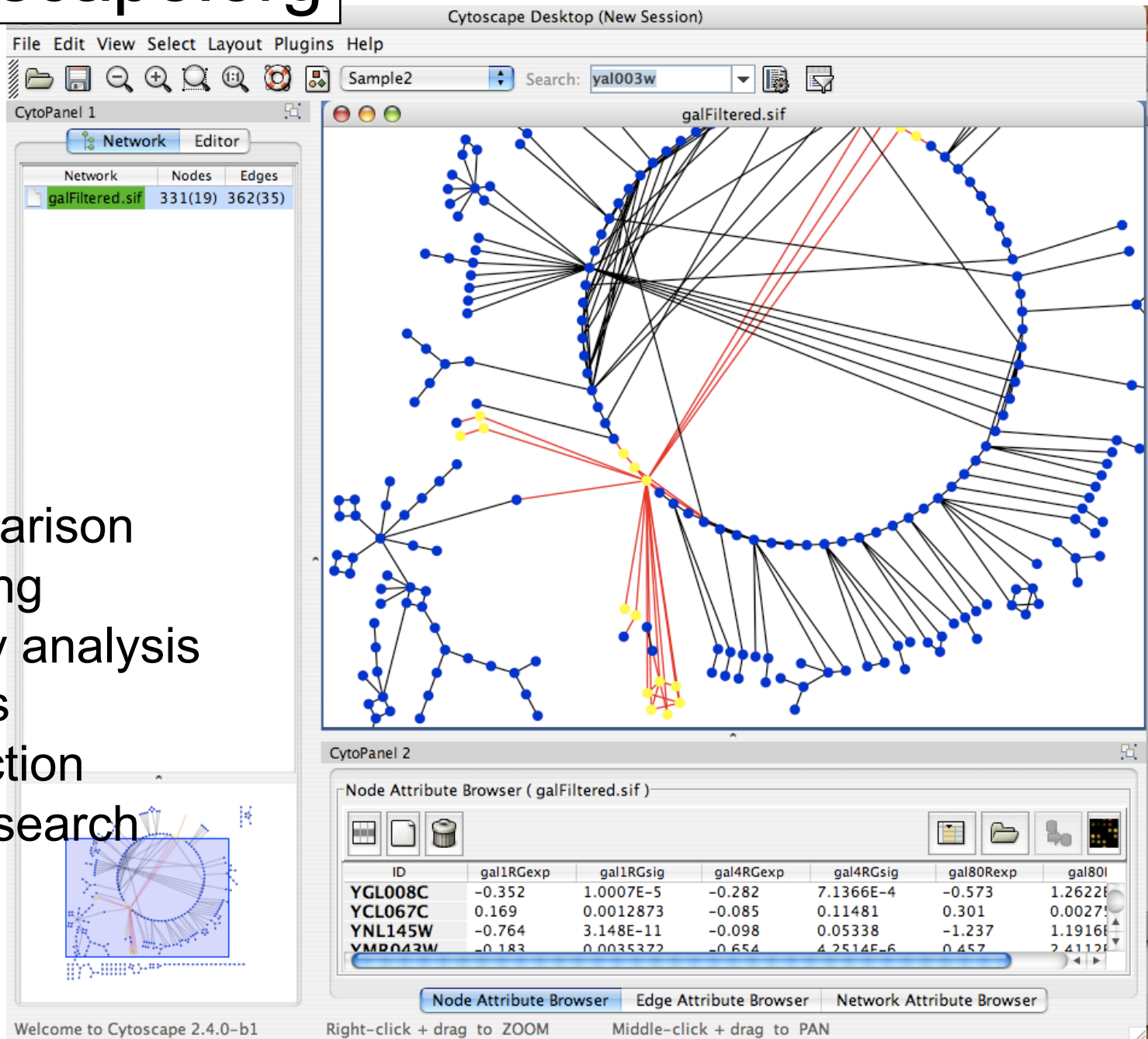
Gene Ontology analysis

Active modules

Complex detection

Network motif search

UCSD, ISB,
Agilent,
MSKCC,
Pasteur



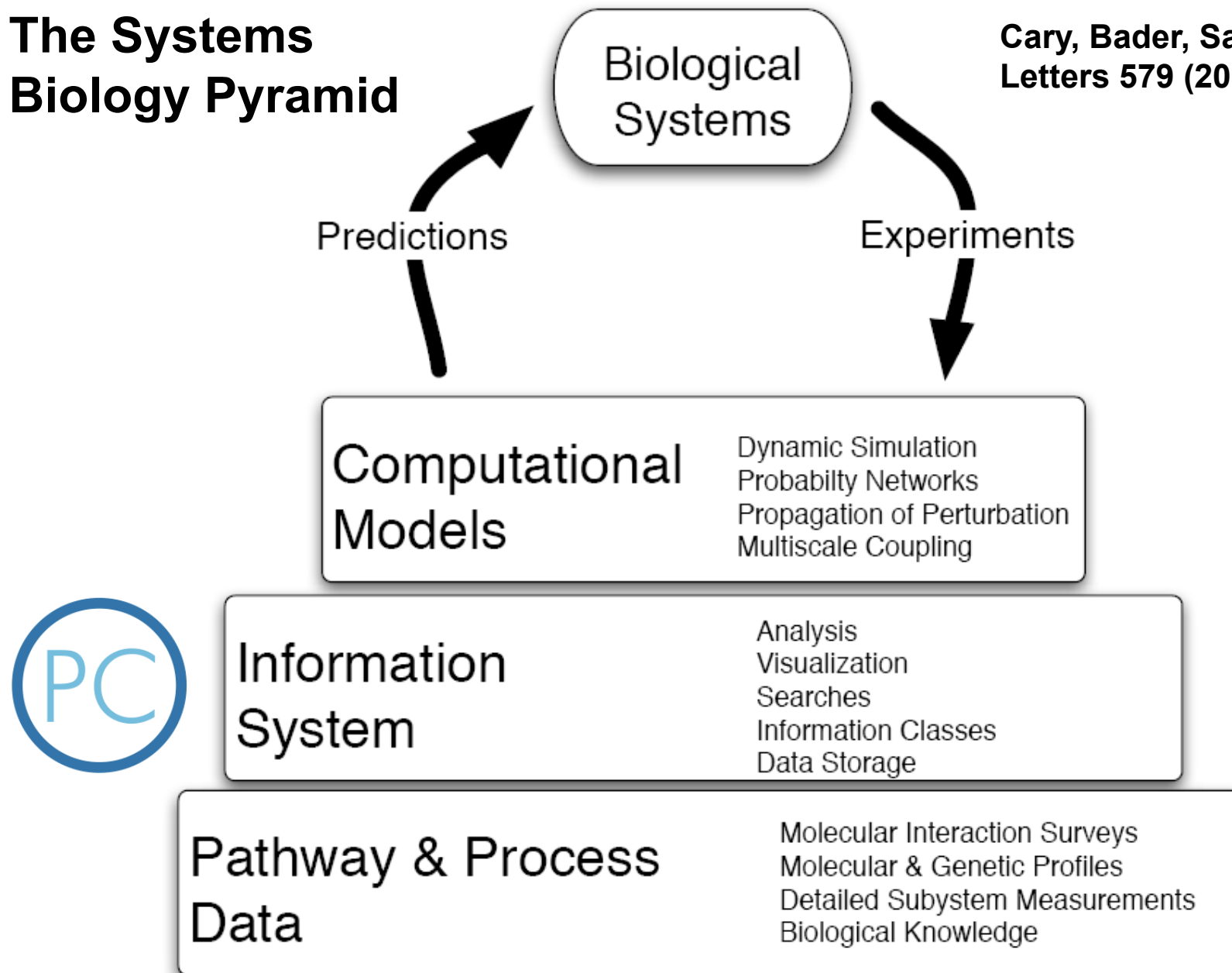
Welcome to Cytoscape 2.4.0-b1

Right-click + drag to ZOOM

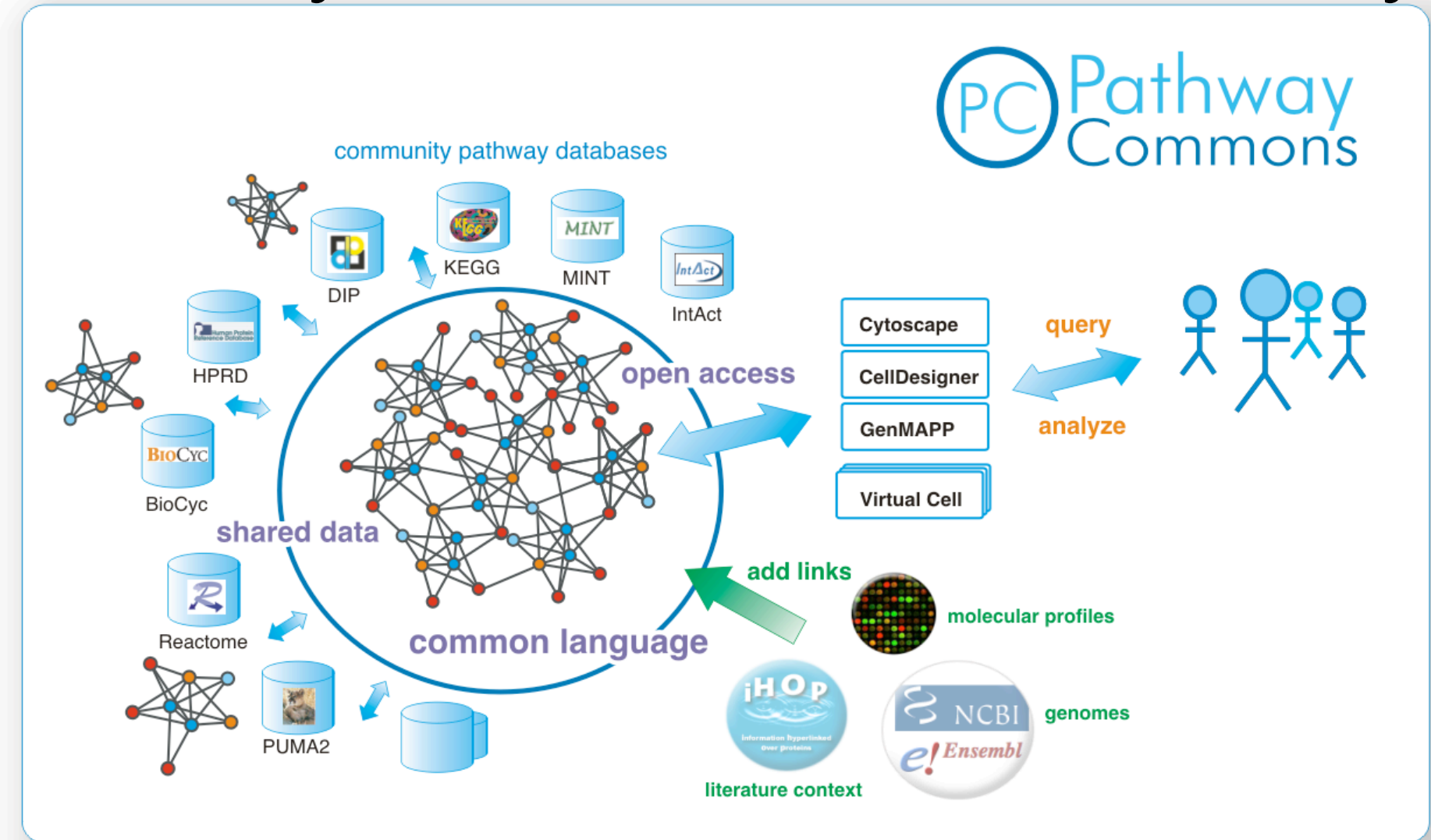
Middle-click + drag to PAN

The Systems Biology Pyramid

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



Pathway Commons: A Public Library

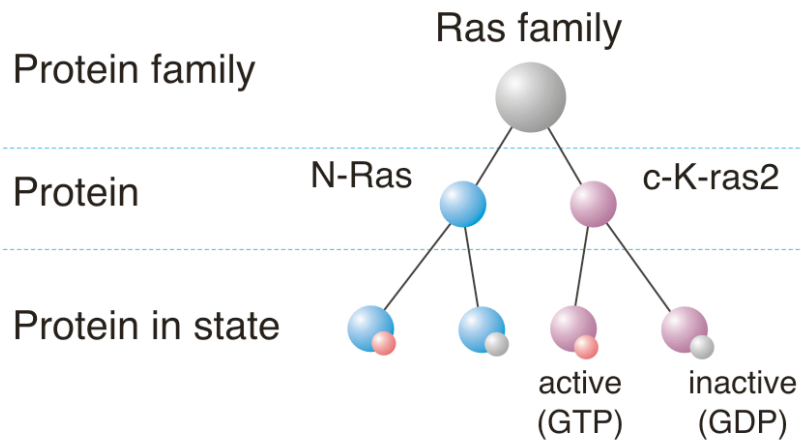


- Books: Pathways
- Lingua Franca: BioPAX OWL
- Index: cPath pathway database software
- Translators: translators to BioPAX

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

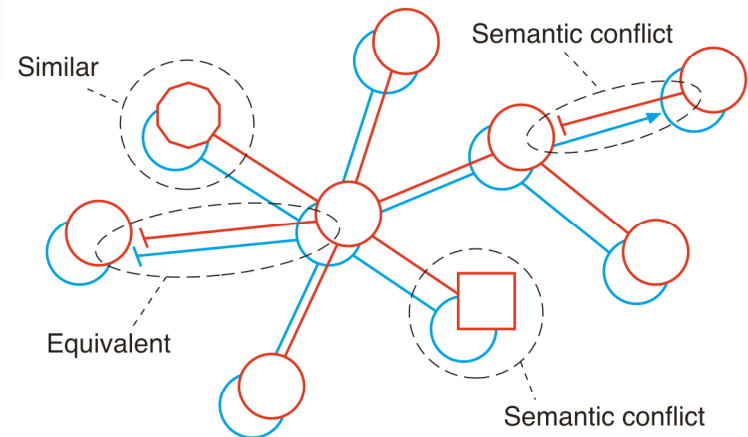
Towards an Integrated Cell Map

- Semantic pathway integration is very hard



Physical entities

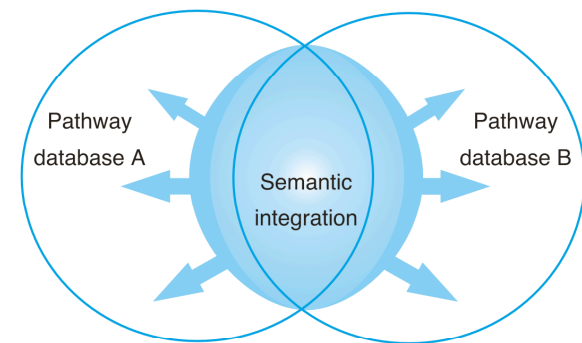
Relationships



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



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](#).

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). ... Under 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

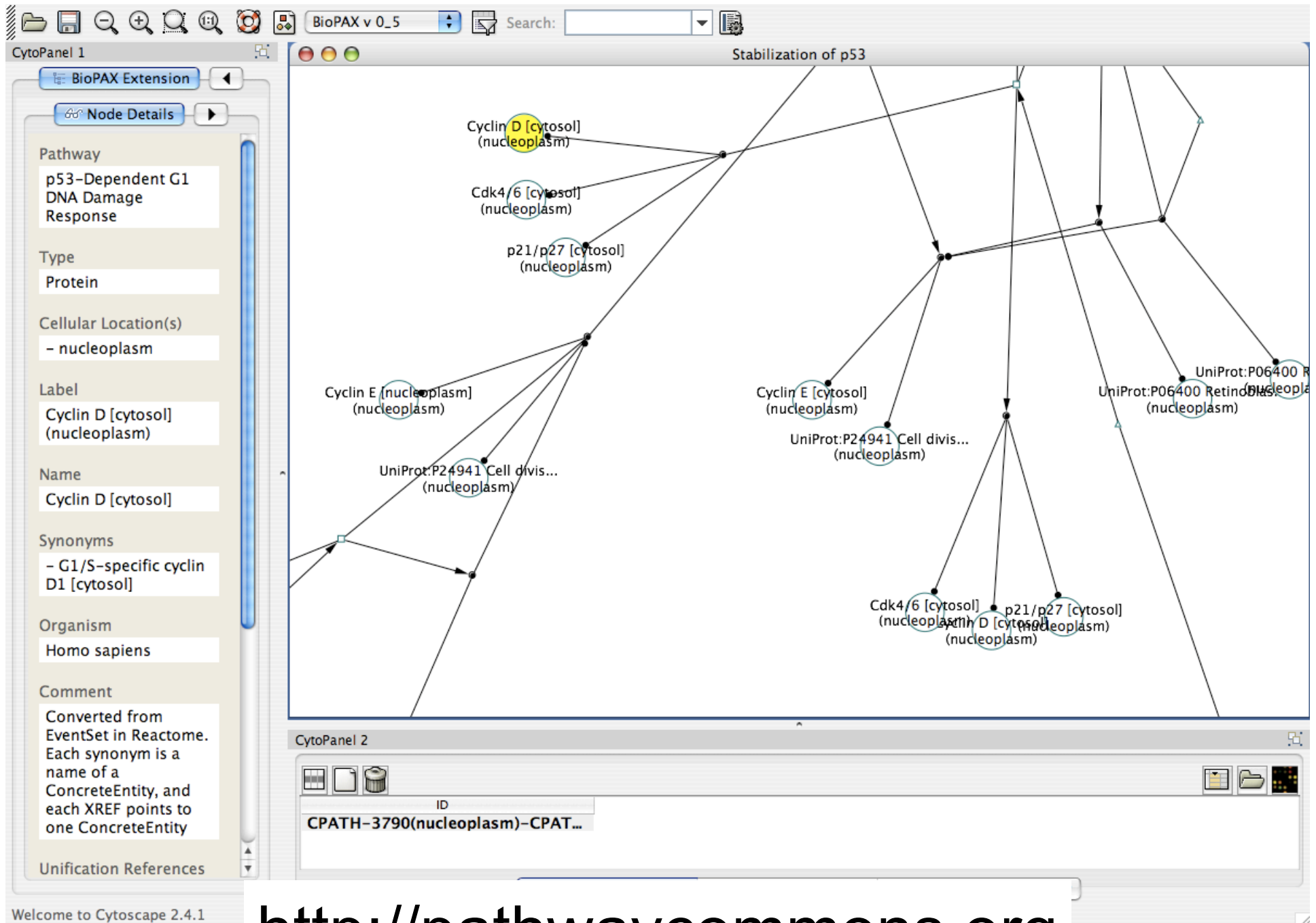
- The G1 arrest induced by DNA damage has been ascribed to the transcription factor and tumor suppressor protein **p53**.

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>



<http://pathwaycommons.org>

Pathway Commons Status

	Cancer Cell Map, Release: 1.0 [19-May-06] Browse
	HPRD [01-Sep-07] Browse
	HumanCyc, Release: 10.5 [18-Sep-06] Browse
	IntAct [14-Dec-07] Browse
	MINT [21-Dec-07] Browse
	NCI / Nature Pathway Interaction Database [28-Jan-08] Browse
	Reactome, Release: 24 [12-Mar-08] Browse

- More Databases
 - iHOP annotation
- Web service API under development
- Neighborhood visualization
- Cytoscape integration

Pathway Commons Quick Stats:

Number of Pathways:	1,111
Number of Interactions:	258,282
Number of Physical Entities:	50,183
Number of Organisms:	445

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

Acknowledgements

Pathway Commons

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

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)

Bader Lab G2N

Chris Tan
David Gfeller
Lianet Lopez
Moyez Dharsee
Shirley Hui

MP

Anastasija
Baryshnikova
Iain Wallace
Laetitia Morrison
ACM
Daniele Merico
Ruth Isserlin
Vuk Pavlovic

Annette Adler (Agilent)

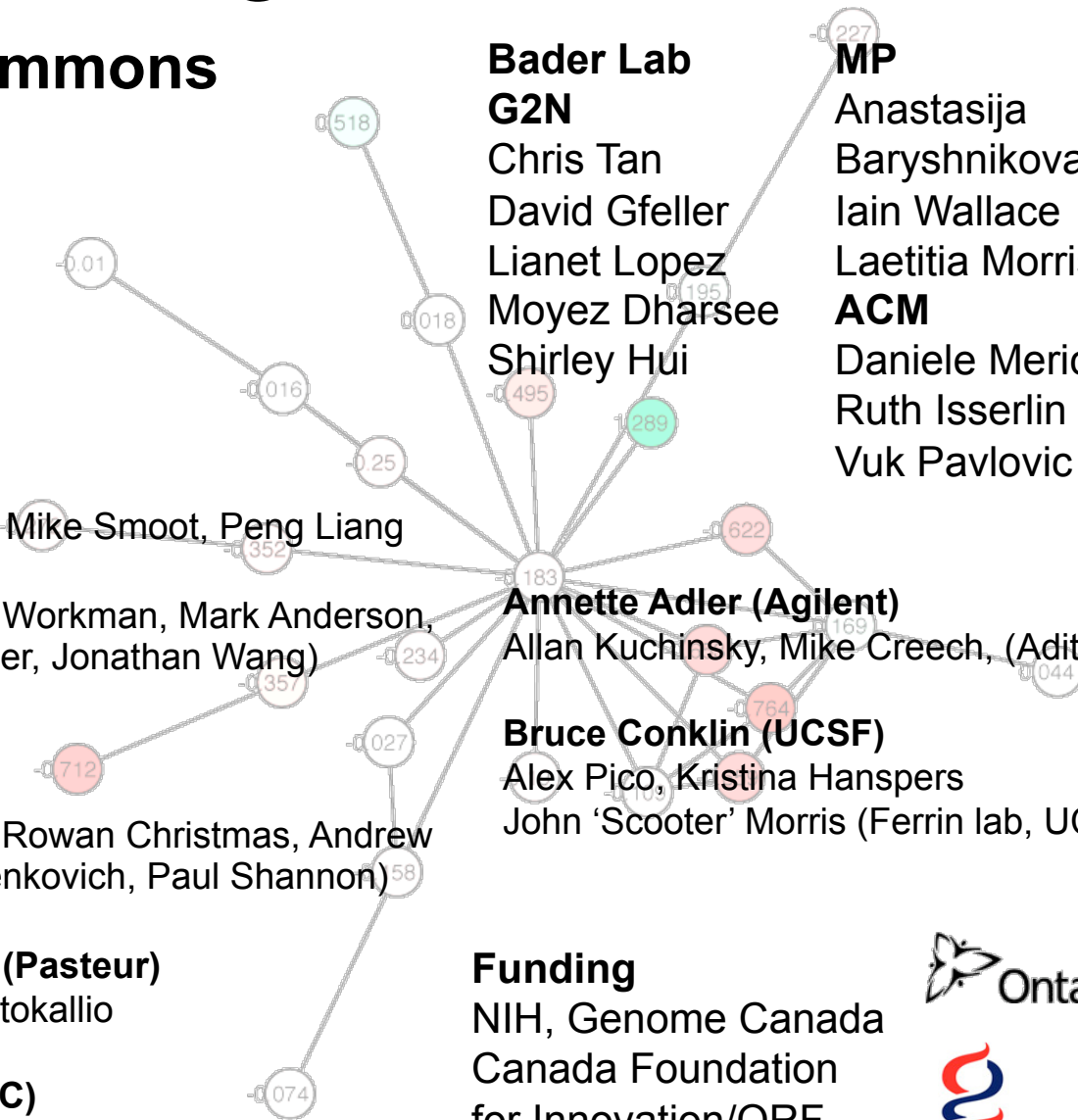
Allan Kuchinsky, Mike Creech, (Aditya Vailaya),

Bruce Conklin (UCSF)

Alex Pico, Kristina Hanspers
John 'Scooter' Morris (Ferrin lab, UCSF)

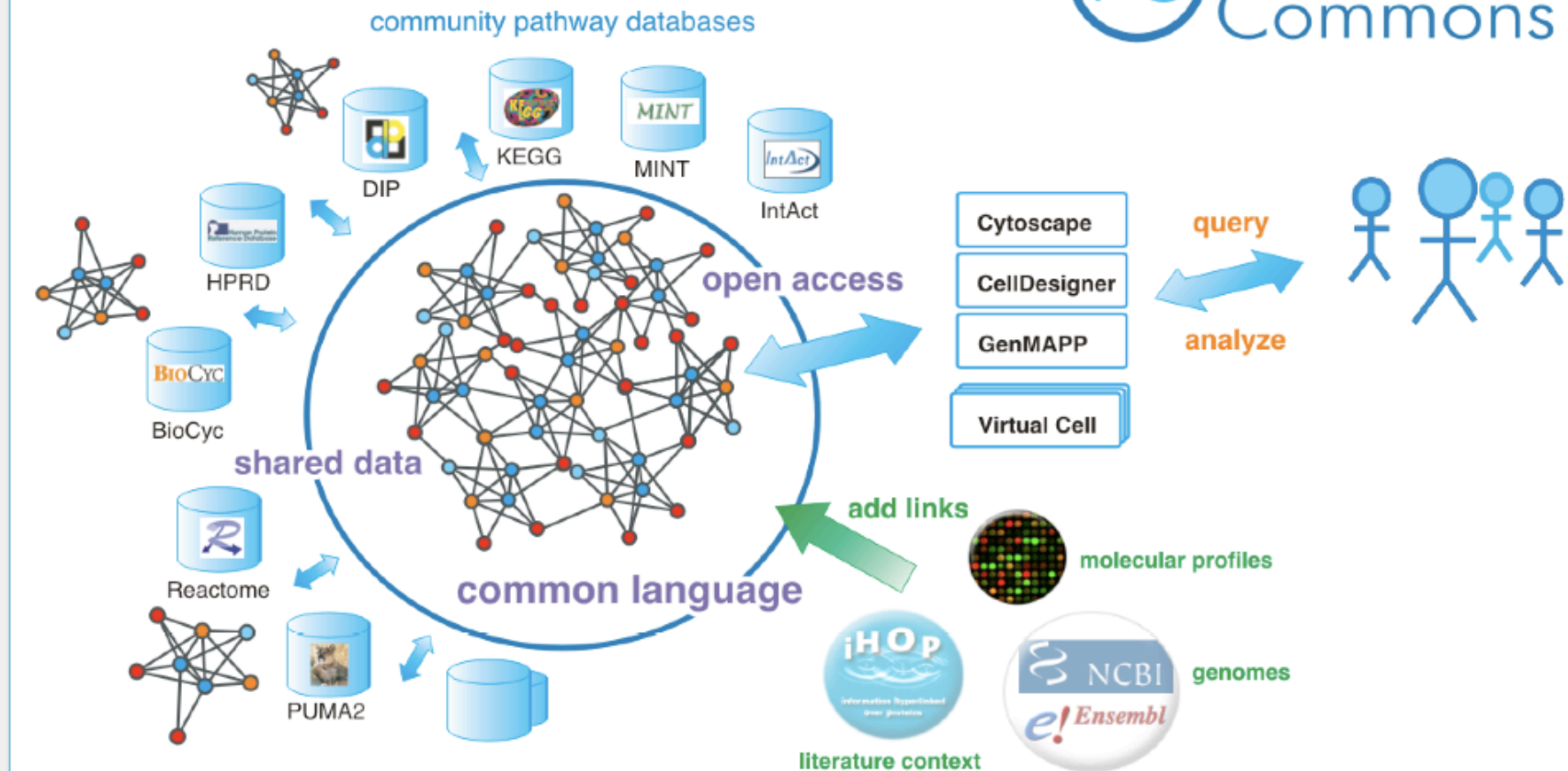
Funding

NIH, Genome Canada
Canada Foundation
for Innovation/ORF



Aim: Convenient Access to Pathway Information

<http://www.pathwaycommons.org>



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