EBRC Engineering Biology Research Consortium

An Assessment of Short-Term Milestones in EBRC's 2019 Roadmap, *Engineering Biology*

Engineering Biology Research Consortium

March 2023

www.ebrc.org

5885 Hollis Street, 4th Floor, Emeryville, CA 94608
Phone: +1.510.871.3272 | Fax: +1.510.245.2223
https://ebrc.org | https://roadmap/ebrc.org

This material is based upon work supported by the National Science Foundation under Award Nos. 1818248 and 2116166. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Executive Summary

The Engineering Biology Research Consortium's 2019 publication of the technical roadmap, *Engineering Biology: A Research Roadmap for the Next-Generation Bioeconomy*, marked a seminal moment for the field of engineering biology. The roadmap extensively cataloged the potential for progress in the field, setting out numerous goals, possible breakthroughs, and ambitious milestones for the following 20 years. As we approached and passed the first milestone timepoint at 2-years post publication, EBRC sought to review progress in the field, compared against the advancements anticipated by the roadmap. This resulting Assessment reports on technical achievements and advancements in addition to barriers, both transient (e.g., impacts of the COVID-19 pandemic) and persistent (e.g., diversity and inclusion in engineering biology), to progress. This information will enable the research community to reflect on its achievements, enable industry to better anticipate nascent and emerging technologies, and support policymakers and funders in identifying priority areas for additional investment and infrastructure to ensure continued advancement.

The Assessment was completed through a series of surveys, discussions with experts and stakeholders, and an extensive literature review, which took place in 2021 and early 2022. The nature of the roadmap and process of appraising myriad published works spanning the field resulted in a primarily qualitative Assessment report. Importantly, the Assessment reflects only a snapshot in time and the knowledge and expertise of its contributors; engineering biology research and biotechnology development advance continuously and at a great pace and thus cannot be exhaustively captured here.

The Assessment reports significant progress in the field and suggests the roadmap has, so far, been a useful predictor of the direction of engineering biology research. Notable technical advancements were achieved in DNA assembly and in host engineering, such as developments in genome engineering in model and non-model organisms. When assessing some bottlenecks and 2-year milestones in the roadmap, certain innovations and approaches were shown to have circumvented those predictions, while still contributing toward later milestones. For example, there was noted to be significant advancement in enzymatic DNA synthesis, thus making further progress in phosphoramidite chemistry synthesis less necessary. Conversely, data integration and other data science capabilities remain a major bottleneck and may be contributing to slower (though still mostly on track) progress in biomolecular engineering. Projects like AlphaFold 2 and advancements in machine learning protocols have the potential to help overcome slowdowns in biomolecular engineering, as the technology and data become more readily available.

The Assessment also examined social considerations and nontechnical dimensions impacting the advancement in engineering biology anticipated by the roadmap. The COVID-19 pandemic, particularly restrictions to in-person activities and supply chain disruptions, had a significant impact on the conduct of research during this period, and the lasting impacts of the pandemic have yet to be seen. Other barriers include regulatory uncertainty and challenges in education, particularly a paucity of comprehensive data science education for trainees. Other dimensions that will have an impact on research advancement going forward include security practices and norms, risk assessment for emerging engineering biology technologies, the capacity to collaborate with the social sciences, and demographic diversity in academia and the research pipeline. Many of these dimensions point to expanding the stakeholder base to enable a robust research enterprise and bioeconomy.

Overall, the Assessment points to steady progress as anticipated by *Engineering Biology* and reinforces the utility of the roadmap as a resource for researchers, policymakers, and industry leaders. Most early milestones have been reasonable predictors of the direction of research, and many, if not all, of the high-level goals and breakthroughs remain viable reference points. Based on the findings of the Assessment, the field of engineering biology is poised to continue its consequential growth and advancement in the coming years, and we look forward to revisiting the roadmap to assess progress in the future.

Assessment of *Engineering Biology*

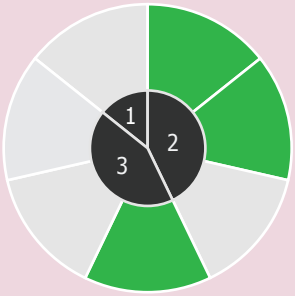
This 2021 Assessment reports on progress in the field relative to EBRC's 2019 technical roadmap, *Engineering Biology: A Research Roadmap for the Next-Generation Bioeconomy* (available at https://roadmap.ebrc.org).

KEY: Central number corresponds to the respective Goal as listed.

- Progress towards this Breakthrough Capability is ahead of roadmap predictions
- Progress has met, or is expected to meet prediction
- Progress is behind

Engineering DNA

1. Manufacture thousands of very long oligonucleotides with high fidelity
2. Many-fragment DNA assembly with simultaneous, high-fidelity sequence validation
3. Precision genome editing at multiple sites simultaneously with no off-target effects



Enzymatic DNA synthesis; Advancements in enzymatic DNA synthesis are enabling the engineering of longer oligomers, improving the efficiency of engineering more organisms.

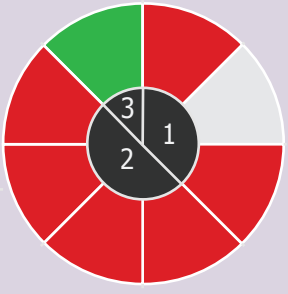
Biomolecular Engineering

1. On-demand design, generation, and evolution of macromolecules for desired functions
2. Special considerations for on-demand design, generation, and evolution of macromolecules that rely on non-canonical/unnatural building blocks
3. Holistic, integrated design of multi-part genetic systems (i.e., circuits and pathways)
4. Integrated design of RNA-based regulatory systems for cellular control and information processing

Protein structure modeling and prediction; Advancements in software and platforms, like AlphaFold 2, are enabling more efficient and accurate prediction and modeling of protein structure.

Data Science

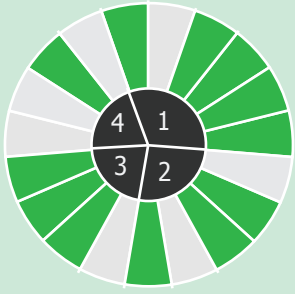
1. Establish a computational infrastructure where easy access to data supports the DBTL process for biology
2. Establish functional prediction through biological engineering design at the biomolecular, cellular, and consortium scale
3. Establish optimal manufacturing processes from the unit-operation to the integrated-screening scale



Computational resources and shared data; Deficiencies in shared and accessible data and a paucity of computational resources are slowing some advancements in engineering biology.

Host Engineering

1. Cell-free systems capable of natural and/or non-natural reactions
2. On-demand production of single-cell hosts capable of natural and non-natural biochemistry
3. On-demand fabrication and modification of multicellular organisms
4. Generation of biomes and consortia with desired functions and ecologies



Genome engineering; New platforms and tools, such as integrases and CRAGE, are enabling greater engineering of host genomes, including in non-model organisms.

Assessment Contributors

This Assessment was made possible by the valuable contribution of many researchers by providing insight and input on milestone progress, technical and nontechnical barriers, unanticipated advancements, and social dimensions and considerations.

Project Leadership

Albert Hinman, Ph.D.
EBRC Postdoctoral Fellow

Emily R. Aurand, Ph.D.
EBRC Director of Roadmapping and Education

Michael Köpke, Ph.D.
Vice President Synthetic Biology, LanzaTech | EBRC Roadmapping Working Group Chair

Technical Theme Section Chairs

Engineering DNA | Gene Editing, Synthesis, and Assembly	**Howard Salis, Ph.D.** Pennsylvania State University	**Rebecca Nugent, Ph.D.** Synthego
Biomolecular Engineering | Biomolecule, Pathway, and Circuit Engineering	**Chang Liu, Ph.D.** University of California, Irvine	**Jesse Zalatan, Ph.D.** University of Washington
Host Engineering | Host and Consortia Engineering	**James Carothers, Ph.D.** University of Washington	**Ute Galm, Ph.D.** Curie Co
Data Science | Data Integration, Modeling, and Automation	**Adam Arkin, Ph.D.** University of California, Berkeley	**Nathan Hillson, Ph.D**. Lawrence Berkeley National Laboratory

Contributors

Adam Silverman, Ph.D., Sherlock Biosciences

Alexander Green, Ph.D., Boston University

Allison Berke, Ph.D., University of California, San Francisco

Andrew Hunt, Northwestern University | SPA Hackathon Facilitator

Ania Baetica, Ph.D., University of California, San Francisco | SPA Hackathon Facilitator

Anna M. Crumbley, Ph.D., U.S. Army DEVCOM Chemical Biological Center

Austin Cole, Ph.D., University of Texas, Austin

Brian Pfleger, Ph.D., University of Wisconsin-Madison

Bridget Luckie, University of California, Berkeley

Bryn Adams, Ph.D., Department of the Army

Carolyn Mills, Ph.D., Northwestern University

Carrie Cizauskas, Ph.D.

Cholpisit Kiattisewee, University of Washington

Danielle Tullman-Ercek, Ph.D., Northwestern University

David Pattie, Ph.D., The MITRE Corporation

David Ross, Ph.D., National Institute of Standards and Technology

Dylan M. Brown, Northwestern University

Elizabeth Strychalski, Ph.D., National Institute of Standards and Technology

Emiley Eloe-Fadrosh, Ph.D., Joint Genome Institute/Lawrence Berkeley National Laboratory

Emily Fulk, Rice University

Ethan Hillman, Ph.D., University of Michigan

Georg K. A. Wachter, Imperial College London

Heidi Klumpe, Ph.D., Boston University

James Diggans, Ph.D., Twist Bioscience

Janet Matsen, Ph.D., Zymergen

Jayan Rammohan, Ph.D., National Institute of Standards and Technology

Jeffrey J. Tabor, Ph.D., Rice University

Joseph Harman, Ph.D., University of Washington

Julie Ming Liang, Northwestern University

Kersh Thevasundaram, Ph.D., UC Berkeley

Kevin Solomon, Ph.D., University of Delaware

Kok Zhi Lee, Ph.D., University of Washington in St. Louis

Lin Su, Ph.D., University of Cambridge

Marilene Pavan, Ph.D., Lanzatech Inc.

Mark Blenner, Ph.D., University of Delaware

Mark Styczynski, Ph.D., Georgia Institute of Technology

Matthew Amrofell, Washington University in St. Louis

Michael Andrew Crone, Ph.D., Imperial College London

Michael Haro, Lockheed Martin

Michael Sheets, Boston University

Myint Toe, Katholieke Universiteit Leuven

Nicole Buan, Ph.D., University of Nebraska, Lincoln

Nigel Mouncey, Ph.D., Lawrence Berkeley National Laboratory

Nolan Kennedy, Northwestern University

Paul Freemont, Ph.D., Imperial College London

Richard Murray, Ph.D., Caltech

Ritu Shrestha, Ph.D., Pacific Northwest National Lab

Roel Bovenberg, Ph.D., DSM

Sean Leonard, Ph.D., University of Texas at Austin

Sean Sullivan, Tufts University

Sebastian Oehm, University of Cambridge

Svetlana Ikonomova, Ph.D., Northwestern University

Tae Seok Moon, Ph.D., Washington University in St. Louis

Tara Deans, Ph.D., University of Utah

Taylor Nichols, Northwestern University

Vincent Noireaux, Ph.D., University of Minnesota

William Grubbe, University of Chicago

Yan Zhang, Ph.D., California Institute of Technology

Yue Han, Georgia Institute of Technology

EBRC Leadership and Staff – all our activities are a team effort, and we are grateful to the EBRC Team for their assistance and support with this project.

India Hook-Barnard, Ph.D.
Executive Director

Becky Mackelprang, Ph.D.
Associate Director for Security Programs

Kaitlyn Duvall
Research Assistant

Sifang Chen, Ph.D.
Postdoctoral Fellow

Angela Lee
Administrator

Doug Friedman, Ph.D.
President

Introduction

Established in 2016, the Engineering Biology Research Consortium (EBRC) seeks to advance engineering biology to address national and global needs. In support of this mission, EBRC develops technical research roadmaps to showcase cutting-edge research and identify challenges and opportunities to applying engineering biology-enabled technologies. The importance of engineering biology research roadmapping was highlighted by the 2015 National Academies of Science, Engineering, and Medicine report, *Industrialization of Biology* (National Academies of Sciences, Engineering, and Medicine, 2015) and EBRC undertakes this work through ongoing support from the National Science Foundation and other stakeholders. By imagining the future of engineering biology through the process and publication of technical roadmaps, EBRC provides a strategic resource for a wide variety of stakeholders: for academic researchers, roadmaps can provide project ideas and areas for cross-disciplinary collaboration; for industry, roadmaps can motivate new products and avenues of innovation; for government and policymakers and those that support research advancement, roadmaps can highlight potential areas for new programs and investment and draw attention to policy and regulatory needs; and for students and trainees, roadmaps can generate career opportunities and inspire applications for the concepts they are learning and the tools they are developing.

Engineering Biology: A Research Roadmap for the Next-Generation Bioeconomy

In June 2019, EBRC published our inaugural research roadmap – *Engineering Biology: A Research Roadmap for the Next-Generation Bioeconomy* (Engineering Biology Research Consortium, 2019). This roadmap represents the culmination of more than three years of discussion and more than 15 months of scoping, drafting, review and revision to assess the status and potential of engineering biology. The roadmap recognizes the potential to leverage biology as a technology and attempts to capture the overwhelming complexity of natural biological properties that could be engineered and scaled to produce biobased products and solutions for commercial, national, and/or societal objectives. The roadmap categorizes engineering biology research into four, often overlapping, **technical themes**: Gene Editing, Synthesis, and Assembly (also referred to by the theme's short title, "Engineering DNA"); Biomolecule, Pathway, and Circuit Engineering ("Biomolecular Engineering"); Host and Consortia Engineering ("Host Engineering"); and Data Integration, Modeling, and Automation ("Data Science"). The roadmap also envisions **applications and impacts** of engineering biology across five **sectors**: Industrial Biotechnology; Health & Medicine; Food & Agriculture; Environmental Biotechnology; and Energy. Through the design-build-test-learn (DBTL) process, the basic research capabilities and tools developed in the technical themes forms the foundation for the application of those capabilities in the five sectors (Figure 1).

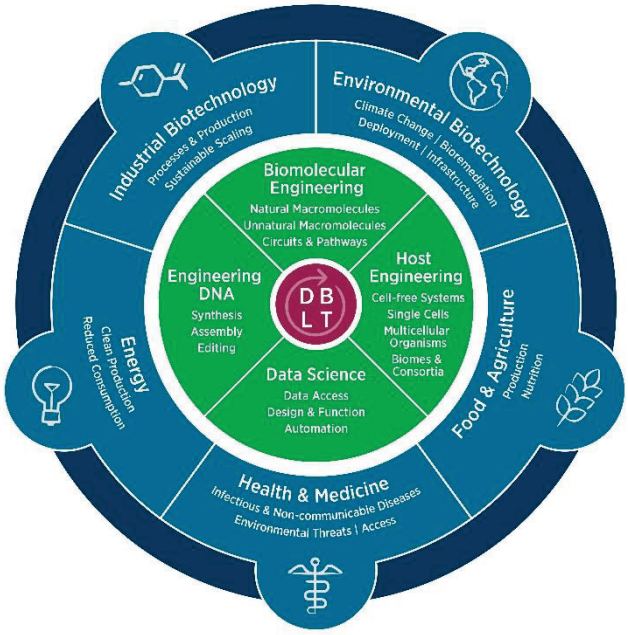


Figure 1. Technical Themes and Application Sectors from *Engineering Biology, A Research Roadmap for the Next-Generation Bioeconomy*. The graphic depicts the four technical themes of the roadmap and the key technologies encompassed in each theme in green. The outer blue layer depicts the roadmap's five application sectors where advancements from the four technical themes can impact major societal challenges. The Design, Build, Test, Learn (DBTL) cycle is central and critically important for research and scaling engineering biology. Research and innovation represented in each of these layers intertwines and forms the basis for a sustainable, resilient, next-generation bioeconomy.

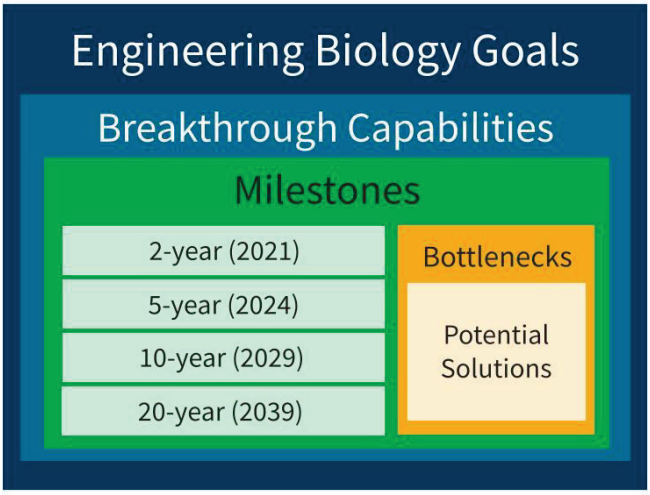


Figure 2. Hierarchy of Roadmap Elements. Each technical theme of *Engineering Biology* is organized into a hierarchy of elements, starting with three-to-four big picture goals comprised of substantial breakthrough capabilities. The breakthrough capabilities each have a number of ambitious milestones at 2-, 5-, 10-, and 20-years post-publication, and each milestone is detailed with specific technical bottlenecks and representative potential solutions.

Engineering Biology is organized hierarchically and matrixed, such that each application and impact sector addresses how tools and technologies from each technical theme can help to address a major societal challenge. The four technical themes employ a bottom-up approach, describing tool and technology innovations needed to meet specific **2-, 5-, 10-, and 20-year milestones**. Collectively, milestones pave the way to achieving a **breakthrough capability**, which represents a significant achievement in engineering biology. Synchronously, these breakthrough capabilities come together under a high-level **goal**, which concisely states a major capacity for the field (Figure 2). Conversely, the application and impact sectors provide a top-down view of how significant societal challenges might be overcome, in part, with solutions from engineering biology.

Assessing Engineering Biology

In 2019, *Engineering Biology* marked a 'first of its kind' technical roadmap for the discipline; existing strategies and roadmaps at the time focused primarily on policy opportunities or suggested end-point technologies, without laying out the technical capabilities necessary to achieve them. Because *Engineering Biology* provides stepwise milestone paths toward achieving technical objectives, technical progress toward these objectives can be tracked and monitored over time. Two years after the release of *Engineering Biology*, EBRC set out to assess progress towards the roadmap's first, 2-year milestones. Notably, the 2-year milestones represent what was largely anticipated to occur; specifically, expected outcomes of research that was ongoing or proposed when the roadmap was being drafted. Later milestones (at 5-, 10- and 20-years) were designed to be more ambitious and speculative, requiring investment, resources, and/or foundational technologies not yet realized. This Assessment presents the findings from tracking engineering biology from 2019 to 2021, providing an evaluation of the field's progress, an acknowledgement of barriers to progress, and commentary on new, noteworthy, or unanticipated research directions and achievements. This Assessment breathes new life into *Engineering Biology*, celebrating the progress that has been made toward the milestones and highlighting new directions and persistent challenges and needs.

This Assessment of *Engineering Biology* was driven by community engagement in evaluating the status of research under the four technical themes and through a deep examination of published literature. EBRC recruited a variety of contributors to conduct the Assessment including: individuals who contributed to the development of *Engineering Biology* in 2018-2019, current academic and industry members of EBRC who may not have contributed previously, members of the EBRC Student and Postdoc Association (SPA), and other stakeholders from the community. The Assessment leadership created surveys and facilitated discussions with contributors, asking whether the 2-year milestones – and preemptively, the 5-year milestones – had been achieved, whether research towards them was in progress, or if research was behind schedule. Contributors were also asked to identify technical and nontechnical barriers (in this context, "nontechnical" was defined as factors beyond the active practice of engineering biology) that were known to or could be impacting research progress. EBRC also hosted a "hackathon" for SPA members, inviting participation from trainees at the forefront of engineering biology research. In addition to the surveys and direct input from contributors, an extensive primary literature search was conducted. This literature review considered publications, products,

patents, and other applications of research advancements as evidence of technical milestone completion. More detail about the process of creating the Assessment can be found in Appendix I.

The Assessment consists of two parts. **Part 1 - Technical Progress in Engineering Biology** addresses the four technical themes of *Engineering Biology*, noting progress toward each goal and breakthrough capability and providing literature evidence to support the appraisal of the 2-year (2021) milestones; the Assessment of each theme also includes highlights in technology achievements and barriers to progress. **Part 2 - Social and Nontechnical Dimensions to Advance Engineering Biology** highlights considerations in policy, security, education, and engagement that are influencing, or play a role in, research advancement as anticipated by the roadmap. While the Assessment, particularly the topics and considerations described in Part 2, may reflect trends across the discipline of engineering biology, we have limited the scope to assessing the themes and content of *Engineering Biology*.

Assessment Results: Progress and Barriers

The Assessment revealed that progress in the field was not uniform across the four technical themes – in some cases, progress has fallen behind the timeline set by the roadmap milestones; however, many technical achievements in engineering biology have been in-line with what was anticipated by the roadmap, following the predicted research and technology development paths. Research falling under the theme of Gene Editing, Synthesis, and Assembly exhibited progress largely consistent with what was anticipated, with many of the 2-year milestones being met. Likewise, Host and Consortia Engineering displayed consistent progress, and in some cases, research was ahead-of-schedule with respect to the coming 5-year milestones. Progress under the theme Biomolecule, Pathway, and Circuit Engineering was less consistent, with notably slower technical progress in developing non-canonical or unnatural building blocks for biomolecular engineering. Importantly, research and capabilities under the theme Data Integration, Modeling, and Automation have fallen behind, with major efforts needed to develop accessible and shared computational infrastructure and functional prediction programs. The lack of certain infrastructure and tools, minimal widely-accepted standards and metrics, and paucity of accessible data for research across the discipline are creating bottlenecks to advancement.

As the community considers future progress and how to facilitate a robust engineering biology research ecosystem and industrial bioeconomy, there are also several nontechnical dimensions to consider. The COVID-19 pandemic greatly impacted research in myriad ways and was a significant and wholly unanticipated disruption to many projects and activities. Other dimensions and considerations have persistently impacted research progress, or are likely to in the future, such as the regulation of biotechnology, the extent to which security and the social sciences are integrated into technical research, improvements to multidisciplinary education for the next generation, and the development of a more inclusive and diverse engineering biology community.

Like the roadmap, this Assessment reflects a snapshot in time and the knowledge and expertise of its contributors. Engineering of biology advances every day and thus, while extensive, this Assessment should not be considered comprehensive or conclusive. Since the point of data collection in 2021 and early 2022, we expect there has been further progress in many areas. Some milestones assessed to be “behind schedule” may have since been reached while others may still be encountering barriers to achievement. Furthermore, we have taken the publication of results related to a milestone as evidence towards its achievement, but this demonstration of accomplishment by one laboratory or group cannot be taken as evidence of a universal capacity, widespread adoption, or broad dissemination of a tool or capability across the field. However, overall, this Assessment reflects a state of progress largely in-line with what was anticipated by *Engineering Biology* and can be useful in guiding areas of investment and attention where progress is generally falling behind and towards the milestones yet to come.

Part 1: Technical Progress in Engineering Biology

Over the last three years, we have seen significant advancements in the field of engineering biology, including the synthesis of ever-longer strands of DNA, in our ability to design and predict protein structure, and in the functional capacity of engineered cells. We have also seen some areas of great potential that are still intractable to progress. It is valuable for the research community, investors, policymakers, and other stakeholders to be aware of new technologies that have arisen and barriers preventing further progress. Measuring this technological progress ensures that engineering biology can be more efficiently developed to achieve national and global objectives necessary for a strong, resilient research enterprise. The Assessment of *Engineering Biology* examines the technical progress made since the roadmap's publication in 2019 by assessing the degree of completion of the technical milestones, with a focus on the 2-year milestones that had been anticipated to be reached by 2021. In addition, the Assessment highlights unanticipated research advances in each technical theme and calls out specific barriers to progress.

Each technical theme had different degrees of progress towards the breakthrough capabilities in the time from 2019 to 2021. Engineering DNA exhibited largely consistent progress, apart from stalled developments in delivering genome-editing cargo efficiently in specific cell and tissue types. There was inconsistent progress among Biomolecular Engineering breakthrough capabilities, with notably slower advancements in developing non-canonical or unnatural building blocks. Host Engineering consistently progressed and, in some cases, is ahead of schedule, with only further advancements needed in the ability to grow any host, anytime, in a controlled and regulated setting. Among Data Science 2-year milestones, all remain unfulfilled, resulting in the technical theme with the least progress towards the roadmap predictions since 2019 and with major efforts still needed towards developing accessible computational infrastructure and functional prediction programs.

In Part 1 of this Assessment, advancements and persistent challenges in each technical theme are described, as well as detailed analysis and reporting of the efforts made to achieve each 2-year, 2021 milestones.

Engineering DNA | Gene Editing, Synthesis, and Assembly

The engineering of DNA is a foundational technology that accelerates work in all application areas and underpins advancements in biomolecular and host engineering. As described in *Engineering Biology*, research in Gene Editing, Synthesis, and Assembly ("Engineering DNA") focuses on "the development and advancement of tools to enable the production of chromosomal DNA and the engineering of entire genomes." *Engineering Biology* predicted that the market for synthesized DNA was ripe for disruption; this observation remains true with advancements in several platform technologies since 2019, including enzymatic DNA synthesis and Chassis-independent Recombinase-Assisted Genome Engineering (CRAGE) (Wang et al., 2019). Beyond the original predictions of the roadmap, several technologies that are foundational to engineering genomes have developed, including improvements in long-read nanopore sequencing of DNA (Amarasinghe et al., 2020) and new methods to engineer the DNA of microbes (see e.g., Alam et al., 2021; McCarty & Ledesma-Amaro, 2019; Wang et al., 2021a for review).

Progress in Gene Editing, Synthesis, and Assembly
Goal: Manufacture thousands of very long oligonucleotides with high fidelity.
Breakthrough Capability: Highly efficient oligonucleotide synthesis to increase the number, length, and fidelity of oligonucleotides.
●●●● 2021 Milestone: Robustly synthesize one million 200-mer oligonucleotides with a per-nucleotide error rate of fewer than one in 500 nucleotides.
Goal: Many-fragment DNA assembly with simultaneous, high-fidelity sequence validation.
Breakthrough Capability: Predictive design of DNA sequences for improved assembly of longer, more information-rich DNA fragments.
●●●○ 2021 Milestone: Coupled design of DNA sequences to optimize nucleotide composition to support synthesis, while maintaining genetic system function.
Breakthrough Capability: Methods for one-step, simultaneous assembly and sequence-verification of long DNA fragments.
●●●● 2021 Milestone: Reliable assembly of 10,000 base pair non-clonal DNA fragments.
Breakthrough Capability: Pipelined synthesis, assembly, and functional testing of engineered genetic systems.
●●●○ 2021 Milestone: Achieve desired functionalities in lower-fidelity, error-prone genetic systems.
Goal: Precision genome editing at multiple sites simultaneously with no off-target effects.
Breakthrough Capability: Ability to reliably create any precise, defined edit or edits (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.
●●●● 2021 Milestone: Ability to generate any defined single base pair change in model organisms.

(Table continues)

Goal: Precision genome editing at multiple sites simultaneously with no off-target effects. *(Cont.)*
Breakthrough Capability: Precise, predictable, and tunable control of gene expression for many genes inside diverse cells and organisms across different timescales.
●●●○ 2021 Milestone: Achieve long-lasting gene repression and activation.
Breakthrough Capability: 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.
●●●● 2021 Milestone: Improve editors to function without sequence requirements (such as protospacer adjacent motif (PAM) sequences) with activity comparable to 2019 state-of-the-art capabilities.

Table 1. Assessment of Engineering DNA 2021 Milestone Achievement. Each 2021 milestone was assessed to determine progress towards its achievement. Four filled circles indicates the 2021 has been achieved or is close to complete, three filled circles indicates significant progress towards the 2021 milestone, two filled circles indicates modest progress towards the 2021, and one filled circle indicates only minimal progress towards achieving the 2021 milestone. In Engineering DNA, all 2021 milestones have been achieved or are close to complete (four filled circles) or have seen significant progress towards their achievement (three filled circles).

Highlights of Technology Developments in Engineering DNA

Universal DNA Assembly Toolkits

DNA assembly is a cornerstone technology necessary for biological research and its continued development will accelerate progress in engineering biology. The arsenal of techniques to assemble small DNA molecules to form larger constructs has rapidly expanded in the past two decades, including polymerase chain reaction, restriction endonuclease digestion, ligation, Type II assembly, and Gibson assembly. However, as the field continues to progress, it's important for researchers to have universal DNA assembly toolkits that will work across multiple species. As discovery of non-model organisms increases, it's important that researchers can efficiently domesticate these organisms through universal DNA assembly toolkits, rather than creating new forms of DNA assembly technology each time a new organism is discovered. Over the past few years, DNA assembly technology development has predominantly focused on creating platforms and tools that increase the use of plasmids and reagents *across* species. There is also great benefit in being able to automate repeated experiments in efforts to scale up production or functionality of engineered microbes towards industrial levels. Engineering biology practitioners recognize the benefit of automating DNA assembly as a valuable parameter to reduce the time it takes to complete a research project. Although DNA assembly and automation are addressed in the roadmap, the focus on and advancement of assembly technologies that are able to work universally across different organisms (including for the domestication of non-model organisms) is notable for progress throughout engineering biology.

More Efficient Long-Read Sequencing DNA Technology

Being able to read DNA sequences is fundamentally important for efficient, accurate, and reproducible development of engineering biology technologies. DNA sequencing technology continues to rapidly evolve to be cheaper, more reliable, and achieve a higher throughput over the past few years. However, longer-read sequencing technology is becoming increasingly important. Non-model organisms, especially in early research efforts, require a first draft of their genome to be compiled and assembled. Genome sequencers collectively take different DNA fragments of this newly identified organism and create a lengthy list of DNA reads, which are the inferred disparate sequences of base pairs belonging to all of the sequenced DNA fragments from that organism. Computational programs have to identify how each of those individual DNA reads collectively match with each other to form the hierarchical organization of that organism's genome. The shorter these disparate

fragments are, the harder this matching becomes due to its computational burden; conversely, longer sequence fragments are easier and less computationally challenging to process. Nanopore sequencing can perform longer reads, thus greatly enabling the genome assembly of non-model organisms to make these initial drafts. Additionally, there is growing excitement in recognizing that many nanopore DNA sequencers are portable and can potentially be used for field studies or in situations where domesticating a non-model organism in the laboratory is difficult. Further developments in DNA sequencing technology will continue to significantly accelerate the verification and discovery of synthetic gene constructs. However, there is a growing appreciation that specific forms of DNA technology, such as nanopore sequencing, may play an outsized role in engineering biology progress.

Engineering DNA Barriers to Progress

Miniaturized, Lab-Portable DNA Synthesis Hardware

Synthesized DNA is a critical reagent for many engineering biologists and is regularly ordered from external manufacturers. As automation and high-throughput technologies enable more engineering biology experiments and assays to be run, the demand for synthesized DNA grows. DNA synthesis companies are developing technologies (see e.g., Twist Bioscience Technology; Evonetix) that enable the synthesis of tens of thousands of DNA strands in parallel, decreasing the cost of synthesis and reducing turn-around times. Laboratories could further increase their experimental bandwidths and move more quickly from experimental design to implementation with on-demand access to miniaturized, "benchtop" versions of high-throughput, paralleled DNA synthesis. While benchtop synthesizers have been available for some time, a new generation of equipment has recently entered the market or is expected to be available soon that enables faster and more accurate synthesis of more oligos (see e.g., DNA Script Syntax System). In the future, distributed equipment that can rapidly print longer strands of synthetic DNA without compromising accuracy may become available. This could not only enable laboratories in well-resourced countries to move more quickly, but, in a global context, could enable researchers who live farther from DNA synthesis providers to potentially reduce their turn-around times and costs significantly.

Established Benchmark Standards for Gene Editing

Genetic editing of organisms has an incredibly high potential to revolutionize science and medicine. However, current benchmark guidance lacks precise verification standards for off-target effects. As researchers continue to use genetic editing in laboratories, standards for what constitutes "efficient" genetic editing and how to precisely measure off-target effects will need to be established. *Engineering Biology* highlights the development of gene-editing technology as an area for advancement but does not specify benchmark standards for what would constitute minimal off-target effects, efficiency ratios, genomic toxicity, and other parameters relevant to gene editing experiments. This lack of standardization can have profound implications, such as the reproducibility of biological systems to produce biomolecules, the mitigation of adverse effects from health biotechnologies in humans, and the promotion of sound, reproducible science. A coalition or network of government, industry, academic, and societal stakeholders to discuss proper benchmarks to ensure safe, reproducible standards would provide a potential remedy to this barrier.

Tools for Assembling and Synthesizing DNA with High GC-Content

DNA sequences with high GC-content can be problematic in DNA synthesis due to their intricate secondary structures, mis-priming, and mis-annealing. Shorter-read sequencing technologies prevalent historically, struggled with sequencing and assembling these high-repeat regions, resulting in incomplete genome sequences for many organisms with higher GC-content (such as complex eukaryotes). Robust longer-read sequencing and assembly tools, along with techniques and technologies to identify errors and losses, for high GC-content DNA are still needed, particularly those that can work with non-model organisms. Increased

prevalence of these tools will also help to grow the number of available genome assemblies and genetic datasets for further research.

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

Manufacturing thousands of very long nucleotides with fidelity is a fundamental technology that can provide innumerous applications in engineering biology. Accomplishment of the goal can also enable high-throughput synthesis of large gene clusters (over ten kilobases), which could be used to create enzymes that can deconstruct lignin and cellulose into monomeric products for clean energy. Further, the synthesis of large oligonucleotides will enable safe, reliable, and efficient delivery vectors for gene editing agents, enabling more bio-based manufacturing processes and for human therapeutics.

Breakthrough Capability: Highly efficient oligonucleotide synthesis to increase the number, length, and fidelity of oligonucleotides. This Breakthrough Capability is proceeding as predicted relative to the original roadmap. The Assessment literature review indicates the 2021 milestone has been achieved.

2021 Milestone: Robustly synthesize one million 200-mer oligonucleotides with a per-nucleotide error rate of fewer than one in 500 nucleotides.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Creating longer DNA oligonucleotides with lower error rates allows researchers to engineer biology more efficiently to perform valuable functions and better characterize natural processes. Evolving forms of DNA production, such as enzymatic DNA synthesis, is critical for achieving these abilities and the completion of this milestone. Several companies are beginning to adopt enzymatic DNA synthesis commercially. "Enzymatic DNA synthesis enters a new phase," a perspective piece by Michael Eisenstein (2020), discusses the synthesis of lengthy oligomers and the benefits of enzymatic DNA synthesis for achieving long, error-free sequences. Twist Bioscience reports the ability to regularly produce up to 300 nucleotide oligomers with an error rate of 1 in 2000 nucleotides (Twist Bioscience, 2021). Similarly, on their "Services" web pages, companies Integrated DNA Technologies, DNAScript, and Camena also attest to routine production of more than 200 base pairs of oligomers. While many academic labs may not have the resources or capacity to achieve this rate of production

Advancements in Enzymatic DNA Synthesis to Build Longer DNA Oligomers

Critically important to engineering biology is the ability to synthesize long DNA oligomers. These longer DNA oligomers allow researchers to better control and engineer useful features in biological organisms and systems, and as a result, allows them to expand the possibilities of useful functions that engineering biology can perform. Traditional oligomer manufacturing methods use phosphoramidite synthesis, which involves multiple rounds of a stepwise assembly of chemically modified nucleotides. The efficiency of this method becomes limited as oligomers reach lengths beyond 200 nucleotides, thus practically limiting the length of what oligos can be produced at scale. In the past few years, newer generation DNA oligomer synthesis enzymes, such as terminal deoxynucleotidyl transferase (TdT), have allowed for the synthesis of longer DNA sequences with great efficiency. In contrast to phosphoramidite synthesis, TdT can efficiently synthesize longer DNA strands in a template-independent fashion and be modified to better incorporate chemically-modified nucleotides. This technology can be applied to health and medicine, where it can be used to enhance the efficacy of nucleic acid-based vaccine research and homology-directed recombination in gene editing; additionally, this generation of longer oligonucleotides could better facilitate the development of DNA as a biomolecular storage medium, which can have profound effects in energy and data infrastructure (Eisenstein, 2020). As enzymatic DNA synthesis becomes more widely adopted across industry and academic labs, it is expected to greatly accelerate many areas of research.

quickly and cheaply, the prevalence of businesses that can meet this need easily provides many sources for obtaining longer oligonucleotides.

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

Assembly of DNA fragments with high fidelity ensures that the products of engineering biology experiments are reliable and can provide finely tuned functionalities towards many applications. For example, completing this goal would enable the creation of variant libraries that could be used to validate models of genetic circuits and pathways for industrial purposes. Additionally, the assembly of multiple high-fidelity fragments could allow multi-gene modification in non-model algae and cyanobacteria for long-term carbon storage. Further, it can also facilitate the production of complex, large, functional DNAs and RNAs, such as for biosensor purposes.

Breakthrough Capability: Predictive design of DNA sequences for improved assembly of longer, more information-rich DNA fragments. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates that the 2021 milestone has been reached. Notably, the Assessment suggests that the 2029 milestone "Design algorithms that identify optimal synthesis strategies for assembling megabase-length genetic systems" may be achieved before anticipated.

2021 Milestone: Coupled design of DNA sequences to optimize nucleotide composition to support synthesis, while maintaining genetic system function.

Progress toward this milestone is **significant**, with some research gaps remaining.

Small changes in the nucleotide composition can affect the stability of a DNA molecule as it is synthesized and how it will function in an engineered system. Since 2019, research has focused mainly on creating data science, bioinformatics, and machine learning tools to optimize nucleotide composition to support synthesis and function separately. Synthesis fidelity was explicitly examined by *Halper, et al.* (2020), which developed a machine learning model, called the Synthesis Success Calculator, to determine if a long DNA fragment can be synthesized with a short turnaround time. A key finding from this study revealed that highly repetitive sequences were one of the most important contributors to DNA synthesis failure. Under this circumstance, *Hossain et al.* (2020) provided a useful solution through the development of a Nonrepetitive Parts Calculator to generate thousands of highly nonrepetitive genetic parts for different uses in synthetic biology. Beyond these non-repetitive parts, there has also been a focus on designing optimal promoters for gene expression systems. *Kotopka, et al.* (2020) used a combination of data science approaches to evaluate the ability of promoters to control gene expression in *Saccharomyces cerevisiae*. Similarly, *Gilman et al.* (2019) developed a broadly applicable method to identify promoters in atypical non-model hosts, such as *Geobacillus thermoglucosidasius*, through bioinformatic filtering and machine learning. There have also been further efforts to design sequences that cater to structural DNA part components, such as *Valeri et al.* (2020) creating the Sequence-based Toehold Optimization and Redesign Model (STORM) and Nucleic-Acid Speech (NuSpeak) to characterize and optimize nucleic acid sensors (known as toeholds). Overall, there are many tools available to predict how DNA sequence design corresponds to its synthesis fidelity and future function (Chechik et al., 2020; Li et al., 2022), yet a coupled workflow of integrating these tools together can better satisfy the original prediction of this milestone.

Breakthrough Capability: Methods for one-step, simultaneous assembly and sequence-verification of long DNA fragments. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached. Of note, the Assessment suggests that the 2024 milestone "Reliably assemble and verify 10,000 base pair clonal DNA fragments" and the 2029 milestone "Reliably assemble and verify 100,000 base pair clonal DNA fragments" may be achieved before anticipated.

2021 Milestone: Reliable assembly of 10,000 base pair non-clonal DNA fragments.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Creating 10,000 base-pair non-clonal DNA fragments enables researchers to introduce sequences that bestow useful biological functions, such as complex regulatory elements that can finely tune expression in genetic systems. Since 2019, research has produced several techniques that have enabled scientists and engineers to create larger DNA fragments more efficiently through plasmid construction and verification workflows. To create very large fragments, *Pryor et al.* (2020) developed a data-optimized design workflow for one-pot Golden Gate assembly demonstrating assemblies of up to 35 DNA fragments. Alongside assembly strategies, it is important to develop standards and verification workflows to ensure that assembly was performed with minimal errors. To this end, *Gallegos et al.* (2020) developed an open-source pipeline to create and verify plasmids in engineering biology, while *Lopez et al.* (2019) and *Currin et al.* (2019) capitalized on nanopore technology for sequence verification of DNA assemblies. Additionally, *Ma et al.* (2019) reported a Guanine/Thymine standard for plasmid construction where DNA sequences are defined as standard, reusable parts for combinatorial assembly. Altogether, the combination of available methods allows for the robust assembly of large DNA fragments and many different means to verify their fidelity (Young et al., 2021).

Breakthrough Capability: Pipelined synthesis, assembly, and functional testing of engineered genetic systems. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates that the 2021 milestone has nearly been reached, though more complete design-to-function pipelines or workflows are needed to fully achieve the milestone.

2021 Milestone: Achieve desired functionalities in lower-fidelity, error-prone genetic systems.

Progress toward this milestone is **significant**, with some research gaps remaining.

Combining synthesis, assembly, and functional testing of genetic systems into an automated pipeline can significantly shorten the time and resources required for engineering biology experiments. Research has focused on developing automation software models, as well as verification methods, to evaluate different features of genetic systems to improve fidelity. Automated model and design software can greatly aid a researcher's ability to forecast the efficacy, or design useful features, in a genetic system, as demonstrated by *Chen et al.* (2020) and *Reis and Salis* (2020). There has also been much development on workflows that verify the functionality and fidelity of features useful for genetic systems. *Gallegos et al.* (2020) developed an open-source pipeline for the creation and verification of plasmids in synthetic biology. Likewise, *Currin et al.* (2019) developed a workflow for highly multiplexed sequencing to verify DNA assemblies using nanopore sequencing technologies. *Fu et al.* (2020) used deep learning approaches to propose a novel codon optimization method for enhancing gene expression. In essence, the individual pieces of a pipelined synthesis, assembly, and functional workflow are being rapidly developed or are mostly in place. Evidence of workflows or pipelines that directly connect the modeling of synthesis, assembly, and functional testing to their physical implementation would better satisfy the original prediction of this milestone.

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

Precision genome editing, especially at multiple sites, can greatly promote the ability of researchers to create more complex, finely tuned modifications that can greatly improve the ability of many engineering biology products. For example, completing this goal may enable the ability to identify and remove transporters involved in the movement of harmful heavy metals in food production. Additionally, precision genome editing can provide parallel and error-free genome engineering of mammalian cell lines to identify drugs to treat non-infectious diseases. Similarly, the high-throughput aspects of this technology can also further the engineering of robust soil biomes by simultaneously genetically editing a variety of soil microbes.

Breakthrough Capability: Ability to reliably create any precise, defined edit or edits (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. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates that the 2021 milestone has been reached and the 2029 milestone "Achieve high-efficiency gene insertion or deletion of moderately significant changes (but less than 10 kilobases) via homologous recombination" may be achieved ahead of schedule.

2021 Milestone: Ability to generate any defined single base pair change in model organisms.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

There is a complex biological process for creating base pair changes and edits across DNA. In a concerted series of events, the editor may have to create a physical break in the DNA, subsequently modify or interject nucleotides from a cargo protein, and then ensure the downstream DNA repair events favor the substituted change. Increasing the reliability of DNA editors thus greatly aids engineering biology researchers in being able to manipulate features useful for genetic circuits and engineered organisms, as well as investigate the impact of small mutations on gene expression. Research since 2019 has focused on increasing the ability of editors to make new forms of nucleotide transitions while minimizing off-target effects. To lessen the degree of required reagents and molecules for editing, *Anzalone et al.* (2019) described prime editing, which uses an engineered molecule to edit the genome without the need for double-strand breaks or donor DNA. In addition, several researchers have introduced new forms of base editors to introduce more difficult types of nucleotide or genome modifications. *Zhao et al.* (2020) created several glycosylase base editors that can edit C-to-A transversions in *Escherichia coli* and C-to-G transversions in mammalian cells. In a similar vein, *Zuo et al.* (2020) engineered a cytosine base editor that retains high on-target activity while minimizing off-target effects. Additionally, in their preprint, *Choi et al.* (2021) presented the prime editing method, Prime-Del, that can create precise genome deletions by using paired guide RNAS to target a site of the genome to be deleted and better control the downstream repair event to result in a favorable outcome. Critically important for DNA editing is being able to control the activation and deactivation of the editor, such as in experiments by *Pan et al.* (2021) describing CRISPR-Act 3.0, a highly robust, multiplex, RNA-guided CRISPR activation system that can activate multiple genes in plants. Conversely, *Carlson-Stever et al. (2020)* developed CRISPRoff, a method for light-induced degradation of sgRNA for precise spatio-temporal control that can effectively disable CRISPR editing. Lastly, in efforts to mitigate the potentially adverse effects of DNA editing, *Manzano et al.* (2020) demonstrated how to use ultrafiltration to purify Cas9-RNA complexes to remove potentially harmful excess RNA that can be detrimental to experiments. Comprehensively, there has been much development of useful functionalities for DNA editors for engineering biology, thus satisfying the original prediction of this milestone.

Breakthrough Capability: Precise, predictable, and tunable control of gene expression for many genes inside diverse cells and organisms across different timescales. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been nearly achieved.

2021 Milestone: Achieve long-lasting gene repression and activation.

Progress toward this milestone is **significant**, with some research gaps remaining.

Numerous applications in engineering biology, including cell-free expression systems and genetic circuits, rely on long-lasting gene repression and activation strategies for the timely production of biomolecules. Since 2019, research has focused on technologies that better quantify gene expression, allow a greater magnitude of expression control, or aid in discovery of fundamental molecules responsible for complex expression regulation in model organisms. Several technologies and methods have been developed to achieve longer lasting gene repression. *Reis et al.* (2019) used nonrepetitive extra-long single-guide RNAs to repress up to 13 genes by

3,500-fold. To create longer-lasting repression in mammalian systems, Nuñez et al. (2021) developed CRISPRoff, a programmable epigenetic memory writer consisting of a single dead Cas9 fusion protein to achieve long lasting repression in human cells. Repression strategies have even been explored in the context of gene drives, where Rottinghaus et al. (2022) engineered a CRISPR-based kill switch in *E. coli* and demonstrated that it is able to control the Salt Overly Sensitive (SOS) signaling pathway. To better understand what biological features may control longer lasting gene activation and repression, researchers have also developed several characterization datasets and tools to better measure expression activity. Fontana et al. (2021) identified multiple characteristics of bacteria promoters that impose strict requirements on CRISPR activation sites. Advancing the development of toolkits, DeLorenzo et al. (2021) developed a CRISPR Interference Tool that facilitates gene expression studies in the non-model organism *Rhodococcus opacus*. Similarly, Gurdon et al. (2020) developed a procedure to measure how a transcription factor can stabilize gene expression and cell fate commitment. Although much has been achieved in gene repression and activation there is still room for further progress towards longer gene repression and more acute control of gene activation.

Breakthrough Capability: 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. This Breakthrough Capability is proceeding as predicted relative to the roadmap. Although the literature review indicated that research for the 2021 milestones have been reached, the Assessment indicates that the 2024 milestone "Routine use of editors without detectable off-target effects (less than 0.001% off-target editing)" may not be achieved on the timeline anticipated.

2021 Milestone: Improve editors to function without sequence requirements (such as protospacer adjacent motif (PAM) sequences) with activity comparable to 2019 state-of-the-art capabilities.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Improving editors to function without sequence requirements can greatly increase the applicability of genomic editing to create a larger diversity of genetic changes, and therefore increase the ability of engineered organisms to create useful products. Research in the past two years has focused on enhancing the ability of CRISPR-Cas systems to perform genetic editing and the ability to transfer useful genetic editing cargo into cells. Several CRISPR-Cas9 modifications or variants have been developed that improved editors to function without sequence requirements. Anzalone et al. (2019) described prime editing, which uses an engineered molecule to edit the genome without the need for double-strand breaks or donor DNA, to minimize the need for reagents. Directly promoting PAM-less functionality, Walton et al. (2020) engineered a variant of *Streptococcus pyogenes* Cas9 (SpRY) that exhibits robust activity on a wide range of sites. Detailing a possible tool that can be extrapolated towards editing uses in other organisms, Swarts et al. (2019) described how *Francisella novicida* Cas12a exhibits PAM-independent ssDNA trans-cleavage activity when triggered by binding to a crRNA-guide-complementary ssDNA. We have also seen the development of methods and technologies to better transport DNA editing cargo into cells, thereby increasing genetic editing efficacy. Sun et al. (2020) engineered a DNA nanoclew-based carrier for enhanced delivery of CRISPR-Cas12a RNA ribonucleoprotein to better regulate cholesterol levels, showing increased protection and delivery of the gene editing cargo. These enhanced delivery capabilities were even shown to have potential therapeutic outcomes, as Zhang et al. (2020a) demonstrated with packaged Cas9 nucleases in single-stranded adeno-associated viruses which enhanced the correction of Duchenne muscular dystrophy corrective editing in mice. In yet another form of improving delivery, Liu et al. (2019) uncovered that a boronic acid-rich dendrimer could deliver native protein cargo to the cytosol, including Cas9 protein, with a higher degree of fidelity compared to *status quo* methods. The demonstration of several technologies and methods to improve PAM-less ability to perform edits, as well as increase editing efficiency through better cargo delivery strategies, greatly satisfies the original prediction of this milestone.

Biomolecular Engineering | Biomolecule, Pathway, and Circuit Engineering

Engineering Biology defines Biomolecule, Pathway, and Circuit Engineering ("Biomolecular Engineering") as focusing "on the importance, challenges, and goals of engineering individual biomolecules themselves to have expanded or new functions." The roadmap notes that "successful progress would be demonstrated by production of functional macromolecules on demand from both natural and non-natural building blocks, targeted design of complex circuits and pathways, and control over the dynamics of regulatory systems." *Engineering Biology* stated that biomolecular engineering "historically has been an exercise in building out from what exists in nature to what doesn't." This observation remains true, with several new technologies emerging around building of biomolecules from both canonical and non-canonical components. Since 2019, major progress has been achieved in biomolecular engineering, particularly in protein structure prediction and synthetic immunology. In addition to the roadmap predictions, emergent machine learning technologies have been brought to bear on many biomolecular engineering applications with significant impact.

Progress in Biomolecule, Pathway, and Circuit Engineering	
Goal: On-demand design, generation, and evolution of macromolecules for desired functions.	
Breakthrough Capability: De novo prediction of RNA structure, protein structure, and complexes of DNAs/RNAs and proteins from primary sequence and the ability to make accurate predictions of mutability and effect of mutations from structure.	
●●●○	2021 Milestone: Reliably predict (greater than a 50% success rate) the structure of 300-amino acid proteins and 200-nucleotide RNA domains within 5 Ångstroms from primary sequence.
●●○○	2021 Milestone: Improve force-field and backbone-sampling algorithms and include capabilities to capture force-fields of post-transcriptionally- and post-translationally-modified nucleosides and amino acids.
Breakthrough Capability: De novo design and/or prediction of macromolecular dynamics and dynamic macromolecular structures.	
●●●○	2021 Milestone: Improving computational models of RNA dynamics that can incorporate experimental data.
Breakthrough Capability: High-throughput integrated computational, experimental, and evolutionary schemes for refinement of desired biomolecule functions including enzymatic activity and binding.	
●●●○	2021 Milestone: Durable and high-mutation-rate in vivo continuous DNA mutagenesis and evolution systems in model organisms.
Goal: Special considerations for on-demand design, generation, and evolution of macromolecules that rely on non-canonical/unnatural building blocks.	
Breakthrough Capability: PCR, reverse transcription, cellular replication, and transcription of fully unnatural nucleotide-containing genes of up to 400 base pairs.	
●●○○	2021 Milestone: Identification of "missing" functionality or functionalities in A-T-G-C base pairs.

(Table Continues)

Goal: Special considerations for on-demand design, generation, and evolution of macromolecules that rely on non-canonical/unnatural building blocks. *(Continued)*
Breakthrough Capability: Expanded genetic code systems for translation of >100-amino acid proteins containing fully-unnatural amino acids, and proteins with at least four, distinct unnatural amino acid building blocks.
●●●○ 2021 Milestone: Create proteins that are capable of gaining new, therapeutically-useful activities through unnatural amino acids.
Goal: Holistic, integrated design of multi-part genetic systems (i.e., circuits and pathways).
Breakthrough Capability: Design of highly-stable, large genetic systems (genomes) with targeted expression levels in a host organism or cell type, incorporating system-wide effects.
●●●● 2021 Milestone: Incorporate gene expression interactions into predictable design of prokaryotic genetic systems.
Breakthrough Capability: Ability to rationally engineer sensor suites, genetic circuits, metabolic pathways, signaling cascades, and cell differentiation pathways.
●●●○ 2021 Milestone: Reliable engineering of genetic circuits with more than ten regulators for sophisticated computations.
Goal: Integrated design of RNA-based regulatory systems for cellular control and information processing.
Breakthrough Capability: Porting nucleic acid strand displacement technology into cellular systems with RNA instantiations.
●●●● 2021 Milestone: RNA implementation of strand displacement cascades in bacteria.
Breakthrough Capability: Porting successes in computationally designed bacterial RNA-based genetic regulators into eukaryotic and mammalian systems.
●●○○ 2021 Milestone: First generation eukaryotic RNA-based gene regulators that utilize RNA:RNA interactions and/or strand-displacement and achieve 10-fold change in gene expression.
●●●● 2021 Milestone: Creation of RNA modification machinery that allows programmable site-specific modifications of RNA, focusing on naturally abundant modifications (N6-methyl adenosine, 2'-O-methylation, pseudouridine).

Table 2. Assessment of Biomolecular Engineering 2021 Milestone Achievement. Each 2021 milestone was assessed to determine progress towards its achievement. Four filled circles indicates the 2021 has been achieved or is close to complete, three filled circles indicates significant progress towards the 2021 milestone, two filled circles indicates modest progress towards the 2021, and one filled circle indicates only minimal progress towards achieving the 2021 milestone. In Biomolecular Engineering, the 2021 milestones have been achieved or are close to complete (four filled circles), or have seen significant (three filled circles) or modest progress (two filled circles) towards their achievement.

Highlights of Technology Developments in Biomolecular Engineering

Advancements in Protein Structure Modeling and Prediction

Over the past few decades, there has been considerable development in tools that can predict protein structure from an amino acid sequence – though the predictions are often limited by accuracy, speed, and relatively inefficient homology analysis methods. To expedite the time required to obtain three-dimensional models of protein structures, researchers had previously developed modeling algorithms for protein structure prediction

Health & Medicine Application: Protein Engineering for Synthetic Immunology

Immunotherapy is the treatment of a disease by activating or suppressing the immune system. This activation or suppression happens through the coordinated regulation of several endogenous biomolecules. As these biomolecules can be actively engineered, there has been significant interest in producing synthetic biomolecules that can potentially regulate the immune system to generate precise therapeutic responses. In turn, this would ultimately expand a clinician's ability to control the immune system and provide better treatment outcomes. The past two years have seen a rapid advancement in synthetic protein engineering, especially when coupled with already-powerful immunological treatments such as chimeric antigen receptor (CAR) T-cell therapies (Cox and Blazeck, 2021). For example, Choe et al. (2021) developed synNotch CAR-T cells, implementing a synthetic Notch-CAR circuit in T cells, to treat problematic mesothelioma, ovarian cancer, and glioblastoma cancers in mouse models. The ability of these synNotch CAR-T cells to treat these diseases was found to be more effective than traditional CAR-T cell therapy strategies, including an enhanced ability to limit toxicity to healthy tissue and prevent tumor escape. The study represents a growing trend in engineering biology towards synthetic immunology therapies and improved medical biotechnologies.

and the past two years have shown advancements in such modeling. Amongst one of the most powerful developments is AlphaFold 2, an artificial intelligence program developed by Alphabet and Google's neural network DeepMind (Jumper, 2021). AlphaFold 2 uses an artificial intelligence deep learning technique to predict protein structure, building upon and showing incredible improvements over its predecessor, AlphaFold 1. In 2020's Critical Assessment of Techniques for Protein Structure Prediction (CASP) competition, a benchmark that measures structure prediction efficiency determined that AlphaFold2 correctly predicted the structure of about 60% of the proteins in their line-up with no previously known structural information (with predictions achieving a global distance test scoring above 90, out of 100) (Service, 2020). AlphaFold 2's open-source software and proteome database are published and can now be accessed at https://github.com/deepmind/alphafold and https://alphafold.ebi.ac.uk/, respectively. In addition to this incredible achievement, parallel efforts by *Baek et al.* (2021) have developed the artificial intelligence program RoseTTAfold to generate high-quality protein structure predictions, to predict protein:protein complex structures, and to solve x-ray crystallography and cryo–electron microscopy modeling problems. The combination of these two achievements, along with concurrent research efforts, has drastically accelerated fundamental research and contributed to the advancement of engineering biology (Service, 2021). The AlphaFold database of over 200 million protein structure predictions is now freely available to all researchers (https://alphafold.ebi.ac.uk/).

Machine Learning to Refine Biomolecular Function

Machine learning has been widely predicted to catalyze advancements in engineering biology, especially by incorporation into tools that predict protein structure. Beyond this application, machine learning algorithms have been applied to identifying potential properties of unnatural amino acids (Giannakoulias et al., 2021), recommending strains for design-test-build-learn cycles in metabolic engineering (Radivojević et al., 2020), and determining the likely effectiveness of RNA "toe-hold" sequences to respond to desired target sequences (Angenent-Mari et al., 2020) For example, *Wu et al.* (2019) incorporated machine learning to explore how multiple simultaneous mutations would impact directed protein evolution experiments. Although *Engineering Biology* forecasted a vital role of machine learning in predicting structure-function relationships for biomolecules, machine learning continues to find *other* applicable roles in biomolecular engineering outside structure prediction. (For more about machine learning advancements for engineering biology, see **Data Science**.)

Biomolecular Engineering Barriers to Progress

Fragile Genetic Circuits Susceptible to Mutations

Engineered genetic circuits can be prone to stability issues over time due to metabolic burden and toxicity, leading to selective evolutionary pressure against the incorporated circuit. Current research solutions are examining a myriad of approaches to improve circuit robustness, including how to use sequencing technologies to better monitor mutations in genetic circuits and the design of more stable genetic circuits by insulating critical DNA sequences (Yannick Ouedraogo et al., 2023; Costello & Badran, 2021; Şimşek et al, 2022). It's important to recognize that, as many of these genetic circuits must be scaled for industrial purposes, that their fragility constitutes a manufacturing risk. One approach to help overcome this barrier is to establish measures and standards of resilience that can be used as benchmarks for engineered circuit research.

Biomolecular Engineering Goal: On-demand design, generation, and evolution of macromolecules for desired functions.

The design, generation, and evolution of macromolecules allows practitioners to engineer changes in macromolecular structure to dictate useful downstream functions. There are several potential applications that can be realized through the completion of this goal, including the engineering of biological polymers that are durable and biodegradable, such as novel or redesigned plastics. The discovery of new macromolecule characteristics could enable the at-will design of non-natural pathways for the *de novo*, model-based creation of proteins, producing novel products and materials that do not exist in nature.

Breakthrough Capability: De novo prediction of RNA structure, protein structure, and complexes of DNAs/RNAs and proteins (from primary sequence) and the ability to make accurate predictions of mutability and effect of mutations from structure. This Breakthrough Capability is close to meeting the pace of predictions relative to the original roadmap. The Assessment literature review indicates that one of the 2021 milestones have not been reached, while a second 2-year milestone is on track.

2021 Milestone: Reliably predict (greater than a 50% success rate) the structure of 300-amino acid proteins and 200-nucleotide RNA domains within 5 Ångstroms from primary sequence.

Progress toward this milestone is **significant**, with some research gaps remaining.

Recent success with protein structure prediction has not yet translated to RNA structures, although there have been major incremental improvements. *Sato et al. (2021)* discussed their algorithm improvements with MXfold2, which achieves robust predictions of RNA secondary structures by addressing obstacles commonly seen with overfitting of data. Likewise, *Townshend et al. (2021)* introduced a machine learning approach that incorporates a scoring function, the Atomic Rotationally Equivariant Scorer (ARES) to better identify accurate RNA structure models. Some new tools for RNA structure prediction are also being created; *Singh et al. (2019)* created SPOT-RNA, a software that uses deep contextual learning for base-pair prediction including non-canonical and non-nested (pseudoknot) base pairs for RNA structure modeling. Additive strategies to make these RNA prediction models more accurate are developing as well, with *Kappel et al. (2020)* showing how to use cryo-electron microscopy to resolve maps of RNA-only systems, which can subsequently be combined with other modeling and mapping technologies to establish structures of RNA molecules. Although the current arsenal of *protein* structure predictors satisfies the 2021 benchmark of progress towards these milestones, RNA structure predictors still need to see improvement.

2021 Milestone: Improve force-field and backbone-sampling algorithms and include capabilities to capture force-fields of post-transcriptionally- and post-translationally-modified nucleosides and amino acids.

Progress toward this milestone is **modest**, with significant research gaps remaining.

Computational models that estimate the forces between and within atoms and molecules, and related backbone sampling models, help engineering biology practitioners understand the biomolecular interactions critical for engineering useful functions. One of the most popular tools to calculate the force-field of a biomolecule is SIRAH 2.0 (so named for its lineage and function, the South-(A)merican Initiative for a Rapid and Accurate Hamiltonian) developed by *Machado et al.* (2019). Many ongoing efforts are working to improve force-field algorithms compatible with SIRAH 2.0 and similar models. For example, *Garay et al.* (2020) presented a set of topologies and interaction parameters for the most common protein post-translational modifications for more accurate SIRAH modeling. Researchers are also developing simulations to better understand how modified nucleotides are affecting biomolecular structural forces. For instance, *Hurst and Chen* (2021) used alchemical and temperature replica exchange molecular dynamics (TREMD) on RNA duplexes to probe the structural effects of modified and mutant nucleotides. There has also been a focus on better measuring force-fields of intrinsically disordered proteins (proteins that lack a fixed or ordered three-dimensional structure). In their characterization study, *Rieloff and Skepö* (2020) examined how phosphorylation of an N-terminal fragment of intrinsically disordered proteins affect conformational changes measured by AMBER ff99SB-ILDN and CHARMM36m force-fields. Progress in intrinsically disordered protein research has contributed towards understanding complex biophysical phenomena. *Perdikari et al.* (2021) developed a coarse-grain model (similar to SIRAH) that better characterizes the liquid-liquid phase separation properties conferred by post-translationally modified intrinsically disordered proteins. Finally, understanding the orientations, residues, and geometry of biomolecules is contributing to more accurate protein structure prediction efforts. For instance, *Yang et al.* (2021) improved the accuracy and speed of protein structure predictions by implementing a deep residual network (a computational method used for deep learning and task models) that predicts residue orientation, residue distances, and minimizes the energy configuration of the proposed biomolecular structure. While there have been some advancements, progress still needs to be made to comprehensively capture force-fields and understand the effects of post-transcriptional and post-translational modifications.

Breakthrough Capability: De novo design and/or prediction of macromolecular dynamics and dynamic macromolecular structures. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates the 2021 milestones have been reached.

2021 Milestone: Improving computational models of RNA dynamics that can incorporate experimental data.

Progress toward this milestone is **significant**, with some research gaps remaining.

Molecular folding dynamics greatly affect stability, function, and biocompatibility of RNA molecules; thus, many engineering biology researchers are trying to better capture RNA dynamics through improvements in predictive computational models. Since 2019, research has focused on improving or creating models that can better incorporate experimental data. Some of these studies have focused on how to more efficiently process data for RNA dynamic simulations. For instance, *Xu et al.* (2022) described a computational strategy that better measures RNA cotranscriptional folding by classifying the molecule into 'partitions' that better model the folding kinetics. Additionally, newer methodologies have emerged such as Reconstructing RNA Dynamics from Data (R2D2), presented by *Yu et al.* (2021), a method that computationally models cotranscriptional folding pathways from selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-seq) data. Lastly, technology to measure RNA folding dynamics has been incorporated into current research priorities, including those of immediate global significance: *Bottaro et al.* (2021) predicted the structure and dynamics of the five 5' RNA stem loops of SARS-CoV-2 through molecular dynamic simulations, identifying structural features potentially relevant for function and drug design. In summary, there have been marked improvements on incorporating experimental data into computational models for RNA dynamics, but these improvements are rather limited and further development to incorporate experimental data could still greatly improve the field.

Breakthrough Capability: High-throughput integrated computational, experimental, and evolutionary schemes for refinement of desired biomolecule functions including enzymatic activity and binding. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached.

2021 Milestone: Durable and high-mutation-rate in vivo continuous DNA mutagenesis and evolution systems in model organisms.

Progress toward this milestone is **significant**, with some research gaps remaining.

Directed evolution, which applies an artificial selection force to a biological system to guide it towards a desired outcome, is a powerful strategy to generate valuable biomolecules or engineer unique functions. Since 2019, research has created several platforms for continuous directed evolution by diversifying the types of organisms that can be directly evolved or creating new methods of artificial selection. For instance, *Miller et al.* (2020) described a protocol for phage-assisted continuous evolution (PACE) to enable the continuous evolution of bacteria species through the use of bacteriophages, creating a platform much faster than conventional strategies. In a similar note, *Cravens et al.* (2021) used TaRgeted *In vivo* Diversification ENabled by T7 RNAP (TRIDENT) to perform continual, and inducible diversification at genes for engineered biological systems. Additionally, *English et al.* (2019) developed Viral Evolution of Genetically Actuating Sequences (VEGAS), a platform for directed evolution in mammalian cells. Various strategies also show the use of directed evolution as a promising vehicle to generate useful biomolecules for industrial settings, such as *Rix et al.* (2020)'s use of the continuous directed evolution platform, OrthoRep, to generate promiscuous enzyme variants of the *Thermotoga maritima* tryptophan synthase β-subunit to perform useful secondary functions. Finally, there are also platforms that integrate existing directed evolution programs to make them more effective and streamlined. A notable example includes the contribution of *Zhong et al.* (2020)'s Automated Continuous Evolution (ACE), a platform that pairs with Orthorep and eVOLVER (an automated culture device for regulating growth conditions) to make directed evolution experiments easier for researchers. In summary, there is an abundance of research activity towards directed evolution platforms. Better control and higher rates over the mutational preferences of *in vivo* continuous DNA mutagenesis systems would satisfy the need for greater progress to mark the completion of this milestone.

Biomolecular Engineering Goal: Special considerations* for on-demand design, generation, and evolution of macromolecules that rely on non-canonical/unnatural building blocks.

*Note from *Engineering Biology*: The design, generation, and evolution of macromolecules containing unnatural building blocks relies on the achievement of the same capabilities as the production of natural macromolecules. This Goal reflects the special considerations necessary for the utilization of unnatural building blocks.

New forms of biotechnology are beginning to use modified or unnatural building blocks to confer impactful properties for a myriad of applications. Critical to expanding the use of unnatural or modified building blocks is to understand how they affect the fidelity of mainstay laboratory techniques (such as PCR) or how feasibly these building blocks can be synthesized in host systems. Understanding these effects can enable several forms of downstream applications, including some wholly unique to engineered biology. Application of non-canonical amino acids in macromolecules are useful in research to investigate protein:protein interactions and clarify biological circuits and pathways, and to control cellular processes through novel and unique post-translational modifications. The subsequent proteins and enzymes can be valuable for environmental sensing and signaling or creating therapeutic proteins to precisely control activity. (See *Young & Schultz*, 2018 and *Adhikari et al.*, 2021 for review of such applications.)

Breakthrough Capability: PCR, reverse transcription, cellular replication, and transcription of fully unnatural nucleotide-containing genes of up to 400 base pairs. This Breakthrough Capability is not meeting the pace of predictions relative to the roadmap. The Assessment literature review indicates the 2021 milestone has not been reached.

2021 Milestone: Identification of ''missing'' functionality or functionalities in A-T-G-C base pairs.

Progress toward this milestone is **modest**, with research gaps remaining.

The canonical nucleotides adenine, thymine, guanine, uracil, and cytosine can be modified to incorporate specialized chemical functionalities using metal chelators and novel functional groups. Research since 2019 has focused on understanding the effects of these modifications on fundamental genetic, biochemical, and material properties. While there has been some significant development in understanding newer or "missing" functionalities of canonical A-T-G-C and U base pairs, based on the Assessment literature review, there has been comparatively more effort towards studies on incorporating unnatural amino acid or nucleotide building blocks, regardless of function. Recent development has been applied toward understanding how the incorporation of nontraditional nucleotides, or modified natural nucleotides, may affect base pair interactions. For example, *Antczak et al.* (2019) presented RNAvista, a database that predicts an extended RNA structure for canonical and non-canonical interactions between base pairs. Similarly, *Flamme et al.* (2020) demonstrated how enzymatic addition of metal cations into nucleic acids can form chromium-mediated metal base pairs for a myriad of applications, including the synthesis of nanowires, energy charge-transfer devices, and the expansion of the genetic alphabet. Current and future uses of modified nucleic acids are further discussed in a review article by *Duffy et al.* (2020), who summarized the use of modified nucleic acids in replication, evolution, and next-generation therapeutics. Researchers are also performing more finely tuned studies on the nanoscale structure of nucleic acid polymers, providing valuable data to better understand how relatively tiny structural forces can confer biological function. In this instance, *Shekaari and Jafari* (2019) modeled the DNA nanobio structure at the base-pair level using statistical mechanics and elucidate what factors are involved with structural formation. Further comprehensive investigation using automated or screening technologies can drive further progress toward this milestone.

Breakthrough Capability: Expanded genetic code systems for translation of >100-amino acid proteins containing fully-unnatural amino acids, and proteins with at least four distinct unnatural amino acid building blocks. This Breakthrough Capability is not meeting the pace of the predictions relative to the roadmap. The Assessment literature review indicates the 2021 milestone "Create proteins that are capable of gaining new, therapeutically-useful activities through unnatural amino acids" has not been achieved to the extent anticipated by the roadmap. It is noted, however, that critical developments in the past year have paved the path for this capability.

2021 Milestone: Create proteins that are capable of gaining new, therapeutically-useful activities through unnatural amino acids.

Progress toward this milestone is **significant**, with some research gaps remaining.

Unnatural amino acids are defined as amino acids not among the twenty found in nature; they can function as structurally similar analogs or their structure can differ significantly from canonical amino acids. There have been several instances of researchers demonstrating use of unnatural amino acids to contribute to therapeutic activity. For example, *Robertson et al. (2021)* examined how the removal of specific cellular transfer RNAs in *Escherichia coli* enables the efficient synthesis of candidate non-canonical amino acids and creates viral resistance in the host organism. Similarly, *Shi et al.* (2021) illustrated that the incorporation of unnatural amino acids can partially restore endogenous protein expression in cases of adverse nonsense mutations. Researchers are also finding potential diagnostic use of unnatural amino acids, as *Zerfas et al.* (2020)

exemplified with a set of improved fluorescent probes that can monitor proteasome activity in live cells. Several efforts have also been underway to recode the genome to expand the capabilities of organisms to create new amino acids. For instance, *Fredens et al.* (2019) created a recoded and refactored *Escherichia coli* strain to show that the number of codons used to produce canonical amino acids can be reduced, creating a 61-codon organism. Similarly, *Fischer et al.* (2020) performed a systematic analysis of unnatural codons to identify nine that can produce an unnatural protein with nearly complete incorporation of an encoded non-canonical amino acid, effectively creating the first 67-codon organism. In summary, while there has not been significant dedicated research toward using unnatural amino acids to create specific proteins or large complexes that perform a precise biological function to restore an adverse condition, there have been many promising discoveries in using unnatural amino acids to create potentially useful diagnostic and therapeutic tools. Comprehensive investigation into the possible therapeutic opportunities of unnatural amino acids through automated, screening, or data science technologies can greatly improve progress towards this milestone.

Biomolecular Engineering Goal: Holistic, integrated design of multi-part genetic systems (i.e., circuits and pathways).

Larger and more comprehensive genetic systems can allow researchers to create more useful functionalities such as sensors, genetic circuits, transporters, metabolic pathways, organelle compartments, and orthogonal expression systems. This increased functionality can enable myriad technologies affecting several application and impact sectors, such as the rapid design and production of custom enzymes and enzyme pathways used in industrial biotechnology, or engineering plants to contain a higher lignin content and lower cellulose/hemicellulose content for greater biomass stimulation for energy applications. Potential impacts in food and agriculture also exist, such as designing multi-part systems to improve the specificity and properties of enzymes involved in provitamin biosynthesis to increase agricultural yield.

Breakthrough Capability: Design of highly-stable, large genetic systems (genomes) with targeted expression levels in a host organism or cell type, incorporating system-wide effects. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates that research for the 2021 milestone has been reached.

2021 Milestone: Incorporate gene expression interactions into predictable design of prokaryotic genetic systems.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Multiple genetic expression inputs, such as changes in transcription, translation, and mRNA decay, can interact with each other and drastically affect the function of an engineered organism. Since 2019, research has focused on how to build systems to better control multiple inputs of gene expression and modify different organisms, including eukaryotes, to accommodate genetic circuit designs more effectively. Several efforts have created engineered bacteria strains, genetic modification tools, or computational strategies that precisely control genetic expression across species. For example, *Meyer et al.* (2019) created *Escherichia coli* "Marionette" strains to generate twelve high-performance small-molecule biosensors to more tightly control gene expression systems. Strategically implementing CRISPR-Cas systems for controlling gene expression, *Tickman et al.* (2021) developed design principles for engineering multiple layers of CRISPR-Cas activation and inactivation in genetic circuits regulated by guide-RNAs for cell-free and bacterial systems. Similarly, *Kiattisewee et al.* (2021) used design principles learned in *E. coli* to implement a CRISPR-Cas activation system in *Pseudomonas putida* and regulate biosynthesis in the biopterin and mevalonate pathways. Researchers have also examined how to design computational strategies for genetic expression control, as *Glasgow et al.* (2019) designed binding sites at the interface of protein heterodimers to create a generalized computation design strategy for modular protein sense-response systems. Although this milestone was specific towards improvements in prokaryotes, it should be noted that much research in the past year also focused on gene

expression modification tools that can be used inclusively across many systems. A notable example is *Hossain et al.* (2020)'s development of the Nonrepetitive Parts Calculator to generate thousands of highly nonrepetitive genetic parts. Future efforts are likely to focus on how to begin to apply these technologies towards more complex eukaryotes and increase the number of genetic regulators.

Breakthrough Capability: Ability to rationally engineer sensor suites, genetic circuits, metabolic pathways, signaling cascades, and cell differentiation pathways. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached.

2021 Milestone: Reliable engineering of genetic circuits with more than ten regulators for sophisticated computations.

Progress toward this milestone is **significant**, with some research gaps remaining.

Genetic circuits can perform increasingly complex and sophisticated tasks with a higher number of internal regulators. These regulators have to intricately account for many biological feedback and workflow mechanisms to adjust the transformation of a chemical or biological input into a useful product. Since 2019, research has investigated how to improve the functionality and diversity of existing regulator classes and prescribe generalizable rules on how complex circuits can be designed. A lot of this development has taken the form of generating new genetic circuit parts, or use of design software to generate parts, in select host organisms. *Taketani et al.* (2020) used the genetic circuit design software, Cello, to design and combine regulatory circuit parts for *Bacteroides thetaiotaomicron*, a human-associated bacterium that holds promise for gut microbiome therapy. Additionally, *Chen et al.* (2020) developed nine insulated gene expression logic gates in *Saccharomyces cerevisiae* that use RNA polymerase flux as the signal carrier in automated genetic circuit design. Some of these advancements have exploited the CRISPR-Cas system for regulatory purposes. For example, *Wu et al.* (2020) created a programmable biosensor using CRISPR inactivation for genetic circuits in *Bacillus subtilis*, potentially identifying a strategy to automatically control key metabolic modules in other microbial species. There has also been a focus on developing parts for mammals: *Muldoon et al.* (2021) engineered multifunctional proteins with transcriptional and posttranscriptional control for mammalian cells. Finally, there has also been the development of methods and procedures to measure how genetic circuits can impact the host organism. *Borujeni et al.* (2020) used RNA sequencing to debug and quantify a genetic circuit's impact on a host by measuring RNA polymerase movement and ribosome usage. While there has been much development towards increasing the number of regulators within genetic circuits, it has not reached a point of development where researchers can *reliably* engineer ten or more regulators to effectively work every time. Further discovery of universal principles of sophisticated genetic circuit design to identify host-specific parts can drive further progress toward this milestone.

Biomolecular Engineering Goal: Integrated design of RNA-based regulatory systems for cellular control and information processing.

Historically, RNA-based regulatory systems have offered many benefits over their protein counterparts, particularly when incorporated into computational strategies for studying and designing nucleic acids. They are also utilized as modular and programmatic mechanisms to regulate circuits, capitalizing on their secondary structure. This increased functionality can offer numerous advantages for impactful and complex engineering biology applications. For example, this increased functionality can be leveraged to evolve organisms with more efficient photosynthetic light-harvesting, enabling feedstock crops for biofuels to reduce global energy consumption (Beckmann et al., 2009). Alternatively, in health and medicine applications, RNA-based regulatory systems can enable the development of highly effective therapeutics for genetic diseases and other illnesses (see for review *Burnett & Rossi*, 2012 and *Zhu*, 2022).

Breakthrough Capability: Porting nucleic acid strand displacement technology into cellular systems with RNA instantiations. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached.

2021 Milestone: RNA implementation of strand displacement cascades in bacteria.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

RNA strand-displacement reactions can build complex and modular systems into bacterial genetic circuits. The past two years of research has focused on developing new RNA regulatory systems and measuring the effectiveness of existing regulators in diverse bacterial species. Some of these advancements have focused on understanding how the design of RNA molecules can affect logic and sensing capabilities. For example, *Kim et al.* (2019) reported using *de novo* RNA design to develop translation-repressing riboregulators through toehold and three-way junction repressors to achieve up to 300-fold expression changes. Other advancements have focused on refining combinatory protein-RNA regulatory systems for increased control of genetic circuits, such as the use of naturally occurring, self-cleaving ribozymes to create gate complexes and strand-displacement circuits for an autonomous, continuous expression system reported in *Bae et al.* (2021). *Oesinghaus and Simmel* (2019) engineered a Cas12a-based DNA processing complex that can be triggered by single-stranded RNA molecules to function as a strand-displacement logic gate. There have also been advancements towards implementing RNA regulatory systems in non-canonical bacterial species. For example, *Strobel et al.* (2019) established the mechanism for ZMP/ZTP riboswitch antitermination in *Clostridium beijerinckii* by determining the cotranscriptional folds and rearrangements that modulate its activity. In summary, there are many different implementations of RNA regulators for both canonical and non-canonical species. Further work to examine how RNA secondary structure-based conformations impact further circuit tunability could expand upon this progress.

Breakthrough Capability: Porting successes in computationally designed bacterial RNA-based genetic regulators into eukaryotic and mammalian systems. This Breakthrough Capability is not meeting the pace of the predictions relative to the roadmap. The Assessment suggests that the 2021 Milestone "First generation eukaryotic RNA-based gene regulators that utilize RNA:RNA interactions and/or strand-displacement and achieve 10-fold change in gene expression" has not yet been achieved, and the 2024 milestone "Second generation eukaryotic RNA-based gene regulators that are suitable for computational design to create libraries that are highly-orthogonal and high-performing, achieving 100's-fold change in gene expression" may not be achieved on the timeline predicted by the roadmap.

2021 Milestone: First generation eukaryotic RNA-based gene regulators that utilize RNA:RNA interactions and/or strand-displacement and achieve 10-fold change in gene expression.*

*For the purpose of this Assessment, we report only on publications that *meet the 10-fold expression change* benchmark in porting bacterial RNA-based gene regulators to eukaryotic organisms.

Progress toward this milestone is **modest**, with significant research gaps remaining.

Bacterial RNA regulators use molecules and structures, such as small transcription activating RNA (STARs) and toeholds, to regulate transcription and translation. Porting these regulators into eukaryotes is not necessarily direct and can require some modification to achieve similar efficacy as in prokaryotes. Research since 2019 has focused on how to increase the fold-change expression of different bacterial regulators in a wide variety of eukaryotic organisms. There are a few studies which begin to show promise in the porting of these technologies to this 10-fold change benchmark. For example, *Finke et al.* (2021) engineered a series of tetracycline-induced synthetic riboswitches to control gene expression for human cell cultures and *C. elegans*, achieving up to 16.9-fold higher expression with some of their constructs. In another instance, *Takahashi and*

Yokobayashi (2019) used a riboswitch-controlled vesicular stomatitis vector to repress as much as 26.8-fold gene expression in mammalian cells. And *Oesinghaus and Simmel* (2021) achieved well above 10-fold higher changes in gene expression in mammalian cells through activation of Cas12a guide RNAs by a strand displacement circuit (or "mechanism"). There are also numerous research articles and reports that do not strictly meet the 10-fold benchmark requirement for porting bacterial RNA-based gene regulators to eukaryotic systems yet represent forward and significant progress towards this milestone. Although there have been some reports of success with more complex model organisms, the successful porting of these technologies has largely been limited to cell cultures. The field would benefit from systematic studies to understand why some presumably generalizable bacterial systems fail in eukaryotic cells.

2021 Milestone: Creation of RNA modification machinery that allows programmable site-specific modifications of RNA, focusing on naturally abundant modifications (N6-methyl adenosine, 2'-O-methylation, pseudouridine).

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

RNA modifications, including N^6-methyladenosine, 2'-O-methylation, and pseudouridine, can strongly regulate transcriptional processes. The solo or combinatorial contribution of these modifications with other forms of RNA-regulator technologies can greatly improve the ability of engineering biology practitioners to create systems that perform precise molecular functions. As such, research since 2019 has investigated mechanisms that allow the incorporation of RNA modifications into different biological systems. For instance, *Liu et al.* (2019) developed an N^6-methyladenosine (m^6A) modification tool using a fusion CRISPR-Cas9 and an m^6A methyltransferase protein, together with a PAMmer to target RNA, that can perform precise-reversible single-site RNA methylations. *Qu et al.* (2019) created a tool called leveraging endogenous ADAR for programmable editing of RNA (LEAPER), that uses short engineered RNAs that recruit enzymes to create precise adenine to inosine modifications in a broad spectrum of human cell types. Summarily, there have been several positive instances of developing new forms of RNA technologies that can deliver precise modifications at will. Further progress can include new modification types and further examine how these modifications can transitively add to other RNA regulators for precise fine tuning.

Host Engineering | Host and Consortia Engineering

The engineering of host cells, organisms and systems, and consortia is fundamental to most applications of engineered biology. Host and Consortia Engineering ("Host Engineering") is defined in *Engineering Biology* as "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." The roadmap predicted that there was a "wealth of potential" to harness the ability of traditional and new model organisms to engineer useful functions; this observation remains true with several platform technologies emerging in the past two years, including the use of integrases to perform genome modifications in non-model microbes. Outside of the original prognostications of the roadmap, technologies foundational to Host and Consortia Engineering have developed, including standards for culturing consortia of microbes and guides for non-model organism domestication.

Progress in Host and Consortia Engineering	
Goal: Cell-free systems capable of natural and/or non-natural reactions.	
Breakthrough Capability: Ability to build reproducible and comparable cell-free systems for practical applications in bioengineering and biomanufacturing from multiple organisms, including non-model hosts.	
●●●○	2021 Milestone: 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).
Breakthrough Capability: Ability to build a cell, including the molecular subsystems that enable the processes of DNA replication, transcription, translation, energy regeneration, and membrane construction.	
●●●○	2021 Milestone: Demonstrated ability to synthesize all components encoded by a minimal or synthetic cell using cell-free systems.
Breakthrough Capability: Long-lasting, robust, and low-cost cell-free system for protein synthesis and biomanufacturing.	
●●●○	2021 Milestone: Identify reagent instabilities in cell-free systems across multiple organisms and all biological kingdoms.
Breakthrough Capability: Ability to use cell-free systems to inform cellular design of genetic parts and circuits.	
●●●○	2021 Milestone: Ability to use next-generation sequencing read-outs to quantitatively map performance of genetic designs in cell-free systems.
Breakthrough Capability: Decentralized, portable, on-demand sensing and manufacturing using cell-free systems.	
●●●●	2021 Milestone: Ability to use safe lysates low in endotoxin for sensing and manufacturing objectives.
Breakthrough Capability: Ability to manufacture any targeted glycosylated protein or metabolite using cell-free biosynthesis.	
●●●●	2021 Milestone: Ability to build modular, versatile cell-free platforms for glycosylation pathway assembly.

(Table continues)

Goal: On-demand production of single-cell hosts capable of natural and non-natural biochemistry.	
Breakthrough Capability: Ability to grow any host, anytime, in a controlled and regulated setting.	
●●●●	2021 Milestone: Establish protocols for the development of media that support cellular viability for non-model organisms.
●●●○	2021 Milestone: Robust screening of useful hosts beyond model organisms.
Breakthrough Capability: Routine domestication of non-model organisms through DNA delivery and genetic modification.	
●●●○	2021 Milestone: 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).
●●●○	2021 Milestone: Develop high-throughput methods that can be done in parallel for DNA delivery (using standard methods) into non-model hosts.
●●●○	2021 Milestone: Establish a suite of gene-editing tools for the rapid insertion and/or deletion of genetic elements in diverse primary mammalian cells.
●●●●	2021 Milestone: Characterize basic DNA parts for expression strength in non-model organisms, specifically a larger library of plants.
Breakthrough Capability: Ability to build and control small molecule biosynthesis inside cells by design or through evolution.	
●●●●	2021 Milestone: Identify model organisms for performing specific types of chemistries or organisms that have native precursor biosynthesis pathways for specific classes of molecules.
●●●○	2021 Milestone: Precise temporal control of gene expression for well-studied systems.
Breakthrough Capability: Spatial control over, or organization of, metabolic pathways in cells and construction of unnatural organelles.	
●●●●	2021 Milestone: Tools to target heterologous proteins to various subcellular compartments.
Breakthrough Capability: Production and secretion of any protein with the desired glycosylation or other post-translational modifications.	
●●●●	2021 Milestone: One or more microbial hosts capable of producing laboratory-scale quantities of a single glycoform of a desired protein.
Goal: On-demand fabrication and modification of multicellular organisms.	
Breakthrough Capability: Ability to control differentiation and de-differentiation of cells within a population.	
●●●●	2021 Milestone: On-demand, reproducible functionalization of simple micro-tissues or micro-consortia made up of two or more engineered cell types.

(Table continues)

Goal: On-demand fabrication and modification of multicellular organisms. *(Continued)*	
Breakthrough Capability: Ability to characterize and control the three-dimensional (3D) architecture of multicellular systems.	
●●●○	2021 Milestone: Characterize existing tissue components and standardize measurements to evaluate function.
Breakthrough Capability: Ability to achieve stable non-heritable changes in somatic cells.	
●●●○	2021 Milestone: Routine delivery of biomolecule "effectors" (i.e., DNA, RNA, proteins) into slowly-dividing or non-dividing cells.
Breakthrough Capability: Ability to make predictable and precise, targeted, heritable changes through germline editing.	
●●●●	2021 Milestone: Complete sequence of select host genomes to allow design of targets for gene editing.
●●●○	2021 Milestone: Define and validate tissue-specific DNA parts in plants.
Goal: Generation of biomes and consortia with desired functions and ecologies.	
Breakthrough Capability: Ability to control cell-to-cell communication between different species.	
●●○○	2021 Milestone: Tightly-controlled promoter-response regulator systems that enable intra- and inter-species cellular communication.
Breakthrough Capability: 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).	
●●○○	2021 Milestone: 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.
Breakthrough Capability: Ability to control and/or define the function of an engineered microbial community/biome.	
●●●●	2021 Milestone: Ability to combine species with specialized functions to enable the production of desired products.
Breakthrough Capability: 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.	
●●○○	2021 Milestone: Use of existing technologies (including metagenomics, transcriptomics, proteomics, and mass spectrometry) to characterize functions of microbial communities from a broad range of environments.

Table 3. Assessment of Host Engineering 2021 Milestone Achievement. Each 2021 milestone was assessed to determine progress towards its achievement. Four filled circles indicates the 2021 has been achieved or is close to complete, three filled circles indicates significant progress towards the 2021 milestone, two filled circles indicates modest progress towards the 2021, and one filled circle indicates only minimal progress towards achieving the 2021 milestone. In Host Engineering, the 2021 milestones have

been achieved or are close to complete (four filled circles), or have seen significant (three filled circles) or modest progress (two filled circles) towards their achievement.

Highlights of Technology Developments in Host Engineering

Chassis-independent Recombinase-Assisted Genome Engineering (CRAGE)

Identifying, modifying, and testing the ability of different microbes to produce valuable biomolecules and metabolites is essential for the commercialization and scale-up of engineering biology research. Our inability to rapidly screen and study several microbes simultaneously for their capacity to produce useful biomolecules has been recognized as a significant barrier. Based on the power of bacteria to horizontally-transfer genetic information, chassis-independent recombinase-assisted genome engineering (CRAGE) is the single-step integration of large biosynthetic gene clusters into the genomes of bacteria with high efficiency and accuracy (Wang et al., 2019). These biosynthetic gene clusters often produce secondary metabolites that are unnecessary for a microbe's survival but can give the microbe a competitive advantage in the face of environmental or induced pressures. Microbes vary in their ability to synthesize products from these biosynthetic gene clusters, thus there is a need to identify which microbes are most tolerant and best able to efficiently express these clusters in industrial settings. Further, this tool could enable researchers to introduce biosynthetic gene clusters to a larger variety of microbes to perform screening and comparative studies on biomolecular synthesis.

Use of Integrases to Edit Genomes in Non-Model Organisms

Since 2019, the growing enthusiasm for the domestication of non-model organisms has included those that natively exhibit phenotypes well-equipped for industrial purposes. Often, these non-model systems lack the toolsets that make their study amenable. Particularly missing is chromosome modification machinery to create useful genetic mutations in non-model organisms. Although new advances in genome editing technology, such as CRISPR-Cas9, show promise for working across a myriad of species, many organisms face obstacles in being able to port it rapidly. A remedy over the past few years has emerged with site-specific DNA integrases. These integrases operate by catalyzing a recombination event between two specific DNA sequences and can often be adapted towards different species. Certain families of integrases, such as large serine recombinases and serine integrases, can function on a broad range of organisms because they do not require molecular machinery from the host to perform the recombination event. This relaxed requirement enables integrases to more readily modify the genomes of non-model organisms, such as *Pseudomonas putida* (Martin-Pascual, 2021). Although there is a sustained need to develop universal toolkits that enable genome editing across *any* desired species, integrases have allowed some headway for researchers to begin to probe non-model organisms.

Using Cellular Fusion Techniques to Create Useful Hybrid Host Organisms

Although the roadmap frequently mentioned the fusion of proteins as an essential tool to advance engineering biology, there was no specific focus on fusing of *cells* from disparate organisms. Since 2019, researchers have investigated the fusion of cells to create a hybrid organism with an admixture of valuable properties from both organisms. Researchers have focused mainly on fusing cells that contain similar biosynthesis pathways that, when combined, can synergistically produce a desired molecule. Recent discoveries highlight some exciting findings involving engineering biology through cellular fusion events. For example, *Foster et al.* (2021) created a dynamic genome-scale metabolic modeling framework that evaluates the changes in properties in a fusion event between *Clostridium acetobutylicum* and *Clostridium ljungdahlii*, and with their model, forecasted improvements with ethanol and isopropanol yields. In their preprint, *Shitut et al.* (2021) described a protocol for generating heterokaryotic cells through bacterial cell-cell fusion, demonstrating control over the specificity of cell fusion events through synthetic membrane-associated lipopeptides. Finally, *Ding et al.* (2021) described a fusion between yeast spheroplasts and mammalian BHK-21 cells to recover *Sindbis* virus particles, an ordinarily complicated procedure that uses expensive laboratory reagents. Cellular fusion offers an intriguing possibility

to combine the different properties of organisms, adding a tremendous capability in the toolkit for researchers to modify organisms. However, understanding which organism pairs are amenable to cell fusion events and predicting if a hybrid organism would have the necessary biomolecules to function correctly are essential pieces of knowledge that researchers will have to investigate to make the most use of this technique.

Host Engineering Barriers to Progress

Tools and Procedures for Non-Model Organism Domestication

Many of the Breakthrough Capabilities found in the Host and Consortia Engineering technical theme have seen significant progress towards their achievement. One area with less progress is the ability to "grow any host, anytime, in a controlled and regulated setting." The roadmap anticipated that several advances would be needed to domesticate non-model organisms for different uses in research and application. Before advanced metagenomic characterization or genomic editing can take place for non-model organisms, many precedent characterization technologies, protocols, and techniques needed to be developed. As a result of this bottleneck, unanticipated advancements have been realized in fundamental analytical methods and resources, including in basic microscopy, cytology, and compatible tool development, such as plasmid creation. Not only can these advancements be applied to existing model organisms, but they are also well-suited to be adapted to domesticate non-model systems. Subsequently, many labs have risen to the challenge of domesticating new species, and several research groups have included fundamental characterization analyses in their published reports. However, the headway towards growing and engineering *any* organism – model species or newly discovered – has been slower than anticipated.

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

In place of using intact cells, cell-free systems use components derived from cellular extracts and lysates. These *in vitro,* non-living systems offer increased flexibility and control for researchers to harness biological components to perform modular tasks, including the manufacture of proteins and small molecules that are toxic to living cells, and enable rapid and high-throughput prototyping of biological parts. Optimizing cell-free systems' productivity could enable the rapid and sensitive detection of pathogens, human health biomarkers, and environmental contaminants. Because of their flexibility and independence from some constraints of typical biological organisms, cell-free systems also possess the opportunity for on-demand manufacturing of proteins, nucleic acids, and small molecule therapeutics and vaccines for more widespread deployment of lifesaving medicines. And by capturing these capabilities, we further enable and build on our capacity for bottom-up construction of wholly synthetic cells.

Breakthrough Capability: Ability to build reproducible and comparable cell-free systems for practical applications in bioengineering and biomanufacturing from multiple organisms, including non-model hosts.
This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached. Of note, the Assessment suggests that the 2029 milestone "Complete library of user-defined reaction components for use in a customizable cell-free system" may be achieved ahead of schedule.

2021 Milestone: 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).

Progress toward this milestone is **significant**, with some research gaps remaining.

The behavior of cell-free systems can depend highly on how they are generated. Since 2019, research has begun to dissect some of the critical harvesting features that significantly affect system behavior and how to identify variability in performance between laboratories. For example, *Cole et al.* (2019) measured the variability of synthesized cell-free systems across different laboratories, uncovering that the laboratory site

where the system was prepared, the operators conducting the synthesis, and the process and conditions of reagent preparation, all contributed significantly to variability. Reaction geometry is shown to strongly affect cell-free protein productivity. *Sakamoto et al.* (2018) showed that larger surface-to-volume ratios of cell free system in emulsion droplets ranging from 10-100μm exhibit negative effects on protein production, demonstrating that confinement alone can alter the yield of cell-free expression. And *Rasor et al.* (2023) used multiomics tools to study the impacts of extract preparation on gene expression and production of proteins and metabolites. Several publications went beyond understanding variabilities to remedies and solutions for many of these issues. *Silverman et al.* (2019) determined that the normally constrained expression of genes from the bacterial σ70 promoter can be alleviated with ribosomal runoff reactions followed by dialysis, offering a generalized view of how downstream extract procedures can impact performance. *Hershewe et al.* (2021) investigated the impact of extract preparation protocols on the activity of exogenous enzymes expressed in cell-free extracts, finding that different lysis methods resulted in different concentrations and sizes of inverted membrane vesicles. Towards the informed creation of more standardized systems, *Miguez et al.* (2019) outlined a novel approach using metabolomics to calibrate performance and inform system design and in follow-up work, further investigated different extract preparation settings greatly affect the metabolic profile of cell-free systems (Miguez et al., 2021). Furthermore, *Contreras-Llano* et al. (2020) showed that the proteome of cell extract can be reprogrammed via implementing genetic circuits in host strain to improve productivity. *Garcia et al.* (2021) further showed that extract proteome can be optimized for metabolite production via selective removal of enzymes in competing pathways. An exhaustive understanding of *all* of the general effects that can affect cell-free system behavior remains to be undertaken, but the past two years have demonstrated remarkable headway in identifying several causal factors and solutions to remedy variability issues. More research can continue to identify factors contributing to variability and markers that practitioners can use in calibration and adjustment.

Breakthrough Capability: Ability to build a cell, including the molecular subsystems that enable the processes of DNA replication, transcription, translation, energy regeneration, and membrane construction. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached.

2021 Milestone: Demonstrated ability to synthesize all components encoded by a minimal or synthetic cell using cell-free systems.

Progress toward this milestone is **significant**, with some research gaps remaining.

Building a synthetic cell requires the essential molecular components to perform processes such as DNA replication, transcription, translation, energy regeneration, and membrane construction. There have been several good demonstrations toward the construction of cell-free systems that demonstrate several components encoded by minimal or synthetic cells, yet a cell-free system that can synthesize *all* the components encoded by a minimal or synthetic cell remains to be reached. In a few pertinent examples of this capability, *Eto et al.* (2022) demonstrated fatty acid synthesis in cell-free systems and incorporation into a mother membrane, paving the way for synthetic cell membrane growth and division, *Garenne et al.* (2020) developed a new method for regulating the shape of synthetic cells, and *Berhanu* et al. (2019) generated artificial cells capable of synthesizing ATP from light. Some of these developments have focused on identifying promising biological material that could be used as potential chassis for potential cell-free systems. In their preprint, *Wei et al.* (2020) described how subsets of isolated mincells (anucleate cells devoid of heritable genetic material but capable of gene expression) have enough gene expression capacity to replicate known prokaryotic proteomes. Researchers have also created toolsets and other platforms towards achieving this milestone. For example, *Karim et al.* (2020) developed modular plasmids that facilitate the expression of candidate-desired enzymes in cell-free systems, demonstrating a streamlined framework for testing biosynthetic pathways *in vitro.* Relevant to minimal cells, *Rees-Garbutt et al.* (2020) developed algorithms that

allow users to design-test-build-learn cycle minimal genomes, and in the process, uncovered candidate minimal genomes of the bacterium *Mycoplasma genitalium*. Additionally, *Lavickova et al.* (2020) and *Wei and Endy* (2021) developed a framework for evaluating the capability of a cell-free system to functionally regenerate life-essential activity using the model cell-system Protein synthesis Using Recombinant Elements (PURE); *Libicher et al.* (2020) also used the PURE system to achieve self-encoded DNA replication of more than 116kb, exceeding the length of the smallest know bacterial genome (112kb).

Breakthrough Capability: Long-lasting, robust, and low-cost cell-free system for protein synthesis and biomanufacturing. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached. Interestingly, the Assessment suggests that the 2039 milestone "Robust and scalable production of cell-free systems that last for weeks" may be achieved ahead of schedule.

2021 Milestone: Identify reagent instabilities in cell-free systems across multiple organisms and all biological kingdoms.

Progress toward this milestone is **significant**, with some research gaps remaining.

Although the most common types of cell-free systems are derived from *Escherichia coli*, researchers are interested in how other biological organism extracts can infer a broader array of properties beneficial for producing desired biomolecules and engineering biological functions. A significant issue with this is that the factors affecting reagent stability of non-*Escherichia coli*-based cell-free systems are less understood relative to their *Escherichia coli* counterparts. This has led a number of research labs to explore and optimize non-*Escherichia coli* systems (see for example Yim et al., 2019 and Zhang et al. 2020b). Additional research prior to 2019 suggests a promising opportunity for production of stable cell-free translation systems employs (hyper)thermophilic bacteria and archaea which naturally grow at >60C and produce enzymes that can be stable for years (Zhou et al., 2012; Uzawa et al., 1993; Endoh et al., 2006). In research published since 2019, further efforts have worked toward providing a valuable set of characteristics for researchers to troubleshoot when creating non-*Escherichia coli*-based systems. For example, *Vezeau and Salis* (2021) analyzed how macromolecular crowding reagents and salts control the time delay, dynamics, and productivity of *in vitro* transcription and translation of cell-free systems. *Miguez et al.* (2021) also investigated endogenous metabolism in cell extract proceeds independently of active gene expression and drain available "energy" in cell-free system. Furthermore, *Lee et al.* (2020) revealed that lyophilized cell-free lysate systems exhibit increased tolerance to various organic solvents. There have also been promising discoveries into different procedures that can prolong the activity of cell-free systems, such as *Gregorio et al.* (2019), who identified unique additive formulations that can stabilize lyophilized *Escherichia coli* extracts for a longer shelf life at room temperature. *Guzman-Chavez et al.* (2022) were able to develop a low-cost cell-free system by modifying additives and developing a drying process not dependent on lyophilization. These studies have identified several major factors affecting reagent stabilities, yet more exhaustive research of further instabilities that affect cell-free systems *derived from any member of the biological kingdom* remains to be undertaken. Other efforts are needed to characterize commonly problematic species and determine some of the unifying causal factors affecting reagent instability.

Breakthrough Capability: Ability to use cell-free systems to inform cellular design of genetic parts and circuits. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates that significant progress has been made toward the 2021 and the 2024 milestone "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" and the 2029 milestone "Ability to identify new genetic circuits in cell-free systems for any bacterial host to facilitate forward engineering in cells" may be achieved ahead of schedule.

2021 Milestone: Ability to use next-generation sequencing read-outs to quantitatively map performance of genetic designs in cell-free systems.

Progress toward this milestone is **significant**, with some research gaps remaining.

In the past two years, research has focused on producing and characterizing next-generation sequencing data on many cell-free system activities and functions, including transcription and translation, and creating workflows, platforms, and pipelines to use these data to inform cell-free system design. For example, through a series of DNA synthesis and multiplexed reporter assays, Park et al. (2021) developed a *Streptomyces albidoflavus* cell-free expression system that can rapidly characterize regulatory sequences affecting biosynthetic gene cluster expression, useful for identifying features that regulate the synthesis of biomolecules. Yim et al. (2019) described a robust *in vitro* approach, DNA Regulatory element Analysis by cell-Free Transcription and Sequencing (DRAFTS), to multiplex measurements of transcription activity from regulatory sequences for extracted cellular lysates. Lashkevitch et al. (2020) developed a C-terminally extended luciferase-based system (CTELS) to assay translation termination events in protein biosynthesis, uncovering how 3' UTR and inhibitors affect the release of polypeptide release from the ribosome. Finally, Marshall and Noireaux (2019) created an ordinary differential equation (ODE)-based model to evaluate how transcription and translation rates are affected by complex regulatory networks. Horvath et al. (2020) have also built a sequence-specific dynamic model of cell-free protein synthesis in *Escherchia coli* extract and found that protein synthesis was only 12% energy efficient and uncovered other key pathways affecting system productivity. While modeling approaches have improved substantially to predict performance, there is not overwhelming evidence that we can quantitatively describe the performance of parts or circuits *in vitro*. Research that examines how other forms of genomics technology, such as Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq), can inform cell-free system design, and an overall greater number and breadth of well-parameterized parts, could enable further progress.

Breakthrough Capability: Decentralized, portable, on-demand sensing and manufacturing using cell-free systems. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached. As an indicator of significant progress in this area, the Assessment suggests that the 2024 milestone "Demonstrate portability (such as two-year storage of freeze-dried reactions without loss of functionality) of cell-free systems," the 2024 milestone "Increase productivity and rate of cell-free reactions," the 2029 milestone "Point-of-care cell-free protein production system ready for validation by the Food and Drug Administration (FDA)", and the 2039 milestone "Point-of-care cell-free protein therapeutic and vaccine production system ready for validation by the Food and Drug Administration (FDA)" may all be achieved ahead of schedule.

2021 Milestone: Ability to use safe lysates low in endotoxin for sensing and manufacturing objectives.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Endotoxins – lipopolysaccharides from the cell membrane of gram-negative bacteria – can drastically hinder the ability of cell-free systems to perform useful functions, just as they would in cell-based research. Since 2019, research has discovered various strategies for keeping cell-free systems robust, even in the presence of endotoxins and other pollutants, with much progress in creating and determining the most efficient methods to develop endotoxin-free cell-free protein systems. Wilding et al. (2018) evaluated three different pre-expression endotoxic removal strategies for *Escherichia coli*-based systems, demonstrating that cell-free extract generation from ClearColi cells was able to clear endotoxins while retaining high synthesis capabilities. As a follow-up to this article, Hunt et al. (2019) further streamlined the protocols for adapting ClearColi cells for cell-free systems by using autoinduction media, producing time-efficient high yields of the FDA-approved therapeutic protein crisantaspase as a demonstration. Along with the removal of endotoxins, the stabilization of endotoxin free cell-

free reactions has been achieved through lyophilization. Lyophilization, the removal of water from a frozen product under vacuum, can often better preserve biological molecules but comes at the cost of adversely affecting cell-free systems. *Wilding et al.* (2019) demonstrated an antiplasticized sugar glass lyoprotected, lyophilized cell-free protein system is superior to traditional lyophilization methods for preserving system activity. *Guo et al.* (2020) also identified the protective role of metal cofactors on enzyme activity in lyophilized transcription-translation systems. Cell-free translation systems from archaea are promising avenues for development as archaea are not pathogenic, do not produce endotoxin, and thermophilic archaea are already used to produce heat, salt, and pH-stable enzymes for industrial use (Ruggero et al., 1993; Endoh et al., 2006). Further research that examines novel ways to streamline or automate these processes is likely to drive further progress in this space.

Breakthrough Capability: Ability to manufacture any targeted glycosylated protein or metabolite using cell-free biosynthesis. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap and the Assessment literature review indicates the 2021 milestone has been reached. The Assessment suggests that the 2024 milestone "Production of bacterial glycoconjugate vaccines in cell-free systems" and 2029 milestone "Expanded set of enzymes capable of glycosylating metabolites *in vitro*" may be achieved ahead of schedule.

2021 Milestone: Ability to build modular, versatile cell-free platforms for glycosylation pathway assembly.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Glycans, which are complex sugar moieties, can be added to amino acid side chains during enzyme-catalyzed protein modification. Protein glycosylation is a key post-translational modification and can significantly alter protein stability, immunogenicity, and protein activity. Therefore, the ability to manufacture glycosylated proteins has been identified as incredibly important in engineering biology, especially in producing therapeutic biomolecules. Since 2019, research has focused on creating cell-free systems that can recapitulate the glycosylation pathway by creating new tools to synthesize these proteins. *Kightlinger et al.* (2019) developed Glycosylation Pathway assembly by Rapid *In vitro* Mixing and Expression (GlycoPRIME), a cell-free biosynthesis platform that allows modular construction of protein glycosylation pathways, demonstrated by an impressive construction of 37 putative protein glycosylation pathways and creation of 23 unique glycan motifs. *Aquino et al.* (2021) created the microfluidic platform, Glycosylation-on-a-Chip, that can be used for the mechanistic dissection of protein glycosylation pathways and for small-batch glycoprotein manufacturing. These platforms' success and similar efforts have yielded patents on glyco-production technology, such as *Jewett et al.* (2021)'s patent for the recombinant production of N-glycosylated proteins using prokaryotic cell lysates. There have also been systematic efforts to identify strategies that increase the efficiency of existing glycoprotein systems. *Warfel et al.* (2023) were able to construct low-cost thermostable cell-free reactions using maltodextrin as both a lyoprotectant and energy source, causing reactions to be reduced from $5/mL using PEP systems to under $2/mL with the minimal maltodextrin system; they were able to then produce bactericidal antibodies using this cheap thermostable system. For example, *Hershewe et al.* (2021) characterized and described enrichment processes for native membrane vesicles in *Escherichia coli*-based cell-free expression systems, improving the synthesis of N-linked and O-linked glycoproteins as a demonstration. *Stark et al.* (2021) successfully manufactured a glycoconjugate vaccine against *Francisella tularensis* using a cell-free system. Researchers have made remarkable progress toward this milestone with numerous platforms and strategies to enrich the production of glycoproteins; further research efforts can investigate new strategies to synthesize these proteins and focus on how to scale up production.

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

One of the scientific challenges of engineering biology is creating and transforming organisms for valuable functions and solving engineering obstacles such as production and scale-up. Producing organisms efficiently enables researchers to employ their resulting technologies to increasingly complex societal applications. For example, completing this goal could allow the synthesis of many genes and regulatory components required to create enzymes and cells to degrade biomass and process by-products, likely major components of sustainable bioprocesses. Likewise, increasing the production of microbes capable of ammonium oxidation, denitrification, and polyphosphate accumulation can enable efficient wastewater fermentation for the safe remediation of environmental contaminants; microbial remediation can be used to transform wastes and contaminants into renewable fuels and fertilizers that further promote sustainable water, energy, and agriculture. And, as a component of healthcare, achieving this goal would contribute to increased production of cell-expressed reporters for rapid, reliable diagnostics to detect viral infections.

Breakthrough Capability: Ability to grow any host, anytime, in a controlled and regulated setting. Overall, this Breakthrough Capability is close to meeting the pace of the predictions relative to the roadmap. The Assessment literature review indicates that the 2021 Milestone "Establishing protocols for the development of media that support cellular viability for non-model organisms" has largely been achieved, though greater dissemination of existing repositories of media compositions would benefit the engineering biology research community. Significant progress has been made toward the other 2021 milestones.

2021 Milestone: Establish protocols for the development of media that support cellular viability for non-model organisms.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

There has been significant development of platforms and databases for researchers to understand viability requirements for organisms. One challenge in non-model organism domestication is overcoming limitations to achieving fluxes, syntrophy, quorum sensing, and crossfeeding interactions that are virtually impossible to predict; however, even these limitations are close to being overcome (Oberhardt et al., 2015 and see e.g., Imachi et al., 2022). Many databases pertaining to the informed domestication of a wide variety of organisms have been developed and maintained, such as the Synthetic Biology Knowledge System (SBKS), Tripal, BioMaster, and NanDeSyn (Mante et al., 2021, Staton et al., 2021, Wang et al., 2021b, and Gong et al., 2020, respectively). In their preprint article, *Kailash et al.* (2019) developed an open-source, nonprofit organism database, ChassiDex, to gather a repository of massive amounts of data for synthetically created organisms, including maintenance, transformation protocols, vectors, and BioBrick parts. There has also been a focus on examining how environmental conditions affect the growth of specific organisms for relevant industrial biotechnology applications. Phenotype microarrays, such as those from BioLog, have long been used to characterize aerobic bacteria and fungi; however, more research is needed to develop similar high-throughput phenomics for non-model anaerobes and archaea (Walter et al., 2016; Cashman et al., 2017). *Catlett et al.,* (2020) uncovered metabolic feedback inhibition resulting in changes to metabolic flux in *Bacteroides*, and in another study from the same group, successfully used phenotyping and machine learning to discover complex metabolic cross-feeding and syntrophic interactions between human symbionts *Bacteroides* and *Methanobrevibacter* (Catlett et al., 2020). For example, *Burdette et al.* (2021) elucidated the effect of growth medium components on secretion titers via the type III secretion system in *Salmonella enterica*, discovering an optimized growth medium for increased recombinant protein secretion. *Wilken et al.* (2020) developed an automated Arduino-based automatic pressure evaluation system to quantify the growth of non-model anaerobes in culture. And quite recently, *Imachi et al.* (2022) enriched *Lokiarchaea* using a continuous-flow down-flow hanging sponge bioreactor. Additionally, in a nod to biodiversity and sustainability efforts, *Roger et*

al. (2021) examined a series of models and methods to maximize the culture formation of corals. While cultivation of non-model organisms has significantly advanced, what is still lacking is the ability to design synthetic biology and metabolic engineering strategies in non-model organisms.

2021 Milestone: Robust screening of useful hosts beyond model organisms.

Progress toward this milestone is **significant**, with some research gaps remaining.

Although traditional model organisms, such as *Escherichia coli*, are powerful and flexible in their ability to be engineered, undomesticated species encompass vast potential for efficiently producing valuable biomolecules. Since 2019, potent methods, especially those using data science approaches, have enabled the identification of promising bacteria and microbes for industrial settings, though these tools still need to be deployed towards diverse clades to identify promising biological organisms. Since 2019, research has focused on different selection and analysis strategies to discover beneficial new model organisms to create certain classes of biomolecules. To identify promising candidate organisms, *Yim et al.* (2019) developed the DNA Regulatory element Analysis by cell-Free Transcription and Sequencing (DRAFTS) approach to perform interspecies analysis of transcriptional profiles from bacterial regulatory sequences to better understand which hosts serve as new chassis. Additionally, *Gilman et al.* (2019) developed a toolset to rapidly discover and design practical promoter sets for atypical microbial organisms for industrial applications, uncovering several useful promoters for *Geobacillus thermoglucosidasius*. There has also been a wealth of research devoted to identifying, selecting, and recovering valuable organisms (see e.g., Imachi et al., 2022). Multiple microfluidic droplet screening platforms have been developed to cultivate and screen microbes in complex communities (Kehe et al., 2019; Watterson et al., 2020). *Gilmore et al.* (2019) used metagenomic sequencing as part of a top-down enrichment guide to select communities of microbes and create co-cultures that can digest lignocellulose to produce methane-rich gas. Although demonstrated in model organism *Escherichia coli, Meksiriporn et al.* (2019) described a genetic selection strategy to isolate post-translationally phosphorylated proteins that can be employed in other organisms. And *Meng et al.* (2021) presented a powerful approach that leverages integrated conjugated elements in *Bacillus subtilis* XPORT strains to screen bacterial consortia for strains amenable to pathway engineering functions. Future research can apply existing methods and develop additional screening tools, to further progress toward this milestone.

Breakthrough Capability: Routine domestication of non-model organisms through DNA delivery and genetic modification. This Breakthrough Capability is likely proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates that research toward the 2021 milestones has been significant; the 2021 milestone "Characterize basic DNA parts for expression strength in non-model organisms, specifically a larger library of plants" has been achieved. The 2024 milestone "Establish robust temporal and/or spatial control of gene expression in mammalian cells" and 2029 milestone "Develop high-throughput, targeted editing and rapid genome-evolution tools that couple genetic changes to phenotypic changes" may be achieved within a shorter time period than anticipated by the roadmap.

2021 Milestone: 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).

Progress toward this milestone is **significant**, with some research gaps remaining.

Species can vary significantly in their ability to tolerate different DNA delivery technologies, necessitating cataloging techniques that maximize a researcher's ability to perform DNA editing with any given organism. Many technologies for delivering DNA have been well-described in molecular biology for some time. Innovations in this space include adapting tools to new organisms, multiplexing, and combining reporters and sensors. Since 2019, research has created database repositories that catalog DNA delivery technologies across several

species. For example, Bernabé-Orts et al. (2019) assessed the genomic editing efficacy of *Acidaminococcus, Lachnospiraceae,* and *Streptococcus* CRISPR-Cas12a variants across several plant model species (including *Nicotiana benthamiana, Solanum lycopersicum,* and *Arabidopsis thaliana*). Rubin et al. (2020) presented a generalizable strategy for editing select atypical microbial genomes using a combination of environmental transformation sequencing (ET-seq) to identify microbial candidates and a DNA-editing all-in-one RNA-guided CRISPR–Cas transposase (DART) to carry out the genome edits. Likewise, there have been several efforts to catalog mammalian CRISPR-Cas9 editing tools for engineering biology applications to generate knockout strains, increasing transgene expression, and gene silencing; see Giulano et al. (2019), Zhan et al. (2020), and He et al. (2020) as examples. Several inclusive platforms documenting aspects of DNA transformation for engineering biology purposes have emerged, though a consensus platform has not yet been determined. In their preprint article, Kailash et al. (2019) developed an open-source, nonprofit organism database, ChassiDex, to gather a repository of massive amounts of data, including maintenance, transformation protocols, vectors, and BioBrick parts, for synthetically created organisms. Likewise, similar databases on the informed domestication of a wide variety of synthetic organisms are also being developed and maintained, such as the Synthetic Biology Knowledge System (SBKS), Tripal, and BioMaster (Mante et al., 2021, Staton et al., 2021, and Wang et al., 2021b, respectively). In summary, researchers have begun to rigorously examine and catalog how several DNA transformation technologies, including CRISPR/Cas9, work for engineering biology purposes. However, a *de-facto* consensus platform to encompass all these protocols and methodologies has not emerged. Consortium-led efforts to create, elevate, or retrofit these platforms to house these developments better can enable further research progress.

2021 Milestone: Develop high-throughput methods that can be done in parallel for DNA delivery (using standard methods) into non-model hosts.

Progress toward this milestone is **significant**, with some research gaps remaining.

Among the core toolsets for researchers across engineering biology is being able to, at high-throughput, edit the genomes of multiple organisms to create valuable properties for myriad applications. Species can vary in their ability to uptake DNA from genome editing technologies, making it essential for engineering biology researchers to discover compatible high-throughput methods of delivering DNA to atypical, non-model hosts. Introducing DNA into cells and integration into host genomes is routine – *Agrobacterium* and *Escherichia coli* have been used for many years to transfer plasmids across domains. Several platforms have emerged in the past two years that enable high-throughput DNA delivery. For example, Wang et al. (2019) described the use of chassis-independent recombinase-assisted genome engineering (CRAGE) to enable single-step integration of biosynthetic gene clusters into diverse groups of bacteria. Along with delivering DNA at high throughput, this platform also allows researchers to screen for bacterial hosts that can synthesize useful metabolites. Brophy et al. (2018) engineered an integrative and conjugative element from *Bacillus subtilis* (ICE*Bs1*) to work with a donor strain (XPORT) that can facilitate the transfer of DNA to undomesticated bacteria, demonstrating 10^{-1} to 10^{-7} conjugation events per donor on over thirty strains. And Demirer et al. (2019) chemically functionalized high-aspect ratio nanomaterials to efficiently deliver DNA in *Eruca sativa* (arugula), *Triticum aestivum* (wheat), and *Gossypium hirsutum* (cotton), enabling a strategy for species-independent and passive delivery of DNA into plant cells. Several DNA delivery technologies have been cultivated specifically for organism clades. For example, Swafford et al. (2020) developed a procedure that efficiently delivers high molecular payloads to lesser-known, parasitic chytrid fungi that do not have many basic molecular genetic tools available. Further, Poliner et al. (2020) designed an extensive vector toolkit and screening strategy for the oil-accumulating microalgae *Nannochloropsis oceanica* CCMP1779, enabling the combinatorial expression of transgenes in a practical protist chassis. Rather than DNA delivery, limitations persist for transformations in non-model hosts in coding/decoding, regulation, replication, epigenetic modifications, and innate immunity mechanisms.

Expanded use of bisulfite sequencing and engineering intermediate permissive hosts is needed, as in a recent example from *Riley et al.* (2019).

2021 Milestone: Establish a suite of gene-editing tools for the rapid insertion and/or deletion of genetic elements in diverse primary mammalian cells.

Progress toward this milestone is **significant**, with some research gaps remaining.

Primary cells are those taken directly from living tissue and established for growth *in vitro.* As these cells represent a state more similar to *in vivo* physiology compared with immortalized cell lines, recent work focuses on edits in these cell types to achieve high editing efficiency, minimal off-target editing events, and low amounts of toxicity as research aim to demonstrate genome editing technologies for health or medical applications. CRISPR-Cas9 technology has proven critical in developing gene-editing tools to engineer primary mammalian cells, with much of the focus centering on human cell lines. For example, *Hultquist et al.* (2018) created a streamlined, high-throughput, multiplex platform that can deliver CRISPR-Cas9 ribonucleoproteins to CD4+ cells through nucleofection. *Shahbazi et al.* (2019) used nano-formulations of CRISPR complexes and gold nanoparticles to create a monodispersed solution of genome editors that can localize to primary human hematopoietic cell nuclei and avoid the typical adverse effects of lysosomal entrapment and toxicity. *Sercin et al.* (2019) described a solid-phase transfection platform enabling CRISPR-based screens in primary human cells, including untransformed and cancer cell lines for potential targeted therapeutic strategies. To increase the efficacy of delivery into target cells, *Mangeot et al.* (2019) engineered murine leukemia virus-like particles loaded with Cas-sgRNA ribonucleotides (Nanoblades) to perform genome editing in primary human and mouse cell lines, demonstrating that it can be used for homology-directed repair or to mediate transcriptional regulation. Most of these gene-editing tools have focused on human primary cell types; although these cells represent some of the most significant opportunities for advancing human health, further efforts will be needed to examine if these technologies can edit other mammalian cell lines. Further steps to extrapolate these technologies to other forms of primary mammalian cells can drive additional progress in this space.

2021 Milestone: Characterize basic DNA parts for expression strength in non-model organisms, specifically a larger library of plants.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

DNA parts can come in many forms, including components that can precisely control gene expression, timing, and strength. This finely tuned gene expression can benefit many engineering functions, especially when an expressed gene needs to respond autonomously to a feedback mechanism or stimulus. Among the many organisms that make up the engineering biology portfolio, DNA parts for plants remain sparse and are needed to provide more efficient forms of regulatory control for a broader swath of engineered systems. Since 2019, several research efforts have created inclusive platforms to rapidly identify new DNA parts for plants or have created parts that enable sophisticated functions. For example, *Belcher et al.* (2020) leveraged regulatory systems from *Saccharomyces* to develop a library of activators, repressors, and promoters to modulate expression strength in plant model systems, validating their system in *Nicotiana benthamiana* and *Arabidopsis thaliana.* Towards parts that could aid strategies for abiotic stress resistance, *Yang et al.* (2021) designed osmotic-related and salt stress-inducible synthetic promoters for hybrid poplar trees. *Dudley et al.* (2021) created an automated workflow for DNA assembly and cell-free expression of plant proteins, creating a platform that rapidly accelerates typical design-build-test-learn cycles. And *Bernabé-Orts et al.* (2020) developed a whole-plant memory switch based on the bacteriophage φC31 site-specific integrase, creating a part that allows precise shifts between on-or-off transcriptional states on two genes of interest. DNA parts are critical for researchers to engineer more sophisticated and powerful functions for plants and non-model organisms, and the past two years of research have created numerous workflows, tools, and sophisticated parts for

practitioners to use. Further research can continue to develop sophisticated DNA parts and begin to apply these inventions in atypical plant hosts and systems.

Breakthrough Capability: Ability to build and control small molecule biosynthesis inside cells by design or through evolution. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review suggests that the 2021 milestones have been reached. Of note, the Assessment suggests that the 2029 milestone "Software and hardware for optimizing titer, rate, and yield of any product produced by any host" may be achieved ahead of schedule.

2021 Milestone: Identify model organisms for performing specific types of chemistries or organisms that have native precursor biosynthesis pathways for specific classes of molecules.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Organisms may have intrinsic characteristics that make them more apt for producing specific types of biomolecules or for performing particular functions. Identifying which organisms can complete these specialized forms of chemistries or have native precursor pathways to synthesize these molecules is a challenge. Since 2019, several tools, workflows, and platforms have emerged to make this identification easier. A presentation from the Agile Biofoundry demonstrated a more formalized process to identify new microbial hosts for industrial bioengineering and reduce the scale-up time for these organisms to be ready for industrial use (Dale & Guss, 2019). Similarly, *Gilman et al.* (2019) developed a data science toolset to rapidly discover and design functional promoter sets for atypical microbial organisms in industrial applications, uncovering several useful promoters for *Geobacillus thermoglucosidasius* as a demonstration. In addition, *Gilmore et al.* (2019) described an enrichment guide to select communities of microbial consortia that can digest lignocellulose to produce methane-rich gas, offering a top-down approach to creating co-cultures of useful communities. And *Yim et al.* (2019) described a robust *in vitro* approach, DNA Regulatory element Analysis by cell-Free Transcription and Sequencing (DRAFTS), to multiplex measurements of transcription activity from regulatory sequences for extracted cellular lysates. Several studies have also showcased the role that specific organisms can play in being able to produce valuable molecules. For instance, *Krüger et al.* (2020) created a cell-free system derived from *Clostridium autoethanogenum*, thereby greatly facilitating a researcher's ability to prototype genetic parts in an organism that can efficiently convert low-cost feedstocks (such as industrial flue gasses) into useful biobased products. *Gülck et al.* (2020) engineered *Nicotiana benthamiana* and *Saccharomyces cerevisiae* to acquire high and pure quantities of pharmaceutically relevant cannabinoids. There has been much progress in developing platforms to identify beneficial organisms for industrial applications or cultivating specific organisms for the intended use. Further research can continue to develop these tools and deploy them to expand the use of atypical model hosts in engineered biotechnologies.

2021 Milestone: Precise temporal control of gene expression for well-studied systems.

Progress toward this milestone is **significant**, with some research gaps remaining.

The production of some biomolecules is toxic to the hosts themselves; in these cases, controlling the timing of gene expression and production of these potentially toxic biomolecules can be incredibly valuable. Beyond toxicity concerns, controlling the timely expression of genetic systems can enable a myriad of more complex, sophisticated, and practical functions for many applications. Since 2019, several research advances have focused on how to design intricate expression systems in engineering biology applications. Coordinating the spatial expression of genes in synthetic constructs can sometimes be challenging due to the limited diffusion range of signaling molecules. *Kim et al.* (2019) provided one solution to this issue by generating coordinated oscillations through a positive feedback loop in microbial consortia to amplify and propagate cellular signals. Similarly, towards informed expression design, *Alnahhas et al.* (2019) used a series of differently-shaped microfluidic traps to manage co-cultures of *Escherichia coli* and examine their ability to intracellularly

communicate and control gene expression, producing a mathematical model that predicts how microfluidic device conditions affect intracellular communication ability. Studies have even begun in biotechnologies oriented towards health and medicine, such as *Israni et al.* (2021) who created a toolkit that enables the orthogonal control of gene expression in synthetic cellular systems based on human cells and, significantly, facilitated FDA-approved molecules as gene expression modulators. Finally, there have also been efforts to identify and screen for organisms with regulatory features amenable towards the precise synthesis of valuable biomolecules. *Wang et al.* (2019) described the use of Chassis-independent Recombinase-Assisted Genome Engineering (CRAGE) to enable single-step integration of biosynthetic gene clusters into groups of bacteria; along with delivering DNA at high throughput, this technology also allows the screening of organisms that can provide the timely or increased expression of genes required for biomolecule synthesis. There have been numerous efforts to more temporally- or spatially-control gene expression. Future efforts in this space can continue to build on these strategies and cultivate new methods that can provide finely-tuned mechanisms for producing valuable molecules or functions.

Breakthrough Capability: Spatial control over, or organization of, metabolic pathways in cells and construction of unnatural organelles. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached.

2021 Milestone: Tools to target heterologous proteins to various subcellular compartments.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Sometimes engineered bio-based machinery needs to be contained and function in a particular part of a cell or organelle; in these cases, careful mechanisms are generally designed to facilitate their safe passage from one area or organelle to another. Directing the localization of these components, generally through heterologous proteins, to various subcellular compartments is critical in performing localized engineered functions within a cell. The combination of tools to characterize sequences necessary for organelle targeting and strategies to mitigate current delivery issues provides a wide berth of disparate approaches that each provide progress towards this milestone. Since 2019, several developments have focused on creating data science tools to discover characteristics of sequences useful for directing proteins. For example, *Armenteros et al.* (2019) made TargetP 2.0, a machine learning software capable of detecting N-terminal sequence signals that direct peptides to various organelles, such as chloroplasts, mitochondria, and the secretory pathway. Other studies have focused on strategies specific to industrial application issues or delivery to particular cellular bodies. For instance, *Zelmer et al.* (2020) constructed biocompatible polymer vehicles that bypass nuclear pore complexes, usually a problematic entity that can prevent the effective transfer of chemo- or gene-based therapies in nuclei. Cytosolic expression of certain chemicals, such as norcoclaurine synthase, can be toxic to species such as *Saccharomyces cerevisiae* and restrict the production of valuable substances like (*S*)-reticuline; *Grewal et al.* (2020) alleviated this issue by discovering a strategy to target norcoclaurine synthase to the peroxisome efficiently. Additionally, *Li et al.* (2021) performed an exhaustive study to characterize how hexamer protein interactions and angles can dictate the morphology of bacterial microcompartments, a critical cellular body that often houses enzymes and proteins beneficial for energy and chemical production. While most of the above represents examples in eukaryotes, cutting-edge research is ongoing in metabolosome and virion nanomaterials engineering, which could be applied to breakthrough technologies for vaccines, biosensors, and microbiome engineering. Further research can build on these efforts and begin to identify useful, atypical processes for organelle targeting used by non-model organisms.

Breakthrough Capability: Production and secretion of any protein with the desired glycosylation or other post-translational modifications. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached. The Assessment suggests that the 2039 milestone "Ubiquitous control of post-translational modification (including

glycosylation of multiple sites with multiple sugars) in a diverse array of hosts" may be achieved ahead of schedule.

2021 Milestone: One or more microbial hosts capable of producing laboratory-scale quantities of a single glycoform of a desired protein.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Glycans, complex sugar moieties, can be added to amino acid side chains in glycosylation. Protein glycosylation plays a vital role in post-translational modification and can significantly alter protein stability, immunogenicity, and protein activity. Therefore, manufacturing glycosylated proteins has been essential to engineering biology, especially for producing therapeutic biomolecules. There has been much development of microbial *and* mammalian hosts to synthesize glycoforms of specific proteins for various purposes, with significant research focused on creating platforms to develop different glycoforms of proteins. For example, Du et al. (2021) engineered an *Escherichia coli* strain to produce O-glycosylated proteins using a plasmid system that dually contains glycosylation machinery (derived from *Campylobacter jejuni*) and components that can target cellular proteins; further work on this platform is aiming to create more complex glycoform modifications. Tytgat et al. (2019) developed a glycoengineering platform in the *Escherichia coli* cytoplasm ("Glycoli") that uses a site-specific polypeptide glycosyltransferase and a modulable glycosyltransferase to create a variety of multivalent glycostructures. Natarajan et al. (2020) developed a series of orthogonal pathways for the attachment of cancer-associated mucin-type glycans (Tn, T, sialyl-Tn, and sialyl-T) on human-associated proteins an *Escherichia coli* system, providing a platform that can perform other diverse forms of glycosylation. Finally, Chang et al. (2019) engineered synthetic circuits in Chinese hamster ovary cells to perform timely *N*-linked glycosylation of Immunoglobulin G, which alleviates a previous manufacturing barrier in controlling glycosylation for monoclonal antibody production. Further research can examine how to scale up several of these efforts and continue discovering platforms to synthesize other protein glycoforms.

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

Relative to cell-free systems and single-celled organisms, multicellular organisms encompass several unique challenges for engineering functions and molecular synthesis pathways. Chiefly among them is that the synthetic tissues, systems, and platforms can be composed of highly divergent cell types, making targeted modification difficult. Although problematic in this sense, the increased complexity also allows for the opportunity to exploit sophisticated properties characteristic of multicellular organisms, namely pattern development, architecture, and population ratios. The on-demand fabrication of multicellular host systems can therefore enable a wide variety of different applications. For example, completing this goal could enable engineered hosts for converting agricultural wastes into commodity products that can sustain production yield and efficiency under a wide range of stress conditions. Modifying multicellular organisms, such as plants, can also allow the engineering of oil crops to be drought-tolerant and not require significant fertilizer inputs to produce biofuels. Multicellular modification in exemplary animal systems can also further develop patient-matched disease models for making personalized drugs and treatments.

Breakthrough Capability: Ability to control differentiation and de-differentiation of cells within a population. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates that the 2021 milestone has been reached.

2021 Milestone: On-demand, reproducible functionalization of simple micro-tissues or micro-consortia made up of two or more engineered cell types.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Tissues are groups of similar cells that collectively function together, while consortia are defined as multiple bacteria or microbial organisms living symbiotically with one another. In both instances, disparate organisms/cells occupy the same space, may compete for the same resources, and respond to the activity of organisms in the immediate environment. Therefore, controlling the functionalization of a microbial consortium or tissue precisely can be incredibly challenging. Since 2019, there has been a flurry of research activity investigating ways to control and measure co-culture populations in experiments. For example, Burmeister et al. (2021) created an approach to optochemically control co-cultures by engineering two *Corynebacterium glutamicum* strains, one that cannot synthesize a substrate required for its survival and another that can take photoactivatable-IPTG to produce this substrate. In application of this method, researchers can use optochemical manipulation (by activating the IPTG through illumination and therefore stimulating the required substrate) and medium controls (by supplying unmodified IPTG) to examine co-culture interactions and control colony growth. Dihn et al. (2020) developed a quorum sensing-based growth-regulation circuit that can regulate co-culture populations used in fermentation systems, demonstrating a 60% titer increase in a naringenin-producing co-culture relative to co-cultures using the *status quo* inoculum-based approach for population control. Additionally, Toda et al. (2020) discovered a method to convert fluorescent tag molecules into synthetic morphogens (molecules that can govern the local development of tissue pattern formation) to localize cell differentiation to particular areas. There has also been a lot of research activity demonstrating the power of co-culture strategies towards different applications and impacts. Horner et al. (2019) showed how to differentiate human mesenchymal stem cells to differing phenotypes in a 3D-spatially-regulated manner using a mechanical gradient, successfully creating a graduated tissue of different cell types that be used to reconstruct more biologically-representative tissue environments. Similarly, in regards to structure, Murphy et al. (2019) generated a 3D organoid endometrium model consisting of epithelial and stromal cells, better recapitulating the natural physiology and the ability to study endometrial and pregnancy diseases. Flores et al. (2019) co-cultured wild-type and ethanologenic (LY180) strains of *Escherichia coli* to break down complex lignocellulose-derived sugars to ethanol, achieving higher ethanol titer (46 g L^{-1}), productivity (488 mg L^{-1} h^{-1}), and yield (~90% of theoretical maximum) compared to monocultures. Finally, VanArsdale et al. (2020) created a co-culture system comprising "catalytic" and "reagent" engineered-transducer cells from *Pseudomonas aeruginosa* that can process molecular cues and produce an electrochemical output that researchers can record on a device. Future research can continue these efforts and develop frameworks and universal guidance on co-culture methodology.

Breakthrough Capability: Ability to characterize and control the three-dimensional (3D) architecture of multicellular systems. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap; however, the Assessment literature review indicates that the 2021 milestone has not yet been achieved. The Assessment suggests that the 2029 milestone "Create modular, synthetic communication circuits that can be implemented in tissues to allow for control of new or existing cellular communication systems" and 2039 milestone "Bottom-up design and construction of whole organs at the centimeter-length scale" may be achieved ahead of schedule.

2021 Milestone: Characterize existing tissue components and standardize measurements to evaluate function.

Progress toward this milestone is **significant**, with some research gaps remaining.

Tissue composition and geometry are critical parameters for defining the three-dimensional architecture of a multicellular system, however both suffer from inconsistent and unstandardized reporting across the literature. For example, it is worth highlighting advances in spatial transcriptomics to better understand the spatial and temporal organization and interactions of multicellular structures (Ren et al., 2022; Ben-Moshe et al., 2022). Much research since 2019 has focused on developing improvements in hydrogel and organoid technology to better recreate the three-dimensional environment of tissues. Brassard et al. (2020) introduced a method to

generate three-dimensional organoids through a bioprinting process that allows self-organization on extracellular matrices and precise control over geometry and cellular density, generating centimeter macro-tissues for engineered purposes. Alternatively, Guo et al. (2019) developed a modular hydrogel cross-linker that can be functionalized with small peptides and large macromolecules, therefore enabling researchers to enhance hydrogel functionality to crosslink and organize in response to biological stimuli. There has also been development of alternative and novel strategies to build more sophisticated engineered multicellular architecture. For instance, Blackison et al. (2021) generated a platform that can create *in vitro* "xenobots" (biological robots) derived from *Xenopus laevis* that exhibit self-organization through cilia present on their surface; the authors further created a computational model that can help understand how these collective behaviors can assemble in developmental patterns typical of natural biological scaffolds. Kriegman et al. (2020) developed a pipeline that designs novel living systems *in silico* by assisting the user in defining the cellular behaviors needed for desired functionalities and using an evolutionary algorithm to craft a building-block blueprint of cellular structure. And Bücher et al. (2022) described a bottom-up approach toward the synthetic construction of target-specific, cytotoxic immune cells for the bioinspired construction of effector immune cells from basic building blocks, giving a detailed characterization of these cells by microfluidics, electron and light microscopy, dynamic light scattering, and flow cytometry. In summary, there has been much development on different approaches to building the higher-order structures of tissue architecture, yet there remain significant gaps in being able to rapidly characterize the geometrical or matrix composition of the desired tissue and immediately couple an engineering design strategy to synthesize it. Further research towards these characterization and translation efforts can progress this space.

Breakthrough Capability: Ability to achieve stable non-heritable changes in somatic cells. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap and the Assessment literature review indicates the 2021 milestone has been reached. The Assessment suggests that the 2029 milestone "Ability to generate cell states that are stable and effective after the inducer/effector is removed in certain model tissues" and the 2029 milestone "Ability to generate cell states that are stable and effective after the inducer/effector is removed in certain model tissues" may be achieved ahead of schedule.

2021 Milestone: Routine delivery of biomolecule ''effectors'' (i.e., DNA, RNA, proteins) into slowly-dividing or non-dividing cells.

Progress toward this milestone is **significant**, with some research gaps remaining.

Researchers need to be able to edit, probe, and engineer synthetic tissues to respond to stimuli and their environment; however, this can be especially difficult with tissues comprised of slowly dividing cells, where nucleic acid and protein effectors have a limited time window to edit the genome before little (if any) cell divisions take place. Since 2019, several developments have focused on enhancing nanoparticle or platform delivery systems to improve effector introduction. For instance, Lee et al. (2020) created a versatile polymeric-protein nanocomposite platform that can deliver proteins to the cytosol with very high efficiency (90%), enabling a new approach to potential therapeutic delivery. Similarly, Han et al. (2020) developed a template-mediated supramolecular assembly strategy to synthesize protein–polyphenol nanoparticles to escape endosomal encapsulation and deliver effector cargo straight to the cellular cytosol. Several researchers have also focused on constructing bio-inspired delivery platforms to encapsulate or deliver engineered biological systems. Ganar et al. (2021) described a procedure for their invention of actinosomes, natural cell-sized, porous containers crafted using the interactions between biomolecular condensates and actin cytoskeleton, that researchers can use to encapsulate cell-free translation machinery for engineering biology applications. Additionally, Staufer et al. (2021) developed a synthetic, bottom-up procedure to generate extracellular vesicles of a user-defined composition of RNA, lipids, and proteins to better understand how extracellular vesicles can be used for molecular signaling. Although this current technology isn't "routine", numerous strategies, from nanoparticles to biologically-inspired encapsulation, provide accessible solutions for researchers aiming to

deliver nucleic acid or protein effectors to slowly dividing cells. Research in this area can further benefit from developing these strategies and identifying new bio-inspired cargo delivery methods.

Breakthrough Capability: Ability to make predictable and precise, targeted, heritable changes through germline editing. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. Although the Assessment suggests that the 2021 Milestone "Define and validate tissue-specific DNA parts in plants" has not yet been achieved, it indicates that the 2024 milestone "Ability to domesticate engineered biological parts to confer immune tolerance in immunocompetent organisms," both the 2029 milestone "Ability to coordinate engineered multicellular functions in intact organisms via orthogonal communication systems," and 2029 milestone "On-demand gene editing of organisms with desired traits," and the 2039 milestone "Routine, on-demand, efficient germline editing for any targeted hosts of interest at high-throughput scale" may be achieved ahead of schedule.

2021 Milestone: Complete sequence of select host genomes to allow design of targets for gene editing.

Progress toward this milestone is close to **complete**, with minimal research gaps remaining.

Although genetic sequencing technology has rapidly progressed in the past couple of decades, numerous higher-order model (and non-model) organisms still do not have their entire genome sequenced, largely because of size and complexity. This lack of sequence makes several biological manipulations difficult, especially mainstay techniques such as genome editing. Although the research roadmap did not select specific organisms to have an assembly of their genome published, much has been devoted to organizing the existing genomic data for model organisms and creating tools to increase the efficacy of gene editing target selection in hosts that lack complete genome information. Research on classical model organisms (including humans) continues to develop high-quality resources to assist sequencing more species. For example, the National Institute of Standards and Technology and Joint Initiative of Metrology in Biology lead "Genome in a Bottle" (https://www.nist.gov/programs-projects/genome-bottle), a program that provides human haplotype samples, documents high-confidence variance calls for reference datasets, and establishes criteria for problematic genome variants. In their commentary, the *Alliance of Genome Resources* (2019) described their publicly-accessible database that shares curated data, including reference genomes, for classical and budding model organisms commonly studied by the basic research community. It should also be noted that several research consortia, public-private partnerships, and coalitions engaged in developing the references and standards for genome editing in non-model organisms (reviewed by Che et al., 2019). Several developments have also focused on how to increase the efficiency of gene-editing tools in non-model hosts, especially those that lack genomic data. For example, *Sun et al.* (2019) developed CRISPR-local, a local single-guide RNA (sgRNA) design tool for non-reference plant genomes for energy, environmental biotechnology, and food and agricultural research. Additionally, *Wilken et al.* (2020) presented a codon-optimization strategy to aid genetic engineering tools by uncovering unique characteristics of anaerobic gut fungi properties for use in biomass research. There is much excitement and demand among researchers to sequence non-model organisms and create genome assemblies, as well as government and commercial suppliers willing to provide these services. This research area can be bolstered by public-private partnerships of engineering biology researchers identifying specific, high-value organisms to be sequenced, as science funders and policymakers are willing to finance these demands but need more targeted guidance.

2021 Milestone: Define and validate tissue-specific DNA parts in plants.

Progress toward this milestone is **significant**, with some research gaps remaining.

Defining and validating tissue-specific DNA parts in plants is challenging. It requires understanding plant gene expression control mechanisms conserved across species, in addition to species-specific mechanisms. Nonetheless, there has been tremendous progress in this space since 2019. *Belcher et al.* (2020) built a library

of transcriptional regulators to create DNA parts for plant systems; as a demonstration, they validated their repressors, activators, and enhancers in *Nicotiana benthamiana* and *Arabidopsis thaliana*. Dudley et al. (2021) designed a workflow for plant protein characterization that automates processes for DNA assembly and cell-free expression system construction, even bypassing the need to perform protein purification for functional assays on potential genetic parts. Among these platform developments for plant parts, several researchers have also focused on characterizing or developing tools for potential model species. For example, Dugé de Bernonville et al. (2020) described their strategy of using whole-genome bisulfite sequencing and RNA-seq to understand the complex regulatory pathways of the medicinal plant *Catharanthus roseus* used to generate monoterpene indole alkaloids. Decaestecker et al. (2019) described a CRISPR-based tissue-specific knockout system for *Arabidopsis thaliana* that enhances creation of precise mutations in the organism and analysis of location-specific effects, without affecting fertility and reproduction. Finally, Feder et al. (2020) developed a fruit-specific CRISPR-based knockout systems in *Solanum lycopersicum* (tomato), demonstrating their technology by targeting a Green Fluorescent Reporter protein in addition SIEZ2, a gene involved in plant morphology. Significant progress has been made toward the generation of platforms and parts for plant systems, though there is still much work to be done to use these tools in different plant hosts. Ongoing research in this space is expected to yield advanced toolkits and platforms for generating plant parts and to apply existing systems to additional species and tissues.

Host Engineering Goal: Generation of biomes and consortia with desired functions and ecologies.

Among the obstacles to engineering biomes and consortia are having to account for potential ecological competition, barriers to interspecies communication, and problematic symbiotic relationships between species. However, the rewards of successful efforts can be extraordinary, developing microbial consortia or complex biomes with predictable composition, dynamics, and function. Achieving this goal will enable researchers to create biological systems that can act as multiplexing sensors capable of analyzing multiple environmental cues and providing measurable responses (or combinations of responses). Completing this goal can enable stable, engineered microbial cultures that can enrich soils, support gut microbiomes, and restore damaged ecosystems.

Breakthrough Capability: Ability to control cell-to-cell communication between different species. This Breakthrough Capability is not meeting the pace of the predictions relative to the roadmap and the Assessment suggests the 2021 milestones have not been reached.

2021 Milestone: Tightly-controlled promoter-response regulator systems that enable intra- and inter-species cellular communication.

Progress toward this milestone is **modest**, with significant research gaps remaining.

Many circuits, systems, and platforms use regulatory components to control gene expression pathways to synthesize valuable molecules or perform useful functions. Within an engineered consortia or biome, these regulatory controls need to accept molecular inputs or stimuli stemming from multiple species; therefore, having inclusive regulator systems capable of inter- and intra-cellular communication is paramount. Since 2019, there has been some research devoted to controlling this form of communication; while quite a bit of study has been devoted to quorum sensing within species, there have been far fewer engineers dedicated to manipulating cell-to-cell communication between species, in part because greater basic science understanding of interspecies communication is needed. Stephens et al. (2019) designed a synthetic co-culture controller consisting of a cell-based signal translator and growth-controller module capable of autonomously regulating the population composition; they additionally refined their results into a mathematical model capable of predicting population trajectories of the system. Miano et al. (2020) engineered an inducible quorum sensing controller that can tune the bacteria dynamics of an engineered system at both the population and community

level, demonstrating their technology with strains equipped with genetic cargo for synthetic circuits. Towards demonstrating inter-species communication control, *Wellington, et al.* (2019) measured activity of quorum-sensing receptors across species, uncovering robust non-self signal response and suggesting that promiscuous receptors can selectively react to interspecies cooperation and competition signals. Further research can continue developing new forms of controllers to work across species boundaries.

Breakthrough Capability: 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). This Breakthrough Capability is not meeting the pace of predictions relative to the roadmap; the Assessment literature review indicates the 2021 milestone has not yet been reached.

2021 Milestone: 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.

Progress toward this milestone is **modest**, with significant research gaps remaining.

Researchers that want to be able to engineer consortia or biomes inspired by nature need a sound understanding of what components and species constitute the environment they are trying to generate or recreate. Researchers have begun to catalog the species in these environments, primarily using metagenomics tools. For example, *Nayfach et al.* (2021) published a genomic catalog of over 10,000 genomes from different microbiome constituents representing many different terrain and ocean habitats. This information can be used to identify species amenable to valuable secondary metabolite synthesis. *Chen et al.* (2021) analyzed the microbiota of swine from 787 gut microbiomes to create an extended pig-integrated gene catalog (PIGC), uncovering over four million unknown proteins in the process. *Li et al.* (2020) used metagenomic sequencing to identify the microbiota present in the bovine rumen, identifying over thirteen thousand non-redundant prokaryotic genes, some of which are likely instrumental in the breakdown of plant polysaccharides. As a last example, *Ma et al.* (2021) constructed a microbial gene catalog of species involved with anaerobic digestion from 56 full-scale biogas plants across China, uncovering how feedstocks (chicken, cow, or pig manure) affected the microbial composition present in each plant. Despite this progress with metagenomics however, comprehensive characterization of biomes and consortia regarding proteins or metabolites is far behind. Further investigation efforts will be needed to catalog a wider array of Earth's different environments and biomes beyond their genomic composition. Particular bottlenecks include extracting RNA from environmental samples; increased sampling, throughput and fidelity for proteomics research, and improved separation and recovery of metabolites.

Breakthrough Capability: Ability to control and/or define the function of an engineered microbial community/biome. This Breakthrough Capability is proceeding as predicted relative to the roadmap. The Assessment literature review indicates the 2021 milestone has been reached.

2021 Milestone: Ability to combine species with specialized functions to enable the production of desired products.

Progress toward this milestone is close to **complete,** with minimal research gaps remaining.

Many microbial species can encompass unique characteristics that endow useful properties for creating valuable biomolecules. Researchers have investigated how they can combine disparate species and the properties of multiple systems to achieve a greater degree of synthetic output. Since 2019, researchers have discovered many helpful co-culture combinations and strategies to combine different species under a singular system, providing numerous examples of clever strategies, approaches, and toolkits to make more constructive use of co-cultures. For example, *Castro et al.* (2019) co-cultured the fungus *Aureobasidium pullulans* and

Saccharomyces cerevisiae to purify fructooligosaccharides in non-prebiotic sugar mixtures, creating an improved yield of ethanol and fructooligosaccharides in the process. Federson et al. (2020) created a co-culture of the cyanobacterium *Synechococcus elongatus* and heterotrophic bacterium *Pseudomonas putida* to degrade the environmental pollutant 2,4-dinitrotoluene, exploiting the photo-metabolomic properties of the co-culture to create improved degradation activity. Guo et al. (2019) coupled a biosensor with an *Escherichia coli* co-culture, each with a molecular pathway to synthesize 4-hydroxybenzoate and tyrosine, respectively, to efficiently produce phenol from glucose substrate compared to monoculture strains. Zhang et al. (2021) co-cultured *Gluconobacter oxydans* and *Escherichia coli* to produce 3,4-dihydroxybutyric acid (typically created with hazardous chemicals and harsh reaction conditions) from natural xylose molecular precursors. An interesting example of research toward this milestone was work by Foster et al. (2021) who fused two microbial cells to create a hybrid with new synthetic properties. They created a dynamic genome-scale metabolic modeling framework that evaluates the changes in properties in a fusion event between *Clostridium acetobutylicum* and *Clostridium ljungdahlii* and with their model, forecasted improvements in ethanol and isopropanol yields as well as growth kinetics. And in a human gut microbiome example, Clark et al. (2021) developed a data-driven model-guided approach to design consortia for optimal butyrate production. More research needs to be undertaken to understand how the fusion events systematically alter the original properties of the host. In summary, there have been many numerous examples that demonstrate the power of combining, or even fusing, different species together to create enhanced functionalities for biologically engineered products. Researchers can improve progress in this space by creating toolkits, models, and algorithms that suggest the compatibility of multiple strains to be co-cultured and the potential biosynthetic benefits of their combination.

Breakthrough Capability: 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. This Breakthrough Capability is not yet meeting the pace of predictions relative to the roadmap; the Assessment literature review indicates the 2021 milestone has not yet been reached. However, evidence suggests that the 2024 milestone "Characterize how select microbiomes respond to changes in the environment, including the addition of toxins, the introduction of new organisms (pathogens or commensals), and the selective removal of species from the community" may be achieved ahead of schedule.

2021 Milestone: Use of existing technologies (including metagenomics, transcriptomics, proteomics, and mass spectrometry) to characterize functions of microbial communities from a broad range of environments.

Progress toward this milestone is **modest**, with significant research gaps remaining.

The various environments and biomes in nature encompass a wide degree of helpful biological diversity for synthesizing biomolecules or engineering useful functions in biological systems. However, to apply this knowledge, there must be a basic understanding of the microbial characteristics of these environments. Since 2019, there are many examples of research using -omics, primarily metagenomics, technologies to characterize communities for engineering biology applications. Sheth et al. (2019) created Metagenomic Plot Sampling by sequencing (MaPS-seq), "a culture-independent method to characterize the spatial organization of a microbiome at micrometer-scale resolution," identifying robust spatial associations of *Bacteroidales* taxa in the gut as a demonstration. Engelberts et al. (2020) used an integrative metagenome-assembled genomics approach to map 259 microbiome symbionts of the marine sponge model system *Ircinia ramosa*, uncovering how critical cellular functions, like carbon fixation, were spread across the various taxa. In their methods article, Roy et al. (2021) developed a guide to showcase how multiple computational tools can enable the storing, visualization, and leverage of multi-omics data collection, demonstrating their guide's utility by using it to create a machine learning algorithm to design new strains for isoprenol production. Additionally, Amarelle et al. (2019) described their strategy of using metagenomic approaches to "mine" transcriptional terminators and other genetic parts, demonstrating their approach in *Pseudomonas putida* and identifying four sequences that

inhibit transcription in *Pseudomonas putida*, *Escherichia coli*, *Burkholderia phymatum*, and Antarctic *Pseudomonas* strains. Among the methods that the roadmap contributors predicted in framing this milestone, several researchers were motivated by other biological principles to characterize complex environments for engineering biology. Inspired by complex genetic interaction models, *Sanchez-Gorostiaga* (2019) designed a quantitative framework that describes how the amylolytic rate (the enzymatic splitting of starch into soluble products) is affected by the combinatorial assemblage rate of different soil bacteria, providing a model to measure how complex communities affect engineered function. Inspired by ecological principles, *Fedeorec et al.* (2021) exploited aspects of amensalism (the association between organisms of two different species in which one is inhibited or destroyed and the other is unaffected) and competitive exclusion to create a tunable, stable two-strain consortium in which one of the strains secretes a toxin in response to inhibitory competition. Additionally, there have been some developments to improve other technologies besides metagenomics to characterize consortia. For instance, *Aakko et al.* (2020) provided a proof of concept for data-independent acquisition metaproteomics, which better enables the integration of mass spectrometry into metagenomic data through increased accuracy and more consistent quantification of the data, regardless of its source. While there has been much development towards applying metagenomics analysis to understand better how consortia and biomes can be engineered to provide valuable biomolecules or functions, further research is needed to apply other methodologies and examine how other ecological principles can assist characterization efforts.

Data Science | Data Integration, Modeling, and Automation

The Data Integration, Modeling, and Automation technical theme ("Data Science") highlights capabilities that could dramatically amplify the ability to perform complex analyses and predictions, leading the way to advanced modeling and automation. Data science for engineering biology faces the challenge of "the need for novel and more robust computational tools and models." These computational tools, such as simulating potential experimental outcomes, designing optimal pathways to synthesize biomolecules, and creating streamlined manufacturing processes, are critical to navigating the inherent complexity of biological organisms, and the past few years of research have undoubtedly sought to fill this need. Along with the original assumptions of the roadmap, a greater range of technologies foundational to data science for engineering biology have also emerged, including the rapid advancement of machine learning technologies devoted to single-cell analyses. Unfortunately, the data science capabilities hoped for in the roadmap have largely been unattained.

Progress in Data Integration, Modeling, and Automation
Goal: Establish a computational infrastructure where easy access to data supports the DBTL process for biology.
Breakthrough Capability: Established standard and accessible repositories for biomanufacturing data and analysis methods. ●○○○ 2021 Milestone: Have developed a system of robust communication between academia and industry surrounding engineering biology data access and needs. ●●●○ 2021 Milestone: Develop findable, accessible, interoperable, and reusable (FAIR) data standards and open repositories for engineering biology.
Breakthrough Capability: Common computational infrastructure for finding biological data and common APIs for search and analysis. *This Breakthrough Capability does not have any associated 2021 milestones.*
Breakthrough Capability: End-to-end, industry-normed design software platforms for engineered biological systems. *This Breakthrough Capability does not have any associated 2021 milestones.*
Goal: Establish functional prediction through biological engineering design at the biomolecular, cellular, and consortium scale.
Breakthrough Capability: Fully-automated molecular design from integrated, large-scale design data frameworks. ●●●○ 2021 Milestone: Structure- and comparative analysis-based libraries for automated directed evolution, with feedback of large-scale results to algorithms.
Breakthrough Capability: Use of enzyme promiscuity prediction algorithms to design biosynthetic pathways for any molecule (natural or non-natural). ●●○○ 2021 Milestone: Retro-biosynthesis software that can identify any biological or biochemical route to any organic molecule.

(Table continues)

Goal: Establish functional prediction through biological engineering design at the biomolecular, cellular, and consortium scale. *(Continued)*	
Breakthrough Capability: Scalable, data-driven host design for complex environments that enable high-level production of natural biomolecules.	
●●●○	2021 Milestone: Ability to make and screen multiple host mutations for epistasis mapping and synthetic interactions, making large-scale host optimization possible.
●●○○	2021 Milestone: Better data on physiology and fitness in deployment environments suitable for informing designs in validated lab-scale simulations that meet activity, persistence, and ecological impact goals.
Breakthrough Capability: Enabled design of functional, self-supporting ecosystems.	
●●○○	2021 Milestone: Data-driven tools for selecting organisms for synthetic assemblies to achieve resistant, resilient activity.
●●○○	2021 Milestone: Direct data collection for the most important communities in human, agriculture, and complex bioreactor work sufficient for informing design.
●●○○	2021 Milestone: Modeling tools to identify cross-organismal networks and ecological interactions.
Goal: Establish optimal manufacturing processes from the unit-operation to the integrated-screening scale.	
Breakthrough Capability: Standardized informatics tools, data, and automation platforms for efficient and collaborative use and integration of data in order to develop novel products more quickly.	
●●●○	2021 Milestone: Establish communications and networks to develop democratized platforms for data exchange and automation across industry and academia.

Table 3. Assessment of Data Science 2021 Milestone Achievement. Each 2021 milestone was assessed to determine progress towards its achievement. Four filled circles indicates the 2021 has been achieved or is close to complete, three filled circles indicates significant progress towards the 2021 milestone, two filled circles indicates modest progress towards the 2021, and one filled circle indicates only minimal progress towards achieving the 2021 milestone. In Data Science, none of the 2021 milestones have been achieved or are close to complete (four filled circles).

Highlights of Technology Developments in Data Science

Machine Learning Analysis for Biomolecular and Single-Cell Technologies

Engineering biology needs computational tools to analyze the vast amounts of data produced, especially when automated methods and parallelized experiments can create data around the clock. Machine learning has proven crucial to addressing this increased need, especially with the ability to infer knowledge and develop tools to predict protein structure and develop mathematical models from single-cell transcriptomic and epigenetic data. (For more information about machine learning for biomolecular engineering, see **Highlights of Technology Developments in Biomolecular Engineering**.) Advancements in machine learning capabilities have included data normalization, the classification of different cell types, deciphering gene regulatory networks, and the inter-operationalizing of multiple data sources (Raimundo, 2021). In particular, the number of single-cell RNA sequencing tools has drastically increased to the point where anthologies of available tools (along with descriptions of how to use them) have developed (Zappia, 2018). Given these incredible analysis mechanisms, it will be essential to create high-quality, standardized datasets to ensure reproducibility across the field and address the limitations of single-cell approaches, such as batch effects and dropouts.

Data Science Barriers to Advancement

High-Throughput Automation for Non-Model Organism Domestication

High-throughput and automated procedures, such as those that introduce genomic edits or pathways into organisms, can greatly increase the utility of the organism to produce valuable molecules or perform valuable functions. While not entirely high-throughput, progress has been made in automation platforms for organismal engineering; see the development of CRAGE by *Wang et al.* (2019) and follow-on work by *Liu et al.* (2020) as examples. And although there has been much development in automation and high-throughput procedures for canonical organisms such as *Escherichia coli* and *Saccharomyces cerevisiae*, these workflows do not generally extend to non-model organisms, exacerbating the barriers to non-model domestication.

Simulations, Projections, and Modeling to Guide Experimental Planning

Due to the multitude of factors that can affect an organism's function, the number of experimental alterations that can affect a system can be incredibly high and therefore costly and time-consuming to properly investigate. To address this barrier, simulations, projections, and models that can accurately predict functional outcomes are viewed as a critical technology for engineering biology. Advanced modeling could help to predict the health of different species under co-culturing conditions, the effect of genetic perturbations on biochemical production, or the growth rate of an organism in a response to a change in the environment. Newer modeling approaches take advantage of machine learning and data science capabilities to predict behavior, thereby saving researcher time, cost, and energy by allowing them to design changes most likely to achieve their goal (Chao, 2020; Ching, 2018). In their review, *Ortero-Muras and Carbonell* (2020) described automated engineering of synthetic metabolic pathways, specifically highlighting optimal experimental design approaches for biomanufacturing. In their survey analysis, *Tellechea-Luzardo et al.* (2022) delineate challenges and technical solutions for building automated pipelines for biofoundries working to develop optimized biotechnological systems. Recognized as a widespread bottleneck, this has also led to numerous efforts, including community contests, to create artificial intelligence systems that can inform engineering and systems biology research. One such example is the Nobel Turing Challenge to "develop a highly autonomous AI system that can perform top-level science, indistinguishable from the quality of that performed by the best human scientists" (Kitano, 2021). Continued development of simulations, projections, and modeling to guide experimental planning across all technical themes is necessary to improve and speed progress across the field.

Publicly Accessible and Shared Data for Engineering Biology Research

Large datasets play a pivotal role in understanding the nuanced characteristics that dictate how biological systems perform useful functions. For example, metagenomics (the study of a collection of genetic material from a mixed community of organisms) helps produce data to understand how microbial consortia interact under environmental conditions. The sharing of large analytical datasets between academic, government, and industry researchers is commonly cited as a major hurdle to engineering biology progress. Coordination between these stakeholders is necessary to support advancement of accessible data sets, as are incentives, particularly for private industry, to share non-proprietary data. The roadmap highlighted the importance of data systems that follow the FAIR (findable, accessible, interoperable, and reusable) data standards. As engineering biology looks to automation and data science approaches to enhance analyses, researchers have a ripe opportunity to incorporate FAIR data standards into their work and create open repositories to share data tools.

A combination of context-dependent and universal policies are likely to play a role in remediating this issue. Incentives to share data as a requirement to participate in common-interest initiatives, such as BioMADE and the Agile Biofoundary, could potentially catalyze standards of sharing across diverse research bodies. Clarity and requirements for data sharing by funders and publishers could also help to strengthen data sharing practices (National Academies of Science, Engineering, and Medicine, 2021). Further, publicly accessible data collection "moonshots," funded by governments or private entities, could also enable equitable sharing of

useful data among researchers. There are increased efforts to support data aggregation across large projects, such as the National Institutes of Health Common Fund Data Ecosystem (CFDE; https://commonfund.nih.gov/dataecosystem). The CFDE and other efforts will eventually help to accelerate discovery across diverse, heterogeneous data.

Data Science Goal: Establish a computational infrastructure where easy access to data supports the DBTL process for biology.

Design-Build-Test-Learn cycles require data to make informed decisions and course adjustments during the engineering process. Easy-to-access computational infrastructure acts as one of the core nodes to facilitate this data transfer. Several components of this infrastructure, such as biomanufacturing data repositories, standard application programming interfaces (APIs), and end-to-end software platforms, can significantly expedite engineering biology progress and promote accessible tools for researchers across the globe. Computational infrastructure for engineering biology can enable researchers to estimate the robustness of circuits and pathways to genetic, host, and environmental contexts and aid in creating predictive scale-up for models. Supportive computational infrastructure can also help better understand (analyze, model, and predict) microbial consortia in natural systems and how they interact and evolve over time and under different conditions for engineering biomes. Further, it can aid the development of protein libraries correlated to biomarkers used to identify promising context-specific sensors and treatments.

Breakthrough Capability: Established standard and accessible repositories for biomanufacturing data and analysis methods. This Breakthrough Capability is not meeting the pace of the predictions relative to the roadmap. The Assessment literature review indicates the 2021 milestones have not been reached. Additionally, the Assessment anticipates that the 2024 milestone "Biomanufacturing-specific data standards and repositories" may not be achieved in the timeframe anticipated by the roadmap.

2021 Milestone: Have developed a system of robust communication between academia and industry surrounding engineering biology data access and needs.

Progress toward this milestone is **minimal**, with significant gaps remaining.

High-quality engineering biology data is required across academia, industry, and government research, sometimes of the same datasets, to accomplish their mission and objectives. Although these groups may share this requirement, data sharing among them is not guaranteed, and entities may lack the incentive to prioritize sharing. Since 2019, numerous initiatives have stressed the importance of creating infrastructure to share these resources, along with the beginnings of preliminary platforms to provide communication between these groups. In their commentary, *Hillson et al.* (2019) discussed the importance of biofoundries, exemplifying the Global Biofoundry Alliance, to enable the rapid design, construction, and testing of organisms to scale engineering biology activities to solve industrial and societal needs. Similarly, *Farzaneh and Freemont* (2021) commented on how biofoundries serve as strategic institutes that can facilitate data sharing standards due to their position as a central focal point for collaborations. *Holowko et al.* (2020) outlined several technical and operational considerations for biofoundry stand-up, including "drivers for establishment, institutional models, funding and revenue models, personnel, hardware and software, data management, interoperability, client engagement, and biosecurity issues." The Global Biodata Collection (https://globalbiodata.org/) serves as a forum for institutions and government bodies that fund research and data infrastructure to coordinate and share tools and approaches to manage biological data and grow data resources. Similarly, the Bioindustrial Manufacturing and Design Ecosystem (https://biomade.org/) is a U.S. Manufacturing Innovation Institute launched in 2021, that aims to develop infrastructure to facilitate data management and analysis and a robust, collaborative data exchange as part of its efforts to "realize the economic promise of industrial biotechnology" (BioMADE). Several research communities have also gathered to create more uniform engineering biology standards for data sharing. For example, as part of the European Commission's

BioRoboost: Fostering Synthetic Biology Standardisation Through International Collaboration, Baldwin (2020) released a report on the gaps, challenges, and opportunities related to synthetic biology standards (readers can follow this project at https://standardsinsynbio.eu). Finally, *Brown et al.* (2020) discussed how the data standard, Synthetic Biology Open Language (SBOL), has evolved to enable researchers to engineer multicellular systems and better document functionality. There has been much initial progress in understanding the value of biofoundries, data sharing incentives, and standards for a productive engineering biology enterprise; however, *robust* communication on these issues concerning government, industry, and academia remains to be reached. Consortia-led efforts to inclusively engage participants of the rising global bioeconomy could provide significant forward progress toward this milestone. Institutions like BioMADE, or EBRC ourselves, could potentially play this role.

2021 Milestone: Develop findable, accessible, interoperable, and reusable (FAIR) data standards and open repositories for engineering biology.

Progress toward this milestone is **significant**, with some research gaps remaining.

Findable, accessible, interoperable, and reusable, or FAIR, data standards are a series of "guidelines for those wishing to enhance the reusability of their data holdings," especially by enhancing the ability of machines to find and use the data alongside people automatically (Wilkinson et al., 2016). As engineering biology looks to automation and data science approaches to enhance analyses, researchers have a ripe opportunity to incorporate FAIR data standards into their work and create open repositories to share data tools. In their correspondence article, *Sansone et al.* (2019) discussed their development of FAIRsharing, a resource that links community-driven standards, databases, repositories, and data policies across various academic disciplines, including those relevant to engineering biology. *Waltemath et al.* (2020) reported on the outcomes of the 10th Computational Modeling in Biology Network (COMBINE), an event focused on systems and synthetic biology standards, and discussed proceedings on FAIR data sharing and computational model standardization. Open repositories and tools have continued to develop under the guidance of FAIR principles. For example, *Yeoh et al.* (2021) presented SynBioPython, an open-sourced Python package that provides standard software solutions to help with engineering biology applications, including batch DNA design, sample and data tracking, data analysis, and more. *Madsen et al.* (2019) described updates to the Synthetic Biology Open Language (SBOL), an open-source standard for the electronic exchange of information on the structural and functional aspects of biological designs, detailing changes to representing sequence modifications, attachments of experimental data, and describing numerical parameters of experiments. Also, *Plahar et al.* (2021) described the implementation of BioParts, a search engine incorporated into the Inventory of Composable Elements (ICE) of the iGEM Registry of Biological Parts, that can identify parts available in the public domain for use in engineering biology. *Torre et al.* (2018) continued updates to Datasets2Tools, a massive repository of over 6,800 RNA-seq and proteomic datasets, 4900 tools, and 31,5000 analyses for use in bioinformatics research and analysis. *Lüders et al.* (2022) developed ODEbase, a database with 662 models in the Systems Biology Markup Language (SBML) format. And *Malik-Sheriff et al.* (2020) summarized the progress in BioModels over the past 15 years and pointed out future directions for the open-source repository with curated models; while more focused on systems biology, BioModels can serve as a valuable resource for engineering biology. There has been significant progress in utilizing tools, discussions, and repositories concerning data under FAIR auspices; however, the incorporation of FAIR practices remains to be propagated extensively throughout individual academic, government, and industry laboratories. Progress towards this milestone can be improved by a vital adoption of FAIR data sharing principles by more stakeholders.

Breakthrough Capability: Common computational infrastructure for finding biological data and common APIs for search and analysis. ***There were no 2021 milestones for this Breakthrough Capability***. The Assessment suggests that this Breakthrough Capability may meet predictions. Progress has been made towards the 2024 milestone “Produce a common library of open design tools, built upon standard APIs, and supported by

portable/virtualized execution environments to demonstrate best-practice interoperable biomanufacturing software" including, for example, SynBioHub (https://synbiohub.org) which is a repository for biological construct design, and the movement to the cloud of searchable sequence data stored by the National Center for Biotechnology Information (NCBI; see https://ncbiinsights.ncbi.nlm.nih.gov/2020/02/24/sra-cloud/ for further detail).

Breakthrough Capability: End-to-end, industry-normed design software platforms for engineered biological systems. ***There were no 2021 milestones for this Breakthrough Capability.*** The Assessment suggests that the 2024 milestone "Develop industry-accepted, sharable assessments of current data tools and uses in reducing cost and increasing reliability of executing the DBTL cycle" may not be achieved when anticipated and that this Breakthrough Capability may not be meeting the pace of the predictions relative to the roadmap. Without robust policies and incentives for public-private data sharing, and for stronger, widely accepted metrics and standards for engineering biology, this Breakthrough Capability will likely be very difficult to achieve.

Data Science Goal: Establish functional prediction through biological engineering design at the biomolecular, cellular, and consortium scale.

Due to the dynamic complexity of biological systems, data science and machine learning approaches are needed to tackle prediction of biomolecular, cellular, and consortia properties that can affect biological function. Having access to these tools can enable numerous applications of biotechnology, though a few are provided here in example: completing this goal would allow for the techno-economic and life cycle analysis models to determine the sustainability of energy production; enhanced functional prediction can also further develop modeling and bioinformatics to predict how novel biological therapeutics may affect individual patients; and, functional predication can also facilitate the development of novel analytics tools to manipulate holistic microbial ecosystem functions by incorporating biological and environmental data for commercial systems.

Breakthrough Capability: Fully-automated molecular design from integrated, large-scale design data frameworks. This Breakthrough Capability is not meeting the pace of the predictions relative to the roadmap. Despite this, the Assessment literature review indicates the 2021 milestone has nearly been reached, though some research towards achieving it remains. However, the Assessment suggests that the 2024 milestone "Automated designs for integrated manufacturing to enable more successful, iterated workflows" and 2024 milestone "Large-scale design data generation to inform next-generation algorithms for molecular design" may not be achieved on the timeline predicted.

2021 Milestone: Structure- and comparative analysis-based libraries for automated directed evolution, with feedback of large-scale results to algorithms.

Progress toward this milestone is **significant**, with some research gaps remaining.

Directed evolution is an important strategy to engineer genes of interest to produce valuable molecules or perform practical functions. Since 2019, researchers have been examining different approaches to accelerate and simplify the directed evolution process by combining automated or machine learning technologies. For instance, *Wu et al.* (2021) described a method to incorporate machine learning processes into directed evolution workflows to predict how multiple mutations can affect the empirical fitness landscape for proteins, validating their strategy of the GB1, a human binding protein that has already undergone strenuous directed evolution processes. *Zhong et al.* (2020) created Automated Continuous Evolution (ACE), a platform that pairs Orthorep (an *in vivo,* scalable, continuous evolution system from *Ravikumar et al.* (2018)) and eVOLVER (an automated culture device from *Wong et al.* (2018) that regulates growth conditions) to directly evolve genes of interest in an automatic, feedback-controlled environmental setup. *DeBenedictis et al.* (2021) created Phage-and-Robotics-Assisted Near-Continuous Evolution (PRANCE), an "automation platform for the continuous

directed evolution of biomolecules that enables real-time activity-dependent reporter and absorbance monitoring of up to 96 parallel evolution experiments." And *Radivojević et al.* (2020) created the Automated Recommendation Tool (ART; available at https://art.lbl.gov/), which uses machine learning, sampling-based optimization, and probabilistic modeling to recommend future strains amenable to the most efficient biosynthesis pathways for making desired molecules. Several researchers have also examined how Orthorep and PRANCE can be used for different purposes, including phage-assisted continuous evolution and plant protein evolution (Miller et al., 2020 and García-García et al., 2021, respectively). The Assessment uncovered that many researchers felt that this milestone had only modest progress and that significant research gaps remained. On the contrary, several machine-learning and automated technologies that use existing directed evolution platforms, such as Orthorep and PRANCE, were discovered in literature assessments. This discordance suggests that although these tools exist, they may not be widely known by the research community, and efforts can enhance development in this space by amplifying their availability and providing training on their use.

Breakthrough Capability: Use of enzyme promiscuity prediction algorithms to design biosynthetic pathways for any molecule (natural or non-natural). The Assessment literature review indicates the 2021 milestones have not been reached and that this Breakthrough Capability is not meeting the pace of the predictions relative to the roadmap. Additionally, the 2024 milestone "Data integration for certain classes of enzymes and pathways and predictable host-specific expression in model organisms" may not be achieved on the timeline predicted by the roadmap.

2021 Milestone: Retro-biosynthesis software that can identify any biological or biochemical route to any organic molecule.

Progress toward this milestone is **modest**, with significant research gaps remaining.

With a near-infinite number of potential chemicals that engineered organisms can synthesize, researchers need software that can use metabolic pathway data to inform genetic circuit or biosynthesis design strategies. For example, *Price et al.* (2020) created GapMind, a web-based tool that annotates amino acid pathways in bacteria and archaea using precompiled, experimentally-validated databases. In their follow-up preprint article, *Price et al.* (2021) described expanding the datasets of GapMind to include biosynthesis strategies for different carbon sources, including glucosamine, citrulline, myoinositol, lactose, and phenylacetate. *Ricart et al.* (2019) created retro-biosynthetic analysis of nonribosomal peptides (rBAN), a computational tool that identifies the monomer substituents of nonribosomal peptides, therefore giving a researcher critical insights into the origin, biosynthesis, and bioactivities of the molecule. *Radivojević et al.* (2020) created the Automated Recommendation Tool (ART), which uses machine learning, sampling-based optimization, and probabilistic modeling to recommend future strains amenable to the most efficient biosynthesis pathways for making desired molecules. Developed by *von Kamp et al.* (2020), MEMO is a computational approach to find smallest metabolic modules with specific stoichiometric and thermodynamic constraints. Finally, *Lee et al.* (2021) developed stepwise classification of unknown regulation (SCOUR), a "machine learning framework that applies established algorithms to identify regulatory interactions in metabolic systems based on metabolic data," reducing the time it can take to identify and validate metabolic regulatory interactions. The Assessment uncovered that many researchers felt that this milestone had only modest progress and that significant research gaps remained. Although there are several promising technologies that, with their continued development, can undoubtedly identify biosynthetic routes for molecule synthesis, efforts to improve this technology and train researchers on these forms of technologies can enhance progress in this space.

Breakthrough Capability: Scalable, data-driven host design for complex environments that enable high-level production of natural biomolecules. This Breakthrough Capability is not meeting the pace of the predictions relative to the roadmap. The Assessment suggests that the 2024 milestones "Thematic design rules for host

system engineering inferred from data," "Tools to acquire and transfer data to a novel host to inform both genetic-domestication and prediction and determination of function," and "Novel design tools to support host design for more complex, natural (non-laboratory) environments" may not be achieved on the timeline anticipated by the roadmap, though some progress has been made in these areas.

2021 Milestone: Ability to make and screen multiple host mutations for epistasis mapping and synthetic interactions, making large-scale host optimization possible.

Progress toward this milestone is **significant**, with some research gaps remaining.

Engineering biology research must triage potential mutations or combinations thereof that can drastically enable, or limit, the production of valuable biomolecules or creation of molecules with beneficial characteristics. Several platforms and tools have emerged since 2019 that couple machine-learning and automation technologies toward mutation mapping. For example, *Wu et al.* (2021) described a method to incorporate machine learning processes into directed evolution workflows to predict how multiple mutations can affect the empirical fitness landscape for proteins, validating their strategy of the GB1, a human binding protein that has already undergone strenuous directed evolution processes. *Iwai et al.* (2018) presented an automated flow-based/digital microfluidic platform that incorporates multiplex electroporation for genome editing and dual optical detection of fluorophores, thereby enabling a very efficient automated screening platform to recover mutated strains. *Biwas et al.* (2021) engineered a machine-learning tool that can use a limited number of functionally assayed mutant sequences (as few as 24) to build a fitness landscape to screen potential mutations *in silico* for protein synthesis. Finally, *Shroff et al.* (2020) developed a deep-learning tool that uses a 3D convolutional neural network that associates amino acids with their micro-chemical-environment to help identify desired gain-of-function mutations for protein engineering. The creation of several different tools that use automation, machine-, or deep-learning technologies to measure how single or multiple mutations affect biosynthesis shows that significant progress has been made towards this goal. Researchers can further progress in this space by providing standardized datasets and experimental workflows for potential users of these tools to gain familiarity with the technology.

2021 Milestone: Better data on physiology and fitness in deployment environments suitable for informing designs in validated lab-scale simulations that meet activity, persistence, and ecological impact goals.

Progress toward this milestone is **modest**, with significant research gaps remaining.

Biological systems will be exposed to a wide variety of diverse environments that can drastically affect physiology and fitness. Researchers aim for their biological systems to have robust activity even in the face of environmental pressure, and they look to lab simulations to test their systems and inform design strategies. Since 2019, research has focused on developing platforms, frameworks, and toolsets that can aid researchers in evaluating fitness and physiological conditions. *Lui et al.* (2021) introduced the Framework for Integrated, Conceptual, and Systematic Microbial Ecology (FICSME), incorporating diverse data types to discern a microbial system's biological, chemical, and physical drivers to understand how they affect the local ecosystem. *Thompson et al.* (2019) demonstrated how random barcode transposon sequencing (RB-TnSeq) can measure the metabolic fitness profiles of thousands of genes in parallel by a demonstration in *Pseudomonas putida* and discovered critical pathway enzymes for lysine metabolism. Similarly, *Thorgersen et al.* (2021) used RB-TnSeq and activity-based metabolomics to uncover how contaminant metal particles, including the cation Al^{3+}, the oxyanion CrO_4^{2-}, and the oxycation UO_2^{2+}, affected the fitness of the metal-tolerant facultative anaerobe *Pantoea* sp. strain MT58, providing an in-depth examination of how bacteria are affected by environmental metals. In their preprint, *Henriques et al.* (2021) compared the metabolism of *Saccharomyces cerevisiae* and *Saccharomyces uvarum* in a wine fermentation setting using a new modeling framework that incorporates both genomic, metabolomic, and kinetic data to better understand physiological capabilities.

Many researchers have also examined how they can further exploit different metabolic conditions and processes to stimulate production from biological systems. For instance, *Du et al.* (2019) manipulated the circadian metabolism of the cyanobacterium *Synechocystis* sp. PCC 6803 (using their Find Reactions Usable in Tapping Side-Products (FRUITS) algorithm) to design a growth- and fitness-coupled strategy that can produce fumarate around the clock. Technologies such as RB-TnSeq and frameworks such as FICSME have shown promising approaches that can better determine how environmental conditions can affect biological system fitness and *vice versa*. Although this is a good start, there is much room to create novel, more universal toolsets that can characterize the fitness of synthetic systems and immediately translate that information to inform design strategy.

Breakthrough Capability: Enabled design of functional, self-supporting ecosystems. This Breakthrough Capability is not meeting the pace of the predictions relative to the roadmap and the Assessment literature review indicates that the 2021 milestones have not been reached.

2021 Milestone: Data-driven tools for selecting organisms for synthetic assemblies to achieve resistant, resilient activity.

Progress toward this milestone is **modest**, with significant research gaps remaining.

Biological organisms exhibit many characteristics that can affect their compatibility with one another in an engineered ecosystem. As more organisms are added to an assembly, or if modifications to the environment occur, an exponentially increasing number of interactions must be accounted for. Machine learning, artificial intelligence, and simulations are best equipped to parse through these complex datasets and identify organisms that can achieve resilient, robust activity within an ecosystem. Since 2019, research has primarily focused on creating tools and repositories to better inform organism selection for practical characteristics. For instance, *Radivojević et al.* (2020) developed the Automated Recommendation Tool (ART), which uses machine learning, sampling-based optimization, and probabilistic modeling to recommend future strains amenable to the most efficient biosynthesis pathways for making desired molecules. *Seaver et al.* (2020) described the release of the ModelSEED biochemistry base, which encompasses valuable information for annotations, constructions, comparisons, and analyses of metabolic models for fungi, plants, and microbes, impressively accumulating data on over 33,978 compounds and 36,645 reactions of metabolic pathways. *Shroff et al.* (2020) developed a deep-learning tool that uses a 3D convolutional neural network to associate amino acids with their micro-chemical-environment to help identify advantageous gain-of-function mutations that promote biosynthesis. And *Roy et al.* (2021) presented a step-by-step tutorial on how to store, visualize, and leverage multi-omic data to predict the outcomes of engineering biology experiments, demonstrating their workflow by correctly predicting and validating a strain that increases isoprenol production by 23%. In a specific demonstration of integrated multi-omic analysis, *Pomraning et al.* (2021) used proteomic and metabolomic measurements to examine how to improve the ability of the filamentous fungus *Aspergillus pseudoterreus* to enhance the production of 3-hydroxypropionic acid (3HP) production, a valuable polymer precursor, identifying a number of metabolic pathways and co-products that impacted 3HP production. Finally, *Chen et al.* (2019) created an automated "cells-to-peptides" sample preparation workflow, including cell lysis, protein precipitation, resuspension, quantification, normalization, and tryptic digestion, to assay the proteomes of gram-negative bacteria and fungi in high-throughput (up to 96 samples from cell pellets to the initiation of the tryptic digestion step in two hours). Although there are several promising tools that, with their continued development, can undoubtedly assist researchers in identifying organisms for robust assembly design, several gaps remain towards broad application of these technologies.

2021 Milestone: Direct data collection for the most important communities in human, agriculture, and complex bioreactor work sufficient for informing design.

Progress toward this milestone is **modest**, with significant research gaps remaining.

Organisms and biological systems are inherently complex and dynamic, reliant on inputs and interactions from their surroundings. Transferring these organisms into a lab for study often disturbs their natural reactions and processes. Thus, tools need to be developed to study organisms in their natural environments and collect data so as to recapitulate those environments for further engineering. *Zengler et al.* (2019) discussed the importance of developing fabricated microbial ecosystems (EcoFABs) that have standardized workflows, computational tools, data standards, and computational models to aid reproducible analysis of novel microbial communities. To this end, *Lui et al.* (2021) introduced the Framework for Integrated, Conceptual, and Systematic Microbial Ecology (FICSME) for incorporating diverse data types to discern a microbial system's biological, chemical, and physical drivers to understand how they affect the local ecosystem, critically offering guidance on how researchers can better pursue field studies. There have been many examples of researchers collecting and characterizing microbes from different environments and offering a repository of their multi-omics data. *Nayfach et al.* (2021) published a genomic catalog on over 10,000 microbiome genomes of different terrain habitats and oceans and demonstrated the use of this information to identify species amenable to valuable secondary metabolite synthesis. *Danko et al.* (2021) cataloged over four thousand metagenomic samples from urban microbiomes and identified which have antimicrobial resistance genes in the first worldwide catalog of the urban microbial ecosystem. In their perspective article, *Brooks and Alper* (2021) analyzed platforms that better equip researchers performing engineering biology research outside of the laboratory in resource-limited or off-the-grid scenarios, especially for bioproduction, biosensing, and closed-loop therapeutic delivery analyses. And *Wilpiszeski et al.* (2020) used in-field bioreactors to examine how geochemical conditions (in this case, contaminated groundwater in Oak Ridge Reservation, TN) affected microbial communities by collecting their DNA for 16S rRNA amplicon sequencing and assaying their cell counts, total proteins, anions, cations, trace metals, organic acids, bicarbonate, pH, oxidation/reduction potential, dissolved oxygen, and conductivity. There is a lot of growing enthusiasm for the potential of conducting more field studies to collect untraditional microbes for engineering biology contexts directly. Although there have been many discussions on the current capabilities and frameworks to perform these procedures correctly, development in this space can be improved by actual execution, especially if the purpose of the study is to examine the feasibility of field-deployed bioreactors.

2021 Milestone: Modeling tools to identify cross-organismal networks and ecological interactions.

Progress toward this milestone is **modest**, with significant research gaps remaining.

The environmental and ecological interactions of an (engineered) organism can be complex and require intensive data science tools to disentangle and process the many variables involved with their characterization. *Ibrahim et al.* (2021) provides a review of the microbial interactions in microbial communities as well as modeling approaches for biotechnological applications. Since 2019, several researchers have focused on creating tools to identify cross-organismal networks and interrogate aspects of their ecological interactions. For example, *Liao et al.* (2020) developed a validated framework that models community dynamics and metabolic exchanges of multi-strain *Escherichia coli* communities, presumably allowing other researchers to quantify cross-feeding interactions in ecosystems better. *Diener et al.* (2020) presented MICOM, a "customizable metabolic model of the human gut microbiome" that allows researchers to better infer how a microbial community corresponds to ecosystem function, demonstrating their technology by uncovering differences in metabolic interaction networks between healthy and diabetic individuals. *Kosina et al.* (2021) introduced Biofilm Interaction Mapping and Analysis (BIMA), a tool that helps deconstruct interspecific interactions in biofilm co-cultures or consortia; they demonstrated the power of their tool by identifying four genes of

importance that are discordant between *Pseudomonas stutzeri* RCH2 (a strain found in chromium-contaminated soils) and other *Pseudomonas* strains. Additionally, databases and repositories continue to develop and collect many metabolic and ecological analyses important for cross-organism interactions. Seaver et al. (2020) described the release of the ModelSEED biochemistry base, which encompasses valuable information for annotations, constructions, comparisons, and analyses of metabolic models for fungi, plants, and microbes, impressively accumulating data on over 33,978 compounds and 36,645 reactions of metabolic pathways. Danko et al. (2021) cataloged over four thousand metagenomic samples from urban microbiomes and identified antimicrobial resistance genes in the first worldwide catalog of the urban microbial ecosystem. Baldini et al. (2019) created a toolbox to model microbe-microbe and host-microbe metabolic interactions and microbial communities using genome-scale metabolic reconstructions and metagenomics data. Dukocski et al. (2021) demonstrated the use of the most recent version of computation of microbial ecosystems in time and space (COMETS), which takes modular environmental and biochemical inputs to simulate the spatiotemporal dynamics of the ecosystem. Although there are promising starts to tools that enable the study of cross-organismal networks and ecological interactions in multi-strain *Escherichia coli* communities, human gut microbiomes, and biofilms, there is still a significant need to develop inclusive toolsets that can detangle interactions across biological systems, particularly those comprising an engineered component.

Data Science Goal: Establish optimal manufacturing processes from the unit-operation to the integrated-screening scale.

Although engineering biology encompasses the core characteristic of discovering scientific knowledge, it must also contend with engineering principles such as scale-up, design, and interoperability. Optimal manufacturing processes are critical to these hallmarks, from unit-by-unit operation to fully integrated screening scales. Accomplishing this goal can enable many applications and impacts universal to every sector, though it has a tremendous significance in industrial biotechnology. One example is the better prediction of media components, additives, and environmental conditions that promote the growth of non-model production hosts from genomic data. Another example is automation being able to screen new candidate hosts for fast growth and desired production rates in industrial biomanufacturing settings. Additionally, enabling artificial intelligence or machine learning approaches can help researchers to predict how to assemble systems under tight production goals and constraints.

Breakthrough Capability: Standardized informatics tools, data, and automation platforms for efficient and collaborative use and integration of data in order to develop novel products more quickly. This Breakthrough Capability is proceeding ahead of schedule relative to the roadmap. The Assessment literature review indicates that the 2021 milestone has been reached, and in contrast to many other Data Science milestones, the Assessment suggests that all other milestones under this Breakthrough Capability may be achieved ahead of schedule.

2021 Milestone: Establish communications and networks to develop democratized platforms for data exchange and automation across industry and academia.

Progress toward this milestone is **significant**, with some research gaps remaining.

Accessible platforms that promote the exchange of data, workflows, and automation between industrial, academic, and government researchers are critical for promoting a healthy engineering biology enterprise. The difficulty, however, is that different stakeholders may have the incentive to withhold public access to their resources, or democratized platforms/repositories for sharing information may not exist. Since 2019, engineering biology practitioners have begun to form foundries, coalitions, and platforms to better communicate the importance of sharing information. In their comment article, Hillson et al. (2019) discussed the importance of biofoundries to enable rapid design, construction, and testing of organisms and subsequent data for biotechnology purposes. As a direct example of the services biofoundries can provide, the Edinburgh

Genome Foundry (2020) continues to develop Plateo, a python library used to assist in planning, running, and checking microplate laboratory experiments. In progress towards data automation, *Steel et al.* (2020) introduced Chi.Bio, a parallelized, open-source platform where researchers can automate the measurement and control of culturing in bulk for *in vivo* biological studies. Similarly, *Storch et al.* (2020) developed DNA-BOT, an open-source software package that automates the building of extensive DNA construct libraries for synthetic biology. Coalitions and platforms continue to develop that enable communication and networks to promote industrial processing. For example, launched in 2021, the Bioindustrial Manufacturing and Design Ecosystem (BioMADE) is a Manufacturing Innovation Institute and part of the Manufacturing USA network, and "is working to build a sustainable, domestic end-to-end bioindustrial manufacturing ecosystem that will enable domestic bioindustrial manufacturing at all scales, develop technologies to enhance U.S. bioindustrial competitiveness, de-risk investment in relevant infrastructure, and expand the biomanufacturing workforce to realize the economic promise of industrial biotechnology." *Dileo et al.* (2022) discussed the importance of "a network of interoperable, highly automated, and interconnected research facilities at the local, regional, and national levels (a BioNet) that will enable rapid execution of projects through coordinated efforts, produce a fully developed biology as-technology ecosystem, and enhance equity by making cutting edge technologies for engineering biology available to researchers that would otherwise not have such access." Although this milestone, in its nature, is subjective towards what constitutes 'established communications and networks,' researchers generally feel much enthusiasm about the several communities and biofoundries being founded to begin to address data exchange and automation issues. Further development in this space must start to produce technical solutions, standardized practices, and agreements for data sharing and automation.

Part 2: Social and Nontechnical Dimensions to Advance Engineering Biology

While evaluating progress towards the milestones in *Engineering Biology*, Assessment contributors were asked to identify nontechnical (i.e., not directly the technical practice of engineering biology) barriers and social and economic considerations that affect or impact the progress of their research. With the goal of engineering biology to address national and global challenges, it is important to make investments and strategies to overcome barriers and to incorporate or address nontechnical considerations early and often. Barriers to progress highlighted by contributors included emergent challenges related to the COVID-19 pandemic, including the availability of research supplies and limits to collaboration and in-person activities. While the impacts of COVID-19 continue to lessen, they still present vulnerabilities to conducting research and to advancing tools and products of engineering biology. Other barriers are more persistent, including insufficient regulatory clarity and muddled approval processes for genetically engineered organisms, minimal data science education and training to support engineering biology practice, and a deficiency of diverse perspective and engagement in the field. Furthermore, there are several nontechnical dimensions that are, or could be, limiting to research and innovation as envisioned by the roadmap and beyond, or that can influence the decisions made about which research to pursue and how. Adequate investment, infrastructure, and resources are needed to support research, education, and workforce development. Clear and nimble policy and regulations can ensure that the most impactful tools and technologies make it into the economy. And to ensure that these products and solutions are beneficial for humanity and the planet, we must incorporate effective and proactive risk assessment and security and safety considerations. Underpinning this all is an inclusive and equitable research enterprise. We advocate on behalf of routine and consistent incorporation of these nontechnical dimensions, because without the support and influence of sufficient resources, effective policy, and broad engagement, the best advancements in engineering biology cannot be realized.

Impacts of the COVID-19 Pandemic

Stressed Supply Chains and Capabilities

Many standard laboratory supplies, such as pipettes, genetic extraction kits, chemicals, buffers, and gloves, are necessary for engineering biology research. During the COVID-19 pandemic, these supplies were redirected toward SARS-CoV-2 diagnostic testing, limiting their availability for everyday use. Although the scientific community recognized that these supplies were urgently needed to curb the effects of the pandemic, the lack of these essential supplies severely hampered and delayed ongoing research. It was especially difficult for researchers to obtain supplies from international markets (Woolston, 2021). Even nearly three years since the beginning of the pandemic, the Food and Drug Administration's research supply availability list still describes shortages of several essential research supplies, including, including general purpose reagents, gloves, tubes and pipettes (U.S. Food and Drug Administration). Strategies and efforts to better secure research supply chains, both immediate and long-term, will help to curb the potential of future delays.

Limits to In-Person Interaction

In their perspective article on the impact to scientific careers, *Woolston (*2021) surveyed over 3,500 researchers to ask how COVID-19 affected their work. The top five responses were challenges in: discussing ideas with advisors, collecting data, collaborating internally, conducting laboratory experiments, and supervising colleagues. The COVID-19 quarantine and social distancing policies prevented and/or limited many engineering biology researchers from being able to physically go into their laboratory space to perform experiments. Further, there were significant cancellations and on-going reductions in the number of scientific conferences and events, which scientists use as opportunities to form or advance collaborations, discover new science, and build career networks. Although these restrictions were necessary for public health and individual safety, and video-conferencing and other social media and web-based collaboration platforms have been used to

ameliorate the impacts (Kobel and Stegle, 2020), limitations on in-person participation in research and related activities has resulted in widespread setbacks to active research.

While this is anticipated to be a temporary disruption, albeit with lasting impacts, a potential solution to overcome future similar circumstances is to identify experimental workflows and develop new technologies that allow engineering biology experiments to be done remotely, including through advanced automation. Identifying strategies and implementing practices that support efficient and safe research and more meaningful mentorship and collaboration through virtual or remote interactions will make the field more robust to future pandemics or other disruptions and may have added benefits of making research and collaborations more accessible to people with travel or movement limitations or restrictions.

Policy for Engineering Biology Advancement

Regulatory Clarity and Streamlined Approval Pipelines for Genetically Engineered Organisms

To ensure that any engineered biotechnologies are safe for the public, environment, and society, government agencies are charged with regulating commercial products and processes. In the United States, genetically engineered organisms can be regulated by the USDA, EPA, FDA, and other bodies depending on the organism, modifications made, and its application (National Academies of Sciences and Engineering, 2016). While regulations have become more clear around engineered genes and products in plants, there is still significant uncertainty around genetically engineered microbes. Particularly in industry, many researchers are proactive in identifying procedures and precautions for consumer use, but the lack of precise guidance remains a critical barrier to progress. Regulatory agencies encourage innovators to meet with them early and often during product development to streamline the regulatory process; however, navigating these agencies and the offices within can be challenging, as it can be unclear whom to contact, particularly for truly new products. Proactive efforts to identify future regulatory concerns and cross-agency task forces to comprehensively examine biotechnology regulations can greatly promote safe and clear guidelines (National Academies of Sciences and Engineering, 2016). Progress is being made, such as the newly formed Unified Website for Biotechnology Regulation. Still, precise and clear guidelines accessible to researchers during product development could streamline processes for industry and ease the burden on regulators.

Strategies for Policy and Investment

The roadmap envisions numerous tools and technologies to help address pressing national and societal challenges; however, none of this can be accomplished without effective policies and adequate funding for research, education, and infrastructure. Through grant award mechanisms, institutional and individual outreach, and other avenues, engineering biology researchers have the task of informing policymakers and science funders about the potential applications of their technologies in a way that best serves public or funder-inspired goals. One such avenue that has been pursued is technical research roadmapping. EBRC's roadmaps serve to speak on behalf of the contributors what the community sees as future areas of importance and innovation that would benefit from – or might only be realized with – public or private investment or supportive policies and regulations. EBRC has published several subsequent research roadmaps across the engineering biology landscape, including roadmaps on the specific topics of microbiomes (EBRC, 2020), materials (EBRC, 2021), and for climate and sustainability (EBRC, 2022). While these technical roadmaps do not extensively identify areas of needed investment or lay out explicit policy recommendations, they serve as a reference point for policymakers and federal granting agencies, among others, where investment might be anticipated to have significant impact.

Other groups and institutions have also highlighted the importance of strategic investment in engineering biology, how policy can incentivize or limit innovation, and/or how advancements in engineering biology can contribute to bioeconomic growth. *Kitney et al.* (2019) examined the importance of policy actions toward advancing a strategic, sustainable bioeconomy, explicitly discussing how public-private biofoundries de-risk

research investment, investment in breakthrough technologies will provide the most benefit to the field, and the importance of harmonized technical standards between academic, government, and industry for innovative progress. More recently, the Schmidt Futures Bioeconomy Task Force released *The U.S. Bioeconomy: Charting a Course for a Resilient and Competitive Future,* which goes into detail on what it would take in the U.S. and worldwide to maximize the benefits of the bioeconomy, including investment and advancements in research, workforce development, regulatory standards, and manufacturing capacity (Hodgson et al., 2022). A 2020 report from McKinsey & Company outlined the innovative potential for biotechnology and engineering biology research to address global challenges, create a circular economy, and secure U.S. supply chains (Chui et al., 2020). Notably included in this report are quantitative metrics about the potential impact of engineering biology on the economy, including a claim that "60 percent of the physical inputs to the global economy could, in principle, be produced biologically." Finally, in a 2022 report from MITRE, *Dileo et al.* (2022) describe several policy initiatives to advance a competitive U.S. bioeconomy, including a proposed network (a BioNET) to help standardize manufacturing processes and help democratize research practices for engineering biology practitioners.

Security and Safety in Engineering Biology

While working toward the broad and far-reaching benefits of engineering biology, consideration should also be given to potential nefarious, accidental, or unintended outcomes of its development and/or use. To ensure the benefit of biotechnologies can be realized while minimizing and mitigating any associated risks, stakeholders throughout the engineering biology community should strive to integrate best safety and security practices proactively and preemptively into their work and institutions. In addition to physical security and laboratory practices, researchers and other stakeholders should reserve time to recognize, consider, and discuss how engineering biology tools could inadvertently or intentionally be used to cause harm to people or the planet. Doing so fosters and upholds a culture where safety and security practices are not simply complied with, but where members of the field intellectually engage with the implications of their own work. In such an environment, members of the field can work together to identify innovative approaches to governance and research needs for preventing and mitigating undesirable outcomes without hampering progress toward solutions to major societal challenges.

Incorporating Risk Evaluation into Engineering Biology Research and Development

The continually evolving landscape of engineering biology necessitates ongoing discussion and evaluation of governance mechanisms and a willingness to experiment with new risk evaluation and management approaches (Evans et al., 2020). Several papers outline approaches or tools for considering the potential positive and negative consequences and implications of research, although it is challenging to identify the extent to which they are used by the research community. *Cummings and Kuzma* (2017) reported the development of a societal risk evaluation scheme (SRES) that can improve a researcher's ability to anticipate the risks of synthetic biology products. Along with incorporating the typical risk-benefit factors of environmental and health consequences, this scheme also includes reversibility, manageability, expected levels of public concern, and uncertainty. *Burgiel et al.* (2021) discussed the proceedings of a workshop dedicated to reviewing the planning and implementation of genetic interventions, including those from synthetic biology, involved with conservation efforts. They suggested that such interventions have several attributes—such as the severity of unwanted outcomes and the degree of certainty that the desired outcome will be achieved— that can be identified and scored from low/least to high/most concern and weighted to set acceptability limits. *Trump et al.* (2021) recommended that social scientists with diverse expertise be engaged early in the process of technology development to help assess elements of risk. Such involvement necessitates some transparency from practitioners but can enable the identification and minimization of any potential downstream harms. Doing so aligns with and fulfills guiding principles in the Statement of Ethics for Engineering Biology Research released by EBRC, which suggests that engineering biology stakeholders should “seek to create products or

processes that benefit people, society, or the environment" and "consider and weigh the benefits of research against potential harms" (Mackelprang et al., 2021). Future work might look to evaluate if, how, or when such tools and approaches are implemented and the impact they have on researcher decisions and outcomes.

Building a Culture that Prioritizes Safety and Security

The nature and magnitude of safety and security hazards resulting from engineering biology research and the development of associated products may change as a given technology develops and as other tools and technologies expand or narrow vulnerabilities and the ease of their exploitation. Therefore, research practitioners should evaluate their research for safety and security concerns on an on-going basis, building a generative safety and security culture within the field (National Research Council, 2014). To support the development of such a culture, EBRC developed and hosts "Malice Analysis" workshops, which "train researchers and others associated with engineering biology to critically evaluate research for potential security concerns." EBRC has also suggested that such a culture can be fostered by prompting researchers to consider the security implications of their work periodically through the research lifecycle, such as at the publication stage (Mackelprang et al., 2022).

Importantly, the engineering biology stakeholder community should recognize the distinctions between safety and security and attend to each. While many safety and security considerations overlap, there are important differences in the potential impacts of each and the prevention and mitigation efforts that may be appropriate. Thus, while it is convenient to discuss them together, the identification and implementation of best practices and standards should consider and attend to each.

Multidisciplinary Engineering Biology Education

Engineering biology requires an understanding of chemistry, biomolecular physics and signal processing, cellular biology, and bioinformatics and data science, among other concepts, to engineer useful features into living systems. Despite this, most education, even at the advanced undergraduate and graduate levels, is siloed by discipline. To enable advancements and innovations within fundamental research, much less application and product development, there is a need to better educate future engineering biology researchers and biotechnology leaders across fields. Incentives to develop programs and curricula, with both formal and informal training opportunities, that incorporate cross-disciplinary learning and experience will help to meet this need.

Data Science Education and Training

As an example with particular relevance to areas where the Assessment indicates we are falling behind, a dearth of curriculum, instruction material, and resources for engineering biology data science remains a critical barrier for trainees (Delebecque and Philip, 2015). Traditional laboratory methods, namely trial and error, can make the process of finding effective perturbations and changes a time-consuming process; machine learning and data science methods can accelerate the manipulation of organisms and general progress of research. Many graduate students and postdoctoral trainees wish to use data science methods to advance their research, but a lack of formal education in data science, or open-source resources specific to engineering biology, leaves many unable to acquire the needed training. Preparing more instructors to teach these disciplines (Emery et al., 2021), as well as the development of curricula, can help address this need.

Engagement

Strategies to Incorporate Social Science into Technical Research

To fully realize the potential of engineering biology to address national and global needs, the research community must consider the broader social dimensions of research products and processes. As innovations in engineering biology become commercial products and applications, facets such as societal impact, technology

accessibility, and responsible research grow in importance. For the field to maximally impact many global challenges, the technical research community will benefit from collaboration and insight from the social sciences. However, when responding to questions about incorporating social and nontechnical dimensions into their research, many technical researchers indicated significant interest but often did not know what forms of social science expertise or mediums of engagement were appropriate or possible, nor the timing for initiating such engagement.

Recently, there have been several critical examinations of engagement and best practices between the social and technical research communities. In a review of the co-evolution of synthetic biology and the social sciences, *Trump et al.* (2019) note how these fields intersected early on in the establishment of engineering biology-related research, with the social sciences providing expertise in addressing risk assessment, governance, and public engagement needs. Now and in the future, social scientists and technical researchers should continue to work in partnerships to address these needs as co-producers of knowledge invested in responsible innovation (Balmer et al., 2015). There are, however, inherent and perceived challenges with integrating social science into the technical research process. For example, *Taylor and Woods* (2019) interviewed senior scientists involved with synthetic biology projects and report that many interpret the construct of "Responsible Research and Innovation" as risk-avoidance. Such a narrow view fails to recognize the enormous benefits of incorporating social dimensions and nontechnical considerations into technical research. To such an end, consortia and networks of public, academic, industry, and government stakeholders can work together to identify effective strategies for integrating the social science and technical communities.

Diversity, Equitable Opportunities, and an Inclusive Culture in Engineering Biology

In order to achieve the enormous potential of engineering biology imagined by the roadmap, the research process and development of products must reflect the diversity of individuals that will be impacted by the resulting biotechnologies. The research community recognizes that some communities remain underrepresented in engineering biology. However, there is overwhelming interest in increasing overall diversity, equity, inclusivity, and accessibility (DEIA) within the field and engaging individuals from underrepresented communities, particularly early in the education pipeline. Still, many persistent barriers to participation must be broken down, including access to engineering biology education and training, a paucity of diverse mentors, and historical entrenchment of perspectives on who does and does not belong in the academic and research enterprise. These are systemic challenges that go far beyond engineering biology, but individual and collective actions can still make a meaningful difference toward overcoming these barriers.

Funding and incentives for engineering biology education, training, and research opportunities can focus on underserved populations, including for Historically Black Colleges and Universities (HBCUs) and Minority-Serving Institutions (MSIs). Engineering biology researchers and leaders can engage with coalitions promoting underrepresented trainees in STEM, including conferences and organizations like AfroBiotech, the Annual Biomedical Research Conference for Minority Students (ABRCMS), and the Society for the Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS), in order to promote awareness and increase engagement. And established engineering biology institutions and organizations, including EBRC, can ensure that their members, participants, and collaborators are recognizing the importance of DEIA, and taking continuous, active steps to increase representation.

References

Aakko, J., Pietilä, S., Suomi, T., Mahmoudian, M., Toivonen, R., Kouvonen, P., Rokka, A., Hänninen, A., & Elo, L. L. (2020). Data-Independent Acquisition Mass Spectrometry in Metaproteomics of Gut Microbiota—Implementation and Computational Analysis. *Journal of Proteome Research*, *19*(1), 432–436. https://doi.org/10.1021/acs.jproteome.9b00606

Adhikari, A., Raj Bhattarai, B., Aryal, A., Thapa, N., Kc, P., Adhikari, A., Maharjan, S., B. Chanda, P., P. Regmi, B., & Parajuli, N. (2021). Reprogramming natural proteins using unnatural amino acids. *RSC Advances*, 11(60), 38126–38145. https://doi.org/10.1039/D1RA07028B

Alam, K., Hao, J., Zhang, Y., & Li, A. (2021). Synthetic biology-inspired strategies and tools for engineering of microbial natural product biosynthetic pathways. *Biotechnology Advances*, *49*, 107759. https://doi.org/10.1016/j.biotechadv.2021.107759

Alnahhas, R. N., Winkle, J. J., Hirning, A. J., Karamched, B., Ott, W., Josić, K., & Bennett, M. R. (2019). Spatiotemporal Dynamics of Synthetic Microbial Consortia in Microfluidic Devices. *ACS Synthetic Biology*, *8*(9), 2051–2058. https://doi.org/10.1021/acssynbio.9b00146

Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., & Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biology*, *21*(1), 30. https://doi.org/10.1186/s13059-020-1935-5

Angenent-Mari, N. M., Garruss, A. S., Soenksen, L. R., Church, G., & Collins, J. J. (2020). A deep learning approach to programmable RNA switches. *Nature Communications*, 11(1), Article 1. https://doi.org/10.1038/s41467-020-18677-1

Antczak, M., Zablocki, M., Zok, T., Rybarczyk, A., Blazewicz, J., & Szachniuk, M. (2019). RNAvista: A webserver to assess RNA secondary structures with non-canonical base pairs. *Bioinformatics*, *35*(1), 152–155. https://doi.org/10.1093/bioinformatics/bty609

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., & Liu, D. R. (2019a). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, *576*(7785), 149–157. https://doi.org/10.1038/s41586-019-1711-4

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., & Liu, D. R. (2019b). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, *576*(7785), Article 7785. https://doi.org/10.1038/s41586-019-1711-4

Aquino, A. K., Manzer, Z. A., Daniel, S., & DeLisa, M. P. (2021). *Glycosylation-on-a-chip: A flow-based microfluidic system for cell-free glycoprotein biosynthesis* (p. 2021.09.24.461663). bioRxiv. https://doi.org/10.1101/2021.09.24.461663

Armenteros, J. J. A., Salvatore, M., Emanuelsson, O., Winther, O., Heijne, G. von, Elofsson, A., & Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance*, *2*(5). https://doi.org/10.26508/lsa.201900429

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., van Dijk, A. A., Ebrecht, A. C., … Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, *373*(6557), 871–876. https://doi.org/10.1126/science.abj8754

Baldini, F., Heinken, A., Heirendt, L., Magnusdottir, S., Fleming, R. M. T., & Thiele, I. (2019). The Microbiome Modeling Toolbox: From microbial interactions to personalized microbial communities. *Bioinformatics*, *35*(13), 2332–2334. https://doi.org/10.1093/bioinformatics/bty941

Baldwin, G. (2020). *S*tandards in synthetic biology: Gaps, challenges and opportunities. *BioRoboost*. https://standardsinsynbio.eu/wp-content/uploads/2020/09/Deliverable-1.1.pdf

Beale, S. (2019, October 21). EBRC Report Offers a 20-Year Synthetic Biology Roadmap That Could Lead to New Diagnostic Technologies for Clinical Laboratories, Pathologists. *Dark Daily*. https://www.darkdaily.com/2019/10/21/ebrc-report-offers-a-20-year-synthetic-biology-roadmap-that-could-lead-to-new-diagnostic-technologies-for-clinical-laboratories-pathologists/

Beckmann, J., Lehr, F., Finazzi, G., Hankamer, B., Posten, C., Wobbe, L., & Kruse, O. (2009). Improvement of light to biomass conversion by de-regulation of light-harvesting protein translation in Chlamydomonas reinhardtii. *Journal of Biotechnology*, 142(1), 70–77. https://doi.org/10.1016/j.jbiotec.2009.02.015

Belcher, M. S., Vuu, K. M., Zhou, A., Mansoori, N., Agosto Ramos, A., Thompson, M. G., Scheller, H. V., Loqué, D., & Shih, P. M. (2020). Design of orthogonal regulatory systems for modulating gene expression in plants. *Nature Chemical Biology*, *16*(8), Article 8. https://doi.org/10.1038/s41589-020-0547-4

Bell, N. M., Mankowska, S. A., Harvey, S. A., & Stemple, D. (2019). *A highly accurate, enzymatic, de novo synthesis and gene assembly technology*. https://www.semanticscholar.org/paper/A-highly-accurate%2C-enzymatic%2C-de-novo-synthesis-and-Bell-Mankowska/62c90d5f904c1c7b7e1af9ac9fe3f7291bca101a

Ben-Moshe, S., Veg, T., Manco, R., Dan, S., Papinutti, D., Lifshitz, A., Kolodziejczyk, A. A., Bahar Halpern, K., Elinav, E., & Itzkovitz, S. (2022). The spatiotemporal program of zonal liver regeneration following acute injury. *Cell Stem Cell*, *29*(6), 973-989.e10. https://doi.org/10.1016/j.stem.2022.04.008

Bernabé-Orts, J. M., Casas-Rodrigo, I., Minguet, E. G., Landolfi, V., Garcia-Carpintero, V., Gianoglio, S., Vázquez-Vilar, M., Granell, A., & Orzaez, D. (2019). Assessment of Cas12a-mediated gene editing efficiency in plants. *Plant Biotechnology Journal*, *17*(10), 1971–1984. https://doi.org/10.1111/pbi.13113

Berhanu, S., Ueda, T., & Kuruma, Y. (2019). Artificial photosynthetic cell producing energy for protein synthesis. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-09147-4

Bevacqua, R. J., Dai, X., Lam, J. Y., Gu, X., Friedlander, M. S. H., Tellez, K., Miguel-Escalada, I., Bonàs-Guarch, S., Atla, G., Zhao, W., Kim, S. H., Dominguez, A. A., Qi, L. S., Ferrer, J., MacDonald, P. E., & Kim, S. K. (2021). CRISPR-based genome editing in primary human pancreatic islet cells. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-22651-w

Biswas, S., Khimulya, G., Alley, E. C., Esvelt, K. M., & Church, G. M. (2021). Low-N protein engineering with data-efficient deep learning. *Nature Methods*, *18*(4), Article 4. https://doi.org/10.1038/s41592-021-01100-y

Blackiston, D., Lederer, E., Kriegman, S., Garnier, S., Bongard, J., & Levin, M. (2021). A cellular platform for the development of synthetic living machines. *Science Robotics*, *6*(52), eabf1571. https://doi.org/10.1126/scirobotics.abf1571

Bottaro, S., Bussi, G., & Lindorff-Larsen, K. (2021). Conformational Ensembles of Noncoding Elements in the SARS-CoV-2 Genome from Molecular Dynamics Simulations. *Journal of the American Chemical Society*, *143*(22), 8333–8343. https://doi.org/10.1021/jacs.1c01094

Brassard, J. A., Nikolaev, M., Hübscher, T., Hofer, M., & Lutolf, M. P. (2021). Recapitulating macro-scale tissue self-organization through organoid bioprinting. *Nature Materials*, *20*(1), Article 1. https://doi.org/10.1038/s41563-020-00803-5

Brooks, S. M., & Alper, H. S. (2021). Applications, challenges, and needs for employing synthetic biology beyond the lab. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21740-0

Brophy, J. A. N., Triassi, A. J., Adams, B. L., Renberg, R. L., Stratis-Cullum, D. N., Grossman, A. D., & Voigt, C. A. (2018). Engineered integrative and conjugative elements for efficient and inducible DNA transfer to undomesticated bacteria. *Nature Microbiology*, *3*(9), Article 9. https://doi.org/10.1038/s41564-018-0216-5

Brown, B., Bartley, B., Beal, J., Bird, J. E., Goñi-Moreno, Á., McLaughlin, J. A., Mısırlı, G., Roehner, N., Skelton, D. J., Poh, C. L., Ofiteru, I. D., James, K., & Wipat, A. (2020). Capturing Multicellular System Designs Using Synthetic Biology Open Language (SBOL). *ACS Synthetic Biology*, *9*(9), 2410–2417. https://doi.org/10.1021/acssynbio.0c00176

Burdette, L. A., Wong, H. T., & Tullman-Ercek, D. (2021). An optimized growth medium for increased recombinant protein secretion titer via the type III secretion system. *Microbial Cell Factories*, *20*(1), 44. https://doi.org/10.1186/s12934-021-01536-z

Burnett, J. C., & Rossi, J. J. (2012). RNA-based therapeutics: Current progress and future prospects. *Chemistry & Biology*, 19(1), 60–71. https://doi.org/10.1016/j.chembiol.2011.12.008

Cao, Z., Yu, J., Wang, W., Lu, H., Xia, X., Xu, H., Yang, X., Bao, L., Zhang, Q., Wang, H., Zhang, S., & Zhang, L. (2020). Multi-scale data-driven engineering for biosynthetic titer improvement. *Current Opinion in Biotechnology*, *65*, 205–212. https://doi.org/10.1016/j.copbio.2020.04.002

Carlson, R., & Wehbring, R. (2020). *Two Worlds, Two Bioeconomies: The Impacts of Decoupling US-China Trade and Technology Transfer*. https://apps.dtic.mil/sti/citations/AD1116887

Carlson-Stevermer, J., Kelso, R., Kadina, A., Joshi, S., Rossi, N., Walker, J., Stoner, R., & Maures, T. (2020). CRISPRoff enables spatio-temporal control of CRISPR editing. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-18853-3

Carr, P. A., Bobrow, J., Comolli, J. C., Guido, N. J., Nargi, F. E., Thorsen, T. A., Walsh, D. I., Walsh, M. E., Wick, S. T., & Cabrera, C. R., (2020). Synthetic Biology. *Lincoln Laboratory Journal*, 24(1), 213-234. https://www.ll.mit.edu/sites/default/files/page/doc/2020-07/11_Synthetic_Bio.pdf

Cashman, M., Catlett, J. L., Cohen, M. B., Buan, N. R., Sakkaff, Z., Pierobon, M., & Kelley, C. A. (2017). BioSIMP: Using Software Testing Techniques for Sampling and Inference in Biological Organisms. *2017 IEEE/ACM 12th International Workshop on Software Engineering for Science (SE4Science)*, 2–8. https://doi.org/10.1109/SE4Science.2017.9

Castro, C. C., Nobre, C., De Weireld, G., & Hantson, A.-L. (2019). Microbial co-culturing strategies for fructo-oligosaccharide production. *New Biotechnology*, *51*, 1–7. https://doi.org/10.1016/j.nbt.2019.01.009

Catlett, J. L., Carr, S., Cashman, M., Smith, M. D., Walter, M., Sakkaff, Z., Kelley, C., Pierobon, M., Cohen, M. B., & Buan, N. R. (2022). Metabolic Synergy between Human Symbionts Bacteroides and Methanobrevibacter. *Microbiology Spectrum*, *10*(3), e01067-22. https://doi.org/10.1128/spectrum.01067-22

Chang, M. M., Gaidukov, L., Jung, G., Tseng, W. A., Scarcelli, J. J., Cornell, R., Marshall, J. K., Lyles, J. L., Sakorafas, P., Chu, A.-H. A., Cote, K., Tzvetkova, B., Dolatshahi, S., Sumit, M., Mulukutla, B. C., Lauffenburger, D. A., Figueroa, B., Summers, N. M., Lu, T. K., & Weiss, R. (2019). Small-molecule control of antibody N-glycosylation in engineered mammalian cells. *Nature Chemical Biology*, *15*(7), Article 7. https://doi.org/10.1038/s41589-019-0288-4

Chao, J. (2020, September 25). *Machine Learning Takes on Synthetic Biology: Algorithms Can Bioengineer Cells for You*. News Center. https://newscenter.lbl.gov/2020/09/25/machine-learning-takes-on-synthetic-biology-algorithms-can-bioengineer-cells-for-you/

Che, S., & Men, Y. (2019). Synthetic microbial consortia for biosynthesis and biodegradation: Promises and challenges. *Journal of Industrial Microbiology and Biotechnology*, *46*(9–10), 1343–1358. https://doi.org/10.1007/s10295-019-02211-4

Chechik, L., Martin, O., & Soutoglou, E. (2020). Genome Editing Fidelity in the Context of DNA Sequence and Chromatin Structure. *Frontiers in Cell and Developmental Biology*, *8*. https://www.frontiersin.org/articles/10.3389/fcell.2020.00319

Chen, C., Zhou, Y., Fu, H., Xiong, X., Fang, S., Jiang, H., Wu, J., Yang, H., Gao, J., & Huang, L. (2021). Expanded catalog of microbial genes and metagenome-assembled genomes from the pig gut microbiome. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21295-0

Chen, Y., Guenther, J. M., Gin, J. W., Chan, L. J. G., Costello, Z., Ogorzalek, T. L., Tran, H. M., Blake-Hedges, J. M., Keasling, J. D., Adams, P. D., García Martín, H., Hillson, N. J., & Petzold, C. J. (2019). Automated "Cells-To-Peptides" Sample Preparation Workflow for High-Throughput, Quantitative Proteomic Assays of

Microbes. *Journal of Proteome Research*, *18*(10), 3752–3761. https://doi.org/10.1021/acs.jproteome.9b00455

Chen, Y., Zhang, S., Young, E. M., Jones, T. S., Densmore, D., & Voigt, C. A. (2020). Genetic circuit design automation for yeast. *Nature Microbiology*, *5*(11), Article 11. https://doi.org/10.1038/s41564-020-0757-2

Chinese Academy of Sciences, Shanghai Bioengineering Society. (2019). (Chinese). Verbatim Chinese Translation of the Roadmap. https://www.ssbt.org.cn/upload/20190919163.

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., Ferrero, E., Agapow, P.-M., Zietz, M., Hoffman, M. M., Xie, W., Rosen, G. L., Lengerich, B. J., Israeli, J., Lanchantin, J., Woloszynek, S., Carpenter, A. E., Shrikumar, A., Xu, J., … Greene, C. S. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface*, *15*(141), 20170387. https://doi.org/10.1098/rsif.2017.0387

Choe, J. H., Watchmaker, P. B., Simic, M. S., Gilbert, R. D., Li, A. W., Krasnow, N. A., Downey, K. M., Yu, W., Carrera, D. A., Celli, A., Cho, J., Briones, J. D., Duecker, J. M., Goretsky, Y. E., Dannenfelser, R., Cardarelli, L., Troyanskaya, O., Sidhu, S. S., Roybal, K. T., … Lim, W. A. (2021). SynNotch-CAR T cells overcome challenges of specificity, heterogeneity, and persistence in treating glioblastoma. *Science Translational Medicine*, *13*(591), eabe7378. https://doi.org/10.1126/scitranslmed.abe7378

Choi, J., Chen, W., Suiter, C. C., Lee, C., Chardon, F. M., Yang, W., Leith, A., Daza, R. M., Martin, B., & Shendure, J. (2021). *Precise genomic deletions using paired prime editing* (p. 2020.12.30.424891). bioRxiv. https://doi.org/10.1101/2020.12.30.424891

Clark, R. L., Connors, B. M., Stevenson, D. M., Hromada, S. E., Hamilton, J. J., Amador-Noguez, D., & Venturelli, O. S. (2021). Design of synthetic human gut microbiome assembly and butyrate production. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-22938-y

Clarke, L., & Kitney, R. (2020). Developing synthetic biology for industrial biotechnology applications. *Biochemical Society Transactions*, *48*(1), 113–122. https://doi.org/10.1042/BST20190349

Cole, S. D., Beabout, K., Turner, K. B., Smith, Z. K., Funk, V. L., Harbaugh, S. V., Liem, A. T., Roth, P. A., Geier, B. A., Emanuel, P. A., Walper, S. A., Chávez, J. L., & Lux, M. W. (2019). Quantification of Interlaboratory Cell-Free Protein Synthesis Variability. *ACS Synthetic Biology*, *8*(9), 2080–2091. https://doi.org/10.1021/acssynbio.9b00178

Commonwealth Scientific and Industrial Research Organisation. (2021). A National Synthetic Biology Roadmap. https://www.csiro.au/-/media/Services/Futures/Synthetic-Biology-Roadmap.pdf.

Congressional Research Service, Gallo, Marcy E. (2021). The Bioeconomy: A Primer. https://crsreports.congress.gov/

Contreras-Llano, L. E., Meyer, C., Liu, Y., Sarker, M., Lim, S., Longo, M. L., & Tan, C. (2020). Holistic engineering of cell-free systems through proteome-reprogramming synthetic circuits. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-16900-7

Costello, A., & Badran, A. H. (2021). Synthetic Biological Circuits within an Orthogonal Central Dogma. *Trends in Biotechnology*, *39*(1), 59–71. https://doi.org/10.1016/j.tibtech.2020.05.013

Cox, J. R., & Blazeck, J. (2022). Protein engineering: A driving force toward synthetic immunology. *Trends in Biotechnology*, *40*(4), 509–521. https://doi.org/10.1016/j.tibtech.2021.09.005

Cravens, A., Jamil, O. K., Kong, D., Sockolosky, J. T., & Smolke, C. D. (2021). Polymerase-guided base editing enables in vivo mutagenesis and rapid protein engineering. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21876-z

Curach, N. (2021). Building capacity for the engineering of biology in Australia. *Engineering Biology*, *5*(2), 43–47. https://doi.org/10.1049/enb2.12009

Currin, A., Swainston, N., Dunstan, M. S., Jervis, A. J., Mulherin, P., Robinson, C. J., Taylor, S., Carbonell, P., Hollywood, K. A., Yan, C., Takano, E., Scrutton, N. S., & Breitling, R. (2019). Highly multiplexed, fast and

accurate nanopore sequencing for verification of synthetic DNA constructs and sequence libraries. *Synthetic Biology*, *4*(1), ysz025. https://doi.org/10.1093/synbio/ysz025

Dale, T., & Guss, A. (2019). Agile BioFoundry – Host Onboarding and Development. 63. http://agilebiofoundry.org/wp-content/uploads/2019/03/Host-Onboarding.pdf

Danko, D., Bezdan, D., Afshin, E. E., Ahsanuddin, S., Bhattacharya, C., Butler, D. J., Chng, K. R., Donnellan, D., Hecht, J., Jackson, K., Kuchin, K., Karasikov, M., Lyons, A., Mak, L., Meleshko, D., Mustafa, H., Mutai, B., Neches, R. Y., Ng, A., … Zubenko, S. (2021). A global metagenomic map of urban microbiomes and antimicrobial resistance. *Cell*, *184*(13), 3376-3393.e17. https://doi.org/10.1016/j.cell.2021.05.002

DeBenedictis, E. A., Chory, E. J., Gretton, D. W., Wang, B., Golas, S., & Esvelt, K. M. (2022). Systematic molecular evolution enables robust biomolecule discovery. *Nature Methods*, *19*(1), Article 1. https://doi.org/10.1038/s41592-021-01348-4

DeBenedictis, E. A., Chory, E. J., Gretton, D., Wang, B., & Esvelt, K. (2020). *A high-throughput platform for feedback-controlled directed evolution* (p. 2020.04.01.021022). bioRxiv. https://doi.org/10.1101/2020.04.01.021022

Decaestecker, W., Buono, R. A., Pfeiffer, M. L., Vangheluwe, N., Jourquin, J., Karimi, M., Van Isterdael, G., Beeckman, T., Nowack, M. K., & Jacobs, T. B. (2019). CRISPR-TSKO: A Technique for Efficient Mutagenesis in Specific Cell Types, Tissues, or Organs in Arabidopsis. *The Plant Cell*, *31*(12), 2868–2887. https://doi.org/10.1105/tpc.19.00454

Delebecque, C. Philip, J. (2015). Training for synthetic biology jobs in the new bioeconomy. https://www.science.org/content/article/training-synthetic-biology-jobs-new-bioeconomy

DeLorenzo, D. M., Diao, J., Carr, R., Hu, Y., & Moon, T. S. (2021). An Improved CRISPR Interference Tool to Engineer Rhodococcus opacus. *ACS Synthetic Biology*, *10*(4), 786–798. https://doi.org/10.1021/acssynbio.0c00591

Demirer, G. S., Zhang, H., Matos, J. L., Goh, N. S., Cunningham, F. J., Sung, Y., Chang, R., Aditham, A. J., Chio, L., Cho, M.-J., Staskawicz, B., & Landry, M. P. (2019). High aspect ratio nanomaterials enable delivery of functional genetic material without DNA integration in mature plants. *Nature Nanotechnology*, *14*(5), Article 5. https://doi.org/10.1038/s41565-019-0382-5

Diener, C., Gibbons, S. M., & Resendis-Antonio, O. (2020). MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. *MSystems*, *5*(1), e00606-19. https://doi.org/10.1128/mSystems.00606-19

Dileo, J., Rambhia, K., Downs, M., Rowell, J., & Kennedy, C. (2022). Maintaining U.S. Leadership in Advanced Biotechnology & Growing the Bioeconomy. https://www.mitre.org/publications/technical-papers/maintaining-us-leadership-in-advanced-biotechnology-growing-the-bioeconomy

Ding, L., Brown, D. M., & Glass, J. I. (2021). Rescue of Infectious Sindbis Virus by Yeast Spheroplast-Mammalian Cell Fusion. *Viruses*, *13*(4), Article 4. https://doi.org/10.3390/v13040603

Dinh, C. V., Chen, X., & Prather, K. L. J. (2020). Development of a Quorum-Sensing Based Circuit for Control of Coculture Population Composition in a Naringenin Production System. *ACS Synthetic Biology*, *9*(3), 590–597. https://doi.org/10.1021/acssynbio.9b00451

Dixon, T. A., & Pretorius, I. S. (2020). Drawing on the Past to Shape the Future of Synthetic Yeast Research. *International Journal of Molecular Sciences*, *21*(19), Article 19. https://doi.org/10.3390/ijms21197156

Dixon, T. A., Williams, T. C., & Pretorius, I. S. (2021). Sensing the future of bio-informational engineering. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-020-20764-2

Dixon, T. A., Williams, T. C., & Pretorius, I. S. (2022). Bioinformational trends in grape and wine biotechnology. *Trends in Biotechnology*, *40*(1), 124–135. https://doi.org/10.1016/j.tibtech.2021.05.001

DNA Script. (2020, February 27). *DNA Script Presents New Data Demonstrating the Feasibility of the Company's Enzymatic Synthesis Technology*. https://www.businesswire.com/news/home/20200227005340/en/DNA-

Script-Presents-New-Data-Demonstrating-the-Feasibility-of-the-Company%E2%80%99s-Enzymatic-Synthesis-Technology

Du, T., Buenbrazo, N., Kell, L., Rahmani, S., Sim, L., Withers, S. G., DeFrees, S., & Wakarchuk, W. (2019). A Bacterial Expression Platform for Production of Therapeutic Proteins Containing Human-like O-Linked Glycans. *Cell Chemical Biology*, *26*(2), 203-212.e5. https://doi.org/10.1016/j.chembiol.2018.10.017

Du, W., Jongbloets, J. A., Guillaume, M., van de Putte, B., Battaglino, B., Hellingwerf, K. J., & Branco dos Santos, F. (2019). Exploiting Day- and Night-Time Metabolism of Synechocystis sp. PCC 6803 for Fitness-Coupled Fumarate Production around the Clock. *ACS Synthetic Biology*, *8*(10), 2263–2269. https://doi.org/10.1021/acssynbio.9b00289

Dudley, Q. M., Cai, Y.-M., Kallam, K., Debreyne, H., Carrasco Lopez, J. A., & Patron, N. J. (2021). Biofoundry-assisted expression and characterization of plant proteins. *Synthetic Biology*, *6*(1), ysab029. https://doi.org/10.1093/synbio/ysab029

Duffy, K., Arangundy-Franklin, S., & Holliger, P. (2020). Modified nucleic acids: Replication, evolution, and next-generation therapeutics. *BMC Biology*, *18*(1), 112. https://doi.org/10.1186/s12915-020-00803-6

Dugé de Bernonville, T., Maury, S., Delaunay, A., Daviaud, C., Chaparro, C., Tost, J., O'Connor, S. E., & Courdavault, V. (2020). Developmental Methylome of the Medicinal Plant Catharanthus roseus Unravels the Tissue-Specific Control of the Monoterpene Indole Alkaloid Pathway by DNA Methylation. *International Journal of Molecular Sciences*, *21*(17), Article 17. https://doi.org/10.3390/ijms21176028

Dukovski, I., Bajić, D., Chacón, J. M., Quintin, M., Vila, J. C. C., Sulheim, S., Pacheco, A. R., Bernstein, D. B., Riehl, W. J., Korolev, K. S., Sanchez, A., Harcombe, W. R., & Segrè, D. (2021). A metabolic modeling platform for the computation of microbial ecosystems in time and space (COMETS). *Nature Protocols*, *16*(11), Article 11. https://doi.org/10.1038/s41596-021-00593-3

Edinburgh-Genome-Foundry/Plateo. (2021). [Python]. Edinburgh Genome Foundry. https://github.com/Edinburgh-Genome-Foundry/Plateo (Original work published 2017)

Eiben, C. B., de Rond, T., Bloszies, C., Gin, J., Chiniquy, J., Baidoo, E. E. K., Petzold, C. J., Hillson, N. J., Fiehn, O., & Keasling, J. D. (2019). Mevalonate Pathway Promiscuity Enables Noncanonical Terpene Production. *ACS Synthetic Biology*, *8*(10), 2238–2247. https://doi.org/10.1021/acssynbio.9b00230

Eisenstein, M. (2020). Enzymatic DNA synthesis enters new phase. *Nature Biotechnology*, *38*(10), Article 10. https://doi.org/10.1038/s41587-020-0695-9

Emery, N. C., Crispo, E., Supp, S. R., Farrell, K. J., Kerkhoff, A. J., Bledsoe, E. K., O'Donnell, K. L., McCall, A. C., & Aiello-Lammens, M. E. (2021). Data Science in Undergraduate Life Science Education: A Need for Instructor Skills Training. *BioScience*, *71*(12), 1274–1287. https://doi.org/10.1093/biosci/biab107

Endoh, T., Kanai, T., Sato, Y. T., Liu, D. V., Yoshikawa, K., Atomi, H., & Imanaka, T. (2006a). Cell-free protein synthesis at high temperatures using the lysate of a hyperthermophile. *Journal of Biotechnology*, *126*(2), 186–195. https://doi.org/10.1016/j.jbiotec.2006.04.010

Engelberts, J. P., Robbins, S. J., de Goeij, J. M., Aranda, M., Bell, S. C., & Webster, N. S. (2020). Characterization of a sponge microbiome using an integrative genome-centric approach. *The ISME Journal*, *14*(5), Article 5. https://doi.org/10.1038/s41396-020-0591-9

Engineering Biology Research Consortium (2019). Engineering Biology: A Research Roadmap for the Next-Generation Bioeconomy. *Engineering Biology Research Consortium.* https://roadmap.ebrc.org/2019-roadmap/ . doi: 10.25498/E4159B.

Engineering Biology Research Consortium (2022). Engineering Biology for Climate & Sustainability: A Research Roadmap for a Cleaner Future. *Engineering Biology Research Consortium.* https://roadmap.ebrc.org/engineering-biology-for-climate-sustainability/ . doi: 10.25498/E4SG64.

Engineering Biology Research Consortium. (2020). Microbiome Engineering: A Research Roadmap for the Next-Generation Bioeconomy. *Engineering Biology Research Consortium.* https://roadmap.ebrc.org/2020-roadmap-microbiomes/

Engineering Biology Research Consortium. (2021). Engineering Biology & Materials Science. *Engineering Biology Research Consortium.* https://roadmap.ebrc.org/2021-roadmap-materials/

Engineering biology-opening a new era of life science development. (2021, August). *The Pandora Report.* https://pandorareport.org/2019/06/21/pandora-report-6-21-2019/

English, J. G., Olsen, R. H. J., Lansu, K., Patel, M., White, K., Cockrell, A. S., Singh, D., Strachan, R. T., Wacker, D., & Roth, B. L. (2019). VEGAS as a Platform for Facile Directed Evolution in Mammalian Cells. *Cell*, *178*(3), 748-761.e17. https://doi.org/10.1016/j.cell.2019.05.051

Espah Borujeni, A., Zhang, J., Doosthosseini, H., Nielsen, A. A. K., & Voigt, C. A. (2020). Genetic circuit characterization by inferring RNA polymerase movement and ribosome usage. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-18630-2

Eto, S., Matsumura, R., Shimane, Y., Fujimi, M., Berhanu, S., Kasama, T., & Kuruma, Y. (2022). Phospholipid synthesis inside phospholipid membrane vesicles. *Communications Biology*, *5*(1), Article 1. https://doi.org/10.1038/s42003-022-03999-1

Farzaneh, T., & Freemont, P. S. (2021). Biofoundries are a nucleating hub for industrial translation. *Synthetic Biology*, *6*(1), ysab013. https://doi.org/10.1093/synbio/ysab013

Feder, A., Jensen, S., Wang, A., Courtney, L., Middleton, L., Van Eck, J., Liu, Y., & Giovannoni, J. J. (2020). Tomato fruit as a model for tissue-specific gene silencing in crop plants. *Horticulture Research*, *7*, 142. https://doi.org/10.1038/s41438-020-00363-4

Fedeson, D. T., Saake, P., Calero, P., Nikel, P. I., & Ducat, D. C. (2020). Biotransformation of 2,4-dinitrotoluene in a phototrophic co-culture of engineered Synechococcus elongatus and Pseudomonas putida. *Microbial Biotechnology*, *13*(4), 997–1011. https://doi.org/10.1111/1751-7915.13544

Fedorec, A.J.H., Karkaria, B.D., Sulu, M., Barnes, C.P. (2021). Single strain control of microbial consortia. *Nature Communications, 12*, 1977. https://www.nature.com/articles/s41467-021-22240-x

Finke, M., Brecht, D., Stifel, J., Gense, K., Gamerdinger, M., & Hartig, J. S. (2021). Efficient splicing-based RNA regulators for tetracycline-inducible gene expression in human cell culture and C. elegans. *Nucleic Acids Research*, *49*(12), e71. https://doi.org/10.1093/nar/gkab233

Fischer, E. C., Hashimoto, K., Zhang, Y., Feldman, A. W., Dien, V. T., Karadeema, R. J., Adhikary, R., Ledbetter, M. P., Krishnamurthy, R., & Romesberg, F. E. (2020). New codons for efficient production of unnatural proteins in a semisynthetic organism. *Nature Chemical Biology*, *16*(5), Article 5. https://doi.org/10.1038/s41589-020-0507-z

Flamme, M., Levi-Acobas, F., Hensel, S., Naskar, S., Röthlisberger, P., Sarac, I., Gasser, G., Müller, J., & Hollenstein, M. (2020). Enzymatic Construction of Artificial Base Pairs: The Effect of Metal Shielding. *ChemBioChem*, *21*(23), 3398–3409. https://doi.org/10.1002/cbic.202000402

Flores, A. D., Ayla, E. Z., Nielsen, D. R., & Wang, X. (2019). Engineering a Synthetic, Catabolically Orthogonal Coculture System for Enhanced Conversion of Lignocellulose-Derived Sugars to Ethanol. *ACS Synthetic Biology*, *8*(5), 1089–1099. https://doi.org/10.1021/acssynbio.9b00007

Fontana, J., Dong, C., Kiattisewee, C., Chavali, V. P., Tickman, B. I., Carothers, J. M., & Zalatan, J. G. (2020). Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-15454-y

Foster, C., Charubin, K., Papoutsakis, E. T., & Maranas, C. D. (2021). Modeling Growth Kinetics, Interspecies Cell Fusion, and Metabolism of a Clostridium acetobutylicum/Clostridium ljungdahlii Syntrophic Coculture. *MSystems*, *6*(1), e01325-20. https://doi.org/10.1128/mSystems.01325-20

Fredens, J., Wang, K., de la Torre, D., Funke, L. F. H., Robertson, W. E., Christova, Y., Chia, T., Schmied, W. H., Dunkelmann, D. L., Beránek, V., Uttamapinant, C., Llamazares, A. G., Elliott, T. S., & Chin, J. W. (2019). Total synthesis of Escherichia coli with a recoded genome. *Nature*, *569*(7757), Article 7757. https://doi.org/10.1038/s41586-019-1192-5

Frisvold, G. B., Moss, S. M., Hodgson, A., & Maxon, M. E. (2021). Understanding the U.S. Bioeconomy: A New Definition and Landscape. *Sustainability*, *13*(4), Article 4. https://doi.org/10.3390/su13041627

Frith, K. (2020). 20-Year Synthetic Biology Research Roadmap: Implications for Vaccine Development and Future Research. www.nursingcenter.com/journalarticle?Article_ID=5583415&Journal_ID=3332683&Issue_ID=5582773

Fu, H., Liang, Y., Zhong, X., Pan, Z., Huang, L., Zhang, H., Xu, Y., Zhou, W., & Liu, Z. (2020). Codon optimization with deep learning to enhance protein expression. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-74091-z

Gallegos, J. E., Rogers, M. F., Cialek, C. A., & Peccoud, J. (2020). Rapid, robust plasmid verification by de novo assembly of short sequencing reads. *Nucleic Acids Research*, *48*(18), e106. https://doi.org/10.1093/nar/gkaa727

Ganar, K. A., Leijten, L., & Deshpande, S. (2022). Actinosomes: Condensate-templated proteinaceous containers for engineering synthetic cells. *Biophysical Journal*, *121*(3), 15a. https://doi.org/10.1016/j.bpj.2021.11.2626

Garay, P. G., Barrera, E. E., & Pantano, S. (2020). Post-Translational Modifications at the Coarse-Grained Level with the SIRAH Force Field. *Journal of Chemical Information and Modeling*, *60*(2), 964–973. https://doi.org/10.1021/acs.jcim.9b00900

Garcia, D. C., Dinglasan, J. L. N., Shrestha, H., Abraham, P. E., Hettich, R. L., & Doktycz, M. J. (2021). A lysate proteome engineering strategy for enhancing cell-free metabolite production. *Metabolic Engineering Communications*, *12*, e00162. https://doi.org/10.1016/j.mec.2021.e00162

García-García, J. D., Van Gelder, K., Joshi, J., Bathe, U., Leong, B. J., Bruner, S. D., Liu, C. C., & Hanson, A. D. (2022). Using continuous directed evolution to improve enzymes for plant applications. *Plant Physiology*, *188*(2), 971–983. https://doi.org/10.1093/plphys/kiab500

Gargantilla Becerra, A., & Lahoz-Beltra, R. (2020). A Microbial Screening in Silico Method for the Fitness Step Evaluation in Evolutionary Algorithms. *Applied Sciences*, *10*(11), Article 11. https://doi.org/10.3390/app10113936

Garenne, D., Libchaber, A., & Noireaux, V. (2020). Membrane molecular crowding enhances MreB polymerization to shape synthetic cells from spheres to rods. *Proceedings of the National Academy of Sciences*, *117*(4), 1902–1909. https://doi.org/10.1073/pnas.1914656117

Ge, Z., Liu, J., Guo, L., Yao, G., Li, Q., Wang, L., Li, J., & Fan, C. (2020). Programming Cell–Cell Communications with Engineered Cell Origami Clusters. *Journal of the American Chemical Society*, *142*(19), 8800–8808. https://doi.org/10.1021/jacs.0c01580

Georgia Tech. (2019). Georgia Tech Scientist Helps Create Roadmap for Engineering Biology. *Georgia Tech College of Sciences.* https://cos.gatech.edu/news/georgia-tech-scientist-helps-create-roadmap-engineering-biology

Giannakoulias, S., Shringari, S. R., Ferrie, J. J., & Petersson, E. J. (2021). Biomolecular simulation based machine learning models accurately predict sites of tolerability to the unnatural amino acid acridonylalanine. *Scientific Reports,* 11(1), Article 1. https://doi.org/10.1038/s41598-021-97965-2

Gilman, J., Singleton, C., Tennant, R. K., James, P., Howard, T. P., Lux, T., Parker, D. A., & Love, J. (2019). Rapid, Heuristic Discovery and Design of Promoter Collections in Non-Model Microbes for Industrial Applications. *ACS Synthetic Biology*, *8*(5), 1175–1186. https://doi.org/10.1021/acssynbio.9b00061

Gilmore, S. P., Lankiewicz, T. S., Wilken, St. E., Brown, J. L., Sexton, J. A., Henske, J. K., Theodorou, M. K., Valentine, D. L., & O'Malley, M. A. (2019). Top-Down Enrichment Guides in Formation of Synthetic Microbial Consortia for Biomass Degradation. *ACS Synthetic Biology*, *8*(9), 2174–2185. https://doi.org/10.1021/acssynbio.9b00271

Giuliano, C. J., Lin, A., Girish, V., & Sheltzer, J. M. (2019). Generating Single Cell–Derived Knockout Clones in Mammalian Cells with CRISPR/Cas9. *Current Protocols in Molecular Biology*, *128*(1), e100. https://doi.org/10.1002/cpmb.100

Glasgow, A. A., Huang, Y.-M., Mandell, D. J., Thompson, M., Ritterson, R., Loshbaugh, A. L., Pellegrino, J., Krivacic, C., Pache, R. A., Barlow, K. A., Ollikainen, N., Jeon, D., Kelly, M. J. S., Fraser, J. S., & Kortemme, T. (2019). Computational design of a modular protein sense-response system. *Science*, *366*(6468), 1024–1028. https://doi.org/10.1126/science.aax8780

Gong, Y., Kang, N. K., Kim, Y. U., Wang, Z., Wei, L., Xin, Y., Shen, C., Wang, Q., You, W., Lim, J.-M., Jeong, S.-W., Park, Y.-I., Oh, H.-M., Pan, K., Poliner, E., Yang, G., Li-Beisson, Y., Li, Y., Hu, Q., … Xu, J. (2020). The NanDeSyn database for Nannochloropsis systems and synthetic biology. *The Plant Journal*, *104*(6), 1736–1745. https://doi.org/10.1111/tpj.15025

Greco, F. V., Pandi, A., Erb, T. J., Grierson, C. S., & Gorochowski, T. E. (2021). Harnessing the central dogma for stringent multi-level control of gene expression. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21995-7

Gregorio, N. E., Kao, W. Y., Williams, L. C., Hight, C. M., Patel, P., Watts, K. R., & Oza, J. P. (2020). Unlocking Applications of Cell-Free Biotechnology through Enhanced Shelf Life and Productivity of E. coli Extracts. *ACS Synthetic Biology*, *9*(4), 766–778. https://doi.org/10.1021/acssynbio.9b00433

Grewal, P. S., Samson, J. A., Baker, J. J., Choi, B., & Dueber, J. E. (2021). Peroxisome compartmentalization of a toxic enzyme improves alkaloid production. *Nature Chemical Biology*, *17*(1), Article 1. https://doi.org/10.1038/s41589-020-00668-4

Grewal. 2020. Pathway and Organelle Engineering for Production of Useful Chemicals in Yeast—ProQuest. (n.d.). https://www.proquest.com/docview/2462981439?pq-origsite=gscholar&fromopenview=true

Gülck, T., Booth, J. K., Carvalho, Â., Khakimov, B., Crocoll, C., Motawia, M. S., Møller, B. L., Bohlmann, J., & Gallage, N. J. (2020). Synthetic Biology of Cannabinoids and Cannabinoid Glucosides in Nicotiana benthamiana and Saccharomyces cerevisiae. *Journal of Natural Products*, *83*(10), 2877–2893. https://doi.org/10.1021/acs.jnatprod.0c00241

Guo, J. L., Kim, Y. S., Xie, V. Y., Smith, B. T., Watson, E., Lam, J., Pearce, H. A., Engel, P. S., & Mikos, A. G. (2019). Modular, tissue-specific, and biodegradable hydrogel cross-linkers for tissue engineering. *Science Advances*, *5*(6), eaaw7396. https://doi.org/10.1126/sciadv.aaw7396

Guo, X., Zhu, Y., Bai, L., & Yang, D. (2020). The Protection Role of Magnesium Ions on Coupled Transcription and Translation in Lyophilized Cell-Free System. *ACS Synthetic Biology*, *9*(4), 856–863. https://doi.org/10.1021/acssynbio.9b00508

Guo, X., Li, Z., Wang, X., Wang, J., Chala, J., Lu, Y., & Zhang, H. (2019). De novo phenol bioproduction from glucose using biosensor-assisted microbial coculture engineering. *Biotechnology and Bioengineering*, *116*(12), 3349–3359. https://doi.org/10.1002/bit.27168

Gurdon, J. B., Javed, K., Vodnala, M., & Garrett, N. (2020). Long-term association of a transcription factor with its chromatin binding site can stabilize gene expression and cell fate commitment. *Proceedings of the National Academy of Sciences*, *117*(26), 15075–15084. https://doi.org/10.1073/pnas.2000467117

Guzman-Chavez, F., Arce, A., Adhikari, A., Vadhin, S., Pedroza-Garcia, J. A., Gandini, C., Ajioka, J. W., Molloy, J., Sanchez-Nieto, S., Varner, J. D., Federici, F., & Haseloff, J. (2022). Constructing Cell-Free Expression Systems for Low-Cost Access. *ACS Synthetic Biology*, *11*(3), 1114–1128. https://doi.org/10.1021/acssynbio.1c00342

Halper, S. M., Hossain, A., & Salis, H. M. (2020). Synthesis Success Calculator: Predicting the Rapid Synthesis of DNA Fragments with Machine Learning. *ACS Synthetic Biology*, *9*(7), 1563–1571. https://doi.org/10.1021/acssynbio.9b00460

Han, Y., Zhou, J., Hu, Y., Lin, Z., Ma, Y., Richardson, J. J., & Caruso, F. (2020). Polyphenol-Based Nanoparticles for Intracellular Protein Delivery via Competing Supramolecular Interactions. *ACS Nano*, *14*(10), 12972–12981. https://doi.org/10.1021/acsnano.0c04197

He, F., Ni, N., Zeng, Z., Wu, D., Feng, Y., Li, A. J., Luu, B., Li, A. F., Qin, K., Wang, E., Wang, X., Wu, X., Luo, H., Zhang, J., Zhang, M., Mao, Y., Pakvasa, M., Wagstaff, W., Zhang, Y., … Fan, J. (2020). FAMSi: A Synthetic Biology Approach to the Fast Assembly of Multiplex siRNAs for Silencing Gene Expression in Mammalian Cells. *Molecular Therapy - Nucleic Acids*, *22*, 885–899. https://doi.org/10.1016/j.omtn.2020.10.007

Henriques, D., Minebois, R., Mendoza, S., Macías, L. G., Pérez-Torrado, R., Barrio, E., Teusink, B., Querol, A., & Balsa-Canto, E. (2021). *A multi-phase multi-objective genome-scale model shows diverse redox balance strategies in yeasts* (p. 2021.02.11.430755). bioRxiv. https://doi.org/10.1101/2021.02.11.430755

Hernandez Bücher, J. E., Staufer, O., Ostertag, L., Mersdorf, U., Platzman, I., & Spatz, J. P. (2022). Bottom-up assembly of target-specific cytotoxic synthetic cells. *Biomaterials*, *285*, 121522. https://doi.org/10.1016/j.biomaterials.2022.121522

Hershewe, J. M., Warfel, K. F., Iyer, S. M., Peruzzi, J. A., Sullivan, C. J., Roth, E. W., DeLisa, M. P., Kamat, N. P., & Jewett, M. C. (2021). Improving cell-free glycoprotein synthesis by characterizing and enriching native membrane vesicles. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-22329-3

Hillson, N., Caddick, M., Cai, Y., Carrasco, J. A., Chang, M. W., Curach, N. C., Bell, D. J., Le Feuvre, R., Friedman, D. C., Fu, X., Gold, N. D., Herrgård, M. J., Holowko, M. B., Johnson, J. R., Johnson, R. A., Keasling, J. D., Kitney, R. I., Kondo, A., Liu, C., … Freemont, P. S. (2019). Building a global alliance of biofoundries. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-10079-2

Hodgson, A., Alper, J., Maxon, M.E. (2022). *The U.S. Bioeconomy: Charting a Course for a Resilient and Competitive Future*. Schmidt Futures. https://doi.org/10.55879/d2hrs7zwc

Holowko, M. B., Frow, E. K., Reid, J. C., Rourke, M., & Vickers, C. E. (2021). Building a biofoundry. *Synthetic Biology*, *6*(1), ysaa026. https://doi.org/10.1093/synbio/ysaa026

Horner, C. B., Maldonado, M., Tai, Y., Rony, R. M. I. K., & Nam, J. (2019). Spatially Regulated Multiphenotypic Differentiation of Stem Cells in 3D via Engineered Mechanical Gradient. *ACS Applied Materials & Interfaces*, *11*(49), 45479–45488. https://doi.org/10.1021/acsami.9b17266

Horvath, N., Vilkhovoy, M., Wayman, J. A., Calhoun, K., Swartz, J., & Varner, J. D. (2020). Toward a genome scale sequence specific dynamic model of cell-free protein synthesis in Escherichia coli. *Metabolic Engineering Communications*, *10*, e00113. https://doi.org/10.1016/j.mec.2019.e00113

Hossain, A., Lopez, E., Halper, S. M., Cetnar, D. P., Reis, A. C., Strickland, D., Klavins, E., & Salis, H. M. (2020). Automated design of thousands of nonrepetitive parts for engineering stable genetic systems. *Nature Biotechnology*, *38*(12), Article 12. https://doi.org/10.1038/s41587-020-0584-2

Hultquist, J. F., Hiatt, J., Schumann, K., McGregor, M. J., Roth, T. L., Haas, P., Doudna, J. A., Marson, A., & Krogan, N. J. (2019). CRISPR–Cas9 genome engineering of primary CD4+ T cells for the interrogation of HIV–host factor interactions. *Nature Protocols*, *14*(1), Article 1. https://doi.org/10.1038/s41596-018-0069-7

Hunt, J. P., Zhao, E. L., Soltani, M., Frei, M., Nelson, J. A. D., & Bundy, B. C. (2019). Streamlining the preparation of “endotoxin-free” ClearColi cell extract with autoinduction media for cell-free protein synthesis of the therapeutic protein crisantaspase. *Synthetic and Systems Biotechnology*, *4*(4), 220–224. https://doi.org/10.1016/j.synbio.2019.11.003

Hurst, T., & Chen, S.-J. (2021). Deciphering nucleotide modification-induced structure and stability changes. *RNA Biology*, *18*(11), 1920–1930. https://doi.org/10.1080/15476286.2021.1882179

Ibrahim, M., Raajaraam, L., & Raman, K. (2021). Modelling microbial communities: Harnessing consortia for biotechnological applications. *Computational and Structural Biotechnology Journal*, *19*, 3892–3907. https://doi.org/10.1016/j.csbj.2021.06.048

Imachi, H., Nobu, M. K., Miyazaki, M., Tasumi, E., Saito, Y., Sakai, S., Ogawara, M., Ohashi, A., & Takai, K. (2022). Cultivation of previously uncultured microorganisms with a continuous-flow down-flow hanging sponge (DHS) bioreactor, using a syntrophic archaeon culture obtained from deep marine sediment as a case study. *Nature Protocols*, *17*(12), Article 12. https://doi.org/10.1038/s41596-022-00735-1

Industrial Biotechnology Innovation and Synthetic Biology Accelerator (IBISBA). (n.d.). *An infrastructure in support of the development of industrial biotechnology as key technology for the industry of the future and for supporting the bioeconomy and the circular economy*. https://hal.inrae.fr/hal-03153762/document\

Israni, D. V., Li, H.-S., Gagnon, K. A., Sander, J. D., Roybal, K. T., Joung, J. K., Wong, W. W., & Khalil, A. S. (2021). *Clinically-driven design of synthetic gene regulatory programs in human cells* (p. 2021.02.22.432371). bioRxiv. https://doi.org/10.1101/2021.02.22.432371

Iwai, K., Ando, D., Kim, P. W., Gach, P. C., Raje, M., Duncomb, T. A., Heinemann, J. V., Northen, T. R., Martin, H. G., Hillson, N. J., Adams, P. D., & Singh, A. K. (2018). Automated flow-based/digital microfluidic platform integrated with onsite electroporation process for multiplex genetic engineering applications. *2018 IEEE Micro Electro Mechanical Systems (MEMS)*, 1229–1232. https://doi.org/10.1109/MEMSYS.2018.8346785

Iwai, K., Sustarich, J., Kim, P. W., Gach, P. C., Raje, M., Heinemann, J. V., Duncombe, T. A., Deng, K., Northen, T. R., Hillson, N. J., Adams', P. D., & Singh, A. K. (n.d.). Hybrid droplet/digital microfluidic platform with integrated optical detection system for high-throughput genetic engineering. 2.

Jewett, M. C., Stark, J. C., Delisa, M. P., & Jaroentomeechai, T. (2021). *Method for Rapid in Vitro Synthesis of Glycoproteins Via Recombinant Production of N-Glycosylated Proteins in Prokaryotic Cell Lysates* (Patent No. 20210139940). https://www.freepatentsonline.com/y2021/0139940.html

Jungbluth, S., Tully, B. J., Arkin, A. P., & Wood-Charlson, E. M. (2020, February 19). *Towards a FAIR-compliant ocean and environmental genome database*. Ocean Sciences Meeting 2020. https://agu.confex.com/agu/osm20/meetingapp.cgi/Paper/657631

Kailash, B. P., Karthik, D., Shinde, M., Damaraju, N., Muthukrishnan, A. B., Prasad, S. B., & Jayaraman, G. (2019). *ChassiDex: A microbial database useful for synthetic biology applications* (p. 703033). bioRxiv. https://doi.org/10.1101/703033

Kappel, K., Zhang, K., Su, Z., Watkins, A. M., Kladwang, W., Li, S., Pintilie, G., Topkar, V. V., Rangan, R., Zheludev, I. N., Yesselman, J. D., Chiu, W., & Das, R. (2020). Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nature Methods*, *17*(7), Article 7. https://doi.org/10.1038/s41592-020-0878-9

Karim, A. S., Liew, F. (Eric), Garg, S., Vögeli, B., Rasor, B. J., Gonnot, A., Pavan, M., Juminaga, A., Simpson, S. D., Köpke, M., & Jewett, M. C. (2020). Modular cell-free expression plasmids to accelerate biological design in cells. *Synthetic Biology*, *5*(1), ysaa019. https://doi.org/10.1093/synbio/ysaa019

Kehe, J., Kulesa, A., Ortiz, A., Ackerman, C. M., Thakku, S. G., Sellers, D., Kuehn, S., Gore, J., Friedman, J., & Blainey, P. C. (2019). Massively parallel screening of synthetic microbial communities. *Proceedings of the National Academy of Sciences*, *116*(26), 12804–12809. https://doi.org/10.1073/pnas.1900102116

Kiattisewee, C., Dong, C., Fontana, J., Sugianto, W., Peralta-Yahya, P., Carothers, J. M., & Zalatan, J. G. (2021). Portable bacterial CRISPR transcriptional activation enables metabolic engineering in Pseudomonas putida. *Metabolic Engineering*, *66*, 283–295. https://doi.org/10.1016/j.ymben.2021.04.002

Kightlinger, W., Duncker, K. E., Ramesh, A., Thames, A. H., Natarajan, A., Stark, J. C., Yang, A., Lin, L., Mrksich, M., DeLisa, M. P., & Jewett, M. C. (2019). A cell-free biosynthesis platform for modular construction of protein glycosylation pathways. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-12024-9

Kim, J. K., Chen, Y., Hirning, A. J., Alnahhas, R. N., Josić, K., & Bennett, M. R. (2019). Long-range temporal coordination of gene expression in synthetic microbial consortia. *Nature Chemical Biology*, *15*(11), Article 11. https://doi.org/10.1038/s41589-019-0372-9

Kim, J., Zhou, Y., Carlson, P., Teichmann, M., Simmel, F. C., Silver, P. A., Collins, J. J., Lucks, J. B., Yin, P., & Green, A. A. (2018). *De-Novo-Designed Translational Repressors for Multi-Input Cellular Logic* (p. 501783). bioRxiv. https://doi.org/10.1101/501783

Kosina, S. M., Rademacher, P., Wetmore, K. M., de Raad, M., Zemla, M., Zane, G. M., Zulovich, J. J., Chakraborty, R., Bowen, B. P., Wall, J. D., Auer, M., Arkin, A. P., Deutschbauer, A. M., & Northen, T. R. (2021). Biofilm Interaction Mapping and Analysis (BIMA) of Interspecific Interactions in Pseudomonas Co-culture Biofilms. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fmicb.2021.757856

Kotopka, B. J., & Smolke, C. D. (2020). Model-driven generation of artificial yeast promoters. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-15977-4

Kriegman, S., Blackiston, D., Levin, M., & Bongard, J. (2020). A scalable pipeline for designing reconfigurable organisms. *Proceedings of the National Academy of Sciences*, *117*(4), 1853–1859. https://doi.org/10.1073/pnas.1910837117

Krüger, A., Mueller, A. P., Rybnicky, G. A., Engle, N. L., Yang, Z. K., Tschaplinski, T. J., Simpson, S. D., Köpke, M., & Jewett, M. C. (2020). Development of a clostridia-based cell-free system for prototyping genetic parts and metabolic pathways. *Metabolic Engineering*, *62*, 95–105. https://doi.org/10.1016/j.ymben.2020.06.004

Lashkevich, K. A., Shlyk, V. I., Kushchenko, A. S., Gladyshev, V. N., Alkalaeva, E. Z., & Dmitriev, S. E. (2020). CTELS: A Cell-Free System for the Analysis of Translation Termination Rate. *Biomolecules*, *10*(6), Article 6. https://doi.org/10.3390/biom10060911

Lavickova, B., Laohakunakorn, N., & Maerkl, S. J. (2020). A partially self-regenerating synthetic cell. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-20180-6

Lee, E. D., Aurand, E. R., & Friedman, D. C. (2020). Engineering Microbiomes—Looking Ahead. *ACS Synthetic Biology*, *9*(12), 3181–3183. https://doi.org/10.1021/acssynbio.0c00558

Lee, J. Y., Nguyen, B., Orosco, C., & Styczynski, M. P. (2021). SCOUR: A stepwise machine learning framework for predicting metabolite-dependent regulatory interactions. *BMC Bioinformatics*, *22*(1), 365. https://doi.org/10.1186/s12859-021-04281-7

Lee, K. J., Kang, D., & Park, and H.-S. (2019). Site-Specific Labeling of Proteins Using Unnatural Amino Acids. *Molecules and Cells*, *42*(5), 386–396. https://doi.org/10.14348/molcells.2019.0078

Lee, M. S., Raig, R. M., Gupta, M. K., & Lux, M. W. (2020). Lyophilized Cell-Free Systems Display Tolerance to Organic Solvent Exposure. *ACS Synthetic Biology*, *9*(8), 1951–1957. https://doi.org/10.1021/acssynbio.0c00267

Lee, Y.-W., Luther, D. C., Goswami, R., Jeon, T., Clark, V., Elia, J., Gopalakrishnan, S., & Rotello, V. M. (2020). Direct Cytosolic Delivery of Proteins through Coengineering of Proteins and Polymeric Delivery Vehicles. *Journal of the American Chemical Society*, *142*(9), 4349–4355. https://doi.org/10.1021/jacs.9b12759

Lewis-Burke Associates LLC. (2021). The Federal Bioeconomy Landscape: Opportunities Related to Biotechnology, Synthetic Biology, and Engineering Biology. https://design.umn.edu/sites/design.umn.edu/files/2021-07/lewis-burke-bioeconomy-landscape-2021.pdf

Li, J., Zhong, H., Ramayo-Caldas, Y., Terrapon, N., Lombard, V., Potocki-Veronese, G., Estellé, J., Popova, M., Yang, Z., Zhang, H., Li, F., Tang, S., Yang, F., Chen, W., Chen, B., Li, J., Guo, J., Martin, C., Maguin, E., … Morgavi, D. P. (2020). A catalog of microbial genes from the bovine rumen unveils a specialized and diverse biomass-degrading environment. *GigaScience*, *9*(6), giaa057. https://doi.org/10.1093/gigascience/giaa057

Li, C., Chu, W., Gill, R. A., Sang, S., Shi, Y., Hu, X., Yang, Y., Zaman, Q. U., & Zhang, B. (2022). Computational tools and resources for CRISPR/Cas genome editing. *Genomics, Proteomics & Bioinformatics*. https://doi.org/10.1016/j.gpb.2022.02.006

Li, Y., Kennedy, N. W., Li, S., Mills, C. E., Tullman-Ercek, D., & Olvera de la Cruz, M. (2021). Computational and Experimental Approaches to Controlling Bacterial Microcompartment Assembly. *ACS Central Science*, *7*(4), 658–670. https://doi.org/10.1021/acscentsci.0c01699

Liao, C., Wang, T., Maslov, S., & Xavier, J. B. (2020). Modeling microbial cross-feeding at intermediate scale portrays community dynamics and species coexistence. *PLOS Computational Biology*, *16*(8), e1008135. https://doi.org/10.1371/journal.pcbi.1008135

Libicher, K., Hornberger, R., Heymann, M., & Mutschler, H. (2020). In vitro self-replication and multicistronic expression of large synthetic genomes. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-14694-2

Liu, C., Wan, T., Wang, H., Zhang, S., Ping, Y., & Cheng, Y. (2019). A boronic acid–rich dendrimer with robust and unprecedented efficiency for cytosolic protein delivery and CRISPR-Cas9 gene editing. *Science Advances*, *5*(6), eaaw8922. https://doi.org/10.1126/sciadv.aaw8922

Liu, H., Robinson, D. S., Wu, Z.-Y., Kuo, R., Yoshikuni, Y., Blaby, I. K., & Cheng, J.-F. (2020). Bacterial genome editing by coupling Cre-lox and CRISPR-Cas9 systems. *PLOS ONE*, *15*(11), e0241867. https://doi.org/10.1371/journal.pone.0241867

Liu, X.-M., Zhou, J., Mao, Y., Ji, Q., & Qian, S.-B. (2019). Programmable RNA N6-methyladenosine editing by CRISPR-Cas9 conjugates. *Nature Chemical Biology*, *15*(9), Article 9. https://doi.org/10.1038/s41589-019-0327-1

Lopez, R., Chen, Y.-J., Dumas Ang, S., Yekhanin, S., Makarychev, K., Racz, M. Z., Seelig, G., Strauss, K., & Ceze, L. (2019). DNA assembly for nanopore data storage readout. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-10978-4

Losoi, P. S., Santala, V. P., & Santala, S. M. (2019). Enhanced Population Control in a Synthetic Bacterial Consortium by Interconnected Carbon Cross-Feeding. *ACS Synthetic Biology*, *8*(12), 2642–2650. https://doi.org/10.1021/acssynbio.9b00316

Lüders, C., Sturm, T., & Radulescu, O. (2022). ODEbase: A repository of ODE systems for systems biology. *Bioinformatics Advances*, *2*(1), vbac027. https://doi.org/10.1093/bioadv/vbac027

Lui, L. M., Majumder, E. L.-W., Smith, H. J., Carlson, H. K., von Netzer, F., Fields, M. W., Stahl, D. A., Zhou, J., Hazen, T. C., Baliga, N. S., Adams, P. D., & Arkin, A. P. (2021). Mechanism Across Scales: A Holistic Modeling Framework Integrating Laboratory and Field Studies for Microbial Ecology. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fmicb.2021.642422

Ma, S., Jiang, F., Huang, Y., Zhang, Y., Wang, S., Fan, H., Liu, B., Li, Q., Yin, L., Wang, H., Liu, H., Ren, Y., Li, S., Cheng, L., Fan, W., & Deng, Y. (2021a). A microbial gene catalog of anaerobic digestion from full-scale biogas plants. *GigaScience*, *10*(1), giaa164. https://doi.org/10.1093/gigascience/giaa164

Ma, S., Jiang, F., Huang, Y., Zhang, Y., Wang, S., Fan, H., Liu, B., Li, Q., Yin, L., Wang, H., Liu, H., Ren, Y., Li, S., Cheng, L., Fan, W., & Deng, Y. (2021b). A microbial gene catalog of anaerobic digestion from full-scale biogas plants. *GigaScience*, *10*(1), giaa164. https://doi.org/10.1093/gigascience/giaa164

Ma, X., Liang, H., Cui, X., Liu, Y., Lu, H., Ning, W., Poon, N. Y., Ho, B., & Zhou, K. (2019). A standard for near-scarless plasmid construction using reusable DNA parts. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-11263-0

Machado, M. R., Barrera, E. E., Klein, F., Sóñora, M., Silva, S., & Pantano, S. (2019). The SIRAH 2.0 Force Field: Altius, Fortius, Citius. *Journal of Chemical Theory and Computation*, *15*(4), 2719–2733. https://doi.org/10.1021/acs.jctc.9b00006

Mackelprang, R., Aurand, E. R., Bovenberg, R. A. L., Brink, K. R., Charo, R. A., Delborne, J. A., Diggans, J., Ellington, A. D., Fortman, J. L. "Clem," Isaacs, F. J., Medford, J. I., Murray, R. M., Noireaux, V., Palmer, M. J., Zoloth, L., & Friedman, D. C. (2021). Guiding Ethical Principles in Engineering Biology Research. *ACS Synthetic Biology*, *10*(5), 907–910. https://doi.org/10.1021/acssynbio.1c00129

Madsen, C., Moreno, A. G., P, U., Palchick, Z., Roehner, N., Atallah, C., Bartley, B., Choi, K., Cox, R. S., Gorochowski, T., Grünberg, R., Macklin, C., McLaughlin, J., Meng, X., Nguyen, T., Pocock, M., Samineni, M., Scott-Brown, J., Tarter, Y., … Wipat, A. (2019). Synthetic Biology Open Language (SBOL) Version 2.3. *Journal of Integrative Bioinformatics*, *16*(2). https://doi.org/10.1515/jib-2019-0025

Malik-Sheriff, R. S., Glont, M., Nguyen, T. V. N., Tiwari, K., Roberts, M. G., Xavier, A., Vu, M. T., Men, J., Maire, M., Kananathan, S., Fairbanks, E. L., Meyer, J. P., Arankalle, C., Varusai, T. M., Knight-Schrijver, V., Li, L., Dueñas-Roca, C., Dass, G., Keating, S. M., … Hermjakob, H. (2020). BioModels—15 years of sharing computational models in life science. *Nucleic Acids Research*, *48*(D1), D407–D415. https://doi.org/10.1093/nar/gkz1055

Mangeot, P. E., Risson, V., Fusil, F., Marnef, A., Laurent, E., Blin, J., Mournetas, V., Massouridès, E., Sohier, T. J. M., Corbin, A., Aubé, F., Teixeira, M., Pinset, C., Schaeffer, L., Legube, G., Cosset, F.-L., Verhoeyen, E., Ohlmann, T., & Ricci, E. P. (2019). Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-018-07845-z

Mante, J., Hao, Y., Jett, J., Joshi, U., Keating, K., Lu, X., Nakum, G., Rodriguez, N. E., Tang, J., Terry, L., Wu, X., Yu, E., Downie, J. S., McInnes, B. T., Nguyen, M. H., Sepulvado, B., Young, E. M., & Myers, C. J. (2021). Synthetic Biology Knowledge System. *ACS Synthetic Biology*, *10*(9), 2276–2285. https://doi.org/10.1021/acssynbio.1c00188

Manzano, I., Taylor, N., Csordas, M., Vezeau, G. E., Salis, H. M., & Zydney, A. L. (2021). Purification of Cas9—RNA complexes by ultrafiltration. *Biotechnology Progress*, *37*(2), e3104. https://doi.org/10.1002/btpr.3104

Mao, N., Aggarwal, N., Poh, C. L., Cho, B. K., Kondo, A., Liu, C., Yew, W. S., & Chang, M. W. (2021). Future trends in synthetic biology in Asia. *Advanced Genetics*, *2*(1), e10038. https://doi.org/10.1002/ggn2.10038

Marshall, R., & Noireaux, V. (2019). Quantitative modeling of transcription and translation of an all-E. coli cell-free system. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-48468-8

Martone, M., Nakamura, R. (2021). Changing the Culture of Data Management and Sharing – A Workshop. https://www.nationalacademies.org/event/04-29-2021/changing-the-culture-of-data-management-and-sharing-a-workshop

Matthews, N. (2021). *The role of sustainability in the UK synthetic biology programme*. SocArXiv. https://doi.org/10.31235/osf.io/vj9nm

Matthews, N., Stamford, L., & Shapira, P. (2021). *The role of business in constructing sustainable technologies: Can the Silicon Valley model be aligned with sustainable development?* SocArXiv. https://doi.org/10.31235/osf.io/sh9an

Mauranyapin, N. P., Terrason, A., & Bowen, W. P. (2021). *Quantum Biotechnology* (arXiv:2111.02021). arXiv. https://doi.org/10.48550/arXiv.2111.02021

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

Meksiriporn, B., Ludwicki, M. B., Stephens, E. A., Jiang, A., Lee, H.-C., Waraho-Zhmayev, D., Kummer, L., Brandl, F., Plückthun, A., & DeLisa, M. P. (2019a). A survival selection strategy for engineering synthetic binding proteins that specifically recognize post-translationally phosphorylated proteins. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-09854-y

Meksiriporn, B., Ludwicki, M. B., Stephens, E. A., Jiang, A., Lee, H.-C., Waraho-Zhmayev, D., Kummer, L., Brandl, F., Plückthun, A., & DeLisa, M. P. (2019b). A survival selection strategy for engineering synthetic binding proteins that specifically recognize post-translationally phosphorylated proteins. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-09854-y

Meng, D., Mukhitov, N., Neitzey, D., Lucht, M., Schaak, D. D., & Voigt, C. A. (2021). Rapid and simultaneous screening of pathway designs and chassis organisms, applied to engineered living materials. *Metabolic Engineering*, *66*, 308–318. https://doi.org/10.1016/j.ymben.2021.01.006

Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J., & Voigt, C. A. (2019). Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors. *Nature Chemical Biology*, *15*(2), Article 2. https://doi.org/10.1038/s41589-018-0168-3

Miano, A., Liao, M. J., & Hasty, J. (2020). Inducible cell-to-cell signaling for tunable dynamics in microbial communities. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-15056-8

Miguez, A. M., McNerney, M. P., & Styczynski, M. P. (2019). Metabolic Profiling of Escherichia coli-Based Cell-Free Expression Systems for Process Optimization. *Industrial & Engineering Chemistry Research*, *58*(50), 22472–22482. https://doi.org/10.1021/acs.iecr.9b03565

Miguez, A. M., Zhang, Y., Piorino, F., & Styczynski, M. P. (2021). Metabolic Dynamics in Escherichia coli-Based Cell-Free Systems. *ACS Synthetic Biology*, *10*(9), 2252–2265. https://doi.org/10.1021/acssynbio.1c00167

Miller, S. M., Wang, T., & Liu, D. R. (2020). Phage-assisted continuous and non-continuous evolution. *Nature Protocols*, *15*(12), Article 12. https://doi.org/10.1038/s41596-020-00410-3

Mirts, E. N., Bhagi-Damodaran, A., & Lu, Y. (2019). Understanding and Modulating Metalloenzymes with Unnatural Amino Acids, Non-Native Metal Ions, and Non-Native Metallocofactors. *Accounts of Chemical Research*, *52*(4), 935–944. https://doi.org/10.1021/acs.accounts.9b00011

Morris, A. (2019). Scientists Chart Course Toward the Next-generation Bioeconomy. *Northwestern Engineering.* https://www.mccormick.northwestern.edu/news/articles/2019/06/scientists-chart-course-toward-the-next-generation-bioeconomy.html

Muldoon, J. J., Kandula, V., Hong, M., Donahue, P. S., Boucher, J. D., Bagheri, N., & Leonard, J. N. (2021). Model-guided design of mammalian genetic programs. *Science Advances*, *7*(8), eabe9375. https://doi.org/10.1126/sciadv.abe9375

Natarajan, A., Jaroentomeechai, T., Cabrera-Sánchez, M., Mohammed, J. C., Cox, E. C., Young, O., Shajahan, A., Vilkhovoy, M., Vadhin, S., Varner, J. D., Azadi, P., & DeLisa, M. P. (2020). Engineering orthogonal human O-linked glycoprotein biosynthesis in bacteria. *Nature Chemical Biology*, *16*(10), Article 10. https://doi.org/10.1038/s41589-020-0595-9

National Academies of Sciences, Engineering, and Medicine. (2015). Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals. *National Academies Press*. https://doi.org/10.17226/19001

National Academies of Sciences, Engineering, and Medicine. (2017). *Preparing for Future Products of Biotechnology.* https://doi.org/10.17226/24605

National Academies of Sciences, Engineering, and Medicine. (2020). Safeguarding the Bioeconomy. Washington, DC: *The National Academies Press*. https://doi.org/10.17226/25525 .

National Research Council. 2014. Safe Science: Promoting a Culture of Safety in Academic Chemical Research. Washington, DC: *The National Academies Press.* https://doi.org/10.17226/18706

Nayfach, S., Roux, S., Seshadri, R., Udwary, D., Varghese, N., Schulz, F., Wu, D., Paez-Espino, D., Chen, I.-M., Huntemann, M., Palaniappan, K., Ladau, J., Mukherjee, S., Reddy, T. B. K., Nielsen, T., Kirton, E., Faria, J. P., Edirisinghe, J. N., Henry, C. S., … Eloe-Fadrosh, E. A. (2021). A genomic catalog of Earth's microbiomes. *Nature Biotechnology*, *39*(4), Article 4. https://doi.org/10.1038/s41587-020-0718-6

Neaman Institute for National Policy Research. (2021). Promoting the leadership of Israeli industry in future technologies. (Hebrew). https://www.neaman.org.il/Files/Promoting%20the%20leadership%20of%20Israeli%20industry%20in%20future%20technologies%20-%20Report_20210503135601.201.pdf

Nuñez, J. K., Chen, J., Pommier, G. C., Cogan, J. Z., Replogle, J. M., Adriaens, C., Ramadoss, G. N., Shi, Q., Hung, K. L., Samelson, A. J., Pogson, A. N., Kim, J. Y. S., Chung, A., Leonetti, M. D., Chang, H. Y., Kampmann, M., Bernstein, B. E., Hovestadt, V., Gilbert, L. A., & Weissman, J. S. (2021). Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell*, *184*(9), 2503-2519.e17. https://doi.org/10.1016/j.cell.2021.03.025

O'Hare, R. (2019). Synthetic biology roadmap could set research agenda for next 10 years. *Imperial News.* https://www.imperial.ac.uk/news/191643/synthetic-biology-roadmap-could-research-agenda/

Oberdick, J. (2019). Penn State professor co-chairs roadmap to guide synthetic biology investments. *Penn State News.* https://www.psu.edu/news/academics/story/penn-state-pHerofessor-co-chairs-roadmap-guide-synthetic-biology-investments/

Oberhardt, M. A., Zarecki, R., Gronow, S., Lang, E., Klenk, H.-P., Gophna, U., & Ruppin, E. (2015). Harnessing the landscape of microbial culture media to predict new organism–media pairings. *Nature Communications*, *6*(1), Article 1. https://doi.org/10.1038/ncomms9493

Oesinghaus, L., & Simmel, F. C. (2019). Switching the activity of Cas12a using guide RNA strand displacement circuits. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-09953-w

Oesinghaus, L., & Simmel, F. C. (2021). Controlling Gene Expression in Mammalian Cells Using Multiplexed Conditional Guide RNAs for Cas12a**. *Angewandte Chemie International Edition*, *60*(44), 23894–23902. https://doi.org/10.1002/anie.202107258

Ohmann, A., Göpfrich, K., Joshi, H., Thompson, R. F., Sobota, D., Ranson, N. A., Aksimentiev, A., & Keyser, U. F. (2019). Controlling aggregation of cholesterol-modified DNA nanostructures. *Nucleic Acids Research*, *47*(21), 11441–11451. https://doi.org/10.1093/nar/gkz914

Pan, C., Wu, X., Markel, K., Malzahn, A. A., Kundagrami, N., Sretenovic, S., Zhang, Y., Cheng, Y., Shih, P. M., & Qi, Y. (2021). CRISPR–Act3.0 for highly efficient multiplexed gene activation in plants. *Nature Plants*, *7*(7), Article 7. https://doi.org/10.1038/s41477-021-00953-7

Park, J., Yim, S. S., & Wang, H. H. (2021). High-Throughput Transcriptional Characterization of Regulatory Sequences from Bacterial Biosynthetic Gene Clusters. *ACS Synthetic Biology*, *10*(8), 1859–1873. https://doi.org/10.1021/acssynbio.0c00639

Perdikari, T. M., Jovic, N., Dignon, G. L., Kim, Y. C., Fawzi, N. L., & Mittal, J. (2021). A predictive coarse-grained model for position-specific effects of post-translational modifications. *Biophysical Journal*, *120*(7), 1187–1197. https://doi.org/10.1016/j.bpj.2021.01.034

Plahar, H. A., Rich, T. N., Lane, S. D., Morrell, W. C., Springthorpe, L., Nnadi, O., Aravina, E., Dai, T., Fero, M. J., Hillson, N. J., & Petzold, C. J. (2021). BioParts—A Biological Parts Search Portal and Updates to the ICE Parts Registry Software Platform. *ACS Synthetic Biology*, *10*(10), 2649–2660. https://doi.org/10.1021/acssynbio.1c00263

Poliner, E., Clark, E., Cummings, C., Benning, C., & Farre, E. M. (2020). A high-capacity gene stacking toolkit for the oleaginous microalga, Nannochloropsis oceanica CCMP1779. *Algal Research*, *45*, 101664. https://doi.org/10.1016/j.algal.2019.101664

Pomraning, K. R., Dai, Z., Munoz, N., Kim, Y.-M., Gao, Y., Deng, S., Kim, J., Hofstad, B. A., Swita, M. S., Lemmon, T., Collett, J. R., Panisko, E. A., Webb-Robertson, B.-J. M., Zucker, J. D., Nicora, C. D., De Paoli, H., Baker, S. E., Burnum-Johnson, K. E., Hillson, N. J., & Magnuson, J. K. (2021). Integration of Proteomics and Metabolomics into the Design, Build, Test, Learn Cycle to Improve 3-Hydroxypropionic Acid Production in Aspergillus pseudoterreus. *Frontiers in Bioengineering and Biotechnology*, *9*. https://www.frontiersin.org/articles/10.3389/fbioe.2021.603832

Popescu, S. (2019). Engineering Biology – A Research Roadmap for the Next-Generation Bioeconomy. *The Pandora Report.* https://pandorareport.org/2019/06/21/pandora-report-6-21-2019/

Price, M. N., Deutschbauer, A. M., & Arkin, A. P. (2020). GapMind: Automated Annotation of Amino Acid Biosynthesis. *MSystems*, *5*(3), e00291-20. https://doi.org/10.1128/mSystems.00291-20

Price, M. N., Deutschbauer, A. M., & Arkin, A. P. (2021). *GapMind for Carbon Sources: Automated annotations of catabolic pathways* (p. 2021.11.02.466981). bioRxiv. https://doi.org/10.1101/2021.11.02.466981

Pryor, J. M., Potapov, V., Kucera, R. B., Bilotti, K., Cantor, E. J., & Lohman, G. J. S. (2020). Enabling one-pot Golden Gate assemblies of unprecedented complexity using data-optimized assembly design. *PLOS ONE*, *15*(9), e0238592. https://doi.org/10.1371/journal.pone.0238592

Qu, L., Yi, Z., Zhu, S., Wang, C., Cao, Z., Zhou, Z., Yuan, P., Yu, Y., Tian, F., Liu, Z., Bao, Y., Zhao, Y., & Wei, W. (2019). Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. *Nature Biotechnology*, *37*(9), Article 9. https://doi.org/10.1038/s41587-019-0178-z

Radivojević, T., Costello, Z., Workman, K., & Garcia Martin, H. (2020). A machine learning Automated Recommendation Tool for synthetic biology. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-18008-4

Rasor, B. J., Chirania, P., Rybnicky, G. A., Giannone, R. J., Engle, N. L., Tschaplinski, T. J., Karim, A. S., Hettich, R. L., & Jewett, M. C. (2023). Mechanistic Insights into Cell-Free Gene Expression through an Integrated -Omics Analysis of Extract Processing Methods. *ACS Synthetic Biology*, *12*(2), 405–418. https://doi.org/10.1021/acssynbio.2c00339

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

Rees-Garbutt, J., Chalkley, O., Landon, S., Purcell, O., Marucci, L., & Grierson, C. (2020). Designing minimal genomes using whole-cell models. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-14545-0

Reis, A. C., Halper, S. M., Vezeau, G. E., Cetnar, D. P., Hossain, A., Clauer, P. R., & Salis, H. M. (2019). Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nature Biotechnology*, *37*(11), Article 11. https://doi.org/10.1038/s41587-019-0286-9

Reis, A. C., & Salis, H. M. (2020). An Automated Model Test System for Systematic Development and Improvement of Gene Expression Models. *ACS Synthetic Biology*, *9*(11), 3145–3156. https://doi.org/10.1021/acssynbio.0c00394

Ren, H., Walker, B. L., Cang, Z., & Nie, Q. (2022). Identifying multicellular spatiotemporal organization of cells with SpaceFlow. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-31739-w

Ricart, E., Leclère, V., Flissi, A., Mueller, M., Pupin, M., & Lisacek, F. (2019). rBAN: Retro-biosynthetic analysis of nonribosomal peptides. *Journal of Cheminformatics*, *11*(1), 13. https://doi.org/10.1186/s13321-019-0335-x

Rieloff, E., & Skepö, M. (2020). Phosphorylation of a Disordered Peptide—Structural Effects and Force Field Inconsistencies. *Journal of Chemical Theory and Computation*, *16*(3), 1924–1935. https://doi.org/10.1021/acs.jctc.9b01190

Riley, L. A., Ji, L., Schmitz, R. J., Westpheling, J., & Guss, A. M. (2019). Rational development of transformation in Clostridium thermocellum ATCC 27405 via complete methylome analysis and evasion of native restriction–modification systems. *Journal of Industrial Microbiology and Biotechnology*, *46*(9–10), 1435–1443. https://doi.org/10.1007/s10295-019-02218-x

Rix, G., Watkins-Dulaney, E. J., Almhjell, P. J., Boville, C. E., Arnold, F. H., & Liu, C. C. (2020). Scalable continuous evolution for the generation of diverse enzyme variants encompassing promiscuous activities. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-19539-6

Robinson, S. A., Hartman, E. C., Ikwuagwu, B. C., Francis, M. B., & Tullman-Ercek, D. (2020). Engineering a Virus-like Particle to Display Peptide Insertions Using an Apparent Fitness Landscape. *Biomacromolecules*, *21*(10), 4194–4204. https://doi.org/10.1021/acs.biomac.0c00987

Rottinghaus, A. G., Ferreiro, A., Fishbein, S. R. S., Dantas, G., & Moon, T. S. (2022). Genetically stable CRISPR-based kill switches for engineered microbes. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-28163-5

Roy, S., Radivojevic, T., Forrer, M., Marti, J. M., Jonnalagadda, V., Backman, T., Morrell, W., Plahar, H., Kim, J., Hillson, N., & Garcia Martin, H. (2021). Multiomics Data Collection, Visualization, and Utilization for Guiding Metabolic Engineering. *Frontiers in Bioengineering and Biotechnology*, *9*. https://www.frontiersin.org/articles/10.3389/fbioe.2021.612893

Rubin, B. E., Diamond, S., Cress, B. F., Crits-Christoph, A., Lou, Y. C., Borges, A. L., Shivram, H., He, C., Xu, M., Zhou, Z., Smith, S. J., Rovinsky, R., Smock, D. C. J., Tang, K., Owens, T. K., Krishnappa, N., Sachdeva, R., Barrangou, R., Deutschbauer, A. M., ... Doudna, J. A. (2022). Species- and site-specific genome editing in complex bacterial communities. *Nature Microbiology*, *7*(1), Article 1. https://doi.org/10.1038/s41564-021-01014-7

Ruggero, D., Creti, R., & Londei, P. (1993). In vitro translation of archaeal natural mRNAs at high temperature. *FEMS Microbiology Letters*, *107*(1), 89–94. https://doi.org/10.1111/j.1574-6968.1993.tb06009.x

Sakamoto, R., Noireaux, V., & Maeda, Y. T. (2018). Anomalous Scaling of Gene Expression in Confined Cell-Free Reactions. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-25532-3

Sanchez-Gorostiaga, A., Bajić, D., Osborne, M. L., Poyatos, J. F., & Sanchez, A. (2019). High-order interactions distort the functional landscape of microbial consortia. *PLOS Biology*, *17*(12), e3000550. https://doi.org/10.1371/journal.pbio.3000550

Sanders, R. (2019, June 19). *Scientists chart course toward a new world of synthetic biology*. Berkeley News. https://news.berkeley.edu/2019/06/19/scientists-chart-course-toward-a-new-world-of-synthetic-biology/

Sansone, S.-A., McQuilton, P., Rocca-Serra, P., Gonzalez-Beltran, A., Izzo, M., Lister, A. L., & Thurston, M. (2019). FAIRsharing as a community approach to standards, repositories and policies. *Nature Biotechnology*, *37*(4), Article 4. https://doi.org/10.1038/s41587-019-0080-8

Sato, K., Akiyama, M., & Sakakibara, Y. (2021). RNA secondary structure prediction using deep learning with thermodynamic integration. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21194-4

Savitskaya, J., Protzko, R. J., Li, F.-Z., Arkin, A. P., & Dueber, J. E. (2019). Iterative screening methodology enables isolation of strains with improved properties for a FACS-based screen and increased L-DOPA production. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-41759-0

Schmidt Futures Task Force on Synthetic Biology and the Bioeconomy. (2021). Public And Private Funding Opportunities To Advance A Circular U.S. Bioeconomy And Maintain U.S. Biotechnology Competitiveness Interim Report. *Schmidt Futures.* https://www.schmidtfutures.com/our-work/task-force-on-synthetic-biology-and-the-bioeconomy/

Seaver, S. M. D., Liu, F., Zhang, Q., Jeffryes, J., Faria, J. P., Edirisinghe, J. N., Mundy, M., Chia, N., Noor, E., Beber, M. E., Best, A. A., DeJongh, M., Kimbrel, J. A., D'haeseleer, P., McCorkle, S. R., Bolton, J. R., Pearson, E., Canon, S., Wood-Charlson, E. M., ... Henry, C. S. (2021). The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes. *Nucleic Acids Research*, *49*(D1), D575–D588. https://doi.org/10.1093/nar/gkaa746

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A. W. R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K., & Hassabis, D. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, *577*(7792), Article 7792. https://doi.org/10.1038/s41586-019-1923-7

Serçin, Ö., Reither, S., Roidos, P., Ballin, N., Palikyras, S., Baginska, A., Rein, K., Llamazares, M., Halavatyi, A., Winter, H., Muley, T., Jurkowska, R. Z., Abdollahi, A., Zenke, F. T., Neumann, B., & Mardin, B. R. (2019). A solid-phase transfection platform for arrayed CRISPR screens. *Molecular Systems Biology*, *15*(12), e8983. https://doi.org/10.15252/msb.20198983

Service, R. (2021). Science's 2021 Breakthrough of the Year: AI brings protein structures to all. *Science*. https://www.science.org/content/article/breakthrough-2021

Shahbazi, R., Sghia-Hughes, G., Reid, J. L., Kubek, S., Haworth, K. G., Humbert, O., Kiem, H.-P., & Adair, J. E. (2019). Targeted homology-directed repair in blood stem and progenitor cells with CRISPR nanoformulations. *Nature Materials*, *18*(10), Article 10. https://doi.org/10.1038/s41563-019-0385-5

Shekaari, A., & Jafari, M. (2019). Statistical mechanical modeling of a DNA nanobiostructure at the base-pair level. *Physica A: Statistical Mechanics and Its Applications*, *518*, 80–88. https://doi.org/10.1016/j.physa.2018.11.054

Sheth, R. U., Li, M., Jiang, W., Sims, P. A., Leong, K. W., & Wang, H. H. (2019). Spatial metagenomic characterization of microbial biogeography in the gut. *Nature Biotechnology*, *37*(8), Article 8. https://doi.org/10.1038/s41587-019-0183-2

Shi, N., Yang, Q., Zhang, H., Lu, J., Lin, H., Yang, X., Abulimiti, A., Cheng, J., Wang, Y., Tong, L., Wang, T., Zhang, X., Chen, H., & Xia, Q. (2022). Restoration of dystrophin expression in mice by suppressing a nonsense mutation through the incorporation of unnatural amino acids. *Nature Biomedical Engineering*, *6*(2), Article 2. https://doi.org/10.1038/s41551-021-00774-1

şimşek, E., Yao, Y., Lee, D., & You, L. (2022). Toward predictive engineering of gene circuits. *Trends in Biotechnology*. https://doi.org/10.1016/j.tibtech.2022.11.001

Shitut, S., Shen, M.-J., Claushuis, B., Derks, R. J. E., Giera, M., Rozen, D., Claessen, D., & Kros, A. (2021). *Generating heterokaryotic cells via bacterial cell-cell fusion* (p. 2021.09.01.458600). bioRxiv. https://doi.org/10.1101/2021.09.01.458600

Shroff, R., Cole, A. W., Diaz, D. J., Morrow, B. R., Donnell, I., Annapareddy, A., Gollihar, J., Ellington, A. D., & Thyer, R. (2020). Discovery of Novel Gain-of-Function Mutations Guided by Structure-Based Deep Learning. *ACS Synthetic Biology*, *9*(11), 2927–2935. https://doi.org/10.1021/acssynbio.0c00345

Silverman, A. D., Kelley-Loughnane, N., Lucks, J. B., & Jewett, M. C. (2019). Deconstructing Cell-Free Extract Preparation for in Vitro Activation of Transcriptional Genetic Circuitry. *ACS Synthetic Biology*, *8*(2), 403–414. https://doi.org/10.1021/acssynbio.8b00430

Singh, J., Hanson, J., Paliwal, K., & Zhou, Y. (2019). RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-13395-9

Smith, E., & Diggans, J. (2020). Next Steps to Grow the Bioeconomy. *Health Security*, *18*(4), 297–302. https://doi.org/10.1089/hs.2020.0012

Stark, J. C., Jaroentomeechai, T., Moeller, T. D., Hershewe, J. M., Warfel, K. F., Moricz, B. S., Martini, A. M., Dubner, R. S., Hsu, K. J., Stevenson, T. C., Jones, B. D., DeLisa, M. P., & Jewett, M. C. (2021). On-demand biomanufacturing of protective conjugate vaccines. *Science Advances*, *7*(6), eabe9444. https://doi.org/10.1126/sciadv.abe9444

Staton, M., Cannon, E., Sanderson, L.-A., Wegrzyn, J., Anderson, T., Buehler, S., Cobo-Simón, I., Faaberg, K., Grau, E., Guignon, V., Gunoskey, J., Inderski, B., Jung, S., Lager, K., Main, D., Poelchau, M., Ramnath, R., Richter, P., West, J., & Ficklin, S. (2021). Tripal, a community update after 10 years of supporting open source, standards-based genetic, genomic and breeding databases. *Briefings in Bioinformatics*, *22*(6), bbab238. https://doi.org/10.1093/bib/bbab238

Staufer, O., Dietrich, F., Rimal, R., Schröter, M., Fabritz, S., Boehm, H., Singh, S., Möller, M., Platzman, I., & Spatz, J. P. (2021). Bottom-up assembly of biomedical relevant fully synthetic extracellular vesicles. *Science Advances*, *7*(36), eabg6666. https://doi.org/10.1126/sciadv.abg6666

Steel, H., Habgood, R., Kelly, C. L., & Papachristodoulou, A. (2020). In situ characterisation and manipulation of biological systems with Chi.Bio. *PLOS Biology*, *18*(7), e3000794. https://doi.org/10.1371/journal.pbio.3000794

Stephens, K., Pozo, M., Tsao, C.-Y., Hauk, P., & Bentley, W. E. (2019). Bacterial co-culture with cell signaling translator and growth controller modules for autonomously regulated culture composition. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-12027-6

Storch, M., Haines, M. C., & Baldwin, G. S. (2020). DNA-BOT: A low-cost, automated DNA assembly platform for synthetic biology. *Synthetic Biology*, *5*(1), ysaa010. https://doi.org/10.1093/synbio/ysaa010

Strobel, E. J., Cheng, L., Berman, K. E., Carlson, P. D., & Lucks, J. B. (2019). A ligand-gated strand displacement mechanism for ZTP riboswitch transcription control. *Nature Chemical Biology*, *15*(11), Article 11. https://doi.org/10.1038/s41589-019-0382-7

Sun, J., Liu, H., Liu, J., Cheng, S., Peng, Y., Zhang, Q., Yan, J., Liu, H.-J., & Chen, L.-L. (2019). CRISPR-Local: A local single-guide RNA (sgRNA) design tool for non-reference plant genomes. *Bioinformatics*, *35*(14), 2501–2503. https://doi.org/10.1093/bioinformatics/bty970

Sun, W., Wang, J., Hu, Q., Zhou, X., Khademhosseini, A., & Gu, Z. (2020). CRISPR-Cas12a delivery by DNA-mediated bioresponsive editing for cholesterol regulation. *Science Advances*, *6*(21), eaba2983. https://doi.org/10.1126/sciadv.aba2983

Swafford, A. J. M., Hussey, S. P., & Fritz-Laylin, L. K. (2020). High-efficiency electroporation of chytrid fungi. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-71618-2

Swarts, D. C., & Jinek, M. (2019). Mechanistic Insights into the cis- and trans-Acting DNase Activities of Cas12a. *Molecular Cell*, *73*(3), 589-600.e4. https://doi.org/10.1016/j.molcel.2018.11.021

Takahashi, K., & Yokobayashi, Y. (2019). Reversible Gene Regulation in Mammalian Cells Using Riboswitch-Engineered Vesicular Stomatitis Virus Vector. *ACS Synthetic Biology*, *8*(9), 1976–1982. https://doi.org/10.1021/acssynbio.9b00177

Taketani, M., Zhang, J., Zhang, S., Triassi, A. J., Huang, Y.-J., Griffith, L. G., & Voigt, C. A. (2020). Genetic circuit design automation for the gut resident species Bacteroides thetaiotaomicron. *Nature Biotechnology*, *38*(8), Article 8. https://doi.org/10.1038/s41587-020-0468-5

The Alliance of Genome Resources Consortium. (2019). The Alliance of Genome Resources: Building a Modern Data Ecosystem for Model Organism Databases. *Genetics*, *213*(4), 1189–1196. https://doi.org/10.1534/genetics.119.302523

Thompson, M. G., Blake-Hedges, J. M., Cruz-Morales, P., Barajas, J. F., Curran, S. C., Eiben, C. B., Harris, N. C., Benites, V. T., Gin, J. W., Sharpless, W. A., Twigg, F. F., Skyrud, W., Krishna, R. N., Pereira, J. H., Baidoo, E. E. K., Petzold, C. J., Adams, P. D., Arkin, A. P., Deutschbauer, A. M., & Keasling, J. D. (2019). Massively Parallel Fitness Profiling Reveals Multiple Novel Enzymes in Pseudomonas putida Lysine Metabolism. *MBio*, *10*(3), e02577-18. https://doi.org/10.1128/mBio.02577-18

Thorgersen, M. P., Xue, J., Majumder, E. L. W., Trotter, V. V., Ge, X., Poole, F. L., Owens, T. K., Lui, L. M., Nielsen, T. N., Arkin, A. P., Deutschbauer, A. M., Siuzdak, G., & Adams, M. W. W. (2021). Deciphering Microbial Metal Toxicity Responses via Random Bar Code Transposon Site Sequencing and Activity-Based Metabolomics. *Applied and Environmental Microbiology*, *87*(21), e01037-21. https://doi.org/10.1128/AEM.01037-21

Tianwei et al. (2021). Accelerate the promotion of green biomanufacturing to help achieve carbon neutrality. *Progress in Chemical Industry*. https://doi.org/10.16085/j.issn.1000-6613.2021-0179

Tickman, B. I., Burbano, D. A., Chavali, V. P., Kiattisewee, C., Fontana, J., Khakimzhan, A., Noireaux, V., Zalatan, J. G., & Carothers, J. M. (2022). Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems. *Cell Systems*, *13*(3), 215-229.e8. https://doi.org/10.1016/j.cels.2021.10.008

Toda, S., McKeithan, W. L., Hakkinen, T. J., Lopez, P., Klein, O. D., & Lim, W. A. (2020). Engineering synthetic morphogen systems that can program multicellular patterning. *Science*, *370*(6514), 327–331. https://doi.org/10.1126/science.abc0033

Torre, D., Krawczuk, P., Jagodnik, K. M., Lachmann, A., Wang, Z., Wang, L., Kuleshov, M. V., & Ma'ayan, A. (2018). Datasets2Tools, repository and search engine for bioinformatics datasets, tools and canned analyses. *Scientific Data*, *5*(1), Article 1. https://doi.org/10.1038/sdata.2018.23

Townshend, R. J. L., Eismann, S., Watkins, A. M., Rangan, R., Karelina, M., Das, R., & Dror, R. O. (2021). Geometric deep learning of RNA structure. *Science*, *373*(6558), 1047–1051. https://doi.org/10.1126/science.abe5650

Twist Bioscience. (n.d.). *Oligo Pools for High Throughput Screens*. https://www.twistbioscience.com/products/oligopools?tab=specs

Tytgat, H. L. P., Lin, C., Levasseur, M. D., Tomek, M. B., Rutschmann, C., Mock, J., Liebscher, N., Terasaka, N., Azuma, Y., Wetter, M., Bachmann, M. F., Hilvert, D., Aebi, M., & Keys, T. G. (2019). Cytoplasmic glycoengineering enables biosynthesis of nanoscale glycoprotein assemblies. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-13283-2

University of California, Irvine. (2019). UCI biomedical engineer co-authors guide for federal investing in synthetic biology research. *University of California, Irvine News.* https://news.uci.edu/2019/06/19/uci-biomedical-engineer-co-authors-guide-for-federal-investing-in-synthetic-biology-research/

U.S. Food & Drug Administration. (Updated 12 December 2022). *Medical Device Shortages During the COVID-19 Public Health Emergency.* https://www.fda.gov/medical-devices/coronavirus-covid-19-and-medical-devices/medical-device-shortages-during-covid-19-public-health-emergency

Uzawa, T., Yamagishi, A., Ueda, T., Chikazumi, N., Watanabe, K., & Oshlma, T. (1993). Effects of Polyamines on a Continuous Cell-Free Protein Synthesis System of an Extreme Thermophile, Thermus thermophilus. *The Journal of Biochemistry*, *114*(5), 732–734. https://doi.org/10.1093/oxfordjournals.jbchem.a124245

Valeri, J. A., Collins, K. M., Ramesh, P., Alcantar, M. A., Lepe, B. A., Lu, T. K., & Camacho, D. M. (2020). Sequence-to-function deep learning frameworks for engineered riboregulators. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-18676-2

Valk, V. D., & Marloes, R. (2020). Metrics of making knowledge in a wilder Anthropocene: The roles and implications of produced knowledge in the governance of synthetic biology for biodiversity conservation. *Master Thesis Series in Environmental Studies and Sustainability Science*. http://lup.lub.lu.se/student-papers/record/9012559

VanArsdale, E., Hörnström, D., Sjöberg, G., Järbur, I., Pitzer, J., Payne, G. F., van Maris, A. J. A., & Bentley, W. E. (2020). A Coculture Based Tyrosine-Tyrosinase Electrochemical Gene Circuit for Connecting Cellular Communication with Electronic Networks. *ACS Synthetic Biology*, *9*(5), 1117–1128. https://doi.org/10.1021/acssynbio.9b00469

Vezeau, G. E., & Salis, H. M. (2021). Tuning Cell-Free Composition Controls the Time Delay, Dynamics, and Productivity of TX-TL Expression. *ACS Synthetic Biology*, *10*(10), 2508–2519. https://doi.org/10.1021/acssynbio.1c00136

Voigt, C. A. (2020). Synthetic biology 2020–2030: Six commercially-available products that are changing our world. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-20122-2

Vojvoda, E. J., Burrington, L. R., & Oza, J. P. (2022). Chapter 19—Development of next-generation diagnostic tools using synthetic biology. In V. Singh (Ed.), *New Frontiers and Applications of Synthetic Biology* (pp. 287–330). Academic Press. https://doi.org/10.1016/B978-0-12-824469-2.00026-9

von Kamp, A., & Klamt, S. (2020). MEMO: A Method for Computing Metabolic Modules for Cell-Free Production Systems. *ACS Synthetic Biology*, *9*(3), 556–566. https://doi.org/10.1021/acssynbio.9b00434

Walter, M. E., Ortiz, A., Sondgeroth, C., Sindt, N. M., Duszenko, N., Catlett, J. L., Zhou, Y., Valloppilly, S., Anderson, C., Fernando, S., & Buan, N. R. (2016). High-throughput mutation, selection, and phenotype screening of mutant methanogenic archaea. *Journal of Microbiological Methods*, *131*, 113–121. https://doi.org/10.1016/j.mimet.2016.10.010

Waltemath, D., Golebiewski, M., Blinov, M. L., Gleeson, P., Hermjakob, H., Hucka, M., Inau, E. T., Keating, S. M., König, M., Krebs, O., Malik-Sheriff, R. S., Nickerson, D., Oberortner, E., Sauro, H. M., Schreiber, F., Smith, L., Stefan, M. I., Wittig, U., & Myers, C. J. (2020). The first 10 years of the international coordination network for standards in systems and synthetic biology (COMBINE). *Journal of Integrative Bioinformatics*, *17*(2–3). https://doi.org/10.1515/jib-2020-0005

Walton, R. T., Christie, K. A., Whittaker, M. N., & Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science*, *368*(6488), 290–296. https://doi.org/10.1126/science.aba8853

Wang, Y., Liu, Y., Zheng, P., Sun, J., & Wang, M. (2021a). Microbial Base Editing: A Powerful Emerging Technology for Microbial Genome Engineering. *Trends in Biotechnology*, *39*(2), 165–180. https://doi.org/10.1016/j.tibtech.2020.06.010

Wang, B., Yang, H., Sun, J., Dou, C., Huang, J., & Guo, F.-B. (2021b). BioMaster: An Integrated Database and Analytic Platform to Provide Comprehensive Information About BioBrick Parts. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fmicb.2021.593979

Wang, G., Zhao, Z., Ke, J., Engel, Y., Shi, Y.-M., Robinson, D., Bingol, K., Zhang, Z., Bowen, B., Louie, K., Wang, B., Evans, R., Miyamoto, Y., Cheng, K., Kosina, S., De Raad, M., Silva, L., Luhrs, A., Lubbe, A., … Yoshikuni, Y. (2019). CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. *Nature Microbiology*, *4*(12), Article 12. https://doi.org/10.1038/s41564-019-0573-8

Wang, X. (2021c). A retrospective on the intellectual adventures of think tanks in biosecurity before and after the COVID-19 pandemic outbreak. *Journal of Biosafety and Biosecurity*, *3*(2), 155–162. https://doi.org/10.1016/j.jobb.2021.10.002

Warfel, K., Williams, A., Wong, D., Sobol, S., Desai, P., Li, J., Chang, Y., DeLisa, M., Karim, A., and Jewett, M. (2023). A Low-Cost, Thermostable, Cell-Free Protein Synthesis Platform for On-Demand Production of Conjugate Vaccines. *ACS Synthetic Biology*. Retrieved February 22, 2023, from https://pubs.acs.org/doi/full/10.1021/acssynbio.2c00392

Watterson, W. J., Tanyeri, M., Watson, A. R., Cham, C. M., Shan, Y., Chang, E. B., Eren, A. M., & Tay, S. (2020). Droplet-based high-throughput cultivation for accurate screening of antibiotic resistant gut microbes. *ELife*, *9*, e56998. https://doi.org/10.7554/eLife.56998

Wellington, S., & Greenberg, E. P. (2019). Quorum Sensing Signal Selectivity and the Potential for Interspecies Cross Talk. *MBio*, *10*(2), e00146-19. https://doi.org/10.1128/mBio.00146-19

Wei, E., & Endy, D. (2021). *Experimental tests of functional molecular regeneration via a standard framework for coordinating synthetic cell building* (p. 2021.03.03.433818). bioRxiv. https://doi.org/10.1101/2021.03.03.433818

Wei, E., Jackson-Smith, A., & Endy, D. (2020). *Minicells as a potential chassis for engineering lineage-agnostic organisms* (p. 2020.07.31.231670). bioRxiv. https://doi.org/10.1101/2020.07.31.231670

Wilding, K. M., Hunt, J. P., Wilkerson, J. W., Funk, P. J., Swensen, R. L., Carver, W. C., Christian, M. L., & Bundy, B. C. (2019). Endotoxin-Free E. coli-Based Cell-Free Protein Synthesis: Pre-Expression Endotoxin Removal Approaches for on-Demand Cancer Therapeutic Production. *Biotechnology Journal*, *14*(3), 1800271. https://doi.org/10.1002/biot.201800271

Wilding, K. M., Zhao, E. L., Earl, C. C., & Bundy, B. C. (2019). Thermostable lyoprotectant-enhanced cell-free protein synthesis for on-demand endotoxin-free therapeutic production. *New Biotechnology*, *53*, 73–80. https://doi.org/10.1016/j.nbt.2019.07.004

Wilken, St. E., Leggieri, P. A., Kerdman-Andrade, C., Reilly, M., Theodorou, M. K., & O'Malley, M. A. (2020). An Arduino based automatic pressure evaluation system to quantify growth of non-model anaerobes in culture. *AIChE Journal*, *66*(12), e16540. https://doi.org/10.1002/aic.16540

Wilken, St. E., Seppälä, S., Lankiewicz, T. S., Saxena, M., Henske, J. K., Salamov, A. A., Grigoriev, I. V., & O'Malley, M. A. (2020). Genomic and proteomic biases inform metabolic engineering strategies for

anaerobic fungi. *Metabolic Engineering Communications*, *10*, e00107. https://doi.org/10.1016/j.mec.2019.e00107

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), Article 1. https://doi.org/10.1038/sdata.2016.18

Wilpiszeski, R. L., Gionfriddo, C. M., Wymore, A. M., Moon, J.-W., Lowe, K. A., Podar, M., Rafie, S., Fields, M. W., Hazen, T. C., Ge, X., Poole, F., Adams, M. W. W., Chakraborty, R., Fan, Y., Nostrand, J. D. V., Zhou, J., Arkin, A. P., & Elias, D. A. (2020). In-field bioreactors demonstrate dynamic shifts in microbial communities in response to geochemical perturbations. *PLOS ONE*, *15*(9), e0232437. https://doi.org/10.1371/journal.pone.0232437

Wong, B. G., Mancuso, C. P., Kiriakov, S., Bashor, C. J., & Khalil, A. S. (2018). Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. *Nature Biotechnology*, *36*(7), Article 7. https://doi.org/10.1038/nbt.4151

Woolston, C. (2021). 'Does anyone have any of these?': Lab-supply shortages strike amid global pandemic. *Nature*. https://doi.org/10.1038/d41586-021-00613-y

Wright, R. C., Moss, B. L., & Nemhauser, J. L. (2022). The Systems and Synthetic Biology of Auxin. *Cold Spring Harbor Perspectives in Biology*, *14*(1), a040071. https://doi.org/10.1101/cshperspect.a040071

Wu, Y., Chen, T., Liu, Y., Tian, R., Lv, X., Li, J., Du, G., Chen, J., Ledesma-Amaro, R., & Liu, L. (2020). Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in Bacillus subtilis. *Nucleic Acids Research*, *48*(2), 996–1009. https://doi.org/10.1093/nar/gkz1123

Wu, Z., Kan, S. B. J., Lewis, R. D., Wittmann, B. J., & Arnold, F. H. (2019). Machine learning-assisted directed protein evolution with combinatorial libraries. *Proceedings of the National Academy of Sciences*, *116*(18), 8852–8858. https://doi.org/10.1073/pnas.1901979116

Xu, X., Jin, L., Xie, L., & Chen, S.-J. (2022). Landscape Zooming toward the Prediction of RNA Cotranscriptional Folding. *Journal of Chemical Theory and Computation*, *18*(3), 2002–2015. https://doi.org/10.1021/acs.jctc.1c01233

Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S., & Baker, D. (2020). Improved protein structure prediction using predicted interresidue orientations. *Proceedings of the National Academy of Sciences*, *117*(3), 1496–1503. https://doi.org/10.1073/pnas.1914677117

Yang, Y., Lee, J. H., Poindexter, M. R., Shao, Y., Liu, W., Lenaghan, S. C., Ahkami, A. H., Blumwald, E., & Stewart Jr, C. N. (2021). Rational design and testing of abiotic stress-inducible synthetic promoters from poplar cis-regulatory elements. *Plant Biotechnology Journal*, *19*(7), 1354–1369. https://doi.org/10.1111/pbi.13550

Yannick Ouedraogo, S., Zhou, X., Chen, H., Chen, F., & Ma, C. (2023). Recent advances in biosensors and sequencing technologies for the detection of mutations. *Microchemical Journal*, *185*, 108306. https://doi.org/10.1016/j.microc.2022.108306

Yeoh, J. W., Swainston, N., Vegh, P., Zulkower, V., Carbonell, P., Holowko, M. B., Peddinti, G., & Poh, C. L. (2021). SynBiopython: An open-source software library for Synthetic Biology. *Synthetic Biology*, *6*(1), ysab001. https://doi.org/10.1093/synbio/ysab001

Yim, S. S., Johns, N. I., Park, J., Gomes, A. L., McBee, R. M., Richardson, M., Ronda, C., Chen, S. P., Garenne, D., Noireaux, V., & Wang, H. H. (2019). Multiplex transcriptional characterizations across diverse bacterial species using cell-free systems. *Molecular Systems Biology*, *15*(8), e8875. https://doi.org/10.15252/msb.20198875

Young, D. D., & Schultz, P. G. (2018). Playing with the molecules of life. *ACS Chemical Biology*, 13(4), 854–870. https://doi.org/10.1021/acschembio.7b00974

Young, R., Haines, M., Storch, M., & Freemont, P. S. (2021). Combinatorial metabolic pathway assembly approaches and toolkits for modular assembly. *Metabolic Engineering*, *63*, 81–101. https://doi.org/10.1016/j.ymben.2020.12.001

Yu, A. M., Gasper, P. M., Cheng, L., Lai, L. B., Kaur, S., Gopalan, V., Chen, A. A., & Lucks, J. B. (2021). Computationally reconstructing cotranscriptional RNA folding from experimental data reveals rearrangement of non-native folding intermediates. *Molecular Cell*, *81*(4), 870-883.e10. https://doi.org/10.1016/j.molcel.2020.12.017

Zelmer, C., Zweifel, L. P., Kapinos, L. E., Craciun, I., Güven, Z. P., Palivan, C. G., & Lim, R. Y. H. (2020). Organelle-specific targeting of polymersomes into the cell nucleus. *Proceedings of the National Academy of Sciences*, *117*(6), 2770–2778. https://doi.org/10.1073/pnas.1916395117

Zengler, K., Hofmockel, K., Baliga, N. S., Behie, S. W., Bernstein, H. C., Brown, J. B., Dinneny, J. R., Floge, S. A., Forry, S. P., Hess, M., Jackson, S. A., Jansson, C., Lindemann, S. R., Pett-Ridge, J., Maranas, C., Venturelli, O. S., Wallenstein, M. D., Shank, E. A., & Northen, T. R. (2019). EcoFABs: Advancing microbiome science through standardized fabricated ecosystems. *Nature Methods*, *16*(7), Article 7. https://doi.org/10.1038/s41592-019-0465-0

Zerfas, B. L., Coleman, R. A., Salazar-Chaparro, A. F., Macatangay, N. J., & Trader, D. J. (2020). Fluorescent Probes with Unnatural Amino Acids to Monitor Proteasome Activity in Real-Time. *ACS Chemical Biology*, *15*(9), 2588–2596. https://doi.org/10.1021/acschembio.0c00634

Zhan, H., Ding, M., Zhou, Q., Li, A., Cai, Z., Huang, W., & Liu, Y. (2020). Improving transgene expression and CRISPR-Cas9 efficiency with molecular engineering-based molecules. *Clinical and Translational Medicine*, *10*(6), e194. https://doi.org/10.1002/ctm2.194

Zhang, H., Zhang, C., Li, Z., Li, C., Wei, X., Zhang, B., & Liu, Y. (2019). A New Method of RNA Secondary Structure Prediction Based on Convolutional Neural Network and Dynamic Programming. *Frontiers in Genetics*, *10*. https://www.frontiersin.org/articles/10.3389/fgene.2019.00467

Zhang, J., Chen, Y., Fu, L., Guo, E., Wang, B., Dai, L., & Si, T. (2021). Accelerating strain engineering in biofuel research via build and test automation of synthetic biology. *Current Opinion in Biotechnology*, *67*, 88–98. https://doi.org/10.1016/j.copbio.2021.01.010

Zhang, Y., Li, H., Min, Y.-L., Sanchez-Ortiz, E., Huang, J., Mireault, A. A., Shelton, J. M., Kim, J., Mammen, P. P. A., Bassel-Duby, R., & Olson, E. N. (2020a). Enhanced CRISPR-Cas9 correction of Duchenne muscular dystrophy in mice by a self-complementary AAV delivery system. *Science Advances*, *6*(8), eaay6812. https://doi.org/10.1126/sciadv.aay6812

Zhang, L., Liu, W.-Q., & Li, J. (2020b). Establishing a Eukaryotic Pichia pastoris Cell-Free Protein Synthesis System. *Frontiers in Bioengineering and Biotechnology*, *8*. https://www.frontiersin.org/articles/10.3389/fbioe.2020.00536

Zhang, Y., Liu, Y., Zhu, J., Xiao, D., Xu, P., Ma, C., Gao, C., & Lü, C. (2021). Coculture of Gluconobacter oxydans and Escherichia coli for 3,4-Dihydroxybutyric Acid Production from Xylose. *ACS Sustainable Chemistry & Engineering*, *9*(32), 10809–10817. https://doi.org/10.1021/acssuschemeng.1c02511

Zhao, D., Li, J., Li, S., Xin, X., Hu, M., Price, M. A., Rosser, S. J., Bi, C., & Zhang, X. (2021). Glycosylase base editors enable C-to-A and C-to-G base changes. *Nature Biotechnology*, *39*(1), Article 1. https://doi.org/10.1038/s41587-020-0592-2

Zhong, Z., Wong, B. G., Ravikumar, A., Arzumanyan, G. A., Khalil, A. S., & Liu, C. C. (2020). Automated Continuous Evolution of Proteins in Vivo. *ACS Synthetic Biology*, *9*(6), 1270–1276. https://doi.org/10.1021/acssynbio.0c00135

Zuo, E., Sun, Y., Yuan, T., He, B., Zhou, C., Ying, W., Liu, J., Wei, W., Zeng, R., Li, Y., & Yang, H. (2020). A rationally engineered cytosine base editor retains high on-target activity while reducing both DNA and RNA off-target effects. *Nature Methods*, *17*(6), Article 6. https://doi.org/10.1038/s41592-020-0832-x

Zhou, Y., Asahara, H., Gaucher, E. A., & Chong, S. (2012). Reconstitution of translation from Thermus thermophilus reveals a minimal set of components sufficient for protein synthesis at high temperatures and functional conservation of modern and ancient translation components. *Nucleic Acids Research*, *40*(16), 7932–7945. https://doi.org/10.1093/nar/gks568

Zhu, Y., Zhu, L., Wang, X., & Jin, H. (2022). RNA-based therapeutics: An overview and prospectus. *Cell Death & Disease*, 13(7), 644. https://doi.org/10.1038/s41419-022-05075-2

Appendices

Appendix I: Process of Creating the Assessment

The Assessment incorporated several methods to measure research progress and identify barriers and other considerations to engineering biology advancement. Contributors to the Assessment were solicited through numerous avenues, including EBRC website announcements, newsletter postings, and presentations at EBRC Roadmapping Working Group and Annual meetings. In total, approximately 75 individuals contributed to the Assessment, including academic researchers, industry professionals, government researchers, and graduate students/postdoctoral fellows (a full list of contributors with affiliations can be found on page 2).

Surveys, created through Qualtrics, queried participants about progress towards milestones, technical and non-technical barriers to research advancement, unanticipated or particularly notable research advancements, and impactful social dimensions. An example survey and the full results of the surveys are available upon request (email roadmapping@ebrc.org). There were 42 unique contributors to the surveys, and their responses were used to identify the foundational publications and topics used for further assessment. The Assessment incorporated a "hackathon" targeted towards graduate students and postdoctoral trainees to collect data on progress and examine the various barriers faced specifically by young researchers. Students and postdocs worked in small groups to complete the survey and then participated in a plenary discussion about major themes and social considerations.

The survey and hackathon results were used to establish a preliminary literature review to identify research articles published since *Engineering Biology*'s release in 2019 that marked progress towards the technical milestones. The Assessment considers publications, products, patents, and other applications of the forecasted research advancements as evidence of technical milestone completion. Altogether, over 300 publications were selected for this report.

EBRC also organized a series of four workshops, each devoted to the individual technical themes. Workshop participants included academic researchers, industry professionals, government researchers, graduate students, and postdoctoral fellows. Like the survey, the workshops were used to solicit qualitative and quantitative information about milestone progress, barriers affecting research progress, and social dimensions. Participants collaborated in groups specific to each technical theme to brainstorm on different elements presented in the roadmap.

Section chairs moderated these workshop discussions along with providing their own insight and deeper review of progress. Once an early draft of each technical theme's progress was complete, the draft was provided to the technical theme chairs for review and feedback. After incorporating this feedback into the draft, a refined version was then provided to all contributors and EBRC members for feedback. The Assessment was published following final incorporation of comments and feedback and copyediting.

Appendix II: Impacts of *Engineering Biology*

Use of *Engineering Biology* in Education, Research and Development, and Policy and Investment

Engineering Biology charted the status and potential of engineering biology and provided researchers and other stakeholders with technical challenges and opportunities in the near and long term. To understand how *Engineering Biology* was able to prove useful as a resource, the Assessment asked how readers used the roadmap and identified how the roadmap was cited across several publications. This knowledge is critical for understanding the roadmap's impact and to strategically develop future roadmaps with different audiences in mind. Contributors noted that the roadmap is a source for unifying the engineering biology field to address everyday needs and goals and remarked on their use of the roadmap in academic training *(Education)*, in government, industry, and nonprofit research strategy development *(Research and Development Strategy)*, and as components of policymaking and investments *(Policy and Investment)* as detailed in the table below.

Education	• When asked how trainees might organize an impactful scientific career in engineering biology, Academic advisors referred graduate students and postdocs to the roadmap. • Trainees cited the roadmap as an excellent onboarding reference when transitioning into the engineering biology field. • Instructors cited the high-level organization of the roadmap as inspiration for their engineering biology syllabi and coursework.
Research and Development Strategy	• Industry representatives reported using the roadmap as an organizational tool when outlining their product goals. • Data platforms, such as KBase, cited the roadmap as instrumental to predicting what forms of datasets, analysis tools, and resources engineering biology practitioners need. • Researchers noted the roadmap framing for breakthrough capabilities and goals is useful when generating project ideas, particularly at the stage of grant writing.
Policy and Investment	• Several international bioeconomy strategic documents cite or reference *Engineering Biology* in plans for policy and/or fundamental research investment. • Researchers noted funding award opportunities, such as the ARPA-E ECOSynBio program, strongly share strategic goals and may be inspired by *Engineering Biology*.

Appendix II Table. Reported Roadmap Use by Stakeholders. The Assessment asked contributors how they personally, or direct anecdotes of how others, have used the roadmap. In academic and educational settings, the roadmap was used to teach and learn about the field, and for professional development. In industry and research planning settings, the roadmap was used to make strategic decisions about projects and products. The roadmap has also influenced funding and policy decisions in the U.S. and globally.

References to *Engineering Biology* in Public Works

The Assessment also examined how the roadmap has been referenced across policy documents, professional and academic publications (including theses, dissertations, and preprints), press releases, and popular media. To identify these citations, the full publication title or DOI of the roadmap (below) was used in publication

search engines such as Google Scholar and Scopus. Given that many government/policy publications do not provide references for their material, the below list likely underscores how referenced the roadmap is. Nonetheless, this information is valuable for understanding how the roadmap is being used. This information is meant to guide future efforts to make roadmaps more impactful.

> ***Engineering Biology* (2019) Citation**
>
> Engineering Biology Research Consortium (2019). Engineering Biology: A Research Roadmap for the Next-Generation Bioeconomy. Retrieved from https://roadmap.ebrc.org. DOI: 10.25498/E4159B.

Research and Investment Strategy

Governments use many different strategies to identify the current landscape of biotechnology efforts or inform prioritized or recommended areas of research. Government agencies often look to the publications and guidance from nonprofit organizations and stakeholder groups to understand important areas for investment or where policy is needed or could be most effective. The Assessment collected the following references to understand how *Engineering Biology* impacts recommendations on U.S. and international policy and funding for biotechnology research.

United States

- National Academies of Sciences, Engineering, and Medicine. (2020). Safeguarding the Bioeconomy. Washington, DC: The National Academies Press. https://doi.org/10.17226/25525.
- Lewis-Burke Associates LLC. (2021). The Federal Bioeconomy Landscape: Opportunities Related to Biotechnology, Synthetic Biology, and Engineering Biology. Retrieved December 21, 2021, from https://design.umn.edu/sites/design.umn.edu/files/2021-07/lewis-burke-bioeconomy-landscape-2021.pdf.
- Congressional Research Service, Gallo, Marcy E. (2021). The Bioeconomy: A Primer. Retrieved from https://crsreports.congress.gov/.
- Schmidt Futures Task Force on Synthetic Biology and the Bioeconomy. (2021). Public And Private Funding Opportunities To Advance A Circular U.S. Bioeconomy And Maintain U.S. Biotechnology Competitiveness Interim Report. Retrieved from https://www.schmidtfutures.com/our-work/task-force-on-synthetic-biology-and-the-bioeconomy/

International

- **Australia.** Commonwealth Scientific and Industrial Research Organisation. (2021). A National Synthetic Biology Roadmap. Retrieved from https://www.csiro.au/-/media/Services/Futures/Synthetic-Biology-Roadmap.pdf.
- **China.** Chinese Academy of Sciences, Shanghai Bioengineering Society. (2019). Verbatim Chinese Translation of the Roadmap. Retrieved from https://www.ssbt.org.cn/upload/20190919163.
- **EU.** Industrial Biotechnology Innovation and Synthetic Biology Accelerator (IBISBA). (n.d.). An infrastructure in support of the development of industrial biotechnology as key technology for the industry of the future and for supporting the bioeconomy and the circular economy. Retrieved from https://hal-amu.archives-ouvertes.fr/GENOMIQUE-METABOLIQUE/hal-03153762v1.
- **Israel.** (Hebrew) Neaman Institute for National Policy Research. (2021). Promoting the leadership of Israeli industry in future technologies. Retrieved from https://www.neaman.org.il/Files/Promoting%20the%20leadership%20of%20Israeli%20industry%20in%20future%20technologies%20-%20Report_20210503135601.201.pdf.

Academic Publications and Perspectives

- Carr, P. A., Bobrow, J., Comolli, J. C., Guido, N. J., Nargi, F. E., Thorsen, T. A., Walsh, D. I., Walsh, M. E., Wick, S. T., & Cabrera, C. R., (2020). Synthetic Biology. Lincoln Laboratory Journal, 24(1), 213-234. https://www.ll.mit.edu/sites/default/files/page/doc/2020-07/11_Synthetic_Bio.pdf
- Cao, Z., Yu, J., Wang, W., Lu, H., Xia, X., Xu, H., Yang, X., Bao, L., Zhang, Q., Wang, H., Zhang, S., & Zhang, L. (2020). Multi-scale data-driven engineering for biosynthetic titer improvement. Current Opinion in Biotechnology, 65, 205–212. https://doi.org/10.1016/j.copbio.2020.04.002
- Clarke, L., & Kitney, R. (2020). Developing synthetic biology for industrial biotechnology applications. Biochemical Society Transactions, 48(1), 113–122. https://doi.org/10.1042/BST20190349
- Curach, N. (2021). Building capacity for the engineering of biology in Australia. Engineering Biology, 5(2), 43–47. https://doi.org/10.1049/enb2.12009
- Dixon, T. A., & Pretorius, I. S. (2020). Drawing on the Past to Shape the Future of Synthetic Yeast Research. International Journal of Molecular Sciences, 21(19), 7156. https://doi.org/10.3390/ijms21197156
- Dixon, T. A., Williams, T. C., & Pretorius, I. S. (2022). Bioinformational trends in grape and wine biotechnology. Trends in Biotechnology, 40(1), 124–135. https://doi.org/10.1016/j.tibtech.2021.05.001
- Dixon, T. A., Williams, T. C., & Pretorius, I. S. (2021). Sensing the future of bio-informational engineering. Nature Communications, 12(1), 388. https://doi.org/10.1038/s41467-020-20764-2
- Frisvold, G. B., Moss, S. M., Hodgson, A., & Maxon, M. E. (2021). Understanding the U.S. Bioeconomy: A New Definition and Landscape. Sustainability, 13(4), 1627. https://doi.org/10.3390/su13041627
- Frith, K. (2020). 20-Year Synthetic Biology Research Roadmap: Implications for Vaccine Development and Future Research. Retrieved December 21, 2021, from www.nursingcenter.com/journalarticle?Article_ID=5583415&Journal_ID=3332683&Issue_ID=5582773
- Grewal. 2020. Pathway and Organelle Engineering for Production of Useful Chemicals in Yeast—ProQuest. (n.d.). Retrieved January 28, 2022, from https://www.proquest.com/docview/2462981439?pq-origsite=gscholar&fromopenview=true
- Holowko, M. B., Frow, E. K., Reid, J. C., Rourke, M., & Vickers, C. E. (2021a). Building a biofoundry. Synthetic Biology, 6(1), ysaa026. https://doi.org/10.1093/synbio/ysaa026
- Lee, E. D., Aurand, E. R., & Friedman, D. C. (2020). Engineering Microbiomes—Looking Ahead. ACS Synthetic Biology, 9(12), 3181–3183. https://doi.org/10.1021/acssynbio.0c00558
- Mackelprang, R., Aurand, E. R., Bovenberg, R. A. L., Brink, K. R., Charo, R. A., Delborne, J. A., Diggans, J., Ellington, A. D., Fortman, J. L. "Clem," Isaacs, F. J., Medford, J. I., Murray, R. M., Noireaux, V., Palmer, M. J., Zoloth, L., & Friedman, D. C. (2021). Guiding Ethical Principles in Engineering Biology Research. ACS Synthetic Biology, 10(5), 907–910. https://doi.org/10.1021/acssynbio.1c00129
- Mao, N., Aggarwal, N., Poh, C. L., Cho, B. K., Kondo, A., Liu, C., Yew, W. S., & Chang, M. W. (2021). Future trends in synthetic biology in Asia. Advanced Genetics, 2(1), e10038. https://doi.org/10.1002/ggn2.10038
- Matthews, N. (2021). The role of sustainability in the UK synthetic biology programme. *SocArXiv.* https://doi.org/10.31235/osf.io/vj9nm
- Matthews, N., Stamford, L., & Shapira, P. (2021). The role of business in constructing sustainable technologies: Can the Silicon Valley model be aligned with sustainable development? *SocArXiv.* https://doi.org/10.31235/osf.io/sh9an

- Mauranyapin, N. P., Terrason, A., & Bowen, W. P. (2021). *Quantum Biotechnology. ArXiv:2111.02021*. http://arxiv.org/abs/2111.02021
- Smith, E., & Diggans, J. (2020). Next Steps to Grow the Bioeconomy. Health Security, 18(4), 297–302. https://doi.org/10.1089/hs.2020.0012
- Tianwei et al. (2021). Accelerate the promotion of green biomanufacturing to help achieve carbon neutrality. Progress in Chemical Industry. https://doi.org/10.16085/j.issn.1000-6613.2021-0179
- Valk, V. D. & Marloes, R. (2020). Metrics of making knowledge in a wilder Anthropocene: The roles and implications of produced knowledge in the governance of synthetic biology for biodiversity conservation. Master Thesis Series in Environmental Studies and Sustainability Science. http://lup.lub.lu.se/student-papers/record/9012559
- Voigt, C. A. (2020). Synthetic biology 2020–2030: Six commercially-available products that are changing our world. Nature Communications, 11(1), 6379. https://doi.org/10.1038/s41467-020-20122-2
- Vojvoda, E. J., Burrington, L. R., & Oza, J. P. (2022). Chapter 19—Development of next-generation diagnostic tools using synthetic biology. In V. Singh (Ed.), New Frontiers and Applications of Synthetic Biology (pp. 287–330). Academic Press. https://doi.org/10.1016/B978-0-12-824469-2.00026-9
- Wang, X. (2021). A retrospective on the intellectual adventures of think tanks in biosecurity before and after the COVID-19 pandemic outbreak. Journal of Biosafety and Biosecurity, 3(2), 155–162. https://doi.org/10.1016/j.jobb.2021.10.002
- Wright, R. C., Moss, B. L., & Nemhauser, J. L. (2022). The Systems and Synthetic Biology of Auxin. Cold Spring Harbor Perspectives in Biology, 14(1), a040071. https://doi.org/10.1101/cshperspect.a040071
- Zhang, J., Chen, Y., Fu, L., Guo, E., Wang, B., Dai, L., & Si, T. (2021). Accelerating strain engineering in biofuel research via build and test automation of synthetic biology. Current Opinion in Biotechnology, 67, 88–98. https://doi.org/10.1016/j.copbio.2021.01.010

Press Releases and Popular Media (by date)

- Sanders. (2019). Scientists chart course toward a new world of synthetic biology. Berkeley News. https://news.berkeley.edu/2019/06/19/scientists-chart-course-toward-a-new-world-of-synthetic-biology/
- Morris. (2019). Scientists Chart Course Toward the Next-generation Bioeconomy. Retrieved from https://www.mccormick.northwestern.edu/news/articles/2019/06/scientists-chart-course-toward-the-next-generation-bioeconomy.html.
- University of California, Irvine. (2019). UCI biomedical engineer co-authors guide for federal investing in synthetic biology research. Retrieved from https://news.uci.edu/2019/06/19/uci-biomedical-engineer-co-authors-guide-for-federal-investing-in-synthetic-biology-research/
- Georgia Tech. (2019). Georgia Tech Scientist Helps Create Roadmap for Engineering Biology. Retrieved from https://cos.gatech.edu/news/georgia-tech-scientist-helps-create-roadmap-engineering-biology
- Oberdick, J. (2019). Penn State professor co-chairs roadmap to guide synthetic biology investments. Retrieved from https://www.psu.edu/news/academics/story/penn-state-professor-co-chairs-roadmap-guide-synthetic-biology-investments/
- O'Hare, R. (2019). Synthetic biology roadmap could set research agenda for next 10 years. Imperial News. Retrieved December 21, 2021, from https://www.imperial.ac.uk/news/191643/synthetic-biology-roadmap-could-research-agenda/
- Beale, S. (2019). EBRC Report Offers a 20-Year Synthetic Biology Roadmap That Could Lead to New Diagnostic Technologies for Clinical Laboratories, Pathologists. Dark Daily. https://www.darkdaily.com/2019/10/21/ebrc-report-offers-a-20-year-synthetic-biology-roadmap-that-could-lead-to-new-diagnostic-technologies-for-clinical-laboratories-pathologists/

- Popescu, S. (2019). Engineering Biology – A Research Roadmap for the Next-Generation Bioeconomy. The Pandora Report. Retrieved from https://pandorareport.org/2019/06/21/pandora-report-6-21-2019/
- Engineering biology-opening a new era of life science development. (2021, August). Retrieved from https://pandorareport.org/2019/06/21/pandora-report-6-21-2019/

Appendix III: Anticipating Progress in Engineering Biology towards the Breakthrough Capabilities

In the Assessment of *Engineering Biology*, we focused our analysis on achievement or progress towards the 2-year milestones that were anticipated to be reached in 2021. Along with this analysis, we considered progress – both published works and through the impressions and response from the research community – towards later, 5-year milestones and their achievement in 2024. This information may provide a higher-level insight for stakeholders that are looking at engineering biology progress long term or more holistically. The tables below summarize a qualitative assessment of where the research stands for achieving the Breakthrough Capabilities, which represents a capacity collective of all the milestones (2-, 5-, 10-, and 20-years). Progress towards the Breakthrough Capabilities was noted to be:

- **On or ahead of schedule** (green +), meaning that most or all of the (2-year) 2021 milestones have been achieved and/or significant progress has been made toward the (5-year) 2024 milestones, such that they might be achieved prior to 2024;
- **Consistent progress** (gray ✓), indicating that some or many 2021 milestones have been achieved and that there may be some progress towards the 2024 milestones, which are anticipated to be achieved in or around 2024; or
- **Inconsistent progress** (red –), indicating that some or most of the 2021 milestones have not yet been achieved and that there has been little progress towards the 2024 milestones.

Progress in Gene Editing, Synthesis, and Assembly	
Goal: Manufacture thousands of very long oligonucleotides with high fidelity.	
Breakthrough Capability: Highly efficient oligonucleotide synthesis to increase the number, length, and fidelity of oligonucleotides.	✓
Goal: Many-fragment DNA assembly with simultaneous, high-fidelity sequence validation.	
Breakthrough Capability: Predictive design of DNA sequences for improved assembly of longer, more information-rich DNA fragments.	+
Breakthrough Capability: Methods for one-step, simultaneous assembly and sequence-verification of long DNA fragments.	+
Breakthrough Capability: Pipelined synthesis, assembly, and functional testing of engineered genetic systems.	✓
Goal: Precision genome editing at multiple sites simultaneously with no off-target effects.	
Breakthrough Capability: Ability to reliably create any precise, defined edit or edits (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.	+
Breakthrough Capability: Precise, predictable, and tunable control of gene expression for many genes inside diverse cells and organisms across different timescales.	✓
Breakthrough Capability: 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.	✓

Appendix III Table 1. Assessment of progress towards Engineering DNA Breakthrough Capabilities. **Green (+)** indicates on or ahead of schedule (green +), meaning that most or all of the (2-year) 2021 milestones have been achieved and/or significant

progress has been made toward the (5-year) 2024 milestones, such that they might be achieved prior to 2024. **Gray (✓)** indicates that some or many 2021 milestones have been achieved and that there may be some progress towards the 2024 milestones, which are anticipated to be achieved in or around 2024. **Red (-)** indicates that some or most of the 2021 milestones have not yet been achieved and that there has been little progress towards the 2024 milestones.

Progress in Biomolecule, Pathway, and Circuit Engineering	
Goal: On-demand design, generation, and evolution of macromolecules for desired functions.	
Breakthrough Capability: De novo prediction of RNA structure, protein structure, and complexes of DNAs/RNAs and proteins from primary sequence and the ability to make accurate predictions of mutability and effect of mutations from structure.	✓
Breakthrough Capability: De novo design and/or prediction of macromolecular dynamics and dynamic macromolecular structures.	✓
Breakthrough Capability: High-throughput integrated computational, experimental, and evolutionary schemes for refinement of desired biomolecule functions including enzymatic activity and binding.	✓
Goal: Special considerations for on-demand design, generation, and evolution of macromolecules that rely on non-canonical/unnatural building blocks.	
Breakthrough Capability: PCR, reverse transcription, cellular replication, and transcription of fully unnatural nucleotide-containing genes of up to 400 base pairs. *Milestones behind schedule:* • **2021:** Identification of "missing" functionality or functionalities in A-T-G-C base pairs.	-
Breakthrough Capability: Expanded genetic code systems for translation of >100-amino acid proteins containing fully-unnatural amino acids, and proteins with at least four, distinct unnatural amino acid building blocks.	✓
Goal: Holistic, integrated design of multi-part genetic systems (i.e., circuits and pathways).	
Breakthrough Capability: Design of highly-stable, large genetic systems (genomes) with targeted expression levels in a host organism or cell type, incorporating system-wide effects.	✓
Breakthrough Capability: Ability to rationally engineer sensor suites, genetic circuits, metabolic pathways, signaling cascades, and cell differentiation pathways.	✓
Goal: Integrated design of RNA-based regulatory systems for cellular control and information processing.	
Breakthrough Capability: Porting nucleic acid strand displacement technology into cellular systems with RNA instantiations.	✓

(Table continues)

Goal: Integrated design of RNA-based regulatory systems for cellular control and information processing. *(Continued)*	
Breakthrough Capability: Porting successes in computationally designed bacterial RNA-based genetic regulators into eukaryotic and mammalian systems. *Milestones behind schedule:* • **2021:** First generation eukaryotic RNA-based gene regulators that utilize RNA:RNA interactions and/or strand-displacement and achieve 10-fold change in gene expression. • **2024:** Second generation eukaryotic RNA-based gene regulators that are suitable for computational design to create libraries that are highly-orthogonal and high-performing, achieving 100's-fold change in gene expression.	–

Appendix III Table 2. Assessment of progress towards Biomolecular Engineering Breakthrough Capabilities. **Green (+)** indicates on or ahead of schedule (green +), meaning that most or all of the (2-year) 2021 milestones have been achieved and/or significant progress has been made toward the (5-year) 2024 milestones, such that they might be achieved prior to 2024. **Gray (✓)** indicates that some or many 2021 milestones have been achieved and that there may be some progress towards the 2024 milestones, which are anticipated to be achieved in or around 2024. **Red (-)** indicates that some or most of the 2021 milestones have not yet been achieved and that there has been little progress towards the 2024 milestones.

Progress in Host and Consortia Engineering	
Goal: Cell-free systems capable of natural and/or non-natural reactions.	
Breakthrough Capability: Ability to build reproducible and comparable cell-free systems for practical applications in bioengineering and biomanufacturing from multiple organisms, including non-model hosts.	+
Breakthrough Capability: Ability to build a cell, including the molecular subsystems that enable the processes of DNA replication, transcription, translation, energy regeneration, and membrane construction.	✓
Breakthrough Capability: Long-lasting, robust, and low-cost cell-free system for protein synthesis and biomanufacturing.	+
Breakthrough Capability: Ability to use cell-free systems to inform cellular design of genetic parts and circuits.	+
Breakthrough Capability: Decentralized, portable, on-demand sensing and manufacturing using cell-free systems.	+
Breakthrough Capability: Ability to manufacture any targeted glycosylated protein or metabolite using cell-free biosynthesis.	+
Goal: On-demand production of single-cell hosts capable of natural and non-natural biochemistry.	
Breakthrough Capability: Ability to grow any host, anytime, in a controlled and regulated setting.	✓
Breakthrough Capability: Routine domestication of non-model organisms through DNA delivery and genetic modification.	+

(Table continues)

Goal: On-demand production of single-cell hosts capable of natural and non-natural biochemistry. *(Cont.)*	
Breakthrough Capability: Ability to build and control small molecule biosynthesis inside cells by design or through evolution.	+
Breakthrough Capability: Spatial control over, or organization of, metabolic pathways in cells and construction of unnatural organelles.	✓
Breakthrough Capability: Production and secretion of any protein with the desired glycosylation or other post-translational modifications.	+
Goal: On-demand fabrication and modification of multicellular organisms.	
Breakthrough Capability: Ability to control differentiation and de-differentiation of cells within a population.	✓
Breakthrough Capability: Ability to characterize and control the three-dimensional (3D) architecture of multicellular systems.	+
Breakthrough Capability: Ability to achieve stable non-heritable changes in somatic cells.	+
Breakthrough Capability: Ability to make predictable and precise, targeted, heritable changes through germline editing.	+
Goal: Generation of biomes and consortia with desired functions and ecologies.	
Breakthrough Capability: Ability to control cell-to-cell communication between different species.	✓
Breakthrough Capability: 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).	✓
Breakthrough Capability: Ability to control and/or define the function of an engineered microbial community/biome.	✓
Breakthrough Capability: 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.	+

Appendix III Table 3. Assessment of progress towards Host Engineering Breakthrough Capabilities. **Green (+)** indicates on or ahead of schedule (green +), meaning that most or all of the (2-year) 2021 milestones have been achieved and/or significant progress has been made toward the (5-year) 2024 milestones, such that they might be achieved prior to 2024. **Gray (✓)** indicates that some or many 2021 milestones have been achieved and that there may be some progress towards the 2024 milestones, which are anticipated to be achieved in or around 2024. **Red (-)** indicates that some or most of the 2021 milestones have not yet been achieved and that there has been little progress towards the 2024 milestones.

Progress in Data Integration, Modeling, and Automation	
Goal: Establish a computational infrastructure where easy access to data supports the DBTL process for biology.	
Breakthrough Capability: Established standard and accessible repositories for biomanufacturing data and analysis methods. *Milestones behind schedule:* • **2021:** Have developed a system of robust communication between academia and industry surrounding engineering biology data access and needs. • **2024:** Biomanufacturing-specific data standards and repositories.	–
Breakthrough Capability: Common computational infrastructure for finding biological data and common APIs for search and analysis.	✓
Breakthrough Capability: End-to-end, industry-normed design software platforms for engineered biological systems. *Milestones behind schedule:* • **2024:** Develop industry-accepted, sharable assessments of current data tools and uses in reducing cost and increasing reliability of executing the DBTL cycle.	–
Goal: Establish functional prediction through biological engineering design at the biomolecular, cellular, and consortium scale.	
Breakthrough Capability: Fully-automated molecular design from integrated, large-scale design data frameworks. *Milestones behind schedule:* • **2024:** Automated designs for integrated manufacturing to enable more successful, iterated workflows. • **2024:** Large-scale design data generation to inform next-generation algorithms for molecular design.	–
Breakthrough Capability: Use of enzyme promiscuity prediction algorithms to design biosynthetic pathways for any molecule (natural or non-natural). *Milestones behind schedule:* • **2021:** Retro-biosynthesis software that can identify any biological or biochemical route to any organic molecule. • **2024:** Data integration for certain classes of enzymes and pathways and predictable host-specific expression in model organisms.	–

(Table continues)

Goal: Establish functional prediction through biological engineering design at the biomolecular, cellular, and consortium scale. *(Continued)*	
Breakthrough Capability: Scalable, data-driven host design for complex environments that enable high-level production of natural biomolecules. *Milestones behind schedule:* • **2021:** Better data on physiology and fitness in deployment environments suitable for informing designs in validated lab-scale simulations that meet activity, persistence, and ecological impact goals. • **2024:** Thematic design rules for host system engineering inferred from data. • **2024:** Tools to acquire and transfer data to a novel host to inform both genetic-domestication and prediction and determination of function. • **2024:** Novel design tools to support host design for more complex, natural (non-laboratory) environments.	–
Breakthrough Capability: Enabled design of functional, self-supporting ecosystems. *Milestones behind schedule:* • **2021:** Data-driven tools for selecting organisms for synthetic assemblies to achieve resistant, resilient activity. • **2021:** Direct data collection for the most important communities in human, agriculture, and complex bioreactor work sufficient for informing design. • **2021:** Modeling tools to identify cross-organismal networks and ecological interactions.	–
Goal: Establish optimal manufacturing processes from the unit-operation to the integrated-screening scale.	
Breakthrough Capability: Standardized informatics tools, data, and automation platforms for efficient and collaborative use and integration of data in order to develop novel products more quickly.	+

Appendix III Table 4. Assessment of progress towards Data Science Breakthrough Capabilities. **Green (+)** indicates on or ahead of schedule (green +), meaning that most or all of the (2-year) 2021 milestones have been achieved and/or significant progress has been made toward the (5-year) 2024 milestones, such that they might be achieved prior to 2024. **Gray (✓)** indicates that some or many 2021 milestones have been achieved and that there may be some progress towards the 2024 milestones, which are anticipated to be achieved in or around 2024. **Red (-)** indicates that some or most of the 2021 milestones have not yet been achieved and that there has been little progress towards the 2024 milestones.