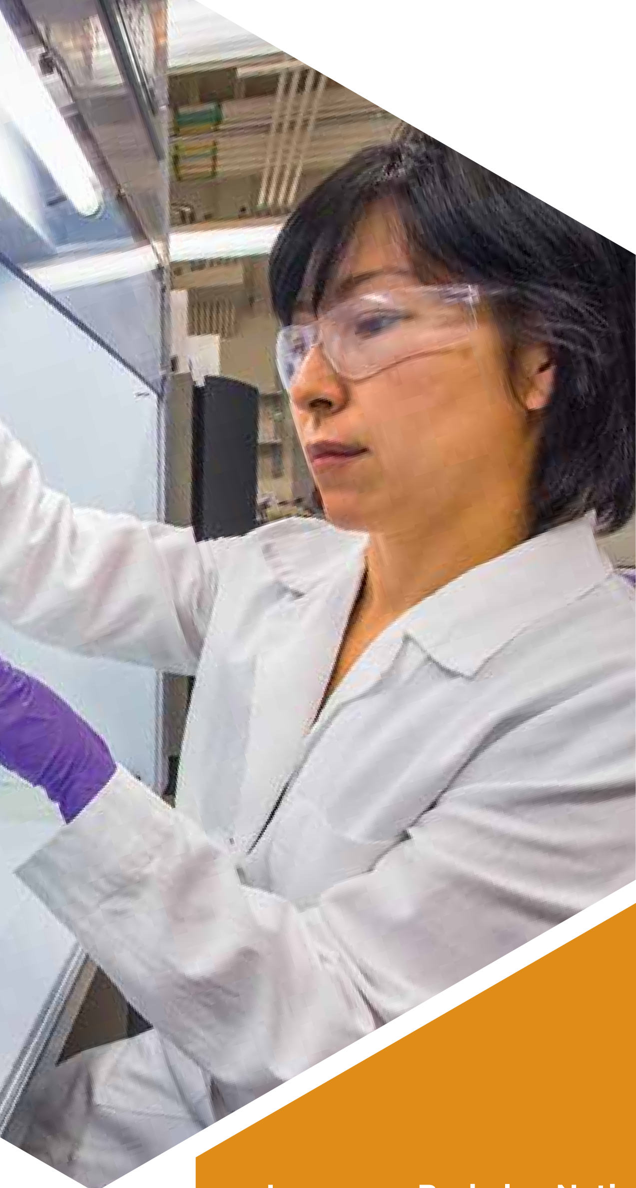


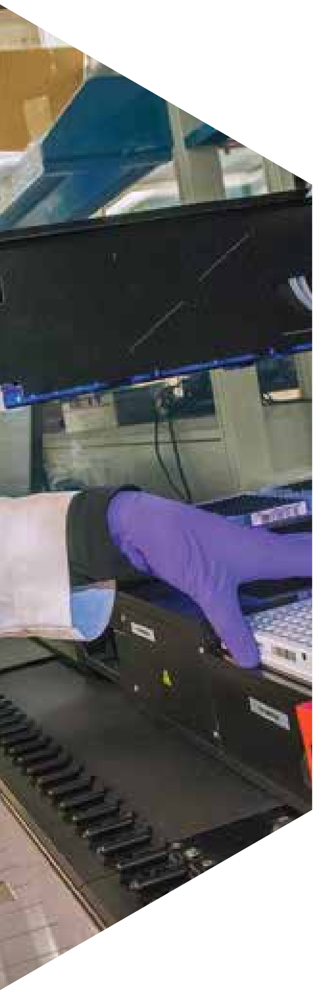
Lawrence Berkeley National Laboratory

# BIOSCIENCES 10-YEAR SCIENTIFIC STRATEGIC PLAN **2013–2023**

2014

BioSciences


JGI
ALBA



Lawrence Berkeley National Laboratory

# BIOSCIENCES 10-YEAR SCIENTIFIC STRATEGIC PLAN **2013-2023**

## REVISION 2014

BioSciences

# INTRODUCTION

As Associate Laboratory Director for Biosciences, I am fortunate to lead a part of Berkeley Lab that has 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 bioscience capabilities, teams, and approaches to address global-scale challenges.

Biological research is changing dramatically as a consequence of the development of powerful technologies along with new concepts and methods derived from integration of physical sciences, mathematics, computational sciences, and engineering. Advances in biological sciences hold tremendous promise for realizing solutions for the many major problems confronting the nation and the world, such as securing access to affordable, sustainable sources of energy, as well as 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 unique assets, and our ability to work effectively in multidisciplinary teams.

In developing this *10-year Scientific 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, 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.

In this 2014 version of the 2013 plan, our vision has been updated, refined, and pared down slightly to maximize success of achieving the 10-year goals, in response to recommendations made by the Biosciences Expert Advisory Committee, our group of external advisors. Also in this 2014 version, we have articulated our new cross-cutting initiatives (Pages 40 & 41) that integrate our research teams in energy, environment, health, and biomanufacturing in new ways.

To lead the implementation of the Biosciences Strategic Plan, a group of talented, collaborative researchers has been selected, and they contributed significantly to this updated version of our plan. For Energy, they are Henrik Scheller, Chia-Lin Wei, and Ian Sharp. For Environment, they are Harry Beller, Peter Nico, and Charles Koven. For Health, they are Susan Celniker, Mark LaBarge, and Trent Northen, and for Biomanufacturing, they are Nathan Hillson, Sarah Richardson, and Caroline Ajo-Franklin.

To enhance collaboration among Bioscience research teams and to maximize opportunities for successful achievement of our goals, we plan to relocate all Biosciences researchers to the Berkeley Lab campus. The first building, the Integrative Genomics Building, is proposed to house the DOE Joint Genome Institute and the Systems Biology Knowledgebase and will bring together these two complementary and synergistic DOE programs.

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 develop 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 2014

Lawrence Berkeley National Laboratory's (Berkeley Lab's) *Biosciences 10-Year Scientific Strategic Plan* describes the vision for a national future strengthened by biological research achievements, and provides guidance for biosciences research activities at the Laboratory. 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. Our plan lays out 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. In this 2014 version of the 2013 plan, our vision has been updated, refined, and pared down slightly to maximize success in achieving our 10-year Goals, as a response to recommendations made by the Biosciences Expert Advisory Committee, our group of external advisors. Also in the past year, we have selected a group of talented collaborative researchers to lead the implementation of activities aimed at achieving our 10-year Goals. Also in this 2014 version, we have articulated our new cross-cutting initiatives (Pages 40 & 41) that integrate our research teams in energy, environment, health, and biomanufacturing in new ways.

# 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 Scientific Strategic Plan* for biosciences 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 and cost-effectively transform the energy in sunlight and readily available $CO_2$ into liquid transportation fuels using biological or bio-inspired approaches?

### Environment Research

What will be the impact of climate change and land management on environmental systems, and what will be the implications for water resources and energy strategies? Can we manipulate environmental systems to remediate contaminants, sequester carbon, and support agricultural productivity?

### Health Research

How does the environment control the health of complex biological systems? What is the impact of environmental changes and exposures, particularly caused by anthropogenic impacts and perturbations, on biological systems and human health? Can we use this knowledge 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 energy, health, and environment challenges 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 the Laboratory 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, computations, 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 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 sources of energy and new materials that require less energy to produce and that 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 new opportunities 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, environment, and health. A significant impediment to progress has been the distribution of biosciences effort among five sites in Berkeley, Emeryville, and Walnut Creek. Accordingly, an important element of this 10-year *Strategic Plan* is to consolidate all biosciences activities at a single site.

While basic biological research is the foundation for future technologies, it also provides the evidence for informed policy making on topics of critical importance. Results of biosciences research efforts can provide guidance for decisionmakers 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 is sensitive to security and the implications of human genomic sequencing on privacy and predictions; on the consequences of biomanufacturing innovations and who controls production; on lessening potential volatility of production; and on how innovations disrupt some professions and create new ones. We pride ourselves on building trust with the public through transparency and direct conversation.

# BIOSCIENCES AT BERKELEY LAB

While Lawrence Berkeley National Laboratory may be best known for its physics, the biological sciences have been a 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 by-products 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 current 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 pathbreaking 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 (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 an 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 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 (JBEI) is one of three national centers created by 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 to 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 plan.

# 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


ENVIRONMENT
ENERGY
BIOMANUFACTURING

# STRATEGIC GOALS

## Biosciences for Energy

Develop cost-competitive biological and bio-inspired energy solutions capable of reducing U.S. dependence on fossil fuels.

## Biosciences for the Environment

Develop predictive understanding of ecosystem processes and their interactions with climate change and land use to improve environmental quality and the efficient use of resources.

## Biosciences for Health

Develop and apply a predictive, multiscale, integrative understanding of biological responses to environmental challenges that will improve human and biosphere health, and drive economic growth.

## Biosciences for Biomanufacturing

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

# BIOSCIENCES FOR ENERGY

## 10-year Goal

Develop cost-competitive biological and bio-inspired energy solutions capable of reducing U.S. dependence on fossil fuels.

## 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, we have begun to tap less-desirable sources 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. We believe that successful production of scalable alternative sources of transportation fuel will be achieved most rapidly through an integrated, team-science approach that has been the hallmark of the Laboratory since its inception.

Biology has the potential to produce energy renewably, particularly liquid hydrocarbon fuels with the high energy density needed by our transportation infrastructure. However, biological mechanisms are relatively inefficient at capturing sunlight energy and transforming it to hydrocarbons. A better understanding of photosynthesis, cell-wall synthesis in plants, and the hydrocarbon-forming reactions in all organisms would make it possible to build predictive models of these processes. These models could be used to engineer plants and algae to capture sunlight more efficiently, to engineer plants to accumulate sugars in cell walls that are more easily digested, and to engineer

10-YEAR
### Energy Goal

Develop cost-competitive biological and bio-inspired energy solutions capable of reducing U.S. dependence on fossil fuels.

### Energy Research Strategies to Achieve Goal

**Cellulosic biofuels.** Derive energy from biomass with new technologies.

**Microbial biofuels.** Engineer photosynthetic microorganisms to produce fuels directly from sunlight, $CO_2$, and water.

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

microorganisms to convert sugars more efficiently into drop-in biofuels compatible with the transportation fleets of today and tomorrow. 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 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 cellulosic 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 deconstruction, and hydrocarbon biosynthesis (see *Fuels from Photosynthesis*, Page 40). We can then harness this knowledge to meet key strategic objectives in the areas of cellulosic biofuels production, microbial biofuels development, and artificial photosynthesis.

## Energy Research Strategies

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

- **Microbial biofuels.** Engineer photosynthetic microorganisms to produce fuels directly from sunlight, $CO_2$, and water.

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

To achieve the Energy Goal by 2023, our approach employs three areas of strategic focus: production of fuels from cellulosic biomass (cellulosic biofuels), microbial production of fuels directly from sunlight and $CO_2$ (microbial biofuels), and nonbiological production of liquid fuels directly from sunlight and $CO_2$ (artificial photosynthesis). These areas were chosen because (1) we believe 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 at Berkeley Lab. These strategies will be executed in parallel.

### **Cellulosic 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 tolerance to stress.

- Greatly improved biomass extraction and degradation strategies.

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

Efforts on these three fronts are under way at Berkeley Lab, primarily funded through JBEI and individual research programs funded by Biological and Environmental Research in the Office of Science at DOE, but supported also through collaborative interactions between JBEI, DOE JGI, and the DOE Systems Biology Knowledgebase (KBase) as well as the Advanced Biofuels Process Demonstration Unit 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 from these plants, and efficient biological conversion of the resulting sugars to advanced biofuels. This team of nearly 200 scientists has already engineered model plants with more cellulose (the source of fermentable sugar) and less 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.

DOE JGI's efforts have effectively complemented those of JBEI. 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 cellulosic biofuels strategy, providing the bases for molecular understandings of relevant pathways, and sources of molecular tools for engineering strategies. In addition, 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, 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 their 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, there remain many inefficiencies in converting sugars to drop-in biofuels.

### **Microbial biofuels:** Engineer photosynthetic organisms to directly produce fuels from sunlight, $CO_2$ and water

The goal of the cellulosic biofuels strategy is to produce fuels from simple sugars derived from plant cell walls. However, sugars derived from cellulosic biomass are complex and often contain contaminants that in the past have prevented them from being efficiently fermented by microorganisms. Therefore, in addition to feeding microorganisms sugars derived from cellulosic biomass to create fuels, Berkeley Lab is also focused on another pathway — the direct microbial conversion of sunlight and $CO_2$ into fuels.

The potential for direct conversion of sunlight to fuels in microorganisms has long been conceptually attractive, but major unsolved challenges continue to impede this effort.

Efforts to identify and exploit species that might be useful for this purpose have found limited success. Photosynthetic bacteria produce very limited amounts of lipids and while many species of microalgae do accumulate storage lipids, they do so only as a response to nutrient-stress that severely limits overall productivity. Efforts to engineer these organisms are hampered by a dearth of required tools with which to do the engineering and more importantly by a lack of basic understanding of the metabolic pathways involved and how they are regulated. This systems-level understanding is needed to carry out the large-scale pathway engineering that will ultimately be required for the development of production strains.

In pursuing this strategy, Berkeley Lab can leverage JBEI's significant experience with microbial fuel synthesis from sugar substrates. Similarly, the DOE JGI is a world leader in algal genomics and has begun a new program in microalgal epigenomics. Finally, the Laboratory's current genome editing and engineering will be further developed — as part of the Biomanufacturing effort, discussed later — to become critically important to the success of our 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 our 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 lack of a deep understanding of photosynthesis on temporal and spatial scales, have prevented an ability to design 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, new prospects for the control and manipulation of the intricate processes of artificial photosynthesis. New nanotechnologies allow Berkeley Lab researchers to work at the scale of atoms and molecules and their hierarchical assembly to functional structures; to synthesize new catalysts to accelerate relevant chemical reactions; and to mass-produce nano-engineered materials capable of absorbing light and $CO_2$ while wicking away newly created fuel for subsequent collection and storage.

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 of nanotechnology, drawing scientists from around the world to its world-class instruments, materials, technical expertise, and training. In addition, the Laboratory's Advanced Light Source and other spectroscopic facilities 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. This X-ray facility is 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. 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. Finally, the Laboratory has a dedicated team at its Joint Center for Artificial Photosynthesis, a DOE Energy Innovation Hub established in 2010 to develop a manufacturable solar-fuels generator made of Earth-abundant elements that will use only sunlight, water, and carbon dioxide as inputs to robustly produce fuel.

## 2023 10-year goal achievement measured by:

### Cellulosic biofuels

- Understand fundamental elements of plant biology that underlie biomass yields and adaptation to stress.
  - 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 fixation occurs in model legumes.
  - Understand the molecular interactions between plants and mycorrhizal fungi and their role in phosphate assimilation and stress tolerance.
  - Probe the variety of biomass-degradation strategies employed by microbes in natural environments.

- 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 nitrogen fixation into nonfixing model plants and crops to reduce dependency on energetically and environmentally costly nitrogen fertilizers.
  - Engineer plants and microbes to stimulate nitrogen fixation by nonleguminous plants.
  - Engineer model plants and crops to grow with 25% less water.
  - Engineer lignin in biomass crops to be an economically useful polymer.

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

- Develop deconstruction processes that produce cellulose and hemicellulose devoid of lignin contaminants.

- Develop inexpensive enzyme cocktails that effectively depolymerize cellulose, hemicellulose, and lignin without product inhibition and inhibition from contaminants resulting from the deconstruction process.

- Understand the synergies between bacteria and fungi in the breakdown of lignocellulose.

- Engineer fuel-producing microorganisms whose communities are resistant to invasion by contaminating organisms and that are genetically constrained to growth in industrially defined conditions.

- 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.
    - Engineer microbes to tolerate biomass deconstruction inhibitors and fuel products.
    - Engineer microorganisms to produce fuels under anaerobic conditions.

### Microbial biofuels

- Use a systems-level understanding of lipid synthesis and accumulation to engineer microalgae that produce storage lipids in the absence of nutrient stress.
    - Develop a systems-level understanding of the microalgal stress response.
    - Engineer regulators of the algal stress response to drive inducible lipid accumulation that does not require nutrient starvation and that minimizes interference with cell growth.
    - Engineer hydrocarbon fuels synthesis in a model alga.
    - Refactor storage-lipid and fuel pathways for transfer from a model to a production species.

- Increase photosynthetic efficiency of microalgae and cyanobacteria by 20%.
    - Provide a systems-level understanding of photosynthesis in cyanobacteria and microalgae.
    - Use an understanding of highly efficient microalgal photosynthesis to improve photosynthetic efficiency in cyanobacteria.

- Apply systems approaches to other phenotypic traits of interest to the algal biofuels community.
    - Engineer algae to use novel (cheaper) sources of nitrogen and phosphorus.
    - Engineer pumps to export hydrocarbons from the cell to avoid dewatering algae, one of the most expensive aspects of algal biofuel production.
    - Engineer salt-tolerant algae that will require less fresh water.

### Artificial photosynthesis

- Develop an artificial photosynthetic system that produces a liquid fuel from light energy and $CO_2$.

- 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 ALS-U.

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

- Incorporate methods for self-healing or repair to enhance longevity of artificial photosystems.

- Redeploy the photosynthetic apparatus in a previously nonphotosynthetic host.

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

## 2018 – Five-year milestones for energy research strategies

### Cellulosic biofuels

- Elucidated cell-wall biosynthesis and assembly through identification of new genes, alleles, and metabolic pathways controlling cell-wall recalcitrance, sugar and lignin content, and fermentation inhibitors.
  - Engineered secondary cell walls to have more C6 sugars and fewer C5 sugars.
  - Engineered cell wall to have easily cleavable lignin.
  - Engineered biomass traits from model plant systems to potential bioenergy crops.

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

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

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

- Engineered hydrocarbon biosynthetic pathways and associated transporters into microbes to convert sugars to transportation fuels.

- Discovered and engineered new enzymes and cofactors for biomass deconstruction that are tolerant of pretreatment regimens.

- Discovered efficient enzymes capable of depolymerizing lignin into aromatic hydrocarbons.

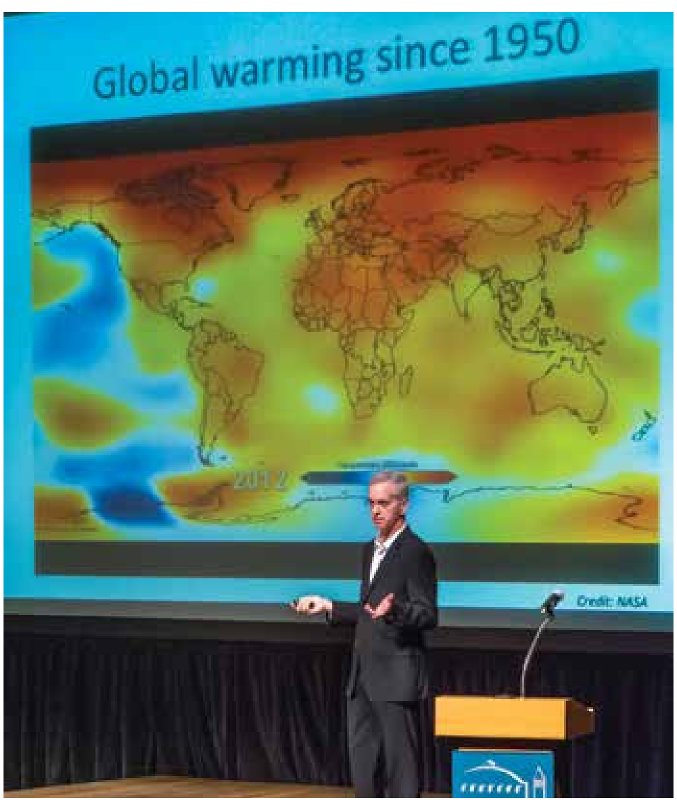
- Described native hydrocarbon biosynthetic pathways in plants and microbes.

- Developed models to predict how modifications to secondary cell-wall biosynthesis and degradation improve biomass yields.

- Developed predictive models to describe release of sugars from the plant secondary cell wall.
- Developed predictive models to describe metabolic fluxes and used them to predict bottlenecks in biosynthetic pathways in microorganisms.
- Developed retrosynthesis software to predict all possible pathways to a hydrocarbon product.
- Advanced knowledge of drought tolerance in model crops through identification of new genes, alleles, and metabolic pathways.

### Microbial biofuels

- Systems-level understanding of the algal stress response used to engineer fuel production in algae.
  - Identified master regulators controlling storage-lipid accumulation in response to nutrient stress in the model algae *Chlamydomonas reinhardtii*.
  - Produced testable computational models of the transcriptional regulatory network that controls lipid accumulation in *C. reinhardtii*.
  - Developed CRISPR/Cas9-based methods for genome engineering in algae.
  - Generated a complete inventory of lipid intermediates and biosynthetic pathways in *C. reinhardtii*.
  - Discovered novel metabolic pathways and enzymes for hydrocarbon production.




- Used a systems-level understanding of photosynthesis to improve efficiency of microalgae and cyanobacteria.
  - Used high-throughput mutagenesis to provide a complete understanding of photosynthetic pathways in microalgae and cyanobacteria.
  - Employed mathematical models to predict how to increase efficiency in photosynthetic organisms.
  - Used genome-editing tools to test these predictions.

- Explored ancillary traits of interest to the algal biofuels industry.
  - Identified the components of nitrogen, sulfur, and phosphate sensing in algae.
  - Described mechanisms of salt tolerance in algae.
  - Developed strategies for hydrocarbon export from algae.

### 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 to the extent that these can be recapitulated or improved in nonbiological photosystems.

- Developed mathematical models of photosynthetic processes to gain a better understanding of the limits of photosynthetic efficiency.

- Explored options for reaching $mA/cm^2$ 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.

- Performed 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.

- Expanded the solar fuels R&D network and establish collaborations to begin creating a clearinghouse for integrating and benchmarking components that can be used in the solar fuels system.

# BIOSCIENCES FOR THE ENVIRONMENT

### 10-year Goal

Develop predictive understanding of ecosystem processes and their interactions with climate change and land use to improve environmental quality and the efficient use of natural resources.

### Background and Motivation

Although Earth's environment provides global natural resources of critical importance for sustaining life on the planet, we understand little about the complex physical, chemical, and biological interactions that occur within its ecosystems. These interactions regulate the geochemical flux of most life-critical elements, control the production of food and renewable energy products, and purify water for much of the biosphere. They regulate atmospheric greenhouse gases — potentially amplifying or attenuating global changes in temperature and water levels. The interactions and feedbacks among humans, plants, microbes, and the environment ultimately shape the world in which we live. However, a growing human population whose numbers and lifestyles drive an ever-increasing demand for resources — including clean water, food, land, and energy — is currently reshaping these interactions on a global scale. Higher temperatures associated with climate change will be accompanied by unprecedented changes in the water supplies, natural resources, and agricultural and ecological systems that are critical to the nation. Many consequences have already emerged, such as increased average temperatures; melting of ice caps; sea level rise; permafrost thaw; and increased incidences of extreme weather, droughts, wildfires, and pest invasions. Land-use and water-use conflicts increase, and energy production (including hydraulic fracturing and bioenergy) threatens both water quality and sustainable water resources. The urgency of developing scientific approaches to secure energy while maintaining or improving environmental quality is becoming increasingly evident.

10-YEAR
## Environment Goal

Develop predictive understanding of ecosystem processes and their interactions with climate change and land use to improve environmental quality and the efficient use of resources.

---

### Environment Research Strategies to Achieve Goal

**Metabolic potential of natural systems.** Quantify and harness the metabolic potential of natural systems, including specific microbially catalyzed processes and biotic (microbe-microbe, microbe-plant) interactions that modulate metabolism under a variety of environmental conditions.

**Biogeochemical cycles and their controls.** Determine abiotic and biotic controls on biogeochemical cycles in representative terrestrial systems and associated feedbacks with climate and land use.

**Climate and environmental change.** Advance process knowledge and computational capabilities that enable accurate predictions of integrated atmospheric system dynamics and changes over time. Use capabilities to guide development of climate-resilient energy and environmental solutions.

To address such 21st century scientific challenges, it is imperative to consider the complexity of the natural environment as an integrated system, which necessitates quantification of biological, geochemical, hydrological, and atmospheric processes and interactions that occur across great time and space scales (see *Microbes to Biomes*, Page 40). Central to this challenge is the development of a quantitative and predictive understanding of how biological components influence the planet's life-support system and how such influence is expected to evolve with climate and land-use changes. For example, microbial and plant communities and their interactions control the production of food and biofuel feedstocks and regulate the flux of major greenhouse gases to the atmosphere. However, an accurate understanding of how these communities interact and function in dynamic environmental systems is missing — under current and global change conditions, in response to humans' direct perturbations, and up to scales where environmental or energy systems are managed (such as watersheds). Accurate prediction of global climate changes that drive these ecosystem responses requires consideration of induced hydrological extreme events and the transport of the dozens of biogeochemical species that interact with biotic and abiotic components of the ecosystem and across different scales of resolution, from molecules to the watershed scale.

Berkeley Lab brings extensive expertise to the environmental component of the Biosciences Strategic Plan, including a longstanding reputation for pathbreaking work in molecular environmental microbiology, soil and subsurface biogeochemistry, hydrogeology and geophysics, and climate sciences. Our scientists draw on this expertise, as well as on the resources available at computational centers and user facilities such as DOE JGI and the Advanced Light Source.

Our research strategies are aimed to address first-order environmental challenges, such as the extent to which terrestrial environments can continue to serve as greenhouse gas sinks over the coming centuries, how these environments attenuate contaminants, how climate change will impact energy production and water resources, how to achieve mineral and food security in the face of rising population growth, and the extent to which microbe-plant interactions co-evolve under global change conditions.

## Environment Research Strategies

- **Metabolic potential of natural systems.** Quantify and harness the metabolic potential of natural systems, including specific microbially catalyzed processes and biotic (microbe-microbe, microbe-plant) interactions that modulate metabolism under a variety of environmental conditions.

- **Biogeochemical cycles and their controls.** Determine abiotic and biotic controls on biogeochemical cycles in representative terrestrial systems and associated feedbacks with climate and land use.

- **Climate and Environmental Change.** Advance process knowledge and computational capabilities that enable accurate predictions of integrated atmospheric system dynamics and changes over time. Use capabilities to guide development of climate-resilient energy and environmental solutions.

To achieve the Environment Goal by 2023, our research strategies focus on advancing predictive understanding along three themes: (1) metabolic potential of natural systems, (2) biogeochemical cycles and their controls, and (3) climate and environmental change.

We believe that advancing and combining these elements are necessary to develop new classes of environmental and energy solutions. We also believe that 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. Environmental bioscience contributions to the Berkeley Lab Microbes to Biomes Initiative (see *Microbes to Biomes*, Page 40) is an example of a grand challenge that requires advances in and combining of the three Environmental strategies described below.

**Metabolic potential of natural systems:** Quantify and harness the metabolic potential of natural systems, including specific microbially catalyzed processes and biotic (microbe-microbe, microbe-plant) interactions that modulate metabolism under a variety of environmental conditions.

Microbes and plants in terrestrial environments carry out a wide range of key biogeochemical processes. Microbial and plant communities are extremely sensitive and reactive to environmental stresses, and their responses in turn impact the environment. Despite their critical role, very little is known about the manner in which microbial communities are organized, the primary determinants of how this organization adapts and evolves, how dynamic such populations are, and the nature of interactions between the microbes and their physical/chemical environment. Not only is the ability to predict the microbial metabolic potential extremely limited, but microbe-plant interactions under changing environmental conditions are not well understood, including changes in water stress, nutrient availability, temperature, or $CO_2$. This lack of understanding limits the ability to manipulate organisms and environmental systems for environmental or energy benefits.

New developments in DNA sequencing, bioinformatics, and systems biology have made it more tractable to quantify how complex plant-microbe communities function and interact in situ in terrestrial environments. This new era of genomics analysis amplifies the possibility of quantifying how information stored in genomes may be translated into predicting the biogeochemical functioning of larger systems.

Our efforts to understand the metabolic potential of natural microbial systems are multifaceted and interconnected. One approach focuses on quantifying microbial metabolic potential in dynamic soil and subsurface systems and on development of models that predict the influence of microbial community interactions on integrated system behavior. Another seeks to advance increasingly high-resolution multiscale imaging of molecules and microbes for the discovery of mechanisms underlying the subcellular, cellular, and intercellular networks of metabolite, protein, RNA, and DNA molecules that drive macroscopic biogeochemical processes. Advances in environmental biosensing (see *Biosensors for Environment and Health,* Page 41) and the Berkeley Synchrotron Infrared Structural Biology Program augment these efforts through real-time imaging of microbes under environmentally relevant conditions. The Systems Biology Knowledgebase (KBase) provides a data and computing resource to enable 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. This strategy is also supported by DOE JGI's vision as a next-generation genome science facility through its commitment to sequencing and decoding of the unexplored “dark matter” of microbial genomes, and its ability to rapidly synthesize and assemble this genetic material into tame laboratory organisms and characterize resultant activities. Starting with solid successes in determining metabolic potential of single microbial isolates and simple microbial communities, Berkeley Lab will broaden to include explorations of the metabolic potential of more complex microbial communities and microbe-plant associations.

**Biogeochemical cycles and their controls:** Determine abiotic and biotic controls on biogeochemical cycles in representative terrestrial systems and associated feedbacks with climate and land use.

Of the components of the Earth system that contribute to biogeochemical cycling, the terrestrial environment is perhaps the most complex, as it hosts a multitude of interactions and processes among plants, animals, microorganisms, minerals, migrating fluids, and dissolved constituents. The cycling of carbon, nutrients, and water is driven by plant productivity, microbial metabolism, and interactions of elements with climate, the soil matrix, and the subsurface. These processes interact within a heterogeneous physical framework and over scales ranging from nanometers to kilometers. These processes are influenced by global atmospheric change as well as by anthropogenic land and resource use and other impacts of economic activities. Although challenging, developing a robust understanding of terrestrial biogeochemical cycling is critical for developing new solutions for sustainable water resources, water quality, agricultural production, and energy production.

Berkeley Lab is well-positioned to explore these challenges through a diverse array of approaches: quantifying key controls on deep soil carbon turnover through in situ sensing (see *Biosensors for Environment and Health*, Page 41); manipulation experiments (warming, moisture, $CO_2$); measuring ecosystem $CO_2$ fluxes at AmeriFlux and other field study sites set up at 100+ locations across the nation; developing process knowledge and simulation capabilities to quantify coupled hydrological-biogeochemical fluxes and transformations in watershed systems and associated with energy production activities; and through the use of laboratory and field approaches to noninvasively characterizing controls on biogeochemical cycling through the use of laboratory and field studies.

**Climate and environmental change:** Advance process knowledge and computational capabilities that enable accurate predictions of integrated atmospheric system dynamics and changes over time. Use capabilities to guide development of climate-resilient energy and environmental solutions.

One of the most urgent challenges facing the world today is ensuring adequate water supply and quality to meet human and ecosystem needs. Climate change is expected to lead to unprecedented changes in the water supply, including flood, drought, and heat waves. Furthermore, all aspects of biogeochemical cycles, from food, fuel, and pharma production to the distribution of biomes on the planet, will be altered by climate change, and lead to climate feedbacks that are highly uncertain. Reliable predictions of environmental and climate change are needed to guide the development of adaptation and mitigation strategies relevant to water, food, health, and energy.

Many scientific and technological gaps prohibit development of a predictive understanding of climate and environmental change. Examples of such gaps include the behavior of the integrated system spanning both spatial scales (e.g., from microbes to biomes; see *Microbes to Biomes,* Page 40) and timescales (e.g., from fast fluxes to slower ecosystems interactions); between ecosystems and climate; limitations in the resolution and accuracy of models used to predict change; and a limited ability to integrate an understanding of Earth and human systems as needed to develop sustainable management of the Earth's resources.

Berkeley Lab teams are working to provide advanced capabilities to project climate change, to predict the influence of such change on biogeochemical functioning, to detect and attribute climate extremes, to simulate climate-water-energy system interactions, and to quantify cloud dynamics. This strategy is greatly supported by the Laboratory's high-performance computing capabilities.

## 2023 10-year goal achievement measured by:

The 10-year Environment goals strive 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 will require joint consideration of all three Environment Strategies. Many goals will be met through investigations carried out at environmentally relevant field-study sites and through synthesis of cross-site data sets.

### Metabolic potential of natural systems

- Through discovery of novel genes, enzymes, syntrophic associations, and/or microbe-abiotic interactions, develop a mechanistic understanding of at least two biochemical/biogeochemical processes of clear environmental importance, including C-N-S-P cycling or contaminant fate.

- Develop a robust approach for gene or metabolic pathway discovery that could be applicable across a wide range of environmental processes by combining biological (e.g., omic) data, chemical/geochemical data, modeling, and informatics methods from field and laboratory experiments to infer metabolic pathways in complex systems. In addition, include high-throughput reductionist approaches to functional genomics as applied to model and complex systems to discover novel gene functions of importance for key biosystem processes.

- Discover how the metabolic lifestyles/pathways of key microorganisms modulate in response to environmental changes, as wells as how communities adapt and evolve as a whole.

- Develop strategies to manipulate the environment or metabolic potential for beneficial purposes, such as soil carbon fixation, contaminant remediation, and increased bioavailability of critical nutrients for food and fuel crops.

### Biogeochemical cycles and their controls

- Discover controls on soil carbon turnover and the contributions of deeper soils, the vadose zone, groundwater, and hyporheic zone processes to "whole system" carbon cycling in select yet representative managed and natural ecosystems.

- Discover how and to what extent information stored in microbial genomes influences biogeochemical cycling and thus system functioning at the larger terrestrial environment scale.
- Predict the effect of coupled hydrological-biogeochemical fluxes associated with global change or land use, such as how climate-induced hydrological changes within a watershed, or growing biofuel crops on degraded lands, affect biogeochemical cycling in those environments.
- Use predictive understandings to develop strategies to remediate and sustain terrestrial environments and their services.

### Climate and environmental change

- Develop robust methods for detecting and attributing climate extremes.
- Identify and predict the onset of tipping points in terrestrial systems and regional-scale climate phenomena that lead to new environmental states.
- Predict the ability of portions of the Arctic permafrost, temperate, and tropical forest systems to continue to serve as carbon sinks over the next century.
- Develop energy and agricultural strategies that perform well under water- and land-use stresses.
- Develop adaptive responses, including to extreme and abrupt climate changes, that increase robustness of natural systems and protect lives and infrastructure.

## 2018 – Five-year milestones for environment research strategies

Near-term efforts largely consist of advancing understanding as well as measurement and predictive capabilities targeted at key components of the system. Within this time frame, many advances associated with the metabolic potential and biogeochemical cycling strategies will be achieved through conducting experiments or assessing samples collected from natural and relevant field-study sites in the Arctic as well as arid and tropical environments. These sites are natural laboratories to bring together various kinds of expertise as needed to understand the biogeochemical, plant, microbial, and hydrological processes involved in elemental and water cycling and ecosystem behavior under varied environmental conditions.

### Metabolic potential of natural systems

- Developed and applied new state-of-the-art imaging technologies to quantify how physical, chemical, and microbiological processes interact in soils to influence C and N cycling, from the nanometer to aggregate scales.
- Elucidated the metabolic lifestyles of soil and subsurface microbial community members and the functions of specific genes thought to be important for community metabolism through a combination of systems biology, physiological analysis, functional genomics, and high-throughput microbial isolation.

- Identified a major knowledge gap in the understanding of a microbial metabolic process that has significant environmental consequences (such as microbial degradation of macromolecular forms of carbon; microbial P solubilization in the rhizosphere; linkages between C, S, N, and/or P cycles; and poorly understood chemolithotrophic processes) and designed focused experiments or projects to fill that gap.

- Developed and applied high-throughput approaches to expand our understanding of novel enzyme families in natural systems.

### Biogeochemical cycles and their controls

Predicted how metabolic potential and biogeochemical cycling vary with changing environmental conditions (i.e., temperature, moisture) at two representative field study sites.

- Discovered key controls on soil and deeper subsurface carbon mobility under varied hydrological and temperature conditions through targeted in situ and ex situ manipulation experiments conducted within two representative ecosystems.

- Discovered terrestrial controls of geomorphology, vegetation, and hydrogeochemistry on the spatial distribution of microbial communities and their propensities to respond to environmental stresses at two relevant field-study sites.

- Made significant advances in measuring and predicting the controls on plant growth and the terrestrial carbon sink, including effects of rising atmospheric $CO_2$ in old-growth tropical forests.

- Developed new approaches for remotely characterizing and monitoring in situ biogeochemical transitions or integrated system behavior using geophysical and isotope methods.

- Developed a “data ecosystem” computational infrastructure that enabled integration of microbial and plant systems biology modeling with macroscale environmental system data sets.

### Climate and environmental change

- Advanced and coupled mechanistic component models, including soil/subsurface models, ecosystem demography models, and multiscale or global climate models.

- Achieved quantified reduction of uncertainty in Earth system model predictions of the sensitivity of climate to greenhouse gas emissions, including physical and biogeochemical feedbacks that operate on multiple spatial and timescales.

- Advanced ability to perform analysis of climate change through developing exascale solutions to climate’s big-data challenges.

- Accelerated ability to explore sustainable energy options through development of computational capabilities that jointly consider climate-water-energy systems.

# 2014 CROSS-CUTTING INITIATIVES

## Microbes to Biomes

Microbes are the most abundant and diverse forms of life on Earth, occupying every living system in our biosphere. The interactions of microbes with one another and with their environment are critical to the health and well-being of their host biome, whether in an aquifer, a cropland, or the human body. How do microbial communities interact with each other and their environment from the molecular to the biome scale? How do external factors, such as disease or climate change, influence these interactions and outcomes? Through taking advantage of cross-laboratory expertise and world-class facilities, researchers contributing to Berkeley Lab's Microbes to Biomes Initiative are seeking to understand relationships between microbes and their host biomes that are vital to our planet's future.

## Fuels from Photosynthesis

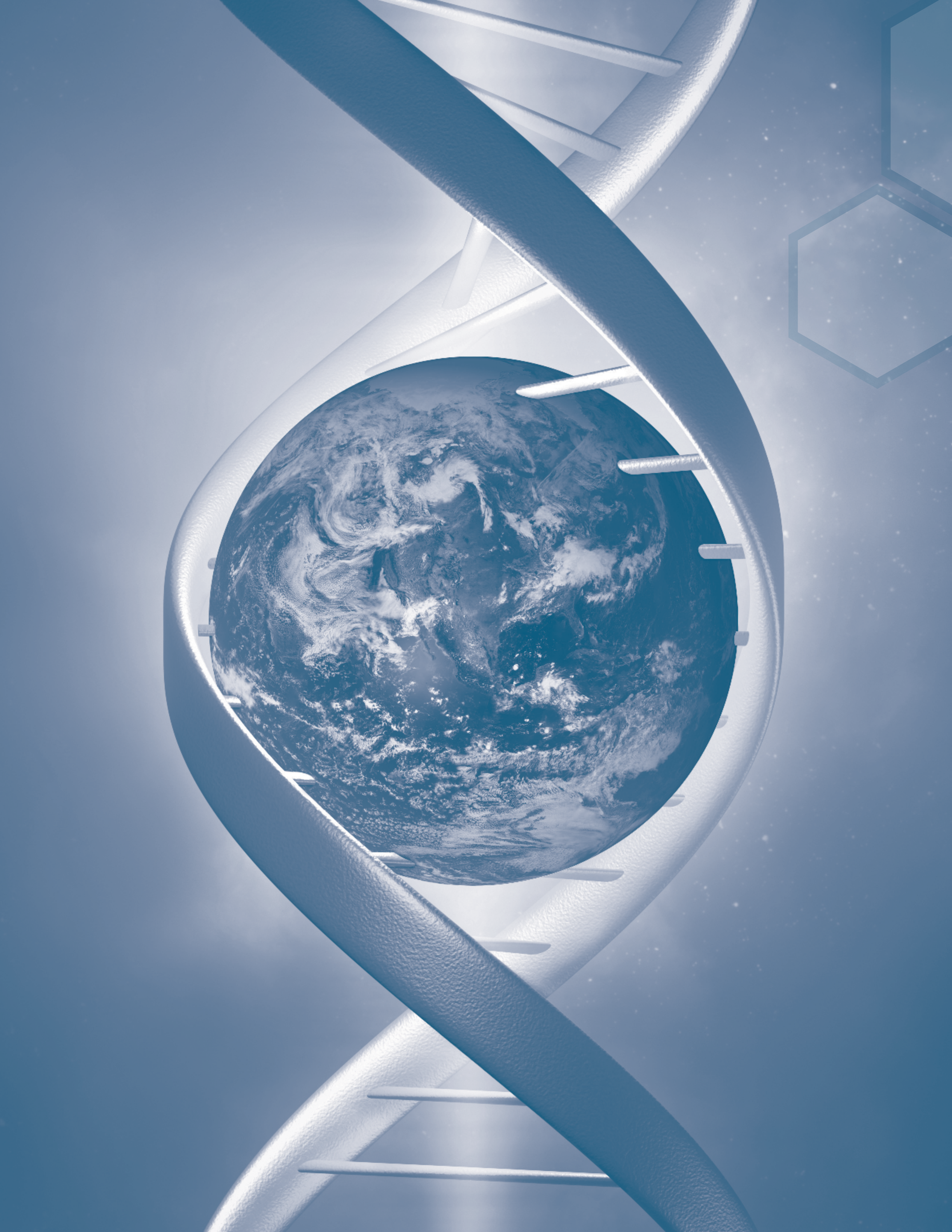
The unifying theme of Berkeley Lab's energy strategy is the desire to convert solar energy to transportation fuels. Plants, microalgae, and cyanobacteria differ in important details of photosynthesis, including efficiency of light harvest, energy transfer, and downstream metabolic events, and we want to understand these details in order to engineer and improve fuel-producing processes in each. This desire extends beyond the plants and microbes themselves — the Laboratory is uniquely positioned to couple its strength in biology with similar strengths in engineering and materials science to create bio-inspired or bio-hybrid photosynthetic fuel cells.

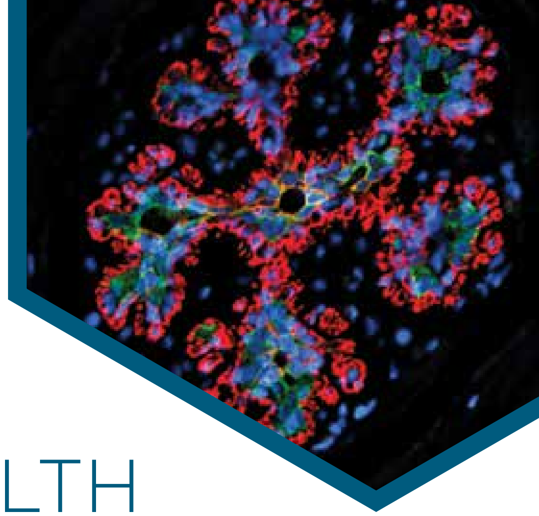
## Advanced Biomanufacturing Initiative

Microbes and plants represent an incredible diversity of life and contain a majority of the genetic potential of the Earth. The biomanufacturing and agricultural Industries have taken advantage of their extraordinary abilities to create biofuels, lubricants, pharmaceuticals, food polymers, and other materials in increasingly efficient and renewable ways. Biomanufacturing alone already accounts for more than $190 billion in U.S. revenues each year. Yet only a small fraction of the possible applications of biomanufacturing have been achieved because design and implementation of new function in plants and microbes is still inefficient. However, thanks in part to key innovations by Berkeley Lab scientists, the technologies for designing, building, testing, and optimizing such systems are evolving rapidly. The Advanced Biomanufacturing Initiative seeks to build a center of excellence — an open facility for creative, collaborative research — to transform the manufacturing process using synthetic biology and machine learning. This center would supports an ever-improving infrastructure that vastly lowers the barriers to harnessing biological function for applications in energy, the environment, and health.

## Biosensors for Environment and Health

An understanding of how environmental challenges impact biological systems requires an ability to accurately measure dynamic responses in natural, living systems at multiple levels — molecules, cells, and organisms. To enable this, Berkeley Lab plans to develop next-generation biosensors that, with systems biology approaches through KBase and modeling and simulation with Berkeley Lab's Computing Sciences, will provide real-time assessment of molecular changes and predictions of possible implications of those changes. For example, biosensors that quantitate levels of metabolites, or individual bacterial species, will help elucidate how soil microbes respond to biogeochemical changes brought about by climate-change land use. Similarly, sensors deployed in the guts of eukaryotic model systems or humans could be used to measure and understand how host cells and the resident microbiome respond to environmental stresses. Such an endeavor requires multidisciplinary strengths found at Berkeley Lab — materials science, engineering, chemistry, and computing — and promises to provide new avenues of bioscience exploration currently not possible.





# BIOSCIENCES FOR HEALTH

## 10-year Goal

Develop and apply a predictive, multiscale, integrative understanding of biological responses to environmental challenges that will improve human and biosphere health, and drive economic growth.

## Background and Motivation

Organisms have complex responses to natural and anthropogenic changes or challenges in their environment. External factors that can influence biological health include diet; temperature and climate; water and air quality; time; and exposure to chemicals, radiation, nanomaterials, and energy production byproducts. Most chronic diseases — including cancer and cardiovascular and neurodegenerative disease — are caused by adverse gene-environment interactions, with the environmental component playing a major role. Environmental factors mediate their effects by altering molecules, cells, and physiological processes inside organisms; this internal chemical milieu continually fluctuates during life due to changes in external and internal sources, so social and environmental factors must be considered when evaluating impacts on fitness and disease.

Berkeley Lab aims to address scientific and societal challenges in a comprehensive and sustainable 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. The ability to assess the genetic contributions to health and disease, the development of technologies for quantifying the presence and amounts of biomolecules (omics) (see *Biosensors for Environment and Health*, Page 41), and computational methods for integrating diverse and large data sets have blossomed in recent decades. However, quantitative assessment of the

### 10-YEAR Health Goal

Develop and apply a predictive, multiscale, integrative understanding of biological responses to environmental challenges that will improve human and biosphere health, and drive economic growth.

### Health Research Strategies to Achieve Goal

**Biological responses to environmental challenges.** Develop and deploy model systems to study molecular, cellular, and organismal responses to environmental challenges.

**Impact of environmental challenges on human biology.** Build platforms that rapidly elucidate the immediate and long-term consequences of environmental challenges on human health and disease, and develop innovative bio-solutions.

**Technologies to accurately assess the impact of environmental challenges on biological systems.** Advance and integrate bioimaging, rapid functional analyses, and computational approaches to measure biological responses to environmental factors.

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. 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 the biosphere. This understanding will allow us to balance economic growth through the development of biological solutions to pressing societal problems, with the policy reforms and targeted interventions required to maintain the fitness of humans and the biome. The size, complexity, and multifaceted nature of this problem require the attention of a national laboratory.

Innovation for sustainable economic growth, while maintaining the fitness of humans and the biosphere, is one of the greatest scientific challenges we face. Human health is tightly coupled to the health of the biosphere, especially the microbial communities that perform a variety of essential ecosystem services, including elemental cycling, clean air and water, abundant food supply, and nutrient transformations. Disruption of these services as a result of environmental challenges (e.g., climate change) will have profound impacts on human health and the global economy. Unraveling the complexities of biological responses to the many internal and external environmental challenges we face — and their impact on the health of humans and the biosphere — requires technical and integrative advances allowing a deep, multiscale, integrated knowledge of mechanistic and phenotypic responses to these factors. For example, determining the roles of environmental factors in cancer initiation and progression requires understanding, predicting, and mediating complex reciprocal interactions among multiple levels of biological function in both host organisms and their microbiomes (see *Microbes to Biomes*, Page 40). 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, a microbial community, 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 effects on organismal fitness.

Integrated, cross-disciplinary, basic, and applied research fits perfectly with the mission and unique capabilities of Berkeley Lab. 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 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, toxins and thirdhand smoke in model systems, and the effects of exposures on the etiology of diseases such as cancer. In addition, partner faculty at UC Berkeley, UCSF, and other Bay Area institutions bring considerable expertise and experience in molecular epidemiological studies of human health and disease, which complements the Berkeley Lab "Big Science Missions" and capabilities in elucidating basic biological mechanisms.

Together, Berkeley Lab and its partner Bay Area institutions are poised to integrate basic, mechanistic information about the responses of biological systems to environmental challenges, with an accurate measurement and understanding of the extent of exposure to environmental factors and their impact on health and disease. Achieving this quantitative and integrative understanding will address a pressing national need to accurately predict — and mediate — the impact of environmental challenges on biological systems, which is essential as reliance on biological solutions and the bioeconomy grows. Our efforts will result in the development of safe, sustainable energy and materials, reduced exposures to harmful environmental factors, improved public health and personalized medical interventions, and an increase in our understanding of the impact on other organisms that contribute to biome fitness. Given sustained effort, we will be world leaders 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 with resulting major positive impacts on the economy.

## Health Research Strategies

- **Biological responses to environmental challenges.** Develop and deploy model systems to study molecular, cellular, and organismal responses to environmental challenges.

- **Impact of environmental challenges on human biology.** Build platforms that rapidly elucidate the immediate and long-term consequences of environmental challenges on human health and disease, and develop innovative bio-solutions.

- **Technologies to accurately assess the impact of environmental challenges on biological systems.** Advance and integrate bioimaging, rapid functional analyses, and computational approaches required to measure biological responses to environmental factors.

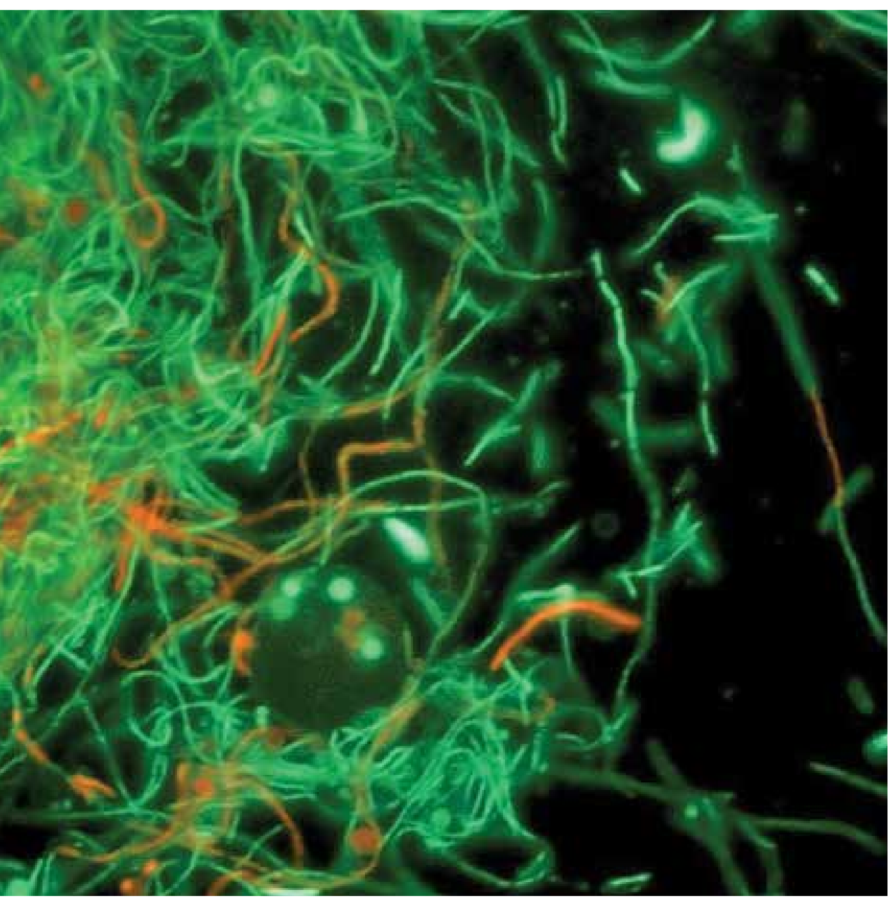
The integrated and multilevel network nature of biological data requires a national laboratory to harness biology for productive advances while avoiding confounding factors. To achieve our Health Goal by 2023, we will focus on three strategies: (1) building a cross-disciplinary platform that provides a comprehensive, integrated understanding of positive and negative biological responses to environmental challenges in model systems, covering levels of function from molecules to organisms; (2) assessing the impact of environmental challenges on human health and disease; and (3) integrating multiscale data using bioimaging, genomics, proteomics, metabolomics, and computational technologies to rapidly phenotype and quantitatively interrogate complex, dynamic biological systems.

Our 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, and Biomanufacturing). Through these integrated efforts we will develop mechanistic understandings of these processes that will ultimately inform prevention, diagnosis, and treatment. The Health component will also play key roles in developing a suite of technologies for rapid assessment and imaging of phenotypes that will be an important tool throughout Biosciences Area.

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 biosensors (see *Biosensors for Environment and Health*, Page 41) and bio-solutions to assess risk and mitigate health problems related to environmental exposures — and are achievable in a 10-year time span. The technical and intellectual strengths of the Biosciences Area and additional Berkeley Lab Divisions from Areas outside of Biosciences will be brought together to define mechanisms that enable advances in the productive application of renewable biological resources to protect and improve the economy. As a result, we also will develop novel methods to connect nanoscale molecules to 1-million-fold larger scales in organisms to enable informed mesoscale biology. By integrating nano- and organism scales, we will enable a predictive, multiscale understanding of responses to environmental challenges for economic growth, and the fitness of humans and the biome. In addition, this effort will provide advanced training of the required workforce, and new solutions that will contribute to the development of the economy.

**Biological responses to environmental challenges:** Develop and deploy model systems to study molecular, cellular, and organismal responses to environmental challenges

This strategy aims to develop a focused, mechanistic understanding of how specific environmental challenges — global climate change, byproducts of energy production, aging, diet, toxins, and industrial chemicals — impact the fitness of humans and other organisms relevant to economic success. An important first step is defining how exposures affect biological systems in specific model organisms, from macromolecular complexes to biological outcomes. With this understanding in hand, we aim 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 our 10-year Goal. Environmental factors can affect different types of molecules, cells, and tissues 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, proteins) in different cells and tissues, plus 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 (see *Microbes to Biomes*, Page 40).

By identifying key biological mechanisms in model organisms, rather than merely cataloging relationships between input and output, we aim to predict responses of biological organisms in hypothetical situations 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? The focus on manipulable, isogenic experimental systems allows us to control the genotype and environment, and to perform 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 our next strategy.

Specifically, we plan to focus on the well-established model systems, i.e., *Drosophila melanogaster* (fruit flies) and *Mus musculus* (mouse). The strengths of these systems include the deep biological insights generated over the past century and the ability to leverage sophisticated genetic, epigenetic, cell biological, and developmental tools for manipulation and measurements under conditions in which the environment can be controlled. In these tractable systems, we aim 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 in both the host and its resident microbiome (see *Microbes to Biomes*, Page 40; and *Biosensors for Environment and Health*, Page 41). The manipulability of these models will also allow us to correlate molecular changes with cell, tissue, and organismal phenotypes, ranging from cellular structures to complex 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 our approaches to elucidating environmental impacts on human biology and health.

**Impact of environmental challenges on human biology:** Build platforms that rapidly elucidate the immediate and long-term consequences of environmental challenges on human health and disease, and develop innovative bio-solutions

National decisions about investments in sustainable solutions for energy and climate change are interwoven with the health of the biome and impacts on health and the changing demographics of the U.S. population. Retiring baby boomers and the increasing cost of medical diagnosis and treatment are problems our nation needs to tackle at a different level than that being approached by current medical research. One approach to solving these problems is personalized medicine, whose goal is to target diagnostic and therapeutic strategies based on an individual's genetic predispositions. This approach is equally applicable to biome health, where personalized metrics for each biome are required to assess and minimize the impact of environmental exposures. In both cases, DNA sequence tells only a small part of the story; environmental factors have an equal or greater impact on health. Thus, prevention and treatment require a deeper understanding of the impact of environmental factors, integrated with information about spatial structure of the cells and their environment, genomics, epigenomics, metabolomes, proteomes, and microbiomes (see *Microbes to Biomes*, Page 40).

A significant knowledge gap exists regarding the composition and levels of biologically active chemicals in humans (the human internal "exposome"), and how these active chemicals are influenced by external exposures at different stages of life. We plan to elucidate the impact of environmental factors on human fitness and disease by using a two-stage approach. First, discoveries in model systems about the molecular, cellular, tissue, and organismal responses to environmental challenges will be used to evaluate the relevance to human fitness using state-of-the-art human tissue-mimetic technologies developed at Berkeley Lab. Second, in collaboration with investigators at other Bay Area research institutions, we aim to measure various omes (see *Biosensors for Environment and Health*, Page 41) in human blood and other human tissues and compare resultant profiles of diseased and healthy individuals to pinpoint discriminating molecules to ultimately identify causal exposures and sources. Identifying genetic, epigenetic, metabolomic, and microbiome factors that enhance or reduce the impact of specific environmental exposures will help identify individuals and populations at risk from specific types of exposures. Desired outputs of this strategy are: sound scientific data that government agencies can use to establish appropriate risk policies to mediate the effects of chemical, microbial, or radiation exposures; and the identification of biomarkers that can predict risk from an exposure or reveal a person's history of exposure.

**Technologies to accurately assess the impact of environmental challenges on biological systems:** Advance and integrate bioimaging, rapid functional analyses, and computational approaches to measure biological responses to environmental factors

Although technologies for deciphering biology, from molecules to organisms, have advanced considerably, there are major problems that need to be addressed to achieve our long-term goal. To achieve our 10-year Health Goal and advance our parallel research efforts in Energy, Environment, and Biomanufacturing, significant technological advancements are needed in three interrelated areas: bioimaging; genomics and phenomics; and computational biology.

### Integrated Bioimaging

Biological systems display unique behaviors, including self-organization across temporal and spatial scales ranging from atoms to organisms. Bioimaging provides for understanding the behavior and functions of molecular components in the context of cells, tissues, organisms, and communities. For example, to understand how a multi-species microbial community living in the termite hindgut degrades plant material into sugars for bioenergy production, we need information about the microbial species present, their location relative to the plant material, what enzymes are being secreted and what cellular pathways in the microbe regulates the secretion, the local reaction kinetics and by-products, the ultrastructural changes in the plant material, etc. However, most advances in bioimaging technologies are focused on a single imaging modality rather than an integrated application of multiple imaging modalities that complement and augment one another over the spatial and temporal resolutions involved in the biological responses to environmental challenges. Furthermore, biologists often begin an imaging project without considering what is needed to process and analyze large amounts of data, which results in suboptimal outcomes and conclusions.

Our 10-year vision addresses environmental challenge responses and other biological questions by integrating multimodality, multiscale, and multidisciplinary bioimaging approaches and by employing a novel "co-design" strategy in which experts in the biological investigation co-design, co-execute, and co-analyze experiments in close collaboration with imaging technology experts, chemists, and materials scientists who develop imaging probes, and computational scientists and mathematicians who excel at data management, analysis, visualization, and modeling. In this way, new imaging technologies will provide spatial and temporal experimental feedback critical to accurately employing major advances in high-performance computation and simulation codes.

To enable integrated bioimaging, coordinated technological innovations are needed to combine unimodal imaging platforms into multimodal imaging systems; develop sample preparation and containment methods that are compatible with multiple imaging modalities; generate multifunctional probes and labeling chemistries to provide contrast across multiple imaging modalities; develop image analysis, registration, and image-feature measurement algorithms to quantify features and relate information across modalities; and create visualization, modeling, and interaction systems for experimenters to interpret the knowledge embedded in the data.

### Genomics and Phenomics

High-throughput functional and phenotypic analyses that can be used to measure and monitor environmental exposures and biological responses to those exposures are central to Berkeley Lab's 10-year Biosciences efforts. Two major new research thrusts are the development of (1) high-throughput and personal sensor technologies that provide rapid and, ideally, real-time monitoring of exposure and response (see *Biosensors for Environment and Health*, Page 41); and (2) development of analytical approaches that provide detailed insights into the genetic and physiological state of the biological system and links to bioimaging and computational efforts. These new efforts promise to bring together new, large-scale research teams to maximize advancements toward our 10-year Health Goal.

### Computational Biology

The ability to define, predict, and control mesoscale biology, which bridges length and timescales of nanometers and millionths of a second in molecules — up to meters and days for organisms — is essential for knowledge to be harnessed into next-generation technology opportunities, societal benefits, and scientific advances for national challenges in energy, economy, and environment. Working closely with scientists in the Computing Sciences Area, we aim to develop the integrative tools and approaches required to advance our research programs and achieve the 10-year Health Goal. Specifically, methods are needed to manage and transfer large imaging and genomics data sets; develop image analysis, registration, and image-feature measurement algorithms to quantify features and relate information across modalities; and create visualization, modeling, and interaction systems for experimenters to interpret the knowledge embedded in imaging, genomics, and phenotypic data, "phenomics."

In addition, effective tools are needed for modeling and predicting specific impacts of environmental challenges on fitness and health. Computational simulations have achieved considerable power and accuracy when limited to narrow ranges of space and time, but expanding computational models beyond these limits yields unreliable outcomes due to small errors, omissions, and force-field limitations that largely preclude useful biological predictions. This limitation impacts energy, environment, and bio-engineering progress and solutions for national needs. Experimental feedback on spatial and temporal information is critical to enable accurate computational experiments and predictive models of responses to environmental challenges at multiscales. So our efforts aim to integrate appropriate experimental data to improve the accuracy of computational models. This integration of multiscale observations with theory and computational models will be enabled by advanced imaging and genomics and phenomics technologies, and will leverage key Berkeley Lab assets including existing tools and databases for predictive systems biology, e.g., KBase, NERSC, the Energy Sciences Network (ESnet), and the Open Microscopy Environment (OME).

## 2023 10-year goal achievement measured by:

### Biological responses to environmental challenges

- Using tractable model systems, identify the key components and mechanisms that regulate the impact of five environmental factors (diet, temperature/climate, time/age, chemicals, and radiation) on interrelated functions of molecules, cells, microbial communities, tissues, and organisms.
    - Quantify the 4-D dynamic responses to these environmental challenges using advanced imaging, genomics, phenomics, and computational approaches to generate an integrated understanding of the effects on molecules (DNA, RNA, protein, metabolites, epigenetic components), as well as phenotypes exhibited by cells, microbial communities, tissues, and organisms.
    - Determine the mechanisms that regulate molecular and phenotypic responses to the five environmental stressors.

- Determine how the effects of the five major environmental challenges are modulated by genetic and epigenetic variation in model systems.
    - Identify how DNA sequence differences affect individual responses to the environmental challenges using genetic mapping.
    - Elucidate how epigenetic variations influence responses to environmental stressors.

- Understand whether and how the effects of the five major environmental stressors are transmitted through cell divisions and transgenerationally to progeny.
    - Determine which acute environmental challenges produce impacts on biological functions across cellular and organismal generations.
    - Identify the epigenetic mechanisms responsible for transgenerational inheritance of environmental stress.

- Elucidate the role of prototypic community interactions within biological systems, and how they are reciprocally affected by the five environmental challenges.
    - Correlate host/microbial community composition with functional responses to environmental challenges in model systems.
    - Identify resident microbial communities that impact host fitness and define their interactions with one another and hosts.
    - Manipulate reciprocal interactions between microbes and model organisms to produce benefits to fitness in response to the five environmental challenges.

- Integrate multiscale imaging and computation to produce predictive models for the effects of environmental challenges on microbiome/eukaryotic biological systems.
    - Produce robust computational models that accurately simulate and predict responses to environmental challenges at different biological levels.
    - Identify predictive bioindicators for fitness that include genes, epigenetic markers, proteins, metabolites, and microbiome components.

## Impact of environmental challenges on human biology

- Determine whether environmental challenges shown to impact model organisms similarly affect human cells and tissues.
  - Develop four different human biomimetic tissue-culture systems (breast, skin, esophagus, gut) fabricated from normal primary human cells and extracellular matrices, and validate their relevance to in vivo tissues at the levels of architecture, gene, and protein expression patterns.
  - Quantify the impacts of the five exposures (identified as having impact in model systems) on human biomimetic tissues using genomics, phenomics, and bioimaging.
  - Incorporate discoveries about specific environmental challenges into exposome/epidemiological studies of human responses (below).

- Identify human biological response markers for six variables: obesity, smoking, age, migrant lifestyle, diet, and pregnancy.
  - Complete exposome studies on human blood samples from epidemiological studies.
  - Identify the 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.
  - Formulate a list of bioindicators for human health and disease that includes genes, epigenetic markers, proteins, metabolites, and microbiome components.
  - Predict how manipulating host and microbiome properties positively or negatively impact 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.
  - Identify individuals at risk for harm from specific environmental challenges using phenotypic data, genetic predispositions, and other biological assays.
  - Utilize links between genetic and epigenetic variation, microbiome composition, and environmental responses to develop personalized therapeutic and prevention strategies against two major diseases.
  - Foster development of clinical trials focused on disease prevention by working with clinical and advocate partners.
  - Inform the development of bioeconomy technologies, including energy production and use, to minimize impact to human health.

## Technologies to accurately assess the impact of environmental challenges on biological systems

### *Bioimaging*

- Demonstrate and employ a co-design strategy in which experts in the biological problem that is being addressed design and execute the entire experiment together with imaging experts, chemists, material scientists, computational scientists, and mathematicians.

- Combine unimodal imaging systems, including cross-scale imaging equipment, probes, and sample preparation, to produce novel and effective multimodal devices (e.g., mass spectrometry plus fluorescence imaging to simultaneously image metabolites and proteins).

- Apply multiscale imaging platforms to generate correlated structural and functional information, enabling quantitative modeling in a diversity of systems, from microbes to humans to biomes (see *Microbes to Biomes*, Page 40).

### *Phenomics*

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

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

### *Computation*

- Work with Berkeley Lab Computing Sciences Area to create and test cutting-edge enabling tools for managing, analyzing, visualizing, and modeling multiscale bioimaging data.

- Generate a comprehensive understanding of environmental responses by computationally integrating 4-D quantitative imaging, omics, and phenotypic data at multiple biological scales.

- Create computational models at multiple scales that accurately simulate responses to environmental challenges.

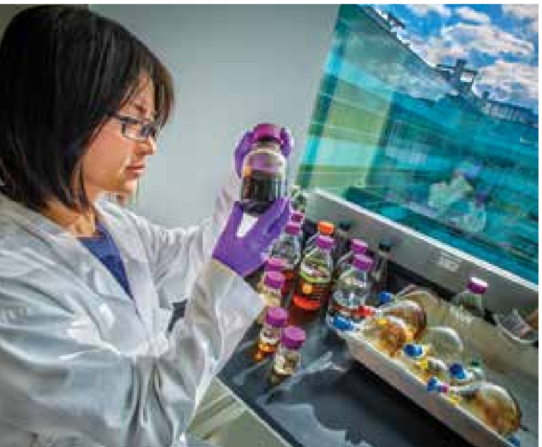
## 2018 – Five-year milestones for health research strategies

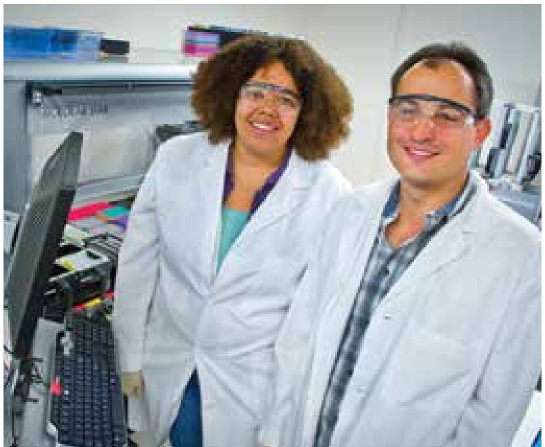
### Biological responses to environmental challenges

- Identified genetic, epigenetic, transcriptomic, metabolomic, proteomic, microbiome, and phenotypic responses to two environmental challenges (diet and toxins) in model microbial communities and host eukaryotic organisms (*Drosophila* 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 transgenerationally.

- Demonstrated technologies for high-throughput characterization of macromolecular





complexes acting in environmental responses.

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

### Impact of environmental challenges on human health

- Developed two human biomimetic tissue culture systems (breast and skin) and validated their in vivo relevance based on tissue architecture and gene expression patterns.

- Determined which two environmental challenges shown to impact model organisms also affect human biomimetic tissues, and quantified the response using genomics, phenomics, and bioimaging.

- Completed phenotypic analysis 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 includes genes, epigenetic markers, transcripts, proteins, metabolites, and microbiome components.

### Technologies to accurately assess the impact of environmental challenges on biological systems

***Bioimaging***

- Demonstrated and applied advanced multiscale bioimaging: combined unimodal imaging systems, sample preparation, multiscale probes, label-free systems.

- Produced and validated effective tools for managing, analyzing, visualizing, and modeling large multiscale bioimaging data sets (see *Computation*).

- Demonstrated the effectiveness of the co-design approach in assessing the impact of two environmental challenges.

***Phenomics***

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

- Demonstrated and applied technologies for high-throughput functional and phenotypic analyses.

***Computation***

- Demonstrated and applied methods for analyzing, visualizing, and integrating large, quantitative data sets.

- Created computational models at multiple scales that accurately simulate responses to environmental challenges.


RACK 500282 FOR 6mm
FOR USE IN THE COVARIS

# BIOSCIENCES FOR BIOMANUFACTURING

## 10-year Goal

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

## 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 our basic understanding of biology. We founded sequencing programs that can quickly and reliably tell us what useful genes exist in ecosystems all over the world. We developed analytic technologies that can predict and pinpoint novel metabolites in almost any sample we can gather. We developed a detailed understanding of biomolecular function and cellular physiology that enabled the design of reliable tools for effectively expressing pathways and engineering hosts to maximize productivity. We advanced the synthesis of arbitrary DNA sequences, obviating the need for a physical template and easing the creation of larger recombinant constructs. And we build knowledge systems that allow us to learn from our experience and computationally control the design and manufacturing process. We are building the systems that manufacture the organisms that in turn manufacture the chemicals and materials needed by our society.

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. Yet while costs have fallen and gains have been made,

### 10-YEAR Biomanufacturing Goal

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

### Biomanufacturing Research Strategies to Achieve Goal

**Tools to design, construct, and debug biology.** Develop computer-aided design tools, technologies to create error-free DNA/genomes, and analytical methods to analyze engineered biological systems.

**Designed biological systems.** Engineer controllable, trackable, robust biological systems (prokaryotes, archaea, 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.

significant technological barriers still limit our ability to reliably create biological solutions to some of the most pressing global problems of energy, environment, and health. We have lofty goals — biofuels, novel antibiotics, helpmate bacteria, and modern materials — but we lack the infrastructure to reach them. Our version of engineering and manufacture has not been as predictive, as scalable, or as computationally supported as the more established disciplines.

Berkeley Lab's Biosciences Area can close the gap. We will develop a science and technology program focused on the challenges of creating an advanced biomanufacturing facility (see *Advanced Biomanufacturing Initiative*, Page 41). We will create scalable infrastructure and knowledge systems and foster a network of expertise. We will seek innovations in predictive engineering and in biological systems design. In doing so, we will improve and broaden the capabilities of biological engineering and bolster our strategic missions in health, energy, and environment.

Berkeley Lab has the critical mass of researchers and infrastructure necessary to form a biomanufacturing power center. Our core capabilities in genomics, microbiology, materials science, 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 tools, technologies to create error-free DNA/genomes, and analytical methods to analyze engineered biological systems.

- **Designed biological systems.** Engineer controllable, trackable, robust biological systems (prokaryotes, archaea, 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, we've developed an approach that employs three areas of strategic focus: the development of tools to design, construct, and debug biological systems; the design of biological systems; and the creation of biodirected materials and bionanosciences. We believe these areas are scientifically tractable within a 10-year span; that they will meet the long-term national need for novel biomanufacturing solutions; and that they 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 tools, technologies to create error-free DNA/genomes, and analytical methods to analyze engineered biological systems

The technical ability to edit and insert DNA into organisms has inspired visions of a new era of "synthetic biology" in which 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 required by the growing complexity of the desired applications. Moreover, once we have constructed a biological system, we lack tools to debug the system and improve upon it. Hence, we need tools to design biological systems, to construct very long DNA with high fidelity, and to debug biological systems once they are constructed.

### Biological Computer-Aided Design (BioCAD)

Advanced engineering relies on sophisticated computational 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 there are active efforts to move toward more sophisticated, functional databases and modeling tools (such as available in the DOE Systems Biology Knowledgebase [KBase]), tools and standards for the simulation and automated design, construction, and iterative improvement of biological systems remain in their infancy. A fully developed biological design automation framework would do much to democratize and advance biomanufacturing broadly.

### High-fidelity, Long DNA Synthesis

Once a biological system or subsystem has been designed, the information to implement the design must be encoded in DNA. The ability to synthesize DNA is foundational to all cellular and biomolecular engineering. Although the cost of DNA synthesis has dropped recently, the speed of synthesis and length of the DNA sequence synthesized is not keeping pace, nor is the error rate decreasing. The remaining challenges are largely in the development of scalable technologies that enable the rapid, error-free assembly of shorter synthesized fragments into very long, designed sequences and specifically diversified libraries that can subsequently be screened for a desired function.

### Biological Debugging Tools

Biological engineering is hindered by extremely long optimization and troubleshooting cycles. When a newly engineered biological system fails, the reason is often difficult to determine. 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 the Laboratory 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. Finally, DOE JGI scientists are developing methods for high-throughput, long DNA synthesis.

**Designed biological systems:** Engineer controllable, trackable, robust biological systems (prokaryotes, archaea, 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. We call the process of selecting for pliability, safety, and utility in organisms domestication, and the only downside so far has been the amount of time it takes to accomplish.

We have domesticated certain microorganisms. Bacteria and fungi involved in fermentative food and pharmaceutical production are the best understood, safest, and most manipulated microorganisms on the planet. In the past few decades, sequencing technology has begun to tell us exactly what genetic changes are correlated with domestication, and genetic manipulation technology has given us the ability to make those changes directly. Biological engineers have taken advantage of these modern tools to rapidly customize the domesticated organisms further, engineering 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 left us with an impressive array of genetic and genomic engineering tools — most of which only work in *Escherichia 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 we can actually engineer.

There is an alternative. Every natural product we seek to manufacture comes from a living organism, some of which may be quite amenable to domestication. That is, rather than re-engineering desirable pathways into a pliable but basically incompatible host, we could engineer suitability into already productive organisms. Adding new bacteria, archaea, fungi, and even plants to the stable of tractable hosts would not only advance the science of genetic manipulation, but would also significantly broaden the range of natural products we are able to manufacture.

Microbiology, botany, and synthetic biology are Berkeley Lab strengths. For example, Laboratory scientists have been central in developing CRISPR/Cas9 technologies for scalable, cross-kingdom engineering and regulation of microbes, plants, and mammalian cells. We have engineered microbes and plants to produce active pharmaceutical ingredients, advanced biofuels, substrates, and commodity and specialty chemicals and engineered them to be amenable to the processes of production such as deconstruction of plant biomass and consumption of raw cellulosic hydrolysates. The engineering of controllable, trackable, and robust biological systems for key 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 pharmaceutical dose to specific disease sites. In another case, a viral platform has been created for biologically assembling sophisticated materials at the nanoscale such as gold and silver noble-metal wires with high aspect ratios and diameters below 50 nm that can be used as cathodes for lithium ion batteries. Nucleic acids also have 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, a user facility that has assembled state-of-the-art tools for doing nanoscience, 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 it 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:
  - Efficient pathway retrosynthesis and host engineering for optimized production under industrially relevant conditions.
  - Integrating functional genomics data into the design process.
  - Controlled tissue and biomaterial manufacture.
  - Characterizations results analysis to optimize next design cycles.

- Attain capabilities for rapid genome/chromosome synthesis and transplantation:
  - Error-free synthesis of a microbial genome or any chromosome of any organism in one month for less than $1,000.
  - Transplantation of this object into key hosts for soil, water, and fermenter microbiology (including algae), and into key plants.

- Develop simulation capabilities for biomanufacturing systems:
  - Small-scale simulations of large-scale reactors to aid in strain optimization.
  - Mesocosm simulations for environmental and agricultural engineering.

- Develop detailed and extensible techno-economic models for manufacturing processes, applications, and risks:
  - Model-informed selection of strategies for synthesis and characterization.
  - Data-driven models to assess containment and potential ecological impacts.

### 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, CRISPR-Cas systems.
  - Create models of physiological changes induced by domestication and scale-up of manufacture drawn from previously domesticated hosts.
  - Develop means for identifying and tracking engineered organisms in the complex environments.
  - Develop means for removing engineered organisms from complex environments.
  - Create models for predicting and assessing the environmental impact of engineered organisms.

- Nominate a range of key hosts for biomanufacturing, agriculture, environmental remediation, water support, and human health:
  - Determine active and potentially tamable hosts compatible with key application environments.
  - Identify host properties that could be engineered for improved robustness, productivity, and optimized energy input.
  - Develop robust genetic toolkits for each host.

- Develop the ability to biosynthesize arbitrary products on demand at predictable yields

### Biodirected materials and bionanosciences

- Re-engineer biological self-assembly to control formation of protein-nanomaterials, protein-compartments, and organelles.

- Invent new routes for the design of biohybrid systems that mechanically or electronically interface active biological elements with polymeric and inorganic materials.

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

- Achieve ability to interface biological components to electronic apparatus to control their activity; transfer energy between chemical, electrical, and light systems.

## 2018 – Five-year milestones for biomanufacturing strategies

### Tools to design, construct, and debug biology

- Developed bioCAD/CAM infrastructure comprising one or more tools each for:
  - Efficient pathway retrosynthesis and host engineering for production under industrially relevant conditions.
  - Integrating functional genomics data into the design process.
  - Characterization results analysis to optimize next design cycles.

- Developed small-scale simulation of a large-scale reactor to aid strain optimization.

- Designed, implemented, and optimized biomanufacture of two key molecules.

- Developed capabilities for rapid microbial genome synthesis and transplantation:
  - Synthesized an error-free microbial genome in one month or less.
  - Booted up a microbial genome (functionally self-sufficient) in less than a year.

### Designed biological systems

- Designed a broad host range domestication protocol for use in nonmodel organisms

- Domesticated one previously intractable host.
  - Identified host systems responsive to domestication and manufacture scale-up.
  - Tested means for identifying and tracking the engineered host in different environments and assessed its impact.
  - Improved biosynthesis of several products at previously unattainable yields.

- Engineered plants that produce modified lignin that can be easily transformed into a useful commodity chemical.

### Biodirected materials and bionanosciences

- Prototyped a biohybrid system that interfaces active biological elements with other chemical, physical, or electronic materials.

- Articulated an architecture to biosynthesize silica nanostructures using a single biological entity.

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

# BIOSCIENCE CAPABILITIES

Berkeley Lab's integrated biosciences program benefits from the expertise of a large staff of leading researchers, access to world-class facilities, and the organizational strength of the Laboratory's divisions and affiliated research institutes. Capabilities are vast, spanning biosciences divisions, mission-focused centers, national user facilities, and the synergistic teams that collaborate on national-scale biosciences research efforts. The Biosciences Area draws from four divisions: Physical Biosciences, Life Sciences, Genomics, and Earth Sciences. Biosciences researchers interact routinely with colleagues throughout Berkeley Lab, which was founded on E.O. Lawrence's vision of team science. As such, each of the 4,200 employees of Berkeley Lab is a potential resource and an asset in the Laboratory's quest to bring science solutions to the world.

## Berkeley Lab Biosciences Divisions

**The Physical Biosciences Division** brings capabilities in structural, molecular, computational, and systems biology to provide fundamental insights into biological processes. The division houses major programs, including Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA), which is probing how individual microbes and microbial communities translate environmental signals into behavior; the Berkeley Center for Structural Biology, which maintains world-class resources for high-resolution crystallographic imaging; and the DOE Knowledgebase (KBase), which provides a platform for investigators to analyze and integrate their data with the growing mass of public data relevant to research in energy and environment.

**The Life Sciences Division** focuses on studying the fundamental biological processes related to human health, with an emphasis on cancer and neurodegenerative diseases. To address these complex biological questions, the division's multidisciplinary staff develops and maintains key capabilities in multiscale imaging, metabolomics, epigenomic analyses, structural biology, computational biology, radiation biology, and organism and cellular microenvironment engineering. These critical capabilities are also leveraged to address questions in bioenergy development and in the interplay between the environment and biological systems.

**The Earth Sciences Division** addresses DOE, national, and global needs to develop sustainable uses of Earth's natural energy resources and stewardship of critical environmental systems. The division's primary focus areas include carbon science and climate change, environmental remediation, energy resources, and fundamental Earth sciences. It has key capabilities in bioscience for environmental applications, including environmental microbiology, omics technologies for understanding complex ecosystems, and trait-based modeling. One of the division's strengths is integration of environmental bioscience with other areas of geoscience, including hydrology, biogeochemistry, and geophysics.

**The Genomics Division** is exploring the wealth of new information flowing from the characterization and analysis of genome sequences, from such diverse organisms as humans to the most primitive of soil microbes. While it is well known that DNA encodes the basic blueprint of life, it is not known how best to interpret most of this information. To address this question, laboratories within the division are developing computational, biochemical, genetic, and imaging methods to decipher the complex sequence motifs that control RNA transcription, DNA replication, and chromosome structure. The division is also home to Berkeley Lab's complement of the workforce at DOE JGI.

## Mission-Focused Centers

**The Joint BioEnergy Institute (JBEI)** is one of three Bioenergy Research Centers created by 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 to integrate various data types to build new knowledge and share their findings with others.

**The Integrated Bioimaging Initiative** brings together the imaging expertise and technology at Berkeley Lab to address some of the current limitations in performing integrated imaging, analysis, and visualization. Critical to this effort is the ability to visualize and analyze biological processes that function over multiple spatial and temporal scales and can only be addressed with multiple imaging modalities. The initiative plans to co-locate these capabilities in a single bioimaging facility in partnership with other regional institutions.

**Berkeley Lab's Computing Sciences Area** — consisting of the National Energy Research Supercomputing Center (NERSC), the Computing 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 the Energy Sciences Network (ESnet), which provides a data highway for all laboratories affiliated with DOE.

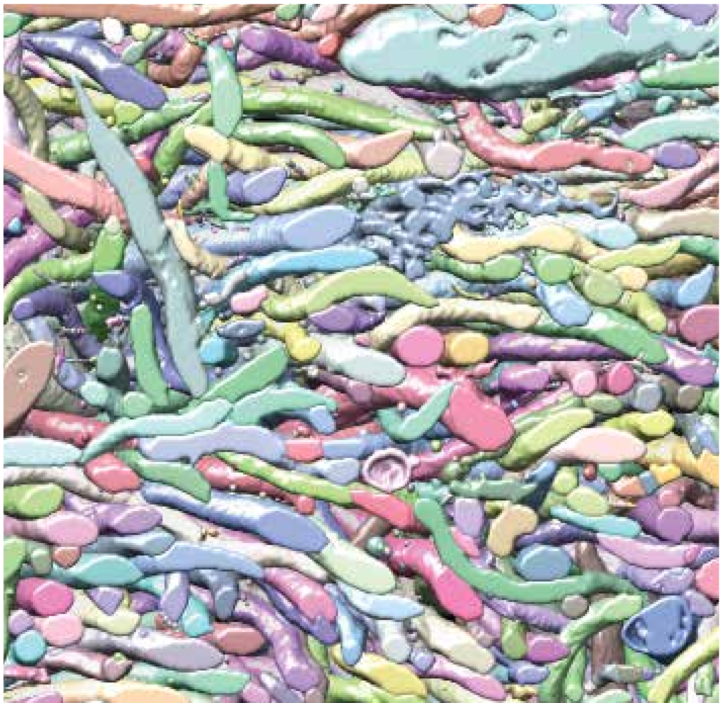
## National User Facilities

**The Advanced Light Source (ALS)** is one of the world's premier synchrotron radiation facilities and hosts a community of eight macromolecular crystallography (MX) beamlines for biological research. Beamline 12.3.1 is unique in providing both crystallography and automated small-angle X-ray scattering (SAXS) capabilities that provide accurate shape and assembly information about functioning macromolecules in solution. Other unique capabilities of the ALS especially valued in biosciences research include tomographic imaging of whole cells (the National Center for X-ray Tomography), infrared spectroscopy of live biological samples (the Berkeley Synchrotron Infrared Structural Biology Program), and spectroscopy of environmental samples (the Advanced Biological and Environmental X-ray Spectroscopy facility). Collectively, these resources are essential to research carried out by the Berkeley Lab Biosciences divisions.

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

**The Advanced Biofuels Process Demonstration Unit (ABPDU)** is a state-of-the-art user facility for testing and developing emergent biofuels technologies. The facility 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 National Energy Research Scientific Computing Center (NERSC)** is DOE's most scientifically productive supercomputing center, where more than 4,000 users log on to the systems from laboratories and universities across the country. Each year, NERSC users

generate about 1,500 scientific papers based on their use of these machines. The fastest at NERSC — the second-most powerful supercomputer in the United States — is clocked at 1.05 quadrillion calculations per second. Cutting-edge physics, materials science, and chemistry would be impossible without this kind of computer horsepower. Scientists use NERSC to model climate change, visualize reactions in biofuels and fusion, model clean combustion, and simulate the birth and death of stars.

**The Energy Sciences Network (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 engineers are developing a new network technology that will boost data transmission rates to 100 gigabytes per second — 10 times faster than today's.

## 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 DOE JGI are used to support JBEI, ENIGMA, and low-dose radiation research by Life Sciences and environmental research by Earth Sciences. JBEI is also a user of the ABPDU for development of novel biomass deconstruction methods. 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, we expect new collaborations and capabilities to emerge. To enhance collaboration among Biosciences research teams, we plan to relocate all Biosciences researchers to the Berkeley Lab campus. The first building, the Integrative Genomics Building, is proposed to house DOE JGI and KBase, bringing together these two complementary and synergistic DOE programs.

# BIOSCIENCES IMPLEMENTATION LEADERSHIP

**ENERGY Mentor: Eddy Rubin**
*Division Director*
Genomics

**ENVIRONMENT Mentor: Susan Hubbard**
*Division Director*
Earth Sciences

**HEALTH Mentor: Gary Karpen**
*Division Director*
Life Sciences

**BIOMANUFACTURING Mentor: Adam Arkin**
*Division Director*
Physical Biosciences

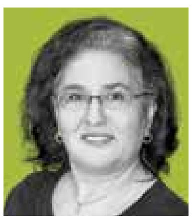
- CELLULOSIC BIOFUELS: Henrik Scheller
- MICROBIAL BIOFUELS: Chia-Lin Wei
- ARTIFICIAL PHOTOSYNTHESIS: Ian Sharp

- METABOLIC POTENTIAL OF NATURAL SYSTEMS: Harry Beller
- BIOGEOCHEMICAL CYCLES AND THEIR CONTROLS: Peter Nico
- CLIMATE AND ENVIRONMENTAL CHANGE: Charles Koven

- BIOLOGICAL RESPONSES TO ENVIRONMENTAL CHANGES: Susan Celniker
- IMPACT OF ENVIRONMENTAL CHALLENGES ON HUMAN BIOLOGY: Mark LaBarge
- TECHNOLOGIES TO ACCURATELY ASSESS THE IMPACT OF ENVIRONMENTAL CHALLENGES ON BIOLOGICAL SYSTEMS: Trent Northen

- TOOLS TO DESIGN, CONSTRUCT, AND DEBUG BIOLOGY: Nathan Hillson
- DESIGNED BIOLOGICAL SYSTEMS: Sarah Richardson
- BIODIRECTED MATERIALS AND BIONANOSCIENCES: Caroline Ajo-Franklin

# BIOSCIENCES AREA OFFICE



**Olga Diaz**
*Biosciences Administrator*

**Lorrayne Leong**
*Executive Assistant*

**Mary Maxon**
*Head of Strategic Planning and Development*

# GLOSSARY

| | |
|---|---|
| **ALS** | Advanced Light Source |
| **ALS-U** | ALS-Upgrade |
| **ABPDU** | Advanced Biofuels Process Demonstration Unit |
| **BioCAD** | biological computer-aided design |
| **CAM** | computer-aided manufacturing |
| **$CO_2$** | carbon dioxide |
| **CRD** | Computational Research Division |
| **CT** | computed tomography |
| **DOE** | Department of Energy |
| **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 |
| **OME** | Open Microscopy Environment |
| **PET** | positron emission tomography |
| **SND** | Scientific Networking Division |
| **SAXS** | small-angle X-ray scattering |
| **UC** | University of California |
| **UCSF** | University of California at San Francisco |

CSO 14-PB-0526