# BIOLOGICAL RESEARCH AT THE DEPARTMENT OF ENERGY: LEVERAGING DOE'S UNIQUE CAPABILITIES TO RESPOND TO THE COVID-19 PANDEMIC

September 11, 2020

A Hearing of the
Subcommittee on Energy
Committee on Science, Space, and Technology
United States House of Representatives

Testimony of
Dr. Mary Maxon
Associate Laboratory Director for Biosciences
Lawrence Berkeley National Laboratory

## Introduction

Chairwoman Johnson, Ranking Member Lucas, Chairwoman Fletcher, Ranking Member Weber, and distinguished members of the committee, thank you for holding this important hearing and for inviting me to testify about the Department of Energy's (DOE) Office of Science's biology research and development enterprise. And, thank you for your strong and unwavering support for science and discovery, and for your commitment to developing the next diverse generation of scientists, engineers, and other critical STEM practitioners.

I am Mary Maxon, and I am the Associate Laboratory Director for Biosciences at Lawrence Berkeley National Laboratory (Berkeley Lab). Berkeley Lab was founded in 1931 and has been managed ever since by the University of California for the DOE. Our researchers develop sustainable energy and environmental solutions, create useful new materials, advance the frontiers of computing, and probe the mysteries of life, matter, and the universe. Over 14,000 scientists from around the world rely on the Lab's national scientific user facilities for their own discovery science.

My testimony represents my views only and does not represent the views or positions of the Office of Science or of the Department of Energy.

Throughout my career – first as a microbiology student and as a researcher within academia and industry, later as a policymaker in government, and today in research management – I've witnessed the progress made in biosciences. It has been extraordinary. Things possible today would have flabbergasted me as a young Ph.D. candidate at the University of California. Among the important facts that I did not know as a young student, but that I know now, is that many of the advances in biology would not have happened without the DOE's Office of Science and without our national laboratories.  Although understood by too few, DOE's Office of Biological and Environmental Research is critical to the nation's innovation engine and to its international economic competitiveness.  I will strive today to describe just how important DOE's role is in driving biological innovation, addressing challenges in energy and environmental sustainability, and building the nation's biobased economy.

Today, my testimony will attempt to tell this story by addressing:

1. *How early, mission-driven federal investments created the DOE biology enterprise we benefit from today;*
2. *How DOE and the nation is leveraging the investments made in facilities and human expertise to advance today's mission priorities; and*
3. *How the Office of Science Biological and Environmental Research program is delivering solutions to society, including responding to national crises such as the Coronavirus pandemic, and helping to drive the nation's bioeconomy.*

In overview, the foundation of DOE's biosciences capabilities has been and continues today to be built on: investing in early career talent and entrusting them with solving tough problems; bringing researchers, engineers, and operations together in dynamic teams; and inventing novel, unique and useful new technologies that advance discovery and invention -- including scientific user facilities.  These tools are sophisticated, often large in scale or scope, and cost-prohibitive for academia or industry to build and maintain. No other federal agency or international entity can match the Department's record on this score.

Perhaps the most important assets at the national laboratories, however, are the men and women who conduct the research, as well as those who provide the much needed administrative, financial, technical, and health and safety support. Although the narrative of the brilliant solo scientist persists in today's culture, in fact most scientific research is conducted by multi-disciplinary teams of people rather than single investigators. Scientific discovery is fueled by creativity and perseverance, and extraordinary progress is most often made when diverse perspectives allow problems to be seen from a variety of different angles.  Cultivating talent and promoting inclusion and diversity are central to the creation of a successful work environment. Among the national labs, Berkeley Lab was the first to publish its workforce diversity demographics. We know that successful innovation depends on the ability to create a community that brings together people with diverse backgrounds and different approaches to problem-solving.

The Coronavirus Pandemic and the National Virtual Biotechnology Laboratory

The Department's ability to respond effectively to the coronavirus pandemic illustrates the power of and value to the nation of biology research capabilities, expertise, and resources at the national laboratories. With emergency funding from the Congress, the Department of Energy, the Office of Science, including its Biological Environmental Research program (BER) have, pretty much on the fly, organized a highly effective and sophisticated response to the pandemic.

Leveraging the remarkable breadth of expertise, capabilities, and resources from every national laboratory, the Department, under the leadership of Office of Science Director Chris Fall, established the National Virtual Biotechnology Laboratory (NVBL) to coordinate and expedite SARS-CoV-2 and COVID-19 related research across the Office of Science. NVBL is supporting the development of innovations in testing capabilities, identifying new targets for medical therapeutics, providing epidemiological and logistical support, and addressing supply chain bottlenecks by harnessing extensive additive manufacturing capabilities. Examples of this incredible response will be included throughout this testimony and will illustrate the Department's deep bench of biosciences expertise and unique scientific user facilities.

## Biology at DOE: A History

DOE's involvement in biological research has a grand and fascinating history, and there are many incredible examples from across the national laboratory complex. These examples reflect the importance that biology has played for the Department, from understanding and addressing the impacts of radiation exposure for humans and its effects on natural environments, to how biology drives energy solutions and creates new economic opportunities for the nation's bioeconomy. Much of the biology research at Lawrence Berkeley, Oak Ridge, Argonne, Pacific Northwest, Brookhaven, and at other national labs began by tackling the challenges of a post-WWII world and by leveraging the unique resources and expertise that existed because of significant federal investments made during the war and afterward. However, because I represent Berkeley Lab and know its stories best, my testimony will focus on its particular role in the history and future of biological research within the Department of Energy.

The Advent of Nuclear Medicine

A foundational moment in the building of DOE's biological capabilities took place on August 26, 1931, 89 years ago, when the University of California took a big gamble on its then youngest full professor: Ernest Lawrence, the namesake of Berkeley Lab. They built for him a lab where he could scale up the experimental tool he invented and for which he would later win the Nobel Prize – the cyclotron, a circular particle accelerator in which beams of electrons smash into materials at 99.9999% the speed of light to reveal their atomic, molecular and chemical secrets. Lawrence's cyclotrons are the great grandfathers of the Large Hadron Collider and light sources such as the Advanced Photon Source at Argonne and the Advanced Light Source at Berkeley Lab.

As Lawrence's cyclotrons grew and became more powerful, their transformational value for science solidified. The team around them got bigger, too, and included experts from many fields: machinists; chemical, electrical, and civil engineers; electricians; chemists, physicists, and materials scientists; as well as accountants, secretaries, and other administrative staff. Big team science was under way, and a new scientific research paradigm was launched.

Throughout the 1930s, 40s, and 50s, cyclotrons were facilitating the discovery of new elements at an amazing rate – among them technetium, neptunium, and plutonium. (Sixteen elements on the periodic table were discovered at Berkeley Lab – more than any other institution in the world.) Producing new radioactive isotopes became a matter of rote. Lawrence understood the potential of the cyclotron and these products beyond basic physics. He turned to family, his younger brother John, an M.D., to explore and capitalize on the power of the cyclotron for biological research – a move that changed modern medicine forever and laid a critical piece of the foundation for DOE's focus on biosciences.

John Lawrence joined his brother at Berkeley in 1935 after teaching at Yale where he studied the effects of radiation on the pituitary gland. John had become interested in the potential use of cyclotron-produced radioisotopes and nuclear radiation in the treatment of cancer when he and a colleague discovered that a beam of neutrons had a much more destructive effect on tumors than an equivalent dose of x-rays. In 1937, John Lawrence used the radioisotope phosphorus-32 to successfully treat polycythemia vera, a bone marrow disorder. In 1939, he used beams of energized neutrons to treat a patient with leukemia, the first treatment of cancer with beams from a particle accelerator. These advances attracted funding to build new cyclotrons and state of the art research labs from renowned institutions and philanthropists such as the National Cancer Institute and steel magnate William Donner, whose son Joseph had died of cancer in 1929.

Although the Manhattan Project dominated Lawrence's lab and cyclotrons for a time, John, his Berkeley colleagues, and researchers at other national labs continued to build the field of nuclear medicine and advance other medical processes and treatments during the war and over the following decades. Among their many accomplishments were:

- the use of particle beams to successfully treat acromegaly (the abnormal growth of the hands, feet, and face) and Cushing's disease;
- establishing world leading research capabilities across several national labs focused on the effects of radiation on humans;
- the invention of the Anger Camera – a revolutionary medical imaging tool invented by Hal Anger in 1957 - that is still the predominant nuclear medicine imaging machine;
- identifying the roles that HDL and LDL cholesterol may play in heart disease[1]; and
- the discovery and production at Brookhaven National Laboratory of the isotope technetium-99m (Tc-99m). Used to create images from inside the human body, Tc-99m is used in tens of millions of medical diagnostic procedures annually, making it the most commonly used medical radioisotope in the world.[2]

Although DOE ended its direct role in funding a broad program of nuclear medicine research, it continues to fund important radioisotope research, development, and production as part of a key federal program that produces isotopes that are otherwise unavailable or in short supply for U.S. science, medicine, and industry. Additionally, health research-focused organizations like the National Institutes of Health and the Howard Hughes Medical Institute fund important research and facilities at the national laboratories to advance their mission objectives, including the development and operation of specialized biomedical beamlines at light sources.

Unlocking the Mysteries of Photosynthesis

Even during the heady days of advances in medicine at the national laboratories, biological research wasn't just limited to human health. Presaging today's DOE biosciences portfolio, which is primarily focused on non-human biology, Ernest Lawrence is reputed to have told biochemist Melvin Calvin in September 1956 that it was "time to do something useful with radioactive carbon." Scientists understood the potential use of one of Carbon's isotopes, Carbon-14, as a "tracer," a substance that can be tagged to organic molecules and used to follow those molecules through different stages of a chemical process. Understanding the mysteries of photosynthesis was the objective.

Using carbon-14 as the tracer, Calvin organized a team of researchers to map the path that carbon takes through a plant during photosynthesis. Over the next few years, in the process of creating their map, Calvin and his team showed that sunlight acts on the chlorophyll in a plant to fuel growth, rather than on carbon dioxide as was previously believed. For leading the research that deciphered the photosynthetic process, Calvin received the 1961 Nobel Prize in chemistry. The Calvin-Benson-Bassham Cycle we learn about in school bears his name.

[1] https://academic.oup.com/jn/article/128/2/439S/4724045
[2] https://www.bnl.gov/newsroom/news.php?a=24796

## **DOE's Office of Science Biological and Environmental Research: The Modern Era**

At the beginning of the national laboratory system, and as evidenced at the creation of the Atomic Energy Commission (AEC) in 1946, DOE's mission imperatives and the national laboratories' capabilities in biological research and development were well understood within government. The enabling legislation for the AEC specifically called out the "utilization of fissionable and radioactive materials for medical, biological, and health purposes." Congressional direction continued to support this view, or slight variations on it, in reauthorization bills over the years. The inclusion of biosciences within the repertoire of the Department was codified again, with a particular focus on energy and the environment, in the legislation that created the Department of Energy in 1977 and in subsequent legislation ever since. The Engineering Biology Research and Development Act sponsored by this Committee and passed by the House on December 9, 2019, is another recent example of Congressional intent that the Department play a critical role in biosciences for the nation.

With roots stretching back to the beginning of the national laboratory complex, the Office of Sciences' Biological and Environmental Research program (BER) is today one of the world's leading supporters of non-human biological research and development – that is, the biology of microbes, plants, and other non-human organisms. As described below in more detail, the BER of today supports world leading scientific user facilities and funds cutting edge biosciences research activities at the labs and at universities across the nation.

### Sequencing the Human Genome

Because of BER's deep experience in managing biological research and its leadership role in driving large, interdisciplinary initiatives such as designing and building particle accelerators, it was no surprise that the nation turned to DOE, and later to the National Institutes of Health (NIH), when setting eyes on the daunting task of sequencing the human genome. Deep understanding within the Congress of the national labs' capabilities in addressing this grand challenge helped as well.

BER had begun seriously considering the Department's interest in deciphering the human genome as early as 1984 as it became clearer that new analytical tools and scientific methods were needed to advance mission-focused research in biological mutations caused by radiation. A DOE-sponsored meeting in December 1984, later dubbed the Alta Conference for its Utah location, convened leading geneticists of the day to discuss new approaches to genomics research and to begin to map a path forward. BER followed on with additional conversations, workshops, and road-mapping and eventually by funding more genomics research across its laboratories. DOE and the national labs were more than ready when the national human genome initiative took off.

Although there were skeptics of DOE's role in the project and a lot of jockeying within the Administration and the Congress for control and funding, the Department and NIH signed a Memorandum of Understanding in 1988 about how to move forward collaboratively. With strong and united support within the Congress and the Administration, the federal Human Genome Project, with a goal of determining the exact order of the DNA letters and the genes encoded in the human genetic blueprint, was established in 1990.[3]

DOE's part of the Human Genome Project was eventually focused on a collaboration among the three University of California managed national laboratories, Lawrence Berkeley, Lawrence Livermore, and Los Alamos and their genome research laboratories. This collaboration was formerly organized as the Joint Genome Institute (JGI) in 1997. The significant economies of scale achieved by bringing these efforts together under one roof enabled the JGI to be the first to publish the sequence analysis of the target chromosomes 5, 16, and 19, in the journal *Nature*. JGI's contribution represented 13% of the total Human Genome Project.

With the completion of the Human Genome Project in 2003, leaders at DOE and within the Congress and the Administration understood that the specialized tools, expertise, and capabilities of JGI could be powerfully leveraged to address grand challenges and grand opportunities in energy solutions and environmental sustainability. So, in 2004, the Department established JGI as a national user facility, similar to the light and neutron sources at national labs that this Committee has examined in previous hearings, to focus on non-human health genomics grand challenges and opportunities.

The Joint Genome Institute

JGI, now a Berkeley Lab-managed user facility, has grown to serve a community of 2,000 users annually from almost every state and from around the world. Like the other national user facilities, researchers submit proposals for using the JGI and are awarded access through a peer-reviewed competitive process. In addition to direct users, there are nearly 18,500 data users worldwide who in 2019 downloaded 2.16 million files and a total of 400 terabytes of data. The vast majority of JGI proposals are supported under the auspices of its Community Science Program (CSP) to characterize organisms (plants and microbes) relevant to the DOE science mission areas of bioenergy, global carbon cycling, and biogeochemistry.

---

[3] https://plato.stanford.edu/entries/human-genome/

The JGI further extends its capabilities through the "Facilities Integrating Collaborations for User Science" (FICUS) initiative. FICUS enables a "one-stop shopping" approach whereby an applicant, in one proposal submission, can request multiple complementary technical resources at JGI and in partnership with the Environmental Molecular Sciences Laboratory (EMSL). EMSL is another national user facility supported by the BER located at Pacific Northwest National Laboratory (PNNL). Through FICUS a user could access metabolomics technologies along with protein characterization and imaging technologies at EMSL and DNA sequencing and synthesis resources at the JGI. The JGI-EMSL FICUS collaboration, now in its eight year, has supported over 70 joint proposals. Owing to this success, the FICUS initiative has been extended to include the National Energy Research Scientific Computing center to focus on data analysis projects.

Over 15 years, the JGI transformed from an organization that was responsible for one project (sequencing its three human chromosomes) to a national user facility completing over 43,500 projects and generating over 2,100 peer-reviewed scientific publications. Here are some selected examples from the diversity of projects that JGI enables.

**Revolutionized our understanding of ecosystems.** Using its powerful DNA sequencing capacity and novel computational analysis strategies, JGI has shed light on the interactions in numerous complex ecosystems, enabling comparisons of nutrient cycling in agricultural and prairie soils and drought tolerance of plants grown under these different conditions. In addition, providing the means to gauge greenhouse gas emissions produced by microbes in Arctic permafrost and Bay Area wetlands, and determining the microbes capable of coping with oil spills and natural seeps in the Gulf of Mexico.

**Cataloging the diversity of life**. There are more microbes in a handful of soil than there are stars in the Milky Way, but we know far less about the former. In 2007, JGI began working to change that by establishing GEBA, a worldwide Genome Encyclopedia of Bacteria and Archaea. Currently, the information gathered by the GEBA initiative along with others has funneled into the Integrative Microbial Genomes (IMG) data system that currently has over 61 billion genes and counting. This repository represents a springboard for adding new branches to the microbial tree of life and understanding the function of the newly characterized species in our environment.

**Coronavirus genomes.** With the goal of developing computational and machine learning/artificial intelligence approaches to study the evolutionary patterns of SARS-CoV-2, JGI researchers are using their genomics expertise to analyze over 80,000 coronavirus genomes to investigate genome similarities and differences. Knowledge gained through this research may have useful purposes in addressing the coronavirus pandemic, including diagnostics, vaccines, and possible treatments. JGI received NVBL support for exceptional operations to continue research during the pandemic this past year

**Sequencing plant species.** The poplar was the first tree ever sequenced. This was done at the JGI in 2006. We now have the genetic sequences of hundreds of different plant species that form the foundational knowledge needed to develop bioenergy crops with desired traits. This genomic knowledge is being used by the DOE Bioenergy Research Centers (BRCs) to develop optimal plant species and conversion processes for the efficient production of biofuels and bioproducts that can bolster and catalyze growth in the US bioeconomy. The BRCs are among the JGI's largest dedicated partnerships.

The Bioenergy Research Centers

Established by BER in 2007, the BRCs conduct basic science but with "use inspired" goals and objectives. Simply put, the BRCs' holy grail is to develop viable integrated approaches for the production of advanced biofuels and bioproducts from biomass that are cost equivalent to and/or better performing than existing alternatives. Jay Keasling, the CEO of the Joint BioEnergy Institute (JBEI), the Berkeley Lab led BRC, says that the ultimate goal is to use every single carbon atom present in the biomass feedstock, be it found in switchgrass, poplar, sorghum, or pine, to make affordable and scalable biofuels and bioproducts. Achieving this would create significant new economic opportunities for American farmers and for the nation's rural communities and catalyze rapid growth in the US bioeconomy.

There are four BRCs, they are:

**Center for Advanced Bioenergy and Bioproducts Innovation (CABBI)**, led by the University of Illinois at Urbana-Champaign.
**Center for Bioenergy Innovation (CBI)**, led by Oak Ridge National Laboratory
**Great Lakes Bioenergy Research Center (GLBRC)**, led by the University of Wisconsin—Madison in partnership with Michigan State University.
**Joint BioEnergy Institute (JBEI)**, led by DOE's Lawrence Berkeley National Laboratory

Each center has a unique focus, yet they work collaboratively on shared scientific and technology challenges, and in syncing up with industry to stay focused on scientific avenues that will most likely scale and ultimately become commercially transferable.

The BRCs' record of accomplishment speaks for itself. Since 2007 they have collectively published 3,424 academic papers, disclosed 728 inventions, and filed 521 patent applications.

Nineteen companies have spun out of the BRCs, including twelve from JBEI. Since its founding, JBEI has contributed many scientific achievements, including:

- *Engineering bioenergy crops to increase sugar-containing polymers and decrease lignin in plant cell walls*
- *Developing a feedstock agnostic pretreatment technology based on ionic liquids that is affordable and scalable*
- *Developing optimized microbial routes for the conversion of biomass-derived sugars and lignin-derived intermediates into advanced, "drop-in" blendstocks for gasoline, diesel, and jet fuels*

Here are some selected examples from the diversity of projects that JBEI has enabled:

**Lower-cost, higher-performing biofuel.** In terms of meeting the cost and performance goals, one example is the cost of isopentenol, a leading advanced drop-in biofuel candidate. Ten years ago, the cost of one gallon produced in the lab was approximately $300,000. Today, it's closer to $3.00 per gallon. Isopentenol has an energy density close to that of gasoline and burns more efficiently than ethanol in certain combustion engines. It can also be chemically converted into an advanced aviation biofuel that enables more efficient jet engine combustion that can extend the range of airplanes. Performance has also improved for other leading biofuel candidates.

**Better adjuvants for possible Coronavirus vaccines.** Adjuvants are used in some vaccines to help create a stronger immune response. Adjuvants help vaccines work better. Many vaccine adjuvants are extracted from plants and are in short supply. Berkeley Lab scientists are working with a pharmaceutical company to engineer yeast to produce a very effective vaccine adjuvant so that there will be a sufficient supply of the adjuvant for the billions of doses that will be needed to address the coronavirus pandemic. The yeast that will produce the vaccine is a variant of the yeast that JBEI originally engineered to produce a variety of biofuels.

**"Bio-advantaged" chemicals.** Lygos, one of JBEI's twelve spin-off companies, produces "bio-advantaged" chemicals, where chemicals from petroleum cannot compete with those produced through biochemistry. JBEI scientific discoveries and technology development underpin much of Lygos' capabilities. The company is currently commercializing its first product, malonic acid, and its platform for synthetic biology is being exploited to develop a number of new products. Malonic acid, traditionally made from petroleum feedstocks, is an important precursor to many valuable products. The Lygos malonic acid is not only made from a renewable feedstock, but also employs a biomanufacturing process that results in the production of less hazardous waste and fewer exposures for workers to toxic chemicals.

Still, the BRCs have work to do. A major objective is using their versatile platform technologies and leveraging their expertise in biofuels to accelerate research into high-value biobased chemicals and bioproducts, such as those driving toward commercialization at Lygos. Developing these high-value products can drive economies of scale that could help to simultaneously bring down the costs of biofuels and speed their delivery to consumers. In addition to being an important part of a more sustainable economic model for biofuels, bioproducts have the additional benefit of storing carbon in durable goods and products such as paints, flooring, furniture, recyclable plastics, etc.

**Today and Tomorrow: Bringing Big Science to Biology**

Today, research assets at all of the national laboratories are leveraged for biosciences discovery and invention. These assets are not just limited to those funded by BER, like JGI, the BRCs, EMSL, and others. Biological research at the labs leverage investments made by the Basic Energy Sciences, Fusion Sciences, High Energy Physics, and Advanced Scientific Computing Research programs among others within the DOE Office of Science. Collectively, the expertise, national scientific user facilities, and other capabilities of the Office of Science serve as research and technology development platforms for researchers from academia and industry from throughout the US and the world. They also represent, figuratively and literally, a machine learning ecosystem where results, failures, and successes are shared, learned from, and disseminated.

Imaging Biology to Drive Advances

A leading example that isn't supported by BER funding but rather by the Office of Science's Basic Energy Sciences program are the bioscience capabilities at DOE's light sources. These descendents of Lawrence's cyclotrons are invaluable to the nation's biological research community of academics, industry, and national lab researchers. As the Committee has heard in previous hearings, almost every pharmaceutical treatment on the market today has been developed utilizing light sources to determine the structure of potential drug targets. Collectively, the light sources have determined the structures of tens of thousands of proteins. Leading pharmaceutical companies utilize the Advanced Light Source at Berkeley Lab to develop their candidate drugs, for example. The examples of direct implications for drug development are too numerous to list, but include successful treatments for melanoma and vaccine development for ebola.

In addition, new femtosecond X-ray free electron laser sources, another type of light source, such as the Linear Collider Light Source at SLAC National Laboratory, are now used to collect data from millions of protein crystals per day, resulting in multi-terabyte datasets. For context, a femto-second is one quadrillionth of a second. Yes, it is hard to fathom. These experiments have also been essential in explaining key features of photosynthesis, and will ultimately help lead to improved alternative energy sources.

There are many examples of how these unique resources are being leverage to address the coronavirus pandemic:

**Coronavirus structures.** These key national assets have been put to use during the current pandemic helping scientists discover the structures of many of the 27 individual proteins in the Coronavirus that cause COVID-19. This can improve diagnostics and treatments by identifying weaknesses in structures and opportunities to disrupt the inner workings of the virus that control how it infects, replicates, and spreads -- helping to speed the development of vaccines and therapeutics. The DOE synchrotron light sources received NVBL support for exceptional operations to continue research during the pandemic this past year, partially supporting these efforts.

**In situ protein responses**. Beyond the light sources, BER's investment in fundamental biological science and Pacific Northwest National Laboratory's (PNNL) long-standing expertise in molecular research and computational biology has made it possible to explore protein structure characteristics of the Coronavirus and how they can be disrupted. Researchers at PNNL are using robust computational capabilities and a technique called nuclear magnetic resonance spectroscopy to see a single protein in incredible detail, even looking at specific segments of the protein and monitoring responses of those proteins during experiments.

<u>Automation and Computation</u>

Because of BER’s investments and the national laboratories’ ingenuity, genome sequencing and other biology research processes that once took days, weeks, months, or years now happen in seconds, minutes, hours, or days. Robots perform intricate chemical manipulations by the thousands an hour, replacing the hands-on work of eager postdocs working over days. These tools quickly weed out the combinatorial wheat from the chaff, freeing researchers up for more productive tasks and speeding science along its way.

New imaging and characterization tools now produce terabytes of biological data instead of the gigabytes of previous methods. The revolution in cryo-electron microscopy (using frozen samples in microscopes that use electrons instead of light) for biological systems has led to the routine generation of terabytes of data per microscope per day. For context, one terabyte is equal to a trillion bytes – roughly 1,400 compact discs worth of information. A new detector at the Berkeley Lab Molecular Foundry's National Center for Electron Microscopy (NCEM) enables scientists to record atomic-scale images in microseconds, or millionths of a second – 60 times faster than possible with previous electron detectors. According to Andrew Minor, NCEM facility director, this new detector, known as the "4D Camera" (for Dynamic Diffraction Direct Detector) is the fastest electron detector ever made.

This abundance of data - its exponential growth and complexity - is a big challenge that also presents a great opportunity for scientific and technological advancement. The computing capabilities at the DOE Office of Science's large computing centers supported by the Office of Science's Advanced Scientific Computing Research program – the National Energy Research Scientific Computing Center (NERSC) at Berkeley Lab and the Leadership Computing facilities at Argonne and Oak Ridge national laboratories, interconnected with ASCR's high-speed network facility, the Energy Sciences Network (ESnet) – are tremendous resources for biology. These supercomputers and DOE's deep bench of computational resources, applied mathematics, software development, and networking capabilities are world class and move biosciences forward by leaps and bounds.

The NVBL and the COVID-19 HPC Consortium, a collaboration among industry, academic, and government HPC centers, are aggressively leveraging Office of Science and other DOE computing resources, such as the supercomputers at the NNSA labs, to address the coronavirus pandemic. Examples include:

**Understanding 55 million years of coronaviruses.** To ascertain whether COVID-19 may have characteristics similar to the seasonal flu, researchers at Berkeley Lab have been studying how SARS-CoV-2 fits into the history of coronavirus evolution, about 55 million years of diversification. If ancient coronaviruses have swapped genetic material frequently, there is good reason to believe that they will continue to do so. To do this, scientists are utilizing NERSC to computationally generate billions of possible histories that can each explain the data, and then analyze these scenarios to determine which are most likely to have actually taken place. The results so far are in line with the emerging scientific consensus that while recombination has taken place among coronaviruses, it has not happened recently or often, and is unlikely to be a major factor in how SARS-CoV-2 adapts to human hosts.

**Identifying possible COVID-19 treatments quickly**. At Oak Ridge and Argonne national laboratories, DOE's light sources and leadership computing capabilities have been leveraged to quickly respond to the coronavirus pandemic. One line of inquiry investigated the potential for drugs that typically treat ulcers and acid reflux for their potential antiviral properties. Through computer simulations of SARS-CoV-2 viral proteins and human proteins, the team rapidly ruled out these drugs as possible anti-viral treatments in COVID-19 patients without the need for testing them further. These efforts continue as part of a multi-lab NVBL effort to identify promising lead compounds for COVID-19 therapeutics.

Although large-scale computing is critical, more and more we understand that there is a rapidly growing need for mid-range, interactive computing, referred to by some as edge-computing, that is not currently best served by the large compute capabilities at DOE's high-performance computing centers. One reason is the rapid growth in omics data (which reflect genes working interconnectedly to express a function or trait), and another is the growing need for data processing and analysis capabilities on site – at the JGI and at NCEM for instance. The 4D Camera mentioned previously generates images at 100,000 frames per second and consequently requires high-speed networks with 400Gbps to 1 Tbps capacity like ESnet to access a supercomputer center like NERSC to process these images in near real-time. On-site, edge computing at NCEM could more efficiently prepare the data for delivery over ESnet. The same is true at JGI and at EMSL.

The Explosion in Data Volumes and Complexity

Another Challenge created by so much data is its distribution and dissemination. There has certainly been progress in providing biological data to researchers online and in ways that are transparent, user friendly, and useful. However, large gaps persist. Information that would be critical to scientists, although published and publicly available, may often exist in an unorganized, uncatalogued way – hard to access and manipulate. It may have well been hidden due to the difficulty of searching through thousands of seemingly unrelated scientific papers and disconnected databases. BER is tackling this challenge head on by leading the development of user-friendly databases as tools to store, catalogue, analyze, and make easily accessible critical biological information. A key feature of this effort is to make data FAIR - findable, accessible, interoperable, and reusable. Essentially FAIR allows many more discoveries to be made in addition to those for which the data were originally gathered.

One example is BER's Systems Biology Knowledgebase (KBase). KBase is an open software and data platform that enables researchers to predict and ultimately design biological functions. KBase's unified data model allows users to perform integrated analyses across plants, microbes, and their communities with a wide range of tools that interoperate across the tree of life, and to publish their data, methods, results, and thoughts in consistent, citable, executable, and reusable ways that allow scientists to build on the work of others. KBase's openness enables external developers to integrate their analysis tools, facilitating distribution and comparative tool analysis.

A second example is the newly established National Microbiome Data Collaborative (NMDC). As you may have heard, especially in relation to your gut and other human biology, microbiomes play important and often determinant roles in the environments in which they exist. This is also true for plants, their growth and resiliency, the health of other minute organisms, and for larger environmental ecosystems. But, most microbes and how they work together in microbial communities, within microbiomes, remain clouded in mystery.

The mystery is understandable. As mentioned previously in the discussion about JGI, a handful of soil may contain as many microbes as there are stars in our galaxy – perhaps a staggering two hundred billion. Just that statistic alone demonstrates the size of the data problem. But, add to that the need to decipher how the constituent microbes communicate, how they strategize, and how or why they act independently or in concert, and ultimately what is the impact of their actions on a plant, a field of biocrops, or an entire watershed, and one can begin to get a better sense of the challenge. We've made progress, however, and over the past few decades microbiome data have grown exponentially. But the sheer amount of data available and the challenge of analyzing and interpreting it productively still present a significant bottleneck.

The NMDC, a partnership led by Berkeley Lab along with Pacific Northwest, Los Alamos, and Oak Ridge national laboratories, is developing an open-access data framework to capture and share microbiome data and analyses. The goal is to facilitate more efficient use of microbiome data, regardless of the purpose for which they may be originally generated, for a wide variety of important applications in energy, environment, health, and agriculture. To tackle this data integration challenge, NMDC leverages DOE's existing data-science resources and the high-performance computing systems mentioned above. The NMDC's guiding principles are: making data FAIR; connecting data and compute resources; and engendering strong community engagement that supports open science and shared ownership.

A collaboration with NMDC and the BioScales project at Oak Ridge National Laboratory aims to understand how genes influence ecosystem-level processes in plant-microbiome systems. For example, understanding the genes that allocate carbon to plant root systems can ultimately enable engineering of plants to increase biological carbon capture and create more resilient feedstocks for biofuels and bioproducts production. A major objective is to engineer the more efficient uptake of nitrogen to reduce the need to use nitrogen fertilizers that cause environmental problems including toxic algal blooms.

Microbiome R&D: The Next Frontier?

BER's establishment of the NMDC is a reaction to a science community need, but it is also a response to the federal government's call for a more focused and collaborative all-of-government approach to microbiome research and technology development. Recognizing that siloes among the federal research agencies limit the sharing of data and hamper collaboration and joint efforts to tackle similar science and technology challenges in microbiome R&D, the White House established the Microbiome Interagency Working Group in February of 2016. And, in April of 2018, 21 federal agencies, including USDA, DOE, DOD, NASA, and NSF, but also including USAID, the Veterans Administration, and the FBI among others, released the *Interagency Strategic Plan for Microbiome Research FY 2018-2022*. The plan recommended three "areas of focus to transform microbiome discoveries to solutions":

1. *Supporting interdisciplinary and collaborative research to enable a predictive understanding of the function of microbiomes in diverse ecosystems to enhance public health, food, and environmental security and grow new bioeconomy product areas.*
2. *Developing platform technologies to generate critical insights and to improve access to and sharing of microbiome data across ecosystems.*
3. *Expanding the microbiome workforce through educational opportunities, citizen science, and public engagement.*

The Interagency Strategic Plan proposes coordination in microbiome research activities across 21 U.S. government agencies, "describing the interagency objectives, structure and operating principles, and research focus areas." The goal is to identify capabilities that may be leveraged across government and "represents an opportunity to increase Federal efficiency."

DOE BER, with strong support from the Congress, is doing its part to fulfill the goals of the Interagency Strategic Plan, but more is needed to ensure that the United States retains its current leadership in this critical area of research and technology development.

<u>Fabricated Ecosystems to Speed Biology-Based Solutions</u>

As evidenced through the successful programs and scientific output at JGI, the BRCs, EMSL, KBase, and now NMDC, BER and the national labs are making great strides in managing the large amounts of and ever-increasing complexity of data. However, significant challenges in biology remain, including leveraging experiments, sensing, data, computation, and other capabilities in real time, in situ, to reproducibly observe biology in action and to more quickly test hypotheses and learn from failures and success.

This need has led to the creation and utilization of fabricated, in situ, hyper-sensored environments -- environments created in the lab that can replicate conditions in the field.

Trent Northen, a Presidential Early Career Award in Science and Engineering awardee and a rising star in the world of microbial science at Berkeley Lab, has designed and proselytizes the EcoFAB initiative. EcoFAB aims to build model laboratory ecosystems at laboratories around the world to study environmental microbiomes and learn more about soil carbon cycling and how to develop low input sustainable crop production. To meet this goal, EcoFAB has developed a small, hand-sized lab-scale device, a fabricated ecosystem, to study soil-plant-microbe interactions in a reproducible and controllable system.

Fabricated ecosystems, such as the small EcoFABs, or much larger systems like EcoPODS currently the size of phone booths and in prototype production, hold promise for replicating important ecosystem dynamics. Observing microbiomes, plants, and other organisms under different conditions, scientists will be able to define principles for microbial community assembly and structure, understand the functions of genes, microbes, and metabolomes, and predict microbiome health. EcoFABs and EcoPODs have the potential to bring together communities of scientists working on shared systems, enabling more effective knowledge transfer and the ability to build upon previous findings. By expanding our understanding of the assembly, structure, and functions of microbiomes, significant scientific and technical advances will be made in studies of biomedical technologies, environmental health, agriculture, energy, and nutrient cycling.

Currently the United States leads the development of this new, larger-scale approach to biological research. With its focus on novel technology creation and open, accessible, and shared capabilities and learnings, these fabricated ecosystem research tools are exemplars of BER's and the national laboratories' secret sauce – the approach to science that began with Ernest and his brother John, and with Melvin Calvin and his team of collaborators. DOE is investing significant resources in this novel approach. The Office of Sciences' Science Laboratories Infrastructure program is funding the construction of a new research building at Berkeley Lab to house these fabricated ecosystems in an integrated way. The Biological & Environmental Program Integration Center (BioEPIC) is currently in the advanced planning stage.

With BER's support and through the support of this and other Congressional committees, the U.S. can retain its leadership and reap the benefits of developing these new resources, and of being first in the delivery of new science, new technologies, new processes, and new products.

**<u>Fueling the United States Bioeconomy</u>**

Today, the biological mission drivers for DOE, the Office of Science, and BER of creating a more diversified energy future and ensuring environmental sustainability are as important as ever.

They also offer even more potential for driving economic growth and advancing the nation's international innovation competitiveness. Because of the BER's decades of critical investments in biology R&D at the national laboratories and at universities, DOE is playing an outsized role in the development of the nation's bioeconomy. The "bioeconomy," which encompasses a broad range of bio-based, bio-derived, and bio-inspired products and production processes, is a potential juggernaut of economic growth and has appropriately been recognized by the federal government, and governments around the world, as a key industry of the future.

Biology already plays an integral role in the U.S. economy. The foundation of every modern biomanufacturing effort is an important gene or suite of genes that were decoded and purposefully re-coded using genomics technologies such as those at the JGI as previously mentioned. According to a January 2020 National Academies of Sciences, Engineering, and Medicine report, the US bioeconomy is a set of economic activities valued at $959B and is growing at a significant rate. Fueled by advances in life sciences, engineering, biotechnology, and computing and information sciences, the US bioeconomy is unique in the world. DOE, through BER and also through applied research programs within its Office of Energy Efficiency and Renewable Energy (EERE), is making incredible contributions to bio driven economic development.

EERE's Bioenergy Technologies Office (BETO), which funds research and development for industrially relevant transformative bioenergy technologies, is enabling economically and environmentally sustainable, domestically produced, biofuels, bioproducts, and biopower. BETO accomplishes this through a variety of mechanisms, including support for the national labs to address pre-competitive challenges where the solutions will be of broad benefit to industry. Some examples of applied biobased research at DOE include:

**The Agile BioFoundry (ABF).** The ABF, a collaboration project between seven national labs, is driving down the cost and time to develop new bioprocessing technologies through an integrated Design-Build-Test-Learn cycle using machine learning. Based on the design of electronic fabs, ABF provides state of the art biology tools and capabilities as a platform for industry -- from startups to large well-established companies.

**The Advanced Biofuels and Bioproducts Process Development Unit (ABPDU).** The ABPDU at Berkeley Lab was established in a previous time of crisis to de-risk and scale up early stage biofuels and bioproducts technologies for industry. Funded by the American Recovery and Reinvestment Act, the ABPDU staff have worked with over 60 partners to transition their early-stage technologies to commercialization ready, resulting in at least 12 products in the market today and over $450M in follow-on investment for those partners.

**Addressing the Coronavirus pandemic at the ABPDU.** True to their mission to accelerate products to market, the ABPDU team assisted three new company partners in the fight against the SARS-CoV-2 virus and COVID-19 with development of new testing reagents, a possible treatment, and nutritional support for patients. The ABPDU stands ready as a resource for the biomanufacturing community and is poised to quickly adapt to any challenge.

**First biotechnology-based jet fuel.** BETO's investment in bioenergy research and PNNL's long-standing expertise in catalysis science has resulted in an ethanol-based jet fuel that has sustained transcontinental commercial flight. PNNL scientists and their industrial partner LanzaTech jointly developed the first jet fuel from industrial waste gas. This partnership will receive a 2020 Industrial Research Institute Interchange (IRI) Award for its breakthrough. The innovative conversion process fueled a Virgin Atlantic commercial passenger flight from Orlando to London in late 2018 and, in late 2019, an All Nippon Airways flight which flew from Seattle to Tokyo. Using a combination of biotechnology coupled with catalysis, the team proved carbon can be recycled and used for commercial flight.

Biology has the potential, however, to play an even larger role in enhancing U.S. international competitiveness if the nation undertakes a more sophisticated and integrated approach to strategic planning and collaboration that includes increasing targeted investments and developing clear goals and objectives. The international competition is moving fast and our nation's leadership is not assured.

Appointed by the U.S. Department of State, I serve as the US delegate to the Organisation of Economic Cooperation and Development's Biotechnology, Nanotechnology, and Converging Technologies Working Party. Serving in this role has given me a front row seat in the development of the bioeconomies of other nations. A number of countries have recognized the importance of undertaking a focused and concerted approach and have developed strategic plans and detailed roadmaps with clear goals, objectives and milestones. Once the lead country in embracing biotechnology as a driver of a bioeconomy strategy, the US now has considerable competition. Several other countries are evaluating the power of biotechnology as they consider updating their bioeconomy strategies from biomass-focused to include biotechnology in an effort to reboot economies ruined by the coronavirus pandemic.

An example of a plan with considerable investment behind it is China's 12th Five Year Plan, which calls for hundreds of billions of dollars in funding for research and development in biopharmaceutical, bioengineering, bio agriculture, and biomanufacturing R&D. The plan aims to strike the right balance between the seed corn of basic science and the technology development needed for commercial application.

The United Kingdom's (UK) roadmap, first issued in 2012, looks at the opportunities and challenges of biotechnologies from basic science to real world applications, regulatory considerations, and health, safety and environmental issues. Although funding allocated to these efforts in the UK is significantly lower than in China, their access to top talent, a focused approach and clear deliverables are helping to build a strong foundation of scientific leadership and entrepreneurial progress. The UK undertook a major update to their bioeconomy strategy in 2018 and doubled down on using biotechnology to solve societal problems, including addressing the plastics problem. Japan also renewed its plan in 2019 to feature biotechnology.

Again, demonstrating DOE's leadership and recognizing the importance of the nation's biosciences and biomanufacturing enterprises, Secretary of Energy Brouillette, Under Secretary for Science Paul Dabbar, and Office of Science Director Chris Fall tasked the national laboratories with organizing a BioManufacturing XLab, a summit, to highlight the national laboratories' biosciences and biomanufacturing capabilities for leading companies, large and small, from across the country. Organized and hosted by Berkeley Lab for DOE at the Cal football stadium, the Bio XLab attracted hundreds of attendees from companies across the biomanufacturing industry and featured talks from Emily LeProust, the CEO and co-founder of Twist Biosciences, a company that synthesizes DNA, Magalie Guilhabert, VP, Head of Microbial Research Technology at Bayer CropScience where they are developing new methods to improve our food crops; and Alta Charo, Professor of Law and Bioethics at University of Wisconsin-Madison, one of the foremost bioethicists working on issues of biotechnology. Speakers like Paul Dabbar, and Daniel Simmons, the Assistant Secretary for Energy Efficiency and Renewable Energy, emphasized DOE's commitment to discovery and development for new biomanufacturing technologies. Our national lab technology transfer experts devised a new innovation for this XLab event- bundles of available intellectual property across the national laboratories, organized by topic, for a "one-stop shop" for potential partners and licensees. This key innovation solves a crucial problem for industry. IP developed at the national labs is challenging to find, let alone coordinate. The XLab is still generating positive collaborations over six months after the event.

Importantly, it was at this January biomanufacturing XLab event that Director Chris Fall announced the concept of the NVBL. To prepare for future biological incidents, intentional or natural such as the current coronavirus pandemic, the DOE national laboratories have the bio-related core capabilities and are now more prepared to respond collaboratively in the event of a future biological crisis as a consequence of the establishment of the NVBL.

## **Conclusion**

The need to develop novel energy and environmentally sustainable solutions, and to build the nation's bioeconomy, grows daily.  Stressors, including the need to reduce and store carbon at humongous scales, develop crops for a changing climate and more resilient bio crops, and secure leadership of the global bioeconomy demand quick action and quick delivery to society of transformative technologies and processes.  The nation or nations that are first to do so, and to do so responsibly, will reap the societal and economic benefits of being first. The United States must retain its lead.

Increased opportunities and greater expectations bring more complex challenges.  And again, as in years past, BER's and the DOE national laboratories' ability to attract a diverse workforce of the best and brightest and to imagine, build, and maintain new sophisticated research tools, facilities, and expertise is keeping the Department and the nation at the vanguard of biosciences discovery.

Thank you, again, for inviting me to testify at this important hearing.  And, thank you for your strong support for science and technology development, and for a diverse and inclusive STEM workforce.