

Project Title: Systems Level Regulation of Rhythmic Growth Rate and Biomass Accumulation in Grasses

Progress Report: 2012-2013

Principal Investigators: Samuel Hazen, Steve Kay, and John Mullet

Project Summary:

Critical to the development of renewable energy sources from biofuels is the improvement of biomass from energy feedstocks, such as sorghum and maize. To facilitate this improvement under the current DOE award, we initiated a series of experiments designed to identify and manipulate molecular networks involved with diurnal growth regulation in sorghum and the emerging monocot model, *Brachypodium distachyon*. The specific goals of this project include 1) characterize the growth and gene expression patterns under diurnal and circadian conditions, 2) select transcription factors associated with growth and build a cis-regulatory network in yeast, and 3) perturb these transcription factors *in planta* using transgenic *Brachypodium* and sorghum, and characterize the phenotypic outcomes as they relate to biomass accumulation. A better understanding of diurnally regulated growth behavior in grasses may lead to species-specific mechanisms highly relevant to future strategies to optimize energy crop biomass yield.

Recent Progress:

Our previous year's work has defined cyclic growth and biomass accumulation in *Brachypodium distachyon* to be strongly temperature dependent. Further, our results indicate a lack of regulation by the circadian clock on growth, contrary to observations in other models such as *Arabidopsis*[1,2]. These phenotypic descriptions have informed the experimental design of a major transcriptome study (outlined in Aim 1), which will quantify relative abundance of all mRNA species over one circadian (LDHC_LLHH) and two diurnal (LDHC_LDHC, LDHC_LLHC) time courses. For each time course, samples were harvested every 3.5 hours for 48 hours, and sequencing libraries were generated from polyadenylated mRNAs extracted from whole growing leaf tissue in 10 day old seedlings. Preliminary data from the first 12 samples in the LDHC_LDHC time course has been fully processed, generating 35-40 million paired end reads on a illumina HiSeq 2000 instrument. 75% of these reads successfully mapped as proper mate-pairs to the *Brachypodium* genome, and abundances for each gene were estimated and normalized using the edgeR and DESeq software packages[3,4]. Transcript abundance estimates were then matched against 12 cosine-wave functions, each having a period of 24 hours and a phase shift of 0, 2, ... , 22 hours relative to ZT0, using the HAYSTACK algorithm[5]. From this analysis (using stringent criteria in HAYSTACK), more than 2,000 diurnally regulated transcripts were identified, among them the clock genes, BdTOC1, BdLHY, BdGI, BdFKF1, and at least 70 other transcription factors (Fig. 1). Notably, many fewer cycling genes were identified as compared to *Arabidopsis*[6], however this may be a function of an incomplete repository of cycling models used to classify expression patterns. We will further mine this dataset to identify the full complement of cycling genes. The currently identified diurnally regulated genes were sorted into 12 different networks based on their time of peak expression. We then ran the ELEMENT[5] program through each of these networks to find cis-elements potentially enriched in genes which are co-expressed during the same phase. This analysis identified the known evening element, AAATATC[7], as well as the core ABRE (ACGTG) and some potentially novel cis-elements which warrant further investigation.

In the immediate future, we will perform similar analyses on the circadian and temperature-driven data sets, which are currently being processed at the core sequencing facility at USC. We hope to compare cycling gene sets with the aim of classifying them as circadian, temperature-regulated, and

light-regulated transcripts, based on their RNAseq profiles. As many enzymes display diurnal expression patterns, we will also conduct pathway analysis on genes associated with primary and secondary metabolism, which may identify major regulatory nodes gating biomass accumulation to certain temperatures. These genes, as well as transcription factors which show a strong temperature dependence will be given priority in downstream analysis. Once a prioritized list has been generated, we will leverage a yeast transcription factor library to identify components responsible for cyclic expression of target genes. We have applied this approach with a small library of transcription factors looking for regulators of secondary cell wall biosynthesis genes using both promoter fragments and bioinformatically-derived motifs as prey. Since the award period began, we have also initiated a collaboration with Todd Mockler's lab at the Donald Danforth Plant Science Center, who is cloning the full complement of ~2000 *Brachypodium* transcription factors in a yeast expression system suitable for yeast one- and two-hybrid screening. This resource will prove crucial to our construction of a transcription factor network regulating biomass accumulation (outlined in Aim 2). Concurrently, we will also continue to generate genetic tools to knock-down and overexpress transcription factors and enzymes showing expression profiles suggesting involvement in biomass accumulation and growth.

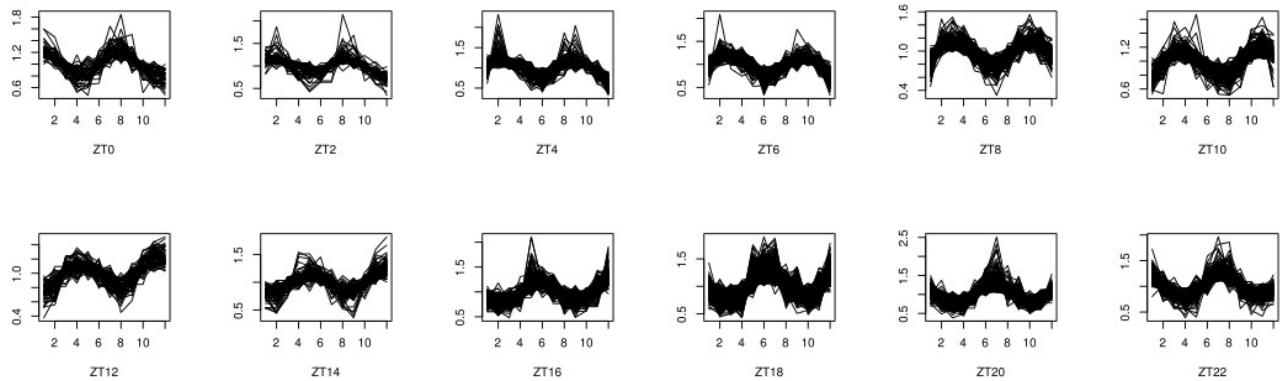


Figure 1: Diurnally regulated *Brachypodium* transcriptome. Normalized, Log₂-transformed RNAseq data from an LDHC time course (12 hours day at 28°C, 12 hours night at 12°C) were run through the HAYSTACK software using 12 cosine-wave models, a correlation cutoff of 0.7, and a fold-change cutoff (peak to trough) of 1.5. Cycling genes were separated into subnetworks based on phase of expression.

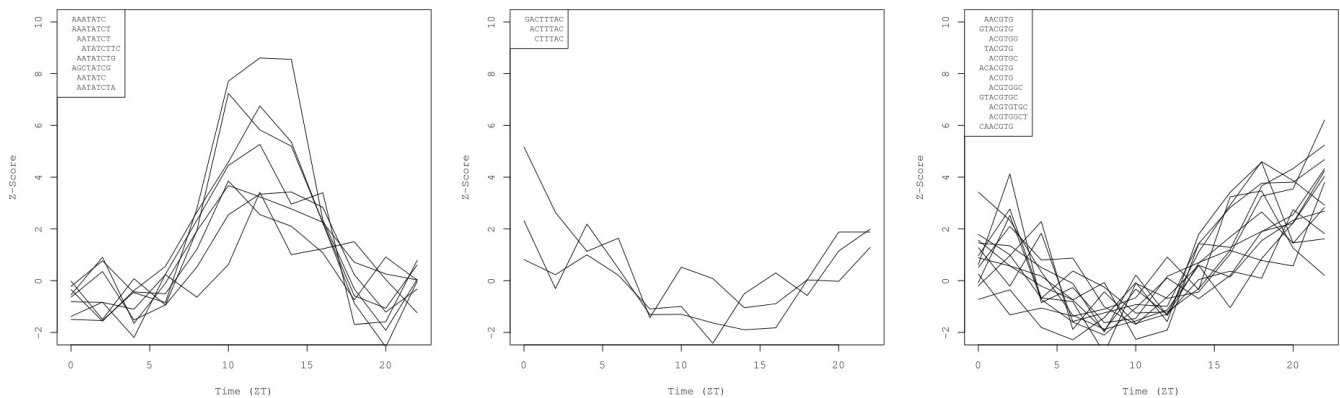


Figure 2: Cis-elements overrepresented in cyclic pattern. Promoters from cycling genes (Figure 1) were examined for cis-elements enriched in certain phase bins using the ELEMENT algorithm. Shown are variations of the evening element (Core: ATATC), an ABRE-Like element[8] (Core: ACGTG) and a novel CTTTAC motif which is enriched in morning-peaking genes).

1. Poire R, Wiese-Klinkenberg A, Parent B, Mielewczik M, Schurr U, et al. (2010) Diel time-courses of leaf growth in monocot and dicot species: endogenous rhythms and temperature effects. *J Exp Bot* 61: 1751–1759. doi:10.1093/jxb/erq049.
2. Nozue K, Covington MF, Duek PD, Lorrain S, Fankhauser C, et al. (2007) Rhythmic growth explained by coincidence between internal and external cues. *Nature* 448: 358–361. doi:10.1038/nature05946.
3. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11: R106. doi:10.1186/gb-2010-11-10-r106.
4. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma Oxf Engl* 26: 139–140. doi:10.1093/bioinformatics/btp616.
5. Mockler TC, Michael TP, Priest HD, Shen R, Sullivan CM, et al. (2007) The DIURNAL project: DIURNAL and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harb Symp Quant Biol* 72: 353–363. doi:10.1101/sqb.2007.72.006.
6. Michael TP, Mockler TC, Breton G, McEntee C, Byer A, et al. (2008) Network Discovery Pipeline Elucidates Conserved Time-of-Day–Specific cis-Regulatory Modules. *Plos Genet* 4: e14. doi:10.1371/journal.pgen.0040014.
7. Harmer SL, Hogenesch JB, Straume M, Chang H-S, Han B, et al. (2000) Orchestrated Transcription of Key Pathways in Arabidopsis by the Circadian Clock. *Science* 290: 2110–2113. doi:10.1126/science.290.5499.2110.
8. Hobo T, Asada M, Kowyama Y, Hattori T (1999) ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. *Plant J* 19: 679–689. doi:10.1046/j.1365-313x.1999.00565.x.