



## OPEN ACCESS

## EDITED BY

Sarah C. Davis,  
Ohio University, United States

## REVIEWED BY

Kelly Maribel Monja,  
Scientific Research Center of Yucatán  
(CICY), Mexico  
Fulgencio Alatorre-Cobos,  
Centro de Investigación Científica de  
Yucatán, Mexico

## \*CORRESPONDENCE

Rafael Vasconcelos Ribeiro,  
✉ rvr@unicamp.br  
Gonçalo Amarante Guimarães Pereira,  
✉ goncalo@unicamp.br

RECEIVED 07 May 2023

ACCEPTED 30 August 2023

PUBLISHED 20 September 2023

## CITATION

Raya FT, de Carvalho LM, José J,  
da Cruz LP, Almeida RL, Delevatti HA, da  
Silveira NM, da Silva SF, Pissolato MD,  
de Oliveira AB, dos Reis WJV,  
de Abreu LGF, Gutiérrez J, Carazzolle MF,  
Soares ACF, Nieto Sotelo J, Ribeiro RV  
and Pereira GAG (2023), Rescuing the  
Brazilian *Agave* breeding program:  
morphophysiological and molecular  
characterization of a new germplasm.  
*Front. Chem. Eng.* 5:1218668.  
doi: 10.3389/fceng.2023.1218668

## COPYRIGHT

© 2023 Raya, de Carvalho, José, da Cruz,  
Almeida, Delevatti, Silveira, da Silva,  
Pissolato, de Oliveira, dos Reis, de Abreu,  
Gutiérrez, Carazzolle, Soares, Nieto  
Sotelo, Ribeiro and Pereira. This is an  
open-access article distributed under the  
terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Rescuing the Brazilian *Agave* breeding program: morphophysiological and molecular characterization of a new germplasm

Fabio Trigo Raya<sup>1,2</sup>, Lucas Miguel de Carvalho<sup>2</sup>, Juliana José<sup>1</sup>, Larissa Prado da Cruz<sup>2</sup>, Rafael Leonardo Almeida<sup>2</sup>, Heliur Alves de Almeida Delevatti<sup>2</sup>, Neidiquele Maria Silveira<sup>2</sup>, Simone Ferreira da Silva<sup>2</sup>, Maria Dolores Pissolato<sup>2</sup>, Adriele Bárbara de Oliveira<sup>1</sup>, Wagner José Villela dos Reis<sup>1</sup>, Luís Guilherme Furlan de Abreu<sup>1</sup>, Jesús Gutiérrez<sup>3</sup>, Marcelo Falsarella Carazzolle<sup>1</sup>, Ana Cristina Fermino Soares<sup>4</sup>, Jorge Nieto Sotelo<sup>3</sup>, Rafael Vasconcelos Ribeiro<sup>2\*</sup> and Gonçalo Amarante Guimarães Pereira<sup>1\*</sup>

<sup>1</sup>Department of Genetics, Evolution, Microbiology and Immunology, Laboratory of Genomics and BioEnergy (LGE), University of Campinas (Unicamp), Campinas, Brazil, <sup>2</sup>Department of Plant Biology, Laboratory of Crop Physiology (LCroP), University of Campinas (Unicamp), Campinas, Brazil, <sup>3</sup>Botanical Garden, Institute of Biology, National Autonomous University of Mexico, Mexico City, Mexico, <sup>4</sup>Center of Agricultural, Environmental and Biological Sciences, Federal University of Recôncavo of Bahia (UFRB), Cruz das Almas, Brazil

*Agaves* have been a valuable resource in dryland areas for centuries, providing fibers (sisal), food, and beverages. However, the advent of synthetic fibers has led to a decrease in research on *Agave*, resulting in the cessation of breeding programs in Brazil. With the rise of climate change, there is renewed interest in *Agave* for its potential as a biofuel feedstock in semiarid regions. Since 2016, we have been collecting *Agave* accessions throughout the country and retrieving what is left of Brazil's original breeding program to establish a new germplasm bank. Here, we evaluated 21 of those accessions growing in the field. We used molecular markers and morphophysiological traits to characterize the plants. Based on the *Mayahuelin* molecular marker, we were able to reconstruct a phylogeny for the Brazilian accessions. The morphophysiological traits explained 34.6% of the phenotypic variation in the dataset, with physiological traits such as leaf water content, effective quantum efficiency of photosystem II ( $\Phi_{PSII}$ ), and specific leaf mass (SLM) as the most significant traits. Specifically, we evaluated nine *Agave* species and found that the physiological traits, rather than the morphological ones, were the most significant. Leaf water content was negatively correlated with specific leaf mass, which could be used as a marker for selecting cultivars with higher biomass accumulation. Interestingly,  $\Phi_{PSII}$  and chlorophyll content were negatively correlated, suggesting photochemical adaptations throughout the rosette. Molecular and phenotypic data suggest that *A. amaniensis*, which is frequently considered a synonym of *A. sisalana*, is effectively another species. Overall, this study provides valuable information on the physiological traits of Brazilian *Agave* accessions and is a starting point for

selecting more productive and climate-resilient cultivars for biorenewables production.

#### KEYWORDS

*Agave*, Biofuels, genetic resources, natural fibers, sisal

## 1 Introduction

*Agave* is a genus of succulent monocotyledonous plants in the family Asparagaceae, subfamily Agavoideae. They are native to the Americas, particularly to arid and semiarid regions from the southwestern United States to northern South America. Mexico is the main biodiversity center for this genus, where approximately 70% of the species can be found (Eguiarte et al., 2021). Also, *Agave* is an important crop in many parts of the world, particularly in arid and semiarid regions where other crops may not be suitable for cultivation. Agaves are used for various purposes, including the production of fibers, known as sisal, for rope, twine, and textiles, as well as to produce alcoholic beverages, such as tequila and mezcal (Raya et al., 2022). Apart from Mexico, Brazil and East Africa have been the most extensive regions for *Agave* plantations, especially for fiber production. Brazil imported its first *Agave sisalana* seedlings to the state of Bahia in 1903, and since the post-war period, it has become one of the leading exporters of sisal fiber (Medina, 1954; Silva and Beltrão, 1999). By the 1960s, Brazil became the largest producer of sisal fiber in the world, a position it holds to this day (FAO, 2023) despite the significant global and local decrease in production. In the 1990s, sisal fiber production in Brazil reached 300, 200 Mg annually (Silva and Beltrão, 1999). However, the current annual production has decreased to approximately 98,000 Mg, which accounts for roughly 36% of the world's annual sisal fiber production (FAO, 2023).

In Brazil, the *Agave* breeding program was led by the Agronomic Institute (IAC, Campinas) and its main objective was to obtain more productive fiber cultivars (Medina, 1954). To achieve this, the IAC introduced over seven species with fiber production potential in the country, including *A. amaniensis*, *A. cantala*, *A. letonae*, *A. zapupe*, and *A. fourcroydes* (Medina, 1954; Medina, 1956; Medina, 1959; Ciaramello et al., 1975). IAC was responsible for hybridization studies using *A. angustifolia* and *A. amaniensis* (Salgado et al., 1979; Azzini et al., 1989), generating lines that can be 10 times more productive than the common *A. sisalana* (Salgado et al., 1979), and, to this date, it remains the only institution that has developed Brazilian *Agave* varieties. However, the advent of synthetic fibers with more competitive prices, led to a decrease in sisal production in Brazil. Consequently, national investment in *Agave* research decreased dramatically, leading to the end of the country's only breeding program. A few of the materials developed by IAC were transferred to the germplasm bank of the Brazilian Agricultural Research Corporation—Embrapa, in Monteiro PB, Brazil (Suinaga et al., 2007; Souza et al., 2018; Monja-Mio et al., 2019; Raya et al., 2021), while some remained at IAC in Campinas.

Even though Brazil is still the world's largest sisal fiber producer, the lack of investment in the sector has brought profound consequences. As of today, Brazil's fiber productivity only ranks 12th globally, and it is five times lower than the leading country in the ranking (FAO, 2023). The sisal production system in Brazil

operates in a decentralized and semi-manual fiber extraction process (Silva and Beltrão, 1999; Broeren et al., 2017), which makes adoption of cultivars with higher fiber content, like the *Agave* hybrid 11648 [(*A. amaniensis* x *A. angustifolia*) x *A. amaniensis*], difficult. Farmers use undefined *A. sisalana* plants with almost no agricultural management (Azzini et al., 1989; Silva and Beltrão, 1999; Suinaga et al., 2007) and the rise of diseases, mainly sisal bole rot, is threatening their production (Duarte et al., 2018; Soares et al., 2020; Quintanilha-Peixoto et al., 2022).

Globally, agaveculture is currently undergoing a revolution. Agaves are emerging as a promising feedstock for biorefineries, particularly in regions with limited water supply (Escamilla-Treviño, 2012; Raya et al., 2022). Biorefineries are now able to produce a variety of products from agaves, including agavins (a probiotic-like inulin), syrup, pharmaceuticals, cosmetics, and nanocellulose (Morán et al., 2008; Laborde Aguirre et al., 2010; Alejandra et al., 2013; Laborde Aguirre et al., 2013; Gutierrez Uribe et al., 2017; Jean, 2015; Jeong et al., 2020). Furthermore, field trials are being conducted to explore the potential of agaves as a biofuel feedstock (Holtum et al., 2011; Davis et al., 2017; Yan et al., 2020; Parascanu et al., 2021). As a result, agaves are increasingly becoming a valuable natural resource for sustainable development and economic growth. With the emerging expectations for agaves as a source of biomolecules, fuels, and other value-added compounds in a world haunted by climate change (Corbin et al., 2015; Palomo-Briones et al., 2018; Parascanu et al., 2021; Raya et al., 2022), it is evident that there is a need to recover and restart the *Agave* breeding programs in Brazil and worldwide.

To achieve this, since 2016 we have been collecting *Agave* accessions throughout the country and retrieving what is left of Brazil's original breeding program to establish a new germplasm bank. Assessing long-cycle plants like *Agave* in a new collection presents challenges. Nonetheless, the integration of molecular and physiological approaches into an *Agave* breeding program surpasses traditional selection based solely on visual traits. It not only deepens our understanding of phenotypic diversity among various accessions but also has the potential to be incorporated into early selection strategies in due course (Mir et al., 2012; Furbank et al., 2019; Natarajan et al., 2019; Davis, 2022). Likewise, the utilization of molecular markers has played an important role in characterizing potential cultivars and comprehending genetic diversity in germplasm collections (Hasan et al., 2021). *Agave*, being a genus with limited available genetic resources, has recently seen the emergence of a novel marker based on the *Mayahuelin* gene (Lledías et al., 2020). This marker exhibits promise in deciphering phylogenetic relationships within the genus, particularly within the Rigidae section. In this study, we undertook the initial assessment of 21 potential breeding materials being grown in the field. We employed the *Mayahuelin* marker and examined 34 morphophysiological traits to acquire insights into the prevailing accessions in Brazil. The data

TABLE 1 Brazilian *Agave* accessions and their original collection sites.

Collection id	Taxon	Locality
CFL10	<i>A. sisalana</i>	Campo Formoso BA
VNA13	<i>A. sisalana</i>	Várzea Nova BA
VNM14	<i>A. sisalana</i>	Várzea Nova BA
LBM15	<i>A. sisalana</i>	Jacobina BA
CCA20	<i>A. sisalana</i>	Conceição do Coité BA
SGJ21	<i>A. sisalana</i> f. <i>marginata</i>	Conceição do Coité BA
IAC1	<i>A. amaniensis</i>	Campinas SP
IAC2	<i>A. sisalana</i> f. <i>marginata</i>	Campinas SP
IAC3	<i>A. amaniensis</i>	Campinas SP
IAC4	<i>Agave</i> sp	Campinas SP
IAC5	<i>A. sisalana</i> f. <i>marginata</i>	Campinas SP
LGE2	<i>A. sisalana</i>	Monteiro PB
LGE3	H11648 (( <i>A. amaniensis</i> x <i>A. angustifolia</i> ) x <i>A. amaniensis</i> )	Monteiro PB
LGE4	<i>A. attenuata</i>	Campinas SP
LGE5	<i>A. wercklei</i>	Campinas SP
LGE6	<i>A. americana</i>	São Paulo SP
LGE7	<i>A. angustifolia</i> var. <i>marginata</i>	Elias Fausto SP
LGE8	<i>A. desmettiana</i>	Cordeirópolis SP
LGE10	<i>A. tequilana</i>	Holambra SP
LGE16	<i>A. desmettiana</i>	Campinas SP
AFR	<i>A. fourcroydes</i> var. <i>remigio</i>	Monteiro PB

generated here not only highlight the primary factors that should guide the *Agave* breeding program in the coming years but also provide a deeper understanding of the crop's underlying biology.

## 2 Materials and methods

### 2.1 Plant materials and growth conditions

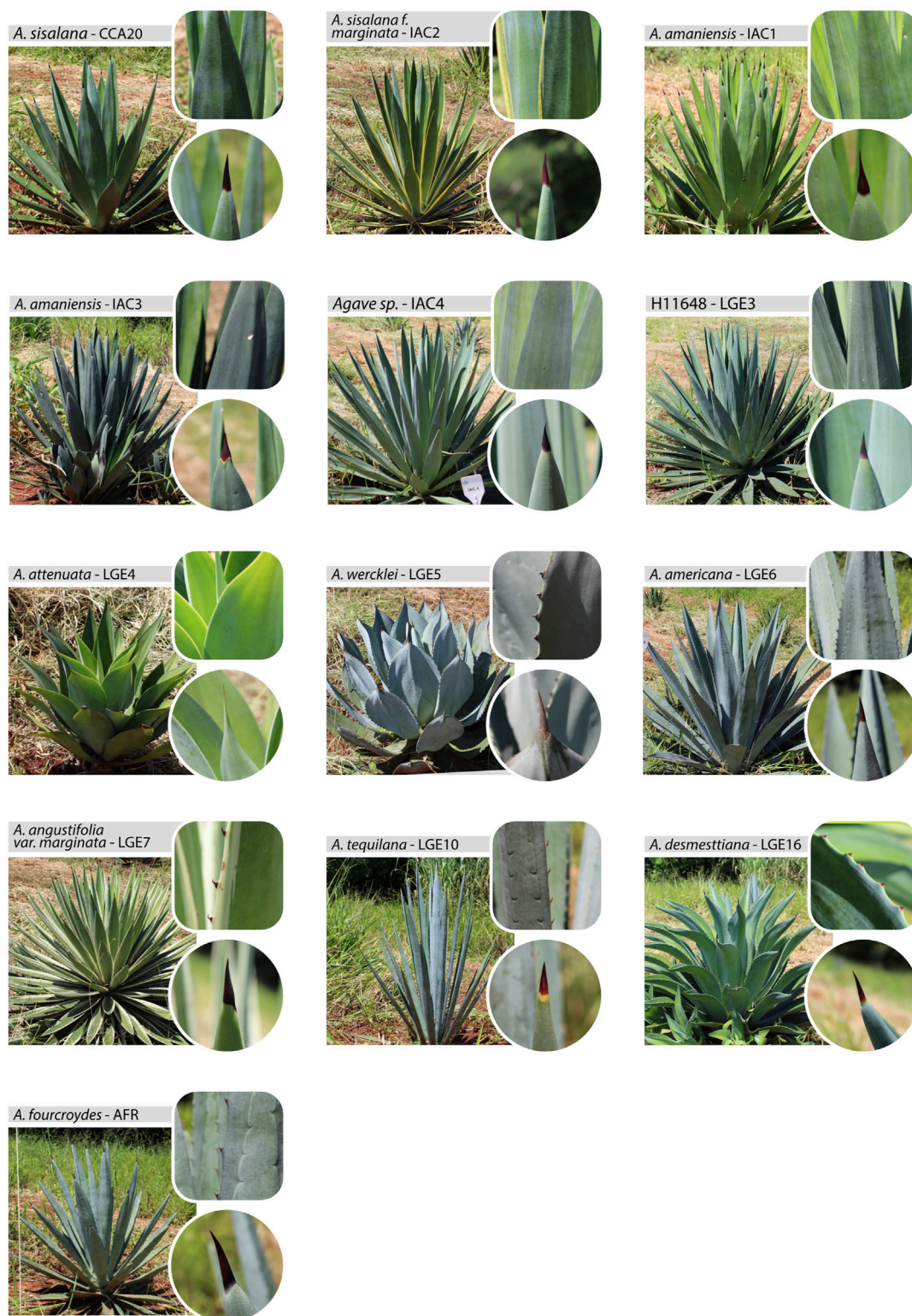
Twenty-one *Agave* accessions were sampled from the new germplasm bank of the University of Campinas (Unicamp, Campinas SP, Brazil), located at 22.8196°S, 47.0605°W, and 650 m a.s.l. The germplasm bank is composed of materials collected from the main sisal production areas in the State of Bahia, Brazil, plants recovered from the Agronomic Institute (IAC, Campinas SP, Brazil), and accessions donated by Brazilian Agricultural Research Corporation (Embrapa, Monteiro PB, Brazil), and by plant collectors. Table 1 summarizes the materials used in this study and their collection sites, while Figure 1 illustrates the phenotypic diversity within the germplasm accessions analyzed. Only materials growing in the field for 11 months were sampled. For each accession, two fully developed leaves of the same plant were analyzed, one youthful leaf from the internal part of the rosette (L1) and one mature leaf from middle part (L2) (Gentry, 1982; Lowman and Box, 1983; Luján et al., 2009). Five

biological replicates growing side-by-side were sampled per accession. Samples were collected in December 2021, at the beginning of the rainy season.

### 2.2 DNA extraction, *mayahuelin* sequencing, and phylogenetic inference

Genomic DNA was extracted from all accessions using the Sahu et al. (2012) protocol. For *Mayahuelin* gene amplification, primers MHL\_F2 (GTGAAATTTGAGGTCAACCT) and MHL\_R2 (TTGGTAGTTAACTTCAACGTT) were used (Ledías et al., 2020), generating approximately 730 bp amplicons. Polymerase chain reactions (PCR) were carried out using ECRA HIFI DNA Polymerase (ECRA biotec), in the following conditions: initial denaturation at 98°C for 35 s; 35 cycles of denaturation at 98°C for 10 s; annealing at 56°C for 30 s; extension at 72°C for 30 s; and final extension at 72°C for 10 min. Initially, the *Mayahuelin* amplicons were sequenced using the Sanger technique. As the sequences generated by this method were found to be unreliable (see Section 3.1 and Section 4.1), we prepared amplicon libraries using the same *Mayahuelin* primer set and sequenced them with an Illumina HiSeq2500, which generated 2 × 250 bp paired-end reads.

The reads for each amplicon were assembled into contigs using IDBA-UD (Peng et al., 2012) with the parameter "--similar 0.99". To



**FIGURE 1**

Phenotypic diversity within the germplasm accessions analyzed. As both *A. sisalana* and *A. desmestiana* have morphologically similar accessions, only one representative of each form was chosen. In the larger square, the rosettes of each representative individual can be observed, while in the smaller square and circle, the details of the leaf margins and apical spine are represented.

cluster contig sequences with high similarity, we used MMseqs2 software version 13.45111 (Steinberger and Söding, 2017) in easy-cluster mode, configured with "--min-seq-id 0.99".

Then, the representative sequence for each cluster was subjected to orthology tests against the known mayahuelin paralog sequences of *A. tequilana* (Lledías et al., 2020) (see also Supplementary Table S1).

TABLE 2 Morphophysiological traits evaluated.

Parameter	Description	Unit
ASL	Apical spine length	cm
C	Chromaticity	Unitless
Chl	Total chlorophyll content of fresh weight	mg g <sup>-1</sup> FW
DALS	Mean distance between last lateral and apical spines	cm
Edema	Edema [Present (Yes-Y) or Absent (No-N)]	Unitless
F <sub>v</sub> /F <sub>m</sub>	Maximum quantum efficiency of photosystem II	Unitless
°h	Hue angle	Unitless
CI	Color index calculated by °h/(L*xC)	Unitless
L*	Lightness	Unitless
LC	Leaf curvature [Plane (PL), Concave (CV), and Convex (CX)]	Unitless
LL	Leaf length	cm
LW	Leaf width	cm
LWR	Leaf length-to-width ratio	Unitless
MDBT	Minimum distance between teeth	mm
NPQ	Non-photochemical quenching of chlorophyll fluorescence	Unitless
qP	Photochemical quenching of chlorophyll fluorescence	Unitless
RD	Rhizome distance [Absent (A), Medium (M), Distant (D), and Medium and Distant (RMD)]	Unitless
Roughness	Roughness [Smooth (SM) Adaxial (AD), Abaxial (AB), and Both (BO)]	Unitless
SLM	Specific leaf mass	g DW cm <sup>-2</sup>
SUC	Succulence	mg H <sub>2</sub> O cm <sup>-2</sup>
TBW	Tooth basal width	mm
TP	Presence of teeth [Present (Yes-Y), Absent (No-N), or escape (E)]	Unitless
TL	Tooth length	mm
TN	Number of teeth in 10 cm	Unitless
V	Variation [Marginata (MAR) or Normal (NOR)]	Unitless
WC	Water content	%
Φ <sub>PSII</sub>	Effective quantum efficiency of photosystem II	Unitless
Ψ <sub>s</sub>	Osmotic potential	MPa

We used PhyML v3.0 (Guindon et al., 2010) to perform Maximum likelihood inference and selected only the sequences clustered with the reference mayahuelin (aC630\_3) for phylogenetic analysis.

To reconstruct the molecular phylogeny of *Agave*, the *Mayahuelin* nucleotide sequences of all Brazilian accessions were aligned together with 33 other *Mayahuelin* representatives previously described for Agavoideae (Lledías et al., 2020). The description for every sample used in the phylogenetic analysis is in Supplementary Table S2. Global alignment was done with MAFFT v7.271 (Yamada et al., 2006) with a maximum of 100 iterations using the L-INS-I algorithm, which is recommended for sequences with a conserved domain and long gaps (Katoh and Standley, 2013). The alignment was manually inspected, and the phylogeny was carried out both with a full

alignment and with a trimmed version. The best-fit nucleotide substitution model was tested using IQtree v2.1.2 (Minh et al., 2020) and chosen using the Akaike criteria (Akaike, 1974). The nucleotide substitution model was tested both for considering one model for the whole alignment and for different models for each position of the codons. Maximum likelihood phylogenies were inferred using IQtree with the selected substitution model and 1000 bootstraps to test branch support.

## 2.3 Morphophysiological characterization

The morphophysiological parameters evaluated to characterize the *Agave* accessions are listed in Table 2. First, all categorical traits

(RD, Edema, LC, Roughness, Teeth, and V) were obtained from plants in the field. Then, the leaves L1 and L2 were harvested and the quantitative morphological parameters (ASL, DALs, LL, LWR, MDBT, TBW, TL, TN, and LW) were measured. Chlorophyll fluorescence emission of such leaves was measured with a modulated fluorometer PAM-2000 (Heinz Walz GmbH, Effeltrich, Germany) according to Almeida et al. (2021), and the photochemical parameters were estimated following Maxwell and Johnson (2000). Under room temperature,  $\Phi_{PSII}$ , qP and NPQ were obtained from light adapted leaves ( $\lambda = 710$  nm; PAR  $\sim 15,000$   $\mu\text{mol m}^{-2} \text{s}^{-1}$ ; 0.8 s), and  $F_v/F_m$  from leaves acclimated to artificial darkness for 30 min.

The leaf colorimetric analysis was performed with the chroma meter CR-300 (Konica Minolta, Japan) after a white calibration. The parameters analyzed in the middle portion of leaf blade were lightness ( $L^*$ , zero = black and 100 = white),  $a^*$  (−80 to zero = green and from zero to +100 = red), and  $b^*$  (−100 to zero = blue and zero to +70 = yellow). Such data were used to estimate the hue angle ( $h$ ), the chromaticity (C) and the color index [ $\text{hlc} = h/(L^* \times C)$ ], following McGuire (1992). Leaf discs were collected and incubated at room temperature with 3 mL of 80% acetone for 8 days in the dark. Chlorophyll a, b and total were quantified using the method of Arnon (1949) modified by the suppression of the steps of discs grinding and centrifugation.

From each leaf, five discs of 0.5 cm<sup>2</sup> (1.5 cm<sup>2</sup> of leaf area, LA) were extracted and weighed immediately for the fresh weight (FW). Such leaf discs were then submerged in distilled water for 24 h under dark and cool conditions ( $\sim 4^\circ\text{C}$ ) and weighted again for the turgid weight (TW). Then, the leaf discs were dried for 24 h at 60°C to obtain its dry weight (DW). The values obtained were used to calculate water content (WC), succulence (SUC), and specific leaf mass (SLM) as described below:

$$WC = \frac{(FW - DW)}{FW} \times 100 \quad (1)$$

$$SUC = \frac{(TW - DW)}{LA} \quad (2)$$

$$SLM = \frac{DW}{LA} \quad (3)$$

The leaf osmotic potential ( $\Psi_s$ ) was estimated using a Wescor HR-33T dewpoint microvoltmeter using the thermocouple psychrometer mode of operation (Wescor Inc., Logan, United States) connected to C-52 sample chambers. Frozen leaf discs (1 cm<sup>2</sup>) were placed inside the chamber and, after a 1 h period of equilibration,  $\Psi_s$  was measured.

## 2.4 Data integration and analysis

To comprehensively analyze the phenotypic variation across the *Agave* accessions, several statistical analyses were performed, with the R software, version 4.1.0 (R Core Team, 2022). The Shapiro-Wilk test was used with a significance level of 0.05 to test for normal distribution of all quantitative variables. The Mann-Whitney *U* test was conducted to compare the quantitative variables of all plants and determine if there was a significant difference between leaves (variables named with suffix “1” for L1 and “2” for L2). Additionally, the Kruskal–Wallis test was performed to determine

whether there was a significant difference in a qualitative variable across all plants, followed by the Dunn *post hoc* test using *rstatix* v0.7.2 package (Kassambara, 2023) to further examine any significant differences found.

To cluster the quantitative and qualitative variables, a factor analysis of mixed data (FAMD) was performed using the FactoMineR package (Lê et al., 2008). The analysis included 15 quantitative variables and six qualitative variables, with all contributions for all dimensions calculated using the factoextra package (Kassambara and Mundt, 2020). To examine the relationships between the variables, a correlation matrix was created using the *corrplot* v0.92 (Wei et al., 2021) package, with a significance level of 0.05 and results visualized using a color-coded matrix.

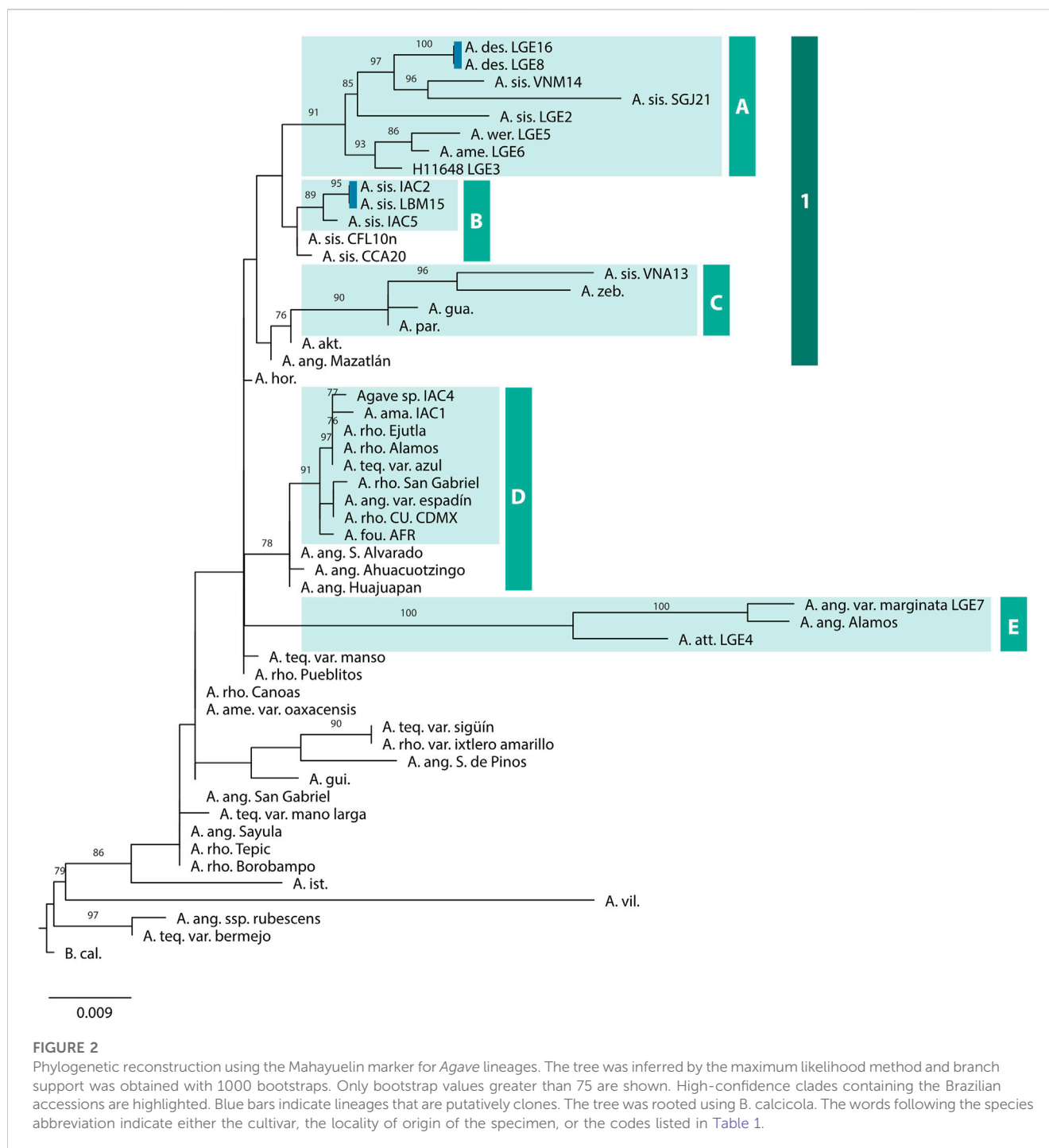
A dendrogram analysis was conducted to examine the relationships between the *Agave* accessions using quantitative biometrical and physiological features. The hierarchical clustering function in the *stats* package (R Core Team, 2022) was used to cluster the plants based on their similarities. The resulting dendrogram was generated using the *dist* and *hclust* function and visualized using the *dendextend* package (Galili, 2015). The dendrogram was colored to indicate the different clusters of plants identified in the analysis, with the distance cutoff used to define the clusters determined using the cophenetic correlation coefficient.

Only the results of the most relevant characteristics for FAMD or correlation analysis are presented in the following section, while additional characteristics can be found in the Supplementary Figure S3 and Supplementary Tables S3, S4.

## 3 Results

### 3.1 Phylogenetic reconstruction of the Brazilian *Agave* accessions based on *Mayahuelin* DNA sequences

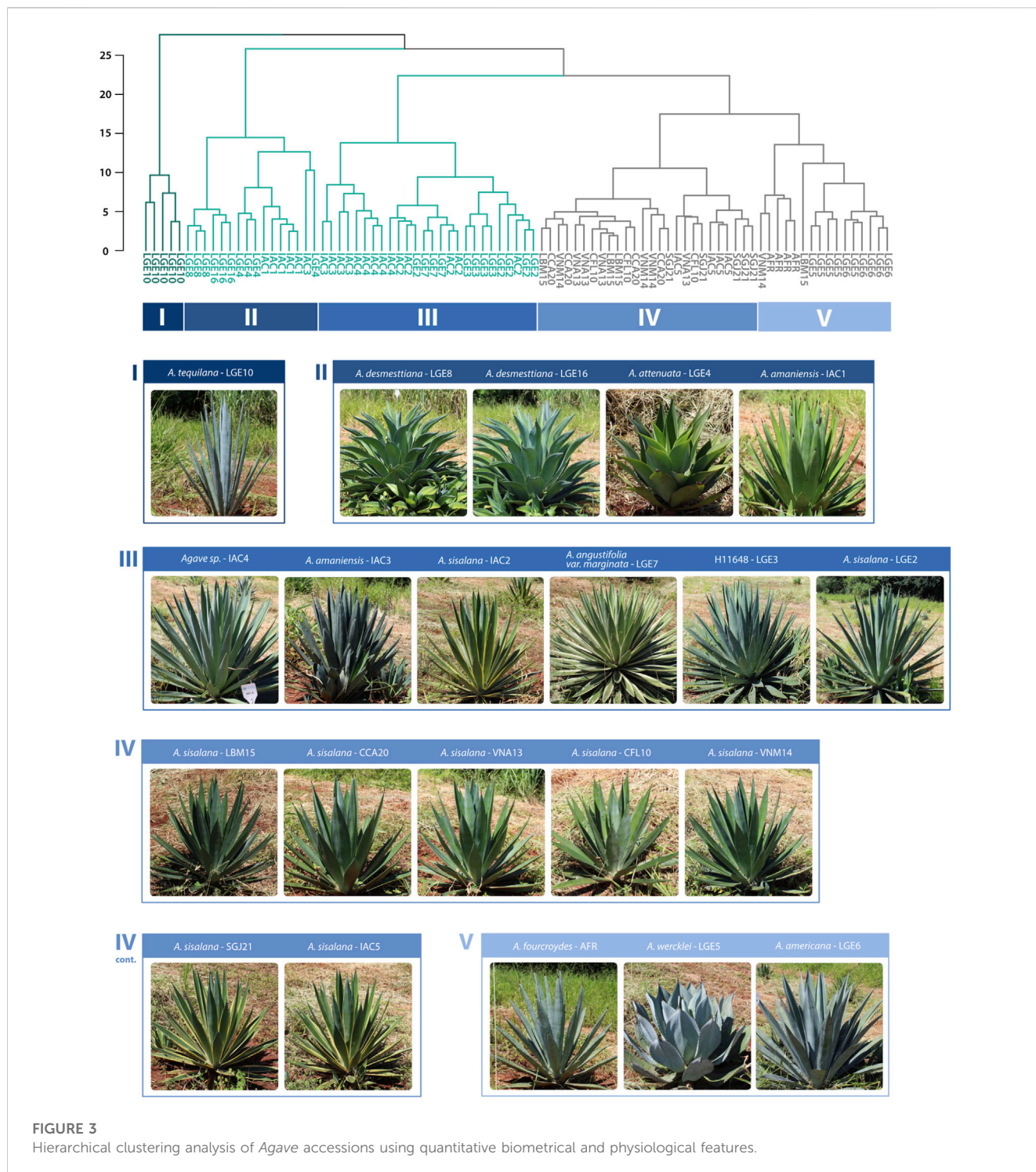
We encountered several challenges when attempting to apply the *Mayahuelin* molecular marker to our dataset. The Sanger sequencing method yielded low-quality sequences and was prone to contamination by paralogous genes, leading to imprecise and unreliable results that had to be discarded. To overcome these issues, we attempted an alternative approach by isolating RNA from the central spike and conducting RT-PCR. However, our efforts to detect the expression of *Mayahuelin* in *A. sisalana* proved to be unsuccessful. Finally, we used a next-generation sequencing strategy to isolate the *Mayahuelin* gene in our samples. For all the samples, we found the amplification of *Mayahuelin* paralogs. Among them, RIP36 (GAHU01014938.1) and RIP37 (MT015974.1) were the most common, but we also discovered several potential new RIPs that require further exploration. RIP36 and RIP37 are novel RIPs identified by transcriptome and genome data mining of *A. tequilana* var. *azul* and other *Agave* species (Gutiérrez and Nieto-Sotelo, unpublished data). Notably, we were unable to detect the *Mayahuelin* ortholog (aC630\_3; MT015973.1) among the assemblies for *A. tequilana* LGE10 and *A. amaniensis* IAC3, due to the significant enrichment of RIP36. Therefore, we were not able to analyze the phylogenetic relationship of these accessions.



We reconstructed three phylogenies: one using the full alignment and a nucleotide substitution model fitted for the whole alignment and a nucleotide substitution model fitted for the whole alignment, and a third one using the full alignment and fitting a nucleotide substitution for each codon position. The tree showing higher branch supports was the first one and it is presented herein (Figure 2). The other two phylogenies are shown as (Supplementary Figure S1). The selected phylogeny was similar to the previous *Mayahuelin* tree reported by Lledías et al. (2020). Five high-confidence (>85 bootstraps) clades were identified for the

Brazilian accessions, and two putative sets of clones were identified (Figure 2). These sets include *A. desmettiana* LGE8 and LGE16, as well as *A. sisalana* IAC2 and LBM15. As expected *Beschorneria calcicola* was resolved as a separate lineage.

The *A. sisalana* samples were resolved into a single lineage (clade 1) but were distributed across three groups (A, B, and C). Group A was composed of *A. desmettiana* LGE16, *A. desmettiana* LGE8, *A. sisalana* VNM14, *A. sisalana* SGJ21, *A. sisalana* LGE2, *A. wercklei* LGE5, *A. americana* LGE6, and H11648 LGE3. Group B includes *A. sisalana* IAC2, *A. sisalana* LBM15, and *A. sisalana* IAC5, while Group C was composed of *A. sisalana* VNA13, *A. zebra* (AG8876),



*A. guadalajarana* (JNS58), and *A. parryi* (AG4968). The other fiber-producing cultivars were grouped in the high-confidence Group D, which includes *Agave* sp. IAC4, *A. amaniensis* IAC1, *A. rhodacantha* Ejutla (AG10959), *A. rhodacantha* Alamos (AG9173), *A. tequilana* var. *azul*, *A. rhodacantha* San Gabriel (JNS40), *A. angustifolia* var. *espadin* (JNS63), *A. rhodacantha* CU CDMX (JE06), and *A. fourcroydes* AFR. Finally, the last high-confidence group, E, was composed of *A. angustifolia* var. *marginata* LGE7, *A. angustifolia* Alamos (DS1682), and *A. attenuata* LGE4.

### 3.2 Phenotypic variation of the morphophysiological traits

Factor analysis of mixed data (FAMD) revealed that our selected morphophysiological traits (Table 2) explained 34.6% of the phenotypic variation in the dataset (Supplementary Figure S2), with colorimetry (chromaticity, C; hue angle,  $^{\circ}h$ ; and color index, CI), water content (WC), roughness, leaf length-to-width ratio (LWR), effective quantum efficiency of



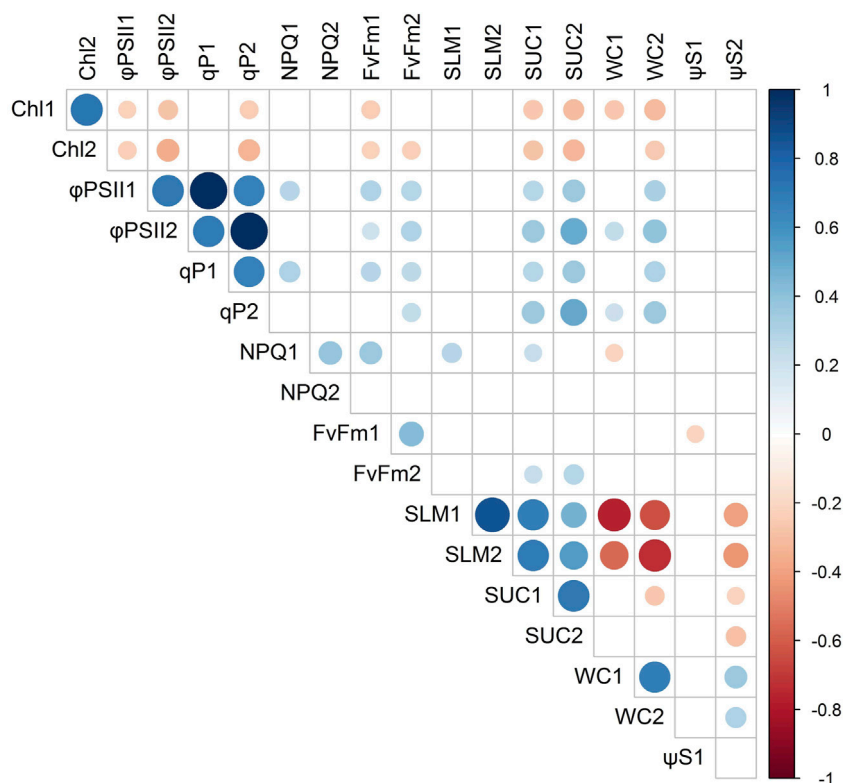


FIGURE 4

Correlation matrix of physiological traits in *Agave* with a significance level of 0.05. See Table 2 for abbreviation of physiological traits.

PSII ( $\Phi_{PSII}$ ), and specific leaf mass (SLM) as the most significant traits.

The dendrogram obtained from the hierarchical clustering analysis using quantitative features (biometrical and physiological) presented a clear separation of five major clusters (Figure 3), revealing distinct phenotypic characteristics among the accessions. Cluster I was composed solely of *A. tequilana* LGE10. In Cluster II, we found *A. amaniensis* IAC1, *A. attenuata* LGE4, and both *A. desmettiana* accessions. In this cluster, it is still possible to observe that despite being very close accessions, LGE8 and LGE16 do not mix in the grouping. In Cluster III, two larger subdivisions can be observed: one composed of *A. amaniensis* IAC3 and *Agave* sp. IAC4, and the other where the samples of *A. sisalana* LGE2, H11648 LGE3, *A. angustifolia* LGE7, and *A. sisalana* IAC2 are mixed. In Cluster IV, only *A. sisalana* from Bahia were found, except for *A. sisalana* IAC5. In Cluster V, predominantly samples of *A. fourcroydes* AFR, *A. wercklei* LGE5, and *A. americana* LGE6 are present, with LGE5 and LGE6 being closer together.

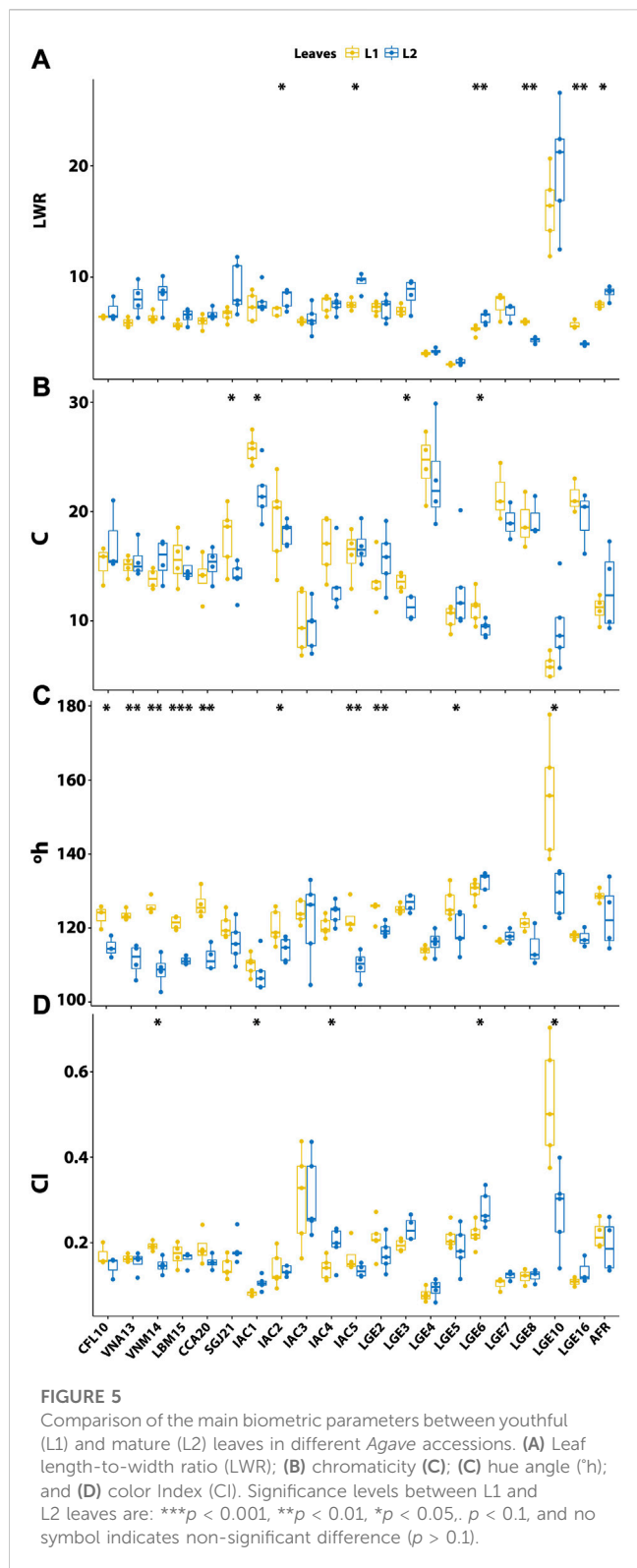
Correlation analysis was performed to investigate the relationship among physiological traits in the studied accessions (Figure 4). Except for  $\Psi_s$ , there was a positive correlation for all characteristics between L1 and L2.  $\Phi_{PSII}$  showed a significant positive correlation with photochemical quenching (qP) ( $r = 0.99$ ,  $p < 0.05$ ) but was negatively correlated with total chlorophyll content (Chl) ( $r = -0.36$ ,  $p < 0.05$ ). SLM presented a significant positive correlation with succulency (SUC) ( $r = 0.64$ ,  $p < 0.05$ ) and negative correlation with WC ( $r = -0.69$ ,  $p < 0.05$ ),

indicating that plants with higher leaf mass per area should have more SUC and that the higher the leaf density the lower the capacity to store water.

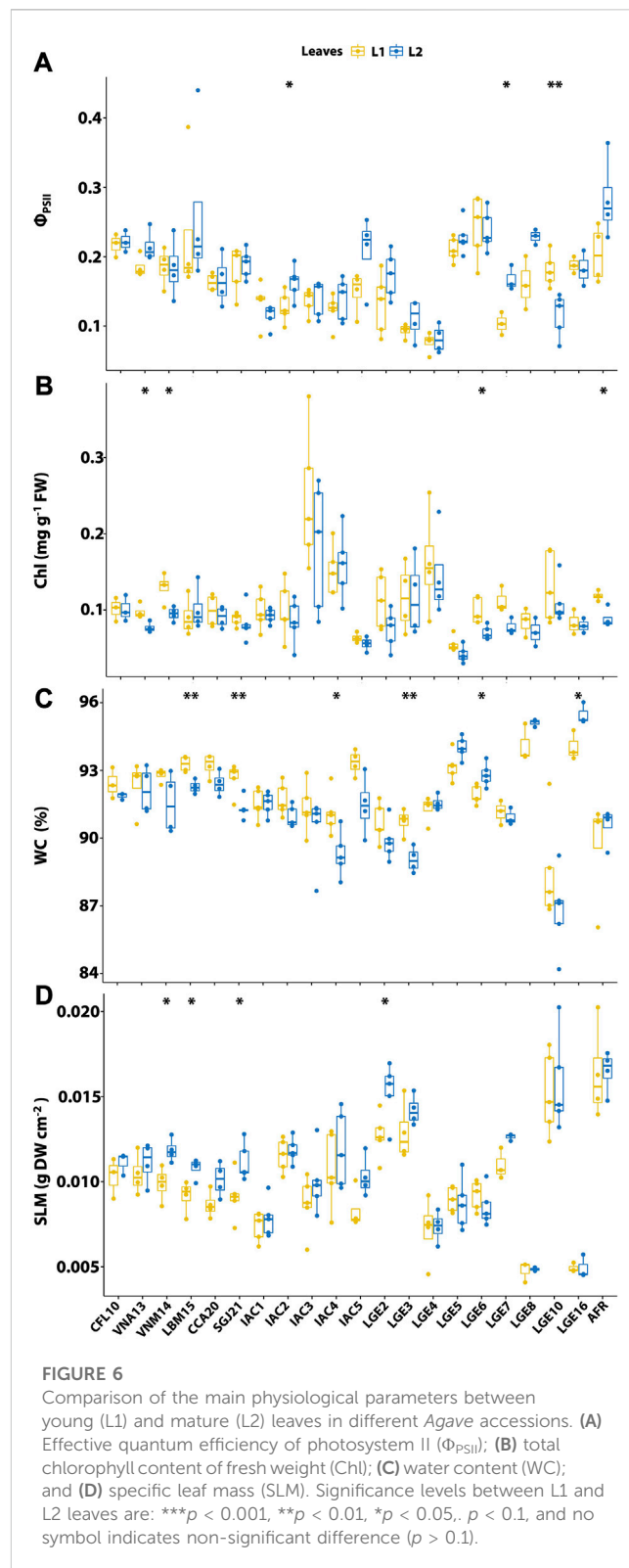
### 3.2.1 Biometry

Except for the *A. desmettiana* materials (LGE8 and LGE16), all accessions had a lower length-to-width ratio (LWR) in L1 (Figure 5A). Among these, *A. tequilana* LGE10 had the highest LWR, while *A. wercklei* LGE5 had the lowest ratio. In general, materials developed for fiber production, such as *A. sisalana*, *A. fourcroydes*, *A. amaniensis*, and H11648 had intermediate LWR values without significant differences, according to the Dunn's test. This indicates a clear selection for longer and wider leaves from which well-developed fibers can be obtained (Medina, 1954; Raya et al., 2021). Although the majority of accessions lacked teeth, approximately 13% of the materials displayed "Escape" (Supplementary Table S3), a trait marked by sporadic residual teeth appearing irregularly along one of the leaf margins. Leaf texture or roughness was important in characterizing the accessions and was the only categorical feature to emerge as one of the most influential traits in FAMD.

Phenotypical diversity among the accessions was highlighted by colorimetry. When comparing within the rosette, a significant difference in chromaticity was found between L1 and L2 only in four materials (SGJ21, IAC1, LGE3, and LGE6; Figure 5B), where L1 presented greater chromaticity in all cases. The Dunn's test indicated no differences in C1 between the accessions of *A. sisalana*, but the two accessions of *A. amaniensis* (IAC1 and



IAC3) showed contrasting phenotypes; IAC1 had the highest chromaticity among the entire dataset while IAC3 had one of the lowest. Other noteworthy differences in C were observed for *A. tequilana* LGE10, which had the lowest value, and *A. attenuata* LGE4, which had one of the highest values. The hue angle was one of the traits that showed most significant differences between L1 and



L2 (Figure 5C), with L1 being greater than L2 in all cases. Among the *A. sisalana* samples, only SGJ21 did not show significant differences between L1 and L2. In contrast to chromaticity, *A. tequilana* LGE10 had the highest hue angle, while *A. amaniensis* IAC1 had the lowest. Finally, CI also exhibited differences within the rosette (Figure 5D), with *A. sisalana* VNM14 and *A. tequilana*

LGE10 showing higher values in L1, while *A. amaniensis* IAC1, *Agave* sp. IAC4, and *A. americana* LGE6 were significantly higher in L2. Considering L1, *A. tequilana* LGE10 showed the highest CI and *A. amaniensis* IAC1 and *A. attenuata* LGE4 were the lowest. *A. tequilana* LGE10 was significantly higher than *A. sisalana* VNA13, *A. sisalana* SGJ21, *A. amaniensis* IAC, *A. sisalana* IAC2, *Agave* sp. IAC4, *A. attenuata* LGE4, and both *A. desmettiana*. No significant interactions for CI was found among *A. sisalana*, but *A. sisalana* LGE2 and H11648 LGE3 were higher than *A. amaniensis* IAC1. For L2, fewer significant interactions were found; however, *A. amaniensis* IAC1 and *A. attenuata* LGE4 remained the lowest values within the dataset.

### 3.2.2 Photochemical activity

There was a great variation in  $\Phi_{PSII}$  among *Agave* accessions (Figure 6A), with values ranging from 0.07 for *A. attenuata* LGE4 to 0.3 for *A. fourcroydes* AFR. However, only tree accessions displayed significant differences between L1 and L2. For these, *A. sisalana* IAC2 and *A. angustifolia* LGE7 showed higher  $\Phi_{PSII}$  in L2, while *A. tequilana* LGE10 showed higher  $\Phi_{PSII}$  in L1. Among the accessions for industrial applications, H11648 showed the lowest  $\Phi_{PSII}$  values, being significantly lower when compared to *A. sisalana* (CFL10, VNA13, VNM14, LBM15, SGJ21, and IAC5), *A. tequilana* LGE10, and *A. fourcroydes* AFR according to the Dunn's test. The most significant differences in  $\Phi_{PSII}$  were found in L2. In this leaf, *A. amaniensis* (IAC1 and IAC3), and *Agave* sp. IAC4 were less efficient than *A. fourcroydes* AFR, *A. wercklei* LGE5, *A. americana* LGE6, *A. desmettiana* LGE8, *A. sisalana* CFL10, and *A. sisalana* IAC5.

Significant variation in total chlorophyll content (Chl) was observed between L1 and L2 (Figure 6B) in only four accessions: *A. sisalana* VNA13, *A. sisalana* VNM14, *A. americana* LGE6, and *A. fourcroydes* AFR. In all cases, L1 had higher Chl. In contrast to  $\Phi_{PSII}$ , *A. amaniensis* IAC3 had the highest Chl. Compared to the average in *A. sisalana* accessions, Chl in *A. amaniensis* IAC3 was approximately 2.5 times higher. According to the Dunn's test, this accession was significantly higher than *A. sisalana* LBM15, SGJ21, and IAC5, both *A. desmettiana*, and *A. wercklei* LGE5. Among the accessions with lower Chl, we can highlight *A. sisalana* IAC5 and *A. wercklei* LGE5, with significantly lower Chl than almost all other accessions. It should be noted that IAC5 was the only *A. sisalana* accession that showed significant differences compared to the others, specifically in the comparison against VNM14 in L1.

### 3.2.3 Leaf water content and specific leaf mass

All the accessions had a WC greater than 85%, with *A. desmettiana* having the highest at 96% (Figure 6C). The WC varied within the *Agave* rosette, and six accessions showed a significant difference between L1 and L2. The materials developed for fiber, such as *A. sisalana* LBM15, *A. sisalana* SGJ21, *Agave* sp. IAC4, and H11648 LGE3, had higher WC in L1, while the accessions *A. americana* LGE6 and *A. desmettiana* LGE16 had higher WC in L2. According to the Dunn's test, there was a significant difference between the *A. sisalana* accessions and *Agave* sp. IAC4, H11648 LGE3, and *A. tequilana* LGE10 in both L1 and L2. Additionally, in L1, differences were found between the accessions of *A. sisalana* LBM15, CCA20, and IAC5, and *A. fourcroydes* AFR, and *A. sisalana* LGE2. The accessions of *A.*

*amaniensis* (IAC1 and IAC3) and *Agave* sp. IAC4 were similar, being significantly lower than *A. desmettiana* (LGE8, and LGE16) and *A. wercklei* LGE5, and were higher than *A. tequilana* LGE10. Finally, *A. fourcroydes* AFR and *A. tequilana* LGE10 had the lowest WC and were not statistically different.

In contrast to WC, *A. tequilana* LGE10 and *A. fourcroydes* AFR had the highest SLM, while both *A. desmettiana* accessions had the lowest (Figure 6D). Statistical differences regarding SLM were only found within L1 and L2 for the *A. sisalana* accessions (VNM14, LBM15, SGJ21, and LGE2), with the SLM being higher in L2. Compared to *A. amaniensis* IAC1 and IAC3 in L1, *A. tequilana* LGE10, *A. fourcroydes* AFR, *A. sisalana* LGE2, and H11648 LGE3 were significantly higher. While in L2, *A. amaniensis* IAC3 had smaller SLM only in comparison to *A. tequilana* LGE10 and *A. fourcroydes* AFR. Additionally, *A. sisalana* LGE2 and H11648 LGE3 showed higher SLM than *A. attenuata* LGE4, *A. wercklei* LGE5, *A. americana* LGE6, and both *A. desmettiana*.

## 4 Discussion

### 4.1 Using *Mayahuelin* to trace the history of *Agave* accessions in Brazil

*Mayahuelin* is a type I ribosome-inactivating protein (RIP) that was first identified in the central spike of *A. tequilana* var. *azul* (Lledías et al., 2020). Initially, the *Mayahuelin* gene was identified as a promising molecular marker, exhibiting good amplification across the *Agave* genus and even *Beschorneria*. Further analysis revealed that *Mayahuelin* possessed sufficient resolution to address evolutionary questions in *Agave* (Lledías et al., 2020). However, we encountered several obstacles while attempting to utilize this molecular marker, especially in *A. sisalana*. Paralog amplification and lack of expression of *Mayahuelin* jeopardized the use of traditional strategies. To solve this issue, we employed a next-generation sequencing approach, which enabled us to successfully isolate the *Mayahuelin* gene in our samples.

The limited applicability of the *Mayahuelin* marker in *A. sisalana* may be attributed to the origin of this species. *A. sisalana* is an irregular pentaploid ( $2n = 5x = 149$ ) (Queiroz et al., 2012) and believed to be a hybrid, possibly between *A. angustifolia* and *A. kewensis* (Gentry, 1982), although this has not been yet confirmed. This complex genome composition increases the likelihood of off-target amplification during PCR, and since there is no control over which copy of the gene will be amplified due to polyploidy, it can result in varying outcomes in phylogenetic inference. Our analysis demonstrated this phenomenon, as *A. sisalana* samples were placed in three groups, with only one group composed exclusively of *A. sisalana* samples (Figure 2). Interestingly, our *A. sisalana* samples were not grouped according to their collection site, indicating a possible exchange of materials between the sisal-producing regions in Bahia. In the phylogeny, it is possible to identify two independent origins of variegation in *A. sisalana*, one from *A. sisalana* VNM14 and SGJ21 and another from IAC2 and IAC5 with LBM15. Medina (1955) first identified the marginata form of *A. sisalana* in Araraquara SP Brazil in 1939 and it is plausible that IAC2 and

IAC5 are descendants of that plant. Additionally, IAC2, IAC5, and LBM15 exhibit a prominent deformation on the margins of the leaf abaxial face, appearing as a localized roughness (Supplementary Figure S4). This reinforces the hypothesis that these materials are closely related. While this deformation can also be observed in other *A. sisalana* plants, it is usually not as evident.

Among the groups with *A. sisalana* samples, group A had the most Brazilian accessions and displayed two clear subdivisions. The first subdivision grouped the two samples of *A. desmettiana* with the samples of *A. sisalana*. Originally, Sisalanae was considered a subsection within Rigidae (Berger, 1915). Later, Gentry (1982) classified Sisalanae as its own group and included *A. desmettiana* and *A. kewensis* in it. In a recent study, Jiménez-Barron et al. (2020) used the ITS region to infer the phylogeny of Agavoideae, placing *A. sisalana* close to *A. tequilana*, but with insufficient resolution to classify *A. desmettiana*. According to our data, *A. desmettiana* is closely related to *A. sisalana*. Specifically, *A. sisalana* accessions LGE2, VNM14, and SGJ21 in group A exhibit a ten-nucleotide gap at 666bp in the *Mayahuelin* gene, resulting in a premature stop codon. The trimmed sequence has been deposited at NCBI, while the complete sequence of these three accessions can be found in the Supplementary Data S1. This sequence gap is likely the main reason why these *A. sisalana* samples were not included in group B. Furthermore, this may be one of the reasons why it was not possible to isolate *Mayahuelin* from RNA in *A. sisalana* samples.

The second subdivision includes samples of H11648, *A. americana*, and *A. wercklei*. Perhaps this is an indicative that Sisalanae might be related to the *A. americana* group. Xu et al. (2022) used a limited set of chloroplast genomes and also found H11648 to be close to *A. americana*. Except for the fact that it is a thornless plant (Gentry, 1982), there is hardly any information available about the *A. angustifolia* used to generate H11648. Therefore, the possibility of this parent being of hybrid origin or even misidentified cannot be ruled out. With regards to the proximity of *A. americana* and *A. wercklei* in the phylogeny, Gentry (1982) reported that Trelease's specimens of *A. wercklei* from Costa Rica exhibited characteristics that were consistent with the Americanae group, which corroborates our findings. Given the various biological processes of *Agave*, such as their long life cycle and the easiness of hybridization with more distantly related groups, it was expected that issues regarding resolution and agreement between phylogenies would arise (Eguiarte et al., 2021), especially possible contradictions between markers originated either from nuclear or plastid genes (Rothfels, 2021).

*A. amaniensis* is a species of unknown origin that was discovered in an abandoned area of the East Africa Agricultural Research Station in Amani, Tanzania in 1929 (Nowell, 1933; Medina, 1954). Known as Blue Sisal, *A. amaniensis* became popular locally due to its fast growth and high fiber yield (Medina, 1954; Ciaramello et al., 1975). *A. amaniensis* was used in breeding by hybridization with *A. angustifolia* (Salgado et al., 1979) and is the parental line of the most widely cultivated hybrid cultivar in the world, H11648 (Monja-Mio et al., 2019). Although Blue Sisal was once considered a variation of *A. sisalana* and later classified as a synonym (WFO, 2023), our data and recent studies indicate that *A. amaniensis* is a distinct species. We initially identified two IAC materials as *A. amaniensis* (IAC1 and IAC3). According to the *Mayahuelin* phylogenetic inference (Figure 2), the IAC1 and

IAC4 samples were grouped in the same clade as *A. tequilana* var. *azul*, *A. rhodachanta* Ejutla, and *A. rhodachanta* Alamos, and not with any of our *A. sisalana* samples. IAC4 may be a remnant line of the IAC hybridization experiments conducted in 1958 (Salgado et al., 1979). Also, *A. amaniensis* is diploid ( $2n = 60$ ) unlike *A. sisalana* (Doughty, 1936; Queiroz et al., 2012), and recent chloroplast analysis has also shown that these two species are not the same (Xu et al., 2022). From a morphophysiological perspective, the IAC1, IAC3, and IAC4 samples showed contrasting behaviors when compared to *A. sisalana*, particularly in terms of chlorophyll content (Figure 6B). Thus, our data indicates that a taxonomic classification review of *A. amaniensis* is necessary.

*A. fourcroydes* is the most used *Agave* species in Mexico for fiber production, and it was first introduced to Brazil in 1906 (Shcolz, 1959). Despite its superior productivity and robustness when compared to *A. sisalana* (Ciaramello et al., 1975; Suinaga et al., 2007), it has never been grown on a commercial scale in Brazil due to the presence of teeth (Medina, 1954; Silva and Beltrão, 1999; Raya et al., 2021). In Mexico, *A. fourcroydes* has been reported to have at least five traditional cultivars (Colunga-García Marín et al., 1999) and variable ploidy (Robert et al., 2008). Nevertheless, in Brazil, only two closely related accessions, "Remigio" and "Cabinho", have been classified (Souza et al., 2018), although it is not clear how they relate to the Mexican cultivars. In our phylogenetic inference, the *A. fourcroydes* sample was grouped with *A. angustifolia*, a species, that is, supposed to be its wild ancestor (Colunga-García Marín et al., 1999), but was closer to *A. rhodachanta* samples, highlighting the intertwined nature of the *A. angustifolia* complex (Rivera-Lugo et al., 2018; Lledías et al., 2020). Curiously, we observed that *A. angustifolia* var. *marginata*, a highly widespread variety across the globe (Gentry, 1982), is very closely related to the *A. angustifolia* from Alamos, Southern Sonora. Finally, it is important to highlight that polytomy is extremely common in *Agave*, and relying solely on morphological characteristics for evaluation, especially from the vegetative phase, often leads to misidentification.

## 4.2 Morphophysiological implications to understand *Agave* complexities

Although we analyzed 34 traits in two leaf types, the percentage of phenotypic variation explained by our data may not be considered high compared to other studies (Morillo-Coronado et al., 2021; Basile et al., 2022). However, this is to be expected when working with a diverse germplasm. Despite being a highly complex genus, *Agave* has often been treated agronomically as a single crop (Long et al., 2015; Davis et al., 2017; Monja-Mio et al., 2019; Raya et al., 2022). Specifically, we evaluated nine *Agave* species and found that the physiological traits, rather than the morphological ones, were the most significant. Moreover, we observed significant physiological differences within the plant rosette. As the *Agave* leaves mature, they undergo a transition from C3 to CAM (Crassulacean Acid Metabolism) photosynthesis (Nobel, 1988; Luján et al., 2009; Abraham et al., 2016). In our dataset, we were able to observe some physiological gradient; youthful leaves present higher water and chlorophyll content and lower photochemical activity (Figure 6 and Supplementary Figure S3). The only exception was *A. tequilana* LGE10, which presented significantly lower  $\Phi_{PSII}$  and  $qP$  in the mature leaf.

Interestingly, chlorophyll content was negatively correlated with  $\Phi_{\text{PSII}}$ . Positive correlation between Chl and  $\Phi_{\text{PSII}}$ , despite being common (Tremblay et al., 2012; Singh et al., 2017), is not a rule. The relationship between these two traits is complex, and a negative correlation has already been reported in rice and barley (Li et al., 2006; Peng et al., 2021). Excess chlorophyll can enhance light capture and cause energetic pressure at PSII level, which can lead to an increase in the production of reactive oxygen species (ROS), turning plants more prone to photooxidative stress and damage to PSII (Tanaka and Tanaka, 2006). In *Agave*, excessive energy dissipation strategies have already been proven to be a major adaptation factor under high light conditions (Campos et al., 2014; Aragón-Gastélum et al., 2020). Since youthful leaves are not completely CAM (Abraham et al., 2016), they likely have higher photosynthetic rates and higher light saturation point. Under such conditions, plants are benefited from extra chlorophyll content and light capture. Another hypothesis is that accessions that accumulate less Chl could use nutrients, such as nitrogen, for the synthesis of other proteins involved in CO<sub>2</sub> fixation rather than in photochemistry.

Water accumulation is one of the main drought resistance mechanisms in *Agave* (Nobel, 1988). However, excessive WC may not be a desirable trait for biomass production, since it increases financial and energy costs related to transport and processing (Lewandowski and Heinz, 2003; Agger et al., 2014). In our dataset, WC ranged from 85% to 96% (Figure 6C), which is significantly higher than the WC in the whole biomass of other energy crops, such as sugarcane (60%–80%) and *Miscanthus* (20%–40%) (Lewandowski and Heinz, 2003; Dinh et al., 2019). In an environment where water is a clear limiting factor, the water captured by agaves can be of great importance (Raya et al., 2022) and still not be a limiting factor for other industrial processes, such as biogas production. However, considering that we found up to 96% of WC, it is possible to optimize the *Agave* water status to obtain plants with higher biomass density. Our analysis revealed a significant negative correlation between WC and SLM (Figure 4). Herein, SLM is a trait that reflects the amount of dry matter present per unit of leaf area. Considering that leaves make up 72%–83% of the dry weight in *Agave* plants (Nobel, 1988), SLM may be a useful metric for screening accessions with less water content and potential higher-quality biomass, which can be more easily converted to value-added products. Interestingly, we found that among the materials analyzed here, those commonly used for industrial purposes, such as *A. sisalana* LGE2, H11648 LGE3, *A. tequilana* LGE10, and *A. fourcroydes* AFR, have higher SLM than less domesticated agaves, such as *A. attenuata* LGE4, *A. wercklei* LGE5, *A. americana* LGE6, and *A. desmettiana* LGE8, and LGE16. Contrary to what one might expect, the materials developed for fibers that presented higher SLM were equivalent to *A. tequilana*. It is possible that higher SLM values correlate with superior photosynthate accumulation capacity leading to higher stem biomass, a favorable trait in *A. tequilana* for tequila production. However, this correlation remains to be assessed in the future. Moreover, higher SLM values in fiber-producing cultivars could have been selected to produce more commercial fiber, which is long or extra-long fibers, not higher total fiber content (Medina, 1954).

Although we observed some occasional variations among the accessions of *A. sisalana*, the samples from the producing regions of

Bahia showed homogeneous behavior in this initial analysis (Figure 3), which was expected as the materials spread around the world and introduced in Brazil originated from the clones selected by Perrine in 1833 (Trejo-Torres et al., 2018). However, it is noteworthy that the *A. sisalana* originating in germplasm banks, such as *A. sisalana* LGE2 and IAC2, did not cluster with the accessions from the producing regions, which presented an apparently higher  $\Phi_{\text{PSII}}$  (Figure 6A). This separation could be a consequence of constant selection pressure applied by farmers or the environment. The cluster analysis even suggests that the samples from Bahia are more similar to *A. wercklei* LGE5, *A. americana* LGE6, and *A. fourcroydes* AFR than those of germplasm origin.

*A. tequilana* was the sample that differed the most from the others, as it was placed alone in a single cluster. This species has been considered one of the most productive *Agave*, with a relatively short cycle, high soluble solid content (up to 38 °Brix), and bioethanol yield that can be compatible with sugarcane (Valenzuela-Zapata and Nabhan, 2004; Gil-Vega et al., 2006; Garcia-Moya et al., 2011; Yan et al., 2020; Raya et al., 2022). The phenotypic contrast that we observed might help us to understanding the agronomical performance of this species, such as the higher photochemical activity in the youthful leaf instead of the mature one. However, it is important to emphasize that the accession of *A. tequilana* analyzed in this study, according to the nursery, was cultivated from seed and, therefore, may have different performance from *A. tequilana* var. *azul* clones cultivated in Mexico and abroad. Also, this *A. tequilana* accession did not have a better  $\Phi_{\text{PSII}}$  performance than *A. sisalana* or *A. fourcroydes*. It is clear that *Agave* growth patterns can change according to the species, and biomass production usually escalates from the third year onward (Nobel, 1988; Yan et al., 2020). However, a thorough comparative study on the physiological changes during their growth phases has not been conducted yet among important cultivars. While it is possible that the *A. tequilana* metabolism undergoes significant changes over time, we should not overlook the possibility that *A. sisalana* and *A. fourcroydes* have been selected or adapted to the Brazilian climate conditions in the nearly 120 years since their introduction.

#### 4.3 Perspectives for *Agave* research and breeding

The great diversity and complexity of the genus *Agave* is a factor that greatly expands the applications of this crop, but it brings some challenges. Most agaves are monocarpic and some species can take more than 50 years to reach maturity (Eguiarte et al., 2021), which limits significantly the possibilities of crossing cycles and progeny evaluation. One way to increase the efficiency of this process is through marker-assisted selection. However, there is limited information on functional genomics in *Agave* and no reference genome was published yet (Raya et al., 2022). Therefore, the development of reliable molecular markers is still hindered, as the case with the use of *Mayahuelin*. Although this locus shows promise, further in-depth studies are needed to identify the most divergent regions in relation to its paralogs. Limited available genomic information makes the development of primers with low off-target chances more challenging. The development of a user-friendly markers with sufficient resolution to understand the

relationship between cultivars or species remains a challenge to be overcome, but it might be the only way to truly comprehend the complex taxonomy of the *Agave* genus.

As a powerful tool for plant breeders working with long-cycle species, physiological traits can enable early selection and expedite the breeding process. Measuring physiological parameters in *Agave* leaves can pose challenges due to their unique morphological and temporal characteristics, especially considering the CAM photosynthetic pathway. *Agave* leaves typically have thick succulent tissues with a waxy cuticle, which can make it difficult to obtain accurate measurements of gas exchange. Additionally, variable leaf anatomy, presence of teeth, and even huge dimensions can further complicate the process of collecting representative leaf samples and conducting proper physiological measurements. In this regard, chlorophyll fluorescence may be a useful tool to understand *Agave* photosynthesis as it can be easily measured under different leaf anatomies. Yet, informative experiments are still needed to elucidate the cross-talk of chlorophyll fluorescence analysis with both nocturnal carboxylation, mediated by phosphoenolpyruvate carboxylase (PEPC), and diurnal carboxylation, mediated by Rubisco. Another measurement that can be useful is colorimetry. Despite being highly variable depending on the nutritional and growth conditions, color is one of the most important factors for cultivar identification (Gentry, 1982; Colunga-GarcíaMarín et al., 1996; Figueredo et al., 2014; Figueredo-Urbina et al., 2021). In our data, colorimetric parameters were among the most significant for sample separation. These measurements can aid in initial screening to easily identify different accessions or abnormal physiological conditions.

At this moment, it is difficult to determine which accessions are showing the greatest potential as biorenewable feedstocks. However, it can be observed that among these, *A. fourcroydes* AFR exhibits one of the highest  $\Phi_{PSII}$  and SLM values, potentially indicating a greater capacity for biomass accumulation. An important issue regarding the selection and domestication of agaves is the presence of lateral teeth. While this is not a hindrance in Mexico, it has been a fundamental factor in the diffusion of agaves in other parts of the world (Medina, 1954; Shcolz, 1959; Silva and Beltrão, 1999). In fact, there is a clear difference in the adoption of plants such as *A. sisalana* and H11648 at a worldwide level, compared to *A. fourcroydes*, which has its commercial production limited to Mexico only. Regardless of the industrial application proposed for agaves, teeth pose a risk to the physical integrity of workers and make crop management difficult, and the selection of toothless cultivars or reduced teeth should be as important as sugar or fiber content. There are reports of cultivars with smooth edges in Mexico, such as *A. fourcroydes* var. *espiculata* and *A. tequilana* var. *manso*, which could have more prominence (Dewey, 1929; Shcolz, 1959; Valenzuela-Zapata and Nabhan, 2004). It can also be pointed out that agaves which are more labor-intensive in terms of management for cultivation and extraction, have smaller or absent teeth. This is the case with pulque cultivars of *A. salmiana*, which clearly show reduced teeth and gigantism as a strong mark of domestication (Gentry, 1982; Álvarez-Ríos et al., 2020). The presence of teeth in the progenies of *A. sisalana* has been a limiting factor for the improvement of this species, and one of the reasons why fiber improvement research focused on hybridization between *A. amaniensis* and *A. angustifolia* (Medina, 1954; Salgado et al., 1979; Silva and Beltrão, 1999; Monja-Mío et al., 2019). *A.*

*sisalana* exhibits instability with respect to this trait (Supplementary Tables S3, S4), often showing frequent “Escape” or reversion to its Armata form. Studies and identification of the genes responsible for teeth formation will be crucial for obtaining elite cultivars, especially considering the possible prospects of gene editing using CRISPR (Zaidi et al., 2018; Gosavi et al., 2020; Li et al., 2020; Zhou et al., 2020).

Before selecting specific characteristics, it is important to conduct on-site screening of different cultivars to identify their productive potential in each region (Davis, 2022). Nevertheless, considering the scenario of biorefineries, it may be important to develop cultivars with intermediate characteristics between *A. tequilana* and *A. sisalana*, that is, plants with a good stem (“piña”) size and sugar content while preserving the characteristics of commercial fiber, serving both the fiber and bioenergy markets. Additionally, breeding for resistance to pests and diseases is crucial. Currently, Brazilian sisal production is threatened by sisal bole rot disease, a poorly understood pathosystem that results in the complete death of *A. sisalana* (Duarte et al., 2018; Soares et al., 2020; Quintanilha-Peixoto et al., 2022). Another point of concern is the *Agave* snout weevil (*Scyphophorus acupunctatus*), one of the main pests of *Agave*, which caused a significant impact in the first field trial in the United States and has been reported in Brazil (Medina, 1954; Davis et al., 2017).

The global importance of *Agave* has been underestimated. Despite its decline in popularity due to synthetic fibers, its potential as a biorenewable feedstock is once again being recognized in face of global warming. Therefore, it is crucial to continue investing in *Agave* research and development to unlock its full potential for a more sustainable and secure future. To fully comprehend the diverse *Agave* genus, it is necessary to discover basic information and develop efficient analytical methodologies. Specifically, tools that facilitate easy identification and evaluation of potential materials are required, considering the challenges associated with studying agaves and their wide-ranging characteristics. Overall, this study provides valuable information on the morphophysiological traits of Brazilian *Agave* accessions and their phylogenetic relationships with other relevant *Agave* cultivars. It highlights the need for continued research and breeding programs to develop more productive and climate-resilient cultivars for sustainable biofuel and fiber production.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession numbers can be found in the article/Supplementary Material.

## Author contributions

FR: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing–Original draft preparation, Writing–Review and editing, Visualization. LuC: Methodology, Software, Formal analysis, Writing–Original draft preparation, Writing–Review and editing, Visualization. JJ: Methodology, Software, Formal analysis, Writing–Original draft preparation,

Writing–Review and editing, Visualization. LaC: Writing–Original draft preparation, Writing–Review and editing. RA: Investigation; HD: Investigation; NS: Investigation; SS: Investigation.; MP: Investigation; AO: Investigation; WR: Investigation; LA: Investigation; JG: Software, Validation, Formal analysis. MC: Methodology, Software, Formal analysis, Resources, Writing–Original Draft. AS: Resources, Writing–Original draft preparation, Writing–Review and editing, Supervision. JS: Resources, Writing–Original draft preparation, Writing–Review and editing, Supervision. RR: Conceptualization, Writing–Original draft preparation, Writing–Review and editing, Resources, Project administration, Funding acquisition. GP: Conceptualization, Writing–Original draft preparation, Writing–Review and editing, Resources, Project Leadership, Funding acquisition. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brazil (CAPES)* (grant number 88882.329521/2019-01); *Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)* (grant number 2020/02524-0, 2019/12914-3); *Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)—Nexus Project: Sisal-Caatinga Integration* (grant number 441625/2017-7); *Brazilian Agave development program (BRAVE)* (research agreements 5902/5903); *Fundo De Apoio Ao Ensino, Pesquisa E Extensão—FAEPEX* (grant number 2582/20), and *Programa de Apoio a Proyectos de Investigación e Innovación Tecnológica (PAPIIT)* from UNAM-Mexico (grant number IN222223). RR and GP are fellows of the National Council for Scientific and Technological Development (CNPq, Brazil).

## References

- Abraham, P. E., Yin, H., Borland, A. M., Weighill, D., Lim, S. D., De Paoli, H. C., et al. (2016). Transcript, protein and metabolite temporal dynamics in the CAM plant Agave. *Nat. Plants* 2, 16178. doi:10.1038/nplants.2016.178
- Agger, J. W., Nilsen, P. J., Eijnsink, V. G. H., and Horn, S. J. (2014). On the determination of water content in biomass processing. *BioEnergy Res.* 7, 442–449. doi:10.1007/s12155-013-9388-2
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Contr.* 19, 716–723. doi:10.1109/TAC.1974.1100705
- Alejandra, G. U. J., Liliana, S. Z., and Othon, S. S. R. (2013). *Agave syrup extract having anticancer activity*. Available at: <https://lens.org/041-823-850-889-841>.
- Almeida, R. L., Silveira, N. M., Pacheco, V. S., Xavier, M. A., Ribeiro, R. V., and Machado, E. C. (2021). Variability and heritability of photosynthetic traits in Saccharum complex. *Theor. Exp. Plant Physiol.* 33, 343–355. doi:10.1007/s40626-021-00217-x
- Álvarez-Ríos, G. D., Pacheco-Torres, F., Figueredo-Urbina, C. J., and Casas, A. (2020). Management, morphological and genetic diversity of domesticated agaves in Michoacán, México. *J. Ethnobiol. Ethnomed.* 16, 3. doi:10.1186/s13002-020-0353-9
- Aragón-Gastélum, J. L., Ramírez-Benítez, J. E., González-Durán, E., González-Salvatierra, C., Ramírez-Tobías, H. M., Flores, J., et al. (2020). Photochemical activity in early-developmental phases of agave angustifolia subsp. tequilana under induced global warming: implications to temperature stress and tolerance. *Flora* 263, 151535. doi:10.1016/j.flora.2019.151535
- Arnon, D. I. (1949). Copper enzymes in isolated chloroplasts. Polyphenoloxidase in beta vulgaris. *Plant Physiol.* 24, 1–15. doi:10.1104/pp.24.1.1
- Azzini, A., Ciaramello, D., Salgado, A. L. de B., and Zullo, M. A. T. (1989). Caracterização tecnológica de híbridos de sisal. *Bragantia* 48, 113–124. doi:10.1590/S0006-87051989000100011
- Basile, B., Mataffo, A., Forlani, M., and Corrado, G. (2022). Diversity in morphometric, pomological, and fruit-quality traits of apricot (*prunus armeniaca*) traditional varieties: implications for landrace differentiation at regional scale. *Diversity* 14, 608. doi:10.3390/d14080608
- Berger, A. (1915). *Die agaven*. Jena: Verlag Von Gustav Fischer.
- Broeren, M. L. M., Dellaert, S. N. C., Cok, B., Patel, M. K., Worrell, E., and Shen, L. (2017). Life cycle assessment of sisal fibre – exploring how local practices can influence environmental performance. *J. Clean. Prod.* 149, 818–827. doi:10.1016/j.jclepro.2017.02.073
- Campos, H., Trejo, C., Peña-Valdivia, C. B., García-Nava, R., Conde-Martínez, F. V., Cruz-Ortega, M., et al. (2014). Photosynthetic acclimation to drought stress in Agave salmiana Otto ex Salm-Dyck seedlings is largely dependent on thermal dissipation and enhanced electron flux to photosystem I. *Photosynth. Res.* 122, 23–39. doi:10.1007/s11120-014-0008-6
- Ciaramello, D., Castro, G. A. de P., and Petinelli, A. (1975). Estudio comparativo entre especies de agave. *Bragantia* 34, 195–201. doi:10.1590/S0006-87051975000100011
- Colunga-GarcíaMarín, P., Coello-Coello, J., Eguiarte, L. E., and Piñero, D. (1999). Isozymatic variation and phylogenetic relationships between henequén (Agave fourcroydes) and its wild ancestor A. angustifolia (Agavaceae). *Am. J. Bot.* 86, 115–123. doi:10.2307/2656960
- Colunga-GarcíaMarín, P., Estrada-Loera, E., and May-Pat, F. (1996). Patterns of morphological variation, diversity, and domestication of wild and cultivated populations of agave in Yucatan, Mexico. *Am. J. Bot.* 83, 1069–1082. doi:10.1002/j.1537-2197.1996.tb12805.x
- Corbin, K. R., Byrt, C. S., Bauer, S., DeBolt, S., Chambers, D., Holtum, J. A. M., et al. (2015). Prospecting for energy-rich renewable raw materials: agave leaf case study. *PLoS One* 10, e0135382. doi:10.1371/journal.pone.0135382

## Acknowledgments

We express our gratitude to Dr. Alisson Chiorato and Dr. Sérgio Carbonell of the Center of Grains and Fibers at IAC for granting us permission to collect the remaining *Agave* materials at the institution. Finally, we dedicate this work to Júlio Cesar Medina, Anisio Azzini, Antonio Luiz de Barros Salgado, and Dirceu Ciaramello, IAC researchers whose essential contributions were instrumental in the consolidation of sisal production in Brazil.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fceng.2023.1218668/full#supplementary-material>

- Davis, S. C. (2022). Agave americana: characteristics and potential breeding priorities. *Plants* 11, 2305. doi:10.3390/plants11172305
- Davis, S. C., Kuzmick, E. R., Niechayev, N., and 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, 314–325. doi:10.1111/gcbb.12324
- Dewey, L. H. (1929). Agave fourcroydes var. espiculata. *J. Washingt. Acad. Sci.* 19, 416.
- Dinh, T. H., Takaragawa, H., Watanabe, K., Nakabaru, M., and Kawamitsu, Y. (2019). Leaf photosynthesis response to change of soil moisture content in sugarcane. *Sugar Tech.* 21, 949–958. doi:10.1007/s12355-019-00735-8
- Doughty, L. R. (1936). Chromosome behaviour in relation to genetics of Agave. *J. Genet.* 33, 198–205. doi:10.1007/BF02982532
- Duarte, E. A. A., Damasceno, C. L., de Oliveira, T. A. S., Barbosa, L. de O., Martins, F. M., de Silva, J. R. Q., et al. (2018). Putting the mess in order: aspergillus welwitschiae (and not A. niger) is the etiological agent of sisal bole rot disease in Brazil. *Front. Microbiol.* 9, 1227. doi:10.3389/fmicb.2018.01227
- Guiarte, L. E., Jiménez Barrón, O. A., Aguirre-Planter, E., Scheinvar, E., Gámez, N., Gasca-Pineda, J., et al. (2021). Evolutionary ecology of agave: distribution patterns, phylogeny, and coevolution (an homage to Howard S. Gentry). *Am. J. Bot.* 108, 216–235. doi:10.1002/ajb2.1609
- Escamilla-Treviño, L. L. (2012). Potential of plants from the genus agave as bioenergy crops. *Bioenergy Res.* 5, 1–9. doi:10.1007/s12155-011-9159-x
- FAO (2023). *FAO statistics division*. Available at: <http://www.fao.org/faostat/en/> (Accessed February 10, 2023).
- Figueredo, C. J., Casas, A., Colunga-GarcíaMarín, P., Nassar, J. M., and González-Rodríguez, A. (2014). Morphological variation, management and domestication of ‘maguey alto’ (Agave inaequidens) and ‘maguey manso’ (A. hookeri) in Michoacán, México. *J. Ethnobiol. Ethnomed.* 10, 66. doi:10.1186/1746-4269-10-66
- Figueredo-Urbina, C. J., Álvarez-Ríos, G. D., García-Montes, M. A., and Octavio-Aguilar, P. (2021). Morphological and genetic diversity of traditional varieties of agave in Hidalgo State, Mexico. *PLoS One* 16, e0254376. doi:10.1371/journal.pone.0254376
- Furbank, R. T., Jimenez-Berni, J. A., George-Jaeggli, B., Potgieter, A. B., and Deery, D. M. (2019). Field crop phenomics: enabling breeding for radiation use efficiency and biomass in cereal crops. *New Phytol.* 223, 1714–1727. doi:10.1111/nph.15817
- Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* 31, 3718–3720. doi:10.1093/bioinformatics/btv428
- García-Moya, E., Romero-Manzanares, A., and Nobel, P. S. (2011). Highlights for agave productivity. *GCB Bioenergy* 3, 4–14. doi:10.1111/j.1757-1707.2010.01078.x
- Gentry, H. S. (1982). *Agaves of continental north America*. Tucson: University of Arizona Press.
- Gil-Vega, K., Díaz, C., Nava-Cedillo, A., and Simpson, J. (2006). AFLP analysis of Agave tequilana varieties. *Plant Sci.* 170, 904–909. doi:10.1016/j.plantsci.2005.12.014
- Gosavi, G., Yan, F., Ren, B., Kuang, Y., Yan, D., Zhou, X., et al. (2020). Applications of CRISPR technology in studying plant-pathogen interactions: overview and perspective. *Phytopathol. Res.* 2, 21. doi:10.1186/s42483-020-00060-z
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi:10.1093/sysbio/syq010
- Hasan, N., Choudhary, S., Naaz, N., Sharma, N., and Laskar, R. A. (2021). Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. *J. Genet. Eng. Biotechnol.* 19, 128. doi:10.1186/s43141-021-00231-1
- Holtum, J. A. M., Chambers, D., Morgan, T., and Tan, D. K. Y. (2011). Agave as a biofuel feedstock in Australia. *GCB Bioenergy* 3, 58–67. doi:10.1111/j.1757-1707.2010.01083.x
- Jean, P. (2015). *Active ingredient with cutaneous application obtained from metschnikowia agaves and uses for improving the state of the skin*. Available at: <https://lens.org/043-503-327-571-345>.
- Jeong, P. E., Chul, S. W., Woo, L. D., Eun, L. J., Sun, Y. T., and Ju, K. N. (2020). *Cosmetic composition comprising enzyme-treated extract of agave syrup as active ingredient*. Available at: <https://lens.org/097-139-203-731-408>.
- Jiménez-Barrón, O., García-Sandoval, R., Magallón, S., García-Mendoza, A., Nieto-Sotelo, J., Aguirre-Planter, E., et al. (2020). Phylogeny, diversification rate, and divergence time of agave sensu lato (asparagaceae), a group of recent origin in the process of diversification. *Front. Plant Sci.* 11, 536135. doi:10.3389/fpls.2020.536135
- Kassambara, A., and Mundt, F. (2020). *factoextra: Extract and visualize the results of multivariate data analyses*. Available at: <http://www.sthda.com/english/rpkgs/factoextra>.
- Kassambara, A. (2023). *rstatix: Pipe-Friendly framework for basic statistical tests*. Available at: <https://rpkgs.datanovia.com/rstatix/>.
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi:10.1093/molbev/mst010
- Laborde Aguirre, A. E., Valencia Gallegos, J. A., Hernández Valdéz, J. S., del Real Laborde, J. I., and Laborde Cancino, M. de J. S. (2013). *Proceso para preparar una mezcla termoplástica de fibras residuos de agave y aditivos oxo-biodegradativos para preparar artículos de plástico biodegradables*. US20130099029A1, 15.
- Laborde Aguirre, A. E., Valencia Gallegos, J. A., Sergio, H. V. J., del Real Laborde, J. I., and Laborde Cancino, M. de J. S. (2010). *Proceso para preparar una mezcla termoplástica polimérica de fibras residuos de agave y aditivos oxo-biodegradativos para preparar artículos de plástico biodegradables*. WO2011155814A1, 37.
- Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* 25. doi:10.18637/jss.v025.i01
- Lewandowski, I., and Heinz, A. (2003). Delayed harvest of miscanthus—Influences on biomass quantity and quality and environmental impacts of energy production. *Eur. J. Agron.* 19, 45–63. doi:10.1016/S1161-0301(02)00018-7
- Li, Q., Sapkota, M., and van der Knaap, E. (2020). Perspectives of CRISPR/Cas-mediated cis-engineering in horticulture: unlocking the neglected potential for crop improvement. *Hortic. Res.* 7, 36. doi:10.1038/s41438-020-0258-8
- Li, R., Guo, P., Michael, B., Stefania, G., and Salvatore, C. (2006). Evaluation of chlorophyll content and fluorescence parameters as indicators of drought tolerance in barley. *Agric. Sci. China* 5, 751–757. doi:10.1016/S1671-2927(06)60120-X
- Lledías, F., Gutiérrez, J., Martínez-Hernández, A., García-Mendoza, A., Sosa, E., Hernández-Bermúdez, F., et al. (2020). Mayahuelin, a type I ribosome inactivating protein: characterization, evolution, and utilization in phylogenetic analyses of agave. *Front. Plant Sci.* 11, 573. doi:10.3389/fpls.2020.00573
- Long, S. P., Karp, A., Buckeridge, M. S., Davis, S. C., Jaiswal, D., Moore, P. H., et al. (2015). “Feedstocks for biofuels and bioenergy.” in *Bioenergy & sustainability: Bridging the gaps* (Paris: SCOPE).
- Lowman, M. D., and Box, J. D. (1983). Variation in leaf toughness and phenolic content among five species of Australian rain forest trees. *Austral Ecol.* 8, 17–25. doi:10.1111/j.1442-9993.1983.tb01515.x
- Luján, R., Lledías, F., Martínez, L. M., Barreto, R., Cassab, G. J., and Nieto-Sotelo, J. (2009). Small heat shock proteins and leaf cooling capacity account for the unusual heat tolerance of the central spike leaves in Agave tequilana Weber. *Plant, Cell and Environment* 32, 1791–1803. doi:10.1111/j.1365-3040.2009.02035.x
- Maxwell, K., and Johnson, G. N. (2000). Chlorophyll fluorescence—A practical guide. *J. Exp. Bot.* 51, 659–668. doi:10.1093/jexbot/51.345.659
- McGuire, R. G. (1992). Reporting of objective color measurements. *HortScience* 27, 1254–1255. doi:10.21273/HORTSCI.27.12.1254
- Medina, J. C. (1954). *O sisal*. USA: Secretaria da Agricultura do Estado de São Paulo.
- Medina, J. C. (1956). Observações preliminares sobre agave zapupe trelease. *Bragantia* 15, XXIII. doi:10.1590/S0006-87051956000100037
- Medina, J. C. (1959). *Plantas Fibrosas da Flora Mundial*. China: Instituto Agronômico de Campinas.
- Medina, J. C. (1955). Variação em Agave sisalana. *Bragantia* 14, XXVII. doi:10.1590/S0006-87051955000100036
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi:10.1093/molbev/msaa015
- Mir, R. R., Zaman-Allah, M., Sreenivasulu, N., Trethowan, R., and Varshney, R. K. (2012). Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. *Theor. Appl. Genet.* 125, 625–645. doi:10.1007/s00122-012-1904-9
- Monja-Mio, K. M., Herrera-Alamillo, M. A., Sánchez-Teyer, L. F., and Robert, M. L. (2019). “Breeding strategies to improve production of agave (agave spp.)” in *Advances in plant breeding strategies: Industrial and food crops* (Cham: Springer International Publishing), 319–362. doi:10.1007/978-3-030-23265-8\_10
- Morán, J. I., Alvarez, V. A., Cyras, V. P., and Vázquez, A. (2008). Extraction of cellulose and preparation of nanocellulose from sisal fibers. *Cellulose* 15, 149–159. doi:10.1007/s10570-007-9145-9
- Morillo-Coronado, A. C., Manjarres Hernández, E. H., and Forero-Mancipe, L. (2021). Phenotypic diversity of morphological characteristics of pitahaya (selenicereus megalanthus haw.) germplasm in Colombia. *Plants* 10, 2255. doi:10.3390/plants10112255
- Natarajan, S., Basnayake, J., Wei, X., and Lakshmanan, P. (2019). High-throughput phenotyping of indirect traits for early-stage selection in sugarcane breeding. *Remote Sens.* 11, 2952. doi:10.3390/rs11242952
- Nobel, P. S. (1988). *Environmental biology of agaves and cacti*. Cambridge: Cambridge University Press.
- Nowell, W. (1933). Agave amaniensis: A new form of fibre-producing agave from Amani. *Bull. Misc. Inf. R. Gard. Kew*, 465. doi:10.2307/4113414
- Palomo-Briones, R., López-Gutiérrez, I., Islas-Lugo, F., Galindo-Hernández, K. L., Munguía-Aguilar, D., Rincón-Pérez, J. A., et al. (2018). Agave bagasse biorefinery: processing and perspectives. *Clean. Technol. Environ. Policy* 20, 1423–1441. doi:10.1007/s10098-017-1421-2



- Parascanu, M. M., Sanchez, N., Sandoval-Salas, F., Carreto, C. M., Soreanu, G., and Sanchez-Silva, L. (2021). Environmental and economic analysis of bioethanol production from sugarcane molasses and agave juice. *Environ. Sci. Pollut. Res.* 28, 64374–64393. doi:10.1007/s11356-021-15471-4
- Peng, J., Feng, Y., Wang, X., Li, J., Xu, G., Phonenasay, S., et al. (2021). Effects of nitrogen application rate on the photosynthetic pigment, leaf fluorescence characteristics, and yield of indica hybrid rice and their interrelations. *Sci. Rep.* 11, 7485. doi:10.1038/s41598-021-86858-z
- Peng, Y., Leung, H. C. M., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi:10.1093/bioinformatics/bts174
- Queiroz, S. R. de O. D., Ortolani, F. A., Mataqueiro, M. F., Osuna, J. T. A., and Moro, J. R. (2012). Análise cromossômica em bulbilhos de sisal (*Agave spp.*) cultivados em diferentes municípios baianos, Brasil. *Acta Bot. Bras.* 26, 842–848. doi:10.1590/S0102-33062012000400013
- Quintanilha-Peixeoto, G., Marone, M. P., Raya, F. T., José, J., Oliveira, A., Fonseca, P. L. C., et al. (2022). Phylogenomics and gene selection in *Aspergillus welwitschiae*: possible implications in the pathogenicity in agave sisalana. *Genomics* 110517, 110517. doi:10.1016/j.ygeno.2022.110517
- R Core Team (2022). *R: A language and environment for statistical computing*. Available at: <https://www.r-project.org/>.
- Raya, F. T., de Abreu, L. G. F., Marone, M. P., Salvador, M. de A., Bressiani, J. A., Laborde, J. I. del R., et al. (2022). in *New feedstocks for bioethanol production: Energy cane and agave liquid biofuels: Bioethanol*. Editors C. R. Soccol, G. A. G. Pereira, C. Dussap, and L. P. S. Vandenberghe (Germany: Springer), 431–455. doi:10.1007/978-3-031-01241-9\_18
- Raya, F. T., Marone, M. P., Carvalho, L. M., Rabelo, S. C., de Paula, M. S., Campanari, M. F. Z., et al. (2021). Extreme physiology: biomass and transcriptional profiling of three abandoned agave cultivars. *Ind. Crops Prod.* 172, 114043. doi:10.1016/j.indcrop.2021.114043
- Rivera-Lugo, M., García-Mendoza, A., Simpson, J., Solano, E., and Gil-Vega, K. (2018). Taxonomic implications of the morphological and genetic variation of cultivated and domesticated populations of the *Agave angustifolia* complex (Agavoideae, Asparagaceae) in Oaxaca, Mexico. *Plant Syst. Evol.* 304, 969–979. doi:10.1007/s00606-018-1525-0
- Robert, M. L., Lim, K. Y., Hanson, L., Sanchez-Teyer, F., Bennett, M. D., Leitch, A. R., et al. (2008). Wild and agronomically important *Agave* species (Asparagaceae) show proportional increases in chromosome number, genome size, and genetic markers with increasing ploidy. *Bot. J. Linn. Soc.* 158, 215–222. doi:10.1111/j.1095-8339.2008.00831.x
- Rothfels, C. J. (2021). Polyploid phylogenetics. *New Phytol.* 230, 66–72. doi:10.1111/nph.17105
- Salgado, A. L. de B., Caramello, D., and Azzini, A. (1979). Melhoramento de Agave por hibridação. *Bragantia* 38, 1–6. doi:10.1590/S0006-87051979000100001
- Sahu, S. K., Thangaraj, M., and Kathiresan, K. (2012). DNA extraction protocol for plants with high levels of secondary metabolites and polysaccharides without using liquid nitrogen and phenol. *ISRN Mol. Biol.* 2012, 1–6. doi:10.5402/2012/205049
- Shcolz, H. (1959). *Sisal (problemas técnicos)*. Ceará: Banco do nordeste do brasil. New York: Escritório Técnico de Estudos Econômicos do Nordeste.
- Silva, O. R. R. F., and Beltrão, N. E. de M. (1999). O agronegócio do sisal no brasil. Brasília, DF: Embrapa.
- Singh, S. K., Reddy, V. R., Fleisher, D. H., and Timlin, D. J. (2017). Relationship between photosynthetic pigments and chlorophyll fluorescence in soybean under varying phosphorus nutrition at ambient and elevated CO<sub>2</sub>. *Photosynthetica* 55, 421–433. doi:10.1007/s11099-016-0657-0
- Soares, A. C. F., Santos, J. S., de Sousa, R. A., Barbosa, L. de O., Duarte, E. A. A., and Góes-Neto, A. (2020). “Sisal: podridão vermelha e o biocontrole,” in *Tópicos em microbiologia agrícola*. Editors A. C. F. Soares, N. S. Evangelista-Barreto, and P. A. S. Marbach (China: Cruz das Almas: EDUFRRB), 276.
- Souza, S. C. de, Cavalcanti, J. J. V., Ramos, J. P. C. R., Alves, I., Santos, R. C. dos, and Lima, L. M. de (2018). Genetic divergence in *Agave* accessions through ISSR markers and phenotypic traits. *Afr. J. Agric. Res.* 13, 526–533. doi:10.5897/AJAR2017.12913
- Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* 35, 1026–1028. doi:10.1038/nbt.3988
- Suinaga, F. A., Silva, O. R. R. F., Coutinho, W. M., Cartaxo, W. V., and Costa, L. B. (2007). Avaliação agrônômica de Oito genótipos de Sisal (*agave spp.*). *Comun. Técnico.* 2007
- Tanaka, A., and Tanaka, R. (2006). Chlorophyll metabolism. *Curr. Opin. Plant Biol.* 9, 248–255. doi:10.1016/j.pbi.2006.03.011
- Trejo-Torres, J. C., Gann, G. D., and Christenhusz, M. J. M. (2018). The yucatan peninsula is the place of origin of sisal (*agave sisalana*, asparagaceae): historical accounts, phylogeography and current populations. *Bot. Sci.* 96, 366–379. doi:10.17129/botsci.1928
- Tremblay, N., Wang, Z., and Cerovic, Z. G. (2012). Sensing crop nitrogen status with fluorescence indicators. A review. *Agron. Sustain. Dev.* 32, 451–464. doi:10.1007/s13593-011-0041-1
- Valenzuela-Zapata, A. G., and Nabhan, G. P. (2004). *¡Tequila!: A natural and cultural history*. Tucson: University of Arizona Press.
- Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., Zemla, J., et al. (2021). *corrplot: Visualization of a correlation matrix*. Available at: <https://github.com/taiyun/corrplot>.
- WFO (2023). *Agave amaniensis trel. & nowell. World flora online data*. Available at: <http://www.worldfloraonline.org/taxon/wfo-0000754118> (Accessed March 18, 2023).
- Xu, B., Tan, S., Qin, X., Huang, X., Xi, J., Chen, H., et al. (2022). The complete chloroplast genome of *agave amaniensis* (asparagales: asparagaceae: agavoideae). *Mitochondrial DNA Part B* 7, 1519–1521. doi:10.1080/23802359.2022.2109440
- Yamada, S., Gotoh, O., and Yamana, H. (2006). Improvement in accuracy of multiple sequence alignment using novel group-to-group sequence alignment algorithm with piecewise linear gap cost. *BMC Bioinforma.* 7, 524. doi:10.1186/1471-2105-7-524
- Yan, X., Corbin, K. R., Burton, R. A., and Tan, D. K. Y. (2020). Agave: A promising feedstock for biofuels in the water-energy-food-environment (WEFE) nexus. *J. Clean. Prod.* 261, 121283. doi:10.1016/j.jclepro.2020.121283
- Zaidi, S. S.-A., Mukhtar, M. S., and Mansoor, S. (2018). Genome editing: targeting susceptibility genes for plant disease resistance. *Trends Biotechnol.* 36, 898–906. doi:10.1016/j.tibtech.2018.04.005
- Zhou, J., Li, D., Wang, G., Wang, F., Kunjal, M., Joldersma, D., et al. (2020). Application and future perspective of CRISPR/Cas9 genome editing in fruit crops. *J. Integr. Plant Biol.* 62, 269–286. doi:10.1111/jipb.12793