

# Synthetic Biology Open Language (SBOL)

Chris J. Myers

University of Utah

COMBINE & ERASysAPP Tutorial

September 14, 2014

# Synthetic Biology

- *Synthetic biology* extends *genetic engineering* by adding:
  - *Standards*
  - *Abstraction*
  - *Decoupling*

(source: Drew Endy)

# Synthetic Biology

- *Synthetic biology* extends *genetic engineering* by adding:
  - *Standards*
  - *Abstraction*
  - *Decoupling*

(source: Drew Endy)

# Importance of Standards

## 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 switches published four years earlier in a paper by Gardner *et al.*<sup>2</sup>. The same year, Basu *et al.*<sup>3</sup> deposited their construct sequences for programmed pattern formation into GenBank<sup>3</sup>. 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>

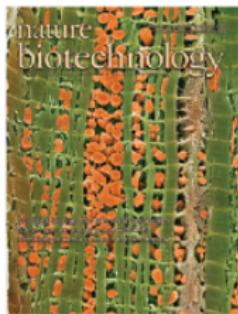
<sup>1</sup>Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, Virginia, USA. <sup>2</sup>Department of Bioengineering, QB3: California Institute for Quantitative Biological Research, University of California, Berkeley, California, USA.

<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. <sup>5</sup>Biomedical and Health Informatics, University of Washington, Seattle, Washington, USA. <sup>6</sup>BIOFAB, Emeryville, California, USA. <sup>7</sup>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. & Voigt, C.A. *Nature* **434**, 1120–1124 (2005).



# Synthetic Biology Open Language (SBOL)

- In 2008, a small group of researchers proposed the development of the *synthetic biology open language* (SBOL), an open-source standard for the exchange of genetic designs.
- In 2011, the first version of the SBOL core data model was released.
- In 2013, the first version of the SBOL Visual standard was released.
- Leveraging libSBOLj, a java-based library for SBOL's core data model, 18 software tools now support SBOL.

# SBOL Community

- SBOL Chair - Herbert Sauro (Washington)
- SBOL Editors - Bryan Bartley (Washington), Kevin Clancy (Life Tech.), Jacqueline Quinn (Google), Goksel Misirli (Newcastle), and Nicholas Roehner (Utah).
- Past SBOL Editors - Michal Galdzicki (U. of Washington), Ernst Oberortner (Boston), Matthew Pocock (Newcastle), Cesar Rodriguez (Autodesk), and Mandy Wilson (VBI).
- SBOL Developers Group includes 80 members from 29 organizations.
  - Academia - Boston, ETH Zurich, Imperial College London, Newcastle, Stanford, Berkeley, Kerala, Montreal, Utah, Washington, VBI, etc.
  - Industry - Agilent, Amyris, Autodesk, BBN, Clark & Parsia, DNA 2.0, Genome Compiler, JBEI, Life Technologies, etc.

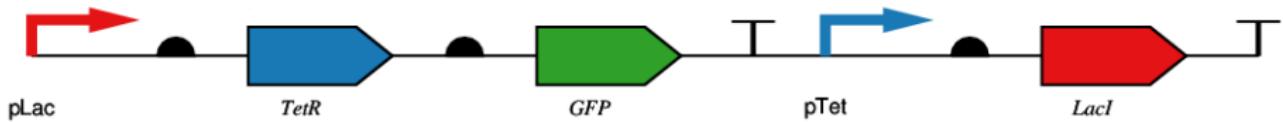
# SBOL and COMBINE

- *SBOL Developers Group* meet twice a year to discuss next steps.
- SBOL recently joined the COMBINE community of standards and participated in the COMBINE Workshop at USC in August.
- There are many potential synergies between SBOL and COMBINE standards, such as SBML, SBGN, and BioPax.

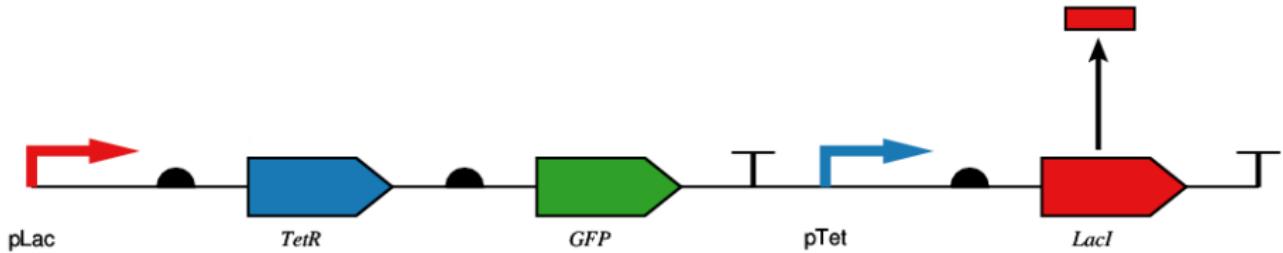
# SBOL Visual

 promoter	 origin of replication
 cds	 primer binding site
 ribosome entry site	 blunt restriction site
 terminator	 sticky restriction site
 operator	 5' overhang
 insulator	 3' overhang
 ribonuclease site	 assembly scar
 rna stability element	 signature
 protease site	 user defined
 protein stability element	

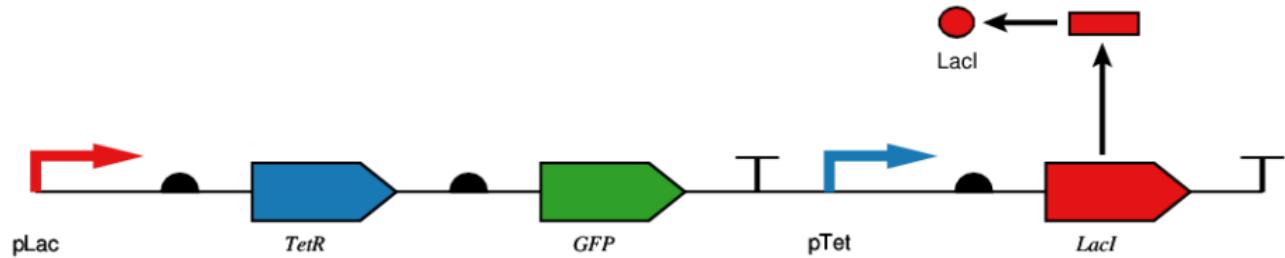
# Genetic Toggle Switch (Gardner et al. 2000)



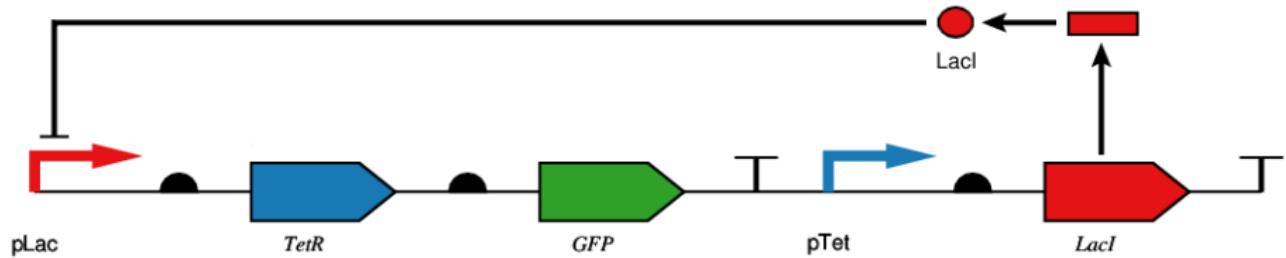
# Genetic Toggle Switch (Gardner et al. 2000)



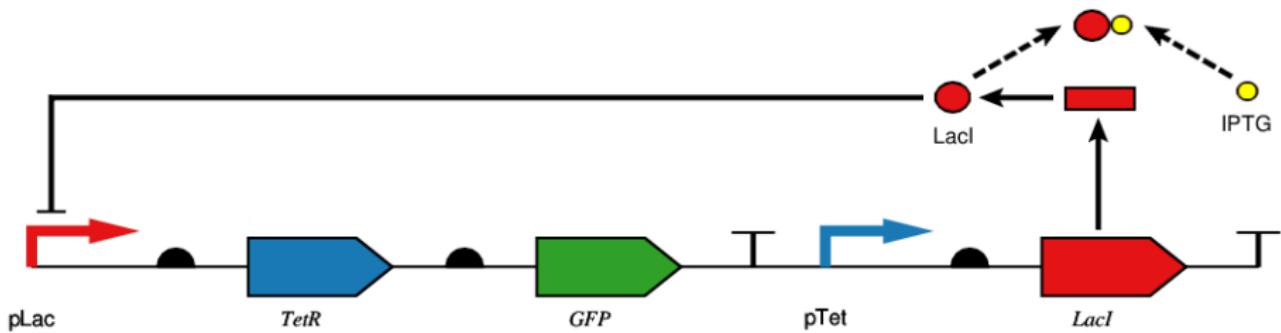
# Genetic Toggle Switch (Gardner et al. 2000)



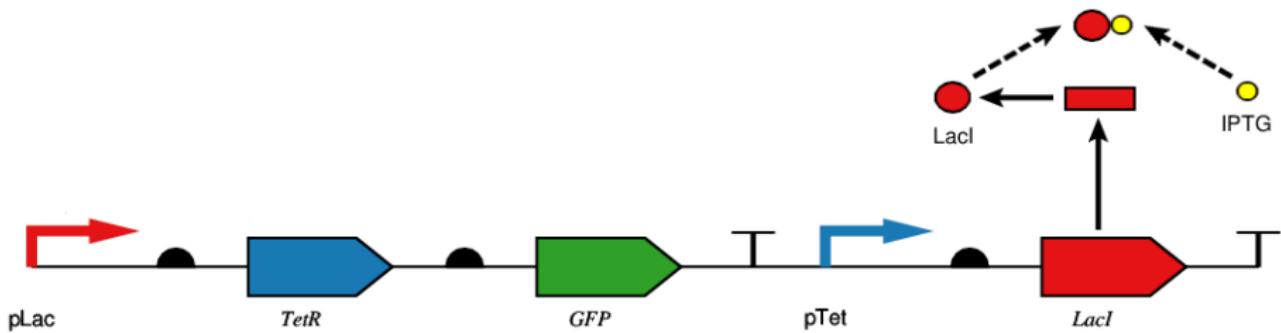
# Genetic Toggle Switch (Gardner et al. 2000)



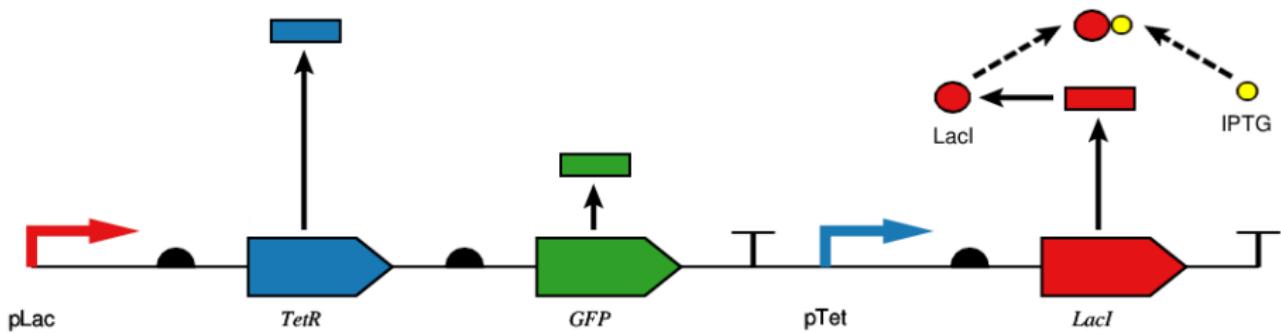
# Genetic Toggle Switch (Gardner et al. 2000)



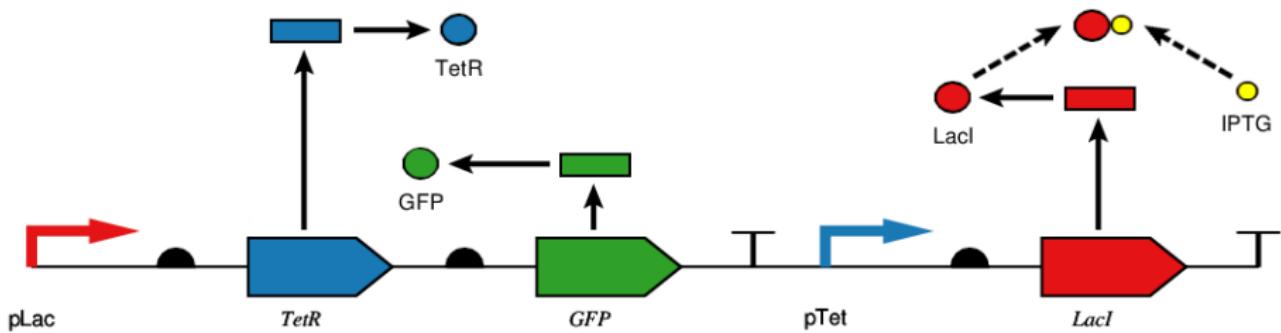
# Genetic Toggle Switch (Gardner et al. 2000)



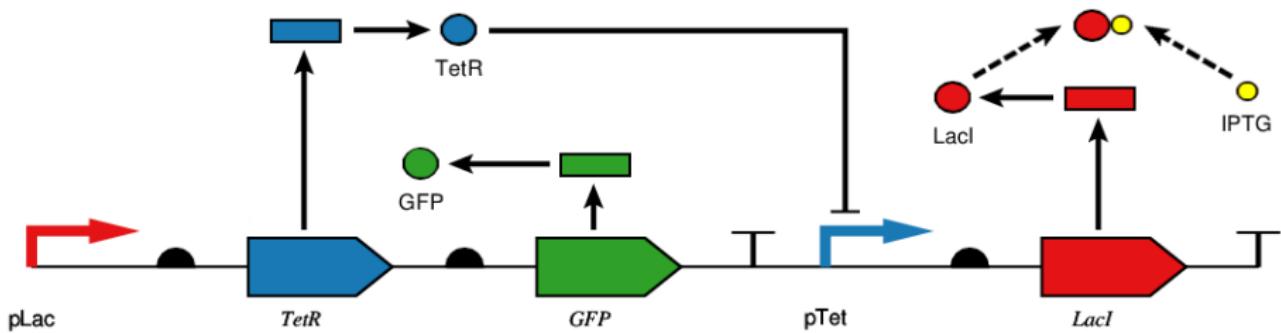
# Genetic Toggle Switch (Gardner et al. 2000)



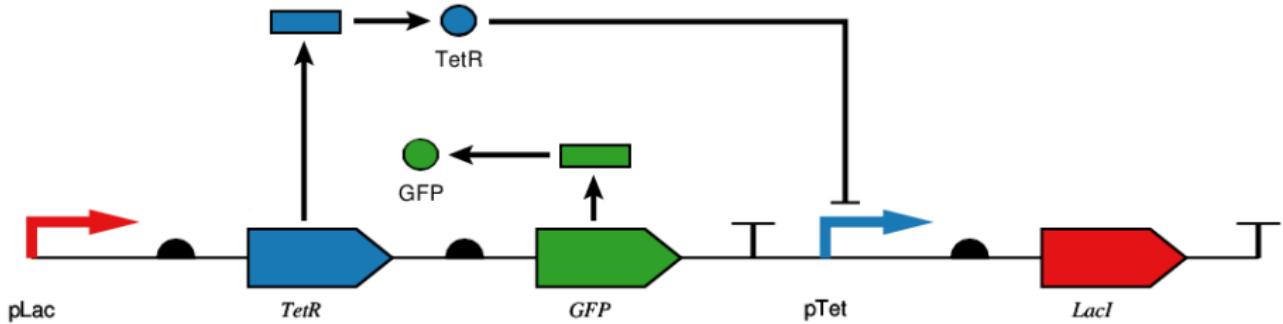
# Genetic Toggle Switch (Gardner et al. 2000)



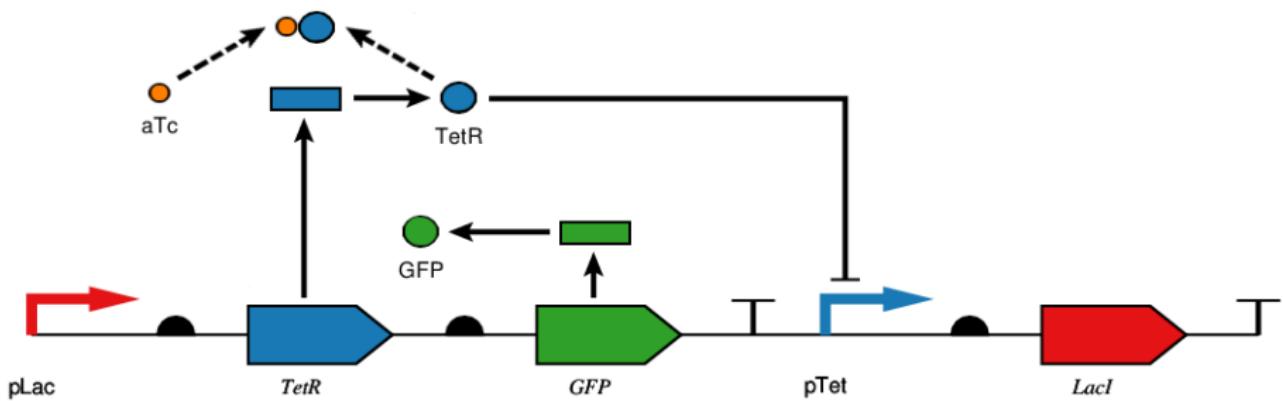
# Genetic Toggle Switch (Gardner et al. 2000)



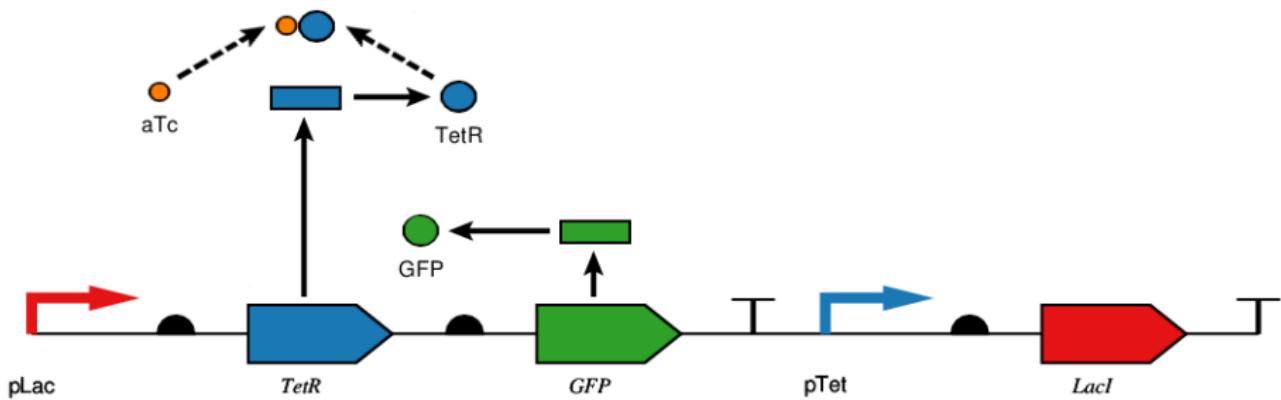
# Genetic Toggle Switch (Gardner et al. 2000)



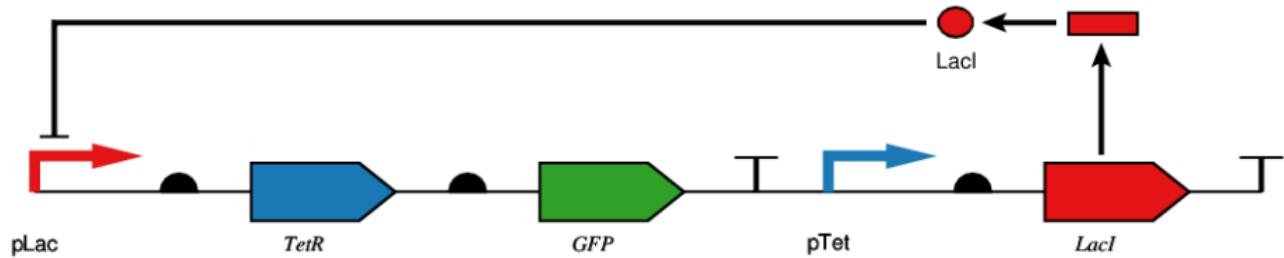
# Genetic Toggle Switch (Gardner et al. 2000)



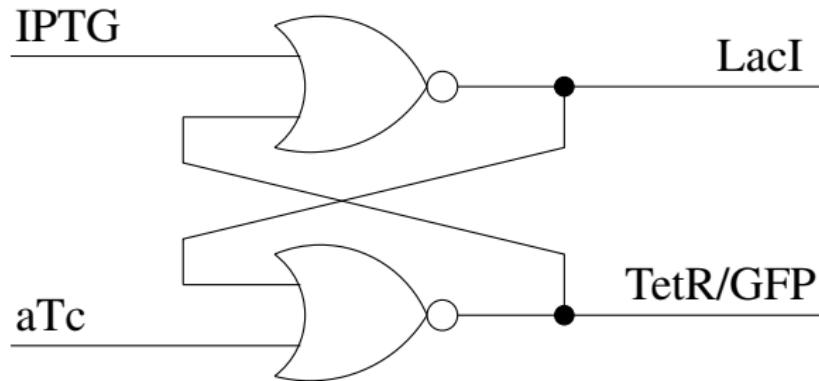
# Genetic Toggle Switch (Gardner et al. 2000)



# Genetic Toggle Switch (Gardner et al. 2000)

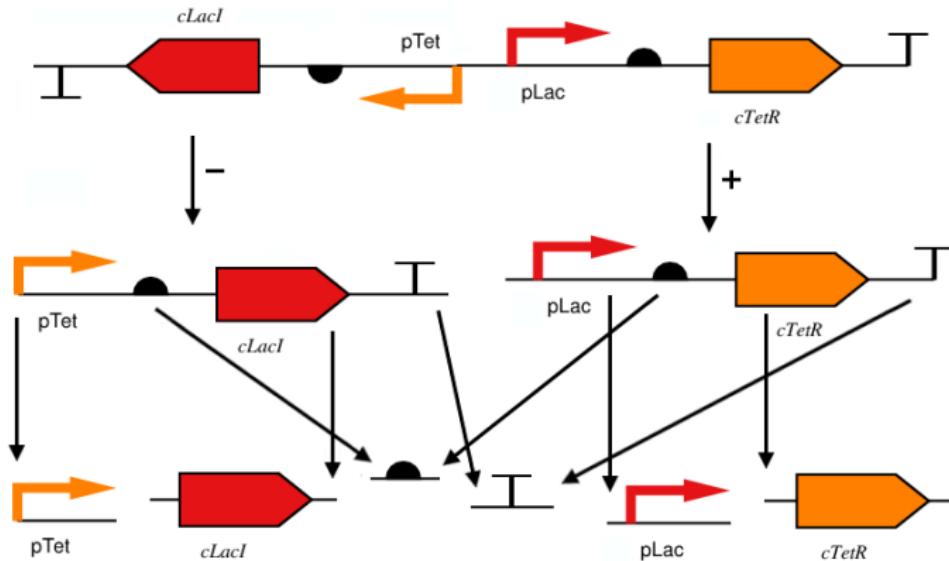


# Genetic Toggle Switch (SR Latch) Logic Diagram



# SBOL Version 1.1

- Version 1.1 specifies the hierarchical composition of *DNA components*.



Galdzicki et al., Nature Biotechnology (2014).

# SBOL Demonstration

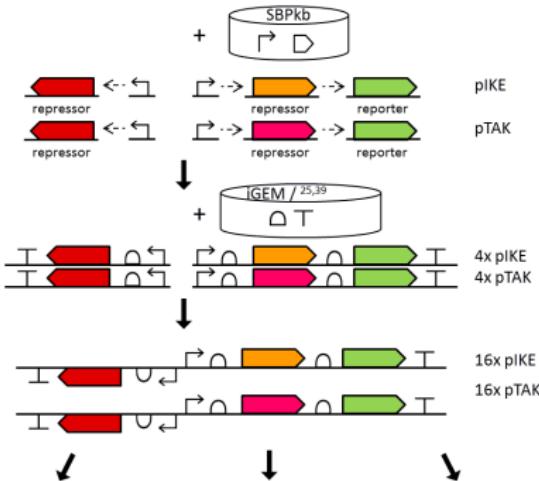


BOSTON  
UNIVERSITY

Boston



Utah



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.

*life*  
technologies™

Life Technologies

Received design file for codon optimization and gene synthesis.

Newcastle  
University

Newcastle

Stored design and model files in repository for dissemination.

jbei  
Joint BioEnergy Institute

Joint BioEnergy Institute

Stored design files in repository for dissemination.

Galdzicki et al., Nature Biotechnology (2014)

# University of Washington's SBOL Designer

(Loading SBOLDesigner.mov)

# Boston University's Eugene

(Loading Eugene.mov)

# University of Utah's iBioSim



(Loading iBioSim.mov)

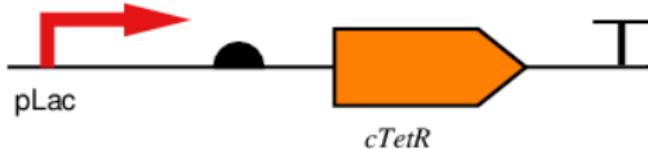
# Newcastle's Virtual Parts Repository

(Loading VirtualParts.mov)

# JBEI's ICE Repository

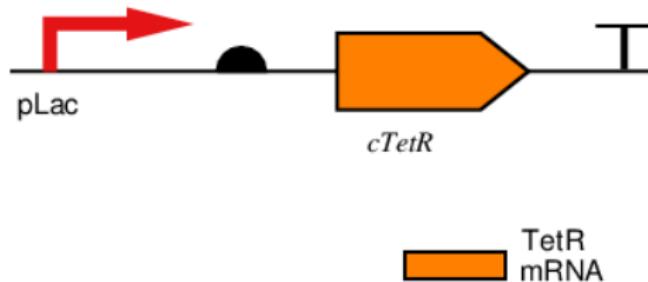
(Loading JBEIICE.mov)

# Increasing SBOL's Structural Range



Roehner et al., ACS Synthetic Biology (2014)

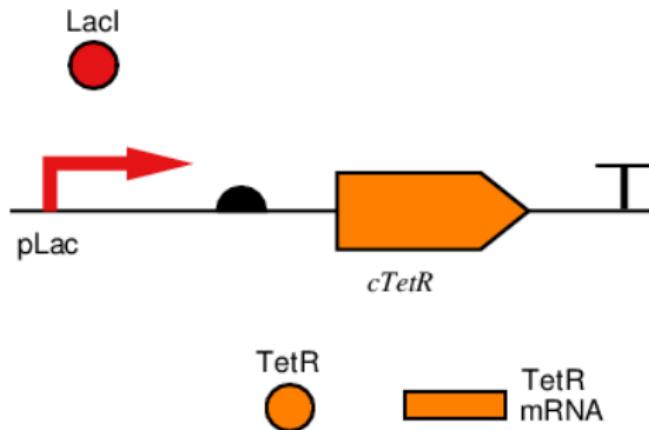
# Increasing SBOL's Structural Range



- RNA components (mRNA, tRNA, siRNA)

Roehner et al., ACS Synthetic Biology (2014)

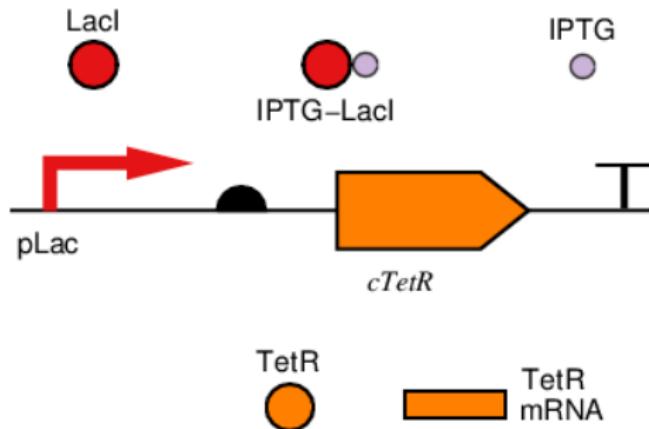
# Increasing SBOL's Structural Range



- RNA components
- Protein components (TFs, enzymes)

Roehner et al., ACS Synthetic Biology (2014)

# Increasing SBOL's Structural Range

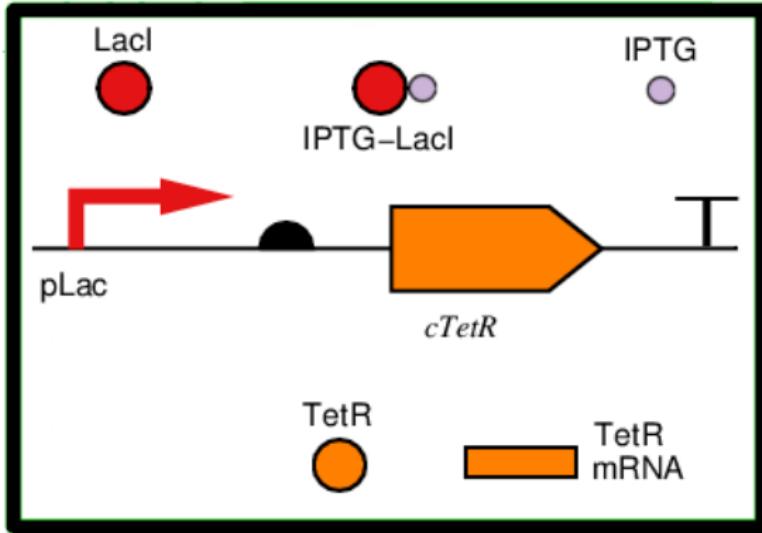


- RNA components
- Protein components
- Other Components (small molecules, complexes, light, temperature, etc.)

Roehner et al., ACS Synthetic Biology (2014)

# Increasing SBOL's Functional Range

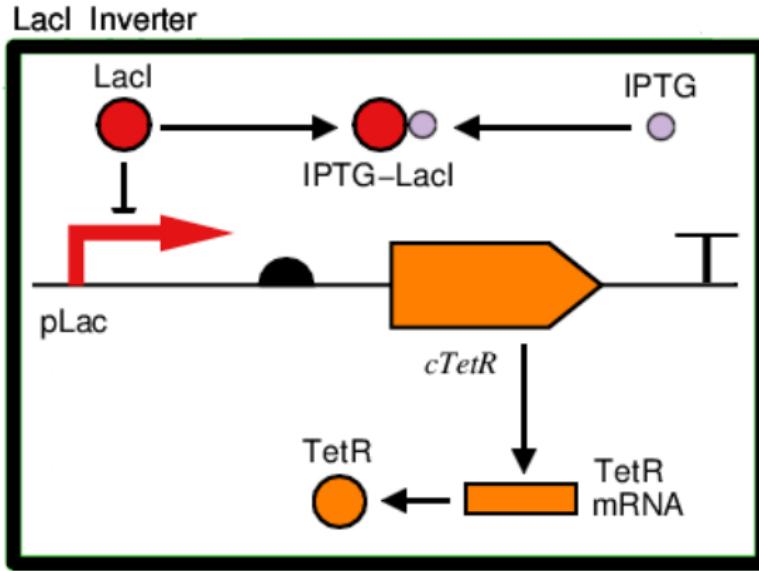
LacI Inverter



- Modules (logic gates, latches, oscillators, sensors, transducers, pathways, cascades)

Roehner et al., ACS Synthetic Biology (2014)

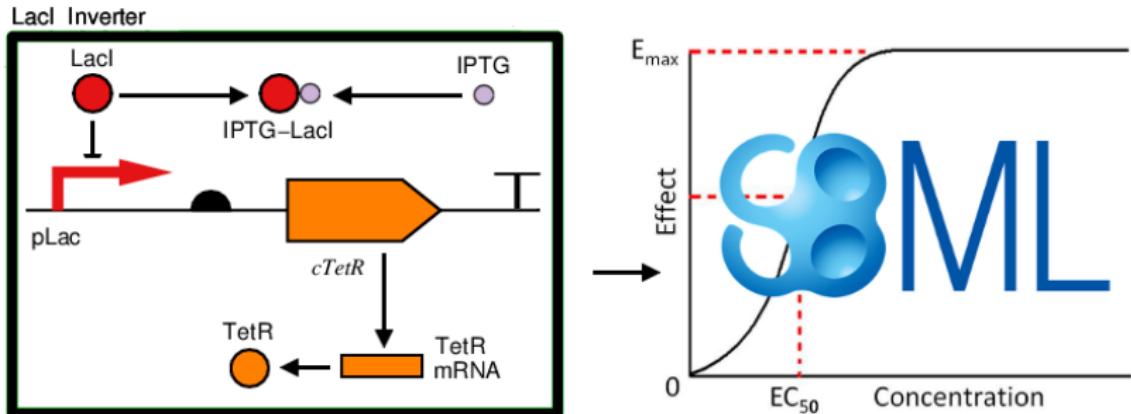
# Increasing SBOL's Functional Range



- Modules
- Interactions (activation, repression, complexation, transcription, translation, phosphorylation)

Roehner et al., ACS Synthetic Biology (2014)

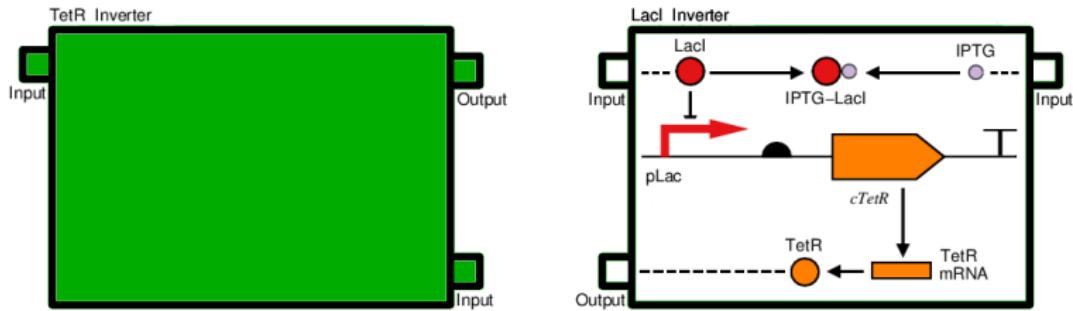
# Increasing SBOL's Functional Range



- Modules
- Interactions
- Models (SBML, CellML, Matlab, etc.)

Roehner et al., ACS Synthetic Biology (2014)

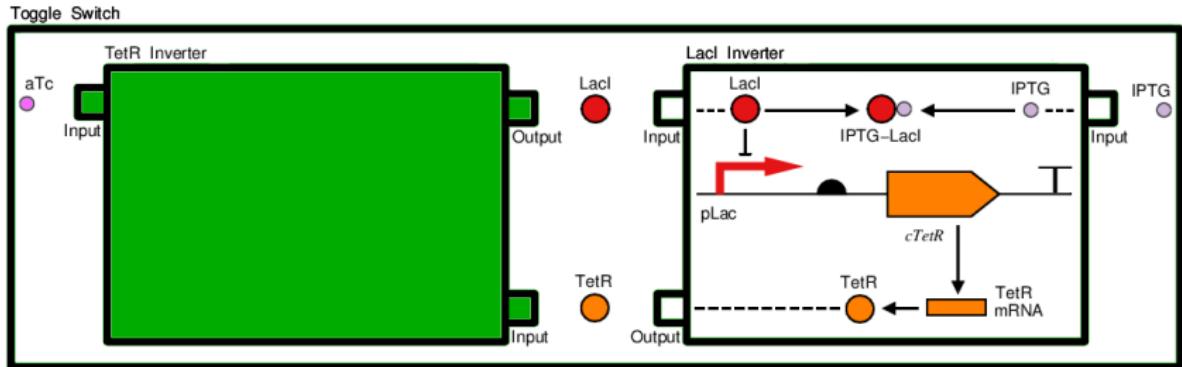
# Increasing SBOL's Hierarchy Support



- Ports

Roehner et al., ACS Synthetic Biology (2014)

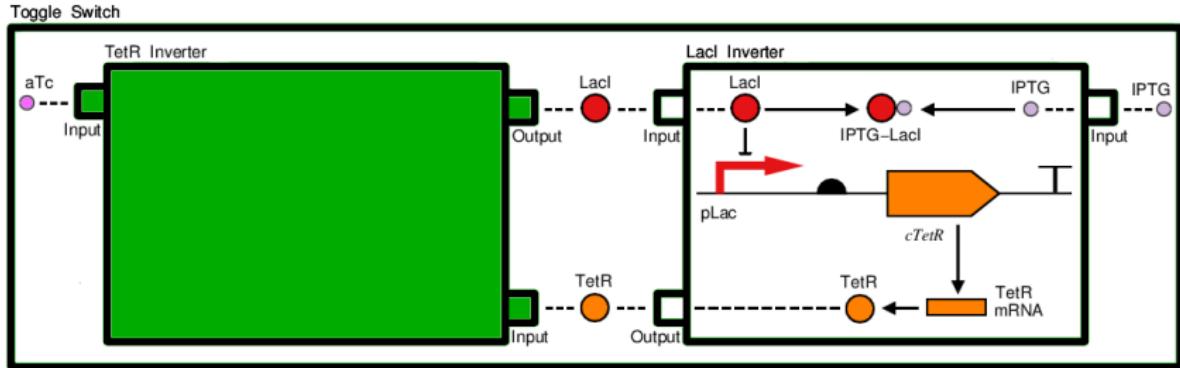
# Increasing SBOL's Hierarchy Support



- Ports
- Instantiation

Roehner et al., ACS Synthetic Biology (2014)

# Increasing SBOL's Hierarchy Support



- Ports
- Module Instantiation
- Port Mapping

Roehner et al., ACS Synthetic Biology (2014)

## Demo Session 1(a)

- Tour of SBML resources on SBML.org (Mike Hucka).
- Tour of the BioModels repository (Nicolas LeNovére).
- Demonstration of iBioSim, a tool that uses the SBML/SED-ML/SBOL standards and connects to the BioModels repository (Chris Myers).

# Installations of iBioSim

- Can follow along demo of iBioSim on your own computer, if you like.
- iBioSim requires Java Version 1.6 or higher installed on your system.
- MacOS - open iBioSim-2.7-MacOS.dmg, copy iBioSim.app into your applications directory.
- Windows - execute iBioSim-2.7-Setup.exe, follow instructions.
- Linux - execute iBioSim-2.7-Linux-x86\_64-Install, follow instructions.
- Native Java version:
  - If installations fails, fallback is to use a native Java version which supports most activities though not all.
  - Unzip iBioSim-2.7-Java.zip in directory you wish to install.
  - iBioSim.jar should be executable, but may require explicit command:  
`java -jar iBioSim.jar`
  - iBioSim.mac64 script makes it look better on Mac environment.
- Can help with installation during the breaks.

# Acknowledgments (SBOL)



Agilent Technologies

CLARK PARSIA



**Raytheon**

BBN Technologies

 AUTODESK

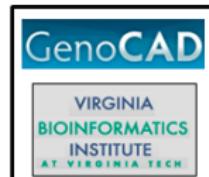
*life*  
technologies™



 DNA 2.0  
 GENOME  
COMPILER  
TOOLS FOR DESIGNING LIFE



National Human  
Genome Research  
Institute



Supported by National Science Foundation Grants  
CCF-1218095, DBI-1356041, and DBI-1355909.

# Acknowledgments (iBioSim)



Nathan Barker



Scott Glass



Kevin Jones



Hiroyuki Kuwahara



Curtis Madsen



Nam Nguyen



Tramy Nguyen



Tyler Patterson



Nicholas Roehner



Jason Stevens



Leandro Watanabe



Zhen Zhang



Supported by National Science Foundation Grants ECCS-0331270, CCF-07377655, CCF-0916042, CCF-1218095, and DBI-1356041.