Lawrence Berkeley National Laboratory

BIOSCIENCES SCIENTIFIC STRATEGIC PLAN SUMMARY REPORT 2013–2023

2023

BioSciences

Lawrence Berkeley National Laboratory

BIOSCIENCES
SCIENTIFIC STRATEGIC PLAN
SUMMARY REPORT
2013–2023

BioSciences

REFLECTIONS FROM BIOSCIENCES' ASSOCIATE LABORATORY DIRECTOR

I would like to start out by thanking Jay Keasling and Mary Maxon for having the vision to introduce strategic planning to the Biosciences Area over a decade ago. By taking an approach that included clear aspirations, and defined milestones and outcomes, I think it has convinced researchers that this is a worthwhile activity. Indeed, the strategic plan has been an essential part of obtaining funding for important new programs in the Area, including the Agile BioFoundry, National Microbiome Data Collaborative, and the Microbial Community Analysis & Functional Evaluation in Soils (m-CAFEs) Science Focus Area. A strategic plan is nothing without the staff, both research and administrative, to enact it. My deepest thanks go out to everyone in the Area who has worked tirelessly over the last decade to deliver on the plan, with 66% of the Goals met. This is an amazing achievement and speaks to the strength and diversity of Biosciences research.

The importance of strategic planning to the Area led to the creation of the Biosciences Strategic Programs Development Group, led by Katy Christiansen. This group has been an essential part of the Area's growth over the last decade and for developing programs for future funding. I'd like to thank the entire program development team for their efforts over the years to help people develop new ideas, report on progress, and communicate the strategic plan to the broader community.

Moving forward, we have the opportunity to develop our strategic plan for the next 10 years, which will help guide the amazing, diverse research in the Biosciences Area. Since the first plan we have seen many changes, including a reorganization of the Area and changes in Area and Lab leadership. However, our science is vibrant and extremely well positioned to address important societal challenges, such as a reduced carbon world, a sustainable bioeconomy, biology for a resilient environment, and biopreparedness and health. It is also clear that making progress on these major challenges will require us to develop new technologies and research approaches that help drive science forward across the world.

In this new strategic plan, it is critical that we take a more holistic approach, where we also incorporate Inclusion, Diversity, Equity, and Accountability (IDEA) and research infrastructure in our thinking. These things might even be separate documents, but both are critical to our long-term success. I look forward to integrating our new strategic plan with those from the other Areas into the broader plan for the entire Berkeley Lab. Of course, our Area always in the business of discovery and our latest advances can be found at biosciences.lbl.gov.

Paul Adams
Associate Laboratory Director for Biosciences

TABLE OF CONTENTS

MISSION

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

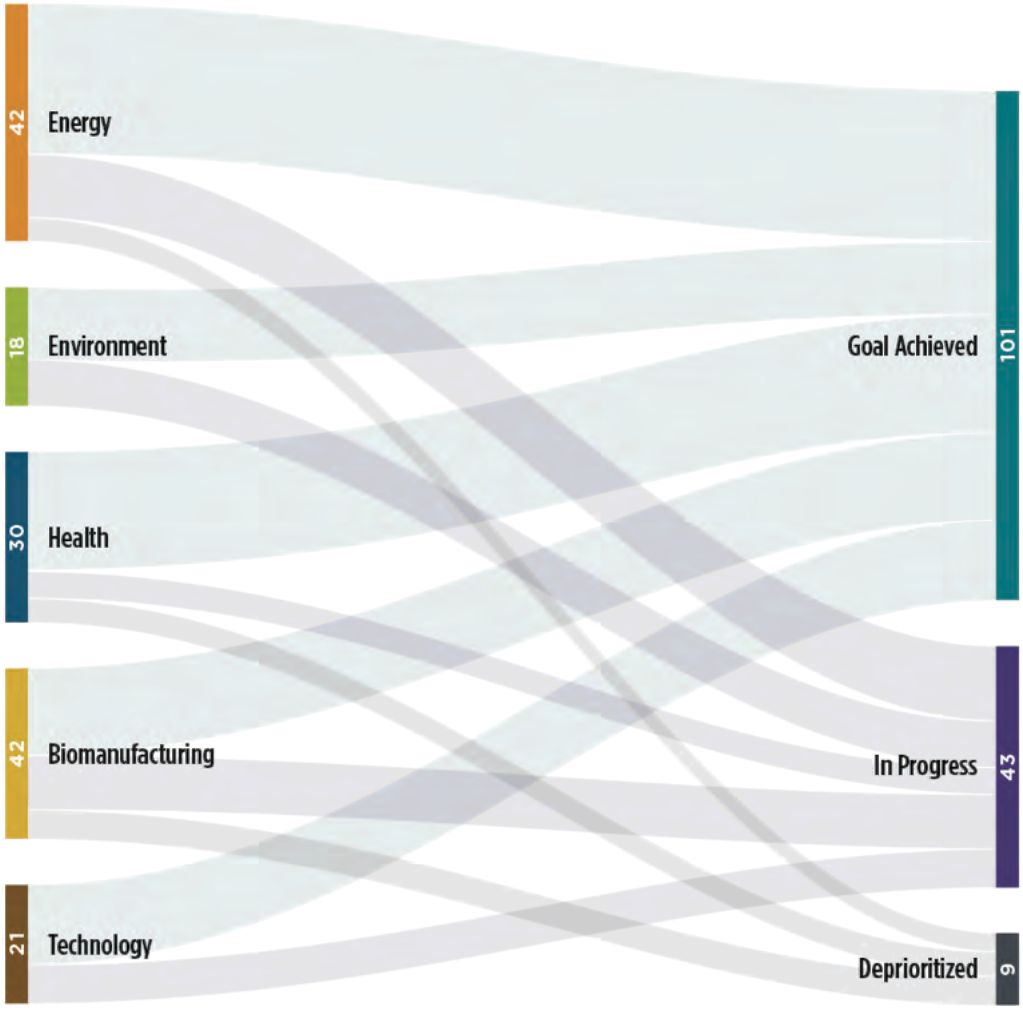
VISION

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


ENVIRONMENT
ENERGY
TECHNOLOGIES
BIOMANUFACTURING

ASSESSING OUR PROGRESS

Since setting our strategic goals in 2013, we've completed 66% of our 10-year metrics. Fully completed Goals were assessed by a final result or product available to the public through a publication, report, or patent. Of the remaining targets, 28% are in progress and 6% were deprioritized. A detailed synopsis of the metrics is provided in the Goal Completion Appendix. We are proud to meet 66% of our audacious metrics over a decade of focused planning. Our percent achieved indicates that we pushed ourselves both in setting and meeting ambitious targets.



CHALLENGES AND LESSONS LEARNED

Similar to the research enterprise, we've found strategic plans to be an iterative, reflective process. Some of the challenges we encountered were internal-facing, others were global or regional effects. Over the past 10 years, the Biosciences Area has undergone significant cultural change, from a somewhat loosely structured research organization acting responsively to funding opportunities, to an organized, collaborative group that looks to the future and sets strategic priorities for prospective research and development. This transition was not without growing pains along the way.

The first version of the Biosciences Strategic Plan introduced strategic planning to the Area. It also introduced the concept of a standing advisory board to provide advice on our priorities on an annual basis. In our first lesson, this advisory committee provided the feedback that the plan needed to be refined for implementation and recommended a framework for that. As a result, we developed a team consisting of Mentors to oversee each theme (Energy, Environment, Health, Biomanufacturing, and Technology) and Strategy Leads. Mentors are experienced leaders with a broad understanding of a research area, while Strategy Leads are subject matter experts in the research fields tasked with developing and monitoring implementation plans for their strategies.

Over the past decade, we have experimented with various ways of monitoring our progress and implementation of our strategic plan. From these different approaches, we have learned that we should be expansive in our definition of progress beyond the traditional metrics of success for research, publications, and patents. We also consider the formation of new research programs, the development of new instrumentation and equipment, and new collaborations as measures of progress. We've learned that we need to target our milestones and metrics such that they can be understood and measured to demonstrate progress, rather than establishing a set of disconnected desired endpoints.

In 2014, we began the process of re-envisioning the composition and organization of the Biosciences Area. Over the history of biosciences research at Berkeley Lab, our

organization had developed organically and our research Divisions represented collections of diverse research projects more than cohesive entities. The Strategic Plan identified the true centers of gravity for the Area and offered a new way to organize our work based on research directions and the scale of our research: from the atomic level of biological mechanisms to the complex interactions of organisms with their environments. We undertook a reorganization that amounted to a complete "melt and repour" of our research Divisions; this established new Divisions brought together based on common scientific interests and created a separate organization structure for the Joint Genome Institute (JGI) User Facility. This enabled the Area to more fully implement our strategic plan and achieve our strategic goals. At the same time, Berkeley Lab established the Earth and Environmental Sciences Area and this reorganization created respectful research boundaries between our Areas.

Another learning from this 10-year plan was to build our strategies around teams rather than individual researchers. When entire strategies are the specialized topic of one or a few individuals, the Goal is not only vulnerable to the individual success of the researcher (including funding, threat of outside opportunities, and career development) but also lacks the vibrancy of an interdisciplinary, team science-based approach. We began to course-correct in our strategic plan during the 2019 refresh and will take this lesson learned into our future plans.

When our strategic plan was first envisioned in 2012, there were significant changes in biological research on the horizon. The discovery and development of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) gene editing tools opened new possibilities in synthetic biology, biomanufacturing, and application of genetic discovery. We needed to refresh our plan in 2016 to reflect the new innovations enabled by CRISPR and how our approaches would be transformed with this emerging technology. While there is no assured way to future-proof strategic plans, we learned that we must embrace agility and respond thoughtfully to new opportunities.

Scientific research does not occur in isolation from other global, national, regional, and local challenges. We have become more agile in using our strategies to address these problems. As climate change accelerates, our research to reduce future carbon emissions

and to develop solutions to capture and store carbon in long-lived forms becomes more urgent. In California, we have seen some of the largest wildfires ever over the past 10 years and we have expanded our research to develop mitigation strategies to reduce or prevent future blazes. We have also experienced issues with our aging infrastructure, presenting hardships in achieving our Goals. Our team science approach includes operations professionals in key roles so that we can address problems quickly, minimizing impacts to research. Climate change impacts more than our research and our facilities, there is an emotional toll on our work as well. We see the impacts on biodiversity firsthand, feel the sense of urgency deep in our bones, and understand intimately what we have to lose.

Finally, the impact of the COVID-19 pandemic was far-reaching in Biosciences. The effects in our Area were similar to those faced by other research institutions, but no less significant to day-to-day operations. In the early days of the pandemic, our Area grappled with safely permitting on-site work, maintaining living organisms used in research, and easing our staff's adaptation to working at home and low staff-density for on-site work. As Berkeley Lab transitioned to more stable access, we learned how to prioritize experiments and think deeply about the new realities of team science. While the pandemic did negatively impact the Area—particularly for those who were adapting to changes in child or elder care—we also made important breakthroughs during this period. Our researchers led and contributed to COVID monitoring in wastewater and investigated the structural interactions between SARS-CoV-2 proteins and the human body, facilitating the development of therapeutics that received emergency use authorization. The pandemic showed us that contributions to science can take place both at home and on-site. As Biosciences enters its next 10 years, we look to the lessons learned about global threats, resilience and adaptation, and creativity in the face of immense challenges.

BIOSCIENCES' PARTNERSHIPS

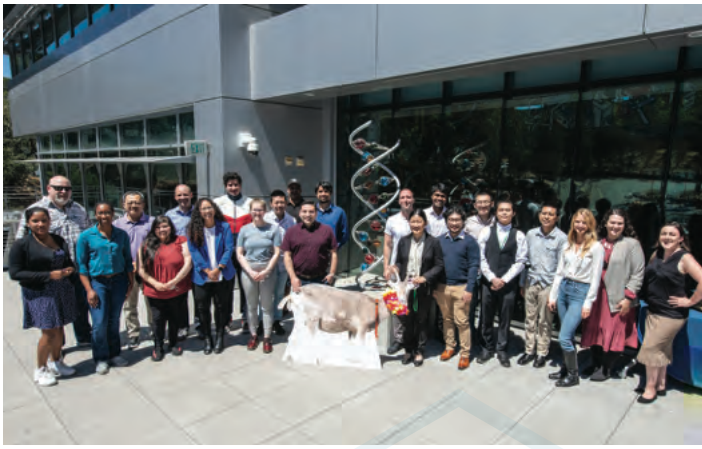
We collaborate with our sister Areas at Berkeley Lab in both long-term and emerging initiatives. The Energy Sciences Area (ESA) operates the Advanced Light Source (ALS) and Molecular Foundry, which have close collaborations with Biosciences. A program example of ESA/Biosciences collaboration is the Liquid Sunlight Alliance (LiSA) where Biosciences and ALS researchers work together on photosynthesis discovery using structural biology tools. With our Computing Sciences Area colleagues, we have been engaged in deep program development collaboration to kick-start initiatives in artificial intelligence/machine learning, quantum computing, and lab automation/self-driving labs. Much of this work has focused on establishing working relationships and common understanding between our disparate research identities. We anticipate this effort will bear fruit in the next round of Berkeley Lab investments for multi-Area proposals. The research focus of the Earth and Environmental Sciences Area (EESA) dovetails with the Environment strategy of this plan. We work together in the Department of Energy (DOE) Science Focus Areas, such as Ecosystems & Networks Integrated with Genes & Molecular Assemblies (ENIGMA) and Microbial Community Analysis & Functional Evaluation in Soils (m-CAFEs). We've already seen deeper collaboration with our EESA colleagues in preparation for the opening of the Biological and Environmental Program Integration

Center (BioEPIC) building. The Energy Technologies Area works with the Biosciences Area on sustainability benchmarking tools, most prominently the Life-cycle Assessment tools used in the Joint BioEnergy Institute (JBEI) to assess the cost and environmental impact of biofuels and biochemicals, described in the Energy section. Our Physical Sciences Area colleagues collaborate with us in laser-based technologies, such as the Berkeley Lab Laser Accelerator (BELLA) Center, used as a potential tumor treatment, described in the Health section.

Biosciences engages with many National Labs. A leading example is JBEI, which includes members from Brookhaven, Lawrence Livermore, Pacific Northwest, and Sandia National Laboratories. In particular, Pacific Northwest National Lab researchers bring their expertise in fungal biotechnology in the Deconstruction Division while Lawrence Livermore and Sandia National Labs scientists and engineers develop microfluidic devices in the Technology Division.

Biosciences has various types of partnerships with universities including research, outreach, and educational engagements. The Joint Genome Institute (JGI) has proudly partnered with UC Merced, a minority-serving institution in the agriculture-rich Central Valley, in a summer internship program since 2014. Students build skills while pursuing research with JGI personnel ranging from data visualization to metagenomics to phage-plant interactions.

Many of Biosciences' technologies make it out of the lab and into industry through technology transfer, licensing, and industry incubation to new business. Two Biosciences programs, the Agile BioFoundry (ABF) and the Advanced Biofuels and Bioproducts Process Development Unit (ABPDU) work directly with industrial partners in their research focus. As an example, ABPDU has collaborated with more than 75 companies since its founding, in industries ranging from waste management (Recology) to chemical production (Zymo Chem) to agriculture (California Almond) to scale bio-based processes from lab bench to industrial scale.



HISTORY OF BIOSCIENCES — THEN AND NOW

Throughout our nearly 90-year history of biological research at Berkeley Lab, early roots set the foundation for achieving our 2013-2023 strategic goals. Berkeley Lab made its name in physics and radiation research beginning at its conception in 1931. Biology research began soon thereafter in 1935 when the founder and namesake of the Lab, Ernest Orlando Lawrence, invited his brother, physicist and physician John Lawrence, to join the growing cohort.

J. Lawrence's work focused on the effects of radiation on human health; he was the first to use radioisotopes to treat cancer and other diseases. In 1937, he had his first breakthrough when he used the radioisotope phosphorus-32 to successfully treat a bone marrow disorder (polycythemia vera). Built in 1942, the Donner Lab was established through a donation from William Donner, a steel magnate and philanthropist, to pursue cancer research. Beginning in 1972, Mina Bissell's research in cell biology revealed the importance of the extracellular matrix around malignant cancer cells. Her work with her team led to the development of cancer screenings based on the conditions of the extracellular matrix. Bissell quietly retired in 2022, but her research legacy continues to grow in our Health Goals, including collaborations with Strategy Lead Antoine Snijders. Using a genetic approach to human health, Strategy Lead Len Pennacchio and colleagues evaluate human genetic variation to assess impact on disease.

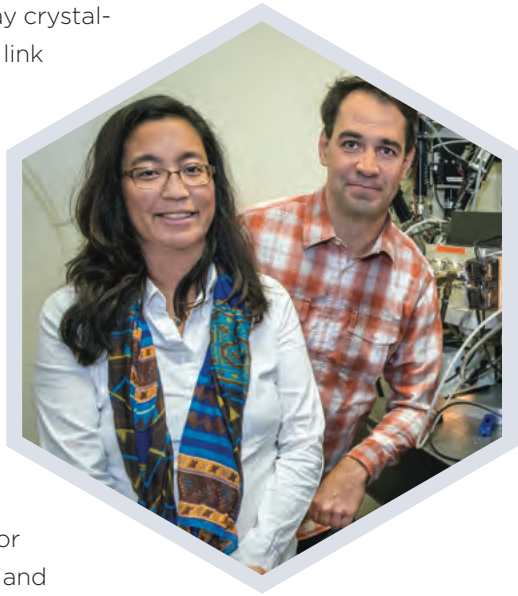
Beginning in 1945, E.O. Lawrence encouraged Melvin Calvin to pursue biological research with radioactive carbon. Calvin and Andrew Benson used carbon-14 as a tracer molecule to track the movement of carbon through plants. Their work in the biochemistry of photosynthesis yielded a Nobel Prize in 1961, and set the stage for current biochemistry and plant research. Today, research in Biosciences improves plants for renewable energy and materials by enhancing photosynthesis, increasing biomass yields, and making biomass more readily convertible to useful feedstocks. Strategy Leads Henrik Scheller, Jan Kern, Setsuko Wakao, and colleagues investigate photosynthetic systems and fine-tune the molecular machinery of grasses and algae for increased productivity, resilience to environmental stress, and deconstruction for biomanufacturing precursors and biomaterials

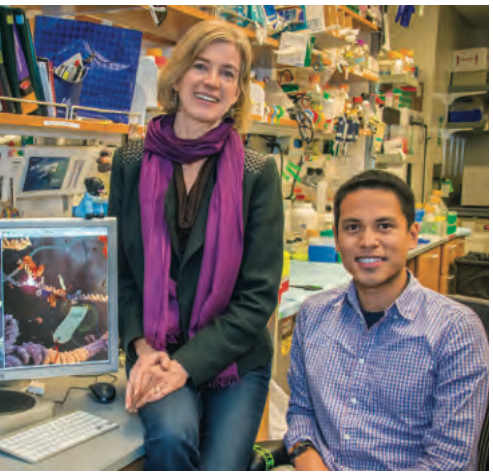
The Department of Energy (DOE) and the National Institutes of Health created the Office for Human Genome Research in 1988 and work began on the Human Genome Project in 1990. Prior accomplishments, such as sequencing of the model organism *Drosophila melanogaster* genome by Strategy Mentor Susan Celniker and colleagues, led to the selection of Berkeley Lab as one of the sequencing hubs. During the massive effort to sequence the human genome, Berkeley Lab's Human Genome Center merged with Lawrence Livermore and Los Alamos National Labs' sequencing efforts to form the DOE Joint Genome Institute (JGI) in 1997. The JGI team sequenced three human chromosomes (5, 16, and 19) between 1997 and 2004, while gradually expanding its scope to include species relevant to energy and the environment such as bacteria, fungi, and plants. In 2006, JGI officially relaunched as a User Facility providing genomics capabilities to the research community pursuing DOE mission–relevant goals in energy and biogeochemistry. Today, researchers including Strategy Leads Benjamin Cole, Rex Malmstrom, and Kateryna Zhalnina and colleagues use and expand on JGI tools to link genes to function, understand plant stress at the cellular level, and explore the deep unknowns of soil microorganisms. In addition to sequencing for the wider scientific community, JGI presently offers DNA synthesis, metabolomics, and other specialized services.

Researchers in Biosciences were early advocates of the importance of high-performance computing for data analysis. The DOE National Energy Research Scientific Computing Center (NERSC) user facility, located at Berkeley Lab, proved a valuable partner in addressing the computational challenges posed by large-scale sequencing. Two computational programs of note are the DOE Systems Biology Knowledgebase (KBase, 2012) and the National Microbiome Data Collaborative (NMDC, 2019). KBase is a software and data science platform that enables predicting and designing biological functions. NMDC is a community-supported microbiome database providing access to data pipelines and integration. Computing and data processing again play dominant roles in our "Eco" suite of technologies, which includes robotic handling and maintenance of small plants/microbes in Fabricated Ecosystems (EcoFABs, 2018,) developed by Strategy Mentor Trent Northen, and monitoring and environmental control of mature plants in the EcoPOD (2021). The Twin Ecosystems pilot project (2022) aims to replicate environmental conditions through probes at field sites and adjusts laboratory conditions to match.

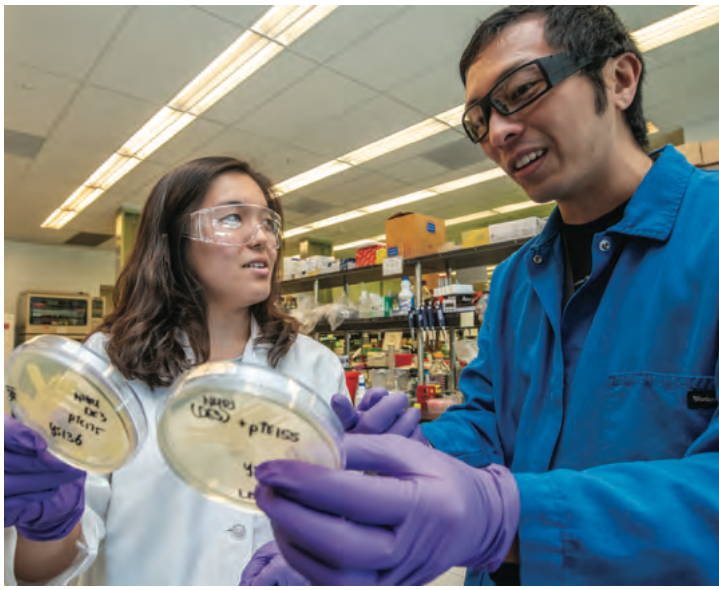
The Advanced Light Source (ALS), built in the dome that originally housed E.O. Lawrence's 184-inch cyclotron, was completed in 1993 after five years of construction. New beamlines and adjacent

technologies opened up research possibilities for Biosciences. In particular, X-ray crystallography enabled deep understanding of protein structures, another important link from gene to function. In the Berkeley Synchrotron Infrared Structural Biology imaging program, Biosciences researchers use synchrotron infrared technologies to image living tissues and cells. Strategy Lead Petrus Zwart develops software and workflows to analyze these images, and creates feedback loops to direct an instrument to prioritize experimentally relevant hot spots. This innovation saves instrument time and data processing. Using infrared spectroscopy methods, Strategy Lead Cynthia McMurray found that some diseased cells have a vibrational profile that allows researchers to identify Alzheimer disease at a cellular level. Strategy Lead Susan Tsutakawa uses small angle X-ray scattering to determine conformation of proteins in solution. Tsutakawa and colleagues recently used this technique to reveal conformations of individual components in a protein complex that binds with the SARS-CoV-2 virus. Connecting the ALS and photosynthesis, Strategy Mentor Junko Yano and colleagues use X-ray spectroscopy and crystallography to investigate metalloenzymes involved in photosynthesis.





Synthetic biology became an important offshoot of Berkeley Lab research in the mid-2000s. Then-Berkeley Lab Director and Nobel Laureate Steven Chu and Biosciences Senior Faculty Scientist Jay Keasling laid groundwork for synthetic biology research and biomanufacturing innovations, culminating in the Joint BioEnergy Institute (JBEI, 2007), Advanced Biofuels and Bioproducts Process Development Unit (ABPDU, 2011), and the Agile BioFoundry (ABF, 2016). In 2008, Jennifer Doudna was granted internal investment funds through her Berkeley Lab affiliation, which she used to research CRISPR DNA strands and the cas1 protein, setting the foundation for her Nobel Prize-winning research with Emmanuelle Charpentier and changing the course of synthetic biology. In 2018, we re-visioned this plan because of the impacts of their discoveries. Today, Strategy Lead Thomas Eng and colleagues use CRISPR techniques to engineer non-model microbes for biomanufacturing. Strategy Lead Steve Singer engineers microorganisms for lignin deconstruction and conversion of carbon to bioproducts. At the ABPDU, Strategy Lead Deepti Tanjore and team work with industry and academic partners to scale-up their synthetic biology–based technologies.



Strategy Mentor Nathan Hillson and team work with industry leaders to build synthetic biology engineering pipelines to shorten and build resiliency in the path from bench discovery to mature technology.

The roots of the scientific achievements in Biosciences today can be traced back to a long history of success. Each of these discoveries relied on the dedication of scientists, engineers, operations staff, and students. In our Lab's early years, those who had access to research time and funds (or had wide recognition) didn't reflect the full breadth and wisdom of the nation. In our last decade, we have deliberately aimed to be a more open, inclusive, and supportive organization towards those who have historically been underrepresented in research. Strategy Mentor Blake Simmons is the executive sponsor of the Latin American and Native American Employee Resource Group and has brought Inclusion, Diversity, Equity, and Accountability (IDEA) recognitions and announcements to Biosciences for many years. Many others across the Area work to build a more inclusive workforce for the improvement of our science's quality and applicability.

As seen through our own history, science requires the creativity of everyone, and everyone should have the ability to pursue the wonder of biology.

For the latest in Berkeley Lab Biosciences, visit us at

biosciences.lbl.gov

TIMELINE OF BIOSCIENCES ROOTS AND ADVANCES

1935	Beginning of biology research at Berkeley Lab
1941	Donner Lab for cancer studies built
1945	Calvin and Benson begin photosynthesis research
1972	Bissel begins cancer research
1990	Human Genome Project begins
1990	Fruit Fly (*Drosophila melanogaster*) genome published
1993	Advanced Light Source construction completed
1997	Joint Genome Institute (JGI) established
1998	Berkeley Center for Structural Biology (BCSB) established
1999	Computational Crystallography Initiative (CCI) established
2001	Structural Cell Biology of DNA Repair Machines (SBDR)
2003	Human genome published
2007	JBEI established
2009	ENIGMA Science Focus Area established
2011	Biosciences Area formed
2011	ABPDU established
2011	California Consortium on Thirdhand Smoke Exposure and Human Health established
2011	KBase established

envisioned 2013	Biosciences Strategic Plan 2013-2023 formed
2015	Reorganization of Biosciences Area
2016	ABF established
2016	EcoFAB research begins
2017	Integrative Genomics Building groundbreaking
2017	m-CAFEs Scientific Focus Area established
begins 2017	EcoPOD research begins, Version 1 unit arrived 2020, Version 2 arrived 2023
2019	NMDC established
2019	Biosciences Strategic Plan to include CRISPR and other advances refreshed
2019	Integrative Genomics Building (IGB) occupied
2021	BioEPIC Building groundbreaking
late 2024	BioEPIC Building occupancy expected

Reflections from Former Associate Laboratory Director, **JAY KEASLING**

I became the Associate Laboratory Director for Biosciences in 2010, at a time when our research portfolio had expanded into new directions. A review of the Biosciences Area in 2011 highlighted that this growth had changed the way we thought about our science and required a fresh approach to our future planning. From this feedback, we initiated a first-of-its-kind effort at Berkeley Lab to develop an Area-wide Strategic Plan that connected across our Divisions and User Facility, and that centered around core themes relevant to our aspirations in Energy, Environment, Health, and Biomanufacturing. This effort kicked off in earnest in 2012, and by 2013 we had the first version of the Biosciences Strategic Plan. Some of the ideas we developed in the plan were completely new, while others focused existing efforts on meeting new goals.

As the Biosciences Strategic Plan evolved and took root within the Area, we saw that it was time to reorganize our Area structure to support the implementation of the plan. This was a large undertaking, but we were able to build a new framework for our research that centered our science and positioned us for future growth. I am very proud that during this reorganization we were able to give our emerging leaders opportunities to manage research Divisions, more effectively implement our operations across the Area, and minimize impacts to ongoing research. The foundation of the Biosciences Strategic Plan was critical for the success of this reorganization.

It was my privilege to begin strategic planning in Biosciences and it has been gratifying to see how it has led to the creation of new research programs and directions, supported the career development of many of our scientists, and enabled Berkeley Lab's leadership in biological research for the Department of Energy and other funders. I am thrilled by the progress we have made so far and ready for more outstanding accomplishments from Biosciences in the future.

Jay Keasling
Former Associate Laboratory Director

Reflections from Former Associate Laboratory Director, **MARY MAXON**

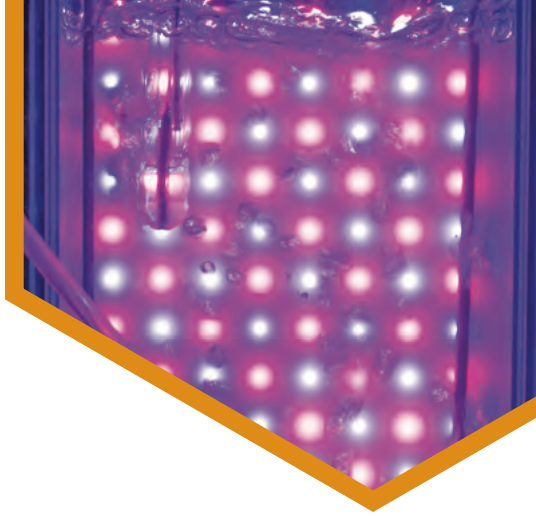
In October 2012, long-range scientific strategic planning with 10 year goals had not been implemented at Berkeley Lab. ALD Jay Keasling had the fortitude to give it a try and tapped me to lead and execute the necessary activities that could give rise to a vision for the Biosciences Area. Many researchers were skeptical that it would be a good use of time. They claimed they didn't have time for these activities, that their jobs were to respond to funding calls and conduct research, not dreaming up visionary ideas that might go nowhere.

Having worked in public and private grantmaking spheres in state, federal, and philanthropic entities, I knew that future funding opportunities came from people with vision—and that with some coaching, the talented and adventurous members of the Biosciences Area could rise to the task. And they did! The activities were designed to be inclusive from the start, with all 800+ Biosciences staff included in survey activities and focus groups to source the goals that benefit a national lab, not a company or university. Hundreds of ideas were shared. The most challenging part was deciding what goals would make the final cut and appear in the finished product, along with the specific 5- and 10-year metrics of success that would prove that we knew what victory would look like when we got there.

We established the Biosciences Expert Advisory Committee, an external body with diverse expertise, for three reasons: to review our new plan; to evaluate our progress annually; and to serve as third-party advocates for us ad hoc. They got their first look at the draft plan in December 2012, provided helpful input, and in May 2013, the Bioscience Strategic Plan was posted. I conducted dozens of briefings thereafter, with Berkeley Lab and UC Berkeley leadership teams, congressional committee staff, DOE program managers, and Federal Advisory Committees. Within three years, it was clear that some of those people had listened—some of our text appeared in funding opportunity announcements! Even better, there were multiple examples where DOE funds came directly to us without proposal competition, based on the widespread support for our vision across the national lab complex. At the time of my departure from Berkeley Lab in August 2021, approximately $200 million had been invested by DOE in our ideas, including in partnership with multiple national labs, directly linked to our Strategic Plan.

With all of our success, I can't wait to see what the Biosciences Area will aspire to achieve in the next 10 years.

Mary Maxon
Former Associate Laboratory Director



ENERGY

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

Assessment

Biosciences' leadership in the understanding and use of biology for the development of new approaches for energy is well known. Many of the programs in this section of our Strategic Plan were established or in planning at the time of the initial activities and more ambitious concepts were developed as a result. Through dedication and perseverance, researchers have successfully completed this Goal. Notably, integrated research efforts at the Joint BioEnergy Institute (JBEI) have reduced barriers around production of bio-based fuels. To meet our aspirations for Alternative Fuels, we have invested in new capabilities to use carbon in its gaseous form and joined a research consortium dedicated to integrating biology and electrochemistry to develop new fuels using renewable energy. Finally, research on photosynthetic systems, both natural and artificial, gave us new insights into how that energy from the sun can be transformed into fuels, products, and other potential applications.

Lignocellulosic biofuels: Derive fuels and coproducts from biomass with new technologies.

The Lignocellulosic Biofuels Strategy aimed to derive energy from terrestrial biomass with new technologies. The United States has abundant biomass resources; over a billion tons of feedstocks could be mobilized to replace 30 percent of U.S. needs for transportation fuels without significant impacts on human food and livestock feed production. To effectively utilize these resources, several breakthroughs had to be made: improved biomass with

10-YEAR

Energy Goal

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

Energy Research Strategies to Achieve Goal

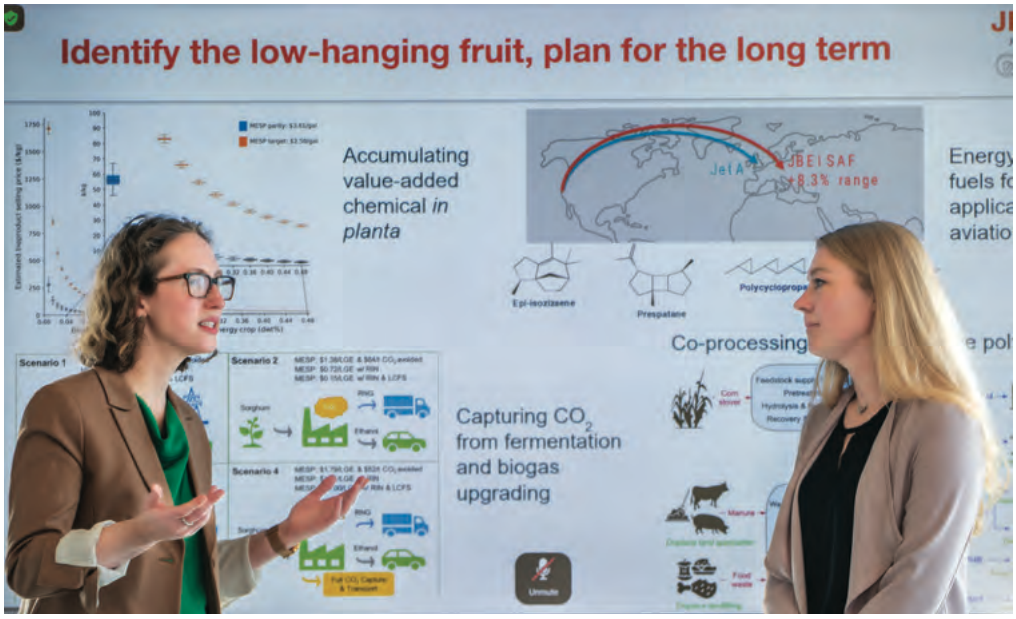
Lignocellulosic Biofuels
Derive fuels and coproducts from biomass with new technologies

Alternative Biofuels
Directly convert one carbon (C1) feedstocks to fuels using microorganisms

Artificial and Engineered Photosynthesis
Use bio-inspired reactions to create fuels directly from atmospheric carbon dioxide and sunlight

greater amounts of fermentable sugars and valuable lignin intermediates, greater tolerance to stress, and improved nutrient acquisition; greatly improved biomass extraction and breakdown strategies; engineered microorganisms capable of converting biomass sugars and lignin to high energy density fuels and additives; and the development and demonstration of economically viable and scalable production processes for an advanced biofuel and coproduct.

Berkeley Lab Biosciences has made significant advances in all the areas of the Strategy, with most of the efforts accomplished by JBEI. However, important results also have been obtained by others, including the Advanced Biofuels and Bioproducts Process Development Unit (ABPDU) and the Agile BioFoundry (ABF). Initial efforts in Biosciences were focused on biofuels, but over the course of the 10-year period a shift occurred, placing more emphasis on bioproducts that can increase the value of biorefinery operations and thereby reduce the costs of biofuels. JBEI was already established at the start of our Strategic Plan as one of DOE's Bioenergy Research Centers; it continued its work throughout the last 10 years, and was recently renewed for fiscal years (FY) 2023–27. JBEI's current focus is on sorghum as a bioenergy crop, deconstruction using distillable ionic liquids, and conversion into fuels and bioproducts by the bacterium *Pseudomonas putida* or the fungus *Rhodosporidium toruloides*. Researchers have directed their efforts toward the development of integrated methods to break down lignin from plant biomass to usable intermediates and develop microbes that can convert these to fuels and products. These efforts have been aided by targeted engineering of the bioenergy crops to modify the content and composition of lignin and other compounds present in plants. Through the use of computational modeling, we predict that these combined efforts over the last decade could lead to a decrease in minimum biofuels



Corinne Scown (pictured left) and her team have led the Techno-Economic Analysis (TEA) and Life Cycle Assessment (LCA) efforts at JBEI. Their models are an integral part of the design, build, test, learn approach taken in Energy research. They built the BioSiting tool (available online) which recommends sites for biorefineries across the US based on natural resources and existing infrastructure.

selling price from more than $150/gallon to around $25/gallon. A biorefinery that would make a coproduct from the biomass stream could further decrease the price, reaching a minimum selling price of $2.50/ per gallon.

Alternative biofuels: Directly convert one carbon (C1) feedstocks to fuels using microorganisms.

Decarbonization has become a key goal of the U.S. government. Biological conversion of carbon dioxide and methane to long-lived materials will be an important strategy for climate change mitigation. Incorporating these molecules into biomanufacturing will provide a low-carbon route to fuels and chemicals that will be independent of fossil fuels and provide a complement to the conversion of biomass to fuels and chemicals. Future work will focus on improving gas uptake and metabolism, developing engineering strategies for bioproducts, and establishing the viability of gas conversion through breakthroughs in bioprocessing.

The Alternative Biofuels Strategy supports the energy pillar by developing cost-competitive and environmentally sustainable routes to petroleum replacements by using gases (hydrogen, carbon dioxide, carbon monoxide, and methane) as substrates. The 10-year metrics for this strategy were to develop a systems-level understanding to identify key bottlenecks that limit fuel production by microbes, to use that information to engineer microbes to produce fuels from gas feedstocks, and to scale production through bioreactor and process development.

Since 2013, Berkeley Lab investments have advanced our ability to engineer and scale production of fuels and products from gases, while catalyzing additional federal and private-sector funding. For example, Biosciences researchers demonstrated that adaptation of a methane-utilizing bacterium to grow in the presence of lipids resulted in an evolved strain with increased capacity to produce fatty-acid derived molecules. Funding from the Energy Biosciences Institute led to re-designing of the carbon dioxide uptake system using transporters obtained from carbon-concentrating machinery in sulfur-oxidizing bacteria and cyanobacteria. These engineered hosts are the basis for flexible systems for carbon dioxide–based biomanufacturing. Subsequent work used a systems biology approach to identify determinants of growth on hydrogen carbon dioxide.

Additional institutional investments led to the development of a bioelectrochemical reactor for the synthesis of complex, multi-carbon products (e.g., drop-in biofuels, biomaterials, and commodity chemicals) directly from renewable power and carbon dioxide. Biosciences staff also participate as members of the DOE-funded CO2 Reduction and Upgrading for e-Fuels Consortium, linking biology and electrochemistry to further expand research and capabilities in the use of gas inputs. While single-carbon molecules have been a focus for Biosciences, researchers have also engineered the laboratory bacterium *E. coli* to enhance bioproduction using hydrogen gas.

The investment in using gases as feedstocks for biomanufacturing has supported the recruitment and retention of researchers and enabled their professional development in this rapidly expanding field of biological engineering. This work demonstrates the importance of multidisciplinary team science; biologists, chemists, and engineers are all needed contributors to advance research for this Strategy.

Early career scientists **Justin Panich** and **Sara Tejedor Sanz** are integral to the research in the Alternative Biofuels Strategy. As a postdoctoral associate, Panich engineered bacterial strains that demonstrated improved carbon dioxide uptake and developed E. coli strains that showed hydrogen-enhanced bioproduction. He was a key member of a collaborative team with UC Berkeley that received DOE funding for a project focused on using hydrogen gas for carbon-efficient bioproduction. Tejedor Sanz studied bioelectrochemical systems as a postdoctoral associate at the Molecular Foundry and brought her expertise in bioelectrochemistry and bioprocessing to the ABPDU, where she received a Berkeley Lab investment to design a new type of bioelectrochemical reactor. Panich is the metabolic engineering lead and Tejedor Sanz is the bioprocess lead for the CO_2 Reduction and Upgrading to e-Fuels Consortium that aims to develop sustainable aviation fuel through a combination of electrochemistry and biology.

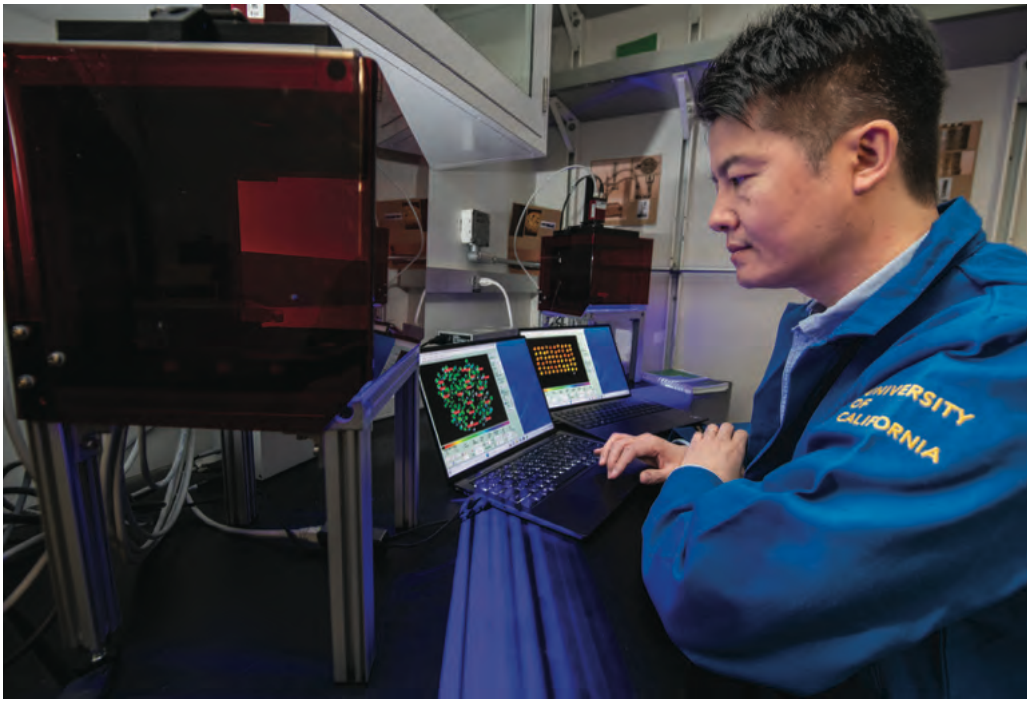
Artificial and engineered photosynthesis: Use bio-inspired reactions to create fuels directly from atmospheric carbon dioxide and sunlight.

In Biosciences many activities focused on different aspects of photosynthetic energy conversion and significant progress has been made on most of the metrics for our 10-year Strategy. Building on Berkeley Lab's historic role in photosynthesis research, efforts for this Strategy focused on both natural and artificial photosynthesis, as well as biohybrid systems that combine biological systems with novel materials. Photosynthesis is nature's way of converting sunlight into chemical energy and carbon-based products, opening up possibilities to harness photosynthesis—or develop new systems based on nature — to provide additional routes to fuels, chemicals, and materials.

Berkeley Lab researchers play an integral role in the Liquid Sunlight Alliance (LiSA), an Innovation Hub that seeks to establish the science principles by which coupled microenvironments directly generate liquid fuels from sunlight, water, carbon dioxide, and nitrogen. The precursor to LiSA, the Joint Center for Artificial Photosynthesis, operated from 2010–20 based on pioneering research conducted at Berkeley Lab. These researchers and collaborators developed a solar fuel generator called SolarCatMesh that creates limitless quantities of green hydrogen using sunlight and water, producing only oxygen as a byproduct. The discovery was recognized with a R&D 100 Award in 2022. In addition to LiSA, complementary work has been done to understand the principles underlying artificial photosynthesis and catalysis processes driven by light.

Biosciences researchers are expanding our understanding of and ability to harness natural photosynthesis to improve that process, as well as to control the biological processes that create inorganic materials for future energy needs. Systematic genetic and spectroscopic studies to understand photoprotection yielded a detailed picture of the different mechanisms that protect plants and algae from stress and damage under different light conditions.This understanding led to continuous improvements in yields of different crops by reducing the time needed for recovery from high-light stress. An improvement of 15% in biomass yield in tobacco plants and seed yield improvements of up to 33% in soybeans were obtained by engineering the photoprotection mechanisms in these plants.

Biological organisms display complex structures, some of which contain inorganic components. This is readily apparent in the diverse shapes of diatoms, single-celled algae that have cell walls composed of silica. Biosciences staff led a team of researchers in using genetic engineering targeting the process by which diatoms' outer skeletons (frustules) form and biomineralize to design and functionalize novel bionanomaterials. These can potentially be used as scaffolds for catalysts, separation reactions, and in photonics applications, opening up opportunities for renewable sources of inorganic materials from algae.



Masakazu (Masa) Iwai is a Biosciences Research Scientist who joined Berkeley Lab as a postdoctoral fellow in 2015. Iwai's research focuses on investigating how photosynthetic organisms utilize light energy efficiently using a combination of biochemistry, bioimaging, and molecular genetics approaches. He was involved in several important studies over the last six years, collaborating with nearly all groups within the bioenergetics department. Highlights include his contributions to recent work of elucidating initial charge separation steps in oxygenic photosynthesis, studies on the organization of light-harvesting antennas in photosynthetic protein complexes, and work on photoprotection mechanisms and how they can be modified to improve the productivity of crop plants. His current work is focusing on understanding how light-harvesting antenna complex proteins regulate energy transfer within photosynthetic proteins and membranes. He is very strongly involved in mentoring and training of undergraduate students and was one of the essential people ensuring that the Niyogi lab could continue functioning during the COVID limitations.

Frances Houle displays the R&D100 Award (above) for the SolarCatMesh solar fuel generator (left) in 2022.

Image credit: (above) courtesy F. Houle; (left) K. Waczak

ENVIRONMENT

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

Assessment

In the first iteration of our Strategic Plan, we partnered with our colleagues in what has since become the Earth and Environmental Sciences Area (EESA) to develop strategies for our Environment Goal. Since the creation of EESA in 2015, we have focused our strategies on organismal biology. The two strategies presented here represent achievement of this new Goal. Much of the work described in these strategies was initiated in the last eight years from the concepts articulated in the 2016 Biosciences Strategic Plan refresh. We have established new research programs to investigate organisms in their environments, from the microscale to a scale that reasonably approximates a real-world environment. In addition to the long-standing programs Ecosystems & Networks Integrated with Genes & Molecular Assemblies (ENIGMA) and DOE Systems Biology Knowledgebase (KBase), the Microbial Community Analysis & Functional Evaluation in Soils (m-CAFEs) program and additional avenues of research have expanded the environmental biology portfolio within Biosciences.

Predictive Understanding of Environmental Organisms

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

Life on land is exceptionally diverse and has an enormous impact on global biogeochemical processes. The interactions among microorganisms, plants, and animals drive nutrient flow, change substrate composition, and

10-YEAR

Environment Goal

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

Environment Research Strategies to Achieve Goal

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

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

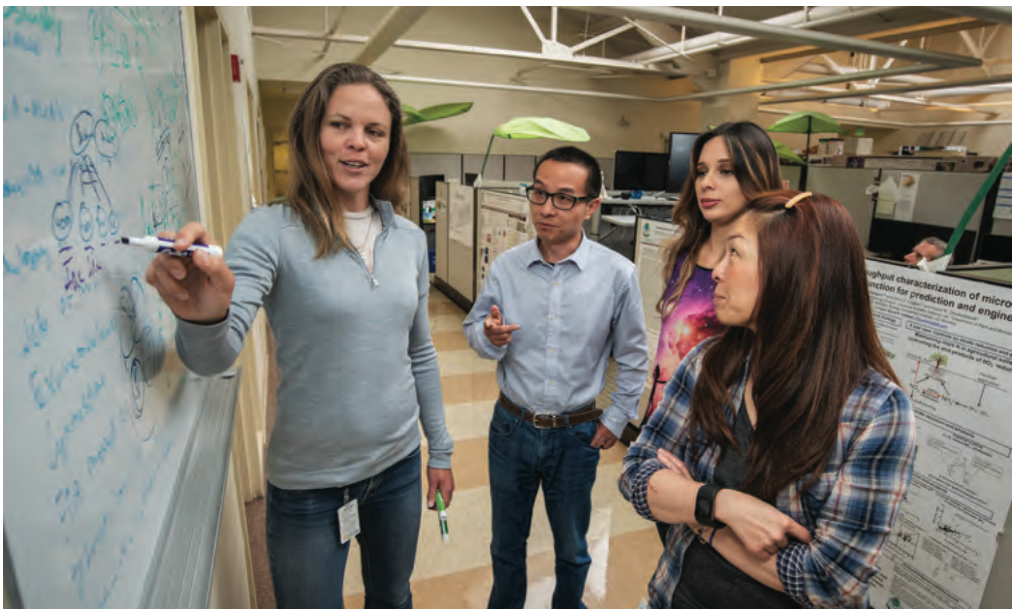
influence organismal evolution. Steady and rapid advances in molecular profiling technology (i.e., DNA sequencing, metabolomics, and proteomics) has enabled scientists to probe these environments with ever-increasing depth and richness. Yet a persistent frontier exists in understanding the functional roles of all the genes, metabolites, and proteins that comprise these environmental organisms.

Internal funding has played a critical role in developing advanced technologies to rapidly characterize molecular mechanisms. These include: projects to develop plant-scale fabricated ecosystems; novel technologies for investigating plant/microbial symbiosis; accessing novel functions within communities; understanding rules for microbial community assembly; and as well as multipartite systems with plants, bacteria, and viruses. The Department of Energy Biological and Environmental Research (DOE BER) supported ENIGMA Science Focus Area has continued to lead the way in the development of functional genomic tools and predictive models. The DOE BER supported KBase has helped democratize bioinformatic capabilities, and now the National Microbiome Data Collaborative (NMDC) program is providing new-to-the-world capabilities for data sharing. In addition to these major programs, two DOE Early Career research proposals have been awarded in the last 10 years to researchers working on predictive understanding of viruses and plants.

The ENIGMA program, through a massive collaborative effort, has used synthetic microbial communities to understand how the chemical environment (e.g., pH, metals, and other toxic compounds) influences large-scale biogeochemical processes driven by microorganisms, and has also developed groundbreaking technologies that enabled functional gene prediction for thousands of unknown bacterial genes. This has contributed to a growing database of functional assignments through our Fitness Browser web-based tool. To leverage these and other genomic resources for engineering, researchers at Berkeley Lab have also developed Chassis-independent, Recombinase-Assisted, Genome Engineering (CRAGE) to tame difficult-to-manipulate environmental bacteria, enabling the construction of large, multi-gene biosynthetic circuits.

The Gene Atlas project has generated thousands of genome and gene expression datasets across 18 diverse plant species to generate tens of thousands of new functional predictions and is freely accessible on the Joint Genome Institute's (JGI) Phytozome portal. The Epigenetic Control of Drought Response in Sorghum (EPICON) project has explored the widespread functional changes incurred during drought by sorghum, a leading bioenergy crop. JGI and other Biosciences researchers have adapted novel single-cell and spatial transcriptomics technologies to profile cell type–specific expression for plant genes, using these tools to probe responses to drought and nutrient limitations in important bioenergy crops.

Computational resources and consortia continue to play a strong role in organizing, synthesizing, and democratizing the large amounts of data generated under these initiatives, including KBase, NMDC, and the Gene Ontology consortium. These resources are necessary for building a predictive understanding of organisms, allowing researchers to store, classify, and analyze their work with additional data. This additional context and the opportunity to build on this knowledge enables new insights into the workings of organisms.



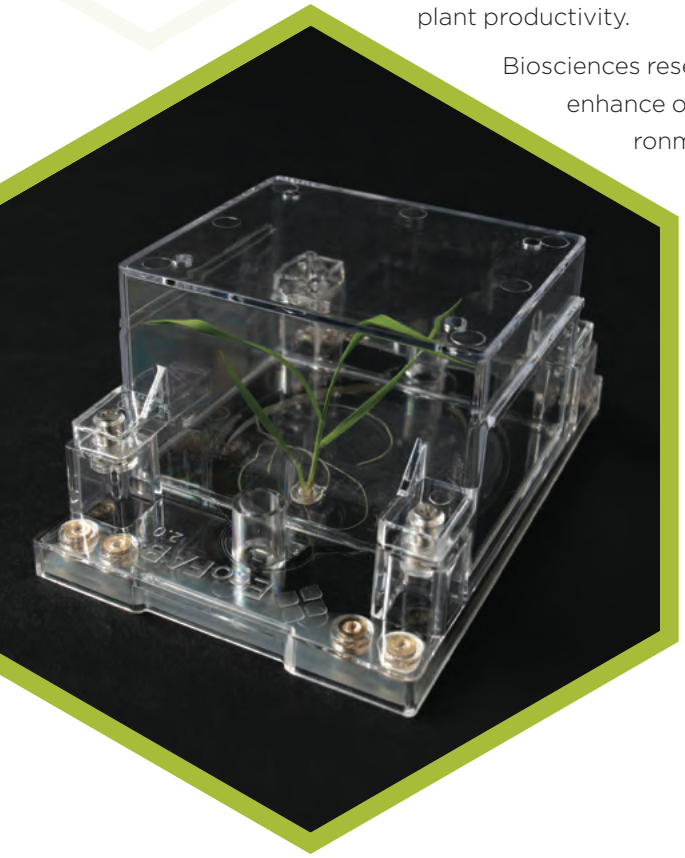
Elisha Wood-Carlson (left) joined Berkeley Lab in 2018. KBase had recently gone through their 3 year renewal process, with a recommendation to ensure the platform had deeper engagement with the broader research community. Trained as a scientist and in community engagement, Wood-Carlson was recruited to fill the new KBase community engagement role. She coordinates across a lot of projects, people, and research interests; finds intersections and possible collaboration points, and invests energy to bring ideas to reality - ultimately enabling researchers to be more efficient. The scientific community engagement role is new to the Biosciences Area, but it isn't a new concept. Many community facing projects have informal ways of making sure data and discoveries are made more findable, accessible, interoperable, and reusable (FAIR), increasing the impact of Biosciences Area research.

Molecular Ecosystems Biology-based Solutions

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

Plants are the primary carbon input into soils, represent the majority of biomass on earth, and are promising sources of renewable fuels and materials. Microbes control soil carbon and nutrient transformations, impact plant productivity, and support ecosystem health and resilience.

Over the past decade, Biosciences teams made crucial progress in studying microbial interactions with plant's environments in diverse ecosystems (grasslands, marginal soils, contaminated sites), at different scales (large field studies, mesoscale greenhouse experiments, lab-scale Fabricated Ecosystems (EcoFABs), and microscopy), with a variety of plant species (bioenergy crops: switchgrass, *Sorghum*, *Camelina*; annual model plants: *Brachypodium*, *Arabidopsis*, *Avena*; and algae). Researchers have analyzed thousands of microbes and microbial communities spanning organisms (bacteria, archaea, fungi, protists and viruses) and community complexities (single isolates, synthetic communities, complex natural microbial communities). Advanced genetic and systems biology tools have been used to uncover critical mechanisms that govern plant-microbe-environment and microbe-microbe-environment interactions and how this affects carbon and nutrient cycling and plant productivity.



Biosciences researchers have pioneered fabricated ecosystem capabilities to enhance our understanding of the molecular basis of plant-microbe-environment interactions and microbiome-driven plant growth promotion through manipulation of the genetic, organismal, and abiotic system components. These capabilities will be central to the new Biological and Environmental Program Integration Center (BioEPIC) building that is currently under construction. Specifically, biosciences has pioneered single-plant scale fabricated laboratory ecosystems, EcoFABs and meter-scale EcoPODs. Through multiple workshops and participating in sessions at the annual American Society for the Advancement of Science meeting, they have brought together a scientific community and built consensus on development of fabricated ecosystems. We've demonstrated the EcoFAB technologies robustness through multiple ring trials. These capabilities are in use and have been disseminated these capabilities as part

of the m-CAFEs SFA, Twin Ecosystems project, and Trial Ecosystems for the Advancement of Microbiome Science (TEAMS) program. EcoFABs have been successfully used to evaluate plant phenotypes, metabolites, and assembly of synthetic communities in the rhizosphere.

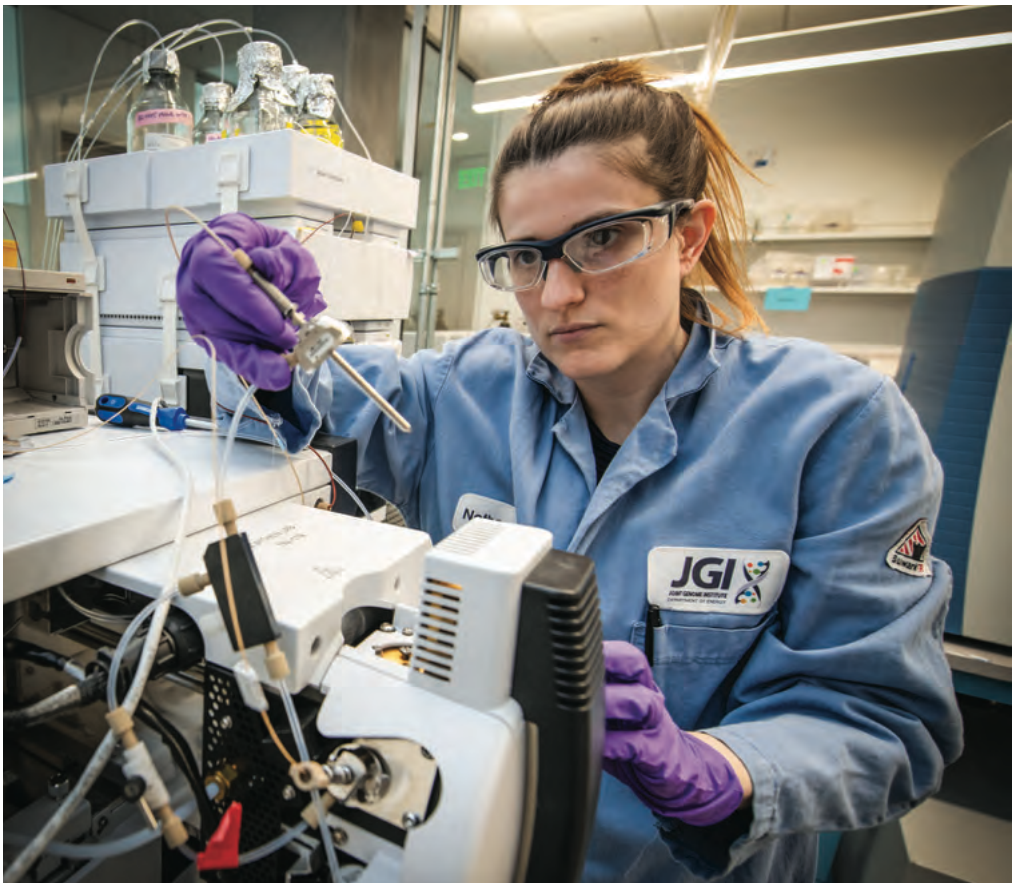
The Biosciences team has recently completed EcoPOD pilot experiments with natural and synthetic soil microbiomes and replicated field drought conditions to improve understanding of microbiome response to environmental perturbations. The Twin Ecosystems project aims to optimize in situ sensor technology and high throughput image analysis to explore temporal and spatial heterogeneity and dynamics of the plant-microbe-soil-atmosphere system from lab to field scales. Given the utility of the first-generation EcoPOD, the m-CAFEs science focus area purchased a next-generation EcoPOD with improved capabilities for microbiome experiments and gas analyses. This is an important step towards developing the final EcoPODs for the new BioEPIC facility.

To understand how to harness plants and their microbiomes for ecosystem services, for example, to achieve a quantitative increase in soil organic carbon storage and improve water and nutrients availability in soils, rhizosphere metabolites, plant exudates and microbial community dynamics have been studied in bioenergy crop switchgrass in marginal soils and model bioenergy plant *Panicum hallii*. Keystone metabolites in the rhizosphere of mediterranean grass Avena and switchgrass that have a significant impact on microbiome assembly and plant productivity have been identified. Biosciences researchers integrated trophic interactions between bacteria, archaea, protists and viruses to elucidate mechanisms controlling carbon stability and nutrient flow in the rhizosphere. The DOE and Berkeley Lab made significant investments to support research of understudied players in biogeochemical cycles, such as viruses, archaea, protists, and methylotrophs. The data generated in these studies will be used in modeling efforts to accurately predict nutrient cycling and biotic interactions in complex ecosystems and identify factors facilitating low-input growth of bioenergy crops.

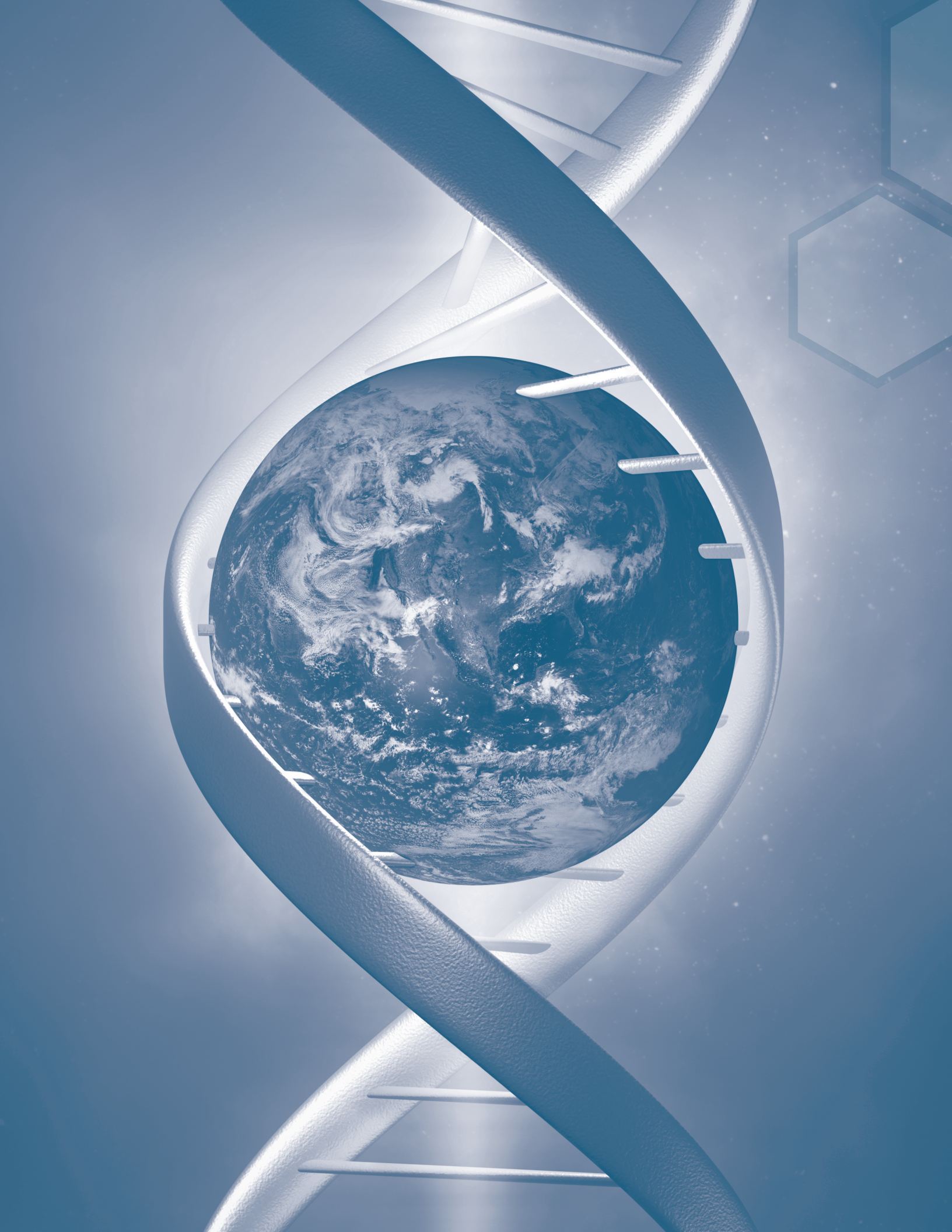
Development of predictive models has been another important component of our strategy. A major advance in this area has been the development of a mechanistic model that integrates field and microbial processes by the ENIGMA science focus area program. This Framework for Integrated, Conceptual, and Systematic Microbial Ecology (FICSME) model can be used to predict native community dynamics and mechanistically account for the material and energy flow at contaminated field sites. This model is being used in combination with extensive metagenomic data and detailed field measurements from a contaminated field site. ENIGMA researchers have pioneered new computational tools and methods for better assembly of genomes from metagenomes and integrate these with diverse data to accurately predict nutrient cycling and biotic interactions in complex ecosystems. Significant progress has been made on integrating dozens of environmental measurements with genomics data for subsurface samples (groundwater and sediment) to understand nutrient cycling by microbes.

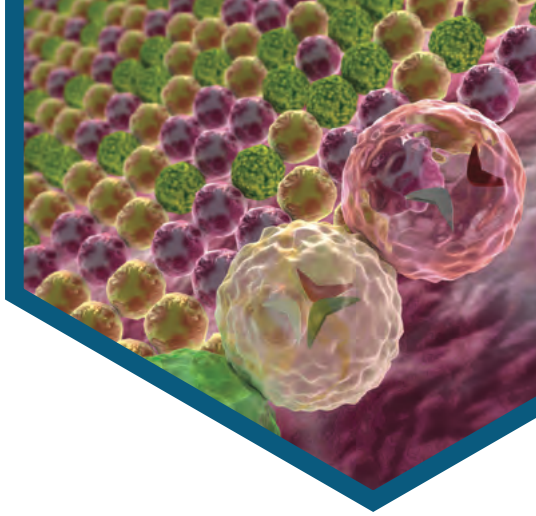
Biosciences developed several computational learning tools to enable mechanistic predictions and discoveries from the integrative analysis of complex, multimodal, multi-scale data collected on artificial and natural ecosystems. A functional zonation approach based on machine learning clustering to identify functionally distinct spatial regions at terrestrial-aquatic interfaces has been developed. Berkeley Lab made internal investments to develop machine learning to produce actionable knowledge for COVID-19 response, and developed algorithms contributing to better understanding the impact of biotic and ecological parameters on COVID-19 outcomes.

Many genes in environmental organisms don't have apparent functions when the organisms are studied in isolation. The m-CAFEs program has recently developed microbial community editing to directly test gene functions within fabricated ecosystems. This work is now being extended to model synthetic rhizosphere communities. This capability in combination with fabricated ecosystems will provide a powerful tool for testing predictive models and advancing a molecular ecosystems biology understanding of important microbial communities.



Susanne (Suzie) Kosina has worked in the Northen Metabolomics group since 2014. She is part of a team that generates functional solutions to some of the environmental problems facing the world today such as climate change and loss of healthy soils used for food production. She uses mass spectrometry to study the molecular changes occuring in plant-microbe-soil based ecosystems. Kosina says, "Analyzing mass spectrometry based metabolomics data is a bit like solving a puzzle — it is very satisfying to help uncover mechanisms of metabolic processes."





HEALTH

10-year Goal: Develop and apply a predictive, multiscale, and integrative appreciation of how individual variation affects responses to environmental challenges in order to improve human and biome health, and drive responsible economic growth

Assessment

The origins of health-related research at Berkeley Lab date back to the Lab's origins as a physics laboratory, where cyclotrons were used to create radioactive isotopes for diagnostics and treatments. In more recent years, much of the work in Biosciences has focused on large-scale team science capabilities, including genomics, structural biology, and other multidisciplinary approaches. Over the last 10 years, we have been able to achieve our aims in understanding fundamental human biology at genomic, molecular, and systems scales. Additionally, through internal and external investments, we have brought this knowledge to bear on questions related to how organisms respond to environmental challenges. Ongoing work in Biosciences, along with the onset of the COVID-19 pandemic and need for additional scientific research, enabled accomplishment of this goal.

Multiscale Understanding of Human Biology

Develop and deploy functional genomics, physiological monitoring, imaging, and computational modeling to enable foundational insights into human biology for applications in health and biodefense

The Joint Genome Institute (JGI) at Berkeley Lab played a critical role in the Human Genome Project, sequencing three human chromosomes. JGI's involvement was initiated by DOE's efforts to leverage its investments in large-scale facilities and team science. Following its completion (at a cost of >$1B), the human genome sequence proved to be riddled with unknowns—including the purpose of the 98% of the genome that does not

10-YEAR

Health Goal

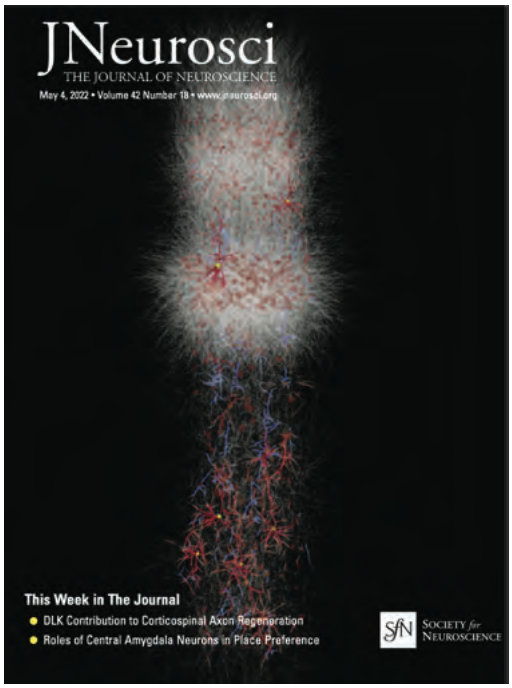
Develop and apply a predictive, multiscale, and integrative appreciation of how individual variation affects responses to environmental challenges in order to improve human and biome health, and drive responsible economic growth

Health Strategies to Achieve Goal

Multiscale Understanding of Human Biology
Develop and deploy functional genomics, physiological monitoring, imaging, and computational modeling to enable foundational insights into human biology for applications in health and biodefense

Biological Responses to Environmental Challenges
Design and integrate experimental and computational approaches to understand how individual genetics, epigenetics, and microbiomes impact molecular, cellular, and organismal responses to environmental challenges, and to identify biomarkers for disease risk assessment, as well as prevention and mitigation strategies

encode genes/proteins. Through studies at Berkeley Lab, we now know that a major function of this "dark matter" of our genome are hundreds of thousands of sequences that regulate the expression of our mere 20,000 genes. This was accomplished by developing and deploying functional and computational genomic tools to provide critical insights into human genome function both in health and disease.

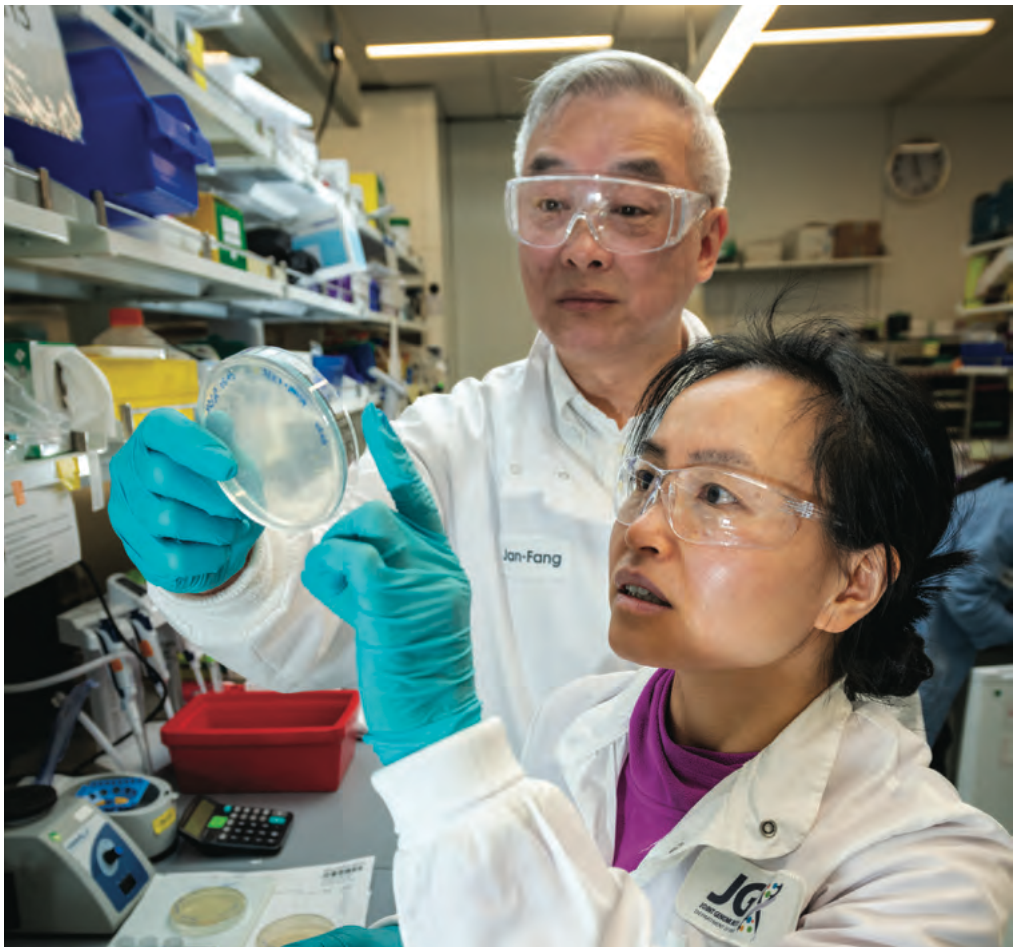


Specifically, Berkeley Lab researchers leveraged comparative genomics to prioritize evolutionarily conserved sequences with the expectation that critical functions in the human genome are likely conserved across species. Subsequently, next-generation sequencing approaches were developed to enrich for sequences containing signatures of gene regulatory elements. Importantly, for both of these approaches, *in vivo* assays in model organisms were developed and used to assess sequences thought to contain regulatory activity. Indeed, through these combined approaches, Biosciences researchers were able to identify and prove the vast landscape of gene regulatory function embedded in noncoding DNA.

These findings were published in >100 research articles and presented in ~100 scientific conferences/seminars since 2015. Biosciences also hosts the VISTA Enhancer Browser, which is used by tens of thousands of users worldwide.

To understand how signals are transmitted in the brain, researchers used electrocorticography, a method of using electrodes on the brain, to investigate the location of electrical signals and their sources in neurons. They found that the signals were located in a cortical column in the brain and were generated by neurons in specific locations. These results highlight the possibility of understanding how microscopic sources produce mesoscale signals in the brain, and how certain stimuli activate signaling. During this past decade, nearly all of these publications were the result of collaborations with external scientists in the UC system (including UC San Francisco, UC Davis, and UC San Diego) and external institutions (including National Institutes of Health (NIH), Cold Springs Harbor, and Stanford University). This led to the creation of the neural systems and data science lab.

Berkeley Lab was a member of an ambitious NIH project, the Encyclopedia of DNA Elements (ENCODE), a result of discoveries and expertise in human gene regulation. The goal of this program was to develop a better understanding of human genome function to enable an improved annotation of our genomes to ultimately relate DNA sequence changes to their impact on function. ENCODE studies identified more than 500,000



Jan-Fang Cheng (pictured behind) joined Berkeley Lab in 1989, arriving with newly appointed Charles Cantor, director of the Human Genome Center early, to pursue genetic research. Cheng and Cantor knew that the completion of the first human reference genome would not only be a landmark achievement in human genetics and medicine, but also the associated technologies developed would greatly change how we conduct biosciences research. The Human Genome Center later on became a part of the Joint Genome Institute.

In the past decade at the Joint Genome Institute, Cheng and his teams have developed ways to sequence genomes from eukaryote and prokaryote single cells, developed high throughput DNA synthesis pipelines to design and write genes and complex synthetic libraries, and developed a novel technology to engineer micro-organisms. The national laboratory system is all about team sciences and two changes in the last ten years have improved collaboration, according to Cheng: Biosciences buildings are coming up at the former Bevatron site and people are moving back to the main lab site and increasing awareness of Diversity, Inclusion, Equity, and Accountability and the core stewardship value of the laboratory.

potential regulatory regions in mammalian genomes, which can now be studied for their possible roles in a wide spectrum of human diseases. This program ended in 2022 and ushered in a new related program, Impact of Genomic Variation on Function, of which Berkeley Lab is a member.

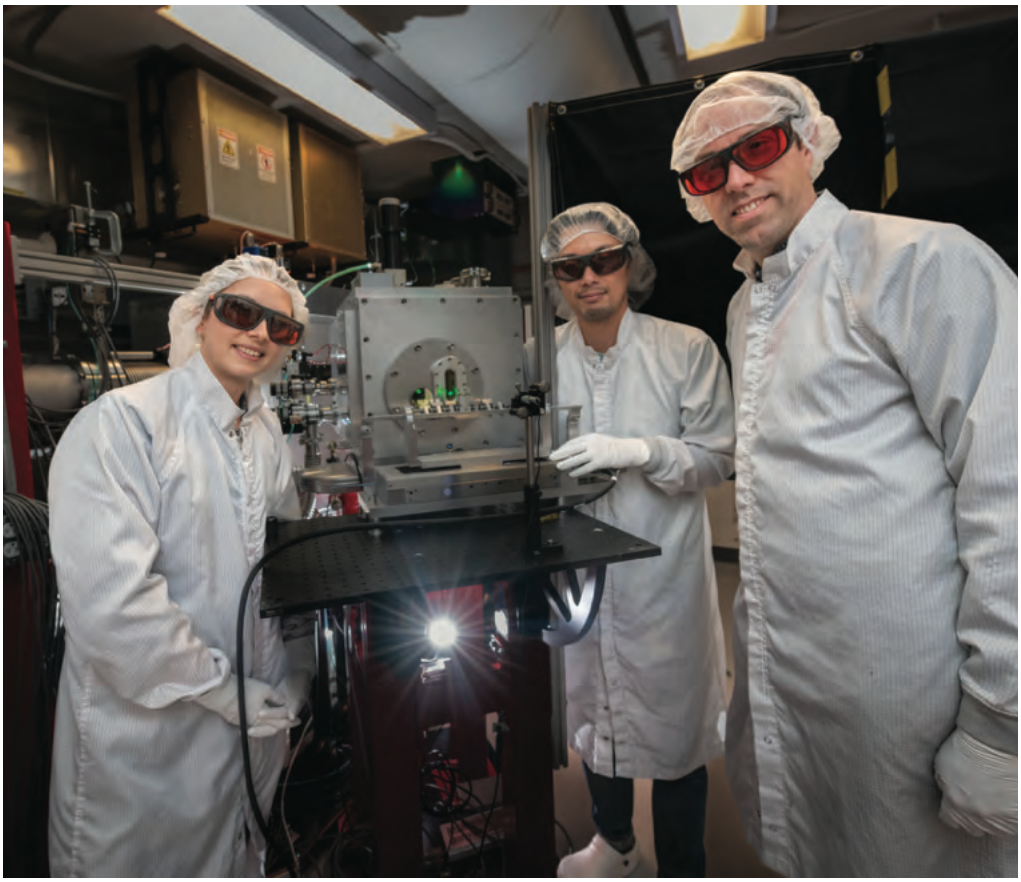
Biological Responses to Environmental Challenges

Design and integrate experimental and computational approaches to understand how individual genetics, epigenetics, and microbiomes impact molecular, cellular, and organismal responses to environmental challenges, and to identify biomarkers for disease risk assessment, as well as prevention and mitigation strategies

Biosciences research aims to understand the contribution of host genetics, environmental exposures, and their interactions in disease risk. Using a systems biology approach, together with model organisms and unique capabilities, scientists have uncovered how genetic susceptibility contributes to the variability in disease sensitivity along with environment and lifestyle. Through the use of mouse lines with genetic variability, researchers have identified a range of metastatic potential for breast cancer and are identifying genetic factors to control metastasis. Additionally, this same approach has been applied to investigate how thirdhand smoke (THS), tobacco smoke residues in indoor environments that remain, react, and/or re-emit from materials and/or re-suspend from surfaces even long after smoking has stopped, affects human health. These studies have shown that exposure to THS causes significant DNA damage in human cells and enhances tumorigenic traits. Moreover, early life exposure to THS caused persistent immunological alterations in mice and resulted in an increase in cancer incidence and alterations in anxiety and memory. This suggests that genetic backgrounds significantly influence THS-induced cancer and other disease development. Research in Biosciences also has generated basic information about the impact of pesticides on flies and mice and chemically induced colon cancer in mice.

Machine learning approaches can be used as an additional tool to predict health risks associated with environmental exposures. In many biomedical studies, human samples can be extremely limited and unbalanced due to various factors; animal models have therefore been developed for unbiased knowledge discovery and hypothesis generation. However, the translation from animal model risk to human risk prediction is an emerging challenge. Biosciences researchers developed an advanced machine-learning system to enable cross-species biomarker learning from animal models for health risk assessment in humans. With this technique, they demonstrated that the biomarkers and subtypes identified from one mouse mammary tumor cohort

stratifies an independent mouse mammary tumor cohort into groups with different metastasis risk, and the human breast cancer cohort into two groups with different overall survival associated with specific molecular features. Future optimization and deployment of the system will enable cross-species biomarker detection from the mouse model toward the assessment of environmental risk (e.g., low-dose radiation, environmental chemical exposure) on humans.



New capabilities for understanding the fundamental biology of health, along with prevention and treatment of diseases have been developed in the last 10 years. Among them, a collaboration between Biosciences staff and physicists at Berkeley Lab has investigated new approaches for radiotherapy treatments for tumors. Using particle acceleration based on ultra-high-intensity lasers, currently developed at the Berkeley Lab Laser Accelerator (BELLA) facility, a multidisciplinary team of **BELLA and Biosciences researchers** demonstrated an increase in cell survival after irradiation compared to conventional radiotherapy techniques. This team is also studying the genetic determinants of radiation induced toxicity in blood.


RACK 500282 FOR 6mm
FOR USE IN THE COVARIS

BIOMANUFACTURING

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

Assessment

Biomanufacturing in the United States has made great strides during the last 10 years and many advances can be attributed to Biosciences researchers. In March 2023, the White House released biotechnology and biomanufactur-ing goals for the nation — demonstrating the maturation of biomanufacturing from nascent concept to national target. Berkeley Lab researchers were key contributors to these new goals. Led by the Advanced Biofuels and Bioproducts Process Development Unit (ABPDU) and the Agile Biofoundry (ABF), our researchers streamlined scale-up and developed tools for industry, academia, and national lab technologies. We advanced Designed Biological Systems, primarily by developing and using Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) tools. We've engineered functions in new organisms, broadening the biological tool kit available for biomanufacturing. A decade ago, we set extremely ambitious Biodirected Materials and Bionanosciences metrics for ourselves. While we were not as successful as anticipated, mostly due to staffing challenges, we have renewed our efforts with recent hires and internal investments.

Tools to Design, Construct, and Debug Biology: Develop computer-aided design and fabrication tools, computational and analytical methods to model and learn from engineered biological systems and processes, and retrosynthetic infrastructure providing optimized pathways to key molecular hubs

Advances in computational methods and their applications in biological engineering have allowed us to develop cutting-edge tools to advance

10-YEAR

Biomanufacturing Goal

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

Biomanufacturing Research Strategies to Achieve Goal

Tools to Design, Construct, and Debug Biology
Develop computer-aided design and fabrication tools, computational and analytical methods to model and learn from engineered biological systems and processes, and retrosynthetic infrastructure providing optimized pathways to key molecular hubs

Designed Biological Systems
Engineer and scale-up predictable, controllable, trackable, robust biological systems (prokaryotes, archaea, eukaryotes, microbiomes) 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

biomanufacturing goals. In the past 10 years, machine learning has been applied in many facets of our lives, including manufacturing. However, advantages of machine learning and artificial intelligence are not fully realized in biomanufacturing, creating the need for Berkeley Lab researchers to develop several essential tools. Designing pathways and hosts for biofuel and bioproduct production using machine learning is a first step to this end. While our understanding of scale-up hurdles has substantially increased, access to scale-up infrastructure remains very limited. Berkeley Lab's ABPDU has contributed significantly by working with more than75 companies to help bring 17 products to market. The ABPDU built a first-of-its-kind pressurized fermentation reactor with internal investment that is available for industry to test their lab-scale strains and predict performance in 10,000 L+ commercial-scale reactors (pictured on page 48). However, access to these capabilities are still resource-limited as a single bioreactor is used to test a single experimental condition and there is very little understanding of commercial-scale behavior. Techno-economic and life-cycle analyses became more prominent with many funding agencies and even consumers prioritizing the benefits of low greenhouse gas emissions and water and land use in biomanufacturing-based products.

Researchers expanded biological computer-aided design and manufacturing (BioCAD/CAM) infrastructure in the last 10 years. Work conducted at the Joint BioEnergy Institute (JBEI) and the ABF led to machine learning–based (i) Automated Recommendation Tool to provide a set of recommended strains to be built along with probabilistic predictions of their production levels, and (ii) approach to predict metabolic pathway dynamics from time-series multi-omics data. Recent further development of engineering biology software platforms Design, Implementation, Validation Automation (DIVA) and Experiment Data Depot (EDD) focused on improving accessibility for the visually impaired. We also have developed a suite of design software to exploit the modular reconfigurability of polyketide synthase enzymes as a flexible system to access millions of molecules biologically. These software include the ClusterCAD database of PKS parts and the RetroTide PKS retrobiosynthesis software. These efforts meet our goals to develop BioCAD/ CAM infrastructure for synthetic biology to enable pathway retrosynthesis and host engineering for optimized production titers, rates, and yields.

Reactor design and scale-up requires advanced simulation and control capabilities for industrial-scale biomanufacturing. A self-driving bioreactor, with artificial intelligence–based real-time optimization of fermentation, can cut process development timelines for novel hosts and targets to half. Developing sensors that can assess intracellular changes in microbial cultures from bioreactors can offer the predictive power necessary to enable algorithms for self-driving capability. Berkeley Lab research funding led to a collaboration

project with UC Berkeley Department of Chemistry to develop quantum sensing capabilities for microbial hosts. The nanodiamond based high-throughput system was deployed with yeast cells and was tested successfully with high sensitivity for gadolinium. Flow cytometry and other imaging capabilities are being deployed to be offered to industrial collaborators.

Techno-economic models enhance our approach to designing robust systems with sustainability metrics. Developing and running process simulation tools can be costly and computationally intensive, limiting their broader applicability. An automated machine-learning approach was used to develop surrogate models based on conventional process simulation models and demonstrated for several high-value biofuels and bioproducts with well-established processes from lignocellulosic feedstocks. Two new publicly available web-based software tools funded by the BioEnergy Technologies Office (BETO), Feedstock to Function and Bio-Cradle-to-Grave (BioC2G), enable users to sort through a database of 10,000 potential molecules and then explore potential locations and process configurations for scaling up their production.

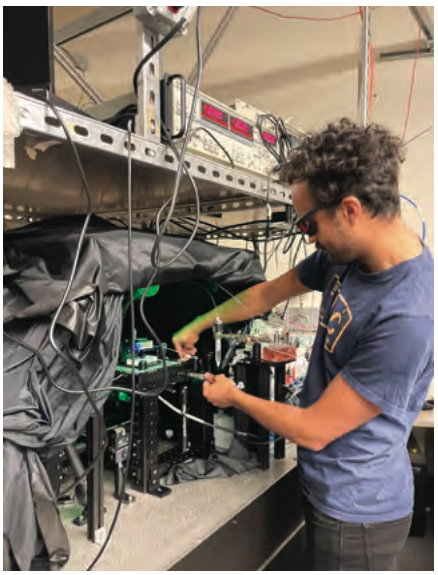


Image credit: courtesy Z. Jones

Zachary Jones is a Postdoctoral Scholar at the ABPDU. Jones led the quantum sensing Berkeley Lab research funding project that implemented nanodiamonds to measure metabolites in bioreactors in real-time. This project required building fluidic systems and analytical instrumentation in order to make measurements on live yeast cells. Jones put together a customized microscope and began proof of concept quantum sensing measurements in microfluidic droplets. Jones's efforts enabled a collaboration with an industry partner to commercialize the technology and branch into several interesting directions.

Nawa Baral, a Project Scientist, and **Tyler Huntington**, a Software Developer, worked together to develop an entirely new machine-learning-enabled surrogate modeling approach for biorefineries. They iterated on sampling strategies for the original process models and auto-machine-learning packages for development of high-performing surrogate models. This has generated enormous interest in the Techno-Economic Analysis community and they are now engaging with other teams across the country to train them in employing this approach to their own processes.

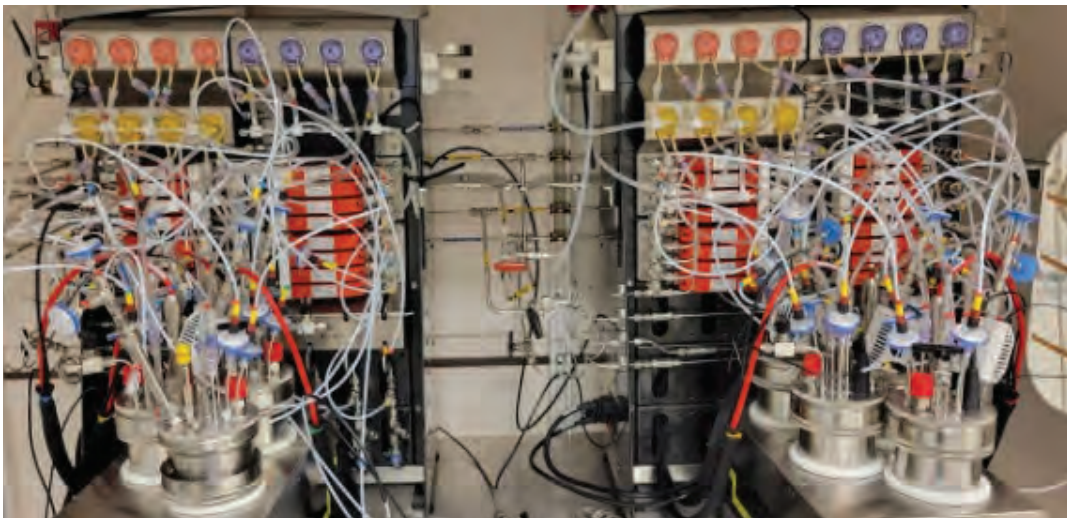


Photo credit: E. Sunstrom

Designed Biological Systems:

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

In the past 10 years, groundbreaking advances in genetic engineering have unlocked entire new microorganisms for biotechnology. Of greatest importance are the tools and applications blossoming from Berkeley Lab and UC Berkeley scientist Jennifer Doudna's and longtime collaborator Emmanuelle Charpentier's 2013 publication describing a new advance in programmable nucleases called CRISPR. This technique promised the ability to edit any genome, but its general applicability was still theoretical. At Berkeley Lab, CRISPR has revolutionized how biological systems are modified and repurposed to suit the experimentalist's constraints. For their contributions to developing the CRISPR technology, Doudna and Charpentier were recognized with the 2020 Nobel Prize in Chemistry. Doudna and colleagues use her discovery for applications in microbial enzyme discovery (JBEI) and microbial community editing (Microbial Community Analysis & Functional Evaluation in Soils, m-CAFEs).

A broader range of host microorganisms are needed to meet our biotechnology challenges and we are developing tools to advance microbial domestication. Our general strategy is to modify environmental microbial isolates that harbor known polyketide gene clusters. This approach is functional in 60 phylogenetically diverse non-model microbial species. In one case study, researchers engineered genes to produce secondary metabolites for health and biomanufacturing applications using Cre-lox-based technology. Another genetic tool for biomanufacturing, mutant fitness information, can iteratively inform strain design. In the lab, our researchers showed a non-model species identified from this screen improved production of a bioproduct from lignin. Furthermore, our researchers demonstrated it is possible to selectively target and modify a specific genetic locus in a particular microbe while it is still in a mixed microbial community, eliminating the need for isolation and purification.

In the last decades, non-model organisms have been used in pilot-scale bioprocess demonstration with promising results. ABF researchers used *Rhodosporidium toluloides*, a nonconventional yeast host, to produce fatty alcohols and observed high titers once they optimized cultivation and processing techniques. At JBEI, researchers paired a multiplex CRISPR-based strategy with computational modeling techniques to demonstrate "growth coupled"

strain design in *Pseudomonas putida*, with final product titers exceeding 40% maximum theoretical yield.

An expanded suite of software tools is needed for designing and predicting microbiomes. The Systems Biology Knowledgebase (KBase), has released new tools that can take a user-provided list of taxa, such as an engineered synthetic community, and determine the traits the community might be capable of containing, based on the (meta)genome information provided. The ABF has developed a Host Onboarding Tool to support the creation of the tiered ranking system to describe the state of microbial engineering. The tool incorporates information about host organisms that have been selected for onboarding and tracks the progress of each organism through the tier system. Additional information about each organism, including publications, associated parts, and experimental data, can be obtained in the tool.

Leah Sloan has worked in Biosciences since 2015. Sloan works in JBEI and the ABPDU, inspired to do her part to combat climate change. As an administrator and program manager, she has skills that are useful in many industries. She is proud that she is on teams working to protect the health of the planet through sustainable biomaterials. Her own career has advanced during her time at Berkeley Lab; she started at the Lab as an Administrator, performing routine operations support functions, and was encouraged to build her skill set with a Berkeley Extension business certification program in project management. She is now a Program Manager overseeing many projects and partnerships. In her role, she gets to witness people from all over the globe also experience growth in the Biosciences Area. Sloan loves watching young Research Assistants and graduate students work with their peers and mentors over the years to build their knowledge and experience. Sloan says, “We are not just developing technology at the Lab; we are building scientific leaders”.

Agnitsch in 2014

Agnitsch in 2023

Mary Agnitsch is a long-time facilities staff member who joined Biosciences in 2006. She keeps our labs humming and our glassware squeaky clean. Biosciences Area research is made possible by Agnitsch and facilities staff members who bring dedication, consistency, and long-term knowledge to keep facilities in working order.

Biodirected Materials and Bionanosciences

Goal: Couple biological components to chemical and physical systems to biosynthesize desired mineral/metal nanostructures

This strategic Goal's target product is organic-inorganic hybrid in nature and we have also had mixed success in meeting our targets. More than half of the milestones set under this strategy were in the area of developing mechanically or electronically biohybrid systems using biological and inorganic material interfaces. This was largely successful in completion (90%), resulting in seven papers published in high-impact journals by the groups of Caroline Ajo-Franklin, Heinz Frei, and Corie Ralston that describe the development of biohybrid electron transfer systems.. However, these research capacities and corresponding Strategic Plan metrics didn't develop as anticipated due to the departure of Ajo-Franklin and the retirement of Frei.

Since the departure of key personnel, we've recently renewed efforts to meet our metrics. New collaborative teams have formed with the aim of understanding the synthesis of naturally occurring inorganic materials. Progress has been made for predictive identification of novel pathways, but lack of funding did not allow further development,

release of software, or publication. Research begun in FY20 and FY21 is underway in developing biological fabrication platforms for synthesis of complex genome-encoded composites and specialized bionanomaterials using diatom frustules. However, multi-omic data has yet to be generated from either of the biomineralization internal investments; more time and funding will be needed to achieve current milestones. Research on calcium oxide crystals commonly found in plant tissues using the bioenergy model plant *Spirodela polyrhiza* (aka duckweed) is also ongoing. The Ercius group has made progress in high-resolution imaging and characterizing of calcareous and silicon biominerals, with an example shown below.

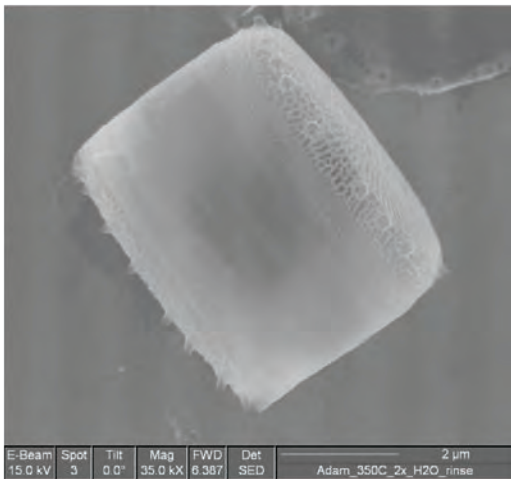
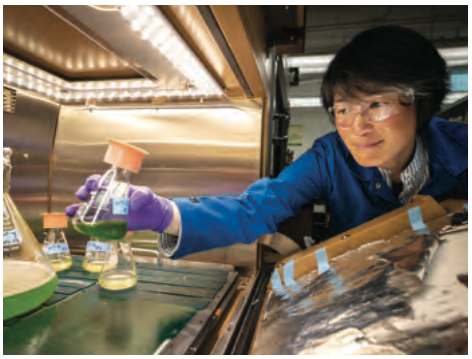


Image credit: Y.-W. Byeon, P. Ercius

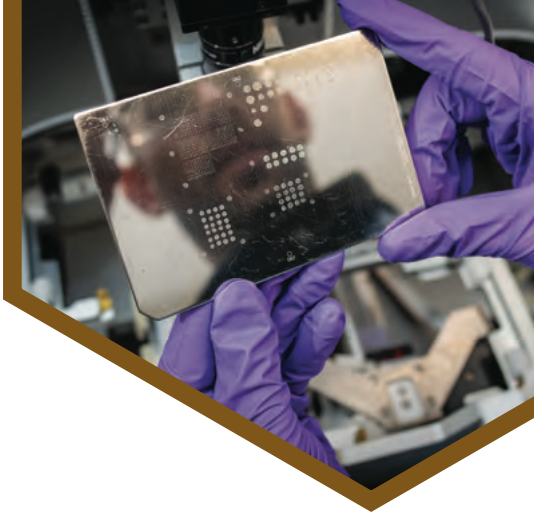
The scanning electron microscopy image, depicting the entire structure of the diatom frustule, reveals biomineral structures that could have applications in biomanufacturing. The Joint Genome Institute (JGI) and the Molecular Foundry made a strategic hire in 2022 to fill the gap between genomes and the emergent properties of nanoscale processes.



Setsuko Wakao, a Biosciences Research Scientist, launched their algal bio-mineralization project through a JGI New Investigator Community Science Program (CSP) that allowed small-scale transcriptome sequencing and preliminary genome evaluation for a new biomineralizing alga, Synura. Its success has led to the awarding of a full CSP, in which more genomes and transcriptomes from algae spanning a branch of silica biomineralizing of stramenopiles will be sequenced. The data generated from this project will serve as a resource to be mined for biomineralization genes for identifying candidate genes for engineering silica structures in these algae.

Crysten Blaby-Haas was recruited as a joint hire between JGI and Molecular Foundry to bridge nanoscience and genomics. Her expertise in genomics, bioinformatics, prediction and character-ization of protein function will not only be invaluable for this challenge and making Berkeley Lab highly competitive in the biomaterials field. Her additional experience in program development (Quantitative Plant Science Initiative at Brookhaven National Lab) will be invaluable for this challenge.


91



TECHNOLOGIES

10-year Goal: Develop a technology infrastructure to measure, predict and control biological systems for solving energy, environmental, and health challenges

Structural Biology

Assessment

Bioscience researchers have contributed to revolutionary technology developments in the United States, rising to solve today's national problems. An important element of our technology portfolio is based on X-ray science, leveraging the Advanced Light Source (ALS) synchrotron and the associated structural biology knowledge on data collection and analysis. Biosciences scientists have also implemented self-driving automation to increase efficiency and enabled remote data collection during the COVID-19 pandemic when site access was restricted.

Biosciences has successfully advanced in two of its three milestones, creating methods for analyzing structural and functional data from the molecular to the cellular level and for taking data structure inputs from synchrotron X-ray and electron microscopy. Correlating data from different sources and integrating functional genomics with structural biology methods is an active area of research, which now takes advantage of advances in machine-learning structure prediction algorithms. Biosciences researchers, in collaboration with the ALS and computational scientists, are in position to leverage current advances in machine-learning algorithms and the anticipated ALS upgrade to correlate data from different sources and integrate structure with functional genomics. Innovations in self-driving automation for data collection and improved data analysis pipelines will provide cutting-edge user facilities and software that will help solve the nation's scientific problems in the biological sciences and health.

10-YEAR
Technologies Goal

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

Technologies for Biosciences to Achieve Goal

Structural biology

Bioimaging across scales

Functional Genomics

Mathematics, Informatics, and Computing

Small Angle X-ray Scattering (SAXS) for Dynamic Structures

The Structurally Integrated BiologY for Life Sciences (SIBYLS) SAXS beamline is a global leader in capturing dynamic motions and correlating SAXS to atomic models. Empowering drug screening and optimization of RNA-based vaccine delivery systems, the beamline scientists have innovatively 3D-printed sample cells to accelerate high-throughput data collection. SIBYLS now has a new data collection mode for challenging proteins, connecting protein purification inline with data collection. With increased efficiency, less material required, and improved data quality, SIBYLS is poised to take advantage of the advances in structure prediction accuracy and increase its applicability to a wide range of biological problems.

X-ray Free Electron Laser (XFEL) Technology and Methods Development

In the last eight years, XFEL emerged with the promise of revolutionizing crystallographic data collection in the study of protein structures. The powerful capabilities of the XFEL also led to multiple challenges in sample preparation, data collection, and data analysis. Biosciences researchers developed hardware and software that not only resolved these issues, but also applied them to several important enzyme systems—leading, for example, to insights into how plants transform the power of the sun into biological energy. This work also opened up new opportunities for small molecule crystallography, with implications ranging from metabolites to drug discovery, to spur scientists to collaborate with the University of Connecticut through Advanced Scientific Computing Research (ASCR) funding.

Protein Crystallography of Membrane Proteins

The Berkeley Center for Structural Biology (BCSB) team built the Gemini beamline, which brings high flux and microbeam capability to the ALS, supported by internal, industrial, academic, and medical nonprofit funds. This achievement clears the way for transformational research of membrane proteins. Crystals of membrane proteins are very small, making data collection difficult. Microbeams are required to focus the X-rays only on the crystal to detect diffraction data. Self-driving, automated crystal screening, to pinpoint the exact position of small crystals, is another major capability development implemented by Biosciences researchers.

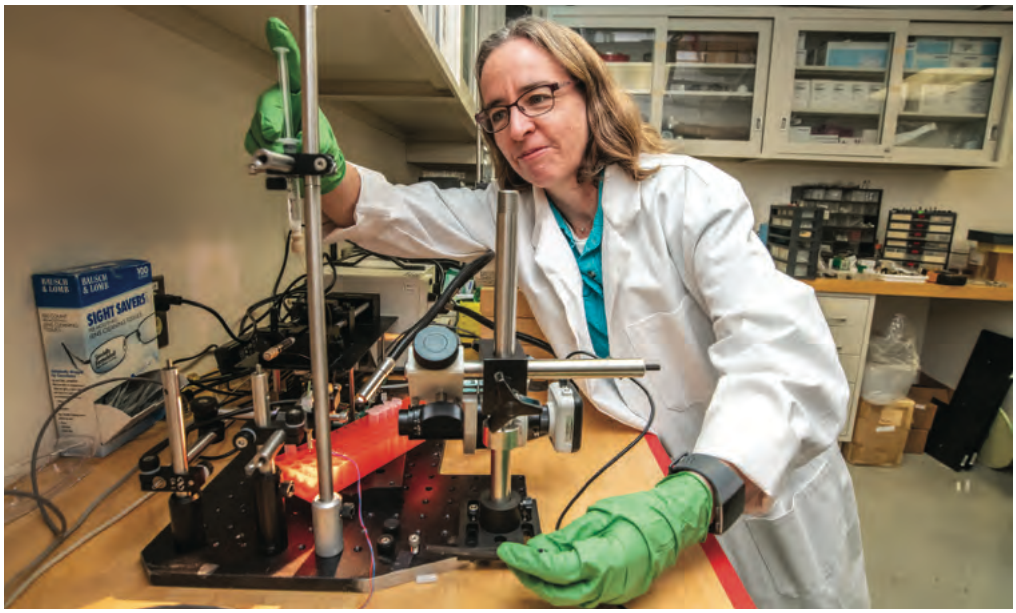
X-ray Tomography for Cellular Analysis

The National Center for X-ray Tomography (NCXT) is the only synchrotron facility in the world that enables quantitative imaging of structures within whole cells through the lens of X-ray absorption, at higher resolution than typical light microscopes. NCXT scientists have added correlated fluorescence imaging, a transformative feature that allows localiza-

tion of specific proteins to organelles and other cellular structures, enabling testing of structure-function relationships in intact cells. For example, NCXT scientists identified novel connections within the chromatin organization in the nucleus, with implications for replication and regulation.

Cryo-electron Microscopy (Cryo-EM)

The Biosciences Phenix team made a tremendous contribution to this field, as it emerged from low-resolution shapes to high-resolution atomic information. They created methods to not only improve the accuracy of atomic models but also ensure the rigor with developments and implementation of quality and identification of model errors.



X-ray Footprinting Capability for Solution Structure

X-ray footprinting, a structural biology method not as well-known as macromolecular crystallography, has been used to probe the protein surface and characterize structural allosteric mechanisms. The Biosciences runs one of only two X-ray footprinting beamlines in the world, and the past ten years have been focused at ramping up the throughput and capabilities at Berkeley, the program led by **Corie Ralston**. Opening up opportunities for remote and even automated data collection, these developments have increased usage over ten-fold. Significantly, implementation of fluorescence in line with X-ray footprinting enables multi-modal measurements of structure.

Bioimaging Across Scales

Assessment

Bioimaging has delivered new technologies for visualizing biological phenomena across time and space. Innovation and technological developments in this central pillar are outlined in the Strategic Plan and the Goals have largely been met. These include faster instrumentation with higher resolution, methods for label-free and non-destructive imaging that preserves biological integrity, bright probes, new analysis pipelines, and methods to register multi-scale images. These developments emerge hand-in-hand with new computational tools and machine learning. Better resolution microscopes, brighter probes, integrated computational methods and automation allow next-generation technology to link structure and function to the emergence of phenotypes from biological systems.

Internal investment has served a prominent role in advancing innovations in multiple areas: a deep-learning framework for a sensorless microscope aberration recognition and correction system; 3D single-cell mapping to characterize carbon cycling activity in the rhizosphere; third-wave machine learning for the biosciences; interactive machine learning for tomogram segmentation and annotation; imagining a root system for plant-microbe interactions; high-resolution optical microscopy; noninvasive and nondestructive infrared and Raman imaging; improving computational speed from programming, statistical, and mathematical methods. All achievements have applications across disciplines.

The bioimaging pillar provides a springboard for developing faster and better imaging strategies. Building on these innovations, we envision new instrumentation that leverages the power of novel machine learning, powerful new computational methods, and automation for rapid high-resolution and high-content information across scales, time, and space. Automation in some cases required novel hardware inventions that optimized sample measurements. For example, coupling IR (micrometer) and fluorescence (nanometer) by new instrumentation integrates non-invasive imaging and allows probing in the same sample over a tenfold range of resolution. These advances will be coupled with continued development of brighter probes and new chemistries

The goal of a virtual bioimaging center remains a work in progress. Unlike the ALS and electromagnetic facilities, which provide a common home and dedicated facilities for structural work (crystallography, cryo-EM, and soft X-rays), optical imaging is diverse, multifaceted, and often does not share common instrumentation. Yet, the talent in this area is enormous and the capabilities to be developed would benefit Berkeley Lab. Advances in bioimaging would be enhanced by facilitating shared resources in a more centralized manner in collaboration with computational scientists, imaging biologists, and engineers.

Speed Scaling in Multiphoton Fluorescence Microscopy

Next-generation optical microscopy methods allow for higher resolution, greater depth, and faster time scales. Imaging technologies were developed to understand neural circuit computation in the visual pathways using the mouse primary visual cortex and superior colliculus as model systems.

Photothermal Instrumentation Linking Infrared and Optical Imaging

Photothermal infrared imaging allows for dual analysis of Fourier transform infrared and fluorescence imaging in the same sample. This technique has already been demonstrated to be 100x more sensitive than other modes of optical photothermal infrared and has a spatial resolution limited by the diffraction-limited spot of the visible fluorescence of the sample rather than the much larger diffraction-limited spots of the infrared excitation. Moreover, because the detected probe is fluorescence, the infrared image and fluorescence image are perfectly correlated with one another, allowing the possibility of creating multiple fluorescent probes in a biological sample and probing infrared spectra of subcellular components.

Advanced Molecular Imaging using Positron Emission Tomography (PET)/ Computed Tomography (CT)

This methodology combines *in silico* modeling of tracer kinetics with lifetime image reconstruction for time-of-flight (TOF) PET.

Relating Cellular Signaling Timescales to Single-molecule Kinetics in First-passage Time Analysis

Combination of techniques in optical microscopy and spectroscopy with materials fabrication methods and cell biology enables the direct observation and physical manipulation of living reaction systems, down to the single molecule level.

Non-linear Optical Spectroscopy

This methodology combines *in silico* modeling of tracer kinetics and lifetime image reconstruction for TOF PET.

Development of Photodetectors and Probes for Imaging Techniques Significantly Contributed to Increased Sensitivity

Open, flexible, scalable, and high-performance electronics systems have been applied in radionuclide imaging; that is called OpenPET. In addition, the development of various probes for imaging has advanced the field. That includes: synthetic nanoparticle-polymer conjugates, genetically-encoded photoactivatable calcium sensors, novel PET probes, and transforming rhodamine dyes — a new chemical tool to integrate advanced microscopy, spectroscopy, and cell biology.

Computational tools

MirroRx technology has been developed as a platform for using infrared light to image cells and to differentiate and predict brain disease and cancer without the use of labels. A deep-learning framework has been applied for a sensorless microscope aberration recognition and correction system.

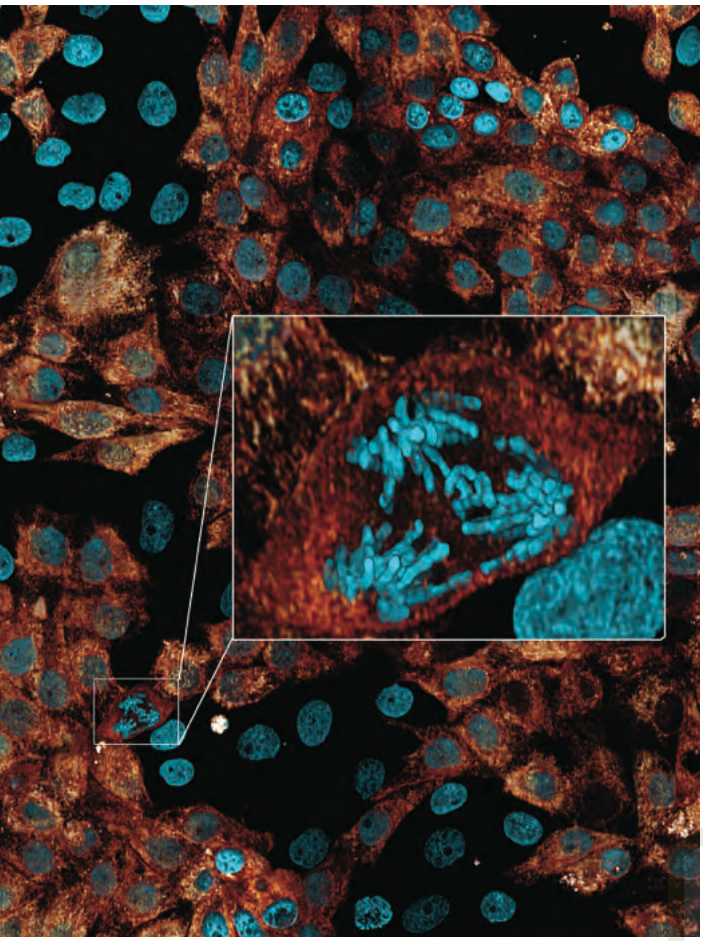


Image credit: G. Upadhyayula

The image is a projection of a 3D volume showing a cancer cell undergoing three-way division. Normal cell division results in two daughter cells, and three-way (or higher) divisions are incredibly rare events, even for cancer cells. Large field-of-view imaging using the Multimodal Optical System with Adaptive Imaging Correction (MOSAIC) makes it possible to catch and visualize these rare events in extreme detail. The endoplasmic reticulum is in orange color, and the chromosomes are in cyan.

Aris Polyzos (pictured center), is an expert in optical imaging. He headed the Lifesciences Advanced Imaging Center at Berkeley Lab and provided guidance in the use of confocal and superresolution microscopy. Based on his expertise, he was invited to lead the microscopy portion in a module of the Cold Spring Harbor course on Single Cell analysis in 2013-2015. Polyzos played an essential role as part of a team of scientists, including **Cynthia McMurray** (pictured right) to develop MirroRx technology, which was a recipient of the 2022 R&D100 award.
Image credit: courtesy C. McMurray

Functional Genomics

Assessment

Technological advances that reveal how different phenotypes emerge from genomes, and tools that enable genetic modifications to produce specific phenotypes, are needed to comprehend — and eventually precisely control—biological systems. For the past 10 years, the Biosciences Area has developed new strategies for understanding the functions of bacterial, fungal, and plant genes; exploring interactions in complex communities, including between plants and their microbiomes; and enabling synthetic biology approaches for modifying organisms to tailor their activities. These technological developments include molecular biology methods, computational analysis tools, and new instrumentation pipelines.

New functional genomics tools have become essential components of several Biosciences projects, and have helped secure new and continuing funding for the Area. In addition, capabilities such as high-throughput metabolomics, BONCAT+FACS, and SIP metagenomics are now available to scientists worldwide through the Joint Genome Institute (JGI). The Biosciences Area delivered new technologies for all seven of the functional genomic targets outlined in the 10-year Strategic Plan.

Random Barcode Transposon-site Sequencing (RB-Tn-Seq)

A high-throughput approach for functional screening of bacterial and fungal gene knockout mutants which allows for mass analysis. Analyzing dozens of microbial isolates in hundreds of culturing conditions using this 'loss-of-function' methodology has provided insights into the functional roles of thousands of genes.

Clustered Regularly Interspaced Short Palindromic Repeats activation/ interference (CRISPRa/i) functional screening.

A synthetic biology process that enables the generation of large CRISPRa and CRISPRi libraries to activate and inactivate genes in bacteria and viruses. The pipeline is enhanced by guide RNA and Sequence Extraction Tool (gRNA-SeqRET), an online tool for the genome-wide design of guide RNAs for CRISPRa/i applications.

Double Barcoded Shotgun Expression Library Sequencing (Dub-seq)

A gain-of-function approach for exploring attributes of exogenous DNA sequences introduced into host microbes. This method provided phenotypic information on hundreds of E. coli genes, including resistance to various types of phage.

DNA-editing All-in-one RNA-guided CRISPR-Cas Transposase (DART) and environmental transformation sequencing (ET-Seq)

This method combines tools to target genome editing of bacteria within complex communities. *In situ* genome editing of specific microbial species will enable precise microbiome manipulations.

Bioorthogonal non-canonical amino acid tagging plus flow sorting (BONCAT+FACS) and stable isotope probing (SIP) metagenomics

These are tools that link in situ functional activity and genome sequences of uncultivated microbes.

High-throughput metabolomics pipeline

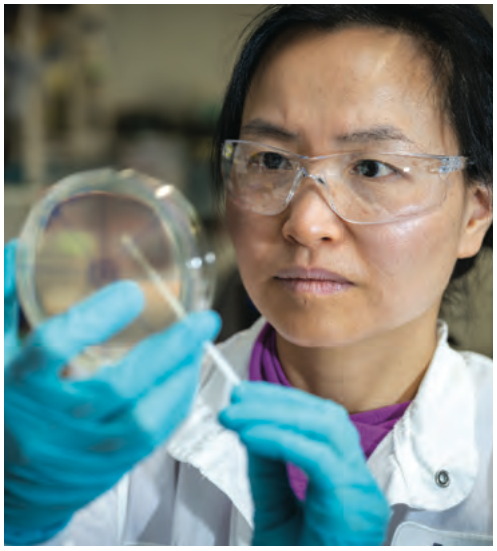
This infrastructure for processing >10K samples per year provides direct measurements of the molecules produced and consumed in biological systems. These measurements are enhanced by the computational tool Metabolites and Gene Integration, which establishes links between detected metabolites and gene annotations.

JGI Plant Gene Atlas

This resource captures and analyzes thousands of RNA-seq datasets to assign functional descriptors to the unknown genes in 17 DOE flagship plant species.

Single-cell transcriptomics (scRNA-seq)

The technique reveals tissue-specific gene expression patterns obscured in traditional RNA-seq methods, including distinct expression patterns driven by plant-microbe interactions.



Transposon mutagenesis is a powerful approach for studying bacterial gene function, and when coupled with next-generation-sequencing (e.g. RB-Tn-seq), it enables high-throughput genotype–phenotype mapping. However, transposon mutagenesis systems must often be modified to work in different species, and finding an effective system through trial and error can be time-consuming and expensive. Rather than testing different systems one at a time, **Hualan Liu** developed a strategy named "magic pools" for simultaneously testing hundreds of different transposon delivery vectors. This approach, developed as part of her postdoctoral research, dramatically reduced the time and cost to identify effective transposon systems for different bacteria. The magic pools method has been adopted by a number of academic and industrial labs where it's been used to study a wide variety of bacterial species. After publishing her work on magic pools, Hualan joined the JGI as a Research Scientist in the synthetic biology group where she continues to develop new functional genomic technologies.

Mathematics, Informatics, and Computing

Assessment

Mathematics, informatics, and computing play a crucial role in the evolution of biology from a descriptive science to a quantitative discipline that relies on and is driven by data. Advanced data analysis tools form the basis of modern genomic research, drug development, and the understanding of cells, tissues, and microbial communities. The Biosciences Area has a history of combining biology and analytics and has developed many technologies used worldwide. Entities such as the JGI, Systems Biology Knowledgebase (KBase), National Microbiome Data Collaborative (NMDC), and the Computational Crystallography Initiative (CCI), among others, are creating new computational methods for various subdisciplines of biosciences research.

It is clear that the integration of artificial intelligence/ machine learning (AI/ML) into the biological and life sciences has been transformative, leading to more efficient processes, improved results, and a greater understanding of complex biological systems. For example, autonomous experiments, in conjunction with robotics and advanced hardware, have streamlined the process of scientific discovery by allowing for the automation of tedious and time-consuming tasks. Solving the protein folding problem, which has been a major challenge in the field of structural biology for decades, has opened up new avenues for drug discovery and potentially transformed the way structures are determined using experimental methods. Advanced image processing methods have improved our ability to analyze large amounts of data, leading to new insights into the workings of biological systems.

We have accomplished over 80% of our Strategic Plan metrics. However, since our last refresh in 2019, the field of AI/ML has undergone major changes that have had a strong impact on our research. Significant investments in the advancement of AI and ML have led to a number of groundbreaking developments, such as autonomous experiments; the de novo design of pharmaceuticals; the solution to the long-standing protein folding problem; and the implementation of advanced image processing methods. Ultimately, the purpose of new mathematics, algorithms, and computing resources, including AI/ML advances, is to provide the scientific community with the tools and resources it needs to push the boundaries of our understanding of biology.

Advancing energy and environmental sciences

JGI, KBase, and NMDC combine tools and services for the wider research public to advance and streamline research. JGI's computing strategy includes increasing algorithmic and pipeline efficiency and transitioning to exascale-aware hardware and software solutions. KBase offers an open science platform for collaborative analytics that provides access to public databases, JGI data and tools, and National Energy Research Scientific Computing Center (NERSC) high-performance computation. KBase enables researchers to study, model, and predict gene, protein, organism, and environmental interactions, leading to better understanding, engineering, and control of biological systems from the micro to ecosystem scales. NMDC offers microbiome researchers access to standardized multi-omics microbiome data and standardized bioinformatic workflows, underpinned by an ontological metadata schema for data interoperability and reuse.

Collaboration with the Center for Advanced Mathematics for Energy Research Applications (CAMERA)

CAMERA at Berkeley Lab, jointly funded by Advanced Scientific Computing Research (ASCR) and Basic Energy Science (BES) in 2012, has been working closely with Biosciences researchers. CAMERA's focus is on developing new mathematics and algorithms to improve current methods or enable the analysis of novel experimental techniques. The Computational Biosciences Group, a collaboration between the Computing Sciences and Biosciences Areas, uses data analytics, statistical machine learning, high-performance computing, and theory/mechanistic models to address energy, environmental, and health issues.

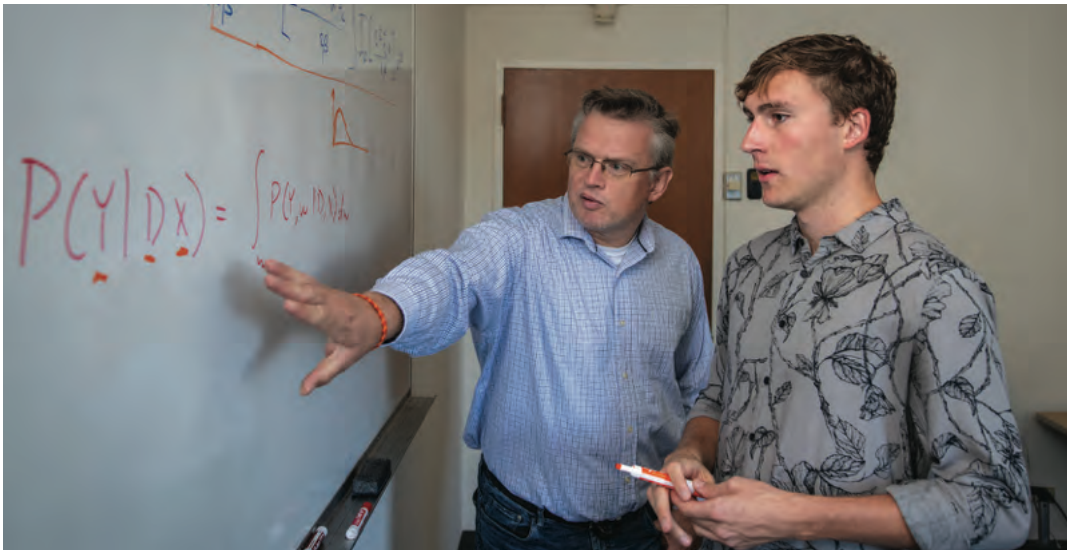
Autonomous Laboratories

The development of methodologies for autonomous laboratories is a major breakthrough, achieved through the collaborative efforts of several groups at Berkeley Lab. This ground-breaking work has advanced the field, highlighting the importance of interdisciplinary collaboration in achieving significant results. CAMERA has been working in collaboration with the Biosciences Area to develop innovative tools for autonomous experiments at light sources. This collaboration has resulted in an orders-of-magnitude increase in efficiency, thanks to the implementation of a Gaussian processing approach. This

approach has proven to be highly effective. At the Joint BioEnergy Institute (JBEI), a similar package developed specifically to design experiments in synthetic biology demonstrated a competitive edge in optimizing strains for the biological production of specific chemicals.

Working with large facilities

Methods are being developed to maximize information from weak data, including experiments from free-electron lasers, electron microscopes, super-resolution microscopy, and other techniques. New algorithms that integrate multiple types of experimental information and experimental data with simulations will be critical for bridging the gap between small/fast and large/slow processes of biological phenomena. Automated knowledge generation from complex data and interpretable statistical machine-learning models will be the most significant breakthroughs.



gpCAM, a software package that enables autonomous discovery, was developed by CAMERA researchers led by Marcus Noack.
gpCAM has been successfully deployed at the Berkeley Synchrotron Infrared Structural Biology (BSISB) Imaging Program. BSISB is an infrared beamline funded by DOE's Office of Biological and Environmental Research at the Advanced Light Source, a BES user facility at Berkeley Lab, led by **Hoi-Ying Holman** (pictured above). "*gpCAM* was designed to assist decision-making processes in any kind of experiment, with minimal effort," Noack says. With *gpCAM*, an experiment that used to take over 8 hours can now be completed within 30 minutes. "This is a game changer," says co-author and Berkeley Lab staff scientist **Petrus (Peter) Zwart** (pictured left), who worked with a team to further develop gpCAM for biogeochemical and geobiological applications.

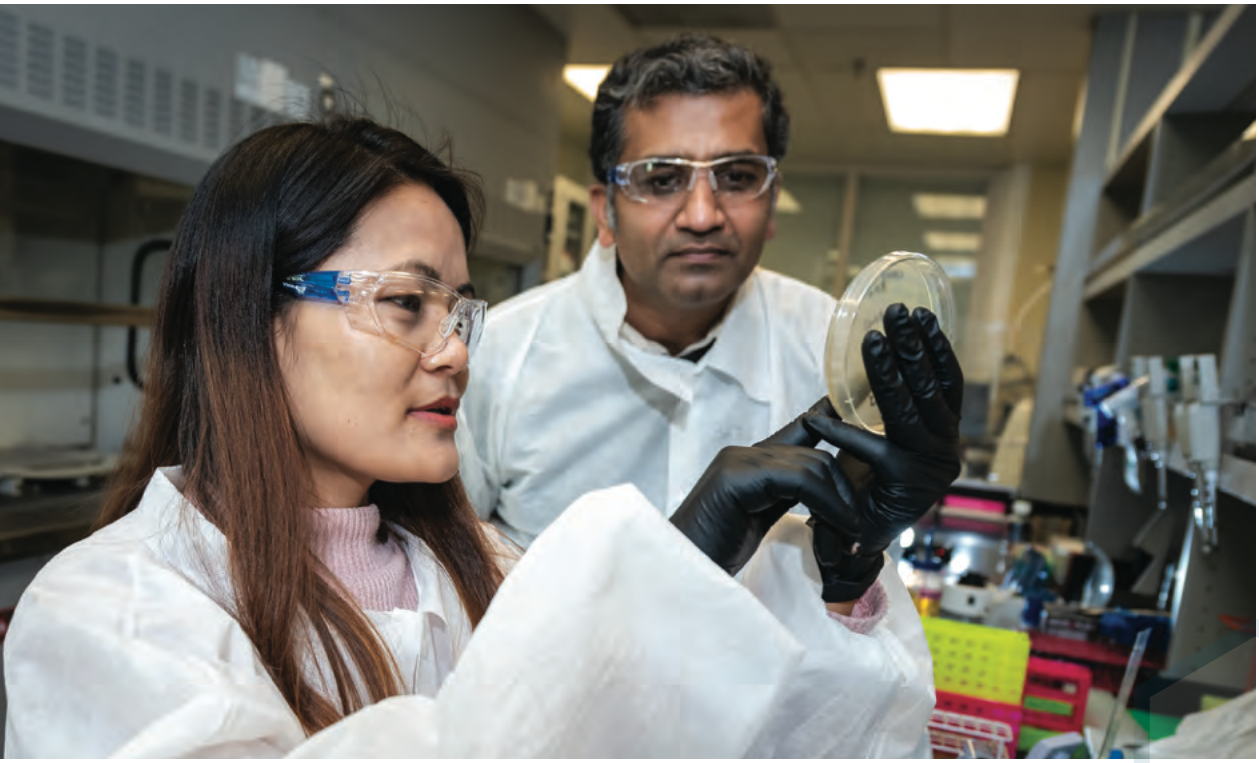
LOOKING FORWARD FOR BIOSCIENCES

Berkeley Lab Biosciences is not an organization to rest on its laurels. As we look forward to our next decade, we see important domains of growth for the Area. Some are continued efforts from our 2013–23 plan and others are building on nascent efforts. We see opportunities in Research, Capabilities, and People and Community.

Research

Climate change continues to be a top challenge. Our approach includes a replacement of fossil fuels with biologically enhanced products that can nest the generation of fuels and products with other environment-improving or industrially relevant services. Further, we confirm and renew our commitment to understanding carbon cycling and sequestration in natural and built ecosystems, such as agricultural and urban lands

We also see the importance of biosecurity and biopreparedness for both natural and engineered threats. This could include the identification of a threat type, monitoring, and mitigation applied in environmental health, human health, and chemicals, materials, and resource use. We are enthusiastic about the federal support for emerging research areas including soil carbon enhancement, bio-based aviation fuels, and increased computing for biosciences.



Capabilities

In order to advance our research goals, we see that gains in computation will advance biology. This includes thoughtful changes to how data is recorded for downstream processing (Findable, Accessible, Interoperable, and Reusable (FAIR) principles) through fully self-driving labs. Additionally, we believe that closer collaboration with our computer science colleagues will drive innovation in both fields.

Our final capability goal is to have a single hub of Biosciences through the completion of the Bayview campus. Co-location of all of our business units will improve connection and collaboration. The currently leased buildings limit the range of research (e.g., ceiling height for large-scale equipment, requirements for gas and chemical safety). Finally, the proposed Biological Genome Engineering and Manufacturing (BioGEM) building includes classroom space that would be a physical home to our deep commitment to training the next generation of science, technology, engineering, and math (STEM) professionals.

People and Community

We recognize that our success in achieving our strategic goals is tightly linked to the success of our people and our community. Team-based research with an interdisciplinary approach is the science of the present and the future. The complexity of the biological world and the deep research questions requires vastly disparate types of expertise and backgrounds. We see the need to not only retain and hire personnel in all fields of biology, management, engineering, and operations, but also to form strong relationships with our colleagues in other scientific fields. We know that group success is closely tied to respect and trust and we want to develop strategies and approaches to unite people across their differences.

Part of that approach is to incorporate outreach activities and initiatives more deeply into our 2023–33 goals. We have both a professional and ethical commitment to increasing access to government energy and scientific resources to everyone in America. Our targets include training for the next generation of STEM professionals; partnering with institutions historically underrepresented in federally funded research programs; and increasing visibility and access for communities who have been historically ignored in outreach. Additionally, we have a strong commitment to our local community. This manifests in a variety of ways, including: place-based research that can deepen Lab-region bonds through outreach; local challenges that have national implications (such as wetlands research, environmental impacts on health, or soil contamination); and advances that could enhance industry.

BIOSCIENCES SUMMARY REPORT: 2013–2023 GOAL COMPLETION APPENDIX

	Key:
√	Goal Achieved
+	In Progress
x	Deprioritized

Energy

Lignocellulosic Biofuels

√	• Understand fundamental elements of plant biology that underlie biomass yields and adaptation to stress by focusing on key gaps in knowledge of primary plant physiology and in plant-microbe interaction relevant to plant resilience ◦ Elucidate how the secondary cell wall of bioenergy crops is synthesized ◦ Understand mechanisms of water conservation and drought tolerance in plants ◦ Advance the understanding of how nitrogen fixation occurs in model plants. ◦ Understand the molecular interactions between plants and beneficial microbes and their role in nutrient acquisition and stress tolerance ◦ Engineer bioenergy crops for reduced inputs (water and nitrogen/phosphate fertilizers), enhanced tolerance to stress, improved sugar yields, and facilitated production of useful compounds from lignin, including fuels ◦ Engineer plants and microbes to stimulate nitrogen fixation ◦ Engineer model plants and crops to grow with less water ◦ Engineer lignin in biomass crops to be an economically useful polymer ◦ Increase the C6 to C5 sugar ratio
+	• Develop predictive models that will facilitate engineering of optimized secondary cell wall synthesis and saccharification by specific genetic manipulations

✓	• Develop deconstruction processes that enable efficient utilization of all plant biomass components (cellulose, hemicellulose, lignin) ◦ Fractionate biomass so the individual components can be separated into different streams ◦ Develop pretreatment technologies that allow for direct conversion of all the plant polymer components in one bioreactor
+	• Develop inexpensive methods to depolymerize plant polymers to intermediates for microbial conversion ◦ Develop a suite of enzymatic cocktails that depolymerize cellulose, hemicellulose, and lignin without product inhibition and inhibition from contaminants resulting from the deconstruction process
✓	◦ Discover chemical catalysts that selectively depolymerize plant polymers to monomers that can be converted to fuels by fermentation • Understand the mechanisms of lignocellulose deconstruction by microbial communities ◦ Probe the variety of biomass degradation strategies employed by microbes in natural environments ◦ Establish a synthetic microbial community that deconstructs one lignocellulosic substrate
+	◦ Understand synergies between fungi and bacteria in a mixed eukaryote-prokaryote microbial community ◦ Define the pathways for deconstruction and metabolism of lignin in one synthetic or adapted microbial community
✓	• Engineer microorganisms tailored for consumption of deconstructed biomass (sugars and aromatic compounds) and production of drop-in biofuels at high yield ◦ Develop metabolic pathways for production of hydrocarbons with fuel properties equivalent to those found in petroleum-based gasoline, diesel, and jet fuels, and engineer these pathways into diverse hosts
+	◦ Use metabolic engineering to generate microorganisms that rapidly convert multiple substrates to desired end products simultaneously
✓	◦ Develop machine learning approaches to optimize metabolic flux in engineered organisms
+	◦ Develop genetic tools, including biosensors, and machine learning algorithms to enable multi-gene engineering in diverse hosts to rapidly engineer complex genetic traits such as fuel yield and stress tolerance
✓	◦ Engineer microbes to tolerate or detoxify biomass deconstruction inhibitors and fuel products, and other stresses introduced during bioreactor scale-up
+	◦ Engineer microorganisms to produce fuels under anaerobic conditions
x	◦ Engineer fuel-producing microorganisms whose communities are resistant to invasion by contaminating organisms and that are genetically constrained to growth in industrially defined conditions.

√	◦ Develop and demonstrate bench- and pilot-scale processes for fuel production
+	◦ Identify, implement, and prototype new unit operations for more flexible and efficient production of precursor biomolecules and finished fuel derivative

Alternative Biofuels

√	• Use a systems-level understanding to identify key bottlenecks that will limit fuel production in the metabolism of photosynthetic, methanotrophic, and chemosynthetic microbes ◦ Identify key regulatory elements that sense and respond to the presence of gas feedstocks in potential fuel production hosts
+	◦ Develop computational methods to predict the metabolism of C1-converting microbes
x	◦ Combine computational predictions and high-throughput automation to develop and predict beneficial changes in metabolism using machine learning
√	• Engineer photosynthetic, methanotrophic, and chemosynthetic microbes to produce fuels from gas feedstocks. ◦ Develop a synthetic biology toolbox leveraging broader methods being developed in the Biomanufacturing Goals (transformation, promoters, genome integration of heterologous pathways) for at least three C1 conversion hosts ◦ Produce energy-dense biofuels from gas feedstocks in a photoautotrophic, methanotrophic, or chemoautotrophic host ◦ Improve the total carbon conversion in host microbe C1 feedstock use by 10%, which will approach theoretical maximums for substrate conversion
+	• Scale fuel production from gas feedstocks through bioreactor and process development ◦ Demonstrate production of an energy-dense fuel from a gas feedstock at pre-pilot scale ◦ Design bioreactors in collaboration with Berkeley Lab researchers that maximize C1 conversion to fuel by improving mass transfer ◦ Integrate product separation with C1 conversion to improve fuel yield

Artificial and Engineered Photosynthesis

√	• Develop scalable artificial photosynthetic systems based on bio-inspired engineered and biohybrid approaches that produce energy-dense liquid fuels from light energy and CO2, well beyond CO or formate currently attainable • Couple predictive models with advanced nanoscale engineering to improve artificial photosynthesis • Improve the biomass productivity of natural photosynthetic systems • Use new levels of understanding of photosynthesis to predict ways it can be improved in plants and microbes

X	• Redeploy the photosynthetic apparatus in a previously non-photosynthetic host with superior coverage of the solar spectrum, improved conversion efficiency to fuel products, and durability
√	• 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 in conjunction with the development and application of advanced characterization methods. That includes using time-resolved X-ray crystallography and X-ray spectroscopy at X-ray free electron lasers, and multidimensional X-ray spectroscopy at the synchrotron and X-ray free electron laser facilities, in combination with time-resolved 2D electronic and vibrational spectroscopies

Environment

Predictive Understanding of Environmental Organisms

√	• Use model ecosystems (e.g., EcoFABs and EcoPODs) to discover the functions of genes and metabolites that mediate microbial interactions
+	◦ Discover the molecular mechanisms by which microbial communities in groundwater and sediment are altered by anthropogenic contaminants
√	◦ Enable data-driven prediction of gene functions for thousands of microbes and metagenomes per year and tens to hundreds of plants ◦ Develop methods for transferring knowledge about specific biomolecules, organisms and ecosystems to related systems so that exploration of new biology is accelerated, and to better understand the behavior of complex ecosystems ◦ Characterize the metabolism, stress responses, and interactions of diverse environmental microbial isolates (such as fungi, algae, bacteria, archaea, and viruses) from relevant field sites using a multi-omics approach including sequencing, mutant phenotyping, and genome engineering
+	◦ Develop a multi-modal and multi-scale computational infrastructure to accurately and rapidly predict microbial metabolism, gene regulation, and stress response for microorganisms in key environments, such as: contaminated sediments, soils, wetlands, deserts, agroecosystems, and grasslands ◦ Integrate diverse high-throughput data types, including genomics, transcriptomics, imaging, epigenomic, and metabolomics approaches and computational learning technologies to discover the roles of novel genes, proteins, regulatory sequences, and metabolites in plant and metazoan responses to environmental change ◦ Use multi-omics measurements in controlled ecosystems of model plants and bioenergy crops to understand the molecular processes through which plants interact with, select, and maintain beneficial bacteria and fungi to improve plant fitness

+ ◦ Pioneer the development and application of *in situ* and *in vivo* sensors that provide spatially and temporally defined measurements of cellular responses to environmental changes and, secondly, link it to synthetic processors that amplify the adaptive molecular signals enhancing organismal fitness

Molecular Ecosystems Bbiology-based Solutions

√
- Create and validate a laboratory-based model plant-soil-microorganism ecosystem (an "EcoFAB") to investigate the molecular basis of microbiome-driven plant growth promotion
- Implement and further optimize in situ sensor technology to explore temporal and spatial heterogeneity and dynamics of the plant-microbe-soil-atmosphere system from lab to field scales
- Develop, test and refine model predictions of native community dynamics to mechanistically account for the material and energy flow at contaminated field sites
- Use model ecosystems, along with genome and community editing, to identify the key microbial functions and cultivated and uncultivated organisms that stabilize microbial communities, and design interventions that decrease energy and nutrient inputs necessary to achieve key ecosystem service goals
- Develop and apply computational learning strategies that integrate detailed imaging, genomic, and metabolic data from model ecological models to accurately predict nutrient cycling and/or biotic interactions in complex ecosystems

+
- Understand how to harness plants and their microbiomes to achieve a quantitative increase in soil organic carbon storage based on computational predictions from simulated ecosystem experiments

√
- Demonstrate the ability to design rhizosphere communities within fabricated ecosystems that improve the low-input growth of an important bioenergy crop
- Identify new metabolites exuded by plants that help select and maintain beneficial microbes, which, in turn, improve the low-input productivity of plant biomass production
- Develop computational learning tools to enable mechanistic predictions and discoveries from the integrative analysis of complex, multimodal, multi-scale data collected on artificial and natural ecosystems—including deep learning and causal inference modalities suitable for low-sample regimes

Environment

Multiscale Understanding of Human Biology

√
- Develop and apply large-scale functional genomics technologies to elucidate genome functions relevant to human biology, evolution, and disease
  - ◦ Understand how noncoding sequence variation impacts human development, evolution, phenotypic variation, and disease susceptibility through integrative

	use of genomic, computational, and model organism strategies
√	◦ Develop and harness innovative CRISPR-Cas genome editing strategies in model organisms to inform structural and functional analyses of the human genome
	• Integrate multiple approaches to understand brain functioning in health and disease
	◦ Elucidate the cellular basis of the signals measured by human brain monitoring technologies (e.g., electrocorticography)
+	◦ Understand how distributed cortical circuits are coordinated to give rise to complex perceptions and behaviors
√	◦ Create mathematical frameworks to recover networks that give rise to neural population dynamics
+	◦ Use combined epigenetic profiling and model organisms to link psychiatric genetic findings to cell type-specific gene regulation
	• Use of computational methods to translate knowledge from model organisms and model systems for predictive disease diagnosis, prevention, therapy, and risk management in order to advance human health and bioresilience:
	◦ Developing methods to predict the pathogenicity of variants in the human genome using machine learning
√	◦ Develop computational learning frameworks to discover the epistatic architecture of complex traits from population data
+	◦ Predicting causal associations between genomic variants and cellular and whole- organism phenotypes, leveraging experimental data from model organisms
X	◦ Developing causal models of gene function in the context of metabolic and signaling pathways in human cells and organ systems
√	◦ Predicting relationship between exposome and bioresilience of humans and other organisms ◦ Develop transfer learning strategies to translate and unify knowledge across model systems

Biological responses to environmental challenges

√	• Identify the keystone components that mediate the impact of environmental challenges (prioritized based on human epidemiological studies and model system discoveries) on molecules, cells, microbial communities, tissues and organisms in tractable model systems
	◦ Phenotypic responses to environmental challenges exhibited by cells, microbial communities, tissues, and organisms ◦ Dynamic responses of molecules (DNA, RNA, protein, metabolites and their extracellular and tissue fluxes) and phenotypes using advanced imaging, genomics, phenomics and computational approaches

✓
- Determine how the effects of environmental challenges are modulated by genetic, physiological and epigenetic variation in model systems
  - Genetic mapping to identify DNA sequence differences that affect individual responses to environmental challenges
  - Cellular and tissue damage responses that vary by physiological status (gender, age, diet, etc.)
- Elucidate the role of prototypic community interactions within biological systems, and how they are reciprocally affected by environmental challenges
  - Model host/microbial community composition and functional responses to environmental challenges that identify conserved and species-specific mechanisms
  - Identifying microbial communities that impact fitness and define their interactions with each other and hosts
  - Identifying changes in brain function and behavior due to reciprocal flux of gut-microbial and metabolites along the gut-blood-brain axis
  - Demonstrate the ability to manipulate reciprocal interactions between microbes and model organisms that produce benefits to fitness in response to environmental challenges
- Determine if environmental challenges shown to impact model organisms similarly affect human cells and populations and tissues
  - Validation of human biomimetic tissue culture systems (e.g. breast, skin) fabricated from normal primary human cells and extracellular matrices, with respect to relevance to *in vivo* tissues at the levels of architecture and gene and protein expression patterns
  - Quantifying the biological effects of exposures (identified as having impact in model systems) on human biomimetic tissues
  - Integrating discoveries about specific environmental challenges into exposome/epidemiological studies in human populations
  - Initiation of strategies for disease prevention, therapy, and risk management based on individual predispositions and systemic responses to environmental challenges
- Integrate mechanistic insights from human and model system studies to develop computational models to predict the effects of environmental challenges on human health
  - Generating a list of bioindicators for human health and disease that includes genes, epigenetic markers, proteins, metabolites and their fluxes, microbiome, and physiological components
  - Through applying standardized ontologies and workflows, link human microbiome data to environmental microbiome data within NMDC to enable cross-study investigations and modeling to extend understanding of microbial processes
  - Successfully predicting how manipulating host and microbiome properties positively or negatively impact human bioresilience

√	• 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 ◦ Deploy deep learning algorithms that integrate omics, phenotypic and exposome data that more effectively identify sensitive populations and biomarkers ◦ Developing personalized therapeutic and prevention strategies by utilizing links between genetic, microbiome composition, and environmental responses

Biomanufacturing

Tools to Design, Construct, and Debug Biology

√	• Develop a BioCAD/CAM infrastructure comprising tools for: ◦ Pathway retrosynthesis and host engineering for optimized production titers, rates, and yields under industrially-relevant conditions ◦ Integrating functional genomics data into the design process ◦ Learning from characterization results to inform the design process
+	• Develop simulation and control capabilities for biomanufacturing systems ◦ Small-scale (< 2 L) physical and computational simulations of large-scale (> 100 L) reactors to aid in strain optimization ◦ Sensors to measure spatially different conditions in a bioreactor, for application in the computational simulations ◦ New control modalities (physico-chemical or metabolic) for real-time optimization of fermentation performance
√	• Develop detailed and extensible techno-economic models for manufacturing processes and applications ◦ Model-informed selection of strategies for synthesis and characterization
+	◦ Holistic consideration of bioprocesses, from feedstock to reactor to downstream processing, integrating chemical and biological process steps, and the path to production at scale
X	• Develop retrosynthetic infrastructure ◦ Establish optimized routes to 30 key retrosynthetic molecular intermediates ◦ Create automated workflows to test retrosynthetic computational predictions

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 functional genomics analyses of 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 platforms for assessing the environmental impact of engineered organisms • Nominate a range of key hosts for biomanufacturing, agriculture, environmental remediation, water support, and human health ◦ A host database serving bacterial information and protocols related to cultivation and transformation
x	◦ Determine active and potentially tamable hosts compatible with key application environments
√	◦ Identify host properties that can be engineered for improved robustness, productivity, and optimized energy throughput
+	◦ Develop robust genetic toolkits for each host: expressing and integrating plasmids, media recipes and transformation protocols, characterized regulatory parts, selection and screening markers • Establish robust tools for plant engineering ◦ A host database containing genome information (size, ploidy, gene expression, etc.) and protocols related to cultivation and transformation
√	◦ Development and characterization of “universal” plant expression tool kits
+	◦ Generation of public part libraries and associated databases
√	◦ Development of a pipeline for rapid and efficient trait stacking • Develop and demonstrate bench- and pilot-scale fermentation processes for novel bioproducts ◦ Quantify the impact of gene expression and metabolism on the scale-up of biomanufacturing fermentations ◦ Identify, prototype, and deploy new protocols for more flexible and efficient production of precursor biomolecules and bioproducts • Develop infrastructure to design and engineer microbiomes ◦ Tools to construct microbiomes to a given functional or compositional specification ◦ Algorithms to predict expected microbiome behavior

Biodirected Materials and Bionanosciences

√	• Invent new routes for the design of biohybrid systems that mechanically or electronically interface active biological elements with polymeric and inorganic materials ◦ Demonstrate electron transfer between intracellular and extracellular redox active species along a molecularly defined path

√	◦ Create at least two molecularly defined pathways that operate at different redox potentials • Achieve ability to interface biological components to electronic apparatus to control their activity ◦ Develop methods to electronically modulate catabolic fluxes ◦ Develop methods to electronically stimulate intracellular reactions ◦ Interconvert electrical energy with chemical and/or light energy • Build infrastructure to understand charge transfer at the abiotic/biotic interface ◦ Develop methods to structurally characterize redox active molecules at the abiotic/biotic interface
x	◦ Develop spectroscopic methods to characterize charge transfer and dynamics energetics ◦ Create basic framework relating structure, energetic and dynamics of charge transfer
+	• Discover the biological basis underlying the biosynthesis of naturally-occurring inorganic materials and inorganic-organic composites ◦ Develop computational pipelines for mining of genomes and metagenomes for genes involved in the transport, biosynthesis, and assembly of inorganic biomaterials ◦ Build out a data platform for predictive methods to identify novel pathways and biomaterials ◦ Conduct multi-omics studies, at the unicellular and nanostructural levels, on native biological systems that produce or modify inorganic materials to develop our understanding of the plant and microbial processes that direct the synthesis, transport, modification, assembly, and storage of inorganic biominerals • Achieve ability to biosynthesize architecturally specified, possibly self-healing, mineral/metal nanostructures and mesostructures on demand, using biological entities ◦ Develop computational systems biology and biodesign tools for a systems-level understanding and forward engineering of inorganic material synthesisDevelop synthetic biology tools for engineering biomineralization including controlling transport, synthesis, spatial patterning, and timing ◦ Invent new technologies to support high-throughput or massively parallel determinations of the function of inorganic material biosynthetic pathways ◦ Intentionally align structural and functional tools to characterize inorganic biominerals

Foundational Technology Development

Structural Biology

√	• Constitution of methods for correlating and analyzing structural and functional data from the molecular to the cellular level
	• Creation of computational analysis programs capable of data structure input from synchrotron X-ray and electron microscopy methods
	• Development of methods to integrate advances in functional genomics with structural biology methods

Bioimaging across scales

+	• Establishment of a center for integrated bioimaging
√	• Availability of probes, labeling chemistries, and label and label-free approaches that provide contrast across multiple imaging modalities
	• Creation of visualization, modeling, and interaction systems for experimenters to efficiently extract knowledge from data
	• Development of new sample preparation methods compatible with multi-model imaging

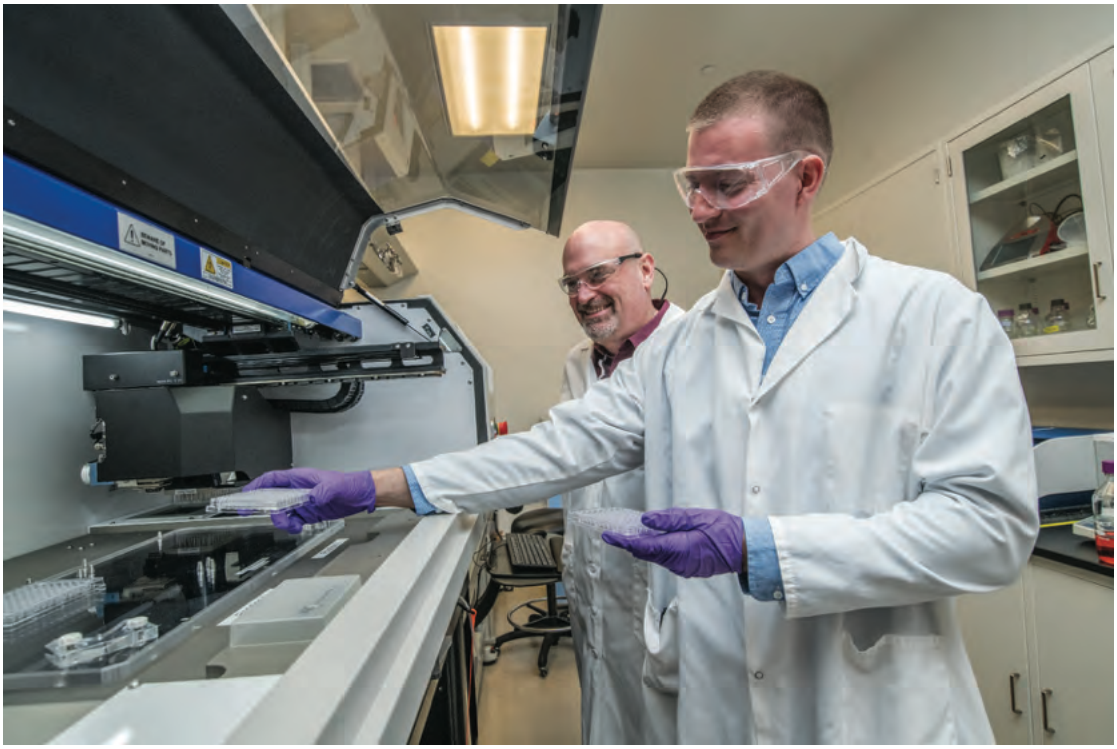
Functional Genomics

+	• Demonstrate technologies for rapid phenotyping, especially microfluidic chip-based automation and mass spectrometers with enhanced analytical chemistry
√	• Vastly improve performance of metabolomics for energy, materials, and environmental research
	• Develop and apply at-scale, new genetic approaches to determine the function of genes from diverse organisms (bacteria, archaea, viruses, fungi, algae, plants)
	• Pioneered community editing technologies (CRISPR, phage, optogenetic, etc.) to determine the functions of cells, organisms, proteins, pathways, and metabolites within specific environmental contexts
+	• Develop, integrate, and apply new nondestructive modalities with functional genomic measurements to gain insights into functional genomic processes within relevant spatiotemporal contexts and scales
√	• Develope isotope-based methods, coupled to systems biology tools, for *in situ* characterization of important metabolic processes
+	• Demonstrate utilization of genetically targeted, *in vivo* cellular manipulation techniques (e.g., optogenetics) to control biological function

Mathematics, Informatics, and Computing

√	• Improve mathematical and computational approaches for addresses challenging problems in structural biology, genomics, bioinformatics and multiscale modeling

√	• The development of novel mathematical and computational approaches enabling new scientific approaches or solving outstanding problems in biophysics, biochemistry and genomics • In collaboration with NERSC and the Computing Research Division, redesign existing computational approaches that can make optimal use of novel hardware, such as exascale computers • The development of mathematical and algorithmic approaches that allows the joint analyses of multimodal methods, operating on length scales from atoms to cells to organisms
+	• Innovation of dynamical systems approaches to understand biological time-series data towards revealing the governing equations of biological systems
√	• Creation and application of statistical machine learning methods to extract small sets of features predictive of complex biological phenomena to enable understanding and causal interventions • An increase in the integration of data analytics to drive the design and adaptive control of experiments in the Biosciences

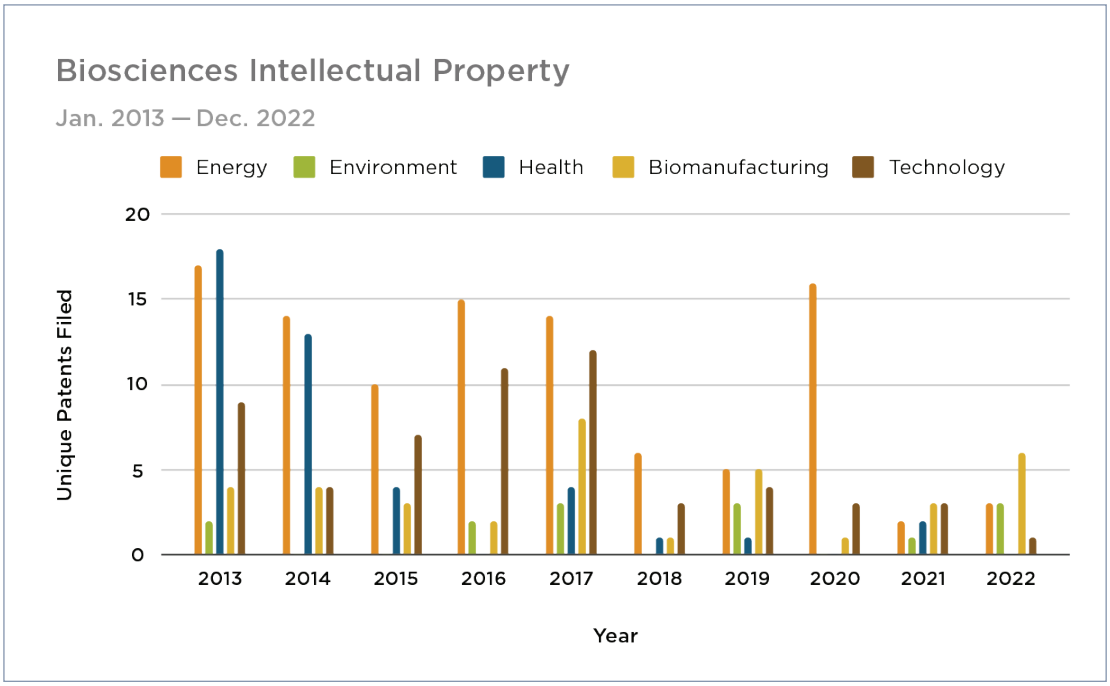


Technology Transition and Translation Appendix

The Biosciences Area values transitioning our discoveries, inventions, and tools beyond the walls of our laboratory. We have a strong history of filing intellectual property (IP) claims. Our portfolio of IP includes invention and software patents and records of inventions (ROI), pooled by Strategic Plan Pillar below. Many of our IP has been recognized by the R&D100 awards, listed below.

Biosciences R&D100 Awards

- Breaking the Mass Spec Logjam, (2013)
- Virus Power, (2013)
- A Closer Look at 3-D Cell Cultures, (2014)
- Efficient Bioengineering, (2014)
- A Portable Biosensor, (2015)
- Compact Dynamic Beamstop, (2016)
- Double Barcoded Shotgun Expression Library Sequencing (Dub-seq), (2017)
- SolarCatMesh, (2022)
- MirroRx technology, (2022)



Strategic Plan Implementation Team

Staff members from across the Biosciences Area are responsible for the implementation of the Biosciences Strategic Plan, analysis of progress, and the creation of this report. Strategy Leads are experts in related disciplines who coordinate research underlying the strategy. For the 10-year Summary Report, Strategy Leads compiled accomplishments across Berkeley Lab, drafted summaries of progress, and assessed the outlook for achieving the 10-year metrics outlined in our strategic plan. Mentors guided Strategy Leads through the process and provided valuable feedback. The Biosciences staff who have contributed to implementation of the Strategic Plan since 2013 are listed below. We thank all contributors for their dedication to the strategic plan and for their thoughtful effort on this report.

Paul Adams[a]
Caroline Ajo-Franklin
Adam Arkin
Kristofer Bouchard[a]
Susan Celniker[a,c]
Katy Christiansen[a]
Sylvain Costes
Benjamin Cole[a,b]
Adam Deutschbauer
Diane Dickel
Thomas Eng[a,b]
Michael Espinosa[a]
Lida Gifford[a]
Nathan Hillson[a,c]
Nikki Humphreys[a]
Lauren Jabusch[a]
Gary Karpen
Jay Keasling[a]
Jan Kern[a,b]
Mark LaBarge
Greta Lorge[a]
Rex Malmstrom[a,b]
Mary Maxon[a]
Cynthia McMurray[a,b]
Don Medley
Aindrila Mukhopadhyay
Vivek Mutalik
Trent Northen[a,c]
Len Pennachio[a,b]
Corie Ralston
Sarah Richardson
Eddy Rubin
Henrik Scheller[a,b]
Ian Sharp
Blake Simmons[a,c]
Esther Singer
Steven Singer[a,b]
Antoine Snijders[a,b]
Deepti Tanjore[a,b]
Susannah Tringe[a]
Susan Tsutakawa[a,b]
Setsuko Wakao[a,b]
Chia-Lin Wei
Junko Yano[a,c]
Kateryna Zhalnina[a,b]
Petrus (Peter) Zwart[a,b]

[a] Contributor to this report
[b] Current Biosciences Strategic Plan strategy lead
[c] Current Biosciences Strategic Plan mentor

We thank the Berkeley Lab Visual Archive and Thor Swift for images, unless otherwise credited. Gratitude to Berkeley Lab Creative Services for their assistance with layout.

Acronyms and Abbreviations

Federal Organizations and Funding Bodies

DOE Department of Energy

BER Biological and Environmental Research, a program of the Department of Energy Office of Science

BES Basic Energy Science, a program of the Department of Energy Office of Science

ASCR Advanced Scientific Computing Research, a program of the Department of Energy Office of Science

BETO Bioenergy Technologies Office, a program of the Department of Energy Office of Energy Efficiency and Renewable Energy

NIH National Institutes of Health

Berkeley Lab Terms

Berkeley Lab Lawrence Berkeley National Laboratory

ALS Advanced Light Source, a user facility

BELLA Berkeley Lab Laser Accelerator

LiSA Liquid Sunlight Alliance, a science program

IDEA Inclusion Diversity Equity Accountability

NERSC National Energy Research Scientific Computing Center, a user facility

CAMERA Center for Advanced Mathematics for Energy Research Applications, a science center

Internal Investment Also known as Laboratory Directed Research and Development (LDRD) funds

JGI Joint Genome Institute, a user facility

JBEI Joint BioEnergy Institute, a Bioenergy Research Center

BioEPIC Biological and Environmental Program Integration Center, an under-construction building

ENIGMA Ecosystems & Networks Integrated with Genes & Molecular Assemblies, a DOE BER science focus area

m-CAFEs Microbial Community Analysis & Functional Evaluation in Soils, a DOE BER science focus area

ABF Agile BioFoundry, a science program

ABPDU Advanced Biofuels and Bioproducts Process Development Unit, a science program

KBase DOE Systems Biology Knowledgebase, a science program

NMDC National Microbiome Data Collaborative, a science program

BCSB Berkeley Center for Structural Biology, a science program

CCI Computational Crystallography Initiative, a science program

SBDR Structural Cell Biology of DNA Repair Machines, a science program

SIBYLS Structurally Integrated BiologY for Life Sciences, a science program

NCXT National Center for X-ray Tomography, a science center

BSISB Berkeley Synchrotron Infrared Structural Biology, a science program

Technologies and Other Terms

AI/ML Artificial Intelligence/Machine Learning

BioC2G Bio-Cradle-to-Grave

BioCAD/CAM Biological Computer-Aided Design and Manufacturing

BONCAT+FACS Bioorthogonal Non-Canonical Amino acid Tagging plus Flow Sorting

COVID-19 Coronavirus Disease 2019

CRISPR Clustered Regularly Interspaced Short Palindromic Repeats

CRISPRa/i Clustered Regularly Interspaced Short Palindromic Repeats activation/interference

CryoEM Cryo-electron Microscopy

CT Computed Tomography

DART DNA-editing All-in-one RNA-guided CRISPR-Cas Transposase

DIVA Design, Implementation, Validation Automation

Dub-Seq Double Barcoded Shotgun Expression Library Sequencing

EcoFAB Fabricated Ecosystems

EcoPOD "Pilot-Scale" Ecosystems

EDD Experiment Data Depot

ET-Seq Environmental Transformation Sequencing

FAIR Findable, Accessible, Interoperable, and Reusable

FY Fiscal Year

gRNA-SeqRET guideRNA and Sequence Extraction Tool

LCA Life Cycle Assessment

MOSAIC Multimodal Optical System with Adaptive Imaging Correction

PET Positron Emission Tomography

R&D100 Research and Development 100 Award

RB-Tn-Seq Random Barcode Transposon-site Sequencing

SAXS Small Angle X-ray Scattering

scRNA-seq Single-cell RNA Sequencing

SIP Stable Isotope Probing

STEM Science, Technology, Engineering, Mathematics

TEA Techno-Economic Analysis

TOF Time of Flight

UC University of California (various locations)

XFEL X-ray Free Electron Laser

For the latest in Berkeley Lab Biosciences, visit us at

biosciences.lbl.gov

23-BAO-3074