Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

A wise consistency: engineering biology for conformity, reliability, predictability

Adam Paul Arkin[1,2]

The next generation of synthetic biology applications will increasingly involve engineered organisms that exist in intimate contact with humans, animals and the rest of the environment. Examples include cellular and viral approaches for maintaining and improving health in humans and animals. The need for reliable and specific function in these environments may require more complex system designs than previously. In these cases the uncertainties in the behavior of biological building blocks, their hosts and their environments present a challenge for design of predictable and safe systems. Here, we review systematic methods for the effective characterization of these uncertainties that are lowering the barriers to predictive design of reliable complex biological systems.

Addresses
[1] Department of Bioengineering, University of California, 2151 Berkeley Way, Berkeley, CA 94704-5230, United States
[2] Physical Biosciences Division, E. O. Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Mailstop 955-512L, Berkeley, CA 94720, United States

Corresponding author: Arkin, Adam Paul (aparkin@lbl.gov)

Current Opinion in Chemical Biology 2013, **17**:893–901

This review comes from a themed issue on **Synthetic biology**

Edited by **Adam P Arkin** and **Martin Fussenegger**

For a complete overview see the Issue and the Editorial

Available online 20th November 2013

1367-5931

http://dx.doi.org/10.1016/j.cbpa.2013.09.012

Introduction

''A foolish consistency is the hobgoblin of little minds'' -Ralph Waldo Emerson, in his essay ''Self-Reliance''

Aspects that differentiate synthetic biology from other fields of molecular biotechnology are the ambition to formalize and scale the complexity of design of new function in biology, and for such designs to reliably and predictably operate as specified. The application areas preexist the field: biosynthesis of valuable chemicals for materials and medicine; production of plants for food, energy and ecological control; engineering of genetic, viral and cellular approaches for health maintenance and improvement; microbial communities for soil and water improvement; and many others.

The areas in which design of predictable and reliable complex biological function is likely to be most important involve engineering biology for applications in the less controlled conditions that obtain beyond the bioreactor such as viral and cellular therapies for medicine or microbial and plant applications for agriculture. Yet these are the applications most in need of synthetic biology, at least according to a recent report of the World Economic Forum put forward an analysis of global risks [1,2].

These applications involve engineered organisms that exist in intimate contact with humans, animals and the rest of the environment. As such, issues of reliability and trust become paramount in addition to the effect of the technology. Reliability and predictability are central not only to trust between technologists and society wherein risk needs a rational actuarial basis but also among the technologists themselves. One designer must trust that reusable systems designed by another will operate as advertised.

Ten years ago, the most immediate barriers to an efficient design-build-test cycle were finding the proper biological parts, cloning and/or synthesizing them, and assembling and inserting them into cells. While these barriers remain, their heights have been significantly lowered by innovations in DNA sequencing, synthesis, assembly and scaling functional assays. The combination is enabling rapid creation and screening of many variants of a design. For some applications it is now possible to screen large libraries for the proper pathway and host variations to produce a target molecule to a given level with increasing efficacy. However, many applications are complex enough that this is not an option. The initial designs must be implemented with parts that work predictably enough to produce systems with that function very close to specification, and safely, so that there is minimal need for testing many variants semi-randomly. Here, the barriers concern the unpredictable operation of biological parts in different contexts — that is, in different configurations with other parts, in different hosts and in different environments. We will start by reviewing a

few key emerging complex biomedical applications that are aimed squarely beyond the bioreactor then describe systematic approaches to achieving reliable function despite variable context.

Example applications from present and future synthetic biological medicine

While all applications can benefit from more predictable operation of synthetic biological systems in deployment environments, few applications challenge this possibility like those in medicine. There have been some startling successes in using organisms as medicine. These include adoptive immunotherapy with engineered T-cells to cure certain types of cancer [3•,4], engineered bacteria and oncolytic viruses for cancer [5,6], viral gene therapy for blindness [7,8] and hemophilia [9], and fecal transplants that harbinger designed communities for inflammation [10,11]. In some cases, the success of these applications might argue that there is not a need for complex design — that a combination of finding the correct natural starting points and modest modifications for our own purposes will be sufficient.

However, as increasing specificity and long term reliability are needed, more sophisticated designs are being proposed. For example, Xie et al. demonstrated a multi-input RNAi logic circuit to be delivered as a gene therapy that would very specifically determine if an infected cell were a particular cancer type only then deliver a molecular therapeutic [12]. Anderson and colleagues built up several steps toward the bottom-up design of a tumor-destroying bacterium that, theoretically, would specifically invade target tumor cells after successful aggregation in the tumor necrotic region, then escape the vacuole and deliver a therapy to the cytosol or nucleus of the target cell [13–15]. Other complicated designs involve sophisticated control systems, composed from logic gates, oscillators and feedback systems, for homeostatic stem-cell differentiation into islet-like cells for diabetes [16•] or designs for what amount to viral parasites that interfere with the propagation of HIV inside hosts with implications for in-host viral quasispecies competitions and transmission of the engineered virus [17–19,20•]. None of these are fully working applications as yet. Clearly, with more 'moving parts', needs for high specificity of function, and persistence in complex competitive environments, they have been harder to implement and these designs would benefit from a degree of trustworthy engineering beyond what we can currently deliver effectively.

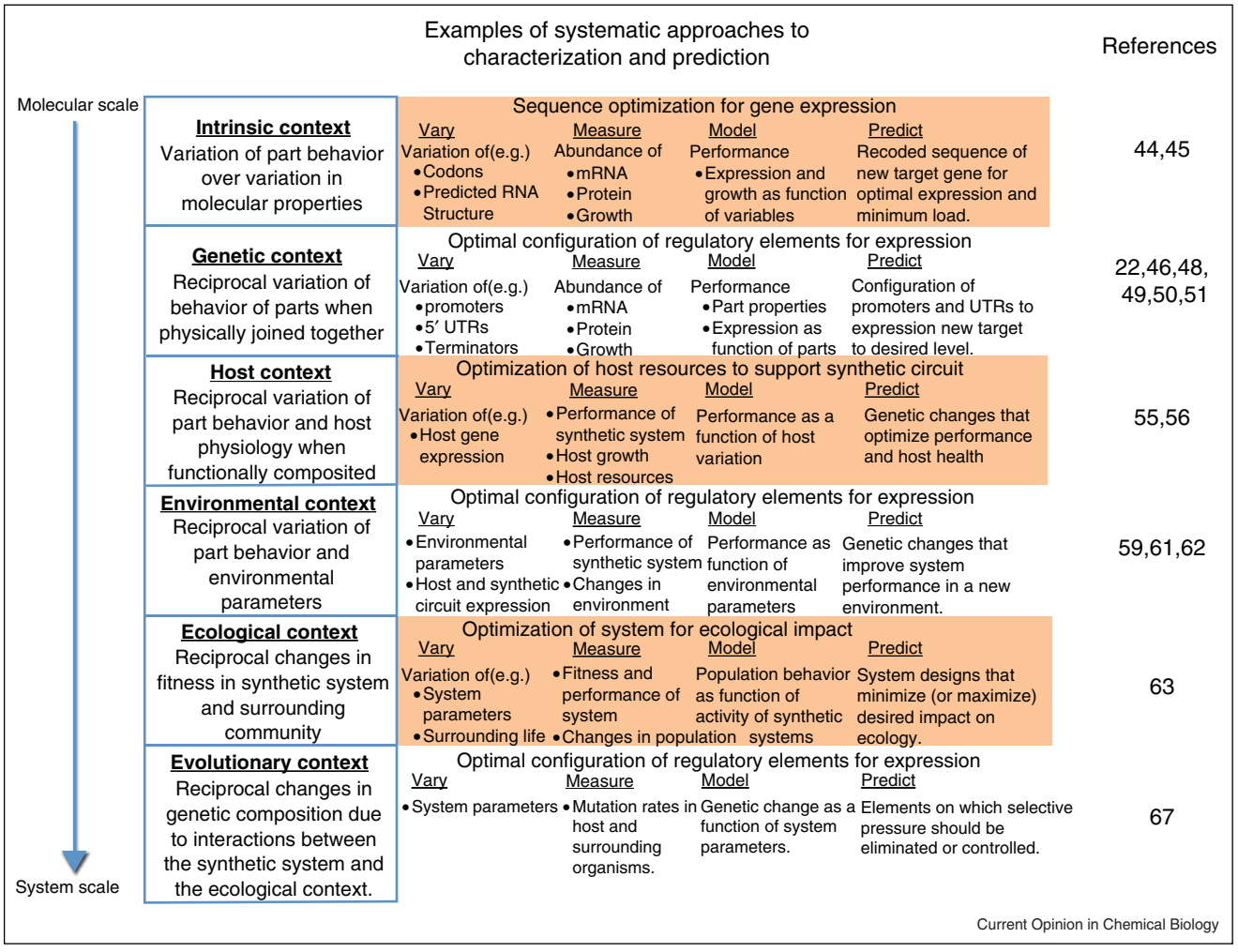
Quantifying reliable function across contexts

Most skepticism of the synthetic biology agenda stems from the criticism that there is too much unknown about the biological system to be engineered and the effects of and on the environment in which it is to be deployed for a predictable engineering approach to be possible. While it is likely true that the levels of uncertainty in biological engineering will be larger than in any other engineering discipline, we argue that it is not a hopeless venture and systematization of the field will enable predictably functioning designs.

One of the controversial tenets of some synthetic biologists is that a reliable engineering field rests, at least in part, on the community agreeing to use well-characterized and 'standardized' parts and hosts. We, and others, have reviewed why this is so elsewhere and outlined much of the desiderata for such parts including tunability, orthogonality, scalability and more [21]. For gene expression in particular there has been an efflorescence of such families of standardized parts or modular strategies for creation of scalable functional regulators. Most of these affect transcription or translation initiation [22••,23–26] or elongation [27–31] though emerging standards are beginning to include elements that mediate transcriptional termination [32,33], orthogonal protein–protein interactions for controlling metabolic pathway flux [34] and signaling [35] and targeted elements for controlling transcript [36] and protein degradation. The results of these have been the ability to predictably create circuits of increasing complexity but even these remain relatively small (2–5 input logic gates and memory circuits [37,38•,39,40]). Ideally, each of these families provides not only building blocks for complex circuits but also represents controlled variations of key performance variables, such as promoter strength, that can be used in formal design-of-experiment protocols to rationally search a parameter space for optimal function [41]. Since the behavior of even these small circuits can be sensitive to changes in media/environment, host background, and configuration of elements on a replicon, characterization of their variable behavior across contexts is necessary.

We, and others, have begun to define and dissect the different areas of uncertainty, context effects, and design approaches to characterize and control their effects [42,43]. In Figure 1, we define six distinct levels of context effect (intrinsic, genetic, host, environmental, ecological, and evolutionary). Below, we review systematic approaches to characterize and design against these effects. We choose, though, to leave out the study of intrinsic context since this can be fairly specific to the molecule involved. However, issues such as methods for sequence optimization for expression control [44], standard elements for affecting molecular folding and solubility [45], and another of other innovations in molecular engineering to affect transport, degradation, and activity are becoming more standard and are worthy of a review of their own. For the others, we focus on systematic methods that aim to elucidate and control general mechanisms of context effects or provide enough data that models can aid in predictable design.

Figure 1



Contexts effects and approaches to their characterization. Context effects are unintended interactions among elements of a synthetic biological system, their host, and their surrounding environment. Here we present one possible categorical breakdown of these effects and outline the approaches to their characterization.

Systematic design and quantification of genetic context

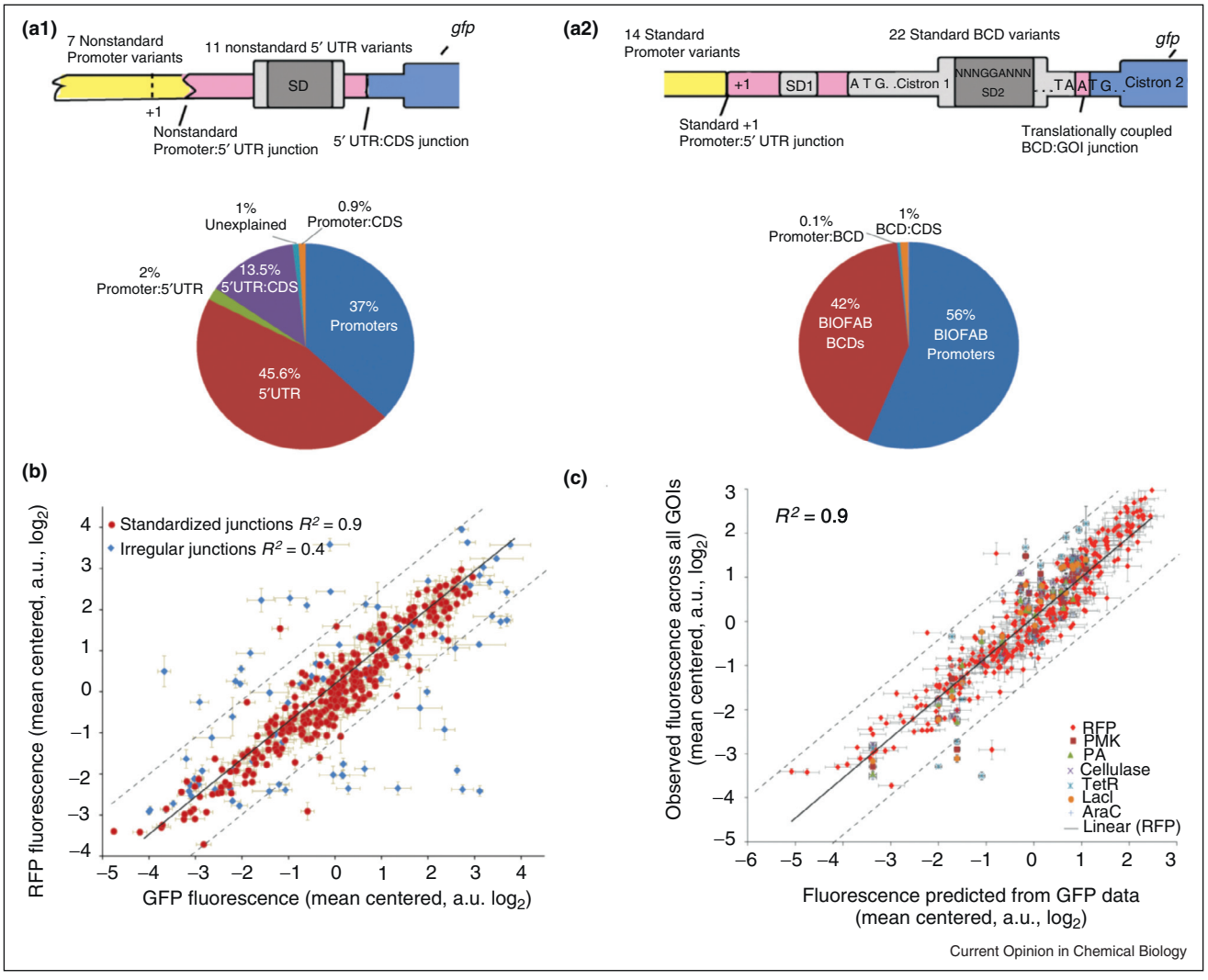
The genetic context of a part comprises those mechanisms that change the key properties of a biological part when it is physically interconnected on the same molecule. For example, the expression of an open reading frame is affected by the presence of a promoter upstream of it, but it is also affected by local DNA structure, epigenetic marks, and structural interactions of its RNA with other elements encoded on the transcript. These interactions are reciprocal and the insertion of an ORF can affect the function of surrounding elements [42,46••].

Recently, systematic approaches to quantify and control these sorts of interactions in the bacterium *Escherichia coli* have emerged. Salis et al. developed the ribosome binding site calculator, a method based on thermodynamic structure predictions of interactions among the ribosome its binding sequence and the local structure around the gene start, to predict 5′UTR and coding sequence variants that will yield a desired relative expression level [47,48]. While very useful, this method still has a wide amount of variability in prediction and does not permit reuse of standard translation initiation elements. Kosuri et al. recently demonstrated the use of large scale gene synthesis to explore over twelve thousand combinations of promoters and 5′UTRS driving gene expression and measured the variable effects of mRNA production, stability and translation [49]. They confirm the importance RNA structural interactions and argue that using this technology one can simply screen for the desired expression level. However, when the designed circuit becomes large such screening would become prohibitively costly. In a complementary approach,

Mutalik et al. performed a factorial ANOVA analysis of a controlled library of widely used but nonstandardized promoters, 5′UTRs, and genes to quantify the intrinsic strengths and context variability of these elements and showed that all elements interacted with each other to affect mRNA and protein production (Figure 2A.1) [46••]. They showed that certain elements were less genetically context sensitive than others (a measure of part quality).

Mutalik et al. then showed how embedding variants of a Shine–Dalgarno sequence inside a short cistron translated just upstream of a target sequence breaks up RNA structures could lead to highly predictable expression across a number of genes (an effect amplified by also using standardized promoters with defined + 1 locations) (Figure 2A.2) [22••]. These highly controlled junctions between standard regulatory elements improved the R^2 of the correlation between the relative expression of

Figure 2



An example of characterization and control of genetic context effects. **(A)** In one study, it was determined that use of nonstandardized promoters and 5′-UTRs that affect translation initiation and transcript stability to drive expression of different genes led to unintended interactions that affected the apparent properties of each element. An ANOVA analysis, a result of which is shown in the pie chart, indicates the amount that each part and the interactions among them affects expression of a target gene. Pie chart is derived from Figure 3 of Ref. [46••]. **(B)** In a follow-on study, design and use of standardized genetic parts to express different genes led to virtual elimination of unintended interactions as shown by a pie-chart similar to that in (A) (derived from Figure 4e from Ref. [22••]). **(C)** Correlation of expression of different fluorescent proteins expressed from the same promoter and UTR combinations. Standardized parts and junctions between them (as shown in (B)) lead to far reliable gene expression across a diversity of gene types (derived from Figure 4d from Ref. [22••]). **(D)** Predicted versus observed expression of a diversity of genes based on models of standard expression parts shows ~93% chance to obtain an expected normalized relative expression for a given gene to within two-fold of a target level (derived from Figure 4f from Ref. [22••]).

different genes driven by the same promoter/UTR combinations from 0.4 to 0.9 (Figure 2B). The method achieves an ~93% chance to obtain an expected normalized relative expression for a given gene to within two-fold of a target level, which represents an ~87% reduction in forward-engineering expression error compared to the error rates of previously best available methods (Figure 2C). Along similar lines, Qi et al. used a CRISPR-associated RNA cleavage protein [50] and Lou et al. used a ribozyme [51] to create controlled, physically separated blocks on the transcript to remove structural interactions on the transcript and improve predictable function of regulatory and gene encoding elements therein. With the CRISPR-protein csy4, Qi et al. showed improved predictability of expression of genes in different positions in an operon [50] and Liu et al. showed composition of multiple regulatory elements to create a 4-input NOR gate on a single transcript [38•].

Systematic quantification of host context

Any addition of replicable DNA to a host cell necessarily impacts the host's physiology. There is at least a small effect of carrying and replicating this DNA. It might disrupt local replicon structures changing the expression of neighboring genes, and the activities encoded in the DNA might affect host physiology through competition for resources, interference with other host biomolecules, and designed interactions. Reciprocally, the ability to express heterologous DNA is dependent on possibly variant host resources, and expressed function might be dependent on particular host subsystems that may vary thereby affecting designed function. The load effects can change the fitness of the synthetic system thereby coupling to issues with evolutionary context.

Metabolic engineers have long dealt with specific issues of host interaction including cofactor and carbon flux balancing to ensure host growth while maximizing flux to a pathway of interest. Designers of regulation have begun to consider, for example, the asymmetrical load between the ON and OFF state of genetic switches which can lead to undesirable growth differences of cells in the different switch states. New switch designs using DNA inversion, for example, can maintain symmetrical low-load ON and OFF states leading to increased fitness of the host and longer-times to mutational failure [52,53,54].

Recently, approaches for systematically discovering unknown mechanisms of host context have been developed to supplement these knowledge-based approaches. Cardinale et al. show that variation in cloning strain background can affect expression of a three gene probe cassette in *E. coli* that is largely explainable by changes in host growth and ribosomal availability (Figure 3A) but that when that same cassette is passed into 88 deletion strains of *E. coli* BW25113 there seem to be more specific effects of each gene deletion on circuit performance (Figure 3B) [55••]. Specific metabolic and signaling genes, when deleted had large positive and negative effects (respectively) on expression of all three fluorescent proteins of the probe while a couple differentially affected expression of at least one of the proteins. Key subsystems that generically and specifically affect heterologous circuit function were thereby identified and mapped to subelements of the synthetic circuit. In a complementary approach, Woodruff et al. [56] created a library of millions of overexpressed genome fragments in an ethanol production strain and subjected it to a growth selection to quantitatively map variation of host genes to improvements in ethanol tolerance and production. They identified that membrane and osmotic stress were important limiting issues for the strain and that a single host gene that when overexpressed led up to a 75% improvement relative to the parent production strain.

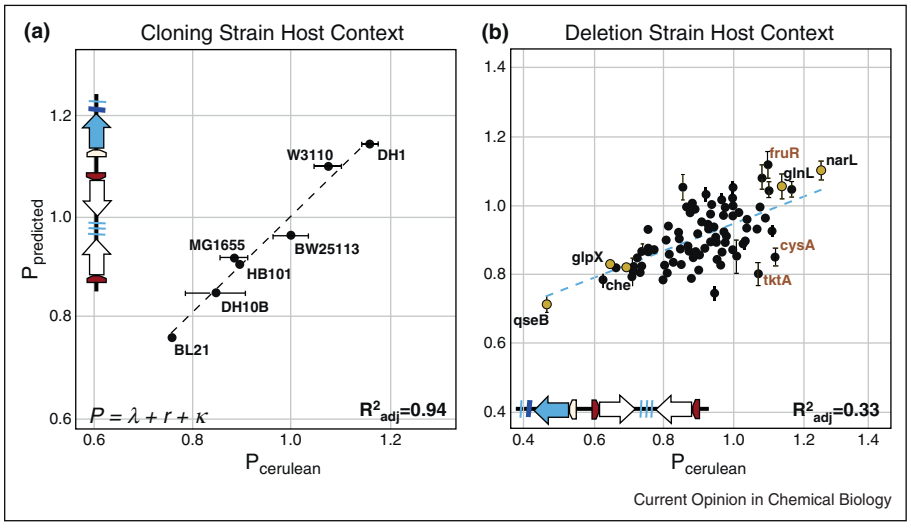
Other genome scale techniques for measuring macromolecular interaction and metabolic profiles will add more data that should aid in improving strain performance. Formal methods to transform these data into models of biological parts and their interactions suitable to drive design decisions remains to be developed.

Systematic quantification of environmental context

Host and environmental context are intimately linked because the major (unintended) effects of environment on a heterologous circuit are likely to arise *via* effects on host physiology. Sometimes, if the environment of deployment is known and static one can design or select circuits that operate well under those conditions. In metabolic engineering, there is the oft-cited problem that the biosynthetic pathways engineered in the laboratory often work poorly in the scaled-reactors that are necessary for economic production [57,58]. To demonstrate some issues, Moser et al. characterized how small synthetic circuits operate in different industrially relevant conditions and showed how changes in fermentation process affect host growth and resources thereby differentially affecting synthetic logic circuits in the host cell [59]. A recent industrial example of the challenge is the conversion of biosynthetic production of 1,3-propanediol, a precursor for many industrial products, from 'specialty' to commodity scale required the optimization of over 70 genes off-pathway before sufficient production in industrially relevant environments was achieved [60]. Many separate groups working over many years achieved the optimization. In more complex environments beyond the bioreactor we can imagine that the issues of designing predictable and reliable function are compounded.

Formal methods for discovering the interaction between host and heterologous genes and environmental conditions should lead to principles of design by which desirable synthetic function is maintained in the face

Figure 3



An example of characterization of host context for a simple probe circuit. A plasmid bearing three simple monocistronic expression cassettes for different fluorescent proteins (mCherry, mVenus, and mCerulean) was transformed either into six different standard laboratory strains of *Escherichia coli* or into 88 deletion strains derived from *E. coli* BW25113. The effect of each genetic background on protein expression was measured. **(A)** Predicted versus observed protein production of mCerulean based on a linear regression model relating inferred protein production to a combination of ribosome availability, lag time, and growth rate (inset equation). For the cloning strains it was possible to explain differences in protein production by strain to strain differences in ribosome availability. (R^2 = 0.94). The small inset is a representation of the probe circuit. Points are labeled by strain name. **(B)** For the deletion strains however, effects seemed more gene specific and the model from (A) could only explain 33% of the production variability. The method identified host subsystems and genes that both generically and specifically affect heterologous gene expression. Figures are derived from Figures 1D and 2D of Ref. [55••] respectively.

of variable conditions. One approach is to systematically vary both environmental conditions and gene expression to map the interactions between environmental components and each gene that affect fitness and designed phenotype. Skerker et al. used large-scale insertional mutagenesis of the ethanol producing bacterium, *Zymomonas mobilis*, to discover the genes that affect tolerance to and productivity in cellulosic hydrolysates that can be feedstocks for industrial fermentation [61••]. Such plant hydrolysates also contain many compounds that inhibit microbial growth and fermentation. By mapping how every gene in this organism conferred fitness in both purified components and mixtures, 44 genes were identified to be key determinants of performance and linked to particular classes of chemical stressor. It was possible to infer from this gene set that the real hydrolysates contained an inhibitory compound, methylglyoxal, that had not been detected previously. The information was used to target genes for strain improvement. In a related approach Sandoval et al. used barcoded promoter mutation libraries to map the effect of increased or decreased expression of nearly every gene in *E. coli* onto growth in several model environments (cellulosic hydrolysate, low pH, and high acetate). They identified more than 25 mutations that improved growth rate 10–200% for several different conditions and pointed to subsystems of importance to tolerance to hydrolysate [62••]. The Sandoval study, however, also demonstrated how difficult it could be to combine knowledge of these different mechanisms together to vastly improve strain performance because of a type of buffering epistasis among effects of the different genes.

Systematic quantification of ecological context

Because there are few applications wherein it is currently feasible to release synthetic organisms into open ecologies there have been scarce studies quantifying the biological basis of persistence of synthetic organisms in complex ecologies or the impact of the synthetic organism thereon. There are not yet rigorous metrics based on definitions of environmental health for how much it is permissible to perturb an ecology through introduction of an organism. However, we have progressed to the point where it is increasingly possible to map interactions between an introduced microbe and the surrounding ecology using metagenomic and associated functional techniques. In a recent study of a long term experiment mapping how a genetically modified microorganism (Pseudomonas fluorescens HK44 engineered for degradation of polycyclic aromatic hydrocarbons) and its DNA survive and propagate in realistic environments, a metagenomic analysis of a soil lysimeter community tracked the changes in microbial population composition over time. After fourteen years, while the

engineered microbe population had declined below detectability and could not be cultured, signatures of its specific DNA did survive and might be associated by transfer to other microbes [63••]. The authors did not specifically conclude how the surrounding microbial population dynamics were different between populations exposed and not exposed to HK44 but the study demonstrated the technical feasibility of addressing this question.

In a mammalian context, similar metagenomic approaches were used to track how the gut microbial population in a patient suffering from Clostridium difficile-associated disease changed after treatment by fecal transplant from a healthy donor [64]. The study demonstrated how the population overall change and stabilized to resemble the healthy microbial population, repopulating with key missing taxa, and alleviating symptoms. While there were no engineered microbes in this particular treatment, the study is a harbinger for how to track and understand the effects of engineered probiotics and other components of the human microbiome.

Systematic quantification of evolutionary context

Evolutionary context concerns how quickly a synthetic organism is selected out of a population or accumulates fitness-enhancing mutations, some of which might change the designed behaviors, in a given environment as a consequence of bearing specific synthetic elements. A goal is to map how inclusion of a specific heterologous DNA sequence into an organism will affect its fitness across environments and how properties of that sequence will affect the mutation rates across the genome. Knowledge of mechanisms of mutation has provided rules of thumb for design. For example, it is known that introduction of repetitive elements into a design invites a higher rate of their recombination and thus mutation of circuit function, an effect that has been recently used in a positive sense to direct mutations to improve circuit function by introduction of repeats into RBS spacer regions to target tuning of translational efficiency [65]. Approaches to prevent heterologous circuit loads from causing evolutionary pressure on the host and thus selection for loss of function have been demonstrated including using switch elements whose state-maintenance requires minimal energy to maintain state [54] and designs that effectively couple expression of a costly element to that of an essential element [66].

There are few systematic studies of how different environments and part designs collude to affect host fitness and mutation rates. Sleight et al. studied how similarity between two homologous terminators leads to differing rates of deletion of the region between [67••]. They found that removing all homology between the terminators increased the evolutionary half-life in a given environment 170-fold compared to identical terminators and that the evolutionary half-life of the circuit decreased exponentially with increasing expression of the intervening gene. Given the systematic methods for measuring environmental context above, and the ability to construct and measure large libraries of configurations and variations of synthetic parts, it should be possible to scale studies to derive quantitative principles linking intrinsic, genetic and evolutionary context to evolutionary rates.

Conclusions

The approaches above suggest a program by which the uncertainties that challenge complex and trustworthy design in synthetic biology might be overcome. Systematic characterization of host biology and synthetic biological part operation across contexts can lead to discovery of mechanisms, both generic and specific, that affect reliable operation of heterologous circuitry and will form a knowledgebase sufficient for predictive design. Most such characterization, to date, has been for engineered bacteria and we need to extend these methodologies to mammalian circuitry.

The scale necessary for such systematic characterization may call for large-scale scientific programs to collect these data on parts and designs for specific challenge applications. For an efficient design, build, test and learn cycle such programs would need defensible laboratory simulations of deployment environments that allow efficient capture of the effects at each level of context above and a suite of measurement tools to capture the physiological state of the cells, the interactions with the nonliving and living members of its environment, and the fitness and mutational effects therein. To serve this, standard experimental designs and computational frameworks need to be developed that properly parameterize and assess predictive models of function of single biological parts and whole systems under context uncertainty. If this can be accomplished then the barriers to design and implementation of the complex biological systems that may be necessary to solve problems beyond the bioreactor will be significantly lowered.

Acknowledgements

This work was supported by a grant from the Department of Energy grant number DE-FOA-0000640. APA would like to acknowledge V.K. Mutalik for his help with Figure 2.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. World Economic Forum, Insight Report: Global Risks 2012; 2012. Geneva.

2. Global Agenda Council on Emerging Technologies, The top 10 emerging technologies for 2012. The World Economics Forum Blog 2012. Available from: http://forumblog.org/2012/02/the-2012-top-10-emerging-technologies/.

3. Grupp SA *et al.*: **Chimeric antigen receptor-modified T cells for acute lymphoid leukemia**. *New Engl J Med* 2013, **368**:1509-1518.
• The paper describes an approach to engineering T-cells that cured patients of acute lymphoid leukemia.

4. Porter DL *et al.*: **Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia**. *New Engl J Med* 2011, **365**:725-733.

5. Yu W, Fang H: **Clinical trials with oncolytic adenovirus in China**. *Curr Cancer Drug Targets* 2007, **7**:141-148.

6. Donnelly O *et al.*: **Live viruses to treat cancer**. *J R Soc Med* 2013, **106**:310-314.

7. Maguire AM *et al.*: **Safety and efficacy of gene transfer for Leber's congenital amaurosis**. *New Engl J Med* 2008, **358**:2240-2248.

8. Bainbridge JW *et al.*: **Effect of gene therapy on visual function in Leber's congenital amaurosis**. *New Engl J Med* 2008, **358**:2231-2239.

9. Nathwani AC *et al.*: **Adenovirus-associated virus vector-mediated gene transfer in hemophilia B**. *New Engl J Med* 2011, **365**:2357-2365.

10. Eiseman B *et al.*: **Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis**. *Surgery* 1958, **44**:854-859.

11. van Nood E *et al.*: **Duodenal infusion of donor feces for recurrent Clostridium difficile**. *New Engl J Med* 2013, **368**:407-415.

12. Xie Z *et al.*: **Multi-input RNAi-based logic circuit for identification of specific cancer cells**. *Science* 2011, **333**:1307-1311.

13. Huh JH *et al.*: **Modular design of a synthetic payload delivery device**. *ACS Synth Biol* 2013.

14. Anderson JC, Voigt CA, Arkin AP: **Environmental signal integration by a modular AND gate**. *Mol Syst Biol* 2007, **3**:133.

15. Anderson JC *et al.*: **Environmentally controlled invasion of cancer cells by engineered bacteria**. *J Mol Biol* 2006, **355**:619-627.

16. Miller M *et al.*: **Modular design of artificial tissue homeostasis: robust control through synthetic cellular heterogeneity**. *PLoS Comput Biol* 2012, **8**:e1002579.
• A wonderful example of an analysis for the design for a complex circuit to control the differentiation of an engineered stem-cell into a pancreatic islet-like cell to treat diabetes.

17. Tebas P *et al.*: **Antiviral effects of autologous CD4 T cells genetically modified with a conditionally replicating lentiviral vector expressing long antisense to HIV**. *Blood* 2013, **121**:1524-1533.

18. Levine BL *et al.*: **Gene transfer in humans using a conditionally replicating lentiviral vector**. *Proc Natl Acad Sci U S A* 2006, **103**:17372-17377.

19. Weinberger LS, Schaffer DV, Arkin AP: **Theoretical design of a gene therapy to prevent AIDS but not human immunodeficiency virus type 1 infection**. *J Virol* 2003, **77**:10028-10036.

20. Rouzine IM, Weinberger LS: **Design requirements for interfering particles to maintain coadaptive stability with HIV-1**. *J Virol* 2013, **87**:2081-2093.
• A very nice analysis of the design requirements to manage a complex multiscale interaction between a designed interfering particle (a designer virus) and its target, HIV to maximize therapeutic effect.

21. Lucks JB *et al.*: **Toward scalable parts families for predictable design of biological circuits**. *Curr Opin Microbiol* 2008, **11**:567-573.

22. Mutalik VK *et al.*: **Precise and reliable gene expression via standard transcription and translation initiation elements**. *Nat Methods* 2013, **10**:354-360.
•• A demonstration of how to engineer against discovered mechanisms of genetic context effects to achieve reliable and predictable function.

23. Temme K *et al.*: **Modular control of multiple pathways using engineered orthogonal T7 polymerases**. *Nucleic Acids Res* 2012, **40**:8773-8781.

24. Blount BA *et al.*: **Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology**. *PLoS ONE* 2012, **7**:e33279.

25. Tamsir A, Tabor JJ, Voigt CA: **Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'**. *Nature* 2011, **469**:212-215.

26. Khalil AS *et al.*: **A synthetic biology framework for programming eukaryotic transcription functions**. *Cell* 2012, **150**:647-658.

27. Takahashi MK, Lucks JB: **A modular strategy for engineering orthogonal chimeric RNA transcription regulators**. *Nucleic Acids Res* 2013, **41**:7577-7588.

28. Qi L *et al.*: **Engineering naturally occurring trans-acting non-coding RNAs to sense molecular signals**. *Nucleic Acids Res* 2012, **40**:5775-5786.

29. Mutalik VK *et al.*: **Rationally designed families of orthogonal RNA regulators of translation**. *Nat Chem Biol* 2012, **8**:447-454.

30. Lucks JB *et al.*: **Versatile RNA-sensing transcriptional regulators for engineering genetic networks**. *Proc Natl Acad Sci U S A* 2011, **108**:8617-8622.

31. Qi LS *et al.*: **Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression**. *Cell* 2013, **152**:1173-1183.

32. Chen YJ *et al.*: **Characterization of 582 natural and synthetic terminators and quantification of their design constraints**. *Nat Methods* 2013, **10**:659-664.

33. Cambray G *et al.*: **Measurement and modeling of intrinsic transcription terminators**. *Nucleic Acids Res* 2013, **41**:5139-5148.

34. Dueber JE *et al.*: **Synthetic protein scaffolds provide modular control over metabolic flux**. *Nat Biotechnol* 2009, **27**:753-759.

35. Thompson KE *et al.*: **SYNZIP protein interaction toolbox: in vitro and in vivo specifications of heterospecific coiled-coil interaction domains**. *ACS Synth Biol* 2012, **1**:118-129.

36. Wei KY, Chen YY, Smolke CD: **A yeast-based rapid prototype platform for gene control elements in mammalian cells**. *Biotechnol Bioeng* 2013, **110**:1201-1210.

37. Moon TS *et al.*: **Genetic programs constructed from layered logic gates in single cells**. *Nature* 2012, **491**:249-253.

38. Liu CC *et al.*: **An adaptor from translational to transcriptional control enables predictable assembly of complex regulation**. *Nat Methods* 2012.
• A demonstration of how control of context on an RNA transcript among a number of complex novel controllers permits construction of a 4-input NOR gate.

39. Shis DL, Bennett MR: **Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants**. *Proc Natl Acad Sci U S A* 2013, **110**:5028-5033.

40. Chen D, Arkin AP: **Sequestration-based bistability enables tuning of the switching boundaries and design of a latch**. *Mol Syst Biol* 2012, **8**:p620.

41. Guimaraes JC, Liu CC, Arkin AP: **From biological parts to circuit design**. In *Synthetic Biology: Tools and Applications.* Edited by Zhao H. Urbana: Elsevier Inc.; 2013:63-78.

42. Cardinale S, Arkin AP: **Contextualizing context for synthetic biology — identifying causes of failure of synthetic biological systems**. *Biotechnol J* 2012, **7**:856-866.

43. Kittleson JT, Wu GC, Anderson JC: **Successes and failures in modular genetic engineering**. *Curr Opin Chem Biol* 2012, **16**:329-336.

44. Gustafsson C *et al.*: **Engineering genes for predictable protein expression**. *Protein Exp Purif* 2012, **83**:37-46.

45. Santner AA *et al.*: **Sweeping away protein aggregation with entropic bristles: intrinsically disordered protein fusions enhance soluble expression**. *Biochemistry* 2012, **51**:7250-7262.

46. Mutalik VK *et al.*: **Quantitative estimation of activity and quality for collections of functional genetic elements**. *Nat Methods* 2013, **10**:347-353.

Demonstration of an approach to systematically characterize genetic context effects. Similar in spirit to the characterization of intrinsic context in Ref. [44].

47. Salis HM: **The ribosome binding site calculator**. *Methods Enzymol* 2011, **498**:19-42.

48. Salis HM, Mirsky EA, Voigt CA: **Automated design of synthetic ribosome binding sites to control protein expression**. *Nat Biotechnol* 2009, **27**:946-950.

49. Kosuri S *et al.*: **Composability of regulatory sequences controlling transcription and translation in *Escherichia coli***. *Proc Natl Acad Sci U S A* 2013.

50. Qi L *et al.*: **RNA processing enables predictable programming of gene expression**. *Nat Biotechnol* 2012, **30**:1002-1006.

51. Lou C *et al.*: **Ribozyme-based insulator parts buffer synthetic circuits from genetic context**. *Nat Biotechnol* 2012, **30**:1137-1142.

52. Ham TS *et al.*: **A tightly regulated inducible expression system utilizing the fim inversion recombination switch**. *Biotechnol Bioeng* 2006, **94**:1-4.

53. Bonnet J *et al.*: **Amplifying genetic logic gates**. *Science* 2013, **340**:599-603.

54. Bonnet J, Subsoontorn P, Endy D: **Rewritable digital data storage in live cells via engineered control of recombination directionality**. *Proc Natl Acad Sci U S A* 2012, **109**:8884-8889.

55. •• Cardinale S, Joachimiak MP, Arkin AP: **Effects of genetic variation on the *E. coli* host–circuit interface**. *Cell Rep* 2013, **4**:231-237.
Demonstration of a systematic approach to characterize host context.

56. Woodruff LB, Boyle NR, Gill RT: **Engineering improved ethanol production in *Escherichia coli* with a genome-wide approach**. *Metabolic Eng* 2013, **17**:1-11.

57. Schmidt FR: **Optimization and scale up of industrial fermentation processes**. *Appl Microbiol Biotechnol* 2005, **68**:425-435.

58. Lara AR *et al.*: **Living with heterogeneities in bioreactors: understanding the effects of environmental gradients on cells**. *Mol Biotechnol* 2006, **34**:355-381.

59. Moser F *et al.*: **Genetic circuit performance under conditions relevant for industrial bioreactors**. *ACS Synth Biol* 2012, **1**:555-564.

60. Lee SY, Mattanovich D, Villaverde A: **Systems metabolic engineering, industrial biotechnology and microbial cell factories**. *Microbial Cell Fact* 2012, **11**:p156.

61. •• Skerker JM *et al.*: **Dissecting a complex chemical stress: chemogenomic profiling of plant hydrolysates**. *Mol Syst Biol* 2013, **9**:674.
Demonstration of how to systematically characterize environmental context effects to discover key genetic targets and environmental components important for system function and viability.

62. •• Sandoval NR *et al.*: **Strategy for directing combinatorial genome engineering in *Escherichia coli***. *Proc Natl Acad Sci U S A* 2012, **109**:10540-10545.
Genome-scale technology to combinatorially vary host context to discover performance limitations.

63. •• Layton AC *et al.*: **Ameliorating risk: culturable and metagenomic monitoring of the 14 year decline of a genetically engineered microorganism at a bioremediation field site**. *J Bioremed Biodegr* 2012, **S**:9-17.
One of the few studies to formally characterize ecological context effects of synthetic systems.

64. Khoruts A *et al.*: **Changes in the composition of the human fecal microbiome after bacteriotherapy for recurrent *Clostridium difficile*-associated diarrhea**. *J Clin Gastroenterol* 2010, **44**:354-360.

65. Egbert RG, Klavins E: **Fine-tuning gene networks using simple sequence repeats**. *Proc Natl Acad Sci U S A* 2012, **109**:16817-16822.

66. Yang S, Sleight SC, Sauro HM: **Rationally designed bidirectional promoter improves the evolutionary stability of synthetic genetic circuits**. *Nucleic Acids Res* 2013, **41**:pe33.

67. •• Sleight SC, Bartley BA, Sauro HM: **Synthetic systems as microbial threats: predictability of loss-of-function mutations in engineered systems**. In *The Science and Applications of Synthetic and Systems Biology: Workshop Summary*. Edited by Choffnes ER, Relman DA, Pray L. Washington, DC: National Academies Press; 2011:394-416.
A nice approach to characterize evolutionary context through designed variations in circuit sequence and environment.