Host and Consortia Engineering

(Host Engineering)

5885 Hollis Street, 4th Floor, Emeryville, CA 94608
Phone: +1.510.871.3272 Fax: +1.510.245.2223

This material is based upon work supported by the National Science Foundation under Grant No. 1818248.

Host and Consortia Engineering

Summary

Host and Consortia Engineering spans the development of cell-free systems, synthetic cells, single-cell organisms, multicellular tissues and whole organisms, and microbial consortia and biomes. Development of robust cell-free systems capable of diverse reactions, domestication and use of many single-cell hosts, targeted modification of multicellular organisms, and manipulation of microbial consortia.

Introduction and Impact

Engineering biology has delivered new tools to engineer microorganisms, plants, and animal cell lines. There are now entirely new ways to construct hosts to perform tasks that nature cannot accomplish. While many of these efforts have focused on 'traditional' hosts represented by model microbes like *E. coli* and *S. cerevisiae*, there is a wealth of potential if the unique capabilities of a broader range of microbes can be harnessed for useful purposes. These might include microbes that are photosynthetic, such as cyanobacteria (Markley, Begemann, Clarke, Gordon, & Pfleger, 2015), that can utilize non-sugar feedstocks such as methane or lignocellulose (Haitjema, Solomon, Henske, Theodorou, & O'Malley, 2014; Sundstrom & Criddle, 2015), or that can be engineered to produce and secrete complex macromolecules more efficiently than model hosts. The possibility that stable multi-organism consortia and biomes of defined compositions could be constructed is particularly tantalizing.

Cell-free biology has been a staple of life science research for more than 50 years. More recent technological achievements have created cell-free gene expression systems that can produce protein at titers reaching grams/liter, that can be constructed from non-model organisms, and that are greatly minimizing the time needed to prototype systems and circuits via the design-build-test cycle. The model-driven construction of complex cell-free systems as hosts could result in programmable hosts for advanced biosensing, for on-demand biomanufacturing, and even for the bottom-up construction of synthetic cells.

At present, engineering complex functions in non-model hosts remains difficult because many of the tools and approaches developed for model organisms cannot be applied with the same efficacy in non-model organisms. The profound impact that the development of similar tools and approaches for engineering non-model organisms would bring to two of the world's most important industrial sectors, energy and chemical production, more than justify attention to these challenges. It will be crucial to ensure that sufficient attention is given to biosecurity and biodefense risks that will rise along with improved capabilities for engineering diverse microbes (National Academies of Sciences, Engineering, and Medicine, Division on Earth and Life Studies, Board on Life Sciences, Board on Chemical Sciences and Technology, & Committee on Strategies for Identifying and Addressing Potential Biodefense Vulnerabilities Posed by Synthetic Biology, 2018). The pharmaceutical industry has long biomanufactured fermentation-based natural products in microbes and therapeutic proteins from mammalian cell culture, both of which could be dramatically improved through advancements in host engineering.

Compared to engineering in single cell hosts, the state-of-the-art for engineering multicellular systems and organisms is less well-developed. To date, these approaches have

been primarily aligned with natural reproduction, where genetically identical cells and tissues are created by editing the gametes or embryos of plants or animals. Gene editing methodologies introducing biochemical and molecular changes have already resulted in plants and animals with desirable characteristics that may be difficult to obtain through traditional breeding techniques. Further developing the capacity to reliably, and selectively, edit and modify multicellular eukaryotes could be transformative for a broad range of environmental and agricultural applications.

Host and Consortia Engineering focuses on the advancement of tools and technologies required for the characterization and engineering of host cells and organisms, and the integration and interaction of these systems and the environment. This includes developing methods, tools, and models to: 1) generate synthetic cells and cell-free systems to accomplish tasks and processes that cannot be accommodated by existing natural hosts; 2) enable organismal transformation, modification, and reprogramming of cellular chemistry, biology, and transport; 3) predict and integrate inputs and outcomes from environmental signals; and 4) enable the control, definition, and determination of differentiation, three-dimensional architecture, and other aspects of complex multicellular systems and biomes.

Transformative Tools and Technologies

Cell-free systems

Cell-free biology is the activation of complex biological processes without using intact living cells. While used for more than 50 years across the life sciences as a foundational research tool, a recent technical renaissance has made possible high-yielding cell-free gene expression systems (that produce protein in excess of grams/liter), the development of cell-free platforms from non-model organisms, and multiplexed strategies for rapidly assessing biological design-build-test cycles. These advances provide exciting opportunities to profoundly transform engineering biology through new approaches to model-driven design of genetic circuits, fast and portable sensing of compounds, on-demand biomanufacturing, building cells from the bottom up (i.e., synthetic cells), and next-generation educational kits. Key opportunities lie in understanding, harnessing, and expanding the capabilities of biological systems. For example, through the use of cell-free systems to inform cellular design, the efficiency of DNA synthesis can be amplified so that many different genes (encoding many different biomolecules) can be synthesized. The ability of cell-free systems to transcribe and translate a piece of DNA without the need to clone it into a specific vector and transforming into an organism (with all the limits associated with DNA transformation efficiencies) enables cell-free systems to shorten the time to testing, thus speeding up the overall design-build-test cycle, and enabling the scalable prototyping of gene function. However, to date, there are limited numbers of large datasets available that allow comparison of part performance between the cell-free environment and in cells (*in vivo*). One need is to make data and models available to the community so that others can build and test improved models leveraging already developed systems and data. In another direction of research, there is a need to investigate the use of cell-free systems in manufacturing. Imagine how rapid access to vaccines and therapeutics in remote settings could change lives, and how new biomanufacturing paradigms suitable for use in low resource

settings might promote better access to costly drugs through decentralized production. “Just-add-water” freeze-dried, cell-free systems could offer a disruptive approach to emerging and re-emerging diseases threats. It is a paradigm shifting concept.

Tools for engineering and characterization of host organisms

Today, we have a number of host organisms for which we have a satisfactory understanding of their metabolism and sufficient genetic tools that we can use for reliable engineering. However, there are a number of applications that beg for more suitable chassis. There is a need to develop new tools for existing organisms, as well as entirely new platform organisms, and capabilities compatible with high-throughput, data-driven workflows that are becoming increasingly favored in industrial biotechnology. Key capabilities needed across organisms include reliable transformation methods for plasmid delivery and genome integration, and well-characterized genetic parts (including promoters and terminators) to regulate gene expression (Johns et al., 2018). Predictive models for gene and protein expression-timing and -levels are also needed. Through the successful engineering of a broader library of host cells and multicellular organisms, we can increase the number of reporters and tools to better understand biology, establish new living sensors and sentinels, and expand the production of polymers, metabolites, and numerous other products.

Host onboarding and transformations

One area of host engineering where progress still needs to be achieved is the on-boarding of engineered genetic sequences, circuits, and pathways into host cells. Crucial to this is a detailed understanding of the central dogma machinery and a systems-level understanding of host physiology such that genes can be reliably expressed and (synthetic) pathways incorporated without negatively impacting fitness. This will include an understanding and prediction of the endogenous gene regulator elements, including transcription factors, important DNA cis elements, regulatory RNAs, and the role of chromatin and epigenetic markings. Furthermore, there is the need for fully sequenced and annotated genomes for the majority of organisms; this can be extended to fully annotated metabolic pathways and enzyme activities. Advancements in genetic transformations (or viral transductions) and the ability to manipulate the genome, and ultimately, the ability to transplant chromosomes, would enable more robust design, control, and/or domestication of host organisms and their functional cellular machinery.

HOST AND CONSORTIA ENGINEERING

Goal	Breakthough Capability	Milestone

Cell-free systems capable of natural and/or non-natural reactions.

2 Years	5 Years	10 Years	20 Years
Ability to build reproducible and comparable cell-free systems for practical applications in bioengineering and biomanufacturing from multiple organisms, including non-model hosts.			
Complete characterization of the general effects of cell-growth harvest conditions and extract preparation parameters on bacterial cell-free extract behavior.	Complete standardization of common-use bacterial cell-free system.	Complete library of user-defined reaction components that could be used in a customizable cell-free system.	Consistent ability to generate cell-free systems from any organism or a subset of organisms that make all types of desired products.
Ability to build a cell, including the molecular subsystems that enable the processes of DNA replication, transcription, translation, energy regeneration, and membrane construction.			
Demonstrated ability to synthesize all components encoded by a (minimal/synthetic) cell using cell-free systems.	Demonstrate and design a minimal genome that could support the construction of a cell, including regulation.	Ability to have ribosomes make ribosomes in a cell-free system.	Engineer compartmentalization and communication strategies for the design of (synthetic) cells.
	Ability to build metabolic modules capable of supporting long-lasting energy regeneration.	Expand the chemistry of living systems to make chemical reactions not possible with biological chemistry alone.	Replace test tubes with chemically-defined, standardized micro-vesicles to compartmentalize processes.
Long-lasting, robust, and low-cost cell-free system for protein synthesis and biomanufacturing.			
Identify reagent instabilities in cell-free systems across multiple organisms and all biological kingdoms.	Alleviate reagent instabilities and prolong the half-life of cell-free reagents from a few hours to several days using inexpensive substrates.	Stabilize catalysts to facilitate cell-free reactions on the order of weeks.	Robust and scalable production of cell-free systems that last for weeks.
	Avoid inhibition (poisoning) of cell-free reactions by byproducts or the desired products.		
Ability to use cell-free systems to inform cellular design of genetic parts and circuits.			
Ability to use next-generation sequencing read-outs to quantitatively map performance of genetic designs in cell-free systems.	Ability to identify new genetic parts in cell-free systems for any bacterial host to facilitate forward engineering in cells.	Ability to identify new genetic circuits in cell-free systems for any bacterial host to facilitate forward engineering in cells.	Ability to identify new genetic circuits in cell-free systems for any eukaryotic host to facilitate forward engineering in cells.
			Accelerate the development of any non-model host into useful chassis organisms for engineering biology with cell-free systems.
Decentralized, portable, on-demand sensing and manufacturing using cell-free systems.			
Ability to use safe lysates low in endotoxin for sensing and manufacturing objectives.	Demonstrate portability (two-year storage duration of freeze-dried reactions without loss of functionality) of cell-free systems.	Point-of-care cell-free protein production system ready for validation by the Food and Drug Administration.	Point-of-care cell-free protein therapeutic and vaccine production system ready for validation by the Food and Drug Administration.
	Increase productivity and rate of cell-free reactions.		

Ability to manufacture any targeted glycosylated protein or metabolite using cell-free biosynthesis.			
Ability to build modular, versatile cell-free platforms for glycosylation pathway assembly.	Expanded set of glycosylation enzyme-variants that efficiently install eukaryotic glycans.	Expanded set of enzymes capable of glycosylating metabolites *in vitro*.	Ability to produce any glycosylated protein therapeutics and vaccines at the point-of-care in less than one week.
	Production of bacterial glycoconjugate vaccines in cell-free systems.	Cell-free pipelines to produce and assess the functionality of diverse, human glycosylated protein therapeutics	

On-demand production of single-cell hosts capable of natural and non-natural biochemistry.

Ability to grow any host, anytime, in a controlled and regulated setting.			
Establish protocols for the development of media that support cellular viability for non-model organisms.	Develop robust, high-throughput screens for rapidly assaying useful properties in libraries of organisms.		
Robust screening of useful chassis beyond model organisms.	Use output of high-throughput screens/sensors and computer control to amplify a signal or expand a cell line that produces a product of interest.		

Routine domestication of non-model organisms through DNA delivery and genetic modification.			
Catalog and assay current methodologies and tools for carrying out DNA delivery in microbial/mammalian systems and plant systems.	Development of well-characterized and robust insertion sites in plant genomes.	Develop high-throughput, targeted editing and rapid-genome-evolution tools that couple genetic changes to phenotypic changes.	Routine genetic manipulation of any non-model host in less than one week from first isolation.
Develop high-throughput methods that can be done in parallel for DNA delivery (using standard methods) into non-model hosts.	Develop high-throughput, genome-wide editing tools for non-model organisms.		
Establish a suite of gene-editing tools for the rapid insertion and/or deletion of genetic elements in diverse primary mammalian cells.	Establish robust temporal and/or spatial control of gene expression in mammalian cells.	Develop universal approaches to transforming any plant.	
Characterize basic DNA parts for expression strength in non-model organisms.	Develop broad-host-range vectors for a variety of model and non-model organisms.		

Ability to build and control small molecule biosynthesis inside cells by design or through evolution.			
Identify model organisms for performing specific types of chemistries or organisms that have native precursor biosynthesis pathways for specific classes of molecules.	Construct a limited number of model host organisms for synthesizing all-natural products.	Software and hardware for optimizing titer, rate, and yield of any product produced by any host.	On-demand construction of single cell organisms for production of nearly any molecule of interest, including organic chemicals and polymers.
Precise temporal control of gene expression for well-studied systems.	Construction of single-cell organisms for production of unnatural derivatives of natural products.		
	Temporal control over multiplexed regulation of many genes in parallel.		

2 Years	5 Years	10 Years	20 Years

Spatial control over, or organization of, metabolic pathways in cells and construction of unnatural organelles.			
Tools to target heterologous proteins to various subcellular compartments.	Inducible synthesis of organelles.	Methods and tools to reprogram transport of metabolites and compartmentalization of biochemical reactions.	
		Alter chemical conditions within the organelle/microcompartment.	
	Gain-control for selective permeability in and out of the organelle.	Multiple orthogonal organelles/ microcompartments in the same cell for compartmentalizing different parts of a pathway.	

Production and secretion of any protein with the desired glycosylation or other post-translational modifications.		
One or more microbial hosts capable of producing laboratory-scale quantities of a single glycoform of a desired protein.	A few microbial hosts capable of secreting functional versions of proteins with no post-translational modifications.	Ubiquitous control of post-translational modification in a diverse array of hosts.

On-demand fabrication and modification of multicellular organisms.

Ability to control differentiation and de-differentiation of cells within a population.	
On-demand, reproducible functionalization of simple micro-tissues or micro-consortia made up of two or more engineered cell types.	Programmable and regulatable pathways that can be induced to differentiate or de-differentiate somatic cells.

Ability to characterize and control the three-dimensional architecture of multicellular systems.			
Characterize existing tissue components and standardize measurements to evaluate function.	Identification of novel 3D scaffold designs that can lead to desirable cellular properties.	Create modular, synthetic communication circuits that can be implemented in tissues to allow for control of new or existing cellular communication systems.	Bottom-up design and construction of whole organs at the centimeter-length scale.

Ability to achieve stable non-heritable changes in somatic cells.			
Routine delivery of biomolecule "effectors" (i.e., DNA, RNA, proteins) into slowly-dividing or non-dividing cells.	Generation of effective artificial epigenetic chromosomal states and maturation of the emerging field of chromatin engineering.	Ability to generate cell states that are stable and effective after the inducer/effector is removed in certain model tissues.	Nimble adaptation of somatic cell engineering technologies to any natural tissue at any developmental stage.

2 Years **5 Years** **10 Years** **20 Years**

Ability to make predictable and precise, targeted, heritable changes through germline editing.			
Complete sequence of select host genomes to allow design of targets for gene editing.	Efficient germline transformation systems developed in targeted hosts.	Ability to coordinate engineered multicellular functions in intact organisms via orthogonal communication systems.	Routine, on-demand, efficient germline editing for any targeted host of interest at high-throughput scale.
	Ability to deliver transgene constructs to most (>90%) somatic cells in a higher organism to rapidly prototype transgenic phenotypes.		
	Temporally controlled transgene expression that works on the scale of generations.		
Define and validate tissue-specific DNA parts in plants.	Efficient gene editing in differentiated cells.	On-demand gene editing of organisms with desired traits.	
	Ability to domesticate engineered biological parts to confer immune tolerance in immunocompetent organisms.		

Generation of biomes and consortia with desired functions and ecologies.

Ability to control cell-to-cell communication between different species.			
Tightly-controlled promoter-response regulator systems that enable intra- and inter-species cellular communication.	Synthetic cell-to-cell communication elements and networks that function in a broad range of host organisms.	Signal-response pathways that function in synthetic communities of 5-10 organisms, employing a variety of pathway types and host species.	Ability to produce engineered microbes that can reliably invade and coexist within a complex community and manipulate the consortium/ biome function and behavior.
Ability to characterize, manipulate, and program three-deminsional architecture of the biome.			
Use of existing technologies to better understand the species composition and collective components of microbial communities and consortia.	Non-destructive, 3D visualization of microbial communities from a broad range of environments.	Ability to manipulate the 3D architecture of natural or engineered communities using external inputs.	Programmed communities that self-assemble into a desired 3D architecture.
Ability to control and/or define the function of an engineered microbial community/biome.			
Ability to combine species with specialized functions to enable the production of desired products.	Assembly of consortia to produce desired molecules/products, considering community-level metabolic flux.	Plug-and-play assembly of consortia to produce desired molecules/products from specific starting materials, considering community level metabolic flux and organism-to-organism communication.	On demand assembly of consortia that are programmed to respond dynamically.
Targeted modification of an existing microbiome to enable new functions or address dysbiosis through the addition, removal, or reorganization of the community members.			
Use of existing technologies to characterize functions of microbial communities from a broad range of environments.	Characterize how select microbiomes respond to changes in the environment.	Predictive models of microbiome function and response to a broad range of environmental and ecological changes.	Ability to modify an existing biome or consortia as desired.

2 Years **5 Years** **10 Years** **20 Years**

Roadmap Elements

Goal 1: Cell-free systems capable of natural and/or non-natural reactions.

[Current State-of-the-Art]: Cell-free synthetic biology is emerging as a transformative approach to understand, harness, and expand the capabilities of natural biological systems. The foundational principle is that complex biomolecular transformations are conducted without using intact cells. Instead, crude cell lysates, or extracts, are used, which provides a unique freedom-of-design to control biological systems for a wide array of applications. For example, cell-free protein synthesis (CFPS) systems have been used to decipher the genetic code, prototype genetic circuits (Moore et al., 2018; Takahashi et al., 2015) and metabolic pathways (Karim & Jewett, 2016), enable portable diagnostics (Pardee et al., 2016; Wen et al., 2017), facilitate on-demand biomolecular manufacturing (Pardee et al., 2016), produce antibody therapeutics at the commercial scale (Yin et al., 2012), and enable advances in education (Huang et al., 2018; Stark et al., 2019, 2018). The recent surge of applications has revitalized interest in cell-free systems, especially in areas where limits imposed by the organism may impede progress. Despite these advances, several barriers limit advancement of the field. Specifically, there are opportunities to: (i) standardize lysate generation approaches, (ii) enable decentralized manufacturing of complex therapeutics and vaccines, (iii) establish design principles for genetically-encoded biosensors to rationalize their engineering for addressing global sustainable development goals (e.g., food and water security) including portable and on-demand strategies, (iv) reduce costs of cell-free reactions by enabling long-lived protein expression, (v) generate large datasets and quantitative models to allow comparison of part performance between the cell-free environment and in cells (*in vivo*), (vi) synthesize more complex classes of proteins such as glycoproteins, (vii) construct synthetic cellular machines (e.g., ribosomes) and biosynthetic modules to both understand life and lead to new manufacturing paradigms, and (viii) build cells from the bottom up.

[Breakthrough Capability 1]: Ability to build reproducible and comparable cell-free systems for practical applications in bioengineering and biomanufacturing from multiple organisms, including non-model hosts.

- **2 years: Complete characterization of the general effects of cell-growth harvest conditions and extract preparation parameters on bacterial cell-free extract behavior (e.g., protein synthesis and native genetic regulators).**
    - [Bottleneck]: Inability to produce cell-free systems in a scalable and standardized fashion, protocols for cell-free systems generation vary from lab-to-lab, and reactions are carried out at very small volumes (1-10µl) that limit the extent to which we can characterize the systems; for example, we lack an understanding of how the metabolic state of an extract impacts energy usage, protein synthesis, and DNA/RNA regulation.
        - [Potential Solution]: Arrive at a standardized bacterial cell-free system extract generation and reaction protocol that is robust, inexpensive, and would allow their routine scalable production; this includes identification of energy mixtures robust to various extract preparation procedures.

        - [Potential Solution]: Use -omics tools (e.g., mass spectrometry, next generation sequencing) to understand batch-to-batch and lab-to-lab lysate quality and composition variability.
        - [Potential Solution]: Define an equipment set for bacterial lysate preparation (i.e., lysis technique).
- **5 years: Complete standardization of common-use bacterial cell-free system.**
    - [Bottleneck]: *Ad hoc* use of different extract preparation procedures, cell-free reaction conditions, and reporter construct architectures lead to challenges in making reproducible and comparable cell-free systems.
        - [Potential Solution]: Design cell-free system physiochemical composition that is robust to different process parameters, or identify compositions for each defined extract preparation method.
        - [Potential Solution]: Identify and implement measurement techniques needed to facilitate reproducibility (e.g., standard assay for quantifying activity across labs).
        - [Potential Solution]: Rigorously characterize and understand how reaction geometries and surface area-to-volume impact cell-free system performance.
        - [Potential Solution]: Create a plasmid repository specific to bacterial cell-free systems (including T7 constructs and *E. coli* promoters), with a description of performance (such as protein expression kinetics).
        - [Potential Solution]: Develop a quantitative ‘culture’ to cell-free system practices.
- **10 years: Complete library of user-defined reaction components for use in a customizable cell-free system.**
    - [Bottleneck]: Gaps in understanding critical components and how they vary across species and kingdoms (e.g., prokaryotes vs. eukaryotes).
        - [Potential Solution]: Assess critical parameters with design-of-experiments and machine learning across strains and procedures.
        - [Potential Solution]: Develop an online tool to customize the components of a cell-free reaction for a given organism and approximate RNA/protein output, reveal components that should have been included, and application-specific assembly (metabolic engineering vs. protein synthesis vs. circuits).
    - [Bottleneck]: Ability to regenerate energy and cofactors with native metabolism is constrained in non-*E. coli* platforms.
        - [Potential Solution]: Assess metabolic pathways that could be activated to facilitate energy regeneration.
        - [Potential Solution]: Develop exogenous cofactor regeneration modules to drop in and out when native ones are missing.
        - [Potential Solution]: Incorporate genome-scale models to predict energy and cofactor regeneration systems most suitable for a new host.

- **20 years: Consistent ability to generate cell-free systems from <u>any organism or a subset of organisms that make all types of desired products</u>, including all biological kingdoms and DNA programmed cell-free systems at-scale.**
    - [Bottleneck]: Gaps in understanding critical components of host lysates that make each unique and which to use for a specific, desired product.
        - [Potential Solution]: Identify the critical species-specific components (using -omics approaches) and experimentally validate these components (including transcription factors, accessory proteins for translation, and metabolic modules).
    - [Bottleneck]: Production of lysate relies on scaled-up culturing of sometimes recalcitrant cells.
        - [Potential Solution]: Understand which extract preparation protocol parameters should/can be tuned to match a new organism (i.e., trace the physiology of an organism to the extract preparation parameters).
    - [Bottleneck]: Identify units of operation required for the scale-up and scale-out production of cell-free systems.
        - [Potential Solution]: Identify quality assurance and quality control metrics at each unit of operation that would allow the scale-up of the process.

[Breakthrough Capability 2]: Ability to build a cell, including the molecular subsystems that enable the processes of DNA replication, transcription, translation, energy regeneration, and membrane construction.

- **2 years: Demonstrated ability to synthesize all components encoded by a minimal or synthetic cell using cell-free systems.**
    - [Bottleneck]: Multiple minimal genomes necessary for building a cell have been proposed but none have yet been demonstrated to work in this context.
        - [Potential Solution]: Define a set of possible minimal genomes that could enable self-replication of a minimal cell and make these component parts.
        - [Potential Solution]: Test each component of a minimal cell individually (including metabolic modules, regulation by most types of regulators, etc.).
- **5 years: Demonstrate and design a minimal genome that could support the construction of a cell, including regulation.**
    - [Bottleneck]: We do not yet know how to build genomes *de novo*, and native genomes have built-in regulation that may not be suitable for engineered biological systems.
        - [Potential Solution]: Create and refactor modularized pathways for building a minimal cell.
        - [Potential Solution]: Show examples of a synthetic self-regulating gene cluster.
        - [Potential Solution]: Exploit CRISPR as a potential universal regulation mechanism to regulate synthetic minimal genomes.

- **5 years: Ability to build metabolic modules capable of supporting long-lasting energy regeneration.**
    - [Bottleneck]: Typical cell-free reactions that serve as the basis for minimal cells use an energy-rich environment of nucleotide triphosphates and high energy phosphate bond donors for chemical energy; these systems are short-lived, expensive, and can lead to inhibitory byproducts.
        - [Potential Solution]: Create generalized approaches for cost-effective, long-lived energy regeneration modules that can integrate with synthetic cells, which may require compartmentalized systems.
        - [Potential Solution]: Integrate efficient physical ATP regeneration systems in synthetic cells, such as photosynthesis.
- **10 years: Ability to have ribosomes make ribosomes in a cell-free system.**
    - [Bottleneck]: Constructing ribosomes built completely of *in vitro* synthesized parts has remained elusive.
        - [Potential Solution]: Develop conditions that facilitate co-synthesis of all ribosome proteins and ribosomal RNA, first in extracts, then in purified systems.
- **10 years: Expand the chemistry of living systems to make chemical reactions not possible with biological chemistry alone.**
    - [Bottleneck]: The palette of biological chemistry is smaller than chemistry.
        - [Potential Solution]: Build hybrid biological-chemical systems.
        - [Potential Solution]: Expand the chemistry of genetically-encoded systems.
- **20 years: Engineer compartmentalization and communication strategies for the design of synthetics cells.** (For related reading, please see Goal 2, Breakthrough Capability 4: Spatial control over (or organization of) metabolic pathways in cells and construction of unnatural organelles.)
    - [Bottleneck]: Lack a precise physical understanding of how the composition of the compartment influences encapsulation, transport, and retention of various types of cargo.
        - [Potential Solution]: Quantify the relationship between compartment composition and encapsulation efficiency, permeability, and stability as a function of compartment cargo.
    - [Bottleneck]: Communication strategies for synthetic cells are not yet well-developed.
        - [Potential Solution]: Identify and implement mechanisms for communication between engineered compartments that allows efficient, reliable, multi-channel signaling between synthetic cells.
        - [Potential Solution]: Identify complementary signaling modules with reliable performance in a cell-free environment.

- **20 years: Replace test tubes with chemically-defined, standardized micro-vesicles to compartmentalize processes.**
    - [Bottleneck]: Inability to generate stable liposomes with the correct size and membrane-permeability to generate micro-vesicles (synthetic cells) with different chemical environments.
        - [Potential Solution]: Use microfluidics for precise production of chemically-defined vesicles.
        - [Potential Solution]: Exploit non-natural compartments, such as block copolymers, as a means to make mechanically robust compartments with negligible non-specific permeability.

[Breakthrough Capability 3]: Long-lasting, robust, and low-cost cell-free system for protein synthesis and biomanufacturing.

- **2 years: Identify reagent instabilities in cell-free systems across multiple organisms and all biological kingdoms.**
    - [Bottleneck]: Stability of reagents is not well understood in non-*E. coli*-based cell-free systems.
        - [Potential Solution]: The full spectrum of metabolic activity during cell-free reactions over time can be characterized in multiple systems, including those emerging cell-free systems, to identify reagent instabilities.
- **5 years: Alleviate reagent instabilities and prolong the half-life of cell-free reagents from a few hours to several days using inexpensive substrates.**
    - [Bottleneck]: High-energy phosphate compounds are still commonly used to fuel cell-free systems (e.g., PEP, 3PGA); these compounds are expensive, lead to inhibitory phosphate concentrations, and only enable bursts of ATP regeneration instead of long-lived energy regeneration.
        - [Potential Solution]: Central metabolism, non-phosphorylated energy substrates, or light driven approaches (e.g., from bacteriorhodopsin), among others, could be used to enable long-lived energy regeneration without inhibitory byproducts; should be tested in extracts made from organism across all kingdoms and identify key factors (such as, concentrations, additional reagents, etc.) for robust alternative energy activation.
    - [Bottleneck]: Cofactor regeneration and balancing strategies preclude long-term activation of energy metabolism to fuel metabolic activity.
        - [Potential Solution]: Molecular purge valves implemented in crude extracts could maintain redox balance.
        - [Potential Solution]: Develop schemes for balancing ATP, and derivatives including NAD(P) and FAD.
        - [Potential Solution]: Create non-natural cofactors, and requisite engineered enzymes, that have better stability properties.
    - [Bottleneck]: Reagent instabilities limit reaction time.
        - [Potential Solution]: Genomic modifications to extract source strains can be carried out to stabilize substrates.

- **5 years: Avoid inhibition (poisoning) of cell-free reactions by byproducts or the desired products.**
    - [Bottleneck]: Small molecule byproducts of metabolism, such as phosphate, inhibit cell-free reactions.
        - [Potential Solution]: Engineer systems to avoid the accumulation of chemical inhibitors.
        - [Potential Solution]: Develop product siphoning strategies and dialysis-like strategies to remove inhibitors from the system.
        - [Potential Solution]: Engineer molecular complexes that are resistant to byproducts and chemical products.
- **10 years: Stabilize catalysts to facilitate cell-free reactions on the order of weeks.**
    - [Bottleneck]: Cell-free reactions terminate because substrates and cofactors are depleted, byproducts accumulate to inhibit the reaction, or the catalysts become inactivated; of these, enzyme stability represents a significant technical and economic hurdle to technology development in this space.
        - [Potential Solution]: Characterize catalyst instabilities such as, tolerance to byproducts and inhibitors.
        - [Potential Solution]: Increase stability of enzymes involved in cell-free systems.
        - [Potential Solution]: Generate design criteria for choosing catalysts (i.e., identifying select organisms from which to derive lysates and enzymes) for desired applications.
        - [Potential Solution]: Develop reactor designs and bioprocesses that continuously replenish the source of catalysts.
- **20 years: Robust and scalable production of cell-free systems that last for weeks.**
    - [Bottleneck]: Without enzymes or lysates that are stable on the order of weeks, significant fractions of carbon will otherwise be used in generating the biocatalysts required of these systems.
        - [Potential Solution]: Integrate knowledge of the system and innovations to facilitate long-lived reactions.

[Breakthrough Capability 4]: Ability to use cell-free systems to inform cellular design of genetic parts and circuits.

- **2 years: Ability to use next-generation sequencing read-outs to quantitatively map performance of genetic designs in cell-free systems.**
    - [Bottleneck]: Sequence-function data are limited by colorimetric and fluorescent read-outs.
        - [Potential Solution]: Develop next-generation, deep-sequencing-based approaches for monitoring transcription and translation in cell-free reactions, including transcription factor metabolite interactions.

- **5 years: Ability to identify new genetic parts in cell-free systems (including promoters, ribosome binding sites, and terminators) for any bacterial host to facilitate forward engineering in cells.**
    - [Bottleneck]: There are limited numbers of sufficiently-large datasets available that allow comparison of genetic part performance and the development of modeling frameworks between the cell-free environment and *in vivo*.
        - [Potential Solution]: Develop cell-free systems for 20 industrially-relevant organisms that could form a testbed to establish libraries of new genetic parts and how to accelerate design.
        - [Potential Solution]: Develop a repository of genetic parts for cell-free systems, including performances.
        - [Potential Solution]: Develop a quantitative modeling platform specific to cell-free systems that takes into account the advantages and limitations of cell-free expression.
- **10 years: Ability to identify new genetic circuits in cell-free systems for any bacterial host to facilitate forward engineering in cells.**
    - [Bottleneck]: Resource limitations in cell-free systems (e.g., energy and cofactor regeneration) constrain the construction of multi-gene systems encoded by complex genetics, as well as time-dynamics needed to assess their function.
        - [Potential Solution]: Create robust, long-lived cell-free systems that can be routinely used for activating and characterizing multi-gene circuits.
    - [Bottleneck]: There are limited numbers of sufficiently-large datasets available that allow comparison of genetic circuit performance and the development of modeling frameworks between the cell-free environment and *in vivo*.
        - [Potential Solution]: Develop cell-free systems for 20 industrially-relevant organisms which could form a testbed to establish libraries of new genetic circuits.
- **20 years: Ability to identify new genetic circuits in cell-free systems for any eukaryotic host to facilitate forward engineering in cells.**
    - [Bottleneck]: Lack of knowledge of how to activate essential components in eukaryotic cell-free systems.
        - [Potential Solution]: Make more lysates from eukaryotes and eukaryotic-like systems to be able to assess essential components.
    - [Bottleneck]: Transcription and translation are typically combined in cell-free systems requiring viral sequences (e.g., internal ribosome entry site) for translation initiation rather than a 5'cap and polyA tail.
        - [Potential Solution]: Develop strategies to compartmentalize transcription to better mimic the natural process.
- **20 years: Accelerate the development of any non-model host into useful chassis organisms for engineering biology with cell-free systems.**
    - [Bottleneck]: Sufficient transcription and translation activity is necessary to assess genetic designs and metabolic pathways.

        - [Potential Solution]: Streamline and provide generalized approaches to enable sufficient cell-free activity for gene expression and biosynthesis from diverse species.
        - [Potential Solution]: Use mass spectrometry to determine the proteome composition and metabolite composition of non-model organisms to accelerate the optimization of cell-free systems.
    - [Bottleneck]: Expand cell-free systems to poorly explored ares, such as extremophiles, for engineering biology far from standard conditions.
        - [Potential Solution]: Select a set of extremophiles that can be grown in laboratories and easily lysed.

[Breakthrough Capability 5]: *Decentralized, portable, on-demand sensing and manufacturing using cell-free systems.*

- **2 years: Ability to use safe lysates low in endotoxin for sensing and manufacturing objectives.**
    - [Bottleneck]: Toxicity of cell-free components is not well understood.
        - [Potential Solution]: Evaluate toxicity of cell-free components.
    - [Bottleneck]: Lipopolysaccharides, also known as lipoglycans and endotoxins, are large molecules consisting of a lipid and a polysaccharide composed of O-antigen which can cause toxicity of products manufactured using bacteria.
        - [Potential Solution]: Cost-effective strategies to remove endotoxin is required to ensure safety.
        - [Potential Solution]: Create bacterial species that are detoxified by design.
- **5 years: Demonstrate portability (such as two-year storage of freeze-dried reactions without loss of functionality) of cell-free systems.**
    - [Bottleneck]: A major limitation of traditional sensors and centralized medicine manufacturing is that the products must be refrigerated; cell-free systems can be freeze-dried for potential room-temperature storage and distribution, but suffer from activity loss in certain conditions.
        - [Potential Solution]: Create stable, freeze-dried systems for storage without a cold-chain by using cryoprotectants and process modifications.
        - [Potential Solution]: Demonstrate multiple reaction formats (i.e., pellets, gels, paper, etc.) stable for 2-5 years.
- **5 years: Increase productivity and rate of cell-free reactions.**
    - [Bottleneck]: Manufacturing medicines in rapid response to emerging and re-emerging threats requires fast protein synthesis rates and reactions that can be completed in minutes to hours.
        - [Potential Solution]: For translation outputs, increase the overall catalyst concentration in the reaction to enhance reaction rates.
        - [Potential Solution]: When possible, develop transcriptional outputs (such as for sensors) which offer a significant speed improvement (minutes versus hours).

- **10 years: Point-of-care cell-free protein production system ready for validation by the Food and Drug Administration (FDA).**
    - [Bottleneck]: Completely automated operation platform to manufacture and purify proteins suitable for the FDA has not yet been fully validated.
        - [Potential Solution]: Build integrated units and measure the repeatability (of the same unit) and reproducibility (between units) of the system.
- **20 years: Point-of-care cell-free protein therapeutic and vaccine production system ready for validation by the Food and Drug Administration (FDA).**
    - [Bottleneck]: Point-of-care synthesis and administration of glycoprotein therapeutics and vaccines requires robust production and purification methods.
        - [Potential Solution]: Develop simple, portable purification systems which can reliably produce FDA-compliant vaccines and therapeutics from cell-free production systems and demonstrate efficiency in animal models.

[Breakthrough Capability 6]: Ability to manufacture any targeted glycosylated protein or metabolite using cell-free biosynthesis.

- **2 years: Ability to build modular, versatile cell-free platforms for glycosylation pathway assembly.**
    - [Bottleneck]: Many of the most important components of glycosylation pathways are associated with cellular membranes and cannot be recapitulated easily in cell-free systems.
        - [Potential Solution]: Develop and optimize efficient strategies use of oligosaccharyltransferases (including eukaryotic versions) to transfer pre-built sugars from lipid-linked oligosaccharides *in vitro.*
        - [Potential Solution]: Develop methods to select and assemble a set of soluble enzymatic tools capable of producing therapeutically-relevant glycoproteins with desired properties *in vitro* from simple, commercially available activated-sugar building blocks.
- **5 years: Expanded set of glycosylation enzyme-variants that efficiently install eukaryotic glycans.**
    - [Bottleneck]: Synthesis of complex human glycans in cell-free systems is constrained by the available set of well characterized enzymes; existing characterizations do not provide sufficient tools to go from a set of sugar monomers and a design to a glycoprotein of interest.
        - [Potential Solution]: Expand the glycoengineering toolkit by characterizing glycoslytransferases to assemble sets of well-characterized enzymes that can reliably produce desired glycoproteins.
        - [Potential Solution]: Engineer existing glycosylation enzymes for desired activities when naturally occurring enzymes with desired functionalities are not available (e.g., engineering of the bacterial oligosaccharyltransferase to accept the eukaryotic core glycan would enable the efficient production of the eukaryotic core glycan in bacterial systems).

- **5 years: Production of bacterial glycoconjugate vaccines in cell-free systems.**
    - [Bottleneck]: The production of glycoconjugate vaccines against bacteria require the culturing of pathogenic strains recalcitrant to bioengineering and the use of non-specific conjugation chemistry.
        - [Potential Solution]: Express diverse bacterial O-antigen pathways for characterization and production of lipid-linked oligosaccharides required for vaccines and use bacterial oligosaccharyltransferases to site-specifically attach these oligosaccharides *in vitro*.
        - [Potential Solution]: Use the greater control afforded by cell-free protein synthesis systems to produce FDA-approved vaccine carrier and antigen proteins, many of which cannot be produced outside of pathogenic organisms.
- **10 years: Expanded set of enzymes capable of glycosylating metabolites *in vitro*.**
    - [Bottleneck]: Glycosylation of therapeutically-relevant metabolites (including many antibiotics) is often required for desirable pharmacokinetic/dynamic properties, but many of the enzymes which perform these glycosylation activities are unknown or can only be expressed in their native host strain.
        - [Potential Solution]: Use cell-free protein synthesis to screen the activity of metabolite-targeting glycosyltransferases on existing bioactive compound libraries to understand their specificities and improve the therapeutic utility of small molecule products.
    - [Bottleneck]: Substrate limitations lead to inherent inefficiencies.
        - [Potential Solution]: Develop metabolic models of lysate hosts to identify genetic knockouts that enhance glycan production and carry out such modifications.
- **10 years: Cell-free pipelines to produce and assess the functionality of diverse, human glycosylated protein therapeutics.**
    - [Bottleneck]: Development timelines of glycoprotein therapeutics are slowed by the need to produce these products in mammalian cell lines.
        - [Potential Solution]: Develop, optimize, and implement strategies to create more than ten unique and homogeneous glycan structures on proteins by cell-free methods (e.g., trimannose core eukaryotic glycan, afucosylated galactose-terminated core glycan, fucosylated sialic-acid terminated core glycan, biantennary glycan, etc.); once synthesized, these products can studied and optimized using therapeutic functionality assays.
- **20 years: Ability to produce any glycosylated protein therapeutics and vaccines at the point-of-care in less than one week.**
    - [Bottleneck]: Intentional engineering of glycan structures and the synthesis of novel structures is constrained by identifying sets of enzymes for the manufacture of the sugar structures.
        - [Potential Solution]: Develop cell-free systems capable of rapidly and robustly producing any defined glycoprotein on-demand.

Goal 2: On-demand production of single-cell hosts capable of natural and non-natural biochemistry.

[Current State-of-the-Art]: Current tools and technologies for on-demand production of organisms are limited by the number and scope of transformation capabilities, continuous and rapid production capability, and the lack of secure public repositories for academics and industry containing the necessary organismal design and characterization information. Today, we have a number of host microbes for which we have a satisfactory, though not extensive, understanding of their metabolism and sufficient genetic tools that we can use for reliable engineering. Engineering of plant and animal cells is expanding, especially given particular applications (e.g., CAR-T cell engineering), but still faces significant bottlenecks.

[Breakthrough Capability 1]: Ability to grow any host, anytime, in a controlled and regulated setting.

- **2 years: Establish protocols for the development of media that support cellular viability for non-model organisms.**
    - [Bottleneck]: Inability to grow most organisms on earth in non-natural environments (i.e., inside the lab) and a lack of knowledge of conditions for growing novel or non-model organisms.
        - [Potential Solution]: Database collection and mapping: a better understanding of how to find what is important for organismal growth and survival, potentially including the ability to apply artificial intelligence to recognize patterns.
        - [Potential Solution]: Develop a general protocol for identifying viable media and create databases of media that work with known organisms; leverage machine learning using various databases (e.g., 16S rRNA profiling to identify evolutionary relationships between known and unknown species) to help generate suggestions for potential media.
- **2 years: Robust screening of useful hosts beyond model organisms.**
    - [Bottleneck]: There are a few hosts that are the primary chassis for engineering applications, but they are not always the best choice for biosynthesis of all potential products.
        - [Potential Solution]: Expand beyond well-studied model organisms to those that are closely related including the development of tools for genetic manipulation and transformation, as necessary.
    - [Bottleneck]: Inability to identify specific organisms that could be beneficial for the production of specific products, metabolites, intermediates, and specific catalytic reactions.
        - [Potential Solution]: Development of rapid assays to map desired activity to potential capability.
        - [Potential Solution]: Inexpensive metabolomics analysis that lends itself to profiling of all metabolites in a cell. Current solutions that are ongoing

include developing of specific sensors (i.e., RNAs) to detect specific metabolites (though this is low-throughput).

- **5 years: Develop robust, high-throughput screens for rapidly assaying useful properties in libraries of organisms.**
    - [Bottleneck]: Availability of high-throughput screens for functions/products that cannot be screened by color or that cannot be selected using growth/viability assays.
        - [Potential Solution]: Incorporate protein or RNA sensors into cells for a particular small molecule that the cell may be engineered to produce or consume.
- **5 years: Use output of high-throughput screens/sensors and computer control to amplify a signal or expand a cell line that produces a product of interest.**
    - [Bottleneck]: There are few (or no) systems for simultaneously measuring the output of a sensor and then using a computer to expand production of the desired product.
        - [Potential Solution]: Develop biosensors to detect one or more particular desired outputs and gene expression systems that allow for computer control (e.g., light, inducer-repressor).

[Breakthrough Capability 2]: Routine domestication of non-model organisms through DNA delivery and genetic modification. (For related reading, please see *Gene editing, Synthesis, and Assembly.)*

- **2 years: Catalog and assay current methodologies and tools for carrying out DNA delivery in microbial/mammalian systems (e.g., viral vectors, conjugations, biochemical methods) and plant systems (e.g., *Agrobacterium*-, biolistic-, nanomaterial-based methods).**
    - [Bottleneck]: These methods have not been systematically compared among organisms resulting in a lack of clarity as to when one approach may be superior, or even viable, compared to another.
        - [Potential Solution]: Better aggregation of the information available for these different organisms, especially if a standard set of similar broad host vectors can be used (as a control) to standardize these.
        - [Potential Solution]: Develop bacteriophages that can be useful for engineering large number of organisms.
        - [Potential solution]: Understand the biological basis for what enables some species (e.g., plants) to be more amenable to transformation and genetic modification.
- **2 years: Develop high-throughput methods that can be done in parallel for DNA delivery (using standard methods) into non-model hosts.**
    - [Bottleneck]: Currently a limitation in the cloning process.
        - [Potential Solution]: Methods exist, but they could be improved and more widespread.

- **2 years: Establish a suite of gene-editing tools for the rapid insertion and/or deletion of genetic elements in diverse primary mammalian cells.**
    - [Bottleneck]: Genome-editing tools, particularly CRISPR/Cas technology, have enabled efficient genetic modification of a variety of immortalized cell lines, but primary mammalian cells are often more difficult to engineer with high efficiency and at scale.
        - [Potential Solution]: Develop non-toxic gene-delivery methods (viral or non-viral) utilizing reagents and equipment that are compatible with clinical manufacturing and/or high-volume cell modification.
- **2 years: Characterize basic DNA parts for expression strength in non-model organisms, specifically a larger library of plants.**
    - [Bottleneck]: Basic characterized promoters with characterized expression strengths has not yet been carried out in a systematic manner. Previous efforts have been piecemeal and the transferability of parts (e.g., promoters) between different plant species has not been well explored.
        - [Potential Solution]: A large-scale, community driven project to standardize and characterize parts will dramatically advance the state of plant engineering.
- **5 years: Development of well-characterized and robust insertion sites in plant genomes.**
    - [Bottleneck]: The majority of plant engineering efforts rely on random insertion of transgenes into the genome, resulting in the necessity to screen and characterize transformants -- a very laborious process when working with plants.
        - [Potential Solution]: Develop CRISPR-based genome-editing tools that reliably get targeted insertions with high efficiency.
- **5 years: Develop high-throughput, genome-wide editing tools for non-model organisms.**
    - [Bottleneck]: Gene editing tools are not always specific or work at all in certain organisms.
        - [Potential Solution]: Screen/develop new CRISPR-based genome editing proteins.
- **5 years: Establish robust temporal and/or spatial control of gene expression in mammalian cells.**
    - [Bottleneck]: Relatively few transcriptional, post-transcriptional, or translational regulatory devices exist for robust gene-expression control in mammalian cells; the vast majority of systems make use of inducible promoters that are often leaky or require precise tuning, which cannot be done in many practical applications such as patient-derived cells.
        - [Potential Solution]: Develop and catalogue a suite of core promoters and response elements, RNA-based regulatory devices, protein-degradation tags, among others, and record standardized, quantitative information on their performance (e.g., basal expression level with enzymatic vs. fluorescent reporters, fold induction, degradation rate, etc.).

- **5 years: Develop broad-host-range vectors for a variety of model and non-model organisms.**
    - [Bottleneck]: Lack of broad-host-range vectors that function for many different organisms requires the development of vectors specific for each organism.
        - [Potential Solution]: Engineer new host vectors targeting broader range of organisms; information from NCBI (e.g., plasmids sequences) could potentially be used to get guidance as to functional capabilities/parts that can be effective in different vector design for different organisms.
- **10 years: Develop high-throughput, targeted editing and rapid genome-evolution tools that couple genetic changes to phenotypic changes.**
    - [Bottleneck]: It is difficult to evolve non-model (and frankly, model) organisms to a desired phenotype.
        - [Potential Solution]: Develop general biosensors for particular desired phenotypes for use in non-model organisms.
        - [Potential Solution]: Develop mutator proteins coupled to sensors to evolve non-model organisms until they achieve some desired phenotype.
- **10 years: Develop universal approaches to transforming any plant.**
    - [Bottleneck]: Current plant transformation techniques work on a limited subset of plants.
        - [Potential Solution]: Better understand the biological basis for barriers in transformation in plants.
        - [Potential Solution]: Develop novel approaches that are species or host agnostic in incorporating and delivering DNA.
        - [Potential Solution]: Develop methods that are tissue-culture-independent.
- **20 years: Routine genetic manipulation of any non-model host in less than one week from first isolation.**
    - [Bottleneck]: It can take years to make a non-model into a host for heterologous gene expression.
        - [Potential Solution]: Use the tools developed to achieve previous milestones to address this bottleneck.

[Breakthrough Capability 3]: Ability to build and control small molecule biosynthesis inside cells by design or through evolution.

- **2 years: Identify model organisms for performing specific types of chemistries or organisms that have native precursor biosynthesis pathways for specific classes of molecules.**
    - [Bottleneck]: Academics and companies have constructed heterologous hosts for a limited number of chemical classes, but there are many other chemical classes that need hosts with precursor pathways constructed; additionally, some of these precursor production hosts may not be ideal for particular applications or environments.
        - [Potential Solution]: Use bioinformatics and screening to identify organisms that might be particularly useful for producing a chemical under a particular environmental condition or that already has precursor

pathways for a particular class of chemicals; build a database of those organisms.
        - [Potential Solution]: Development of robust plant hosts in order to bridge the gap before trying to engineer plant metabolic pathways into microbes.
- **2 years: Precise temporal control of gene expression for well-studied systems.**
    - [Bottleneck]: Many desired chemicals are toxic to growth of the producing host. Avoidance of this results in separation of growth and production phases which has been achieved by changing the media conditions (Clark & Blanch, 1997); however, such approaches are costly, may require the introduction of extra chemicals that are difficult to remove from the desired products, and cannot easily accommodate cellular heterogeneity, or be used to fine tune the shift from growth to production phases.
        - [Potential Solution]: Develop several types of expression systems for controlling the timing of gene expression and thus the timing of chemical production; ultimately, decouple growth and reproduction from energy and carbon metabolism and product generation (Venayak, von Kamp, Klamt, & Mahadevan, 2018).
    - [Bottleneck]: The necessary genetic regulatory elements should be sufficiently orthogonal to the host to permit tuning, but nonetheless responsive to changes in cell state.
        - [Potential Solution]: Identifying, and then re-engineering, diverse sets of protein and RNA regulators capable of activating or repressing gene expression in response to targeted changes in growth and nutrient state would expand these capabilities (Hsiao, Cheng, & Murray, 2016; Weinhandl, Winkler, Glieder, & Camattari, 2014); integrating the responses of these regulatory elements, perhaps through the use of engineered information processing networks, would permit the construction of circuitry for fine-tuning the timing of gene expression in microbial hosts.
- **5 years: Construct a limited number of model host organisms for synthesizing all-natural products.**
    - [Bottleneck]: There is a need for hosts capable of producing all natural products found in nature; there is no need for a single host to produce all of the natural products, but rather at least one host for every class of natural products.
        - [Potential Solution]: Use the database of possible hosts, genetic control systems and modular pathways to develop a range of hosts for each class of natural products; these hosts should produce high levels of the precursors to natural products.
- **5 years: Construction of single-cell organisms for production of unnatural derivatives of natural products.**
    - [Bottleneck]: There are few engineered cells that produce unnatural derivatives of natural products and the methods for engineering organisms to produce unnatural natural products are nascent.

        - [Potential Solution]: Construct hosts with the modular natural precursor pathways and engineered or evolved enzymes to create hosts that can produce unnatural natural products.
        - [Potential Solution]: Construct hosts with promiscuous enzymes (either natural, engineered, or evolved enzymers) and feed unnatural precursors that are incorporated into the final product.
- **5 years: Temporal control over multiplexed regulation of many genes in parallel.**
    - [Bottleneck]: It remains difficult to target gene expression to a particular growth phase or culture time in order to maximize production of a desired natural or unnatural product.
        - [Potential Solution]: Develop a catalog of promoters (including timing and gene expression strength) that are activated in various phases of cultivation.
        - [Potential Solution]: Identify promoters that are activated at the same time or with the same environmental queue and that can be used with multiple genes (i.e., multi-step biosynthetic pathway).
- **10 years: Software and hardware for optimizing titer, rate, and yield of any product produced by any host.**
    - [Bottleneck]: It is challenging and time-consuming to optimize the titer, rate and yield of a desired product in any host.
        - [Potential Solution]: Develop a software suite that considers the metabolic pathway(s) and host and incorporate a wide variety of measurements and can predict that changes that need to be made to the metabolic pathway add the host to optimize titer, rate, and yield.
- **20 years: On-demand construction of single cell organisms for production of nearly any molecule of interest, including organic chemicals and polymers.**
    - [Bottleneck]: Reliance on nature's single-cell organisms that are not ideally suited for producing a particular chemical.
        - [Potential Solution]: Use retrobiosynthesis software as well as genetic control system design software to *de novo* design microbial cells that can produce nearly any organic chemical; the cells would be designed for the processing environment to enable inexpensive purification of the final product.

[Breakthrough Capability 4]: Spatial control over, or organization of, metabolic pathways in cells and construction of unnatural organelles.

- **2 years: Tools to target heterologous proteins to various subcellular compartments.**
    - [Bottleneck]: There is a need for modular targeting tags that can be appended to any peptide sequence as N- or C-terminal fusions with efficient targeting to a desired compartment. The challenges are finding sequences that provide modularity; in other words, the ability to perform targeting on any protein sequence of interest, and compartmentalize a high percentage of total expressed protein, preferably at high expression levels.

        - [Potential Solution]: Most organelles have identified targeting sequences, but few satisfy the dual requirements of high efficiency and modularity; high-throughput screens for peptide sequences can be conducted using enzyme sequestration assays (DeLoache, Russ, & Dueber, 2016).
        - [Potential solution]: Bacterial organelles, known as microcompartments, have several known targeting tags for cargo enzymes and encapsulation is tunable with expression levels (Jakobson et al., 2016); however, distinct tags may use the same mechanism for loading, making it difficult to reliably control targeting of multiple enzymes into the same compartment. Additional biophysical characterization of existing and putative tags can be carried out using existing technologies to identify those with unique loading mechanisms or interaction partners.
- **5 years: Inducible synthesis of organelles.** Synthetic organelles (or compartments) that can be made on-demand and/or provide the means to compartmentalize different heterologous protein cargos would allow the engineer to mimic tissue compartmentalization in a single-cell microbe.
    - [Bottleneck]: Achieving sufficient understanding of the formation of organelles and microcompartments.
        - [Potential Solution]: Placing necessary organelle biogenesis factors under the expression control of a heterologous, inducible promoter should provide on-demand organelle production.
        - [Potential Solution]: Inducible control of bacterial microcompartment-forming proteins leads to some well-formed particles, but other factors must be identified to make sufficient quantities and do so robustly. These factors may be environmental factors or additional chaperone proteins, and can be identified by using a high-throughput screen, such as a flow cytometry-based assay that links cellular fluorescence to successful compartment formation (Kim & Tullman-Ercek, 2014).
- **5 years: Gain-control for selective permeability in and out of the organelle.**
    - [Bottleneck]: Many organelles and microcompartments will naturally have permeability to molecules that are desired to be partitioned either outside or inside the organelle (e.g., pathway intermediates); these organelles and the microcompartments will need to be modified to have lower permeability to these molecules.
        - [Potential Solution]: Selective transport of desired metabolites (e.g., pathway substrate, cofactors) must be engineered into the organelle/microcompartment by manipulating the proteins that govern the movement of small molecules across the membrane/shell.
- **10 years: Methods and tools to reprogram transport of metabolites and compartmentalization of biochemical reactions.**
    - [Bottleneck]: Prokaryotes and eukaryotes contain multiple compartments that can be utilized for sequestering sensitive chemistries or metabolites; unfortunately, it is not always clear which compartments to use.

        - [Potential Solution]: Catalogue organelles/microcompartments with favorable properties for desired applications/products (a suite of organelles will likely be of value to meet the needs of varied applications).
    - [Bottleneck]: Tools for engineering organelles and microcompartments are not readily available.
        - [Potential Solution]: Develop gene expression tools for a variety of different organelles in eukaryotes or for microcompartments in prokaryotes.
        - [Potential Solution]: Engineer multiple pathways to function in concert within the same organelle/microcompartment for optimal performance (e.g., cofactor balancing, co-substrate production, condensation reactions).
    - [Bottleneck]: Identifying parameters to optimize and means of optimization of microcompartment morphology.
        - [Potential Solution]: Genetically control the morphology of organelle/microcompartment (e.g., size, number) to achieve increased capacity for product cargo. Low capacity for enzymes is a limitation for many organelles; a solution would be the genetic manipulation of, for example, the shell protein expression levels or an organelle biogenesis pathway, to increase this capacity.
- **10 years: Alter chemical conditions within the organelle/microcompartment.**
    - [Bottleneck]: Altering the chemical environment of a compartment's lumen would require successful completion of efficient import of heterologous protein, reduction of small molecule permeability, and, for organelles, the functional trafficking of membrane proteins to the organelle's membrane.
        - [Potential Solution]: Combination of heterologous small molecule transport, enzyme, and compartmentalization of both enzymes and small molecule substrate, intermediates, and products will allow alteration of the chemical environment of the lumen.
- **10 years: Multiple orthogonal organelles/microcompartments in the same cell for compartmentalizing different parts of a pathway.**
    - [Bottleneck]: Achieving compartmentalization of different cargo in distinct organelles or microcompartments within the same cell demands the ability to generate such compartments with distinct "addresses" or protein import machinery recognizing targeting sequences orthogonal to the other native and synthetic organelles in the cell.
        - [Potential Solution]: Multiple natural organelles using different protein importomers can be utilized; however, the lumenal conditions are likely not optimal for the needs of each compartmentalized activities. Alternatively, engineered organelles with the ability to import orthogonal, distinct pools of protein in separate organelles should be achievable by either performing directed evolution on the native protein importomers or targeting heterologous importomers native to other organelles.

        - [Potential Solution]: Bacteria employ multiple distinct microcompartments within the same cell, and these are encoded by distinct operons which includes genes to make the protein shell boundary; it is feasible, but yet to be demonstrated, that these distinct microcompartments could be engineered simultaneously using existing techniques.

[Breakthrough Capability 5]: Production and secretion of any protein with the desired glycosylation or other post-translational modifications. (National Research Council (US) Committee on Assessing the Importance and Impact of Glycomics and Glycosciences, 2012)

- **2 years: One or more microbial hosts capable of producing laboratory-scale quantities of a single glycoform of a desired protein.**
    - [Bottleneck]: Compared to protein synthesis, post-translational modification and machinery is not well understood and any given protein may exist in the cell in multiple glycoforms, making advanced study challenging; this problem is particularly prevalent in the production of biologic pharmaceuticals (e.g., biosimilars) where the complex array of glycosylation patterns can cause differences in efficacy and safety (Sethuraman & Stadheim, 2006).
        - [Potential Solution]: Understand the pathways involved in glycosylation of the protein of interest and develop these pathways to enable tightly controlled synthesis of a single glycoform.
- **5 years: A few microbial hosts capable of secreting functional versions of proteins with no post-translational modifications.**
    - [Bottleneck]: There are relatively few bacterial hosts that have well developed protein secretion systems. Those that exist are primarily in Gram positive hosts, and are not necessarily compatible with all desired protein products; gram negative hosts, on the other hand, have an outer membrane that must also be crossed before exiting the cell, and secreting proteins across both the inner and outer membranes remains a challenge at the efficiencies required for commercial application.
        - [Potential Solution]: Develop new bacterial hosts with machinery specifically designed for protein secretion.
    - [Bottleneck]: The systems that do exist are either required for cellular function and thus can only be repurposed to a limited extent, or are highly regulated and difficult to activate or keep activated.
        - [Potential Solution]: Utilize systems biology to understand and remove the regulatory controls on secretion systems.
- **20 years: Ubiquitous control of post-translational modification (including glycosylation of multiple sites with multiple sugars) in a diverse array of hosts.**
    - [Bottleneck]: Post-translational modifications are often considered the “analog” control in the cell; as such, they can carefully tune how an individual protein interacts with its environment. The lack of understanding of this control makes affecting post-translational modification for desired use particularly challenging. The ability to tune the *analog* portion of the cell, however, can allow for more

efficient systems, better synthetic yields, and a broader set of uses than would otherwise be possible.

Goal 3: On-demand fabrication and modification of multicellular organisms.

[Current State-of-the-Art]: Currently for multicellular systems and organisms, our technology is closely aligned with natural reproduction: we edit gametes or embryos, and rely on natural processes to differentiate genetically-identical cells into tissues. Gene editing methodology allows substantial improvements and inclusion of novel biochemical and molecular changes. Today our engineering abilities in plants are limited to stable integration of small genetic circuits (fewer than 200 kb). Examples of engineered modification of animals include genome editing of chicken embryos to produce virus resistance (Looi et al., 2018; Sid & Schusser, 2018), and the inactivation of porcine endogenous retroviruses in pigs for human organs transplants (Niu et al., 2017; Ross & Coates, 2018), but significant work is necessary before we are able to selectively edit and modify multicellular eukaryotes with confidence and consistency.

Generally, we need a better understanding of cell-to-cell interactions and to establish stable modifications within multicellular systems. Current state-of-the-art in multicellular engineering includes tools and technologies for some plants (Farré et al., 2014) and fungi (such as *Aspergillus* (Lubertozzi & Keasling, 2009)), but most advances toward this goal have occurred in the engineering of a single cell type within a multicellular organism (for example, the introduction of the Polled trait into dairy cattle breeds (Van Enennaam, 2018)) and germline engineering. An emerging technology in multicellular system engineering is cell-scaffolding (loading specialized cells onto engineered matrices) and enabled control over the three-dimensional shape and structure of a system. Advancements in engineering for tissue- and organ-on-a-chip technologies are also helping to bring about advancements in this area.

For related reading, please see *Biomolecule, Pathway, and Circuit Engineering*, ***Goal 3: Holistic, integrated design of multi-part genetic systems (i.e., circuits and pathways)***.

[Breakthrough Capability 1]: Ability to control differentiation and de-differentiation of cells within a population.

- **2 years: On-demand, reproducible functionalization of simple micro-tissues or micro-consortia made up of two or more engineered cell types.**
    - [Bottleneck]: Variations in cell culture quality between trials and institutions.
        - [Potential Solution]: Investigate incubation conditions to identify and mitigate environmental sources of variability in cell behavior and growth.
        - [Potential Solution]: Identify informative biomarkers.
- **5 years: Programmable and regulatable pathways that can be induced to differentiate or de-differentiate somatic cells.**
    - [Bottleneck]: Gaps in understanding which genes and networks can be altered to control cell behavior.
        - [Potential Solution]: Experimental tools to co-regulate any desired set of multiple gene targets, such as via engineered transcription factors.

[Breakthrough Capability 2]: Ability to characterize and control the three-dimensional (3D) architecture of multicellular systems.

- **2 years: Characterize existing tissue components and standardize measurements to evaluate function.**
    - [Bottleneck]: Inconsistent reporting of matrix and cell performance resulting from differences in composition and tissue geometry.
        - [Potential Solution]: Generate a library of known cell types, matrices, and exogenous signalling molecules and characterize all combinations under identical conditions and geometries; measurements should include characteristics such as stress-strain response, degradation rates in serum, and immunogenicity.
    - [Bottleneck]: Discovery of novel exogenous signalling molecules to regulate cell and tissue behavior.
        - [Potential Solution]: Evaluate specific factors (via high-throughput chemical screening) that can be supplemented to existing 3D matrices/scaffolds to induce drastic changes in cellular behavior (such as morphology, differentiation, tissue composition, matrix alignment).
- **5 years: Identification of novel 3D scaffold designs that can lead to desirable cellular properties.**
    - [Bottleneck]: Limited capacity for nutrient delivery (~100μm by diffusion alone).
        - [Potential Solution]: Develop synthetic microvascular networks, either by self-assembly of endothelial cells and pericytes or 3D patterning of tissues; these networks must be able to support cell growth within the scaffold and be robust to changes in tissue composition (including the introduction of additional cell types, mechanical forces, or chemical factors).
    - [Bottleneck]: Heterogeneous cell seeding within a large scaffold.
        - [Potential Solution]: Enable and advance formation and production of extracellular matrix and methods to improve seeding (for example, forced flow of cells into tissue).
- **10 years: Create modular, synthetic communication circuits that can be implemented in tissues to allow for control of new or existing cellular communication systems.**
    - [Bottleneck]: Cell heterogeneity within tissue and ability to target only the cell type(s) of interest.
        - [Potential Solution]: Link integration or expression of a circuit to expression of a synthetic molecule in the target cell type.
    - [Bottleneck]: Robustness against perturbations, including the various signalling molecules expressed in tissue.
        - [Potential Solution]: Build in redundant control to genetic circuits, and leverage advances in biomolecular design to use components with large induction ratios and minimal cross-talk with other circuit components.

- **20 years: Bottom-up design and construction of whole organs at the centimeter-length scale.**
    - [Bottleneck]: Nutrient delivery in organs.
        - [Potential Solution]: Assemble synthetic arterioles and venuoles to interface with capillaries, then couple these synthetic vascular networks with cell culture media perfusion strategies.
    - [Bottleneck]: Large-scale assembly of substructures into complete tissues.
        - [Potential Solution]:
    - [Bottleneck]: Understanding the principles of organ *design*; most efforts to date focus on recapitulation of existing organs or a reduced set of functions performed by a given organ.
        - [Potential Solution]: Prototyping synthetic organs to substitute for, complement, or enhance native organ function in a manner beyond recapitulation of evolved biology.

[Breakthrough Capability 3]: Ability to achieve stable non-heritable changes in somatic cells.

- **2 years: Routine delivery of biomolecule "effectors" (i.e., DNA, RNA, proteins) into slowly-dividing or non-dividing cells.**
    - [Bottleneck]: Lack of technologies for homogenous delivery of macromolecules into tissues.
        - [Potential Solution]: Further development of cell-penetrating nanoparticles and exosomes.
- **5 years: Generation of effective artificial epigenetic chromosomal states and maturation of the emerging field of chromatin engineering.**
    - [Bottleneck]: Incomplete functional characterization of natural chromatin.
        - [Potential Solution]: Engineered platforms to rapidly interrogate hundreds of structural, enzymatic, and synthetic chromatin proteins.
    - [Bottleneck]: Uncertain causal relationship between genome/epigenome and cell behavior.
        - [Potential Solution]: Coordination of statistical genome/ epigenome association studies (GWAS/ EWAS) with experimental reconstruction of states to test and validate associations.
- **10 years: Ability to generate cell states that are stable and effective after the inducer/effector is removed in certain model tissues.**
    - [Bottleneck]: Gaps in understanding cell "homeostasis" and how biochemical processes inside cells are interconnected and reinforce each other.
        - [Potential Solution]: Development of experimentally-supported predictive (systems biology) models to predict the long-term impact of an artificial perturbation.
    - [Bottleneck]: Gene-editing dependency; too much focus on transcriptional regulation in the nucleus.
        - [Potential Solution]: New tools to control self-perpetuating "post-translational" states (such as RNA and protein modification).

        - [Potential Solution]: Advances in organelle engineering (especially for mitochondria).
- **20 years: Nimble adaptation of somatic cell engineering technologies to any natural tissue at any developmental stage.**
    - [Bottleneck]: Gaps in understanding cell-type and tissue-type-specific barriers that enable cells to resist conversions.
        - [Potential Solution]: Further advancement of systems biology methods to quickly identify appropriate target genes, proteins, and molecular networks.

[Breakthrough Capability 4]: Ability to make predictable and precise, targeted, heritable changes through germline editing.

- **2 years: Complete sequence of select host genomes to allow design of targets for gene editing.**
    - [Bottleneck]: Genetic variation.
        - [Potential Solution]: Sequencing of specific transformation target lines.
- **2 years: Define and validate tissue-specific DNA parts in plants.**
    - [Bottleneck]: There has been a dearth of plant DNA parts (e.g., promoters) that have been systematically characterized. Although many genes have been described from transcriptome datasets as tissue-specific, the validation and characterization of their tissue specificity will be required for future plant synthetic biology efforts.
        - [Potential Solution]: Systematic characterization of various tissue-specific promoters in various plant species.
- **5 years: Efficient germline transformation systems developed in targeted hosts.**
    - [Bottleneck]: Transformation systems are limited and optimization is slow; efficiency is such that the molecular analysis burden is high.
        - [Potential Solution]: Increase transformation efficiency through new vector design components that will stimulate cell division during the time the DNA is introduced into the cell OR enable improvement in high-throughput molecular analysis platforms to screen for those with the correct edits.
- **5 years: Ability to deliver transgene constructs to most (>90%) somatic cells in a higher eukaryote organism to rapidly prototype transgenic phenotypes.**
    - [Bottleneck]: Higher eukaryotes have relatively long timescales of organism development, making phenotype development too slow for effective research.
        - [Potential Solution]: Improve somatic cell nuclear transfer to embryos to speed multi-locus genome engineering for non-model organisms with long generation times.
    - [Bottleneck]: Existing gene delivery technologies reach only a subset of cells in an intact organism.
        - [Potential Solution]: Enhance gene delivery technologies to approach organism-scale delivery.

- **5 years: Temporally controlled transgene expression that works on the scale of generations.** For example, kill switches that are activated only after a defined number of generations.
    - [Bottleneck]: Robust molecular time-keeping methods.
        - [Potential Solution]: Design and implementation of robust synthetic cell cycle oscillators and other molecular timers.
    - [Bottleneck]: Gene expression platforms that confer stable expression across multiple cell divisions (such as in primary cells).
        - [Potential Solution]: Development of stable, controllable, heritable extra-genomic expression platforms, including artificial chromosomes.
    - [Bottleneck]: Spontaneous silencing of transgene constructs being expressed over long periods of time.
        - [Potential Solution]: Synthetic epigenetic mechanisms that interfere with or block natural silencing mechanisms.
- **5 years: Efficient gene editing in differentiated cells.**
    - [Bottleneck]: DNA folding is a physical blockade against gene-editing enzymes.
        - [Potential Solution]: Engineering the editing enzymes and/or helper proteins to unfold DNA.
    - [Bottleneck]: Bacterial CRISPR particles induce an immunogenic response.
        - [Potential Solution]: Discovery of non-immunogenic variants.
        - [Potential Solution]: Development of “coating” particles or chemical tags.
    - [Bottleneck]: Some cell types carry heterogeneous, naturally-modified genomes (such as immune cells).
        - [Potential Solution]: Delivery of gene expression cassettes that are not integrated into the chromosome.
- **5 years: Ability to domesticate engineered biological parts to confer immune tolerance in immunocompetent organisms.**
    - [Bottleneck]: Introduction of foreign proteins can induce immune rejection in immunocompetent organisms; the rules governing how such rejection is elicited by synthetic biology parts and how it may be circumvented are not yet clear.
        - [Potential Solution]: Generation toolboxes of “stealthed” parts that are unlikely to elicit immune rejection.
        - [Potential Solution]: Development of strategies for inducting active immune tolerance of synthetic biology parts.
- **10 years: Ability to coordinate engineered multicellular functions in intact organisms via orthogonal communication systems.**
    - [Bottleneck]: Generating synthetic analogs of coordinated processes, such as wound healing or immune protection, will likely require communication between engineered cells; co-opting native cell-cell signaling mechanisms is likely to exhibit cross-talk and cross-regulation with native systems.
        - [Potential Solution]: Generation of libraries of mutually orthogonal synthetic signaling molecules and receptors that can confer coordination across various length scales within an organism.

- **10 years: On-demand gene editing of organisms with desired traits.**
    - [Bottleneck]: Gene editing efficiency is low for multiple edits; limitations in what sequences can be edited due to CRISPR target-recognition constraints.
        - [Potential Solution]: Develop new CRISPR or other engineered enzymes that have expanded recognition sequences and efficiencies.
- **20 years: Routine, on-demand, efficient germline editing for any targeted hosts of interest at high-throughput scale.**
    - [Bottleneck]: Different and diverse transformation systems needed across species.
        - [Potential Solution]: Develop a process that is automated from preparation of embryos to transformation to selection/identification of successfully-edited embryos.

Goal 4: Generation of biomes and consortia with desired functions and ecologies.
Note: We make the distinction between biomes and multicellular organisms (see ***Goal 3: On-demand fabrication and modification of multicellular organisms***) through the definition of a biome as containing organisms (including multicellular organisms) with different genomes.

[Current State-of-the-Art]: While a few microbiome systems are well characterized, such as rhizobium for nitrogen fixation, we are still struggling to understand and how and why consortia of microbes cooperate in nature. Systems with mutual metabolic dependencies (synthetic heterotrophs) have enabled the construction of engineered consortia that are stable in laboratory settings (Pacheco, Moel, & Segrè, 2019). Our ability to produce synthetic interactions is possible with some ongoing efforts; for example, a small number of synthetic microbial consortia have been created as model systems, consisting of 2-3 different organisms (Kong, Meldgin, Collins, & Lu, 2018; McCarty & Ledesma-Amaro, 2019). Bioremediation and waste-water treatment demonstrate the principles that consortia can be used industrially, while probiotics and fecal microbe transplants demonstrate the principle that the composition of gut flora can be manipulated. Industrial startups in this space are emerging at a rapid pace, but our ability to make targeted changes, such as adding or removing a single organism, in an existing microbiome are very limited. Overall, our ability to understand and manipulate systems with specific functions or to remediate biomes and consortia that cease to function as desired is very limited.

[Breakthrough Capability 1]: Ability to control cell-to-cell communication between different species.

- **2 years: Tightly-controlled promoter-response regulator systems that enable intra- and inter-species cellular communication.**
    - [Bottleneck]: Limited technologies for the exchange of biochemical information within a population of cells.
        - [Potential Solution]: Use available cell-cell communication regulators to enable cell-cell communication in broader range of organisms.
        - [Potential Solution]: Expand communication systems by using a broader range of natural quorum-sensing/communication modules (such as acyl-

homoserine lactones, autoinducer-2, peptide-based signaling, and metabolic signaling).

- **5 years: Synthetic cell-to-cell communication elements and networks that function in a broad range of host organisms.**
    - [Bottleneck]: Current synthetic communication systems have been engineered to function in a limited set of organisms.
        - [Potential Solution]: Identify cell-cell communication elements in non-traditional hosts, characterize, and modify for use in synthetic circuits.
        - [Potential Solution]: Engineered membranes (specifically, receptors) to transmit information via cell-cell contact.
        - [Potential Solution]: Engineered transmission of secreted biomolecules and exosomes.
- **10 years: Signal-response pathways that function in synthetic communities of 5-10 organisms, employing a variety of pathway types and host species.**
    - [Bottleneck]: Cross-talk between communication elements.
        - [Potential Solution]: Combine communication modules in a manner that minimizes cross-talk; employ metabolic signaling to coordinate behavior across population as needed.
    - [Bottleneck]: Feasibility of enabling signaling between all community members.
        - [Potential Solution]: Design networks with essential community-level coordination using a limited set of communication modules.
- **20 years: Ability to produce engineered microorganisms that can reliably invade and coexist within a complex community and manipulate the consortium/biome function and behavior.**
    - [Bottleneck]: Ecological understanding of complex microbiomes, rules of coexistence, cooperation, and competition.
        - [Potential Solution]: Characterization of broad ranging natural, as well as synthetic, communities during environmental and ecological changes.
        - [Potential Solution]: Employ cell-cell communication or metabolic interactions to enable engineered cells to be accepted into/required by the target community.

[Breakthrough Capability 2]: Ability to characterize, manipulate, and program the three-dimensional (3D) architecture of a biome (i.e., the "ecosystem" of a natural or manipulated biome containing multiple species).

- **2 years: Use of existing technologies (including metagenomics, transcriptomics, proteomics, and mass spectrometry) to better understand the species composition and collective components of microbial communities and consortia.**
    - [Bottleneck]: Need data from diverse ecosystems, environments to build predictive models.
        - [Potential Solution]: Generate and include data from many environments so that models can integrate information with respect to the community, what they are doing, and the environment they are inhabiting.

- **5 years: Non-destructive, 3D visualization of microbial communities from a broad range of environments.**
    - [Bottleneck]: New technologies needed to report and visualize 3D structures and functions of consortia.
        - [Potential Solution]: Adapt imaging, sequencing, -omics technologies to characterize natural or engineered systems and their dynamics.
        - [Potential Solution]: Develop new reporter systems, such as cell-based sensors, to assess and quantify function and/or 3D organization.
- **10 years: Ability to manipulate the 3D architecture of natural or engineered communities using external inputs (such as molecules, temperature, or pH).**
    - [Bottleneck]: Limited understanding of how communities respond dynamically to environmental changes, especially in non-homogeneous systems.
        - [Potential Solution]: Characterization of how natural and engineered communities respond to environmental changes; build spatio-temporal models that incorporate genomic, functional, and environmental outcomes.
    - [Bottleneck]: Ability to add sensing and actuation capabilities to any cell type in a community setting.
        - [Potential Solution]: Targeted gene editing approaches that deliver only to a specific organismal-member of a community.
- **20 years: Programmed communities that self-assemble into a desired 3D architecture.**
    - [Bottleneck]: Tools that enable desired stratification and self-organization (or reorganization) of microbial communities.
        - [Potential Solution]: Use strategies from multi-organismal cell development, such as the generation and sensing of gradients and motility.
        - [Potential Solution]: Specific binding between cells using extracellularly-displayed proteins to build or lock in specific levels of organization.

[Breakthrough Capability 3]: Ability to control and/or define the function of an engineered microbial community/biome.

- **2 years: Ability to combine species with specialized functions to enable the production of desired products.**
    - [Bottleneck]: Growth rates and conditions ideal for production may vary between species.
        - [Potential Solution]: Apply synthetic ecological approaches, including identifying optimal growth conditions for the community, such as engineering mutualistic interactions to control community composition.
- **5 years: Assembly of consortia to produce desired molecules/products, considering community-level metabolic flux.**
    - [Bottleneck]: Ideal division of labor within the consortium is difficult to predict.

        - [Potential Solution]: Develop metabolic engineering approaches to separate processes (such as reducing metabolic load, balancing redox and cofactor use), ideally amongst community members.
- **10 years: Plug-and-play assembly of consortia to produce desired molecules/products from specific starting materials, considering community level metabolic flux and organism-to-organism communication.** For example, developing consortia of different microbial species that are grown/fermented together to create a desired product.
    - [Bottleneck]: Optimal growth/production conditions of the individual community members are likely to be different.
        - [Potential Solution]: Screen for, or predict, conditions that are optimal for the community; tune relative population densities through inoculation ratios and via feedback (cell-to-cell communication).
        - [Potential Solution]: Engineer community members to function optimally under target bioreactor/process conditions.
- **20 years: On-demand assembly of consortia that are programmed to respond dynamically, such that they can use different feedstocks, metabolize toxins or toxic byproducts, or produce different products in response to endogenous (system) or exogenous (user) cues.**
    - [Bottleneck]: New strategies needed for holistic engineering of consortia that can work under a broad range of conditions off the shelf.
        - [Potential Solution]: Develop reliable building blocks (organisms, communication modules, sensors, enzymes, metabolic processes) that can be recombined reliably and adapted for specific applications.

[Breakthrough Capability 4]: Targeted modification of an existing microbiome to enable new functions or address dysbiosis - at the host, community, or environment level - through the addition, removal, or reorganization of the community members.

- **2 years: Use of existing technologies (including metagenomics, transcriptomics, proteomics, and mass spectrometry) to characterize functions of microbial communities from a broad range of environments.**
    - [Bottleneck]: Need data from diverse ecosystems, environments to build predictive models.
        - [Potential Solution]: Generate and include data from many environments so that models can integrate information with respect to the community, including major functions and surrounding ecosystem.
- **5 years: Characterize how *select* microbiomes respond to changes in the environment, including the addition of toxins, the introduction new organisms (pathogens or commensals), and the selective removal of species from the community.**
    - [Bottleneck]: Ability to selectively remove species from a biome.
        - [Potential Solution]: Targeted anti-microbials.

- **10 years: Predictive models of microbiome function and response to a broad range of environmental and ecological changes.**
    - [Bottleneck]: Need to be able to undertake modeling and comparative pathway analysis to determine most robust, prioritized, and resilient systems.
        - [Potential Solution]: Controlled laboratory experiments or observational studies where microbiome function is determined as a function of community composition and environment.
        - [Potential Solution]: Machine learning approaches to determine whether there are patterns between microbiomes of interest (with respect to both form and function); follow-up with strategies designed to improve mechanistic understanding, as needed.
- **20 years: Ability to modify an existing biome or consortia as desired.**
  Biome/consortia modifications include: 1) adding functions such as the ability to sense the environment and coordinate responses for defined outcomes (pathogen defense, substrate transformation, and biosensing), and 2) manipulating composition of host-associated communities to address dysbiosis or add new functions.
    - [Bottleneck]: Need controlled invasion of non-native organism with desired properties or the ability to reintroduce a host-associated organism after engineering.
        - [Potential Solution]: *In situ* gene editing to add function to a community with minimal disruption.
    - [Bottleneck]: Ability to selectively add and remove community members, including engineered cells, and have them persist in a community for a desired length of time and at a desired population density.
        - [Potential Solution]: Kill switches (to remove organisms), auxotrophy/complementation and shared metabolisms (to retain organisms).
    - [Bottleneck]: Need an integrated understanding of microbial community ecology and function; ability to predict how adding a new member with desired functions will affect community health and stability.
        - [Potential Solution]: Focus on developing organisms that can integrate into host consortia and deliver the required functions; recent evidence suggests that function is more important that what organism is carrying it out.

References

Alford, R. F., Leaver-Fay, A., Jeliazkov, J. R., O'Meara, M. J., DiMaio, F. P., Park, H., Shapovalov MV, Renfrew PD, Mulligan VK, Kappel K, Labonte JW, Pacella MS, Bonneau R, Bradley P, Dunbrack RL, Das R, Baker D, Kuhlman B, Kortemme T, Gray, J. J. (2017). The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *Journal of Chemical Theory and Computation*, *13*(6), 3031–3048. https://doi.org/10.1021/acs.jctc.7b00125

Ali, H., & Khan, E. (2018). Trophic transfer, bioaccumulation, and biomagnification of non-essential hazardous heavy metals and metalloids in food chains/webs—Concepts and implications for wildlife and human health. *Human and Ecological Risk Assessment: An International Journal*, 1–24. https://doi.org/10.1080/10807039.2018.1469398

AlQuraishi, M. (2019). AlphaFold at CASP13. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btz422

Badran, A. H., & Liu, D. R. (2015). In vivo continuous directed evolution. *Current Opinion in Chemical Biology*, *24*, 1–10. https://doi.org/10.1016/j.cbpa.2014.09.040

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S., & Milo, R. (2011). The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*, *50*(21), 4402–4410. https://doi.org/10.1021/bi2002289

Bier, E., Harrison, M. M., O'Connor-Giles, K. M., & Wildonger, J. (2018). Advances in Engineering the Fly Genome with the CRISPR-Cas System. *Genetics*, *208*(1), 1–18. https://doi.org/10.1534/genetics.117.1113

Blind, M., & Blank, M. (2015). Aptamer selection technology and recent advances. *Molecular Therapy. Nucleic Acids*, *4*, e223. https://doi.org/10.1038/mtna.2014.74

Boeing, P., Leon, M., Nesbeth, D. N., Finkelstein, A., & Barnes, C. P. (2018). Towards an Aspect-Oriented Design and Modelling Framework for Synthetic Biology. *Processes (Basel, Switzerland)*, *6*(9), 167. https://doi.org/10.3390/pr6090167

Cambray, G., Guimaraes, J. C., & Arkin, A. P. (2018). Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in Escherichia coli. *Nature Biotechnology*, *36*(10), 1005–1015. https://doi.org/10.1038/nbt.4238

Carlson, P. D., & Lucks, J. B. (2019). Elements of RNA design. *Biochemistry*, *58*(11), 1457–1459. https://doi.org/10.1021/acs.biochem.8b01129

Carothers, J. M., Goler, J. A., Juminaga, D., & Keasling, J. D. (2011). Model-driven engineering of RNA devices to quantitatively program gene expression. *Science*, *334*(6063), 1716–1719. https://doi.org/10.1126/science.1212209

Carothers, J. M., Oestreich, S. C., Davis, J. H., & Szostak, J. W. (2004). Informational complexity and functional activity of RNA structures. *Journal of the American Chemical Society*, *126*(16), 5130–5137. https://doi.org/10.1021/ja031504a

Chappell, J., Westbrook, A., Verosloff, M., & Lucks, J. B. (2017). Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. *Nature Communications*, *8*(1), 1051. https://doi.org/10.1038/s41467-017-01082-6

Cherry, K. M., & Qian, L. (2018). Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks. *Nature*, *559*(7714), 370–376. https://doi.org/10.1038/s41586-018-0289-6

Clark, D. S., & Blanch, H. W. (1997). *Biochemical Engineering (Chemical Industries)* (2nd ed., p. 716). New York, New York: Crc Press.

Costello, Z., & Martin, H. G. (2018). A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *Npj Systems Biology and Applications*, *4*, 19. https://doi.org/10.1038/s41540-018-0054-3

Cox, J. C., Hayhurst, A., Hesselberth, J., Bayer, T. S., Georgiou, G., & Ellington, A. D. (2002). Automated selection of aptamers against protein targets translated in vitro: from gene to aptamer. *Nucleic Acids Research*, *30*(20), e108. https://doi.org/10.1093/nar/gnf107

Cuperus, J. T., Groves, B., Kuchina, A., Rosenberg, A. B., Jojic, N., Fields, S., & Seelig, G. (2017). Deep learning of the regulatory grammar of yeast 5' untranslated regions from 500,000 random sequences. *Genome Research*, *27*(12), 2015–2024. https://doi.org/10.1101/gr.224964.117

Das, R., Karanicolas, J., & Baker, D. (2010). Atomic accuracy in predicting and designing noncanonical RNA structure. *Nature Methods*, *7*(4), 291–294. https://doi.org/10.1038/nmeth.1433

Davey, J. A., Damry, A. M., Goto, N. K., & Chica, R. A. (2017). Rational design of proteins that exchange on functional timescales. *Nature Chemical Biology*, *13*(12), 1280–1285. https://doi.org/10.1038/nchembio.2503

de Kok, S., Stanton, L. H., Slaby, T., Durot, M., Holmes, V. F., Patel, K. G., Platt D, Shapland EB, Serber Z, Dean J, Newman JD, Chandran, S. S. (2014). Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synthetic Biology*, *3*(2), 97–106. https://doi.org/10.1021/sb4001992

Dehingia, M., Adak, A., & Khan, M. R. (2019). Ethnicity-Influenced Microbiota: A Future Healthcare Perspective. *Trends in Microbiology*, *27*(3), 191–193. https://doi.org/10.1016/j.tim.2019.01.002

DeLoache, W. C., Russ, Z. N., & Dueber, J. E. (2016). Towards repurposing the yeast peroxisome for compartmentalizing heterologous metabolic pathways. *Nature Communications*, *7*, 11152. https://doi.org/10.1038/ncomms11152

Doudna, J. A., & Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 1258096. https://doi.org/10.1126/science.1258096

Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature*, *346*(6287), 818–822. https://doi.org/10.1038/346818a0

Engler, C., Kandzia, R., & Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *Plos One*, *3*(11), e3647. https://doi.org/10.1371/journal.pone.0003647

Espah Borujeni, A., Mishler, D. M., Wang, J., Huso, W., & Salis, H. M. (2016). Automated physics-based design of synthetic riboswitches from diverse RNA aptamers. *Nucleic Acids Research*, *44*(1), 1–13. https://doi.org/10.1093/nar/gkv1289

Espah Borujeni, A., & Salis, H. M. (2016). Translation initiation is controlled by RNA folding kinetics via a ribosome drafting mechanism. *Journal of the American Chemical Society*, *138*(22), 7016–7023. https://doi.org/10.1021/jacs.6b01453

Esvelt, K. M., Carlson, J. C., & Liu, D. R. (2011). A system for the continuous directed evolution of biomolecules. *Nature*, *472*(7344), 499–503. https://doi.org/10.1038/nature09929

Fan, W., Guo, Q., Liu, C., Liu, X., Zhang, M., Long, D., Xiang, Z., Zhao, A. (2018). Two mulberry phytochelatin synthase genes confer zinc/cadmium tolerance and accumulation in transgenic Arabidopsis and tobacco. *Gene*, *645*, 95–104. https://doi.org/10.1016/j.gene.2017.12.042

Farré, G., Blancquaert, D., Capell, T., Van Der Straeten, D., Christou, P., & Zhu, C. (2014). Engineering complex metabolic pathways in plants. *Annual Review of Plant Biology*, *65*, 187–223. https://doi.org/10.1146/annurev-arplant-050213-035825

Galdzicki, M., Clancy, K. P., Oberortner, E., Pocock, M., Quinn, J. Y., Rodriguez, C. A., Roehner N, Wilson ML, Adam L, Anderson JC, Bartley BA, Beal J, Chandran D, Chen J, Densmore D, Endy D, Grünberg R, Hallinan J, Hillson NJ, Johnson JD, Kuchinsky A, Lux M, Misirli G, Peccoud J, Plahar HA, Sirin E, Stan GB, Villalobos A, Wipat A, Gennari JH, Myers CJ, Sauro, H. M. (2014). The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nature Biotechnology*, *32*(6), 545–550. https://doi.org/10.1038/nbt.2891

Gantz, V. M., & Bier, E. (2015). Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science*, *348*(6233), 442–444. https://doi.org/10.1126/science.aaa5945

Gantz, V. M., Jasinskiene, N., Tatarenkova, O., Fazekas, A., Macias, V. M., Bier, E., & James, A. A. (2015). Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(49), E6736–43. https://doi.org/10.1073/pnas.1521077112

Gibson, D. G. (2011). Enzymatic assembly of overlapping DNA fragments. *Methods in Enzymology*, *498*, 349–361. https://doi.org/10.1016/B978-0-12-385120-8.00015-2

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318

Gilbert, J. A., & Melton, L. (2018). Verily project releases millions of factory-reared mosquitoes. *Nature Biotechnology*, *36*(9), 781–782. https://doi.org/10.1038/nbt0918-781a

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi, L. S. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, *154*(2), 442–451. https://doi.org/10.1016/j.cell.2013.06.044

Goldsmith, M., & Tawfik, D. S. (2017). Enzyme engineering: reaching the maximal catalytic efficiency peak. *Current Opinion in Structural Biology*, *47*, 140–150. https://doi.org/10.1016/j.sbi.2017.09.002

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews. Genetics*, *17*(6), 333–351. https://doi.org/10.1038/nrg.2016.49

Green, A. A., Silver, P. A., Collins, J. J., & Yin, P. (2014). Toehold switches: de-novo-designed regulators of gene expression. *Cell*, *159*(4), 925–939. https://doi.org/10.1016/j.cell.2014.10.002

Grunwald, H. A., Gantz, V. M., Poplawski, G., Xu, X.-R. S., Bier, E., & Cooper, K. L. (2019). Super-Mendelian inheritance mediated by CRISPR-Cas9 in the female mouse germline. *Nature*, *566*(7742), 105–109. https://doi.org/10.1038/s41586-019-0875-2

Haitjema, C. H., Solomon, K. V., Henske, J. K., Theodorou, M. K., & O'Malley, M. A. (2014). Anaerobic gut fungi: Advances in isolation, culture, and cellulolytic enzyme discovery for biofuel production. *Biotechnology and Bioengineering*, *111*(8), 1471–1482. https://doi.org/10.1002/bit.25264

Halperin, S. O., Tou, C. J., Wong, E. B., Modavi, C., Schaffer, D. V., & Dueber, J. E. (2018). CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature*, *560*(7717), 248–252. https://doi.org/10.1038/s41586-018-0384-8

Ham, T. S., Dmytriv, Z., Plahar, H., Chen, J., Hillson, N. J., & Keasling, J. D. (2012). Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic Acids Research*, *40*(18), e141. https://doi.org/10.1093/nar/gks531

Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S. N., Richelle, A., Heinken, A., Haraldsdóttir HS, Wachowiak J, Keating SM, Vlasov V, Magnusdóttir S, Ng CY, Preciat G, Žagare A, Chan SHJ, Aurich MK, Clancy CM, Modamio J, Sauls JT, Noronha A, Bordbar A, Cousins B, El Assal DC, Valcarcel LV, Apaolaza I, Ghaderi S, Ahookhosh M, Ben Guebila M, Kostromins A, Sompairac N, Le HM, Ma D, Sun Y, Wang L, Yurkovich JT, Oliveira MAP, Vuong PT, El Assal LP, Kuperstein I, Zinovyev A, Hinton HS, Bryant WA, Aragón Artacho FJ, Planes FJ, Stalidzans E, Maass A, Vempala S, Hucka M, Saunders MA, Maranas CD, Lewis NE, Sauter T, Palsson BØ, Thiele I, Vlasov, V. (2019). Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nature Protocols*, *14*(3), 639–702. https://doi.org/10.1038/s41596-018-0098-2

Higgins, S. A., & Savage, D. F. (2018). Protein science by DNA sequencing: how advances in molecular biology are accelerating biochemistry. *Biochemistry*, *57*(1), 38–46. https://doi.org/10.1021/acs.biochem.7b00886

Hillson, N., Caddick, M., Cai, Y., Carrasco, J. A., Chang, M. W., Curach, N. C., Bell DJ, Le Feuvre R, Friedman DC, Fu X, Gold ND, Herrgård MJ, Holowko MB, Johnson JR, Johnson RA, Keasling JD, Kitney RI, Kondo A, Liu C, Martin VJJ, Menolascina F, Ogino C, Patron NJ, Pavan M, Poh CL, Pretorius IS, Rosser SJ, Scrutton NS, Storch M, Tekotte H, Travnik E, Vickers CE, Yew WS, Yuan Y, Zhao H, Freemont, P. S. (2019). Building a global alliance of biofoundries. *Nature Communications*, *10*(1), 2040. https://doi.org/10.1038/s41467-019-10079-2

Hooker, S. E., Woods-Burnham, L., Bathina, M., Lloyd, S. M., Gorjala, P., Mitra, R., Nonn L, Kimbro KS, Kittles, R. (2019). Genetic ancestry analysis reveals misclassification of commonly used cancer cell lines. *Cancer Epidemiology, Biomarkers & Prevention*. https://doi.org/10.1158/1055-9965.EPI-18-1132

Hoshika, S., Leal, N. A., Kim, M.-J., Kim, M.-S., Karalkar, N. B., Kim, H.-J., Bates AM, Watkins NE, SantaLucia HA, Meyer AJ, DasGupta S, Piccirilli JA, Ellington AD, SantaLucia J, Georgiadis MM, Benner, S. A. (2019). Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science*, *363*(6429), 884–887. https://doi.org/10.1126/science.aat0971

Hsiao, V., Cheng, A., & Murray, R. M. (2016). *Design and application of stationary phase combinatorial promoters*. MurrayWiki. Retrieved from http://www.cds.caltech.edu/~murray/preprints/hcm16-seed_s.pdf

Huang, A., Nguyen, P. Q., Stark, J. C., Takahashi, M. K., Donghia, N., Ferrante, T., Dy AJ, Hsu KJ, Dubner RS, Pardee K, Jewett MC, Collins, J. J. (2018). BioBits™ Explorer: A modular synthetic biology education kit. *Science Advances*, *4*(8), eaat5105. https://doi.org/10.1126/sciadv.aat5105

Hughes, R. A., & Ellington, A. D. (2017). Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harbor Perspectives in Biology*, *9*(1). https://doi.org/10.1101/cshperspect.a023812

Jakobson, C. M., Chen, Y., Slininger, M. F., Valdivia, E., Kim, E. Y., & Tullman-Ercek, D. (2016). Tuning the catalytic activity of subcellular nanoreactors. *Journal of Molecular Biology*, *428*(15), 2989–2996. https://doi.org/10.1016/j.jmb.2016.07.006

Jeske, L., Placzek, S., Schomburg, I., Chang, A., & Schomburg, D. (2019). BRENDA in 2019: a European ELIXIR core data resource. *Nucleic Acids Research*, *47*(D1), D542–D549. https://doi.org/10.1093/nar/gky1048

Johns, N. I., Gomes, A. L. C., Yim, S. S., Yang, A., Blazejewski, T., Smillie, C. S., Smith MB, Alm EJ, Kosuri S, Wang, H. H. (2018). Metagenomic mining of regulatory elements enables programmable species-selective gene expression. *Nature Methods*, *15*(5), 323–329. https://doi.org/10.1038/nmeth.4633

Jones, H. P., Holmes, N. D., Butchart, S. H. M., Tershy, B. R., Kappes, P. J., Corkery, I., Aguirre-Muñoz A, Armstrong DP, Bonnaud E, Burbidge AA, Campbell K, Courchamp F, Cowan PE, Cuthbert RJ, Ebbert S, Genovesi P, Howald GR, Keitt BS, Kress SW, Miskelly CM, Oppel S, Poncet S, Rauzon MJ, Rocamora G, Russell JC, Samaniego-Herrera A, Seddon PJ, Spatz DR, Towns DR, Croll, D. A. (2016). Invasive mammal eradication on islands results in substantial conservation gains. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(15), 4033–4038. https://doi.org/10.1073/pnas.1521179113

Karim, A. S., & Jewett, M. C. (2016). A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. *Metabolic Engineering*, *36*, 116–126. https://doi.org/10.1016/j.ymben.2016.03.002

Kedzierska, K., Valkenburg, S. A., Doherty, P. C., Davenport, M. P., & Venturi, V. (2012). Use it or lose it: establishment and persistence of T cell memory. *Frontiers in Immunology*, *3*, 357. https://doi.org/10.3389/fimmu.2012.00357

Kim, E. Y., & Tullman-Ercek, D. (2014). A rapid flow cytometry assay for the relative quantification of protein encapsulation into bacterial microcompartments. *Biotechnology Journal*, *9*(3), 348–354. https://doi.org/10.1002/biot.201300391

Kong, W., Meldgin, D. R., Collins, J. J., & Lu, T. (2018). Designing microbial consortia with defined social interactions. *Nature Chemical Biology*, *14*(8), 821–829. https://doi.org/10.1038/s41589-018-0091-7

Kosuri, S., & Church, G. M. (2014). Large-scale de novo DNA synthesis: technologies and applications. *Nature Methods*, *11*(5), 499–507. https://doi.org/10.1038/nmeth.2918

Kuo-chen, C., & Shou-ping, J. (1974). Studies on the rate of diffusion-controlled reactions of enzymes. Spatial factor and force field factor. *Scientia Sinica*, *27*(5), 664–680.

Kyrou, K., Hammond, A. M., Galizi, R., Kranjc, N., Burt, A., Beaghton, A. K., Nolan, T., Crisanti, A. (2018). A CRISPR-Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nature Biotechnology*, *36*(11), 1062–1066. https://doi.org/10.1038/nbt.4245

Lane, R. S., & Quistad, G. B. (1998). Borreliacidal Factor in the Blood of the Western Fence Lizard (Sceloporus occidentalis). *The Journal of Parasitology*, *84*(1), 29. https://doi.org/10.2307/3284524

Leistra, A. N., Amador, P., Buvanendiran, A., Moon-Walker, A., & Contreras, L. M. (2017). Rational modular RNA engineering based on in vivo profiling of structural accessibility. *ACS Synthetic Biology*, *6*(12), 2228–2240. https://doi.org/10.1021/acssynbio.7b00185

Leistra, A. N., Curtis, N. C., & Contreras, L. M. (2019). Regulatory non-coding sRNAs in bacterial metabolic pathway engineering. *Metabolic Engineering*, *52*, 190–214. https://doi.org/10.1016/j.ymben.2018.11.013

Li, M. Z., & Elledge, S. J. (2007). Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nature Methods*, *4*(3), 251–256. https://doi.org/10.1038/nmeth1010

Linshiz, G., Stawski, N., Poust, S., Bi, C., Keasling, J. D., & Hillson, N. J. (2013). PaR-PaR laboratory automation platform. *ACS Synthetic Biology*, *2*(5), 216–222. https://doi.org/10.1021/sb300075t

Long, M. R., Ong, W. K., & Reed, J. L. (2015). Computational methods in metabolic engineering for strain design. *Current Opinion in Biotechnology*, *34*, 135–141. https://doi.org/10.1016/j.copbio.2014.12.019

Looi, F. Y., Baker, M. L., Townson, T., Richard, M., Novak, B., Doran, T. J., & Short, K. R. (2018). Creating disease resistant chickens: A viable solution to avian influenza? *Viruses*, *10*(10). https://doi.org/10.3390/v10100561

Lubertozzi, D., & Keasling, J. D. (2009). Developing Aspergillus as a host for heterologous expression. *Biotechnology Advances*, *27*(1), 53–75. https://doi.org/10.1016/j.biotechadv.2008.09.001

Ma, S., Saaem, I., & Tian, J. (2012). Error correction in gene synthesis technology. *Trends in Biotechnology*, 30(3), 147–154. https://doi.org/10.1016/j.tibtech.2011.10.002

Markley, A. L., Begemann, M. B., Clarke, R. E., Gordon, G. C., & Pfleger, B. F. (2015). Synthetic biology toolbox for controlling gene expression in the cyanobacterium Synechococcus sp. strain PCC 7002. *ACS Synthetic Biology*, *4*(5), 595–603. https://doi.org/10.1021/sb500260k

Martin, R. W., Des Soye, B. J., Kwon, Y.-C., Kay, J., Davis, R. G., Thomas, P. M., Majewska NI, Chen CX, Marcum RD, Weiss MG, Stoddart AE, Amiram M, Ranji Charna AK, Patel JR, Isaacs FJ, Kelleher NL, Hong SH, Jewett, M. C. (2018). Cell-free protein synthesis from genomically recoded bacteria enables multisite incorporation of noncanonical amino acids. *Nature Communications*, *9*(1), 1203. https://doi.org/10.1038/s41467-018-03469-5

McCarty, N. S., & Ledesma-Amaro, R. (2019). Synthetic biology tools to engineer microbial communities for biotechnology. *Trends in Biotechnology*, *37*(2), 181–197. https://doi.org/10.1016/j.tibtech.2018.11.002

McDermott, J., & Hardeman, M. (2018). Increasing Your Research's Exposure on Figshare Using the FAIR Data Principles. *Figshare*. https://doi.org/10.6084/m9.figshare.7429559.v2

Medema, M. H., van Raaphorst, R., Takano, E., & Breitling, R. (2012). Computational tools for the synthetic design of biochemical pathways. *Nature Reviews. Microbiology*, *10*(3), 191–202. https://doi.org/10.1038/nrmicro2717

Molteni, M. (2019, March 10). 23andMe's New Diabetes Test Has Experts Asking Who It's For | WIRED. Retrieved May 21, 2019, from https://www.wired.com/story/23andmes-new-diabetes-test-has-experts-asking-who-its-for/

Moore, S. J., MacDonald, J. T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., Kylilis N, Bell DJ, McClymont DW, Jensen K, Polizzi KM, Biedendieck R, Freemont, P. S. (2018). Rapid acquisition and model-based analysis of cell-free transcription-translation reactions from nonmodel bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(19), E4340–E4349. https://doi.org/10.1073/pnas.1715806115

Morrell, W. C., Birkel, G. W., Forrer, M., Lopez, T., Backman, T. W. H., Dussault, M., Petzold CJ, Baidoo EEK, Costello Z, Ando D, Alonso-Gutierrez J, George KW, Mukhopadhyay A, Vaino I, Keasling JD, Adams PD, Hillson NJ, Garcia Martin, H. (2017). The Experiment Data Depot: A Web-Based Software Tool for Biological Experimental Data Storage, Sharing, and Visualization. *ACS Synthetic Biology*, *6*(12), 2248–2259. https://doi.org/10.1021/acssynbio.7b00204

Muthusaravanan, S., Sivarajasekar, N., Vivek, J. S., Paramasivan, T., Naushad, M., Prakashmaran, J., Gayathri V, Al-Duaij, O. K. (2018). Phytoremediation of heavy metals: mechanisms, methods and enhancements. *Environmental Chemistry Letters*, *16*(4), 1–21. https://doi.org/10.1007/s10311-018-0762-3

Nahar, N., Rahman, A., Nawani, N. N., Ghosh, S., & Mandal, A. (2017). Phytoremediation of arsenic from the contaminated soil using transgenic tobacco plants expressing ACR2 gene of Arabidopsis thaliana. *Journal of Plant Physiology*, *218*, 121–126. https://doi.org/10.1016/j.jplph.2017.08.001

Naran, K., Nundalall, T., Chetty, S., & Barth, S. (2018). Principles of immunotherapy: implications for treatment strategies in cancer and infectious diseases. *Frontiers in Microbiology*, *9*, 3158. https://doi.org/10.3389/fmicb.2018.03158

National Academies of Sciences, Engineering, and Medicine, Division on Earth and Life Studies, Board on Life Sciences, Board on Chemical Sciences and Technology, & Committee on Strategies for Identifying and Addressing Potential Biodefense Vulnerabilities Posed by Synthetic Biology. (2018). *Biodefense in the age of synthetic biology*. Washington (DC): National Academies Press (US). https://doi.org/10.17226/24890

National Research Council (US) Committee on Assessing the Importance and Impact of Glycomics and Glycosciences. (2012). *Transforming glycoscience: A roadmap for the future*. Washington (DC): National Academies Press (US). https://doi.org/10.17226/13446

National Research Council (US) Committee on Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals, Board on Chemical Sciences and Technology, Board on Life Sciences, Division on Earth and Life Studies. (2015). *Industrialization of biology: A roadmap to accelerate the advanced manufacturing of chemicals.* Washington (DC): National Academies Press (US). https://doi.org/10.17226/19001

National Research Council (US) Committee on Scientific Evaluation of the Introduction of Genetically Modified Microorganisms and Plants into the Environment. (1989). *Field testing genetically modified organisms: framework for decisions*. Washington (DC): National Academies Press (US). https://doi.org/10.17226/1431

Nielsen, A. A. K., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., Ross D, Densmore D, Voigt, C. A. (2016). Genetic circuit design automation. *Science*, *352*(6281), aac7341. https://doi.org/10.1126/science.aac7341

Niu, D., Wei, H.-J., Lin, L., George, H., Wang, T., Lee, I.-H., Zhao HY, Wang Y, Kan Y, Shrock E, Lesha E, Wang G, Luo Y, Qing Y, Jiao D, Zhao H, Zhou X, Wang S, Wei H, Güell M, Church GM, Yang, L. (2017). Inactivation of porcine endogenous retrovirus in pigs using CRISPR-Cas9. *Science*, *357*(6357), 1303–1307. https://doi.org/10.1126/science.aan4187

Pacheco, A. R., Moel, M., & Segrè, D. (2019). Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nature Communications*, *10*(1), 103. https://doi.org/10.1038/s41467-018-07946-9

Palacino, J., Swalley, S. E., Song, C., Cheung, A. K., Shu, L., Zhang, X., Van Hoosear M, Shin Y, Chin DN, Keller CG, Beibel M, Renaud NA, Smith TM, Salcius M, Shi X, Hild M, Servais R, Jain M, Deng L, Bullock C, McLellan M, Schuierer S, Murphy L, Blommers MJ, Blaustein C, Berenshteyn F, Lacoste A, Thomas JR, Roma G, Michaud GA, Tseng BS, Porter JA, Myer VE, Tallarico JA, Hamann LG, Curtis D, Fishman MC, Dietrich WF, Dales NA, Sivasankaran, R. (2015). SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nature Chemical Biology*, *11*(7), 511–517. https://doi.org/10.1038/nchembio.1837

Palluk, S., Arlow, D. H., de Rond, T., Barthel, S., Kang, J. S., Bector, R., Baghdassarian HM, Truong AN, Kim PW, Singh AK, Hillson NJ, Keasling, J. D. (2018). De novo DNA synthesis using polymerase-nucleotide conjugates. *Nature Biotechnology*, *36*(7), 645–650. https://doi.org/10.1038/nbt.4173

Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., Ferrante T, Ma D, Donghia N, Fan M, Daringer NM, Bosch I, Dudley DM, O'Connor DH, Gehrke L, Collins, J. J. (2016). Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell*, *165*(5), 1255–1266. https://doi.org/10.1016/j.cell.2016.04.059

Pearl, J. (2018). Theoretical Impediments to Machine Learning With Seven Sparks from the Causal Revolution. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining' ' - WSDM '18* (pp. 3–3). New York, New York, USA: ACM Press. https://doi.org/10.1145/3159652.3176182

Plesa, C., Sidore, A. M., Lubock, N. B., Zhang, D., & Kosuri, S. (2018). Multiplexed gene synthesis in emulsions for exploring protein functional landscapes. *Science*, *359*(6373), 343–347. https://doi.org/10.1126/science.aao5167

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, *152*(5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

Qian, L., & Winfree, E. (2011). Scaling up digital circuit computation with DNA strand displacement cascades. *Science*, *332*(6034), 1196–1201. https://doi.org/10.1126/science.1200520

Rai, P. K., Lee, S. S., Zhang, M., Tsang, Y. F., & Kim, K.-H. (2019). Heavy metals in food crops: Health risks, fate, mechanisms, and management. *Environment International*, *125*, 365–385. https://doi.org/10.1016/j.envint.2019.01.067

Ravikumar, A., Arrieta, A., & Liu, C. C. (2014). An orthogonal DNA replication system in yeast. *Nature Chemical Biology*, *10*(3), 175–177. https://doi.org/10.1038/nchembio.1439

Ravikumar, A., Arzumanyan, G. A., Obadi, M. K. A., Javanpour, A. A., & Liu, C. C. (2018). Scalable, Continuous Evolution of Genes at Mutation Rates above Genomic Error Thresholds. *Cell*, *175*(7), 1946–1957.e13. https://doi.org/10.1016/j.cell.2018.10.021

Reva, B. A., Finkelstein, A. V., & Skolnick, J. (1998). What is the probability of a chance prediction of a protein structure with an rmsd of 6 A? *Folding & Design*, *3*(2), 141–147. https://doi.org/10.1016/S1359-0278(98)00019-4

Richardson, S. M., Mitchell, L. A., Stracquadanio, G., Yang, K., Dymond, J. S., DiCarlo, J. E., Lee D, Huang CL, Chandrasegaran S, Cai Y, Boeke JD, Bader, J. S. (2017). Design of a synthetic yeast genome. *Science*, *355*(6329), 1040–1044. https://doi.org/10.1126/science.aaf4557

Ross, M. J., & Coates, P. T. (2018). Using CRISPR to inactivate endogenous retroviruses in pigs: an important step toward safe xenotransplantation? *Kidney International*, *93*(1), 4–6. https://doi.org/10.1016/j.kint.2017.11.004

Schellenberger, J., Lewis, N. E., & Palsson, B. Ø. (2011). Elimination of thermodynamically infeasible loops in steady-state metabolic models. *Biophysical Journal*, *100*(3), 544–553. https://doi.org/10.1016/j.bpj.2010.12.3707

Seelig, G., Soloveichik, D., Zhang, D. Y., & Winfree, E. (2006). Enzyme-free nucleic acid logic circuits. *Science*, *314*(5805), 1585–1588. https://doi.org/10.1126/science.1132493

Sethuraman, N., & Stadheim, T. A. (2006). Challenges in therapeutic glycoprotein production. *Current Opinion in Biotechnology*, *17*(4), 341–346. https://doi.org/10.1016/j.copbio.2006.06.010

Shih, S. C. C., Goyal, G., Kim, P. W., Koutsoubelis, N., Keasling, J. D., Adams, P. D., Hillson, N. J., Singh, A. K. (2015). A versatile microfluidic device for automating synthetic biology. *ACS Synthetic Biology*, *4*(10), 1151–1164. https://doi.org/10.1021/acssynbio.5b00062

Sid, H., & Schusser, B. (2018). Applications of gene editing in chickens: A new era is on the horizon. *Frontiers in Genetics*, *9*, 456. https://doi.org/10.3389/fgene.2018.00456

Smith, H. O., Hutchison, C. A., Pfannkoch, C., & Venter, J. C. (2003). Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(26), 15440–15445. https://doi.org/10.1073/pnas.2237126100

Stark, J. C., Huang, A., Hsu, K. J., Dubner, R. S., Forbrook, J., Marshalla, S., Rodriguez F, Washington M, Rybnicky GA, Nguyen PQ, Hasselbacher B, Jabri R, Kamran R, Koralewski V, Wightkin W, Martinez T, Jewett, M. C. (2019). BioBits Health: Classroom Activities Exploring Engineering, Biology, and Human Health with Fluorescent Readouts. *ACS Synthetic Biology*, *8*(5), 1001–1009. https://doi.org/10.1021/acssynbio.8b00381

Stark, J. C., Huang, A., Nguyen, P. Q., Dubner, R. S., Hsu, K. J., Ferrante, T. C., Anderson M, Kanapskyte A, Mucha Q, Packett JS, Patel P, Patel R, Qaq D, Zondor T, Burke J, Martinez T, Miller-Berry A, Puppala A, Reichert K, Schmid M, Brand L, Hill LR, Chellaswamy JF, Faheem N, Fetherling S, Gong E, Gonzalzles EM, Granito T, Koritsaris J, Nguyen B, Ottman S, Palffy C, Patel A, Skweres S, Slaton A, Woods T, Donghia N, Pardee K, Collins JJ, Jewett, M. C. (2018). BioBits™ Bright: A fluorescent synthetic biology education kit. *Science Advances*, *4*(8), eaat5107. https://doi.org/10.1126/sciadv.aat5107

Stephens, N., Di Silvio, L., Dunsford, I., Ellis, M., Glencross, A., & Sexton, A. (2018). Bringing cultured meat to market: Technical, socio-political, and regulatory challenges in cellular agriculture. *Trends in Food Science & Technology*, *78*, 155–166. https://doi.org/10.1016/j.tifs.2018.04.010

Sundstrom, E. R., & Criddle, C. S. (2015). Optimization of Methanotrophic Growth and Production of Poly(3-Hydroxybutyrate) in a High-Throughput Microbioreactor System. *Applied and Environmental Microbiology*, *81*(14), 4767–4773. https://doi.org/10.1128/AEM.00025-15

Takahashi, M. K., Chappell, J., Hayes, C. A., Sun, Z. Z., Kim, J., Singhal, V., Spring KJ, Al-Khabouri S, Fall CP, Noireaux V, Murray RM, Lucks, J. B. (2015). Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems. *ACS Synthetic Biology*, *4*(5), 503–515. https://doi.org/10.1021/sb400206c

Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, *249*(4968), 505–510. https://doi.org/10.1126/science.2200121

Van Enennaam, A. (2018, June 12). Use of Gene Editing to Introduce the Polled Trait into Elite Germplasm. Retrieved February 26, 2019, from https://www.dairyherd.com/article/use-gene-editing-introduce-polled-trait-elite-germplasm/

Venayak, N., von Kamp, A., Klamt, S., & Mahadevan, R. (2018). MoVE identifies metabolic valves to switch between phenotypic states. *Nature Communications*, *9*(1), 5332. https://doi.org/10.1038/s41467-018-07719-4

Villa, J. K., Su, Y., Contreras, L. M., & Hammond, M. C. (2018). Synthetic biology of small rnas and riboswitches. *Microbiology Spectrum*, *6*(3). https://doi.org/10.1128/microbiolspec.RWR-0007-2017

Watkins, A. M., Geniesse, C., Kladwang, W., Zakrevsky, P., Jaeger, L., & Das, R. (2018). Blind prediction of noncanonical RNA structure at atomic accuracy. *Science Advances*, *4*(5), eaar5316. https://doi.org/10.1126/sciadv.aar5316

Weinhandl, K., Winkler, M., Glieder, A., & Camattari, A. (2014). Carbon source dependent promoters in yeasts. *Microbial Cell Factories*, *13*, 5. https://doi.org/10.1186/1475-2859-13-5

Wen, K. Y., Cameron, L., Chappell, J., Jensen, K., Bell, D. J., Kelwick, R., Kopniczky M, Davies JC, Filloux A, Freemont, P. S. (2017). A Cell-Free Biosensor for Detecting Quorum Sensing Molecules in P. aeruginosa-Infected Respiratory Samples. *ACS Synthetic Biology*, *6*(12), 2293–2301. https://doi.org/10.1021/acssynbio.7b00219

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A.,

Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJ, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PA, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone SA, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. https://doi.org/10.1038/sdata.2016.18

Yang, K. K., Wu, Z., & Arnold, F. H. (2018). Machine learning-guided directed evolution for protein engineering. Retrieved from https://arxiv.org/abs/1811.10775v2

Yin, G., Garces, E. D., Yang, J., Zhang, J., Tran, C., Steiner, A. R., Roos C, Bajad S, Hudak S, Penta K, Zawada J, Pollitt S, Murray, C. J. (2012). Aglycosylated antibodies and antibody fragments produced in a scalable in vitro transcription-translation system. *MAbs*, *4*(2), 217–225. https://doi.org/10.4161/mabs.4.2.19202

You, M., & Jaffrey, S. R. (2015). Designing optogenetically controlled RNA for regulating biological systems. *Annals of the New York Academy of Sciences*, *1352*, 13–19. https://doi.org/10.1111/nyas.12660

Zhong, Z., & Liu, C. C. (2019). Probing pathways of adaptation with continuous evolution. *Current Opinion in Systems Biology*. https://doi.org/10.1016/j.coisb.2019.02.002