Pacific Northwest
NATIONAL LABORATORY

PNNL-30351

The Consortium for Advanced Sorghum Phenomics (CASP)

September 2020

G.C. Jansson

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <https://www.ntis.gov/about>
Online ordering: http://www.ntis.gov

PNNL-30351

The Consortium for Advanced Sorghum Phenomics (CASP)

September 2020

G.C. Jansson

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Final Scientific/Technical Report

Agency to which Report is Submitted:	Advanced Research Projects Agency - Energy
FOA Name:	TRANSPORTATION ENERGY RESOURCES FROM RENEWABLE AGRICULTURE (TERRA)
FOA Number:	**DE-FOA-0001211**
Report:	Research Performance Progress Report
Award:	
Prime Recipient:	Pacific Northwest National Laboratory (PNNL)
Project Title:	**THE CONSORTIUM FOR ADVANCED SORGHUM PHENOMICS (CASP)**
Principal Investigator Name, Title, and Contact Info:	Christer G. Jansson Director of Plant Sciences EMSL, PNNL christer.jansson@pnnl.gov 509-371-6516
Prime Recipient DUNS Number:	**78576738**
Date of Report:	June 25, 2020
Reporting Period:	October 1, 2015 – December 14, 2018

Public Executive Summary

The goal of CASP was to accelerate breeding of biomass sorghum [*Sorghum bicolor* (L.) Moench] by identifying genotypes exhibiting high yield under well-watered, pre- or post-drought and/or salinity-stress conditions. We did this by combining high-throughput, non-invasive drone phenotyping with genomics and molecular profiling. Field-based phenotyping utilized a multi-modal sensor suite of LiDAR, multispectral cameras, and thermal cameras mounted on a commercial drone to detect traits required for yield prediction and selection of drought and saline tolerant lines of sorghum. Traits of interest included plant height (PH), leaf area index (LAI), wet biomass (BMW), and biomass at 65% moisture (BM65) and were measured from emergence to harvest on a weekly basis over three growing seasons. The final output were measurements of traits on a plot-by-plot basis, identified by the plot ID used by the Proprietary data processing software enabled raw field data to be turned into plant traits and delivered to the PNNL and JGI within the same workday.

Field sites

The CASP field sites were located at the UC Kearny Agricultural Research & Extension (KARE) Center and Westside Research and Extension Center (WREC), CA. The lack of summer rainfall in this area offers an excellent opportunity for well-controlled field-based drought experiments.

BRT drone flying over one of the sorghum plots at the CASP field sites in CA.

Sorghum Accessions from the Sorghum Conversion Program (SCP) were selected based on race and working group information, pedigree, seed origins, and pre-evaluations under stress responses and represents most of the genetic diversity found within sorghum. Additional accessions were selected from Chromatin's biomass breeding program based on chemical composition screening. All SCP accessions will be acquired from the Germplasm Resources Information Network's (GRIN) seed bank. From this diverse set of preselected sorghum germplasm, we chose 648 sorghum genotypes that were planted in control, pre- and post-flowering drought nurseries on two different field and soil types in KARE and WREC for a total of 3,888 research plots. A salinity nursery was added to the evaluations at WREC. Seeds of the 648 varieties were planted in a randomized complete block design in ten rows, 30 ft long with 30" row spacing, comparable to normal planting regimes for grain sorghum. The field trial was repeated for three seasons (2015 – 2018).

Image-based phenotyping by UAV

Drone data for PH, LAI, BMW, and BM65 were collected on every location once per week during the growing seasons. For the drought study, drone phenotyping yielded a total of more than 500,000 phenotypic observations for associations with genome-wide association studies (GWAS) and metabolite and protein phenotyping. Algorithms for analyzing data enable prediction of plant growth potential, even at the early stages of growth. The high accuracy and high precision of the drone data was demonstrated by ground phenotyping in a random-block design; R^2 by cross validation with ground-truthing = 0.89 - 0.98

Genotyping

Genotyping by sequencing (GBS) and GWAS based on drone data from the 648 genotypes yielded 213 high quality, reliable GWAS peaks for single-nucleotide polymorphisms (SNPs) associated with biomass and/or drought tolerance traits. The large number of associations are due to the accuracy of the phenotyping over the entire growing seasons and provide a broader knowledge of the genetics underlying drought tolerance and biomass accumulation than any previous study. This validates the overall approach of the CASP project and provides numerous targets for breeding and more detailed studies. Significantly, the majority of the loci were significant on multiple dates and some were only significant early or late in the season.

Metabolomics

A subset of 120 genotypes based on drone and GWAS data were selected for molecular phenotyping. A total of 3,800 samples (8 biological replicates for each genotype sampled from well-watered, pre-flowering drought, or post-flowering drought conditions) were collected and analyzed by GC-MS/LC-MS, a next-generation SPE-IMS-MS platform, and FTICR-MS, which yielded a data set of ~20,000 metabolite features, the lion's share of which are unknowns.

The data and leaf samples generated under CASP represent an unprecedented resource for sorghum research and breeding. The molecular profiling, in combination with GWAS, has the potential to provide mechanistic understanding for genes and pathways responsible for high biomass accumulation in sorghum on marginal land. This information can be utilized for identifying sorghum cultivars with desirable phenotypes solely based on genomic, protein and/or metabolite data.

Acknowledgements

We acknowledge ARPA-E for financially supporting the research, and for all team mebers at PNNL, JGI, BRT, KARE and WREC who participated in the work.

Table of Contents

Accomplishments and Objectives

Table 1. Key Milestones and Deliverables.

Task	Key Milestones and Deliverables
1. T2M	Q1: Build value model. **Actual Performance:** (1/2016) Milestone was met. Q1: Form advisory board. **Actual Performance:** (1/2016) Milestone was met. Q2: Revise T2M plan and identify point of contact for T2M. **Actual Performance:** (4/2016) Milestone was met. **Actual Performance:** 4/2016) Milestone was met. Q3: Update Value model based on industry input **Actual Performance:** (7/2016) Milestone was met. Q4: T2M report update **Actual Performance:** (10/2016) Milestone was met. Q4: xpand value model. **Actual Performance:** (10/2016) Milestone was met. Q4: Create industry market value chain report on promising markets. **Actual Performance:** (10/2016) Milestone was met. Q5: Build Model for value of terminal biomass prediction. **Actual Performance:** (1/2017) Milestone was met. Q5: Submit industry engagement plan. **Actual Performance:** (1/2017) Milestone was met. Q6: Finalize service agreement with an industry partner (e.g. Syngenta) for proof-of-concept (PoC) project. **Actual Performance:** (4/2017) Milestone was met. Q6: Create initial T2M plan for identifying commercially relevant salt and/or drought tolerant inbred lines and public lines. Develop an initial plan for testing a path to market for these technologies. **Actual Performance:** (4/2017) Milestone was met. Q6: Create initial T2M plan and path forward for commercialization of identified molecular biomarkers for drought tolerance and/or biomass accumulation in sorghum. **Actual Performance:** (4/2017) In progress. Q7: Deliver updated T2M Plan with refined path to market based on discussions with industry advisors and technical learnings. **Actual Performance:** (7/2017) Milestone was met. Q7: Update cost models in M1.7 and M1.8. **Actual Performance:** (7/2017) In progress. Q7: Progress against engagement plan approved by ARPA-E, reported every other quarter. **Actual Performance:** (7/2017) Milestone was met. Q8: Update T2M Plan. **Actual Performance:** 10/2017) Milestone was met. Q8: Report on Proof-of-Concept project with industry partner. **Actual Performance:** 10/2017) Milestone was met. Q8: Update plan for commercialization of sorghum germplasm with improved drought resistance. **Actual Performance:** 10/2017) Milestone was met. Q8: Update plan for commercialization of molecular biomarkers for drought tolerance

	and/or biomass accumulation in sorghum. **Actual Performance:** 10/2017) Milestone was met. Q8: Build value model for Biomarker Assays. **Actual Performance:** 10/2017) Milestone was met. Q9: Demonstrate continuing external market interest for 2018. **Actual Performance:** (1/2018) Milestone was met. Q9: Progress against engagement plan reported. Q10: Deliver analyses estimating the value of hybrid development progress and update options for commercial implementation. The value of future salt and drought tolerant hybrid pipeline will be assessed based on potential new markets for salt tolerant bioenergy, forage, and grain hybrids worldwide. **Actual Performance:** (4/2018) In progress. Q11: Deliver updated T2M Plan with refined path to market based on discussions with industry advisors and technical learnings. **Actual Performance: (**10/2018) No longer applicable. The original tech-to-market plan targeting plant breeders has changed since the acquisition of Blue River by John Deere. Q11: Identify specific field trials for PoC project with an industry partner (e.g. Syngenta) **Actual Performance: (**10/2018) No longer applicable. Q11: Deliver analyses estimating the value of hybrid development progress and update options for commercial implementation. The value of future salt and drought tolerant hybrid pipeline will be assessed based on potential new markets for salt tolerant bioenergy, forage, and grain hybrids worldwide. **Actual Performance:** (10/2018) Chromatin is currently evaluating their field experiments and correlating with drone data. Q11: Deliver schedule and next steps post-project for commercializing salt and drought tolerant products. **Actual Performance:** (10/2018) No longer applicable. Q11: Create plan and path forward for commercialization of identified molecular biomarkers for high-salinity tolerance and/or biomass accumulation in sorghum (Biomarker-S). **Actual Performance:** (10/2018) The salinity field trial at WREC was sampled and metabolite analyses are ongoing. Q11: Recommendation for development of assay protocols and metrics for Biomarker-D and/or Biomarker-S applications in sorghum breeding. **Actual Performance:** (10/2018) Metabolite features correlating with biomass accumulation and draught stress have been obtained. Q12: port on PoC project with industry partner. **Actual Performance:** (2/2019) No longer applicable. Q12: Commercialization of sorghum germplasm with improved drought and salt tolerance. **Actual Performance:** (2/2019) KARE and WREC are working with statisticians to develop a stress index ranking based on regression analyses of high-throughput phenotypic data collected on weekly basis from drone fly-byes. Q12: Deliver schedule and next steps post-project for commercializing salt and drought tolerant products. **Actual Performance:** (2/2019) No longer applicable. Q12: Commercialization of deliverables from molecular phenotyping. **Actual Performance:** (2/2019) No longer applicable. Q12: Build value model for Biomarker Applications. **Actual Performance:** (2/2019) No longer applicable.

2. Sensor Platform	Q1: Build UAS system capable of collecting, storing, & transmitting field phenotypic data. **Actual Performance:** (1/2016) Milestone was met. Q1: Provide ARPA-e with updated sensor calibration protocol. **Actual Performance:** (1/2016) Milestone was met. Q2: Identification of target phenotypes and sensors. **Actual Performance:** (4/2016) Milestone was met. Q3: None Q4: None. Q5: Demonstrate data capture rate of 1x/day, 2x/week in the field with 70% reliability. **Actual Performance:** (1/2017) Milestone was met. Q5: Autonomously fly using GPS waypoints, covering at least one field per day (>2 acre, ~900 plots) and two fields per week. **Actual Performance:** (1/2017) Milestone was met. Q5: LiDAR and multispectral imaging achieves target spatial resolutions for year 1 (see Robotics and Sensing Workbook). **Actual Performance:** (1/2017) Milestone was met. Q5: For 10 consecutive measurements on five different targets, achieve multispectral reflectance repeatability and LiDAR repeatability at target metric (see Robotics and Sensing Workbook). **Actual Performance:** (1/2017) Milestone was met. Q5: Georeferenced data is sufficiently accurate to extract plot-level phenotypes (RMSE < 0.3 m horizontal positional accuracy). **Actual Performance:** (1/2017) Milestone was met. Q5: Provide ARPA-E with updated sensor calibration protocol. **Actual Performance:** (1/2017) Milestone was met. Q6: Integrate thermal camera into UAS system (mechanical, electrical, firmware integration) and characterize water stress accuracy using ground-truthing. **Actual Performance:** (4/2017) Milestone was met. Q6: ARPA-E approval of updated Robotics and Breeding workbook containing target phenotypes and associated sensors. **Actual Performance:** (4/2017) Milestone was met. Q7: None. Q8: Integrate thermal camera in to UAS system and characterize water stress accuracy using ground truthing. **Actual Performance:** (10/2017) Milestone was met. Q8: Autonomously fly throughout the season, demonstrating coverage at least once per week of target field (>2 acre). **Actual Performance:** (10/2017) Milestone was met. Q8: Collect UAS data at two corn breeding locations in Midwest with industry partner. **Actual Performance:** (10/2017) Milestone was met. Q8: ARPA-E approval of the sensor suite and target phenotypes for the next generation design of the UAS system. **Actual Performance:** (10/2017) No changes to sensor suite. Q8: Provide ARPA-E with updated sensor calibration protocol. **Actual Performance:** (10/2017) No changes to sensor calibration protocols. Q9: Demonstrate data capture rate of 2x/day, 2x/week at a vehicle speed of at least 1.5m/s with 85% reliability. Collect data at a second sorghum site at Chromatin near Lubbock, TX with 85% reliability, in addition to KARE and WREC. **Actual Performance:** (1/2018) Milestone was met. Q9: Thermal, LiDAR, and multispectral imaging achieves target spatial resolutions for year 2.

	Actual Performance: (1/2018) Milestone was met. Q9: Achieve canopy temperature measurement repeatability for 10 consecutive measurements. **Actual Performance:** (1/2018) Milestone was met. Q10: None. Q11: Collect data at a second sorghum site at Chromatin near Lubbock, TX, with 85% reliability (in addition to KARE and WREC sites near Fresno, CA). Reliability is defined as 8.5 out of 10 collection days result in all sensors producing usable data. **Actual Performance:** (10/2018) Milestone was met. Q11: Performer Input: No longer applicable, please see Milestone 1.9. Collect UAS data at two corn (maize) breeding locations in the Midwest, in partnership with Syngenta or other industry partner. **Actual Performance:** (10/2018) No longer applicable. Q12: Collect UAS data at two corn (maize) breeding locations in the Midwest, in partnership with Syngenta or other industry partner. **Actual Performance:** (2/2019) No longer applicable.
3. Data Analytics	Q1: None. Q2: None. Q3: None. Q4: Demonstrate data transfer from UAS to field office to server. **Actual Performance:** (1/2017) Milestone was met. Q5: Demonstrate image data is extracted from the correct plot (requires both airborne and ground-based plot map to be accurate) with 95% reliability. **Actual Performance:** (1/2017) Milestone was met. Q5: Yearly rank sensors and data streams, cost benefit analysis of each sensor/phenotype (e.g. pareto optimization front). **Actual Performance:** (1/2017) Milestone was met. Q5: Validate that high-throughput height and leaf area measurement accuracies are equivalent to those of current best practices in field-based (not high-throughput) phenotyping [$R2 > 0.95$ and RMSE < 10% for plot-average height and LAI]. **Actual Performance:** (1/2017) Milestone was met. Q5: Achieve correlation of terminal harvest biomass equivalent to current research methods ($R2 > 0.95$, Gitelson et al 2003 GRL). **Actual Performance:** (1/2017) Milestone was met. Q5: Prediction of heights and leaf area distributions of 3D canopy structure at target metric (see Robotics and Sensing workbook). **Actual Performance:** (1/2017) In progress. Q5: Predict terminal biomass yield of individual lines from early season field data (first half of growing season); Validate algorithms using field data that match current state of the art growth models. **Actual Performance:** (1/2017) In progress. Q5: Process and store genomics data. All bioinformatics/genomics code targeted for efficient parallelization and cloud based options. **Actual Performance:** (1/2017) Milestone was met. Q6: Update the genomics strategy and goals for the project in the broader TERRA effort. **Actual Performance:** (4/2017) In progress. Q7: Determine population structure, and relevant GWAS/GS algorithms for selected sorghum. **Actual Performance:** (7/2017) Milestone was met. Q8: Validate stand count and early plot quality in field corn.

Actual Performance: (10/2017) Milestone was met.
Q8: Discover >100 new alleles and genetic markers linked to biomass levels or other key phenotypes.
Actual Performance: (10/2017) Milestone was met
Q9: Process 1 week's data within 48 hours within target accuracy metrics.
Actual Performance: (1/2018) Milestone was met.
Q9: Yearly rank sensors and data streams, cost benefit analysis of each sensor/phenotype.
Actual Performance: (1/2018) Milestone was met.
Q9: Drought tolerance metrics: validate canopy temperature and leaf water potential have correlations $R^2>0.80$.
Actual Performance: (1/2018) In progress.
Q9: Achieve >= 10% improvement in accuracy (RMSE) of terminal harvest biomass compared to year 1.
Actual Performance: (1/2018) In progress.
Q9: Demonstrate >= 10% improvement in predictive accuracy compared to year 1.
Actual Performance: (1/2018) In progress.
Q9: Create GS model to predict performance from genotype and select lines for future crosses. Identify at least one marker useful for breeding decisions for each phenotype.
Actual Performance: (1/2018) In progress.
Q9: Repeat GWAS analysis with each year of field data to identify the most robust genetic associations and genetic associations influenced by annual environmental variation.
Actual Performance: (1/2018) In progress.
Q10: None.
Q11: Drought tolerance metrics: validate that canopy temperature measurements and leaf water potential estimates from UAS have correlations of $R^2 > 0.80$ (equivalent or better than current research). These phenotypes are new to commercial field-based phenotyping and thus have no current best practice metric.
Actual Performance: (10/2018) We have validated that the canopy temperature measurements are well correlated to field measurements. Due to a change in priorities of Syngenta, our industry partner at the time, the thermal imaging in general and particularly the leaf water potential calibration was deemed lower priority.
Q11: Validate lodging as a trait in field corn (see Robotics and Breeding workbook for quantitative metrics).
Actual Performance: (10/2018) In July 2017, exploratory data was collected at a corn field in Stanwood, IA that had experiencedlodging due to a weather event. However, due to the inherent time lag in scouting the lodged field and deploying the team to the site, most of the field had started recovering from the lodging event. The data was planned to be used toexplore how to model lodging as a trait. Due to the change in focus after the Deere acquisition and decision not to continue with a Syngenta PoC project in 2018, the milestone is no longer applicable.
Q11: Achieve ≥ 10% improvement in accuracy (RMSE) of terminal harvest biomass compared to year 1. Biomass will bevalidated as described in the Robotics and Breeding workbook.
Actual Performance: (10/2018) Milestone was met.
Q11: Demonstrate ≥ 10% improvement in predictive accuracy compared to year 1.
Actual Performance: (10/2018) Due to a change in priorities of Syngenta, our industry partner at the time, we have not investigated predicting terminal biomass from early season field data.

	Q11: Create GS model to predict performance from genotype and select lines for future crosses. Identify at least one marker useful for breeding decisions for each phenotype. **Actual Performance:** (10/2018) This milestone has been delayed due to the fact that the postdoc working on this left for an industry position and it took 8 months to hire a replacement. Q12: Develop algorithms to estimate select phenotypes. **Actual Performance:** (2/2019) We have validated that the canopy temperature measurements are well correlated to field measurements. Due to a change in priorities of Syngenta, our industry partner at the time, the thermal imaging in general and particularly the leaf water potential calibration was deemed lower priority. Q12: Demonstrate ≥ 10% improvement in predictive accuracy compared to year 1. **Actual Performance:** (2/2019) Milestone was met. Q12: Create GS model to predict performance from genotype and select lines for future crosses. Identify at least one marker useful for breeding decisions for each phenotype. **Actual Performance:** (2/2019) This milestone has been delayed due to the fact that the postdoc working on this left for an industryposition and it took 8 months to hire a replacement. Q12: Identify robust genetic associations and environmentally variable genetic associations. **Actual Performance:** (2/2019) Milestone was met.
4. Germplasm Development	Q1: Increase proposed germplasm for use. **Actual Performance:** (1/2016) Milestone was met. Q1: Identify and begin preparation of screening sites at Westside and Kearney REC Centers. **Actual Performance:** (1/2016) Milestone was met Q1: Send photoperiod sensitive sorghum lines to Winter Nursery in Puerto Rico for seed increase. **Actual Performance:** (1/2016) Milestone was met Q1: Select sorghum lines for field sampling. **Actual Performance:** (1/2016) Milestone was met Q2: Participate in TERRA program genetics coordination plan to determine the sorghum germplasm to characterize. **Actual Performance:** (4/2016) Milestone was met. Q3: Identify and begin preparation of screening sites at Westside and Kearney REC Centers. **Actual Performance:** (7/2016) Milestone was met. Q3: Planting of screening nurseries for initial screening and evaluation of selected germplasm sources for pre- and post-flowering. **Actual Performance:** (7/2016) Milestone was met. Q3: Genotype reference sorghum lines and relevant accessions. **Actual Performance:** (7/2016) Milestone on track to be completed (was completed in August). Q4: SNP profiling of > 1,000 sorghum lines by APEK1 GBS or other metrics agreed by the TERRA genetics coordinating committee (not including lines genotyped by other teams). **Actual Performance:** (10/2016) On track for completion (. Extracted DNA from 648 lines, to be sent out for genotyping). Q4: Metabolite and protein profiles from sorghum lines during non-stressed conditions and during drought acclimation.

Actual Performance: (10/2016) Metabolite profiles obtained. Due to the reduced award size, we decided not to pursue protein profiling.

Q5: 1st Harvest yield and plant data, process for quality parameters and characterization, and perform initial statistical analyses.

Actual Performance: (1/2017) Milestone was met.

Q5: Bag, bulk, and store seed of PS lines if necessary.

Actual Performance: (1/2017) Milestone was met.

Q5: SNP profiling of > 1,000 sorghum lines by APEK1 GBS or other metrics agreed by the TERRA genetics coordinating committee (not including lines genotyped by other teams).

Actual Performance: (1/2017) Milestone was met.

Q6: Based on initial phenotypic screening, planting of second year of both drought and salt tolerant screening nurseries.

Actual Performance: (4/2017) In progress.

Q6: Identify commercial hybrid from Chromatin to add to screening nurseries for drought and salinity testing. Additionally, identify known cultivars as baseline for drought and salinity tolerance. Performed in Y2 and Y3.

Actual Performance: (4/2017) Milestone met.

Q6: Identification of > 5 metabolite and/or protein biomarkers from drought-stress field trials that correlate with drought tolerance and/or terminal biomass yield ($r>0.7$).

Actual Performance: (4/2017) In progress.

Q7: Based on initial phenotypic screening, planting of second year of both drought and salt tolerant screening nurseries.

Actual Performance: (7/2017) Milestone met.

Q7: Identification of > 5 metabolite and/or protein biomarkers from drought-stress field trials that correlate with drought tolerance and/or terminal biomass yield ($r>0.7$).

Actual Performance: (7/2017) In progress.

Q8: Establish nurseries to screen for salinity tolerance.

Actual Performance: (10/2017) Milestone met.

Q8: Identify drought tolerant cultivars with less than 40% reduction in biomass yield compared with irrigated (control) treatments.

Actual Performance: (10/2017) Milestone met.

Q8: Harvest nurseries at Chromatin, KARE, and WREC.

Actual Performance: (10/2017) Milestone met.

Q8: Isolate RNA from lines that we will propose to sequence through the JGI CSP.

Actual Performance: (10/2017) In progress.

Q8: Validation of biomarker results ($r>0.7$).

Actual Performance: (10/2017) In progress.

Q9: Identify commercial hybrids from Chromatin to add to screening nurseries for drought and salinity testing.

Actual Performance: (1/2018) In progress.

Q9: Metabolite and protein profiles from sorghum lines during non-stressed conditions and during acclimation to salinity stress.

Actual Performance: (1/2018) Sampling at Westside could not be managed due to logistic reasons and lack of manpower. (Sampling at KARE required all available resources).

Q10: Identify commercial hybrids from Chromatin to add to screening nurseries for drought and salinity testing.

Actual Performance: (4/2018) In progress.

Q11: Establish nurseries to screen for salinity tolerance

Actual Performance: (10/2018) Milestone was met.

	Q11: Harvest nurseries at Chromatin (TX), KARE (CA), and WREC (CA). **Actual Performance:** (10/2018) Nurseries at KARE and WREC will be harvest completely by the first week of November. Foragesamples from the Exotic parents, which are photoperiod sensitive have been harvested, weighed, dried, and weighed again and dried samples are being shipped to Chromatin for NRI analyses for compositional characteristics related tobioenergy production. Q11: Metabolite and protein profiles from sorghum lines during non-stressed conditions and during acclimation to salinity stress. **Actual Performance:** (10/2018) Milestone was met. Q11: Identification of ≥5 metabolite and/or protein biomarkers from salt-stress field trials that correlate with high-salinity tolerance and/or terminal biomass yield (r>0.7). **Actual Performance:** (10/2018) In progress. Q11: Validation of biomarker results (r>0.7). **Actual Performance:** (10/2018) In progress. Q11: Validation of biomarker results for drought and/or salt-stress treatments (r>0.8) as part of a protocol for biomarker-assisted breeding. **Actual Performance:** (10/2018) Eight leaf replicates from 120 Sorghum genotypes (~4,000 leaf samples) were harvested at two time points (pre-and post-flowering drought) from KARE drought nurseries in 2017. All 4,000 leaf samples were lyophilized ground and cataloged for various downstream analysis. A subset of 40 high-priority genotypes were identified based on the drone-derived wet-biomass (BWET) values from well-watered controls (BWETWW) and pre-flowering droughtexposed plants. Q12: Identify drought tolerant cultivars **Actual Performance:** (2/2019) Milestone was met. Q12: Identify saline tolerant cultivars. **Actual Performance:** (2/2019) In progress. Q12: Metabolite and protein profiles from sorghum lines during non-stressed conditions and during acclimation to salinity stress. **Actual Performance:** (2/2019) Samples have beemn collected and are stored in the freezers. However, no funding remains for analyses. Q12: Untargeted metabolite and protein profiling of field samples from different developmental stages under non-stressed and drought conditions. **Actual Performance:** (2/2019) We processed and analyzed ~ 4,000 samples from the 2017 field trial. Spearman Rank correlations between 11,579 metabolite features and drone data for biomass accumulation revealed 40 highly correlated features (ρ>0.7 or <-0.7); all but one are unknown. Q12: Validation of biomarker results (r>0.7). **Actual Performance:** (2/2019) No longer applicable. Q12: Validation of biomarker results for drought and/or salt-stress treatments (r>0.8) as part of a protocol for biomarkerassisted breeding. **Actual Performance:** (2/2019) No longer applicable.

Project Activities

The goal of CASP was to accelerate breeding of biomass sorghum [*Sorghum bicolor* (L.) Moench] by identifying genotypes exhibiting high yield under well-watered, pre- or post-drought and/or salinity-stress conditions. We did this by combining high-throughput, non-invasive drone phenotyping with genomics and molecular profiling. Genotyping by sequencing (GBS) and GWAS based on drone data from the 648 genotypes yielded 213 high quality, reliable GWAS peaks for single-nucleotide polymorphisms (SNPs) associated with biomass and/or drought tolerance traits. A subset of 120 genotypes based on drone and GWAS data were selected for molecular phenotyping, which yielded data sets of around 12,000 metabolite. The molecular profiling, in combination with GWAS, has the potential to provide mechanistic understanding for genes and pathways responsible for high biomass accumulation in sorghum on marginal land. This information can be utilized for identifying sorghum cultivars with desirable phenotypes solely based on genomic, protein and/or metabolite data.

Project Outputs

Publications

- Handakumbura P.P. *et al.* Plant iTRAQ-based proteomics. (2017) **Curr. Protocols Plant Biol.** 2, 158-172, https://doi.org/10.1002/cppb.20052
- Spindel, J. E. et al. Association mapping by aerial drone reveals 213 genetic associations for Sorghum bicolor biomass traits under drought. (2018) **BMC Genomics** 19, https://doi.org/10.1186/s12864-018-5055-5.
- Photosynthesis in Sorghum by Combining C_3 and C_4 Metabolism. (2019) *In* **Oxygen Production and Reduction in Artificial and Natural Systems**. Ruban, A.V. Ed. Chapter 19, pp 397-412. World Scientific Publishers, https://www.worldscientific.com/doi/abs/10.1142/9789813276925_0019

Status Reports

Quarterly reports to ARPA-*E.*

Follow-On Funding

N/A

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov