Integrative Biology

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/ibiology

Insight box:

We discuss different types of biological uncertainty that plague the predictable design of gene circuits. We classify these uncertainties into three categories: incomplete characterization of biological parts, unintended interactions between components and the host, and stochastic dynamics. We suggest different methods and review their effectiveness at minimizing each type of uncertainty. Our unique perspective on this topic can help guide novel circuit design. Contextualizing different problems that can cause circuit failure will expedite their efficient identification. Moreover, providing potential strategies can reduce time and resources used to optimize circuits, which could help speed up scientific discovery.

Addressing biological uncertainties in engineering gene circuits

Carolyn Zhang[1+], Ryan Tsoi[1+], Lingchong You[1,2*]

[1]Department of Biomedical Engineering, Duke University, Durham, North Carolina, 27708, USA

[2]Center for Genomic and Computational Biology, Duke University, Durham, North Carolina, 27708, USA

[+]These authors contributed equally

*Corresponding author. Department of Biomedical Engineering, Duke University, CIEMAS 2355, 101 Science Drive, Box 3382, Durham, North Carolina, 27708, USA

Tel.: +1 (919)660-8408; Fax: +1 (919)668-0795; E-mail:you@duke.edu

Abstract

Synthetic biology has grown tremendously over the past fifteen years. It represents a new strategy to develop biological understanding and holds great promise for diverse practical applications. Engineering of a gene circuit typically involves computational design of the circuit, selection of circuit components, and test and optimization of circuit functions. A fundamental challenge in this process is the predictable control of circuit function due to multiple layers of biological uncertainties. These uncertainties can arise from different sources. We categorize these uncertainties into incomplete quantification of parts, interactions between heterologous components and the host, or stochastic dynamics of chemical reactions and outline potential design strategies to minimize or exploit them.

Introduction

Synthetic biology has shown great promise in contributing to our basic understanding of biology [1] and creating novel systems with practical applications [2, 3]. While there are many facets to synthetic biology, we focus on the engineering of genetic circuits. From the

development of gene networks as biosensors [4] to the incorporation of complex regulatory modules in model organisms [5], synthetic circuits have the potential for applications in biological research [6-9]. Despite past successes, the predictable design and implementation of these circuits remains a fundamental challenge. This limitation can be attributed to the many layers of uncertainty that emerge throughout the engineering process.

Engineering genetic circuits has often been compared to programming [10], where the cell is the computer and the gene circuits are introduced software programs. From this perspective, building a gene circuit is like inserting a small script into an operating system without full understanding of the context. Despite knowing the programming language, an incomplete understanding of the operating system provides a layer of uncertainty similar to introducing a gene circuit. The program must be written without syntax errors, must not hinder underlying operations that maintain the system, and must have variables that do not overlap with those that already exist. Ideally, a gene circuit must use the correct parts, must not inhibit the growth of the host, and must be orthogonal to native processes. However, these conditions are difficult to realize due to multiple layers of uncertainties, which are often challenging to anticipate. Here, we discuss some common uncertainties that confound predictable engineering of gene circuits in living cells as well as strategies to alleviate or take advantage of the impact of such uncertainties.

Sources of uncertainty

1. Incomplete characterization or quantification of biological components

In typical engineering disciplines, the building blocks are often well defined. For example, in electrical engineering the basic parameters associated with various components are well-documented [11]. In comparison, synthetic biology lacks the systematic quantification of parts fundamental to other engineering fields (**Figure 1A**). Even for model organisms such as

Escherichia coli, which has been a workhorse for microbiology, biotechnology, and gene circuit engineering, the kinetic properties of many biological components are poorly characterized (**Figure 1B**). For example, when constructing OR gates, Tamsir et. al. found that the unexpected transcriptional interference between two promoters caused one design to only respond to a single input [12]. The challenge increases drastically in higher organisms where precise genetic control is constrained by their greater complexity and a lack of well-quantified biological components [13-15].

2. Unintended interactions between circuit components and the chassis

A challenge in providing a comprehensive characterization of parts is their context dependence. In electronic circuit design, engineers achieve modularity of parts through the spatial segregation of components — for example, two transistors will never interact or share signals without direct connection by wires. In synthetic biology, however, a circuit component well characterized in one species or strain can behave unpredictably when introduced into another due to unintended interactions with native parts [16] (**Figure 1C)**. For example, expression of an algal nucleotide transporter for the uptake of unnatural nucleotides caused growth inhibition of *E. coli*, which the authors attributed to toxic effects of expressing heterologous membrane proteins [17].

Synthetic gene circuits rely on endogenous host machinery and resources such as ribosomes, polymerases, and other enzymes to carry out their designed function (**Figure 1D**). Since the host relies on the same pool of resources to maintain native processes, synthetic networks draw from this limited pool. Often referred to as the metabolic burden, this titration of resources can interfere with both the cell and the circuit. Indeed, heterologous gene expression can drastically inhibit growth in bacteria depending on specific genes and their expression levels [18-20]. Similarly, Karim et. al. observed that growth of *Saccharomyces cerevisiae* decreased

up to 25% when expressing different selection markers on plasmids depending on the origin of replication, the promoter, and the yeast strain [21]. Moreover, cellular growth and gene expression are intertwined by resource allocation constraints resulting in growth reduction [22].

Certain components or functions may be toxic to the host, which can occur when burden is too high or new genes are introduced from a different kingdom or species. For example, expression products of more than 15,000 genes from 393 microbial genomes inhibited growth of *E. coli* [23]. Forming the basis of antimicrobial development, toxic compounds have been found in organisms from plants [24] to fungi [25]. While these compounds can display inhibitory effects in some species, they can have limited effects in other organisms. The restriction endoribonuclease RegB, which is highly toxic to *E. coli*, exhibits no detectable toxicity in *S. cerevisiae* [26]. In some situations, gene mutations can have various host-dependent effects. The R436-S mutant form of the GyrB protein promotes temperature-sensitivity in *Salmonella enterica* but is lethal to *E. coli* [27]. In other cases, insertion of a foreign gene into the chromosome can result in unexpected cellular toxicity [28]. These interactions can impose selection pressure that causes genetic instability -- the loss of circuit function after prolonged circuit activation [29-31].

3. Stochastic dynamics

Even with precise measurements of the component parameters, predictable engineering of circuit dynamics is confounded by the randomness (noise) associated with cellular processes [32]. Ultimately, this noise results from the stochastic nature of reactions between small numbers of molecules (**Figure 1E**). In a bacterial cell, for example, many proteins are present in tens or hundreds of molecules. Gene expression noise from transcriptional bursting in *Bacillus subtilis* resulted in 30-250 GFP molecules per cell depending on the strength of the promoters [33]. At such small numbers, the relative magnitudes of fluctuations are large: it approximately

scales with the inverse of the square root of the average molecular number [34]. Thus, a protein with 100 copies per cell experiences at least a 10% fluctuation in protein concentration unless this fluctuation is suppressed by specific regulatory mechanisms. This is often referred to as "intrinsic" noise, as it reflects the baseline noise even if all the rate constants are unchanging. In addition, the rate constants themselves may fluctuate, as they are determined by available resources or other (often global) processes like transcription and translation (**Figure 1F**). This is typically referred to as "extrinsic" noise [35] and can be the dominant source of variability for highly expressed genes or in eukaryotes [36]. Extrinsic and intrinsic noise can be distinguished through various mathematical and experimental methods [36-38].

Regardless of its sources, cell-to-cell variability imposes a fundamental constraint on the performance of many gene circuits. For instance, a pioneering synthetic oscillator, the repressilator, exhibits highly variable oscillatory dynamics in growing cells: fewer than half of the cells exhibited oscillations, which were largely asynchronous between cell lineages [39]. The authors hypothesized that the lack of coherent oscillations was in part due to stochastic gene expression. Indeed, a recent study shows that cell-size control can lead to slow fluctuations and even transient oscillations in gene expression, in the absence of additional feedback regulation [40]. Overcoming this limitation has become a focus of subsequent oscillator designs.

Addressing uncertainty in gene circuit engineering

1. Documenting and characterizing parts

Circuit design relies on choosing optimal components to generate desired function. Part varieties has improved recently with the quantification of synthetic and natural terminators [41] and synthetic libraries of promoters in different host contexts [42-48]. Likewise, in mammalian cells, promoter libraries [49] and RNAi libraries for gene knockouts [50] have expanded the repertoire of potential components that can be incorporated into synthetic circuits. Ideally, a

complete catalog for biological parts would minimize uncertainty by providing a quantitative understanding of circuit components in different host contexts [51]. This would be a useful tool to design complex systems consisting of multiple gene circuits [16]. For example, one can imagine using a consortium of microorganisms for sophisticated sensing and processing functionality.

The ability to rapidly quantify the large number of biological parts is constrained by a dependence on living organisms as a chassis for gene circuits. This dependence has spurred research into engineering cell-free systems. Cell free systems utilize enzymes or metabolites isolated from microorganisms to simulate cellular reactions [52, 53]. They also serve as the foundation for efforts to develop minimal cells – encapsulated cell-free gene expression systems. In a minimal cell, the host contains the minimal number of genes necessary to survive, decreasing the opportunities for cross-interactions [54-56]. The former has been used for paper-based biosensors [4], while the latter has been explored for their potential as bioreactors [57]. Cell-free systems provide a more controlled environment and are easy to use [58], which can provide a foundation for rapid quantification of biological parts (**Figure 2A**). Libraries of parts can be contained on DNA and characterized using simple inducible systems expressing a reporter. With a standardized protocol, this strategy could present a consistent, high-throughput technique for systematic quantification of circuit components. Still, we note that these measurements only provide *initial* guidance on circuit design, especially for those implemented in living cells.

2. Quantifying effects of parts on the chassis

2.1 Computational models

Modeling can evaluate the impact of uncertainty on circuit function and the robustness of a circuit against these uncertainties. Typically, synthetic gene circuits are modeled

deterministically or stochastically [59, 60] to describe the temporal or spatial dynamics of circuit components or cells. Sensitivity or bifurcation analysis of models can allow one to estimate the parameter space that results in a desired function, such as varying input levels or promoter characteristics [44]. A large parameter space indicates that a circuit is less sensitive to perturbations, resulting in predictable and reliable function. In such models, the impact of metabolic burden or potential crosstalk can be evaluated [61] (**Figure 2B**), which provides another test for circuit performance.

Typical models can only describe metabolic burden or crosstalk in a lumped manner. In contrast, one can potentially gain insights into these uncertainties by using whole-cell models, which attempt to more comprehensively describe the known components and pathways. For example, Purcell et. al. used a whole-cell model of *Mycoplasma genitalium* to examine the burden of introducing heterologous circuits like the Goodwin oscillator [62]. However, increases in predictive power from whole-cell models come at the cost of being more computationally intensive. Moreover, it can introduce details that obscure the underlying mechanisms behind biological uncertainty, complicating further analysis [63].

In such cases, minimal models that account for allocation of limited resources can be used to evaluate the impact of metabolic burden. This creates a trade-off, where a resource committed to one process reduces its availability for other processes [63-65]. In one example, a model by Scott et. al. accounted for resources allotted to ribosome synthesis, defined as the RNA/protein ratio [22]. Empirical correlations between growth rates and RNA/protein ratio could predict the effect of cell proliferation on gene expression and vice versa [22]. Another model, which treated overall promoter activity as a fixed resource, showed that gene expression and growth rate are inherently correlated due to resource constraints [66]. While the previous models accounted for allocation of a single resource, these can be extended to include multiple

trade-offs. For example, WeiBe et. al. constructed a model that comprised three-tradeoffs: limited levels of cellular energy, finite ribosome levels, and a maximum cell mass [63]. Their model could predict trade-offs generated by introducing synthetic circuits among the circuit's function, its induction level, and the host's growth rates [63]. These models provide the ability to link mechanistic changes to phenotypic alterations without the complexity of whole-cell models.

2.2 Real-time quantification of burden

Real-time sensors that detect specific analytes would enable a more precise quantification of burden. To this end, Ceroni et. al. evaluated strong titration effects by tracking the production capacity in *E. coli* using a 'capacity monitor' [67]. This real-time sensor measures the capacity for GFP expression, which is used to assay burden (decreased growth rate) caused by circuit activation. The authors analyzed a library of synthetic constructs using this sensor and determined that increasing strength of the ribosome binding site (RBS) increased burden and the rate of mutations [67]. This is due to the titration of ribosomes, which can reduce available translational machinery for other processes like cell division. In contrast, circuits with weaker RBSs, which do not titrate ribosomes as strongly, maintained cell growth rates [67]. Real-time sensors integrated with downstream regulatory mechanisms could potentially be used to dynamically modulate uncertainty. Dahl et. al. designed such a system using stress-response promoters sensitive to accumulation of farnesyl pyrophosphate (FPP) and HMG-COA to control production of these intermediates in isoprenoid biosynthesis [68]. This strategy improved product yields, reduced intermediate accumulation, and increased cell growth.

2.3 *Exploiting effects due to uncertainty*

In some cases, interactions between hosts and gene circuits can result in unexpected phenotypes. In *E. coli*, growth inhibition by a simple positive feedback circuit using non-cooperative T7 polymerase resulted in bistable gene expression [20]. Another example is the

sharing of transcription factors among different promoter binding sites. While this may negatively affect native pathways, this biological uncertainty can be exploited to generate novel dynamics. For instance, titration of repressors by strongly competing binding sites generated an ultrasensitive response of a reporter (YFP) to increasing repressor concentrations [69]. As the number of repressors exceeds the number of competitive binding sites, the repressor becomes available to regulate the reporter. This creates a Hill-like "threshold response" such that a sufficiently high concentration of the repressor is required to bind the promoter and inhibit transcription. This transition becomes sharper with increasing strengths of competing binding sites, analogous to increasing Hill cooperativity [69]. This presents a potential strategy for controlling gene expression by tuning competitive binding sites to modulate transcriptional activity of heterologous circuits.

3. Optimized engineered parts

In some cases, the desired components have yet to be discovered, and those readily available cannot be used in a predictable manner. By tuning their parameters, one can engineer parts to exhibit desired functions. Strategies for this include using components from other organisms (**Figure 2C**) [70-73], directed evolution [26, 44, 74-77], rational design [78-81], and constructing de novo enzymes through computer-aided design [82-86]. For example, multiple studies have designed genetic devices to maintain modularity in synthetic circuits utilizing the fast reaction rates of phosphotransfers [87, 88]. One such device increased the capability for dynamic responses of a transcriptional cascade to temporal inputs in the presence of a titration effect [87]. In another example, Segall-Shapiro et. al. minimize toxicity of phage RNA polymerases by dividing the enzyme into four fragments [89]. These are co-expressed according to a resource allocator that sets the maximal transcriptional capacity available based on concentrations of a specific fragment [89]. These techniques share the common goal of generating components with optimal parameters.

Recent advances in genome editing techniques such as CRISPR facilitate the design and implementation of gene circuits [90-93]. For example, CRISPR has been used for targeted DNA degradation to prevent transfer of engineered DNA [94]. This function can be exploited to maintain genetic stability of gene circuits where plasmids with specific mutations are targeted for degradation [95]. CRISPR has also been used to construct gene circuits including Boolean-logic gates. Using dCas9 and guide RNAs, Nielsen et. al. designed a set of NOT gates with high on-target repression and minimal off-target interactions [96]. The authors combined these gates into complex circuits and coupled them to a native *E. coli* regulatory network to control sugar utilization, chemotaxis, and phage resistance. Finally, CRISPR-mediated cleavage can provide reliable regulation of multi-gene operons [97]. In this manner, promoters, RBSs, cis-regulatory elements, and riboregulators can also be combined into new operons with programmable gene expression [97]. RNA processing via CRISPR presents an innovative strategy for both improvement of existing gene circuits and bottom-up construction of new circuits.

4. Designs to account for stochastic dynamics

4.1 Coping with noise

Specific network motifs can resist effects of noise on circuit function (**Figure 2D**). For example, negative feedback with minimal time delay has been shown to reduce noise in gene expression [98, 99]. Specifically, negative feedback shifts noise towards higher frequencies then acts as a low-pass filter to prevent this noise from generating an output [100]. However, noise reduction is most effective for intermediate strengths of repression; above this, external noise from larger fluctuations in plasmid concentrations increases the variance of expression levels [101]. In contrast, positive feedback can either intensify or diminish the impact of noise depending on circuit parameters. Specifically, noise near an ultrasensitive transition can result in an abrupt shift between states, intensifying effects of noise [102]. In contrast, in a system with

hysteresis, the output is buffered against noisy inputs, which could otherwise cause spurious transitions between states [103].

Feedforward loops, which incorporate both positive and negative regulation, also have the capability to attenuate noise. Incoherent feedforward loops (IFFL) can filter out noise as a band-pass filter [104] and output gene expression is robust against varying copy numbers of the plasmid encoding the circuit [105]. Coherent feedforward loops (FFL) reject transient inputs and respond only to persistent stimuli [106]. However, additional control loops show diminishing returns, and adding a particular regulatory mechanism does not guarantee noise reduction [107]. Reaction rates constrain the ability for these motifs to attenuate noise, which at most decreases by a quartic root of the number of signal birth events [107]. Modulating parameters of these networks like gene copy number can also minimize intrinsic noise. This has been shown in *E. coli*, where having a high transcription rate coupled with a low translation rate is more likely to reduce noise than having a low transcription rate coupled with a high translation rate. [108].

Noise-resistant oscillators appear prevalent in natural systems despite constant extrinsic and intrinsic fluctuations. Design principles learned from circadian oscillations [109] and ultradian oscillations [110] have been used to minimize the impact of noise in synthetic circuits. A critical requirement for generating oscillations is negative feedback with sufficient time delay. Intertwining negative with positive feedback can enhance the tenability and noise resistance of these oscillations. Since the repressilator, several new oscillators have been engineered to generate more robust oscillations at the single cell level by incorporating these design principles [111, 112].

Aside from intracellular regulatory motifs, cell-cell communication represents another mechanism for noise reduction [113] or coordination of dynamics between single cells. In

multicellular organisms, intercellular coupling can be used to maintain synchronized oscillations, such as in periodic somite segmentation in vertebrate embryos [114]. Similarly, synthetic oscillators in bacteria using population-based strategies, such as quorum sensing, have exhibited robust, synchronous oscillations [31, 115-117].

4.2 Exploiting noise

Despite the stochastic nature of noise, this uncertainty can be utilized in the development of circuits with complex functions [118]. For example, noisy expression of transcription factors can be exploited to generate bistable gene expression without the need for cooperative binding [119]. At the population level, noise can be utilized to produce heterogeneous expression with a single circuit design. While undesirable for predictable circuit functions, this stochasticity is crucial for adaptation to changing environments [120, 121].

In addition, noise can be used to generate a wide dynamic range of stimulation, allowing for high-throughput characterization of circuit dynamics [122]. For example, noisy gene expression mediated by viral vectors revealed biphasic signaling dynamics of transcription factor E2F1 in mammalian cells [123]. This large variability coupled with single-cell analysis provided high-throughput quantification of the impact of Myc, another transcription factor, on E2F1. In addition, stochastic fluctuations associated with a particular genetic circuit can be used to estimate its kinetic parameters [124]. Population distributions from cell-to-cell variability can be used to distinguishing between different cancer types [125] and generate fingerprints to identify various network parameters [126]. Similarly, Slack et. al. demonstrate a method to characterize cellular phenotypes using their heterogeneous responses to perturbations [127].

Conclusion

Synthetic biology aims to create tunable systems that display predictable functions through the use of genetic circuits. However, unexpected failures plague circuit design, as even the best-characterized organisms remain far from being fully characterized. These failures have been attributed to biological uncertainty, which broadly encompasses any cellular process that cannot be predicted nor controlled directly.

While not all uncertainty can be removed, engineering strategies to confront potential pitfalls are desirable to advance the field. To this end, contextualizing different uncertainties is necessary towards establishing guiding principles for circuit design. Innovations in the field have provided design constraints and methods to limit the effect of uncertainty. We attempt to organize these strategies into different types and review their effectiveness in improving circuit function. Our categorization of biological uncertainty and its solutions is a first step in developing a general methodology for minimizing uncertainty in gene circuits.

Acknowledgments

We thank Hannah Meredith, Tatyana Sysoeva, Zhuojun Dai, and Kui Zhu for critical reading of the draft manuscript. This work was partially supported by the National Science Foundation (LY, CBET-0953202), the National Institutes of Health (LY, R01GM098642, R01GM110494), the Army Research Office (#W911NF-14-1-0490), a David and Lucile Packard Fellowship (LY), and the National Science Foundation Graduate Research Fellowship (CZ).

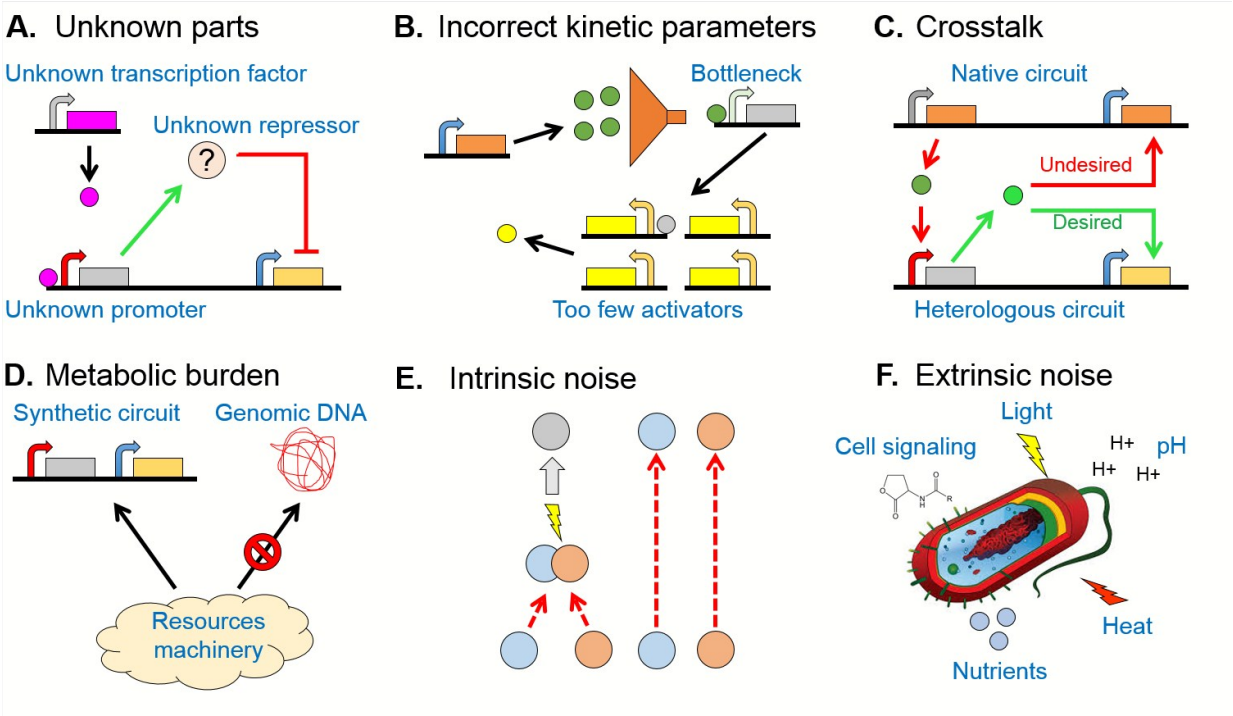

Figure 1. Sources of uncertainty that can lead to circuit failure. (A) Unknown parts. Many components are not known for the construction of circuits. **(B) Incorrect kinetic parameters.** With few components characterized, incorrect kinetic parameters can result in circuit failure or unintended circuit function. **(C) Crosstalk.** The lack of physical separation results in interactions with the potential to disrupt circuit function through crosstalk. **(D) Metabolic burden**. Exogenous synthetic circuits rely on the limited pool of endogenous cellular machinery to function, which can result in metabolic burden.**(E) Intrinsic noise**. Intrinsic noise can be attributed to stochastic chemical kinetics and the small size of cells, resulting in fluctuating metabolite concentrations. **(F) Extrinsic noise.** Extrinsic noise can be due to external perturbations such as the cell signaling molecules, light, pH, nutrients, or heat.

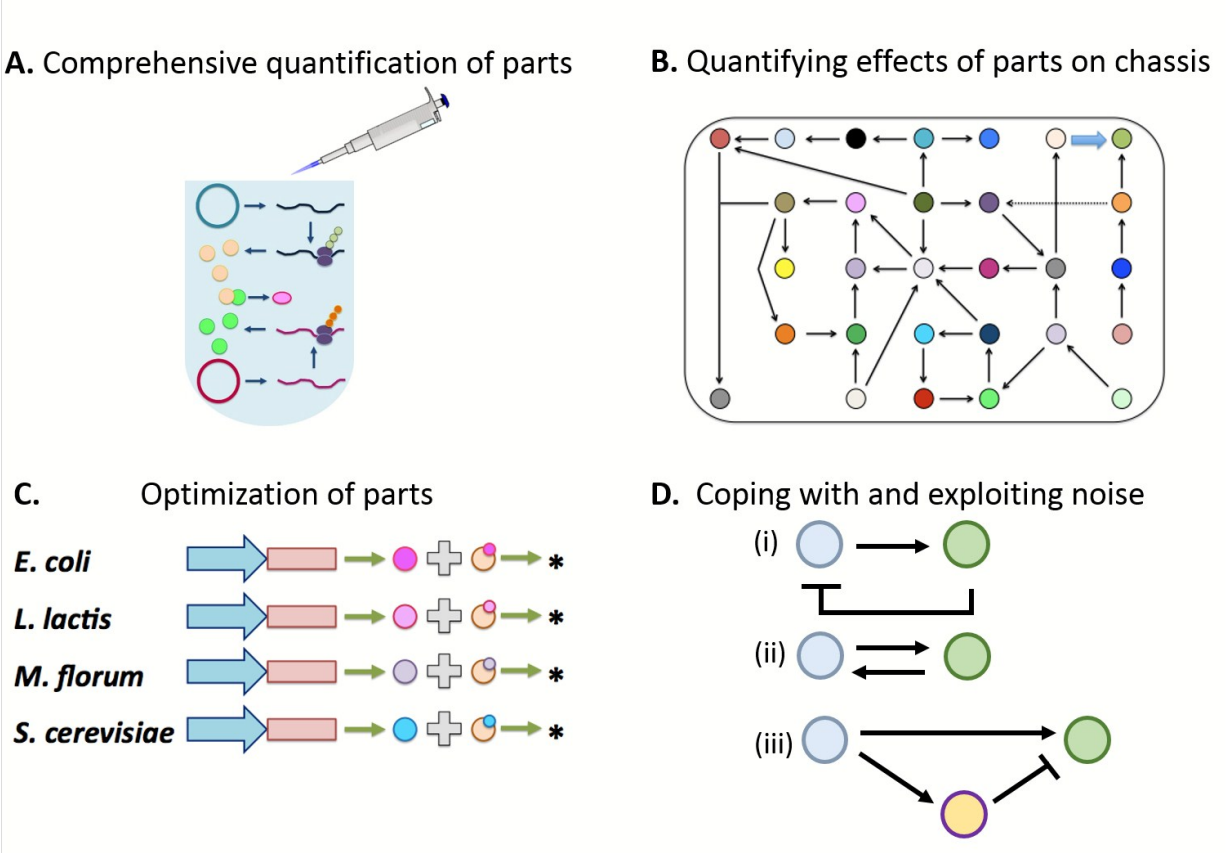

Figure 2. Strategies to address to biological uncertainty. **(A) Comprehensive quantification of parts.** The documentation of parts can alleviate the lack of components for circuit implementation. **(B) Quantifying effects of parts on chassis.** Understanding how different components impact the host can provide a method for reducing the genetic instability that results for parts that induce a high level of burden. **(C) Optimization of parts.** The use of orthogonal parts from different organisms can reduce the chances of cross-interactions with endogenous host machinery. **(D) Coping with and exploiting noise.** Specific network motifs have been shown to decrease the impact of noise on circuits. (i) Negative feedback (ii) Positive feedback (iii) Incoherent feedforward loop.

References

1. Sprinzak, D. and M.B. Elowitz, *Reconstruction of genetic circuits.* Nature, 2005. **438**(7067): p. 443-8.
2. Bar-Even, A., et al., *Design and analysis of synthetic carbon fixation pathways.* Proc Natl Acad Sci U S A, 2010. **107**(19): p. 8889-94.
3. Carothers, J.M., J.A. Goler, and J.D. Keasling, *Chemical synthesis using synthetic biology.* Curr Opin Biotechnol, 2009. **20**(4): p. 498-503.
4. Pardee, K., et al., *Paper-based synthetic gene networks.* Cell, 2014. **159**(4): p. 940-54.
5. Wroblewska, L., et al., *Mammalian synthetic circuits with RNA binding proteins for RNA-only delivery.* Nat Biotechnol, 2015. **33**(8): p. 839-41.
6. Chen, A.Y., C. Zhong, and T.K. Lu, *Engineering living functional materials.* ACS Synth Biol, 2015. **4**(1): p. 8-11.
7. Farzadfard, F. and T.K. Lu, *Synthetic biology. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations.* Science, 2014. **346**(6211): p. 1256272.
8. Yang, L., et al., *Permanent genetic memory with >1-byte capacity.* Nat Methods, 2014. **11**(12): p. 1261-6.
9. Slomovic, S. and J.J. Collins, *DNA sense-and-respond protein modules for mammalian cells.* Nat Methods, 2015.
10. Andrianantoandro, E., et al., *Synthetic biology: new engineering rules for an emerging discipline.* Mol Syst Biol, 2006. **2**: p. 2006 0028.
11. *Reliability Characterisation of Electrical and Electronic Systems*. 2015: Woodhead Publishing.
12. Tamsir, A., J.J. Tabor, and C.A. Voigt, *Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'.* Nature, 2011. **469**(7329): p. 212-5.
13. Fussenegger, M., et al., *Autoregulated multicistronic expression vectors provide one-step cloning of regulated product gene expression in mammalian cells.* Biotechnol Prog, 1997. **13**(6): p. 733-40.
14. Greber, D. and M. Fussenegger, *Mammalian synthetic biology: engineering of sophisticated gene networks.* J Biotechnol, 2007. **130**(4): p. 329-45.
15. Haellman, V. and M. Fussenegger, *Synthetic Biology-Toward Therapeutic Solutions.* J Mol Biol, 2015.
16. Moon, T.S., et al., *Genetic programs constructed from layered logic gates in single cells.* Nature, 2012. **491**(7423): p. 249-53.
17. Malyshev, D.A., et al., *A semi-synthetic organism with an expanded genetic alphabet.* Nature, 2014. **509**(7500): p. 385-8.
18. Gorgens, J.F., et al., *The metabolic burden of the PGK1 and ADH2 promoter systems for heterologous xylanase production by Saccharomyces cerevisiae in defined medium.* Biotechnol Bioeng, 2001. **73**(3): p. 238-45.
19. van Rensburg, E., et al., *The metabolic burden of cellulase expression by recombinant Saccharomyces cerevisiae Y294 in aerobic batch culture.* Appl Microbiol Biotechnol, 2012. **96**(1): p. 197-209.
20. Tan, C., P. Marguet, and L. You, *Emergent bistability by a growth-modulating positive feedback circuit.* Nat Chem Biol, 2009. **5**(11): p. 842-8.

21. Karim, A.S., K.A. Curran, and H.S. Alper, *Characterization of plasmid burden and copy number in Saccharomyces cerevisiae for optimization of metabolic engineering applications.* FEMS Yeast Res, 2013. **13**(1): p. 107-16.
22. Scott, M., et al., *Interdependence of cell growth and gene expression: origins and consequences.* Science, 2010. **330**(6007): p. 1099-102.
23. Kimelman, A., et al., *A vast collection of microbial genes that are toxic to bacteria.* Genome Res, 2012. **22**(4): p. 802-9.
24. Cammue, B.P., et al., *Gene-encoded antimicrobial peptides from plants.* Ciba Found Symp, 1994. **186**: p. 91-101; discussion 101-6.
25. Mygind, P.H., et al., *Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus.* Nature, 2005. **437**(7061): p. 975-980.
26. Saida, F., *Overview on the expression of toxic gene products in Escherichia coli.* Curr Protoc Protein Sci, 2007. **Chapter 5**: p. Unit 5 19.
27. Champion, K. and N.P. Higgins, *Growth rate toxicity phenotypes and homeostatic supercoil control differentiate Escherichia coli from Salmonella enterica serovar Typhimurium.* J Bacteriol, 2007. **189**(16): p. 5839-49.
28. Montini, E., et al., *The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy.* J Clin Invest, 2009. **119**(4): p. 964-75.
29. Sleight, S.C., et al., *Designing and engineering evolutionary robust genetic circuits.* J Biol Eng, 2010. **4**: p. 12.
30. Wu, F., D.J. Menn, and X. Wang, *Quorum-sensing crosstalk-driven synthetic circuits: from unimodality to trimodality.* Chem Biol, 2014. **21**(12): p. 1629-38.
31. Marguet, P., et al., *Oscillations by minimal bacterial suicide circuits reveal hidden facets of host-circuit physiology.* PLoS One, 2010. **5**(7): p. e11909.
32. Paulsson, J., *Summing up the noise in gene networks.* Nature, 2004. **427**(6973): p. 415-8.
33. Ferguson, M.L., et al., *Reconciling molecular regulatory mechanisms with noise patterns of bacterial metabolic promoters in induced and repressed states.* Proc Natl Acad Sci U S A, 2012. **109**(1): p. 155-60.
34. Tsimring, L.S., *Noise in biology.* Rep Prog Phys, 2014. **77**(2): p. 026601.
35. Swain, P.S., M.B. Elowitz, and E.D. Siggia, *Intrinsic and extrinsic contributions to stochasticity in gene expression.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12795-800.
36. Raser, J.M. and E.K. O'Shea, *Control of stochasticity in eukaryotic gene expression.* Science, 2004. **304**(5678): p. 1811-4.
37. Hilfinger, A., M. Chen, and J. Paulsson, *Using temporal correlations and full distributions to separate intrinsic and extrinsic fluctuations in biological systems.* Phys Rev Lett, 2012. **109**(24): p. 248104.
38. Elowitz, M.B., et al., *Stochastic gene expression in a single cell.* Science, 2002. **297**(5584): p. 1183-6.
39. Elowitz, M.B. and S. Leibler, *A synthetic oscillatory network of transcriptional regulators.* Nature, 2000. **403**(6767): p. 335-8.
40. Tanouchi, Y., et al., *A noisy linear map underlies oscillations in cell size and gene expression in bacteria.* Nature, 2015. **523**(7560): p. 357-60.

41. Chen, Y.J., et al., *Characterization of 582 natural and synthetic terminators and quantification of their design constraints.* Nat Methods, 2013. **10**(7): p. 659-64.
42. Redden, H. and H.S. Alper, *The development and characterization of synthetic minimal yeast promoters.* Nat Commun, 2015. **6**: p. 7810.
43. Jensen, P.R. and K. Hammer, *Artificial promoters for metabolic optimization.* Biotechnol Bioeng, 1998. **58**(2-3): p. 191-5.
44. Ellis, T., X. Wang, and J.J. Collins, *Diversity-based, model-guided construction of synthetic gene networks with predicted functions.* Nat Biotechnol, 2009. **27**(5): p. 465-71.
45. Alper, H., et al., *Tuning genetic control through promoter engineering.* Proc Natl Acad Sci U S A, 2005. **102**(36): p. 12678-83.
46. Davis, J.H., A.J. Rubin, and R.T. Sauer, *Design, construction and characterization of a set of insulated bacterial promoters.* Nucleic Acids Res, 2011. **39**(3): p. 1131-41.
47. Rudge, T.J., et al., *Characterization of intrinsic properties of promoters.* ACS Synth Biol, 2015.
48. Rhodius, V.A., et al., *Design of orthogonal genetic switches based on a crosstalk map of sigmas, anti-sigmas, and promoters.* Mol Syst Biol, 2013. **9**: p. 702.
49. Tornoe, J., et al., *Generation of a synthetic mammalian promoter library by modification of sequences spacing transcription factor binding sites.* Gene, 2002. **297**(1-2): p. 21-32.
50. Silva, J.M., et al., *Second-generation shRNA libraries covering the mouse and human genomes.* Nat Genet, 2005. **37**(11): p. 1281-8.
51. Greener, A., S.M. Lehman, and D.R. Helinski, *Promoters of the broad host range plasmid RK2: analysis of transcription (initiation) in five species of gram-negative bacteria.* Genetics, 1992. **130**(1): p. 27-36.
52. Stech, M., et al., *A continuous-exchange cell-free protein synthesis system based on extracts from cultured insect cells.* PLoS One, 2014. **9**(5): p. e96635.
53. Jun, S.Y., S.H. Kang, and K.H. Lee, *Continuous-exchange cell-free protein synthesis using PCR-generated DNA and an RNase E-deficient extract.* Biotechniques, 2008. **44**(3): p. 387-91.
54. Murtas, G., et al., *Protein synthesis in liposomes with a minimal set of enzymes.* Biochem Biophys Res Commun, 2007. **363**(1): p. 12-7.
55. Kuruma, Y., et al., *A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells.* Biochim Biophys Acta, 2009. **1788**(2): p. 567-74.
56. Kobori, S., et al., *A controllable gene expression system in liposomes that includes a positive feedback loop.* Mol Biosyst, 2013. **9**(6): p. 1282-5.
57. Noireaux, V. and A. Libchaber, *A vesicle bioreactor as a step toward an artificial cell assembly.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(51): p. 17669-17674.
58. Lopatkin, A.J. and L. You, *Synthetic biology looks good on paper.* Cell, 2014. **159**(4): p. 718-20.
59. Gillespie, D.T., *Exact stochastic simulation of coupled chemical reactions.* The Journal of Physical Chemistry, 1977. **81**(25): p. 2340-2361.
60. Gillespie, D.T., *Stochastic simulation of chemical kinetics.* Annu Rev Phys Chem, 2007. **58**: p. 35-55.

61. Neubauer, P., H.Y. Lin, and B. Mathiszik, *Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in Escherichia coli.* Biotechnol Bioeng, 2003. **83**(1): p. 53-64.
62. Purcell, O., et al., *Towards a whole-cell modeling approach for synthetic biology.* Chaos, 2013. **23**(2): p. 025112.
63. Weisse, A.Y., et al., *Mechanistic links between cellular trade-offs, gene expression, and growth.* Proc Natl Acad Sci U S A, 2015. **112**(9): p. E1038-47.
64. Scott, M., et al., *Emergence of robust growth laws from optimal regulation of ribosome synthesis.* Mol Syst Biol, 2014. **10**: p. 747.
65. Maitra, A. and K.A. Dill, *Bacterial growth laws reflect the evolutionary importance of energy efficiency.* Proc Natl Acad Sci U S A, 2015. **112**(2): p. 406-11.
66. Keren, L., et al., *Promoters maintain their relative activity levels under different growth conditions.* Mol Syst Biol, 2013. **9**: p. 701.
67. Ceroni, F., et al., *Quantifying cellular capacity identifies gene expression designs with reduced burden.* Nat Methods, 2015. **12**(5): p. 415-8.
68. Dahl, R.H., et al., *Engineering dynamic pathway regulation using stress-response promoters.* Nat Biotechnol, 2013. **31**(11): p. 1039-46.
69. Brewster, R.C., et al., *The transcription factor titration effect dictates level of gene expression.* Cell, 2014. **156**(6): p. 1312-23.
70. Cameron, D.E. and J.J. Collins, *Tunable protein degradation in bacteria.* Nat Biotechnol, 2014. **32**(12): p. 1276-81.
71. Grilly, C., et al., *A synthetic gene network for tuning protein degradation in Saccharomyces cerevisiae.* Mol Syst Biol, 2007. **3**: p. 127.
72. Holland, A.J., et al., *Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells.* Proc Natl Acad Sci U S A, 2012. **109**(49): p. E3350-7.
73. Stanton, B.C., et al., *Systematic transfer of prokaryotic sensors and circuits to mammalian cells.* ACS Synth Biol, 2014. **3**(12): p. 880-91.
74. Tang, S.Y., H. Fazelinia, and P.C. Cirino, *AraC regulatory protein mutants with altered effector specificity.* J Am Chem Soc, 2008. **130**(15): p. 5267-71.
75. Lee, S.K., et al., *Directed evolution of AraC for improved compatibility of arabinose- and lactose-inducible promoters.* Appl Environ Microbiol, 2007. **73**(18): p. 5711-5.
76. Cobb, R.E., N. Sun, and H. Zhao, *Directed Evolution as a Powerful Synthetic Biology Tool.* Methods (San Diego, Calif.), 2013. **60**(1): p. 81-90.
77. Mishima, N., et al., *Insertion of stabilizing loci in vectors of T7 RNA polymerase-mediated Escherichia coli expression systems: a case study on the plasmids involving foreign phospholipase D gene.* Biotechnol Prog, 1997. **13**(6): p. 864-8.
78. Lutz, S., *Beyond directed evolution--semi-rational protein engineering and design.* Curr Opin Biotechnol, 2010. **21**(6): p. 734-43.
79. Pavelka, A., E. Chovancova, and J. Damborsky, *HotSpot Wizard: a web server for identification of hot spots in protein engineering.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W376-83.
80. Kuipers, R.K., et al., *3DM: systematic analysis of heterogeneous superfamily data to discover protein functionalities.* Proteins, 2010. **78**(9): p. 2101-13.

81. Li, Y., et al., *A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments.* Nat Biotechnol, 2007. **25**(9): p. 1051-6.
82. Mandell, D.J. and T. Kortemme, *Computer-aided design of functional protein interactions.* Nat Chem Biol, 2009. **5**(11): p. 797-807.
83. Chevalier, B.S., et al., *Design, activity, and structure of a highly specific artificial endonuclease.* Mol Cell, 2002. **10**(4): p. 895-905.
84. Karanicolas, J., et al., *A de novo protein binding pair by computational design and directed evolution.* Mol Cell, 2011. **42**(2): p. 250-60.
85. Kortemme, T., et al., *Computational redesign of protein-protein interaction specificity.* Nat Struct Mol Biol, 2004. **11**(4): p. 371-9.
86. Siegel, J.B., et al., *Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction.* Science, 2010. **329**(5989): p. 309-13.
87. Mishra, D., et al., *A load driver device for engineering modularity in biological networks.* Nat Biotechnol, 2014. **32**(12): p. 1268-75.
88. Del Vecchio, D., A.J. Ninfa, and E.D. Sontag, *Modular cell biology: retroactivity and insulation.* Mol Syst Biol, 2008. **4**: p. 161.
89. Segall-Shapiro, T.H., et al., *A 'resource allocator' for transcription based on a highly fragmented T7 RNA polymerase.* Mol Syst Biol, 2014. **10**: p. 742.
90. Dahlman, J.E., et al., *Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease.* Nat Biotechnol, 2015. **33**(11): p. 1159-1161.
91. Nissim, L., et al., *Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells.* Mol Cell, 2014. **54**(4): p. 698-710.
92. Polstein, L.R. and C.A. Gersbach, *A light-inducible CRISPR-Cas9 system for control of endogenous gene activation.* Nat Chem Biol, 2015. **11**(3): p. 198-200.
93. Wright, A.V., et al., *Rational design of a split-Cas9 enzyme complex.* Proc Natl Acad Sci U S A, 2015. **112**(10): p. 2984-9.
94. Caliando, B.J. and C.A. Voigt, *Targeted DNA degradation using a CRISPR device stably carried in the host genome.* Nat Commun, 2015. **6**: p. 6989.
95. Moore, R., et al., *CRISPR-based self-cleaving mechanism for controllable gene delivery in human cells.* Nucleic Acids Res, 2015. **43**(2): p. 1297-303.
96. Nielsen, A.A. and C.A. Voigt, *Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks.* Mol Syst Biol, 2014. **10**: p. 763.
97. Qi, L., et al., *RNA processing enables predictable programming of gene expression.* Nat Biotechnol, 2012. **30**(10): p. 1002-6.
98. Becskei, A. and L. Serrano, *Engineering stability in gene networks by autoregulation.* Nature, 2000. **405**(6786): p. 590-3.
99. Shimoga, V., et al., *Synthetic mammalian transgene negative autoregulation.* Mol Syst Biol, 2013. **9**: p. 670.
100. Austin, D.W., et al., *Gene network shaping of inherent noise spectra.* Nature, 2006. **439**(7076): p. 608-11.
101. Dublanche, Y., et al., *Noise in transcription negative feedback loops: simulation and experimental analysis.* Mol Syst Biol, 2006. **2**: p. 41.
102. Hasty, J., et al., *Noise-based switches and amplifiers for gene expression.* Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2075-80.

103. Becskei, A., B. Seraphin, and L. Serrano, *Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion.* EMBO J, 2001. **20**(10): p. 2528-35.
104. Guantes, R., J. Estrada, and J.F. Poyatos, *Trade-offs and noise tolerance in signal detection by genetic circuits.* PLoS One, 2010. **5**(8): p. e12314.
105. Bleris, L., et al., *Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template.* Mol Syst Biol, 2011. **7**: p. 519.
106. Ghosh, B., R. Karmakar, and I. Bose, *Noise characteristics of feed forward loops.* Phys Biol, 2005. **2**(1): p. 36-45.
107. Lestas, I., G. Vinnicombe, and J. Paulsson, *Fundamental limits on the suppression of molecular fluctuations.* Nature, 2010. **467**(7312): p. 174-8.
108. Ozbudak, E.M., et al., *Regulation of noise in the expression of a single gene.* Nat Genet, 2002. **31**(1): p. 69-73.
109. Mihalcescu, I., W. Hsing, and S. Leibler, *Resilient circadian oscillator revealed in individual cyanobacteria.* Nature, 2004. **430**(6995): p. 81-5.
110. Isomura, A. and R. Kageyama, *Ultradian oscillations and pulses: coordinating cellular responses and cell fate decisions.* Development, 2014. **141**(19): p. 3627-36.
111. Stricker, J., et al., *A fast, robust and tunable synthetic gene oscillator.* Nature, 2008. **456**(7221): p. 516-9.
112. Tigges, M., et al., *A tunable synthetic mammalian oscillator.* Nature, 2009. **457**(7227): p. 309-12.
113. Tanouchi, Y., et al., *Noise reduction by diffusional dissipation in a minimal quorum sensing motif.* PLoS Comput Biol, 2008. **4**(8): p. e1000167.
114. Horikawa, K., et al., *Noise-resistant and synchronized oscillation of the segmentation clock.* Nature, 2006. **441**(7094): p. 719-23.
115. Balagadde, F.K., et al., *Long-term monitoring of bacteria undergoing programmed population control in a microchemostat.* Science, 2005. **309**(5731): p. 137-40.
116. Chen, Y., et al., *SYNTHETIC BIOLOGY. Emergent genetic oscillations in a synthetic microbial consortium.* Science, 2015. **349**(6251): p. 986-9.
117. Danino, T., et al., *A synchronized quorum of genetic clocks.* Nature, 2010. **463**(7279): p. 326-30.
118. Eldar, A. and M.B. Elowitz, *Functional roles for noise in genetic circuits.* Nature, 2010. **467**(7312): p. 167-73.
119. To, T.L. and N. Maheshri, *Noise can induce bimodality in positive transcriptional feedback loops without bistability.* Science, 2010. **327**(5969): p. 1142-5.
120. Balazsi, G., A. van Oudenaarden, and J.J. Collins, *Cellular decision making and biological noise: from microbes to mammals.* Cell, 2011. **144**(6): p. 910-25.
121. Wu, M., et al., *Engineering of regulated stochastic cell fate determination.* Proc Natl Acad Sci U S A, 2013. **110**(26): p. 10610-5.
122. Li, B. and L. You, *Predictive power of cell-to-cell variability.* Quantitative Biology, 2013. **1**(2): p. 131-139.
123. Wong, J.V., et al., *Viral-mediated noisy gene expression reveals biphasic E2f1 response to MYC.* Mol Cell, 2011. **41**(3): p. 275-85.
124. Cox, C.D., et al., *Using noise to probe and characterize gene circuits.* Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10809-14.

125. Li, B. and L. You, *Stochastic sensitivity analysis and kernel inference via distributional data.* Biophys J, 2014. **107**(5): p. 1247-55.
126. Munsky, B., B. Trinh, and M. Khammash, *Listening to the noise: random fluctuations reveal gene network parameters.* Mol Syst Biol, 2009. **5**: p. 318.
127. Slack, M.D., et al., *Characterizing heterogeneous cellular responses to perturbations.* Proc Natl Acad Sci U S A, 2008. **105**(49): p. 19306-11.

Table of Contents Entry:

We discuss biological uncertainties that complicate predictable engineering of gene circuits and potential strategies to address these uncertainties.

Graphic