

Bi9690en Synthetic Biology – Lecture 3

Karel Říha, CEITEC MU, A26/209
e-mail: karel.riha@ceitec.muni.cz

 @karlriha



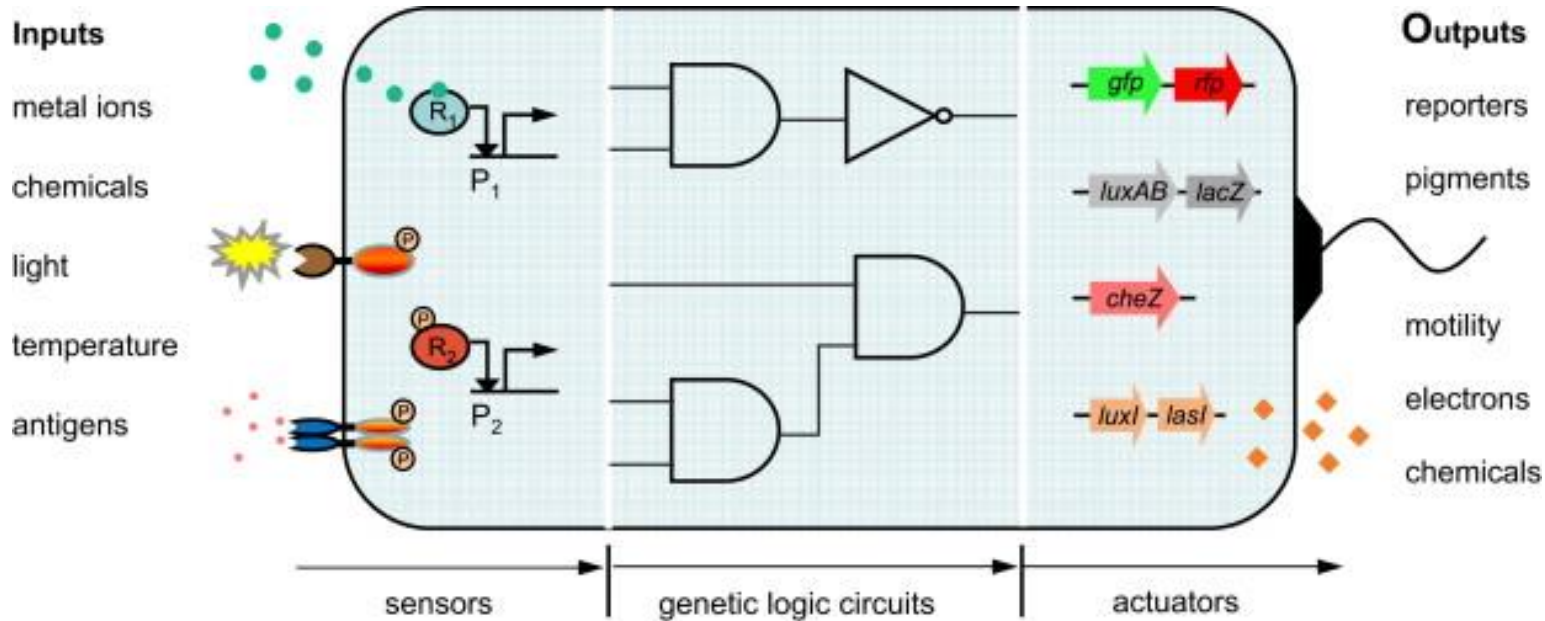
Basic concepts in synthetic biology

Cellular programming and logic gates

The synthetic biology design cycle



Cells are replicating living computers



Cells live in an ever-changing environment and continuously sense, process and react to environmental signals using their inherent signaling and gene regulatory networks.



Synthetic biology: basic concepts

Synthetic biology is in many aspects similar to **electric engineering**: cellular decision-making processes share basic operations with electronic control circuits. Intra- and extracellular information is collected by sensors that communicate the input signal states into a network, which processes the data according to logic and arithmetic operations. These operations result in decisions that are finally executed by output signals.



- Synthetic biology follows a hierarchical structure, building up systems from smaller components.
- An important aspect of synthetic biology is the application of systemic design. This approach is based on the engineering principles of **modularity**, **characterization** and **standardization**.



Synthetic biology: abstraction hierarchy and modularity

Engineering

Synthetic Biology

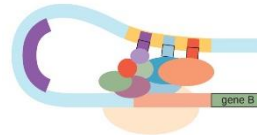
Parts



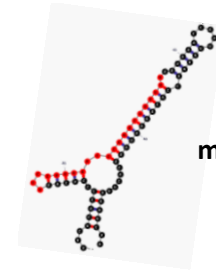
protein



promoter

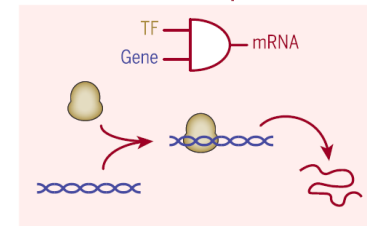
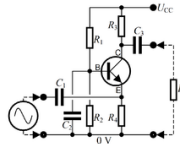


miRNA



Basic biological functions

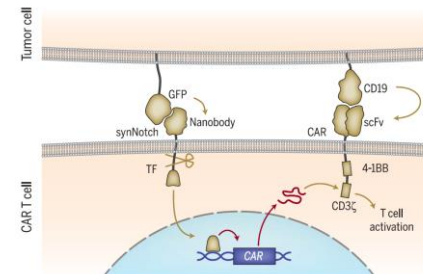
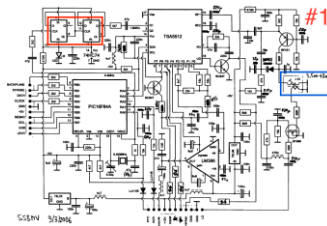
Devices



Inducible protein expression module

Combination of parts performing a human defined function

Systems



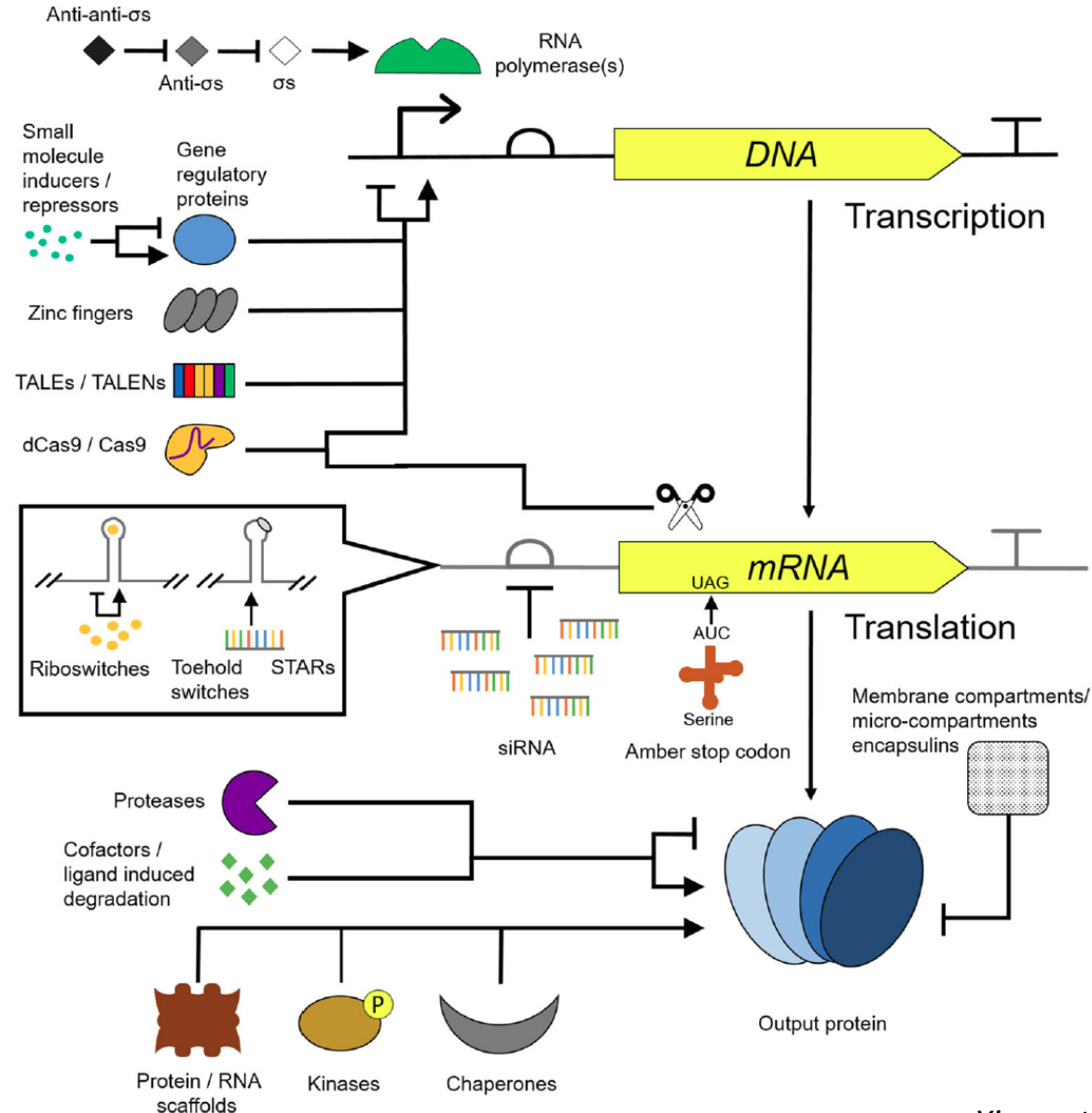
Synthetic therapeutic program to specifically kill a cancer cell

Combination of devices



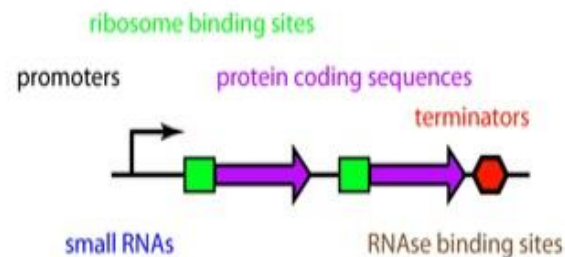
Tools (parts) for regulating gene expression

Fig. 3 Expanded toolbox for engineering complex gene regulation programs. These include using proteins that affect DNA transcription and RNA translation through protein-DNA and protein-RNA base pair binding. Also shown is the ability to use RNA secondary structure and base pair binding to control mRNA translation initiation. Protein activity can also be controlled by other proteins, through protein-protein interactions or enzymatic reactions that modulate activity. The activity of many regulators can be controlled by small molecule ligands/cofactors. (σ = sigma factors, STARs = small transcriptional activating RNAs, siRNA = small interfering RNA, TALE(N)s = transcription activator-like effector (nuclease))

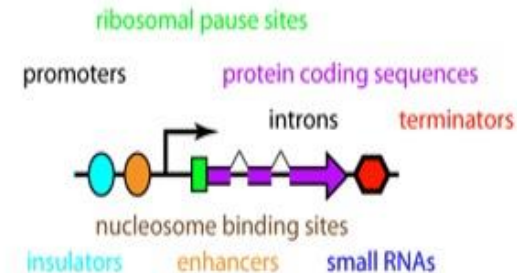


Parts

- **Promoters:** recruits RNA polymerase and other accessory proteins to prime the transcription of mRNA. Promoters have different strength and can be constitutive or regulated (inducible, repressible).
- **Terminators:** signal termination of transcription, polyA signals.
- **Ribosome binding sites:** recruit ribosome for initiation of translation (Shine-Delgarno nad Kozak sequences). These sequences affect efficiency of translation and hence protein production.
- **Translational riboswitches:** a regulatory segment within a mRNA that can bind a small molecule, which in turn affects translation (eg lysin riboswitch).
- **Protein coding sequences:** encode transcriptional activators and repressors, transcription factors, sensors (e.g. cryptochromes for light perception, receptors for chemical ligands), signaling molecules (kinases, proteases), protein scaffolds and „output“ proteins (antibiotic resistance, GFP, LUX, suicidal proteins).
- **Factors affecting RNA stability:** bacterial mRNA has very short half life (average 2 min in E.coli). In contrast, stability of eukaryotic mRNA can range from minutes to days. *Cis*-elements affecting RNA stability: length of 3' UTR, hairpins, RNA binding motives, introns.
- **siRNAs, miRNAs:** affect gene expression on postranscriptional or translational level.



Prokaryotic Genetic Parts



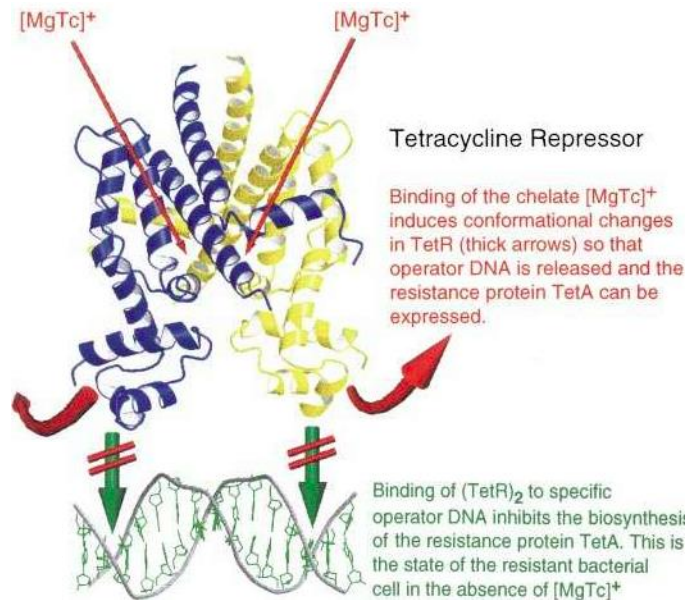
Eukaryotic Genetic Parts



Orthogonality in synthetic biology

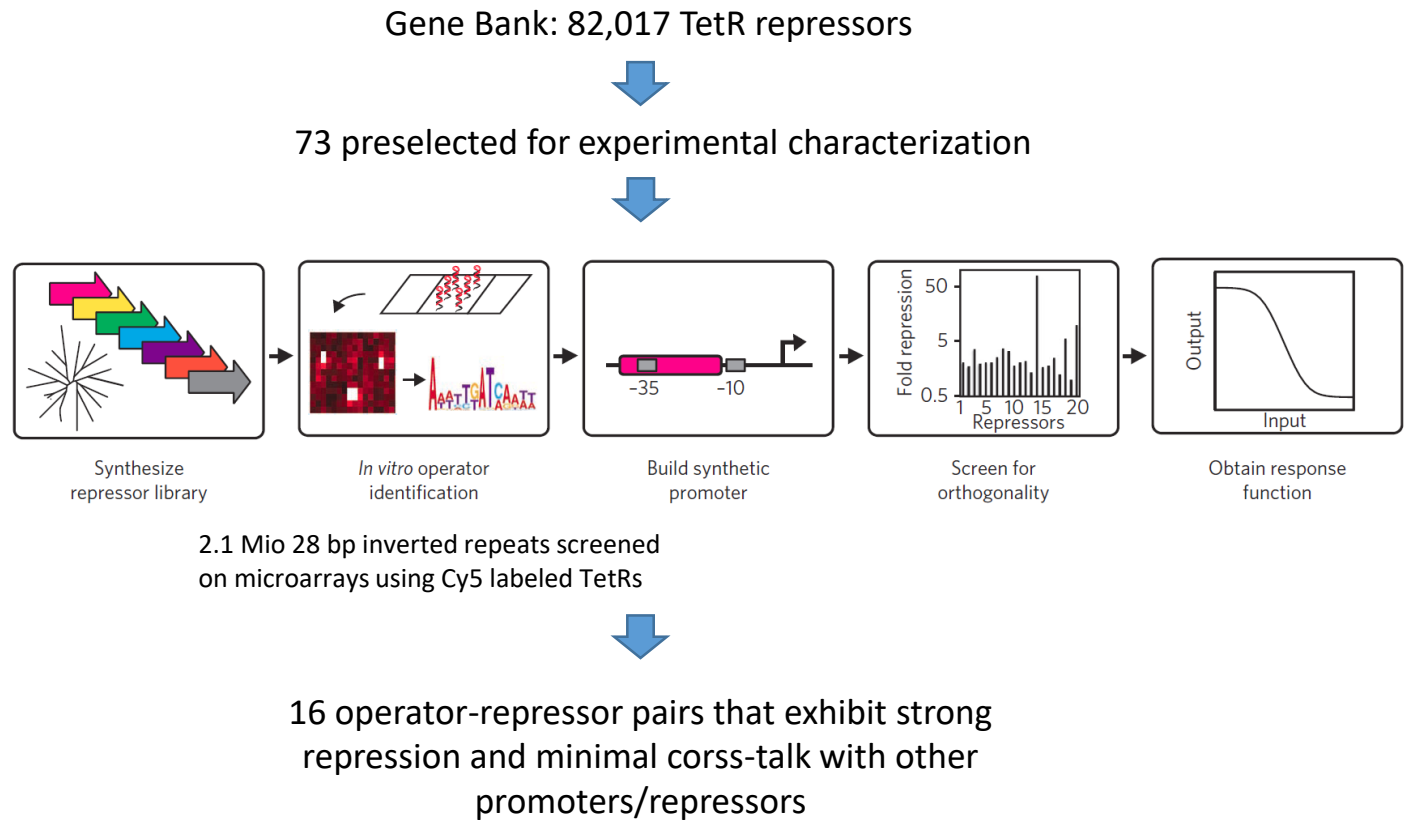
Synthetic biology approaches commonly introduce heterologous gene networks into a host to predictably program cells, with the expectation of the synthetic network being **orthogonal** (non-interfering) to the host background and to other synthetic networks. It also implies context independent performance of a synthetic network.

TetR repressor



- Relatively short operator sequence
- Often used in synthetic genetic circuits
- Thousands of homologues that exhibit binding specificities to disparate operators are available

Mining of prokaryotic genomes for orthogonal TetR repressors



Synthetic transcription regulators based on TALE and CRISPR/Cas9

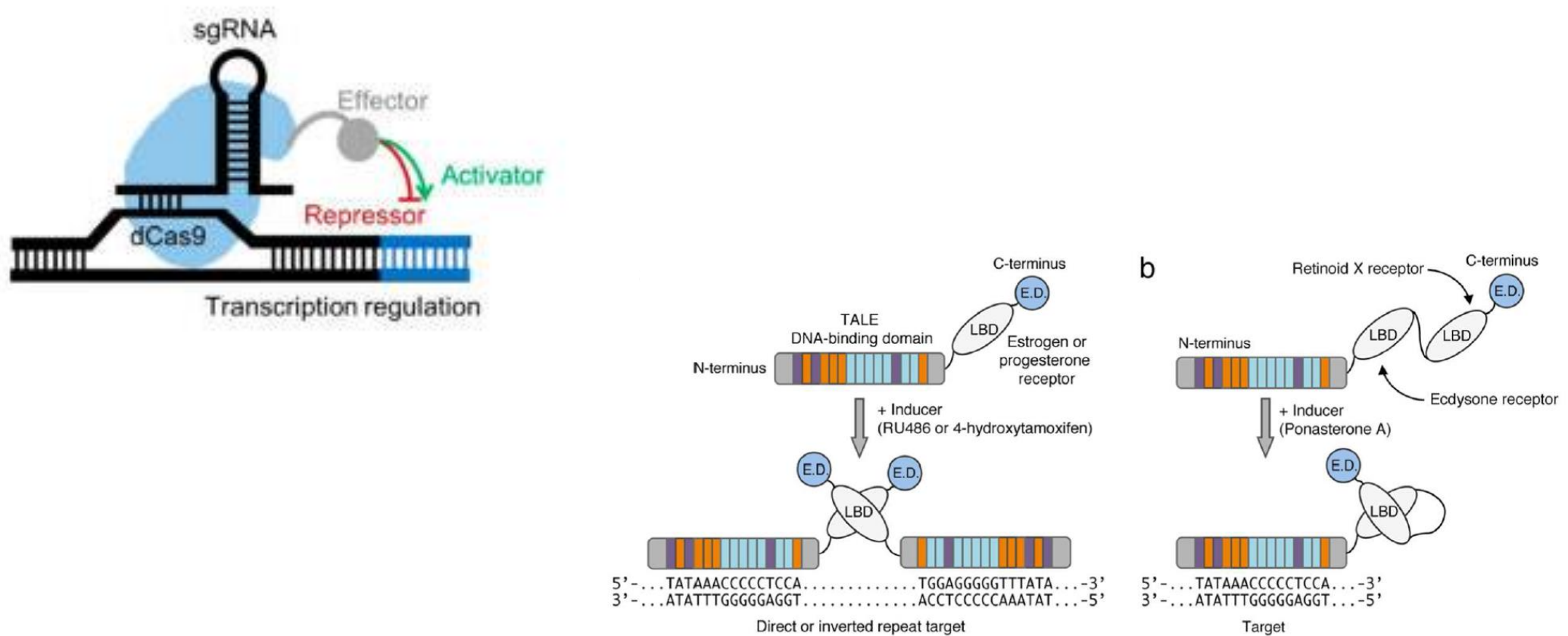
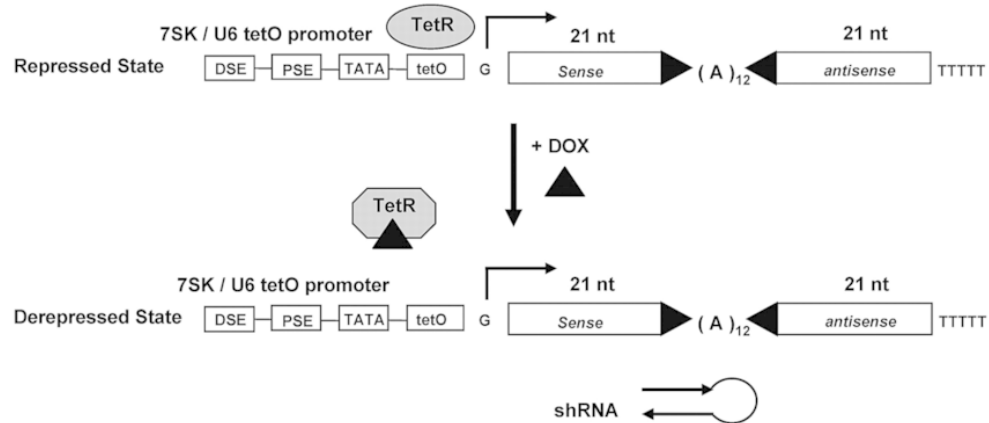


Figure 1. Ligand-inducible TALE transcription factors. (a) TALE-TF proteins fused to ligand-binding domains (LBDs) from the estrogen receptor (ER) or progesterone receptor (PR) undergo intermolecular dimerization in response to 4-hydroxytamoxifen (4-OHT) or RU486, respectively, and up-regulate gene activation from DNA sequences that contain two direct or inverted repeat TALE binding sites. (b) TALE-TF proteins fused to the chimeric single-chain retinoid X- α /ecdysone (RXE) LBD undergo intramolecular rearrangement in response to ponasterone A (PonA) and up-regulate gene activation from target DNA that contains only a single TALE binding site. E.D. indicates effector domain.

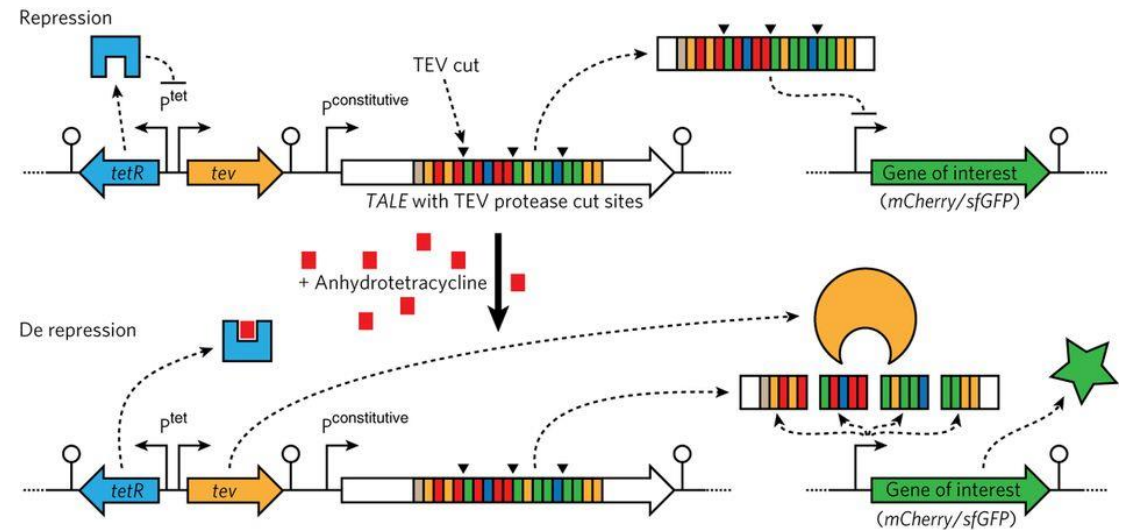
Devices

- Combination of parts performing a human defined function.



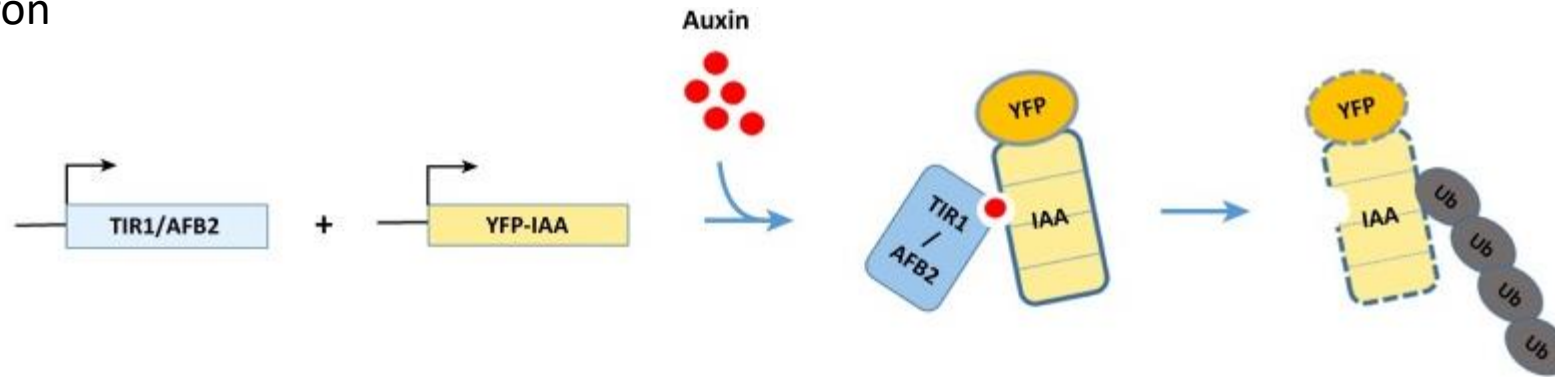
Tetracycline regulated siRNA production

Tetracycline regulated protein production via proteolytic cleavage

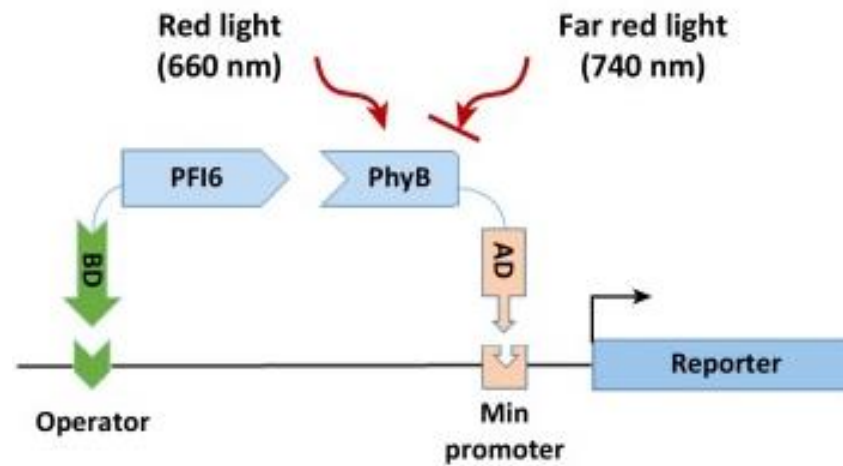


Devices

Auxin inducible degron



Red light regulated expression



Synthetic logic gates and cellular computation

A logic gate in **electronics**, is a physical device which is implemented with a Boolean function based on input and output signals (0 and 1). It executes a logical function on **one or more inputs** that produce a **single output**. Logic gates are used for storing the data that can be constructed by connecting several gates in a Flip-flops circuit which is a central building block of digital electronics systems in computers and communications.

In **biological systems**, logic gates are synthetic gene circuits programmed to permit the expression of an output protein only when a strictly defined signature of input signals is matched. Genetic elements interact with regulatory proteins to switch a **gene ON or OFF** while RNA or protein concentration can serve as input or output. One of the key approaches of synthetic biology is to **reprogram** the decision-making gene networks in order to implement them as **logic gates** in living systems.

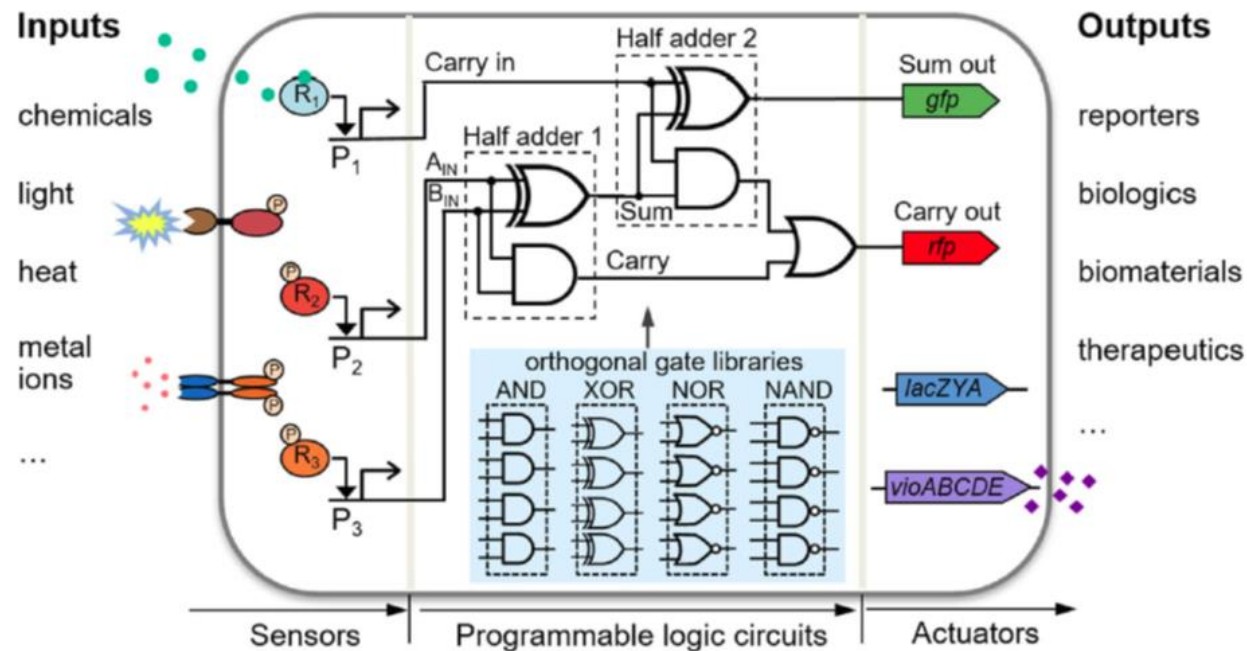


Fig. 1 Programmable cellular computation with scalable signal processing capacity. To achieve large-scale control of cellular behaviour, an expanded library of versatile orthogonal genetic regulatory blocks and associated wiring principles are needed. For example, a genetic 1-bit full adder program adds binary numbers, it has 3 inputs and 2 outputs, and can be constructed from 5 modular

logic gates that are wired in 3 layers and selected from well-characterized orthogonal gate libraries. The genetic circuits can be coupled to modular input genetic sensors and output actuators to achieve complex decision making for a variety of human desired applications

Logic gates: the AND gate

Input A	Input B	Output AB
0	0	0
1	0	0
0	1	0
1	1	1

The **AND** gate gives a high output (1) only if both the inputs are high.

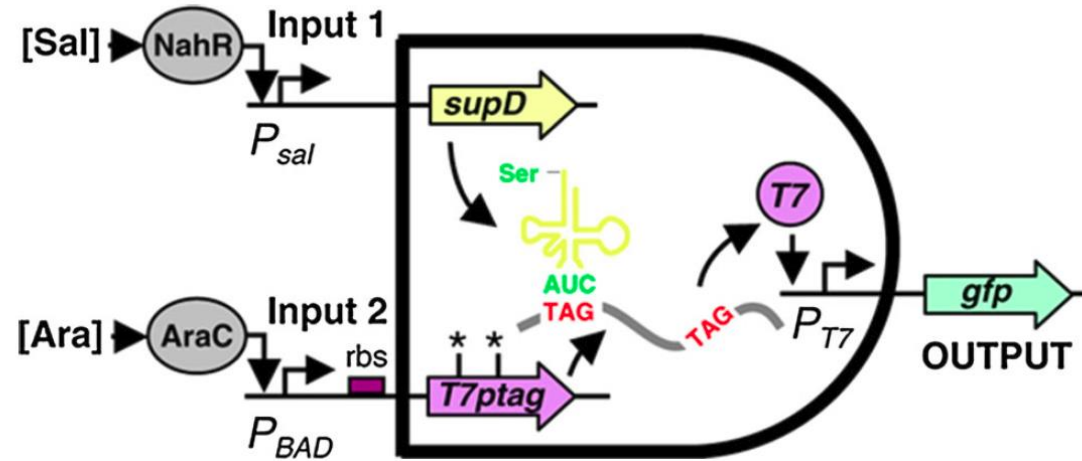
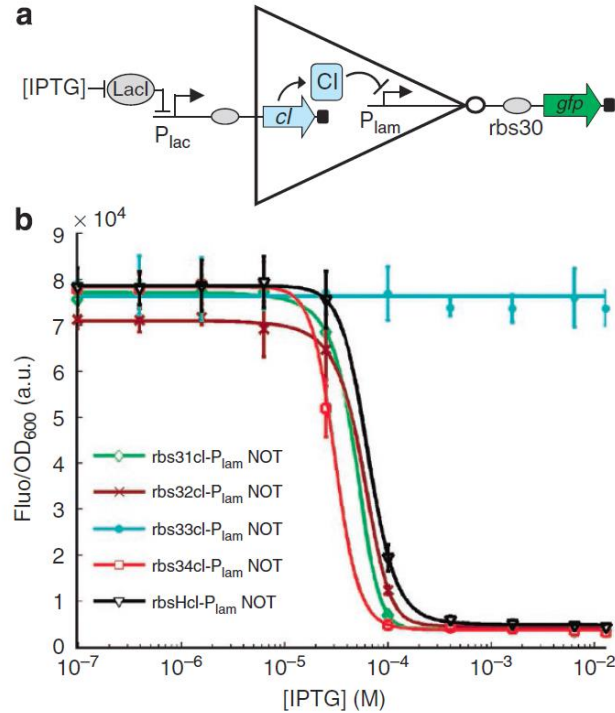


Fig. 2 Design of AND gate based on two promoters. The first promoter was linked with transcription of amber suppressor tRNA *supD* and second T7 RNA polymerase. Polymerase was modified to contain two amber stop codons and translated as serine when *supD* was transcribed and T7 pol was expressed when both *SupD* and *T7ptag* mRNA are present. Figure reproduced with permission from Anderson et al. (2007)

Singh, Syst Synth Biol 2014



The NOT gate module



- A **NOT** gate takes a single input and inverts it, so 0 becomes 1; 1 becomes 0.
- Even these simple gates can perform signal-processing functions, for example, converting a dark sensor into a light sensor.

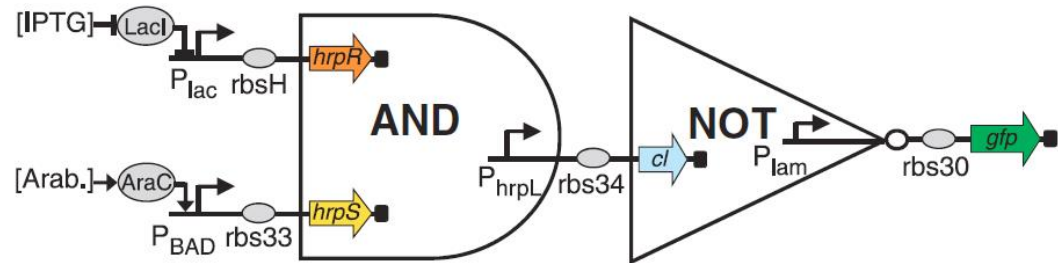
The modular NOT gate was designed on the basis of the *cl*/*P_{lam}* repressor module consisting of lambda gene *cl* and its regulatory *P_R* promoter.



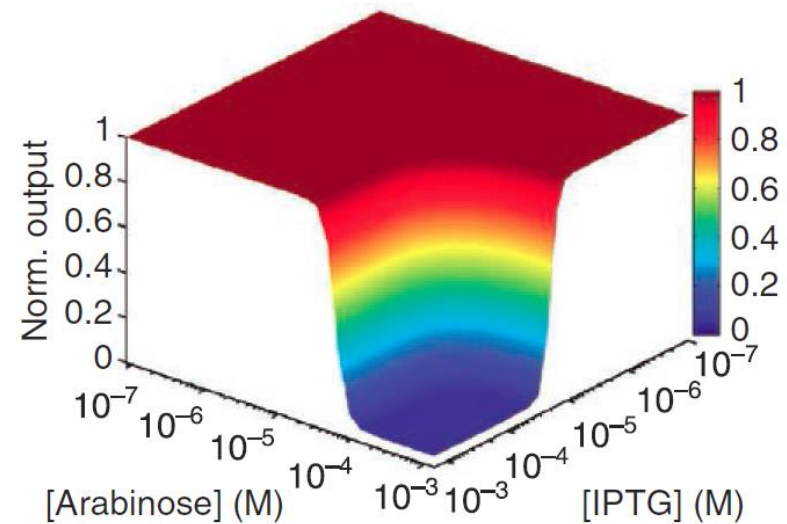
The NAND gate

Input A	Input B	Output AB
0	0	1
1	0	1
0	1	1
1	1	0

The output of **NAND** gate is high if any of inputs are low.



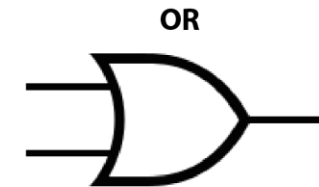
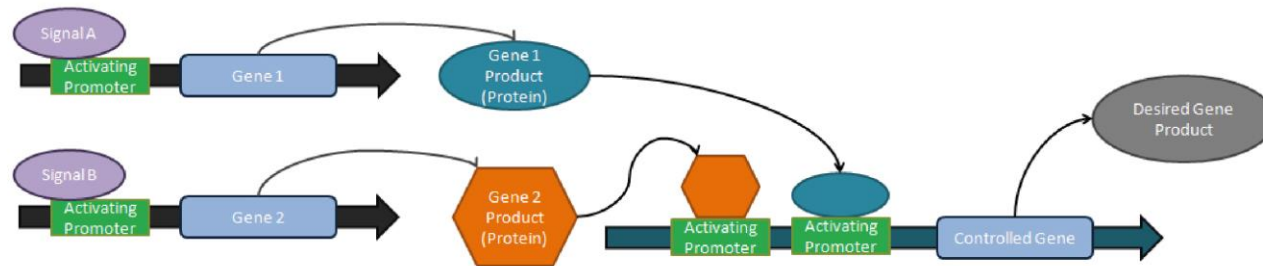
This NAND gate results from the combination of AND and NOT gates. The AND gate is derived from hetero-regulation module from *Pseudomonas syringae*. The device comprises two co-activating genes *hrpR* and *hrpS* controlled by separate promoter inputs, and a σ_{54} -dependent *hrpL* promoter driving the output. The *hrpL* promoter is activated only when both genes are expressed.



The OR gate

Input A	Input B	Output A + B
0	0	0
1	0	1
0	1	1
1	1	1

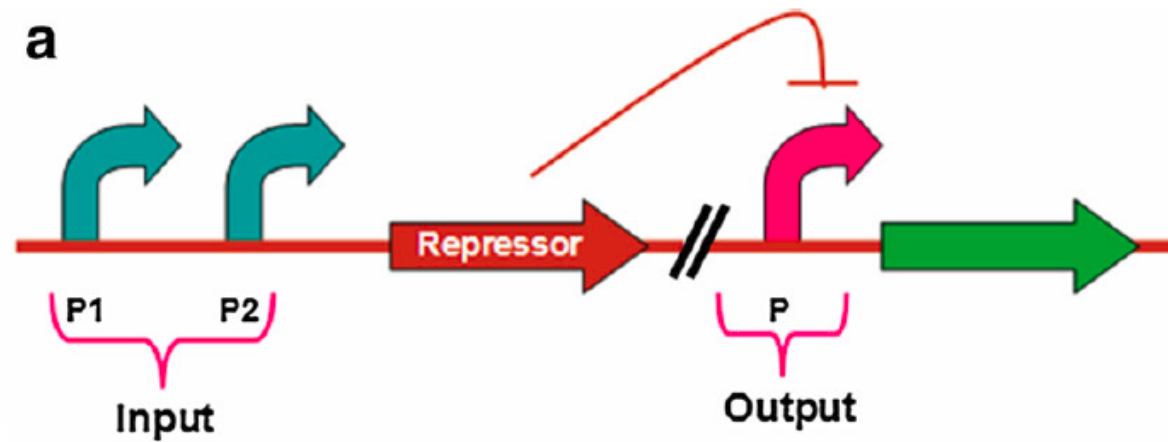
An **OR** gate gives output 1 as long as either (or both) of the inputs are 1.



The NOR gate

Input A	Input B	Output A + B
0	0	1
1	0	0
0	1	0
1	1	0

A **NOR** gate is equivalent to OR gate followed by a NOT gate. The outputs of NOR gates are low, if any of the inputs are high.

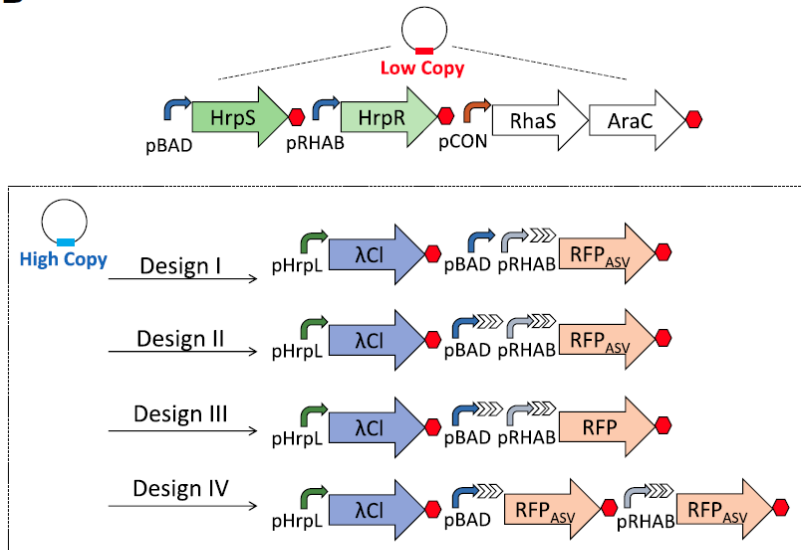


The XOR gate

An **exclusive-OR (XOR)** gate gives a high output only if either input is present.

Input A	Input B	Output A + B
0	0	0
1	0	1
0	1	1
1	1	0

B



C

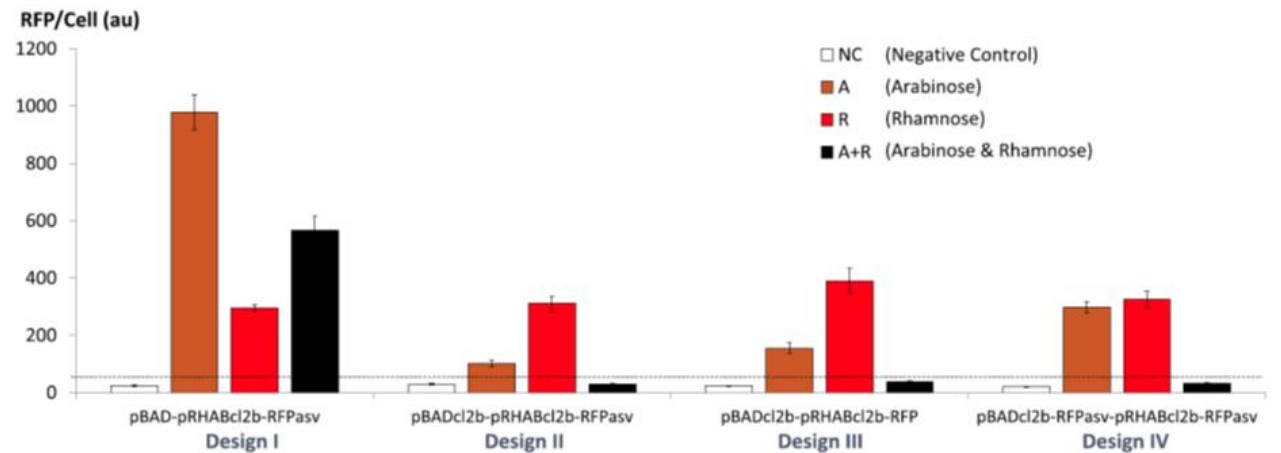
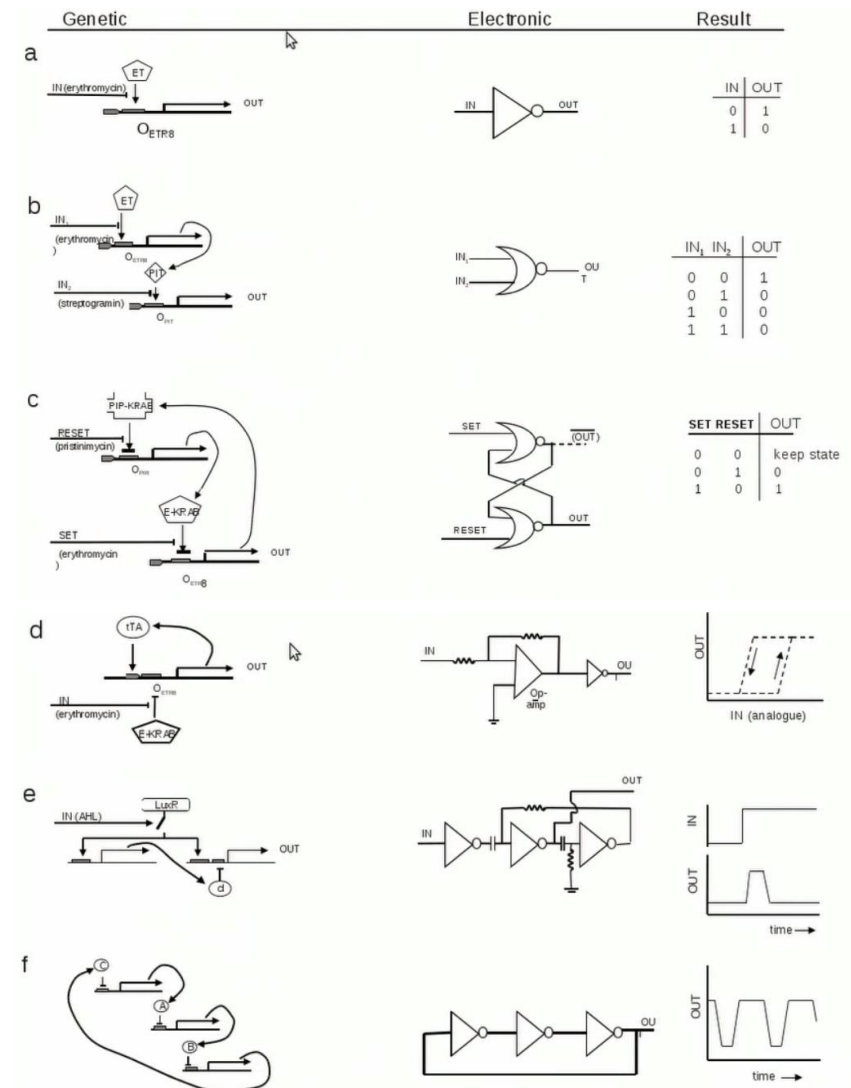
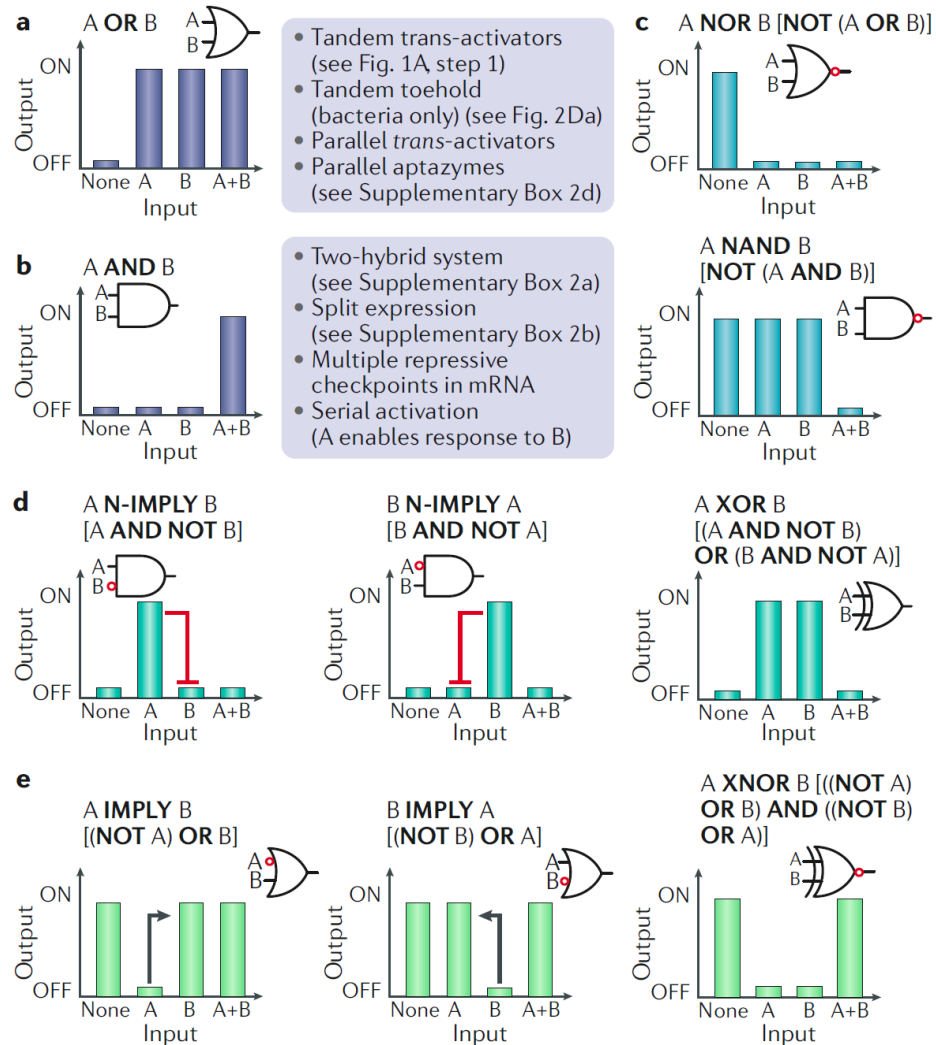


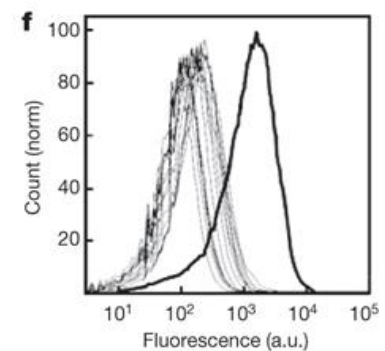
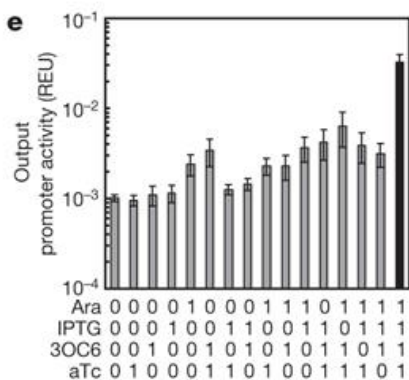
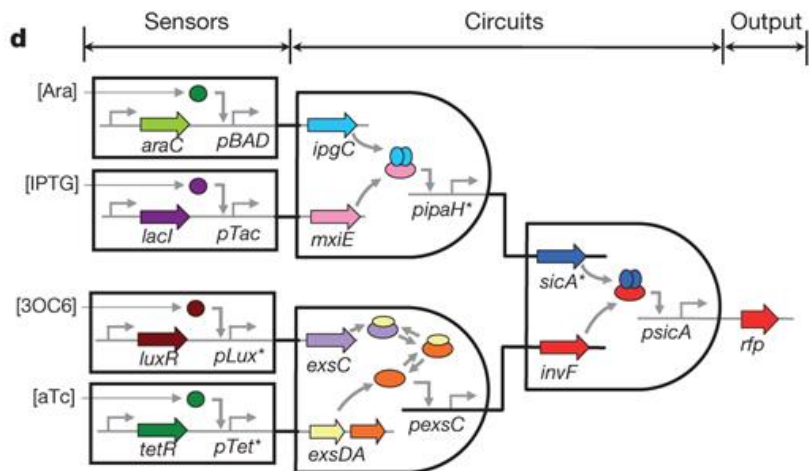
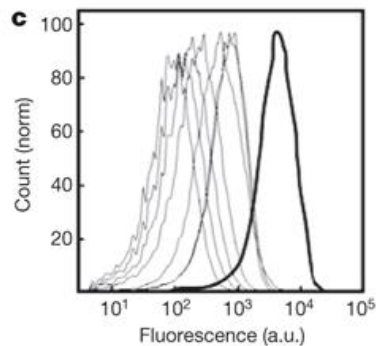
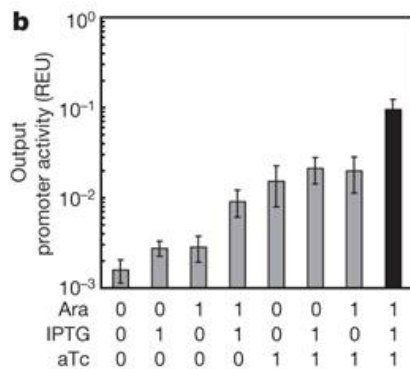
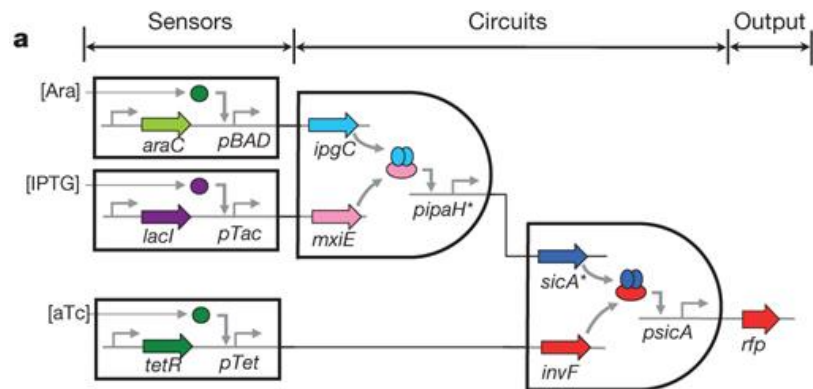
Fig. 5 Design and characterisation of biological XOR gates. **a** The logic output of XOR gate. **b** Genetic blueprint of four biological XOR gate designs. The XOR gate comprises serially layered AND, NOT and OR gates. HrpRS transcription factors are carried in a low copy plasmid, while pHrpL-λCI and distinct modules of OR gates with lambda repressor binding sites expressing RFP reporter are carried in high copy plasmids. Design I comprises tandem promoters with repressor binding sites downstream of pRHAB promoter and an RFP reporter engineered with the ASV protein degradation tag. Designs II and III comprise tandem promoters with repressor binding sites downstream of each promoter and RFP with and without the ASV degradation tag, respectively. Design IV is modified from design II with RFP expressed in two disparate transcripts. **c** Digital performance of various designs of biological XOR gates at steady state. **d** The steady state profile of XOR gate IV for various concentrations of arabinose (input A) and rhamnose (input B). Error bars represent the standard deviation of four independent experiments



Assembly of logic gates allows building circuits with complex behaviour

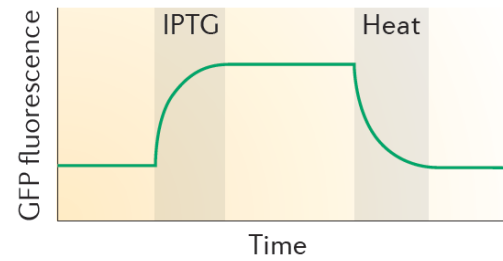
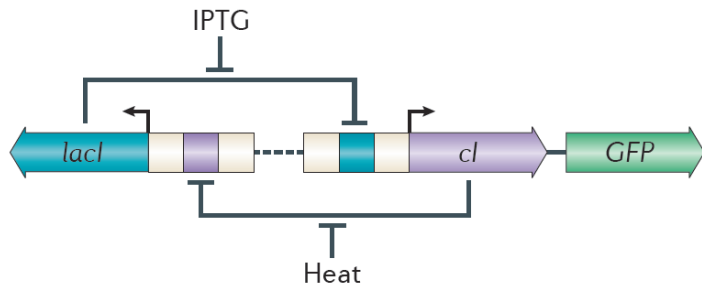


A multilayer AND gates

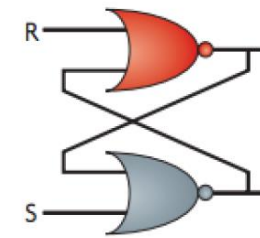


Toggle switch and repressilator

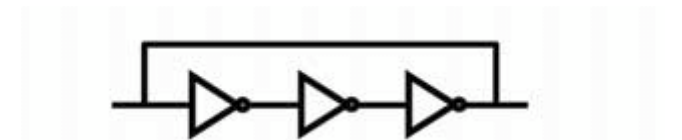
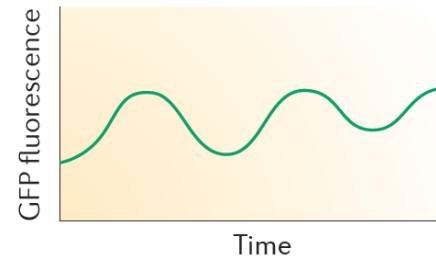
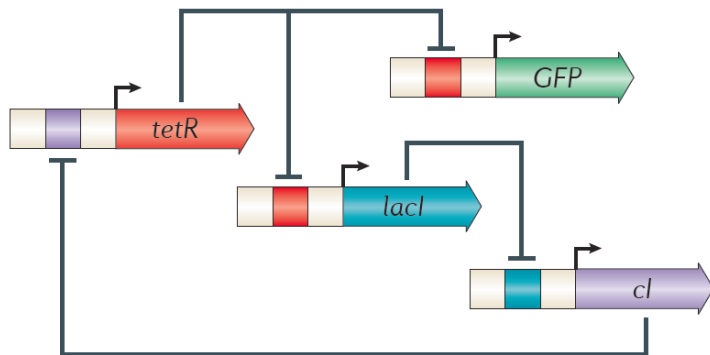
Toggle switch



Reset-set latch

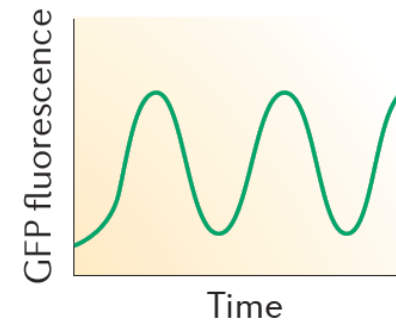
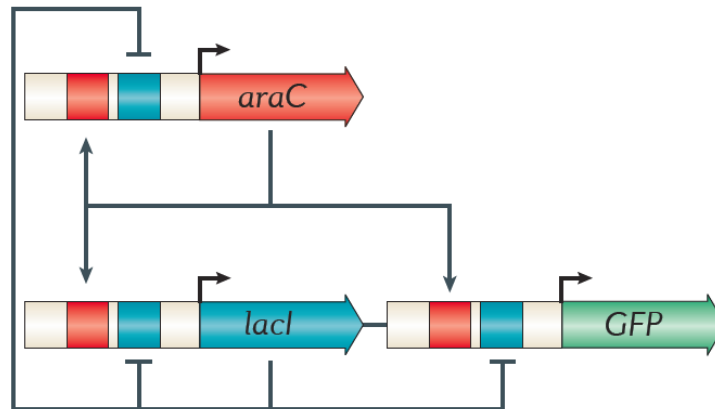
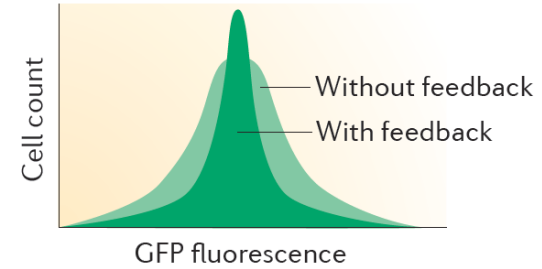


Repressilator



Relaxation oscillator uses autoregulatory feedback

Negative feedback loop rescues the noise and results in a narrow expression distribution



Registry of parts for synthetic biology

The basic concept in engineering is using a combination of standard parts to produce standard devices, which are then combined to produce standard systems.

MIT registry for the International Genetically Engineered Machine (iGEM)

http://parts.igem.org/Main_Page

parts.igem.org/Main_Page

IGEM wiki tools search PRODUCTION 2017 SERVER login

Registry of Standard Biological Parts

tools catalog repository assembly protocols help search

iGEM 2019 Begins!

Add and Document Parts

Start adding and documenting your parts now! Your parts should be well characterized and measured, and follow the Registry's requirements.

Sample Submissions

iGEM teams do not need to send samples of their parts this year. We want teams to focus on the documentation of their parts! Teams must follow 2019 requirements for parts, including BioBrick RFC10 or Type IIS compatibility.

Registry Updates

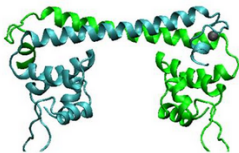
The Registry will be undergoing updates (some major, some minor) over the next few months. If you notice any issues with functionality, please let us know at [hq \(at\) igem \(dot\) org](mailto:hq@igem.org).

Featured Part

Metal Binding and Sensing Parts

Every year, a number of iGEM teams complete a variety of biosensors and bioremediation projects that involve metal-binding and metal-sensing. Their focus may be on several pollutants or just one. iGEM teams have worked with metals like nickel, mercury, lead, arsenic, copper, amongst others.

We've put together a collection of projects and DNA parts that are responsible for both metal binding and metal sensing.



DNA Synthesis Offer: IDT

IDT is once again generously offering 20 kb of DNA as gBlocks® Gene Fragments free of charge to each iGEM 2019 team! Click here to go to IDT's partner offers page for more info.

2019 DNA Distribution

The iGEM 2019 DNA Distribution has started shipping to registered and approved iGEM teams! Be sure to read through the 2019 Distribution Handbook for storage instructions and how to use your kit!

Collections

We've updated the Registry part collections. Users can discover new parts and collections and build upon what previous iGEM teams and labs have achieved.

- Well Documented Parts
- Frequently Used Parts
- Plant Chassis
- Bacillus subtilis*
- Reporter Proteins
- Anderson Promoters

Registry Help

Before starting your projects, be sure to read through our help pages. If you can't find an answer to your question, contact [hq \(at\) igem . org](mailto:hq@igem.org).

Useful help topics:

- BioBrick Prefix and Suffix
- Assembly Standards
- Assembly Compatibility
- 3A Assembly Overview
- Requesting Parts
- Add a Part

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.

Protocols

Be sure to read through the Registry recommended protocols. They've been tested using the DNA Distribution Kit and BioBrick parts. See the Registry

- Standard parts must be thoroughly characterized and their performance well described
- Data on standardized parts are organized in registries of parts for synthetic biology
- iGEM Registry provides a resource of available biological parts that have been user-tested and characterized for users developing synthetic biology projects.
- iGEM Registry is an open community that runs on the "Get & Give & Share" philosophy. Users get parts, samples, data, and tools to work on their synthetic biology projects. They'll give back to the Registry the new parts they've made, as well as data and experience on new and existing parts.
- iGEM Registry contains about 20,000 parts
- The parts on the iGEM Registry adhere to the BioBrick standard allowing them to be assembled together creating new longer and more complex parts, while still maintaining the structural elements of the standard. This allows the engineer to focus on design instead of assembly.
- BioBrick Assembly Standard 10 is based on restriction cloning

```
Prefix                                     Suffix
5' - GAATTC GCGGCCGC T TCTAGA G ...part... T ACTAGT A GCGGCCGC CTGCAG - 3'
EcoRI      NotI      XbaI                                     SpeI      NotI      PstI
```

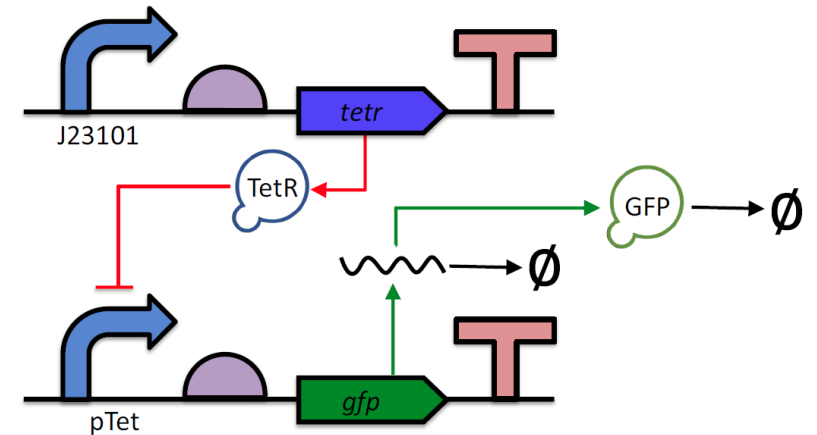
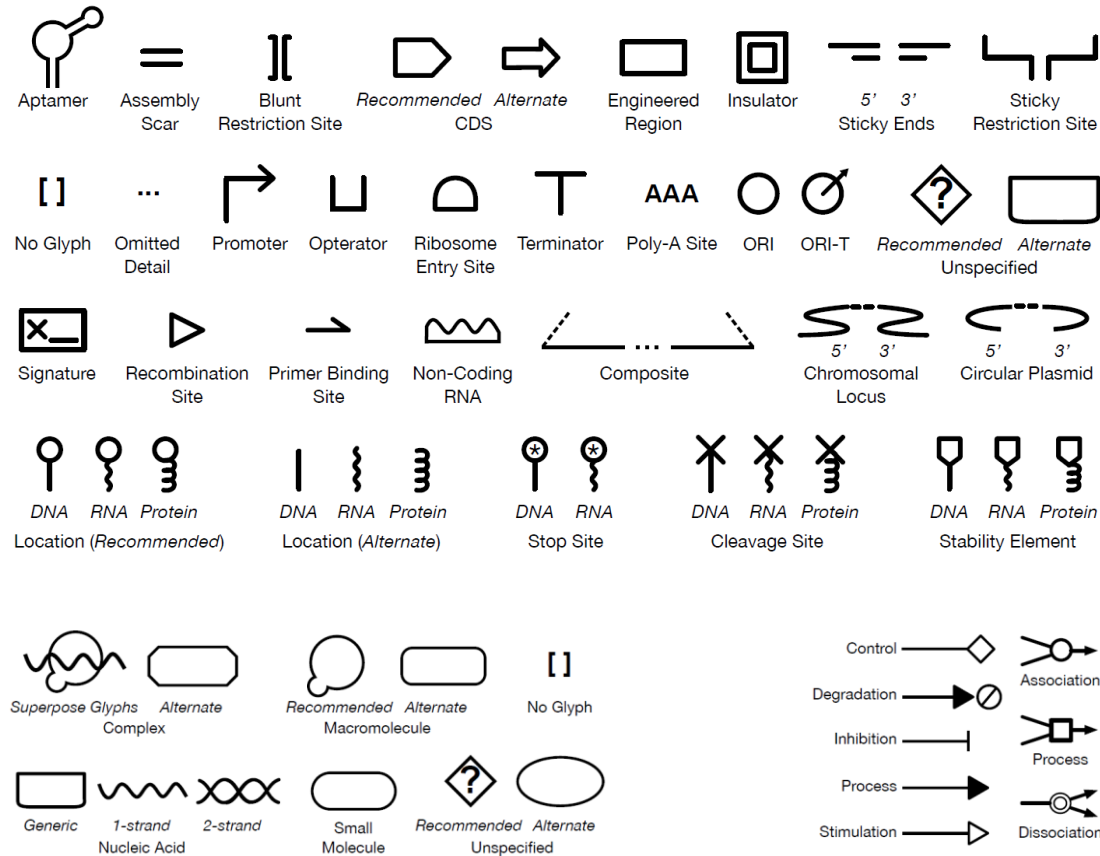
- Other registries: SYNbis Database http://synbis.bg.ic.ac.uk/synbis2/Welcome_Page.html



Standardization: The Synthetic Biology Open Language (SBOL)

<https://sbolstandard.org/>

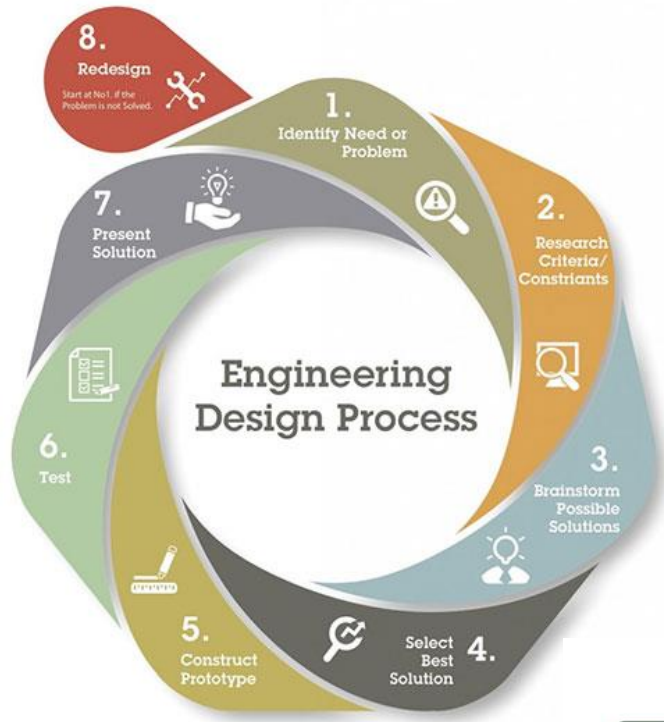
- SBOL is an open standard for the representation of in silico biological designs.
- SBOL also provides schematic glyphs to graphically depict genetic designs called SBOL Visual.



The top functional unit produces the TetR protein constitutively, under control of promoter J23101. TetR represses the pTet promoter, which is regulating production of GFP. The diagram of GFP production explicitly includes the intermediate mRNA and the degradation of both the mRNA and protein products.



The synthetic biology design cycle



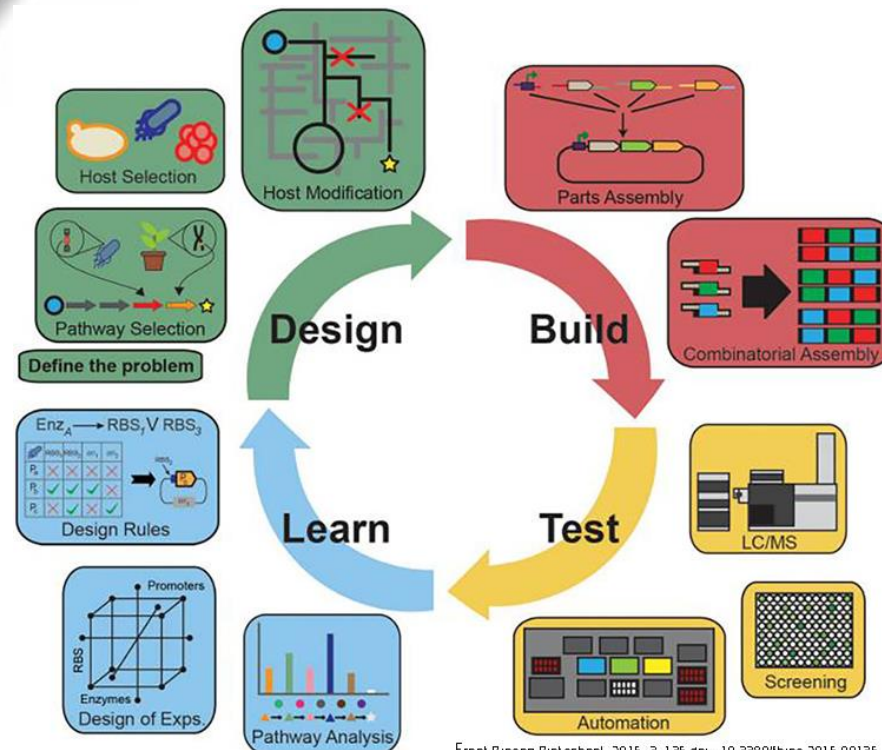
Specification: formal definition of the desired function and design of a target genetic system.

Design: the set of decisions needed to determine the constructs and hosts, to be used and/or modifications to the host to be made. Involves also creating a plan for composing the DNA constructs from their elements.

Build: Implementing DNA assembly plan and construction of the biological system.

Test: design and implementation of experiments for characterizing engineered systems and accompanying analysis and data interpretation.

Learn: include approaches to allow for revision of designs based on experimental outcomes.

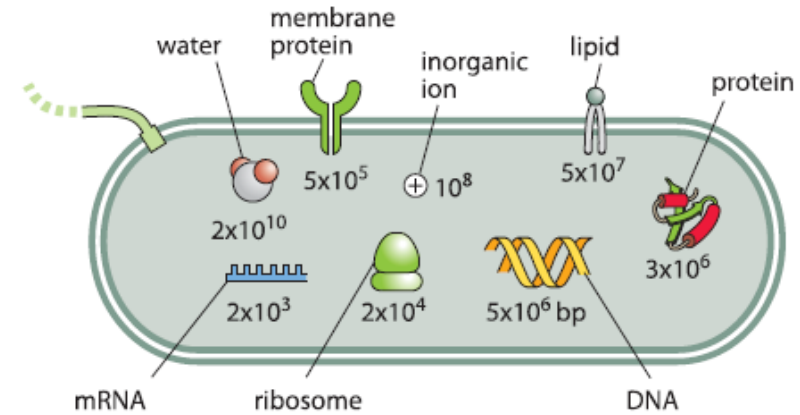


Synthetic biology - complexity of biological systems

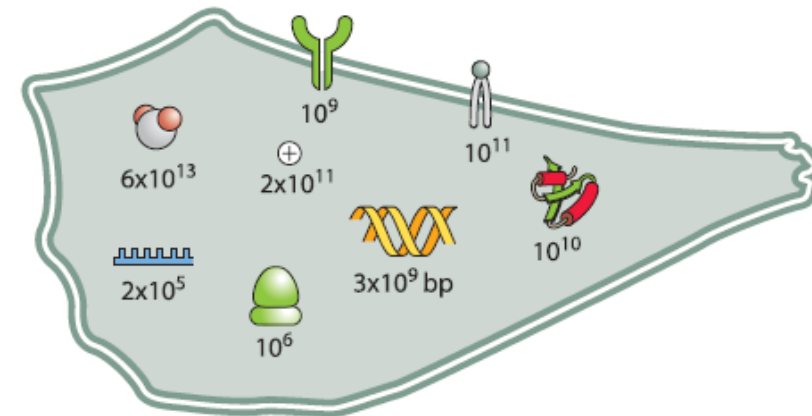
A single car has about 30,000 parts, counting every part down to the smallest screws.



(A) bacterial cell (specifically, *E. coli*: $V \approx 1 \mu\text{m}^3$; $L \approx 1 \mu\text{m}$; $\tau \approx 1$ hour)

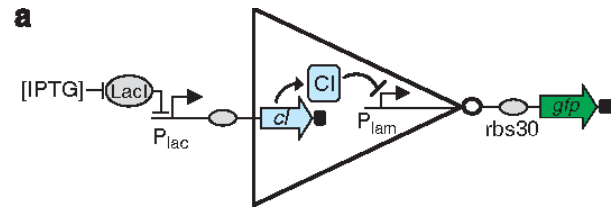


(C) mammalian cell (specifically, HeLa: $V \approx 3000 \mu\text{m}^3$; $L \approx 20 \mu\text{m}$; $\tau \approx 1$ day)

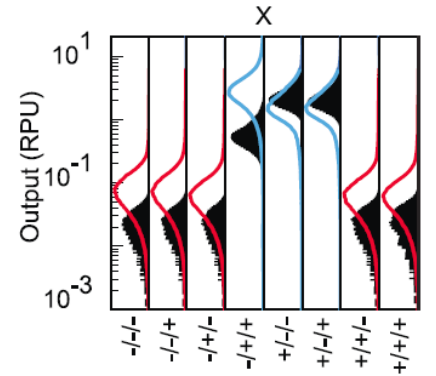
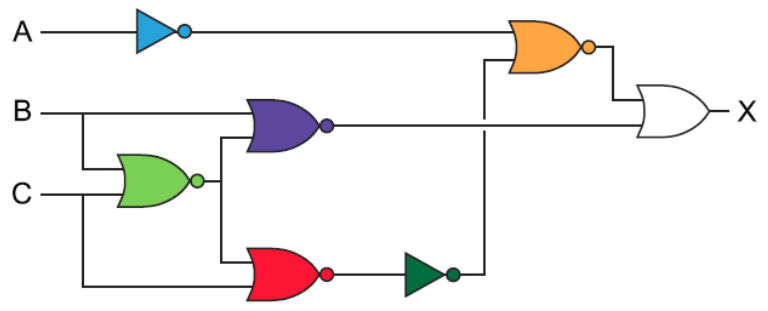


Design efforts are confronted with the complexity and unpredictability of engineering inside living cells.

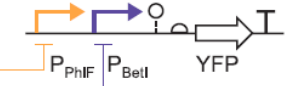
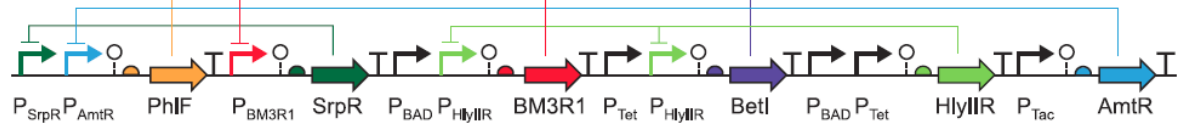




0x1C



50 parts



Biological Computer Aided Design (BioCAD)

- BioCAD assists the *de novo* design and selection of existing genetic components to achieve a desired biological activity, as part of an integrated design-build-test cycle.
- BioCAD tools facilitate the design of larger systems from smaller genetic parts by providing users with visual, textual or programming-language-like interfaces, or automatically generating designs from intended function.
- As the field moves towards real-world applications, tools that can adequately predict functionality from design will be indispensable.
- Predictability of part behavior, particularly in different contexts, is still a major issue in synthetic biology design.

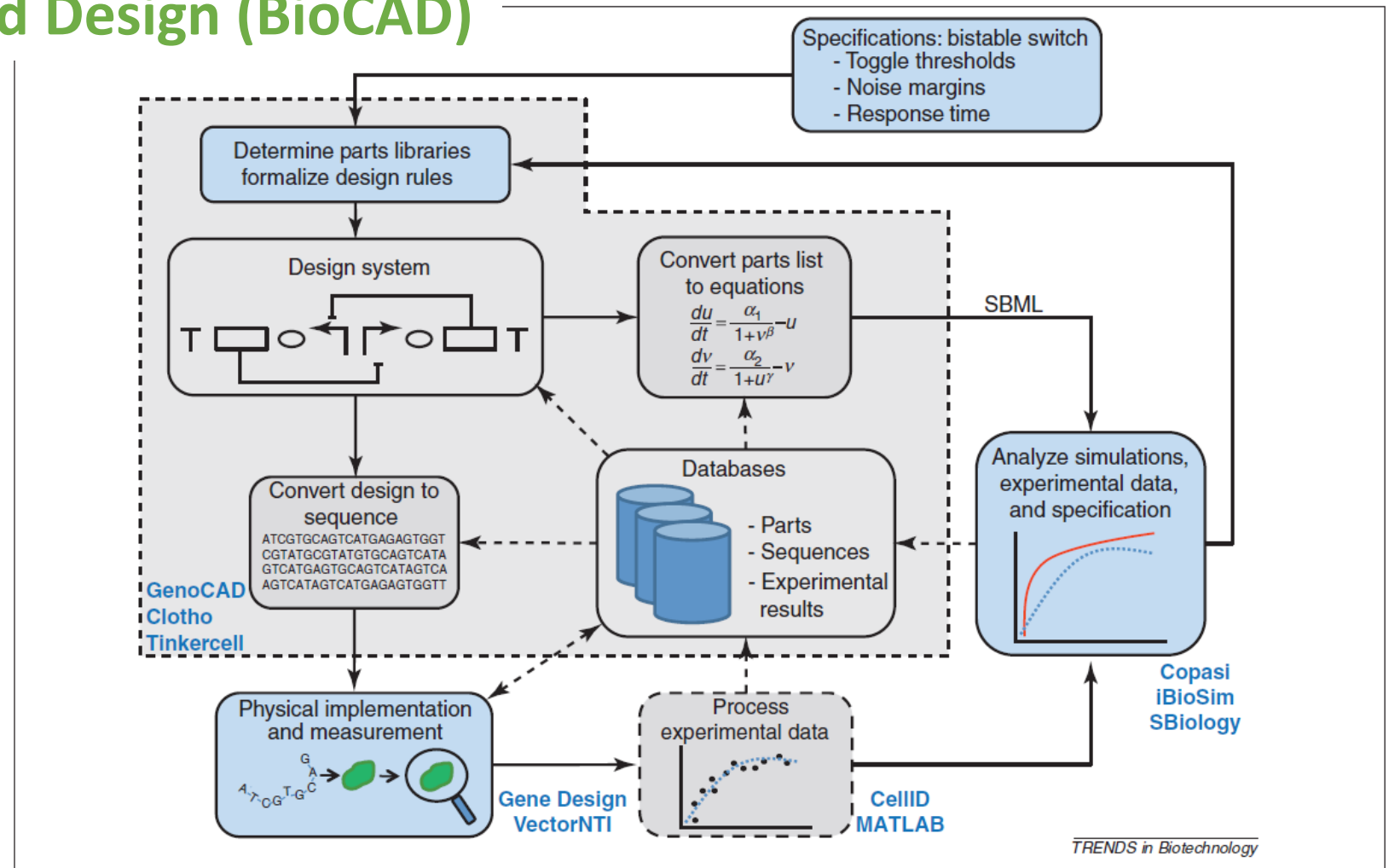


Figure 1. GDA design flow. Synthetic biology projects typically rely on iterative workflows composed of different tasks. Emerging GDA tool chains rely on numerous software applications that support different phases of the project workflow. The development of a genetic switch [72] will first involve expression of the design objective as a list of quantitative requirements: input toggle thresholds, noise margins, switching response time, etc. Once the objective is specified, it is possible to develop a list of genetic parts useable for the project. The choice of biological parts will involve factors such as use of the parts in prior projects, quality of the data characterizing the parts function, or intellectual property considerations. Formalization of design rules often takes place in parallel with parts library development. Design rules may cover issues such as whether it is acceptable to have polycistronic expression cassettes or if the design should be split between different plasmids. Only after parts have been selected and a strategy has been agreed on is it possible to start designing constructs. In the fabrication phase, the construct is assembled, usually by combining *de novo* gene synthesis and cloning of existing DNA sequences. Operators use molecular biology software suites to facilitate assembly or to order the sequence from a gene synthesis company. Experimentalists insert the synthetic DNA molecule into the host of choice and collect phenotypic data. Experimental data are then processed, for example by reducing microscopy images to time series of quantitative data. Performance is evaluated by considering simulations, experimental data and the original specifications. At nearly every stage, software interacts with databases to reuse past work or to store current work for future use. The shaded area delimited by dashes denotes stages facilitated by synthetic biology CAD software, whereas other stages are handled by more general purpose software. Text in blue indicates examples of software that provide assistance in each stage.

Genetic circuit design automation

Alec A. K. Nielsen,¹ Bryan S. Der,^{1,2} Jonghyeon Shin,¹ Prashant Vaidyanathan,²
 Vanya Paralanov,³ Elizabeth A. Strychalski,³ David Ross,³
 Douglas Densmore,² Christopher A. Voigt^{1*}

Cello design specification

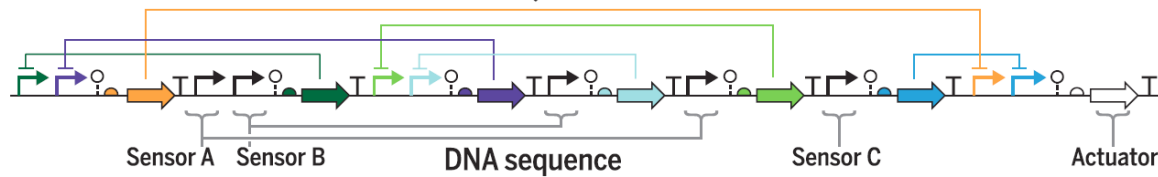
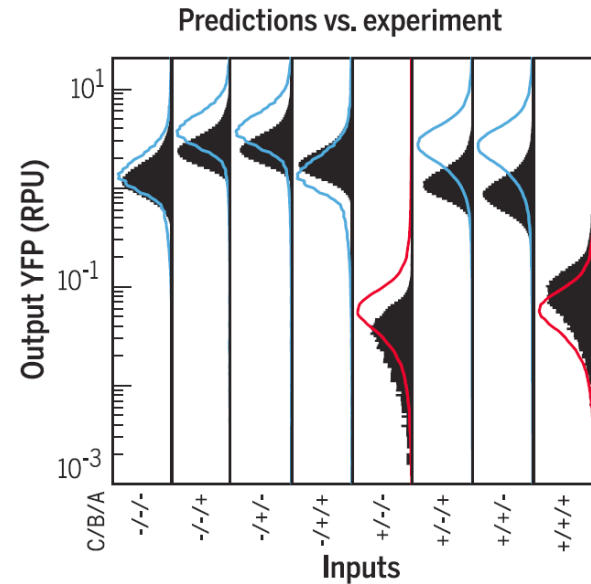
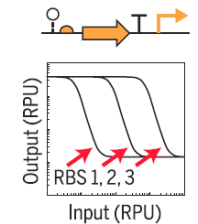
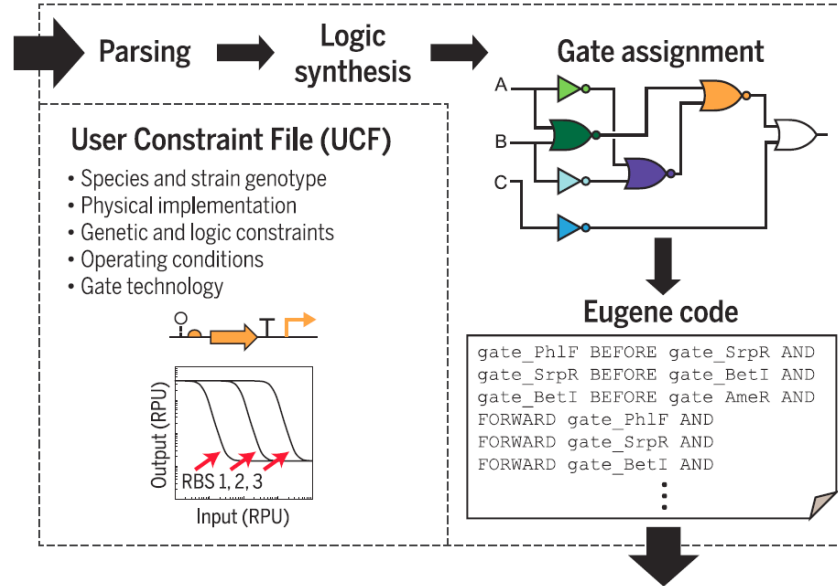
Sensors			
name	low	high	promoter sequence
A	0.003	2.8	AACGATCGTTGGCTGTGTTGACAATT
B	0.001	4.4	TACTCCACCGTTGGCTTTTTTCCTA
C	0.008	2.5	ACTTTTCATACTCCCGCCATTTCAGAG

```

Verilog
module 0xF6(output out, input A,B,C);
always@(C,B,A)
begin
case({C,B,A})
3'b000: {out} = 1'b1;
3'b001: {out} = 1'b1;
3'b010: {out} = 1'b1;
3'b011: {out} = 1'b1;
3'b100: {out} = 1'b0;
3'b101: {out} = 1'b1;
3'b110: {out} = 1'b1;
3'b111: {out} = 1'b0;
endcase
end
endmodule
    
```

Actuators	
name	sequence
YFP	ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGT

Run



Genetic programming using Cello. A user specifies the desired circuit function in Verilog code, and this is transformed into a DNA sequence. An example circuit is shown (0xF6); red and blue curves are predicted output states for populations of cells, and solid black distributions are experimental flow cytometry data. The outputs are shown for all combinations of sensor states; plus and minus signs indicate the presence or absence of input signal. RBS, ribosome binding site; RPU, relative promoter unit; YFP, yellow fluorescent protein.

