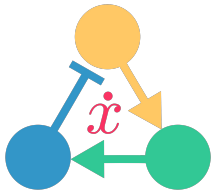


Standards and databases for sharing models in systems biology

Nicolas Le Novère
n.lenovere@gmail.com

Babraham Institute

EMBL-EBI

“You should not develop standards and easy to use modelling software. It allows biologists to write models, and they don't know how to do it properly.”

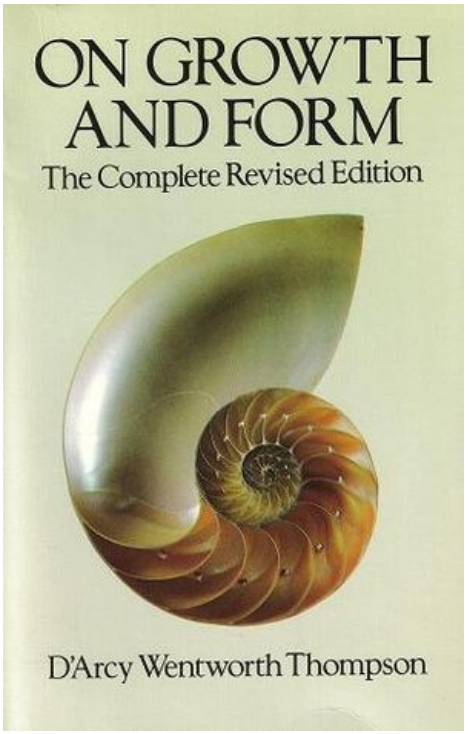
influential British biomathematician, EU Syst Biol centres meeting, 2007

“By developing BioModels you harm the cause of modelling in biology. My students do not learn how to make a model, they download it ready to use instead”

Famous theoretical biologist, BioScope 2006

Why using mathematical models?

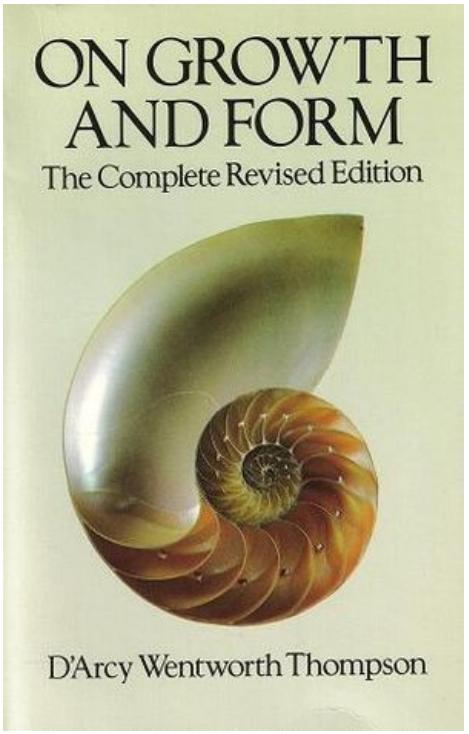
Describe



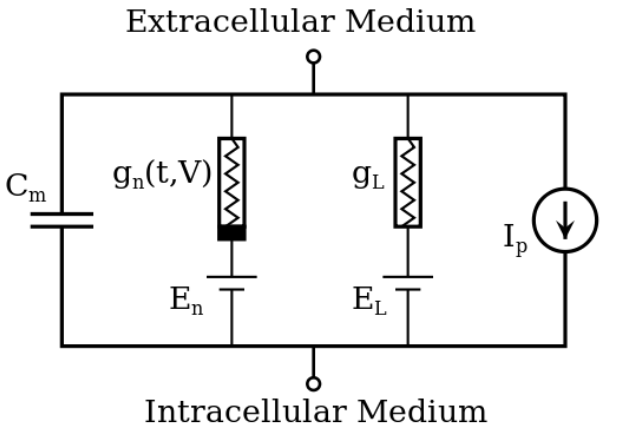
1917

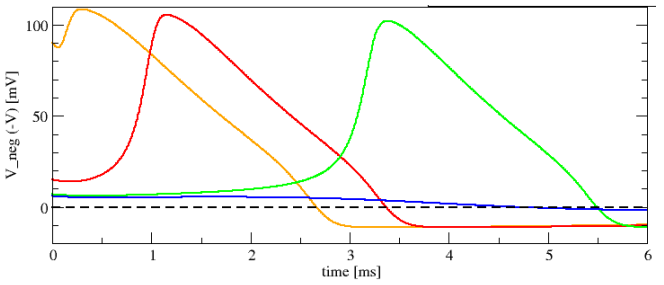
Why using mathematical models?

Describe	Explain



1917

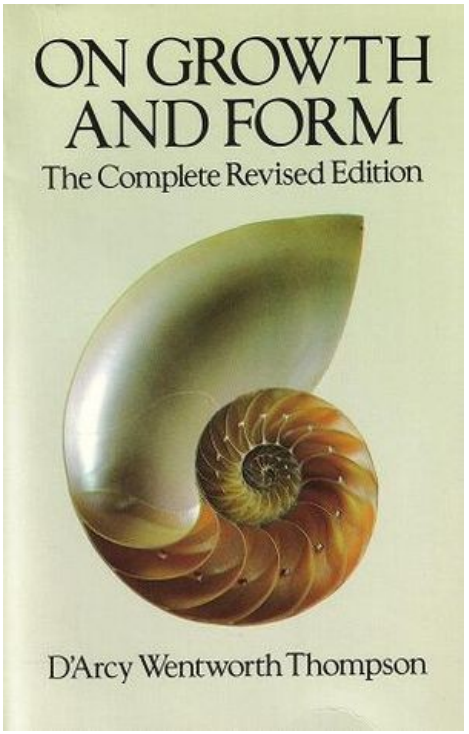


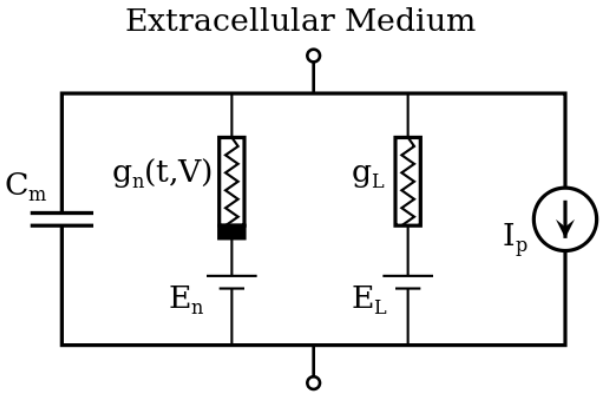


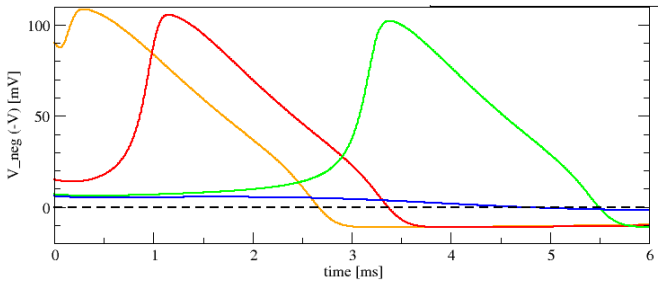
1952

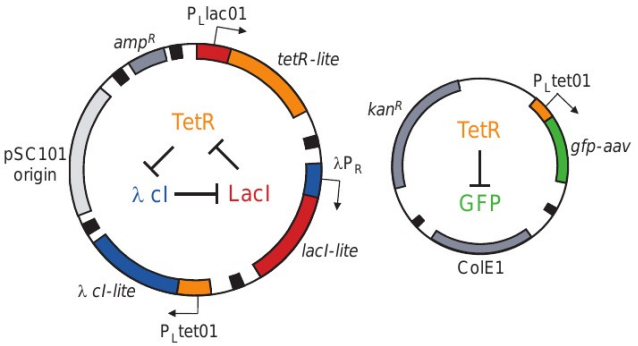
Why using mathematical models?

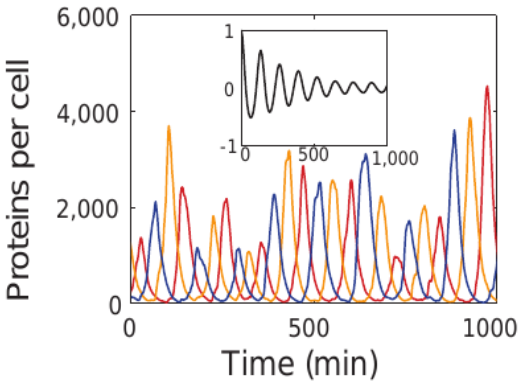
Describe	Explain	Predict











1917 1952 2000

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

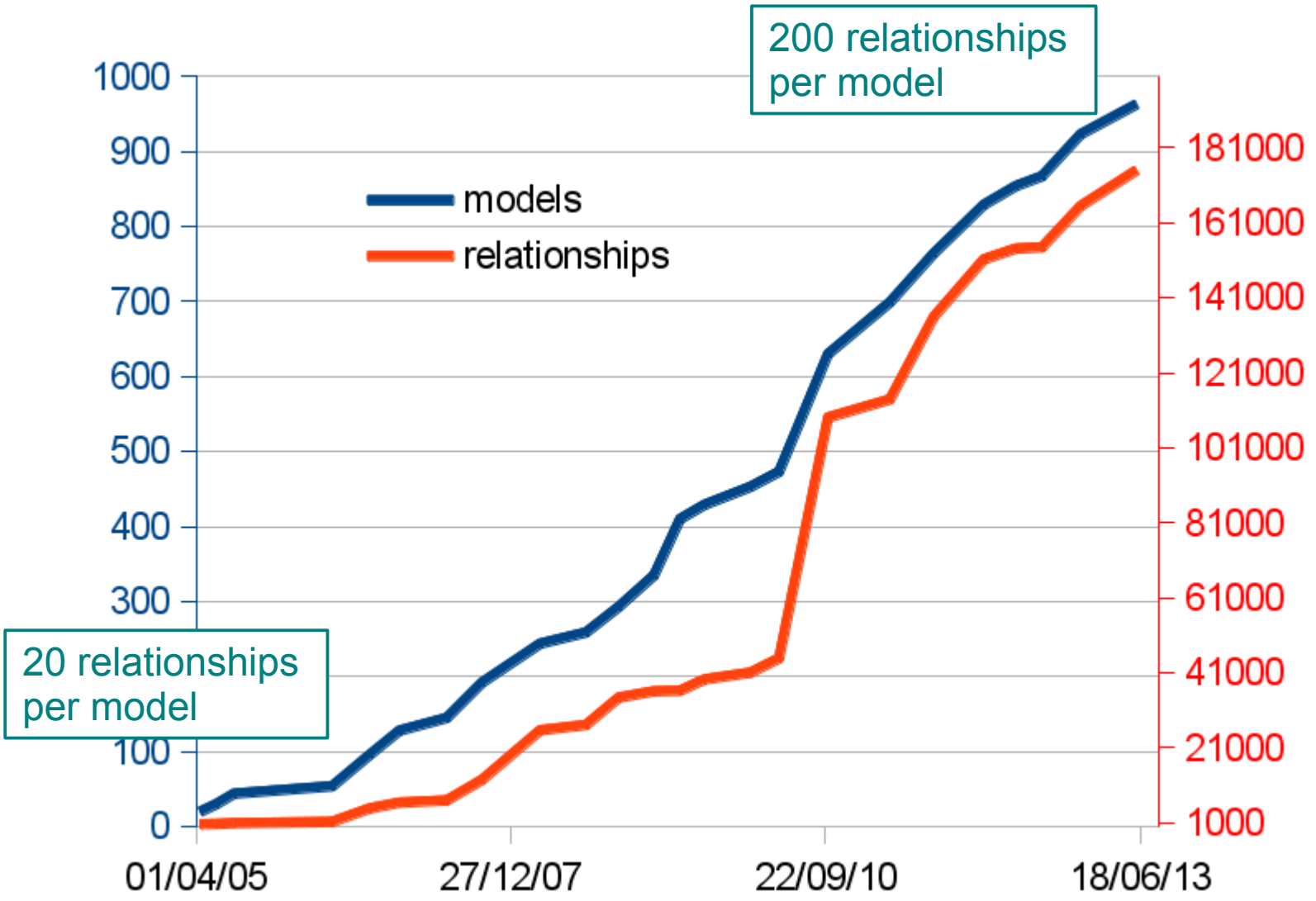
What is a mathematical model?

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

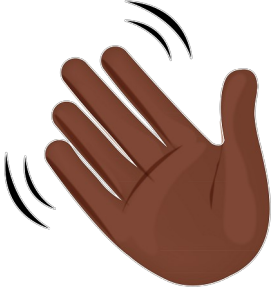
Therefore we need to share

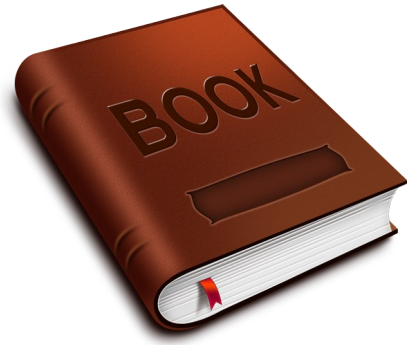
Model descriptions

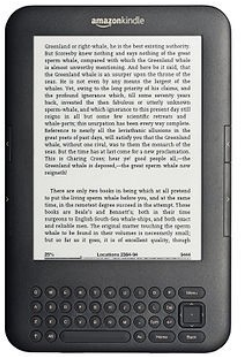
Simulation descriptions

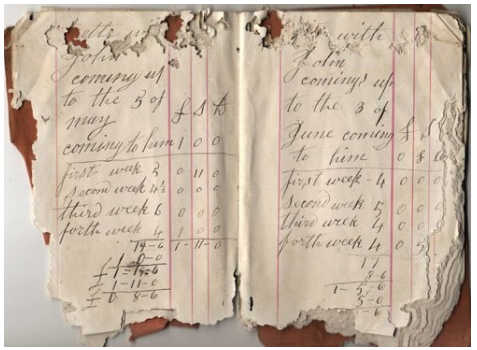
Parametrisations

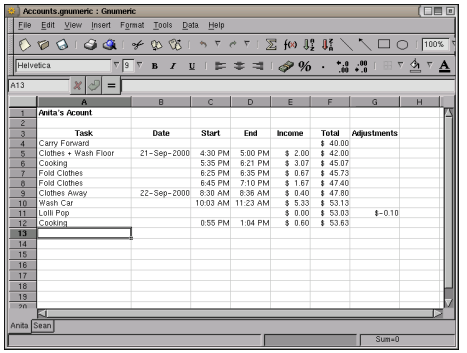
What are standards good for (1)?









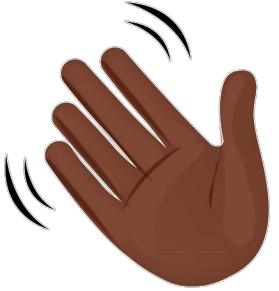


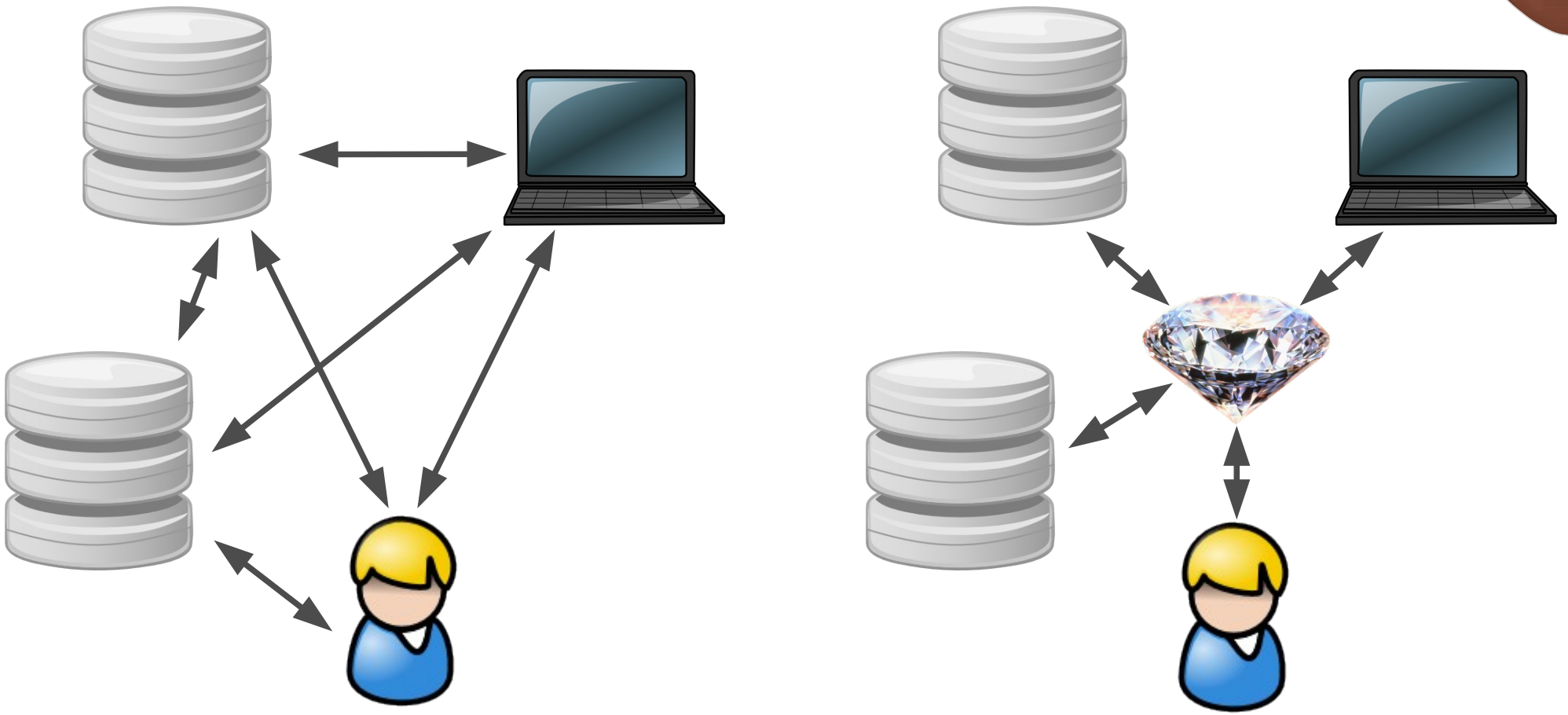
$$\int_1^t \frac{dx}{x}$$

```
<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

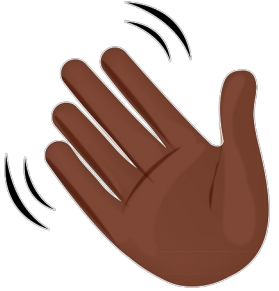
What are standards good for (2)?

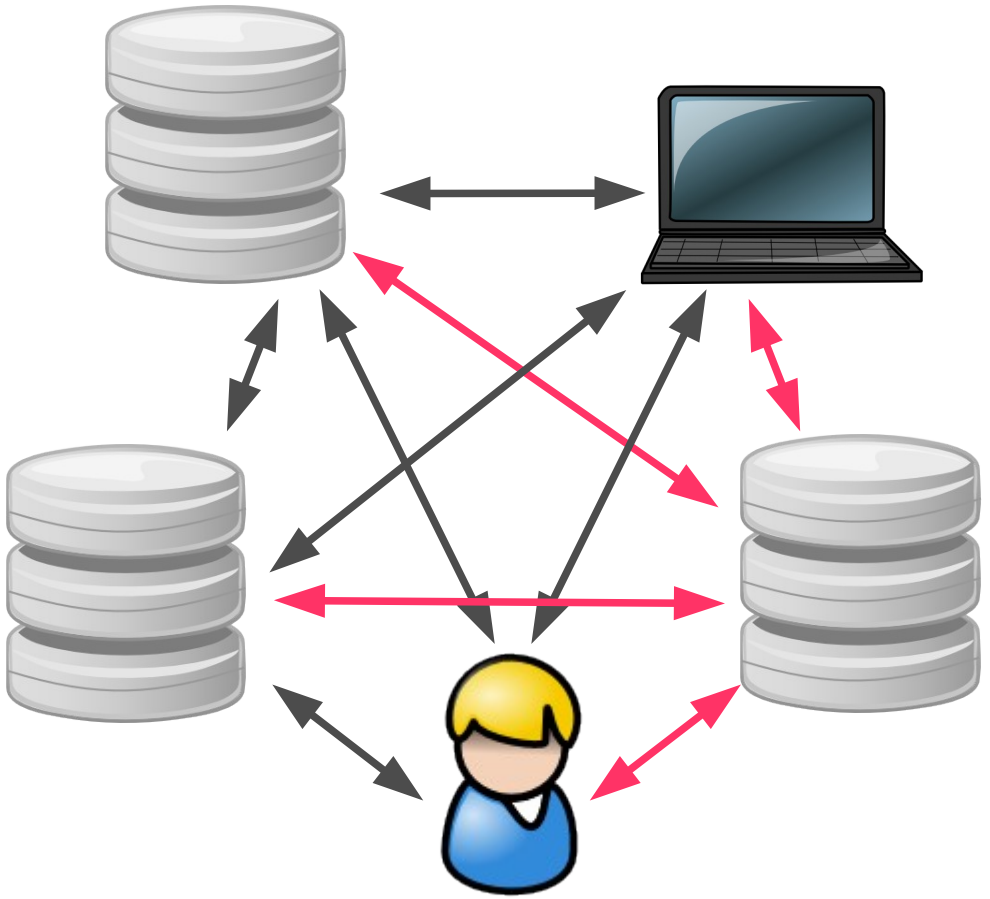


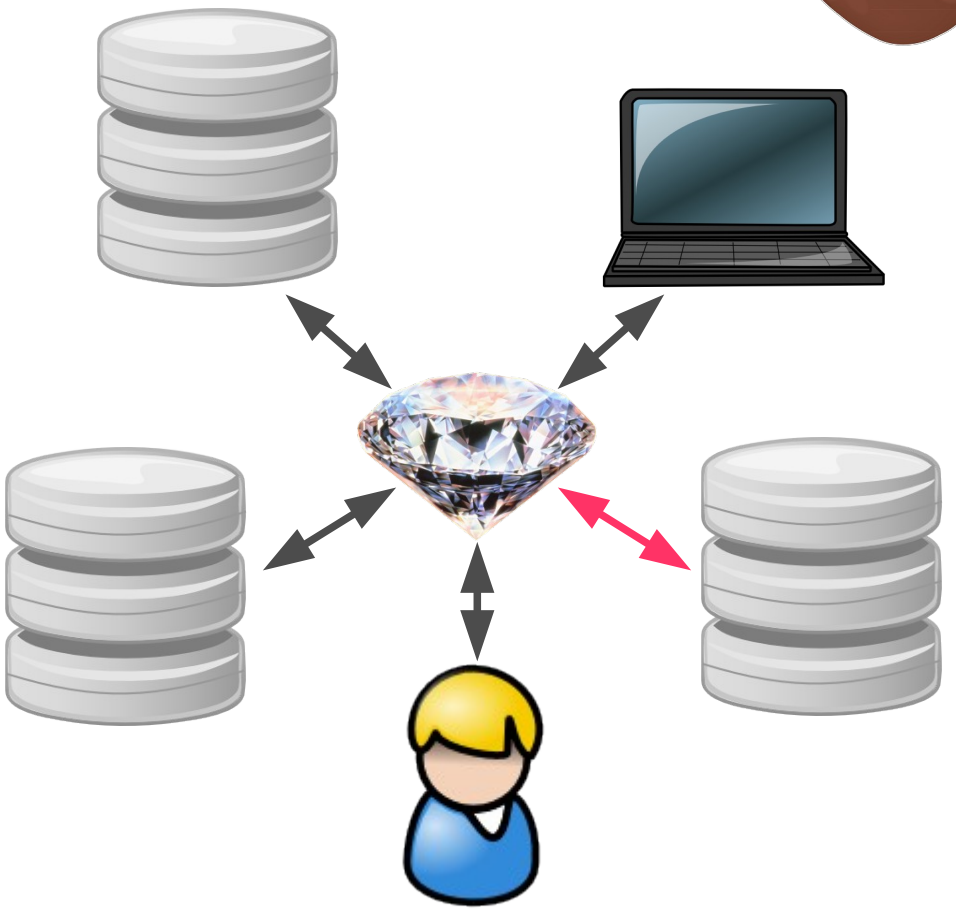


N tools require N conversions for exchange and not N(N-1)/2

What are standards good for (2)?

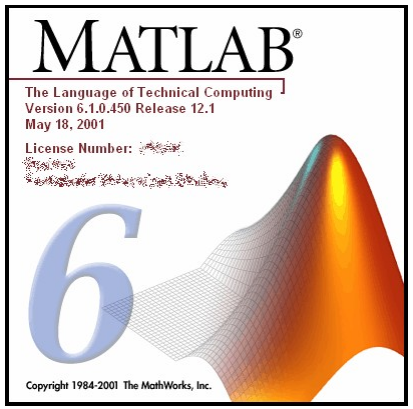


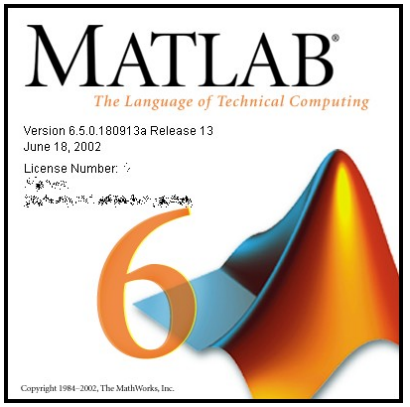


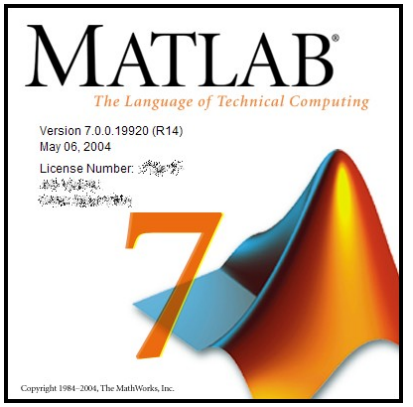


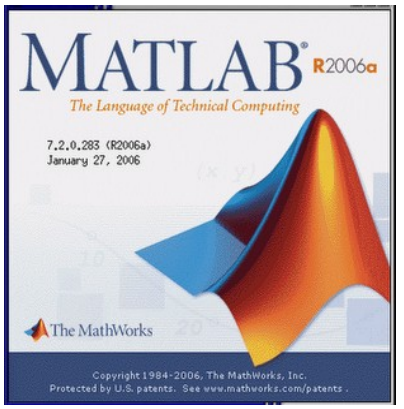
N tools require N conversions for exchange and not N(N-1)/2

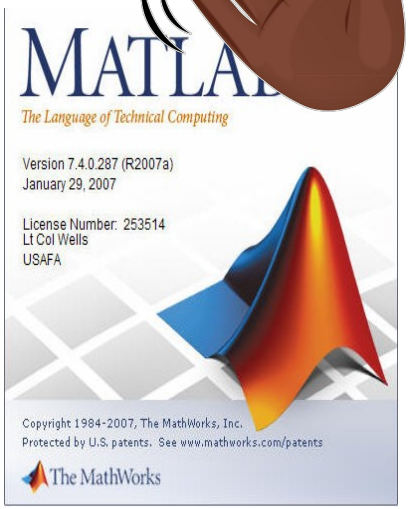
What are standards good for (3)?

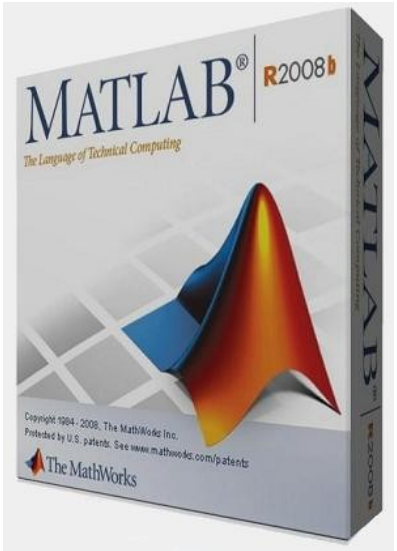


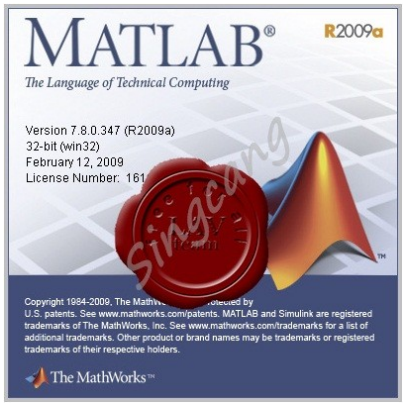


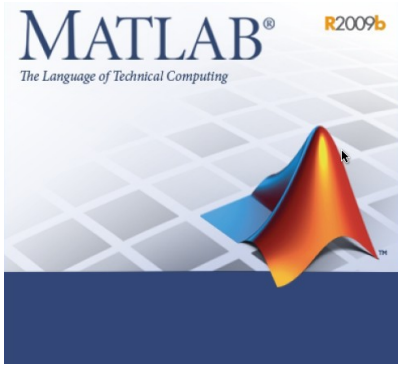


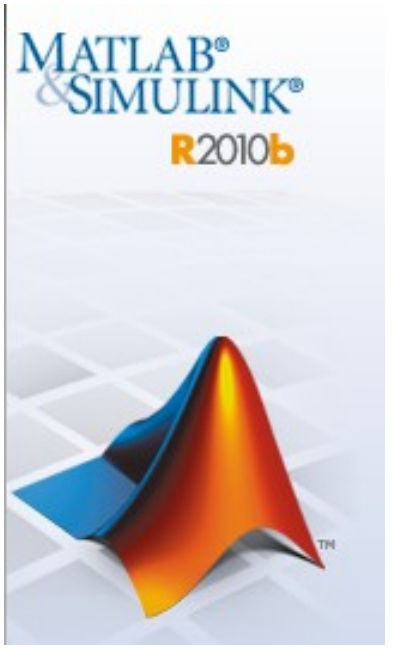


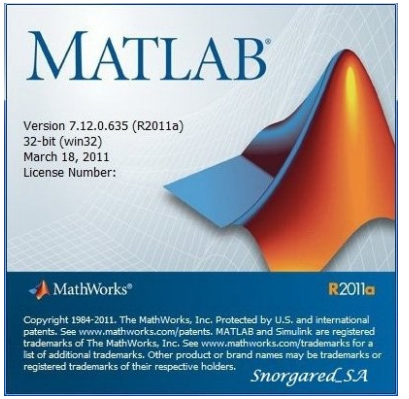






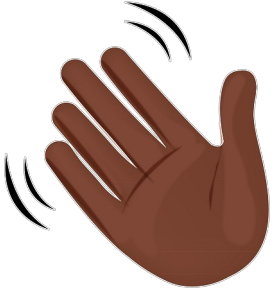


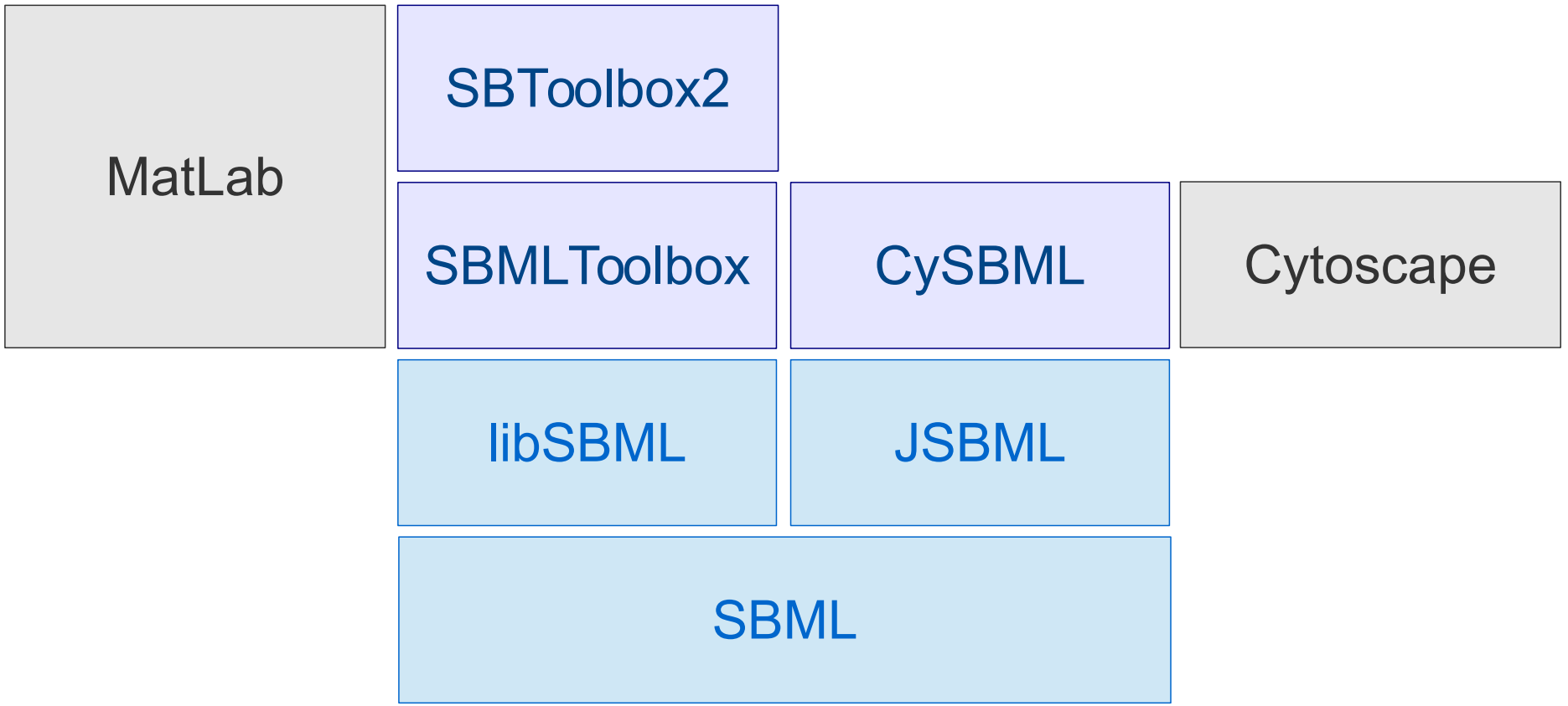




Open standards are more stable than proprietary formats

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 matrix of standards

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

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

SBML Software Guide

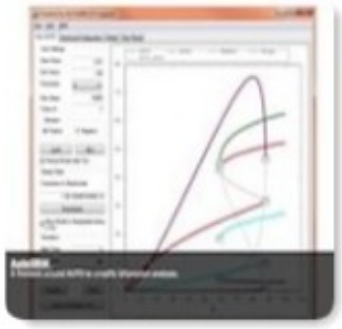
The following pages describe SBML-compatible software packages known to us. We offer different ways of viewing the information, all drawn from the same underlying data collected from the systems' developers via our **software survey**. The *Matrix* provides a table listing all known software and a variety 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: **262**.

Go to the SBML Software Matrix

Go to the SBML Software Summary

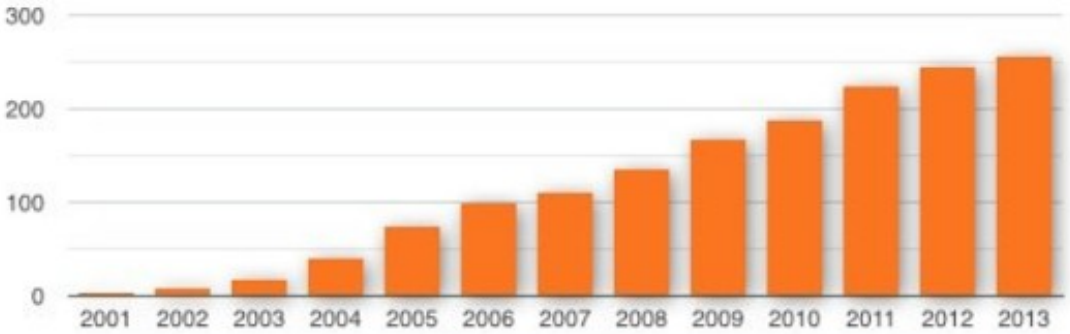
Go to the SBML Software Showcase



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

Historical trend

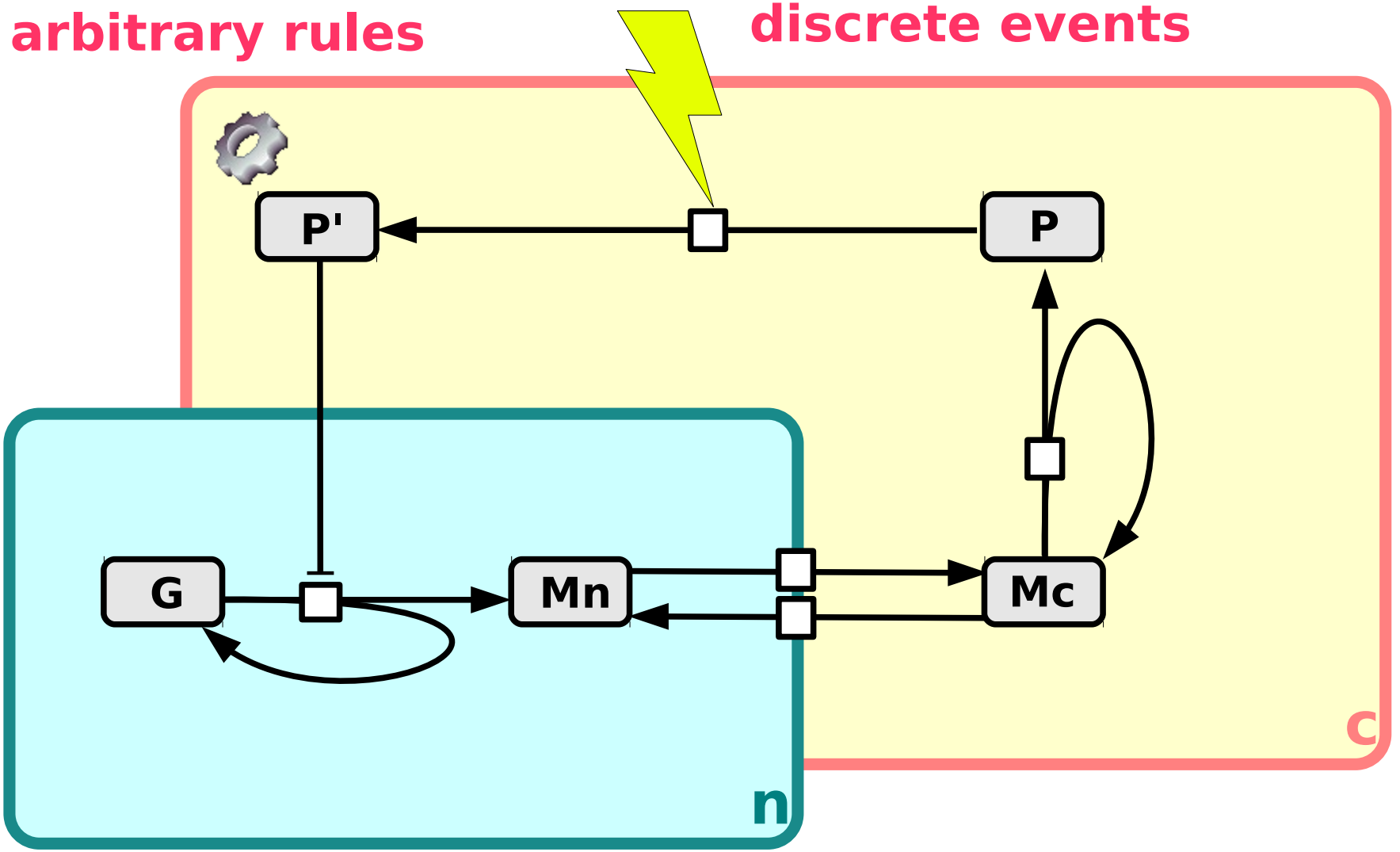
The following graph shows the total number of known SBML-compatible software packages each year, as counted by the SBML Team. The counts shown are for approximately the middle of each year.



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 inspect them.

What can we encode in SBML (core)?



Why the Extensible Markup Language (XML)?

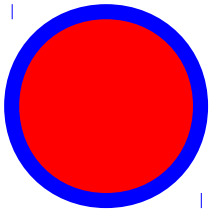
- **HTML**

```
A <strong>strong word</strong> and an
<a href=”http://www.w3.org/”>hyperlink</a>
```

A **strong word** and an hyperlink

- **SVG**

```
<circle r="100" fill="red"
stroke="blue" stroke-width="10"  />
```



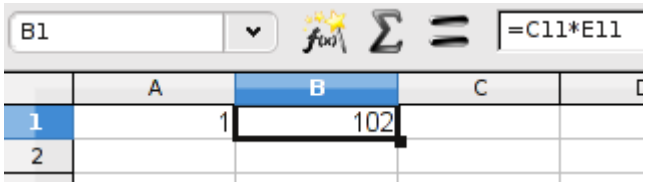
- **MathML**

```
<apply>
  <int/>
  <bvar><ci> x </ci></bvar>
  <lowlimit><cn> 0 </cn></lowlimit>
  <uplimit><ci> a </ci></uplimit>
  <apply><ci> f </ci><ci> x </ci></apply>
</apply>
```

$$\int_0^a f(x)\,dx$$

- **Excel**

```
<row r="1">
  <c r="A1"><v>1</v></c>
  <c r="B1">
    <f>C11*E11</f>
    <v>102</v>
  </c>
</row>
```



Why the Extensible Markup Language (XML)?

- Easy to define and validate
  - Rapid prototyping, processing tools can be generated and thrown away
- Existence of a very large toolkit
  - Libraries in every programming languages
  - A very large number of description formats in life sciences are in XML
- Associated technologies
  - Definition: XML Schema, Schematron (themselves XML)
  - Conversion: XSLT (using XSL in XML)
  - Linking: XPath and XQuery

Structure of SBML

```
<?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>
```

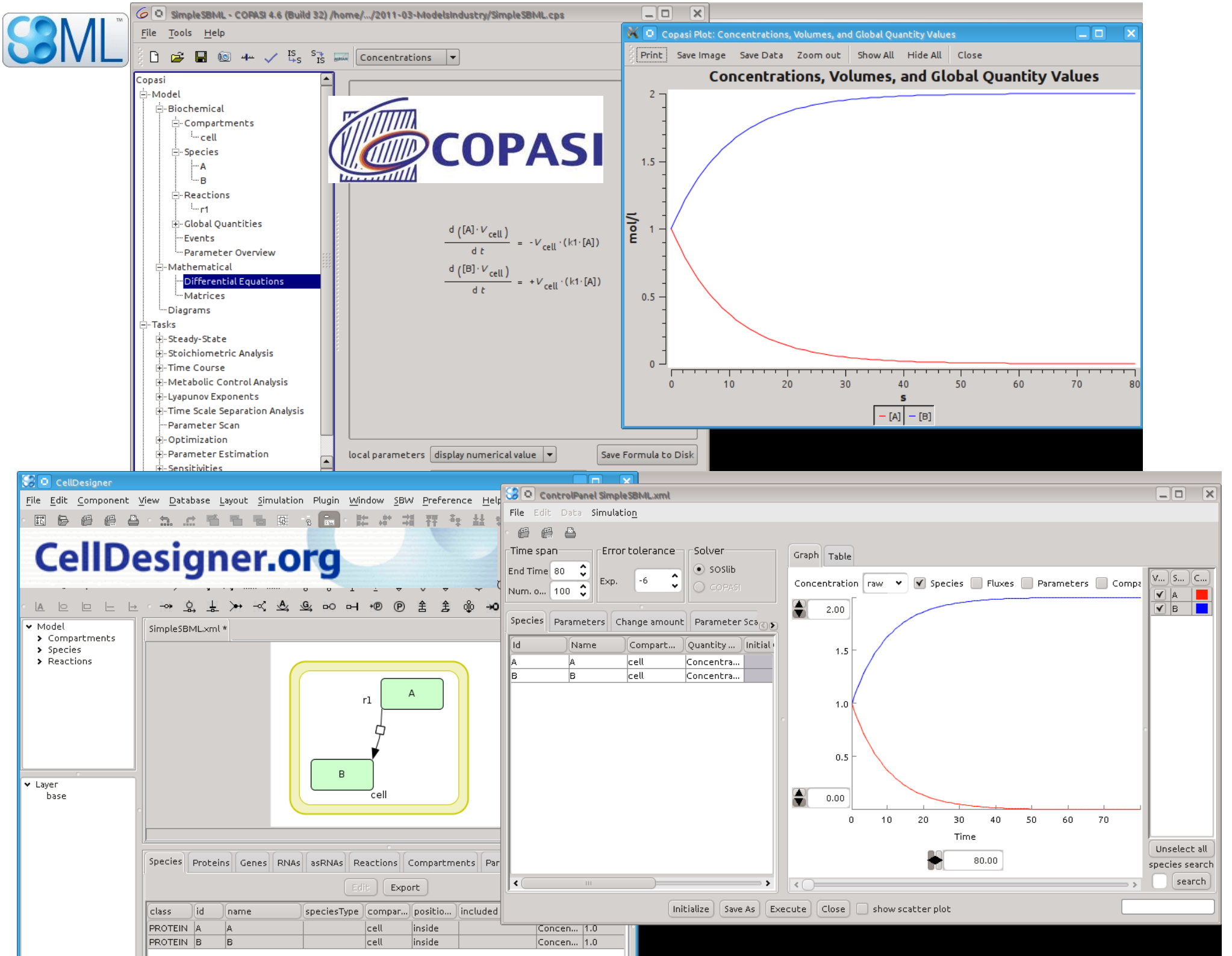
variables

relationships

```
<?xml 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>
```

MathML

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 more realistic example ...

```
<species
    id="A"
    name="a-tubulin"
    compartment="cell"
    initialAmount="1000"
    substanceUnits="item"
    hasOnlySubstanceUnits="true"
    boundaryCondition="true"
    constant="false"
    charge="0"
    metaid="PX"
    sboTerm="SBO:0000245" >
  <notes>
    <body xmlns="http://www.w3.org/1999/xhtml">
      <p>One of the components of a microtubule</p>
    </body>
  </notes>
  <annotation>
    <rdf:RDF
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
        xmlns:bqmodel="http://biomodels.net/model-qualifiers/"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
      <rdf:Description rdf:about="#PX">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P68370"/>
            <rdf:li rdf:resource="urn:miriam:obo.go:GO%3A0045298"/>
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

A more realistic example ...

```
<species·
    id="A"·
    name="a-tubulin"
    compartment="cell"
    initialAmount="1000"
    substanceUnits="item"
    hasOnlySubstanceUnits="true"
    boundaryCondition="true"
    constant="false"
    charge="0"
    metaid="PX"
    sboTerm="SBO:0000245" >
  <notes>
    <body xmlns="http://www.w3.org/1999/xhtml">
      <p>One of the components of a microtubule</p>
    </body>
  </notes>
  <annotation>
    <rdf:RDF·
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"·
        xmlns:bqmodel="http://biomodels.net/model-qualifiers/"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
      <rdf:Description rdf:about="#PX">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P68370"/>
            <rdf:li rdf:resource="urn:miriam:obo.go:GO%3A0045298"/>
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

biological semantics: macromolecule

XHTML

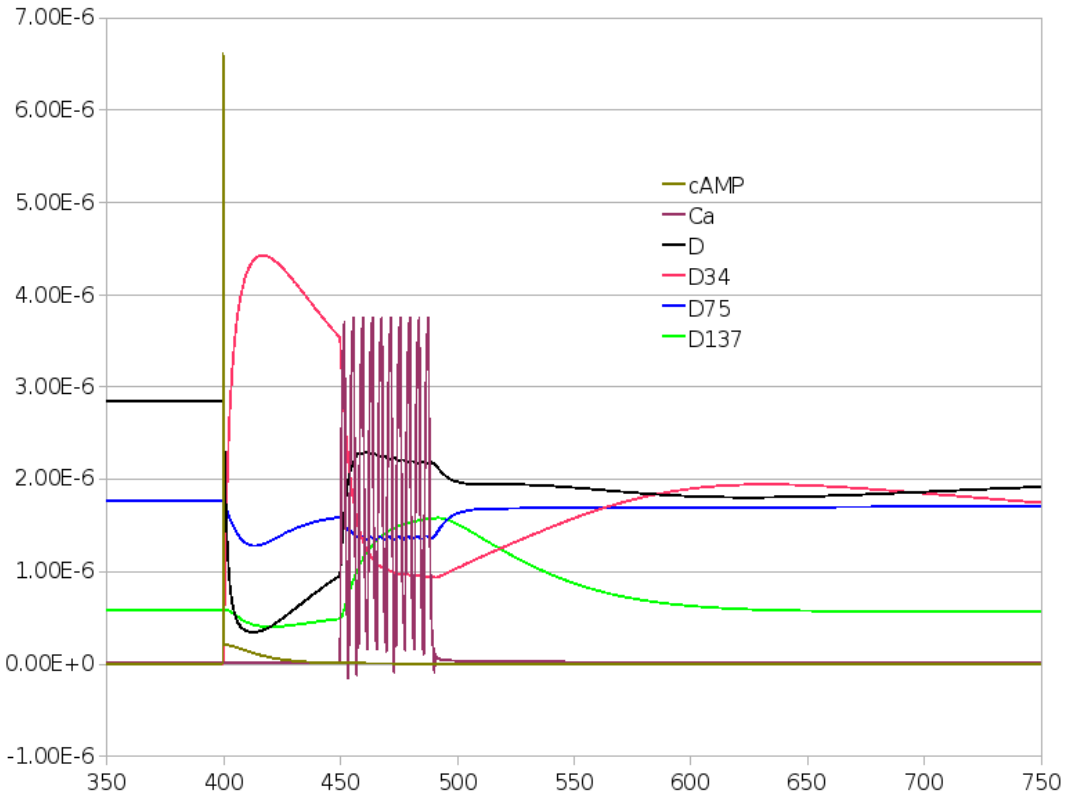
RDF

SBML is not limited to biochemistry!

- A species is a pool of entities participating to a reaction, **not always** a **chemical** entity
  - It can be a pool of molecules
  - It can be a pool of cells
  - It can be a pool of organs
  - It can be a population of organism
- Rate Rules can describe the temporal evolution of any quantitative parameter, e.g. transmembrane voltage, tumour size etc.
- Events can describe any discontinuous change, e.g. neurotransmitter release, repolarisation, cell division etc.

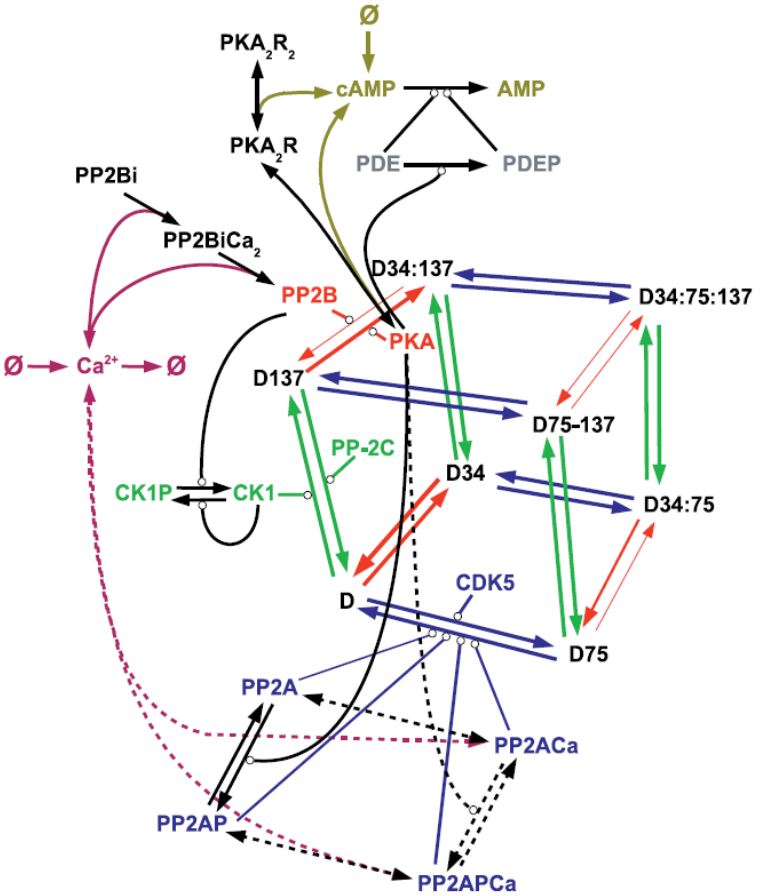
➔ SBML (Core) is about process descriptions

Biochemical models



Fernandez et al. DARPP-32 is a robust integrator of dopamine and glutamate signals
PLoS Comput Biol (2006) 2: e176.

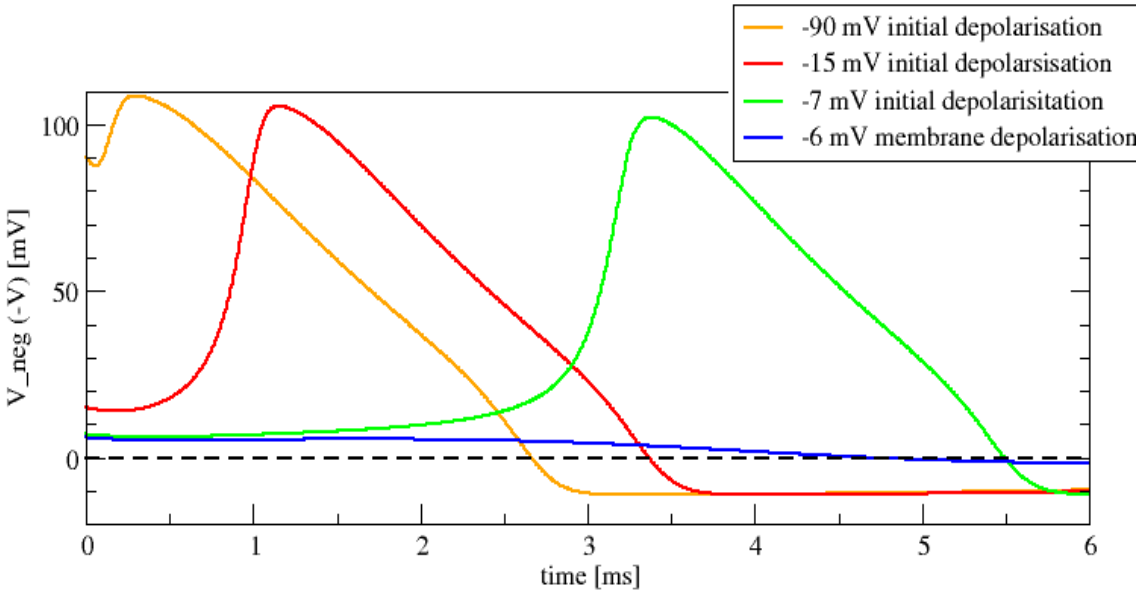
BIOMD0000000153



reaction:

$$v_{\text{on}} = k_{\text{on}} \times [\text{D}] \times [\text{CDK5}] \times \text{Vol}$$

Conductance-based model

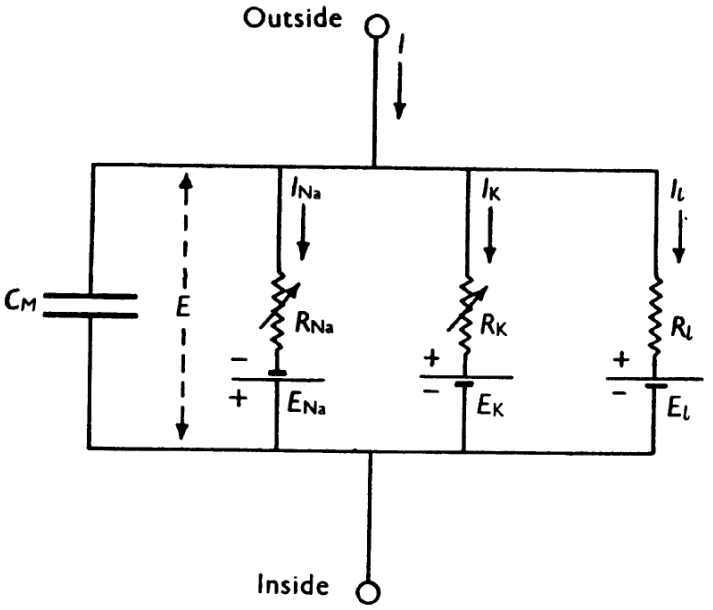


Hodgkin AL, Huxley AF.
A quantitative description of membrane current and its application to conduction and excitation in nerve.
J Physiol (1952) 117:500-544.

BIOMD0000000020

rate rule:

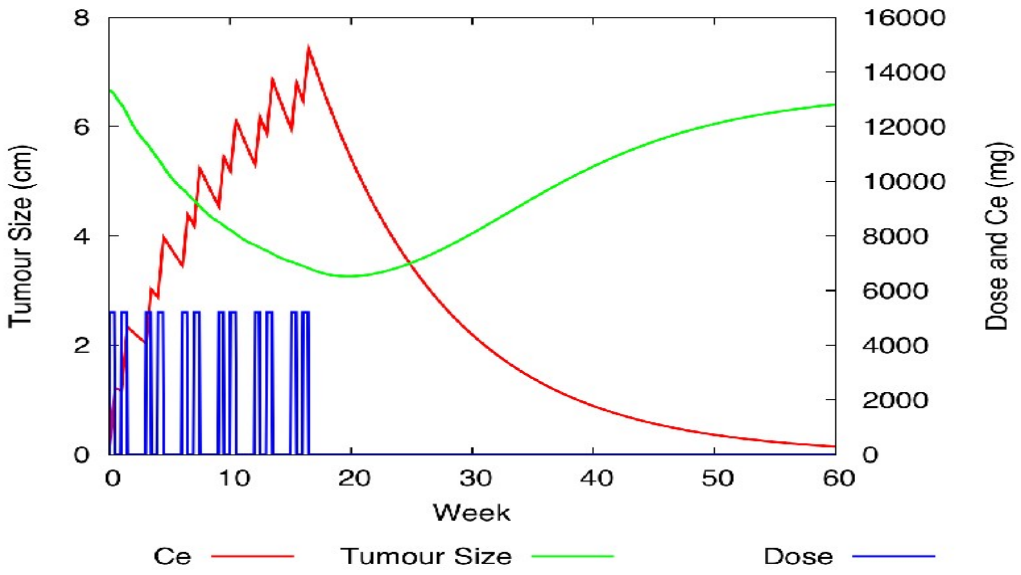
$$\frac{dv}{dt} = \frac{I - (i_{\mathrm{Na}} + i_{\mathrm{K}} + i_{\mathrm{L}})}{C_{\mathrm{m}}}$$



assignment rule:

$$i_{\mathrm{Na}} = g_{\mathrm{Na}} \times m^3 \times h \times (V - E_{\mathrm{Na}})$$

Pharmacometrics models



Tham et al (2008) A pharmacodynamic model for the time course of tumor shrinkage by gemcitabine + carboplatin in non-small cell lung cancer patients. *Clin Cancer Res*. 2008 14(13): 4213-8.

BIOMD0000000234

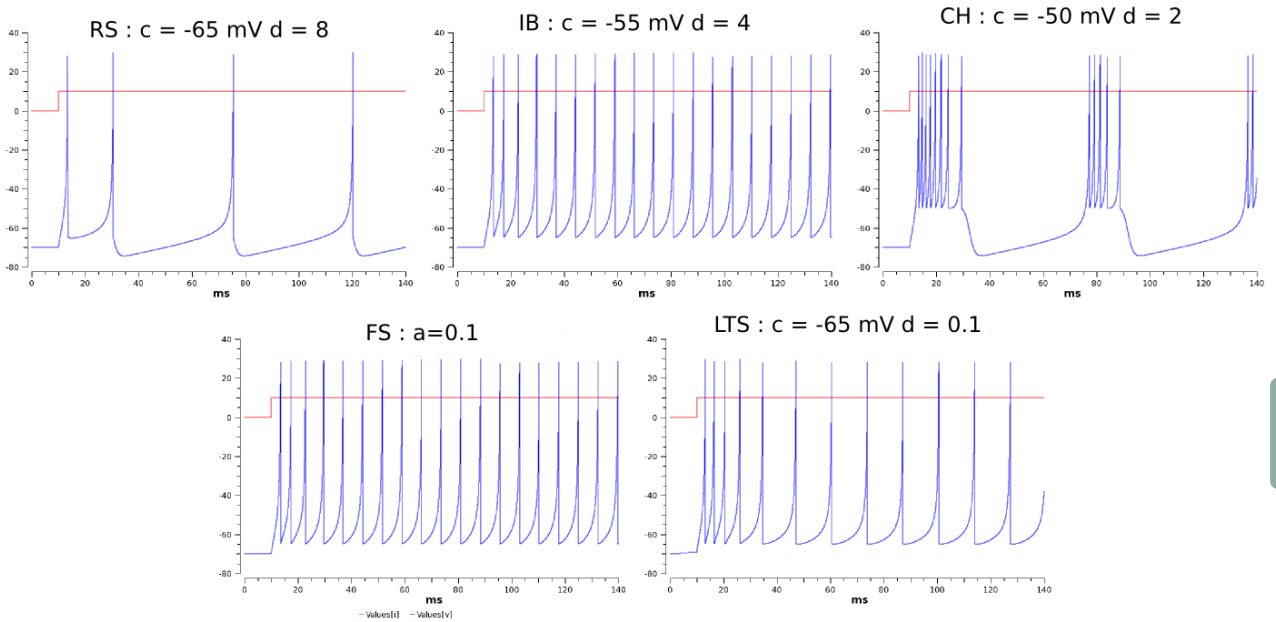
rate rule:

$$\frac{dSize}{dt} = (Rate_{in} \times Effect - K_{over} \times Size) \times Size$$

assignment rule:

$$Effect = 1 - \frac{E_{max} \times Ce}{Amt_{50} + Ce}$$

Single-compartment neurons



Izhikevich EM. Simple model of spiking neurons. *IEEE Trans Neural Netw* (2003) 14(6):1569-1572.

BIOMD0000000127

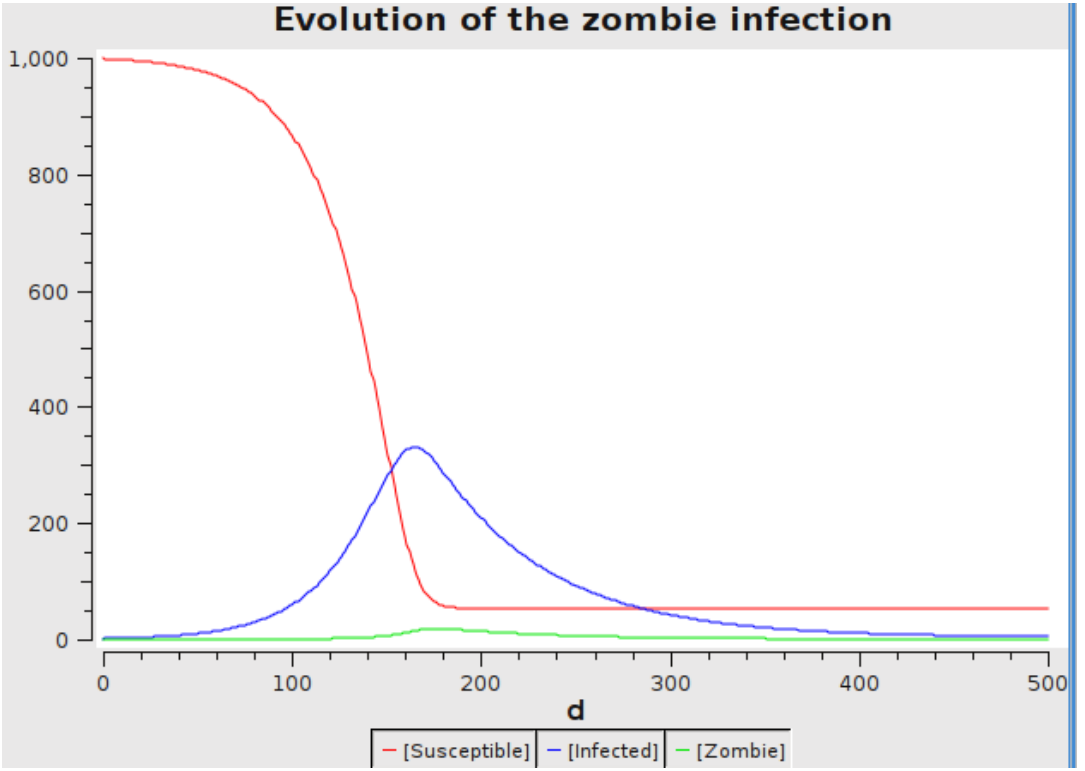
rate rule:

$$\frac{dv}{dt} = 0.04^2 + 5 \times V + 140 - U + i$$

event:

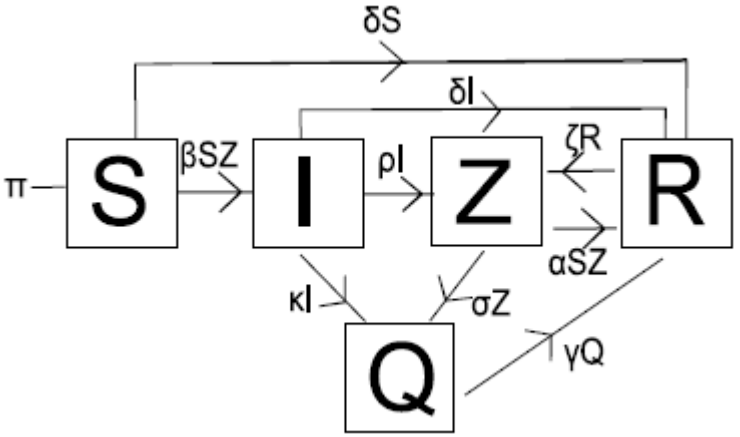
$$\text{when } v > V_{\text{thresh}} \begin{cases} v = c \\ U = U + d \end{cases}$$

Spread of infection diseases ...



Munz P et al. When zombies attack!: Mathematical modelling of an outbreak of zombie infection. in "Infectious Disease Modelling Research Progress", (2009)133-150

MODEL1008060001



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

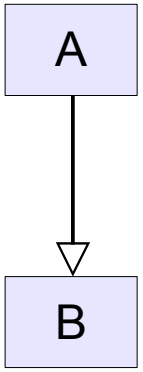
A not so simple SBML file (Recon2)

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

MODEL1109130000

- 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

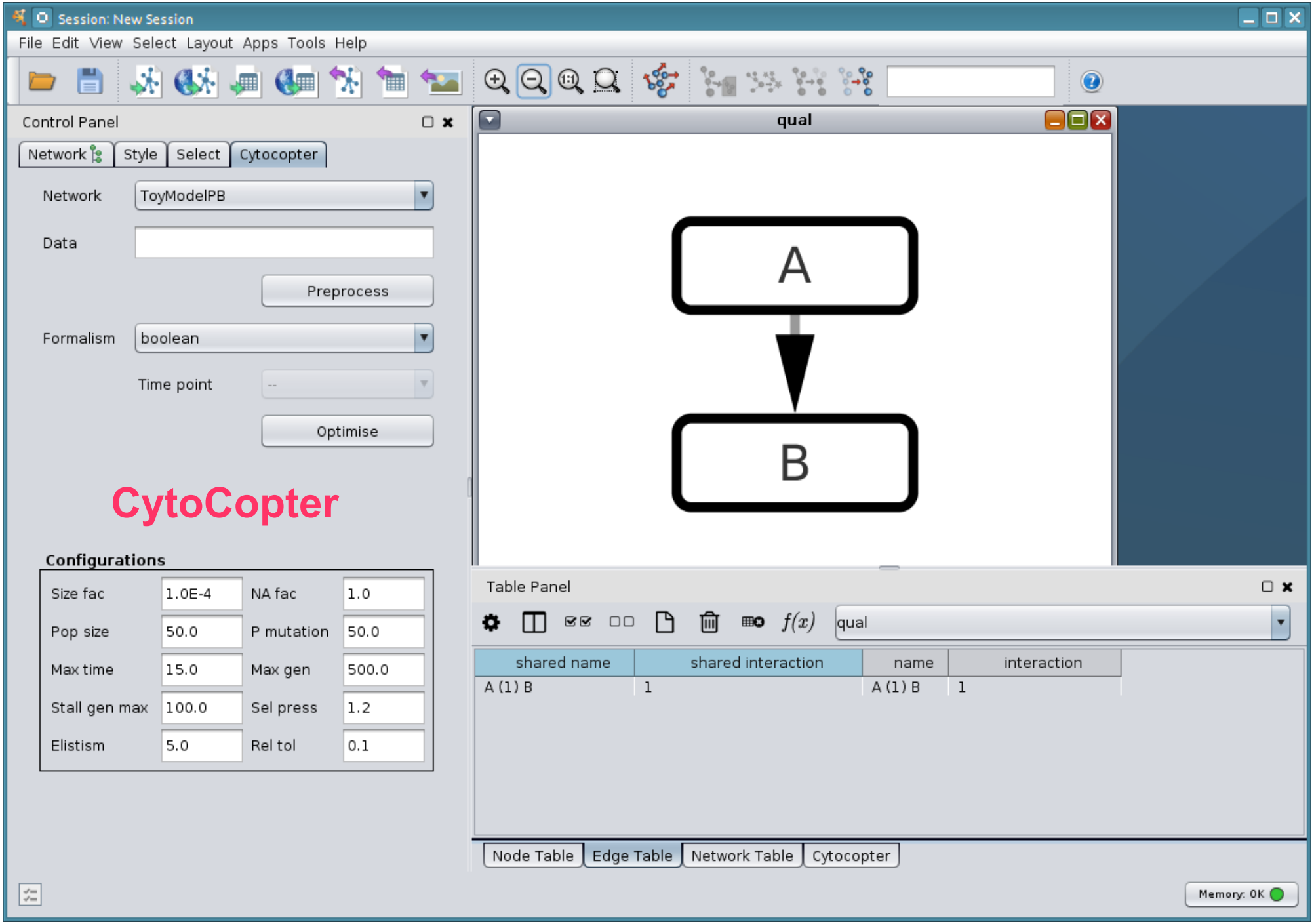
SBML Level 3 is modular



$$A \geqslant 1 \Rightarrow B = 1$$

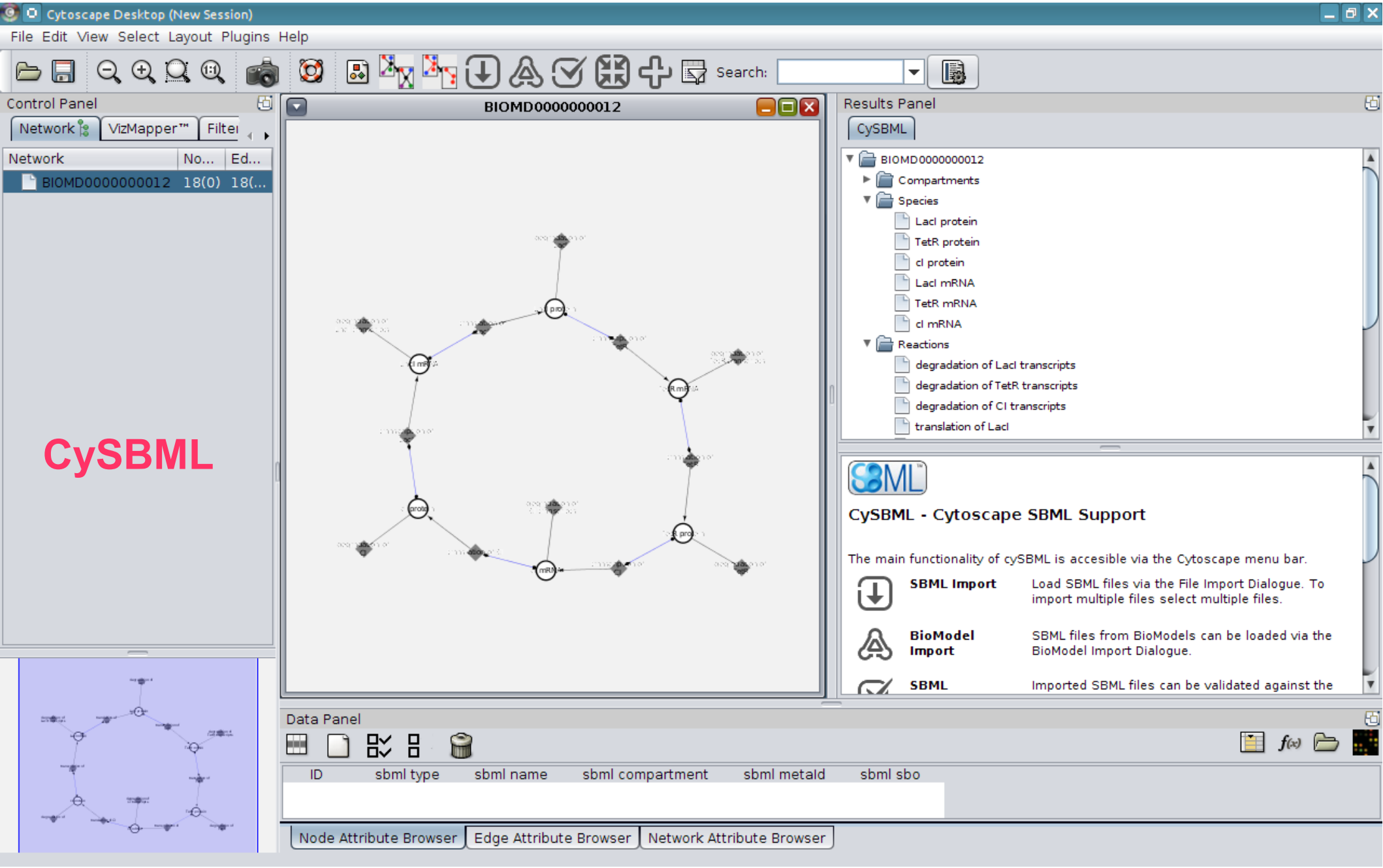
Logical model with SBML Qual

```
<?xmlversion="1.0" encoding="UTF8"?>
<sbml xmlns="http://www.sbml.org/sbml/level3/version1/core" level="3" version="1"
      xmlns:qual="http://www.sbml.org/sbml/level3/version1/qual/version1" qual:required="true">
  <model id="example">
    <listOfCompartments>
      <compartment id="cytosol" name="cytosol" constant="true"/>
    </listOfCompartments>
    <qual:listOfQualitativeSpecies>
      <qual:qualitativeSpecies qual:compartment="cytosol" qual:constant="false"
                               qual:id="A" qual:maxLevel="2"/>
      <qual:qualitativeSpecies qual:compartment="cytosol" qual:constant="false"
                               qual:id="B" qual:maxLevel="1"/>
    </qual:listOfQualitativeSpecies>
    <qual:listOfTransitions>
      <qual:transition qual:id="tr_B">
        <qual:listOfInputs>
          <qual:input qual:id="theta_B_A" qual:qualitativeSpecies="A" qual:sign="positive"
                      qual:thresholdLevel="1" qual:transitionEffect="none"/>
        </qual:listOfInputs>
        <qual:listOfOutputs>
          <qual:output qual:transitionEffect="assignmentLevel" qual:qualitativeSpecies="B"/>
        </qual:listOfOutputs>
        <qual:listOfFunctionTerms>
          <qual:functionTerm qual:resultLevel="1">
            <math xmlns="http://www.w3.org/1998/Math/MathML">
              <apply>
                <geq/>
                <ci>A</ci>
                <ci>theta_B_A</ci>
              </apply>
            </math>
          </qual:functionTerm>
          <qual:defaultTerm qual:resultLevel="0"/>
        </qual:listOfFunctionTerms>
      </qual:transition>
    </qual:listOfTransitions>
  </model>
</sbml>
```



Babraham Institute

EMBL-EBI


Cytoscape Desktop (New Session)
File Edit View Select Layout Plugins Help
Search:
Control Panel
Network
VizMapper™
Filter
Network
No...
Ed...
BIOMD0000000012 18(0) 18(...
BIOMD0000000012
Results Panel
CySBML
BIOMD0000000012
Compartments
Species
LacI protein
TetR protein
cI protein
LacI mRNA
TetR mRNA
cI mRNA
Reactions
degradation of LacI transcripts
degradation of TetR transcripts
degradation of CI transcripts
translation of LacI
CySBML - Cytoscape SBML Support
The main functionality of cySBML is accesible via the Cytoscape menu bar.
SBML Import
Load SBML files via the File Import Dialogue. To import multiple files select multiple files.
BioModel Import
SBML files from BioModels can be loaded via the BioModel Import Dialogue.
SBML
Imported SBML files can be validated against the
Data Panel
ID
sbml type
sbml name
sbml compartment
sbml metaId
sbml sbo
Node Attribute Browser
Edge Attribute Browser
Network Attribute Browser

CySBML

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

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

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

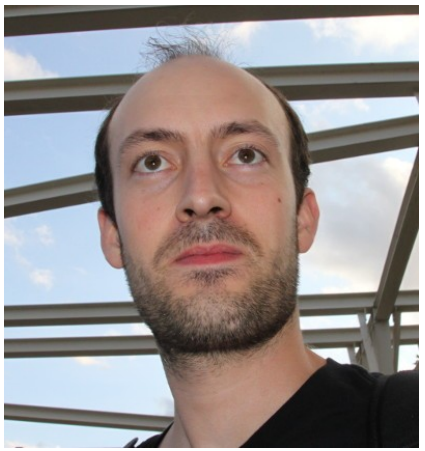
1. Annotation unambiguously model constituent to data
2. Link to external information as a triplet {collection, identifier, qualifier}
3. Annotation written as a Uniform Resource Identifier
4. Identifier considered within framework of the collection.
5. Collection namespace and record identifier in one 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

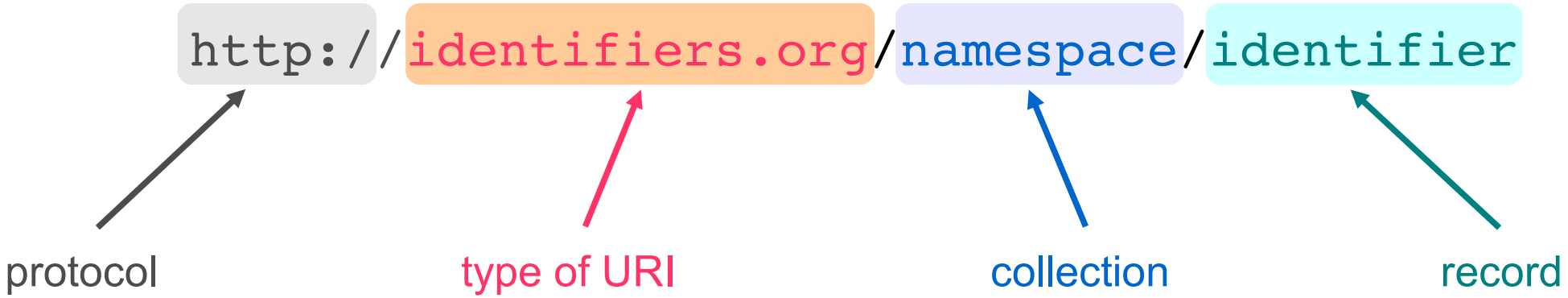


Camille Laibe

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

MIRIAM Registry

Search

Examples: ontology, enzyme, Japan, EMBL

Categories & tags

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

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>
```

Data collection: *Enzyme Nomenclature*

Overview | Miscellaneous | RDF/XML | Turtle

General information

Recommended name	**Enzyme Nomenclature**
Alternative name(s)	Enzyme Classification EC code EC
Description	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+$
Registry identifier	MIR:00000004

• protein • enzyme • classification • taxonomy

Identification schemes

Namespace	ec-code
Root URL	http://identifiers.org/ec-code/
Root URN	urn:miriam:ec-code:

Available as HTML or RDF, through browsers or Web Services

Other root URI(s) ⊞

Physical locations (resources)

Resource MIR:00100308	Description	ExploreEnz at Trinity College
	Access URL	http://www.enzyme-database.org/query.php?ec=$id [Example: 1.1.1.1]
	Institution	Trinity College, Dublin, Ireland
	Website	http://www.enzyme-database.org/
Resource MIR:00100002	Description	KEGG Ligand Database for Enzyme Nomenclature
	Access URL	http://www.genome.jp/dbget-bin/www_bget?ec:$id [Example: 1.1.1.1]
	Institution	Kyoto University Bioinformatics Center, Japan
	Website	http://www.genome.jp/dbget-bin/www_bfind?enzyme
Resource MIR:00100003	Description	Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
	Access URL	http://enzyme.expasy.org/EC/$id [Example: 1.1.1.1]
	Institution	Swiss Institute of Bioinformatics, Switzerland
	Website	http://enzyme.expasy.org/
Resource	Description	IntEnZ (Integrated relational Enzyme database)
	Access URL	http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=$id [Example: 1.1.1.1]

Alternative ways to access information for 1.1.1.1 (from Enzyme Nomenclature) can be found at http://info.identifiers.org/ec-code/1.1.1.1. Close

EMBL-EBI Services Research Training About us

IntEnz

Search

Examples: 1.1.1.1, alcohol%, alpha-glucosidase

Advanced

Home Browse Submit Download Documentation About Feedback

EC 1 - Oxidoreductases

EC 1.1 - Acting on the CH-OH Group of Donors

EC 1.1.1 - With NAD^+ or $NADP^+$ as acceptor

EC 1.1.1.1 - Alcohol dehydrogenase

IntEnz view | ENZYME view

XML

IntEnz Enzyme Nomenclature

EC 1.1.1.1

Names

Accepted name: alcohol dehydrogenase

Other names: ADH
NAD-dependent alcohol dehydrogenase
NAD-specific aromatic alcohol dehydrogenase
NADH-alcohol dehydrogenase
NADH-aldehyde dehydrogenase

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

Identifiers.org URI for identifying *1.1.1.1* from Enzyme Nomenclature.

Information about 1.1.1.1 from Enzyme Nomenclature can be accessed from any of the following locations:

KEGG Ligand Database for Enzyme Nomenclature
Kyoto University Bioinformatics Center

Japan

(Uptime: 100%)

IntEnZ (Integrated relational Enzyme database)
European Bioinformatics Institute

United Kingdom

(Uptime: 100%)

Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
Swiss Institute of Bioinformatics

Switzerland

(Uptime: 99%)

ExploreEnz at Trinity College
Trinity College, Dublin

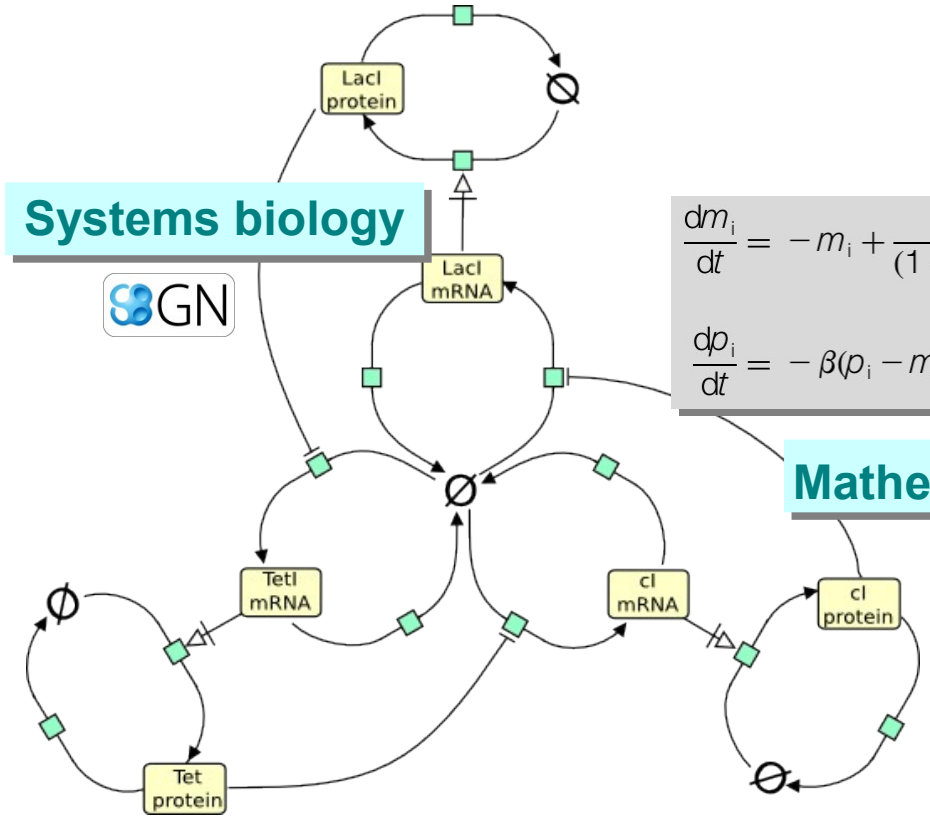
Ireland

(Uptime: 100%)

Powered by MIRIAM Registry

Information also available in: RDF/XML

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



Systems biology

SBGN

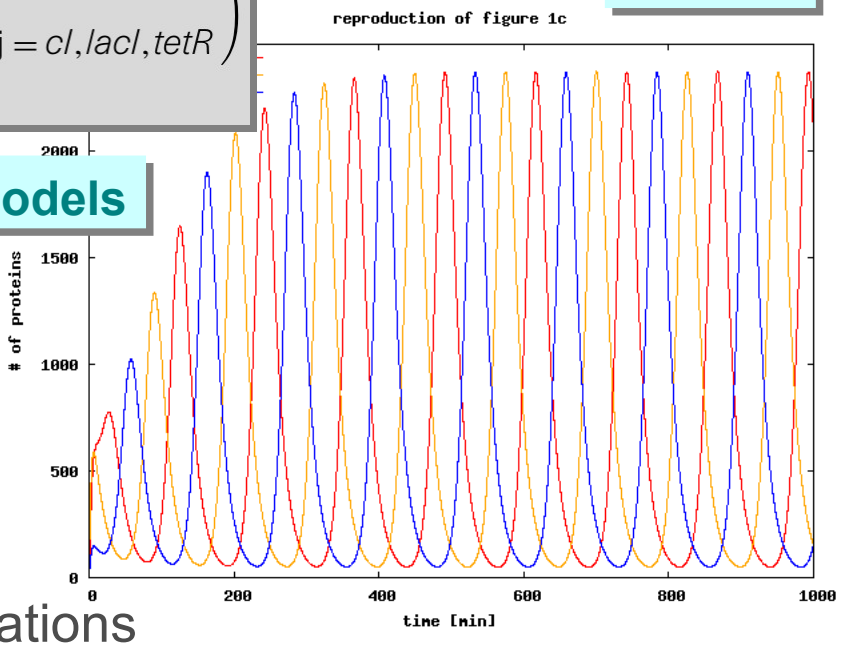
$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{(1+p_j^n)} + \alpha_0 \qquad \begin{pmatrix} i = lacI, tetR, cI \\ j = cI, lacI, tetR \end{pmatrix}$$

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

Mathematical models

SBML

curated



>140 000 models, >10 millions mathematical relations

Deposition advised by >300 journals, database ~1000 citations

1 million page requests per year

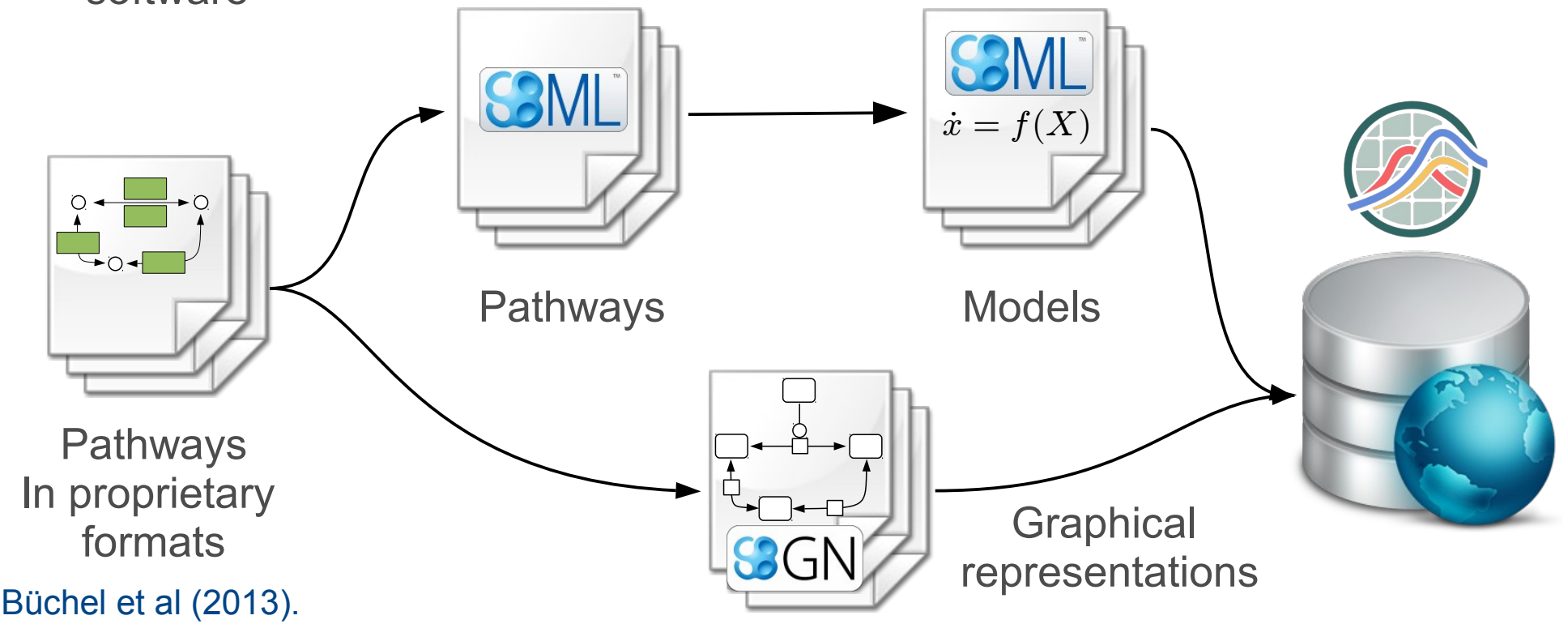
Submission in SBML and CellML; Export in SBML, CellML, XPP, SciLab, BioPAX, Octave, PDF, VCML, SBGN

Live demo ...

Was it worth it?

What interoperability can do: the 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

EMBL-EBI

Mihai Glont
Henning Hermjakob
Sarah Keating
Camille Laibe
Nicolas Rodriguez
Julio Saez-Rodriguez
Martijn van Iersel
Michael Schubert

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Andreas Dräger
Roland Keller
Matthias Rall
Clemens Wrzodek
Andreas Zell
Finja Bünel
Florian Mittag

MANCHESTER
1824

Douglas Kell
Pedro Mendes
Neil Swainston

INSTITUTO GULBENKIAN DE CIÊNCIA

Claudine Chaouyia

IPK
GATERSLEBEN

Falk Schreiber
Tobias Czauderna

Heidelberg Institute for
Theoretical Studies

HITS

Martin Golebiewski
Wolfgang Müller

MONASH University

Michael Wybrow

CALIFORNIA INSTITUTE OF TECHNOLOGY
1891

Michael Hucka

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

+

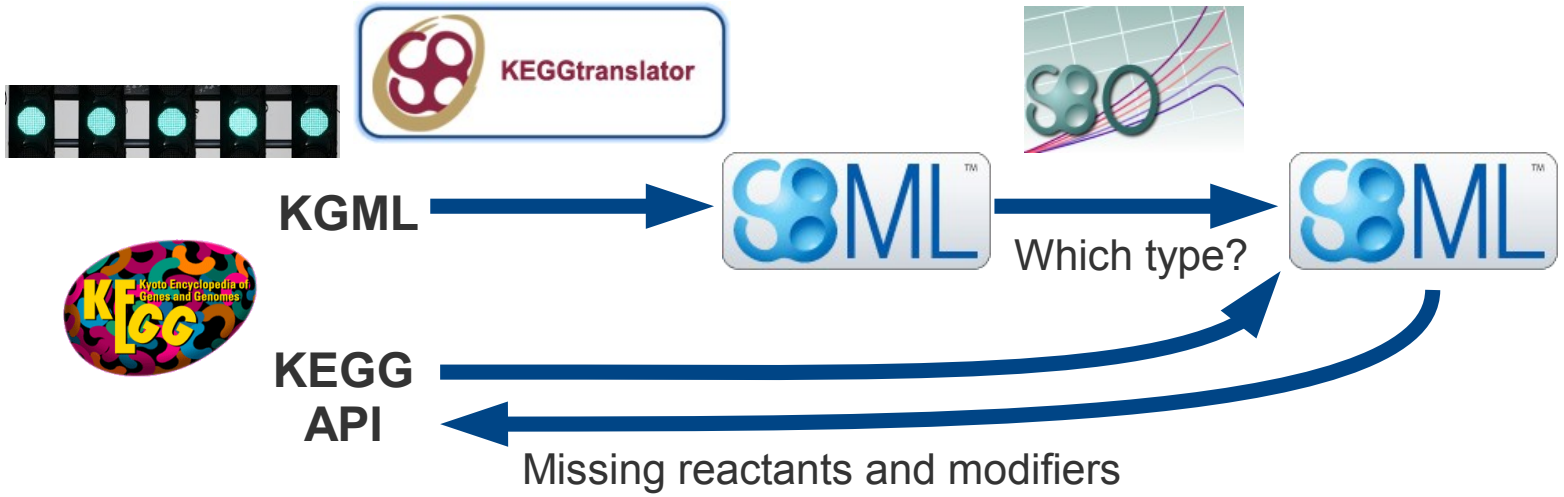
Flux Balance Analysis
of whole genome reconstructions

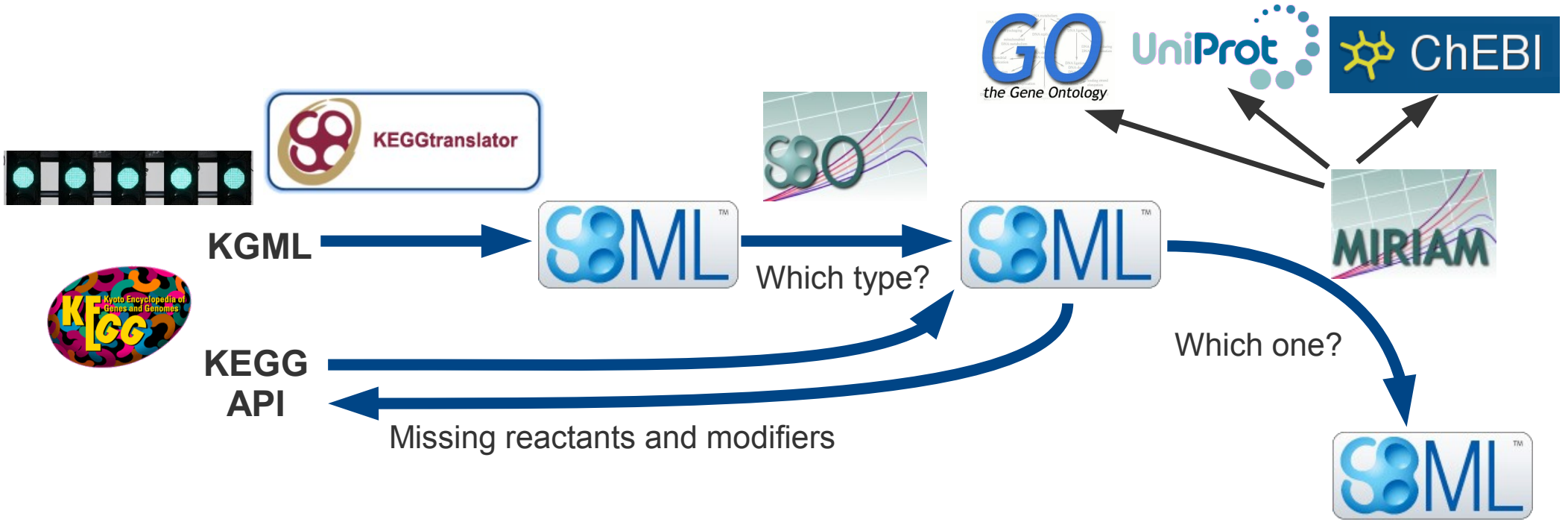
KEGGtranslator

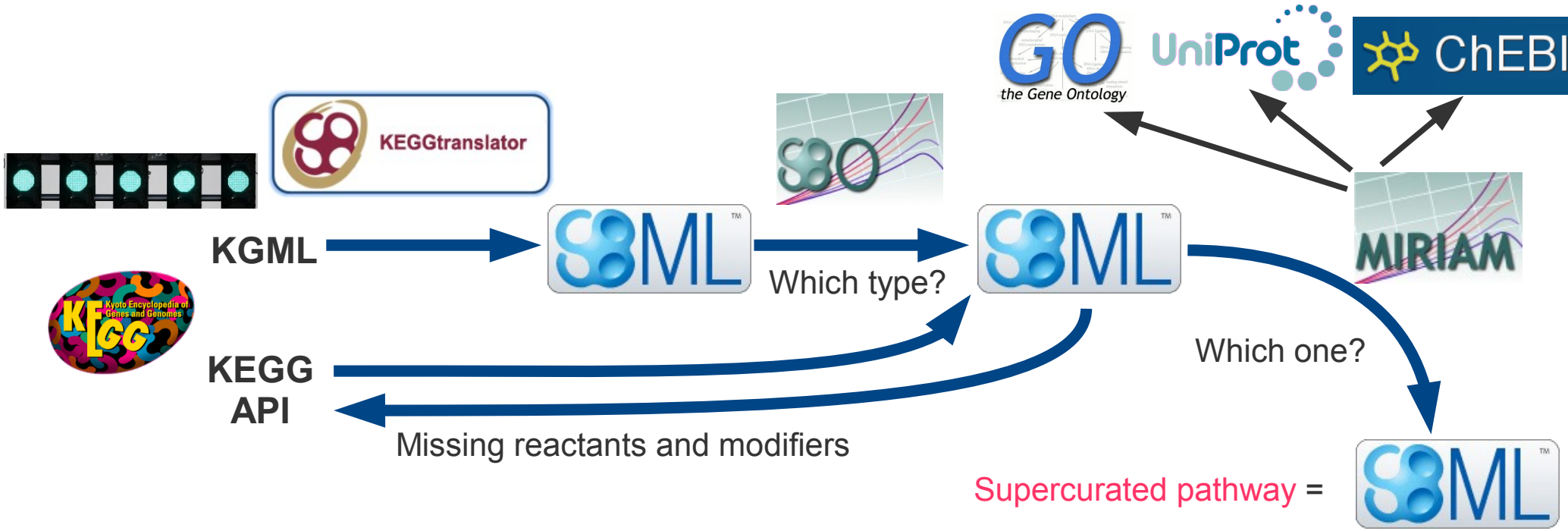
KGML → SBML

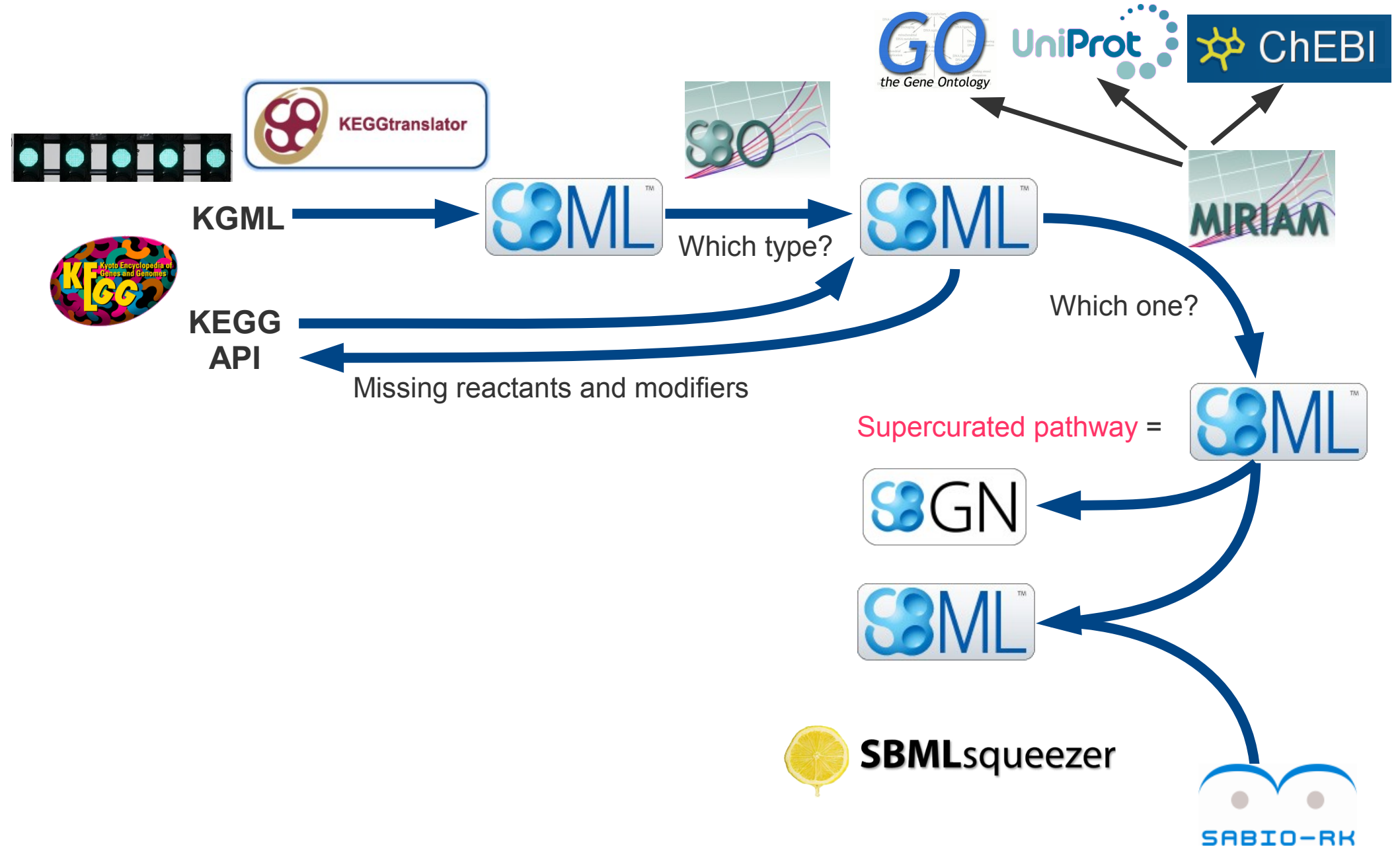
Core

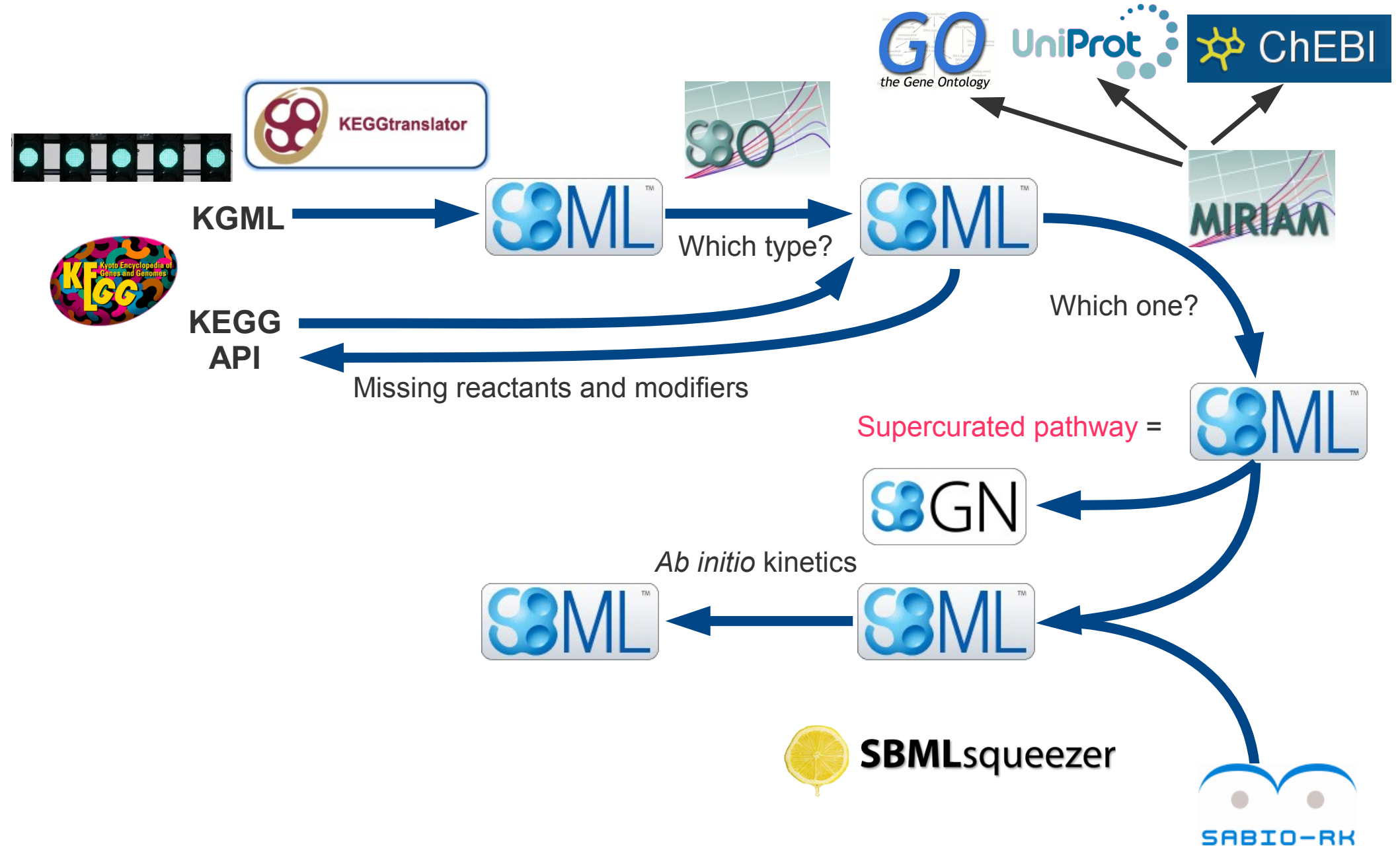

KEGGtranslator
KGML
SBML
Which type?
SBML
Kyoto Encyclopedia of Genes and Genomes

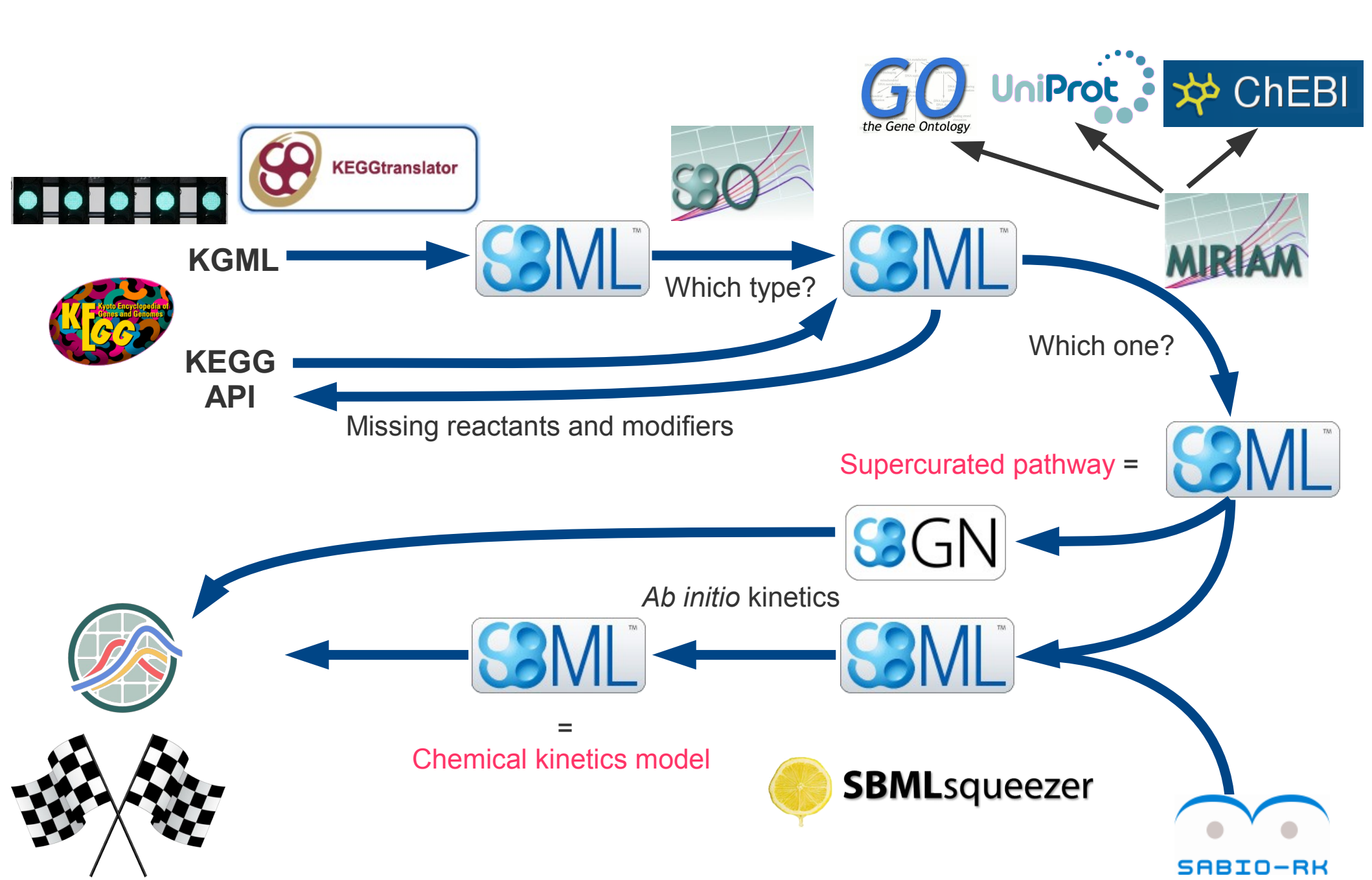

KEGGtranslator
KGML
Which type?
KEGG
API
Missing reactants and modifiers


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


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?
Supercurated pathway =
SBML
SBGN
SBML
SBMLsqueezer
SABIO-RK


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


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

I am grateful and indebted to

All the collaborators listed during the presentation

Editors of COMBINE standards

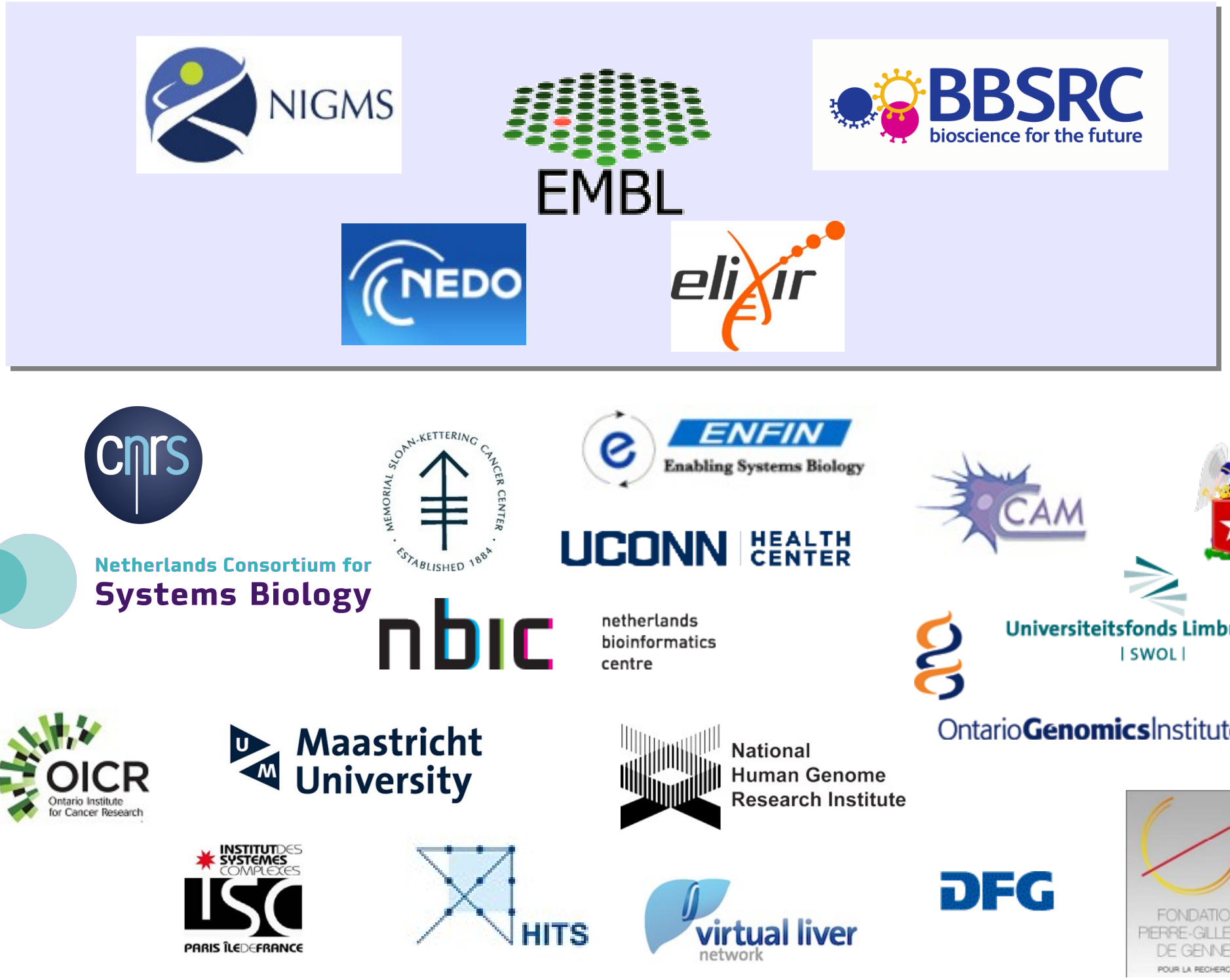
Developers of BioModels Database

Developers of related ontologies and software

Organisers of meetings and efforts to support our standards

The community of Computational Systems Biology

You for staying awake


NIGMS
EMBL
BBSRC
bioscience for the future
NEDO
elixir
cnrs
MEMORIAL SLOAN-KETTERING CANCER CENTER
ESTABLISHED 1884
ENFIN
Enabling Systems Biology
CAM
UCONN HEALTH CENTER
Netherlands Consortium for
Systems Biology
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