Welcome to the Synthetic Biology Open Language (SBOL) Developers Workshop

> Newcastle University Newcastle Upon Tyne August 2016







### Programme

- 09:00-09:30 Overview of SBOL Anil Wipat
- 09:30-10:00 SBOL Visual Jake Beal
- 10:00-11:00 SBOL Data Model Chris Myers
- 11:00-11:30 Coffee
- 11:30-12:00 SBOL Tools and Resources
  - SBOLStack / SBOL Shorthand / VisBOL James McLaughlin
  - SBOLDesigner / iBioSim Michael Zhang
  - SBOL Validator / ICE Zach Zundel
- 12:00-01:00 Libraries
  - Java library Chris Myers
  - JavaScript library James McLauglin
  - C++ & Python libraries Bryan Bartley
- 13:00-14:00 Lunch
- 14:00-15:30 SBOL Tutorial (CRISPR/CAS9 example) Meher Samineni
- 15:30-16:00 Coffee
- 16:00-17:30 User problems Anil Wipat (moderator)
- 19:00 Dinner







# SBOL – promoting data sharing and exchange in Synthetic Biology

Anil Wipat (on behalf of the SBOL community) ICOS, Computing Science, Centre for Synthetic Biology and BioEconomy, Newcastle University UK (ico2s.org)







### Standards





Туре	Port Image	Connector Image	
Type A	4.5mm x 12.0mm		
Type B	7.3mm x 8.5mm		
Mini-A	3.0mm x 6.8mm		
Mini-B	3.0mm x 6.8mm		







### Standards



The nice thing about standards is that there are so many of them to choose from.

(Andrew S. Tanenbaum)

izquotes.com







### Standards









# What is SBOL – why is it needed?

- Synthetic Biology is engineering biology
- Requires movement of data & information
  - \* Between people, software & machines
  - Across geographic and political boundaries
- \* Standards needed to ensure interoperability
- SBOL a standard way to represent information about synthetic biology designs















### **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<sup>1</sup> disclosed the sequences of the toggle biotechnology switches published four years earlier in a paper by Gardner et al.<sup>2</sup>. The same year, Basu et al.3 deposited their construct sequences for programmed pattern formation into GenBank3. 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<sup>8</sup>. 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<sup>9</sup>. 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<sup>1</sup>, J Christopher Anderson<sup>2</sup>, Deepak Chandran<sup>3</sup>, Douglas Densmore<sup>4</sup>, Michal Galdzicki<sup>5</sup>, Matthew W Lux<sup>1</sup>, Cesar A Rodriguez<sup>6</sup>, Guy-Bart Stan<sup>7</sup> & Herbert M Sauro<sup>3</sup>

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

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

J. Peccoud, C. J. Anderson, D. Chandran, D. Densmore, M. Galdzicki, M. W. Lux, C. A. Rodriguez, G. Stan, and H. M. Sauro, "Essential information for synthetic DNA sequences," *Nature Biotechnology*, vol. 29, iss. 1, pp. 22-22, 2011.







02011 Nature America, Inc. All rights reserved







### SBOL v 1

### nature biotechnology

Home | Current issue | News & comment | Research | Archive 💌 Authors & referees 💌 About the journal

home > archive > issue > computational biology > perspective > full text

NATURE BIOTECHNOLOGY | COMPUTATIONAL BIOLOGY | PERSPECTIVE < 《 昌 日本語要約

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

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 v 1



Designed four expression cassettes, leaving RBS and terminator components unspecified.

Generated four variant designs of each cassette using six RBS and six terminator components.

Assembled and modeled sixteen variant designs of the toggle switch.

Joint BioEnergy Institute Joint BioEnergy Institute

Stored design files in repository for dissemination.













### SBOL visual

BBF RFC 93

SBOL Visual

### Synthetic Biology Open Language Visual (SBOL Visual), version 1.0.0

Jacqueline Quinn, Jacob Beal, Swapnil Bhatia, Patrick Cai, Joanna Chen, Kevin Clancy, Nathan Hillson, Michal Galdzicki, Akshay Maheshwari, Umesh P, Matthew Pocock, Cesar Rodriguez, Guy-Bart Stan, Drew Endy

March 21, 2013







# SBOL visual

- Open source
   graphical notation
   for SynBio designs
- \* SBOL 1 currently,SBOL 2 in progress
- \* Spec: DOI 1721.1/78349

→ promoter		<ul> <li>primer binding site</li> </ul>
◯ cds	1	restriction site
ribosome entr	ry site ][	blunt restriction site
terminator	L	5' sticky restriction site
operator		3' sticky restriction site
insulator		= 5' overhang
ribonuclease s	site —	- 3' overhang
Orna stability el	ement 🗕	assembly scar
★ protease site	×	signature
<b>Q</b> protein stabili	ty element	user defined
O origin of replic	cation	







### Visual Paper

PLOS BIOLOGY

COMMUNITY PAGE

### SBOL Visual: A Graphical Language for Genetic Designs

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

1 Autodesk Research, Autodesk Inc., San Francisco, California, United States of America, 2 Chemical Science and Engineering, Kobe University, Kobe, Japan, 3 Information and Knowledge Technologies, Raytheon BBN Technologies, Cambridge, Massachusetts, United States of America, 4 Electrical and Computer Engineering, Boston University, Boston, Massachusetts, United States of America, 5 School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom, 6 Fuels Synthesis and Technologies Divisions, Joint BioEnergy Institute, Emeryville, California, United States of America, 7 Lawrence Berkeley National Lab, Berkeley, California, United States of America, 8 Synthetic Biology Unit, ThermoFisher Scientific, Carlsbad, California, United States of America, 9 Arzeda Corp, Seattle, Washington, United States of America, 10 Babraham Institute, Cambridge, United Kingdom, 11 Stanford University School of Medicine, Stanford, California, United States of America, 12 School of Computing Science, Newcastle University, Newcastle upon Tyne, United Kingdom, 13 Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, Utah, United States of America, 14 Department of Computational Biology & Bioinformatics, University of Kerala, Kerala, India, 15 Turing Ate My Hamster LTD, Newcastle upon Tyne, United Kingdom, 16 Department of Biomedical Sciences, College of Medicine, Florida State University, Tallahassee, Florida, United States of America, 17 Computer Science, Brunel University, London, United Kingdom, 18 Department of Bioengineering, Centre for Synthetic Biology and Innovation, Imperial College London, South Kensington Campus, London, United Kingdom, 19 Centre for Synthetic Biology of Fine and Specialty Chemicals (SYNBIOCHEM), University of Manchester, Manchester, United Kingdom, 20 Bioengineering, University of Washington, Seattle, Washington, United States of America

These authors contributed equally to this work.

Creative Commons Attribution License, which permits \* hsauro@uw.edu









GOPEN ACCESS

Published: December 3, 2015

Citation: Quinn JY, Cox RS III, Adler A, Beal J,

Graphical Language for Genetic Designs. PLoS Biol

13(12): e1002310. doi:10.1371/journal.pbio.1002310

Copyright: © 2015 Quinn et al. This is an open access article distributed under the terms of the

Bhatia S, Cai Y, et al. (2015) SBOL Visual: A

## Example tools for the generation of **SBOL** Visual



\* VisBOL

Newcastle

niversitv

- PigeonCAD \*
- **DNAplotLib** \*

http://visbol.org/ http://pigeoncad.org/ https://github.com/VoigtLab/dnaplotlib







## SBOL 2 specification

### BBF RFC 108: Synthetic Biology Open Language (SBOL) Version 2.0.0

Bryan Bartley Jacob Beal Kevin Clancy Goksel Misirli Nicholas Roehner **Editors:** 

University of Washington, USA Raytheon BBN Technologies, USA ThermoFisher Scientific, USA Newcastle University, UK Boston University, USA

Herbert Sauro

University of Washington, USA

editors@sbolstandard.org

Chair:

#### Additional authors, by institution:

Ernst Oberortner Curtis Madsen, Matthew Pocock, Anil Wipat Tramy Nguyen, Zhen Zhang, Chris Myers John H. Gennari Michael Bissell DOE Joint Genome Institute, USA Newcastle University, UK University of Utah, USA University of Washington, USA Amyris, Inc., USA

Version 2.0.0







### SBOL 2 in practise

### Synthetic Biology-

#### Research Article

pubs.acs.org/synthbio

#### Sharing Structure and Function in Biological Design with SBOL 2.0

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

<sup>†</sup>Department of Electrical and Computer Engineering, Boston University, Boston, Massachusetts 02215, United States

<sup>‡</sup>Raytheon BBN Technologies, Cambridge, Massachusetts 02138, United States

<sup>§</sup>Thermo Fisher Scientific, Carlsbad, California 92008, United States

<sup>1</sup>Department of Bioengineering, University of Washington, Seattle, Washington 98195, United States

School of Computing Science, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.

<sup>¶</sup>Institute for Research in Immunology and Cancer, University of Montreal, Montreal, Quebec H3T 1J4, Canada

"U.S. Department of Energy Joint Genome Institute, Walnut Creek, California 94598, United States

<sup>V</sup>Turing Ate My Hamster, Ltd., Newcastle upon Tyne NE27 0RT, U.K.

<sup>®</sup>Amyris, Inc., Emeryville, California 94608, United States

Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, Utah 84112, United States

▲Department of Bioengineering, University of Utah, Salt Lake City, Utah 84112, United States

Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, Washington 98195, United States

Supporting Information

ABSTRACT: The Synthetic Biology Open Language (SBOL) is a standard that enables collaborative engineering of biological systems across different institutions and tools. SBOL is developed through careful consideration of recent synthetic biology trends, real use cases, and consensus among leading researchers in the field and members of commercial biotechnology enterprises. We demonstrate and discuss how a set of SBOL-enabled software tools can form an integrated, cross-organizational workflow to recapitulate the design of one of the largest published genetic circuits to date, a 4-input AND sensor. This design encompasses the structural components of the system, such as its DNA, RNA, small molecules, and proteins, as well as the interactions between these components that determine the system's behavior/function. The demonstrated workflow and resultion circuit design illustrate







### SBOL 2.0









## SBOL Community



**SBOL Developers Group** includes 116 members from more than 30 organizations.









- \* **SBOL Chair** Anil Wipat (Newcastle) formerly Herbert Sauro (U. of Washington)
- \* **SBOL Steering Committee** Jake Beal, Kevin Clancy, Doug Densmore, John Genari, Nathan Hilson, Chris Myers, Herbert Sauro
- \* **SBOL Editors** Tramy Nguyen (Utah), Jacob Beal (BBN Technologies), Robert Sidney Cox (Kobe), Raik Grunberg (Montreal) and James McLaughlin (Newcastle).
- Past SBOL Editors Michal Galdzicki (U. of Washington), Ernst Oberortner (JGI/BU), Matthew Pocock (Newcastle), Jacqueline Quinn (Google), Cesar Rodriguez (Autodesk), Nick Roehner (BU) and Mandy Wilson (VBI), Bryan Bartley (U. of Washington), Goksel Misirli (Newcastle).













### Infrastructure & tools

### Libraries (libSBOL)

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. Libraries are implemented in:

- Java
- Javascript
- C/C++
- Python

All libraries are under active development. The latest release of the Java, and Javascript libraries now support SBOL 2.0. In addition a beta release of the C/C++ libraries supporting SBOL 2.0 are now available. Python bindings for SBOL 1.1 and SBOL 2.0 are currently available.







## SBOL and SBML



- \* Misirli et al., (2011)
- Roehner et al., ACS
   Synthetic Biology (2013)
- Roehner et al., ACS
   Synthetic Biology (2014)
- Nguyen/Myers, IWBDA (2015)



### SyntheticBiology-----

Generating Systems Biology Markup Language Models from the Synthetic Biology Open Language

Nicholas Roehner, $*'^{\dagger}$  Zhen Zhang,<sup>‡</sup> Tramy Nguyen,<sup>‡</sup> and Chris J. Myers<sup>‡</sup>

<sup>†</sup>Department of Bioengineering, University of Utah, Salt Lake City, Utah 84112, United States <sup>‡</sup>Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, Utah 84112, United States

Lacl Inverter

### A Converter from the Systems Biology Markup Language to the Synthetic Biology Open Language

Tramy Nguyen Dept. of Electrical and Computer Engineering tramy.nguyen@utah.edu Chris J. Myers Dept. of Electrical and Computer Engineering myers@ece.utah.edu







**Research Article** 

pubs.acs.org/synthbio

## Infrastructure - applications









Software	Туре	Visual	Data Model
ICE	Repository	Yes	Yes
SBOL Hub	Repository	No	Yes
GenBank Converter	Helper	No	Yes
MoSeC	Helper	No	Yes
DNAplotlib	Visualization	Yes	Yes
Pigeon	Visualization	Yes	No
VisBOL	Visualization	Yes	Yes
Benchling	DNA Design	No	Yes
DeviceEditor	DNA Design	Yes	Yes
Eugene	DNA Design	Yes	Yes
GeneGenie	DNA Design	No	Yes
GenoCAD	DNA Design	Yes	Yes
SynBad	DNA Design	Yes	Yes
TeselaGen	DNA Design	Yes	Yes
VectorEditor	DNA Design	No	Yes
Cello	Circuit Design	Yes	Yes
iBioSim	Circuit Design	No	Yes
RetroPath	Circuit Design	No	Yes
SBROME	Circuit Design	No	Yes
TinkerCell	Circuit Design	Yes	Yes
BioCompiler	Circuit Design	Yes	Yes
Clotho 3.0	Assembly	Yes	Yes
Raven	Assembly	Yes	No



S

-



Pigeon Design Visualizer for Synthetic Biology



**DeviceEditor** 

SBOL Desi**de** 











# Publishing









### ACS Synthetic Biology recommends SBOL



Viewpoint pubs.acs.org/synthbio

SB®

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

Nathan J. Hillson,\*<sup>,†,‡,§,||</sup> Hector A. Plahar,<sup>†,‡,||</sup> Jacob Beal,<sup>\*,⊥</sup> and Ranjini Prithviraj<sup>#</sup>

<sup>†</sup>Fuels Synthesis and Technology Divisions, DOE Joint BioEnergy Institute (JBEI), Emeryville, California 94608, United States
 <sup>‡</sup>Biological Systems and Engineering Division, Lawrence Berkeley National Lab, Berkeley, California 94720, United States
 <sup>§</sup>DOE Joint Genome Institute, Walnut Creek, California 94598, United States
 <sup>II</sup>Synthetic Biology Engineering Research Center, Emeryville, California 94608, United States
 <sup>⊥</sup>Raytheon BBN Technologies, Cambridge, Massachusetts 02138, United States
 <sup>#</sup>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.







# SBOL is not just about design

SB®L



# SBOL is not just about design

- \* AND
- Work with other standards









### SBOL & COMBINE

### **c**ambine

- \* SBOL is part of something bigger
- \* **COMBINE** an initiative to coordinate systems biology standards
- \* Enables interaction with communities developing other standards
  - \* SBOL Visual/SBGN
  - \* SBOL/SBML/SED-ML
  - \* SBOL/BioPAX
- \* Provides
  - \* Specification infrastructure
  - \* Metadata and annotations
  - \* COMBINE archive
  - \* Repositories
  - \* Compliance testing





### SBOL & COMBINE

### **c**ambine

Home > COMBINE 2016

### **COMBINE 2016**

Dates: September 19-24, 2016 Location: Newcastle upon Tyne, UK Hosted by: <u>Newcastle University</u>

The "Computational Modeling in Biology" Network (COMBINE) is an initiative to coordinate the development of the various community standards and formats in systems biology and related fields.

COMBINE is a workshop-style event with oral presentation, posters, and breakout sessions. The five meeting days will include talks about the <u>COMBINE standards and associated or related standardization efforts</u>, as well as presentations of tools using these standards. Oral presentations will be selected from the submitted abstracts. In addition to oral presentations, poster sessions will allow people to inform each other about their software and other projects in a setting that fosters interaction and in-depth discussion.

The 2016 COMBINE meeting will be held in Newcastle upon Tyne, UK from September 19 to 24, 2016: the main Workshop will run 19-23 September and an SBGN Workshop will be held on September 24.







### Acknowledgements





Supported by National Science Foundation Grants CCF-1218095, DBI-1356041, and DBI-1355909 and EPSRC EP/J02175X/1

Thank you to the ICOS group, Newcastle Centre for Synthetic Biology and the BioEconomy (CSBB), and Centre for Health and Bioinformatics (CHaBi) for sponsoring dinner





