Pacific Northwest
NATIONAL LABORATORY

Persistence Control
@PNNL

DOE BSSD Performance Management Metrics Summary Report

SFA Laboratory Research Manager: Jon Magnuson[1], Jon.Magnuson@pnnl.gov
SFA Principal Investigator: Robert G. Egbert[1], Robert.Egbert@pnnl.gov
[1]**Pacific Northwest National Laboratory**, Richland, WA 99354

Investigators: Robert G. Egbert[1], Jayde A. Aufrecht[1], Adam M. Deutschbauer[2], Joshua R. Elmore[1], Adam M. Guss[3], Pubudu P. Handakumbura[1], Caroline S. Harwood[6], Vivian S. Lin[1], Ryan S. McClure[1], Jason E. McDermott[1], Ernesto S. Nakayasu[1], William C. Nelson[1], Devin Coleman-Derr[4], Enoch Yeung[5]

Participating Institutions: [1]**Pacific Northwest National Laboratory**, Richland, WA 99354; [2]Lawrence Berkeley National Laboratory, 94720; [3]Oak Ridge National Laboratory, Oak Ridge, TN 37831; [4]University of California, Berkeley, CA 94704; [5]University of California, Santa Barbara, CA 93106; [6]University of Washington, Seattle, WA 98195.

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <https://www.ntis.gov/about>
Online ordering: http://www.ntis.gov

FY23 Metric: Report on the development of genomic approaches to securely design organisms with targeted beneficial functions for bioenergy and bioeconomy applications.

Introduction

The potential to employ emerging genetic and computational tools to design and securely deploy synthetic rhizosphere microbiomes, defined communities of microbes that colonize plant roots and benefit plant growth, offers tremendous opportunity to realize highly productive and stress-tolerant biomass cropping systems. A critical obstacle to realizing this vision is understanding fundamental principles of microbiome persistence in complex environments, notably amid the dynamics of root exudation and microbial colonization. Further, we need predictive tools to assess the risks associated with the deployment or unintended release of engineered microbes in plant and soil ecosystems and to mitigate those risks through effective containment strategies. To meet these challenges, high-throughput genetic manipulation and bioinformatic tools provide a platform to elucidate the genetic elements underpinning rhizosphere community function, to create beneficial communities, and to control their environmental persistence.

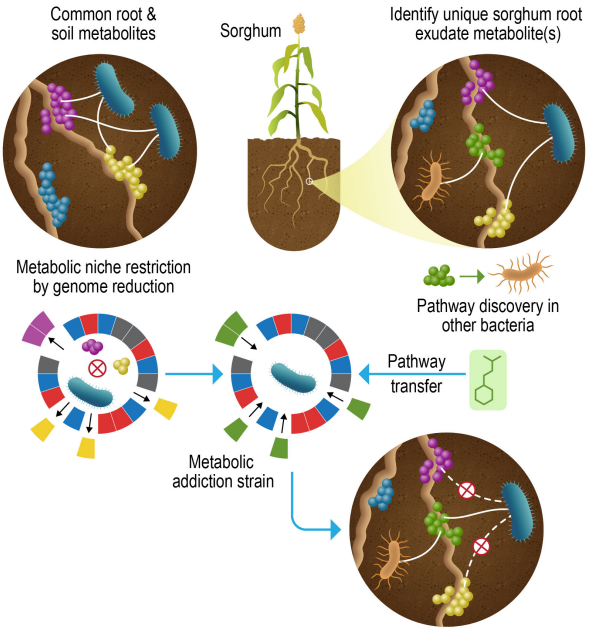

Figure 1. The Persistence Control SFA aims to control the proliferation of engineered microbes by establishing metabolic addiction of microbe to a bioenergy crop supported by the engineered microbiome.

The Persistence Control SFA investigates approaches to reshape the environmental niche of native sorghum rhizosphere bacterial isolates by creating an engineered metabolic addiction to root exudate compounds. We aim to understand the genes and networks that underpin environmental persistence of microbiomes to create dependencies of engineered microbes on plants such that the microbes can no longer persist in the environment in the absence of the target plant (**Figure 1**). Our strategy to control the environmental persistence of engineered microbes is to establish metabolic addiction to plant root exudate compounds by installing compound-specific catabolic pathways while removing pathways used for scavenging nutrients in soil microbiome environments. This strategy requires exquisite control of genome content to, for instance, eliminate metabolic niches that allow survival cross-feeding of nutrients by other microbiome members as well as to introduce complete pathways to thrive in the plant rhizosphere, but not in the rhizosphere of other plants. It also demands a deep understanding of plant or microbe interactions that could overcome control through complementation.

To help accomplish the goals of the Secure Biosystems Design program, the PerCon SFA team draws expertise from across three DOE Office of Science National Laboratories – Pacific Northwest National Laboratory, Lawrence Berkeley National Laboratory, and Oak Ridge National Laboratory – and three Universities: the University of California Berkely, the University of California Santa Barbara, and the University of Washington. To understand the genes and networks that control environmental fitness we have developed and integrated research strategies that draw from our team's expertise in functional genomics, synthetic biology, microbial ecology, chemical biology, bioinformatics, machine learning, and plant-microbe interactions. In this report, we share a selection of capabilities developed for persistence control that accelerate the design and testing of microbial isolates with new beneficial functions.

This report details advances made by the Persistence Control SFA during its inaugural funding cycle from 2020-2023. It highlights research outcomes that are available from the quarterly reports focusing on the following performance metrics:

1) **Describe overall approaches to securely designing new functions into organisms for bioenergy and bioeconomy applications.** (Q1 Performance Management Metrics Report) In this report, we provided an overview of seven experimental and computational approaches the Persistence Control SFA uses to interrogate microbial function and plant-microbe interactions: a strain-agnostic method for high-throughput genetic manipulation of bacteria, including rhizosphere isolates [1]; a methodology to perform genome-wide mutant fitness assays in soil environments; two types of synthetic microbial communities derived from competitive co-enrichment and from defined isolates [2] cultivated on a data-driven synthetic sorghum growth medium; a scalable software pipeline to cluster and search protein libraries for shared functions using amino acid recoding schemes [3]; a laboratory plant cultivation platform to interrogate spatially-resolved plant-microbe interactions in soil [4]; an activity-based chemical probe that mimics the structure of sorgoleone [5], an important biological nitrification inhibitor compound produced in sorghum roots that is the focus of the SFA's efforts to enhance the rhizosphere persistence niche of engineered bacteria; and a multi-omics platform to discover genes responsible for bacterial metabolism of sorghum root exudate compounds[6, 7].

2) **Report on progress towards designing new secure functions into specific microbial isolates in the laboratory.** (Q2 Performance Management Metric Report) This report focused on the SFA team's development and application of Serine recombinase-assisted genome engineering (SAGE) [1]. Despite recent development of tools such as CRISPR-Cas gene editing, efficient integration of exogenous DNA with well-characterized functions remains limited to model bacteria. SAGE is an easy-to-use, highly efficient, and genome integration technology that enables selection marker-free, site-specific genome integration of up to 10 DNA constructs, often with efficiency on par with or superior to replicating plasmids, the molecular biology standard for heterologous DNA transfer. The advantage of SAGE is the stability and use cases it provides that are not possible with plasmid-based systems. These applications include stable propagation of microbial functions in complex conditions without antibiotics, which is particularly useful when culturing engineered hosts with laboratory microbiome samples or in soil or rhizosphere environments. We demonstrated SAGE by characterizing genome integration efficiency in five bacteria that span multiple taxonomy groups and biotechnology applications and by identifying more than 95 heterologous promoters in each host with consistent transcription across environmental and genetic contexts. We anticipate that SAGE will rapidly expand the number of industrial and environmental bacteria compatible with high-throughput genetics and synthetic biology.

3) **Provide a summary of methods used to characterize and test modified microorganisms for secure use.** (Q3 Performance Management Metric Report) Our third quarter report detailed synthetic microbial community testbeds we developed from sorghum-associated soil as platforms to

characterize and test modified microorganisms for persistence control phenotypes. First, we summarized our efforts to naturally evolve a synthetic microbial community that grows on a synthetic sorghum exudate medium, finding that the diversity in community composition after enrichment experiments determines the colonization potential of an engineered plant-growth promoting bacterium that is not native to sorghum soil. Second, we described our efforts to create a synthetic microbial community from defined sorghum-associated bacterial isolates [2]. We found that the defined community promotes the growth of sorghum plants in laboratory and field contexts, though the community does not appear to be enriched in the field rhizosphere after multiple weeks. We are following up on these studies to better understand community and plant-microbe interactions that drive environmental persistence dynamics.

4) **Report on capabilities that could accelerate the design and testing of microbial isolates with new beneficial functions.** (Q4 Performance Management Metrics Report) In this report, we documented our advances in discovering genes that are essential to utilize root exudate compounds in the rhizosphere that contribute to biological nitrification inhibition. The ability to control the relative abundance and environmental persistence of plant growth promoting microbes in the rhizosphere could drastically increase the resilience of plants grown in challenging environments while reducing requirements for nutrients and water. The SFA team has discovered genes responsible for bacterial catabolism of sorgoleone and methyl 3-(4-hydroxyphenyl) propionate (MHPP), two root exudate compounds that inhibit biological nitrification for the bioenergy crop sorghum. Better understanding of these and related catabolic pathways will provide better understanding of mechanisms and dynamics of biological nitrification inhibition (BNI) in soil and rhizosphere environments and establish means to better control the persistence of engineered microbes in the rhizosphere by creating tunable metabolic niches for plant-specific metabolites.

Below we summarize highlights from the four subject reports. More details can be found by referencing these and prior Performance Management Reports on the BER Genomic Science Program website: https://genomicscience.energy.gov/doe-performance-metrics/.

Overall approaches to securely designing new functions into organisms for bioenergy and bioeconomy applications

Understanding the genetic factors that control microbial persistence in soil and rhizosphere environments and applying these scientific discoveries to control the persistence of engineered functions requires new experimental and computational methods. To demonstrate persistence control principles in the sorghum rhizosphere, our SFA must develop high-throughput genetic tools for non-model bacterial hosts, synthetic microbial consortia that represent the phylogenetic and functional diversity of native communities, computational methods to estimate the potential of a microbial community to complement engineered niche restrictions, high-throughput methods to discover and characterize catabolic pathways for root exudate compounds, and model fabricated plant-microbe ecosystems that are predictive of field dynamics. The SFA's primary approaches are summarized below:

1) We developed a bacterial genome engineering tool called Serine recombinase-Assisted Genome Engineering, or SAGE, with which we sought to achieve parity for engineering model, non-model, and undomesticated bacterial strains [8]. SAGE utilizes up to 10 site-specific serine recombinases that each catalyze DNA recombination between two distinct attachment sequences (*attP* and *attB* sites). This process is unidirectional, which unlike many similar technologies generates a stably modified chromosome with no further steps. For further reference, see BER Research Highlight on the published research (https://science.osti.gov/ber/Highlights/2023/BER-2023-06-a).

2) Understanding the genes and associated functions that contribute to environmental fitness phenotypes in microbes is critical to developing persistence control strategies for engineered microbial functions in rhizosphere environments. We have developed an approach to apply high-throughput functional genomics methods [9] to understand which genes contribute to fitness in sterile and live soil environments. This understanding will drive genome reduction efforts to establish metabolic addiction phenotypes for our engineered hosts wherein they colonize the sorghum rhizosphere but are unfit in soil without sorghum.

3) The rhizosphere represents a dynamic and complex interface between plant hosts and the microbial community found in the surrounding soil. To understand microbial rhizosphere persistence at the community level and to provide a platform for genetic and metabolic complementation studies for engineered persistence control strains, we constructed an inoculation formula using a synthetic community of individual isolates that we hypothesized would improve colonization efficiency and the relationship with the host plant, *Sorghum bicolor*. The community was assembled from bacterial isolates that were identified to potentially play a role in community cohesion through network analysis or identified to benefit from host-specific exudate compounds. We observed that the synthetic community promotes growth of above- and below-ground plant phenotypes compared to uninoculated controls, both in the lab and when applied to plants grown in the field [2]. These results demonstrate the potential utility of synthetic communities to support crop performance even in the absence of persistence and the need for a deeper mechanistic understanding of community control of host fitness in agricultural contexts.

4) Analysis of protein functional annotation assignment has revealed that, on average, only ~50-60% of genes are assigned any functional annotation using current standard annotation techniques, and only about half of those receive specific functional assignment [10, 11]. To address this knowledge gap, we have developed a scalable pipeline for building models of protein functional families, searching those families against novel sets of protein sequences (e.g., from metagenomes), and clustering novel sets of protein sequences for the purpose of classification into functional families. The framework, called Snekmer, is based on protein fingerprinting, whereby protein sequences transformed using one of several amino acid reduction approaches to make functional patterns more apparent, and are treated as a profile comprising overlapping subsequences of a fixed length, or kmer [3]. Snekmer models are constructed for each protein family using machine learning to identify patterns that uniquely and specifically identify the family of interest. We have applied this approach to develop models for each of the sorghum utilization genes identified in the project, and applied those models to search for other examples in large-scale genome databases. For further reference, see BER Research Highlight on the published research (https://science.osti.gov/ber/Highlights/2023/BER-2023-04-b).

5) To understand the colonization dynamics of our engineered hosts and synthetic communities, we developed a platform to investigate spatially resolved root-microbiome interactions using Rhizogrids. A Rhizogrid is an innovative plant cultivation platform to integrate taxonomic and metabolomic data with a three-dimensional root cartography workflow to enable unprecedented spatial mapping of root exudates and microbes using plants grown in soil [4]. We designed Rhizogrids to anchor a pot-grown root system and to provide 3D coordinates for root and rhizosphere samples harvested from the pot. With our approach, we can link microbial taxa and metabolites corresponding to specific locations in the root system. By linking molecular-level measurements to spatial information for the root system we can begin to understand microenvironment effects on rhizosphere colonization and interrogate colonization dynamics using engineered rhizosphere isolates and synthetic microbial communities. The resulting information is critical to shed light into persistence control principles in field-relevant laboratory environments and, more broadly, interactions between plants and microbes in the rhizosphere which can play significant roles in the resilience of plants to climate change.

6) To complement our ongoing work involving isolation and characterization of microbes capable of using sorgoleone as a substrate for metabolic addiction, we designed and synthesized a novel chemical probe based on the structure of sorgoleone to enable ABPP of proteins that interact with sorgoleone. This probe features the benzoquinone head group of sorgoleone and a lipid-like tail which incorporates a diazirine photocrosslinking group and alkyne "click" tag for further attachment of reporter groups such as a fluorophore or biotin to allow for fluorescence detection or streptavidin enrichment, respectively, of probe-labeled proteins. We have applied the sorgoleone diazirine alkyne (SoDA) probe to *Acinetobacter pittii* SO1, a novel sorgoleone catabolizing isolate we isolated from sorghum field soil, to profile sorgoleone-interacting proteins. We identified proteins that are known to bind benzoquinones, as well as lipid metabolizing proteins, transporters, and transcriptional regulators, which are now being explored further using structural prediction and biochemical approaches [5].

7) Despite substantial research on the production and characteristics of many sorghum-specific root exudate compounds, the metabolic pathways through which microbes consume these compounds are unknown and must be discovered to create metabolic dependence pathways to transfer into persistence control hosts. Bacteria capable of using sorghum-specific exudates are largely unknown. We isolated organisms that consume sorghum-specific exudates by enriching for growth in media containing the exudate as the sole carbon source with sorghum field soil as the inoculant. We have found that combining transcriptomic data with genome-wide mutant fitness data for these hosts with unique metabolic profiles is a powerful method to filter out uninvolved genes and identify critical genes with no differential gene expression [6, 7].

Designing new secure functions into specific microbial isolates in the laboratory

The lack of high-throughput genetic systems for non-model hosts remains a substantial barrier to understanding what drives environmental persistence in rhizosphere microbiomes. Here, we summarize efforts from the Persistence Control SFA team to break down these barriers to introduce engineered functions in rhizosphere hosts, a critical step towards establishing persistence control principles.

1) We established serine recombinase-assisted genome engineering (SAGE) [1], utilizing phage recombinase enzymes and attachment site 'landing pads' to efficiently insert up to 10 recombinant DNA molecules into a single bacterial host. The attachment sites for the phage recombinases is introduced into the host using traditional insertion methods (e.g., homologous recombination, transposase insertion) and the SAGE recombinase & cargo DNA molecules are introduced by electroporation.

2) SAGE DNA integration was demonstrated for five bacteria representing broad functional and taxonomic diversity (**Figure 2**). Three hosts were Pseudomonads from soil environments, which included two novel isolates from sorghum fields. One host is the model purple photosynthetic bacterium *Rhodopseudomonas palustris* used to study nitrogen fixation and hydrogen production. One host is the model actinomycete *Rhodococcus jostii* used for bioremediation. The ready application of SAGE to these and additional hosts by collaborating labs suggest that efficient genome integration can be accomplished for virtually any bacteria.

3) Residual DNA (e.g. antibiotic resistance markers or replication origins) from SAGE-integrated plasmids can be removed to enable multiple SAGE integration cycles. SAGE-integration DNA is designed with attachment sites for a phage recombinase that when expressed in a SAGE-compatible host is excised from the chromosome. The excision recombinase can be transiently expressed or selectively propagated and then removed using temperature shifts or counterselection methods.

4) SAGE-integrated DNA is genetically stable and requires no antibiotics for maintenance. Fluorescence gene expression from a SAGE-integrated gene cassette remained consistent over twenty generations of growth compared to similar genetic constructs on plasmid DNA which were lost after only ten generations. The stable expression of introduced functions is extremely important to evaluating persistence control in soil and rhizosphere contexts where antibiotic use is deleterious to the microbial community or impractical at field scales.

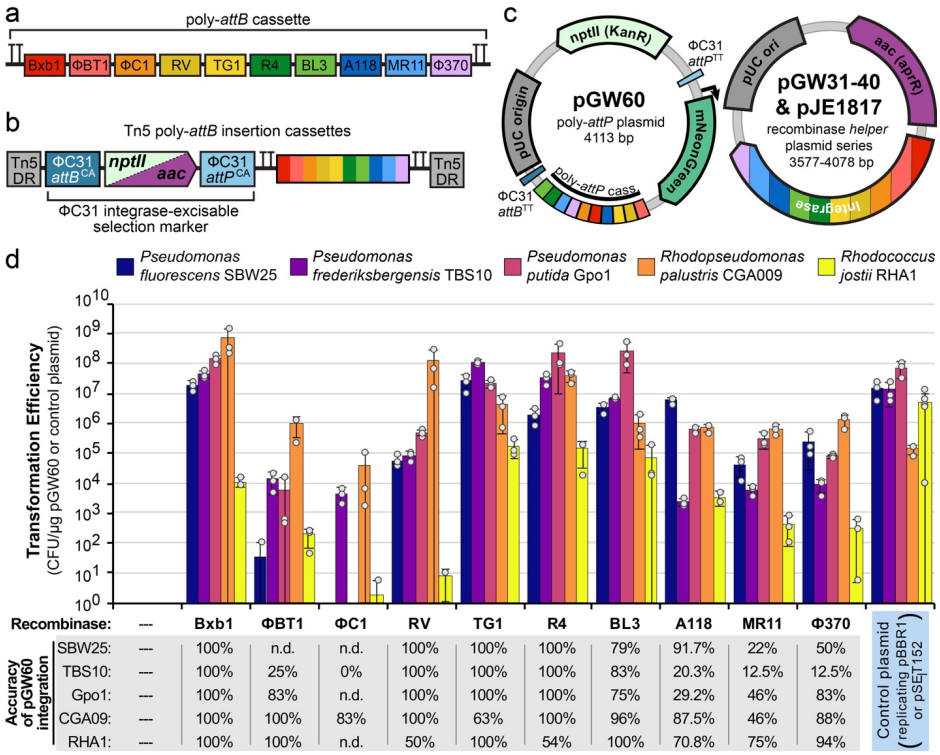

Figure 2. SAGE enables stable, highly efficient integration of plasmid DNA into the genomes of engineered bacteria. (a) Diagram of genome-integrated 10x poly-attachment site cassette. (b) Diagram of transposase-mediated insertion cassette. (c) Plasmid maps of SAGE plasmids used for efficiency experiments in panel d. (d) Transformation efficiency when the poly-attP target plasmid pGW60 is transformed with or without an integrase-expressing helper plasmid, or when the positive control plasmid is transformed.

5) SAGE enables multiplexed evaluation of genetic variants. We adapted a method of combined nucleic acid barcode sequencing of DNA and RNA [12] to assess the activity of hundreds of transcriptional promoters in parallel. Our analysis resulted in a toolkit of promoters for each of our five hosts, which will improve genetic circuit and pathway design from reliable genetic parts. This approach could be extended to evaluate complex engineered functions, including combinatorial multi-stage integration libraries, from DNA cassettes expressed stably from the genome.

Methods used to characterize and test modified microorganisms for secure use

The Persistence Control SFA has advanced methods to characterize and test modified microorganisms for secure use by establishing two types of synthetic microbial communities derived from a sorghum field site as testbeds to assess the performance of strains engineered for persistence control. The scientific community lacks a clear understanding of what drives the colonization dynamics of microbial communities by individual species. Our progress in developing and testing the community dynamics of these synthetic communities will directly support SFA science goals and be a resource for other scientists working to understand plant-microbe interactions and microbial plant growth promotion traits.

1) One effective approach to investigate microbial interactions is to reduce community complexity by developing simple, experimentally tractable microbial communities [13, 14]. We developed five distinct synthetic sorghum communities (SSC) by passaging a sorghum field soil sample on nutrient agar media supplemented with 20 readily procured carbon and nitrogen compounds detected in

sorghum metabolomics data. The SSCs stabilized after multiple passages to be represented by dozens of distinct strains that span five to seven bacterial phyla.

2) The SSCs can be reproducibly generated through cold storage and revival. Replicate community samples stored in glycerol or dimethyl sulfoxide could be reproducibly reconstituted. Reproducible production of synthetic communities is a critical requirement for using these communities to test interactions with engineered persistence control hosts.

3) The colonization efficiency of an introduced bacterium depends on SSC composition. We introduced *Pseudomonas fluorescens* SBW25, a plant growth promoting bacterium, to each SSC culture and assessed the population fraction of SBW25 over multiple passages on agar plates. SBW25 reliably colonized only one of the five SSC communities (**Figure 3**).

4) Network analysis of community composition for SSC colonization by SBW25 suggest positive and negative co-abundance interactions. We observed positive interactions with *Pseudomonas* and *Klebsiella* species and negative interactions with *Pseudaminobacter*, *Pseudorhodoferax*, and *Shinella* species. We are investigating mechanisms that drive this differential colonization phenotype as well as how to use genetic engineering to conditionally enhance or suppress the colonization of engineered microbes in rhizosphere environments.

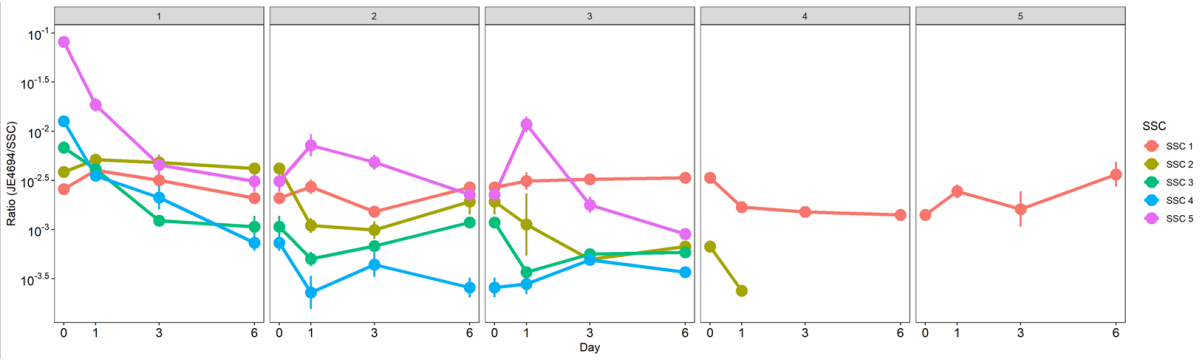

Figure 3. Differential colonization of a synthetic community by an engineered bacterium. Population fraction of introduced *Pseudomonas fluorescens* SBW25 (JE4694) in each synthetic sorghum community (SSC). Plots represent ratios of JE4694 which expresses a fluorescent protein to the non-fluorescent SSC community. JE4694 was cultured independently with each SSC over five culture passages (grey header). Ratios were measured using flow cytometry.

5) We formulated a defined synthetic community (dSynCom) from individual sorghum field bacterial isolates. We constructed this community from evidence of co-abundance in field conditions and the ability of these microbes to utilize specific root exudate compounds known to rhizosphere microbiome composition.

6) The community dynamics of the dSynCom is strongly impacted by cultivation conditions. When grown on nutrient agar plates supplemented with known sorghum exudates, many species were not detected. By contrast, when grown in the presence of sorghum plants in lab, the community has much more even representation (**Figure 4**).

7) The dSynCom promotes the growth of sorghum in the laboratory and the field. The dSynCom treatment improves sorghum growth by measures of shoot fresh weight and dry weight compared to mock treatments. Despite similar growth promotion outcomes observed in a field trial, the dSynCom members notably do not appear enriched in the field environment. We are actively investigating

multiple hypotheses related to the field outcome, including a role the dSynCom may play in priming the rhizosphere microbiome or in eliciting an immune response by the plants.

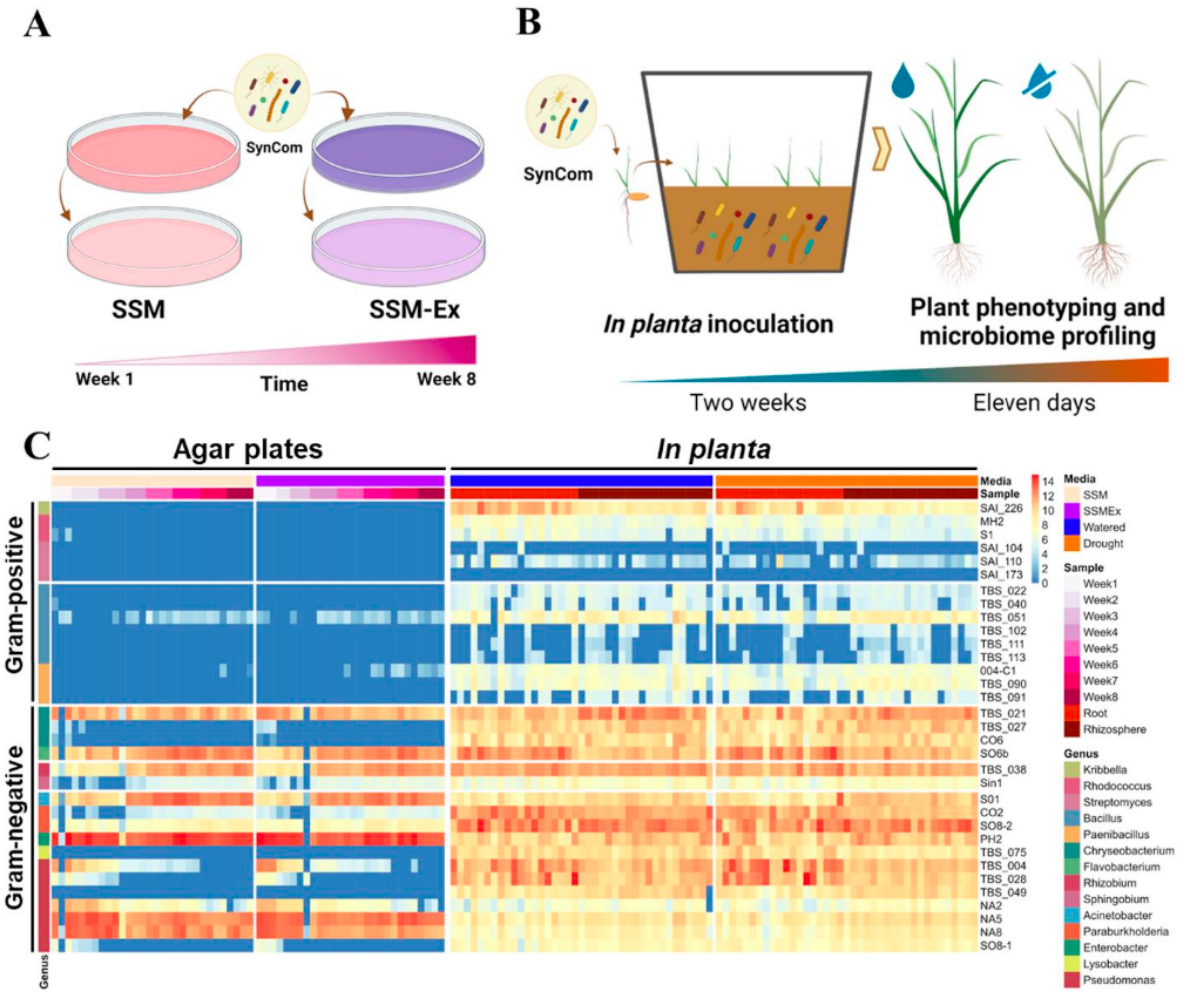

Figure 4. dSynCom colonizes the rhizosphere and roots of sorghum. (A) Schematic diagram of dSynCom community analysis on agar plates. Dilutions of the dSynCom were plated on agar plates with Synthetic Soil Media (SSM) or SSM with sorghum exudates (SSM-Ex) and passaged weekly. (B) Schematic representation of the laboratory dSynCom colonization of sorghum. Surface disinfected sorghum seeds were germinated and treated with mock or dSynCom in sterile conditions. Plants were watered or drought stressed. C) Heat map of the abundance of the dSynCom members identified across treatments *in vitro* and *in planta* experiments. The abundance is presented in log2 scale for n=174 samples for *in vitro* experiments and n=154 samples for the *in planta* experiment.

Accelerating the design and testing of microbial isolates with new beneficial functions

Agriculture is a major source of greenhouse gases (GHG) on the planet [15]. Although application of N-fertilizers increases crop yields [16], a large proportion of applied N-fertilizer is ultimately lost to the atmosphere as N_2O, a potent GHG (**Figure 5**) [17, 18]. Nitrifying microbes catalyze the oxidation of reduced forms of nitrogen (NH_3, NH_4^+) into nitrate (NO_3^-) and NO_3^- is converted to the potent greenhouse gas N_2O by denitrifying bacteria that are widely prevalent in soil [19]. One approach to reduce nitrification in agriculture is the co-application of chemical nitrification inhibitors with fertilizers [20, 21]. These chemicals generally inhibit activity of the highly conserved ammonium oxidase (AMO) and

hydroxylamine oxidoreductase (HAO) enzymes that perform biological nitrification. These enzymatic functions are limited to a small collection of archaeal (AOA) and bacterial (AOB) lineages, and blocking these activities inhibits the growth of these microbes. Another approach is to limit the application of chemical fertilizers. For this, researchers and several companies are developing engineered microbial and plant [22] systems that can be deployed in agricultural soils to enhance fixation of elemental nitrogen (N_2) into biologically available nitrogen (NH_3) [22-25]. Supplementing soils with these engineered N_2-fixing bacteria can substantially reduce the amounts of chemical N-fertilizer required for robust crop yields. While each approach is effective, a deeper understanding of the mechanisms driving biological nitrogen cycles is necessary to address the complex challenge of reducing nitrification.

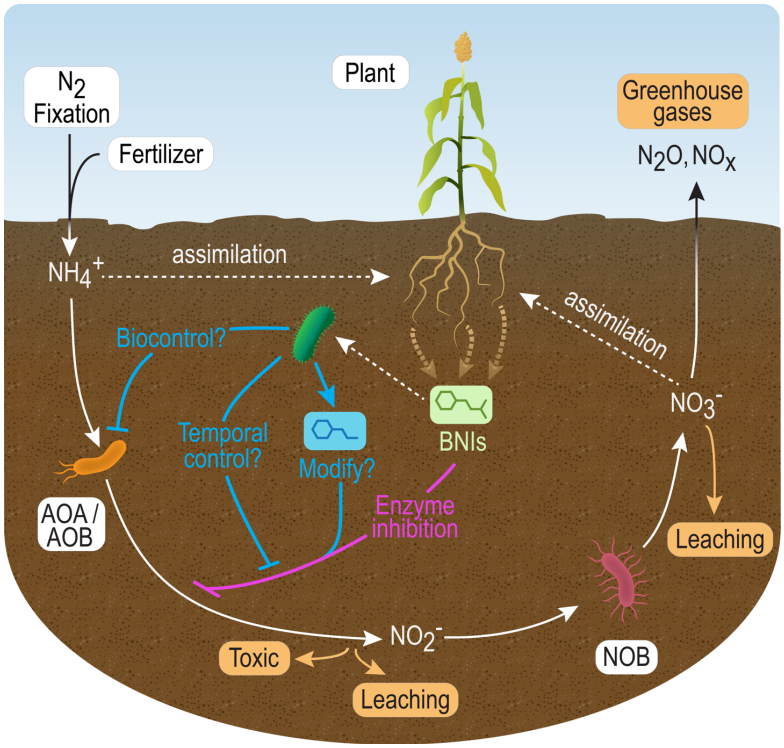

Fig. 5. Biological inhibition of the nitrification cycle. Ammonium oxidizing archaea (AOA) and bacteria (AOB) convert excess N in the soil into forms that leach into soils and are released as nitrogenous greenhouse gases. Plant roots exude biological nitrification inhibitor chemicals (BNIs) to prevent this loss of soil nitrogen. To date the main known BNI mode-of-action is the inhibition of ammonia oxidizing enzymes. A more complete understanding of the nitrogen cycle will require investigation of microbes capable of consuming BNIs. For example, enriched microbes may enhance inhibition through biocontrol of AOA/AOB growth or exert temporal control over enzyme inhibition through mineralization or modification of BNI compounds.

The Persistence Control SFA has advanced the field of bacterial metabolism of plant root exudates through investigating catabolism of two sorghum exudate compounds that act as biological nitrification inhibitors (BNI):

1) We have identified microbes that can use phenylpropanoid methyl esters (PPMEs), a class of BNIs, as growth substrates. This class of phenolic BNI is produced by a number of plants [26-28], including sorghum which produces MHPP (see Q4 summary above). To identify genes involved in PPME utilization in the host *Pseudomonas fluorescens* SBW25, we co-evaluated genome-wide mutant fitness assays and differential gene expression datasets. These two data types are complementary and allow us to narrow down which genes are critical to the pathway by looking at the intersection of fitness and expression phenotypes.

2) We characterized a metabolic pathway for PPME degradation in *Pseudomonas fluorescens* SBW25 and identified a novel phenylpropanoid methyl esterase enzyme that performs the first step in PPME catabolism. The combination of genome-wide functional genomics data, differential transcriptomics data, and the generation of single-gene mutants in SBW25 allowed us to devise a putative catabolic pathway for PPMEs that funnels into a characterized pathways for *p*-coumaric acid catabolism.

3) We showed this metabolic pathway can be transferred into multiple heterologous hosts, including two sorghum rhizosphere isolates. This demonstrates the ability to discover and transfer plant root exudate pathways to metabolically naïve microbes, an initial step to creating an engineered microbial dependence on root exudate compounds.

4) We isolated three microbes from sorghum field soil that catabolize the highly expressed allelochemical and biological nitrification inhibitor sorgoleone as a sole carbon source. These three isolates from distinct genera (*Acinetobacter*, *Burkholderia*, and *Pseudomonas*) share a common four-gene operon we named *srgABCD*. Transcriptomic analyses in the three hosts showed this gene cluster is significantly overexpressed in sorgoleone growth conditions compared to an acetate growth control.

5) Gene knockout studies in the Pseudomonad isolate dubbed *Pseudomonas sorgoleonovorans* SO81 confirmed that the conserved gene cluster is essential for growth. Genome-wide fitness studies revealed additional genes that contribute to host fitness when cultivated on sorgoleone. Combined with the transcriptomic analysis, we have a collection of eight genes to investigate a complement of genes that provide sorgoleone catabolism phenotypes for heterologous hosts.

6) To better understand the distribution and prevalence of the *srg* cluster in environmental isolates, we leveraged Snekmer protein family models [3] and the GFOBAP collection of 3,837 genomes derived from soil- and plant-associated environments [29] to search for the presence of *srg* genes. We found an enrichment for the *srg* cluster among actinobacteria, primarily among the genus *Streptomyces*. Further, we compared *Streptomyces* genomes from isolates derived from sorghum and poplar fields (**Figure 6**) to find a significant enrichment for *srg* genes among sorghum isolates. The presence of *srg* enabled growth on sorgoleone in all strains from our isolate collection when tested; no sorghum isolates we tested that lack the *srg* cluster could utilize sorgoleone for growth.

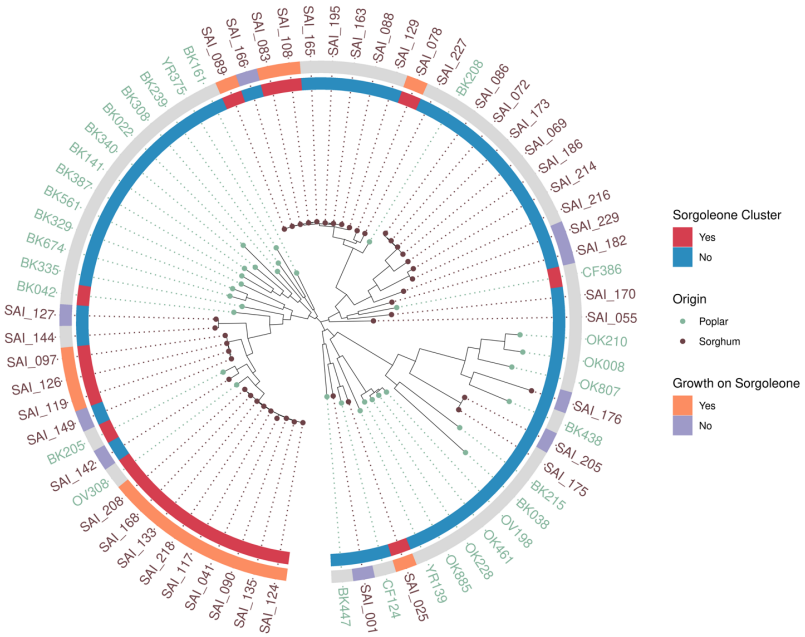

Figure 6. Phylogenetic analysis and growth of selected *Streptomyces* strains on sorgoleone. Outside to inside, the rings represent strain-plant association, growth phenotype on sorgoleone for a subset of strains, and the presence/absence of the *srg* gene cluster from whole-genome sequence data. The tree at the center represents phylogenetic distance.

Conclusions and Next Steps

Our SFA progress positions us to address our decadal vision to *achieve predictive understanding of rhizosphere microbiome community structure and function to enable the design of secure, co-dependent plant–microbe ecosystems that underpin highly productive biomass cropping systems through dynamic provision of plant-growth-promoting traits*. The establishment of our defined synthetic community provides a collection of strains with known contributions to sorghum plant health from which to develop genetic tools to test our hypotheses, to test the limits of persistence control approaches for containing engineered strains in complex environments, and to understand synthetic community colonization and growth promotion dynamics in laboratory and native rhizosphere environments. Our comprehensive functional genomics datasets for multiple *Pseudomonas* isolates provide well-studied and hypothetical gene targets for genome reduction to reduce environmental fitness in bulk soil. We can apply our discovery of a conserved pathway for sorgoleone catabolism to fully elucidate the pathway and provide a new niche in the sorghum rhizosphere to engineered strains.

Understanding how microbial persistence in soil and rhizosphere environments is affected by microbiome and plant interactions is an important scientific frontier with broad translational potential. Over the next three years, the PerCon SFA has scoped out a research challenge to *create fundamental understanding of microbial gene functions for survival and plant-growth promotion in soil and rhizosphere environments to establish design principles for controlling the persistence of engineered plant-growth-promoting rhizosphere microbiomes.* We have structured our research plan into three overarching investigation areas with integrated research aims to (1) demonstrate how metabolic pathways, regulatory systems, and engineered functions can be manipulated to enhance fitness and enforce the confinement of diverse bacteria in field-simulant media, (2) identify interaction mechanisms within microbial communities and apply this understanding to evaluate the efficacy of persistence control strategies to restrict metabolic and genetic exchange, and (3) determine how spatial localization and developmental timing affect metabolite production and bacterial colonization in the rhizosphere and how colonization dynamics affect host phenotypes.

Our proposed work directly addresses BER mission objectives and aspects of recent BER workshops aimed at understanding Gene Function across Taxa [30], Genomes to Structure and Function [31], and Artificial Intelligence and Machine Learning for Bioenergy Research [32]. Further, it advances the Secure Biosystems Design program mission [33] by (1) understanding the genetic determinants of microbial growth and survival in complex soil–plant ecosystems; (2) developing microbial engineering approaches to control persistence by modulating the fitness of rhizosphere microbes outside the laboratory; and (3) establishing innovative biosystems design approaches to increase crop resilience by engineering persistence of beneficial rhizosphere communities.

Creating persistence control phenotypes for rhizosphere environments requires functional annotation of thousands of genes across bacteria taxa, which will provide high-throughput data resources for new functional annotation models. Related, we anticipate our machine learning models to understand mechanisms of genetic and metabolic complementation through improved protein functional annotation will help bridge gaps between protein structure and function. More broadly, introducing secure microbial functions to the rhizosphere will promote plant productivity and stress tolerance in support of strengthening the U.S. bioeconomy[34, 35]. Approaches to reinforce the niche of an engineered microbe in a rhizosphere should promote complementary goals of increased biomass crop yields and minimized risks to native ecosystems through responsible deployment of engineered functions. We anticipate our research findings will extend to cropping systems beyond sorghum. Sustained commitment to the development of secure biosystems design for the rhizosphere will position researchers, policymakers, and industry to harness the full potential of microbiomes and designer ecosystems for future cropping needs and environments.

References

1. Elmore, J.R., et al., *High-throughput genetic engineering of nonmodel and undomesticated bacteria via iterative site-specific genome integration.* Sci Adv, 2023. **9**(10): p. eade1285.
2. Garcia, C.F., et al., *Defined synthetic microbial communities colonize and benefit field-grown sorghum.* bioRxiv, 2023: p. 2023.05.30.542977.
3. Chang, C.H., et al., *Snekmer: a scalable pipeline for protein sequence fingerprinting based on amino acid recoding.* Bioinform Adv, 2023. **3**(1): p. vbad005.
4. Handakumbura, P.P., A. Rivas Ubach, and A.K. Battu, *Visualizing the Hidden Half: Plant-Microbe Interactions in the Rhizosphere.* mSystems, 2021. **6**(5): p. e0076521.
5. Kroll, J.O., et al., *Profiling sorghum-microbe interactions with a specialized photoaffinity probe identifies key sorgoleone binders in Acinetobacter pittii.* bioRxiv, 2023: p. 2023.05.31.542302.
6. Oda, Y., et al., *Sorgoleone degradation by sorghum-associated bacteria; an opportunity for enforcing plant growth promotion*. 2023, Cold Spring Harbor Laboratory.
7. Wilson, A., et al., *A novel phenylpropanoid methyl esterase enables catabolism of aromatic compounds that inhibit biological nitrification.* bioRxiv, 2023: p. 2023.06.02.543320.
8. Elmore, J.R., et al., *High-throughput genetic engineering of non-model and undomesticated bacteria via iterative site-specific genome integration* Science Advances (in press), 2023.
9. Price, M.N., et al., *Mutant phenotypes for thousands of bacterial genes of unknown function.* Nature, 2018. **557**(7706): p. 503-509.
10. Salzberg, S.L., *Next-generation genome annotation: we still struggle to get it right.* Genome Biol, 2019. **20**(1): p. 92.
11. Lobb, B., et al., *An assessment of genome annotation coverage across the bacterial tree of life.* Microb Genom, 2020. **6**(3).
12. Yim, S.S., et al., *Multiplex transcriptional characterizations across diverse bacterial species using cell-free systems.* Molecular Systems Biology, 2019. **15**(8).
13. McClure, R., et al., *Development and Analysis of a Stable, Reduced Complexity Model Soil Microbiome.* Front Microbiol, 2020. **11**: p. 1987.
14. Zegeye, E.K., et al., *Selection, Succession, and Stabilization of Soil Microbial Consortia.* mSystems, 2019. **4**(4).
15. Gilbert, N., *Summit urged to clean up farming.* Nature, 2011. **479**(7373): p. 279-279.
16. Yang, X.L., et al., *Optimising nitrogen fertilisation: A key to improving nitrogen-use efficiency and minimising nitrate leaching losses in an intensive wheat/maize rotation (2008-2014).* Field Crops Research, 2017. **206**: p. 1-10.
17. Subbarao, G.V., et al., *Suppression of soil nitrification by plants.* Plant Science, 2015. **233**: p. 155-164.
18. Moss, B., *Water pollution by agriculture.* Philosophical Transactions of the Royal Society B-Biological Sciences, 2008. **363**(1491): p. 659-666.
19. Stein, L.Y. and M.G. Klotz, *The nitrogen cycle.* Current Biology, 2016. **26**(3): p. R94-R98.
20. Prasad, R. and J.F. Power, *Nitrification Inhibitors for Agriculture, Health, and the Environment.* Advances in Agronomy, Vol 54, 1995. **54**: p. 233-281.
21. Abalos, D., et al., *Meta-analysis of the effect of urease and nitrification inhibitors on crop productivity and nitrogen use efficiency.* Agriculture Ecosystems & Environment, 2014. **189**: p. 136-144.
22. Yan, D., et al., *Genetic modification of flavone biosynthesis in rice enhances biofilm formation of soil diazotrophic bacteria and biological nitrogen fixation.* Plant Biotechnol J, 2022.

23. Bloch, S.E., et al., *Harnessing atmospheric nitrogen for cereal crop production.* Current Opinion in Biotechnology, 2020. **62**: p. 181-188.
24. Li, Q. and S.F. Chen, *Transfer of Nitrogen Fixation (nif) Genes to Non-diazotrophic Hosts.* Chembiochem, 2020. **21**(12): p. 1717-1722.
25. Larrea-Alvarez, M. and S. Purton, *The Chloroplast of Chlamydomonas reinhardtii as a Testbed for Engineering Nitrogen Fixation into Plants.* International Journal of Molecular Sciences, 2021. **22**(16).
26. Subbarao, G.V., et al., *Biological nitrification inhibition (BNI) activity in sorghum and its characterization.* Plant and Soil, 2013. **366**(1-2): p. 243-259.
27. Gopalakrishnan, S., et al., *Nitrification inhibitors from the root tissues of Brachiaria humidicola, a tropical grass.* Journal of Agricultural and Food Chemistry, 2007. **55**(4): p. 1385-1388.
28. Yuan, S.Z., et al., *Defense Responses, Induced by p-Coumaric Acid and Methyl p-Coumarate, of Jujube (Ziziphus jujuba Mill.) Fruit against Black Spot Rot Caused by Alternaria alternata.* Journal of Agricultural and Food Chemistry, 2019. **67**(10): p. 2801-2810.
29. Levy, A., et al., *Genomic features of bacterial adaptation to plants.* Nature Genetics, 2018. **50**(1): p. 138-150.
30. Buell, R., et al., *Breaking the Bottleneck of Genomes: Understanding Gene Function Across Taxa Workshop Report*. 2019, USDOE Office of Science (SC).
31. DOE Biological and Environmental Research Program, *Genomes to Structure and Function Workshop Report 2022*. 2022.
32. Zhao, H., et al., *Artificial Intelligence and Machine Learning for Bioenergy Research: Opportunities and Challenges*. 2023, U.S. DOE Office of Science and Office of Energy Efficiency and Renewable Energy.
33. DOE Biological and Environmental Research Program. *Secure Biosystems Design*. 2023 June 7, 2023]; Available from: https://genomicscience.energy.gov/secure-biosystems-design/.
34. Biomass R&D Board, *Federal Activities Report on the Bioeconomy*. 2016: CreateSpace Independent Publishing Platform.
35. Langholtz, M.H., B.J. Stokes, and L.M. Eaton, *2016 Billion-Ton Report: Advancing Domestic Resources for a Thriving Bioeconomy.*

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov