Genomics-Enabled Plant Biology for Determination of Gene Function

Summary of projects awarded in 2022 under Funding Opportunity Announcement DE-FOA-0002601

Funded Projects

- Elucidating the Genetic Components of the Physiological and Metabolic Processes Governed by the TORC Regulatory Module in Poplar

- Using an Evolutionary Perspective to Discover and Predict Stress-Associated Gene Functions

- Harnessing Regulatory Variation to Elucidate Drought-Resilience Mechanisms in Sorghum

- Validation of an Acyltransferase Toolbox for Plant Biomass Engineering

- A Systems Framework to Enhance the Potential of *Camelina* as an Oilseed Crop

- *Populus* and *Sorghum* Gene Function in Biomass Development

- Discovering Transcriptional Regulators of Photosynthesis in Energy Sorghum to Improve Productivity

- Functional Analysis of Genes Encoding Ubiquitin-Proteasome System Components Affecting Complex Traits Influencing Biomass Resilience and Productivity in Poplar

- WINTR: Winter Transcriptome Regulation in Poplar

- Functional Characterization of Glycosyltransferase in Duckweed to Enable Predictive Biology

- Phenotypic and Molecular Characterization of Nitrogen-Responsive Genes in Sorghum

Plants have been developed to be productive on lands that are unsuitable for traditional agricultural crops, and these plants have the potential to play a significant role as an alternative source of renewable feedstocks for fuels and chemicals. Plant scientists and breeders recognize that potentially bigger gains in productivity and sustainability may be realized in such crops by leveraging multiomics approaches, systems biology, and computational biology to better understand the regulation of critical plant processes at the molecular level. However, understanding molecular function remains a challenge in plants due to several unique features, such as segmental and whole genome duplications, differences in mode of reproduction, and domestication. To overcome these obstacles, the U.S. Department of Energy's Biological and Environmental Research (BER) program selected 11 projects in fiscal year 2022 aimed at (1) understanding regulatory elements that control plant gene expression, (2) developing approaches to understand the biological mechanisms that lead to differential gene expression of different members of gene families, and (3) dissecting multigene traits and understanding their component parts. The awards are sponsored by BER's Genomic Science program (GSP).

Experimental characterization of gene function continues to be a significant bottleneck, hampering scientists' ability to fully exploit the potential of available omics data. This limits direct understanding of how genes regulate organismal function and ultimately the ability to accurately assign and validate gene

Sorghum growing in a hot arid environment under the field scanner gantry at the University of Arizona's Maricopa Agricultural Center. [Courtesy University of Arizona]

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

function in nonmodel organisms. By elucidating complex regulatory processes at the molecular, cellular, and organismal levels, a better understanding of plants as a whole and more accurate predictions of plant behavior under varying conditions can be achieved.

With this research, GSP's goal is to address the challenges and opportunities in associating gene(s) to function (i.e., genotype to phenotype) in plant systems relevant to BER missions in energy and the environment. More efficient, high-throughput methods for interpreting experimental evidence need to be developed for accurate determination of gene function. While experimental approaches are unlikely to scale with advances in omics technologies, concentrated development and application of new approaches are required to tackle this critical knowledge gap of genomics space.

The new plant biology projects encompass the following topical areas:

- Comprehensive investigations of regulatory elements and mechanisms (e.g., transcription factors, cis-regulatory regions, enhancers, insulators, epigenetic regulation, and alternative splicing mechanisms) to validate regulation of gene expression.

- Approaches to functionally understand gene families to elucidate regulatory differences among members, subfunctionalization, and functional divergence, particularly in large gene families.

- Dissection of multigene traits (e.g., yield) into their genetic components and characterization and validation of molecular genetic functions for these components, with an aim toward understanding their impacts on the final phenotype.

2022 Awards

Elucidating the Genetic Components of the Physiological and Metabolic Processes Governed by the TORC Regulatory Module in Poplar

- **Principal Investigator:** Gary Coleman (University of Maryland)

- **Co-Investigators:** Yiping Qi, Ed Eisenstein (University of Maryland); Victor Busov, Hairong Wei (Michigan Technological University)

Understanding the genetic basis of complex traits related to resource use efficiency and abiotic stress is necessary to improve lignocellulose crops for a sustainable biobased economy. Nutrient sensing and signaling is a fundamental mechanism that modulates cellular activities mediating growth, development, and biomass accrual. This project will investigate the molecular processes involved in sensing and signaling of the poplar target of rapamycin (TOR) protein complex (TORC) to gain a better understanding of the nutrient response. Using CRISPR gene editing, genomics, biochemical, and computational approaches, the role(s) and possible functional differences between the two poplar TOR genes will be deciphered, specifically in relation to carbon and nitrogen signaling. The results will elucidate the regulatory and signaling

networks and hubs involved in the nitrogen response in poplar and uncover how TOR and Rho of plants (ROP) guanosine triphosphate binding proteins (GTPases) modulate other processes such as environmental stress and biomass yield.

Using an Evolutionary Perspective to Discover and Predict Stress-Associated Gene Functions

- **Principal Investigator:** José Dinneny (Stanford University)

- **Co-Investigators:** Maheshi Dassanayake (Louisiana State University); Song Li (Virginia Tech University)

Innovations in gene function allow wild plants to inhabit environments that are commonly stressful to domesticated crops. The goal of this project is to identify such innovations by defining their regulatory and physiological context across 11 Brassicaceae genomes, including bioenergy crops and crop wild relatives, using genomic approaches and artificial intelligence algorithms. Signaling pathways that control the expression of genes during environmental stress will be characterized and used to predict how genetic variation alters these stress tolerance programs. Computational approaches that utilize gene duplication events and evolutionary theory will identify

genetic states associated with the gain or loss of stress tolerance within the Brassicaceae family. Predicted gene functions will be validated using advanced gene editing approaches. Together, these approaches will define gene functions through a comparative approach and identify pathways that can be leveraged for the improvement of stress tolerance in bioenergy crops.

Harnessing Regulatory Variation to Elucidate Drought-Resilience Mechanisms in Sorghum

- **Principal Investigator:** Andrea Eveland (Donald Danforth Plant Science Center)
- **Co-Investigators:** Todd Mockler (Donald Danforth Plant Science Center); Duke Pauli (University of Arizona); Brian Dilkes (Purdue University)

This project aims to define and functionally characterize genes related to drought-stress tolerance in sorghum as well as variations on gene regulation that drive phenotypic plasticity. Leveraging a field-based phenotyping infrastructure in Maricopa, Ariz., the team will (1) develop an integrative multi-omics framework for drought response in sorghum and its genetic diversity to predict functions of genes and their regulatory components, (2) dissect the morphophysiological processes by which sorghum responds to drought at the whole-plant level and link these back to genetic loci, and (3) experimentally validate predictions of gene function using molecular and genetic assays and targeted gene editing. Assessment of genotype to phenotype in response to drought over the life cycle of the plant will enable the development of predictive growth models and the discovery of genes and their functions as well as facilitate
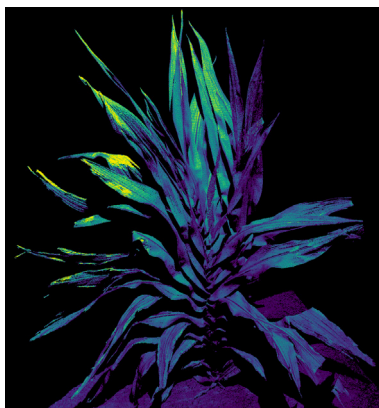
3D representation of a single sorghum plant collected with a laser scanner. Single plant point clouds are extracted for ~7,000 plants over the season to evaluate plant architecture during plant growth and development. [Courtesy University of Arizona]

development of improved, climate-resilient varieties for production on marginal soils with limited inputs.

Validation of an Acyltransferase Toolbox for Plant Biomass Engineering

- **Principal Investigator:** Brian Fox (University of Wisconsin)
- **Co-Investigators:** John Ralph (University of Wisconsin); Shawn Mansfield (University of British Columbia)

The goal of this project is to expand the understanding and utility of the acyl-CoA ligases and BAHD acyltransferases, whose coordinated action creates diverse valuable bioproducts, by systematically evaluating BER's priority bioenergy plants. An annotated and curated list of ~450 genes from these two families will be created, followed by functional validation using *in vitro* biochemistry and screening of 10 to 30 selected genes encoding best-in-class enzymes. These genes will be subjected to a thorough genomic, compositional, and viability analysis along with the content, composition, and yield of modified cell wall polymers such as zip-lignin, clip-off aromatics, and additional metabolites. The result will be an experimentally validated toolbox of BAHD acyltransferase and acyl-CoA ligase genes detailing their *in vitro* reactivities and abilities to modify the properties of bioenergy crop plants, providing an experimentally validated knock-out/knock-in list for performing precision engineering of plant metabolism to obtain new and improved properties.

A Systems Framework to Enhance the Potential of *Camelina* as an Oilseed Crop

- **Principal Investigator:** Erich Grotewold (Michigan State University)
- **Co-Investigator:** Patrick Edger (Michigan State University)

Camelina sativa has emerged as one of the leading commercially viable options for biofuel and bioproduct production in the United States. Spring and winter varieties show superior abiotic stress tolerance, traits that could be utilized in breeding programs to greatly enhance its performance and productivity. However, a major limitation of widespread adoption of *Camelina* as an industrial oilseed crop is its modest yield, which is further hindered by a poor understanding of gene regulatory networks, genetic diversity, and gene content across *Camelina* accessions. This project will address these obstacles by (1) characterizing the genetic variation, gene expression,

and chromatin accessibility across varieties and growth conditions, and (2) developing the tools to understand and manipulate *Camelina* gene expression. These studies will identify key genes and genomic regions to target in breeding efforts for more resilient and higher yielding cultivars while providing tools to understand and manipulate *Camelina* gene expression.

Populus and *Sorghum* Gene Function in Biomass Development

- **Principal Investigator:** Matias Kirst (University of Florida)
- **Co-Investigators:** Kelly Balmont, Wilfred Vermerris (University of Florida); Sushmita Roy (University of Wisconsin)

Plant biomass is made of several cell types, each of which has different physical and chemical characteristics and functions. Uncovering the function of the genes involved in the development of cell types that compose plant biomass creates the opportunity to manipulate biomass composition and make it more suitable to produce energy. By applying single-cell transcriptomics to the analysis of *Populus* and *Sorghum* stem tissue, this project will characterize the genes involved in determining the cellular composition of plant biomass. Genes that control evolution of one cell type to another will be identified, and their function will be validated to confirm their role in biomass development. The results will provide new routes to modify biomass properties in *Populus* and *Sorghum,* and consequently, the value of their biomass for bioenergy production.

Discovering Transcriptional Regulators of Photosynthesis in Energy Sorghum to Improve Productivity

- **Principal Investigator:** Laurie Leonelli (University of Illinois)
- **Co-Investigator:** Stephen Long (University of Illinois)
- **Collaborator:** Matthew Brooks (U.S. Department of Agriculture)

The aim of this project is to enhance understanding of regulatory factors that control photosynthesis throughout the sorghum canopy and to identify molecular targets for optimizing photosynthetic efficiency to boost overall sorghum yields. An innovative high-throughput approach to validate transcription factor (TF)-target gene interactions *in planta* will be utilized to investigate how those interactions influence gene expression at different positions and light environments of the sorghum canopy. Tracking gene expression in response to the light environment will allow determination of which genes are differentially expressed and consequently implicated in the loss of photosynthetic efficiency observed in self-shaded sorghum. Gene expression will be integrated with validated TF-target gene data through gene regulatory network modeling to reveal key TFs involved in regulation of photosynthesis. These TFs will be the targets for future efforts to enhance overall plant yield and improve sorghum's potential as a bioenergy crop.

Functional Analysis of Genes Encoding Ubiquitin-Proteasome System Components Affecting Complex Traits Influencing Biomass Resilience and Productivity in Poplar

- **Principal Investigator:** Nitzan Shabek (University of California, Davis)
- **Co-Investigators:** Andrew Groover (U.S. Forest Service); Justin Walley (Iowa State University)

In this project, key components of ubiquitin-proteasome regulation of protein levels will be functionally characterized in poplar. Data indicate that ubiquitin-mediated processes affect key aspects of wood formation, including regulation of the diameter of water-conducting cells, which impacts growth rates and drought tolerance. A proximity labeling strategy will be used to determine gene function by identifying proteins that interact with a given gene product. Additional systems biology approaches will be employed to integrate gene expression, protein-protein interactome, and phenotypic data. Altogether, this project will provide predictive models of physiological and developmental processes critical for enhanced biomass yield and properties for bioenergy applications.

WINTR: Winter Transcriptome Regulation in Poplar

- **Principal Investigator:** Chung-Jui Tsai (University of Georgia)
- **Co-Investigators:** Wellington Muchero, Priya Ranjan (Oak Ridge National Laboratory)

Woody biomass growth of trees comprises a significant contribution to the supply of renewable feedstocks for biofuels and biomaterials. While most studies of growth have been performed during the summer or indoors, field-grown trees are physiologically active year-round, with physiological and metabolic adjustments conferring

winter protection in temperate deciduous species. The aim of this project is to investigate how molecular control of the winter latent state is linked to perennial woody biomass productivity. A multipronged approach integrating RNA-Seq, genome-wide association studies, expression quantitative trait loci mapping, high-precision CRISPR genome editing, and gene network modeling will be used to investigate transcriptome regulation in woody stem tissues of *Populus* during the winter. Of particular interest are genome duplicates that exhibit either winter-biased expression or divergent seasonal expression in both species. Understanding how seasonal growth dynamics impact woody biomass productivity will offer new targets for bioenergy crop improvement.

Functional Characterization of Glycosyltransferase in Duckweed to Enable Predictive Biology

- **Principal Investigator:** Breeanna Urbanowicz (University of Georgia)

- **Co-Investigators:** Kelley Moremen, Maria Peña, Pradeep Prabhakar (University of Georgia); Yannick Bomble (National Renewable Energy Laboratory)

Glycosyltransferases (GTs) catalyze the formation of glycosidic linkages to produce complex carbohydrates. This project will use a multidisciplinary, high-throughput biochemical and computational biology approach to study carbohydrate metabolic processes in duckweed, a promising bioenergy crop. The role of enzymatic microenvironments will be assessed through a combined proteomic and computational biology approach. This combined

data will be used to populate a deep-learning framework to predict plant GT function. Functional validation achieved through this research will be used to assign gene function and study plant processes at the systems level to efficiently link genome sequence with gene function.

Phenotypic and Molecular Characterization of Nitrogen-Responsive Genes in Sorghum

- **Principal Investigator:** Jinliang Yang (University of Nebraska)

- **Co-Investigators:** Thomas Clemente, Yufeng Ge, James Schnable (University of Nebraska); Sanzhen Liu (Kansas State University); Kankshita Swaminathan, Jane Grimwood, Avinash Sreedasyam (HudsonAlpha Institute for Biotechnology)

The inefficient use of nitrogen fertilizer increases the agricultural burden on the environment and ecosystems. Adding to this problem, inorganic nitrogen fertilizer is one of the most expensive and energy-intensive agricultural inputs in crop production, especially for sorghum. Improving sorghum's nitrogen use efficiency as an energy crop will enhance farmer profitability and alleviate environmental impacts. Using an integrated approach that includes gene editing, phenotyping, genomics, and population-wide transcriptomic analyses, newly edited and newly discovered nitrogen-responsive genes will be characterized under high-nitrogen and low-nitrogen conditions. Data will be analyzed through population genetic and comparative genomic approaches. Identification and functional validation of candidate genes will facilitate improving nitrogen use efficiency in sorghum.

Contact and Websites