

Open PHACTS

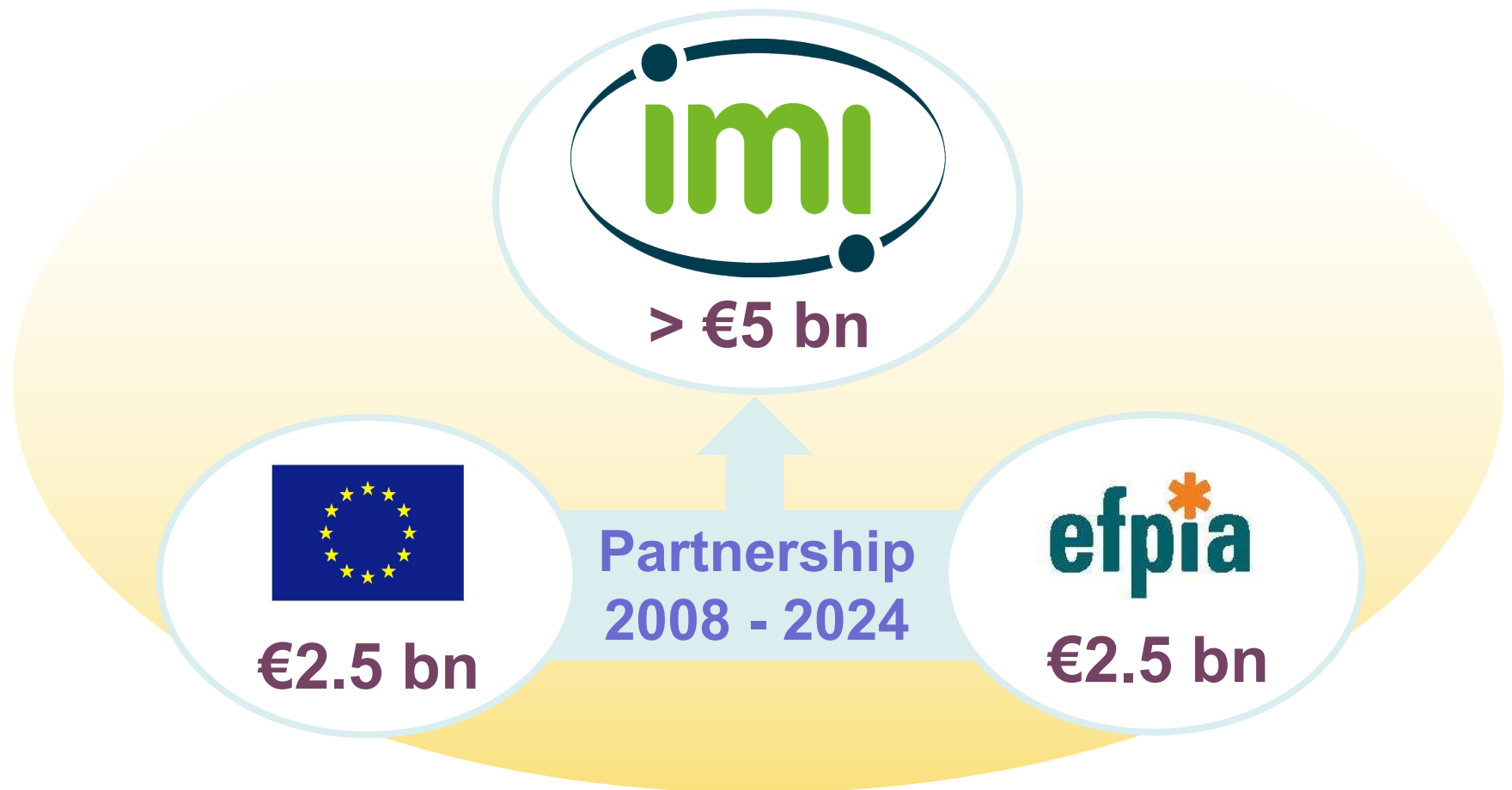
Open Pharmacological Space

@Chris Evelo
Dept Bioinformatics - BiGCaT
Maastricht University

chris.evelo@maastrichtuniversity.nl



IMI – Europe's partnership for health





How do R&D companies use public data?



Literature Patents
Genbank PubChem



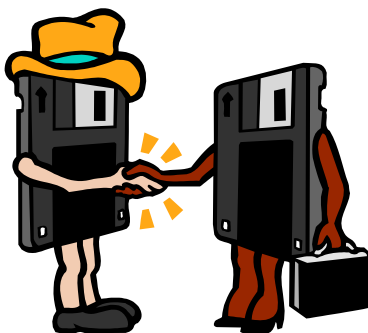
Databases



Downloads



Data Analysis



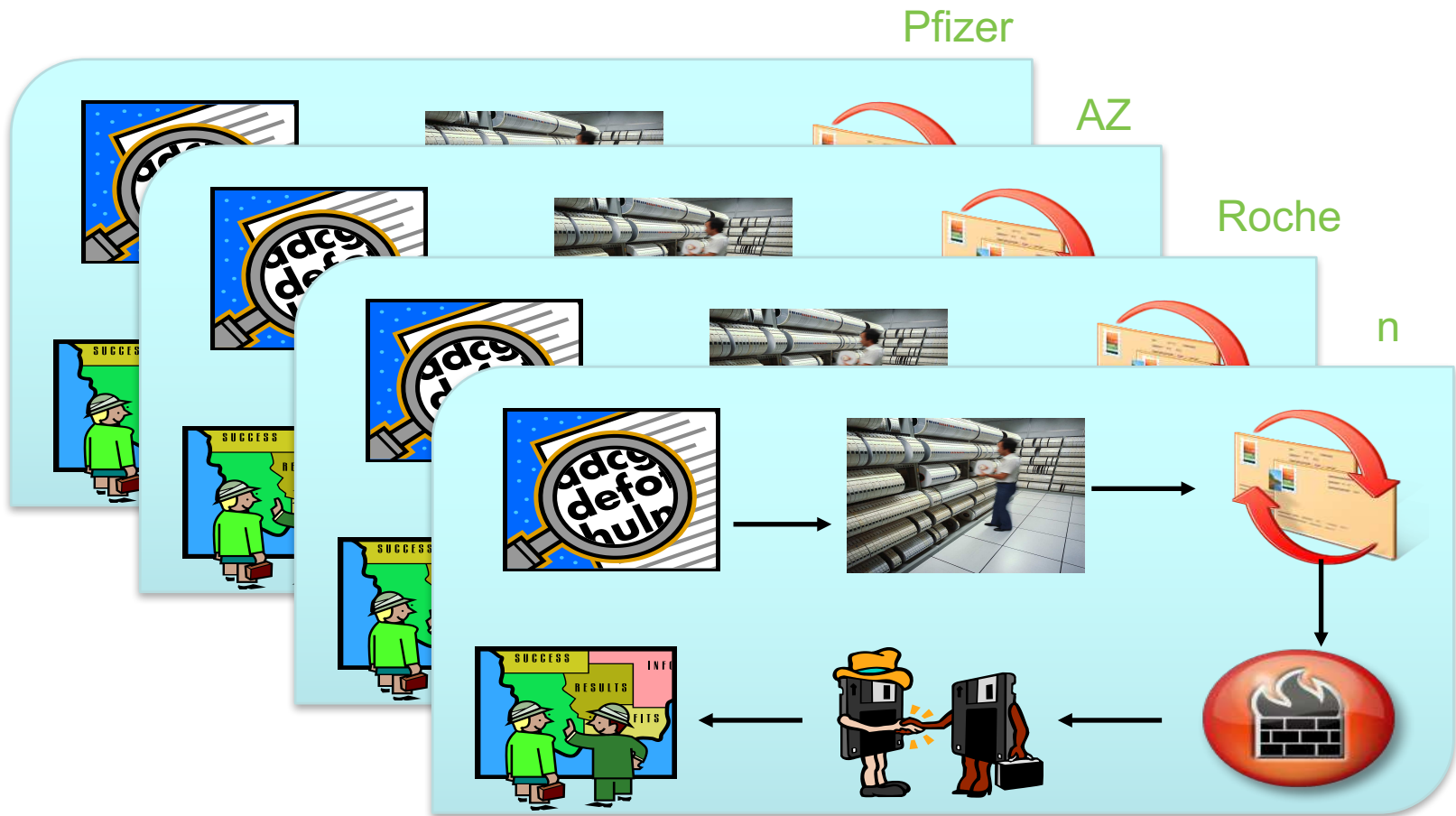
Data Integration



Firewalled Databases



How do pharma companies use public data?





Open PHACTS Mission:
Integrate Multiple Research
Biomedical Data Resources
Into A Single **Open & Free**
Access Point

...and make it sustainable in the long term



"What is the selectivity profile of known p38 inhibitors?"



"Let me compare MW, logP and PSA for known oxidoreductase inhibitors"



"Find me compounds that inhibit targets in NFkB pathway assayed in only functional assays with a potency <math><1 \mu\text{M}</math>"



ChEMBL

DrugBank

Gene
Ontology

Wikipathways

GeneGo

ChEBI

UniProt

UMLS

neXtProt

GVKBio

ConceptWiki

ChemSpider

DisGeNet

TrialTrove

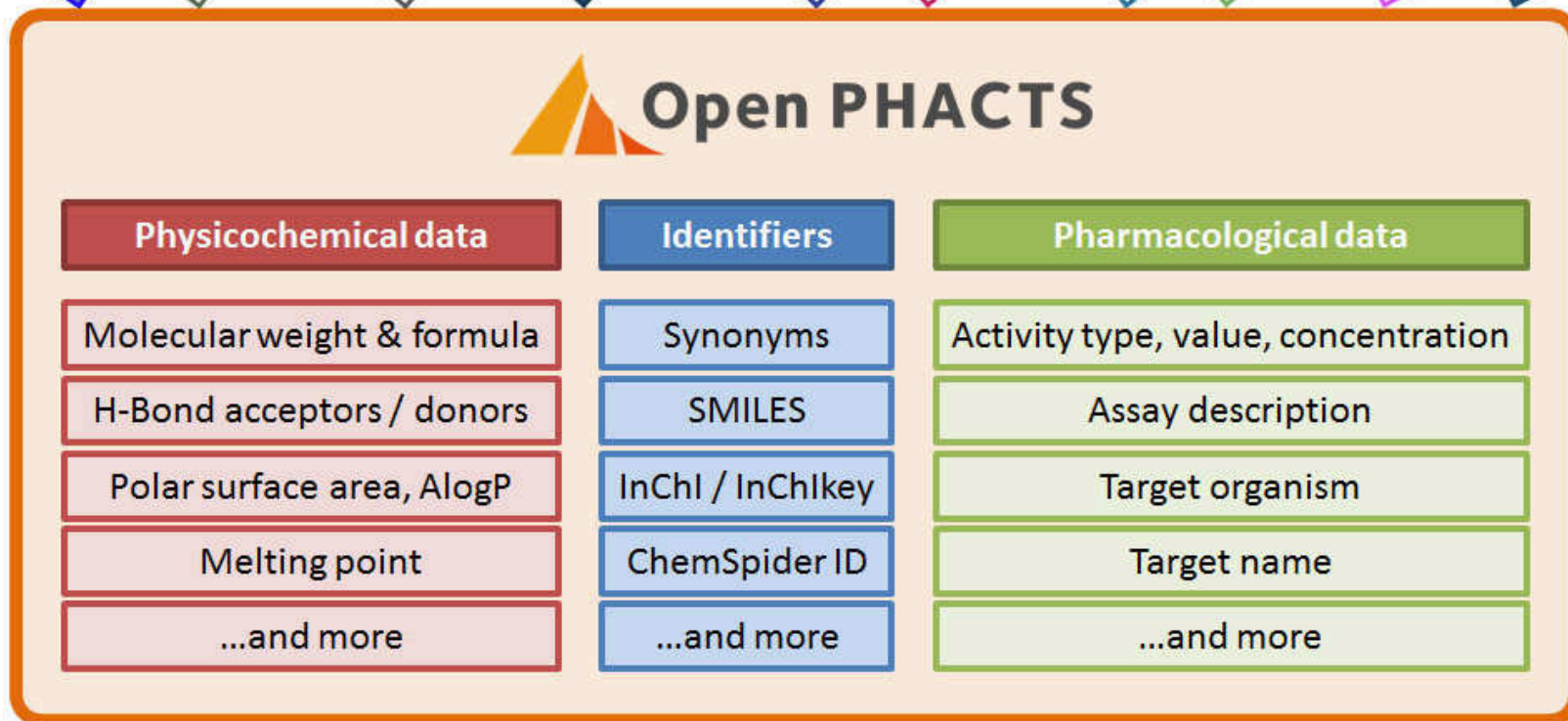
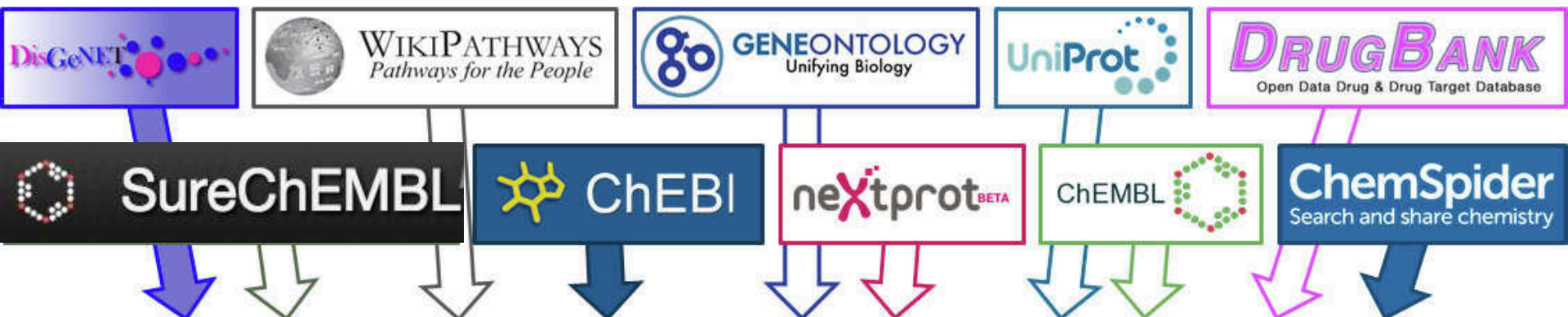
TR
Integrity

ChEMBL
Target Class

ENZYME

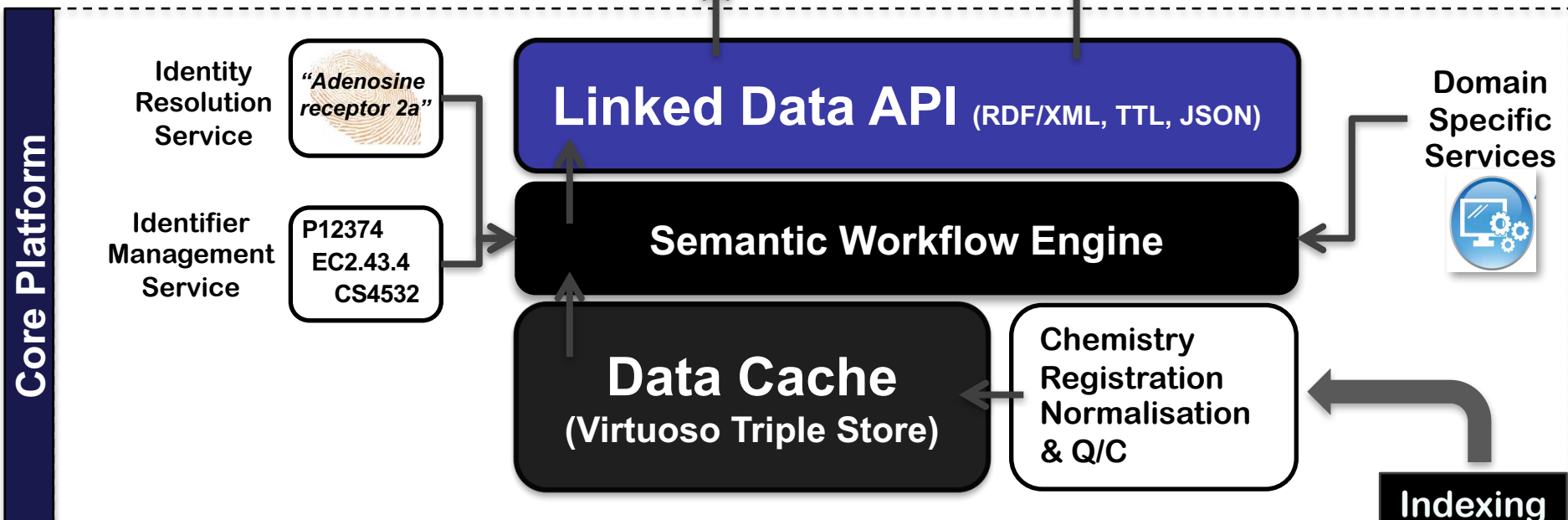
FDA adverse
events

SureChEMBL





Apps



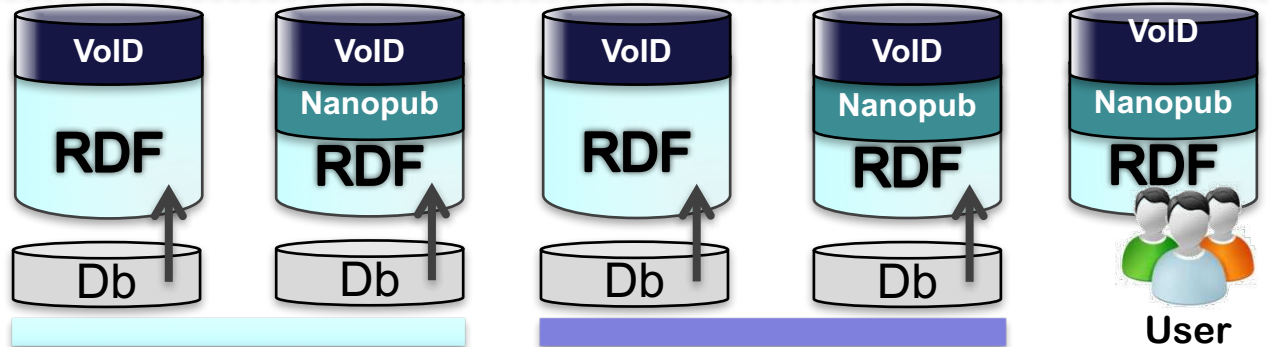
Domain Specific Services



Indexing



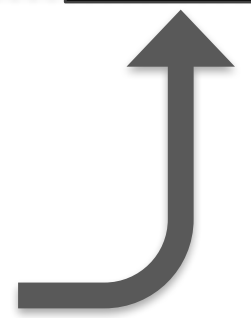
Public Ontologies



Public Content

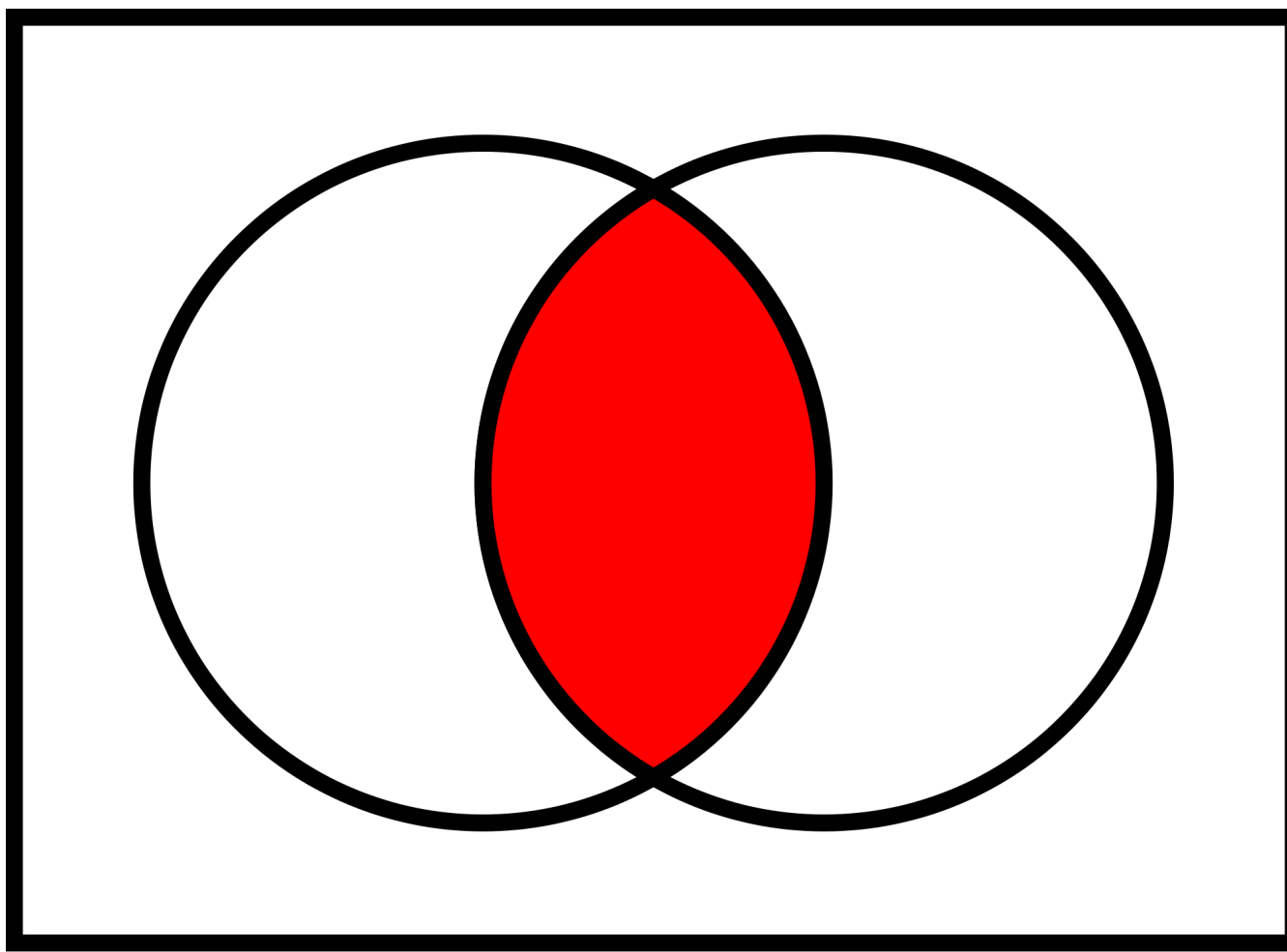
Commercial

User Annotations



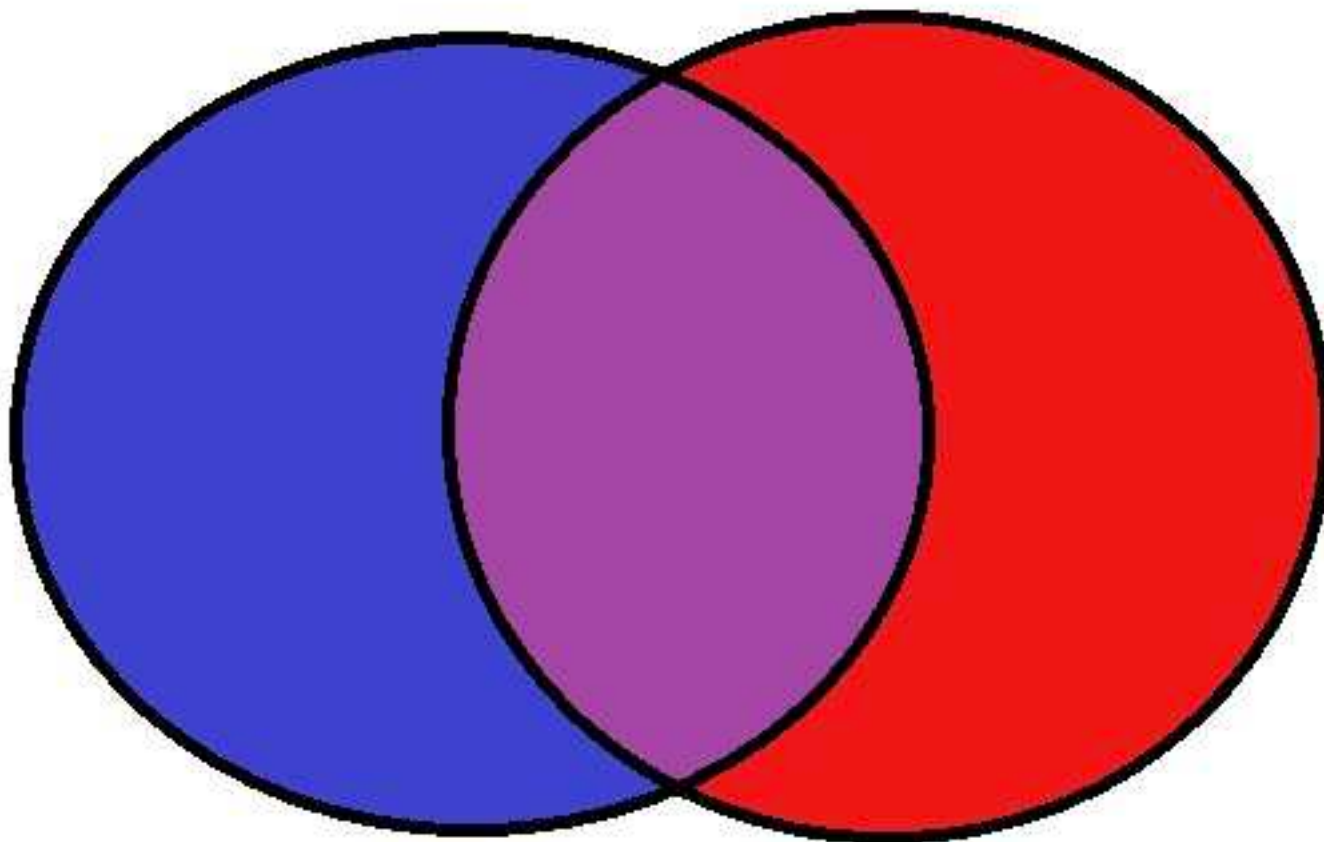


Link one resource to another



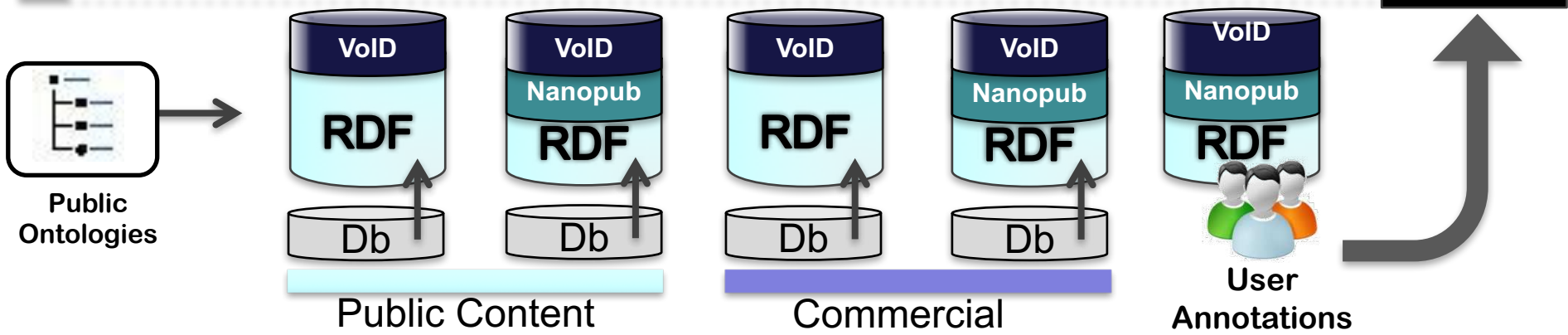
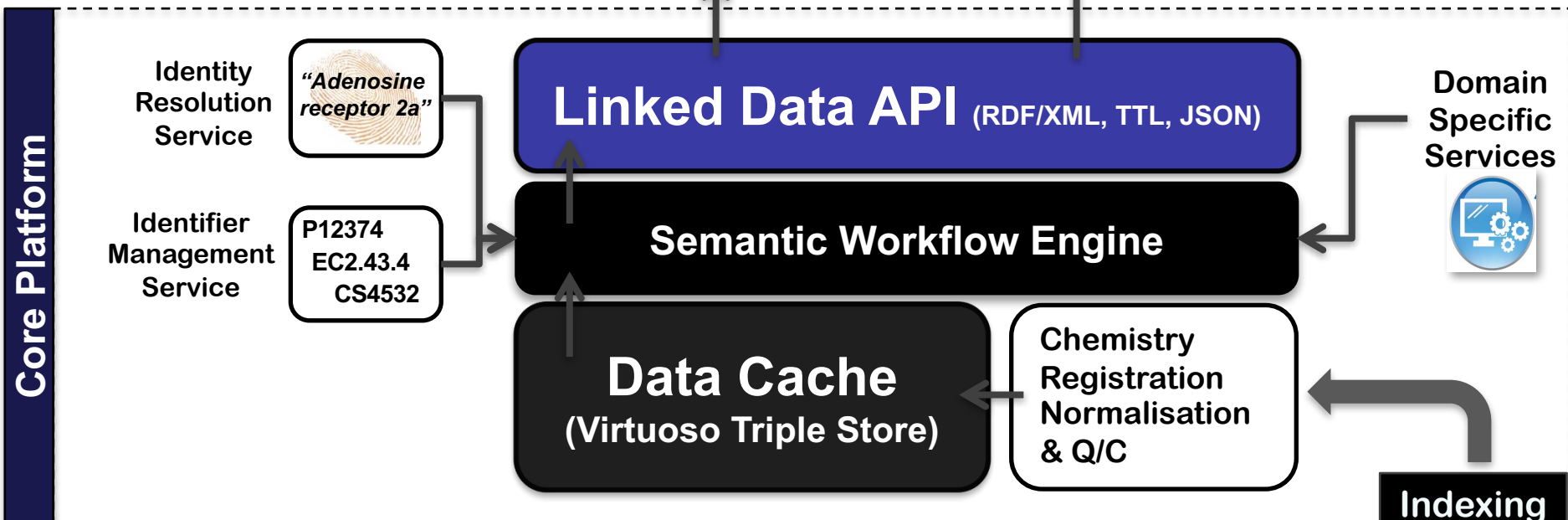


Or use both and map





Apps





We need mapping

- **ID Mapping**

BridgeDb based, installable webservice or Docker VM

Derby database or semantic linkset

Links created from curated resources (ENSEMBL, HMDB/ChEBI, RHEA)

Uses Identifiers.org, can do transitive, can be used using lenses

- **Concept (text) mapping**

Conceptwiki approach abandoned, indexing not full solution

Missing...

- **Ontology searches (OLS) and mapping**

Lots of manual work (NCBO, Bioportal)

- **Chemistry registration**

Finds structures and substructures (based on names, IDs, INCHI, drawn structure)

Not (yet?) Open Source

- **Sequence registration**

Technically simple, not there (Natasa Przulj at NetBio SIG)



The nutrigenomics request to bioinformatics



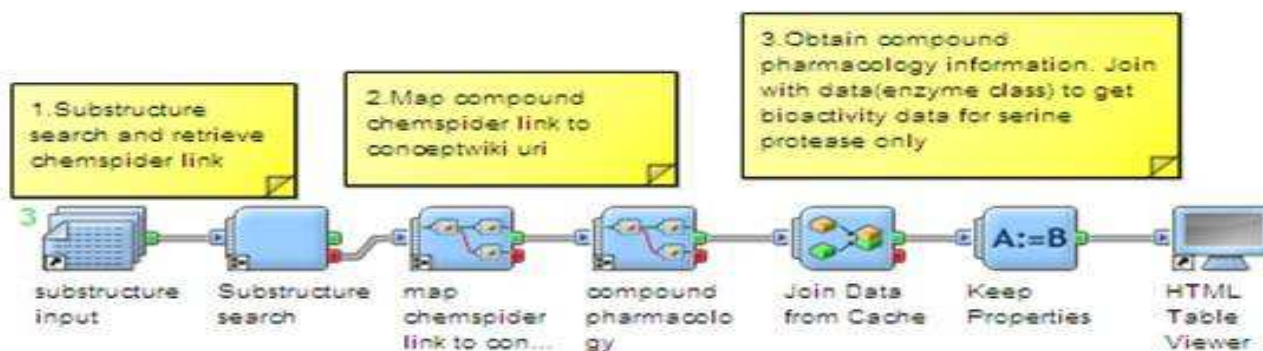
KNIME

Table View - 0:31 - Interactive Table (7 x 6)

Name	Inchi	Activity	Units	Relation	Target
.. Sorafenib	MLDQTXFUGDVEO-UHFFFAOYSA...	3400	nM	=	Serine/threonine-protein kinase PLK4
.. Sorafenib	MLDQTXFUGDVEO-UHFFFAOYSA...	250	nM	=	MAP kinase signal-integrating kinase 2
.. Sorafenib	MLDQTXFUGDVEO-UHFFFAOYSA...	5,4	uM	=	HCT-116 (Colon carcinoma cells)
.. Sorafenib	MLDQTXFUGDVEO-UHFFFAOYSA...	1700	nM	=	Ephrin type-B receptor 1
.. Sorafenib	MLDQTXFUGDVEO-UHFFFAOYSA...	3300	nM	=	Dual specificity mitogen-activated protein kinase kin.
.. Sorafenib	MLDQTXFUGDVEO-UHFFFAOYSA...	6200	nM	=	Cyclin-dependent kinase 5

Workflow Diagram:

- File Reader**: Simply gets the URL [I dont know how to get it to start otherwise!]
- Java Snippet**: Fetch JSON from web
- Get Name and Inchi**: Name & Inchi Grabber
- Get Activity**: Now turn the activity JSON into rows
- Activity Parser**: For each activity row, extract the columns we want
- Column Filter**: Tidy Up: Remove Processing Columns Now
- Interactive Table**: Node 31









UPDATE, UPDATE, UPDATE

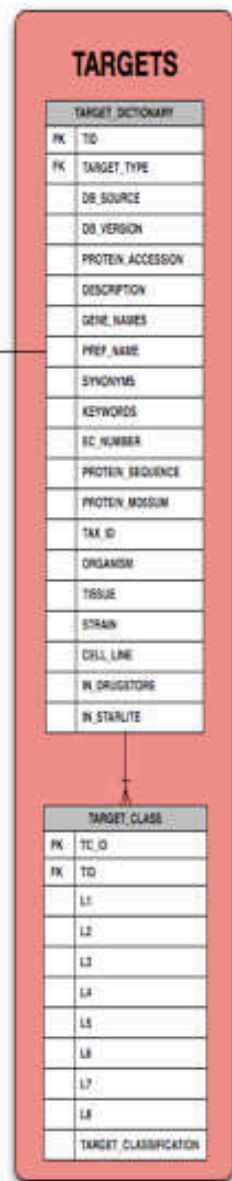
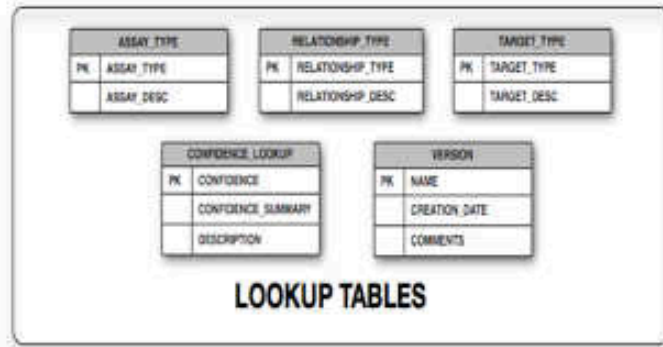
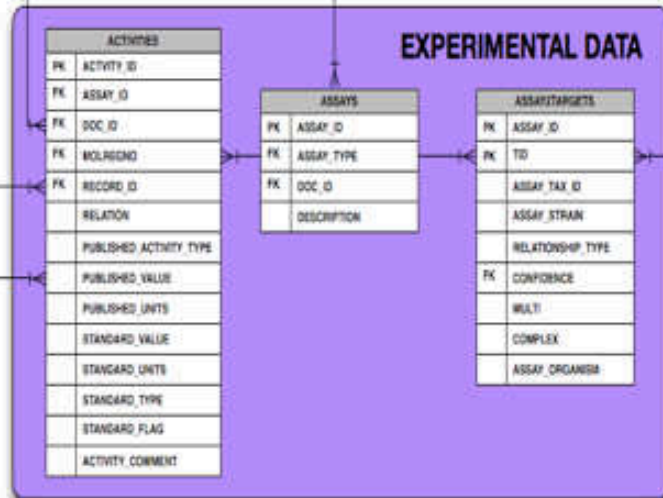
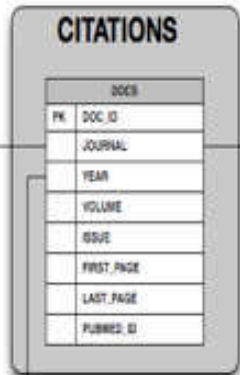
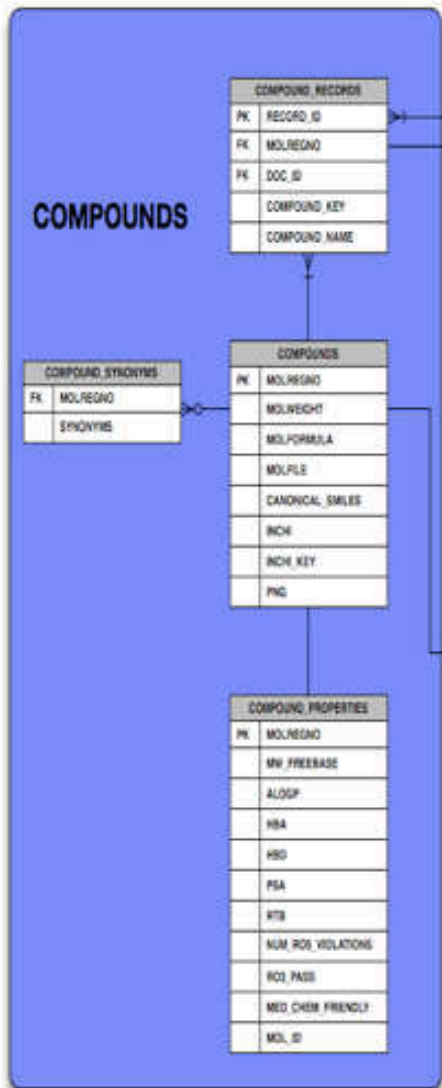
- ❖ Regular data updates as the core data refreshes
- ❖ API updates aligned to new business questions and changes
- ❖ Workstreams to add further new data – see later
- ❖ New release May 2016 2.1
 - SureChEMBL and Pathways update
- ❖ Further updates planned for summer 2016



DATA IS ALWAYS EVOLVING

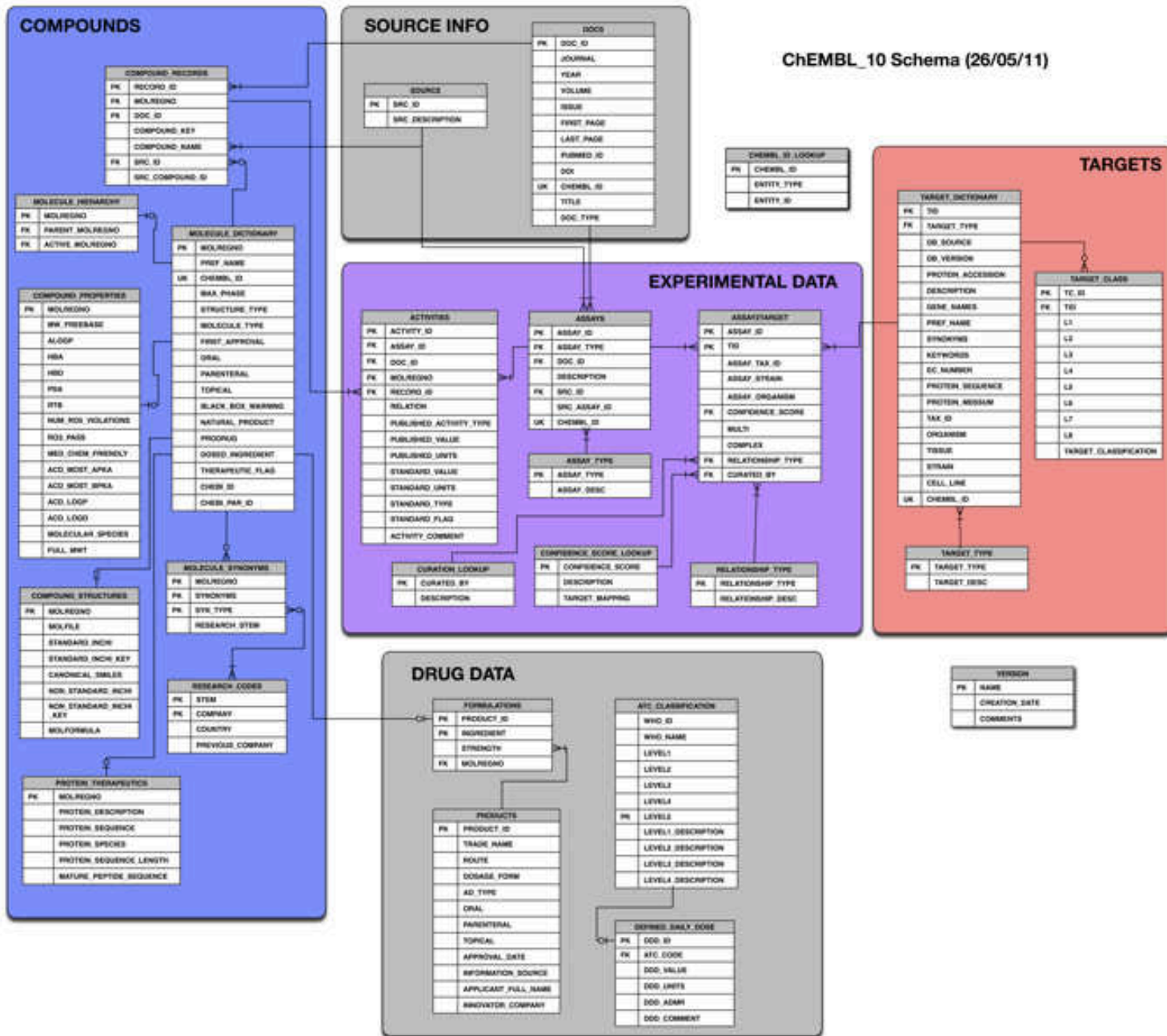


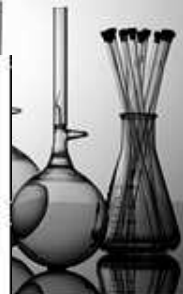
ChEMBL01 2009



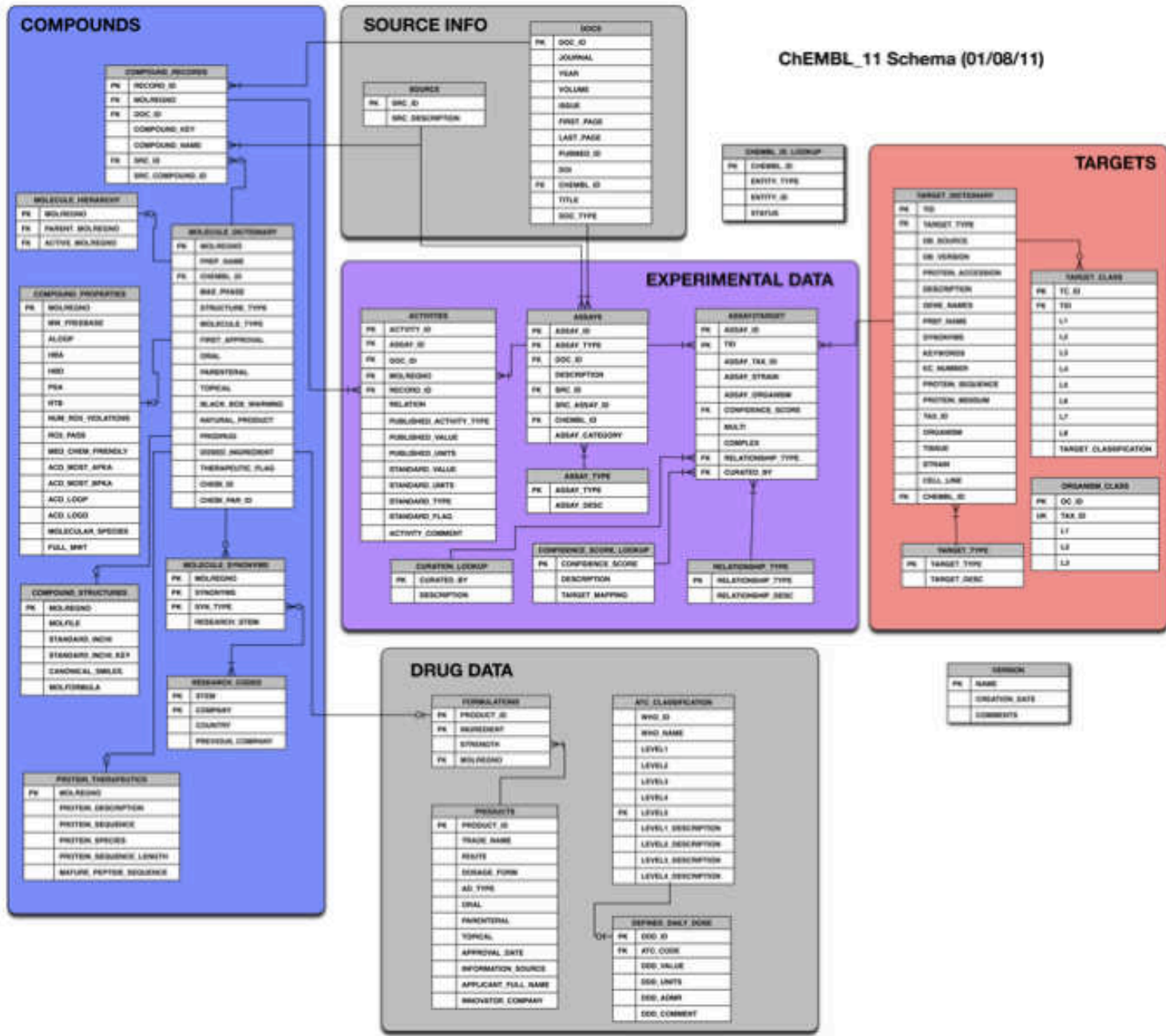


ChEMBL_10 Schema (26/05/11)





ChEMBL_11 Schema (01/08/11)





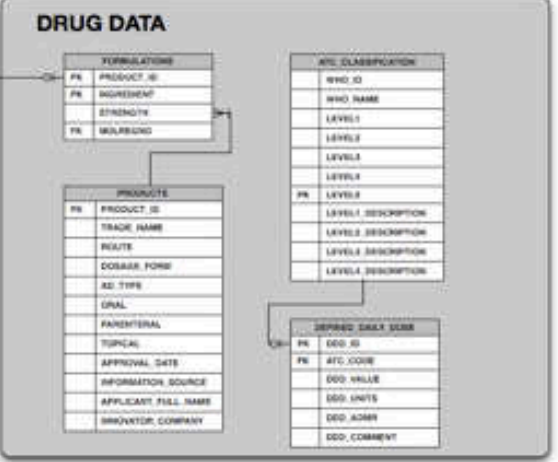
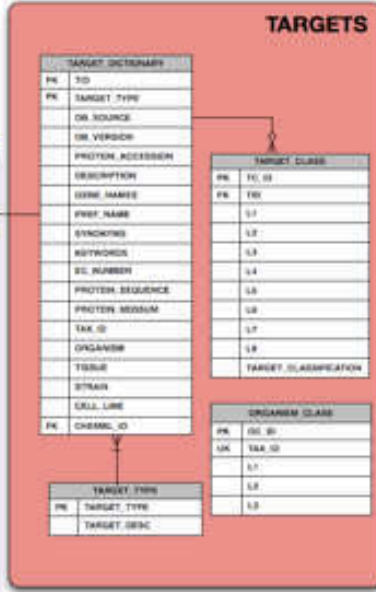
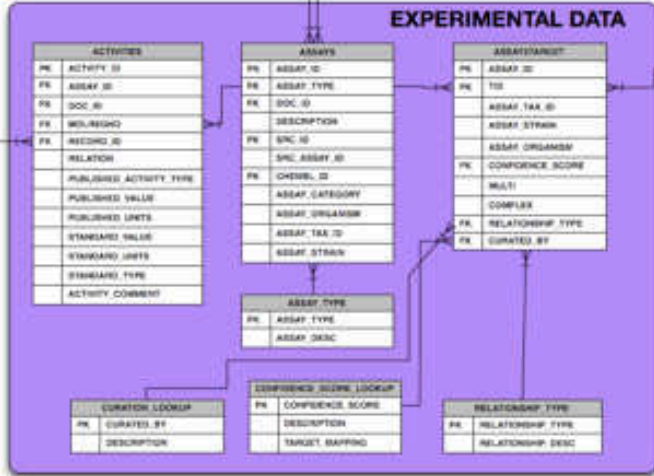
COMPOUNDS



SOURCE INFO



ChEMBL_12 Schema (18/11/11)





COMPOUNDS

COMPOUND RECORDS	
PK	RECORD_ID
FK	MOLREGNO
FK	DOC_ID
	COMPOUND_KEY
	COMPOUND_NAME
FK	SRC_ID
	SRC_COMPOUND_ID

MOLECULE HIERARCHY	
FK	MOLREGNO
FK	PARENT_MOLREGNO
FK	ACTIVE_MOLREGNO

COMPOUND PROPERTIES	
FK	MOLREGNO
	MW
	PRELOG
	ALOGP
	HBIA
	HBII
	PSA
	RTS
	RUN_RCN_VIOLATION
	FOI_PASS
	REG_OEM_FRENELT
	ADD_MOET_BPH4
	ADD_MOET_BPH4
	ADD_LODF
	ADD_LOGE
	MOLCULAR_SPECIES
	FULL_MWT

COMPOUND STRUCTURES	
FK	MOLREGNO
	MOLFILE
	STANDARD_INCHI
	STANDARD_INCHI_KEY
	CANONICAL_SMILES
	MOLFORMULA

PROTEIN THERAPEUTICS	
FK	MOLREGNO
	PROTEIN_DESCRIPTION
	PROTEIN_SEQUENCE
	PROTEIN_SPECIES
	PROTEIN_SEQUENCE_LENGTH
	NATURE_PROTEIN_SEQUENCE

MOLECULE DICTIONARY	
FK	MOLREGNO
	PREF_NAME
FK	CHEMBL_ID
	SOURCE_TYPE
	MOLECULE_TYPE
	FIRST_APPROVAL
	ORAL
	PARENTERAL
	TOPICAL
	BLACK_BOX_WARNING
	NATURAL_PRODUCT
	PROGRAM
	DDOCS_APPROVED?
	THERAPEUTIC_FLAG
	CHEM_ID
	CHEM_PFX_ID

MOLECULE SYNONYMS	
FK	MOLREGNO
FK	SYNONYM
FK	SYN_TYPE
	RESEARCH_STEM

RESEARCH CODES	
FK	STEM
FK	COMPANY
	COUNTRY
	PREVIOUS_COMPANY

SOURCE INFO

SOURCE	
FK	SRC_ID
	SRC_DESCRIPTION

BOOK	
FK	DOC_ID
	JOURNAL
	YEAR
	VOLUME
	ISSUE
	FIRST_PAGE
	LAST_PAGE
	PUBMED_ID
	DOI
FK	CHEMBL_ID
	TITLE
	DOC_TYPE

CHEMBL_ID LOOKUP	
FK	CHEMBL_ID
	ENTITY_TYPE
	ENTITY_ID
	STATUS

ChEMBL_13 Schema (21/02/12)

EXPERIMENTAL DATA

LEADS EXP	
FK	ACTIVITY_ID
	SEQ
	SEQ

ACTIVITIES	
FK	ACTIVITY_ID
FK	ASSAY_ID
FK	DOC_ID
FK	MOLREGNO
FK	RECORD_ID
	RELATION
	PUBLISHED_ACTIVITY_TYPE
	PUBLISHED_VALUE
	PUBLISHED_UNITS
	STANDARD_VALUE
	STANDARD_UNITS
	STANDARD_TYPE
	ACTIVITY_COMMENT

ASSAYS	
FK	ASSAY_ID
FK	ASSAY_TYPE
FK	DOC_ID
FK	SRC_ID
FK	SRC_ASSAY_ID
FK	CHEMBL_ID
	ASSAY_CATEGORY
	ASSAY_ORGANSIM
	ASSAY_TAX_ID
	ASSAY_STRAIN

ASSAY TYPE	
FK	ASSAY_TYPE
	ASSAY_DESC

CONFIDENCE SCORE LOOKUP	
FK	CONFIDENCE_SCORE
	DESCRIPTION
	TARGET_MAPPING

ASSAY/VARIET	
FK	ASSAY_ID
FK	TID
FK	CONFIDENCE_SCORE
	MULTI
	COMPLEX
FK	RELATIONSHIP_TYPE
FK	CURATED_BY

CURATOR LOOKUP	
FK	CURATED_BY
	DESCRIPTION

RELATIONSHIP TYPE	
FK	RELATIONSHIP_TYPE
	RELATIONSHIP_DESC

TARGETS

TARGET DICTIONARY	
FK	TID
FK	TARGET_TYPE
	OR_SOURCE
	OR_VERSION
	PROTEIN_ACCESSION
	DESCRIPTION
	GENE_NAME
	PROT_NAME
	SYNONYMS
	KEYWORDS
	EC_NUMBER
	PROTEIN_SEQUENCE
	PROTEIN_RESIDUM
	TAX_ID
	ORGANISM
	TISSUE
	STRAIN
	CELL_LINE
FK	CHEMBL_ID

TARGET CLASS	
FK	TID
FK	TID
	S1
	S1
	S1
	S1
	S1
	S1
	S1
	S1
	S1
	S1

ORGANISM CLASS	
FK	ORG_ID
FK	TAX_ID
	S1
	S1
	S1

TARGET TYPE	
FK	TARGET_TYPE
	TARGET_DESC

VERSION	
FK	NAME
	CREATION_DATE
	COMMENTS

DRUG DATA

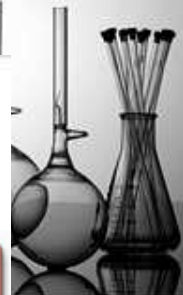
FABULATIONS	
FK	PRODUCT_ID
FK	INGREDIENT
FK	MOLREGNO

PRODUCTS	
FK	PRODUCT_ID
	TRADE_NAME
	ROUTE
	DOSEGE_FORM
	AD_TYPE
	ORAL
	PARENTERAL
	TOPICAL
	APPROVAL_DATE
	INFORMATION_SOURCE
	APPLICANT_FULL_NAME
	RESECTOR_COMPANY

ATC CLASSIFICATION	
FK	WHO_ID
	WHO_NAME
	LEVEL1
	LEVEL2
	LEVEL3
	LEVEL4
	LEVEL5
	LEVEL1_DESCRIPTION
	LEVEL2_DESCRIPTION
	LEVEL3_DESCRIPTION
	LEVEL4_DESCRIPTION

SEPMED DAILY DOSE	
FK	DDD_ID
FK	ATC_CODE
	DDD_MILLE
	DDD_UNITS
	DDD_ADMIN
	DDD_COMMENT





ChEMBL_14 Schema (27/06/12)

COMPOUNDS

COMPOUND RECORDS	
PK	RECORD_ID
FK	MOLEKIND
FK	MOLEKIND
FK	DOC_ID
	COMPOUND_KEY
	COMPOUND_NAME
FK	SRC_ID
FK	SRC_COMPOUND_ID

MOLEKULE HIERARCHY	
PK	MOLEKIND
FK	PARENT_MOLEKIND
FK	ACTIVE_MOLEKIND

MOLEKULE DICTIONARY	
PK	MOLEKIND
	PREF_NAME
FK	CHEMBL_ID
	SMX_PHASE
	STRUCTURE_TYPE
	MOLEKULE_TYPE
	FIRST_APPROVAL
	ORAL
	PARENTERAL
	TOPICAL
	BLACK_BOX_WARNING
	NATURAL_PRODUCT
	PRODRUG
	WORK_IN_PROGRESS
	THERAPEUTIC_FLAG
	CHEMBL_ID
	CHEMBL_PAK_ID

COMPOUND PROPERTIES	
PK	MOLEKIND
	MW_FREEMASS
	ALOGP
	TPSA
	MRD
	PSA
	RTS
	NUM_BOL_VIOLATIONS
	REEL_PAGE
	MED_CHEM_FRIENDLY
	ADD_MOST_APR4
	ADD_MOST_APR4
	ACD_LOGP
	ACD_LOGD
	MOLECULAR_SPECIES
	FULL_MWT

COMPOUND STRUCTURES	
FK	MOLEKIND
	MOLEFILE
	STANDARD_INCHI
	STANDARD_INCHI_KEY
	CANONICAL_SMILES
	MOLEFORMULA

MOLEKULE SYNONYMS	
PK	MOLEKIND
FK	SYNONYM
FK	SYN_TYPE
	RESEARCH_STEM

RESEARCH CODES	
FK	STEM
FK	COMPANY
	COUNTRY
	PREVIOUS_COMPANY

PROTEIN THERAPEUTICS	
FK	MOLEKIND
	PROTEIN_DESCRIPTION
	PROTEIN_SEQUENCE
	PROTEIN_SPECIES
	PROTEIN_SEQUENCE_LENGTH
	NATURE_PEPtide_SOURCE

SOURCE INFO

SOURCE	
FK	SRC_ID
	SRC_DESCRIPTION

SOURCE	
PK	DOC_ID
	JOURNAL
	YEAR
	VOLUME
	ISSUE
	FIRST_PAGE
	LAST_PAGE
	FURNISH_ID
	DOI
FK	CHEMBL_ID
	TITLE
	DOC_TYPE

CHEMBL_ID LOOKUP	
FK	CHEMBL_ID
	ENTITY_TYPE
	ENTITY_ID
	STATUS

EXPERIMENTAL DATA

LIGAND EFF	
FK	ACTIVITY_ID
	BEI
	SEI

ACTIVITIES	
FK	ACTIVITY_ID
FK	ASSAY_ID
FK	DOC_ID
FK	MOLEKIND
FK	RECORD_ID
	RELATION
	PUBLISHED_ACTIVITY_TYPE
	PUBLISHED_VALUE
	PUBLISHED_UNITS
	STANDARD_VALUE
	STANDARD_UNITS
	STANDARD_TYPE
	ACTIVITY_COMMENT

ASSAYS	
PK	ASSAY_ID
FK	SRC_ID
FK	DOC_ID
	DESCRIPTION
FK	SRC_ASSAY_ID
FK	CHEMBL_ID
	ASSAY_CATEGORY
	ASSAY_ORGANISM
	ASSAY_TAX_ID
	ASSAY_STRAIN

ASSAY TYPE	
PK	ASSAY_TYPE
	ASSAY_DESC

ASSAYTARGET	
FK	ASSAY_ID
FK	TID
FK	CONFIDENCE_SCORE
	MULTI
	COMPLEX
FK	RELATIONSHIP_TYPE
FK	CURATED_BY

CONFIDENCE SCORE LOOKUP	
FK	CONFIDENCE_SCORE
	DESCRIPTION
	TARGET_MAPPING

CURATOR LOOKUP	
FK	CURATED_BY
	DESCRIPTION

RELATIONSHIP TYPE	
FK	RELATIONSHIP_TYPE
	RELATIONSHIP_DESC

TARGETS

TARGET DICTIONARY	
FK	TID
FK	TARGET_TYPE
	OR_SOURCE
	OR_VERSION
	PROTEIN_ACCESSION
	GENE_NAME
	PREF_NAME
	SYNONYMS
	KEYWORDS
	EC_NUMBER
	PROTEIN_SEQUENCE
	PROTEIN_MOSMUM
	TAX_ID
	ORGANISM
	TISSU
	STRAIN
	CELL_LINE
FK	CHEMBL_ID

TARGET CLASS	
FK	TID
FK	TID
	L1
	L2
	L3
	L4
	L5
	L6
	L7
	L8
	TARGET_CLASSIFICATION

ORGANISM CLASS	
FK	DOC_ID
FK	TAX_ID
	L1
	L2
	L3

TARGET TYPE	
FK	TARGET_TYPE
	TARGET_DESC

VERSION	
FK	NAME
	CREATION_DATE
	COMMENTS

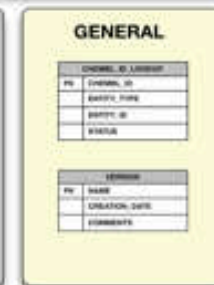
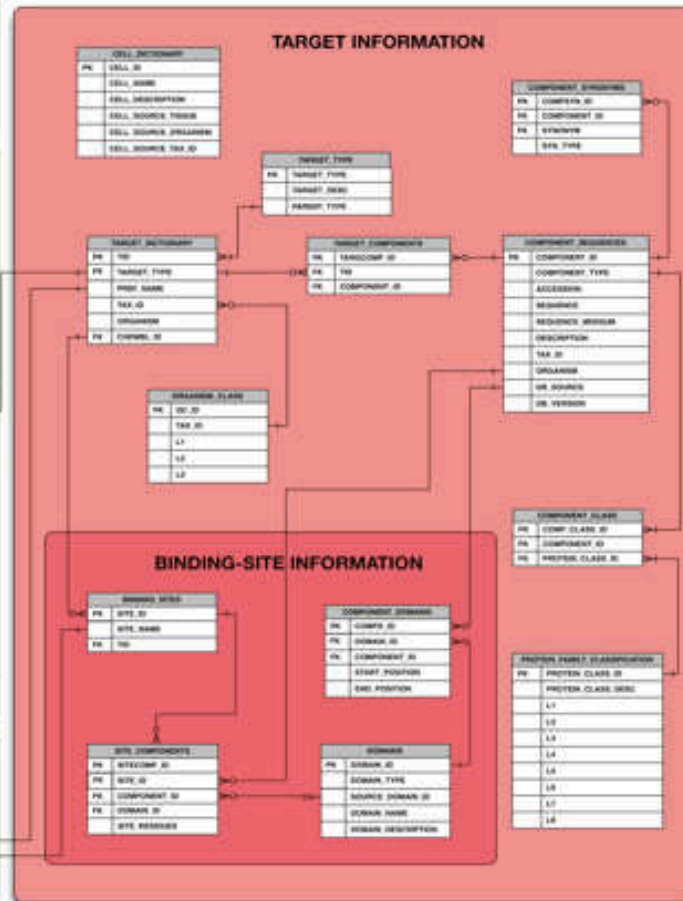
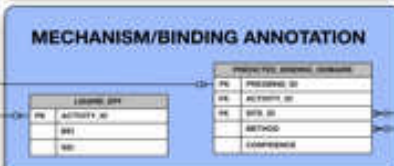
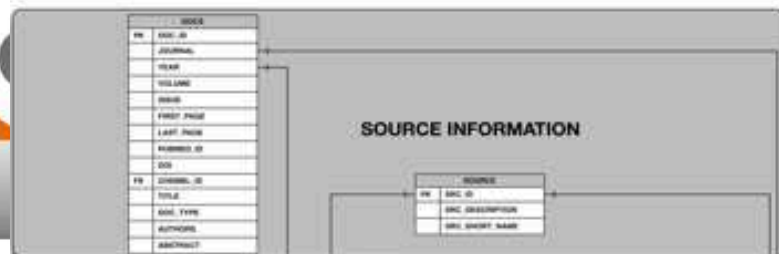
DRUG DATA

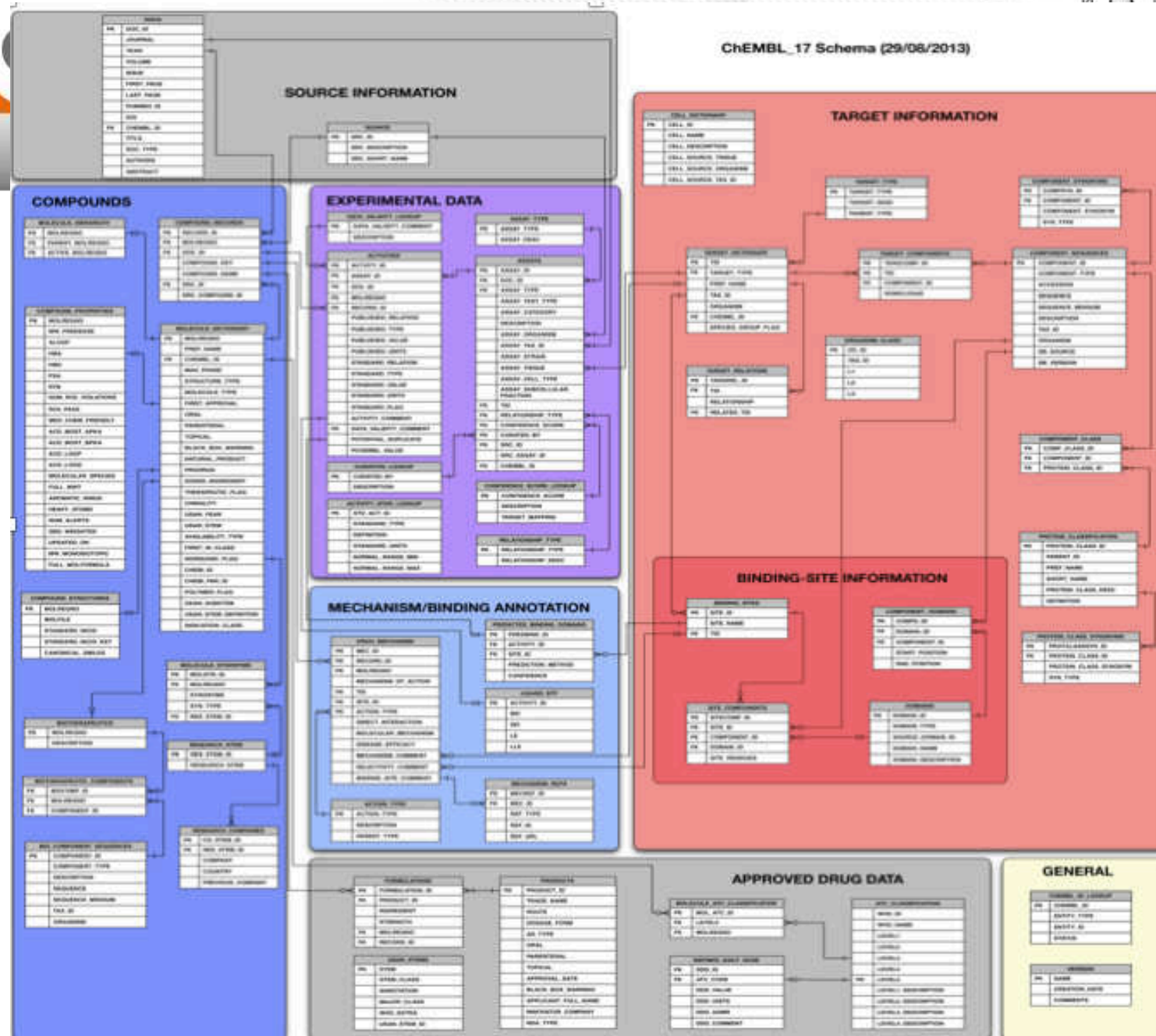
FURNISHMENTS	
FK	PRODUCT_ID
FK	INGREDIENT
FK	STRENGTH
FK	MOLEKIND

PRODUCTS	
FK	PRODUCT_ID
	TRADE_NAME
	ROUTE
	DOSEAGE_FORM
	AD_TYPE
	ORAL
	PARENTERAL
	TOPICAL
	APPROVAL_DATE
	INFORMATION_SOURCE
	APPLICANT_FULL_NAME
	INNOVATOR_COMPANY

ATC CLASSIFICATION	
FK	WHO_ID
	WHO_NAME
	LEVEL1
	LEVEL2
	LEVEL3
	LEVEL4
	LEVEL5
FK	LEVEL5
	LEVEL1_DESCRIPTION
	LEVEL2_DESCRIPTION
	LEVEL3_DESCRIPTION
	LEVEL4_DESCRIPTION
	LEVEL5_DESCRIPTION

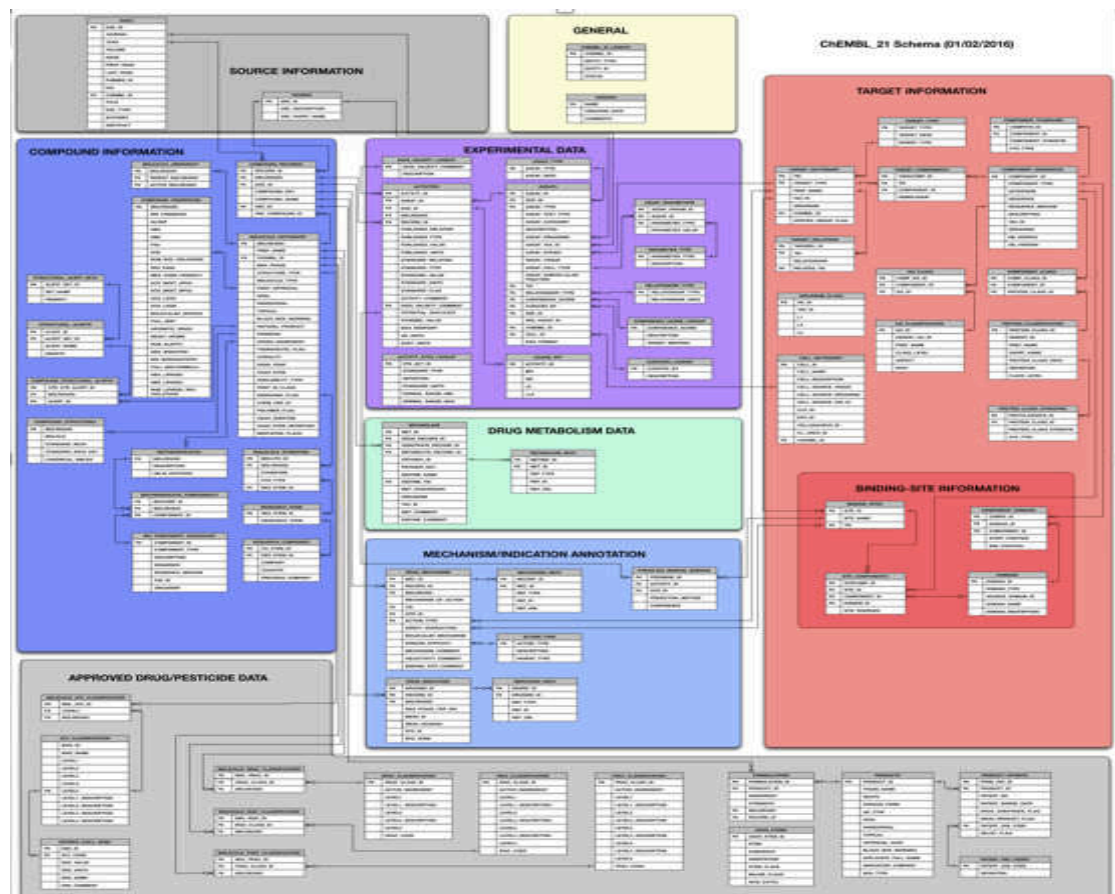
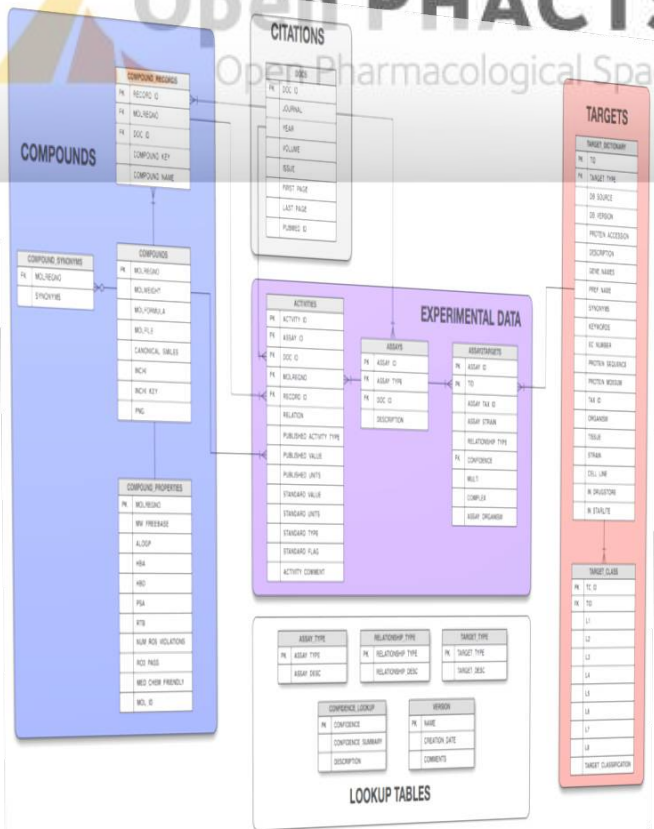
SCHEDULED DAILY DOSE	
FK	DDO_ID
FK	ATC_CODE
	DDO_VALUE
	DDO_UNITS
	DDO_ADMN
	DDO_COMMENT





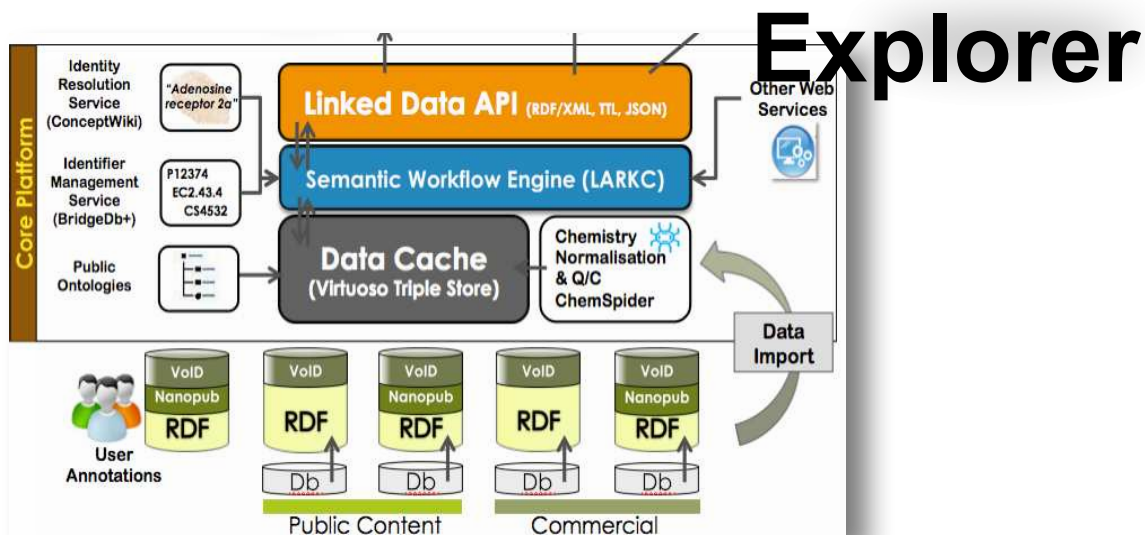
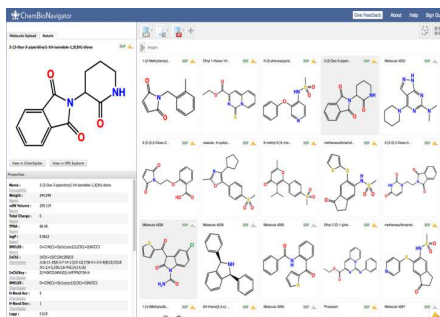
Open PHACTS

Open Pharmacological Space



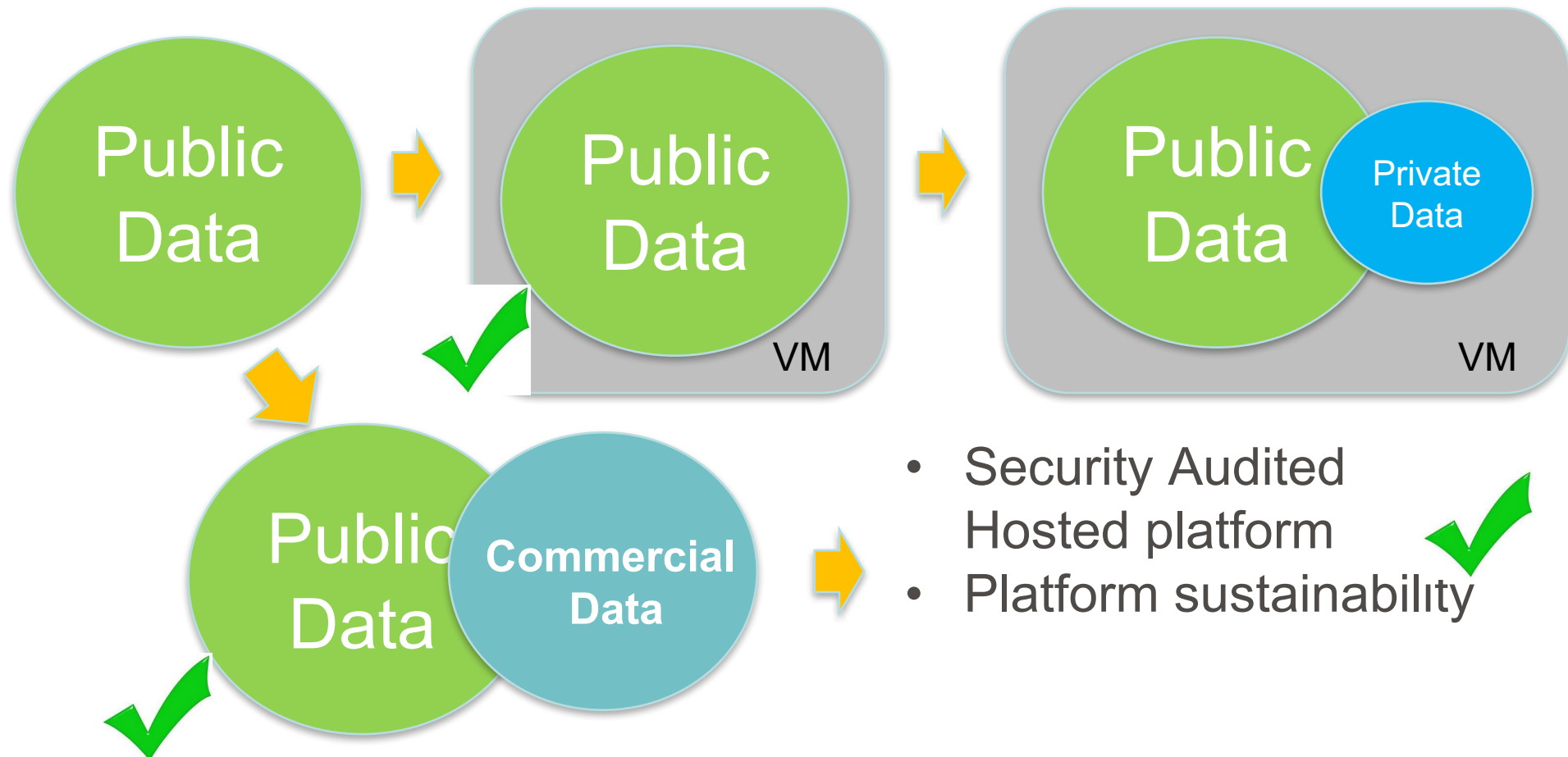


Open PHACTS Expanding EcoSystem





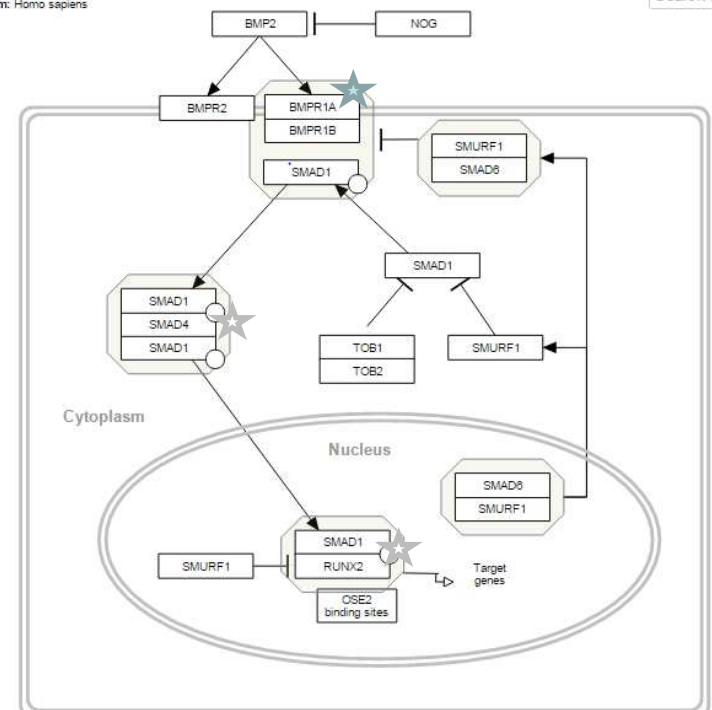
Open PHACTS Evolution - Platform



For a given target and pathway, what proteins/genes are upstream or downstream?

- ✓ WikiPathway interaction editing standardised
- ✓ WikiPathways RDF Data Model Enhanced
- ✓ Signalling and Metabolic Pathways/Interactions handled
- ✓ Higher Confidence Pathways loaded into Open PHACTS
- ✓ Existing pathway API calls rebuilt for the new data model
- ✓ Up or downstream API Call (in development)

Title: BMP Signalling and Regulation
Organism: Homo sapiens



QU: What is downstream of BMPR1A in the BMP pathway?

1 node

A: SMAD1, SMAD4

QU: What is downstream of SMAD1?

1 node

A: SMAD1, SMAD4, RUNX2

QU: What is downstream of SMAD4?

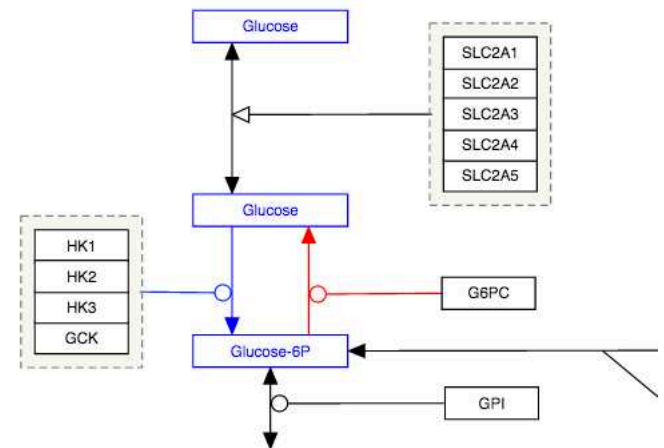
1 node

A: SMAD1, RUNX2



Modelling three different interaction standards:

- Plain drawn connectors (arrows, lines)
- MIM (Molecular Interaction Map)
In WikiPathways Applet
- SBGN
PathVisio plugin
(with WikiPathways plugin)
From Reactome





PATENTS

SureChEMBL
Open Patent Data



- ✦ SureChEMBL is perhaps the only freely available, large-scale, comprehensive and *live* resource of chemistry extracted from the patent literature
- ✦ SureChEMBL automatically annotates, normalises and indexes chemistry found in the full text, images and attachments (i.e. mol files) of patent documents

With Open PHACTS support:

- ✦ Chemical and biological annotations will be semantically integrated
 - biological annotations added directly to the patent documents using Termite



PATENTS

SureChEMBL
Open Patent Data

EMBL-EBI 

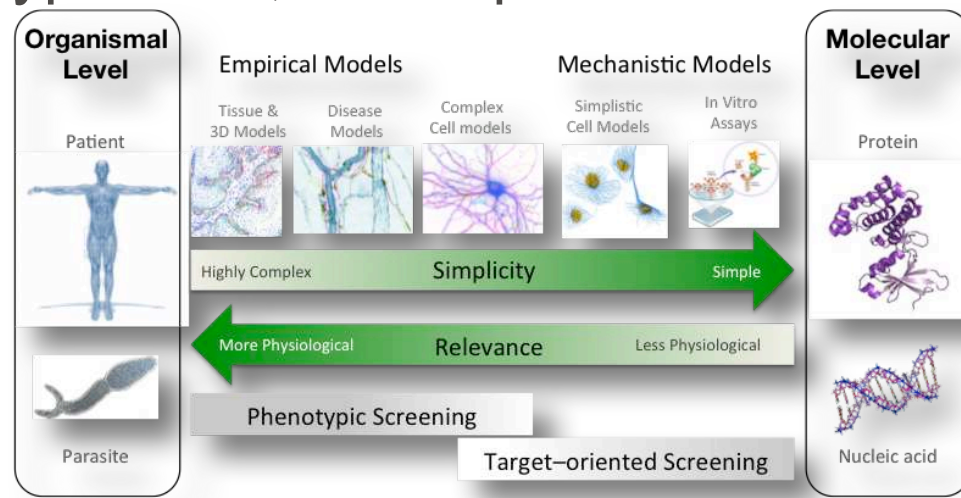
What can I do now with Patents?

- ✦ Through linked Queries:
 - Patent to Compounds
 - Compounds to Patents
 - Disease to Patent & reverse
 - Target to Patent & reverse
- ✦ As well as being possible to link the data further
- ✦ Extend based on user feedback



PHENOTYPIC SCREENING

- ❖ Science meeting Feb 2015 defining requirements, pre-competitive group plus other IMI projects
- ❖ NCATS Phenotypic data will be available
- ❖ Full data set from the Lilly sponsored work
- ❖ Integrated into Open PHACTS
- ❖ Further workflows on Phenotypic data, developed with partners





Open PHACTS Foundation engaged in:

 **EUTOXRISK**



The Open PHACTS Foundation

OPF is a not-for-profit membership organisation, supporting the Open PHACTS Discovery Platform:

A sustainable, open, vibrant and interoperable information infrastructure for applied life science research and development.

To reduce the barriers to drug discovery in industry, academia and for small businesses, the Open PHACTS Discovery Platform provides tools and services to interact with multiple integrated and publicly available data sources. To integrate this data, extensive cross-referencing of scientific concepts is needed across all databases.

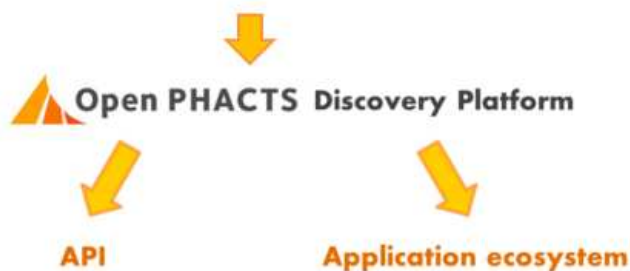
The Open PHACTS Foundation ensures the sustainability of the Open PHACTS Discovery Platform infrastructure and acts as a hub for relevant scientific research and development.



ChEMBL



The free chemical database

**DRUGBANK**
Open Data Drug & Drug Target DatabaseWIKIPATHWAYS
Pathways for the People

Key Resources

[Open PHACTS API](#)[Open PHACTS Repository](#)

Subscribe to the Foundation Newsletter

[Subscribe](#)

Contact us

Email:
info@openphactsfoundation.org

Twitter: [@Open PHACTS](#)



Open PHACTS Foundation Routes to Access

Access Route	Open API services	Unlimited API services	Unlimited API, RDF and Link sets	Open PHACTS Virtual Machine
Full OPF Member	✓	✓	✓	✓
Licensor/ Reseller*	✓	✓	✓	
Licensor (Own Use)	✓	✓	✓	
High volume API Licensor	✓	✓		
Open Access API Consumer	✓			
Open Data Non-commercial	✓		**	

*3rd parties must have own agreement with OPF

** talk to us for collaborative proposals – non commercial use



Come and collaborate



- ✦ New projects
- ✦ Improve our code and services
- ✦ Open Innovation projects
- ✦ Webinars
- ✦ New ideas for data services and workflows

Acknowledgements



GlaxoSmithKline – Coordinator
Universität Wien – Managing entity
Technical University of Denmark
University of Hamburg, Center for
Bioinformatics
BioSolveIT GmbH
Consorti Mar Parc de Salut de Barcelona
Leiden University Medical Centre
Royal Society of Chemistry
Vrije Universiteit Amsterdam
Novartis
Merck Serono
H. Lundbeck A/S
Eli Lilly
Netherlands Bioinformatics Centre
Swiss Institute of Bioinformatics
ConnectedDiscovery
EMBL-European Bioinformatics Institute
Janssen Esteve Almirall
OpenLink Scibite
The Open PHACTS Foundation
Spanish National Cancer Research Centre
University of Manchester
Maastricht University
Aqnowledge
University of Santiago de Compostela
Rheinische Friedrich-Wilhelms-Universität
Bonn
AstraZeneca
Pfizer



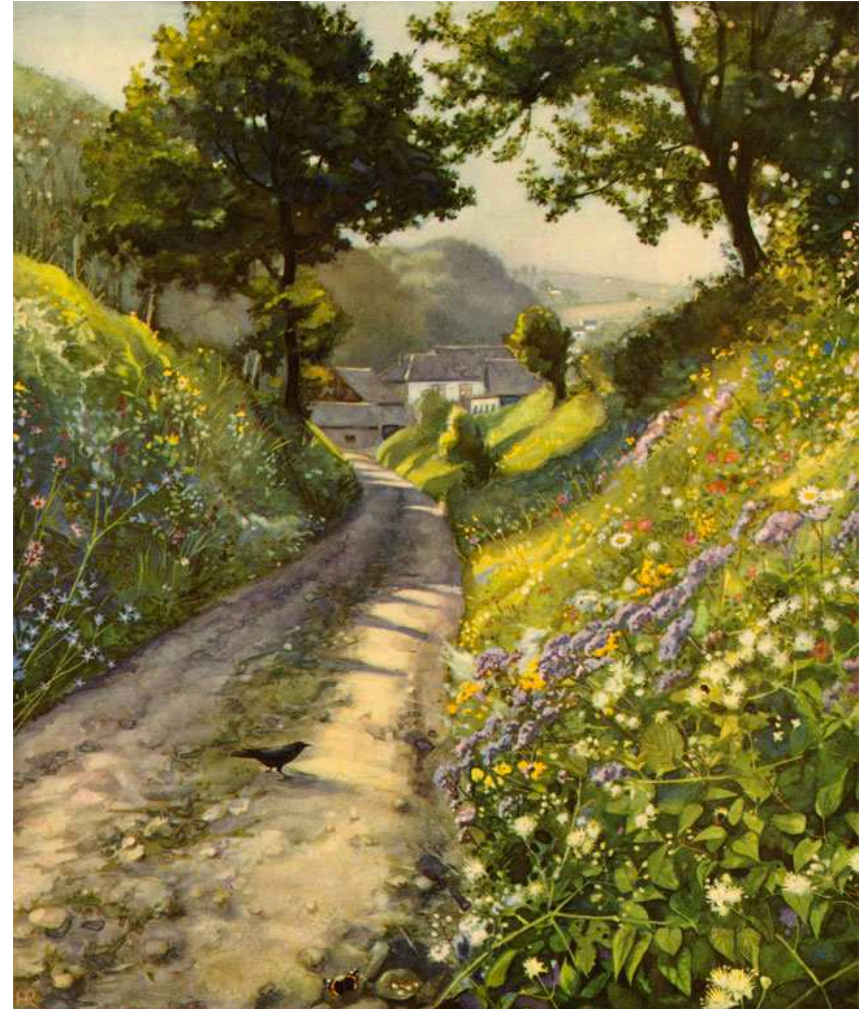
info@openphactsfoundation.org



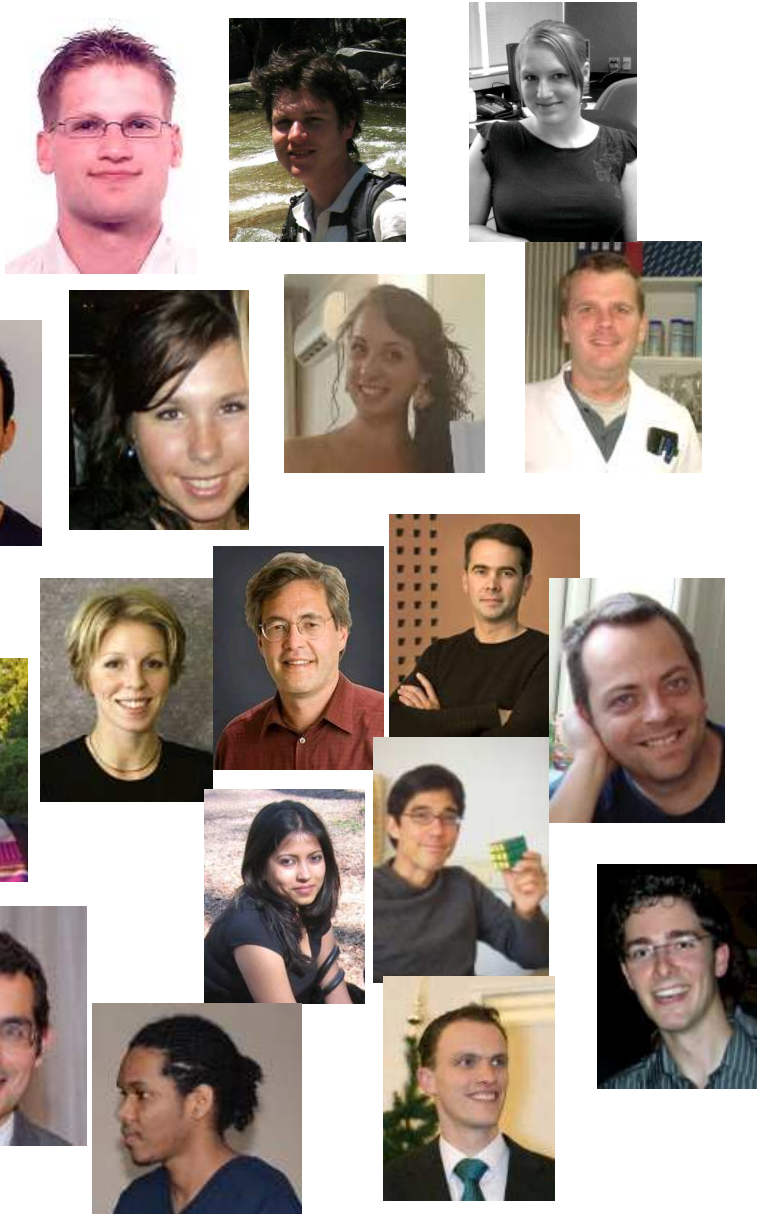
@Open_PHACTS

Using biological pathways as hubs in networks of things we know about biology

Prof. Chris Evelo
Department Bioinformatics – BiGCaT
Maastricht University
@Chris_Evelo



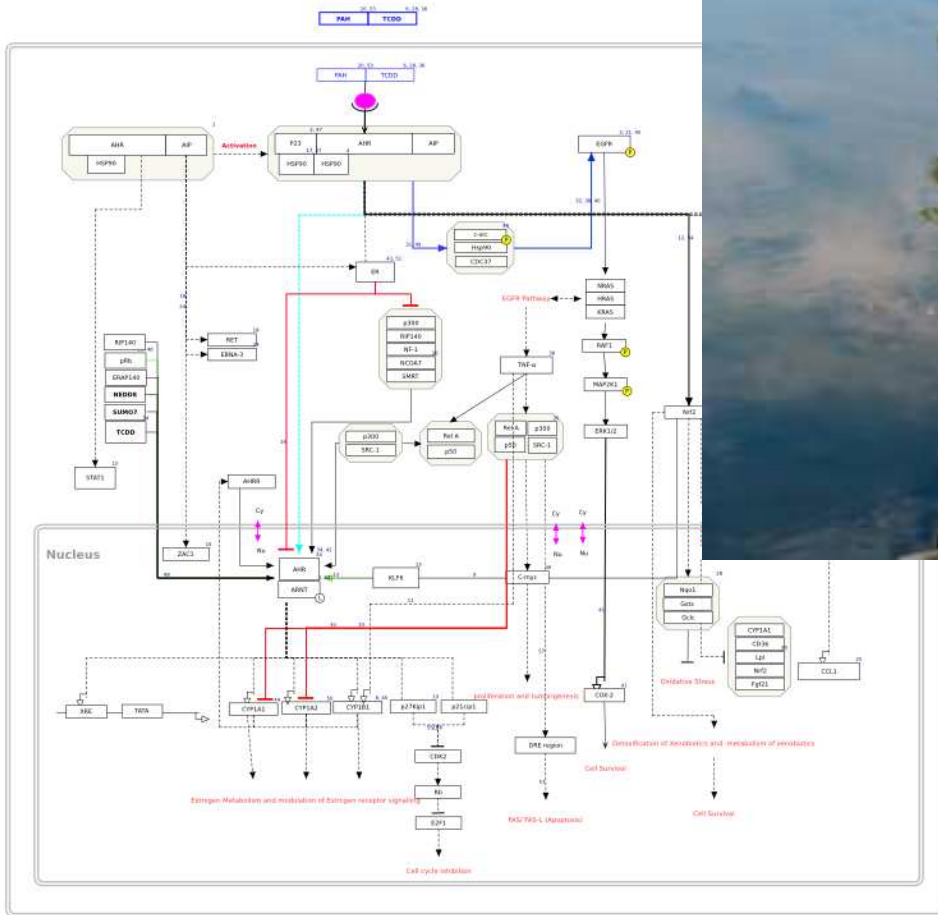
Thanks!



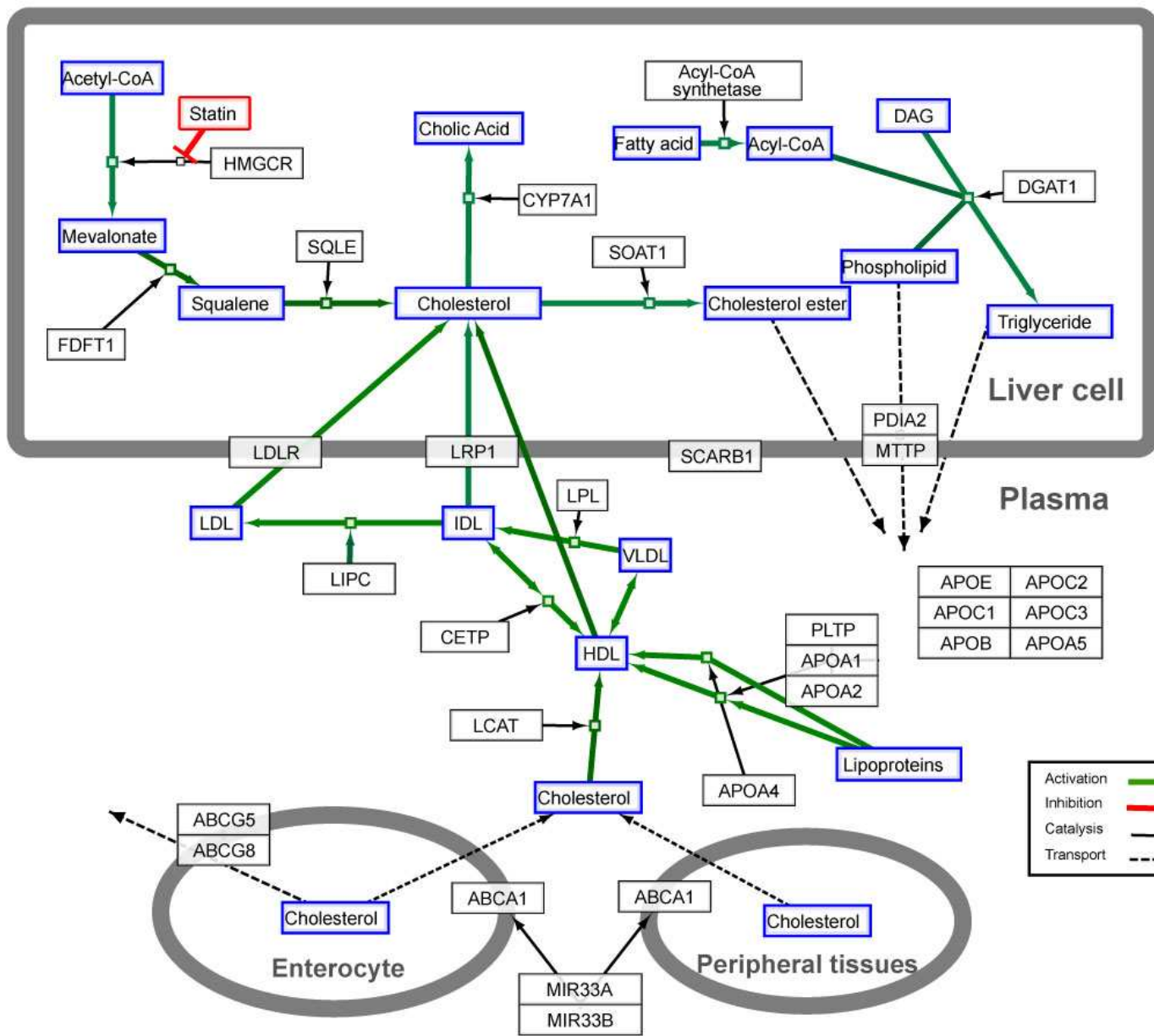
Funding. IOP, NBIC, NuGO, NCSB. Transnational University. NuGO and Microgenet, Open PHACTS. Agilent thought leader grant.

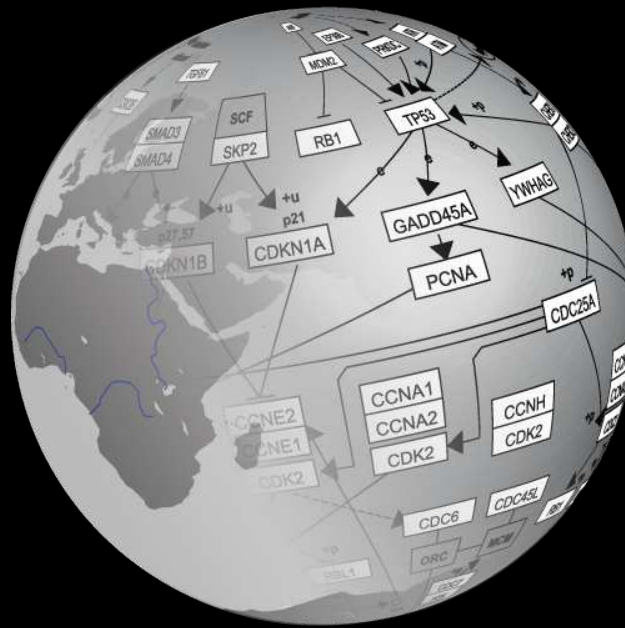
Paths in nature

Tiller, And Hydroxide Receptor
Dependent, more details



↑ Translocation
 P2Z Plasma membrane
 Cy Cytoplasm
 EC Extracellular
 ER Endosome
 ER Endoplasmic reticulum
 GO Golgi apparatus
 M Mitochondria
 Nucleus





WIKIPATHWAYS AND HOW TO CHANGE THE WORLD (OR AT LEAST A SMALL CORNER OF THE WORLD)

Chris Evelo
Alexander Pico

Maastricht University
Gladstone Institutes

The Netherlands
San Francisco, CA

WikiPathways:Statistics

This page contains several statistics and graphs to indicate

Summary statistics

Last update: 2014/10/31

Number of pathways

- Public pathways: **1910**
- Private pathways: **8**

Number of registered users

- Non-editing users: **9008**
- At least 1 edit: **583**
- At least 1 edit (excluding test/tutorial pathways): **425**

Number of edits

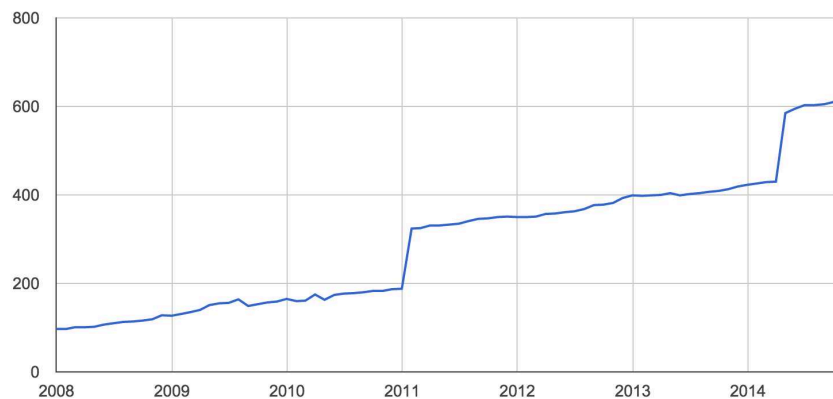
- User edits: **18668**
- User edits (excluding test/tutorial pathways): **16837**
- Bot edits: **11065**

Content growth

Pathways per species

The following graph shows the number of pathways over time and gives an indication of the growth of the content. Pathways marked for testing/tutorials are excluded from this graph.

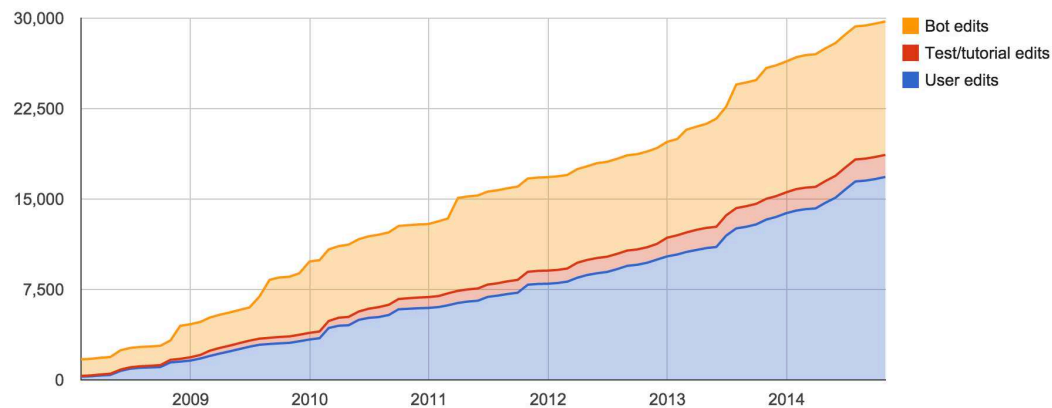
Show pathway counts for



Cumulative number of edits

The cumulative number of edits, divided into three categories:

- User edits: edits done by users, excluding edits to pathways marked as test/tutorial.
- Test/tutorial edits: edits to pathways marked as test/tutorial.
- Bot edits: edits done by bots (automated scripts for maintenance and batch upload).



WikiPathways, a web of communities

portal discussion view source history

Portal:ExRNA

Welcome to the ExRNA Pathway Portal

This portal highlights pathway content relevant to the extracellular RNA research community.

The Common Fund strategic planning process has developed a program on Extracellular RNA Communication. This program aims to discover fundamental biological principles about the mechanisms of exRNA generation, secretion, and transport; to identify and develop a catalog of exRNA found in normal human body fluids; and to investigate the potential for using exRNAs in the clinic as therapeutic molecules or biomarkers of disease. To find out more about the program, visit exrna.org.

[Pathway Curation](#) [Featured Pathways](#)

To the right you'll see a rotating display of featured pathways related to external RNA biology. Where did these pathways come from? They came from people like you! The editor added and added to using the pathway drawing and annotation tools

portal discussion view source history

Portal:Reactome

Reactome Pathways

Welcome to the Reactome Portal

The Reactome Portal at WikiPathways maintains pathway content periodically synced with the Reactome collection.

Reactome is an open-source, open access, manually curated and peer-reviewed pathway database. Pathway annotations are authored by expert biologists, in collaboration with Reactome editorial staff. The rationale behind Reactome is to convey the rich information in the visual representations of biological pathways familiar from textbooks and articles in a detailed, computationally accessible format. The core unit of the Reactome data model is the reaction. Entities (nucleic acids, proteins, complexes and small molecules) participating in reactions form causal chain of events or pathways.

Learn more about the Reactome Website and Pathway Analysis tools

- Read About Reactome
- See our User guide

[Reactome website](#)

Community Curation Initiative

In collaboration with WikiPathways, we have established a custom pathway data exchange pipeline. Reactome pathways are now incorporated into WikiPathways, an open, collaborative platform dedicated to the curation of biological pathways. We encourage you to edit, curate and create pathway content here. The Reactome team will periodically assess and selectively incorporate changes back into our centralized collection.

Feel free to propose edits on pathway discussion pages or [contact us](#) directly with any questions.

News and Notes

All Pathways

Apoptotic execution phase (Homo sapiens)

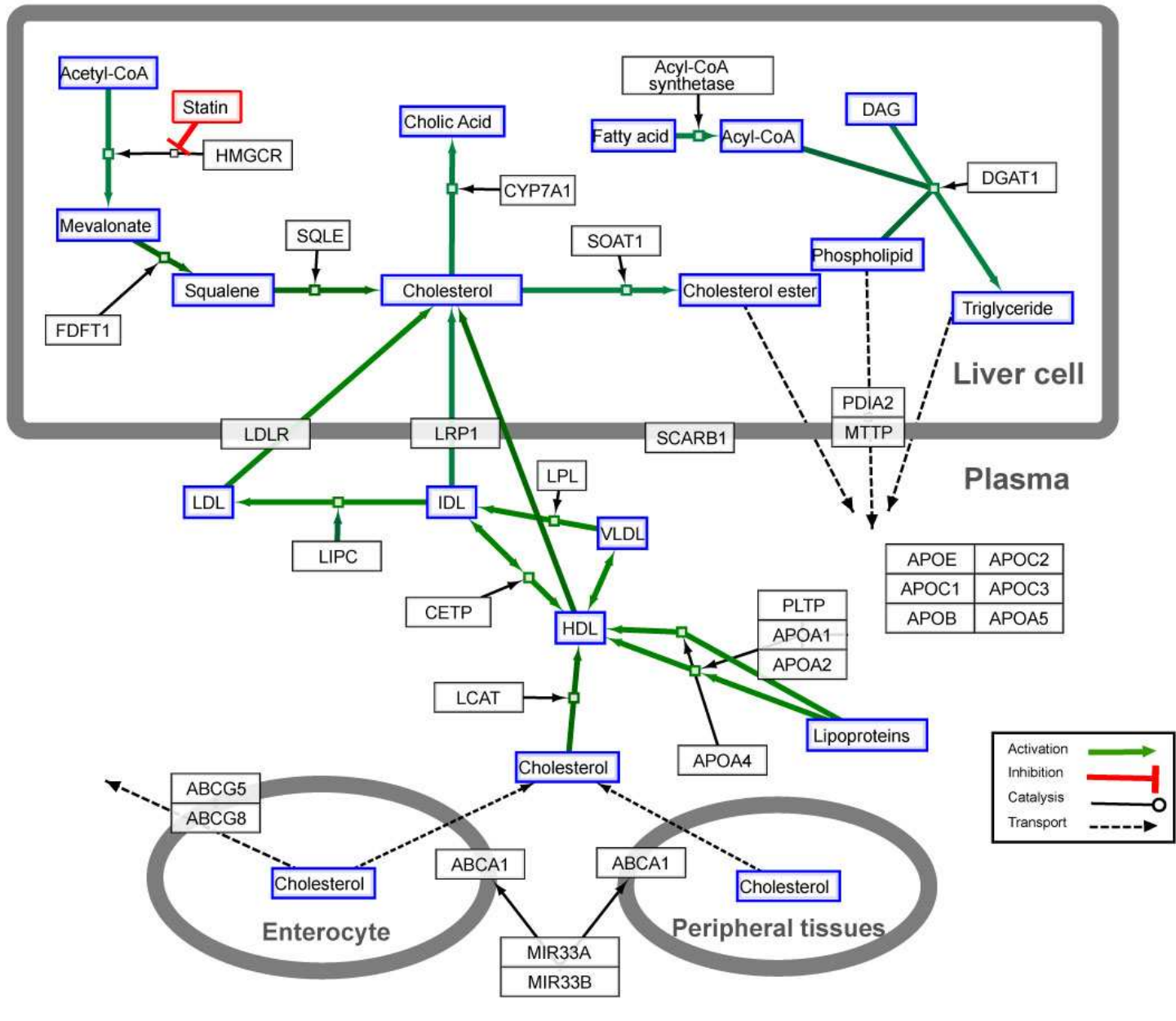
Apoptotic execution phase

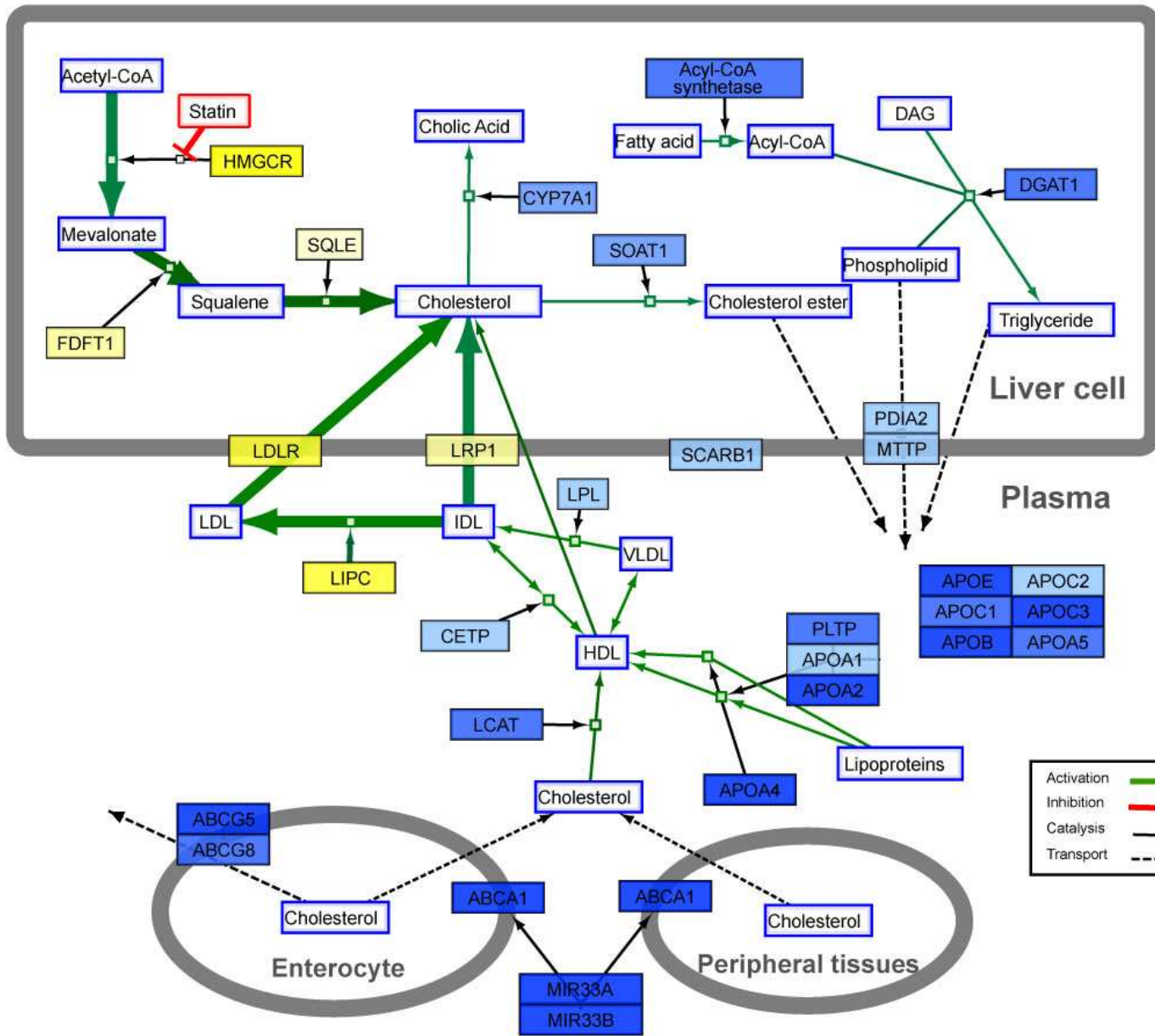
[View all Pathways](#) for this Portal

Download all Reactome pathways by species:

- Homo sapiens

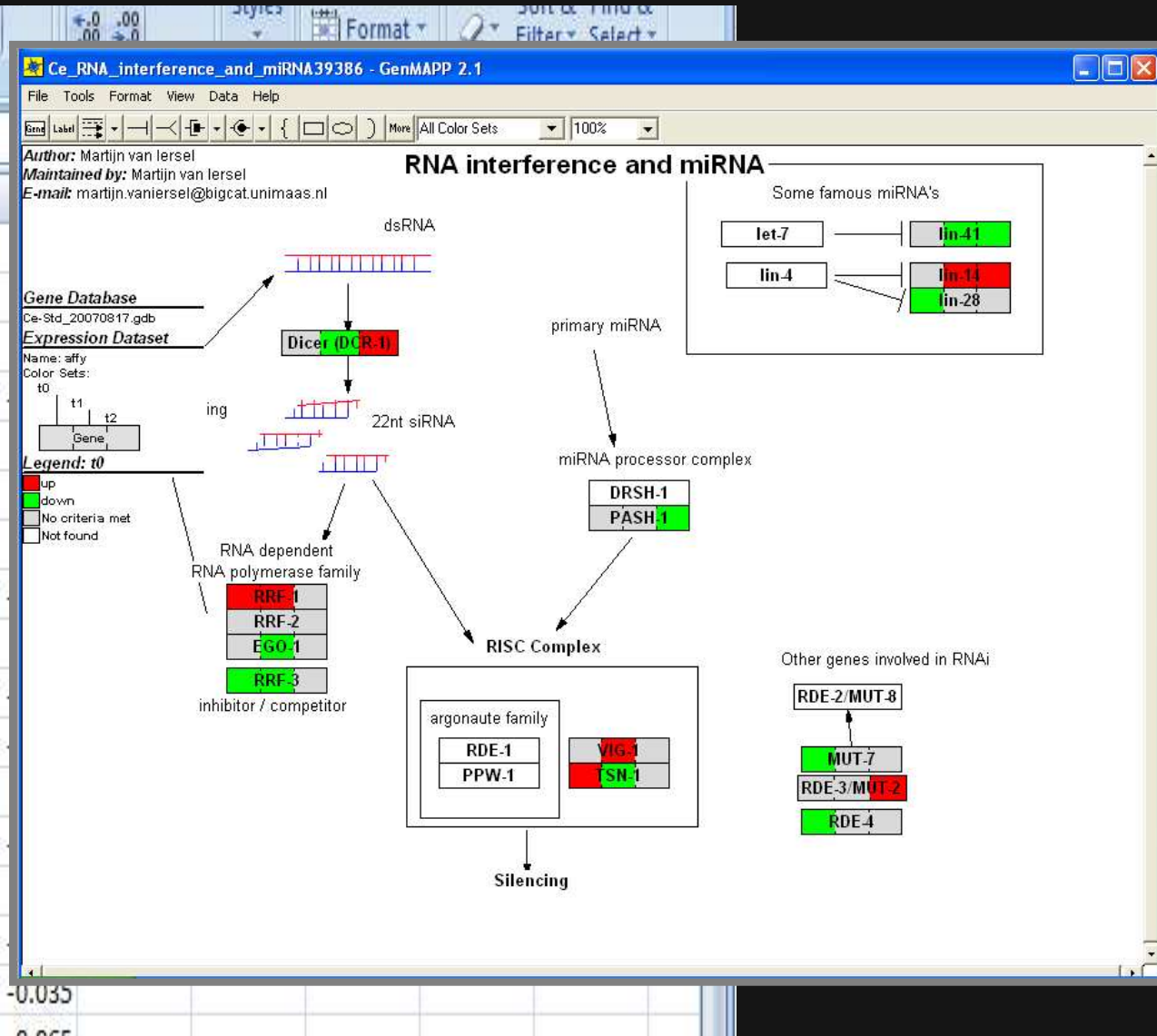
- [CIRM portal](#)
- [ExRNA portal](#)
- [GenMAPP portal](#)
- [Micronutrient portal](#)
- [NetPath portal](#)
- [Plants portal](#)
- [Reactome portal](#)
- [WormBase portal](#)



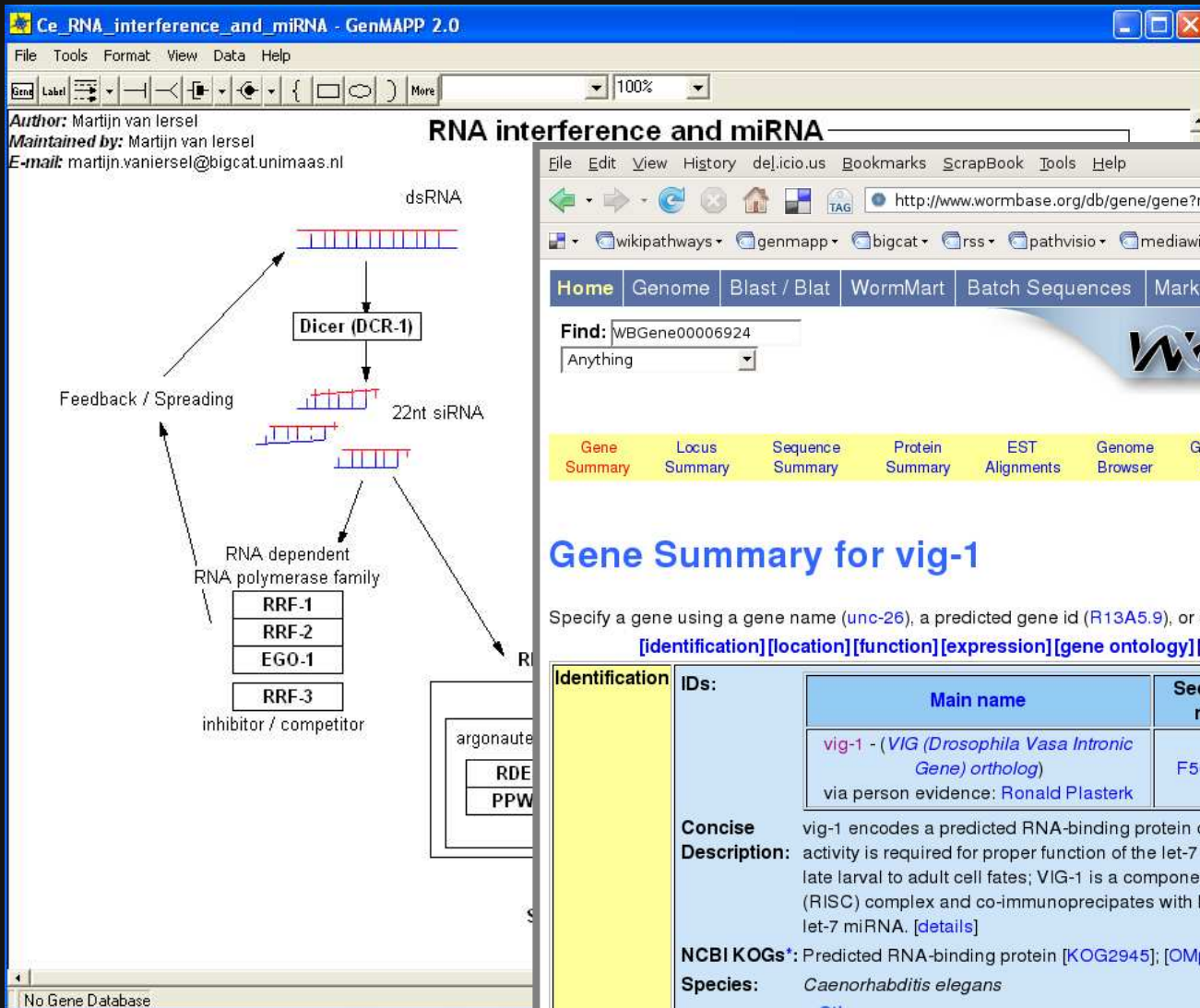


How to do data visualization?

A20					
	A	B	D	E	
1	ID	System	t0	t1	t2
2	177140_s_X		0.693	-2.856	
3	186328_at_X		1.327	3.497	
4	192151_at_X		3.863	-3.457	
5	187703_s_X		1.16	-3.352	
6	193913_s_X		-0.435	3.356	
7	173437_at_X		-2.36	-1.633	
8	173452_s_X		-2.43	0.035	
9	no affy	X	-1.878	2.211	
10	178221_at_X		-1.758	-3.223	
11	173832_at_X		-3.386	1.695	
12	184510_at_X		0.109	-1.045	
13	192371_s_X		-3.075	-0.404	
14	172670_x_X		3.642	3.428	
15	176778_s_X		-0.112	1.072	
16	175832_s_X		1.875	-3.817	
17	178112_s_X		2.247	2.247	0.035
18	178112_s_X		2.247	2.247	0.035



Connect to Genome Databases



File Edit View History del.icio.us Bookmarks ScrapBook Tools Help

http://www.wormbase.org/db/gene/gene?name=WBGene0006924

Home Genome Blast / Blat WormMart Batch Sequences Markers Genetic Maps Submit Searches Site Map

Find: WBGene0006924

Anything

Gene Summary Locus Summary Sequence Summary Protein Summary EST Alignments Genome Browser Genetic Map Nearby Genes Tree Display Bibliography XML Schema Acedb Image

Gene Summary for vig-1

Specify a gene using a gene name (*unc-26*), a predicted gene id (R13A5.9), or a protein ID (CE02711)

[identification][location][function][expression][gene ontology][alleles][similarities][reagents][bibliography]

Identification	IDs:	Main name	Sequence name	Other name(s)	WB Gene ID
		vig-1 - (<i>VIG (Drosophila Vasa Intronic Gene) ortholog</i>) via person evidence: Ronald Plasterk	F56D12.5	2B613 (inferred automatically)	WBGene0006924

Concise Description: vig-1 encodes a predicted RNA-binding protein orthologous to *Drosophila* VIG (*Vasa Intronic Gene*); vig-1 activity is required for proper function of the let-7 miRNA in vivo and thus, for regulating the transition from late larval to adult cell fates; VIG-1 is a component of the 250 kDa RNA-induced silencing complex (RISC) complex and co-immunoprecipitates with both TSN-1, the *C. elegans* Tudor-SN ortholog, and the let-7 miRNA. [details]

NCBI KOGs*: Predicted RNA-binding protein [KOG2945]; [OMpre_WH000977]

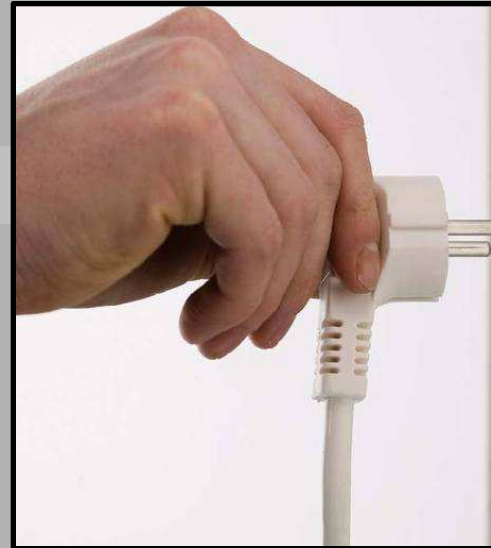
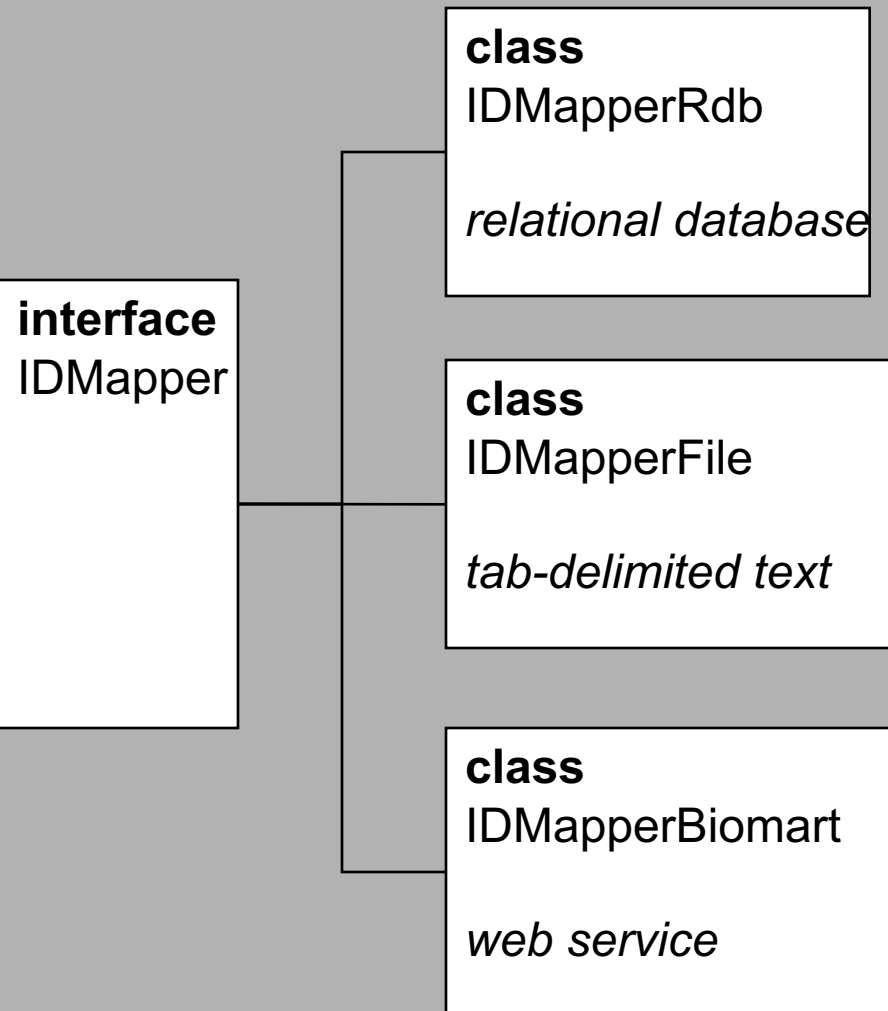
Species: *Caenorhabditis elegans*

Other sequences: Other sequences

NCBI: [AceView: 2B613]

Done 4472 1087.thomaskelder@gmail.com

BridgeDb: Abstraction Layer



Backpages link to databases

WikiPathways editor - cytochrome P450

Zoom: 100%

Title: cytochrome P450
Organism: Homo sapiens

Cytochrome P450 genes

CYP1A1	CYP3A4	CYP11A1
CYP1A2	CYP3A5	CYP11B1
CYP1B1	CYP3A7	CYP11B2
CYP2A6	CYP3A43	CYP17A1
CYP2A7	CYP4A11	CYP19A1
CYP2A13	CYP4A22	CYP20A1
CYP2B6	CYP4B1	CYP21A2
CYP2C8	CYP4F2	CYP24A1
CYP2C9	CYP4F3	CYP26A1
CYP2C18	CYP4F8	CYP26A1
CYP2C19	CYP4F11	CYP26B1
CYP2D6	CYP4F12	CYP26C1
CYP2E1	CYP4F22	CYP27A1
CYP2F1	CYP4V2	CYP27B1
CYP2G1P	CYP4X1	CYP27C1
CYP2J2	CYP4Z1	CYP39A1
CYP2R1	CYP51A1	CYP46A1
CYP2S1	CYP7A1	CYP51A1
CYP2U1	CYP7B1	
CYP2W1	CYP8B1	

GeneProduct information

Gene ID:	ENSG00000100197
Gene Symbol:	CYP2D6
Description:	cytochrome P450, family 2, subfamily D, polypeptide 6 [Source:HGNC Symbol;Acc:2625]
Chr:	22

Cross references

- [A_23_P143734](#), Agilent
- [A_23_P155123](#), Agilent
- [ENSG00000100197](#), Ensembl Human
- [1565](#), Gene Wiki
- [CYP2D6](#), HGNC
- [0006350750](#), Illumina
- [ILMN_1740648](#), Illumina
- [ILMN_2383124](#), Illumina
- [IP100433508](#), IPI
- [IP100456699](#), IPI
- [IP100868752](#), IPI
- [IP100873186](#), IPI
- [IP100873251](#), IPI
- [IP100943274](#), IPI
- [1565](#), Entrez Gene
- [124030](#), OMIM
- [608902](#), OMIM
- [2F9Q](#), PDB
- [NM_000106](#), RefSeq
- [NM_001025161](#), RefSeq
- [NP_000097](#), RefSeq
- [NP_001020332](#), RefSeq
- [P10635](#), Uniprot/TrEMBL
- [Q007T9](#), Uniprot/TrEMBL

You could do this for gene lists

The screenshot shows the WikiPathways editor interface for the 'cytochrome P450' pathway in Homo sapiens. A central window displays a grid of 30 Cytochrome P450 genes, with CYP2D6 highlighted. To the right, a 'GeneProduct information' panel provides details for CYP2D6, including its Gene ID (ENSG00000100197), Gene Symbol (CYP2D6), Description (cytochrome P450, family 2, subfamily D, polypeptide 6 [Source:HGNC Symbol;Acc:2625]), and Chromosome (22). Below this, a 'Cross references' section lists various database identifiers for CYP2D6, such as Agilent, Ensembl, Gene Wiki, HGNC, Illumina, IPI, OMIM, and RefSeq.

Cytochrome P450 genes

CYP1A1	CYP3A4	CYP11A1
CYP1A2	CYP3A5	CYP11B1
CYP1B1	CYP3A7	CYP11B2
CYP2A6	CYP3A43	CYP17A1
CYP2A7	CYP4A11	CYP19A1
CYP2A13	CYP4A22	CYP20A1
CYP2B6	CYP4B1	CYP21A2
CYP2C8	CYP4F2	CYP24A1
CYP2C9	CYP4F3	CYP26A1
CYP2C18	CYP4F8	CYP26B1
CYP2C19	CYP4F11	CYP26C1
CYP2D6	CYP4F12	CYP27A1
CYP2E1	CYP4F22	CYP27B1
CYP2F1	CYP4V2	CYP27C1
CYP2G1P	CYP4X1	CYP39A1
CYP2J2	CYP4Z1	CYP46A1
CYP2R1	CYP51A1	CYP51A1
CYP2S1	CYP7A1	
CYP2U1	CYP7B1	
CYP2W1	CYP8B1	

GeneProduct information

Gene ID:	ENSG00000100197
Gene Symbol:	CYP2D6
Description:	cytochrome P450, family 2, subfamily D, polypeptide 6 [Source:HGNC Symbol;Acc:2625]
Chr:	22

Cross references

- [A_23_P143734](#), Agilent
- [A_23_P155123](#), Agilent
- [ENSG00000100197](#), Ensembl Human
- [1565](#), Gene Wiki
- [CYP2D6](#), HGNC
- [0006350750](#), Illumina
- [ILMN_1740648](#), Illumina
- [ILMN_2383124](#), Illumina
- [IPI00433508](#), IPI
- [IPI00456699](#), IPI
- [IPI00868752](#), IPI
- [IPI00873166](#), IPI
- [IPI00873251](#), IPI
- [IPI00943274](#), IPI
- [1565](#), Entrez Gene
- [124030](#), OMIM
- [608902](#), OMIM
- [2F9Q](#), PDB
- [NM_000106](#), RefSeq
- [NM_001025161](#), RefSeq
- [NP_000097](#), RefSeq
- [NP_001020332](#), RefSeq
- [P10635](#), Uniprot/TrEMBL
- [Q007T9](#), Uniprot/TrEMBL

Regulator Interaction in MiPaSt PathVisio plugin

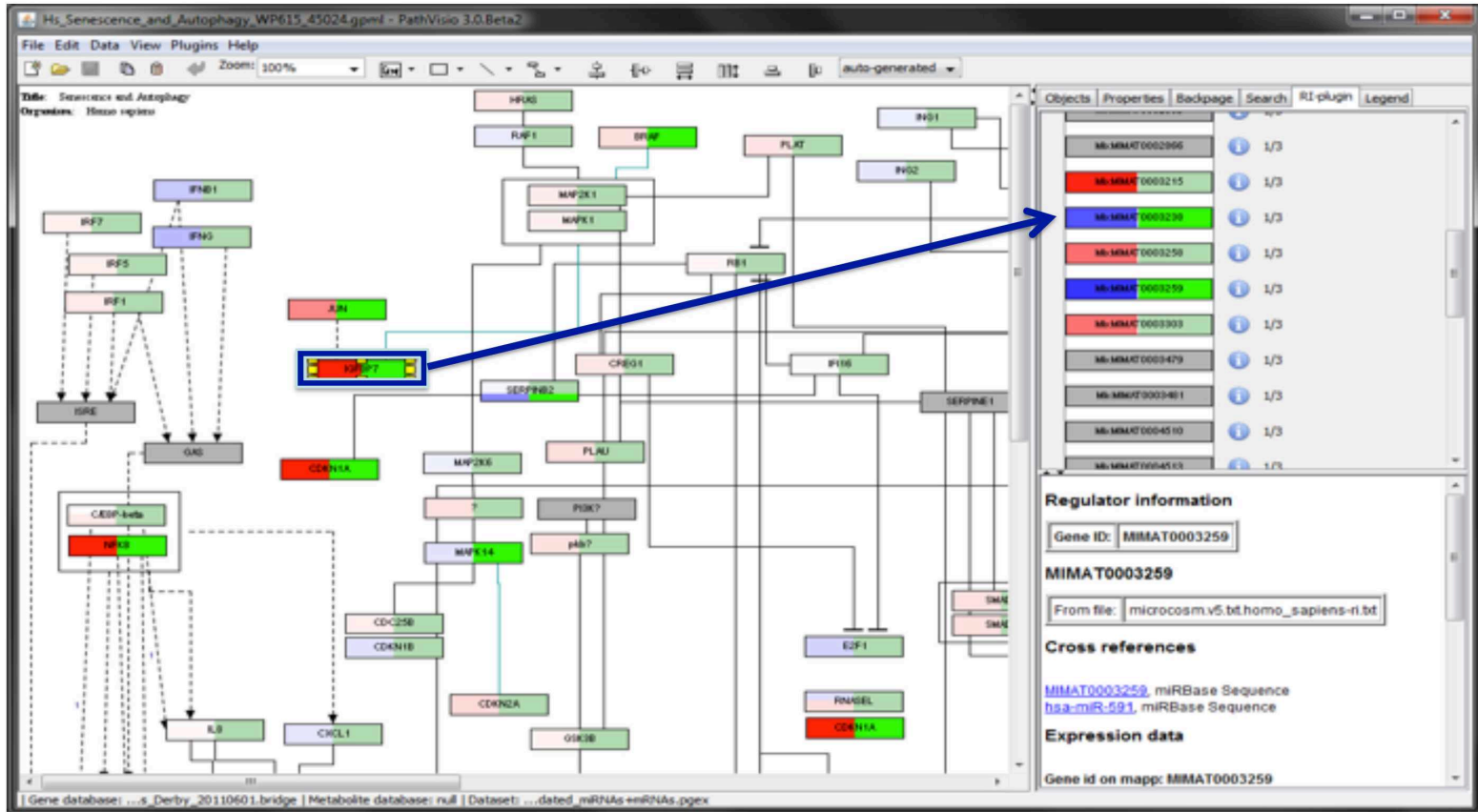


Figure 1: The regulatory interaction plugin in action. For a selected gene, indicated by the blue box, and its mapped miRNA interactions are shown in the side panel. The side panel also shows the additional information on the regulators, such as the regulator ID and references to databases.

Work done by Christian Oertlin.

The WikiPathways tissue analyzer (teaser)

TissueAnalyzer
 Visit the Main Page [ctrl-alt-z]

The dataset come from the [E-MTAB-1733](#) experiment. And it have been analysed through the TissueAnalyzer plugin on [PathVisio](#)

Select tissue:

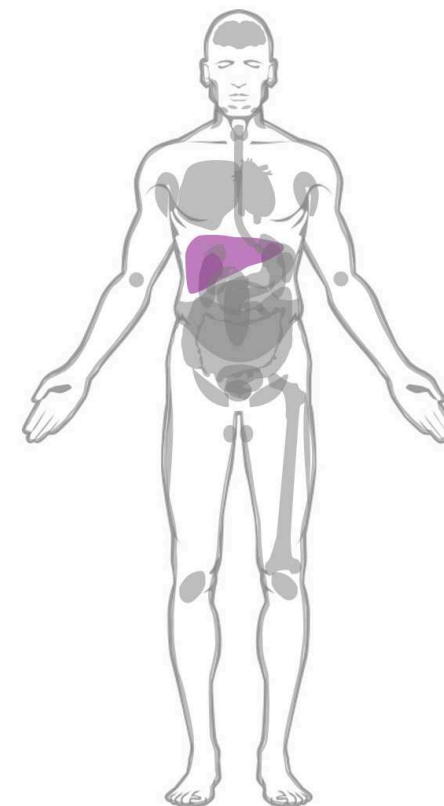
Show common pathways



Gradient color scale



Viewer	Pathway name	Median expression	Mean expression	Ratio active gene measured	(%)
<input type="radio"/>	Nicotine_Metabolism	8.31	6.96	5/6	83
<input type="radio"/>	Polyol_Pathway	7.23	7.06	3/4	75
<input type="radio"/>	Complement_Activation,_Classical_Pathway	7.22	7.22	9/17	52
<input type="radio"/>	Diclofenac_Metabolic_Pathway	7.21	7.36	2/4	50
<input type="radio"/>	Tamoxifen_metabolism	7.11	6.51	10/19	52
<input type="radio"/>	Synthesis_and_Degradation_of_Ketone_Bodies	6.94	6.35	3/5	60
<input type="radio"/>	Arachidonate_Epoxygenase_-_Epoxyde_Hydrolase	6.54	6.71	3/7	42
<input type="radio"/>	Benzo(a)pyrene_metabolism	6.49	6.67	4/9	44
<input type="radio"/>	Folate-Alcohol_and_Cancer_Pathway	6.49	6.05	4/8	50
<input type="radio"/>	Complement_and_Coagulation_Cascades	6.41	6.52	24/51	47
<input type="radio"/>	Phase_I_biotransformations,_non_P450	6.36	6.21	2/7	28
<input type="radio"/>	Glycine_Metabolism	6.21	4.6	1/3	33
<input type="radio"/>	Blood_Clotting_Cascade	6.18	5.58	8/24	33
<input type="radio"/>	Mitochondrial_LC-Fatty_Acid_Beta-Oxidation	6.14	6.0	5/16	31
<input type="radio"/>	Alanine_and_aspartate_metabolism	6.08	5.35	5/12	41
<input type="radio"/>	Amino_acid_conjugation_of_benzoic_acid	5.95	4.97	2/4	50
<input type="radio"/>	Quercetin_and_Nf-kB-AP-1_induced_cell_apoptosis	5.95	5.4	3/13	23



Work done by Jonathan Melus

WikiPathways, a house of webs?



Pathways in Cytoscape

F1000Research » Articles



SOFTWARE TOOL

WikiPathways App for Cytoscape: Making biological pathways amenable to network analysis and visualization [v1; ref status: indexed, <http://f1000r.es/3ij>]

Martina Kutmon^{1*}, Samad Lotia^{2*}, Chris T Evelo¹, Alexander R Pico²

+ Author affiliations

* Equal contributors

+ Grant information

Views

525

Downloads

141

Get PDF

Get XML

Cite

Track

Email

Share



This article is included in the [Cytoscape App Collection](#)

Abstract

In this paper we present the open-source WikiPathways app for Cytoscape (<http://apps.cytoscape.org/apps/wikipathways>) that can be used to import biological pathways for data visualization and network analysis. WikiPathways is an open, collaborative biological pathway database that provides fully annotated pathway diagrams for manual download or through web services. The WikiPathways app allows users to load pathways in two different views: as an annotated pathway ideal for data visualization and as a simple network to perform computational analysis. An example pathway and dataset are used to demonstrate the functionality of the WikiPathways app and how they can be combined and used together with other apps. More than 2000 downloads between its first release in August 2013 and the submission of the paper in May 2014 highlight the importance and adoption of the app in the network biology field.

Figure 1. The Cardiac Hypertrophic Response pathway loaded as a pathway.

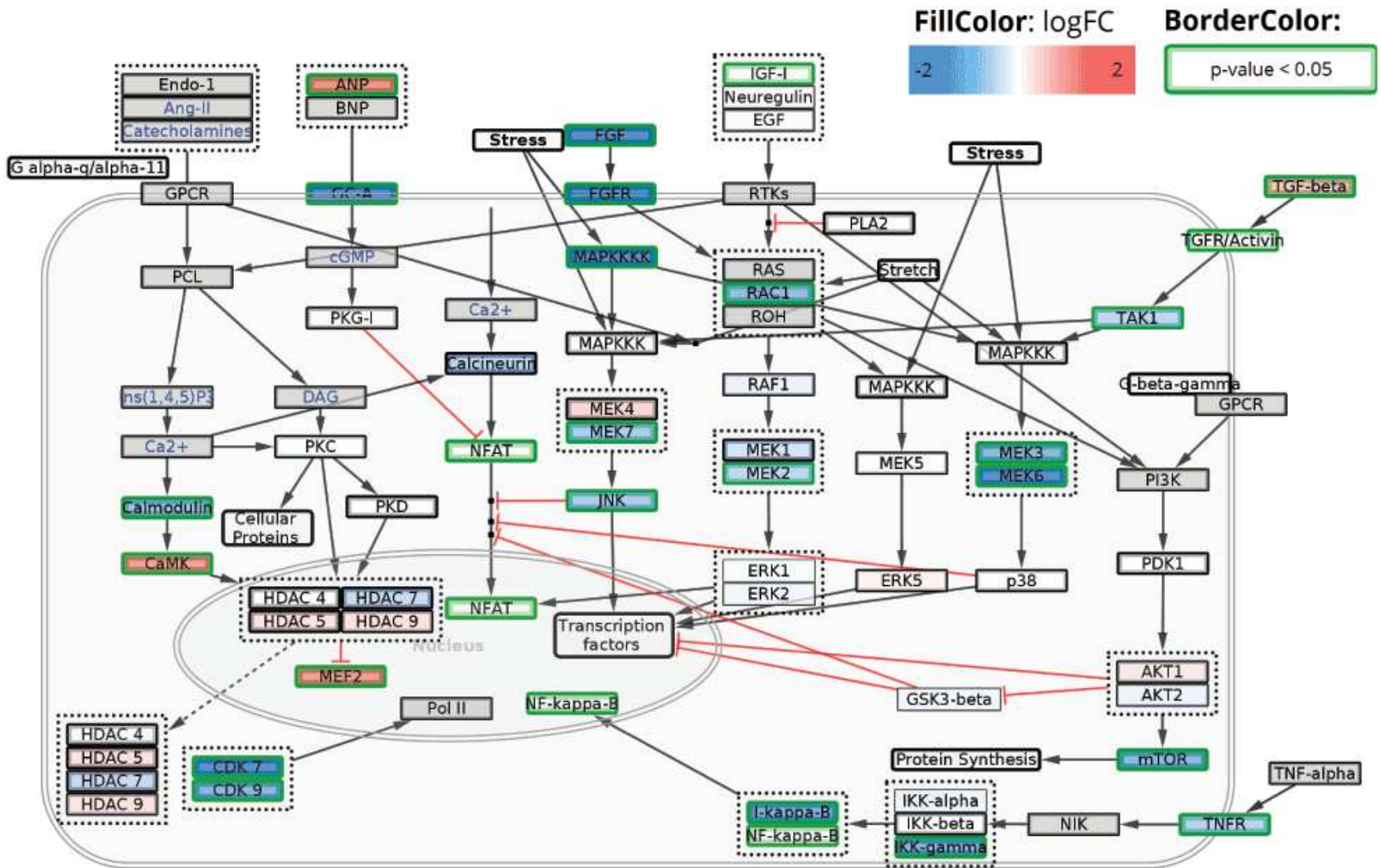
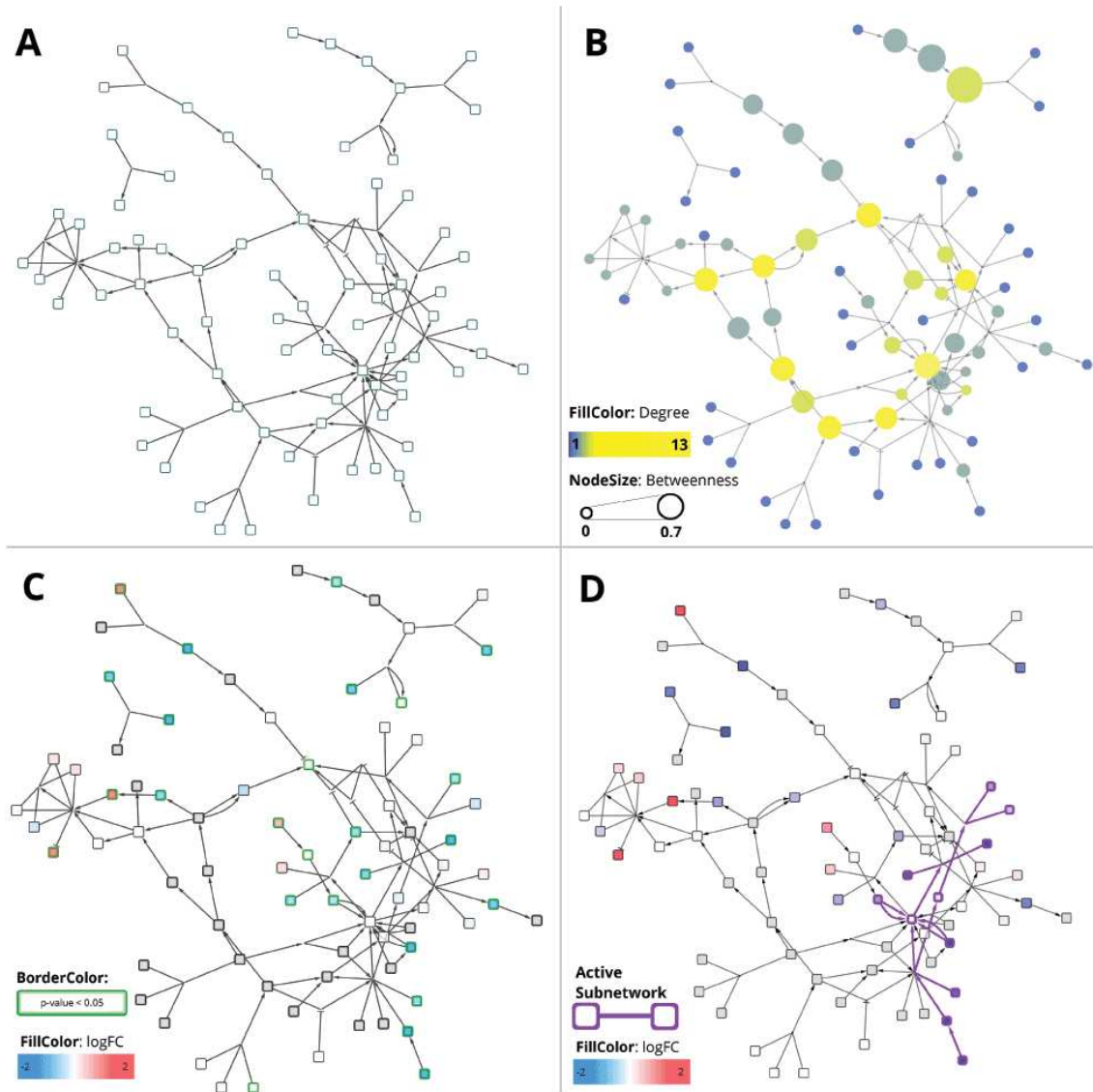


Figure 2. The Cardiac Hypertrophic Response pathway loaded as a network.



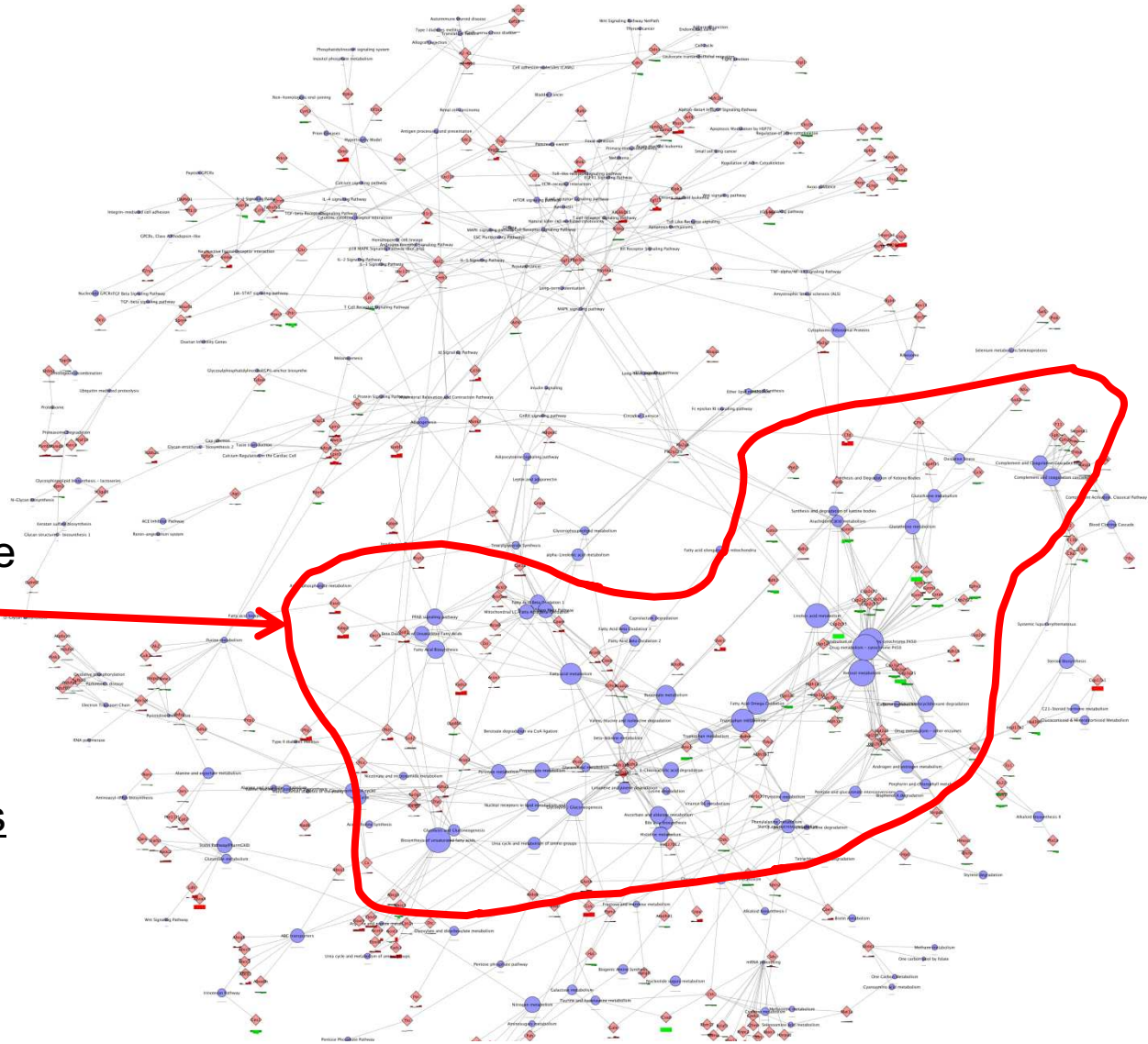
Cytoscape visualization used to group

PPS1
Liver

All pathways

Pathways with high z-score grouped together.

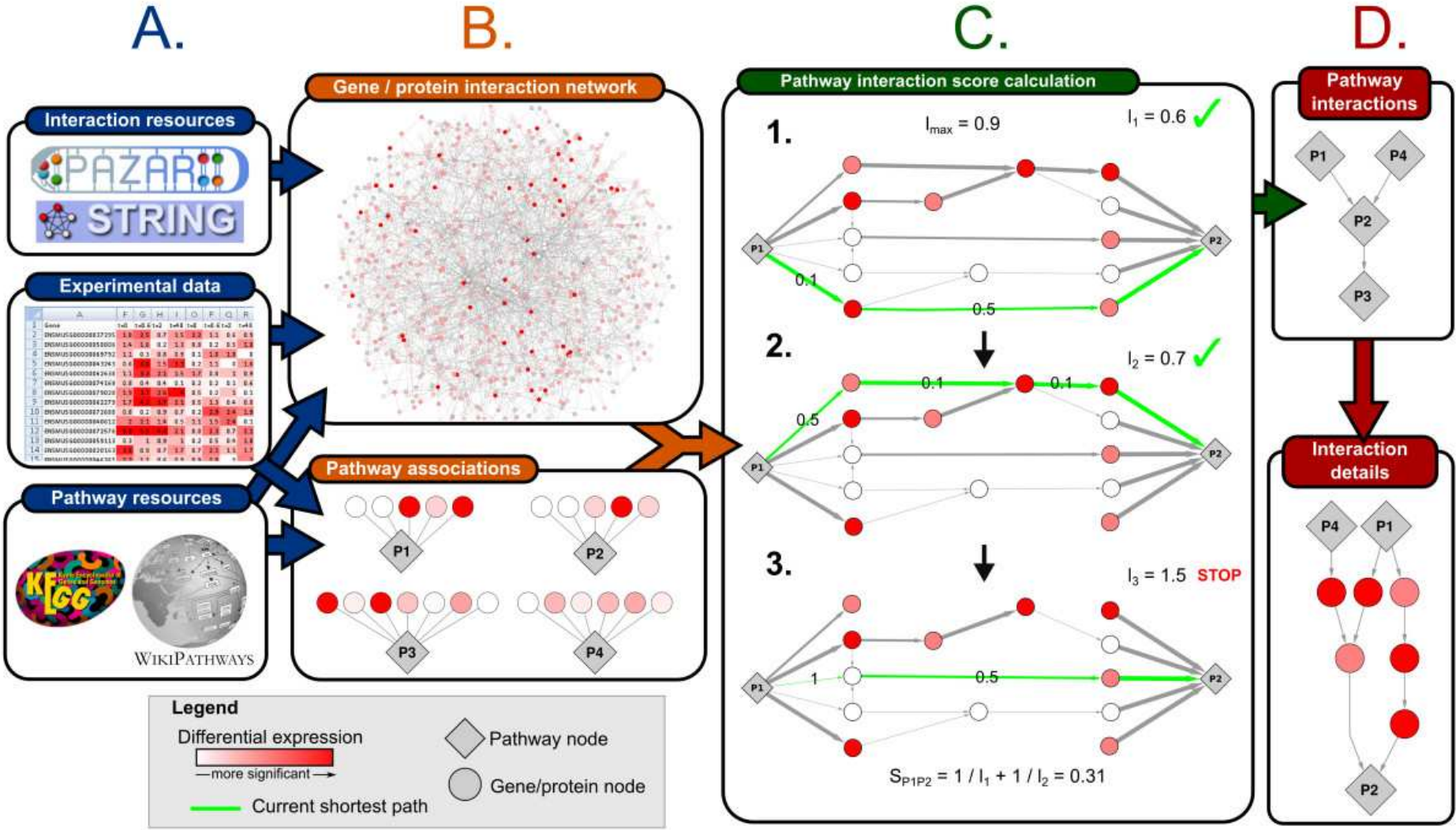
Explains why there are relatively few significant genes, but many pathways with high z-score.



Robert Caesar et al (2010) A combined transcriptomics and lipidomics analysis of subcutaneous, epididymal and mesenteric adipose tissue reveals marked functional differences. PLoS One 5: 7. e11525

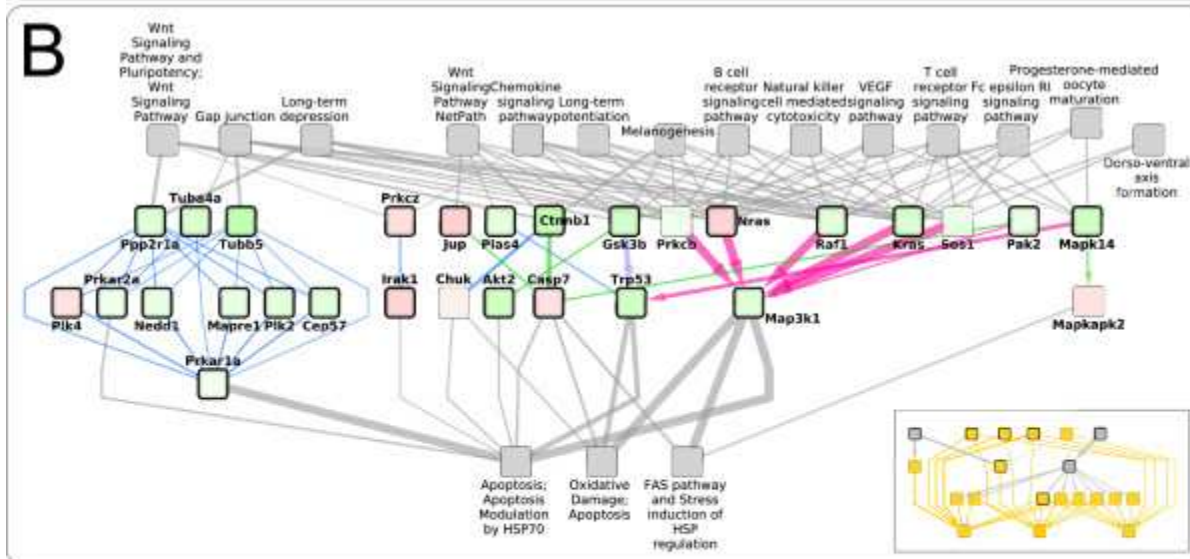
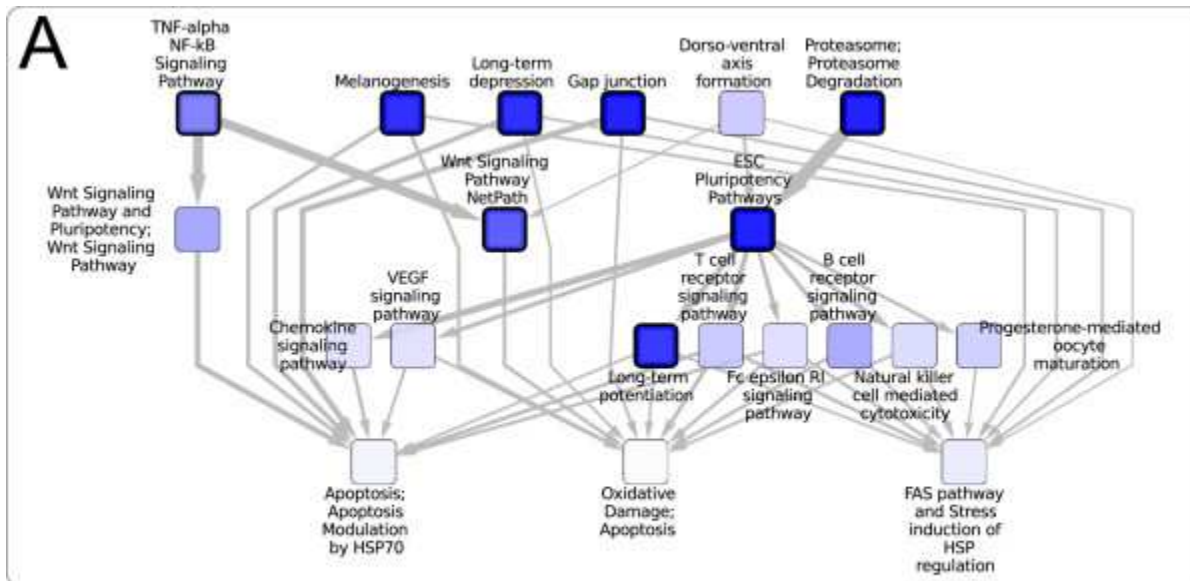
<http://dx.doi.org/doi:10.1371/journal.pone.0011525>

Pathway interactions and what causes them



Thomas Kelder, Lars Eijssen, Robert Kleemann, Marjan van Erk, Teake Kooistra, Chris Evelo (2011) Exploring pathway interactions in insulin resistant mouse liver.

BMC Systems Biology 5: 127 Aug. <http://dx.doi.org/doi:10.1186/1752-0509-5-127>



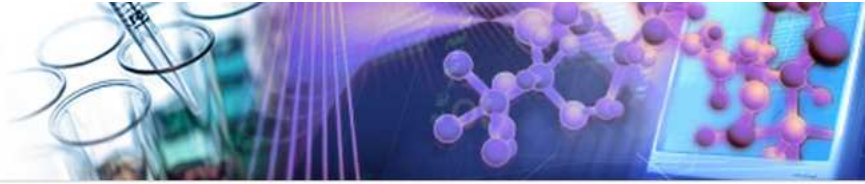
Pathway interactions and detailed network visualization for the interactions with three apoptosis related pathways for the comparison between HF and LF diet at t = 0. A: Subgraph of the pathway interaction network, based on incoming interactions to three stress response and apoptosis pathways with the highest in-degree. Pathway nodes with a thick border are significantly enriched ($p < 0.05$) with differentially expressed genes. B: The protein interactions that compose the interactions between the three apoptosis related pathways and their neighbors in the subgraph as shown in box A (see inset, included interactions are colored orange). Protein nodes have a thick border when their encoding genes are significantly differentially expressed ($q < 0.05$).



Regulation resources

DRUGBANK

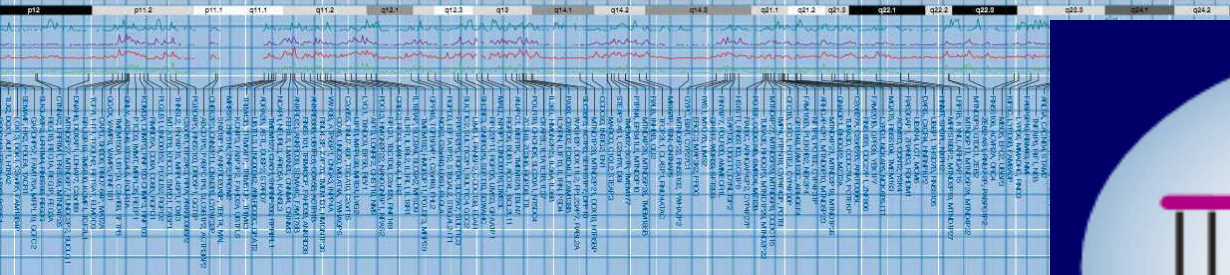
Open Data Drug & Drug Target Database



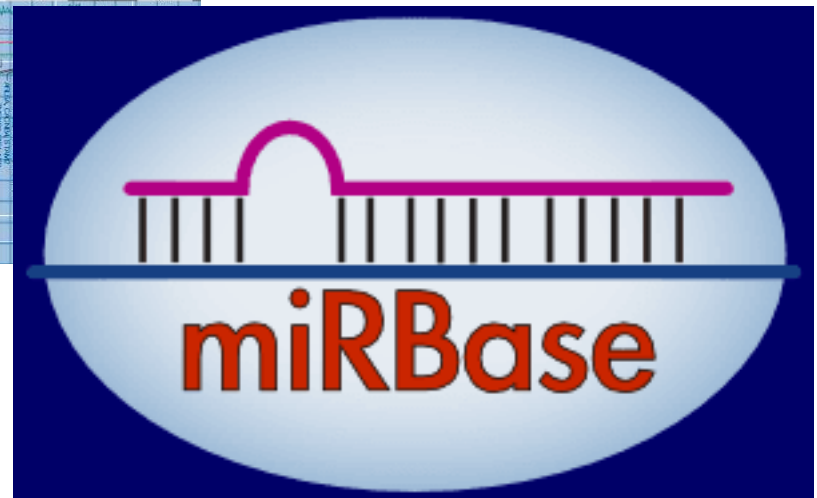
ENCODE
ENCYCLOPEDIA OF DNA ELEMENTS



transcription factor
encyclopedia



miRTarBase
ORF AAAA



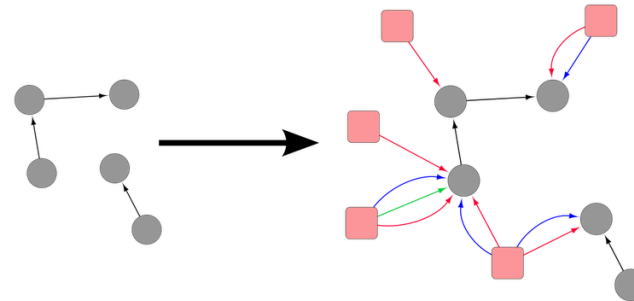
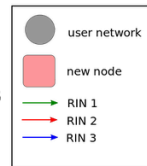
CyTargetLinker

a generic Cytoscape app to build regulatory networks

[Home](#)[Tutorials](#)[Download](#)[RegINs](#)[Source Code](#)[How to cite](#)[Citations](#)[Team](#)[Contact](#)

News

- **CyTargetLinker publication in PLoS One:** [How to cite CyTargetLinker.](#)
- **ICSB 2013:** CyTargetLinker was accepted for a presentation at the International Conference on Systems Biology 2013 in Copenhagen on 2 September 2013. Check out my slides [here](#).
- **Cytoscape 3 app:** CyTargetLinker 3.0.1 was released on 15 August 2013. This is the first version for Cytoscape 3.
- **NetBio SIG 2013:** CyTargetLinker was accepted for a presentation at the Network Biology SIG of the ISMB 2013 in Berlin on 19 July 2013.
- **NBIC Application Showcase Prize:** CyTargetLinker won the application showcase price at Netherlands Bioinformatics Conference 2013 in Lunteren.



Tutorials

[How to use CyTargetLinker.](#)

Learn how to use CyTargetLinker by going through one of our tutorials.

Downloads

[Get CyTargetLinker.](#)

Download CyTargetLinker and the relevant regulatory interaction networks (RINs).

Regulatory Interaction Networks (RegINs)

We are providing RegIN collections per species. If you are looking for a specific version of a RegIN, please check our [archive](#). If you are missing a regulatory interaction database or species, please [contact](#) us and we will try to create the network files.

Due to license issues, we are not providing the data from [miRecords](#) as RegINs, but you can run our conversion script for this resource to create the RegINs yourself. Check out the our [conversion scripts](#). If you have any questions, please [contact](#) us.

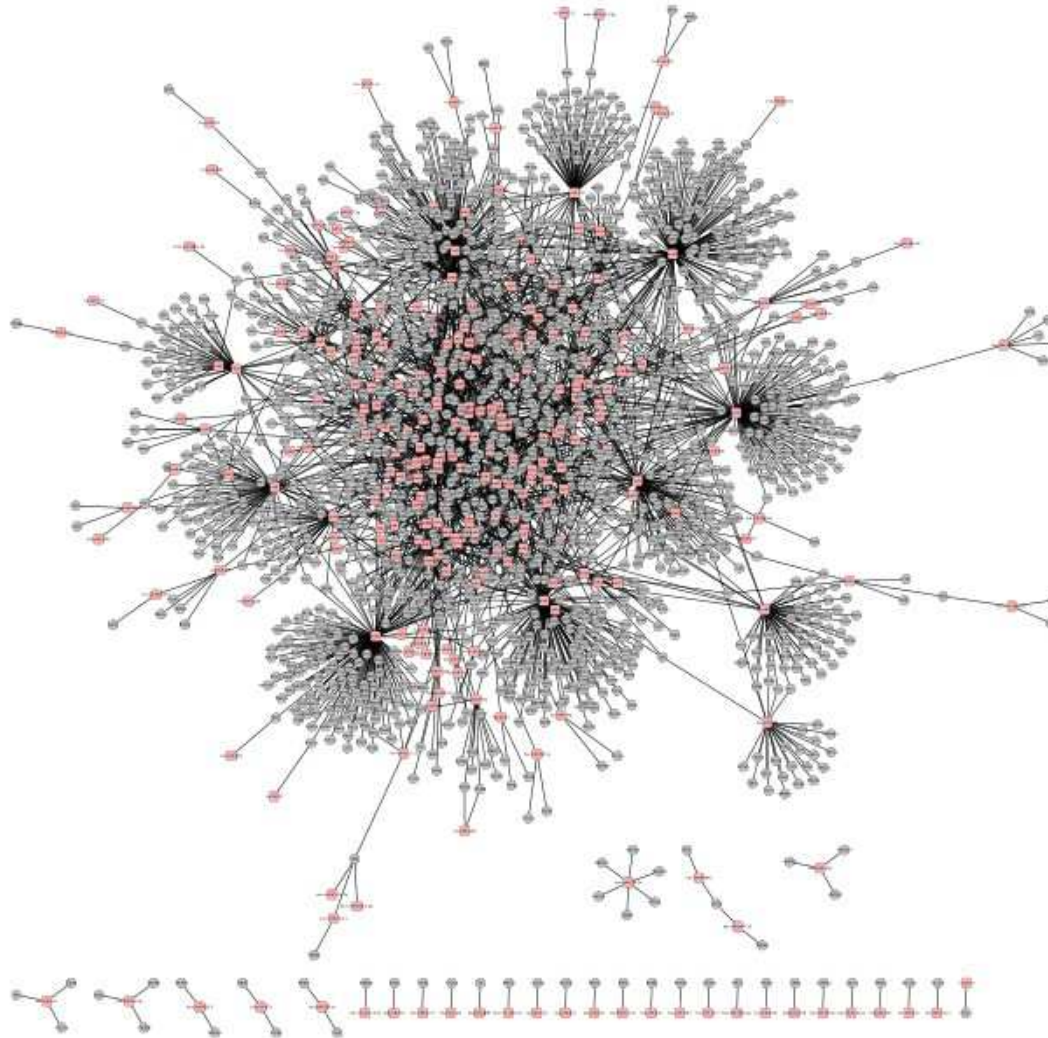
Do you want to create your own RegIN? [Find our more.](#)

Collection	Interaction Types	Databases
Homo sapiens	microRNA-target, TF-target, drug-target	MicroCosm 5 (541,039) TargetScan 6.2 (511,040) miRTarBase 4.4 (20,942) ENCODE (24,111 proximal + 18,240 distal) TFe (1,088) DrugBank 3 (9,570)
Mus musculus	microRNA-target	MicroCosm 5 (494,822) TargetScan 6.2 (186,431) miRTarBase 4.4 (9,216)
Rattus norvegicus	microRNA-target	MicroCosm 5 (511,057) miRTarBase 4.4 (285)
Caenorhabditis elegans	microRNA-target	MicroCosm 5 (53,729) miRTarBase 4.4 (3,213)
Danio rerio	microRNA-target	MicroCosm 5 (121,992) miRTarBase 4.4 (104)

Databases

Database	Interaction Types	RegINs
TransmiR	Transcription factor – microRNA interactions	TransmiR RegINs
miRBase	primary transcript – miRNA	miRBase transcripts

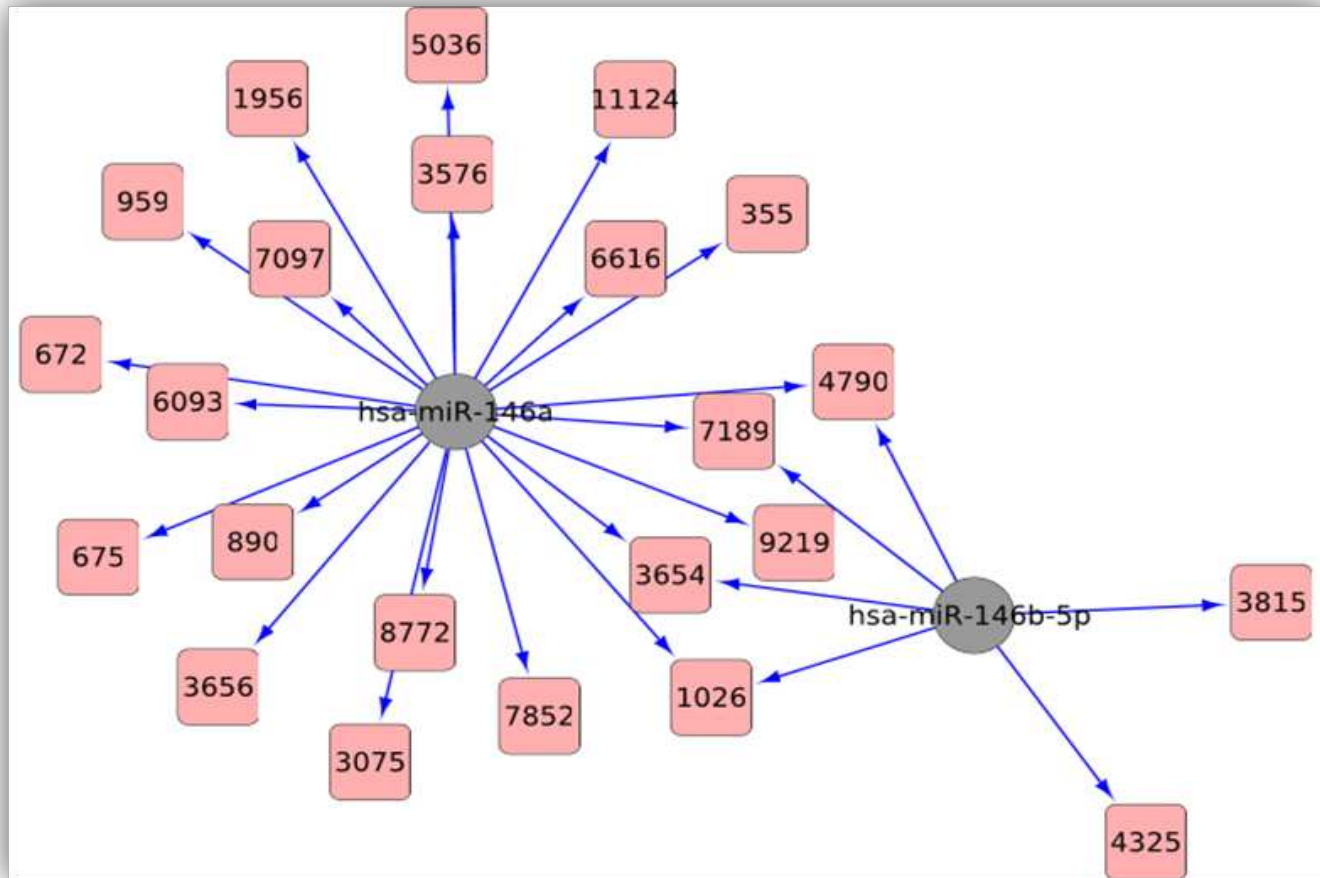
miRTarBase as a target interaction network



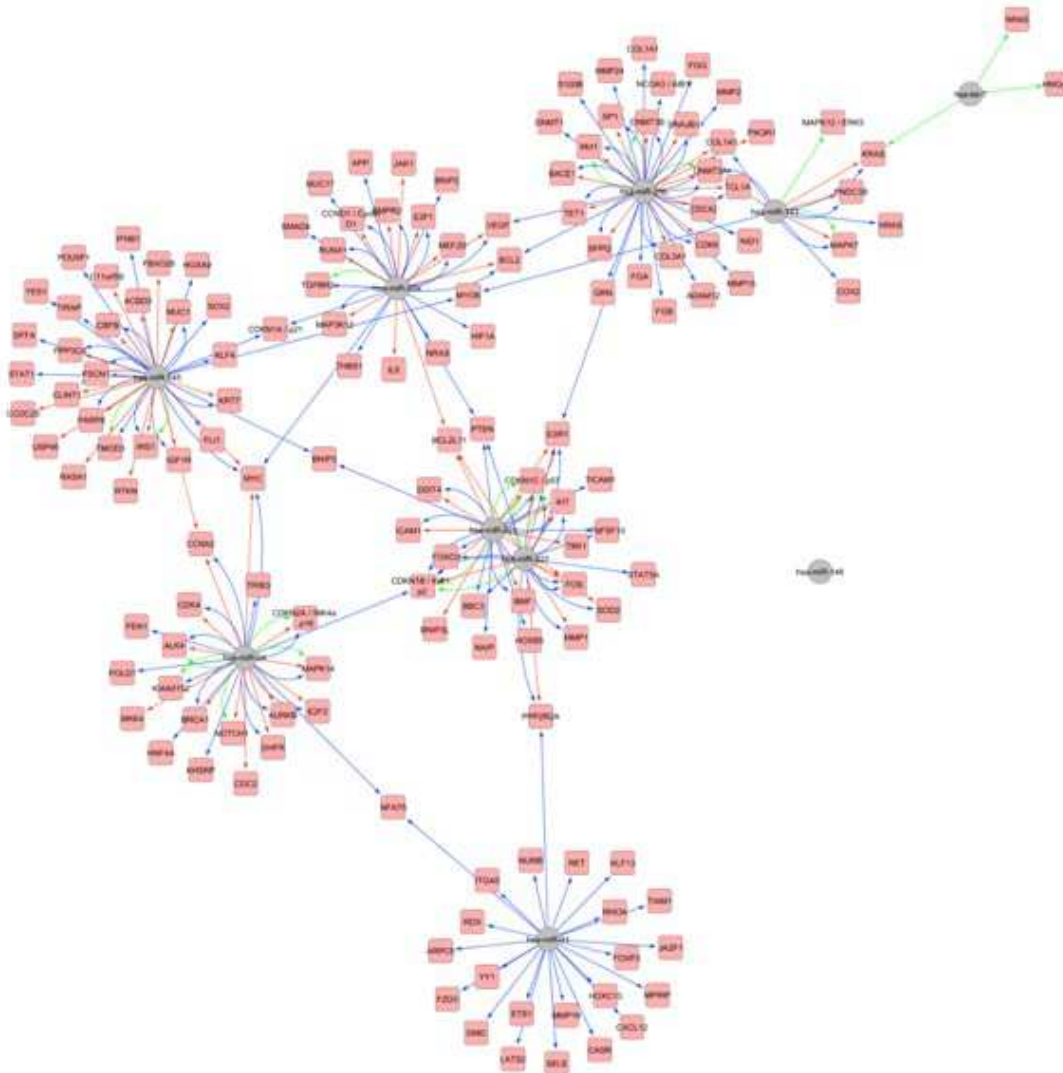
Collection of miRNA-target gene interactions in the miRTarBase database with 1,715 genes, 286 miRNAs and 2,817 interactions.

MiRNAs of Interest

miRNA target information from mirTarBase



miRNAs associated with colorectal cancer extended with validated target genes



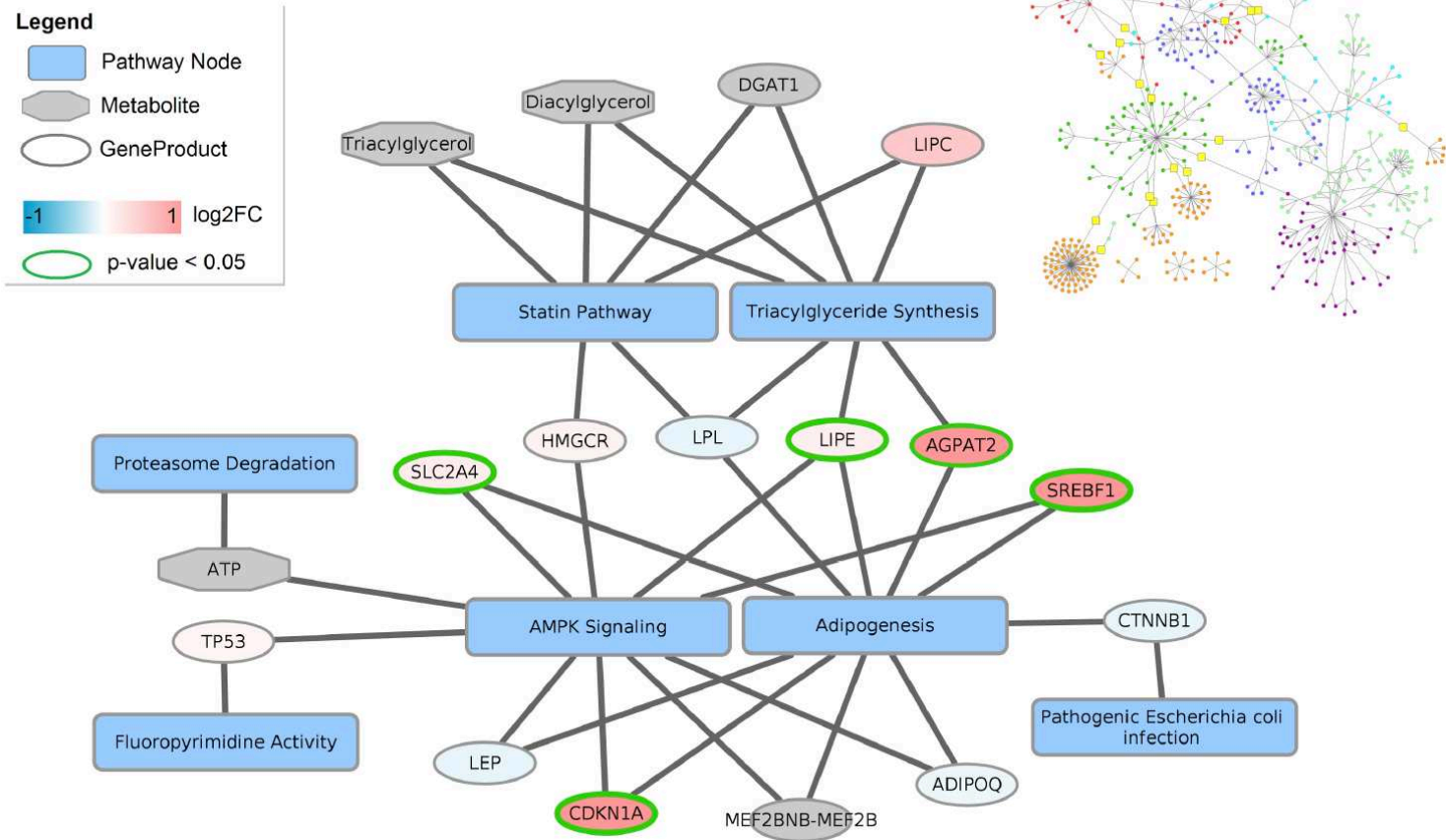


Use It

Now!

Integrate altered pathways into one combined network

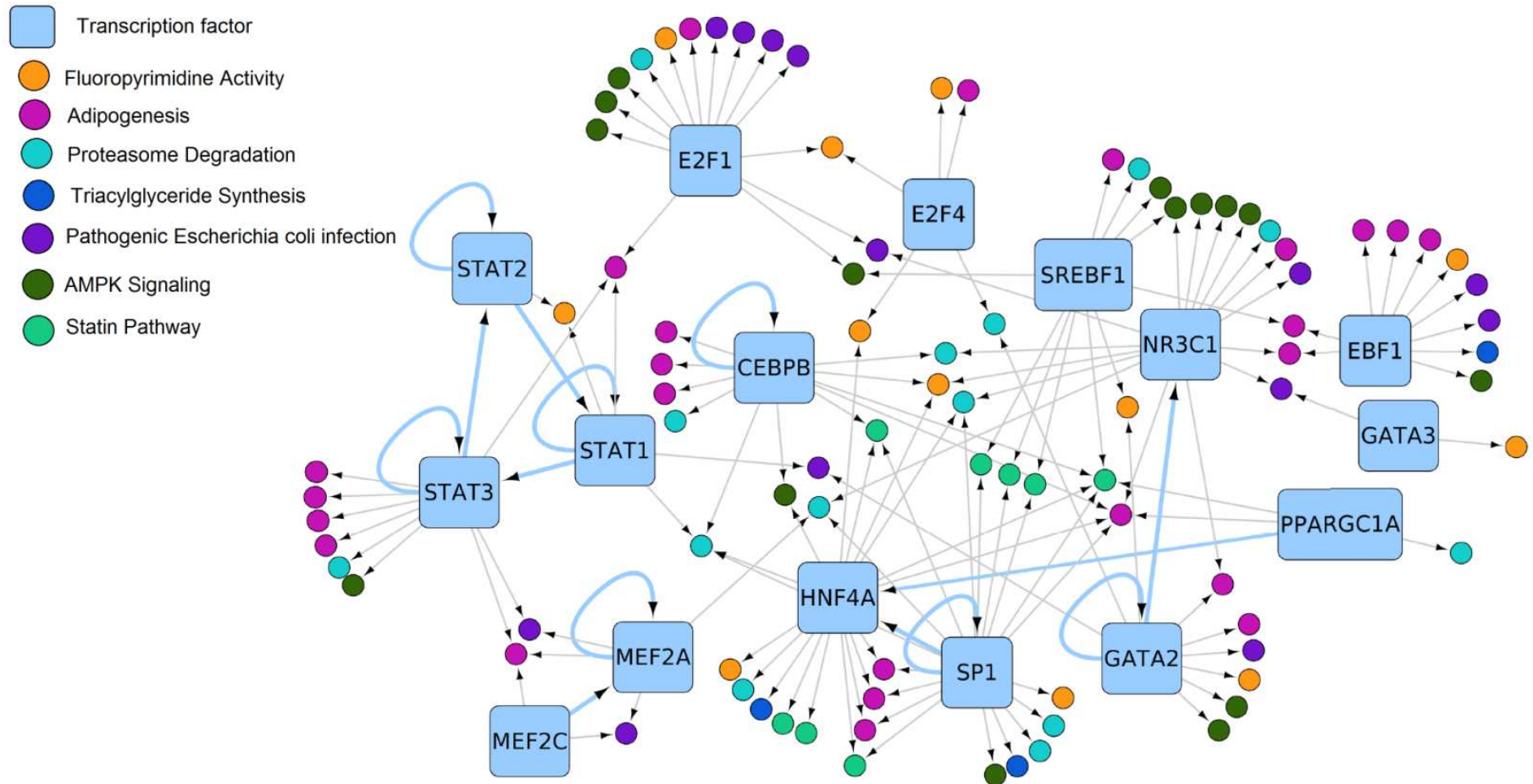
The pathways were merged into one large network using Cytoscape (WikiPathways and BridgeDb App). 14 genes and 4 metabolites were identified to link the different pathways to each other.



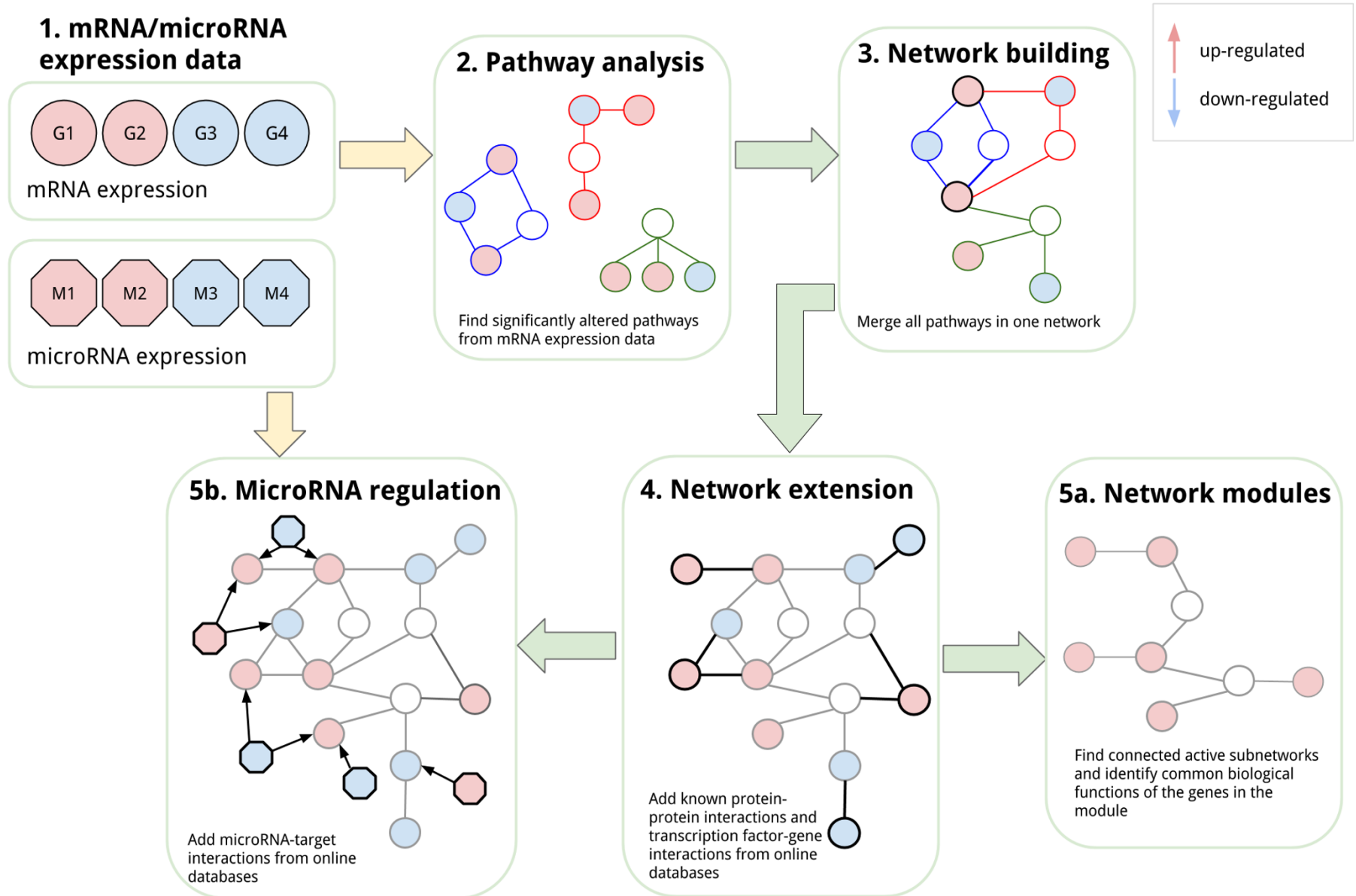
All linking genes are known to be key players in the development of diabetes and/or NAFLD. AGPAT2, CDKN1A and SREBF1 are significantly upregulated in the fatty, diabetic liver.

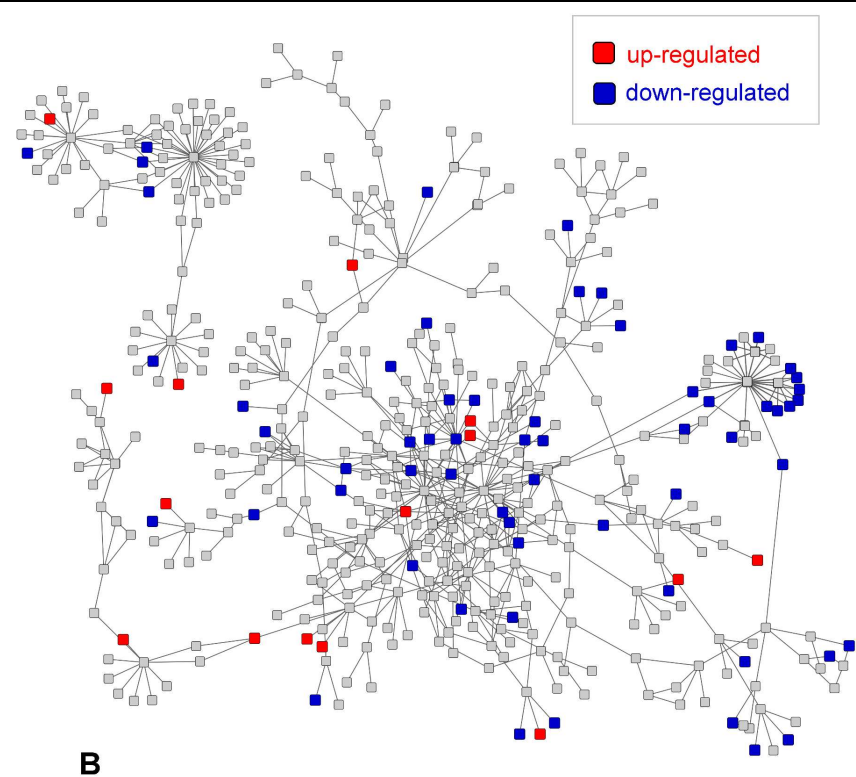
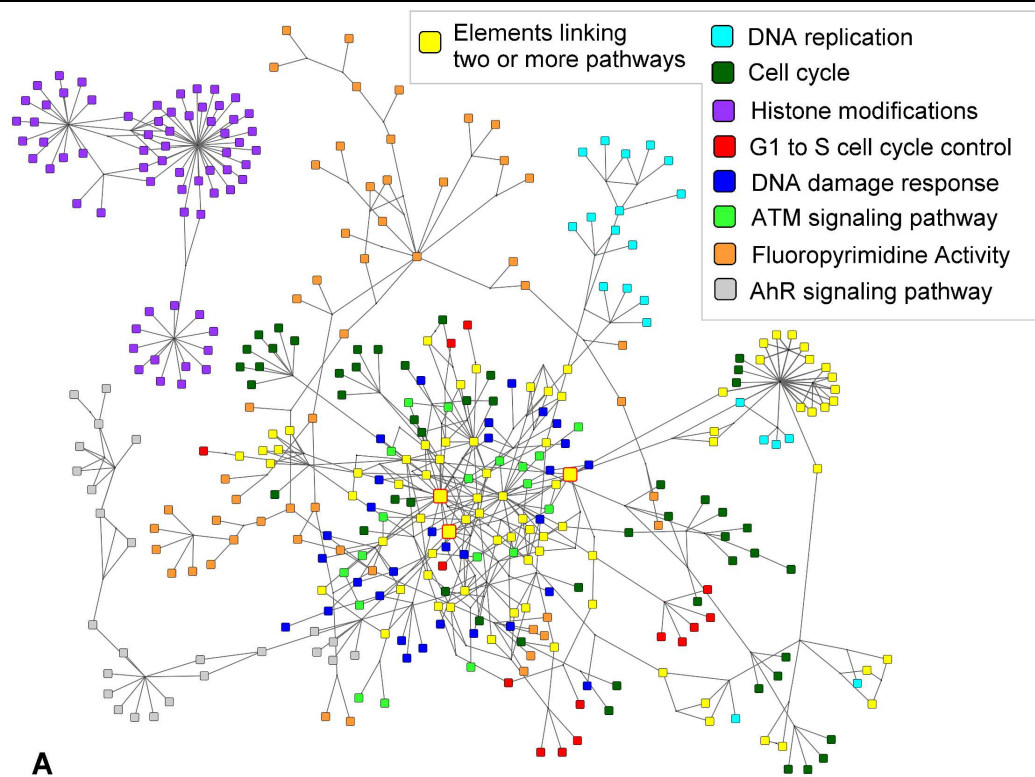
Transcriptional regulation

Using CyTargetLinker and the ENCODE proximal TF-target gene network, 16 transcription factors were identified in the selected pathways. Our analysis showed that TFs can be considered additional links between pathways and adding the regulatory interactions increases the overall connectivity of the network significantly.



Integrative network-based analysis of mRNA and microRNA expression in vitamin D3-treated cancer cells





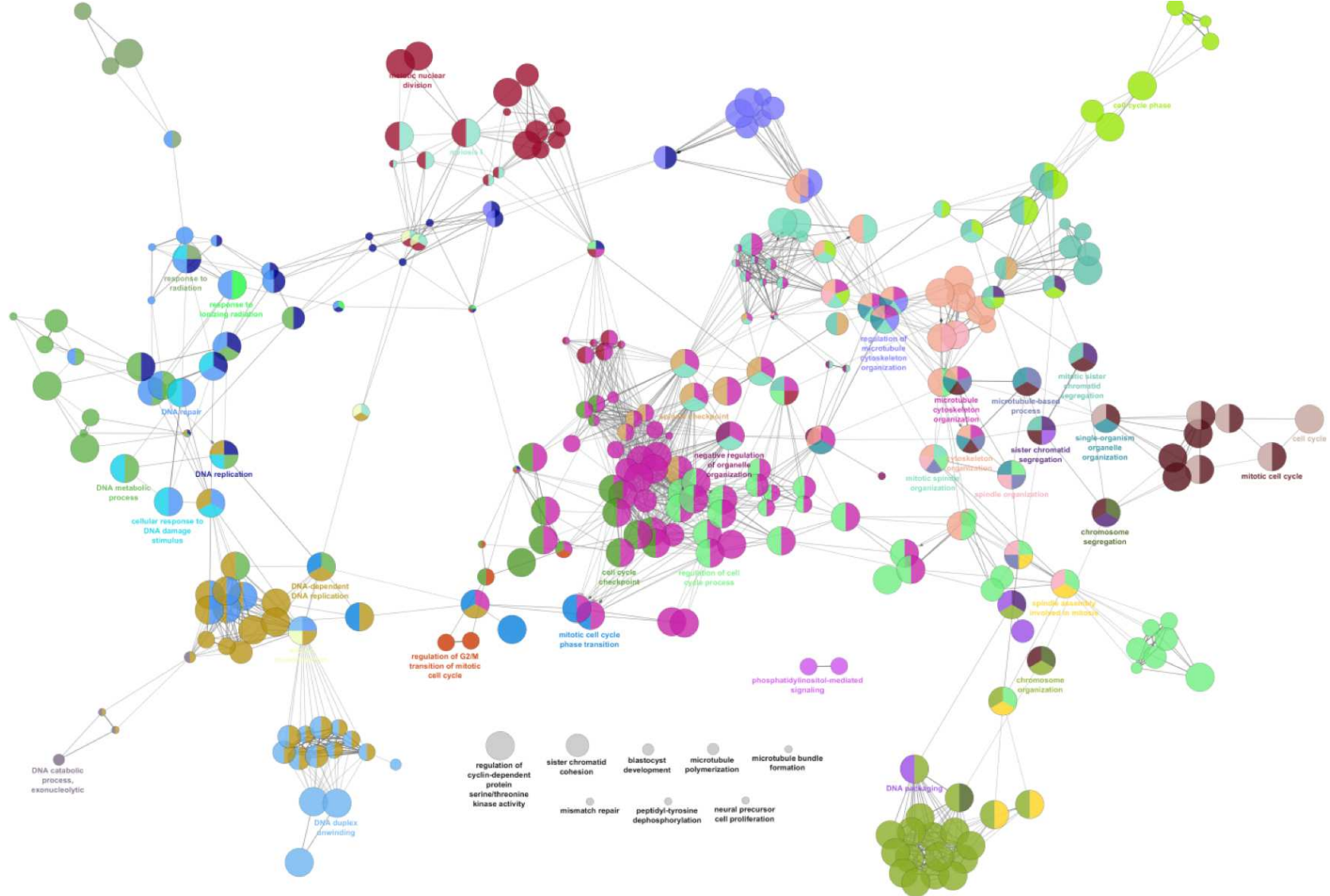
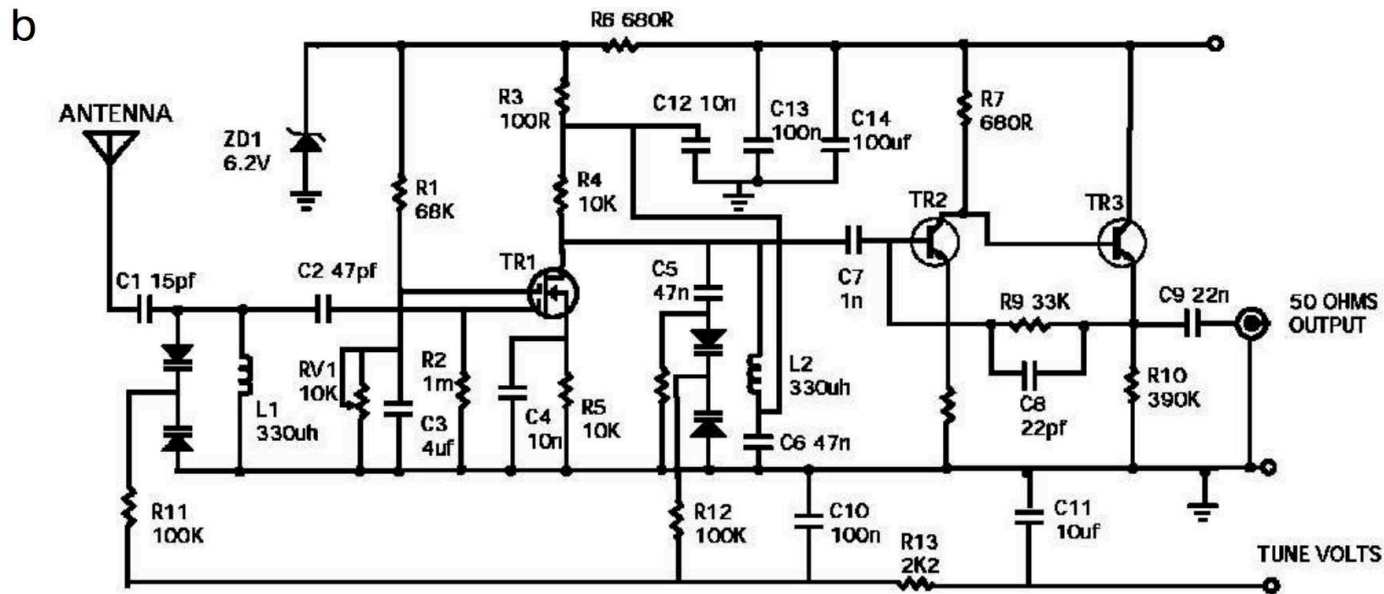


Fig 1: ClueGO network for highest scoring active module. The highest scoring active module in the network with 193 down-regulated genes was identified using the jActiveModules app. Then the ClueGO app was used to find relevant GO processes and a network of connected GO terms was created. Each node represents a GO biological process and the colors represent the GO group. In total, 34 GO groups and 8 GO processes not assigned to a group (shown in gray) are present in the network. Per group one representing GO biological process is named in the figure.

US advantage?



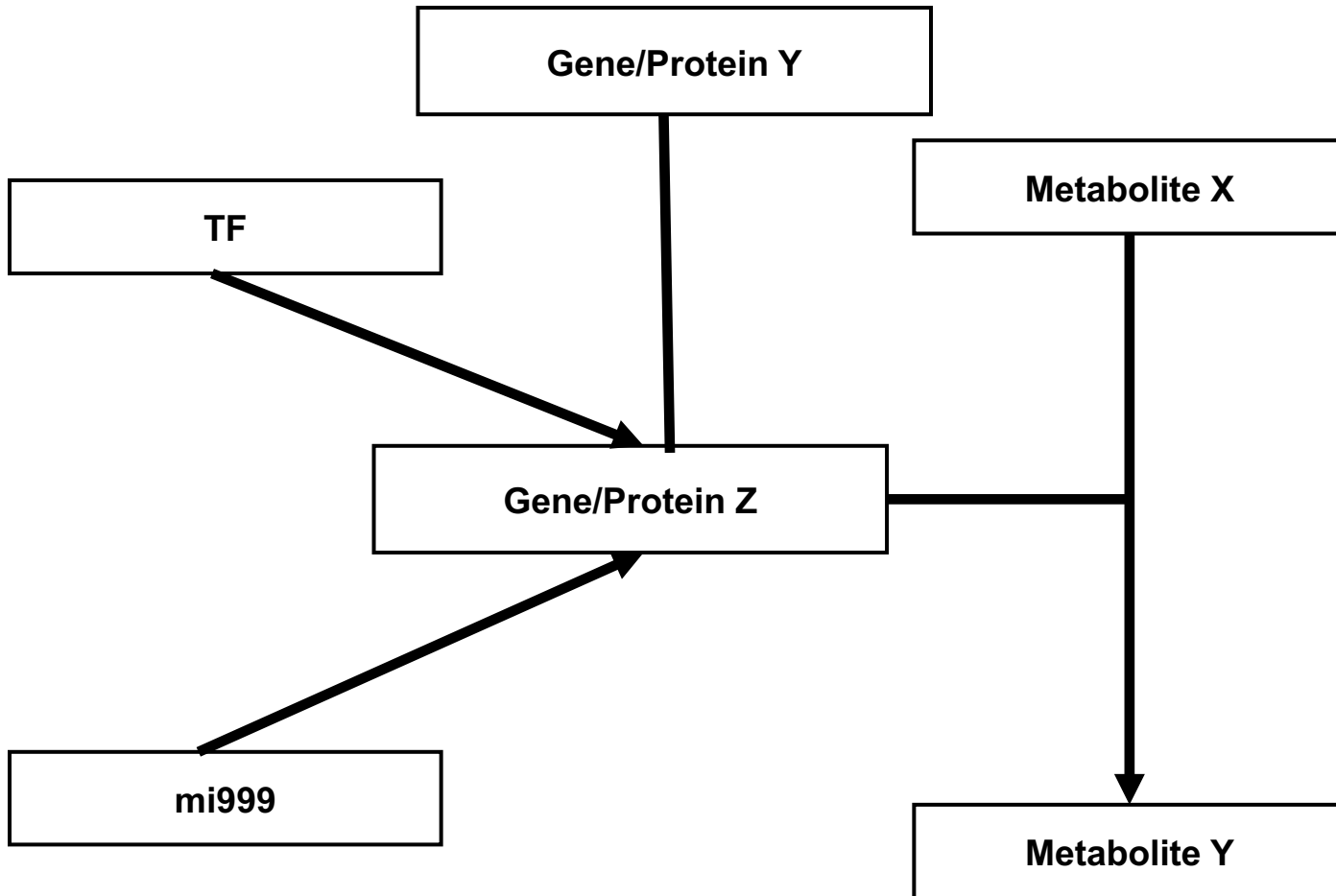
Can a biologist fix a radio?

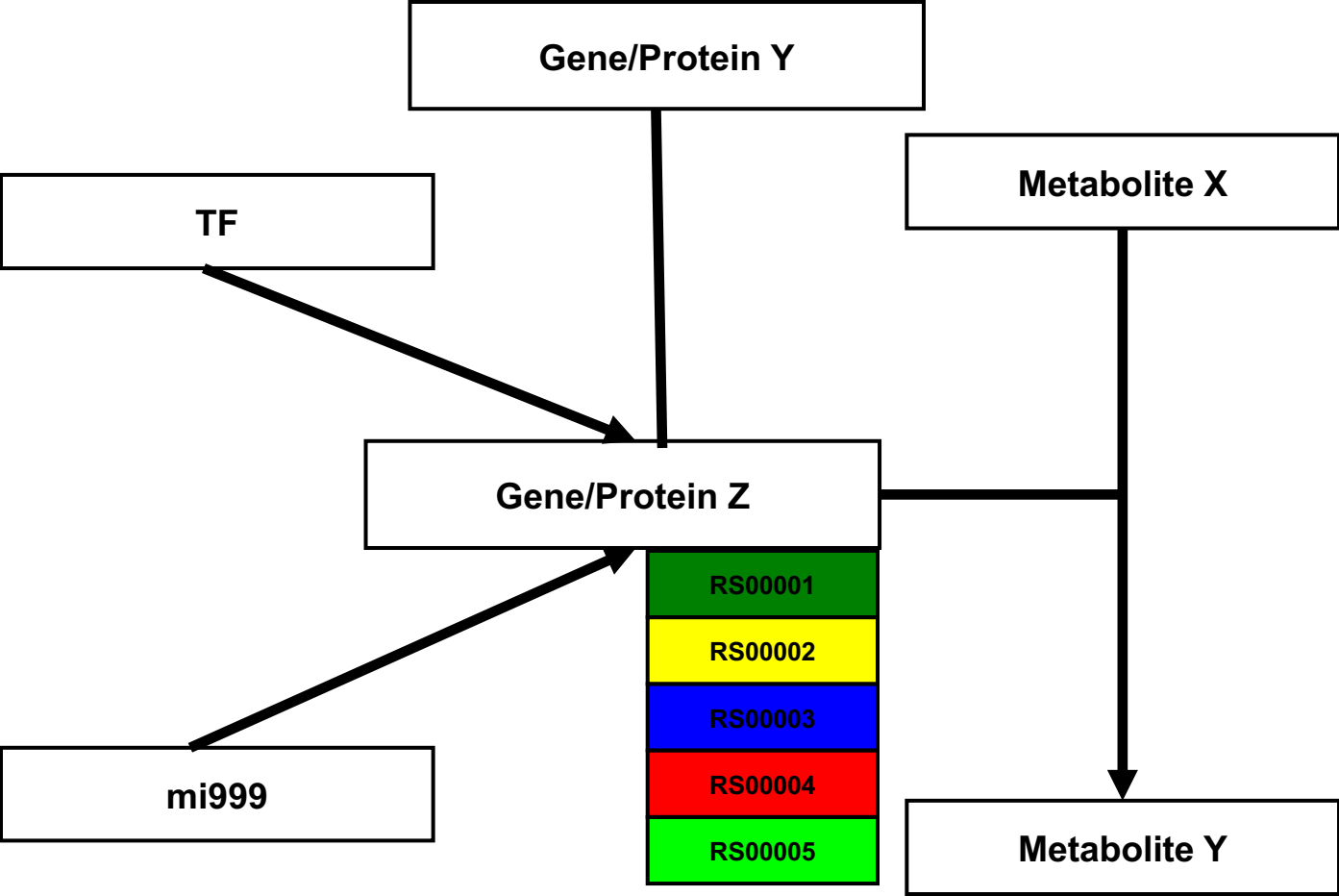


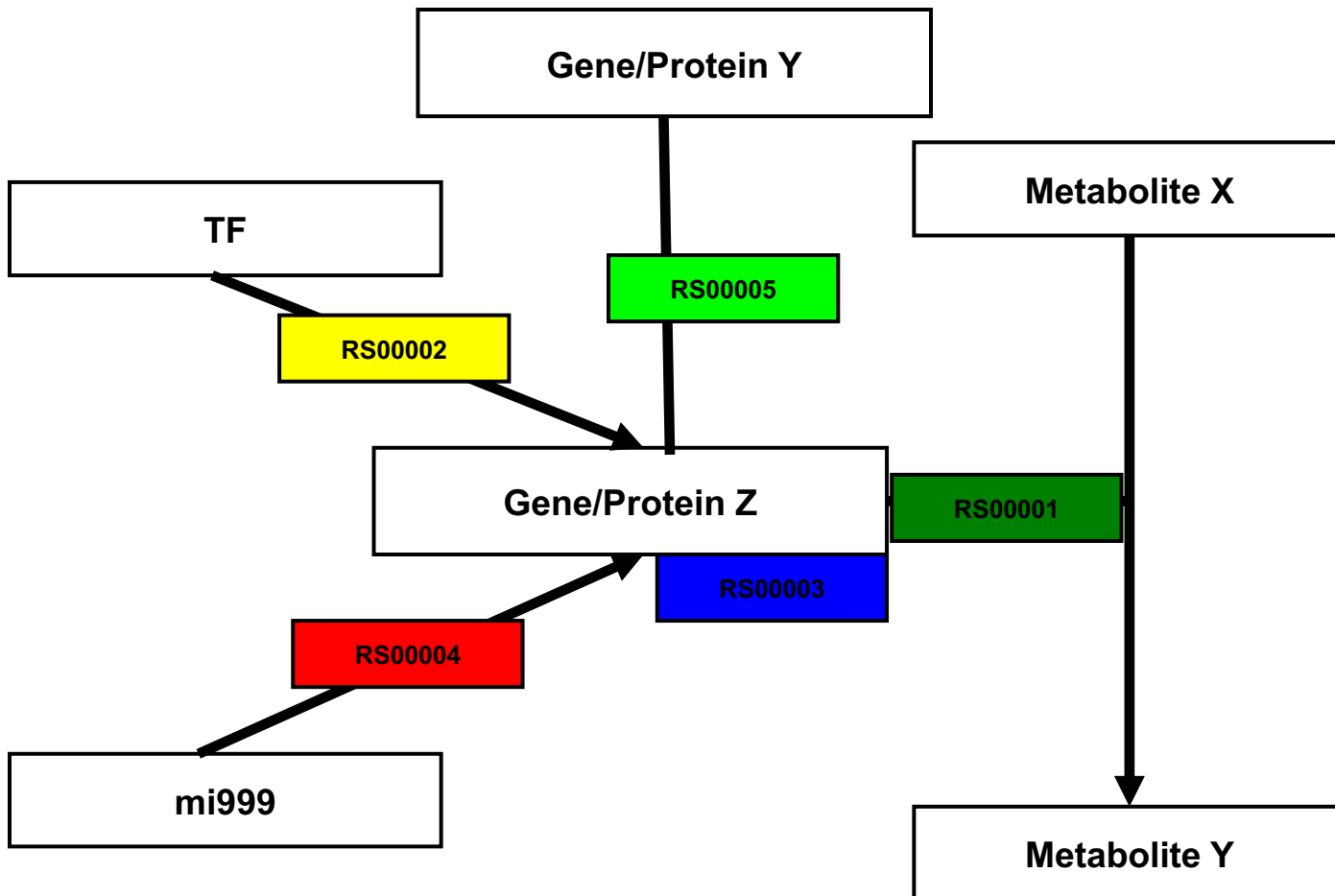
Lazebnik, Y. (2002). Can a biologist fix a radio? - or, what I learned while studying apoptosis. *Cancer Cell* 2, 179-182.

SNPs in pathways









Functionalize SNPs

Unknown function (attribute to gene)

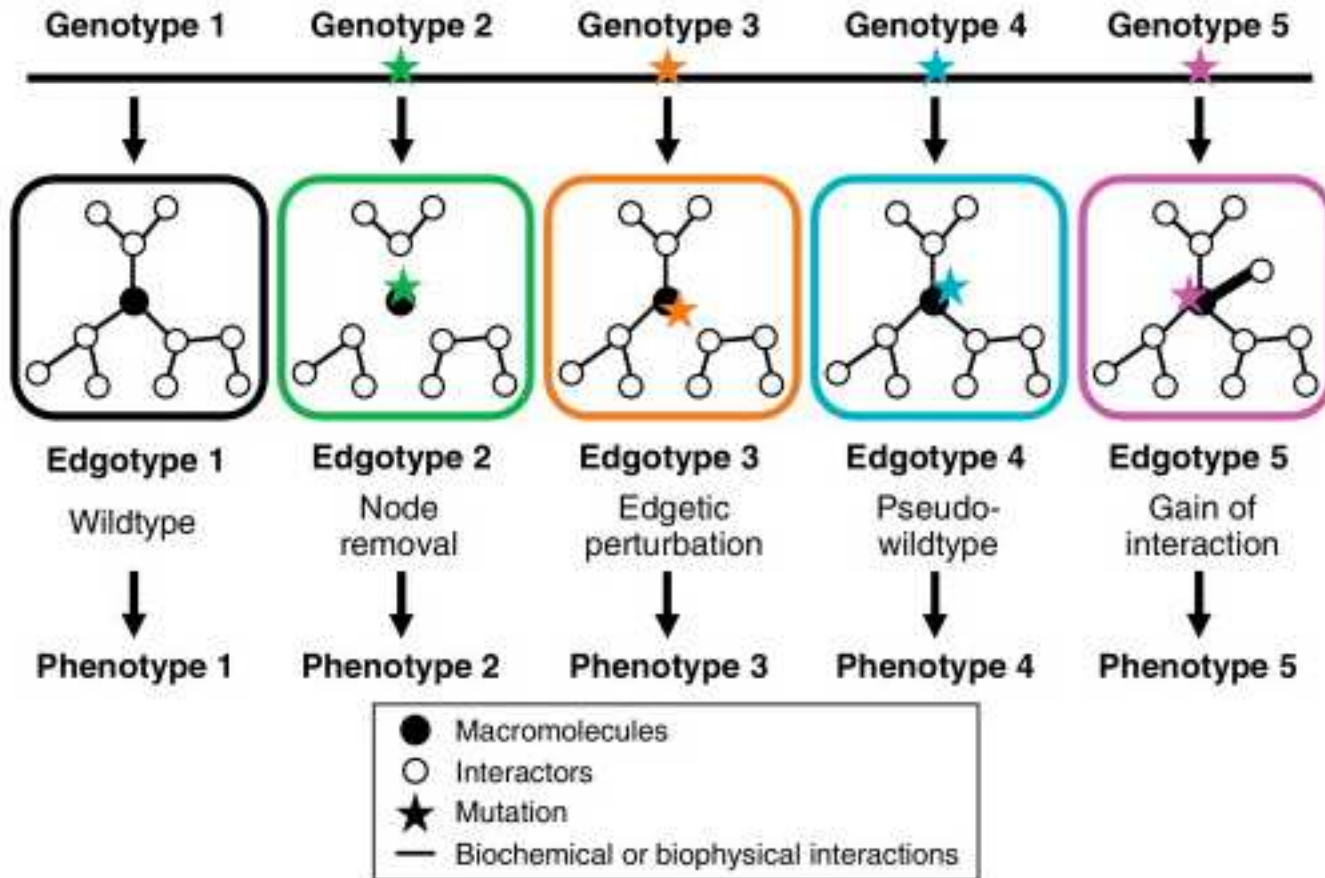
In miRNA binding site

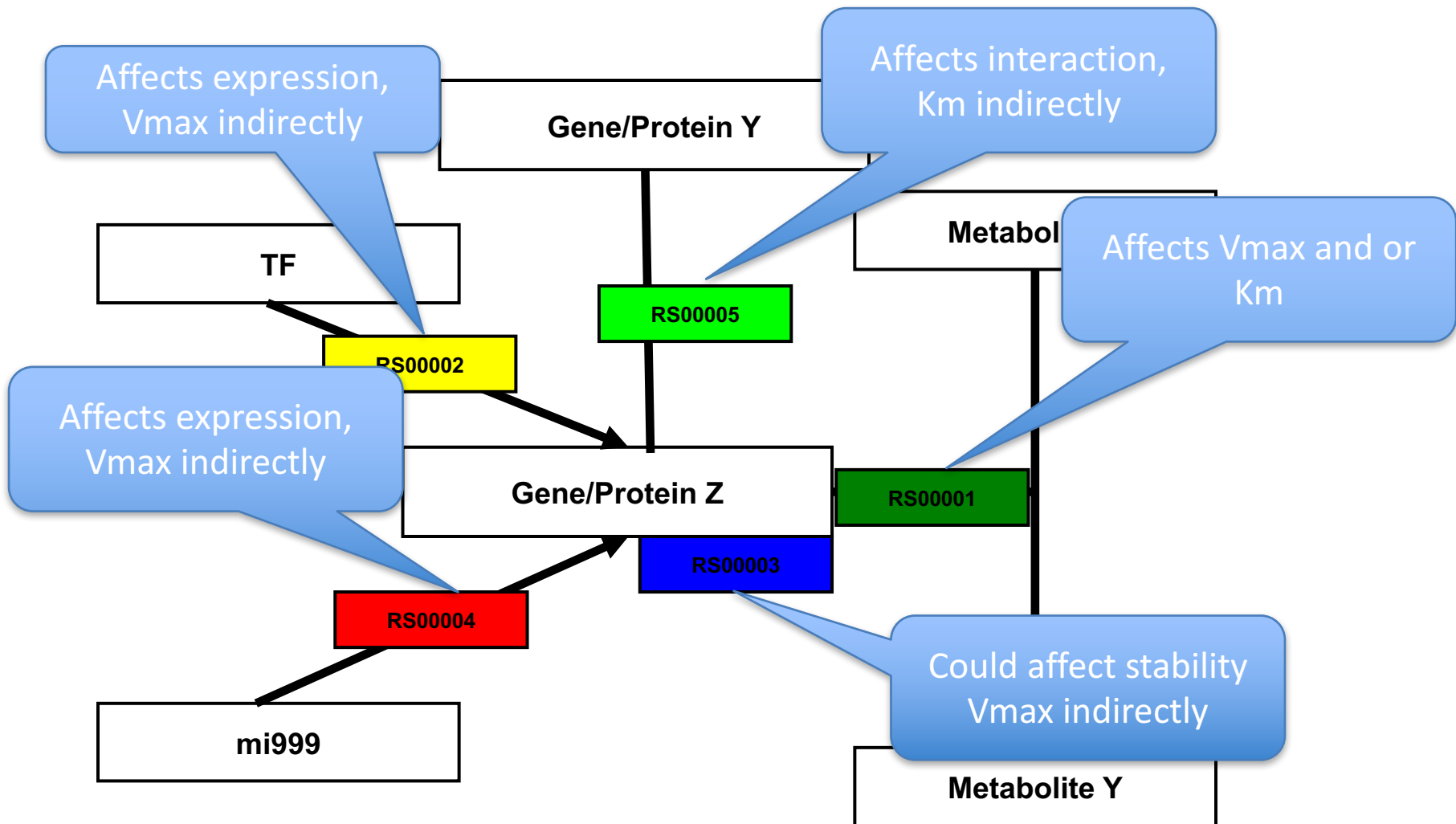
In TF binding site

Changing protein functionality (coding)

Changing protein interactions (coding)

Edgetics





Functionalize SNPs

Unkown function (attribute to gene)

In miRNA binding site

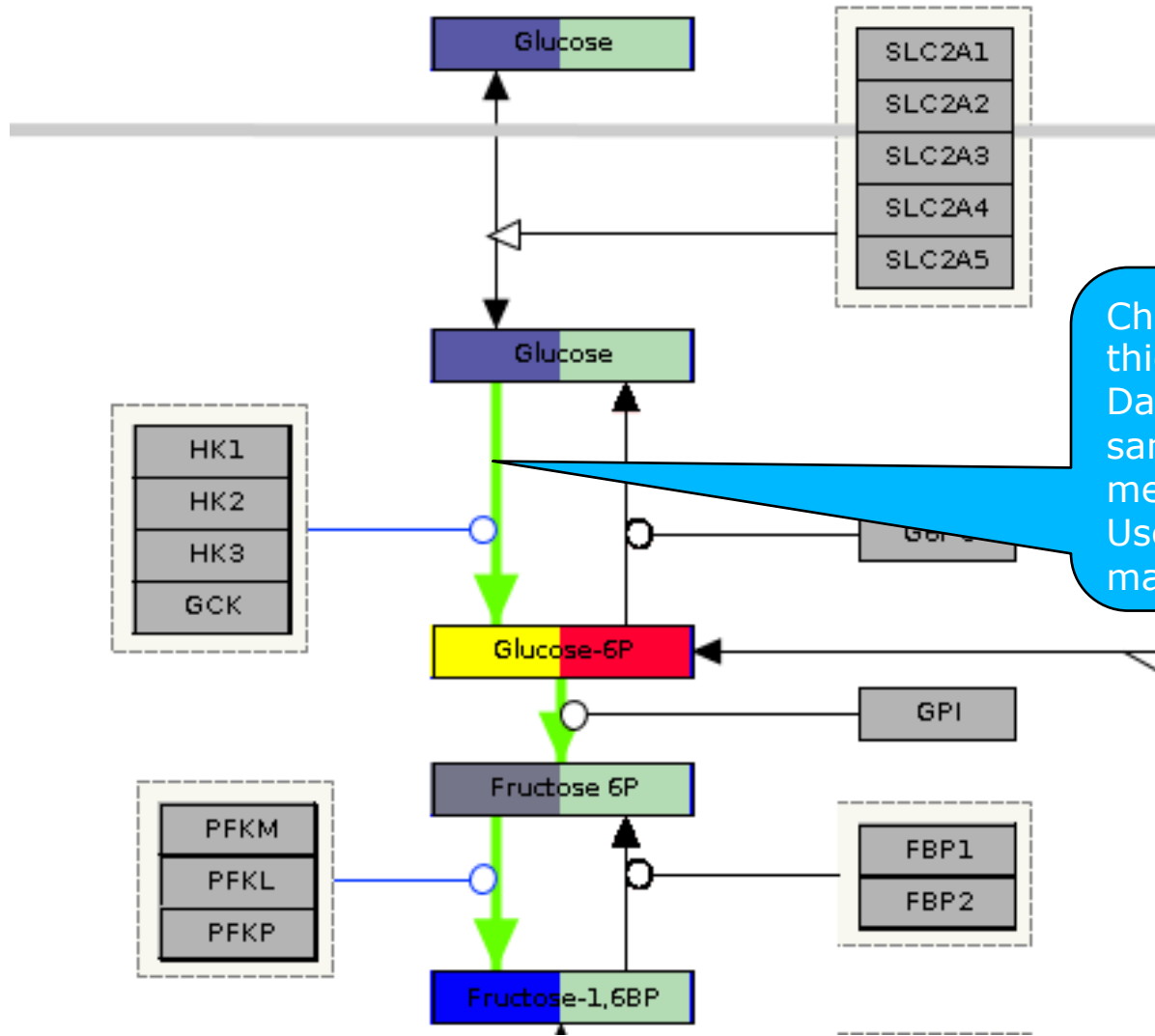
In TF binding site

Changing protein functionality (coding)

Changing protein interactions (coding)

Connect to modelling allows using SNPS

Fluxomics visualized on pathway



Change colours and line thickness.
Data mapping basically the same as for gene and metabolite data.
Uses RHEA as identifier mapping database.

Using SBML models as pathways

Search

PathSBML

Visualize data SBGN diagrams of SBML models

[Home](#) [Source Code](#) [Documentation](#) [Tutorials](#) [Blog](#) [Team](#) [Contact](#)

Home

PathSBML Plugin:

The PathSBML plugin for PathVisio which allows the researchers to visualize and overlay data on the same pathway models they use in simulation experiments. Exporting these models as GPML would also allow them to be uploaded to wikipathways for community curation.

Functionalities :

1. **Importing the SBML models.**
2. **Validation of SBML models.**
3. **Visualization of SBML elements.**
4. **Bio model integration – Importing the models from the bio model database.**
5. **Bio model search by Bio model name, Publication name/id, Person name, Taxonomy id, Uniprot id, CHEBI Id and GO Id.**
6. **Applying layout to the imported SBML model.**
7. **Visualization of the annotations of an SBML model.**

Poster presented by Anwasha Bohler.

PathSBML PathVisio plugin

Import/export models (e.g. from biomodels.org)

Combined analysis & visualization of omics and modelling results.

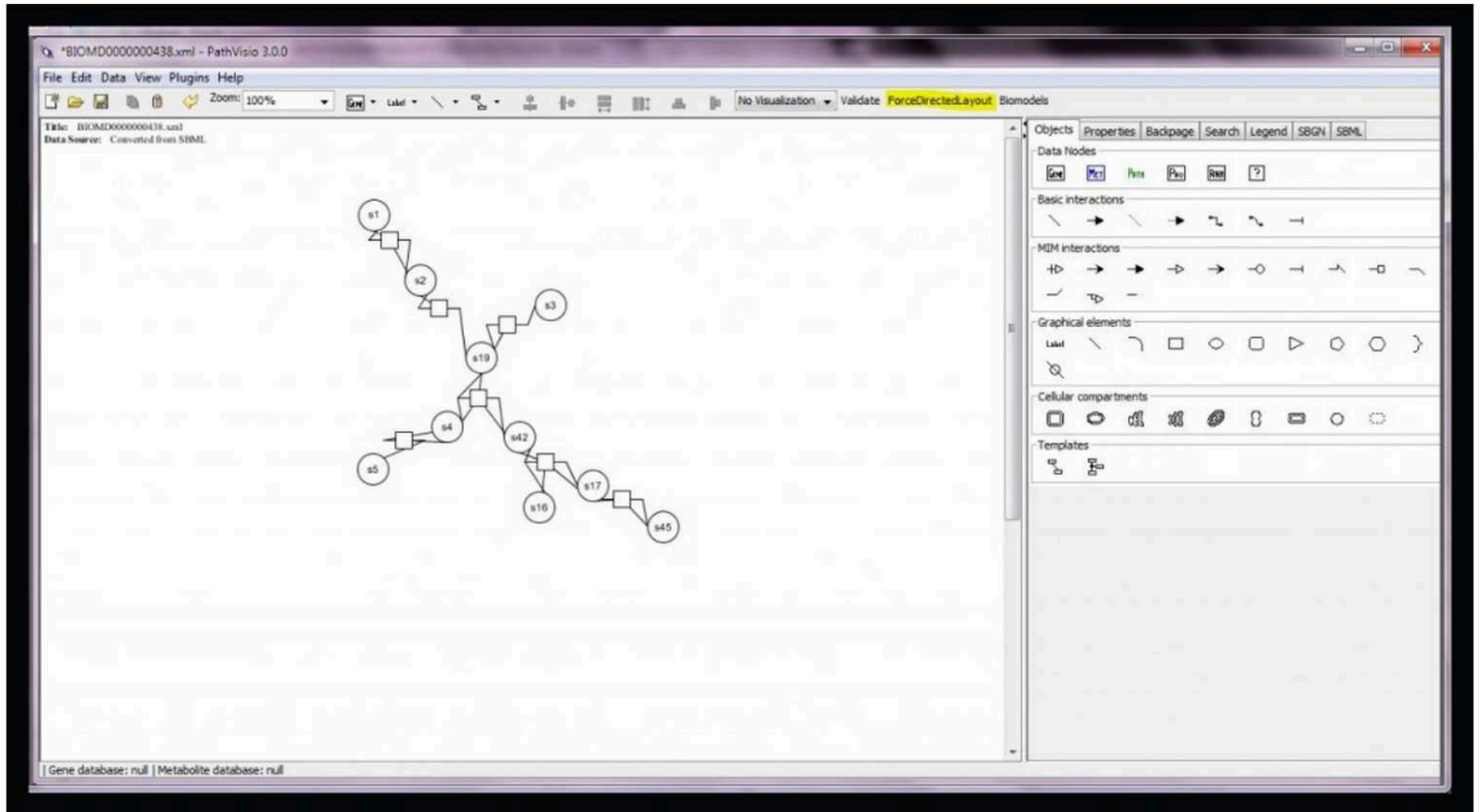
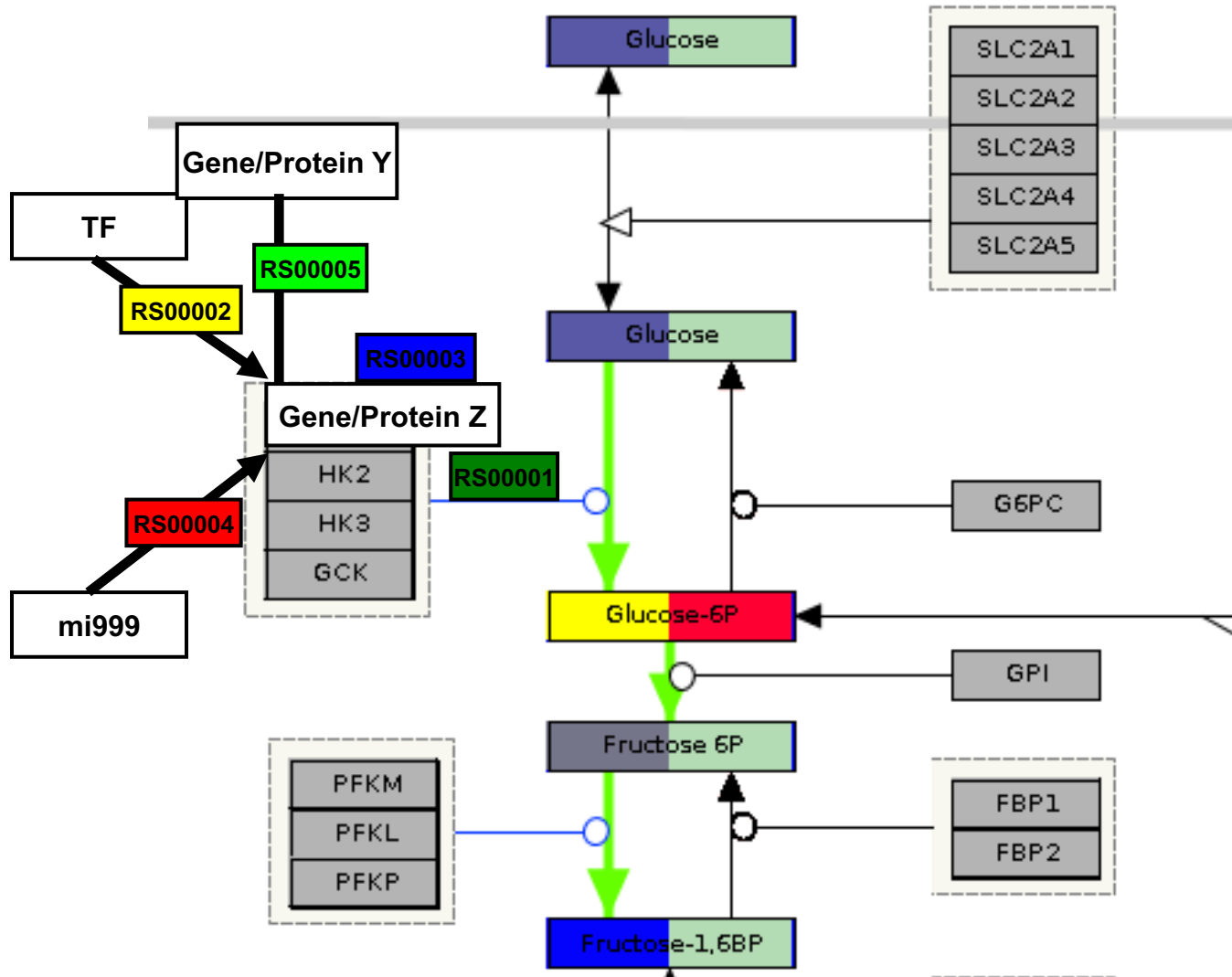


Figure2: After clicking on Force directed Layout button.

Connect to modelling allows using SNPS

Fluxomics visualized on pathway





Hand-drawn by [unreadable]