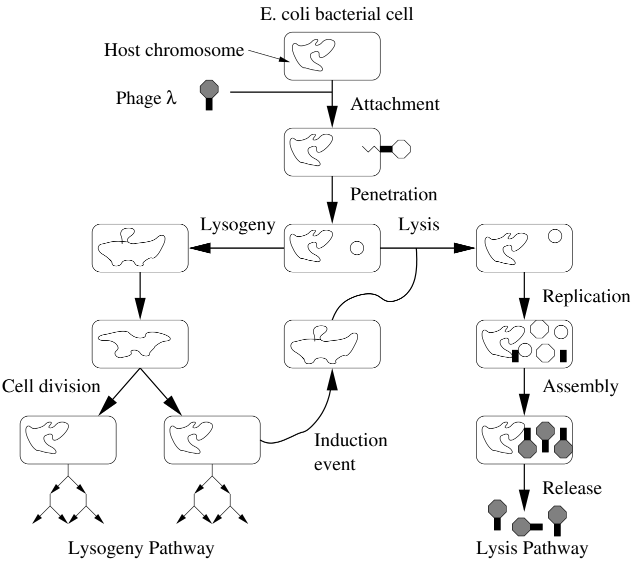
A Standard-Enabled Workflow for Synthetic Biology

Chris J. Myers

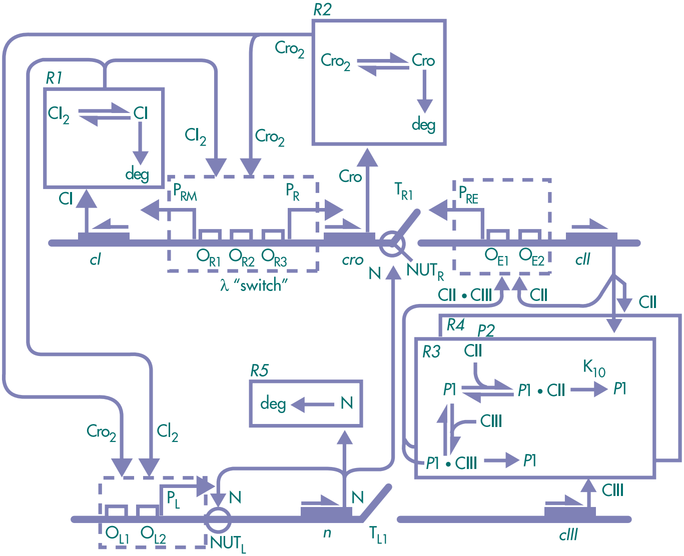
University of Utah

Computational Methods in Systems Biology
September 14, 2018

Phage λ Developmental Pathways

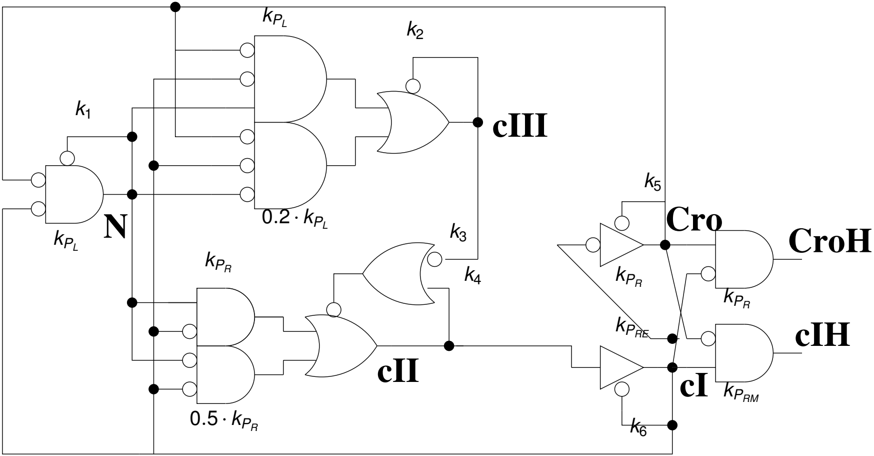


Phage λ Decision Circuit



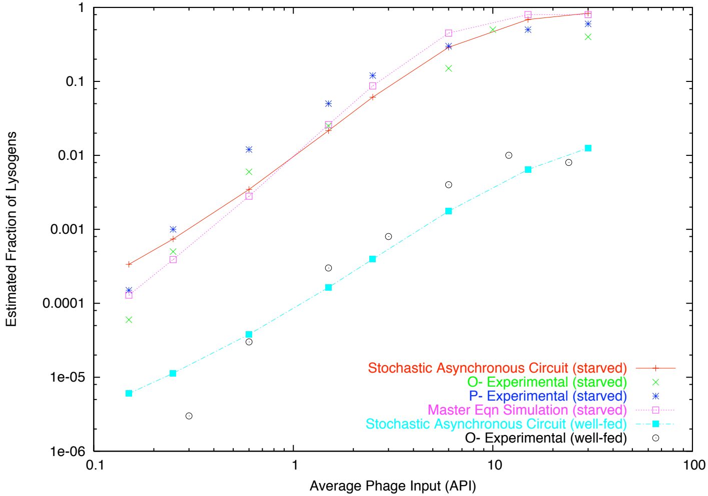
YGG 01-0052

Stochastic Asynchronous Circuit Model for Phage λ



Kuwahara et al., *Computational Methods in Systems Biology* (2005)
Kuwahara et al., *Transactions on Computational Systems. Biology* (2006)

Stochastic Asynchronous Circuit Results



SAC results generated in only 7 minutes.

Kuwahara et al., *Computational Methods in Systems Biology* (2005)

Kuwahara et al., *Transactions on Computational Systems. Biology* (2006)

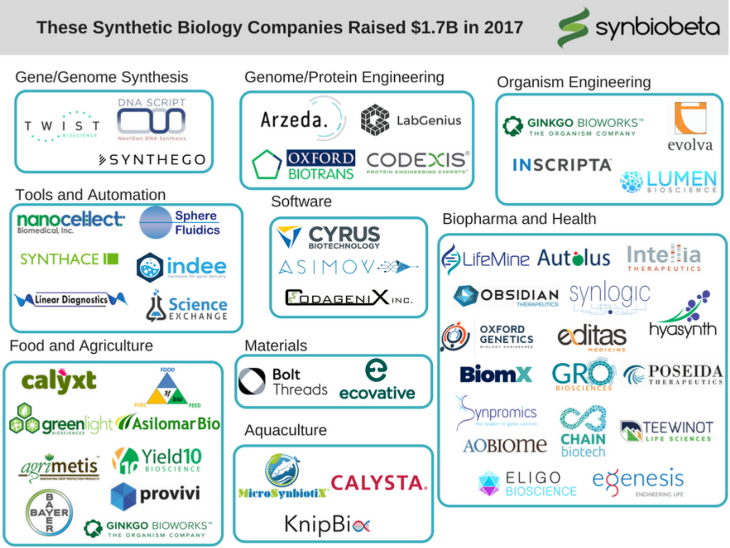
Systems Biology Versus Synthetic Biology

Drew Endy

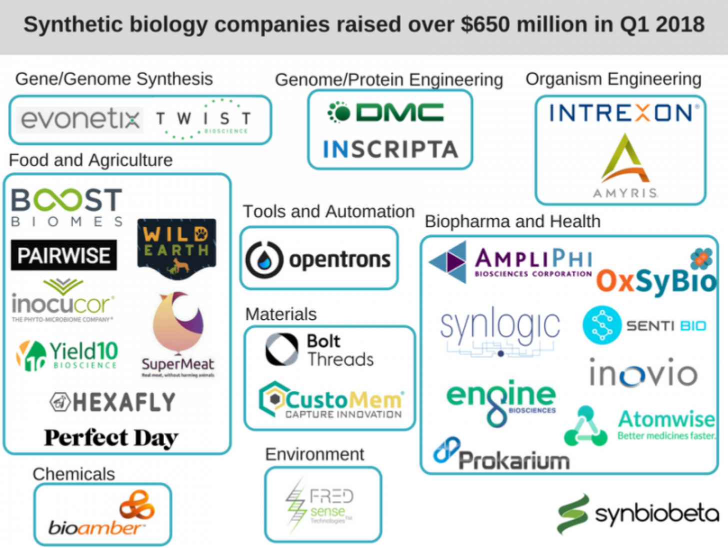
Synthetic Biology Startups

These 33 Synthetic Biology Companies
Raised More Than $900 Million in 2016

Synthetic Biology Startups

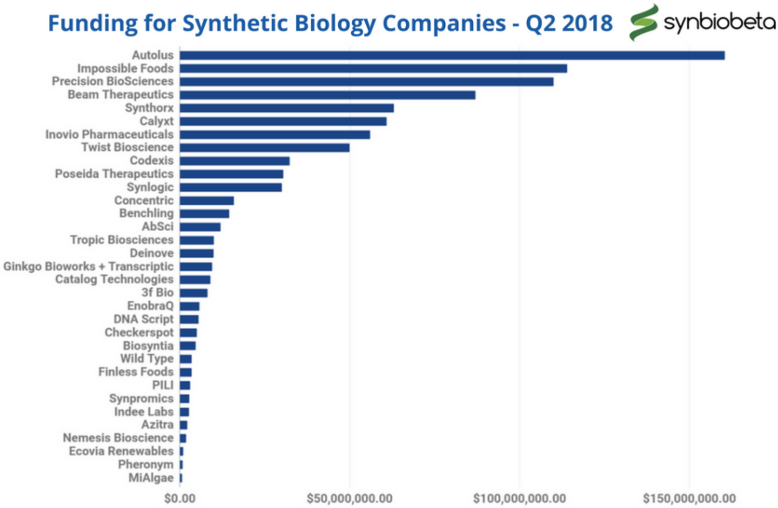


Synthetic Biology Startups



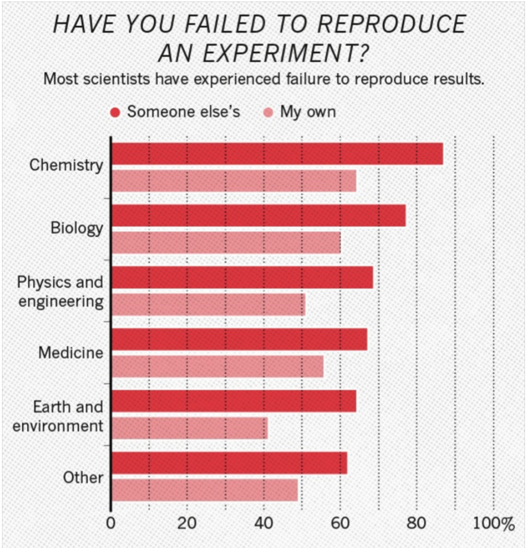
Synthetic biology companies raised over $650 million in Q1 2018

Synthetic Biology Startups



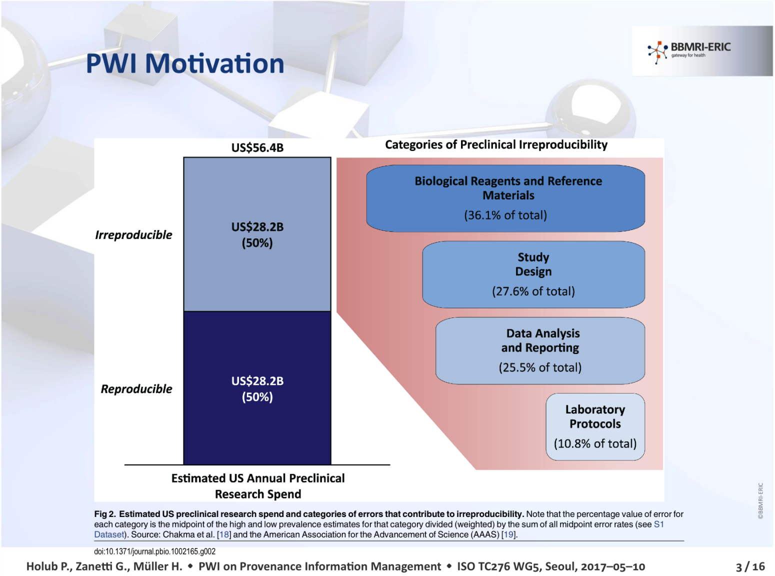
These 33 Synthetic Biology Companies
Raised $925 Million in 2018 Q2

Reproducibility Crisis



(V. Simonyan, Center for Biologics Evaluation and Research FDA, USA)

Reproducibility Crisis



PWI Motivation

BBMRI-ERIC
gateway for health

US$56.4B

Irreproducible US$28.2B (50%)

Reproducible US$28.2B (50%)

Estimated US Annual Preclinical Research Spend

Categories of Preclinical Irreproducibility

- **Biological Reagents and Reference Materials** (36.1% of total)
- **Study Design** (27.6% of total)
- **Data Analysis and Reporting** (25.5% of total)
- **Laboratory Protocols** (10.8% of total)

Fig 2. Estimated US preclinical research spend and categories of errors that contribute to irreproducibility. Note that the percentage value of error for each category is the midpoint of the high and low prevalence estimates for that category divided (weighted) by the sum of all midpoint error rates (see S1 Dataset). Source: Chakma et al. [18] and the American Association for the Advancement of Science (AAAS) [19].

doi:10.1371/journal.pbio.1002165.g002

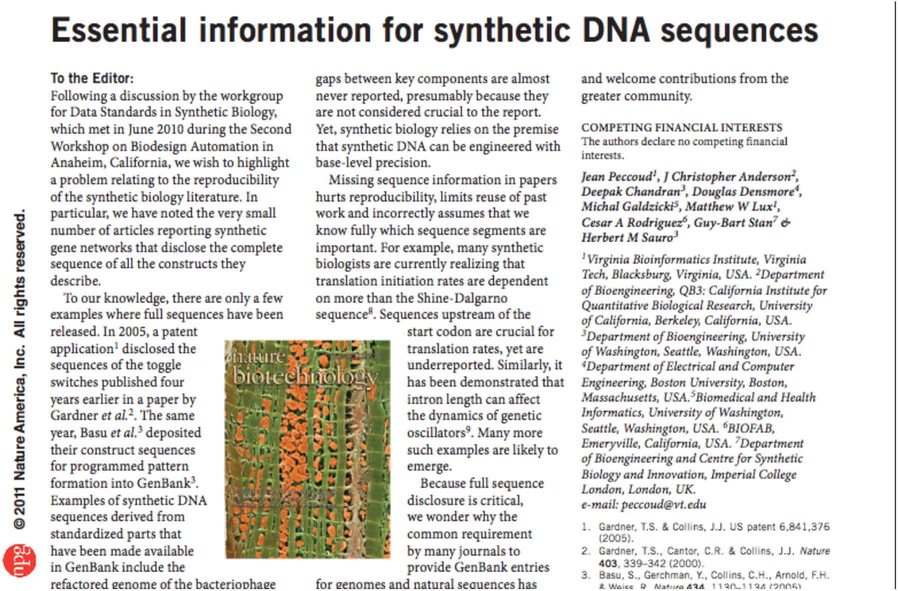
©BBMRI-ERIC

Holub P., Zanetti G., Müller H. • PWI on Provenance Information Management • ISO TC276 WG5, Seoul, 2017–05–10 3 / 16

Reproducibility Crisis

> *An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete ... set of instructions [and data] which generated the figures.*
>
> *– David Donoho, 1998*

Reproducibility in Synthetic Biology



Essential information for synthetic DNA sequences

To the Editor:
Following a discussion by the workgroup for Data Standards in Synthetic Biology, which met in June 2010 during the Second Workshop on Biodesign Automation in Anaheim, California, we wish to highlight a problem relating to the reproducibility of the synthetic biology literature. In particular, we have noted the very small number of articles reporting synthetic gene networks that disclose the complete sequence of all the constructs they describe.

To our knowledge, there are only a few examples where full sequences have been released. In 2005, a patent application[1] disclosed the sequences of the toggle switches published four years earlier in a paper by Gardner *et al.*[2]. The same year, Basu *et al.*[3] deposited their construct sequences for programmed pattern formation into GenBank[3]. Examples of synthetic DNA sequences derived from standardized parts that have been made available in GenBank include the refactored genome of the bacteriophage

gaps between key components are almost never reported, presumably because they are not considered crucial to the report. Yet, synthetic biology relies on the premise that synthetic DNA can be engineered with base-level precision.

Missing sequence information in papers hurts reproducibility, limits reuse of past work and incorrectly assumes that we know fully which sequence segments are important. For example, many synthetic biologists are currently realizing that translation initiation rates are dependent on more than the Shine-Dalgarno sequence[8]. Sequences upstream of the start codon are crucial for translation rates, yet are underreported. Similarly, it has been demonstrated that intron length can affect the dynamics of genetic oscillators[9]. Many more such examples are likely to emerge.

Because full sequence disclosure is critical, we wonder why the common requirement by many journals to provide GenBank entries for genomes and natural sequences has

and welcome contributions from the greater community.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Jean Peccoud[1], J Christopher Anderson[2], Deepak Chandran[3], Douglas Densmore[4], Michal Galdzicki[5], Matthew W Lux[1], Cesar A Rodriguez[6], Guy-Bart Stan[7] & Herbert M Sauro[3]

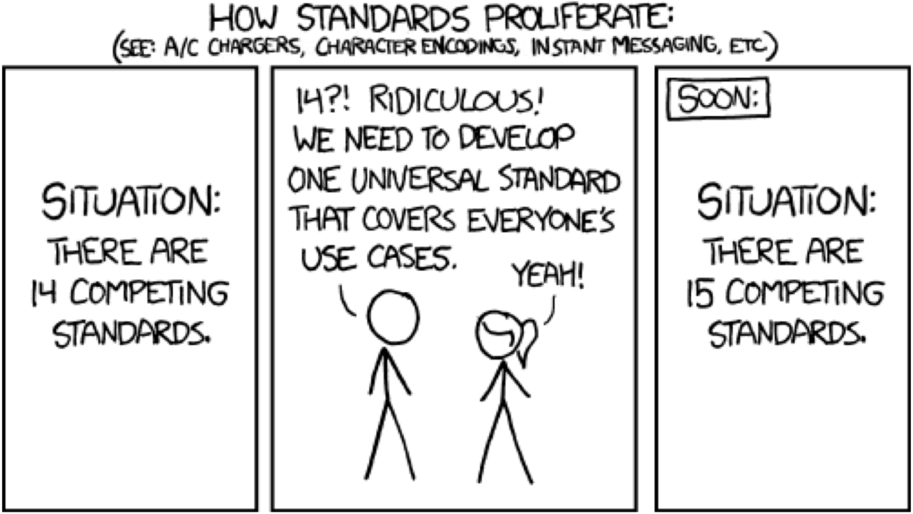
[1]Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, Virginia, USA. [2]Department of Bioengineering, QB3: California Institute for Quantitative Biological Research, University of California, Berkeley, California, USA. [3]Department of Bioengineering, University of Washington, Seattle, Washington, USA. [4]Department of Electrical and Computer Engineering, Boston University, Boston, Massachusetts, USA. [5]Biomedical and Health Informatics, University of Washington, Seattle, Washington, USA. [6]BIOFAB, Emeryville, California, USA. [7]Department of Bioengineering and Centre for Synthetic Biology and Innovation, Imperial College London, London, UK.
e-mail: peccoud@vt.edu

1. Gardner, T.S. & Collins, J.J. US patent 6,841,376 (2005).
2. Gardner, T.S., Cantor, C.R. & Collins, J.J. *Nature* **403**, 339–342 (2000).
3. Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H. & Weiss, R. *Nature* **434**, 1130–1134 (2005).

Standards to the Rescue

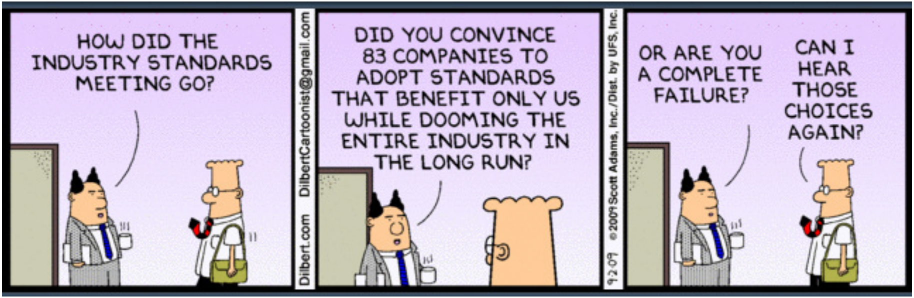
(source `https://www.eaglegenomics.com/do-data-standards-really-matter/`)

Word of Warning



(source `xkcd.com`)

Proprietary Standards



Coordination of Standard Development in Systems/Synthetic Biology

http://co.mbine.org

- COmputational Modeling in BIology NEtwork
- Tasks and Actions:
  - Organizes joint standards meetings: HARMONY & the COMBINE Forum
  - Provides training in application of standards (COMBINE tutorials)
  - Coordinates standards development
  - Develops common procedures and tools
  - Provides a recognized voice

COMBINE History

Enabling technologies

2003
libSBML

2005
MIRIAM, SBO, BioModels qualifiers

2006
PaxTools

2007
MIASE, KiSAO

2011
Identifiers.org

September 2014
COMBINE Archive

Representation formats

March 2001
SBML Level 1

August 2001
CellML 1.0, NeuroML

June 2003
SBML Level 2

July 2004
BioPAX Level 1

December 2005
BioPAX Level 2

August 2008
SBGN PD L1

September 2009
SBGN ER L1, SBGN AF L1

March 2010
SED-ML Level 1

July 2010
BioPAX Level 3

October 2010
SBML Level 3

October 2011
SBOL v1

March 2013
SBOLvisual v1

July 2015
SBOL v2

1999

Influential meetings

April 1999
NATO workshop, proposing to create
a language to encode metabolic models

April 2000
Start of SBML at the
1st "ERATO Kitano" workshop,

August 2002
Start of BioPAX project at the
4st Biopathway consortium meeting

July 2003
1st SBML hackathon

October 2005
Start of SBGN project at
the BioPAX face 2 face meeting

2006
Decision to create a language for synth biol designs

January 2008
Okinawa superhackathon
SBGN, BioPAX, SBO, MIRIAM

April 2008
1st SBOL meeting

April 2009
Waiheke combined meeting
CellML, SBGN, BioPAX, SBO, MIASE

2010

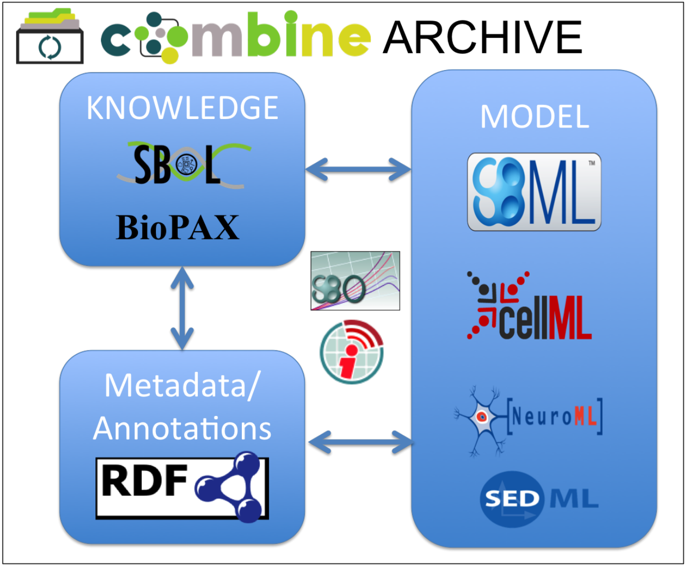
Creation of COMBINE

October 2010
1st COMBINE forum

April 2011
1st HARMONY hackathon

Myers et al., *2017 Winter Simulation Conference* (2017).

COMBINE Overview



combine ARCHIVE

KNOWLEDGE

SBOL

BioPAX

MODEL

SBML

CellML

NeuroML

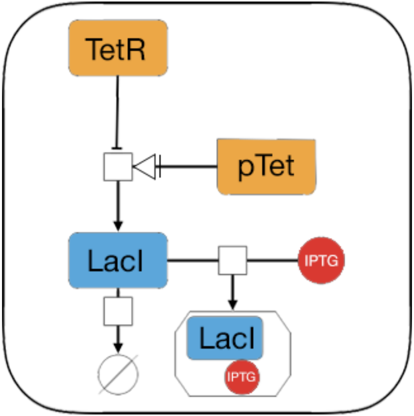
SED ML

Metadata/ Annotations

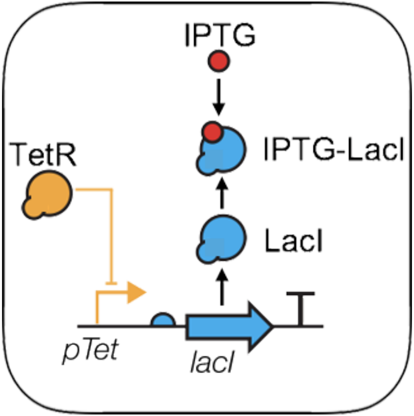
RDF

SBO

COMBINE Visualization Standards

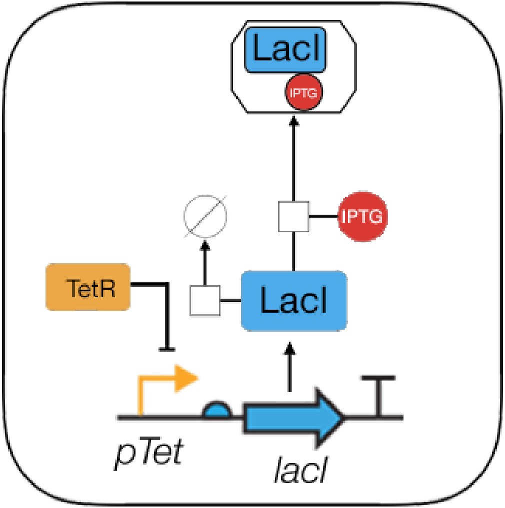


Systems Biology Graphical Notation
(SBGN)



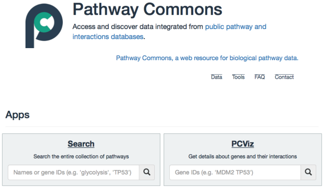
SBOL Visual
(SBOLv)

COMBINE Visualization Standards



Unified SBGN/SBOLv Diagram

COMBINE Repositories



Pathway Commons

Access and discover data integrated from public pathway and interactions databases.

Pathway Commons, a web resource for biological pathway data.

Apps

Search

Search the entire collection of pathways

PCViz

Get details about genes and their interactions

`http://www.pathwaycommons.org`

SynBioHub

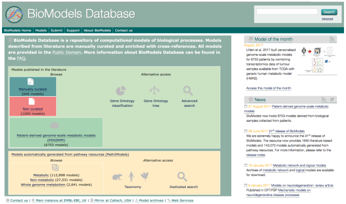
SynBioHub is a *design repository* for people designing biological constructs. It enables DNA and protein designs to be uploaded, then provides a shareable link to allow others to view them. SynBioHub also facilitates searching for information about existing useful parts and designs by combining data from a variety of sources.

Search for useful parts and designs

Upload your design for safekeeping

Share designs for publication or collaboration

`https://synbiohub.org`



BioModels Database

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

- Manually curated (640 models)
- Non curated (1000 models)
- Patient-derived genome-scale metabolic models (PDGSMM) (6753 models)

Alternative access

- Gene Ontology classification
- Gene Ontology tree
- Advanced search

Models automatically generated from pathway resources (Path2Models)

Browse

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

Alternative access

- Taxonomy
- Dedicated search

Model of the month

August, 2017

Uhlen et al. 2017 built personalised genome-scale metabolic models for 6753 patients by combining transcriptomics data of tumour samples available from TCGA with generic human metabolic model (HMR2).

Access this model of the month.

News

21 August 2017 Patient-derived genome-scale metabolic models

BioModels now hosts 6753 models derived from biological samples collected from patients.

28 June 2017 31st release of BioModels

We are extremely happy to announce the 31st release of BioModels. The resource now provides 1640 literature-based models and 143,070 models automatically generated from pathway resources. For more information, please refer to the release notes.

16 January 2017 Metabolic network and logical models

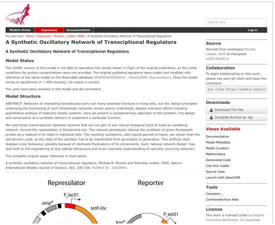
Archives of metabolic network and logical models are available for download.

9 January 2017 Models on neurodegeneration: review article

Published in CPT:PSP Mechanistic models on neurodegenerative disease processes

Contact us | Main instance at EMBL-EBI, UK | Mirror at Caltech, USA | Model archives | Web Services

`http://biomodels.net`



You are here: Home / Exposures / Elowitz, Leibler, 2000 / A Synthetic Oscillatory Network of Transciptional Regulators

A Synthetic Oscillatory Network of Transciptional Regulators

A Synthetic Oscillatory Network of Transciptional Regulators

Model Status

This CellML version of this model is not able to reproduce the results shown in Fig1c of the original publication, as the initial conditions for protein concentrations were not provided. The original published equations were scaled and modified with reference to the same model on the Biomodels database (BIOMD0000000012 - Elowitz2000_Repressilator). Once the model comes to equilibrium (t > 400 minutes,) its output is correct.

The units have been checked in this model and are consistent.

Model Structure

ABSTRACT: Networks of interacting biomolecules carry out many essential functions in living cells, but the 'design principles' underlying the functioning of such intracellular networks remain poorly understood, despite intensive efforts including quantitative analysis of relatively simple systems. Here we present a complementary approach to this problem: the design and construction of a synthetic network to implement a particular function.

We used three transcriptional repressor systems that are not part of any natural biological clock to build an oscillating network, termed the repressilator, in Escherichia coli. The network periodically induces the synthesis of green fluorescent protein as a readout of its state in individual cells. The resulting oscillations, with typical periods of hours, are slower than the cell-division cycle, so the state of the oscillator has to be transmitted from generation to generation. This artificial clock displays noisy behaviour, possibly because of stochastic fluctuations of its components. Such 'rational network design' may lead both to the engineering of new cellular behaviours and to an improved understanding of naturally occurring networks.

The complete original paper reference is cited below:

A synthetic oscillatory network of transcriptional regulators, Michael B. Elowitz and Stanislas Leibler, 2000, *Nature: International Weekly Journal of Science*, 403, 335-338. PubMed ID: 10659856

Source

Derived from workspace Elowitz, Leibler, 2000 at changeset cb2913da4514.

Collaboration

To begin collaborating on this work, please use your git client and issue this command:

```
git clone https://models.physio
```

Downloads

Views Available

- Documentation
- Model Metadata
- Model Curation
- Mathematics
- Generated Code
- Cite this model
- Source View
- Launch with OpenCOR

Tools

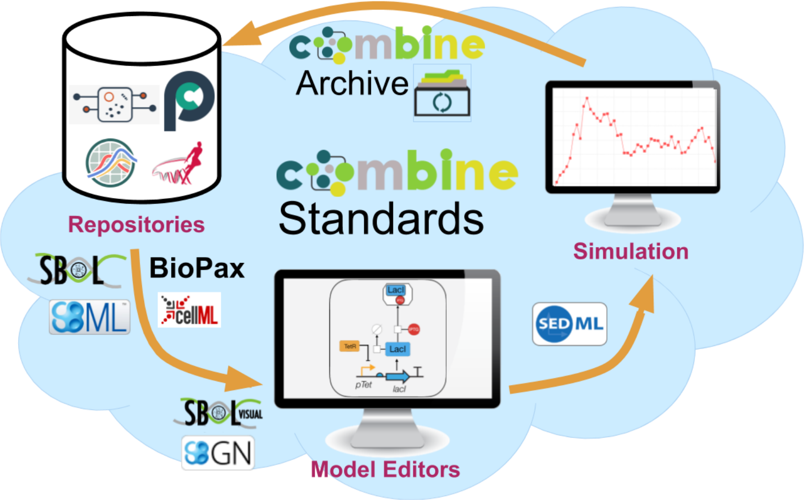
- Compare...
- CombineArchive Web

License

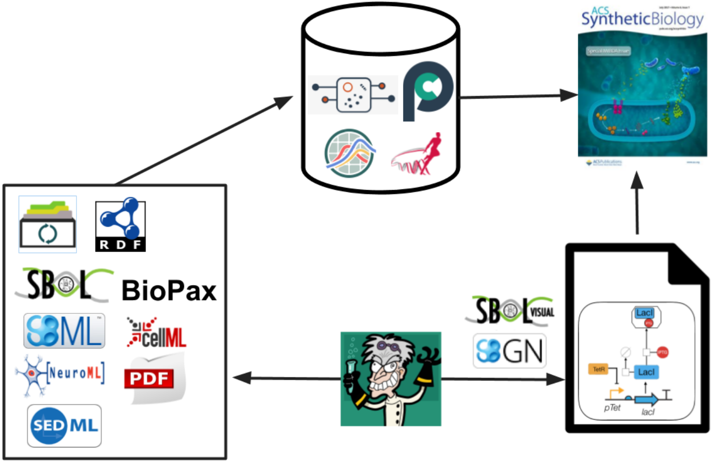
This work is licensed under a Creative Commons Attribution 3.0 Unported

`https://models.physiomeproject.org`

Standard Enabled Systems/Synthetic Biology Workflow



Journal Workflow for Reproducibility



Invitation

- You are invited to join the COMBINE community.
- Contact the COMBINE Coordinators or standard editors to join the appropriate mailing lists.
- You are also invited to the upcoming COMBINE Forum:

Boston University
October 8-12

Synthetic Biology Open Language (SBOL)
Version 1 Released in 2011

NATURE BIOTECHNOLOGY | COMPUTATIONAL BIOLOGY | PERSPECTIVE

日本語要約

The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology

Michal Galdzicki, Kevin P Clancy, Ernst Oberortner, Matthew Pocock, Jacqueline Y Quinn, Cesar A Rodriguez, Nicholas Roehner, Mandy L Wilson, Laura Adam, J Christopher Anderson, Bryan A Bartley, Jacob Beal, Deepak Chandran, Joanna Chen, Douglas Densmore, Drew Endy, Raik Grünberg, Jennifer Hallinan, Nathan J Hillson, Jeffrey D Johnson, Allan Kuchinsky, Matthew Lux, Goksel Misirli, Jean Peccoud, Hector A Plahar, Evren Sirin, Guy-Bart Stan, Alan Villalobos, Anil Wipat, John H Gennari, Chris J Myers & Herbert M Sauro

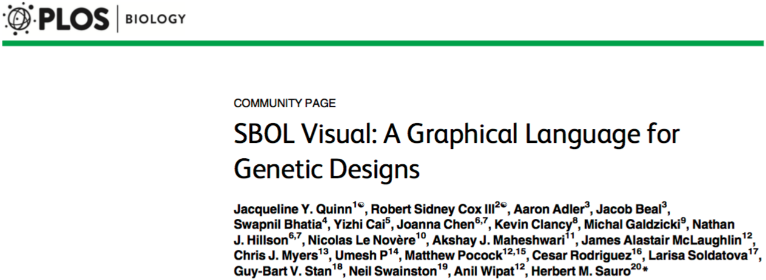
⊟ Show fewer authors

Affiliations | Contributions | Corresponding author

Nature Biotechnology **32**, 545–550 (2014) | doi:10.1038/nbt.2891

Received 09 November 2013 | Accepted 20 December 2013 | Published online 06 June 2014

SBOL Visual Version 1 Released in 2013

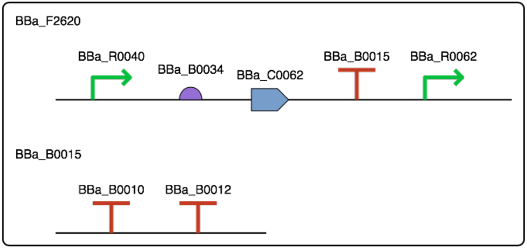


PLOS | BIOLOGY

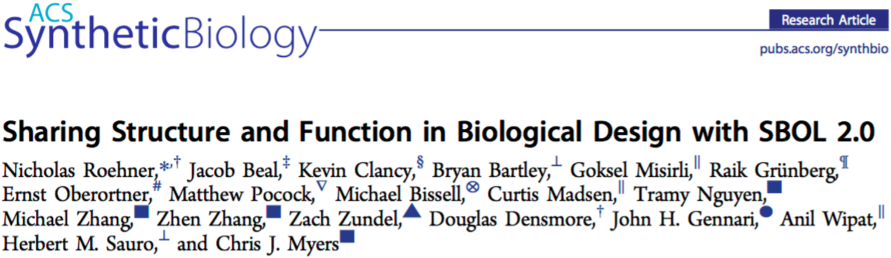
COMMUNITY PAGE

SBOL Visual: A Graphical Language for Genetic Designs

Jacqueline Y. Quinn[1]☯, Robert Sidney Cox III[2]☯, Aaron Adler[3], Jacob Beal[3], Swapnil Bhatia[4], Yizhi Cai[5], Joanna Chen[6,7], Kevin Clancy[8], Michal Galdzicki[9], Nathan J. Hillson[6,7], Nicolas Le Novère[10], Akshay J. Maheshwari[11], James Alastair McLaughlin[12], Chris J. Myers[13], Umesh P[14], Matthew Pocock[12,15], Cesar Rodriguez[16], Larisa Soldatova[17], Guy-Bart V. Stan[18], Neil Swainston[19], Anil Wipat[12], Herbert M. Sauro[20]*



SBOL Version 2 Released in 2015



ACS Synthetic Biology

Research Article

pubs.acs.org/synthbio

Sharing Structure and Function in Biological Design with SBOL 2.0

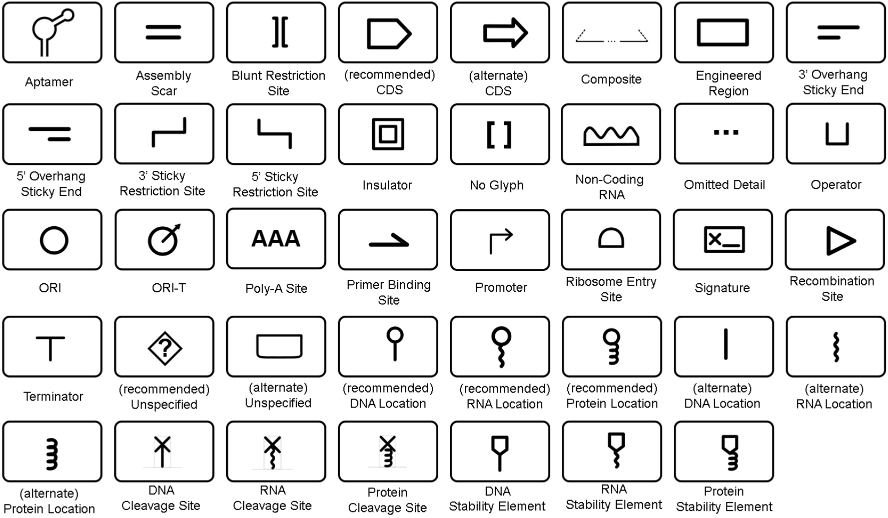
Nicholas Roehner,[*,†] Jacob Beal,[‡] Kevin Clancy,[§] Bryan Bartley,[⊥] Goksel Misirli,[∥] Raik Grünberg,[¶] Ernst Oberortner,[#] Matthew Pocock,[∇] Michael Bissell,[⊗] Curtis Madsen,[∥] Tramy Nguyen,[■] Michael Zhang,[■] Zhen Zhang,[■] Zach Zundel,[▲] Douglas Densmore,[†] John H. Gennari,[●] Anil Wipat,[∥] Herbert M. Sauro,[⊥] and Chris J. Myers[■]

SEE ALSO:

- Roehner et al., *ACS Synthetic Biology* (2014)
- Bartley et al., *Journal of Integrative Bioinformatics* (2015)
- Beal et al., *Journal of Integrative Bioinformatics* (2016)
- Cox et al., *Journal of Integrative Bioinformatics* (2017)

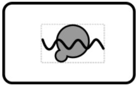
SBOLv Version 2: Nucleic Acid Glyphs (2017)

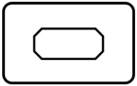
Nucleic Acid Glyphs

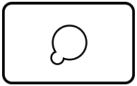


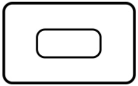
SBOLv Version 2: Molecular Species Glyphs (2017)

Molecular Species Glyphs

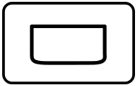

Superpose Glyphs Complex

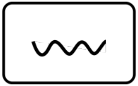

(alternate) Complex


(recommended) Macromolecule

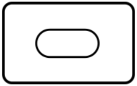

(alternate) Macromolecule


No Glyph

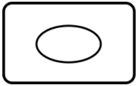

Nucleic Acid (Generic)


Nucleic Acid (1-Strand)

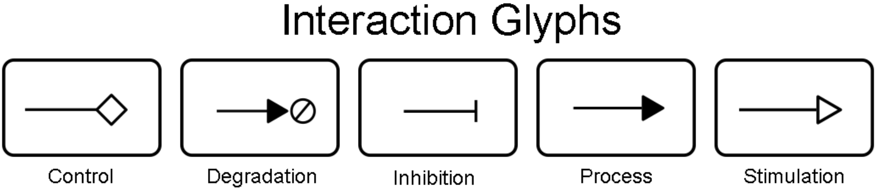

Nucleic Acid (2-Strand)


Small Molecule

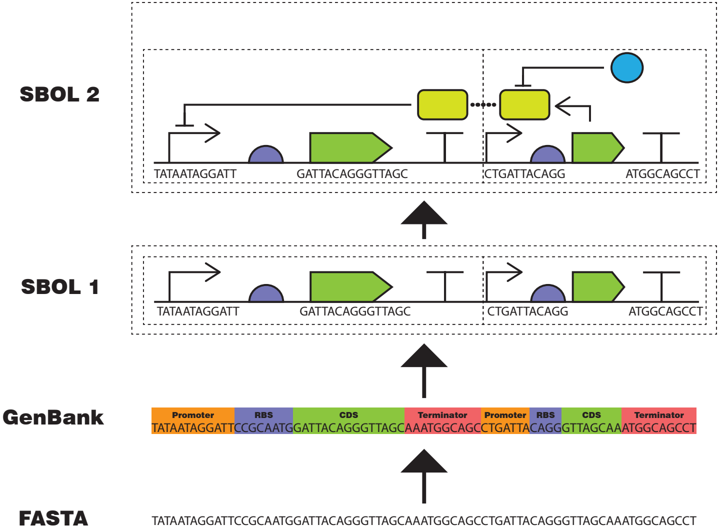

(recommended) Unspecified


(alternate) Unspecified

SBOLv Version 2: Interaction Glyphs (2017)



SBOL Evolution



SBOL 2

TATAATAGGATT GATTACAGGGTTAGC CTGATTACAGG ATGGCAGCCT

SBOL 1

TATAATAGGATT GATTACAGGGTTAGC CTGATTACAGG ATGGCAGCCT

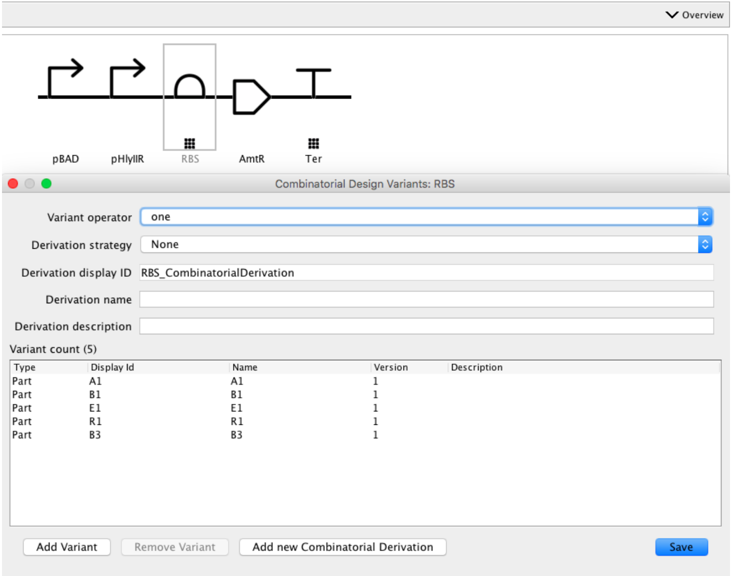
GenBank

Promoter	RBS	CDS	Terminator	Promoter	RBS	CDS	Terminator
TATAATAGGATT	CCGCAATG	GATTACAGGGTTAGC	AAATGGCAGC	CTGATTA	CAGG	GTTAGCAA	ATGGCAGCCT

FASTA

TATAATAGGATTCCGCAATGGATTACAGGGTTAGCAAATGGCAGCCTGATTACAGGGTTAGCAAATGGCAGCCT

SBOL Version 2.2: Combinatorial Derivations



Variant operator: one

Derivation strategy: None

Derivation display ID: RBS_CombinatorialDerivation

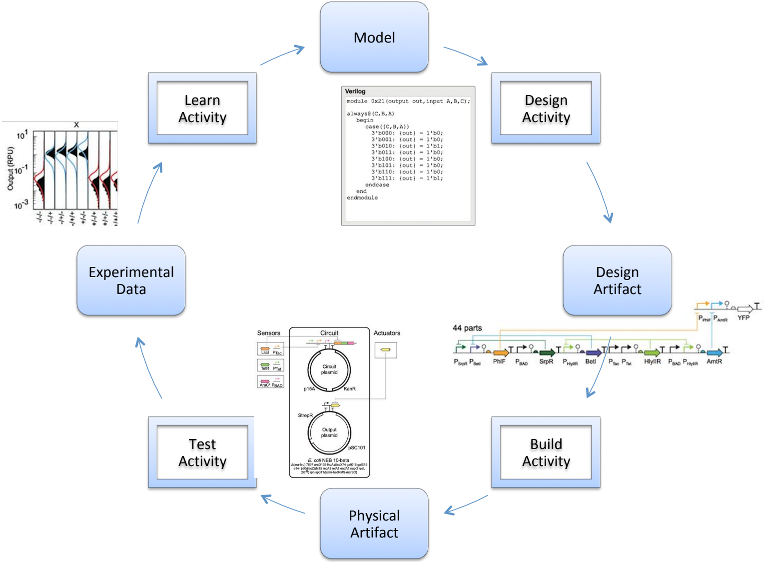
Derivation name:

Derivation description:

Variant count (5)

Type	Display Id	Name	Version	Description
Part	A1	A1	1	
Part	B1	B1	1	
Part	E1	E1	1	
Part	R1	R1	1	
Part	B3	B3	1	

SBOL Version 2.2: Design-Build-Test-Learn (Prov-O)



Model

Design Activity

Design Artifact

Build Activity

Physical Artifact

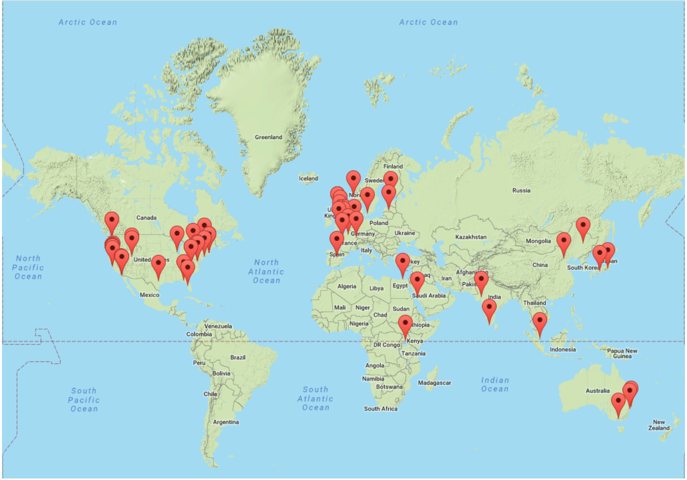
Test Activity

Experimental Data

Learn Activity

```
Verilog
module 0x21(output out,input A,B,C);

always@(C,B,A)
   begin
      case({C,B,A})
        3'b000: {out} = 1'b0;
        3'b001: {out} = 1'b0;
        3'b010: {out} = 1'b1;
        3'b011: {out} = 1'b0;
        3'b100: {out} = 1'b0;
        3'b101: {out} = 1'b0;
        3'b110: {out} = 1'b0;
        3'b111: {out} = 1'b1;
      endcase
   end
endmodule
```

SBOL Community



128 people from 16 countries

Representing 43 universities and 29 companies and government labs

Organizations Supporting SBOL

NSF

EPSRC
Engineering and Physical Sciences
Research Council

U.S. DEPARTMENT OF ENERGY Office of Science

FLOWERS

JGI JOINT GENOME INSTITUTE DEPARTMENT OF ENERGY

jbei Joint BioEnergy Institute

National Human Genome Research Institute

NATIONAL LIBRARY OF MEDICINE

BDAC BioDesign Automation Consortium

Agilent Technologies

AMYRIS

AUTODESK.

CLARK PARSIA

DNA 2.0

Genome Compiler

life technologies

Raytheon BBN Technologies

teselagen BIOTECHNOLOGY

ThermoFisher SCIENTIFIC

Berkeley UNIVERSITY OF CALIFORNIA

BOSTON UNIVERSITY

Imperial College London

Newcastle University

THE UNIVERSITY OF UTAH

W UNIVERSITY of WASHINGTON

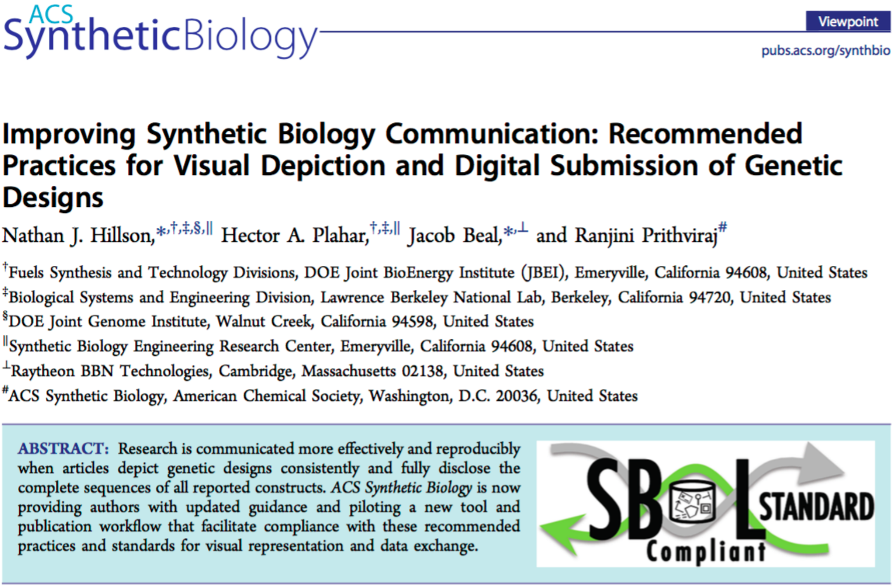
BERKELEY LAB

BU CoSBi

I(CO)$_2$S

Current support for the development of SBOL provided by National Science Foundation Grants DBI-1356041 and DBI-1355909, and the Engineering and Physical Sciences Research Council under Grant Number EP/J02175X/1.

ACS Synthetic Biology Recommends Use of SBOL in 2016



ACS SyntheticBiology

Viewpoint

pubs.acs.org/synthbio

Improving Synthetic Biology Communication: Recommended Practices for Visual Depiction and Digital Submission of Genetic Designs

Nathan J. Hillson,[*,†,‡,§,||] Hector A. Plahar,[†,‡,||] Jacob Beal,[*,⊥] and Ranjini Prithviraj[#]

[†] Fuels Synthesis and Technology Divisions, DOE Joint BioEnergy Institute (JBEI), Emeryville, California 94608, United States

[‡] Biological Systems and Engineering Division, Lawrence Berkeley National Lab, Berkeley, California 94720, United States

[§] DOE Joint Genome Institute, Walnut Creek, California 94598, United States

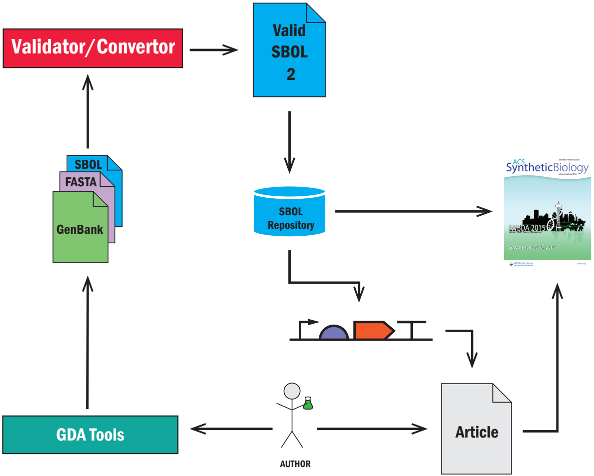
[||] Synthetic Biology Engineering Research Center, Emeryville, California 94608, United States

[⊥] Raytheon BBN Technologies, Cambridge, Massachusetts 02138, United States

[#] ACS Synthetic Biology, American Chemical Society, Washington, D.C. 20036, United States

ABSTRACT: Research is communicated more effectively and reproducibly when articles depict genetic designs consistently and fully disclose the complete sequences of all reported constructs. *ACS Synthetic Biology* is now providing authors with updated guidance and piloting a new tool and publication workflow that facilitate compliance with these recommended practices and standards for visual representation and data exchange.

ACS Synthetic Biology SBOL Workflow



Validator/Convertor

Valid SBOL 2

SBOL

FASTA

GenBank

SBOL Repository

GDA Tools

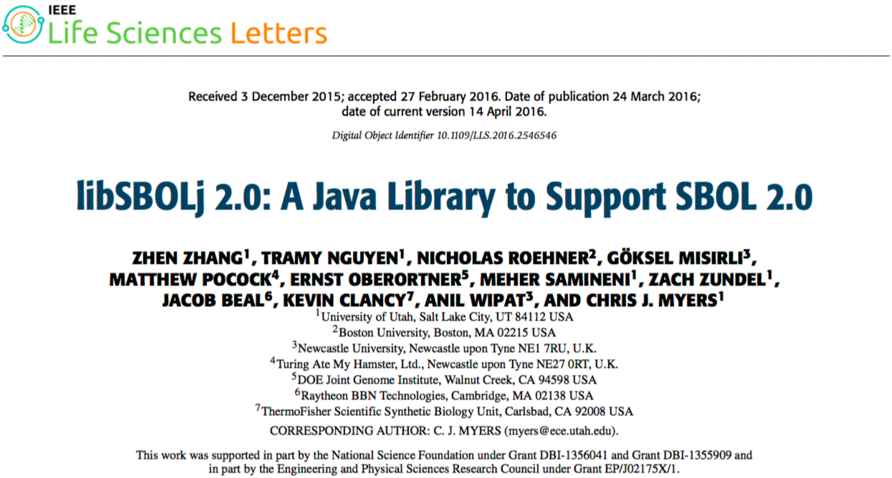
AUTHOR

Article

Hillson et al., *ACS Synthetic Biology* (2016)

Zundel et al., *ACS Synthetic Biology* (2017)

libSBOLj Version 2 Released in Late 2015



IEEE
Life Sciences Letters

Received 3 December 2015; accepted 27 February 2016. Date of publication 24 March 2016;
date of current version 14 April 2016.

Digital Object Identifier 10.1109/LLS.2016.2546546

libSBOLj 2.0: A Java Library to Support SBOL 2.0

ZHEN ZHANG[1], TRAMY NGUYEN[1], NICHOLAS ROEHNER[2], GÖKSEL MISIRLI[3], MATTHEW POCOCK[4], ERNST OBERORTNER[5], MEHER SAMINENI[1], ZACH ZUNDEL[1], JACOB BEAL[6], KEVIN CLANCY[7], ANIL WIPAT[3], AND CHRIS J. MYERS[1]
[1]University of Utah, Salt Lake City, UT 84112 USA
[2]Boston University, Boston, MA 02215 USA
[3]Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.
[4]Turing Ate My Hamster, Ltd., Newcastle upon Tyne NE27 0RT, U.K.
[5]DOE Joint Genome Institute, Walnut Creek, CA 94598 USA
[6]Raytheon BBN Technologies, Cambridge, MA 02138 USA
[7]ThermoFisher Scientific Synthetic Biology Unit, Carlsbad, CA 92008 USA
CORRESPONDING AUTHOR: C. J. MYERS (myers@ece.utah.edu).

This work was supported in part by the National Science Foundation under Grant DBI-1356041 and Grant DBI-1355909 and in part by the Engineering and Physical Sciences Research Council under Grant EP/J02175X/1.

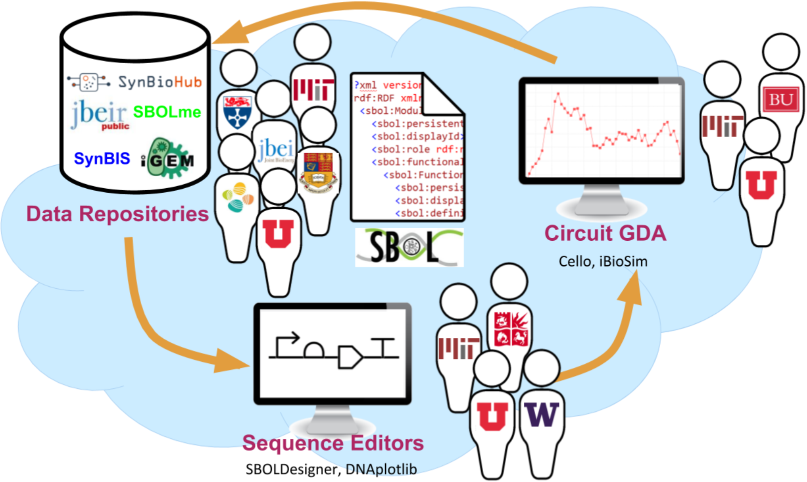
Other SBOL libraries available for C++, Python, and Javascript

Software Supporting SBOL

	Function					**SBOL**		
Name	R	S	V	G	M	1	2	v
Benchling		●				●		
BOOST		●				●	●	
Cello				●			●	
DeviceEditor		●	●			●		●
DNAPlotLib			●			●		●
Eugene		●				●		●
Finch		●	●	●			●	●
GenoCAD		●	●					●
GeneGenie		●				●		
Graphviz			●					●
ICE	●		●			●	●	●
iBioSim		●	●	●	●	●	●	●
j5		●						
MoSeC		●			●	●		
Pigeon			●					●
Pinecone		●						●
Pool Designer		●				●	●	
Proto BioCompiler			●	●		●		●
SBOLDesigner		●	●			●	●	●
SBOLme	●						●	
ShortBol		●		●			●	
SynBioHub	●		●			●	●	●
Tellurium					●		●	
TeselaGen		●	●			●		●
TinkerCell			●	●	●	●		●
VisBOL			●				●	●
VirtualParts	●				●		●	

`http://sbolstandard.org/software/tools/`

Synthetic Biology Workflow Using SBOL



Data Repositories

Circuit GDA

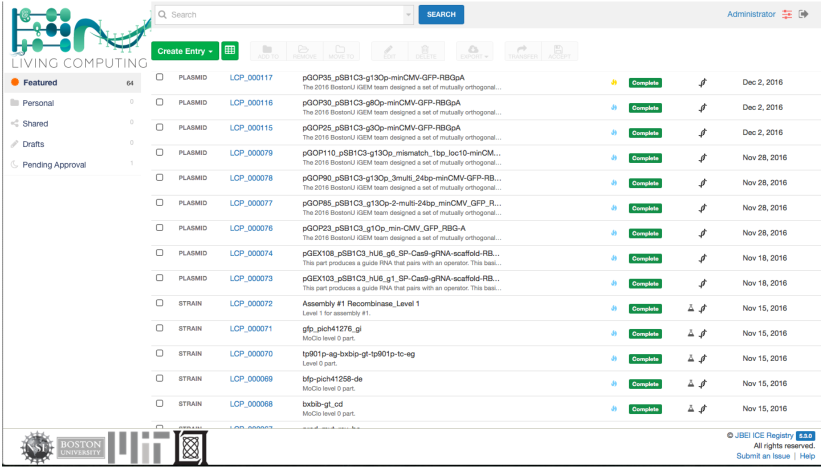
Cello, iBioSim

Sequence Editors

SBOLDesigner, DNAplotlib

Myers et al., *Biochemical Society Transactions* (2017).

Data Repositories (ICE)

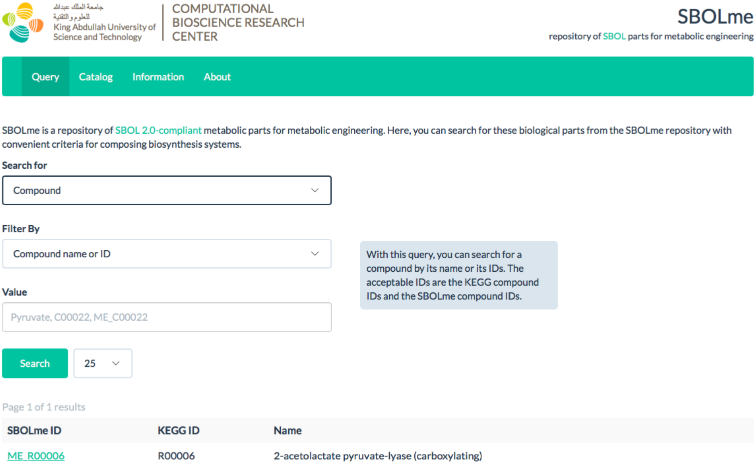


☐	PLASMID	LCP_000117	pGOP35_pSB1C3-g13Op-minCMV-GFP-RBGpA The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	Dec 2, 2016
☐	PLASMID	LCP_000116	pGOP30_pSB1C3-g8Op-minCMV-GFP-RBGpA The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	Dec 2, 2016
☐	PLASMID	LCP_000115	pGOP25_pSB1C3-g3Op-minCMV-GFP-RBGpA The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	Dec 2, 2016
☐	PLASMID	LCP_000079	pGOP110_pSB1C3-g13Op_mismatch_1bp_loc10-minCM... The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	Nov 28, 2016
☐	PLASMID	LCP_000078	pGOP90_pSB1C3-g13Op_3multi_24bp-minCMV-GFP-RB... The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	Nov 28, 2016
☐	PLASMID	LCP_000077	pGOP85_pSB1C3_g13Op-2-multi-24bp_minCMV_GFP_R... The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	Nov 28, 2016
☐	PLASMID	LCP_000076	pGOP23_pSB1C3_g1Op_min-CMV_GFP_RBG-A The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	Nov 28, 2016
☐	PLASMID	LCP_000074	pGEX108_pSB1C3_hU6_g6_SP-Cas9-gRNA-scaffold-RB... This part produces a guide RNA that pairs with an operator. This basi...	Complete	Nov 18, 2016
☐	PLASMID	LCP_000073	pGEX103_pSB1C3_hU6_g1_SP-Cas9-gRNA-scaffold-RB... This part produces a guide RNA that pairs with an operator. This basi...	Complete	Nov 18, 2016
☐	STRAIN	LCP_000072	Assembly #1 Recombinase_Level 1 Level 1 for assembly #1.	Complete	Nov 15, 2016
☐	STRAIN	LCP_000071	gfp_pich41276_gi MoClo level 0 part.	Complete	Nov 15, 2016
☐	STRAIN	LCP_000070	tp901p-ag-bxbip-gt-tp901p-tc-eg Level 0 part.	Complete	Nov 15, 2016
☐	STRAIN	LCP_000069	bfp-pich41258-de MoClo level 0 part.	Complete	Nov 15, 2016
☐	STRAIN	LCP_000068	bxbib-gt_cd MoClo level 0 part.	Complete	Nov 15, 2016

© JBEI ICE Registry 5.3.0
All rights reserved.
Submit an Issue | Help

Ham et al., *Nucleic Acid Research* (2012)

Data Repositories (SBOLme)



SBOLme is a repository of SBOL 2.0-compliant metabolic parts for metabolic engineering. Here, you can search for these biological parts from the SBOLme repository with convenient criteria for composing biosynthesis systems.

Search for: Compound

Filter By: Compound name or ID

Value: Pyruvate, C00022, ME_C00022

With this query, you can search for a compound by its name or its IDs. The acceptable IDs are the KEGG compound IDs and the SBOLme compound IDs.

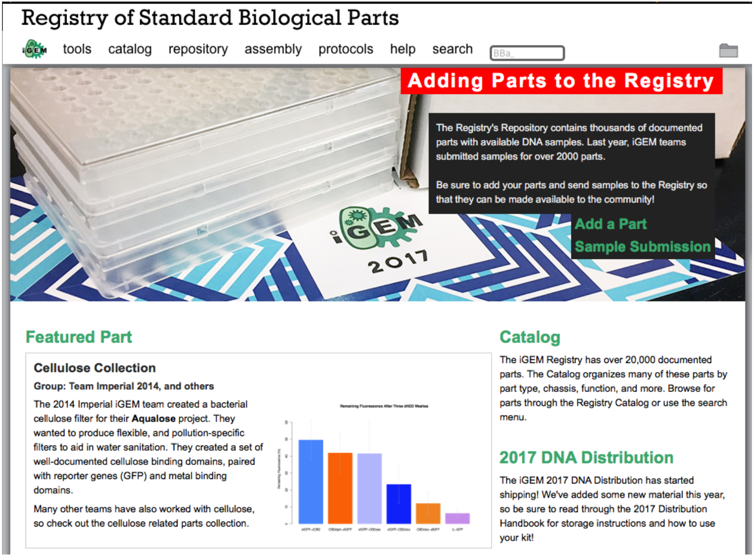
Page 1 of 1 results

SBOLme ID	KEGG ID	Name
ME_R00006	R00006	2-acetolactate pyruvate-lyase (carboxylating)

Kuwahara et al., *ACS Synthetic Biology* (2017)

Includes 28,437 chemical compounds, 6,883 enzyme classes, 9,909 metabolic reactions, and 3,173,238 proteins from 3,908 organisms.

iGEM Registry of Standard Biological Parts (BioBricks)



Registry of Standard Biological Parts

Adding Parts to the Registry

The Registry's Repository contains thousands of documented parts with available DNA samples. Last year, iGEM teams submitted samples for over 2000 parts.

Be sure to add your parts and send samples to the Registry so that they can be made available to the community!

Featured Part

Cellulose Collection

Group: Team Imperial 2014, and others

The 2014 Imperial iGEM team created a bacterial cellulose filter for their **Aqualose** project. They wanted to produce flexible, and pollution-specific filters to aid in water sanitation. They created a set of well-documented cellulose binding domains, paired with reporter genes (GFP) and metal binding domains.

Many other teams have also worked with cellulose, so check out the cellulose related parts collection.

Catalog

The iGEM Registry has over 20,000 documented parts. The Catalog organizes many of these parts by part type, chassis, function, and more. Browse for parts through the Registry Catalog or use the search menu.

2017 DNA Distribution

The iGEM 2017 DNA Distribution has started shipping! We've added some new material this year, so be sure to read through the 2017 Distribution Handbook for storage instructions and how to use your kit!

`http://parts.igem.org`

International Genetically Engineered Machine (iGEM) Competition

Started in 2004 with 5 teams and 31 participants.

In 2017: 310 teams with nearly 5400 participants from 44 countries.

International Genetically Engineered Machine (iGEM) Competition

Started in 2004 with 5 teams and 31 participants.

In 2017: 310 teams with nearly 5400 participants from 44 countries.

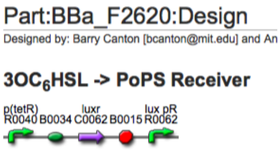
International Genetically Engineered Machine (iGEM) Competition

Started in 2004 with 5 teams and 31 participants.

In 2017: 310 teams with nearly 5400 participants from 44 countries.

Conversion of the iGEM Registry to SBOL

Conversion of the iGEM Registry to SBOL



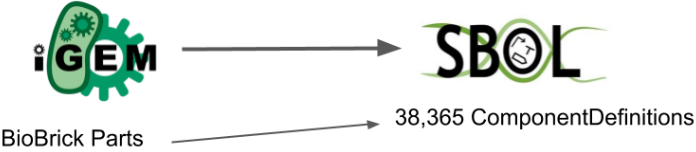
Part:BBa_F2620:Design

Designed by: Barry Canton [bcanton@mit.edu] and An

3OC$_6$HSL -> PoPS Receiver

ComponentDefinition

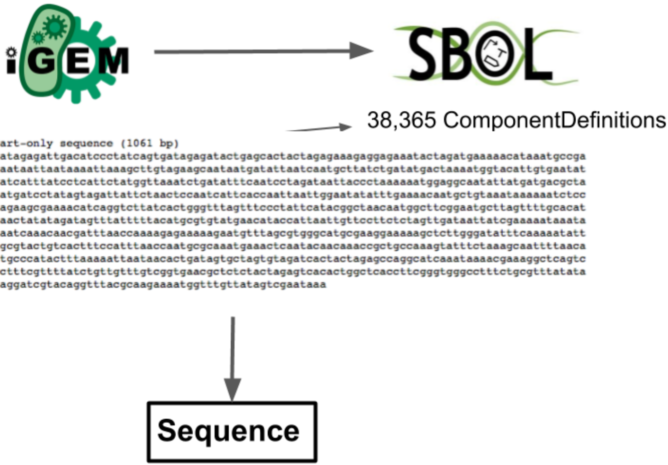
Conversion of the iGEM Registry to SBOL



38,365 ComponentDefinitions

BioBrick Parts

Conversion of the iGEM Registry to SBOL



38,365 ComponentDefinitions

```
>BBa_F2620 Part-only sequence (1061 bp)
tccctatcagtgatagagattgacatccctatcagtgatagagatactgagcactactagagaaagaggagaaatactagatgaaaaacataaatgccga
cgacacatacagaataattaataaaattaaagcttgtagaagcaataatgatattaatcaatgcttatctgatatgactaaaatggtacattgtgaatat
tatttactcgcgatcatttatcctcattctatggttaaatctgatatttcaatcctagataattaccctaaaaaatggaggcaatattatgatgacgcta
atttaataaaatatgatcctatagtagattattctaactccaatcattcaccaattaattggaatatatttgaaaacaatgctgtaaataaaaaatctcc
aaatgtaattaaagaagcgaaaacatcaggtcttatcactgggtttagtttccctattcatacggctaacaatggcttcggaatgcttagttttgcacat
tcagaaaaagacaactatatagatagtttatttttacatgcgtgtatgaacataccattaattgttccttctctagttgataattatcgaaaaataaata
tagcaaataataaatcaaacaacgatttaaccaaaagagaaaaagaatgtttagcgtgggcatgcgaaggaaaaagctcttgggatatttcaaaaatatt
aggttgcagtgagcgtactgtcactttccatttaaccaatgcgcaaatgaaactcaatacaacaaaccgctgccaaagtatttctaaagcaattttaaca
ggagcaattgattgcccatactttaaaaattaataacactgatagtgctagtgtagatcactactagagccaggcatcaaataaaacgaaaggctcagtc
gaaagactgggcctttcgttttatctgttgtttgtcggtgaacgctctctactagagtcacactggctcaccttcgggtgggcctttctgcgtttatata
ctagagacctgtaggatcgtacaggtttacgcaagaaaatggtttgttatagtcgaataaa
```

Sequence

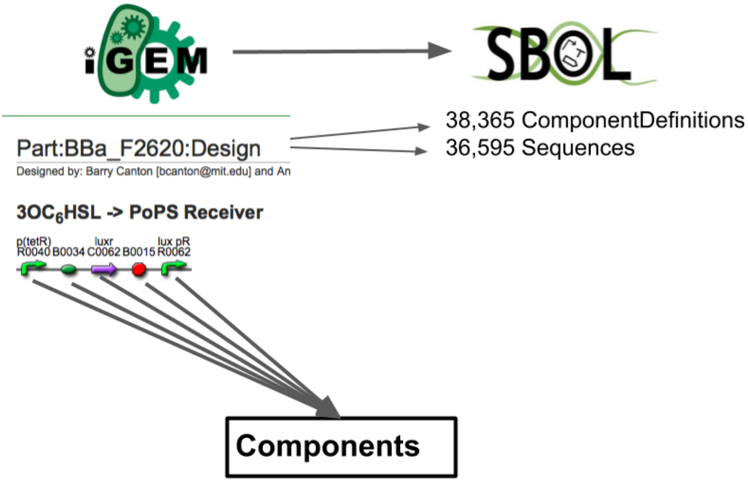
Conversion of the iGEM Registry to SBOL

iGEM → SBOL

BioBrick Parts → 38,365 ComponentDefinitions

BioBrick Parts → 36,595 Sequences

Conversion of the iGEM Registry to SBOL



38,365 ComponentDefinitions
36,595 Sequences

Part:BBa_F2620:Design

Designed by: Barry Canton [bcanton@mit.edu] and An

3OC$_6$HSL -> PoPS Receiver

Components

Conversion of the iGEM Registry to SBOL

iGEM → SBOL

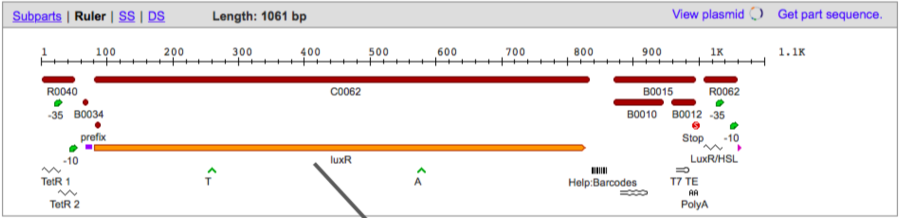
BioBrick Parts → 38,365 ComponentDefinitions

BioBrick Parts → 36,595 Sequences

BioBrick Uses → 70,493 Components

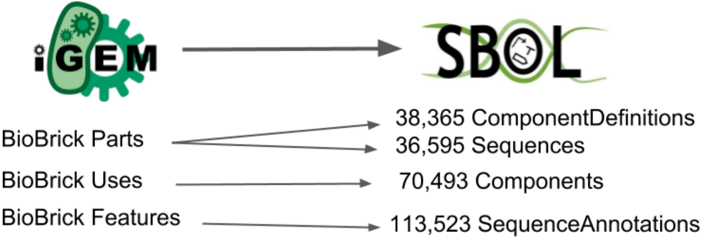
Conversion of the iGEM Registry to SBOL

38,365 ComponentDefinitions



SequenceAnnotations

Conversion of the iGEM Registry to SBOL



BioBrick Parts → 38,365 ComponentDefinitions, 36,595 Sequences

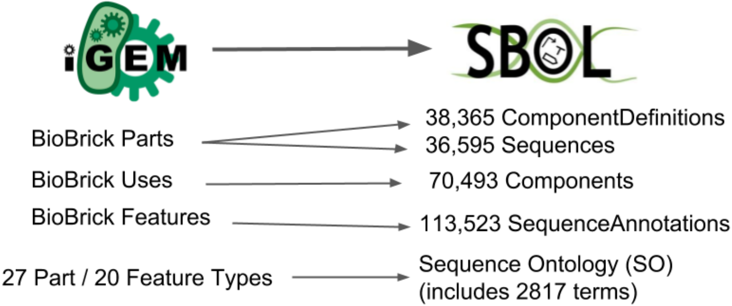
BioBrick Uses → 70,493 Components

BioBrick Features → 113,523 SequenceAnnotations

Conversion of the iGEM Registry to SBOL

iGEM Part/Feature Type	**SequenceOntology (SO) Term**
Coding	CDS (SO:0000316)
Device	Engineered Region (SO:0000804)
Primer	Primer (SO:0000112)
Protein Domain	Polypeptide Domain (SO:0000417)
RBS	Ribosome Entry Site (SO:0000139)
Regulatory	Promoter (SO:0000167)
Tag	Tag (SO:0000324)
Terminator	Terminator (SO:0000141)
etc.	etc.

Conversion of the iGEM Registry to SBOL



iGEM → SBOL

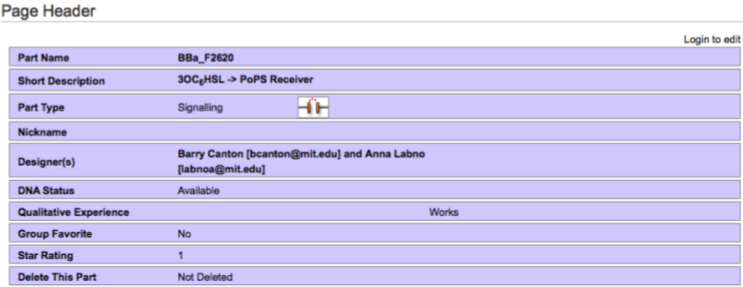
BioBrick Parts → 38,365 ComponentDefinitions, 36,595 Sequences

BioBrick Uses → 70,493 Components

BioBrick Features → 113,523 SequenceAnnotations

27 Part / 20 Feature Types → Sequence Ontology (SO) (includes 2817 terms)

Conversion of the iGEM Registry to SBOL



Page Header

Login to edit

Part Name	BBa_F2620
Short Description	3OC$_6$HSL -> PoPS Receiver
Part Type	Signalling
Nickname	
Designer(s)	Barry Canton [bcanton@mit.edu] and Anna Labno [labnoa@mit.edu]
DNA Status	Available
Qualitative Experience	Works
Group Favorite	No
Star Rating	1
Delete This Part	Not Deleted

Dublin Core/iGEM Annotations

Conversion of the iGEM Registry to SBOL

iGEM → SBOL

BioBrick Parts → 38,365 ComponentDefinitions

BioBrick Parts → 36,595 Sequences

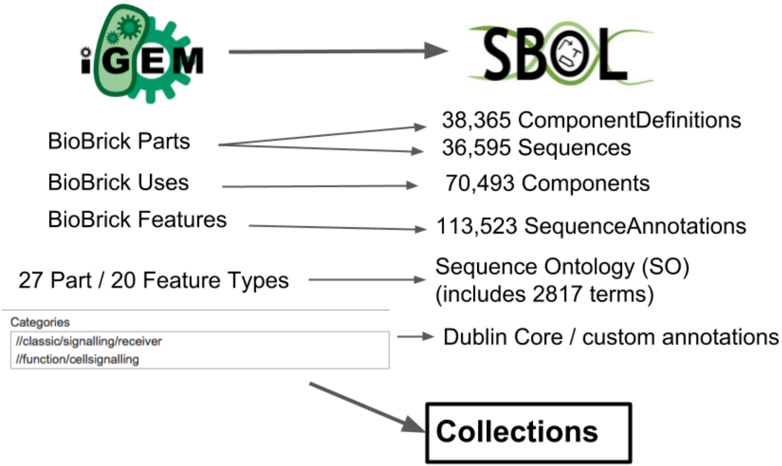
BioBrick Uses → 70,493 Components

BioBrick Features → 113,523 SequenceAnnotations

27 Part / 20 Feature Types → Sequence Ontology (SO) (includes 2817 terms)

Additional metadata → Dublin Core / iGEM annotations

Conversion of the iGEM Registry to SBOL



BioBrick Parts → 38,365 ComponentDefinitions, 36,595 Sequences

BioBrick Uses → 70,493 Components

BioBrick Features → 113,523 SequenceAnnotations

27 Part / 20 Feature Types → Sequence Ontology (SO) (includes 2817 terms)

Categories

//classic/signalling/receiver

//function/cellsignalling

→ Dublin Core / custom annotations

→ **Collections**

Conversion of the iGEM Registry to SBOL

iGEM → SBOL

BioBrick Parts	38,365 ComponentDefinitions 36,595 Sequences
BioBrick Uses	70,493 Components
BioBrick Features	113,523 SequenceAnnotations
27 Part / 20 Feature Types	Sequence Ontology (SO) (includes 2817 terms)
Additional metadata	Dublin Core / custom annotations
Categories	372 Collections

Additional Features of SBOL

- Tractable hierarchical representation of part-subpart relations.
- Systematic detection of inconsistencies in parts with SBOL tooling.
  - Found 521 instances of mis-aligned annotations.
  - Found 2285 instances of composite BioBricks with inconsistent sequences with respect to their basic BioBricks.
- Builds on standard ontologies, such as the Sequence Ontology.
- Enables powerful search queries using Semantic Web technology.
- Can convey versioning and rich provenance information via PROV-O.
- Enables users of the registry to take advantage of the emerging ecosystem of SBOL-enabled software tools.

Data Repositories (SynBioHub)

SynBioHub is a *design repository* for people designing biological constructs. It enables DNA and protein designs to be uploaded, then provides a shareable link to allow others to view them. SynBioHub also facilitates searching for information about existing useful parts and designs by combining data from a variety of sources.

Search for useful parts and designs

Upload your design for safekeeping

Share designs for publication or collaboration

Newcastle University

James McLaughlin
Anil Wipat

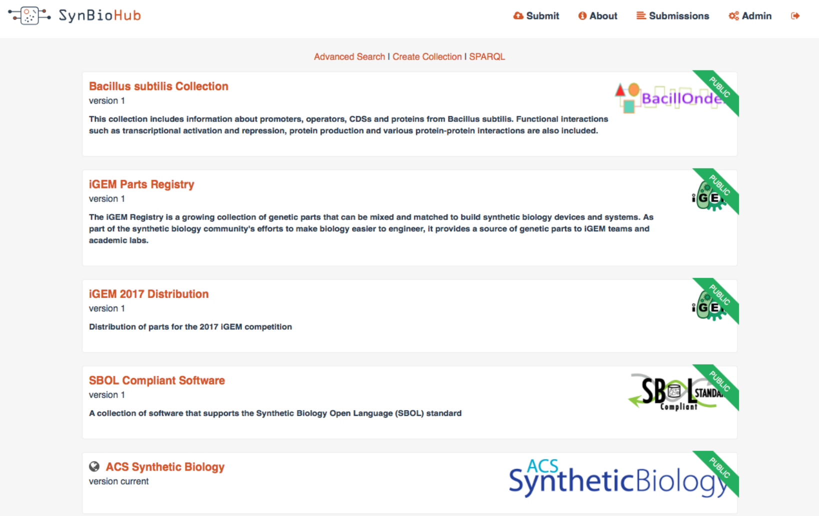
THE UNIVERSITY OF UTAH®

Zach Zundel
Chris Myers

Version 1.0 released
June 14, 2017

McLaughlin et al., *ACS Synthetic Biology* (2018).

Reference Instance (`https://synbiohub.org`)



Advanced Search | Create Collection | SPARQL

Bacillus subtilis Collection
version 1

This collection includes information about promoters, operators, CDSs and proteins from Bacillus subtilis. Functional interactions such as transcriptional activation and repression, protein production and various protein-protein interactions are also included.

iGEM Parts Registry
version 1

The iGEM Registry is a growing collection of genetic parts that can be mixed and matched to build synthetic biology devices and systems. As part of the synthetic biology community's efforts to make biology easier to engineer, it provides a source of genetic parts to iGEM teams and academic labs.

iGEM 2017 Distribution
version 1

Distribution of parts for the 2017 iGEM competition

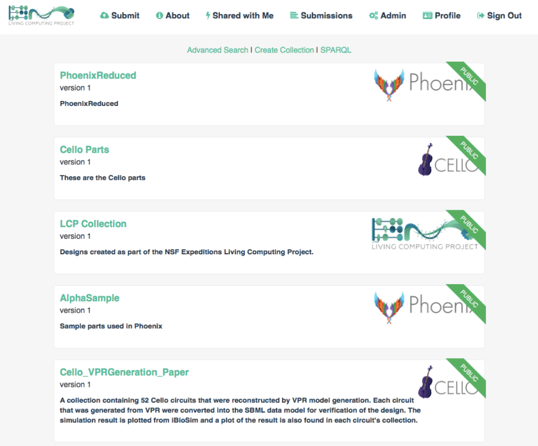
SBOL Compliant Software
version 1

A collection of software that supports the Synthetic Biology Open Language (SBOL) standard

ACS Synthetic Biology
version current

McLaughlin et al., *ACS Synthetic Biology* (2018).

NSF Expeditions Living Computing Project
(`https://synbiohub.programmingbiology.org`)



PhoenixReduced
version 1
PhoenixReduced

Cello Parts
version 1
These are the Cello parts

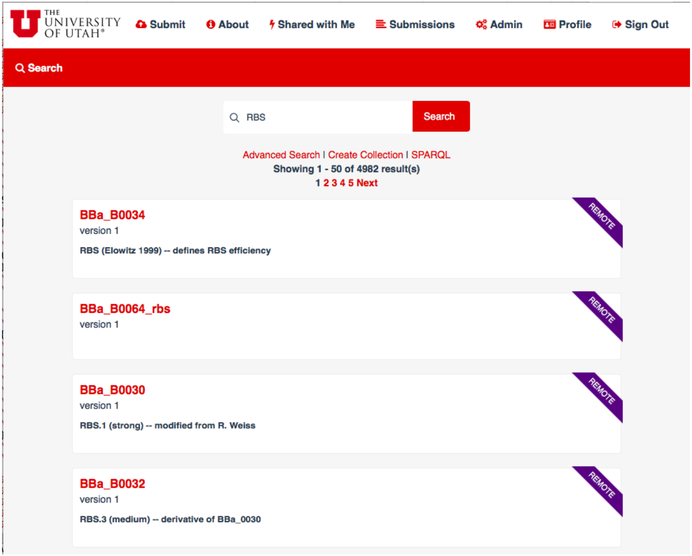
LCP Collection
version 1
Designs created as part of the NSF Expeditions Living Computing Project.

AlphaSample
version 1
Sample parts used in Phoenix

Cello_VPRGeneration_Paper
version 1
A collection containing 52 Cello circuits that were reconstructed by VPR model generation. Each circuit that was generated from VPR were converted into the SBML data model for verification of the design. The simulation result is plotted from iBioSim and a plot of the result is also found in each circuit's collection.

McLaughlin et al., *ACS Synthetic Biology* (2018).

SBOLExplorer (`https://synbiohub.utah.edu`)



Advanced Search | Create Collection | SPARQL

Showing 1 - 50 of 4982 result(s)

1 2 3 4 5 Next

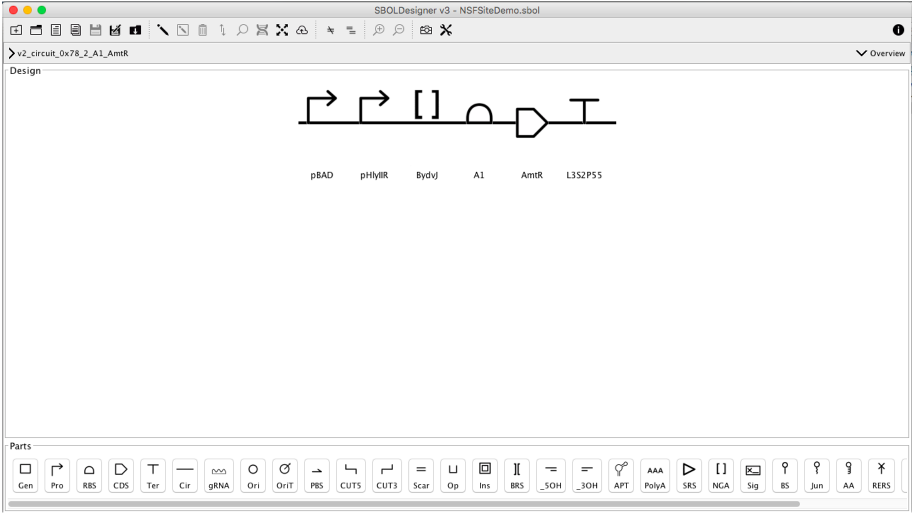
BBa_B0034
version 1
RBS (Elowitz 1999) -- defines RBS efficiency

BBa_B0064_rbs
version 1

BBa_B0030
version 1
RBS.1 (strong) -- modified from R. Weiss

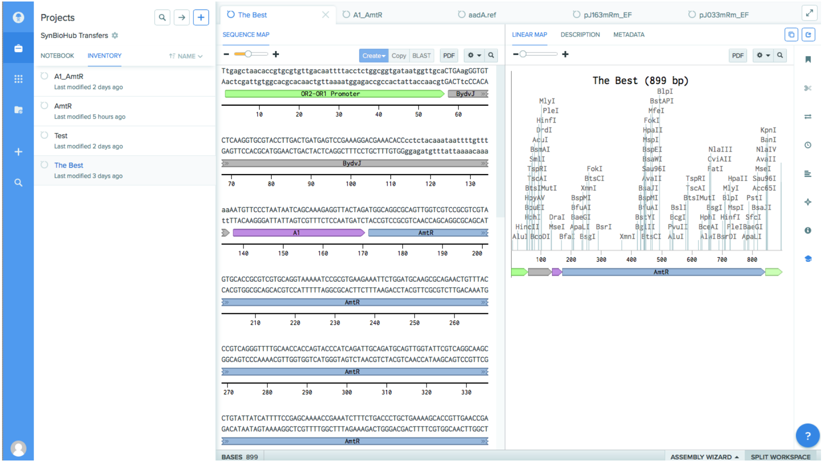
BBa_B0032
version 1
RBS.3 (medium) -- derivative of BBa_0030

Sequence Editors (SBOLDesigner)



Zhang et al., *ACS Synthetic Biology* (2017)

Sequence Editors (Benchling)

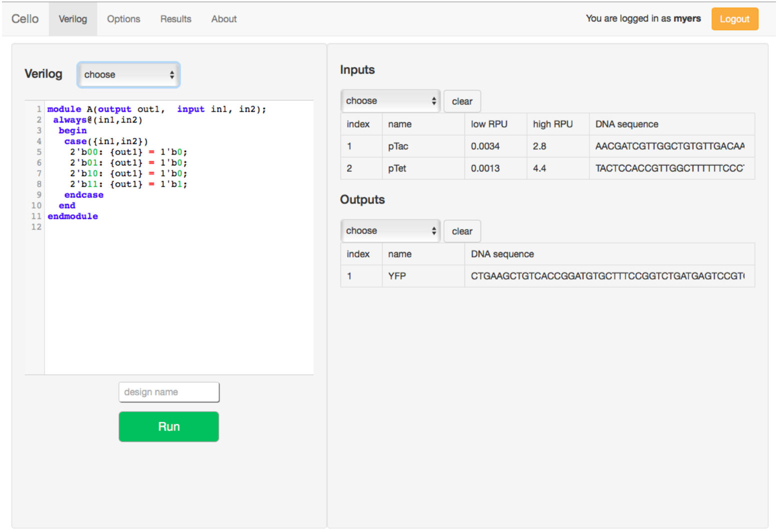


```
TtgagctaacaccgtgcgtgttgacaattttacctctggcggtgataatggttgcaCTGAagGGTGT
AactcgattgtggcacgcacaactgttaaaatggagaccgccactattaccaacgtGACTtcCCACA
       10        20        30        40        50        60

CTCAAGGTGCGTACCTTGACTGATGAGTCCGAAAGGACGAAACACCcctctacaaataattttgttt
GAGTTCCACGCATGGAACTGACTACTCAGGCTTTCCTGCTTTGTGGgagatgtttattaaaacaaa
 70        80        90        100       110       120       130

aaAATGTTCCCTAATAATCAGCAAAGAGGTTACTAGATGGCAGGCGCAGTTGGTCGTCCGCGTCGTA
ttTTACAAGGGATTATTAGTCGTTTCTCCAATGATCTACCGTCCGCGTCAACCAGCAGGCGCAGCAT
    140       150       160       170       180       190       200

GTGCACCGCGTCGTGCAGGTAAAAATCCGCGTGAAGAAATTCTGGATGCAAGCGCAGAACTGTTTAC
CACGTGGCGCAGCACGTCCATTTTTAGGCGCACTTCTTTAAGACCTACGTTCGCGTCTTGACAAATG
        210       220       230       240       250       260

CCGTCAGGGTTTTGCAACCACCAGTACCCATCAGATTGCAGATGCAGTTGGTATTCGTCAGGCAAGC
GGCAGTCCCAAAACGTTGGTGGTCATGGGTAGTCTAACGTCTACGTCAACCATAAGCAGTCCGTTCG
270       280       290       300       310       320       330

CTGTATTATCATTTTCCGAGCAAAACCGAAATCTTTCTGACCCTGCTGAAAAGCACCGTTGAACCGA
GACATAATAGTAAAAGGCTCGTTTTGGCTTTAGAAAGACTGGGACGACTTTTCGTGGCAACTTGGCT
```

Other sequence editors that support SBOL:

DeviceEditor, J5, VectorEditor (JBEI), DNAPlotLib (MIT/UW/Bristol), Eugene (Boston), GenoCAD (VBI), BOOST (JGI), etc.

Circuit GDA Tools (Cello)



```
module A(output out1,  input in1, in2);
 always@(in1,in2)
  begin
   case({in1,in2})
    2'b00: {out1} = 1'b0;
    2'b01: {out1} = 1'b0;
    2'b10: {out1} = 1'b0;
    2'b11: {out1} = 1'b1;
   endcase
  end
endmodule
```

Inputs

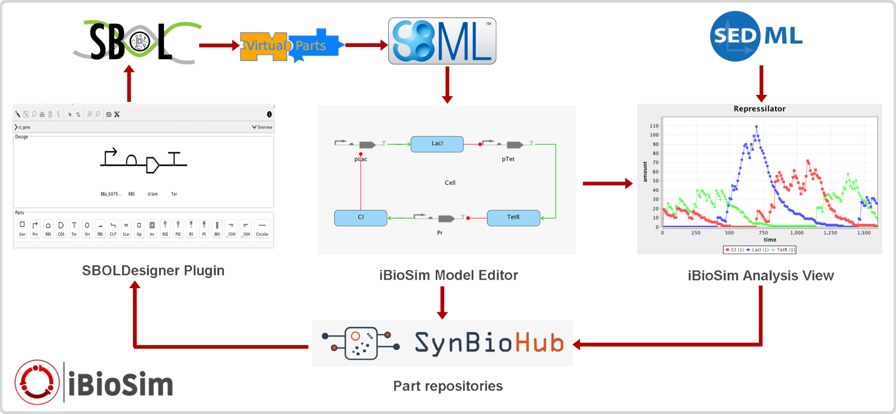
index	name	low RPU	high RPU	DNA sequence
1	pTac	0.0034	2.8	AACGATCGTTGGCTGTGTTGACAA
2	pTet	0.0013	4.4	TACTCCACCGTTGGCTTTTTTCCC

Outputs

index	name	DNA sequence
1	YFP	CTGAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGT

Nielsen et al., *Science* (2016)

Circuit GDA Tools (iBioSim)



SBOLDesigner Plugin

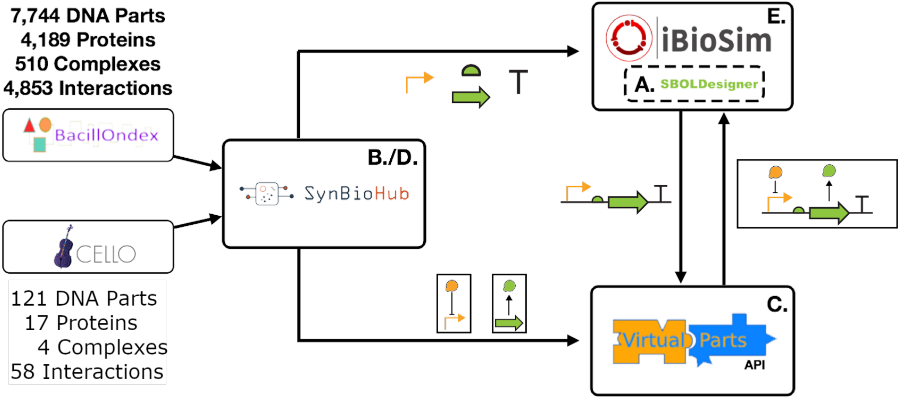
iBioSim Model Editor

iBioSim Analysis View

Part repositories

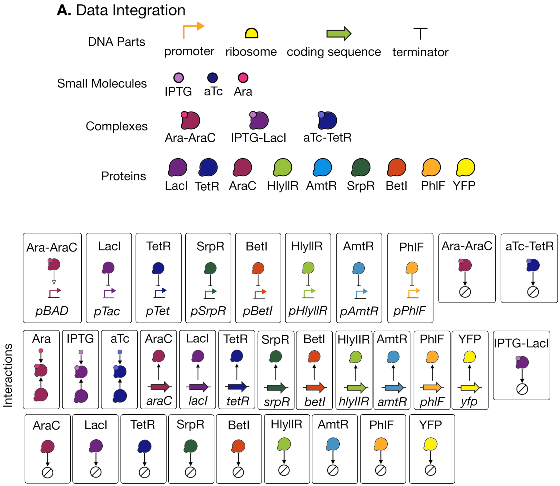
Myers et al., *Bioinformatics* (2009)
Madsen et al., *IEEE Design & Test* (2012)
Watanabe et al., *ACS Synthetic Biology* (2018)

Model Generation Workflow



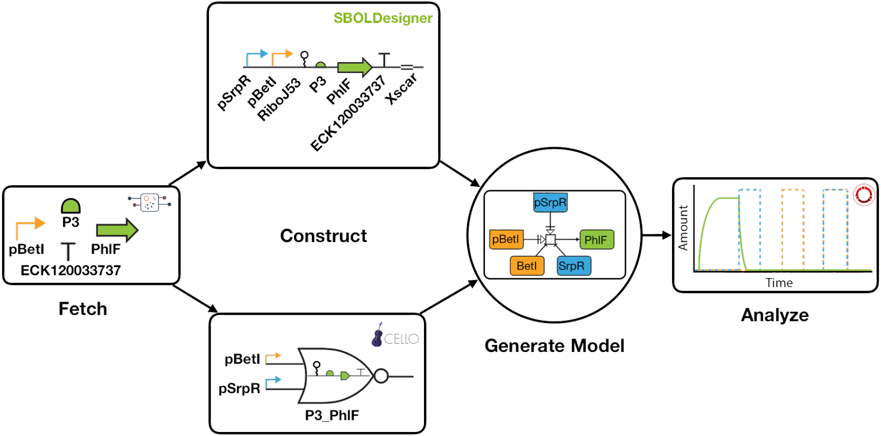
Mısırlı et al., *ACS Synthetic Biology* (2018).

Data Integration: Cello Part Library



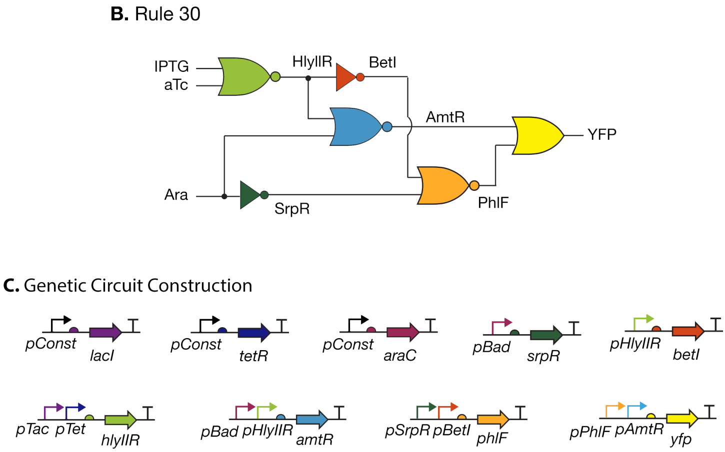
Mısırlı et al., *ACS Synthetic Biology* (2018).

Genetic Circuit Construction



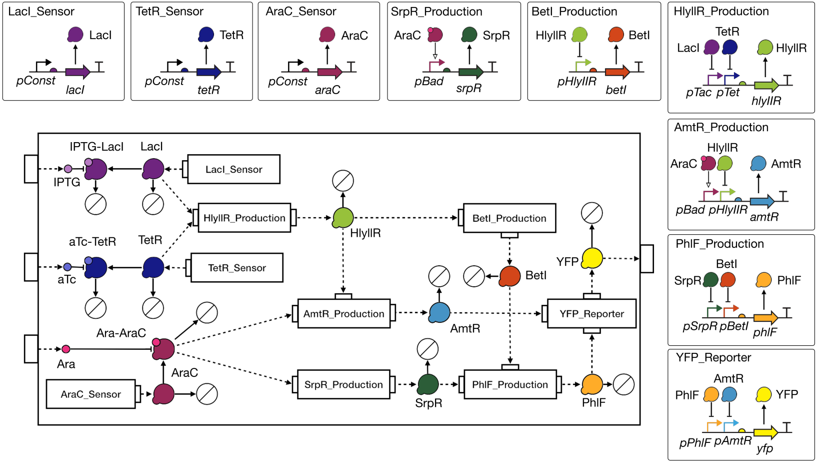
Mısırlı et al., *ACS Synthetic Biology* (2018).

Genetic Circuit Construction: Rule 30 Example



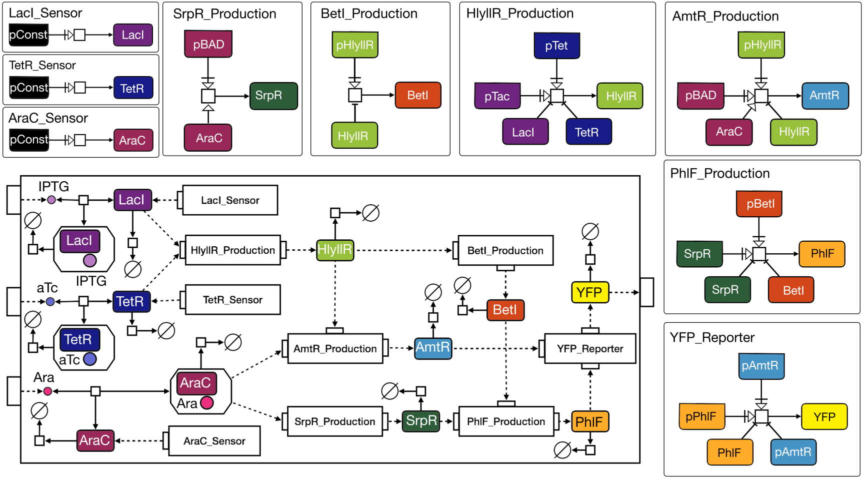
Mısırlı et al., *ACS Synthetic Biology* (2018).

Enriched SBOL Representation: Rule 30 Example



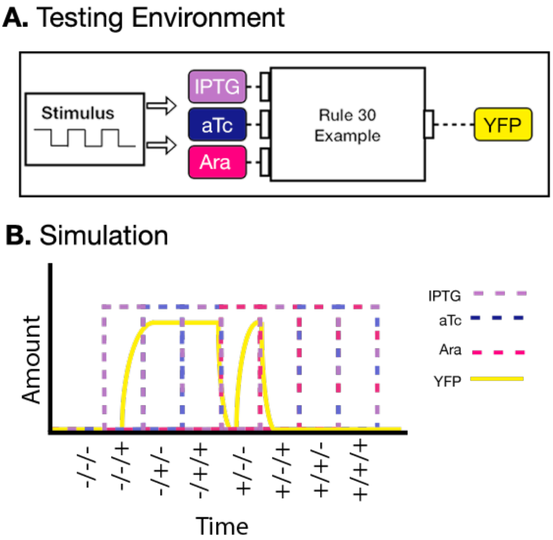
Mısırlı et al., *ACS Synthetic Biology* (2018).

Dynamic SBML Model: Rule 30 Example



Mısırlı et al., *ACS Synthetic Biology* (2018).

Simulation: Rule 30 Example



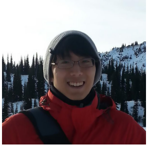
Mısırlı et al., *ACS Synthetic Biology* (2018).

More Information

- Standards are an important enabler for data sharing and reproducibility.
- Ultimate goal should be a complete standard-enabled workflow.
- Much more information is available from: `http://www.sbolstandard.org/`.
- Check out the SBOL Standard Youtube channel.
- SBOL is also on Facebook, Twitter, and Vimeo.

Acknowledgements (University of Washington)

Dr. Bryan Bartley
(Now with BBN/Raytheon)



Kiri Choi

Dr. Herbert Sauro

Acknowledgements (Newcastle University)

James McLaughlin

Dr. Goksel Misirli
(Now with Keele University)

Dr Angel Goni-Moreno

Prof. Anil Wipat

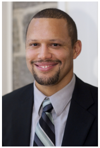
Acknowledgements (Boston University)

Dr. Timothy Jones

Dr. Curtis Madsen

Dr. Nicholas Roehner
(Now with BBN/Raytheon)

Prashant Vaidyanathan



Prof. Douglas Densmore

Acknowledgements (Other Institutions)

Dr. Jacob Beal
(BBN/Raytheon)

Dr. Thomas Gorochowski
(University of Bristol)

Dr Hiroyuki Kuwahara
(KAUST)

Dr. Ernst Oberortner
(Joint Genome Institute)

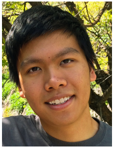
Acknowledgements (University of Utah)


Pedro Fontanarrosa


Tramy Nguyen


Meher Samineni


Leandro Watanabe


Michael Zhang


Zach Zundel

Supported by National Science Foundation Grants CCF-1218095, DBI-1356041, CCF-1748200, and CCF-1522074 (sub-award from BU), and DARPA FA8750-17-C-0229 (sub-award from MIT).