U.S. Department of Energy
Office of Biological and
Environmental Research

Biological Systems Science Division

STRATEGIC PLAN

U.S. DEPARTMENT OF ENERGY | Office of Science
Office of Biological and Environmental Research

October 2015

U.S. Department of Energy

Office of Science

Office of Biological and Environmental Research

Biological Systems Science Division

Unless noted, email addresses for BER staff follow the format: firstname.lastname@science.doe.gov

Todd Anderson, Director, 301.903.3213

Genomic Science Program

Catherine Ronning, 301.903.9549

Pablo Rabinowicz, 301.903.0379

Kent Peters, 301.903.5549

Dawn Adin, 301.903.0570

Ramana Madupu, 301.903.1398

Roland Hirsch, 301.903.9009

Bioenergy Research

Bioenergy Research Centers
Kent Peters, 301.903.5549

Systems Biology Research for Bioenergy
Dawn Adin, 301.903.0570

Plant Feedstocks Genomics
Catherine Ronning, 301.903.9549

Sustainable Bioenergy Research
Catherine Ronning, 301.903.9549

Biosystem Design

Pablo Rabinowicz, 301.903.0379

Carbon Cycling and Biogeochemical Research

Dawn Adin, 301.903.0570

Computational Biology

Ramana Madupu, 301.903.1398

Bioimaging Technology Development

Prem Srivastava, 301.903.4071

Amy Swain, 301.903.1828

Facilities and Infrastructure

Joint Genome Institute

Daniel Drell, 301.903.4742

Structural Biology Infrastructure

Amy Swain, 301.903.18280

Human Research Subjects Protection

Elizabeth White, 301.903.7693

U.S. Department of Energy Office of Biological and Environmental Research

Biological Systems Science Division

STRATEGIC PLAN

October 2015

science.energy.gov/ber/research/bssd/

U.S. DEPARTMENT OF ENERGY | Office of Science

Contents

From Genomes to Predictive Biology

Encoded in the genomes of plants, microbes, and their communities are biological principles that offer a wealth of potential for advancing U.S. Department of Energy (DOE) priorities in developing renewable energy sources, understanding possible effects of Earth system change, and establishing leadership in biotechnology development.[1–2] To harness this potential, the Office of Biological and Environmental Research's (BER) Biological Systems Science Division (BSSD) within the DOE Office of Science supports basic research to understand the fundamental nature of biological processes relevant to DOE energy and environmental challenges.

A central feature of BSSD research is using genome sequences as a foundational basis for attaining a mechanistic, predictive understanding of diverse plant and microbial systems across a range of scales, from molecules to small ecosystems (see Fig. 1. Multiscale Exploration of Biological Function, this page). This emphasis on genomics traces back to the instrumental role BER and collaborating entities played in pioneering the genome sequencing technology that culminated in

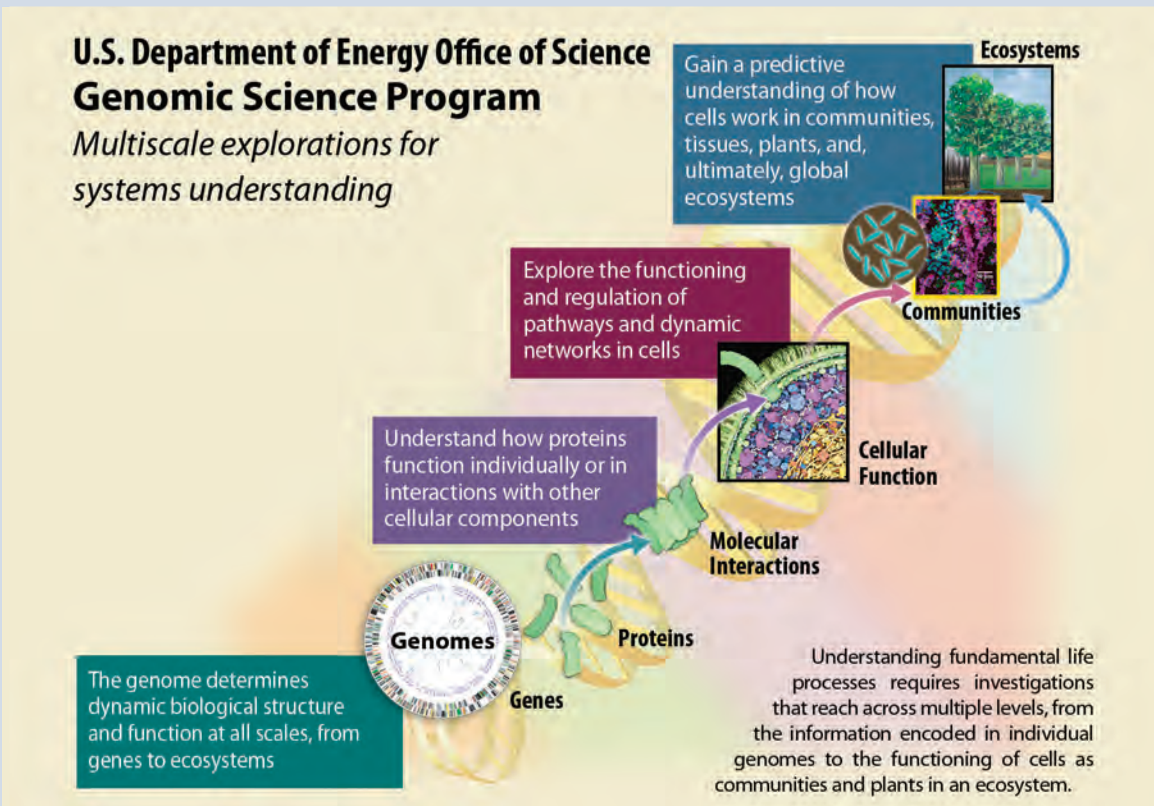

Fig. 1. Multiscale Exploration of Biological Function.

[1] *U.S. Department of Energy Strategic Plan 2014–2018.* www.energy.gov/sites/prod/files/2014/04/f14/2014_dept_energy_strategic_plan.pdf

[2] *DOE Genomic Science Program 2014 Strategic Plan.* genomicscience.energy.gov/strategicplan/

the production of the human genome sequence in 2003. Since then, genome sequencing capacity has progressed dramatically in a relatively short time. In fact, the vast volume of genome sequences and associated "omics" data from an ever-increasing range of organisms is presenting new challenges for meaningfully interpreting this information at a pace comparable to data production. Interpreting and leveraging this wealth of information will require new and efficient ways to accelerate knowledge generation and dissemination within the research community. This document outlines a high-level overview of BSSD's current science efforts and a strategy for incorporating new capabilities to propagate this knowledge for solutions to DOE missions.

Research Portfolio

BSSD supports a unique combination of interdisciplinary, team-oriented research groups; advanced analysis capabilities at DOE scientific user facilities; and access to high-performance computing. A major element of the BSSD portfolio is the Genomic Science program, which supports the bulk of the division's bioenergy and environmental research. To advance this research, the program relies on the insights and talents of a multidisciplinary research community from academia, industry, and the DOE national laboratories.

In addition, BSSD supports genome sequencing and analysis technologies at the DOE Joint Genome Institute (JGI), as well as structural biology capabilities at DOE synchrotron light (and neutron) sources. The division also has access to high-performance computational systems via the National Energy Research Scientific Computing Center (NERSC) and the DOE Systems Biology Knowledgebase (KBase). These capabilities recently have been supplemented with a new Bioimaging Technology development program (science.energy.gov/ber/bioimaging-technology/).

BSSD's combination of scientific expertise coupled with advanced enabling capabilities and high-performance computational systems constitutes a unique and world-class research program in plant and microbial systems biology (see sidebar, Applying the Power of Systems Biology, this page).

Basic Research Goals

The overall goal of BSSD is to "provide the necessary fundamental science to understand, predict, manipulate, and design biological processes that underpin innovations for bioenergy and bioproduct production and to enhance the understanding of natural environmental processes relevant to DOE" (see sidebar, BSSD Overarching Goal and Objectives, p. 3). At its most fundamental level, the division collectively seeks to answer a set of basic questions important to understanding biology:

Applying the Power of Systems Biology

The ongoing "omics" revolution continues to generate new and increasingly sophisticated experimental systems biology approaches and analytical techniques that can be brought to bear on Department of Energy mission goals. A useful description of systems biology is provided in the National Academy of Sciences report, *A New Biology for the 21st Century*:

"Systems biology seeks a deep quantitative understanding of complex biological processes through dynamic interaction of components that may include multiple molecular, cellular, organismal, population, community, and ecosystem functions. It builds on foundational large-scale cataloguing efforts (e.g., genomics, proteomics, metabolomics, etc.) that specify the 'parts list' needed for constructing models. The models relate the properties of parts to the dynamic operation of the systems they participate in."

BSSD Overarching Goal and Objectives

***Overarching Goal:** Provide the necessary fundamental science to understand, predict, manipulate, and design biological processes that underpin innovations for bioenergy and bioproduct production and to enhance the understanding of natural environmental processes relevant to DOE.*

Objectives

- Provide a basic understanding of plant and microbial biology to underpin the production of biofuels and bioproducts from sustainable plant biomass resources.
- Develop the fundamental understanding of genome biology needed to design, modify, and optimize plants, microbes, and biomes for beneficial purposes.
- Gain a predictive understanding of biological processes controlling the flux of materials (e.g., carbon, nutrients, and contaminants) in the environment and how these processes impact ecosystem function.
- Develop the enabling computational, visualization, and characterization capabilities to integrate genomic data with functional information on biological processes.
- Broaden the integrative capabilities within and among DOE user facilities to foster a more interdisciplinary approach to BER-relevant science and aid interpretation of plant, microbe, and microbial community biology.

1. What information is encoded in a genome sequence, and how is this information translated to function?
2. What interactions (e.g., among cells in a microbial community or within plant tissues) regulate the function of living systems, and how can these interactions be understood dynamically and predictively?
3. How do living systems interact with their environment and respond to changes (e.g., in temperature, water and nutrient availability, and ecological interactions among microbial communities and plants in microbiomes or small ecosystems)?
4. What organizing principles need to be understood to combine or design biological processes for beneficial purposes?

Answering these basic questions across multiple species in varied environments (both natural and in the laboratory) and at scales ranging from molecules to ecosystems provides opportunities to learn from biology and use or design biological processes for numerous societal benefits.

Interdisciplinary and Integrative Research Approaches

Today, the scope of BSSD basic science programs extends to the myriad microbial and plant species that harbor metabolic capabilities that could be harnessed for DOE missions in bioenergy and environment. The challenge is to understand how genomic information is translated to function in these biological systems and how these systems are coupled together over a range of cellular, multicellular, and small ecosystem scales. The breadth and complexity of this research require interdisciplinary approaches and new computational strategies for integrating and effectively working with the large-scale and diverse data emanating from each program in the BSSD portfolio:

Genomic Science Program

- Bioenergy Research
  - Bioenergy Research Centers
  - Systems Biology Research for Bioenergy
  - Plant Feedstocks Genomics
  - Sustainable Bioenergy Research
- Biosystems Design
- Carbon Cycling and Biogeochemical Research
- Computational Biology

Bioimaging Technology Development

Facilities and Infrastructure

- DOE Joint Genome Institute
- Structural Biology Infrastructure

Maximizing Insights Through Interdisciplinary Efforts

Interdisciplinary collaborations are a hallmark of BSSD research, and the division will continue to promote such efforts across a range of modalities—from small projects (including those led by single investigators) to large comprehensive science teams within academia and the DOE national laboratories. Additionally, BSSD will strive to cultivate an integrative, collaborative atmosphere across its entire portfolio through focused workshops and principal investigator meetings that promote a free exchange of scientific ideas among researchers and program staff from different disciplines. These efforts will maximize the potential for new and translational insights.

Integrating and Leveraging Data and Experiments Through Computing

A key counterpart to BSSD interdisciplinary research is the development of new approaches to efficiently combine high-performance computational techniques with experimental analyses to enable iterative hypothesis-based research on whole biological systems (see. Fig. 2. Integrating Experiments, Computing, and Technologies for Predictive Understanding, p. 5). Rapid advances in omics technologies are providing increasingly detailed insights into genome organization and expression, and new experimental capabilities are yielding equally detailed information on overall biological function. The ever-increasing complexity and diversity of these datasets require developing high-performance computational approaches to integrate this information across molecular, multicellular, and small ecosystem scales and to simplify the ability to work with and interpret the data.

As analyses become more complex, capabilities also are needed for sharing, reproducing, and propagating scientific results among researchers. Common computational platforms openly accessible to the research community (e.g., KBase) could supply the frameworks needed to facilitate these interactions, resulting in new knowledge generation.

Current Research Priorities

BSSD basic science challenges in bioenergy, biosystems design, and environment are guided by national needs to develop renewable sources of energy that can replace products currently derived from fossil fuels and to better predict the impacts of Earth and environmental system change. Outlined below are long-term goals for each area of emphasis within the BSSD research portfolio, including bioenergy, biosystems design, environmental research, enabling capabilities, and user facility technologies and integration. The goals for each area align with and address overall division objectives.

Also described for each area are important existing or future linkages with other entities both within and external to DOE that pursue complementary research activities. In addition, BSSD will maintain communications with the National Academies, scientific societies, and nonprofit organizations to coordinate on science themes of mutual interest. The division also will continue to coordinate research activities under the umbrella of the National Science and Technology Council.

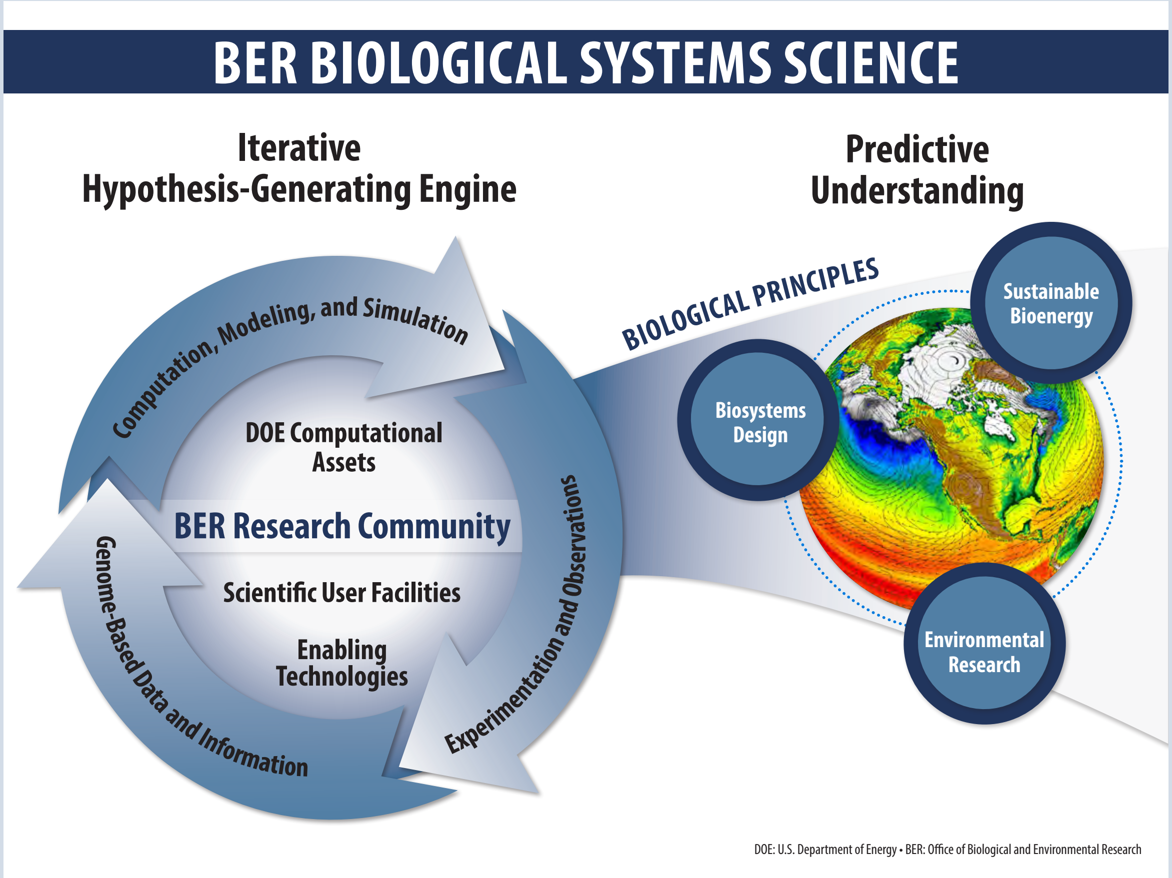

Fig. 2. Integrating Experiments, Computing, and Technologies for Predictive Understanding.

Bioenergy

***Goal:** Provide a basic understanding of plant and microbial biology to underpin the production of biofuels and bioproducts from sustainable plant biomass resources.*

Bioenergy and Bioproducts Research

Today's liquid fuels and many of the chemical products that form the basis of modern society are derived from petroleum and other fossil fuel resources, which are precious and ultimately limited in quantity. Producing fuels and chemicals from renewable resources lessens the dependence on petroleum for modern-day energy sources, paving the way toward a more sustainable long-term energy future.

BSSD basic research efforts in bioenergy and bioproducts seek to provide the fundamental science needed to underpin the development and use of renewable and sustainable biomass as a source of fuels and products.[2–5] The scope of research extends from gaining a basic understanding of plant and microbial biology necessary for developing and

[3] *Lignocellulosic Biomass for Advanced Biofuels and Bioproducts.* genomicscience.energy.gov/biofuels/lignocellulose/BioenergyReport-February-20-2015LR.pdf

[4] *Basic Research Opportunities in Genomic Science to Advance the Production of Biofuels and Bioproducts from Plant Biomass.* genomicscience.energy.gov/biofuels/BER-Bioenergy-WhitePaper-Final.pdf

[5] *Industrialization of Biology: A Roadmap to Accelerate the Advanced Manufacturing of Chemicals.* www.nap.edu/catalog/19001/industrialization-of-biology-a-roadmap-to-accelerate-the-advanced-manufacturing/

converting dedicated bioenergy crops to fuels and products to obtaining fundamental insights into the sustainability of bioenergy crop production.

Much of the division's recent efforts in bioenergy research have been performed within the DOE Bioenergy Research Centers (genomicscience.energy.gov/centers/). These large interdisciplinary centers have provided key insights into the fundamental biology needed to spur broader development of biofuels from lignocellulosic plant biomass. Continued bioenergy research within BSSD will seek to expand on this body of knowledge to extend the range of biofuels and bioproducts that could be produced from biomass feedstocks. The challenge is to develop the underlying fundamental science needed to produce a wide variety of fuels and products from renewable and sustainable biomass resources as replacements for those currently generated from petroleum.

Subgoal: Provide the basic science to enable a sustainable and commercially viable lignocellulosic biomass-derived advanced biofuels and bioproducts industry.

This goal will be achieved by:

- Elucidation of the regulation of gene networks, proteins, and metabolites to improve plant feedstock productivity and sustainability.
- Comparative approaches to enhance knowledge of the structure, function, and organization of plant genomes, leading to innovative strategies for feedstock characterization, breeding, manipulation, and improvement.
- Characterization of plant germplasm collections and advanced breeding lines of bioenergy crops to discover and deploy valuable alleles for key bioenergy traits.
- Development of new cultivars of regionally adapted bioenergy feedstock crops in public breeding programs using innovative approaches to identify desirable traits and accelerate trait integration.
- Research into the complex interactions between bioenergy feedstock plants and their environment and the influence of these processes on plant growth and development, expression of bioenergy-relevant traits, and adaptation to changing environments.
- New and fundamental knowledge of biomass deconstruction processes to convert a broader range of biomass types into a range of precursor substrates readily convertible into biofuels and bioproducts.
- Increased rates of biomass degradation and expanded utilization of biomass components (e.g., hemicellulose and lignin monomers).
- Decreased susceptibility to product inhibition by toxins released during biomass breakdown or high titers of synthesized biofuels, thus increasing the efficiency and yield of biofuel production.
- Increased tolerance to physicochemical stresses associated with industrial-scale biofuels production (e.g., elevated temperature, altered pH levels, or fuel toxicity).
- Modified functional properties that result in increased yield of biofuel compounds or synthesis of a broader range of molecules that can be used as next-generation biofuels and related bioproducts.

Fundamental bioenergy research efforts within BSSD will be coordinated with DOE programs and governmental other agencies engaged in bioenergy research, including the DOE Advanced Research Projects Agency–Energy (ARPA-E), DOE Office of Energy Efficiency and Renewable Energy (EERE), U.S. Department of Agriculture (USDA), National Science Foundation (NSF), and U.S. Department of Defense (DOD).

Sustainability Research

BSSD's basic science investments in plant and microbial research are advancing the production of biofuels and bioproducts from domestic lignocellulosic sources. As these efforts progress toward development of new dedicated bioenergy crops, novel strategies will be needed to ensure sustainable large-scale crop production.[2, 6] BSSD's systems biology approaches to bioenergy production will build on extensive knowledge of plant and microbial systems and combine these efforts to examine plant-microbe interactions to gain insights into sustainable bioenergy crop cultivation. Key to this effort will be gaining a predictive understanding of plant-microbe interactions leading to improved growth and yield of dedicated bioenergy crops across a range of marginal soil types and geographic regions. An important aspect of this research will be to enable development of robust feedstocks and associated microbial communities that require few external inputs (i.e., nutrients), are resistant to environmental change, and have minimal ecosystem impacts.

Subgoal: Develop new approaches to bioenergy agriculture that cost-effectively provide high yields of biomass on marginal lands requiring few or no inputs with plants highly adaptable to changing environmental conditions and having minimal to no impacts on the ecosystem.

This goal will be accomplished by:

- Understanding the molecular and physiological mechanisms that control crop vigor, resource use efficiency, and resilience to increase biomass productivity under changing and occasionally suboptimal conditions.
- Characterizing the extent to which microbes contribute to plant performance, stress tolerance, and adaptation to changing environmental conditions and maximizing positive plant-microbe interactions in cropping systems by developing cultivars that take advantage of these beneficial interactions.
- Using genomic knowledge of soil microbial communities to predict the response of key biogeochemical processes (e.g., carbon stabilization, denitrification, nitrous oxide fluxes, methane oxidation, and leaching losses) to episodic environmental events such as freeze-thaw cycles, prolonged drought, and rainfall.
- Manipulating specific feedstock plant traits to enhance soil carbon sequestration via changes to the quantity, quality, and location of belowground carbon inputs and promoting nutrient and soil conservation through changes in rooting patterns and architecture.
- Developing process-based, multiscale models that take into account key variables such as genomic properties of biofuel crops, rhizosphere community functions, soil carbon stabilization, nutrient cycling, and greenhouse gas production and that accurately predict plant performance and ecosystem processes under changing environments.

BSSD efforts in sustainability research will be closely coordinated with USDA programs (particularly for access to field sites) and DOE ARPA-E and EERE projects. Additionally, as insights are gained on key variables affecting plant growth and yields across small ecosystems, potential opportunities may emerge to incorporate predictive understanding of this information into the larger-scale ecosystem research and modeling efforts conducted by BER's Climate and Environmental Sciences Division (CESD). Such efforts could contribute to regional-scale impact assessments of Earth and environmental system change on bioenergy plant growth and yield, information needed by planners at a more local scale.

Biosystems Design

***Goal:** Develop the fundamental understanding of genome biology needed to design, modify, and optimize plants, microbes, and biomes for beneficial purposes.*

The ability to manipulate the genomes of plants and microbes is expanding rapidly. Biosystems design offers the potential to take advantage of biochemical mechanisms developed in natural systems and apply them in beneficial ways for a broad range of

[6] *Research for Sustainable Bioenergy: Linking Genomic and Ecosystem Sciences.* genomicscience.energy.gov/sustainability/sustainability13report.shtml

uses in bioenergy and a host of industrial applications. In fact, gaining an understanding of the design principles involved in manipulating microbial and plant genomes is key to developing innovations for a biobased economy.[2,5,7]

Building on BSSD's long history of genomic science, the division will gain insights into the underlying principles governing genome organization and regulation for the purpose of designing new beneficial functions and capabilities into organisms. Biotechnology as a research endeavor is advancing quickly, and although new techniques are enabling researchers to more easily manipulate plant and microbial genomes for desired uses, significant barriers for expanding this capability still exist. BSSD will seek to provide the necessary fundamental understanding and capabilities to overcome these barriers and spur broader innovation in biosystems design for a host of DOE-relevant purposes. These efforts will lead to new innovations in biotechnology with implications for a broader biobased economy through development of:

- Better understanding of fundamental biological design principles to determine the foundational laws governing living systems and better define the "solution space" available to biosystems designers.
- Genome-scale engineering tools enabling improved integration of large synthetic DNA constructs, high-throughput gene editing and recoding technologies, and incorporation of alternate amino acids and other synthetic design assemblies.
- New DNA synthesis and assembly technologies for cost-effective, high-efficiency synthesis and assembly of large DNA constructs.
- Diverse platform microbes and plants for which there is sufficient baseline understanding of their systems biology properties and available molecular genetics tools on which to base genome-scale engineering approaches, providing a broader array of functional capabilities and growth characteristics.
- Minimal cell and *in vitro* systems using highly streamlined platform organisms or entirely cell-free *in vitro* systems with fully defined pathways and functional components that facilitate easier modification and increased process control.
- Advanced computationally aided design tools for biological systems (e.g., BioCAD) to facilitate *in silico* design of synthetic parts and systems, accommodate a wider suite of biological parts and platform organisms, and enable efficient workflows for construction and experimental validation of large numbers of genetic design variants.
- Biocontainment mechanisms permitting incorporation of multiple redundant safeguards at each step of the design process, thereby preventing inadvertent release or deliberate misuse of synthetic biological systems.

BSSD efforts in biosystems design will be coordinated with agencies conducting complementary basic science efforts in this area, including NSF and DOD's Defense Advanced Research Projects Agency program.

Environmental Research

Goal:* *Gain a predictive understanding of biological processes controlling the flux of materials (carbon, nutrients, and contaminants) in the environment and how these processes impact ecosystem function.

BSSD's Genomic Science program has been at the forefront of research to gain a broad understanding of microorganisms found in the environment, their functional attributes, and their impact on local and global biogeochemical cycles.[2] Advances in genome sequencing technology have greatly expanded these efforts not only to uncover new metabolic capabilities within individual microbes but also to begin to mechanistically understand

[7] *Biosystems Design: Report from the July 2011 Workshop.* genomicscience.energy.gov/biosystemsdesign/biosystemsdesign.shtml

how microbial communities function.[2] BSSD's environmental research efforts combine this genome-enabled information with experimental approaches on function to gain a predictive understanding of biological processes in the environment. These studies will continue to explore a range of environments to discover new metabolic processes of potential use in BSSD's bioenergy research, but also to gain a predictive understanding of how microbial communities impact the cycling of carbon, nutrients, and DOE-relevant contaminants. These efforts are tied to obtaining a better understanding of how microbial communities impact the flux of materials in the environment, respond to Earth system changes, and affect plant-microbe interactions influencing the sustainable growth and yield of bioenergy plants. This research builds on BSSD's genomic science legacy and seeks to integrate mechanistic understanding of cellular metabolism with external environmental and microbial community interactions across a range of observational scales, from processes occurring within single cells to microbiome or small ecosystem scales. Major objectives for environmental research include:

- Systems biology studies on microbes, microbial consortia, and microbe-plant interactions involved in large-scale terrestrial carbon cycling processes, with particular emphasis on giving organisms that have been determined to play key roles the status of experimentally tractable model organisms.
- Determination of the role of microbial communities as key points of integration between major biogeochemical cycles (e.g., carbon, nitrogen, sulfur, and phosphorus) and the impact of these interconnections on the rate and magnitude of carbon cycle processes in changing ecosystems.
- Further development of environmental omics approaches (e.g., metagenomics, metatranscriptomics, metaproteomics, and community-scale metabolomics) and associated data integration tools to investigate *in situ* microbial community functional processes involved in carbon cycling.
- Development of omics-enabled techniques for imaging and analysis of microbial community structure and function in terrestrial environments, with particular emphasis on tools permitting high-resolution and quantitative measurements in technically challenging settings such as soils, sediments, and key interfacial environments (e.g., decaying organic material, mineral aggregates, and plant roots).
- The ability to translate predictive understanding of biological processes occurring within individual organisms to complex communities operating at ecosystem scales. As BSSD environmental research matures, the division will seek opportunities to couple these efforts with CESD environmental programs to achieve this objective.

Enabling Capabilities for Systems Biology Research

Goal:* *Develop the enabling computational, visualization, and characterization capabilities to integrate genomic information with functional information on biological processes.

BSSD seeks to accelerate the pace of genomic science by developing more efficient ways to integrate a broad range of omics information to test hypotheses and validate simulations of key biological processes relevant to DOE missions. The division will support the development of complementary enabling capabilities in areas including high-performance computing (HPC), imaging, and characterization.

Specifically needed are new technologies that can evaluate whole biological systems. As genomic information becomes more widely available, novel high-throughput experimental techniques will be required for testing multiple hypotheses on these systems and for identifying and manipulating gene function. Such capabilities hold the potential not only to speed understanding of individual metabolic processes, but also to place those processes in the context of larger biological systems (e.g., cells, plant tissues, or microbial communities).

Similarly, new capabilities to visualize biological processes within and among cells are needed to validate assumptions on the spatiotemporal relationships of such processes in and among cells or plant tissues. New imaging, measurement, and characterization technologies will provide a vital check on simulations and predictions of cellular processes and will enable researchers to image gene expression in the dynamic environment of living cells. These capabilities will be crucial for gaining a predictive understanding of complex biological systems.

New technologies often spur discovery but also can simplify and accelerate the interpretation, analysis, and dissemination of data within the larger research community. New open-access cyberinfrastructure platforms that rely on HPC are needed so that researchers can readily access, analyze, update, and share results of genomic research with the broader scientific community.

Leveraging High-Performance Computational Platforms

BSSD's approach to systems biology research requires the integration of a vast range of omics data from numerous disparate sources. The sheer volume of information available is itself a significant bottleneck to conducting hypothesis-based research on complex biological systems. HPC offers a way to help simplify working effectively with this large-scale data.

Open-access cyberinfrastructure platforms operating in the cloud on DOE systems provide researchers with access to many of the world's most powerful computers. These machines offer the ability to rapidly automate and analyze extremely large and complex datasets efficiently and in ways that could greatly streamline systems biology investigations. Such platforms also can enable researchers to share large-scale analyses reproducibly with collaborators and extrapolate omics results across an ever-increasing phylogenetic spectrum.

The integration of open-access cyberinfrastructure platforms on HPC systems into BSSD's genomic science endeavor is crucial for accelerating knowledge generation and dissemination within the research community. All these needs constitute the motivating ideas behind KBase, a major BER investment to provide a common cyberinfrastructure for genomic science. KBase is an ambitious step toward integrating HPC into genomic research, an effort that BSSD will pursue within its portfolio. Specific BSSD objectives include support for developing:

- Open-access cloud-based platforms to enable researchers to conduct complex, diverse, and integrated genomic-based analyses on high-performance computers.
- User-friendly interfaces to simplify the analysis of complex, large-scale omics data and to assemble the results in an experimentally tractable format.
- Capabilities to reproducibly share large-scale omics analyses with other researchers.
- New computationally efficient approaches for analyzing large omics datasets.
- Methods for assembling omics data into whole-cell simulations to allow iterative *in silico* behavior prediction and hypothesis development under varied environmental conditions.
- Methods for assembling models of multicellular organisms and communities to gain a predictive understanding of plant and microbial community function.
- Progress toward integration of biological and ecological or environmental models.

BSSD will coordinate with the DOE Office of Advanced Scientific Computing Research (ASCR) to address these goals and the unique computational challenges posed by large-scale genomics research and biological process simulation. Closer ties with ASCR programs are needed to understand and explore the potential for incorporating HPC and new computational approaches into biological research, not only for discovery science but also to simplify the research community's ability to work with extremely large and diverse datasets. BSSD also will coordinate data integration efforts with CESD to permit the incorporation of genome-based predictive understanding of microbial and plant systems into larger-scale models of ecosystem function.

Molecular-Scale Science and Bioimaging Technologies

While genomics provides a window into the genetic potential and expression of genes in a microorganism or plant cell, understanding the spatial and temporal functional dynamics of cellular processes occurring *in vivo* is also essential. Visualizing how processes are regulated, expressed, and organized within the dynamic context of a living cell is crucial to gaining a systems perspective on how cells function, particularly in response to environmental stimuli, signaling from other cells, or engineered genomic changes.

To understand how genetic information is translated to function, BSSD is seeking development of new multifunctional, multiscale imaging and measurement technologies to visualize the spatiotemporal expression and function of biomolecules, intracellular structures, and the flux of materials across cellular compartments. These technologies will provide information on the expression, modification, function, and degradation of biomolecules within living cells. This information is crucial for validating hypotheses about cellular function, improving simulations of the cellular environment, and designing new functions into plants, microorganisms, or biomes. BSSD will seek new technologies to:

- Combine structural information about biomolecules with computation and bioinformatics to infer function, improve genome annotation, and/or design new functions.
- Visualize the spatial and temporal dynamics of expressed biomolecules within or between living plant or microbial cells and their communities.
- Use visualization techniques to validate predictions and simulations of cellular dynamics.
- Develop multifunctional bioimaging technology capabilities and scientific infrastructure for measuring, analyzing, and modeling whole-cell and multicellular biological systems.
- Develop *in situ*, dynamic, and nondestructive approaches to multifunctional imaging, quantitative flux measurements, and multiscale integrative analysis of biological systems.

These efforts combine molecular-scale science capabilities within BSSD's structural biology component, the division's new effort in Bioimaging Technology development, and technologies available at the DOE Environmental Molecular Sciences Laboratory (EMSL) and other DOE user facilities. Combining structural information on biomolecules with visualization techniques and computer simulation offers the potential to gain a deeper understanding of cellular metabolism or engineered metabolic processes occurring *in vivo*. BSSD will seek engagement of the scientific community to help chart a path forward for this important portfolio element.

Integration of User Facility Capabilities

Goal: Broaden the integrative capabilities within and among DOE user facilities to foster a more interdisciplinary approach to BER-relevant science and aid interpretation of plant, microbe, and microbial community biology.

DOE operates a unique portfolio of national scientific user facilities with a broad array of analytical capabilities freely available to the scientific community. For BSSD programs, these facilities could provide added value to researchers if the ability to combine access to one or more capabilities could be offered to users simultaneously. Such collaborations could promote the types of interdisciplinary research needed to address complex BSSD basic science objectives.

Within BER, collaborations among user facilities are already ongoing. For example, DOE's JGI and EMSL issue joint solicitations to scientists seeking access to multiple capabilities. These collaborative activities enable users to combine unique capabilities from both user facilities to address interdisciplinary research questions. Specifically, the JGI-EMSL collaboration enables users to couple advanced genome sequencing and analysis capabilities at JGI with the advanced proteomics and imaging capabilities at EMSL. Other combinations are certainly possible between these two centers, and similar collaborative approaches could be taken with other user facilities. BSSD will explore this potential among the BER user facilities and seek to develop comparable collaborations with other DOE user facilities where applicable.

Genome Sequence Production and Interpretation (DOE Joint Genome Institute)

The DOE JGI is the premier DNA sequencing facility devoted to understanding DOE-relevant plants, microbes, and microbial communities. BSSD's efforts in genomic science depend on JGI and other sequencing centers to produce the genome sequences that serve as the foundational data on which to base a systems biology research program. As a user facility, JGI has been an essential resource for the genomic science community and has continued to adapt new technology to its user services portfolio. JGI is transitioning to focus more on the interpretation of genomes rather than just their production. As sequencing costs continue to decrease (and capacity continues to increase), JGI is engaging the research community to incorporate new innovative and high-throughput technologies to provide additional interpretative value to the genomic sequences it produces. As BSSD programs shift toward interpreting and designing new genomic functions, JGI is providing an array of new capabilities, including computational techniques to analyze and compare genome sequences, technologies for single-cell analysis, DNA synthesis services, and techniques to aid metabolic engineering of organisms.

JGI remains a prominent and crucial feature of the BSSD portfolio and will continue to support BSSD bioenergy and environmental science goals through:

- Continued leadership in genome sequencing technologies adapted to address the highly complex genomes of plants and microbial communities (metagenomics).
- Development of large-scale DNA synthesis and genomic manipulation capabilities to enable users to understand genome organization, regulation, manipulation, and design.
- New analysis techniques to compare genomic data for improved structural and functional genomic annotations.
- Development of new high-throughput phenotyping technologies to support genomic functional annotation with experimental data.
- Incorporation of HPC into genome analysis capabilities.
- Collaborations and joint user access activities with other DOE user facilities that have complementary capabilities for interpreting gene and genome function.

BSSD will continue to work with JGI to develop a range of capabilities for genomic science, either in house or in combination with other DOE user facilities. The division will work collaboratively with JGI and EMSL to further promote joint user access to capabilities at each facility. Similar joint solicitations could be explored to leverage, for example, the structural biology capabilities at DOE's synchrotron light and neutron sources, as well as other DOE user facilities.

BSSD will seek to facilitate collaborative activities among DOE user facilities not only to increase the visibility of the centers among a broader range of BER scientists, but also to promote the interdisciplinary use of these facilities in BSSD research.

Summary

These are exciting times for biology. Rapid advances in basic biological science and biotechnology are beginning to provide solutions to major health, energy, environmental, and economic challenges. BSSD science has been and continues to be at the forefront of providing leading-edge biological science for DOE and is well positioned to continue to make major scientific contributions in the future. With its unique combination of scientific and technological talent supported within academia and DOE national laboratories, advanced world-class scientific capabilities at DOE user facilities, and access to the most powerful computational systems in the world, BSSD research has an auspicious future. This document describes current areas of emphasis within the BSSD portfolio but also outlines the division's plan to develop more integrative and efficient use of the capabilities available within its programs and the larger DOE complex. Working with the research community to leverage these assets will help provide the basic science underpinning solutions to DOE energy and environmental challenges.

Acronyms

ARPA-E	DOE Advanced Research Projects Agency–Energy
ASCR	DOE Office of Advanced Scientific Computing Research
BER	DOE Office of Biological and Environmental Research
BSSD	BER Biological Systems Science Division
CESD	BER Climate and Environmental Sciences Division
DOD	U.S. Department of Defense
DOE	U.S. Department of Energy
EERE	DOE Office of Energy Efficiency and Renewable Energy
EMSL	DOE Environmental Molecular Sciences Laboratory
HPC	high-performance computing
JGI	DOE Joint Genome Institute
KBase	DOE Systems Biology Knowledgebase
NERSC	National Energy Research Scientific Computing Center
NSF	National Science Foundation
USDA	U.S. Department of Agriculture