

Agave transcriptomes and microbiomes for bioenergy research

Stephen Gross ^{1,2}, Jeffrey Martin ^{1,2}, June Simpson ³, Laila Partida-Martínez ³, Gretchen North ⁴, Yuri Peña-Ramirez ⁵, Aidee Orozco-Hernandez ⁵, Kristen DeAngelis ⁶, 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., Amatitán, Jalisco, Mexico

6 University of Massachusetts, Amherst, MA

^aTo whom correspondence may be addressed. E-mail: smgross@lbl.gov

March 12, 2012

ACKNOWLEDGMENTS:

The work conducted by the US Department of Energy (DOE) Joint Genome Institute is supported by the Office of Science of the DOE under Contract Number DE-AC02-05CH11231. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government, or any agency thereof, or the Regents of the University of California.

DISCLAIMER:

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California

Agave transcriptomes and microbiomes for bioenergy research



Stephen Gross ^{1,2}, Jeffrey Martin ^{1,2}, June Simpson ³, Laila Partida-Martínez ³, Gretchen North ⁴, Yuri Peña-Ramírez ⁵, Aidee Orozco-Hernandez ⁵, Kristen DeAngelis ⁶, 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., Amatián, Jalisco, Mexico
⁶ University of Massachusetts, Amherst, MA



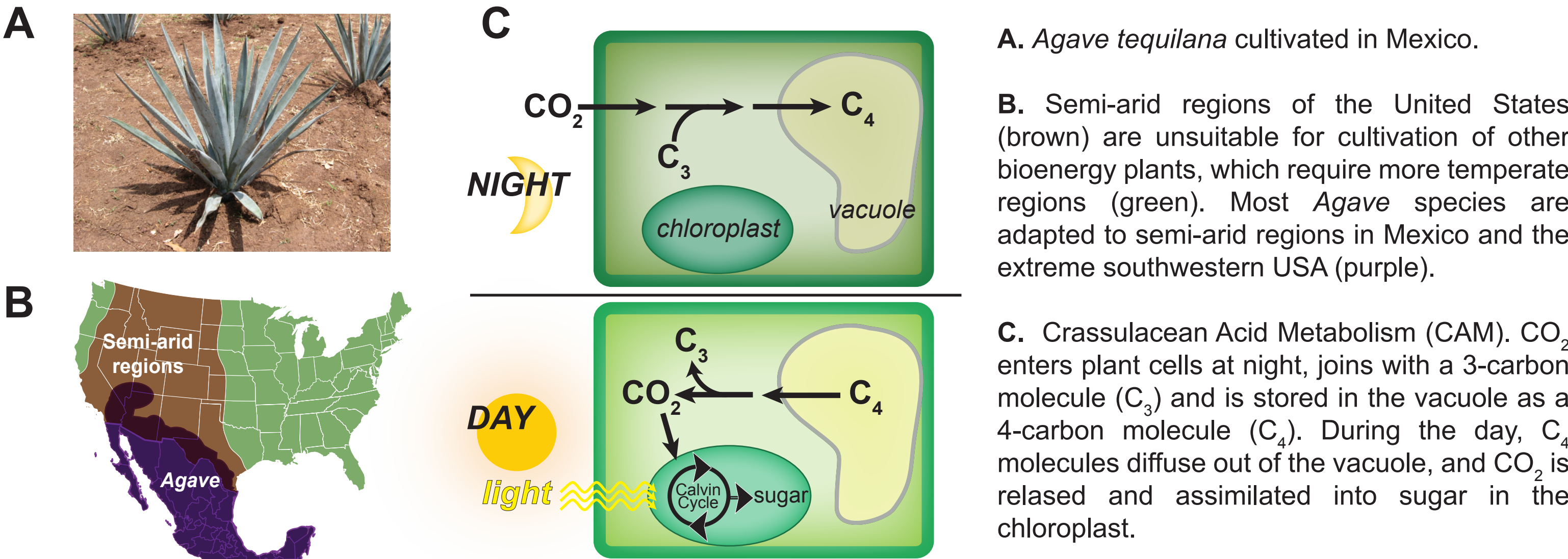
Abstract

Development of new lignocellulosic bioenergy feedstocks minimizing impacts on staple food production and withstanding abiotic stresses anticipated from climate change is key to building a future with sustainable liquid transportation fuels. Because of their exceptional ability to thrive on nutrient-poor soils in arid, hot environments with minimal water and nitrogen requirements, *Agave* species have recently been proposed as an additional bioenergy feedstock. While the physiological mechanisms enabling agaves to survive in their native arid environments are understood, a paucity of sequence information prohibits powerful sequence-based analyses of *Agave* adaptations to abiotic stress. Additionally, as microbes associated with agaves remain unstudied, the role microbe communities may have in augmenting stress resistance is unclear. To address these issues, we are constructing *de novo* reference transcriptomes for *Agave tequilana*, 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 the reference transcriptomes, we plan to explore gene regulatory and physiological responses to drought and heat in controlled greenhouse experiments. In parallel, we are investigating the microbiomes of cultivated *A. tequilana*, and wild *A. deserti* and *A. salmiana* using sequence-based microbial community profiling techniques. Select microbiomes will be chosen for deep metagenome sequencing in order to understand microbial genes and pathways conferring additional stress resistance to agaves. Endophytic microbes living within agave tissues will be targeted for single-cell genome sequencing. Taken together, our work builds a robust platform to accelerate discovery of plant adaptations to abiotic stress and further development of *Agave* species as a bioenergy feedstocks.

Overview

I. Agave can supplement other bioenergy feedstocks

Agave species are adapted to their native habitat in arid regions of Mexico and the United States. *Agave* thus holds promise as a biofuel feedstock [1], capable of growing on marginal lands where other proposed 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 hot day, minimizing water loss through evapotranspiration.



II. Agaves are productive with minimal resources

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

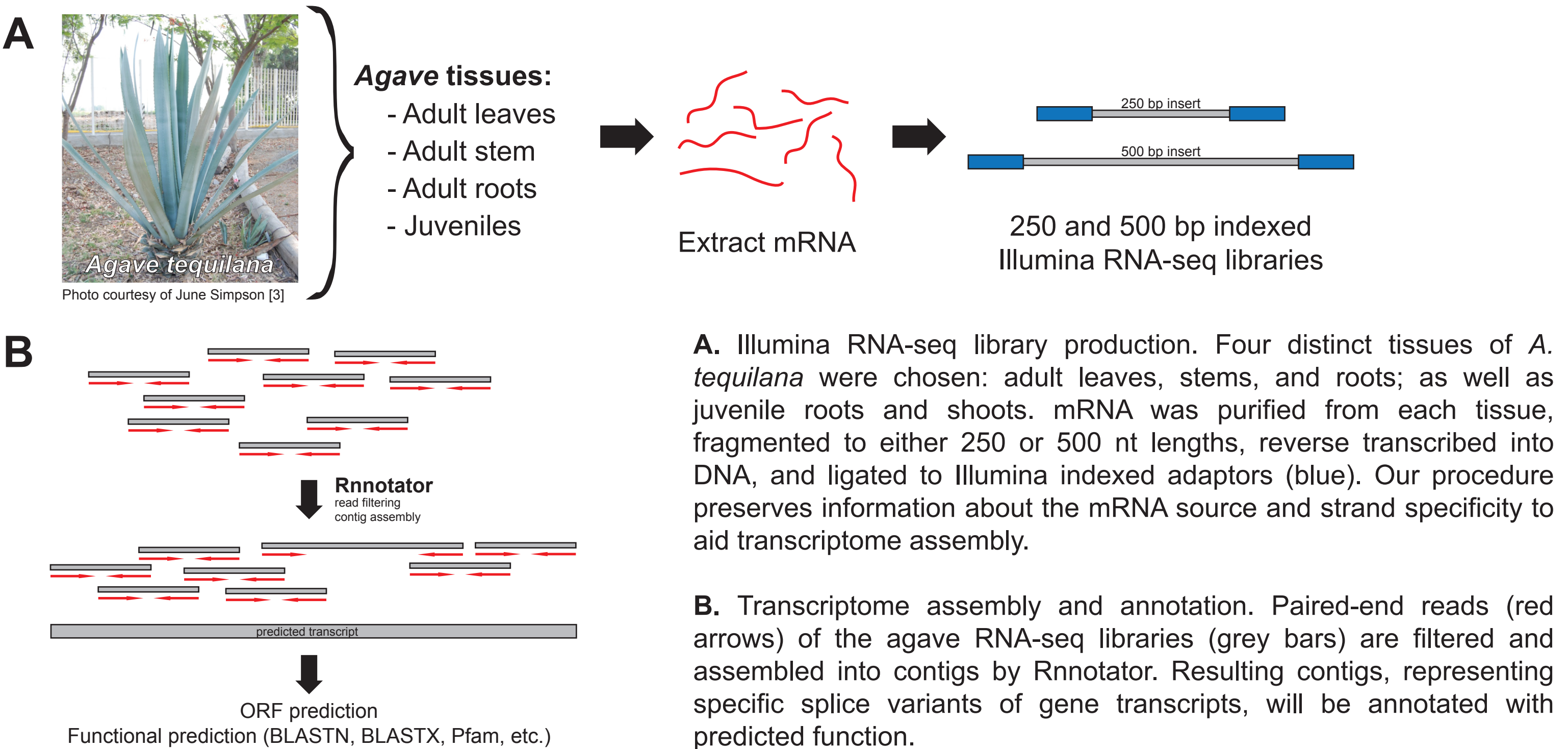
Feedstock	Inputs			Outputs	
	Water (cm yr ⁻¹)	Drought tolerance	Nitrogen (kg ha ⁻¹ yr ⁻¹)	Dry biomass (Mg ha ⁻¹ yr ⁻¹)	Ethanol (liters yr ⁻¹)
Corn grain	50–80	low	90–120	7–10	2900
Corn stover				3–6	900
<i>Miscanthus</i>	75–120	low	0–15	15–40	4600–12,400
Poplar coppice	70–105	moderate	0–50	5–11	1500–3400
<i>Agave spp.</i>	30–80	high	0–12	10–34	3000–10,500

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 on par with *Miscanthus*. Table is modified from reference [4].

Building sequence resources for Agave

De novo assembly of Agave transcriptomes

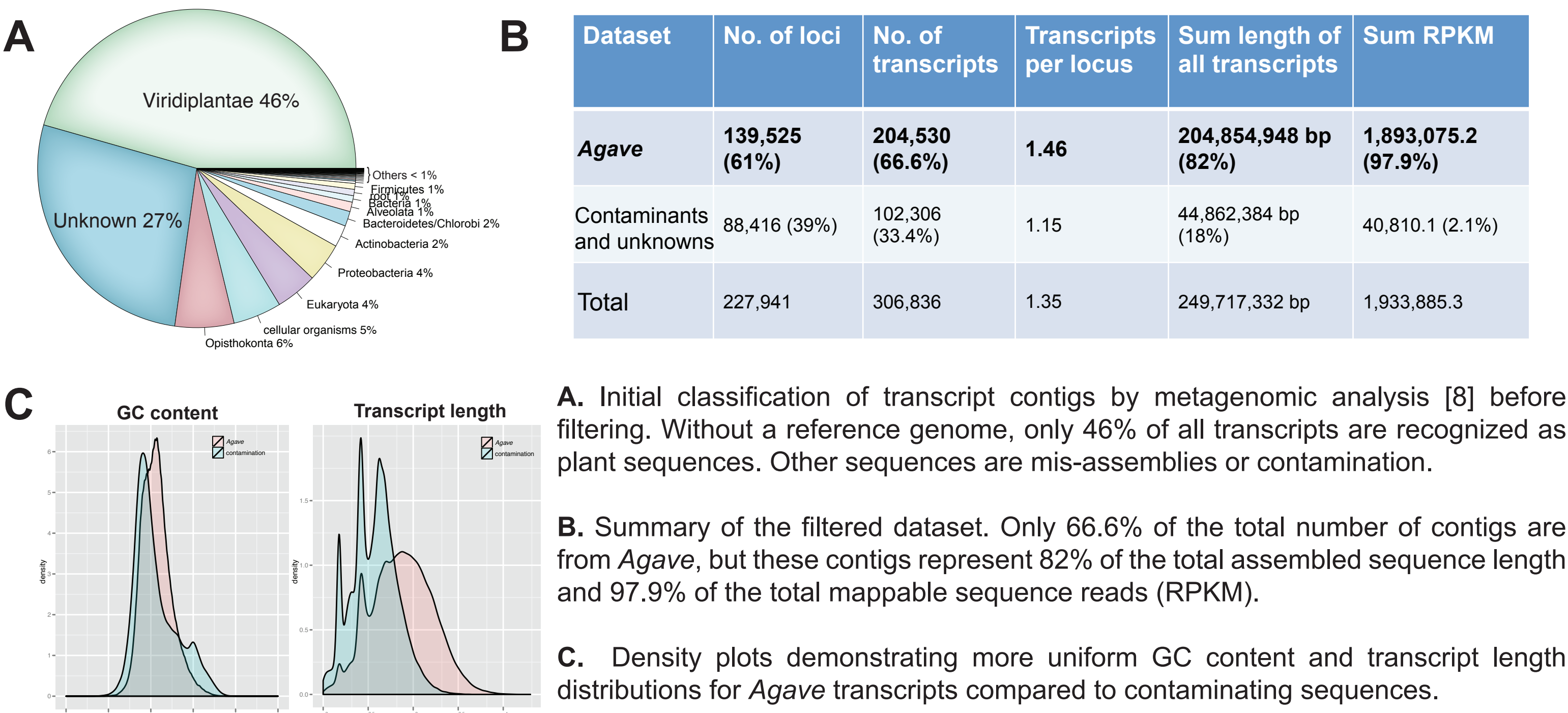
Without sequence information, molecular studies of agaves are difficult. To address this need, we chose *A. tequilana*, which is currently cultivated for tequila production, as our reference species. As agaves have large genomes (~4–7 Gb) [5, 6], we sequenced the protein-coding transcriptome using Illumina RNA-seq technologies. A new assembly pipeline developed at JGI, Rnnotator [7], was used to assemble RNA-seq data for *de novo* transcriptome assembly. The transcriptome of *Agave deserti* is currently being sequenced using similar techniques.



Filtering the Agave tequilana transcriptome

Removing contaminating sequences

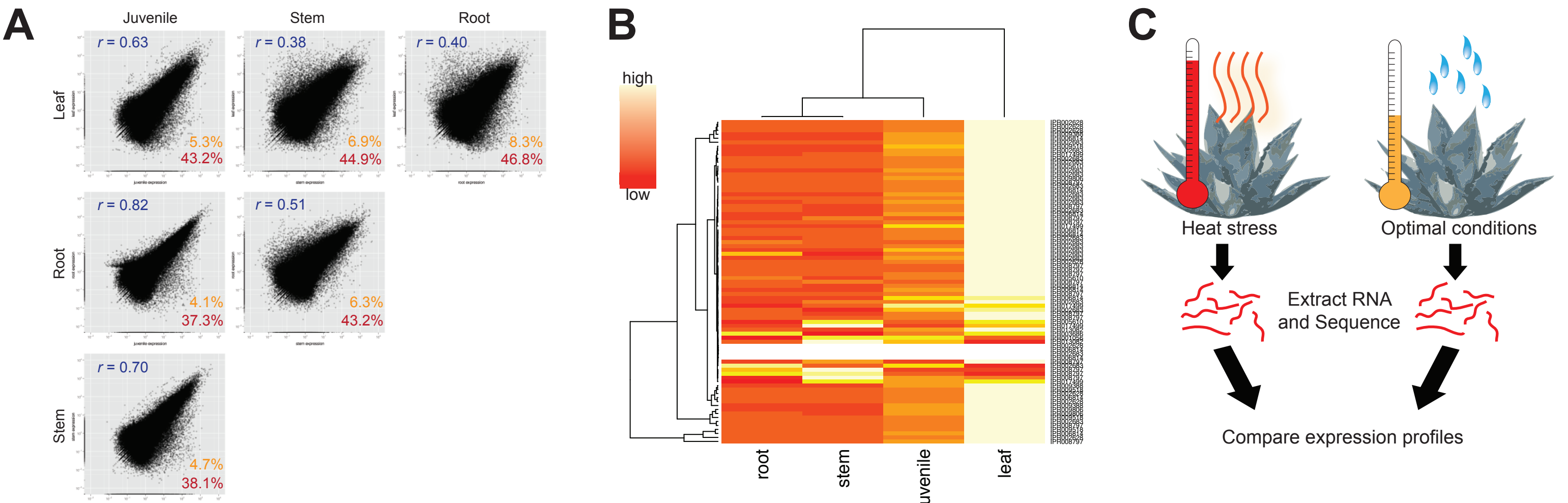
Deep transcriptome sequencing can assemble sequences from microorganisms associated with *Agave*. Without a reference genome, identifying contaminating sequences can be difficult. Using a combination of metagenomics tools [8] and expression data, transcripts were binned into either an “*Agave*” dataset or “contamination” dataset.



Expression profiling of A. tequilana

Adding functional data to annotation

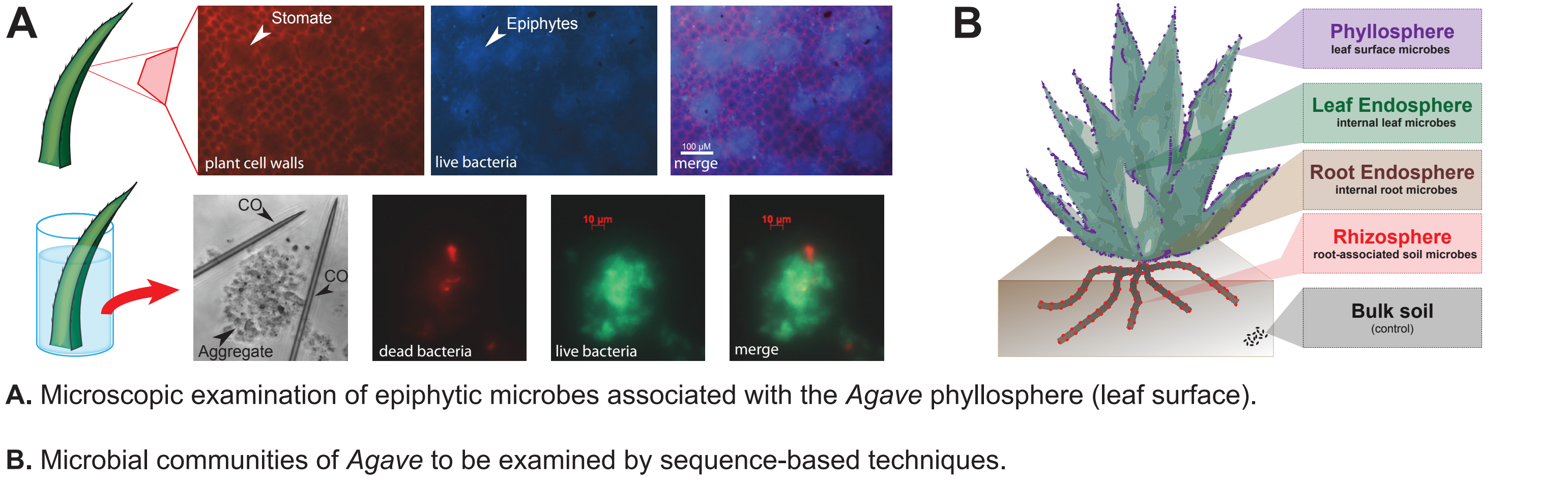
With a reference transcriptome in place, expression profiling experiments can be initiated. Using data from the *A. tequilana* transcriptome, tissue-specific expression profiles can be studied. Additional experiments are in progress to understand transcriptome responses to heat and drought stress.



Agave microbiomes and adaptations to stress

Discovering the Agave microbial community

With support from the JGI Community Sequencing Program (CSP), we have initiated studies of the *A. tequilana*, *A. salmiana* and *A. deserti* microbiomes. Using sequencing-based approaches, we aim to understand the community of microbes associated with agaves both in the wild and under cultivation. 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 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 tequilana*. *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 tequilana* Weber. *Caryologia* 56, 37–46, (2003).
[6] Robert, M. L. *et al.* Wild and agronomically 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] Huson, D. H. *et al.* MEGAN analysis of metagenomic data. *Genome Research* 17, 377–386, (2007).