

***Miscanthus sinensis* Genome: Updates on the Genome Assembly and Analysis of the Transcriptome and Small RNA**

Mohammad B. Belaffif^{1*} (mbelaffif@hudsonalpha.org), Brandon T. James,¹ Adam Session,² Therese Mitros,³ Stephen P. Moose,⁴ Daniel S. Rokhsar,³ and **Kankshita Swaminathan**¹

¹HudsonAlpha Institute for Biotechnology, Huntsville, AL; ²Joint Genome Institute, Walnut Creek, CA; ³University of California, Berkeley; and ⁴University of Illinois at Urbana-Champaign, Urbana

Project Goals:

Generated a reference genome for *Miscanthus sinensis* to enable comparative genomics in the *Andropogoneae* grasses.

- Chromosome-scale assembly of the *M. sinensis* genome
- Characterization of subgenome-specific signatures that reflect recent allotetraploidy
- Annotation of both gene and small RNA producing regions to lay the foundation for functional genomics and biotechnology

Understanding key genes and networks regulating stem carbon partitioning, nutrient remobilization, and perenniality

- Identify genes with expression patterns that can be used to engineer pathways in a tissue-specific manner
- Understanding the spatiotemporal dynamics of gene regulatory networks governing carbon and nitrogen partitioning

Miscanthus is a perennial C4 grass which exhibits a number of desirable traits for sustainable biomass production, especially in temperate regions. The availability of a high quality, well annotated, reference genome for *Miscanthus* will provide a foothold for understanding the genetic basis of traits of interest and facilitate genomics enabled breeding. Here we present an update of our progress in providing a chromosome scale genome assembly for *Miscanthus sinensis* (available on Phytozome https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Msinensis_er).

The draft genome presents an almost complete coverage of the 19 chromosomes of the *M. sinensis* genome, with 91% of assembled sequence placed on chromosomes. We have identified LTR retrotransposons that allow us to identify an allotetraploidy within the *Miscanthus* clade and mark the two subgenomes. Our current annotation predicts a total of 67,789 gene models, representing gene coding regions, 82% of which have transcriptional evidence. To better understand the developmental biology *Miscanthus* and mine genes of interest, we have generated both RNA-seq and small RNA-seq data from *Miscanthus x giganteus*, a natural derived triploid of *M. sinensis* and *M. sacchariflorus*. We used the transcriptomic data to annotate *M. sinensis*

genome. By leveraging the high-quality genome of *M. sinensis*, we were able to identify genes preferentially expressed in leaves, stems or rhizomes and are in the process of obtaining a more detailed spatiotemporal expression for these genes using promoter-reporter fusions.

Small RNAs play an important role in plant growth and development and are implicated in important phenomena like fine regulation of gene function and heterosis. We used small RNA sequencing to characterize the types of small RNA and regions producing them. We have categorized the small RNA producing regions into six types: microRNA, putative microRNA, natsiRNA, putative natsiRNA, putative cncRNA, and phasiRNA. We have identified 430 high confidence microRNA or putative microRNA producing regions. Most of these are conserved between the two subgenomes of *M. sinensis*. We are in the process of combining the expression data from both the transcriptomic and small RNA to build a composite gene regulatory network to better understand gene regulation and their role in development in *Miscanthus*.

This work was funded by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018420). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Energy.