MODELLING

Synthetic biology: understanding biological design from synthetic circuits

Shankar Mukherji and Alexander van Oudenaarden‡*

Abstract | An important aim of synthetic biology is to uncover the design principles of natural biological systems through the rational design of gene and protein circuits. Here, we highlight how the process of engineering biological systems — from synthetic promoters to the control of cell–cell interactions — has contributed to our understanding of how endogenous systems are put together and function. Synthetic biological devices allow us to grasp intuitively the ranges of behaviour generated by simple biological circuits, such as linear cascades and interlocking feedback loops, as well as to exert control over natural processes, such as gene expression and population dynamics.

Modularity
A property of a system such that it can be broken down into discrete subparts that perform specific tasks independently of the other subparts.

Bioremediation
The treatment of pollution with microorganisms.

**Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.*
‡Department of Physics and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
Correspondence to A.v.O.
e-mail: avano@mit.edu
doi:10.1038/nrg2697
Published online
10 November 2009

One of the most astounding findings of the Human Genome Project was that our genome contains as many genes as that of *Drosophila melanogaster*. This finding begged the question: how do you get one organism to look like a fly and another like a human with the same number of genes? One possibility is that the rich repertoire of non-protein-coding sequences found in the genomes of complex organisms adds many new parts with which to generate complexity[1]. However, a decade of research has put forward the rather different idea that instead of looking at the length of the parts list as the determinant of organismal complexity, we should look at how those parts fit together[2,3]. From this perspective, complexity arises from novel combinations of pre-existing proteins, and the ability to evolve new phenotypes rests on the modularity of biological parts.

In addition to natural examples of modularity[3], strong evidence to support this post-genomic view of biology has come from the synthesis of new biological systems. Rational synthesis of biological systems can hint at the natural history of how a particular system came to acquire its properties[4,5]. More often, however, we use synthetic circuits to explore, in a hands-on fashion, the set of design principles that determine the structure and operation of biological systems.

The core aim of synthetic biology is to develop and apply engineering tools to control cellular behaviour by using precisely characterized parts, such as *cis*-regulatory elements, to achieve desired functions. An important direction, for example, has been to engineer cells with practical applications in the areas of bioremediation[6], biosensing[7] and biofuel production[8,9], or even with potential clinical applications[10–12]. In this Review, however, we focus on how synthetic circuits help us to understand how natural biological systems are genetically assembled and how they operate in organisms from microbes to mammalian cells. In this light, synthetic circuits have been crucial as simplified test beds in which to refine our ideas of how similarly structured natural networks function, and they have served as tools for controlling natural networks. We highlight the contribution of synthetic biology to the generation of increasingly quantitative descriptions of gene expression and signal transduction, to uncovering the diversity of behaviours that can arise from positive and negative feedback systems, and to advances in the rational control of spatial organization and cell–cell interactions. We pay particular attention to recent progress in using synthetic systems to uncover novel aspects of cell biology, such as how cells decide to undergo apoptosis and the molecular basis for communication between the endoplasmic reticulum and mitochondria. We aim to show that synthetic biological approaches have given us many insights into how the simple building blocks that underlie complex natural systems work, in addition to basic tools with which to quantitatively characterize natural phenomena, both of which are crucial for the field to progress towards the analysis and complete control of natural circuits.

Quantitative descriptions of gene expression

The first step in assembling a biological circuit is to gather the component parts. In cells, circuits are accomplished by gene expression, and so a great deal of effort

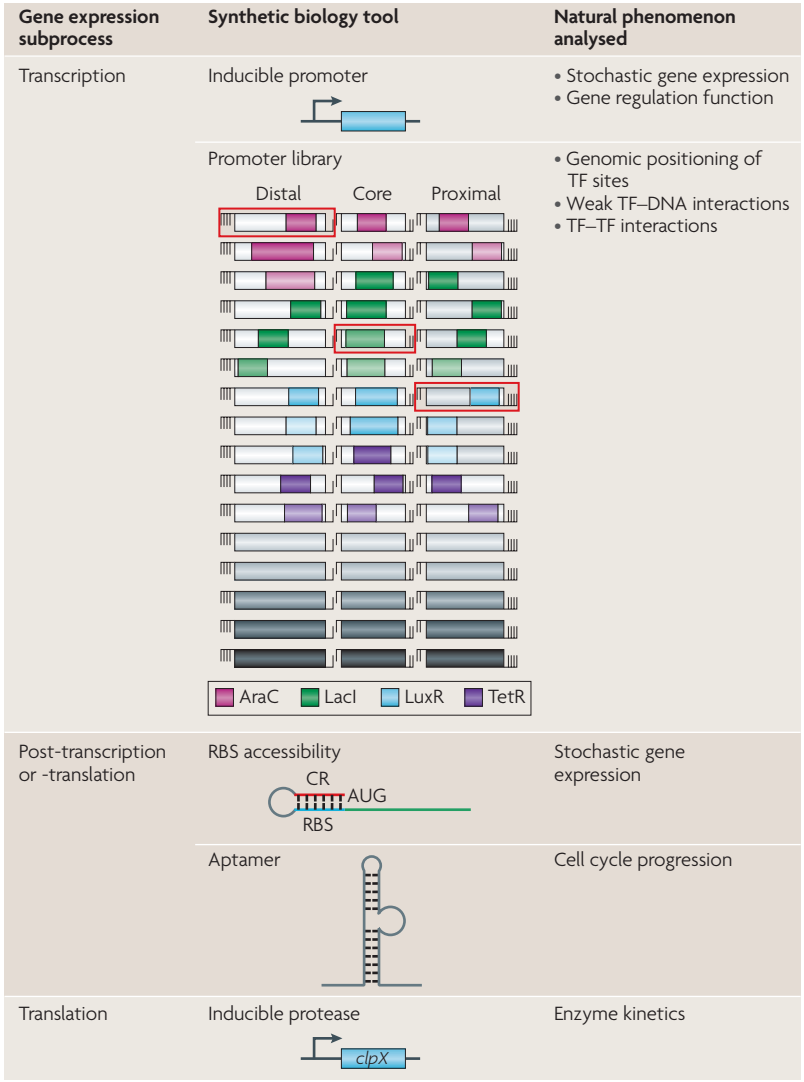


Figure 1 | **Controlling the flow of information from DNA to proteins using synthetic elements.** The diagram shows the transcriptional and post-transcriptional processes in gene expression that can be manipulated by synthetic biology tools, with some example applications. The differences in shading reflect variations in the strength of the input from the four regulators (for example, dark pink represents strong input, and light pink represents weak input). CR, complementary region to the RBS; RBS, ribosome binding site; TF, transcription factor. Promoter library diagram is reproduced, with permission, from REF. 24 © (2007) Macmillan Publishers Ltd. All rights reserved. RBS accessibility diagram is reproduced, with permission, from REF. 38 © (2004) Macmillan Publishers Ltd. All rights reserved. Aptamer diagram is reproduced, with permission, from REF. 34 © (2001) Elsevier.

in synthetic biology has gone into investigating the rules surrounding the expression of genes, particularly the processes of transcription and translation. The precise measurements afforded by artificially constructed systems allow us to transform qualitative notions of transcriptional repression, transcriptional activation and post-transcriptional regulation into quantifiable effects — such as the precise relationship between promoter architecture and the rate of transcription, and the exact degradation rate specified by a given sequence motif.

Motif
A subcircuit that is embedded in a larger network and that is found to be statistically overrepresented in that larger network when compared with a random network with similar graphical properties.

Transcriptional regulation. The earliest contributions of synthetic biology to understanding natural biological processes include detailed, quantitative measurements of transcriptional regulation, which build on a foundation laid 50 years ago in the groundbreaking work of researchers such as Jacob and Monod[13]. Synthetic constructs have been used to map out the transfer function that relates the input concentrations of transcription factors (TFs)[14,15] and inducers[16] to the output concentrations of reporter genes[14,17,18], single mRNA molecules[19,20] or single proteins[21]. Many of these constructs have also been used to measure the mean output of the transcriptional process and the higher-order moments (such as the variance) in organisms ranging from *Escherichia coli* and *Bacillus subtilis* to mammalian cells. Single-molecule studies in these model organisms have directly established that mRNA and proteins are produced in bursts of activity[22].

A key question in the study of transcriptional regulation is how the architecture of promoters affects transcriptional activity. For example, below we describe several studies that have shown how the number and genomic positions of TF binding sites affect transcriptional activity. Given the combinatorial control of gene expression, it is also crucial to study how multiple TFs interact with DNA and with each other to tune mRNA production. Endogenous promoters use all of these parameters to specify either a desired transcription rate or a Boolean function, such as an AND gate that allows transcription to occur only when all TF binding sites in the promoter are occupied.

Promoter library studies in bacteria and eukaryotes. The experimental breakthrough that allowed quantitative measurements of the transcriptional power of different promoter architectures was the use of combinatorial promoter libraries[23]. Libraries of promoters that drive reporter proteins, such as luciferase or fluorescent proteins, allow for an unbiased measurement of transcriptional activity over the space of possible promoters — such an unbiased method can be used to ascertain rules that describe the responsiveness of a promoter to TFs. Earlier work used randomly mutated promoters to draw inferences about the functional subparts of the promoter, such as the TATA box; by contrast, the construction of combinatorial promoter libraries involves identifying specific operator sites that bind TFs and randomly ligating them together in a way that shuffles their relative positions and copy numbers (FIG. 1). The studies highlighted below have combined such promoter libraries and modelling to show that the strength of a promoter is determined largely by the position of TF binding sites with respect to key promoter elements, such as the TATA box, and with respect to each other.

The simplest case is to understand how the positioning of a single operator affects the expression of a promoter. In bacteria, operators are classified as being in the core, proximal or distal regions of the promoter (FIG. 1). Working in *E. coli*, Cox *et al.*[24] and Kinkhabwala and Guet[25] independently observed that repressors can effectively repress expression from all three promoter

subregions. Cox *et al.* showed that the strength of repression is greatest when the repressor site is in the core region of the promoter, less strong when in the proximal region and weakest when in the distal region. Conversely, activators work only in the distal site and have no effect in the core and proximal sites[24]. Both studies go on to develop simple models of promoter activity by taking into account the binding reactions of TFs to DNA in thermodynamic equilibrium.

It was expected that the situation would be far more subtle in eukaryotes, in which chromatin structure can strongly influence expression levels[26]. However, even in *Saccharomyces cerevisiae*, 49% of the variation in expression in the promoter library could be explained by a simple thermodynamic model that incorporated just TF–DNA and TF–TF interactions[27] — interactions that were also suggested in theoretical work[28]. More surprisingly, Gertz *et al.* provided evidence that weak binding sites, which are important for bacterial transcription, can also be important in eukaryotes. Focusing on the TF multicopy inhibitor of Gal1 (Mig1), Gertz *et al.* showed that repression from one weak and one strong Mig1 binding site can be as effective as two strong Mig1 binding sites. This is particularly crucial given that 24% of all yeast promoters contain putative weak Mig1 binding sites.

The promoter library studies open the way to considering some general questions in transcriptional control. For example, the theoretical frameworks in the *E. coli* and yeast studies differ slightly: the *E. coli* studies do not require TF–TF interactions and frame the issue mostly in the language of Boolean logic, whereas the yeast studies make heavy use of TF–TF interactions, particularly in the analysis of weak binding sites. Future single-molecule studies of transcriptional control can help to resolve the relative importance of TF–DNA and TF–TF interactions in generating transcriptional activity. Furthermore, the fact that simple equilibrium binding explains much, but not all, of the effect of promoter architecture on expression level suggests that the next goal should be to track down the source of the remaining variation. Genomic location can make an important contribution to expression and expression fluctuations[29], perhaps by affecting local chromatin context. Knowing how to apportion the variation to these different effects will be particularly helpful when these studies are extended to mammalian systems, in which there is considerably less control over where synthetic transgene constructs are integrated into the genome.

Post-transcriptional and post-translational regulation. Although much of the early work in synthetic biology focused on transcriptional regulation, substantial progress has also been made in incorporating post-transcriptional effects on RNA and proteins into synthetic circuits. At the RNA level, for example, mutagenesis screens based on synthetic constructs have been used to determine the sequences that are recognized by RNA-editing enzymes to change adenine into inosine[30]. Furthermore, as regulatory RNAs have been increasingly recognised as important drivers of gene expression, synthetic circuits have included elements from the RNA interference pathway[31], aptamers[32–34] and riboswitches[35,36] to control the flow of genetic information[37].

Synthetic circuits that involve enzymatic RNAs have mostly been developed as platforms for tuning gene expression, but many of these platforms can easily be extended to understand natural biological phenomena. In the study by Grate and Wilson, for example, an aptamer is used to control the expression of cyclin B2 (Clb2), a key regulator of the cell cycle, in a tetramethylrosamine (Tmr)-dependent manner[34]. The authors slowed the speed of the cell cycle by adding Tmr; this method can be useful for measuring how the level of Clb2 affects the speed at which the cell cycle progresses while keeping all transcriptional feedback constant.

Synthetic studies have also directly modified how mRNA is translated into proteins and how long proteins persist before being degraded. Several experiments in bacterial systems, especially those studying the stochastic nature of gene expression, have altered the translation rate by mutating ribosomal binding sites (RBSs)[17,38]. Apart from showing another possible layer of quantitative regulation of gene expression, studies involving RBS variants provided early evidence that *E. coli* cells could tune the stochasticity in the expression level of a given gene independently of its mean. Lastly, Grilly *et al.* have developed a circuit that controls the degradation of a target protein using the ClpXP protease machinery from *E. coli*[39]. Typically, models of gene expression treat protein degradation as an exponential decay process, with the decay being due to the growth of cell volume over time. Regulated proteolysis, however, can depend on the formation of enzyme–substrate complexes as intermediates on the way to degradation. In finding that the degradation follows Michaelis–Menten kinetics, Grilly *et al.* made one of the few quantitative comparisons of specific protease activity with models of enzyme kinetics.

Taken together, these results point to some interesting similarities between transcription and translation — both are inherently noisy processes that can be quantitatively modulated by specific sequence elements, such as RBSs and protease-recognition sites. Future studies can use the ideas and methods from the study of transcription, such as combinatorial library approaches, to systematically explore the process of translation.

Integrating transcriptional and post-transcriptional control. The two approaches described above — using a natural inducible promoter or engineering specific promoter architectures to tune transcriptional activity, and using specific sequence sites to tune translational yield — can be combined to achieve precise and flexible control over gene expression[17,31]. An example of this combined approach for studying natural processes in mammalian cells can be found in recent work in which Tet- and Lac-controlled regulation was adapted and combined with RNAi for use in HeLa cells[40] (FIG. 2).

As synthetic biology begins to create more realistic systems that contain many moving parts, demand will increase for circuits that control every step of the process

Transfer function
A mathematical or graphical representation of the relationship between the input and output of a system.

Higher-order moment
For a probability distribution, a number that characterizes the shape of the distribution, as opposed to the mean.

Variance
The second-order moment of a probability distribution; it characterizes the width of the distribution.

Boolean function
A special class of transfer function that takes binary values as inputs, performs a logical operation and yields binary values as outputs.

Combinatorial promoter library
A collection of promoters that is constructed by randomly ligating together promoter subregions, such as the sequence between −35 and −10 from the start codon, taken from different promoters. Such random ligation of subregions allows for the combinatorial generation of novel promoters from a small number of parts.

Aptamer
A short nucleic acid or peptide sequence that specifically binds to a target molecule.

Riboswitch
A segment of an mRNA molecule that specifically binds a target molecule; riboswitches are closely related to aptamers.

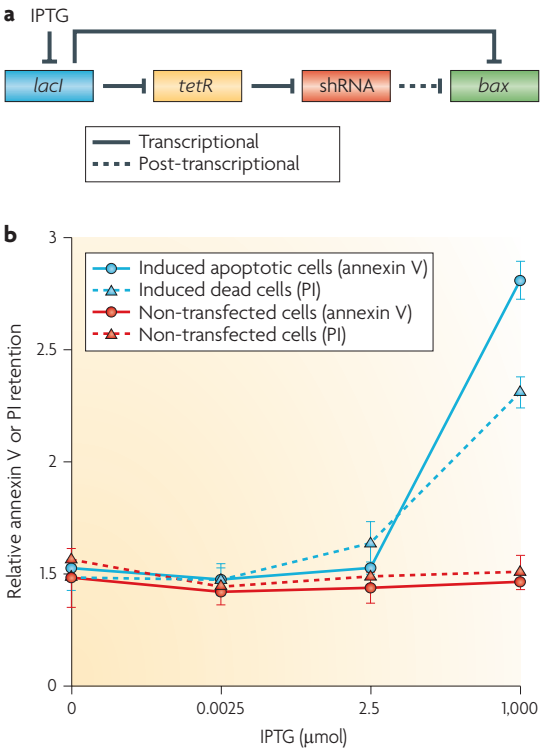


Figure 2 | **An integrated transcription and translation circuit for controlling gene expression in mammalian systems. a** | The output of the genetic switch created by Deans *et al.*[40] can be monitored by a *GFP* reporter or a gene of interest; here, the gene of interest that we focus on is *bax*, a pro-apoptotic gene. *bax* is under the transcriptional control of the Lac repressor (LacI) and under the translational control of a short hairpin RNA (shRNA), which itself is under transcriptional control of the TetR repressor. In the 'off' state, LacI inhibits transcription of *bax*. Additionally, LacI inhibits transcription of the *tetR* repressor; this allows the transcription of the shRNA, which goes on to inhibit translation of *bax* by cleaving its mRNA. The result of this dual-layered repression is the creation of an 'off' state with complete repression; in the initial characterization, each mode of repression alone reduced reporter levels by about 80%, leaving a basal expression of 20%, but the combination resulted in >99% repression. The circuit can be tunably activated by adding varying amounts of isopropyl-β-D-thio-galactoside (IPTG), which blocks the effects of LacI. **b** | The fraction of cells that undergo apoptosis is determined by the Bax expression levels. Data obtained by tuning Bax with IPTG, as described above, offer some tantalizing clues as to the fundamental molecular biology underlying the apoptosis pathway. In particular, the data are consistent with the idea that the decision to undergo apoptosis (assessed by the retention of propidium iodide (PI) dye (which stains dead cells) relative to the retention of annexin V (which stains apoptotic cells) compared with the transfection protocol alone) is determined by reaching a threshold level of Bax. Although the Bax threshold data are not conclusive, the result shows the power of this technique, which allows the rational tuning of gene expression and the examination of the consequences. Part **b** is reproduced, with permission, from REF. 40 © (2007) Elsevier.

that turns DNA sequences into proteins. Such layered circuits can help to show why certain regulatory schemes are used over others for controlling gene expression in a given context. For example, gene expression in natural systems can be attenuated by epigenetic silencing, transcriptional repressors or post-transcriptional regulators, such as microRNAs (either alone or in concert with other molecules); this leads to the questions of why a system uses one system rather than the other, and to what extent different layers of regulation generate collective effects that no one layer can accomplish. One area that will be increasingly under study, and that may help to unravel the issues surrounding layered circuits, is the dynamics of the different steps that contribute to expression; the studies highlighted above almost exclusively focus their attention on steady-state behaviour. Although intuition tells us that TFs act slowly compared with post-transcriptional elements, such as regulatory RNAs (which presumably do not have to be transported back to the nucleus and then locate a specific genomic locus), there is currently a lack of data that would enable us to turn these intuitive notions into quantitative facts.

Rewiring genetic and signalling pathways

Engineering cellular pathways has allowed insights into two key properties: the precise measurement and control of the input–output relationship of a pathway, and the functional architecture of the pathway constituents themselves. In particular, engineering signalling pathways has provided insights into the functional significance of specific protein sequences and structures by showing exactly which protein domains and which amino acid residues are responsible for mediating specific interactions along the pathway.

The challenges of rewiring pathways. Initially, pathway engineering was primarily explored in the context of metabolism[41]. Metabolic engineering typically involved the use of genetic screens and directed evolution to maximize targeted metabolic fluxes. Synthetic efforts in boosting metabolic fluxes have begun to pay off, as shown in a recent study in which a synthetic protein scaffold was used to draw metabolic enzymes spatially closer to each other[42] — however, it should be noted that this study did not involve any pathway rewiring. By contrast, the rational rewiring of pathways involves specific manipulations of the components of the system to achieve a desired outcome. The most crucial aspect of protein and gene structure that synthetic biologists use to rewire pathways is the inherent modularity of many proteins[43] (signalling proteins, for example, typically have dedicated domains for recognizing binding partners that act independently of other functional domains). Most rewiring studies therefore focus on signal transduction and genetic cascades (BOX 1). There are fewer examples of achieving metabolic control through specifically designed changes in protein sequence[44,45]. Changes in the structure of an allosteric site in a metabolic enzyme are more prone to alter the active site than changes in the allosteric site in a signalling protein[46].

Basal expression
The level of transcription that occurs in the absence of an inducer.

Directed evolution
A cyclic sequence of steps, including modification, selection and amplification. It is used, typically *in vitro*, to enrich for proteins or nucleic acids that show properties that are desired by the researcher but that are not necessarily found in nature.

Metabolic flux
The rate of turnover of metabolites in a metabolic pathway.

Allosteric site
A region of an enzyme that is physically distinct from the active site and that can induce conformational changes, usually by binding small molecules, to affect the accessibility or efficiency of the active site.

Box 1 | **Rationally rewiring the input–output relationship of signalling pathways**

As the bottom left of the figure shows, membrane proteins (light blue) can be engineered to have sensors (green) and can be made to interact with adaptors (red), which can in turn be made to interact with other adaptors (dark blue). More formally, the input–output relationship can be controlled in two ways: by changing the stimulus that a receptor is triggered by (shown in part **a** of the figure) or by changing the transducing molecules that the receptor uses to pass the information from the environment to the cellular interior (shown in parts **b** and **c**).

Chimeric receptors

Chimeric receptors are an example of the first type of change. As shown in part **a**, a chimeric receptor can be made so that, for example, a light signal triggers the salt signal transducer OmpR (the 'rewiring' is depicted as a red arrow) instead of its normal transducer Rcp1 (wild-type interactions are depicted as black arrows). Indeed, in principle, chimeric receptors can be made so that any stimulus triggers any response regulator (other possibilities are shown as grey arrows). Although chimeric receptors have been used previously[104], chimeric photoreceptors in particular are exceptional for allowing much higher sensitivity measurements and for avoiding crosstalk effects.

In the case of *Escherichia coli*, the rewiring is accomplished by transcriptionally fusing the cytosolic signal transduction domain of the pathway sensor — the histidine kinase domain of EnvZ — to cyanobacterial phytochrome 1 (Cph1), resulting in a system in which the response regulator of EnvZ, OmpR, can be triggered by light[50]. The pathway activity is measured by placing the *lacZ* gene, the product of which creates a black compound, under the control of the OmpR-depdendent *ompC* promoter.

The response to a light gradient input serves as a precise measurement of the transfer function of the pathway (example in part **a**). The transfer curve seems to indicate that the pathway operates in a threshold linear manner, although whether this is due to the phytochrome sensor rather than the pathway needs to be explored. Such thresholding could serve to protect the cell from overreacting to small signals.

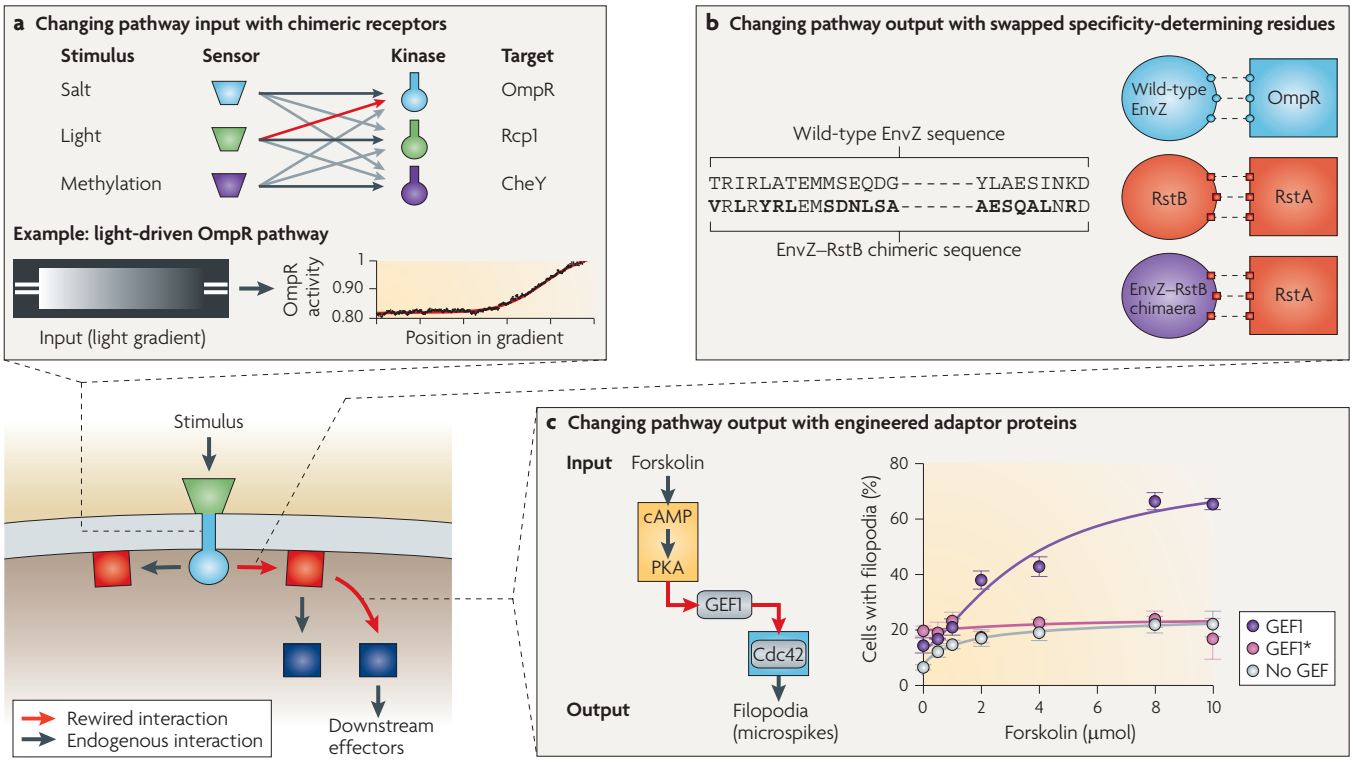
Shimizu-Sato *et al.* applied similar principles in yeast, but instead fused a galactose 4 (Gal4)-binding domain (GBD) to the red-light-absorbing phytochrome Pr and a Gal-activating domain (GAD) to the the binding partner of Pr, phytochrome interacting factor 3 (PIF3), therefore bypassing the Gal signalling cascade[51]. Any gene of interest can therefore be controlled by placing it under the control the *gal1* promoter and exposing the cells to red light instead of Gal.

Swapping specificity-determining residues

Once activated, the signal from the sensor must be specifically transduced to affect specific downstream processes. By studying covariance among residues from interacting proteins, statistical scores, such as mutual information, can be used to predict which residues determine the specificity of the interaction. As shown in part **b**, specificity-determining residues from the protein RstB (shown in bold) were substituted into EnvZ, resulting in an EnvZ–RstB chimeric protein. As a result, phosphotransfer occurred between EnvZ–RstB and RstA rather than OmpR (which is the normal partner of EnvZ)[63].

Engineering adaptor proteins

Finally, a great deal of signal processing takes place in between the triggering of a sensor by the environment and the output of the pathway, especially in eukaryotes. One major intermediate in eukaryotes is the class of proteins known as guanine exchange factors (GEFs), which control morphological pathways. Yeh *et al.* swapped wild-type GEFs that control the formation of filopodia and lamellipodia for synthetic GEFs that could be induced by the small molecule forskolin and that could generate novel morphological outputs[67] (part **c**). Specifically, they substituted an autoinhibitory domain in the wild-type GEFs with a protein kinase A (PKA)-responsive inhibitory domain, PDZ. Placing an endogenous pathway under tunable control allows crucial aspects of cell biology to be characterized in quantitative detail. Interestingly, Yeh *et al.* found that the morphological output was only manifested probabilistically — it is the fraction of cells that display either filopodia (shown here) or lamellopodia (not shown) that increases with increasing forskolin. GEF1*, synthetic GEF1. Part **a**, lower panel, is reproduced, with permission, from REF. 50 © (2005) Macmillan Publishers Ltd. All rights reserved. Part **b** is modified, with permission, from REF. 63 © (2008) Elsevier. Part **c** is reproduced, with permission, from REF. 67 © (2007) Macmillan Publishers Ltd. All rights reserved.



This property allows for the regulation of metabolic fluxes by effects such as allostery, but the relative lack of modularity also makes it difficult to forward engineer new behaviours by altering one domain but leaving the others unchanged.

Even in signalling systems, researchers are presented with severe challenges. Among the major limitations in understanding the signal propagation characteristics of many pathways is confusion over what cue triggers the cascade and whether the cue affects other processes that are taking place in the cell. For example, many organisms have dedicated signalling systems for relaying information about an osmotic shock to the cell, but the presence of abundant osmolyte will affect numerous processes besides signalling, such as global TF binding[47]. The examples described below show how techniques that both specifically and sensitively activate a selected cascade can allow researchers to focus on pathway behaviour independently of such off-target effects.

Manipulating the sensors. One of the most direct ways of rewiring the input–output relationship of a pathway is by directly changing the cue that the pathway sensor responds to. If the cue is chosen so that its level can be directly modulated, the pathway transfer functions can be measured in a similar way to the method described above for promoters. For example, Armbruster *et al.* generated a G protein-coupled receptor that responded to a pharmacologically inert compound, which could then be titrated to measure the pathway response[48], and Anderson *et al.* engineered sensors that can detect changes in tumour-related microenvironments[12]. Alternatively, the ligands that drive pathway activity can be manipulated: Cironi *et al.* linked together epidermal growth factor (EGF) and mutated forms of interferon-α 2a (IFNα2a) so that the only cells that could correctly respond to the IFNα2a signal were those that co-expressed the EGF receptor[49].

A particularly striking example of how sensor rewiring can shed light on the operation of a cascade *in vivo* in a sensitive and specific manner can be found in the use of chimeric photoreceptors (BOX 1). Two studies used light itself as the cue to drive a signalling system[50,51]; this approach is unlike traditional implementations of light-driven systems[52–54], such as those that use light to activate a small molecule, which then activates a desired biological process[55]. Levskaya *et al.* engineered the *E. coli* EnvZ–OmpR two-component system to respond to light, whereas Shimizu-Sato *et al.* fused a phytochrome and its binding partner to selected pathway proteins in the *S. cerevisiae* galactose pathway. Levskaya *et al.* proceeded to map out the input–output transfer function with high precision by exposing a lawn of rewired bacteria to a light gradient. The measurements of transfer function in the bacteria suggested that a threshold level of the environmental cue is needed before pathway activity is triggered. Although careful titration of an osmolyte would have allowed precise measurement of the transfer function, such as through the use of microfluidic devices[56], matching the sensitivity of a simple light gradient is difficult to accomplish. Furthermore, matching the specificity afforded by using light to drive pathway activity is probably impossible. However, given the ease with which we can deliver precisely controlled light signals to cells compared with delivering chemical signals, the Levskaya *et al.* and Shimizu-Sato *et al.* studies can be easily extended to tasks such as measuring Bode plots, as was recently done for the yeast osmo-response system[57,58].

Manipulating sensor–transducer interactions. Swapping the sensor in a signalling pathway is a way to engineer the input side of the input–output relationship, whereas changing the identity of the molecules that carry the signal from the sensor to downstream effectors can affect the output side. In fact, given the high degree of sequence homology between many sensor–transducer pairs, there is great interest in developing a detailed description of sensor–transducer interactions to understand the multiple ways in which pathways prevent crosstalk[59] — for example, by using scaffold proteins[60], mutual inhibition[61] and kinetic insulation[62].

This is the basic strategy that was followed by Skerker *et al.* to rewire the EnvZ–OmpR system[63]. This study made heavy use of the large amount of sequence data available for two-component systems to computationally detect individual amino acid residues that co-vary between cognate pairs. Specifically, they calculated the mutual information between all possible pairs of residues from sensors and response regulators and found the pairs that maximized mutual information. These pairs were thought to be the specificity-determining residues. Remarkably, they then substituted a given sensor's specificity-determining residues for a different sensor's specificity-determining residues, keeping all other residues intact, and thereby activated the latter sensor's pathway with the former sensor's trigger. Furthermore, they performed the same rewiring feat by substituting specificity-determining residues in the response regulator (BOX 1).

For now, the relative paucity of sequence data precludes the use of this technique for other systems, such as eukaryotic homologues of two-component systems. Nevertheless, this study provides a framework that goes beyond crude domain-level protein engineering all the way to molecular details. A particularly enticing possibility is the unification of a bioinformatically guided rewiring approach with data on crystal structure, especially structures of protein–protein complexes. Using a crystal structure of a complex made up of proteins similar to EnvZ and OmpR, Skerker *et al.* showed that the specificity-determining residues for the sensor kinase and the response regulator probably occur at the interface of the two proteins, which suggests that the co-evolving residues interact physically rather than allosterically. Combining structural and rewired pathway data can indicate how to explore further the numerous systems in which docking site interactions have been identified[64,65]. Synthetic pathways and crystallography together can be key in unravelling the fundamental biophysical interactions that underlie signal transduction.

Osmotic shock
A sudden change of the osmotic pressure gradient generated by the balance of the concentration of dissolved molecules inside and outside the cell.

Two-component system
The dominant architecture of environmental signal transduction systems in bacteria. It consists of a sensor kinase that transforms the environmental signal to a phosphate signal, and a cognate response regulator that further transmits the signal to the ultimate effector molecules.

Microfluidic device
A device in which fluids are conveyed to samples in channels with diameters in the order of 1 μm; these chambers can be used to precisely and dynamically control the microenvironment to which cells are exposed.

Bode plot
A special class of transfer function that relates the frequency of the input, such as a stimulus that triggers a signalling cascade, to the output of the system, such as the amplitude of the response.

Scaffold protein
An element of a signal transduction pathway that simultaneously binds multiple members of the pathway. Scaffold proteins increase the local concentrations of pathway proteins and therefore increase the probability of them interacting.

Mutual inhibition
A network architecture that consists of two interacting pathways in which the output of each pathway inhibits the activity of the other pathway.

Kinetic insulation
A mechanism in which a signal is transduced through a particular pathway based on the temporal profile of the signal; for example, a transient signal can be interpreted by the cell as using one particular pathway, whereas a slowly varying signal can be interpreted as using a different pathway.

Manipulating the intermediate transducers. Altering the way in which a sensor interacts with its environmental cues and its immediate downstream signalling partner is the most obvious way to manipulate signal transduction. The next most obvious way is to follow the signal and tackle the intermediate transducers in the pathway. For example, Howard *et al.* took the pro-apoptotic Fadd death domain and fused it to Grb2 and ShcA, which are members of the receptor tyrosine kinase (RTK) pathway; as a result, RTK-triggered signals could be used to drive apoptosis[66].

At the adaptor level, one key target for pathway engineering is the family of guanine nucleotide exchange factors (GEFs) that regulate the actin cytoskeleton through the Rho family of GTPases[67]. Yeh *et al.* exploited the presence of an autoinhibitory domain in GEFs that can be swapped for an inhibitory domain that is under the control of a small molecule. The authors swapped the wild-type GEFs that control the formation of filopodia and lamellipodia for synthetic GEFs that could be induced by the small molecule forskolin (BOX 1). They then 'daisy-chained' two GEFs in series and showed that the combined, and therefore longer, GEF system was more sensitive to inducers and displayed a sharper separation between 'on' and 'off' states. These results are in line with previous synthetic studies that examined the sensitivity and sharpness of transcriptional cascades as the cascade length was varied[68]. As seen above in the case of apoptosis in the RNAi switch, placing an endogenous pathway — morphological in this case — under tunable control allows crucial aspects of cell biology to be characterized in quantitative detail.

Connecting pathway rewiring to evolvability. Another interesting and complementary theme that emerges from rewiring studies is how differently rewired circuits can yield the same output. The library of combinatorially synthesized gene networks constructed by Guet *et al.* contains instances of systems that have different connectivity properties but the same Boolean truth table, and those that have the same connectivity properties but different Boolean truth tables[69]. Along these lines, Isalan *et al.* showed that randomly rewiring the transcriptional network of *E. coli* results in growth defects in only 5% of the rewirings, a level of tolerance that is difficult to replicate in man-made systems[70]. The phenomenon of a circuit being rewired but having its logic maintained seems to have occurred in the evolution of the mating type switch in yeast — alpha genes in *Candida albicans* activate the alpha mating type, whereas alpha genes in *S. cerevisiae* repress the alpha mating type[71]. Theoretical studies on the evolvability of biochemical networks suggest that networks that are wired differently but produce the same output constitute a 'neutral space', which allows flexibility in the design of networks and therefore eases the way for phenotypic changes to take place[72,73]. Continuing on the theme of using rewired pathways to highlight system flexibility, Antunes *et al.* transplanted a bacterial two-component system from *E. coli* into the eukaryotic plant *Arabidopsis thaliana*. The bacterial transcriptional activator managed to cross into the nucleus to drive gene expression, which has fuelled speculation that pathway evolution can be driven by horizontal gene transfer between organisms from different kingdoms[74].

Synthetic feedback networks

Synthesis has uncovered several rules that govern how DNA is turned into proteins and how proteins interact to generate diverse phenotypes without the need for a combinatorial explosion in the number of genes. However, in the examples considered above, the flow of information is largely an ordered sequence of events: diverse outcomes in these systems resulted from combinatorial rearrangements of modular parts. Furthermore, the complexity of naturally occurring cellular networks is often dominated by feedback and feedforward loops. By incorporating these features, synthetic circuits have taught us about the dynamics and systems-level functions of more complex molecular interactions.

Initial work in this area primarily focused on the identification[75] and experimental characterization of simple motifs that occur frequently in genetic and signalling networks. In this first generation of synthetic biology studies, the mimicking of natural systems has confirmed theoretical expectations that positive feedback systems can be bistable[76–79] and that negative feedback systems are noise resistant[80] and can speed up circuit dynamics[81]. More recently, engineered feedback loops have been extended to signalling and metabolic systems: the generation of novel protein–protein and genetic interactions has allowed the exploration of how signalling pathways set their sensitivity to inputs and how they tune their kinetics[82,83]. One concrete way in which synthetic circuits are helping us to approach more complicated interaction networks is by serving as benchmarks against which theoretical and computational tools can be tested[84,85] (BOX 2).

Oscillatory behaviours. Here, we give an example of how the above techniques can be used to arrange biological parts into a biologically relevant dynamic system: an oscillator (BOX 3). Cells display a range of oscillatory behaviours. Some oscillators have tunable periods, such as the dependence of the cell cycle period on the nutrient levels available, whereas others are more robust to changes in parameters, such as the circadian oscillator. Examples include oscillatory signalling from nuclear factor-κB, which oscillates to control gene expression[86], and the p53–murine double minute 2 negative feedback loop, which oscillates to drive the DNA damage response[87].

How can a robust but tunable oscillator be constructed in a living cell? The construction of *in vivo* oscillators provides an example of how the interplay between the analysis of naturally occurring systems, modelling and the construction of synthetic systems can yield insights into biological phenomena. The story began with the observation that the simplest oscillator design, a delayed negative feedback loop, could not sustain oscillations beyond a small number of periods when operating in a cell. Instead, naturally occurring

Boolean truth table
The table of inputs and outputs that specifies a certain Boolean function.

Bistable
A property of a dynamical system in which two discrete states of the system are stable; in a biological setting, bistability implies that a system will persist in a given state even if the stimulus that drove it to that state is removed.

Box 2 | **Synthetic circuits aid the modelling of biological systems**

One of the most important functions of synthetic circuits is in building and refining analytic and computational models of biological systems. When modelling a gene or protein circuit, a series of choices must be made. The first choice is the scale at which the input–output relationship is to be measured — typically this choice boils down to whether one wants to view the system as a Boolean logic operator or a dynamical system. The dynamical system framework can be further broken down along two dimensions, depending on whether spatial or stochastic effects need to be taken into account. Spatial effects can usually be ignored when the biochemical reactions that make up the system occur on timescales that are slower than the time it takes to mix the reactants by diffusion. Stochastic effects can usually be ignored if the dynamical variables of the system can be represented as continuous rather than discrete entities — that is, when we are interested in the concentrations of a molecule rather than the number of molecules. Synthetic circuits have been used to explore all of these issues in some detail.

Until recently, the choice of modelling methodology was based on a best guess of which effects were important to include, along with post-hoc comparison of the model with data. Detailed comparisons of different modelling paradigms have been lacking. Cantone *et al.*[84] and Ellis *et al.*[85] have offered the field some guidance through the introduction of benchmark networks — that is, networks that have defined topologies and that interact only minimally with endogenous systems — against which to test proposed modelling methods. In particular, Cantone *et al.* created a relatively sophisticated synthetic transcriptional network of five genes that serves as an 'oracle' that is queried by different perturbations (such as overexpression of the network genes and induction by transcriptional inducers). Finally, they tested methods based on ordinary differential equations, Bayesian inference and information theory to uncover the connectivity of the network; they found that differential equation and Bayesian inference approaches were better at uncovering the functional relationships than the information theory-based approach, as expected for such a small network. Cantone *et al.* therefore provide an example of how synthetic circuits can be helpful in refining our understanding of large-scale biological systems by improving the algorithms we use to analyse genomic and proteomic data sets.

oscillators hinted at the crucial role of interlocking positive feedback in maintaining a robust oscillator, and this system was used in the genetic oscillators recently synthesized by Stricker *et al.*[88] and Tigges *et al.*[89] (BOX 3).

As the studies in BOX 3 show, oscillators are among the simplest *in vivo* systems that can be used to understand interactions between different types of feedback loop. Simple motifs are broadly understood by systems biologists, but the interactions between these motifs are not. It is worth considering that even for interlocking positive and negative feedback loops, multiple behaviours are possible when the parameters of the system are varied and stochastic effects are included. For example, in the yeast galactose-utilization pathway, the negative feedback loop effectively counteracts the positive feedback loop and limits the parameter space over which the system is bistable[90]. Beyond two or three loops, however, we are usually at a loss to describe the system — especially a natural one that may contain even more interactions than are being accounted for. Synthetic circuits are helping us to systematically understand how motifs interact to generate ever-richer behaviour.

Bayesian inference
A method in which observations are used to calculate the probability that a particular hypothesis about the data is true, such as whether two genes in a network interact.

The ultimate goal: spatiotemporal control

If there is one context in which all of the various biological processes that are tackled by synthetic biologists come together it is in the engineering of spatiotemporal interactions, both intracellular and intercellular. Engineering cell–cell interactions in a rational manner requires the manipulation of communication devices (signalling pathways), the specification of the desired transcriptional responses and their strength using promoters, and the arrangement of these elements in a circuit architecture that robustly encodes the desired function. If we hope to systematically improve our understanding of the functional compartments of the cell, cell development and ecology, it is imperative that we integrate the lessons learned from diverse areas of synthetic biology.

Uncovering intra- and intercellular processes in an organism. Perhaps the most striking feature of the eukaryotic cell is its organization into functional subcompartments, including the nucleus for genetic material, mitochondria for respiration and the endoplasmic reticulum for protein production. For the eukaryotic cell to accomplish its tasks, the behaviour of these compartments must be coordinated in space and time. A recent study in *S. cerevisiae* has yielded new insights into how mitochondria and the endoplasmic reticulum communicate by using a genetic screen coupled with a synthetic construct that is designed to specifically tether the two organelles[91]. Kornmann *et al.* found that the synthetic tether complemented mutations in maintenance of mitochondrial morphology 1 (Mmm1), mitochondrial distribution and morphology 10 (Mdm10), Mdm12 and Mdm34, therefore identifying these four proteins as constituents of a complex that ties the organelles together and allows the exchange of phospholipids (which are needed by the mitochondrial membranes) and calcium (which acts as a signalling molecule between the two).

Two properties that we still cannot reliably engineer are the dynamics of a circuit and spatial control. Both of these behaviours have one major biological process in common: development. In anticipation of one day tackling developmental processes and other intercellular pathways, some groups have designed circuits to spatiotemporally control gene expression. For example, using a network that mimicks naturally occurring feedforward circuits, Basu *et al.* have designed cells that can respond to the signal acyl-homoserine lactone (AHL) from nearby cells but ignore equal concentrations of this signal from faraway cells[92]. This feat is accomplished by a key property of the feedforward network in the signal-receiving cells — it responds not only to the concentration of the signal but also to the rate of increase of that concentration. Signal-sending cells that are near to signal-receiving cells increase the rate of AHL concentration more rapidly than signal-sending cells that are distant from signal-receiving cells. Basu *et al.* built on this work to create a circuit that could respond to only a narrow range of AHL signals, and in so doing recapitulated another feature of developmental processes — the ability of cells to behave like a band filter[93].

The exquisite coordination that is a hallmark of development almost certainly requires the use of networks that can act as genetic timers and counters. Friedland *et al.* have provided a design for a network that constitutively pumps out GFP mRNA transcripts

Box 3 | How to build a robust genetic oscillator

The simplest way to achieve oscillation is through the use of a delayed negative feedback loop[105]. Imagine that a system is constructed with two genes, *A* and *B*, and that protein A activates the transcription of *B*, whereas protein B inhibits the transcription of *A*. Turning on gene *A* leads to a build up of protein A, but also of protein B. After some time, enough protein B builds up to cause the levels of protein A to decrease: this results in a decrease in the levels of protein B, which allows the levels of protein A to rise, and so on.

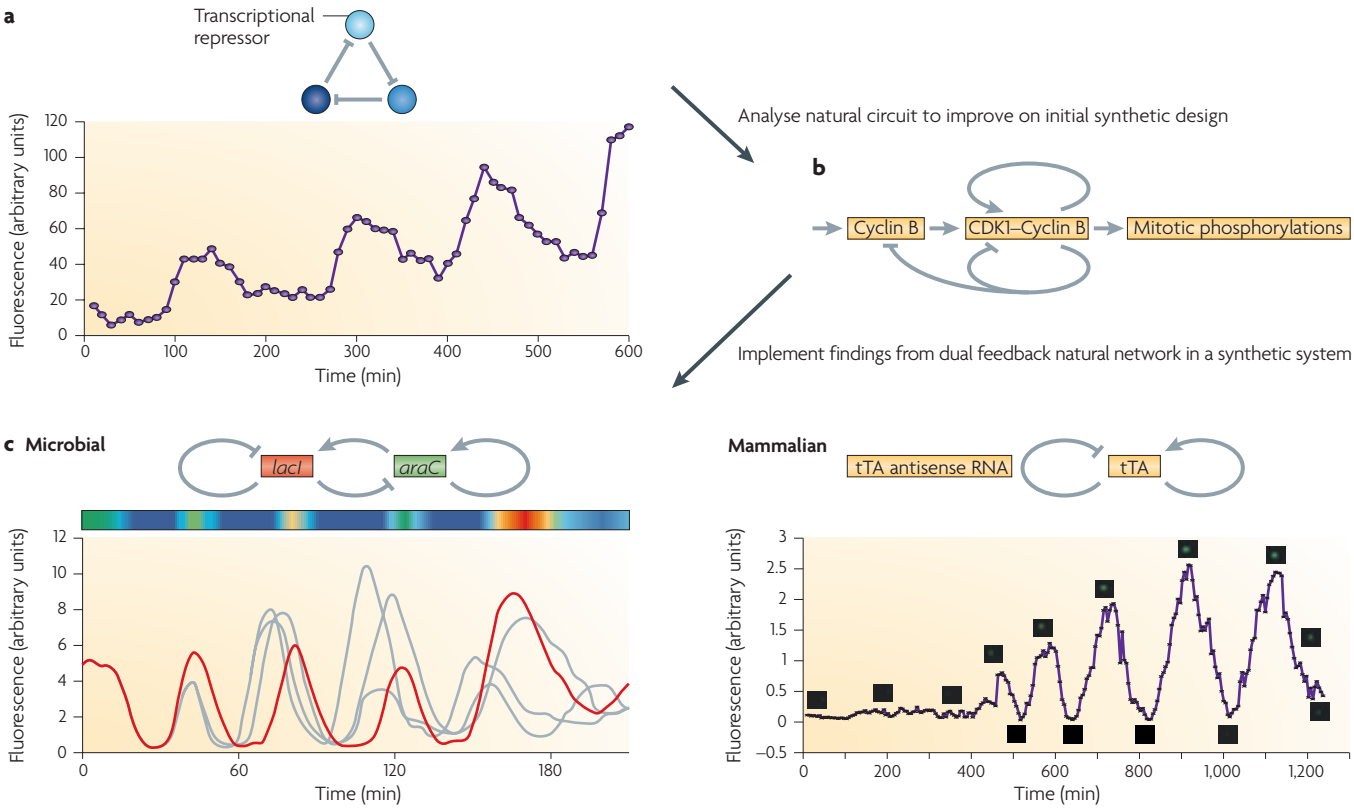
However, when a simple negative feedback circuit, such as the one described above, is constructed, the oscillations are in general not robust. In the repressilator built by Elowitz and Leibler[106], which consists of a cycle of three transcriptional repressors and a fluorescent protein read-out (see the figure, part **a**), the oscillators fall out of phase and 'damp out' after a small number of cycles. Swinburne *et al.* engineered an autoinhibitory circuit in which the delay timescale in the negative feedback was set by the length of an intron engineered into the construct (not shown); they also found that even for a given intron length, the oscillation period varied widely from cell to cell[107]. The source of the damping in both cases can be found in the stochastic nature of gene expression: random amounts of protein produced at random times result in uncoordinated behaviour that causes the components that make up the oscillator to fall out of phase. The synthetic genetic oscillator was missing a key feature.

A strong hint as to the identity of that key feature was provided by the analysis of naturally occurring oscillators. In particular, the cell cycle oscillator contained interlocking positive feedback loops in addition to the core negative feedback loop that was generally assumed to generate the oscillations (part **b**). Experiments in the cell cycle of frog embryos, along with computational simulations, suggested that the positive feedback loops could stabilize two states that the system would cycle between through the negative feedback loop[108–110], creating a relaxation oscillator. Could something as simple as positive feedback be responsible for robustness in genetic oscillators in organisms as diverse as bacteria and mammals? And can positive feedback enable cells to independently tune the amplitude and frequency of the oscillations?

Two recent studies, in agreement with earlier work[111], indicate that coupling positive and negative feedback is indeed sufficient to ensure stable oscillations. Stricker *et al.* implemented a transcriptional circuit in *Escherichia coli* that drives a fluorescent reporter[88] (part **c**). The circuit integrates the output of an autoregulatory positive feedback loop driven by the *araC* gene and a negative feedback loop in which AraC turns on *lacI* expression, which leads to *araC* transcriptional repression. Tigges *et al.*, working in mammalian cells, used a tetracycline-regulated transctivator (tTA) to drive transcriptional positive autofeedback and negative feedback mediated by tTA-dependent transcription of a tTA antisense RNA[89]; the state of the circuit is measured by the tTA-dependent transcription of *GFP*. The plots in part **c** show the steady cycling of the fluorescent reporter levels in each circuit over time.

Experimentally, Stricker *et al.* observed that the dual feedback oscillator was robust to a number of perturbations, including changes in inducer level and temperature. These features could not be adequately described by their initial modelling of this circuit[112] — it was only through the addition of various biological steps in the negative feedback loop, such as transcription factor–DNA binding and multimerization, that the model could reproduce the robustness of the oscillator to parameter changes. The authors concluded that from the point of view of the operation of the oscillator, what matters is not the details of which processes make up the negative feedback loop but instead that the negative feedback loop includes a delay; by contrast, the positive feedback loop only ensures robustness and tunability.

The system built by Tigges *et al.* shares many of these details, with the delay in the negative feedback coming from post-transcriptional repression of the circuit's transcriptional activator, but the system itself is sensitive to molecular details, such as the relative ratios of the circuit components — for some ratios of circuit components, oscillations were abolished. Part **a** is reproduced, with permission, from REF. 106 © (2000) Macmillan Publishers Ltd. All rights reserved. Part **c** (microbial) is reproduced, with permission, from REF. 88 © (2008) Macmillan Publishers Ltd. All rights reserved. Part **c** (mammalian) is reproduced, with permission, from REF. 89 © (2009) Macmillan Publishers Ltd. All rights reserved.



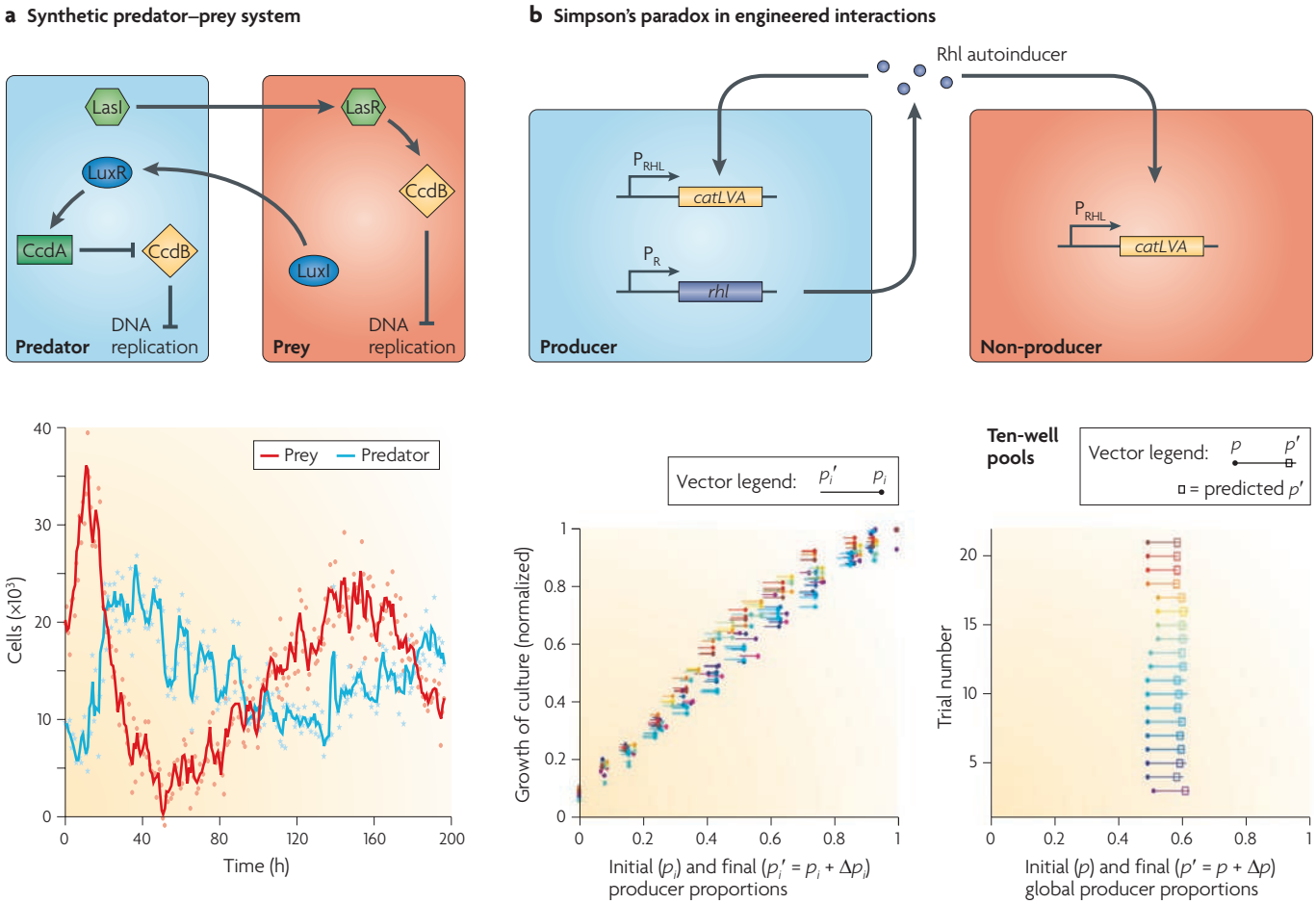


Figure 3 | **Using synthetic circuits to engineer cell–cell interactions.** Studies of ecology and evolution are often dependent on carefully characterizing the interactions of different organisms. In a natural setting, however, such data collection often proves to be noisy at best and impossible at worst. At the same time, mathematical models in theoretical ecology and evolutionary biology are among the most sophisticated in all of the life sciences. Laboratory-scale experiments on cellular interactions could quantitatively test some of the remarkable predictions and open the way to new theory. **a** | Among the most elementary interactions in nature is the predator–prey interaction. The prey in this case produces the quorum-sensing pathway protein LuxI, which is engineered to drive a transcriptional cascade in the predator that produces CcdA, which inhibits the DNA replication inhibitor CcdB, thereby allowing the predator to replicate. The predator produces the quorum-sensing pathway protein LasI, which activates CcdB in a LasR-dependent manner in the prey. CcdB expression in the prey prevents it from replicating. The graph shows that the cyclic dynamic is similar in style to that of genetic oscillators: high levels of prey lead to low levels of CcdB and therefore high levels of predator; high levels of predator lead to high levels of CcdB and therefore low levels of prey, which subsequently leads to high levels of CcdB in predators, and so on. As shown in Balagadde *et al.*[100], predator–prey interactions can result in limit cycle oscillations around an unstable fixed point of the dynamics; these interactions are most commonly studied in the framework of the Lotka–Volterra model. **b** | Simpson's paradox is a statistical phenomenon that captures the fact that even if the producer of a common good grows at a slower rate in all given subpopulations than a non-producer, it can nevertheless make up an increasing fraction of the population as a whole. Although Simpson's paradox usually arises as a result of misinterpretation of data, natural populations can in fact display heterogeneities in sample size that often underlie the paradox. The particular implementation in Chuang *et al.*[102] casts bacteria that generate the rhamnolipid (*rhl*) autoinducer as the producers. Although Rhl is only expressed by producers, which consequently grow more slowly, both producers and non-producers use this Rhl, which is rewired to activate the synthesis of a chloramphenicol resistance gene called *catLVA*. The graphs depict how the fraction of producers changes over time in different subpopulations. Dots represent the initial fraction of producers in a given subpopulation, and the length of the line emanating from each dot represents the change in the fraction of producers in that given subpopulation. As shown in the left graph, in which each data point represents a particular subpopulation, the fraction of producers decreases as the cells are exposed to chloramphenicol. But as shown in the right graph, in which each data point represents a pooled set of subpopulations from a given experimental run, the fraction of producers actually increases in the global population, therefore illustrating Simpson's paradox. The graph in part **a** is reproduced, with permission, from REF. 100 © (2008) Macmillan Publishers Ltd. All rights reserved. The graphs in part **b** are reproduced, with permission, from REF. 102 © (2009) American Association for the Advancement of Science.

Relaxation oscillator
An oscillator made up of two states and characterized by cycles of relatively long persistence in a state followed by rapid transitions to the other state.

Limit cycle oscillation
A periodic solution to a set of differential equations that is characterized by either attracting or repelling nearby solutions.

Lotka–Volterra model
A first-order nonlinear set of ordinary differential equations that are used to model the interactions between predators and prey. The model is most well known for admitting periodic solutions in which predator numbers rise and fall with prey numbers after a specified lag time.

Box 4 | **Creating extremely complex synthetic systems**

In the future the synthetic circuits deployed in cells will grow in complexity and will integrate multiple cellular processes, as has been done for genetic regulation and metabolism[83]. There is likely to be increasing overlap between synthetic biology and large-scale cell biology owing to the creation of synthetic organelles, the *in vivo* construction of which will be guided by synthetic regulatory networks. Progress along these fronts is currently limited by many of the same obstacles found across the subdisciplines of biology: we are still in need of more ways to specifically modulate the expression levels of genes of interest and the activity states of pathways of interest, and we require more sensitive techniques (ideally at single-molecule resolution) to measure the abundance of mRNAs, proteins and specifically modified proteins in live cells.

Methodological advances will be particularly useful for tightly constraining models of biological networks. Obstacles that occur when synthetic circuits are adapted from the blackboard to the cell can often be traced to the fact that the system under study does not behave as the initial modelling indicated. This, in turn, is usually due to the fact that the systems are under-determined, meaning that many different models can usually describe the circuit data. Higher resolution data, both in terms of abundances of the relevant molecules and as a function of time, will constrain the space of possible models substantially and should allow for more rational, predictable design processes.

Assuming these technical obstacles are overcome, in a future in which man-made circuits increasingly look like their byzantine natural counterparts, it is not unreasonable to expect nearly synthetic or fully synthetic cells to make their appearance. At these extreme levels of complexity, it may prove difficult or even unhelpful to mechanistically model the relevant systems. However, comparing the performance of natural cells or circuits with their synthetic equivalents in a rigorous fashion – perhaps through the formulation of a Turing test for synthetic biology – will probably prove useful, as differences in performance can point to possible design principles.

that are translationally inhibited but can have their inhibition removed by a transactivating RNA (taRNA)[94]; the transcription of the taRNA is inducible by arabinose, so the network output, in the form of discrete amounts of GFP, represents pulses of arabinose.

Finally, Isalan *et al.* have gone as far as building a mock-up of a realistic *D. melanogaster* embryo; they modelled the syncytium as a collection of paramagnetic beads coated with DNA, in which genetic networks analogous to the gap gene system could be placed[95]. Interestingly, this 'minimal embryo' led the authors to suggest that pattern formation in the real embryo requires activator molecules to propagate faster than inhibitors, which implies that the gap system is a reaction–diffusion system that uses a mechanism unlike Turing instabilities to lay down patterns. As the authors point out, this is hardly surprising given that the gap system uses non-homogeneous initial conditions in the form of spatially localized components deposited in the insect egg, and as the activator is not autocatalytic. Whether these observations hold true in natural embryos remains to be seen.

Modelling ecological interactions. As is the case with the band filter circuits described above, most synthetic circuits involved in cell–cell communication make use of the quorum-sensing pathway[96] (FIG. 3a). These circuits usually borrow components from organisms like *Vibrio fischeri*, although attempts at incorporating other systems have also been successful[97,98]. Examples of using such systems to study natural phenomena are more limited. Balagadde *et al.*, by adapting an earlier design[99], used the quorum-sensing proteins to drive the expression of an antibiotic to create a synthetic predator–prey system[100], and Brenner *et al.* used a similar system to study the ability of cells to signal in the context of a biofilm[101]. Chuang *et al.* recently used engineered circuits for cell–cell interactions to study the evolutionary phenomenon of Simpson's paradox (FIG. 3b), in which the cells that provide a useful product to the population make up a diminishing fraction of the population but nevertheless increase in absolute number by promoting population growth[102]. Gore *et al.* provide another example of synthetic ecology in their study of the evolutionary game dynamics underlying sucrose metabolism in yeast[103]. The study establishes that sucrose metabolism can be thought of as a snowdrift game, in which cells that metabolize sucrose (cooperators) and those that do not (cheaters) stably coexist in a population, thereby showing how competition between different alleles can actually promote diversity in a population.

Studies such as these on fundamental aspects of ecology and evolution are difficult to carry out in natural environments owing to the multiplicity of confounding factors, but synthetically engineered populations provide a way to cleanly separate different effects. Studies on engineered populations not only highlight the ability to connect the molecular details of a network to population-level effects but also the utility of abstracting away from such details and focusing on general cell–cell interactions. Taking sucrose metabolism from Gore *et al.* as an example, it was possible to predict population-level responses to changes in the cost of cooperation based on the theoretical characterization of the interaction between cheaters and cooperators; no direct knowledge of the molecular details was needed. Indeed, this approach of constructing synthetic systems that are dedicated to characterizing how cells interact can be useful in cases such as cancer dynamics, in which the underlying molecular details are either poorly understood or exceedingly complicated but in which population-level measurements are both feasible and relevant to understanding the phenomenon.

Perspectives

The synthetic biology community has made great strides in working out some of the most basic features of regulatory networks and cellular pathways. We are exerting greater control over the process of gene expression, and we have a wealth of information regarding the effects of network topology on system function. Topological details, such as connectivity, cascade length and feedback structure, have been explored. But there is much work to do before we can treat biological circuits in the same way as we treat electronic ones (BOX 4).

Looking back on the various examples of circuits and processes that synthetic biologists have examined, we can see that the usefulness of synthetic circuits can be measured in three different dimensions. First, synthetic circuits can serve as easily manipulable toy models that we can characterize in exacting quantitative detail to give insights into how similarly structured

Turing test
In computer science, a hypothetical test that is meant to decide whether a machine is displaying intelligent behaviour.

Syncytium
A collection of cytoplasm that contains several nuclei.

Reaction–diffusion
A class of mathematical models in which the concentrations of the molecules being modelled are tracked in space as well as time, taking into account the chemical transformations that the molecules can undergo and their diffusive motion.

Turing instability
A mathematical condition in reaction–diffusion systems in which differences in the diffusion of activating and inhibiting morphogenic molecules result in pattern formation; particular patterns form when inhibitors diffuse faster than autoactivators.

Quorum-sensing pathway
A signalling pathway used by microbes to determine the abundance of related and unrelated microbes in the local environment through the exchange of specific small molecules.

natural networks operate. Second, synthetic circuits can be used to control natural networks to make discoveries about the molecular and cellular biology that underlies important physiological processes. Third, on a more conceptual level, synthetic systems provide clear evidence that one can generate complexity by rearranging even well-known parts, therefore bolstering claims of the evolvability of natural systems.

Although we are still very far from rationally assembling a living organism from scratch, and far from understanding all of the design principles under which biological networks operate, the first generation of synthetically designed systems has highlighted the need to unite tools from disparate processes — from transcriptional regulation to signal transduction — to approach fundamental questions in modern biology.

1. Mattick, J. S. RNA regulation: a new genetics? *Nature Rev. Genet.* **5**, 316–323 (2004).
2. Davidson, E. H. *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* (Academic Press, 2006).
3. Prud'homme, B. P., Gompel, N. & Carroll, S. B. Emerging principles of regulatory evolution. *Proc. Natl Acad. Sci. USA* **104** (Suppl. 1), 8605–8612 (2007).
4. Bridgham, J. T., Carroll, S. M. & Thornton, J. W. Evolution of hormone–receptor complexity by molecular exploitation. *Science* **312**, 97–101 (2006). **This study used the rational synthesis of a resurrected glucocorticoid receptor to explore the evolution of the receptor–ligand pair.**
5. Rapp, M., Seppala, S., Granseth, E. & von Heijne, G. Emulating membrane protein evolution by rational design. *Science* **315**, 1282–1284 (2007).
6. Gilbert, E. S., Walker, A. W. & Keasling, J. D. A constructed microbial consortium for biodegradation of the organophosphorus insecticide parathion. *Appl. Microbiol. Biotechnol.* **61**, 77–81 (2003).
7. Rajendran, M. & Ellington, A. D. Selection of fluorescent aptamer beacons that light up in the presence of zinc. *Anal. Bioanal. Chem.* **390**, 1067–1075 (2008).
8. Steen, E. J. *et al.* Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microb. Cell Fact.* **7**, 36 (2008).
9. Waks, Z. & Silver, P. A. Engineering a synthetic dual organism system for hydrogen production. *Appl. Environ. Microbiol.* **75**, 1867–1875 (2009).
10. Khosla, C. & Keasling, J. D. Metabolic engineering for drug discovery and development. *Nature Rev. Drug Discov.* **2**, 1019–1025 (2003).
11. Ro, D. *et al.* Production of the antimalarial drug precursor artemisinc acid in engineered yeast. *Nature* **440**, 940–943 (2006). **Perhaps the most striking example of a pathway that has been successfully engineered to meet a design goal, in this case the production of the antimalarial compound arteminisin.**
12. Anderson, J. C., Clarke, E. J., Arkin, A. P. & Voigt, C. A. Environmentally controlled invasion of cancer cells by engineered bacteria. *J. Mol. Biol.* **355**, 619–627 (2006).
13. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).
14. Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S. & Elowitz, M. B. Gene regulation at the single-cell level. *Science* **307**, 1962–1965 (2005).
15. Pedraza J. & van Oudenaarden, A. Noise propagation in gene networks. *Science* **307**, 1965–1968 (2005). **Reference 14 introduced the concept of the gene regulation function (GRF) and showed that the GRF fluctuates from cell to cell, whereas reference 15 used a synthetic transcriptional cascade to measure and model how these fluctuations can be attributed to noise intrinsic to the expression of a gene and GRF fluctuations from upstream components.**
16. Setty, Y., Mayo, A. E., Surette, M. G. & Alon, U. Detailed map of a *cis*-regluatory input function. *Proc. Natl Acad. Sci. USA* **100**, 7702–7707 (2003).
17. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31**, 69–73 (2002).
18. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
19. Golding, I., Paulsson, J., Zawilski, S. M. & Cox, E. C. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025–1036 (2005).
20. Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309 (2006).
21. Cai, L., Friedman, N. & Xie, X. S. Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358–362 (2006). **References 19–21 are three single-molecule studies that directly visualized the burst-like nature of transcription and translation, capturing the stochastic nature of these basic processes in great detail.**
22. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008).
23. Hammer, K., Mijakovic, I. & Jensen, P. R. Synthetic promoter libraries — tuning of gene expression. *Trends Biotechnol.* **24**, 53–55 (2006).
24. Cox, R. S., Surette, M. G. & Elowitz, M. B. Programming gene expression with combinatorial promoters. *Mol. Syst. Biol.* **3**, 145 (2007).
25. Kinkhabwala, A. & Guet, C. C. Uncovering *cis* regulatory codes using synthetic promoter shuffling. *PLoS ONE* **3**, e2030 (2008).
26. Segal, E. & Widom, J. From DNA sequence to transcriptional behaviour: a quantitative approach. *Nature Rev. Genet.* **10**, 443–456 (2009).
27. Gertz, J., Siggia, E. D. & Cohen, B. A. Analysis of combinatorial *cis*-regulation in synthetic and genomic promoters. *Nature* **457**, 215–218 (2009).
28. Buchler, N. E., Gerland, U. & Hwa, T. On schemes of combinatorial transcription logic. *Proc. Natl Acad. Sci. USA* **100**, 5136–5141 (2003).
29. Becskei, A., Kaufmann, B. B. & van Oudenaarden, A. Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nature Genet.* **37**, 937–944 (2005).
30. Pokharel, S. & Beal, P. A. High-throughput screening for function adenosine to inosine RNA editing systems. *ACS Chem. Biol.* **1**, 761–765 (2006).
31. Beisel, C. L., Bayer, T. S., Hoff, K. G. & Smolke, C. D. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Mol. Syst. Biol.* **4**, 224 (2008).
32. Win, M. N. & Smolke, C. D. Higher-order cellular information processing with synthetic RNA devices. *Science* **322**, 456–460 (2008).
33. Werstruck, G. & Green, M. R. Controlling gene expression in living cells through small molecule–RNA interactions. *Science* **282**, 296–298 (1998).
34. Grate, D. & Wilson, C. Inducible regulation of the *S. cerevisiae* cell cycle mediated by an RNA aptamer–ligand complex. *Bioorg. Med. Chem.* **9**, 2565–2570 (2001).
35. Suess, B., Fink, B., Berens, C., Stentz, R. & Hillen, W. A. A theophylline responsive riboswitch based on helix slipping controls gene expression *in vivo*. *Nucleic Acids Res.* **32**, 1610–1614 (2004).
36. Desai, S. K. & Gallivan, J. P. Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J. Am. Chem. Soc.* **126**, 13247–13254 (2004).
37. Davidson, E. A. & Ellington, A. D. Synthetic RNA circuits. *Nature Chem. Biol.* **3**, 23–28 (2007).
38. Isaacs, F. J. *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nature Biotechnol.* **22**, 841–847 (2004).
39. Grilly, C., Stricker, J., Pang, W. L., Bennett, M. R. & Hasty, J. A synthetic gene network for tuning protein degradation in *Saccharomyces cerevisiae*. *Mol. Syst. Biol.* **3**, 127 (2007).
40. Deans, T. L., Cantor, C. R. & Collins, J. J. A tunable genetic switch based on RNAi and repressor proteins for regulating gene expression in mammalian cells. *Cell* **130**, 363–372 (2007).
41. Martin, C. H., Nielsen, D. R., Solomon, K. V. & Prather, K. L. Synthetic metabolism: engineering biology at the protein and pathway scales. *Chem. Biol.* **16**, 277–286 (2009).
42. Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnol.* **27**, 753–759 (2009).
43. Janin, J. & Chothia, C. Domains in proteins: definitions, location, and structural principles. *Methods Enzymol.* **115**, 420–430 (1985).
44. Rothlisberger, D. *et al.* Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
45. Kaplan, J. & DeGrado, W. F. *De novo* design of catalytic proteins. *Proc. Natl Acad. Sci. USA* **101**, 11566–11570 (2004).
46. Bhattacharyya, R. P., Remenyi, A., Yeh, B. J. & Lim, W. A. Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. *Ann. Rev. Biochem.* **75**, 655–680 (2006).
47. Proft, M. & Struhl, K. MAP kinase-mediated stress relief that precedes and regulates the timing of transcriptional induction. *Cell* **118**, 351–361 (2004).
48. Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S. & Roth, B. L. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl Acad. Sci. USA* **104**, 5163–5168 (2007).
49. Cironi, P., Swinburne, I. A. & Silver, P. A. Enhancement of cell type specificity by quantitative modulation of a chimeric ligand. *J. Biol. Chem.* **283**, 8469–8476 (2008).
50. Levskaya, A. *et al.* Engineering bacteria to see light. *Nature* **438**, 441–442 (2005).
51. Shimizu-Sato, S., Huq, E., Tepperman, J. M. & Quail, P. H. A light-switchable gene promoter system. *Nature Biotechnol.* **20**, 1041–1044 (2002).
52. Cruz, F. G., Koh, J. T. & Link, K. H. Light activated gene expression. *J. Am. Chem. Soc.* **122**, 8777–8778 (2000).
53. Cambridge, S. B., Geissler, D., Keller, S. & Curten, B. A caged doxycycline analogue for photoactivatable gene expression. *Angew. Chem. Int. Ed.* **45**, 2229–2231 (2006).
54. Dugave, C. & Demange, L. *Cis–trans* isomerization of organic molecules and biomolecules: implications and applications. *Chem. Rev.* **103**, 2475–2532 (2003).
55. Young, D. D. & Deiters, A. Photochemical control of biological processes. *Org. Biomol. Chem.* **5**, 999–1005 (2007).
56. Taylor, R. J. *et al.* Dynamic analysis of MAPK signaling using a high-throughput microfluidic single-cell imaging platform. *Proc. Natl Acad. Sci. USA* **106**, 3758–3763 (2009).
57. Mettetal, J. T., Muzzey, D., Gomez-Uribe, C. & van Oudenaarden, A. The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*. *Science* **319**, 482–484 (2008).
58. Hersen, P., McClean, M. N., Mahadevan, L. & Ramanathan, S. Signal processing by the HOG MAP kinase pathway. *Proc. Natl Acad. Sci. USA* **105**, 7165–7170 (2008).
59. Ubersax, J. A. & Ferrell, J. E. Mechanisms of specificity in protein phosphorylation. *Nature Rev. Mol. Cell Biol.* **8**, 530–541 (2007).
60. Harris, K. *et al.* Role of scaffolds in MAP kinase pathway specificity revealed by custom design of pathway-dedicated signaling proteins. *Curr. Biol.* **11**, 1815–1824 (2001).
61. McClean, M. N., Mody, A., Broach, J. R. & Ramanathan, S. Cross-talk and decision making in MAP kinase pathways. *Nature Genet.* **39**, 409–414 (2007).
62. Behar, M., Dohlman, H. G. & Elston, T. C. Kinetic insulation as an effective mechanism for achieving pathway specificity in intracellular signaling networks. *Proc. Natl Acad. Sci. USA* **104**, 16146–16151 (2007).

63. Skerker, J. M. *et al.* Rewiring the specificity of two-component signal transduction systems. *Cell* **133**, 1043–1054 (2008).
This study highlights the combined use of bioinformatic analysis, structural data and rewired pathways for determining the molecular basis of pathway specificity in bacterial two-component systems.
64. Tatebayashi, K., Takekawa, M. & Saito, H. A docking site determining specificity of Pbs2 MAPKK for Ssk2/Ssk22 MAPKKKs in the yeast HOG pathway. *EMBO J.* **22**, 3624–3634 (2003).
65. Remenyi, A., Good, M. C. & Lim, W. A. Docking interactions in protein kinase and phosphatase networks. *Curr. Opin. Struct. Biol.* **16**, 676–685 (2006).
66. Howard, P. L., Chia, M. C., Del Rizzo, S., Liu, F. F. & Pawson, T. Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins, *Proc. Natl Acad. Sci. USA* **100**, 11267–11272 (2003).
67. Yeh, B. J., Rutigliano, R. J., Deb, A., Bar-Sagi, D. & Lim, W. A. Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. *Nature* **447**, 596–600 (2007).
68. Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl Acad. Sci. USA* **102**, 3581–3586 (2005).
69. Guet, C. C., Elowitz, M. B., Hsing, W. & Leibler, S. Combinatorial synthesis of genetic networks. *Science* **296**, 1466–1470 (2002).
This paper explores a library of possible network topologies in a way that highlights the impressive flexibility of even simple gene networks, notably showing that networks of identical topology can have different behaviours.
70. Isalan, M. *et al.* Evolvability and hierarchy in rewired bacterial gene networks. *Nature* **452**, 840–845 (2008).
This study demonstrates the ability of *E. coli* transcriptional networks to tolerate new connections and topologies and, in a limited number of cases, to exploit them to achieve higher fitness.
71. Tsong, A. E., Tuch, B. B., Li, H. & Johnson, A. D. Evolution of alternative transcriptional circuits with identical logic. *Nature* **443**, 415–420 (2006).
72. Wagner, A. *Robustness and Evolvability in Living Systems* (Princeton Univ. Press, 2007).
73. Gerhart, J. & Kirschner, M. The theory of facilitated variation. *Proc. Natl Acad. Sci. USA* **104** (Suppl. 1), 8582–8589 (2007).
74. Antunes, M. S. *et al.* Engineering key components in a synthetic eukaryotic signal transduction pathway. *Mol. Syst. Biol.* **5**, 270 (2009).
75. Alon, U. Network motifs: theory and experimental approaches. *Nature Rev. Genet.* **8**, 450–461 (2007).
76. Maeda, Y. T. & Sano, M. Regulatory dynamics of synthetic gene networks with positive feedback. *J. Mol. Biol.* **359**, 1107–1124 (2006).
77. Becskei, A. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20**, 2528–2535 (2001).
78. Ozbudak, E., Thattai, M., Lim, H. N., Shraiman, B. I. & van Oudenaarden, A. Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427**, 737–740 (2004).
79. Isaacs, F. J., Hasty, J. & Collins, J. J. Prediction and measurement of an autoregulatory genetic module. *Proc. Natl Acad. Sci. USA* **100**, 7714–7719 (2003).
80. Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).
81. Rosenfeld, N., Elowitz, M. B. & Alon, U. Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.* **323**, 785–793 (2002).
82. Bashor, C. J., Helman, N. C., Yan, S. & Lim, W. A. Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. *Science* **319**, 1539–1543 (2008).
This paper transformed the perception of scaffold proteins as being passive aggregators of pathway components to that of being active players in specifying pathway dynamics, through features such as participation in feedback loops.
83. Fung, E. *et al.* A synthetic gene-metabolic oscillator. *Nature* **435**, 118–122 (2005).
84. Cantone, I. *et al.* A yeast synthetic network for *in vivo* assessment of reverse-engineering and modeling approaches. *Cell* **137**, 172–181 (2009).
85. Ellis, T., Wang, X. & Collins, J. J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nature Biotechnol.* **27**, 465–471 (2009).
86. Nelson, D. E. *et al.* Oscillations in NF-κB signaling control the dynamics of gene expression. *Science* **306**, 704–708 (2004).
87. Geva-Zatorsky, N. *et al.* Oscillations and variability in the p53 system. *Mol. Syst. Biol.* **2**, 2006.0033 (2006).
88. Stricker, J. *et al.* A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516–519 (2008).
89. Tigges, M., Marquez-Lago, T. T., Stelling, J. & Fussenegger, M. A tunable synthetic mammalian oscillator. *Nature* **457**, 309–312 (2009).
References 88 and 89 achieved, in microbial and mammalian contexts, respectively, the long-sought-after design and implementation of robust, tunable synthetic genetic oscillators.
90. Acar, M., Becskei, A. & van Oudenaarden, A. Enhancement of cellular memory by reducing stochastic transitions. *Nature* **435**, 228–232 (2005).
91. Kornmann, B. *et al.* An ER–mitochondria tethering complex revealed by a synthetic biology screen. *Science* **325**, 477–481 (2009).
92. Basu, S., Mehreja, R., Thiberge, S., Chen, M. T. & Weiss, R. Spatiotemporal control of gene expression with pulse-generating networks. *Proc. Natl Acad. Sci. USA* **101**, 6355–6360 (2004).
An excellent example of a circuit that responds to the dynamics of an input signal rather than the steady state input signal.
93. Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130–1134 (2005).
94. Friedland, A. *et al.* Synthetic gene networks that count. *Science* **324**, 1199–1202 (2009).
95. Isalan, M., Lemerle, C. & Serrano, L. Engineering gene networks to emulate *Drosophila* embryonic pattern formation. *PLoS Biol.* **3**, e64 (2005).
96. Tanouchi, Y., Pai, A. & You, L. Decoding biological principles using gene circuits. *Mol. Biosyst.* **5**, 695–703 (2009).
97. Bulter, T. *et al.* Design of artificial cell–cell communication using gene and metabolic networks. *Proc. Natl Acad. Sci. USA* **101**, 2299–2304 (2004).
98. Chen, M.-T. & Weiss, R. Artificial cell–cell communication in yeast *Saccharomyces cerevisiae* using signaling elements from *Arabidopsis thaliana*. *Nature Biotechnol.* **23**, 1551–1555 (2005).
99. You, L., Cox, R. S., Weiss, R. & Arnold, F. H. Programmed population control by cell–cell communication and regulated killing. *Nature* **428**, 868–871 (2004).
100. Balagadde, F. K. *et al.* A synthetic *Escherichia coli* predator–prey ecosystem. *Mol. Syst. Biol.* **4**, 187 (2008).
101. Brenner, K., Karig, D. K., Weiss, R. & Arnold, F. H. Engineered bidirectional communication mediates a consensus in a microbial biofilm consortium. *Proc. Natl Acad. Sci. USA* **104**, 17300–17304 (2007).
102. Chuang, J. S., Rivoire, O. & Leibler, S. Simpson's paradox in a synthetic microbial system. *Science* **323**, 272–275 (2009).
103. Gore, J., Youk, H. & van Oudenaarden, A. Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459**, 253–256 (2009).
104. Kwon, O., Georgellis, D. & Lin, E. C. C. Rotational on–off switching of a hybrid membrane sensor kinase Tar–ArcB in *Escherichia coli*. *J. Biol. Chem.* **278**, 13192–13195 (2003).
105. Conrad, E. D. & Tyson, J. T. in *System Modeling in Cellular Biology* (eds Szallasi, Z., Stelling, J. & Periwal, V.) 116–118 (MIT Press, 2006).
This paper is contained within a useful reference on mathematical techniques that are used to model biological circuits.
106. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335–338 (2000).
107. Swinburne, I. A., Miguez, D. G., Landgraf, D. & Silver, P. A. Intron length increases oscillatory periods of gene expression in animal cells. *Genes Dev.* **22**, 2342–2346 (2008).
108. Pomerening, J. R., Sontag, E. D. & Ferrell, J. E. Building a cell cycle oscillator: hysterisis and bistability in the activation of Cdc2. *Nature Cell Biol.* **5**, 346–351 (2003).
109. Pomerening, J. R., Kim, S. Y. & Ferrell, J. E. Systems-level dissection of the cell-cycle oscillator: bypassing positive feedback produces damped oscillations. *Cell* **122**, 565–578 (2005).
110. Tsai, T. Y. *et al.* Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science* **321**, 126–129 (2008).
111. Atkinson, M. R., Savageau, M. A., Myers, J. T. & Ninfa, A. J. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell* **113**, 597–607 (2003).
112. Hasty, J., Dolnik, M., Rottschafer, V. & Collins, J. J. Synthetic gene network for entraining and amplifying cellular oscillations. *Phys. Rev. Lett.* **88**, 148101 (2002).

Acknowledgements
We apologize to our colleagues whose work was not discussed. This work was supported by grants from the US National Institutes of Health and the National Science Foundation.

FURTHER INFORMATION
Shankar Mukherji's homepage:
http://web.mit.edu/mukherji/www
Alexander van Oudenaarden's homepage:
http://web.mit.edu/biophysics
***Nature Reviews Genetics* series on Modelling:**
http://www.nature.com/nrg/series/modelling
ALL LINKS ARE ACTIVE IN THE ONLINE PDF