

Synthetic Biology Open Language Applications

Presenter: Dr Göksel Mısırlı

School of Computing and Mathematics



Keele
UNIVERSITY

SBOL Editors: Hasan Baig (Habib University), Vishwesh Kulkarni (University of Warwick), Curtis Madsen (Boston University), James McLaughlin (Newcastle University), Zach Palchick (Zymergen)

SBOL Chair: Chris Myers (University of Utah)

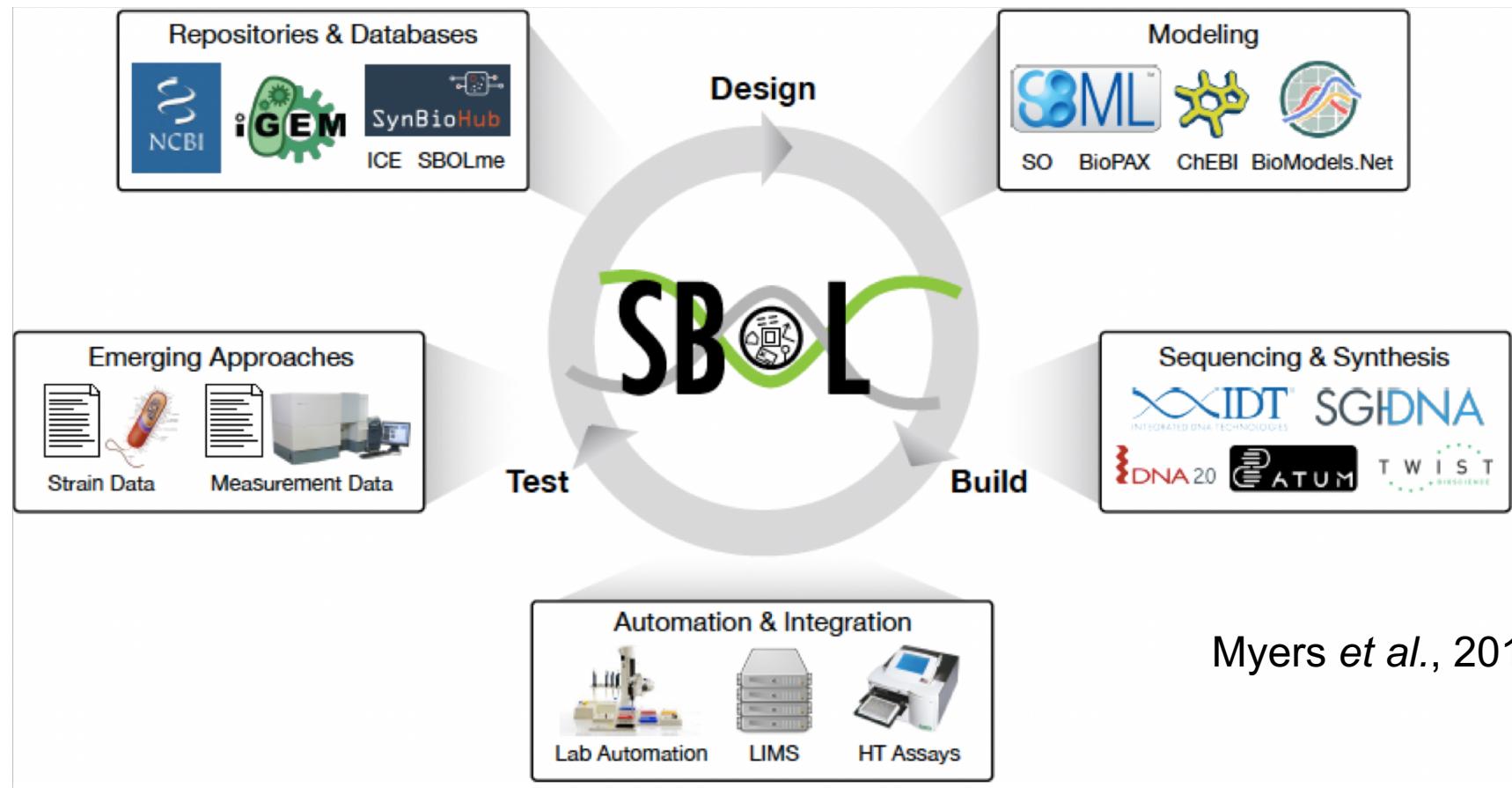
SBOL Workshop

July 8, 2019



Slides were provided by Jacob Beal, Chris Myers and Anil Wipat
These slides were adapted and new slides were added by Göksel Mısırlı

Collaborating across different stages of (automation-assisted) synbio workflows



- Unambiguous **exchange** of information
- **Reuse** of information and existing tools
- **Extendible** data model to incorporate application specific data
- **Integrating** with existing technologies to **store, query** and **visualise** data

ACS Synthetic Biology has officially adopted SBOL as publication standard

- **SBOL Visual** is the recommended graphical notation for depicting genetic constructs
- **SBOL 2.0 Data Model** is the preferred format for nucleic acid sequences

Viewpoint
pubs.acs.org/synthbio

ACS Synthetic Biology

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.



- Manuscript submission, review, and production process is linked to **SBOL-enabled repositories**
- Joint BioEnergy Institute (JBEI) has set up the initial repository

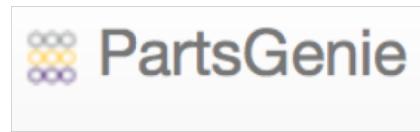
Applications



iBioSim



CELLO



Gene Designer

Shortbol



poolDesigner

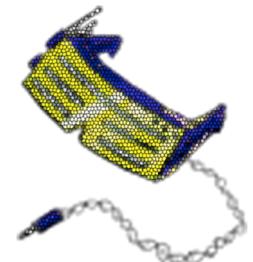
Pinecone



SBOLDesigner



GoldenBraid 3.0



DeviceEditor

Applications - continued

APPLICATION	KNOWLEDGE MANAGEMENT	SEQUENCE DESIGN	CIRCUIT DESIGN	MODELING	VISUALIZATION	SBOL VISUAL	IM-PORT SBOL1	IM-PORT SBOL2	EX-PORT SBOL1	EXPORT SBOL2
Pigeon Visualizer					★	★				
DNAPlotlib					★	★	★	★	★	★
SBOLDesigner	★	★			★	★	★	★	★	★
ICE	★	★			★	★		★		★
j5		★		★		★	★	★		★
DeviceEditor	★	★	★		★	★	★		★	
Finch	★	★	★		★	★				★
Pool Designer		★					★	★	★	
Virtual Parts Repository	★		★	★		★	★	★	★	★
MoSeC		★	★						★	

Data Repositories: Inventory of Composable Elements (ICE)



Search SEARCH

Create Entry grid icon

Administrator exit icon

Category	Type	ID	Description	Status	Actions	Date
Featured	PLASMID	LCP_000117	pGOP35_pSB1C3-g13Op-minCMV-GFP-RBGpA The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	✓	Dec 2, 2016
Personal	PLASMID	LCP_000116	pGOP30_pSB1C3-g8Op-minCMV-GFP-RBGpA The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	✓	Dec 2, 2016
Shared	PLASMID	LCP_000115	pGOP25_pSB1C3-g3Op-minCMV-GFP-RBGpA The 2016 BostonU iGEM team designed a set of mutually orthogonal...	Complete	✓	Dec 2, 2016
Drafts	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
Pending Approval	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

NSF BOSTON UNIVERSITY JBEI ICE Registry 5.3.0 All rights reserved. Submit an Issue | Help

Data Repositories: SynBioHub

The screenshot shows the homepage of the SynBioHub website. At the top, there is a navigation bar with a logo, the text "SynBioHub", a "Submit" button, and a "Login or Register" button. Below the navigation bar is a large orange header with the text "Home". The main content area features a large "SynBioHub" logo with a circuit board icon. A descriptive text block explains that SynBioHub is a design repository for biological constructs, enabling users to upload DNA and protein designs and search for existing parts. Below this are three main sections: "Search for useful parts and designs" with a search bar and "Browse Design Collections" button; "Upload your design for safekeeping" with an "Upload" icon and "Submit a Design" button; and "Share designs for publication or collaboration" with a "Share" icon and "Manage Submissions" button.

- Supports storing and retrieving native SBOL and other related data
- Unified syntax and semantics via SBOL
- Data can be federated across multiple instances
 - A web of registries
- Promotes data integration

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

Data Repositories: SBOLme

The screenshot shows the SBOLme repository interface. At the top left is the King Abdullah University of Science and Technology logo and the Computational Bioscience Research Center name. At the top right is the SBOLme logo and tagline. A navigation bar below has tabs for Query (selected), Catalog, Information, and About. The main content area contains a search form with dropdowns for 'Search for' (set to 'Compound') and 'Filter By' (set to 'Compound name or ID'). A value input field contains 'Pyruvate, C00022, ME_C00022'. To the right, a tooltip explains that the query can search by compound name or ID, mentioning KEGG and SBOLme IDs. Below the form are buttons for 'Search' and a dropdown for page size ('25'). A success message from 'Kuwahara et al., ACS Synthetic Biology (2017)' is displayed. At the bottom, a table shows the first result: SBOLme ID ME_R00006, KEGG ID R00006, and Name 2-acetolactate pyruvate-lyase (carboxylating).

جامعة الملك عبد الله
العلوم والتكنولوجيا
King Abdullah University of
Science and Technology

COMPUTATIONAL
BIOSCIENCE RESEARCH
CENTER

SBOLme
repository of **SBOL** parts for metabolic engineering

Query Catalog Information About

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.

Search 25

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)

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

Sequence Editor: Benchling

The screenshot shows the Benchling Sequence Editor interface with several tabs open:

- Projects**: Shows a list of projects including "SynBioHub Transfers", "A1_AmtR" (selected), "aadA.ref", "pJ163mRm_EF", and "pJ033mRm_EF".
- INVENTORY**: Shows the status of projects: "A1_AmtR" (Last modified 2 days ago), "AmtR" (Last modified 5 hours ago), "Test" (Last modified 2 days ago), and "The Best" (Last modified 3 days ago).

The main area displays the sequence map for "The Best" construct, which is 899 bp long. The map includes:

- OR2-OR1 Promoter**: Located between bases 10 and 60.
- BydvJ**: A gene fragment located between bases 70 and 130.
- A1**: A gene fragment located between bases 140 and 170.
- Amtr**: A gene fragment located between bases 170 and 200.
- Amtr**: A gene fragment located between bases 210 and 260.
- Amtr**: A gene fragment located between bases 270 and 330.
- Amtr**: A gene fragment located between bases 270 and 330.

The right side of the interface shows a detailed view of the "Amtr" gene fragment from base 100 to 800, listing restriction enzymes and their cleavage sites:

Enzyme	Cleavage Site (Base)	Enzyme	Cleavage Site (Base)
MlyI	100	BspAPI	800
PleI	105	MfeI	810
HinfI	110	FokI	815
DrdI	115	HpaII	820
AcuI	120	MspI	825
BsmAI	125	BspEI	830
SmlI	130	BsaWI	835
TspRI	135	Sau96I	840
TscAI	140	FokI	845
BtsIMutI	145	BtsCI	850
HoyAV	150	XmnI	855
BpuE1	155	BsaI	860
HchI	160	BspMI	865
DraI	165	BpuI	870
BaeGI	170	BcgI	875
HincII	175	BstYI	880
MseI	180	BglII	885
ApalI	185	PvuII	890
BsrI	190	EcoRI	895
AluI	195	XmnI	900
BcoDI	200	BtsCI	905
BfaI	205	AluI	910
BsgI	210		

At the bottom, there are buttons for "ASSEMBLY WIZARD" and "SPLIT WORKSPACE".

Sequence Editor: SBOLDesigner

The screenshot shows the SBOLDesigner interface. At the top, there's a toolbar with various icons for file operations like new, open, save, and search. Below the toolbar, the title bar displays 'GeneticInverter' and 'Overview'. The main workspace is titled 'Design' and contains a schematic diagram of a genetic circuit. The circuit consists of four components connected in series: a promoter (pAmtR), an RBS (A1), a gene construct (Betl), and a terminator (L3S2P55). Below the circuit, each component is labeled with its name. At the bottom of the interface, there's a 'Parts' section containing a grid of icons representing different biological parts, such as Genes, Promoters, Ribosome Binding Sites, CDSs, Termination sites, Origins of replication, PBSs, CUT sites, Scars, Operons, Insertions, RSEs, PSEs, RSs, PSs, BRSs, 5OH ends, 3OH ends, and Circular constructs.

Zhang et al., ACS Synthetic Biology (2017)

Other sequence editors that support SBOL: DeviceEditor, J5, VectorEditor (JBEI), DNAPlotLib (MIT/UW/Bristol), Eugene (Boston), GenoCAD (VBI), BOOST (JGI), etc.

Circuit Design & Modelling: Cello

The screenshot shows the Cello web application interface. On the left, a Verilog code editor displays the following code:

```
1 module A(output out1,  input in1, in2);
2 always@(in1,in2)
3 begin
4   case({in1,in2})
5     2'b00: {out1} = 1'b0;
6     2'b01: {out1} = 1'b0;
7     2'b10: {out1} = 1'b0;
8     2'b11: {out1} = 1'b1;
9   endcase
10 end
11 endmodule
12
```

On the right, there are two tables: "Inputs" and "Outputs".

Inputs

index	name	low RPU	high RPU	DNA sequence
1	pTac	0.0034	2.8	AACGATCGTGGCTGTGTTGACAA
2	pTet	0.0013	4.4	TACTCCACCGTTGGCTTTTTCCCC

Outputs

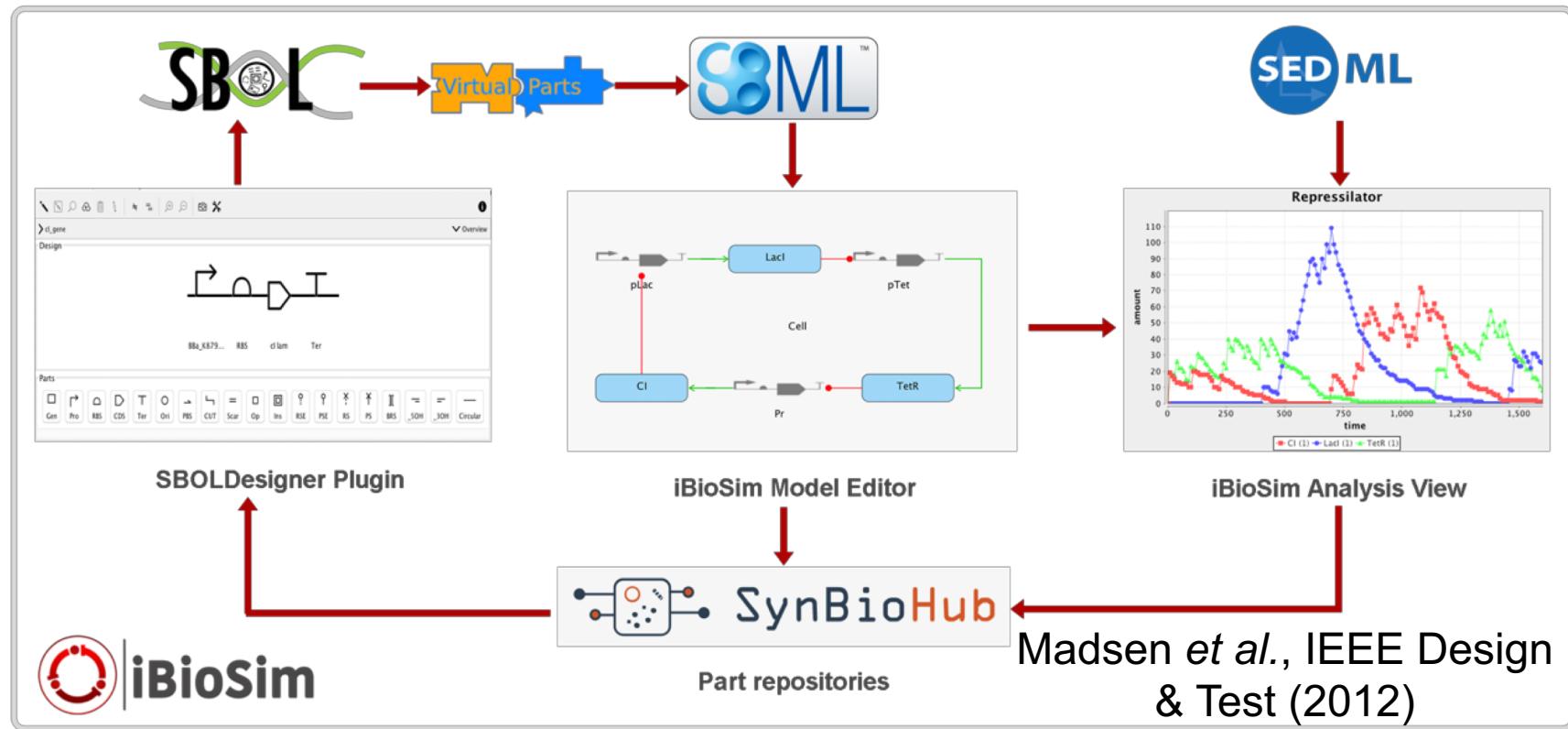
index	name	DNA sequence
1	YFP	CTGAAGCTGTCACCGGATGTGCTTCCGGTCTGATGAGTCCGT

Below the code editor is a "design name" input field and a green "Run" button.

Nielsen *et al.*,
Science (2016)

- Compiles combinational designs described in Verilog into genetic circuits
- Uses a gate library of 2-input NOR gates and inverters

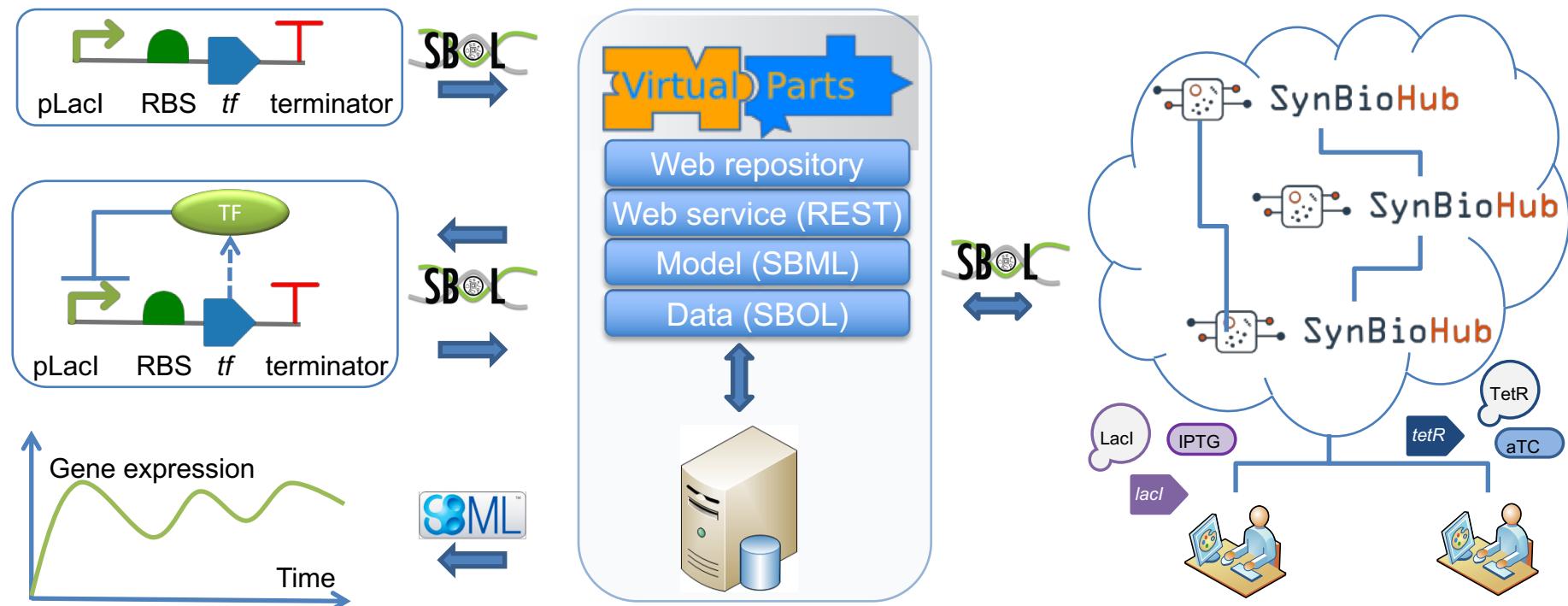
Circuit Design & Modelling: iBioSim



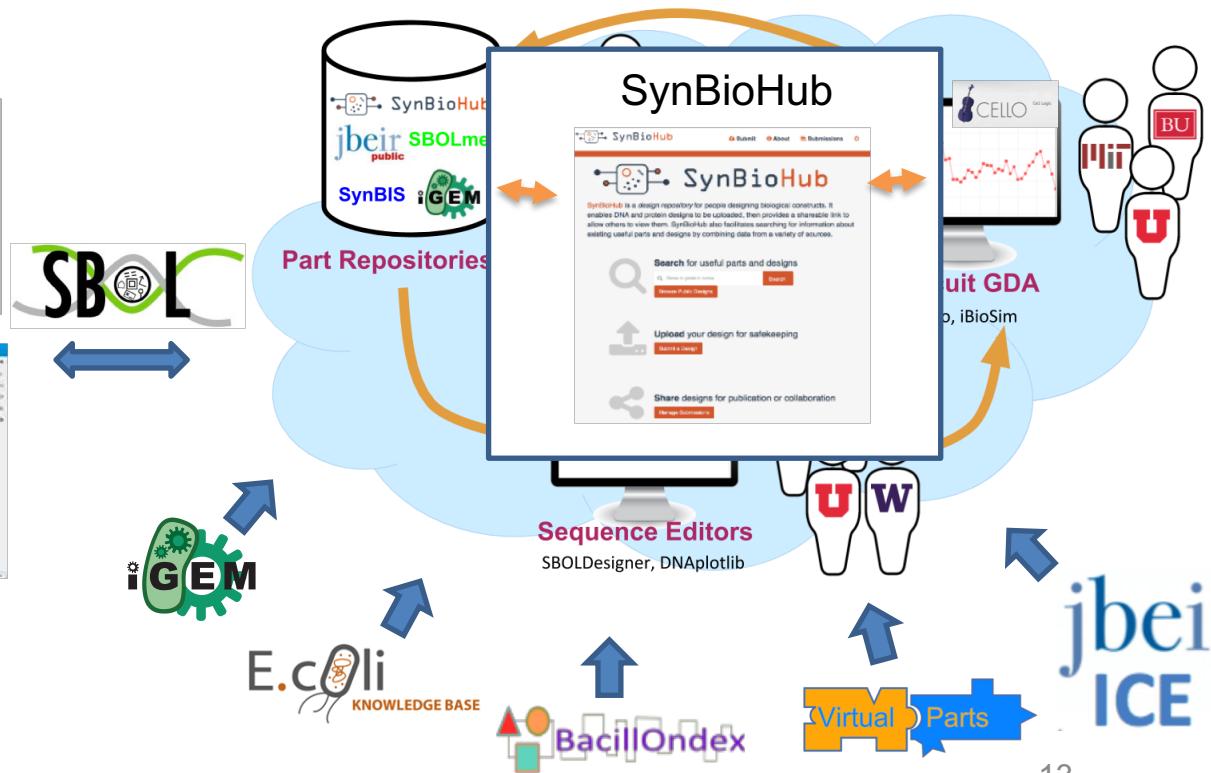
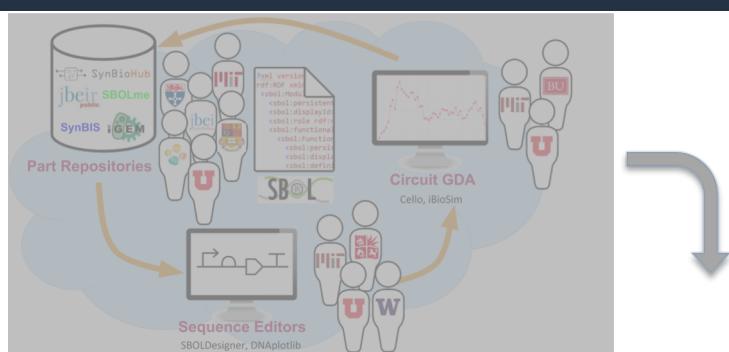
- Supports modeling, analysis, and design of genetic circuits
- Modelling uses SBML, analysis uses SED-ML, and design uses SBOL
- Includes SBOLDDesigner as plugin providing a connection to SynBioHub
- Provides converters between SBML and SBOL

Circuit Design & Modelling: Virtual Parts Repository

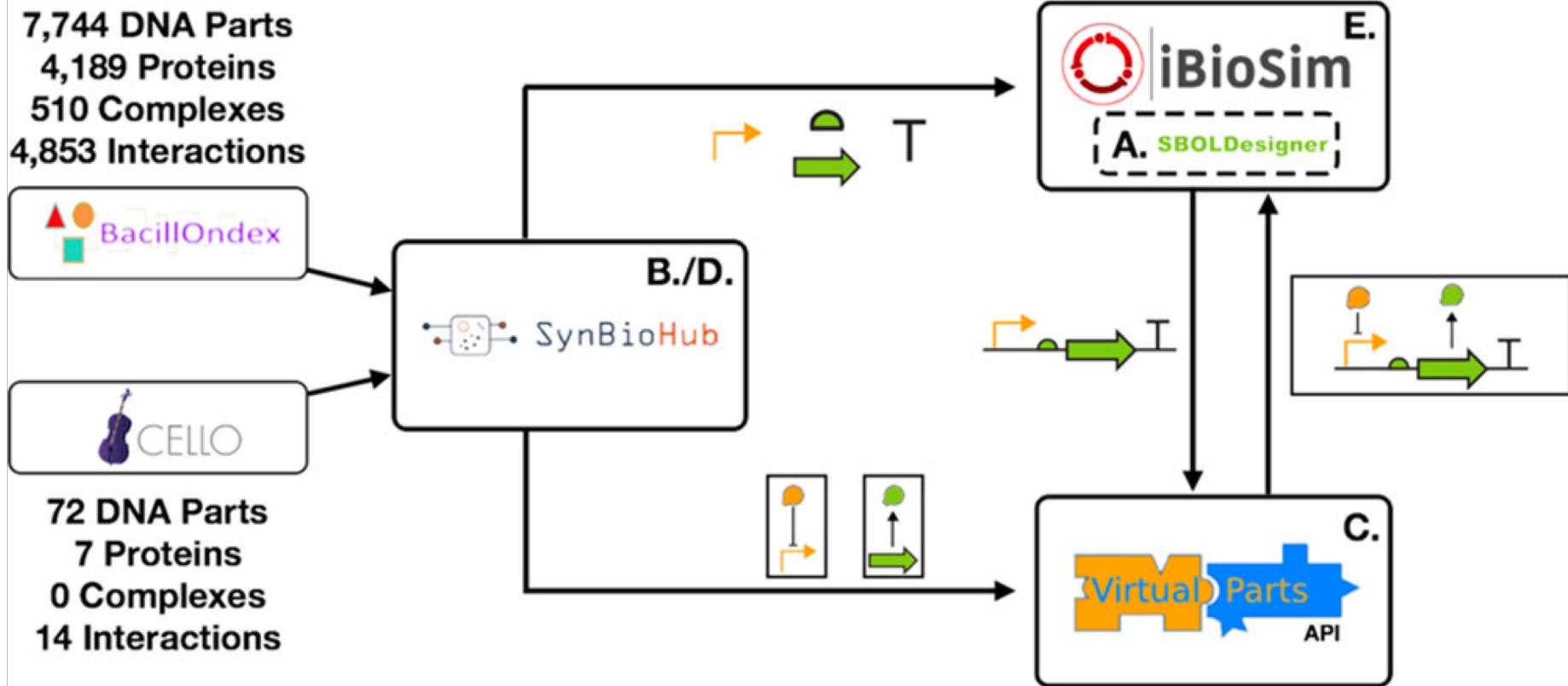
- A repository of modular and composable models of biological parts and interactions
- Provides a modelling abstraction for genetic circuit designs specified in SBOL files



Workflows: SynBioHub as a central hub for synbio data management and workflows



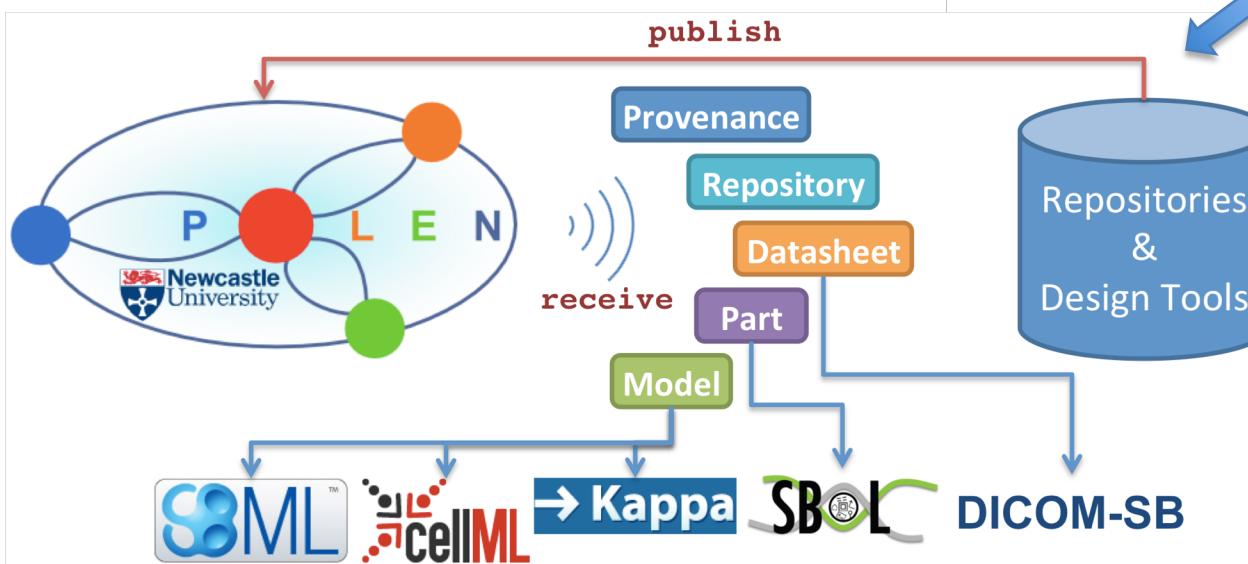
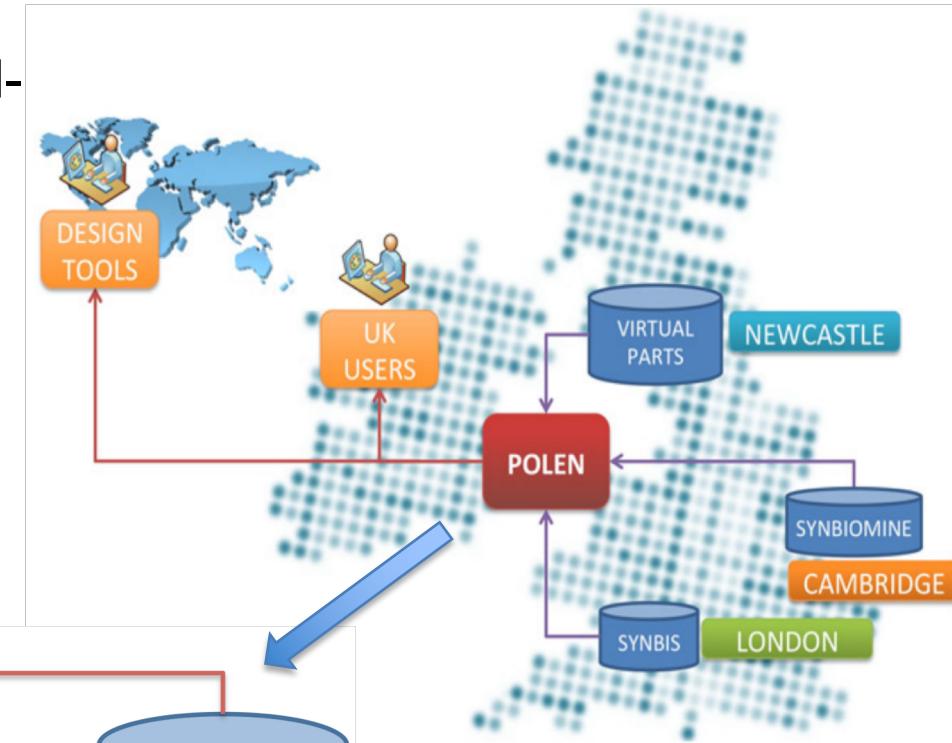
Workflows: Automated generation of models of genetic designs



Misirli & Nguyen *et al.*, 2018
ACS Synthetic Biology

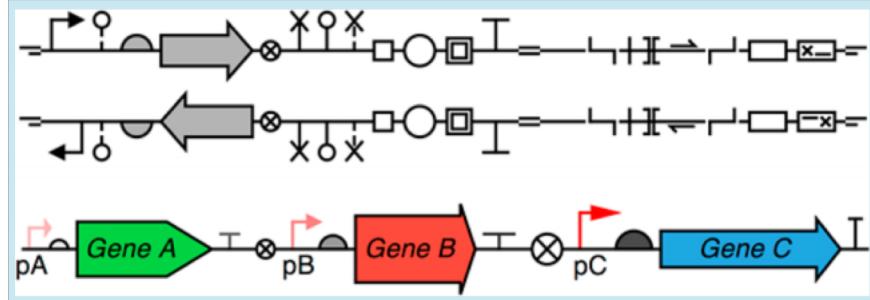
Workflows: POLEN - PrOtocol for Linking External Nodes

- Light-weight, asynchronous, Cloud-based messaging system
- Messages can be published/received for a particular topic
- Messages include URIs that point to actual resources in standard formats



Misirli et al., 2017
IET Engineering Biology

Visualisation examples



DNAplotlib

VisBOL is a web-based tool for visualizing genetic designs. It includes a design interface with tabs for SBOL, PIGEON, and GENBANK, and a display list. It also features a preview area showing a genetic circuit with components like pLaci, BBA_J61101 RBS, tetR, BBA_J61101 RBS, and gfp. The interface includes font and scale controls.

PartsGenie is a tool for designing genetic parts. It shows a library of parts (e.g., hexokinase-encoding part) and a detailed view of a specific part's sequence, which includes a promoter, a coding region, and a terminator.

PartsGenie

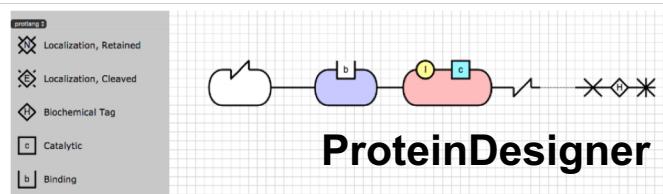
pigeoncad.org is a design visualizer for synthetic biology. It features a drawing of a pigeon and the text "Pigeon Design Visualizer for Synthetic Biology". Below the drawing are color swatches for codes 1 through 14. A detailed genetic circuit diagram is shown at the bottom.

box: A list of type [name] [color] [n] where type is one of p, P, c, r, t, T, s, o, >, <, | or -; a blank line; followed by name ind | rep name There are 14 colors labeled as the third item in each line. Even numbers are darker versions of the odd numbered colors in the following order of hues: blue, green, red, orange, purple! as a fourth item removes the label for that element. Characters < and > before p and c give inverse p and c.

or codes:



or codes:



ProteinDesigner

Open source libraries

Software libraries which import and export SBOL files are freely available at the [Synthetic Biology Data Exchange](#) on GitHub under the Apache 2.0 license.

Library	Programming Language
libSBOLj	Java
sboljs	JavaScript
libSBOL	C++
pySBOL	Python
semanticSBOL	OWL
SBOL-Validator	Online validation

- Developer support
- Online documentation
- Getting started tutorials
- Sample projects
- Code examples

<http://sbolstandard.org/software/libsbol>
<https://github.com/SynBioDex/>

Acknowledgments

SBOL Developers

sbol-dev@googlegroups.com

SBOL Editors

editors@sbolstandard.org

SBOL Libraries

<http://sbolstandard.org/software/libsbol>

Synthetic Biology Data

Exchange Group

<https://github.com/SynBioDex/>

<http://sbolstandard.org>



EPSRC

Engineering and Physical Sciences
Research Council



CLARK PARSIA



Raytheon
BBN Technologies



Berkeley
UNIVERSITY OF CALIFORNIA



THE UNIVERSITY OF UTAH

W
UNIVERSITY of
WASHINGTON



I(C₂O)₂S