

Transcriptomic and Metagenomic Exploration of Agave Drought and Stress Tolerance



Stephen Gross^{1,2}, Jeffrey Martin^{1,2}, June Simpson³, Laila Partida-Martínez³, Gretchen North⁴, Yuri Peña-Ramírez⁵, Aidee Orozco-Hernandez⁵, Kristen DeAngelis⁶, Tanja Woyke^{1,2}, Scott Clingenpeel^{1,2}, Susannah Tringe^{1,2}, Zhong Wang^{1,2}, and Axel Visel^{1,2}

¹ Department of Energy Joint Genome Institute, Walnut Creek, CA
² Genomics Division, Lawrence Berkeley National Lab, Berkeley, CA
³ CINVESTAV, Irapuato, Guanajuato, Mexico
⁴ Occidental College, Los Angeles, CA
⁵ Brown-Forman, Inc., Amantlan, Jalisco, Mexico
⁶ University of Massachusetts, Amherst, MA



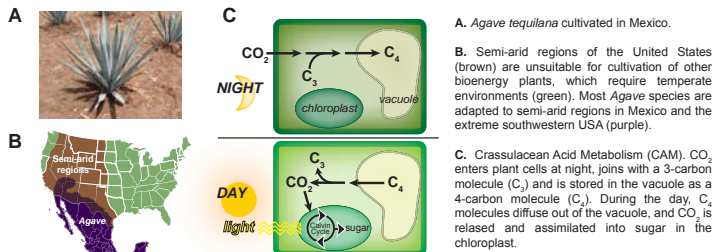
Abstract

Understanding drought and heat tolerance in plants will be critical for sustained agriculture and bioenergy production in a changing global climate. Because of their exceptional ability to thrive in arid, hot environments with minimal soil nitrogen, *Agave* species have been identified as a candidate bioenergy feedstock and provide a model for studies of adaptations to drought and heat. While some physiological mechanisms underlying these traits have been studied in detail, the molecular basis of the extreme heat and drought tolerance in these plants remains unclear. We have constructed *de novo* reference transcriptomes for *Agave tequiliana*, an economically important species cultivated in Mexico for spirit distillation, and *A. deserti*, an extremely thermo- and drought-tolerant species native to the Colorado Desert, using 368 and 185 gigabasepairs, respectively, of Illumina short-read sequence data. These data sets enable cross-species comparative analysis of protein-coding sequences and tissue-specific gene expression levels. Based on these reference transcriptomes, we also perform quantitative expression profiling of agaves subjected to heat and drought stress in controlled greenhouse experiments, enabling a transcriptome-wide understanding of *Agave* stress responses. Complementary to plant-endogenous mechanisms, *Agave* plants are associated with microbes that may play important roles in mitigating abiotic stress in arid resource-limited environments. Using targeted sequence-based microbial community profiling, we study the microbiomes of cultivated *A. tequiliana*, and wild *A. deserti* and *A. salmiana*. We also use deep metagenomic sequencing to gain further insight into *Agave* rhizosphere and phyllosphere microbial communities, and perform isolation and single-cell genome sequencing of individual microbial cells residing within *Agave* tissues. These studies aim to identify microbial species, genes, and pathways conferring additional stress resistance to agaves. Taken together, our work builds a robust platform to accelerate discovery of plant and plant-associated microbial adaptations to major abiotic stresses.

Overview

I. Agave can supplement other bioenergy feedstocks

Agave species, adapted to their native habitat in arid regions of Mexico and the United States, hold promise as a biofuel feedstock [1], capable of growing on marginal lands where other bioenergy plants cannot. The ability of agaves to withstand hot and arid conditions relies upon Crassulacean Acid Metabolism (CAM)—a specialized form of photosynthesis allowing agaves to keep leaf stomata (pores) closed during the day, minimizing water loss through evapotranspiration.



II. Agaves are productive with minimal resources

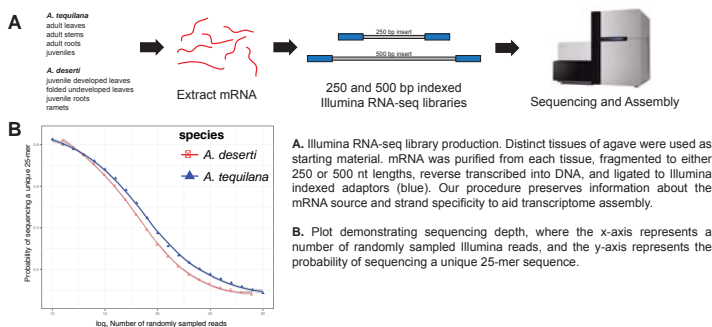
Agaves are capable of producing lignocellulosic biomass with little water and nitrogen inputs. Species *A. salmiana* and *A. mapisaga* have been reported to produce up to 40 metric tonnes (Mg) of dry biomass per hectare per year [2].

Feedstock	Inputs		Outputs		Comparison of inputs (water and nitrogen) and outputs (biomass and ethanol) of agaves and other biofuel feedstock species. Though agaves are harvested at several years of age, their annualized growth rate is comparable to <i>Miscanthus</i> . Table is modified from reference [4].
	Water (cm yr ⁻¹)	Drought tolerance	Nitrogen (kg ha ⁻¹ yr ⁻¹)	Dry biomass (Mg ha ⁻¹ yr ⁻¹)	
Corn grain	50–80	low	90–120	7–10	2900
Corn stover				3–6	900
<i>Miscanthus</i>	75–120	low	0–15	15–40	4600–12,400
Poplar coppice	70–105	moderate	0–50	5–11	1500–3400
Agave spp.	30–80	high	0–12	10–34	3000–10,500

Building Agave reference transcriptomes

De novo assembly of Agave transcriptomes

Without sequence information, molecular studies of agaves are difficult. To address this need, we chose *A. tequiliana*, which is currently cultivated for tequila production, and *A. deserti*, an extremely drought and heat tolerant agave, as our reference species. Agaves have large genomes (~4–7 Gb) [5, 6], so we focused on sequencing the protein-coding transcriptome using Illumina RNA-seq and *de novo* transcriptome assembly with Rnnotator [7], a *de novo* transcriptome assembly pipeline developed at JGI.

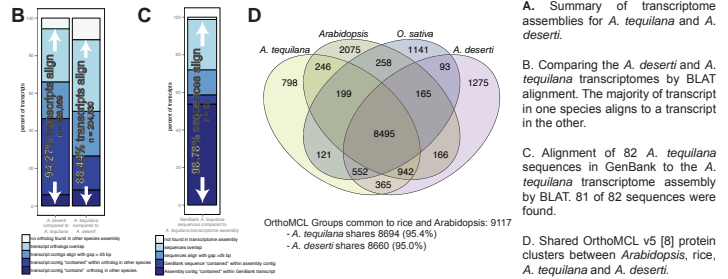


Robust transcriptome assemblies

A resource for Agave molecular biology

Transcriptome assemblies were filtered for artifacts and sequences from associated microorganisms. The resulting *Agave* transcriptomes, while not a 100% complete representation of *Agave* genes, contain a significant number of loci and encode the majority of proteins within other plant species.

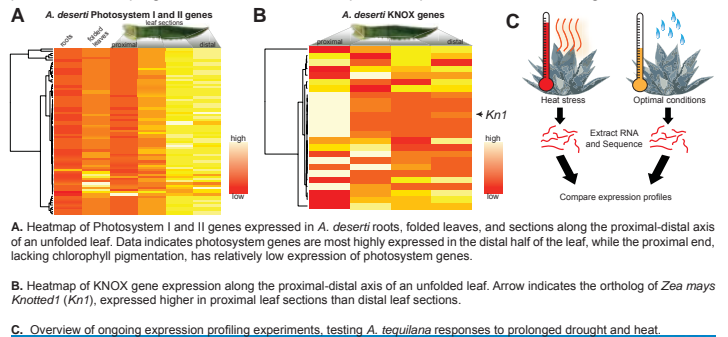
Species/dataset	Raw sequence	No. of loci	No. of transcripts	N50 transcript length	Assembled transcriptome size
<i>A. tequiliana</i>	368 Gbp	139,525	204,530	1387 nt	204.8 Mbp
<i>A. deserti</i>	185 Gbp	88,718	128,869	1323 nt	125.0 Mbp



Expression profiling of agaves

Adding functional data to annotation

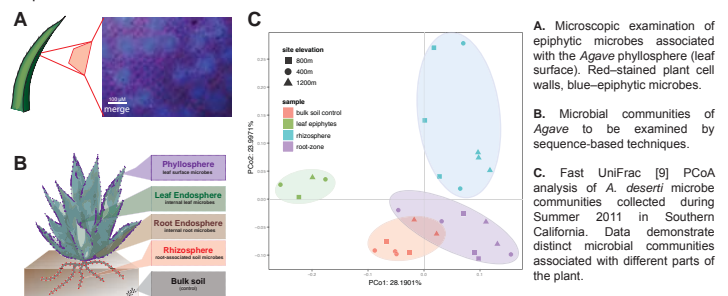
With reference transcriptomes in place, expression profiling experiments can be initiated. Using data from the transcriptome assemblies, tissue-specific expression profiles can be studied. Additional experiments are in progress to understand transcriptome responses to heat and drought stress.



Agave microbiomes and adaptations to stress

Discovering the Agave microbial community

We have initiated sequence-based studies of the *A. tequiliana*, *A. salmiana* and *A. deserti* microbiomes. Ultimately, we aim to identify microbes conferring stress and disease resistance to agaves. With both transcriptome and microbe sequences in hand, we will have a strong foundation for powerful plant-microbe interaction studies.



Acknowledgements and References

This work performed at the U.S. Department of Energy Joint Genome Institute was supported in part by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH112. *Agave* microbiome work is supported in part by the 2012 DOE JGI Community Sequencing Program.

[1] Davis, A. S. *et al.* The global potential for *Agave* as a biofuel feedstock. *GCB Bioenergy* 3, 68–78, (2011).
 [2] Nobel, P. S. *et al.* High annual productivity of certain agaves and cacti under cultivation. *Plant Cell Environ* 15, 329–35, (1992).
 [3] Simpson, J. *et al.* Genomic resources and transcriptome mining in *Agave tequiliana*. *GCB Bioenergy* 3, 25–36, (2011).
 [4] Somerville, C. *et al.* Feedstocks for lignocellulosic biofuels. *Science* 329, 790–2, (2010).
 [5] Palomino, G. *et al.* Nuclear genome size analysis of *Agave tequiliana* Weber. *Caryologia* 56, 37–46, (2003).
 [6] Robert, M. L. *et al.* Wild and agriculturally important *Agave* species (Asparagaceae) show proportional increases in chromosome number, genome size, and genetic markers with increasing ploidy. *Bot J Linn Soc* 158, 215–22, (2008).
 [7] Martin, J. *et al.* Rnnotator: an automated *de novo* transcriptome assembly pipeline from stranded RNA-seq reads. *BMC Genomics* 11, 663, (2010).
 [8] Chen, F. *et al.* OrthoMCL-DB: querying a comprehensive multi-species collection of orthologs. *Nucleic Acids Res* 34, D363–8, (2006).
 [9] Hamady, M. *et al.* Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J* 4, 17–27, (2010).