Identifying transcriptionally active regulatory DNase I hypersensitive sites in switchgrass (*Panicum virgatum* L.) genotypes AP13 and VS16

Ryne Davis[1], Mollee Crampton[2] and Venu (Kal) Kalavacharla[2,3]

[1]Department of Health and Science, Delaware Technical and Community College, Georgetown, DE 19947
[2]Molecular Genetics & EpiGenomics Laboratory, College of Agriculture Related Sciences, Delaware State University, Dover, DE, 19901
[3]Center for Integrated Biological and Environmental Research (CIBER), Delaware State University, Dover, DE, 19901

I.) Background

Switchgrass (*Panicum virgatum* L.) is a perennial C4 grass native to North America that has potential to be used as a viable renewable energy source capable of producing bioethanol through bioconversion utilizing simultaneous saccarification and fermentation (SSF) (Casler *et al*., 2011 and Shen *et al*., 2013). Bioethanol produced from cellulosic feedstocks like switchgrass can provide a clean energy source that can potentially reduce greenhouse gas emissions by up to 86% when compared with gasoline (Shen *et al*., 2013) . Congressionally mandated renewable fuel standards would require 35 billion gallons of ethanol-equivalent biofuels in addition to one-billion gallons of biodiesel by 2022, however, without any significant technological breakthroughs it is improbable that the United States will be able to meet these federal mandates (Shen *et al*., 2013). Research into switchgrass increased in 1992 when the USDOE chose it as the herbaceous model for the Biofuels Feedstock Development Program (BFDP). This species was selected for several reasons including: broad species adaptation, high biomass yields, tolerance to marginal land conditions (low fertility and drought), and simple seed processing and handling. Improvements in management and breeding have also produced significant increases in biomass yield and conversion efficiency (Casler *et al*., 2011). Switchgrass can be divided into two primary ecotypes: upland and lowland. The two genotypes that are used in this research are AP13 and VS16. AP13 is a high-yielding, tetraploid, lowland switchgrass clone, originating from the cultivar "Alamo". VS16 is an upland switchgrass clone originating from the cultivar "Summer" (Casler *et al*., 2011). DNase I Hypersensitive Sites (DHSs) consist of open-state chromatin regions that are detected by DNase I digestion (Madrigal and Krajewski, 2012). Disrupted regions of nucleosome are hypersensitive to DNase I and are correlated with gene activation in eukaryotes. DNase I can be applied in order to discriminatorily digest nucleosome-depleted DNA, presumptively near transcription factors (TFs), while regions of DNA that are tightly wound in nucleosomes and high-order structures are resistant to digestion. By mapping DHSs, different types of regulatory components can be identified, including: promoters, enhancers, silencers, insulators, and locus control regions. DNase I digestion and high-throughput sequencing (DNase-seq) provides a high-throughput method of identifying DHSs across a whole genome by taking DNase-digested fragments and sequencing them using high-throughput techniques. A single experiment utilizing DNase-seq is capable of identifying the majority of active regulatory regions from any type of cell from species that have genomes that have been sequenced (Song and Crawford, 2010).

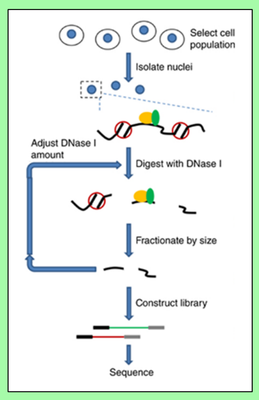
II.) Objectives

i) Use fragmented chromatin to prepare a DNase-seq library that can be sequenced and mapped back to reference genome
ii) Identify active site locations responsible for gene regulation
iii) Use DNase-seq data to help identify different active regulatory sites between genotypes AP13 and VS16

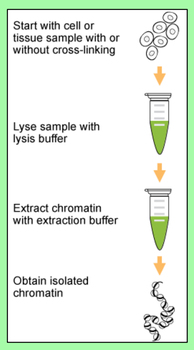
III.) Methods and Materials

DNase I digestion and high-throughput sequencing (DNase-seq)

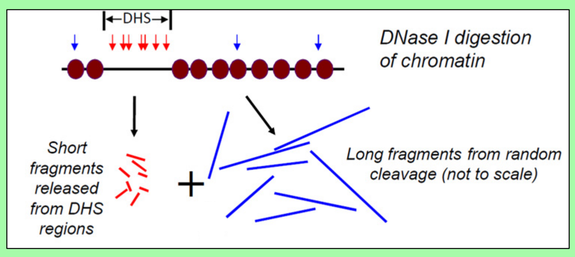
- Identify majority of active regulatory regions with a single experiment
- Identify DHS's across a whole genome by sequencing digested fragments



1) Nuclear Isolation
   1) Isolate chromatin from plant tissue
   2) Formaldehyde used for cross-linking to stabilize DNA-protein complexes



2) DNase Treatment
   - DNase I used as indicator of chromatin state ("open" or "closed")
   - DNase I cleavage sites more frequently occur in DHS regions
   - DNase I digestion releases short fragments from DHS regions and larger fragments that result from random cutting of genomic DNA
   - Short fragments are then separated from larger fragments using electrophoresis size selection on agarose gel



3) DNA Library Preparation
   - Size selected DNA purified 350-500bp
   - Prepare sequencing library following TruSeq® Nano DNA Sample Preparation HS protocol

4) Illumina HiSeq 2500 Sequencing
   - High-throughput sequencing system capable of genome-wide and ultra-deep sequencing
   - Capable of vast ranges of applications and study sizes

III.) Preliminary Results

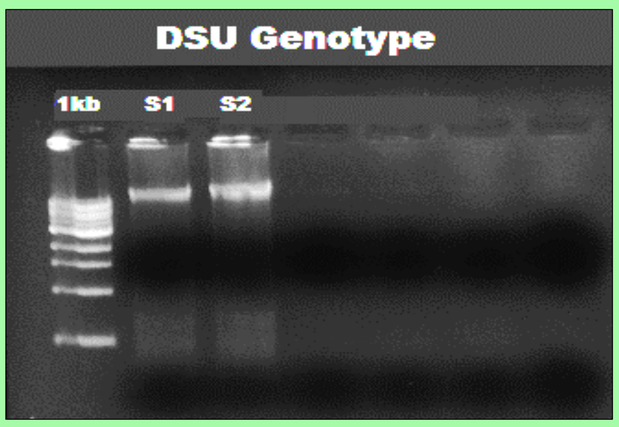


Figure 1: Isolated nuclear DNA from DSU switchgrass genotype samples (S1 and S2)

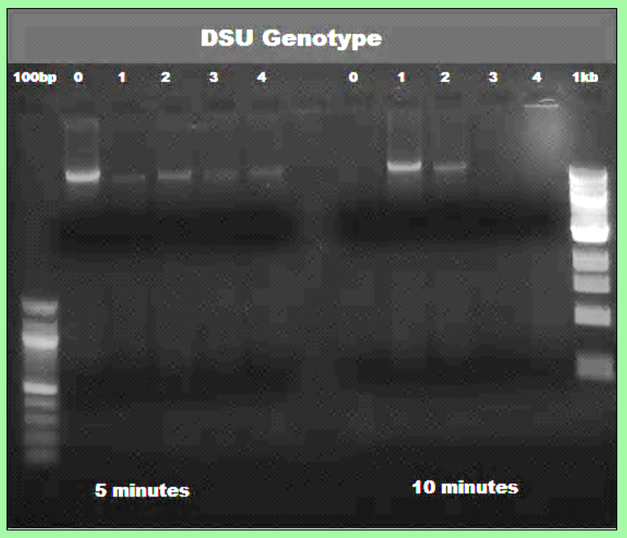


Figure 2: DNase I treated nuclear DNA from DSU switchgrass genotype sample 2. Sample was treated with DNase I at 5 and 10 minute intervals. Numbers given (0-4) indicate DNase volume in µl.

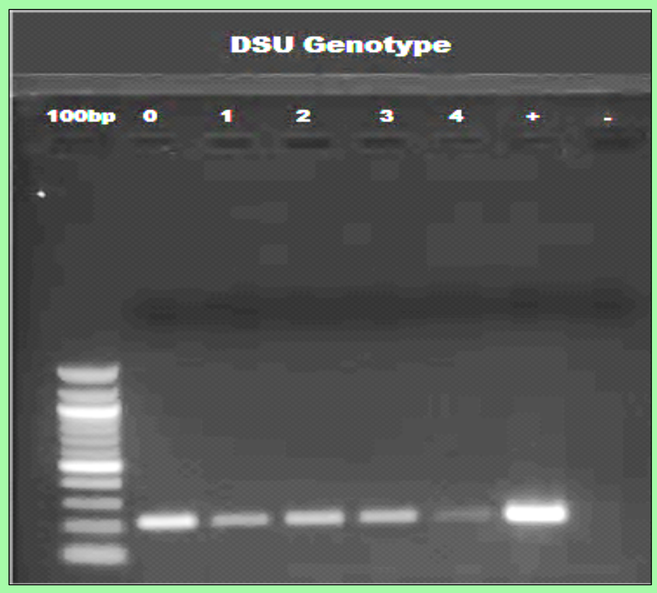


Figure 3: Positive amplification of DNase I treated nuclear DNA from sample 2 using AP13CTG11779 primers. Sample was DNase-treated for 10 minutes. Numbers given (0-4) indicate DNase volume in µl. + is positive control (AP13 gDNA), - is negative control (water).

IV.) Discussion and Future Work

i) DNase-seq should help identify active regulatory sites responsible for gene expression
ii) PCR results confirm that nuclear isolation and DNase treatment yielded quality fragments ready to undergo size selection and sequencing.
iii) Better understanding of the biological processes occurring at these sites and how gene expression differs between the two genotypes
iv) Future work may consist of a similar project, except the same genotypes will be analyzed after being exposed to abiotic stress factors (drought, salinity, etc.)
v) Compare results to other epigenetic studies (i.e. DNA methylation and ChIP-seq)

V.) References

Casler, M., Tobias, C., Kaeppler, S., Buell, R., Wang, Z., Cao, P., . . . Ronald, P. (2011). The Switchgrass Genome: Tools and Strategies. The Plant Genome, 4(3), 273-282. doi:10.3835/plantgenome2011.10.0026

Epigenetics.(n.d.). Retrieved June 12, 2015, from http://www.nature.com/subjects/epigenetics?WT.ac=search_subjects_epigenetics#research-and-reviews

Madrigal, P., and Krajewski, P. (2012). Current bioinformatic approaches to identify DNase I hypersensitive sites and genomic footprints from DNase-seq data. Frontiers in GENETICS, 3(230), 1-3. doi:10.3389/fgene.2012.00230

Shen, H., Poovaiah, C., Ziebell, A., Tschaplinski, T., Pattathil, S., Erica, G., . . . Richard, D. (2013). Enhanced characteristics of genetically modified switchgrass (Panicum virgatum L.) for high biofuel production. Biotechnology for Biofuels 2013, 6(71). doi:10.1186/1754-6834-6-71. Retrieved June 10, 2015.

Song, L., and Crawford, G. (2010). DNase-seq: A High-Resolution Technique for Mapping Active Gene Regulatory Elements across the Genome from Mammalian Cells. Cold Spring Harbor Laboratory Press, 2010(2), 1-11. doi:10.1101/pdb.prot5384

Google Images

VI.) Acknowledgments

I would like to acknowledge and thank the Northeast Woody/Warm-season Biomass Consortium (NEWBio), grant number: 2012-68005-19703 and Delaware EPSCoR, grant number: EPS-1301765, Delaware State University, the Department of Agriculture and Natural Resources, and everyone involved with the Molecular Genetics and EpiGenomics laboratory.