The Journal of Animal & Plant Sciences, 29(5): 2019, Page: 1308-1317
ISSN: 1018-7081

# GENETIC STUDIES ON MORPHOLOGICAL VARIABILITY AND CARBOHYDRATE CONTENT OF *AGAVE DURANGENSIS* AS POTENTIAL TOOLS FOR CONSERVATION AND PROPAGATION PROGRAMS

E. A. Delgado-Alvarado, J. A. Ávila-Reyes, N. Almaraz-Abarca, L. S. González-Valdez, J. N. Uribe-Soto, H. M. Monreal-García, V. Hernández-Vargas, N. Naranjo-Jiménez, and R. Torres-Ricario*

Interdisciplinary research center for Integrated Regional Development, Durango, Instituto Politécnico Nacional, Sigma 119, Colonia 20 de Noviembre II, Durango, Mexico, 34220.

*Corresponding author's email: rtorresr@ipn.mx

## ABSTRACT

*Agave durangensis* is a wild endemic species of Mexico, which supports a regional mescal industry. The reproduction of *A. durangensis* is mainly sexual, conferring it a high variability. For the mescal manufacturing, *A. durangensis* is collected from wild populations, affecting their reproductive success and demographic performance. We evaluated the morphological variation, carbohydrate content, and genetic structure of *A. durangensis* plants coming from a single mother plant, to identify lines with some agroindustry potential. Three different morphological and genetic groups were distinguished. According to Mantel tests (r = 0.71), statistical correlations between morphological (Mahalanobis distance) and genetic (Fst) distances (r= 0.71), as well as between carbohydrate content similarity matrix and morphological traits (r= 0.477) were found. The high correlation values suggested that the use of genetic markers combined with morphological traits could be useful as a tool for the selection of high carbohydrate production lines to establish cultivars, which would have conservational implications.

**Key words:** *Agave durangensis,* genetic structure, microsatellites, morphological, carbohydrate content.

## INTRODUCTION

The genus *Agave* (Agavaceae) hosts about 175 species, of which 135 are endemic to Mexico (Montañez-Soto *et al*., 2011), because of this, Mexico is considered as center of origin of the genus (Almaraz-Abarca *et al*., 2013). Different species of *Agave* has been used for several purposes, such as food, medicinal, construction, ornamental, textile, and raw material for manufacturing the distilled alcoholic beverages like tequila, and mescal (Nava-Cruz *et al*., 2015). Mescal can be obtained from the fermented sugars of cooked stems from 53 different *Agave* species. Currently, around 37 of them are harvested from wild populations without management programs (Torres *et al*., 2015). One of these species is *Agave durangensis*, which is mainly extracted from wild populations (Almaraz-Abarca *et al*., 2013). The reproduction of *A. durangensis* is majorly by seeds, which confers it a high genetic and morphological variability that makes difficult the delimitation of species, but represents an essential source of valuable alleles for the selection of plants with outstanding features for the establishment of commercial plantations (Almaraz-Abarca *et al*., 2013).

The increased acceptance of the mescal from *A. durangensis*has resulted in irrational exploitation of this plant, affecting seed and pollen production and decreasing gene flow among populations (Félix-Valdez *et al*., 2016) and has revealed the need to establish cultivars. A previous study carried out by Almaraz-Abarca *et al*., (2013) revealed that several wild populations of *A. durangensis* showed an advanced diversification, and highlighted the necessity of developing conservation and protection programs for the maintenance of genetic variability of natural populations.

Several propagation and conservation programs for some *Agave* species had been based only on morphological and phenotypical traits, but recently, the use of molecular markers such AFLPS, ISSRs, RADPS and SSR has improved these studies (Colunga-GarcíaMarín and May-Pat, 1997; Figueredo and Nassar, 2011; Rodríguez-Garay *et al*., 2009). An important aspect for the improvement and conservation of crops is the selection of progenitors bearing desirable attributes. The levels of bioproduct accumulation have been criteria used to identify progenitors of economically important plants; such is the case of carotenoid in carrots and tocopherol in *Brassica napus* L., among others (Fritsche *et al*., 2012; Jourdan *et al*., 2015). For *Agave* species, the amount of carbohydrates accumulated in the stems is an important feature to select progenitors, but for wild populations, this accumulation and its possible relation with morphologic and molecular traits has been poorly studied (Michel-Cuello *et al*., 2008). The present study aimed to evaluate the genetic variability, carbohydrate contents, and morphological variations of six populations of *A. durangensis*, coming from just one plant progenitor, to evaluate the variability generated by a single plant and to

assess the correlation between carbohydrate contents and morphologic and genetic variations.

## MATERIALS AND METHODS

**Plant material and morphological analysis:** One hundred and thirty-two individuals of *A. durangensis* (three years old), coming from seeds of a single progenitor plant (collected in October 29, 2009) from La Parrilla, Durango, Mexico (23° 42' 21.8" N, 104° 12' 39.1" W) were cultivated in a nursery Durango, Mexico (24° 03' 03" N, 104° 36' 38" W). Thirty-months old plants were classified in six lines according to their morphological features. The lines were called L1, L2, L3, L4, L5, and L6. Five morphological characters were measured, and three ratios were also calculated using those characters (Figueredo-Urbina *et al*., 2017). The evaluated morphological characters were: Plant Height (PH), Rosette Diameter (RD), Leaf Length (LL), Leaf Medium Width (LMW) and Length of Apical Spine (LAS); meanwhile, the calculated ratios were: Number or Lateral Spines/ 10 cm (LS10), Plant Height/Rosette Diameter (PH/RD), and Leaf Length/Leaf Medium Width (LL/LMW). Random plants (unknown age) from a wild population of *A. durangensis* (D) was sampled from Nombre de Dios, Durango, Mexico (24° 02' 51"N; 104°15' 24"W) and analyzed as same as cultivated ones.

**Carbohydrate contents:** Samples (2 x 2 cm) of stem and leaves were obtained using a Pressler drill. Water-soluble carbohydrates (WSC) were obtained mixing 1 g of fresh tissue with 10 mL of bi-distilled water for 30 min at 60 °C with constant stirring, as described by Montañez-Soto *et al*., (2011).

Total carbohydrate contents (TCC) were determined mixing 100 µL of the WSC with 200 µL of 1M anthrone (prepared in $H_2SO_4$), heated at 100 °C for 10 min with constant stirring; the reaction was stopped by ice immersion for five min. The absorbance at 625 nm was registered using a Genesys-10s UV-vis spectrometer (Thermo Scientific, Waltham, MA).Carbohydrate content was obtained from a standard curve of sucrose with the following equation:[Sucrose concentration (mg/ml) = $(ABS_{625}+0.0002)/14338$], $R^2=0.995$.

Free reducing sugars (FRS) were determined by mixing 100 µL of WSC with 100 µL of 1 % DNS reagent (40 mM 3,5-Dinitrosalicylic acid, 1M $KNaC_4H_4O_6 \cdot 4H_2O$and 0.4 M NaOH). Mixtures were shaken for 5 min at 100 °C, and then placed in ice for 5 min (Montañez-Soto *et al*., 2011). The absorbance at 540 nm was registered and FRS contents were calculated from a standard curve of fructose with the following equation: [Fructose concentration (mg/ml) = $(ABS_{540}+0.0851)/3.4564$] $R^2=0.999$.

The fructooligosaccharides (FOS) contents were calculated from the difference of the TCC minus the FSC, according to Arrizon *et al*., (2010).

**DNA extraction and SSRs analysis:** DNA was extracted from foliar tissues using a method based on CTAB (Keb-Llanes *et al*., 2002).The used primers and PCR amplification conditions were, according to Lyndsay *et al*., (2012). The total PCR mixture was formed by 20 µL, consisted of 4 µL of 5X PCR Buffer (Promega, Madison, WI, USA), 2.8 µL of 25 mM $MgCl_2$ (Promega, Madison, WI, USA), 0.4 µL of 10 mM dNTPs mix (Promega, Madison, WI, USA), 0.2 µL of 10 mM primer (IDT technologies), 0.2 U of Taq polymerase (5 U/µL) (Promega, Madison, WI, USA), 11.2 µL of free nuclease and RNase $H_2O$, and 1 µL of DNA solution (25 ng/µL). The PCR conditions were: One cycle of 95 °C for 5 min, 40 cycles of 95°C for 30 s, 57-60 °C (annealing temperature depending on each primer) for 60 s, and 72 °C for 45 s, with a final step of 72°C for 10 min. PCR products were separated in 6% polyacrylamide gels at 160 V and stained with 10X SYBR® Gold (Invitrogen, USA), during 30 min. The gels were observed under UV light on a UVP transilluminator CANON; then the fragments were recorded using a gel image software PyElph (Pavel and Vasile, 2012).

**Data analysis:** To identify whether there were any statistical differences either of morphological or carbohydrate variables between the seven populations, a one-way analysis of variance (ANOVA) was performed, measuring 22 individuals from each population. Tukey's multiple mean comparison was conducted for test of significant differences. A Principal Component Analysis (PCA), using Pearson correlations, was done for both morphological and carbohydrate data to evaluate the contribution of each trait to variation (Mamuris *et al.,* 1998). A similarity matrix for each type of data was build using Mahalanobis distances, according to Legendre and Legendre (2012).

Polymorphic SSRs bands were scored as presence (1) or absence (0) and used for the construction of a binary matrix. Evidence of null alleles due to homozygote excess, large allele dropouts and scoring errors were checked using the software micro-checker 2.2.3 (Van Oosterhout *et al*., 2004). Allele number (Na), Effective allele number (Ne), Expected heterozygosity (He), and Polymorphic Index Content (PIC), were calculated using Friedman method at $\alpha=0.05$, with the software Infogen (Balzarini & Di Rienzo, 2003). Inbreeding or fixation index (Fis) and coefficient of differentiation (Fst) were calculated using the software Genepop (Rousset, 1995, 2008). A genetic similarity matrix was constructed using the Fst values. Analysis of Molecular Variance (AMOVA) (Excoffier *et al*., 1992) was conducted to find genetic differences among populations, within individuals, and among populations

within groups, with the software Arlequin v 3.5.2.2 (Excoffier and Lischer, 2010). The genetic structure was estimated using a Bayesian clustering model with a pre burn-in process of 5000 and 50000 runs and $10^6$ periods of the Markov Chain Monte Carlo algorithm with the program STRUCTURE v 2.3.4 (Pritchard *et al*., 2000). The optimum group number ($k$) was determined from 1 to 10 with ten runs for each $k$ value, calculated with the ΔK method that uses the maximum value of the posterior probability ln P($K$) described by Evanno (2005) with the online platform Structure Harvester (Earl, 2012).

The correlation among the three groups of variables was determined using a Mantel test at α=0.05, as reported by Diniz-Filho *et al.,* (2013) using Pearson correlations between similarity matrices previously calculated, employing the software XLSTAT (Addinsoft, 2019).

## RESULTS AND DISCUSSION

**Morphological variation:** The descriptive statistics of the eight morphological characters evaluated for six cultivated lines and one wild population of *A. durangensis*are is given in Table 1. Significance differences were found in seven of the eight morphological traits (PH/RD index was the only trait that showed no differences). Plant height (PH) ranged from 29.22 cm (line 5) to 46.79 cm (Line 2), and grouped the six cultivated lines into three groups, D formed an independent group. The first group is formed by line 1 and line 2, the second group by lines 3 and 4 and the third one by lines 5 and 6. The same groups were formed for LL and partially by RD. The rest of the evaluated characters showed no apparent differences among cultivated lines, neither between lines and wild individuals. The wild individuals of *A. durangensis* showed similar PH values to those reported by Almaraz-Abarca *et al.*, (2013), who indicated an average value of 94.69 cm for three wild populations. The observed variation of the morphological characters in both cultivated and wild individuals of *A. durangensis* gave an insight of the natural variety of this species, which makes it challenging to develop harvest strategies, as plants of the same age displayed considerable morphological differences. Some authors suggested that the morphological variation in *Agave* species is due to the heterogeneous environment where these species occur and to the indirect expression of their genomic variation (Figueredo-Urbina *et al*., 2017). The current results revealed that a single plant has the potential to generate a population highly variable, considering that cultivated individuals were grown under the same conditions, the environmental context has less important participation than the genetic one in determining the morphological variation of *A. durangensis*.

**Table 1. Morphological characters measured for cultivated (L1 to L2) lines of *Agave durangensis* and one wild population (D) of *Agave durangensis.***

| Variable | L1 | L2 | L3 | L4 | L5 | L6 | D |
|---|---|---|---|---|---|---|---|
| PH (cm)* | 41.6 ± 4.3 c | 46.7 ± 7.4 c | 35.1 ± 5.8 b | 38.3 ± 6.5 b | 29.2 ± 3.2 a | 29.3 ± 3.1 a | 99.2 ± 18.1 d |
| RD (cm)* | 57.8 ± 12.4 b | 60.8 ± 11.5 b | 50.4 ± 6.6 b | 49.7 ± 8.4 b | 39.6 ± 5.6 a | 38.3 ± 6.2 a | 127.1 ± 28.5 c |
| LL (cm)* | 48.8 ± 3.9 c | 47.1 ± 6.2 c | 33.8 ± 3.6 b | 37.5 ± 3.7 b | 26.7 ± 3.1 a | 29.0 ± 2.5 a | 68.1 ± 11.9 d |
| LMW (cm)* | 15.0 ± 1.7 b | 13.7 ± 1.5 b | 14.0 ± 1.8 b | 10.7 ± 1.6 a | 11.5 ± 1.2 a | 9.5 ± 1.2 a | 20.1 ± 3.7 c |
| LS10 (#)* | 7.5 ± 2.7 d | 5.4 ± 1.5 b | 6.1 ± 1.4 c | 7.4 ± 1.8 d | 6.7 ± 1.5 d | 7.3 ± 2.2 d | 4.8 ± 0.7 b |
| LAS (cm)* | 3.7 ± 0.8 a | 4.6 ± 0.8 b | 3.7 ± 0.6 a | 3.7 ± 0.8 a | 3.3 ± 0.6 a | 3.9 ± 0.8 a | 5.7 ± 1 c |
| PH/RD (ratio) | 0.7 ± 0.1 a | 0.7 ± 0.1 a | 0.7 ± 0.1 a | 0.7 ± 0.1 a | 0.7 ± 0.1 a | 0.7 ± 0.1 a | 0.7 ± 0.1 a |
| LL/LMW (ratio)* | 2.6 ± 0.3 a | 3.4 ± 0.4 c | 2.4 ± 0.3 a | 3.5 ± 0.5 c | 2.3 ± 0.2 a | 3.0 ± 0.3 b | 3.4 ± 0.5 c |

Mean value ± standard deviation, $^*p \leq 0.05$, Different letters in the same line mean statistical differences among populations.

The results of a PCA considering eight morphological attributes are shown in Figure 1. The first component accounted for 47.78% of the total variance, whereas the second one accounted for 18.11%. Two well-identified groups were formed along the principal component axis, separating the cultivated from the wild individuals. The first component was related to PH (0.414), LL (0.416), LMW (0.398), and RD (0.411), being the ones with the highest significant contribution, based on the cosine square determinations ($p \leq 0.05$). These characters could be important morphological markers for distinguishing lines of *A. durangensis*. The second component was mainly influenced by the indexes PH/RD (0.728) and LL/LMW (0.636). Previous studies showed that LL and LMW were associated to the identification of elite individuals of *A. fourcroydes* and that these attributes may play an important role for the genetic improvement of this species (González *et al*., 2003). LL and LMW, besides lateral and apical spines, also were valuable markers to distinguish between species of *Agave*, as between *A. angustifolia* and *A. tequilana* (Rodríguez-Garay *et al*., 2009).

**Carbohydrate contents:** Carbohydrate contents of stem and leaves are shown in Table 2. Significant differences

were revealed between cultivated and wild individuals of *A. durangensis* ($p \leq 0.05$). Leaves of cultivated individuals showed a range of TCC from 68.9 mg/g (L5) to 98.6 mg/g (L6), in stems this range was from 12.0 mg/g (L1) to 27.7 mg/g (L2). The FRS contents ranged from 6.7 mg/g (L4) to 12.9 mg/g (L1) in leaf, whereas from 5.1 mg/g (L3) to 11.7 mg/g (L6) in stems. Meanwhile, the FOS content ranged from 60.5 mg/g (L5) to 90.1 mg/g (L6) in leaves, and 3.5 mg/g (L1) to 21.6 mg/g (L2) in stem.

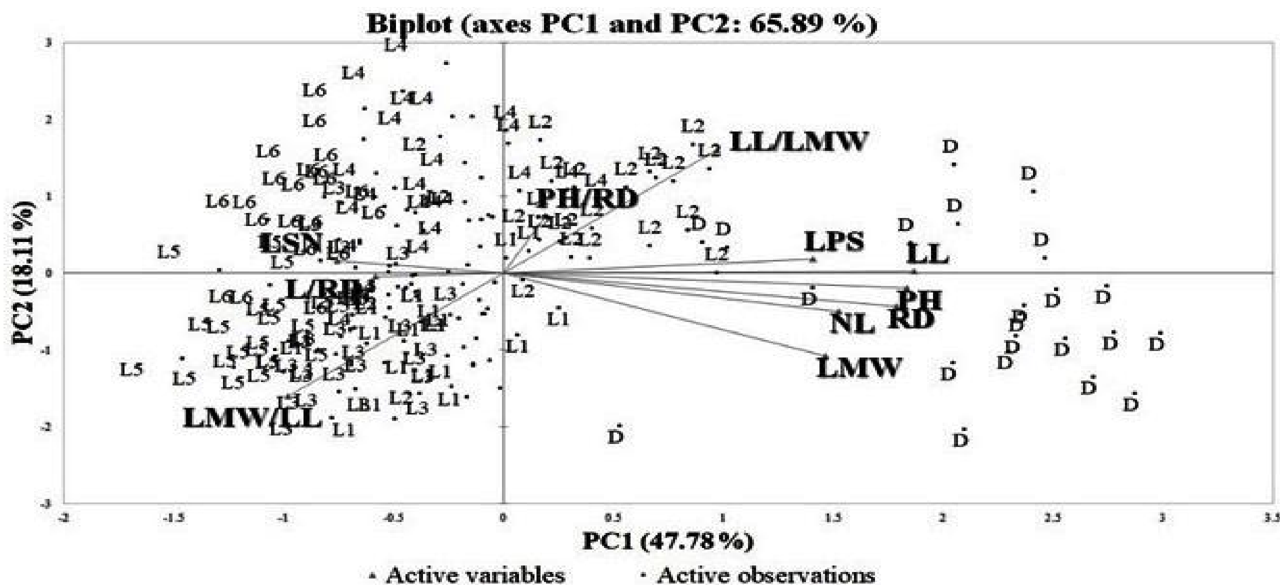


**Figure 1. Biplot of the first two axes of the Principal Components Analysis based on eight morphological measurements, comparing cultivated (six lines: L1 - L6) and one wild individuals (D) of *Agave durangensis***

**Table 2. Carbohydrate content of leaf and stem of wild and cultivated *Agave durangensis* populations.**

| Line | Leaf | | | Stem | | |
|---|---|---|---|---|---|---|
| | TCC (mg/g tissue) | FRS (mg/g tissue) | FOS (mg/g tissue) | TCC (mg/g tissue) | FRS (mg/g tissue) | FOS (mg/g tissue) |
| 1 | 80.2 ± 23.8 a | 12.9 ± 2.8 b | 67.3 ± 24.9 a | 12.0 ± 0.9 c | 8.5 ± 0.8 ab | 3.5 ± 1.3 c |
| 2 | 95.4 ± 20.2 a | 12.3 ± 2.3 b | 83.4 ± 50.4 a | 27.7 ± 4.3 a | 6.0 ± 1.2 c | 21.6 ± 4.4 a |
| 3 | 79.5 ± 35.4 a | 10.6 ± 3.9 bc | 68.8 ± 33.3 a | 16.6 ± 4.8 b | 5.1 ± 2.5 c | 11.5 ± 4.6 b |
| 4 | 94.5 ± 44.1 a | 6.7 ± 1.7 d | 87.8 ± 44.2 a | 21.9 ± 6.6 a | 6.9 ± 3.4 bc | 15.0 ± 5.1 b |
| 5 | 68.9 ± 31.6 a | 9.3 ± 3.8 cd | 60.5 ± 34.6 a | 16.2 ± 6.2 b | 5.5 ± 1.3 bc | 10.4 ± 6.8 b |
| 6 | 98.6 ± 37.0 a | 8.4 ± 3.2 cd | 90.1 ± 39.4 a | 21.3 ± 9.6 a | 11.7 ± 4.7 a | 9.5 ± 6.2 bc |
| D | 24.5 ± 5.0 b | 18.3 ± 4.2 a | 6.2 ± 3.5 b | ND | ND | ND |

Mean value ± standard deviation, $^*p \leq 0.05$, Different letters in the same column mean statistical differences among populations, ND = Not Determined

The results of a PCA based on the carbohydrate contents (Figure 2) showed that the first component represented 95.60 % of the total variance, being the variables FOS and TSC those with more weight on the component. A nonspecific distribution was observed in the PCA over the six cultivated lines of *A. durangensis*, contrary to the observed by the morphological analysis. TCC leaf and FOS leaf showed a significant correlation (Table 3) ($r = 0.944$, $p \leq 0.0001$), and FOS stem and FOS stem also showed a significant correlation ($r = 0.899$, $p \leq 0.0001$). No correlation between leaf and stem carbohydrate was shown.

**Genetic variability and genetic structure:** of the seven original microsatellites used to test genetic diversity and structure of cultivated and wild individuals of *A. durangensis*, no evidence of genotyping errors were present, according to the microchecker analysis. Genetic descriptive statistics are shown in table 4. According to Friedman´s method, statistical differences were present among the evaluated lines for all genetic measured indexes ($p \leq 0.05$). Cultivated individuals of *A. durangensis* showed average values of Na of 16.80 and wild population of 16.14 respectively, using the same markers as Lindsay *et al.,* (2012) who reported Na values

from 1 to 16 for two populations of *Agave palmeri*. The expected heterozygosity (He) was 0.76 for the wild population, whereas, for the cultivated lines the He ranged from 0.75 to 0.90 (mean= 0.73), revealing no excess of homozygotes among the individuals under study. No differences were observed between wild and cultivated *A. durangensis* individuals according to the He values, suggesting a current domestication process. These results agreed with those reported by Figueredo *et al*., (2015), who informed no differences in the allelic frequency level using SSRs markers between wild and cultivated populations of *Agave inaequidens*, probably because of the short time of domestication. This effect is explained with the results of the fixation index (Fis), which were lower than zero (Fis = -0.08 for cultivated and Fis =-0.12 for wild), indicating an excess of heterozygotes.

Nevertheless, the L3 individuals showed values of Fis = 0.02, slightly higher than zero, which could be indicating a possible inbreeding. The current results are in agreement with those of Parker *et al.,* (2014) for a cultivated population of *Agave parryi*, which showed positive values of Fis, contrary to the negative values showed from wild populations var *huachucensis*. Cultivated lines of *A. durangensis* showed a range of Fst values from 0.03 (L3) to 0.08 (L4), with a mean of Fst = 0.05, indicating a strong genetic similarity among them. Values of Fst for both cultivated and wild populations of *A. durangensis* were similar to other *Agave* species as previously reported for several authors such *A. subsimplex* (0.084), *A. deserti* (0.135), *A. cerulata* (0.098), *A. angustifolia* (0.175) (Navarro-Quezada *et al*., 2003; Sánchez-Teyer *et al*., 2009). The AMOVA (Table 5) based on alleles frequencies showed that the major genetic variation of *A. durangensis* was found within individuals (84.33%) and not among groups (0.77 %), and the genetic variation among population within groups was of 6.14 %. This last issue agreed with both Fis and Fst values, indicating a little genetic differentiation among cultivated lines.

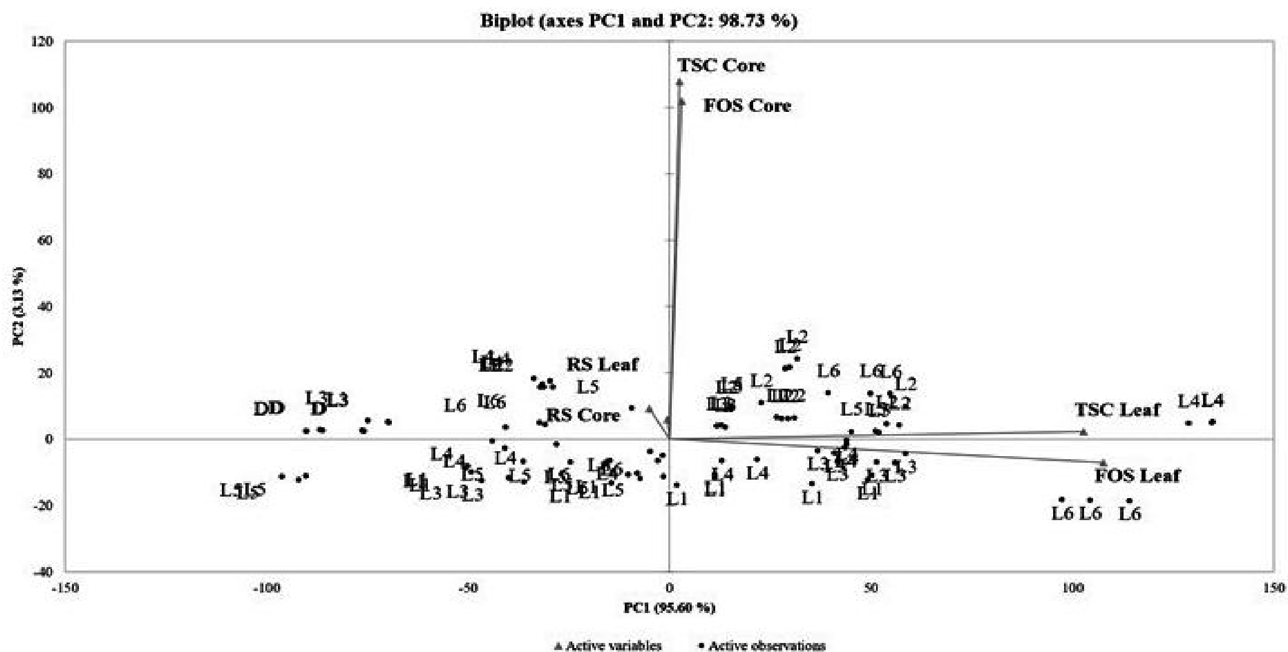


**Figure 2. Biplot of the first two axes of the Principal Components Analysis based on Carbohydrate content measurements, comparing cultivated (L1 - L6) and one wild population (D) of *Agave durangensis***

**Table 3. Pearson correlation matrix (Below the diagonal) and p values (Above diagonal) between carbohydrate content of cultivated lines and a wild population of *Agave durangensis.***

| Variables | TCC Leaf | FRS Leaf | FOS Leaf | TCC Stem | FRS Stem | FOS Stem |
|---|---|---|---|---|---|---|
| TCC Leaf | | 0.213 | **≤ 0.0001** | 0.188 | 0.579 | 0.103 |
| FRS Leaf | -0.137 | | **0.027** | 0.484 | 0.252 | 0.839 |
| FOS Leaf | **0.994** | **-0.242** | | 0.225 | 0.505 | 0.115 |
| TCC Stem | 0.145 | 0.077 | 0.134 | | **0.004** | **≤ 0.0001** |
| FRS Stem | -0.061 | 0.126 | -0.074 | **0.308** | | 0.208 |
| FOS Stem | 0.179 | 0.023 | 0.173 | **0.899** | -0.139 | |

**Values in **bold** indicate statistical significance with values of $p \leq 0.05$.

**Table 4. Summary of genetic variability estimates from wild and six cultivated morphological lines of *Agave durangensis* based on seven microsatellite loci.**

| | Na | Ne | He | Fis | Fst | PIC |
|---|---|---|---|---|---|---|
| **Cultivated lines** | | | | | | |
| ***A.durangensis*** | | | | | | |
| 1 | 18.43$^{ab}$ | 8.00$^{a}$ | 0.90$^{ab}$ | -0.02$^{a}$ | 0.04$^{ab}$ | 0.75$^{a}$ |
| 2 | 19.43$^{ab}$ | 7.78$^{a}$ | 0.87$^{abc}$ | -0.01$^{a}$ | 0.06$^{ab}$ | 0.73$^{a}$ |
| 3 | 12.71$^{d}$ | 7.07$^{b}$ | 0.80$^{bc}$ | 0.02$^{a}$ | 0.03$^{c}$ | 0.69$^{b}$ |
| 4 | 17.00$^{cd}$ | 8.35$^{a}$ | 0.75$^{c}$ | -0.26$^{d}$ | 0.08$^{ab}$ | 0.75$^{a}$ |
| 5 | 17.57$^{bc}$ | 8.28$^{a}$ | 0.77$^{c}$ | -0.11$^{c}$ | 0.05$^{ab}$ | 0.74$^{a}$ |
| 6 | 15.71$^{cd}$ | 8.42$^{a}$ | 0.82$^{bc}$ | -0.04$^{b}$ | 0.06$^{b}$ | 0.75$^{a}$ |
| Mean | 16.8±2.36 | 7.9±0.50 | 0.73±0.23 | -0.08±0.16 | 0.05±0.01 | 0.73±0.02 |
| Wild population | | | | | | |
| *A. durangensis* | 16.14±4.63$^{bc}$ | 8.14±2.01$^{a}$ | 0.76±0.20$^{c}$ | -0.12±0.24$^{c}$ | 0.13±0.08$^{d}$ | 0.73±0.10$^{a}$ |

Na=number of alleles, Ne=effective number of alleles, He=expected heterozygosity, Fis=fixation index, PIC= Polymorphic Index Content.

**Table 5. Analysis of Molecular Variance using F statistics for six cultivated lines and one wild population of *Agave durangensis*. $^{**}p \leq 0.05$.**

| Source of variation | Sum of Squares | Variance Component | Percentage Variation | F statistics |
|---|---|---|---|---|
| Among groups | 44.52 | 0.02 | 0.77 | 0.093** |
| Among populations within groups | 50.28 | 0.21 | 6.14 | 0.061** |
| Among individuals within populations | 571.25 | 0.30 | 8.74 | 0.007** |
| Within individuals | 496.15 | 2.92 | 84.33 | 0.156** |
| Total | 1162.58 | 3.46 | | |

Genetically similar groups were determined using an admixture model-based clustering analysis implemented in the software STRUCTURE. According to Evanno´s test (Evanno *et al*., 2005), the most informative number of subpopulations (*K*) was four, with a value of ΔK = 0.408 (Figure 3A). The inferred population structure is presented in Figure 3B, in which the proportion (*p*) of each individual to belong to one or more of the four possible sub-populations is displayed. Defined groups represent different sub-populations (Fst = 0.033, p≤0.05). Wild *A. durangensis* formed a single group. The observed sub-population structure indicated that the cultivated populations of *A. durangensis* could be clearly grouped in three sub-populations. A similar discrimination between wild and cultivated forms of a single species, had been reported for other *Agave* species as *A. angustifolia*, using ISSRs markers; both wild and cultivated populations having similar genetic diversity indexes (He= 0.442 and He= 0.428, respectively) but different values of Weir & Cockerham index(θ) (an equivalent for Fst statistics) (θ= 0.405 and θ= 0.212 for cultivated and wild populations, respectively) (Vargas-Ponce *et al.,* 2009). Cultivated populations of *A. angustifolia*, showed higher genetic differentiation compared to wild one, this difference may be the result of the domestication process to which the former has been submitted for a long time. Wild and cultivated forms of *Agave hookeri*, *cupreata*, and *inaequidens* also have been distinguished using SSR markers (Figueredo-Urbina *et al.,* 2017). Molecular markers, as AFLP, served to discriminate between varieties of *A. tequilana* (Gil-Vega *et al.,* 2006). To the best of our knowledge, the genetic structure of *A. durangensis* (either cultivated or wild) has not been assessed. The results of the present study could sustain the implementation of control, protection, and propagation strategies, providing morphological and genetic tools to distinguish groups with outstanding attributes.

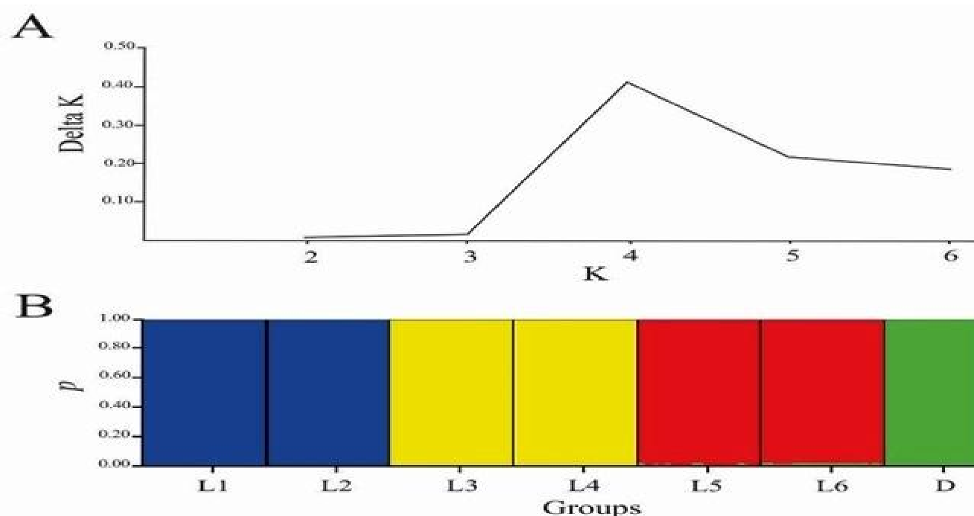


**Figure 3. (A) Estimated number of genetic groups (K) for cultivated and wild *Agave durangensis* using Delta K, calculated using Evanno's method. (B) The inferred structure formed by four genetic clusters (blue, yellow, red, green), represents the proportion of individuals in each genetic group.**

**Correlation among variable groups:** The Mantel test showed a positive and significant correlation between morphological and genetic variations ($r= 0.713$, $p \leq 0.001$, $R^2= 0.314$) based on the matrix of similarity distance, determined from morphological traits, and Fst genetic identity estimated by SSR markers (Fig. 3A). In agreement with the present results, Figueredo-Urbina *et al*., (2017) reported a close relation between morphological and genetic features for *A. hookeri* and *A. inaequidens*. Both are cultivated species, used for economic endings, the knowledge about wild populations is scare. Nevertheless, for other *Agave* species, low correlations between morphological and molecular traits were found as *A. tequilana* and *A. angustifolia* suggesting that both characteristics are independent and that molecular markers should not replace morphological characterization; however, they may be an important complement (Rodríguez-Garay *et al*., 2009). In contrast with the present results, Almaraz-Abarca *et al.,* (2013) reported incongruences among genetic and morphological variability on three wild populations of *A. durangensis* using ISSR markers. Some authors attributed a low correlation between morphological and genetic markers to the type of primer used, because the sequences might be originated from the intron regions of the genome (Hashemifar and Rahimmalek, 2018).

In the current study, the three independent groups formed for the cultivated individuals is an indicator of the morphological variability that is present in wild populations of *A. durangensis*. According to García-Mendoza *et al*., (2007), one single *Agave* plant is capable of producing around 65,000 seeds, the high genetic information contained in them favors the conservation and protection of the species (Ramírez-Tobías *et al*., 2012, 2014). However, this variability may decrease due to the overexploitation and to the fact that the more prominent individuals are harvested; the same has been reported for *Agave potatorum* Zucc., *A. seemanniana*, *A. nussaviorum* subsp. nussaviorum*, A. angustifolia* Haw., *A. marmorata* Roezl., and *A. karwinskii* Zucc., var. Americana (Félix-Valdez *et al*., 2016; Porras Ramírez *et al*., 2016). These harvest patterns could contribute to decreasing the genetic pool that eventually, could fragment and loss of the genetic identity of the species.

Mantel test also showed a significant correlation between carbohydrate content and morphological traits ($r = 0.571$, $p = 0.003$, $R^2 = 0.238$), FRS has similar behavior to that of the morphological features and Fst (Fig. 4B). No correlation was found between the carbohydrate content and the genetic variability of *A.durangensis* determined with SSR markers ($r = 0.219$, $p = 0.121$, $R^2 = 0.047$) (Fig. 4C). It is needed to evaluate an association between carbohydrate content and the variability of other DNA sequences, as some authors had reported an association using carbohydrate-associated genes, like 1-FFT and 6G-FFt (Suárez-González *et al*., 2014).

Finally, we evaluated the influence of the carbohydrate content over the correlation between genetic and morphological traits, showing no effect ($r = 0.399$, $p = 0.061$) (Fig. 4D).

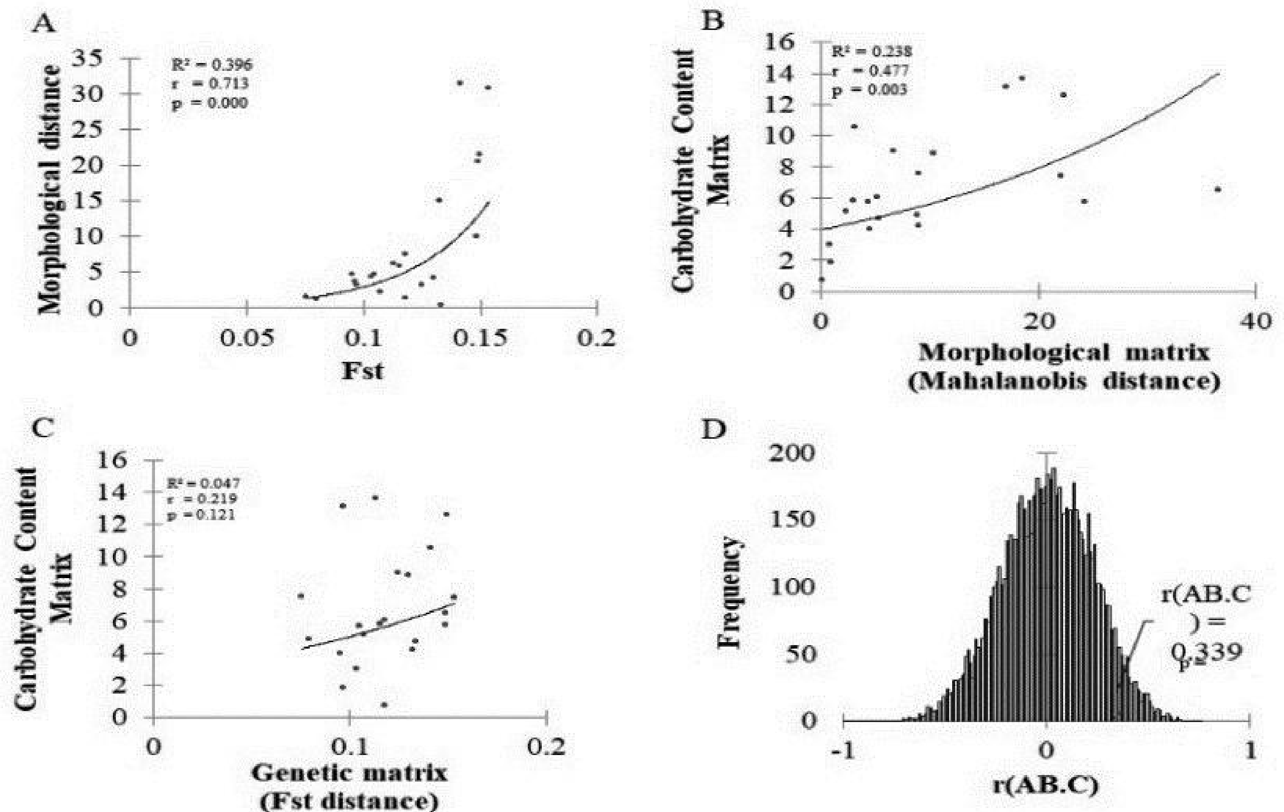


**Figure 4. Mantel test of different similarity matrices obtained from cultivated and wild *Agave durangensis*. Effect sizes were reported as correlation values (r), and *p* values were computed with 10000 permutations. (A) Correlation between morphological similarity and pairwise Fst matrices. (B) Correlation between carbohydrates (Free Reducing Sugars) similarity and morphology similarity matrices. (C) Correlation between carbohydrates similarity and pairwise Fst matrices. (D) Effect of the carbohydrates similarity matrix over the correlation between morphological similarity and pairwise Fst matrices of *Agave durangensis***

**Conclusion:** Morphological characters can discriminate between wild and cultivated individuals of *Agave durangensis.* A single plant can generate a highly variable offspring, allowing to distinguish six different morphological lines. But the most important result is the significant relation between morphologically related individuals and genetic variables; this is a valuable relation, which has important conservation and breeding implications. Furthermore, the correlation between carbohydrate content and morphological characters could be an important tool to identify outstanding individuals for mescal production. The current results showed the genetic potential of this species and the valuable potential of the morphological, chemical, and molecular markers used to develop propagation and conservation programs.

**Acknowledgements:** Authors thank Commission operation and promotion of the academic activities of the National Polytechnic Institute and Council state of science and technology of the State of Durango for the grants SIP20195620, SIP20195272 and SIP2017-RE/063 provided for the development of the present research.

## REFERENCES

Addinsoft (2019). XLSTAT statistical and data analysis solution. Long Island, NY, USA. https://www.xlstat.com

Almaraz-Abarca, N., E. A. Delgado-Alvarado, M. I. Torres-Morán, J. Herrera-Corral, J. A. Ávila-Reyes, N. Naranjo-Jiménez and J. N. Uribe-Soto (2013). Genetic variability in natural populations of *Agave durangensis* (Agavaceae) revealed by morphological and molecular traits. The Southwestern Naturalist, 58 (3): 314-324.

Arrizon, J., S. Morel, A. Gschaedler and P. Monsan (2010). Comparison of the water-soluble carbohydrate composition and fructan structures of *Agave tequilana* plants of different ages. Food Chemistry, 122 (1): 123-130.

Balzarini, M. and J. Di Rienzo (2003). Info-Gen: Software for statistical analysis of genetic data. Faculty of Agricultural Science. Universidad Nacional de Córdoba. Argentina.

Colunga-GarcíaMarín, P., and F. May-Pat (1997). Morphological variation of henequen (*Agave fourcroydes*, Agavaceae) germplasm and its wild ancestor (*A. angustifolia*) under uniform growth conditions: diversity and domestication. American J. Botany, 84 (11):1449-1465.

Diniz-Filho, J. A. F., T. N. Soares, J. S. Lima, R. Dobrovolski, V. L. Landeiro, M. P. D. C. Telles, T. F. Rangel and L. M. Bini (2013). Mantel test in population genetics. Genetics and Molecular Biology, 36 (4): 475-485.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4 (2): 359-361.

Evanno, G., S. Regnaut and J. Goudet (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14 (8): 2611-2620.

Excoffier, L., P. E. Smouse and J. M. Quattro (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, 131 (2): 479-491.

Excoffier, L., and H. E. Lischer (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10 (3): 564-567.

Félix-Valdez, L.I., O. Vargas-Ponce, D. Cabrera-Toledo, A. Casas, A. Cibrian-Jaramillo and L. de la Cruz-Larios (2016). Effects of traditional management for mescal production on the diversity and genetic structure of *Agave potatorum* (Asparagaceae) in central Mexico. Genetic Resources and Crop Evolution, 63 (7): 1255-1271.

Figueredo-Urbina, C.J., A. Casas and I. Torres-García (2017). Morphological and genetic divergence between *Agave inaequidens*, *A. cupreata* and the domesticated *A. hookeri*. Analysis of their evolutionary relationships. PloS one, 12 (11): e0187260.

Figueredo, C. J., A. Casas, A. González-Rodríguez, J. M. Nassar, P. Colunga-GarcíaMarín and V. Rocha-Ramírez (2015). Genetic structure of coexisting wild and managed agave populations: implications for the evolution of plants under domestication. AoB Plants, 7: https://doi.org/10.1093/aobpla/plv114

Figueredo, C. J., and J. M. Nassar (2011). Population genetics of *Agave cocui*: evidence for low genetic diversity at the southern geographic limit of genus *Agave*. J. Heredity, 102 (3): 306-314.

Fritsche, S., X. Wang, J. Li, B. Stich, F. J. Kopisch-Obuch, J. Endrigkeit, G. Leckband, F. Dreyer, W. Friedt, J. Meng and C. Jung (2012). A candidate gene-based association study of tocopherol content and composition in rapeseed (*Brassica napus*). Frontiers in Plant Sci., 3:129

García-Mendoza, A. (2007). Los agaves de México. Ciencias (87): 14-23.

Gil-Vega, K., C. Díaz, A. Nava-Cedillo and J. Simpson (2006). AFLP analysis of *Agave tequilana* varieties. Plant Science, 170 (4): 904-909.

González, G., S. Alemán and D. Infante (2003). Asexual genetic variability in *Agave fourcroydes* II: selection among individuals in a clonally propagated population. Plant Science, 165 (3), 595-601.

Hashemifar, Z. and M. Rahimmalek (2018). Genetic structure and variation in *Perovskia abrotanoides* Karel and *P. atriplicifolia* as revealed by molecular and morphological markers. Scientia Horticulturae, (230): 169-177.

Jourdan, M., Gagné, S., C. Dubois-Laurent, M. Maghraoui, S. Huet, A. Suel, L. Hamama, M. Briard, D. Peltier and E. Geoffriau (2015). Carotenoid content and root color of cultivated carrot: a candidate-gene association study using an original broad unstructured population. PloS one, 10 (1): e0116674.

Keb-Llanes, M., G. González, B. Chi-Manzanero and D. Infante (2002). A rapid and simple method for small-scale DNA extraction in Agavaceae and other tropical plants. Plant Molecular Biology Reporter, 20 (3): 299-300.

Legendre, P., and L. F. Legendre (2012). Numerical ecology (24): Elsevier Radarweg 29, P O Box 211, 1000 AE Amsterdam, The Netherland.

Lindsay, D. L., C. E. Edwards, M. G. Jung, P. Bailey and R. F. Lance (2012). Novel microsatellite loci for *Agave parryi* and cross-amplification in *Agave palmeri* (Agavaceae). American J. Botany, 99 (7):e295-e297.

Mamuris, Z., P. Apostolidis, P. Panagiotaki, A. J. Theodorou and C. Triantaphyllidisdir (1998). Morphological variation between red mullet populations in Greece. J. Fish Biology, 52 (1): 107-117.

Michel-Cuello, C., B. I. Juárez-Flores, J. R. Aguirre-Rivera and J. M. Pinos-Rodríguez (2008). Quantitative characterization of nonstructural carbohydrates of mezcal *Agave* (*Agave salmiana* Otto ex Salm-Dick). J. Agri. Food Chemistry, 56 (14): 5753-5757.

Montañez-Soto, J., J. Venegas-González, M. Vivar-Vera and E. Ramos-Ramírez (2011). Extraction, characterization and quantification of fructans contained in the head and leaves of *Agave tequilana* Weber blue. Bioagro, 23 (3):199-206

Nava-Cruz, N.Y., M. A. Medina-Morales, J. L. Martinez, R. Rodriguez and C. N. Aguilar (2015). Agave biotechnology: an overview. Critical Reviews in Biotechnology, 35 (4): 546-559.

Navarro-Quezada, A., R. González-Chauvet, F. Molina-Freaner and L. Eguiarte (2003). Genetic differentiation in the *Agave deserti* (Agavaceae) complex of the Sonoran Desert. Heredity, 90 (3): 220-227.

Parker, K.C., D. W. Trapnell, J. Hamrick and W. C. Hodgson (2014). Genetic and morphological contrasts between wild and anthropogenic populations of *Agave parryi* var. *huachucensis* in south-eastern Arizona. Annals of Botany, 113 (6): 939-952.

Pavel, A.B., and C. I. Vasile (2012). PyElph-a software tool for gel images analysis and phylogenetics. BMC Bioinformatics, 13 (1): 9.

Porras Ramírez, E., Y. Sosa Marcos, F. Palma-Cruz and Cisneros (2016). Phenotyping the genetic diversity of wild *Agave* species that coexist in the same spatial region. Notulae Botanicae Horti Agrobotanici Cluj-Napoca, 44 (2): 640-648

Pritchard, J.K., M. Stephens and P. Donnelly (2000). Inference of population structure using multilocus genotype data. Genetics, 155 (2): 945-959.

Ramírez-Tobías, H.M., C. B. Peña-Valdivia, R. Aguirre J. A. Reyes-Agüero, A. B. Sánchez-Urdaneta and G. S. Valle (2012). Seed germination temperatures of eight Mexican Agave species with economic importance. Plant Species Biology, 27 (2): 124-137.

Ramírez-Tobías, H. M., C. B. Peña-Valdivia, C. Trejo and H. Vaquera (2014). Seed germination of *Agave* species as influenced by substrate water potential. Biological Research, 47 (1): 11.

Rodríguez-Garay, B., J. Lomelí-Sención, E. Tapia-Campos, A. Gutiérrez-Mora, J. García-Galindo, J. Rodríguez-Domínguez, D. Urbina-López and I. Vicente-Ramírez (2009). Morphological and molecular diversity of *Agave tequilana* Weber var. Azul and *Agave angustifolia* Haw. var. Lineño. Industrial Crops and Products, 29 (1): 220-228.

Rousset, F. (1995). GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicalism. J. Heredity, (83): 239.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources, 8 (1): 103-106.

Sánchez-Teyer, F., S. Moreno-Salazar, M. Esqueda, A. Barraza and M. Robert (2009). Genetic variability of wild *Agave angustifolia* populations based on AFLP: A basic study for conservation. J. Arid Environments, 73 (6): 611-616.

Suárez-González, E.M., M. G. López, J. P. Délano-Frier and J. F. Gómez-Leyva (2014). Expression of the 1-SST and 1-FFT genes and consequent fructan accumulation in *Agave tequilana* and *A. inaequidens* is differentially induced by diverse (a) biotic-stress related elicitors. J. Plant Physiology, 171 (3-4): 359-372.

Torres, I., A. Casas, E. Vega, M. Martínez-Ramos and A. Delgado-Lemus (2015). Population dynamics and sustainable management of mescal agaves in central Mexico: *Agave potatorum* in the Tehuacán-Cuicatlán Valley. Economic Botany, 69 (1): 26-41.

Van Oosterhout, C., W. F. Hutchinson, D. P. Wills and P. Shipley (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Resources, 4 (3): 535-538.

Vargas-Ponce, O., D. Zizumbo-Villarreal, J. Martínez-Castillo, J. Coello-Coello and P. Colunga-GarcíaMarín (2009). Diversity and structure of landraces of *Agave* grown for spirits under traditional agriculture: a comparison with wild populations of *A. angustifolia* (Agavaceae) and commercial plantations of *A. tequilana*. American J. Botany, 96 (2): 448-457.