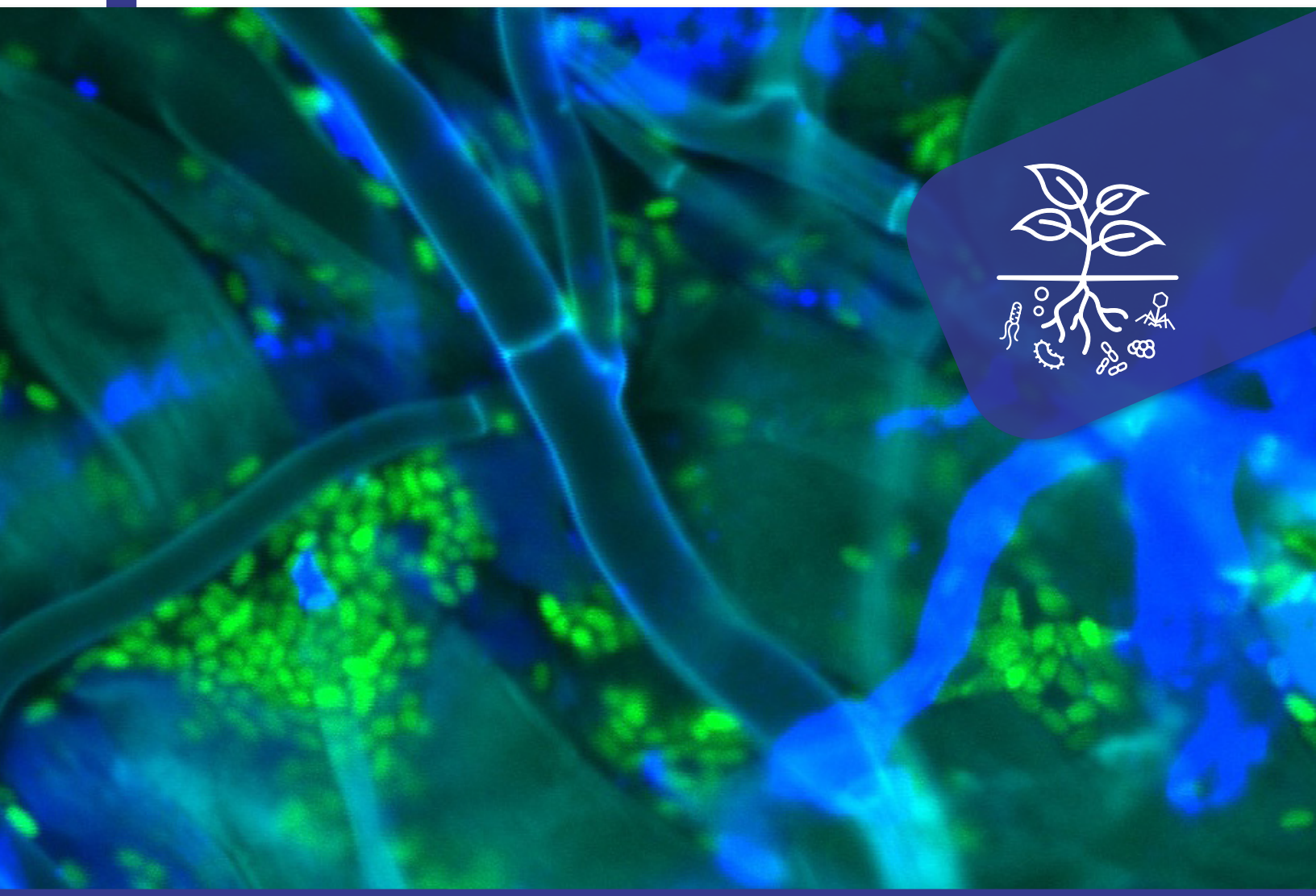
DOE/SC-2025

Engineering Microbial Communities

Frontier Science for the Bioeconomy Workshop Series



U.S. DEPARTMENT of ENERGY | Office of Science
Biological and Environmental Research Program

Engineering Microbial Communities Workshop

December 16–18, 2024

Convened by
U.S. Department of Energy Office of Science
Biological and Environmental Research Program

Workshop Co-Chairs

Melissa Cregger
Oak Ridge National Laboratory

Jonathan (Joff) Silberg
Rice University

Organizer

Boris Wawrik
U.S. Department of Energy

Writing Team

Melissa Cregger (co-chair), Oak Ridge National Laboratory; Jonathan (Joff) Silberg (co-chair), Rice University; Greg Bonito, Michigan State University; Chris Henry, Argonne National Laboratory; Kirsten Hofmockel, Pacific Northwest National Laboratory; Michelle O'Malley, University of California–Santa Barbara; Lingchong You, Duke University

About BER

The Biological and Environmental Research (BER) program's mission is to support transformative science and scientific user facilities to achieve a predictive understanding of complex biological, Earth, and environmental systems in support of DOE's vision to advance innovative solutions for U.S. energy expansion and national security challenges. BER's fundamental research, conducted at DOE national laboratories and other research institutions, plays a unique role in ensuring national leadership in biotechnology and the ability to understand and predict the interdependencies involving energy and the environment over a wide range of conditions.

Front cover image: *Rhizopus oryzae* MC20 and *Enterobacter ludwigii* FCP2-01 expressing green fluorescent protein (green) on the surface of a *Brachypodium distachyon* root, grown in plate culture. Fungal cell walls are stained with wheat germ agglutinin (blue), while both fungal and plant cell walls are stained with Calcofluor White Stain M2R (teal). [Courtesy Michigan State University]

Suggested citation: U.S. DOE. 2026. *Engineering Microbial Communities: Frontier Science for the Bioeconomy Workshop Series,* DOE/SC-2025. U.S. Department of Energy Office of Science. https://doi.org/10.2172/2589381.

DOE/SC-2025

Engineering Microbial Communities

Workshop Report

Frontier Science for the Bioeconomy Workshop Series

April 2026

U.S. DEPARTMENT of ENERGY | Office of Science
Biological and Environmental Research Program

Frontier Science for the Bioeconomy

The *Engineering Microbial Communities* workshop report is one of a four-part Frontier Science for the Bioeconomy series hosted by the DOE Biological and Environmental Research program's Biological Systems Science Division. The series defines frontiers of plant science, agriculture, synthetic biology, biobased materials, and environmental microbiome science that will unlock the promise of an innovative, resilient U.S. bioeconomy.

Engineering Microbial Communities

Vision: Understand the assembly, function, and behavior of microbiomes and how to manipulate them to facilitate microbial energy solutions and advance their utility across the bioeconomy.

Microbial Design for a Developing Bioeconomy

Vision: Harness and leverage the diverse genetic and metabolic potential of microbes as platforms to efficiently produce biofuels, bioproducts, and biomaterials for a thriving bioeconomy.

Plant Design for a Developing Bioeconomy

Vision: Understand and manipulate potential bioenergy crops to efficiently generate biofuels, bioproducts, and biomaterials under variable abiotic stresses.

Resilient Bioenergy Crop Production

Vision: Determine, predict, and improve the functioning of plants and their microbiomes in the environment to optimize biomass feedstock production.

Contents

Executive Summary

In nature, biological systems are shaped by complex interactions of diverse microorganisms such as bacteria, archaea, fungi, and viruses living within communities called microbiomes (Berg et al. 2020; Prescott 2017). These collective interactions result in emergent community properties that can be leveraged for beneficial purposes such as bioenergy and biomolecule production. Given this potential and the immensity of microbial genomic diversity, the U.S. Department of Energy's (DOE) Biological and Environmental Research (BER) program has long invested in research to better understand the biology of environmental microbes and microbiomes.

The strategic priorities highlighted in this workshop report are expected to enable transformative research that realizes the full potential of microbiome engineering across complex, heterogeneous environments.

BER BSSD Microbiome Research Vision and Goals

Within BER, the Biological Systems Science Division (BSSD) seeks to achieve fundamental scientific breakthroughs needed to predict, manipulate, and design biological systems that underpin energy innovations and enhance understanding of natural processes. BSSD's vision includes developing the knowledge necessary to engineer microbiomes, thereby facilitating microbial solutions for challenging environmental problems and advancing the utility of microbiomes across the broader U.S. bioeconomy. In this context, BSSD's fundamental basic science goals include:

- Understanding the principles underlying the assembly, structure, stability, and collective function of microbial populations and communities and how these principles are derived from genome-encoded traits.
- Using genome-based approaches to design and manipulate microbiomes and advance their ability to produce or manufacture biofuels and bioproducts.
- Establishing how microorganisms interact with each other and with plant hosts to leverage and manipulate these relationships for improved bioenergy feedstock production.
- Elucidating the functions of microbial and phage dark matter in microbiomes to improve prediction of genes, pathways, and proteins of unknown function across uncharacterized environmental communities.
- Leveraging the power of emerging artificial intelligence (AI) approaches to understand the biology of microbial cells, populations, and complex communities, so they can be manipulated for beneficial outcomes.

About the Workshop

In support of these efforts, BSSD convened a 3-day Engineering Microbial Communities virtual workshop in December 2024 (see Appendix A: Workshop Agenda, p. 51). BSSD encouraged the 35 participants from academia, national laboratories, government, and industry to think deeply about the potential of microbiome engineering to achieve DOE missions (see Appendix B: Workshop Participants, p. 54). The workshop was one of a four-part Frontier Science for the Bioeconomy series hosted by BSSD (see p. ii). The series defines frontiers of plant science, agriculture,

synthetic biology, biobased materials, and environmental microbiome science that can unlock an innovative, resilient U.S. bioeconomy.

Workshop discussions focused on (1) defining state-of-the-art technologies to design, build, test, and learn (DBTL) from microbiome experiments; (2) exploring the potential of microbiome design to solve BSSD-relevant science challenges; and (3) identifying the fundamental knowledge and technical innovations needed to realize the real-world applications and benefits of microbiome engineering.

This report is organized into six chapters, which outline the current state-of-the-art in microbiome research, knowledge gaps, technical limitations, research opportunities, and needs that emerged from these discussions. Chapter 1 introduces microbiomes and discusses how microbiome engineering can provide new services across a range of economic sectors (see p. 1). Chapter 2 describes how microbiomes can be engineered using natural isolates to provide scalable technologies that advance the bioeconomy (see p. 11). Chapter 3 covers genetic programming technologies and the ways that synthetic biology can advance microbiome engineering (see p. 19). Chapter 4 outlines challenges in translating laboratory findings to real-world settings and the research needed to engineer microbiomes with predictive functions across scales (see p. 29). Chapter 5 highlights how modeling and AI are poised to advance predictive microbiome engineering (see p. 39). Chapter 6 discusses resources needed to accelerate microbiome research and realize the full potential of engineered microbiomes (see p. 47).

Engineering Microbiomes for National Needs

A strong consensus among workshop participants is that research targeting current knowledge gaps will create new opportunities for programming microbiomes to provide beneficial services across the U.S. bioeconomy. These opportunities include engineering microbiomes to support:

- Advanced biomanufacturing and bioproduction, such as the extraction of critical minerals and the valorization of complex feedstocks to useful chemicals, materials, and products.
- Distributed sensing, waste upcycling, and remediation strategies that support growth of a circular bioeconomy.
- Introduction and manipulation of specific microbial traits to improve plant and soil health, decrease reliance on energy-intensive chemical fertilizers, and support crop and ecosystem stress resilience.

Underlying this potential are rapid advancements in microbiome engineering. Key developments include wide-ranging strategies to engineer microbiomes as well as diverse systems for monitoring their behaviors at different scales and with varying levels of complexity. In addition, precision gene-editing tools are enabling genomic engineering and genetic circuit design rules to facilitate the programming of complex dynamic behaviors in laboratory strains. Other advances include computational models that anticipate the metabolism of individual microbes and capture cell–cell interactions in communities. Although in their infancy, initial applications of AI and machine learning (ML) hold great potential for accelerating predictive microbiome engineering.

Knowledge Gaps

Wide-ranging knowledge gaps currently limit the ability to predictively engineer communities. For example, researchers need a better mechanistic understanding of how genomic sequences translate to the production of biomolecules underlying phenotypes of individual cells, cell–cell interactions, and emergent community behaviors. Another limitation is the ability to translate research findings from the Petri dish to the environmental scale. The need to anticipate microbial activities in real heterogeneous, environmental settings is challenged by a lack of robust models that go beyond predicting spatiotemporal behaviors of pairs of laboratory strains in homogeneous laboratory settings.

Key knowledge gaps include:

- Unifying theory of how communities function as integrated systems with clockwork parts to enable engineering for real-world applications.
- Approaches to measure and predict the impact of individual microbes on community structure, function, and stability.
- Understanding of biotic and abiotic controls on community structures, functions, and stabilities in real, heterogeneous environments.
- Knowledge of the mechanisms underlying cross-kingdom interactions among bacteria, fungi, microeukaryotes, archaea, and plants that control community composition.
- Insights into the ecology of mobile DNA and RNA exchanged across the tree of life, including environmental DNA, viruses, and conjugative vectors.
- Knowledge of the sequence-function relationships for natural and engineered genetic parts and regulatory elements, including their dynamic activities.
- Control theory for anticipating how DNA-encoded parts can be composed into Boolean logic gates for engineering dynamic functions in communities.
- Foundational understanding of fungal ecology, gene transfer, and regulatory networks as well as interactions with plants and other microbes.
- Knowledge enabling biomolecular insights from laboratory studies to scale to field applications, including insight into the ways communities evolve at different scales.
- Limited fundamental understanding of the effects of spatial complexity and environmental heterogeneity on community functions across environments.
- Strategy to anticipate temporal and emergent behaviors of engineered microbiomes in different settings and across length scales.
- Modeling strategies that accurately anticipate emergent community behaviors by incorporating critical aspects of environmental complexity.
- Establishment of important biotic and abiotic variables to leverage for microbiome modeling that reduce the complexity of the vast parameter space.
- Robust strategies for integrating data from different experiments within individual models and synergizing models that consider different parameter and length scales.
- AI and ML computational approaches to enable the discovery and prediction of emergent properties of microbiomes to derive engineering design rules.
- Approaches for decreasing the data dimensionality required to train AI/ML models that predict microbiome structures, functions, and stabilities.

Technical Limitations

Technical innovations can accelerate microbiome engineering and strengthen connections between each step in the DBTL engineering paradigm. Current technical limitations include (1) generating the big data necessary to create AI/ML models that can anticipate emergent behaviors of engineered microbiomes and (2) identifying the correct characterization data to train models. Standards are necessary to deal with microbiome complexity and variation across settings. For example, efforts to simplify comparisons across experiments and datasets will require biological standards (e.g., community standards), material standards (e.g., artificial soils), and computational standards (e.g., data formats and metadata). Other critical needs underpinning precise genetic programming of microbiomes include synthetic biology tools and well-specified genetic parts to simplify genetic manipulations across the full diversity of microbes found in nature. Additionally, nondisruptive sensors are needed to report on dynamic community functions.

Key technical limitations include:

- Strategies to engineer microbial communities to synthesize target biomolecules through

multispecies interactions, thereby gaining efficiencies in biomanufacturing and achieving better bioprocess applications.

- Model microbial communities that can be reproducibly interrogated in parallel studies to understand controls on microbiome structure and function arising from interactions among species and with the environment.
- High-throughput technologies to detect the host range of mobile DNA and genomic editing achieved in microbes that take up engineered DNA.
- High-throughput screening and phenotyping methods that accurately capture the heterogeneity and complexity of natural environments.
- Genetic-editing tools and parts for undomesticated, non-model microbes and their hosts (e.g., plants), such as technologies to enable efficient DNA uptake, chromosomal modifications, and genetic programming.
- High-throughput, cross-scale platforms that leverage automated and self-driving workflows to accelerate the microbiome engineering DBTL cycle and enable generation of big data for training models.
- High-throughput platforms that leverage automation to develop well-specified genetic regulatory elements, protein coding sequence, and genetic circuits that can be used across diverse microbes found in environmental communities.
- Protocols and technologies for scaling findings from the laboratory to real-world environmental settings, which present greater complexity and variation.
- Sensor families for *in situ* monitoring of microbiome functions including genetically encoded sensors that convert the detection of user-defined inputs (e.g., chemicals and metabolites) into easy-to-detect outputs that can be observed nondisruptively.
- AI-assisted experiment platforms that accelerate the DBTL cycle of microbiome engineering using AI-driven decisions.
- Tools for real-time and nondestructive sensing and monitoring of microbiome functions at any scale to continuously monitor the molecular signals and processes controlling microbiome functions.
- Control theory that informs the design of genetic circuits that endow communities with user-defined stability, robustness, and dynamic functions across a range of environmental conditions.
- Robust safety measures and mitigation tools to constrain the growth of engineered microbiomes to user-defined operational environments for finite durations.

Research Opportunities

Numerous research opportunities could advance microbiome engineering and lead to innovations in bioenergy, biomanufacturing, agriculture, and environmental resilience. One example is the development of well-characterized genetic parts for engineering microbiomes using synthetic biology. Such resources—including atlases for DNA regulatory elements, coding sequences, and mobile DNA—would accelerate the DBTL cycle.

Another opportunity lies in technology investments that increase the throughput of phenotyping by generating dynamic data about microbiome functions in testbeds of different complexities and scales ranging from laboratories to greenhouses to field sites. Such technologies are critical for training AI/ML models for microbiome engineering and for rigorously validating those models across different application sites.

New modeling opportunities could transform microbiome engineering into a scalable, data-driven discipline with real-world impact. Further advancements are possible by better leveraging existing DOE resources such as the Joint Genome Institute; the Environmental Molecular Sciences Laboratory; the Nanoscale Science Research Centers; and world-class computing capabilities, including the Argonne Leadership Computing

Facility, Oak Ridge Leadership Computing Facility, and the National Energy Research Scientific Computing Center.

Additional key research opportunities include:

- Designing microbial communities to deliver reproducible, dynamic functions under natural, variable environmental conditions.
- Developing strategies to engineer communities to present stable and resilient behaviors and services in the face of competition with native microbes.
- Establishing strategies to engineer interactions between bacteria, fungi, and plants to improve plant health, agricultural production, and ecosystem resilience.
- Constructing a DNA mobilome atlas and toolkit to inform and advance genetic programming, microbiome adaptation, and safety feature development.
- Creating a genetic parts foundry with standardized components to enable predictive genetic engineering across undomesticated archaea, bacteria, and fungi.
- Leveraging distributed field sites to understand biotic and abiotic controls on microbiome functions.
- Developing a framework for multispecies bioproduction to leverage interspecies metabolic networks for greater process efficiencies.
- Advancing multiscale modeling to anticipate dynamic microbiome functions—including ecosystem resilience, productivity, and soil health—in real-world environments.
- Leveraging AI agents to control and drive microbiome engineering laboratories and to integrate microbiome big data across the community.
- Developing AI/ML models and physics-based control theory for genetic circuit design that uses molecular details to anticipate complex community behaviors.

Path Forward

The strategic priorities highlighted in this workshop report are expected to enable transformative research that realizes the full potential of microbiome engineering across complex, heterogeneous environments. The workshop highlighted the enormous opportunities that exist for microbiome engineering to accelerate DOE missions in energy security and innovation. Capitalizing on these opportunities will require investments in foundational ecology, systems biology, and synthetic biology research.

Among workshop participants, there was a strong consensus that development of standardized tools and datasets will be necessary to improve integration across experimental scales and biological domains. Additionally, predictive engineering will require a deeper understanding of gene flow within communities, microbe–microbe interactions, host–microbe interactions, and environmental constraints that influence engineered microbiome stability and function. Finally, emerging AI and ML approaches hold great promise for synergizing with all aspects of the DBTL microbiome engineering cycle.

Chapter 1

Harnessing Environmental Microbiomes To Advance the Bioeconomy

Microbes are ubiquitous, with total abundances exceeding 10^{30} cells on Earth (Flemming and Wuertz 2019). They form complex communities in almost every niche on the planet, ranging from soil, plant tissues, the atmosphere, oceans, and groundwater to the built environment and the human gut (see Fig. 1.1, this page). At the centimeter scale, microbial communities present extraordinary biological complexity. For example, 1 gram of soil is estimated to contain $\geq 10^4$ microbial species (Roesch et al. 2007). These microbes include bacteria, archaea, fungi, and protists—making soil the world's most biodiverse ecosystem and largest reservoir of genomic information and biotechnological resources. For this reason, microbiome engineering research and discovery frequently focus on microbial interactions within soil or rhizosphere-associated processes. However, general principles derived from such work are applicable to all microbiomes and can broadly transform research across the bioeconomy where complex biological interactions play a significant role.

Fig. 1.1. Microbiomes Are Everywhere. Microbial communities can be found aboveground on plant stems and leaves and in belowground environments such as bulk soil and around plant roots. Microbiomes are found in and on vertebrates who in turn transport these communities. They also are in and on built infrastructure, such as the pipes used to transport fossil fuels and chemicals. Microbiomes shape large-scale environmental processes, affect public health, and play an important role in industrial and biotechnology applications. Advances in the ability to measure, analyze, and manipulate microbiomes continue to provide exciting opportunities for scientific discovery.

1.1 Microbial Complexity in Heterogeneous Habitats

The spatiotemporal variability in microbial community composition and function is extraordinarily complex. This is especially true in heterogeneous habitats such as soil where hot

spots of activity arise from nutrient inputs in the form of plant root exudates, organic matter decomposition, weathering, and fertilization (Kuzyakov and Blagodatskaya 2015). Additional heterogeneity stems from variation in an environment's chemical properties (e.g., electron donors, nutrient content, hydration, and pH) and physical properties (e.g., texture, aggregation, and permeability), which vary dynamically across micro- and macroscales (Quesada et al. 2010). This heterogeneity can lead to complex feedbacks in interactions among microbes and with chemical and physical environmental properties, resulting in emergent and difficult-to-predict responses within communities.

1.2 Omics-Enabled Insights

Microbial communities catalyze millions of chemical reactions, often simultaneously, and they do this at room temperature and ambient pressure, making them more complex and efficient than any human-built chemical refinery. Their remarkable adaptability and functionality enable them to inhabit most environments on Earth and to efficiently utilize available energy through complex cell–cell and cell–environment interactions. Although harnessing this catalytic potential currently remains out of reach, advances in genomic technology and data science are creating opportunities to more efficiently leverage the beneficial activities of microbial communities, both in environmental and industrial settings.

While microbial interactions and their emergent corollaries are usually treated as a black box, systems biology now allows for the routine characterization of environmental microbiomes. Omics-based approaches can provide powerful insights into what microbes are present in environmental samples, what biochemical reactions each microbe might catalyze, which biomolecules the community is synthesizing, and how the community is impacting the surrounding environment (Daniel 2005; Bundy 2008; Jansson and Hofmockel 2018).

1.3 Engineering Approaches

Researchers currently use numerous approaches to engineer microbiomes (see Fig. 1.2, this page). In the

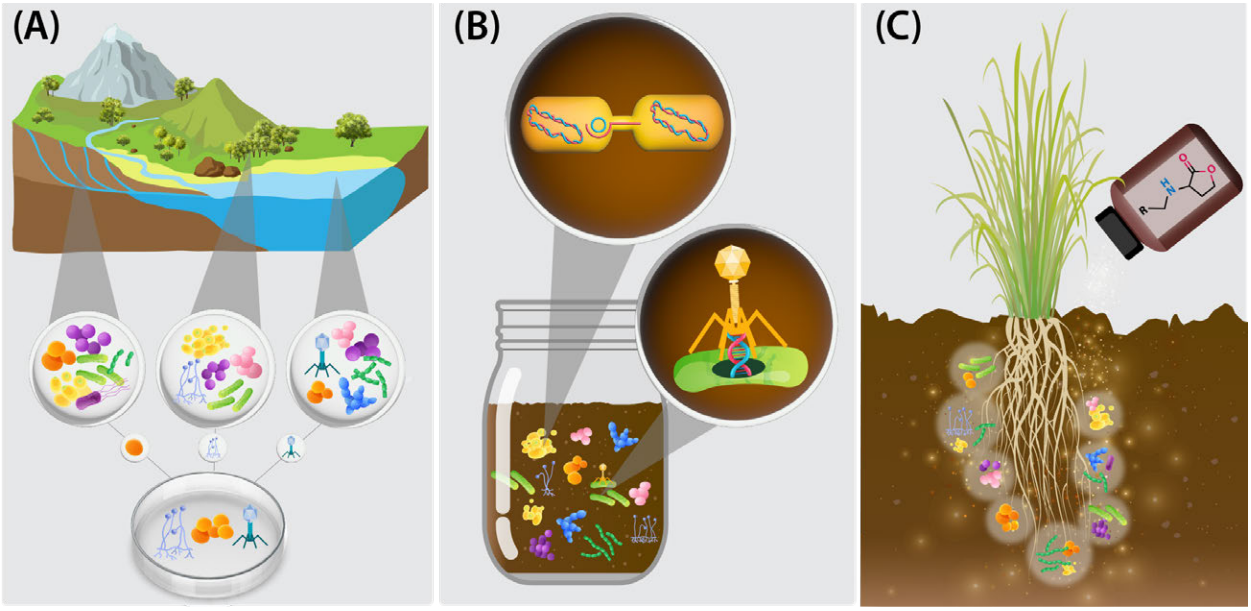


Fig. 1.2. Three Strategies for Engineering Microbiomes. (A) Microbial communities can be isolated from the environment, and microbes from these communities can be mixed to create synthetic communities (SynComs). SynComs bring together the functional characteristics of microbes to create microbiome behaviors that are distinct from those found in nature. **(B)** One or more microbes in a community can be genetically modified by introducing natural or engineered DNA to endow those microbes with new genetic traits, resulting in beneficial behaviors. This DNA can be introduced by programming cells to exchange DNA using a process called conjugation or by introduction of a bacterial virus, either native or engineered. **(C)** Chemicals or root exudates can be altered to affect the structure and function of the microbiome.

first approach, simplified communities are assembled with decreased taxonomic complexity compared to those found in natural environments; the resulting microbial mixtures are referred to as a constructed community or a synthetic community (SynCom). These simplified communities can be constructed in multiple ways. Isolates from natural environments can be mixed in different combinations (Karkaria et al. 2021), communities can be created by decreasing the concentrations of microbes in a microbiome via dilution (Philippot et al. 2013), or consortia can be enriched from environmental samples (McClure et al. 2020).

In the next approach, communities can be modified by introducing new genetic materials. For example, phages can be added to selectively remove or suppress specific microbes from complex communities and to augment the function of specific members of a community (Tanaka et al. 2024). This approach has the potential to enable the precise manipulation of microbiome composition and function without broadly disrupting ecosystem balance (Tanaka et al. 2024). Phages and microbes can also be genetically programmed *in vitro* and introduced to communities to present new traits or to share DNA with native microbes to endow them with new behaviors (Charbonneau et al. 2020). Further, one or more microbes can be genetically reprogrammed directly *in situ* (Rubin et al. 2021).

Lastly, prebiotics, chemicals, biomolecules, plant exudates, and plant engineering are harnessed to manipulate root, rhizosphere, and soil microbiomes (Vila et al. 2020; Griffin et al. 2024).

1.4 The Potential of Engineered Microbial Communities

Understanding the design rules for microbial community assembly would enable microbiome-based solutions for a range of bioprocess applications across several economic sectors, including agriculture, public health, defense, and environmental management. As illustrated in Fig. 1.3, p. 4, a deeper understanding of microbiome engineering principles could support a growing bioeconomy in the following ways:

- Elucidating fundamental biological, ecological, and evolutionary principles that accelerate biodesign.
- Producing energy, chemicals, and materials from biomass feedstocks and other waste streams.
- Detecting and degrading harmful toxins and contaminants.
- Improving and monitoring soil and plant health while minimizing energy inputs for agriculture.
- Enhancing ecosystem resilience, such as agrosystems that produce food, fuel, and other bioproducts.
- Detecting and mitigating microbial threats in the environment.

A grand challenge for microbiology and microbiome science is to develop a unifying theory that explains how natural and engineered microbial communities function as integrated systems in different environmental settings.

Elucidating Biological Principles

Engineered microbiomes represent powerful experimental systems for understanding the principles underlying microbial behaviors fundamental to biotechnology, plant–microbe interactions, and resilient ecosystems. Current knowledge about microbial community interactions and their emergent phenotypes is largely constrained by the biotic, physical, chemical, and spatiotemporal conditions examined in individual studies, rather than by a unifying theory that captures how communities function as integrated systems (Agoussar and Yergeau 2021; Leggieri et al. 2021). Identifying the governing mechanisms that translate across this vast parameter space is a grand challenge in microbiome research. New microbial engineering and modeling approaches provide exciting opportunities to illuminate the genomic basis of organismal phenotypes, multispecies interactions, and emergent

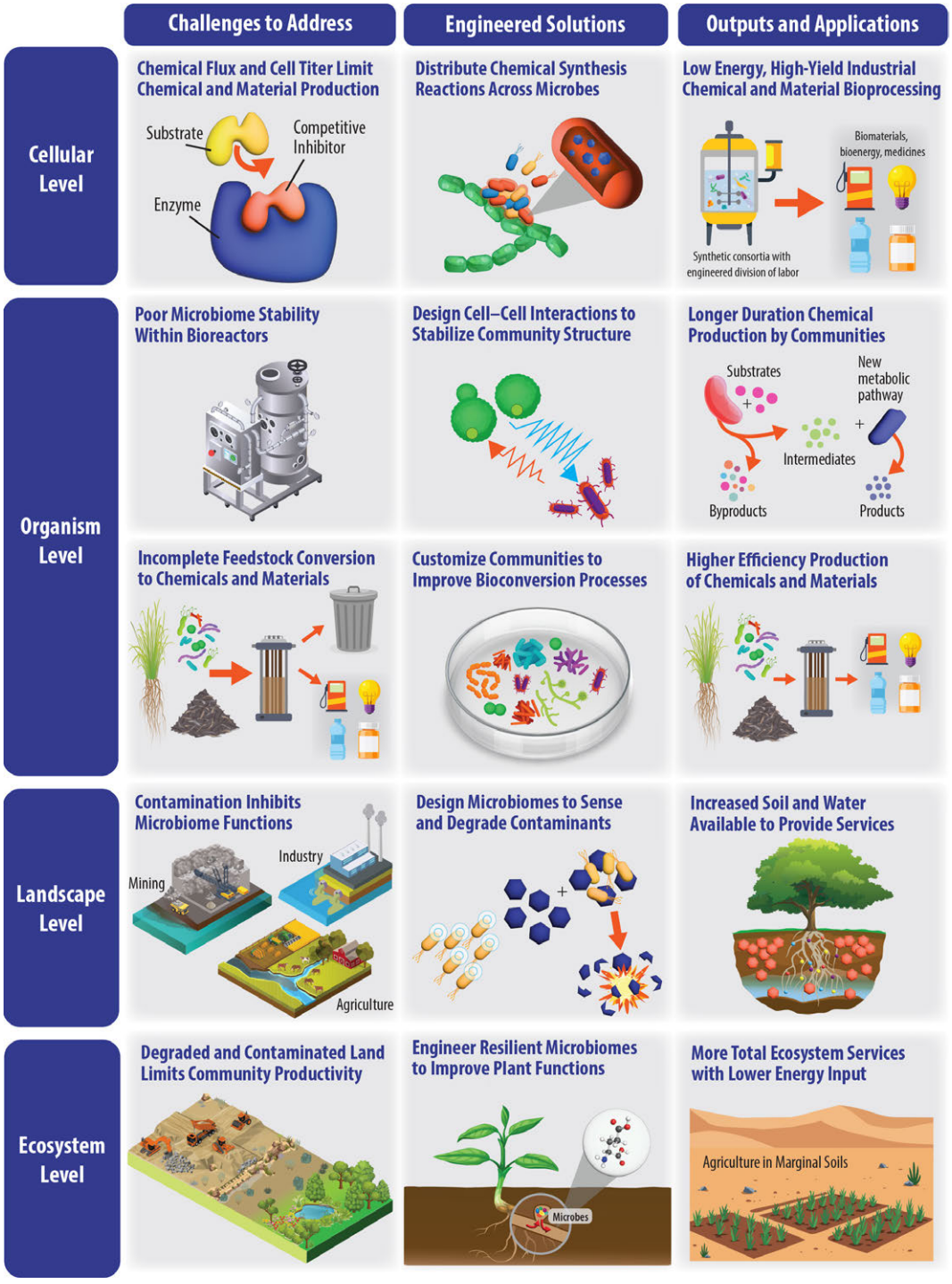


Fig. 1.3. Microbial Communities Demonstrate Potential to Transform the Bioeconomy. At every scale, engineered microbial approaches offer solutions that yield beneficial bioeconomic outputs and applications. At the **Cellular Level**, microbial genes and metabolic pathways can be engineered to produce chemicals using diverse feedstocks. At the **Organism Level**, synthetic cocultures and communities can be designed to efficiently produce diverse bioproducts in bioreactors. At the **Landscape Level**, small-scale field deployments of engineered microbiomes can sense and degrade soil and water contaminants. At the **Ecosystem Level**, landscape-scale microbial applications (1) improve soil and plant health, resulting in decreased energy inputs for bioenergy crop production, and (2) improve the resilience of ecosystems that provide valuable services.

behaviors in microbial communities, which are the foundation of resilient microbial systems and innovative biotechnology solutions (van den Berg et al. 2022).

Engineered microbial consortia can serve as simplified, tractable models of natural ecosystems, allowing researchers to test ecological theories under controlled conditions. For example, engineered communities with defined compositions enable the isolation of specific ecological interactions from environmental noise and variability, shedding light upon community assembly rules that govern natural microbiomes in biotechnology-relevant settings (Johns et al. 2016; Rodríguez-Ramos et al. 2025).

Microbial communities can also be designed and manipulated to understand how their genomic contents and epigenetic states underlie their emergent functions and dynamics. Mechanisms that allow microbes to persist, interact, and function can be identified and demonstrated using engineered microbial systems (McClure et al. 2022). In addition, understanding these microbial mechanisms can enable the identification of organisms that provide the same community services and the establishment of functional redundancies in communities (Louca 2018). These engineered consortia can demonstrate how higher-order interactions promote stable coexistence in multispecies communities (Kelsic et al. 2015). Such findings must be extended to understand how communities exhibit resilience to stressors and disturbances common to bioproduction and environmental ecosystems.

Scientists cannot yet predict how the addition or elimination of individual microbial community members impacts the dynamic structure, function, and stability of microbiomes. To leverage multispecies microbial communities for targeted engineering approaches, a basic understanding of the plug-and-play parameters for mixing microbes is needed. Such knowledge would support the design of stable, engineered communities that can survive in both natural and biotechnology environments while behaving in predictable and controllable ways.

In addition, realizing how genomic content translates into microbe–microbe and microbe–host interactions in the environment can help researchers better comprehend dispersal, invasion, and persistence, which is critical for managing bioinoculants in the wild. Similarly, understanding evolution, mutation, and adaptation via horizontal gene transfer (HGT) elements will be important for maintaining the functions of chassis organism stabilities within any environment where they are applied, including bioproduction facilities.

The use of engineered microbiomes to understand and manipulate microbial properties will aid in developing communities that provide a precise function and persist only in specific environmental contexts for defined durations. To elucidate these and other biological principles, the design-build-test-learn (DBTL) cycle must be employed with engineered microbiomes (see Fig. 1.4, p. 6). Through this iterative process, principles can be elucidated for augmenting community functions by altering community structure and through genetic programming of microbes in communities. Precision control over microbiomes may require the development of safety switches or other novel biocontainment strategies to prevent unintended release and uncontrolled growth of engineered microorganisms (Lee et al. 2018).

Identifying the ecological principles underlying the diverse microbial activities found in microbiomes is fundamental to engineering predictable and reproducible safety outcomes that support the production of chemicals and materials, accelerate environmental remediation, and facilitate interactions that benefit plant and ecosystem health and resilience.

Producing Energy, Chemicals, and Materials

Biomanufacturing commercially valuable products is dominated by applications that use a small number of industrial workhorse strains including *Escherichia coli*, *Saccharomyces cerevisiae*, and *Pichia pastoris* (Zhang, C. et al. 2022). Typically, targeted metabolic engineering within individual strains is used to enable production of proteins, enzymes, or small molecules. While examples of industrial-scale bioproduction (>1,000-liter reactors) exist—such as fermentation to ethanol, lactic acid, and butanediol—many processes are hard

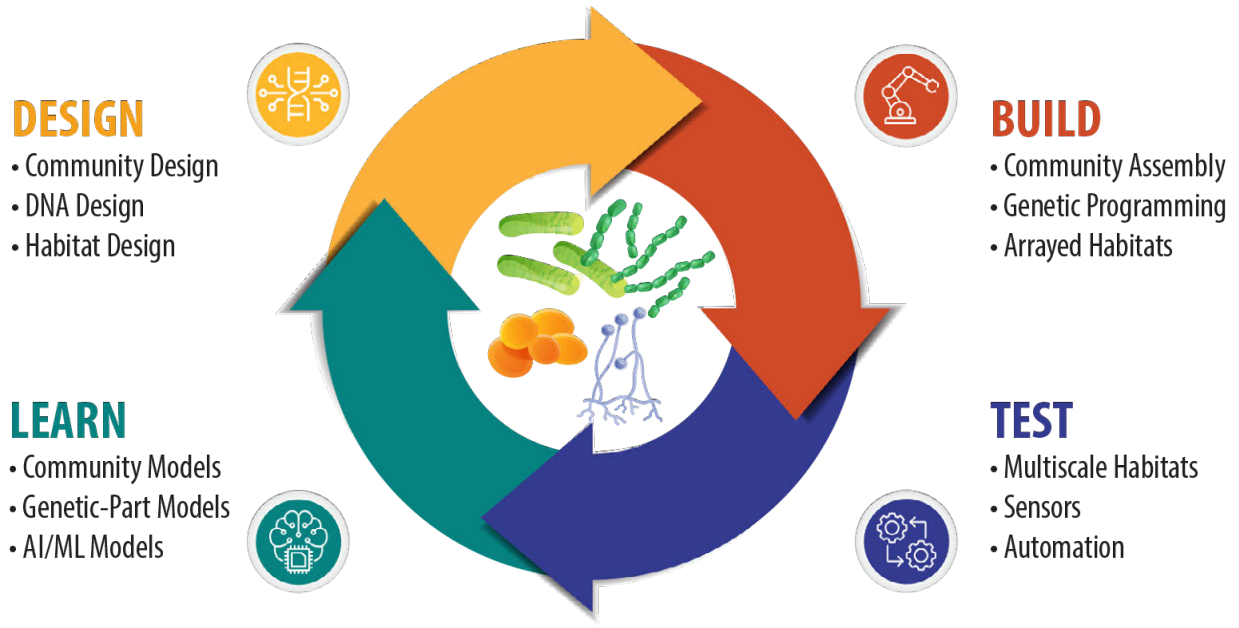


Fig. 1.4. The Microbiome Design-Build-Test-Learn Cycle. To **Design** microbial communities, the first step is to determine the set of microbes that will be incorporated into the sample, the engineered DNA that will be delivered to the microbiome, the strategy for achieving efficient DNA delivery, and the habitat and conditions that will be used to assess whether the desired microbiome functions have been achieved. This design is performed using computational models. To **Build** microbial communities, microbial samples are generated *in vitro*. DNA is delivered when genetic programming is incorporated into the design process, and samples are arrayed to allow for parallel analysis of multiple community samples. High-throughput community analysis is required as biomolecular, cellular, and microbiome design spaces are vast. To **Test** microbial community functions, their dynamic functions are monitored in habitats that span a wide range of length scales, including centimeter- to meter- to field-scale settings. At each length scale of characterization, different microbiome functions are monitored using distributed sensors that are compatible with automation and high-throughput screening. To **Learn** from the testing of microbial communities, the models originally used for design are refined using data generated from high-throughput experiments. Models are needed to accurately anticipate community-level behaviors across dynamic environmental settings and to anticipate the function of genetically encoded parts used to program communities. Artificial intelligence (AI) and machine learning (ML) are promising new tools for achieving model accuracy.

to scale beyond the laboratory (Carruthers and Lee 2022). In addition, the array of chemicals capable of being produced by a single organism is limited by the thermodynamic and metabolic versatility of individual hosts targeted for metabolic engineering, and the accumulation of toxic byproducts can limit cellular yields (Ling et al. 2014). Bioproduction-capable microbes are optimized to grow on a narrow range of substrates, such as sugars, that limit technoeconomic feasibility for many production processes. Despite some progress in product synthesis via variable mixed substrates, current approaches remain limited to a few examples (Ni and Prather 2024).

Notably, microbes in nature almost exclusively exist as communities with tightly coupled metabolic dependencies (Gralka et al. 2020). It is through such interactions that microbial communities can achieve remarkable metabolic feats, often thriving at the edge of thermodynamic constraints (Großkopf and Soyer 2016). Microbial communities may thus offer solutions to the limitations of single-chassis bioproduction, particularly when seeking to valorize complex feedstocks and waste to useful chemical compounds, materials, and products.

Microbial communities are inherently agile, resilient, and capable of synthesizing an array of complex

products through metabolite exchange. Engineering such complex dependencies is beyond current technologies but might offer opportunities to design bioprocesses that overcome limitations of single, engineered strains. Through a design-based, divide-and-conquer approach, microbial communities could upcycle waste via enzyme deconstruction specialists that valorize molecules as complex as lignocellulose (Lankiewicz et al. 2022) or synthetic plastics and polymers (Bergeson and Alper 2024). Once a substrate's complexity has been decreased, other community members could be used to convert metabolic waste products into sugars, monomers, or fermentation products, which could be upgraded to value-added compounds by yet another group of microbes (Agena et al. 2024). Similarly, microbial communities could be leveraged to control gas fermentation from lignocellulose or waste remediation (Fackler et al. 2021) or to extract and concentrate high-value rare earth elements from dilute sources. Microbial communities could also be programmed to grow into regenerative biocomposite structures with new properties and functionalities that go beyond known living materials (McBee et al. 2021).

Detecting and Degrading Harmful Toxins and Contaminants

DOE has a significant environmental legacy arising from decades of nuclear weapons production and energy research. Long-term stewardship thus remains critical to mitigate health risks and enable economic redevelopment (U.S. DOE 2025). Beyond DOE's legacy, pollution associated with U.S. industrial activity is considerable. For example, more than 1,000 sites are included in the Superfund National Priorities List, an effort created to help the U.S. Environmental Protection Agency determine which sites may require significant cleanup efforts (U.S. EPA 2025). Thousands of additional smaller or less-documented polluted sites also exist. The scale of contamination at all these locations is an environmental burden that requires sustained investment and research into new technology to ensure safe, permanent solutions. In addition, municipal waste is expected to grow by 70% over the next 25 years, further amplifying these challenges and driving a need for a circular approach that supports robust production practices and efficient recycling (Kaza et al. 2018).

In this context, biotechnology could broadly enable distributed, real-time chemical detection and monitoring. Environmental microbes have evolved to regulate activities by sensing their local conditions using protein sensors (Stock et al. 2000), with an estimated 1 to 2% of all microbial genes representing such sensor systems (Salvado et al. 2015). In environments like soil, more than 1 million different sensor systems can be found in a handful of material. This natural sensor diversity represents an emerging opportunity to create families of real-time sensors using synthetic biology. Cells can be genetically programmed to couple their sensed information to easy-to-detect outputs including visual, chemical, and electrical signals (Del Valle et al. 2021). Critically, this sensing capacity can be paired with engineered responses (e.g., *in situ* contaminant degradation) to enable autonomous environmental management. Further, cells and communities can be programmed to couple this sensing to new functions, such as the conversion of harmful chemicals and materials into nontoxic forms or their upcycling into valuable products.

Notably, environmental contaminants and toxins have very distinct chemical structures, which can present complex spatial and temporal dynamics in their movement through the environment (Vörösmarty et al. 2010). Contaminants include chemicals such as solvents (e.g., benzene, toluene, trichloroethylene, and dioxins), heavy metals (e.g., arsenic, mercury, and lead), and polyfluoroalkyl substances (i.e., forever chemicals), each of which presents unique remediation challenges. In addition, materials such as synthetic polymers have no natural surrogates and pose a growing challenge (Cottom et al. 2024). A one-size-fits-all approach to contaminant remediation is thus not feasible given the large diversity of chemical structures, recalcitrance, and potential lack of naturally occurring enzymes that catalyze chemical activation steps. Because of their diverse sensing abilities and enormous catalytic potential, microbiomes represent a unique opportunity to quickly and affordably engineer solutions tailored to individual remediation needs.

Recent breakthroughs in enzyme and cellular engineering promise to accelerate remediation using microbiome engineering. The discovery of genome-engineering techniques, such as CRISPR-Cas systems (Knott and Doudna 2018), enables the introduction of genetic modifications into the chromosomes of microbes, augmenting their metabolic capabilities. Scientists could engineer complex microbial communities to degrade multiple contaminants and colocalize the beneficial degradation and processing capabilities of different microbes to achieve numerous goals simultaneously. Recent work has demonstrated that such engineered communities can maintain their functional capacity even under complex and fluctuating environmental conditions. Environmental bacteria engineered to enhance degradation perform robustly over multiple years in simulated environments, underscoring the potential for stable, long-term pollutant remediation *in situ* (Chemla et al. 2022). Furthermore, *de novo* enzyme designs that leverage artificial intelligence advances have begun to revolutionize enzyme discovery and can be used to engineer enzymes that target recalcitrant materials, such as polyethylene terephthalate (Lu et al. 2022), which accounts for more than 10% of global solid waste (MacLeod et al. 2021).

Improving and Monitoring Soil and Plant Health

Microbial community engineering presents a potential transformative opportunity for enhancing soil and plant health. It reduces energy requirements for managing soils including decreasing the need for fertilization and irrigation (Capdevila-Cortada 2019), which account for more than 10% of the energy consumed in the chemical industry (Lim et al. 2021). Innovative genomics- and ecology-informed approaches can potentially guide microbiome engineering, enhancing protective and adaptive functions to plants and soils in complex and dynamic environments. Microbiomes also could boost bioenergy productivity through opportunities aimed at controlling the activities of diverse archaea, bacteria, fungi, protists, and viruses in soil communities (Roesch et al. 2007).

By promoting specific plant–microbe interactions or introducing specific microbial traits and functions into ecosystems, engineered microbial communities could improve plant and soil health through a range of mechanisms. Communities could enhance nutrient availability and uptake (i.e., nitrogen mineralization; Jilling et al. 2018), confer increased stress- and disease-resistance to plants (Vannier et al. 2019), and improve plant–water relations (Colica et al. 2014). Engineered microbes also could foster symbiotic relationships and interactions that facilitate soil and bioenergy crop health and productivity. Notably, microbiome engineering strategies could improve biological nitrogen fixation in nonleguminous plants and thus decrease reliance on chemical fertilizers (Wen et al. 2021; Bloch et al. 2020), which consume >1% of global energy annually. Strategies could target phosphorus mineralization and solubilization (Alori et al. 2017), thereby lessening environmental impacts and promoting sustainable agricultural practices. The recent discovery of a nitrogen-fixing organelle in a marine alga represents a new opportunity to consider how to manipulate eukaryotes to improve biological nitrogen fixation (Coale et al. 2024). Further, microbial interventions to trigger beneficial strains to bolster crop resilience against environmental stresses could enhance nutrient availability and resistance to soil toxins and pollutants.

Beyond increasing soil services, such as bioenergy productivity and bioconversion efficiencies, microbiome engineering could play a pivotal role in stabilizing soil organic matter and enhancing overall soil health (Bhattacharyya et al. 2022). Soil quality often is degraded through loss of organic matter, changes in pH, increases in toxic metals and salinity, and loss of nutrients (Qadir and Schubert 2002). By minimizing degradation, microbiome engineering could improve the ability of soil to cycle nutrients and retain water (Indoria et al. 2020), which is crucial for facilitating plant growth and agricultural yield in arid climates and during drought conditions. Microbes could be engineered to support community and plant resistance to pollutants and to enhance crop health and productivity. Engineered microbiome technologies that multiplex services could not only enhance soil health, but also foster safer environments for crops and bioenergy ecosystems and report on organic matter storage or remediation in soils (Keenor et al. 2021).

Enhancing Agrosystem and Ecosystem Resilience

Ecosystems are undergoing increasingly dynamic shifts due to many factors including extreme events, natural resource extraction, habitat degradation, and the introduction of invasive species, pathogens, and pests (Weiskopf et al. 2020; Staudt et al. 2013). Such shifts lead to changes in ecosystem structure and function, often limiting the ability of native organisms to persist within a given habitat (Weiskopf et al. 2020; Franklin et al. 2016; Van Nuland et al. 2023) and necessitating tools that enhance ecosystem resilience. Resilience is the capacity of an ecosystem to maintain its identity and function despite disturbance (Delettre 2021) and is influenced by complex, multiscale interactions among organisms and their environments (Thorogood et al. 2023).

Microbiome engineering strategies to enhance resilience in one ecosystem may be ineffective or even detrimental in another. A lack of understanding of the drivers of the inherent complexity of microbial systems currently limits the development of useful, scalable, and adaptable tools that support ecosystem responses (Silverstein et al. 2023). Microbiome engineering offers a promising path forward to understanding how ecosystem resilience can be achieved. By manipulating microbial community composition and function, beneficial interactions might be enhanced to support ecosystem stability under stress.

In natural and managed systems, targeted microbiome interventions could mitigate the effects of invasive species, buffer against pathogen outbreaks, and accelerate recovery in degraded landscapes (Peixoto et al. 2022). Engineered microbial communities could be tailored to support conservation goals, stabilize native communities' response to change, aid in species migrations, and improve establishment success in restoration and bioremediation projects (Barnes and Tringe 2022).

The potential for microbiome engineering expands even further when combined with advances in plant genetic engineering. By co-engineering plant traits that modulate root exudation in parallel with microbial communities that respond to these chemical cues, it is possible to design plant–microbe partnerships tailored to specific environmental challenges or stimuli. Such approaches could facilitate the expansion of species into new or degraded habitats, support the improvement of soil quality and formation, and enable the development of productive, resilient bioenergy landscapes. Proximity engineering of microbial fitness offers an additional layer of control by ensuring that engineered microbes can function only in close association with specific entities in the environment (e.g., plant hosts), providing built-in biocontainment. By restricting microbial activity to defined spatial contexts (Lee et al. 2018; Jones et al. 2024), proximity engineering minimizes the potential for unintended spread or ecological disruption. With coordinated efforts across synthetic biology, ecology, and data science, microbiome engineering could revolutionize ecosystem management, restoration, and adaptation in the face of accelerating global changes.

1.5 Synergies Beyond DOE Relevance

Microbiomes are an inherent and necessary element of all systems composed of, or influenced by, biological processes. Advances that overcome the grand challenge of targeted microbiome engineering will therefore have beneficial impacts far beyond the areas described above. For example, the development of a unifying theory that captures how communities function as integrated clockwork-like systems will provide insights that could be leveraged to advance human health solutions to acute, infectious, and chronic disease.

Understanding how to engineer microbiomes for resilience may offer novel approaches to stabilize beneficial microbes on human skin or in the human gut, which face frequent perturbations from diet, infection, and therapeutics (Langdon et al. 2016). Similarly, tools and frameworks developed for environmental microbiome manipulation (e.g., precision editing of community composition or targeted removal of harmful taxa using phages or antimicrobials) might translate to precision abilities to mitigate intended and unintended releases of harmful organisms. Other applications include monitoring agricultural crops, extracting critical minerals and materials, mitigating corrosion, terminating harmful algal blooms, or recycling waste.

Chapter 2

Strategies for Assembling and Characterizing Synthetic Microbial Communities

Microbiome engineering has advanced rapidly due to the development of tools enabling synthetic community (SynCom) design and multiomic profiling that links microbial identity to function. This chapter focuses on progress in engineering microbial communities using natural communities and isolates and identifies the knowledge gaps, technical limitations, and research opportunities to advance the growing bioeconomy.

2.1 State of the Art in Community Design

Optimizing microbial communities by creating combinations of cells that do not occur in nature is one way to engineer microbiomes without genetic manipulation. Researchers use multiple strategies to build these SynComs but with mixed results that limit opportunities to understand the underlying processes, interspecies interactions, and environmental contexts shaping microbial community structure, function, and stability.

Current Engineering Approaches

Approaches for generating SynComs are highly valuable in microbiome engineering because they strike a balance between ecological realism and experimental tractability. These engineered microbial consortia provide a well-defined framework for dissecting microbial interactions, functional redundancy, and emergent properties arising from interspecies dynamics. The potential to reproduce these community structures and functions across replicates and laboratories makes them ideal for comparative studies and high-throughput screening. Once a functional consortium is identified, its key members can be sequenced and recombined into new SynComs with altered compositions. This enables fine-tuning of community structure and function—facilitating rational design principles for microbiome engineering aimed at improving plant productivity, nutrient cycling, or stress resilience in managed ecosystems. These SynCom approaches bridge a gap between natural microbiome complexity and isolate collections, offering a pathway to leverage microbial functions for biotechnological applications. There are currently three methods to design SynComs (see Fig. 2.1, p. 12):

- Isolation and Reassembly
- Community Dilution
- Community Enrichment

Isolation and Reassembly

Microbial isolate libraries are systematically constructed for specific ecosystems or host organisms (e.g., plants), providing a foundational resource for studying microbial interactions in a controlled manner. These libraries are typically derived through extensive culturing efforts that target the diverse microbial communities associated with a particular niche such as soil or the plant rhizosphere. For instance, researchers have assembled culture collections from the root zones of

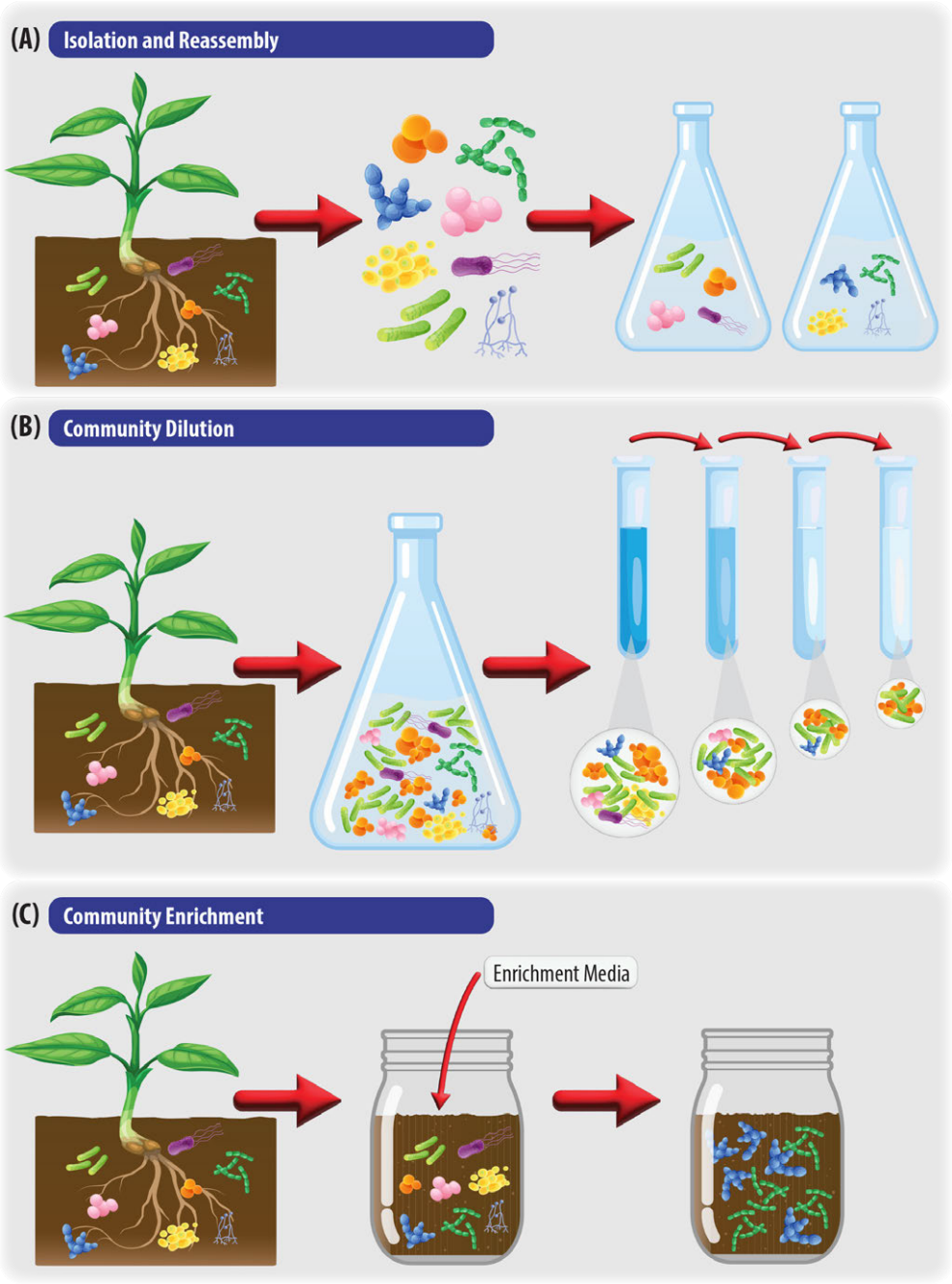


Fig. 2.1. Engineering Synthetic Communities Using Natural Isolates. (A) Isolation and Reassembly. Microbes can be isolated from environmental communities through culturomics, a process in which individual microbes are cultured in isolation. Isolates can then be reassembled into communities with different compositions. The number of possible reassembled communities increases combinatorially with the number of unique microbes in the original community. **(B) Community Dilution.** A natural community can be isolated and diluted to create communities with different combinations of microbes. In contrast to isolation and assembly, which is limited to microbes that can be cultured, this approach can be applied to microbes in any natural community because it does not require culturomics and the need to grow each community member individually before community assembly. **(C) Community Enrichment.** Communities can be grown in an enrichment medium to select for combinations of microbes that grow optimally under certain environmental conditions. With this approach, microbiomes that cooperate to achieve the targeted function are enriched from the natural community, thereby diluting out other microbes.

model species like *Arabidopsis* (Gates et al. 2023) as well as DOE-relevant energy crops including *Panicum virgatum*, *Sorghum bicolor*, and *Populus trichocarpa* (Carper et al. 2021). The goal of these culturing efforts is to capture both the phylogenetic and functional breadth of the native microbiome while maintaining cultivability under environmentally relevant conditions for downstream applications. Once assembled, these isolate libraries can be characterized genomically and phenotypically to gain insights into the mechanisms that underlie the collective function of the community and to inform the design of additional SynComs.

Assembled libraries of microbial isolates are not static. They are reused and iteratively refined to address a wide array of biological questions. By recombining isolates in SynCom experiments, researchers can assess how specific taxa or functions contribute to community assembly, community function, host performance, or resilience under environmental stress. These reusable culture collections facilitate reproducibility and scalability in microbiome research by enabling the comparison of results across laboratories and studies. In some cases, libraries may be expanded or tailored to include additional taxa identified through amplicon or metagenomic surveys, ensuring they remain representative of dynamic microbial populations and communities. New isolation techniques using reverse genomics and genome-informed antibody engineering enable expansion of these collections to organisms believed to be previously unculturable (Cross et al. 2019). Isolate libraries serve as both a tool for hypothesis testing and for enhancing host and ecosystem function.

Community Dilution

Communities of microbes can be created through dilution of natural communities (Hol et al. 2015). This dilution-to-extinction community design approach is widely used to explore links between microbial diversity and ecosystem function. By serially diluting inocula, community complexity is decreased, thereby allowing researchers to examine the functional consequences of microbial diversity loss. However, the microbial community that reassembles after inoculation is often shaped more by environmental filtering during incubation than by the initial dilution, which can limit the predictability of final community composition. Functional outcomes (e.g., effects on plant productivity) also tend to depend on specific characteristics of the original microbial community, with diversity loss sometimes leading to saturating or inconsistent responses. These dynamics underscore both the utility and limitations of dilution-to-extinction for constructing SynComs. While effective in reducing diversity and enriching for dominant taxa, the method is influenced by environmental context and should be applied with careful consideration of its constraints (Yan et al. 2015).

Community Enrichment

Culturing is performed to enrich specific community functions such as organic matter decomposition, nutrient cycling, or host resilience (McClure et al. 2022). Selecting for communities from complex environmental samples through cultivation under conditions that favor a desirable trait provides a powerful strategy for generating simplified, yet ecologically relevant, microbial consortia. Other examples include enriching communities on specific substrates (Zegeye et al. 2019) or to tolerate specific stressors (Carrell et al. 2022). These enriched consortia are not only easier to maintain and manipulate than the original highly diverse starting microbiomes, they can also retain key metabolic capabilities for relevant ecosystem services. Enrichment culturing can be a step toward SynCom design.

Community-Scale Measurements

The history of microbial genomics reflects progression from gene-based surveys to ecosystem-scale, genome-resolved insights. Early ribosomal RNA (rRNA) gene surveys revealed the existence and dominance of previously unknown microbial lineages, such as SAR11 (i.e., *Candidatus Pelagibacter*) in ocean ecosystems (Britschgi and Giovannoni 1991). Cultivation of representative strains allowed for whole genome sequencing and the discovery of fundamental properties of microbial physiology, for example the streamlining of genomes found in organisms adapted to oligotrophic environments (Connon and Giovannoni 2002). Advances in metagenomics, metatranscriptomics, proteomics, and

metabolomics continue to extend the reach of microbial ecology beyond cultivable organisms. Progress has revealed fascinating features of natural microbial populations, such as previously hidden metabolic capabilities like proteorhodopsin-mediated phototrophy, the ubiquity of auxiliary metabolic genes in viruses, and novel carbon-fixation pathways (Béjà et al. 2000). Single-cell genomics and genome-resolved metagenomics have expanded understanding of genome evolution, community interactions, and niche differentiation in complex microbial assemblages. Together, these developments have transformed microbial genomics into a powerful framework for linking genome content to ecosystem processes.

Current state-of-the-art approaches used to measure microbial communities rely on a sophisticated arsenal of techniques. High-throughput sequencing remains central to these efforts. Targeted rRNA gene and internal transcribed spacer (ITS) sequencing provide rapid insights into community composition (Quast et al. 2012), while untargeted metagenomic sequencing approaches enable more comprehensive characterization and improved taxonomic resolution and functional gene identification (Blanco-Míguez et al. 2023). These DNA-based strategies are complemented by multiomics approaches (e.g., metatranscriptomics, metaproteomics, and metabolomics), which collectively allow researchers to move beyond microbial identification to understanding function. These advances offer unprecedented opportunities to link genomic potential with activity across spatiotemporal scales; they are poised to accelerate discovery and enable deeper insights into microbial adaptation, function, and interactions in natural environments (Srivastava et al. 2024).

2.2 Knowledge Gap

Microbial community design is both a significant opportunity and a challenge for advancing bioproduction, waste valorization, and environmental applications. While SynComs offer powerful experimental platforms, current approaches for microbiome engineering remain largely trial-and-error due to the immense complexity and limited understanding of natural microbiome dynamics.

Community Composition Controls

Researchers do not understand the controls on community composition enough to create perturbations (e.g., stable additions or removals of even a single organism) that alter community structure in predictable ways for defined durations or functions. Similarly, microbe–microbe interactions and the molecular and abiotic cues that regulate these interactions are not well characterized. This challenge is compounded by limited knowledge of the inherent spatial and temporal variability in environmental communities and their responses to changing niche structures. While context-specific demonstration of community phenotypes is possible, transferability across niches is limited due to an incomplete understanding of how community environmental context (e.g., soil type and abiotic gradients) affects cell–cell interactions. The inability to distinguish between additive versus emergent community properties results in poor predictability of community stability or change in the face of shifting abiotic conditions, the influence of predation, phage infection, gene transfer, evolution, or environmental stress. Predictability gaps are exacerbated by measurement limitations (e.g., the inability to monitor basic properties like cellular carbon and electron flow) and by incomplete knowledge of genotype-to-phenotype relationships. Substantial portions of microbial and phage genomes and phenomes remain functionally uncharacterized and are often referred to as microbial dark matter.

2.3 Technical Limitations

Consensus on Model Communities for Parallel Studies

Experiment standardization is crucial for advancing microbiome research and its diverse applications, but there is no community consensus on how to achieve standardization. By establishing standardized microbial communities with defined compositions that include different microbes (e.g., bacteria, fungi, and

archaea), researchers could improve the reproducibility of experiments.

The inherent complexity and spatiotemporal variability of the natural microbial communities that SynComs aim to mimic pose significant challenges to the type of high-throughput experimentation common in other research fields (e.g., using *Escherichia coli* as a model microbe or *Arabidopsis* as a model plant). Yet, standardized and rich data are precisely what is needed to enable promising artificial intelligence (AI) and machine learning (ML) techniques to yield advances in predictive microbiome engineering. Standardized communities could be leveraged to understand specific, ecosystem-relevant functions such as nitrate reduction, salt tolerance, metal accumulation, primary production, and organic matter degradation via experiments that seek modular introduction or deletion of specific taxa or genetically encoded functions. Similarly, additive and nonlinear impacts of mixing different microbes could be studied more rigorously.

While researchers are unlikely to agree on a single dominant model community, there is value in developing a small number of hierarchically organized model systems tailored to specific hosts and environmental contexts. For example, a model community designed for poplar may differ from one optimized for sorghum or for soil settings lacking plants. Rather than endorsing a single community for every use case, the research community would benefit from a framework that helps guide decisions and balance generalizability with ecological relevance. The availability of such model systems at multiple levels of abstraction will enable parallel studies by different laboratories, in different soils, and with varying hosts to provide critical insights into the context-dependent functions of those communities.

High-Throughput Community Characterization

The inherent variability of natural ecosystems, including the composition and chemistry of diverse soil types and fluctuating environmental conditions, complicates predictions of microbial community dynamics and functions. A critical technical limitation is the paucity of high-throughput screening or phenotyping methods and data needed to accurately reflect complex natural environments and evaluate the performance of microbial communities under realistic conditions to drive AI-based approaches. Many current automation and data collection systems are constrained to liquid cultures or plate-based assays, which don't capture the spatial heterogeneity and complexity of natural environments, nor do they sufficiently reflect the network of interspecies and interkingdom interactions. A lack of adequate high-throughput screening tools restricts the ability to assess how microbial communities interact with their surroundings and respond to environmental perturbations, ultimately hampering the development of effective community design strategies.

Standardized and rich data are needed to enable promising artificial intelligence and machine learning techniques to yield advances in predictive microbiome engineering.

Moreover, a lack of standardized protocols, community resources, and metadata standards can lead to inconsistencies in experimental outcomes and reproducibility. Advances within microbiome engineering are also hampered by a scarcity of comprehensive databases that catalog known or presumed metabolic capabilities and ecological roles of various microbial strains and communities. The field is further limited by general uncertainty about the functions and activities of many microbes observed in the environment. While metagenomic sequencing and other omic technologies have advanced, translating vast amounts of sequence, chemical, and biomolecular data into practical applications remains a challenge.

2.4 Research Opportunities

Create Ecologically Relevant Communities

To make microbiome engineering successful in natural environments, microbial community designs

must be ecologically relevant and tailored to function effectively under variable real-world conditions. Engineered communities must be resistant to fluctuations in abiotic factors (e.g., moisture and temperature) and be able to withstand biotic interactions including competition with native microbes and interactions with host plants and soil fauna. These communities should be designed to enhance specific ecosystem processes and functions including nitrogen fixation, host growth, host stress resilience, soil health, and biomass conversion. Community engineering should be informed by both field-based ecological data and laboratory experiments, ensuring communities retain their desired function outside of controlled conditions. Model communities should be widely accessible to the scientific community to support reproducibility and transferability across laboratories, as well as testing and validation across diverse conditions. Public isolate libraries, annotated community datasets, standardized metadata, and open-source protocols for culturing, formulating, and deploying such communities are essential to enable broader collaboration and accelerate innovation. These innovations are vital for supporting feedstock production and agriculture on bioenergy lands.

New community resources and standards specifically tailored to microbiome engineering are urgently needed. There is demand to establish reproducible SynComs with well-defined microorganisms to serve as model systems. Just as *E. coli* and *Arabidopsis* are standard models, sets of reference SynComs for varied environments (e.g., a simplified plant, soil, or gut community) are required to enable consistent experiments to generate model data. A recent community consensus paper recommended defining reference SynComs and creating standard protocols and benchmark data for working with them (Northen et al. 2024). Developing a suite of artificial soil substrates and experimental growth systems is also important for reproducibility and sampling critical properties. Having a standard set of artificial soils would allow researchers to test engineered communities under comparable conditions and to extrapolate their findings across the vast parameter space of soils in the environment.

Establish Cross-Kingdom Multidomain Interactions

SynCom construction must move beyond single-domain models (i.e., those involving organisms from a single kingdom, like bacteria) to address multi-domain interactions, particularly among bacteria, fungi, archaea, and host plants. Many current models, SynComs, and engineering targets are limited to bacterial interactions, yet fungal partners are critical for nutrient uptake, drought resilience, and organic matter stabilization. Fungi account for about three times more biomass than bacteria in some environments like topsoil (He et al. 2020). Establishing model microbiomes, such as standardized communities with defined ecological roles, can support research into cross-domain interactions and serve as baselines for both experimental and computational studies. Advanced modeling techniques, including digital twins (Sizemore et al. 2024), offer a promising way to simulate these complex interactions. These virtual replicas can incorporate multidomain dynamics to predict outcomes across environments and scales, optimizing community performance before scaling to field implementation. ML and AI can further support this effort by integrating large, multimodal datasets and identifying parameters that control the emergent behaviors of microbiomes. By fostering interdisciplinary collaborations among microbiologists, ecologists, engineers, and data scientists, the field can develop holistic, scalable strategies for engineering microbiomes that are resilient, functional, and context-aware.

Achieve Greater Reproducibility

A massive challenge with microbiome engineering studies seeking to affect community function is the inability to introduce new reproducible functions into natural communities through experimental manipulations. Currently, it is not possible to bring a community from the environment into the laboratory and maintain community composition under similar growth conditions. Furthermore, despite advances in building and using SynComs, significant challenges remain for effectively measuring and interpreting data from these simplified communities. Conventional methods often provide only snapshots of community

states rather than continuous monitoring of dynamic processes, limiting the ability to understand the mechanisms underlying temporal community development and responses to perturbations.

The integration of multiomic datasets remains computationally challenging, particularly for communities with many members and when attempting to connect genomic potential to actual phenotypic outcomes. Most measurement techniques lack sufficient spatial resolution to capture the microscale heterogeneity that profoundly influences microbial interactions in natural environments.

To scale community assembly beyond laboratory-scale studies, reproducibility is paramount. Advanced tools such as 3D-printed coculture platforms and microfluidic devices are being developed to study microbial activity in controlled, replicable settings (Birer-Williams et al. 2021; Leung et al. 2012). These platforms enable precise measurement of microbial processes and metabolisms at micro- to macroscales, facilitating mechanistic and process-based research that is reproducible across laboratories and institutions. Reproducibility hinges on developing standardized protocols for community design, testing, and data collection. Coordinated and replicated experiments across laboratories, combined with the sharing of well-characterized SynComs, will help establish common baselines for performance and comparability. This coordination involves creating biobanks containing genome-annotated and phenotypically characterized strains, including both isolates and functional consortia. Such biobanks would enable broader access and foster community-level data sharing and reproducibility.

Emerging technologies can enhance reproducibility by allowing continuous and nondestructive monitoring and quantification of microbial activities over time and across space. These technologies include noninvasive imaging using fluorescence and quantum dots (van't Padje et al. 2021), X-ray computed tomography (Ghosh et al. 2023), isotope tracing, and biosensors. AI and ML tools can help analyze the resulting large datasets and reveal key variables influencing microbiome structure, stability, and functions across different SynCom trials, increasing predictive capacity.

Chapter 3

Using Genetic Programming To Engineer Microbial Communities

Microbes can be genetically programmed using synthetic DNA to exhibit new traits and capabilities, such as processing metabolites for biofuel and chemical synthesis (Smanski et al. 2016), sensing chemicals (Nguyen et al. 2021), removing toxic chemicals (Rylott and Bruce 2020), growing living materials with unique properties (Tang et al. 2020), and processing information as computers do (Bencurova et al. 2023). While tools are emerging to standardize genetic programming of laboratory organisms to address long-standing questions in microbiome engineering, there are significant knowledge gaps and technical limitations to programming undomesticated microbes and microbial communities. This chapter reviews state-of-the-art genetic programming technologies (see Fig. 3.1, p. 20), details existing challenges limiting microbiome engineering, and highlights opportunities to unlock the full potential of synthetic biology.

3.1 State of the Art in Genetic Programming

Genetic programming can be achieved in some microbes and at times within multiple microbes within a community. Current approaches for genetic programming can change how microbes arrange themselves in space, the partnerships they form, and their function in the group. However, predicting the effect these changes will have on individual microbes grown in a flask is difficult. Even harder is anticipating the effect of changes within the context of a microbiome grown under real environmental conditions.

Nucleic Acid Introduction

Engineered DNA can be introduced into microbes using horizontal gene transfer (HGT) methods, such as conjugation (Sørensen et al. 2005; Brophy et al. 2018), transduction (Thomas and Nielsen 2005), and transformation (Lorenz and Wackernagel 1994). In addition, some microbes are thought to exchange synthetic DNA via vesicle exchange, and synthetic vesicles can be leveraged to deliver engineered DNA to some microbes (Metcalf et al. 1997). However, the controls on these mechanisms of DNA uptake remain poorly defined for most microbes and microbial communities.

Two strategies have emerged for mapping organisms in a community that take up mobile DNA via different gene transfer mechanisms: culture-dependent and culture-independent. In culture-dependent HGT, selectable markers or protein reporters are coded into mobile DNA and introduced into the community. Cells taking up the DNA are then enriched to determine which ones participated in gene transfer. This method cannot be applied to unculturable microbes.

In culture-independent HGT, chemical strategies are used to link mobile DNA and host nucleic acids. These strategies include (1) high-throughput chromosome conformation capture (Hi-C; Yaffe and Relman 2019) and (2) emulsion, paired isolation, and concatenation

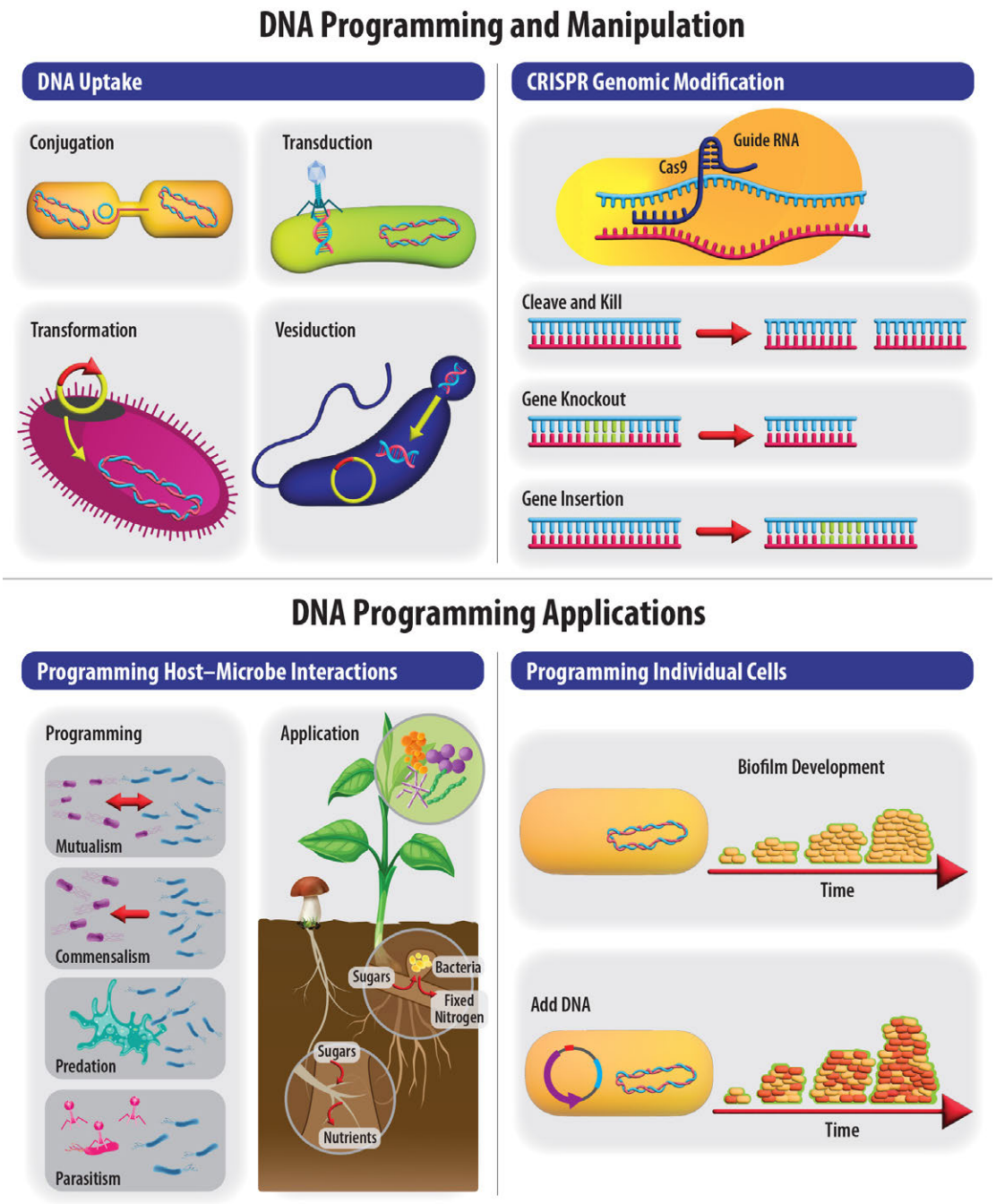


Fig. 3.1. Engineering Microbiomes Using Synthetic Biology. (A) DNA Uptake Approaches. Shown are four methods to introduce engineered DNA into cells within communities: conjugation mediated by cell–cell interactions, phage virus infection through transduction, transformation via plasmids, and delivery of lipid-encapsulated DNA via vesiduction. Efficiencies of each process are not known for most microbes in environmental communities. **(B) Genomic Modification Strategies.** Diverse biomolecules can be coded into engineered DNA, which is then introduced into communities. For example, encoded proteins like Cas9 and their guide RNAs can be used to achieve gene editing in cells to remove microbes from the community, delete one or more genes, or introduce engineered DNA into genomes. **(C) Programming Host–Microbe Interactions.** Engineered DNA delivered to cells in communities can be used to alter (left) microbe–microbe interactions and (right) cell–host interactions. **(D) Programming Dynamic Behaviors.** Behaviors can be programmed into communities using synthetic biology, such as the creation of patterned growth over time. Currently, these dynamic behaviors can only be achieved in the laboratory using a pair of microbes, illustrating the need for improved control theory to enable new environmental applications.

PCR (epicPCR; Spencer et al. 2015). In addition, genetic circuits can be coded into mobile DNA and used to barcode host DNA or RNA. Examples are environmental transformation sequencing (ET-seq; Rubin et al. 2021) and RNA-addressable modification (RAM; Kalvapalle et al. 2025). Culture-independent HGT suffers from poor signal sensitivity.

Precision Genome Editing

Extraordinary advances in genome editing have been enabled by the discovery of CRISPR-Cas systems (Ko et al. 2020; Volk et al. 2022; Jakočiūnas et al. 2016). Early CRISPR-Cas systems achieved targeted DNA insertion in organisms using RNA-guided CRISPR-Cas transposase (Strecker et al. 2019) and DNA-editing-all-in-one RNA-guided CRISPR-Cas transposase (DART; Rubin et al. 2021). These approaches can achieve targeted, marker-free, efficient integration of DNA fragments of >10 kilobases in some microbes. Genome integrations mediated by these systems have been multiplexed to achieve deletions and insertions at multiple genomic loci. Similarly, recombinase-assisted genome engineering can achieve site-specific genome integration of engineered DNA (Santos et al. 2013), such as chassis-independent recombinase-assisted genome engineering (CRAGE) and serine-integrase-assisted genome engineering (SAGE) systems (Wang et al. 2019). Particularly exciting is the potential opportunity to modify microbes *in situ* via integration of engineered DNA in non-model bacteria and the removal of the marker (Elmore et al. 2023). Synergistically with gene editing, several gene regulation approaches have been developed using CRISPR-Cas systems, such as CRISPR interference (CRISPRi; Hren et al. 2025), which can be used to rapidly discover essential genes and fitness-improved knockdowns in high-throughput experimental formats.

Host Genetic Engineering

Hosts can be targeted for genetic engineering to influence the composition and function of their associated microbiomes (Haldar and Sengupta 2015). Synthetic microbe–plant communication pathways are providing a foundation for programmable interkingdom interactions (Boo et al. 2024). Host-to-microbe signaling has been engineered to allow plants to influence microbial activity by releasing signals that activate microbial gene expression, suppress undesirable activities, or recruit beneficial symbionts that improve plant response to shifts in abiotic characteristics in the environment (Griffin et al. 2024). In addition, plant receptors have been discovered that enhance interactions with beneficial microorganisms, such as mycorrhizal fungi, which promote mutualism while limiting pathogen colonization (Bashyal et al. 2025). Biosensors have been engineered in plants for microbe-to-plant communication where the plant responds dynamically to microbial inputs (Liu et al. 2024). These successes suggest that similar approaches could be applied to undomesticated bioenergy crops to open new avenues for optimizing microbial contributions to productivity in challenging environments (Farrar et al. 2014).

Microbes as Sensors

A surge in protein sensor understanding has expanded the diversity of genetically encoded parts that can be leveraged to engineer microbial biosensors to report environmental conditions (Mehrotra 2016; Rogers 2006; Su et al. 2011; Del Valle et al. 2021). Protein sensors have been described for diverse primary and secondary metabolites, chemical intermediates in biogeochemical cycles, metal ions, and other common environmental parameters, such as osmolytes, cell–cell signals, light, pH, and temperature (Del Valle et al. 2021). Recent efforts have extended sensing from small molecules to macromolecules like DNA (Cheng et al. 2001). Biosensors most frequently produce visual signals as outputs (e.g., fluorescent proteins). Because these signals present challenges in opaque environmental materials like soils, alternatives have been developed including volatile metabolites, gas vesicles, and protein wires. Among the newest offerings are hyperspectral imaging reporters for nondisruptive sensing of microbial reporters in fields (Chemla et al. 2025). While many biosensors provide real-time information on chemical detection, others are capable of recording and storing information about transient environmental chemicals (Essington et al. 2024), which can be read weeks after sensing. Biosensors are moving towards more *in vivo* applications (Slomovic et al. 2015), but significant

work is needed to translate these into tools for engineering microbiomes.

3.2 Knowledge Gaps

Control Theory for Assembly and Function

Innovations have yielded strategies to use genetic engineering to achieve surface functionalization of microbes by displaying proteins or DNA to direct spatial assembly (Custódio and Mano 2016). In addition, a key conceptual advancement emerging from microbiome research is the focus on engineering functional modules rather than individual strains, recognizing that community stability may depend more upon metabolic network structure than specific taxonomic composition (Louca et al. 2016). This latter approach acknowledges the importance of functional redundancy and resilience, while aiming for predictable community behaviors and outputs. Emerging frameworks including optogenetics, metabolic division of labor, and modular design, are just beginning to enable predictable engineering of microbial consortia by coordinating interspecies interactions, resource use, and functional outputs through rational design principles (Hoffman et al. 2022). Control theory and emerging technologies for controlling spatial organization, such as 3D bioprinting of microbial communities in hydrogels, are enabling researchers to investigate how physical proximity and specific cell–cell signals influence metabolic exchanges and community dynamics (Kumar and Foster 2022; Chen et al. 2015). While technical approaches are emerging for achieving spatiotemporal control over microbe–microbe associations, they are in the very early stages of development and only involve a handful of proof-of-concept studies.

Ecology of Mobile Genetic Elements

The ecology of mobile genetic elements remains poorly defined. The mechanisms by which DNA plasmids are exchanged, environmental DNA is taken up, and bacterial viruses infect cells have been extensively studied (Sørensen et al. 2005; Thomas and Nielsen 2005; Lorenz and Wackernagel 1994). In addition, databases have been created that catalog mobile DNA found in the environment (Dunivin et al. 2019; Wang et al. 2024); some include metadata about host taxonomy and geographic distribution (Camargo et al. 2024). While these resources are available to support microbiome engineering, recent computational analyses of mobile genetic elements suggest that most of Earth's mobile DNA is absent from these databases and remains unexplored (Graham et al. 2024).

The host range for even the best characterized mobile DNA remains challenging to predict in microbial communities, yet it can be highly specific. It is still difficult to anticipate which members of a microbiome can participate in gene transfer, how the rates of exchange of different DNA sequences vary across community members, and how variation in host-defense systems within community members affects the stability and persistence of mobile DNA following gene transfer.

Knowledge is so limited that it is impossible to predict how changes and mutations within a single plasmid component (e.g., changes in the origin of replication) affect host range. Also unclear is how DNA acquired through gene transfer modulates cell fitness and how those effects vary across microbes and environments. Genes introduced into microbes via gene transfer may at times enable microbes to access scarce, previously inaccessible nutrients, which allows them or their hosts to expand into new niches. Cross-kingdom gene transfer has the potential to improve microbial host productivity similar to that observed with organelle-to-nucleus DNA transfer in plants (Feyissa et al. 2024). Alternatively, DNA acquired via gene transfer could create fitness burdens that lead to the local extinction of microbes from the population or mutation within engineered DNA. Thus, a strong understanding of gene transfer and the DNA-encoded traits exchanged is especially critical when considering how best to engineer microbiomes using synthetic biology for field-scale applications.

Abiotic and Biotic Effects on Gene Transfer

When communities are situated in natural settings, such as soil, gene transfer may impact microbiome functions. A better understanding of how abiotic chemicals and physical materials in the environment affect gene transfer is needed to design effective

strategies for gene transfer, DNA circuit persistence, and biocontainment (Lee et al. 2018). Particularly for soil microbiomes, it remains unclear how gene transfer is affected by variation in physical (e.g., texture) or chemical (e.g., nutrients) properties. In addition to understanding the host range of different mobile DNA, gene transfer dynamics need to be understood (e.g., the ways that gene transfer rates vary with different types of mobile DNA and ecological stress). Critical to understanding gene transfer *in situ* are measurements across soil parameter space, which is vast compared to human gut microbiomes where much gene transfer research is performed. Soil parameter space includes planktonic, biofilm-associated, and dormant cells that experience dramatic environmental changes on a range of time scales.

Information on mechanisms underlying the relationship between gene transfer, community resilience, functional redundancy, and multilevel selection remains scarce (van Vliet and Doebeli 2019). Such knowledge is required to inform strategies for engineering communities that adapt to changing conditions or outcompete indigenous microbes in field conditions.

Fungal Community Engineering

Fungal community engineering is hindered by knowledge gaps in fungal gene transfer, ecology, and evolution. Compared to bacteria, gene transfer in fungi remains poorly understood *in situ* (Liu et al. 2024). There are examples of gene transfer in budding yeast (Kawai et al. 2010; Moriguchi et al. 2013), filamentous fungi (de Groot et al. 1998), and arbuscular mycorrhizal fungi (Helber and Requena 2007). However, researchers cannot directly engineer fungal communities even though these microbes account for half of the biomass in many communities (e.g., soils). RNA and DNA mycoviruses are widespread across fungal taxa and can modulate host phenotypes, yet their ecology, transmission mechanisms, and impact on community function are unresolved (Rosewich and Kistler 2000; Gabriel et al. 2015). Additionally, understanding of fungal–fungal, fungal–microbe, and fungal–plant interactions is far less developed than that of bacterial systems, particularly regarding how gene flow, competition, cooperation, and selection shape community function over time. Functional redundancy, multilevel selection, and ecological feedbacks within fungal communities are essential for predicting how these microbes respond to engineering. Even for relatively well-studied fungi, regulatory networks and ecological functions are insufficiently mapped to support rational design (Idnurm et al. 2017). As a result, efforts to engineer fungal communities in natural environments are hindered by a lack of foundational ecological and evolutionary insight.

Genetic Parts and Regulatory Elements Identification

Identification of genetic parts and regulatory elements (e.g., genes, promoters, terminators, and autonomously replicating sequences) is a critical knowledge gap for engineering non-model microbes found in microbiomes. Construction of novel pathways in any microbe requires knowledge about promoter activities needed to drive gene expression; most genetic circuits require multiple well-characterized promoters. As promoters can be constitutive or conditional, there is a need to understand how environmental conditions affect promoter activities. If proteins are also expressed, information is required about translation initiation sequences and the coding DNA sequences (CDS) being expressed.

Some methods exist to predict promoter sequences and operons upstream of prokaryotic genes (Cassiano and Silva-Rocha 2020), but they cannot be robustly applied to eukaryotes like fungi and plants because identifying genetic parts is challenging (Ledesma-Dominguez et al. 2024). Methods have been reported to biochemically identify promoters and associated transcription factors [e.g., DNA affinity purification sequencing (DAP-seq) and chromatin immunoprecipitation sequencing (ChIP-seq)] in non-model microbes (Baumgart et al. 2021). However, these methods are not suitable for identifying promoters of different transcriptional strength, particularly those that are catabolite responsible (e.g., inducible). In addition, these methods do not provide insight into

The ability to predict the activities of engineered DNA regulatory elements (e.g., promoters, translation-initiation sites, and terminators) across non-model microbes is a critical need for genetically programming environmental microbiomes.

the activity of completely synthetic regulatory elements in microbes.

The ability to predict the activities of engineered DNA regulatory elements (e.g., promoters, translation-initiation sites, and terminators) across non-model microbes is a critical need for genetically programming environmental microbiomes. The knowledge gap limits the ability to reliably express genome-editing tools in undomesticated microbes and to create synthetic genetic circuits with dynamic functions, such as the ability to synthesize chemicals that control microbe–microbe interactions. For example, sigma factor sequence–structure–function relationships are not understood sufficiently in microbiomes to anticipate how efficiently a given natural or synthetic promoter might be transcribed within different members of a community whose sigma factors have different protein sequences. Similarly, researchers cannot anticipate the length of transcripts due to the inability to predict the efficiencies of transcriptional terminators. Neither can scientists anticipate protein levels, which are determined by translation-initiation and protein-degradation rates. Furthermore, researchers cannot predict substrate specificities of enzymes or their dynamic functions and interactions with other macromolecules, even though they can annotate the CDS (Bharti and Grimm 2021). Advances in AI have transformed the ability to predict protein structure using AlphaFold (Jumper et al. 2021), but these models remain limited in their abilities to anticipate dynamic functions of the vast uncharacterized genes in environmental microbe and phage genomes.

3.3 Technical Limitations

Genome-Editing Tools and Genetic Parts for Non-Model Organisms

A large fraction of environmental microbes and communities remains recalcitrant to genomic modification using emerging genome-editing tools (Fatma et al. 2020). Researchers need simple technologies for delivering DNA, RNA, and proteins (e.g., CRISPR-Cas tools) into undomesticated organisms found in communities such as bacteria, fungi, protozoa, and their viruses (Riley and Guss 2021). For some organisms, technologies for cargo delivery are needed to overcome internal defense mechanisms including those that inhibit DNA uptake (Riley and Guss 2021). In addition, there is a need to accurately predict how genetic regulatory elements function that control the synthesis (e.g., promoter, terminator, and translation-initiation site activities) and degradation of cargo delivered to cells. These challenges exist due to the lack of reliable technologies for overcoming transport limitations associated with microbes having complex membrane systems and glycan-rich cell walls. Also needed are simple technologies to reliably deliver DNA programs to specific cellular compartments (e.g., the nucleus or mitochondria) in the case of eukaryotes like fungi.

There is demand for a genetic parts toolbox for undomesticated microbes, such as promoters, ribosomal binding sites, terminators, and genetic-logic gates with predictable functions. Researchers lack models for predicting the efficacy of promoters (both inducible and constitutive) across most microbes. These tools are required to achieve precision control over protein expression, especially proteins that perform gene editing, and to allow pathways of genes to be expressed at desired times and locations at user-defined levels. Such dynamic gene expression is critical to programming predictable time-dependent behaviors in microbiomes.

Within eukaryotes, safe sites within genomes need to be identified to incorporate foreign DNA (Peterson et al. 2021) and to understand DNA repair mechanisms (e.g., homologous recombination and non-homologous end joining) that could be exploited

to improve editing tools (Wilken et al. 2019). Such knowledge will aid in understanding not only the function of DNA upon genome modification but also the potential for spread via HGT. Also required are genetic circuits that control cell–cell interactions, such as competition, syntrophy, signaling, quorum sensing, and energy flow. Other needs include circuit architectures capable of remaining functional over intended timescales despite rapid microbial evolution, which may lead to loss or gain of new traits.

In fungi, gene transfer using both transformation and bacteria-to-yeast conjugation can be reliably achieved in laboratory strains of the budding yeast *Saccharomyces cerevisiae* (Kawai et al. 2010; Moriguchi et al. 2013). With filamentous and arbuscular mycorrhizal fungi, *Agrobacterium tumefaciens* is increasingly used to achieve DNA uptake (de Groot et al. 1998; Helber and Requena 2007), mirroring approaches leveraging this bacterium for genetic modification in plants (Chilton et al. 1977).

In contrast, the engineering of intact, diverse fungal communities remains technically challenging due to key limitations in current tools. While genome-editing technologies such as CRISPR-Cas9 have expanded the capabilities for manipulating individual fungal species, their application across the broad phylogenetic and ecological diversity of fungi observed in environmental communities is limited by transformation barriers, inefficient delivery systems, and a lack of species-specific genetic parts (Malci et al. 2022).

Many fungi, particularly non-model organisms involved in complex environmental or symbiotic interactions, are difficult to cultivate or genetically manipulate using standard laboratory techniques (Liu et al. 2024). These challenges are compounded by limited genomic and transcriptomic resources for many ecologically important taxa. Though current tools are effective for engineering some individual strains, scaling these approaches to design or manipulate diverse, naturally occurring fungal communities remains a major challenge in synthetic ecology and fungal biotechnology.

Monitoring Sensors

To monitor microbiome functions, sensors are needed to capture information about the ecological niche spaces occupied by communities, microbiome functions, and chemicals that serve as proxies for community behaviors. Genomic sequencing has identified more than 10 million unique sensor proteins in nature (Park et al. 2023), yet only a handful have been used to build biosensors. Building and refining sensors for new molecular targets remain arduous. The research community does not have simple platforms to support high-throughput screening of uncharacterized sensors, even though strategies for grafting sensor domains have been reported that accelerate characterization (Schmidl et al. 2019).

Workflows are needed that couple newly characterized sensors to outputs compatible with *in situ* monitoring. There is also demand to diversify outputs enabling sensor multiplexing in microbiomes, such as coupling diverse sensing input to the production of different metabolites that could be detected in parallel using hyperspectral imaging (Chemla et al. 2025). Such sensors could be used for engineered microbiome characterization at different scales by allowing researchers to monitor the diverse environmental parameters that microbes are sensing in real time to regulate their behaviors. Such information, namely the ways that microbes perceive their environment, will be important for understanding how to engineer reliable microbiome functions.

High-Throughput Detection and Editing Technologies

High-throughput technologies are needed to detect DNA uptake and edit genomes in microbiomes. Monitoring the efficiency by which nucleic acid payloads are delivered to microbiomes requires sensitive, high-throughput technologies to enable rapid detection of nucleic acid uptake, such as phage transduction host range and efficiency. Existing technologies enable detection of mobile DNA uptake across microbiomes, such as Hi-C (Yaffe and Relman 2019), ET-seq (Rubin et al. 2021), and RAM (Kalvapalle et al. 2025). However, these approaches have limited sensitivity for reporting

on the host range of DNA uptake in microbial communities. Technologies that function robustly across archaea, bacteria, and fungi are needed to provide information on both host range and transfer efficiency.

There is also demand to simplify and increase the throughput of genomic-editing approaches. When applied in communities, the introduction of new functions depends on how genomic changes are directed, whether off-target impacts occur, and how multiple edits across single and multiple taxa integrate. Microfluidic approaches have become useful to analyze the success of genetic edits in single microbes (Gach et al. 2016) and the interaction of different microbes (edited or not) within a single droplet (Hsu et al. 2019). Microbial communities can be rapidly prototyped, assembled, and even cultivated in model environments via droplet microfluidics, which enables conditions that promote assembly and stability of synthetic communities (SynComs) to be screened over thousands of conditions (Leung et al. 2012; Yu 2022; Hsu et al. 2019). However, these devices are often limited to small monodisperse populations of microbes (e.g., simple bacteria and yeasts). Generally, they cannot scale to genetically engineered communities with more complex spatial interactions, especially where biofilms are formed or filamentous growth occurs.

3.4 Research Opportunities

DNA and RNA Mobilome Atlas

Construction of a mobilome atlas would inform genetic programming of environmental microbes. Some existing databases for plasmids and bacteriophages catalog mobile DNA in the environment (Wang et al. 2024; Camargo et al. 2024; Graham et al. 2024). However, these resources are not designed to guide the introduction of engineered DNA into microbes found in any environmental sample. Multiple tools are now available to map gene transfer host range in communities (Rubin et al. 2021; Kalvapalle et al. 2025). These could be refined and used to catalog how mobile DNA (e.g., bacteriophage and plasmids) with different characteristics affects the host range and transfer efficiencies in microbiomes having varying compositions.

The DNA-encoded characteristics of individual microbes in a community are expected to affect their participation in gene transfer. Examples include the structures of their surface receptors and their host-defense system types (Koonin et al. 2001; Makarova et al. 2013). As such, parallel systems biology analyses of microbiomes targeted for gene transfer studies offer an opportunity to establish the mechanistic controls on gene transfer host range and efficiency. In addition, studies that examine gene transfer across different environmental conditions, such as those varying in chemical and physical properties, represent an opportunity to understand how gene transfer is affected by abiotic controls and transport processes in the environment. Such models would be useful for anticipating gene transfer in native communities, guiding microbiome engineering *in situ*, and establishing biocontainment systems with the necessary fidelity.

Cross-Kingdom Genetic Engineering Tools

The microbial host range of genome-editing tools (e.g., CRISPR-Cas9) and RNA-guided synthetic biology tools (e.g., CRISPRi) is limited. New tools are needed that function across a wider range of organisms (e.g., archaea, bacteria, and eukaryotes) and viruses found in different ecological niches. Also needed are tools that selectively edit or regulate genes in a subset of the microbes within a community and within specific genome regions while limiting off-target effects. Other opportunities include tuning the expression of native transcripts using CRISPRi methods (Larson et al. 2013) and using these tools in high-throughput formats to elucidate the roles that different genes play in controlling the fitness of individual microbes in a community across different environmental conditions. These types of precision control could be achieved by coding tools into mobile DNA with constrained host ranges or by controlling their synthesis using conditional expression. This opportunity illustrates the importance of developing atlases for mobile DNA and regulatory elements.

Additional needs include methods to enrich and retain genetic edits within community members. These capabilities are critical for studies that examine

the spatiotemporal impacts of such modifications in communities. Strategies that perform genetic selection via antibiotics cannot be used in microbiomes, as they alter community structures. New selection strategies that endow edited microbes with the ability to thrive on specific crucial substrates are needed, enabling the community to thrive only if the substrate and genome edits are retained. Beyond targeting microbiomes for genome editing, another opportunity lies in performing cis-genic engineering of plant hosts and microbes. Such approaches offer regulatory flexibility and functional precision, allowing for more sophisticated multidomain designs.

Genetic Parts Foundry

To accelerate the creation of DNA circuits that endow microbes with new functions in communities, atlases of regulatory parts (e.g., promoters, terminators, and ribosomal binding sites) need to be created, enabling predictive genetic engineering across microbiomes. Models that predict sequence–function relationships in laboratory microbes transformed synthetic biology (Salis et al. 2009; LaFleur et al. 2022; Chen et al. 2013). To extend these tools, both natural and synthetic regulatory elements must be characterized across a greater diversity of archaea, bacteria, and fungi. For example, by evaluating the activities of synthetic promoters across diverse microbes in parallel, researchers could establish how these biological parts vary in function as they are applied in different microbial chassis with distinct sigma factors.

Pairing promoter strength measurements across diverse microbes with systems biology data on the evolution of sigma factors (Feklístov et al. 2014) would accelerate efforts to understand the mechanisms underlying promoter activity variation. Similar approaches could be applied to terminator- and translation-initiation sequences. The safe use of synthetic microbes in diverse environments requires generating robust tools that constrain where they grow and persist. Modular safety features, which could be used in both domesticated and undomesticated microbes, would enable precision control over when and where engineered microbes persist.

Microbial Sensor Atlas and Sensor Foundry

Genetically encoded sensors, or biosensors, can provide spatiotemporal information on engineered microbiome functions *in situ*. While many microbes devote a large percentage of their genes to sensing, scientists do not understand what most sensors detect in microbiomes. Thus, there is an opportunity to build an atlas of biosensors capable of reporting on all possible environmental parameters. To realize this, researchers need to understand the molecules sensed by natural sensors and the DNA binding specificity of those systems. This parallel challenge limits the speed at which researchers can map how natural sensor systems convert chemical information in the environment into biochemical information within cells.

This complex challenge can be overcome with two-component systems (TCS), which are signal transduction pathways that enable bacteria to sense and respond to diverse physical, chemical, and biological stimuli outside and inside the cell (Lazar and Tabor 2021). Researchers can use TCS to graft natural sensor domains onto characterized DNA-binding domains (Schmidl et al. 2019). This grafting approach could be scaled up to create libraries of natural sensor proteins derived from environmental microbiomes. These could be screened against chemical libraries to map specificities. Protein design could be used to tune, extend, and understand sequence controls on sensing and DNA molecular recognition.

Fundamental insight into the sequence–structure–function relationships in sensor proteins would yield tools for monitoring microbiome functions. This knowledge could ultimately enable the prediction of the natural sensor functions in microbiomes, which regulate microbial responses to changing conditions. Another opportunity exists to develop output modules for sensors that enable multiplexed and distributed sensing of diverse environmental conditions.

Finally, there is demand to develop memory-based sensors that record transient environmental information during microbiome incubations, which can be read out at the end of an incubation. Memory-based systems could be developed to store large amounts of information within individual microbes in a microbiome, much like the hard drives on computers.

Chapter 4

Scaling Microbiome Engineering to Complex and Dynamic Environments

Advancing microbiome engineering will require a suite of cutting-edge platforms that enable comprehensive testing, monitoring, and scaling across diverse environmental conditions. This chapter (1) outlines the current state of the art in testing and characterization of engineered microbial communities, which range from highly controlled microfluidic devices and mesocosms to complex, heterogeneous field environments, and (2) considers performance across spatial and temporal scales.

A grand challenge for microbiome engineering lies in translating findings from controlled laboratory systems to dynamic real-world settings where environmental and spatial heterogeneity and temporal variability often limit predictability and performance. This requires integrative, cross-disciplinary strategies and the development of novel infrastructure including advanced sensor technologies, distributed field networks, and collaborative research frameworks. Together, these innovations are essential to realize scalable microbiome engineering and support the growth of a robust bioeconomy.

4.1 State of the Art in Testing and Characterization

The characterization of engineered microbiomes generates spatiotemporal data about the composition, function, and stability of communities across different environmental conditions. By learning from the effects of different environments on community properties, rules could be developed to improve translatability and control. Current approaches do not capture the necessary data across different habitats to enable this predictive control over microbiome function.

Testing Habitats

Diverse platforms enable the testing of engineered microbiomes (see Fig. 4.1, p. 30) across scales and offer differing degrees of control and complexity to study cell–cell, cell–host, and cell–environment interactions.

Microscale Habitats

At the microscale, microfluidic devices enable continuous high-resolution characterization of microbial function, chemical gradients, and interactions with roots or other microbes under tightly controlled conditions (Aufrecht et al. 2022; Gupta et al. 2024). Though simplified compared to natural environments, these systems can mimic soil pore structures and root exudation patterns, offering mechanistic insights into microbial colonization and competition in confined spaces. Advanced microfluidic approaches, such as droplet-based platforms and chip-based technologies (Kehe et al. 2019), allow for high-throughput exploration of microbial community composition by testing thousands of combinations in parallel (Kehe et al. 2019). A complementary approach is artificial intelligence (AI)–guided robotic culturomic platforms, which enable large-scale isolation and analysis

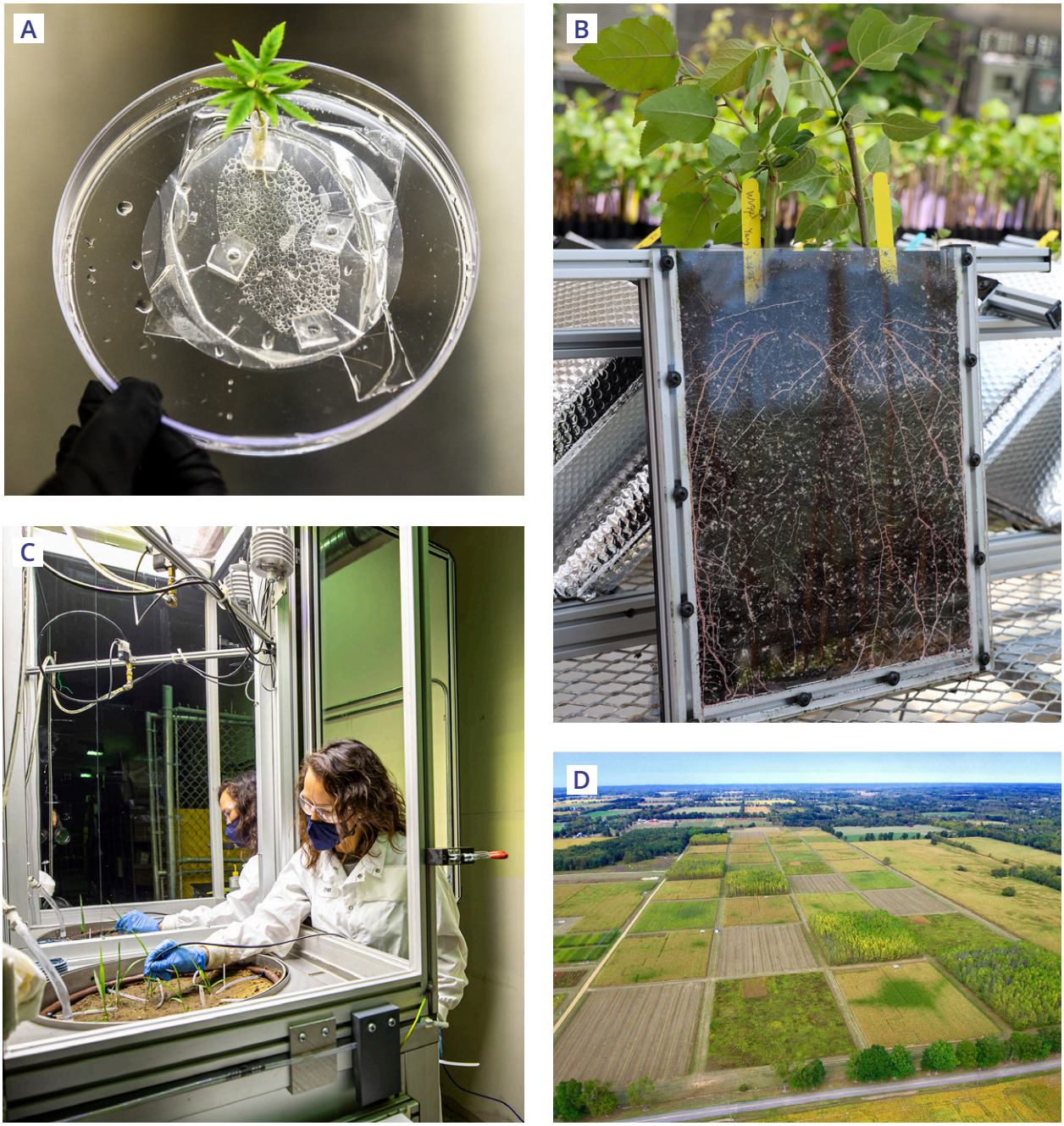


Fig. 4.1. Platforms for Testing Microbiome Engineering Span Micro- to Macroscales. (A) Microfluidic devices such as the rhizosphere-on-a-chip platform (RhizoChip) provide a microscale soil environment that retains the physical structure of natural soils and can be used to map the molecular environment of roots. [Courtesy Environmental Molecular Sciences Laboratory] **(B)** A Rhizobox soil observation system enables visualization of root growth and interactions with rhizosphere soil. [Courtesy Oak Ridge National Laboratory] **(C)** A scientist inspects plants growing inside an EcoPOD. EcoPODs are enclosed environments enabling intensive monitoring and manipulation of replicated plant−soil−microbe−atmosphere interactions over the complete plant life cycle. [Courtesy Lawrence Berkeley National Laboratory] **(D)** An aerial view of different treatments in a field research site at the Great Lakes Bioenergy Research Center and the Kellogg Biological Station's Long Term Ecological Research Cellulosic Biofuels Research Experiment. This long-term study examines the impact of different cropping systems on ecosystem services related to biofuel crops. [Courtesy Kevin Kahmark, Michigan State University]

of microbial colonies and accelerate the discovery and characterization of previously unculturable organisms (Kehe et al. 2019).

Both dynamic measurement techniques and spatial tools are needed to understand the collective function of microbial communities. Dynamic measurement techniques (e.g., stable isotope probing and volatilomics) can reveal and quantify metabolic activities (Brown et al. 2021). Spatial tools, such as imaging mass spectrometry, can uncover the organization of metabolic processes in structured microbial consortia (Kertesz and Cahill 2021). Fabricated platforms can be used to scale between microcosms and mesocosms and facilitate multiomic and phenotypic measurements and reproducibility (Zwart et al. 2025).

Mesoscale Habitats

At the mesoscale, controlled environmental chambers aim to bridge the gap between laboratory and field conditions (Yee et al. 2021). Such systems allow for precise manipulation of soil composition, moisture levels, temperature, and microbial communities, enabling experiments under seminatural, yet controlled conditions. Chambers also provide a valuable platform for studying plant–microbe–environment interactions over longer timescales while maintaining reproducibility (Singer et al. 2021). Current chamber systems support research on microbiome function and plant performance, facilitating controlled testing before field deployment. Although useful for targeted hypothesis testing, chamber systems lack the full complexity and scale of natural field environments, limiting the realism of plant–microbe–environment interactions that can be measured. Specific constraints of chamber systems include simplified biotic interactions, limited spatial scales, high operational costs, limited ecological relevance, and low experimental throughput.

Greenhouses

Scaling up in space and complexity, controlled greenhouses can regulate abiotic conditions such as temperature, humidity, and light cycles, enabling microbiome engineering experiments in whole-plant and plant–soil systems. Greenhouses can further integrate automated plant phenotyping, which enables high-throughput monitoring of key plant traits like growth, root architecture, stress responses, and nutrient acquisition over time (Feyissa et al. 2024; NAPPN 2025; ORNL 2025). Semiautonomous systems provide a scalable environment for evaluating microbiome-driven plant performance under controlled conditions. While such phenotyping systems are transformative for linking genotype to phenotype, access remains limited with only a small number of high-throughput facilities available to the broader research community. Because these platforms primarily rely on imaging technologies to infer plant traits, advancing methods to more accurately translate image data into physiological information is essential. Current systems also need greater integration of real-time and nondestructive microbial measurement capabilities (e.g., volatilomics or metabolomics).

Field Sites

Long-term field sites are critical for testing engineered microbiomes in natural environments where environmental conditions cannot be controlled but perhaps experimentally manipulated (Hanson and Walker 2019; Chuckran et al. 2024; Taylor et al. 2024). Such sites offer the opportunity to evaluate microbiome stability, ecological interactions, and longer-term impacts on microbiome and host plant performance across diverse, dynamically changing environmental conditions. Although field studies inherently introduce variability, they remain critical for assessing the real-world feasibility of microbiome engineering approaches.

Sensors

A wide range of sensors can be used to monitor microbiome functions (see Table 4.1, p. 32). Omics-enabled systems biology methods are widely used to measure who is present in a microbiome (Daniel 2005), what macromolecules are produced (Bashiardes et al. 2016), and what ensemble of small molecules are synthesized (Bundy et al. 2008). These techniques usually require destructive sampling, though some omics approaches are compatible with dynamic measurements. For example, volatilomics reports on the ensemble of volatile chemicals released into the headspace of microbiome atmospheres (Meredith and Tfaily 2022). Similarly, hyperspectral imaging represents a noninvasive technique that captures spectral

Table 4.1 Biosensors for Nutrient Cycling, Metals, Cell–Cell Signaling, and Environmental Parameters

	Analyte	Sensor Name	Mechanism	Mode of Prior Use	Reference
Carbon	CO	CooA	TR	*Rhodospirillum rubrum*	Roberts et al. 2001
	CO_2	RegAB*	TCS	*Rhodobacter capsulatus*	Ganesh et al. 2020
	CH_3OH	FlhS/EnvZ	TCS	*Escherichia coli*	Selvamani et al. 2017
	CH_20	FrmR	TR	*E. coli*	Rohlhill et al. 2017
	CH_3X	Ada	TR	*E. coli*	Moser et al. 2013
Nitrogen	NH_4^+	Km-Amt^	TCS	*Kuenenia stuttgartiensis*	Pflüger et al. 2018
	NO_3^-	NarXL	TCS	*E. coli*	DeAngelis et al. 2005
	NO_2^-	NarXL	TCS	*E. coli*	DeAngelis et al. 2005
	NO	NorR	TR	*E. coli*	Archer et al. 2012
	N_2	NtrBC*	TCS	*E. coli*	Ganesh et al. 2020
	Amino acids	GFP	Translation	Cell-free	Jang et al. 2017
Phosphorous	PO_4^{3-}	PstS	PS	Purified protein	Solscheid et al. 2015
	ATP	Plg2	PS	Purified protein	Branchini et al. 2015
Metal Ions	Mn(II)	MntR Mn	TR	*Bacillus subtilis*	Huang et al. 2017
	Zn(II)	CzcSR	TCS	*Pseudomonas putida*	Liu et al. 2012
	Ni(II)	RcnR	TR	*E. coli*	Cayron et al. 2017
	Cu(II)	CueR	TR	*P. putida*	Li et al. 2014
	Fe(III)	BasSR	TCS	*E. coli*	Hagiwara et al. 2004
	Fe(II)	BqsSR	TCS	*Pseudomonas aeruginosa*	Kreamer et al. 2012, 2015
Gases	O_2	ANR	TR	*Pseudomonas fluorescens*	Højberg et al. 1999
	H_2	HupUV	TCS	*R. capsulatus*	Elsen et al. 2003
Osmolytes	Osmotic stress	*proU* promoter	TR	*P. putida, Pantoea agglomerans*	Herron et al. 2010
	Glycine betaine	CFP, GBP, YFP	PS	*E. coli*	Ahmad et al. 2016
Signals	Ethylene	SynEtr1	TCS	*Synechocystis sp.*	Lacey and Binder 2016
	Salicylic acid	NahR	TR	*E. coli*	Meyer et al. 2019
	GBAP	FscAC	TCS	*Enterococcus faecalis*	Sturme et al. 2002
pH	5 to 7	RstA	TCS	*E. coli*	Hoynes-O'Connor et al. 2017
	5.5 to 7	VirA-ChvE	TCS	*Agrobacterium tumefaciens*	Gao and Lynn 2005
	6 to 8	SO_4387, SO_4388	TCS	*E. coli*	Schmidl et al. 2019
Temperature	27 to 37°C	CspA	TR	*E. coli*	Hoynes-O'Connor et al. 2017
	32 to 46°C	TlpA, Tcl	TR	*E. coli*	Piraner et al. 2017

Recent advances in synthetic biology have expanded the range of inputs that can be sensed. The modules listed include both natural and engineered sensors, most of which use a single transcriptional regulator (TR), two-component system (TCS), or a protein switch (PS) to convert environmental information to biochemical information. For each sensor, the table notes whether the mode of prior use has been within a cell-free system, or isolated biomolecule. **Asterisks (*)** indicate sensors that respond to inputs indirectly; **carets (^)** denote that the sensing mechanism is not fully characterized. [Adapted under a Creative Commons Attribution License (CC BY) from Del Valle, I., et al. 2021. "Translating New Synthetic Biology Advances for Biosensing into the Earth and Environmental Sciences," *Frontiers in Microbiology* **11**, 618373.]

data and serves as a fingerprint for specific chemicals *in situ* (Lu et al. 2020).

The integration of biologically produced optical or volatile reporters—particularly those responsive to metabolite concentrations or stress signals—offers critical tools for tracking dynamic changes within engineered systems. Nonbiological sensors for distributed measurements, such as soil sensors and plant-wearable devices have also been described (Yin et al. 2021). Typically, these are used to report on temperature, moisture, pH, nutrients, fertilizer, and gases (e.g., carbon dioxide, methane, and nitrous oxide). However, regardless of measurement type, current sensor technology can monitor only a small fraction of the chemicals processed by microbiomes, obfuscating most real-time microbial metabolic interactions underlying the collective and emergent properties of microbiomes. This limitation is particularly problematic for engineering efforts that require dynamic regulation of microbial activity, an essential function that is difficult to implement and tune without real-time, spatially resolved feedback.

Addressing this limitation, microfluidic technologies are increasingly designed to include new capabilities for spatially resolved, real-time measurements of microbiomes. Microfluidic devices compatible with micron- to centimeter-scale imaging have been used to visualize biosensors reporting on environmental parameters (Zhu et al. 2022). Field-deployable microfluidic chips now allow on-site analysis of community nutrient dynamics (Hong et al. 2025). Furthermore, engineered soil habitats integrated with imaging approaches are facilitating the study of microbiomes associated with plants and fungi at intermediate spatial scales (Nagel et al. 2012; Sasse et al. 2019). These integrated platforms complement biosensor approaches, providing critical windows into microbiome function across diverse spatial and temporal contexts. Together, these technologies support the development and fine-tuning of dynamic regulatory systems, an essential step toward ensuring that engineered microbes can respond appropriately to fluctuating environments.

4.2 Knowledge Gaps

Scaling from the Laboratory to the Field

Engineering approaches that succeed in laboratory settings frequently fail when deployed in open systems like soil environments, where native microbes are better adapted to local conditions and outcompete introduced synthetic communities (SynComs). There is no way to scale molecular-level insights from biomolecules (e.g., DNA, RNA, and proteins) into phenotypic variations across all levels of biological organization—from single cells to populations to communities of diverse microbiomes, including those associated with plant hosts. This challenge is multidimensional, spanning biological scales (from pure culture to microbiomes), spatial scales (from micrometers in the laboratory to meters in the field), and temporal scales (from enzyme-mediated behaviors occurring over seconds to long-term, seasonal and multiyear impacts). This scaling challenge is particularly evident in the contrast between controlled, closed systems like bioreactors and the stochastic, dynamic nature of field environments.

Fundamental knowledge learned about biological phenotypes at one scale often does not directly translate to another scale, necessitating integrative experimental approaches and interpretive frameworks. Overcoming this knowledge gap will require experiments across scales that compare spatial and temporal variation. Also needed are models that bridge across scales or design experimental habitats that recapitulate key aspects of environmental complexity (especially spatial structure) at different scales. Addressing the scaling challenge also will require studies that consider the evolution of microbiomes. Particularly important is how mutation rates and population size contribute to the phenotypic trait variation observed. Some variation could arise because larger-scale experiments enable microbial communities to sample more mutations and adapt more readily within a given setting.

System Complexity and Environmental Heterogeneity

Another challenge is anticipating how the chemical and physical characteristics of the environment affect microbiome structure, stability, and function.

Although engineered communities can show specific beneficial functions in controlled settings, such as improved nutrient uptake or plant-disease resistance (Martinez-Feria et al. 2024), these benefits often fail to materialize in the field. Outcomes can be altered by factors such as soil type, environmental heterogeneity, and dynamic contingencies like moisture and interactions with other soil biota.

Environmental gradients add additional complexity. For example, chemical gradients that arise in the rhizosphere create hot spots of metabolic activity (Kuzyakov and Blagodatskaya 2015). In bulk soil, redox and pH gradients occur with increasing depth in soil settings and across soil aggregates. These gradients and hot spots can exhibit complex spatiotemporal variation that affects microbial growth, composition, and behavior. In soil, this variation can be amplified by irrigation, weather, climate, bioperturbation, or other hard-to-predict variables.

Root-exudate dynamics and microbial interactions and secretions in natural soils are key to understanding plant–microbe signaling and cell–cell interactions, which affect microbiome structure and function including soil carbon and nutrient dynamics. These dynamic interactions differ dramatically from those observed in laboratory experiments, but replicating such environmental contexts when evaluating the behaviors of engineered microbiomes is currently intractable. This knowledge gap complicates the deployment of functional synthetic or engineered microbial communities at larger scales.

Achieving precision microbiome engineering across diverse settings and timescales remains a grand challenge in microbiome engineering.

Temporal Dynamics of Microbiomes

Microbiomes are dynamic and context-dependent. Short-term fluctuations are often driven by stochastic processes while long-term trends emerge from multilevel selection, dispersal, genetic drift, and environmental filtering (Shade et al. 2013; van Vliet and Doebeli 2019; Argiroff et al. 2024; Dove et al. 2021). Microbiome engineering therefore must consider a wide spectrum of temporal scales. These range from diffusion-limited processes that regulate protein activities, such as allosteric enzymes and sensors, to microbial growth (days to weeks in the laboratory), to environmental applications (days to seasons to years). Achieving precision microbiome control across these settings and timescales remains a grand challenge in microbiome engineering.

The temporal scale of microbiome shifts in the laboratory, where conditions are stable and experiments are often short term, differs significantly from shifts in the field where environmental variables introduce continual perturbations. These differences necessitate normalization strategies to align laboratory findings with real-world microbial dynamics. Another challenge is that key time points for sampling microbial activity can be missed due to constraints and costs of disruptive sampling. These constraints, along with the stochastic nature of some ecological drivers of microbial activity (e.g., storms), limit sampling resolution.

Multiomic readouts of DNA, RNA, protein, and metabolite profiles often are mismatched, exhibiting asynchronous temporal patterns even within a single organism. Fundamental barriers to scaling microbiome engineering include acquiring and integrating time-resolved datasets that bridge space and time. Microbiomes must be studied across multiple temporal scales, from rapid microbial turnover events to long-term ecosystem shifts. Longitudinal datasets are particularly valuable, providing insights into how microbiomes respond to environmental change, engineered interventions, and host factors over extended periods. Further, sampling at microbial-relevant scales could help integrate spatial heterogeneity and dynamics.

Prediction of Evolution Across Scales

Microbes are continuously experiencing genomic changes due to DNA replication errors, DNA damage, and recombination (Darmon and Leach 2014). Many of these DNA-encoded changes lead to deleterious

phenotypic changes, which cause microbes to be purged through natural selection. Neutral phenotypic changes become propagated within a population, and beneficial changes become fixed. These changes can improve cell fitness by allowing a protein to function under new conditions, altering the flux through a specific metabolic pathway or allowing an enzyme to access new substrates. When these new traits are acquired through mutation, interactions between the evolved microbe and other members of a community can change.

Although mutation sampling is known to be proportional to the size of a microbial population, researchers cannot anticipate sequence–structure–function relationships accurately enough to predict how mutations affect protein stability, activity, and substrate specificity. This critical knowledge gap limits the ability to project the effects of most mutations on the emergent properties of a microbiome at any length scale.

Large genetic changes arising from horizontal gene transfer (HGT) can allow microbes to quickly make significant fitness jumps. As illustrated by the rapid spread of antibiotic resistance, HGT can swiftly improve the fitness of a microbe under certain growth conditions (Lerminiaux and Cameron 2019). However, the research community lacks the ability to predict how the exchange of mobile DNA (e.g., plasmids and viruses) through HGT affects microbial function. The sequencing of numerous conjugative plasmids and viruses has revealed the diversity of genes encoded by these mobile DNA, which can be large. Despite this information, researchers often cannot anticipate the encoded functions nor their potential impact on microbial fitness when acquired via HGT. Neither can they predict how a given environment might affect mutation rates or the spread of DNA, as materials such as soil differentially sorb viruses and DNA depending on composition (Zhuang and Jin 2003).

4.3 Technical Limitations

High-Throughput, Cross-Scale Measurement Platforms

Ecologically relevant SynComs represent a frontier in microbiome research, enabling the design of microbial assemblages for specific ecological functions. These communities can be engineered to mimic or simplify the complexity and functionality of natural microbiomes, which is invaluable for applications in agriculture, bioenergy, and environmental remediation. However, despite significant utility, numerous challenges remain. These include (1) the difficulty of scaling from the laboratory to real-world applications and (2) ensuring that the design and formulation of engineered communities created from isolates are compatible, stable, and functionally robust across diverse conditions.

A significant hurdle to using simplified communities is the inherent complexity of natural ecosystems. Microbial communities in soils are influenced by a multitude of factors, including soil type, moisture, and environmental conditions, as well as plant hosts and other ecological interactions (Delgado-Baquerizo et al. 2025). When attempting to scale engineered communities for field applications, accounting for these variables becomes crucial. The transition between controlled laboratory conditions to the heterogeneous environments of natural ecosystems often leads to unpredictable outcomes. For example, competition from indigenous microbial populations or unsuitable environmental conditions in the field might prevent an engineered community that performed well in a laboratory setting from thriving. Furthermore, the large biomass required for community deployment must be preserved and stored until application, ensuring that the inoculum remains viable.

Effective protocols for formulating and applying engineered communities at field scales are under development and require significant refinement. Current technologies cannot replicate the full temporal and spatial heterogeneity of natural systems, limiting their utility and the successful translation of community-based approaches to real-world applications. To overcome these challenges, it is critical to standardize storage protocols that preserve inoculum viability, implement pilot studies to evaluate application methods across diverse environments, and invest in technologies that better simulate natural spatiotemporal variability. Also needed are field trials to benchmark

performance under realistic conditions and interdisciplinary collaborations to integrate advances in microbiology, engineering, and environmental modeling, which will accelerate the refinement and deployment of effective community-based solutions.

Real-Time Measurements of Microbial and Ecosystem Dynamics

Parallel measurements of environmental fluctuations and microbial function at a scale relevant to community interactions remain a significant limitation. There is a pressing need for biosensors, genetic recorders, and persistent field sensors to track microbial activity, plant responses, and environmental variables over time and space. Current environmental sensors are often insufficient for microbiome engineering applications. This is particularly true for *in situ*, real-time monitoring of microbial and ecosystem dynamics, which requires distributed information across broad scales of length (micrometer to meter to field) and time (day to week to season).

The challenge of integrating temporal and spatial data underscores the importance of continuous, high-resolution environmental sensing to capture long-term microbe–plant–environment interactions. Distributed and persistent field sensors are critical tools for tracking nutrient cycling, drought responses, and disease suppression—all of which are essential for understanding microbiome function in real-world conditions. Sensors are needed at all scales of microbiome measurements, ranging from small habitats in the laboratory to greenhouse settings to the field. Ideally, they would provide dynamic data nondestructively, although sensors that record and remember transient activities *in situ* (e.g., memory biosensors) could be applied because they can be retrieved after incubations to report on past and temporal microbiome activities (Essington et al. 2024). While early examples of such technologies have been reported, ensuring they provide the desired information is a challenge.

Controlling Microbe Persistence

A wide range of biocontainment approaches have been described, including metabolic dependencies (Asin-Garcia et al. 2022), genetic kill switches (Stirling et al. 2017), genetically recoded organisms (Nyerges et al. 2023), and entangled genes (Blazejewski et al. 2019). While promising, their performance in natural settings and across diverse environmental conditions where engineered microbes might be used remains difficult to assess. Relevant test-bed systems would be useful for benchmarking the performance of safety measures when used to control individual or combinations of microbes within a microbiome and across the spectrum of soil types and conditions (Gillett et al. 2025). As such, safety measures designed to function in environmental microbes that take up engineered DNA need to be engineered to function robustly across a broad diversity of transformable environmental microbes. Furthermore, there is a technical need to understand the fate of DNA following engineered cell death and the persistence of such DNA in the environment. It also remains unclear how biocontainment in engineered cells is affected by processes like cell dormancy, which can be prevalent in environmental microbiomes (Blagodatskaya and Kuzyakov 2013).

4.4 Research Opportunities

Scaling Up Testing Through Interdisciplinary Collaboration and Distributed Field Sites

Cross-Scale Collaborative Teamwork

A grand challenge in microbiome engineering is to connect the results from laboratory-scale habitat measurements to greenhouse studies and, ultimately, to environmental applications and to understand variation across these scales. One way to meet this challenge is to encourage collaboration and interdisciplinary team research that spans length and complexity scales. Such teams would leverage expertise ranging from molecular-scale understanding of biological mechanisms to field applications and would include translators who can effectively interpret findings across domains.

Another opportunity lies in developing ecological theories that specifically capture scalable processes in evolving microbiome systems. These efforts may require more generalizable frameworks for considering microbiome ecology rather than more in-depth

knowledge of specific metabolic or genetic mechanisms, since broader principles may initially be more likely to scale across different environments and conditions.

Advances in Data Sharing

Improved data-sharing frameworks can also accelerate progress by enabling researchers to build upon each other's findings digitally without the friction of human-mediated data transfers. These frameworks require identifying the most significant biotic and abiotic metadata across different microbiome engineering studies to allow for meaningful comparisons. AI and machine learning (ML) approaches represent an opportunity to understand how to integrate complex datasets across scales and to identify patterns that might inform more robust microbiome engineering designs and solutions. Future funding models that explicitly support these types of cross-scale collaborative efforts and teamwork are critical.

Expanded Field Sites

To translate microbiome engineering into applications, scientists will ultimately need a framework for field-based testing, which is essential for assessing the real-world feasibility, stability, and scalability of engineered microbiomes. There is demand for dedicated reference field sites for designated studies to investigate engineered microbiome stability, ecological interactions, and long-term impacts on ecosystem function across diverse soils, climates, and plant species. Strategic access to long-term field sites with appropriate regulatory permits would provide an opportunity to conduct rigorous experiments to capture the complexity of real-world ecosystems and their inherent spatial and temporal heterogeneity. By assessing microbiome performance under these variable environmental conditions, the robustness of engineered microbial communities could be evaluated. Such field-based testing is fundamental for rigorous testing of safety features developed to determine when and where genetically engineered microbiomes persist.

A dedicated network of field sites is needed for microbiome engineering that mirrors the Long Term Ecological Research (LTER) Network supported by the National Science Foundation, which enables thousands of scholars to study ecological processes *in situ*. Expanding distributed field networks available for microbiome engineering studies, including SynComs, would allow researchers to bridge the gap between controlled experimentation and large-scale application. This would also remove the onus of personally maintaining these sites, which is not possible for individual researchers and many academic institutions.

Such field sites represent an opportunity to develop shared infrastructure not only for performing field experiments, but also to leverage next-generation field sensors capable of measuring real-time changes in pH, nutrient fluxes, and carbon and nutrient dynamics at the soil aggregate scale. This infrastructure would revolutionize the ability to study plant–microbe–soil interactions, providing the level of detail necessary to link microbial function with ecosystem processes across temporal and spatial scales. The integration of real-time environmental monitoring technologies into field-based high-throughput phenotyping at these sites would further enhance the ability to track plant–microbe interactions, soil health, and ecosystem function over time.

Chapter 5

Advancing Microbiome Engineering Using Modeling and Artificial Intelligence

To realize the full potential of engineered microbiomes, models are needed that can leverage multiomic and environmental data to predict and control the collective dynamic functions of microbial communities across time and space. This chapter details state-of-the-art approaches currently used for modeling microbiome function, including genetic programming, metabolic, and ecological models, as well as models for guiding the genetic engineering of microbiomes. Applications of artificial intelligence (AI) and machine learning (ML) are also discussed, along with key knowledge gaps and technical limitations that hinder predictive modeling of microbiomes. Finally, opportunities are presented to advance microbiome engineering through development of multiscale modeling frameworks, self-driving laboratories, and integrative AI tools.

5.1 State of the Art in Microbiome Modeling

Diverse models have been described for anticipating the emergent functions of individual microbes, synthetic microbial communities (SynComs), and microbes associated with hosts. However, these models cannot yet be used to reliably anticipate microbiome structure, function, and stability in any natural environment. Each model has unique strengths and weaknesses that must be overcome to predict community behaviors across space and time in real operational environments.

Metabolic Models

Genome-scale metabolic models (GEMs) use the network of biochemical reactions within individual microbes to predict growth and metabolic outputs (Gu et al. 2019). Community metabolic models (CMMs) for microbiomes can be created by combining individual GEMs. In CMMs, each microbe is represented as a separate compartment in the model, and the exchange of nutrients is simulated by the transport of metabolites into and out of a shared extracellular compartment (Colarusso et al. 2021). CMMs have identified complementary metabolisms between strains in which one microbe produces nutrients that another needs (Stolyar et al. 2007). These models have informed cultivation approaches by identifying metabolic needs of uncultivated strains (Scarborough et al. 2020). While some models such as Computation of Microbial Ecosystems in Time and Space (COMETS; Harcombe et al. 2014) capture explicit spatial structure in communities, they have limited accuracy with non-model organisms and larger consortia, since they rely on well-curated pathways and steady-state assumptions.

Ecological Models

Classic approaches like the generalized Lotka-Volterra (gLV) equations (Wangersky 1978) and

consumer-resource models (Lafferty et al. 2015) have been widely used to model dynamic cell–cell interactions. Although these models can fit experimental data, complex microbiomes pose unique challenges because interactions are often context-dependent and involve higher-order interactions beyond pairwise. In practice, gLV model assumptions break down when applied to natural ecosystems because they lack (1) sufficient data to parameterize the model, (2) the ability to capture heterogeneity, and (3) the representation of nonstatic environmental parameters. Synthetic biology has begun to provide control theory for ecological modeling—a field of engineering and applied mathematics that deals with the behavior of dynamical systems (Angulo et al. 2019). Simple microbial communities can now be programmed to present complex phenotypes, such as emergent genetic oscillations (Chen et al. 2015), defined social interactions (Kong et al. 2018), division of labor (Zhang et al. 2021), and proportion control (Grandel et al. 2025). Currently, these models and theories can be applied only to simple laboratory organisms under a narrow set of conditions.

Genetic Programming Models

Models for genetic engineering can be divided into two classes: (1) models that design components of genetic circuits such as protein-coding sequences (Chen et al. 2013), promoters (LaFleur et al. 2022), ribosomal binding sites (RBS; Salis et al. 2009), and terminators; and (2) models that anticipate how the assembly of different components can be leveraged to execute DNA-coded functions in cells (Chen et al. 2013). Beyond enabling genetic programming, these models address a key limitation of GEM models by accounting for context-dependent gene expression and enabling dynamic predictions of metabolic function, which GEMs cannot capture because they rely on static, steady-state assumptions. Existing models for genetic programming have been trained only on model organisms, such as laboratory strains of *Escherichia coli* grown in rich medium. As such, these models struggle to anticipate component functions in non-model microbes and as environmental conditions change from ideal laboratory settings.

Artificial Intelligence and Machine Learning Applications

AI and ML can be used to derive meaning from microbiome data (Tourab et al. 2024) by simplifying the identification of microbes in samples via imaging (Zhang, J. et al. 2022), distinguishing diseased versus healthy host states (Oh and Zhang 2020; Marcos-Zambrano et al. 2021; Lee and Rho 2022; Liñares-Blanco et al. 2022), and decoding pattern formation rules in engineered microbiomes (Lu et al. 2024). Deep learning models now outperform classical ecology models in replicating microbial time-series data and linking community structure to ecosystem processes like decomposition (Ridenhour et al. 2017; Thompson et al. 2019; Borowiec et al. 2022; Pichler and Hartig 2023), illustrating their potential for microbiome engineering. Furthermore, AI-enabled advances in protein-structure prediction (i.e., AlphaFold) have accelerated the ability to design enzymes for genetic programming (Jumper et al. 2021). These successes represent proof-of-concept, but the reliability of predictions in other microbiome engineering scenarios remains limited.

5.2 Knowledge Gaps

Incorporating Complexity into Models To Predict Emergent Properties

Dynamic properties in engineered microbiomes, such as oscillations or traveling growth waves, remain difficult to represent using current modeling approaches. Addressing this challenge requires overcoming key uncertainties, including how to define organisms or symbionts in microbiome models. Determining whether the models should focus on individual strains, species aggregates, or functional guilds within communities remains a question. Functional redundancy makes species-centric approaches less effective, and although intraspecies genomic diversity and competition can influence community stability and function, they are rarely modeled.

Another limitation of existing models is their inability to adequately handle environments where spatial structure, nutrient gradients, heterogeneity, and physical constraints shape community dynamics.

Most pairwise microbe–microbe interactions are still unknown, and cell–cell interactions depend on more than just the identities and abundances of organisms. The spatial locations of microbes are also critical. As community complexity increases, modeling higher-order and context-dependent interactions within heterogeneous environments becomes difficult, especially since these interactions grow combinatorially and include spatial feedbacks such as microbial-driven pH shifts and redox gradients.

Temporal variation, including dynamic changes in gene expression, weather, and diurnal cycles, introduces additional complexity that models must capture. Viruses also significantly impact microbiome dynamics yet are excluded from most current models, as is evolution through mutation, recombination, and horizontal gene transfer (HGT). Because engineered microbiomes are subject to selective pressures from changing environments, models need to incorporate adaptive processes, such as DNA mutations, gene transfer, and virus-mediated genetic exchange.

Developing and Curating Optimal Datasets

Microbiomes exist in a complex, multidimensional design space shaped by species composition, gene content, and environmental conditions that are heterogeneous and dynamically changing. The number of possible community configurations increases combinatorially with species count, making exhaustive exploration of this organismal space infeasible. This is similar to the challenges biochemists face in navigating protein sequence space (Romero and Arnold 2009). For example, a 10-species community can have thousands of cell–cell interactions and an astronomically large number of possible subcommunity combinations due to the way microbes colonize distinct niches within samples and host tissues.

This combinatorial explosion limits the effectiveness of trial-and-error experimentation because sampling most potential cell–cell interactions is impossible. In modeling, this results in underdetermined systems with many unknown parameters such as interaction coefficients, which cannot be uniquely estimated from limited data. A major knowledge gap is understanding how to construct models that effectively navigate this space. New dimensionality-reduction strategies that preserve critical information are currently needed to advance microbiome modeling.

Another challenge in modeling microbiomes is the absence of a clear definition for what constitutes high-quality reference datasets to parameterize and benchmark models. In contrast to genomics, where such datasets are well established (The Reference Genome Group of the Gene Ontology Consortium 2009), microbiome models are built using a wide spectrum of data types. These range from rRNA gene surveys that reveal community composition, to metagenomic data that captures metabolic potential, to dynamic datasets that reflect community functions in time and space. It is unclear what data types comprise the minimal information needed to enable accurate predictions.

There is a critical need to determine the relative importance of different parameters, such as taxonomic composition and structure, microbial abundances, multiomic data (e.g., transcripts, proteins, or volatile organic compounds), and metabolic fluxes. Likewise, key environmental metadata—such as pH, nutrient concentrations, spatial gradients, heterogeneity, and hydration—must be identified. Developing datasets that comprehensively capture both abiotic and biotic variables will be essential for advancing predictive modeling and enabling rational microbiome engineering.

Understanding Sequence–Function Relationships

A major barrier to modeling is the inability to predict gene expression *in situ*, where environmental stresses and cell–cell interactions are known to influence promoter activity, leading to unpredictable expression outcomes. This knowledge gap significantly limits the rational design of gene circuits in undomesticated microbes and natural microbiomes. Addressing this challenge will require the generation of large, well-annotated datasets documenting the activities of native and synthetic regulatory elements (e.g., promoter and RBS) across diverse organisms and conditions.

AI and ML approaches could be applied to develop models that map DNA sequences to expression outcomes. Metagenome mining combined with multiplexed testing of regulatory elements using barcoded constructs and sequencing represents a promising strategy. Advanced models, potentially incorporating deep learning to identify regulatory motifs, will be essential for designing gene circuits with predictable behavior across multiple microbial species. Although theoretical frameworks (i.e., control theory) exist to describe complex community properties such as oscillations (Stricker et al. 2008), these frameworks remain largely limited to SynComs containing pairs of well-characterized laboratory microbes.

All models would benefit from community-wide data standards that support greater consistency in reporting, thereby enabling integration of data from different experiments and platforms into models.

Overcoming Mismatches Between Experimental Data and Models

Calibrating models using experimental microbiome data from different studies is challenging because the types of data and metadata reported vary. All models would benefit from community-wide data standards that support greater consistency in reporting, thereby enabling integration of data from different experiments and platforms into models. More consistent gene annotation and naming conventions for metabolites and reactions across databases would support these efforts. Other needs include improved understanding of the causes of data variation between laboratories (even when similar approaches are used) and strategies to integrate data across experiments by independent research teams.

Additional efforts should be made to better align models and experimental measurements. For example, the heterogeneity found in environmental microbiome samples should be captured in models, which often assume well-mixed conditions. Models should more rigorously consider cells in different metabolic states (e.g., planktonic, biofilm, and dormant) that are common *in situ*. The dynamic nature of necromass formation needs to be considered as well because it impacts environmental heterogeneity and alters the environment where cells are growing. Virus contributions to gene transfer and microbial dynamics must also be accounted for, as these affect microbial community phenotypes and necromass formation.

Minimal Data Dimensionality To Train Models

Beyond developing models that better capture real-world conditions, simple strategies are needed to validate model predictions across scales. Matching molecular predictions (e.g., protein expression) with community-level behaviors in complex samples (e.g., metabolite production) can sometimes require complex analysis such as metabolite-labeling experiments. In addition, AI models tend to require substantial amounts of data for training and validation, necessitating both high-quality datasets and sufficient data volumes. Some models may not require massive amounts of characterization data for effective predictions. Currently, there is a need to understand the minimal data that must be collected to enable accurate predictions for each type of model that is being developed to predict community composition, function, and stability. By establishing this minimal data dimensionality, arduous characterization efforts can be streamlined to more rapidly and affordably acquire critical data for training models.

5.3 Technical Limitations

AI-Assisted Experimental Platforms

Automated experimental platforms guided by real-time modeling have the potential to accelerate the design-build-test-learn (DBTL) cycle by integrating automation that performs high-throughput experimentation, adaptive sampling, and AI-driven decision-making to create an active learning pipeline (see Fig 1.4, p. 6). When applied to protein design, self-driving laboratories have accelerated the discovery of biomolecules with new functions by decreasing the number of variants that must be characterized to reach desired

products (Singh et al. 2025). To apply AI-assisted experimental platforms to microbiome engineering, several technical challenges must be addressed:

- Establish high-throughput workflows with low variances for different engineering goals, including mixing isolates to create SynComs with user-defined functions and genetically programming microbes and communities with new traits.
- Facilitate complex, multivariate measurements using biosensors that can be multiplexed to report on above- and belowground system properties and observe key functions.
- Capture context-dependent metabolisms across different scales of length and time (i.e., days to months).
- Develop soil and microbiome standards that enable researchers to benchmark the performance of engineered soil microbiomes across biological, chemical, and physical parameter spaces.
- Develop hardware platform innovations optimized for working with spatially structured samples, such as soils and biofilms, which present many high-value microbiome engineering opportunities and applications.

To solve these challenges, tight integration of operational sample handling and monitoring must be coupled with multidimensional environmental and imaging data collection.

Control Theory for Genetic Programming

There is a need for computational tools grounded in control theory that anticipate the dynamic phenotypes encoded by engineered genetic programs as well as their stability and performance in the context of diverse environmental conditions. Computational tools grounded in control theory cannot currently guide the design or assembly of synthetic DNA parts (e.g., promoters, RBS, coding DNA sequences, and terminators) for non-model organisms. There is a need to extend the techniques and tools developed for anticipating DNA component functions in model organisms, such as RBS and promoter calculators, to non-model organisms. Furthermore, there is a need to extend control theory developed to inform the design of genetic circuits that encode dynamic behaviors (e.g., oscillations) from pairs of laboratory microbes to more complex communities in real heterogeneous environments.

5.4 Research Opportunities

Multiscale Modeling Frameworks To Capture Heterogeneity

A shift from static community abstractions toward dynamic, spatially grounded modeling is essential to support predictive engineering of communities in real-world, complex environments. As such, an opportunity exists to develop multiscale modeling frameworks that capture the spatial and temporal complexities of microbial communities. Potential advances include:

- Creating agent-based models that simulate the behavior of individual cells and organisms in space and time while capturing heterogeneity in growth, diffusion, and interactions (Nagarajan et al. 2022).
- Developing partial differential equation models that describe the spatial and temporal evolution of metabolite concentrations, signaling gradients, and biomass densities.
- Developing hybrid models that couple ecological population dynamics with metabolic or regulatory network models, which can be parameterized with experimental imaging or microfluidic data.
- Using ML methods to guide model discovery by inferring spatial interaction kernels or local rules from time-lapse measurements from microscopy or sensor-rich datasets.

Such approaches will require algorithmic innovation and new types of data that capture key observables in heterogeneous microbiome samples.

Microbiome–Host Interaction Models that Anticipate Dynamics

Emerging modeling frameworks include host processes to support the engineering of beneficial microbial communities. In human health research, GEMs simulate both host and microbial metabolism together (Rodenburg et al. 2021). In plant systems,

these models represent an opportunity to predict how engineering changes in a microbiome might alter host behaviors, such as rhizosphere metabolite levels. Models that capture root exudation profiles and microbial uptake could be used to predict how plants sculpt the rhizosphere microbiome and how that microbiome influences plant growth (Roque-Malo et al. 2020).

The predictive power of such models could be enhanced by establishing which host factors are crucial to include (e.g., the minimum complexity required for predictions) and integrating host data (e.g., genomic and physiological) with microbiome models. AI could further assist by relating host and microbiome patterns, such as linking shifts in microbiome genes with host health indicators to identify causal relationships. Another opportunity involves development of standardized host–microbiome model platforms with customizable host compartments. As multiomic data from host-associated microbiomes become more available, such integrated models can be trained and validated.

Minimum Data Dimensionality for Prediction

Models are needed that capture essential features without accounting for every microbiome parameter in full detail. A recent study showed that microbial community outcomes can be described with relatively few interaction terms, illustrating the extent of negligible higher-order interactions (Arya et al. 2023). Such approaches represent an opportunity to use compressive sensing to predict community composition from limited experiments.

AI can help reduce data dimensionality. Autoencoder neural networks, which have been used to compress microbial growth dynamics into low-dimensional representations (Baig et al. 2023), provide an opportunity to compress multiomic data into a few latent variables that correlate with microbiome functions of interest. Such models not only would simplify computation, they also would be easier to interpret and visualize.

Additionally, instead of trying to model every species interaction, researchers could model community-level emergent properties (e.g., total fermentation rate and nutrient cycling efficiency) as functions of a few community traits. Metabolite-centric or function-centric models could map thousands of genes to a handful of pathways that control the microbiome engineering outcome.

Finally, guild-based models represent an opportunity to decrease complexity by grouping species into functional guilds (e.g., cellulose degraders and nitrogen fixers) using AI in a way that aligns with genomic inference and ecological theory.

Coarse-Grained Models for Gene Circuit Design

A major challenge with genetic programming is the context sensitivity of gene circuits (Stone et al. 2024). DNA circuits that behave predictably in monocultures frequently break down in complex environments. Even when detailed models of transcriptional and translational regulation exist, they are rarely sufficient to predict dynamics of circuit behavior in spatially structured environments. New control theory is currently needed for gene circuit design.

One promising opportunity to address this challenge is to develop coarse-grained models that include molecular-level details and emergent properties at population and community scales. Rather than modeling the strengths of promoters and RBS, researchers can model the fraction of cells in a population carrying a particular function or the abundance and persistence of mobile DNA that encodes that function (Son et al. 2025). These models incorporate parameters such as plasmid loss, gene-transfer rates, and fitness costs to predict dynamic functions. This approach could guide the design of circuits whose performance depends on how genetic material spreads and persists in a population.

Such models offer two key advantages. First, they provide flexible control. By modulating gene frequency or copy number at the population level, circuits can achieve orders-of-magnitude shifts in output that would be difficult to attain through promoter engineering. Second, they enhance stability. Instead of overloading individual cells, the burden of circuit execution can be

spread across many community members, improving robustness and long-term function (Wang et al. 2022).

Coarse-grained frameworks also provide an opportunity to align with control theory design principles, allowing synthetic biologists to implement feedback regulation and memory at the level of community composition. By coupling gene circuit activity to environmental selection or resource access, such systems have the potential to maintain homeostasis, exhibit hysteresis (a lag between input and output in a system upon a change in direction), and shift between discrete functional states depending on ecological or chemical inputs. Further opportunities to advance this area of research include the development of:

- Circuit architectures that are modular and population-aware by leveraging feedback, quorum sensing, and gene transfer as layers of control beyond intracellular logic.
- Testing platforms that mimic real-world dynamics including fluctuating nutrient availability, spatial gradients, and community interactions.
- Modeling approaches that abstract away unnecessary molecular details while still capturing key dependencies and enabling predictions across conditions.
- Quantitative metrics of robustness, including context-sensitivity analysis, host burden evaluation, and dynamic range estimation over time.

Achieving robust dynamic behavior in gene circuits will ultimately require a shift from fine-grained, cell-centric design to population-informed, coarse-grained modeling approaches that embrace the complexity of the microbiome as a design constraint. These frameworks represent an opportunity to bridge the mechanistic rigor of synthetic biology and the emergent dynamics of microbiome ecology.

AI-Driven Microbiome Engineering Laboratories

Microbiome engineering can leverage active learning by using AI agents to identify both the microbial community variants to mix and the environmental perturbations to test and then use the resulting microbiome behaviors to inform subsequent modeling rounds. For example, communities could be designed *in silico* by predicting metabolic dependencies from existing omics data. Those communities can then be built and tested in habitats of varying complexities to evaluate hypotheses. Results would feed back into the model to refine the next DBTL cycle.

Such iterative approaches have been explored using a synthetic human gut community. A recent study used a long short-term memory framework to model time-dependent changes in species abundance and the production of key health-relevant metabolites in a human gut SynCom (Baranwal et al. 2022). This model outperformed a Lotka-Volterra model based on ecological theory and could guide the design of synthetic microbiomes with dynamic functions. Such models represent an opportunity to develop self-driving laboratories for microbiome engineering. These laboratories could be integrated with automation to increase efficiency of the DBTL cycle far more effectively than by human intuition and statistics alone, as the models learn from mistakes and surprises. In parallel with AI tool development, innovative strategies are needed to reap the benefits of AI tools without risking misuse.

Self-driving laboratories for microbiome engineering could be integrated with automation to increase efficiency of the DBTL cycle far more effectively than by human intuition and statistics alone.

Integration Across Data Types

Microbiome research produces disparate, difficult-to-integrate data types, ranging from complex multiomic community snapshots to community behavior dynamics to environmental contexts. AI and ML tools (e.g., network analysis, manifold learning, and graph neural networks) represent an opportunity to identify connections across data types that classical analyses might miss. Recent reviews affirm this opportunity, noting that the

analysis of different datasets together have the potential to uncover functional genes and microbial activities linked to environmental traits (Mathieu et al. 2022).

One opportunity to make connections across data is by constructing knowledge graphs for microbiomes. These graphs use networks to connect microbial genes, species, and environmental parameters. AI reasoning over such graphs could be used to generate testable hypotheses about cell–cell interactions in communities, such as potential species in a community that suppress other organisms through the biosynthesis of specific metabolites. Another opportunity is the use of AI to harmonize data from different sources, addressing quality control issues and noise when combining datasets. Moving forward, the development of pipelines where different microbiome engineering datasets are integrated with minimal human bias would enable models to find emergent patterns that underlie microbiome function.

Chapter 6

Resources Needed To Advance Microbiome Engineering

DOE possesses a uniquely powerful ecosystem of national user facilities and computational platforms that have played a key role in advancing microbiome science for more than two decades. Available resources span molecular, organismal, and ecosystem scales enabled by a data ecosystem maintained through the DOE Joint Genome Institute; National Microbiome Data Collaborative (NMDC); and DOE's Systems Biology Knowledgebase (KBase), which provides an integrated computational environment for multiomic analysis and modeling.

DOE's leadership in high-performance computing (HPC), exemplified by exascale systems like Frontier, Aurora, and El Capitan, enables predictive simulations of microbial networks and gene flow at unprecedented resolution. In addition to these capabilities, DOE's light and neutron sources are uniquely powerful resources for imaging and structural analysis of biomolecules.

While these resources are considerable, a range of investments in infrastructure, workforce development, and coordinated programs are needed to drive microbiome engineering forward and bridge experimental and computational domains for future innovation. Advancing the field will require a coordinated investment in shared resources, standardized systems, and interoperable data frameworks to ensure that research is reproducible, scalable, and comparable across laboratories.

Advancing the field will require a coordinated investment in shared resources, standardized systems, and interoperable data frameworks to ensure that research is reproducible, scalable, and comparable across laboratories.

6.1 Community Standards for Microbiome Engineering

Well-Defined Microbial Communities

The research community lacks well-defined, reproducible microbial communities that can serve as model systems across laboratories and experiments. Just as *Escherichia coli* and *Arabidopsis* are standard models for microbial and plant geneticists, respectively, reference communities would enable consistent experiments across laboratories to build and test models. A recent community consensus paper emphasized the importance of defining reference synthetic communities (SynComs) and creating standardized protocols and benchmark datasets for working with these communities (Northen et al. 2024). Ideally, isolates included in SynComs would have accompanying reference genomes and phenotypic characterizations.

Standardized Soils

Also needed are standardized soils to enable comparative analysis of microbiome function across laboratories (Del Valle et al. 2022). A defined set of artificial soils that sample a range of environmental conditions (e.g., sand, silt, and clay types) and incorporate relevant chemical and physical matrix properties would allow researchers to test engineered communities under controlled, comparable conditions. Such standards are fundamental for facilitating studies that relate microbial behaviors to the vast abiotic parameter space found in real environments.

Model Host Organisms

Model host organisms should be developed as shared community resources. The research community already is using accessible wild-type plants like *Arabidopsis* spp. and *Brachypodium* spp. However, model host organisms specifically engineered for microbiome research—such as those genetically modified for reporter activity or controlled exudation of microbiome-relevant compounds—either don't exist in standardized forms or are distributed informally through individual laboratory mutant or transgenic collections. Establishing such engineered host systems, particularly for DOE-relevant crops, including sorghum, switchgrass, and poplar, would be invaluable for testing ecological theories, benchmarking models, and evaluating engineered microbiome activities in reproducible ways.

Consistent Data Standards and Centralized Data Repositories

Consistent data standards are needed to support cross-platform modeling efforts. These standards should ensure that experiments capture critical data for every sample including (1) biological variables such as community composition and multiomic outputs including transcripts and proteins; (2) chemical parameters such as metabolite production, dissolved organic matter, pH, and mineral profiles; and (3) physical properties like temperature and soil aggregation.

Also needed are centralized data repositories with uniform metadata structures and file formats. These repositories would enable the aggregation of high-quality, interoperable datasets required for robust artificial intelligence (AI) and machine learning (ML) applications in microbiome engineering.

6.2 Microbiome Data Atlases and Biobanks of Strains, Genetic Parts, and Biosensors

Although biologists have generated immense amounts of microbiome data, the research community lacks large, curated databases mapping the diversity of microbiome components and functions. Resources such as an atlas of metabolic capabilities across isolates and a catalog of microbial interactions under various conditions are not available but would provide a knowledge baseline that accelerates the design of new communities.

Also needed are biobanks of microbial strains, genetic parts, and biosensors. Developing these would require maintaining collections of culturable isolates useful for microbiome work, including representatives of key functional groups, isolate mixtures, and DNA parts (e.g., broad-host-range vectors, phages, regulatory sequences, and coding sequences). Easy access to these resources would accelerate microbiome engineering because researchers could request a standardized SynCom or a suite of verified genetic tools for non-model microbes rather than isolating or constructing their own for each experiment. Part of this vision includes a framework for sharing strains and data similar to how model organism stock centers operate.

6.3 Automated Microbiome Engineering Workflows

Microbiome engineering must traverse a massive parameter space. Communities often contain thousands of microbes, each having thousands of genes whose activities vary dynamically. The robustness and agility of microbiomes often lead to different activities and biological functions as environmental conditions change. This limits the ability to construct models to guide microbiome engineering. Automated workflows offer a promising way to strategically sample this vast

parameter space, as AI can be coupled with these workflows to reveal the critical drivers of microbial community assembly, stability, and productivity for the next design-and-build phase.

Efforts are critically needed to identify the conditions under which automation provides the greatest benefit. Automating the design-build-test-learn (DBTL) cycle can improve reproducibility in the test phase and generate high-quality metadata to support AI applications (see Fig. 1.4, p. 6). Automation can provide additional DBTL cycle benefits including:

- Isolation and identification of previously uncultured microbes.
- Engineering of SynComs using different isolate combinations.
- Studies of mobile DNA host range in communities.
- Mapping the behaviors of genetic parts across strains.
- Genomic engineering of isolates and microbiomes.
- Aerobic and anaerobic manipulations that replicate host environments.
- Characterization of engineered community functions.

These different automation applications are expected to require unique device solutions to accelerate microbiome engineering.

Automation also presents opportunities to increase understanding of engineered microbiome functions when applied to phenotyping microbiomes that associate with plant hosts. For example, automation can monitor both above- and belowground plant and microbial traits. Aboveground opportunities include automation that leverages emerging tools for obtaining dynamic information without disrupting plant-associated microbiome samples (e.g., hyperspectral imaging and volatilome analysis). Belowground opportunities include expanding the automation of root-zone imaging, *in situ* omic profiling, and biosensor reporting.

6.4 Adaptive Self-Driving Laboratories with Next-Generation Sensors

Integrating automation into adaptable self-driving laboratories can be used to create microbiome research platforms for culturing and characterizing communities across diverse environmental conditions. These capabilities would enable a predictive understanding of how microbiome manipulations underlie complex emergent properties.

Recent culturomics successes have illustrated the potential of self-driving laboratories. ML has been used to leverage genomic data observed from automated experiments and colony morphology to maximize the diversity of isolated microbes and enable targeted picking of specific genera (Huang et al. 2023). Similar innovations in automation were developed to advance engineered microbiome testing to seamlessly integrate miniaturized and noninvasive sensors into habitats. These self-driving laboratories should be extended to habitats containing host-associated microbiomes. Such automation would benefit from inclusion of sensors that allow users to monitor critical metabolites and nutrient fluxes, cell–cell signaling, energy flow, and the processing of labile and refractory materials that underlie biogeochemical cycling.

The habitats in self-driving systems should be coupled to closed-loop algorithms capable of sampling key experimental conditions in real time including temperature, humidity, precipitation, and nutrient fluxes. Some automation capabilities already exist within DOE, such as the Agile BioFoundry (Asin-Garcia et al. 2024), and have been recently targeted for development by the National Science Foundation (NSF 2025). These resources are still sparse, limiting the number of researchers who can leverage them to accelerate the DBTL microbiome engineering cycle.

6.5 Targeted High-Performance Computing Resources

Using computational models to drive the DBTL cycle for microbiome engineering is expected to be computationally intensive, necessitating cutting-edge, HPC

resources. These resources are expected to become a research bottleneck as microbiome engineering experiments generate larger, terabyte-scale multidimensional datasets—such as genomes, transcriptomes, proteomes, and metabolomes with metadata across diverse samples and time points within complex materials. Meeting this challenge will require optimizing algorithms for HPC environments, including making algorithms efficient for simulating the level of DNA complexity found in microbiomes. Also needed are user-friendly HPC interfaces allowing researchers to more easily run an ensemble of community models. By developing improved infrastructure, biologists can pursue more ambitious microbiome modeling, such as exhaustive *in silico* screening of community configurations or real-time analysis of microbiome data from field experiments.

Appendix A

Workshop Agenda

All times Eastern

Day 1: Monday, December 16, 2024

Defining the State of the Art for Microbiome Engineering

12:00 p.m.	**Welcome** ***Speakers:*** Dorothy Koch, Associate Director, U.S. Department of Energy Biological and Environmental Research (BER) program; and Todd Anderson, Director, BER Biological Systems Science Division
12:10 p.m.	**Opening Remarks and BER Workshops Overview** ***Speaker:*** Boris Wawrik, Program Manager, Genomic Science Program
12:25 p.m.	**Summary of Brainstorming Results and Workshop Assignments** ***Speakers:*** Melissa Cregger (Workshop Co-Chair), Oak Ridge National Laboratory; and Jonathan (Joff) Silberg (Workshop Co-Chair), Rice University
12:40 p.m.	**Q&A Session**
1:00 p.m.	**Breakout 1: Define State-of-the-Art Technologies for Microbiome Designing and Building** This session will explore cutting-edge technologies for designing and building complex microbial communities in DOE-relevant research. Participants will discuss both current tools and emerging technologies, with a focus on those that show promising future applications. The session aims to identify key advancements, challenges, and opportunities for innovation in microbial community research. Expected outcomes include a list of relevant technologies and potential areas for future research and collaboration. (Design for 20 minutes; Build for 20 minutes)
1:45 p.m.	**Break**
2:15 p.m.	**Breakout 1 Report-Out**
2:45 p.m.	**Breakout 2: Define State-of-the-Art Technologies for Microbiome Testing and Learning** This session will focus on defining state-of-the-art technologies for testing and learning in engineered microbiomes, covering both current tools and emerging advancements. Participants will explore promising new technologies on the horizon and their potential impact. The discussion will aim to identify key challenges and opportunities in advancing microbiome engineering research. (Test for 20 minutes; Learn for 20 minutes)

3:30 p.m. **Breakout 2 Report-Out**

3:55 p.m. **Closing Remarks**
Speaker: Jonathan (Joff) Silberg

Day 2: Tuesday, December 17, 2024

Using Microbiome Scenarios to Understand Fundamental Barriers to Innovation and Needs to Stimulate Microbiome Engineering Research

12:00 p.m. **Welcome and Day 2 Overview**

Breakout sessions on Day 2 will focus on solving real-world problems through collaborative interdisciplinary research. Topics include carbon-negative bioenergy production, advancing a circular bioeconomy, and understanding microbial interactions. Each session will emphasize practical solutions and future research opportunities to address key environmental and industrial challenges. Participants will work together to identify actionable strategies for these scenarios, specifically focusing on fundamental research that is needed to realize solutions.

12:15 p.m. **Breakout Session 3: Scenario 1—Carbon-Negative Bioenergy Production**

12:55 p.m. **Breakout 3 Report-Out**

1:20 p.m. **Break**

1:30 p.m. **Breakout Session 4: Scenario 2—Circular Bioeconomy**

2:10 p.m. **Breakout 4 Report-Out**

2:35 p.m. **Break**

2:45 p.m. **Breakout Session 5: Scenario 3—Microbial Interactions**

3:25 p.m. **Breakout 5 Report-Out**

3:50 p.m. **Closing Remarks**
Speaker: Melissa Cregger

Day 3: Wednesday, December 18, 2024

Grand Challenges with Microbiome Engineering and Translation to Real Applications

12:00 p.m. **Welcome and Day 3 Overview**

12:15 p.m. **Breakout Session 6: Common Challenges**

This session will address three key challenges: (1) scaling from lab to field, (2) advancing models and predictions, and (3) accelerating the design-build-test-learn (DBTL) cycle. Groups 1 through 3 will focus on scaling. Group 4 will focus on improving models. Groups 5 through 7 will focus on speeding up the DBTL cycle. These sessions aim to identify strategies to overcome these challenges through collaborative research.

12:45 p.m.	**Breakout 6 Report-Out**
1:15 p.m.	**Break**
1:45 p.m.	**Breakout Session 7: Additional Challenges** In this session, participants will explore fundamental challenges within their areas of expertise. Goals are to identify what is particularly difficult to achieve today and to collaboratively brainstorm fundamental research over the next decade that is required to realize solutions.
2:30 p.m.	**Breakout 7 Report-Out**
3:00 p.m.	**Breakout Session 8: Potential Synergies and Opportunities** In the final breakout of the workshop, we will reflect on all we have discussed over the three days and identify critical interdisciplinary interactions that are needed to drive innovation. We will look for near-term opportunities and ways to translate research from other fields to DOE-relevant problems.
3:20 p.m.	**Breakout 8 Report-Out**
3:45 p.m.	**Conclusion and Next Steps** ***Speakers:*** Melissa Cregger and Jonathan (Joff) Silberg
4:00 p.m.	**Adjourn Workshop**

Appendix B

Workshop Participants

Organizers

Melissa Cregger*, *Oak Ridge National Laboratory*

Jonathan (Joff) Silberg*, *Rice University*

Boris Wawrik, *U.S. Department of Energy*

Attendees

Bryn Adams, *U.S. DEVCOM Army Research Laboratory*

Michaeline Albright, *Allonnia*

Greg Bonito*, *Michigan State University*

Otto Cordero, *Massachusetts Institute of Technology*

Jennifer Doudna, *University of California–Berkeley*

Chris Dundas, *University of Georgia*

Mary Dunlop, *Boston University*

Robert Egbert, *Pacific Northwest National Laboratory*

Matthew Fields, *Montana State University*

Michael Guarnieri, *National Laboratory of the Rockies*

Will Harcombe, *University of Minnesota*

Chris Henry*, *Argonne National Laboratory*

Kirsten Hofmockel*, *Pacific Northwest National Laboratory*

David Karig, *Clemson University*

Christopher Lawson, *University of Toronto*

Scott Lenaghan, *University of Tennessee–Knoxville*

Xiaoxia (Nina) Lin, *University of Michigan*

Costas Maranas, *Pennsylvania State University*

Jennifer Martiny, *University of California–Irvine*

Joshua Michener, *Oak Ridge National Laboratory*

Mark Mimee, *University of Chicago*

Vivek Mutalik, *Lawrence Berkeley National Laboratory*

Trent Northen, *Lawrence Berkeley National Laboratory*

Michelle O'Malley*, *University of California–Santa Barbara*

Jennifer Pett-Ridge, *Lawrence Livermore National Laboratory*

Aaron Robinson, *Los Alamos National Laboratory*

Carlotta Ronda, *University of California–Berkeley*

Howard Salis, *The Pennsylvania State University*

Hyun-Seob Song, *University of Nebraska–Lincoln*

Rhona Stuart, *Lawrence Livermore National Laboratory*

Ophelia Venturelli, *Duke University*

Kelly Wrighton, *Colorado State University*

Lingchong You*, *Duke University*

Jizhong Zhou, *University of Oklahoma*

* Writing Team

Appendix C

Breakout Group Assignments

Day 1: Breakout Sessions 1-2

Group 1
Group Leader: Vivek Mutalik

Report Delegate: Chris Henry

Participants: Aaron Robinson, Howard Salis, Rhona Stuart

Group 2
Group Leader: Michaeline Albright

Report Delegate: Costas Maranas

Participants: Greg Bonito, Christopher Lawson, Joshua Michener, Ophelia Venturelli

Group 3
Group Leader: Jennifer Martiny

Report Delegate: Jennifer Pett-Ridge

Participants: Mark Mimee, Hyun-Seob Song, Lingchong You

Group 4
Group Leader: Trent Northen

Report Delegate: Robert Egbert

Participants: Otto Cordero, Melissa Cregger, Scott Lenaghan, Jizhong Zhou

Group 5
Group Leader: Jonathan (Joff) Silberg

Report Delegate: Kelly Wrighton

Participants: Chris Dundas, Matthew Fields, David Karig, Xiaoxia (Nina) Lin

Group 6
Group Leader: Michelle O'Malley

Report Delegate: Kirsten Hofmockel

Participants: Bryn Adams, Michael Guarnieri, Will Harcombe

Day 2: Breakout Sessions 3-5

Group 1
Group Leader: Joshua Michener

Report Delegate: Otto Cordero

Participants: Kirsten Hofmockel, Jennifer Martiny, Howard Salis, Michelle O'Malley

Group 2
Group Leader: Jennifer Pett-Ridge

Report Delegate: Greg Bonito

Participants: Christopher Lawson, Hyun-Seob Song, Lingchong You

Group 3
Group Leader: Chris Dundas

Report Delegate: Matthew Fields

Participants: Michaeline Albright; David Karig; Scott Lenaghan; Carlotta Ronda (Session 5 only)

Group 4
Group Leader: Xiaoxia (Nina) Lin

Report Delegate: Robert Egbert

Participants: Costas Maranas (Sessions 3 and 4 only), Jonathan (Joff) Silberg, Kelly Wrighton, Jizhong Zhou

Group 5
Group Leader: Michael Guarnieri

Report Delegate: Rhona Stuart

Participants: Melissa Cregger, Vivek Mutalik, Aaron Robinson

Group 6
Group Leader: Chris Henry

Report Delegate: Trent Northen

Participants: Will Harcombe, Ophelia Venturelli

Day 3: Breakout Sessions 6-7

Group 1
Group Leader: Kirsten Hofmockel

Report Delegate: Aaron Robinson

Participants: Michaeline Albright, Matthew Fields

Group 2
Group Leader: Rhona Stuart

Report Delegate: Trent Northen

Participants: Greg Bonito, Melissa Cregger, Jennifer Martiny

Group 3
Group Leader: Kelly Wrighton

Report Delegate: Jennifer Pett-Ridge

Participants: Will Harcombe, David Karig, Jizhong Zhou

Group 4
Group Leader: Costas Maranas

Report Delegate: Chris Henry

Participants: Xiaoxia (Nina) Lin, Hyun-Seob Song, Lingchong You

Group 5
Group Leader: Bryn Adams

Report Delegate: Howard Salis

Participants: Michael Guarnieri, Michelle O'Malley

Group 6
Group Leader: Robert Egbert

Report Delegate: Jonathan (Joff) Silberg

Participants: Joshua Michener, Mark Mimee, Carlotta Ronda

Group 7
Group Leader: Ophelia Venturelli

Report Delegate: Chris Dundas

Participants: Christopher Lawson, Scott Lenaghan, Vivek Mutalik

Day 3: Breakout Session 8

Group 1
Group Leader: Joshua Michener

Report Delegate: Jennifer Martiny

Participants: Kirsten Hofmockel, Howard Salis

Group 2
Group Leader: Jennifer Pett-Ridge

Report Delegate: Greg Bonito

Participants: Bryn Adams, Christopher Lawson, Jennifer Martiny, Hyun-Seob Song, Lingchong You

Group 3
Group Leader: Chris Dundas

Report Delegate: Matthew Fields

Participants: Michaeline Albright, David Karig, Scott Lenaghan

Group 4
Group Leader: Xiaoxia (Nina) Lin

Report Delegate: Robert Egbert

Participants: Jonathan (Joff) Silberg, Jizhong Zhou

Group 5
Group Leader: Michael Guarnieri

Report Delegate: Rhona Stuart

Participants: Melissa Cregger, Vivek Mutalik, Aaron Robinson

Group 6
Group Leader: Chris Henry

Report Delegate: Trent Northen

Participants: Will Harcombe, Mark Mimee, Ophelia Venturelli

Appendix D

References

Agena, E., et al. 2024. "Evaluating the Feasibility of Medium-Chain Oleochemical Synthesis Using Microbial Chain Elongation," *Journal of Industrial Microbiology and Biotechnology* **51**, kuae027. DOI:10.1093/jimb/kuae027.

Agoussar, A., and E. Yergeau. 2021. "Engineering the Plant Microbiota in the Context of the Theory of Ecological Communities," *Current Opinion in Biotechnology* **70**, 220–5. DOI:10.1016/j.copbio.2021.06.009.

Ahmad, M., et al. 2016. "Live Cell Monitoring of Glycine Betaine by FRET-based Genetically Encoded Nanosensor," *Biosensors and Bioelectronics* **86**, 169–75. DOI:10.1016/j.bios.2016.06.049.

Alori, E. T., et al. 2017. "Microbial Phosphorus Solubilization and Its Potential for Use in Sustainable Agriculture," *Frontiers in Microbiology* **8**. DOI:10.3389/fmicb.2017.00971.

Angulo, M. T., et al. 2019. "A Theoretical Framework for Controlling Complex Microbial Communities," *Nature Communications* **10**. DOI:10.1038/s41467-019-08890-y.

Archer, E. J., et al. 2012. "Engineered *E. coli* that Detect and Respond to Gut Inflammation through Nitric Oxide Sensing," *ACS Synthetic Biology* **1**, 451–7. DOI:10.1021/sb3000595.

Argiroff, W. A., et al. 2024. "Seasonality and Longer-Term Development Generate Temporal Dynamics in the *Populus* Microbiome," *mSystems* **9**(3). DOI:10.1128/msystems.00886-23.

Arya, S., et al. 2023. "Sparsity of Higher-Order Landscape Interactions Enables Learning and Prediction for Microbiomes," *Proceedings of the National Academy of Sciences of the USA* **120**(48), e2307313120. DOI:10.1073/pnas.2307313120.

Asin-Garcia, E., et al. 2022. "Phosphite Synthetic Auxotrophy as an Effective Biocontainment Strategy for the Industrial Chassis *Pseudomonas putida*," *Microbial Cell Factories* **21**, 156. DOI:10.1186/s12934-022-01883-5.

Asin-Garcia, E., et al. 2024. "A Snapshot of Biomanufacturing and the Need for Enabling Research Infrastructure," *Trends in Biotechnology*. DOI:10.1016/j.tibtech.2024.10.014.

Aufrecht, J. et al. 2022. "Hotspots of Root-Exuded Amino Acids are Created within a Rhizosphere-on-a-Chip," *Lab on a Chip* **22**, 954–63. DOI:10.1039/D1LC00705J.

Baig, Y., et al. 2023. "Autoencoder Neural Networks Enable Low Dimensional Structure Analyses of Microbial Growth Dynamics," *Nature Communications* **14**, 7937. DOI:10.1038/s41467-023-43455-0.

Baranwal, M., et al. 2022. "Recurrent Neural Networks Enable Design of Multifunctional Synthetic Human Gut Microbiome Dynamics," *eLife* **11**, e73870. DOI:10.7554/eLife.73870.

Barnes, E. M., and S. G. Tringe. 2022. "Exploring the Roles of Microbes in Facilitating Plant Adaptation to Climate Change," *The Biochemical Journal* **479**(3), 327–35. DOI:10.1042/BCJ20210793.

Bashiardes, S., et al. 2016. "Use of Metatranscriptomics in Microbiome Research," *Bioinformatics and Biology Insights* **10**, 19–25. DOI:10.4137/BBI.S34610.

Bashyal, S. et al. 2025. "A Plant CLE Peptide and Its Fungal Mimic Promote Arbuscular Mycorrhizal Symbiosis via CRN-Mediated ROS Suppression," *Proceedings of the National Academy of Sciences of the USA,* **122**(16), e2422215122. DOI:10.1073/pnas.2422215122.

Baumgart, L. A., et al. 2021. "Persistence and Plasticity in Bacterial Gene Regulation," *Nature Methods* **18**, 1499–505. DOI:10.1038/s41592-021-01312-2.

Béjà, O., et al. 2000. "Bacterial Rhodopsin: Evidence for a New Type of Phototrophy in the Sea," *Science* **289**(5486), 1902–6. DOI:10.1126/science.289.5486.1902.

Bencurova, E., et al. 2023. "DNA Storage–From Natural Biology to Synthetic Biology," *Computational and Structural Biotechnology Journal* **21**, 1227–35. DOI:10.1016/j.csbj.2023.01.045.

Berg, G., et al. 2020. "Microbiome Definition Re-Visited: Old Concepts and New Challenges," *Microbiome* **8**, 103. DOI:10.1186/s40168-020-00875-0.

Bergeson, A. R., and H. S. Alper. 2024. "Advancing Sustainable Biotechnology Through Protein Engineering," *Trends in Biochemical Science* **49**(11). DOI:10.1016/j.tibs.2024.07.006.

Bharti, R., and D. G. Grimm. 2021. "Current Challenges and Best-Practice Protocols for Microbiome Analysis," *Briefings in Bioinformatics* **22**(1), 178–93. DOI:10.1093/bib/bbz155.

Bhattacharyya, S. S., et al. 2022. "Soil Carbon Sequestration–An Interplay Between Soil Microbial Community and Soil Organic Matter Dynamics," *Science of the Total Environment* **815**, 152928. DOI:10.1016/j.scitotenv.2022.152928.

Birer-Williams, C. M. C., et al. 2021. "SubTap, a Versatile 3D Printed Platform for Eavesdropping on Extracellular Interactions," *mSystems* **6**. DOI:10.1128/msystems.00902-21.

Blagodatskaya, E., and Y. Kuzyakov. 2013. "Active Microorganisms in Soil: Critical Review of Estimation Criteria and Approaches," *Soil Biology and Biochemistry* **67**, 192–211. DOI:10.1016/j.soilbio.2013.08.024.

Blanco-Míguez, A., et al. 2023. "Extending and Improving Metagenomic Taxonomic Profiling with Uncharacterized Species Using MetaPhlAN 4," *Nature Biotechnology* **41**, 1633–44. DOI:10.1038/s41587-023-01688-w.

Blazejewski, T. et al. 2019. "Synthetic Sequence Entanglement Augments Stability and Containment of Genetic Information in Cells," *Science* **365**, 595–8. DOI:10.1126/science.aav5477.

Bloch, S. E., et al. 2020. "Biological Nitrogen Fixation in Maize: Optimizing Nitrogenase Expression in a Root-Associated Diazotroph," *Journal of Experimental Botany* **71**(15), 4591–603. DOI:10.1093/jxb/eraa176.

Boo, A. et al. 2024. "Synthetic Microbe-to-Plant Communication Channels," *Nature Communications* **15**. DOI:10.1038/s41467-024-45897-6.

Borowiec, M. L., et al. 2022. "Deep Learning as a Tool for Ecology and Evolution," *Methods in Ecology and Evolution* **13**(8), 1640–60. DOI:10.1111/2041-210X.13901.

Branchini, B. R., et al. 2015. "An Enhanced Chimeric Firefly Luciferase-Inspired Enzyme for ATP Detection and Bioluminescence Reporter and Imaging Applications," *Analytical Biochemistry* **484**, 148–53. DOI:10.1016/j.ab.2015.05.020.

Britschgi, T. B., and S. J. Giovannoni. 1991. "Phylogenetic Analysis of a Natural Marine Bacterioplankton Population by rRNA Gene Cloning and Sequencing," *Applied and Environmental Microbiology* **57**(6). DOI:10.1128/aem.57.6.1707-1713.1991.

Brophy, J. A. N., et al. 2018. "Engineered Integrative and Conjugative Elements for Efficient and Inducible DNA Transfer to Undomesticated Bacteria," *Nature Microbiology* **3**, 1043–53. DOI:10.1038/s41564-018-0216-5.

Brown, R. W., et al. 2021. "Volatile Organic Compounds (VOCs) Allow Sensitive Differentiation of Biological Soil Quality," *Soil Biology and Biochemistry* **156**. DOI:10.1016/j.soilbio.2021.108187.

Bundy, J. G., et al. 2008. "Environmental Metabolomics: A Critical Review and Future Perspectives," *Metabolomics* **5**, 3–21. DOI:10.1007/s11306-008-0152-0.

Camargo, A. P., et al. 2024. "IMG/PR: A Database of Plasmids from Genomes and Metagenomes with Rich Annotations and Metadata," *Nucleic Acids Research* **52**(1), D164–73. DOI:10.1093/nar/gkad964.

Capdevila-Cortada, M. 2019. "Electrifying the Haber-Bosch," *Nature Catalysis* **2**, 1055. DOI:10.1038/s41929-019-0414-4.

Carper, D. L., et al. 2021. "Cultivating the Bacterial Microbiota of Populus Roots," *mSystems* **6**(3). DOI:10.1128/msystems.01306-20.

Carrell, A. A., et al. 2022. "Habitat-Adapted Microbial Communities Mediate *Sphagnum* Peatmoss Resilience to Warming," *New Phytologist* **234**(6), 2111–25. DOI:10.1111/nph.18072.

Carruthers, D. N., and T. S. Lee. 2022. "Translating Advances in Microbial Bioproduction to Sustainable Biotechnology," *Frontiers in Bioengineering and Biotechnology* **10**. DOI:10.3389/fbioe.2022.968437.

Cassiano, M. H. A., and R. Silva-Rocha. 2020. "Benchmarking Bacterial Promotor Prediction Tools: Potentialities and Limitations," *mSystems* **5**(4). DOI:10.1128/msystems.00439-20.

Cayron, J., et al. 2017. "Pushing the Limits of Nickel Detection to Nanomolar Range Using a Set of Engineered Bioluminescent *Escherichia coli*," *Environmental Science and Pollution Research* **24**, 4–14. DOI:10.1007/s11356-015-5580-6.

Charbonneau, M. R., et al. 2020. "Developing a New Class of Engineered Live Bacterial Therapeutics to Treat Human Diseases," *Nature Communications* **11**, 1738. DOI:10.1038/s41467-020-15508-1.

Chemla, Y., et al. 2022. "Parallel Engineering of Environmental Bacteria and Performance Over Years Under Jungle-Simulated Conditions," *PLOS One* **17**(12), e0278471. DOI:10.1371/journal.pone.0278471.

Chemla, Y., et al. 2025. "Hyperspectral Reporters for Long-Distance and Wide-Area Detection of Gene Expression in Living Bacteria," *Nature Biotechnology*. DOI:10.1038/s41587-025-02622-y.

Chen, Y.-J., et al. 2013. "Characterization of 582 Natural and Synthetic Terminators and Quantification of their Design Constraints," *Nature Methods* **10**, 659–64. DOI:10.1038/nmeth.2515.

Chen, Y., et al. 2015. "Emergent Genetic Oscillations in a Synthetic Microbial Consortium," *Science* **349**(6251), 986–9. DOI:10.1126/science.aaa3794.

Cheng, Y. Y., et al. 2001. "Programming Bacteria for Multiplexed DNA Detection," *Nature Communications* **14**. DOI:10.1038/s41467-023-37582-x.

Chilton, M.-D., et al. 1977. "Stable Incorporation of Plasmid DNA into Higher Plant Cells: the Molecular Basis of Crown Gall Tumorigenesis," *Cell Press* **11**(2). DOI:10.1016/0092-8674(77)90043-5.

Chuckran, P. F., et al. 2024. "Metatranscriptomes of California Grassland Soil Microbial Communities in Response to Rewetting," *Microbial Resource Announcements* **13**(6). DOI:10.1128/mra.00322-24.

Coale, T. H., et al. 2024. "Nitrogen-Fixing Organelle in a Marine Alga," *Microbiology* **384**(6692), 217–22. DOI:10.1126/science.adk1075.

Colarusso, A. V., et al. 2021. "Computational Modeling of Metabolism in Microbial Communities on a Genome-Scale," *Current Opinion in Systems Biology* **26**, 46–57. DOI:10.1016/j.coisb.2021.04.001.

Colica, G., et al. 2014. "Microbial Secreted Exopolysaccharides Affect the Hydrological Behavior of Induced Biological Soil Crusts in Desert Sandy Soils," *Soil Biology and* Biochemistry **68**, 62–70. DOI:10.1016/j.soilbio.2013.09.017.

Connon, S. A., and S. J. Giovannoni. 2002. "High-Throughput Methods for Culturing Microorganisms in Very-Low-Nutrient Media Yield Diverse New Marine Isolates," *Applied and Environmental Microbiology* **68**(8). DOI:10.1128/AEM.68.8.3878-3885.2002.

Cottom, J. W., et al. 2024. "A Local-to-Global Emissions Inventory of Macroplastic Pollution," *Nature* **633**, 101–8. DOI:10.1038/s41586-024-07758-6.

Cross, K. L., et al. 2019. "Targeted Isolation and Cultivation of Uncultivated Bacteria by Reverse Genomics," *Nature Biotechnology* **37**, 1314–21. DOI:10.1038/s41587-019-0260-6.

Custódio, C. A., and J. F. Mano. 2016. "Cell Surface Engineering to Control Cellular Interactions," *ChemNanoMat* **2**(5), 376–84. DOI:10.1002/cnma.201600047.

Daniel, R. 2005. "The Metagenomics of Soil," *Nature Reviews Microbiology* **3**, 470–78. DOI:10.1038/nrmicro1160.

Darmon E., and D. R. F. Leach. 2014. "Bacterial Genome Instability," *Microbiology and Molecular Biology Reviews* **78**(1), 1–39. DOI:10.1128/MMBR.00035-13.

de Groot, M. J. A., et al. 1998. "*Agrobacterium tumefaciens*–Mediated Transformation of Filamentous Fungi," *Nature Biotechnology* **16**, 839–42. DOI:10.1038/nbt0998-839.

DeAngelis, K. M., et al. 2005. "Two Novel Bacterial Biosensors for Detection of Nitrate Availability in the Rhizosphere," *Applied and Environmental Microbiology* **71**, 8537–47. DOI:10.1128/AEM.71.12.8537-8547.2005.

Del Valle, I., et al. 2021. "Translating New Synthetic Biology Advances for Biosensing Into the Earth and Environmental Sciences," *Frontiers in Microbiology* **11**. DOI:10.3389/fmicb.2020.618373.

Del Valle, I., et al. 2022. "Artificial Soils Reveal Individual Factor Controls on Microbial Processes," *Environmental Microbiology* **7**(4). DOI:10.1128/msystems.00301-22.

Delettre, O. 2021. "Identity of Ecological Systems and the Meaning of Resilience," *Journal of Ecology* **109**(9), 3147–56. DOI:10.1111/1365-2745.13655.

Delgado-Baquerizo, M., et al. 2025. "Integrating Ecological and Evolutionary Frameworks for SynCom Success," *New Phytologist* **246**(5), 1922–33. DOI:10.1111/nph.70112.

Dove, N. C., et al. 2021. "Assembly of the *Populus* Microbiome is Temporally Dynamic and Determined by Selective and Stochastic Factors," *mSystems* **6**(3). DOI:10.1128/msphere.01316-20.

Dunivin, T. K., et al. 2019. "RefSoil+: a Reference Database for Genes and Traits of Soil Plasmids," *mSystems* **4**(1). DOI:10.1128/msystems.00349-18.

Elmore, J. R., et al. 2023. "High-Throughput Genetic Engineering of Nonmodel and Undomesticated Bacteria via Iterative Site-Specific Genome Integration," *Science Advances* **9**(10). DOI:10.1126/sciadv.ade1285.

Elsen, S., et al. 2003. "Interaction between the H2 Sensor HupUV and the Histidine Kinase HupT Controls HupSL Hydrogenase Synthesis in *Rhodobacter capsulatus*," *Journal of Bacteriology* **185**, 7111–9. DOI:10.1128/JB.185.24.7111-7119.2003.

Essington, E. A., et al. 2024. "An Autonomous Microbial Sensor Enables Long-Term Detection of TNT Explosive in Natural Soil," *Nature Communications* **15**, 10471. DOI:10.1038/s41467-024-54866-y.

Fackler, N., et al. 2021. "Stepping on the Gas to a Circular Economy: Accelerating Development of Carbon-Negative Chemical Production from Gas Fermentation," *Annual Review of Chemical and Biomolecular Engineering* **12**, 439–70. DOI:10.1146/annurev-chembioeng-120120-021122.

Farrar, K., et al. 2014. "Understanding and Engineering Beneficial Plant–Microbe Interactions: Plant Growth Promotion in Energy Crops," *Plant Biotechnology Journal* **12**(9), 1193–206. DOI:10.1111/pbi.12279.

Fatma, Z., et al. 2020. "Recent Advances in Domesticating Non-Model Microorganisms," *Biotechnology Progress* **36**(5), e3008. DOI:10.1002/btpr.3008.

Feklístov, A., et al. 2014. "Bacteria Sigma Factors: A Historical, Structural, and Genomic Perspective," *Annual Review of Microbiology* **68**, 357–76. DOI:10.1146/annurev-micro-092412-155737.

Feyissa, B. A., et al. 2024. "An Orphan Gene *BOOSTER* Enhances Photosynthetic Efficiency and Plant Productivity," *Developmental Cell* **60**(5). DOI:10.1016/j.devcel.2024.11.002.

Flemming, H.-C., and S. Wuertz. 2019. "Bacteria and Archaea on Earth and their Abundance in Biofilms," *Nature Reviews Microbiology* **17**, 247–260. DOI:10.1038/s41579-019-0158-9.

Franklin, J., et al. 2016. "Global Change and Terrestrial Plant Community Dynamics," *Proceedings of the National Academy of Sciences of the USA* **113**(14), 3725–34. DOI:10.1073/pnas.1519911113.

Gabriel, S. A., et al. 2015. "50-Plus Years of Fungal Viruses," *Virology* **479-80**, 356–68. DOI:10.1016/j.virol.2015.02.034.

Gach, P. C., et al. 2016. "A Droplet Microfluidic Platform for Automating Genetic Engineering," *ACS Synthetic Biology* **5**(5). DOI:10.1021/acssynbio.6b00011.

Ganesh, I., et al. 2020. "Gas-Sensing Transcriptional Regulators," *Biotechnology Journal* **15**, 1900345. DOI:10.1002/biot.201900345.

Gao, R., and Lynn, D. G. 2005. "Environmental pH Sensing: Resolving the VirA/VirG Two-Component System Inputs for Agrobacterium Pathogenesis," *Journal of Bacteriology* **187**, 2182–9. DOI:10.1128/JB.187.6.2182-2189.2005.

Gates, A. D., et al. 2023. "A *Streptomyces* Consortium Contributes Distinct Microbial Interactions During *Arabidopsis thaliana* Microbiome Assembly," *Phytobiomes Journal* **7**(4), 2471–906. DOI:10.1094/PBIOMES-11-22-0081-R.

Ghosh, T., et al. 2023. "Application of X-Ray Computed Tomography in Soil and Plant—A Review," *Frontiers in Environmental Science* **11**. DOI:10.3389/fenvs.2023.1216630.

Gillett, D.L., et al. 2025. "A Roadmap to Understanding and Anticipating Microbial Gene Transfer in Soil Communities," *Microbiology and Molecular Biology Reviews* e00225-24. DOI:10.1128/mmbr.00225-24.

Graham, E. B., et al. 2024. "A Global Atlas of Soil Viruses Reveals Unexplored Biodiversity and Potential Biogeochemical Impacts," *Nature Microbiology* **9**, 1873–83. DOI:10.1038/s41564-024-01686-x.

Gralka, M., et al. 2020. "Trophic Interactions and the Drivers of Microbial Community Assembly," *Current Biology* **30**(19), R1176–88. DOI:10.1016/j.cub.2020.08.007.

Grandel, N. E., et al. 2025. "Long-Term Homeostasis in Microbial Consortia via Auxotrophic Cross-Feeding." Preprint, *bioRxiv*, January 9, 2025. DOI:10.1101/2025.01.08.631749.

Griffin, C., et al. 2024. "Engineering Plant–Microbe Communication for Plant Nutrient Use Efficiency," *Current Opinion in Biotechnology* **88**, 103150. DOI:10.1016/j.copbio.2024.103150.

Großkopf, T., and O. S. Soyer. 2016. "Microbial Diversity Arising from Thermodynamic Constraints," *The ISME Journal* **10**(11), 2725–33. DOI:10.1038/ismej.2016.49.

Gu, C., et al. 2019. "Current Status and Applications of Genome-Scale Metabolic Models," *Genome Biology* **20**, 121. DOI:10.1186/S13059-019-1730-3.

Gupta, K., et al. 2024. "EcoFAB 3.0: A Sterile System for Studying Sorghum that Replicates Previous Field and Greenhouse Observations," *Frontiers in Plant Science* **15**. DOI:10.3389/fpls.2024.1440728.

Hagiwara, D., et al. 2004. "A Genome-wide View of the *Escherichia coli* BasS–BasR Two-Component System Implicated in Iron-Responses," *Bioscience, Biotechnology, and Biochemistry* **68**, 1758–67. DOI:10.1271/bbb.68.1758.

Haldar, S., and S. Sengupta. 2015. "Plant–Microbe Cross-Talk in the Rhizosphere: Insight and Biogeochemical Potential," *The Open Microbiology Journal* **9**, 1–7. DOI:10.2174/1874285801509010001.

Hanson, P. J., and A. P. Walker. 2019. "Advancing Global Change Biology Through Experimental Manipulations: Where Have We Been and Where Might We Go?" *Global Change Biology* **26**(1), 287–99. DOI:10.1111/gcb.14894.

Harcombe, W. R., et al. 2014. "Metabolic Resource in Individual Microbes Determines Ecosystem Interactions and Spatial Dynamics," *Cell Reports* **7**(4), 1104–15. DOI:10.1016/j.celrep.2014.03.070.

He, L., et al. 2020. "Global Biogeography of Fungal and Bacterial Biomass Carbon in Topsoil," *Soil Biology and Biochemistry* **151**. DOI:10.1016/j.soilbio.2020.108024.

Helber, N., and N. Requena. 2007. "Expression of the Fluorescence Markers DsRed and GFP Fused to a Nuclear Localization Signal in the Arbuscular Mycorrhizal Fungus *Glomus intraradices*," *New Phytologist* **177**(2). DOI:10.1111/j.1469-8137.2007.02257.x.

Herron, P. M., et al. 2010. "Micro-Scale Water Potential Gradients Visualized in Soil around Plant Root Tips using Microbiosensors," *Plant, Cell & Environment* **33**, 199–210. DOI:10.1111/j.1365-3040.2009.02070.x.

Hoffman, S. M., et al. 2022. "Optogenetics Illuminates Applications in Microbial Engineering," *Annual Review of Chemical and Biomolecular Engineering* **13**, 373–403. 10.1146/annurev-chembioeng-092120-092340.

Højberg, O., et al. 1999. "Oxygen-Sensing Reporter Strain of *Pseudomonas fluorescens* for Monitoring the Distribution of Low-oxygen Habitats in Soil," *Applied and Environmental Microbiology* **65**, 4085–93. DOI:10.1128/AEM.65.9.4085-4093.1999.

Hol, W. H. G., et al. 2015. "Context Dependency and Saturating Effects of Loss of Rare Soil Microbes on Plant Productivity," *Frontiers in Plant Science* **6**. DOI:10.3389/fpls.2015.00485.

Hong, Y., et al. 2025. "A Novel Capacity Coupled Contactless Conductivity Detection (C^4D) Microfluidic Chip Integrated 3D Microelectrodes for On-Site Determination of Soil Nutrients," *Computers and Electronics in Agriculture* **219**. DOI:10.1016/j.compag.2024.108829.

Hoynes-O'Connor, A., et al. 2017. "Enabling Complex Genetic Circuits to Respond to Extrinsic Environmental Signals," *Biotechnology and Bioengineering* **114**, 1626–31. DOI:10.1002/bit.26279.

Hren, A., et al. 2025. "High-Density CRISPRi Screens Reveal Diverse Routes to Improved Acclimation in Cyanobacteria," *Proceedings of the National Academy of Sciences of the United States of America* **122**(12). DOI:10.1073/pnas.2412625122.

Hsu, R. H., et al. 2019. "Microbial Interaction Network Inference in Microfluidic Droplets," *Cell Systems* **9**(3), 229–42. DOI:10.1016/j.cels.2019.06.008.

Huang, D., et al. 2012. "A Genetic Bistable Switch Utilizing Nonlinear Protein Degradation," *Journal of Biological Engineering* **6**, 9. DOI:10.1186/1754-1611-6-9.

Huang, X., et al. 2017. "*Bacillus subtilis* MntR Coordinates the Transcriptional Regulation of Manganese Uptake and Efflux Systems," *Molecular Microbiology* **103**(2), 253–68. DOI:10.1111/mmi.13554.

Huang, Y., et al. 2023. "High-Throughput Microbial Culturomics using Automation and Machine Learning," *Nature Biotechnology* **41**, 1424–33. DOI:10.1038/s41587-023-01674-2.

Idnurm, A., et al. 2017. "A Silver Bullet in a Golden Age of Functional Genomics: the Impact of *Agrobacterium*-Mediated Transformation of Fungi," *Fungal Biology and Biotechnology* **4**, 6. DOI:10.1186/s40694-017-0035-0.

Indoria, A. K., et al. 2020. "Ch. 18: Hydraulic Properties of Soil Under Warming Climate." In *Climate Change and Soil Interactions*, 473–508. Eds. Pietrzykowksi, M., and M. N. V. Prasad. Elsevier, Netherlands. DOI:10.1016/B978-0-12-818032-7.00018-7.

Jakočiūnas, T., et al. 2016. "CRISPR/Cas9 Advances Engineering of Microbial Cell Factories," *Metabolic Engineering* **34**, 44–59. DOI:10.1016/j.ymben.2015.12.003.

Jang, Y.-J., et al. 2017. "Complementary Cell-Free Translational Assay for Quantification of Amino Acids," *Analytical Chemistry* **89**, 9638–42. DOI:10.1021/acs.analchem.7b01956.

Jansson, J. K., and K. S. Hofmockel. 2018. "The Soil Microbiome—From Metagenomics to Metaphenomics," *Current Opinion in Microbiology* **43**, 162–68. DOI:10.1016/j.mib.2018.01.013.

Jilling, A., et al. 2018. "Minerals in the Rhizosphere: Overlooked Mediators of Soil Nitrogen Availability to Plants and Microbes," *Biogeochemistry* **139**, 103–22. DOI:10.1007/s10533-018-0459-5.

Johns, N. I., et al. 2016. "Principles for Designing Synthetic Microbial Communities," *Current Opinion in Microbiology* **31**, 146–53. DOI:10.1016/j.mib.2016.03.010.

Jones, E. M., et al. 2024. "Synthetic Microbiology in Sustainability Applications," *Nature Reviews Microbiology* **22**(6), 345–59. DOI:10.1038/s41579-023-01007-9.

Jumper, J., et al. 2021. "Highly Accurate Protein Structure Prediction with AlphaFold," *Nature* **596**, 583–9. DOI:10.1038/s41586-021-03819-2.

Kalvapalle, P. B., et al. 2025. "Information Storage Across a Microbial Community Using Universal RNA Barcoding," *Nature Biotechnology*. DOI:10.1038/s41587-025-02593-0.

Karkaria, B. D., et al. 2021. "Automated Design of Synthetic Microbial Communities," *Nature Communications* **12**, 672. DOI:10.1038/s41467-020-20756-2.

Kawai, S., et al. 2010. "Transformation of *Saccharomyces cerevisiae* and Other Fungi: Methods and Possible Underlying Mechanism," *Bioengineered Bugs* **1**(6), 395–403. DOI:10.4161/bbug.1.6.13257.

Kaza, S., et al. 2018. *What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050.* World Bank Publications. hdl.handle.net/10986/30317

Keenor, S. G., et al. 2021. "Capturing a Soil Carbon Economy," *Royal Society Open Science* **8**(4), 202305. DOI:10.1098/rsos.202305.

Kehe, J., et al. 2019. "Massively Parallel Screening of Synthetic Microbial Communities," *Proceedings of the National Academy of Sciences of the USA* **116**(26), 12804–9. DOI:10.1073/pnas.1900102116.

Kelsic, E. D. et al. 2015. "Counteraction of Antibiotic Production and Degradation Stabilizes Microbial Communities," *Nature* **521**, 516–19. DOI:10.1038/nature14485.

Kertesz, V., and J. F. Cahill. 2021. "Spatially Resolved Absolute Quantitation in Thin Tissue by Mass Spectrometry," *Analytical and Bioanalytical Chemistry* **413**, 2619–36. DOI:10.1007/s00216-020-02964-3.

Knott, G. J., and J. A. Doudna. 2018. "CRISPR-Cas Guides the Future of Genetic Engineering," *Science* **361**(6405), 866–9. DOI:10.1126/science.aat5011.

Ko, Y.-S., et al. 2020. "Tools and Strategies of Systems Metabolic Engineering for the Development of Microbial Cell Factories for Chemical Production," *Chemical Society Reviews* **49**, 4615–36. DOI:10.1039/D0CS00155D.

Kong, W., et al. 2018. "Designing Microbial Consortia with Defined Social Interactions," *Nature Chemical Biology* **14**, 821–9. DOI:10.1038/s41589-018-0091-7.

Koonin, E. V., et al. 2001. "Horizontal Gene Transfer in Prokaryotes: Quantification and Classification," *Annual Review of Microbiology* **55**, 709–42. DOI:10.1146/annurev.micro.55.1.709.

Kreamer, N. N. K., et al. 2012. "BqsR/BqsS Constitute a Two-Component System that Senses Extracellular Fe(II) in *Pseudomonas aeruginosa*," *Journal of Bacteriology* **194**, 1195–204. DOI:10.1128/JB.05634-11.

Kreamer, N. N., et al. 2015. "The Ferrous Iron-Responsive BqsRS Two-Component System Activates Genes that Promote Cationic Stress Tolerance," *mBio* **6**, e02549. DOI:10.1128/mBio.02549-14.

Kumar, R. K., and K. R. Foster. 2022. "3D Printing of Microbial Communities," *Microbial Technology* **16**(3), 489–93. DOI:10.1111/1751-7915.14168.

Kuzyakov, Y., and E. Blagodatskaya. 2015. "Microbial Hotspots and Hot Moments in Soil: Concept & Review," *Soil Biology and Biochemistry* **83**, 184–99. DOI:10.1016/j.soilbio.2015.01.025.

Lacey, R. F., and Binder, B. 2016. "Ethylene Regulates the Physiology of the Cyanobacterium *Synechocystis* sp. PCC 6803 via an Ethylene Receptor," *Plant Physiology* **171**, 2798–809. DOI:10.1104/pp.16.00602.

Lafferty, K. D., et al. 2015. "A General Consumer-Resource Population Model," *Science* **349**(6250), 854–7. DOI:10.1126/science.aaa6224.

LaFleur, T. L., et al. 2022. "Automated Model-Predictive Design of Synthetic Promoters to Control Transcriptional Profiles in Bacteria," *Nature Communications* **13**, 5159. DOI:10.1038/s41467-022-32829-5.

Langdon, A., et al. 2016. "The Effects of Antibiotics on the Microbiome throughout Development and Alternative Approaches for Therapeutic Modulation," *Genomic Medicine* **8**, 39. DOI:10.1186/s13073-016-0294-z.

Lankiewicz, T. S., et al. 2022. "Enzyme Discovery in Anaerobic Fungi (Neocallimastigomycetes) Enables Lignocellulosic Biorefinery Innovation," *Microbiology and Molecular Biology Reviews* **86**(4). DOI:10.1128/mmbr.00041-22.

Larson, M .H., et al. 2013. "CRISPR Interference (CRISPRi) for Sequence-Specific Control of Gene Expression," *Nature Protocols* **8**, 2180–96. DOI:10.1038/nprot.2013.132.

Lazar, J. T., and J. J. Tabor. 2021. "Bacterial Two-Component Systems as Sensors for Synthetic Biology Applications," *Current Opinion in Systems Biology* **28**, 100398. DOI:10.1016/j.coisb.2021.100398.

Ledesma-Dominguez, L., et al. 2024. "DeepReg: A Deep Learning Hybrid Model for Predicting Transcription Factors in Eukaryotic and Prokaryotic Genomes," *Scientific Reports* **14**, 9155. DOI:10.1038/s41598-024-59487-5.

Lee, J. W., et al. 2018. "Next-Generation Biocontainment Systems for Engineered Organisms," *Nature Chemical Biology* **14**, 530–7. DOI:10.1038/s41589-018-0056-x.

Lee, S. J., and M. Rho. 2022. "Multimodal Deep Learning Applied to Classify Healthy and Disease States of Human Microbiome," *Scientific Reports* **12**. DOI:10.1038/s41598-022-04773-3.

Leggieri, P. A., et al. 2021. "Integrating Systems and Synthetic Biology to Understand and Engineer Microbiomes," *Annual Review of Biomedical Engineering* **23**, 169–201. DOI:10.1146/annurev-bioeng-082120- 022836.

Lerminiaux, N. A., and A. D. S. Cameron. 2019. "Horizontal Transfer of Antibiotic Resistance Genes in Clinical Environments," *Canadian Journal of Microbiology* **65**(1), 34 –44. DOI:10.1139/cjm-2018-0275.

Leung, K., et al. 2012. "A Programmable Droplet-Based Microfluidic Device Applied to Multiparameter Analysis of Single Microbes and Microbial Communities," *Proceedings of the National Academy of Sciences of the USA* **109**(20), 7665–70. DOI:10.1073/pnas.1106752109.

Li, P.-S., et al. 2014. "Construction and Optimization of a *Pseudomonas putida* Whole-Cell Bioreporter for detection of Bioavailable Copper," *Biotechnology Letters* **36**, 761–6. DOI:10.1007/s10529-013-1420-2.

Lim. J., et al. 2021. "Ammonia and Nitric Acid Demands for Fertilizer Use in 2050," *ACS Energy Letters* **6**(10). DOI:10.1021/acsenergylett.1c01614.

Liñares-Blanco, J., et al. 2022. "Machine Learning Based Microbiome Signature to Predict Inflammatory Bowel Disease Subtypes," *Frontiers in Microbiology* **13**. DOI:10.3389/fmicb.2022.872671.

Ling, H., et al. 2014. "Microbial Tolerance Engineering Toward Biogeochemical Production: From Lignocellulose to Products," *Current Opinion in Biotechnology* **29**, 99–106. DOI:10.1016/j.copbio.2014.03.005.

Liu, F., et al. 2024. "Uneven Distribution of Prokaryote-Derived Horizontal Gene Transfer in Fungi: A Lifestyle-Dependent Phenomenon," *mBio* **16**(1). DOI:10.1128/mbio.02855-24.

Liu, P., et al. 2012. "Construction and Application of a Zinc-Specific Biosensor for Assessing the Immobilization and Bioavailability of Zinc in Different Soils," *Environmental Pollution* **164**, 66–72. DOI:10.1016/j.envpol.2012.01.023.

Liu, Y., et al. 2024. "A Small Secreted Protein Serves as a Plant-Derived Effector Mediating Symbiosis between *Populus* and *Laccaria bicolor*," *Horticulture Research* **11**(10). DOI:10.1093/hr/uhae232.

Lorenz, M. G., and W. Wackernagel. 1994. "Bacterial Gene Transfer by Natural Genetic Transformation in the Environment," *Microbiological Reviews* **58**(3). DOI:10.1128/mr.58.3.563-602.1994.

Louca, S. 2018. "Function and Functional Redundancy in Microbial Systems," *Nature Ecology and Evolution* **2**(6), 936–43. DOI:10.1038/s41559-018-0519-1.

Louca, S., et al. 2016. "High Taxonomic Variability Despite Stable Functional Structure Across Microbial Communities," *Nature Ecology & Evolution* **1**, 0015. DOI:10.1038/s41559-016-0015.

Lu, B., et al. 2020. "Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture," *Remote Sensing* **12**(16), 2659. DOI:10.3390/rs12162659.

Lu, H., et al. 2022. "Machine Learning-Aided Engineering of Hydrolases for PET Depolymerization," *Nature* **602**, 663–7. DOI:10.1038/s41586-022-04599-z.

Lu, J., et al. "Decoding Pattern Formation Rules by Integrating Mechanistic Modeling and Deep Learning," Preprint, *bioRxiv*, September 2, 2024. DOI:10.1101/2024.09.02.610872.

MacLeod, M., et al. 2021. "The Global Threat from Plastic Pollution," *Science* **373**(6550), 61–5. DOI:10.1126/science.abg5433.

Makarova, K. S., et al. 2013. "Comparative Genomics of Defense Systems in Archaea and Bacteria," *Nucleic Acids Research* **41**(8), 4360–77. DOI:10.1093/nar/gkt157.

Malci, K., et al. 2022. "Standardization of Synthetic Biology Tools and Assembly Methods for *Saccharomyces cerevisiae* and Emerging Yeast Species," *ACS Synthetic Biology* **11**(8). DOI:10.1021/acssynbio.1c00442.

Marcos-Zambrano, L. J., et al. 2021. "Applications of Machine Learning in Human Microbiome Studies: A Review on Feature Selection, Biomarker Identification, Disease Predication and Treatment," *Frontiers in Microbiology* **12**. DOI:10.3389/fmicb.2021.634511.

Martinez-Feria, R., et al. 2024. "Genetic Remodeling of Soil Diazotrophs Enables Partial Replacement of Synthetic Nitrogen Fertilizer with Biological Nitrogen Fixation in Maize," *Scientific Reports* **14**. DOI:10.1038/s41598-024-78243-3.

Mathieu, A., et al. 2022. "Machine Learning and Deep Learning Applications in Metagenomic Taxonomy and Functional Annotation," *Frontiers in Microbiology* **13.** DOI:10.3389/fmicb.2022.811495.

McBee, R. M., et al. 2021. "Engineering Living and Regenerative Fungal–Bacterial Biocomposite Structures," *Nature Materials* **21**, 471–8. DOI:10.1038/s41563-021-01123-y.

McClure, R., et al. 2020. "Development and Analysis of a Stable, Reduced Complexity Model Soil Microbiome," *Frontiers in Microbiology* **11**. DOI:10.3389/fmicb.2020.01987.

McClure, R., et al. 2022. "Interaction Networks are Driven by Community-Responsive Phenotypes in a Chitin-Degrading Consortium of Soil Microbes," *mSystems* **7**(5). DOI:10.1128/msystems.00372-22.

Mehrotra, P. 2016. "Biosensors and Their Applications–A Review," *Journal of Oral Biology and Craniofacial* Research **6**(2), 153–9. DOI:10.1016/j.jobcr.2015.12.002.

Meredith, L. K., and M. M. Tfaily. 2022. "Capturing the Microbial Volatilome: An Oft Overlooked 'Ome'," *Trends in Microbiology* **30**(7), 622–31. DOI:10.1016/j.tim.2021.12.004.

Metcalf, W. W., et al. 1997. "A Genetic System for Archaea of the genus *Methanosarcina*: Liposome-Mediated Transformation and Construction of Shuttle Vectors," *Proceedings of the National Academy of Sciences the USA* **94**(6), 2626–31. DOI:10.1073/pnas.94.6.2626.

Meyer, A. J., et al. 2019. "*Escherichia coli* "Marionette" Strains with 12 Highly Optimized Small-Molecule Sensors," *Nature Chemical Biology* **15**, 196–204. DOI:10.1038/s41589-018-0168-3.

Moriguchi, K., et al. 2013. "Transkingdom Genetic Transfer from *Escherichia col*i to *Saccharomyces cerevisiae* as a Simple Gene Introduction Tool," *Applied and Environmental Microbiology* **79**(14). DOI:10.1128/AEM.00770-13.

Moser, F., et al. 2013. "Genetic Sensor for Strong Methylating Compounds," *ACS Synthetic Biology* **2**(10), 614–24. DOI:10.1021/sb400086p.

Nagarajan, K., C. Ni, and T. Lu. 2022. "Agent-Based Modeling of Microbial Communities," *ACS Synthetic Biology* **11**(11), 3564–74. DOI:10.1021/acssynbio.2c00411.

Nagel, K. A., et al. 2012. "GROWSCREEN-Rhizo is a Novel Phenotyping Robot Enabling Simultaneous Measurements of Root and Shoot Growth for Plants in Soil-Filled Rhizotrons," *Functional Plant Biology* **39**(11). DOI:10.1071/FP12023.

NAPPN. 2025. "North American Plant Phenotyping Network." Accessed August 2025. North American Plant Phenotyping Network. plantphenotyping.org

Nguyen, P. Q., et al. 2021. "Wearable Materials with Embedded Synthetic Biology Sensors for Biomolecule Detection," *Nature Biotechnology* **39**, 1366–74. DOI:10.1038/s41587-021-00950-3.

Ni, C., and K. L. J. Prather. 2024. "Consistent Biosynthesis of D-Glycerate from Variable Mixed Substrates," *Metabolic Engineering* **82**, 41–48. DOI:10.1016/j.ymben.2024.01.001.

Northen, T. R., et al. 2024. "Community Standards and Future Opportunities for Synthetic Communities in Plant–Microbiota Research," *Nature Microbiology* **9**, 2774–84. DOI:10.1038/s41564-024-01833-4.

NSF. 2025. "NSF Invests in BioFoundries to Drive Advances Across Science and Engineering." Accessed August 2025. National Science Foundation. nsf.gov/news/nsf-invests-biofoundries-drive-advances-across-science

Nyerges, A., et al. 2023. "A Swapped Genetic Code Prevents Viral Infections and Gene Transfer," *Nature* **615**, 720–7. DOI:10.1038/s41586-023-05824-z.

Oh, M., and L. Zhang. 2020. "DeepMicro: Deep Representation Learning for Disease Prediction Based on Microbiome Data," *Scientific Reports* **10**, 6026. DOI:10.1038/s41598-020-63159-5.

ORNL. 2025. "Welcome to the Advanced Plant Phenotyping Laboratory (APPL)." Accessed August 2025. Oak Ridge National Laboratory. www.ornl.gov/appl

Park, H., et al. 2023. "A Bacterial Sensor Taxonomy Across Earth Ecosystems for Machine Learning Applications," *mSystems* **9**(1). DOI:10.1128/msystems.00026-23.

Peixoto, R. S., et al. 2022. "Harnessing the Microbiome to Prevent Global Biodiversity Loss," *Nature Microbiology* **7**, 1726–35. DOI:10.1038/s41564-022-01173-1.

Peterson, D., et al. 2021. "Advances in Agrobacterium *Transformation and Vector Design Result in High-Frequency Targeted Gene Insertion in Maize," Plant Biotechnology Journal* **19**(10), 2000–10. DOI:10.1111/pbi.13613.

Pflüger, T., et al. 2018. "Signaling Ammonium Across Membranes through an Ammonium Sensor Histidine Kinase," *Nature Communications* **9**,164. DOI:10.1038/s41467-017-02637-3.

Philippot, L., et al. 2013. "Loss in Microbial Diversity Affects Nitrogen Cycling in Soil," *The ISME Journal* **7**, 1609–19. DOI:10.1038/ismej.2013.34.

Pichler, M., and F. Hartig. 2023. "Machine Learning and Deep Learning–A Review for Ecologists," *Methods in Ecology and Evolution* **14**(4), 994–1016. DOI:10.1111/2041-210X.14061.

Piraner, D. I., et al. 2017. "Tunable Thermal Bioswitches for *in vivo* Control of Microbial Therapeutics," *Nature Chemical Biology* **13**, 75–80. DOI:10.1038/nchembio.2233.

Prescott, S. L. 2017. "History of Medicine: Origin of the Term Microbiome and Why It Matters," *Human Microbiome Journal* **4**, 24–25. DOI:10.1016/j.humic.2017.05.004.

Qadir, M., and S. Schubert. 2002. "Degradation Processes and Nutrient Constraints in Sodic Soils," *Land Degradation & Development* **13**(4), 275–94. DOI:10.1002/ldr.504.

Quast, C., et al. 2012. "The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools," *Nucleic Acids Research* **41**(D1), 590–6. DOI:10.1093/nar/gks1219.

Quesada, C. A., et al. 2010. "Variations in Chemical and Physical Properties of Amazon Forest Soils in Relation to Their Genesis," *Biogeosciences* **7**, 1515–41. DOI:10.5194/bg-7-1515-2010.

Ridenhour, B. J., et al. 2017. "Modeling Time-Series Data from Microbial Communities," *The ISME Journal* **11**(11), 2526–37. DOI:10.1038/ismej.2017.107.

Riley, L. A., and A. M. Guss. 2021. "Approaches to Genetic Tool Development for Rapid Domestication of Non-Model Microorganisms," *Biotechnology for Biofuels* **14**, 30. DOI:10.1186/s13068-020-01872-z.

Roberts, G. P., et al. 2001. "CooA: A Heme-Containing Regulatory Protein that Serves as a Specific Sensor of Both Carbon Monoxide and Redox State," *Progress in Nucleic Acid Research and Molecular Biology* **67**, 35–63. DOI:10.1016/S0079-6603(01)67024-7.

Rodenburg, S. Y. A., et al. 2021. "Uncovering the Role of Metabolism in Oomycete–Host Interactions Using Genome-Scale Metabolic Models," *Frontiers in Microbiology* **12**. DOI:10.3389/fmicb.2021.748178/full.

Rodríguez-Ramos, J., et al. 2025. "Environmental Matrix and Moisture Influence Soil Microbial Phenotypes in a Simplified Porous Media Incubation," *mSystems* **10**(3). DOI:10.1128/msystems.01616-24.

Roesch, L. F. W., et al. 2007. "Pyrosequencing Enumerates and Contrasts Soil Microbial Diversity," *The ISME Journal* **1**(4), 283–90. DOI:10.1038/ismej.2007.53.

Rogers, K. R. 2006. "Recent Advances in Biosensor Techniques for Environmental Monitoring," *Analytica Chimica Acta* **568**(1–2), 222–31. DOI:10.1016/j.aca.2005.12.067.

Rohlhill, J., et al. 2017. "Sort-Seq Approach to Engineering a Formaldehyde-Inducible Promoter for Dynamically Regulated *Escherichia coli* Growth on Methanol," *ACS Synthetic Biology* **6**, 1584–95. DOI:10.1021/acssynbio.7b00114.

Romero, P. A., and F. H. Arnold. 2009. "Exploring Protein Fitness Landscapes by Directed Evolution," *Nature Reviews Molecular Cell Biology* **10**, 866–76. DOI:10.1038/nrm2805.

Roque-Malo, S., et al. 2020. "Modeling the Role of Root Exudation in Critical Zone Nutrient Dynamics," *Water Resources Research* **56**(8). DOI:10.1029/2019WR026606.

Rosewich, U. L., and H. C. Kistler. 2000. "Role of Horizontal Gene Transfer in the Evolution of Fungi," *Annual Review of Phytopathology* **38**, 325–63. DOI:10.1146/annurev.phyto.38.1.325.

Rubin, B. E., et al. 2021. "Species- and Site-Specific Genome Editing in Complex Bacterial Communities," *Nature Microbiology* **7**, 34–47. DOI:10.1038/s41564-021-01014-7.

Rylott, E. L., and N. C. Bruce. 2020. "How Synthetic Biology Can Help Bioremediation," *Current Opinion in Chemical Biology* **58**, 86–95. DOI:10.1016/j.cbpa.2020.07.004.

Salis, H., et al. 2009. "Automated Design of Synthetic Ribosome Binding Sites to Control Protein Expression," *Nature Biotechnology* **27**, 946–50. DOI:10.1038/nbt.1568.

Salvado, B. et al. 2015. "A Survey of HK, HPt, and RR Domains and Their Organization in Two-Component Systems and Phosphorelay Proteins of Organisms with Fully Sequenced Genomes," *PeerJ* **13**(3), e1183. DOI:10.7717/peerj.1183.

Santos, C. N. S., et al. 2013. "Implementation of Stable and Complex Biological Systems Through Recombinase-Assisted Engineering," *Nature Communications* **4**, 2503. DOI:10.1038/ncomms3503.

Sasse, J., et al. 2019. "Multilab EcoFAB Study Shows Highly Reproducible Physiology and Depletion of Soil Metabolites by a Model Grass," *New Phytologist* **222**(2), 1149–60. DOI:10.1111/nph.15662.

Scarborough, M. J., et al. 2020. "Diagnosing and Predicting Mixed-Culture Fermentations with Unicellular and Guild-Based Metabolic Models," *mSystems* **5**(5). DOI:10.1128/msystems.00755-20.

Schmidl, S. R., et al. 2019. "Rewiring Bacterial Two-Component Systems by Modular DNA-Binding Domain Swapping," *Nature Chemical Biology* **15**, 690–8. DOI:10.1038/s41589-019-0286-6.

Selvamani, V., et al. 2017. "Engineering Chimeric Two-Component System into *Escherichia coli* from *Paracoccus denitrificans* to Sense Methanol," *Biotechnology and Bioprocessing Engineering* **22**, 225–30. DOI:10.1007/s12257-016-0484-y.

Shade, A., et al. 2013. "A Meta-Analysis of Changes in Bacterial and Archaeal Communities with Time," *The ISME Journal* **7**(8), 1493–506. DOI:10.1038/ismej.2013.54.

Silverstein, M. R., et al. 2023. "Environmental Microbiome Engineering for the Mitigation of Climate Change," *Global Change Biology* **29**(8), 2050–66. DOI:10.1111/gcb.16609.

Singer, E., et al. 2021. "Novel and Emerging Capabilities that Can Provide a Holistic Understanding of the Plant Root Microbiome," *Phytobiomes Journal* **5**(2). DOI:10.1094/PBIOMES-05-20-0042-RVW.

Singh, N., et al. 2025. "A Generalized Platform for Artificial Intelligence-Powered Autonomous Protein Engineering." Preprint, *bioRxiv*, February 14, 2025. DOI:10.1101/2025.02.12.637932.

Sizemore, N., et al. 2024. "A Digital Twin of the Infant Microbiome to Predict Neurodevelopmental Deficits," *Science Advances* **10**(15). DOI:10.1126/sciadv.adj0400.

Slomovic, S., et al. 2015. "Synthetic Biology Devices for *in vitro* and *in vivo* Diagnostics," *Proceedings of the National Academy of Sciences of the USA* **112**(47), 14429–35. DOI:10.1073/pnas.1508521112.

Smanski, M., et al. 2016. "Synthetic Biology to Access and Expand Nature's Chemical Diversity," *Nature Reviews Microbiology* **14**, 135–49. DOI:10.1038/nrmicro.2015.24.

Solscheid, C., et al. 2015. "Development of a Reagentless Biosensor for Inorganic Phosphate, Applicable over a Wide Concentration Range," *Biochemistry* **54**, 5054–62. DOI:10.1021/acs.biochem.5b00449.

Son, H.-I., et al. 2025. "Population-Level Amplification of Gene Regulation by Programmable Gene Transfer," *Nature Chemical Biology.* DOI:10.1038/s41589-024-01817-9.

Sørensen, S. J., et al. 2005. "Studying Plasmid Horizontal Transfer *in situ*: A Critical Review," *Nature Reviews Microbiology* **3**, 700–10. DOI:10.1038/nrmicro1232.

Spencer, S. J., et al. 2015. "Massively Parallel Sequencing of Single Cells by epicPCR Links Functional Genes with Phylogenetic Markers," *The ISME Journal* **10**, 427–36. DOI:10.1038/ismej.2015.124.

Srivastava, U., et al. 2024. "Chapter 2–Types of Omics Data: Genomics, Metagenomics, Epigenomics, Transcriptomics, Proteomics, Metabolomics, and Phenomics." In *Integrative Omics: Concept, Methodology, and Application*, pp. 13–34. Eds. Gupta, M. K., P. Katara, S. Mondal, and R. L. Singh. Integrative Omics Academic Press.

Staudt, A., et al. 2013. "The Added Complications of Climate Change: Understanding and Managing Biodiversity and Ecosystems," *Frontiers in Ecology and the Environment* **11**(9), 494–501. DOI.10.1890/120275.

Stirling, F., et al. 2017. "Rational Design of Evolutionarily Stable Microbial Kill Switches," *Molecular Cell* **68**(4), 686–97. DOI:10.1016/j.molcel.2017.10.033.

Stock, A. M., et al. 2000. "Two-Component Signal Transduction," *Annual Review of Biochemistry* **69**(1), 183–215. DOI:10.1146/annurev.biochem.69.1.183.

Stolyar, S., et al. 2007. "Metabolic Modeling of a Mutualistic Microbial Community," *Molecular Systems Biology* **3**(92). DOI:10.1038/msb4100131.

Stone, A., et al. 2024. "Context-Dependent Redesign of Robust Synthetic Gene Circuits," *Trends in Biotechnology* **42**(7), 895–909. DOI:10.1016/j.tibtech.2024.01.003.

Strecker, J., et al. 2019. "RNA-Guided DNA Insertion with CRISPR-Associated Transposes," *Science* **365**(6448), 48–53. DOI:10.1126/science.aax9181.

Stricker, J., et al. 2008. "A Fast, Robust and Tunable Synthetic Gene Oscillator," *Nature* **456**, 516–9. DOI:10.1038/nature07389.

Sturme, M. H. J., et al. 2002. "Cell to Cell Communication by Autoinducing Peptides in Gram-Positive Bacteria," *Antonie Van Leeuwenhoek* **81**(1–4), 233–43. DOI:10.1023/A:1020522919555.

Su, L., et al. 2011. "Microbial Biosensors: A Review," *Biosensors and Bioelectronics* **26**(5), 1788–99. DOI:10.1016/j.bios.2010.09.005.

Tanaka, T., et al. 2024. "Precise Microbiome Engineering Using Natural and Synthetic Bacteriophages Targeting an Artificial Bacterial Consortium," *Frontiers in Microbiology* **15**. DOI:10.3389/fmicb.2024.1403903.

Tang, T.-C., et al. 2020. "Materials Design by Synthetic Biology," *Nature Reviews Materials* **6**, 332–50. DOI:10.1038/s41578-020-00265-w.

Taylor, G., et al. 2024. "Harnessing the Power of Poplar Tree Natural Genetic Variation for the Development of Future Sustainable Biofuels and Bioproducts: A Droughted Marginal-Land Experiment for Multi-Disciplinary Investigations." Preprint, *bioRxiv*, January 15, 2024. DOI:10.1101/2024.01.11.575272.

The Reference Genome Group of the Gene Ontology Consortium. 2009. "The Gene Ontology's Reference Genome Project: A Unified Framework for Functional Annotation across Species," *PLOS Computational Biology* **5**(7), e1000431. DOI:10.1371/journal.pcbi.1000431.

Thomas, C. M., and K. M. Nielsen. 2005. "Mechanisms of, and Barriers to, Horizontal Gene Transfer Between Bacteria," *Nature Reviews Microbiology* **3**, 711–21. DOI:10.1038/nrmicro1234.

Thompson, J., et al. 2019. "Machine Learning to Predict Microbial Community Functions: An Analysis of Dissolved Organic Carbon from Litter Decomposition," *PLOS One* **14**(7). DOI:10.1371/journal.pone.0215502.

Thorogood, R., et al. 2023. "Understanding and Applying Biological Resilience, From Genes to Ecosystems," *npj Biodiversity* **2,** 16. DOI:10.1038/s44185-023-00022-6.

Tourab, H., et al. 2024. "The Use of Machine Learning and Explainable Artificial Intelligence in Gut Microbiome Research: A Scoping Review," Preprint, *TechRxiv*, October 15, 2024. DOI:10.36227/techrxiv.172902810.02756061.

U.S. DOE. 2025. "Soil & Groundwater Remediation." Accessed June 2025. U.S. Department of Energy. energy.gov/em/soil-groundwater-remediation

U.S. EPA. 2025. "Superfund: National Priorities List (NPL)." Accessed August 2025. U.S. Environmental Protection Agency. www.epa.gov/superfund/superfund-national-priorities-list-npl

van den Berg, N. I., et al. 2022. "Ecological Modelling Approaches for Predicting Emergent Properties in Microbial Communities," *Nature Ecology & Evolution* **6**, 855–65. DOI:10.1038/s41559-022-01746-7.

Van Nuland, M. E., et al. 2023. "Above- and Belowground Fungal Biodiversity of Populus Trees on a Continental Scale," *Nature Microbiology* **8**, 2406–19. DOI:10.1038/s41564-023-01514-8.

van Vliet, S., and M. Doebeli. 2019. "The Role of Multilevel Selection in Host Microbiome Evolution," *Proceedings of the National Academy of Sciences of the USA* **116**(41), 20591–7. DOI:10.1073/pnas.1909790116.

van't Padje, A., et al. 2021. "Temporal Tracking Of Quantum-Dot Apatite Across *in vitro* Mycorrhizal Networks Shows How Host Demand Can Influence Fungal Nutrient Transfer Strategies," *The ISME Journal* **15**(2), 435–49. DOI:10.1038/s41396-020-00786-w.

Vannier, N., et al. 2019. "Microbiota-Mediated Disease Resistance in Plants," *PLOS Pathogens* **15**(6), e1007740. DOI:10.1371/journal.ppat.1007740.

Vila, A. V., et al. 2020. "Impact of Commonly Used Drugs on the Composition and Metabolic Function of the Gut Microbiota," *Nature Communications* **11**, 362. DOI:10.1038/s41467-019-14177-z.

Volk, M. J., et al. 2022. "Metabolic Engineering: Methodologies and Applications," *Chemical Reviews* **123**(9), 5521–70. DOI:10.1021/acs.chemrev.2c00403.

Vörösmarty, C. J., et al. 2010. "Global Threats to Human Water Security and River Biodiversity," *Nature* **467**, 555–61. DOI:10.1038/nature09440.

Wang, G., et al. 2019. "CRAGE Enables Rapid Activation of Biosynthetic Gene Clusters in Undomesticated Bacteria," *Nature Microbiology* **4**, 2498–510. DOI:10.1038/s41564-019-0573-8.

Wang, R. H., et al. 2024. "PhageScope: A Well-Annotated Bacteriophage Database with Automatic Analyses and Visualizations," *Nucleic Acids Research* **52**(1), D756–61. DOI:10.1093/nar/gkad979.

Wang, T., et al. 2022. "Horizontal Gene Transfer Enables Programmable Gene Stability in Synthetic Microbiota," *Nature Chemical Biology* **18**, 1245–52. DOI:10.1038/s41589-022-01114-3.

Wangersky, P. J. 1978. "Lotka-Volterra Population Models," *Annual Review of Ecology and Systematics* **9**,189–218. DOI:10.1146/annurev.es.09.110178.001201.

Weiskopf, S. R., et al. 2020. "Climate Change Effects on Biodiversity, Ecosystems, Ecosystem Services, and Natural Resource Management in the United States," *Science of the Total Environment* **733**, 137782. DOI:10.1016/j.scitotenv.2020.137782.

Wen, A., et al. 2021. "Enabling Biological Nitrogen Fixation for Cereal Crops in Fertilized Fields," *ACS Synthetic Biology* **10**(12), 3264–77. DOI:10.1021/acssynbio.1c00049.

Wilken, S.-E., et al. 2019. "Linking 'Omics' to Function Unlocks the Biotech Potential of Non-Model Fungi," *Current Opinion in Systems Biology*. DOI:10.1016/j.coisb.2019.02.001.

Yaffe, E., and D. A. Relman. 2019. "Tracking Microbial Evolution in the Human Gut Using Hi-C Reveals Extensive Horizontal Gene Transfer, Persistence and Adaption," *Nature Microbiology* **5**, 343–53. DOI:10.1038/s41564-019-0625-0.

Yan, Y., et al. 2015. "Revisiting the Dilution Procedure Used to Manipulate Microbial Biodiversity in Terrestrial Systems," *Applied and Environmental Microbiology* **81**(13). DOI:10.1128/AEM.00958-15.

Yee, M. O., et al. 2021. "Specialized Plant Growth Chamber Designs to Study Complex Rhizosphere Interactions," *Frontiers in Microbiology* **12**. DOI:10.3389/fmicb.2021.625752.

Yin, H., et al. 2021. "Soil Sensors and Plant Wearables for Smart and Precision Agriculture," *Advanced Materials* **33**(20). DOI:10.1002/adma.202007764.

Yu, Y. 2022. "Emerging Microfluidic Technologies for Microbiome Research," *Frontiers in Microbiology* **13**. DOI:10.3389/fmicb.2022.906979.

Zegeye, E. K., et al. 2019. "Selection, Succession, and Stabilization of Soil Microbial Consortia," *mSystems* **4**(4). DOI:10.1128/msystems.00055-19.

Zhang, C., et al. 2022. "Microbial Utilization of Next-Generation Feedstocks for the Biomanufacturing of Value-Added Chemicals and Food Ingredients," *Frontiers in Bioengineering and Biotechnology* **10**. DOI:10.3389/fbioe.2022.874612.

Zhang, J., et al. 2022. "A Comprehensive Review of Image Analysis Methods for Microorganism Counting: From Classical Image Processing to Deep Learning Approaches," *Artificial Intelligence Review* **55**, 2875–944. DOI:10.1007/s10462-021-10082-4.

Zhuang, J., and Y. Jin. 2003. "Virus Retention and Transport as Influenced by Different Forms of Soil Organic Matter," *Journal of Environmental Quality* **32**(3), 816–23. DOI:10.2134/jeq2003.0816.

Zhang. R., et al. 2021. "Winner-Takes-All Resource Competition Redirects Cascading Cell Fate Transitions," *Nature Communications* **12**, 853. DOI:10.1038/s41467-021-21125-3.

Zhu, X., et al. 2022. "Microfluidics as an Emerging Platform for Exploring Soil Environmental Processes: A Critical Review," *Environmental Science & Technology* **56**(2). DOI:10.1021/acs.est.1c03899.

Zwart, P. H., et al. 2025. "Hyperspectral Segmentation of Plants in Fabricated Ecosystems," *Frontiers in High Performance Computing* **3**. DOI:10.3389/fhpcp.2025.1547340.

Appendix E

Acronyms and Abbreviations

3D	three dimensional
AI	artificial intelligence
BER	Biological and Environmental Research program
BSSD	Biological Systems Science Division
Cas	CRISPR-associated protein
Cas9	CRISPR-associated protein 9
CDS	coding DNA sequences
ChIP-seq	chromatin immunoprecipitation sequencing
CMM	community metabolic model
COMETS	Computation of Microbial Ecosystems in Time and Space
CRAGE	chassis-independent recombinase-assisted genome engineering
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRi	CRISPR interference
DAP-seq	DNA affinity purification sequencing
DART	DNA-editing-all-in-one RNA-guided CRISPR-Cas transposase
DBTL	Design-Build-Test-Learn (cycle)
DOE	U.S. Department of Energy
epicPCR	emulsion, paired isolation, and concatenation PCR
ET-seq	environmental transformation sequencing
GEM	genome-scale metabolic model
gLV	generalized Lotka-Volterra (equation)
Hi-C	high-throughput chromosome conformation capture
HGT	horizontal gene transfer
HPC	high-performance computing
ITS	internal transcribed spacer
KBase	DOE's Systems Biology Knowledgebase
LTER	Long Term Ecological Research (Network)
ML	machine learning
NMDC	National Microbiome Data Collaborative
PS	protein switch
RAM	RNA-addressable modification
RBS	ribosomal binding sites
rRNA	ribosomal RNA
SAGE	serine-integrase-assisted genome engineering
SynCom	synthetic community
TCS	two-component system
TR	transcriptional regulator