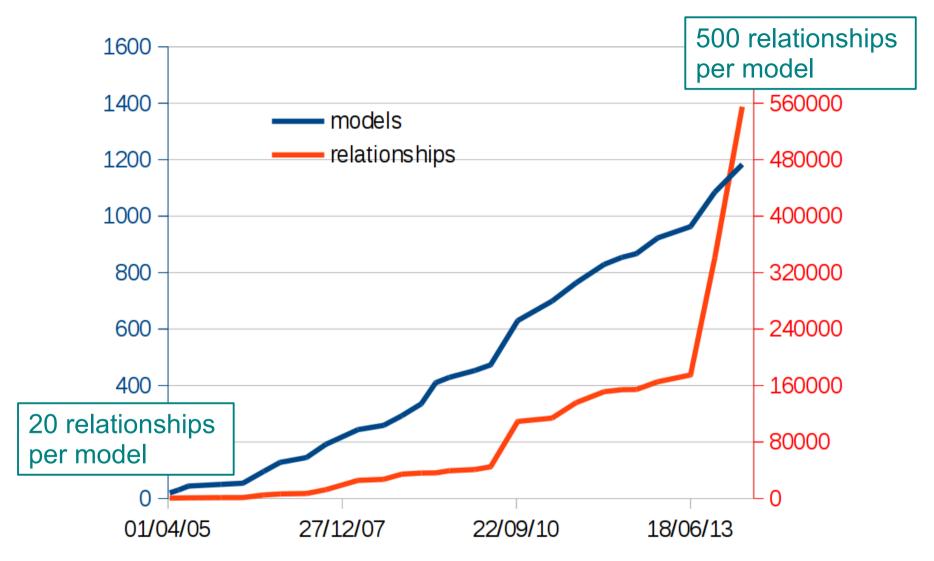




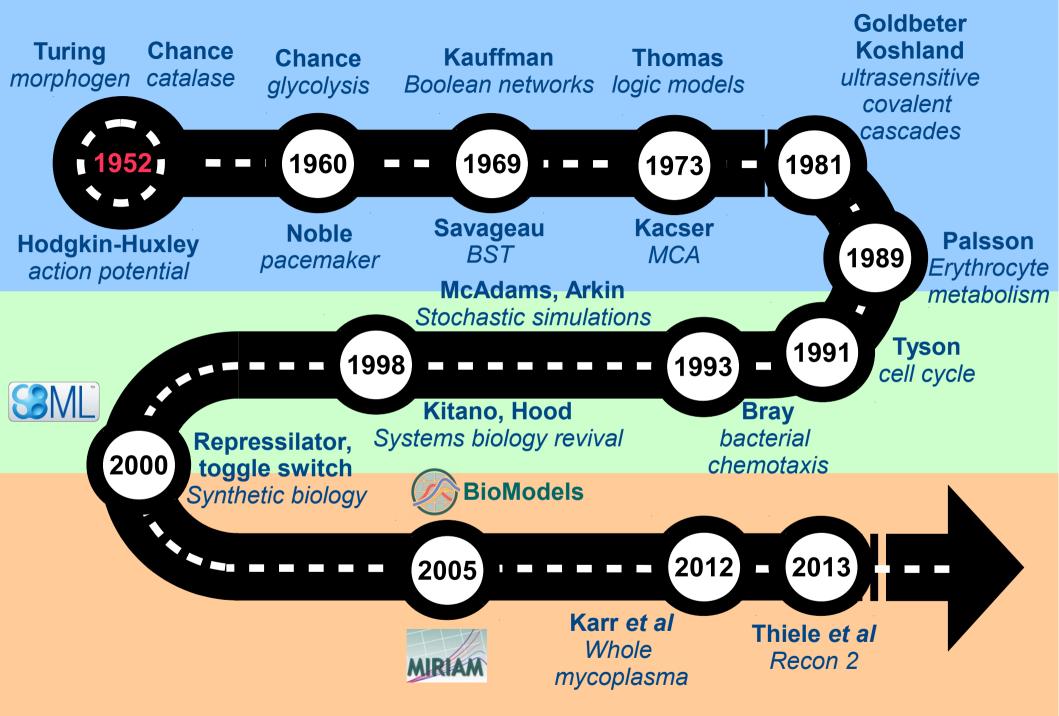
Computational models on the rise





BioModels Database growth (published models branch) since its creation











Improvised

Designed

One off

Many

Unique

Standard

Manually produced

Automated production

One or few artists

Collaboration

Produced in one go

Workflow

Fragile

Robust







We need to

Verify

Re-use

Modify

Build upon

Integrate with

Therefore we need to share

Model descriptions

Simulation descriptions

Parametrisations

Biological meaning



Three types of standards

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



http://co.mbine.org



HARMONY 2015

Standards

Events

Documents

About

Forums

Search

• Home • Help • Sign-in

Coordinating standards for modeling in biology

The 'COmputational Modeling in BIology' NEtwork (COMBINE) is an initiative to coordinate the development of the various community <u>standards and formats</u> for computational models. By doing so, it is expected that the federated projects will develop a set of interoperable and non-overlapping standards covering all aspects of modeling in biology.

Building on the experience of mature projects, which already have stable specifications, software support, user-base and community governance, COMBINE will help foster or support fledgling efforts aimed at filling gaps or new needs. As those efforts mature, they may become part of the <u>core set of COMBINE standards</u>.

One of the initial activities of COMBINE is to coordinate the organization of scientific and technical <u>events</u> common to several standards. Those events, as others related to our field of research are gathered in a calendar.

To receive announcements from COMBINE, subscribe to <u>combine-announce@ebi.ac.uk</u> (Note that the main list of each of the <u>COMBINE standards</u> is already subscriber).

To discuss the goals, organization and operation of COMBINE, subscribe to combine-discuss@ebi.ac.uk.

To report issues about the co.mbine.org website, send a mail to combine-support @ googlegroups.com



A language to describe computational models in biology

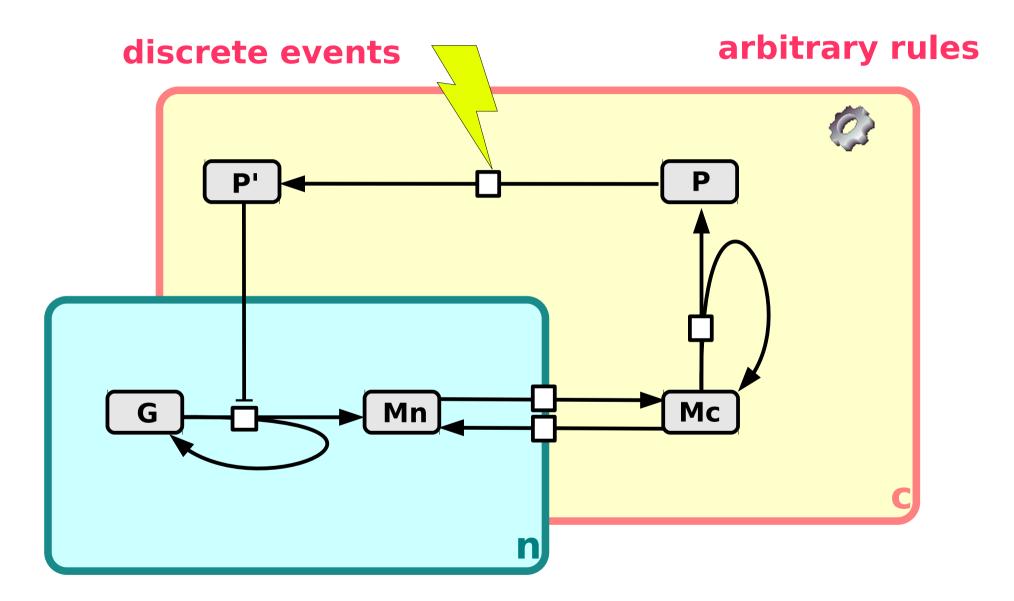
Born in Caltech 2000 Model descriptions John Hiroaki Doyle Kitano Data-models Mike Herbert Hamid Andrew Sauro Bolouri Hucka Finney

Hucka et al. Bioinformatics (2003)





What can we encode in SBML (core)?







Structure of SBML

```
<?xml version="1.0" encoding="UTF-8"?>
           <sbml level="3" version="1".</pre>
                 xmlns="http://www.sbml.org/sbml/level3/version1/core">
              <model>
                <listOfFunctionDefinitions> </-- --> </listOfFunctionDefinitions>
                <listOfUnitDefinitions> </-- --> </listOfUnitDefinitions>
                <list0fCompartments> </-- --> </list0fCompartments>
                <list0fSpecies> <!-- --> </list0fSpecies>
  variables
                <list0fParameters> </-- --> </list0fParameters>
                <list0fInitialAssignments> </-- --> </list0fInitialAssignments>
                t0fRules> </-- --> </list0fRules></-->
                <list0fConstraints> </-- --> </list0fConstraints>
relationships
                <listOfReactions> </-- --> </listOfReactions>
                <list0fEvents> </-- --> </list0fEvents>
              </model>
           </sbml>
```

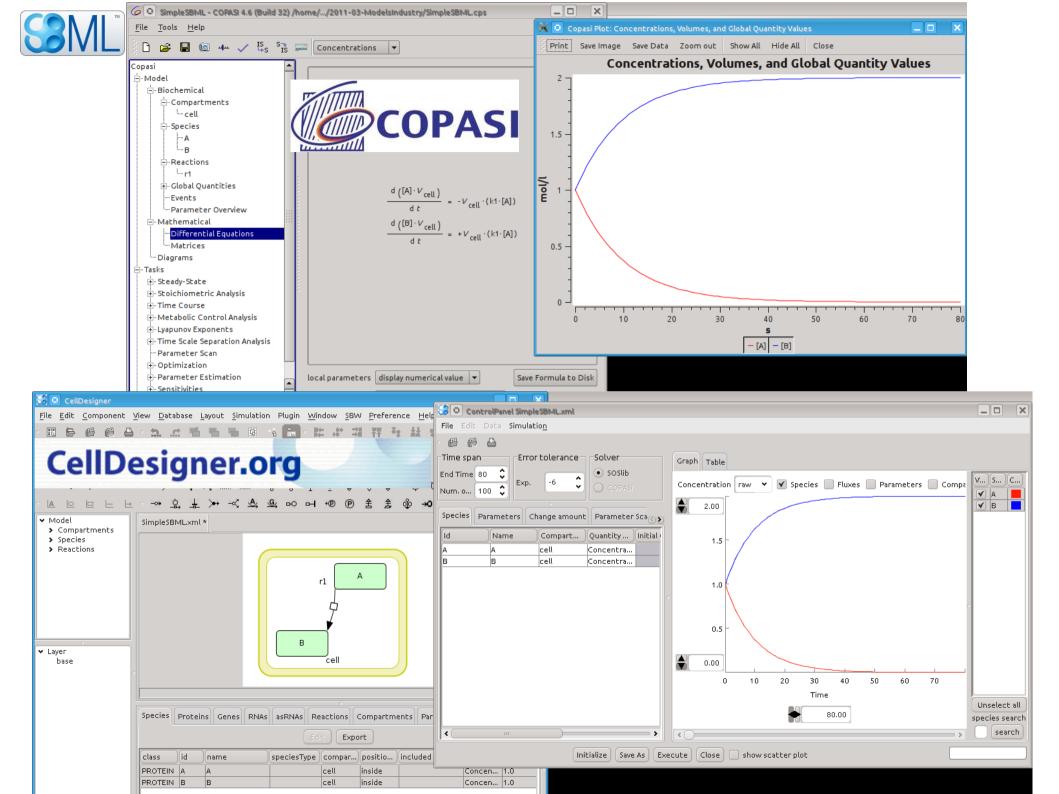




```
<?xml version="1.0" encoding="UTF-8"?>
      <sbml xmlns="http://www.sbml.org/sbml/level2/version4" level="2" version="4">
        <model name="Simple Model">
         <compartment id="cell" size="1" />
         <species id="A" compartment="cell" initialConcentration="1"/>
           <species id="B" compartment="cell" initialConcentration="1"/>
         </listOfSpecies>
         Α
           <parameter id="k1" value="0.1"/>
         </listOfParameters>
         IstOfReactions>
           <reaction id="r1" reversible="false">
                                                    A very simple
           IstOfReactants>
               <speciesReference species="A"/>
                                                        SBML file
             Ist0fProducts>
 В
               <speciesReference species="B"/>
             <kineticLaw>
               <math xmlns="http://www.w3.org/1998/Math/MathML">
                <apply>
                  <times/>
\frac{1}{2} = k1 \times [A]
                  <ci> cell </ci>
                  <ci> k1 </ci>
                  <ci> A </ci>
                </apply>
               </kineticLaw>
           </reaction>
         </listOfReactions>
        </model>
      </sbml>
```

http://sbml.org





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 Hoppe⁶, Swagatika Sahoo¹,
Maike K Aurich¹, Hulda Haraldsdottir¹, Monica L Mo⁷, Ottar Rolfsson¹, Miranda D Stobbe^{8,9},
Stefan G Thorleifsson¹, Rasmus Agren¹⁰, Christian Bölling⁶, Sergio Bordel¹⁰, Arvind K Chavali¹¹,
Paul Dobson¹², Warwick B Dunn^{3,13}, Lukas Endler¹⁴, David Hala¹⁵, Michael Hucka¹⁶, Duncan Hull⁴,
Daniel Jameson^{3,4}, Neema Jamshidi⁷, Jon J Jonsson⁵, Nick Juty¹⁷, Sarah Keating¹⁷, Intawat Nookaew¹⁰,
Nicolas Le Novère^{17,18}, Naglis Malys^{3,19,20}, Alexander Mazein²¹, Jason A Papin¹¹, Nathan D Price²²,
Evgeni Selkov, Sr²³, Martin I Sigurdsson¹, Evangelos Simeonidis^{22,24}, Nikolaus Sonnenschein²⁵, 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 Nielsen¹⁰,
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 exhibit¹. The metabolic-network reconstruction procedure

is now well-established² and has been applied to a growing number of model organisms³. 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 networkwide mechanistic investigations of the genotype-phenotype relationship. A high-quality reconstruction of the metabolic network is thus

A not so simple SBML file (Recon2)

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



MODEL1109130000



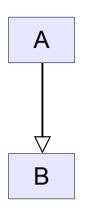




- Core package public specification
- Flux balance constraint public specification
- Qualitative models public specification
- Model composition public specification
- Graph Layout public specification
- Graph rendering specification finalised
- Complex species specification finalised
- Groups specification 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





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

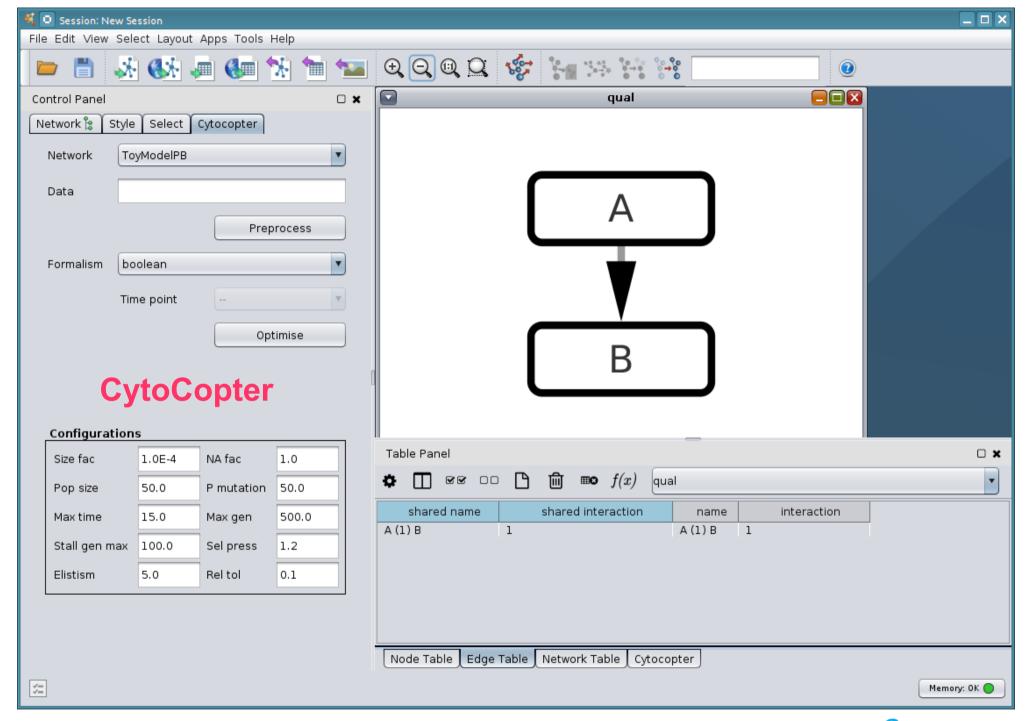
Logical model with SBML Qual

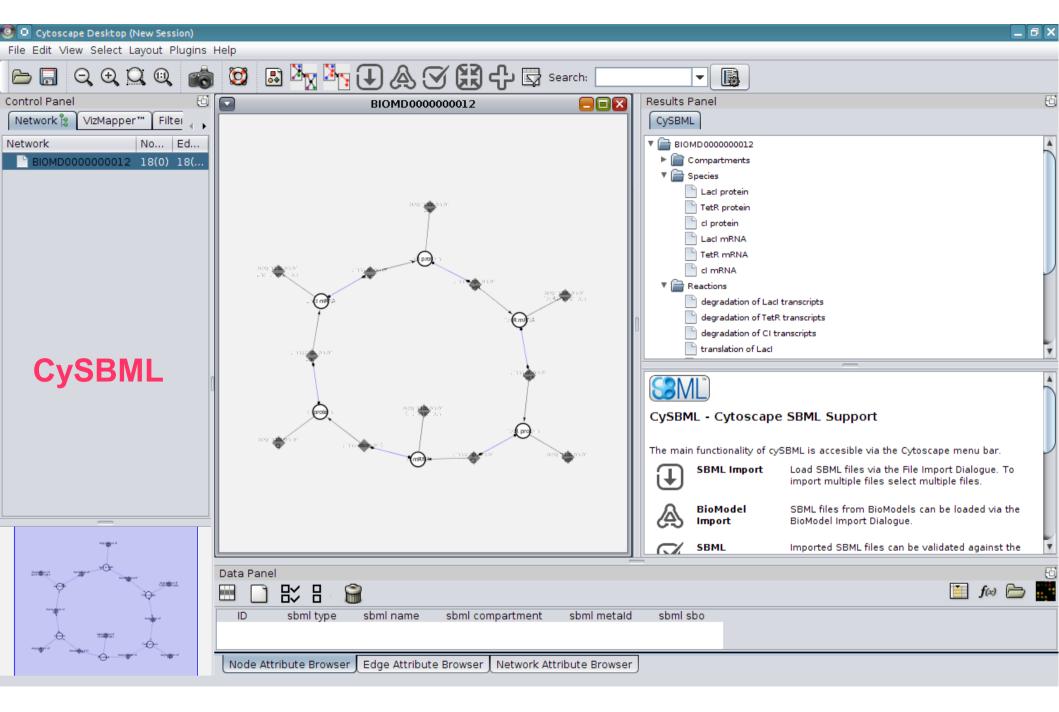
```
xmlns:qual="http://www.sbml.org/sbml/level3/version1/qual/version1" qual:required="true":
  <model id="example">
    <compartment id="cytosol" name="cytosol" constant="true"/>
    </l></l></l></l></l><
    <qual:list0fQualitativeSpecies>
      <qual:qualitativeSpecies qual:compartment="cytosol" qual:constant="false"</pre>
                               qual:id="A" qual:maxLevel="2"/>
      <qual:qualitativeSpecies qual:compartment="cytosol" qual:constant="false"</pre>
                               qual:id="B" qual:maxLevel="1"/>
    </qual:list0fQualitativeSpecies>
    <qual:listOfTransitions>
      <qual:transition qual:id="tr B">
        <qual:listOfInputs>
          <qual:input qual:id="theta B A" qual:qualitativeSpecies="A" qual:sign="positive"</pre>
                      qual:thresholdLevel="1" qual:transitionEffect="none"/>
        </gual:listOfInputs>
        <qual:list0f0utputs>
          <qual:output qual:transitionEffect="assignmentLevel" qual:qualitativeSpecies="B"/>
        </gual:list0f0utputs>
        <qual:listOfFunctionTerms>
          <qual:functionTerm qual:resultLevel="1">
            <math xmlns="http://www.w3.org/1998/Math/MathML">
              <apply>
                <qeq/>
                <ci>A</ci>
                <ci>theta B A</ci>
              </apply>
            </qual:functionTerm>
          <qual:defaultTerm qual:resultLevel="0"/>
        </gual:listOfFunctionTerms>
     </gual:transition>
    </gual:listOfTransitions>
  </model>
</sbml>
```

<sbml xmlns="http://www.sbml.org/sbml/level3/version1/core" level="3" version="1"</pre>



<?xmlversion="1.0" encoding="UTF8"?>







The Systems Biology Markup Language



👺 News Documents Downloads Forums Facilities Community Events About 📘 🔊



Q Google Site Search.

Parent pages: SBML.org

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: 263.

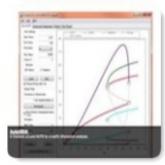
Go to the SBML Software Matrix



Go to the SBML Software Summary



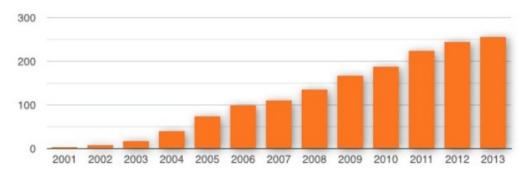
Go to the SBML Software Showcase



Please tell us about additions and updates.

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.



Adding the semantics to the syntax

Minimal requirements

Data-models

Model descriptions

Minimal managements

Born in Heidelberg 2004





Terminologies



Le Novère et al. Nat Biotechnol (2005), Courtot et al. Mol Syst Biol(2011)



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.
- Collection namespace and record identifier in one URI
- 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.
- Collection namespace and record identifier in one URI
- Qualifiers to refine the link between model constituent and external knowledge
- Standard set of valid URIs agreed upon by community

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



identifiers (aka new MIRIAM URIs)





Camille Laibe



Nick Juty

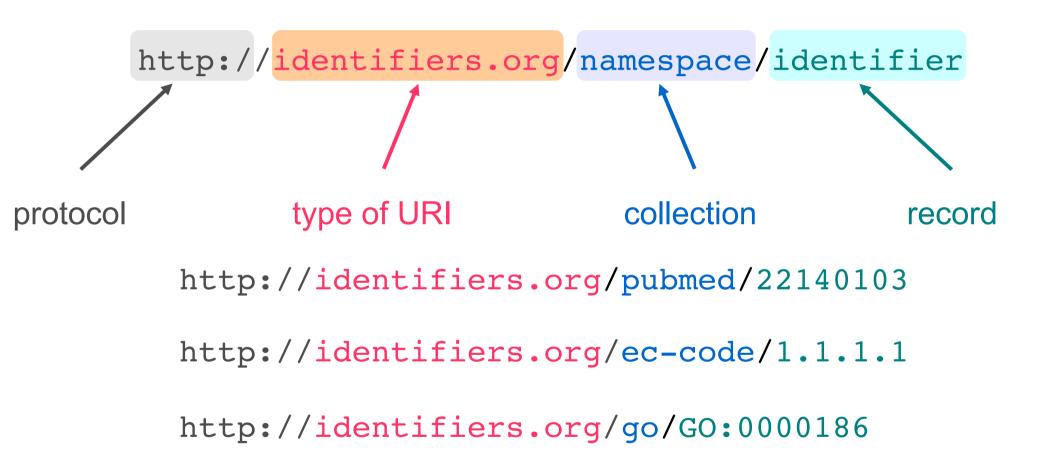


Sarala Wimalaratne

Juty et al. Nucleic Acid Res. (2012)



identifiers (aka new MIRIAM URIs)





MIRIAM Registry

Examples: ontology, enzyme, Japan, EMBL



Home

Browse

Download

Web services

Documentation

Contribute

Identifiers.org

About



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 Identifiers.org.

The core of the Registry is a catalogue of data collections (corresponding to controlled vocabularies, databases, ...), their URIs and the corresponding physical URLs (or resources). These resources are monitored daily to ensure data accessibility and the validity of the resolution mechanism.

Access to the Registry's dataset is made available via exports (XML and RDF) and Web Services (SOAP and REST).

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

Registry statistics

Published

Data collections: 521 (531) Resources: 651 (704) Last update: Aug 12, 2014

Under curation

Data collections: 409 415 Resources: Last update: Jun 30, 2014





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"

Hackathon 2013 Symposium in an (slides, PDF)

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):

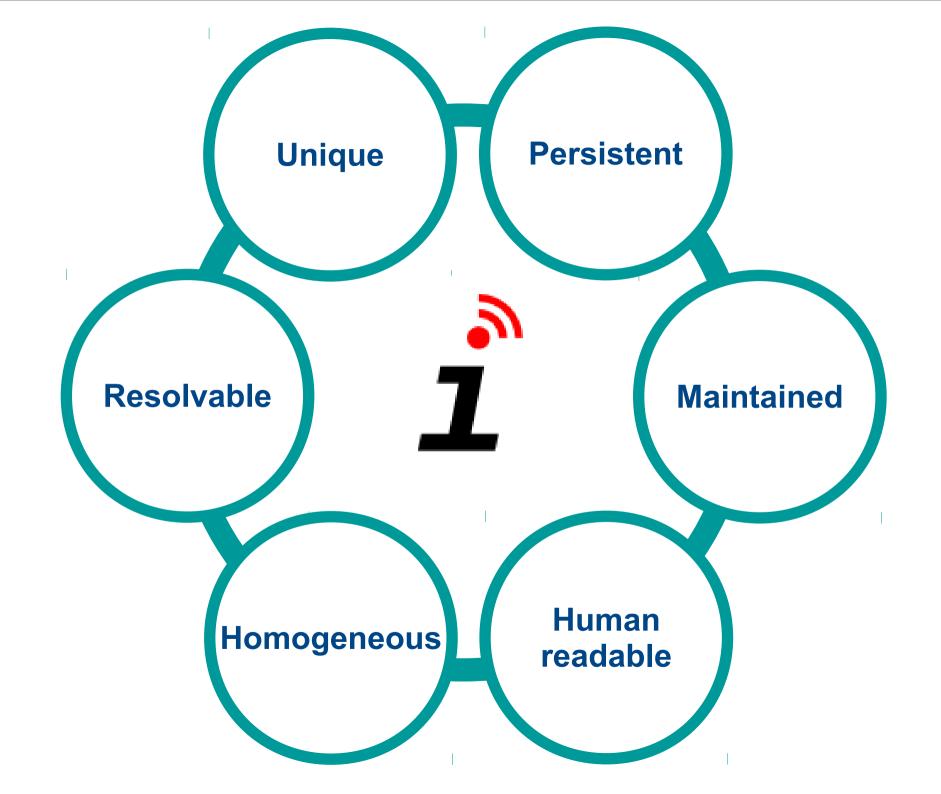
Laibe et al. BMC Syst Biol (2007)

http://www.ebi.ac.uk/miriam/ http://identifiers.org/registry

About the Reaistry









The Systems Biology Graphical Notation



http://sbgn.org/

Le Novère et al (2009





Unambiguous consensual visual notation

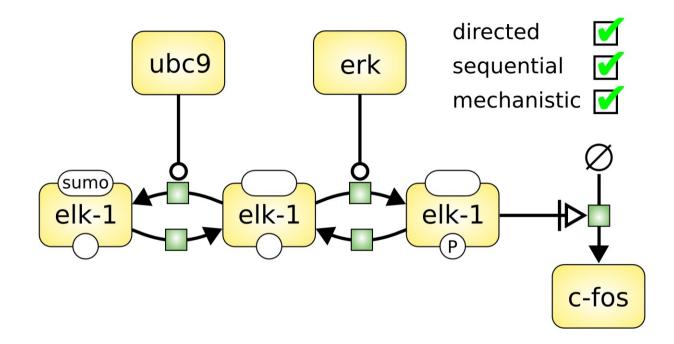
- An unambiguous way of graphically describing and interpreting biochemical and cellular events
- Limited amount of symbols
 Re-use existing symbols

Smooth learning curve

- Can represent logical or mechanistic models, biochemical pathways, at different levels of granularity
- Detailed technical specification, precise data-models, standard API and growing software support
- Developed over ten years by a diverse community, including biologists, modellers, computer scientists etc.



Process Descriptions



- Process modelling
- Biochemistry, Metabolic networks
- Generally within "closed world"
- Subjected to combinatorial explosion



Entity Relationships

ubc9

sequential mechanistic

elk-1

S383

erk

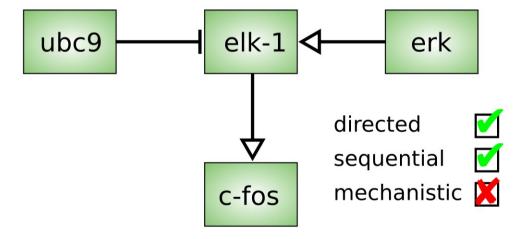
directed
sequential
mechanistic

c-fos

- Rule-based modelling
- Molecular Biology
- "Open world"
- Independent rules: no explosion

Activity Flows

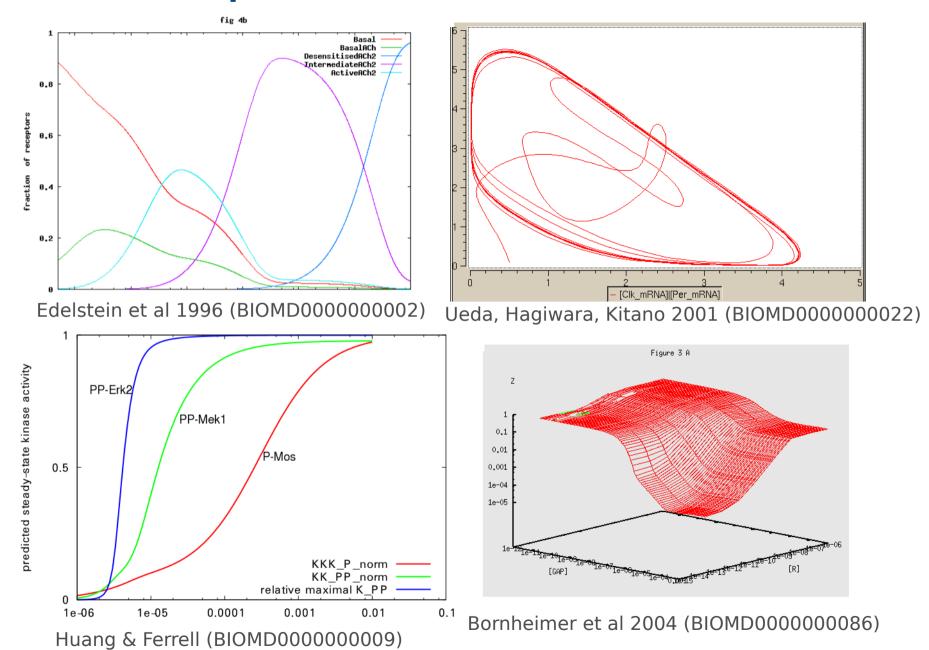
- Logical modelling
- Signalling pathways, gene regulatory networks



Surely, this is enough?

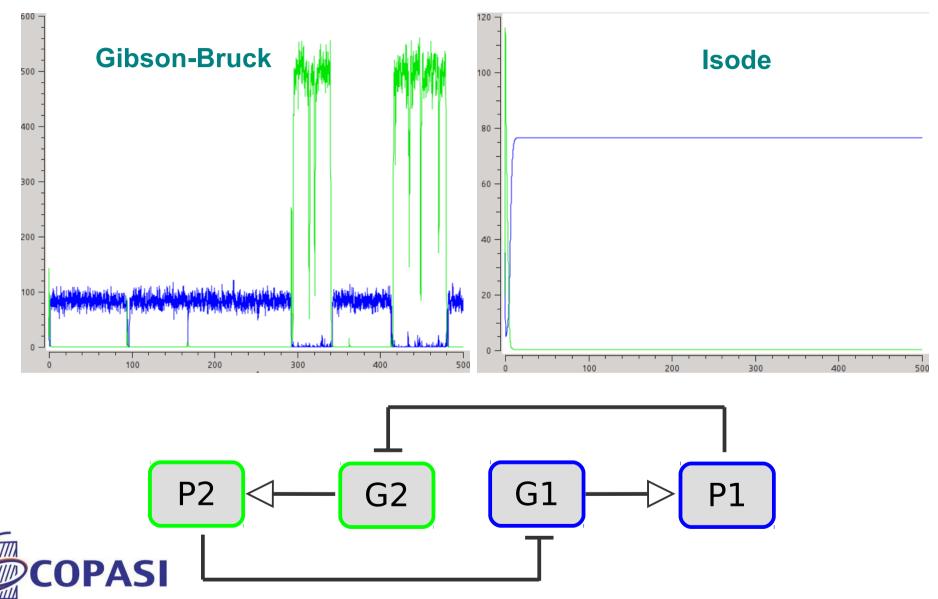


Simulation experiment = model + what to do with it





Choice of algorithm affects behaviour



Description of simulations and analyses

	Model descriptions	Simulations and analysis
Minimal requirements	MIRIAM	MIASE
Data-models	SIML SIGN	SED ML
Terminologies	S30	KISAO



Dagmar Waltemath



Anna Zhukova

Born in Hinxton 2007

Waltemath et al. PloS Comput Biol (2011), BMC Syst Biol (2011), Courtot et al. Mol Syst Biol (2011)



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/"</pre>
      xmlns:math="http://www.w3.org/1998/Math/MathML"
      level="1" version="1">
  <listOfSimulations><!-- --> </listOfSimulations>
  st0fModels>
    <model id="" source="">
      t0fChanges></-- --></list0fChanges></-->
   </model>
  </listOfModels>
  t0fTasks></-- --></list0fTasks></-- -->
  <listOfDataGenerators></-- --></listOfDataGenerators>
  <plot2D />
   <plot3D />
   <report />
  </list0f0utputs>
</sedML>
```



http://sed-ml.org



Flexible model use in SED-ML

```
Any XML
st OfHodels>
 <model id="modell"
        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="modell">

    Modifications before simulations

   <changeAttribute target=</pre>
          "/sbml/model/listOfParameters/parameter[@id='c']/@value" newValue="-50">
     </changeAttribute>
     <changeAttribute target=</pre>
          "/sbml/model/listOfParameters/parameter[@id='d']/@value" newValue="42">
     </changeAttribute>
   </model>
</listOfModels>
```





Etc.

	Model descriptions	Simulations and analysis	Numerical
Minimal requirements	MIRIAM	MIASE	
Data-models	SIML SIGN	SEDML	Christian Knüpfer NuML
Terminologies	S30	KISAO	TEDDY



That looks very useful. Where can I find those?



Search Advanced

BioModels Home Models Submit Support About BioModels Contact us

BioModels Database is a repository of computational models of biological processes. Models described from literature are manually curated and enriched with cross-references. All models are provided in the Public Domain. More information about BioModels Database can be found in the FAQ.

Models published in the literature

Browse



(530 models)



Non curated (655 models) Alternative access



Gene Ontology classification



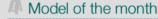
Gene Ontology tree



Advanced

search

Marco Donizelli



August, 2014

A mathematical model describing the molecular mechanisms involved in AD and the effect of immunisation against AB on soluble AB,





Models automatically generated from pathway resources (Path2Models)

Browse

Alternative access



Acknowledgements:

Metabolic (112,898 models) Non-metabolic (27,531 models) Whole genome metabolism (2,641 models)



Taxonomy

Viji Chelliah



25 April 2014 Our recent diabetes review most

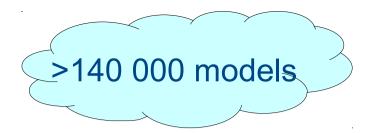


http://www.ebi.ac.uk/biomodels

EMBL

Nature Publishing Group (PSPod) is now online.

Le Novère et al. Nucleic Acid Res (2006), Li et al BMC Syst Biol (2010), Juty et al. Nucleic Acid Res (2015)



>10 millions math relations

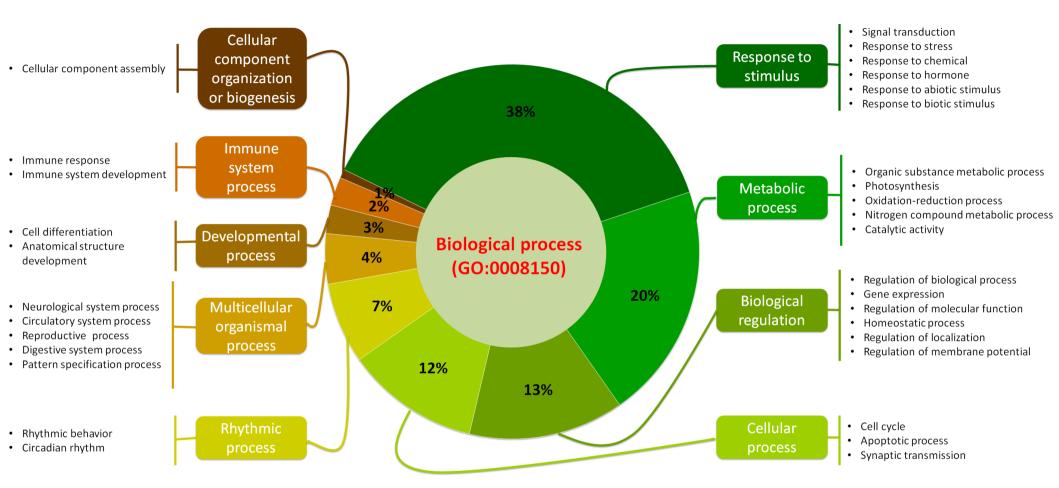
~ 200 millions cross-refs

~ 1000 citations

> 300 journals advise deposition

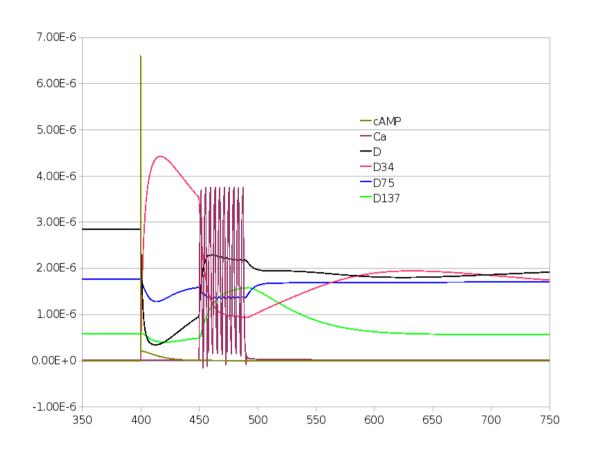
1 million page requests per year







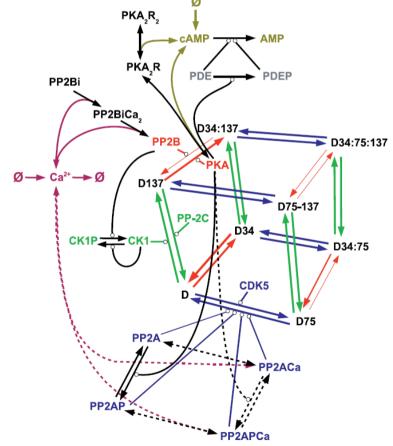
Biochemical models



reaction: $v_{\rm on} = k_{\rm on} \times [{\rm D}] \times [{\rm CDK5}] \times {\rm Vol}$

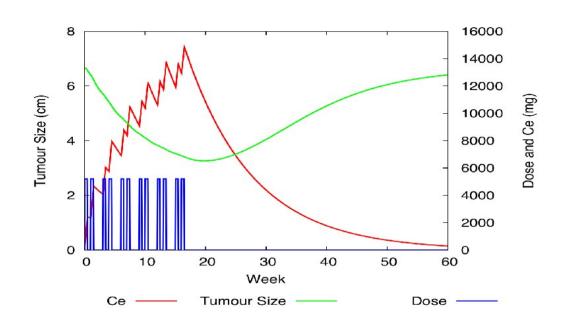
Fernandez et al. DARPP-32 is a robust integrator of dopamine and glutamate signals *PLoS Comput Biol* (2006) 2: e176.







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.



rate rule:

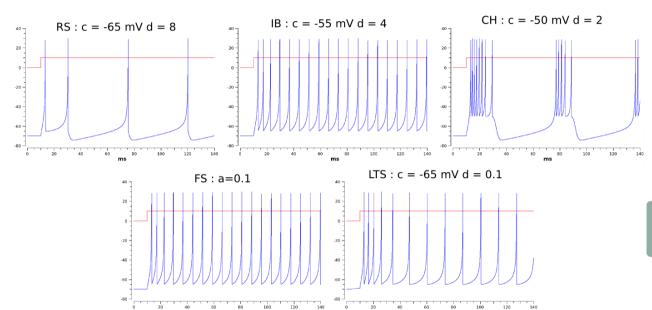
$$\frac{dSize}{dt} = (Rate_{in} \times Effect - K_{over} \times Size) \times Size$$

assignment rule:

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



Neuroscience models



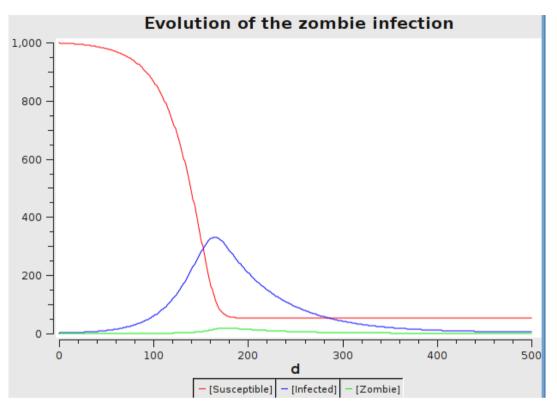
Izhikevich EM. Simple model of spiking neurons. *IEEE Trans Neural Netw* (2003) 14(6):1569-1572.



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



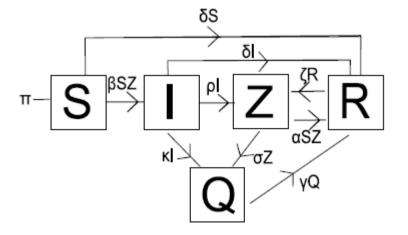
Epidemiology models





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







Was it worth it?



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

Biomathematician, 2007

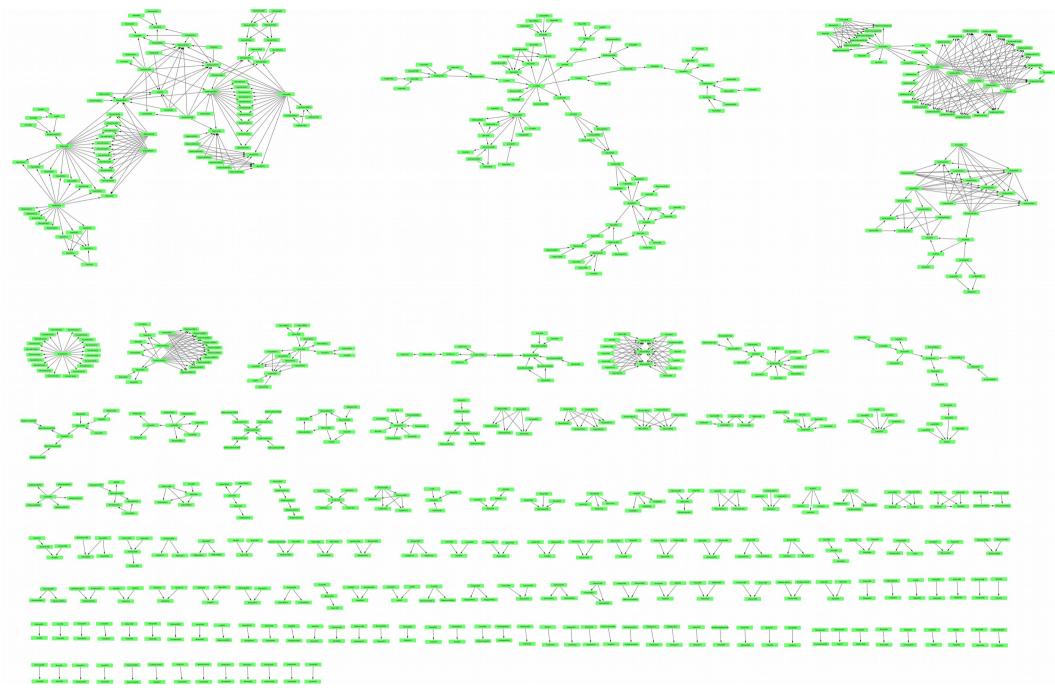
"By developing BioModels you harmed the cause of modelling in biology. My students do not learn how to make a model any more, instead, they download it ready to use."

Theoretical biologist, 2006



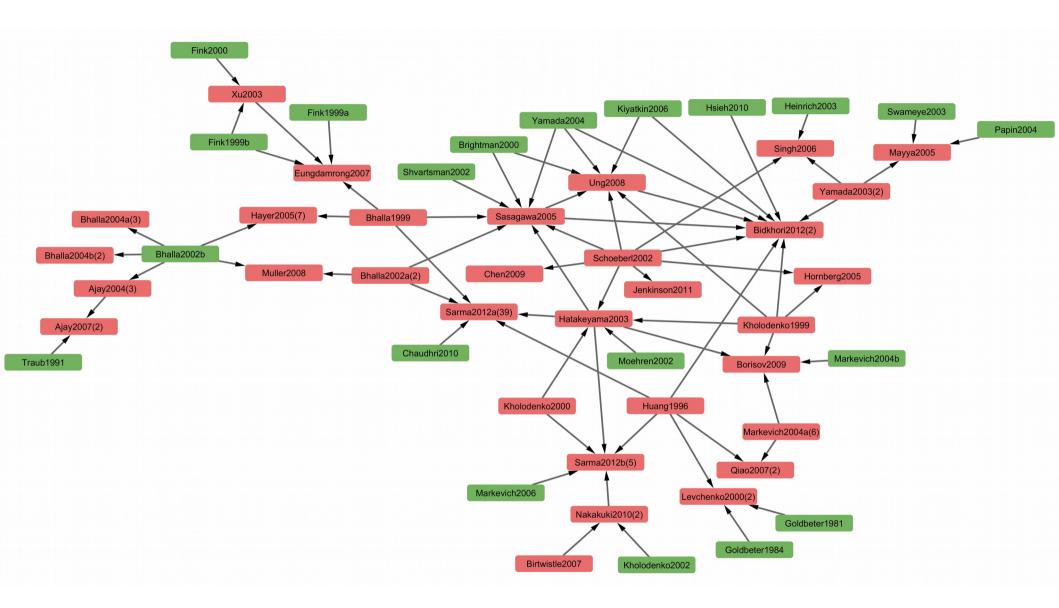
REUSE







Erb receptor signalling





DISCOVER



Clustering models (and data) based on metadata

ATP:protein_phosphotransferase_(non-specific)
RAF_proto-oncogene_serine/threonine-protein_kinase
inactivation_of_MAPKKK_activity
inactivation_of_MAPKK_activity
protein_amino_acid_dephosphorylation
protein_amino_acid_phosphorylation
MAP_kinase_kinase_kinase_kinase_activity
MAP_kinase_kinase_kinase_activity
activation_of_MAPKKK_activity
activation_of_MAPKK_activity
Ras_small_GTPase_Ras_type
mitogen-activated_protein_kinase_kinase_kinase_binding
urn:miriam:reactome:REACT_143
urn:miriam:reactome:REACT_996
urn:miriam:reactome:REACT_525
Mitogen-activated_protein_kinase_mos
urn:miriam:reactome:REACT_525
Mitogen-activated_protein_kinase_1
ATP:protein_phosphotransferase_(MAPKKK-activated)
MAP_kinase_kinase_activity
activation_of_MAPK_activity
Dual_specificity_mitogen-activated_protein_kinase_kinase_1
urn:miriam:reactome:REACT_136
urn:miriam:reactome:REACT_136
urn:miriam:reactome:REACT_2247
urn:miriam:reactome:REACT_2247
urn:miriam:uniprot:Q90W58
phosphoprotein_phosphatase_activity
mitogen-activated_protein_kinase_binding
urn:miriam:reactome:REACT_1780
urn:miriam:reactome:REACT_1780
urn:miriam:reactome:REACT_1495
peptidyl-threonine_phosphorylation
peptidyl-tyrosine_phosphorylation

Schulz et al. Mol Syst Biol (2011)



Ranking and retrieval of models

Model	BioModel	Similarity	<i>P</i> -value	Overlap	<i>P</i> -value	
Huang1996_MAPK_ultrasens	BIOMD00000000009	1.000	<=1e-3	30	0.0e+00	
Levchenko2000_MAPK_noScaffold	BIOMD0000000011	0.930		28	0.0e+00	
Levchenko2000_MAPK_Scaffold	BIOMD000000014	0.874	<=1e-3	26	0.0e+00	
Kholodenko2000_MAPK_feedback	BIOMD0000000010	0.830	<=1e-3	20	0.0e+00	
Markevich2004_MAPK_orderedElementary	BIOMD0000000026	0.749	<=1e-3	16	2.9e-15	
Markevich2004_MAPK_phosphoRandomElementary	BIOMD0000000028	0.692	<=1e-3	15	9.1e-14	
Markevich2004_MAPK_AllRandomElementary	BIOMD0000000030	0.692	<=1e-3	15	9.1e-14	
Markevich2004_MAPK_orderedMM	BIOMD0000000027	0.691	<=1e-3	12	9.8e-10	MAPK
Markevich2004_MAPK_orderedMM2kinases	BIOMD0000000031	0.691	<=1e-3	12	9.8e-10	
Markevich2004_MAPK_phosphoRandomMM	BIOMD0000000029	0.626	<=1e-3	11	1.6e-08	
Hornberg2005_ERKcascade	BIOMD0000000084	0.523	<=1e-3	9	2.7e-06	
McClean2007_CrossTalk	BIOMD0000000116	0.453	<=1e-3	8	2.7e-05	
Kofahl2004 pheromone	BIOMD0000000032	0.441	<=1e-3	12	9.8e-10	
Goldbeter1991 MinMitOscil Explinact	BIOMD00000000004	0.389	<=1e-3	3	1.5e-01	
Brown2004_NGF_EGF_signaling	BIOMD0000000033	0.371	<=1e-3	9	2.7e-06	
Ung2008_EGFR_Endocytosis	BIOMD0000000205	0.363	<=1e-3	8	2.7e-05	
Kim2007 Mnt EDK Crosstelk	RIOMD0000000140	U 3EE	∠ −10.3	10	2.20.07	
Goldheter1991 MinMitOscil	RIOMD0000000003	0.349	<=1e-3	3	1 5e-01	
Sasagawa2005 MAPK	BIOMD0000000049	0.339	<=1e-3	9	2.7e-06	
Swat2004_Mammalian_G1_S_Transition	BIOMD0000000228	0.317	<=1e-3	2	4.0e-01	
Tyson1991_CellCycle_6var	BIOMD0000000005	0.304	<=1e-3	4	4.2e-02	
Goldbeter1995_CircClock	BIOMD0000000016	0.274	<=1e-3	4	4.2e-02	
Novak1997_CellCycle	BIOMD0000000007	0.259	<=1e-3	1	7.6e-01	
Novak2001_FissionYeast_CellCycle	BIOMD0000000111	0.255	<=1e-3	2	4.0e-01	
Leloup1999_CircClock	BIOMD0000000021	0.246	<=1e-3	4	4.2e-02	
Birtwistle2007_ErbB_Signalling	BIOMD0000000175	0.236	<=1e-3	1	7.6e-01	
Neves2008_Cell_Shape	BIOMD000000182	0.222	<=1e-3	6	1.6e-03	
Leloup1998_CircClock_LD	BIOMD0000000171	0.219	<=1e-3	4	4.2e-02	
Veening2008_DegU_Regulation	BIOMD0000000240	0.211	4.0e-03	2	4.0e-01	
Chen2004_CellCycle	BIOMD000000056	0.209	4.0e-03	4	4.2e-02	
Fernandez2006_ModelA	BIOMD000000152	0.207	5.0e-03	2	4.0e-01	
Eicher2006_NEAT_Activation	RIOMD000000123	0.100	7.00.03	2	4.00-01	
Hatakevama2003 MAPK	BIOMD0000000146	0.198	7 0e-03	1	7 6e-01	



Retrieval of models using gene expression

Model	BioModel	Similarity	<i>P</i> -value	Overlap	P-value
Wolf2001 respiratory oscillations	BIOMD0000000090	0.207	<=1e-3	6	6.6e-09
Chassagnole2001_Threonine Synthesis	BIOMD0000000066	0.184	<=1e-3	4	1.5e-05
Curien2009_Aspartate_Metabolism	BIOMD0000000212	0.170	<=1e-3	5	3.6e-07
Curien2003_MetThr_synthesis	BIOMD0000000068	0.141	<=1e-3	2	1.0e-02
Proctor2007_ubiquitine	BIOMD0000000105	0.098	2.0e-03	1	1.4e-01
Curto1998_purineMetabol	BIOMD0000000015	0.063	1.1e-02	2	1.0e-02
Ibrahim2008_Spindle_Assembly_Checkpoint_dissociation	BIOMD000000186	0.057	1.8e-02	0	1.0e+00
Ibrahim2008_Spindle_Assembly_Checkpoint_convey	BIOMD000000187	0.057	1.8e-02	0	1.0e+00
Rodriguez-Caso2006_Polyamine_Metabolism	BIOMD000000190	0.040	7.1e-02	1	1.4e-01
Nijhout2004_Folate_Cycle	BIOMD0000000213	0.032	1.1e-01	1	1.4e-01
Morrison1989_FolateCycle	BIOMD000000018	0.030	1.3e-01	1	1.4e-01
Zatorsky2006_p53_Model3	BIOMD000000154	0.023	2.5e-01	0	1.0e+00
Zatorsky2006_p53_Model6	BIOMD000000155	0.023	2.5e-01	0	1.0e+00
Hunziker2010_p53_StressSpecificResponse	BIOMD0000000252	0.023	2.5e-01	0	1.0e+00
Zatorsky2006_p53_Model5	BIOMD0000000156	0.022	2.7e-01	0	1.0e+00
Zatorsky2006_p53_Model4	BIOMD0000000157	0.022	2.7e-01	0	1.0e+00
Zatorsky2006_p53_Model2	BIOMD000000158	0.022	2.7e-01	0	1.0e+00
Zatorsky2006_p53_Model1	BIOMD000000159	0.022	2.7e-01	0	1.0e+00
Proctor2008_p53_Mdm2_ATM	BIOMD000000188	0.013	4.3e-01	0	1.0e+00
McClean2007_CrossTalk	BIOMD0000000116	0.012	4.7e-01	0	1.0e+00
Proctor2008_p53_Mdm2_ARF	BIOMD000000189	0.012	4.9e-01	0	1.0e+00
Haberichter2007_cellcycle	BIOMD000000109	0.011	5.0e-01	0	1.0e+00
Sasagawa2005_MAPK	BIOMD0000000049	0.006	5.5e-01	0	1.0e+00

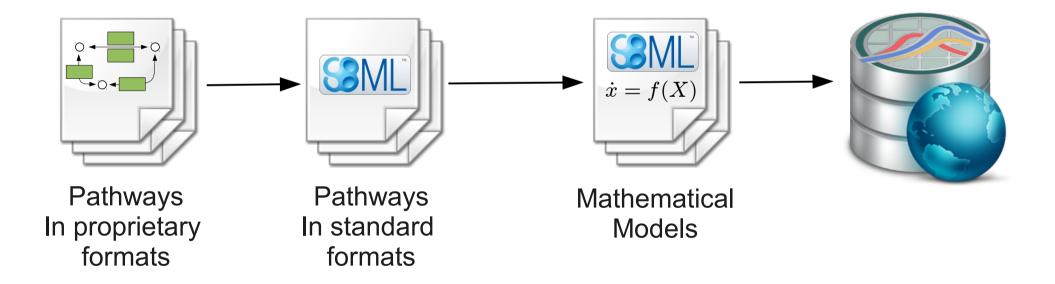


BUILD



From pathways to models ... path2models

- Provide pathways in a standard format
- Re-use existing pathway data to generate biochemically based models
- Provide starting points to build more quantitative models



Büchel et al. (2013)







PathwayInteractionDatabase



Logical models of individual signalling pathways







PathwayInteractionDatabase



Logical models of individual signalling pathways





Chemical kinetics models of individual metabolic pathways







PathwayInteractionDatabase



Logical models of individual signalling pathways





Chemical kinetics models of individual metabolic pathways







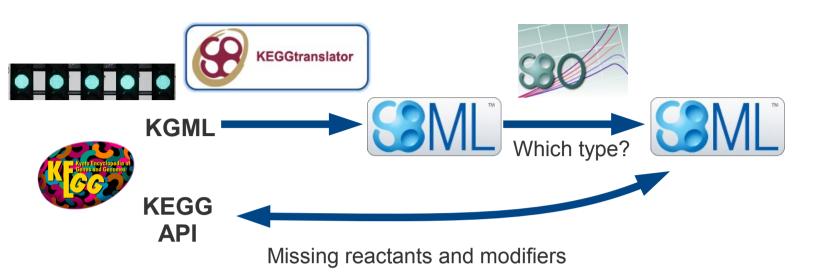
Flux Balance Analysis of whole genome reconstructions

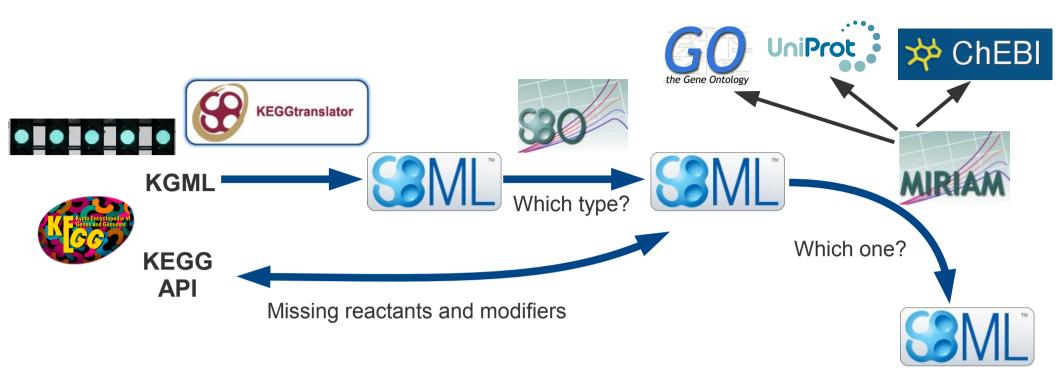


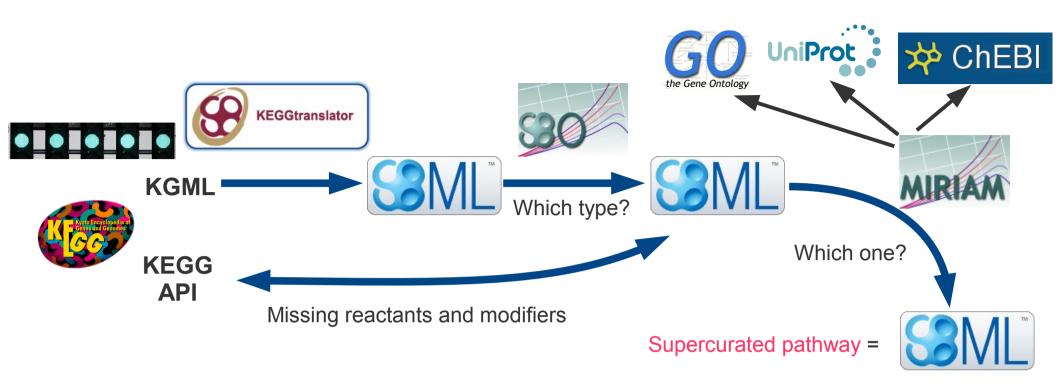


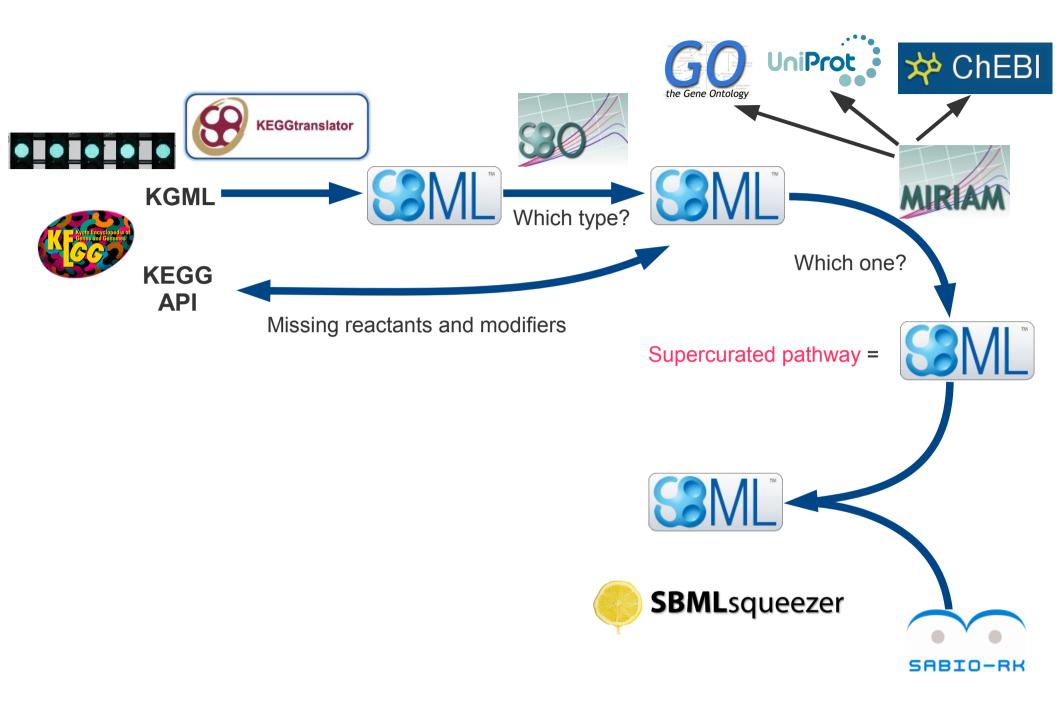




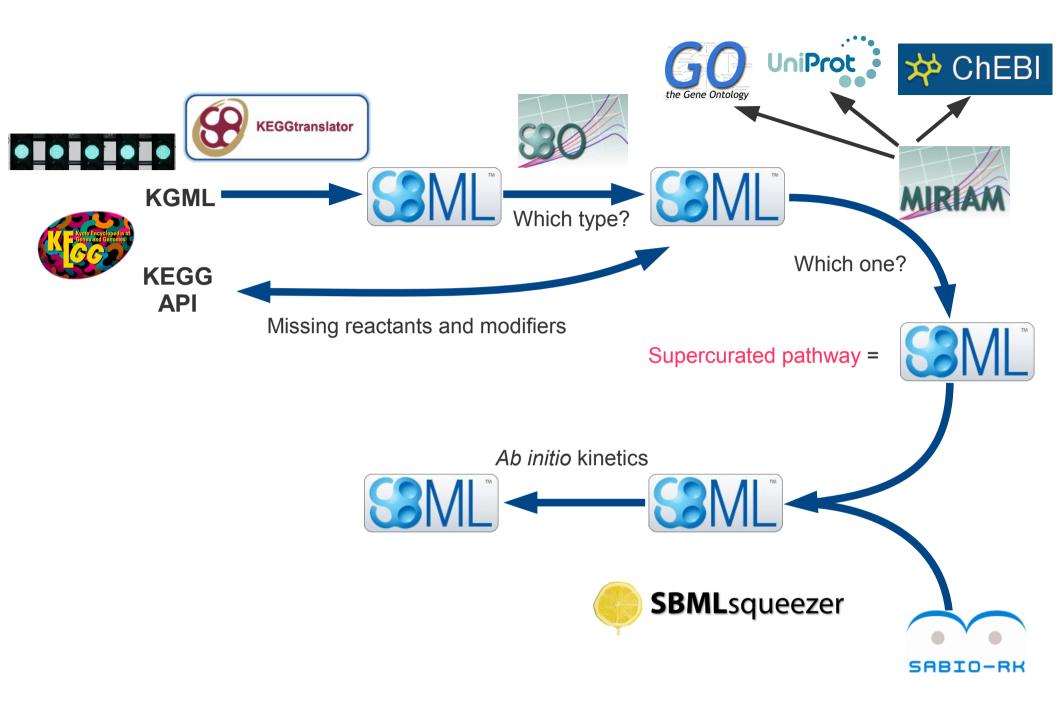




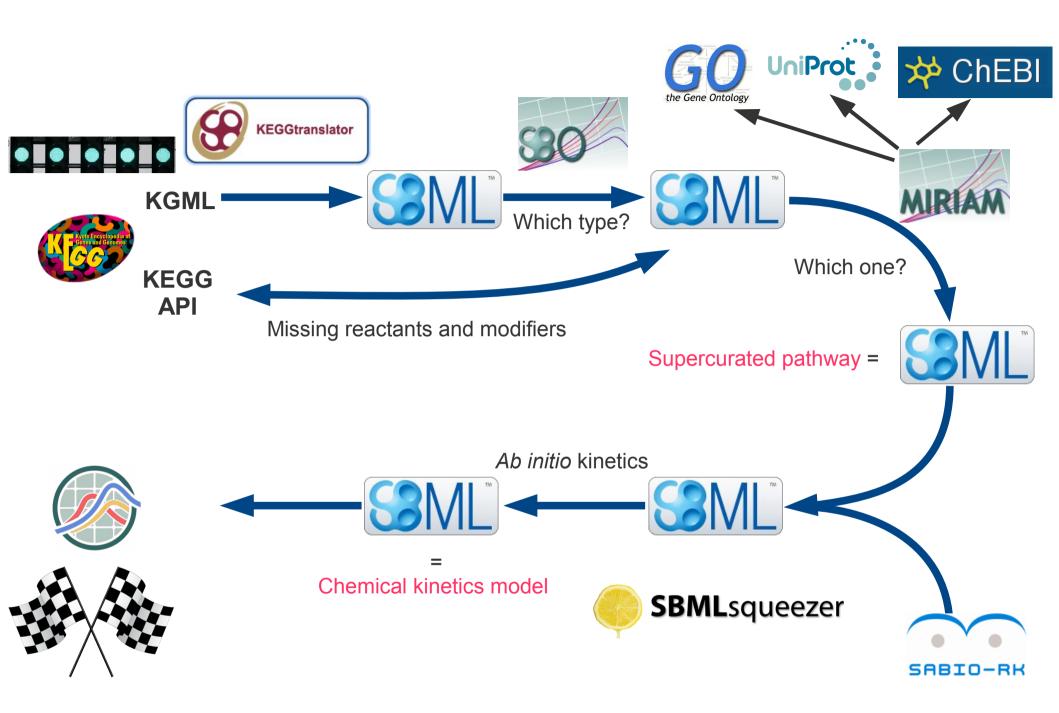








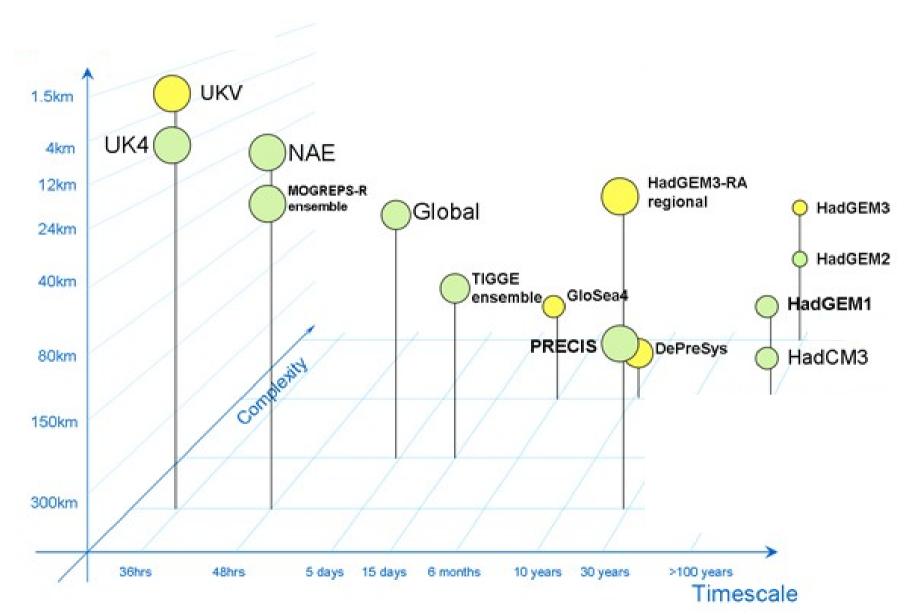


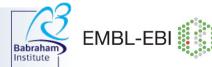




Are-we ready to develop comprehensive models of whole cells, organs and organisms?

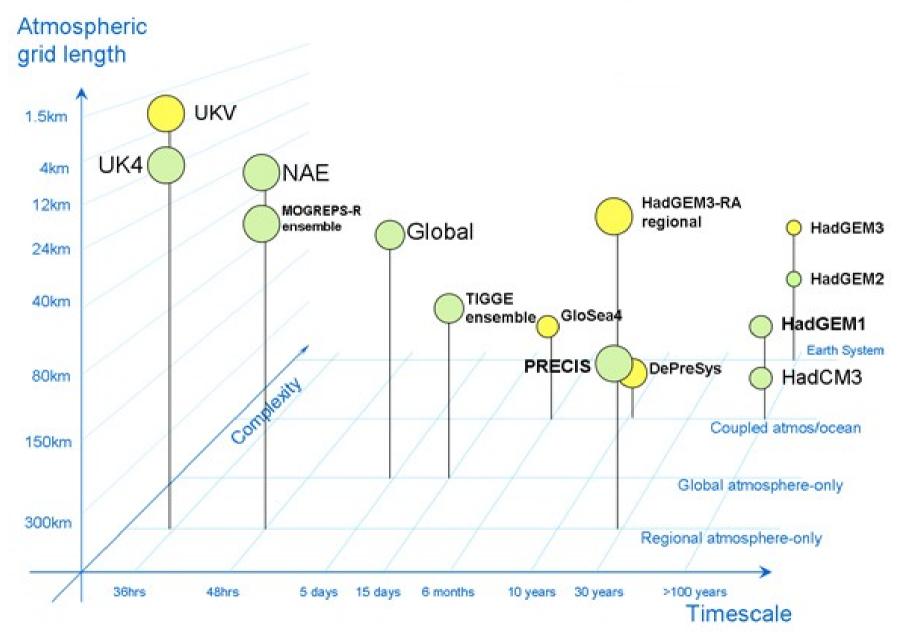






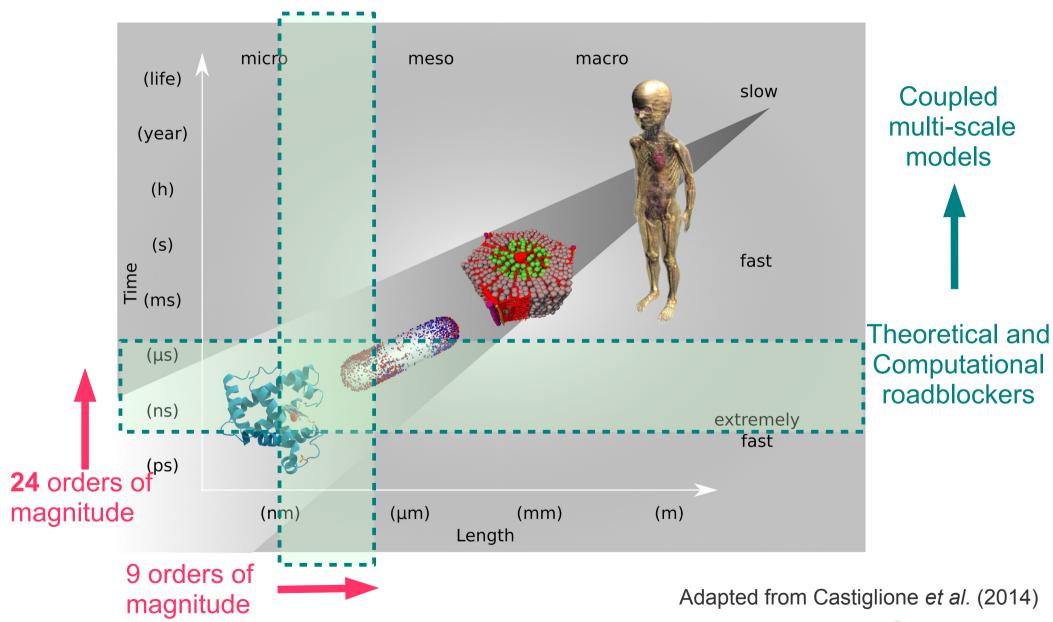


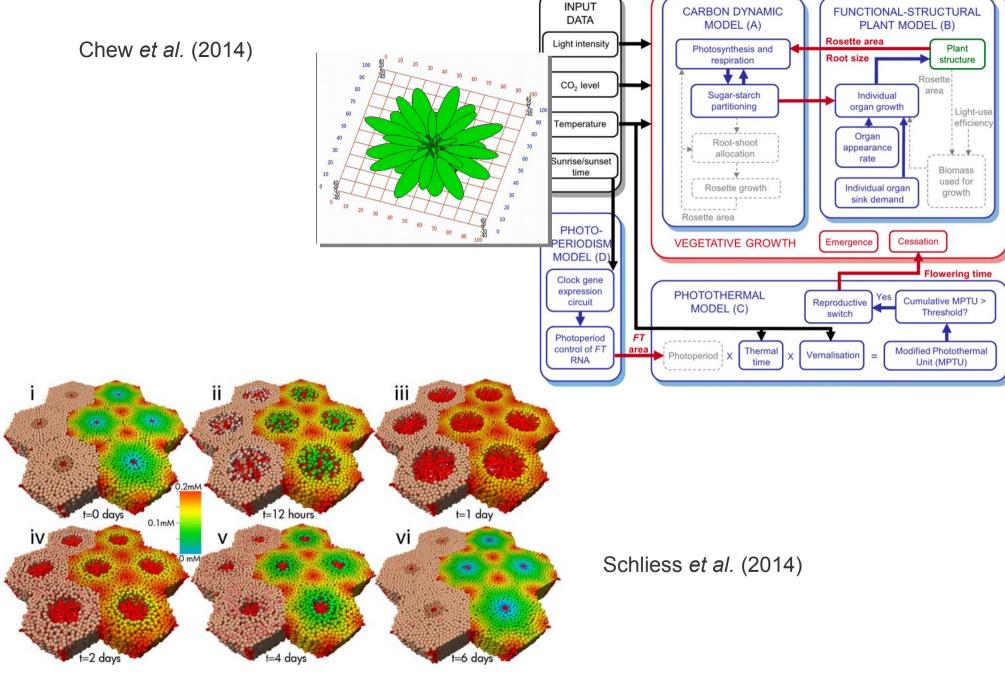
Met Office Seamless Unified Model

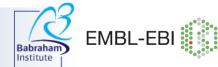




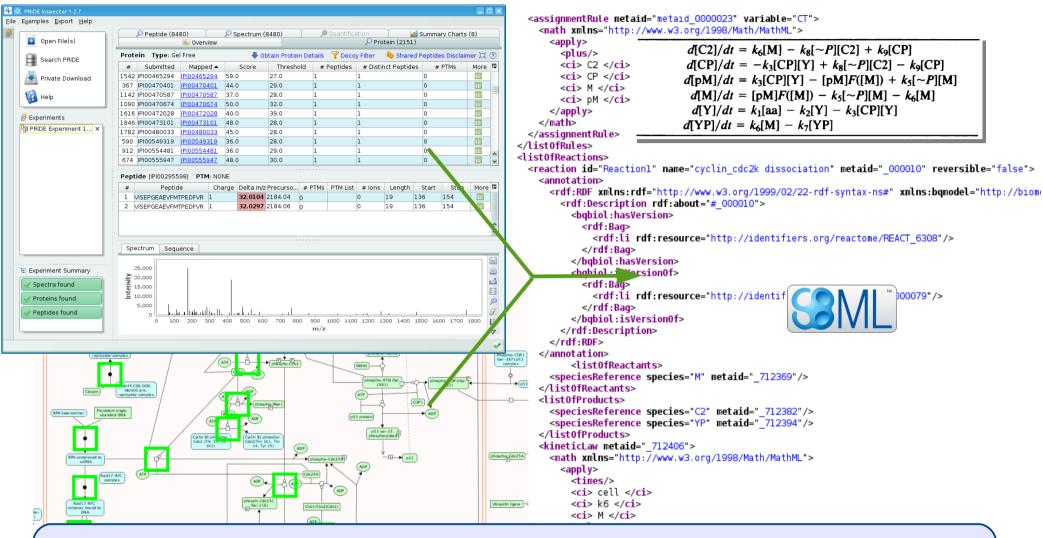
Challenge 1: scales







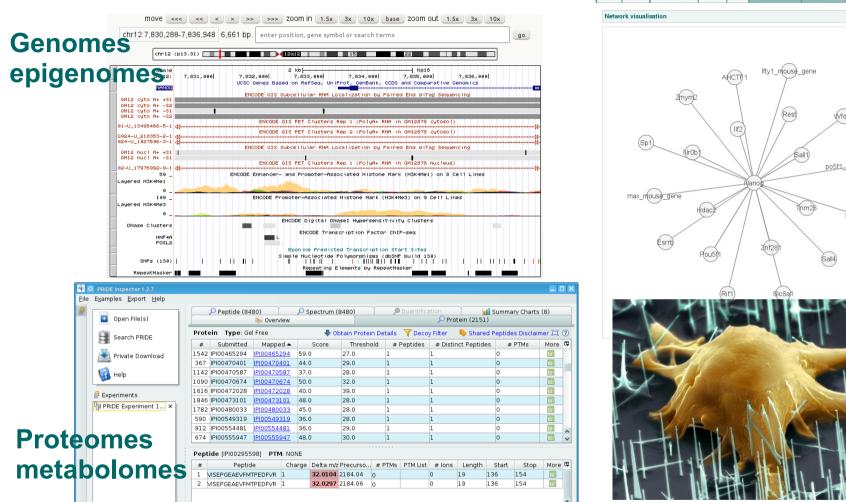
Challenge 2: automatic generation

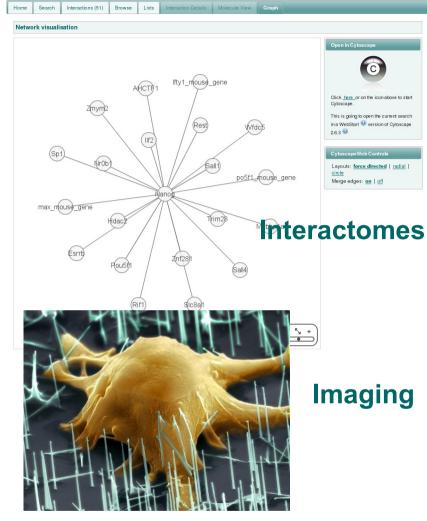


Need to generate model automatically based on omics datasets. Known on small scale. The larger, the more human intervention needed.

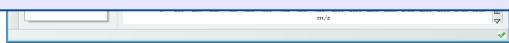


Challenge 3: data heterogeneity



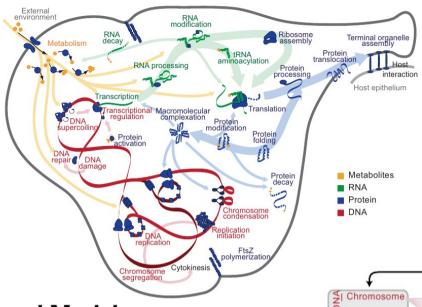


Need multiple types of datasets. Current solutions either for small models (e.g. Karr et al.), or focused on one type of model and data (e.g. metabolic models)





Modularity is mandatory



Theory

A Whole-Cell Computational Model Predicts Phenotype from Genotype

Jonathan R. Karr,^{1,4} Jayodita C. Sanghvi,^{2,4} Derek N. Macklin,² Miriam V. Gutschow,² Jared M. Jacobs,² Benjamin Bolival, Jr.,² Nacyra Assad-Garcia,³ John I. Glass,³ and Markus W. Covert^{2,*}

¹Graduate Program in Biophysics

²Department of Bioengineering

Stanford University, Stanford, CA 94305, USA

3J. Craig Venter Institute, Rockville, MD 20850, USA

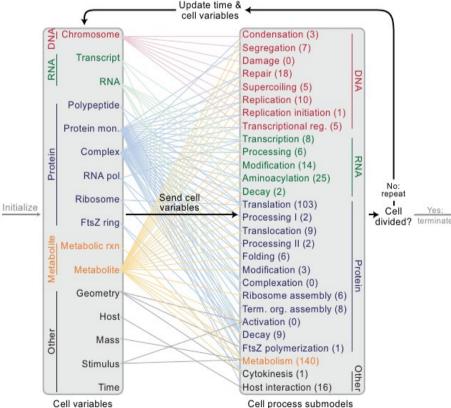
*Correspondence: mcovert@stanford.edu http://dx.doi.org/10.1016/i.cell.2012.05.044

SUMMARY

Understanding how complex phenotypes arise from individual molecules and their interactions is a primary challenge in biology that computational approaches are poised to tackle. We report a whole-cell computational model of the life cycle of the human pathogen Mycoplasma genitalium that includes all of its molecular components and their interactions. An integrative approach to modeling that combines diverse mathematics enabled the simultaneous inclusion of fundamentally different cellular processes and experimental measurements. Our whole-cell model accounts for all annotated gene functions and was validated against a broad range of data. The model provides insights into many previously unobserved cellular behaviors, including in vivo rates of protein-DNA association

First, until recently, not enough has been known about the individual molecules and their interactions to completely model any one organism. The advent of genomics and other hightroughput measurement techniques has accelerated the characterization of some organisms to the extent that comprehensive modeling is now possible. For example, the mycoplasmas, a genus of bacteria with relatively small genomes that includes several pathogens, have recently been the subject of an exhaustive experimental effort by a European consortium to determine the transcriptome (Güell et al., 2009), proteome (Kühner et al., 2009), and metabolome (Yus et al., 2009) of these organisms.

The second limiting factor has been that no single computational method is sufficient to explain complex phenotypes in terms of molecular components and their interactions. The first approaches to modeling cellular physiology, based on ordinary differential equations (ODEs) (Atlas et al., 2008; Browning et al., 2004; Castellanos et al., 2004, 2007; Domach et al., 1984; Tomita et al., 1999), were limited by the difficulty in obtaining the necessary model parameters. Subsequently, alternative





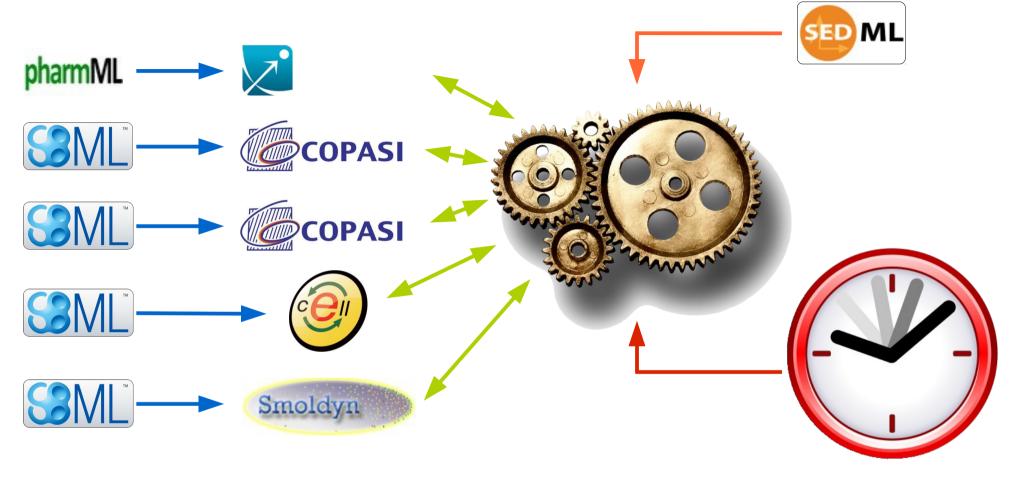
⁴These authors contributed equally to this work

Why is modularity important?

- Different parts can be modelled with different approaches
- Alternative modules representing the same biological process with different granularities or different modelling approaches
- Model families with alternative parts: avoid combinatorial explosion
- Modules can be developed at their own pace; easier versioning
- Distributed maintenance. No single individual or group can master the entirety of the project.



Challenge 4: Single software not sufficient



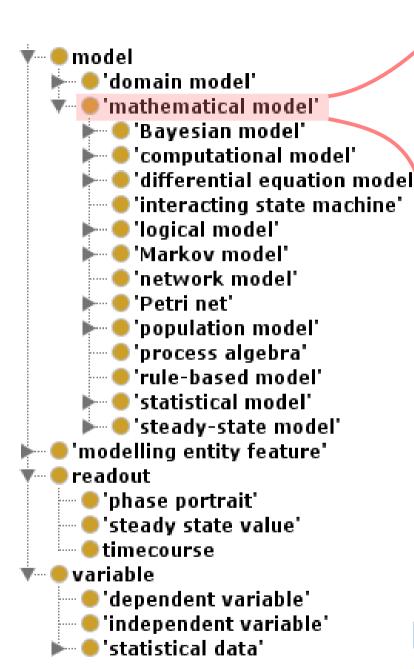
modular model

hybrid simulation system

discrete event detection

adaptive synchronization

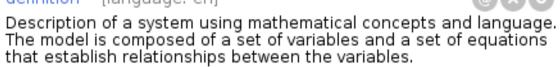




definition [language: en]







example [language: en]



A set of ordinary differential equation describing a physical process.

'preferred label' [language: en]



mathematical model





seeAlso [type: anyURI]

http://en.wikipedia.org/wiki/Mathematical model



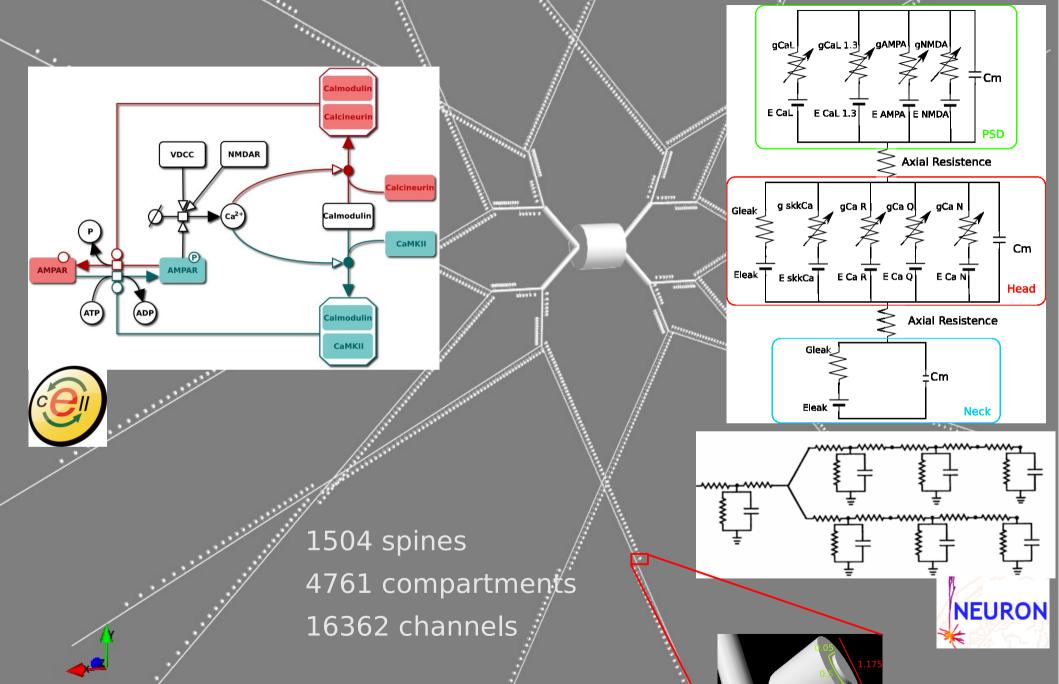


Mathematical Modelling Ontology

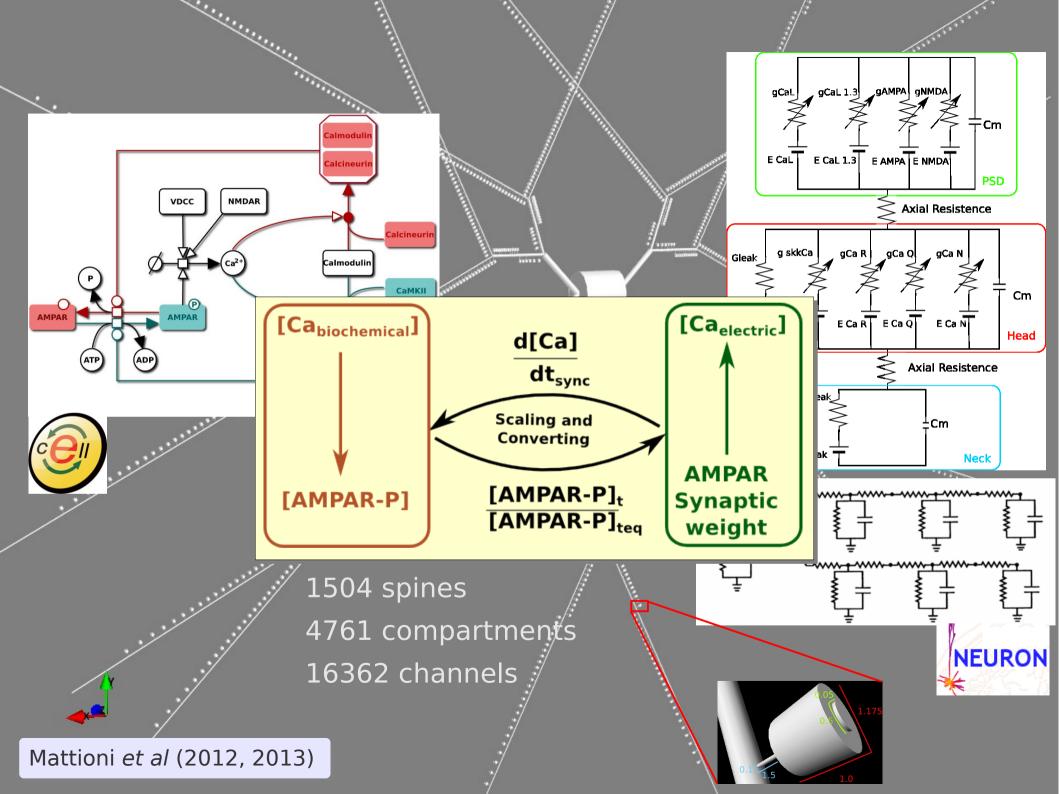
http://sourceforge.net/projects/mamo-ontology



Whole cell: electro-biochemical models



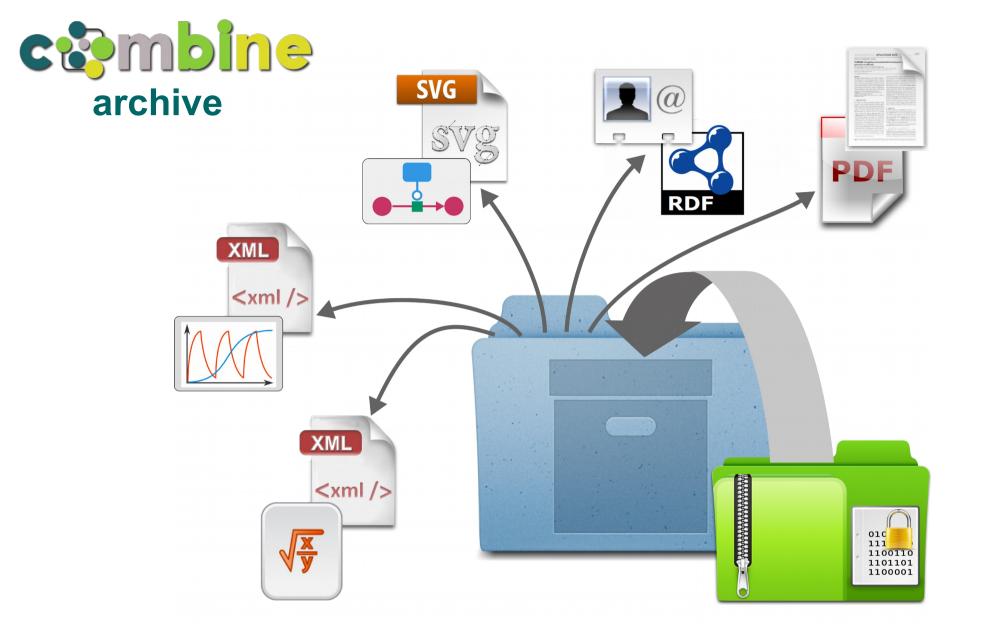
Mattioni et al (2012, 2013)



Challenge 5: reproducibility and virtualisation



Models must be self-contained and portable for everyone to reuse and customise → Virtual machines with models+tools



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

Bergman et al. BMC Syst Biol (2014)



BioModels team

Ishan Ajmera
Raza Ali
Duncan Berenguier
Alexander Broicher
Vijayalakshmi Chelliah
Melanie Courtot
Marco Donizelli

Marine Dumousseau

Lukas Endler Yvon Florent Mihai Glont Enuo He Arnaud Henry

Henning Hermjakob

Gael Jalowicki

Nick Juty

Sarah Keating Christian Knüpfer Nicolas Le Novère

Chen Li Camille Laibe

Stuart Moodie Kedar Nath Natarajan

Jean-Baptiste Pettit Michael Schubert

Maciej Swat Karim Tazibt

Dagmar Waltemath

Sarala Wimalaratne

Anna Zhukova

Standards' editors

Richard Adams Frank Bergman Hamid Bolouri Jonathan Cooper Tobias Czauderna **Emek Demir** Andrew Finney Stefan Hoops Mike Hucka Sarah Keating Hiroaki Kitano Nicolas I e Novère Yukiko Matsuoka Huaiyu Mi **Andrew Miller** Stuart Moodie Ion Moraru **Chris Myers David Nickerson Brett Olivier** Sven Sahle Herbert Sauro Jim Schaff Falk Schreiber Lucian Smith **Anatoly Sorokin** Alice Villeger Katia Wegner Darren Wilkinson

Dagmar Waltemath

Path2Models

Finja Büchel

Claudine Chaouiya Tobias Czauderna Andreas Dräger Mihai Glont Martin Golebievski Henning Hermiakob

Mike Hucka
Douglas Kell
Roland Keller
Camille Laibe
Nicolas Le Novère

Pedro Mendes

Florent Mittag
Wolfang Müller

Matthias Rall Nicolas Rodriguez

Julio Saez-Rodriguez Michael Schubert

Falk Schreiber

Neil Swainston

Martijn van Iersel Clemens Wrzodek

Andreas Zell

The numerous scientists who participated to discussions

The community of Computational Systems Biology



































National Human Genome Research Institute













