CrossMark

Scaling up genetic circuit design for cellular computing: advances and prospects

Yiyu Xiang[1,2] · Neil Dalchau[3] · Baojun Wang[1,2]

Abstract

Synthetic biology aims to engineer and redesign biological systems for useful real-world applications in biomanufacturing, biosensing and biotherapy following a typical design-build-test cycle. Inspired from computer science and electronics, synthetic gene circuits have been designed to exhibit control over the flow of information in biological systems. Two types are Boolean logic inspired TRUE or FALSE digital logic and graded analog computation. Key principles for gene circuit engineering include modularity, orthogonality, predictability and reliability. Initial circuits in the field were small and hampered by a lack of modular and orthogonal components, however in recent years the library of available parts has increased vastly. New tools for high throughput DNA assembly and characterization have been developed enabling rapid prototyping, systematic in situ characterization, as well as automated design and assembly of circuits. Recently implemented computing paradigms in circuit memory and distributed computing using cell consortia will also be discussed. Finally, we will examine existing challenges in building predictable large-scale circuits including modularity, context dependency and metabolic burden as well as tools and methods used to resolve them. These new trends and techniques have the potential to accelerate design of larger gene circuits and result in an increase in our basic understanding of circuit and host behaviour.

1 Introduction

In order to survive and reproduce, cells must sense a wide variety of inputs both external and internal. In response they compute and actuate a number of output functions such as changes to cell morphology (Goranov et al. 2013), or production of proteins and small molecules (Williams et al. 2016). We can exploit these changes for desirable functions such as the manufacture of valuable products or sensing of dangerous environmental toxins (Ro et al. 2006;

Wang et al. 2013a; Bereza-Malcolm et al. 2015; Machado et al. 2016; Bernard and Wang 2017). Many of these properties have useful real-world functions and it is desirable to manipulate them for our own goals.

Historically, the approach to the workflow in biotechnology has been based on bespoke, unique solutions. Often this results in laborious ad-hoc laboratory processes which result in an inability to port solutions from one problem to another, and a missed opportunity to retain valuable information learned through the process. There exists no one precise universally accepted definition of synthetic biology although most do overlap strongly. One definition states that synthetic biology is 'the design and engineering of biologically based parts, novel devices and systems as a well as the redesign of existing, natural biological systems,' (Clarke and Kitney 2016). Broadly speaking, synthetic biology is a rational approach to biotechnology inspired by ideas from engineering and aims to make designing biology easier, faster and more predictable. Key

✉ Baojun Wang
 baojun.wang@ed.ac.uk

1 School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FF, UK

2 Centre for Synthetic and Systems Biology, University of Edinburgh, Edinburgh EH9 3JR, UK

3 Microsoft Research, Cambridge CB1 2FB, UK

� Springer

concepts in this pursuit are standardization, modularity, characterization and orthogonality (Andrianantoandro et al. 2006). The field has attempted to create libraries and repositories of biological parts (Knight 2003), and incorporated the engineering design-build-test-learn (DBTL) workflow in various parts of the literature (Paddon and Keasling 2014; Hutchison et al. 2016; Clarke and Kitney 2016; Cox et al. 2018), industry (Anne Ravanona 2015; Siliconreview Team 2017), and government (Si and Zhao 2016). Ideally one could dependably generate entirely new systems with novel functions and predictable behaviour from a standardized parts list.

A major focus in the field has been on cellular computing, in which response pathways can be co-opted to produce useful biological computing devices which can produce programmable and predictable outputs in response to diverse input signals. One design paradigm has been to use gene circuits to regulate behaviour, of which two methods are digital binary-like and analog models. Other models such as DNA computing through tiling, hybridisation, self-assembly or recombination exist, but this review will focus on genetic circuits. A review covering other aspects of computing with biological parts was done by Moe-behrens (2013).

Digital-like biological parts often resemble logic gates and their discrete binary states found in silicon transistors such as the 1-bit full adder; comprising 5 logic gates wired in 3 layers with 3 inputs and 2 outputs (Fig. 1). Here there is a high contrast between high levels and low levels of output, corresponding to a discrete binary ON or OFF state respectively. As seen in Fig. 2a this means a high and sharp change

in the output signal over a small change in the input signal once it hits a threshold level (represented by a hill coefficient > 1). Often these circuits will be described by Boolean logic, in which all values are reduced to either TRUE (1) or FALSE (0) (Bradley et al. 2016a). Logic operations like AND (where output is TRUE only if both inputs are TRUE), can be represented by a Boolean truth table as well as circuit symbols adapted from electronics as shown in Fig. 2a. Ideally, both the input and the output must be able to be connected to and interact with upstream and downstream components and operate in the intended fashion (be modular), the signal output must be stable, exhibit low noise (random unintended fluctuations), and have a large ON:OFF ratio, or dynamic range (Bradley and Wang 2015). This prevents the signal from being degraded as it propagates through a system. Digital logic is particularly useful in a decision-making circuit such as in natural cell differentiation or apoptosis. The strong state change is ideal for reliable state transitions and signal integration as digital circuits are relatively robust to noise. Early instances included the gene toggle switch (Gardner et al. 2000), the repressilator (Elowitz and Leibler 2000), and the autoregulator (Becskei and Serrano 2000). From there on, binary-like logic gates such as AND (Anderson et al. 2007; Wang et al. 2011), OR (Mayo et al. 2006), and NOR (Guet 2002; Tamsir et al. 2011) were built and combined into more complex circuits, for instance a 4 input AND gate (Moon et al. 2012).

In contrast, analog responses are designed to give a continuous output changing dynamically according to the input. Good transfer functions of analog computation are clear and
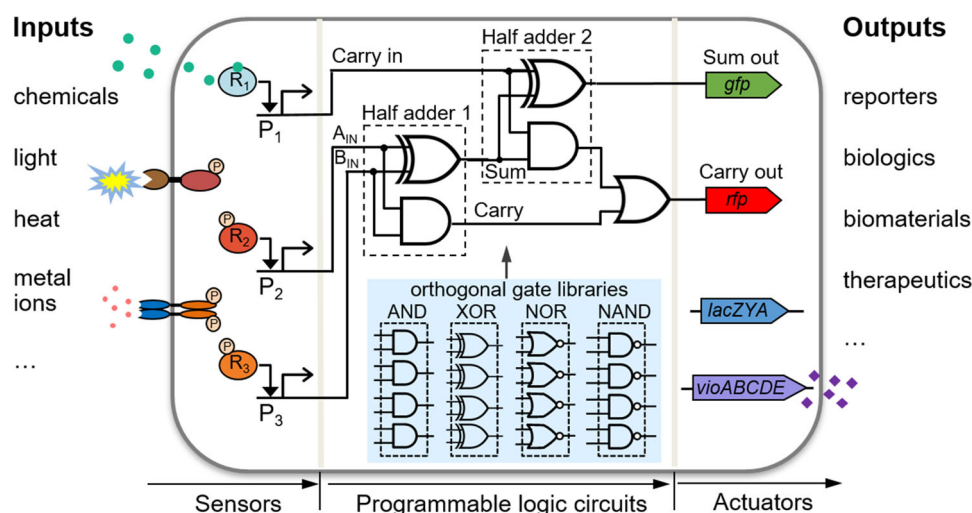

Fig. 1 Programmable cellular computation with scalable signal processing capacity. To achieve large-scale control of cellular behaviour, an expanded library of versatile orthogonal genetic regulatory blocks and associated wiring principles are needed. For example, a genetic 1-bit full adder program adds 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
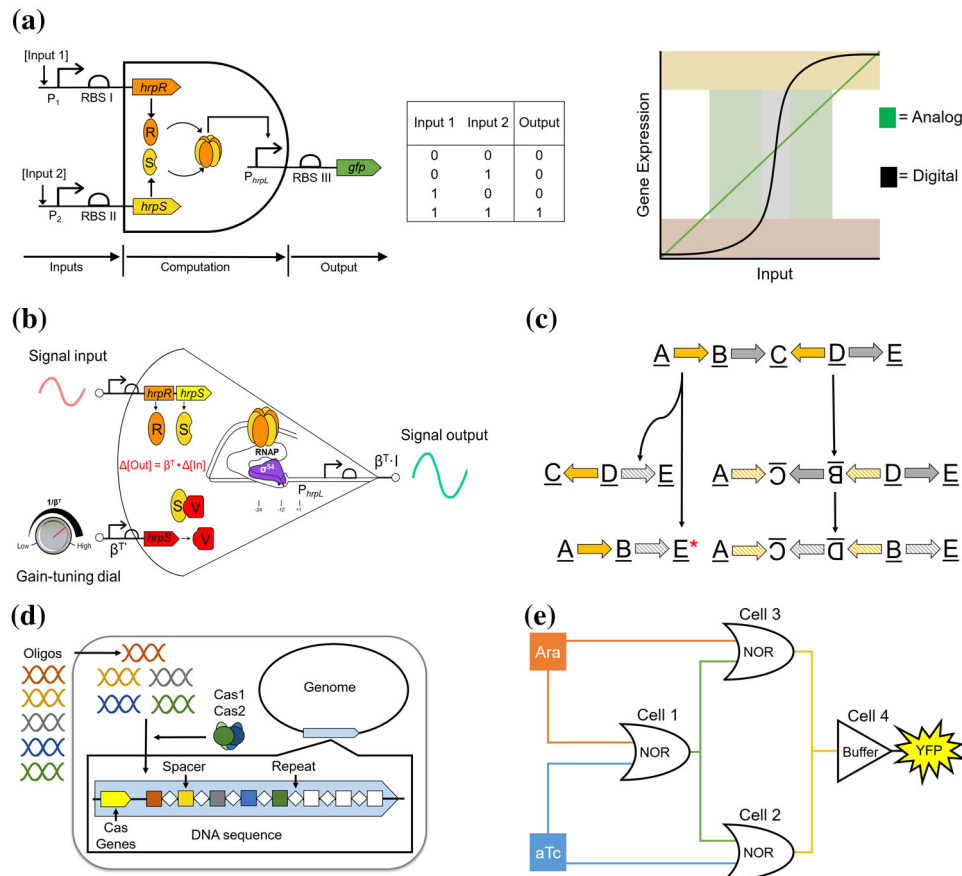
Fig. 2 Versatile cellular computing paradigms enabled by synthetic biology. **a** A two-input AND gate using the σ^{54}-dependent HrpR/HrpS hetero-regulation module and the corresponding truth table. The HrpS and HrpR enhancer-binding proteins expressed from separate inducible input promoters bind to form a heteromeric complex which activates the output σ^{54}-dependent *hrpL* promoter. Also shown is a graph with representative transfer functions of digital (grey) and analog (green) signal responses. Digital signals have a steep sigmoidal response with a large change in expression over a small change in input. Analog signals have a much more graded response. **b** An analog transcriptional signal amplifier designed based on the *hrp* gene regulation module from plant pathogen *P. syringae*. The amplifier achieves different gains and input dynamic ranges by varying the expression levels of the underlying ligand-free ultrasensitive activator proteins (HrpRS) in the device. HrpV binds and sequesters HrpS so that it can no longer bind its co-activator HrpR and is used to modulate the intensity of the amplified signal output. **c** A recombinase-based state machine. Depending on if the att sites face in the same direction or towards each other, recombinases can excise or invert pieces of DNA respectively. By doing so they can record events, and with the correct elements can modulate gene expression. Striped arrows represent post inversion, the new site is sequentially different from the old one. **d** CRISPR-based memory storage. Oligos are sequentially incorporated into CRISPR arrays in the genome by the cas1–cas2 complex. Each oligo encodes information in the DNA sequence. The sequence at which oligos are ordered can be controlled by spacing addition over time. **e** Spatially distributed gates can reuse parts and signalling molecules, using proximity as a way to insulate signal propagation. Here 3 NOR gates effectively replicate a XOR gate using the inducers arabinose (Ara) and anhydrotetracycline (aTc). Signal lines green and yellow represent the quorum sensing molecules *N*-3-oxo-dodecanoyl-homoserine lactone (3OC12-HSL) and *N*-butyryl-homoserine lactone (C4-HSL) respectively and the output is yellow fluorescent protein (YFP). (Color figure online)

well defined, responding to a large range of input, as well as exhibiting low noise and being reliable and modular, again as shown in Fig. 2a. Analog computation is particularly resource efficient, good at generating small autonomous responses to temporal differences, autoregulation and interaction with host metabolism, although analog circuits are susceptible to disruption from noise and temporary perturbations in input (Sarpeshkar 2014). Examples of analog circuits include amplifiers for scaling transcriptional signals in cascaded gene networks (Wang et al. 2014), comparators

that convert signals from analog to digital (Rubens et al. 2016) and networks that allow for reconfigurable inversion of the network transfer function (Lewis et al. 2018).

Potential applications for gene circuits have been hinted at with the proof-of-concept arsenic biosensors (Wang et al. 2013a), and the cancer killing 3-input NOR gate "classifier circuit" (Xie et al. 2011). Cell consortia have been used to build an analog to digital converter detecting smell using olfactory receptors (Müller et al. 2017) whilst memory engineered commensal *E. coli* was able to

function for six months in a mouse gut, sensing and reporting on the presence of inflammation-indicating tetrathionate (Riglar et al. 2017).

Many of the early examples were small in scale and design, often with few synthetic parts. Gander et al. (2017) reviewed the literature and found some of the largest circuits in recent publications only had 7 parts and 6 connections. This is roughly in line with Privman et al. (2008) who surmised that under optimal conditions going over 10 processing steps would be difficult using methods available at the time and would require new noise reduction paradigms. Recently larger circuit examples have been emerging, advances in characterization techniques and the expansion of the library of functional parts has promised a much larger space in which to build more complex circuits. RNA based parts enabled construction of a substantial 12-input disjunctive normal form circuit (Green et al. 2017), whilst a 6-input 1-output Boolean logic look-up table circuit was also demonstrated in mammalian cells. The group created a circuit which receives 4 selection inputs and 2 data inputs, the selection inputs determine which one of sixteen 2 data input logic functions the circuit uses allowing switching of logic in cells on the fly. By normalizing the promoter and recombinases used, they were able to generate 113 circuits, of which 109 were working, the largest collection of functionally unique logic circuits in mammalian cells as of publication (Weinberg et al. 2017). Another large circuit, the 1-bit full adder was functionally constructed in mammalian consortia incorporating 22 separate gates distributed amongst 9 specialized cell types in a complex three-dimensional environment (Ausländer et al. 2017). However construction of such large scale genetic circuits are uncommon, large numbers of logic gates in single cells are scarce and require significant amounts of time and effort to work through an iteration of the design-build-test-learn cycle.

In this article, we will discuss the tools and challenges surrounding the construction of large-scale gene circuits. We will decompose this into the DBTL cycle for clarity: design—the arrangement of reusable components to produce biological programs, build—large scale DNA assembly, test—high throughput characterization and debugging tools, learn—modelling and circuit design automation. We will review common methods enabling control of genetic circuitry and DNA assembly. The latest in advanced characterization and debugging methods using cell free systems, microfluidics, and ribonucleic acid (RNA) sequencing (RNA-seq) will be discussed. We will give insights into automated gene circuit design software and examine the implementation of more complex computing paradigms such as distributed computing and memory integration. Although most of the work has been directed towards genetic circuits in single cells, there has been a significant body of work that has experimented with using cell consortia, separating out circuits into many hosts (Regot et al. 2011; Macía et al. 2012). Integration of memory enables a move away from just combinatorial logic (in which the output is a function of the present inputs), making sequential logic possible (Siuti et al. 2013; Purcell and Lu 2014; Roquet et al. 2016). The challenges in scaling up circuit design and the techniques used to tackle them will be discussed. Focus will be on the obstacles to modularity as well as context effects and metabolic burden.

2 Design: expanded toolbox for engineering complex gene regulation programs

The construction of any large-scale circuit requires a large library of well characterized, orthogonal and modular gates comprising the 'building blocks' of the system. Since the beginning of the field there has been a significant and promising expansion of the molecular toolbox. A large variety of repurposed biochemical tools have been demonstrated to admit some degree of control over cellular state. Many of these tools co-opt biology's central dogma, the expression of a gene and the information flow from DNA to RNA to proteins. These can be broadly grouped into three types; control of DNA transcription, messenger RNA (mRNA) translation or protein–protein interactions. Many of these methods are shown in Fig. 3.

Inducible parts provide useful externally mediated control of systems whilst customizable DNA sequence binding enables a significant expansion of the number of orthogonal parts available for large circuits (Garg et al. 2012; Lohmueller et al. 2012; Qi et al. 2013; Kiani et al. 2014; Nielsen and Voigt 2014; Li et al. 2015).

Control of expression using RNA tends to be less burdensome on host metabolism and can be governed by cleavage (Qi et al. 2012), pair binding (Rinaudo et al. 2007; Anderson et al. 2007; Xie et al. 2011; Wroblewska et al. 2015) or secondary structure (Sharma et al. 2008; Auslander and Fussenegger 2014; Myhrvold and Silver 2015; Chappell et al. 2015; Karagiannis et al. 2016), as folding has been proved to be moderately predictable using software such as NUPACK (Zadeh et al. 2011; Wang and Buck 2012). Protein–protein interactions such as those in the *hrp* (hypersensitive response and pathogenicity) gene regulation system can be utilized to generate versatile multi-input genetic logic gates (Wang et al. 2011; Wang and Buck 2014). Artificially or naturally split inteins can integrate signals (Schaerli et al. 2014) and different proteins can even be fused to each other in order to have hybrid properties (Wang et al. 2013b).

The expansion of parts has been enabled by a variety of tools; genomic part mining was successful in building a

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)
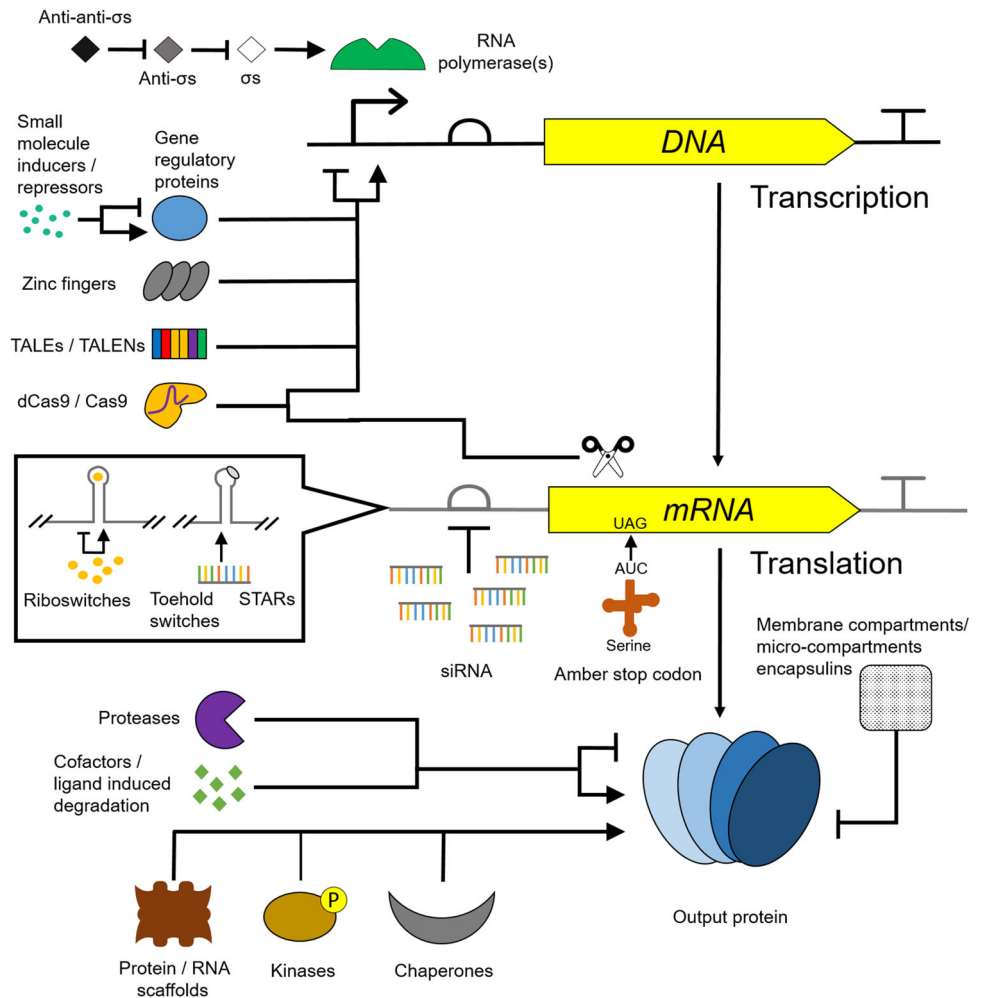
library of 16 orthogonal strongly repressing TetR family repressors (Stanton et al. 2013), 4 different T7 polymerases (Temme et al. 2012) and 20 different sigma factors (Rhodius et al. 2013). Chen et al. (2013) characterized 582 terminators whilst a protein engineering approach using bioinformatics and site directed mutagenesis generated different DNA binding specificities in the cAMP receptor protein (CRP) family (Desai et al. 2009). However, it is likely that the predictable natures of customizable DNA binding proteins and RNA secondary structure-based tools will provide the largest number of orthogonal parts. Didovyk et al. (2016) generated 633 possible orthogonal (to host and circuit) guide RNA's for CRISPR/Cas9 mediated gene regulation through computational screening whilst 180 TALE effectors were designed by Garg et al. (2012). Other studies developed 26 toehold switches (Green et al. 2014), 4 RNA riboregulator/genetic switchboard pairs (Callura et al. 2012), 6 RNA-IN-RNA-OUT families (Mutalik et al. 2012), 5 zinc fingers (Khalil et al. 2012), and 20 sigma factors (Rhodius et al. 2013). From here, one can determine that there is a significant number

of parts, far more than is needed than for most current circuits and yet circuit size has not increased in proportion to the size of the library.

Finally, there is a variety of tuning 'knobs' which although not able to provide control themselves, can be used to adjust the dynamic properties of a system to respond effectively to expected inputs and produce desired outputs depending on the need. These include changing the strength of the promoter sequence, hybrid combinations of promoter sequences (Chen et al. 2018), operator site modification (Ang et al. 2013), ribosome binding site (RBS) modification (Salis et al. 2009), altering plasmid copy number (Guido et al. 2006), using decoy DNA operators (Lee and Maheshri 2012), RNA interference (RNAi), degradation tags (Bonger et al. 2011), or co-expression with sequestering proteins or molecules (Wang et al. 2014). Cooperativity has been improved using oligomerization domains (Hou et al. 2018). Positive feedback loops and signal cascades have improved the ON/OFF ratio in digital-like circuits which is often poor due to an inherent basal level of 'leaky' gene expression even

without the presence of activators or in the presence of repressors (Bradley et al. 2016a). A informative review of available control and tuning methods was covered by Ang et al. (2013), Bradley and Wang (2015) and Bradley et al. (2016b).

3 Build: standard large-scale DNA assembly

Although genes can be chemically synthesized, it is still an expensive solution despite the recent drop in cost and importantly does not incorporate any standardization. Ideally all parts would be characterized, stored in a library and then manipulated at will using a common scalable DNA assembly protocol. Initial attempts were based on lengthy stepwise restriction enzyme mechanisms such as in Bio-Bricks^{TM} (Shetty et al. 2008). The latest methods are termed 'one pot' as multiple fragments can be assembled at once in a defined order saving many man-hours in the laboratory. These include Golden Gate and its derivations; MoClo (modular cloning) and Goldenbraid (Engler et al. 2008; Sarrion-Perdigones et al. 2011; Weber et al. 2011). Gibson assembly (2009), the most popular non-synthesis based method (Kahl and Endy 2013), uses homology of overlapping single stranded DNA which also avoids the necessity of removing forbidden sequences (such as restriction enzyme sites) in the sequences being assembled. Unfortunately, the superseding of BioBricks with formats such as Gibson assembly has moved parts back towards non-modular tailored solutions. Assembled parts cannot be ported to another assembly without new specific primer, and each new PCR reaction itself has bespoke conditions affecting assembly success due to the lack of standardization in flanking sequences. This can also have ramifications in terms of genetic context (discussed later on), and libraries of parts become more complex to reuse. Efforts have been made to add modular prefixes and suffixes; Casini et al. (2014) developed a strategy named MODAL (Modular Overlap-Directed Assembly with Linkers) and similar methods were published elsewhere (Torella et al. 2014). Woodruff et al. (2017) used unique flanking sequences acting as a barcode in a pool of collected constructs that can be retrieved using PCR and subsequently assembled using Golden gate assembly. Biopart Assembly Standard for Idempotent Cloning (BASIC) assembly exploits orthogonal linkers to avoid PCR entirely and achieved over 90% accuracy with a 7-part reaction. Devised a hybrid method, Golden Gate-Gibson (3G) combines overhang assembly with Golden Gate style part libraries (Halleran et al. 2018). Despite the relatively rapid speed of modern assembly methods, construction of large libraries of clones can still take a significant amount of time. It is here that automated construction methods using robotics or microfluidics (discussed later), would greatly aid in speeding up the process as they can potentially run many assemblies at once 24 h a day with minimal human supervision.

4 Test: high-throughput circuit characterization

After the 'build' part of the cycle, the next part is to 'test'; quantifying the characteristics and dynamics of the circuit to inform on future designs and find solutions to any failures that have been encountered (debugging). Scalability in characterization is largely determined by the feasibility of running many concurrent experiments at once, measuring many different properties and gathering precise data from the samples reducing the time needed to complete the DBTL cycle.

Traditionally, this has been dominated by fluorescent gene reporters such as green fluorescent protein (GFP) and red fluorescent protein (RFP). These proteins are used to measure gene expression and quantified using a plate reader or a fluorescent microscope. Beyond measurement of the regulatory sequences of gene expression, they can be fused to other proteins to study protein localization and interaction through Förster resonance energy transfer (FRET) (Selvin 2000). The advent of flow cytometry has enabled the simultaneous measurement of multiple cellular properties in every cell, such as size, granularity and fluorescence analysis through multiple lasers fed by a small current of cells suspended in fluid. Flow cytometers can even separate cells by fluorescence levels known as fluorescent activated cell sorting (FACS) (Tracy et al. 2010). This analysis of single cell data offers a much more precise view of cell state and reveals a much deeper relationship between host-circuit physiology, such as the relationship between fluorescence and cell volume, something that is much more difficult in population level measurements.

Measurement of the fluorescence of a protein encoded downstream of a regulatory sequence is equivalent of measuring the final product of gene expression. This expression measurement combines both transcription and translation together and doesn't differentiate between the two. Pothoulakis et al. (2014) developed the 'spinach' RNA aptamer which fluoresces in the presence of the fluorophore, 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), and is consequently a good option to measure transcription separately. Another limitation is the incorporation of undesirable genetic context (discussed in more detail later on) such as variable untranslated regions (UTR) of DNA that can form secondary structures disrupting translation though themselves not being translated, this can be addressed through the use of insulators that

reduce context effects such ribozymes that cleave the 5′ UTR (Lou et al. 2012). The expression of these protein can also cause variable metabolic load (Bentley et al. 1990), making it hard to measure the expression of many genes in parallel and minor variations in experimental conditions can cause large changes in expression (Kelly et al. 2009; Rudge et al. 2016). To solve this issue, Kelly et al. (2009) normalized promoter activity using a reference promoter resulting in the relative promoter unit (RPU), and Rudge et al. (2016) compared the output of reporter genes concurrently with a control plasmid to find the intrinsic promoter activity, reducing the variance due to extrinsic factors to less than 4%.

High throughput experiments allow us to repeat many experiments in parallel that gather a greater quantity of data in a shorter space of time whilst genome and proteome wide techniques offer a wider view of cell state that singular gene expression experiments cannot practically offer. For example, using RNA-seq we can get a non-invasive snapshot measurement of the RNA levels of single cells and populations, enabling complete analysis of changes when implementing genetic circuits. Based on next generation sequencing, mRNA is cut into small sections and turned into complementary DNA (cDNA) through reverse transcription before sequencing and alignment. RNA-seq has been used to analyze the transcriptome of cells and has been demonstrated to work in situ in mammalian tissues (Lee et al. 2014). Liao et al. (2017) used RNA-seq to determine changes in the host cell transcriptome when an AND-gate circuit is designed under different circuit compositions and in different plasmid copy numbers; finding that higher copy number decreased the orthogonality between the circuit and host gene expression in addition to increasing metabolic load and causing imbalance among the circuit components. Gorochowski et al. (2017) used RNA-seq to measure simultaneously part performance and the state of a three-input one-output circuit comparing 46 parts. They were able to debug failures in the circuit due to antisense promoters, terminator malfunction and media related failure and make informed design decisions, such as including a bi-directional terminator to cease antisense transcription. Limitations in terms of cost and library preparation time are being addressed with simplifying techniques such as RNAtag-seq whereby DNA barcodes are uniquely tagged to allow early pooling of samples before the preparation of the library (Shishkin et al. 2015). Other methods include single molecule RNA fluorescence in situ hybridisation (smFISH), used by Nielsen et al. (2016) to quantify mRNA levels of yellow fluorescent protein (YFP). 25 oligonucleotide probes, each 20 bases in length were fluorescently labelled with TAMRA (carboxytetramethylrhodamine), and binding of multiple probes enables sufficient fluorescence to detect and localize

target mRNA. Similarly whole cell mass spectrometry can be used to attempt to identify changes in the proteome (Ouedraogo et al. 2013).

Prototyping gives meaningful biological information towards the design of the final system yet can be assembled and tested much more quickly. Cell free in vitro systems have all the machinery necessary for basic protein expression but do not require long culturing times, often also omitting the complexity of full host metabolism and thereby being significantly easier to model. They also offer options to monitor dynamics of the system in real time with fluorescent RNA aptamers (Niederholtmeyer et al. 2013) and FRET probes (Norred et al. 2015), also facilitating direct reaction sampling into a HPLC or MS machine (Heinemann et al. 2017). Drawbacks include being much more lacking in some shared cellular resources (Gyorgy and Murray 2016), as well as having potential energy consumption imbalances (Garamella et al. 2016). One example is the *E. coli* transcription-translation based cell free system (TX-TL) which has successfully been used for prototyping promoters (Iyer et al. 2013) and negative feedback loops (Karig et al. 2012). Further advances enabled use of linear DNA through protection from degradation by RecBCD through the addition of GamS protein (Sun et al. 2014). Using this method a 4-piece genetic switch was assembled within 8 h (1 working day), using simple Golden Gate assembly and polymerase chain reaction (PCR) to create 4 linear sequences directly used for testing, although in this case there was a lack of correlation between in vivo and in vitro results (Sun et al. 2014). Another example built on the concept of using whole cell extracts by using microbial consortia to obtain purified translational machinery (Villarreal et al. 2017). Pardee et al. (2014) demonstrated freeze dried paper based cell free systems containing gene networks that can be rehydrated when needed.

A different way to increase scale is to minimize the resource consumption of each experiment and to automate physical tasks. The field of microfluidics deals with the precise manipulation of small amounts of fluids in the micro and nanoliter scale. Discrete volumes of liquid can be packaged into droplets and controlled automatically either as a solution or individually, each droplet functioning as an independent reaction mix with the small scale enabling conserved use of reagents and biological material. Typically these methods are either continuous; using oil and water to generate a controllable liquid stream, or digital; using voltage to control the movement of individual droplets on a conductive material (Huang and Densmore 2014). Shih et al. (2015) demonstrated that it was possible to concurrently run several assemblies at once (Golden Gate, Gibson and yeast), using a hybrid of both microfluidic technologies, assembling a library of 16 plasmids and

performing on chip electroporation into bacteria and yeast. Other systems demonstrated heat shock transformation and the capability to culture the cells on chip (Gach et al. 2016). Analogous to FACS, Baret et al. (2009) sorted cells by fluorescence through fluorescent activated droplet sorting (FADS). Procedures have been developed that can trap single cells in droplets, and provide them with all the nutrients needed to be incubated for days, permitting longer study and performance that cannot be obtained through FACS (Bennett and Hasty 2009).

Uniting cell free and microfluidic technologies can combine the benefits of both, for example, generating many artificial cell-free entities, prototyping in parallel. By example, Schwarz-Schilling et al. (2016) produced functional AND gates and sender circuits in droplets containing cell free systems and bacteria. Fan et al. (2017) use droplet microfluidics to print accurate and small quantities of cell free systems to measure interactions between three genetic factors at a synthetic promoter and used this data generate a model. Wang et al. (2018) used a similar method, only combining a locked nucleic acid probe (measuring mRNA levels) with fluorescent proteins enabling simultaneous measurement of transcription and translation in massively parallel cell free droplet experiments.

5 Learn: biological circuit design automation and modelling

The synthetic biology community has developed software tools that aim to replicate the success of computer aided design (CAD) used in electronic circuit engineering. Design automation has the potential to accelerate biological design by allowing designers to access existing knowledge of biological parts, arrange parts into circuits, design experiments, store and visualise experimental data, and potentially make predictions about circuit behaviours. For circuit construction (build), it could plan out assembly of the physical DNA sequence from the given starting material and include the experimental protocol needed to do so. For testing, software might enable experiments to be designed and simulate a computational model of the system, allowing costly and time-consuming experiments to be replaced, but still give insight into how a system might behave and identify which experiments are critical or contain the most information for guiding design decisions.

Several software tools have emerged over the last 10 years that seek to deliver some of these features, including Genocad (Czar et al. 2009), CellDesigner (Funahashi et al. 2008), Biojade (Goler et al. 2008), SynbioSS (Hill et al. 2008), Tinkercell (Chandran et al. 2009), Visual GEC (Pedersen and Phillips 2009) and Cello (Nielsen et al. 2016), although many of the former projects are dormant.

Some software forgoes biological part data and only operates at the abstracted design level (Bhatia et al. 2017), whilst others are primarily concerned with data storage standards, such as DICOM-SB (Sainz De Murieta et al. 2016), many instances of software use the Synthetic Biology Open Language (SBOL), an open standard for the representation of genetic parts, with common formats for both data and visual symbols (Roehner et al. 2016). Most also offer the export of the models via the Systems Biology Markup Language (SBML) (Finney and Hucka 2003; Hucka et al. 2003), which enable model analysis in more general software platforms, such as Matlab and Copasi.

Cello, one of the latest iterations of gene circuit design software enables circuits to be constructed that compute specified logic functions. The Verilog logic programming language is used to describe circuit function, and a user constraints file to specify parts and organisms to create a searchable design space. Circuits are modified to be compatible with a library containing NOR and NOT gates based on repression. The system is simulated to predict circuit performance, factoring in load, population variability, growth, and connectivity in terms of RNA polymerase (RNAP) flux. Finally, the physical component is designed for assembly, i.e. the circuit contained in one plasmid and the reporter on another with appropriate promoters, terminators and other gene regulatory elements on each (Nielsen et al. 2016).

Figure 4 shows a representative gene circuit design automation flow, inspired by electronic circuit design automation, for an exemplar 3-input and 2-output 1-bit full subtractor that comprises 7 logic gates wired in 5 layers. Initially there would be an abstract level of input of the circuit function, such as using a graphical user interface (GUI,) a truth table, or Boolean algebra, before an optimization of the overall design to reduce the number of parts to a minimum whilst retaining the desired function. Subsequently, logic synthesis would be performed to transform the expression into desired gate level format before technology mapping all the possible circuits from standard well characterized libraries of logic gates and parts, such as the open iGEM parts registry (Mitchell et al. 2011) or SynBioLGDB (Wang et al. 2015b). Potential systems would be modelled for functionality and ranked accordingly, and genetic assembly constructs would be designed for assembly.

One of the major challenges for design automation software is the difficulty in producing mathematical models that are predictive of circuit function. This is challenging for several reasons. The first major challenge is that it is difficult to know how to even write down a mathematical model that captures the nonlinear features of the biochemical interactions of a given circuit. To model a system mechanistically (as opposed to purely statistical models),
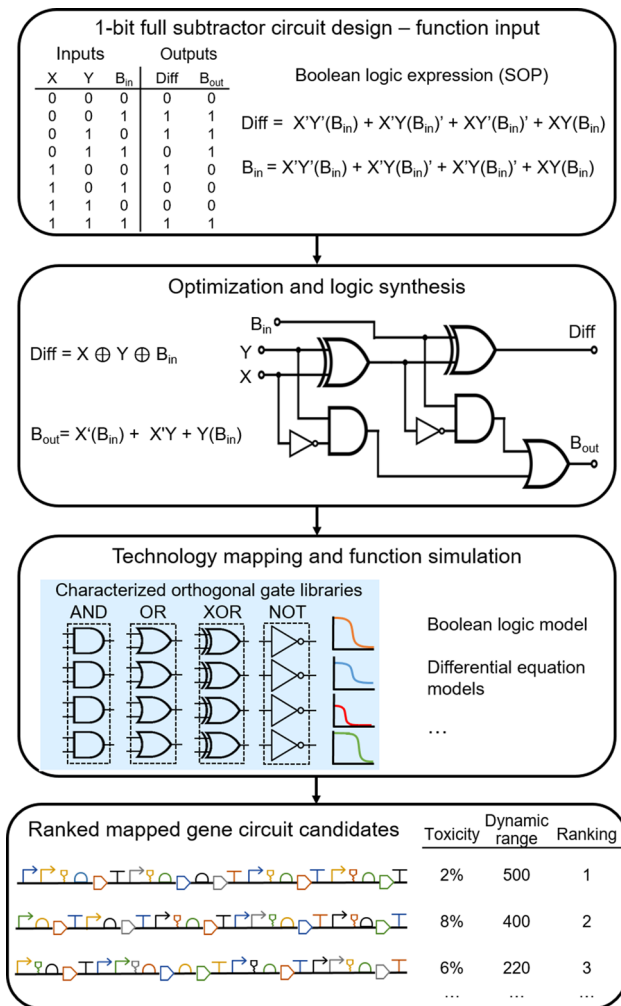
1-bit full subtractor circuit design – function input

Inputs			Outputs	
X	Y	B_{in}	Diff	B_{out}
0	0	0	0	0
0	0	1	1	1
0	1	0	1	1
0	1	1	0	1
1	0	0	1	0
1	0	1	0	0
1	1	0	0	0
1	1	1	1	1

Boolean logic expression (SOP)

$Diff = X'Y'(B_{in}) + X'Y(B_{in})' + XY'(B_{in})' + XY(B_{in})$

$B_{in} = X'Y'(B_{in}) + X'Y(B_{in})' + X'Y'(B_{in})' + XY(B_{in})$

Optimization and logic synthesis

$Diff = X \oplus Y \oplus B_{in}$

$B_{out} = X'(B_{in}) + X'Y + Y(B_{in})$

Technology mapping and function simulation

Characterized orthogonal gate libraries
AND   OR   XOR   NOT

Boolean logic model

Differential equation models

…

Ranked mapped gene circuit candidates

	Toxicity	Dynamic range	Ranking
	2%	500	1
	8%	400	2
	6%	220	3
	…	…	…

Fig. 4 Towards large-scale genetic circuit design automation. Representative design flow of an exemplar digital logic gene circuit is shown. Design input in an abstracted format, in this case a truth table and a Boolean logic expression (SOP = sum of products). The system then performs an optimization in Boolean algebra for the 1-bit full subtractor. Logic synthesis is performed to design the most efficient circuit using a set of preferred types of logic gates; parts and devices are assigned from a characterized gate library (mapping) and modelled to assess feasibility of implementation. The genetic sequences are deconstructed and reconstituted into an optimal sequence design to be used in the target host organism. These are then ranked using important factors when implementing in vivo. Examples used in this case are toxicity effects on the host in terms of percentage growth reduction and dynamic range of the output for the circuit, measured in fold change between the ON and the OFF states

can be determined to describe the system. Beyond intrinsic circuit interactions there are also interactions between circuit components and host resources (which we return to later in this review).

Another challenge relates to the fact that chemical reactions are discrete and stochastic, although modelling these variations can be computationally burdensome. In some cases, deterministic approximations can be sufficient depending on the scope of the answer desired. Simpler still, Boolean approximations can be very effective for describing gene networks where the gene is notionally simply on or off, and when intermediate expression is not functionally relevant. Models could be based on, amongst others, ordinary differential equations (ODEs), partial differential equations (PDEs), stochastic differential equations (SDEs), reaction–diffusion equations, and either stochastic sampling or integration of the chemical master equations (CME) (Chandran et al. 2008; Marchisio and Stelling 2009; MacDonald et al. 2011). How to assess which model structures are most appropriate (multimodel selection) is a problem known to be extremely challenging, even when it is possible to perturb and directly measure many of the circuit components (Marbach et al. 2010; Hayden et al. 2016). Nevertheless, approximate Bayesian methods are gaining traction for deciding between a set of similar models (Toni et al. 2009).

A related (simpler) problem to model selection that is still of practical utility is the problem of inferring model parameters (of a fixed model structure), given observation data (Toni et al. 2009; Golightly and Wilkinson 2011). After all, to make accurate predictions about circuit behaviour, the parameter values must surely be known to some level of accuracy. However, the majority of biological components described thus far in this review have not been characterized in sufficient detail to enable predictive modelling to be commonplace in synthetic biology. With recent advances in machine learning, there are now methods emerging that can handle high-dimensional parameter inference problems, though require the use of large-scale computing resources. Nevertheless, such methods enable us to determine the relationship between the sequence of DNA parts (Alipanahi et al. 2015; Kreimer et al. 2017), and their quantitative behaviour in the cell, including the RBS calculator described above (Salis et al. 2009), but also more generally how protein expression depends on the whole $5'$ untranslated region (Cuperus et al. 2017). But all circuits also incorporate other modes of biochemical regulation (e.g. protein turnover, ligand binding, translocation), which means that methods for inferring parameters for models of specific circuits is just as important. Characterizing sets of related circuits simultaneously is beginning to enable model parameterizations from no prior quantitative information directly from measurements, enabling design

the main elements of the system should be known and their interactions with the genetic and protein elements of the cell specified directly (the intricacy of which depends on the scope of the model). This requires either prior knowledge or the mining of interaction databases, though these are far from exhaustive at present. There is also the consideration of chemical diffusion and spatial arrangement (Endler et al. 2009). Subsequently mathematical formulae

optimisations to be generated (Huynh and Tagkopoulos 2014, 2016; Grant et al. 2016).

6 Towards advanced paradigms in cellular computing

6.1 Memory and data storage

Memory in cells relies on permanent cellular changes in response to temporary inputs, normally genetic, this enables sequential logic over common combinatorial logic. One of the earliest reported devices was the bistable toggle switch, containing two stable gene expression states (Gardner et al. 2000). Subsequent work used recombinases, enzymes that can catalyze DNA excision or flipping depending on the direction of the corresponding attachment (att) sites that flank the DNA sequence of interest (Bonnet et al. 2012). Some recombinases are bidirectional, either inherently or alongside a co-expressed Recombinase Directionality Factor (RDF). However, often permanent reactions are preferred thanks to their inherent stability. Reading the output can be as easy as basic sequencing. Alternatively, there is a possibility of enclosing parts such as promoters and gene coding sequences which can be flipped in and out of the correct gene coding orientation enabling circuit integration for sequential logic. Siuti et al. (2013) were able to create 16 two-input Boolean functions using recombinases that surrounded genetic elements such as promoters and terminators, demonstrating memory stable for 90 cell generations and Yang et al. (2014) reached a recording capacity of 1 byte.

State machines can be one of a number of finite states at any given time, with access to states dependent on predetermined sequence of events triggered by various conditions. A basic 3-state version is shown in Fig. 2c, and the most complex reported has 16 different positions (Roquet et al. 2016). A version of this machine was used to record temporal events in a population, with the distribution of final cell states and spatial location recording the dynamics of any inducer response including pulses (Hsiao et al. 2016). Another example created analog-like memory by generating graded expression of single stranded DNA in response to various signals, co-expressed with a corresponding recombinase targeting specific genomic sites resulting in a mutational response in the genome proportional to the duration and magnitude of the input (Farzadfard and Lu 2014). As mentioned in the introduction, Weinberg et al. (2017) used recombinases to build over 100 different types of multi-input multi-output circuits that include a 1-bit full adder, by taking advantage of recombinases they were able to combine computation into a single layer.

The natural CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and CRISPR associated protein (Cas) system stores DNA from invading species in the genome to generate an immune response. Memory can also be stored by co-opting this approach as shown in Fig. 2d. Recoverable via sequencing, Shipman et al. (2017) assigned colours to the sequences and by distributing the system amongst many cells, they were able to construct a simple image recorder, with a capacity of 2.6 kilobytes. Akin to this method, another group recorded the occurrence of certain metabolites by initiating production of trigger DNA in their presence (Sheth et al. 2017), others modified the sequence of recorder plasmids, shifting and measuring the ratio between modified and unmodified plasmids (Tang and Liu 2018). Another recent example used DNA methylation to record events; repressed by a zinc finger in the off state, induction by heat methylates the DNA and prevents binding thereby activating permanent expression (Maier et al. 2017).

6.2 Distributed computing

Most of the circuits have thus far been localized in a single cell. However natural systems can organize around different cells by using intercellular communication. Analogously much work has also gone into spatially distributing gene circuits across multiple cells, replacing the intracellular 'wires' with synthetic intercellular signalling pathways as seen in Fig. 2e. This allows components to be reused whilst avoiding any potential 'cross talk', enabling scale up of vastly larger systems from a much smaller library of parts (Macía et al. 2012) as combining the outputs of separate gates can functionally recreate the logic behaviour of a single more complex one (Regot et al. 2011; Macía et al. 2012).

In bacteria, signal propagation between cells can be achieved by adapting the quorum sensing molecules, n-acyl homoserine lactones (AHLs), natural molecules secreted by cells that allows for coordinated activity such as biofilm formation based on cell density (Tamsir et al. 2011). Early circuits were small (Regot et al. 2011) but recent ones are much larger, such as the 6-input multiplexer (Macia et al. 2016) and the most complex system to date, a 1-bit full binary adder; incorporating 22 separate gates distributed amongst 9 specialized mammalian cell types in a complex three dimensional environment (Ausländer et al. 2017). Guiziou et al. (2018) created an automated design framework that did not require cell to cell communication using integrase networks distributed across multiple cells. Memory, like other functions, can also be distributed into different cells, the bistable toggle switch was effectively replicated by two cells containing a NOT logic gate, each gate repressing the other when activated, communicating

through inter-cellular signalling (Urrios et al. 2016). This level of distributed computing can be applicable to metabolic engineering for synthetic product production where different populations each handle a different part of the pathway in a division of labour (Tsoi et al. 2018). Other potential applications include using cell–cell signalling to induce structural self-organisation of tissues (Toda et al. 2018) and pattern formation (Basu et al. 2005; Karig et al. 2018).

7 Challenges and outlooks

There are number of challenges which become especially relevant when scaling up the size and complexity of gene circuits for useful functions. Namely the ability of a circuit to sense inputs and generate useful outputs, manage resource consumption, and maintain modularity of parts. In particular, modularity faces challenges in orthogonality, retroactivity and avoiding undesirable behaviour arising from genetic and cellular context. Each problem will be defined, its importance explained, the current state of the art and future prospects examined.

7.1 Modular, robust and well characterized parts

Generally, larger circuits contain more parts with each part and connection representing another point of failure. Therefore parts must be well characterized with robust and predictable behaviour regardless of context, enabling the design of large-scale circuits to be fast, predictable and reliable. Essentially this refers to modularity, whereby parts retain their inherent function and behavioural characteristics irrespective of the conditions that they are placed in Sauro (2008). This enables two key processes to occur, the decomposition of a system into individual parts which can be constructed and tested separately and the subsequent construction of larger systems and from a library of smaller well understood pieces which generate predictable functions. Modularity is difficult because of several overlapping yet distinct challenges common to biological systems: connectivity, retroactivity, orthogonality, and context effects.

Connectivity in this scenario refers to the ability of parts to communicate reliably with other parts, robust signal propagation down a system is important to generate a consistent output, if a signal degrades due to noise or is unable to be propagated it can disrupt function. Therefore it is desirable to protect circuits by maintaining good connectivity.

Context dependency is the phenomena whereby part behaviour becomes dependent or affected by unwanted interactions from the host, environment, or even its own

composition (Cardinale and Arkin 2012). Unlike in electronic circuit design modules are linked by discrete wires, which when layered correctly are unidirectional and do not propagate signals to unintended recipients and have minimal interaction with the surrounding environment. This trait is increasingly important as circuits get larger and more complex, as cross talk leads to noise and unpredictability. Functionality can break down due to these unwanted interactions with the host, system and environment (Kwok 2010; Wang et al. 2011; Wang and Buck 2012; Liu et al. 2018a, b). A part that does not interact significantly with this context can be surmised to be orthogonal (Liu et al. 2018a). Orthogonality therefore, is important for both functionality and modularity.

The problem of context runs deep, even genetically identical cells in the same environment can show variable phenotypes; attributed in part to stochastic gene expression due to the variable nature of small numbers of interacting molecules (Munsky et al. 2012). Synthetic pathways can elicit responses in the host such as stress or simply display toxicity and circuit performance is tied closely to the health of the host, its physiology, the growth rate, and (discussed later) the availability of resources both internal and external, cell volume and even division state (Cardinale and Arkin 2012; Brophy and Voigt 2014; Liao et al. 2017). Context can also extend into environmental factors such as pH or media (Wang et al. 2011). Temperature in particular has been shown to affect the rate of transcription and the secondary structure of DNA and RNA (Cardinale and Arkin 2012). There is also genetic context; expression can be disturbed by the composition of the adjacent DNA sequences resulting in UTRs affecting the secondary structure and translation rate of the mRNA (Reeve et al. 2014). The size and copy number of the host plasmid can also affect behaviour (Liu et al. 2018b). DNA folding and spacing can affect the steric (spatial) ability for transcription factors to bind, sequence homology can cause deleterious effects and even the orientation of genes on the plasmid can modulate expression levels (Yeung et al. 2017). Since replication of the DNA must occur, there will be errors and possible rendering of parts non-functional. Because many parts have a negative effect on cell health and growth, eventually populations will incorporate an increasingly large subsection of non-functional circuits, this is known as genetic instability (Zhang et al. 2016). This is in spite of selection methods with, for example, antibiotics as the cells will still evolve to only retain the minimum number of genes required.

Retroactivity specifically, was defined by Jayanthi et al. (2013) as "the phenomenon by which a downstream system changes the dynamic state of an upstream system in the process of receiving information from the latter". In this case, downstream and upstream are relative to the intended

flow of information (Del Vecchio et al. 2008). Essentially, this means that attaching example part B to receive the output of example part A will change the way part A behaves, this of course scales as a problem the more connections there are. In biology this can occur when upstream regulatory factors bind their downstream targets. This gets worse as the 'load' increases (the number of sites relative to factors), and is magnified in larger circuits, as when signalling molecules are bound they can no longer transfer information.

The question of context, orthogonality, signal strength and modularity, although distinct are also overlapping challenges and have interacting solutions. One of the best ways to maintain a robust signal in digital like circuits is to maintain a large dynamic range, that is, a large difference or ratio between the ON and OFF state. Although by nature analog graded responses are more vulnerable to noise as they have continuous outputs, this remains true for both as a large dynamic range means the relative effect of any noise is smaller providing the scale of noise remains the same. This protects the signal from degradation as it propagates throughout a system, which itself lends itself to modularity, as a strong signal can mean behaviour remains robust throughout different environments. The tuning tools discussed earlier in Sect. 2 are the often used in adjusting response curves and in optimisation to ensure the output of one part can be received and function as a relevant input for the downstream part. They can also affect the dynamic range as well as modulate retroactivity by increasing expression of the component, (Brophy and Voigt 2014). Alternatively, signal strength can be modulated by the addition of other parts such as amplifiers (Wang et al. 2014).

Other solutions to retroactivity have been attempted by borrowing of concepts from control theory and the subsequent addition of feedback and feed forward loops for insulation, although the latter can only be used based on how the disturbance affects the system, thereby being a much more specialized solution (Del Vecchio et al. 2016). The ideal insulator has zero retroactivity to the input and is not functionally affected in terms of output after taking on the load. One possibility is to use phosphorylation-dephosphorylation cycles since they work on a much faster timescale and do not place a large metabolic burden on the host (Del Vecchio et al. 2008).

To avoid crosstalk within a circuit, we must minimize unwanted interaction with the host and other sections of the circuit. This generally means avoiding repeat use of the same parts, in turn requiring proportionally more parts to increase the complexity and scale of a system making the expansion of the library of well characterized orthogonal parts essential. Alternatively, the circuit can be insulated from unwanted interactions, for example, the circuit can be

constructed as to not rely on the host transcriptional machinery (Liu et al. 2018a) or follow the multicellular distributed approach mentioned previously. The former has gained some traction within the community. The phagemid T7 RNAP has been co-opted to separate the transcriptional machinery from the bacterial host Temme et al. (2012). Chan et al. (2005), refactored the T7 RNAP itself by isolating genes through physical separation, removing or standardizing adjacent sequences to the coding region whilst retaining functionality, making it much simpler to model and easier to manipulate. This has further led to the idea of an entirely orthogonal central dogma, conceptualising the addition of orthogonal DNA polymerases, aminoacyl-tRNA synthetases, and ribosomes for replication and translation respectively, (Liu et al. 2018a). Cello has incorporated into its design space strong terminators preventing RNAP read-through and ribozyme binding sequences and secondary structures that can cleave off the UTR to standardize context (Lou et al. 2012; Nielsen et al. 2016). Carr et al. (2017) developed a degenerate insulator screening (DIS) technique to determine exact levels of insulation desired for bacterial promoters. Lengthy DNA sequences can be compressed by sharing regulatory parts, though paradoxically this will take it out of the genetic context it was characterized in, adding more uncertainty (Brophy and Voigt 2014). Lowering expression of and reducing resource consumption as well as reducing the number of repeated sequences and using inducible promoters can provide a reduction in genetic instability, as well as tying the function of the circuit to host health thereby making it advantageous to host survival (Sleight et al. 2010). Noise can be resisted with negative feedback, as well as from feed forward loops which incorporate both positive and negative regulation, whilst cell–cell communication has also been suggested to have robustness to noise (Zhang et al. 2016).

Understandably, for large scale gene circuits, all of these issues are proportionally magnified. The more parts the more points of failure. Ideally, there would be a large number of highly modular components that could easily be assembled together with predictable behaviour, in a sense 'plug in and play'. However this is far from the case: despite the wide array of parts that have been described in the literature many of them are not well characterized enough to facilitate easy reuse. Simply put, the behaviour of many components becomes less predictable as they taken are taken further away from their original context. This is partly down to a lack of standardization in characterization. Protocols vary across groups, equipment also differs and characterization will be subject to a host of specific design factors such as plasmid, strain and reporter choice which often we do not understand well enough to reliably predict behaviour when they are changed. The latter

problem is a result of our general lack of knowledge regarding basic biological system behaviour. Whilst mapping out potential cross reactivity between a small libraries of parts is reasonably feasible, mapping all potential connections and determining all possible interactions with the host is an order of magnitude more complex and can be even higher if accounting for changing environments and different species across time and space. Not only would this be computationally burdensome and difficult to mathematically model, it would also require a heavy amount of accurate and precise data that simply does not exist in the required scale.

Although optimisation steps listed above are possible, the time cost of performing optimization steps in multiple components is vast and any cross talk only increases the time needed as parts respond to multiple unwanted factors and become more difficult to adjust. In a recent pressure test where organisms were to be engineered to produce 10 molecules unknown in advance, Casini et al. (2018) noticed that literature searches and database entries did not produce actionable data,and even standard procedures such as sequence verification and plasmid/oligo design became bottlenecks. In addition, they had to wait 3-8 weeks for DNA synthesis further reducing available bench time suggesting that there is room for improvement across the board.

The solution to these problems will lie in more accurate and standardized initial characterization of parts, improved understanding of basic circuit–circuit and circuit-host interactions to predict behaviour under different conditions and the reduction of man-hours required using high throughput automated design, construction and characterization methods. It is in this context that large scale circuits could benefit from the scale up and automation of microfluidics for tasks such as genetic assembly and high throughput characterization experiments that gather precise single cell data (which offers a much deeper understanding of host-circuit physiology than population averages), cell free systems enabling rapid prototyping, and methods such as RNA-seq giving us a much wider view of cell state. The resulting data can then be fed into computational simulations and models in order to be fed into the next round of the DBTL cycle. This will result in a positive feedback loop of knowledge; as circuits become better characterized, our understanding of systems will increase, further informing our design, our ability to model and predict behaviour and subsequently reducing the time needed to complete the DBTL cycle.

7.2 Generating relevant inputs and outputs

For circuits to have pertinent real-world applications, they must be able to sense relevant phenomena such as the intracellular concentration of a metabolite or extracellular factors such as heavy metals, RNA, DNA, protein, pH, light, oxygen or heat. In addition they must actuate outputs that are valuable to human endeavour. By doing so gene circuits can make the leap from interesting academic problems to useful biotechnological applications.

The generation of novel functional parts often finds its inspiration in already existing natural systems, although a degree of characterization and refining of these parts is necessary to add them to the toolbox (Wang et al. 2015a). Existing proteins have been engineered to sense new metabolites through directed evolution (Collins et al. 2006; Taylor et al. 2016) and some hybrids with novel function have also been developed. A synthetic light-sensitive sensor kinase (Cph1–EnvZ) was made in *E. coli* by fusing the photoreceptor domain of the phytochrome Cph1 protein from Synechocystis to the intracellular signal transduction domain of the *E. coli* EnvZ kinase, yielding a functional sensor chimera (Tabor et al. 2009). Antibody domains have been fused with DNA binding domains and activated via ligand induced dimerization to enable sensing of new molecules (Chang et al. 2018) and chimeric custom proteins have also been demonstrated with modified Notch receptors (Morsut et al. 2016). In some cases sensors can be modified to work in different hosts, as demonstrated with the retooling of TetR family repressors, to work in human embryonic kidney (HEK293) and Chinese hamster ovary (CHO) cells (Stanton et al. 2014). Examples of outputs include useful biological or small molecule products (Paddon and Keasling 2014), simple signalling responses to difficult to detect stimuli (Wang et al. 2013a; Bereza-Malcolm et al. 2015), to the cancer targeting classifier circuits that secrete apoptotic proteins (Xie et al. 2011).

Larger scale circuits will likely include a greater number of these unique sensing and output parts that will enable complex programmable functionality. For example, a bioremediation based system could potentially monitor many environmental inputs and secrete specific enzymes that degrade waste products in response. Circuits would benefit then, from a larger library of unique well characterized and modular parts, the general challenges and solutions of which have already been discussed. In particular, the ability to link novel inputs and outputs would benefit strongly from improved protein engineering techniques in modifying existing functionality or the building of chimeric proteins. In turn this would strongly benefit from deep structure function understanding to avoid time consuming trial and error experimentation (Wang et al. 2013b). Bioinformatics may be able to play a strong role too, in estimating structure and function of candidate proteins from their genetic sequences to narrow the design space (Stanton et al. 2013).

7.3 Metabolic burden

Metabolic burden or load can be understood as the resource consumption required by the engineered system upon the host. The concerns of burden are often the focus of metabolic engineers when optimizing a product producing pathway, however it is also relevant in the construction of gene circuits as resource limitation fundamentally affects system behaviour. Cells have an upper limit of nutrient and energy intake that limits all cellular activity, one of these hard limits can usually be described in terms of ATP. Cells can compensate somewhat by increasing respiration and catabolism but under too much strain there is a sharp drop in total protein production to near 0 and often results in the collapse of the population (Wu et al. 2016). The effect of foreign protein production on the host was spotted early on; increasing amounts of foreign protein production led to decreasing growth rate in *E. coli* (Bentley et al. 1990; Bhattacharya and Dubey 1995). The amino acid content of recombinant proteins has also been shown to affect production levels (Bonomo and Gill 2005) whilst the amount of free ribosomes and RNAPs is also important, itself affected by presence of plasmid DNA (Birnbaum and Bailey 1991). There is evidence that genetic load resembles the equations of Ohm's law for resistance in electrical circuits (Carbonell-Ballestero et al. 2016). Other findings have shown that 'leaky' basal levels of transcription and high plasmid copy number contribute to the protein burden (Lee et al. 2016), with copy number also changing gene circuit expression as well as in the host cell. Increasing copy number increases expression of the receptors to the system input, thereby increasing retroactivity, decreasing the sensitivity and dynamic range of repressor based systems given the same amount of repressor, and vice versa for activator based systems (Wang et al. 2015a; Liu et al. 2018b).

Managing load requires accurate characterization and calculated mitigation. The copy number and general expression levels of the circuit should be as low as is essential for predictable behaviour. If necessary, the circuit can be spread into multiple cells following the principles of distributed computing. RNA based control tends to be the least burdensome on host metabolism; Lapique and Benenson (2017) even combined two orthogonal binding sites into one DNA sequence using recombinases to reversibly express equal amounts of the forward and reverse DNA sequence, thereby generating two separate species of RNA, each with one functional and orthogonal binding site. Ceroni et al. (2015) inserted a constitutively expressed GFP element that would act as a tracker for metabolic change in the host. The Cello design framework manages burden through simulating the load on each cell by factoring in the impact on growth relative to the functional activity of the input promoter (Nielsen et al. 2016). This information can be used by the designer to optimize the circuit (Wu et al. 2016). Liao et al. (2017) created a model that considers different RNA levels, the proteome (dividing it into gene expression apparatus and metabolic machinery), resource partitioning (including ATP and amino acid synthesis) as well as other factors such as growth, copy number and cell volume. The CRISPR-Cas system has been used to attenuate leaky gene expression with T7 RNAP and has been shown to improve growth in systems with previously toxic leaky expression (McCutcheon et al. 2018). Incoherent feedforwards loops (iFFL) have been engineered into promoters using transcription-activator-like effectors (TALEs) which stabilised expression level at different copy numbers (Segall-Shapiro et al. 2018) whilst Lee et al. (2016) created single copy plasmids with stable expression.

Larger circuits mean more components and this will inevitably have a proportionally larger effect on metabolic load. Selecting parts that have minimal resource consumption (such as RNA based tools), and reducing consumption of existing through tuning will constitute a large part of the solution. In the latter case, there are complications as once a part is modified away from its original specifications, it will need to be characterized again. Furthermore, reducing the expression level can have negative effects on signal robustness and increase the susceptibility towards unwanted interactions and noise. The literature has suggested that parts with analog behaviour are significantly more resource efficient and the authors suggest that hybrid devices will likely be common in the future (Sarpeshkar 2014). Parts might be also arranged so as not to overlap on the type of load they produce, for example, distributing load across both transcription and translation, or they might be combined into a single layer that does not require communication between parts for sub-computation as demonstrated earlier (Weinberg et al. 2017). However the authors do note that this means the performance of circuits cannot be predicted based off its constituent parts. Another solution would be to distribute the circuit into different consortia, as discussed earlier; likely to become a common approach as reduction of load from individual parts cannot decrease indefinitely.

Tools that allow us to monitor and predict load will also become increasingly important. Here the related field of metabolic engineering may hold some promise. High throughput experimentation again will allow us to gather a larger amount of data in a shorter space of time and tools such RNA-seq or whole cell mass spectrometry that offer a wide view of cellular gene expression and metabolism to be key in deciphering the interactions between circuit and host (Liu et al. 2018b). Here the related field of metabolic

engineering holds promise, having developed tools such as metabolic flux balance analysis to predict the distribution of important resources such as carbon (Yang et al. 2007). Finally, like as before, as data becomes more readily available and accurate, computational prediction will become increasingly important.

8 Concluding remarks

Gene circuits hold great potential for addressing real-world challenges including applications in biomanufacturing (Si and Zhao 2016), biosensing (Bernard and Wang 2017) and biotherapy (Riglar et al. 2017). Larger scale systems potentially enable more intricate control and the larger circuits thus far discussed have been able to compute more complex functions than the smaller ones. Circuits have been steadily increasing in size, albeit slowly, and the molecular toolbox available to synthetic biologists is now larger than ever before. There has been a significant expansion of orthogonal parts that enable a vast quantity of versatile methods to control behaviour, providing a solid foundation for constructing complex circuits.

However there remains a significant lack of predictability of the behaviour of parts when put together that scales in larger systems preventing regular reuse of all but the most basic parts. Modularity and standardization remain issues for biological components and there are fundamental gaps in our knowledge on basic biological processes that prevent us from accurately predicting changes. Recent advances in characterization techniques enable high throughput experiments providing single cell and genome or proteome levels of data, whilst new methods in microfluidics and cell free systems potentially allow for high speed prototyping of systems in a matter of hours and days instead of weeks. The increase in time efficiency in the laboratory whilst simultaneously gathering larger data sets promises a positive feedback loop that enables increasingly faster iterations of the DBTL cycle that concurrently will result in larger more robust systems as well as a leap in our fundamental understanding of biological interactions. Automated systems can already be seen in industry at the start-up stage, at companies such as Ginkgo bioworks and Zymergen (Anne Ravanona 2015; Silicon-review Team 2017), performing industrial strain engineering with heavy use of robotics, next generation sequencing, automation and software. Some of these companies like Ginkgo, are spin-off companies from universities seeking to capitalize on their proprietary technologies and in 2017 50 synthetic biology companies managed to raise 1.7 billion US dollars in funding (Calvin Schmidt 2018). Both academia and industry could benefit from continued and potential closer collaboration.

Academia is well placed to investigate the basic biochemistry of the systems it engineers, furthering understanding of the relationship between circuit and host and do the groundwork that enables basic modular functional parts whilst industry works to apply the principles to relevant real-world applications. It would be pertinent for industry here to establish a forum for discussion of specific problems that need to be tackled for relevant market needs that academia could potentially cooperate on. Closer partnership will require adoption of model organisms that are more relevant for biotechnology and close collaboration with fields such as chemical engineering that work with relevant techniques in order to bridge the gap between proof of concept and industrially sized production (Moser et al. 2012).

References

Alipanahi B, Delong A, Weirauch MT, Frey BJ (2015) Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nat Biotechnol. https://doi.org/10.1038/nbt.3300

Anderson JC, Voigt CA, Arkin AP (2007) Environmental signal integration by a modular AND gate. Mol Syst Biol 3:133. https://doi.org/10.1038/msb4100173

Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. Mol Syst Biol 2:1–14. https://doi.org/10.1038/msb4100073

Ang J, Harris E, Hussey BJ, Kil R, McMillen DR (2013) Tuning response curves for synthetic biology. ACS Synth Biol 2:547–567. https://doi.org/10.1021/sb4000564

Anne R (2015) Trailblazing women: Reshma Shetty, Co-founder of Gingko Bioworks| HuffPost. In: huffingtonpost.com. https://www.huffingtonpost.com/anne-ravanona/trailblazing-women-reshma_b_7173272.html. Accessed 24 Jun 2018

Auslander S, Fussenegger M (2014) Toehold gene switches make big footprints. Nature 516:5–6. https://doi.org/10.1038/516333a

Ausländer D, Ausländer S, Pierrat X, Hellmann L, Rachid L, Fussenegger M (2017) Programmable full-adder computations in communicating three-dimensional cell cultures. Nat Methods 15:57–60. https://doi.org/10.1038/nmeth.4505

Baret J-C, Miller OJ, Taly V et al (2009) Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. Lab Chip 9:1850. https://doi.org/10.1039/b902504a

Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R (2005) A synthetic multicellular system for programmed pattern formation. Nature 434:1130–1134. https://doi.org/10.1038/nature03461

Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. Nature 405:590–593. https://doi.org/10.1038/35014651

Bennett MR, Hasty J (2009) Microfluidic devices for measuring gene network dynamics in single cells. Nat Rev Genet 10:628–638. https://doi.org/10.1038/nrg2625

Bentley WE, Mirjalili N, Andersen DC, Davis RH, Kompala DS (1990) Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. Biotechnol Bioeng 35:668–681. https://doi.org/10.1002/bit.260350704

Bereza-Malcolm LT, Mann G, Franks AE (2015) Environmental sensing of heavy metals through whole cell microbial biosensors: a synthetic biology approach. ACS Synth Biol 4:535–546. https://doi.org/10.1021/sb500286r

Bernard E, Wang B (2017) Synthetic cell-based sensors with programmed selectivity and sensitivity. In: Prickril B, Rasooly A (eds) Biosensors and biodetection: methods and protocols, volume 2: electrochemical, bioelectronic, piezoelectric, cellular and molecular biosensors, vol 2. Springer, New York, pp 349–363

Bhatia SP, Smanski MJ, Voigt CA, Densmore DM (2017) Genetic design via combinatorial constraint specification. ACS Synth Biol 6(11):2130–2135. https://doi.org/10.1021/acssynbio.7b00154

Bhattacharya SK, Dubey AK (1995) Metabolic burden as reflected by maintenance coefficient of recombinant *Escherichia coli* over-expressing target gene. Biotechnol Lett 17:1155–1160. https://doi.org/10.1007/BF00128377

Birnbaum S, Bailey JE (1991) Plasmid presence changes the relative levels of many host cell proteins and ribosome components in recombinant *Escherichia coli*. Biotechnol Bioeng 37:736–745. https://doi.org/10.1002/bit.260370808

Bonger KM, Chen L, Liu CW, Wandless TJ (2011) Small-molecule displacement of a cryptic degron causes conditional protein degradation. Nat Chem Biol 7:531–537. https://doi.org/10.1038/nchembio.598

Bonnet J, Subsoontorn P, Endy D (2012) Rewritable digital data storage in live cells via engineered control of recombination directionality. Proc Natl Acad Sci 109:8884–8889. https://doi.org/10.1073/pnas.1202344109

Bonomo J, Gill RT (2005) Amino acid content of recombinant proteins influences the metabolic burden response. Biotechnol Bioeng 90:116–126. https://doi.org/10.1002/bit.20436

Bradley RW, Wang B (2015) Designer cell signal processing circuits for biotechnology. New Biotechnol 32:635–643. https://doi.org/10.1016/j.nbt.2014.12.009

Bradley RW, Buck M, Wang B (2016a) Recognizing and engineering digital-like logic gates and switches in gene regulatory networks. Curr Opin Microbiol 33:74–82. https://doi.org/10.1016/j.mib.2016.07.004

Bradley RW, Buck M, Wang B (2016b) Tools and principles for microbial gene circuit engineering. J Mol Biol 428:862–888. https://doi.org/10.1016/j.jmb.2015.10.004

Brophy JAN, Voigt CA (2014) Principles of genetic circuit design. Nat Methods 11:508–520. https://doi.org/10.1038/nmeth.2926

Callura JM, Cantor CR, Collins JJ (2012) Genetic switchboard for synthetic biology applications. Proc Natl Acad Sci 109:5850–5855. https://doi.org/10.1073/pnas.1203808109

Calvin Schmidt (2018) These fifty synthetic biology companies raised $1.7B in 2017—SynBioBeta. In: Synbiobeta. https://synbiobeta.com/fifty-synthetic-biology-companies-raised-1-7b-2017/. Accessed 29 Jun 2018

Carbonell-Ballestero M, Garcia-Ramallo E, Montañez R, Rodriguez-Caso C, Macía J (2016) Dealing with the genetic load in bacterial synthetic biology circuits: convergences with the Ohm's law. Nucleic Acids Res 44:496–507. https://doi.org/10.1093/nar/gkv1280

Cardinale S, Arkin AP (2012) Contextualizing context for synthetic biology—identifying causes of failure of synthetic biological systems. Biotechnol J 7:856–866. https://doi.org/10.1002/biot.201200085

Carr SB, Beal J, Densmore DM (2017) Reducing DNA context dependence in bacterial promoters. PLoS ONE 12:1–15. https://doi.org/10.1371/journal.pone.0176013

Casini A, MacDonald JT, De JJ et al (2014) One-pot DNA construction for synthetic biology: the modular overlap-directed assembly with linkers (MODAL) strategy. Nucleic Acids Res 42:e7. https://doi.org/10.1093/nar/gkt915

Casini A, Chang F-Y, Eluere R et al (2018) A pressure test to make 10 molecules in 90 days: external evaluation of methods to engineer biology. J Am Chem Soc 140:4302–4316. https://doi.org/10.1021/jacs.7b13292

Ceroni F, Algar R, Stan GB, Ellis T (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden. Nat Methods 12:415–418. https://doi.org/10.1038/nmeth.3339

Chan LY, Kosuri S, Endy D (2005) Refactoring bacteriophage T7. Mol Syst Biol 1:E1–E10. https://doi.org/10.1038/msb4100025

Chandran D, Copeland WB, Sleight SC, Sauro HM (2008) Mathematical modeling and synthetic biology. Drug Discov Today Dis Model 5:299–309. https://doi.org/10.1016/j.ddmod.2009.07.002

Chandran D, Bergmann FT, Sauro HM (2009) TinkerCell: modular CAD tool for synthetic biology. J Biol Eng 3:19. https://doi.org/10.1186/1754-1611-3-19

Chang H-J, Mayonove P, Zavala A et al (2018) A modular receptor platform to expand the sensing repertoire of bacteria. ACS Synth Biol 7:166–175. https://doi.org/10.1021/acssynbio.7b00266

Chappell J, Takahashi MK, Lucks JB (2015) Creating small transcription activating RNAs. Nat Chem Biol 11:214–220. https://doi.org/10.1038/nchembio.1737

Chen YJ, Liu P, Nielsen AAK et al (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. Nat Methods 10:659–664. https://doi.org/10.1038/nmeth.2515

Chen Y, Ho JML, Shis DL et al (2018) Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. Nat Commun 9:64. https://doi.org/10.1038/s41467-017-02473-5

Clarke LJ, Kitney RI (2016) Synthetic biology in the UK—an outline of plans and progress. Synth Syst Biotechnol 1:243–257. https://doi.org/10.1016/j.synbio.2016.09.003

Collins CH, Leadbetter JR, Arnold FH (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. Nat Biotechnol 24:708–712. https://doi.org/10.1038/nbt1209

Cox RS, Madsen C, McLaughlin JA et al (2018) Synthetic biology open language (SBOL) version 2.2.0. J Integr Bioinform 15:30–132. https://doi.org/10.1515/jib-2018-0001

Cuperus JT, Groves B, Kuchina A et al (2017) Deep learning of the regulatory grammar of yeast 5′ untranslated regions from 500,000 random sequences. Genome Res 163:1–10. https://doi.org/10.1101/gr.224964.117.7

Czar MJ, Cai Y, Peccoud J (2009) Writing DNA with genoCAD™. Nucleic Acids Res 37:40–47. https://doi.org/10.1093/nar/gkp361

Del Vecchio D, Ninfa AJ, Sontag ED (2008) Modular cell biology: retroactivity and insulation. Mol Syst Biol 4:161. https://doi.org/10.1038/msb4100204

Del Vecchio D, Dy AJ, Qian Y (2016) Control theory meets synthetic biology. J R Soc 13:3–43. https://doi.org/10.1098/rsif.2016.0380

Desai TA, Rodionov DA, Gelfand MS, Alm EJ, Rao CV (2009) Engineering transcription factors with novel DNA-binding specificity using comparative genomics. Nucleic Acids Res 37:2493–2503. https://doi.org/10.1093/nar/gkp079

Didovyk A, Borek B, Hasty J, Tsimring L (2016) Orthogonal modular gene repression in *Escherichia coli* using engineered CRISPR/Cas9. ACS Synth Biol 5:81–88. https://doi.org/10.1021/acssynbio.5b00147

Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403:335–338. https://doi.org/10.1038/35002125

Endler L, Rodriguez N, Juty N et al (2009) Designing and encoding models for synthetic biology. J R Soc Interface 6:S405–S417. https://doi.org/10.1098/rsif.2009.0035.focus

Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. PLoS ONE 3:e3647. https://doi.org/10.1371/journal.pone.0003647

Fan J, Villarreal F, Weyers B et al (2017) Multi-dimensional studies of synthetic genetic promoters enabled by microfluidic impact printing. Lab Chip 17:2198–2207. https://doi.org/10.1039/C7LC00382J

Farzadfard F, Lu TK (2014) Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. Science 346:1256272. https://doi.org/10.1126/science.1256272

Finney A, Hucka M (2003) Systems biology markup language: level 2 and beyond. Biochem Soc Trans 31:1472–1473. https://doi.org/10.1042/bst0311472

Funahashi BA, Matsuoka Y, Jouraku A, Morohashi M, Kikuchi N, Kitano H (2008) A versatile modeling tool for biochemical networks. Proc IEEE 96:1254–1265. https://doi.org/10.1109/JPROC.2008.925458

Gach PC, Shih SCC, Sustarich J et al (2016) A droplet microfluidic platform for automating genetic engineering. ACS Synth Biol 5:426–433. https://doi.org/10.1021/acssynbio.6b00011

Gander MW, Vrana JD, Voje WE, Carothers JM, Klavins E (2017) Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. Nat Commun 8:15459. https://doi.org/10.1038/ncomms15459

Garamella J, Marshall R, Rustad M, Noireaux V (2016) The all *E. coli* TX-TL toolbox 2.0: a platform for cell-free synthetic biology. ACS Synth Biol 5:344–355. https://doi.org/10.1021/acssynbio.5b00296

Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. Nature 403:339–342. https://doi.org/10.1038/35002131

Garg A, Lohmueller JJ, Silver PA, Armel TZ (2012) Engineering synthetic TAL effectors with orthogonal target sites. Nucleic Acids Res 40:7584–7595. https://doi.org/10.1093/nar/gks404

Gibson DG, Young L, Chuang R, Venter JC, Hutchison CA, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6:343–345. https://doi.org/10.1038/nmeth.1318

Goler JA, Bramlett BW, Peccoud J (2008) Genetic design: rising above the sequence. Trends Biotechnol 26:538–544. https://doi.org/10.1016/j.tibtech.2008.06.003

Golightly A, Wilkinson DJ (2011) Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo. Interface Focus. https://doi.org/10.1098/rsfs.2011.0047

Goranov AI, Gulati A, Dephoure N et al (2013) Changes in cell morphology are coordinated with cell growth through the TORC1 pathway. Curr Biol 23:1269–1279. https://doi.org/10.1016/j.cub.2013.05.035

Gorochowski TE, Borujeni AE, Park Y et al (2017) Genetic circuit characterization and debugging using RNA-seq. Mol Syst Biol 13:952. https://doi.org/10.15252/msb.20167461

Grant PK, Dalchau N, Brown JR et al (2016) Orthogonal intercellular signaling for programmed spatial behavior. Mol Syst Biol 12:849. https://doi.org/10.15252/msb.20156590

Green AA, Silver PA, Collins JJ, Yin P (2014) Toehold switches: de-novo-designed regulators of gene expression. Cell 159:925–939. https://doi.org/10.1016/j.cell.2014.10.002

Green AA, Kim J, Ma D, Silver PA, Collins JJ, Yin P (2017) Complex cellular logic computation using ribocomputing devices. Nature 548:117–121. https://doi.org/10.1038/nature23271

Guet C (2002) Combinatorial synthesis of genetic networks. Science 296:1466–1470. https://doi.org/10.1126/science.1067407

Guido NJ, Wang X, Adalsteinsson D et al (2006) A bottom-up approach to gene regulation. Nature 439:856–860. https://doi.org/10.1038/nature04473

Guiziou S, Ulliana F, Moreau V, Leclere M, Bonnet J (2018) An automated design framework for multicellular recombinase logic. ACS Synth Biol 7:1406–1412. https://doi.org/10.1021/acssynbio.8b00016

Gyorgy A, Murray RM (2016) Quantifying resource competition and its effects in the TX-TL system. In: 2016 IEEE 55th conference on decision control CDC 2016, vol 1, pp 3363–3368. https://doi.org/10.1109/CDC.2016.7798775

Halleran AD, Swaminathan A, Murray RM (2018) Single day construction of multigene circuits with 3G assembly. ACS Synth Biol 7:1477–1480. https://doi.org/10.1021/acssynbio.8b00060

Hayden D, Chang YH, Goncalves J, Tomlin CJ (2016) Sparse network identifiability via Compressed sensing. Automatica 68:9–17. https://doi.org/10.1016/j.automatica.2016.01.008

Heinemann J, Deng K, Shih SCC et al (2017) On-chip integration of droplet microfluidics and nanostructure-initiator mass spectrometry for enzyme screening. Lab Chip 17:323–331. https://doi.org/10.1039/C6LC01182A

Hill AD, Tomshine JR, Weeding EMB, Sotiropoulos V, Kaznessis YN (2008) SynBioSS: the synthetic biology modeling suite. Bioinformatics 24:2551–2553. https://doi.org/10.1093/bioinformatics/btn468

Hou J, Zeng W, Zong Y et al (2018) Engineering the ultrasensitive transcription factors by fusing a modular oligomerization domain. ACS Synth Biol 7:1188–1194. https://doi.org/10.1021/acssynbio.7b00414

Hsiao V, Hori Y, Rothemund PW, Murray RM (2016) A population-based temporal logic gate for timing and recording chemical events. Mol Syst Biol 12:869. https://doi.org/10.15252/msb.20156663

Huang H, Densmore D (2014) Integration of microfluidics into the synthetic biology design flow. Lab Chip 14:3459–3474. https://doi.org/10.1039/C4LC00509K

Hucka M, Finney A, Sauro HM et al (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19:524–531. https://doi.org/10.1093/bioinformatics/btg015

Hutchison CA, Chuang R-Y, Noskov VN et al (2016) Design and synthesis of a minimal bacterial genome. Science 351:aad6253. https://doi.org/10.1126/science.aad6253

Huynh L, Tagkopoulos I (2014) Optimal part and module selection for synthetic gene circuit design automation. ACS Synth Biol 3:556–564. https://doi.org/10.1021/sb400139h

Huynh L, Tagkopoulos I (2016) A parts database with consensus parameter estimation for synthetic circuit design. ACS Synth Biol. https://doi.org/10.1021/acssynbio.5b00205

Iyer S, Karig DK, Norred SE, Simpson ML, Doktycz MJ (2013) Multi-input regulation and logic with T7 promoters in cells and cell-free systems. PLoS ONE 8:1–12. https://doi.org/10.1371/journal.pone.0078442

Jayanthi S, Nilgiriwala KS, Del Vecchio D (2013) Retroactivity controls the temporal dynamics of gene transcription. ACS Synth Biol 2:431–441. https://doi.org/10.1021/sb300098w

Kahl LJ, Endy D (2013) A survey of enabling technologies in synthetic biology. J Biol Eng 7:13. https://doi.org/10.1186/1754-1611-7-13

Karagiannis P, Fujita Y, Saito H (2016) RNA-based gene circuits for cell regulation. Proc Jpn Acad Ser B Phys Biol Sci 92:412–422. https://doi.org/10.2183/pjab.92.412

Karig DK, Iyer S, Simpson ML, Doktycz MJ (2012) Expression optimization and synthetic gene networks in cell-free systems. Nucleic Acids Res 40:3763–3774. https://doi.org/10.1093/nar/gkr1191

Karig D, Martini KM, Lu T, DeLateur NA, Goldenfeld N, Weiss R (2018) Stochastic turing patterns in a synthetic bacterial population. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.1720770115

Kelly JR, Rubin AJ, Davis JH et al (2009) Measuring the activity of BioBrick promoters using an in vivo reference standard. J Biol Eng 3:4. https://doi.org/10.1186/1754-1611-3-4

Khalil AS, Lu TK, Bashor CJ et al (2012) A synthetic biology framework for programming eukaryotic transcription functions. Cell 150:647–658. https://doi.org/10.1016/j.cell.2012.05.045

Kiani S, Beal J, Ebrahimkhani MR et al (2014) CRISPR transcriptional repression devices and layered circuits in mammalian cells. Nat Methods 11:723–726. https://doi.org/10.1038/nmeth.2969

Knight T (2003) Idempotent vector design for standard assembly of biobricks. MIT Artificial Intelligence Laboratory; MIT Synthetic Biology Working Group

Kreimer A, Zeng H, Edwards MD et al (2017) Predicting gene expression in massively parallel reporter assays: a comparative study. Hum Mutat 38:1240–1250. https://doi.org/10.1002/humu.23197

Kwok R (2010) Five hard truths for synthetic biology. Nature 463:288–290. https://doi.org/10.1038/463288a

Lapique N, Benenson Y (2017) Genetic programs can be compressed and autonomously decompressed in live cells. Nat Nanotechnol. https://doi.org/10.1038/s41565-017-0004-z

Lee T-H, Maheshri N (2012) A regulatory role for repeated decoy transcription factor binding sites in target gene expression. Mol Syst Biol 8:1–11. https://doi.org/10.1038/msb.2012.7

Lee JH, Daugharthy ER, Scheiman J et al (2014) Highly multiplexed subcellular RNA sequencing in situ. Science 343:1360–1363. https://doi.org/10.1126/science.1250212

Lee JW, Gyorgy A, Cameron DE et al (2016) Creating single-copy genetic circuits. Mol Cell 63:329–336. https://doi.org/10.1016/j.molcel.2016.06.006

Lewis DD, Chavez M, Chiu KL, Tan C (2018) Reconfigurable analog signal processing by living cells. ACS Synth Biol 7:107–120. https://doi.org/10.1021/acssynbio.7b00255

Li Y, Jiang Y, Chen H et al (2015) Modular construction of mammalian gene circuits using TALE transcriptional repressors. Nat Chem Biol 11:207–213. https://doi.org/10.1038/nchembio.1736

Liao C, Blanchard AE, Lu T (2017) An integrative circuit–host modelling framework for predicting synthetic gene network behaviours. Nat Microbiol 2:1658–1666. https://doi.org/10.1038/s41564-017-0022-5

Liu CC, Jewett MC, Chin JW, Voigt CA (2018a) Toward an orthogonal central dogma. Nat Chem Biol 14:103–106. https://doi.org/10.1038/nchembio.2554

Liu Q, Schumacher J, Wan X, Lou C, Wang B (2018b) Orthogonality and burdens of heterologous AND gate gene circuits in E. coli. ACS Synth Biol 7:553–564. https://doi.org/10.1021/acssynbio.7b00328

Lohmueller JJ, Armel TZ, Silver PA (2012) A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. Nucleic Acids Res 40:5180–5187. https://doi.org/10.1093/nar/gks142

Lou C, Stanton B, Chen YJ, Munsky B, Voigt CA (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. Nat Biotechnol 30:1137–1142. https://doi.org/10.1038/nbt.2401

MacDonald JT, Barnes C, Kitney RI, Freemont PS, Stan G-BV (2011) Computational design approaches and tools for synthetic biology. Integr Biol 3:97. https://doi.org/10.1039/c0ib00077a

Machado LFM, Dixon N, Ragauskas AJ et al (2016) Development and substrate specificity screening of an in vivo biosensor for the detection of biomass derived aromatic chemical building blocks. Chem Commun 52:11402–11405. https://doi.org/10.1039/C6CC04559F

Macia J, Manzoni R, Conde N et al (2016) Implementation of complex biological logic circuits using spatially distributed multicellular consortia. PLoS Comput Biol 12:1–24. https://doi.org/10.1371/journal.pcbi.1004685

Macía J, Posas F, Solé RV (2012) Distributed computation: the new wave of synthetic biology devices. Trends Biotechnol 30:342–349. https://doi.org/10.1016/j.tibtech.2012.03.006

Maier JAH, Möhrle R, Jeltsch A (2017) Design of synthetic epigenetic circuits featuring memory effects and reversible switching based on DNA methylation. Nat Commun. https://doi.org/10.1038/ncomms15336

Marbach D, Prill RJ, Schaffter T, Mattiussi C, Floreano D, Stolovitzky G (2010) Revealing strengths and weaknesses of methods for gene network inference. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.0913357107

Marchisio MA, Stelling J (2009) Computational design tools for synthetic biology. Curr Opin Biotechnol 20:479–485. https://doi.org/10.1016/j.copbio.2009.08.007

Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U (2006) Plasticity of the cis-regulatory input function of a gene. PLoS Biol 4:e45. https://doi.org/10.1371/journal.pbio.0040045

McCutcheon SR, Chiu KL, Lewis DD, Tan C (2018) CRISPR-Cas expands dynamic range of gene expression from T7RNAP promoters. Biotechnol J 13:1700167. https://doi.org/10.1002/biot.201700167

Mitchell R, Dori YJ, Kuldell NH (2011) Experiential engineering through iGEM-an undergraduate summer competition in synthetic biology. J Sci Educ Technol 20:156–160. https://doi.org/10.1007/s10956-010-9242-7

Moe-Behrens GH (2013) The biological microprocessor, or how to build a computer with biological parts. Comput Struct Biotechnol J 7:e201304003. https://doi.org/10.5936/csbj.201304003

Moon TS, Lou C, Tamsir A, Stanton BC, Voigt CA (2012) Genetic programs constructed from layered logic gates in single cells. Nature 491:249–253. https://doi.org/10.1038/nature11516

Morsut L, Roybal KT, Xiong X et al (2016) Engineering customized cell sensing and response behaviors using synthetic notch receptors. Cell 164:780–791. https://doi.org/10.1016/j.cell.2016.01.012

Moser F, Broers NJ, Hartmans S et al (2012) Genetic circuit performance under conditions relevant for industrial bioreactors. ACS Synth Biol 1:555–564. https://doi.org/10.1021/sb3000832

Müller M, Ausländer S, Spinnler A et al (2017) Designed cell consortia as fragrance-programmable analog-to-digital converters. Nat Chem Biol 13:309–316. https://doi.org/10.1038/nchembio.2281

Munsky B, Neuert G, van Oudenaarden A (2012) Using gene expression noise to understand gene regulation. Science 336:183–187. https://doi.org/10.1126/science.1216379

Mutalik VK, Qi L, Guimaraes JC, Lucks JB, Arkin AP (2012) Rationally designed families of orthogonal RNA regulators of translation. Nat Chem Biol 8:447–454. https://doi.org/10.1038/nchembio.919

Myhrvold C, Silver PA (2015) Using synthetic RNAs as scaffolds and regulators. Nat Struct Mol Biol 22:8–10. https://doi.org/10.1038/nsmb.2944

Niederholtmeyer H, Xu L, Maerkl SJ (2013) Real-time mRNA measurement during an in vitro transcription and translation reaction using binary probes. ACS Synth Biol 2:411–417. https://doi.org/10.1021/sb300104f

Nielsen AA, Voigt CA (2014) Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. Mol Syst Biol 10:763. https://doi.org/10.15252/msb.20145735

Nielsen AAK, Der BS, Shin J et al (2016) Genetic circuit design automation. Science 352:aac7341. https://doi.org/10.1126/science.aac7341

Norred SE, Caveney PM, Retterer ST et al (2015) Sealable femtoliter chamber arrays for cell-free biology. J Vis Exp 28:e52616–e52616. https://doi.org/10.3791/52616

Ouedraogo R, Daumas A, Capo C, Mege J-L, Textoris J (2013) Whole-cell MALDI-TOF mass spectrometry is an accurate and rapid method to analyze different modes of macrophage activation. J Vis Exp. https://doi.org/10.3791/50926

Paddon CJ, Keasling JD (2014) Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nat Rev Microbiol 12:355–367. https://doi.org/10.1038/nrmicro3240

Pardee K, Green AA, Ferrante T et al (2014) Paper-based synthetic gene networks. Cell 159:940–954. https://doi.org/10.1016/j.cell.2014.10.004

Pedersen M, Phillips A (2009) Towards programming languages for genetic engineering of living cells. J R Soc Interface 6:S437–S450. https://doi.org/10.1098/rsif.2008.0516.focus

Pothoulakis G, Ceroni F, Reeve B, Ellis T (2014) The spinach RNA aptamer as a characterization tool for synthetic biology. ACS Synth Biol 3:182–187. https://doi.org/10.1021/sb400089c

Privman V, Strack G, Solenov D, Pita M, Katz E (2008) Optimization of enzymatic biochemical logic for noise reduction and scalability: How many biocomputing gates can be interconnected in a circuit? J Phys Chem B 112:11777–11784. https://doi.org/10.1021/jp802673q

Purcell O, Lu TK (2014) Synthetic analog and digital circuits for cellular computation and memory. Curr Opin Biotechnol 29:146–155. https://doi.org/10.1016/j.copbio.2014.04.009

Qi L, Haurwitz RE, Shao W, Doudna JA, Arkin AP (2012) RNA processing enables predictable programming of gene expression. Nat Biotechnol 30:1002–1006. https://doi.org/10.1038/nbt.2355

Qi LS, Larson MH, Gilbert LA et al (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152:1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

Reeve B, Hargest T, Gilbert C, Ellis T (2014) Predicting translation initiation rates for designing synthetic biology. Front Bioeng Biotechnol 2:1–6. https://doi.org/10.3389/fbioe.2014.00001

Regot S, Macia J, Conde N et al (2011) Distributed biological computation with multicellular engineered networks. Nature 469:207–211. https://doi.org/10.1038/nature09679

Rhodius VA, Segall-Shapiro TH, Sharon BD et al (2013) Design of orthogonal genetic switches based on a crosstalk map of σs, anti-σs, and promoters. Mol Syst Biol 9:702. https://doi.org/10.1038/msb.2013.58

Riglar DT, Giessen TW, Baym M et al (2017) Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. Nat Biotechnol 35:653–658. https://doi.org/10.1038/nbt.3879

Rinaudo K, Bleris L, Maddamsetti R, Subramanian S, Weiss R, Benenson Y (2007) A universal RNAi-based logic evaluator that operates in mammalian cells. Nat Biotechnol 25:795–801. https://doi.org/10.1038/nbt1307

Ro D-K, Paradise EM, Ouellet M et al (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440:940–943. https://doi.org/10.1038/nature04640

Roehner N, Beal J, Clancy K et al (2016) Sharing structure and function in biological design with SBOL 2.0. ACS Synth Biol 5:498–506. https://doi.org/10.1021/acssynbio.5b00215

Roquet N, Soleimany AP, Ferris AC, Aaronson S, Lu TK (2016) Synthetic recombinase-based state machines in living cells. Science 353:aad8559. https://doi.org/10.1126/science.aad8559

Rubens JR, Selvaggio G, Lu TK (2016) Synthetic mixed-signal computation in living cells. Nat Commun 7:11658. https://doi.org/10.1038/ncomms11658

Rudge TJ, Brown JR, Federici F et al (2016) Characterization of intrinsic properties of promoters. ACS Synth Biol 5:89–98. https://doi.org/10.1021/acssynbio.5b00116

Sainz De Murieta I, Bultelle M, Kitney RI (2016) Toward the first data acquisition standard in synthetic biology. ACS Synth Biol. https://doi.org/10.1021/acssynbio.5b00222

Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. Nat Biotechnol 27:946–950. https://doi.org/10.1038/nbt.1568

Sarpeshkar R (2014) Analog synthetic biology. Philos Trans R Soc A Math Phys Eng Sci 372:20130110. https://doi.org/10.1098/rsta.2013.0110

Sarrion-Perdigones A, Falconi EE, Zandalinas SI et al (2011) GoldenBraid: an iterative cloning system for standardized assembly of reusable genetic modules. PLoS ONE 6:e21622. https://doi.org/10.1371/journal.pone.0021622

Sauro HM (2008) Modularity defined. Mol Syst Biol 4:166. https://doi.org/10.1038/msb.2008.3

Schaerli Y, Gili M, Isalan M (2014) A split intein T7 RNA polymerase for transcriptional AND-logic. Nucleic Acids Res 42:12322–12328. https://doi.org/10.1093/nar/gku884

Schwarz-Schilling M, Aufinger L, Mückl A, Simmel FC (2016) Chemical communication between bacteria and cell-free gene expression systems within linear chains of emulsion droplets. Integr Biol 8:564–570. https://doi.org/10.1039/C5IB00301F

Segall-Shapiro TH, Sontag ED, Voigt CA (2018) Engineered promoters enable constant gene expression at any copy number in bacteria. Nat Biotechnol 36:352–358. https://doi.org/10.1038/nbt.4111

Selvin PR (2000) The renaissance of fluorescence resonance energy transfer. Nat Struct Biol 7:730–734. https://doi.org/10.1038/78948

Sharma V, Nomura Y, Yokobayashi Y (2008) Engineering complex riboswitch regulation by dual genetic selection engineering complex riboswitch regulation by dual genetic. J Chem Soc Perkin Trans I 130:16310–16315. https://doi.org/10.1021/ja805203w

Sheth RU, Yim SS, Wu FL, Wang HH (2017) Multiplex recording of cellular events over time on CRISPR biological tape. Science 0958:eaao0958. https://doi.org/10.1126/science.aao0958

Shetty RP, Endy D, Knight TF (2008) Engineering BioBrick vectors from BioBrick parts. J Biol Eng 2:1–12. https://doi.org/10.1186/1754-1611-2-5

Shih SCC, Goyal G, Kim PW et al (2015) A versatile microfluidic device for automating synthetic biology. ACS Synth Biol 4:1151–1164. https://doi.org/10.1021/acssynbio.5b00062

Shipman SL, Nivala J, Macklis JD, Church GM (2017) CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria. Nature 547:345–349. https://doi.org/10.1038/nature23017

Shishkin AA, Giannoukos G, Kucukural A et al (2015) Simultaneous generation of many RNA-seq libraries in a single reaction. Nat Methods 12:323–325. https://doi.org/10.1038/nmeth.3313

Si T, Zhao H (2016) A brief overview of synthetic biology research programs and roadmap studies in the United States. Synth Syst Biotechnol 1:258–264. https://doi.org/10.1016/j.synbio.2016.08.003

Siliconreview Team (2017) Engineering biology to shape the world around us: Zymergen. In: thesiliconreview.com. http://thesiliconreview.com/magazines/engineering-biology-to-shape-the-world-around-us-zymergen/. Accessed 24 Jun 2018

Siuti P, Yazbek J, Lu TK (2013) Synthetic circuits integrating logic and memory in living cells. Nat Biotechnol 31:448–452. https://doi.org/10.1038/nbt.2510

Sleight SC, Bartley BA, Lieviant JA, Sauro HM (2010) Designing and engineering evolutionary robust genetic circuits. J Biol Eng 4:12. https://doi.org/10.1186/1754-1611-4-12

Stanton BC, Nielsen AAK, Tamsir A, Clancy K, Peterson T, Voigt CA (2013) Genomic mining of prokaryotic repressors for orthogonal logic gates. Nat Chem Biol 10:99–105. https://doi.org/10.1038/nchembio.1411

Stanton BC, Siciliano V, Ghodasara A et al (2014) Systematic transfer of prokaryotic sensors and circuits to mammalian cells. ACS Synth Biol 3:880–891. https://doi.org/10.1021/sb5002856

Sun ZZ, Yeung E, Hayes CA, Noireaux V, Murray RM (2014) Linear DNA for rapid prototyping of synthetic biological circuits in an *Escherichia coli* based TX-TL cell-free system. ACS Synth Biol 3:387–397. https://doi.org/10.1021/sb400131a

Tabor JJ, Salis HM, Simpson ZB et al (2009) A synthetic genetic edge detection program. Cell 137:1272–1281. https://doi.org/10.1016/j.cell.2009.04.048

Tamsir A, Tabor JJ, Voigt CA (2011) Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. Nature 469:212–215. https://doi.org/10.1038/nature09565

Tang W, Liu DR (2018) Rewritable multi-event analog recording in bacterial and mammalian cells. Science 360:eaap8992. https://doi.org/10.1126/science.aap8992

Taylor ND, Garruss AS, Moretti R et al (2016) Engineering an allosteric transcription factor to respond to new ligands. Nat Methods 13:177–183. https://doi.org/10.1038/nmeth.3696

Temme K, Hill R, Segall-Shapiro TH, Moser F, Voigt CA (2012) Modular control of multiple pathways using engineered orthogonal T7 polymerases. Nucleic Acids Res 40:8773–8781. https://doi.org/10.1093/nar/gks597

Toda S, Blauch LR, Tang SKY, Morsut L, Lim WA (2018) Programming self-organizing multicellular structures with synthetic cell-cell signaling. Science 0271:eaat0271. https://doi.org/10.1126/science.aat0271

Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH (2009) Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. J R Soc Interface 6:187–202. https://doi.org/10.1098/rsif.2008.0172

Torella JP, Boehm CR, Lienert F, Chen JH, Way JC, Silver PA (2014) Rapid construction of insulated genetic circuits via synthetic sequence-guided isothermal assembly. Nucleic Acids Res 42:681–689. https://doi.org/10.1093/nar/gkt860

Tracy BP, Gaida SM, Papoutsakis ET (2010) Flow cytometry for bacteria: enabling metabolic engineering, synthetic biology and the elucidation of complex phenotypes. Curr Opin Biotechnol 21:85–99. https://doi.org/10.1016/j.copbio.2010.02.006

Tsoi R, Wu F, Zhang C, Bewick S, Karig D, You L (2018) Metabolic division of labor in microbial systems. Proc Natl Acad Sci 115:2526–2531. https://doi.org/10.1073/pnas.1716888115

Urrios A, Macia J, Manzoni R et al (2016) A synthetic multicellular memory device. ACS Synth Biol 5:862–873. https://doi.org/10.1021/acssynbio.5b00252

Villarreal F, Contreras-Llano LE, Chavez M et al (2017) Synthetic microbial consortia enable rapid assembly of pure translation machinery. Nat Chem Biol 14:1–22. https://doi.org/10.1038/nchembio.2514

Wang B, Buck M (2012) Customizing cell signaling using engineered genetic logic circuits. Trends Microbiol 20:376–384. https://doi.org/10.1016/j.tim.2012.05.001

Wang B, Buck M (2014) Rapid engineering of versatile molecular logic gates using heterologous genetic transcriptional modules. Chem Commun 50:11642–11644. https://doi.org/10.1039/C4CC05264A

Wang B, Kitney RI, Joly N, Buck M (2011) Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. Nat Commun 2:508. https://doi.org/10.1038/ncomms1516

Wang B, Barahona M, Buck M (2013a) A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals. Biosens Bioelectron 40:368–376. https://doi.org/10.1016/j.bios.2012.08.011

Wang B, Barahona M, Buck M, Schumacher J (2013b) Rewiring cell signalling through chimaeric regulatory protein engineering. Biochem Soc Trans 41:1195–1200. https://doi.org/10.1042/BST20130138

Wang B, Barahona M, Buck M (2014) Engineering modular and tunable genetic amplifiers for scaling transcriptional signals in cascaded gene networks. Nucleic Acids Res 42:9484–9492. https://doi.org/10.1093/nar/gku593

Wang B, Barahona M, Buck M (2015a) Amplification of small molecule-inducible gene expression via tuning of intracellular receptor densities. Nucleic Acids Res 43:1955–1964. https://doi.org/10.1093/nar/gku1388

Wang L, Qian K, Huang Y et al (2015b) SynBioLGDB: a resource for experimentally validated logic gates in synthetic biology. Sci Rep 5:1–5. https://doi.org/10.1038/srep08090

Wang S, Majumder S, Emery NJ, Liu AP (2018) Simultaneous monitoring of transcription and translation in mammalian cell-free expression in bulk and in cell-sized droplets. Synth Biol 3:ysy005. https://doi.org/10.1093/synbio/ysy005

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S (2011) A modular cloning system for standardized assembly of multigene constructs. PLoS ONE 6:e16765

Weinberg BH, Pham NTH, Caraballo LD et al (2017) Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. Nat Biotechnol 35:453–462. https://doi.org/10.1038/nbt.3805

Williams TC, Pretorius IS, Paulsen IT et al (2016) Synthetic evolution of metabolic productivity using biosensors. Trends Biotechnol 34:371–381. https://doi.org/10.1016/j.tibtech.2016.02.002

Woodruff LBA, Gorochowski TE, Roehner N et al (2017) Registry in a tube: multiplexed pools of retrievable parts for genetic design space exploration. Nucleic Acids Res 45:1553–1565. https://doi.org/10.1093/nar/gkw1226

Wroblewska L, Kitada T, Endo K et al (2015) Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery. Nat Biotechnol 33:839–841. https://doi.org/10.1038/nbt.3301

Wu G, Yan Q, Jones JA, Tang YJ, Fong SS, Koffas MAG (2016) Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. Trends Biotechnol 34:652–664. https://doi.org/10.1016/j.tibtech.2016.02.010

Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y (2011) Multi-input RNAi-based logic circuit for identification of specific cancer cells. Science 333:1307–1311. https://doi.org/10.1126/science.1205527

Yang S, Liu X, Zhang Y (2007) Metabolic engineering—applications, methods, and challenges. In: Bioprocessing for value-added products from renewable resources, Elsevier, pp 73–118

Yang L, Nielsen AAK, Fernandez-Rodriguez J et al (2014) Permanent genetic memory with > 1-byte capacity. Nat Methods 11:1261–1266. https://doi.org/10.1038/nmeth.3147

Yeung E, Dy AJ, Martin KB et al (2017) Biophysical constraints arising from compositional context in synthetic gene networks. Cell Syst 5(11–24):e12. https://doi.org/10.1016/j.cels.2017.06.001

Zadeh JN, Steenberg CD, Bois JS et al (2011) NUPACK: analysis and design of nucleic acid systems. J Comput Chem 32:170–173. https://doi.org/10.1002/jcc.21596

Zhang C, Tsoi R, You L (2016) Addressing biological uncertainties in engineering gene circuits. Integr Biol 8:456–464. https://doi.org/10.1039/C5IB00275C