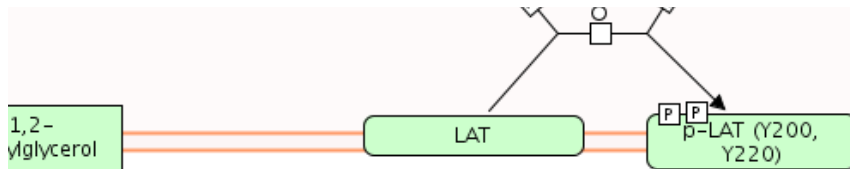
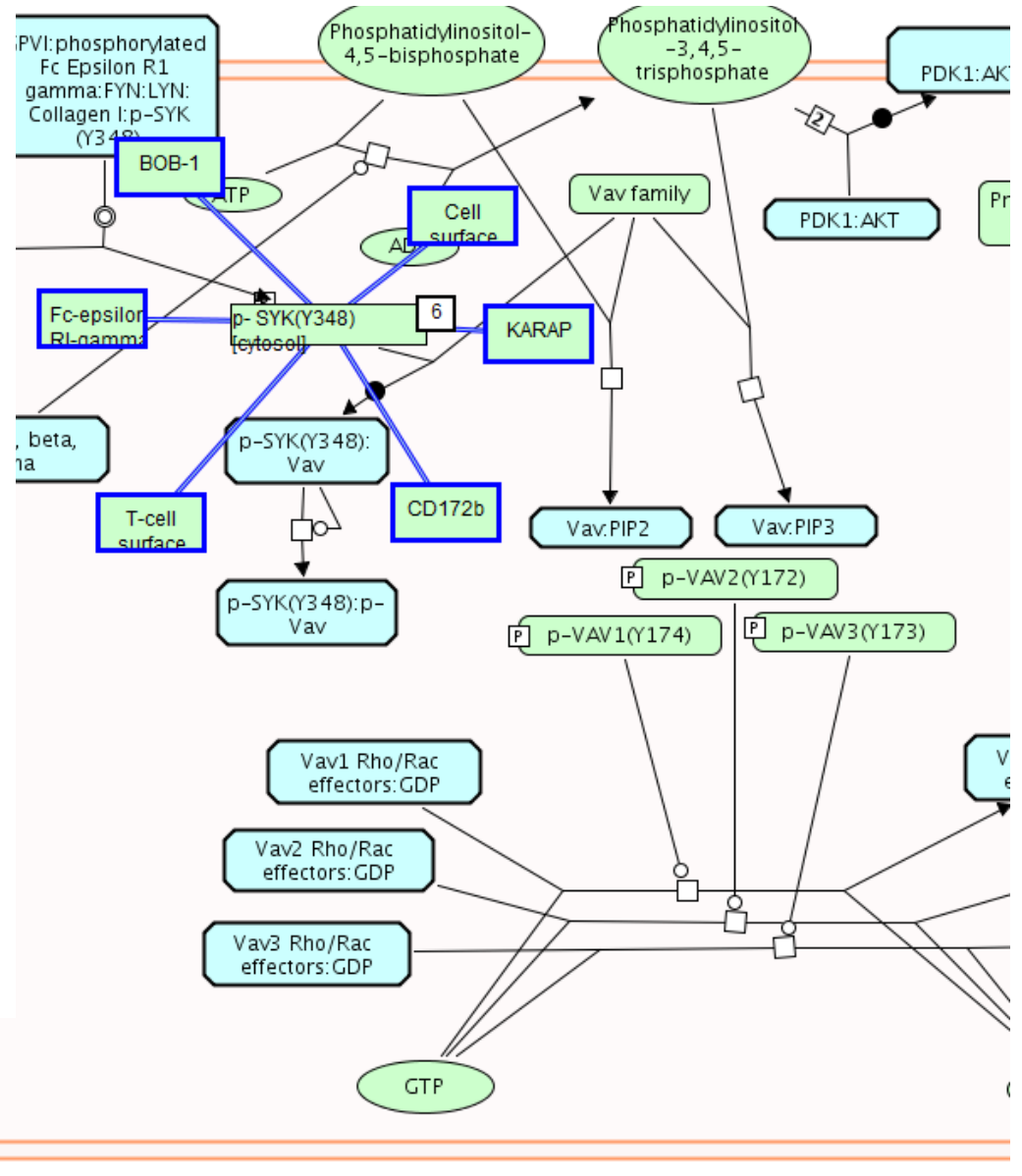




A Database of human biological pathways

Henning Hermjakob
 hhe@ebi.ac.uk

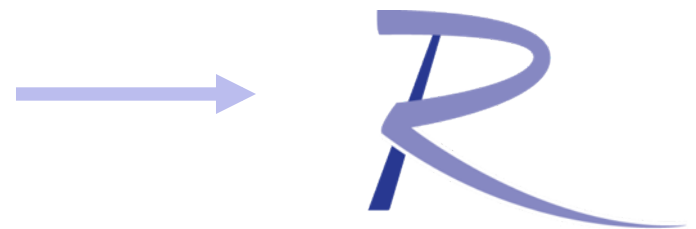


Rationale – Journal information

Nature 407(6805):770-6. The Biochemistry of Apoptosis.

“**Caspase-8** is the key **initiator** caspase in the **death-receptor pathway**. Upon **ligand binding**, death receptors such as **CD95** (Apo-1/Fas) **aggregate** and form **membrane-bound** signalling **complexes** (Box 3). These complexes then recruit, through **adapter proteins**, several molecules of **procaspase-8**, resulting in a high local concentration of zymogen. The induced proximity model posits that under these crowded conditions, the low intrinsic protease activity of procaspase-8 (ref. 20) is sufficient to allow the various proenzyme molecules to mutually cleave and activate each other (Box 2). A similar mechanism of action has been proposed to mediate the activation of several other caspases, including **caspase-2** and the nematode caspase **CED-3** (ref. 21).”

How can I access the pathway described here and reuse it?

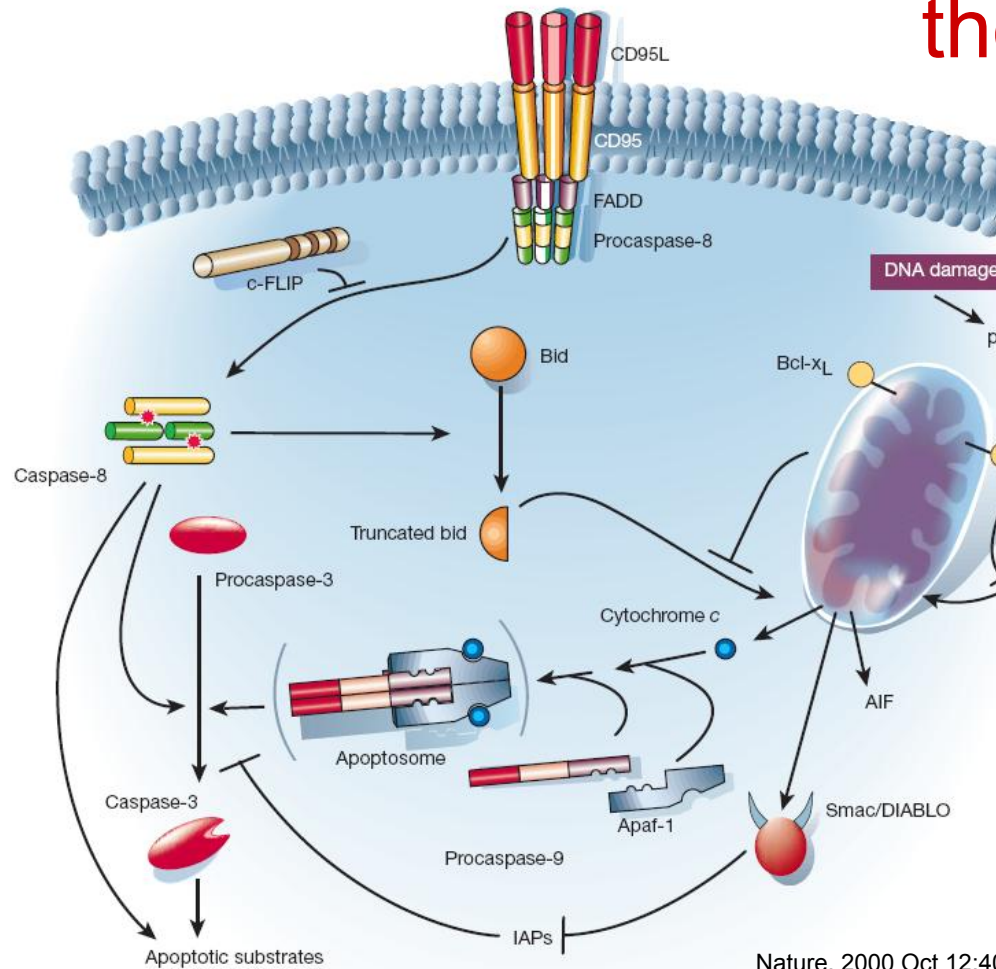


Rationale - Figures

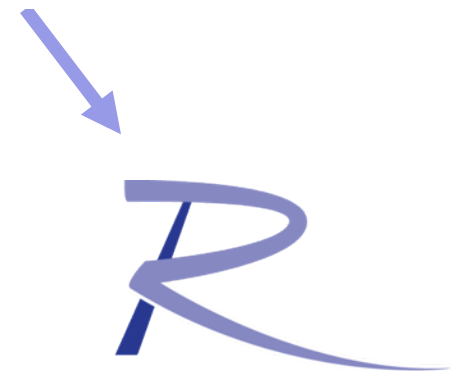
A picture paints a thousand words...

but....

- Just pixels
- Omits key details
- Assumes
- Fact or Hypothesis?

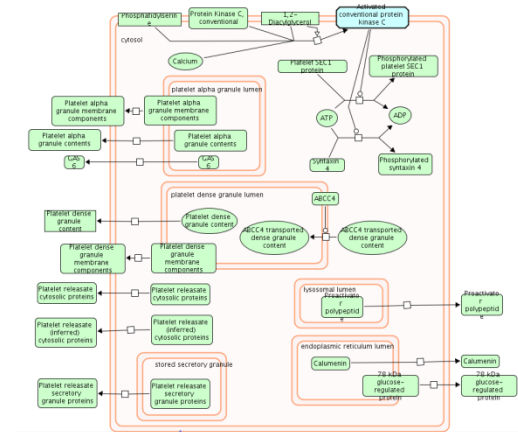


Nature. 2000 Oct 12;407(6805):770-6.
The biochemistry of apoptosis.

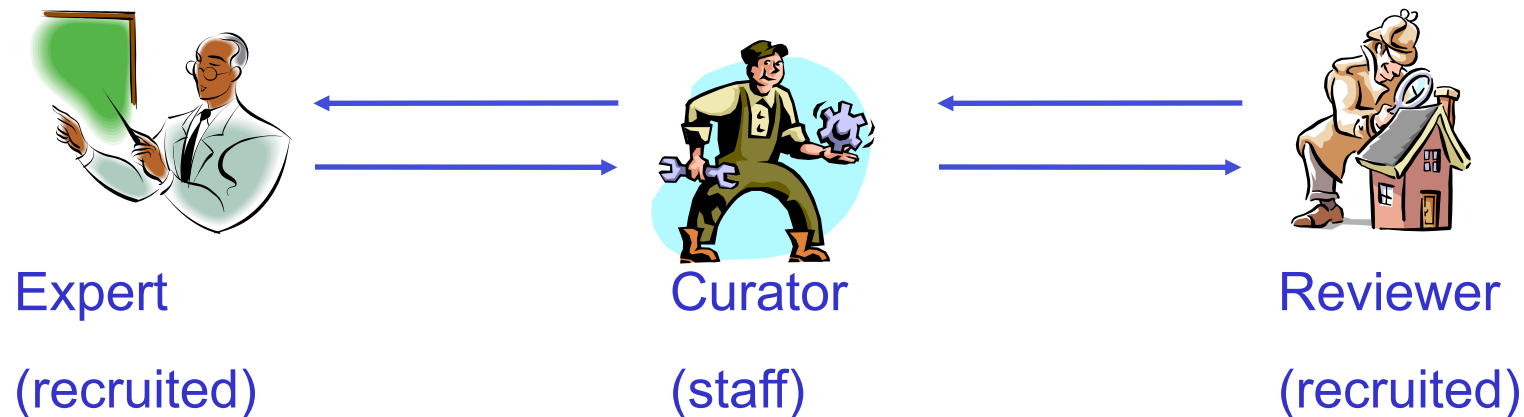


Reactome is...

Free, open source, open data,
curated database of pathways and
reactions in human biology

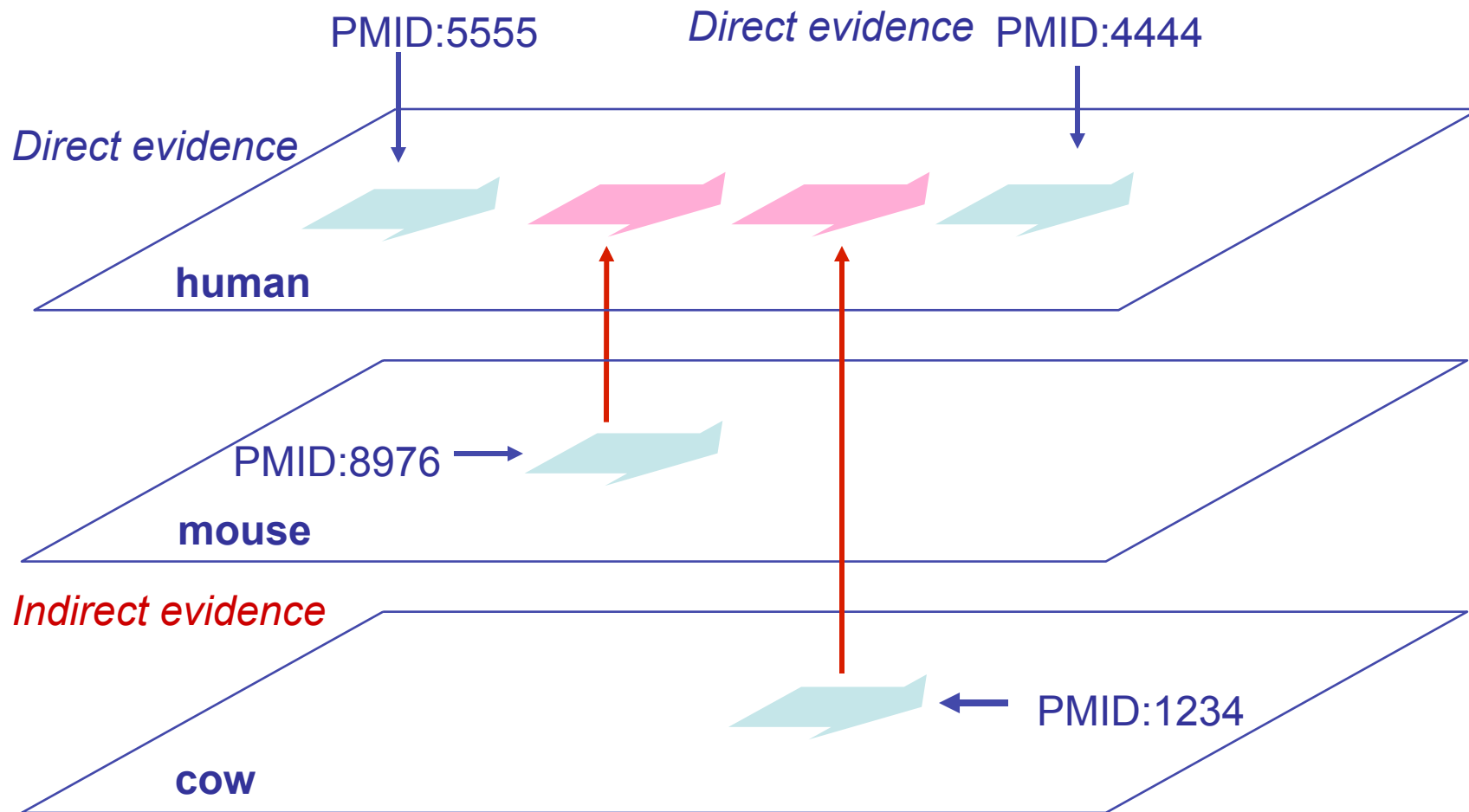


Reactome Curation



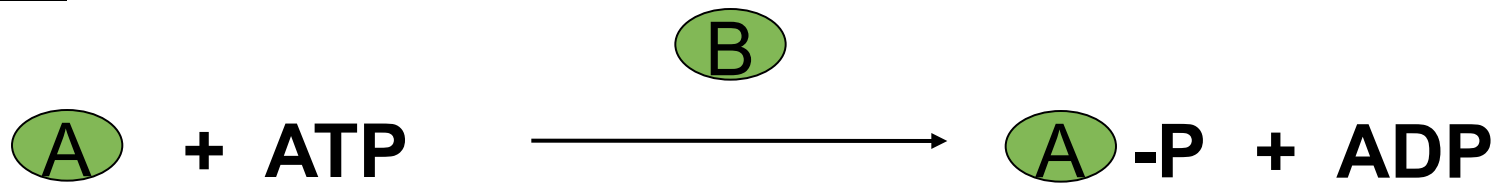
- Expert and curator create outline of new pathway
- Expert fills in details
- Curator enters information into database
- Reviewer (another expert) checks biological correctness

Using model organism data to build pathways – Inferred pathway events

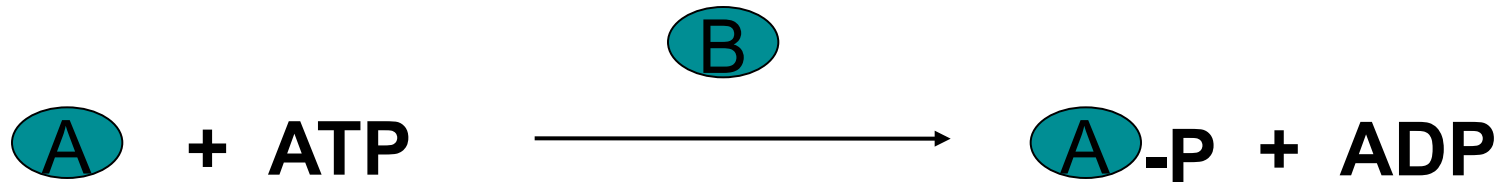


Data Expansion – Projecting to Other Species

Human



Mouse



Drosophila



Reaction not
inferred

No orthologue - Protein not inferred

Coverage – Content, TOC

Topic
Apoptosis [Homo sapiens] <ul style="list-style-type: none"> - Extrinsic Pathway for Apoptosis (DOI) - Intrinsic Pathway for Apoptosis (DOI) - Apoptotic execution phase - Regulation of Apoptosis
Axon guidance [Homo sapiens] (DOI) <ul style="list-style-type: none"> - Semaphorin interactions - NCAM signaling for neurite out-growth - Netrin-1 signaling - Signaling by Robo receptor - L1CAM interactions
Biological oxidations [Homo sapiens] <ul style="list-style-type: none"> - Phase I - Functionalization of compounds - Phase II conjugation (DOI)
Botulinum neurotoxicity [Homo sapiens, Clostridium botulinum] <ul style="list-style-type: none"> - Translocation of BoNT Light chain (DOI) - Proteolytic cleavage of SNARE complex proteins (DOI)
Cell Cycle Checkpoints [Homo sapiens] <ul style="list-style-type: none"> - G1/S DNA Damage Checkpoints (DOI) - G2/M Checkpoints (DOI) - Mitotic Spindle Checkpoint (DOI)
Cell Cycle, Mitotic [Homo sapiens] <ul style="list-style-type: none"> - Mitotic G1-G1/S phases - S Phase (DOI) - Regulation of DNA replication (DOI) - Mitotic G2-G2/M phases - Mitotic M-M/G1 phases - Regulation of mitotic cell cycle
Cell junction organization [Homo sapiens] (DOI) <ul style="list-style-type: none"> - Cell-cell junction organization (DOI) - Cell-extracellular matrix interactions - Type I hemidesmosome assembly
Chromosome Maintenance [Homo sapiens] <ul style="list-style-type: none"> - Nucleosome assembly - Telomere Maintenance

And many more...

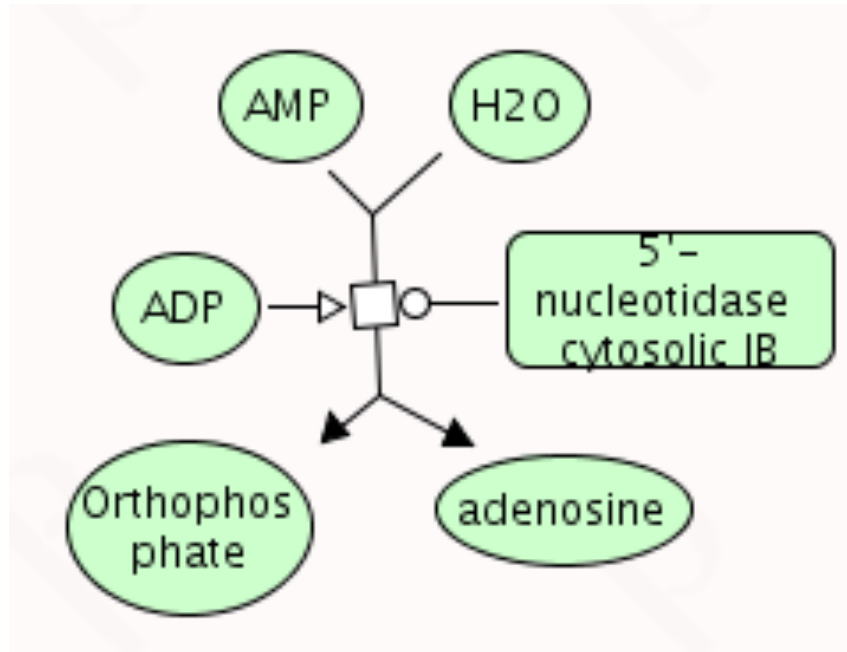
The image displays a grid of small thumbnail images, each representing a different biological pathway or process. Each thumbnail includes a title and a list of associated genes and proteins. For example, one thumbnail titled 'Regulation of energy metabolism' lists genes like 'UCP1', 'UCP2', and 'UCP3'. Another titled 'Cell cycle checkpoints' lists 'p53', 'p21', and 'p27'. The thumbnails are arranged in a grid, with some overlapping, and they cover a wide range of biological topics including metabolism, cell cycle, apoptosis, and signaling.

Planned Coverage – Editorial Calendar

Next Release - Ver 38 : September 2011 [Reactome](#) 

Curator	Pathway Topic	Author	Reviewer
Bruce May	Mitochondrial import	B May	TBA
Bruce May	Carbonic Anhydrases	B May	TBA
Bruce May	Nodal signaling	B May	TBA
Bruce May	Regulatory RNA Pathways (siRNA, post-transcriptional silencing)	B May	TBA
Karen Rothfels	Signaling by FGFR- update	K Rothfels	TBA
Steve Jupe	Interleukin-7 signaling	K Ray	TBA
Steve Jupe	Chromatin modifying enzymes	TBA	TBA
Steve Jupe	Defensins	S Jupe	TBA
Marija Milacic	Signaling by ERBB4	M Milacic	TBA
Marija Milacic	Signaling by constitutively active EGFR	M Milacic	TBA
Marija Milacic	Signaling by ERBB2	M Milacic	TBA
Marija Milacic	Cyclin D associated events in G1 - update - p107 and p130 events	TBA	D MacPherson
Bijay Jassal	O-linked glycosylation of mucins	B Jassal	TBA
Bijay Jassal	Nitric oxide metabolism (revision)	B Jassal	TBA
Bijay Jassal	ABC-family proteins mediated transport (revision)	B Jassal	TBA
Bijay Jassal	Latent infection of Homo sapiens with Mycobacterium tuberculosis	R Stephan	TBA
Bijay Jassal	Post-translational modifications	B Jassal	TBA
Marc Gillespie	Reproduction	M Gillespie	TBA
Steve Jupe	Regulation of complement cascade	S Jupe	TBA
Steve Jupe	Growth hormone receptor signaling	S Jupe	A Herington
Phani Garapati	MHC Class I cross presentation	P Garapati	TBA
Phani Garapati	ISG15 antiviral mechanism	P Garapati	TBA
Phani Garapati	MHC Class II mediated antigen processing & presentation	P Garapati	TBA
Phani Garapati	SCF:c-kit signaling	P Garapati	L Ronnstrand, NW Lukacs
Veronica Shamovsky	Toll Like Receptors 5 and 10 update	F Luo	TBA
Marc Gillespie	Fertilization	M Gillespie	TBA
Bruce May	Mitochondrial Iron-Sulfur Cluster Biogenesis	B May	TBA

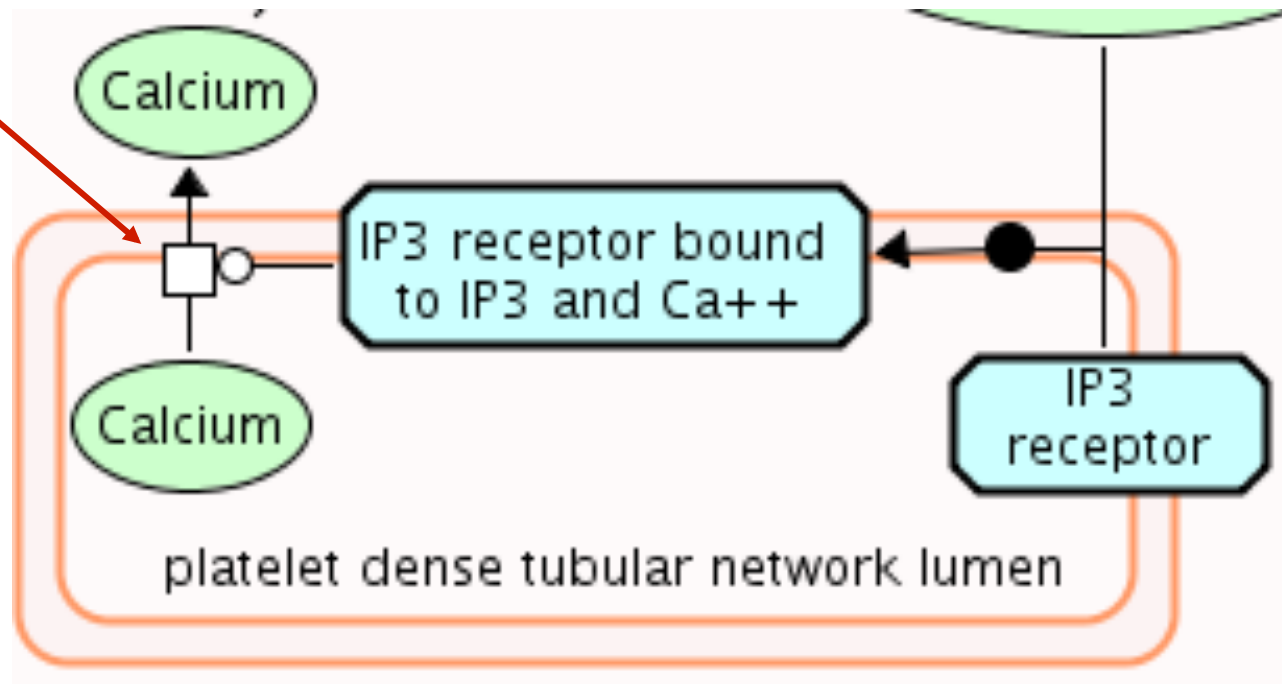
Reaction Example 1: Enzymatic



Reaction Example 2: Transport

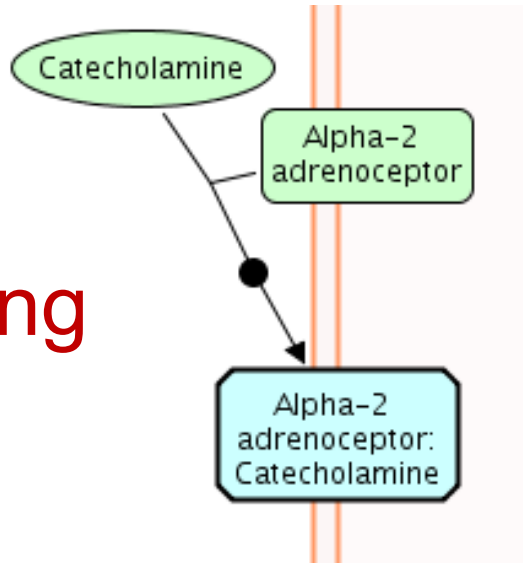
Transport of Ca^{++} from platelet dense tubular system to cytoplasm

REACT_945.4

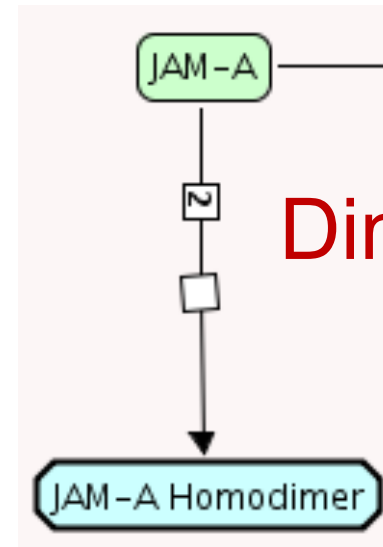


Other Reaction Types

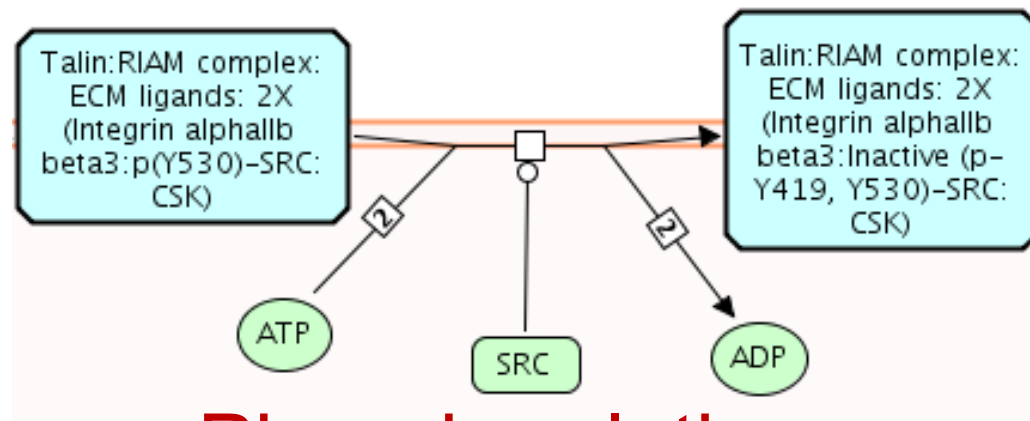
Binding



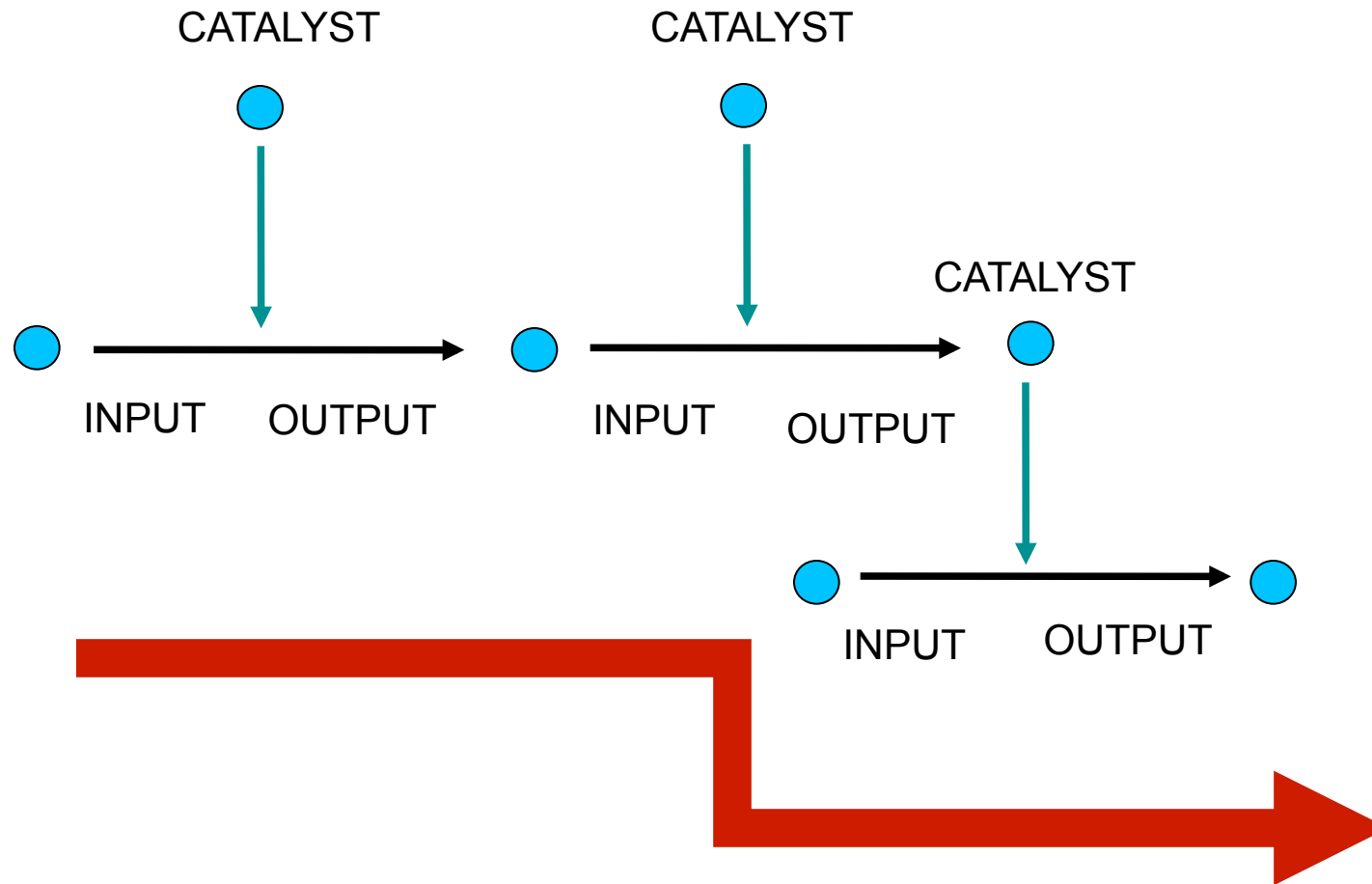
Dimerization



Phosphorylation



Reactions Connect into Pathways



Pathway Analysis

Pathway Analysis

Allows you to analyse a list of protein, gene, expression data or compound identifiers and determine how they are likely to affect pathways. [More...](#)

Paste or upload your data:

[Example](#)

O00139
O00186
O00187
O00204
O00217
O00231
O00232
O00233
O00254
O00267
O00268
O00273
O00303

Select your desired analysis tool

Inhouse services:

- ID mapping and pathway assignment.** Takes your list of IDs and finds the corresponding pathways from Reactome, plus the corresponding UniProt IDs.
- Overrepresentation analysis.** Finds the Reactome pathways in which IDs in your list are strongly enriched - can help to understand the biological context of your data.

15 August 2011

www.reactome.org

Pathway Analysis – Overrepresentation

Statistically over-represented events in hierarchy

Each Event is coloured according to the **un-adjusted, i.e. not corrected for multiple testing**, probability (from **hypergeometric test**) of seeing given number of more genes in this Event by chance. Please note that only those "child" events are shown which have a p-value lower than the "parent" event. The top-level (root) Events are ordered according to the lowest p-value of their components.

Colour key for probabilities:

1e+00 3e-01 1e-01 3e-02 1e-02 3e-03 1e-03 3e-04 1e-04 3e-05 1e-05 3e-06 1e-06 3e-07 1e-07 3e-08 1e-08 3e-09 1e-09 3e-10 >

open all | close all

- Cell Cycle, Mitotic 1.9e-20, 137/294
 - Matching identifiers
- Gene Expression 2.5e-20, 174/411
- DNA Replication 8.4e-16, 90/181
- Signaling by NGF 7.9e-14, 96/210
- Apoptosis 3.8e-15, 71/131
- DNA Repair 2.7e-13, 58/104
- Transcription 9.5e-15, 74/141
- Cell Cycle Checkpoints 1.1e-14, 62/109
- Metabolism of carbohydrates 1.9e-13, 62/114
- Integrin cell surface interactions 3.0e-12, 48/82
- mRNA Processing 3.0e-11, 66/136
- HIV Infection 1.7e-01, 102/408
- Hemostasis 2.3e-10, 101/251
- Chromosome Maintenance 1.3e-04, 25/54
- Signaling by Wnt 3.3e-09, 31/50
- Cdc20:Phospho-APC/C mediated degradation of Cyclin A 5.9e-09, 35/61
- Metabolism of amino acids and derivatives 2.7e-08, 76/187
- Regulation of activated PAK-2p34 by proteasome mediated degradation 2.4e-08, 27/43
- Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins 2.4e-08, 46/94
- Signaling in Immune system 3.7e-04, 104/338
- Signaling by EGFR 2.5e-06, 26/48
 - Matching identifiers
- Transmembrane transport of small molecules 1.0e+00, 29/239
 - Matching identifiers
 - SLC-mediated transmembrane transport 1.0e+00, 29/231
 - Matching identifiers
 - Hexose transport 1.6e-04, 20/40
 - Matching identifiers
 - Glucose transport 6.4e-05, 20/38
 - Pyruvate metabolism and Citric Acid (TCA) cycle 9.3e-05, 19/36
 - Metabolism of proteins 5.5e-01, 53/232

P-val

Matching identifiers
 Q99466 NOTCH4
 Q9UM47 NOTCH3
 Q04721 NOTCH2

Top-level

www.reactome.org

Expression Analysis I



Home About Content Documentation Tools Download Contact Us Outreach

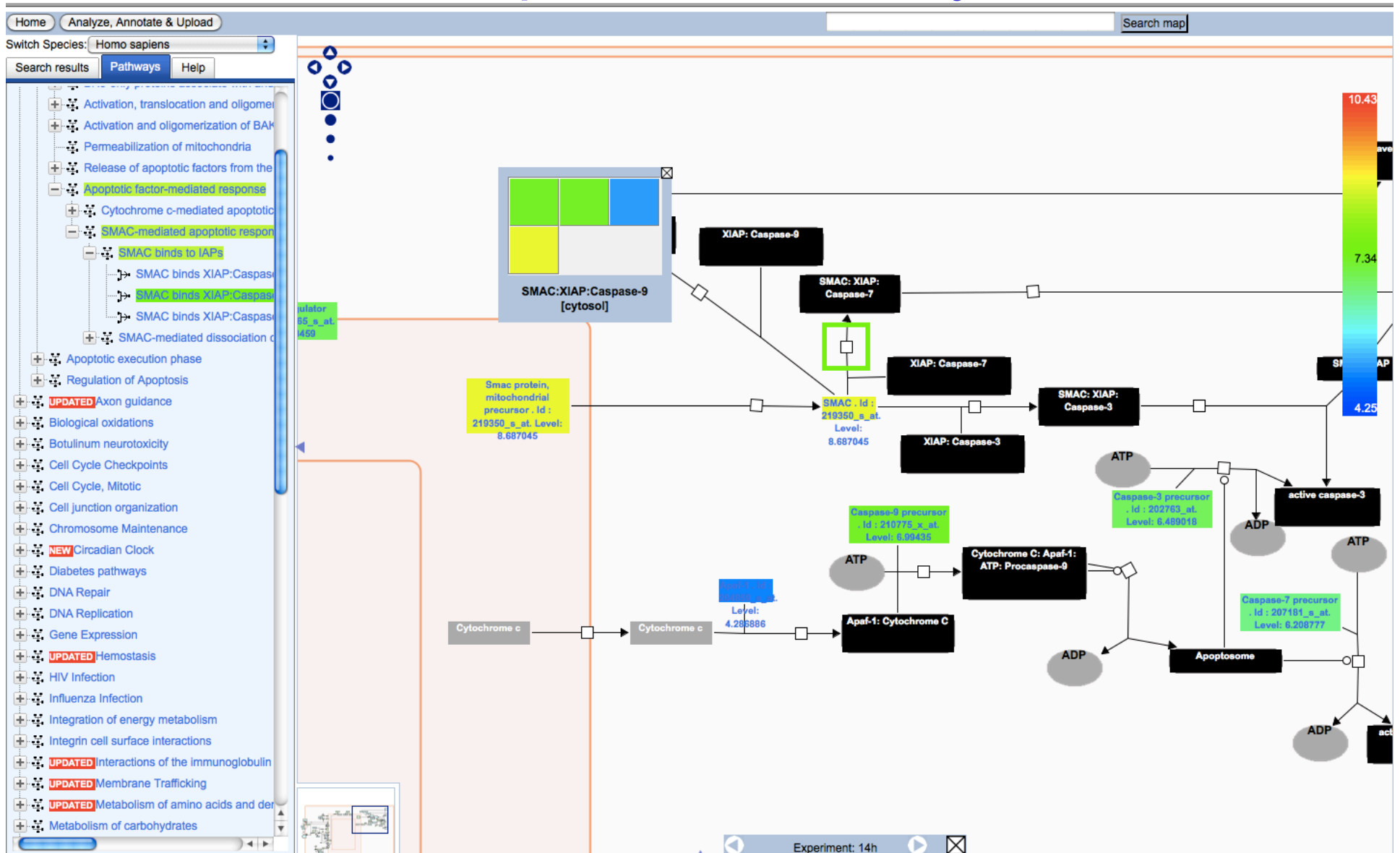
Upload expression data

Takes gene expression data (and also numerical proteomics or metabolomics data) and shows how expression levels affect reactions and pathways in living organisms. May be time-consuming, depending on the number of identifiers you are submitting; less than 5000: a few seconds, 5000 - 10000: a few minutes, 10000 or more: 10 minutes or longer. [More....](#)

Paste or upload your data: [Example](#)

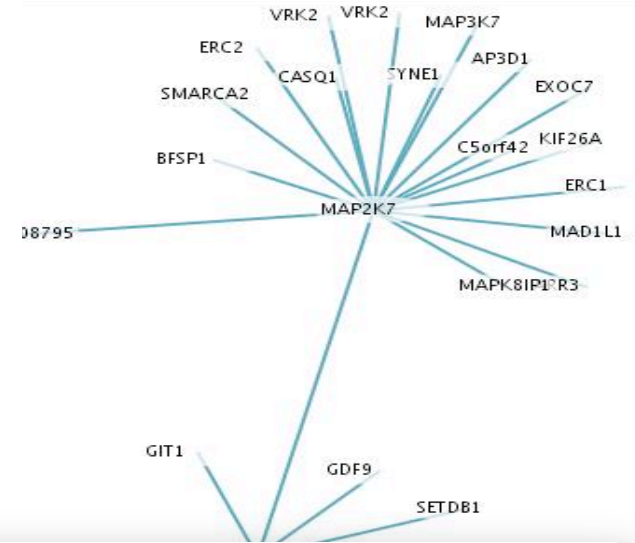
```
#Probeset 10h_control 10h 14h 18h 24h
1053_at 8.040078 7.147358 6.706705 6.794622 7.475157
1729_at 6.869688 6.99104 7.129922 7.112222 7.04721
1861_at 6.437999 6.620092 6.20117 6.407735 5.717815
200000_s_at 9.381569 9.710802 9.874874 9.934639 9.495911
200002_at 12.555275 12.511045 12.564419 12.538642 12.439174
200003_s_at 12.401259 12.054083 12.275169 12.206342 12.015476
200005_at 9.609852 9.099299 9.73072 9.530097 9.194303
200012_x_at 12.486269 12.402275 12.302666 12.256543 12.282444
200014_s_at 10.371458 9.548578 9.978313 9.871472 8.753136
200016_x_at 12.110468 11.913288 11.938524 11.899243 11.458105
200022_at 12.205038 11.927471 12.064725 12.031422 11.932256
200023_s_at 10.377248 9.902753 9.990862 10.248412 9.513486
```

Expression Overlay



IntAct – Molecular Interactions

- Manually curated molecular interaction DB
- 266.855 binary interaction evidences
- Detailed, “deep” curation model



Experiment (100 interactions)

Accession: EBI-2509892
Name: sowa-2009-1
Host organism: [human-293t](#)
Interaction Detection Method: [anti tag coip](#)
Participant Identification Method: [sequence tag](#)
Cross References:

imex	IM-12079	-	imex-primary
pubmed	19615732	-	primary-reference
newt	9606	human	target-species

Annotations:

curation request	Harper Wade: 28th Oct 2009, Department of Pathology, Harvard Medical School, 77 Avenue Louis Pasteur, NRB 940, Boston, MA 02115
confidence-mapping	High-confidence candidate interaction proteins were identified using CompPASS, an automated MS/MS data-processing component, a protein function/annotation component, and an interaction network analysis component. Interactors scored using the D score which incorporates the uniqueness, the abundance of the interactor (TSC), and the reproducibility of the interaction to assign a score to each protein within each IP (the same protein in two different IPs may have distinct D scores since its TSC

POSITIONAL FEATURES

FEATURE TYPE	LABELS	FEATURE ANNOTATIONS	SERVER NAME	EVIDENCE (Category)
signal peptide	UNIPROTKB P0...		uniprot	inferred by curator
mature protein region	Amyloid beta...		uniprot	inferred by curator
sufficient to bind	711-770, 683...		intact	inferred from direct
sufficient to bind	672-713		intact	inferred from direct
active peptide	P3 (42)		uniprot	inferred by curator
extramembrane	Extracellula...		uniprot	inferred by curator
transmembrane	Helical;		uniprot	inferred by curator
alpha helix	UNIPROTKB P0...		uniprot	inferred by curator
beta strand	UNIPROTKB P0...		uniprot	inferred by curator
turn	UNIPROTKB P0...		uniprot	inferred by curator

- Top
- Abstract
- Background
- Results and discussion
- Conclusion
- Materials and methods
- Abbreviations

This article is part of [Informa](#) The Plant Cell, Vol. 22: 997–1005, April 2010, www.plantcell.org © 2010 American Society of Plant Biologists

Research **MINT chalk quality**
Andrew Arnaud Huntley Cesaren
PERSPECTIVE
Special Series on Large-Scale Biology
Mapping Plant Interactomes Using Literature Curated and Predicted Proteins **Research Article**



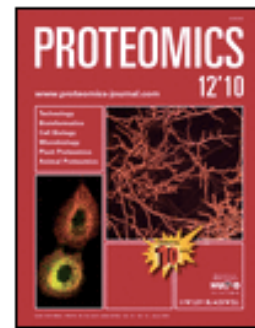
KiYoung Lee,^{a,b,1,2} David T
^a Department of Biomedical Ir
^b Departments of Medicine ar
^c European Molecular Biology

Most cellular processes are proteome. The study of these article informs the plant science they can be useful to research set as an example, we discuss our efforts to add value to the locations.

A domain level interaction network of amyloid precursor protein and Aβ of Alzheimer's disease

Victoria M. Perreau^{1,2,3,7}, Sandra Orchard⁴, Paul A. Adlard³, Shayne A. Bellingham^{3,5,6}, Roberto Cappai⁷, Giuseppe D. Ciccotosto^{3,7}, Tiffany F. Cowie^{1,7}, Peter J. Crouch^{2,3,7}, James A. Duce^{3,7}, Genevieve Evin^{3,7}, Noel G. Faux³, Andrew F. Hill^{3,5}, Ya Hui Hung^{2,3}, Simon A. James^{3,8}, Qiao-Xin Li^{2,3,7}, Su San Mok^{3,7}, Deborah J. Tew^{3,7}, Anthony R. White^{2,7}, Ashley I. Bush³, Henning Hermjakob⁴, Colin L. Masters^{1,2,3}

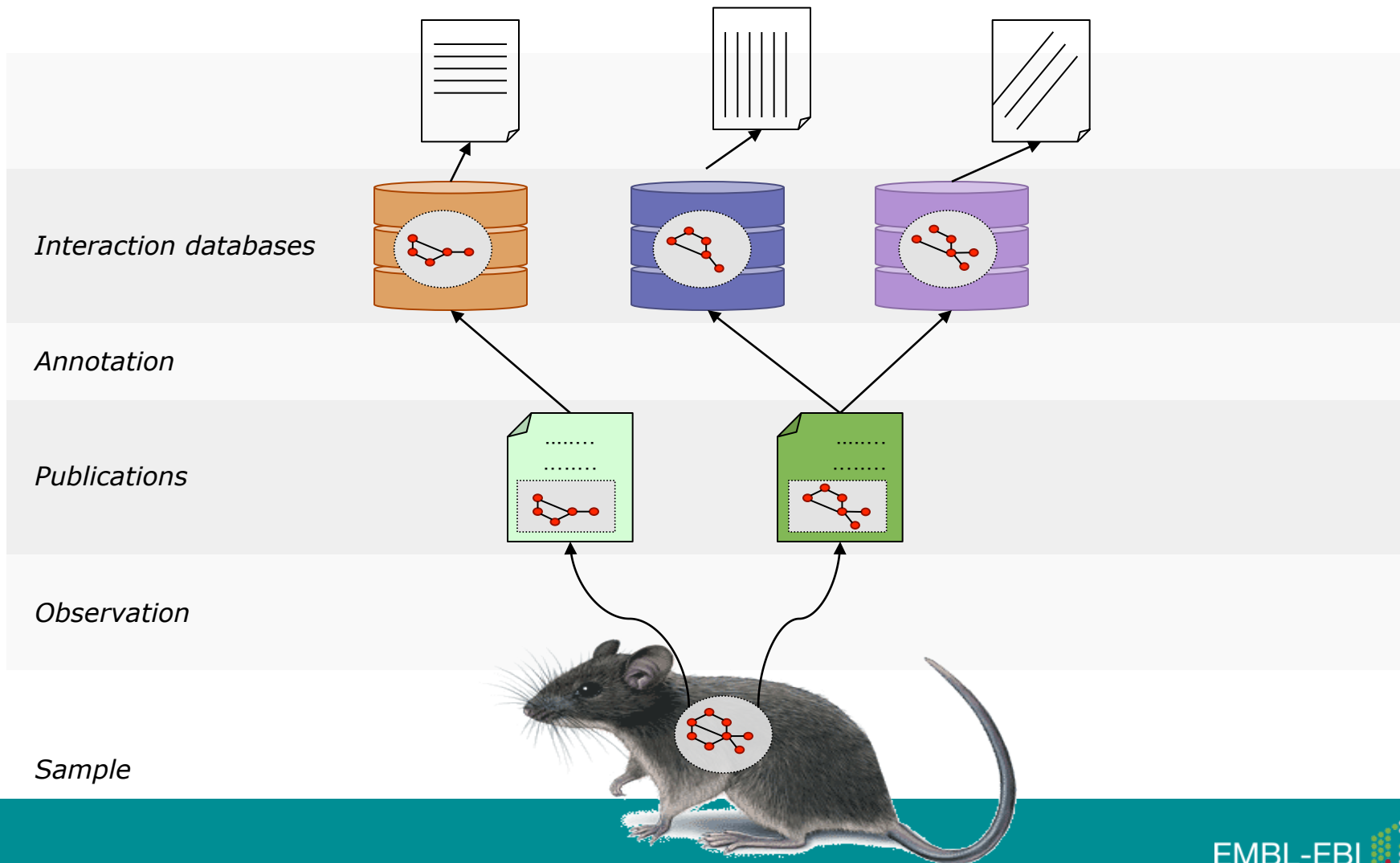
Issue



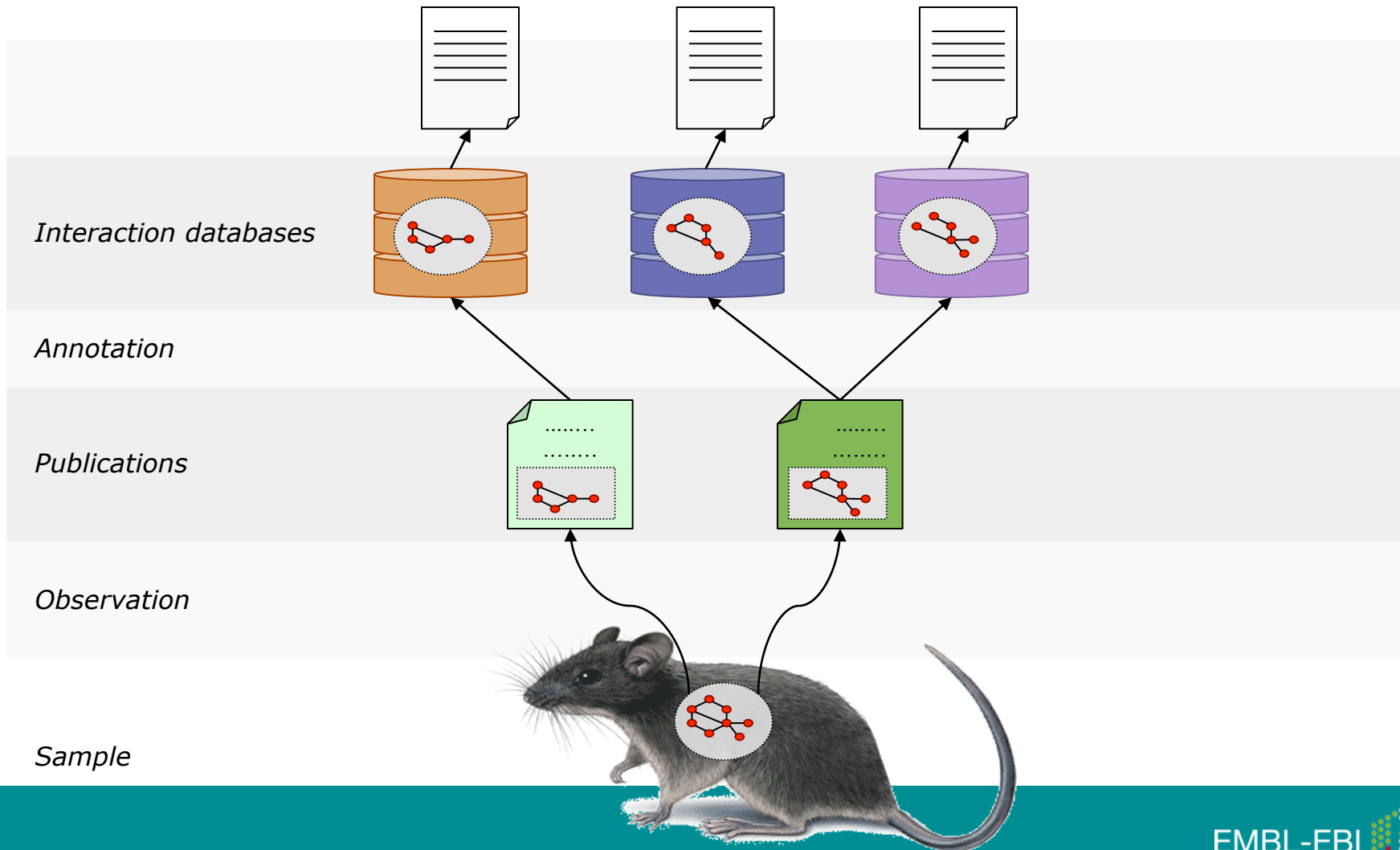
PROTEOMICS
 Volume 10, Issue 12, pages 2377–2395, No. 12 June

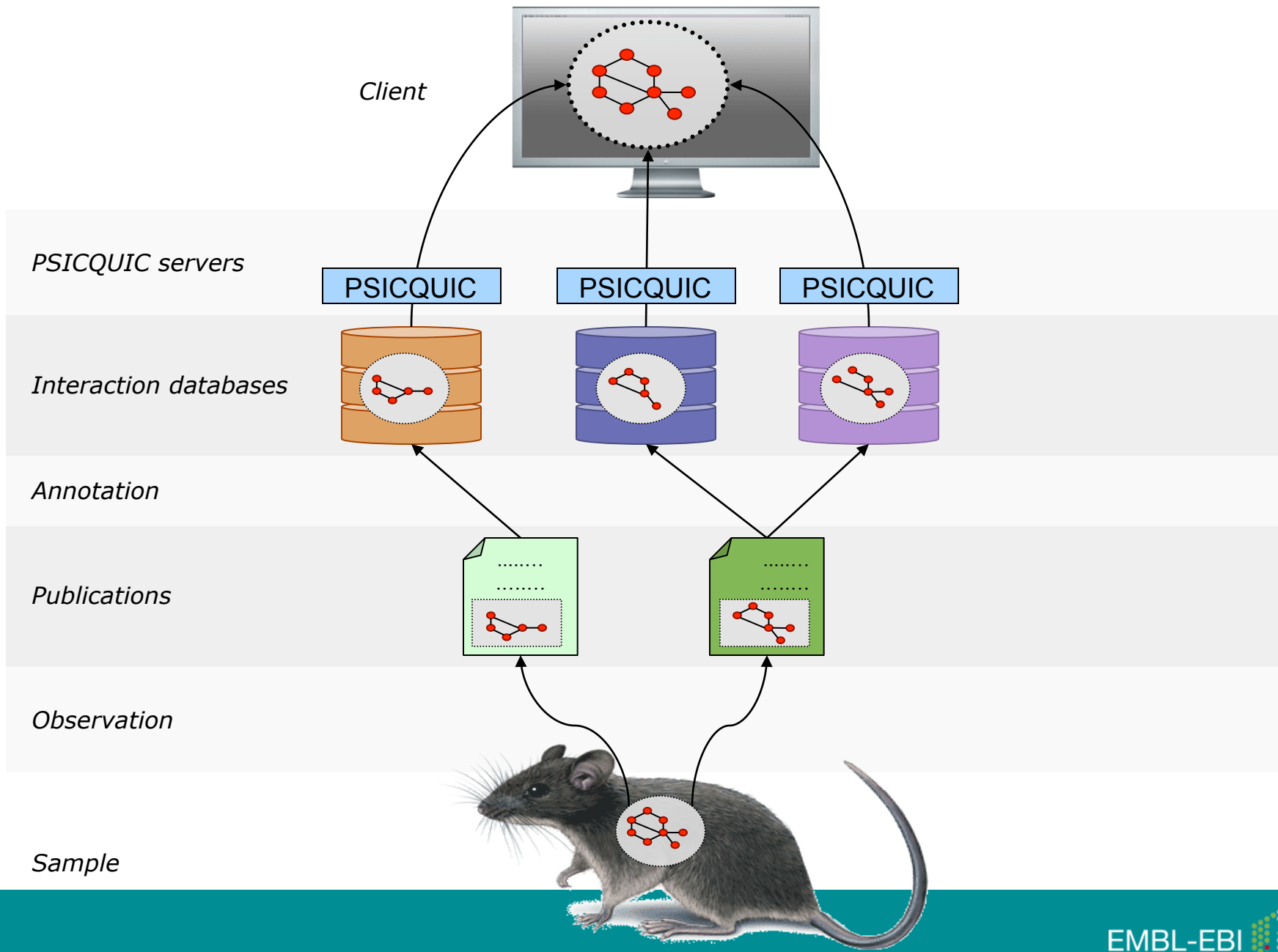
Article first published online: 13 APR 2010

Molecular Interactions 2002



PSI-MI





The PSI Common Query Interface: PSICQUIC

- Common computational interface for querying molecular interaction databases
 - Not limited to protein-protein interactions, also e.g.
 - Drug-target interactions
 - Simplified pathway data

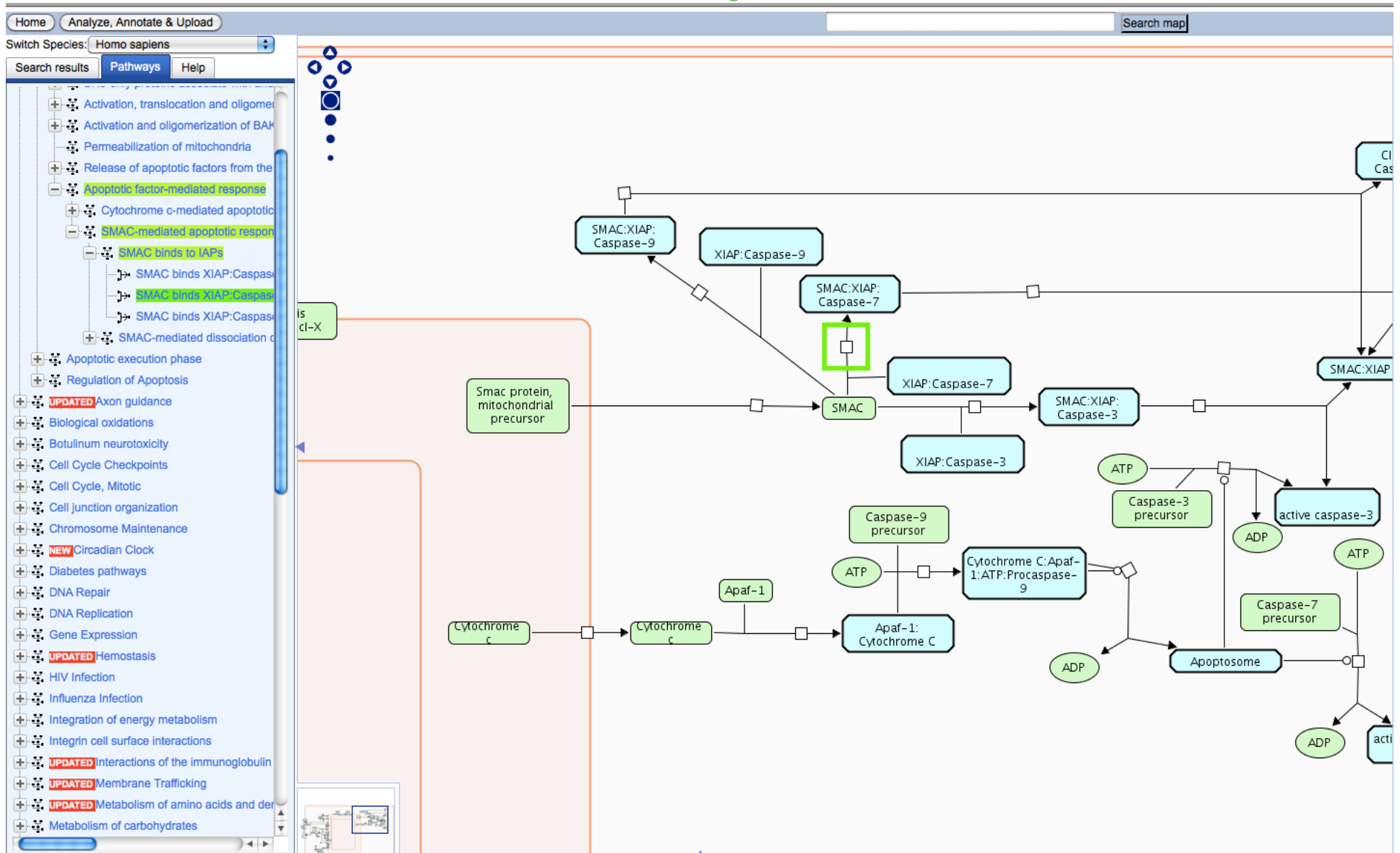
Search: [MIQL syntax reference](#)
[Fields »](#)

Total: **16,313,207** binary interactions

Click on the links below to display the results for each service ([refresh](#))

APID - 416,124	DIP - 107,619	iRefIndex - 1,465,119	Reactome - 74,861
BIND - 192,961	InnateDB - 9,909	MatrixDB - 845	Reactome-FIs - 209,988
BioGrid - 362,355	IntAct - 272,410	MINT - 129,972	STRING - 12,231,763
ChEMBL - 606,455	Interporc - 208,558	MPIDB - 24,268	

Pathway view



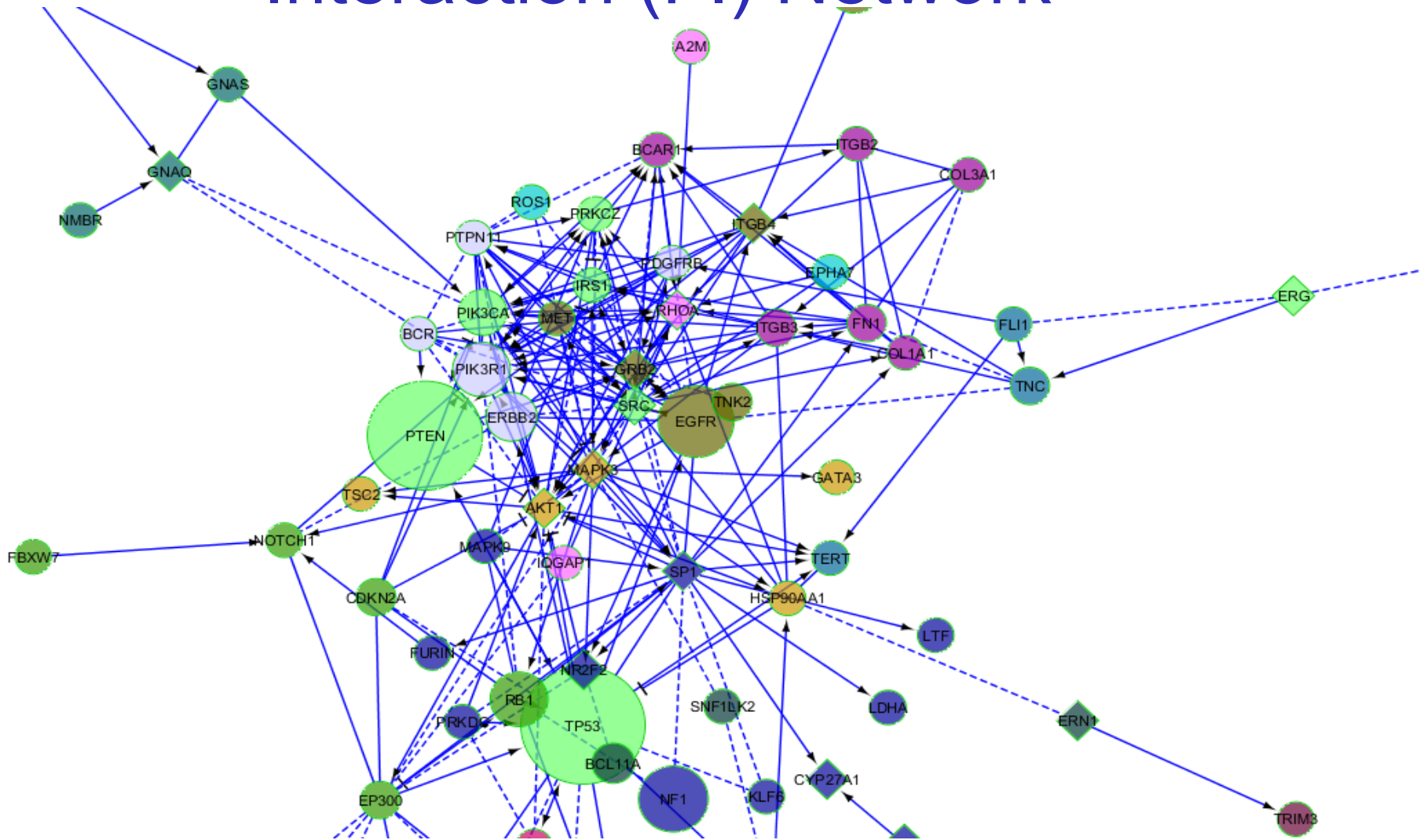
Interaction overlay: IntAct

The screenshot displays the IntAct web interface for pathway analysis. On the left, a search sidebar shows a hierarchical tree of biological processes, with 'SMAC-mediated apoptotic response' and 'SMAC binds to IAPs' highlighted. The main diagram area shows a network of interactions between various proteins and molecules, including SMAC, XIAP, Caspase-9, Caspase-7, Caspase-3, and Apaf-1. Interactions are represented by lines connecting nodes, with some nodes highlighted in blue or green. A control panel on the right, titled 'Analyze, Annotate & Upload', allows users to select the interaction database (IntAct), upload files, and set a confidence level threshold (0.50). It also includes options for coloring interactions and displaying/exporting tables of interactors.

Interaction overlay: ChEMBL

The screenshot displays the Reactome ChEMBL interaction overlay tool. On the left, a sidebar lists various biological pathways, with 'Apoptotic factor-mediated response' and 'SMAC-mediated apoptotic response' expanded. The main area shows a pathway diagram with nodes such as 'Smac protein, mitochondrial precursor', 'Caspase-9 precursor [cytosol]', 'Apaf-1: Cytochrome C', and 'Apoptosome'. Chemical structures are overlaid on several nodes, including SMAC, XIAP, and caspases. A control panel on the right allows for database selection (ChEMBL), file upload, and confidence level threshold setting (0.50). It also includes options for coloring and displaying/exporting interactor tables.

Reactome Functional Interaction (FI) Network



28 August 2011

www.reactome.org

FI Network Analysis Pipeline

Generate Functional Interactions

Select gene list (mutated, over-expressed, down-regulated, amplified or deleted genes)

Project genes of interest onto Reactome F.I. Network

Identify Relevant (e.g. Disease) Subnetwork

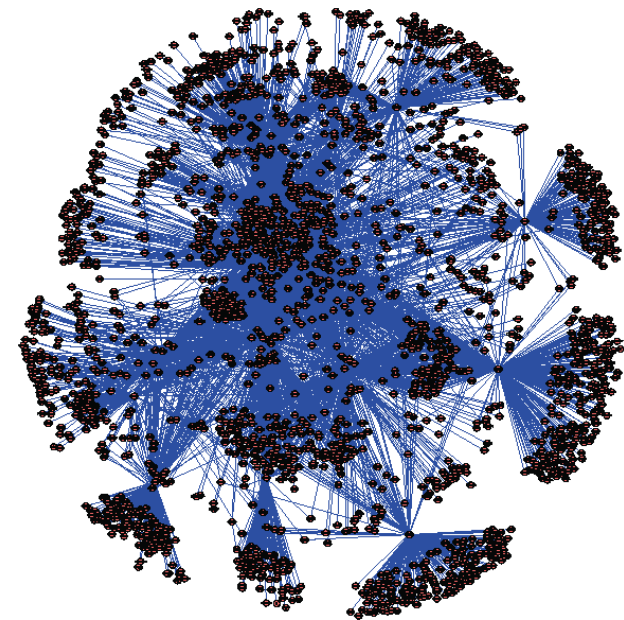
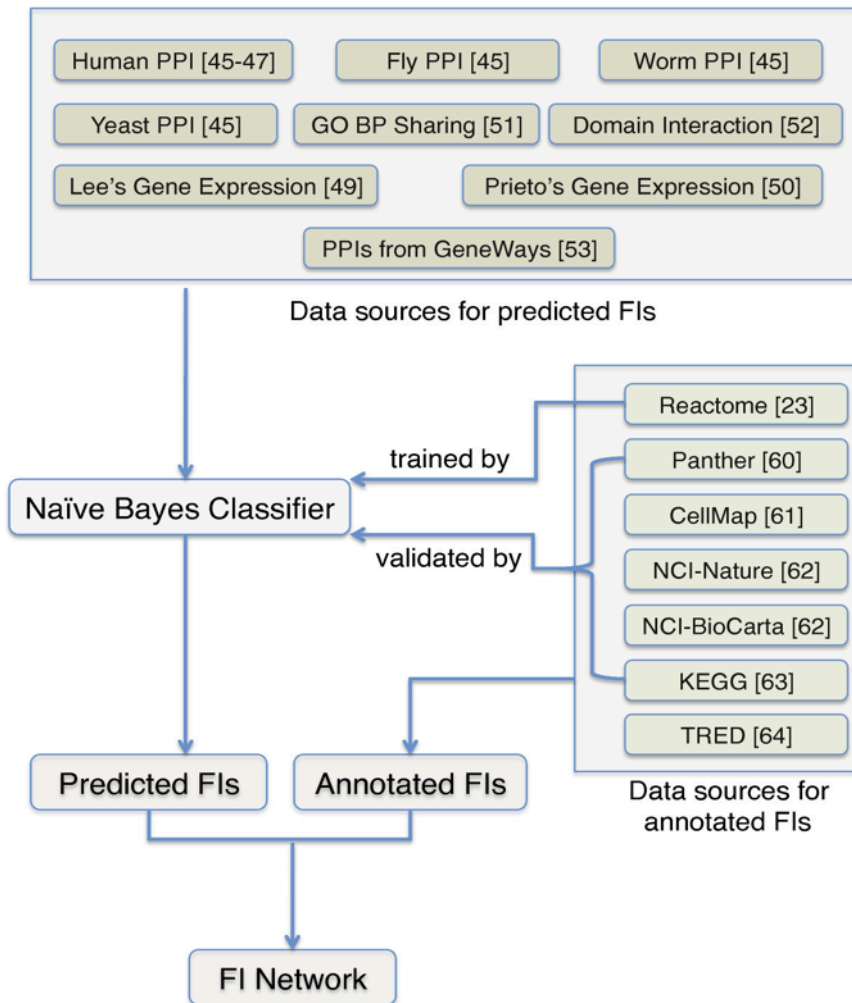
Apply Clustering Algorithms

Apply Pathway/GO Annotation to each cluster

Generate Biological Hypothesis!

**T2D
OvCa
GBM
Breast
Prostate
TCGA
ICGC**

Construction of the Reactome FI Network

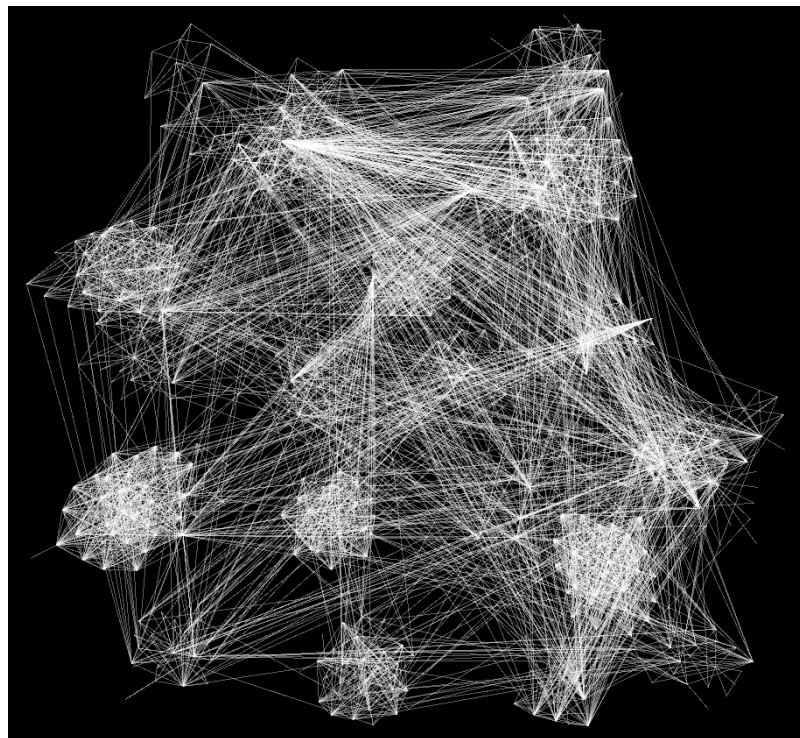


Reactome Functional Interaction (FI) Network [15%]

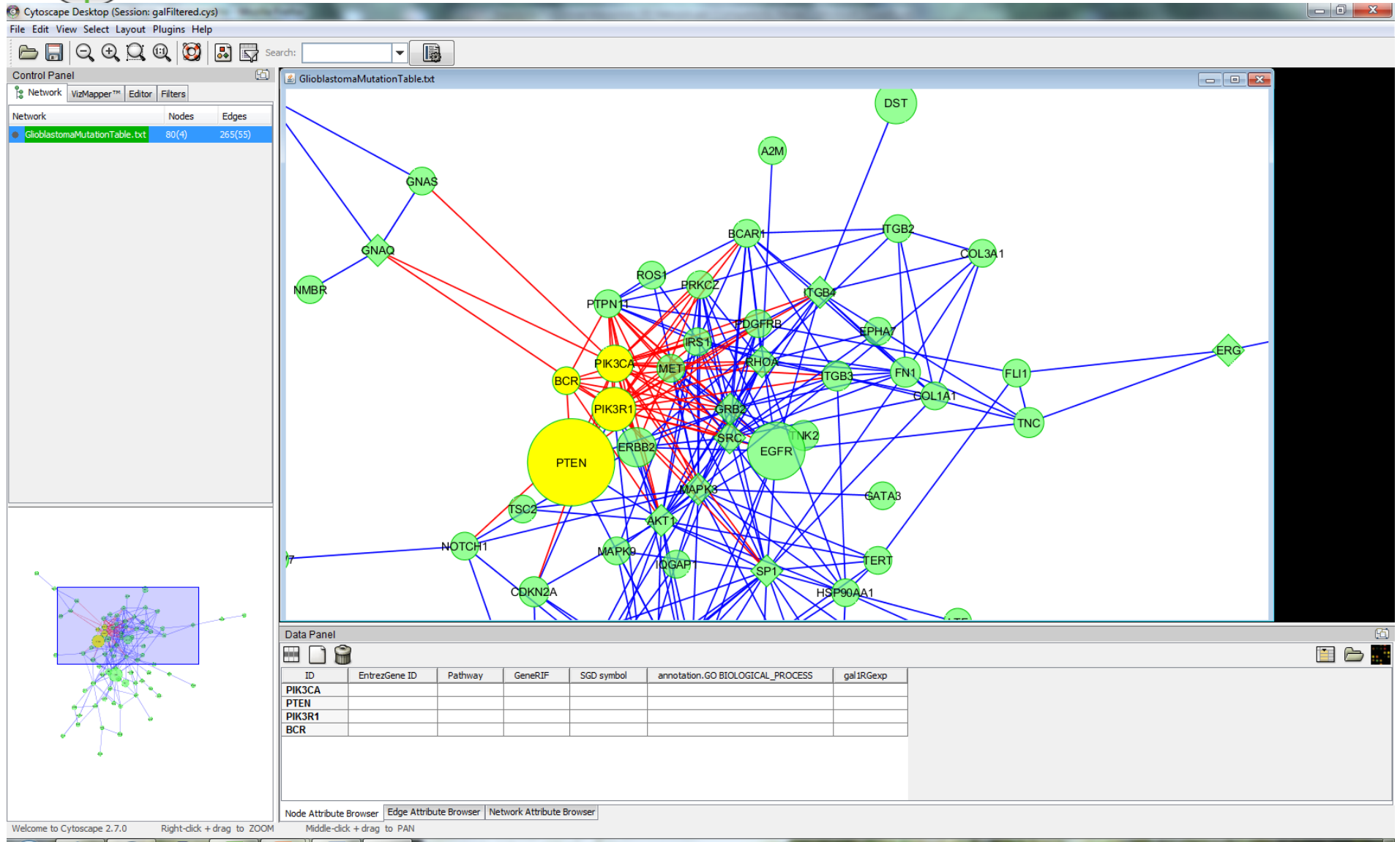
Wu et al. *Genome Biology* 2010

Coverage after Prediction

Source Type	Proteins	SwissProt Proteins (coverage)	Interaction
Pathways	6316	5496 (27%)	98590
Predicted	8345	7546 (37%)	111398
Total	10956	9393 (46%)	209988

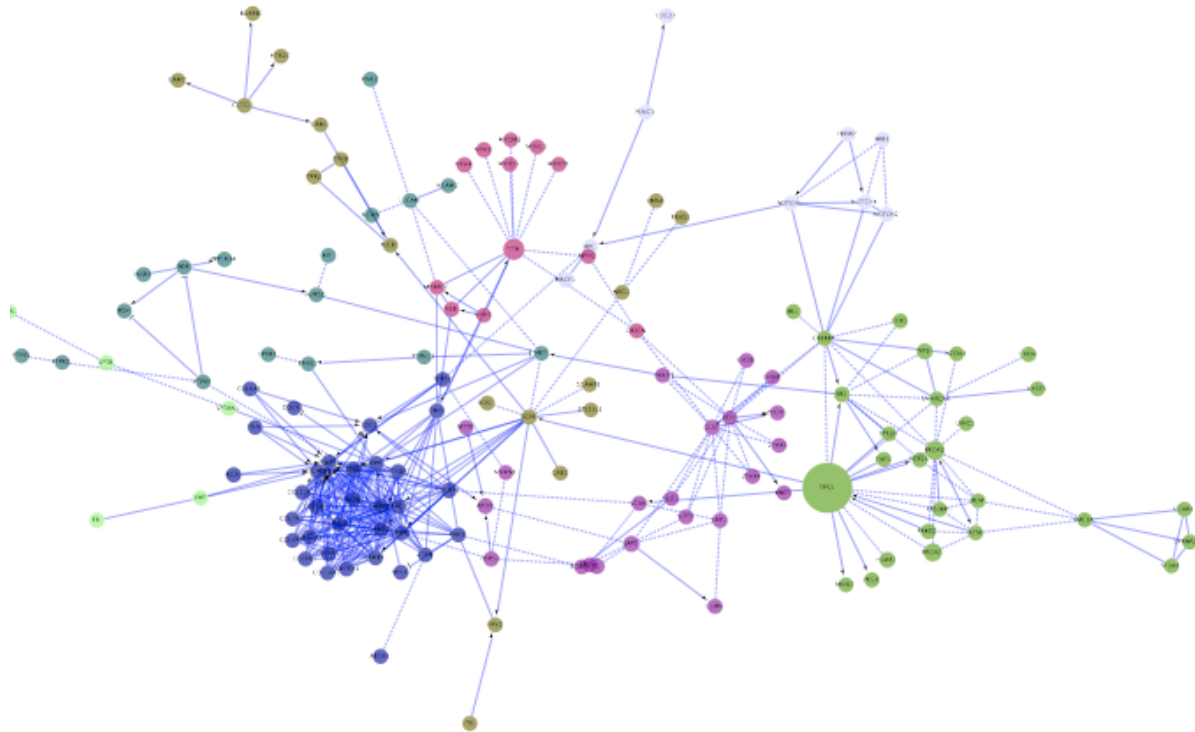


- 10,956 proteins (9,542 genes)
- 209,988 FIs
- 5% of network shown here



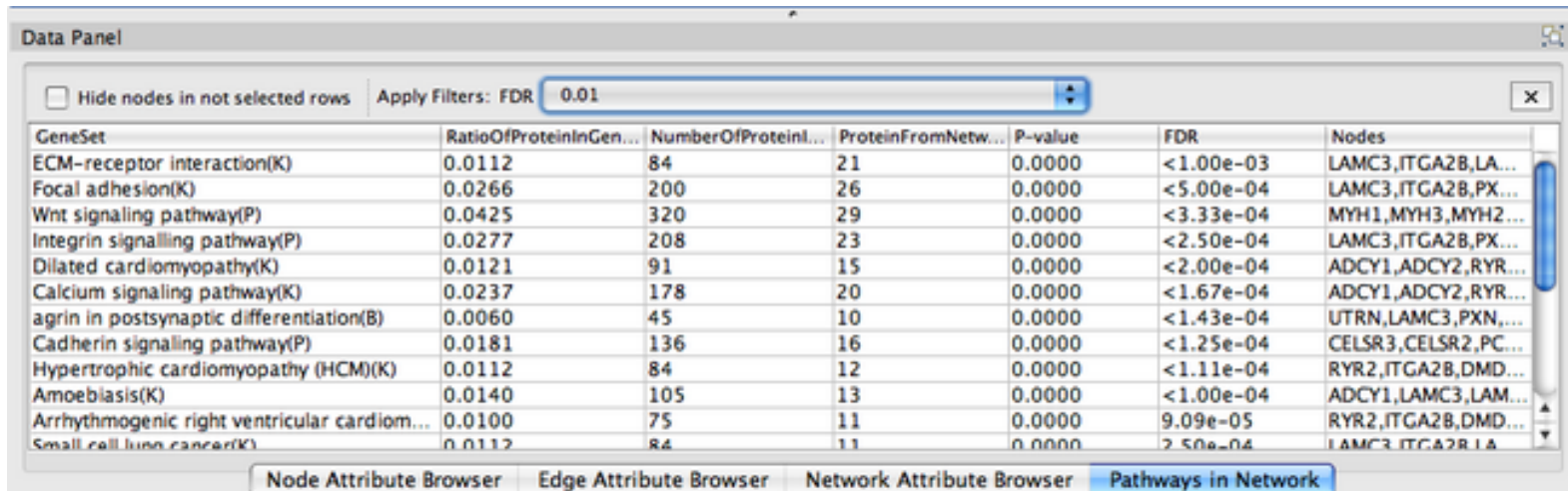
Cluster FI network

- Runs spectral partition based network clustering ([Newman, 2006](#)) on the displayed FI network.
- Nodes in different network modules will be shown in different colours (max 15 colours).



Analyze network functions

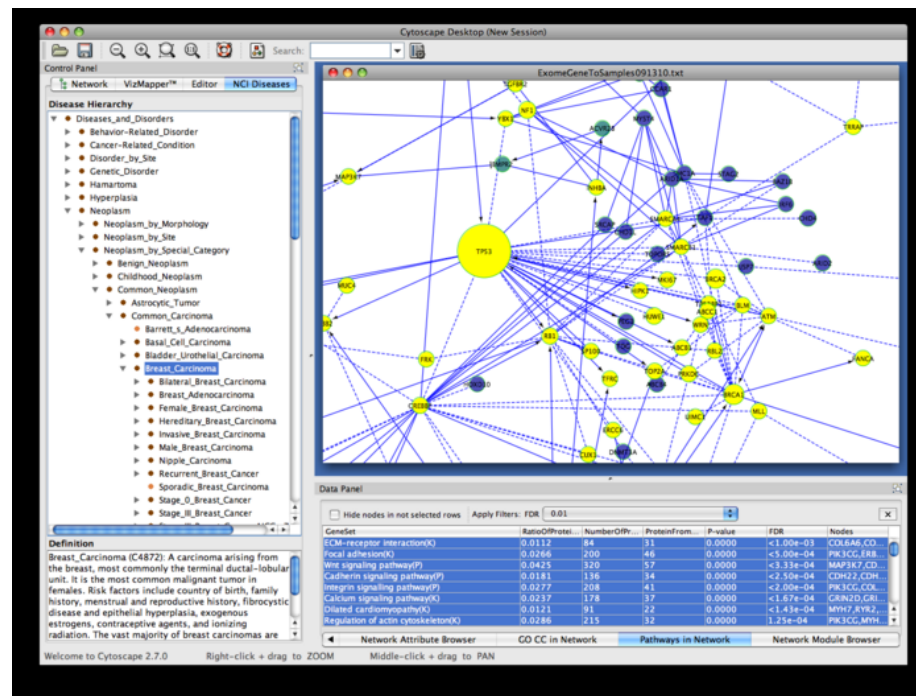
- Pathway or GO term enrichment analysis for the displayed network.
 - Use Data Panel to:
 - Filter enrichment results by FDR value.
 - Check “Hide nodes in not selected rows” to display only nodes in the selected row or rows
 - To analyze an unlinked set of genes choose the "Show genes not linked to others" option when importing data



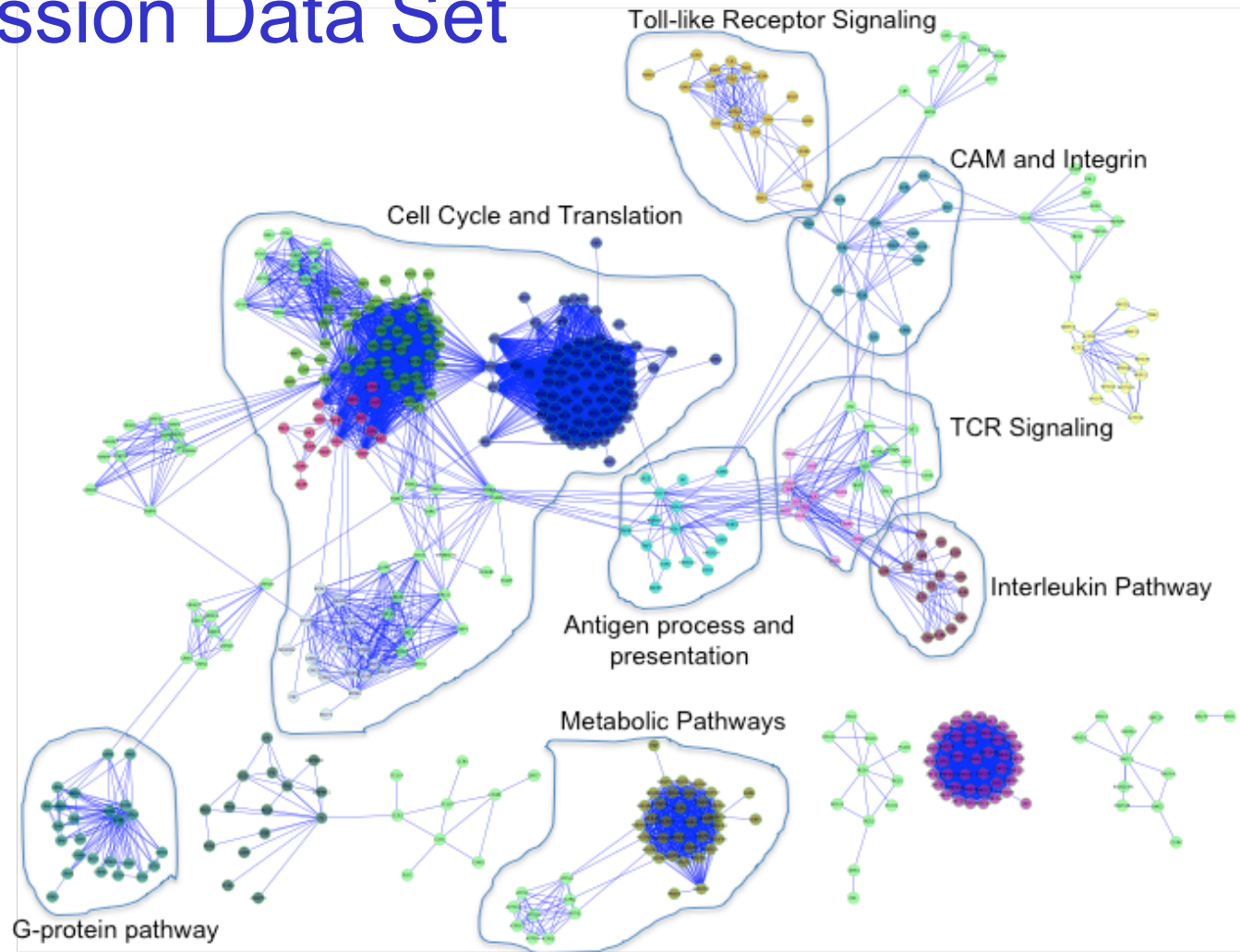
GeneSet	RatioOfProteinInGen...	NumberOfProtein...	ProteinFromNetw...	P-value	FDR	Nodes
ECM-receptor interaction(K)	0.0112	84	21	0.0000	<1.00e-03	LAMC3,ITGA2B,LA...
Focal adhesion(K)	0.0266	200	26	0.0000	<5.00e-04	LAMC3,ITGA2B,PX...
Wnt signaling pathway(P)	0.0425	320	29	0.0000	<3.33e-04	MYH1,MYH3,MYH2...
Integrin signalling pathway(P)	0.0277	208	23	0.0000	<2.50e-04	LAMC3,ITGA2B,PX...
Dilated cardiomyopathy(K)	0.0121	91	15	0.0000	<2.00e-04	ADCY1,ADCY2,RYR...
Calcium signaling pathway(K)	0.0237	178	20	0.0000	<1.67e-04	ADCY1,ADCY2,RYR...
agrin in postsynaptic differentiation(B)	0.0060	45	10	0.0000	<1.43e-04	UTRN,LAMC3,PXN...
Cadherin signaling pathway(P)	0.0181	136	16	0.0000	<1.25e-04	CELSR3,CELSR2,PC...
Hypertrophic cardiomyopathy (HCM)(K)	0.0112	84	12	0.0000	<1.11e-04	RYR2,ITGA2B,DMD...
Amoebiasis(K)	0.0140	105	13	0.0000	<1.00e-04	ADCY1,LAMC3,LAM...
Arrhythmogenic right ventricular cardiom...	0.0100	75	11	0.0000	9.09e-05	RYR2,ITGA2B,DMD...
Small cell lung cancer(K)	0.0112	84	11	0.0000	2.50e-04	LAMC3,ITGA2B,LA...

Load Cancer Gene Index

- This method loads the tree of NCI disease terms in the left panel.
- Select a disease term in the tree to select all diagram objects that have this annotation or one of its sub-terms.



A FI Sub-Network Generated from Significant Network Modules based on a TCGA OV Expression Data Set



Acknowledgements and Resources

- www.reactome.org
 - Lincoln Stein, Peter D'Eustachio, Ewan Birney, Steve Jupe
 - Croft D, et al. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 2011 Jan;39 (Database issue):D691-7.
 - Reactome is supported by NIH grant P41 HG003751 and EU grant LSHG-CT-2005-518254 "ENFIN".
- http://wiki.reactome.org/index.php/Reactome_FI_Cytoscape_Plugin
 - Guanming Wu, Lincoln Stein
 - Wu G, Feng X, Stein L. A human functional protein interaction network and its application to cancer data analysis. *Genome Biol.* 2010;11(5):R53.
- <http://www.ebi.ac.uk/intact>
 - Samuel Kerrien, Bruno Aranda, Sandra Orchard
 - Aranda B, et al. The IntAct molecular interaction database in 2010. *Nucleic Acids Res.* 2010 Jan;38(Database issue):D525-31.
 - IntAct is supported by EU grants SLING 226073, PSIMEx HEALTH-2007-223411, APO-SYS FP7-HEALTH-2007-200767.
- <http://www.ebi.ac.uk/Tools/webservices/psicquic/view>
 - Bruno Aranda, Sam Kerrien, and many participants of the HUPO PSI.
 - Aranda B, et al. PSICQUIC and PSISCORE: accessing and scoring molecular interactions. *Nat Methods.* 2011 Jun 29;8(7):528-9.
 - PSICQUIC is supported by EU grant PSIMEx HEALTH-2007-223411.