Crowdsourcing Open Redesign Engine

SYSU-Software
BACKGROUND

Incompatible parts

Efficiency
BACKGROUND

Bank Redesign
BACKGROUND

Collaboration
Community
BACKGROUND

Crowdsourcing Open Redesign Engine
OUR GOAL

🔧 Design specific mathematical models and algorithms to improve prediction accuracy.

💻 Characterize and standardize circuits to organize standard parts and improve reproducibility of design and redesign.

☁️ Apply crowdsourcing-based design principle to collaboratively resolve design problems and simplify the design procedures.

🧪 Direct wet-lab experiments.
PROJECT

Features

Modeling

$X^2$

Techniques

Wet-lab Validation

Human Practices

?
FEATURES
FEATURES

Online version
https://core.sysusoftware.info
WORKFLOW

All Questions

53 views
Sunlight responсор system
I want a Sunlight responсор system.
2 votes
3 answers

18 views
Bxb1 protein monitoring
I need to assess the expression of Bxb1 protein, but it is hard to monitor by biochemical method. Anyone could help me?
1 votes
1 answers

31 views
Searching for ANDgate
How to set up an and gate? Could anyone help me?
0 votes
3 answers

16 views
AND GATE without T7

10 questions in total
What can CORE design exactly achieve?
CORE Design

Better design experience
FEATURES

Safer

Part Name: AtlPT4
Part type: gene

BBa Name: BBa_K1139100
Risk ground: Low risk(1)

Low risk(1)

Moderate risk(2)
You have finished your circuit, congratulations!

A control system, like a killer switch, may prevent biosafety problems and help your circuit work better in real world application. Are you sure you want to save your work without a control system?
FEATURES

💡 Inspiring
FEATURES

Flexible

Create a custom part

Part Type
Select a type

Part Name
Enter Part Name

BBa Name
Enter BBa Name

Risk ground
Risk ground

Introduction

Cancel  Create
FEATURES

CORE Design

Accurate simulation

Wet-lab assistant
How to make the simulation of circuit more accurate?
FEATURES

Dynamic Performance
The equations describing the change of concentration of protein over time

Dynamic Performance
The equations describing the change of concentration of protein over time
## FEATURES

### Protein Selection

<table>
<thead>
<tr>
<th>Protein:</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Simulation Data

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Simulation Options

- **Simulation Time**
  - Maximum time: Default is 3.0
  - Simulation interval: Default is 0.05
- **Initial Quantity**
  - CI: Default is 0
  - UVR8-TetR: Default is 0
  - GFP: Default is 0
  - TetR: Default is 0
  - Re-Plot
How to conduct the experiment more efficiently?
FEATURES

Current Design: Sunlight sensor

Choose a plasmid: Psb4T5
Choose a circuit: Circuit1
Choose a part: UVR8-tetR (BBa_K909009)
FEATURES

Part Information:

**Part Name:** UVR8-tetR  

**BBa Name:** BBa_K909009

- **Part type:**
- **Bacterium:**
- **Introduction:**
- **No introduction:**
- **Learn more:**
- **NCBI:**
- **Check this:**

This is the coding sequence:

```
atgggattcgcggaggaatgggtgcgccgacgaagttatgcctctctctgaaggtcttcattcatctcccgctgttagcactcgcgtgcctctctctcttagatagcatttttgtttctttggggtcaggagaggagacaccctgtgatatagcgtgttcggaggtatcggtatatgttgcttttggctttgtttgctttgcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggcttttggc
```
FEATURES

- **Safer**

- **Inspiring**  +  **Wet-lab need**  →  Integral solution

- **Flexible**
FEATURES

CORE Bank

Standardized documentation of genetic design
Reference:
**FEATURES**

**OR GATE.**

_descriptions_

This is an OR GATE. The input is L-arabinose or IPTG, the output is GFP. Either IPTG or L-arabinose existing can the GFP express. This device can be set on one or two circuit.

_modeling_

Equations

\[
\frac{dAraC}{dt} = \frac{\alpha_{AraC} \cdot L\text{-arabinose}^n}{K_d + L\text{-arabinose}^n} - k_{deg} \cdot AraC
\]

\[
\frac{dGFP}{dt} = -d \cdot GFP
\]
What does the Bank mean?
Stand on the shoulders of the giants.
Co-development
Collaborative building of genetic design
### Features

**Sunlight responder system**
I want a Sunlight responder system.

**Bxb1 protein monitoring**
I need to assess the expression of Bxb1 protein, but it is hard to monitor by biochemical method. Anyone could help me?

**Searching for ANDgate**
How to set up an and gate? Could anyone help me?

**AND GATE without T7**

---

**Crowdsourcing**
Collaborations  Crowdsourcing

Crowdsourcing is for everyone.
Gamers Decode AIDS Protein That Stumped Researchers For 15 Years In Just 3 Weeks (VIDEO)

FEATURES
I made a synbio design using C.O.R.E.!

Wow! C.O.R.E. is amazing! You'd better try it out :-)

Image:
FEATURES

Crowdsourcing + Synthetic biology = ?
FEATURES

Co-development
Crowdsourcing platform for collaborative design.

“Synbio Hub”

CORE Bank
A standardized genetic design repositories.

CORE Design
Integrated solutions for wet-lab assistance.
MODELING
Qianhui Wan
Integrated system of ODE building rules
\[ \theta = \frac{[L]^n}{K_d + [L]^n} = \frac{[L]^n}{(K_A)^n + [L]^n} = \frac{1}{\left(\frac{K_A}{[L]}\right)^n + 1} \]

**Theorem**

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>Fraction of the ligand-binding sites on the receptor protein occupied by the ligand</th>
</tr>
</thead>
<tbody>
<tr>
<td>[L]</td>
<td>Free (unbound) ligand concentration</td>
</tr>
<tr>
<td>( K_d )</td>
<td>Apparent dissociation constant derived from the law of mass action</td>
</tr>
<tr>
<td>( K_A )</td>
<td>Ligand concentration producing half occupation</td>
</tr>
<tr>
<td>( n )</td>
<td>Hill coefficient</td>
</tr>
</tbody>
</table>
Formulae Form Concerning genes and proteins

\[
\frac{d\alpha}{dt} = \chi_{Promoter} k_\alpha [\alpha_{gene}] - d[\alpha]
\]

Formulae Form of Promotion and Inhibition Relationships

\[
\frac{d[\text{LuxUp}]}{dt} = \beta_{cu} \frac{[\text{LuxU}]}{k_{cu} + [\text{LuxU}]} - d[\text{LuxUp}]
\]
Our Work
**Formulae for two certain parts**

LasR and lasR
\[
\frac{d\text{[LasR]}}{dt} = \chi_{promoter J23119} k_{\text{lasR}} \text{[lasR]} - d\text{[LasR]}
\]

LasR1 and A1i1
\[
\frac{d\text{[LasR]}}{dt} = \chi_{promoter J23119} k_{1} \text{[A1i]} - d\text{[LasR]}
\]

GFP1 and gfp1
\[
\frac{d\text{[GFP]}}{dt} = \chi_{promoter J23119} k_{2} \text{[gfp]} - d\text{[GFP]}
\]

\[
[gfp] = [gfp] \frac{[A1i]^n}{Kd \chi_{promoter J23119} + [A1i]^n}
\]

**Formulae for numerical simulation**

\[
\frac{d\text{[LasR]}}{dt} = \chi_{promoter J23119} k_{1} \text{[lasR]} [A1i] - d\text{[LasR]}
\]

\[
\frac{d\text{[GFP]}}{dt} = \chi_{promoter J23119} k_{2} \text{[gfp]} - d\text{[GFP]}
\]

\[
[gfp] = [gfp] \frac{[LasR]^n}{Kd \chi_{promoter J23119} + [LasR]^n}
\]
Numerical simulation function based on ODEs
TECHNIQUES & ALGORITHM

Shaofan Lai
Extensibility

Detailed Documentations
Extensible and easy-to-use API
Continuous Integration

Bug Tracking Facilities (Github Issues)

Changes Between Releases
Logistic Regression Algorithm
**Largest Set Matching**
Wet-lab VALIDITION

Xiangyue Hu, Xiaoran Yang
Our design - Sunlight responsor system
Plasmid 1
Plasmid 2
Our design - plasmid
Plasmid 3
Experiment results
Left: 1-4: CI-YFP; 5: UVR8-TetR; 6: TetR-GFP

Right: 1-2: Promoter-RBS-TetR-GFP; 3: RBS-TetR-GFP
## WET LAB

<table>
<thead>
<tr>
<th>Hour</th>
<th>Units of fluorescein</th>
<th>OD600</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Whole1</td>
<td>Whole2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2218</td>
<td>1455</td>
</tr>
<tr>
<td>2</td>
<td>3324</td>
<td>3271</td>
</tr>
<tr>
<td>3</td>
<td>1358</td>
<td>2398</td>
</tr>
<tr>
<td>4</td>
<td>5453</td>
<td>5215</td>
</tr>
<tr>
<td>5</td>
<td>4228</td>
<td>4352</td>
</tr>
<tr>
<td>6</td>
<td>5123</td>
<td>4700</td>
</tr>
<tr>
<td>7</td>
<td>6879</td>
<td>3145</td>
</tr>
<tr>
<td>8</td>
<td>7968</td>
<td>8036</td>
</tr>
<tr>
<td>9</td>
<td>13096</td>
<td>13701</td>
</tr>
<tr>
<td>10</td>
<td>27550</td>
<td>27550</td>
</tr>
</tbody>
</table>

Our wet-lab experimental data
Experiment results-
The simulation result from CORE modeling module
Goodness of fit: SSE: 4.627e+06 R-square: 0.9923; Adjusted R-square: 0.979; RMSE: 1115

The results above indicate that our models accord well with experimental data.
CONCLUSION

New Simulation Module and Algorithm
✔ Advanced ’Largest Set Matching’ algorithm
✔ New and efficient models
✔ Dynamic Performance, Static Performance

Crowdsourcing-based Circuit Design and Redesign
✔ Automatic design of larger circuits
✔ Collaboratively resolve design problems
✔ Simplification of design procedure
✔ Reproducibility of design and redesign

Wet-lab Validation
✔ Successfully validated our model with a self-practice wet-lab experiment
HUMAN PRACTICES

All team members
HUMAN PRACTICES

Overview

Scientific and Societal Thinking

Integrated into Project

Project Assessment

Public Engagement
HUMAN PRACTICES

Meet up
HUMAN PRACTICES

What is useful in wet-lab practices?

Even software track should pay attention to biosafety.
— By NUDT_CHINA

Communication between iGEM teams isn’t enough. Neither is reproducibility.
— By Nankai

Is genetic design the only way software helps Synbio?
— By HFUT-China
HUMAN PRACTICES

Parts with & without BBa number
HUMAN PRACTICES

Overview

Scientific and Societal Thinking

Integrated into Project

Project Assessment

Public Engagement
HUMAN PRACTICES

The origins of the parts of the known parts in database

Results of Searching Parts with Red Flags against the Database

<table>
<thead>
<tr>
<th>Parts with Red Flags</th>
<th>Search Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBa_L10001</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K177004</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K177005</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K177010</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K177013</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K177026</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K177029</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K299810</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K299811</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K299812</td>
<td>Not Found</td>
</tr>
<tr>
<td>BBa_K299813</td>
<td>Not Found</td>
</tr>
</tbody>
</table>
Biosafety Poker
HUMAN PRACTICES

Overview

Scientific and Societal Thinking

Integrated into Project

Project Assessment

Public Engagement
Synbio Box
Learn Synthetic Biology in a game.
HUMAN PRACTICES

Garden Party
HUMAN PRACTICES

- Studying wet-lab’s needs in meetups
- Reproducibility in iGEM (a survey)
- Safety and security (consult with experts)
- Intellectual property (considerations)
- Laws and regulations (a talk with student on humanity)
- Ethics (consult with professor)

- Risk assessment (safe parts in database)
- Feasibility assessment (approachable safe design)
- End-user considerations (easy-to-use while extensible)

- Synbio Box: a game on synbio
- Mini lecture to freshmen
- Poker on biosafety
Collaborate with iGEM Software Track and host our software on Heroku.

Improve the function of ‘Synbio Hub’, to better serve the community.

Validate the models with more experiment results.
MEDALS

Bronze The following 6 goals must be achieved:
√ Register for iGEM, have a great summer, and attend the Giant Jamboree.
√ Complete the Judging form.
√ Create and share a Description of the team's project using the iGEM wiki.
√ Present a poster and a talk at the iGEM Jamboree.
√ Create a page on your team wiki with clear attribution of each aspect of your project.
√ Develop and make available, via the iGEM GitHub page.

Silver In addition to the Bronze Medal requirements, the following 3 goals must be achieved:
√ Provide a comprehensive, well-designed User Guide for your software and upload it to your wiki.
√ Develop a well documented library or API for other developers.
√ Demonstrate that you followed best practices in software development.

Gold In addition to the Bronze and Silver Medal requirements, at least two additional goals must be achieved:
√ Demonstrate how your team has identified, investigated and addressed one or more of these issues in the context of your project.
√ Address a problem that you would like to co-develop with a wetlab team.
√ Re-use and further develop a previous iGEM software project.
√ Demonstrate your software at the iGEM Giant Jamboree in the software demo suite.

We fulfilled all of these requirements!
ACKNOWLEDGEMENTS

Instructors
Prof. Jianhua Yang
Prof. Yan Zhang
Prof. Jian Ren
Prof. Yongjun Lu
Prof. Xionglei He
Prof. Shengfeng Huang
Dr. Lingling Zheng
Dr. Yongdong Zhang
Prof. Zhumei He

Special Thank
Prof. Chuan-Hsiung Chang
Thank you,

Any question?
Poster Session
No. 178, Zone 6
Backup Slides

• 1. Libraries
• 2. Wet-lab details
• 3. Modeling
• 4. Github
• 5. Dry-lab Testing
• 6. User Studies
• 7. Collaborations
• 8. Human Practices
• 9. Crowdsourcing
• 10. Source of our protocols
• 11. Why C.O.R.E.
## Libraries

### Backend
- **Framework**
  - Flask, Flask-Login, Flask-Mail, Flask-Script, Flask-SQLAlchemy
- **Unit test**
  - nose
- **Scientific calculation**
  - Numpy, Scipy
- **Automated documentation**
  - Sphinx
- **Others**
  - gevent, Pillow, WTForms

### Frontend
- **Basic styles**
  - Semantic UI
- **Framework**
  - vue.js
- **Graphs**
  - jsPlumb, angularPlasmid.js, highcharts.js, chart.js
- **Formula rendering**
  - Math.js, MathJax
- **Others**
  - fullCalendar, jQuery
Wet-lab details  Dynamic Performance

The equations describing the change of concentration of protein over time

Output (concn [a.u.])

0k  10k  20k  30k

0.00  0.50  1.00  1.50  2.00  2.50  3.00  3.50  4.00  4.50  5.00  5.50  6.00  6.50  7.00  7.50  8.00  8.50  9.00  9.50  10.00

Dynamic Performance

The equations describing the change of concentration of protein over time

Output (concn [a.u.])

0k  10k  20k  30k

0.00  1.00  2.00  3.00  4.00  5.00  6.00  7.00  8.00  9.00  10.00

Series 1: 4.9115 [a.u.]
Modeling

• assumptions
  Leakage
the distribution of RNA polymerases, ribosomes and chemicals
growth and division of cells
limitations
cell types and strains
binding constants
other modeling method
GitHub & “Synbio Hub”

• Same: Promotes collaboration

• Differences:

  – 1. Github for CS students, while “Synbio Hub” for iGEMers and Synbioers.
  
  – 2. “Synbio Hub” utilizes “crowdsourcing” as a way to collaborate.

  – 3. “Synbio Hub” mainly focuses on a standard “bank” for projects with detailed documentation, while GitHub emphasise the importance of collaboration on coding.
Dry-lab Testing

- **Installability test** on MacOS X & Windows 10
- **Risk assessment**
  - 1. No parts of Risk Group 3 or 4 organisms in database
  - 2. No parts with Red Flags in database
User Studies

- Covering students of different background:
- Yu Zhou (team leader of 2015 NJU-China)
- Jiyong Ma (bioinformatics graduates in Ren’s lab)
- Jinyu Li (PhD at Yale University)
- Haoquan Zhao (Master at Harvard University)
User Studies
Collaborations

- A mouse movement tracking tools (NJU-China)
- Mentored high-school teams on project planning
- A garden party (as Human Practices) with SYSU_CHINA
- ...

Human Practices

- Ethics: Consult with Prof. He for advice on bioethics of software
- Intellectual property: well-documented project with attributions.
- Laws and Regulations: consult Jiaming Hui (student of Humanity) and discuss on the public’s consensus on biosafety.
Crowdsourcing

**Sunlight respon sor system.**

I want a Sunlight respon sor system.

dsk

- asked 25 Sep, 2015
- viewed 28
- active 27 Sep, 2015
Crowdsourcing

Brief Description:
It is a sunlight sensor.
A UV sensor. This may help you.

0 comments
Crowdsourcing
Crowdsourcing

**Brief Description:**

toggle switch

**Full Description:**

A toggle switch. This may help you.
Crowdsourcing

**Brief Description:**
This circuit consists of a sunlight (UVB) sensor and a toggle switch. The E. coli will produce GFP protein without sunlight while they will produce YFP protein under sunlight.

A sunlight responser system. This may help you.

1 comments

**dsk** 27 Sep. 2015
That's what I need! It seems functional!
Source of our protocols

• Reference:

• 2. iGEM protocol
# Why choose our software

<table>
<thead>
<tr>
<th></th>
<th>CORE</th>
<th>Genome Compiler</th>
<th>SnapGene</th>
<th>Primer3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Free or not?</strong></td>
<td>√</td>
<td>√</td>
<td>SnapGene Viewer is free, but SnapGene is not.</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>(Open Source)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Designed for Synbio?</strong></td>
<td>√</td>
<td>(Collaborative building of genetic design.)</td>
<td>Interact with Registry, but mainly for genome processing.</td>
<td>For molecular biologists.</td>
</tr>
<tr>
<td></td>
<td>(Collaborative building of genetic design.)</td>
<td></td>
<td></td>
<td>For primer design in wet-lab.</td>
</tr>
<tr>
<td><strong>Online sharing and communication?</strong></td>
<td>√</td>
<td>√ (Comment on project and data)</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td><strong>Based on standard parts?</strong></td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td><strong>Facilitate wet-lab in which aspects?</strong></td>
<td>Integrated solution; modeling, plasmid, protocol management, experiment scheduling</td>
<td>Genome assembly</td>
<td>Inspection of plasmid; endonuclease restriction sites of plasmid.</td>
<td>Primer design</td>
</tr>
<tr>
<td><strong>Modeling?</strong></td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>(mathematical modeling + curve fitting of experimental data)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>