DOE/SC-2029

Plant Design for a Developing Bioeconomy

Frontier Science for the Bioeconomy Workshop Series

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

Plant Design for a Developing Bioeconomy Workshop

March 12–14, 2025

Convened by

U.S. Department of Energy Office of Science

Biological and Environmental Research Program

Workshop Co-Chairs

Gözde S. Demirer
California Institute of Technology

Yiping Qi
University of Maryland

Chung-Jui (CJ) Tsai
University of Georgia

Organizers

Pablo Rabinowicz
U.S. Department of Energy

Vijay K. Sharma
U.S. Department of Energy

Shing Kwok
U.S. Department of Energy

Kari Perez
U.S. Department of Energy

Writing Team

Rebecca Bart, Donald Danforth Plant Science Center; Tessa Burch-Smith, Donald Danforth Plant Science Center; Heather Coleman, Syracuse University; Arjun Khakhar, Colorado State University; Laurie Leonelli, University of Illinois Urbana-Champaign; Seung Yon (Sue) Rhee, Michigan State University; Nathan Springer, Bayer Crop Science; Cătălin Voiniciuc, University of Florida; Daniel Voytas, University of Minnesota–Twin Cities

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. This research supports the U.S. Department of Energy's (DOE) 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 credit: A robotic arm attached to the SMART Plant 1.0 automation system snips a plant tissue sample at Oak Ridge National Laboratory (ORNL). Researchers use the system to harness artificial intelligence and robotics to manipulate and sample living plants, accelerating efforts to optimize traits in bioenergy and agricultural crops. [Courtesy ORNL]

Suggested citation: U.S. DOE. 2026. *Plant Design for a Developing Bioeconomy: Frontier Science for the Bioeconomy Workshop Series,* DOE/SC-2029. U.S. Department of Energy Office of Science. https://doi.org/10.2172/3021282.

DOE/SC-2029

Plant Design for a Developing Bioeconomy

Workshop Report

Frontier Science for the Bioeconomy Workshop Series

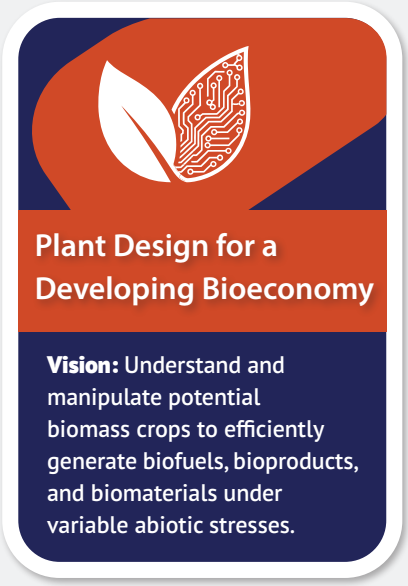
July 2026

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

Frontier Science for the Bioeconomy

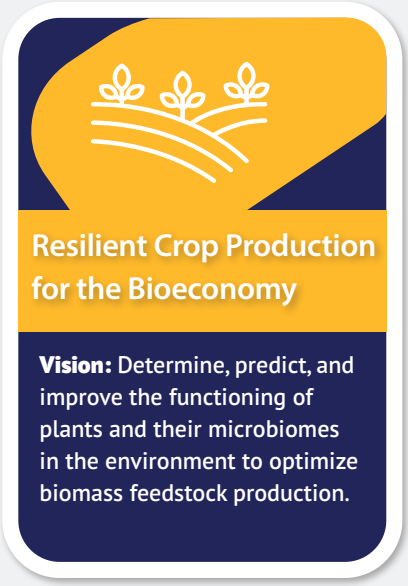
The *Plant Design for a Developing Bioeconomy* workshop report is one of a four-part Frontier Science for the Bioeconomy series hosted by the DOE Biological and Environmental Research program. The series defines frontiers of plant science, agriculture, synthetic biology, biobased materials, and environmental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

The bioeconomy is economic activity driven by research and involving innovation in the life sciences, engineering, and biomanufacturing. It includes industries, products (e.g., chemicals, biomass and seed oil feedstocks, and bioproducts), services, and processes derived from plant and microbial resources.



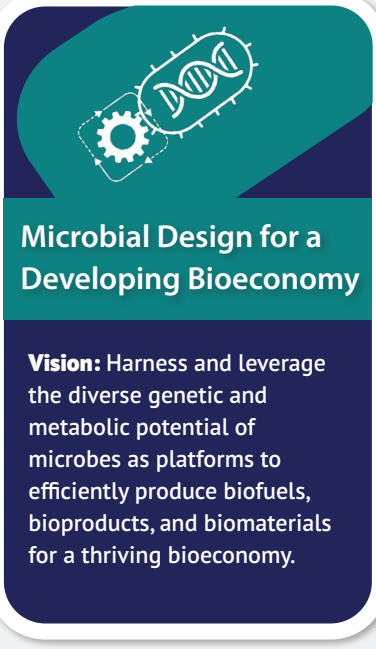
Plant Design for a Developing Bioeconomy

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



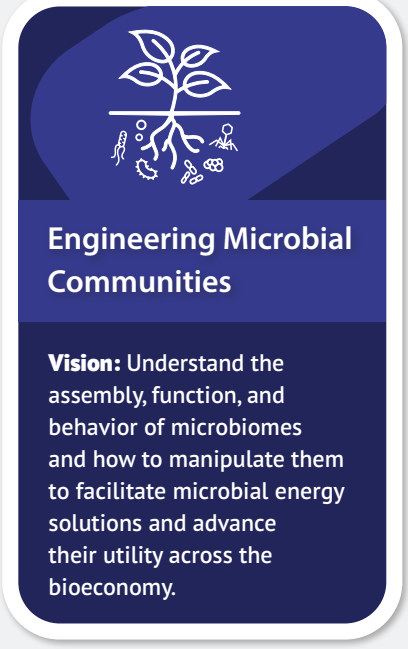
Resilient Crop Production for the Bioeconomy

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



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.



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.

Contents

Executive Summary

Recent advances in fundamental plant biology research, synthetic biology, and artificial intelligence (AI) are unlocking powerful new capabilities in plant biodesign, offering unprecedented potential to reimagine plants as programmable platforms for resource-efficient production of bioenergy, biomaterials, chemicals, and more. The U.S. Department of Energy (DOE) convened the Plant Design for a Developing Bioeconomy virtual workshop on March 12 through 14, 2025, to bring together leaders across plant science, engineering, and computation to assess the current landscape and define a bold vision for future research. Discussions during the workshop built upon findings included in DOE's Biological and Environmental Research (BER) workshop report *Overcoming Barriers in Plant Transformation: A Focus on Bioenergy Crops* (U.S. DOE 2024; genomicscience.energy.gov/plant-transformation). Participants identified critical knowledge gaps, technical barriers, and emerging opportunities in the design and engineering of plant systems to support a robust, resilient domestic bioeconomy aligned with DOE's mission. Key themes that emerged during the workshop are discussed below.

Plants as Chemical Factories: Challenges and Opportunities

A central theme of the workshop was the challenge and opportunity of engineering plants as chemical factories for products ranging from bioplastics to living materials. While the genomic era has yielded rapid improvements in sequencing and gene discovery, plant metabolism remains poorly annotated, particularly in feedstocks relevant to bioenergy and biomaterials. Key barriers include limited understanding of gene and enzyme function, gene regulatory elements, spatial and temporal dynamics of metabolic pathways, and transport mechanisms that govern intracellular product distribution. Bioengineering efforts often face trade-offs between productivity and plant health due to metabolite toxicity or interference with development. To address these challenges, workshop participants emphasized the need for predictive design strategies, including modular regulatory elements, tissue-specific expression systems, and even the creation of synthetic organs or compartments dedicated to specialized product synthesis and storage. Such strategies could isolate engineered pathways from core metabolic functions, enabling greater yields and stability without compromising plant performance.

However, the complexity of multicellular plant systems with dynamic environmental responses presents significant challenges to rapid iteration and optimization. Unlike microbial systems, where genetic designs can be tested and refined in days, plant systems involve long generation times and context-dependent behavior. This makes high-throughput experimental platforms a valuable asset. The development of heterologous expression systems, engineered plant cell lines, and real-time biosensors for metabolite monitoring were identified as essential innovation areas. These systems can support automated trait screening and provide annotated datasets critical for AI-guided design. In parallel, aligning upstream engineering with downstream manufacturing considerations, such as product recovery, purification, and scaling, will be crucial to translating plant designs into economically viable solutions. This alignment will require greater dialogue between academic and industrial stakeholders and the development of standard benchmarks for evaluating plant-based production platforms.

Bioengineered Plants for Environmental Resilience

Beyond metabolic output, workshop discussions highlighted the urgent need to engineer plants for resilience under real-world conditions. Environmental variability, ranging from drought to nutrient stress to microbial dynamics, often limits crop productivity and consistency in field settings. Experimental tools

and assays developed and optimized for simplified or artificial stress conditions in controlled environments are poorly suited to capture such complexity. Bridging this gap requires deeper understanding of stress sensing, signal integration, and plant–microbe interactions. A growing need also exists for understanding common and species-specific natural mechanisms of plant development, resilience, and interactions for more rational and robust plant biodesign. Synthetic biology can offer tools to build programmable circuits that respond to environmental cues with conditional activation of desired traits. However, these systems should be validated to ensure reliable performance across variable field conditions, resource efficiency, and minimal disruptiveness to core plant functions. The development of *in vivo* sensors, standardized field phenotyping protocols, and germplasm repositories with engineered lines adapted to key stressors were all cited as key enablers of future work in this area.

Synthetic Biology and Genome Engineering Breakthroughs To Advance Plant Biodesign

The workshop also emphasized that breakthroughs in synthetic biology and genome engineering are foundational to all aspects of plant biodesign. Participants noted the importance of precise, scalable genome editing technologies that go beyond simple mutagenesis to include targeted DNA insertion, transgene control, and synthetic chromosome design. While tools such as CRISPR-Cas systems and base editors have shown promise, their efficiency and specificity remain limited in many feedstock species. Newer technologies, including but not limited to DNA cutting and pasting methods (e.g., transposon- and recombinase-based integration platforms) and artificial chromosomes (e.g., synthetic centromeres), offer the potential to build modular, orthogonal genetic architectures in plants. These systems could enable entire biosynthetic pathways, developmental programs, or environmental sensing modules to be encoded on synthetic chromosomes, operating alongside the native genome with minimal interference. Although still in early stages, such approaches could radically expand the genetic design space for biomass feedstocks.

Biomolecule delivery remains a critical bottleneck in translating genetic designs into engineered crops. Current transformation protocols are still slow, labor-intensive, reliant on skilled labor, and genotype dependent. Emerging methods using viral vectors, nanoparticles, and regeneration-enhancing factors promise more efficient, tissue culture–free approaches, particularly if they can be adapted for germline delivery and broad host range. Advances in protoplast regeneration, shoot apical meristem targeting, and visual transformation markers may further increase throughput and enable greater automation. Coupling these delivery strategies with stable and archivable cell lines and phenotyping pipelines could form the basis of a distributed plant biofoundry network capable of supporting rapid design-build-test-learn cycles across diverse feedstock species.

Computational Tools To Aid Plant Design and Engineering

Recognizing the growing role of computational tools, the workshop devoted substantial attention to integrating automation, modeling, and AI into plant biodesign. AI tools, including large language models and structure-aware protein design platforms, already assist in literature mining, regulatory sequence design, and hypothesis generation. Future directions should include building plant-specific models capable of predicting gene function, regulatory interactions, and genotype-to-phenotype outcomes across varied environments. Such tools will require expansive high-quality training datasets linking molecular, physiological, and environmental information. A vision emerged of digital plant twins—computational models that can simulate trait expression, environmental responses, and productivity—to support rational design and field deployment. However, these advances hinge on a dramatic increase in the availability of omic and phenotypic data and on making AI tools more user-friendly, modular, and tailored to plant scientists' needs, while also incentivizing interdisciplinary collaboration between computer scientists and biologists.

Education and Workforce Development for Next-Generation Scientists

Finally, the workshop acknowledged that realizing this ambitious vision will depend on training a new generation of scientists equipped to work across disciplines. Traditional training often emphasizes depth over breadth, but future leaders must be fluent in molecular biology, computation, systems design, and plant–environment interactions. Education programs should support integrated, cohort-based training experiences, case-based learning, and meaningful engagement with industry, policy, and field-based research. Design thinking, systems literacy, and collaboration must become central components of scientific training to ensure innovation translates into impact.

The Future of Plant Design

Looking forward, the convergence of synthetic biology, AI, and plant science holds the promise to fundamentally reshape bioenergy and biomanufacturing. Designer crops may soon be able to rapidly and dynamically sense their environment. These genetically enhanced crops could reprioritize metabolic functions in real time to produce high-value materials tailored to local conditions and economic demands, as part of their engineered tasks. With coordinated investment in foundational research, enabling technologies, translational infrastructure, and workforce development, plant systems can be rationally programmed to support a scalable and resilient bioeconomy. Plants, long viewed as passive elements of ecosystems, are now emerging as dynamic platforms for engineered function: programmable biological systems shaped by advances in data science and synthetic biology to meet the evolving needs of a thriving bioeconomy.

The convergence of synthetic biology, artificial intelligence, and plant science holds the promise to fundamentally reshape bioenergy and biomanufacturing.

Chapter 1

Introduction

Recent breakthroughs in biotechnology and artificial intelligence (AI) have presented unprecedented opportunities and challenges for biodesign of next-generation crops, including biomass and bioenergy feedstock species. As new research directions rapidly emerge, assessing the current landscape and identifying areas with the greatest potential to advance innovation will be key to position the United States at the forefront of the global bioeconomy. In light of these needs, the U.S. Department of Energy (DOE) convened the Plant Design for a Developing Bioeconomy virtual workshop on March 12 through 14, 2025, to gather community input on the potential of plant design research that aligns with the mission of the Biological Systems Science Division's (BSSD)[1] Genomic Science Program (GSP), part of the Biological and Environmental Research (BER) program within DOE's Office of Science. The workshop was one of a four-part Frontier Science for the Bioeconomy series hosted by BSSD. The series defines frontiers of plant science, agricultural sustainability, synthetic biology, biobased materials, and environmental microbiome science that can unlock an innovative, resilient U.S. bioeconomy.

1.1 Workshop Focus and Objectives

The workshop aimed to assess how, within DOE and BER missions, plant feedstock systems can be understood and reprogrammed to efficiently produce biofuels, bioproducts, and biomaterials under variable abiotic stresses (see Appendix A: Workshop Agenda, p. 41). Thirty-two experts with diverse expertise (see Appendix B: Workshop Participants, p. 45) attended virtual presentations and breakout sessions organized around the workshop's five high-level focus areas: (1) plant chemical factories for biofuels, biomaterials, and bioproducts; (2) stress resilience and environmental interactions; (3) genome engineering challenges across scales; (4) computational tools, automation, and modeling; and (5) cross-cutting topics (see Appendix C: Workshop Vision and Goals, p. 46).

The workshop's overarching objective was to identify opportunities, roadblocks, and knowledge gaps that must be addressed to realize the full potential of plant design and bioengineering in a growing bioeconomy. Achieving the ambitious goal of a robust bioeconomy over the next 5 to 10 years will require innovative, multidisciplinary approaches spanning molecular biology, genetics, biochemistry, and stress physiology, integrated with synthetic biology, engineering, computational biology, and AI approaches.

As new research directions in plant design emerge, assessing the current landscape and identifying areas with the greatest potential to advance innovation will be key to position the United States at the forefront of the global bioeconomy.

1.2 Workshop Questionnaire

Workshop organizers developed a pre-workshop brainstorming questionnaire (see Appendix D, p. 48) that

[1] BER's BSSD has now been reorganized as part of two new BER divisions.

was sent to experts in the field who were also invited to participate in the virtual workshop. Questions asked respondents to reflect on the fundamental research needed to develop the next generation of genome engineering technologies that would enable effective and efficient manipulation of plant functions to produce value-added chemicals, biofuels, and biomaterials from renewable biomass in support of the emerging U.S. bioeconomy. Responses explored the potential of plant design research, the current state of implementing the design-build-test-learn cycle in plant genetic engineering, existing knowledge gaps, key technical barriers, and possible risks or unintended consequences of expanding bioengineered crop applications.

1.3 Report Structure

This report summarizes the workshop's key outcomes and provides a forward-looking perspective on the future of plant synthetic biology and biosystems design beyond the current state of the art.

Chapter 2 (p. 3) explores how advances in plant biotechnology and metabolic engineering can transform crops into efficient biofactories for fuels, materials, and high-value bioproducts. Chapter 3 (p. 13) examines strategies for engineering biomass crops that can maintain productivity and quality under complex, variable environmental conditions. Chapter 4 (p. 17) outlines an integrated engineering pipeline for next-generation plant biodesign to ultimately support the creation of resource-efficient, high-performing bioenergy feedstocks. Chapter 5 (p. 31) considers how automation, modeling, and AI can accelerate plant biodesign and identifies critical barriers to the implementation of AI-based tools for biodesign. Chapter 6 (p. 37) emphasizes training the next generation of plant design scientists to prepare researchers for the complex, real-world challenges of plant biodesign and bioengineering. Chapter 7 (p. 39) offers final insights for realizing the full potential of plants for bioproduction and enabling scalable, resilient, and high-performing bioengineered crops for a sustainable bioeconomy.

Chapter 2

Plant Chemical Factories for Biofuels, Biomaterials, and Bioproducts

Meeting the grand challenge of a thriving bioeconomy requires bold advances in plant biotechnology to transform biomass feedstock crops into efficient producers of high-value compounds. These include advanced fuel precursors, bioplastics, organic–inorganic composites, biominerals, and living materials with novel properties. This chapter identifies key knowledge gaps and research opportunities essential for enabling robust plant design platforms that support resource-efficient bioproduction.

2.1 Knowledge Gaps in Plant Metabolism and Engineering

The last two decades have brought remarkable advances in nucleic acid sequencing technologies, enabling the sequencing and assembly of increasingly large and complex plant genomes. However, a major barrier to leveraging these genomes for discovery and engineering lies in the limited functional annotation of genes, particularly those involved in specialized metabolic pathways (see Figure 2.1, p. 4). This challenge is compounded by the lack of accurate and comprehensive metabolite databases, which significantly lag behind genomic resources due to limited access to plant products, difficulties in isolation, and the labor-intensive nature of structural identification using metabolomics and nuclear magnetic resonance tools. These challenges are especially true for DOE-relevant feedstock crops, where many genes, enzymes, and specialized metabolites remain poorly annotated. Even with tools like single-cell transcriptomics, genes involved in low-abundance or cell-type-specific pathways often escape detection. Incomplete understanding and annotation of individual metabolic pathways and their regulatory mechanisms limit abilities to model complex pathway interactions and predict competition between pathways.

Differences in plant metabolism can often be attributed to variation in enzyme activity regulation. However, gene regulatory elements, such as promoters and enhancers, are poorly understood. Also underexplored are the roles of epigenetic mechanisms (e.g., DNA methylation and histone modifications), post-translational modifications of structural or regulatory proteins, and messenger RNA (mRNA) modifications (e.g., RNA methylation)—all of which can dynamically influence gene expression and metabolic outcomes. These gaps hamper scientists' ability to model and predict changes in metabolic flux. A related challenge is the lack of predictable, modular regulatory elements for metabolic engineering. Without a robust toolkit, it is difficult to control gene expression with spatial and temporal precision in engineered systems. Currently, transferability of a desired pathway from a wild or undomesticated crop to an experimental crop cannot be predicted and is often determined only after considerable trial and error.

Another major challenge is understanding the compartmentalization of metabolism. Many specialized metabolites are produced in specific cell types or

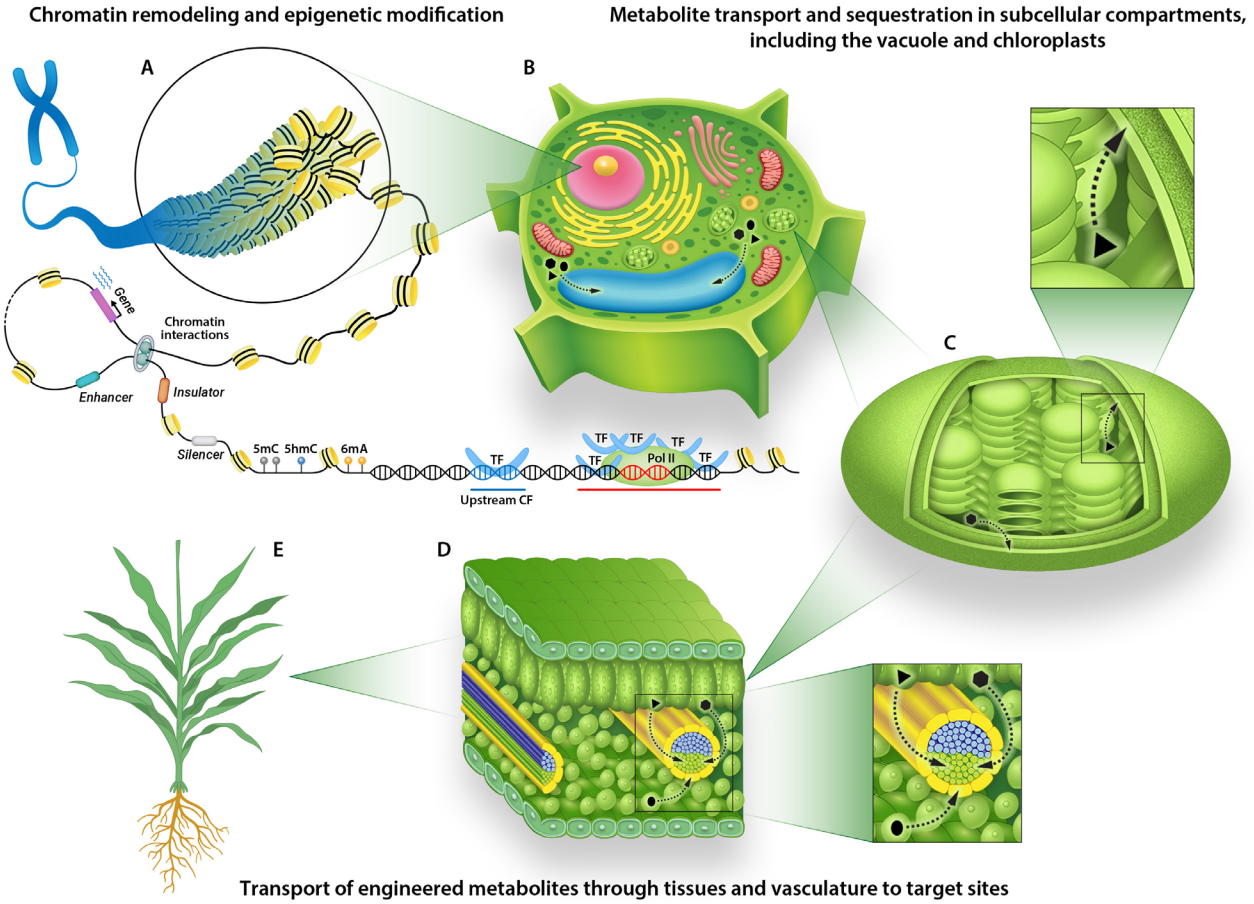


Fig. 2.1. Challenges and Opportunities in Plant Metabolic Engineering. (A) Precision metabolic engineering requires identifying gene regulatory mechanisms that can be harnessed to optimize metabolic gene expression and understanding other layers of genome regulation (e.g., epigenetics). The organization of subcellular compartments and metabolic intermediates is integral to the successful production of bioproducts. For example, metabolites could be imported into the vacuole **(B)** or exported into the chloroplast intermembrane space **(C)** to sequester them from the rest of metabolism. Interrogating and redesigning how products of engineered metabolism can be transported from synthesis sites to the plant vasculature for long-distance transport **(D)** will be critical for ensuring their arrival at optimal sites for harvest and use **(E)**. Black shapes represent products of metabolic engineering.

subcellular compartments, such as plastids or vacuoles, which provide unique biochemical environments. Subcellular sequestration can render otherwise toxic intermediates harmless. Furthermore, once a bioproduct is synthesized, its movement within and between cells and tissues becomes critical. In some cases, the product must be sequestered or retained in specialized cells to avoid toxicity or degradation. In other cases, transporting the product to harvestable tissues is desirable. Since organelles are surrounded by lipid bilayers and cells are encased in macropolymer-rich walls, the movement of these biomolecules typically depends on membrane-bound transporters and plasmodesmata. However, understanding of these transport mechanisms—including cargo specificity, capacity, and regulation—remains extremely limited. This significant knowledge gap constrains the ability to optimize valuable bioproduct distribution.

The relationship between plant metabolism and development is another underexplored area. Many bioproducts, such as wood and fiber, are tightly linked to developmental programs. Some metabolic processes are also sensitive to environmental changes. For example, cell wall macromolecular composition can

DOE Developments in Plant Engineering with Potential Industry Applications

DOE researchers have made significant strides in plant engineering research that offer opportunities to expand industry collaborations and speed the translation of basic research to applications that stimulate the U.S. bioeconomy.

Boosting Lipid Content in Sugarcane

Researchers at the Center for Advanced Bioenergy and Bioproducts Innovation have metabolically engineered sugarcane stems to divert carbon flux from sucrose to oil (triacylglycerol) to boost lipid yields per hectare for biodiesel production. This advancement may support commercial production of biodiesel and other fatty acid derivatives from high-biomass-accumulating crops (Zale et al. 2015).

Rapid Poplar Flowering

Researchers at the Center for Bioenergy Innovation have developed a CRISPR gene editing method to shorten flowering time in poplar trees from about 10 years to just a few months. This contraction of the reproductive cycle accelerates efforts to breed trees with improved traits (Ortega et al. 2022).

Synthetic Transcription Factors

Joint BioEnergy Institute researchers have developed a toolbox of plant transcriptional effectors that enable a wide range of up- and down-regulation of plant gene expression. This approach has inspired at least one startup company—BasidioBio, Inc.—due to easy translation into different expression systems and fusion proteins (LBNL 2023).

Zip-Lignin™

Great Lakes Bioenergy Research Center researchers have made lignin easier and cheaper to break apart by introducing weak bonds, or "zips," into the lignin polymer in plant cell walls. With the potential to reduce costs involved in deconstructing biomass, Zip-Lignin™ technology would exert wide-ranging effects on industries producing paper, bioproducts, and cellulosic biofuels (GLBRC 2018).

shift in response to biotic or abiotic stress. Engineering metabolic pathways without considering developmental timing, tissue specificity, or stress responsiveness can lead to unintended phenotypes or reduced plant fitness. A deeper understanding of these links would enable more precise control over when and where bioproducts are synthesized and help avoid negative developmental consequences from premature or unintended pathway activation.

Finally, many scientists are unaware of the full range of high-value products that industry seeks. For example, green materials such as bioplastics and bionylons are increasingly sought by industry. While microbes have been engineered to produce bioplastic and nylon-like intermediates and polymers, translating these pathways into plant systems remains largely unexplored. Plants are easier and more affordable to scale up compared to microbial fermentation in bioreactors. For many specialized metabolites, intermediates in the biosynthetic pathway may provide a better industrial entry point than the final product, which can be harder to scale due to cellular toxicity, solubility issues, or purification challenges. These intermediates can be efficiently produced in plants and then upgraded through relatively inexpensive chemical steps, making them an attractive option for industrial processing. The potential misalignment between academic research and industrial priorities represents another knowledge gap. Greater awareness of opportunities to translate basic research into industrial applications would help researchers focus on targets with the highest potential for impact and bioeconomy viability (see sidebar, DOE Developments in Plant Engineering with Potential Industry Applications, this page).

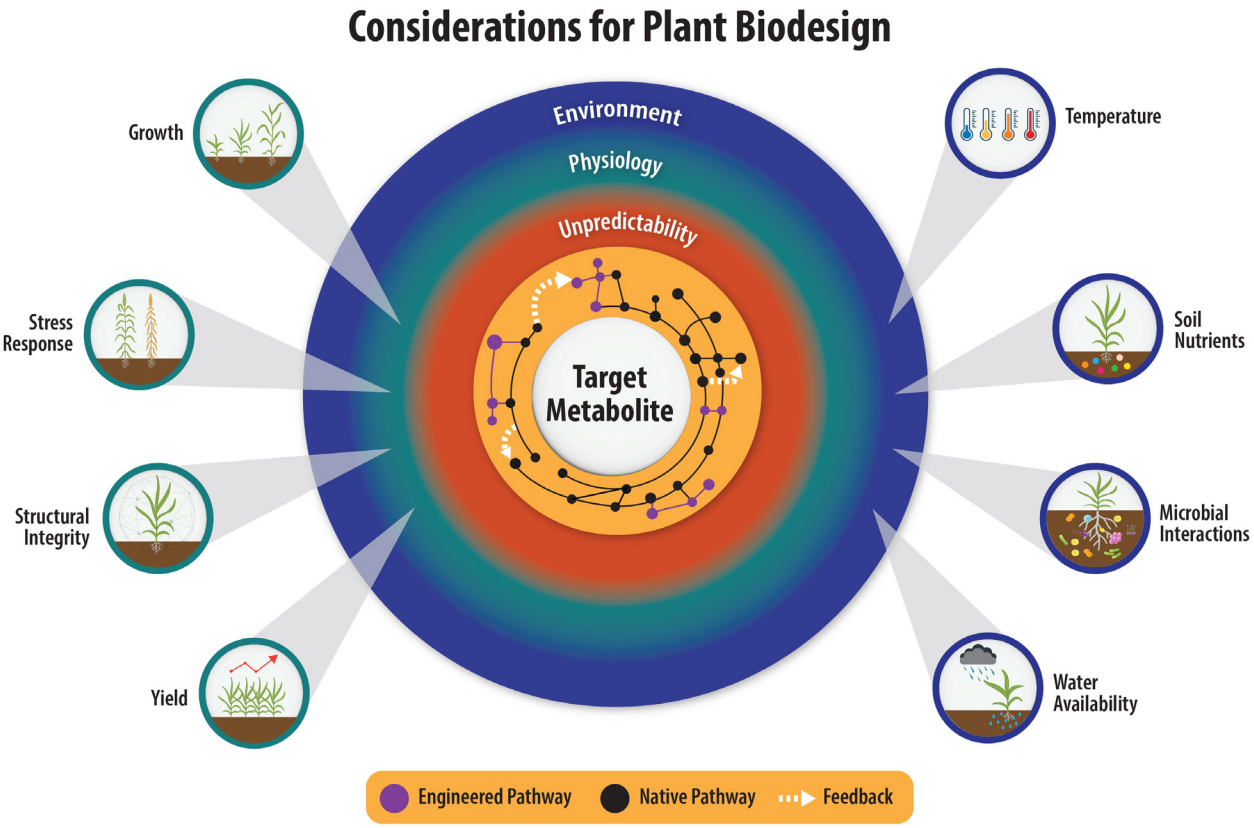


Fig. 2.2. Holistic Considerations for Plant Metabolic Biodesign. The central metabolic network illustrates the integration of engineered modules (purple dots) with native pathways (black dots), including feedback regulation (white arrows). These pathways operate within multiscale metabolic and physiological interactions shaped by environmental conditions. Traits such as growth, stress responses, structural integrity, and yield emerge from interactions between metabolism and physiology and are further modulated by environmental factors, including temperature, soil nutrient availability, microbial interactions, and water status.

2.2 Trade-Offs in Metabolic Engineering and Strategies for Mitigation

Metabolic pathway engineering in plants often introduces trade-offs that can compromise structural integrity, growth, or yield. Overexpression of certain metabolites or the introduction of new biosynthetic pathways can lead to cellular toxicity, disruption of native metabolism, and unintended feedback regulation. These pleiotropic effects are further complicated by the long-term instability of engineered traits, which may diminish over time due to epigenetic reprogramming or metabolic plasticity.

Even well-characterized pathways can behave unpredictably across different tissues or environmental conditions. Moreover, many biosynthetic pathways are modular in nature, where enzymes can function in different combinations to yield a wide range of products, complicating both prediction and regulation. This variability underscores the need for a holistic design approach, one that not only considers the target product but also the broader metabolic and physiological context of bioengineered crops (see Fig. 2.2, this page).

In naturally evolved systems, the synthesis of specialized compounds is tightly regulated through multitiered control mechanisms, such as spatial separation

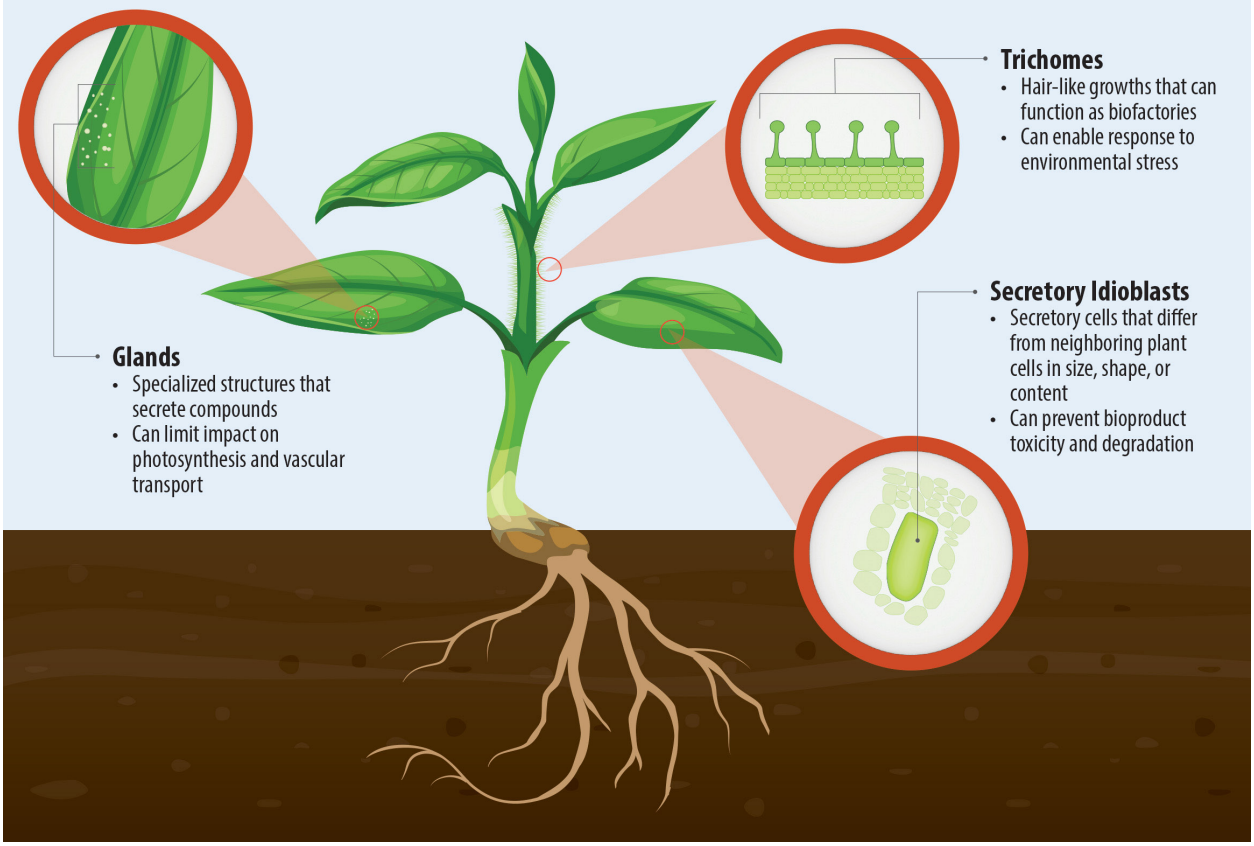


Fig. 2.3. Synthetic Plant Compartments for Bioproduction. Conceptual organ- and tissue-like structures spatially isolate engineered metabolic pathways from essential physiology. Insets illustrate synthetic glands, engineered trichomes, and secretory idioblast-like cells designed to accumulate, store, or secrete high-value products while minimizing trade-offs. Each compartment can integrate tailored transporters, regulatory circuits, and stress buffers to control production, storage, and release.

of pathway steps across distinct cell types or organelles, enzymatic modifications of metabolites (e.g., glycosylation and acylation), and sequestration of intermediates or end products into specialized cell types (e.g., trichomes or seed coat epidermal cells) or compartments (e.g., vacuoles or cell walls). These spatiotemporal mechanisms help minimize toxicity and prevent interference with core metabolism, yet they are not routinely incorporated into metabolic engineering efforts.

To mitigate these trade-offs, precise and context-aware engineering strategies are essential. There is a pressing need for experimentally validated libraries of promoters and regulatory elements that enable fine-tuned spatial and temporal control of gene expression at the organ, tissue, and even single-cell level. Having a diverse set of regulatory elements for each target tissue or cell type is especially important to avoid silencing effects caused by repetitive sequences. Advances in single-cell omics are rapidly expanding understanding of plant regulatory networks. However, translating these insights into functional, modular tools for synthetic biology remains a major challenge.

One promising direction is the *de novo* design of synthetic organs or novel tissues dedicated to bioproduction. By spatially isolating engineered metabolic pathways from essential physiological functions, these synthetic structures can reduce unintended trade-offs such as growth defects, metabolic imbalances, or toxicity (see Fig. 2.3, this page). For example, a synthetic gland-like structure could be engineered in leaves or

Technique Spotlight

Cell-free systems: Cell-free systems are synthetic environments that reproduce key biochemical functions of living cells without the constraints of membranes, chromosomes, or other cellular structures. By removing these barriers, cell-free systems make it easier to program, test, and modify gene circuits and metabolic pathways. These systems can synthesize proteins, produce diverse biomolecules, function as biosensors, and analyze environmental samples, providing a flexible platform for rapid prototyping in biological engineering.

stems to accumulate high-value terpenoids or alkaloids without interfering with photosynthesis or vascular transport. Engineered trichome-like tissues could serve as specialized sites for secreting value-added compounds or storing lipophilic products, mimicking natural strategies. These synthetic compartments could also be designed to include tailored transporters, regulatory circuits, and stress buffers, enabling precise control over production, storage, and release of target molecules. Nature offers a wealth of inspiration for the biodesign of specialized plant tissues, ranging from underground storage organs (e.g., corms and tubers) that help plants survive seasonal stresses, to versatile seed coats that release biopolymers upon hydration, to nonliving conifer cones that open and close depending on humidity. Such innovations could not only enhance crop yield and resilience under adverse conditions but also broaden the spectrum of products that can be synthesized *in planta*.

2.3 Experimental Characterization and High-Throughput Testing

Compared to the simplicity and speed of engineering single-celled microbes, characterizing and validating genetic changes in plants remains complex and time-consuming. Historically, bioengineers have not been well integrated with plant scientists, and high-throughput (HTP) systems for functional gene analysis in plants are still limited. Most plant biology departments also lack the infrastructure and automation expertise needed to support large-scale testing, from molecular to laboratory and field scales.

HTP experimental platforms are urgently needed to support AI-driven bioengineering, as the accuracy of machine learning algorithms and foundational models depends on well-annotated training datasets, which are currently lacking for plants. While cell-free systems offer some utility, they are often too simplified to evaluate plant biodesigns (see Technique Spotlight, this page). Therefore, additional HTP approaches must be developed to accelerate trait improvement in feedstock crops and enhance *in planta* production of bioproducts and biomaterials (see Fig. 2.4, p. 9). For example, subcellular compartments and secretion strategies must be screened to determine how cells can be protected from valuable but toxic compounds, or conversely, how to guard desired chemicals from being modified or degraded. Since bacterial systems lack key eukaryotic processes such as post-translational modifications, they are insufficient for many plant-specific applications. A variety of cellular systems, including yeast, plant cell lines, and hairy root cultures (see Technique Spotlight, p. 10), will likely be needed to establish standards and generate datasets that can prototype genetic designs and predict whole-plant outcomes with the aid of AI.

In addition to evaluating enzymes involved in product synthesis and transport, HTP screens can also test cell-specific and stress-responsive regulatory elements to enable feedstock crops to function as biofactories. However, a persistent challenge is that most engineered pathways do not produce visible

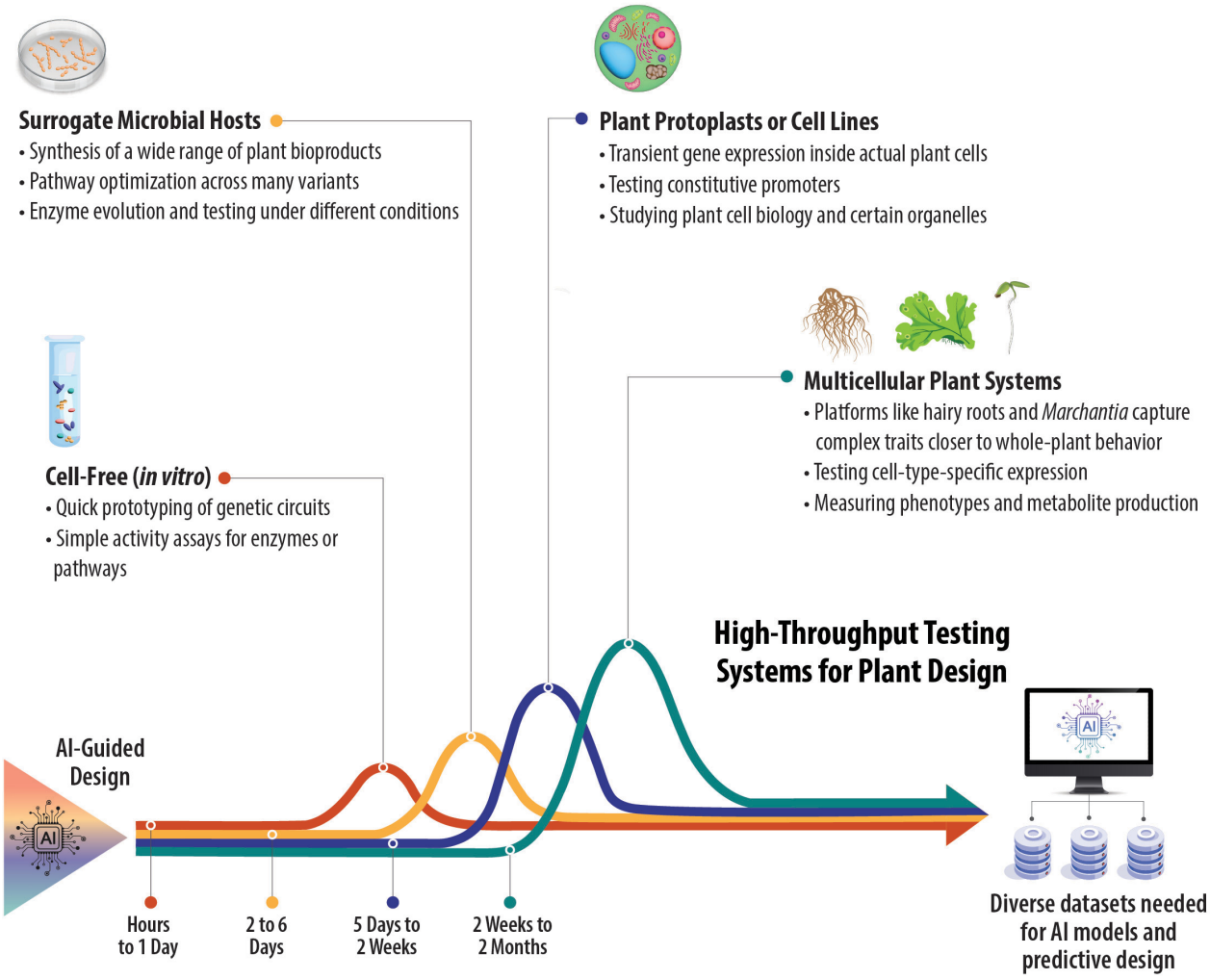


Fig. 2.4. High-Throughput Testing Systems for Plant Design. Testing genetic designs directly in whole plants is slow and complex. Using AI-guided design to develop better high-throughput (HTP) testing platforms, such as cell-free systems, surrogate microbial hosts, plant protoplasts or cell lines, and multicellular plant systems, can help accelerate the design-build-test-learn cycle, allowing researchers to identify the most promising designs before committing to whole plants. Each HTP system has different timelines and testing capabilities, ranging from very fast prototyping to more realistic plant-like outcomes. Cell-free systems are the fastest way to test a design's functionality (hours to 1 day), but this method lacks the complexity of real plant cells. Surrogate microbial hosts offer a mid-speed system for testing how plant pathways function in fast-growing microbes (2 to 6 days), though results may not fully translate to plants. Protoplasts or cell lines provide more realistic plant context than microbes, though this system is slower (5 days to 2 weeks) and less scalable. Multicellular plant platforms, such as hairy roots and *Marchantia*, are slower (2 weeks to 2 months) but more powerful for capturing complex, organ-level traits. By combining these systems, researchers can generate the diverse datasets needed for AI models and predictive design, making plant bioengineering more efficient.

Technique Spotlight

Hairy root cultures: Hairy root cultures are specialized plant tissues generated when *Rhizobium rhizogenes* (formerly known as *Agrobacterium rhizogenes*) transfers its DNA into wounded plant cells, triggering the formation of fast-growing, root-like structures. During infection, the bacterium exploits compounds released at the wound site to enter the plant, integrates segments of its plasmid DNA (known as transfer DNA) into the host genome, and induces these distinctive roots. Because hairy roots have the capacity to produce abundant and often unusual secondary metabolites, they are widely used to study and generate biofuels, biomaterials, and other bioproducts. They also provide a faster alternative to generating stable transgenic lines in certain plant species.

Nanobodies: Nanobodies are exceptionally small, stable antibody fragments derived from certain animals and valued for their robustness under stress. Their compact size and durability make them well suited for expression in plants, where they can function as molecular probes and living sensors or confer other properties, such as enhanced resistance to pathogens. Delivery methods such as *Agrobacterium*-based transformation can be used to introduce nanobody genes into plant cells, after which the plant can take over production of the nanobodies.

phenotypes (e.g., pigments), making them difficult to screen directly. Current analytical techniques are time-consuming, destructive, and require metabolite class–specific methods. HTP cultivation systems must therefore be able to generate sufficient biomass for downstream biochemical and omics analyses, such as proteomics, metabolomics, and transcriptomics, which remain costly and time intensive.

To address this challenge, novel probes like genetically encoded sensors can enable scalable phenotyping and real-time monitoring. For instance, fluorescently labeled carbohydrate-binding modules and nanobodies (see Technique Spotlight, this page) can be developed for live-cell imaging of specific plant polysaccharides, eliminating the need for traditional hydrolytic steps or lengthy immunohistological procedures (Voiniciuc et al. 2018).

To expand fundamental understanding of metabolic pathways, future experiments in heterologous hosts should not only determine individual gene functions but also evaluate *in vivo* bottlenecks such as carbon source competition. Advances in DNA synthesis and modular cloning can facilitate HTP screening of plant genes across multiple species, mutant or cultivar backgrounds, and spatiotemporal contexts, as different chassis may offer metabolic advantages for different target molecules. Many desirable (and unknown) compounds are lineage-specific, requiring discovery across diverse plant taxa. However, biosynthetic precursors are often more widely conserved, enabling pathway engineering in more tractable hosts once species-specific routes are identified. Long-term stability and plasticity of transgenes are critical considerations for feedstock crop transformation. Ultimately, HTP results coupled with predictive AI models should anticipate trade-offs and unintended consequences before committing to full-scale engineering. Retrofitting existing model systems or leveraging natural producers of target metabolites (e.g., *Jatropha* for latex or *Carnauba* for wax) can further accelerate the design-build-test-learn cycle, provided transformation bottlenecks can be overcome. A recent example uses CRISPR-Cas9 to redirect flux from the competing inulin pathway toward natural rubber biosynthesis in rubber dandelion (Ariyaratne et al. 2023), illustrating the potential of pathway reprogramming to enhance target metabolite production.

2.4 From Research to Product Development: Designing for Downstream Success

Despite the diversity and sheer abundance of plant species, the full potential of lignocellulosic biomass remains underutilized due to the recalcitrance of complex polymers and limited means to bioconvert and valorize all its carbon building blocks. Other compounds critical to U.S. energy security do not accumulate in sufficient quantities in current feedstock crops. Establishing a catalog of biosynthetic targets or pathway intermediates with real-world relevance to agriculture and energy security could guide strategic innovation. At the same time, exploratory research into novel or poorly characterized metabolic pathways is essential to avoid overlooking unexpected mechanisms or rare natural products that could unlock entirely new opportunities for next-generation biomanufacturing.

Although regulatory oversight, downstream processing, and biomanufacturing fall outside the scope of basic research, these challenges must be considered when designing plant-based bioproduction strategies. Extraction, purification, and scale-up are central to bridging the "valley of death" between laboratory discovery and commercial application. In this context, evaluating multiple plant biofactories through HTP assays will be essential for identifying cost-effective platforms. Engineering designer crops with built-in biosafety measures can also streamline regulatory approval and enable controlled field deployment and performance validation.

Feedback from downstream considerations can inform upstream engineering priorities and help define realistic goals—whether the aim is to produce intermediates for further processing or final products directly *in planta*. Designing with the end in mind will improve the likelihood of translational success and ensure plant synthetic biology contributes meaningfully to a thriving bioeconomy.

Decades of research in lignin bioengineering have demonstrated the value of structural modifications—rather than just reducing the amount of lignin—for improving biomass processing characteristics without compromising plant growth or structural integrity. Engineering novel monomers, either new-to-nature or new-to-the-target feedstock, offers a promising strategy to increase structural diversity for process-advantaged cell wall engineering. As exemplified by the development of Zip-Lignin, where lignin monomers are linked through more labile bonds (Unda et al. 2022), pathway engineering can introduce atypical monomers that promote such bonds during lignin polymerization, thereby facilitating downstream processing and bioproduction (see sidebar, DOE Developments in Plant Engineering with Potential Industry Applications, p. 5). Similar strategies targeting the structurally complex hemicelluloses, such as side-chain decorations or backbone modifications, could enable the selective separation of matrix polysaccharides without energy or chemically intensive steps. Cellulose fibers can also be decorated *in vivo* through genetically encoded novel precursors or enzymatic modifications to confer new chemical, physical, and thermal functionalities for programmable plant-based materials (Kuperman et al. 2024). Designer glycans could also support the development of engineered living materials (ELMs) with tunable properties, such as embedded protein functions that are activated by external stimuli or the ability to sequester rare minerals. While recent advances in ELMs have largely focused on bacterial polysaccharides, the plant life cycle—from a tiny seed to a majestic tree—continues to inspire the design of complex biomaterials with genetically encoded functions.

Many other industrial processes can also benefit from bioengineering. For instance, the production of optically clear cellulose nanofibers for electronics, biosensors, packaging, textiles, or transparent wood requires lignin removal through harsh chemical treatments. Gene-edited crops with modified cell wall chemotypes (e.g., reduced lignin or lignin-free in select cell types) can be purpose-grown for such applications, provided that potential environmental trade-offs, as discussed here and in the next chapter, are carefully considered.

Chapter 3

Bioengineered Feedstock Design for Environmental Resilience

Maximizing plant productivity and resilience to achieve a robust and sound bioeconomy requires a deep understanding of how plants interact with their environments. Future bioenergy crops must thrive under variable and often suboptimal field conditions, such as fluctuating water availability, nutrient limitations, temperature extremes, and soil heterogeneity, while maintaining yield, biomass quality, and resilience. Despite a pressing need, current tools for studying and engineering plant–environment interactions lag behind those available for controlled environments or single-stress scenarios, leaving critical knowledge and capability gaps. This chapter outlines key gaps, technical limitations, and research opportunities at the interface of plant biology and environmental complexity. Addressing these challenges will enable predictive design of crop genotypes optimized for diverse and dynamic field conditions, a prerequisite for realizing the full potential of a plant-based bioeconomy.

Addressing challenges in bioengineered feedstocks will enable predictive design of crop genotypes optimized for diverse and dynamic field conditions.

3.1 Knowledge Gaps and Technical Limitations

Environmental Complexity and Stress Dynamics

Fundamental questions remain about how plants perceive abiotic stress. For many key stressors—including drought, salinity, and nutrient deficiency—the primary sensors and early signal transduction pathways are not fully defined. Moreover, abiotic and biotic stresses in field settings are highly variable in timing, intensity, and type, often involving multiple concurrent or sequential stressors. Studies producing gene regulatory networks that describe how plants integrate signals to coordinate a response to multiple environmental variables are in development but remain largely nonpredictive, underscoring the complexity of plant–environment interactions. To further complicate matters, stress responses are tightly linked with developmental programs, often resulting in conditional growth–defense trade-offs that manifest only when stresses occur during a critical growth phase. As discussed in Chapter 2: Plant Chemical Factories for Biofuels, Biomaterials, and Bioproducts (see p. 3), these trade-offs can significantly impact the success of engineered traits, particularly when performance is evaluated under controlled conditions. Yet such complexity is often oversimplified in laboratory or greenhouse studies, where experiments typically involve single, well-defined stress treatments that may not fully capture the multifactorial and dynamic nature of field environments (see Fig. 3.1, p. 14). As a result, plant responses observed in controlled experiments do not always translate to field performance. While this underscores the importance of field testing, integrating controlled studies with targeted field evaluation could

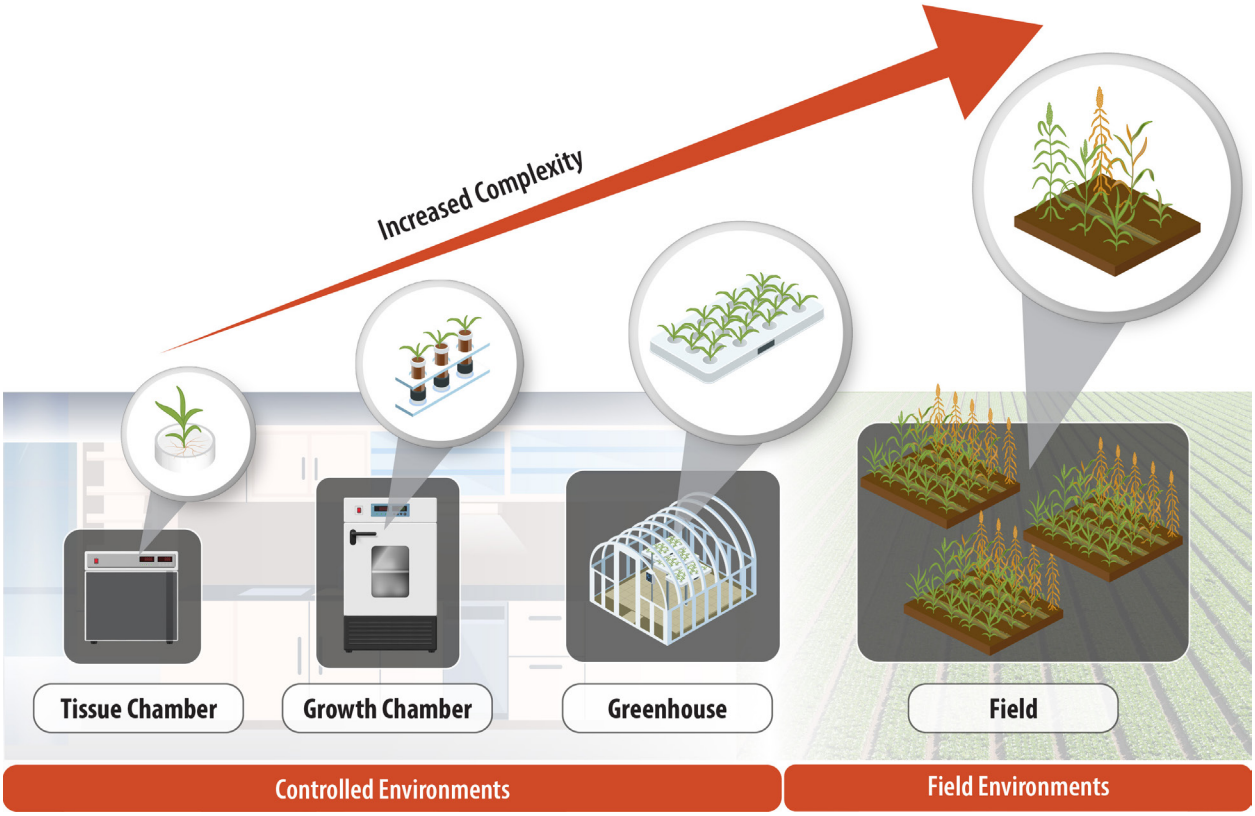


Fig. 3.1. Scaling Plant Design from Controlled Systems to the Field. Plant engineering and phenotyping progress across increasing levels of biological and environmental complexity. Workflows begin in highly controlled settings, from tissue culture and *in vitro* systems, where genetic constructs, synthetic pathways, or engineered chromosomes can be evaluated with maximal experimental control. These designs then advance through growth chambers and greenhouse environments, enabling assessment of whole-plant performance, developmental effects, and genotype-by-environment interactions under semi-controlled conditions. Final validation occurs in the field, where engineered traits are tested within realistic agronomic contexts that capture soil heterogeneity, microbial interactions, environmental variability, and ecosystem-level effects. The arrow denotes increasing biological, environmental, and regulatory complexity along this pipeline.

offer an efficient and scalable approach to assessing environmental interactions and informing trait design. This integrated strategy could also support the development of more predictive models and field-relevant testing platforms to bridge the gap between experimental insights and real-world outcomes.

Impacts of Microbial Communities on Resilience

Beneficial biotic interactions must also be considered when evaluating crop response to stresses encountered in the field. Microbial communities, particularly those associated with plant roots, can enhance crop resilience and nutrient uptake. These kinds of plant–microbe interactions influence plant gene expression and can help plants cope with abiotic stresses like drought. Root exudates—the compounds secreted by roots—have emerged as a key area of study because they shape root-associated microbial communities that confer stress tolerance. However, the genes controlling this interkingdom signaling are not well understood. Interest is growing in engineering these interactions or creating synthetic symbioses to improve plant stress resistance and yield. Yet, significant knowledge gaps persist, especially in identifying beneficial microbes and understanding the mechanisms behind plant–microbe and plant–environment interactions.

Technique Spotlight

RNA switches: RNA switches (also called riboswitches) are RNA-based regulatory elements that change shape in response to specific molecules in the cell's environment, altering how the RNA is translated. This structural shift turns gene expression on or off, enabling precise control of gene activity across many cell types. Because they react directly to environmental cues, RNA switches are powerful tools for engineering new cellular functions and for monitoring how plants respond to changing conditions.

Availability of Molecular Tools for Precision Engineering

Efforts to stabilize yields in variable environments are limited by an incomplete understanding of stress response pathway intersections in plants and the lack of fully characterized genetic components for crop design. Echoing the foundational need discussed in Chapter 2, the ability to engineer stress-resilient plants with minimal fitness costs is severely limited by the lack of validated, modular regulatory elements (e.g., promoters, enhancers, or RNA-based switches; see Technique Spotlight, this page) that can drive precise, condition-specific gene expression in response to environmental inputs. Similarly, epigenetic modifications and interactions with microbes are well understood to influence stress responses, particularly stress memory and acclimation. However, these factors are not well characterized and are rarely integrated into breeding pipelines or predictive models of plant performance.

3.2 Emerging Topics and Strategic Opportunities

Community Standards and Infrastructure

Despite growing interest and investment in environmental resilience, plant biology as a field broadly lacks standardized protocols for defining and measuring abiotic stresses, particularly under field conditions. Developing uniform methods for stress quantification, environmental monitoring, metadata collection, and outcome reporting would enable cross-study comparisons and improve model development. Additionally, addressing the challenge that greenhouse and growth chamber studies do not closely reflect field-informed stress regimes and developing scalable phenotyping platforms would allow for more accurate predictions and subsequent field performance measurements. Intermediate systems, such as semi-controlled outdoor plots with high-resolution monitoring, could help bridge this gap. Expertise in experimental design for complex environments, use of environmental sensors or high-throughput phenotyping tools, and interpretation of field-based data, especially when grounded in plant physiology, are increasingly rare. Strengthening both infrastructure and interdisciplinary expertise is essential to advancing this research vision.

Genetic Resources for Plant Environmental Response Design

Development of synthetic biology components for plants that respond to specific environmental cues (e.g., water potential, ion concentration, heavy metals, and temperature) is needed to enable programmable responses with greater spatial, temporal, and conditional control. These components must be field validated and capable of functioning across diverse genetic backgrounds and environments. Once field validated, they will pave the way for the development of new biosensors (see Technique Spotlight, p. 16) that can monitor and provide real-time readouts of stress perception, hormone levels, and physiological status in the field. *In vivo* sensors could also support trait selection and screen for functional variation. As these synthetic systems become more complex, traits enabling containment (e.g., inducible sterility

Technique Spotlight

Biosensors: Biosensors are biological components that detect specific molecules or events and convert them into an observable signal, such as fluorescence or a color change. The signal typically scales with the concentration of the target substance, such as a biomolecule or microbe, allowing real-time measurement of processes like gene expression or pathogen presence in plants. Through genetic engineering, biosensors can be integrated into living systems to monitor plant health, performance, and environmental interactions.

Phytomining: Phytomining is the use of living plants for extracting valuable minerals from soil by using species that naturally accumulate metals in their tissues. Just as plants take up metals through normal nutrient absorption pathways, certain hyperaccumulator species concentrate these elements in leaves, stems, or sap at exceptionally high levels. Although less than 1% of plants show this trait—and most preferentially accumulate nickel—ongoing research aims to enhance the abilities of natural hyperaccumulators or engineer new plants to expand the range of metals amenable to phytomining.

or niche restriction) will also be vitally important for responsible deployment. These genetic resources and tools may aid the development of hyperaccumulators for phytomining critical minerals like nickel. This research theme was recently funded by the Plant HYperaccumulators TO MIne Nickel-Enriched Soils (PHYTOMINES) within DOE's Advanced Research Projects Agency-Energy (ARPA-E). An exploratory topic, ARPA-E's PHYTOMINES aims to encourage the technological development of domestic phytomining (see Technique Spotlight, this page; ARPA-E 2024). Similar approaches hold promise for engineering plants to mine other critical minerals and rare earth elements (Rylott and van der Ent 2025).

Establishment of Germplasm Repositories

Establishing living repositories as a centralized resource for engineered or vegetatively propagated plant materials would not only help reduce redundancy in research time and cost but also enable more consistent replication of experiments across research sites. Databases like The Arabidopsis Information Resource (TAIR) and the Germplasm Resources Information Network (GRIN) maintain seed stocks for Arabidopsis and non-genetically modified crop species, respectively. However, comparable resources are lacking for most feedstock species—whether wild-collected or genetically modified. Additionally, encouraging the incorporation of a standard set of "benchmark" lines and environments in experimental design, as well as including a core set of genotypes for each crop species and representative stress scenarios, would complement germplasm repositories and help researchers standardize the evaluation of resilience traits across conditions. Together, these coordinated efforts would enhance comparability and efficiency in crop research, while still allowing the flexibility needed to accommodate diverse crop types and regional conditions.

Chapter 4

Synthetic Biology, Genome Engineering, Transformation, and Phenotyping for Next-Generation Plant Biodesign

Technical advances over the past few decades have built important groundwork, enabling early efforts to harness the biosynthetic capacity of plants (see Ch. 2: Plant Chemical Factories for Biofuels, Biomaterials, and Bioproducts, p. 3) and explore stress response pathways (see Ch. 3: Bioengineered Feedstock Design for Environmental Resilience, p. 13). Yet, as previous chapters make clear, many tools remain underdeveloped for precise and scalable plant biodesign. Continued innovation in plant synthetic biology, genome engineering, transformation, and phenotyping is essential to close critical gaps and build an integrated plant engineering pipeline that fully realizes the potential of plant-based bioproduction. This chapter outlines key priorities and emerging opportunities across the plant engineering ecosystem to enable next-generation plant biodesign.

4.1 Plant Synthetic Biology

Plant synthetic biology is the application of quantitative, engineering-driven approaches to design, construct, and optimize functions in plants. While it encompasses plant genome engineering (discussed in detail in the next section), plant synthetic biology extends far beyond genome engineering to include reprogramming gene regulation, constructing novel metabolic and signaling pathways, and integrating synthetic circuits to enable new traits or capabilities in plants. Hence, it has the potential to both rapidly enhance traits in DOE-relevant crops through precise modulation of developmental and metabolic pathways and engineer entirely novel biological functions that expand the utility of plant-derived biomass. Yet, despite impressive scientific advances, the impact of plant synthetic biology remains constrained by a set of core technical and infrastructural barriers. Unlocking the full potential of this field will require coordinated progress across three interconnected priority areas: enabling sophisticated genetic control systems; accelerating biological prototyping; and building the infrastructure needed to support a robust, distributed, and collaborative innovation ecosystem (see Fig. 4.1, p. 18).

Enabling Sophisticated Genetic Control Systems

As discussed in Chapters 2 and 3, engineering traits related to metabolism, biomass growth, nutrient-use efficiency, or stress tolerance remains challenging due to their genetic complexity and the trade-offs they often impose on plant fitness. Many current tools still lack the precision to control when, where, and how genes are turned on or off, often leading to unintended pleiotropic effects. For example, in Arabidopsis a mutation in one gene can affect both flowering time and seed dormancy. To overcome these limitations,

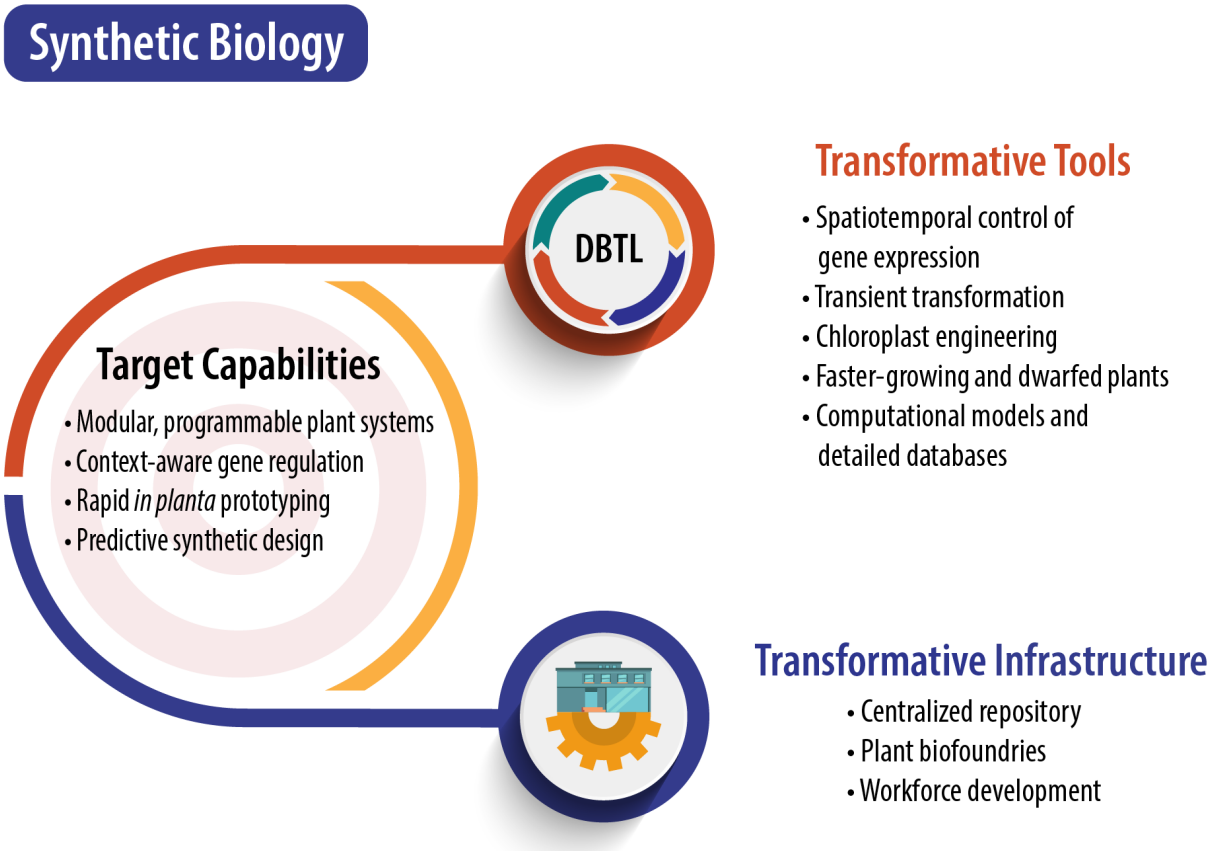


Fig. 4.1. Synthetic Biology for Next-Generation Plant Biodesign. The design-build-test-learn cycle (DBTL) anchors an integrated framework connecting transformative tools and infrastructure to defined target capabilities. Transformative tools—including spatiotemporal control of gene expression, transient transformation platforms, chloroplast engineering, growth-modulated plant systems, and computational modeling—enable modular and programmable plant systems. These advances are supported by transformative infrastructure, such as centralized repositories, plant biofoundries, and workforce development. Together, these components can aid synthetic biology in achieving the key capabilities of context-aware gene regulation, rapid *in planta* prototyping, and predictive synthetic design in plants.

plant synthetic biology must advance genetic control systems that are programmable, tunable, and responsive to environmental or developmental cues.

A compelling example comes from chimeric antigen receptor T-cell therapies, where multilayered control systems enable precise, context-dependent targeting of disease without autoimmunity. Similar strategies in plants could mitigate trade-offs. For example, low-leak inducible systems could be used to activate stress tolerance pathways only in specific tissues and only when needed, such as in response to weather-predicted stress events. These systems must precisely regulate both native and transgene expression across time and tissue types, and they must be compatible with field deployment. They also need to be inexpensive, nontoxic, and stable under fluctuating environmental conditions. Ideally, such systems would also be broadly conserved or minimally host-dependent to ensure interoperability across species.

Examples of recent breakthroughs include re-engineered phytohormone signaling or orthogonal fungal signaling systems (see Technique Spotlight, p. 19) to perceive existing agrochemicals and coupling developmental cues to integrate circuits with tissue-specific promoters. However, improvements are still needed to signal transduction fidelity, target native loci, and translate these

Technique Spotlight

Phytohormone signaling: Phytohormone signaling controls physiological processes in plants, from growth and development to stress responses. Natural hormone pathways can be engineered using a variety of techniques, including genome modification or synthetic chemical and biological applications, to induce morphological or physiological changes.

Orthogonal fungal signaling: Mycorrhizal fungi form a symbiotic relationship with plants involving not only nutrient exchange but also communication with other plants, including warnings about disease and harsh environmental conditions. Tapping into these existing communication networks could enhance warning systems and improve plant robustness.

proof-of-concept studies to feedstock species. Achieving robust spatiotemporal control will be essential for navigating plant genetic complexity and optimizing phenotypes tailored to DOE's mission needs.

Accelerating Biological Prototyping

Having the ability to precisely control gene expression is only half the battle. Testing and refining genetic constructs in crops remain painfully slow and resource-intensive. As noted in Chapter 2, unlike microbes, plants have long generation times, require significant space and care, and develop slowly. These factors severely constrain the pace of innovation.

Faster-growing and compact versions of DOE-relevant crops, akin to mini-maize or micro-tom tomato, could compress growth cycles and reduce space constraints. Transient expression systems and viral vector platforms, which allow gene function to be assessed without stable transformation, have accelerated prototyping in model systems and should be expanded to bioenergy crops. However, cell cultures differ metabolically from whole plants, and the transient nature of these approaches makes studying complex environmental impacts challenging. Some transient approaches include traditional tools like virus-induced gene silencing, *Agrobacterium* infiltration in *Nicotiana benthamiana* leaves, and hairy root cultures, as well as newer platforms like ViparíNama and graft-based transgene delivery (see Technique Spotlight, p. 20). Recently, virus-induced gene editing has successfully achieved tissue culture–free editing in grasses, such as wheat (see Technique Spotlight, p. 20; Qiao et al. 2025).

Developing species- and tissue-specific plant cell lines that mimic *in planta* physiology—and can regenerate into whole plants—would be transformative. These systems could dramatically increase throughput and enable recovery of valuable genotypes from screens. However, care must be taken, as results from protoplasts or transient systems do not always translate to stable lines.

In parallel, expanding engineering capacity in other genetic compartments—particularly chloroplasts—offers underexplored opportunities. Chloroplasts are attractive due to their central role in photosynthesis and their potential to translate bacterial synthetic biology innovations into a plant chassis. Examples of these innovations include programable metabolic modules, such as engineered carbon-fixation pathways, nitrogenase complexes, or high-flux terpenoid and alkaloid biosynthetic circuits originally developed in bacterial hosts. However, chloroplast transformation remains technically challenging, particularly in nonmodel systems and in bioenergy feedstocks.

To maximize the utility of these prototyping tools, they must be integrated with computational models capable of predicting genetic part behavior across species, simulating regulatory networks, and forecasting plant responses under diverse environmental

Continued on p. 21

Technique Spotlight

***Agrobacterium* delivery and infiltration:** *Agrobacterium tumefaciens* is a soil bacterial species that can transfer DNA to host plants via tumor-inducing (Ti) plasmids. Researchers can alter the Ti plasmid to carry genes of interest instead of tumor-inducing ones.

The *Agrobacterium* delivery method infects plant tissues that are then regenerated into stably transformed plants. *Agrobacterium* can be infiltrated into plant leaves using vacuum or syringe methods to transiently express genes carried on its transfer DNA (T-DNA). Upon infiltration, the bacteria release their engineered T-DNA, which enter plant cell nuclei where the genes they carry are expressed. Although some cells may become stably transformed (the T-DNA gets inserted into the plant's genome), transient transformation is more frequent. The infiltration method is commonly used in virus-induced gene silencing studies (further described below).

Graft-based transgene delivery: Graft-based transgene delivery enables transgene-free editing by grafting a wild-type scion (shoot) onto a transgenic rootstock. The transgenic rootstock is engineered, by traditional means, to produce mobile Cas9 and guide RNA (gRNA) transcripts fused to transfer RNA-like sequences. These RNAs move across the graft junction into the scion, where Cas9 is translated and complexed with the gRNA to perform genome editing without integrating foreign DNA into the scion genome. This approach can generate heritable edits in a single generation, while avoiding tissue culture and regeneration steps apart from the initial engineering of the rootstock.

Virus-induced gene silencing: Virus-induced gene silencing (VIGS) is a method of analyzing gene function in plants that leverages the plant's natural antiviral RNA interference (RNAi) pathway. During viral replication, double-stranded RNA (dsRNA) forms and triggers the plant's silencing machinery. If a virus has been engineered to express a fragment of a plant gene, the resulting dsRNA includes that host sequence. The plant's RNAi pathway then degrades RNAs with matching sequences, thereby reducing expression of the endogenous plant gene. Any phenotypic changes or symptoms that result from this gene silencing can provide insight into the gene's function.

Virus-induced gene editing: Virus-induced gene editing (VIGE) uses engineered plant viruses to deliver CRISPR reagents into plant cells, analogous to VIGS. The viral vectors replicate and move systemically, enabling rapid distribution of gRNAs throughout plant tissues to induce targeted edits. Because most plant viruses have strict cargo-size limits, VIGE studies often rely on stable transgenic plants that constitutively express Cas9, with the virus delivering only the gRNA. Newer and more compact Cas variants are beginning to enable full viral delivery of the editing machinery. This approach bypasses traditional tissue culture–dependent transformation.

ViparinNama is a VIGE-based platform that repurposes tobacco rattle virus to deliver regulatory single gRNAs that recruit dead Cas9-based synthetic transcription factors. It enables rapid, tunable modulation of gene expression without altering DNA for fast functional assays and transient rewiring of gene networks.

Continued from p. 19

conditions (see Ch. 5: Harnessing the Power of Automation, Modeling, and Artificial Intelligence to Aid Plant Design and Engineering, p. 31). Training such models effectively will require purposeful generation of well-annotated genotype–phenotype–environment datasets in DOE-relevant feedstock crops.

Building Infrastructure for a Collaborative Innovation Ecosystem

Scaling plant synthetic biology to meet the needs of the bioeconomy will require more than tools—it will require infrastructure that supports a coordinated, distributed research ecosystem. For example, a major bottleneck is the lack of standardized, shareable genetic parts and associated performance data. A centralized repository—including negative or failed results—hosted within existing or new frameworks or facilities could fill this gap and ensure long-term accessibility and stewardship.

Beyond data and tools, a national network of plant biofoundries equipped for construct design, automated transformation, high-throughput (HTP) phenotyping [e.g., Bellwether Foundation Phenotyping Facility at the Donald Danforth Plant Science Center and Oak Ridge National Laboratory's Advanced Plant Phenotyping Laboratory (APPL)], and standardized field testing would dramatically lower barriers to innovation in this field. These facilities could serve academic, government, and private-sector researchers alike. Additionally, having genetic innovations tested in a centralized facility under uniform conditions and protocols would enable more reliable comparison of engineering approaches.

4.2 Plant Genome Engineering

To meet DOE's mission needs in bioenergy and biotechnology, plant genome editing must advance beyond simple mutagenesis to precise, scalable, and programmable genome engineering. Although CRISPR technologies have enabled efficient gene disruption, challenges remain in achieving reliable precision editing, controlled DNA insertion, and species-independent delivery. Investments in improved editing tools, targeted integration strategies, and scalable functional genomics platforms will accelerate the development of environment-adapted and energy-efficient crops optimized for biomass production and environmental performance (see Fig. 4.2, p. 22).

Expanding the CRISPR Toolbox for Precision Editing

CRISPR-Cas systems have made it possible to efficiently target specific chromosomal loci through base-pairing between a guide RNA (gRNA) and the intended DNA sequence. While many variants have been developed for diverse applications, CRISPR-Cas9 from *Streptococcus pyogenes* remains the most widely used. In its canonical form, Cas9 introduces double-strand breaks (DSBs), enabling targeted mutagenesis through error-prone non-homologous end-joining (NHEJ). Multiplexed editing is feasible with multiple gRNAs, and expanded genetic outcomes are possible by fusing Cas9/Cas9 nickase/nuclease-deficient Cas9 (dCas9) to effector domains, such as deaminases, reverse transcriptases, transcriptional regulators, polymerases, or other DNA-modifying enzymes.

While mutagenesis through imprecise NHEJ has been broadly implemented in tractable species, precise sequence alteration remains a key challenge. Base editors coupling dCas9 or Cas9 nickase to deaminases enable adenine-to-guanine and cytosine-to-thymine transitions without inducing DSBs. However, their narrow mutational spectrum limits their utility for biological studies. Prime editors, which use a reverse transcriptase to write new sequences at a targeted site, hold promise, particularly in monocots where they have been successfully deployed. However, editing efficiencies remain too low for routine use in most plant systems. Recent efforts have focused on developing computational and experimental protocols to enhance targeting and editing efficiency. Efficient, broadly applicable tools for precision sequence modification represent a high-priority area for future investment.

A major limitation in current plant genome engineering methods is the lack of control over transgene integration. Standard delivery methods, such as *Agrobacterium*

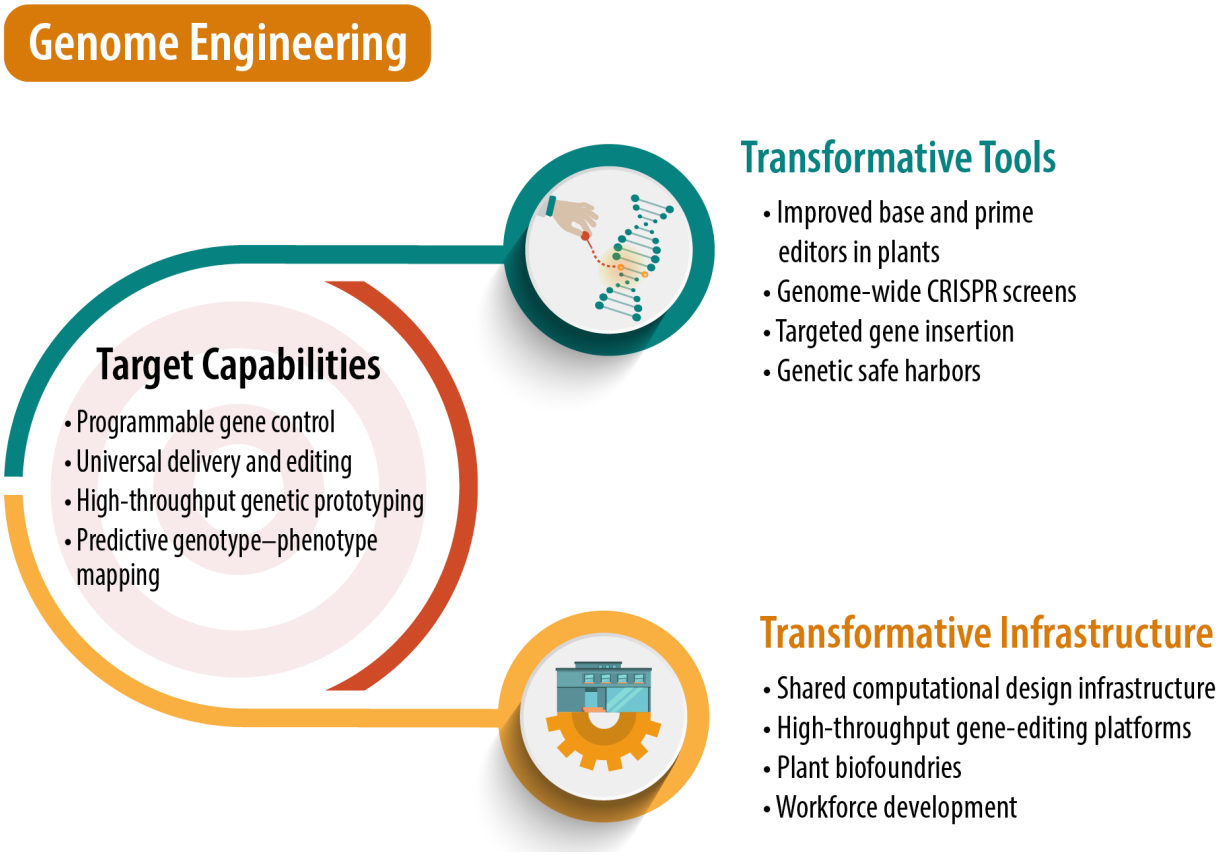


Fig. 4.2. Genome Engineering for Next-Generation Plant Biodesign. An integrated ecosystem of transformative tools and infrastructure is required to achieve programmable and predictive genome engineering in plants. Transformative tools—including improved base and prime editors, genome-wide CRISPR screens, targeted gene insertion technologies, and defined genetic safe harbors—enable precise and scalable manipulation of plant genomes. These advances are supported by transformative infrastructure, such as shared computational design resources, high-throughput gene-editing platforms, plant biofoundries, and workforce development. Together, these elements can drive the target capabilities of universal delivery and editing, high-throughput genetic prototyping, and predictive genotype-to-phenotype mapping.

or biolistics (see Technique Spotlight, p. 23), result in random integration, requiring screening of many independent events to identify those with desired expression. In mammalian systems, many genomic safe harbors have been identified to support consistent transgene expression. Yet only a few examples exist in plants, such as rice. A comparable site-specific transgene insertion platform is a critical unmet goal in plants. Targeted insertion through homology-directed repair (HDR) remains impractically inefficient, while NHEJ-based insertion has low and variable success rates. Site-specific recombinases, which are common in mammalian systems, are underutilized in plants. However, recent work has started to make progress in these technologies for use in plant genome engineering. For example, PrimeRoot editors can enable precise DNA insertions of up to 11 kilobases into plant genomes (Sun et al. 2024.) In addition, transposase-assisted target-site integration methods, such as the CRISPR-associated transposases, offer a promising solution to targeted DNA insertion, but more research is needed to develop robust, species-independent protocols for plants. One recent example fused the rice Pong transposase protein to the Cas9 or Cas12a programmable nucleases and demonstrated sequence-specific targeted insertion in Arabidopsis and soybean (Liu et al. 2024).

Technique Spotlight

Biolistics: Biolistics, also known as particle-mediated gene transfer, is a physical method of delivering DNA into the cell. In this system, a gene gun accelerates dense metal particles coated with DNA at high velocity, so they can penetrate the cell wall and cell membrane. Once inside the cell, the DNA may integrate into the nuclear genome. This method can also deliver DNA into organelles such as chloroplasts and mitochondria.

Biolistics can deliver linear or plasmid DNA of variable size and is not constrained by transfer DNA borders. Biolistics can also deliver messenger RNA or proteins, enabling applications beyond stable transformation. This delivery method can be applied to a wide range of tissues and species, including those recalcitrant to *Agrobacterium*-mediated transformation. However, it can lead to DNA fragmentation and integration of vector backbone into the host.

Advancing a New Frontier in Plant Engineering with Synthetic Chromosomes

Synthetic plant chromosomes represent a transformative alternative to top-down genome engineering, enabling modular, *de novo* construction of orthogonal genetic platforms that function alongside native genomes. Though still in early stages, these technologies could support the stable and heritable introduction of entire biosynthetic pathways—for example, enabling plants to fix atmospheric nitrogen, produce high-value biochemicals like rubber or industrial polymer precursors, degrade environmental pollutants, or recover critical minerals. Synthetic chromosomes could also carry programmable gene circuits that control development or stress responses with high spatial and temporal precision, such as drought-responsive yield enhancement or light-triggered flowering. By insulating genetic modules from native chromatin and regulatory interference, synthetic chromosomes would allow more predictable expression and facilitate trait stacking without unintended interactions. In the longer term, these platforms could support synthetic epigenomes for tunable gene regulation, *in planta* DNA computing for environmental sensing and decision-making, and even adaptive chromosomes that evolve under selective pressure to optimize plant performance in changing field conditions. This technology has the potential to radically accelerate the design of environment-resilient, high-performing crops tailored for bioenergy, carbon storage, and robust biomanufacturing (see sidebar, Enabling Modular, Predictable Plant Design, p. 24).

The most progress in this direction has been made in the yeast model *Saccharomyces cerevisiae*, and similar approaches are being applied to moss, *Physcomitrella patens*. However, the necessity of efficient homologous recombination limits the application of similar approaches to bioenergy crops, which is a current major challenge for plant genome engineering. An alternative is the development of synthetic chromosomes using synthetic centromeres or reappropriation of existing episomal elements. While these approaches have shown promising initial results, progress has been bottlenecked by inefficient delivery of these chromosomes, unstable maintenance over time, and failure of post-delivery plant regeneration. Future research to develop standardized, scalable technologies for targeted DNA integration will enable precise trait deployment, accelerate breeding pipelines, and support the construction of synthetic plant chromosomes.

Enabling Scalable Functional Genomics in Plants

To fully leverage gene editing for functional genomics and crop improvement, HTP platforms are essential. In

Enabling Modular, Predictable Plant Design

Synthetic chromosomes offer a controlled genomic canvas to uncover fundamental design rules of gene expression while enabling robust, modular plant engineering.

Design Rule Discovery

- Test how repetitive DNA influences transcription and translation, including when repetition has no effect.
- Evaluate codon choice, combinations, and positional effects within genes.
- Probe genomic context effects, including proximity to telomeres or ribosomal RNA loci, on gene activity and stability.
- Interrogate 3D genome organization and its influence on pathway output.

Plant Engineering Advances

- Introduce new functions while reducing unintended interactions with endogenous genes.
- Enable stepwise dissection of function by adding or removing genes and regulatory or noncoding elements.
- Create dedicated genomic space and safe harbors to colocate pathway genes, simplifying gene stacking and DNA transfer.

Basic Research Advantages

- Cluster entire biosynthetic pathways into single, modular loci to simplify analysis and redesign (e.g., nitrogen fixation pathways).
- Streamline advanced genome engineering, including large-scale recoding and genetic code expansion for unnatural amino acids.
- Define minimal functional genomes and identify genes required for specific traits such as stress tolerance.
- Decouple pathway behavior from native chromatin noise to improve mechanistic insight.

Breeding and Deployment Impacts

- Enable breeders to move complex traits as single genetic units rather than many dispersed loci.
- Improve trait stability, traceability, and predictability across generations and environments.
- Support genetic containment by reducing recombination with native chromosomes.

mammalian systems, genome-wide CRISPR screens—often conducted in cultured cells using pooled libraries of gRNAs—enable systematic interrogation of gene function by linking gene perturbations to phenotypic outcomes, such as cell survival, morphology, or reporter activity. For example, CRISPR knockout screens have been used to identify genes required for drug resistance or immune signaling. Such approaches remain largely infeasible in plants due to low transformation efficiency; reliance on tissue culture; and the lack of scalable, cell-based assay platforms. However, the development of versatile delivery methods could enable direct editing or gene repression in intact tissues or protoplast populations without stable integration. These advancements would open the door to pooled CRISPR screens in plant cells or tissues, allowing genome-scale perturbation libraries to be screened for traits like stress tolerance, metabolic output, or developmental changes, thereby dramatically accelerating functional genomics and trait discovery in crops.

4.3 Plant Transformation

Plant transformation is foundational to crop biodesign. Yet, even after more than four decades of research, major challenges persist, especially in terms of time, cost, and the limited range of transformable species,

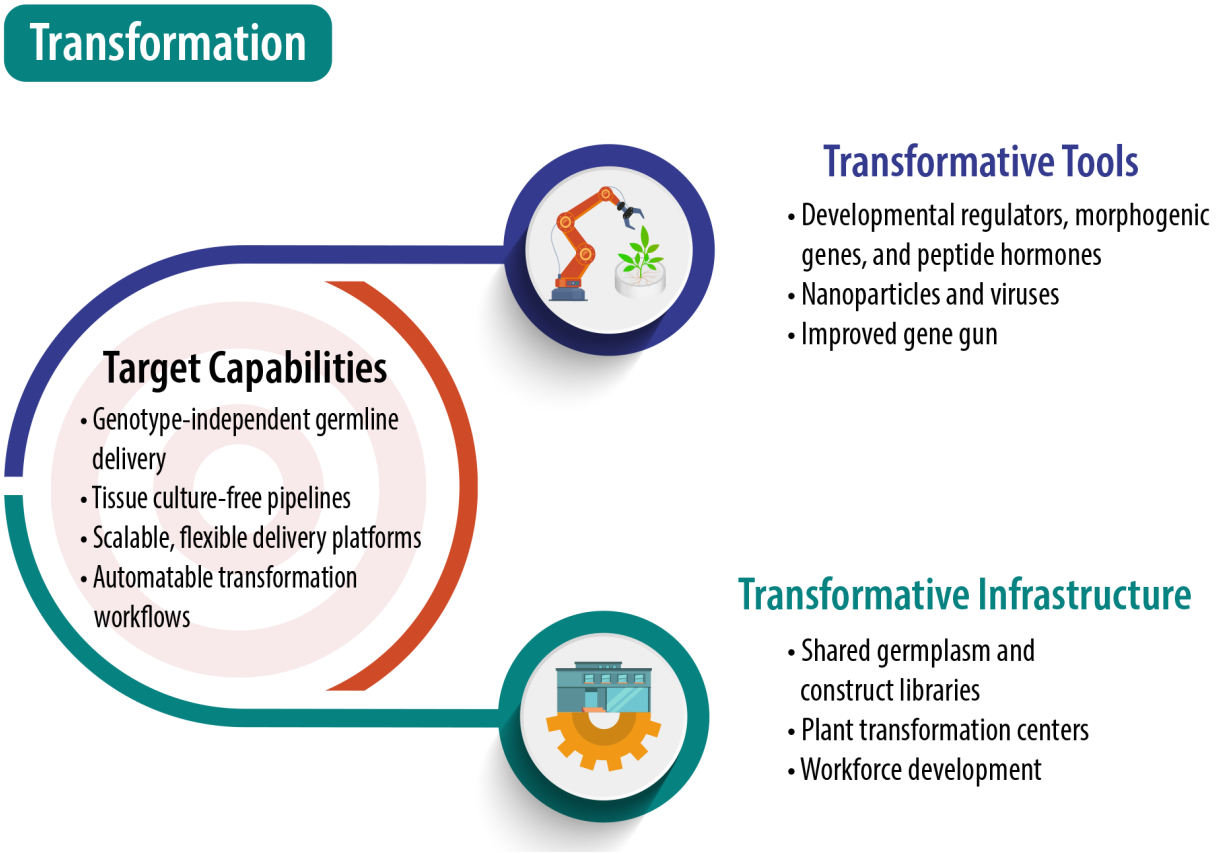


Fig. 4.3. Transformation for Next-Generation Biodesign. Achieving scalable, genotype-independent plant transformation requires coordinated advances in tools and infrastructure aligned with defined target capabilities. Transformative tools—including developmental regulators and morphogenic genes, peptide hormones, nanoparticle- and virus-based delivery systems, and improved gene gun technologies—expand the range and efficiency of DNA and cargo delivery across species and genotypes. These innovations are supported by transformative infrastructure, such as shared germplasm and construct libraries, dedicated plant transformation centers, and workforce development. Together, these elements can enable the target capabilities of genotype-independent germline delivery, streamlined tissue culture pipelines, scalable and flexible delivery platforms, and automatable transformation workflows.

as highlighted in a recent DOE workshop report (U.S. DOE 2024; genomicscience.energy.gov/plant-transformation). Transformation is also limited by reliance on labor-intensive tissue culture systems. These challenges restrict scalability and slow the design-build-test-learn (DBTL) cycle in crop engineering. Overcoming these limitations will require more broadly applicable and scalable delivery and regeneration strategies to support efficient genome modification across diverse crops. Such advancements are critical to not only improving plant transformation but also unlocking the full potential of synthetic biology and genome engineering in plants (see Fig. 4.3, this page).

Advancing Scalable and Genotype-Independent Plant Transformation

Current transformation methods largely rely on tissue culture, where DNA is introduced into individual cells that are subsequently regenerated into whole plants. This process suffers from low efficiency and strict genotype dependence. A major breakthrough has been the use of morphogenic genes and developmental regulators to induce more efficient regeneration and overcome genotype recalcitrance. Still, further gains in efficiency and broader applicability across both monocots and eudicots are needed. More recently, several

Technique Spotlight

Cut-dip-budding: Cut-dip-budding is a tissue culture–free plant transformation system for species with strong root suckering or root-derived shoot regeneration capacity. In this approach, roots or root-adjacent tissues are cut and directly dipped into an *Agrobacterium rhizogenes* solution. The transformed roots are then segmented, and each root segment can regenerate (or sucker) into a stably transformed plant.

peptide hormones have been identified to enhance regeneration, warranting further studies to elucidate their genotype dependence and impact on transformation pipelines. In addition to advancements in the regeneration process, several recent improvements have been made to *Agrobacterium tumefaciens* itself, such as the development of auxotrophic and higher-copy-number variants.

An ideal transformation system would directly deliver biomolecules to the plant germline, bypassing tissue culture (e.g., the current floral dip approach with *Arabidopsis thaliana*). For gene editing in particular, transient delivery without integration is a key objective. RNA viruses have shown promise for germline delivery of genome-editing reagents, enabling heritable and tissue culture–free genome engineering. For instance, Weiss et al. (2025) recently engineered tobacco rattle virus to carry the compact RNA-guided TnpB enzyme ISYmu1 and its guide RNA. This advancement allows for transgene-free editing of *A. thaliana* in a single step, with edits inherited in the subsequent generation. Yet, plant host range and cargo size capacity can still be further improved. Similarly, various nanomaterials, such as carbon- and peptide-based nanoparticles, have been developed for transient expression of genetic cargoes without genome integration in plants, with additional capabilities of chloroplast and mitochondria targeting. Efforts are now underway to adapt these systems for stable and heritable plant transformation. For both viral and nanoparticle-based approaches, achieving efficient, genotype-independent methods that are easy to implement will be essential for broad adoption.

Expanding Scalable Platforms and Transformation Workflows

In addition to viral and nanoparticle-mediated delivery, several other strategies are being developed to improve the scalability and accessibility of plant transformation. Advances in hairy root transformation using improved *Rhizobium rhizogenes* strains, *in planta* delivery of reagents into shoot apical meristem, and optimized protoplast regeneration systems all show promise but require broader translation across plant species and further validation. Emerging methods such as "cut-dip-budding" also warrant deeper investigation (see Technique Spotlight, this page; Cao et al. 2023). The use of visual and seed-sorting markers (e.g., the RUBY reporter system and the Fast Red system) offers a powerful tool for early identification of transformed cells and tissues, potentially enabling automation and increasing the throughput of plant transformation workflows. Finally, the development of long-lived cell cultures with low maintenance—akin to immortal animal cell lines—will be groundbreaking to increase the scale and throughput of plant science and engineering. Although a few suspension culture plant lines are currently available, the range of species needs to be expanded, and the toolkit to transform and work with these suspension cultures needs to be further developed.

4.4 Phenotyping Engineered Crops

Phenotyping links engineered genetic variation to functional performance in plant biodesign, yet as transformation and genome engineering capabilities accelerate, scalable trait measurement—particularly

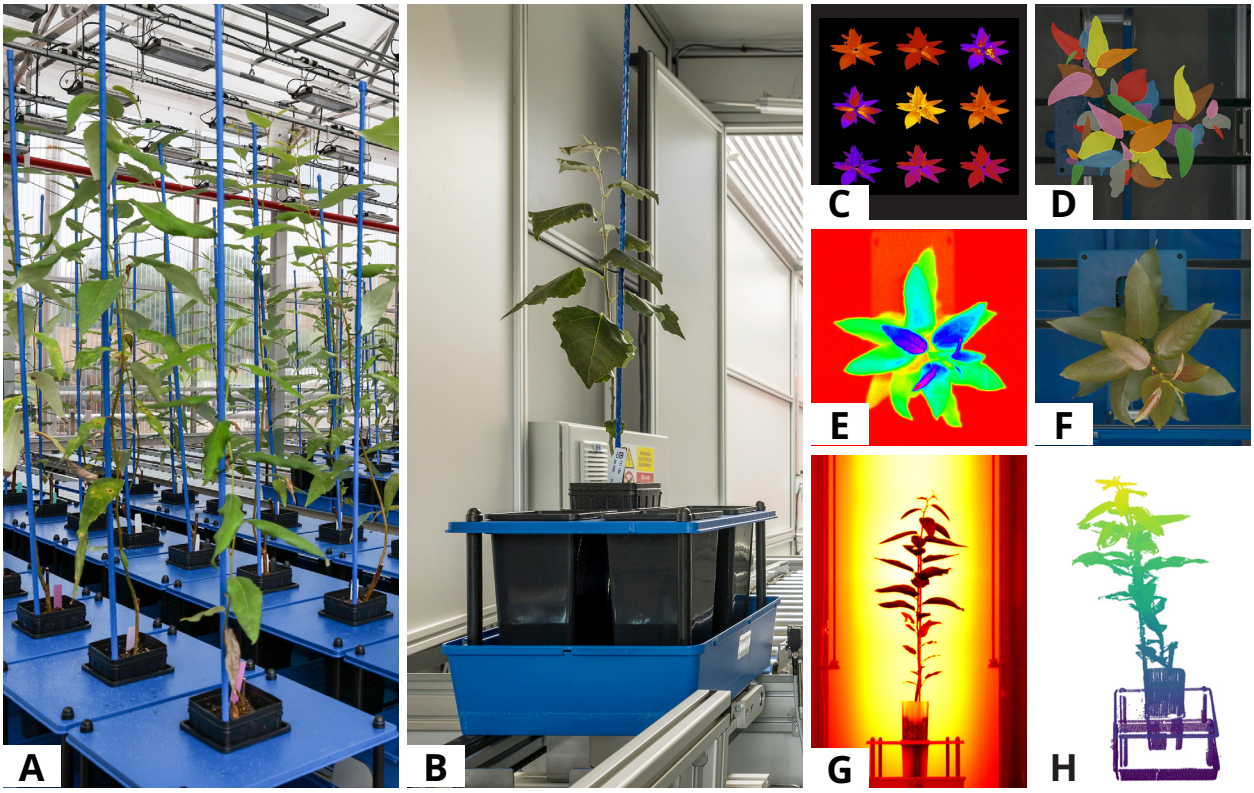


Fig. 4.4. Advanced Plant Phenotyping Laboratory. The Advanced Plant Phenotyping Laboratory (APPL) at Oak Ridge National Laboratory **(A)** offers one of the most diverse suites of imaging capabilities in the world. APPL uses robotics and sensors to automatically move plants through multiple imaging stations **(B)** to measure, or phenotype, plant characteristics using a broad spectrum of light from invisible ultraviolet to visible colors to infrared **(C, D, E, G)**. It can also create digital twins **(F)** for precise growth tracking and generate 3D models **(H)** to analyze plant morphology and growth dynamics. As one of only a few automated phenotyping platforms in existence, APPL's state-of-the-art imaging capabilities are in high demand and access to the facility is limited. [Courtesy Oak Ridge National Laboratory]

under realistic field conditions—has become a central bottleneck. Current approaches remain labor-intensive and environmentally constrained, limiting throughput and predictive capacity. Overcoming these barriers will require more scalable measurement strategies and better integration of data across experimental systems.

Advancing Scalable and High-Throughput Phenotyping

Downstream of transformation, a major bottleneck lies in phenotyping engineered plant lines—particularly under realistic field conditions, where scalability remains a key challenge for many crops. Several critical limitations and opportunities for innovation exist. First, tools for assessing physiological traits (e.g., chlorophyll content and carbon dioxide assimilation) are often labor-intensive and costly. High-throughput, early-stage markers are needed to enable rapid selection and reduce the need to grow plants to maturity. Second, integrating genotype-to-phenotype relationships requires collecting and interpreting large-scale multiomics datasets, including physiological, metabolomic, and proteomic profiles. Third, very few automated phenotyping platforms have been developed—especially for tall or nonmodel plants—and access is limited to existing facilities, such as Danforth's Bellwether and APPL (see Fig. 4.4, this page). Furthermore, most controlled growth environments poorly mimic field-relevant conditions, limiting their utility for feedstock development. Clonally

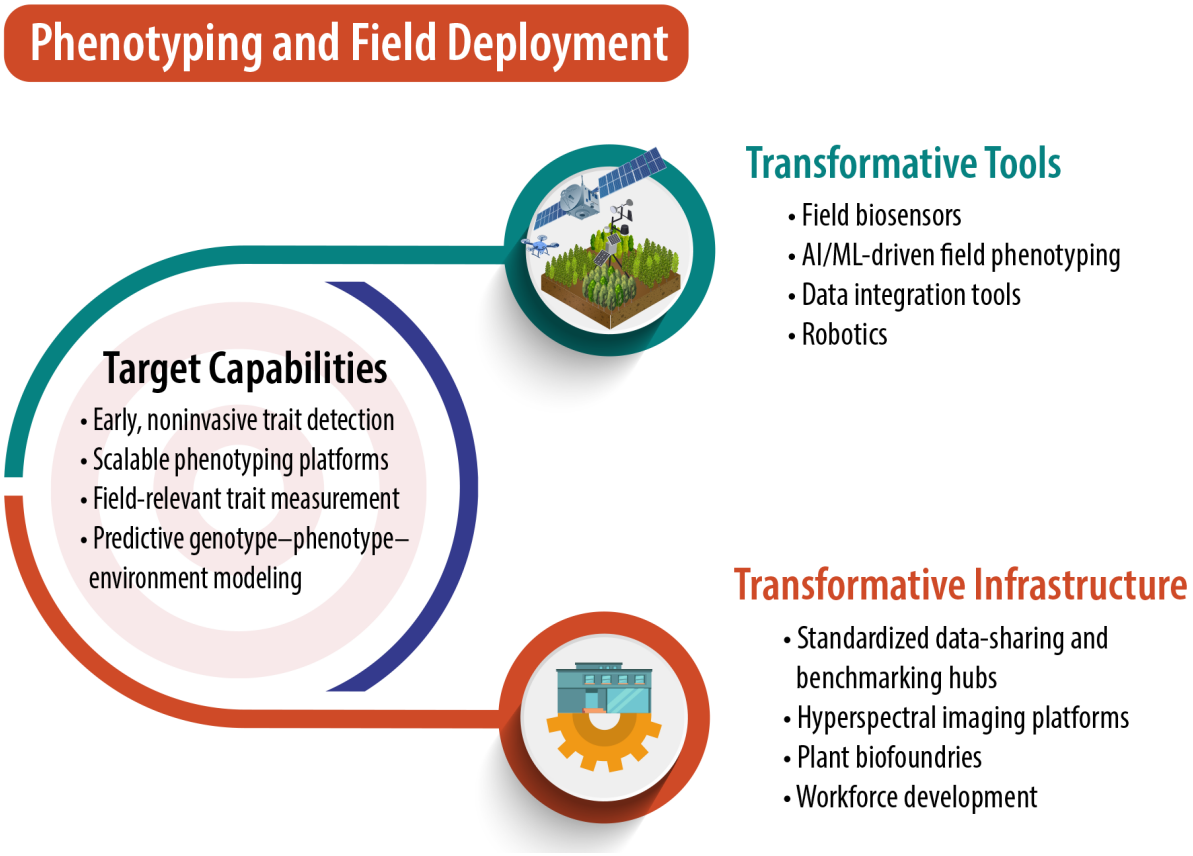


Fig. 4.5. Phenotyping and Field Deployment for Next-Generation Plant Biodesign. Accelerating translation from genotype to field performance requires coordinated development of transformative tools and infrastructure aligned with clearly defined target capabilities. Transformative tools—including field-deployable biosensors, AI- and machine learning (ML)-driven phenotyping, data integration platforms, and robotics—enable continuous and high-resolution monitoring of plant traits in real-world environments. These efforts are supported by transformative infrastructure, such as standardized data-sharing and benchmarking hubs, hyperspectral imaging platforms, plant biofoundries, and workforce development. Together, these components can enable the target capabilities of early and noninvasive trait detection, scalable and field-relevant phenotyping, and predictive genotype–phenotype–environment modeling to guide crop improvement and deployment.

propagated and perennial species introduce further complications due to long generation times and space requirements. Dwarfed or fast-cycling genotypes could play a key role in addressing these limitations.

Scaling Field Phenotyping and Evaluation Infrastructure

Advancements in phenotyping under realistic field conditions are also needed to enable precise, scalable biodesign. Field phenotyping is expensive; limited by regulatory hurdles; and lacks robust, noninvasive tools, such as biosensors or hyperspectral imaging, for monitoring specific crop traits and metabolites. In particular, phenotyping plant–microbiome interactions remains a big challenge. Improved modeling approaches and down sampling based on AI and machine learning could help reduce the scale of required experiments while retaining predictive power. Additionally, the ability to outsource transformation and phenotyping to centralized biofoundry-type facilities equipped with robotics and high-throughput screening platforms would allow researchers to focus more deeply on core scientific questions, accelerating progress in crop biodesign (see Fig. 4.5, this page).

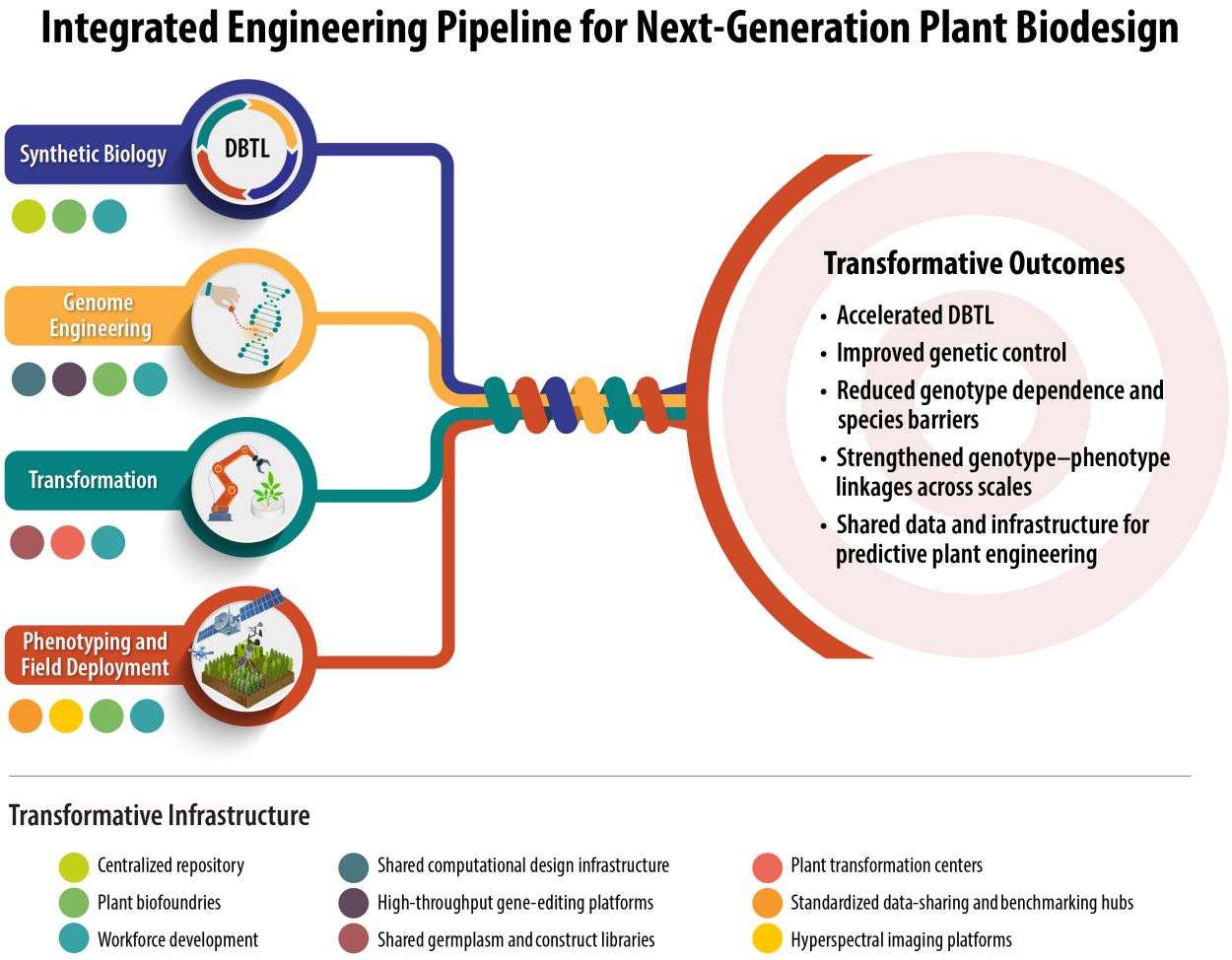


Fig. 4.6. Integrating Synthetic Biology, Genome Engineering, Transformation, and Phenotyping Enables Transformative Outcomes. Synthetic biology, genome engineering, transformation, and phenotyping and field deployment are linked through a coordinated design-build-test-learn (DBTL) cycle to form a unified plant engineering pipeline. Each domain contributes specialized tools and enabling infrastructure, including shared repositories, biofoundries, computational design platforms, gene-editing centers, germplasm libraries, data-sharing hubs, and advanced imaging systems. Their integration converges on transformative outcomes: accelerated DBTL cycles, improved genetic control, reduced genotype and species barriers, strengthened genotype–phenotype linkages across scales, and shared data ecosystems that enable predictive plant engineering.

4.5 An Integrated Engineering Pipeline for Next-Generation Plant Biodesign

Next-generation plant biodesign requires more than isolated technological advances. It demands an integrated, end-to-end engineering framework connecting molecular design, precise genome modification, scalable transformation, and field-relevant phenotyping within a continuous feedback loop (see Fig. 4.6, this page). This pipeline unifies synthetic biology, genome engineering, transformation technologies, and real-world phenotyping under a coordinated DBTL paradigm. By embedding shared computational, biological, and physical infrastructure across stages, the framework reduces fragmentation in the plant engineering

ecosystem and enables systematic iteration rather than one-off innovation. The result is a predictive, scalable, and species-agnostic platform for plant biodesign. By strengthening genotype–phenotype–environment linkages and accelerating iterative learning, this integrated pipeline lays the foundation for programmable plants and data-driven crop improvement at unprecedented speed and precision.

Chapter 5

Harnessing the Power of Automation, Modeling, and Artificial Intelligence To Aid Plant Design and Engineering

Advances in automation, computational modeling, and artificial intelligence (AI) are transforming the possibilities of plant design and engineering. As biological datasets grow in scale and complexity, these tools offer new ways to extract knowledge, generate hypotheses, and accelerate the biodesign cycle. From literature mining and gene function prediction to modeling phenotypes and designing novel biological parts, next-generation AI systems are beginning to enhance every stage of plant research (see Fig. 5.1, p. 32). Realizing their full potential, however, will require purpose-built models, high-quality training data, automated experimentation platforms, and closer collaboration between plant scientists and computational experts. This chapter outlines the emerging capabilities of AI and automation for plant biodesign, describes the critical datasets needed to power these tools, and highlights key barriers that must be addressed to ensure their effective implementation.

From literature mining and gene function prediction to modeling phenotypes and designing novel biological parts, next-generation AI systems are beginning to enhance every stage of plant research.

5.1 AI Tools To Aid Plant Biodesign Research

AI is becoming an essential driver of modern plant biodesign, offering powerful ways to extract knowledge, make predictions, and accelerate the engineering cycle. One of the most immediate applications is literature mining. A large amount of information—domain knowledge and specific design rules for promoter activities and gene and protein functions—is embedded in publications and databases. Open-source large language models (LLMs) trained on scientific literature and curated databases can integrate this knowledge and offer free and easy access to users, thereby augmenting biodesign research. For example, fine-tuning the LLM Llama3-8B on Arabidopsis genes and phenotype data produced PlantGPT, a virtual expert capable of providing knowledge-based output upon user inquiries (Zhang et al. 2025).

AI is also well positioned to help handle large, complex datasets for understanding and prediction in biodesign research. As automated platforms develop key datasets critical for training LLMs, the resulting improved AI

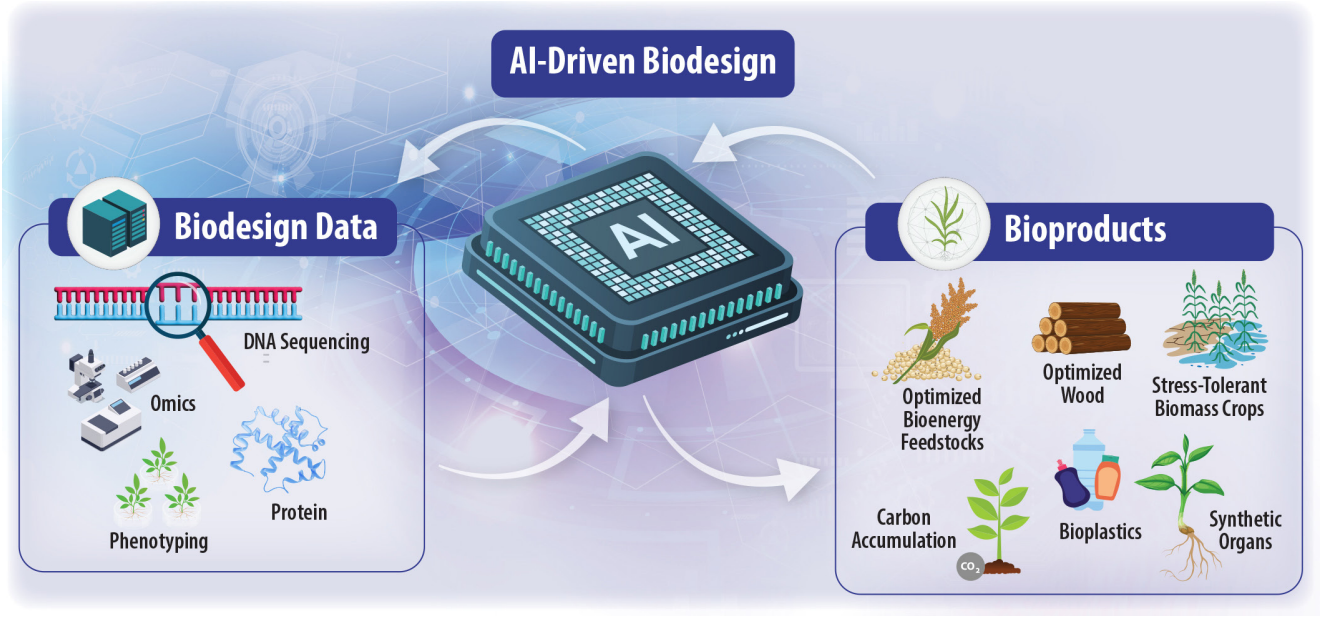


Fig. 5.1. Using AI to Accelerate Next-Generation Plant Biodesign. High-dimensional biodesign data—including DNA sequencing, multiomics profiling, protein characterization, and phenotyping—are integrated through AI to enable predictive design and optimization. Iterative feedback between data generation and AI modeling guides the rational engineering of plants toward diverse bioproduct outcomes, such as optimized bioenergy feedstocks, improved wood traits, stress-tolerant biomass crops, enhanced carbon accumulation, bioplastics production, and synthetic organs. This closed-loop platform illustrates how data-driven modeling accelerates next-generation plant biodesign and sustainable bioinnovation.

tools could be used to help annotate large transcriptomic and proteomic datasets, identifying novel gene functions or regulatory interactions. Because LLMs are predictive by nature, they can propose multiple design options that may warrant experimental testing. LLMs could also enable researchers to transfer trait prediction across species and varieties. New plant DNA-focused models, such as plant DNA LLMs, represent an important step toward building foundational DNA models that can analyze and predict plant genome functions (Liu et al. 2025).

Beyond data analysis, AI tools are increasingly capable of (1) predicting gene function (especially genes and proteins of unknown function) and functional elements of genomes and variant effects; (2) predicting and mapping gene network interactions to chart out components and parts of a network or pathway; (3) integrating omics data from cells to tissues to organs; and (4) forecasting phenotypes. Additionally, computational models, such as *in silico* perturbation tools, would be useful to predict perturbation impacts on molecular and phenotypic levels.

Furthermore, AI can be used to design novel sequences, pathways, and products. Following the 2024 Nobel Prize in Chemistry, which recognized groundbreaking work in AI-driven protein folding and design, more advanced AI tools have been developed to design new and improved proteins. These include structure-based protein clustering (Huang et al. 2024), structure-informed protein language model (PLM; Shanker et al. 2024), and few-shot active learning platforms (e.g., EVOLVEpro) that integrate PLMs (Jiang et al. 2024). Similar efforts could dramatically advance protein design in plants. Additional efforts to further strengthen AI-enabled design include (1) constructing synthetic promoters or other regulatory elements to assist metabolic modeling in gene prediction and pathway editing and (2) developing metabolite data annotation for LLMs (e.g., tokenization of chemical structure data, mass spectrometry

Technique Spotlight

Chemical structure data tokenization: This process converts chemical structure information into small, machine-readable units (tokens) that enable efficient analysis by machine learning and AI models. Because chemical structure encompasses a wide range of features—atomic arrangements, bonding patterns, and stereochemistry—tokenization provides a unified way to encode this structural information, so algorithms can characterize molecules, predict their behavior, and support new molecular design.

Enzyme site geometry: Enzyme site geometry refers to the 3D arrangement of amino acid residues within an enzyme's active site that governs how it binds substrates and catalyzes reactions. By characterizing these spatial patterns, such as residue orientation, distances, and physicochemical environments, computational and machine learning methods can link active-site shape to catalytic function and reaction specificity. Understanding enzyme site geometry improves the prediction accuracy of uncharacterized enzyme activities and supports the discovery or design of more efficient catalysts for building new metabolic pathways.

Digital twins for plant metabolism: Digital twins for plant metabolism are virtual models that replicate a plant's physiological and biochemical state, updated continuously with real-time data from the physical plant. Using sensor inputs and computational modeling, such as functional–structural plant models and metabolic network models, these systems can simulate how a plant's metabolism responds to environmental or genetic changes. Because plants produce a rich array of specialized metabolites with agricultural and industrial value, digital twins offer a powerful framework for analyzing the genes, pathways, and conditions that influence metabolic outputs.

Agentic synthetic biology task design: Agentic synthetic biology task design leverages autonomous AI agents working alongside robotics and computational infrastructure to plan, adapt, and execute laboratory protocols with minimal human intervention. Beyond simple automation, agentic systems can reason about goals and constraints, dynamically modify protocols in response to interim data, and coordinate instruments end to end.

Common synthetic biology procedures—such as construct design and assembly, polymerase chain reaction and cloning, bacterial strain building, and phenotypic characterization—can be orchestrated by agentic workflows to reduce manual labor; increase reproducibility; and generate the large, standardized datasets required for AI-driven analysis. By combining closed-loop experiment planning (e.g., design-build-test-learn cycles) with real-time data assimilation and adaptive decision-making, agentic systems improve throughput and consistency, enabling rapid optimization and evaluation of engineered biological constructs and metabolic pathways within a design-build-work framework.

data, or enzyme active site geometry relative to substrates; see Technique Spotlight, this page). Moreover, generating digital twins for plant metabolism would optimize gene constructs for trait stacking without introducing element incompatibilities, and agentic synthetic biology task design (e.g., cloning strategy design) would help implement biodesign in crops (see Technique Spotlight, this page). Achieving these goals will require multilayered, multiscale datasets that include genomic, transcriptomic, proteomic, and metabolomic information, paired with time-series sensor data on the plant's environment and physical state.

When coupled with automated phenotyping platforms and automated biofoundries—such as Oak

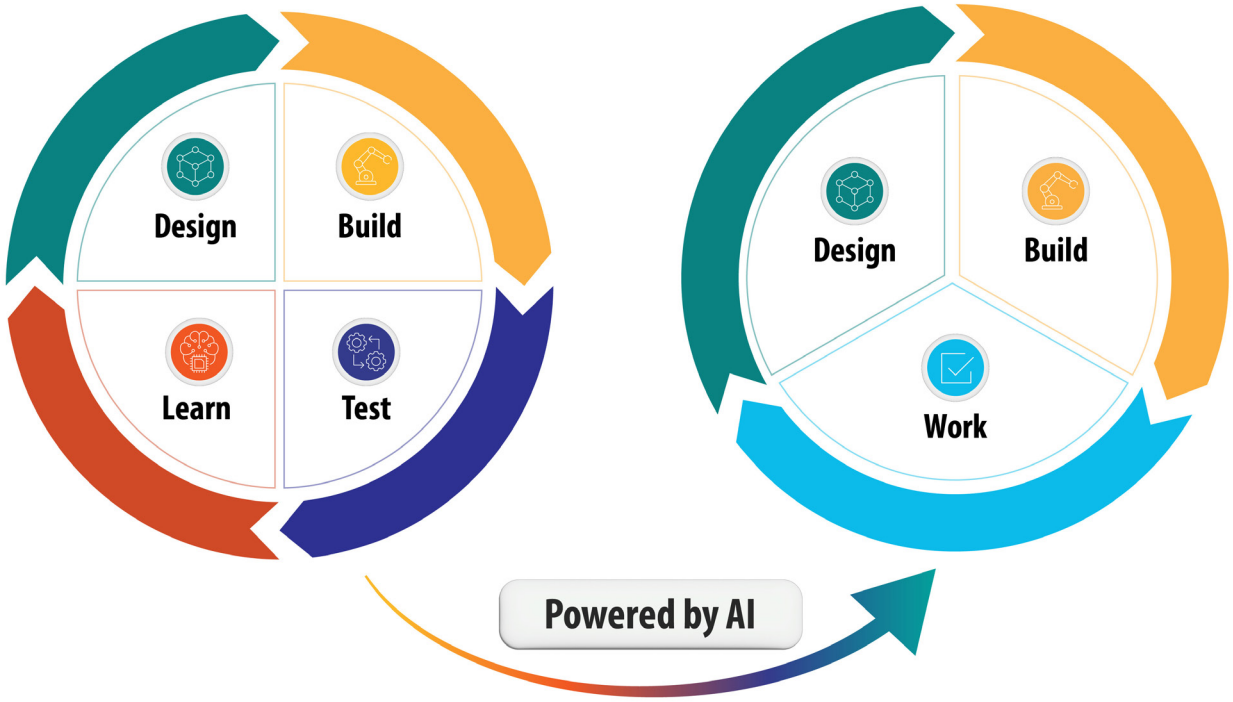


Fig. 5.2. AI-Driven Evolution of the Biodesign Cycle. Traditional design-build-test-learn workflows (left) rely on time- and resource-intensive testing because design quality is limited by available knowledge. As high-quality experimental datasets expand and feed increasingly powerful AI and machine learning models, scientists may be able to predict system behavior with such accuracy that the iterative "test" step becomes far less necessary. In this more streamlined design-build-work cycle (right), designs are generated with the expectation that they will function upon construction, reflecting a transformational shift toward data-driven biodesign.

Ridge National Laboratory's Advanced Plant Phenotyping Laboratory and the fast, automated, scalable, high-throughput pipeline for plant bioengineering (FAST-PB) at the University of Illinois Urbana-Champaign (Dong et al. 2025)—these advanced computational tools are poised to accelerate plant synthetic biology research. Ultimately, these advancements will inspire a new paradigm for plant biodesign, such as streamlining and accelerating the biodesign cycle (see Fig. 5.2, this page).

5.2 Critical Datasets To Enable Trait Engineering with Next-Generation AI Tools

The effectiveness of AI in plant biodesign depends on the availability of large, high-quality, and well-structured datasets. Current tools are often limited by incomplete or inconsistent data formats, missing or low-quality metadata, and weak standardization across studies, underscoring the need for coordinated efforts to generate datasets that can power connections and provide novel inferences. While data standardization and metadata remain important, a key initial use of AI might be to automate accurate metadata generation and harmonize diverse datasets.

Generating training datasets that are relevant to key science questions and can increase plant trait prediction based on genotype and environment will be a key enabler for plant biodesign. A critical step toward this end is developing and accessing large-scale datasets that can connect plant traits to genetic changes for improved development of future crop varieties and cropping systems. An ideal outcome

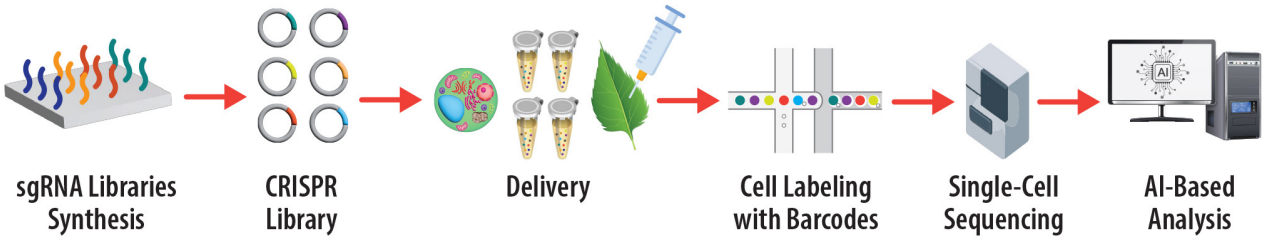


Fig. 5.3. An Illustrative CRISPR-Cas Perturbation-Sequence Pipeline for Scalable Gene Functional Studies in Plants. This figure shows one possible workflow in which single guide RNA (sgRNA) libraries are designed and assembled into CRISPR plasmid libraries to introduce targeted genetic perturbations in plant cells. Perturbed cells are barcoded and profiled by high-throughput single-cell sequencing to measure molecular responses. The resulting perturbation–response datasets can then be analyzed using AI-based models to infer gene function and regulatory relationships at scale. The ordering and implementation of steps may vary depending on experimental design.

could be a set of virtual plant models that can predict trait changes based on genotype, trait interactions, and environment.

Critical datasets for connecting a genotype to a specific phenotype could include many different types of experiments, such as natural diversity screens like genome-wide association studies, massively parallel reporter assays of promoters or coding variants, multiomic cell- and tissue-specific profiles, and genetic perturbation datasets (e.g., perturbation-sequence datasets; see Fig. 5.3, this page). Datasets that improve annotation quality of genomic features and functions can also augment genotype and phenotype connections. Additional critical datasets are also necessary to enable a better understanding of how plant phenotypes are interrelated or influenced by the environment. High-resolution monitoring of temperature, soil, microbes, plant pathogens, and other environmental factors will add key phenotypic knowledge and aid in developing digital twin approaches that can provide insights into expected plant performance across novel environments. Increased trait measurements, including time-series measurements, will further inform how phenotypic shifts or developmental transitions in one trait are related to other traits. Together, the datasets discussed in this section could inform the foundation for next-generation AI tools that can accurately model plant biology; guide trait improvement; and enable predictive, design-driven crop engineering.

5.3 Key Barriers to Broader Implementation of AI-Based Tools for Biodesign

Despite growing enthusiasm for AI-driven innovation in plant biodesign, several significant barriers continue to hinder the broader implementation of AI-based tools in this space. A central challenge is that most current AI tools are designed for generalists or adapted from models built for mammalian or microbial systems. These tools often lack the specificity needed for plant-focused goals, such as precise engineering of traits that involve multigene families, gene regulatory design, or genome editing in diverse crop species. Custom-built models tailored to the intricacies of plant biology—including tissue-specific regulation, noncoding region impacts, and plant–microbe interactions—are urgently needed. Without such specificity, AI applications in biodesign risk generating false positives or biologically irrelevant predictions.

Another major constraint is usability. Many existing AI models are not easily accessible to plant biologists due to steep learning curves, nonintuitive interfaces, or a lack of modularity. To enable wider adoption, AI tools must be more user-friendly and flexible. Modular platforms that allow researchers to plug in their own datasets and ask targeted questions—such as optimizing multigene constructs or predicting phenotypes from genotypes—would significantly

lower entry barriers. Likewise, AI tools capable of using virtual plant models for development and iteration could streamline hypothesis testing and trait prediction, provided they are supported by adequate training datasets.

The development of such AI tools greatly depends on the availability of high-quality, relevant datasets. Unfortunately, there remains a paucity of standardized, large-scale training and characterization datasets that link molecular, physiological, and phenotypic data in plants. Such datasets include multiomics profiles, transcriptomic time series, perturbation assays, and comprehensive phenotype data under diverse environmental conditions. Without these datasets, it is difficult to train robust models or evaluate their accuracy in predicting real-world outcomes.

Compounding these technical challenges is a broader structural issue: the lack of sustained incentives to attract AI talent—particularly from computer science and software development—into plant biology. The plant science community must do more to foster interdisciplinary collaboration, possibly through the creation of a "marketplace of design challenges." Such a platform would allow plant scientists to articulate specific computational needs (e.g., design a promoter that drives expression under drought conditions) and enable AI developers to find and work on these problems. Ultimately, increased collaboration among the plant science community could catalyze innovation, direct talent toward high-impact problems, and help cultivate a generation of AI tools truly designed with plant biodesign in mind.

Chapter 6

Training the Next Generation of Plant Design Scientists

Meeting the grand challenges of plant design—whether for environmental resilience, advanced bioproduction, or energy security—requires a new generation of scientists equipped with diverse and integrative skill sets. As the field continues to evolve, so must its approach to education and workforce development. Preparing future scientists in plant design requires training programs that are intentionally designed to not only promote cross-disciplinary fluency, systems thinking, and design literacy but also cultivate engagement opportunities across sectors of the workforce.

6.1 Cross-Disciplinary Fluency

Future leaders in plant design must be fluent in molecular and cellular biology, plant physiology, omics, field ecology, computational modeling, and engineering principles. They must also embrace the transformative potential of AI as a tool for innovation and intelligent design. Yet, few current training programs are structured to support meaningful integration across these domains. Students and postdocs often receive deep training in one area while remaining peripheral to others that are equally critical to success in plant design and bioengineering.

Bridging this gap will require intentional curriculum design, as well as fellowship and internship programs that foster both technical competency and the collaborative mindset essential for innovation. Training should also prepare scientists to operate in complex, real-world environments, including designing experiments under variable field conditions; deploying sensors; interpreting high-dimensional omics, phenotypic, and environmental data; and collaborating across disciplines and sectors.

Cohort-based training models offer one promising approach. By bringing together interdisciplinary teams of trainees to tackle shared, real-world challenges, these programs can cultivate collaborative problem-solving and a shared vocabulary across fields. Similarly, short-term interdisciplinary bootcamps can provide exposure to complementary skills, such as AI for biologists or field physiology for engineers, helping to bridge knowledge gaps and build confidence in unfamiliar domains.

Training must emphasize systems thinking: the ability to anticipate how a genetic modification might ripple through metabolic trade-offs, developmental pathways, and ecological interactions.

6.2 Systems Thinking and Design Literacy

Plant design is inherently a systems-level challenge. Training must therefore emphasize systems thinking: the ability to anticipate how a genetic modification might ripple through metabolic trade-offs, developmental pathways, and ecological interactions. It also requires design literacy: the capacity to iterate, prototype, and evaluate solutions in the face of uncertainty.

These skills are rarely taught explicitly in scientific training, yet they are essential for navigating the complexity of plant systems and the unpredictability of bioproduction under field environments. Case-based learning, drawing on both successes and failures in synthetic biology and metabolic engineering, can help trainees develop a more nuanced understanding of design trade-offs and constraints. Design studios, capstone projects, and interdisciplinary mentorship can further reinforce these competencies and foster a mindset of thoughtful experimentation.

6.3 Workforce Development and Multisector Engagement

Preparing the next generation of scientists to work across sectors is essential. Many of the most pressing challenges in plant design, such as scaling production and navigating regulatory landscapes, extend beyond the boundaries of academia. Training programs should therefore create opportunities for trainees to engage meaningfully with industry, government, and nonprofit partners.

Industry-sponsored internships offer one such opportunity, providing hands-on experience with real-world challenges and workflows. Joint training programs codeveloped with external partners can help align academic preparation with workforce needs while also exposing trainees to a range of career paths. The goal is to cultivate a generation of scientists who can translate innovation into impact, move between research and application, and contribute to the broader systems in which plant design operates.

Chapter 7

Conclusions

Realizing the full potential of plant systems for bioproduction requires a transformative approach that integrates advances across genomics, metabolism, plant–environment interactions, and data-driven design. Major knowledge gaps in plant metabolic networks currently constrain the ability to rationally design and optimize plants as chemical factories. Overcoming these limitations demands a systems-level understanding of plant biochemical pathways, cellular compartmentalization, and the complex trade-offs between product synthesis and biomass accumulation.

Answering foundational questions—such as which feedstocks or bioproducts can be efficiently produced in specific plant tissues, cell types, or compartments—requires coordinated, interdisciplinary efforts. These include advanced gene expression profiling and genetic engineering toolkits, high-throughput imaging and metabolic platforms, and AI-enabled modeling tools capable of predicting and optimizing complex plant phenotypes under diverse environmental conditions.

A key frontier also lies in understanding and engineering plant–environment interactions. Environmental variability, long viewed as a constraint, should instead become a tunable design parameter. Achieving this will require the development of robust field-testing infrastructure, modular sensing and control systems, and training platforms that foster collaboration across biology, engineering, and computation. Standardized tools and shared datasets will enable reproducibility and scalability, allowing predictive design principles to emerge.

Plant synthetic biology, genome engineering, transformation, and phenotyping sit at the heart of this integrated vision. It provides a strategic framework for building a resilient, domestic bioeconomy that leverages plants as programmable platforms for sustainable fuel, material, and bioproduct synthesis. However, the current scientific foundation must be matched by a bold, coordinated investment in translational infrastructure, community-scale tool development, and scalable manufacturing systems.

Also central to this vision is the integration of large, multimodal datasets that link genotype to phenotype, while accounting for environmental and management interactions. Investments are needed in data generation and interoperability and the development of plant-specific AI and machine learning tools. Collaborative challenge platforms can help bridge the divide between experimental biology and computational design, accelerating innovation at the interface of data science and synthetic biology. Realizing this vision will also require training a new generation of scientists equipped with cross-disciplinary fluency and systems-level design skills.

Together, these efforts will enable a new generation of rationally engineered crops tailored for environmental resilience, bioproduct yield, and scalability—supporting a robust U.S. bioeconomy and global leadership in plant-based solutions.

Realizing the full potential of plant systems for bioproduction requires a transformative approach that integrates advances across genomics, metabolism, plant–environment interactions, and data-driven design.

Appendix A

Workshop Agenda

All times Eastern

Day 1: Wednesday, March 12, 2025

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**

Speaker: Pablo Rabinowicz, Program Manager, Genomic Science Program

12:20 p.m. **Introduction to the Workshop and Logistics**

Speakers: Gözde S. Demirer (Workshop Co-Chair), California Institute of Technology; Yiping Qi (Workshop Co-Chair), University of Maryland; and Chung-Jui (CJ) Tsai (Workshop Co-Chair), University of Georgia

12:30 p.m. **Plenary Presentation: Challenges in Moving Cutting-Edge Science to Growers' Fields**

Speaker: Nathan Springer, Bayer Crop Sciences

(30 minutes; 10 minutes Q&A)

1:10 p.m. **Summary of Brainstorming Responses**

1:30 p.m. **Breakout 1: Plant Chemical Factories for Biofuels, Biomaterials, and Bioproducts**

This session will explore research opportunities and technical innovations aimed at harnessing nature's genetic potential to engineer plants to synthesize advanced biofuel precursors, organic–inorganic composite materials, carbon-storage materials, biominerals, bioplastics, or living materials with new-to-nature properties and functionalities. Participants will consider both experimental and computational perspectives as well as continuously evolving omics capabilities and resources.

2:15 p.m. **Break**

2:35 p.m. **Breakout 1 Report-Out**

3:05 p.m. **Breakout 2: Bioengineered Feedstock Design for Environmental Resilience**

Basic science advancements toward a circular bioeconomy with plant chemical factories must also address plant interactions with the natural environment. This session will focus on knowledge gaps and roadblocks to the success of bioengineered

feedstock under field conditions. The discussion could include, but is not limited to, biodesign considerations related to water and nutrient limitations; other abiotic stresses; their effects on cell wall architecture, sensing, and macromolecule engineering; efficient carbon harvesting and channeling toward designer pathways; and perenniality. Participants will consider both experimental and computational perspectives as well as continuously evolving omics capabilities and resources.

3:50 p.m. **Breakout 2 Report-Out**

4:20 p.m. **Closing Remarks**

Day 2: Thursday, March 13, 2025

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

Speaker: Gözde S. Demirer, California Institute of Technology

12:10 p.m. **Breakout 3: Research Opportunities to Fill Key Knowledge Gaps in Plant Genome Engineering**

This session aims to identify strategies for advancing plant genome engineering and improving the plant design process. Participants will explore current limitations in the genome engineering toolkit, spanning scales from target selection to organismal design. Emerging technologies to be discussed include epigenome editing, large-scale CRISPR screens, synthetic chromosomes, synthetic genomes, and methods for targeted integration of large DNA molecules to enable innovative plant biodesign. The session will also focus on leveraging cutting-edge molecular engineering tools in combination with omics data and predictive models driven by artificial intelligence (AI) to enhance the precision and efficiency of plant engineering.

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

1:25 p.m. **Breakout 4: Advancing Plant Synthetic Biology for Transformative Feedstock Engineering**

This breakout session focuses on the current landscape and future potential of plant synthetic biology, emphasizing the development of advanced tools and standardized parts for engineering regulatory circuits, polygenic traits, metabolic pathways, and more. The discussion aims to tackle key technical challenges and explore strategies to enhance the precision, scalability, and efficiency of synthetic biology applications in feedstock species. Participants will also brainstorm how automation and cutting-edge computational technologies can drive innovation and accelerate progress in plant synthetic biology.

2:10 p.m. **Break**

2:30 p.m. **Breakout 4 Report-Out**

3:00 p.m. **Breakout 5: Plant Transformation Innovation, Characterization, and Phenotyping**

Plant transformation is a cornerstone of the biodesign and engineering workflow, yet significant challenges persist despite over four decades of research, hindering

progress in feedstock crop development. Building on insights from a previous DOE workshop, this breakout session will focus on recent innovations in plant transformation technologies and address bottlenecks in downstream processes, such as the characterization and phenotyping of bioengineered lines. Participants will discuss how to overcome these obstacles and advance high-throughput approaches to accelerate bioengineered feedstock design.

3:45 p.m. **Breakout 5 Report-Out**

4:15 p.m. **Closing Remarks**

Day 3: Friday, March 14, 2025

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

Speaker: Yiping Qi, University of Maryland

12:10 p.m. **Breakout 6: Harnessing the Power of Automation, Modeling, and AI to Aid Plant Design and Engineering**

This session aims to identify strategies for applying automation, modeling, and AI to aid the biodesign of feedstock species. It requires significant time, cost, and effort to practice design-build-test-learn (DBTL) cycles in crop engineering. In the meantime, automation and modeling have been widely used in many engineering fields. In recent years, discoveries and innovations in biology have been fueled by machine learning and AI, such as protein design and functional prediction of DNA elements. Participants will explore the status and limitations in the cross-cutting interface between wet and dry laboratory research and seek promising strategies to boost DBTL by incorporating computation and AI in the process.

12:55 p.m. **Breakout 6 Report-Out**

1:25 p.m. **Breakout 7: Enabling Efficient, Integrative Multitrait Engineering**

This session aims to identify problems and opportunities in integrating multiple traits into feedstock crops. Engineering a desirable trait cannot compromise other traits in the crop. It is more desirable to simultaneously design multiple traits in the target plant. This, however, adds another layer of complexity in feedstock biodesign and engineering. Participants will explore new strategies that enable efficient design, modeling, and testing toward engineering multiple traits rapidly.

2:10 p.m. **Break**

2:30 p.m. **Breakout 7 Report-Out**

3:00 p.m. **Breakout 8: Emerging Synergies and Opportunities for Plant Biodesign Research**

Breakouts 1 through 7 discussed multiple aspects of the biodesign of feedstock crops. These sessions covered plant chemical factories, environmental resilience, plant transformation, plant genome engineering, plant synthetic biology, automation and AI, and cross-cutting multitrait engineering. In this final session, participants will reflect on the previous discussions and identify emerging opportunities, interactions, and inspirations that have not yet been addressed.

3:40 p.m.	**Breakout 8 Report-Out**
4:10 p.m.	**Conclusion and Next Steps**
4:30 p.m.	**Adjourn Workshop**

Appendix B

Workshop Participants

Co-Chairs

Gözde S. Demirer, *California Institute of Technology*

Yiping Qi, *University of Maryland*

Chung-Jui (CJ) Tsai, *University of Georgia*

Participants

Amir Ahkami, *Pacific Northwest National Laboratory*

Ana Alonso, *University of North Texas*

Rebecca Bart, *Donald Danforth Plant Science Center*

Jennifer Brophy, *Stanford University*

Robin Buell, *University of Georgia*

Tessa Burch-Smith, *Donald Danforth Plant Science Center*

Shao-shan Carol Huang, *New York University*

Heather Coleman, *Syracuse University*

Kelly Dawe, *University of Georgia*

José Dinneny, *Stanford University*

Arjun Khakhar, *Colorado State University*

Jung-Youn Lee, *University of Delaware*

Laurie Leonelli, *University of Illinois Urbana-Champaign*

Song Li, *Virginia Polytechnic Institute and State University*

Jianxin Ma, *Purdue University*

Amy Marshall-Colon, *University of Illinois Urbana-Champaign*

Dawn Nagel, *University of California–Riverside*

Mihai Pop, *University of Maryland*

Seung Yon (Sue) Rhee, *Michigan State University*

Julian Schroeder, *University of California–San Diego*

Heike Sederoff, *North Carolina State University*

John Shanklin, *Brookhaven National Laboratory*

Patrick Shih, *University of California–Berkeley*

Keith Slotkin, *Donald Danforth Plant Science Center*

Nathan Springer, *Bayer Crop Science*

Kankshita Swaminathan, *HudsonAlpha Institute for Biotechnology*

John Vogel, *Lawrence Berkeley National Laboratory*

Cătălin Voiniciuc, *University of Florida*

Daniel Voytas, *University of Minnesota–Twin Cities*

Jing-Ke Weng, *Northeastern University*

Xiaohan Yang, *Oak Ridge National Laboratory*

Philipp Zerbe, *University of California–Davis*

Appendix C

Workshop Vision and Goals

The vision of this workshop is to assess how, within the DOE and BER mission, to understand and manipulate potential bioenergy crops to efficiently generate renewable and sustainable biofuels, bioproducts, and biomaterials under variable abiotic stresses. High-level goals of the workshop include but are not limited to the below themes.

Genome Engineering Challenges Across Scales

- Discuss the prospects and bottlenecks for next-generation plant genome engineering technologies, such as large-scale CRISPR screens, synthetic chromosomes, synthetic genomes, and delivery and integration of large DNA molecules for plant biodesign at multiple scales, from molecular to pathway, to cellular, to organismal levels, and from genes to cis regulatory elements to repetitive regions.

Computational, Automation, Modeling

- Articulate specific needs in high-throughput and automation approaches for testing, phenotyping, genome editing, and other technologies.
- Delve into computational modeling and AI to accelerate the design of new functionality and to predict outcomes of engineering approaches.
- Develop standards for generating data that are compatible with machine learning and AI.

Plant Chemical Factories for Biofuels, Biomaterials, Bioproducts

- Explore opportunities for harnessing plants' genetic potential to synthesize advanced biofuel precursors, organic-inorganic composite materials, carbon-storage materials, biominerals, bioplastics, and living materials with new-to-nature properties and functionality. Consider environmental benefits that can be expected from such endeavors.
- Discuss resources to bridge the valley of death for technology translation.

Stress Resilience, Environmental Interactions

- Define the basic science in plant biology that holds promise to advance towards a circular bioeconomy, including but not limited to water and nutrient use/acquisition, efficient carbon harvesting and sequestration, abiotic stress tolerance, cell wall architecture, metabolism, and perenniality.
- Brainstorm how to better support field studies of engineered crops.

Crosscutting Topics

- Analyze the state of the art and innovations needed in omics/systems biology, analytic technologies, and plant genome and epigenome engineering.

- Evaluate the risks posed by growing engineered crops in open environments and what type of safeguards can be built in for biocontainment and prevention of gene flow to address ethical, legal, and societal implications.

In summary, the overarching objective of the workshop is to identify opportunities, roadblocks, and knowledge gaps that need to be addressed to reach the potential of plant design and engineering in developing a sustainable bioeconomy. Meeting the ambitious targets of a robust bioeconomy in the next 5 to 10 years requires innovative, multidisciplinary approaches encompassing the fields of biology, genetics, biochemistry, engineering, and physiology coupled with computational and synthetic biology approaches. The outcomes of the workshop will be captured in a report that is expected to provide a visionary outlook on the future of the plant synthetic biology and biosystems design field beyond the current state-of-the art.

Appendix D

Pre-Workshop Brainstorming Questions

Workshop organizers developed the following questionnaire, which was sent to experts in the field who were also invited to participate in the Plant Design for a Developing Bioeconomy virtual workshop.

Pre-Workshop Questionnaire

The purpose of this virtual brainstorming form is to invite members of the community to provide feedback on the potential of plant design research as it relates to the mission of the Genomic Science Program (GSP) supported by the Biological Systems Science Division (BSSD) of the Biological and Environmental Research (BER) program within the U.S. Department of Energy's (DOE) Office of Science. This brainstorming seeks to consider fundamental research that is needed to develop the next generation of genome engineering technologies that will enable effective and efficient manipulation of plant functions to produce fuels, valuable chemicals, and materials from renewable biomass to support the developing U.S. bioeconomy. The overarching objective of the workshop is to identify opportunities, roadblocks, and knowledge gaps that need to be addressed to reach the potential of plant design and engineering. Meeting the ambitious targets of a robust bioeconomy in the next 5 to 10 years requires innovative, multidisciplinary approaches encompassing the fields of biology, genetics, biochemistry, engineering, and physiology coupled with computational and synthetic biology approaches.

The response to these brainstorming questions should take less than 15 minutes to complete. Please feel free to answer only the subset of questions you are comfortable responding to, given your scholarly expertise. Your responses will be used in anonymized, aggregated form to help frame the discussions at the upcoming BSSD workshop on this topic.

1. As one considers future innovations for Plant Design research, what is the greatest potential of bioenergy crop engineering for addressing challenging scientific, nature resource, and environmental problems?
2. What is the current state-of-the-art in bioenergy crop engineering? Describe cutting-edge technologies that currently support the design-build-test-learn cycle for plant genetic engineering.
3. What are the fundamental barriers and technical challenges to realize the full potential of Plant Design research?
4. What is most needed to stimulate innovation in Plant Design research? What resources, technology, facilities, infrastructure, or capabilities might be needed or would be beneficial?
5. What are the potential risks or unintended consequences (ethical, legal, regulatory, social, and environmental) of expanding designer crop applications? How can we identify and predict them, and how can they be addressed?
6. What other factors not captured above should be considered to advance Plant Design and bioenergy crop engineering research?

Appendix E

Breakout Questions

Breakout 1

Plant Chemical Factories for Biofuels, Biomaterials, and Bioproducts

1. What are the critical knowledge gaps preventing us from effectively utilizing feedstock crops as chemical factories? What metabolic pathways are still difficult to engineer? What about new-to-nature pathways?
2. What are the technical barriers to achieving these goals, considering that some topics will have dedicated breakout sessions on Day 2?
3. Discuss the essential basic science R&D efforts and resources needed to bridge the gap between research and technology translation.

Breakout 2

Bioengineered Feedstock Design for Environmental Resilience

1. What additional traits should be prioritized to ensure performance and success of bioengineered feedstock under variable abiotic conditions? What are the knowledge gaps and opportunities in bioengineering for enhancing abiotic stress resilience?
2. What are the current limitations and design challenges in achieving this goal, considering that some topics will have dedicated breakout sessions on Day 2?
3. How can we better design and support field studies to advance basic scientific knowledge? How can biodesign science improve the success of translational research from model systems to feedstock species and from laboratory or greenhouse settings to field conditions?

Breakout 3

Research Opportunities to Fill Key Knowledge Gaps in Plant Genome Engineering

1. Target selection remains a significant bottleneck due to limited understanding of gene function, regulation, and redundancy. What tools have advanced target selection over the past decade, and how can these be utilized to develop the next generation of plant genome engineering technologies?
2. What are the rate-limiting steps in current plant genome engineering practices, and what critical capabilities are still missing from the toolkit?
3. Considering these existing gaps, which emerging wet and dry laboratory tools hold the greatest potential, and what new tool development should be prioritized to overcome these challenges?
4. How can we effectively integrate molecular engineering tools with computational advances, such as AI and omics data, to streamline and improve plant biodesign workflows?

Breakout 4

Advancing Plant Synthetic Biology for Transformative Feedstock Engineering

1. How can we standardize and expand the library of synthetic biology tools and parts, including tissue- or cell-specific promoters, regulatory elements, synthetic gene circuits, synthetic chromosomes, and minimal genomes, to accelerate plant engineering? What are the major bottlenecks, and what innovative solutions could address them?
2. How can synthetic biology approaches enable the design of dynamic, tunable, and programmable regulatory circuits for precise control of (multi)gene expression in plants? What tools or

systems show the most promise for achieving this level of control?

3. How can synthetic biology tools, such as cell-free systems and computational innovations, be leveraged to prototype and validate new metabolic pathways rapidly before *in planta* implementation?
4. What advancements are needed to streamline and accelerate the design-build-test-learn (DBTL) cycle in plant synthetic biology?

Breakout 5

Plant Transformation Innovation, Characterization, and Phenotyping

1. What are the most promising innovations in plant transformation methods from the past two years, and how can we ensure their broad adoption and routine application, particularly for feedstock species?
2. What are the current challenges in achieving high-throughput phenotyping for hundreds of transformed plants at the whole-plant stage, and what technological or methodological (wet or dry laboratory) innovations are required to address them?
3. What advancements are needed in transformation workflows to support high-throughput CRISPR screens or other large-scale mutagenesis studies, and what steps can be taken to achieve these improvements?
4. How can transformation tools be made more ubiquitous and applicable to a diverse range of cells, tissues, and species? Or is the idea of universal plant transformation unrealistic and efforts should focus on finding rapid ways of transformation pipeline development in each target species?

Breakout 6

Harnessing the Power of Automation, Modeling, and AI to Aid Plant Design and Engineering

1. In what areas have we seen automation and modeling play a role in biodesign research? In what areas could we expand the reach of automation and modeling to aid biodesign, and how?
2. AI-based design tools are being rapidly developed, covering, for example, protein structure prediction and design, protein–protein interactions, and protein–nucleic acid interactions. How can we use these tools efficiently in biodesign research? What are the relevant tools that we need urgently but have not been developed yet?
3. The development of AI tools demands large and high-quality datasets for machine learning. Beyond current omics data, what datasets need to be generated to help the development of next-generation AI tools for biological and trait engineering?
4. How can a large language model (LLM) be used in biodesign, prediction, and trait engineering? What are the opportunities in this area, and how can we materialize these opportunities?

Breakout 7

Enabling Efficient, Integrative Multitrait Engineering

1. What are the knowledge gaps in multitrait engineering? How can these gaps be effectively addressed?
2. What are the major technical limitations to efficiently design multiple traits in a feedstock crop? How can such limitations be overcome?
3. Engineering stress tolerance can often have trade-offs in plant growth, development, or yield. Should there be a general guideline for benchmarking successful multitrait engineering? If so, what would that include?

Breakout 8

Emerging Synergies and Opportunities for Plant Biodesign Research

Breakouts 1 through 7 discussed multiple aspects of the biodesign of feedstock crops. These sessions covered plant chemical factories, environmental resilience, plant transformation, plant genome engineering, plant synthetic biology, automation and AI, and cross-cutting multitrait engineering. In this final session, participants will reflect on the previous discussions and identify emerging opportunities, interactions, and inspirations that have not yet been addressed.

Appendix F

References

Ariyaratne, M., et al. 2023. "CRISPR/Cas9-Mediated Targeted Mutagenesis of Inulin Biosynthesis in Rubber Dandelion," *Journal of the American Society for Horticultural Science* **148**(6), 266–75. DOI:10.21273/JASHS05311-23.

ARPA-E. 2024. "How Could Phytomining Bolster U.S. Critical Mineral Supply Chains?" Department of Energy Advanced Research Projects Agency-Energy. https://arpa-e.energy.gov/news-and-events/news-and-insights/how-could-phytomining-bolster-us-critical-mineral-supply-chains

Cao, X., et al. 2023. "Cut–Dip–Budding Delivery System Enables Genetic Modifications in Plants Without Tissue Culture," *The Innovation* **4**(1), 100345. DOI:10.1016/j.xinn.2022.100345.

Dong, J., et al. 2025. "Enhancing Lipid Production in Plant Cells Through Automated High-Throughput Genome Engineering and Phenotyping," *The Plant Cell* **37**(2), koaf026. DOI:10.1093/plcell/koaf026.

GLBRC. 2018. "Zip-Lignin™," Great Lakes Bioenergy Research Center. https://www.glbrc.org/industry/technologies/zip-lignintm

Huang, J., et al. 2024. "Discovery of Deaminase Functions by Structure-Based Protein Clustering," *Cell* **187**(16), 4426–8. DOI:10.1016/j.cell.2024.07.003.

Jiang, K., et al. 2024. "Rapid *In Silico* Directed Evolution by a Protein Language Model with EVOLVEpro," *Science* **387**(6732), eadr6006. DOI:10.1126/science.adk8946.

Kuperman, O. A., et al. 2024. "Harnessing Precursor-Directed Biosynthesis with Glucose Derivatives to Access Cotton Fibers with Enhanced Physical Properties," *Cell Reports Physical Science* **5**(5), 101963. DOI:10.1016/j.xcrp.2024.101963.

LBNL. 2023. "Synthetic Transcription Factor Systems for Enhanced Control of Gene Expression," Lawrence Berkeley National Laboratory, Intellectual Property Office. https://ipo.lbl.gov/2023/01/18/synthetic-transcription-factor-systems-for-enhanced-control-of-gene-expression/

Liu, G., et al. 2025. "PDLLMs: A Group of Tailored DNA Large Language Models for Analyzing Plant Genomes," *Molecular Plant* **18**(2), 175–8. DOI:10.1016/j.molp.2024.12.006.

Liu, P., et al. 2024. "Transposase-Assisted Target-Site Integration for Efficient Plant Genome Engineering," *Nature* **631**, 593–600. DOI:10.1038/s41586-024-07613-8.

Ortega, M. A., et al. 2022. "*In vitro* Floral Development in Poplar: Insights into Seed Trichome Regulation and Trimonoecy," *New Phytologist* **237**(4), 1078–81. DOI:10.1111/nph.18624.

Qiao, J.-H., et al. 2025. "Transgene- and Tissue Culture-Free Heritable Genome Editing Using RNA Virus-Based Delivery in Wheat," *Nature Plants* **11**, 1252–9. DOI:10.1038/s41477-025-02023-8.

Rylott, E. L., and A. van der Ent. 2025. "Harnessing Hyperaccumulator Plants to Recover Technology-Critical Metals: Where Are We At?" *New Phytologist* **246**(3), 859–66. DOI:10.1111/nph.20449.

Shanker, V. R., et al. 2024. "Unsupervised Evolution of Protein and Antibody Complexes with a Structure-Informed Language Model," *Science* **385**(6704), 46–53. DOI:10.1126/science.adk8946.

Sun, C., et al. 2024. "Precise Integration of Large DNA Sequences in Plant Genomes Using PrimeRoot Editors," *Nature Biotechnology* **42**, 316–27. DOI:10.1038/s41587-023-01769-w.

Unda, F., et al. 2022. "A New Approach to Zip-Lignin: 3,4-Dihydroxybenzoate Is Compatible with Lignification," *New Phytologist* **235**(1), 234–46. DOI:10.1111/nph.18136.

U.S. DOE. 2024. *Overcoming Barriers in Plant Transformation: A Focus on Bioenergy Crops*, DOE/SC-0215. U.S. Department of Energy Office of Science. DOI:10.2172/2335710.

Voiniciuc, C., et al. 2018. "Monitoring Polysaccharide Dynamics in the Plant Cell Wall," *Plant Physiology* **176**(4), 2590–600. DOI:10.1104/pp.17.01776.

Weiss, T., et al. 2025. "Viral Delivery of an RNA-Guided Genome Editor for Transgene-Free Germline Editing in *Arabidopsis*," *Nature Plants* **11**, 967–76. DOI:10.1038/s41477-025-01989-9.

Zale, J., et al. 2015. "Metabolic Engineering of Sugarcane to Accumulate Energy-Dense Triacylglycerols in Vegetative Biomass," *Plant Technology Journal* **14**(2), 661–9. DOI:10.1111/pbi.12411.

Zhang, R., et al. 2025. "PlantGPT: An Arabidopsis-Based Intelligent Agent that Answers Questions about Plant Functional Genomics," *Advanced Science* **12**(30), e03926. DOI:10.1002/advs.202503926.

Appendix G

Acronyms and Abbreviations

3D	three-dimensional
AI	artificial intelligence
ARPA-E	DOE Advanced Research Projects Agency-Energy
BER	DOE Biological and Environmental Research program
BSSD	BER's Biological Systems Science Division
CRISPR	clustered regularly interspaced short palindromic repeats
DBTL	design–build–test–learn
DOE	U.S. Department of Energy
DSBs	double-strand breaks
dsRNA	double-stranded ribonucleic acid
ELMs	engineered living materials
FAST-PB	fast, automated, scalable, high-throughput pipeline for plant bioengineering
GRIN	Germplasm Resources Information Network
gRNA	guide ribonucleic acid
GSP	BER's Genomic Science Program
HDR	homology-directed repair
HTP	high-throughput
LLM	large language model
ML	machine learning
mRNA	messenger ribonucleic acid
NHEJ	non-homologous end-joining
PHYTOMINES	Plant HYperaccumulators TO MIne Nickel-Enriched Soils
PLM	protein language model
RNAi	ribonucleic acid interference
sgRNA	single guide ribonucleic acid
TAIR	The Arabidopsis Information Resource
T-DNA	transfer DNA
VIGE	virus-induced gene editing
VIGS	virus-induced gene silencing