COVER PAGE

DOE STI Product/Report Number	DOE/USDA SC0001490
STI Product Title	THE HUNT FOR GREEN EVERY APRIL: FACTORS AFFECTING FITNESS IN SWITCHGRASS Final report
STI Product Type and Reporting Period	Final Report for Period 09/01/2009 to 08/31/2013
Date of Issuance or Publication	Date Issued/Published December 2014
Author	Gautam Sarath
Sponsoring Organization	PREPARED FOR THE UNITED STATES DEPARTMENT OF ENERGY OFFICE OF SCIENCE
Award/Contract/Financial Number	Work Performed Under Contract No. SC0001490

TABLE OF CONTENTS **PAGE NUMBER**

ABSTRACT

This grant funded work was undertaken to develop fundamental biological knowledge of the factors affecting the complex plant trait "fitness" in switchgrass (*Panicum virgatum* L.), a plant being developed as a biomass crop. Using a diverse range of latitudinally-adapted switchgrass plants, genomic, molecular and physiological studies were performed to track a number of different aspects of plant genetics and physiology over the course of the growing season. Work was performed on both genetically unrelated and genetically related plants. Plants were established in the field from seedlings raised in a greenhouse, or from clones present in other field nurseries. Field grown plants were used as the source of all tissues. The three objectives of this proposal were:(**1**) *Transcript Profiling, Metabolomics, and C and N Partitioning and Recycling in Crowns and Rhizomes of Switchgrass over two growing seasons;* (**2**) *Gene Profiling During Regreening and Dormancy of Bulked Segregants;* (**3**) *Extent of Linkage Disequilibrium in Populations for Adaptation and Fitness Traits Being Developed for Central and Northern USA, that Show Significant Heterosis.* **Objective 1 results**: Plants were labeled using ^{13}CO2 (a stable isotope) using an acrylic chamber constructed specifically for this purpose. Plants became labeled with ^{13}C and label decayed in aerial tissues over the course of the growing season. Varying amounts of ^{13}C were recovered in the rhizomes. These data are being analyzed. Plants were also labeled with ^{15}N-urea. Plants absorbed significant amounts of label that was remobilized to the growing shoots. N-dynamics would suggest that a portion of the ^{15}N absorbed into the crowns and rhizomes is sequestered below ground. Variable amounts of ^{15}N were translocated from the shoots to the roots over the course of the growing season. Polar metabolites extracted from a diverse array of rhizomes were analyzed using GCMS. Data indicated that there was a significant shift in metabolite pools over the course of the growing season, and differences in the levels of specific metabolites could be linked to the progression of dormancy. Several metabolites that accumulate in dormant rhizomes were identified. Some of these metabolites could be potentially linked to winter-survival of switchgrass. Extensive high-throughput sequencing was conducted on crown and rhizome samples collected from field grown plants. Initial work was performed on a Roche 454 system. All later work was performed on an Illumina sequencing-by-synthesis system. Some of these datasets have been published as peer-reviewed papers, other data are currently being analyzed and being readied for publication. **Objective 2 results**: Genetically related but phenotypically divergent plants from an octaploid switchgrass population were grown in a replicated field nursery. Rhizomes were harvested at four different times over the course of the growing season from plants with high winter survival and those with lower winter survival. RNA-Seq was performed on harvested materials. Initial analysis suggests that plants with lowered winter survival experience a greater level of cellular stress in dormant tissues. This aspect of plant function is being probed in greater depth. **Objective 3 results:** A total of 592 individual clones with three clonal replications in a randomized complete block design from each of five populations used in Objective 1 studies were rated for heading date in 2012 and 2014, green-up day of year in 2013, anthesis date in 2012, and yield in 2012, they were also subjected to NIR spectroscopy to derive cell wall composition estimates based on prior NIR calibrations. Plants were genotyped via a genotyping by sequencing (GBS) approach from reduced representation libraries constructed with adaptors that identified each individual. Libraries generated with the restriction enzyme *Pst*I and called SNPs using Samtools after alignment to version 1.1 of the switchgrass genome sequence. A total of approximately 40,000 SNPs were found. These were then further filtered to eliminate markers

with a minor allele frequency of < 0.05. The results of population analysis using STRUCTURE with expected population sizes or cluster numbers (K), clearly shows the hybrid composition of the KxS population and discriminated easily between upland (Summer) and lowland (Kanlow) populations. Under an assumption of 5 distinct populations there were detectable differences in allele frequencies between subpopulations within the three Kanlow populations particularly with respect to Kanlow EM and Kanlow base. We detected 110 SNPs with an allele frequency difference of ≥ 0.2 between Kanlow EM and Kanlow base populations, while 120 SNPs showed an allele frequency difference of ≥ 0.15 between Kanlow N1 and Kanlow base populations. These data are being readied for publication.

Final Report

This is the final report for DOE Grant DOE/USDA SC0001490. The document is subdivided into four sections, with one section devoted to each of three objectives, and a final section listing the outcomes (peer-reviewed manuscripts).

Objective 1: **Transcript Profiling, Metabolomics, and C and N Partitioning and Recycling in Crowns and Rhizomes of Switchgrass over two growing seasons.**

The goal of this objective was to derive molecular and physiological data on the below-ground tissues of switchgrass that contribute to perenniality. These tissues are the crowns and rhizomes from which new tillers are formed. There have been no such detailed cellular investigations performed on these tissues from switchgrass plants so far.

C and N partitioning and recycling

Switchgrass plants with divergent winter-hardiness were selected for stable carbon isotope labeling in the field in June 2010. Six plants per population (4 populations) in two replicates blocks were labeled with ^{13}C with 99-atom % $^{13}CO_2$ at the mid vegetative stage. Labeling was performed using two large plexiglass boxes designed and constructed specifically for these studies Figure 1.

Figure 1. Plexiglass chamber for labeling field-grown switchgrass plants.

Plants were labeled twice, one week apart and materials were collected at distinct times over the growing season, based on the phenological stage of the cv Summer plants. Collected plant materials were divided into two equal portions, one portion was dried and ground to obtain total biomass, the other portion was dried and hand separated into leaves, sheaths, stems and inflorescences. In order to validate the design and construction of the plexiglass chambers used in our field studies, we conducted a few experiments to test these chambers both in the greenhouse and field conditions with or without enclosed plants. Results from both depletion and $^{13}CO_2$ labeling studies revealed that the constructed chambers were almost leak free under both greenhouse and field conditions. We were able to label the enclosed plants in the chamber

twice in quick succession with relative ease under field conditions. However, the density of plants and their size need to be taken into careful consideration for good mixing dynamics of the label within the chamber.

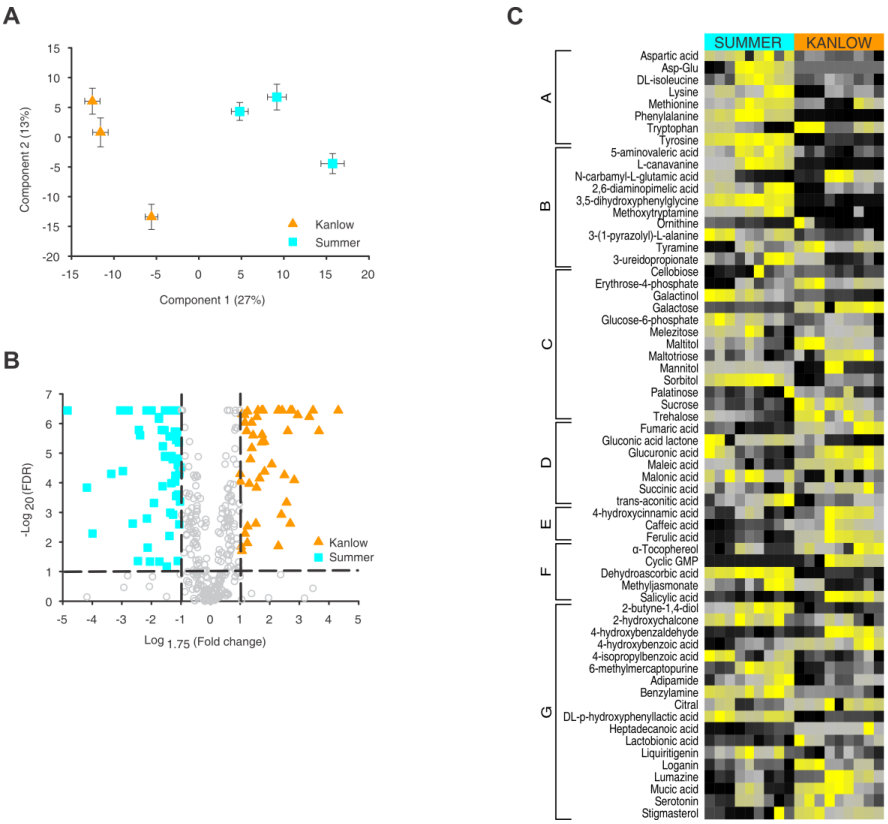
Plant materials were harvested at four different dates extending up to the time after aerial senescence. Materials that were collected have already been analyzed. All plants were well labeled with ^{13}C. In general materials that were collected a week after labeling showed significant labeling in the bulk leaf material and as expected, the label was diluted over the growing period. However, it was observed that materials collected after the active growing season retained approximately 10% of the ^{13}C label. Other analyses indicated that there were no significant varietal differences across the tested switchgrass populations and the labeling of the structural and non-structural carbohydrates was similar to the trend observed with the bulk leaf material. However, the structural cellulosic fractions of the bulk leaf material retained a significant amount (~30%) of ^{13}C label compared to the non-structural starch fraction which retained approximately 15% of the label. In 2011, along with the originally selected 4 populations, a fifth (K x S) was added to more fully understand traits controlling switchgrass winter hardiness. This fifth population was from a bulk cross between an upland and lowland cultivar, specifically selected for improved yields at the Lincoln location. Six plants per population (5 populations) in two replicates blocks were uniformly supplied with ~18 g of N as NH_4NO_3 or urea containing ^{15}N with 10 % final enrichment, dissolved in water and injected into the soil in early May 2011. All harvests have been completed and some of the harvested material have been analyzed for both nitrogen and carbon isotope enrichment observations. In general materials that were collected two weeks after labeling showed a very high ^{15}N label and unlike the ^{13}C label getting diluted rather rapidly during the growing season, the ^{15}N label was heavily retained by the bulk leaf material even after 6-8 weeks after labeling. This indicated that a decent amount of ^{15}N was not only absorbed but get fixed and translocated to a variety of tissues. Due to some technical difficulties, we were unable to finish the entire harvested material but are currently being processed.

Metabolomics

Metabolite profiling of rhizome samples was conducted on polar metabolites using GCMS and subjected to principal component analyses (PCA). PCA of Kanlow and Summer crown and rhizome extracts using GCMS revealed that Summer and Kanlow extracts were clearly differentiated by the first principal component, which explained 27% of the variance. The second component differentiated among the individual genotypes within each cultivar and explained an additional 13% of the variance. Concentration ratios, approximated by major ion count, were coupled with differences that had an FDR <0.05 to generate a volcano plot of the metabolite data. Concentration ratios between the cultivars were considered meaningful if the averaged ion counts for a metabolite had a ratio greater than 1.75 or less than 1.75^{-1}. The stricter criteria resulted in 66 metabolites with sufficient FDR values and ion area ratios in Summer and 47 metabolites in Kanlow. In Summer plants, levels of several amino acids, amine derivatives and some sugar alcohols were significantly higher. In contrast, erythrose-4-phosphate, maltotriose, sucrose and trehalose were detected in greater levels in Kanlow crowns and rhizomes. Further significant differences (based on an FDR <0.05) in the metabolism between Kanlow and Summer rhizomes were evident in the observed levels of several organic acids such as glucuronic acid, a precursor of xylans, and a range of secondary metabolites, including

phenolic acids. Summer tissues also appeared to be in a more oxidized state as compared to Kanlow crowns and rhizomes as they contained significantly higher levels of dehydroascorbic acid, gluconic acid lactone, and lower concentrations of α-tocopherol (Figure. 2).

Figure 2. Metabolite profiling reveals ecotype specific differences in crown and rhizome tissues. (**A**) PCA of overall metabolite profiles observed by GCMS for each of three individual genotypes within each cultivar. Kanlow tissue extracts (orange triangles) are separated from Summer tissue extracts (cyan squares) by the first component. Two-way error bars are shown for each plant that was based on nine separate GCMS runs for each sample. (**B**) Volcano plot showing the $\log_{10}$FDR versus the $\log_{1.75}$ fold change in peak area for major ions for all metabolites detected by GCMS. Significant differences were observed for several metabolites between the two cultivars (Kanlow orange triangles) and Summer (cyan squares). Grey circles are metabolites that failed to show sufficient differences in ion area of had an FDR value of >0.05. (**C**) Heat map shows marked differences in tissue abundances of selected metabolites in Kanlow and Summer crowns and rhizomes. Data are the average of triplicate injections from each of three separate extractions from the three biological replicates of each cultivar. Scale is from high abundance (yellow) to low abundance (black) for each metabolite. Data were subjected to one-way hierarchical clustering based on cultivar, and metabolites were manually reordered into more biologically related groupings (A-G).



These overall results have been extended by metabolite profiling of rhizomes obtained from several field plantings. These additional data indicate a strong accumulation of several sugar alcohols, proline, and related compounds in dormant rhizomes. Many of these compounds are associated with cell protection and dehydration that accompanies dormancy in other plants. It is possible that timely accumulation of these metabolites is crucial to the winter-survival of perenniating tissues.

Transcriptomics

Several different platforms were used for high-throughput sequencing over the grant period. In total we will have deposited over six billion reads from switchgrass tissues into the public databases.

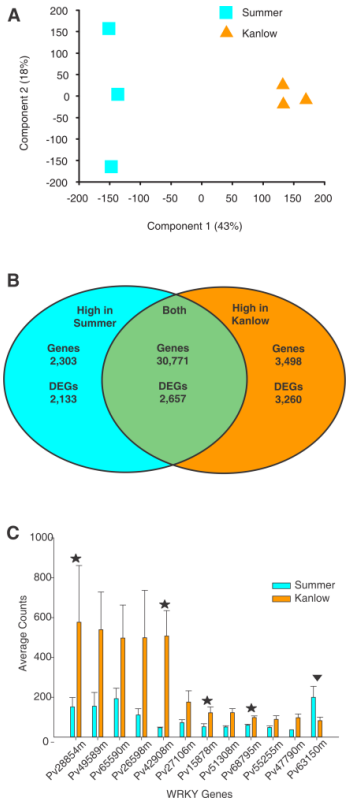
We first started with the Roche 454 FLX pyrosequencing system since this was the major HTS instrument available to us at the start of the grant period, and there was no data available on the switchgrass genome. The crown and rhizome transcriptome of an upland tetraploid switchgrass cultivar cv Summer well adapted to the upper-Midwest was investigated using the Roche 454-FLX pyrosequencing platform. In all approximately 1 million reads consisting of 216 million bases were assembled into 27,687 contigs and 43,094 singletons. Analyses of these sequences revealed minor contamination with non-plant sequences (< 0.5 %), indicating that a majority were for transcripts coded by the switchgrass genome. Blast2Go comparisons resulted in the annotation of ~65 % of the contig sequences and ~40 % of the singleton sequences. Contig sequences were mostly homologous to other plant sequences, dominated by matches to the *Sorghum bicolor* genome. Singleton sequences while displaying significant matches to *Sorghum bicolor* also contained sequences matching non-plant species. Comparisons of the 454 dataset to existing EST collections resulted in the identification of 30,177 new sequences. These new sequences coded for a number of different proteins and a selective analysis of two categories, namely peroxidases and transcription factors resulted in the identification of specific peroxidases and a number of low-abundance transcription factors expected to be involved in chromatin remodeling. KEGG maps for glycolysis and sugar metabolism showed high-levels of transcripts coding for enzymes involved in primary metabolism. The assembly provided significant insights into the status of these tissues, and broadly indicated that there was active metabolism taking place in the crown and rhizomes at the post-anthesis, seed maturation stage of plant development.

We next utilized the Illumina system sequencing-by-synthesis (SBS) Genome Analyzer IIx platform to query the rhizome transcriptomes obtained from plants with contrasting latitudinal adaptation. Each sample yielded approximately 24-27 million reads, of which greater than 77-83 % were mapped to the 0.0 early-release version of the switchgrass genome. Functional annotation of the PviDraft0 genome by Blast2GO resulted in the annotation of 45,487 gene models at an e-value threshold of 1×10^{-15} or lower. EdgeR analysis identified a total of 8,050 genes that were differentially expressed between Kanlow and Summer crown and rhizomes at an FDR <0.05. PCA of this expression dataset effectively separated the three Kanlow samples from the three Summer samples (Figure 3A). Component 1 accounted for 43 % of the variance and separated the two cultivars. Component 2 accounted for approximately 18 % of the variance and separated the three Summer samples. A Venn diagram of the total number of genes found in the Summer and Kanlow datasets is shown in Figure 3B. In total transcripts were detected for 36,572 genes. Of this total ~84 % (30,771 genes) were common to both populations and 2,303 and 3,498 genes were unique to the Summer and Kanlow datasets respectively. About 41 % of the differentially expressed genes (DEGs) had two-fold or greater average expression in Kanlow, and ~26 % had two-fold or greater expression in Summer (Figure 3B).

A more detailed evaluation of the DEGs was performed to identify changes in transcript abundance for three plant-specific classes of transcription factors, namely MYBs, NACs and WRKYs which are known to have multiple functions in cellular processes [36-40]. A scan of the switchgrass genome based on Pfam PF03106 (WRKY), PF00249 (MYB) and PF02365 (NAC)

domains [41], identified 178 WRKYs (~89 WRKYs per diploid genome); 324 MYBs (~162 MYBs per diploid genome); and 239 NACs (~120 NACs per diploid genome). Similar searches for WRKY members in other species gave the following results: Arabidopsis – 72; rice – 94 and 82 in *Brachypodium*. Using a similar approach for MYBs and NACs in other species, we found: Arabidopsis – 256/112 (MYBs/NACs); rice – 233/140 (MYBs/NACs), and Brachypodium – 200/100 (MYBs/NACs). Based on the relative numbers of these genes found in model plants, we expect that approximately 80-90 % of the orthologous switchgrass genes were identified by this search. A total of 199 MYBs (30 DEG) and 108 NACs (20 DEG) were found in the expression datasets. In all 119 WRKYs (~67 % of the all WRKY genes identified in the genome) were found in the Summer and Kanlow datasets. Of these 119 genes, a total of 12 were differentially expressed at an FDR<0.05. A database search indicated that orthologs to four of the WRKY genes upregulated in Kanlow (identified by a star symbol in Figure 3C) were involved in responses to pathogens. In contrast, the ortholog to one WRKY gene upregulated in Summer and has been demonstrated to be an important mediator of cellular responses to senescence (inverted triangle).

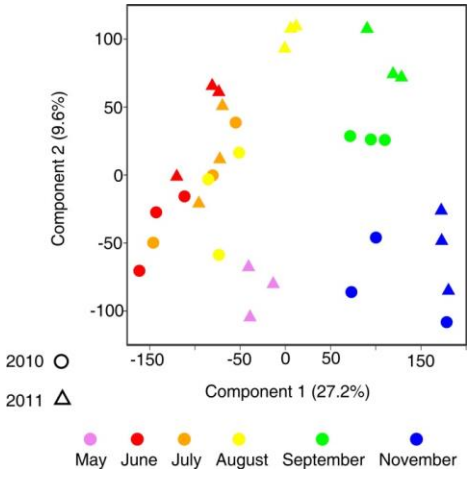
Figure 3. Transcriptomes of Kanlow and Summer plants are different, and show differential enrichment in WRKY genes. (**A**) PCA of transcriptomes of the three biological replicates obtained from each ecotype. Kanlow (orange triangle) were differentiated from Summer (cyan squares) by the first component. (**B**) Venn diagram of numbers of common and highly expressed genes in Summer ($log_2FC < -1$) and Kanlow ($log_2FC > 1$), including DEGs in the Kanlow (orange) and Summer (cyan) datasets. (**C**) Mean counts for twelve transcripts identified as WRKY transcription factors. Stars above orange bars (Kanlow) identify putative switchgrass orthologs of WRKY factors involved in defense response in other systems. The inverted triangle over a cyan bar identifies a WRKY ortholog to an Arabidopsis gene involved in senescence that is upregulated in Summer crowns and rhizomes.



These data showed that differences existed in the metabolite and transcript profiles in the crowns and rhizomes obtained from field-grown switchgrass plants belonging to two populations with divergent photoperiod responses. Crowns and rhizomes obtained from the more southern adapted cultivar Kanlow appeared to be in a "growth-mode" with enrichment in gene-sets associated with biosynthesis and accretion of tissues. In contrast, the northern adapted cultivar Summer appeared to have entered a more quiescent state with greater enrichment of transcripts and metabolites favoring channeling of carbon skeletons through an acetyl-CoA hub. The findings provide an initial framework to understand differences in crown and rhizome metabolism which could be exploited to enhance the latitudinal adaptation of diverse switchgrass populations.

For the last part of the transcriptomics studies, rhizome samples obtained from 4 populations and were subjected to HTS analysis on a Illumina HiSeq 2000 system. Much of these data are still in the analytical pipeline. For the cv Summer dataset obtained from rhizomes collected over two growing seasons, there were approximately 400 M reads, of which 64-66 % could be mapped to the switchgrass genome 1.1 version available at www.phytozome.org. Annotated transcripts from every sample harvested in year 1 and year 2 were separated by PCA (Figure 4). The first component accounted for 27.2 % of the total variance and differentiated individual harvests into three main categories. Transcriptomes at green-up (purple triangles Figure 3) were differentiated from the "growth stage harvests" of June and July and the "senescence/dormant stage harvests" of September and November. During both years, there was overlap between the June and July harvests (Figure 4). Component 2 accounted for 9.6 % of the variance and effectively differentiated the September harvests from the November harvests. Component 2 also indicated that the August harvests from year 1 could be separated from the year 2 harvest. This was consistent with the growth stages of the plants which appeared to be impacted by water availability. During year 1 there was greater total precipitation which appears to have delayed aerial senescence relative to the drier year 2, when plants were senescing at the time of the August harvests (visual observation).

Figure 4. PCA analysis of transcriptomes obtained from field grown cv Summer switchgrass plants over two growing seasons (2010; 2011). Circles are data from 2010 harvests, and triangles are data from 2011 harvests. Each harvest date is color coded.



We also analyzed rhizomes for plant hormones using LCMS. Rhizome GA content, determined as GA_4 equivalents, increased from green-up (May) through July and then significantly declined from a maximum of approximately 102.5 ± 8.3 ng GA_4 g rhizome fresh wt^{-1} to 52.7 ± 6.9 ng GA_4 g rhizome fresh wt^{-1} in tissues harvested after a killing frost. In contrast, tissue ABA content was relatively low for the first 4 harvests (~10-25 ng ABA g rhizome fresh wt^{-1}), then increased significantly to 110.8 ± 7.6 ng ABA g rhizome fresh wt^{-1} in rhizomes harvested in September. Rhizome ABA content had a further significant increase to 180.1 ± 60 ng ABA g rhizome fresh wt^{-1} from the September to the November harvests.

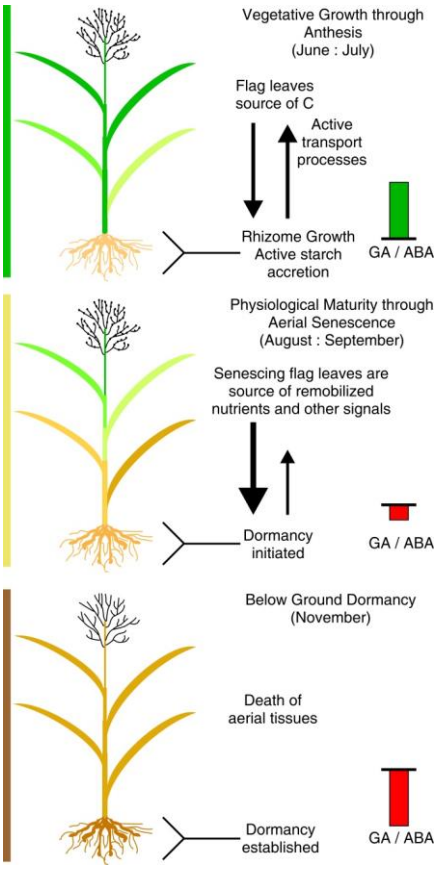
Based on the GA/ABA ratios, August/September appears to be a key transitional point for rhizome exiting from a growth phase and entering dormancy. During this time frame, seeds had reached physiological maturity and aerial senescence, especially of flag leaves, was well underway. Continued increase in tissue ABA content coupled to the decline in tissue GA content appeared to be an important signal for the imposition of dormancy in this upland switchgrass cultivar. There was a greater than 2-fold change in the ABA/GA ratios in rhizomes harvested after a killing frost as compared to the previous harvest. These data indicated that ABA content of rhizomes responds to changes in shoot physiology, and high ABA content reflected both the onset of dormancy (September), and the imposition of full dormancy in the post-frost samples.

Using plant phenology, environmental, hormone and, PCA data, a model for plant growth integrating aerial developmental stages with rhizome metabolic status is proposed (Figure 4). In this model, the focus is more towards developing a framework to interpret the metabolic changes occurring in rhizomes during the onset of dormancy, and those needed to successfully establish and maintain this state. The onset of first visible signs of growth (green-up, May) is followed by a period of active growth during June and July. During this interval, the tiller apical meristem transitions from a vegetative phase to a reproductive phase, followed by the flowering, anthesis and seed set. A key regulatory aspect during these transitory events appears to be the physiological status of the upper leaves, and prominently the flag leaf.

During periods of active shoot and rhizome growth (June/July), rhizome GA/ABA ratios were high, and there was limited variance in the rhizome transcriptomes between these two time periods (Figure 3). These data indicate that many cellular processes were similar between these two harvest dates. Furthermore, maturation of flag leaves into sources for photoassimilates over this part of the growing season, would predict that this is a period of growth of the rhizomes (including roots), and that bidirectional transfer of assimilates and nutrients between the shoots and the below-ground tissues would be active. It is also expected that this is a time of starch accumulation within rhizomes. Subsequent to this time period, seeds would have reached physiological maturity, and tiller senescence, at least as documentable as flag-leaf senescence, would have been initiated. These expected shifts in tissue metabolism, coupled to a large increase in rhizome ABA content could represent the first signs for the onset of dormancy taking place during August/September. The availability of moisture later in the growing season appears to be a driver for delaying aerial senescence and the onset of dormancy in rhizomes (as seen from the PCA analysis) for the 2010 and 2011 transcriptomic datasets (Figure 4), but not the imposition of dormancy in September. The September rhizome transcriptomes for both 2010 and 2011 growing seasons were relatively similar (Figure 4). We predict that the transfer of nutrients from the senescing shoots to the rhizomes had diminished and dormancy-related metabolism became more fully established during September (Figure 5). By the time of a killing frost (late October/early November) shoots had essentially fully senesced and rhizomes were dormant

(Figure 5). This transition appears to be accompanied by continued increase in the ABA/GA ratios and by substantial changes to the transcriptomes (Figure 4). With the loss of nutrient transfer to the rhizomes as a result of the death of the aerial tissues, dormant rhizomes will become dependent on stored C and N compounds for over-wintering and winter survival.

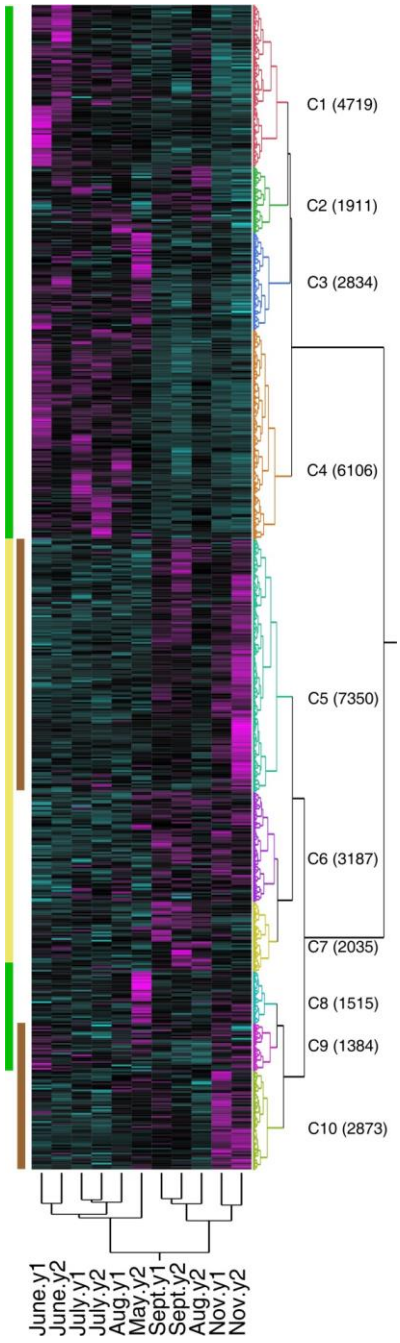
Figure 5. A proposed model of switchgrass aerial and below ground development integrating data from C/N remobilization, plant phenology, transcriptomes and hormone levels in rhizomes. Colored bars on left designate developmental periods, green = growth, yellow = transition to dormancy and aerial senescence, brown = dormant rhizomes, death of aerial parts of the plant. Green and red bars on the right indicate the relative ratios of GA/ABA in rhizomes.



The combined 2-year transcriptomic dataset was analyzed for differentially expressed genes (DEGs), and all of the DEGs identified through a combined analysis of the transcriptomes from both harvest years were subjected to 2-way hierarchical clustering to obtain an overview of developmental patterns of gene expression in rhizomes, and similarities and dissimilarities across the two years of analysis. Except for the august harvests, rhizome transcriptomes collected at equivalent shoot developmental stages for the two years clustered together (Figure 6). Based on these analyses, we could broadly segregate the datasets into 10 clusters (labeled C1-C10; Figure 6). Around 12,800 genes were significantly upregulated during periods of active rhizome metabolism (June and July harvests, green bar, Figure 6). Rhizomes harvested in May contained 4,349 genes that were expressed at significantly higher levels (clusters C3 and C8; Figure 6).

Whereas the transcriptomes from the two September harvests clustered together, the year 1 August rhizome transcriptomes were differentiated from the year 2 August transcriptomes. The post-frost rhizomes harvested in November were similar and contained over 11,607 genes which were expressed at significantly greater levels. These genes were part of three clusters (C5, C9, C10; Figure 6).

Figure 6. All differentially expressed genes (DEGs) in the combined two year transcriptome datasets for cv Summer rhizomes over the course of the growing season(s). The colored bars on the left indicate the developmental stage (see Figure 5). DEG clusters (C1-10) and number of genes (in brackets). Magenta is high expression, cyan is low expression.



There were a number of genes that were significantly upregulated as the rhizomes entered dormancy, many of these genes were regulated by ABA, consistent with the data described above. More notably in dormant rhizomes, the transcriptomes were significantly different than at other times in the growing season. Genes that were differentially expressed in dormant rhizomes coded for a number of proteins that were not present at earlier stages of growth, and ones that appeared to be directly involved in rhizome metabolism. Combined with the data on metabolites, plant hormones and transcriptomes, our research appears to have unlocked many of the keys that underpin rhizome dormancy and winter-hardiness in switchgrass.

Objective 2: **Gene Profiling During Regreening and Dormancy of Bulked Segregants**

Perenniality can be expected to include mechanisms that control partitioning of assimilates between the shoots and crown, rhizomes and roots, as well as mechanisms that affect tiller initiation and dormancy. Internal and external factors that impair this seasonal exchange of metabolites between the above ground and below ground tissues will impact plant fitness, or more specifically winter survival. *Winter survival is therefore a crucial component of our ability to sustainably produce lignocellulosic feedstocks from switchgrass*. We had proposed that impairment of assimilate transport could arise from at least two different mechanisms. **Mechanism 1** involved photoperiod and temperature regulated traits including spring greenup, flowering date, and senescence all of which are associated with C and N cycling processes. Moving northern switchgrass ecotypes south gives them a shorter than normal photoperiod and they flower early and have reduced biomass yield. The opposite occurs when southern ecotypes are moved north and as a result they stay vegetative longer and produce more forage than northern strains moved south. The photoperiod response also appears to be responsible for susceptibility for winter kill. Southern types moved too far north will not survive. It has been assumed but not proven that this is because they stayed vegetative too long and failed to translocate enough carbohydrates and other nutrients to the roots, rhizomes and crowns to ensure winter survival (see Objective 1 results above). Selection of switchgrass for forage digestibility by ruminants has indicated that there are other cellular processes (**mechanism 2**) that affect plant fitness. **Mechanism 2** is apparent in switchgrass plants of populations that have been divergently selected for forage digestibility from a base population without associated selection for fitness traits. In these populations, genotypes with increased digestibility and associated significantly reduced lignin concentrations in biomass show lowered rates of winter-survival, although plant flowering is not affected. Generating these well-characterized genetically-related switchgrass plants through recurrent cycles of selection and breeding has required a significant investment in human and fiscal resources, but they now represent a valuable and unique resource for research on fitness in herbaceous perennials.

Plants were maintained in a clonal nursery and harvested at three times points during a growing season, corresponding to anthesis, physiological maturity and post-killing frost. Four plants each with high survival scores and low survivor scores were collected. In addition, 10 plants with mid survival scores were collected after a killing frost. In total 34 libraries were analyzed by RNA-Seq, yielding an average of 29 million reads per sample, of which an average of 79% mapped to the genome (v 1.1) using TopHat2. Over 65% of the total reads mapped to the annotated gene space in the genome. Analysis on this complex sample set is still in progress.

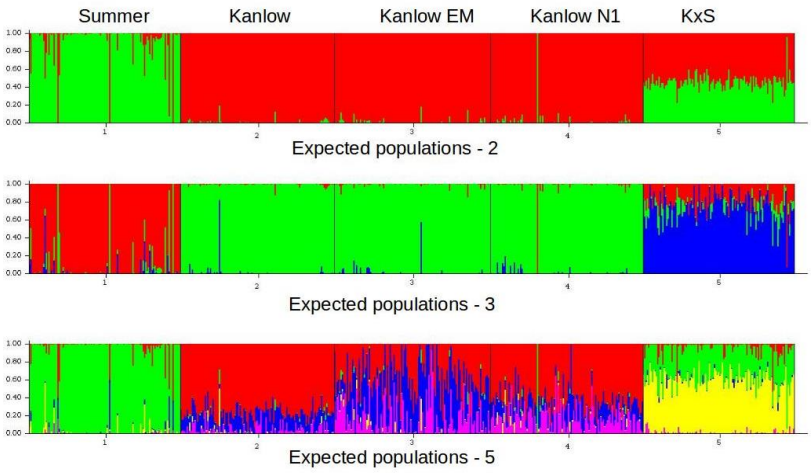
Objective 3: **Extent of Linkage Disequilibrium in Populations for Adaptation and Fitness Traits Being Developed for Central and Northern USA, that Show Significant Heterosis**.

Five switchgrass populations were genotyped and phenotyped. These populations consisted of the upland cultivar Summer, the lowland cultivar Kanlow, and selections from Kanlow selected for early maturity (Kanlow EM) and hardiness in USDA hardiness zone 5A (Kanlow N1). Another population was derived from the F3 of a Kanlow (male) to Summer (female) cross (KxS). These individuals were from the same seed lots used for C and N assimilation studies and were grown together in spaced plantings comprising a total of 592 individual clones with three clonal replications in a randomized complete block design. These plants were rated for heading date in 2012 and 2014, green-up day of year in 2013, anthesis date in 2012, and yield in 2012, they were also subjected to NIR spectroscopy to derive cell wall composition estimates based on prior NIR calibrations .

Plants were genotyped via a genotyping by sequencing (GBS) approach from reduced representation libraries constructed with adaptors that identified each individual. This approach is similar to that described previously. We sequenced libraries generated with the restriction enzyme *Pst*I and called SNPs using Samtools after alignment to version 1.1 of the switchgrass genome sequence. After excluding markers with greater than 20% missing data and those of low quality we obtained approximately 40,000 SNPs. These were then further filtered to eliminate markers with a minor allele frequency of < 0.05. Population structure was determined using the software STRUCTURE using the remaining marker.

The results of population analysis using STRUCTURE with expected population sizes or cluster numbers (K), clearly shows the hybrid composition of the KxS population and discriminated easily between upland (Summer) and lowland (Kanlow) populations (Figure 7). Under an assumption of 5 distinct populations there were detectable differences in allele frequencies between subpopulations within the three Kanlow populations particularly with respect to Kanlow EM and Kanlow base. We detected 110 SNPs with an allele frequency difference of ≥ 0.2 between Kanlow EM and Kanlow base populations, while 120 SNPs showed an allele frequency difference of ≥ 0.15 between Kanlow N1 and Kanlow base populations.

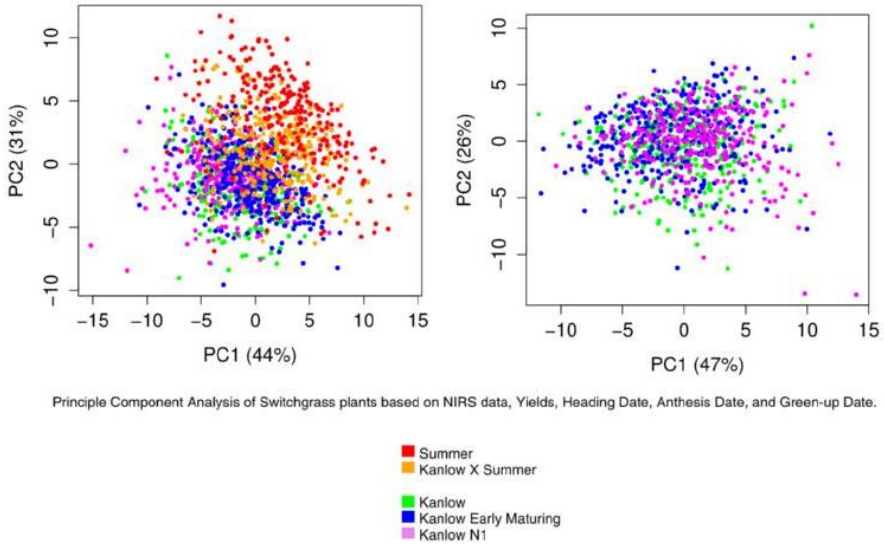
Figure 7: Results of STRUCTURE: Populations were labeled based on known breeding history and include Summer, Kanlow, Kanlow EM (early maturity), Kanlow NI (cold tolerance), and KxS.



PCA was performed on the phenotypic data and allowed separation of populations on principle coordinates 1 and 2 for Summer and KxS population (Figure 8). However, the Kanlow-based populations were not clearly separated along either principle coordinates. At this time we will be testing a mixed model

approach for association analysis in the presence of population structure to determine if there are any significant marker/trait associations. We are also looking at decay of LD with distance for the markers which have been positioned on the current switchgrass genome assembly.

Figure 8: PCA analysis of switchgrass plants based on NIRS data, yield, heading date, anthesis date, and green-up date.



We anticipate that in addition to the 11 publications listed on page 15, there will be at least 5 to 6 more publications that in the pipeline that acknowledge DOE grant support. These manuscripts will be submitted for peer-review in the coming months.

Grant Outcomes –Peer Reviewed Publications Acknowledging Grant Support

1. Saathoff AJ, Tobias CM, Sattler SE, Haas, EE, Twigg P, Sarath G (2011) Switchgrass contains two cinnamyl alcohol dehydrogenase involved in lignin formation. Bioener. Res 4: 120-133

2. Saathoff AJ, Sarath G, Chow E, Dien BS, Tobias CM (2011) Downregulation of cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. PLoS One. 6: e16416

3. Palmer NA, Saathoff AJ, Kim J, Benson A, Tobias CM, Twigg P, Vogel KP, Madhavan S, Sarath G (2011) A first analysis of a tetraploid switchgrass crown and rhizome transcriptome using next-generation sequencing. Bioener Res. 5: 649-661

4. Saathoff AJ, Hass E, Tobias CM, Twigg P, Sattler SE, Sarath G (2012) Switchgrass PviCAD1: Understanding residues important for substrate preferences and activity. Appl Biochem Biotechnol. 168: 1086-1100

5. Young HA, Sarath G, Tobias CM (2012) Karyotype variation is indicative of subgenomic and ecotypic differentiation in switchgrass. BMC Plant Biology. 12: 117

6. Saathoff AJ, Donze T, Palmer NA, Bradshaw J, Heng-Moss T, Twigg P, Tobias CM, Lagrimini M, Sarath G (2013) Towards uncovering the roles of switchgrass peroxidases in plant processes. Frontiers Plant Science. 4: 202.

7. Palmer NA, Saathoff AJ, Waters B, Donze T, Heng-Moss T, Twigg P, Tobias CM, Sarath G (2014) Global changes in mineral transporters in tetraploid switchgrasses (*Panicum virgatum* L.). Frontiers in Plant Science. 4:549

8. Sarath G, Baird LM, Mitchell RB (2014) Senescence, dormancy and tillering in perennial C_4 grasses. Invited review. Plant Science 217-218: 140-151

9. Palmer NA, Donze-Reiner T, Horvath D, Heng-Moss T, Waters B, Tobias CM, Sarath G (2014). Switchgrass (*Panicum virgatum* L) flag leaf transcriptomes reveal molecular signatures of leaf development, senescence, and mineral dynamics. Funct Integr Genomics. E-pub ahead of print.

10. Palmer NA*, Saathoff AJ*, Twigg P, Tobias CM, Vogel KP, Madhavan S, Sarath G (2014) Contrasting metabolism in perenniating structures of upland and lowland switchgrass plants late in the growing season. Plos One 9: e105138

11. A.J. Saathoff, M. Soundararajan, K.P. Vogel, C.M. Tobias, P. Twigg and G. Sarath (2014) A new Simple method for labeling field crops with stable isotope tracers. Australian Journal of Crop Science (in Press).

FULL LEGAL DISCLAIMER

TRADEMARK DISCLAIMER