Title: Exploring Switchgrass Genetic Diversity with Multiple Reference Genomes

Authors: Thomas E. Juenger,[1] John T. Lovell[2,3]* (jlovell@hudsonalpha.org), Jane Grimwood,[3] and Jeremy Schmutz[2,3]

Institutions: [1]Department of Integrative Biology, University of Texas at Austin, Austin, TX, USA; [2]Genome Sequencing Center, HudsonAlpha Institute for Biotechnology, Huntsville, AL, USA; [3]Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Website URL: https://sites.cns.utexas.edu/juenger_lab/switchgrass

Project Goals: Overall, we are striving to improve bioenergy feedstock production by understanding the genetic basis of plant-environment interactions. This goal includes testing for climate adaptation, modeling beneficial and stressful biotic interactions, and exploring the mechanisms of abiotic stress responses. During our earlier work (e.g. [1-2]), we discovered a massive amount of physiological and molecular variation in switchgrass. While this diversity is the raw material that allows breeders to improve feedstock production, making use of this variation is very challenging — the immense DNA differences between some switchgrass genotypes means that traditional methods to explore genetic diversity simply do not work. Under the work presented here, we developed and employed multiple genome resources that span this diversity to provide the foundation for molecular characterization of switchgrass biomass production, stress responses and biotic interactions.

Abstract Text: A single haploid reference genome gives breeders the resources to connect alleles to traits; a significant step towards accelerating crop improvement. However, breeding programs often leverage highly diverged germplasm, which contain large-scale variants that are not readily identified by a single reference genome. For example, in switchgrass, the fast-growing southern lowland AP13 genotype (which serves as the reference genome [1]) is ~1 million years diverged from the cold-tolerant northern upland gene pool. To assist breeding and gene discovery efforts, we have developed four total reference genomes, two northern uplands and two southern lowlands. Each chromosome-scale genome is fully outbred representing both meiotically homologous chromosomes. Here, we present these genomes and a detailed comparison of gene presence absence, structural and copy number variation. These variants can serve as *a priori* targets for ongoing molecular breeding efforts to make switchgrass a more economically and ecologically sustainable biofuel feedstock.

References/Publications
1. Lovell, John T., Alice H. MacQueen, Sujan Mamidi, Jason Bonnette, Jerry Jenkins, Joseph D. Napier, Avinash Sreedasyam, et al. 2021. "Genomic Mechanisms of Climate Adaptation in Polyploid Bioenergy Switchgrass." *Nature* 590 (7846): 438–44.
2. Lowry, David B., John T. Lovell, Li Zhang, Jason Bonnette, Philip A. Fay, Robert B. Mitchell, John Lloyd-Reilley, et al. 2019. "QTL × Environment Interactions Underlie

Adaptive Divergence in Switchgrass across a Large Latitudinal Gradient." *Proceedings of the National Academy of Sciences of the United States of America* 116 (26): 12933–41.