Final Report
Grant Number: DE-FG02-07ER64550

Title: Genome-Wide Analysis of miRNA targets in *Brachypodium* and Biomass Energy Crops

PI: Pamela J. Green

Publications:

1. Jeong, D.H., Schmidt, S.A., Rymarquis, L.A., Park, S., Ganssmann, M., German, M.A., Accerbi, M., Zhai, J., Fahlgren, N., Fox, S.E., Garvin, D.F., Mockler, T.C., Carrington, J.C., Meyers, B.A., and Green, P.J. Parallel analysis of RNA ends enhances global investigation of microRNAs and target RNAs of Brachypodium distachyon. Genome Biology, 14:R145 (2013). PMCID: PMC4053937.

2. Jeong, D.H and Green, P.J. The role of rice microRNAs in abiotic stress responses. J. Plant Biol. 56:187-197 (2013).

3. Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. XRN 5' to 3' exoribonucleases: Structure, mechanisms and functions. Biochem. Biophys. Acta. - Gene Regulatory Mechanisms, Special Issue: RNA Decay mechanisms, Biochimica et Biophysica Acta 1829: 590–603 (2013). PMCID 3742305.

4. Franke, K.R. and Green, P.J. The miRNAs of *Brachypodium*. in Vogel J (Ed.) Genetics and Genomics of *Brachypodium.* Springer, New York, NY. In press.

Websites, database submissions, and other online resources:

https://mpss.udel.edu/dbs/index.php?SITE=brachy_sRNA2
https://mpss.udel.edu/dbs/index.php?SITE=brachy_pare2
National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO) under accession number GSE52441 (GSM1266839 to GSM1266857), GSM506620 and GSM506621
http://brachymirnas.appspot.com

Collaborations: This project fostered collaborations with the labs of Jim Carrington, Todd Mockler, David Garvin, and Blake Meyers who contributed to the miRNA component of publication 1. The Meyers lab set up the first two websites above to provide user-friendly public release of the data.

Research accomplishments:

The goal of this project was to elucidate miRNA targets in *Brachypodium distachyon* and extend that knowledge to bioenergy crops. The research had five objectives. Under this project 1) Parallel Analysis of RNA Ends (PARE) libraries and small RNA libraries were constructed from different tissues and following abiotic stress treatments of *Brachypodium* and bioenergy crops; 2) the libraries were deeply sequenced and matched to genome and cDNA sequences, 3) miRNA targets and new small RNAs that cause target cleavage were identified and evaluated for their functional associations, 4) tissue and stress-regulated changes in target cleavage were identified, and 5) data was released to the public through GEO and user-friendly websites as indicated above. This research is highlighted and aspects are being pursued further as described below.

PARE Enhances Global Investigation of miRNAs and Target RNAs of *Brachypodium distachyon*

Summary. *Brachypodium* has emerged as a model system for temperate grasses and biofuel plants. However, before this work, the global analysis of miRNAs, small RNAs known to be key for eukaryotic gene regulation, had been limited in *Brachypodium*. Prior studies examined only a few samples or relied on computational predictions. Similarly, an in-depth global analysis of miRNA-mediated target cleavage using PARE data was lacking in *Brachypodium*. To enhance knowledge and understanding of miRNAs and target RNAs of *Brachypodium*, 17 small RNA and four PARE libraries were constructed and deeply sequenced with Illumina technology. This, together with the sequencing of four PARE libraries, allowed for the identification of many new miRNAs and the characterization of their roles in target cleavage. The miRNAs identified include both conserved miRNAs that had not been reported in *Brachypodium*, as well as non-conserved miRNAs that were not found in other plants and thus may function in regulatory mechanisms specific to *Brachypodium* or related species. More than 80 new miRNA precursors of conserved and non-conserved miRNAs were uncovered. These provided new insights and confirmed several miRNA precursors that had been recently reported. Evidence that miR162-mediated DCL1 regulation may not function in *Brachypodium* was obtained. As discussed below, a new mode of miR5200 regulation in response to submergence stress was identified. More than 260 targets of new and known miRNAs with PARE sequences at the precise cleavage site were identified and their prominence characterized. Validated miRNA targets were found at all prominence levels and even low prominence was found to be characteristic of some targets. Combining PARE data with the small RNA data identified the miRNAs responsible for initiating hundreds of phased loci, including one of the novel miRNAs. PARE data also supported differential target cleavage in various tissues by miRNA variants.

Identification of *Brachypodium* miRNAs and miRNA precursors. The 17 small RNA libraries resulted in more than 75 million genome-matched small RNA sequences which was the largest such resource generated for *Brachypodium*. These were analyzed by a stringent criteria pipeline linked to a sequence homology pipeline and all miR candidates were visually analyzed to make final identifications. More than 80 new miRNA precursors of conserved and non-conserved miRNAs confirmed 11 conserved and 5 nonconserved precursors that had been recently examined. The miRNA from the new precursors combined with those previously identified, uncovered miRNAs with interesting features. These included miRNAs generated by alternative folding and stem-sharing and multiple cases of miRNA regulation which were conserved in other plants or novel.

miR5200 regulates gene expression of a flowering time regulator under submergence conditions. A submergence-inducible miR5200 was identified that is predicted to target two genes encoding FT (Flowering time locus T)-like proteins in *Brachypodium*. When miR5200 was induced during submergence, Bdi-FTL1 was found to be down-regulated. This is interesting because in both Arabidopsis and rice, flowering is thought to be inhibited by submergence Pena-Castro et al. (Plant J 67:434). Ectopic expression of the rice SUBMERGENCE 1A (SUB1A) gene in Arabidopsis led to flowering inhibition by reducing the level of FT and CONSTANS (CO) mRNAs. Further, ectopic over-expression or submergence-induced expression of SUB1A in rice affected the abundance of mRNAs encoding conserved flowering regulators of CO and FT. Based on work under this project, *Brachypodium* also may contain a regulatory mechanism where submergence represses flowering by down-regulation of Bdi-FTL1 and miR5200 may play a crucial role. Notably, miR5200 was not identified in either rice or Arabidopsis, yet, wheat and barley express miR5200 or related small RNAs at low levels. Thus,

in wheat and barley, miR5200 could be up-regulated under submergence and target FT-like genes, similar to *Brachypodium*.

Power of PARE to identify cleaved targets of *Brachypodium* miRNAs. The strategy to investigate the targets of *Brachypodium* miRNAs was to use PARE to deeply sequence RNA decay intermediates that have a 5′ monophosphate and a 3′ poly(A) tail. This is the structure of the downstream fragments produced from miRNA-guided cleavage. In Arabidopsis, PARE captured the vast majority of these cleavage products which had been previously validated in independent experiments. *Brachypodium* PARE libraries were constructed from RNA from leaves, stems, roots and flowers and sequenced to a total depth of nearly 70 million reads. This yielded a total of 5 million nonredundant genome-matched PARE sequences. These libraries represent a rich resource of PARE (also called RNA degradome) data which previously had been lacking for *Brachypodium*.

To identify miRNA target cleavage events, a set of predicted target cleavage sites for *Brachypodium* miRNAs were identified and the nonredundant PARE sequences were examined for those that precisely match these predictions. psRNATarget and CleaveLand, were used to predict a larger set of potential targets than either program alone. When these were matched to the PARE data, more than 260 of the sites had PARE sequences starting precisely at the predicted cleavage site in at least one of the four tissue libraries. This represents PARE sequences from the predicted cleavage sites for high proportion (81%) of the *Brachypodium* miRNAs examined. Among annotated miRNAs, 85% had PARE cleavage support and 70% did among the newly identified miRNAs discussed above.

A characterization tool was developed to assign a prominence level, from 1 to 4 with 4 being most prominent, to each of the aforementioned precisely positioned PARE sequences. The tool used a set of criteria established based on Arabidopsis miRNA targets with independent experimental validation, and was useful in multiple ways. First, it identified the targets validated by PARE which had the most prominent target cleavage data (Level 3 or 4 targets). Second, since experimentally confirmed targets are present at all levels, by knowing which specific criteria a particular Level 2 or 3 target has passed, potential explanations and future experiments can be identified. For instance, if the PARE sequence at the cleavage is of low abundance, this could be due to low accumulation of the miRNA or the target RNA in the samples examined, or because the miRNA largely inhibits translation and cleavage is less prominent. Moreover, a low prominence level may be a conserved characteristic of the target RNA in some cases. This is clearly evident for AGO1. Cleavage of AGO1 mRNA is guided by miR168; AGO1 is a Level 2 target in Arabidopsis and a Level 1 target in rice and *Brachypodium*. As PARE libraries are sequenced from diverse samples and closely related species, some of the *Brachypodium* targets of new or annotated miRNAs may be classified in higher levels or found to be characteristically Level 1. The published PARE data contributed by this work is extensive but will be augmented with additional *Brachypodium* PARE data from this project (see below) and probably others in the future.

Evidence that Bdi-miR5163b-3p-guided cleavage induces phased siRNAs from an NB-ARC mRNA. As reported in the genome paper, *Brachypodium* contains hundreds of 21- and 24-nucleotide phased small RNA (phasiRNA) loci in panicles (The International *Brachypodium* Initiative, Nature, 2010). Using the same method, more than 1,100 phasiRNA loci were identified from the key tissue small RNA libraries indicated above. PARE sequences were associated with 945 of these loci and initiation of almost half of could be accounted for by cleavage by well-known phasiRNA-initiating miRNAs (miR2118, miR2275, miR390). When other miRNAs of a size (22nt) that has the potential to initiate phased siRNAs were examined. the

PARE data indicated that Bdi-miR5163b-3p, from the list of newly identified miRNAs, targets one locus for phasing. Phased small RNAs were evident in all four tissues for this locus, Bradi4g10171, encoding an NB-ARC-containing R-protein. This protein is a member of the NB-LRR family of disease-resistance genes and their NB domains are known to be targeted for phasiRNA production in plants by several miRNAs including miR2118. Interestingly, Bdi-miR5163b-3p is not similar to these other miRNAs and it targets the LRR domain. Because this miRNA had not been reported in other plants, it may be a *Brachypodium*-specific miRNA that regulates NB-ARC genes and plays a role in disease resistance.

PARE identifies differential cleavage by miRNA family members. Many miRNA family members in plants have slight differences in sequences and differential expression patterns, which could affect their specificity for target regulation (Jeong et al., Plant Cell, 2011; Jeong and Green, Methods 2012; J. Plant. Biol. 2013). Under this project, both the miR166 and the miR156/529 miRNA families of *Brachypodium* were found to have members with sequence variations and differential regulation patterns. miR166f differs slightly in sequence compared to other miR166 family members including six which generate a miRNA of identical sequence, miR166a-eg. The latter was ubiquitously expressed except in mature leaf, whereas miR166f was preferentially expressed in roots. To address whether the differentially expressed miR166 family members also differentially regulate their target genes, cleavage events in a target mRNA for a homeobox domain-leucine zipper transcription factor were examined with PARE. Cleavage of this target guided by miR166f can easily be distinguished from cleavage guided by other miR166 family members because it is offset by 2 nt relative to the other family members. In root tissue, two different PARE sequences were mapped to the cleavage sites of both the root preferential miR166f and the remaining miR166 family members. However, from the panicle library miR166f-mediated target cleavage was detected at only a basal level and the one prominent cleavage was from the remaining family members. This indicates that miR166 family members may regulate HD-ZIP gene expression in a tissue-preferential manner.

Another miRNA family of interest was that encompassing miR156 and miR529 which differentially regulate the SPL14 (Squamosa Promoter Binding Protein 14- Like) gene and are thought to play a critical role in tillering and panicle branching in rice. This project showed that in *Brachypodium*, the expression of miR156 was high in young shoots and downregulated in mature leaves. In the panicle, miR156 levels were lower than in most other tissues in contrast to mi529 which was preferentially expressed in panicles. These expression patterns were similar to those in rice. The effects of miR156 and miR529 on SPL14L, a *Brachypodium* homolog of the rice SPL14 gene, were investigated using PARE. In leaf tissue, a prominent PARE sequence was mapped to the site of miR156-guided cleavage site whereas the PARE sequence from the site of miR529-guided cleavage was found at low abundance. This was in contrast to the situation in panicle in which the PARE sequence from the site of miR529-mediated cleavage was more abundant than that of miR156. The results imply that cleavage of BdiSPL14L during vegetative growth is predominantly miR156-guided, whereas miR529 and miR156 regulate the same gene during reproductive development. That both miRNA families were shown to be differentially expressed and to guide differential target cleavage, highlights the potential general significance of distinct miRNA family members for post-transcriptional control.

Advances in *Brachypodium* miRNA Target Analysis and Relevance to Bioenergy Crops

Following the work described above, we have completed the construction and sequencing of PARE libraries made from at least two biological replicates of *Brachypodium* seedlings subjected to a number of abiotic stress conditions (drought, high salt, heat, cold, submergence,

phosphate starvation, and control). The sequencing of these libraries generated over a billion reads with each library containing an average of over 5.9 million distinct genome-matched sequences. Prompted by the release of the *Brachypodium* v2.1 annotation along with miRBase 21, a new set of potential miRNA targets was generated using the psRNAtarget and Targetfinder algorithms. These target prediction programs yielded nearly 6400 potential miRNA target sites using the newly annotated transcripts and miRNAs as inputs. To search for transcripts undergoing miRNA guided post-transcriptional gene regulation during the abiotic stress responses, a pipeline was designed to analyze PARE libraries from control and stress conditions and identify sequences mapping to predicted target sites and which exhibited a two-fold change in abundance. This analysis yielded nearly 300 candidates from the cold, submergence, drought and heat libraries compared to controls. D-Plots were generated and following manual inspection, approximately half were selected as the most promising for further analysis. RNA-Seq libraries representing two biological replicates for each of the abiotic stress conditions were generated and sequenced. These data were used to characterize the changes in transcript levels of the targets in comparison to the changes observed in PARE sequences mapping to the target cleavage sites. This allows for the selection of miRNA target candidates which exhibit changes in PARE which cannot be explained by changes in transcription as well as provides insight as to whether the miRNA-guided cleavage is most likely a major or minor component of the target transcript's regulation. About 100 transcripts exhibited changes in transcript levels under stress conditions in the same direction ("Direct") as seen in PARE and thus could be explained by a transcriptional change. Approximately 35 transcripts had no statistically significant difference in transcript levels between control and stress conditions and were classified as "Unchanged." In these cases, the miRNAs may have fine tuning roles or may have a greater regulatory impact at an earlier or later stage of the stress response. Seven miRNA targets exhibited changes in RNA-Seq in the opposite direction as PARE and were classified as "Inverse." Cleavage guided by miRNAs is likely to have a stronger regulatory impact for these targets. Consistent with this assessment, a number of the miRNAs predicted to target transcripts classified as "Inverse" have previously been implicated in the plant stress response. Ongoing experiments will facilitate characterization of the levels of miRNA and cleaved target mRNAs at multiple points throughout the stress treatments.

This project also generated many PARE and small RNA libraries from sorghum and switchgrass tissues and stress treatments. These are being used to identify cases in which *Brachypodium* miRNAs and target cleavage events are found in these bioenergy crops. For example, the PARE data indicate conservation in switchgrass of differential cleavage of a SPL-like RNA guided by miR156 and miR529, similar to that seen in *Brachypodium.* Subsequent studies should indicate which stress associations of miRNA-guided cleavage are also conserved. Given that several miRNA and target RNAs showed potential involvement with multiple abiotic stress responses in *Brachypodium*, this project could reveal prime targets for creation of transgenic plants capable of withstanding harsher environmental conditions on multiple fronts. In addition, the PARE and small RNA data sets from this work will provide an excellent resource for future studies in *Brachypodium*, switchgrass and sorghum, not only for miRNA-mediated regulation, but for other types of post-transcriptional control as well.