Transcriptome analysis of three *Agave* fiber-producing cultivars suitable for biochemicals and biofuels production in semiarid regions

Fabio Trigo Raya1; Marina Pupke Marone1; Lucas Miguel Carvalho1; Sarita Candida Rabelo2; Maiki Soares de Paula1; Maria Fernanda Zaneli Campanari1; Luciano Freschi3; Juliana Lischka Sampaio Mayer4; Odilon Reny Ribeiro Ferreira Silva5; Piotr Mieczkowski6; Marcelo Falsarella Carazzolle1,7; Gonçalo Amarante Guimarães Pereira1*

ABSTRACT

- Agaves, which have been grown commercially for fiber or alcoholic beverages, are emerging as a candidate crop for biochemicals and biofuels production in semiarid regions because of their high productivity in low rainfall areas, drought tolerance, and low lignin content.

- In this work, we present the transcriptomic atlas of *Agave sisalana*, *Agave fourcroydes*, and agave hybrid 11648 (*A. amaniensis* x *A. angustifolia*) under prolonged drought in field conditions. Leaf, stem, and root tissues were sequenced, and gene expression profiles were correlated with biomass composition, enzymatic hydrolysis of cell wall carbohydrates, histochemical analysis, and non-structural carbohydrates content.

- Differences in biomass accessibility were attributed to either lignin content or lignin composition, possibly through modification of s/g ratio promoted by changes in Caffeic Acid 3-O-Methyltransferase (COMT) transcript abundance. Unlike most plants, the most highly expressed transcripts do not encode photosynthetic proteins, but rather involved in stress response. Although the three cultivars presented quantitative differences in global
gene expression, they activated a highly overlapping set of genes. The main molecular strategies employed by agave to cope with high-temperature and drought seem to consist in overexpressing HSP and LEA, as well as promoting raffinose accumulation as an osmolyte.

- In conclusion, our data provide vital new genetic information for the study of Agave species and provide new insights into cell wall architecture, recalcitrance, and resistance to abiotic stresses for these species.

**Keywords:** Sisal, RNAseq, biomass, abiotic stress.

1. **BACKGROUND**

Agaves are evergreen xerophyte monocots native to the semiarid regions of North America that are cultivated worldwide as ornamental plants (Nobel, 1994), and as crops for natural fiber and alcoholic beverages production (Davis & Long, 2015; Medina, 1954; Nobel, 1994). Among the agaves grown to obtain fiber, the most widely cultivated is *Agave sisalana* (Sisal), though other taxons such as *Agave fourcroydes* (henequen) and hybrid 11648 (*A. amaniensis* x *A. angustifolia*) are also of great importance (Davis & Long, 2015; Duarte et al., 2018; Medina, 1954). Brazil is currently the world's largest producer and exporter of sisal fiber, representing 70% of the exports and 58% of the global production (Davis & Long, 2015; FAO, 2020). The Brazilian semiarid total area is around 83 million hectares (Projeto MapBiomas, 2019), which is bigger than most European countries, like France or Spain. Only a small portion of this territory is used for agriculture, and, in these areas, sisal is often the only growing crop alternative with attractive economic results (Alvarenga Jr., 2012; Broeren et al., 2017; Silva & Beltrão, 1999). Since the late 1960s, sisal fiber production has been decreasing dramatically, mainly due to competition with synthetic products. Nonetheless, in recent years, an emerging interest in natural fibers has made sisal production rise again (Davis & Long, 2015; FAO, 2020). This interest could be attributed to environmental concerns about microplastic pollution as well as the demand for niche markets (Cesa, Turra, & Baruque-Ramos, 2017; Davis & Long, 2015). Still, only 4% of the harvested leaves are converted into commercial fiber, resulting in a huge amount of bagasse that is dumped back into the field (Leão et al., 2016). This waste is rich in carbohydrates and organic acids, presents high technological potential (as pharmaceuticals, cosmetics and nanocellulose) and could be suited as biorefineries feedstocks (Borland, Griffiths, Hartwell, & Smith, 2009; Branco et al., 2010; Davis, Kuzmick, Niechayev, & Hunsaker, 2017; Dellaert, 2014; Lacerda, de Paula, Zambon, & Frollini, 2012; Michel-Cuello, Juárez-Flores, Aguirre-Rivera, & Pinos-Rodríguez, 2008; Morán, Alvarez, Cypriano, & Vázquez, 2008; Santos, Vieira, Braz-Filho, & Branco, 2015).

The great success of Agave species in hostile environments is largely associated with its photosynthesis, the crassulacean acid metabolism (CAM) (Borland et al., 2009; Yang et al., 2015). CAM is a temporally regulated inorganic carbon concentration mechanism that relies on the primary CO₂ fixation during nighttime, when evapotranspiration rates are reduced, to maximize water use efficiency, which results in up
to 80% less water use when compared to conventional C₃ and C₄ crops (Borland et al., 2014; Stewart, 2015; Yin et al., 2018). Within CAM species, agaves are among the few that can achieve high productivity and, for that, are being considered as bioenergy feedstocks for dryland areas, which are commonly neglected and underused (Borland et al., 2014; Davis et al., 2017; Nobel, 1994; Owen, Fahy, & Griffiths, 2016; Somerville, Youngs, Taylor, Davis, & Long, 2010). Depending on region and species, commercially grown agave can achieve yields ranging from 8.5 to 22 Mg ha⁻¹ (Davis, LeBauer, & Long, 2014); a recent field trial in Arizona (US) found yields between 4.0–9.3 Mg ha⁻¹ yr⁻¹ total biomass with 300–530 mm yr⁻¹ of water input, which were already greater than conventional crops in Arizona, like cotton (1.5 Mg ha⁻¹ yr⁻¹), with less water consumption (Davis et al., 2017). However, theoretical analysis indicates even greater potential productivity (38 Mg (dry) ha⁻¹ yr⁻¹) (Owen et al., 2016). Also, agaves offer other desirable traits as biorenewables and biofuels feedstocks, such as the abundance of non-structural carbohydrates (mainly fructans), high shoot to root ratio, and low lignin content (Borland et al., 2009; Smith, 2008). Among cell wall polymers, lignin is the main cause of recalcitrance and a barrier for lignocellulosic biofuels production (Ragauskas et al., 2014; Simmons, Loqué, & Ralph, 2010). Therefore, understanding how the agave cell wall is formed may be important to find new ways to reduce lignin content without impairing plant growth (Simpson et al., 2011; Stewart, 2015; Vanholme, Ralph, et al., 2010). Furthermore, agaves constitute interesting models for the study of severe abiotic stress responses in plants, including drought and high temperatures, which is relevant for the development of agronomic solutions associated with climate change and bioenergy production (Borland et al., 2014; Stewart, 2015).

Despite their social and economic importance, few studies have been carried out with Agave species and their hybrids, especially at the molecular level. This is mainly because of: (i) lack of basic genetic knowledge; (ii) large genomes, which are estimated to be between 2,940 and 4,704 Mbp; (iii) long life cycles (5-12 years) (Davis & Long, 2015; Simpson et al., 2011). Nonetheless, the Agave genus is rich in genetic diversity, as ploidy varies from 2n to 8n (n=30) even within species, and hybridization has occurred frequently in wild cultivars, contributing to increasing variation in ploidy and complexity of hybrid genomes (Davis & Long, 2015; Simpson et al., 2011). Although DNA-seq data from Agave tequilana has been deposited at NCBI, a reference genome sequence has not been published to date. Furthermore, only six studies have been published on agave transcriptomics, but none of them has explored the molecular basis of cell wall biosynthesis (Abraham et al., 2016; Cervantes-Pérez et al., 2018; Gross et al., 2013; Huang et al., 2018, 2019; McKain et al., 2012; Sarwar et al., 2019; Simpson et al., 2011).

In the present work, we analyze the transcriptomic profile of Agave sisalana, Agave fourcroydes, and agave hybrid 11648 (A. amaniensis x A. angustifolia) in field conditions under prolonged drought, and correlate transcription with non-structural carbohydrates profiles, biomass composition, enzymatic hydrolysis of cell wall carbohydrates, histochemical analysis of leaf anatomy. The valuable resources generated offer a new
window of opportunity to molecular breeders fulfill the emerging expectations for agave use as biorefineries feedstocks for dryland areas and provide insights into abiotic stress resistance and cell wall architecture.

2. METHODS

2.1. Plant material and growth conditions

*A. sisalana*, *A. fourcroydes*, and agave hybrid 11648 (*A. amaniensis* x *A. angustifolia*) (H11648) samples were harvested under field conditions from Embrapa’s germplasm bank, located on the Experimental Station of Monteiro-PB, Brazil (7º 53' south latitude, 37º 07' west longitude, 619 m altitude). Plants were grown without artificial irrigation in non-calcic brown soil and semi-arid climate (BS, according to the Köppen system) (BRASIL, 1972; KÖPPEN & GEIGER, 1936). Rainfall data from the National Institute of Meteorology of Brazil shows that the municipality of Monteiro has suffered from a prolonged drought, not exceeding 150 mm of rainfall, since April 2014 until the sampling date (supplementary material S1). For all cultivars, leaf (central fraction), root, and stem were sampled from seven-year-old healthy adult plants growing side-by-side (n=3). As agave leaves grow during daytime (Abraham et al., 2016), sample harvested at midday were selected for sequencing to better dissect the relationship between abiotic stress response and cell wall biogenesis.

2.2. RNA Extraction and sequencing

Total RNA was extracted according to the protocol described by Zeng & Yang, 2002, with the modifications proposed by Le Provost et al., 2003. The RNA concentration and quality were verified using a Nanodrop 2000 spectrophotometer (Thermo Scientific) and integrity with HT RNA LabChip® Kit (Caliper Life Sciences). mRNA libraries and sequencing were done at the High-Throughput Sequencing Facility of the Carolina Center for Genome Sciences (University of North Carolina at Chapel Hill, USA). The libraries were prepared using the KAPA Stranded mRNA-Seq kit (07962193001) for Illumina platforms following the manufacturer's protocol, using 1 µg total RNA. To facilitate the assembly of transcripts, larger insert sizes (150-180 bp) were generated altering the fragmentation step of the protocol to 7 min at 94°C. The sequencing was done on the Illumina HiSeq 4000 system. Libraries were separated into two pools and each pool was sequenced in a single lane, generating 50bp paired-end reads.

2.3. *De novo* transcriptome assembly, quantification, and annotation

FastQC software was used for quality control (Andrews, 2010). Reads from each cultivar were assembled separately using Trinity v. 2.5.1 (Grabherr et al., 2011). Transcript quantification was performed using kallisto v 0.44.0 (Bray, Pimentel, Melsted, & Pachter, 2016) with 100 bootstraps, returning TPM (Transcripts Per
Million) values representing the transcripts’ abundance. For ORF prediction, we used Transdecoder v. 5.0.2 (Haas et al., 2013) configured to a minimum length of 200 nucleotides. Subsequently, we selected the longest isoform of each locus, considering only those with a TPM value greater than 1, and ORF length longer than 255 nucleotides. Functional annotation assignment was performed using Pannzer2 (Törönen, Medlar, & Holm, 2018). Unannotated proteins were blasted (BLASTp) against the Uniref90 database.

**2.4. Fungal identification and classification**

We have developed an in-house Perl script that used the Uniref90 BLASTp results and a list of fungal genera from Taxonomy DB (NCBI) to separate the fungal transcripts from the agave ones. The criterion to classify as fungal sequence was based on the percentage of fungal genera in the top blast hits considering an e-value <= 1e-10 (it classifies as fungi if at least 80% of the top ten hits are fungi).

To confirm this classification and separate the fungi transcripts between Ascomycota, Basidiomycota and other fungi, we used two strategies: a) Kaiju program v. 1.6.3 (Menzel, Ng, & Krogh, 2016) with fungi database to perform the taxonomy composition analysis and b) BLASTn against three separate databases of coding sequences (CDS) obtained from NCBI (Ascomycota genomes, Basidiomycota genomes and complete genomes of other fungal phyla) that were analyzed by an in-house Perl script, considering the alignment query coverage >= 30% and the e-value <= 1e-20 to separate the fungi groups.

**2.5. Tissue-specific transcripts analysis**

To analyze the specific transcription profile of each tissue, we used the SPM (specificity measure) metric implemented by the software tspex (https://github.com/apcamargo/tspex). Tissue specific genes with SPM >= 0.95 were selected.

**2.6. Orthologous analysis between cultivars**

To compare the transcriptome within cultivars, we first defined orthologous groups using OrthoMCL v. 1.4 (L. Li, 2003) using nucleotide sequences from the assembled transcriptomes.

**2.7. Differential gene expression analysis and GO enrichment**

R package sleuth v. 0.29.0 (Pimentel, Bray, Puente, Melsted, & Pachter, 2017) was used for the differential expression analysis with the Wald test. Pairwise comparisons were made between tissues within each cultivar, and differentially expressed genes (DEG) with a fold change > 2 and FDR ≤ 0.05 were selected. Representative DEG from each tissue (using sets with two lists of genes preferentially expressed in that tissue) were submitted to Gene Ontology enrichment analysis using topGO R package (Alexa, Rahnenfuhrer, & Lengauer, 2006). We considered only enriched Gene Ontology terms with p-value < 0.05 for further analyses.
2.8. Comparative genomics

To identify exclusive and expanded gene families from Agave and other related species, we employed a comparative genomics approach using OrthoFinder v. 2.3.1 (Emms & Kelly, 2015). Single-copy orthologs were aligned with MAFFT v. 7.394 (Katoh & Standley, 2013) configured to use the L-INS algorithm with 1,000 iterations. All alignments were concatenated in a supermatrix that was used for the phylogenetic inference with IQ-TREE v. 1.6.8 (Nguyen, Schmidt, von Haeseler, & Minh, 2015). BadiRate v. 1.35 (Librado, Vieira, & Rozas, 2012) was used to identify expanded gene families using a FDR threshold of 0.05. We compared our data to protein sequences from genomes available at the Phytozome database (Goodstein et al., 2012). We selected Amborella trichopoda (outgroup) v. 1.0, Asparagus officinalis v. 1.1, Eucalyptus grandis v. 2.0, Sorghum bicolor v. 3.1.1, Ananas comosus v. 3, Punica granatum (Qin et al., 2017) and Saccharum spontaneum (Zhang et al., 2018). These species were chosen based on their photosynthetic metabolism, biomass composition and productivity.

2.9. Non-structural carbohydrates profiling

Rhamnose, ribose, fructose, pinitol, galactose, mannitol, sorbitol, glucose, sucrose, maltose, trehalose, melibiose, and raffinose were quantified in leaves and stem using a gas chromatographer (Hewlett-Packard® 6890) connected to a quadrupole mass spectrometer (Hewlett-Packard® model 5973) following the same methodology as Freschi et al., 2010. The initial running condition was 95 °C for 2 min, followed by a gradient up to 320 °C at 8 °C min⁻¹. The column used for separation was an HP-1701 (30 m, I.D. 0.25 mm, 0.5 μm), with helium as the carrier gas, with a flux of 4 mL min⁻¹. The endogenous metabolite concentration was obtained by comparing the peak areas of the chromatograms with commercial standards. All measurements were made in triplicate.

Water extraction for 15 min at 80°C (Mancilla-Margalli & López, 2006) followed by thermally hydrolyzation (80°C for 30h in water) of half of the samples (Michel-Cuello et al., 2008) was performed to estimate of the oligosaccharides and monosaccharides within the cultivars. Oligosaccharides were estimated by comparing the difference between the hydrolyzed and non-hydrolyzed fractions. All the samples were evaluated using reducing sugars content measured with DNS assay (Miller, 1959).

2.10. Determination of structural carbohydrates and lignin in biomass

Cellulose, hemicellulose, lignin, extractives and ash were quantified according to the standardized analytical methods of Sluiter et al., 2016.

2.11. Enzymatic hydrolysis of cell wall carbohydrates
A simple enzymatic hydrolysis was performed to verify the saccharification and accessibility potential of the agave biomass. We performed the enzymatic hydrolysis assays following the Bragatto et al., 2012 protocol with the modifications of Lepikson-Neto et al., 2014. Cellulase enzyme C2730 (Sigma) was used, and the content of reducing sugars was measured with DNS assay (Miller, 1959), as suggested by Lepikson-Neto et al., 2014. Also, for comparison, we added an Energy Cane sample (US59-6; whole plant biomass) processed in parallel.

2.12. Histochemical analysis of Agave leaves

To complement the compositional data, the central fraction of fully developed leaves from one-year-old plants of the three Agave cultivars was sampled to verify the distribution of lignin, pectic compounds, and callose. Leaves cross-sections were stained with phloroglucinol-HCl, for lignin detection (Johansen, 1940), and with ruthenium red, for pectic compounds (Johansen, 1940), and, then, observed under light Olympus BX 51 photomicroscope equipped with an Olympus DP71 camera. For callose, samples were stained with aniline blue (Currier & Strugger, 1956) and observed under an epifluorescence microscope with a UV filter (BP 340 to 380 nm, LP 425 nm).

3. RESULTS

3.1. Transcriptome assembly and annotation

The total number of transcripts obtained by de novo assembly of RNA-seq reads from the different Agave species ranged from 136,692 to 170,474. After the removal of short (<250 bp) and weakly expressed transcripts (TPM<1), we selected one transcript per locus. The total number of transcripts ranged from 23,973 to 26,842. These values are similar to those reported by Gross et al., 2013 in the transcriptomic analysis of other agave species: 34,870 and 35,086 protein-coding loci for A. tequilana and A. deserti, respectively. In addition, Asparagus officinalis, which is the closest species to the agaves with a reference genome available, has 27,656 protein-coding genes (Harkess et al., 2017). Several unannotated proteins ranging from 12.4 to 14.2% were found. All transcriptomic data is available in supplementary table 2 (S2). Remarkably, around 12% of the root transcripts belonged to fungi. The presence of these fungal transcripts was persistent in roots for all cultivars and biological replicates, and 99.8% of those transcripts were root-specific, with SPM > 0.95. Interestingly, 5-6% of the fungal transcriptomes were annotated as heat shock proteins - HSP, and, in some cases, the fungal HSP were as abundant as some of the agave’s HSP (depending on the cultivar, the most expressed fungal HSP ranged from 120 to 413 TPM.). It was possible to identify 46.8-54.7% of the fungal transcripts as Ascomycetes and 16.3-20.6% as Basidiomycetes. The fungal transcripts were excluded from further analysis.
### 3.1.1. Most expressed transcripts

From the transcriptome quantification, we have ranked the ten most highly expressed transcripts in each tissue for each cultivar (Table 1). Surprisingly, a similar set of highly expressed transcripts was identified in all tissues of the three cultivars. One of those transcripts is LEA-5 (late embryogenesis abundant protein 5). LEA is one of the most expressed, with transcription ranging from 2,772 to 15,947. LEA transcripts have been previously found in many vegetative tissues with several functions associated with abiotic stress responses, including heat and salinity in Agave species (Liang et al., 2014; Pedrosa, Martins, Gonçalves, & Costa, 2015; Tamayo-Ordóñez et al., 2016). Other highly expressed transcripts encode Ubiquitin and Polyubiquitin proteins. However, for H1648 leaves and stem, these transcripts were not at the topmost. Ubiquitins are related to proteolysis signaling, which has also been associated with drought stress (Flick & Kaiser, 2012a; Lyzenga & Stone, 2012). Many heat shock proteins and uncharacterized protein transcripts are highly expressed in the three cultivars as well.

### 3.1.2. Tissue-specific transcripts

As expected, our analysis revealed that most of the leaf-specific transcripts are photosynthesis-related (Table 2). Interestingly, in the stem for all cultivars, we have found several homeobox transcripts, which are proteins commonly associated with development (Mukherjee, Brocchieri, & Burglin, 2009). The most expressed root-specific genes are no-hit proteins, and for A. fourcroydes and A. sisalana ubiquitins are present as well. Also, A. sisalana has one root-specific heat shock protein with high expression; and H11648 has a Thaumatin, which is a sweet-tasting protein homologous to pathogenesis-related (PR) protein PR-5 (Min et al., 2003). One member of the PR-5 protein family, Osmotin, is accumulated in cells adapted to osmotic stress (Singh et al., 1987).

### 3.1.3. Phenylpropanoids pathway

The phenylpropanoid pathway is composed of many branches, but all of them share the same common precursors (Fraser & Chapple, 2011). We analyzed the transcriptional profile of the lignin and flavonoids biosynthesis branches (Figure 6a). At least one member of each gene family of the phenylpropanoid pathway was differentially expressed, and here we highlight the main focal points. Phenylalanine Ammonium Lyase (PAL), the first enzyme of the pathway, presented similar behavior in A. fourcroydes and A. sisalana, in both species, the most expressed isozyme gene had higher transcription in the stem. Also, an orphan PAL isozyme gene was found for each species. Notwithstanding, for H11648, the main PAL isozyme gene was not differentially expressed. The first diverging point is the conversion of p-coumaroyl CoA by either Chalcone...
Synthase (CHS) or Hydroxycinnamoyl-coenzyme A Shikimate:Quinate Hydroxycinnamoyltransferase (HCT). For CHS, which is the starting enzyme for the flavonoid biosynthesis, two main encoding genes were found to be differentially expressed. However, the overall expression of these two isozymes was relatively low. Curiously, one of those isozymes presented significantly higher expression at the H11648 stem, although, its transcription was not very high. For HCT, which catalyzes the outset of the lignin branch, four HCT isozyme genes were deferentially expressed, and two were responsible for the majority of the transcription. Whereas for H11648, these two isozyme-encoding gene did not present any relevant differences. For the other cultivars, each HCT isozyme-encoding transcript behaved oppositely, being either more expressed in leaves or stems; a phenomenon particularly evident in A. sisalana. On the phenylpropanoid pathway, there are only two enzymes exclusive to syringyl (S) monomer biosynthesis: Ferulate 5-Hydroxylase (F5H) and Caffeic Acid 3-O-Methyltransferase (COMT). We identified just one F5H isozyme-encoding transcript differentially expressed and more expressed in stems than leaves in all cultivars. On the other hand, three COMT isozymes were found, with two predominantly expressed in stems and third one mainly found in leaves. Interestingly, in general, COMT expression was higher at A. fourcroydes stem.

3.1.4. Exclusive and expanded gene families of Agave

Twelve expanded families were found in the Agave clade using phylogeny and the birth and death models to calculate rates of gains and loss of genes through evolution (see methods 2.8). Among them, we identified a few transcription factors like MYB, Far-Red Impaired Response 1 (FAR1) and Zinc finger BED domain-containing protein DAYSLEEPER (OG0000003, OG0000022, and OG0000084, respectively). FAR1 activates the transcription of Far-Red Elongated Hypocotyl 1 (FHY1) and its homolog, FHY1-Like (FHL), and positively regulates chlorophyll biosynthesis via the activation of Delta-aminolevulinic acid dehydratase 1 (HEMB1) (Hudson, Lisch, & Quail, 2003; Lin et al., 2007; Tang et al., 2012; Wang & Deng, 2002). Two expanded families unannotated (OG0000281 and OG0000512) and one alpha expansin were also detected. No heat shock proteins or other apparent abiotic stress-related families were expanded.

Regarding biomass, we have found some important exclusive families. Gene families that contain at least one representative member shared in the agave cultivars and in the other analyzed species were considered as exclusive families. Among the 149 that are shared with Saccharum spontaneum, 68 are unannotated, and two are directly related to the cell wall: callose synthase 5 (OG0016451) and cellulose synthase (fragment) (OG0013945). In comparison, Eucalyptus grandis 94 exclusive families were found, of which only 48 were annotated. We identified two families related to lignin metabolism: a Caffeoyl-CoA 3-O-methyltransferase (CCoAMT) (OG0014109) and Laccase-14 (LAC14) (OG0013133). We also identified an
endoglucanase (OG0012311) and an alpha-glucosidase (OG0012714). The full list of ortholog genes families generated by the comparative genomics approach is presented at the supplementary table 3 (S3).

### 3.2. Chemical analysis

#### 3.2.1. Non-structural carbohydrates profile

The GC-MS analysis detected all the following saccharides in both leaf and stem: fructose, glucose, sucrose and raffinose (Fig. 7a). Among them, raffinose was the most abundant carbohydrate for every cultivar and tissues sampled. Raffinose is an oligosaccharide related to osmotic and oxidative stress response in plants. Under dry conditions, raffinose acts as a compatible osmolyte avoiding water loss to the environment (Ende, 2013; Nishizawa, Yabuta, & Shigeoka, 2008; Sengupta, Mukherjee, Basak, & Majumder, 2015). In seeds of monocots, such as barley, corn, and sorghum, raffinose concentrations oscillates between 2.6-7.9 mg/g (Kuo, VanMiddlesworth, & Wolf, 1988). In contrast, for *Agave* vegetative tissues, we encountered mean values ranging from 10.96 to 23.38mg/g. Perhaps, the elevated presence of this carbohydrate could be related to the prolonged drought that challenged the plants under field conditions. Interestingly, no significant differences were found between the samples for this carbohydrate. That was not the case for fructose, glucose and sucrose. Glucose content was statistically different in the comparison between *A. fourcroydes* and H11648 leaves, with almost double of this carbohydrate being present in the leaves of H11648. However, there were no significant differences for the other carbohydrates nor between *A. sisalana* and *A. fourcroydes*. Among the leaves of *A. sisalana* and H11648, a difference was detected in glucose and fructose, again almost 50% more was quantified in the leaves of the hybrid.

The highest concentrations of water-soluble carbohydrates were found in leaves regardless of the cultivar (Fig. 7b). *A. fourcroydes* and H11648 contained the highest sugar content. However, their mono and oligosaccharides fractions were contrasting. For *A. fourcroydes*, 26.41% of sugars detected were monosaccharides, while the same category in the H11648 represented 76.11%. Interestingly, the transcriptomic data revealed a higher transcription of fructan-exohydrolases (FEH) in H11648 leaves (Fig. 4b) that may explain the differences found in monosaccharides content. The oligosaccharides fraction for *A. sisalana* leaves was also higher than the monosaccharide. In stem, *A. sisalana* presented higher oligosaccharides content than the other cultivars, when analyzing the transcriptome, we noticed that fructosyltransferases (FT) were more abundant in those samples.

#### 3.2.2. Compositional analysis of *Agave* leaves and stem

Regardless of the cultivar or tissue, the extractives represented the largest fraction of the biomass (Figure 6B). *A. fourcroydes* presented the highest content variation among its tissues, going from 31.2% in
stem to 56.89% in leaf. The other cultivars and their tissues varied little, ranging from 49.3% to 52.94%. However, only *A. fourcroydes* stem was statically different from the other samples.

In general, the stems were more lignified than leaves. Considering leaves, *A. fourcroydes* lignin fraction was significantly lower than *A. sisalana* and H11648. For *A. fourcroydes*, lignin levels in leaves were practically half of those encountered for stem. Also, the stem of this agave cultivar presented the highest lignin content in all samples. No statistical difference was found between *A. sisalana* and H11648 lignin fractions. In relation to the cellulose content, *A. fourcroydes* stem presented the highest percentage (32.25%), with no statistical difference among the other samples. As for hemicellulose mass fraction, similar patterns to lignin were found. Leaves had lower hemicellulose contents and the stems, higher. There is no significant difference between the leaves of the cultivars. Nonetheless, among stems, hemicellulose fraction was significantly higher for *A. fourcroydes*, this cultivar presented 18.81%, which is 2.5x higher than its leaves. Concerning ashes, the highest values were found at *A. sisalana* leaves (11.28%), and the lowest was found at *A. fourcroydes* stem. In general, ash fractions in leaves were almost two times higher than stems.

### 3.2.3. Enzymatic hydrolysis of cell wall carbohydrates

The hydrolysis protocol was effective in *Agave*, and, as expected, differences between samples were observed. Regardless of the agave cultivar, the stem showed higher recalcitrance than the leaves. The leaves of *A. fourcroydes* had the highest saccharification, while the stem of H11648 was the most recalcitrant sample. Compared to Energy Cane, *A. fourcroydes* leaves were 47% more hydrolysable, while *A. sisalana* and H11648 were 37%. Energy Cane hydrolysis was statistically equal to the stems of *A. sisalana* and *A. fourcroydes*. Within stems, *A. sisalana* and H11648 presented similar chemical composition; however, *A. sisalana* stem was the less recalcitrant. Interestingly, even with higher cellulose and lignin levels, the stem of *A. fourcroydes* was more accessible for digestion than the H11648 stem.

### 3.2.4. Microscopic analysis of leaf anatomy and cell wall composition

The three cultivars presented thick cuticle, sunken stomata, and well-developed fiber caps, as previously described for *Agave* (Blunden, Yi, & Jewers, 1973). Pectic compounds deposition was mainly observed in the outer periclinal epidermal walls and mesophyll cells (supplementary material S4 a), as well as fibers around the vascular bundles, which stained intensely (figure 8 a-c). Compared to other cultivars, *A. sisalana* vascular bundle presented more fiber cap cells with thicker cell walls, which may explain the higher fiber quality of this cultivar (Medina, 1959). In contrast to pectin deposition, which was widespread, lignification was detected only in the secondary cell wall of the xylem conducting cells for all cultivars (figure 8 d-f). For the aniline blue staining, although we encountered brighter regions in the cuticle and fiber caps cells, which could indicate callose deposition, autofluorescence impaired the analysis (supplementary material S4 and S5).
4. DISCUSSION

Little is known about the chemical composition of agaves, especially for the cultivars utilized for fiber production. Few studies have been done on the subject, and most of the data available focus on the chemical composition of the fibers itself, not whole leaves or stems, which could be misleading when evaluating agave bioenergetic potential (Davis & Long, 2015; McDougall, Morrison, Stewart, Weyers, & Hillman, 1993; Mylsamy & Rajendran, 2010; Vieira, Heinze, Antonio-Cruz, & Mendoza-Martinez, 2002). Our data revealed some differences and common features between the fiber producing cultivars. In general, extractives represented a major part of the biomass composition, few differences were found on cellulose mass fraction, and agave stems were more lignified than leaves. Extractives, which are nonstructural components of biomass, may include pectin compounds and other waxes, phenolics, resin compounds, inorganic material, non-structural sugars, and nitrogenous material, among others (Pecha & Garcia-Perez, 2015). Using the same NREL’s normatives as we did, Yang et al., 2015 has found a chemical composition in A. tequilana leaves similar to those reported here for the stems. The agaves developed for fiber production seem to have advantages from the industrial perspective considering the contents of lignin in leaves. Biorefineries based on agave leaves were suggested as an interesting alternative due to the possibility of maintaining a constant supply of plant material through the year, unlike stem-based industries in which the harvest occurs every 5-6 years (Davis, Dohleman, & Long, 2011; Davis & Long, 2015; Nobel, 2010; Yang et al., 2015). Compared to other crops, the lignin mass fraction, regardless of the tissue or cultivar, was lower than those found in woody biomass (21-32%, e.g., Populus, Eucalyptus and Pinus) (Ragauskas et al., 2014) and less or equal to found in herbaceous crops (9-18%, e.g., Miscanthus, Panicum, and corn straw) (Ragauskas et al., 2014).

In Agave leaves, lignin deposition occurs only on xylem vessels. Lignin is important for the xylem to hold its structure under negative pressure allowing water transport to occur properly (Campbell & Sederoff, 1996; Kitin et al., 2010). The lack of lignin in other leaf cells may be related to its hydrophobicity, as most of the agave leaf tissue is composed of water-storage cells of the mesophyll, and, in those cells, we encountered deposition of pectic compounds, which are hydrophilic. In addition, a high concentration of pectin in the cell wall was previously described to increase water absorption speed in leaves (Boanares et al., 2018), especially when accumulated in the epidermis. For all three cultivars, we encountered outer periclinal epidermis walls that are thicker with pectin deposition. The abundance of pectic compounds in leaves could contribute to the high percentage of extractives found in the compositional analysis.

Indeed, lignin content was the most interesting aspect of our chemical data, especially when analyzed together with the enzymatic hydrolysis assay. The differences observed on biomass accessibility could be explained by lignin mass fraction and possibly by its composition. The structure of lignin consists basically of three monomeric units: (I) syringyl (S); (II) guacyl (G); and (III) p-hydroxyphenyl (H) (Vanholme,
The individual contribution of each monomer to the composition of this polymer varies significantly between tissues and species (Simmons et al., 2010). Higher rates of the S/G ratio are advantageous for the production of paper, cellulose and bioenergy, since the S-rich lignin is more easily dissociated from the cellulosic material (Lepikson-Neto et al., 2014; Simmons et al., 2010). It is possible that the differences found in the enzymatic hydrolysis of the stems, specifically the comparison between A. fourcroydes and H11648, are closely related to the S/G content. Although A. fourcroydes stem presented higher lignification, its biomass was less recalcitrant than H11648 stem. Moreover, the transcriptomic analysis of the phenylpropanoid pathway revealed that the expression of COMT, one of the main controlling points of the S lignin monomer biosynthesis, was higher in A. fourcroydes, suggesting that this cultivar is richer in S lignin, which would explain the higher accessibility of this tissue. The other recalcitrance oscillations may be partially explained by the differences in lignin mass fraction. In all cases, stems were more recalcitrant than leaves, and our expression data corroborate with this data, as most of the transcripts related to the lignin branch were more expressed in the stems.

Considering agave cell wall composition, it comes to one’s mind how can agaves be able to have such low lignin contents and still maintain their spatial structure supporting leaves that can achieve 1.7 meters long? Our phylogenetic analysis revealed an exclusive family of Callose synthase (CALS) between Saccharum spontaneum, A. sisalana and H11648. S. spontaneum was previously reported as having high expression and expanded families of CALS (Nascimento et al., 2019). Callose, a β-1,3-glucan polymer, is a cell wall component that is connected to a wide variety of plant processes, such as maintenance of the vascular system, plasmodesmata function, pollen and pollen tube development, cell plate formation, as well as biotic and abiotic defense responses (Falter et al., 2015; Schneider, Hanak, Persson, & Voigt, 2016). Although in our data, the exclusive CALS family did not present high expression, other CALS were differentially expressed and present at levels similar to those found for CsL, indicating that callose may play an important role in Agave. Also, low transcription of expansins were found, suggesting that the CALS transcription might not be connected to cell division and elongation activities. Other monocotyledonous energy crops with relatively low lignin content, such as maize and Miscanthus x giganteus, were found to form an unusual outer layer with callose fibrils interknitted in cellulose (Falter et al., 2015). In this context, we hypothesized that callose might be interacting with the agave cell wall and possibly thickening it to compensate for lower lignification through support and structure with the advantage to keep the cell wall elasticity, which is essential to allow the wide range of the water dynamics within the cells (filling and emptying in response to water availability). Fluorescence microscopy with aniline blue staining indicated a possible callose deposition at the fiber caps cells, which surrounds the vascular bundles, and are the main structural support for the leaves (Corbin et al., 2015; Rüggeberg et al., 2008). However, future analyzes are still required to confirm it. Perhaps, a novel strategy...
to genetically engineer crops with reduced lignin content without the typical decreased growth (Cesarino, Araújo, Domingues Júnior, & Mazzafera, 2013; Simmons et al., 2010; Vanholme, Ralph, et al., 2010) should consider improving callose fraction as well. Also, since callose is an easily degradable glucose polymer, it could still be harnessed for biofuels and biochemicals production.

However, low lignin contents do not come without negative collateral consequences, as positive correlations between lignin amount and pathogen resistance have been observed, especially for necrotrophic pathogens (Miedes, Vanholme, Boerjan, & Molina, 2014). Hence the importance of sisal bole rot disease caused by *Aspergillus welwitschiae*, a saprophytic fungus that infects sisal plants and behaves as a classical necrotrophic pathogen (Duarte et al., 2018). Many resistance mechanisms to opportunistic phytopathogens are related to plant secondary metabolites like flavonoids (Treutter, 2006). When analyzing the phenylpropanoid pathway, we were expecting that the flavonoid branch would present higher expression than the lignin one, since these two branches are competitive (Lepikson-Neto et al., 2014; X. Li, Bonawitz, Weng, & Chapple, 2010; Salazar et al., 2013). However, *CHS* transcript abundance was lower than *HCT*. Nevertheless, the phenylpropanoid pathway as a whole appears to be less expressed than other cell wall biosynthetic pathways such as cellulose. Precipitation, micrometeorological factors, and humidity are some of the most important components involved in the epidemiology of fungal pathogens and play a key role in both dispersion and spore germination (Bashi & Rotem, 1974; Buchanan, Gruissem, & Jones, 2015; Cook & Papendick, 1972; Huber & Gillespie, 1992; Rotem & Palti, 1969). It is possible that agaves can afford to express less of the phenylpropanoid pathway by inhabiting environments that are unfavorable to the development of fungal infections and, thus, redirecting valuable resources to abiotic stress resistance mechanisms.

In fact, agaves seem to invest heavily in drought and high-temperature resistance mechanisms. Our transcriptomic data revealed that the most expressed transcripts encode proteins that are well known to be stress-responsive. In *A. fourcroydes* leaves, *Phosphoenolpyruvate carboxylase* responsible for the first step in CAM pathway, appears only in the fifth position of our most expressed transcripts list (Table 1) with an expression twenty-five times lower than the topmost (a Chaperone protein DnaJ); for *A. sisalana* this transcript is the fourth and for H11648 is the twenty-fourth. Since obligatory CAM species cannot count on transpirational cooling during the day and inhabit areas of high-light-intensity, these plants have developed alternative strategies to cope with high-temperature stress (Bita & Gerats, 2013; Borland et al., 2014; Sarwar et al., 2019). Our data suggest that for Agave, the molecular strategies may consist of overexpressing *HSP* and *LEA*, as well as genes related to proteolysis processes. These molecular mechanisms are present in all tissues regardless of the cultivar, yet, differences within cultivars were found at the transcript level. For instance, the most expressed *HSP* for *A. fourcroydes* was an *HSP40*, for *A. sisalana* it was an *HSP70*, and for H11648 a small *HSP*. 
Considering LEA-encoding transcripts, the main difference appeared on expression levels between tissues. In general, LEA5 was more abundant in leaves than stem and roots. Nonetheless, this transcript was one of the most expressed in every sample. Although their precise role has not been defined, LEA proteins may help prevent the formation of damaging protein aggregates due to desiccation or osmotic stresses (GOYAL, WALTON, & TUNNACLIFFE, 2005; Hundertmark & Hincha, 2008; Liu, Chakrabortee, Li, Zheng, & Tunacliffe, 2011). It has been proposed that LEA proteins have different molecular mechanisms than chaperones, and evidence suggests that they can play a role as integral membrane protein (Caramelo & Iusem, 2009; Menze, Boswell, Toner, & Hand, 2009), and even protect mitochondrial membranes against dehydration damage (Caramelo & Iusem, 2009; Tolleter, Hincha, & Macherel, 2010). Also, previous work has correlated LEA expression with abiotic stress resistance in Agave (Tamayo-Ordóñez et al., 2016). Another interesting molecular strategy is proteolysis; we have found a GO term related to this process in leaves, and many ubiquitins highly expressed in all tissues and cultivars. Ubiquitylation has been reported to regulate many aspects of stress-response, being responsible for labeling damaged proteins for proteolysis by the proteasome system (Flick & Kaiser, 2012a). Thereby, the cells can remove proteins damaged by stress that are unable to function properly, preventing the accumulation of potentially harmful aggregates and recycling nutrients, especially nitrogen (Flick & Kaiser, 2012a; Lyzenga & Stone, 2012).

Many plants accumulate compatible osmolytes to endure drought and high salinity stress through osmoprotection, but these compounds vary among species. For example, while the presence of the amino acid proline as a compatible solute is widespread in plants, just a few members of Plumbaginaceae family accumulates β-alanine betaine as an osmolyte (Buchanan et al., 2015; Rathinasabapathi, Fouad, & Sigua, 2001). One important osmolyte is trehalose, whose biosynthetic genes were differentially expressed, as well as an unshared family in A. sisalana. Previous transcriptomic data have suggested that, in Agave sisalana, trehalose may be an osmolyte (Sarwar et al., 2019). However, in our samples, no actual trehalose was detected by the gas chromatography. Instead, our chemical analysis detected the raffinose in every sample tested. We suggest that, in Agave, one of the main osmolytes might be raffinose. This carbohydrate was detected in similar concentration in every sample tested, yet our transcriptomic data revealed clear distinction regarding the raffinose biosynthetic pathways; whereas A. fourcroydes seems to prefer the RafS-mediated synthesis, the other two cultivars seem to utilize the route mediated by GosL. It is plausible that the inclination of A. fourcroydes to use RafS-mediated synthesis could affect cellulose biosynthesis since SUSY and RafS could compete for the same substrate.

Even though leaves presented higher expression of genes related to abiotic stresses, it was in roots that we have found many of the GO terms referred to these responses (e.g., response to oxidative stress; response to stress). Soil surface can have large daily oscillations in temperature, and in desert-like regions, these
temperatures can easily overcome 40ºC (Nobel, 2010; Sattari, Dodangeh, & Abraham, 2017). As agaves present shallow roots systems (Nobel, 1988), this tissue is continuously exposed to harsh conditions and must develop special adaptations to cope with high temperature, salinity, and water deficit. It makes sense that the most abundant transcripts of the root-associated fungi are of proteins related to heat responses, which may indicate that the microbial community is well adapted to such an environment.

5. CONCLUSION

Drought is one of the most important environmental factors that impact plant growth and development. In the eminence of climate change, projections estimate disturbances in rainfall patterns and rise in global temperature, arid and semi-arid environments might become more common, and plants will have to be adapted to those conditions. Therefore, understanding relevant agronomical examples at molecular level can be the key to the development of new technologies. Perhaps, Agave may offer the ideal conditions to decipher the blueprint for biomass production under dry, hot conditions. Our inspection of these plants, thriving under extreme conditions, indicated that abiotic stress mechanisms pervade this crop metabolism. Stress-response genes were the most highly expressed transcripts, and, also, many of the plant processes have secondary functions that act in response to stress, from reserve carbohydrates that can stabilize membranes and protect cells against the effects of dehydration (Livingston, Hincha, & Heyer, 2009), photosynthetic mechanisms which are highly water-use efficient (Stewart, 2015), to cell wall components and adaptations that can control water flux (Boaãares et al., 2018; De Storme & Geelen, 2014). It is possible that the fine-tuning between these auxiliary functions and the main resistance mechanisms is what makes agaves so adapted to marginal ecosystems and still great biomass producers. In summary, our data constitute an important new resource for the study of Agave species and indicate potential mechanisms that could be used to improve the tolerance of other crops to drought and high-temperature conditions.

6. ACKNOWLEDGEMENTS

This work was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001, Center for Computational Engineering and Sciences - FAPESP/Cepid (2013/08293-7), and Fundação de Amparo à Pesquisa do Estado de São Paulo - Brazil (FAPESP) - process numbers 2016/05396-8 and 2017/04900-7.

7. CONFLICT OF INTEREST

The authors declare that they have no competing interests.

8. REFERENCES
Abraham, P. E., Yin, H., Borland, A. M., Weighill, D., Lim, S. D., De Paoli, H. C., … Yang, X. (2016). Transcript, protein and metabolite temporal dynamics in the CAM plant Agave. *Nature Plants*, 2(12), 16178. https://doi.org/10.1038/nplants.2016.178

Alexa, A., Rahnenfuhrer, J., & Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13), 1600–1607. https://doi.org/10.1093/bioinformatics/btl140

Alvarenga Jr., E. R. (2012). Cultivo e aproveitamento do Sisal (Agave sisalana). *Dossiê Técnico*, 24.

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*.

Bashi, E., & Rotem, J. (1974). Adaptation of Four Pathogens to Semi-Arid Habitats as Conditioned by Penetration Rate and Germinating Spore Survival. *Phytopathology*, 64(7), 1035. https://doi.org/10.1094/Phyto-64-1035

Bita, C. E., & Gerats, T. (2013). Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. *Frontiers in Plant Science*, 4. https://doi.org/10.3389/fpls.2013.00273

Blunden, G., Yi, Y., & Jewers, K. (1973). The comparative leaf anatomy of Agave, Beschorneria, Doryanthes and Furcraea species (Agavaceae: Agaveae). *Botanical Journal of the Linnean Society*, 66(2), 157–179. https://doi.org/10.1111/j.1095-8339.1973.tb02167.x

Boanares, D., Ferreira, B. G., Kozovits, A. R., Sousa, H. C., Isaias, R. M. S., & França, M. G. C. (2018). Pectin and cellulose cell wall composition enables different strategies to leaf water uptake in plants from tropical fog mountain. *Plant Physiology and Biochemistry*, 122, 57–64. https://doi.org/10.1016/j.plaphy.2017.11.005

Borland, A. M., Griffiths, H., Hartwell, J., & Smith, J. A. C. (2009). Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands. *Journal of Experimental Botany*, 60(10), 2879–2896. https://doi.org/10.1093/jxb/erp118

Borland, A. M., Hartwell, J., Weston, D. J., Schlauch, K. A., Tsチャplinski, T. J., Tuskan, G. A., … Cushman, J. C. (2014). Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends in Plant Science*, 19(5), 327–338. https://doi.org/10.1016/j.tplants.2014.01.006

Bragatto, J., Segato, F., Cota, J., Mello, D. B., Oliveira, M. M., Buckeridge, M. S., … Driemeier, C. (2012). Insights on How the Activity of an Endoglucanase Is Affected by Physical Properties of Insoluble Celluloses. *The Journal of Physical Chemistry B*, 116(21), 6128–6136. https://doi.org/10.1021/jp3021744

Branco, A., Santos, J. D. G., Pimentel, M. M. A. M., Osuna, J. T. A., Lima, L. S., & David, J. M. (2010). D-Mannitol from Agave sisalana biomass waste. *Industrial Crops and Products*, 32(3), 507–510. https://doi.org/10.1016/j.indcrop.2010.06.025
BRASIL, M. da A. (1972). Levantamento exploratório reconhecimento de solos do Estado da Paraíba. *Boletim Técnico MA/SUDENE, 15*, 670.

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology, 34*(5), 525–527. https://doi.org/10.1038/nbt.3519

Broeren, M. L. M., Dellaert, S. N. C., Cok, B., Patel, M. K., Worrell, E., & Shen, L. (2017). Life cycle assessment of sisal fibre – Exploring how local practices can influence environmental performance. *Journal of Cleaner Production, 149*(2017), 818–827. https://doi.org/10.1016/j.jclepro.2017.02.073

Buchanan, B. B., Gruissem, W., & Jones, R. L. (2015). *Biochemistry & molecular biology of plants* (2nd ed.).

Campbell, M. M., & Sederoff, R. R. (1996). Variation in Lignin Content and Composition (Mechanisms of Control and Implications for the Genetic Improvement of Plants). *Plant Physiology, 110*(1), 3–13. https://doi.org/10.1104/pp.110.1.3

Caramelo, J. J., & Iusem, N. D. (2009). When cells lose water: Lessons from biophysics and molecular biology. *Progress in Biophysics and Molecular Biology, 99*(1), 1–6. https://doi.org/10.1016/j.pbiomolbio.2008.10.001

Cervantes-Pérez, S. A., Espinal-Centeno, A., Oropeza-Aburto, A., Caballero-Pérez, J., Falcon, F., Aragón-Raygoza, A., … Cruz-Ramírez, A. (2018). Transcriptional profiling of the CAM plant Agave salmiana reveals conservation of a genetic program for regeneration. *Developmental Biology, 442*(1), 28–39. https://doi.org/10.1016/j.ydbio.2018.04.018

Cesa, F. S., Turra, A., & Baruque-Ramos, J. (2017). Synthetic fibers as microplastics in the marine environment: A review from textile perspective with a focus on domestic washings. *Science of The Total Environment, 598*, 1116–1129. https://doi.org/10.1016/j.scitotenv.2017.04.172

Cesarino, I., Araújo, P., Domingues Júnior, A. P., & Mazzaferra, P. (2013). An overview of lignin metabolism and its effect on biomass recalcitrance. *Brazilian Journal of Botany, 35*(4), 303–311. https://doi.org/10.1590/s0100-84042012000400003

Cook, R. J., & Papendick, R. I. (1972). Influence of Water Potential of Soils and Plants on Root Disease. *Annual Review of Phytopathology, 10*(1), 349–374. https://doi.org/10.1146/annurev.py.10.090172.002025

Corbin, K. R., Byrt, C. S., Bauer, S., DeBolt, S., Chambers, D., Holtum, J. A. M., … Burton, R. A. (2015). Prospecting for Energy-Rich Renewable Raw Materials: Agave Leaf Case Study. *PLOS ONE, 10*(8), e0135382. https://doi.org/10.1371/journal.pone.0135382

Currier, H. B., & Strugger, S. (1956). Aniline blue and fluorescence microscopy of callose in bulb scales of Allium cepa L. *Protoplasma, 45*(4), 552–559. https://doi.org/10.1007/BF01252676

Davis, S. C., Dohleman, F. G., & Long, S. P. (2011). The global potential for Agave as a biofuel feedstock.
Davis, S. C., Kuzmick, E. R., Niechayev, N., & Hunsaker, D. J. (2017). Productivity and water use efficiency of Agave americana in the first field trial as bioenergy feedstock on arid lands. *GCB Bioenergy, 9*(2), 314–325. https://doi.org/10.1111/gcbb.12324

Davis, S. C., LeBauer, D. S., & Long, S. P. (2014). Light to liquid fuel: theoretical and realized energy conversion efficiency of plants using Crassulacean Acid Metabolism (CAM) in arid conditions. *Journal of Experimental Botany, 65*(13), 3471–3478. https://doi.org/10.1093/jxb/eru163

Davis, S. C., & Long, S. P. (2015). *Sisal/Agave* (V. M. V. Cruz & D. A. Dierig, Eds.). In (pp. 335–349). https://doi.org/10.1007/978-1-4939-1447-0_15

De Storme, N., & Geelen, D. (2014). Callose homeostasis at plasmodesmata: molecular regulators and developmental relevance. *Frontiers in Plant Science, 5*. https://doi.org/10.3389/fpls.2014.00138

Dellaert, S. N. C. (2014). Sustainability Assessment of the Production of Sisal Fiber in Brazil. (2009), 84.

Duarte, E. A. A., Damasceno, C. L., de Oliveira, T. A. S., Barbosa, L. de O., Martins, F. M., de Silva, J. R. Q., … Soares, A. C. F. (2018). Putting the mess in order: Aspergillus welwitschiae (and not A. niger) is the etiological agent of sisal bole rot disease in Brazil. *Frontiers in Microbiology, 9*(JUN).

https://doi.org/10.3389/fmicb.2018.01227

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology, 16*(1), 157.

https://doi.org/10.1186/s13059-015-0721-2

Ende, W. Van den. (2013). Multifunctional fructans and raffinose family oligosaccharides. *Frontiers in Plant Science, 4*. https://doi.org/10.3389/fpls.2013.00247

Falter, C., Zwikowics, C., Eggert, D., Blümke, A., Naumann, M., Wolff, K., … Voigt, C. A. (2015). Glucanocellulosic ethanol: the undiscovered biofuel potential in energy crops and marine biomass. *Scientific Reports, 5*(1), 13722. https://doi.org/10.1038/srep13722

FAO. (2020). Food and Agriculture Organization (2020) Statistical Database, Agriculture. Retrieved from faostat.fao.org

Flick, K., & Kaiser, P. (2012a). Protein degradation and the stress response. *Seminars in Cell & Developmental Biology, 23*(5), 515–522. https://doi.org/10.1016/j.semcdb.2012.01.019

Flick, K., & Kaiser, P. (2012b). Protein degradation and the stress response. *Seminars in Cell & Developmental Biology, 23*(5), 515–522. https://doi.org/10.1016/j.semcdb.2012.01.019

Fraser, C. M., & Chapple, C. (2011). The Phenylpropanoid Pathway in Arabidopsis. *The Arabidopsis Book, 9*, e0152. https://doi.org/10.1199/tab.0152

Freschi, L., Rodrigues, M. A., Tiné, M. A. S., & Mercier, H. (2010). Correlation between citric acid and nitrate metabolisms during CAM cycle in the atmospheric bromeliad Tillandsia pohliana. *Journal of
Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., … Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1), D1178–D1186. https://doi.org/10.1093/nar/gkr944

Goyal, K., Walton, L. J., & Tunnaccliffe, A. (2005). LEA proteins prevent protein aggregation due to water stress. *Biochemical Journal*, 388(1), 151–157. https://doi.org/10.1042/BJ20041931

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29, 644. Retrieved from https://doi.org/10.1038/nbt.1883

Gross, S. M., Martin, J. A., Simpson, J., Abraham-Juarez, M. J., Wang, Z., & Visel, A. (2013). De novo transcriptome assembly of drought tolerant CAM plants, Agave deserti and Agave tequilana. *BMC Genomics*, 14(1), 1–14. https://doi.org/10.1186/1471-2164-14-563

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Harkess, A., Zhou, J., Xu, C., Bowers, J. E., Van Der Hulst, R., Ayyampalayam, S., … Chen, G. (2017). The asparagus genome sheds light on the origin and evolution of a young y chromosome. *Nature Communications*, 8(1). https://doi.org/10.1038/s41467-017-0108-8

Huang, X., Wang, B., Xi, J., Zhang, Y., He, C., Zheng, J., … Yi, K. (2018). Transcriptome Comparison Reveals Distinct Selection Patterns in Domesticated and Wild Agave Species, the Important CAM Plants. *International Journal of Genomics*, 2018, 1–12. https://doi.org/10.1155/2018/5716518

Huang, X., Xiao, M., Xi, J., He, C., Zheng, J., Chen, H., … Yi, K. (2019). De Novo Transcriptome Assembly of Agave H11648 by Illumina Sequencing and Identification of Cellulose Synthase Genes in Agave Species. *Genes*, 10(2), 103. https://doi.org/10.3390/genes10020103

Huber, L., & Gillespie, T. J. (1992). Modeling Leaf Wetness in Relation to Plant Disease Epidemiology. *Annual Review of Phytopathology*, 30(1), 553–577.

Hudson, M. E., Lisch, D. R., & Quail, P. H. (2003). The FHY3 and FAR1 genes encode transposase-related proteins involved in regulation of gene expression by the phytochrome A-signaling pathway. *The Plant Journal*, 34(4), 453–471. https://doi.org/10.1046/j.1365-313X.2003.01741.x

Hundertmark, M., & Hincha, D. K. (2008). LEA (Late Embryogenesis Abundant) proteins and their encoding genes in Arabidopsis thaliana. *BMC Genomics*, 9(1), 118. https://doi.org/10.1186/1471-2164-9-118

Johansen, D. A. (1940). *Plant microtechnique*. New York, NY: McGraw and Hill Book Company.
Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution, 30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kitin, P., Voelker, S. L., Meinzer, F. C., Beeckman, H., Strauss, S. H., & Lachenbruch, B. (2010). Tyloses and Phenolic Deposits in Xylem Vessels Impede Water Transport in Low-Lignin Transgenic Poplars: A Study by Cryo-Fluorescence Microscopy. *Plant Physiology, 154*(2), 887–898. https://doi.org/10.1104/pp.110.156224

KÖPPEN, W., & GEIGER, R. (1936). Handbuch der Klimatologie. Gebr. Borntraeger, 44.

Kuo, T. M., VanMiddlesworth, J. F., & Wolf, W. J. (1988). Content of raffinose oligosaccharides and sucrose in various plant seeds. *Journal of Agricultural and Food Chemistry, 36*(1), 32–36. https://doi.org/10.1021/jf00079a008

Lacerda, T. M., de Paula, M. P., Zambon, M. D., & Frollini, E. (2012). Saccharification of Brazilian sisal pulp: Evaluating the impact of mercerization on non-hydrolyzed pulp and hydrolysis products. *Cellulose, 19*(2), 351–362. https://doi.org/10.1007/s10570-011-9638-4

Le Provost, G., Paiva, J., Pot, D., Brach, J., & Plomion, C. (2003). Seasonal variation in transcript accumulation in wood-forming tissues of maritime pine (Pinus pinaster Ait.) with emphasis on a cell wall glycin-rich protein. *Planta, 217*(5), 820–830. https://doi.org/10.1007/s00425-003-1051-2

Leão, D. A. S., Conceição, M. M., Conrado, L. S., Morais, C. R. S., Souza, A. G., Lima, C. S. S., … Silva, F. L. H. (2016). *Production of Energy—The Second Generation Ethanol and Prospects*. https://doi.org/10.1007/978-3-319-19767-8_9

Lepikson-Neto, J., Nascimento, L. C., Salazar, M. M., Camargo, E. L. O., Cairo, J. P. F., Teixeira, P. J., … Pereira, G. A. G. (2014). Flavonoid supplementation affects the expression of genes involved in cell wall formation and lignification metabolism and increases sugar content and saccharification in the fast-growing eucalyptus hybrid E. urophylla x E. grandis. *BMC Plant Biology, 14*(1), 1–17. https://doi.org/10.1186/s12870-014-0301-8

Li, L. (2003). OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research, 13*(9), 2178–2189. https://doi.org/10.1101/gr.1224503

Li, X., Bonawitz, N. D., Weng, J.-K., & Chapple, C. (2010). The Growth Reduction Associated with Repressed Lignin Biosynthesis in Arabidopsis thaliana Is Independent of Flavonoids. *The Plant Cell, 22*(5), 1620–1632. https://doi.org/10.1105/tpc.110.074161

Liang, J., Zhou, M., Zhou, X., Jin, Y., Xu, M., & Lin, J. (2014). JcLEA, a Novel LEA-Like Protein from Jatropha curcas, Confers a High Level of Tolerance to Dehydration and Salinity in Arabidopsis thaliana. *PLOS ONE, 8*(12), e83056. Retrieved from https://doi.org/10.1371/journal.pone.0083056

Librado, P., Vieira, F. G., & Rozas, J. (2012). BadiRate: estimating family turnover rates by likelihood-
based methods. *Bioinformatics*, 28(2), 279–281. https://doi.org/10.1093/bioinformatics/btr623

Lin, R., Ding, L., Casola, C., Ripoll, D. R., Feschotte, C., & Wang, H. (2007). Transposase-Derived Transcription Factors Regulate Light Signaling in Arabidopsis. *Science*, 318(5854), 1302–1305. https://doi.org/10.1126/science.1146281

Liu, Y., Chakrabortee, S., Li, R., Zheng, Y., & Tunnacliffe, A. (2011). Both plant and animal LEA proteins act as kinetic stabilisers of polyglutamine-dependent protein aggregation. *FEBS Letters*, 585(4), 630–634. https://doi.org/10.1016/j.febslet.2011.01.020

Livingston, D. P., Hincha, D. K., & Heyer, A. G. (2009). Fructan and its relationship to abiotic stress tolerance in plants. *Cellular and Molecular Life Sciences*, 66(13), 2007–2023. https://doi.org/10.1007/s00018-009-0002-x

Lyzenga, W. J., & Stone, S. L. (2012). Abiotic stress tolerance mediated by protein ubiquitination. *Journal of Experimental Botany*, 63(2), 599–616. https://doi.org/10.1093/jxb/err310

Mancilla-Margalli, N. A., & López, M. G. (2006). Water-Soluble Carbohydrates and Fructan Structure Patterns from Agave and Dasylirion Species. *Journal of Agricultural and Food Chemistry*, 54(20), 7832–7839. https://doi.org/10.1021/jf060354v

McDougall, G. J., Morrison, I. M., Stewart, D., Weyers, J. D. B., & Hillman, J. R. (1993). Plant fibres: Botany, chemistry and processing for industrial use. *Journal of the Science of Food and Agriculture*, 62(1), 1–20. https://doi.org/10.1002/jsfa.2740620102

McKain, M. R., Wickett, N., Zhang, Y., Ayyampalayam, S., McCombie, W. R., Chase, M. W., … Leebens-Mack, J. (2012). Phylogenomic analysis of transcriptome data elucidates co-occurrence of a paleopolyploid event and the origin of bimodal karyotypes in Agavoideae (Asparagaceae). *American Journal of Botany*, 99(2), 397–406. https://doi.org/10.3732/ajb.1100537

Medina, J. C. (1954). *O Sisal*. Secretaria da Agricultura do Estado de São Paulo.

Medina, J. C. (1959). *Plantas Fibrosas da Flora Mundial*. Instituto Agronômico de Campinas.

Menze, M. A., Boswell, L., Toner, M., & Hand, S. C. (2009). Occurrence of Mitochondria-targeted Late Embryogenesis Abundant (LEA) Gene in Animals Increases Organelle Resistance to Water Stress. *Journal of Biological Chemistry*, 284(16), 10714–10719. https://doi.org/10.1074/jbc.C900001200

Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, 7(1), 11257. https://doi.org/10.1038/ncomms11257

Michel-Cuello, C., Juárez-Flores, B. I., Aguirre-Rivera, J. R., & Pinos-Rodriguez, J. M. (2008). Quantitative Characterization of Nonstructural Carbohydrates of Mezcal Agave (Agave salmiana Otto ex Salm-Dick). *Journal of Agricultural and Food Chemistry*, 56(14), 5753–5757. https://doi.org/10.1021/jf800158p

Miedes, E., Vanholme, R., Boerjan, W., & Molina, A. (2014). The role of the secondary cell wall in plant
resistance to pathogens. *Frontiers in Plant Science*, 5. https://doi.org/10.3389/fpls.2014.00358

Miller, G. L. (1959). Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Analytical Chemistry*, 31(3), 426–428. https://doi.org/10.1021/ac60147a030

Min, K., Ha, S. C., Hasegawa, P. M., Bressan, R. A., Yun, D.-J., & Kim, K. K. (2003). Crystal structure of osmotin, a plant antifungal protein. *Proteins: Structure, Function, and Bioinformatics*, 54(1), 170–173. https://doi.org/10.1002/prot.10571

Morán, J. I., Álvarez, V. A., Cyras, V. P., & Vázquez, A. (2008). Extraction of cellulose and preparation of nanocellulose from sisal fibers. *Cellulose*, 15(1), 149–159. https://doi.org/10.1007/s10570-007-9145-9

Mukherjee, K., Brocchieri, L., & Burglin, T. R. (2009). A Comprehensive Classification and Evolutionary Analysis of Plant Homeobox Genes. *Molecular Biology and Evolution*, 26(12), 2775–2794. https://doi.org/10.1093/molbev/msp201

Mylsamy, K., & Rajendran, I. (2010). Investigation on Physio-chemical and Mechanical Properties of Raw and Alkali-treated Agave americana Fiber. *Journal of Reinforced Plastics and Composites*, 29(19), 2925–2935. https://doi.org/10.1177/0731684410362817

Nascimento, L. C., Yanagui, K., Jose, J., Camargo, E. L. O., Grassi, M. C. B., Cunha, C. P., … Mieczkowski, P. (2019). Unraveling the complex genome of Saccharum spontaneum using Polyploid Gene Assembler. *DNA Research*, 0(0), 1–12. https://doi.org/10.1093/dnares/dsz001

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, 32(1), 268–274. https://doi.org/10.1093/molbev/msu300

Nishizawa, A., Yabuta, Y., & Shigeoka, S. (2008). Galactinol and Raffinose Constitute a Novel Function to Protect Plants from Oxidative Damage. *Plant Physiology*, 147(3), 1251–1263. https://doi.org/10.1104/pp.108.122465

Nobel, P. S. (1988). *Environmental biology of agaves and cacti*. Cambridge University Press.

Nobel, P. S. (1994). *Remarkable Agaves and Cacti*. Oxford University Press.

Nobel, P. S. (2010). *Desert Wisdom/Agaves and Cacti: CO2, Water, Climate Change*. New York, NY: iUniverse.

Owen, N. A., Fahy, K. F., & Griffiths, H. (2016). Crassulacean acid metabolism (CAM) offers sustainable bioenergy production and resilience to climate change. *GCB Bioenergy*, 8(4), 737–749. https://doi.org/10.1111/gcbb.12272

Pecha, B., & Garcia-Perez, M. (2015). Pyrolysis of Lignocellulosic Biomass. In *Bioenergy* (pp. 413–442). https://doi.org/10.1016/B978-0-12-407909-0.00026-2

Pedrosa, A. M., Martins, C. de P. S., Gonçalves, L. P., & Costa, M. G. C. (2015). Late Embryogenesis Abundant (LEA) Constitutes a Large and Diverse Family of Proteins Involved in Development and
Abiotic Stress Responses in Sweet Orange (Citrus sinensis L. Osb.). *PLOS ONE, 10*(12), e0145785. https://doi.org/10.1371/journal.pone.0145785

Pimentel, H., Bray, N. L., Puente, S., Melsted, P., & Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature Methods, 14*(7), 687–690. https://doi.org/10.1038/nmeth.4324

Projeto MapBiomas. (2019). Projeto MapBiomas: Coleção 3.0 da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil. Retrieved from Projeto MapBiomas: Coleção 3.0 da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil. website: http://mapbiomas.org/

Qin, G., Xu, C., Ming, R., Tang, H., Guyot, R., Kramer, E. M., … Xu, Y. (2017). The pomegranate (Punica granatum L.) genome and the genomics of punicalagin biosynthesis. *The Plant Journal: For Cell and Molecular Biology, 91*(6), 1108–1128. https://doi.org/10.1111/tpj.13625

Ragauskas, A. J., Beckham, G. T., Biddy, M. J., Chandra, R., Chen, F., Davis, M. F., … Wyman, C. E. (2014). Lignin valorization: Improving lignin processing in the biorefinery. *Science, 344*(6185). https://doi.org/10.1126/science.1246843

Rathinasabapathi, B., Fouad, W. M., & Sigua, C. A. (2001). β-Alanine Betaine Synthesis in the Plumbaginaceae. Purification and Characterization of a Trifunctional, S -Adenosyl-l-Methionine-Dependent N -Methyltransferase from Limonium latifolium Leaves. *Plant Physiology, 126*(3), 1241–1249. https://doi.org/10.1104/pp.126.3.1241

Rotem, J., & Palti, J. (1969). Irrigation and Plant Diseases. *Annual Review of Phytopathology, 7*(1), 267–288. https://doi.org/10.1146/annurev.py.07.090169.001411

Rüggeberg, M., Speck, T., Paris, O., Lapiere, C., Pollet, B., Koch, G., & Burgert, I. (2008). Stiffness gradients in vascular bundles of the palm Washingtonia robusta. *Proceedings of the Royal Society B: Biological Sciences, 275*(1648), 2221–2229. https://doi.org/10.1098/rspb.2008.0531

Salazar, M. M., Nascimento, L. C., Camargo, E. L. O., Gonçalves, D. C., Neto, J. L., Marques, W. L., … Pereira, G. A. G. (2013). Xylem transcription profiles indicate potential metabolic responses for economically relevant characteristics of Eucalyptus species. *BMC Genomics, 14*(1), 201. https://doi.org/10.1186/1471-2164-14-201

Santos, J. D. G., Vieira, I. J. C., Braz-Filho, R., & Branco, A. (2015). Chemicals from agave sisalana biomass: Isolation and identification. *International Journal of Molecular Sciences, 16*(4), 8761–8771. https://doi.org/10.3390/ijms16048761

Sarwar, M. B., Ahmad, Z., Rashid, B., Hassan, S., Gregersen, P. L., Leyva, M. D. la O., … Husnain, T. (2019). De novo assembly of Agave sisalana transcriptome in response to drought stress provides insight into the tolerance mechanisms. *Scientific Reports, 9*(1), 396. https://doi.org/10.1038/s41598-018-35891-6
Sattari, M. T., Dodangeh, E., & Abraham, J. (2017). Estimation of daily soil temperature via data mining techniques in semi-arid climate conditions. *Earth Sciences Research Journal, 21*(2), 85–93. https://doi.org/10.15446/esrj.v21n2.49829

Schneider, R., Hanak, T., Persson, S., & Voigt, C. A. (2016). Cellulose and callose synthesis and organization in focus, what’s new? *Current Opinion in Plant Biology, 34*, 9–16. https://doi.org/10.1016/j.pbi.2016.07.007

Sengupta, S., Mukherjee, S., Basak, P., & Majumder, A. L. (2015). Significance of galactinol and raffinose family oligosaccharide synthesis in plants. *Frontiers in Plant Science, 6*(August), 1–11. https://doi.org/10.3389/fpls.2015.00656

Silva, O. R. R. F., & Beltrão, N. E. de M. (1999). *O agronegócio do sisal no Brasil*. Brasília, DF: embrapa.

Simmons, B. A., Loqué, D., & Ralph, J. (2010). Advances in modifying lignin for enhanced biofuel production. *Current Opinion in Plant Biology, 13*(3), 312–319. https://doi.org/10.1016/j.pbi.2010.03.001

Simpson, J., Martínez Hernández, A., Jazmín Abraham Juárez, M., Delgado Sandoval, S., Sánchez Villarreal, A., & Cortés Romero, C. (2011). Genomic resources and transcriptome mining in Agave tequilana. *GCB Bioenergy, 3*(1), 25–36. https://doi.org/10.1111/j.1757-1707.2010.01079.x

Singh, N. K., Bracker, C. A., Hasegawa, P. M., Handa, A. K., Buckel, S., Hermodson, M. A., … Bressan, R. A. (1987). Characterization of Osmotin. *Plant Physiology, 85*(2), 529–536. https://doi.org/10.1104/pp.85.2.529

Sluiter, J. B., Chum, H., Gomes, A. C., Tavares, R. P. A., Azevedo, V., Pimenta, M. T. B., … Mori, M. N. (2016). Evaluation of Brazilian Sugarcane Bagasse Characterization: An Interlaboratory Comparison Study. *Journal of AOAC International, 99*(3), 579–585. https://doi.org/10.5740/jaoacint.15-0063

Smith, A. M. (2008). Prospects for increasing starch and sucrose yields for bioethanol production. *The Plant Journal, 54*(4), 546–558. https://doi.org/10.1111/j.1365-313X.2008.03468.x

Somerville, C., Youngs, H., Taylor, C., Davis, S. C., & Long, S. P. (2010). Feedstocks for lignocellulosic biofuels. *Science, 329*(5993), 790–792. https://doi.org/10.1126/science.1189268

Stewart, J. R. (2015). Agave as a model CAM crop system for a warming and drying world. *Frontiers in Plant Science, 6*(September). https://doi.org/10.3389/fpls.2015.00684

Tamayo-Ordóñez, M. C., Rodríguez-Zapata, L. C., Narváez-Zapata, J. A., Tamayo-Ordóñez, Y. J., Ayil-Gutiérrez, B. A., Barredo-Pool, F., & Sánchez-Teyer, L. F. (2016). Morphological features of different polyploids for adaptation and molecular characterization of CC-NBS-LRR and LEA gene families in Agave L. *Journal of Plant Physiology, 195*, 80–94. https://doi.org/10.1016/j.jplph.2016.03.009

Tang, W., Wang, W., Chen, D., Ji, Q., Jing, Y., Wang, H., & Lin, R. (2012). Transposase-Derived Proteins FHY3/FAR1 Interact with PHYTOCHROME-INTERACTING FACTOR1 to Regulate Chlorophyll
Biosynthesis by Modulating HEMB1 during Deetiolation in Arabidopsis. *The Plant Cell*, 24(5), 1984–2000. https://doi.org/10.1105/tpc.112.097022

Tolleter, D., Hincha, D. K., & Macherel, D. (2010). A mitochondrial late embryogenesis abundant protein stabilizes model membranes in the dry state. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1798(10), 1926–1933. https://doi.org/10.1016/j.bbamem.2010.06.029

Törönen, P., Medlar, A., & Holm, L. (2018). PANNZER2: a rapid functional annotation web server. *Nucleic Acids Research*, 46(W1), W84–W88. https://doi.org/10.1093/nar/gky350

Treutter, D. (2006). Significance of flavonoids in plant resistance: a review. *Environmental Chemistry Letters*, 4(3), 147–157. https://doi.org/10.1007/s10311-006-0068-8

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., & Boerjan, W. (2010). Lignin Biosynthesis and Structure. *Plant Physiology*, 153(3), 895–905. https://doi.org/10.1104/pp.110.155119

Vanholme, R., Ralph, J., Akiyama, T., Lu, F., Pazo, J. R., Kim, H., … Boerjan, W. (2010). Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in Arabidopsis. *The Plant Journal*, 64(6), 885–897. https://doi.org/10.1111/j.1365-313X.2010.04353.x

Vieira, M. C., Heinze, T., Antonio-Cruz, R., & Mendoza-Martinez, A. M. (2002). Cellulose derivatives from cellulosic material isolated from Agave lechuguilla and fourcroydes. *Cellulose*, 9(2), 203–212. https://doi.org/10.1023/A:1020158128506

Wang, H., & Deng, X. W. (2002). Arabidopsis FHY3 defines a key phytochrome A signaling component directly interacting with its homologous partner FAR1. *The EMBO Journal*, 21(6), 1339–1349. https://doi.org/10.1093/emboj/21.6.1339

Yang, X., Cushman, J. C., Borland, A. M., Edwards, E. J., Wullschleger, S. D., Tuskan, G. A., … Holtum, J. A. M. (2015). A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in a hotter, drier world. *New Phytologist*, 207(3), 491–504. https://doi.org/10.1111/nph.13393

Yin, H., Guo, H.-B., Weston, D. J., Borland, A. M., Ranjan, P., Abraham, P. E., … Yang, X. (2018). Diel rewiring and positive selection of ancient plant proteins enabled evolution of CAM photosynthesis in Agave. *BMC Genomics*, 19(1), 588. https://doi.org/10.1186/s12864-018-4964-7

Zeng, Y., & Yang, T. (2002). RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. *Plant Molecular Biology Reporter*, 20(4), 417–417. https://doi.org/10.1007/BF02772130

Zhang, J., Zhang, X., Tang, H., Zhang, Q., Hua, X., Ma, X., … Ming, R. (2018). Allele-defined genome of the autopolyploid sugarcane Saccharum spontaneum L. *Nature Genetics*, 50(11), 1565–1573. https://doi.org/10.1038/s41588-018-0237-2
**Figure legends**

Figure 1. The three cultivars sampled at Embrapa’s germplasm bank and a summary of the methods used in this work. (a) *A. fourcroydes*; (b) *A. sisalana*; and (c) H11648. Using samples of three different individuals of these three cultivars, we correlated chemical and transcriptomic data to provide new insights into cell wall architecture, recalcitrance, and resistance to abiotic stresses.

Figure 2. (a) Venn diagram showing number of shared and exclusive gene families (nucleotides) between the assemblies of the three cultivars. (b) Number of exclusive genes per tissue and cultivars.

Figure 3. Significant (p-value < 0.05) Gene Ontology terms enriched in leaf (a), stem (b) and root (c) for the three cultivars. The dashed line shows the significance threshold adopted.

Figure 4. Gene expression of cellulose and associated carbohydrates pathways. (a) Schematic representation of the pathways, showing the coding genes for each enzyme. Red and grey indicate genes absent in our dataset and expressed at very low levels, respectively. (b) Heatmaps of log2 normalized values, representing the gene expression value in each tissue. AF: *A. fourcroydes*, AS: *A. sisalana*, HY: H11648.

Figure 5. Gene expression of the phenylpropanoids pathway. (a) illustrates the pathway, showing the flavonoids (green arrows) and lignin (blue arrows) branches. (b) Heatmaps of log2 normalized values, representing gene expression value in each tissue for the lignin pathway genes. Flavonoids pathway genes did not present a very high expression and are not shown here. AF: *A. fourcroydes*, AS: *A. sisalana*, HY: H11648.

Figure 6. Maximum likelihood phylogenetic inference for 10 plant species. *Amborella trichopoda* was used as outgroup. Tree was reconstructed using 978 single-copy ortholog groups in IQ-TREE with 1,000 bootstrap randomizations. Phylogeny is scaled in substitutions per positions.

Figure 7. (a) Soluble carbohydrates profile obtained by GCMS; (b) Estimate of the oligosaccharides and monosaccharides fractions; (c) Compositional analysis of Agave leaves and stem, values are shown in percentage; (d) Enzymatic hydrolysis of cell wall carbohydrates. AF: *A. fourcroydes*, AS: *A. sisalana* and HY: H11648.

Figure 8. Lignin and Pectic compounds distribution in *Agave* vascular bundle. (a), (b), and (c) are leaf cross-sections from *A. sisalana*, H11648, and *A. fourcroydes*, respectively, stained with ruthenium red for pectic compounds detection; (d), (e), and (f) are leaf cross-sections stained with phloroglucinol-HCl for lignin detection of *A. sisalana*, H11648, and *A. fourcroydes*, respectively; (x.) xylem; (ph.) phloem; and (f.c.) fiber cap cells. Scale bars = 100 µm
Table 1. The most highly expressed transcripts in each tissue for the three cultivars. The annotation was performed by Pannzer software and BLASTp against Uniref90. Uniref90’s annotations are shown only for Pannzer’s no hits.

| Genotype | Tissue | ID               | Pannzer2 annotation                        | Uniref90 blast annotation / E-value                                                                 | Mean expression (TPM) |
|----------|--------|------------------|--------------------------------------------|-----------------------------------------------------------------------------------------------------|-----------------------|
| AF       | Leaf   | AF_DN51229_c1_g1 | No hits                                    | dnaJ homolog subfamily B member 3-like isoform X2 n=1 Tax = Asparagus officinalis / 9.00E-66           | 50596                |
|          |        | AF_DN43092_c0_g1 | Late embryogenesis abundant protein, LEA-5  |                                                                                                      | 15624                |
|          |        | AF_DN49666_c0_g3 | Pentameric polyubiquitin                    |                                                                                                      | 7246                 |
|          |        | AF_DN51214_c2_g1 | Translationally-controlled tumor protein isogeny |                                                                                                      | 2602                 |
|          |        | AF_DN24667_c0_g1 | Phosphoenolpyruvate carboxylase (Fragment)   |                                                                                                      | 1991                 |
|          |        | AF_DN51074_c3_g1 | Zinc finger protein 1                       |                                                                                                      | 1888                 |
|          |        | AF_DN48042_c1_g1 | Copper chaperone                            |                                                                                                      | 1577                 |
|          |        | AF_DN53035_c4_g1 | No hits                                    |                                                                                                      | 1358                 |
|          |        | AF_DN53103_c2_g1 | Oxoglutarate/iron-dependent dioxygenase      |                                                                                                      | 1290                 |
|          |        | AF_DN50306_c4_g1 | Glyceraldehyde-3-phosphate dehydrogenase C2  |                                                                                                      | 1259                 |
|          | Stem   | AF_DN49666_c0_g3 | Pentameric polyubiquitin                    |                                                                                                      | 11725                |
|          |        | AF_DN51214_c2_g1 | Translationally-controlled tumor protein isogeny |                                                                                                      | 4148                 |
|          |        | AF_DN5961_c0_g1  | No hits                                    | Phloem protein 2-2 n=1 Tax= Apium graveolens Dulce Group / 3.00E-11                                | 3644                 |
|          |        | AF_DN51888_c1_g1 | No hits                                    | Uncharacterized protein n=1 Tax= Oryza meridionalis / 0                                            | 2439                 |
|          |        | AF_DN30460_c0_g1 | heat shock cognate 70 kDa protein-like       |                                                                                                      | 2344                 |
|          |        | AF_DN53587_c2_g1 | Elongation factor 1-alpha                  |                                                                                                      | 2336                 |
|          |        | AF_DN50408_c1_g1 | No hits                                    | protein BPS1, chloroplastic-like n=1 Tax= Phoenix dactylifera / 2.00E-48                           | 2219                 |
|          |        | AF_DN51229_c1_g1 | No hits                                    | dnaJ homolog subfamily B member 3-like isoform X2 n=1 Tax = Asparagus officinalis / 9.00E-66       | 2033                 |
|          |        | AF_DN42307_c0_g1 | No hits                                    |                                                                                                      | 2025                 |
|          | Root   | AF_DN49666_c0_g3 | Pentameric polyubiquitin                    |                                                                                                      | 20476                |
|          |        | AF_DN51888_c1_g1 | No hits                                    | Uncharacterized protein n=1 Tax = Oryza meridionalis / 0                                            | 3728                 |
|          |        | AF_DN43092_c0_g1 | Late embryogenesis abundant protein, LEA-5  5 |                                                                                                      | 3569                 |
|          |        | AF_DN30460_c0_g1 | heat shock cognate 70 kDa protein-like       |                                                                                                      | 3287                 |
|          |        | AF_DN54159_c4_g1 | Molecular chaperone DnaK                    |                                                                                                      | 2499                 |
|          |        | AF_DN51214_c2_g1 | Translationally-controlled tumor protein isogeny |                                                                                                      | 2296                 |
|          |        | AF_DN53841_c5_g1 | Luminal-binding protein 5                   |                                                                                                      | 2154                 |
|          |        | AF_DN51277_c1_g2 | No hits                                    | Uncharacterized protein (Fragment) n=1 Tax= Noccaea caerulescens / 6.00E-36                         | 1905                 |
|          |        | AF_DN53587_c2_g1 | Elongation factor 1-alpha                  | dnaJ homolog subfamily B member 3-like isoform X2 n=1 Tax = Asparagus officinalis / 9.00E-66       | 1694                 |
|          |        | AF_DN51229_c1_g1 | No hits                                    |                                                                                                      | 1696                 |
| Genotype | Tissue | ID               | Pannzer2 annotation                                      | Uniref90 blast annotation / E-value                                                                 | Mean expression (TPM) |
|----------|--------|------------------|----------------------------------------------------------|------------------------------------------------------------------------------------------------------|-----------------------|
|          | Leaf   | AS_DN59331_c3_g1 | Late embryogenesis abundant protein, LEA-5               | -                                                                                                    | 15948                 |
|          |        | AS_DN57028_c4_g2 | No hits                                                  | Uncharacterized protein n=1 Tax=Punica granatum / 7.00E-12                                          | 3897                  |
|          |        | AS_DN56755_c0_g1 | BTB and TAZ domain protein 1                             | -                                                                                                    | 3537                  |
|          |        | AS_DN59551_c3_g1 | Phosphophenolpyruvate carboxylase                        | -                                                                                                    | 3050                  |
|          |        | AS_DN57028_c4_g1 | No hits                                                  | Uncharacterized protein n=1 Tax=Punica granatum / 3.00E-11                                          | 2883                  |
|          |        | AS_DN50032_c0_g1 | Transmembrane protein, putative                          | -                                                                                                    | 2487                  |
|          |        | AS_DN56737_c2_g1 | Fiber polyubiquitin                                      | -                                                                                                    | 2204                  |
|          |        | AS_DN58072_c2_g2 | Senescence-specific cysteine protease                    | -                                                                                                    | 2160                  |
|          |        | AS_DN56737_c5_g1 | Ubiquitin supergroup                                    | -                                                                                                    | 2138                  |
|          |        | AS_DN57241_c2_g1 | ATP-dependent 6-phosphofructokinase                      | -                                                                                                    | 1902                  |
|          | Stem   | AS_DN56737_c5_g1 | Ubiquitin supergroup                                    | -                                                                                                    | 6468                  |
|          |        | AS_DN58948_c5_g1 | Heat shock cognate 70 kDa protein 2                      | -                                                                                                    | 4918                  |
|          |        | AS_DN59331_c3_g1 | Late embryogenesis abundant protein, LEA-5               | -                                                                                                    | 4393                  |
|          |        | AS_DN56737_c2_g1 | Fiber polyubiquitin                                      | -                                                                                                    | 3671                  |
|          |        | AS_DN56911_c0_g1 | No hits                                                  | No hits                                                | 3501                  |
|          |        | AS_DN46558_c0_g1 | S-adenosylmethionine synthase                            | -                                                                                                    | 2641                  |
|          |        | AS_DN51248_c1_g1 | No hits                                                  | No hits                                                | 2407                  |
|          |        | AS_DN50032_c0_g1 | Transmembrane protein, putative                          | -                                                                                                    | 2186                  |
|          |        | AS_DN58780_c3_g2 | Stress responsive NAC transcription factor               | -                                                                                                    | 2124                  |
|          |        | AS_DN53121_c0_g1 | Translationally-controlled tumor protein isogeny         | -                                                                                                    | 2062                  |
|          | Root   | AS_DN56737_c5_g1 | Ubiquitin supergroup                                    | -                                                                                                    | 8062                  |
|          |        | AS_DN58948_c5_g1 | Heat shock cognate 70 kDa protein 2                      | -                                                                                                    | 5371                  |
|          |        | AS_DN56737_c2_g1 | Fiber polyubiquitin                                      | -                                                                                                    | 3814                  |
|          |        | AS_DN56911_c0_g1 | No hits                                                  | No hits                                                | 3762                  |
|          |        | AS_DN46558_c0_g1 | S-adenosylmethionine synthase                            | -                                                                                                    | 3266                  |
|          |        | AS_DN59331_c3_g1 | Late embryogenesis abundant protein, LEA-5               | -                                                                                                    | 3233                  |
|          |        | AS_DN58780_c3_g2 | Stress responsive NAC transcription factor               | -                                                                                                    | 1920                  |
|          |        | AS_DN59592_c7_g1 | Heat shock protein 70 (Fragment)                         | -                                                                                                    | 1880                  |
|          |        | AS_DN53925_c1_g1 | Classical arabinogalactan protein 5                     | -                                                                                                    | 1659                  |
|          |        | AS_DN59334_c4_g1 | heat shock cognate protein 80                           | -                                                                                                    | 1596                  |
| Genotype | Tissue | ID                  | Pannzer2 annotation                                           | Uniref90 blast annotation / E-value                                      | Mean expression (TPM) |
|----------|--------|---------------------|--------------------------------------------------------------|------------------------------------------------------------------------|-----------------------|
| HY       | Leaf   | HY_DN37894_c0_g1    | Late embryogenesis abundant protein, LEA-5                   | -                                                                      | 15322                |
|          |        | HY_DN35134_c1_g1    | Cytosolic class I small heat shock protein type 1 (Fragment) | -                                                                      | 3247                 |
|          |        | HY_DN34855_c0_g1    | No hits                                                      | Uncharacterized protein (Fragment) n=1 Tax=Genlisea aurea / 7.00E-22    | 2582                 |
|          |        | HY_DN39094_c0_g1    | S-adenosylmethionine synthase                                | -                                                                      | 1505                 |
|          |        | HY_DN36990_c0_g1    | No hits                                                      | Uncharacterized protein n=1 Tax=Punica granatum / 1.00E-14              | 1437                 |
|          |        | HY_DN35295_c0_g1    | Glyceraldehyde-3-phosphate dehydrogenase (Phosphorylating)   | -                                                                      | 1390                 |
|          |        | HY_DN37711_c1_g1    | Translationally-controlled tumor protein isogeny             | ethylene-responsive transcription factor RAP2-4-like n=1 Tax=Ipomoea nil / 1.00E-11 | 1385                 |
|          |        | HY_DN38896_c4_g1    | No hits                                                      | -                                                                      | 1322                 |
|          |        | HY_DN35119_c3_g1    | No hits                                                      | No hits                                                                | 1262                 |
|          |        | HY_DN37863_c0_g1    | Salt tolerance zinc finger                                   | -                                                                      | 1176                 |
|          | Stem   | HY_DN37894_c0_g1    | Late embryogenesis abundant protein, LEA-5                   | -                                                                      | 9901                 |
|          |        | HY_DN35134_c1_g1    | Cytosolic class I small heat shock protein type 1 (Fragment) | -                                                                      | 6706                 |
|          |        | HY_DN34855_c0_g1    | No hits                                                      | Uncharacterized protein (Fragment) n=1 Tax=Genlisea aurea / 7.00E-22    | 5130                 |
|          |        | HY_DN39971_c0_g1    | No hits                                                      | No hits                                                                | 2169                 |
|          |        | HY_DN36990_c0_g1    | No hits                                                      | Uncharacterized protein n=1 Tax=Punica granatum / 1.00E-14              | 1804                 |
|          |        | HY_DN37711_c1_g1    | Translationally-controlled tumor protein isogeny             | -                                                                      | 1783                 |
|          |        | HY_DN33602_c0_g1    | Homocysteine S-methyltransferase                             | ethylene-responsive transcription factor RAP2-4-like n=1 Tax=Ipomoea nil / 1.00E-11 | 1196                 |
|          |        | HY_DN38896_c4_g1    | No hits                                                      | -                                                                      | 1184                 |
|          |        | HY_DN33896_c1_g1    | Histone H1 variant                                           | -                                                                      | 1181                 |
|          |        | HY_DN39118_c1_g1    | Chaperone protein dnaJ 2 (Fragment)                         | -                                                                      | 1179                 |
|          | Root   | HY_DN35134_c1_g1    | Cytosolic class I small heat shock protein type 1 (Fragment) | -                                                                      | 4874                 |
|          |        | HY_DN37894_c0_g1    | Late embryogenesis abundant protein, LEA-5                   | -                                                                      | 2773                 |
|          |        | HY_DN27224_c1_g2    | Ubiquitin C                                                 | -                                                                      | 1785                 |
|          |        | HY_DN39331_c4_g1    | Heat shock cognate 70 kDa protein 2                         | -                                                                      | 1742                 |
|          |        | HY_DN34855_c0_g1    | No hits                                                      | Uncharacterized protein (Fragment) n=1 Tax=Genlisea aurea / 7.00E-22    | 1666                 |
|          |        | HY_DN39094_c0_g1    | S-adenosylmethionine synthase                                | -                                                                      | 1660                 |
|          |        | HY_DN39450_c5_g1    | Heat shock protein 70 cognate                                | -                                                                      | 1587                 |
|          |        | HY_DN36990_c0_g1    | No hits                                                      | Uncharacterized protein n=1 Tax=Punica granatum / 1.00E-14              | 1461                 |
|          |        | HY_DN33375_c2_g1    | Heat shock cognate 70 kDa protein 2                         | -                                                                      | 1442                 |
|          |        | HY_DN30193_c0_g1    | SAM synthetase S-adenosylmethionine synthetase (Fragment)    | -                                                                      | 1431                 |
Table 2. Six most expressed tissue specific transcripts for the three cultivars. The annotation was performed by Pannzer software and BLASTp against Uniref90. Uniref90’s annotations are shown only for Pannzer’s no hits.

| Tissue | Genotype | ID | Pannzer2 annotation | Uniref90 blast annotation / E-value | Mean expression (TPM) | SPM |
|--------|----------|----|---------------------|-----------------------------------|------------------------|-----|
| Leaf   | AS       | HY_DN29394_c0_g1 | Ribulose bisphosphate carboxylase small chain | 354 | 0,98 |
|        | AS       | HY_DN34568_c0_g1 | Dihydrodipicolinate synthase | 332 | 0,96 |
|        | HY       | HY_DN38533_c1_g1 | Fructose-bisphosphate aldolase | 319 | 0,95 |
|        | HY       | HY_DN36770_c0_g1 | Alamine glyoxylate aminotransferase (Fragment) | 190 | 0,99 |
|        | HY       | HY_DN33146_c1_g1 | General substrate transporter | 172 | 0,98 |
|        | AS       | HY_DN47573_c0_g1 | Homeobox domain | 51 | 0,99 |
|        | AS       | HY_DN4042_c0_g1 | Aspartic proteinase nepenthesin-1 | 32 | 0,95 |
|        | AS       | HY_DN72852_c0_g1 | Peroxidase | 25 | 0,97 |
| Stem   | AS       | AS_DN54087_c0_g1 | Non-specific lipid-transfer protein 2 | 95 | 0,95 |
|        | AS       | AS_DN51290_c1_g1 | No hits | 53 | 0,99 |
|        | AS       | AS_DN2954_c0_g1 | Homeobox transcription factor KN3 (Fragment) | 27 | 1,00 |
|        | AS       | AS_DN47232_c0_g1 | Non-specific lipid-transfer protein 3 | 19 | 0,99 |
|        | AS       | AS_DN52004_c0_g1 | transcription factor LHW | 14 | 0,96 |
|        | AS       | AS_DN52843_c0_g1 | Remorin | 13 | 0,96 |
|        | HY       | HY_DN33608_c1_g1 | Basic helix-loop-helix transcription factor | 43 | 0,95 |
|        | HY       | HY_DN34445_c0_g1 | Class I KNOX homeobox transcription factor | 14 | 1,00 |
|        | HY       | HY_DN17938_c0_g1 | Transmembrane receptor | 12 | 1,00 |
|        | HY       | HY_DN43876_c0_g1 | Tyrosine kinase | 11 | 0,96 |
|        | HY       | HY_DN34723_c0_g1 | Beta-glucosidase | 9 | 0,96 |
|        | HY       | HY_DN32368_c0_g1 | General substrate transporter | 9 | 0,99 |
| Root   | AS       | AS_DN34326_c0_g1 | No hits | 383 | 1,00 |
|        | AS       | AS_DN34318_c1_g2 | Ubiquitin C | 148 | 1,00 |
|        | AS       | AS_DN36791_c0_g1 | No hits | 139 | 1,00 |
|        | AS       | AS_DN34989_c0_g1 | No hits | 383 | 1,00 |
|        | AS       | AS_DN39474_c0_g1 | No hits | 110 | 1,00 |
|        | AS       | AS_DN49042_c0_g1 | Heat shock cognate 70 kDa protein 2 | 96 | 1,00 |
|        | HY       | HY_DN32923_c0_g1 | No hits | 929 | 1,00 |
|        | HY       | HY_DN33213_c0_g1 | Thaumatin | 338 | 0,97 |
|        | HY       | HY_DN29718_c0_g1 | No hits | 145 | 1,00 |
|        | HY       | HY_DN38652_c1_g1 | Dihydrodipicolinate synthase | 124 | 1,00 |
|        | HY       | HY_DN33235_c0_g1 | No hits | 115 | 1,00 |
|        | HY       | HY_DN21465_c0_g1 | Early flowering protein 1 | 114 | 0,99 |
Figure 1

(d) Leaf and stem

Leaf, stem and root

Biomass composition
Enzymatic hydrolysis
Non-structural carbohydrates

Insights in drought resistance and cell wall architecture

RNA-Seq

Figure 2

(a) Venn diagram of gene expression in A. fourcroydes, A. sisalana, and H11648.

(b) Bar graph showing the number of genes in leaf, stem, root, and all tissues for A. fourcroydes, A. sisalana, and H11648.
Figure 3

a) cysteine biosynthetic process from serine
b) γ-aminobutyric acid biosynthetic process

A. fourcroydes
A. sisalana
H11648

-Log(p-value)

-Log(p-value)

-Log(p-value)

response to high light intensity
superoxide metabolic process
proteolysis
protein peptidyl-prolyl isomerase activity
dephosphorylation
glycerol ether metabolic process
photosynthesis

response to ethylene
response to UV-B
protein phosphorylation

response to jasmonic acid-mediated signaling pathway
response to ethylene
response to jasmonic acid-mediated signaling pathway
response to UV-B
protein phosphorylation

response to molecule of fungal origin
response to water
L-phenylalanine biosynthetic process
hydrogen peroxide catabolic process
induced systemic resistance jasmonic acid-mediated signaling pathway
S-adenosylmethionine biosynthetic process
translation
Figure 4

(a) Enzymatic pathway for the degradation of cell wall polysaccharides.

(b) Heatmap showing the expression levels of enzymes involved in the cell wall degradation pathway.
Figure 5

Figure 6
Figure 7

(a)

(b)

(c)

(d)

| Genotype   | Tissue | Extractives | Ash | Lignin | Cellulose | Hemicellulose |
|------------|--------|-------------|-----|--------|------------|---------------|
| *A. hypogaea* (Yang et al 2015) | Leaf   | 29.0 ± 1.2  | 6.0 ± 0.1 | 13.8 ± 1.3 | 7.0 ± 1.3 | 5.0 ± 1.2 |
| *A. hypogaea* | Leaf   | 56.9 ± 1.2  | 8.3 ± 0.7  | 7.3 ± 0.9  | 20.4 ± 1.5 | 7.5 ± 0.9  |
| *A. hypogaea* | Leaf   | 51.9 ± 1.6  | 11.3 ± 1.1 | 6.4 ± 0.5  | 20.3 ± 1.1 | 6.7 ± 0.4  |
| *H1548*    | Leaf   | 52.9 ± 1.6  | 9.2 ± 1.7  | 8.3 ± 0.5  | 20.9 ± 2.4 | 9.1 ± 0.3  |
| *A. hypogaea* | Stem   | 31.2 ± 1.1  | 3.9 ± 0.1  | 13.8 ± 3.2 | 12.2 ± 7.0 | 18.6 ± 3.5 |
| *A. hypogaea* | Stem   | 49.3 ± 1.5  | 5.2 ± 0.4  | 10.5 ± 3.1 | 21.6 ± 2.8 | 13.0 ± 2.8 |
| *H1548*    | Stem   | 51.6 ± 1.8  | 5.4 ± 0.6  | 10.5 ± 1.9 | 21.2 ± 2.8 | 11.3 ± 1.5 |
Figure 8