Gene Editing, Synthesis, and Assembly

(Engineering DNA)

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.

Gene Editing, Synthesis, and Assembly

Summary

Gene Editing, Synthesis, and Assembly focuses on the development and advancement of tools to enable the production of chromosomal DNA and the engineering of entire genomes. Advancements are needed in the design and construction of functional genetic systems though the synthesis of long oligonucleotides, assembly of multiple fragments, and precision editing with high specificity.

Introduction and Impact

Fundamentally, an organism's sensing, metabolic, and decision-making capabilities are all encoded within their genome, a very long double-stranded DNA molecule. By changing an organism's genome sequence, we have the ability to rationally alter these cellular functions, and thereby engineer them to address a myriad of societal challenges. The ability to rationally alter DNA sequences, combining gene editing, DNA synthesis, and DNA assembly, are therefore considered a cornerstone capability of engineering biology, enabling us to construct engineered genetic systems to reprogram organisms with targeted functions. Advances in gene editing, synthesis, and assembly have significant transformative impacts on all sectors impacted by engineering biology by broadening the complexity and breadth of functionality that can be introduced into an engineered organism.

The market for synthesized DNA is both mature and ripe for disruption. Using existing technologies, several service providers currently synthesize single-stranded DNA molecules (oligonucleotides) and double-stranded DNA molecules (DNA fragments). They actively compete across several criteria, including cost-per-DNA base pair, sequence fidelity, turnaround time, confidentiality of intellectual property, and customer service. However, several early-stage technologies have the potential to dramatically alter the commercial landscape by enabling the manufacture of much longer DNA fragments at significantly reduced costs.

Gene Editing, Synthesis, and Assembly highlights several technological routes to achieving the overall goal of manufacturing mega-base length DNA molecules, and designing genes and genomes with desired functionalities. We also illustrate how new technological developments in one process (e.g., oligonucleotide synthesis, or coupled synthesis and sequencing) can directly lead to improvements in downstream processes (e.g., DNA fragment synthesis).

Transformative Tools and Technologies

Oligonucleotide synthesis technologies

Currently, phosphoramidite-based chemistry is the predominant approach for synthesizing oligonucleotides. Even after significant optimization, per-cycle synthesis yields are about 99.5%, meanwhile synthesis of a 200-nucleotide oligonucleotide has a yield of only 35% (Hughes & Ellington, 2017). New technologies seek to improve this process by: 1) synthesizing thousands of oligonucleotides in parallel, using either on-chip supports or within tiny microtiter wells (Kosuri & Church, 2014); or 2) improving synthesis processivity by replacing the

phosphoramidite-based chemistry, for example, using enzyme catalysis (e.g., terminal deoxynucleotidyl transferases) to extend primers with defined nucleotides (Palluk et al., 2018). Clearly, achieving picomole production of 1000-mer oligonucleotides with error-free sequences would significantly improve the overall DNA assembly protocol.

Technologies for oligonucleotide assembly into non-clonal DNA fragments

Currently, multiple 60- to 200-mer oligonucleotides are assembled into non-clonal DNA fragments using a combination of annealing, ligation, and/or polymerase chain reaction. The cost of synthesizing non-clonal DNA fragments is $0.10 to $0.30 per base pair, depending on size and complexity. DNA fragments between 300 and 1800 base pairs can be synthesized by multiple providers and DNA fragments up to 5800 base pairs can be synthesized by select providers at increased cost. Errors are introduced whenever two oligonucleotides form undesired base pairings, when two oligonucleotides are incorrectly ligated together, or when DNA polymerases extend a synthesized DNA fragment with an incorrect nucleotide. Certain sequence determinants will increase the error rate, resulting in a mixture of undesired fragments. Computational sequence design can reduce the frequency of these errors. Mismatch repair enzymes may be added (with added cost) to eliminate DNA fragments with mis-paired nucleotides, for example, as a result of mis-annealing or DNA polymerase errors. This process has been scaled up to assemble thousands of non-clonal DNA fragments per day. The purification of full-length, error-free DNA fragments remains a challenge. Utilizing longer oligonucleotides (see Oligonucleotide synthesis technologies) would enable the synthesis of longer non-clonal DNA fragments with the same error rate. New technologies utilizing nanopore sequencing have the potential to couple sequencing and purification at single-molecule resolution.

Multi-fragment DNA assembly techniques for clonal genetic systems and genomes

Currently, multiple DNA fragments (300 to 3000 base pairs long) are assembled into large genetic systems (10,000 to 1,000,000 base pairs long) using single-pot DNA assembly techniques that combine cocktails of bio-prospected and/or engineered enzymes, including exonucleases, endonucleases, DNA polymerases, ligases, and/or recombinases (Gibson et al., 2009; Hughes & Ellington, 2017). Enzyme costs are currently about $25 per assembly. Assembled DNA is then introduced into cells for clonal separation and replication. Most assembly techniques have essential sequence determinants, for example, regions of overlapping homology or flanking Type IIs restriction sites (Engler, Kandzia, & Marillonnet, 2008). Errors are introduced when two fragments anneal together at incorrect overlap regions, when two fragments are mis-ligated at incorrect ligation junctions, or when DNA polymerases incorporate incorrect nucleotides during DNA synthesis. Computational sequence design can limit the frequency of errors. A major challenge for DNA assembly is the trial-and-error identification of a full-length, error-free genetic system. For example, an optimized assembly technique with a per-junction efficiency of 90% will assemble a 10-part (3000 base pairs per part) system with 35% yield. At the same per-junction efficiency, assembling a 1,000,000 base pair genome from 3000 base pair DNA fragments will have a miniscule yield of 5.2×10^{-14} %. This limitation to DNA assembly has motivated the synthesis of longer non-clonal DNA fragments

(see Technologies for oligonucleotide assembly into non-clonal DNA fragments). For example, 1,000,000 base pair genomes could be assembled from 10,000 base pair, 30,000 base pair, or 50,000 base pair DNA fragments with a 0.002%, 2.7%, or 11% efficiency, respectively. If longer non-clonal DNA fragments are unavailable, then hierarchical approaches to DNA assembly are required, which increases the number of DNA assembly reactions and verification costs.

Sequencing costs become significant once assembled genetic systems are large and/or assembly yields are exceedingly small. For example, after assembling a 30,000 base pair genetic system with a 35% yield, it is necessary to sequence at least seven clonal isolates to achieve at least a 95% chance of identifying a fully-correct one. At low throughput, this cost is about $1000 (using Sanger sequencing). Using next generation sequencing, this cost can be greatly reduced to about $0.70, but only when a large amount of DNA (2 billion base pairs) is sequenced at the same time (Goodwin, McPherson, & McCombie, 2016). Similarly, if a 1,000,000 base pair genome is assembled from 30,000 base pair fragments with a 2.7% yield, then it would be necessary to sequence 100 clonal isolates to achieve a 93% chance of identifying a fully correct one (about $275 in sequencing costs). Finally, hierarchical DNA assembly can be performed by first assembling and purifying smaller genetic systems (e.g., up to 30,000 base pairs) and then using them to perform a multi-fragment assembly to build larger genetic systems (e.g., up to 35 five times larger than the smaller systems) (Richardson et al., 2017). Hierarchical DNA assembly increases sequencing costs by a multiplier roughly equal to the number of hierarchical cycles. Overall, DNA assembly costs are greatly reduced by utilizing longer non-clonal DNA fragments and by parallelizing operations such that at least 2 billion base pairs of DNA are verified across multiple DNA assembly reactions.

GENE EDITING, SYNTHESIS, AND ASSEMBLY

Goal	Breakthough Capability	Milestone

Manufacture thousands of very long oligonucleotides with high fidelity.

Highly efficient oligonucleotide synthesis to increase the number, length, and fidelity of oligonucleotides.			
Robustly synthesize one million 200-mer oligonucleotides with a per-nucleotide error rate of fewer than one in 500 nucleotides.	Robustly synthesize 1,000-mer oligonucleotides with a per-nucleotide error rate of fewer than one in 1,000 nucleotides.	Reduce per-nucleotide error rates for 1,000-mer oligonucleotide synthesis to fewer than one in 5,000 nucleotides.	Synthesize 10,000-mer oligonucleotides at 99.99% cycle efficiency within one minute with a per-nucleotide error rate of fewer than one in 30,000 nucleotides.

Many-fragment DNA assembly with simultaneous, high-fidelity sequence validation.

Predictive design of DNA sequences for improved assembly of longer, more information-rich DNA fragments.			
Coupled design of DNA sequences to optimize nucleotide composition to support synthesis, while maintaining genetic system function.	Incorporate machine learning to identify poorly-understood problematic sequences and process conditions.	Design algorithms that identify optimal synthesis strategies for assembling megabase-length genetic systems.	Design algorithms for optimal one-pot-assembly of billions of unique genomic/chromosomal variants with defined sequences.
Methods for one-step, simultaneous assembly and sequence-verification of long DNA fragments.			
Reliable assembly of 10,000 base pair non-clonal DNA fragments.	Reliable assembly and verification of 10,000 base pair clonal DNA fragments.	Reliable assembly and verification of 100,000 base pair clonal DNA fragments.	Reliable assembly and verification of 1,000,000 to 10,000,000 base pair clonal DNA fragments.
Pipelined synthesis, assembly, and functional testing of engineered genetic systems.			
Achieve desired functionalities in lower-fidelity, error-prone genetic systems.	Achieve reliable Design-for-Testing in engineered genetic systems.	Achieve readily-swappable modules within large genetic systems.	Achieve one-month Design-to-Test cycles for megabase-length genetic systems.

Precision genome editing at multiple sites simultaneously with no off-target effects.

Ability to reliably create any precise, defined edit(s) (single nucleotide polymorphisms or gene replacement) with no unintended editing in any organism, with edits ranging from a single base change to the insertion of entire pathwa.			
Ability to generate any defined single base pair change in model organisms.	High efficiency editing (> 90%) across the genome with no off-target activity.	High-efficiency gene insertion or deletion of moderately large changes (< 10 kb) via homologous recombination.	Precise, parallel editing or regulatory modifications (10 to 1000 modifications) across model and non-model organisms, including plants and animals.
Precise, predictable, and tunable control of gene expression for many genes inside diverse cells and organisms across different timescales.			
Achieve long-lasting gene repression and activation.	Ability to regulate expression in non-model organisms.	Technologies to monitor and manipulate genetic and epigenetic mechanisms controlling tissue-wide and organism-wide expression levels over time.	Ability to precisely regulate gene expression in whole-body organisms, with single-cell resolution using dynamic or static control.
Ability to reproducibly deliver editing cargo efficiently and specifically to a given target cells or tissues, and control dosage and timing of the editing machinery.			
Improve editors to function without sequence requirements with activity comparable to 2019 state-of-the-art capabilities.	Routine use of editors without detectable off-target effects.	Enhance specificity of delivery modalities for high efficiency (>90% efficient) editing of cells in a defined tissue.	Quantitative, specific, and multiplexed editing of any site, in any cell, in any organism.
2 Years	**5 Years**	**10 Years**	**20 Years**

Roadmap Elements

Goal 1: Manufacture thousands of very long oligonucleotides with high fidelity.

[Current State-of-the-Art]: Existing synthesis chemistries manufacture oligonucleotides up to 200 nucleotides long with cycle efficiencies of 99.5% and yields of 35%. Parallel synthesis of oligonucleotides is carried out on solid supports, producing up to 300,000 oligonucleotides with defined sequences (Hughes & Ellington, 2017; Kosuri & Church, 2014).

[Breakthrough Capability]: Highly efficient oligonucleotide synthesis to increase the number, length, and fidelity of oligonucleotides.

- **2 years: Robustly synthesize one million 200-mer oligonucleotides with a per-nucleotide error rate of fewer than one in 500 nucleotides.**
    - [Bottleneck]: Scaling-up the production of chip-based or semiconductor-based oligonucleotide synthesis chemistries.
        - [Potential Solution]: Microfabrication of nanotiter plates and patterned nanometer-scale chips.
        - [Potential Solution]: Improved process dynamics taking into account inherent stochasticity and improved electronic control of reaction chemistries.
- **5 years: Robustly synthesize 1,000-mer oligonucleotides with a per-nucleotide error rate of fewer than one in 1,000 nucleotides.**
    - [Bottleneck]: Current phosphoramidite-based chemistries have peaked at 99.5% per-nucleotide efficiencies, resulting in only 0.66% yields when producing 1,000-mers; efficiencies must be 99.9% to achieve more than 35% yields and cycle times must also be reduced for commercial scalability.
        - [Potential Solution]: Enzyme-based, non-templated synthesis (e.g., via terminal deoxynucleotidyl transferases) has the potential to achieve greater than 99.9% per-nucleotide efficiencies and synthesis rates exceeding one nucleotide-per-second.
- **10 years: Reduce per-nucleotide error rates for 1,000-mer oligonucleotide synthesis to fewer than one in 5,000 nucleotides.**
    - [Bottleneck]: Non-templated DNA synthesis is currently slow with lower fidelity than templated synthesis and improvements in enzyme substrate selectivity or substrate availability are needed to control sequence-specific synthesis.
        - [Potential Solution]: Significant bioprospecting, rational design, and directed evolution of enzymes responsible for non-templated DNA synthesis can improve selectivity and increase catalytic efficiencies.
- **20 years: Synthesize 10,000-mer oligonucleotides at 99.99% cycle efficiency within one minute with a per-nucleotide error rate of fewer than one in 30,000 nucleotides.**
    - [Bottleneck]: Multiple synergistic improvements are needed, including improved non-templated DNA polymerases, fast substrate switching at the nanoliter scale, multi-nucleotide same-cycle addition, and electronic control of substrate selection.

- [Potential Solution]: Inspiration from natural DNA polymerases, ligases, recombinases, and helicases, working together in a dynamic molecular machine, potentially using non-natural nucleotides for greater specificity.

Goal 2: Many-fragment DNA assembly with simultaneous, high-fidelity sequence validation.

[Current state-of-the-art]: Oligonucleotides are assembled into double-stranded DNA fragments up to 6000 base pairs long using *in vitro* techniques (e.g., polymerase cycling assembly, ligation cycling) as well as *in vivo* techniques (yeast-mediated homologous recombination), producing non-clonal DNA fragments (Gibson, 2011; Li & Elledge, 2007; Richardson et al., 2017; Smith, Hutchison, Pfannkoch, & Venter, 2003). Clonal (isogenic) fragments are then identified using a combination of enzyme-based removal of mismatched base pairs (e.g., MutS) and DNA sequencing (Sanger or NGS). Multiple verified DNA fragments are then assembled together into longer fragments (10,000 to 100,000 base pairs long) using hierarchical approaches employing DNA assembly techniques (e.g., Gibson assembly, ligation cycling reaction, Golden Gate). Megabase length DNA is then assembled from 100,000 base pair fragments using yeast-mediated homologous recombination. More detailed descriptions of commonly used techniques follow:

- **Polymerase Cycling Assembly (PCA)** is a method to assemble larger DNA constructs from shorter oligonucleotides (Smith et al., 2003). PCA is an efficient method for assembling constructs between 200 to 1,000 base pairs in length. The process is similar to PCR, but utilizes a set of overlapping “seed” oligonucleotides that are designed to hybridize to one another leaving gaps that are then filled in using a thermostable DNA polymerase. The oligonucleotides are generally 50 to 100 nucleotides in length to ensure uniqueness in the hybridization with their complement. The reactions are cycled from ~60 and ~95 C^{o} for 15 to 30 cycles. The full-length assembled product is then usually amplified by PCR using two terminal-specific primers. PCA is an efficient method for assembling constructs between 200 and 1,000 base pairs in length and can be performed in individual tubes or multiplexed using microtiter well plates.
- **Emulsion PCA** is a method developed by Sriram Kosuri for highly multiplexing the assembly of larger constructs from small amounts of shorter DNA fragments (Plesa, Sidore, Lubock, Zhang, & Kosuri, 2018). In this method, the oligos required for a given construct are designed with a unique barcode on the terminus which specifically hybridizes with a complementary barcoded attached to a bead from a complex pool of oligonucleotides. The bead mixture is then emulsified into picoliter-sized droplets containing a Type IIs restriction endonuclease (RE), dNTPs, and a thermostable DNA polymerase. The oligonucleotides are released from the bead by the Type II RE and then assembled by PCA through thermal cycling of the emulsion. Using this method, thousands of specific constructs can be assembled in a single emulsion tube depending upon the number uniquely barcoded beads.
- **Ligase Cycling Assembly (LCA)** is a method to assemble larger DNA constructs from shorter oligonucleotides or double-stranded DNA fragments (de Kok et al., 2014). LCA is an efficient method for assembling constructs between 500 and 10,000 base pairs in length. LCA assembly uses shorter, single-stranded bridging oligonucleotides that are complementary to the termini of adjacent DNA fragments that are to be joined using a thermostable ligase. Like PCA, LCA

utilizes multiple temperature cycling to denature, re-anneal, and then ligate the fragments to assemble the larger DNA construct, and can be performed in individual tubes or multiplexed using microtiter well plates.

- **Gibson Assembly** is a method to assemble larger DNA constructs from shorter oligonucleotides or double-stranded DNA fragments (Gibson et al., 2009). Gibson assembly is an efficient method for assembling constructs up to many tens of kilobase-pairs in length. This method, which is isothermal, utilizes up to 15 double-stranded DNA fragments having around 20 to 40 base pair overlaps with the adjacent DNA fragments. The DNA fragments are first incubated with 5' to 3' exonuclease, resulting in single-stranded regions on the adjacent DNA fragments that can anneal in a base pair-specific manner. The gaps are then filled in with a DNA polymerase and the final nicks closed with a DNA ligase. This method can be performed in individual tubes or multiplexed using microtiter well plates.

Importantly, the final fidelity (error rate) of the assembled constructs using the above methods are at the mercy of the quality of the input oligonucleotides. These methods usually incorporate some type of error reduction or correction methods which include removing errored duplexes (mismatches and insertions) with the MutS protein after denaturation and reannealing of the construct, or degradation of the error containing DNA using T7 or CEL endonuclease. (For review please see: 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)

[Breakthrough Capability 1]: Predictive design of DNA sequences for improved assembly of longer, more information-rich DNA fragments.

- **2 years: Coupled design of DNA sequences to optimize nucleotide composition to support synthesis, while maintaining genetic system function.**
    - [Bottleneck]: Many genetic systems contain polymeric sequences, long repeats, and non-canonical DNA structures that inhibit the assembly process.
        - [Potential Solution]: Genetic systems can be rationally designed to eliminate problematic sequence elements, while maintaining their function, thus reducing their "synthesis complexity".
        - [Potential Solution]: Toolboxes of highly non-repetitive genetic parts can be designed and characterized to enable design of non-repetitive genetic systems.
- **5 years: Incorporate machine learning to identify poorly-understood problematic sequences and process conditions.**
    - [Bottleneck]: The complete list of sequence elements that inhibit DNA assembly is not fully known and the process conditions leading to undesired byproducts are not well understood.
        - [Potential Solution]: Machine learning algorithms have the ability to identify problematic DNA sequences and undesired process conditions that lead to inefficient DNA assembly.

- **10 years: Design algorithms that identify optimal synthesis strategies for assembling megabase-length genetic systems.**
    - [Bottleneck]: The functions of some genetic system components are more strictly reliant on problematic sequences and trade-offs between design-for-function versus design-for-synthesis are likely.
        - [Potential Solution]: Design algorithms can identify regions with problematic sequences and identify optimal strategies for mixing and matching megabase-length assembly strategies accounting for these regions.
- **20 years: Design algorithms for optimal one-pot-assembly of billions of unique genomic/chromosomal variants with defined sequences.**
    - [Bottleneck]: Mixtures of oligonucleotides can be used to construct combinatorial libraries of DNA fragments, though assembling those fragment libraries into diversified mega-base genetic systems has not been achieved.
        - [Potential Solution]: Parallel evaluation of sequence design criteria across billions (or trillions) of potential sequence variants can be carried out; as the diversification of libraries increase, the number of sequence variants increases combinatorially.

[Breakthrough Capability 2]: Methods for one-step, simultaneous assembly and sequence-verification of long DNA fragments.

- **2 years: Reliable assembly of 10,000 base pair non-clonal DNA fragments.**
    - [Bottleneck]: The availability of high-fidelity long oligonucleotides, the optimization of process conditions, and the presence of problematic sequences.
        - [Potential Solution]: Higher fidelity 100-mer and 200-mer oligonucleotides.
        - [Potential Solution]: Identification of optimal process conditions and removal (by design) of problematic sequence elements.
- **5 years: Reliable assembly and verification of 10,000 base pair clonal DNA fragments.**
    - [Bottleneck]: Low assembly yields and decoupled sequencing leads to more costly hierarchical processes with higher failure rates.
        - [Potential Solution]: Enzyme-based selection (e.g., via MutS) can eliminate DNA fragments containing errors**.**
        - [Potential Solution]: Approaches using simultaneous DNA synthesis and sequencing can rapidly sort DNA fragments, excluding fragments with errors (for example, using nanopore-based sequencing and dynamic pore flicking).
- **10 years: Reliable assembly and verification of 100,000 base pair clonal DNA fragments.**
    - [Bottleneck]: Reliable, low-cost assembly of clonal 10,000 base pair fragments.
        - [Potential Solution]: Higher efficiency, ten-part assemblies using lower-cost, clonal 10,000 base pair DNA fragments.
        - [Potential Solution]: Extra long read sequencing for verification of 100,000 base pair fragments (e.g., nanopore sequencing).

- **20 years: Reliable assembly and verification of 1,000,000 to 10,000,000 base pair clonal DNA fragments.**
    - [Bottleneck]: Reliable, low-cost assembly of clonal 100,000 base pair fragments.
        - [Potential Solution]: *In vivo*, yeast-mediated assembly of clonal 100,000 base pair fragments into megabase-length genetic systems.
        - [Potential Solution]: Extra long read sequencing for verification of 1,000,000 base pair fragments (e.g., nanopore sequencing).

[Breakthrough Capability 3]: Pipelined synthesis, assembly, and functional testing of engineered genetic systems.

- **2 years: Achieve desired functionalities in lower-fidelity, error-prone genetic systems.**
    - [Bottleneck]: Unpredictable relationship between synthesis and assembly errors versus undesired functional outcomes.
        - [Potential Solution]: Elimination of problematic sequences via rational design and incorporation of robust, mutation-invariant design into genetic systems.
        - [Potential Solution]: Routine application of low-cost -omics technologies to verify the functions of genetic systems (e.g., DNA-Seq, RNA-Seq, Ribo-Seq, and metabolomics).
- **5 years: Achieve reliable Design-for-Testing in engineered genetic systems.**
    - [Bottleneck]: Costly to assay diverse genetic functions to verify desired behaviors.
        - [Potential Solution]: Synthesized and assembled genetic systems can directly incorporate a suite of sensors and genetic circuits for self-testing of genetic system function; sensor-circuit outputs could be tailored for desired high-throughput assays, including surface display, Flow-Seq, and RNA-Seq.
- **10 years: Achieve readily-swappable modules within large genetic systems.**
    - [Bottleneck]: Synthesis of megabase-length genetic systems may contain commonly used and re-used genetic modules.
        - [Potential Solution]: Previously synthesized and assembled genetic modules (of more than 100,000 base pair fragments) can be re-used in downstream processes and models can be developed to predict inter-module interactions and overall system function.
- **20 years: Achieve one-month Design-to-Test cycles for megabase-length genetic systems.**
    - [Bottleneck]: Design algorithms, synthesis chemistries, assembly techniques, simultaneous sequencing, and functional testing must be seamlessly integrated within a commercially viable suite of services with fast turnaround times.
        - [Potential Solution]: A combination of cooperative horizontal service providers and well-integrated vertical service providers operating within a healthy commercial ecosystem.

Goal 3: Precision genome editing at multiple sites simultaneously with no off-target effects.

[Current State-of-the-Art]: A variety of current tools can be used for DNA sequence edits and for non-editing-based genome engineering including gene regulation and chromatin engineering. Transcription activator-like effector nucleases (TALEN)-based or clustered regularly interspaced short palindromic repeats (CRISPR)-based genome engineering techniques introduce site-specific nicks or double-stranded breaks, which are then repaired using natural repair pathway (Doudna & Charpentier, 2014). Additional state-of-the-art editing technologies include adeno-associated virus (AAV)-mediated homologous recombination and meganuclease activity. With CRISPR and TALEN technologies, up to six distinct sites, and up to 15,000 identical sites, have been targeted simultaneously, with efficiencies ranging from 2% to 90%. Gene regulation is achieved through site-specific DNA-binding proteins (zinc-finger proteins, transcription activator-like effectors, and Cas proteins), which fuse to gene regulatory domains to carry out activation or repression of desired genes. In these cases, up to six distinct genes have been targeted for regulation, with repression magnitudes up to 300-fold (knock-down) and activation magnitudes up to 20-fold (knock-up) (L. A. Gilbert et al., 2013; Qi et al., 2013).

[Breakthrough Capability 1]: Ability to reliably create any precise, defined edit(s) (single nucleotide polymorphisms or gene replacement) with no unintended editing in any organism, with edits ranging from a single base change to the insertion of entire pathways.

- **2 years: Ability to generate any defined single base pair change in model organisms.**
    - [Bottleneck]: Performance of current editing technology and known-unknowns (e.g., chromatin, double-stranded break repair) and unknown-unknowns regarding basic biology.
        - [Potential Solution]: Improved base editing enzymes capable of catalyzing all possible nucleotide transitions.
        - [Potential Solution]: Engineered nucleases and recombinases to control repair of double-stranded breaks using either non-homologous end-joining or homologous recombination.
- **5 years: High-efficiency editing (beyond 90%) across the genome with no off-target activity.**
    - [Bottleneck]: A better understanding of canonical protospacer adjacent motif (PAM) specificities, non-canonical R-loop formation, types of double-stranded breaks, the effects of DNA supercoiling, and the double-strand DNA repair pathway mechanisms.
        - [Potential Solution]: A quantitative, predictive understanding of the coupled chromatin, editing, and repair interactions.
        - [Potential Solution]: A suite of genome editors covering all possible nucleotide (PAM) specificities.
        - [Potential Solution]: Genome editors with improved on-target and reduced off-target effects.

        - [Potential Solution]: Design algorithms for predicting single guide RNA (sgRNA) guide RNA sequences, sgRNA concentrations, and genome editor concentrations to achieve desired on-target activities with minimal off-target activities.
    - [Bottleneck]: Need for a better understanding of chromatin effects on editing in higher order systems.
        - [Potential Solution]: Fusion of epigenetic effectors to Cas9.
- **10 years: High-efficiency gene insertion or deletion of moderately large changes (but less than 10 kilobases) via homologous recombination.**
    - [Bottleneck]: The ability to manipulate double-stranded DNA break repair at high efficiency in non-model cells.
        - [Potential Solution]: Single-effector base editors that catalyze insertions and deletions with high efficiency.
        - [Potential Solution]: Improving the targeted delivery of DNA repair templates to nuclei in model and non-model organisms.
        - [Potential Solution]: Inducible control of DNA break repair pathways with significantly higher efficiency.
        - [Potential Solution]: Retron-mediated synthesis of DNA repair templates from co-delivered crRNAs, sgRNAs or RNPs.
    - [Bottleneck]: Ability to efficiently deliver large DNA constructs into cells.
        - [Potential Solution]: Expand viral packaging size.
        - [Potential Solution]: Develop and/or enhance non-viral DNA delivery tools and technologies.
- **20 years: Precise parallel editing or regulatory modifications (10 to 1000 modifications) across model and non-model organisms, including plants and animals.**
    - [Bottleneck]: Co-expressing many crRNAs or sgRNAs within arrays can trigger genetic instability, due the presence of many repetitive sequences.
        - [Potential Solution]: Toolboxes of highly non-repetitive CRISPR genetic parts can be designed and characterized, enabling the design of many-sgRNA arrays that do not trigger genetic instability.

[Breakthrough Capability 2]: Precise, predictable, and tunable control of gene expression for many genes inside diverse cells and organisms across different timescales.

- **2 years: Achieve long-lasting gene repression and activation.**
    - [Bottleneck]: Insufficient quantitative understanding of transcriptional regulation, epigenetic mechanisms, and cross-regulatory interactions.
        - [Potential Solution]: Quantitative characterization of additional genome editors with diverse protein fusions for more potent CRISPRa and CRISPRi.
        - [Potential Solution]: Systematic characterization of gene regulatory effects when changing RNP binding site locations, affinities, and RNP-PIC interactions.

        - [Potential Solution]: Engineered pioneer transcription factors that can reliably generate desired epigenetic states.
- **5 years: Ability to regulate expression in non-model organisms.**
    - [Bottleneck]: A quantitative understanding of promoter-specific, epigenetic-specific, and tissue-specific gene regulatory interactions.
        - [Potential Solution]: Organism- and tissue-specific promoters to express CRISPR components in desired non-model organisms.
        - [Potential Solution]: Organism- and tissue-specific protein fusion domains for potent activation or repression of genes in non-model organisms.
        - [Potential Solution]: Improved delivery and/or expression of CRISPR components into non-model organisms.
- **10 years: Technologies to monitor and manipulate genetic and epigenetic mechanisms controlling tissue-wide and organism-wide expression levels over time.**
    - [Bottleneck]: Insufficient methodologies for measuring and quantifying epigenetic, mRNA level, protein level, and metabolite level changes at genome-wide, single-cell resolution with sufficient precision to be relevant to clinical phenotypes.
        - [Potential Solution]: Single-cell epigenetic modification (histone modification, lncRNA, nucleosome position), protein level, and metabolite level determination.
        - [Potential Solution]: Single-cell, primary cell line measurements quantifying epigenetic modifications, protein levels, and metabolite levels, for example, using clinical tissues.
- **20 years: Ability to precisely regulate gene expression in whole-body organisms, including humans, with single-cell resolution using dynamic or static control.** This capability excludes germline genome editing.
    - [Bottleneck]: Improved delivery and expression of large genetic constructs in primary cell lines.
        - [Potential Solution]: Coupled transfection, genome editing, and genome repair to insert large genetic constructs into site-specific regions within chromosomes.
    - [Bottleneck]: Poorly understood tissue-specific cell states.
        - [Potential Solution]: Predictive models quantifying relationships between genetic, epigenetic states, and cellular phenotypes.
        - [Potential Solution]: Organ-, tissue-, and site-specific chromatin effectors with predictable/controllable effects on chromatin structure, gene expression, and gene accessibility.

[Breakthrough Capability 3]: Ability to reproducibly deliver editing cargo efficiently and specifically to a given target cells or tissues, and control dosage and timing of the editing machinery.

- **2 years: Improve editors to function without sequence requirements (such as protospacer adjacent motif (PAM) sequences) with activity comparable to 2019 state-of-the-art capabilities.**
    - [Bottleneck]: Current editors are large and prone to off-target activity.
        - [Potential Solution]: "Version 2.0" of high-fidelity editors.
        - [Potential Solution]: Smaller editors suitable for enhanced delivery.
        - [Potential Solution]: Regulated editors that are active in a context-dependent fashion (cell type, small molecule regulation, etc.).
- **5 years: Routine use of editors without detectable off-target effects (less than 0.001% off-target editing).**
    - [Bottleneck]: Limit of detection for current sequencing technologies.
        - [Potential Solution]: Technology improvement to lower detection limits and achieve more targeted detection.
        - [Potential Solution]: Reliable assays to discover potential off-target sites.
- **10 years: Enhance specificity of delivery modalities for high efficiency (>90% efficient) editing of cells in a defined tissue.**
    - [Bottleneck]: Current delivery modalities have low cell-type specificity.
        - [Potential Solution]: Large scale development of virus engineering for tropism/specificity.
        - [Potential Solution]: RNPs engineered with cell-type specificity via receptor interactions, and other modalities.
    - [Bottleneck]: Editing *in vivo* often leads to low efficiency of edits.
        - [Potential Solution]: Enhanced viral delivery with long half-life and low immunogenicity (likely needed for every organism of interest, including plant viruses and animal viruses).
        - [Potential Solution]: Viruses with large capacity (up-to or greater than 10 kilobase capacity) needed to deliver editor, guide RNAs, and any donor DNA molecule.
        - [Potential Solution]: Engineered effector complexes (such as RNPs) that can be delivered directly *in vivo* and maintain activity.
- **20 years: Quantitative, specific, and multiplexed editing of any site, in any cell, in any organism.**
    - [Bottleneck]: Specificity, efficiency, genetic stability, and off-target effects all pose challenges.
        - [Potential Solution]: Continued improvement of delivery vectors for plants, non-model animals, and other organisms.
        - [Potential Solution]: Development of toolboxes of non-repetitive CRISPR components to enable highly-multiplexed editing without triggering genetic instability.

- [Potential Solution]: Tools for editing in humans and animals for therapeutics, specifically those that overcome or mask the immune response.

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