UNIVERSITÉ
BORDEAUX
SEGALEN

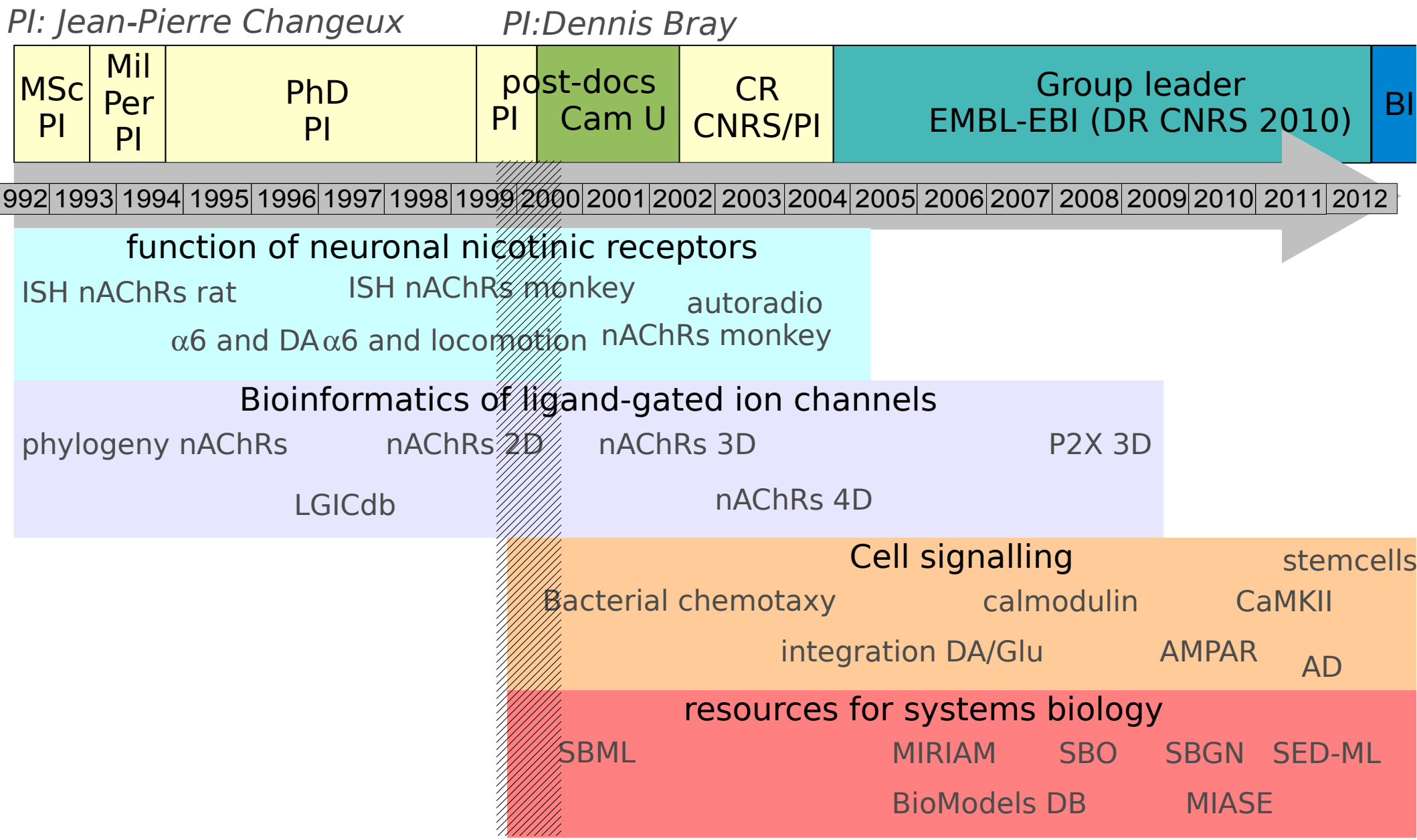
Soutenance de thèse pour l'obtention de l'habilitation à diriger des recherches

Nicolas Le Novère
n.lenovere@gmail.com

Babraham Institute

EMBL-EBI

Career summary



Babraham Institute

EMBL-EBI

Supervision and teaching

- Student supervision
  - PhD supervision: 7, 2 on-going, 5 defended
    - Zhi-Yan Han, research engineer at Institut Curie
    - Dominic Tolle, scientific developer in Berlin,
    - Mélanie Stefan, research fellow in Harvard
    - Lu Li, post-doc at the Babraham Institute
    - Michele Mattioni, security software developer in Cambridge
  - PhD visitors: 6
  - BSc and MSc internships: 23
  - PhD jurys: 12
- Courses
  - Organisation and lectures: 6
  - Lecturer only: 24

Research management

- Reviewer: 34 journals
- Conference committees: 11 (+ standards-related)
- Evaluation panels: AERES (FR), ANR (FR), BBSRC (UK), BMBF (DE), FRM (FR), Juelich (DE)
- Scientific Advisory Boards
  - Projects: BioPAX (NIH), CARMEN (EPSRC), COPASI (BBSRC), ENFIN (EU), Labex EpiGenMed, Plant Models Portal (BBSRC), Virtual Liver Network (BMBF)
  - Institutes: Centre for Organismal Studies (Heidelberg), Institut Curie research centre (Paris), Pole biosanté Rabelais (Montpellier)

Scientific production

- Publications
  - Peer-reviewed journal: 87 (45 open-access)
  - Peer-reviewed conference proceedings: 11 (6 open-access)
  - Chapters, editorials, non peer reviewed etc. : 26 (4 open-access)
  - Technical reports: 18
  - Total citations (Google Scholar Nov 2013): 11629
  - H-index: 46
- Software
  - Maintained software applications: 4
  - Databases: 3
  - Terminologies: 5

UNIVERSITÉ
BORDEAUX
SEGALEN

Normes pour le partage des modèles en biologie des systèmes

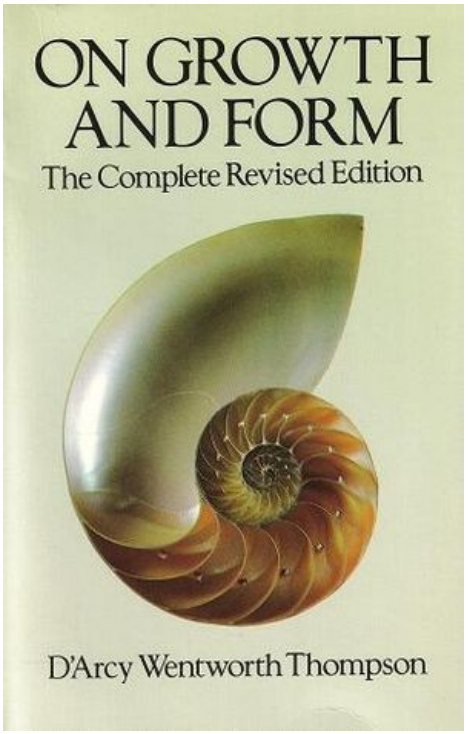
Nicolas Le Novère
n.lenovere@gmail.com

Babraham Institute

EMBL-EBI

Why using mathematical models?

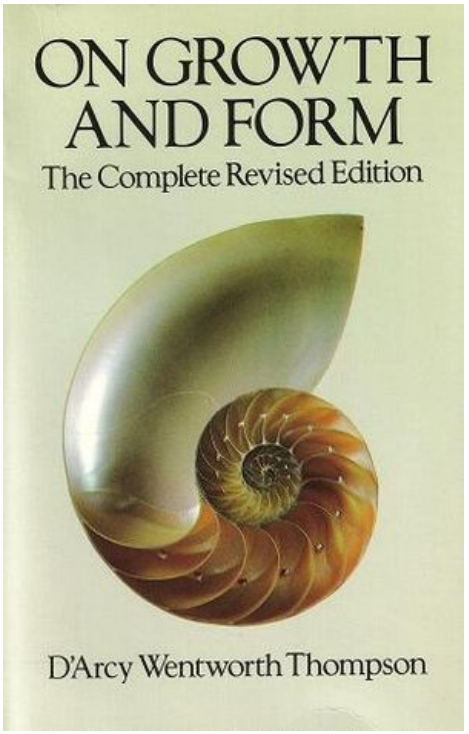
Describe



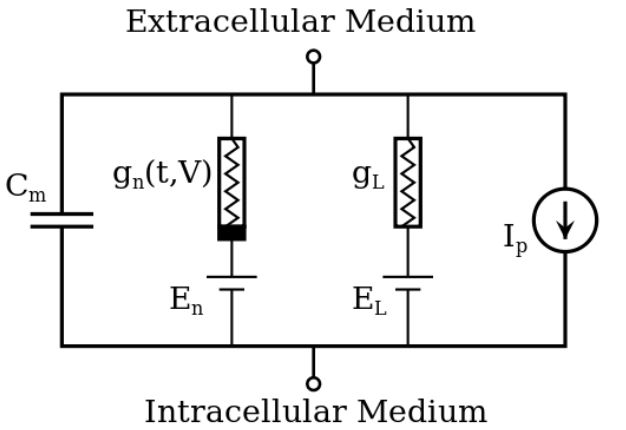
1917

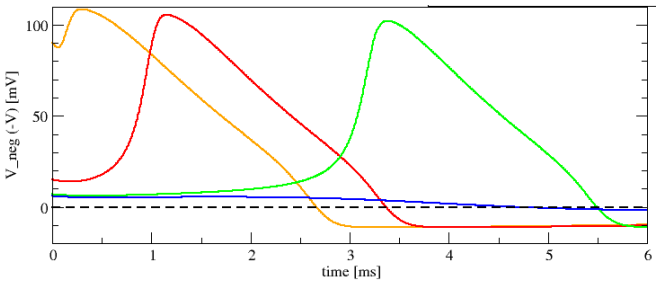
Why using mathematical models?

Describe	Explain



1917





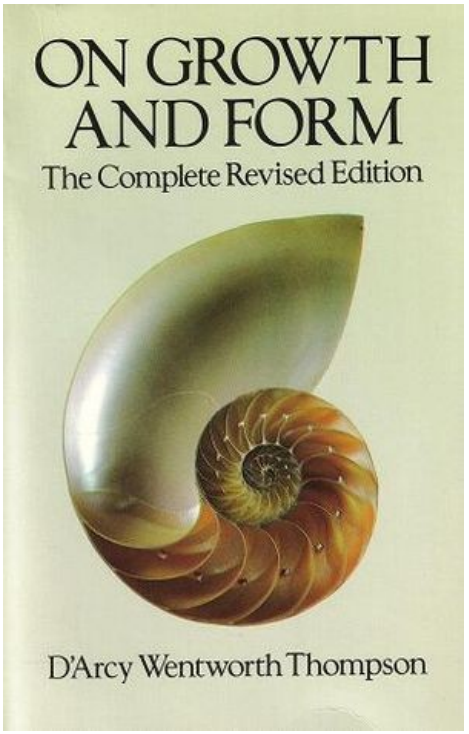
1952

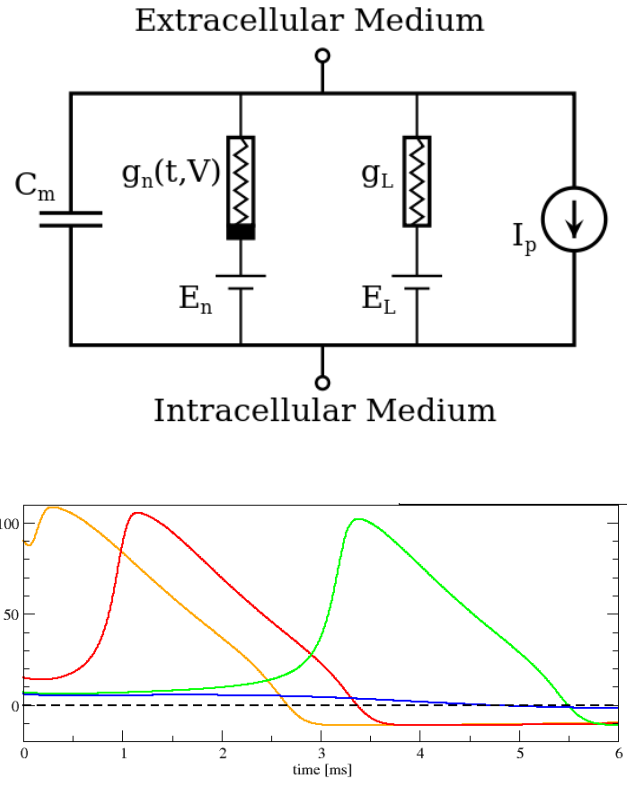
Babraham Institute

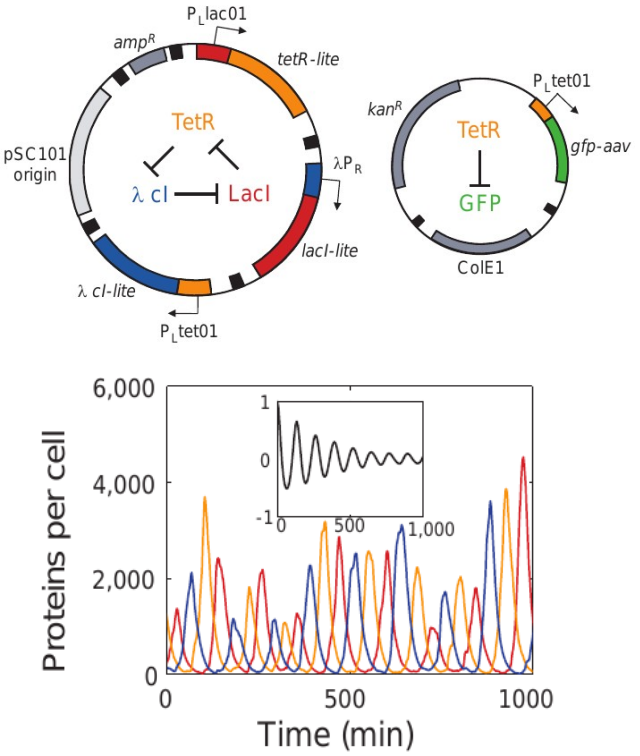
EMBL-EBI

Why using mathematical models?

Describe	Explain	Predict







1917 1952 2000

Babraham Institute EMBL-EBI

What is a mathematical model?

Wikipedia (October 14th 2013): “A mathematical model is a description of a system using mathematical concepts and language.”

What is a mathematical model?

Wikipedia (October 14th 2013): “A mathematical model is a description of a system using mathematical concepts and language.”

variables
[x]
Vmax
Kd
EC_{50}
length
$t_{1/2}$

What we want to know or compare with experiments

What is a mathematical model?

Wikipedia (October 14th 2013): “A mathematical model is a description of a system using mathematical concepts and language.”

variables	relationships
[x]	$K_d = \frac{[A] \cdot [B]}{[AB]}$
Vmax	$d[X]/dt = k \cdot [Y]^2$
Kd	$\sum_i [X]_i - F(t) = 0$
EC_{50}	$k(t) \sim N(k, \sigma^2)$
length	If $\text{mass}_t > \text{threshold}$ then $\text{mass}_{t+\Delta t} = 0.5 \cdot \text{mass}$
$t_{1/2}$	

What we already know or want to test

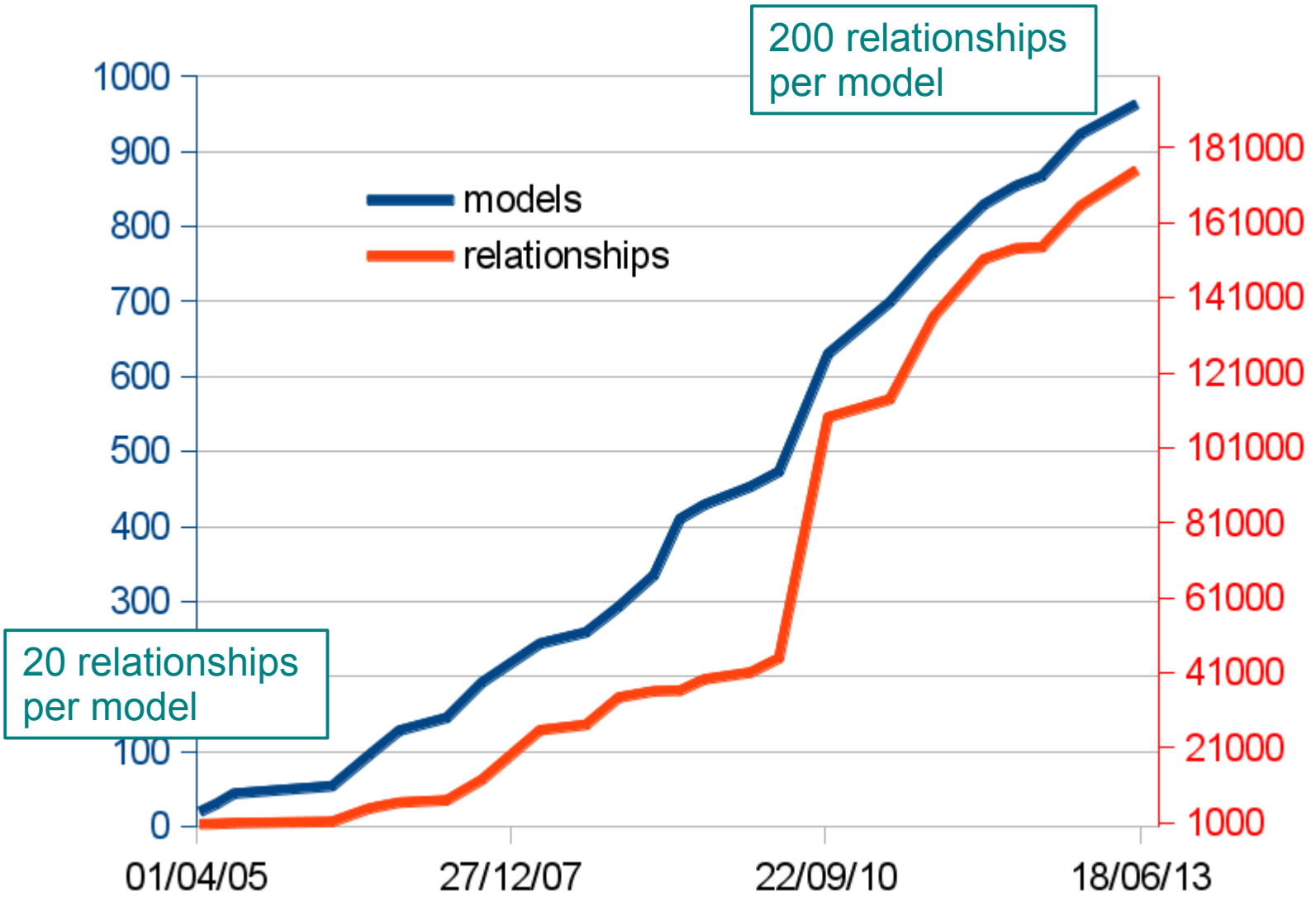
What is a mathematical model?

Wikipedia (October 14th 2013): “A mathematical model is a description of a system using mathematical concepts and language.”

variables	relationships	constraints
[x]	$K_d = \frac{[A] \cdot [B]}{[AB]}$	[x]>0
Vmax	$d[X]/dt = k \cdot [Y]^2$	Energy conservation
Kd	$\sum_i [X]_i - F(t) = 0$	Boundary conditions (v < upper limit)
EC_{50}	$k(t) \sim N(k, \sigma^2)$	Objective functions (maximise ATP)
length	If $\text{mass}_t > \text{threshold}$ then $\text{mass}_{t+\Delta t} = 0.5 \cdot \text{mass}$	Initial conditions
$t_{1/2}$		

The context or what we want to ignore

Computational models on the rise



BioModels Database growth (published models branch) since its creation

We need to

Verify

Re-use

Modify

Build upon

Integrate with

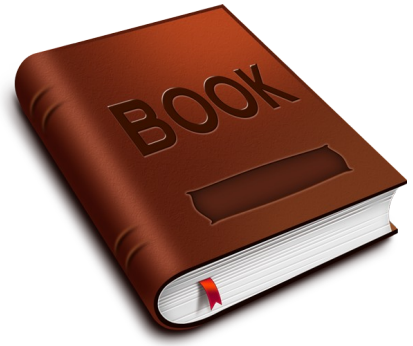
Therefore we need to share

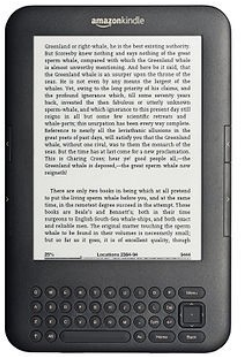
Model descriptions

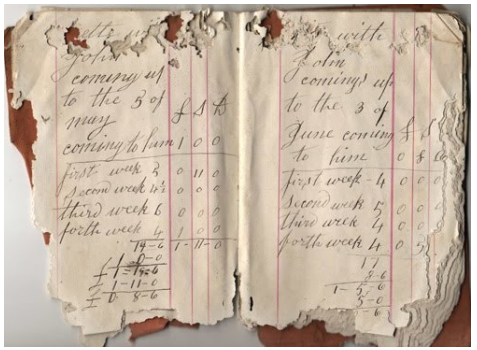
Simulation descriptions

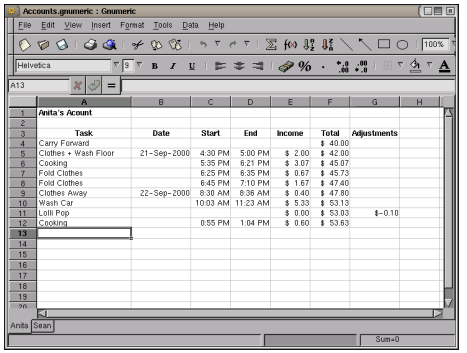
Parametrisations

What are standards good for (1)?







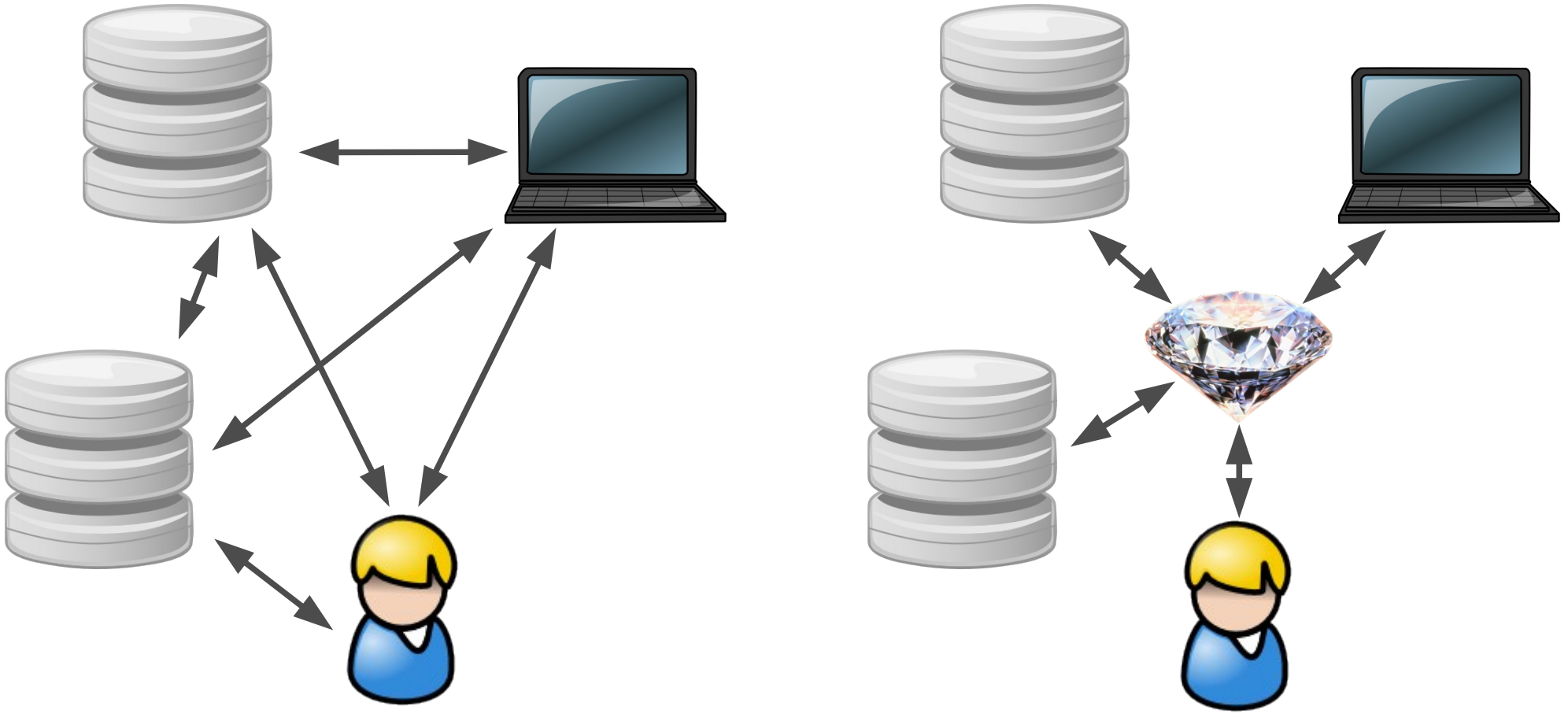


$$\int_1^t \frac{dx}{x}$$

```
<int/>
  <bvar><ci>x</ci></bvar>
  <lowlimit><cn>1</cn></lowlimit>
  <uplimit><ci>t</ci></uplimit>
  <apply>
    <divide/>
      <cn>1</cn>
      <ci>x</ci>
  </apply>
</int>
```

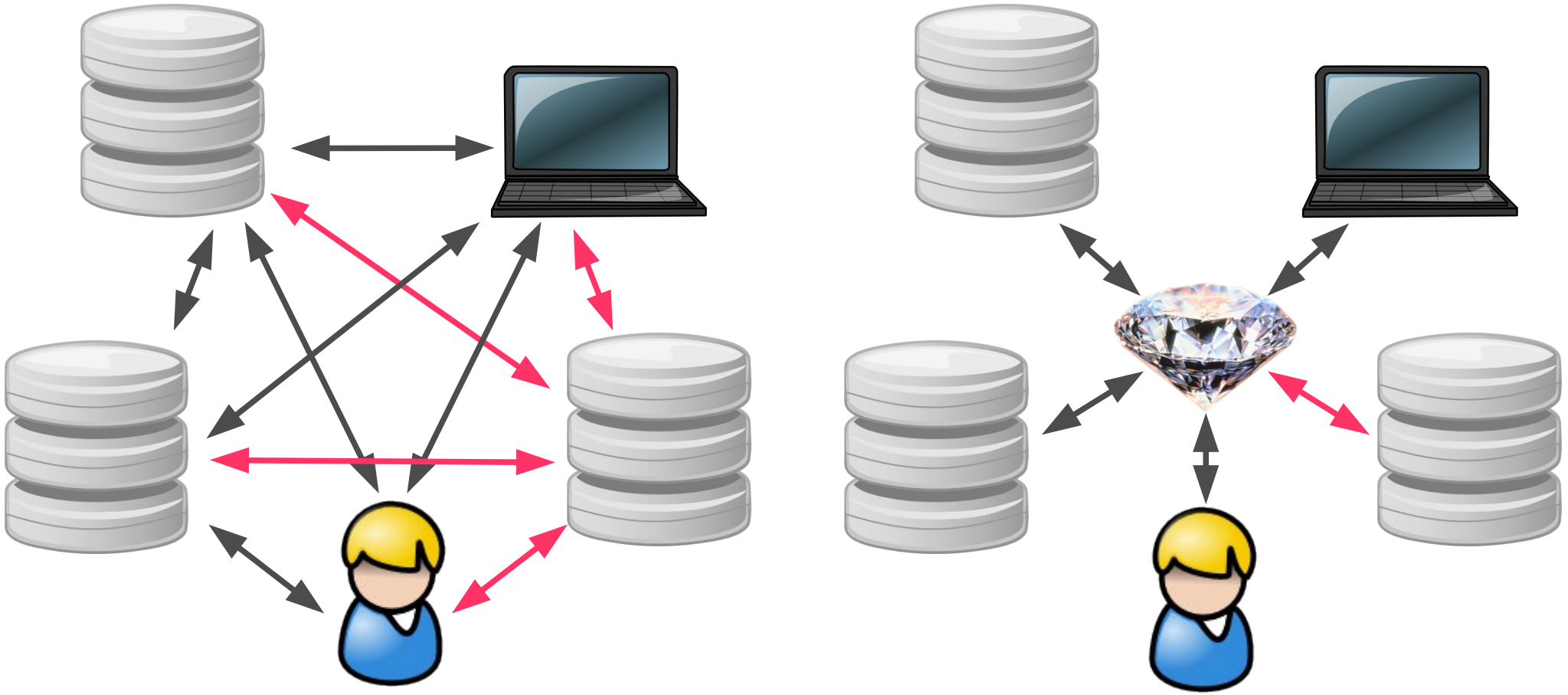
Formal languages can be read and written by computers

What are standards good for (2)?



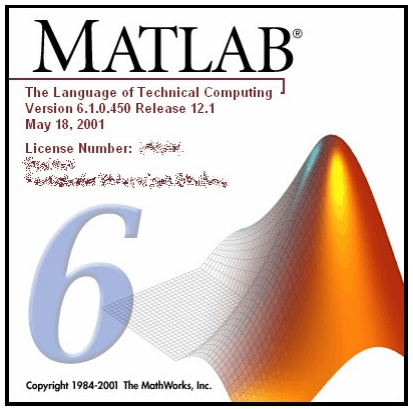
N tools require N conversions for exchange and not N(N-1)/2

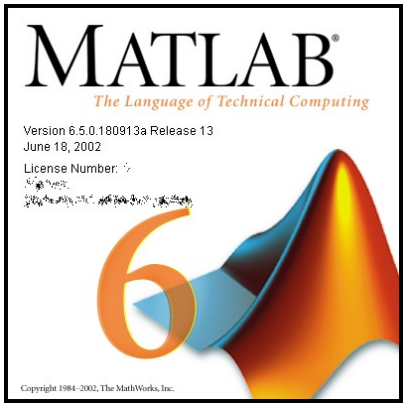
What are standards good for (2)?

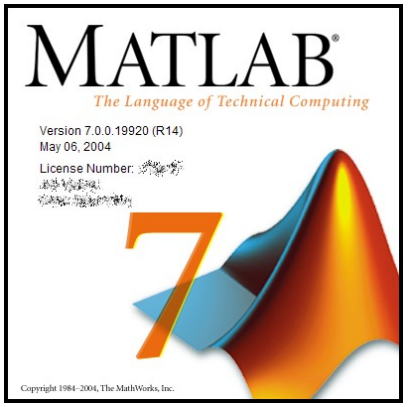


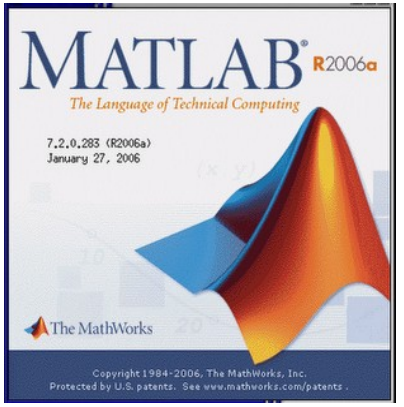
N tools require N conversions for exchange and not N(N-1)/2

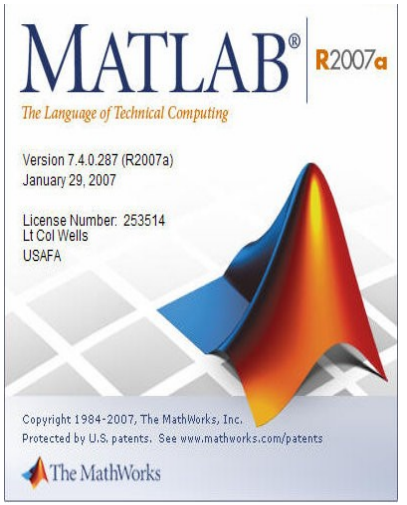
What are standards good for (3)?

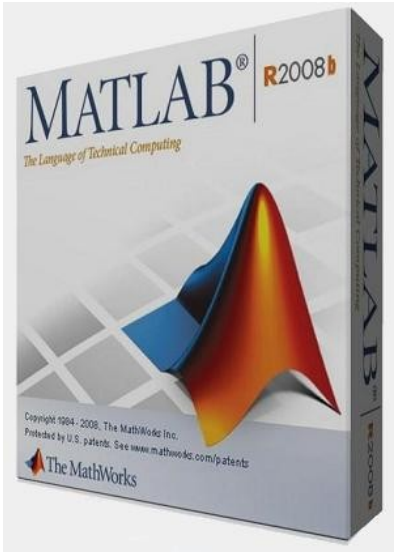


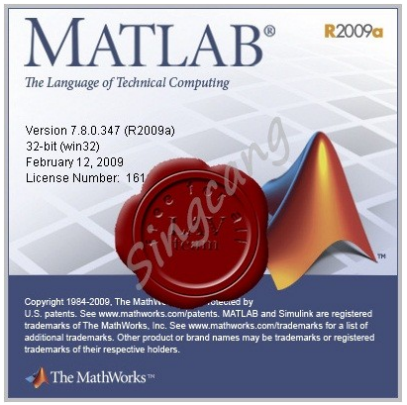


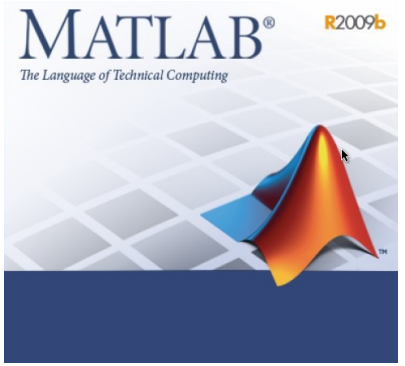


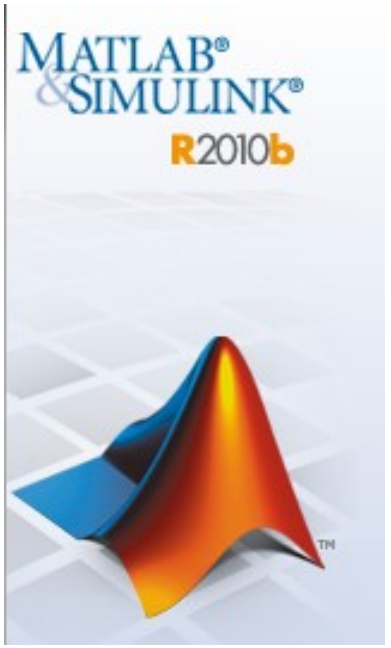


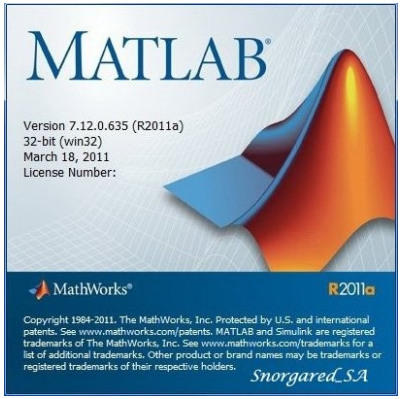






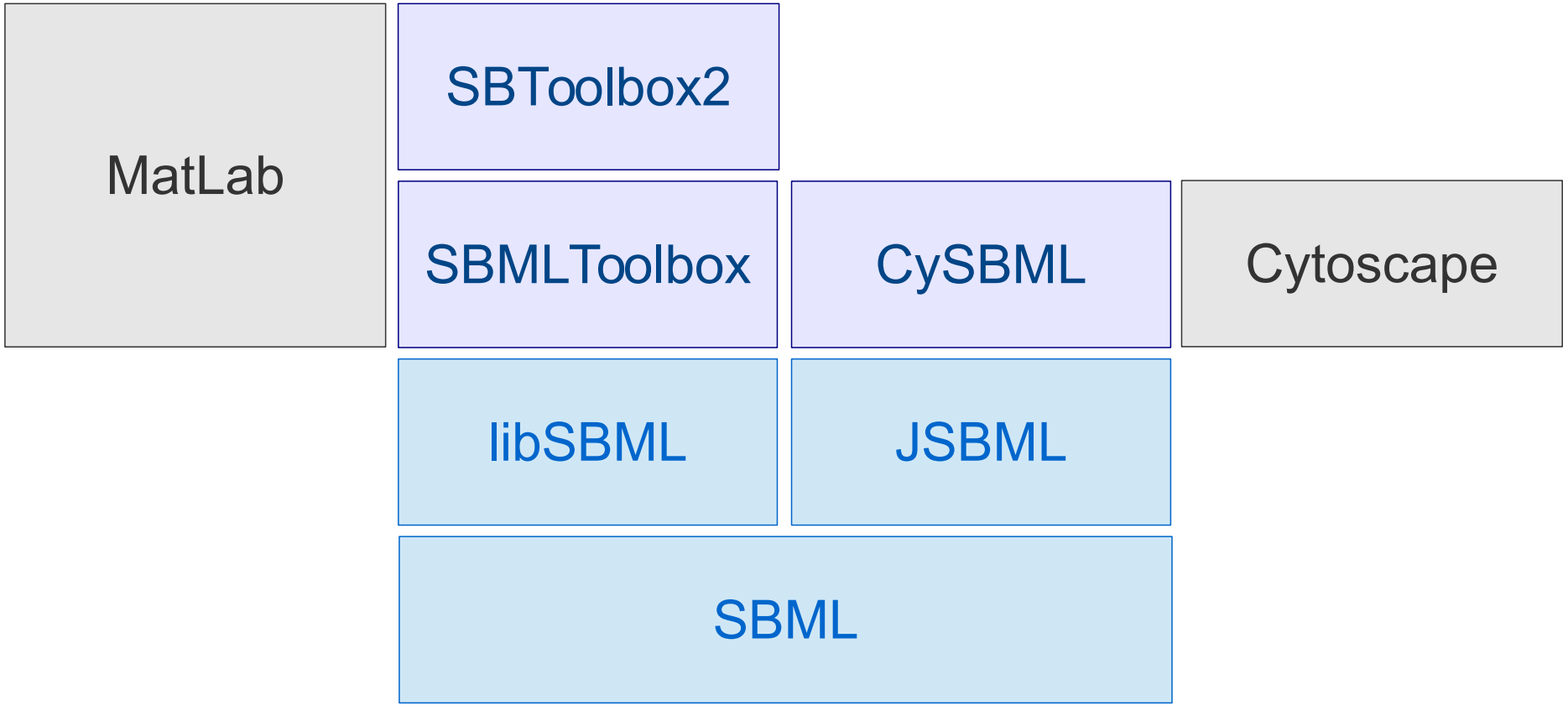






Open standards are more stable than proprietary formats

What are standards good for (4)?



Open standards can be built on. They generate new science

Three types of standards

Minimal requirements	What to encode in order to share experiments and understand results 
Data-models	How to encode the information defined above in a computer-readable manner 
Terminologies	Structured representation of knowledge, with concept definitions and their relationships 

A language to describe computational models in biology

	Model descriptions
Data-models	SBML

Born in Caltech 2000

John Doyle

Hiroaki Kitano

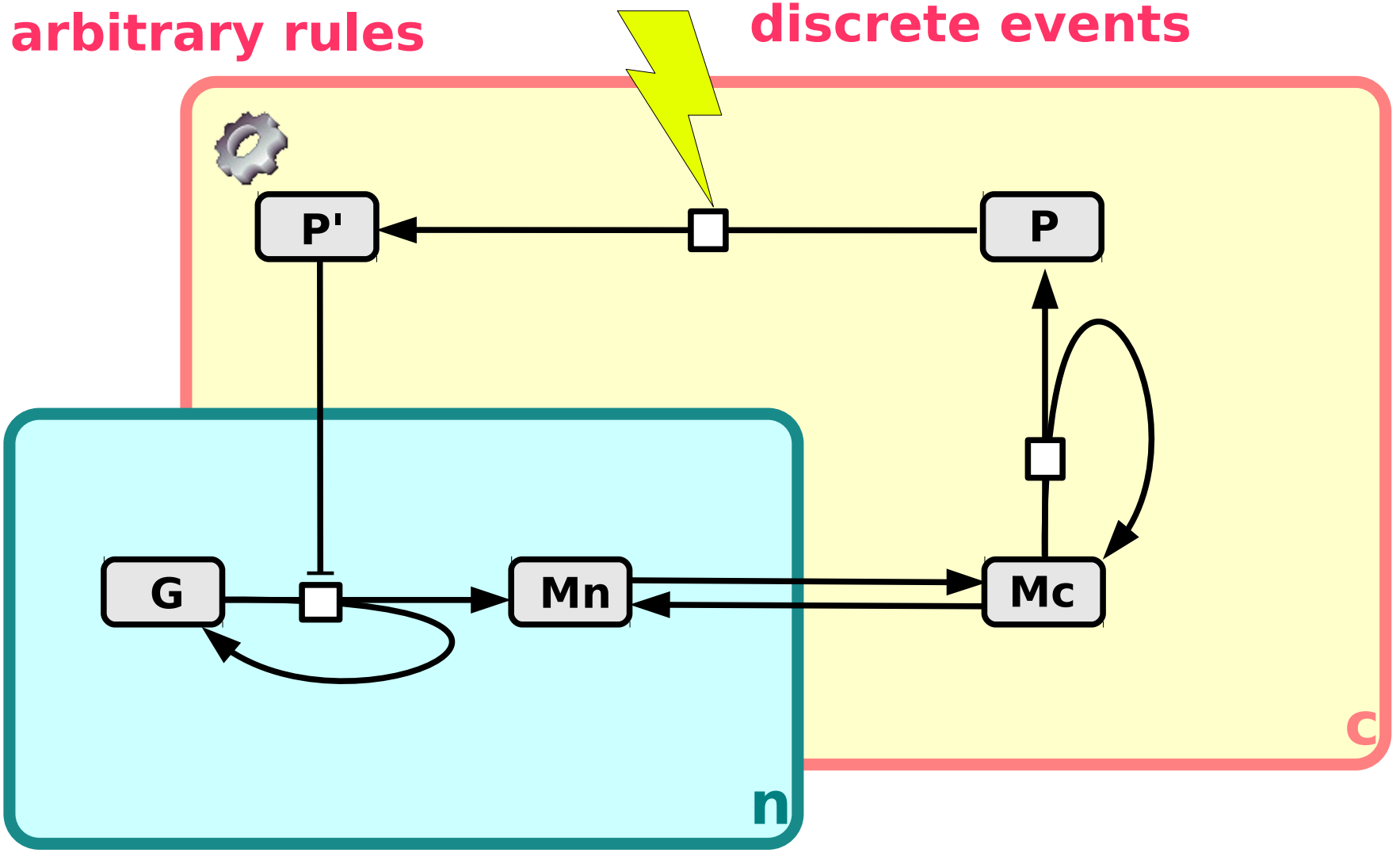
Hamid Bolouri

Mike Hucka

Andrew Finney

Herbert Sauro

What can we encode in SBML (core)?



Systems Biology Markup Language

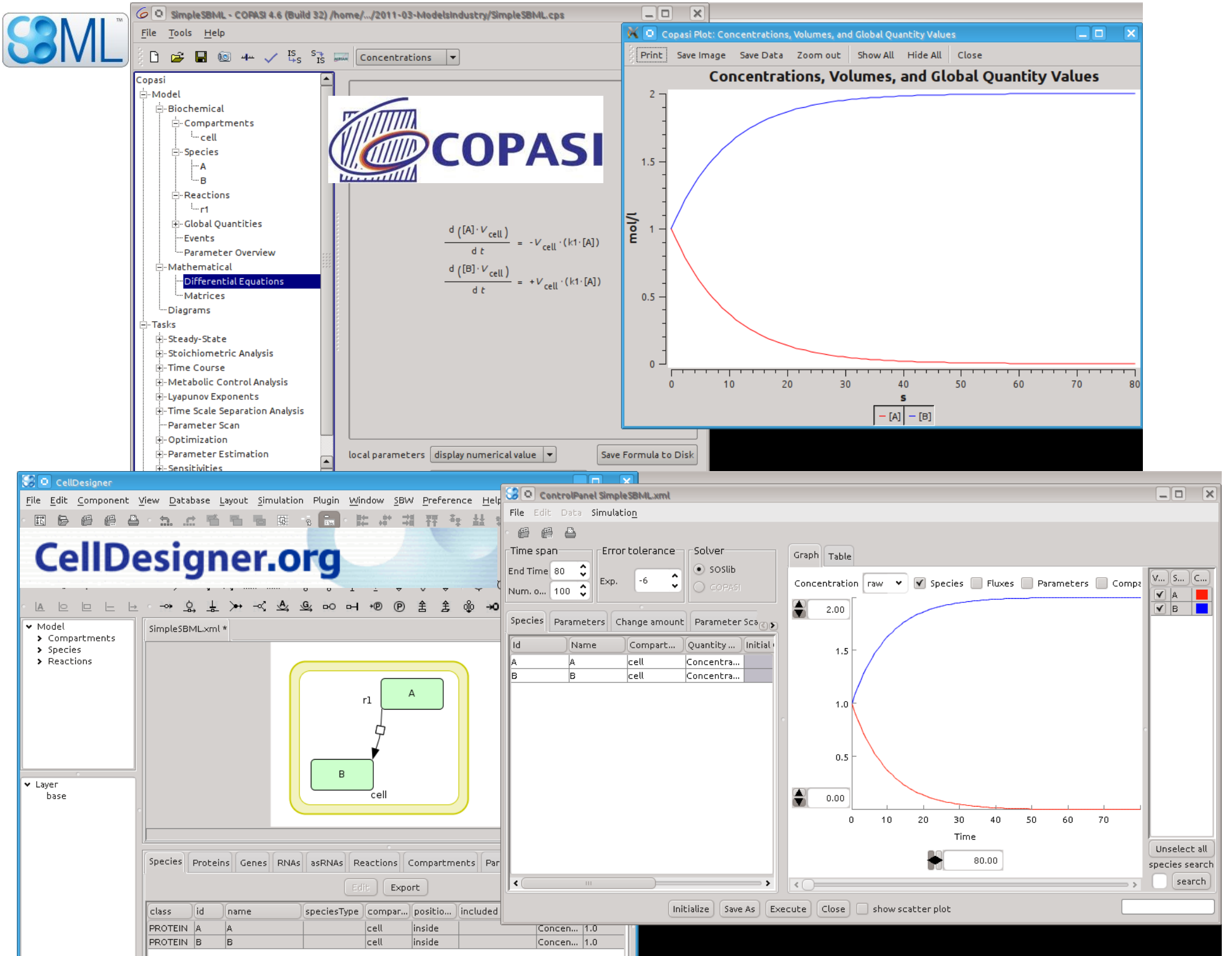
```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="3" version="1"·
      xmlns="http://www.sbml.org/sbml/level3/version1/core">
  <model>
    <listOfFunctionDefinitions> <!-- --> </listOfFunctionDefinitions>
    <listOfUnitDefinitions> <!-- --> </listOfUnitDefinitions>
    <listOfCompartments> <!-- --> </listOfCompartments>
    <listOfSpecies> <!-- --> </listOfSpecies>
    <listOfParameters> <!-- --> </listOfParameters>
    <listOfInitialAssignments> <!-- --> </listOfInitialAssignments>
    <listOfRules> <!-- --> </listOfRules>
    <listOfConstraints> <!-- --> </listOfConstraints>
    <listOfReactions> <!-- --> </listOfReactions>
    <listOfEvents> <!-- --> </listOfEvents>
  </model>
</sbml>
```

variables

relationships

```xml
<?xml version="1.0" encoding="UTF-8"?>
<sbml xmlns="http://www.sbml.org/sbml/level2/version4" level="2" version="4">
  <model name="Simple Model">
    <listOfCompartments>
      <compartment id="cell" size="1" />
    </listOfCompartments>
    <listOfSpecies>
      <species id="A" compartment="cell" initialConcentration="1"/>
      <species id="B" compartment="cell" initialConcentration="1"/>
    </listOfSpecies>
    <listOfParameters>
      <parameter id="k1" value="0.1"/>
    </listOfParameters>
    <listOfReactions>
      <reaction id="r1" reversible="false">
      <listOfReactants>
          <speciesReference species="A"/>
        </listOfReactants>
        <listOfProducts>
          <speciesReference species="B"/>
        </listOfProducts>
        <kineticLaw>
          <math xmlns="http://www.w3.org/1998/Math/MathML">
            <apply>
              <times/>
              <ci> cell </ci>
              <ci> k1 </ci>
              <ci> A </ci>
            </apply>
          </math>
        </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```

A very simple SBML file (A ⟶ B)


SBML
SimpleSBML - COPASI 4.6 (Build 32) /home/.../2011-03-ModelsIndustry/SimpleSBML.cps
File Tools Help
Concentrations
Copasi
Model
Biochemical
Compartments
cell
Species
A
B
Reactions
r1
Global Quantities
Events
Parameter Overview
Mathematical
Differential Equations
Matrices
Diagrams
Tasks
Steady-State
Stoichiometric Analysis
Time Course
Metabolic Control Analysis
Lyapunov Exponents
Time Scale Separation Analysis
Parameter Scan
Optimization
Parameter Estimation
Sensitivities
COPASI
d([A]·V_cell)/dt = -V_cell·(k1·[A])
d([B]·V_cell)/dt = +V_cell·(k1·[A])
local parameters
display numerical value
Save Formula to Disk
Copasi Plot: Concentrations, Volumes, and Global Quantity Values
Print Save Image Save Data Zoom out Show All Hide All Close
Concentrations, Volumes, and Global Quantity Values
mol/l
s
[A] [B]
CellDesigner
File Edit Component View Database Layout Simulation Plugin Window SBW Preference Help
CellDesigner.org
Model
Compartments
Species
Reactions
Layer
base
SimpleSBML.xml *
r1
A
B
cell
Species Proteins Genes RNAs asRNAs Reactions Compartments
Edit Export
class id name speciesType compar... positio... included
PROTEIN A A cell inside Concen... 1.0
PROTEIN B B cell inside Concen... 1.0
ControlPanel SimpleSBML.xml
File Edit Data Simulation
Time span
End Time 80
Num. o... 100
Error tolerance
Exp. -6
Solver
SOSlib
COPASI
Species Parameters Change amount Parameter Sca
Id Name Compart... Quantity ... Initial
A A cell Concentra...
B B cell Concentra...
Graph Table
Concentration raw
Species Fluxes Parameters Compa
2.00
0.00
Time
80.00
V... S... C...
A
B
Unselect all
species search
search
Initialize Save As Execute Close show scatter plot

A community-driven global reconstruction of human metabolism

Ines Thiele[1,2,37], Neil Swainston[3,4,37], Ronan M T Fleming[1,5], Andreas Hoppe[6], Swagatika Sahoo[1], Maike K Aurich[1], Hulda Haraldsdottir[1], Monica L Mo[7], Ottar Rolfsson[1], Miranda D Stobbe[8,9], Stefan G Thorleifsson[1], Rasmus Agren[10], Christian Bölling[6], Sergio Bordel[10], Arvind K Chavali[11], Paul Dobson[12], Warwick B Dunn[3,13], Lukas Endler[14], David Hala[15], Michael Hucka[16], Duncan Hull[4], Daniel Jameson[3,4], Neema Jamshidi[7], Jon J Jonsson[5], Nick Juty[17], Sarah Keating[17], Intawat Nookaew[10], Nicolas Le Novère[17,18], Naglis Malys[3,19,20], Alexander Mazein[21], Jason A Papin[11], Nathan D Price[22], Evgeni Selkov, Sr[23], Martin I Sigurdsson[1], Evangelos Simeonidis[22,24], Nikolaus Sonnenschein[25], Kieran Smallbone[3,26], Anatoly Sorokin[21,27], Johannes H G M van Beek[28–30], Dieter Weichart[3,31], Igor Goryanin[21,32], Jens Nielsen[10], Hans V Westerhoff[3,28,33,34], Douglas B Kell[3,35], Pedro Mendes[3,4,36] & Bernhard Ø Palsson[1,7]

Multiple models of human metabolism have been reconstructed, but each represents only a subset of our knowledge. Here we describe Recon 2, a community-driven, consensus 'metabolic reconstruction', which is the most comprehensive representation of human metabolism that is applicable to computational modeling. Compared with its predecessors, the reconstruction has improved topological and functional features, including ~2× more reactions and ~1.7× more unique metabolites. Using Recon 2 we predicted changes in metabolite biomarkers for 49 inborn errors of metabolism with 77% accuracy when compared to experimental data. Mapping metabolomic data and drug information onto Recon 2 demonstrates its potential for integrating and analyzing diverse data types. Using protein expression data, we automatically generated a compendium of 65 cell type–specific models, providing a basis for manual curation or investigation of cell-specific metabolic properties. Recon 2 will facilitate many future biomedical studies and is freely available at http://humanmetabolism.org/.

An understanding of metabolism is fundamental to comprehending the phenotypic behavior of all living organisms, including humans, where metabolism is integral to health and is involved in much of human disease. High quality, genome-scale 'metabolic reconstructions' are at the heart of bottom-up systems biology analyses and represent the entire network of metabolic reactions that a given organism is known to exhibit[1]. The metabolic-network reconstruction procedure is now well-established[2] and has been applied to a growing number of model organisms[3]. Metabolic reconstructions allow for the conversion of biological knowledge into a mathematical format and the subsequent computation of physiological states[1,4,5] to address a variety of scientific and applied questions[3,6]. Reconstructions enable network-wide mechanistic investigations of the genotype-phenotype relationship. A high-quality reconstruction of the metabolic network is thus

[1]Center for Systems Biology, University of Iceland, Reykjavik, Iceland. [2]Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, University of Iceland, Reykjavik, Iceland. [3]Manchester Centre for Integrative Systems Biology, University of Manchester, Manchester Institute of Biotechnology, Manchester, UK. [4]School of Computer Science, University of Manchester, Manchester, UK. [5]Department of Biochemistry and Molecular Biology, University of Iceland, Reykjavik, Iceland. [6]Computational Systems Biochemistry Group, Charité– Universitätsmedizin Berlin, Berlin, Germany. [7]Department of Bioengineering, University of California, San Diego, La Jolla, California, USA. [8]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands. [9]Netherlands Bioinformatics Centre, Nijmegen, the Netherlands. [10]Department of Chemical and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden. [11]Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia, USA. [12]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK. [13]Centre for Advanced Discovery and Experimental Therapeutics (CADET), Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Sciences Centre, Manchester, UK. [14]Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria. [15]Department of Biology, University of North Texas, Denton, Texas, USA. [16]Computing and Mathematical Sciences Department, California Institute of Technology, Pasadena, California, USA. [17]European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, UK. [18]Babraham Institute, Babraham Research Campus, Cambridge, UK. [19]Faculty of Life Sciences, University of Manchester, Manchester, UK. [20]School of Life Sciences, Gibbet Hill Campus, University of Warwick, Coventry, UK. [21]School of Informatics, University of Edinburgh, Edinburgh, UK. [22]Institute for Systems Biology, Seattle, Washington, USA. [23]Genome Designs, Inc., Walnut Creek, California, USA. [24]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Campus Belval, Esch-sur-Alzette, Luxembourg. [25]School of Engineering and Science, Jacobs University Bremen, Bremen, Germany. [26]School of Mathematics, University of Manchester, Manchester, UK. [27]Institute of Cell Biophysics, Russian Academy of Sciences, Moscow region, Pushchino, Russia. [28]Department of Molecular Cell Physiology, Vrije Universiteit, Amsterdam, the Netherlands. [29]Section Medical Genomics, Department of Clinical Genetics, Vrije Universiteit University Medical Centre, Amsterdam, the Netherlands. [30]Netherlands Consortium for Systems Biology, Amsterdam, The Netherlands. [31]School of Dentistry, The University of Manchester, Manchester, UK. [32]Okinawa Institute Science and Technology, Okinawa, Japan. [33]School of Chemical Engineering and Analytical Science, University of Manchester, Manchester, UK. [34]Swammerdam Institute for Life Sciences, Faculty of Science, University of Amsterdam, Amsterdam, The Netherlands. [35]School of Chemistry, The University of Manchester, Manchester, UK. [36]Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, Virginia, USA. [37]These authors contributed equally to this work. Correspondence should be addressed to I.T. (ines.thiele@gmail.com).

Received 7 September 2012; accepted 19 December 2012; published online 3 March 2013; doi:10.1038/nbt.2488

A not so simple SBML file (Recon2)

- 8 compartments
- 5 063 metabolites
- 2 194 proteins
- 7 440 reactions

- Core package – public specification
- Flux balance constraint – public specification
- Qualitative models – public specification
- Model composition – public specification
- Graph Layout – public specification
- Graph rendering – specification finalised
- Complex species – specification finalised
- Groups - specification finalised
- Distributions and ranges - specification under discussion
- Spatial diffusion – specification under discussion
- Enhanced metadata – specification proposed
- Arrays and sets – specification proposed
- Dynamic structures - discussed

SBML Level 3 is modular

SBML Software Guide

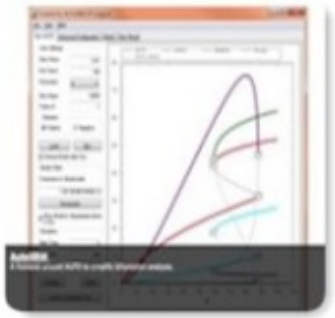
The following pages describe SBML-compatible software packages known to us. We offer different ways of viewing the information, all drawn from the same underlying data collected from the systems' developers via our **software survey**. The *Matrix* provides a table listing all known software and a variety of their features; the *Summary* provides general descriptions of most of the software; and the *Showcase* provides a sequential slideshow of a subset of the software.

Number of software packages listed in the matrix today: **254**.

Go to the SBML Software Matrix

Go to the SBML Software Summary

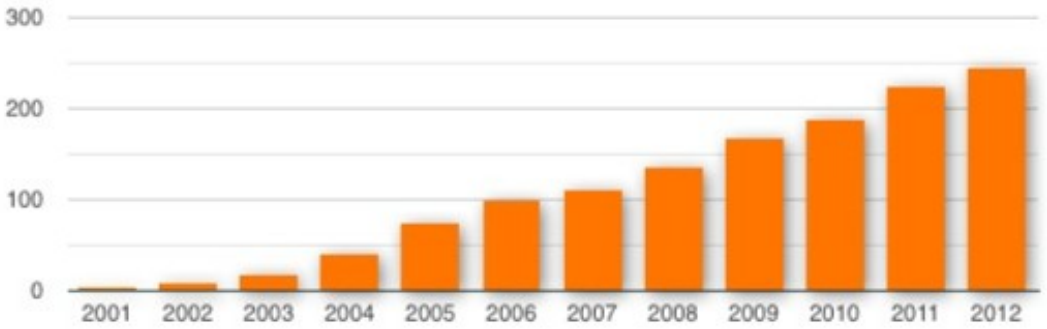
Go to the SBML Software Showcase



Please **use the survey form** to notify us about additions and suggestions.

Historical trend

The following graph shows the total number of known SBML-compatible software packages each year, as counted by the SBML Team. The counts shown are for approximately the middle of each year.

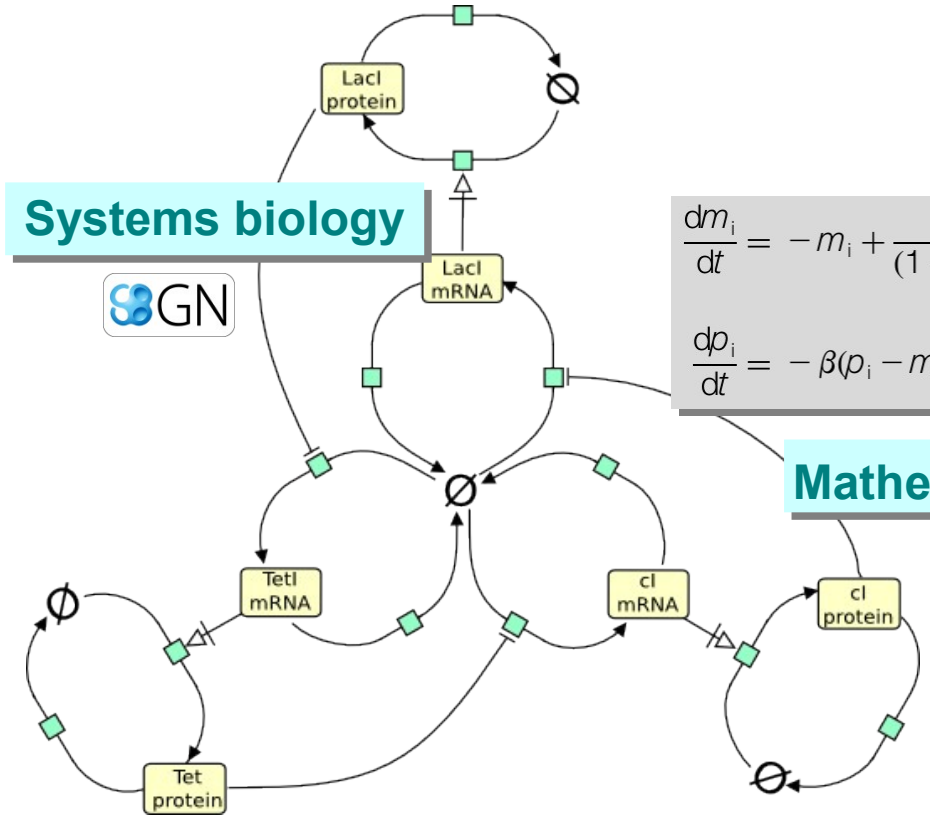


(Note: the flat period in 2007 is an artifact of inadequate record keeping rather than a lull in SBML software development.)

SBML conformance testing

The **SBML Test Suite** provides an operational means of testing SBML support in software simulation and analysis systems. Software authors can choose to make the test results for their software public in the **SBML Test Suite Database**, where you can

BioModels Database – http://www.ebi.ac.uk/biomodels



Systems biology

SBGN

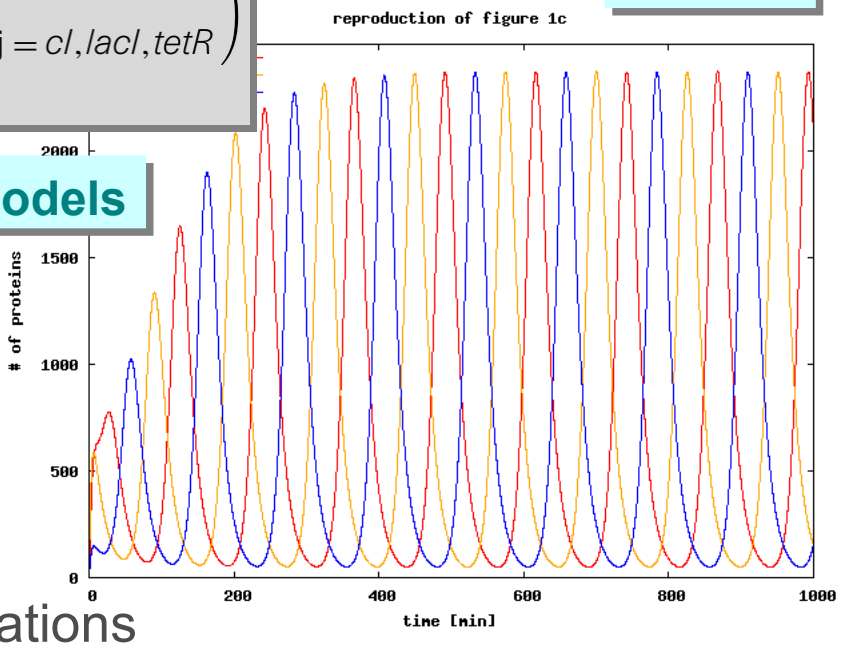
$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{(1+p_j^n)} + \alpha_0 \qquad \begin{pmatrix} i = lacI, tetR, cI \\ j = cI, lacI, tetR \end{pmatrix}$$

$$\frac{dp_i}{dt} = -\beta(p_i - m_i)$$

Mathematical models

SBML

curated



>140 000 models, >10 millions mathematical relations

Deposition advised by >300 journals, database > 600 citations

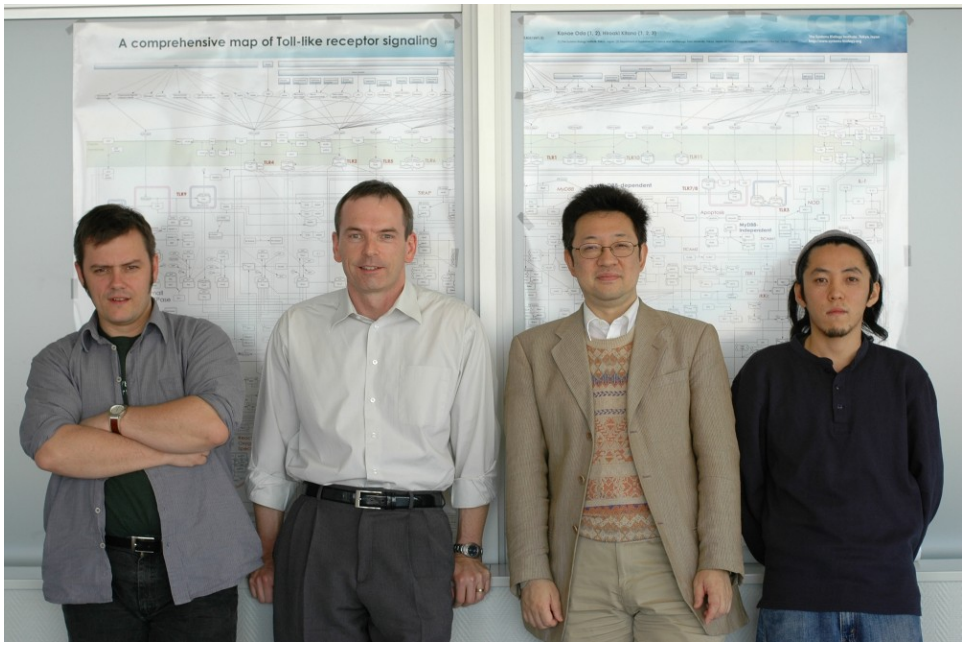
1 million page requests per year

Submission in SBML and CellML; Export in SBML, CellML, XPP, SciLab, BioPAX, Octave, PDF, VCML, SBGN

The interface with all biologists

	Model descriptions
Data-models	SBML™ SBGN

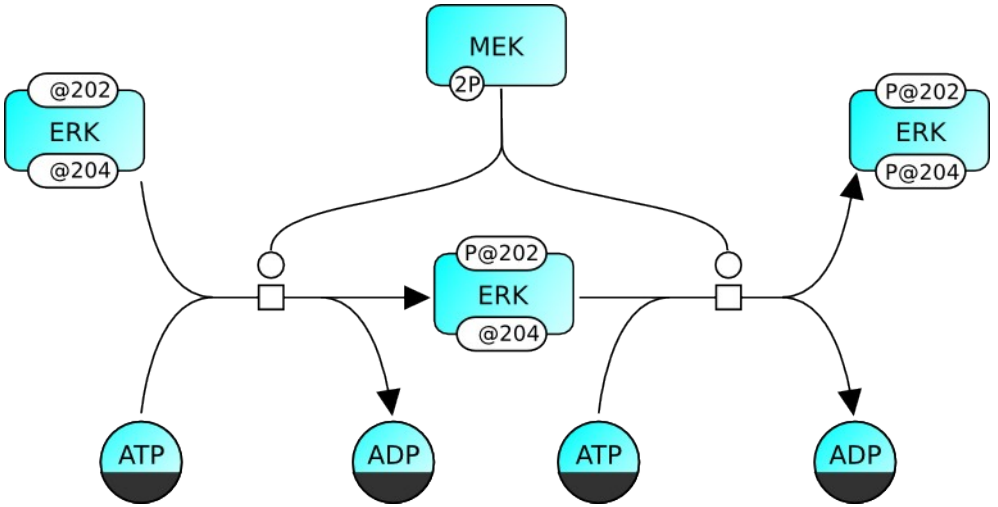
Born in Tokyo 2005



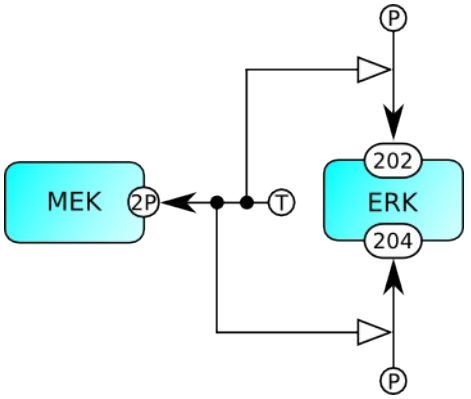
Akira Funahashi

Graph trinity: three languages in one notation

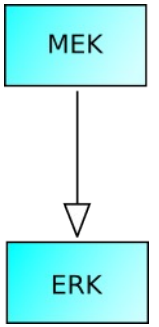
Process Descriptions



Entity Relationships



Activity Flows

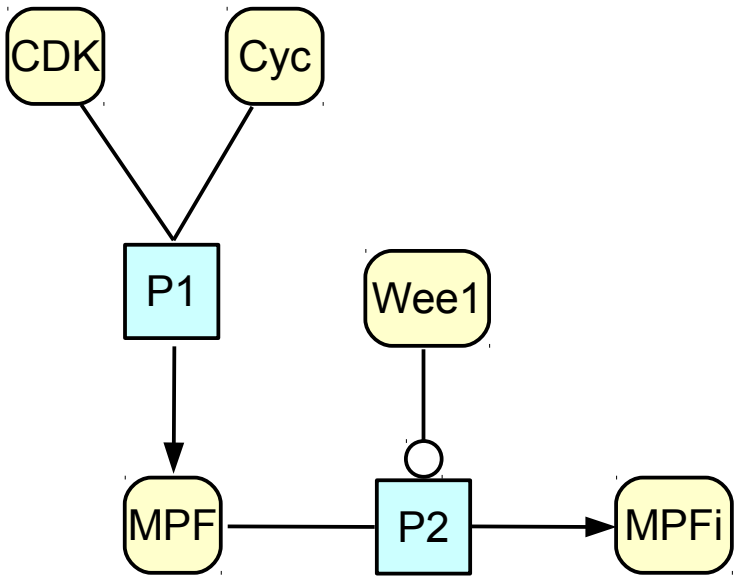


http://sbgn.org

Babraham Institute

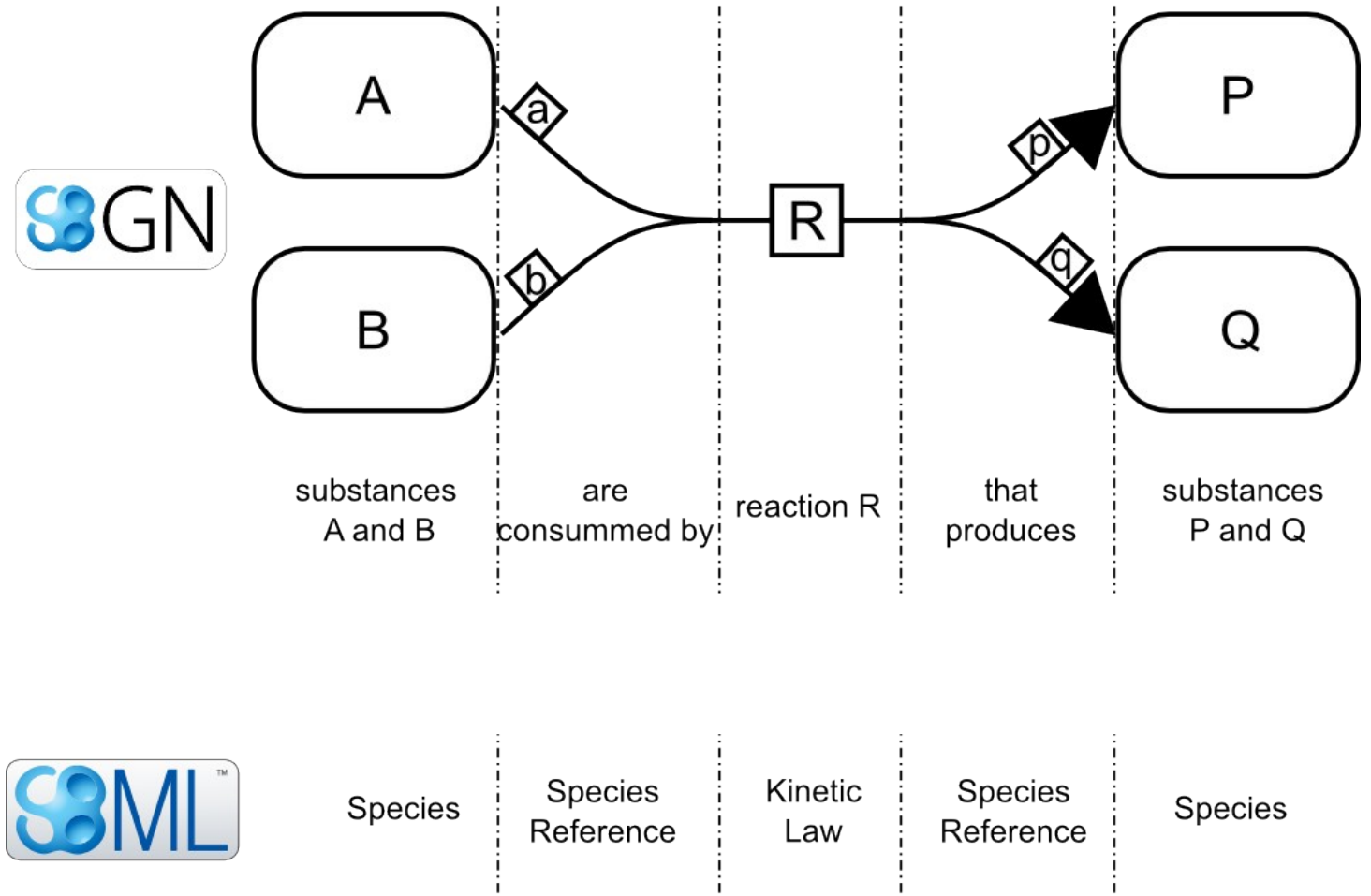
EMBL-EBI

Process Descriptions

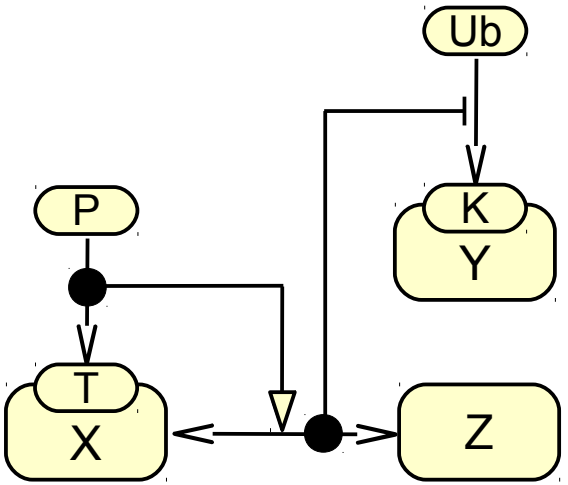


- Biochemistry, Metabolic networks
- Nodes represent populations of reactants and reactions
- Directional
- Sequential
- Mechanistic
- “Close” world
- Subjected to combinatorial explosion
- Process modelling (ODE, SSA)
  SBML core

A biochemical reaction as a process

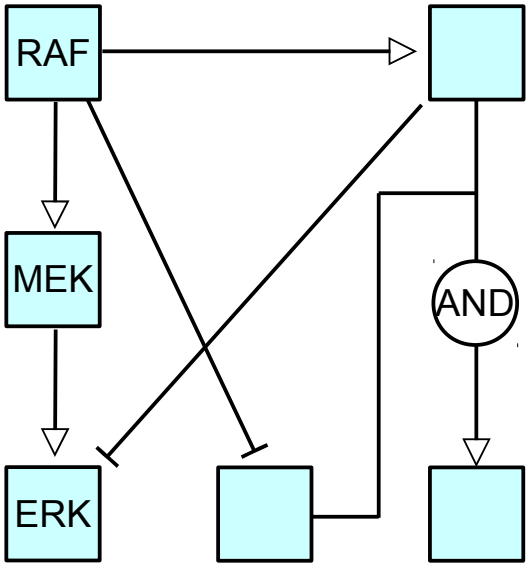


Entity Relationships



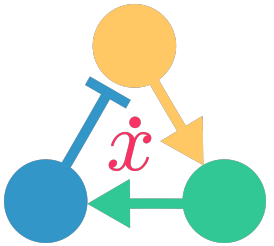
- Molecular Biology
- Nodes represent existing entities
- Directional
- Non-sequential
- Mechanistic
- “Open world”
- Independent rules: no explosion
- Rule-based modelling
  SBML multi

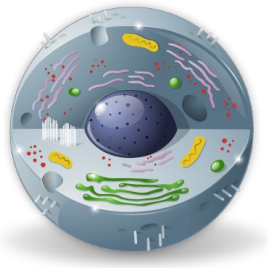
Activity-flows (influence diagrams)



- Signalling pathways, gene regulatory networks
- Nodes represent activities
- Directional
- Sequential
- Non-mechanistic
- Logical modelling SBML qual

Is this sufficient for model sharing and re-use?





Ruminations on Creating a Biomodels.net

Mike Hucka
mhucka@caltech.edu
The SBML Team
Control and Dynamical Systems, MC 107-81
California Institute of Technology, Pasadena, CA 91125, USA
http://www.sbml.org/

DRAFT

May 14, 2004

1 An Introduction

There are currently several project threads that have been floating around and I'd like to propose a project that I think would bring some of them together into a unique and globally useful resource. These threads include the following:

DRAFT

May 14, 2004

BioModels Database

1 An Introduction

There are currently several project threads that have been floating around and I'd like to propose a project that I think would bring some of them together into a unique and globally useful resource. These threads include the following:

MIRIAM annotations

- People have been repeatedly voicing the desire for a better database of SBML models. The existing repository on `sbml.org` was never meant to be a full database, only a stop-gap collection until such time as we can find the resources to do better. It may be time to organize a real database.
- People want databases of models to connect to other databases, especially databases of molecular information, Medline, etc.
- We have connections now to EBI and other groups who have expertise in databases and an eagerness to connect them to computational modeling efforts in systems biology. The time is ripe for leveraging these connections.
- The SBML and CellML groups have a long-standing interest to connect not only to each other's efforts, but to others such as BioPAX.

COMBINE

- We have technology now to translate a substantial portion of models from CellML to SBML. Technology exists for translating SBML Level 1 to CellML, and although it doesn't exist for SBML Level 2, there is good reason to believe it could be implemented. It is therefore highly likely that a significant portion of models in practice can be cross-translated.

SBO

- The CellML group has been working on ontologies for models which may ultimately be an ideal basis for a unifying database schema that could serve as a master schema for storing models in format-independent manner. (It is worth remembering that SBML was never designed to be a database representation. It may not be suitable for that role. Ditto for CellML. Internal to a database, it may make sense to use a root schema from which projections of the contents of a model are then translated out to SBML, BioPAX, CellML, whatever.)

Adding the semantics to the syntax

	Model descriptions
Minimal requirements	MIRIAM
Data-models	SBML
Terminologies	SBO

Born in Heidelberg 2004

Minimal Information Required In the Annotation of Models

Reference correspondence

1. In a public, standardized, machine-readable format
2. Comply with the standard in which it is encoded
3. Clearly related to a single reference description
4. Reflect biological processes
5. Instantiable in a simulation all numbers provided
6. Able to reproduce results

http://co.mbine.org/standards/miriam

Minimal Information Required In the Annotation of Models

Reference correspondence

1. In a public, standardized, machine-readable format
2. Comply with the standard in which it is encoded
3. Clearly related to a single reference description
4. Reflect biological processes
5. Instantiable in a simulation all numbers provided
6. Able to reproduce results

Attribution

1. Has to be named
2. Citation must be provided
3. Model creators details
4. Date and time of creation and last modification
5. Link to precise statement about terms of distribution

http://co.mbine.org/standards/miriam

Minimal Information Required In the Annotation of Models

Reference correspondence

1. In a public, standardized, machine-readable format
2. Comply with the standard in which it is encoded
3. Clearly related to a single reference description
4. Reflect biological processes
5. Instantiable in a simulation all numbers provided
6. Able to reproduce results

Attribution

1. Has to be named
2. Citation must be provided
3. Model creators details
4. Date and time of creation and last modification
5. Link to precise statement about terms of distribution

External resources

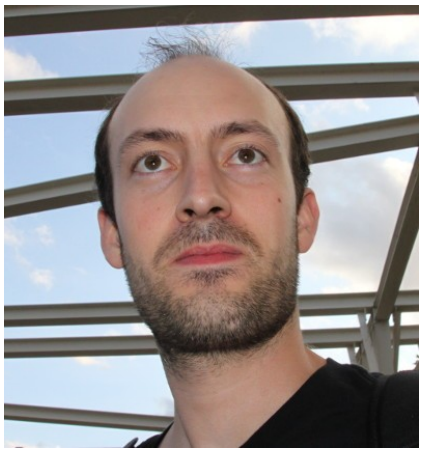
1. Annotation unambiguously model constituent to data
2. Link to external information as a triplet {collection, identifier, qualifier}
3. Annotation written as a Uniform Resource Identifier
4. Identifier considered within framework of the collection.
5. Collection namespace and record identifier in one URI
6. Qualifiers to refine the link between model constituent and external knowledge
7. Standard set of valid URIs agreed upon by community

http://co.mbine.org/standards/miriam

identifiers.org (aka new MIRIAM URIs)

`http://identifiers.org/namespace/identifier`

protocol — type of URI — collection — record

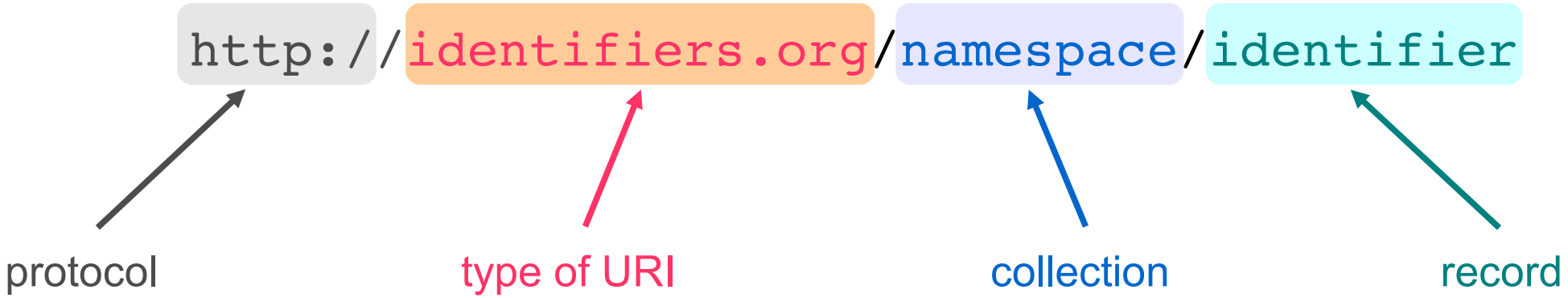


Camille Laibe

Nick Juty

Sarala Wimalaratne

identifiers.org (aka new MIRIAM URIs)



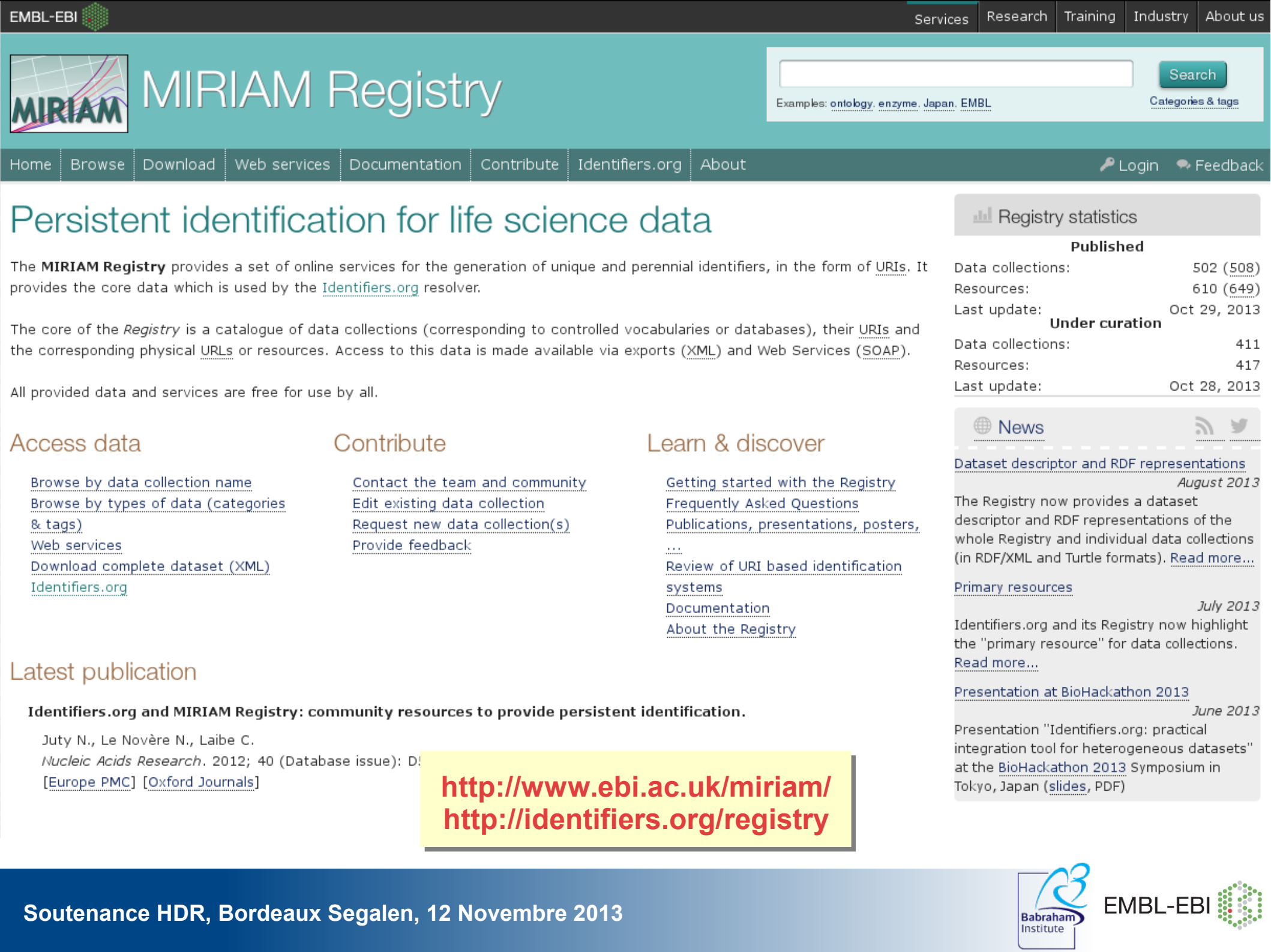
http://identifiers.org/pubmed/22140103

http://identifiers.org/ec-code/1.1.1.1

http://identifiers.org/go/GO:0000186

SBML and MIRIAM cross-references

```
<species id="ca_calmodulin" metaid="cacam">
  <annotation>
    <rdf:RDF
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/">
      <rdf:Description rdf:about="#cacam">
        <bqbiol:hadPart>
          <rdf:Bag>
            <rdf:li rdf:resource="http://identifiers.org/uniprot/62158"/>
            <rdf:li rdf:resource="http://identifiers.org/chebi/CHEBI:29108"/>
          </rdf:Bag>
        </bqbiol:hasPart>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```


EMBL-EBI
Services
Research
Training
Industry
About us
MIRIAM
MIRIAM Registry
Search
Examples: ontology, enzyme, Japan, EMBL
Categories & tags
Home
Browse
Download
Web services
Documentation
Contribute
Identifiers.org
About
Login
Feedback
Persistent identification for life science data
The MIRIAM Registry provides a set of online services for the generation of unique and perennial identifiers, in the form of URIs. It provides the core data which is used by the Identifiers.org resolver.
The core of the Registry is a catalogue of data collections (corresponding to controlled vocabularies or databases), their URIs and the corresponding physical URLs or resources. Access to this data is made available via exports (XML) and Web Services (SOAP).
All provided data and services are free for use by all.
Access data
Browse by data collection name
Browse by types of data (categories & tags)
Web services
Download complete dataset (XML)
Identifiers.org
Contribute
Contact the team and community
Edit existing data collection
Request new data collection(s)
Provide feedback
Learn & discover
Getting started with the Registry
Frequently Asked Questions
Publications, presentations, posters, ...
Review of URI based identification systems
Documentation
About the Registry
Latest publication
Identifiers.org and MIRIAM Registry: community resources to provide persistent identification.
Juty N., Le Novère N., Laibe C.
Nucleic Acids Research. 2012; 40 (Database issue): D
[Europe PMC] [Oxford Journals]
Registry statistics
Published
Data collections: 502 (508)
Resources: 610 (649)
Last update: Oct 29, 2013
Under curation
Data collections: 411
Resources: 417
Last update: Oct 28, 2013
News
Dataset descriptor and RDF representations
August 2013
The Registry now provides a dataset descriptor and RDF representations of the whole Registry and individual data collections (in RDF/XML and Turtle formats). Read more...
Primary resources
July 2013
Identifiers.org and its Registry now highlight the "primary resource" for data collections. Read more...
Presentation at BioHackathon 2013
June 2013
Presentation "Identifiers.org: practical integration tool for heterogeneous datasets" at the BioHackathon 2013 Symposium in Tokyo, Japan (slides, PDF)
http://www.ebi.ac.uk/miriam/
http://identifiers.org/registry
Soutenance HDR, Bordeaux Segalen, 12 Novembre 2013
Babraham Institute
EMBL-EBI

Data collection: *Enzyme Nomenclature*

Available as HTML or RDF, through browsers or Web Services

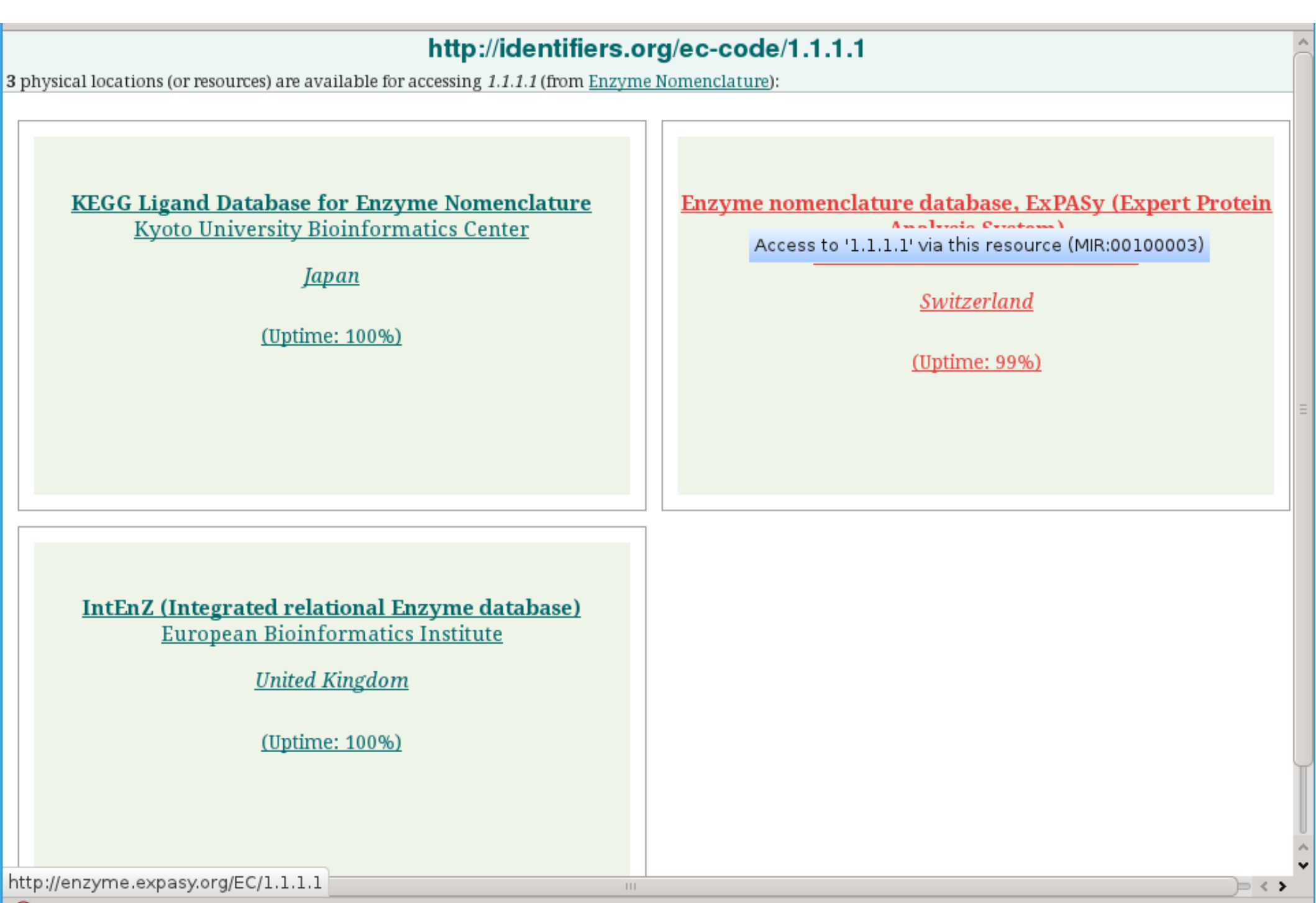
Overview | Categories | Miscellaneous

Overview of the data collection

Name		
Identifier		MIR:00000004
Name		Enzyme Nomenclature
Synonyms		Enzyme Classification
		EC code
		EC
Information		
Definition		The Enzyme Classification contains the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions.
Identifier pattern		^\d+\.-\.-\.-|\d+\.\d+\.-\.-|\d+\.\d+\.\d+\.-|\d+\.\d+\.\d+\.(n)?\d+$
URIs		
Namespace		ec-code
Root URL		http://identifiers.org/ec-code/
Root URN		urn:miriam:ec-code:
Physical Locations		
Resource MIR:00100002	**Access URL**	http://www.genome.jp/dbget-bin/www_bget?ec:**$id** [Example: 1.1.1.1]
	Website	http://www.genome.jp/dbget-bin/www_bfind?enzyme
	Description	KEGG Ligand Database for Enzyme Nomenclature
	Institution	Kyoto University Bioinformatics Center, Japan
Resource MIR:00100001	**Access URL**	http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=**$id** [Example: 1.1.1.1]
	Website	http://www.ebi.ac.uk/intenz/
	Description	IntEnZ (Integrated relational Enzyme database)
	Institution	European Bioinformatics Institute, United Kingdom
Resource MIR:00100003	**Access URL**	http://enzyme.expasy.org/EC/**$id** [Example: 1.1.1.1]
	Website	http://enzyme.expasy.org/
	Description	Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
	Institution	Swiss Institute of Bioinformatics, Switzerland
References		
http://www.chem.qmul.ac.uk/iubmb/enzyme/		
http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMID:10812475]		

Go back to the list of data collections

Suggest modifications to this data collection


http://identifiers.org/ec-code/1.1.1.1
3 physical locations (or resources) are available for accessing 1.1.1.1 (from Enzyme Nomenclature):
KEGG Ligand Database for Enzyme Nomenclature
Kyoto University Bioinformatics Center
Japan
(Uptime: 100%)
Enzyme nomenclature database, ExPASy (Expert Protein
Access to '1.1.1.1' via this resource (MIR:00100003)
Switzerland
(Uptime: 99%)
IntEnZ (Integrated relational Enzyme database)
European Bioinformatics Institute
United Kingdom
(Uptime: 100%)
http://enzyme.expasy.org/EC/1.1.1.1

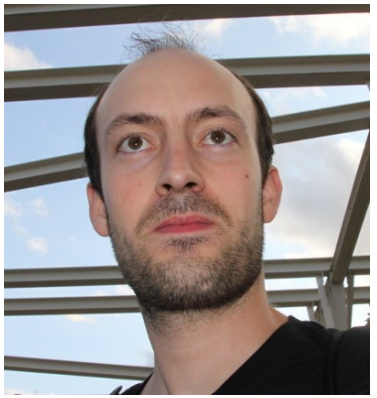
Systems Biology Ontology

- Download
- Recent changes
- Term request
- Edit tree
- Web Services
- FAQ
- News
- Project on SourceForge
- Contact

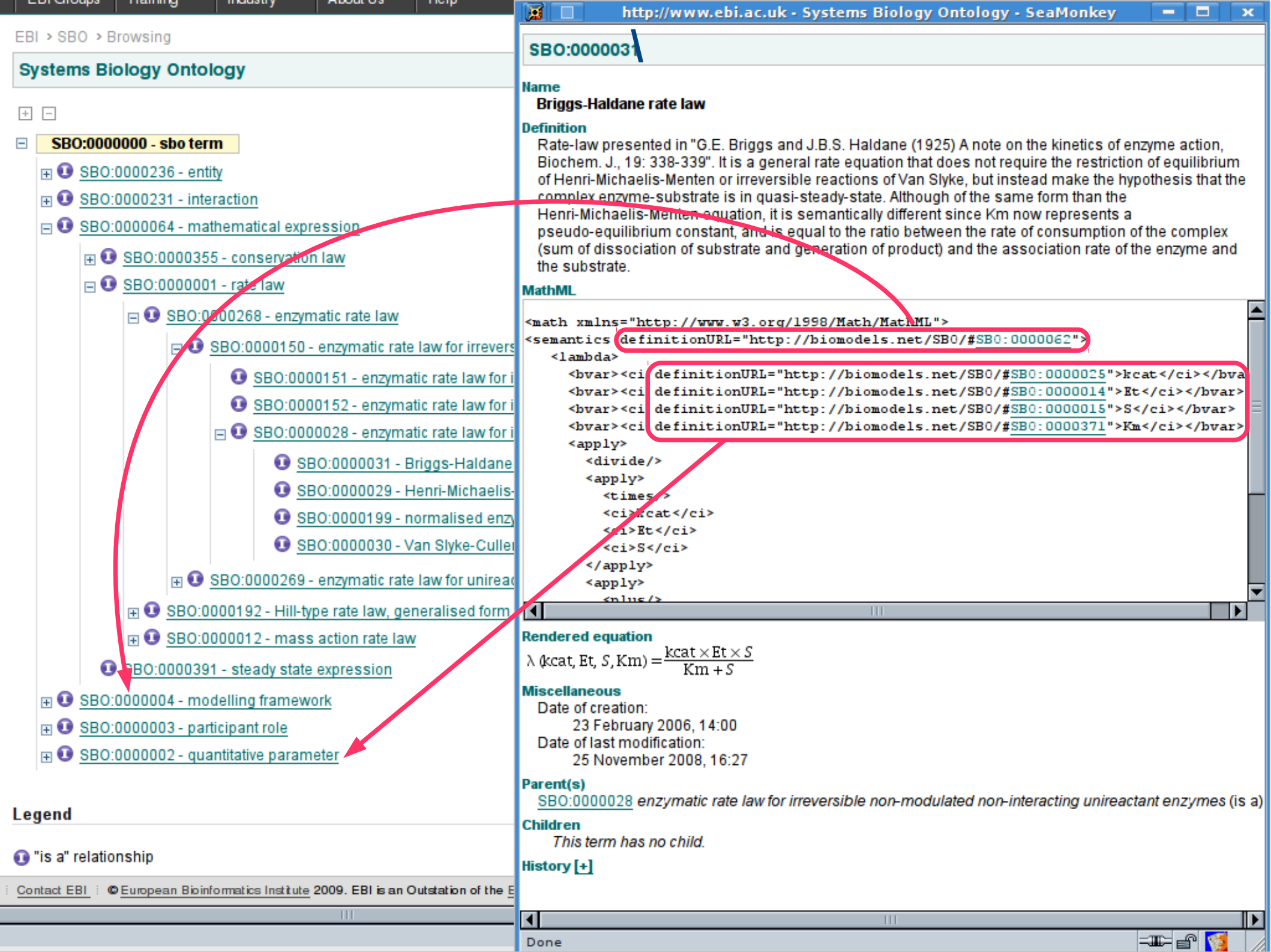
BIOMODELS.NET

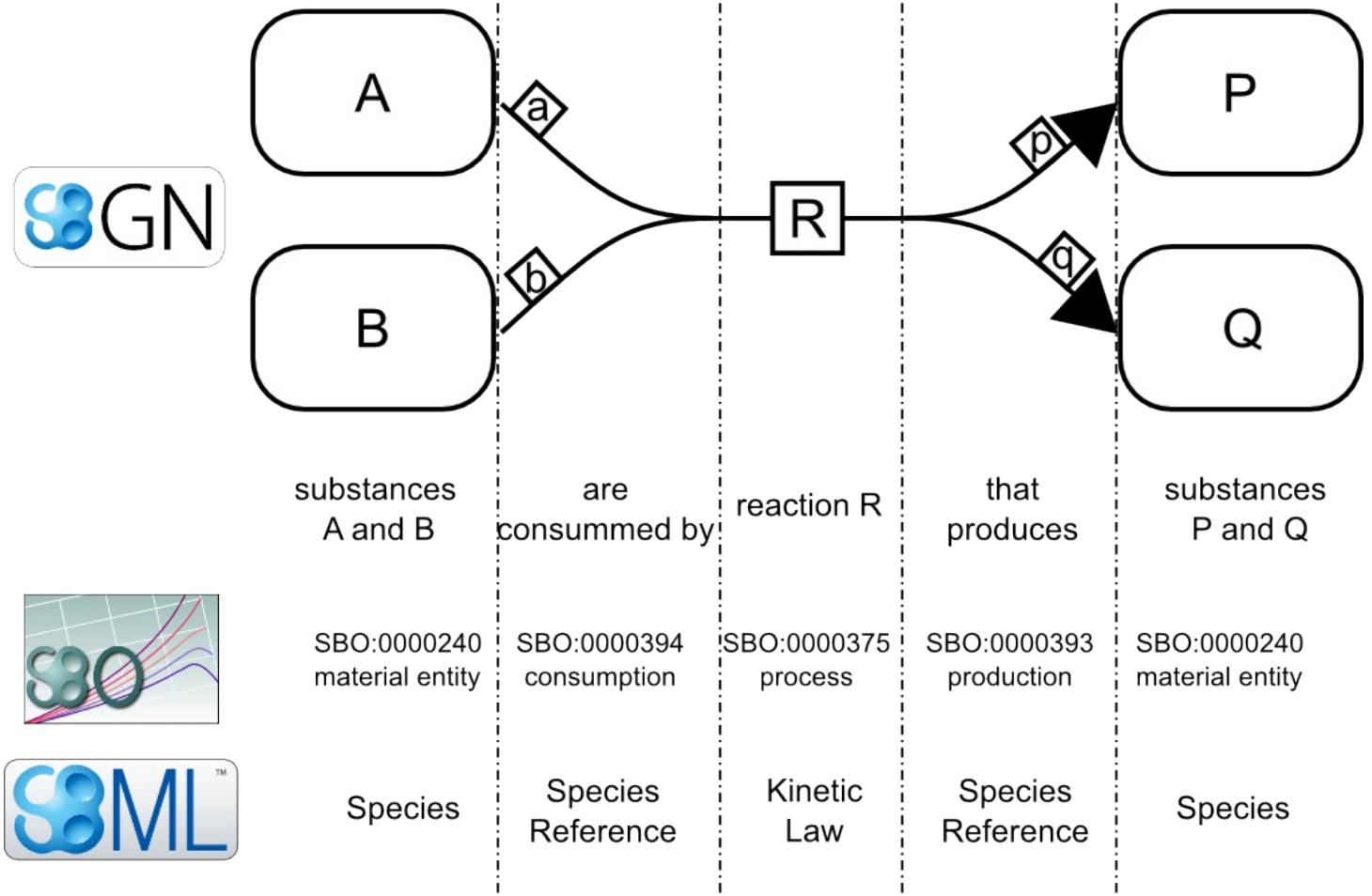
- SBO:0000000 - systems biology representation
  - SBO:0000064 - mathematical expression
    - SBO:0000355 - conservation law
    - SBO:0000474 - convenience function
    - SBO:0000001 - rate law
    - SBO:0000391 - steady state expression
  - SBO:0000544 - metadata representation
  - SBO:0000004 - modelling framework
  - SBO:0000231 - occurring entity representation
  - SBO:0000003 - participant role
  - SBO:0000236 - physical entity representation
  - SBO:0000545 - systems description parameter
    - SBO:0000546 - qualitative systems description parameter
    - SBO:0000002 - quantitative systems description parameter
      - SBO:0000492 - amplitude
      - SBO:0000542 - basic reproductive ratio
      - SBO:0000380 - biochemical coefficient
      - SBO:0000258 - capacitance
      - SBO:0000257 - conductance
      - SBO:0000254 - electrical resistance
      - SBO:0000308 - equilibrium or steady-state characteristic

Mélanie Courtot



http://www.ebi.ac.uk/sbo/


EBI > SBO > Browsing
Systems Biology Ontology
SBO:0000000 - sbo term
SBO:0000236 - entity
SBO:0000231 - interaction
SBO:0000064 - mathematical expression
SBO:0000355 - conservation law
SBO:0000001 - rate law
SBO:0000268 - enzymatic rate law
SBO:0000150 - enzymatic rate law for irrevers
SBO:0000151 - enzymatic rate law for i
SBO:0000152 - enzymatic rate law for i
SBO:0000028 - enzymatic rate law for i
SBO:0000031 - Briggs-Haldane
SBO:0000029 - Henri-Michaelis-
SBO:0000199 - normalised enzy
SBO:0000030 - Van Slyke-Culle
SBO:0000269 - enzymatic rate law for unireac
SBO:0000192 - Hill-type rate law, generalised form
SBO:0000012 - mass action rate law
SBO:0000391 - steady state expression
SBO:0000004 - modelling framework
SBO:0000003 - participant role
SBO:0000002 - quantitative parameter
Legend
"is a" relationship
Contact EBI
© European Bioinformatics Institute 2009. EBI is an Outstation of the E
http://www.ebi.ac.uk - Systems Biology Ontology - SeaMonkey
SBO:0000031
Name
Briggs-Haldane rate law
Definition
Rate-law presented in "G.E. Briggs and J.B.S. Haldane (1925) A note on the kinetics of enzyme action, Biochem. J., 19: 338-339". It is a general rate equation that does not require the restriction of equilibrium of Henri-Michaelis-Menten or irreversible reactions of Van Slyke, but instead make the hypothesis that the complex enzyme-substrate is in quasi-steady-state. Although of the same form than the Henri-Michaelis-Menten equation, it is semantically different since Km now represents a pseudo-equilibrium constant, and is equal to the ratio between the rate of consumption of the complex (sum of dissociation of substrate and generation of product) and the association rate of the enzyme and the substrate.
MathML
<math xmlns="http://www.w3.org/1998/Math/MathML">
<semantics definitionURL="http://biomodels.net/SBO/#SBO:0000062">
<lambda>
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000025">kcat</ci></bva
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000014">Et</ci></bvar>
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000015">S</ci></bvar>
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000371">Km</ci></bvar>
<apply>
<divide/>
<apply>
<times/>
<ci>kcat</ci>
<ci>Et</ci>
<ci>S</ci>
</apply>
<apply>
Rendered equation
$\lambda(\mathrm{kcat}, \mathrm{Et}, S, \mathrm{Km}) = \frac{\mathrm{kcat} \times \mathrm{Et} \times S}{\mathrm{Km} + S}$
Miscellaneous
Date of creation:
23 February 2006, 14:00
Date of last modification:
25 November 2008, 16:27
Parent(s)
SBO:0000028 enzymatic rate law for irreversible non-modulated non-interacting unireactant enzymes (is a)
Children
This term has no child.
History [+]
Done


A
B
a
b
R
p
q
P
Q
substances A and B
are consummed by
reaction R
that produces
substances P and Q
SBO:0000240 material entity
SBO:0000394 consumption
SBO:0000375 process
SBO:0000393 production
SBO:0000240 material entity
Species
Species Reference
Kinetic Law
Species Reference
Species

SBML and SBO

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" /
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

SBML and SBO

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" />
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

functional compartment

simple chemical

simple chemical

enzyme

catalysis

substrate

product

catalyst

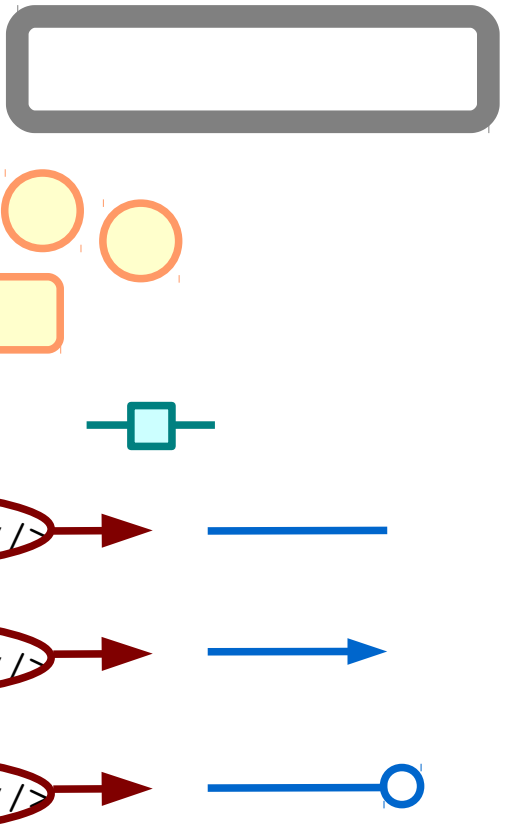
Briggs-Haldane equation

Km

kcat

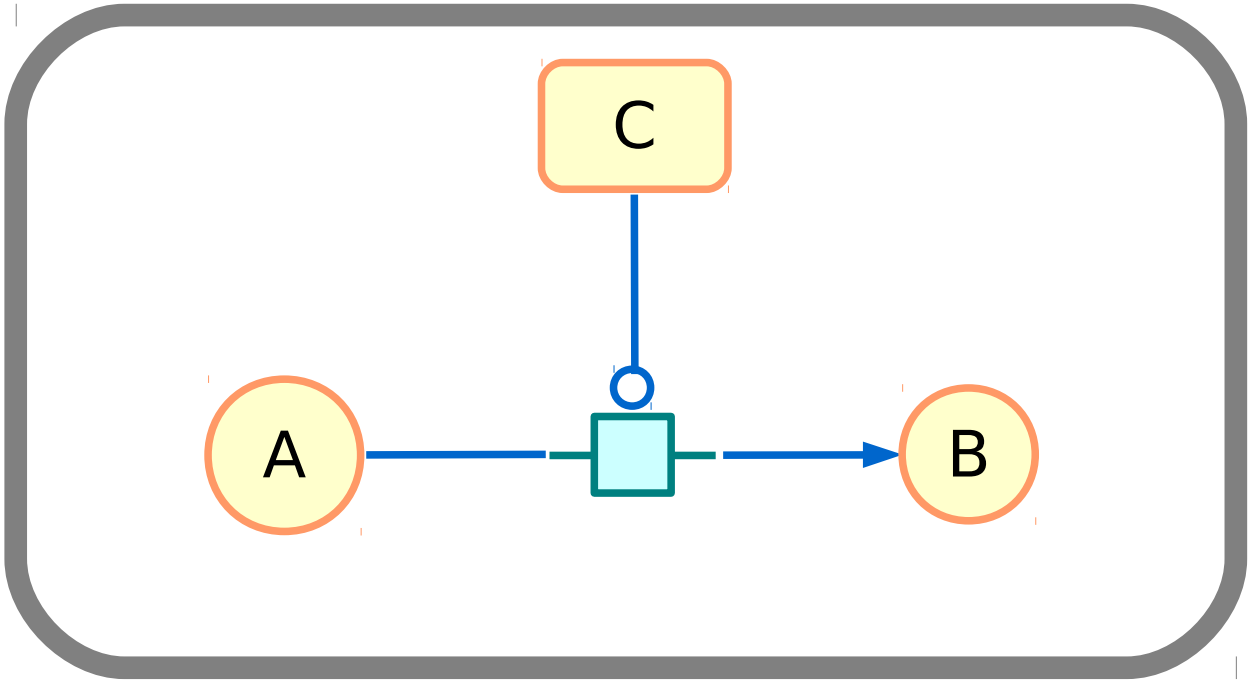
SBML to SBGN conversion using SBO

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" />
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```



http://www.sbgn.org/

SBML to SBGN conversion using SBO



Conversion between modeling frameworks

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" /
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

discrete simulator

$$v1 = \frac{(k_{-1} + V)}{U} \cdot [A] \cdot [C]$$

$$v2 = k_{-1} \cdot [D]$$

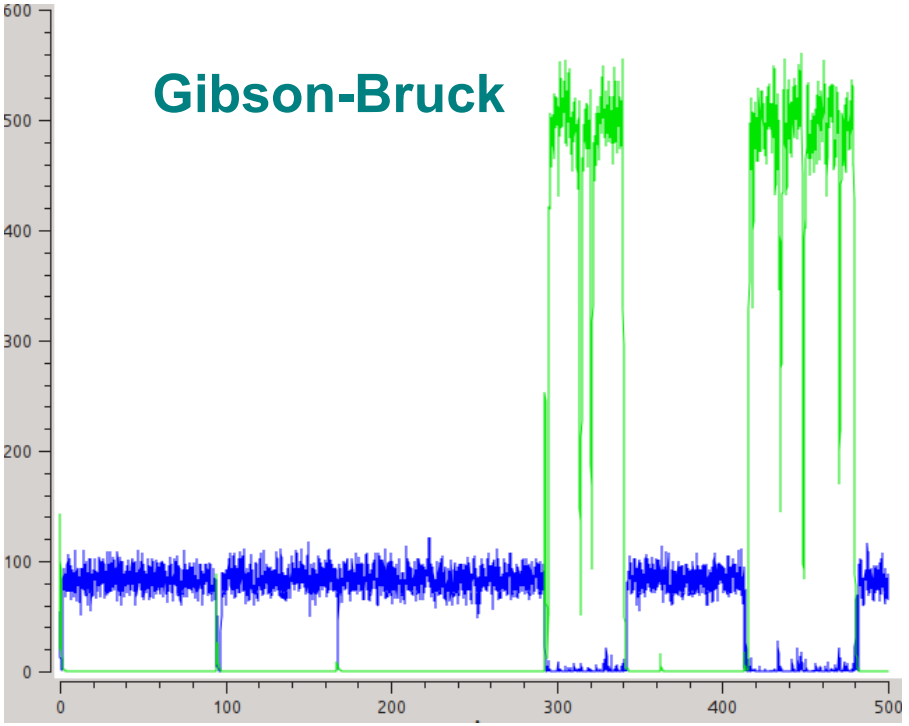
$$v3 = V \cdot [D]$$

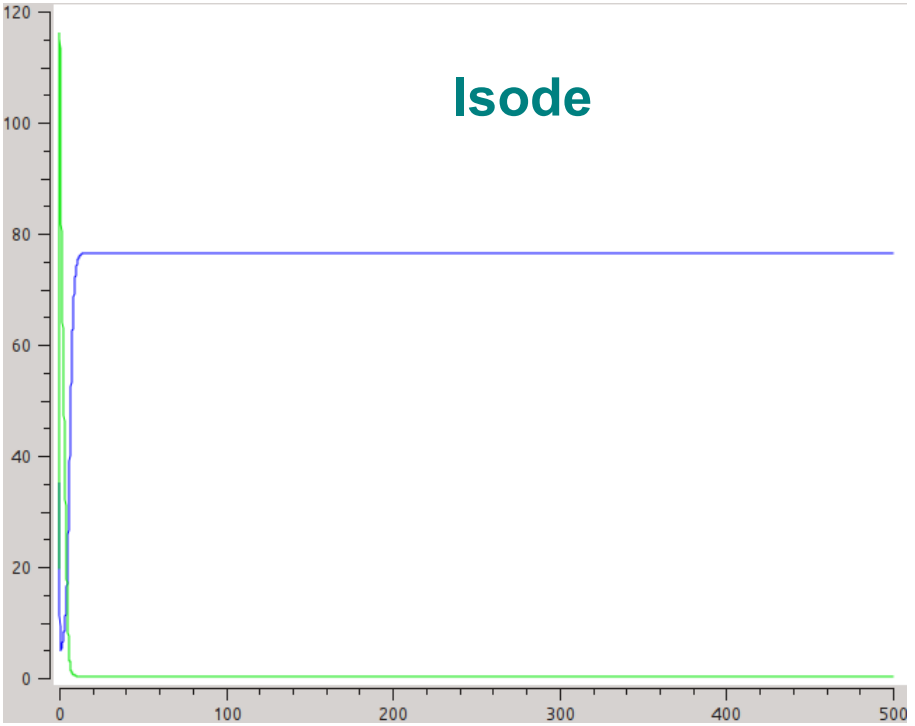
continuous simulator

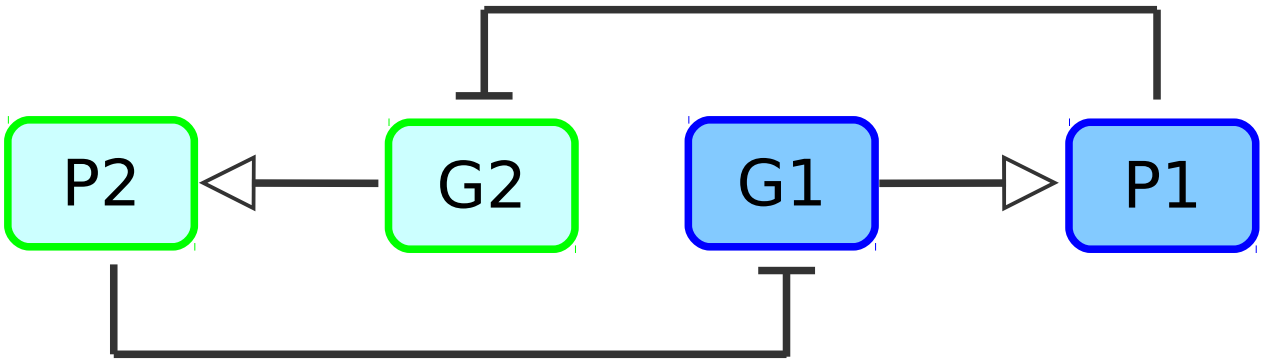
$$\nu = \frac{C \cdot V \cdot [A]}{(U + [A])}$$

Now this is sufficient, surely?

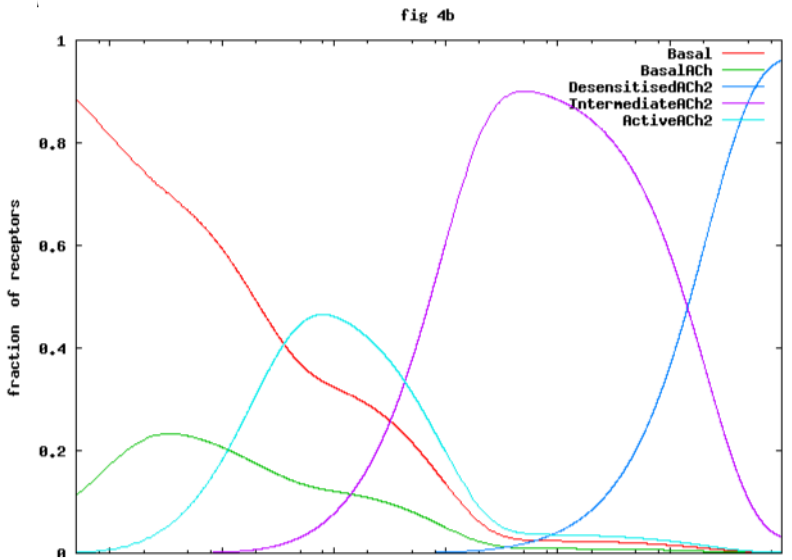
Algorithm choice and multistability



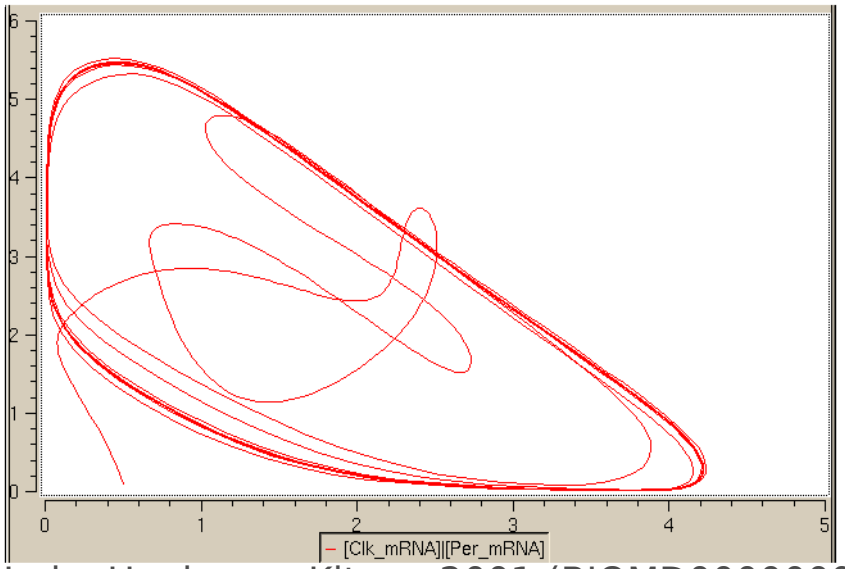




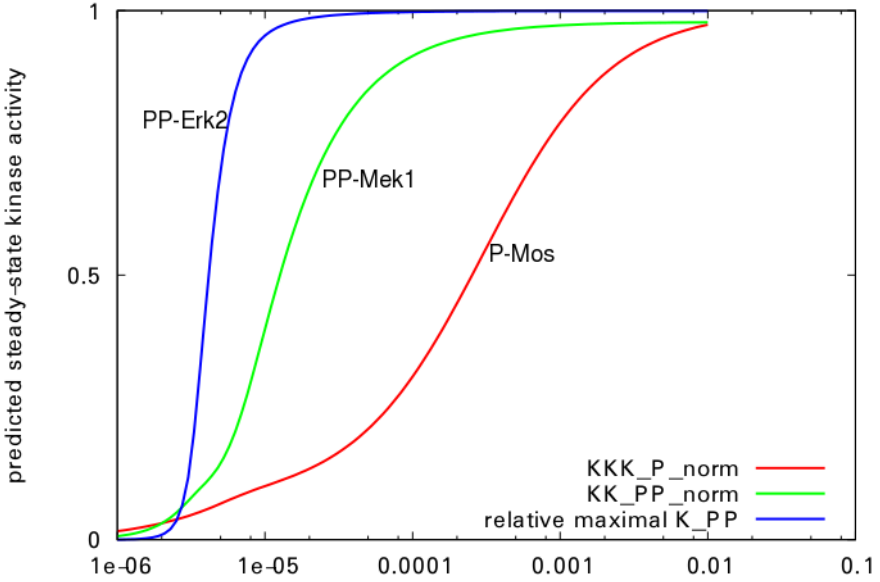
Reproduction of published simulation results



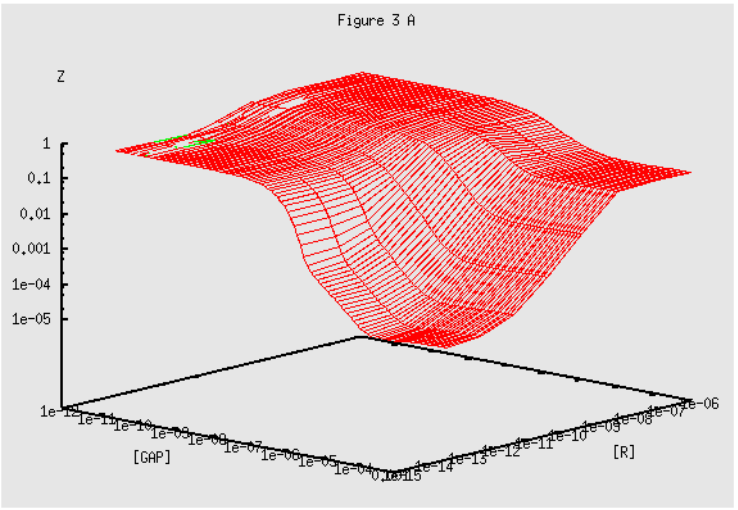
Edelstein et al 1996 (BIOMD0000000002)



Ueda, Hagiwara, Kitano 2001 (BIOMD0000000022)



Huang & Ferrell (BIOMD0000000009)



Bornheimer et al 2004 (BIOMD0000000086)

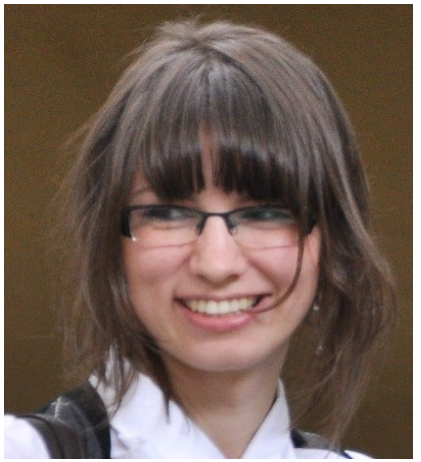
Babraham Institute

EMBL-EBI

Description of simulations and analyses

	Model descriptions	Simulations and analysis
Minimal requirements	MIRIAM	MIASE
Data-models	SBML, SBGN, BioPAX	SED ML
Terminologies	SBO	KISAO

Dagmar Waltemath



Anna Zhukova

Born in Hinxton 2007

Minimal Information About a Simulation Experiment

Models to use

1. Models provided, or precise mean of access
2. Model with all governing equations, parameter values and necessary conditions
3. Standard formats, otherwise code available. If not open code, full description provided
4. Description of modifications required before the execution of the simulation experiment

http://co.mbine.org/standards/miase

Minimal Information About a Simulation Experiment

Models to use

1. Models provided, or precise mean of access
2. Model with all governing equations, parameter values and necessary conditions
3. Standard formats, otherwise code available. If not open code, full description provided
4. Description of modifications required before the execution of the simulation experiment

Simulation steps

1. Simulation steps described, with simulation algorithms, models, order of steps, data processing between steps
2. Information needed for correct implementation of necessary steps
3. If software source-code not available, information needed to reproduce the simulation, and not only repeat it, with algorithms and necessary info (e.g. discretization meth)
4. If divergence are known in different environments or platforms, explanation on how to be run with the specified environment/platform to achieve experiment's purpose

http://co.mbine.org/standards/miase

Minimal Information About a Simulation Experiment

Models to use

1. Models provided, or precise mean of access
2. Model with all governing equations, parameter values and necessary conditions
3. Standard formats, otherwise code available. If not open code, full description provided
4. Description of modifications required before the execution of the simulation experiment

Simulation steps

1. Simulation steps described, with simulation algorithms, models, order of steps, data processing between steps
2. Information needed for correct implementation of necessary steps
3. If software source-code not available, information needed to reproduce the simulation, and not only repeat it, with algorithms and necessary info (e.g. discretization meth)
4. If divergence are known in different environments or platforms, explanation on how to be run with the specified environment/platform to achieve experiment's purpose

Output specification

1. Post-processing steps applied on the raw results to generate the final results, with identification of data to process, nature and order of changes to apply
2. If insights depend on relation between different results, (e.g. plot of one against another), the results to be compared must be specified.

http://co.mbine.org/standards/miase

Simulation Experiment Description Markup Language

```
<?xml version="1.0" encoding="utf-8"?>
<sedML xmlns="http://sed-ml.org/"
       xmlns:math="http://www.w3.org/1998/Math/MathML"
       level="1" version="1">
  <listOfSimulations><!-- --> </listOfSimulations>
  <listOfModels>
    <model id="" source="">
      <listOfChanges><!-- --></listOfChanges>
    </model>
  </listOfModels>
  <listOfTasks><!-- --></listOfTasks>
  <listOfDataGenerators><!-- --></listOfDataGenerators>
  <listOfOutputs>
    <plot2D />
    <plot3D />
    <report />
  </listOfOutputs>
</sedML>
```

http://sed-ml.org

Flexible model use in SED-ML

Any XML

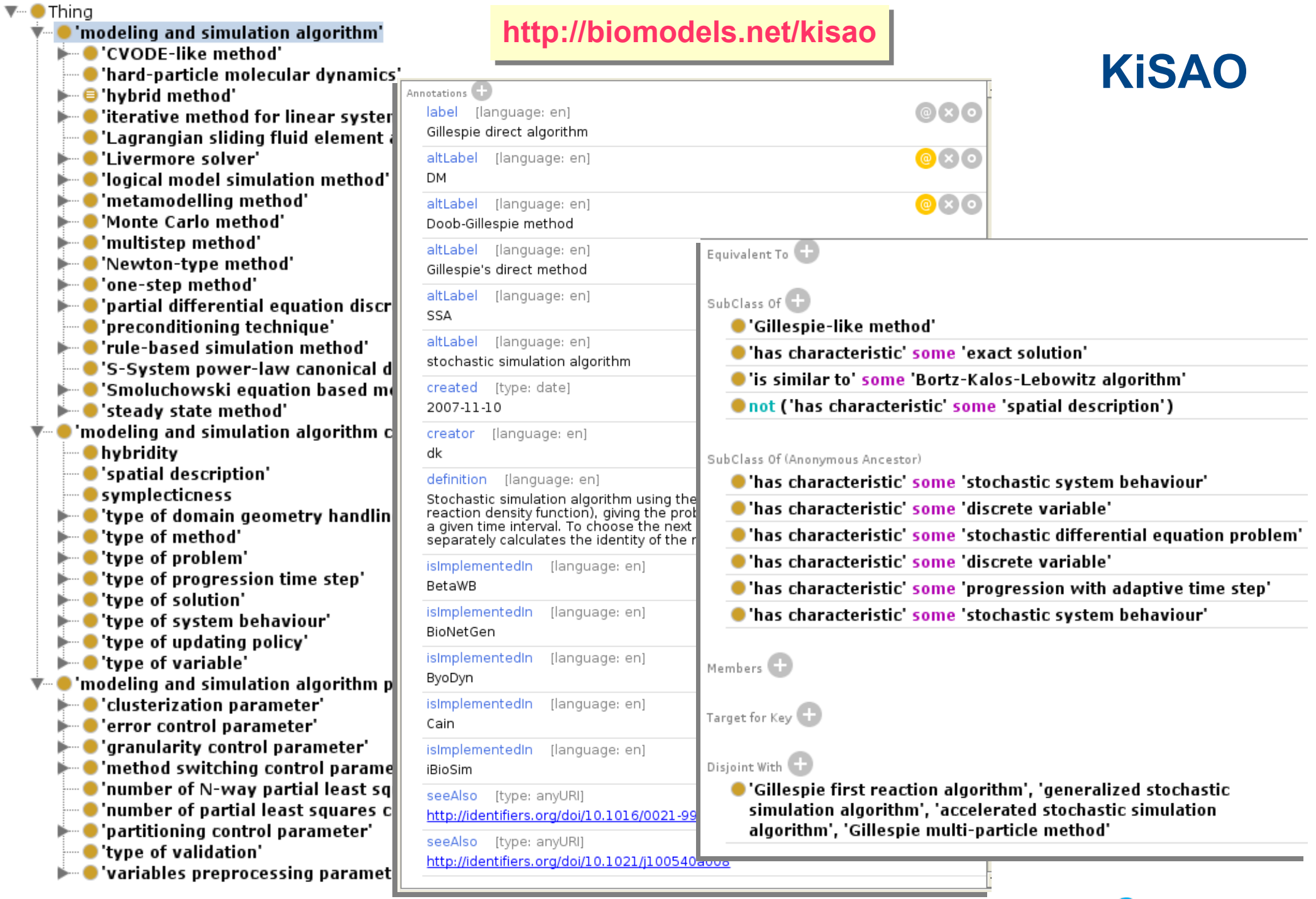
```
<listOfModels>
  <model id="model1"
         name="Regular_Spiking"
         language="http://identifiers.org/combine.specifications/sbml.level-2.version-4.release-1"
         source="http://identifiers.org/biomodels.db/BIOMD0000000127" />
  <model id="model2"
         name="chattering"
         source="model1">
    <listOfChanges>
      <changeAttribute target=
           "/sbml/model/listOfParameters/parameter[@id='c']/@value" newValue="-50">
      </changeAttribute>
      <changeAttribute target=
           "/sbml/model/listOfParameters/parameter[@id='d']/@value" newValue="42">
      </changeAttribute>
    </listOfChanges>
  </model>
</listOfModels>
```

Remote access

Modif before simulations

KiSAO

http://biomodels.net/kisao

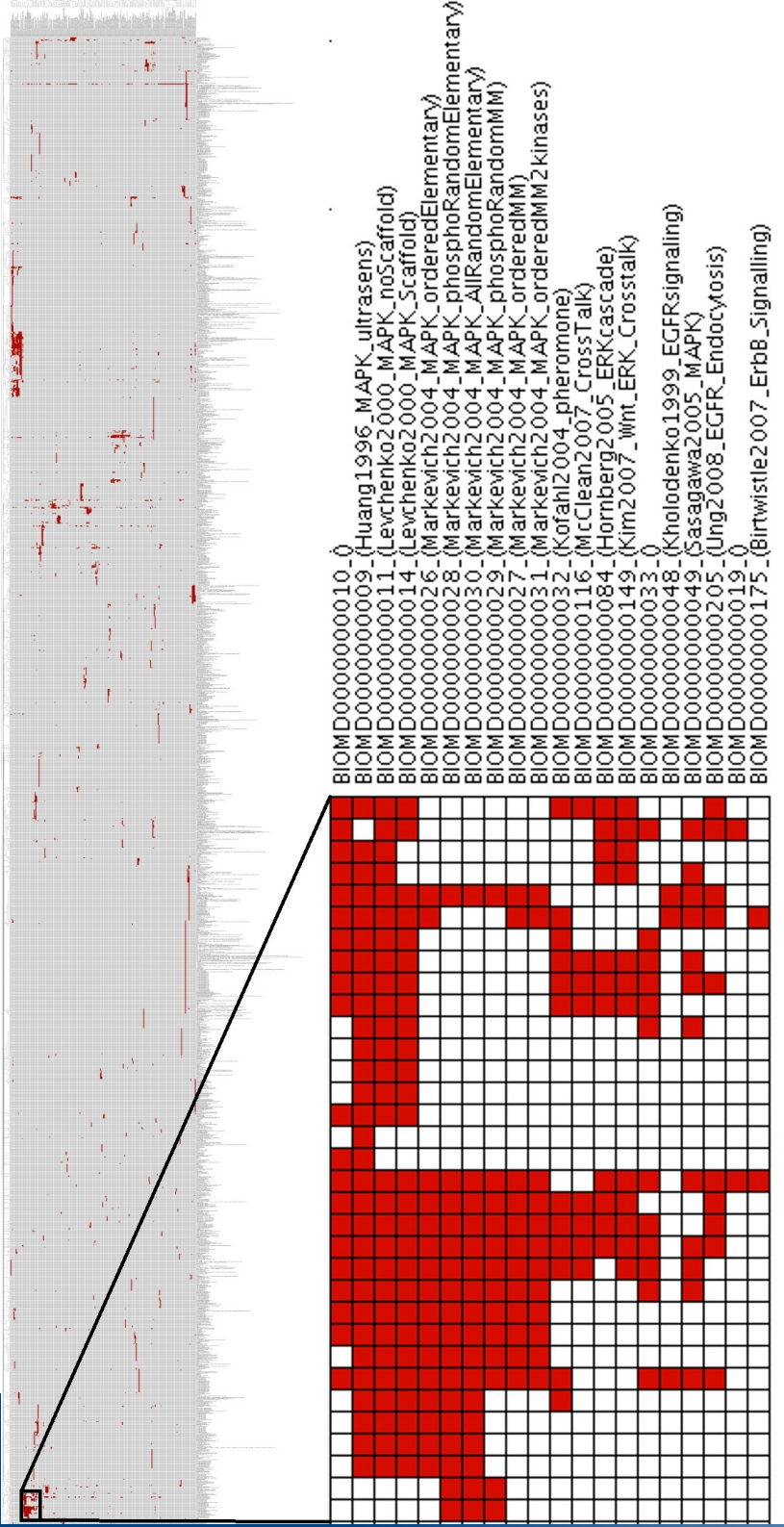


Characterising dynamical behaviours

	Model descriptions	Simulations and analysis	results
Minimal requirements	MIRIAM	MIASE	 Christian Knüpfer
Data-models	SBML, SBGN	SED ML	NuML
Terminologies	SBO	KISAO	TEDDY

Was it worth it?

3 Clustering models (and data) based on metadata



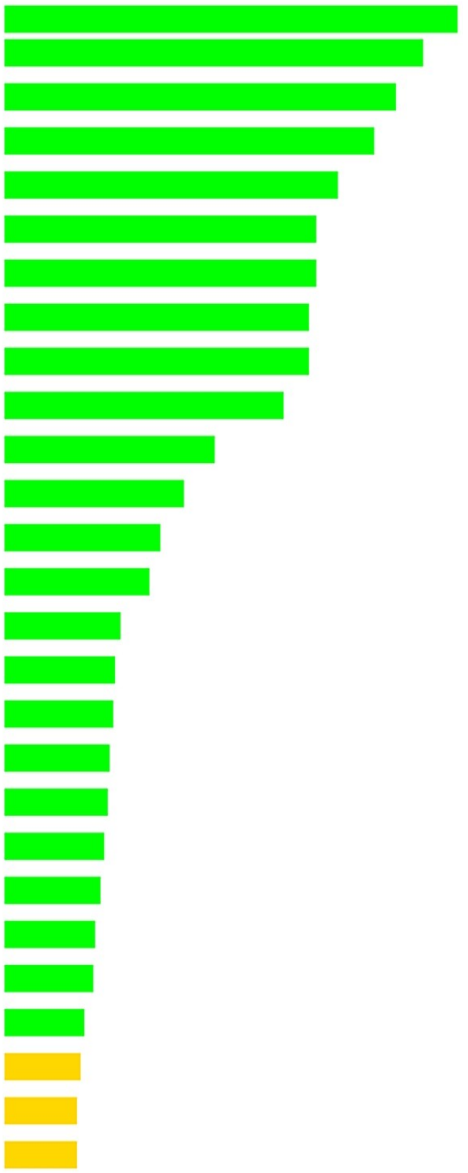
Schulz M., Krause F., Le Novère N., Klipp E., Liebermeister W.
Retrieval, alignment, and clustering of computational models based on semantic annotations.
Molecular Systems Biology (2011), 7: 512

Babraham Institute

EMBL-EBI

Ranking and retrieval of models

Model	BioModel	Score	p-val num
Huang1996_MAPK_ultrasens	BIOMD0000000009	1.000	< 1/1000
Levchenko2000_MAPK_noScaffold	BIOMD0000000011	0.925	< 1/1000
Levchenko2000_MAPK_Scaffold	BIOMD0000000014	0.865	< 1/1000
Kholodenko2000_MAPK_feedback	BIOMD0000000010	0.816	< 1/1000
Markevich2004_MAPK_orderedElementary	BIOMD0000000026	0.737	< 1/1000
Markevich2004_MAPK_phosphoRandomElementary	BIOMD0000000028	0.690	< 1/1000
Markevich2004_MAPK_AllRandomElementary	BIOMD0000000030	0.690	< 1/1000
Markevich2004_MAPK_orderedMM	BIOMD0000000027	0.673	< 1/1000
Markevich2004_MAPK_orderedMM2kinases	BIOMD0000000031	0.673	< 1/1000
Markevich2004_MAPK_phosphoRandomMM	BIOMD0000000029	0.617	< 1/1000
Hornberg2005_ERKcascade	BIOMD0000000084	0.468	< 1/1000
McClean2007_CrossTalk	BIOMD0000000116	0.397	< 1/1000
Kofahl2004_pheromone	BIOMD0000000032	0.348	< 1/1000
Kim2007_Wnt_ERK_Crosstalk	BIOMD0000000149	0.323	< 1/1000
Brown2004_NGF_EGF_signaling	BIOMD0000000033	0.260	< 1/1000
Goldbeter1995_CircClock	BIOMD0000000016	0.244	< 1/1000
Sasagawa2005_MAPK	BIOMD0000000049	0.240	< 1/1000
Ung2008_EGFR_Endocytosis	BIOMD0000000205	0.234	< 1/1000
Leloup1999_CircClock	BIOMD0000000021	0.229	< 1/1000
Goldbeter1991_MinMitOscil_ExplInact	BIOMD0000000004	0.222	< 1/1000
Goldbeter1991_MinMitOscil	BIOMD0000000003	0.214	< 1/1000
Leloup1998_CircClock_LD	BIOMD0000000171	0.201	< 1/1000
Tyson1991_CellCycle_6var	BIOMD0000000005	0.199	< 1/1000
Veening2008_DegU_Regulation	BIOMD0000000240	0.180	< 1/1000
Neves2008_Cell_Shape	BIOMD0000000182	0.168	4.0e-03
Fisher2006_Ca_Oscillation_dpdnt_NFAT_dynamics	BIOMD0000000122	0.164	5.0e-03
Leloup2003_CircClock_DD	BIOMD0000000073	0.163	5.0e-03

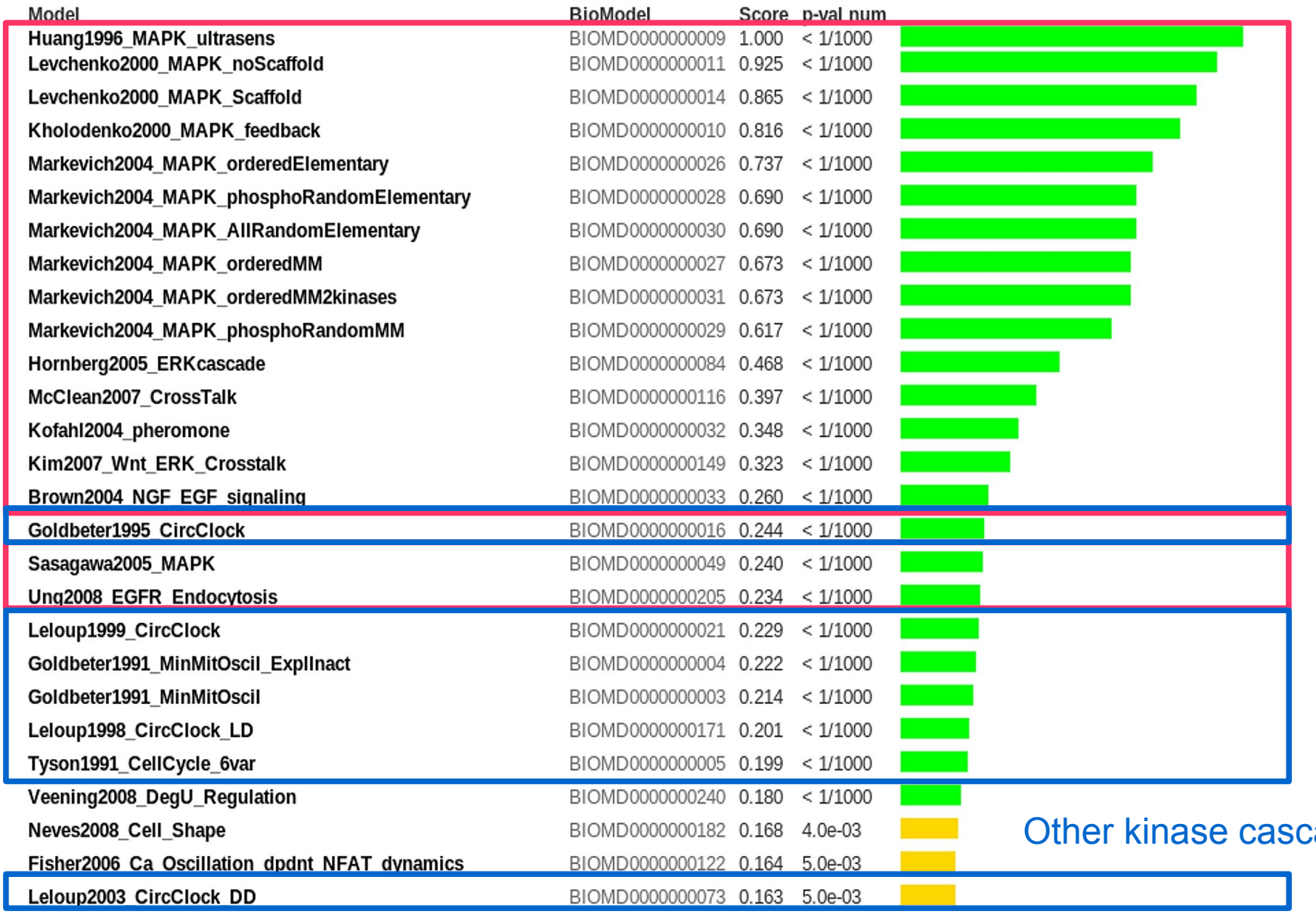


Ranking and retrieval of models

Model	BioModel	Score	p-val num
Huang1996_MAPK_ultrasens	BIOMD0000000009	1.000	< 1/1000
Levchenko2000_MAPK_noScaffold	BIOMD0000000011	0.925	< 1/1000
Levchenko2000_MAPK_Scaffold	BIOMD0000000014	0.865	< 1/1000
Kholodenko2000_MAPK_feedback	BIOMD0000000010	0.816	< 1/1000
Markevich2004_MAPK_orderedElementary	BIOMD0000000026	0.737	< 1/1000
Markevich2004_MAPK_phosphoRandomElementary	BIOMD0000000028	0.690	< 1/1000
Markevich2004_MAPK_AllRandomElementary	BIOMD0000000030	0.690	< 1/1000
Markevich2004_MAPK_orderedMM	BIOMD0000000027	0.673	< 1/1000
Markevich2004_MAPK_orderedMM2kinases	BIOMD0000000031	0.673	< 1/1000
Markevich2004_MAPK_phosphoRandomMM	BIOMD0000000029	0.617	< 1/1000
Hornberg2005_ERKcascade	BIOMD0000000084	0.468	< 1/1000
McClean2007_CrossTalk	BIOMD0000000116	0.397	< 1/1000
Kofahl2004_pheromone	BIOMD0000000032	0.348	< 1/1000
Kim2007_Wnt_ERK_Crosstalk	BIOMD0000000149	0.323	< 1/1000
Brown2004_NGF_EGF_signaling	BIOMD0000000033	0.260	< 1/1000
Goldbeter1995_CircClock	BIOMD0000000016	0.244	< 1/1000
Sasagawa2005_MAPK	BIOMD0000000049	0.240	< 1/1000
Ung2008_EGFR_Endocytosis	BIOMD0000000205	0.234	< 1/1000
Leloup1999_CircClock	BIOMD0000000021	0.229	< 1/1000
Goldbeter1991_MinMitOscil_ExplInact	BIOMD0000000004	0.222	< 1/1000
Goldbeter1991_MinMitOscil	BIOMD0000000003	0.214	< 1/1000
Leloup1998_CircClock_LD	BIOMD0000000171	0.201	< 1/1000
Tyson1991_CellCycle_6var	BIOMD0000000005	0.199	< 1/1000
Veening2008_DegU_Regulation	BIOMD0000000240	0.180	< 1/1000
Neves2008_Cell_Shape	BIOMD0000000182	0.168	4.0e-03
Fisher2006_Ca_Oscillation_dpdnt_NFAT_dynamics	BIOMD0000000122	0.164	5.0e-03
Leloup2003_CircClock_DD	BIOMD0000000073	0.163	5.0e-03

MAPK

Ranking and retrieval of models

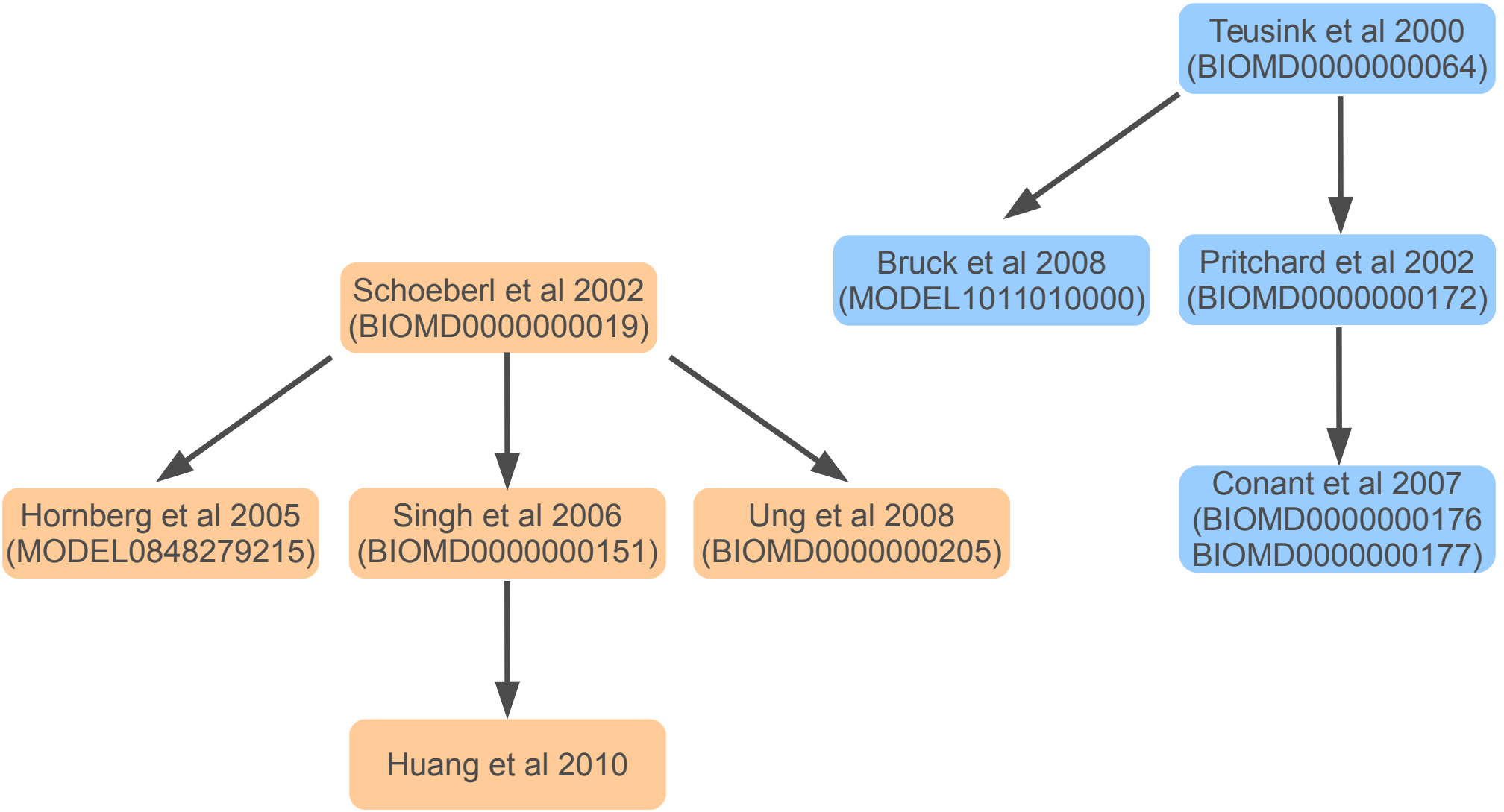


Model	BioModel	Score	p-val num
Huang1996_MAPK_ultrasens	BIOMD0000000009	1.000	< 1/1000
Levchenko2000_MAPK_noScaffold	BIOMD0000000011	0.925	< 1/1000
Levchenko2000_MAPK_Scaffold	BIOMD0000000014	0.865	< 1/1000
Kholodenko2000_MAPK_feedback	BIOMD0000000010	0.816	< 1/1000
Markevich2004_MAPK_orderedElementary	BIOMD0000000026	0.737	< 1/1000
Markevich2004_MAPK_phosphoRandomElementary	BIOMD0000000028	0.690	< 1/1000
Markevich2004_MAPK_AllRandomElementary	BIOMD0000000030	0.690	< 1/1000
Markevich2004_MAPK_orderedMM	BIOMD0000000027	0.673	< 1/1000
Markevich2004_MAPK_orderedMM2kinases	BIOMD0000000031	0.673	< 1/1000
Markevich2004_MAPK_phosphoRandomMM	BIOMD0000000029	0.617	< 1/1000
Hornberg2005_ERKcascade	BIOMD0000000084	0.468	< 1/1000
McClean2007_CrossTalk	BIOMD0000000116	0.397	< 1/1000
Kofahl2004_pheromone	BIOMD0000000032	0.348	< 1/1000
Kim2007_Wnt_ERK_Crosstalk	BIOMD0000000149	0.323	< 1/1000
Brown2004_NGF_EGF_signaling	BIOMD0000000033	0.260	< 1/1000
Goldbeter1995_CircClock	BIOMD0000000016	0.244	< 1/1000
Sasagawa2005_MAPK	BIOMD0000000049	0.240	< 1/1000
Ung2008_EGFR_Endocytosis	BIOMD0000000205	0.234	< 1/1000
Leloup1999_CircClock	BIOMD0000000021	0.229	< 1/1000
Goldbeter1991_MinMitOscil_ExplInact	BIOMD0000000004	0.222	< 1/1000
Goldbeter1991_MinMitOscil	BIOMD0000000003	0.214	< 1/1000
Leloup1998_CircClock_LD	BIOMD0000000171	0.201	< 1/1000
Tyson1991_CellCycle_6var	BIOMD0000000005	0.199	< 1/1000
Veening2008_DegU_Regulation	BIOMD0000000240	0.180	< 1/1000
Neves2008_Cell_Shape	BIOMD0000000182	0.168	4.0e-03
Fisher2006_Ca_Oscillation_dpdnt_NFAT_dynamics	BIOMD0000000122	0.164	5.0e-03
Leloup2003_CircClock_DD	BIOMD0000000073	0.163	5.0e-03

MAPK

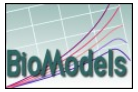
Other kinase cascades

1 Model re-use: e.g. EGFR signalling and glycolysis



2 Standard formats generate new research

- Herrgård et al (2008) A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nature Biotechnol*, 26: 1155-1160



MODEL0072364382: 2152 species, 1857 reactions

  - stoichiometric map, no concentrations, no kinetics

- Dobson et al (2010) Further developments towards a genome-scale metabolic model of yeast. *BMC Syst Biol,* 4:145

MODEL1012110000: 2657 species, 1865 reactions

- Smallbone et al (2010) Towards a genome-scale kinetic model of cellular metabolism. *BMC Syst Biol*, 4:6

MODEL1001200000: 1748 species, 1059 reactions

  - Concentrations and flux from BioModels Database
  - Constraint-based model and simplified linlog kinetics

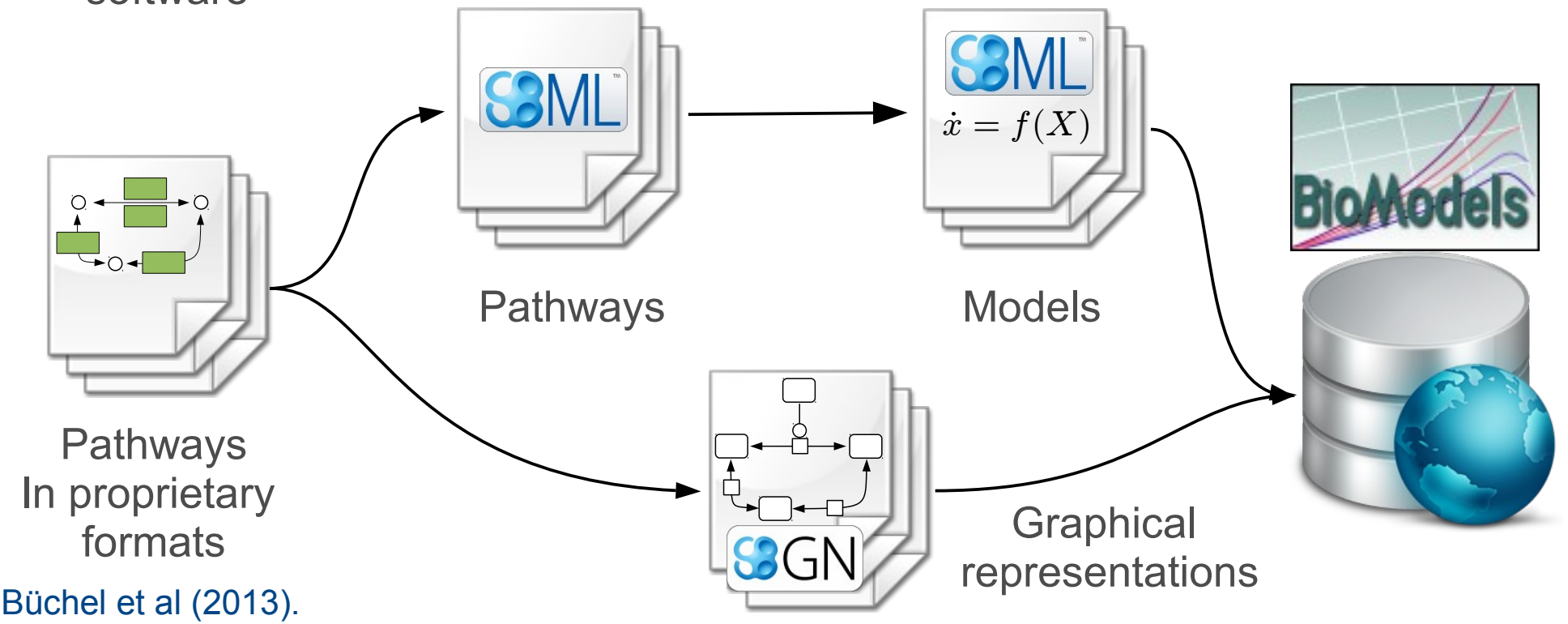
- Li et al (2010) Systematic integration of experimental data and models in systems biology. *BMC Bioinfo,* 11: 582

MODEL1012110001

  - Workflows using experimental kinetic information database (SABIO-RK) plus metabolomics and proteomics database
  - Full quantitative chemical kinetics descriptions

4 Path2Models project

- To re-use existing pathway data to generate biochemically based models
- To provide a starting point to model as many biochemical pathways as possible in as many species as possible
- To provide models in a standard format readable by most systems biology software



Büchel et al (2013).
Path2Models: large-scale generation of computational models from biochemical pathway maps.
BMC Systems Biology (2013) 7: 116

Three parallel workflows

Signalling Pathways

+

Logical models
of individual signalling pathways

Three parallel workflows

Signalling Pathways

+

Logical models
of individual signalling pathways

Metabolic Networks

Chemical kinetics models
of individual metabolic pathways

Three parallel workflows

Signalling Pathways

+

→ **Logical models**
of individual signalling pathways

Metabolic Networks

→ **Chemical kinetics models**
of individual metabolic pathways

Metabolic Networks

+

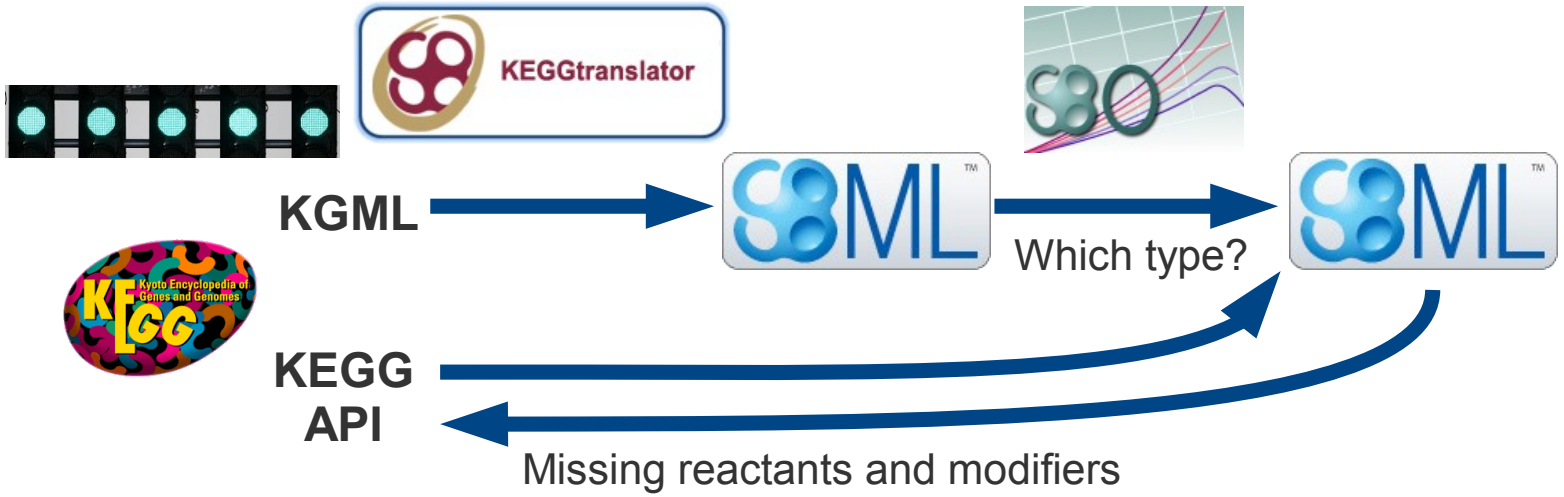
→ **Flux Balance Analysis**
of whole genome reconstructions

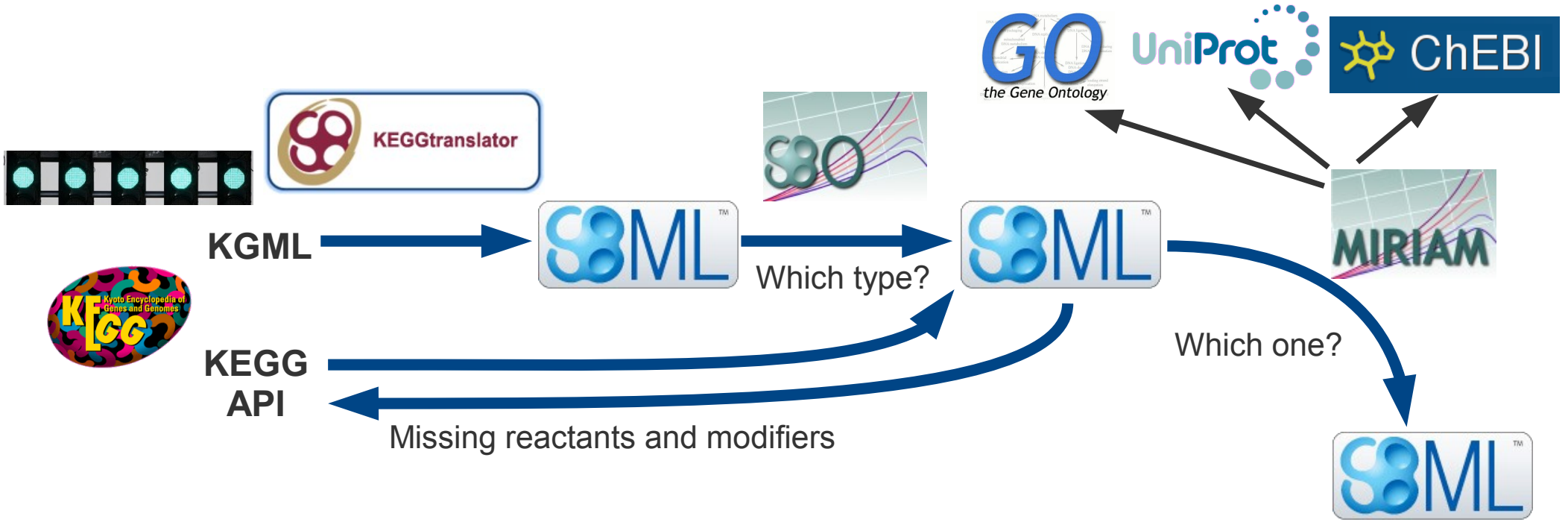
KEGGtranslator

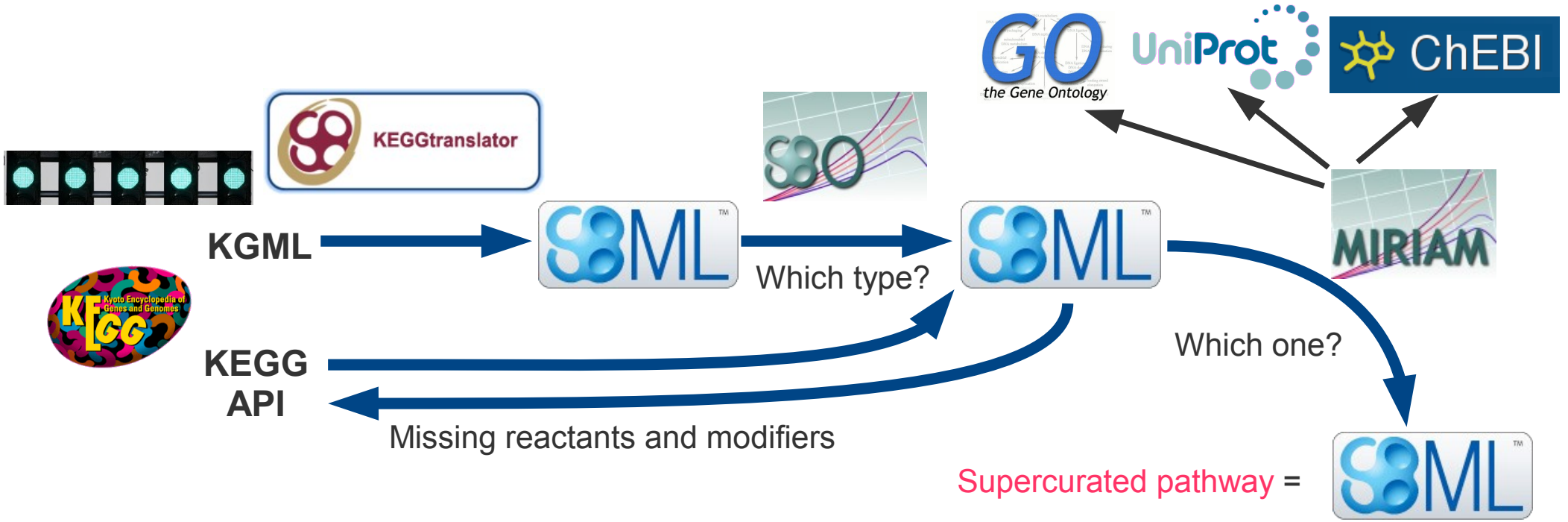
KGML → SBML

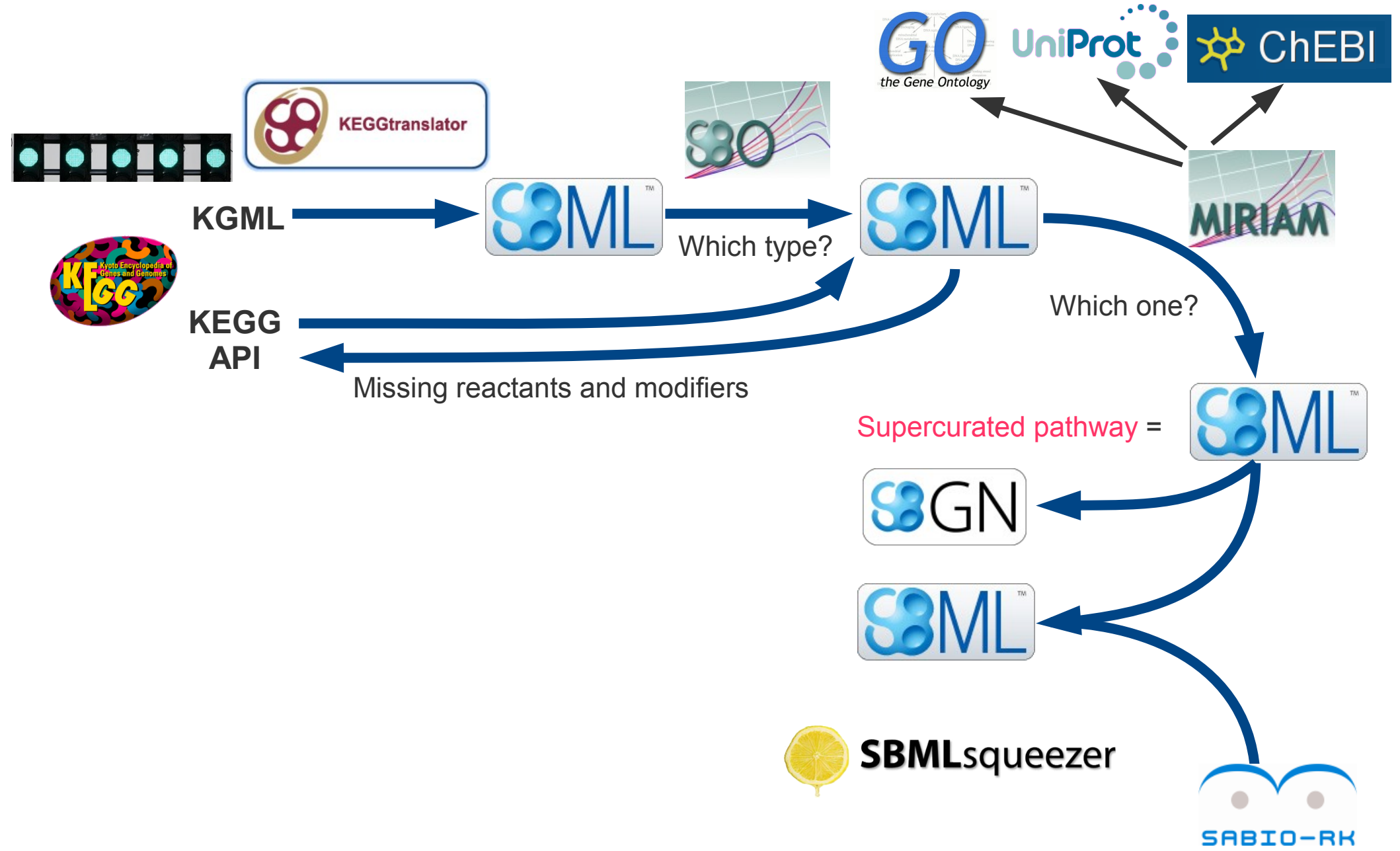
Core


KEGGtranslator
KGML
SBML
Which type?
SBML
KEGG
Kyoto Encyclopedia of Genes and Genomes


KEGGtranslator
KGML
SBML
Which type?
SBML
KEGG
API
Missing reactants and modifiers

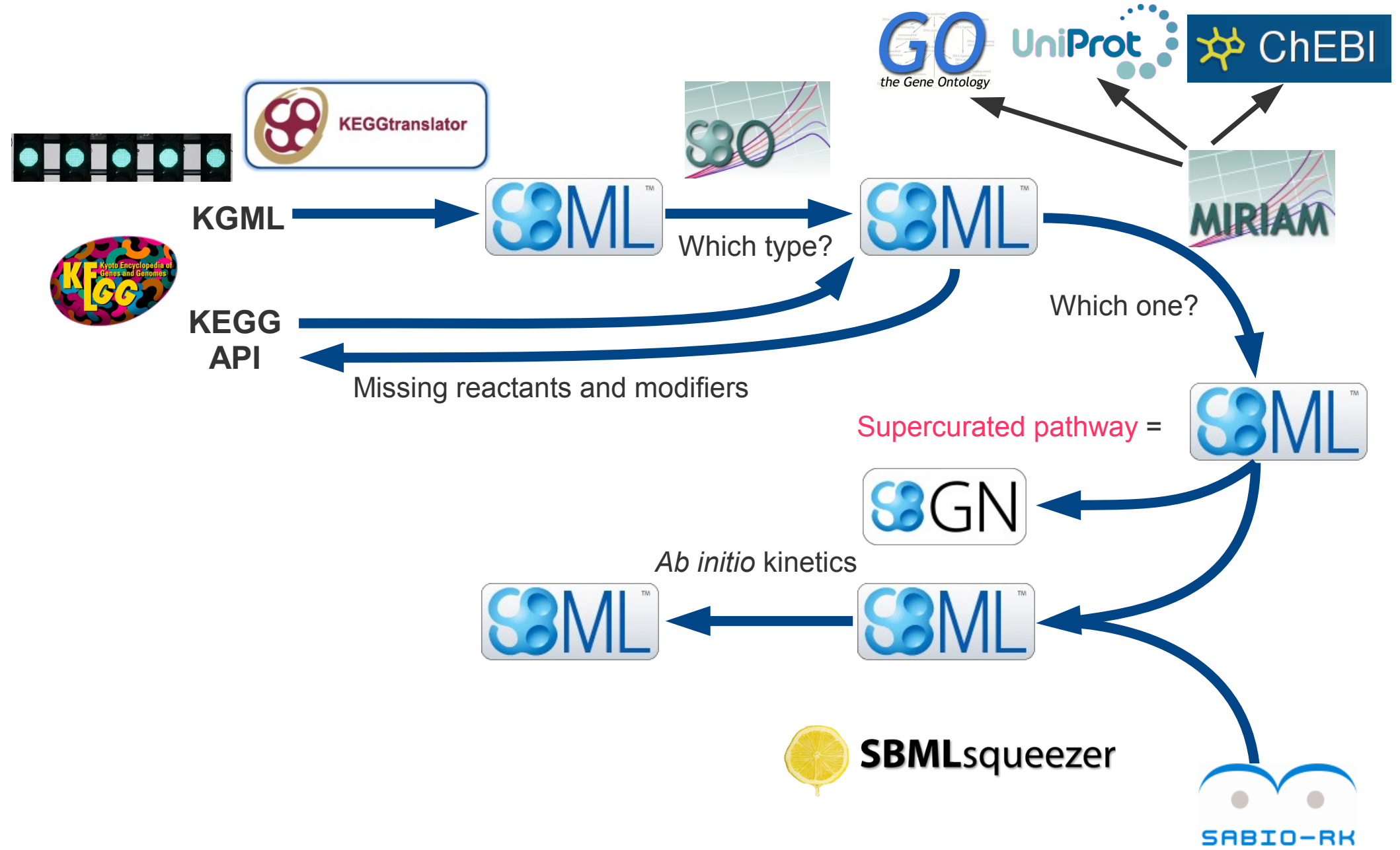

GO
the Gene Ontology
UniProt
ChEBI
KEGGtranslator
SBO
KGML
SBML
Which type?
SBML
MIRIAM
KEGG
Kyoto Encyclopedia of Genes and Genomes
KEGG API
Missing reactants and modifiers
Which one?
SBML

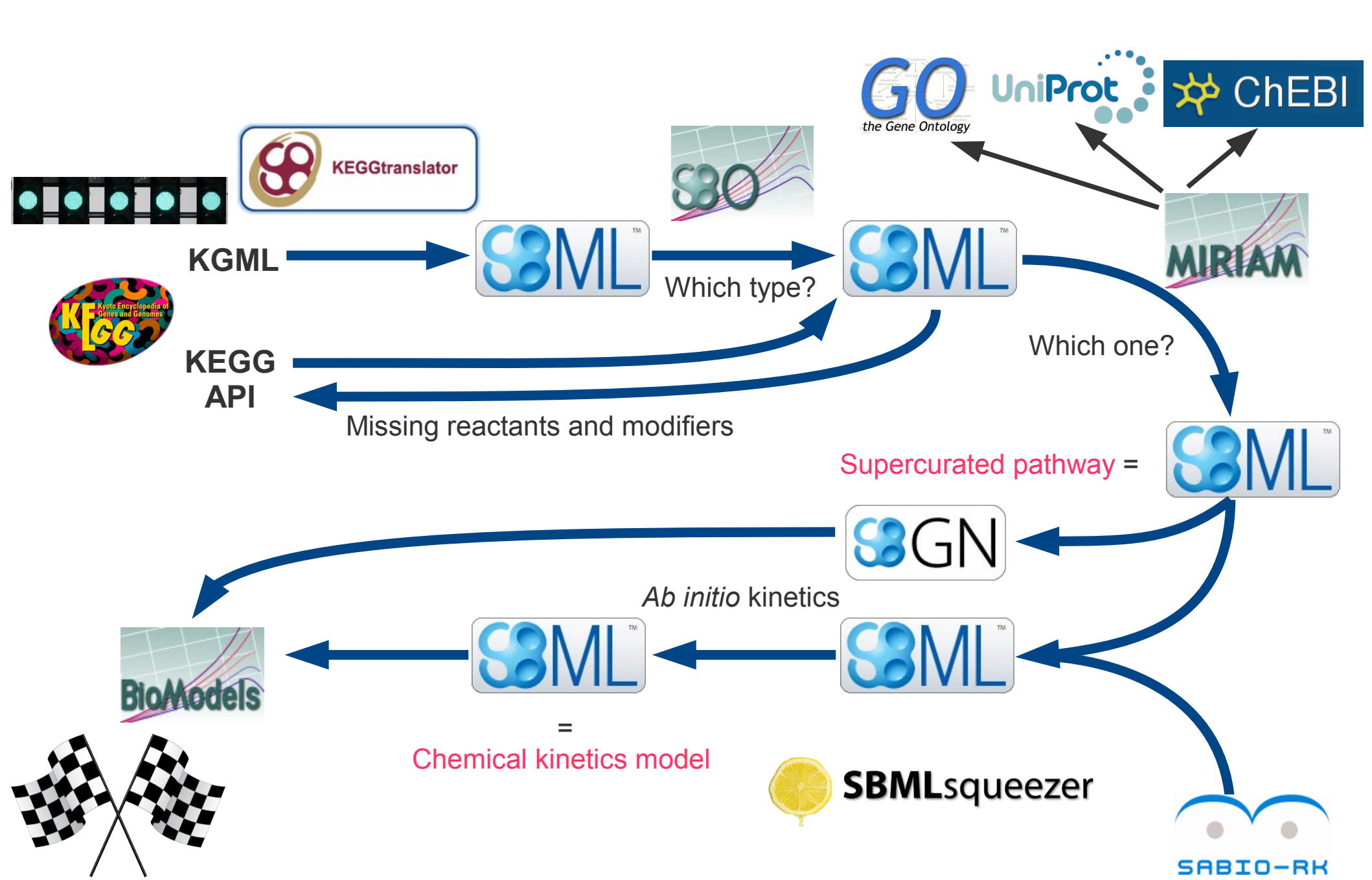

GO
the Gene Ontology
UniProt
ChEBI
KEGGtranslator
SBO
KGML
SBML
Which type?
SBML
MIRIAM
KEGG
API
Which one?
Missing reactants and modifiers
Supercurated pathway =
SBML



Babraham Institute

EMBL-EBI


GO
the Gene Ontology
UniProt
ChEBI
KEGGtranslator
SBO
MIRIAM
KGML
SBML
Which type?
SBML
Which one?
KEGG
API
Missing reactants and modifiers
Supercurated pathway =
SBML
SBGN
Ab initio kinetics
SBML
SBML
SBMLsqueezer
SABIO-RK

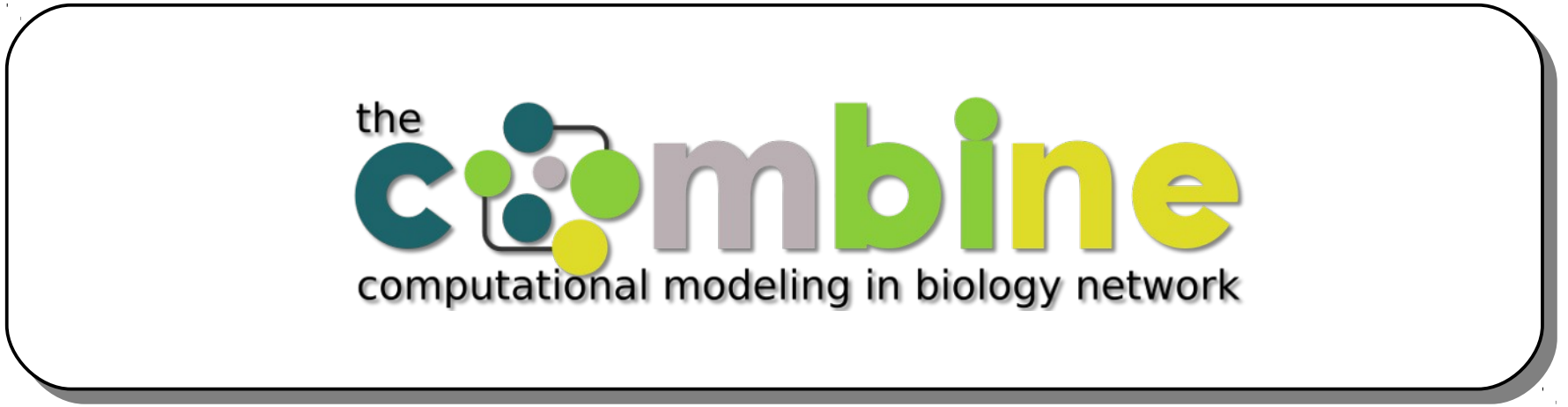

GO
the Gene Ontology
UniProt
ChEBI
KEGGtranslator
SBO
KGML
SBML
Which type?
SBML
MIRIAM
KEGG
API
Which one?
Missing reactants and modifiers
Supercurated pathway =
SBML
SBGN
Ab initio kinetics
BioModels
SBML
SBML
=
Chemical kinetics model
SBMLsqueezer
SABIO-RK

Is it all rosy?

Is it all rosy?

- Heterogeneity
  - Different levels of granularity
  - Different attitude towards implicit and explicit knowledge (e.g. predefined functions, units etc.)
  - Imperfect interoperability (reinvention of wheel; lack of external references)
- Complexity
  - Standards evolve: no clean slate development, plenty of fossils and pseudo-genes
  - All efforts try to cover everything rather than re-use
  - Software support is becoming harder (although helped by good APIs)
- Sustainability
  - Funding: Developing standards is expensive (meetings, dedicated developers etc.)
  - Motivation: Only PIs are stable (avg pos = 2-3 y). Ageing efforts + loss of drive
  - No *a posteriori* involvement of publishers, big agencies or institutes (they only endorse what they develop); standards are constraining

Overarching standardisation structure



Coordinating meetings

- ISCB 2014 end-user workshop (presentations and tutorials)
  - August 2014, Melbourne
  - http://www.icsb14.com/
- COMBINE 2014 (further development of standards)
  - 16 to 20 September 2013, Los Angeles
  - http://co.mbine.org/events/COMBINE_2014
- HARMONY 2014 (development of software support)
  - 22 to 25 April 2014, Manchester
  - http://co.mbine.org/events/HARMONY_2014

COMBINE specifications

View | Edit | Revisions | Access control

Name of the specification	Identifier	Location	node
Biological Pathway Exchange format	biopax	standards/biopax	86
BioPAX Level 1	biopax.level-1	standards/biopax/level-1	89
BioPAX Level 2	biopax.level-2	standards/biopax/level-2	88
BioPAX Level 3	biopax.level-3	standards/biopax/level-3	87
CellML 1.0	cellml.1.0	standards/cellml/1/0	93
CellML 1.1	cellml.1.1	standards/cellml/1/1	94
Systems Biology Graphical Notation	sbgn	standards/sbgn	52
SBGN Activity Flow language	sbgn.af	standards/sbgn/af	66
SBGN AF Level 1	sbgn.af.level-1	standards/sbgn/af/level-1/version-1/0	67
SBGN AF Level 1 Version 1	sbgn.af.level-1.version-1	standards/sbgn/af/level-1/version-1/0	67
SBGN AF Level 1 Version 1.0	sbgn.af.level-1.version-1.0	standards/sbgn/af/level-1/version-1/0	67
SBGN ER	sbgn.er	standards/sbgn/er	57
SBGN ER Level 1	sbgn.er.level-1	standards/sbgn/er/level-1/version-1/2	56
SBGN ER Level 1 Version 1	sbgn.er.level-1.version-1	standards/sbgn/er/level-1/version-1/2	56
SBGN ER Level 1 Version 1.0	sbgn.er.level-1.version-1.0	standards/sbgn/er/level-1/version-1/0	60
SBGN ER Level 1 Version 1.1	sbgn.er.level-1.version-1.1	standards/sbgn/er/level-1/version-1/1	59
SBGN ER Level 1 Version 1.2	sbgn.er.level-1.version-1.2	standards/sbgn/er/level-1/version-1/2	56
SBGN PD	sbgn.pd	standards/sbgn/pd	61
SBGN PD Level 1	sbgn.pd.level-1	standards/sbgn/pd/level-1/version-1/3	62
SBGN PD Level 1 Version 1	sbgn.pd.level-1.version-1	standards/sbgn/pd/level-1/version-1/3	62
SBGN PD Level 1 Version 1.0	sbgn.pd.level-1.version-1.0	standards/sbgn/pd/level-1/version-1/0	65
SBGN PD Level 1 Version 1.1	sbgn.pd.level-1.version-1.1	standards/sbgn/pd/level-1/version-1/1	64
SBGN PD Level 1 Version 1.2	sbgn.pd.level-1.version-1.2	standards/sbgn/pd/level-1/version-1/2	63
SBGN PD Level 1 Version 1.3		vel-1/version-1/3	62
Systems Biology Markup Langu			55
SBML Level 1	sbml.level-1	standards/sbml/level-1/version-2	78
SBML Level 1 Version 1	sbml.level-1.version-1	standards/sbml/level-1/version-1	79

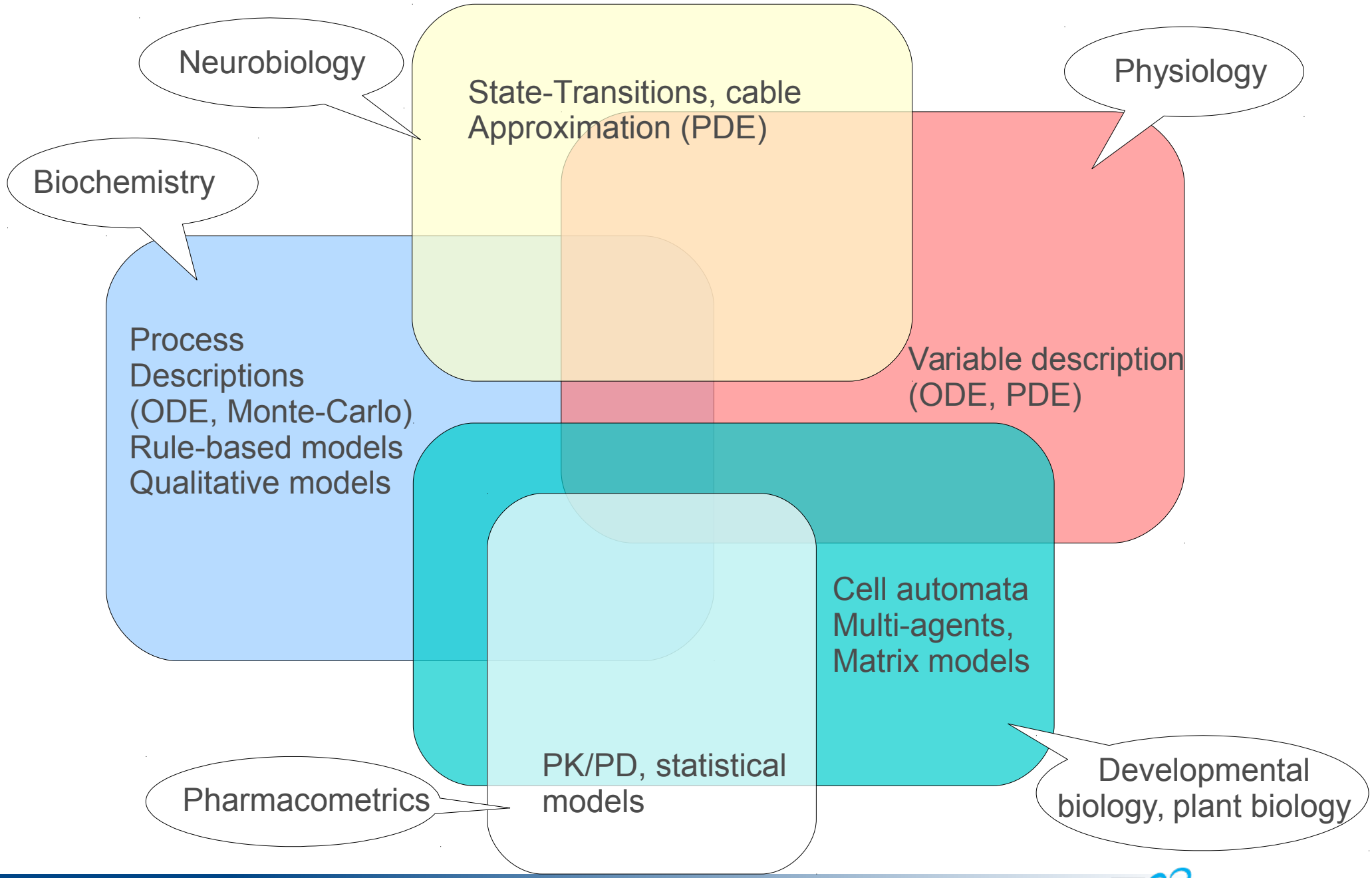
http://co.mbine.org/standards/specifications/

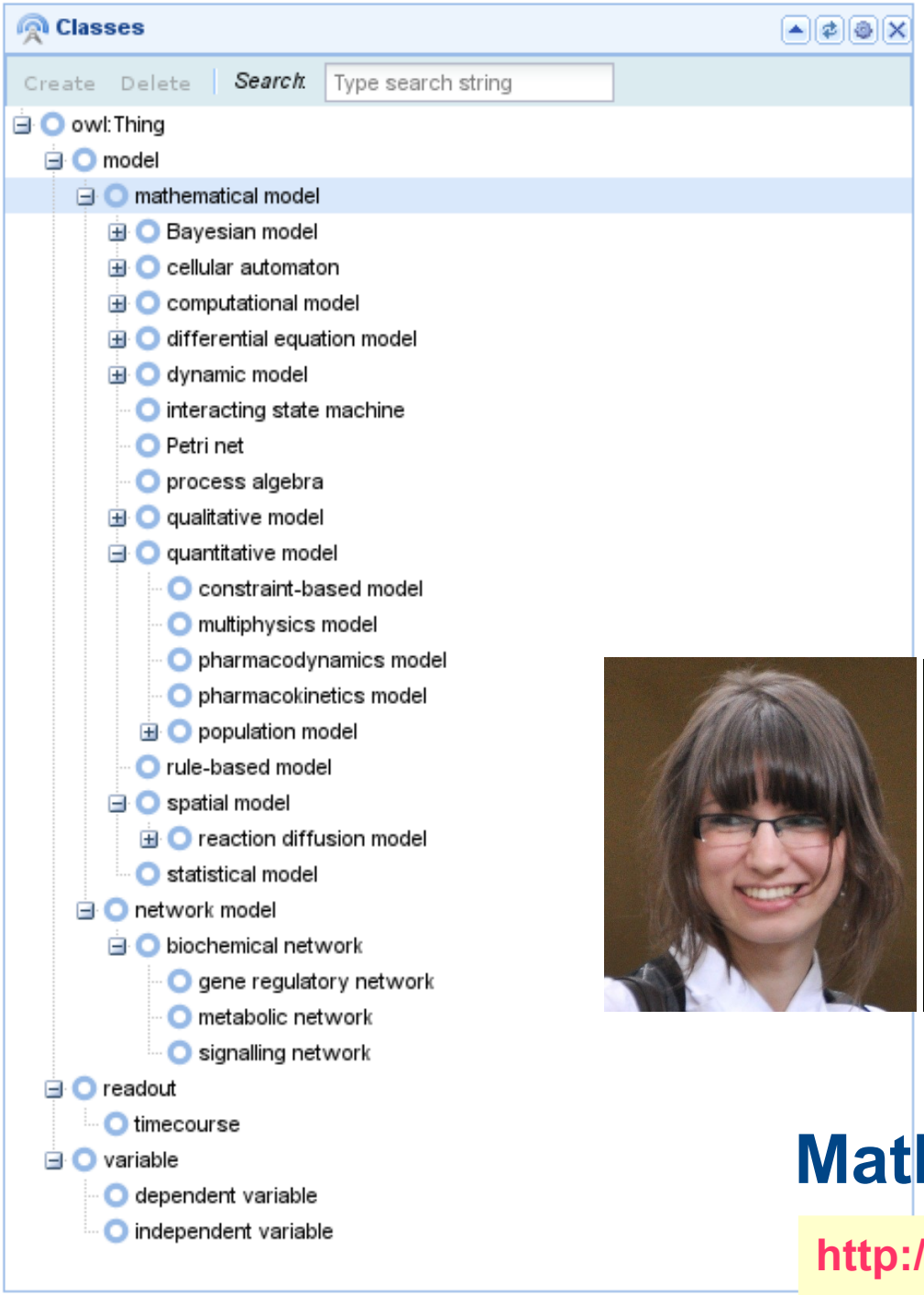
COMBINE archive: “One File to bundle them all”

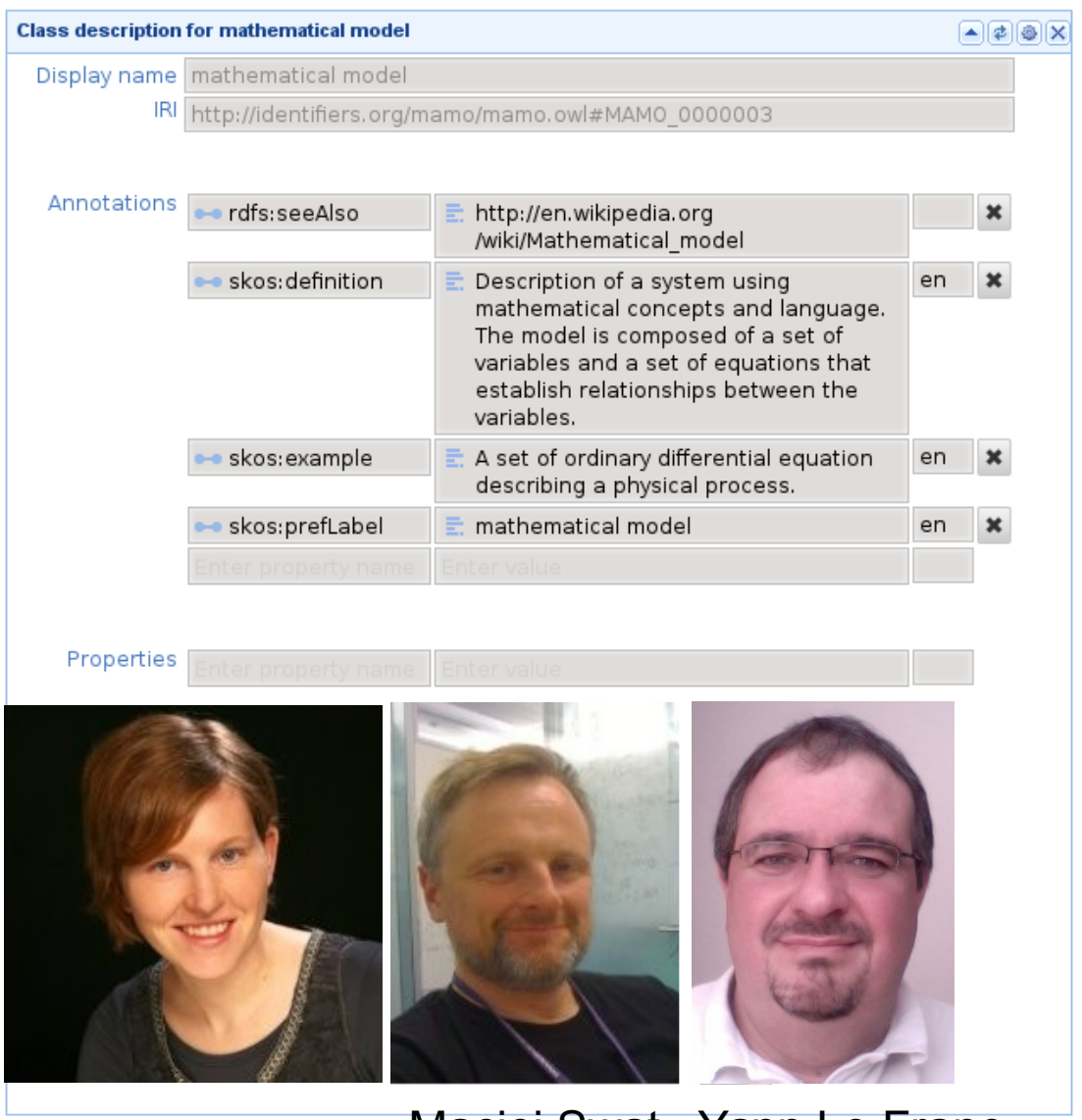
```
<?xml version="1.0" encoding="utf-8"?>
<omexManifest xmlns="http://identifiers.org/combine.specifications/omex-manifest">
    <content location="./manifest.xml"
             format="http://identifiers.org/combine.specifications/omex-manifest"/>
    <content location="./model/model.xml"
             format="http://identifiers.org/combine.specifications/sbml"/>
    <content location="./simulation.xml"
             format="http://identifiers.org/combine.specifications/sed-ml"/>
    <content location="./data.pdf"
             format="application/pdf"/>
    <content location="./metadata.rdf"
             format="http://identifiers.org/combine.specifications/omex-metadata"/>
</omexManifest>
```

http://co.mbine.org/documents/archive

Many alternative modelling approaches







Maciej Swat Yann Le Franc

Mathematical Modelling Ontology

http://sourceforge.net/projects/mamo-ontology

Babraham Institute

EMBL-EBI

I am grateful and indebted to

All the collaborators listed during the presentation

SBGN, SBML, SED-ML editors

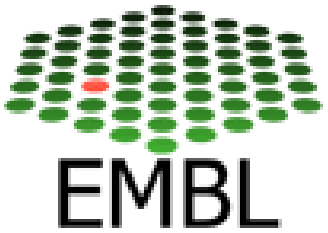
Developers of related ontologies and software

Organisers of meetings and efforts to support our standards

The community of Computational Systems Biology

Frédéric Bois, François Fages, Macha Nikolsi,
Anne Poupon, Denis Thieffry


NIGMS


EMBL


BBSRC
bioscience for the future


NEDO


elixir


cnrs


MEMORIAL SLOAN-KETTERING CANCER CENTER
ESTABLISHED 1884


ENFIN
Enabling Systems Biology


CAM


Netherlands Consortium for
Systems Biology


UCONN HEALTH CENTER


nbic
netherlands
bioinformatics
centre


Universiteitsfonds Limburg
| SWOL |

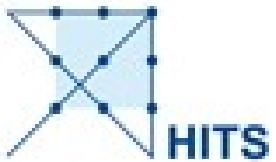

OICR
Ontario Institute
for Cancer Research


Maastricht
University


National
Human Genome
Research Institute


OntarioGenomicsInstitute


INSTITUT DES SYSTEMES COMPLEXES
ISC
PARIS ÎLE-DE-FRANCE


HITS


virtual liver
network


DFG


FONDATION
PIERRE-GILLES
DE GENNES
POUR LA RECHERCHE