

## EPICON Accomplishments

August 15, 2015 – August 14, 2021

### What were the major goals of the project?

Major aims of EPICON research were the identification, in *Sorghum bicolor* (L.) Moench and its associated microbiome in the field, the mechanisms involved in temporal and spatial responses to drought. The C4 crop, sorghum, was chosen for the EPICON studies most importantly for its ability to tolerate drought, while continuing to accumulate biomass and grain. Analysis of plant transcriptional, metabolomic and proteomic data from the plant and metagenomic and metatranscriptomic data from the microbiome is identifying genes and potentially genetic markers, associated with drought tolerance, that can provide avenues for crop improvement. This information is providing a platform for ongoing efforts to validate gene function via gene engineering and editing. Equally important is the information gained allowing identification of bacterial and fungal communities associated with the plant, especially those related to drought response. Together this data provides insights into the dynamic nature of the plant-microbiome relationships, establishing their importance in contributing to the plant's drought tolerance.

Specific goals of the project are listed below.

- I. Determine mechanisms of transcriptional and epigenetic processes over space and time in restructuring regulatory and metabolic landscapes in sorghum during drought by:
  - a. Determining differential gene expression during development and drought using high-throughput transcriptional profiling.
  - b. Characterizing developmentally and drought-induced RNAs and their target genes.
  - c. Mapping transcriptional regulatory circuitry involved in drought response and key regulatory transcription factors.
  - d. Performing metabolomic profiling to better define downstream effects of transcriptional changes.
- II. Explore the role of the microbiome in enhancing drought tolerance in sorghum by:
  - a. Identifying organismal composition of soil, rhizosphere and phyllosphere during plant development and before, during and after pre- and post-flowering drought.
  - b. Determining fluctuations in functional capacity and transcriptional activity of sorghum-associated bacterial and fungal communities during plant development and before, during and after pre-flowering and post-flowering drought.
  - c. Attempting to establish relationships between changes in associated microbial community dynamics with temporal and spatial transcriptional, proteomic and metabolomic profiles *in planta*.
  - d. Using metagenomics and metatranscriptomics to identify bacterial and fungal endophyte relationships and their potential to confer increased drought tolerance.
- III. Explore functions of genes, identified via omic systems biology approaches, by engineering and editing sorghum to determine the role of genes in enhancing drought tolerance.
- IV. Integrate phenotypic, transcriptomic, metabolomic, (proteomic), metagenomic and metatranscriptomic data to develop models to improve understanding of mechanisms sorghum and its microbiome tolerate drought.

### What was accomplished under the goals of EPICON?

#### Specific Aims

**(i) Impact of pre-flowering and post-flowering drought on phenotypic and agronomic performance**

A drought screening nursery at the Kearney Agricultural Research and Extension (KARE) Center located in Parlier, California was created to explore and define transcriptomic control mechanisms and microbial impacts on temporal and spatial responses of *Sorghum bicolor* (L.) Moench to drought under field-scale conditions. Two sorghum genotypes with known pre- and post-flowering drought stress susceptibility and tolerance were selected for evaluation in replicated, controlled field-scale irrigation studies. These involved precise applications of water to mimic pre- and post-flowering stress and control conditions. Field laboratories were established weekly to collect leaf, root and soil samples that were flash-frozen on-site, transported on dry ice to the laboratory where they were stored under minus 80 conditions, ground and shipped to collaborating laboratories for various omic studies: transcriptomics, metabolomics, proteomics, metagenomics and metatranscriptomics.

From an agronomic perspective, three years of phenotypic and agronomic data using field samples indicated control plants outperformed both pre- and post-flowering drought-treated plants of both genotypes in grain and biomass yield and most other phenotypic traits, e.g., plant height, days to flowering (**Table I**). Pre-flowering stress impacted both cultivars to differing extents, compared to plants watered under control conditions. RTx430, characterized as pre-flowering drought-tolerant and post-flowering drought susceptible, outperformed, during pre-flowering drought stress, BTx642, characterized as pre-flowering drought susceptible and post-flowering drought-tolerant (a stay-green variety). Despite eight weeks without water during post-flowering drought, the three-year average showed that lbs/ac of grain was 105% of control for one sorghum variety (BTx642) and 95% for the other (RTx430). For tons/ac of forage, yields were 84% versus control for both varieties. There was a reduction in both yield and 1000-seed weight in RTx430 during post-flowering drought. During eight weeks of pre-flowering drought, RTx430 outperformed BTx642, but both genotypes yielded grain less well than those varieties under controlled watering conditions (56.7% for 430 vs 40.0% for 642).

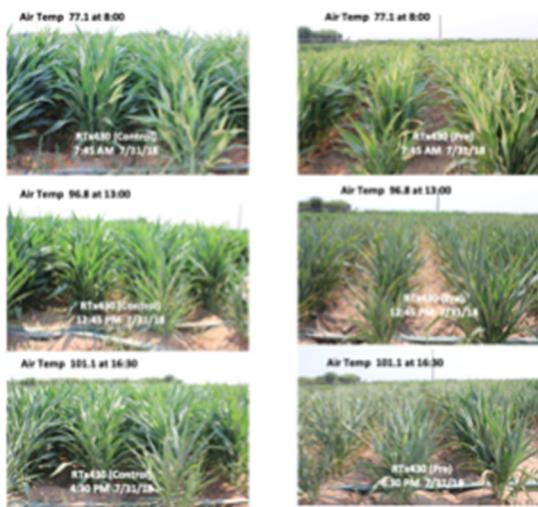
**Table I:** Average agronomic characteristics (2016-2018) for two cultivars at Kearney Agricultural Research & Extension Center grown under control and two drought-imposed stresses.

<b>BTx642</b>	Days to Flowering	Plant Height (cm)	Forage (65%) (t ac <sup>-1</sup> )	Grain (13% moisture) (lbs ac <sup>-1</sup> )	1000 seed wt (g)
Control	59.6	104.0	16.1	3122.5	28.6
Pre-Flowering	66.3	86.7	10.6	1248.8	23.6
Post-Flowering	59.3	88.5	13.5	3295.7	26.3
<b>RTx430</b>					
Control	64.5	120.5	14.7	2917.4	33.5

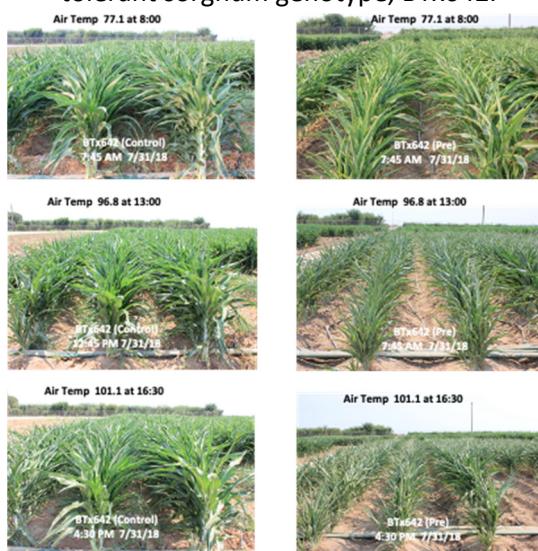
Pre-Flowering	73.1	95.3	10.4	1653.6	30.6
Post-Flowering	64.3	118.2	12.3	2777.1	27.0

California is ideal for research studying drought and its impacts on crops, such as sorghum and maize. Its Mediterranean climate, with little or no rainfall during the summer months, allows for precise, controlled irrigation studies that mimic drought conditions on a field-level scale. **Figures 1 and 2** demonstrate RTx430 and BTx642 responses under varying drought conditions. Not only can basic research trials be conducted, as demonstrated by EPICON data, but replicated field trials and evaluation of large-scale genotype responses can also be performed.

**Figure 1:** Phenotypic response during pre-flowering drought stress in pre-flowering drought-tolerant sorghum genotype, RTx430.



**Figure 2:** Phenotypic response during pre-flowering drought stress in post-flowering drought-tolerant sorghum genotype, BTx642.



### **(ii) Differential gene expression monitored with high-throughput transcript profiling.**

Over the three-year course of the EPICON field studies, gene expression was profiled from approximately two thousand sorghum samples using mRNA-seq, representing one of the largest and most comprehensive field surveys of drought undertaken at the molecular level. A very large proportion of the sorghum transcriptome was observed to respond to drought, many reproducibly so (30-50%) across successive years of field sampling. However, the remainder of the genes exhibited highly variable expression patterns over the three years, possibly attributable to different soil environments and/or different weather patterns.

An initial detailed analysis of data from 2016 (Varoquaux et al. 2019) revealed that >40% of transcriptomes had transcriptional responses to drought under at least one condition, with roots showing larger expression changes than leaves. In addition, a large overlap was found to exist in genes affected by pre-flowering vs. post-flowering drought; however, a subset showed a drought response in only one drought condition. While gene expression changes due to drought were largely shared between the two genotypes, many genes varied in the magnitude or duration of their expression changes. Through *de-novo* assembly and reannotation of the two sorghum genotypes BTx642 and RTx430, using generated Pac-Bio genome sequences for those two varieties, several hundred genes were shown to be present in only one genotype. Also some genes were found to be present in the reference BTx623 genome (to which all of year 1 data were mapped), but were not present in RTx430 or BTx642. Analysis of photosynthesis differences showed that the stay-green genotype, BTx642, maintains higher levels of photosynthetic gene expression during post-flowering drought versus RTx430. Transcriptomic data also suggest that there is a potential role for reactive oxygen species-scavenging and glutathione-S-transferase, a gene associated with an identified stay-green QTL.

Differentially expressed genes were grouped into distinct clusters based on their expression patterns in root and leaf tissue. Within each cluster, many putative transcription factors were also found to have similar patterns, which can be used to construct transcriptional networks governing drought (work ongoing). Several transcriptional patterns under differing water treatments were characterized based on biological function. One such group showed strongest expression decreases in roots versus control under both drought conditions. These genes strongly overlap with genes activated when arbuscular mycorrhiza fungi (AMF) are absent. When plants grown under pre-flowering drought conditions were rewatered, expression of these genes in the more pre-flowering drought-tolerant genotype, RTx430, resumes normal expression quicker than in the post-flowering drought-tolerant variety, BTx642, suggesting quicker resumption of this symbiosis in that genotype. (See Studies of the role of fungi in the response of sorghum to drought).

While these insights (currently derived primarily from a single-year field study) offer an unprecedented glimpse into the molecular responses of sorghum to drought, the fact that many genes exhibit variability (both in magnitude and significance) across subsequent years suggests that simplistic explanations of drought sensing, signaling, and adaptation will be difficult to achieve. Thus, it will be critical to consider reproducible signals emanating from multiple field trials, and in addition to not discount the many environmental variables (e.g., soil type and temperature) as having a combinatorial impact on overall drought responses.

### **(iii) Impacts of differentially methylated regions on transcriptional regulation.**

Bisulfite-seq analysis of ~50 samples from year 1 leaf identified regions with significant differences due to drought. Potential differences in CG methylation levels in 783/1708 (RTx430/BTx642) regions between pre-flowering drought and control; 105/139 regions covered genes; 47/66 had apparent gene expression differences. Analysis of CHG and CHH methylation

revealed small, but widespread, shifts in CHG methylation levels in pre-flowering drought. Analysis of BS-Seq data under control conditions revealed changing methylation patterns during normal plant development with all three forms of methylation having specific regions showing developmental changes, as well as global shifts in CHH methylation levels over time. While some global patterns of methylation change and a small number of differential methylation associated with gene expression were observed, none of the changes were associated with drought-associated or drought-recovery genes. As a result, we decided to focus our attention on the much more easily interpretable and informative transcriptomic and metatranscriptomic aspects, generated by the EPICON study. The latter information provided clear insights into potential mechanistically important drought- and recovery-associated genes.

**(iv) ChiP-Seq and global histone profiling.**

ChIP-seq data exists for H3K4me3 in 66 of 88 samples, with only 37 samples being high quality (> 20k narrow peaks/sample). Several repeated attempts were made to recover missing samples; however, when it was determined that failure was strongly associated with specific tissue-samples, the decision was made to terminate this research direction due to diminishing returns on resources and time commitments. These efforts, however, led to an effective pipeline for the H3K4me3 chromatin mark for extraction, immunoprecipitation and sequencing of field-grown sorghum. However, as the data set was incomplete, and extensive analysis of the existing samples did not provide compelling insights into any role for H3K4me3 in drought or drought-recovery, efforts into further exploration of these data sets did not seem advisable. Untargeted histones of select 2016 leaf samples, analyzed by LC-MS, led to discovery of novel drought-related histone posttranslational modifications. Significant changes in histone terminal clipping in late development for drought-stressed samples is potentially related to epigenetic control via chromatin modifiers (Zhou et al. 2019, 2021). Year 2 samples from additional time points will be analyzed via LC-MS data.

**(v) Metabolomic and proteomic profiling during drought.**

This aspect of EPICON represents one of the largest multiplexed proteomics datasets of sorghum under drought conditions. Most previous proteomics datasets were produced using 2D gels, followed by mass spec analyses; a recently published overview of sorghum under drought stress had very limited proteomic data (Ojolo et al. 2018. *Front Plant Sci* 9:1232 ). One of the most interesting observations was made when combining proteomics with metabolomics data in samples collected in the 2016 field season. This analysis demonstrated that, while metabolite and protein functional category enrichment and depletion in each tissue during either pre- or post-flowering drought remained conserved over several consecutive weeks of drought, individual proteins/enzymes making up the significantly changing functional categories are, for the most part, distinct and unique from time point to time point (i.e., over developmental time).

BTx642, relatively more pre-flowering drought susceptible compared to RTx430, showed unique protein enrichments in organic substance metabolism and catabolism as well as organonitrogen metabolic processes. The elevated organic catabolism could explain increased ROS generated during mitochondrial oxidative metabolism. From these results we hypothesize that RTx430 is better at balancing catabolism for energy production through either satisfactory amelioration of generated ROS or more efficient inhibition of electron capture, transfer or quenching. This results in a ROS stress response that is not as strong as it is in BTx642.

During post-flowering drought, when BTx642 shows greater drought tolerance compared to RTx430 (see “Impact of pre-flowering and post-flowering drought on phenotypic and agronomic performance”), RTx430 has a greater number of functional enrichments, dominance in heat

shock/chaperonins and response to ROS; specifically these were more unique to RTx430. These proteins include those involved in protein folding, response to heat/temperature stimuli, inorganic substances, hydrogen peroxide, and protein complex oligomerization and those related to ATPases. BTx642 was uniquely enriched in proteins involved in ion binding, photosynthesis and ROS processes. That specific array of enzymes enriched during post-flowering drought in BTx642 might make that genotype better able to ameliorate or neutralize ROS than in RTx430.

Significant constitutively increased metabolic markers of drought-stress in RTx430 and BTx642 included L-proline, branched chain amino acids (BCAAs) and aromatic amino acids. Proline is well-known to accumulate under drought stress in plants and is considered to act as an osmolyte and radical/ROS scavenger. The BCAAs, i.e., leucine, isoleucine and valine, and aromatic amino acids, tryptophan, phenylalanine and tyrosine, are hypothesized to function as alternative electron donors in the respiratory chain under stress conditions. They also might act as an alternative source of respiratory substrates and as precursors of many natural products, including alkaloids, phytoalexins, cell wall components, auxins, and other phytohormones.

In coordination with the Coleman-Derr lab, glycerol 3-phosphate (G3P) was found to be highly abundant in roots, especially as the duration of both pre-flowering and post-flowering drought periods increased (see “Studies of the role of bacteria in the response of sorghum to drought”; Xu et al., 2018). G3P enrichment in leaves, was followed by a shift to greater enrichment in roots as drought progressed. This is shown in the G3P accumulation in leaves in the early pre-flowering drought period, followed by a drop to normal levels and a shift to root enrichment during prolonged pre-flowering drought. In post-flowering drought, G3P accumulation increases in the roots as drought is prolonged and this corresponds with correlative decrease in leaves.

Results from our proteomics and metabolomics studies point to several avenues for future study. These include targeted studies of genotype-specific chaperonin/heat shock proteins and of ROS production, coupled with identification of specific mechanisms of ROS amelioration. In addition, work could focus on post-translational ubiquitination/proteasome activity, critical in the turnover of unfolded/heat-damaged proteins. That, along with characterizing other proteolytic activities, would add further insights into processes occurring during sorghum’s response to drought. Finally, characterization of spatially resolved sorghum root exudation during drought could be used, along with studies employing stable isotope labeling, to explore G3P transport mechanisms, e.g., plasmodesmata or other symplastic enhancements, found to be universally enriched in leaves and roots in both RTx430 and BTx642 during drought..

#### **(vi) Studies of the role of bacteria in the response of sorghum to drought.**

Collectively, this work represents one of the first high-resolution, longitudinal crop metagenomics efforts using field derived root-associated bacterial microbiomes. Based on three years of field-based microbiome analysis, new patterns of bacterial root microbiome dynamics under drought stress were discovered and described. Following exposure to drought, the root microbiome of sorghum shifts to an Actinobacterial dominated state, including shifts in both community composition and activity (Xu 2018). Interestingly, the compositional shift is rapidly reversed upon rewetting. This transition has also been reported to occur in additional plant species by other groups (Fitzpatrick 2018, Naylor 2016, Santos-Medellin 2018). This fact, along with the strong phylogenetic signal within the bacterial phyla of Actinobacteria, suggests that this is an ancient and evolutionarily selected for phenomena.

Using a holomic approach, evidence was discovered that shifts in host root metabolism and nutrient status are correlated with changes in the microbiome (Xu et al., 2021). Specifically,

changes in carbohydrate, amino acid, and secondary metabolite were found to be linked to increases in microbiome activity. One specific metabolite, glycerol-3-phosphate (G3P, was found to be produced by the plant at significantly greater levels under drought stress, and that the bacterial microbiome greatly increases activity of transport of this metabolite under the same conditions (Xu et al.,2018). As G3P is a precursor to cell wall biosynthesis, we hypothesized that increases in G3P production may give gram positive bacteria, which have much thicker cell walls, an opportunity for increased growth. Editing efforts were aimed at knocking out two enzymes involved in the biosynthetic pathway to try to understand this relationship more fully (see **Explore via plant transformation function of specific genes in drought tolerance of sorghum**).

Finally, an unexpected connection was discovered between plant iron metabolism, drought stress, and the root microbiome (Xu 2021). Drought was shown to lead to downregulation of iron uptake within the root, and upregulation of iron storage mechanisms. Using a genetics-based approach, evidence was provided that altered iron content in the sorghum root, caused by alterations in iron transport activity during drought, may be partly responsible for the observed shifts in Actinobacterial concentrations. A large collection of sorghum root isolates has been created to help probe Actinobacterial potential for altering host phenotypes in field and lab-based experiments. Ongoing work with this collection will continue to identify ways to manipulate crop yield and fitness under drought stress using approaches based on the microbiome-plant relationship.

#### **(vii) Studies of the role of fungi in the response of sorghum to drought.**

These studies followed five paths. First, efforts were focused on fungi that form obligate, mutualistic, symbiotic relationships with sorghum roots. These arbuscular mycorrhizal fungi (AMF) provide P, K, N and water to the plant in exchange for ~15% of the plant's sugars and fats. Initially AMF were identified as having the strongest ecological succession of any fungi. A signal for temporal change in AMF community similarity was shown to be 40-fold stronger than most recent studies (Gao et al. 2019), likely due to (i) weekly plant samplings from seedling to fruit maturity and (ii) use of the fungal DNA barcode to identify species in an agricultural environment. These studies demonstrated patterns of nestedness and turnover and processes of immigration and extinction. This was the first evidence that AMF species co-exist, rather than co-occur, by demonstrating negative, density-dependent population growth for multiple species. This shows the advantages of using fungi to test basic ecological hypotheses over periods as short as one season.

Second, abundance of AMF was compared with expression of plant genes following imposition of drought. Soil, rhizosphere, root and leaf samples of both genotypes under three watering conditions over 17 weeks revealed that the most important fungi during drought are AMF and AMF community composition was not affected by pre- or post-flowering drought, but abundance declined (Varoquaux et al. 2019). A total of 470 plant genes, expressed only in AMF presence, had strong, coordinated expression reduction during pre-flowering drought, correlating with lower AMF abundance. When water was restored, genes sorghum uses to manage AMF rebounded along with abundance of AMF fungi.

Third, the role of specific sorghum genes in regulating AMF symbiosis was studied by determining gene expression in pathways essential to AMF colonization. Strong, differential correlation was found between the shift from ruderal to competitive AMF, plus sorghum genes whose products (i) produce and release strigolactone signals, (ii) perceive mycorrhizal-lipochitinoligosaccharide signals, (iii) provide plant lipids and sugars to AMF and, (iv) import minerals and water provided by AMF. This demonstrated adaptive strategies, evolved by AMF

and their hosts, providing rationale for selecting for AMF to reduce inputs and maximize yield in commercial agriculture (manuscript in review at *Molecular Ecology*).

Fourth, unlike AMF, other fungi were shown to have strong community composition shifts in root and rhizosphere in pre-flowering drought; changes reverted to control when drought ended. While plant selection is key to assembling AMF communities, assembly for all fungi showed a stochastic force, drift, played a role in assembling fungal communities (Gao et al. 2020).

Lastly, interactions of fungi and bacteria, living in the plant and rhizosphere and important to drought response, were investigated. These populations could mitigate drought stress; however, microbiome stability is poorly understood. Resistance and resilience of fungi and bacteria to drought in agricultural settings were studied using community composition and microbial interactions. In general, drought disrupts microbial networks. Surprisingly, co-occurrence networks among specific rhizosphere fungi and leaf bacteria were dramatically strengthened by pre-flowering drought, also for AMF networks in the rhizosphere. Both pre- and post-flowering drought increased relative frequency of positive correlations (manuscript in revision, *Nature Communications*). Our discoveries raise new questions. i) Do other stresses strengthen microbial networks? ii) What molecular and chemical mechanisms strengthen microbial networks? iii) Can networks mitigate drought stress in agriculture, perhaps by introducing hub taxa?

#### **(viii) Explore via plant transformation function of specific genes in drought tolerance of sorghum**

One of the ultimate goals of EPICON was to use holo-omics to uncover interesting gene targets for functional validation in sorghum. This approach had the potential to generate a large number of hypotheses and to highlight numerous genes of interest (GOIs). To prepare for the effort to probe gene function, using engineering and editing, it was necessary to dramatically improve the speed and efficiency of sorghum transformation and editing. This was accomplished with a transformation method using the morphogenic genes, BBM and WUS, previously shown to improve maize transformation.

Efforts to develop those tools began with a morphogene-assisted transformation (MAT) method, which led to a quicker transformation process, reduced genotype dependence and improved transformation efficiency (Aregawi et al., 2020; Aregawi, Shen et al., 2021). The accelerated transformation time was nearly half the time required with classical callus-based, non-MAT approaches. These efforts also led to expanded numbers of amenable genotypes, including one genotype used in EPICON field trials, BTx642. The other genotype, RTx430, is a standard transformation variety. Transformation efficiency using a classical callus-based transformation method was ~2%, but using a MAT-based approach we achieved efficiencies around 50%, i.e., ~50% of independent regenerated plants were transgenic. To minimize transgene silencing, digital droplet PCR was used to find single-copy transformants, which occurred at >60% frequency. A novel method was used to determine transgene integration independence, based on an *Arabidopsis* method to perform high-throughput mapping.

Another advance in transformation, termed altruistic, was developed in which a gene of interest (GOI), i.e., RFP for proof-of-concept, was introduced in a separate *Agrobacterium* strain from the one with morphogenes in a 9:1 ratio. Because WUS leaves the cell and influences neighboring cells, this leads to cells receiving only the GOI but not the morphogenes, but those cells can be developmentally triggered by diffusible WUS to regenerate a plant. Transformed RFP plants without morphogenes were generated, but efficiencies varied among replicates and large numbers of plants were generated with both the morphogenes and the GOI.

After demonstrating significant improvements in transformation efficiency with the MAT-base approach, CRISPR/Cas9 editing technology was combined with BBM/WUS. This MAT-based approach was successfully used to edit an exemplary target gene, phytoene desaturase. This construct with a maize ubiquitin1 promoter driving a maize codon-optimized Cas9; the *Oryza sativa* U3 promoter driving gRNAs was used to edit the two sorghum *pds* genes. Genotyping revealed a 22.7% editing efficiency.

Based on that success, attention was turned to editing genes identified through EPICON transcriptomic analyses as being involved in drought tolerance. One set of genes was aimed at the relationship between bacteria and sorghum. Using a combination of EPICON metagenomic, transcriptomic, and metabolomic data, ROS and G3P production were chosen as our targets, since both pathways showed clear transcriptomic and metabolic shifts under drought (see Metabolomic and proteomic profiling during drought). The focus was narrowed to three genes: catalase (CAT2) involved in detoxification of hydrogen peroxide, and two biosynthetic enzymes involved in G3P production (GLI1 and GLY1). Two sgRNAs were designed to target these genes and introduced into RTx430 using the MAT-based transformation approach. Of nine CAT2 transformants, genotyping revealed that none were edited. Of the twelve GLI1/GLY1 transformants, three had significant edits in GLI1, as demonstrated with Synthego analysis, none in GLY1. The GLI1 T<sub>0</sub> plants are currently growing in the GH for further analysis.

Another set of genes chosen were transcription factors involved in post-flowering drought in sorghum. A NAC TF, Sobic.006G147400, was shown in EPICON transcriptomic studies to be induced at the end of post-flowering drought and is strongly suspected to control sorghum stem water content. Overexpression of NAC was shown by others to convert a sweet sorghum stem into a dry-type sorghum stem; GWAS studies demonstrated strong linkage of a loss of function mutation at this locus to increased water-content in the stem. A large difference in basal expression of this gene in stay-green BTx642 vs RTx430 was seen in the root in all three years of EPICON data. A CRISPR-Cas9 mutation at this site should yield a sweet sorghum-type stem. NAC-specific gRNAs were put into a MAT-type transformation construct and two transformations were performed on RTx430, with a transformation efficiency ~32% and 28%. From the two experiments a 6.9 % editing efficiency was obtained. However none of the editing events led to a gene knockout; five plants have been moved to the GH. A second TF, Myb, Sobic.009G03g500, shown to be induced at the end of post-flowering drought, is an ortholog of a myb TF that represses stomatal opening. As with NAC, gRNAs were designed and MAT-based transformations were conducted on Tx430. Two transformation experiments were conducted with 47 of 60 regenerated transformed; however, only one plant was edited but did not survive. Efforts are underway to improve editing efficiency through construct and methodology modifications.