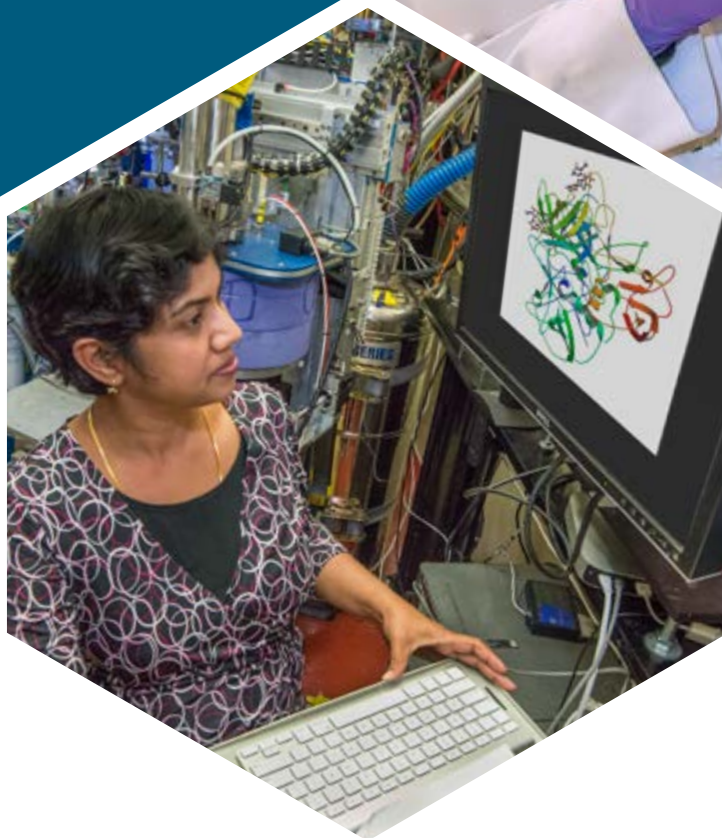**Lawrence Berkeley National Laboratory**

BIOSCIENCES 10-YEAR SCIENTIFIC STRATEGIC PLAN 2013–2023

2016

BioSciences

Lawrence Berkeley National Laboratory

BIOSCIENCES 10-YEAR SCIENTIFIC STRATEGIC PLAN

2013–2023

REVISION 2016

BioSciences

INTRODUCTION

As Associate Laboratory Director for Biosciences, I am fortunate to lead a part of Berkeley Lab with enormous talent, capabilities, and potential. As national laboratories across the country have augmented their historical strengths in physical and computational sciences with biosciences to solve national-scale problems, Berkeley Lab has developed powerful new and continually improving bioscience capabilities, teams, and approaches to address global-scale challenges.

Biological research is changing dramatically as a consequence of powerful technologies developing along with new concepts and methods derived from the integration of physical sciences, mathematics, computational sciences, and engineering. Advances in biological sciences hold tremendous promise for realizing solutions to the many major problems confronting the nation and the world, such as securing access to affordable, sustainable sources of energy, as well as for maintaining the health of the planet and the people who live on it.

Historically, advances in basic science have provided foundations for the development of solutions to economic and social challenges. However, societal challenges are growing ever more complex, and new approaches are needed to tackle these expanding complexities. Because these problems have become too large to be solved by single-investigator laboratories, inter- and multidisciplinary teams offer powerful new synergies with which to attack global-scale problems. Berkeley Lab's Biosciences Area is particularly well suited to solve some of these large problems given our unique history in understanding and solving energy and environmental problems, our remarkable assets, and our ability to work effectively in multidisciplinary teams.

In developing the 10-Year Scientific Strategic Plan (Biosciences Strategic Plan), the Biosciences Area harnessed the intellectual capital of our team to create a shared vision for the future that captures our passions and will guide our research efforts. As we endeavor to solve national-scale problems in energy production and environmental contamination, and negative human health impacts from such contamination, and as we tap into the power of biology as a manufacturing platform, we are aiming high.

After our Biosciences Strategic Plan in 2013, our second major strategic undertaking was the re-organization of the Biosciences Area in 2015 creating a new leadership and management structure to enable Biosciences to achieve its 10-year Goals. Once again, we called upon the 800+ Biosciences members to help develop missions and visions for new

Divisions that better position the Area for program development and implementation of the Biosciences Strategic Plan, and to create an operations structure that reduces costs and better enables research, communication, collaboration, and safety. I'm very pleased with the results of these efforts. In addition to our new Divisions and reduced costs, new cross-divisional groups emerged from our re-organization, including an Area Informatics Group, a Communications Team, an Early Career Scientist Organization, and an Area Seminar Committee, all of which strengthen our Area and foster new collaboration opportunities.

Here's a quick look at our new Divisions:

The **Biological Systems and Engineering Division** aims to advance a mechanistic and predictive understanding of complex biological systems and responses to environmental challenges, and translate this knowledge using engineering principles to develop resilient systems, tools, and processes for the efficient production of fuels, chemicals, materials, tissues, and biohybrids.

The **Environmental Genomics and Systems Biology Division** links genome biology to ecosystem dynamics to develop a mechanistic and predictive understanding of ecosystem function to enable sustainable solutions to energy and environmental challenges.

The **Molecular Biophysics and Integrated Bioimaging Division's** mission is to generate a mechanistic and predictive understanding of biological processes, by developing and applying molecular- and meso-scale visualization and advanced spectroscopies, enabling the control, manipulation, and generation of biological function.

In this 2016 version of the 2013 plan, our vision has been updated to align with our newly re-organized Area and to maximize success of achieving the 10-year Goals. Substantive revisions of this plan include new Environment strategies that drive efforts to achieve the Environment Goal, and a new section, "Technologies for Biosciences," that describes the forward path for the foundational cross-cutting technologies that power the research across our Area.

Our third strategic undertaking, begun on the heels of our re-organization, was the recent development of the Biosciences Campus Vision, an effort driven by the need to co-locate our 800+ Biosciences members at the main Berkeley Lab site to enhance collaboration among Biosciences research teams and maximize opportunities for successful achievement of our Goals through proximity to the main campus assets, such as the Advanced Light Source, the Molecular Foundry, and the National Energy Research Scientific Computing Center, to name a few. I'm thrilled that the first of the five new buildings we've planned, the Integrative Genomics Building, which will house two complementary Department of Energy (DOE) programs — the DOE Joint Genome Institute and the Systems Biology Knowledgebase — will begin construction soon.

As I look to the future and what Biosciences can bring to it, I understand that the achievement of our vision depends not only on excellent science but also on a diverse and passionate workforce supported by informed scientific and public communities. As we pursue the scientific strategies needed to serve the nation and achieve our Goals, we must simultaneously enhance our outreach and education efforts accordingly. Achievement of our vision for the future rests not only on effective research teams, but also on teams that include dedicated educators, students, citizens, and policymakers.

I am confident that the talented teams of bioscientists at Berkeley Lab will achieve many of the Goals outlined in this document and, as a result, improve the U.S. economy and make the world a much better place.

Jay Keasling
Associate Laboratory Director
Lawrence Berkeley National Laboratory

Lawrence Berkeley National Laboratory Biosciences

10-YEAR
SCIENTIFIC STRATEGIC PLAN
2013–2023

REVISION 2016

Lawrence Berkeley National Laboratory's (Berkeley Lab) *Biosciences' 10-Year Scientific Strategic Plan* (Biosciences Strategic Plan) describes the vision for a national future strengthened by biological research achievements, and provides guidance for biosciences research activities at Berkeley Lab. It establishes a framework for maintaining that vision and achieving these Goals from 2013–2023. We describe here the large-scale biological science challenges appropriate for a national laboratory and relevant to Berkeley Lab's own mission and values. The Biosciences Strategic Plan details ambitious Goals and relies on our capacity for multidisciplinary, collaborative research to bring bioscience solutions to the world. Here we sharpen our focus on the uses of bioscience to address the energy needs of our nation, protect the environment, understand and improve health, and develop novel biomanufacturing technologies. Our plan is meant to be both a blueprint and a catalyst for achieving these Goals. This 2016 version of the 2014 plan reflects a vision that has been updated and aligned with our newly re-organized Area and its three new Divisions. Since the previous version, we've added a new section, "Technologies for Biosciences," that describes a forward path for the foundational cross-cutting technologies that power the research across our Area.

TABLE OF CONTENTS

MISSION

Use integrated research teams to solve national challenges in energy, environment, health, and biomanufacturing

VISION

Berkeley Lab's Biosciences Area will lead the nation in using biology to solve energy and environmental challenges

A STRATEGIC FRAMEWORK

Our strategy is grounded in the national laboratories' core mission to carry out basic research that addresses the nation's most pressing science and technological challenges. Our plan is informed and guided by Berkeley Lab values:

• Overarching commitment to pioneering science

• Highest integrity and impeccable ethics

• Uncompromising safety

• Sense of urgency

• Diversity in people and thought

With that mission and these values in mind, more than 800 Biosciences researchers and staff participated in an inclusive process to develop this plan, using surveys, focus groups, formal meetings, and informal discussions. Here we present a strategy to enhance and exploit key competencies associated with our discovery-sciences mission. Our overarching goal is to use basic research in bioscience to discover and illuminate paths toward practical solutions. We plan to accomplish our objectives with a strategy built on a framework of historical strengths and well-defined metrics of future success.

Four Key Scientific Challenges

This 10-year Biosciences Strategic Plan focuses on four research areas: Energy, Environment, Health, and Biomanufacturing. In these areas, Berkeley Lab expertise is deep, the national need is great, and our commitment to solve problems is unshakeable. In each of these areas, our culture of team science and cross-disciplinary research can be brought to bear for maximum efficiency. Here are the challenges in each area.

Energy Research

How can we efficiently, cost-effectively, and sustainably transform the energy in lignocellulose and plentiful gas feedstocks to liquid transportation fuels using biological or bio-inspired approaches?

Environment Research

How can we develop principles to understand and model the dynamic reciprocity between organisms and their environments to predict how they will respond to climate change? How do we manipulate environmental systems to remediate contaminants, sequester carbon, and support agricultural productivity?

Health Research

How do environmental challenges affect the health of diverse organisms in complex biological systems? What are the effects of environmental changes and exposures, particularly due to anthropogenic impacts and perturbations, on biological systems and human health? Can we use fundamental knowledge of biological diversity in response to exposure to enable prediction, prevention, and treatments?

Biomanufacturing Research

Can significantly reducing the cost and increasing the speed of engineering biological systems transform manufacturing in the United States? Can we solve challenges in energy, agriculture, health, and environment with new biomanufacturing approaches?

Progress That Is Measurable

In each of these four research areas, our strategic plan specifies 10-year metrics and five-year milestones to assess progress toward and achievement of the four 10-year Goals, outlined below. Metrics along the full spectrum of research — from basic to applied — were selected to describe our vision for success and underscore our commitment to the breadth of scientific achievement from early discovery to applied solutions for each of our four primary research endeavors.

New learning and discoveries underpin successful approaches to tackle our increasingly complex scientific and societal challenges. Basic research is not intended to lead to immediate commercial benefit, but to new knowledge and theory. Basic research and discovery have historically played a foundational role in technological innovation.

A comprehensive understanding of a biological system enables predictions of how it will respond under certain conditions. It makes possible the reconstruction and redesign of components of the system — capabilities needed to move discoveries closer to solutions that address societal challenges.

At Berkeley Lab, we are dedicated to "bringing science solutions to the world." Technologies developed at Berkeley Lab have generated billions of dollars in revenue and thousands of jobs. Berkeley Lab breakthroughs in energy-sparing technologies, such as more efficient lighting and windows, have also saved billions of dollars for industry and consumers.

With our metrics in place at the outset, we'll track our progress toward success in each of the four research areas along this "discovery-to-solution" paradigm.

Forward Looking, Secured by the Past

The long-term Goals developed in this planning process tap into the core scientific competencies established at the founding of Berkeley Lab over 80 years ago and strengthened over decades. These competencies, which sprang from the Laboratory's focus on physical sciences and the synergistic academic environment offered by the University of California (UC) at Berkeley, evolved early in its history to include state-of-the-art biological science. Berkeley Lab investigators played pioneering roles in recent revolutions in genomics, computation, biophysics, synthetic biology, and imaging that are continuing to change the way biological research is conducted.

Our 10-year strategy will enhance Berkeley Lab's role as a leading center for the use of biosciences to meet national objectives for energy, environment, health, and biomanufacturing. As we execute our 10-year plan, we carry forward a legacy of transformative research. From the diversity of microbes, microbial communities, and plants, we will continue to uncover nature's secrets to gain a deeper insight into how biological systems work, how they interact with each other and with their environment, and how they can be manipulated to harness their processes and products. From the potential encoded in an organism's genome, we will work to define the principles that guide the translation of the genetic code into functional proteins and pathways. We will continue to advance our understanding of the metabolic and regulatory networks that underlie the systems biology of plants, animals, and microbes as they respond to their environments. Inspired by these processes, we will explore biological means to manufacture new and alternate sources of energy and materials that require less energy to produce and can restore balance to natural carbon cycles.

Biosciences at Berkeley Lab has been successful by many measures — high-impact scientific journal articles, promising technologies transferred to industry, multiple spinoff companies — and these successes bring a new opportunity to expand our impact. By extending existing capabilities and combining them in new ways, we hope to make rapid progress toward solving essential challenges in energy, the environment, and health. A significant impediment to progress has been the distribution of biosciences effort between five sites in Berkeley, Emeryville, and Walnut Creek. Accordingly, an important element of this 10-year strategic plan is consolidation of all biosciences activities at a single site to enable increased interactions, communication, and the efficient development of new technological infrastructure.

While basic biological research is the foundation for future technologies, it also provides the evidence for informed policymaking on topics of critical importance. Results of biosciences research efforts can provide guidance for decision-makers who, for example, must assess the present and future impacts of climate change and other environmental challenges, evaluate new avenues to energy independence, and develop new medical technologies.

Our vision for biosciences at Berkeley Lab depends on more than merely setting scientific Goals for the future. To achieve the 10-year Goals outlined in this plan, and to meet the needs of the nation, we must strengthen several supporting activities: the transfer of

promising technologies to the private sector to create public benefit; our outreach efforts to develop an enhanced public understanding of the science that will provide solutions for our future; and the inspiration and education of a new generation of diverse scientists to recognize future national needs and achieve the objectives required to meet them. We must facilitate research at the intersections of diverse scientific disciplines to create environments that promote creative thinking, attract the brightest and most inquisitive scientists, and accelerate transformative discoveries. Furthermore, underpinning our efforts at Berkeley Lab is the leadership we demonstrate in developing new best practices for ensuring that environmental/ecological, biosecurity, social justice, and ethical concerns are considered in advance of research and development. Our work spans security and implications of human genomic sequencing for privacy and predictions, on the consequences of biomanufacturing innovations and who controls production, how the volatility of production may be lessened, and the disruption of certain professions and the inception of others. We pride ourselves on building trust with the public through transparency and direct conversation.

BIOSCIENCES AT BERKELEY LAB

While Berkeley Lab may be best known for its physics, the biological sciences have been part of its DNA almost from the beginning, when founder and namesake Ernest O. Lawrence recruited top-flight scientists to UC Berkeley in the 1930s.

Lawrence's younger brother John, a physicist and physician, is considered the father of nuclear medicine. At Berkeley Lab, John studied the biological effects of the byproducts of the atom smashers Ernest built, and carried out the first successful treatment of human disease with radioisotopes. Today nuclear medicine still plays a central role in the diagnosis and treatment of cancer and other human diseases, and today's health-related scientists at Berkeley Lab are building on these foundations in their research efforts to better understand cancer, DNA repair, genome structure and function, and neurodegenerative diseases.

Biochemist Melvin Calvin used radioactive carbon-14 from a Berkeley Lab cyclotron to map the route that carbon travels through a plant during photosynthesis — research that led to discovery of the "Calvin cycle" and the Nobel Prize in Chemistry in 1961. Today's physical bioscientists and engineers at Berkeley Lab are building on advances in the physical sciences and modern biology, including those of Calvin, to examine, characterize, and mimic biological molecules and molecular functions to create unique biological structures that can then be used to solve some of the 21st century's most difficult fundamental research problems.

Berkeley Lab conducted path-breaking research on medical imaging, including early development of computed tomography (CT) scans and positron-emission tomography, (PET) scans. Cancer studies broadened to include tracking the behavior of healthy and malignant cells in culture and animals, pioneering the development of 3-D human tissue models, defining cancers as diseases of tissue microenvironments, and identifying many of the impacts of radiation on cells and organisms. Studies of heart disease and Alzheimer's disease helped to characterize the role of oxygen radicals in aging and disease. Bioscience research at Berkeley Lab deepened our understanding of what was becoming known as "systems biology."

The extensive work in biological sciences and pioneering studies on mapping and sequencing the model organism *Drosophila melanogaster* genome led to selection of Berkeley Lab as one of five centers for the Human Genome Project, the massive national effort to map and sequence the entire complement of human DNA. Berkeley Lab's Human Genome Center, which was consolidated into the Department of Energy Joint Genome Institute (the DOE JGI) in Walnut Creek, was responsible for sequencing

a significant portion of the human genome. Since that time, the DOE JGI has undertaken a considerable effort to determine the genome sequences of thousands of plants and microorganisms with the aim of using this genomic information to develop solutions to national-scale energy and environment challenges.

Aided by faster computers and more advanced algorithms, studies of gene regulation intensified. Berkeley Lab played a major role in the Model Organism Encyclopedia of DNA Elements Project, which resulted in greatly improved genome annotations and scientific understanding of non-protein coding RNAs, chromatin "landscapes," and genome functions. Rapid sequencing renewed interest in proteins, including how they are structured and how they work. X-ray crystallography at the Advanced Light Source, plus a range of powerful microscopic techniques, revealed structures of important proteins at the highest resolutions ever.

The focus on genetics and molecular biology developed naturally toward the discipline now called synthetic biology, which holds the promise of reducing dramatically the costs and time required to design, build, and characterize biological systems. These innovations have led to focused applications and the creation of a number of spinoff companies.

Under the direction of then Laboratory Director Steven Chu, the Nobel laureate who would become President Barack Obama's Secretary of Energy in 2009, Berkeley Lab embarked on an intensive effort to use the tools of genetics, supercomputers, and microbiology to develop biofuels and new sources of sustainable energy. The Joint BioEnergy Institute is one of three national centers created by the DOE in 2007 to advance the development of biofuels. Building on a legacy of advanced research in biosciences, Berkeley Lab has the infrastructure and expertise to bring biological solutions to the energy, health, and environmental challenges of our time as well as provide the foundational underpinnings for a strong biological manufacturing industry.

In addition to our focus on using science to bring solutions to the world, our strategy also embraces a Berkeley Lab commitment to transferring our knowledge to our surrounding communities. We will continue to combine our research efforts with efforts to reach out to our neighbors. Through workshops, internships, and educational programs at local schools, colleges, and universities, we will promote understanding of science and encourage young people of diverse backgrounds to make a career in biosciences part of their own strategic plans.

THE NEXT 10 YEARS

MISSION

Use integrated research teams to solve national challenges in energy, environment, health, and biomanufacturing

VISION

Berkeley Lab's Biosciences Area will lead the nation in using biology to solve energy and environmental challenges

HEALTH

ENVIRONMENT

ENERGY

BIOMANUFACTURING

TECHNOLOGIES

STRATEGIC GOALS

Biosciences for Energy

Develop/enable cost-competitive (economically-sustainable) and environmentally-sustainable biological and bio-inspired energy solutions capable of reducing U.S. dependence on petroleum.

Biosciences for the Environment

Understand the genetic and molecular mechanisms governing the activities and ecology of organisms and multispecies communities to predict responses and harness microbes and plants for energy and environmental solutions.

Biosciences for Health

Develop and apply a predictive, multiscale, integrative understanding of how biological diversity impacts responses to environmental challenges, to improve human and biome health and drive responsible economic growth.

Biosciences for Biomanufacturing

Develop and demonstrate a scalable, flexible, cost-effective, and sustainable biology-based manufacturing infrastructure driven by applications in energy, health, the environment, and agriculture.

Technologies for Biosciences

Develop a technology infrastructure to measure, predict, and control biological systems for solving energy, environmental, and health challenges.

BIOSCIENCES FOR ENERGY

Background and Motivation

Energy production is the world's largest industry and is based almost exclusively on fossil fuels, whose extraction and subsequent burning pollute land, water, and the atmosphere. As "clean" sources of fossil fuels are depleted, less desirable sources are being tapped that require significantly more energy to produce and may pollute the environment in other ways.

The development of alternative energy sources is a pressing national need, a major mission of DOE's Office of Science, and a central thrust of Berkeley Lab's strategic plan since 2006. This effort now involves all of Berkeley Lab's Biosciences Divisions and receives more than $60 million in annual funding. Successful production of scalable bio-based transportation fuel sources will be most rapidly achieved through an integrated, team-science approach, which has been the hallmark of Berkeley Lab since its inception.

Biology has the potential to produce energy renewably, particularly liquid hydrocarbon fuels with the high energy density needed by the U.S. transportation infrastructure. However, biological mechanisms are relatively inefficient at harnessing sunlight and transforming it to hydrocarbons. Converting the full gamut of the lignocellulosic and non-lignocellulosic material to high-density fuels and replacements to petrochemicals is also necessary to make a meaningful difference in the current paradigm of bioenergy and biomanufacturing. A better understanding of photosynthesis, cell-wall synthesis in plants, and the hydrocarbon-forming pathways in all organisms would make it possible to build predictive models of these processes. Such predictive models could be used to engineer plants to capture sunlight and use *in situ* nutrient resources more efficiently, be resilient to otherwise harsh environments, and alter sugars accumulation in cell walls and enhance their breakdown and use. Such models could

10-YEAR
Energy Goal

Develop/enable cost-competitive (economically-sustainable) and environmentally-sustainable biological and bio-inspired energy solutions capable of reducing U.S. dependence on petroleum.

Energy Research Strategies to Achieve Goal

Lignocellulosic biofuels. Derive energy from biomass with new technologies.

Alternative biofuels. Engineer and scale the direct conversion of gas feedstocks to fuels by photosynthetic, methanotrophic, and chemoautotrophic microorganisms.

Artificial photosynthesis. Use bio-inspired reactions to create fuels directly from atmospheric CO_2 and sunlight.

also contribute to the successful engineering of microorganisms to convert sugars, aromatics, and other carbon sources more efficiently into drop-in biofuels and fuel additives compatible with the transportation fleets of today and tomorrow, and leverage the progress in designed biological systems also critical for biomanufacturing. In addition, these efforts could facilitate design of bio-inspired catalysts that mimic photosynthesis to produce transportation fuels directly from sunlight and CO_2.

The time pressure to produce cost-competitive non-ethanol advanced biofuels derives not just from the increasing environmental effects of fossil fuel combustion, but also from likely political pressure on the biofuels industry given the relatively slow pace of progress in the lignocellulosic biofuels arena.

The Biosciences Area's approach to this problem is to develop a molecular description of the biological processes of photosynthesis, cell-wall synthesis, biomass breakdown, microbial metabolism, and hydrocarbon biosynthesis. This new knowledge can subsequently be used to meet key strategic objectives in the areas of lignocellulosic biofuels production, alternative biofuels development, and artificial photosynthesis

Energy Research Strategies

- **Lignocellulosic biofuels.** Derive energy from biomass with new technologies.

- **Alternative biofuels.** Engineer and scale the direct conversion of gas feedstocks to fuels by photosynthetic, methanotrophic, and chemoautotrophic microorganisms.

- **Artificial photosynthesis.** Use bio-inspired reactions to create fuels directly from atmospheric CO_2 and sunlight.

To achieve the Energy Goal by 2023, Berkeley Lab's approach employs three areas of strategic focus: production of fuels from plant biomass (lignocellulosic biofuels), production of fuels from gas feedstocks (alternative biofuels), and non-biological production of liquid fuels directly from sunlight and CO_2 (artificial photosynthesis). These areas were chosen because (1) they collectively have potential to meet the long-term national need for sustainable cost-competitive alternatives to fossil fuels; (2) they are scientifically tractable within a 10-year span; and (3) they leverage specific facilities, organized research groups, and core competencies that exist, or can be readily assembled, within the Biosciences Area. These strategies will be executed in parallel.

Lignocellulosic biofuels: Derive energy from biomass with new technologies

The cellulosic biofuels strategy outlined by DOE in the 2005 *Billion Ton Study* and 2011 *U.S. Billion-Ton Update* highlighted the amount of feedstocks required for the development of cellulosic biofuels' capacity to replace 30% of U.S. needs for transportation fuels, without significant impacts on human food and livestock feed production. The success of this strategy depends on:

- Improved biomass feedstocks with greater yields of fermentable sugars and valuable lignin intermediates, greater tolerance to stress, and improved nutrient acquisition aimed at reducing energy input and pollution and to allow for wider land use and less volatile production due to environmental variation.

- Greatly improved biomass extraction and breakdown strategies.

- Engineering of microorganisms capable of converting biomass sugars and lignin aromatics to high energy density hydrocarbon fuels and additives compatible with the current and future transportation fleet, including jet aircrafts.

- The development and demonstration of economically viable and scalable production processes for an advanced biofuel and co-product, including feedstock deconstruction, fermentation, chemical catalysis, and product recovery and purificatio.

Efforts on these fronts are underway at Berkeley Lab, primarily funded through the Joint BioEnergy Institute (JBEI) and individual research programs funded by Biological and Environmental Research in the Office of Science at DOE. These efforts are also supported through collaborative interactions between JBEI, the DOE JGI, and the DOE Systems Biology Knowledgebase (KBase), and particularly with the Advanced Biofuels and Bioproducts Process Demonstration Unit (ABPDU), funded by the DOE Office of Energy Efficiency and Renewable Energy.

JBEI focuses on the development of advanced bioenergy plants, efficient chemical and biological processes to extract sugars and lignin aromatics from these plants, and efficient biological conversion of the resulting sugars to advanced biofuels and co-products. This team of nearly 200 scientists has already engineered bioenergy crop plants with more cellulose and less or altered lignin; developed deconstruction processes based on ionic liquids to deliver clean cellulose and hemicellulose that can be more readily depolymerized into sugars using fewer enzymes than other deconstruction processes; and engineered microorganisms to produce drop-in biofuels for gasoline, diesel, and jet engines. In this effort, the broad retrosynthesis infrastructure, progress in tools development for Biomanufacturing would allow the improved prediction of potential pathways to a hydrocarbon product, while progress in designed biological systems would allow development of the corresponding plants and microbes. Current efforts are also directed at utilizing lignin aromatics and producing co-products in addition to fuels, thereby rendering the downstream biorefinery process more cost-competitive.

The DOE JGI's efforts have effectively complemented those of JBEI. The DOE JGI has produced more de novo plant, fungal, and bacterial genomes than any other single genome-sequencing center. Many of these genomes have direct relevance to Berkeley Lab's lignocellulosic biofuels strategy, providing the bases for molecular understandings of relevant pathways, and sources of molecular tools for engineering strategies. In addition, the DOE JGI is a world leader in the field of metagenomics, pioneering genome assembly from complex microbial communities. It has leveraged skills in this arena to mine biomass-degrading activities from unique environments ranging from termite gut to hoatzin gizzard, and from Yellowstone hot springs to tropical rainforests, which offer new avenues for biofuels investigations. Further, the DOE JGI has pioneered large-scale DNA synthesis and assembly in the service of characterizing these enzymatic activities and enabling manufacture of fuel production pathways under its new synthetic biology efforts.

Much remains to be done. Proof-of-concept studies in cell-wall modification, advanced biomass pretreatment (e.g., ionic liquids), and fractionation must be developed and demonstrated; cell-wall-degrading enzyme discovery must be extended; and processes developed and scaled for licensing to the cellulosic-biofuels industry. Finally, several inefficiencies remain in converting sugars to drop-in biofuels. Process efficiency research

and scaling are being addressed in the bench-to-pilot scale development and demonstration environment at the ABPDU in collaboration with JBEI, industry partners, and other national labs.

Alternative biofuels: Engineer and scale the direct conversion of gas feedstocks to fuels by photosynthetic, methanotrophic, and chemoautotrophic microorganisms

The goal of the lignocellulosic biofuels strategy is to produce fuels from depolymerized lignin and sugars derived from plant cell walls. However, sugars and lignin intermediates derived from lignocellulosic biomass are complex and often contain contaminants that in the past have prevented them from being efficiently fermented by microorganisms. Processes to produce biofuels from biomass-derived sugars typically require multiple unit operations that make the bioconversion more expensive. Therefore, in addition to feeding microorganisms sugars and aromatics derived from plant biomass to create fuels, Berkeley Lab is also focused on another set of pathways — the direct microbial conversion of gas feedstocks, such as abundantly available CO_2, synthesis gas ($CO/CO_2/H_2$), and waste gases, such as H_2S and biogas methane, to fuels by engineered photosynthetic, methanotrophic, or chemoautotrophic microorganisms. The potential for fuel production by direct conversion has long been conceptually attractive, but major unsolved challenges continue to impede this effort. Microbial conversion of these gas substrates is ubiquitous in nature, but efforts to exploit species that might be useful for this purpose have found limited success. Efforts to engineer these organisms are hampered by a dearth of required genetic tools with which to do the engineering and a lack of basic understanding of the metabolic pathways involved and how they are regulated. Tools for a range of new hosts are described in the Biomanufacturing section of this document, which along with a systems-level understanding are needed to carry out the large-scale pathway engineering for the development of new production strains. In pursuing this strategy, Berkeley Lab leverages capabilities in the Biosciences Area, both existing and those being advanced in the Technology section, to obtain a systems-level understanding of gas-converting microbes. Berkeley Lab also leverages JBEI's significant experience with microbial fuel production from sugar substrates to transfer these pathways to microorganisms that generate fuels from gas feedstocks. The ABPDU is developing bioreactors and upstream and downstream processes internally and in partnership with JBEI, industry, and other DOE laboratories to scale fuel production from gas feedstocks. Synergistic JGI missions, such as genomics of biofuel feedstocks and discovery strategies for fungal/microbial enzymes, are valuable to this effort, as are JGI's major data resources (e.g., the Integrated Microbial Genomes and Microbiomes [IMG/M] system) that can be mined to achieve these objectives. Berkeley Lab's Biomanufacturing team provides genome editing and strain engineering strategies, along with a foundational retrosynthesis infrastructure, that are highly synergistic and critically important to the success of this strategy.

Artificial photosynthesis: Use bio-inspired reactions to create fuels directly from atmospheric CO_2 and sunlight

Berkeley Lab's Melvin Calvin mapped the route that carbon travels through a plant during photosynthesis — research that led to discovery of the "Calvin cycle" and the Nobel Prize in Chemistry in 1961. Once achieved and scaled up, artificial photosynthesis, a chemical process that replicates the natural process of photosynthesis, could be significantly more efficient than biological fuel production processes, and would not require arable land, agricultural feedstock, or substantial inputs of energy or water. Over the past 50 years, basic research has steadily increased an understanding of the subtle and complex mechanisms behind natural photosynthetic systems as well as in the use of photochemical methods that mimic key steps in the process: splitting water and reducing carbon dioxide. However, significant impediments, such as the inability to control chemical reactions on a nanometer scale and the limited understanding of photosynthesis on temporal and spatial scales, have prevented designing solar-energy-to-fuel conversion systems with the required efficiency, scalability, and sustainability to be economically viable.

In the past 20 years, nanotechnology — the making and manipulation of matter on nanometer scales — has advanced dramatically, and with it, the ability to control and optimize the intricate processes of artificial photosynthesis. New nanotechnologies allow Berkeley Lab researchers to work at the scale of atoms and molecules, to synthesize new catalysts to accelerate relevant chemical reactions, and to assemble molecular and nanostructured components into scalable hierarchical assemblies for the efficient conversion of carbon dioxide and water to solar fuels. Nanotechnology of inorganic materials, together with bioengineering technology that enables the manipulation of biological systems at the genomic level, make it possible to integrate living organisms with inorganic catalysis into nanoscale biohybrids, and take advantage of the complementary strengths of abiotic and biotic systems — a robustness of inorganic systems and the complexity and nuanced function of biological systems — in energy harvesting and chemical production.

Berkeley Lab has committed to this strategy, and is uniquely positioned to address it with a wide variety of assets. In addition to its historical and current scientific expertise, Berkeley Lab has state-of-the-art, unequaled, relevant facilities for studies of artificial photosynthesis. With its Molecular Foundry, a nanoscience facility, Berkeley Lab has become an epicenter for the synthesis of novel nanomaterials and advanced characterization, drawing scientists from around the world to its world-class instruments, materials, technical expertise, and training. Here again, the high-throughput DNA sequencing, synthesis, and analysis provided by the JGI strengthen the foundation of essential synthetic biology aspects of the biohybrid goals. In addition, Berkeley Lab's ALS provides state-of-the-art macromolecular crystallography facilities and a planned national cryo-EM facility, which are used to determine structural features of key proteins and complexes in these systems. Other proximal light sources, like SLAC National Accelerator Laboratory's Linac Coherent Light Source, and laboratory-based unique transient spectroscopies are being used to study natural photosynthesis with the goal of translating discoveries of natural photosynthetic processes into bio-inspired design principles for artificial photosynthesis. These facilities and technologies emerging from them are critical for monitoring photocatalytic processes in action, providing insights that rapidly lead to improved artificial photosystem designs. Additional enabling knowledge for artificial photosynthesis will be gained through research conducted at Berkeley Lab's National Center for Electron Microscopy, which provides cutting-edge instrumentation, techniques,

and expertise for advanced electron beam mesoscale characterization of materials at high spatial resolution, and the new envisioned bioimaging center, which will enable new insight into photosynthetic processes through multimodal imaging. The computing power of the National Energy Research Scientific Computing (NERSC) Center at Berkeley Lab plays an important role in guiding materials selection for optimal systems performance. At Berkeley Lab's Joint Center for Artificial Photosynthesis (JCAP), the DOE Energy Innovation Hub for Fuels from Sunlight (a partnership with the California Institute of Technology), a dedicated team of scientists and engineers is developing a solar powered technology at the testbed level for converting carbon dioxide into liquid transportation fuel.

2023 10-year goal achievement measured by:

Lignocellulosic biofuels

- Understand fundamental elements of plant biology that underlie biomass yields and adaptation to stress by focusing on key gaps in knowledge of primary plant physiology and in plant-microbe interaction relevant to plant resilience.
  - Elucidate how the secondary cell wall of biomass crops is synthesized.
  - Understand mechanisms of water conservation and drought tolerance in plants.
  - Advance the understanding of how nitrogen could occur in model plants.
  - Understand the molecular interactions between plants and mycorrhizal fungi and the roles of these interactions in nutrient acquisition and stress tolerance.

- Engineer biomass crops for reduced inputs (water and nitrogen/phosphate fertilizers), enhanced tolerance to stress, improved sugar yields, and facilitated production of useful compounds from lignin, including fuels
  - Engineer plants and microbes to stimulate nitrogen fixation by non-leguminous plants.
  - Engineer model plants and crops to grow with less water.
  - Engineer lignin in biomass crops to be an economically useful polymer.
  - Increase the C6 to C5 sugar ratio.

- Develop predictive models that will facilitate engineering of optimized secondary cell wall synthesis and saccharification by specific genetic manipulations.

- Develop deconstruction processes that enable efficient utilization of all plant biomass components (cellulose, hemicellulose, and lignin).

- Develop inexpensive methods to depolymerize plant polymers to intermediates for microbial conversion.
  - Develop a suite of enzymatic cocktails that depolymerize cellulose, hemicellulose, and lignin without product inhibition and inhibition from contaminants resulting from the deconstruction process.
  - Discover chemical catalysts that selectively depolymerize plant polymers to monomers that can be converted to fuels by fermentation.

- Understand the mechanisms of lignocellulose deconstruction by microbial communities.
  - Probe the variety of biomass degradation strategies employed by microbes in natural environments.
  - Establish a synthetic microbial community that deconstructs one lignocellulosic substrate.

- Understand synergies between fungi and bacteria in a mixed eukaryote-prokaryote microbial community.
- Define the pathways for deconstruction and metabolism of lignin in one synthetic or adapted microbial community.

• Engineer microorganisms tailored for consumption of deconstructed biomass (sugars and aromatic compounds) and production of drop-in biofuels at high yield.
- Develop metabolic pathways for production of hydrocarbons with fuel properties equivalent to those found in petroleum-based gasoline, diesel, and jet fuels, and engineer these pathways into diverse hosts.
- Use metabolic engineering to generate microorganisms that rapidly convert multiple substrates to desired end products simultaneously.
- Develop genetic tools, including biosensors, to enable multigene engineering in diverse hosts to rapidly engineer complex genetic traits such as fuel yield and stress tolerance.
- Engineer microbes to tolerate or detoxify biomass deconstruction inhibitors and fuel products, and other stresses introduced during bioreactor scale-up.
- Engineer microorganisms to produce fuels under anaerobic conditions.
- Engineer fuel-producing microorganisms whose communities are resistant to invasion by contaminating organisms and that are genetically constrained to growth in industrially-defined conditions.

• Develop and demonstrate bench- and pilot-scale processes for fuel production.
- Understand gene expression and metabolism at different scales of production to enable efficient and predictive scale-up of manufacturing processes.
- Identify, implement, and prototype new unit operations for more flexible and efficient production of precursor biomolecules and finished fuel derivatives.

• Develop predictive models to quantify the potential cost, net energy, greenhouse gas emission, and water impacts of research advancements in feedstock and biofuel production.
- Model the shifts in biomass yield, land requirements, and fuel yield resulting from new engineered feedstocks.
- Identify key drivers of feedstock and biorefinery costs, emissions, and resource requirements to guide research towards more sustainable pathways.
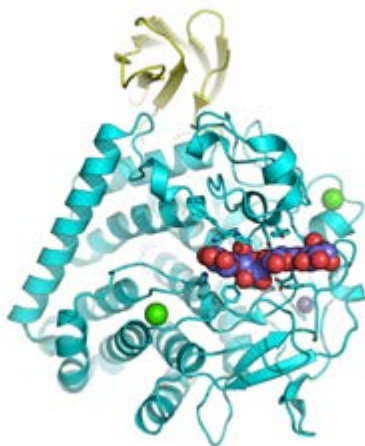
Alternative biofuels

- Use a systems-level understanding to identify key bottlenecks that will limit fuel production in the metabolism of photosynthetic, methanotrophic, and chemosynthetic microbes.
  - Identify key and structurally characterize proteins involved in the metabolic pathways essential for conversion of gas feedstocks.
  - Identify key regulatory elements that sense and respond to the presence of gas feedstocks in potential fuel production hosts.
  - Develop and validate genome-scale metabolic models of a photoautotroph, methanotroph, or chemoautotroph that show promise for conversion of gas feedstocks to fuel.

- Engineer photosynthetic, methanotrophic, and chemosynthetic microbes to produce fuels from gas feedstocks.
  - Develop a synthetic biology toolbox leveraging broader methods being developed in the Biomanufacturing goals, (transformation, promoters, genome integration of heterologous pathways) for at least three gas conversion hosts.
  - Produce hydrocarbon biofuels from gas feedstocks in a photoautotrophic, methanotrophic or chemoautotrophic host.
  - Improve the total carbon conversion in host microbe gas feedstock use by 10%.

- Scale fuel production from gas feedstocks through bioreactor and process development.
  - Demonstrate production of a hydrocarbon fuel from a gas feedstock at pilot scale.
  - Design bioreactors that maximize gas conversion to fuel by improving mass transfer.
  - Integrate product separation with gas conversion to improve fuel yield.

Artificial photosynthesis

- Develop scalable artificial photosynthetic systems based on bioinspired engineered and biohybrid approaches that produce energy dense liquid fuels from light energy and CO_2, well beyond the status quo.

- Couple predictive models with advanced engineering to improve artificial photosynthesis.

- Understand the fundamentals (e.g., excited state and charge transfer dynamics) that govern the multidimensionally (time and space) controlled chemistry in photosynthetic enzymes and artificial systems using time-resolved X-ray crystallography and X-ray spectroscopy at X-ray free electron lasers, and multidimensional X-ray spectroscopy at the synchrotron and X-ray free electron laser facilities, in combination with time resolved 2-D electronic and vibrational spectroscopies.

- Use new levels of understanding of photosynthesis to predict ways it can be improved in plants and microbes.

- Redeploy the photosynthetic apparatus in a previously non-photosynthetic host with superior coverage of the solar spectrum, improved conversion efficiency to fuel products, and durability.

- Develop scalable biohybrid systems that utilize synthesized light-harvesting components to sustainably shuttle electrons and protons to cellular organisms for efficient and selective production of carbon-based fuels.

2018 – Five-year milestones for energy research strategies

Lignocellulosic biofuels

- Cell-wall biosynthesis and assembly elucidated through identification of new genes, alleles, and metabolic pathways controlling cell-wall recalcitrance, sugar and lignin content, and fermentation inhibitors.
    - Secondary cell walls were engineered to have more C6 sugars and fewer C5 sugars.
    - The primary cell wall was engineered to have easily cleavable or altered lignin.
    - Biomass traits have been engineered from model plant systems to potential bioenergy crops.

- Tools developed to determine metabolite levels and metabolic bottlenecks in plants.

- Knowledge of drought tolerance in model crops advanced through identification of new genes, alleles, and metabolic pathways.

- New pretreatment methods developed that reduce cost and efficiently fractionate lignocellulose into targeted lignin and sugar output streams.

- Enzymes discovered and developed for optimal performance under pretreatment/ saccharification conditions (temperature, pressure, and the presence of ionic liquids).

- New enzymes and cofactors discovered and engineered for biomass deconstruction that are tolerant of pretreatment regimens.

- Efficient enzymes capable of depolymerizing lignin into aromatic hydrocarbons discovered to create a strong value stream for lignin.

- Models developed to predict how modifications to secondary cell-wall biosynthesis improve degradation of biomass.

- Predictive models to describe release of sugars from the plant secondary cell wall developed.

- Hydrocarbon biosynthetic pathways and associated transporters engineered into microbes to convert sugars to transportation fuels.

- Native hydrocarbon biosynthetic pathways in plants and microbes described.

- Predictive models to describe metabolic fluxes developed and used to predict bottlenecks in biosynthetic pathways in microorganisms.

- New bench- and pilot-scale unit operations and integrated processes developed and demonstrated for the production of a lignocellulosic biofuel.

Alternative biofuels

- Developed and implemented gas-fed bioreactor processes and associated systems and upstream and downstream unit operations for fuel production to enable demonstration at bench and pilot scale.

- Identified essential genes required for growth and CO_2 metabolism in a photoautotroph.

- Developed synthetic biology toolbox related to hydrocarbon biofuels production in a chemoautotrophic host.

- Converted synthesis gas to hydrocarbon biofuels under anaerobic conditions.

Artificial photosynthesis

- Developed an advanced mechanistic understanding of photosynthesis in plants and microbes.
    - Identified the determinants of efficient photosynthesis in plants, algae, and bacteria.
    - Defined the repair mechanisms of the photosynthetic apparatus.

- Explored options for reaching mA/cm2 currents between artificial systems and cellular organisms.

- Synthesized membranes capable of separating carbon-based fuels from oxygen.

- Demonstrated capability for photocatalyzing conversion of CO_2 to a carbon-based fuel beyond CO and formic acid.

- Designed the first prototype devices for testing components (catalysts, light harvesters, membranes, interfaces, etc.) as an integrated system.

- Demonstrated a photosystem for unassisted CO_2 reduction by H_2O under membrane separation on the nanoscale.

- Developed a mechanistic understanding of light-driven H_2O oxidation on a robust Earth abundant catalyst.

- Performed an analysis of components, materials and chemical inputs, and hardware designs to provide information on manufacturability, life-cycle costs, and reusability to ensure the system's scalability.

.

BIOSCIENCES FOR THE ENVIRONMENT

Background and Motivation

Globally, humans face a confluence of environmental and societal challenges to providing resources for a rapidly growing population while safeguarding vital environmental ecosystem processes. To address such 21st century challenges, bioscientists must drastically improve their mechanistic understanding of environmental organisms and their ecology across dynamic and changing environments. This approach will provide foundational insight into how microbes, plants, and other organisms interact to process nutrients and contaminants, enabling bioscientists to predict how soil processes and water quality change with anthropomorphic and natural inputs. Such insight will provide the foundation for developing urgently needed mechanistic approaches to managing vital nutrient cycles and harnessing beneficial plant-microbe-metazoan interactions, resulting in improved bioenergy and food crops with increased productivity, enhanced drought resistance, and lower inputs (e.g., energy, fertilizer, and pesticides).

The tremendous advance in DNA/genome sequencing now enables inexpensive and rapid sequencing of prokaryote and eukaryote genomes and communities from environmental samples. As a result, researchers continue to accumulate vast amounts of sequence information from all types of organisms and environments. Functional genomic technologies, including gene expression, epigenetic profiling, large-scale protein and metabolite measurements, and powerful imaging modalities, are providing unprecedented views into the activities encoded in organismal and community genomes. Together with sequencing and environmental characterization, these advanced technologies have the potential to provide the scientific foundation to accurately predict the functions of genes, organisms, and whole communities. This approach will transform ecosystem science and enable developing new theories and models

10-YEAR
Environment Goal

Understand the genetic and molecular mechanisms governing the activities and ecology of organisms and multispecies communities to predict responses and harness microbes and plants for energy and environmental solutions.

Environment Research Strategies to Achieve Goal

Predictive Understanding of Environmental Organisms. Discover and deeply characterize the genetic and molecular mechanisms of environmental organisms that drive and respond to environmental changes.

Molecular Ecosystems Biology-Based Solutions. Use biological and environmental characterization of natural and laboratory ecosystems to understand native ecosystem processes, predict responses, and harness plants and microbes for energy and environmental solutions.

to predict microbial community nutrient cycling and environmental responses, and harnessing organisms to promote beneficial outcomes. However, a predictive level of understanding of these ecosystems is still lacking. In fact about half of the genes in the genomes of most environmental microbes remain undefined with respect to their functions. Even less is understood about how to scale from the genes in a genome to the phenotypes and activities of an organism, and then to move from an individual organism and genomes to highly interdependent communities of organisms.

Deploying advanced sequencing, functional genomic, and imaging technologies to study microbe-plant-animal communities in realistic environments is the best way to learn about the function of their genomes. For example, a large fraction of the poorly annotated genes in environmental organisms are believed to function only under specific environmental conditions or in the context of other species with which they interact. Hence, discovering the function of such genes/genomes will require studying them in an ecological context. However, most studies to date have been performed either on isolated organisms or fully complex environmental communities, both of which have major challenges. The same is likely true for a large number of plant genes, which are presumably dedicated to selecting for and maintaining beneficial microbiomes, protecting from pathogens, response to abiotic and biotic stress, and enabling nutrient acquisition. Yet, because of the extreme complexity and undefined nature of natural ecosystems, determining the functions of specific genes/genomes and ecological interactions under field conditions is very difficult.

To address these challenges, Berkeley Lab is pioneering the development of laboratory ecosystems and advanced technologies for characterization and genetic, environmental, and biochemical manipulation of diverse environmental organisms within these environments. This approach provides unprecedented control of and visibility into environmental biological processes. The results obtained will feed directly into developing new theory, data scientific, and modeling tools to enable predictive approaches on biomolecular, organismal, and community functions, activities, and dynamics.

Central to these efforts is Biosciences' rapidly expanding ability to cultivate, characterize, and manipulate diverse organisms and communities within controlled environments. The model fabricated ecosystems (EcoFABs) are being developed based on field studies of native ecosystems, which provide controlled environments needed for genetic, synthetic biology, functional genomic, and imaging approaches to determine the role of specific organisms, genes, metabolites, and environmental factors mediating activities important for ecological attributes. The EcoFABS will enable both systems biology-based discoveries and reduction to causal mechanisms for developing predictive models and new ecological theories that can be tested and refined under laboratory and native ecosystem conditions. The ultimate goal is to enable the development of effective approaches for environmental cleanup, balancing nutrient cycles, and harnessing plants and beneficial microbiomes to build soil carbon and increase agricultural and ecosystem productivity.

Berkeley Lab has extensive expertise and unique capabilities for addressing grand challenges in environmental biology and ecology. For example, the DOE-funded Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) program has a long history of groundbreaking work in advancing fundamental scientific understanding of critical environmental organisms' and microbial communities' genomics that impact contaminated field sites and nutrient cycling. This program has developed state of the art functional genomics, high throughput genetics, systems biology, mass spectrometry, bioimaging, bioinformatics, computational modeling, and molecular environmental microbiology

approaches that greatly enhance Biosciences' ability to achieve the Environment strategies. The new Environmental Genomics and Systems Biology Division brings together experts in microbe, plant and metazoan biology, genomics, biotechnology, data science, and modeling to address these critical challenges in energy and environment.

Biosciences' Environment strategies are highly integrated with Berkeley Lab assets, such as the DOE JGI. The DOE JGI provides advanced sequencing, DNA synthesis, and computational analysis capabilities to the energy and environmental genomics communities. Examples of its activities include the sequencing of DOE mission-relevant genomes, the development and application of advanced "sequence-to-function" capabilities, the sequencing and decoding of the unexplored "dark matter" of microbial genomes, and the synthesis and functional characterization of large DNA constructs to better understand microbial and plant secondary metabolism and plant-microbe interactions.

The Environment strategies will also leverage the Biosciences-led BioFoundry, the recently established multinational lab effort that is aimed at democratizing and accelerating the engineering of biology for desired purposes. Of specific relevance is the BioFoundry's potential to provide synthetic biology tools that allow rapid characterization and manipulation of biological activities. For example, BioFoundry capabilities will enable the synthesis, assembly, and expression of putative secondary metabolite biosynthetic pathways to discover their effects on microbial and plant ecology.

High-resolution multiscale imaging used to examine microbe-microbe and plant-microbe interactions effectively leverages the cutting-edge imaging capabilities of the Advanced Light-Source (ALS). The supercomputing resources required to analyze and model complex systems biology and imaging data sets are available at NERSC. The Systems Biology Knowledgebase (KBase) provides critical data and computing resources enabling researchers to collaboratively generate, test, and share new hypotheses about gene and protein functions and to build and share predictive systems models of microbes and microbial communities, and their interactions with plants and other biotic and abiotic components of ecosystems.

Berkeley Lab's Earth and Environmental Sciences Area (EESA) is an important Berkeley Lab partner in this effort, with complementary expertise in environmental characterization, remote sensing, hydrology, geophysics, isotope chemistry, environmental microbiology, ecology, biogeochemistry, and multi-scale modeling. The EESA, in collaboration with Biosciences scientists working to achieve the Environment Goal, will play a key role in developing laboratory ecosystems that are deeply connected to and accurately recapitulate key aspects of native ecosystems. Biosciences' strategies of incorporating and integrating new knowledge of biological mechanisms obtained through world leading systems biology approaches are synergistic with EESA programs focused on developing multiscale mechanistic biotic-abiotic simulation capabilities. This Berkeley Lab cross-Area collaboration will enable powerful macrosystem-scale predictions of ecosystem function relevant to climate and energy-related challenges.

Together, these strategies and effective partnerships at Berkeley Lab will drive critical advances in understanding the activities and dynamics of interactions of complex microbial communities and microbe-plant associations. Ultimately, this will lead to designed interventions that improve environments and enable the discovery of novel environmental organisms and biomolecular mechanisms that can be harnessed for bioenergy and biomanufacturing.

Environment Research Strategies

- **Predictive understanding of environmental organisms.** Discover and deeply characterize the genetic and molecular mechanisms of environmental organisms that drive and respond to environmental changes.

- **Molecular ecosystems biology-based solutions.** Using biological and environmental characterization of natural and laboratory ecosystems to understand native ecosystem processes, predict responses, and harness plants and microbes for energy and environmental solutions.

Two strategies are the focus to achieve the Environment Goal by 2023: (1) foundational science to develop a predictive understanding of environmental organisms, and (2) developing molecular ecosystems biology-based solutions. Advancing and coupling these elements is necessary to develop new classes of environmental and energy solutions. Berkeley Lab has the potential for a unique contribution to understanding integrated environmental system behavior by linking organized research efforts and expertise in molecular microbiology, microbial ecology, and subsurface and terrestrial ecosystem science with global climate expertise. Biosciences researchers are contributing to and helping lead Berkeley Lab's Microbes-to-Biomes Initiative. As part of this initiative, researchers are developing deep understanding of microbial, plant, and metazoan genomics and biomolecular mechanisms using integrated systems biology approaches to study the dynamics, stabilities, activities, and interactions of organisms with their communities. To facilitate this work, Biosciences is pioneering the development of advanced functional genomic technologies and EcoFABs that together will enable discovery of causal mechanisms through manipulation of genetic, organismal, and abiotic system components. Since these model ecosystems are continually referenced to native ecosystems through field studies and refined until the models accurately represent key aspects of the native environments, they provide a powerful framework for determining the necessary and sufficient components of an ecological process of interest. They also provide test beds for examining important ecological questions, such as functional redundancy, selective and neutral forces, population dynamics, and adaptation to environmental changes.

Biosciences' efforts to understand natural ecosystems are multi-faceted and interconnected. The Area's researchers use measurement of *in situ* dynamics of natural and manipulated systems as a function of environmental constraints to identify the primary environmental drivers of community structure and activities. To discover the mechanisms that drive ecosystem dynamics, Biosciences' researchers integrate technologies for constructing model ecosystems with *in situ* monitoring and manipulation. This work includes spatially and temporally defined sampling coupled to systems biology-based analysis (e.g., stable isotope tracing of metabolite, protein, RNA, and DNA molecules, as well as imaging approaches) of both microbes and their hosts. Controlled manipulations within these systems will allow the discovery of causal mechanisms that predict dynamic responses of organisms and their reciprocal impact on their environment, thus enabling the development of effective environmental solutions (e.g., microbiomes that clean contaminated water and growth promoting microbiomes for bioenergy crops). The mechanisms discovered are captured in predictive models to inform the development of effective policies and harness microbial communities and plants for resilient energy, agricultural, and environmental properties for sound ecological solutions.

Predictive understanding of environmental organisms. Discover and deeply characterize the genetic and molecular mechanisms of environmental organisms that drive and respond to environmental changes.

Microbes and plants in terrestrial environments carry out a wide range of essential biogeochemical processes. In natural ecosystems, microorganisms live in complex communities, interacting with plants, animals, and other microbes. Advances in DNA sequencing now enable rapid and inexpensive interrogation of genomes from both isolated organisms and mixed environmental communities (metagenomics). As a result, Biosciences Area researchers are accumulating vast amounts of sequence information from all types of organisms and environments. The DOE JGI, a user facility within the Area, plays a global leadership role in collecting and analyzing such data sets from organisms and environmental samples that are relevant to the DOE mission. These sequence data, in principle, contain the information necessary to make accurate predictions of organismal metabolism, community interactions, and fitness under controlled conditions. However, vital information on functional attributes of these sequences, and computational models and resources necessary to make accurate predictions are still lacking. For example, approximately half of the genes in any given bacterial genome are of unknown function, and very little is known about the genetics and ecology of viruses, which are numerically the most dominant microbes on the planet. Higher organisms, including plants, fungi, and soil metazoans, also play critical roles in global nutrient cycles and other environmental processes, yet their genomics remain poorly understood. This gap between sequence and function greatly diminishes the utility of sequence information and has led to correlative research (e.g., genes correlated with different states or environments). Bridging the sequence to function gap and generating a mechanistic, molecular understanding of organismal activities is one of the grand challenges in biology in the post-genome era.

Since function is intimately connected with environmental context, simple pure culture laboratory experiments are insufficient for discovering the functions of many genes. For example, a microbial pathway for making a secondary metabolite that is essential for mediating plant interactions only functions in the context of a plant. Thus, Biosciences' functional genomics approach includes characterization of gene functions across environmental conditions and model ecosystems. This approach, coupled with technological innovations in imaging, DNA sequencing, functional genomics, bioinformatics, and systems biology, developed at Berkeley Lab and elsewhere, will allow us to study gene functions in the context of changing environmental conditions.

Similar to the "microbial dark matter" of uncultivated, uncharacterized organisms, a vast "dark biochemistry" in nature exists encoded in the DNA of these uncultivated environmental organisms that must play vital ecological roles yet to be discovered. To address this challenge, Biosciences' researchers use environmental metagenomics integrated with bioinformatics, DNA synthesis, metabolomics, and cheminformatics to allow expression of biosynthetic pathways from uncultured environmental microbes in host organisms to identify metabolites and examine their biochemical ecology in laboratory ecosystems. This thrust is closely aligned with the strategic directions of the DOE JGI, which is actively expanding its portfolio of capabilities to include "sequence-to-function" genome-based technologies and approaches with the goal to make them available to its user communities.

More complete pictures of microbial genomics within realistic environmental contexts will provide important new insights into how microbes adapt and evolve, how dynamic such populations are, and the nature of interactions between the microbes and their physical/chemical environment and with plant and metazoan hosts. This functional genomic and physiological understanding will greatly enhance genomics-based modeling approaches and effective manipulation of organisms and environmental systems for environmental or energy benefits.

Molecular ecosystems biology-based solutions. Using biological and environmental characterization of natural and laboratory ecosystems to understand native ecosystem processes, predict responses, and harness plants and microbes for energy and environmental solutions.

Climate change and anthropogenic inputs into the environment are affecting carbon, nitrogen, and other nutrient cycles in soils. Microbial communities play critical roles in mediating these processes, as well as many other environmental effects of energy production. Improved understanding of the molecular mechanisms mediating microbial, metazoan, and plant ecology is urgently needed to accurately predict native ecosystem processes and harness plant and microbial ecosystems for environmental cleanup and sustainable production of biofuels and other crops.

To address these critical challenges, Biosciences is developing and using model ecosystems connected to field studies and advanced bioinformatics resources to understand mechanistic aspects affecting nutrient cycling and plant-microbiome interactions that enhance plant growth and abiotic stress tolerance. This, combined with advanced genetic and systems biology approaches, will provide new insights into microbial, metazoan, and plant genomics and enable elucidation of causal mechanisms that determine the resilience and change of natural ecosystems in response to climate and anthropogenic inputs. This mechanistic understanding of microbial, soil metazoan, and plant ecosystems will enable the development of new technologies, including defined microbial communities for environmental cleanup and carbon storage, as well as new plant cultivars coupled with specific microbiomes that increase productivity and efficiency (energy, nutrient, and water use efficiency) to enable bioenergy production on marginal land.

This aim is important since nearly all arable land is already under cultivation. Farmland is also lost due to urbanization and soil degradation, making developing approaches for cultivation of bioenergy crops on damaged soils critical. Typically, these degraded soils have very low organic carbon content, which plays a critical role in supporting beneficial microbiomes, increasing water infiltration, preventing erosion, and retaining water and nutrients. Building soil carbon content through bioenergy crop cultivation has tremendous implications for increasing the productivity of damaged soils while decreasing atmospheric carbon. Soils are a major global carbon reservoir and understanding soil carbon cycling requires studying plants, which are the major carbon input, as well as the activities of microbes and soil fauna that transform these inputs, and the abiotic soil factors that control plant and microbial biochemical activities.

Unfortunately, processes governing soil carbon creation and retention are poorly understood. Predictive power regarding how environmental conditions impact soil carbon is lacking, much less how crops and environmental organisms to build soil carbon might be harnessed. Thus, there is great urgency to develop the necessary foundational

understanding of soil carbon and other nutrient cycles to enable accurate predictions, develop beneficial practices, and inform effective policies and land management practices. Biosciences will help address this challenge by extending systems biology approaches into ecosystems to identify causal principles governing carbon cycling and other ecosystem services.

These efforts will be synergistic with the DOE JGI and KBase. The DOE JGI has started to build user communities around topics such as the role of secondary metabolites in plant-microbe interactions and, along with this, substantial expertise in relevant systems and technologies that will cross-fertilize efforts in the divisions. The DOE JGI has also played a global leadership role in producing reference genomes, as well as functional genomic data sets (transcriptomes, epigenomic data, and population-scale data) for many of the plants that are of interest to the DOE. This substantial expertise and firsthand access to massive data resources will facilitate and significantly accelerate research in this area by the Biosciences Divisions. KBase is working to accelerate users' ability to turn complex data into predictions of molecular, organismal, and community function. KBase will be a vital tool to effectively transfer knowledge about specific biomolecules, organisms, and ecosystems to related systems enabling exploration of new biology and development of predictive models. The resulting molecular ecosystems biology-based predictive understanding will enable Berkeley Lab to help develop more accurate models and biology-based solutions for environmental cleanup, soil carbon management, and much needed low-input agriculture.

2023 10-year Goal achievement measured by:

The Environment effort strives to gain a deep scientific understanding of the major drivers and consequences of environmental change arising from both natural variability and human activities and to develop new environmental solutions that consider integrative system behavior — some of which span both Environment strategies. The 10-year Goal will be met through investigations carried out using relevant model laboratory ecosystems integrated with field study sites and synthesis of resulting data sets.

Predictive understanding of environmental organisms

- Use model ecosystems to discover the functions of genes that mediate microbial interactions.

- Discover the molecular mechanisms by which microbial communities in groundwater and sediment are altered by anthropogenic contaminants.

- Enable data-driven prediction of gene functions for thousands of microbes and metagenomes per year and tens to hundreds of plants.

- Develop a computational infrastructure that transfers knowledge about specific biomolecules, organisms, and ecosystems to related systems, accelerating exploration of new biology.

- Characterize the metabolism, stress responses, and interactions of diverse environmental microbial isolates from relevant field sites using a multi-omics approach including sequencing, mutant phenotyping, and genome engineering.

- Develop a genomics-driven computational infrastructure to accurately and rapidly predict microbial metabolism, gene regulation, and stress response for microorganisms in at least five key environments: contaminated sediment, wetlands, deserts, agroecosystems, and grasslands.

- Integrate high-throughput genomic, imaging, epigenomic, and metabolomics approaches and computational technologies to discover the roles of novel genes, proteins, regulatory sequences, and metabolites in plant and metazoan responses to environmental change.

- Discover the molecular mechanisms by which plant growth-promoting microorganisms impact fitness of model grasses through multi-omics measurements in controlled laboratory and natural ecosystems.

Molecular ecosystems biology-based solutions

- Identify key natural *in situ* environmental processes to recapitulate and validate laboratory model ecosystem.

- Develop, test and refine model predictions of native community *in situ* energy and material dynamics to mechanistically account for the material and energy flow amongst ecosystem participants and extract principles of resilience, optimal nutrient and energy utilization.

- Identify the key microbial functions and organisms that stabilize microbial communities and then design interventions in model ecosystems that decrease energy and nutrient inputs necessary to achieve key ecosystem service goals such as carbon sequestration and crop productivity while minimizing soil degradation and run-off of pollutants and toxins.

- Integrate detailed measurement of organisms and environments within the model ecosystems to accurately predict nutrient cycling and or biotic interactions in native ecosystems.

- Demonstrate ability to harness the metabolic capabilities of microorganisms by using a microbial consortia to increase the yield of an important crop under sub-optimal conditions for example, water stress, growth on degraded lands, or low-nutrient inputs.

2018 – Five-year milestones for environment research strategies

Near-term efforts largely consist of advancing the functional genomics of environmental and model organisms and studies on controlled and reproducible model ecosystems that recapitulate important aspects of native ecosystems to link genomics to ecosystem processes. Emphasis is placed on studies that improve understanding of nutrient cycles and biochemical ecology under dynamic environments to harness microbiomes to improve plant productivity and efficiencies (e.g., water and nutrients). The Area will also obtain new insights into the flow of matter and energy through ecosystems to understand how these processes are mediated by specific organisms to predict critical ecosystem features, such as resiliency or productivity.

Predictive understanding of environmental organisms

- Used model ecosystems and bacterial mutant fitness profiling to discover new genes mediating microbial interactions under environmental constraints.

- Developed and applied new functional genomics technologies for discovery and validation of gene and noncoding regulatory DNA functions that impact the fitness of plants and microorganisms under multiple environmental conditions.

- Used exometabolomic profiling of groundwater and soil bacteria to model how soil metabolites shape microbial community structure.

- Discovered the molecular basis of plant-growth promotion for three bacteria, each from a different phylum, to determine if these mechanisms are conserved.

- Used plant mutants with defined microbiomes to discover plant genes that select for beneficial microbiomes.

- Identified over 100 metabolites exuded by plant roots and characterized how these metabolites are used by isolate soil microbes and how they impact microbial community structure.

- Developed technologies for genome editing, mating, and selection to gain genetic mastery of model microorganisms, plants, and metazoans and applied these strategies to understanding key biological processes in the environment.

Molecular ecosystems biology-based solutions

- Constructed model desert ecosystems that recapitulate key aspects of native ecosystems for controlled analysis of microbial metabolic processes.

- Used synthetic biology tools to construct reporter microbes and used these with advanced imaging technologies to determine where and when metabolites are being produced within a microbial community.

- Elucidated the metabolic strategies of key environmental microbial community members and the functions of specific genes that are important for community metabolism through a combination of systems biology, physiological analysis, functional genomics, and high-throughput microbial isolation.

- Used manipulation of environmental variables and spatially defined sequencing to determine the environmental controls on spatial distribution of microbial communities in soil environments.

- Applied understanding of resource competition in a model ecosystem to accurately predict soil metabolite composition and community structure.

- Constructed model agroecosystems that recapitulate key aspects for controlled analysis of the plant-soil-microbiome processes.

- Demonstrated the ability to alter microbiome structure and soil carbon cycling through targeted modification of plant genes.

- Discovered at least one novel microbial metabolite beneficial to the plant host and confirmed through introduction of the biosynthetic pathway into a model plant-associated microbe.

- Developed a "data ecosystem" computational infrastructure that enabled integration of microbial and multi-cellular systems biology modeling for interactions in microbiomes, especially in native and model ecosystems.
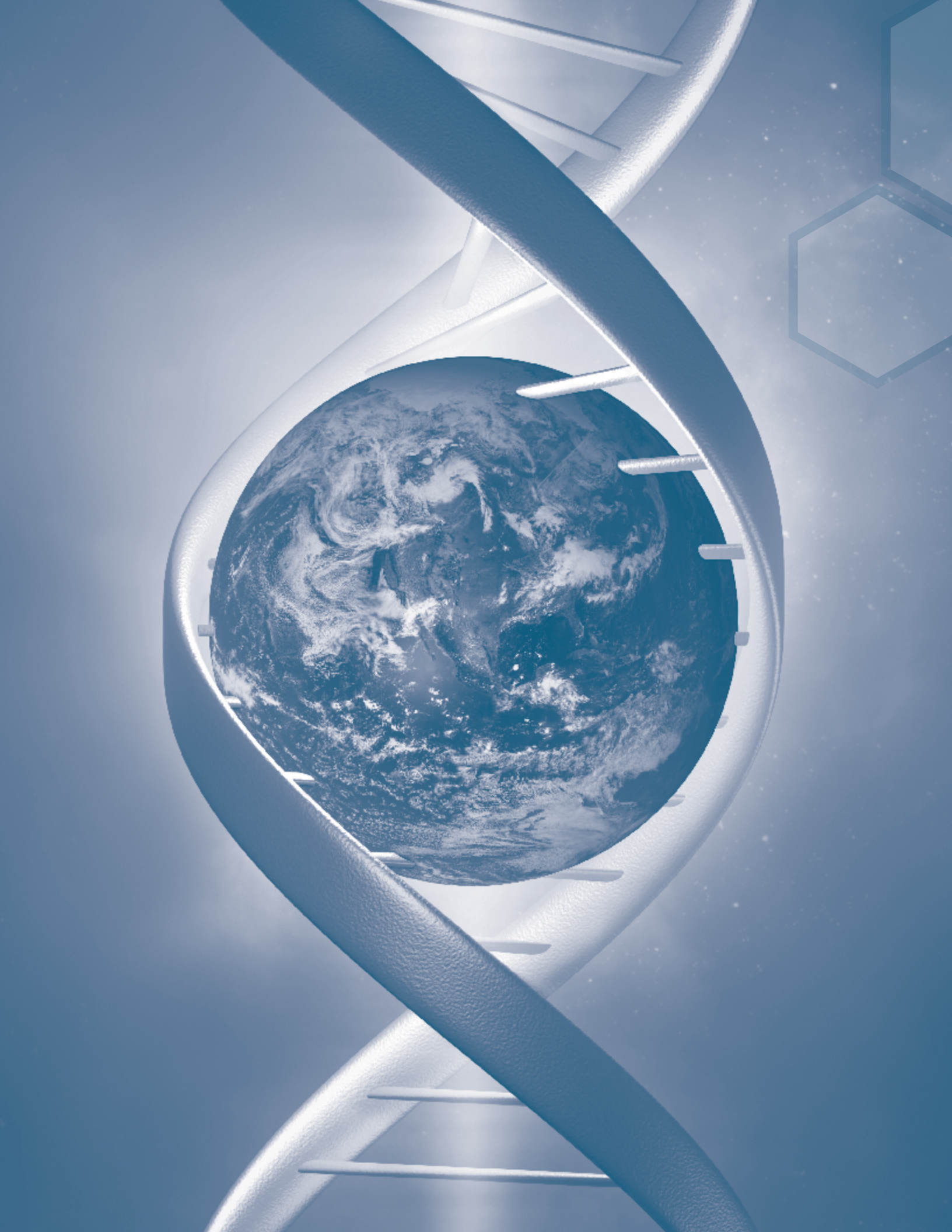
BIOSCIENCES FOR HEALTH

Background and Motivation

Organisms at all evolutionary levels have complex responses to natural and anthropogenic changes or challenges in their environment. External factors that can influence biological health include diet, temperature, climate, water and air quality, chemicals, radiation, nanomaterials, and byproducts from biomanufacturing and energy production. Most chronic diseases in higher organisms — including cancer and cardiovascular and neurodegenerative diseases — are caused by adverse gene-environment interactions, with the environmental component playing a major role. Environmental factors induce effects by altering molecules, cells, and physiological processes inside organisms. For example, in humans, environmental exposures such as radiation, physical trauma, and certain drugs damage brain tissue, leading to behavioral changes and memory loss, anxiety, and neurological diseases. Environmentally-induced defects in neural function impact basic behaviors (finding food, mating behavior, aggression, etc.) that transcend the individual and affect the long term survival of species and entire ecological systems.

A lack of understanding of the biological and health responses to environmental challenges also has a large impact on the U.S. economy. For example, energy production impacts air and water quality by generating airborne contaminants that increase rates of asthma and cancer, and broadly reduce human resiliency. In the United States, an additional $886B would be spent per year if the price of energy reflected the hidden costs associated with lost work time and early mortality due to energy-induced environmental challenges. Improvements in understanding the health consequences of energy production lead to cost-effective regulation: the Clean Air Act generates an estimated net $83B per year for the U.S. economy. Other environmental challenges, such as intentional or unintentional deployment of chemical toxicants, have similar "hidden" economic costs.

10-YEAR Health Goal

Develop and apply a predictive, multiscale, and integrative understanding of how individual variation affects responses to environmental challenges, to improve human and biome health and drive responsible economic growth.

Health Strategies to Achieve Goal

Biological responses to environmental challenges. Develop and deploy model systems to understand how individual genetic, epigenetic, and microbiome variation affect molecular, cellular, and organismal responses to environmental challenges, and to identify risk factors for somatic and neurological diseases.

Impact of environmental challenges on human health. Develop the knowledge and infrastructure to elucidate the immediate and long-term consequences of environmental challenges on health in diverse human populations, and to enable innovative solutions through responsible economic growth.

Berkeley Lab envisions the emergence of a concerted national effort focused on the protection of human health through advances in environmental stewardship. It is therefore of central importance to understand the health consequences of energy production and other environmental challenges, and to develop new remediation strategies and guide technologies toward safe and sustainable solutions. To realize this vision, advances are required in exposure biology, neurology, and the science of individual susceptibility and precision ecotoxicology. One of the greatest scientific challenges is balancing the potential risks associated with innovation and sustainable economic growth with maintaining the fitness of humans and the earth's biomes.

It is well known that genetic background and life history (e.g., gender, age, health, and diet) can modulate tissue and organismal responses to environmental exposures. Scientists have recently come to appreciate how the trillions of microbes present in humans and other species can greatly impact organismal and ecological health. Microbes produce enzymes, vitamins, and anti-microbial substances that contribute to individual susceptibility to environmental toxicants. For example, microbial metabolites produced in the human gut or soil metazoans can be much more or less toxic than their parent compounds. The internal chemical context of organisms continually fluctuates during life, due to changes in external and internal sources, so social and environmental factors also must be considered when evaluating impacts on fitness and disease.

Berkeley Lab aims to address scientific and societal challenges in a comprehensive manner. A major challenge for 21st century biology is to develop a deeper understanding of the types of responses to environmental challenges at many interrelated biological levels, including molecules, cells, cellular communities, tissues, and organisms, and the multitude of timescales associated with dynamics across these different scales. The ability to assess the genetic contributions to health and disease; the development of technologies for quantifying technologies to analyze the genome; transcriptome, metabolome, fluxome, and proteome ("omics"); and computational methods for integrating diverse and large data sets have blossomed in recent years. However, quantitative assessment of the extent of organismal exposure to environmental challenges, and its relation to fitness and disease, have lagged due to technical limitations in assessment and monitoring, and a lack of comprehensive and accurate data. Thus, it is essential to develop an integrative, quantitative, and predictive understanding of the biological responses to environmental challenges, and how they impact the fitness of humans and ecological biomes. Through the development of biologically-informed solutions to pressing societal problems, this knowledge will lead to reduction in risks associated with deploying new cutting edge products, and the introduction of new interventions aimed to maintain human and environmental fitness.

Human health is tightly coupled with the health of the biome, especially the plant and microbial communities that perform a variety of essential ecosystem services, including nutrient cycling and chemical transformations, and provide clean air and water and an abundant food supply. Disruption of these services by environmental challenges (e.g., climate change) will have profound impacts on human health and the global economy. Further, an improved understanding of the potential risks and benefits of chemical influences on individuals and ecosystems will enable new approaches to streamline product design in biomanufacturing, agriculture, and pharmacology. Screening procedures that broadly characterize the biological activities and potential risks of compounds at low cost and high-throughput have the potential to identify new bioactive compounds, including herbicides, antimicrobials, and prebiotics.

Unraveling the complexities of biological responses to the many internal and external environmental challenges ahead — and their impact on the health of humans and the biome — requires technical and integrative advances allowing a deep, multiscale, and integrated knowledge of mechanistic and phenotypic responses to these factors. For example, determining the roles of environmental factors in cancer initiation and progression and neurological diseases requires understanding, predicting, and mediating complex reciprocal interactions among multiple levels of genetic and physiological functions, in both host organisms and their microbiomes. These same skills and disciplines are required to monitor and analyze the health of humans and the biome. Whether the subject is the human body, microbial communities, or a critical insect population, what is needed is a quantitative understanding of the short- and long-term responses to environmental challenges encountered by hosts and their resident microbiomes, and their combined effects on organismal fitness, behavior, and long-term health.

Berkeley Lab has a long history of successfully integrating multiscale, cross-disciplinary approaches to address challenges of similar size and complexity. The Biosciences Area aims to address bioeconomy-related challenges in a comprehensive manner, by leveraging advanced facilities and equipment, a culture of interdisciplinary team science, and a historic and deep level of biological and technical expertise. Relevant disciplines include structural biology; biochemistry; cell and organismal biology; microbial communities; genetics, genomics, and epigenetics; metabolomics and proteomics; multiscale imaging and physiological measurements; and data integration through advanced computational analyses. Berkeley Lab has extensive experience with integrating data from multiple levels of biological function to understand the impact of environmental challenges, such as radiation, climate change, toxicants, and third-hand smoke in model systems, and the effects of exposures on the etiology of diseases such as cancer and neurological disorders. Importantly, the Biosciences Health efforts leverage Berkeley Lab's extensive technical capabilities and national user facilities, including the ALS, the DOE JGI, and NERSC. In addition, partner faculty at the University of California (UC) Berkeley, UC San Francisco, UC Davis, and other Bay Area institutions bring additional expertise and experience in molecular epidemiological studies of human health and disease, which complement the Berkeley Lab "Big Science Missions" and capabilities in elucidating basic biological mechanisms.

Together, Berkeley Lab and partner Bay Area institutions are poised to integrate basic, mechanistic information about the responses of biological systems to environmental challenges with accurate measurement and understanding of the extent and magnitude of exposures and their impact on health and disease. Achieving this mechanistic understanding will address a pressing national need to accurately predict the impact of environmental challenges on biological systems. This knowledge will then be leveraged for risk management, by eliminating or modifying detrimental compounds from the environment, identifying individual humans or organisms that are sensitive to exposure, and developing innovative biosolutions. Berkeley Lab's efforts will result in development of safe, sustainable energy and materials and reduced exposures to harmful environmental factors, improved public health and personalized medical interventions, and increased understanding of the impact on organisms that are essential to biome fitness. Given sustained effort, Berkeley Lab will be a world leader in generating scientific discoveries that have long-term, high-value impacts on improving the fitness of humans and the biome, and the quality of life for humans, resulting in significant positive impacts on the overall economy.

Health Research Strategies

- **Biological responses to environmental challenges.** Develop and deploy model systems to understand how individual genetic, epigenetic, and microbiome variation affects molecular, cellular, and organismal responses to environmental challenges, and to identify risk factors for somatic and neurological diseases.

- **Impact of environmental challenges on human biology.** Develop the knowledge and infrastructure to elucidate the immediate and long-term consequences of environmental challenges on health in diverse human populations, and to enable innovative solutions through responsible economic growth.

National laboratories excel at integrating the multiscale, complex data required to address these important challenges. To achieve the Health Goal by 2023, Berkeley Lab will focus on two strategies: (1) build a cross-disciplinary platform that provides a comprehensive, integrated understanding of positive and negative responses to environmental challenges, using complementary metazoan and microbial model systems that cover levels of function from molecules to organisms, and (2) assess the impact of environmental challenges on human health and disease. These focus areas will integrate multi-scale data using bioimaging, genomics, proteomics, metabolomics, and computational technologies to rapidly phenotype and quantitatively interrogate complex, dynamic biological systems. Technological and intellectual advances at Berkeley Lab and elsewhere will accelerate the success of the Health strategies.
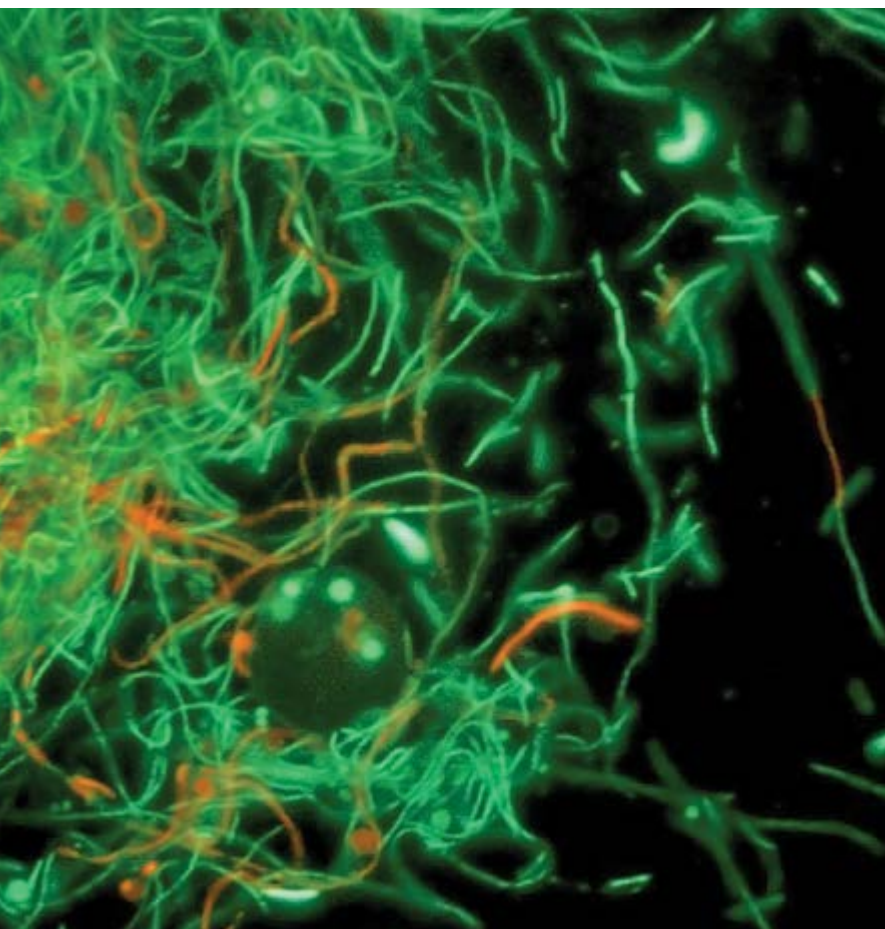
Berkeley Lab's work on understanding human and biome health responses to environmental perturbations will be performed in close integration with the other Biosciences focus areas (Energy, Environment, Biomanufacturing, and Technologies for Biosciences). For example, development of the bioimaging, computational, and functional genomics tools described in "Technologies for Biosciences" will be crucial to the success of this program. Through these integrated efforts, mechanistic understandings will be developed that will ultimately inform prevention, diagnosis, and treatment. The Health component will also play key roles in development of cross-cutting technologies for rapid assessment and imaging of phenotypes that will be utilized across the Biosciences research portfolio.

Together, these strategies address short- and long-term national needs by generating a comprehensive understanding of the impact of environmental challenges on the fitness of biological systems, and developing innovative bio-solutions to

assess hazards and mitigate health problems related to environmental exposures. These strategies are designed to be achievable in a 10-year time span. The technical and intellectual strengths of the Biosciences Area and of other Berkeley Lab scientific areas will be brought together to achieve the advances needed to enable productive application of renewable biological resources that protect and improve the economy. As a result, novel methods will be developed that connect nanoscale to organismal behaviors by collecting multiscale data simultaneously from the same animals or tissues. This will enable a predictive, multiscale understanding of responses to environmental challenges, which will promote responsible and sustainable economic growth, and ensure the fitness of humans and the biome.

Biological responses to environmental challenges: Develop and deploy model systems to understand how biological diversity impacts molecular, cellular, and organismal responses to environmental challenges, and to identify risk factors for somatic and neurological diseases.

This strategy aims to develop a mechanistic understanding of how environmental challenges, specifically chemicals, radiation, nanomaterials, diet, energy production and use, impact the fitness of metazoan organisms. An important first step is defining how exposures affect biological systems in specific model organisms, from macromolecular complexes to biological outcomes. Using diverse model systems will enable us to leverage evolutionary biology, and map multiscale models of chemical exposures onto the phylogenetic tree. A critical advance will be the development of quantitative models linking responses measured in one or a few species or systems to others. A central aim is to predict the effects of these exposures and additional challenges on organism fitness, and to develop solutions that prevent or mediate negative impacts. Technologies for measuring the types, levels, and distributions of biomolecules, and cellular and organismal phenotypes, have advanced considerably. However, the robust application of these technologies in an integrative, comprehensive manner that reveals impacts on interrelated levels of biological mechanisms and functions has lagged considerably.

Due to the enormous complexity and multifaceted nature of biological systems, a holistic approach is needed to achieve the 10-year Goal. Environmental factors can affect different types of molecules, cells, tissues, and organisms independently, but how biological functions are affected by interactions among these components must be incorporated to achieve a true and effective understanding of impact on the fitness of an organism. For example, understanding how toxicants impact fitness requires integrating measurements of changes in the levels and distributions of diverse biomolecules (DNA, RNA, metabolites and their fluxes, and proteins) in different cells and tissues, as well as information about alterations in the composition and activities of resident microbiome communities, combined with quantification of the impact on host cells, tissues, and organismal phenotypes and functions.

By identifying key biological mechanisms in model organisms, instead of merely cataloging relationships between input and output, responses of biological organisms in hypothetical situations can be predicted based upon mechanistic knowledge. For example, can the presence of specific microbiome communities protect against specific toxicants, and is this modulated by the genomic makeup of the individual organisms and the host? To approach this, manipulatable, isogenic experimental systems that allow control of genotypes and environments are required to enable precise measurements of

responses in a way that is impossible with studies of humans. The results will provide a scientific and mechanistic foundation that will inform microbial and human fitness studies described in the second strategy.

This research strategy utilizes three well-established model systems — *Drosophila melanogaster* (fruit flies), *Mus musculus* (mouse) and *Rattus rattus* (rat) — whose strengths have been used to generate deep biological insights over the last century. Most importantly, these model systems provide the ability to leverage sophisticated genomic, genetic, epigenetic, cell biological, and developmental tools for multimodal manipulation and measurements of biological responses to environmental challenges, under conditions where the environment, genetic variation, and microbiome composition can be controlled. For example, the rat has been foundational for understanding changes in brain function and behavior after environmental challenges. It is also a favorite model for testing innovative biosensing and neuroimaging technologies for measuring brain function, neural networks and communication, and behavior at multiple scales. In these tractable systems, Berkeley Lab aims to quantify responses at the molecular level by developing and applying advanced "omics" tools and expertise, focusing on changes to the types, structures, and levels of DNA, RNA, proteins, and metabolites and their fluxes, in both the host and resident gut microbiomes.

The manipulability of these models also will allow us to correlate molecular changes with cell, tissue, and organismal phenotypes, as well as individual variations in responses, ranging from changes in molecular machines and pathways to cellular structures, tissue phenotypes, and animal behaviors. To provide a comprehensive understanding of the health effects of environmental challenges, the advanced computational infrastructure at Berkeley Lab will be employed to manage, analyze, visualize, and integrate the "big data" generated by these studies. Lessons learned from these controllable model systems will also provide mechanistic insights that will inform approaches to elucidating environmental impacts on human biology and health.

Impact of environmental challenges on human biology: Develop the knowledge and infrastructure to elucidate the immediate and long-term consequences of environmental challenges on health in diverse human populations, and to enable innovative solutions through responsible economic growth.

National decisions about investments in sustainable solutions for energy and climate change are interwoven with their impacts on human health and the changing demographics of the U.S. population. Retiring "baby boomers" and the increasing cost of medical diagnosis and treatment are problems the United States needs to tackle at a different level than is currently being approached by medical research. One approach to solving these problems is the new federal Precision Medicine Initiative, the goal of which is to base diagnostic and treatment decisions on individual variability in genes, environment, and lifestyle. Genetic background is known to determine whether or not individuals are susceptible to the bioactive effects of environmental challenges and pharmaceuticals. Properties such as gender, age, diet, and history of environmental exposure also have significant impact on health. Realistic predictions of exposure outcomes in human populations requires the development of higher-order human cell systems that incorporate significant genetic diversity and enable manipulation and measurement of responses in human tissues, organs, and organ systems.

Berkeley Lab plans to elucidate the impact of environmental factors on human fitness and disease by focusing on four staged approaches. First, genetically diverse model organisms will be used to identify molecular-, cellular-, tissue- and organismal-level responses to environmental challenges that are comparable to responses in humans. These experimental tools will be used to identify harmful environmental factors and susceptible individuals, to create countermeasures against environmental exposures such as radiation and chemicals, and to develop environmental sentinels that forewarn of imminent health risks, such as those stemming from deployment or unintentional releases of commercial chemicals into the environment. Second, multi-tissue bodies-on-a-chip systems will be engineered to enable their democratization and wider use for exposure research. These will take advantage of primary human cell culture technologies developed at Berkeley Lab, which provide a unique resource of genetically diverse normal human cell strains, in perpetuity. These will be developed for identifying both deleterious and beneficial effects of environmental exposures, and for defining and measuring the bioactive effects of commercially and pharmaceutically important molecules to human systems. Third, impacts on critical "omes" will be measured in model organisms, body-on-chip systems, easily accessible human tissues such as blood, bodily fluids, and tissues from biopsy to compare resultant profiles of diseased and healthy individuals. Such an approach will allow identification of discriminating molecules, or emergent properties of tissues, which ultimately identify causal exposures and sources. Fourth, real-time physiological functioning will be measured for specific tissues and organs (e.g., brains) in vivo to understand the immediate, intermediate, and long-term effects of exposures and the proposed solutions in live organisms. Exposures to some agents could have beneficial or adaptive responses for some individuals, and these technologies and analyses will enable the identification of individuals who may benefit or be harmed by specific environmental challenges, such as energy products and bioactive molecules.

Desired outputs of this strategy are (1) an understanding of the molecular mechanisms underlying variation in responses to environmental toxicants, (2) identification of variations that predict individual risk for somatic and neurological diseases, (3) candidate strategies for detecting susceptible and resistant individuals to environmental exposures, (4) candidate biomarkers that can predict hazards from an exposure or reveal a person's history of exposure, and (5) sound scientific data that government agencies can use to establish appropriate policies, to mediate the effects of, for example, chemical, microbial, or radiation exposures.

2023 10-year goal achievement measured by:

Biological responses to environmental challenges

- Identify the keystone components that mediate the impact of five environmental challenges (chemicals, radiation, nanomaterials, diet, and energy production and use) on molecules, cells, microbial communities, tissues, and organisms in tractable model systems, as measured by:
  - Phenotypic responses to environmental challenges in three model systems (flies, mice and rats) exhibited by cells, microbial communities, tissues, and organisms).
  - The 4-D dynamic responses of molecules (DNA, RNA, protein, and metabolites and their extracellular and tissue fluxes) and phenotypes using advanced imaging, genomics, phenomics, and computational approaches.

- Determine how the effects of five major environmental challenges are modulated by genetic, physiological, and epigenetic variation in model systems, as measured by:
  - Genetic mapping to identify DNA sequence differences that affect individual responses to environmental challenges.
  - Epigenetic variations that influence responses to environmental stressors.
  - Cellular and tissue damage responses that vary by physiological status (gender, age, diet, etc.).
  - The impact on brain electrophysiology, the blood-brain barrier, and behavior.

- Understand if and how the effects of five major environmental stressors are transmitted through cell divisions and trans-generationally to progeny, as measured by:
  - Environmental challenges applied acutely to a cell or organism that result in trans-generational impacts on biological functions.
  - Defining genetic and epigenetic mechanisms responsible for trans-generational inheritance of environmental stress.
  - Revealing sensitive windows of germ cell and embryonic development responsible for the transmission of heritable changes to offspring.

- Elucidate the role of prototypic community interactions within biological systems, and how they are reciprocally affected by five environmental challenges, as measured by:
  - Model host/microbial community composition and functional responses to environmental challenges that identify conserved and species-specific mechanisms.
  - Identifying microbial communities that impact fitness and define their interactions with each other and hosts.
  - Identifying changes in brain function and behavior due to reciprocal flux of gut-microbial and central nervous system-derived metabolites along the gut-blood-brain axis.
  - Demonstrated ability to manipulate reciprocal interactions between microbes and model organisms that produce benefits to fitness in response to environmental challenges.

- Develop integrative multi-omics strategies to identify mechanisms of impaired memory and increased anxiety in animals exposed to environmental neurotoxicants, and to develop remediation strategies, as measured by:
  - Implementation of micro-fluid sampling biosensors and biomarker panels that measure metabolic fluxes in samples [e.g., blood, cerebral spinal fluid, and central nervous system (CNS) regions].
  - Successful utilization of advanced imaging and computational tools that map the locations of affected CNS cells, and model mechanisms of persistent memory impairment and anxiety.
  - Implementation of rodent platforms that test drug and dietary strategies for protecting against anxiety and memory loss from environmental neurotoxicants.

- Integrate multiscale imaging and computation to enable predictive models for the effects of environmental challenges on five prototypic microbiome/eukaryotic biological systems, as measured by:
  - Implementation of robust computational models that accurately simulate and predict responses to environmental challenges at different biological levels.
  - Identification and implementation of predictive bioindicators for fitness that include genes, epigenetic markers, proteins, metabolites, and microbiome components

Impact of environmental challenges on human biology

- Determine if environmental challenges shown to impact model organisms similarly affect human cells and tissues, as measured by:
  - Validation of human biomimetic tissue culture systems (e.g., breast, liver, cardiac, and skin) fabricated from normal primary human cells and extracellular matrices, with respect to relevance to in vivo tissues at the levels of architecture, gene, and protein expression patterns.
  - Integration of all biomimetic human tissues into single and parallel devices that capture variation in individual genomes.
  - Quantifying the effects of exposures (identified as having impact in model systems) on human biomimetic tissues using genomics, phenomics, and high content imaging.
  - Integrating discoveries about specific environmental challenges into exposome/ epidemiological studies of human responses.
  - Initiation of strategies for disease prevention, therapy, and risk management based on individual predispositions and systemic responses to environmental challenges.
  - Applying methodologies for in vivo physiological monitoring to both model organisms and humans, to bridge the gap between basic knowledge and human health and disease.

- Identify human biological response markers for up to six physiological variables (e.g., obesity, smoking, age, migrant lifestyle, diet, and pregnancy) as measured by successful deconvolving of contributions from genetic, epigenetic, transcriptomic, metabolomic, and microbiome components and variation to biological responses.

- Integrate mechanistic insights from human and model system studies to develop computational models to predict the effects of environmental challenges on human health, as measured by:
  - Generating a list of bioindicators for human health and disease that includes genes, epigenetic markers, proteins, metabolites and their fluxes, microbiomes, and physiological components.
  - Successfully predicting how manipulating host and microbiome properties positively or negatively impacts human nutrition, longevity, and pathogen/disease resistance.

- Use knowledge from studies in humans and model systems to design strategies for disease prevention, therapy, and risk management based on individual predispositions and systemic responses to environmental challenges, as measured by:
  - Using phenotypic data, genetic predispositions, and other biological assays to identify individuals at risk for harm from specific environmental challenges.
  - Deploying deep learning algorithms that integrate omics, phenotypic and exposome data that more effectively identify sensitive populations and biomarkers.
  - Developing personalized therapeutic and prevention strategies against at least two major diseases by utilizing links between genetic and epigenetic variation, microbiome composition, and environmental responses.
  - Working closely with clinical and advocacy partners to initiate development of clinical trials focused on disease prevention.
  - Implementing strategies to minimize the negative impact of bioeconomy technologies on human health while maximizing economic growth.
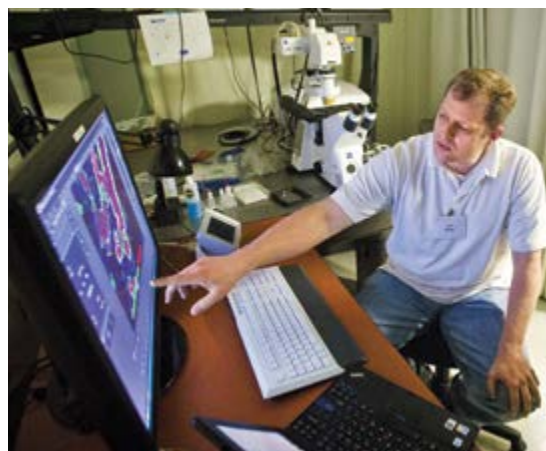
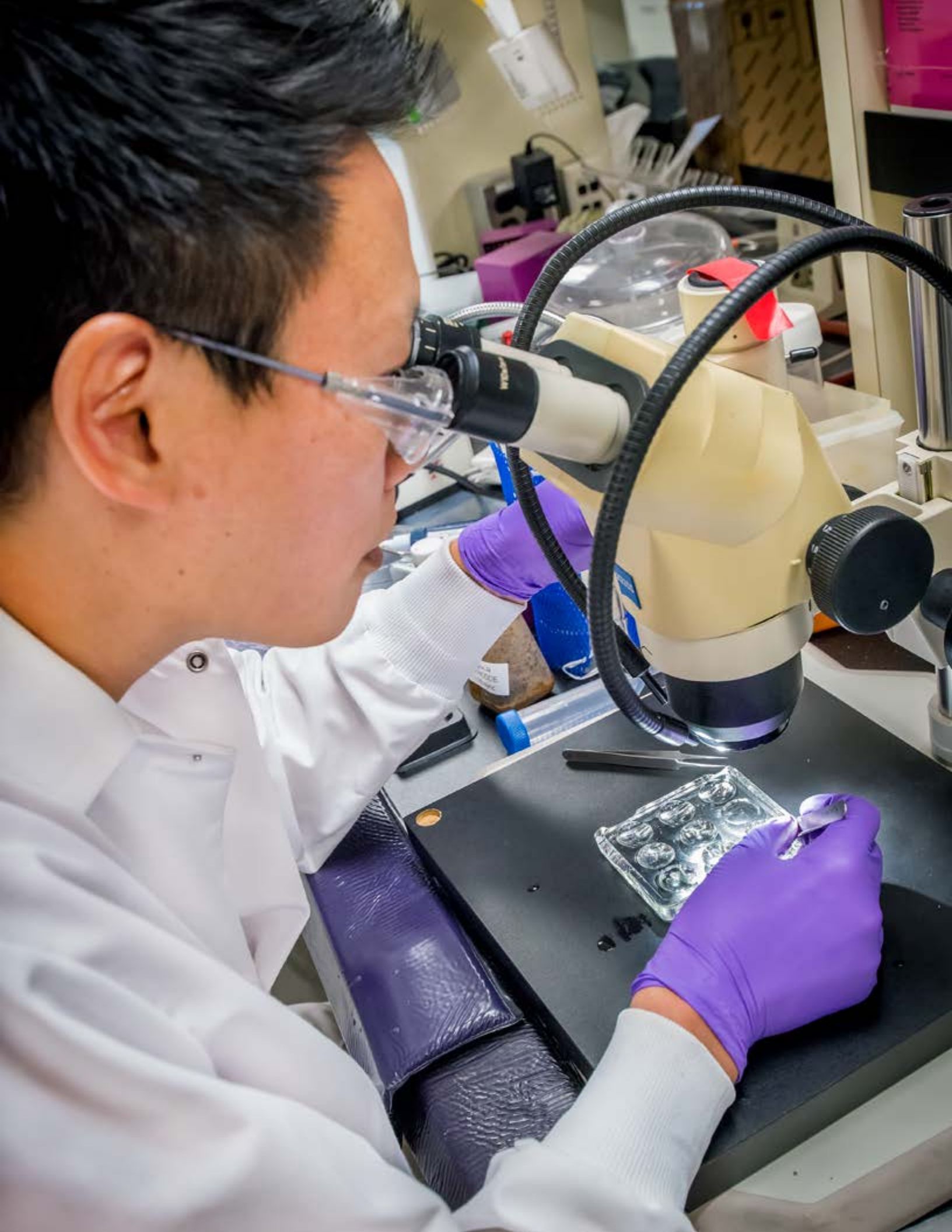## 2018 – Five-year milestones for health research strategies

Biological responses to environmental challenges

- Identified genetic, epigenetic, transcriptomic, metabolomic, proteomic, microbiome, and phenotypic responses (molecules, cells, tissues, and organisms) to two environmental challenges (e.g., anthropogenic pollutants such as heavy metals and synthetic organic compounds) in model microbial communities, biomes, and eukaryotic organisms (*Drosophila*, rats, and mice).

- Determined how the effects of these two environmental challenges are influenced by genetic and epigenetic variation, and whether they are transmitted across cell generations and trans-generationally.

- Demonstrated technologies for high-throughput characterization of macromolecular complexes acting in environmental responses.

- Integrated data to develop and test mechanistic and predictive models of two environmental responses, including microbiome/host interactions.

- Formulated a list of predictive bioindicators for fitness that includes genes, epigenetic markers, proteins, metabolites, and microbiome components.

Impact of environmental challenges on human biology

- Developed prototype body-on-a-chip devices that represent at least two biomimetic tissues or tissue-states (normal vs. diseased), fabricated from single source normal primary human cells and extracellular matrices. Validated relevance to in vivo tissues at the levels of architecture, gene, and protein expression patterns.

- Determined if at least one of the environmental challenges identified in model systems affect human biomimetic tissues in a predictable manner.

- Completed phenotypic analyses of human blood samples, focusing on the impact of diet and different stages of pregnancy, and the roles of genetics, epigenetics, transcripts, metabolites, and the microbiome.

- Formulated a list of key bioindicators for human health and disease that include genes, epigenetic markers, transcripts, proteins, metabolites, and microbiome components.

BIOSCIENCES FOR BIOMANUFACTURING

Background and Motivation

Forty years ago the development of recombinant DNA revolutionized biotechnology. The unprecedented ability to engineer bacteria to produce any molecule that could be encoded in a gene birthed multibillion-dollar industries in pharmaceuticals, materials, chemicals, foods, and fuels. The promise of biomanufacturing inspired four decades of improvement in the basic understanding of biology. Sequencing programs were developed to quickly and reliably identify what useful genes exist in ecosystems all over the world. Analytic technologies were developed to predict and pinpoint novel metabolites in almost any sample. A detailed understanding of biomolecular function and cellular physiology was developed that enabled the design of reliable tools for effectively expressing pathways and engineering hosts to maximize productivity. Synthesis of arbitrary DNA sequences was advanced, obviating the need for a physical template and easing the creation of larger recombinant constructs. And knowledge systems are now being built to allow us to learn from experience and to computationally control the design and manufacturing process. Systems are now being built that manufacture the organisms that in turn manufacture the chemicals and materials needed by society. With biology, solutions to long-term nationally significant challenges are now being created.

Given such radically improved tools, biological engineers have managed the manufacture of more exotic products, moving from single gene drugs like insulin and somatropin to multigene drugs like artemisinin, and moving from single-celled bacteria into plants and higher eukaryotes. And yet while costs have fallen and gains have been made, there are still significant technological barriers that limit the ability to reliably create biological solutions to some of the most pressing global problems of energy, environment, and health. Berkeley Lab has ambitious Goals — biofuels, novel antibiotics, helpmate bacteria, and modern materials — but

10-YEAR Biomanufacturing Goal

Develop and demonstrate a scalable, flexible, cost-effective, and sustainable biology-based manufacturing infrastructure driven by applications in energy, health, environment, and agriculture.

Biomanufacturing Research Strategies to Achieve Goal

Tools to design, construct, and debug biology. Develop computer-aided design and fabrication tools, computational, and analytical methods to model and learn from engineered biological systems and processes, and retrosynthetic infrastructure providing optimized pathways to key molecular hubs.

Designed biological systems. Engineer and scale-up controllable, trackable, and robust biological systems (prokaryotes, archaea, and eukaryotes) for key energy, health, and environmental biomanufacturing applications.

Biodirected materials and bionanosciences. Couple biological components to chemical and physical systems to biosynthesize desired mineral/metal nanostructures.

the infrastructure is lacking to reach them. Engineering and manufacturing have not been as predictive, as scalable, or as computationally supported as the more established scientific disciplines.

Berkeley Lab's Biosciences Area can close the gap by developing a science and technology program focused on the challenges of creating an advanced biomanufacturing facility. Berkeley Lab will create scalable infrastructure and knowledge systems and foster a network of expertise. In seeking innovations in predictive engineering and in biological systems design, Berkeley Lab will improve and broaden the capabilities of biological engineering and bolster its strategic missions in health, energy, and environment as well.

Berkeley Lab has the critical mass of researchers and infrastructure necessary to form a biomanufacturing power center. Core capabilities in genomics, microbiology, materials science, bioprocess development and demonstration, and computation will serve as the basis for a vibrant center for biological innovation that will address these important global challenges.

Biomanufacturing Research Strategies

- **Tools to design, construct, and debug biology.** Develop computer-aided design and fabrication tools, computational and analytical methods to model and learn from engineered biological systems and processes, and retrosynthetic infrastructure providing optimized pathways to key molecular hubs.

- **Designed biological systems.** Engineer and scale-up controllable, trackable, robust biological systems (prokaryotes, archaea, and eukaryotes) for key energy, health, and environmental biomanufacturing applications.

- **Biodirected materials and bionanosciences.** Couple biological components to chemical and physical systems to biosynthesize desired mineral/metal nanostructures.

To achieve the Biomanufacturing Goal by 2023, Berkeley Lab developed an approach that employs three strategies: (1) the development of tools to design, construct, and debug biological systems; (2) the design and scaling of biological systems; and (3) the creation of biodirected materials and bionanosciences. These strategies are scientifically tractable within a 10-year span. They will meet the long-term national need for novel biomanufacturing solutions, and leverage specific facilities, organized research groups, and core competencies within the Biosciences Area at Berkeley Lab. These strategies will be executed in parallel.

Tools to design, construct, and debug biology: Develop computer-aided design and fabrication tools, and computational and analytical methods to model and learn from engineered biological systems and processes, and retrosynthetic infrastructure providing optimized pathways to key molecular hubs.

The technical ability to edit and insert DNA into organisms has inspired visions of a new era of "synthetic biology" where novel genes could be designed and constructed for useful purposes. Today, whole-genome engineering promises to enable the manufacture of increasingly complex genetic designs. However, while the advent of genome-manipulation technologies has the potential to rapidly accelerate this process, progress is severely limited by the lack of knowledge of how to design and control the sophisticated gene networks and metabolic pathways in the context of an overall biomanufacturing

process at relevant pilot, demonstration, and commercial scales. Moreover, once a biological system has been constructed, there are limited tools to debug the system and improve upon it. Hence, tools are needed to design and fabricate biological systems, and model and learn from biological systems and processes once they are constructed, as well as infrastructure providing optimized pathways to key hub molecules in retrosynthetic design space. Critical to these goals is the need for structural and imaging tools to precisely characterize resulting biological systems at the protein-specific level, using cryo-EM and other scattering methods, and the need for imaging systems at the nm level using a new integrated bioimaging center.

Biological Computer-Aided Design (BioCAD)

Advanced engineering relies on sophisticated mathematics, informatics, and computing infrastructure to guide all aspects of system design and manufacture. Electronics design automation revolutionized the electronics industry with tools spanning simulation of silicon materials, physical and logic design of circuitry, and physical layout and manufacturing optimization engines. These billion-dollar-scale infrastructures are based on libraries of knowledge and models of physical principles, standard manufacturing protocols and design elements (physical parts), and design templates for standard applications. Standards for information interchange, algorithmic update and testing, and form factors and interconnects have been specified so that multiple horizontal industries can compete to serve various vertical application industries.

Today bioengineering has few such infrastructures. While there is a maturing genomic information infrastructure in the form of sequence databases (such as those available at the National Center for Biotechnology Information and the significant environmental sequence resources as well as relevant analysis tools at the DOE JGI) and there are active efforts to move towards more sophisticated, functional genomics databases and modeling tools (such as those available in the KBase), and standards for the simulation and automated design, construction, and iterative improvement of biological systems and processes remain in their infancy. Beyond functional genomics data, structural biology and bioimaging present new opportunities for integration into the design process. Gene synthesis expertise and infrastructure also exists at the DOE JGI and can be leveraged for program development in this strategy. A fully developed biological design automation framework would do much to democratize and advance biomanufacturing broadly.

Biological debugging tools

Biological engineering is hindered by extremely long optimization and troubleshooting cycles. When newly engineered biological systems fail, it is often difficult to determine why the system failed. The ability to rapidly diagnose failed biological designs by combining the tools of systems biology with computational analysis could greatly decrease cycle times. Incorporating the results of the debugging process into the computer-aided design tools would further improve future initial designs.

Berkeley Lab's bioscientists have demonstrated leadership in many research areas that relate to the development of tools for biomanufacturing. DOE has funded Berkeley Lab to use the latest functional genomics tools to analyze biological systems in the environment, and these same functional genomics tools can be used to analyze engineered biological systems. Integrating the information from these functional genomics tools into testable

hypotheses is the mission of the KBase, which Berkeley Lab scientists lead. Berkeley Lab scientists have also led the development of nascent bioCAD tools, such as j5, as well as computational and analytical approaches to model and learn from biological systems and processes. Finally, Berkeley Lab scientists have begun to build out a biomanufacturing infrastructure around optimized routes to retrosynthetic targets of opportunity that greatly enable biosynthetic access to downstream molecules.

Designed biological systems: Engineer and scale-up controllable, trackable, robust biological systems (prokaryotes, archaea, and eukaryotes) for key energy, health, and environmental biomanufacturing applications.

Humans have an established history of modifying the natural world to their own ends. The child-friendly dogs in the pet store and the huge juicy carrots in the grocery store are the direct result of human will impressed upon the DNA of promising life forms. The process of selecting for pliability, safety, and utility in organisms is termed domestication, and the only apparent downside so far has been the amount of time it takes to accomplish.

Certain microorganisms have been domesticated. Bacteria and fungi involved in fermentative food, feed, beverage, fuel ethanol, and pharmaceutical production are the best understood, safest, and most manipulated microorganisms on the planet. In the last few decades, sequencing technology has begun to identify exactly what genetic changes are correlated with domestication, and genetic manipulation technology has enabled the ability to make those changes directly. Biological engineers have taken advantage of these modern tools to rapidly customize the domesticated organisms further, engineering and scaling them to produce a broader range of natural products than ever before.

However, successes have been constrained by the limited set of organisms in play. Decades of *Escherichia coli* domestication have provided an impressive array of genetic and genomic engineering tools — most of which only work in *E. coli*. When engineers plan the manufacture of a natural product, they are limited to organisms that are well understood and genetically pliable for historical reasons, but are not necessarily inherently capable of the biochemistries required for production. This has a dramatic impact on design time and difficulty, scalability and product titer, and the range of products that can actually be engineered.

There is an alternative. Every desirable natural product comes from a living organism, often plants or microbes, some of which may be quite amenable to domestication. That is, rather than re-engineering desirable pathways into a pliable but basically incompatible host, suitability could be engineered into already productive organisms. Adding new bacteria, archaea, fungi, and plants that are generally regarded as safe to the stable of tractable hosts would not only advance the science of genetic manipulation, it would significantly broaden the range of natural products amenable to being biomanufactured.

Microbiology, botany, synthetic biology, and process development and scale-up are Berkeley Lab strengths. For example, Berkeley Lab scientists have been central in developing CRISPR/Cas9 technologies for scalable, cross-kingdom engineering and regulation of microbes, plants, and mammalian cells. Microbes have been engineered to produce a vast array of active pharmaceutical ingredients, advanced biofuels, and commodity and specialty chemicals and their intermediates. And microbes have been produced that are holistically amenable to the biomanufacturing processes of plant biomass deconstruction and raw cellulosic hydrolysate consumption. The ABPDU, in conjunction with industry sponsors, has demonstrated commercially relevant

bioprocesses for the production of biofuels, bio-based chemicals, food ingredients, and materials.

The engineering of controllable, trackable, and robust biological systems for biomanufacturing applications will depend upon the identification of likely organisms, the development of domestication protocols, and the creation of novel and potentially idiosyncratic genetic toolkits. No institution is better poised to make this massive contribution to biomanufacturing than Berkeley Lab.

Biodirected materials and bionanosciences: Couple biological components to chemical and physical systems to biosynthesize desired mineral/metal nanostructures.

Molecular self-assembly, the process by which molecules spontaneously adopt a desired arrangement without external guidance, underlies the construction of macromolecular assemblies that enable cells to function. Because of this inherent "programmability," molecular self-assembly has also become fundamental to certain aspects of nanotechnology and mesoscale science, and there has been a recent bloom in the areas of programmable biomolecular assemblies and biodirected materials. For example, researchers have developed highly sophisticated drug-delivery vehicles that decorate and are encapsulated by cell-mimetic materials and allow an unprecedented degree of control over the localization, specificity, timing, and dose of pharmaceuticals to specific disease sites. In another case, a viral platform has been created for biologically assembling sophisticated materials at the nano-scale, such as gold and silver noble-metal wires with high aspect ratios and diameters below 50nm that can be used as cathodes for lithium ion batteries. Nucleic acids have also arisen as nanoscale supramolecular building blocks and so-called DNA origami can self-assemble into arbitrarily-shaped 2-D and 3-D nanomaterials. These applications rely on the long-term development of macromolecular and viral engineering frameworks that provide a foundation for developing new molecular components in new arrangements for new applications, and just a handful of laboratories have technological expertise in using these systems. The rate of new innovations in this area is high, but the translation of the results into industrial application and scaling beyond a few highly skilled laboratories has lagged. Increased investment in macromolecular design, scaling manufacture of cell-mimetic systems, and computational and experimental methods for supramolecular assembly design will greatly enhance biodirected manufacturing capabilities.

At Berkeley Lab, the intersection of biology and nanotechnology is strong and has potential to do what no other research entity can do in this area. Berkeley Lab's Molecular Foundry, which has assembled state-of-the-art tools for enabling nanoscience and hosts users, has a particular strength in the integration of biology and nanotechnology. Berkeley Lab's combination of expertise in synthetic biology and nanotechnology and available tools and hardware for synthesizing and characterizing biomaterials position Berkeley Lab as a strong leader to drive national-scale scientific advances for biomanufacturing.

2023 10-year achievement measured by:

Tools to design, construct, and debug biology

- Develop a bioCAD/CAM infrastructure comprising tools for:
  - Pathway retrosynthesis and host engineering for optimized production titers, rates, and yields under industrially-relevant conditions.

- Integrating functional genomics data into the design process.
- Learning from characterization results to inform the design process.

- Develop simulation capabilities for biomanufacturing systems:
  - Small-scale (< 2 L) physical and computational simulations of large-scale (> 100 L) reactors to aid in strain optimization.
  - Integrate emerging sensor-measurement data to increase the accuracy of reactor modeling approaches

- Develop detailed and extensible techno-economic models for manufacturing processes and applications:
  - Model-informed selection of strategies for synthesis and characterization.
  - Holistic consideration of bioprocesses, from feedstock to reactor to downstream processing, integrating chemical and biological process steps, and the path to production at scale.

- Develop retrosynthetic infrastructure:
  - Establish optimized routes to at least 30 key retrosynthetic molecular intermediates.

Designed biological systems

- Establish a robust protocol for host organism domestication.
  - Identify common barriers to genetic pliability and laboratory cultivation.
  - Develop broad-range tools for host manipulation (i.e. plasmids and CRISPR-Cas systems).
  - Create models of physiological changes induced by domestication and scale-up of manufacturing drawn from functional genomics analyses of previously domesticated hosts.
  - Develop means for identifying and tracking engineered organisms in complex environments.
  - Develop means for removing engineered organisms from complex environments.
  - Create models for predicting and platforms for assessing the environmental impact of engineered organisms.

- Nominate a range of key hosts for biomanufacturing, agriculture, environmental remediation, water support, and human health.
  - Create a host database serving bacterial information and protocols related to cultivation and transformation.
  - Determine active and potentially tamable hosts compatible with key application environments.
  - Identify host properties that can be engineered for improved robustness, productivity, and optimized energy throughput.
  - Develop robust genetic toolkits for each host: expressing and integrating plasmids, media recipes and transformation protocols, characterized regulatory parts, and selection and screening markers.

- Establish robust tools for plant engineering.
  - Create a host database containing genome information (size, ploidy, gene expression, etc.) and protocols related to cultivation and transformation.
  - Development and characterization of "universal" plant expression tool kits.

- Generation of public part libraries and associated databases.
- Development of a pipeline for rapid and efficient trait stacking.

- Develop and demonstrate bench- and pilot-scale fermentation processes for novel bio-products.
    - Quantify the impact of gene expression and metabolism on the scale-up of biomanufacturing fermentations.
    - Identify, prototype, and deploy new protocols for more flexible and efficient production of precursor biomolecules and bioproducts.

Biodirected materials and bionanosciences

- Invent new routes for the design of biohybrid systems that mechanically or electronically interface active biological elements with polymeric and inorganic materials.
    - Demonstrate electron transfer between intracellular and extracellular redox active species along a molecularly-defined path.
    - Create at least two molecularly-defined pathways that operate at different redox potentials.

- Achieve ability to interface biological components to electronic apparatus to control their activity.
    - Develop methods to electronically modulate catabolic fluxes.
    - Develop methods to electronically stimulate intracellular reactions.
    - Interconvert electrical energy with chemical and/or light energy.

- Build infrastructure to understand charge transfer at the abiotic/biotic interface.
    - Develop methods to structurally characterize redox active molecules at the abiotic/biotic interface.
    - Develop spectroscopic methods to characterize charge transfer and dynamics energetics.
    - Create a basic framework that relates structure with energetics and dynamics of charge transfer.

- Achieve the ability to biosynthesize architecturally specified, possibly self-healing, mineral/metal nanostructures and mesostructures on demand, using biological entities.

2018 – Five-year milestones for biomanufacturing strategies

Tools to design, construct, and debug biology

- BioCAD/CAM infrastructure developed comprising one or more tools each for:
    - Pathway retrosynthesis and host engineering for production titers, rates. and yields under industrially relevant conditions.
    - Integrating functional genomics data into the design process.
    - Learning from characterization results to inform the design process.

- Small-scale (< 2 L) physical simulation of a large-scale (> 100 L) reactor developed to aid strain optimization.

- Biomanufacture of two key product molecules designed, implemented, and optimized.

- Biological routes to 10 key retrosynthetic molecular intermediates established.

Designed biological systems

- A broad host range domestication protocol designed for use in non-model organisms.

- Domesticated one previously intractable host.
    - Used functional genomics to identify host systems responsive to domestication and manufacture scale up.
    - Proposed a means for identifying and tracking the engineered host in different environments and assessing its impact.
    - Improved biosynthesis of at least one product at previously unattainable yields.
    - Demonstrate that lessons learned apply to an unrelated host.

- Developed and demonstrated new bench- and pilot-scale unit operations and integrated processes for the production of a novel bioproduct.

- Engineered plants that produce modified lignin that can be easily transformed into a useful commodity chemical.
    - Identified aromatic precursors that can be used in metabolic pathways to efficiently produce chemicals of interests.
    - Manipulated plant metabolism to redirect metabolic flux toward desired products (e.g., commodity chemicals, aromatic precursors, and novel monolignols).
    - Engineered monolignol pathways to produce novel lignin that are efficiently deconstructed and are more readily processed to commodity chemical.
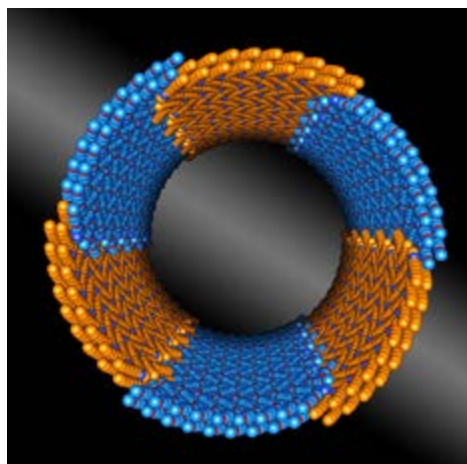
Biodirected materials and bionanosciences

- Prototyped a biohybrid system that interfaces active biological elements with other chemical, physical, or electronic materials.
    - Demonstrated molecularly-defined extracellular electron transfer to metals, metal oxides, and electrodes.
    - Demonstrated molecularly-defined electron transfer from an electrode to intracellular species.
    - Identified at least two electron transfer pathways to extracellular acceptors that operate at different redox potentials.

- Structurally characterized an example of a redox active molecules that functions at the abiotic/biotic interface.

- Developed a prototype multifunctional platform intended for manufacturing of two or more biodirected materials.

TECHNOLOGIES FOR BIOSCIENCES

Background and Motivation

Science is not possible without methods for measurement, visualization, and modeling. The combination of experimental data and computation is what enables prediction and manipulation of the world around us. From the light microscope of the seventeenth century to the particle accelerator of today, the tools of science have grown exponentially in sophistication and capability. The cornerstones of the Biosciences Area technology platform are imaging biological systems from atoms to cells to organisms, functional characterization, decoding and understanding the function of genomes, and computational methods for analysis and prediction.

Today instruments exist to view the molecular world in atomic details, with over 100,000 protein and nucleic acid structures solved. But molecular structure is just the beginning of understanding and controlling biological systems. Rarely do proteins in cells work alone, or in a static configuration. Tools for kinetic analysis, imaging of larger order structures such as complexes or subcellular structures, measuring cellular physiochemistry *in situ*, assembly and analysis of complex plant feedstock genomes, and sequence-based analysis of the composition and function of complex environmental communities of organisms must all be brought to bear to gain a complete picture and true predictive power. Macromolecules are in constant motion, partnering with other molecules in a constant, complicated dance, and thus no matter the "zoom" level, no two complexes, organelles, cells, or organisms are identical. Thus, novel methods must be developed to distinguish and quantify the similarities and differences across systems. The major challenges of this century are multiscale and thus require cross-cutting approaches. No one technology can be used to single-handedly address these challenges. Application of these tools requires an integrated approach, both in instrumentation and analysis.

10-YEAR

TECHNOLOGIES Goal

Develop a technology infrastructure to measure, predict and control biological systems for solving energy, environmental, and health challenges.

Technologies for BioSciences to Achieve Goal

Structural Biology; Bioimaging; Functional Genomics; Mathematics, Informatics, and Computing.

Berkeley Lab houses both the expertise and advanced instrumentation to make this Goal a reality. Over the last two decades at the ALS, crystallography has been developed for atomic level structural information of macromolecules, automated small-angle X-ray scattering to provide accurate shape information about macromolecules in solution, tomographic imaging of whole cells, and infrared spectroscopy of live biological samples. The DOE JGI provides high-quality DNA sequencing, genome assembly, analysis and functional annotation, and DNA synthesis capabilities. The ABPDU includes reactors for biomass pretreatment, controlled-environment fermentation capacity from 3 L to 300 L, and product analysis capabilities. The Molecular Foundry enables research on the nanoscale in a multidisciplinary, collaborative environment. NERSC provides large-scale, state-of-the-art computing, storage, and networking for unclassified research programs. Berkeley Lab is developing a new cutting-edge cryo-electron microscopy facility that will take Berkeley Lab's high-resolution imaging capability to new heights, enabling analysis of samples in their native environment without staining or fixation. The resolution of cryo-EM is now capable of generating structures at near atomic resolution. New resources for mass spectrometry imaging and scanning ion optical imaging will provide new frontiers for integrating imaging with functional properties of cells.

Structural Biology

The workhorse technique of crystallography has been used to delineate over one hundred thousand molecular structures to date. As a mature and accessible method, crystallography has additionally benefitted from internationally accepted databases to store solved structures, and numerous computational programs designed to aide in structure solution and visualization. As has been proved with the crystallographic databases, as more data is shared more widely, more science is enabled more broadly. In the case of protein crystallography, access to a greater number of previously solved structures gives a higher probability of success for future structures. The technique of small-angle scattering is also approaching maturity, with numerous software analysis packages and instruments available worldwide. Another structural method, cryo-electron microscopy (cryo-EM), has significantly benefitted in recent years through advances in detector technology, and this has enabled larger molecular complexes to be visualized at resolutions approaching the atomic level. Other structural methods, such as X-ray tomography and infrared spectroscopy are being applied to ever more diverse biological systems and gaining wider access by the national scientific community. These techniques inform and enhance one another. For instance, crystallographic structures are often used to help refine cryo-EM structures. Scattering data is often essential in determining the molecular envelope for complexes in which crystallography data has supplied the individual subunits.

Biosciences aims to link these structural methods into a unified platform in both shared instrumentation and in data analysis programs and knowledge databases, to gain a full picture of the conformational and functional flexibility of biological components. The first step towards an integrated infrastructure is the physical colocation of instrumentation, beginning with the housing of a cryo-EM facility near the ALS. User programs for electron and X-ray methods will be integrated to allow simultaneous data collection at both facilities on the same systems. Within the synchrotron facility, instrumentation will be developed to pair methods in which data can be collected simultaneously, as in the case of scattering and footprinting, or spectroscopy and crystallography. For samples requiring different preparation states, such as X-ray

tomography and infrared, user programs will be integrated to enable serial data collection and collaborative data sharing. In the area of data integration, common formats will be developed to share cross-technique structural information and to enable teams of scientists with diverse areas of expertise to communicate effectively about a common problem.

10-year Goal achievement measured by:

- Methods for correlating and analyzing structural and functional data from the molecular to the cellular level.

- Computational analysis programs capable of data structure input from all synchrotron X-ray and electron diffraction methods.

Five-year milestones:

- Establishment of a national biosciences cryo-EM facility.

- Integration of scattering and diffraction structural methods at the ALS.

Bioimaging

All biological systems display unique behaviors, including self-organization across temporal and spatial scales ranging from atoms to organisms. Visualization provides perhaps the most powerful basis for understanding the behavior of molecular components in the context of cells, tissues, organisms, and communities. For example, in the termite hindgut, living multispecies microbial communities degrade plant material into sugars for bioenergy production. Harnessing this system promises to improve biomass production, but more information is needed to describe the microbial species present, their location relative to the plant material, what enzymes are being secreted, and what cellular pathways regulate the secretion. In higher organisms, brain cells operate within a complex microenvironment comprising multiple connections and signaling systems, which are altered by environmental toxicants. The responses, however, depend on which network connections are targeted, which neurotransmitters are affected, how fast they respond, the pathways that regulate neurotransmission, and the blood-gut-biome communication.

Despite obvious complexity in biological systems, most advances in bioimaging technologies have been focused on a single modality, which often fails to adequately address the spatial and temporal realities of biological systems, their cell-cell communication, and their ability to respond to challenge. Berkeley Lab plans to address these challenges through advances in instrumentation and integrated computational strategies for large and diverse data arrays, allowing multimodal imaging across a range of length and time scales. This will build on available expertise and instrumentation for imaging cells, tissues, and organisms using fluorescent probes, radioisotopes, and electromagnetic radiation.

10-year goal achievement measured by:

- Availability of probes, labeling chemistries, and label-free approaches that provide contrast across multiple imaging modalities.

- Creation of visualization, modeling, and interaction systems for experimenters to efficiently extract knowledge from data.

- Development of new sample preparation methods compatible with multi-model imaging.

Five-year milestones:

- Establishment of a center for integrated bioimaging.

- Multiple imaging modalities applied to imaging biological systems from the nm to mm length scale and the msec to hour time scale.

- New light microscopy technologies, including new contrast probes and labeling chemistries, integrated into existing imaging systems.

- Model-based algorithms developed for combining information across multiple imaging methods and integrating functional data.

- Collaborations with Berkeley Lab mathematics, informatics, and computing researchers result in big data analysis methods applied to several challenging biological problems.

Functional Genomics

Robust, high-throughput functional genomic analyses are critical to diverse biosciences strategies, including understanding the genomics of microbial communities, using synthetic biology to develop sustainable energy and materials, and measuring and monitoring phenotypes associated with engineered systems and biological responses to environmental exposures. Two major new research thrusts are the development of (1) high-throughput technologies that provide rapid feedback on the performance of engineered systems, real-time monitoring of exposures and responses, and the direct linkage of DNA sequence to function, and (2) analytical approaches that provide detailed insights into the genetic and physiological states of biological systems and link to bioimaging and computation. These new efforts promise to bring large-scale research teams together to achieve our 10-year goals.

10-year achievement measured by:

- Demonstrated technologies for rapid phenotyping, especially mass spectrometers with enhanced analytical chemistry capabilities and front-end microfluidic chip-based automation for efficient quantitative assessments.

- Characterized dynamic processes using high-throughput functional analyses at multiple biological scales for integrative computational analyses and predictive models.

- Discovery of novel gene functions, associations, and new branches of life from critical environmental and uncultivated microorganisms.

- Successful application of rapid sequence-to-function analysis technologies to discover biological activities and enable advanced pathway engineering.

- Developed isotope-based methods, coupled to systems biology tools, for identification of active microbes in complex communities and their modes of exchanging metabolites.

Five-year milestones:

- Demonstrated effectiveness of an automated chip-based mass spectrometry platform with enhanced throughput and analytical chemistry capabilities.

- Demonstrated advanced technologies for high-throughput functional genomic analyses that are tightly integrated with computational resources.

- Applied comparative gene expression coupled with metabolomic, proteomic, and fitness data for biochemical discovery and testing of genomic predictions.

- Demonstrated technologies to identify active metabolic pathways within complex multicellular systems.

Mathematics, Informatics, and Computing

Mathematics, informatics, and computing play a central role as biology continues to evolve from an anecdote-based, descriptive field of science into a quantitative discipline relying on, and driven by, large data sets. Advanced data analysis tools underpin modern drug discovery and genomic analyses, and have been essential in elucidating a fundamental understanding of cell and microbial communities. The Biosciences Area has a rich track record in pursuing these types of studies, having developed a number of key technologies for the analysis of crystallographic data (Phenix), visualization, analysis, and management of Mass Spectroscopic Imaging (OpenMSI), analysis of three-dimensional cell culture images and their response to external perturbations (BioSig3D), and the standardization and optimization of synthetic DNA construction (j5). The DOE JGI and KBase are developing new computational methods to advance biosciences research. The DOE JGI is focused on the generation, processing, and analysis of large genome-scale datasets. The current strategic development directions at the DOE JGI in the computing arena are, for example, focused on (but not limited to) increases in algorithmic and pipeline efficiency, and moving towards an exascale aware configuration of hardware and software solutions such that the DOE JGI can provide a state-of-the-art environment in which to fulfill its mission of advancing energy and environmental sciences via state-of-the-art capabilities in genomics. The KBase platform allows the scientific community to collaboratively drive a data-driven understanding of genes, proteins, and organisms and their interactions, ultimately allowing researchers to predict, engineer, and control biological systems, from a micron scale up to ecosystems.

Future developments will exploit the strong Berkeley Lab computational environment. NERSC provides high-performance computing solutions to the scientific community and is a world leader in accelerating scientific discovery through scientific computing. ESNet is a high-performance unclassified network built to support scientific research, providing computer network capabilities to over 40 DOE research sites, all national laboratories, supercomputing facilities and major scientific instruments. The Center for Advanced Mathematics for Energy Research Applications (CAMERA) is focused on the development

of novel mathematics and algorithms that either significantly improve current approaches or enable the analyses of novel experimental techniques.

Research in mathematics, informatics, and computational sciences will form an integral part of the technology strategies. Methods will be developed to extract maximal information from weak data, enabling the analysis of experiments from free electron lasers and electron microscopes. A key advance will be new algorithms that integrate experimental information of multiple types. Finally, the most significant breakthrough will be methods that enable automated knowledge generation from complex data types, constructing mathematical models that explain a wealth of experimental observations.

10-year achievement measured by:

- An improved overall computational efficiency of major algorithms in structural biology, genomics, bioinformatics, and multiscale modeling.

- The development of novel mathematical and computational approaches enabling new scientific approaches or solving outstanding problems in biophysics, biochemistry, and genomics.

- In collaboration with NERSC and the Computational Research Division of the Computing Sciences Area, the redesign of existing computational approaches that can make optimal use of novel hardware, such as exascale computers.

- The development of a mathematical and algorithmic framework that allows the joint analyses of multimodal methods, operating on length scales from atoms to cells to organisms.

Five-year milestones:

- Novel data analyses methods developed for existing methods, such as X-ray crystallography, small and wide angle X-ray scattering (SAXS / WAXS), and X-ray tomography.

- New applied mathematical and algorithmic methods developed for the analysis of the emerging methods that make use of X-ray free electron lasers (XFEL), near atomic cryo-electron microscopy, electron tomography, and correlated X-ray scattering.

- New methods developed that bring together genomic, imaging, and functional genomic data types across multiple resolution ranges through computational models, in the process generating new knowledge of biological systems.

- Biological activities, at the atomic, cellular, and organismal level predcited through the use of integrated analysis and computation.

BIOSCIENCE CAPABILITIES

Berkeley Lab's integrated biosciences efforts benefit from the expertise of a large staff of leading researchers, access to world-class facilities, and the organizational strength of Berkeley Lab's divisions and affiliated research institutes. Capabilities are vast, and span Biosciences Divisions, mission-focused centers, national user facilities, and the synergistic teams that collaborate on national-scale biosciences research efforts.

Berkeley Lab Biosciences Divisions

In 2015, the Biosciences Area underwent an inclusive, bottom-up re-organization of its structure and developed three new Divisions that align with the strategic Goals of the Area.

The Biological Systems and Engineering Division — advances mechanistic and predictive understanding of complex biological systems as a function of environment and scale and then translate this knowledge using engineering principles to advance resilient and efficient solutions in synthetic biology, oncology, agriculture, biofuels, renewable chemicals, materials, and abiotic-biotic systems.

The Environmental Genomics and Systems Biology Division — links genome biology to ecosystem dynamics through deep understanding of the genetic and molecular mechanisms governing the activities and ecologies of organisms and multispecies communities to predict responses and harness microbes and plants for energy and environmental solutions.

The Molecular Biophysics and Integrated Bioimaging Division — studies the fundamentals of biology with the goal of learning how to manipulate, control, and create biological functions to solve national challenges in energy, environment, health, and biomanufacturing. By using advanced imaging technologies that probe the structure, chemistry, physics, and dynamics of complex systems, the secrets of critical biological processes are being revealed.

Mission-Focused Efforts

The Joint BioEnergy Institute (JBEI) is one of three bioenergy research centers created by the DOE in 2007 to advance the development of transportation fuels from lignocellulosic biomass. Key capabilities at JBEI include basic gene discovery in plants, microbes, and microbial communities; process development for cellulose extraction from biomass; engineering fuel synthesis in microbes; and synthetic biology/biodesign.

The Joint Center for Artificial Photosynthesis (JCAP) is the nation's largest research program dedicated to the development of an artificial solar-fuel technology. Established in 2010 as a DOE energy innovation hub, JCAP employs capabilities in physics, chemistry, materials science, and nanotechnology to find a cost-effective method to produce liquid fuels using only sunlight, water, and carbon dioxide. JCAP is led by the California Institute of Technology with Berkeley Lab as its lead partner.

The Systems Biology Knowledgebase (KBase) is an extensible and scalable open-source software framework and application system to support the analysis of microbes, microbial communities, and plants. Recently launched, KBase will ultimately offer free and open access to data, models, and simulations. This will help scientists and researchers integrate various data types to build new knowledge and share their findings with others.

The BioFoundry is a multi-lab effort created by the DOE in 2016 to enable the realization of a thriving national bioeconomy through advanced biomanufacturing. At the BioFoundry, product development is conducted in an automated Design-Build-Test-Learn integrated fashion that generates robust organisms, genetic circuits, and conversion pathways tunable to industry demand, with predictable process scaling and potential downstream process development impacts addressed at the outset.

Berkeley Lab's Computing Sciences Area — including the National Energy Research Scientific Computing Center (NERSC), the Computational Research Division (CRD), and the Scientific Networking Division (SND) — similarly provides both the infrastructure and the opportunity for collaborative relationships. NERSC is also a designated DOE user facility. Biosciences Area researchers enjoy ready access to two other national user facilities operated by Berkeley Lab: the Molecular Foundry and Energy Sciences Network (ESnet), which provides a data highway for all laboratories affiliated with DOE.

National User Facilities

The Advanced Light Source (ALS) is one of the world's premier synchrotron radiation facilities and hosts hundreds of investigators from the biosciences community every year on its beamlines. Capabilities of the ALS especially valued in biosciences research include macromolecular crystallography (MX) for atomic level structural information of proteins and nucleic acids singly or in complexes, automated small-angle X-ray scattering (SAXS) capabilities that provide accurate shape and assembly information about functioning macromolecules in solution, tomographic imaging of whole cells (the National Center for X-ray Tomography), and infrared spectroscopy of live biological samples (the Berkeley Synchrotron Infrared Structural Biology Program). Collectively, these resources are essential to research carried out by the Berkeley Lab Biosciences Divisions.

The Joint Genome Institute (JGI) is a DOE-funded high-throughput genomics user facility operated by the University of California. It is the world's leader in production of plant and microbial genomes, as well as a pioneer in the burgeoning field of metagenomics — microbial community sequencing and analysis. The DOE JGI provides high-quality DNA sequencing, genome analysis, and DNA synthesis capabilities not readily available to users seeking alternative energy solutions and understanding of environmental phenomena, including JBEI and other Berkeley Lab investigators.

The Advanced Biofuels and Bioproducts Process Demonstration Unit (ABPDU) is a state-of-the-art user facility for testing and developing emergent biofuels technologies. It was funded by DOE to allow laboratory-scale processes for fuels synthesis to be scaled up and commercialized. The facility includes reactors for biomass pretreatment, controlled-environment fermentation capacity from 3 L to 300 L, and product analysis capabilities. The 15,000-square-foot facility is available to bioenergy research centers, DOE-supported researchers, academic institutes, nonprofit research organizations, and companies involved in biofuels R&D production.

The Molecular Foundry, one of five DOE nanoscale science research centers, provides users with access to expert staff and leading-edge instrumentation to enable research on the nanoscale in a multidisciplinary, collaborative environment. Selected through an external peer-reviewed proposal process, nearly 700 users come each year from academic, industrial, and national laboratories, both domestic and international, free of charge. Research that results in over 300 publications each year is organized into seven closely coupled facilities: Inorganic, Organic, and Biological facilities for material synthesis, preparation, and assembly; Nanofabrication, for processing and integration; the National Center for Electron Microscopy and Imaging and Manipulation, for characterization; and Theory, for understanding and predicting material properties.

The National Energy Research Scientific Computing Center (NERSC) is the mission high performance computing facility for the DOE's Office of Science (DOE SC); it has about 6,000 users from universities, national laboratories, and industry. NERSC's primary goal is to accelerate scientific discovery at the DOE SC through high performance computing, related technology development, and data management and analysis. Toward this end,
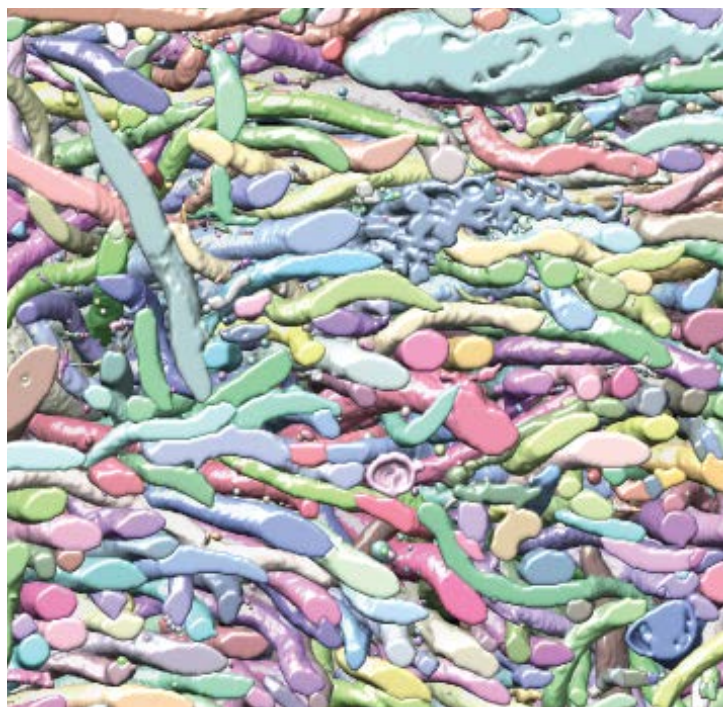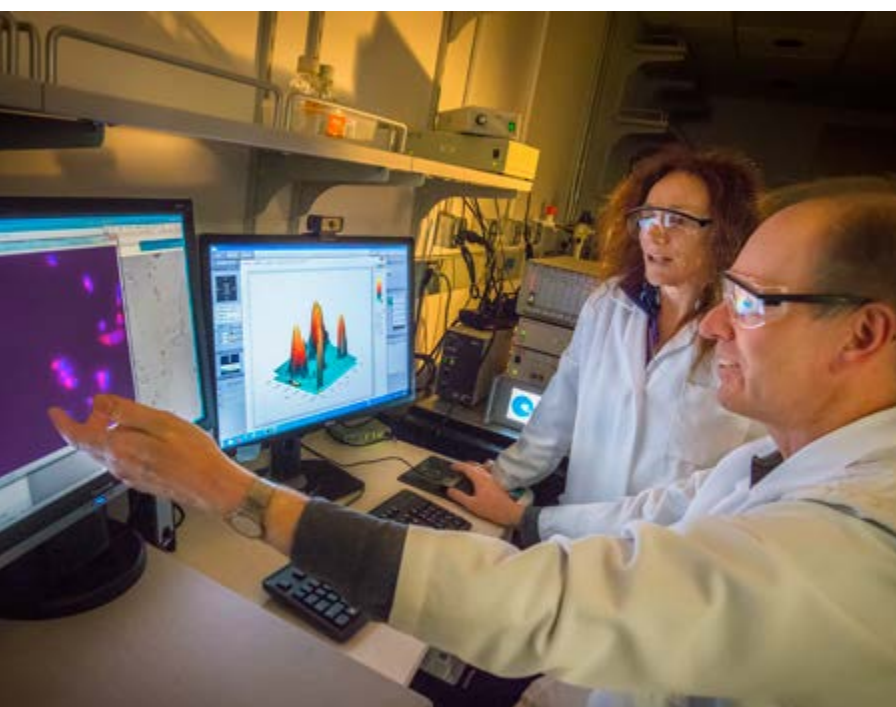
NERSC provides large-scale, state-of-the-art computing, storage, and networking for DOE SC's unclassified research programs in alternative energy sources, climate change, energy efficiency, environmental science, and other science areas.

ESnet is the ultrafast data highway for all DOE national laboratories. If supercomputers like NERSC provide the horsepower for data-intensive science, ESnet provides the connectivity. Large-scale collaborative research is the heart and soul of the modern scientific enterprise. Researchers today share data sets in the petabyte range — a million times the size of files familiar to consumers. The network links tens of thousands of researchers at more than 40 institutions, at high speed and securely. ESnet now provides high-speed bandwidth across the Atlantic to Europe with capacity of 340 Gbps, and provides connectivity to London, Amsterdam and CERN in Geneva.

Collaborative Research and Resources

Biosciences research teams at Berkeley Lab have an exceptional history of productive collaborative interaction. For example, the DNA sequencing capabilities at the DOE JGI are used to support JBEI, ENIGMA, and low-dose radiation research, as well as environmental research by Berkeley Lab's EESA. JBEI is also a user of the ABPDU for development of novel biomass deconstruction methods. The DOE JGI is engaged with KBase to develop and serve analytical tools and public data to scientists studying DOE-relevant problems. During the next 10 years, Berkeley Lab expects new collaborations and capabilities will emerge. To enhance collaboration among Biosciences research teams, Biosciences plans to relocate all its researchers to the Berkeley Lab campus. The first building, the Integrative Genomics Building, is proposed to house the DOE JGI and KBase and will bring together these two complementary and synergistic DOE programs.
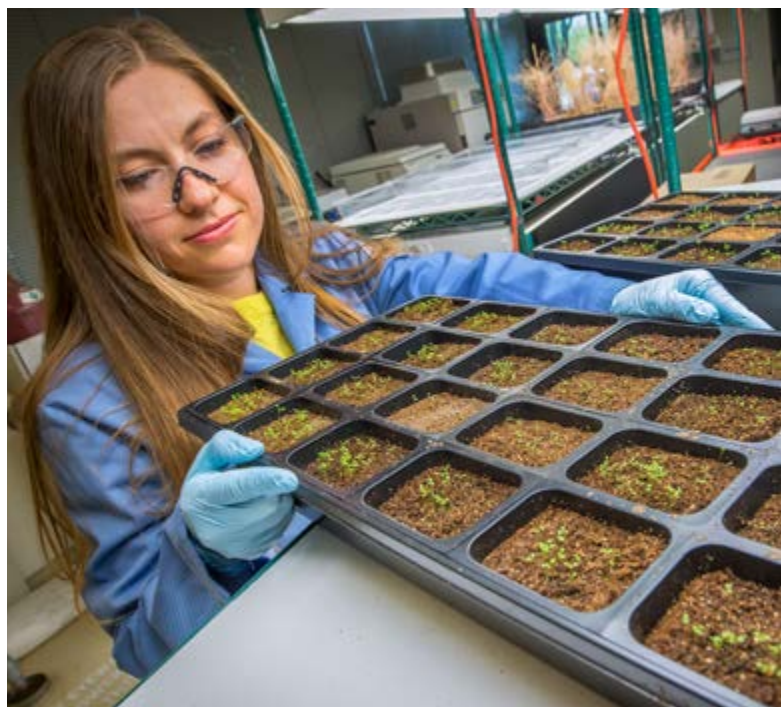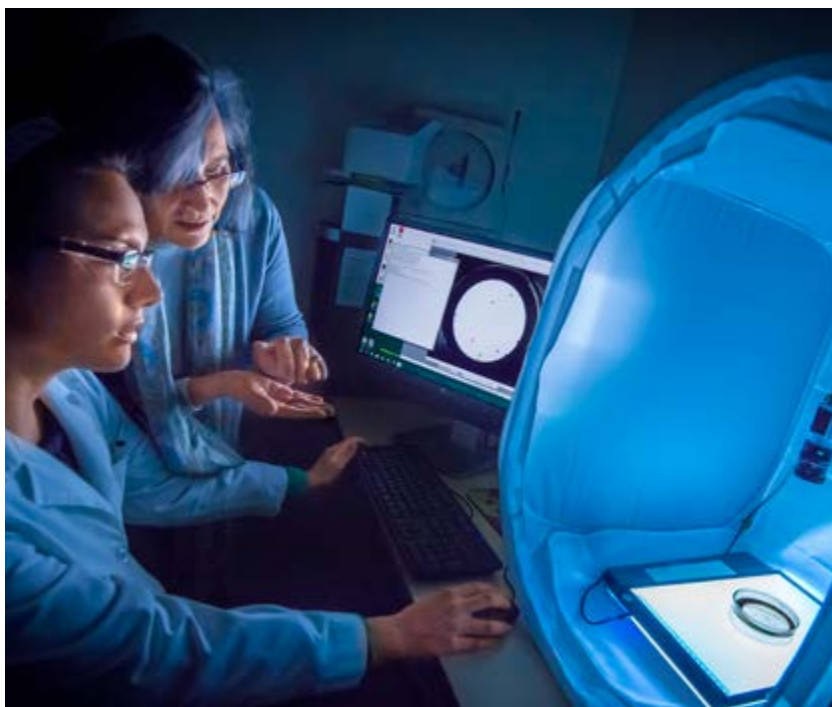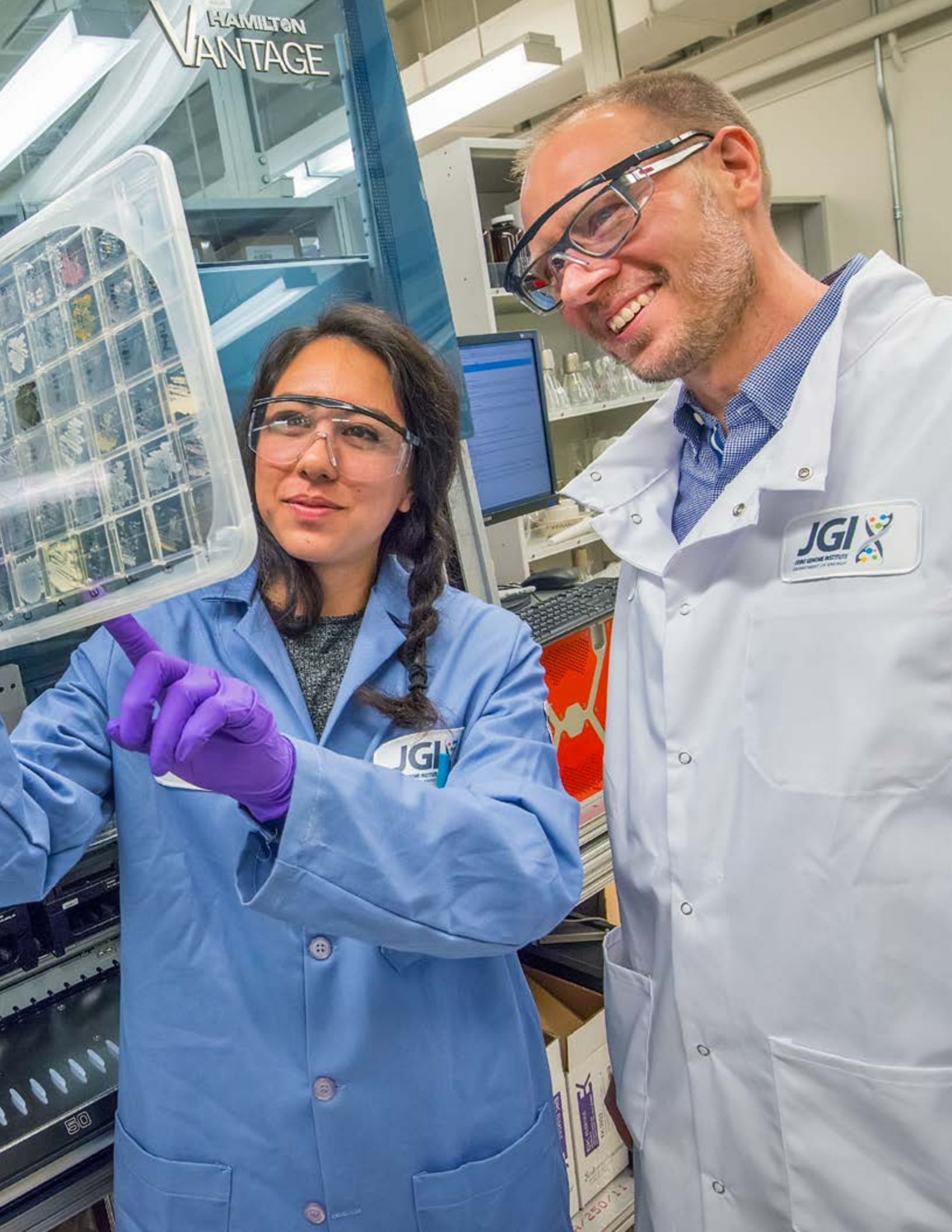
GLOSSARY

ABPDU	Advanced Biofuels and Bioproducts Process Demonstration Unit
ALS	Advanced Light Source
BioCAD	biological computer-aided design
CAM	computer-aided manufacturing
CNS	central nervous system
CO$_2$	carbon dioxide
Cryo-EM	Cryo-electron microscopy
CT	computed tomography
DOE	Department of Energy
DOE SC	Department of Energy's Office of Science
EcoFAB	fabricated ecosystem
EESA	Earth and Environmental Sciences Area
ENIGMA	Ecosystems and Networks Integrated with Genes and Molecular Assemblies
ESnet	Energy Sciences Network
JBEI	Joint BioEnergy Institute
JCAP	Joint Center for Artificial Photosynthesis
JGI	Joint Genome Institute
KBase	Systems Biology Knowledgebase
MX	Macromolecular crystallography
NERSC	National Energy Research Scientific Computing Center
PET	Positron emission tomography
SAXS	small-angle X-ray scattering
UC	University of California