2255-3

2nd Conference on Systems Biology and New Sequencing Techniques" (2-4 November), preceded by Introductory Lectures on "Quantitative Approaches to Biological Problems" (31 October - 1 November)'

31 October - 4 November, 2011

Introduction to Systems Biology

Pascal Kahlem

European Bioinformatics Institute, Hinxton
United Kingdom

Introduction to Systems Biology

Examples from the European Network of Excellence ENFIN

ENFIN
Direction: Ewan Birney
Management: Pascal Kahlem
EU Project Officers: Frederick Marcus - Bernard Mulligan

EMBL - European Bioinformatics Institute
Hinxton, UK

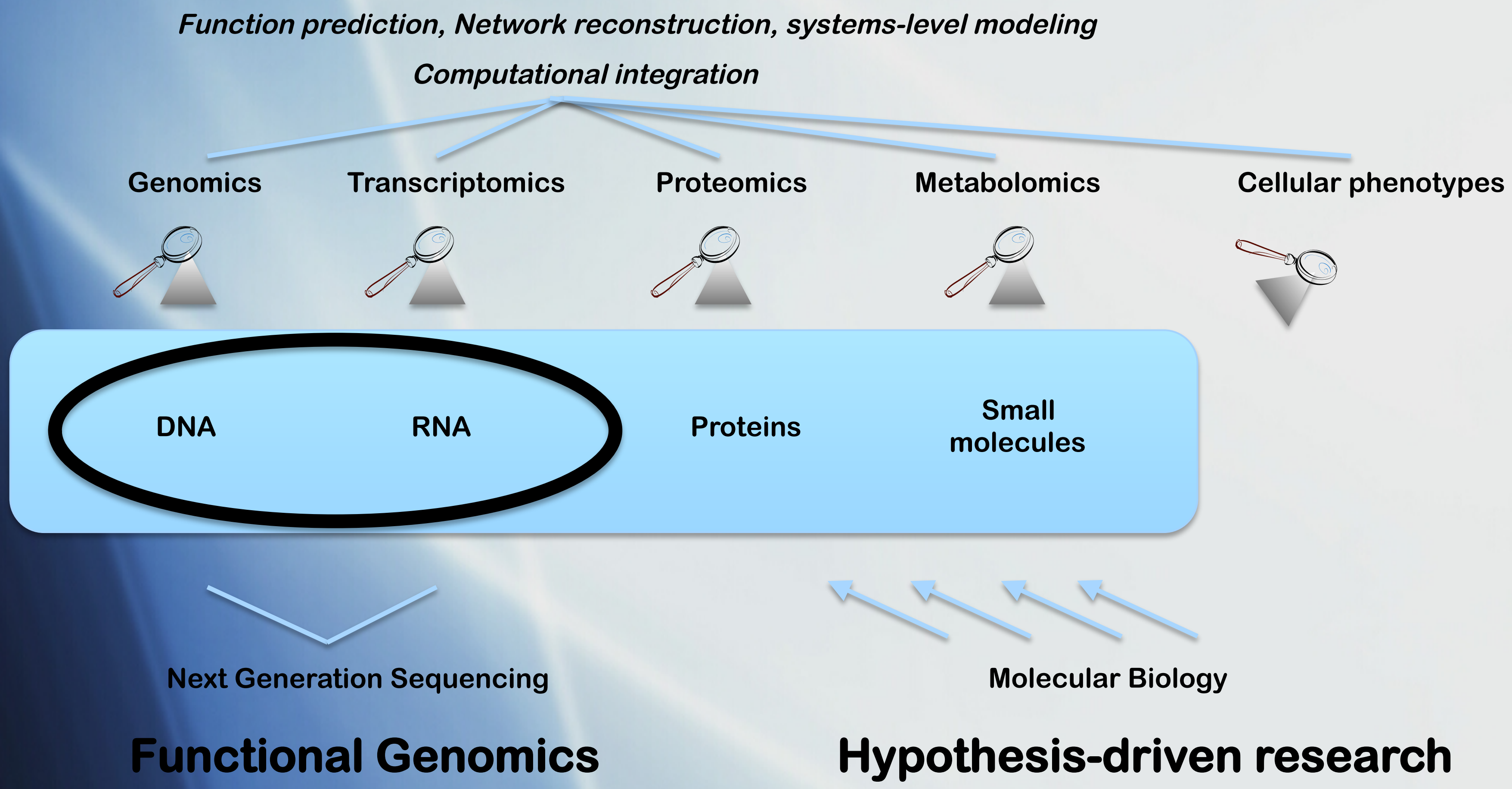
Nov. 2005 - May 2011

A Network of 21 partners

9 Million Euros for 5 years

Ewan Birney
Henning Hermjakob
Jan Ellenberg
Geoffrey Barton
Soren Brunak
Gianni Cesareni
John Hancock
Carl-Henrick Heldin
Edda Klipp
James Adjaye
Erich Nigg
Tomi Makela
Christine Orengo
Christos Ouzounis
Vincent Schachter
Dietmar Schomburg
Eran Segal
Jaak Vilo
Ioannis Xenarios
Alfonso Valencia
Jaap Heringa

Systems Biology



Function prediction, Network reconstruction, systems-level modeling

Computational integration

Genomics

Transcriptomics

Proteomics

Metabolomics

Cellular phenotypes

DNA

RNA

Proteins

Small molecules

Next Generation Sequencing

Molecular Biology

Functional Genomics

Hypothesis-driven research

- **Systems Biology** is the study of the interactions between the components of a biological system, and how these interactions give rise to the function and behaviour of that system (for example, the enzymes and metabolites in a metabolic pathway).

From Wikipedia http://en.wikipedia.org/wiki/Systems_biology

Systems Biology has a wide range of dimensions

STRUCTURE

Gene	Protein	Molecular network	
Organelle	Cell	Cellular Network	
Tissue	Organ	Organism	Ecosystem

FUNCTIONAL

Regulatory	Growth	Metabolic
Electrical	Mechanical	Transport

DATA TO THEORY

Empirical Data	Statistical Modeling	Predictive Modeling
Mathematical	Ontologies	Systems Analysis

EASE OF EXPERIMENTS, MODELS

Simple bacteria: *rapid reproduction*

Eukaryotes: *still relatively simple, rapid reproduction*

Single cells (cell lines, fertilized eggs): *simple system, more difficult to obtain, maintain, artificial environment*

In-vivo studies of model organisms: *complex, more relevant*

Primates, Humans: *highly complex and relevant*

NEED FOR APPROPRIATE TOOLS

High Throughput	Quantitative and Qualitative Results
Single cell data	Organ level data

Systems Biology has a wide range of dimensions

- It is impossible to fully model biological systems, even at a single level.

- It is possible to create useful reductionist models that improve understanding, at several levels:
    - Bioinformatics *(genotype-phenotype)*
    - Systematic biology *(pathways with linear linkage)*
    - Systems biology *(single or interlinked pathways with quantitative and dynamic calculations iterated with experiments)*
    - Physiology with key effects identified *(e.g. Oxford Prof. Noble – Heart modelling)*

Integrative Systems Biology

-Biological data

-Databases

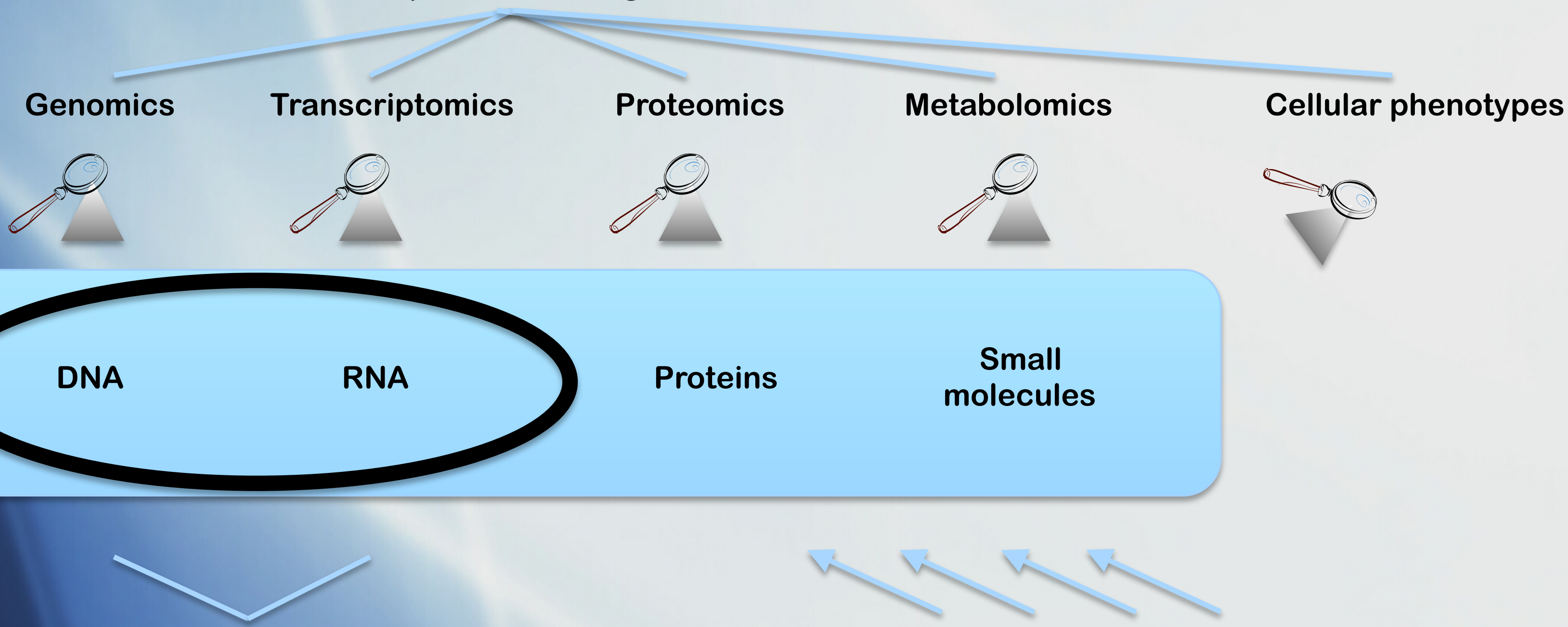
-Analysis tools: Bioinformatics

-Integration into Systems Biology: ENFIN

Systems Biology

Function prediction, Network reconstruction, systems-level modeling

Computational integration



Genomics

Transcriptomics

Proteomics

Metabolomics

Cellular phenotypes

DNA

RNA

Proteins

Small
molecules

Next Generation Sequencing

Molecular Biology

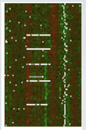
Functional Genomics

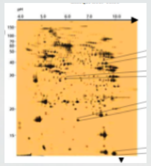
Hypothesis-driven research

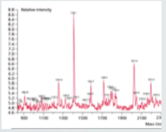
Data acquisition

Technology platforms

Gene expression	DNA microarrays SAGE
Protein levels	2-D gel electrophoresis Mass spectrometry Phosphoproteomics
Small molecule metabolites	Metabolomics
Sugars	Glycomics
Interactomes	Immunopurification Yeast-2-Hybrids







Large-scale perturbation methods

RNAi screens

GFP-tagged recombinant protein expression

Chemicals screens

Data Collection

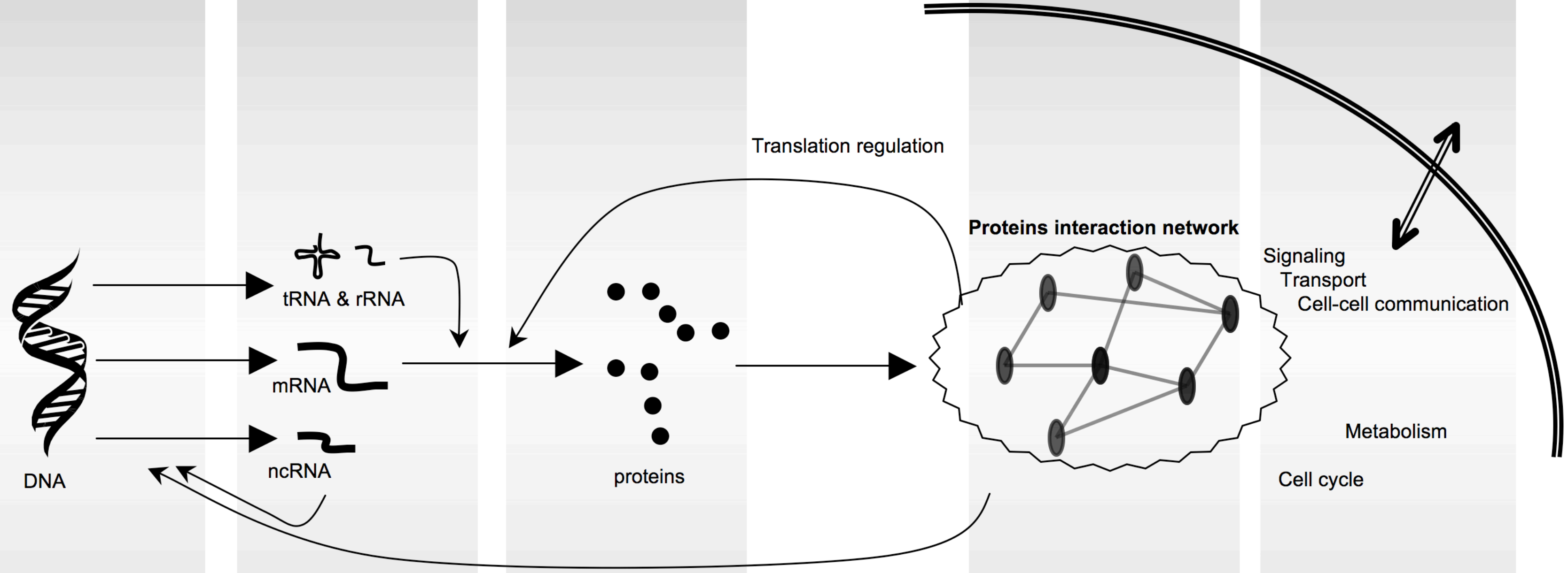
Qualitative Data

Biological Ontologies			
	Small chemical entities	ChEBI	www.ebi.ac.uk/chebi
	Gene Ontology	GO	www.geneontology.org
	Taxonomy	Newt	www.ebi.ac.uk/newt
Literature			
	Genetic diseases	OMIM	www.ncbi.nlm.nih.gov/omim
	Literature	CiteXplore	www.ebi.ac.uk/citexplore
Nucleotides			
	Genomes	Ensembl	www.ensembl.org
	Nucleotide Sequences	EMBL-Bank	www.ebi.ac.uk/embl
Proteins			
	Proteins	UniProt	www.ebi.uniprot.org
	Enzymes	IntEnz	www.ebi.ac.uk/intenz
	Structure	MSD	www.ebi.ac.uk/msd
	Enzymes	BRENDA	www.brenda-enzymes.info
Pathways and Networks			
	Pathways	Reactome	www.reactome.org
	Biological Models	BioModels	www.ebi.ac.uk/biomodels
	Pathways	KEGG	www.genome.jp/kegg
	Pathways	Chillibot	www.chilibot.net
	Pathways	iHOP	www.ihop-net.org/UniPub/iHOP
	Pathways	Panther	www.pantherdb.org/panther
	Pathways	Ingenuity	www.ingenuity.com
	Pathways	GeneGO	www.genego.com
	Pathways	BioCyc	www.biocyc.org
	Transporters	TransportDB	www.membranetransport.org
	Thermodynamics	NIST	xpdb.nist.gov/enzyme_thermodynamic

Quantitative Data

Gene expression	Array Express	www.ebi.ac.uk/aerep
Protein interaction	IntAct	www.ebi.ac.uk/intact
Reaction Kinetics	SABIO-RK	sabiork.villa-bosch.de
Cellular phenotypes	MitoCheck	www.mitocheck.org
Growth phenotypes	ASAP	asap.ahabs.wisc.edu/asap/home.php
Metabolite concentrations	Systomonas	www.systomonas.de
Metabolite concentrations	Escherichia coli Multi-omics Datab	ecoli.iab.keio.ac.jp
Reaction fluxes	Escherichia coli Multi-omics Datab	ecoli.iab.keio.ac.jp
Kinetic Parameters	KMedDB	sysbio.molgen.mpg.de/KMedDB

OVERVIEW



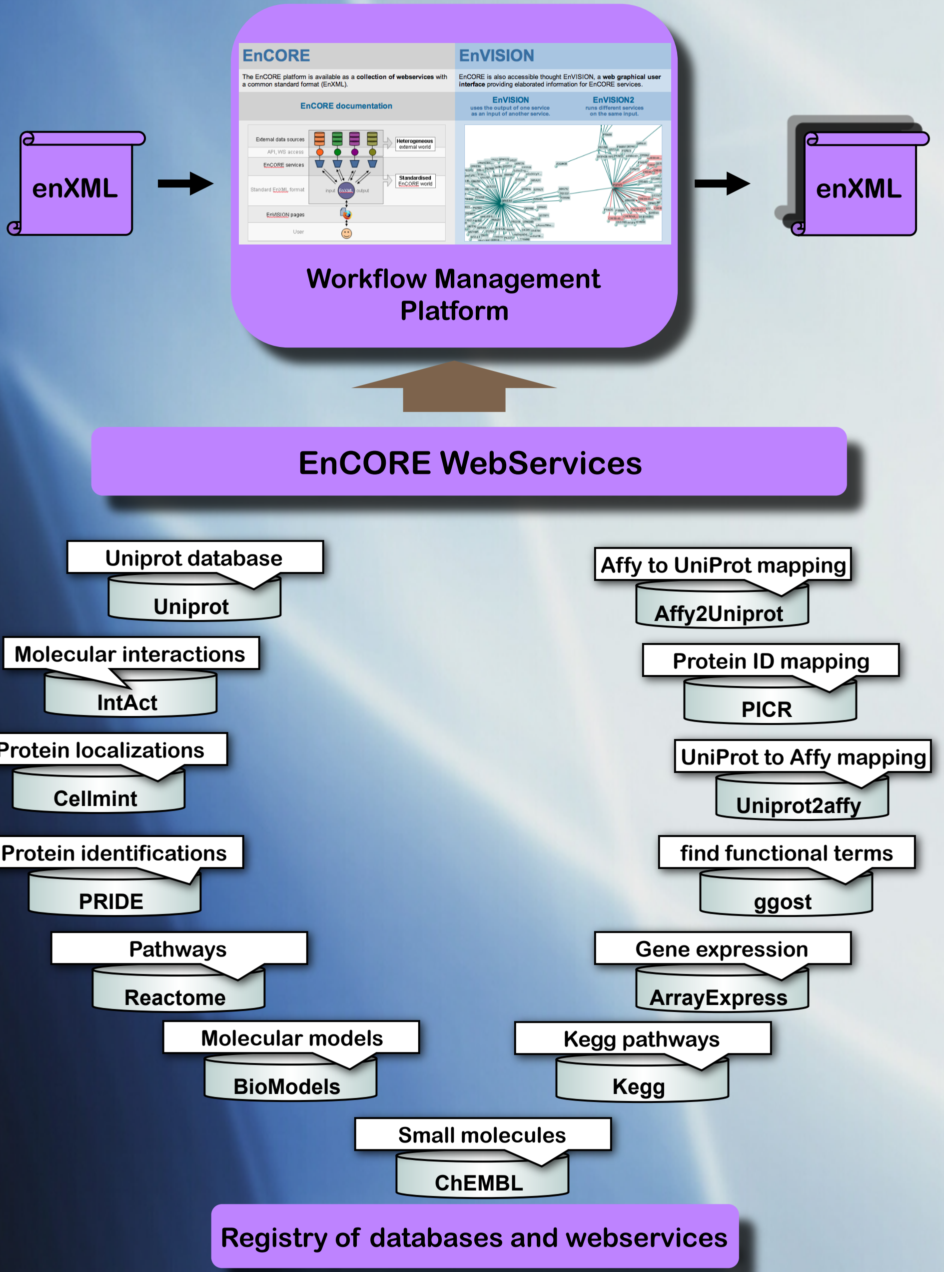
Quantitative		ArrayExpress			SABIO-RK
Qualitative	Ensembl	RNAdb	UniProt, BRENDA	Reactome, MINT, …	

DATABASES

Data Integration

EnCORE - ENVISION

H. Hermjakob (EBI, UK)



EnCORE

The EnCORE platform is available as a **collection of webservices** with a common standard format (EnXML).

EnCORE documentation

EnVISION

EnCORE is also accessible thought EnVISION, a **web graphical user interface** providing elaborated information for EnCORE services.

EnVISION uses the output of one service as an input of another service.

EnVISION2 runs different services on the same input.

Features:

- Open standards (SOAP, WSDL, XML, WS-I Basic Profile 1.0)

- Well-supported technologies (JAX-WS, AJAX)

- Enables data retrieval, analysis tools and algorithms

- Standard interface and data exchange format

- Support of set-based operations

- In-document audit trail

- Modular webservices

Standardization and inter-operability:

- Data standardization:

  Cellular Assays (MIACA)
  Microarray Experiment (MIAME)
  Molecular Interactions (PSI-MI)
  …..

- Standardization of experimental protocols

- Use of controlled vocabulary: Ontologies

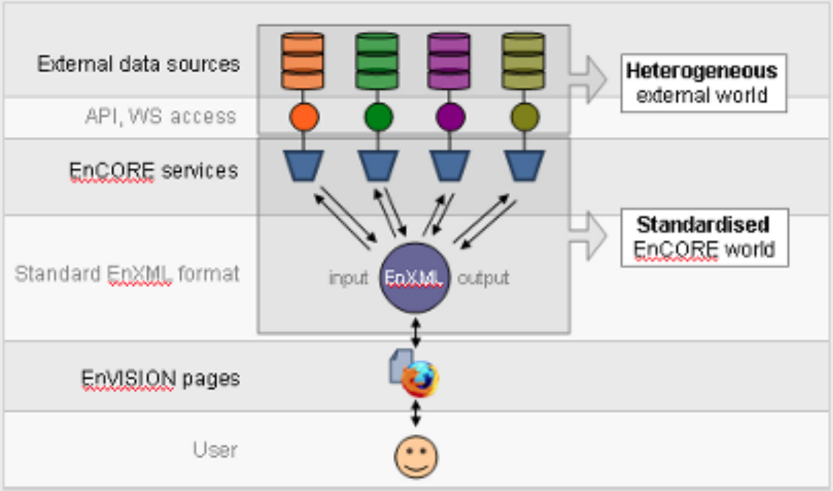
Data Integration

EnCORE - ENVISION

EnCORE

The EnCORE platform is available as a **collection of webservices** with a common standard format (EnXML).

EnCORE documentation



EnVISION

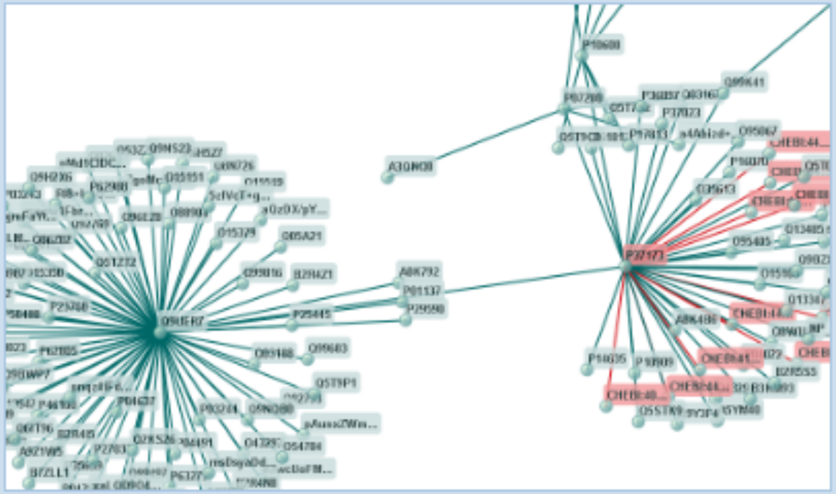
EnCORE is also accessible thought EnVISION, a **web graphical user interface** providing elaborated information for EnCORE services.

EnVISION

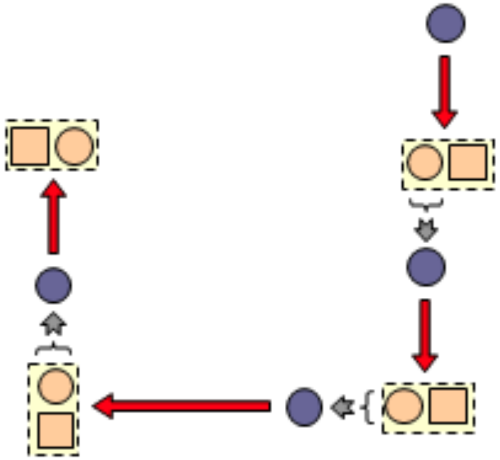
uses the output of one service as an input of another service.

EnVISION2

runs different services on the same input.

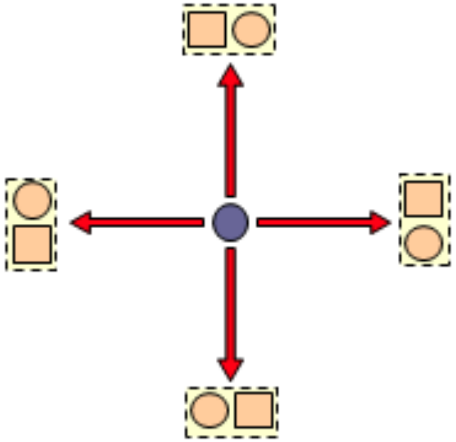


EnVISION



Use the output of one service as an input of another service

EnVISION2

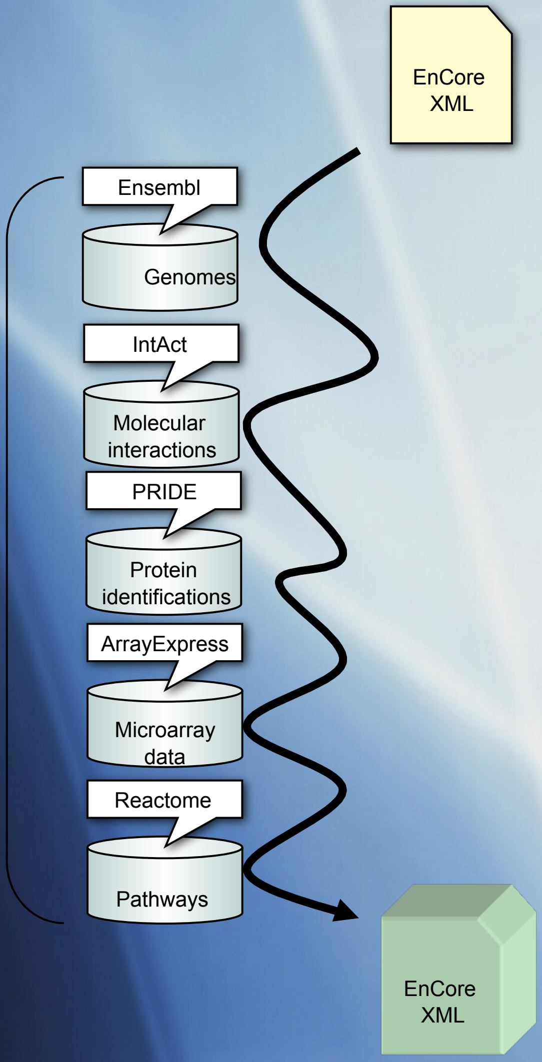


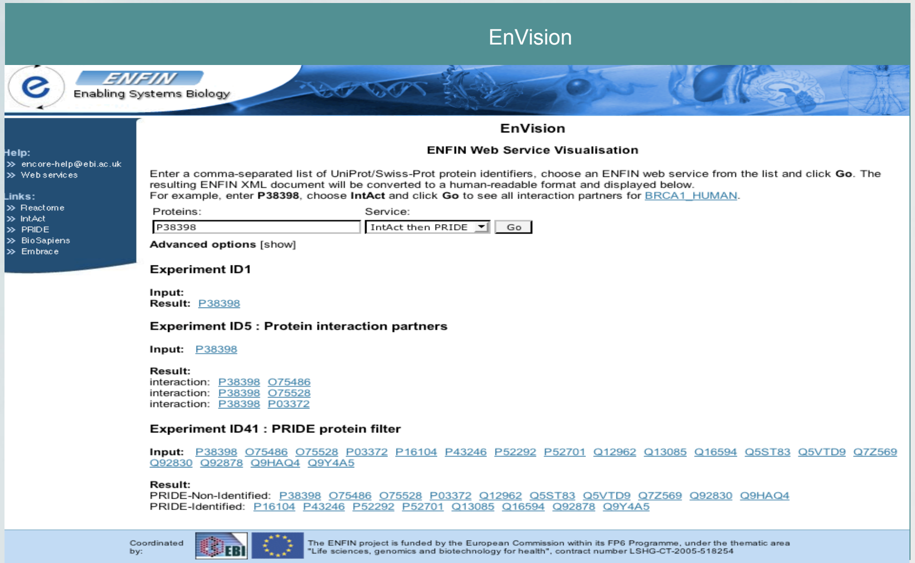
Run different services on the same input

http://www.enfin.org/products/envision/

Workflow Management

EnVISION




EnVision

EnVision

ENFIN Web Service Visualisation

Help:
- encore-help@ebi.ac.uk
- Web services

Links:
- Reactome
- IntAct
- PRIDE
- BioSapiens
- Embrace

Enter a comma-separated list of UniProt/Swiss-Prot protein identifiers, choose an ENFIN web service from the list and click **Go**. The resulting ENFIN XML document will be converted to a human-readable format and displayed below.
For example, enter **P38398**, choose **IntAct** and click **Go** to see all interaction partners for BRCA1_HUMAN.

Proteins: P38398
Service: IntAct then PRIDE

Advanced options [show]

Experiment ID1

Input:
Result: P38398

Experiment ID5 : Protein interaction partners

Input: P38398

Result:
interaction: P38398 O75486
interaction: P38398 O75528
interaction: P38398 P03372

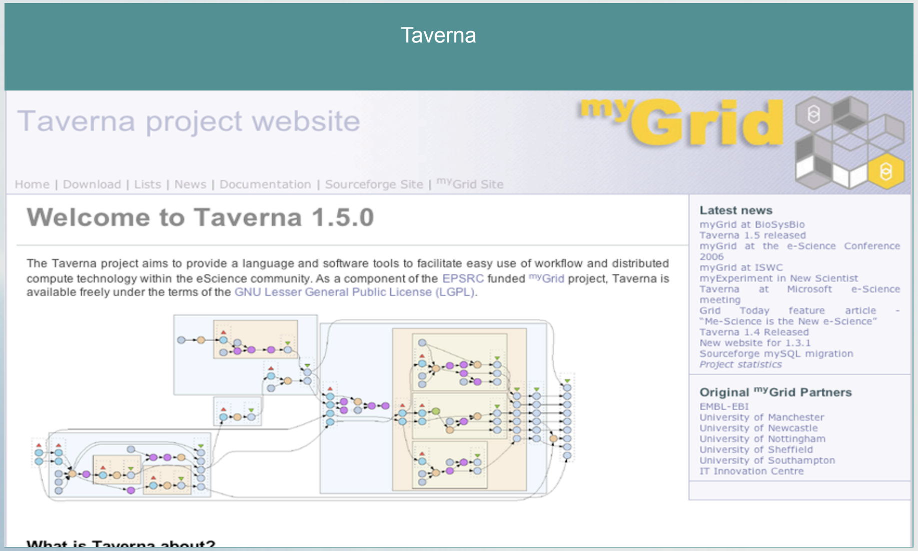
Experiment ID41 : PRIDE protein filter

Input: P38398 O75486 O75528 P03372 P16104 P43246 P52292 P52701 Q12962 Q13085 Q16594 Q5ST83 Q5VTD9 Q7Z569 Q92830 Q92878 Q9HAQ4 Q9Y4A5

Result:
PRIDE-Non-Identified: P38398 O75486 O75528 P03372 Q12962 Q5ST83 Q5VTD9 Q7Z569 Q92830 Q9HAQ4
PRIDE-Identified: P16104 P43246 P52292 P52701 Q13085 Q16594 Q92878 Q9Y4A5

Coordinated by:

The ENFIN project is funded by the European Commission within its FP6 Programme, under the thematic area "Life sciences, genomics and biotechnology for health", contract number LSHG-CT-2005-518254


Taverna

Taverna project website

Welcome to Taverna 1.5.0

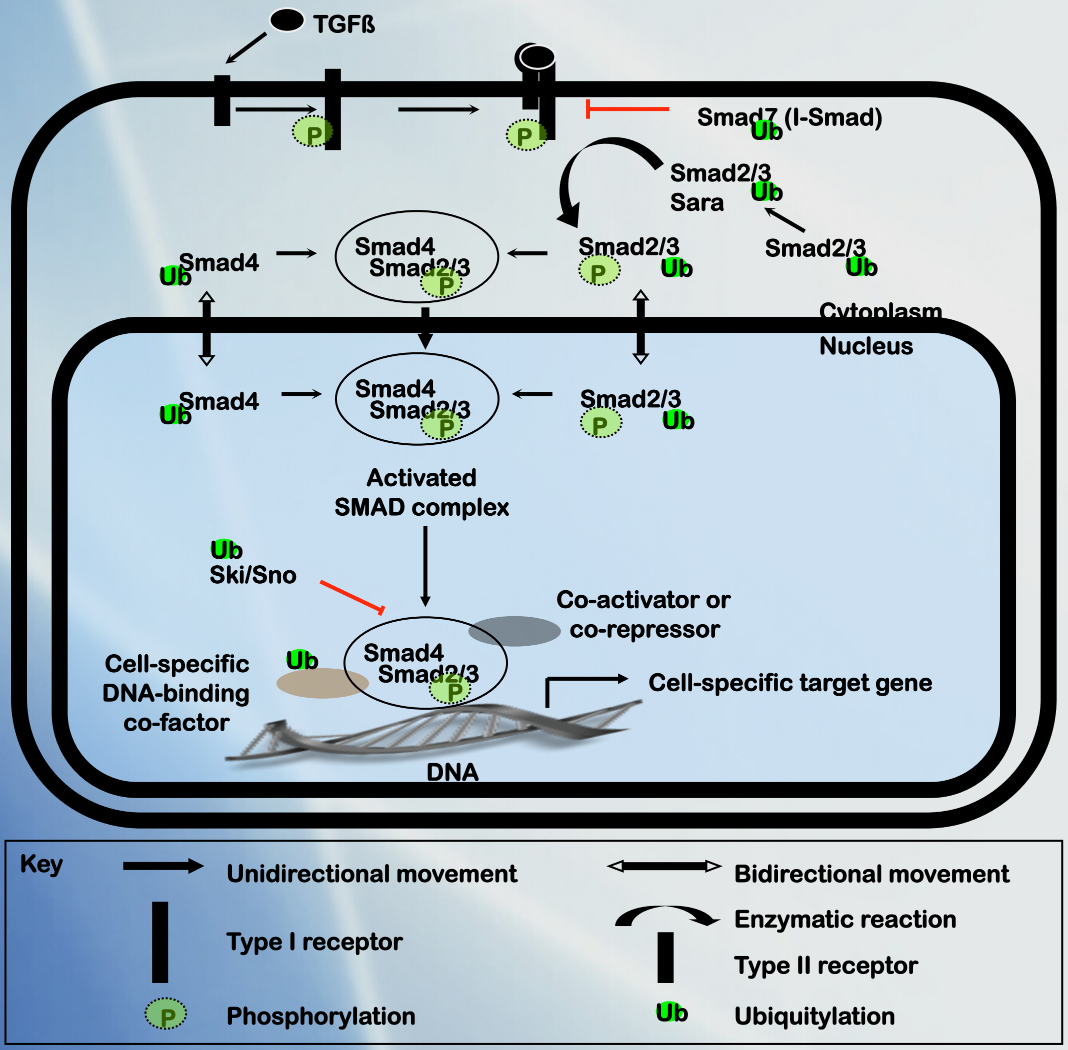
The Taverna project aims to provide a language and software tools to facilitate easy use of workflow and distributed compute technology within the eScience community. As a component of the EPSRC funded myGrid project, Taverna is available freely under the terms of the GNU Lesser General Public License (LGPL).

Latest news
- myGrid at BioSysBio
- Taverna 1.5 released
- myGrid at the e-Science Conference 2006
- myGrid at ISWC
- myExperiment in New Scientist
- Taverna at Microsoft e-Science meeting
- Grid Today feature article - "Me-Science is the New e-Science"
- Taverna 1.4 Released
- New website for 1.3.1
- Sourceforge mySQL migration
- *Project statistics*

Original myGrid Partners
- EMBL-EBI
- University of Manchester
- University of Newcastle
- University of Nottingham
- University of Sheffield
- University of Southampton
- IT Innovation Centre

What is Taverna about?

Example: The TGF-beta pathway



TGFß

P

P

Smad7 (I-Smad)

Ub

Smad2/3

Sara

Ub

Smad2/3

Ub

Ub

Smad4

Smad4

Smad2/3

P

Smad2/3

P

Ub

Cytoplasm

Nucleus

Ub

Smad4

Smad4

Smad2/3

P

Smad2/3

P

Ub

Activated

SMAD complex

Ub

Ski/Sno

Co-activator or

co-repressor

Cell-specific

DNA-binding

co-factor

Ub

Smad4

Smad2/3

P

Cell-specific target gene

DNA

Key

Unidirectional movement

Bidirectional movement

Enzymatic reaction

Type I receptor

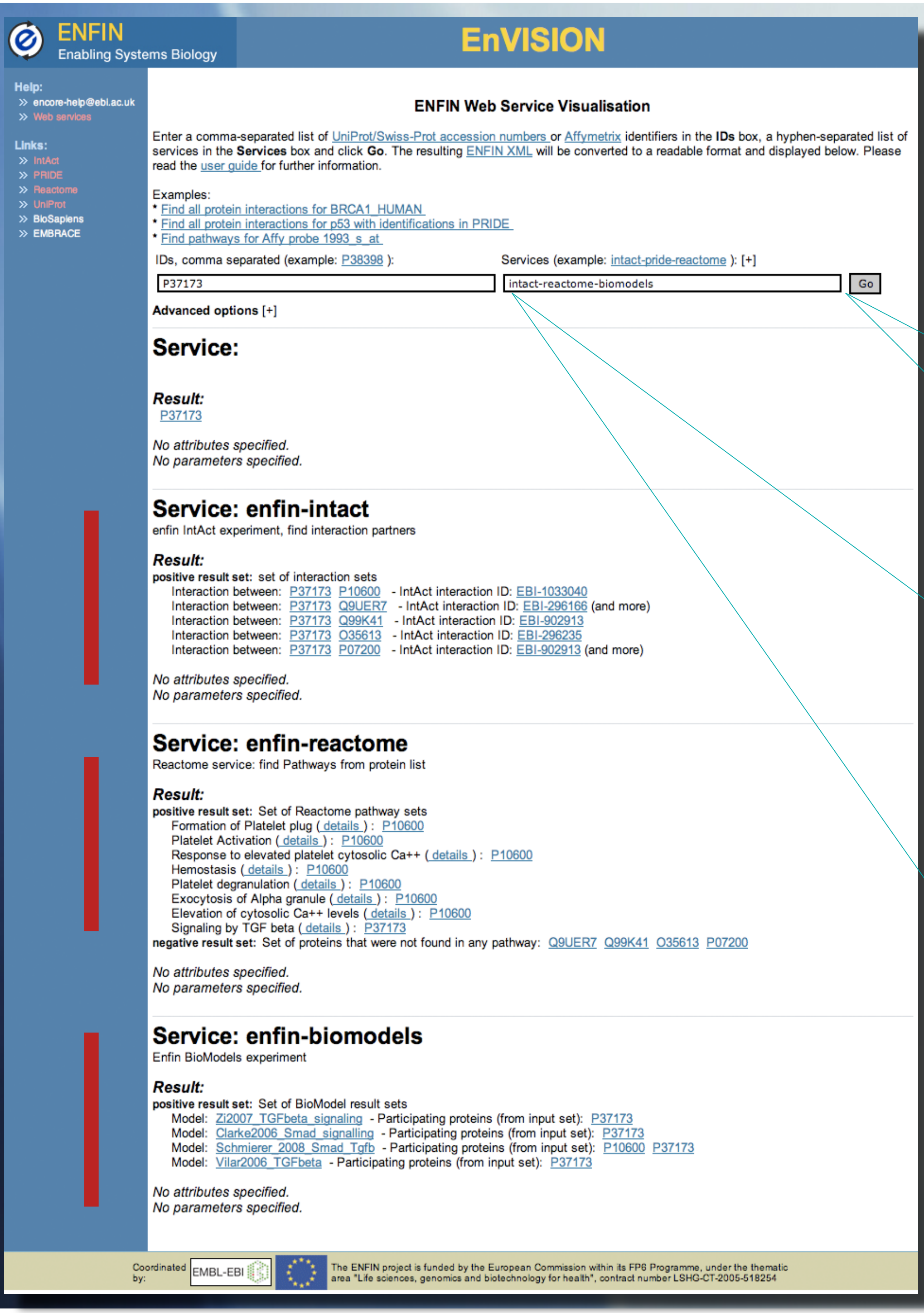
Type II receptor

P Phosphorylation

Ub Ubiquitylation

Workflow Management

EnVISION



ENFIN
Enabling Systems Biology

EnVISION

Help:
- encore-help@ebi.ac.uk
- Web services

Links:
- IntAct
- PRIDE
- Reactome
- UniProt
- BioSapiens
- EMBRACE

ENFIN Web Service Visualisation

Enter a comma-separated list of UniProt/Swiss-Prot accession numbers or Affymetrix identifiers in the **IDs** box, a hyphen-separated list of services in the **Services** box and click **Go**. The resulting ENFIN XML will be converted to a readable format and displayed below. Please read the user guide for further information.

Examples:
- Find all protein interactions for BRCA1_HUMAN
- Find all protein interactions for p53 with identifications in PRIDE
- Find pathways for Affy probe 1993_s_at

IDs, comma separated (example: P38398): P37173

Services (example: intact-pride-reactome): [+] intact-reactome-biomodels

Go

Advanced options [+]

Service:

Result:
P37173

No attributes specified.
No parameters specified.

Service: enfin-intact
enfin IntAct experiment, find interaction partners

Result:
positive result set: set of interaction sets
Interaction between: P37173 P10600 - IntAct interaction ID: EBI-1033040
Interaction between: P37173 Q9UER7 - IntAct interaction ID: EBI-296166 (and more)
Interaction between: P37173 Q99K41 - IntAct interaction ID: EBI-902913
Interaction between: P37173 O35613 - IntAct interaction ID: EBI-296235
Interaction between: P37173 P07200 - IntAct interaction ID: EBI-902913 (and more)

No attributes specified.
No parameters specified.

Service: enfin-reactome
Reactome service: find Pathways from protein list

Result:
positive result set: Set of Reactome pathway sets
Formation of Platelet plug (details) : P10600
Platelet Activation (details) : P10600
Response to elevated platelet cytosolic Ca++ (details) : P10600
Hemostasis (details) : P10600
Platelet degranulation (details) : P10600
Exocytosis of Alpha granule (details) : P10600
Elevation of cytosolic Ca++ levels (details) : P10600
Signaling by TGF beta (details) : P37173
negative result set: Set of proteins that were not found in any pathway: Q9UER7 Q99K41 O35613 P07200

No attributes specified.
No parameters specified.

Service: enfin-biomodels
Enfin BioModels experiment

Result:
positive result set: Set of BioModel result sets
Model: Zi2007_TGFbeta_signaling - Participating proteins (from input set): P37173
Model: Clarke2006_Smad_signalling - Participating proteins (from input set): P37173
Model: Schmierer_2008_Smad_Tgfb - Participating proteins (from input set): P10600 P37173
Model: Vilar2006_TGFbeta - Participating proteins (from input set): P37173

No attributes specified.
No parameters specified.

Coordinated by: EMBL-EBI

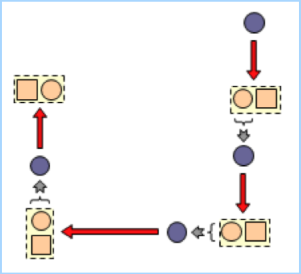
The ENFIN project is funded by the European Commission within its FP6 Programme, under the thematic area "Life sciences, genomics and biotechnology for health", contract number LSHG-CT-2005-518254

Example case:

Query with: **TGF-beta receptor type-2**

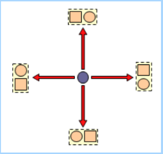
UniProtKB/Swiss-Prot P37173 (TGFR2_HUMAN)

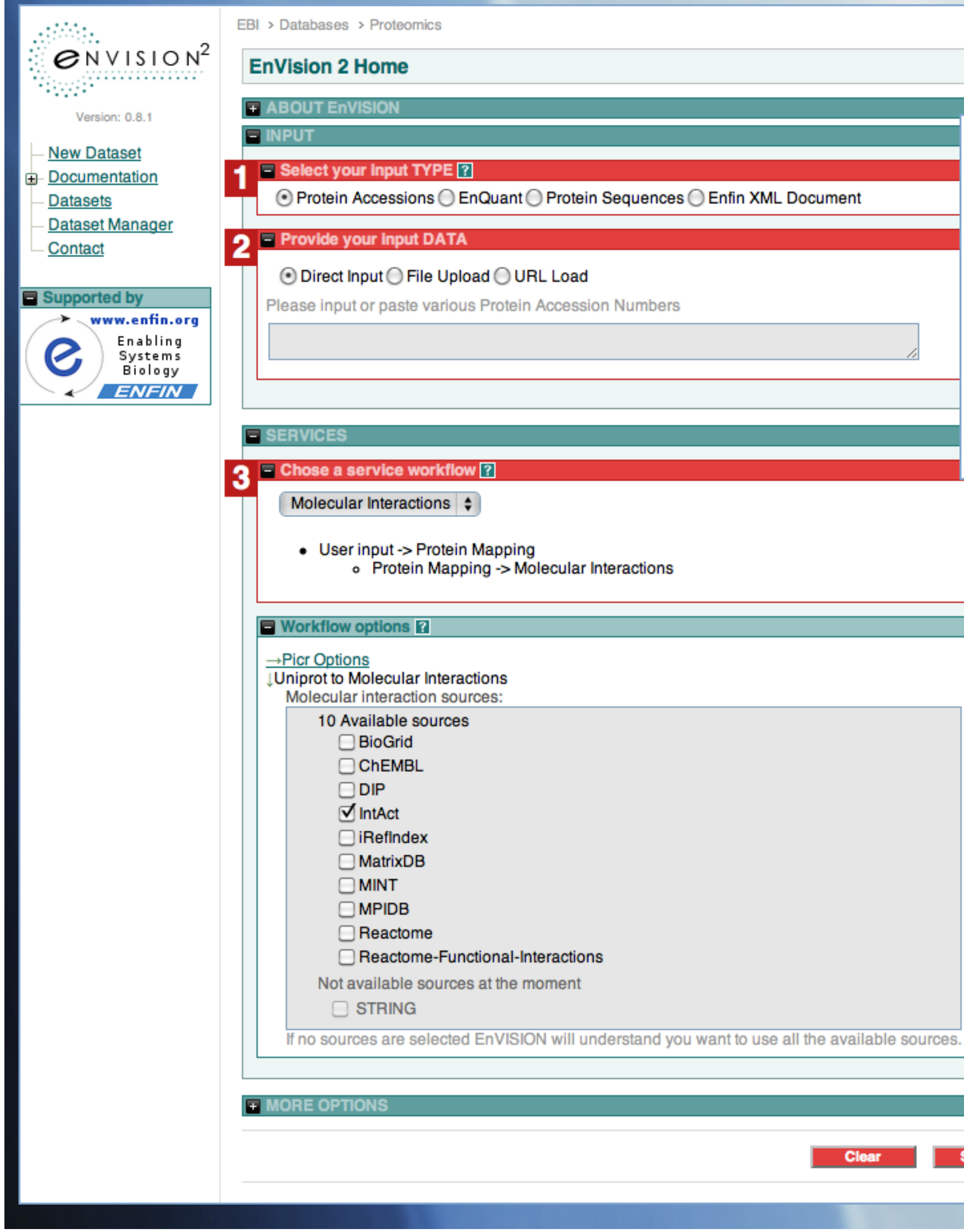
Services (example: intact-pride-reactome): [-]
affy2uniprot - Affy to UniProt mapping
biomodels - find models from the BioModels database
cellmint - groups proteins by localization
ggost - find functional terms
intact - protein interactions
kegg - find KEGG pathways
picr - protein identifier mapping
pride - protein identifications
reactome - find pathways
uniprot2affy - UniProt to Affy mapping
uniprot - retrieve information from UniProt



Workflow Management

EnVISION2





ENVISION2

Version: 0.8.1

- New Dataset
- Documentation
- Datasets
- Dataset Manager
- Contact

Supported by

www.enfin.org Enabling Systems Biology ENFIN

EBI > Databases > Proteomics

EnVision 2 Home

ABOUT EnVISION

INPUT

1 Select your Input TYPE

● Protein Accessions ○ EnQuant ○ Protein Sequences ○ Enfin XML Document

2 Provide your Input DATA

● Direct Input ○ File Upload ○ URL Load

Please input or paste various Protein Accession Numbers

SERVICES

3 Chose a service workflow

Molecular Interactions

- User input -> Protein Mapping
  - Protein Mapping -> Molecular Interactions

Workflow options

→Picr Options

↓Uniprot to Molecular Interactions

Molecular interaction sources:

10 Available sources

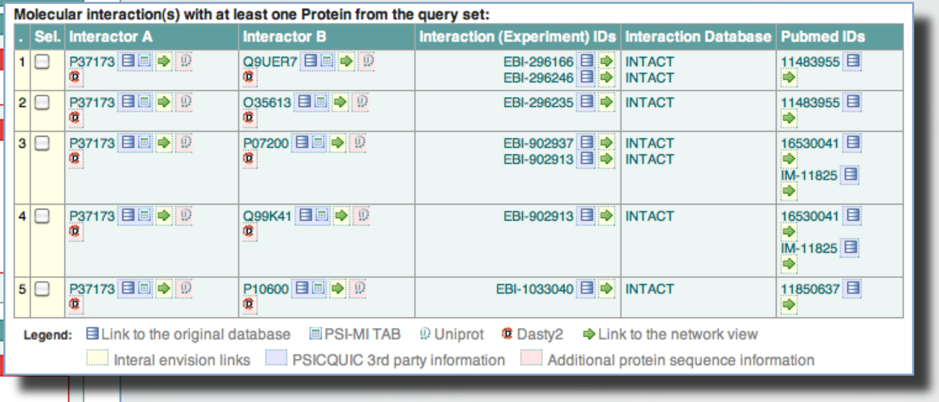
☐ BioGrid
☐ ChEMBL
☐ DIP
☑ IntAct
☐ iRefIndex
☐ MatrixDB
☐ MINT
☐ MPIDB
☐ Reactome
☐ Reactome-Functional-Interactions

Not available sources at the moment

☐ STRING

If no sources are selected EnVISION will understand you want to use all the available sources.

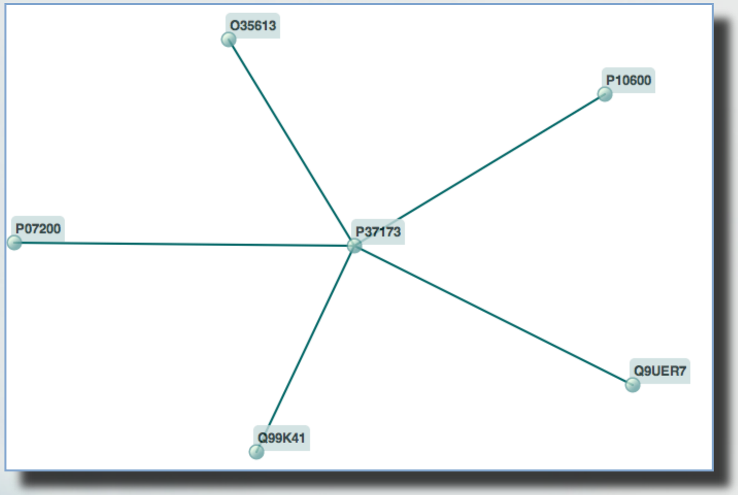
MORE OPTIONS



Molecular interaction(s) with at least one Protein from the query set:

.	Sel.	Interactor A	Interactor B	Interaction (Experiment) IDs	Interaction Database	Pubmed IDs
1	☐	P37173	Q9UER7	EBI-296166 EBI-296246	INTACT INTACT	11483955
2	☐	P37173	O35613	EBI-296235	INTACT	11483955
3	☐	P37173	P07200	EBI-902937 EBI-902913	INTACT INTACT	16530041 IM-11825
4	☐	P37173	Q99K41	EBI-902913	INTACT	16530041 IM-11825
5	☐	P37173	P10600	EBI-1033040	INTACT	11850637

Legend: Link to the original database; PSI-MI TAB; Uniprot; Dasty2; Link to the network view; Interal envision links; PSICQUIC 3rd party information; Additional protein sequence information



Integrative Systems Biology

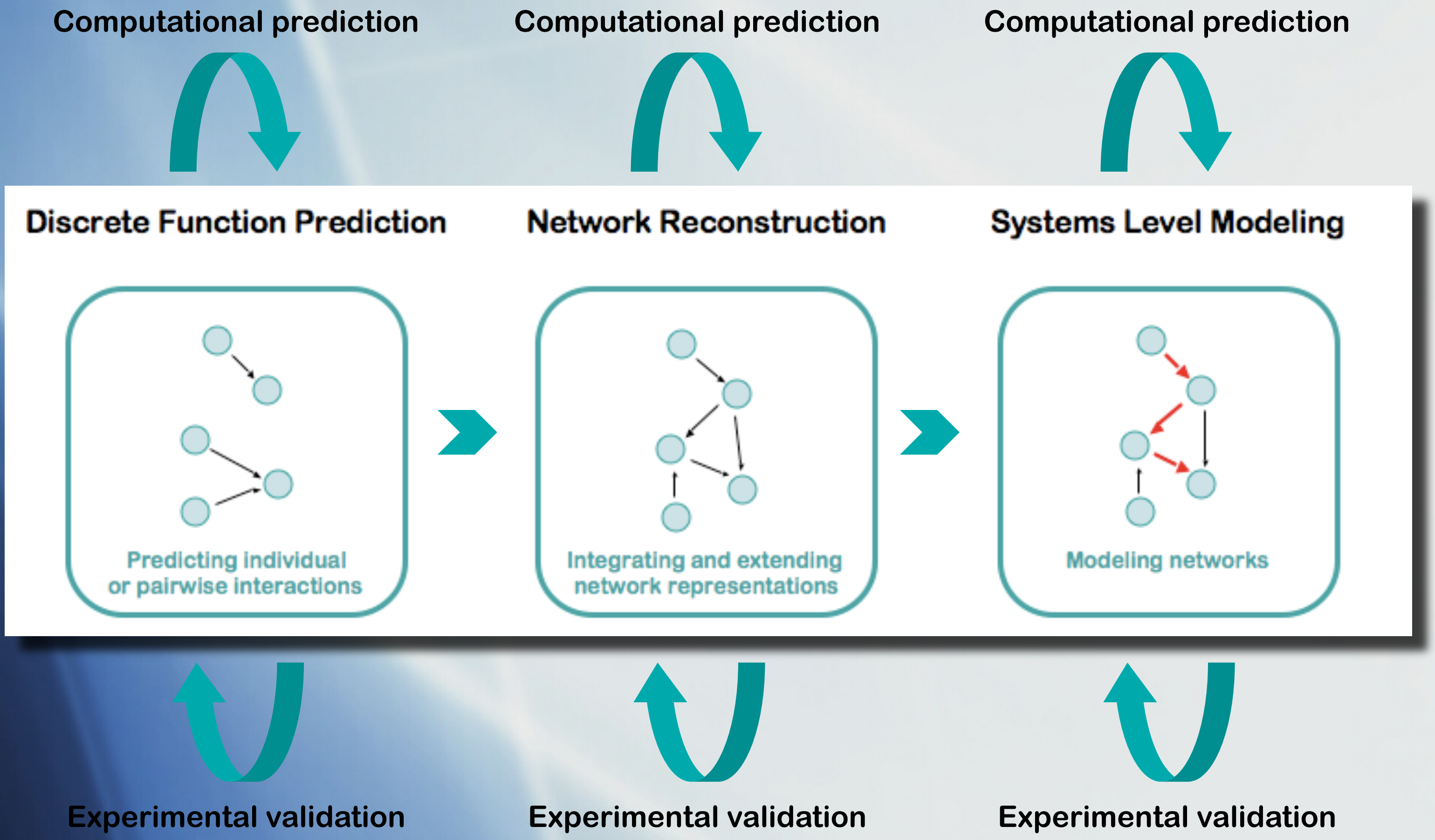
-Biological data

-Databases

-Analysis tools: Bioinformatics

-Integration into Systems Biology: ENFIN

Analysis Platform



Computational prediction
Computational prediction
Computational prediction

Discrete Function Prediction
Network Reconstruction
Systems Level Modeling

Predicting individual or pairwise interactions
Integrating and extending network representations
Modeling networks

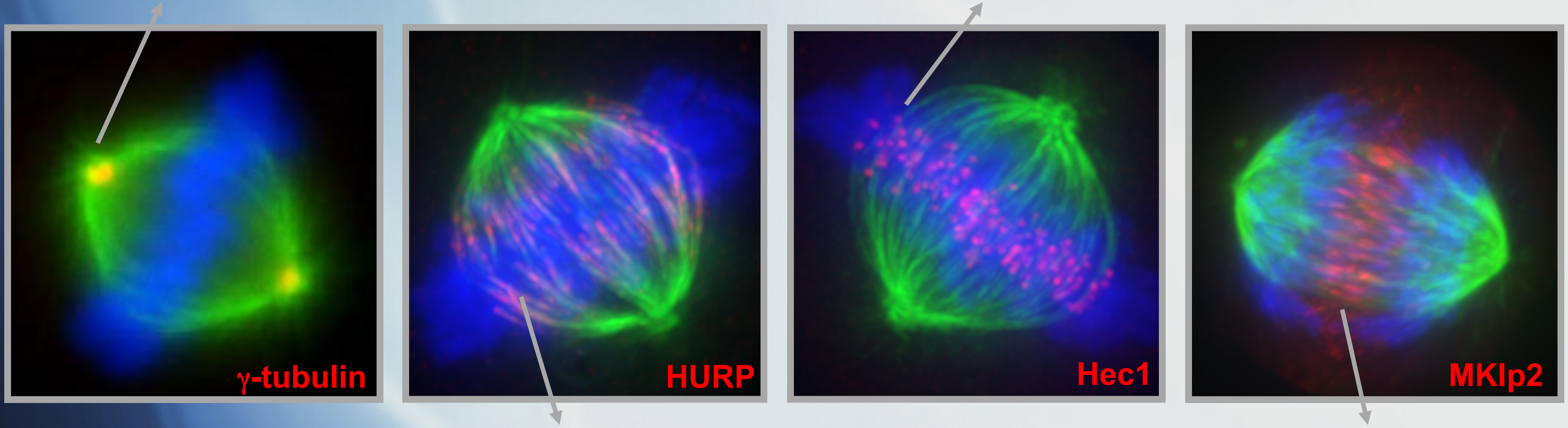
Experimental validation
Experimental validation
Experimental validation

Semantic similarity
Fused-domain architectures
Domain over-representation
Inherited protein interaction
Literature mining iHOP
Vector machine
Neural network

Co-regulated genes (clusters)
Promoter study (modules)
Direct/indirect TF targets

Boolean modeling
Kinetic modeling (ODE)
Constrain-based modeling

Example 1: Where do we expect a protein with a function in mitosis to localize?



Centrosomes
(SPINDLE POLES)

Kinetochores
(CHROMOSOMES)

Microtubules
(SPINDLE)

Microtubules
(CENTRAL SPINDLE)

α-Tubulin → Microtubules

DAPI → DNA

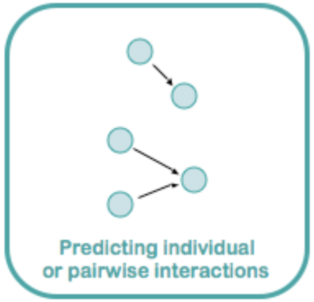
E. Nigg, MPIB, Munich

Analysis Platform

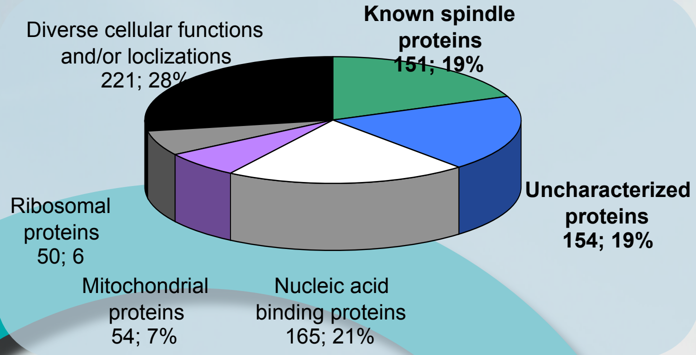
Discrete Function Prediction

E. Nigg (Biozentrum Basel, Switzl.)
A. Valencia (CNIO, Sp.)
C. Orengo (UCL, UK)
S. Brunak (DTU, Denmark)

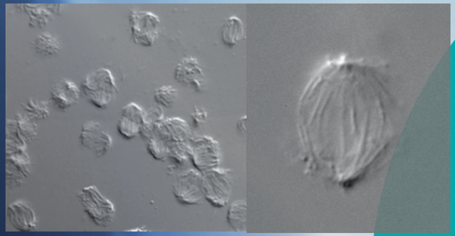
Discrete Function Prediction



Data Analysis



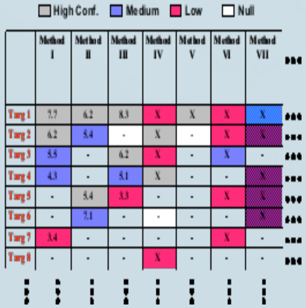
Protein Purification



New Mitotic Spindle Proteins

Function Prediction

Semantic similarity
Fused Domain architectures
Domain over-representation
Inherited Protein Interaction
Literature mining iHOP
Vector Machine
Neural Networks



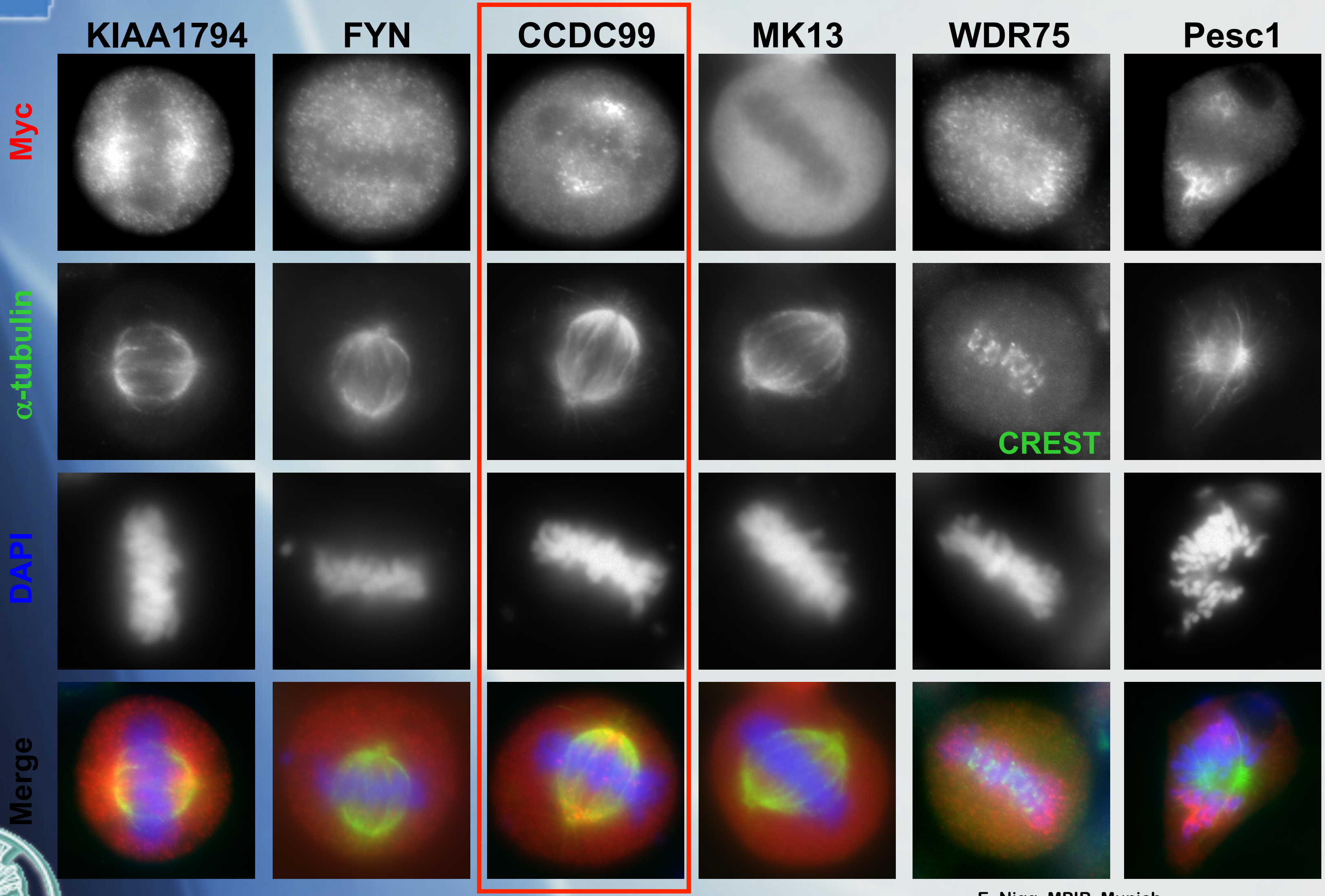
⬜ High Conf. 🟦 Medium 🟥 Low ⬜ Null

	Method I	Method II	Method III	Method IV	Method V	Method VI	Method VII
Targ 1	7.7	6.2	8.3	X	X	X	X
Targ 2	6.2	5.4	-	X	-	X	X
Targ 3	5.5	-	6.2	X	-	X	-
Targ 4	4.3	-	5.1	X	-	-	X
Targ 5	-	5.4	3.3	-	-	X	X
Targ 6	-	7.1	-	-	-	-	X
Targ 7	3.4	-	-	-	-	X	-
Targ 8	-	-	-	X	-	-	-

Validation of Candidates

Silljé and Nigg, Methods 2006
Sauer et al., Mol Cell Proteomics 2005
Santamaria et al., Current Biol. 2008

Protein localization: Immunofluorescence Microscopy



E. Nigg, MPIB, Munich

Analysis with Workflow

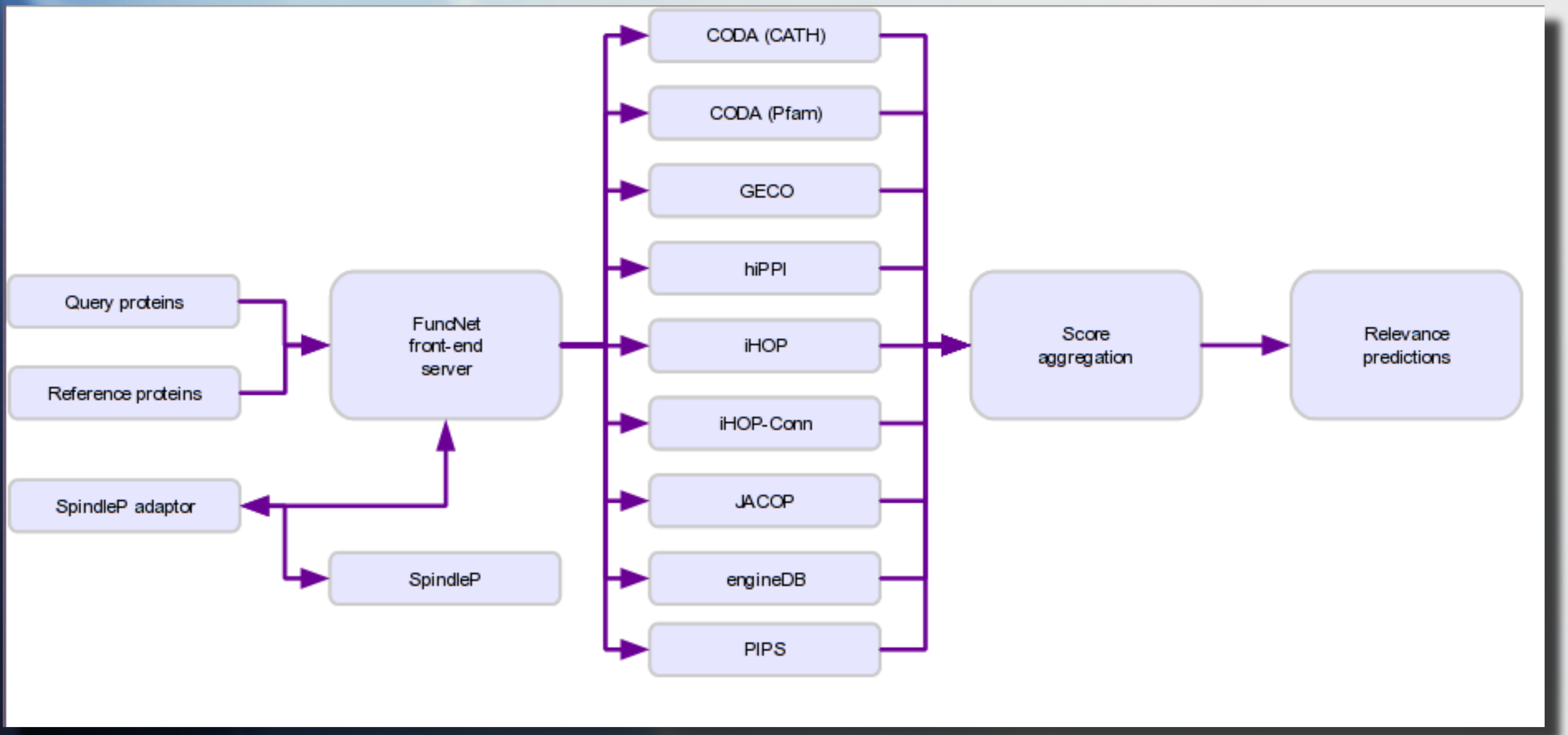
Discrete Function Prediction

C. Orengo (UCL)

FuncNet

FUNCTIONAL ANALYSIS OF PROTEINS

www.funcNet.eu



Kahlem et al., Comptes Rendus Biologies 2009

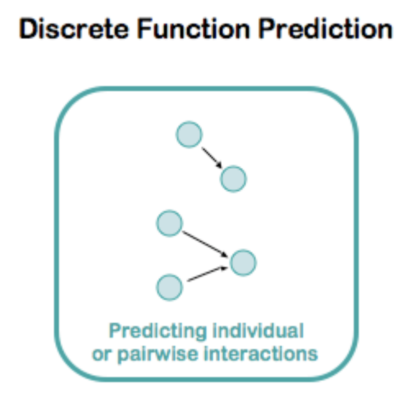
Analysis Platform

Protein Interaction Networks

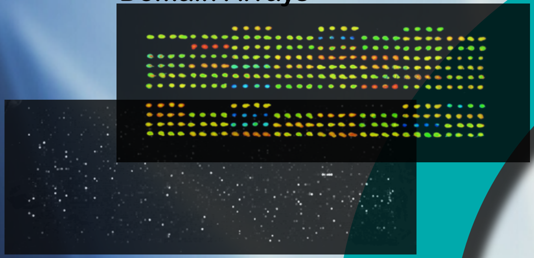
G. Cesareni (Uni Rome, It.)

S. Brunak (DTU, Denmark)

Discrete Function Prediction

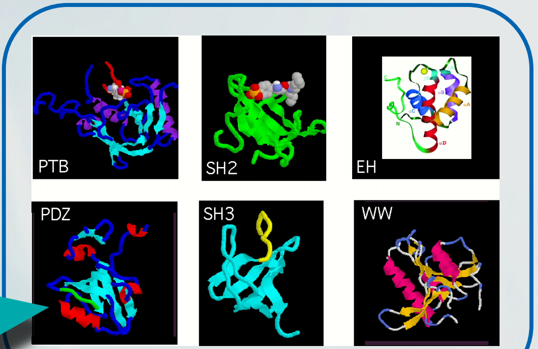


Domain Arrays



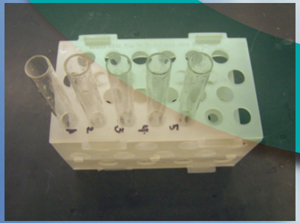
Peptide Arrays

SH2 Domain Interaction Network



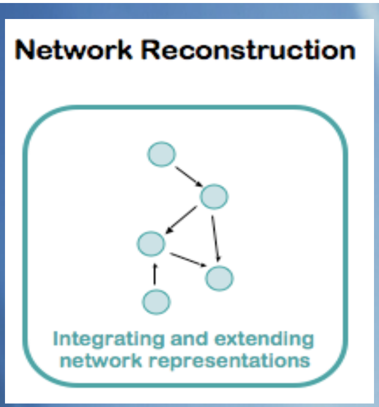
G. Cesareni, Univ. Tor Vergata, Rome, Italy

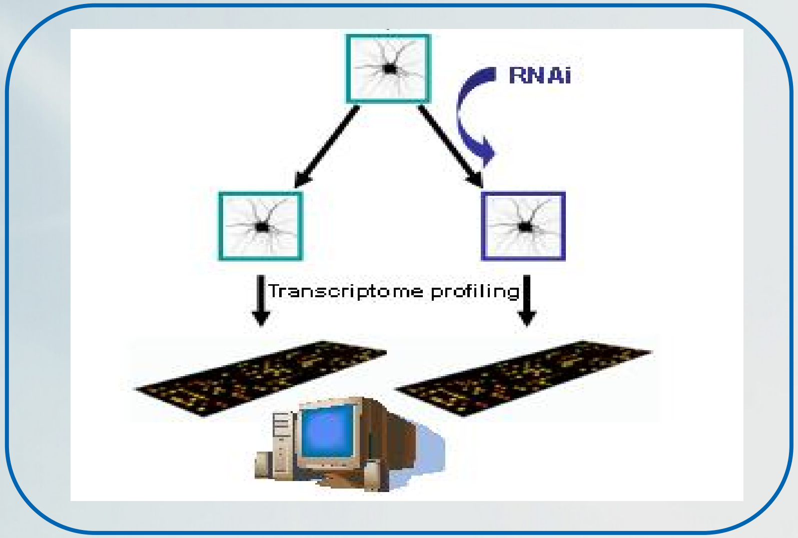
Prediction of new interactions



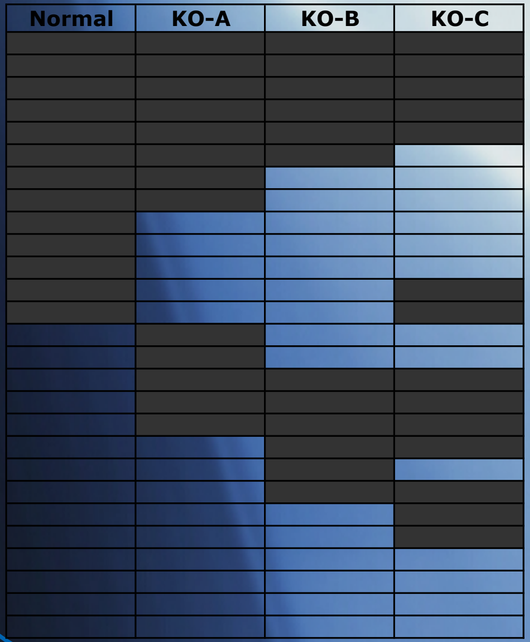
Validation of Candidates:
Pull-down Assay

Kiemer et al., Proteomics 2007





Differential gene expression



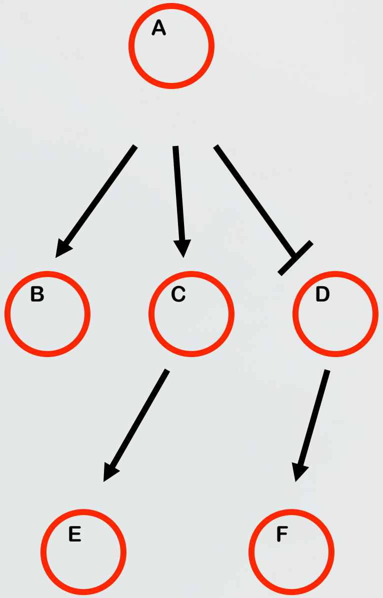
Prediction of regulatory networks

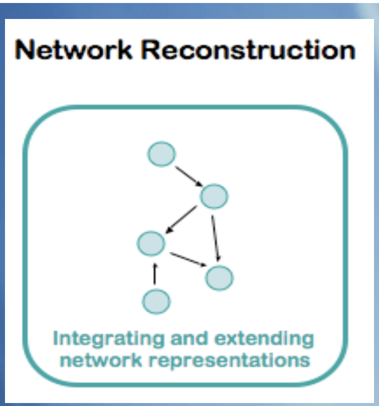
Unchanged expression

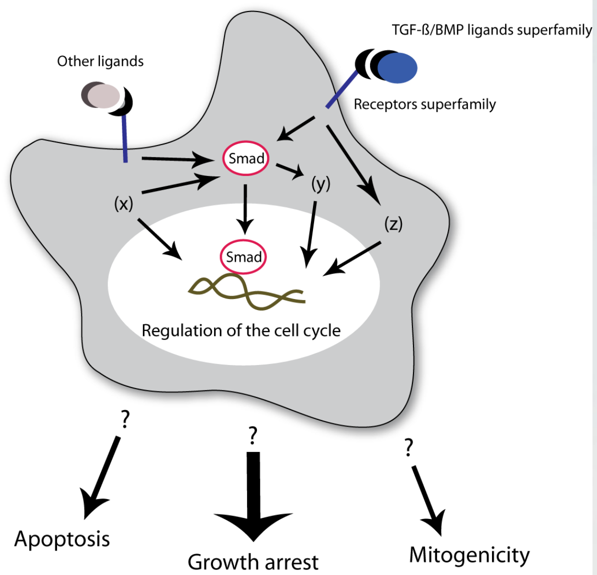
Set of down-regulated genes

Set of up-regulated genes

Unchanged expression







Co-expression of TGFβ target gene(s)

across multiple datasets

http://kotkas.ebc.ee/enfin/tgfb/MEM/matrix.cgi?affy=205430_at&head=25&org=hsapiens&pathway=TGF-BETA&best=10

CLUST(*) --> 02060 00519 08820 08635 08570 01763 04485 06565 00485 01114 00865 02524 00222 02401 01851 01293 00012 00001 02935 02915 00273 08251 00813 07646 05444 05333 02807 00772 002

205430_at(BMP5)
205431_s_at(BMP5)
221037_s_at(SLC25A31)
207489_at(N/A)
216529_at(N/A)
207415_at(PLA2R1)
221137_at(N/A)
215532_x_at(ZNF492)
207479_at(N/A)
216480_x_at(MLLT10)
208045_at(N/A)
215929_at(N/A)
207815_at(PF4V1)
211522_s_at(GNRHR)
216437_at(EPC1)
207596_at(Q9P181_HUMAN)
217375_at(N/A)
220897_at(Q9HAJ0_HUMAN)
204052_s_at(SFRP4)
216408_at(OR2B2)
215743_at(NMT2)
220700_at(N/A)
207261_at(CNGA3)
210661_at(GLRA3)
207294_at(AGTR2)

[send to GOST] [send to URLMAP] [send to TEXT] [get FULL LIST]

	order	display value	display color
Pearson	○	○	○
Rank	●	●	●

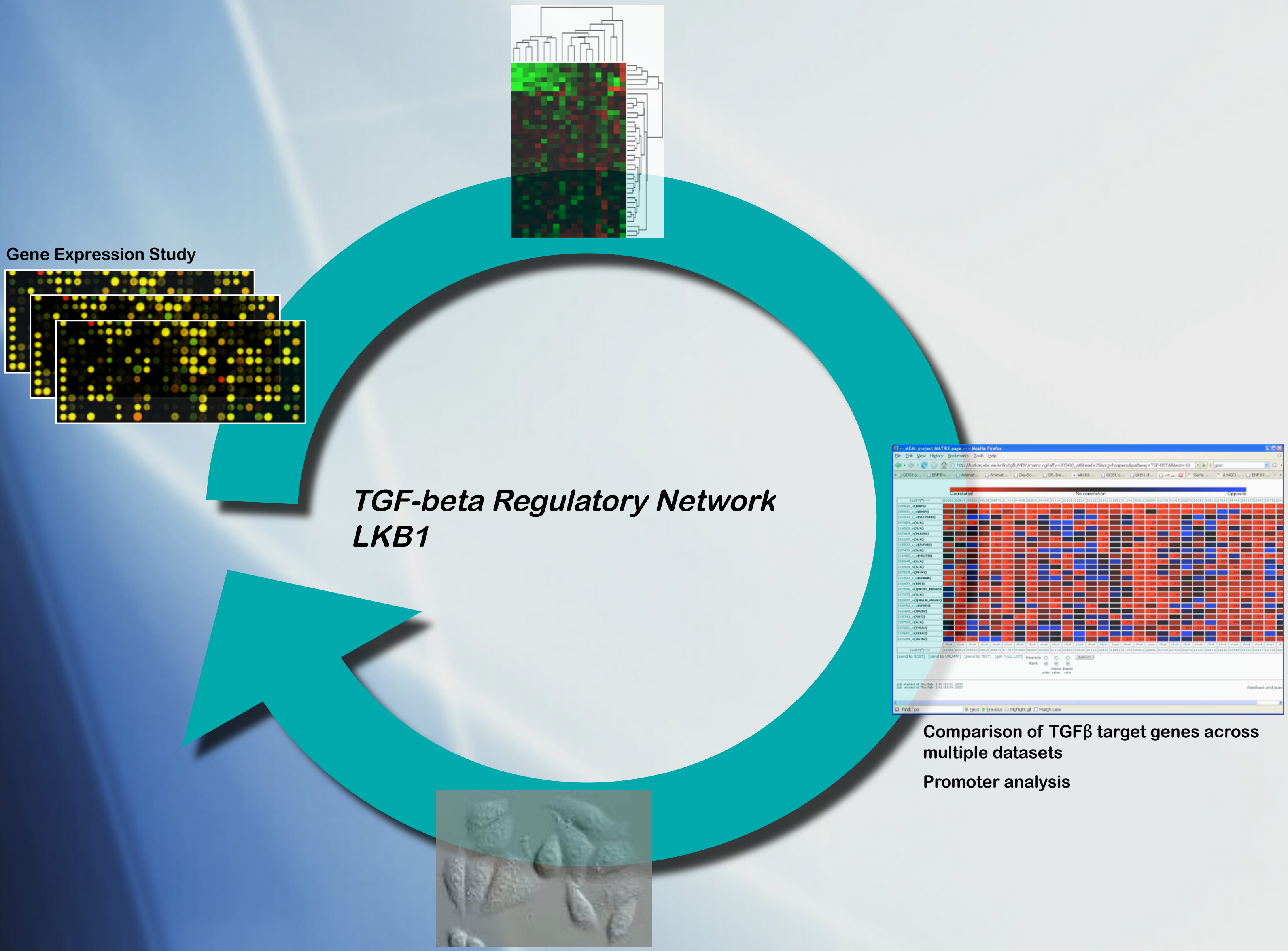
job started at Thu Feb 1 01:13:51 2007
job ended at Thu Feb 1 01:13:54 2007

Find: ppi Next Previous Highlight all ☐ Match case

J. Vilo, EGeen, Estonia

Analysis Platform

Network Reconstruction



Data Analysis

Gene Expression Study

TGF-beta Regulatory Network
LKB1

Comparison of TGFβ target genes across multiple datasets

Promoter analysis

Validation of Candidates: RNAi

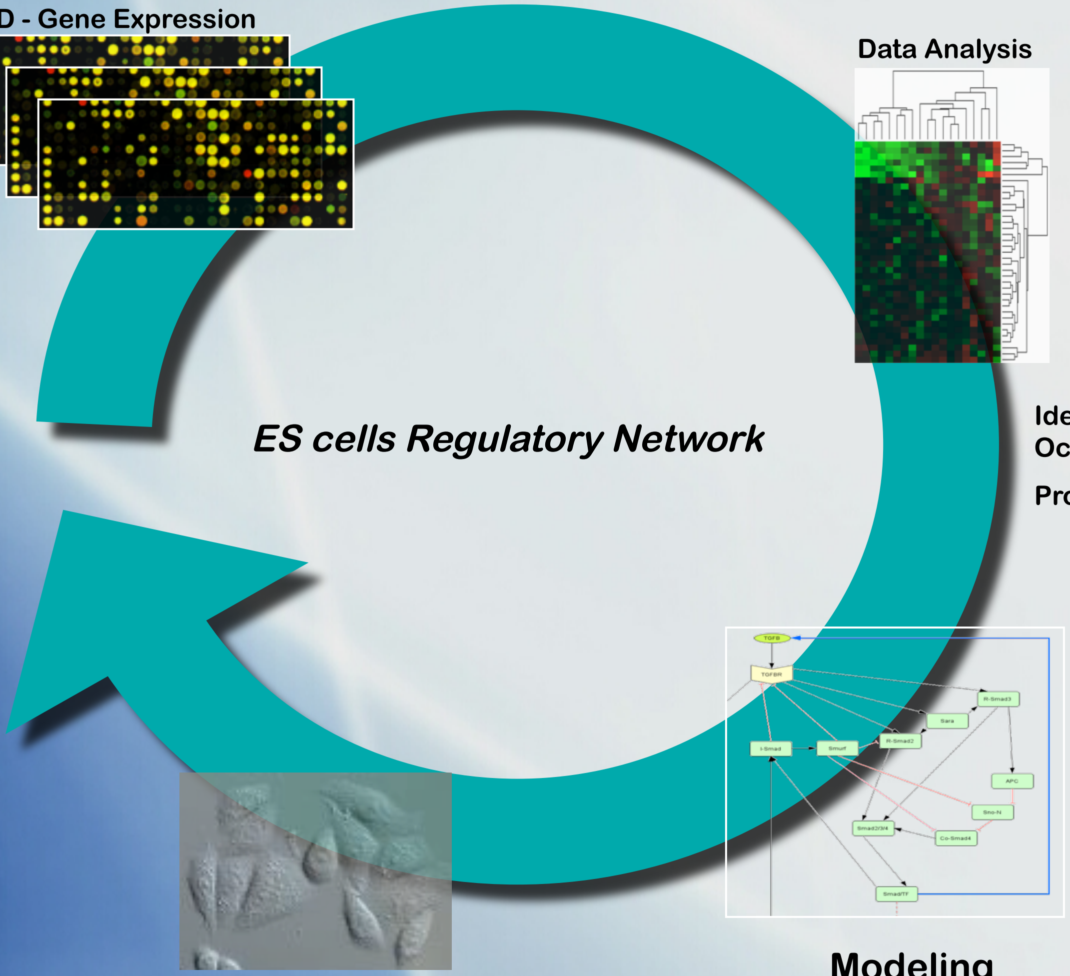
Analysis Platform

Network Reconstruction

J. Adjaye (MPI, Germany)

J. Vilo (QURETEC)

I. Xenarios (SIB, Switzl.)



ChIP-chip

KD - Gene Expression

Data Analysis

ES cells Regulatory Network

Identification of target genes of Oct4

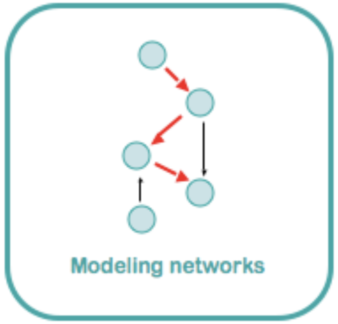
Promoter analysis

Modeling

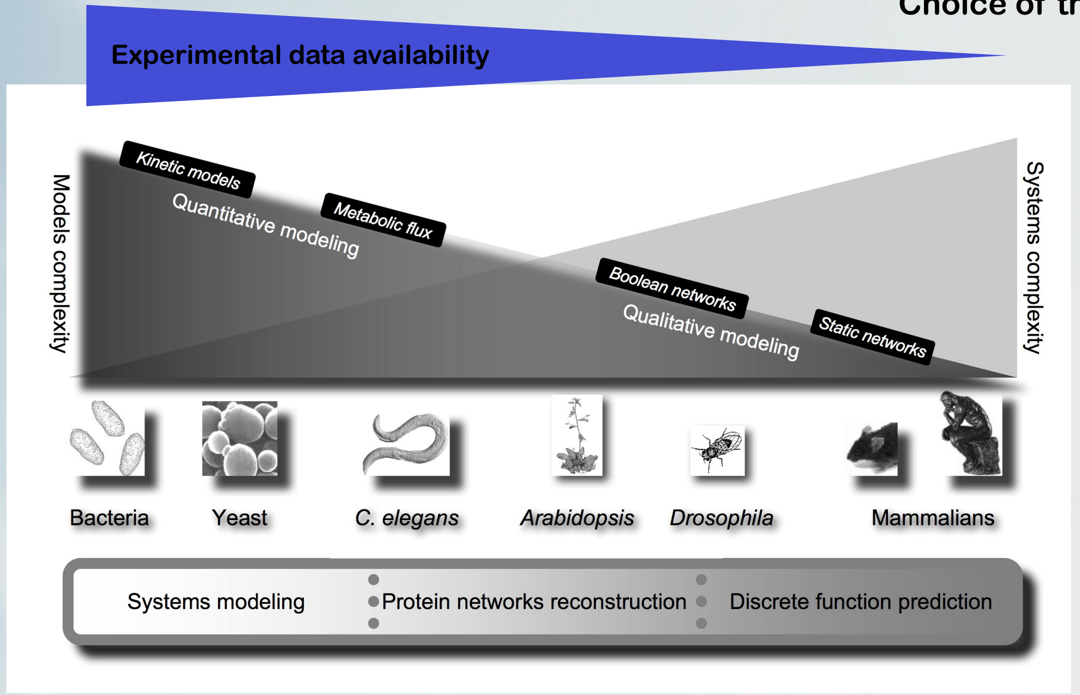
Validation of new targets

Jung et al., PlosONE 2010

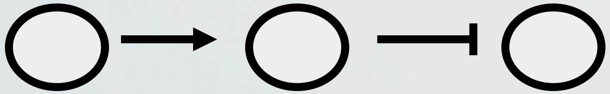
Systems Level Modeling



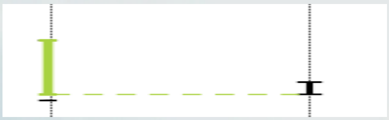
Choice of the system



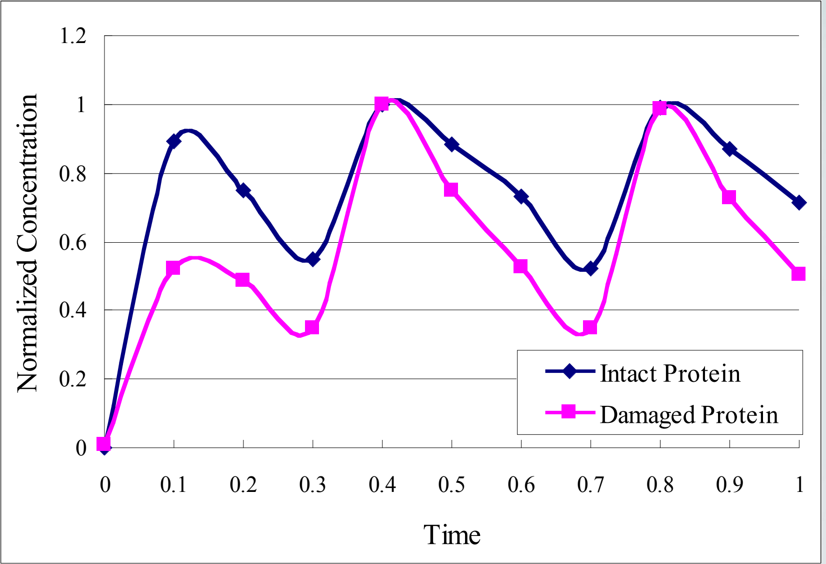
Modeling methods



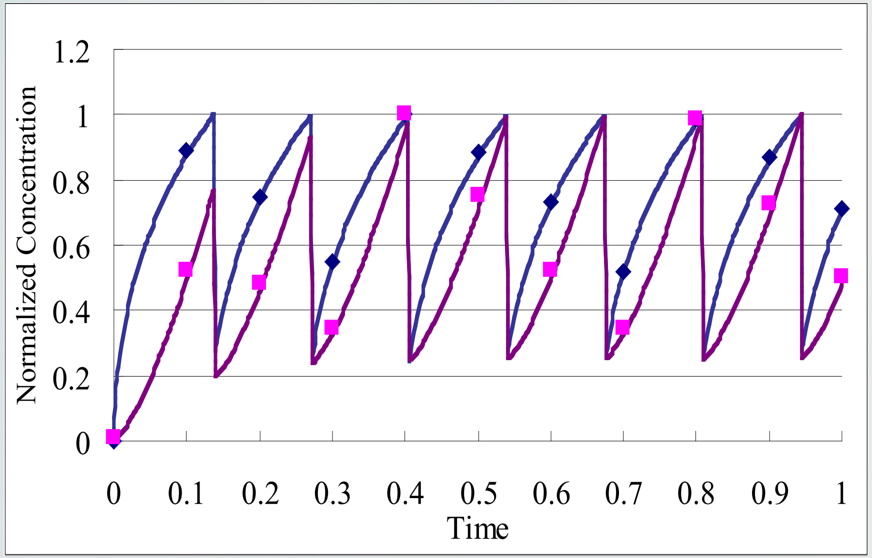
. Kinetic modeling	A	v1	B	v2	C
. Boolean modeling	1		1		0
. Constrain-based modeling					



Need for informatics to understand biology



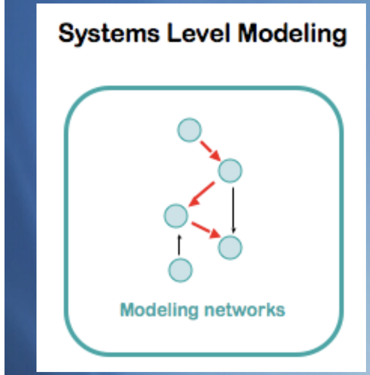
Wrong

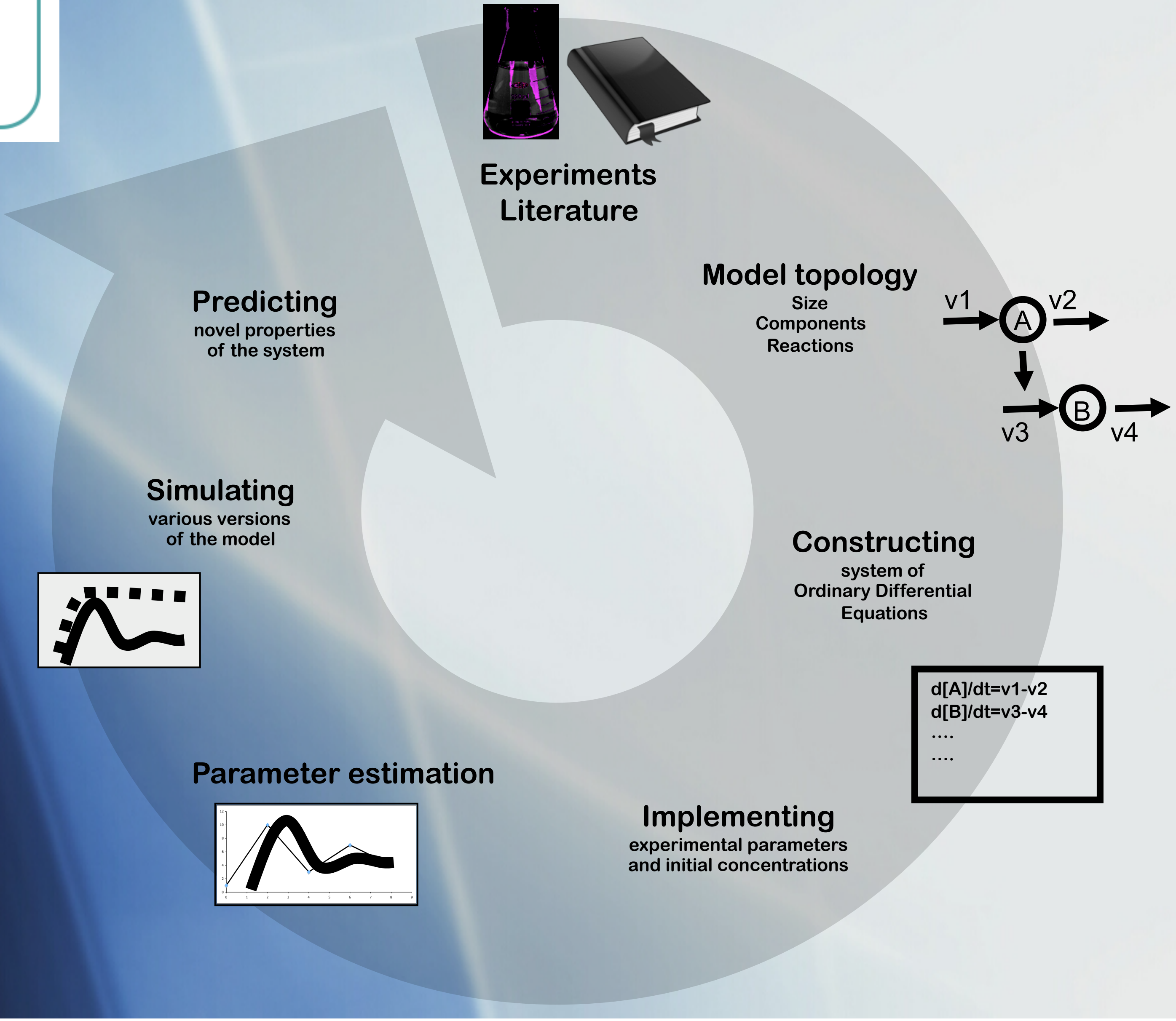


Real

From Zi Z. MPIMG, Berlin, Germany

Constructing a mathematical model of a biological system





Experiments
Literature

Model topology
Size
Components
Reactions

Constructing
system of
Ordinary Differential
Equations

d[A]/dt=v1-v2
d[B]/dt=v3-v4
....
....

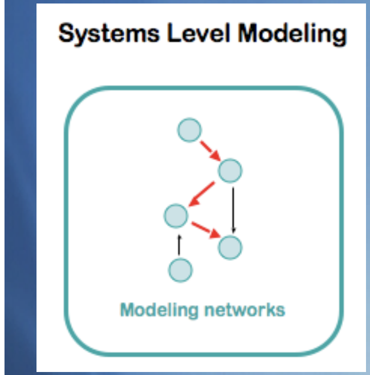
Implementing
experimental parameters
and initial concentrations

Parameter estimation

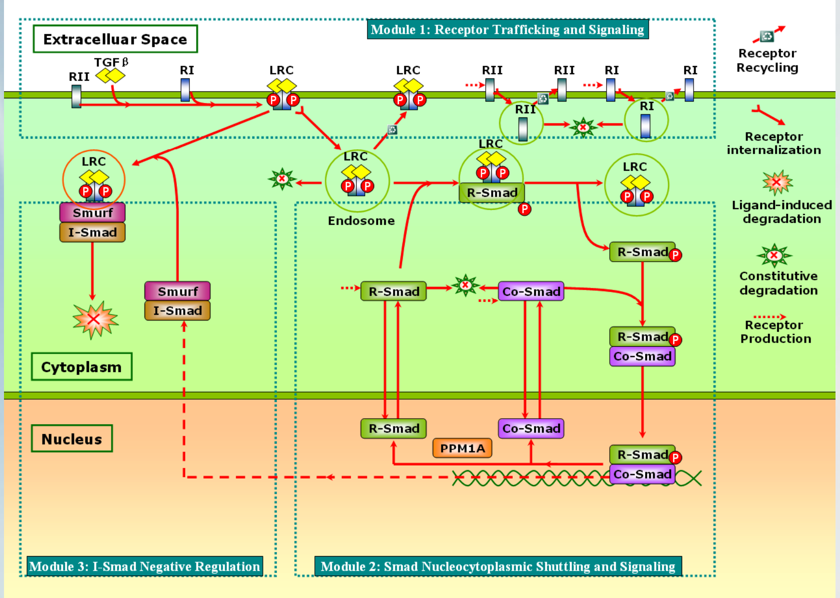
Simulating
various versions
of the model

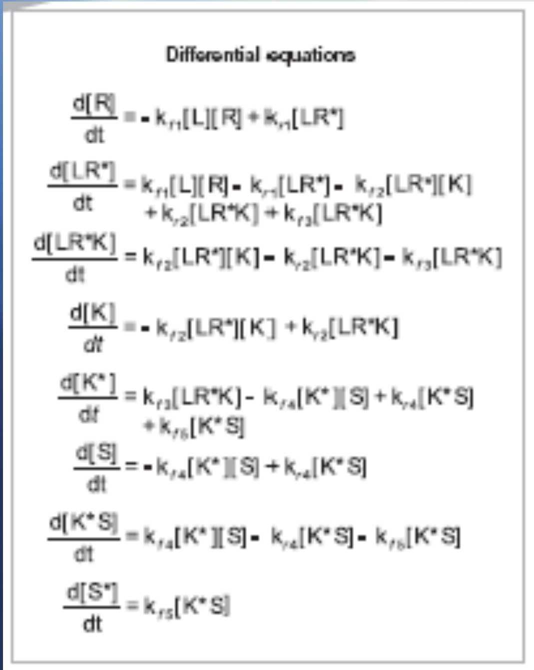
Predicting
novel properties
of the system

Kinetic modeling



Modeling Signaling pathways: TGF-beta





Differential equations

$$\frac{d[R]}{dt} = -k_{f1}[L][R] + k_{r1}[LR^*]$$

$$\frac{d[LR^*]}{dt} = k_{f1}[L][R] - k_{r1}[LR^*] - k_{f2}[LR^*][K] + k_{r2}[LR^*K] + k_{f3}[LR^*K]$$

$$\frac{d[LR^*K]}{dt} = k_{f2}[LR^*][K] - k_{r2}[LR^*K] - k_{f3}[LR^*K]$$

$$\frac{d[K]}{dt} = -k_{f2}[LR^*][K] + k_{r2}[LR^*K]$$

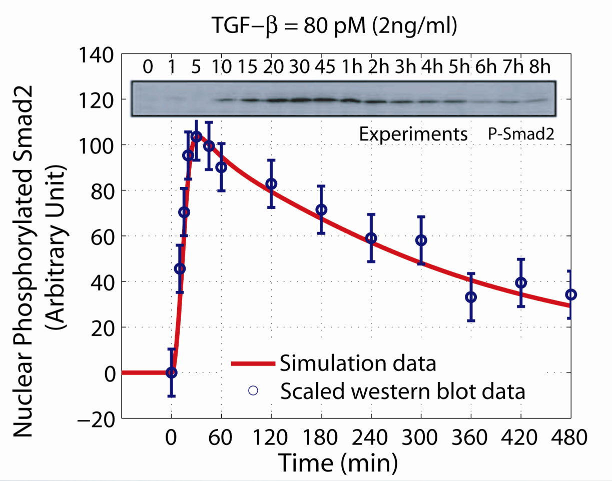
$$\frac{d[K^*]}{dt} = k_{f3}[LR^*K] - k_{f4}[K^*][S] + k_{r4}[K^*S] + k_{f5}[K^*S]$$

$$\frac{d[S]}{dt} = -k_{f4}[K^*][S] + k_{r4}[K^*S]$$

$$\frac{d[K^*S]}{dt} = k_{f4}[K^*][S] - k_{r4}[K^*S] - k_{f5}[K^*S]$$

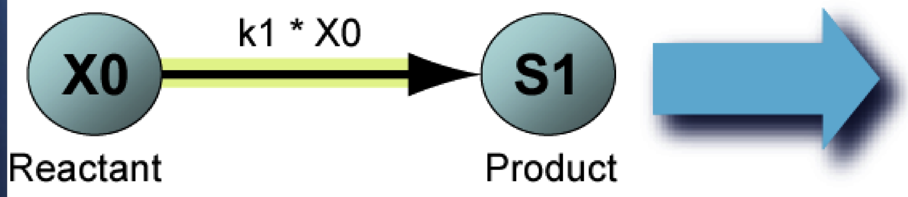
$$\frac{d[S^*]}{dt} = k_{f5}[K^*S]$$

Parameter Estimation



From Zi Z. MPIMG, Berlin, Germany

SBML language



Biochemical reaction

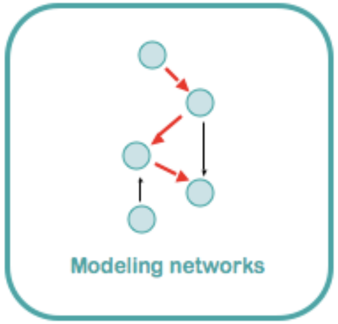
```
<listOfSpecies>
  <species name="X0">
  <species name="S1">
</listOfSpecies>
<listOfReactions>
  <reaction name="reaction_1"/>
    <listOfReactants>
      <speciesReference species="X0"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="S1"/>
    </listOfProducts>
    <kineticLawformula="k1 * X0">
        <listOfParameters>
        </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

SBML representation

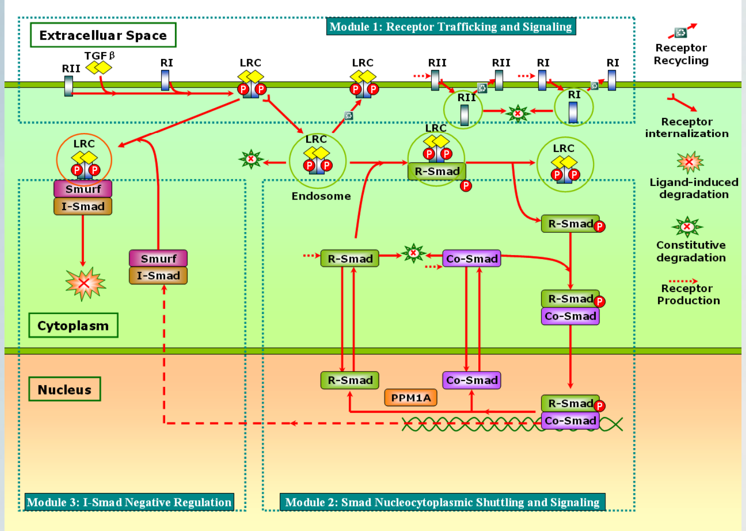
From M. Barberis, MPIMG, Berlin, Germany

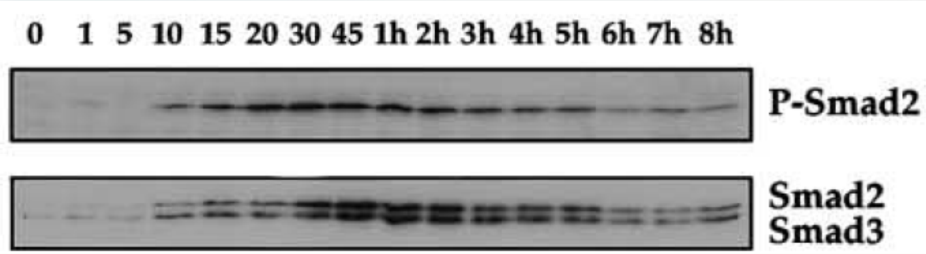
Kinetic modeling

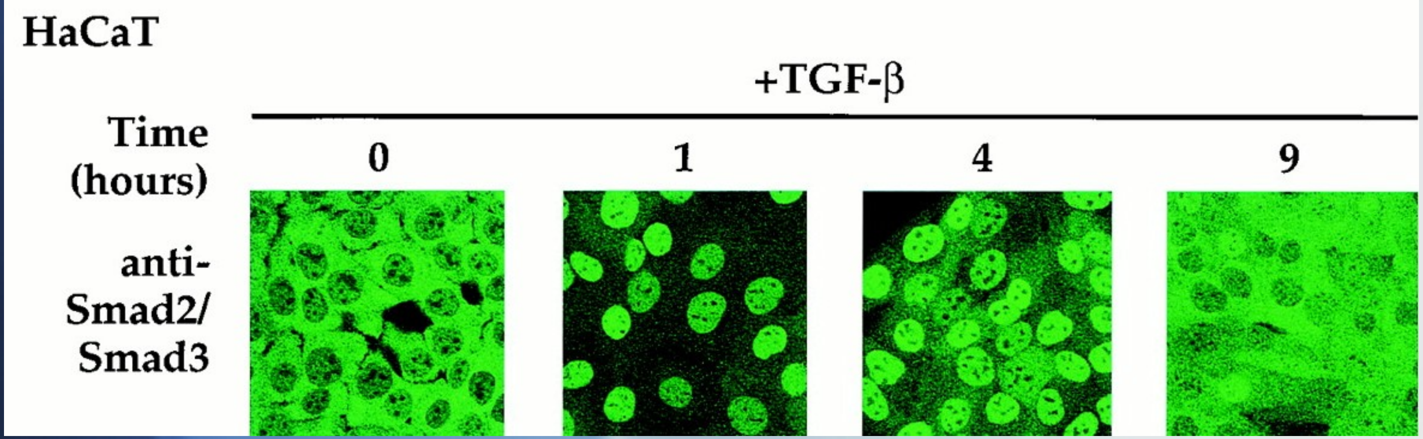
Systems Level Modeling



Modeling Signaling pathways: TGF-beta







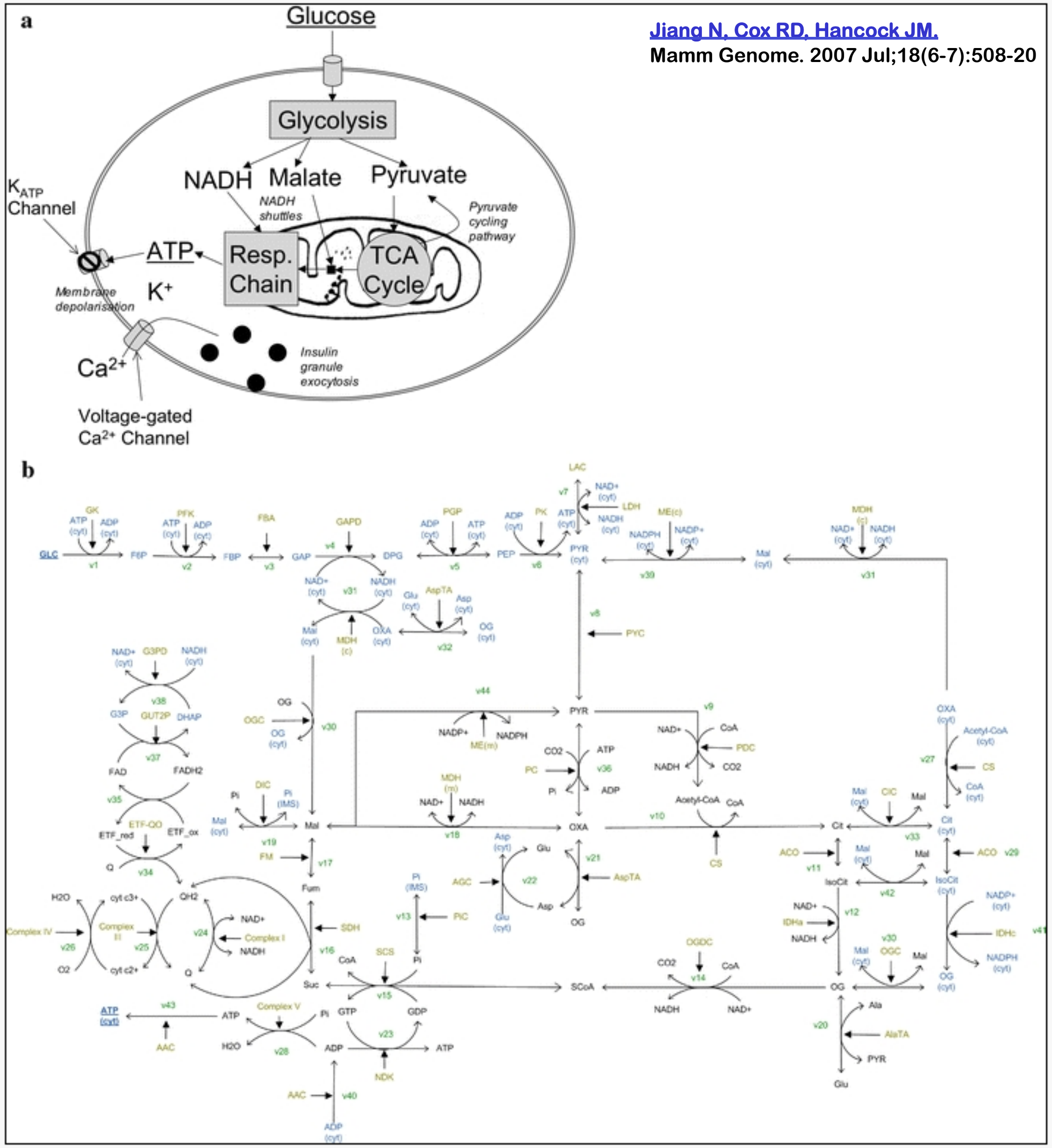
From Zi Z. MPIMG, Berlin, Germany

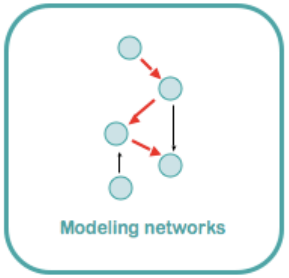
Systems Level Modeling

Modeling networks

A kinetic core model of the glucose-stimulated insulin secretion network of pancreatic beta cells

Jiang N, Cox RD, Hancock JM.
Mamm Genome. 2007 Jul;18(6-7):508-20

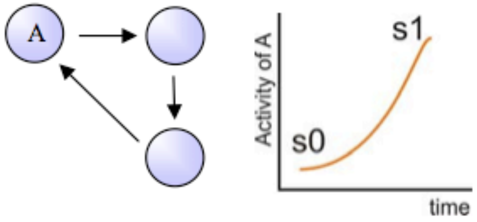




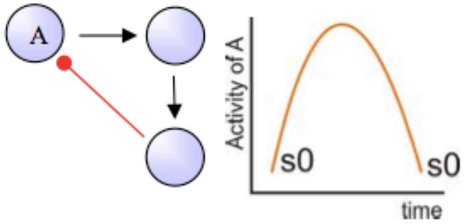
Modeling Signaling pathways: TGF-beta

Boolean modeling

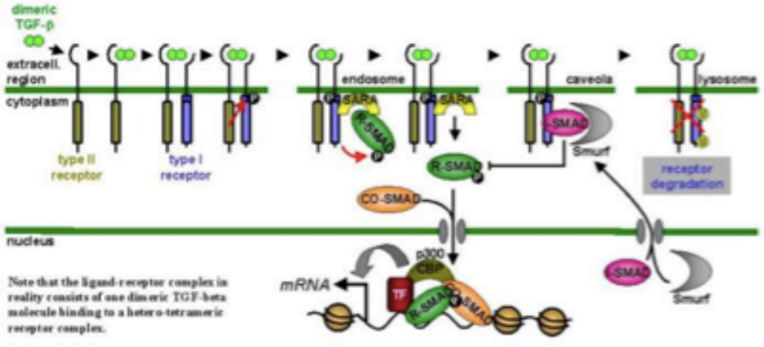
Positive feed-back



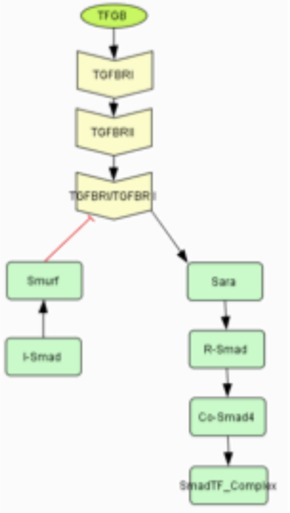
Negative feed-back



Mathematical:	Multiple stationary states	Transitory states, oscillations
Biological:	Differentiation	Homeostasis

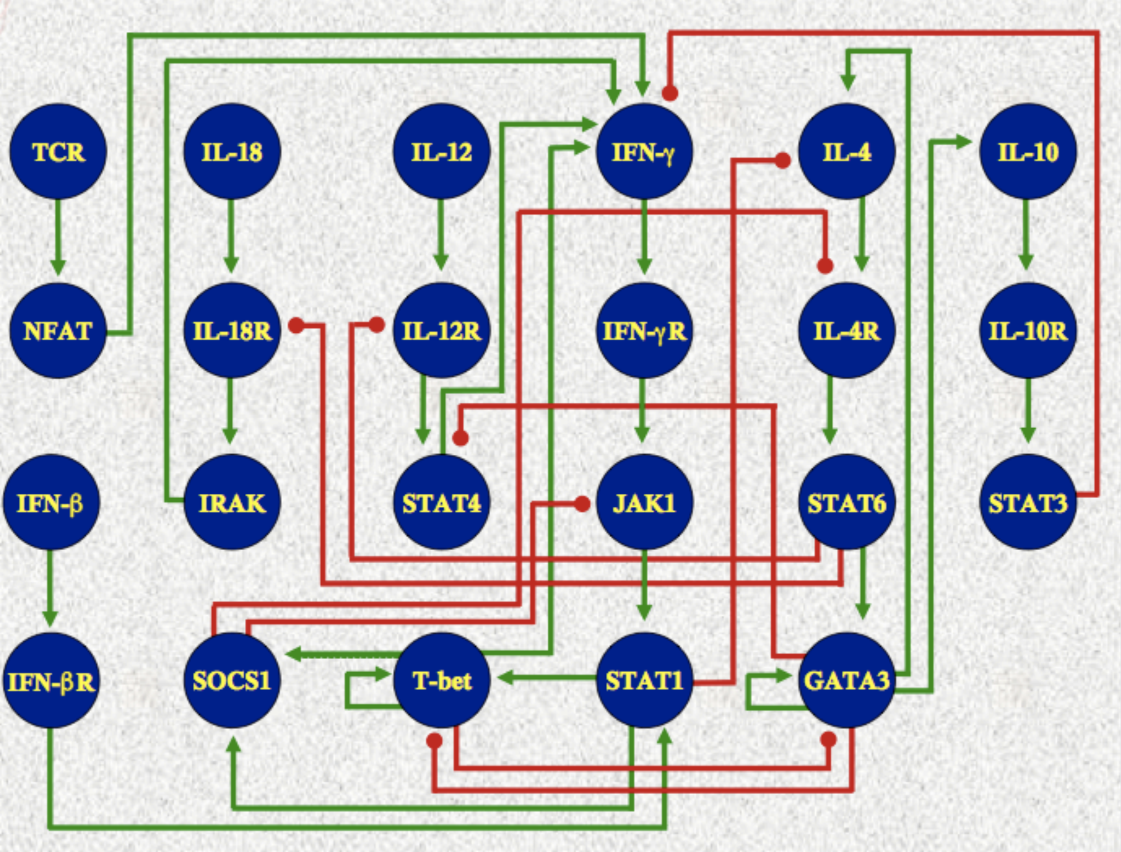


Note that the ligand-receptor complex in reality consists of one dimeric TGF-beta molecule binding to a hetero-tetrameric receptor complex.



From DiCara A., Merk Serono, Geneva, Switzerland

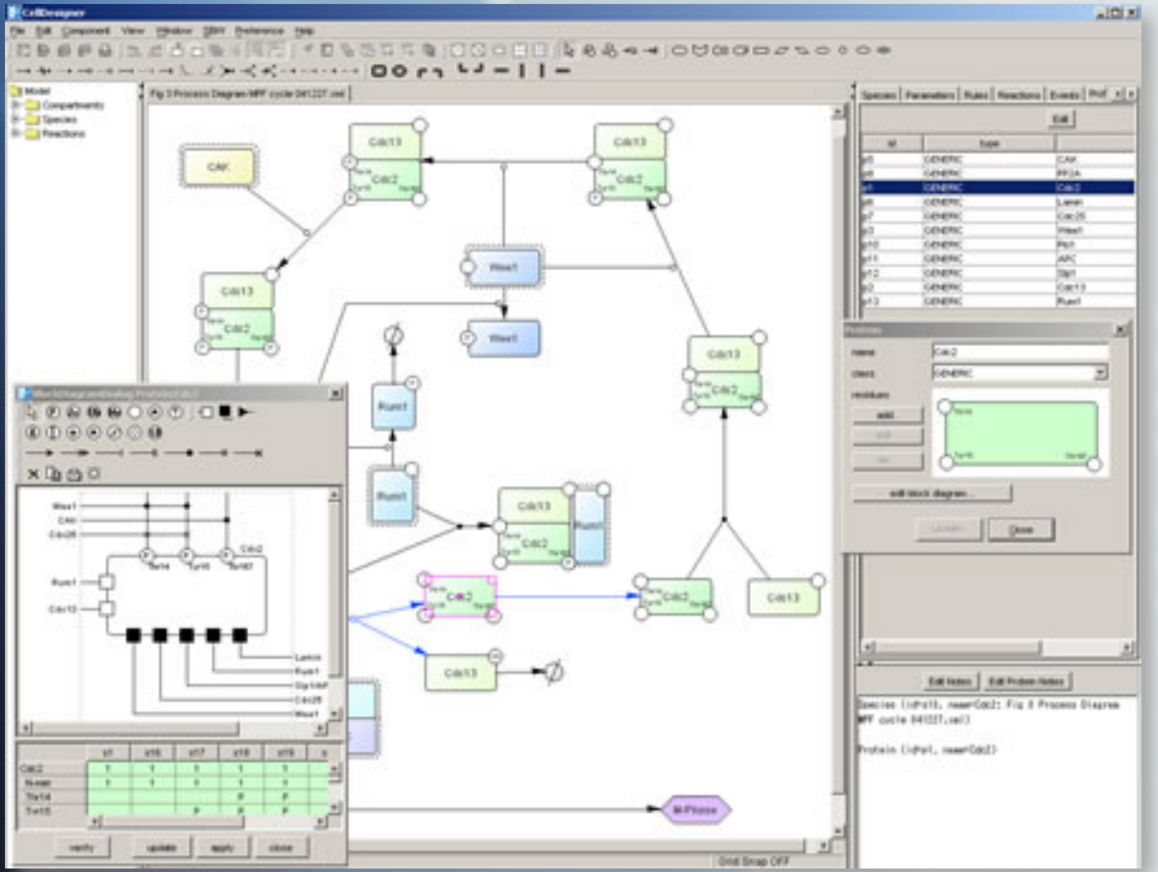
Need for informatics to understand biology



2^{23} = 8.388.608 activation states !!

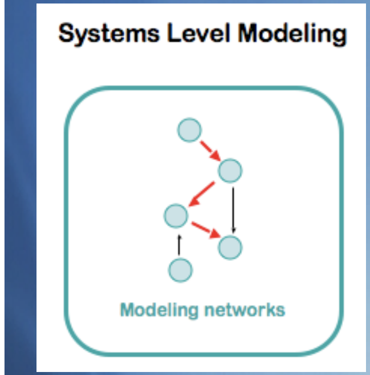
From L. Mendoza, UNAM, Mexico

Model design tools



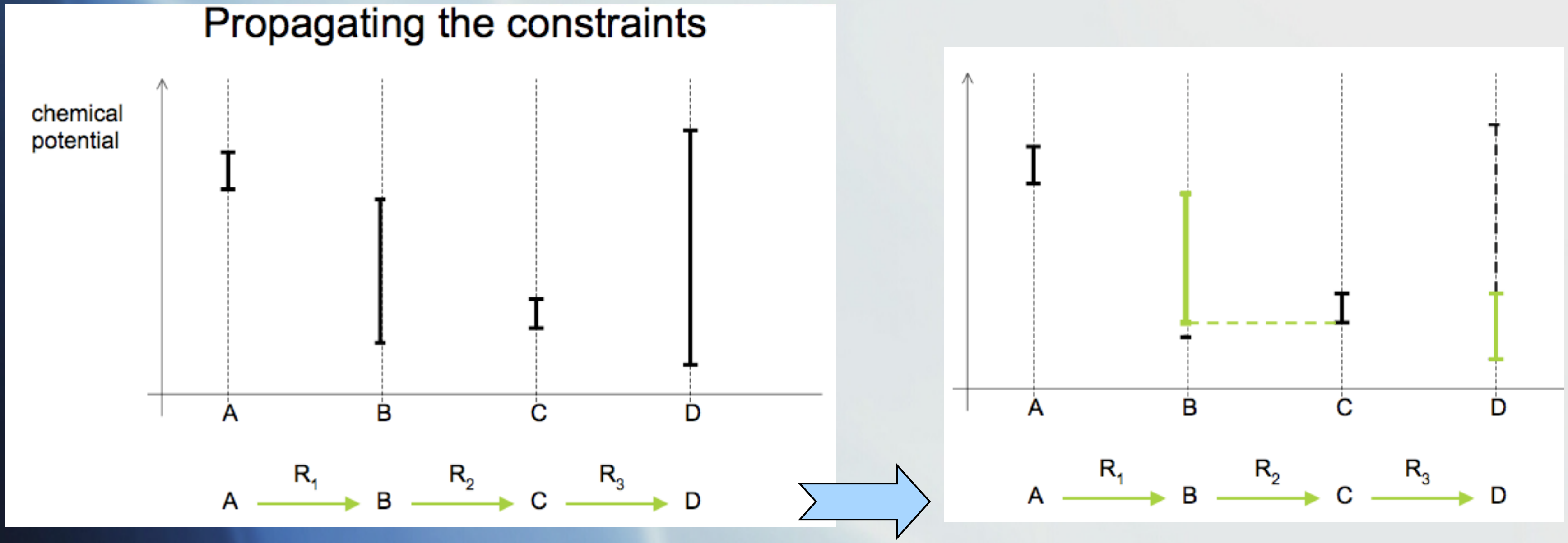
- Full SBML support
- Graphical notation (SBGN)
- Built-in simulator (SBML ODE Solver)
- Integrate with Analysis tool, other simulators through SBW
- Database connection
- Export to PDF, PNG, …
- Freely available
- Supported Environment
  - Windows (2000 or later)
  - Mac OS X
  - Linux

http://www.celldesigner.org



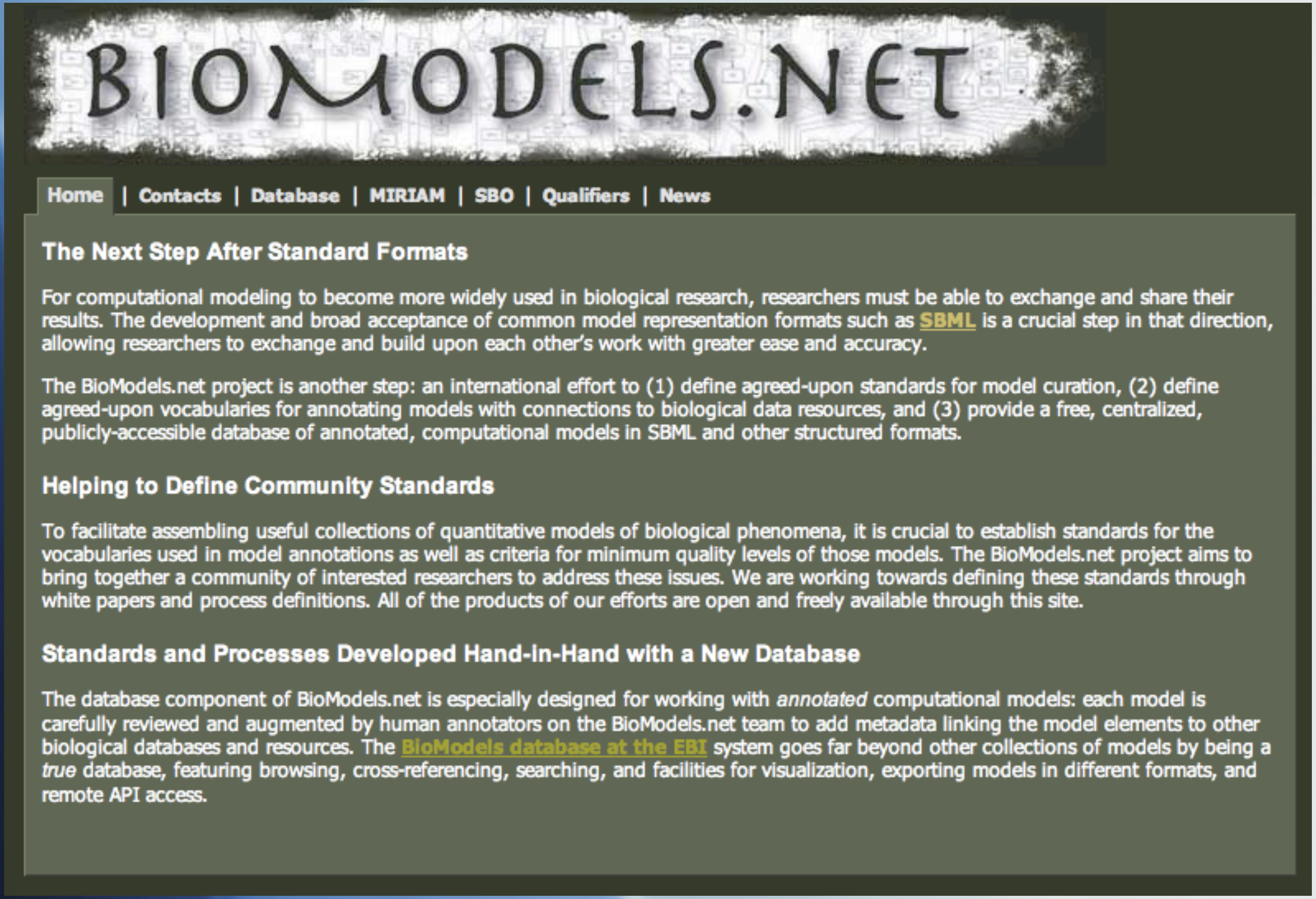
Modeling metabolic networks: Central carbon metabolism in Yeast

Constrain-based modeling


Propagating the constraints

From M. Durot, Genoscope, France

Models database



The Next Step After Standard Formats

For computational modeling to become more widely used in biological research, researchers must be able to exchange and share their results. The development and broad acceptance of common model representation formats such as SBML is a crucial step in that direction, allowing researchers to exchange and build upon each other's work with greater ease and accuracy.

The BioModels.net project is another step: an international effort to (1) define agreed-upon standards for model curation, (2) define agreed-upon vocabularies for annotating models with connections to biological data resources, and (3) provide a free, centralized, publicly-accessible database of annotated, computational models in SBML and other structured formats.

Helping to Define Community Standards

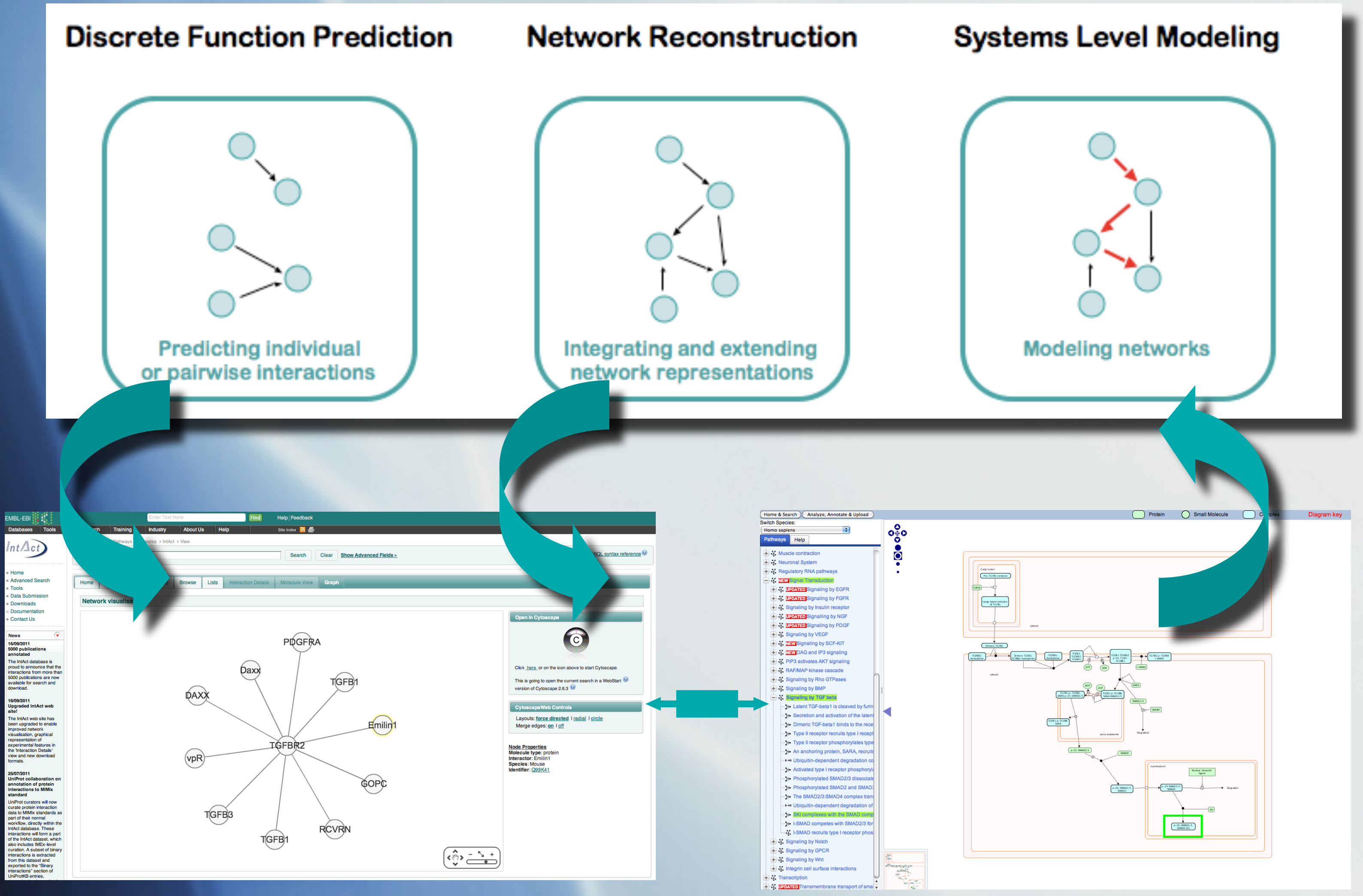
To facilitate assembling useful collections of quantitative models of biological phenomena, it is crucial to establish standards for the vocabularies used in model annotations as well as criteria for minimum quality levels of those models. The BioModels.net project aims to bring together a community of interested researchers to address these issues. We are working towards defining these standards through white papers and process definitions. All of the products of our efforts are open and freely available through this site.

Standards and Processes Developed Hand-in-Hand with a New Database

The database component of BioModels.net is especially designed for working with *annotated* computational models: each model is carefully reviewed and augmented by human annotators on the BioModels.net team to add metadata linking the model elements to other biological databases and resources. The BioModels database at the EBI system goes far beyond other collections of models by being a *true* database, featuring browsing, cross-referencing, searching, and facilities for visualization, exporting models in different formats, and remote API access.

Data Integration

Storage of Predicted Data



News

16/09/2011
5000 publications annotated

The IntAct database is proud to announce that the interactions from more than 5000 publications are now available for search and download.

16/09/2011
Upgraded IntAct web site!

The IntAct web site has been upgraded to enable improved network visualisation, graphical representation of experimental features in the 'Interaction Details' view and new download formats.

25/07/2011
UniProt collaboration on annotation of protein interactions to MIMIx standard

UniProt curators will now curate protein interaction data to MIMIx standards as part of their normal workflow, directly within the IntAct database. These interactions will form a part of the IntAct dataset, which also includes IMEx-level curation. A subset of binary interactions is extracted from this dataset and exported to the "Binary interactions" section of UniProtKB entries,

Click here or on the icon above to start Cytoscape.

This is going to open the current search in a WebStart version of Cytoscape 2.6.3

Layouts: force directed | radial | circle
Merge edges: on | off

Node Properties
Molecule type: protein
Interactor: Emilin1
Species: Mouse
Identifier: Q99K41

- Muscle contraction
- Neuronal System
- Regulatory RNA pathways
- NEW Signal Transduction
  - UPDATED Signaling by EGFR
  - UPDATED Signaling by FGFR
  - Signaling by Insulin receptor
  - UPDATED Signalling by NGF
  - UPDATED Signaling by PDGF
  - Signaling by VEGF
  - NEW Signaling by SCF-KIT
  - NEW DAG and IP3 signaling
  - PIP3 activates AKT signaling
  - RAF/MAP kinase cascade
  - Signaling by Rho GTPases
  - Signaling by BMP
  - Signaling by TGF beta
    - Latent TGF-beta1 is cleaved by furin
    - Secretion and activation of the latent
    - Dimeric TGF-beta1 binds to the rece
    - Type II receptor recruits type I recept
    - Type II receptor phosphorylates type
    - An anchoring protein, SARA, recruits
    - Ubiquitin-dependent degradation co
    - Activated type I receptor phosphoryla
    - Phosphorylated SMAD2/3 dissociate
    - Phosphorylated SMAD2 and SMAD3
    - The SMAD2/3:SMAD4 complex trans
    - Ubiquitin-dependent degradation of
    - SKI complexes with the SMAD comp
    - I-SMAD competes with SMAD2/3 for
    - I-SMAD recruits type I receptor phos
  - Signaling by Notch
  - Signaling by GPCR
  - Signaling by Wnt
  - Integrin cell surface interactions
- Transcription
- UPDATED Transmembrane transport of sma

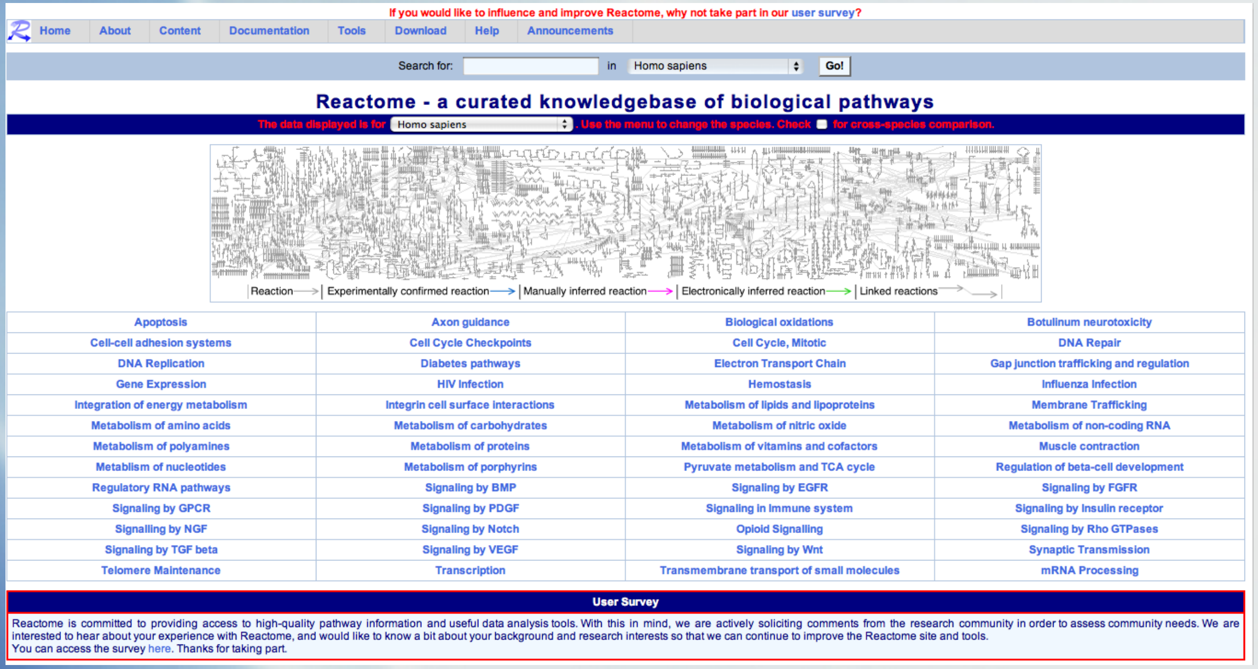
www.ebi.ac.uk/intact

www.reactome.org

www.reactome.org

Reactome

Database of curated biological reactions
NIH
ENFIN
EBI Industry Programme



If you would like to influence and improve Reactome, why not take part in our user survey?

Reactome - a curated knowledgebase of biological pathways

The data displayed is for Homo sapiens. Use the menu to change the species. Check ☐ for cross-species comparison.

Apoptosis	Axon guidance	Biological oxidations	Botulinum neurotoxicity
Cell-cell adhesion systems	Cell Cycle Checkpoints	Cell Cycle, Mitotic	DNA Repair
DNA Replication	Diabetes pathways	Electron Transport Chain	Gap junction trafficking and regulation
Gene Expression	HIV Infection	Hemostasis	Influenza Infection
Integration of energy metabolism	Integrin cell surface interactions	Metabolism of lipids and lipoproteins	Membrane Trafficking
Metabolism of amino acids	Metabolism of carbohydrates	Metabolism of nitric oxide	Metabolism of non-coding RNA
Metabolism of polyamines	Metabolism of proteins	Metabolism of vitamins and cofactors	Muscle contraction
Metablism of nucleotides	Metabolism of porphyrins	Pyruvate metabolism and TCA cycle	Regulation of beta-cell development
Regulatory RNA pathways	Signaling by BMP	Signaling by EGFR	Signaling by FGFR
Signaling by GPCR	Signaling by PDGF	Signaling in Immune system	Signaling by Insulin receptor
Signalling by NGF	Signaling by Notch	Opioid Signalling	Signaling by Rho GTPases
Signaling by TGF beta	Signaling by VEGF	Signaling by Wnt	Synaptic Transmission
Telomere Maintenance	Transcription	Transmembrane transport of small molecules	mRNA Processing

User Survey

Reactome is committed to providing access to high-quality pathway information and useful data analysis tools. With this in mind, we are actively soliciting comments from the research community in order to assess community needs. We are interested to hear about your experience with Reactome, and would like to know a bit about your background and research interests so that we can continue to improve the Reactome site and tools.
You can access the survey here. Thanks for taking part.

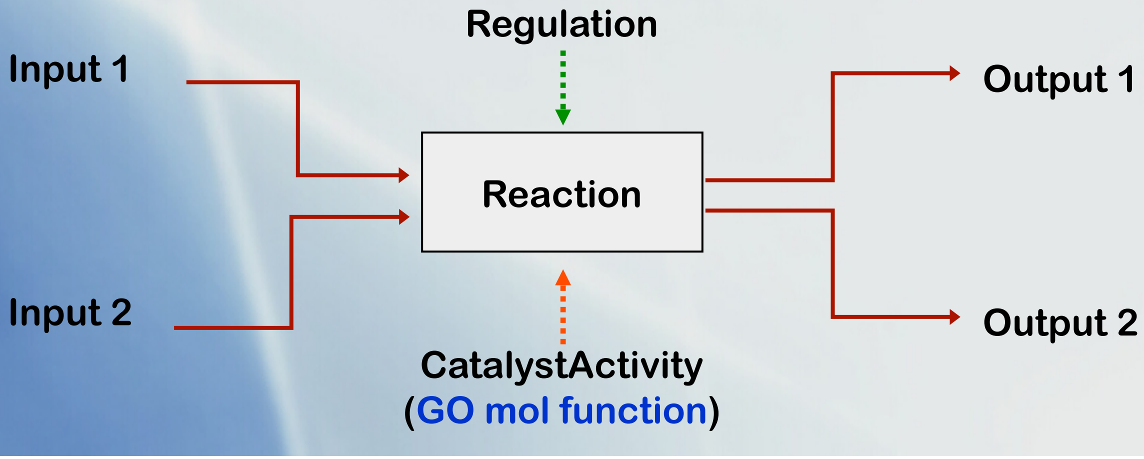
Stats For Homo Sapiens:

Proteins: 5037

Complexes: 3515

Reactions: 3837

Pathways: 1059



Regulation

Input 1

Input 2

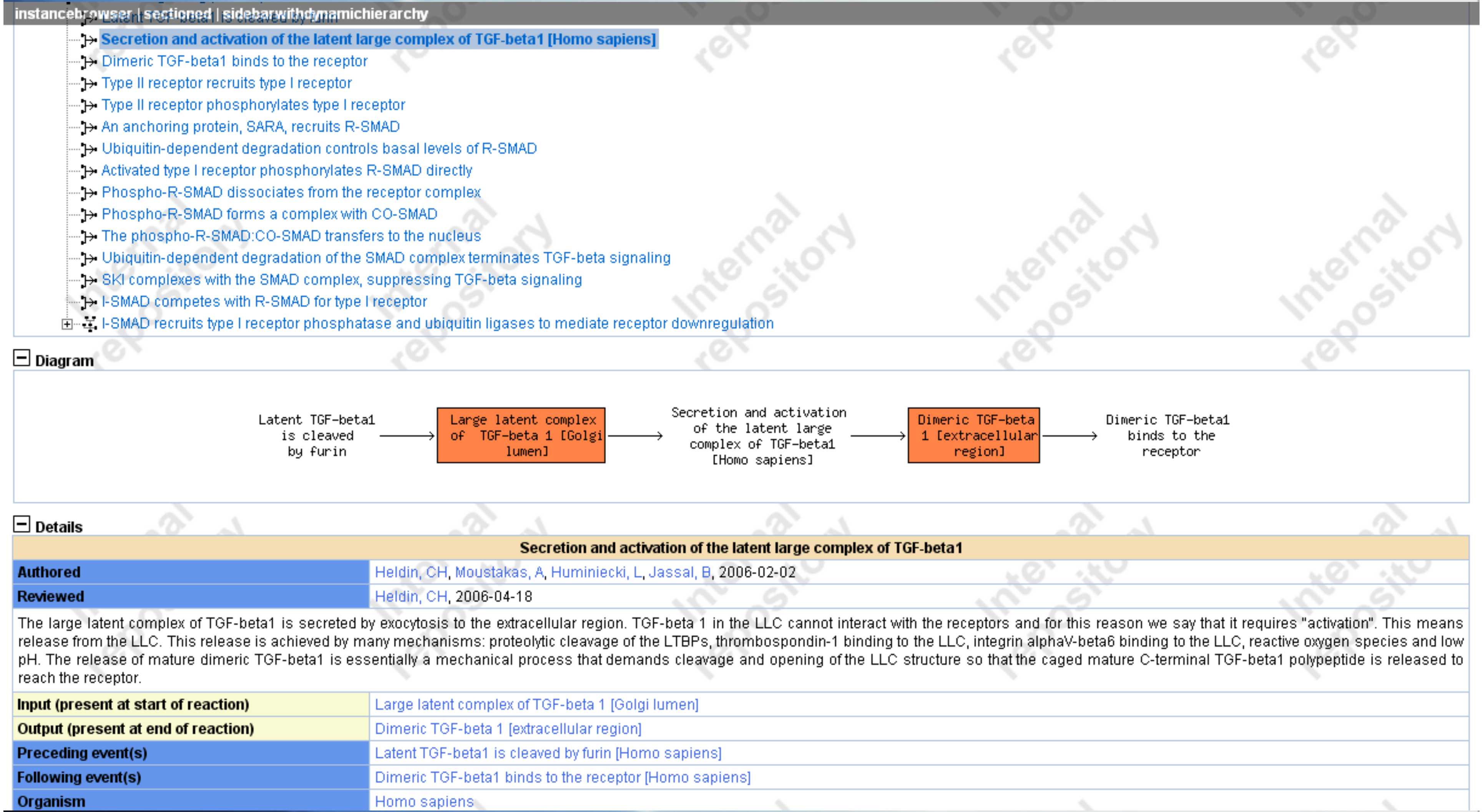
Reaction

Output 1

Output 2

CatalystActivity
(GO mol function)

www.reactome.org



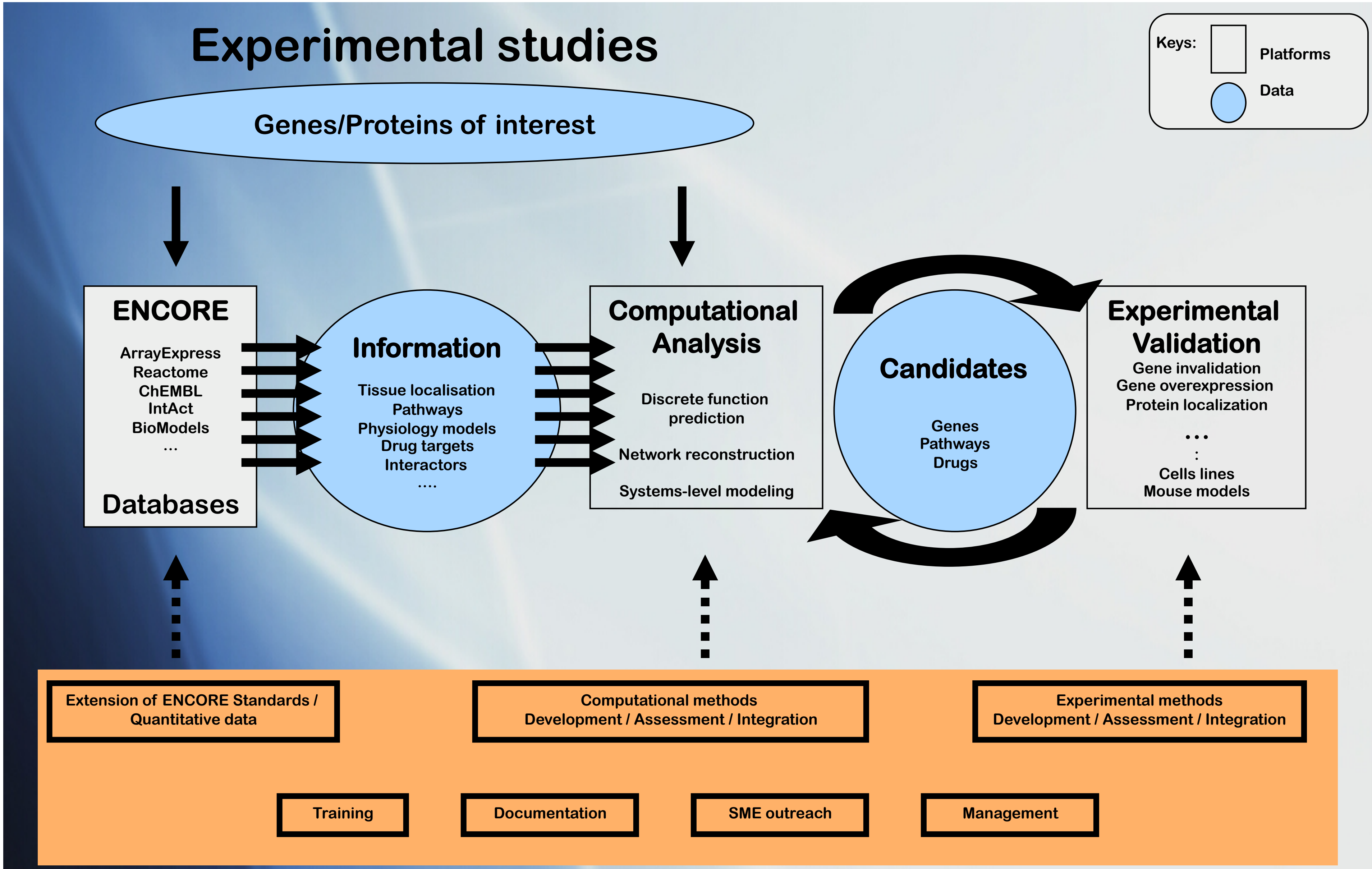
- Latent TGF-beta1 is cleaved by furin
- **Secretion and activation of the latent large complex of TGF-beta1 [Homo sapiens]**
- Dimeric TGF-beta1 binds to the receptor
- Type II receptor recruits type I receptor
- Type II receptor phosphorylates type I receptor
- An anchoring protein, SARA, recruits R-SMAD
- Ubiquitin-dependent degradation controls basal levels of R-SMAD
- Activated type I receptor phosphorylates R-SMAD directly
- Phospho-R-SMAD dissociates from the receptor complex
- Phospho-R-SMAD forms a complex with CO-SMAD
- The phospho-R-SMAD:CO-SMAD transfers to the nucleus
- Ubiquitin-dependent degradation of the SMAD complex terminates TGF-beta signaling
- SKI complexes with the SMAD complex, suppressing TGF-beta signaling
- I-SMAD competes with R-SMAD for type I receptor
- I-SMAD recruits type I receptor phosphatase and ubiquitin ligases to mediate receptor downregulation

Diagram

Details

Secretion and activation of the latent large complex of TGF-beta1	
Authored	Heldin, CH, Moustakas, A, Huminiecki, L, Jassal, B, 2006-02-02
Reviewed	Heldin, CH, 2006-04-18
The large latent complex of TGF-beta1 is secreted by exocytosis to the extracellular region. TGF-beta 1 in the LLC cannot interact with the receptors and for this reason we say that it requires "activation". This means release from the LLC. This release is achieved by many mechanisms: proteolytic cleavage of the LTBPs, thrombospondin-1 binding to the LLC, integrin alphaV-beta6 binding to the LLC, reactive oxygen species and low pH. The release of mature dimeric TGF-beta1 is essentially a mechanical process that demands cleavage and opening of the LLC structure so that the caged mature C-terminal TGF-beta1 polypeptide is released to reach the receptor.	
Input (present at start of reaction)	Large latent complex of TGF-beta 1 [Golgi lumen]
Output (present at end of reaction)	Dimeric TGF-beta 1 [extracellular region]
Preceding event(s)	Latent TGF-beta1 is cleaved by furin [Homo sapiens]
Following event(s)	Dimeric TGF-beta1 binds to the receptor [Homo sapiens]
Organism	Homo sapiens

Experimental studies



Keys:
▭ Platforms
● Data

Genes/Proteins of interest

ENCORE

ArrayExpress
Reactome
ChEMBL
IntAct
BioModels
…

Databases

Information

Tissue localisation
Pathways
Physiology models
Drug targets
Interactors
….

Computational Analysis

Discrete function prediction

Network reconstruction

Systems-level modeling

Candidates

Genes
Pathways
Drugs

Experimental Validation

Gene invalidation
Gene overexpression
Protein localization
…
:
Cells lines
Mouse models

Extension of ENCORE Standards / Quantitative data

Computational methods
Development / Assessment / Integration

Experimental methods
Development / Assessment / Integration

Training

Documentation

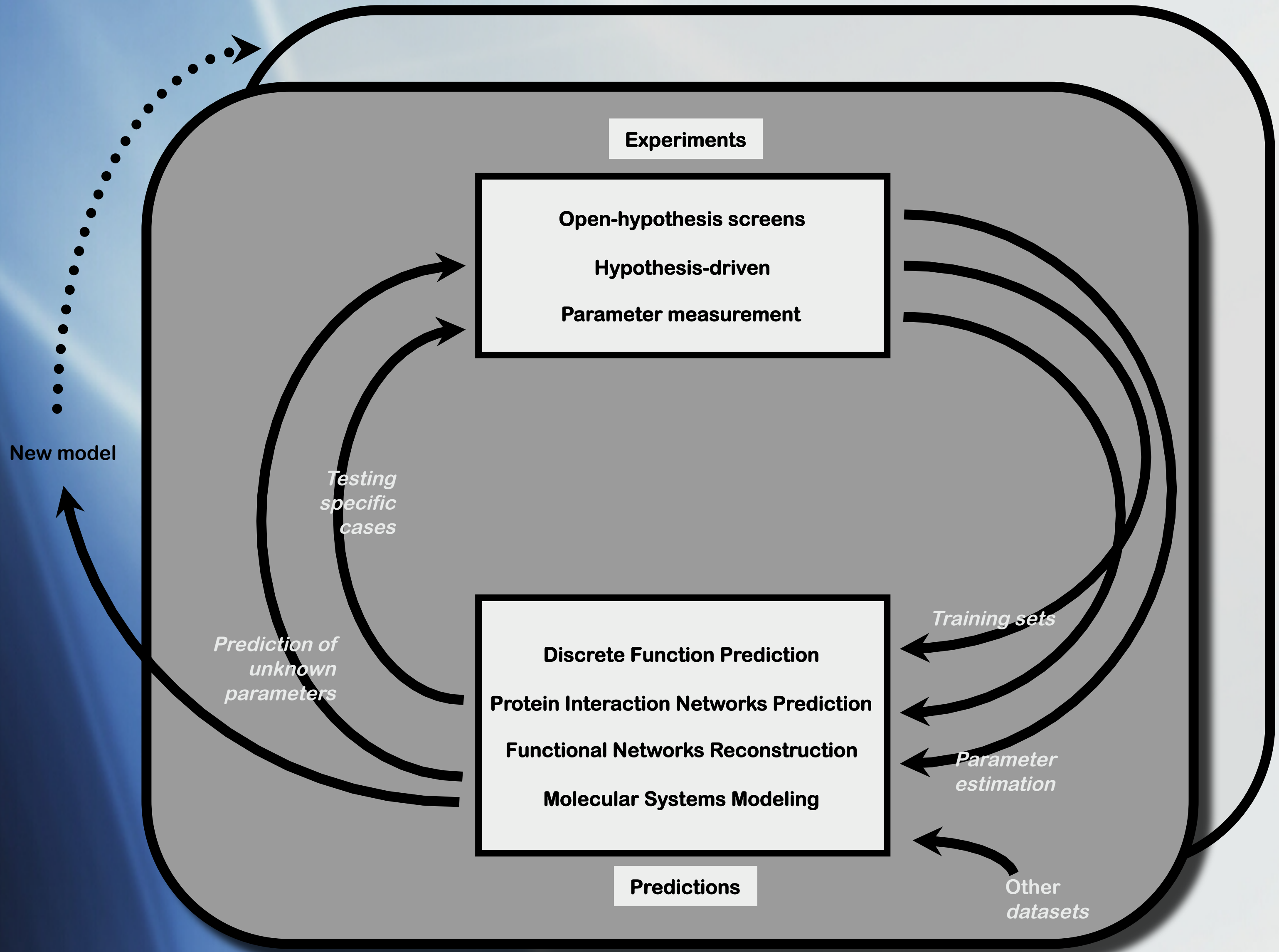
SME outreach

Management

ENFIN
Enabling Systems Biology
www.enfin.org

Analysis Platform

Data Integration



Experiments

Open-hypothesis screens

Hypothesis-driven

Parameter measurement

Predictions

Discrete Function Prediction

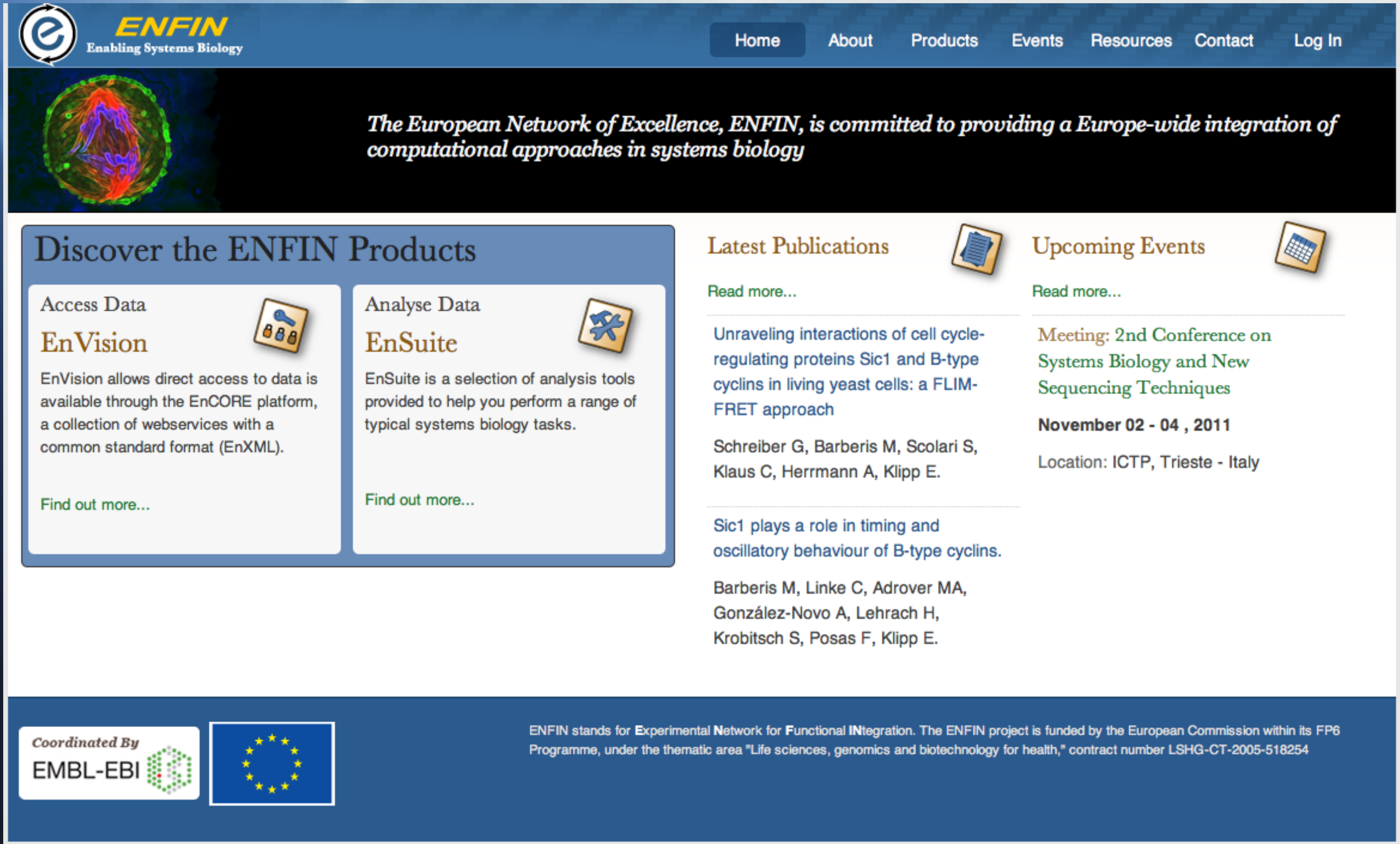
Protein Interaction Networks Prediction

Functional Networks Reconstruction

Molecular Systems Modeling

ENFIN Website: www.enfin.org

Training **Conferences** **100+ Publications** **Documentation**



The European Network of Excellence, ENFIN, is committed to providing a Europe-wide integration of computational approaches in systems biology

Discover the ENFIN Products

Access Data

EnVision

EnVision allows direct access to data is available through the EnCORE platform, a collection of webservices with a common standard format (EnXML).

Find out more...

Analyse Data

EnSuite

EnSuite is a selection of analysis tools provided to help you perform a range of typical systems biology tasks.

Find out more...

Latest Publications

Read more...

Unraveling interactions of cell cycle-regulating proteins Sic1 and B-type cyclins in living yeast cells: a FLIM-FRET approach

Schreiber G, Barberis M, Scolari S, Klaus C, Herrmann A, Klipp E.

Sic1 plays a role in timing and oscillatory behaviour of B-type cyclins.

Barberis M, Linke C, Adrover MA, González-Novo A, Lehrach H, Krobitsch S, Posas F, Klipp E.

Upcoming Events

Read more...

Meeting: 2nd Conference on Systems Biology and New Sequencing Techniques

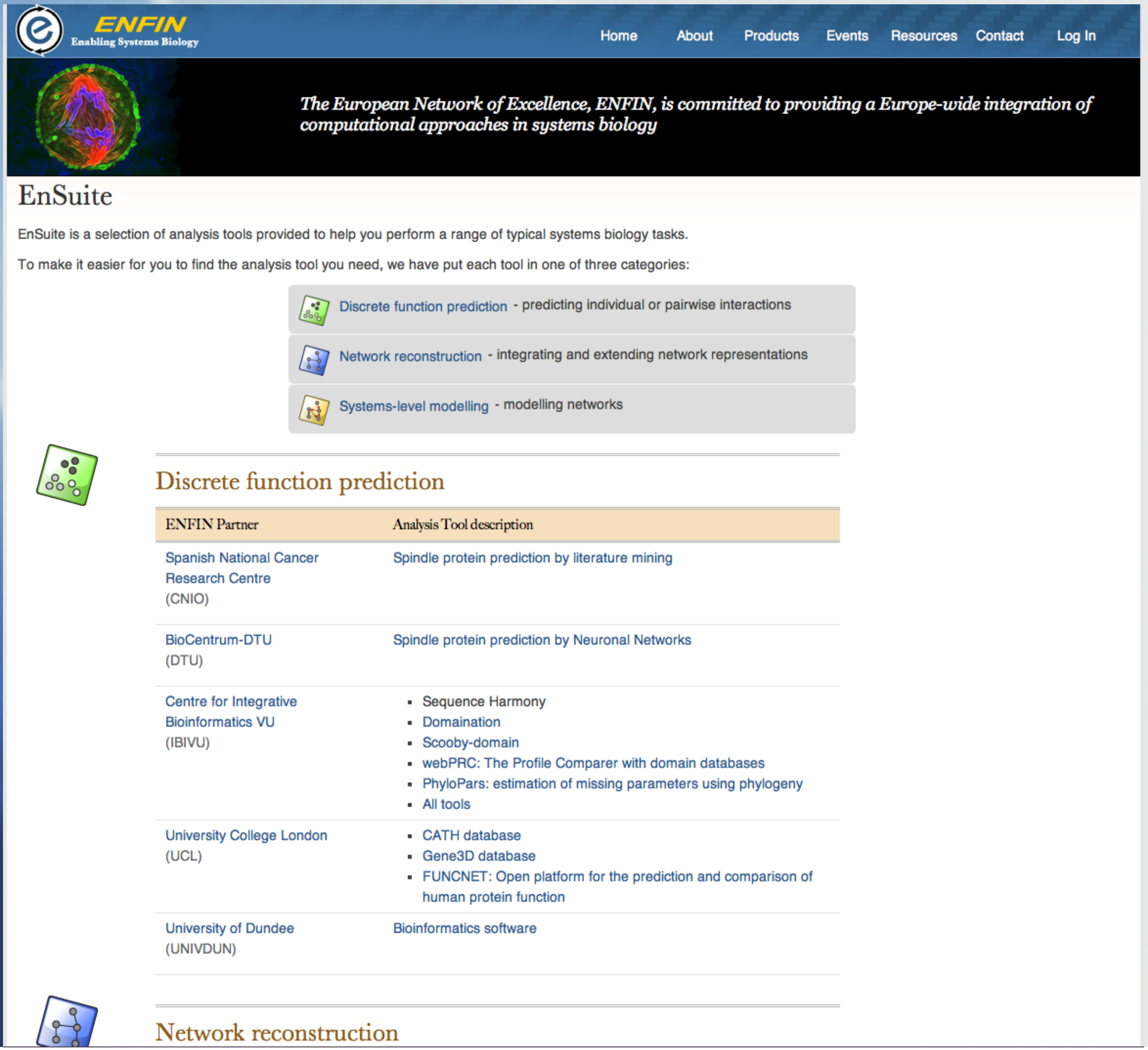
November 02 - 04 , 2011

Location: ICTP, Trieste - Italy

ENFIN stands for **E**xperimental **N**etwork for **F**unctional **IN**tegration. The ENFIN project is funded by the European Commission within its FP6 Programme, under the thematic area "Life sciences, genomics and biotechnology for health," contract number LSHG-CT-2005-518254

ENFIN
Enabling Systems Biology

EnSUITE



The European Network of Excellence, ENFIN, is committed to providing a Europe-wide integration of computational approaches in systems biology

EnSuite

EnSuite is a selection of analysis tools provided to help you perform a range of typical systems biology tasks.

To make it easier for you to find the analysis tool you need, we have put each tool in one of three categories:

- Discrete function prediction - predicting individual or pairwise interactions
- Network reconstruction - integrating and extending network representations
- Systems-level modelling - modelling networks

Discrete function prediction

ENFIN Partner	Analysis Tool description
Spanish National Cancer Research Centre (CNIO)	Spindle protein prediction by literature mining
BioCentrum-DTU (DTU)	Spindle protein prediction by Neuronal Networks
Centre for Integrative Bioinformatics VU (IBIVU)	Sequence Harmony Domaination Scooby-domain webPRC: The Profile Comparer with domain databases PhyloPars: estimation of missing parameters using phylogeny All tools
University College London (UCL)	CATH database Gene3D database FUNCNET: Open platform for the prediction and comparison of human protein function
University of Dundee (UNIVDUN)	Bioinformatics software

Network reconstruction

Acknowledgements

Ewan Birney
Henning Hermjakob
Jan Ellenberg
Geoffrey Barton
Soren Brunak
Gianni Cesareni
John Hancock
Carl-Henrick Heldin
Edda Klipp
Erich Nigg
James Adjaye
Tomi Makela
Christine Orengo
Christos Ouzounis
Vincent Schachter
Dietmar Schomburg
Eran Segal
Jaak Vilo
Ioannis Xenarios
Alfonso Valencia
Jaap Heringa

LUDWIG INSTITUTE FOR CANCER RESEARCH

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

DTU

Centre for Integrative Bioinformatics VU
vrije Universiteit amsterdam

UNIVERSITY OF HELSINKI

GENOSCOPE

TorVergata Università

UCL

EMBL-EBI

MRC Mammalian Genetics Unit

CERTH
The Centre for Research and Technology Hellas

Max Planck Institute for Molecular Genetics

QureTEC

cnio
Centro Nacional de Investigaciones Oncológicas

SIB
Swiss Institute of Bioinformatics

UNIVERSITY OF DUNDEE

HUMBOLDT-UNIVERSITÄT ZU BERLIN

TECHNISCHE UNIVERSITÄT
CAROLO-WILHELMINA
zu Braunschweig