

Genetic diversity and structure of *Magnolia mexicana* (Magnoliaceae): A threatened species in eastern Mexico

Marisol Gutiérrez-Lozano¹, Arturo Sánchez-González^{1*}, Pablo Octavio-Aguilar¹, Dulce María Galván-Hernández¹, José Antonio Vázquez-García²

¹ Centro de Investigaciones Biológicas, Universidad Autónoma del Estado de Hidalgo, Ciudad del Conocimiento; Carretera Pachuca-Tulancingo Km. 4.5, Mineral de la Reforma, C.P. 42184, Hidalgo, México.

² Centro Universitario de Ciencias Biológicas y Agropecuarias, Departamento de Botánica y Zoología, Herbario IBUG, Universidad de Guadalajara, Carretera Guadalajara-Nogales Km. 15.5. Camino Ramón Padilla Sánchez 2100, Nextipac, 45110 Zapopan, Jalisco, México.

* Correspondence author: Arturo Sánchez-González, e-mail: arturosg@uaeh.edu.mx

Abstract

This study analyzes the diversity and genetic structure of seven natural populations of *Magnolia mexicana* (Yoloxóchitl), an endangered and endemic species distributed in the Mexican Sierra Madre Oriental. The objective of the analysis is to generate basic information for the development of management and conservation programs for this species. The variation and genetic structure were estimated based on 12 SSR markers. The results indicate that the expected heterozygosity was 0.63 ± 0.02 and the allelic richness was 1.21 ± 0.05 . The analysis of molecular variance (AMOVA) showed that variation between individuals explains 49 % of the total variation, and variation within individuals explains 36 %. The technical Neighbor joining and the ordination diagram of the canonical discriminant analysis show that the populations are made up of two genetic groups mostly. By contrast with the model Bayesian grouping (obtained with the STRUCTURE program) indicated the formation the six genetics groups. Recent bottleneck effects were observed in all localities, except in Duraznillo, El Cajón, and Zapotla. The relationship between genetic and geographic distances was not significant, which rules out a structure determined by isolation; the results suggest that the diversity and genetic structure of the species are affected by local factors such as habitat fragmentation, land use change, and excessive use of the flowers and fruits of *M. mexicana* by the inhabitants. Therefore, it is a priority to implement management and conservation programs to guarantee their survival.

Keywords: *Extinction risk, Magnoliaceae, plant conservation, population genetics, SSR markers, Yoloxóchitl.*

Introduction

The genus *Magnolia* L. (Magnoliaceae) includes more than 300 tree shrub species, grouped into 15 sections based on their morphology, and whole plastid genome sequences (Wang et al., 2020), one of the them is *Magnolia* sect. *Talauma*; (Juss.) Baill. Figlar (2006), recognized three subsections between this section: *Cubensis*, *Dugandiodendron*, and *Talauma*. The subsection *Talauma* comprises ca. 31 species, exclusive to the Neotropics, and distributed from Mexico to southeastern Brazil, including the Caribbean Island (Vázquez-García et al., 2013). *Magnolia mexicana* (DC.) G. Don (*Magnolia* sect. *Talauma* subsect. *Talauma*) commonly known as Yoloxóchitl, is an evergreen tree that can reach up to 40 m in height and a trunk diameter of 1.35 m. It presents solitary flowers with numerous stamens and a superior ovary composed of carpels in a helical arrangement. The leaves are simple and spirally arranged (Arteaga-Ríos et al., 2020; Gutiérrez-Lozano et al., 2021).

M. mexicana is endemic to eastern Mexico, where it has several traditional uses, mainly medicinal; the flowers and bark are used to treat heart conditions (Waizel-Bucay, 2002). An infusion of the bark is used to treat arterial hypertension, the flowers are used as antispasmodics and to flavor chocolate, and the seeds to treat paralysis. It is also used for ritual purposes and as an ornamental plant in patios, parks, and gardens, due to its exuberant foliage and the beauty and aroma of its flowers (Pennington and Sarukhán, 1998). Recently, the fruits of *M. mexicana* have become an important product for local trade, due to seed extraction for medicinal purposes.

The natural distribution of *M. mexicana* is restricted to the states of Puebla and Veracruz. It grows in hot and humid sub-warm climates in transition zones between mountain cloud

forest (MCF) and evergreen tropical forest (ETF) (Aldaba-Núñez et al., 2021). Because this species faces multiple threats, such as habitat fragmentation and excessive collection of flowers and seeds, coupled with poor regeneration capacity, its populations are highly fragmented and vary substantially in density and occupied area, which is reflected in significant loss of genetic information (Gutiérrez-Lozano et al., 2021). According to its biological characteristics, distribution and survival status, *M. mexicana* is classified in the category of “Endangered” in the Mexican standard NOM-059-SEMARNAT-2010 (SEMARNAT, 2010), “Vulnerable” (VU) in the Red List of the International Union for Conservation of Nature (Rivers et al., 2016), and at a low vulnerability level in the Global Conservation Gap Analysis of *Magnolia* (Linsky et al., 2022).

Molecular markers enable precise estimates of the diversity and genetic structure of species. The use of microsatellites or SSR (simple sequence repeat) markers has some advantages, including high polymorphism, codominance and relatively low cost (Agarwal et al., 2008), which is why they have been widely used in the evaluation of intraspecific genetic diversity and to recognize possible geographic patterns of differentiation in different magnolia species; *Magnolia acuminata*, *M. ashei*, *M. cubensis*, *M. kobus*, *M. schiedeana*, *M. sharpii* and *M. tripelata*, among others (Budd et al., 2015; Hernández et al., 2020; Newton et al., 2008; Quintana-Delgado et al., 2022; Tamaki et al., 2016; Veltjen et al., 2018; von Kohn et al., 2018; Wang et al., 2019). The results of these studies have generated basic data necessary for decision-making and for the formulation of conservation and protection plans for these species.

Genetic variability is the raw material for preservation of the evolutionary potential of the species in the face of changes in the environment hence, in endemic and endangered plant species such as *Magnolia mexicana*, it is essential to generate information about genetic variation within and between populations to enable formulation of management and conservation strategies. Therefore, the objectives of the present study were: 1) to evaluate the diversity and genetic structure of *Magnolia mexicana* in different localities, throughout the natural distribution area of the species, based on SSR markers; and 2) to estimate whether the genetic distance between these localities is associated with the geographic distance (variation in local environmental conditions by distance), and what implications it could have for management and conservation of this species.

Materials and methods

Study zone and data collection

The distribution of natural populations of *M. mexicana* is limited to the southern Sierra Madre Oriental, Mexico (Vázquez-García et al., 2013). Its ideal habitat is the transition zone between the evergreen tropical forest (ETF) and the montane cloud forest (MCF). The ETF is characterized by its high density of plant species. It has a type A climate with total annual rainfall between 1,100 and 1,300 mm. The average annual

temperature varies between 22 and 26°C, while the lowest temperature is 18 °C, with no frost (Pennington and Sarukhán, 1998). The MCF has a C(f) climate, temperate with rain all year round, with the presence of mist, so humidity is high, between 55 and 65 %. The total annual precipitation ranges between 1,048 and 2,385 mm and the average annual temperature between 17.1 and 20.7°C. The elevation distribution interval of the MCF is 600 to 2,500 m a.s.l (Ruiz-Jiménez et al., 2012).

The seven *M. mexicana* localities analyzed are distributed in the states of Puebla and Veracruz; five of them were located based on references (research articles and the NaturaLista website; <https://www.naturalista.mx/>), and the rest were recorded for the first time, based on field exploration in localities near known presence sites. The study locations were: 1. El Cajón (Xicotepéc de Juárez), 2. Las Margaritas (Hueytamalco), 3. La Tuza (Tuzamapan de Galena), 4. Duraznillo (Quimixtlan) in Puebla; 5. Vista Hermosa (Cosautlán de Carvajal), 6. Chocamán (Córdoba) and 7. Zapotla (Zongolica) in Veracruz (Figure 1, Table 1).

During May and June of 2019 and 2020, leaf tissue was collected from 6 to 13 individuals per population, depending on the population size, since in some locations there were few trees. The samples obtained were placed in plastic bags to which silica beads were added for desiccation, and they were stored in a freezer at –20 °C until they were processed in the laboratory.

Extraction and Amplification

To obtain DNA, the cetyltrimethylammonium bromide protocol was used (Doyle and Doyle, 1987), with modifications to eliminate secondary metabolites. 0.5 grams of tissue was pulverized with liquid nitrogen to a fine powder which was placed in a 2 mL microtube. Then 1 mL of 2× CTAB-PVP buffer and 4 µL of β-mercaptoethanol were added to the microtube and the solution was incubated at 37°C for 60 minutes, after which 10 µL of proteinase K and 10 µL of RNA'se (1 mg /1Lm) were added.

Subsequently, 600 µL of chloroform-octanol solution (24:1 v/v) was added to the preparation and it was centrifuged at 12,000 rpm for 10 minutes. The supernatant was recovered in a new 1.5 mL microtube with 400 µL of isopropanol. The sample was frozen-precipitated at –20°C for 12 hours, after which time it was centrifuged at 12,000 rpm for 10 minutes to obtain the pellet, which was washed with 1 mL of 70 % ethanol and centrifuged at 14,000 for 10 minutes. Finally, the pellet was dried and resuspended with 50 µL TE buffer and stored at –20 °C.

DNA quality was determined by 1 % agarose gel electrophoresis, with MIDORI Green Advance staining (Nippon Genetics, Japan). Subsequently, the DNA was quantified by spectrophotometry at 260 nm (A260) and the purity was obtained by means of the A260/A280 and A260/A230 ratio. The purified and quantified samples were stored in a freezer at –20°C. Twelve microsatellite markers (MA39_159, MA39_182, MA39_185, MA39_236, MA39_342, MA39_348, MA40_282, MA41_076, MA41_373, MA42_028, MA42_231 and MA42_471; Table 2) developed by Veltjen et al. (2018) for *M. lacandonica*, *M. mayae*, *M. dealbata* and *M. cubensis* subsp. *acunae* were used.

Table 1

General characteristics of the study localities and sample sizes of *Magnolia mexicana* populations

Municipality	Locality	Sample size	Latitude (N)	Longitude (W)	Elevation (m)	AAT (°C)	TAP (mm)
Puebla state							
Xicotepec de Juárez	El Cajón	10	20° 18'48.0"	-97° 56'11.1"	826	25	2,300
Tuzamapan de Galena	La Tuza	6	20° 03'58.9"	-97° 34'66.6"	626	26	3,000
Hueytamalco	Las Margaritas	11	19° 59'0.18"	-97° 18'96.4"	614	26	2,500
Quimixtlan	Duraznillo	11	19° 16'53.5"	-97° 00'27.7"	1,205	20	1,700
Veracruz state							
Cosautlán de Carvajal	Vista Hermosa	8	19° 19'35.3"	-96° 59'14.4"	1,236	22	2,000
Córdoba	Chocamán	8	19° 00'29.7"	-97° 01'54.7"	1,360	19	1,844
Zongolica	Zapotla	13	18° 38'34.59"	-98° 58'44.60"	1,141	26	2,700

Average annual temperature (AAT), total annual precipitation (TAP)

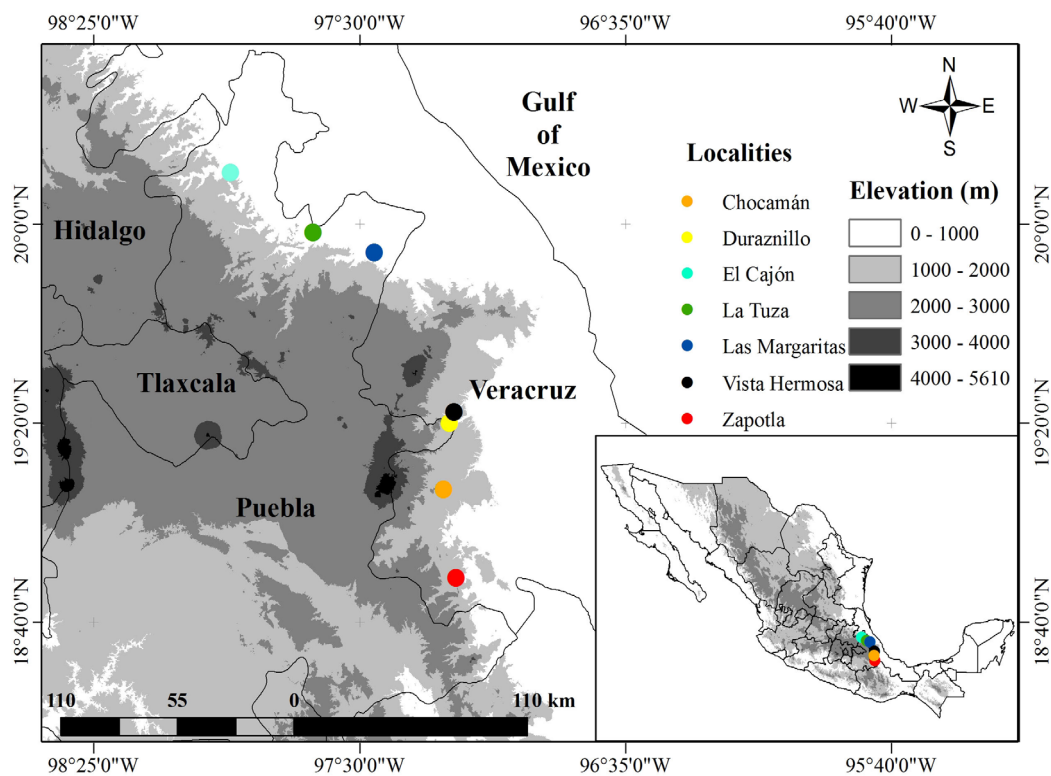


Figure 1

Geographic locations of *Magnolia mexicana* populations in the Sierra Madre Oriental, Mexico.

Table 2

Molecular markers used for genetic analysis of *M. mexicana* localities (proposed by Veltjen et al., 2018).

Locus	Forward Sequence	Reverse Sequence	Number of alleles	Structure	Ta (observed)	Allele size (bp)
MA39_159	F: ATCAGGAGTGTAACGCCACC	R: GCGGAGCTCGTTAGATCCTC	3	(TC)16	54.6	146–178 [139]
MA39_182	F: CTACACGGGTGAAGCCTACC	R: GGCCGTAATCAGAGTCCACC	3	(TC)12	64	144–148 [129]
MA39_185	F: CGGGTGTGTAGATGACGCT	R: AAGACACGGAATGGGACGAG	2	(AG)15	59.8	231–358 [209]
MA39_236	F: GTAGCCATGTGGGTCTGTCC	R: AGTTGGTAGGGCACATGTCC	2	(CT)13	58.6	148–158 [126]
MA39_342	F: TCCCTTCAGTCTTCACACGC	R: AAAGGAGCGTTGAGTGGTGG	3	(TC)14	62	164–208 [146]
MA39_348	F: GTAGAGCTCCCATGCCTCAC	R: GGGCTGTCTACTGGATGGAC	2	(TC)17	56.2	150–180 [119]
MA40_282	F: TCTCTTCCCTCCGTCCTCC	R: TCTCCGGCTTCATGTCGTC	4	(GA)15	54.6	128–166 [116]
MA41_076	F: AACAAACGCTGGGTGATGGAA	R: TGGAGTTGACGCCTCTAGGA	3	(GA)26	59.2	169–209 [176]
MA41_373	F: GCGCCAATCAGAACACAAC	R: GGGAAAGAGCTTCTTCGCCA	3	(CT)16	54	166–207 [165]
MA42_028	F: GGATCGTCTCCGCCATTCT	R: TTCCGTACGATGCTCCCATG	3	(CT)33	56.1	129–147 [151]
MA42_231	F: GGGTGCGAAATGTGCATCAA	R: GGGCCAGTGAGCATTAGAGC	2	(AG)14	59.2	152–194 [131]
MA42_471	F: TGATGAAGAGCCCAGATCGTC	R: TGGCCTTGTCTCCATACGT	3	(GA)16	57.2	163–230 [153]

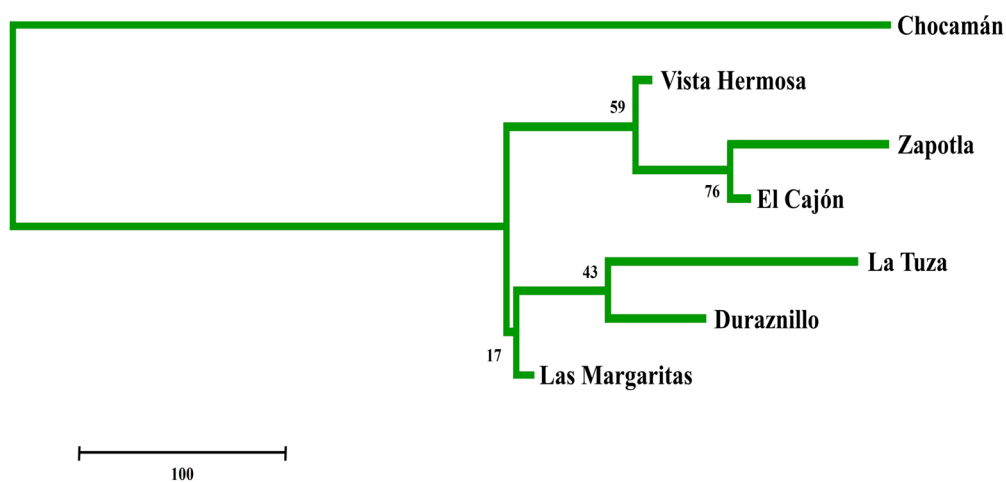


Figure 2

Neighbor-joining grouping based on Nei distances for the seven localities of *Magnolia mexicana*. The number at each node indicates the bootstrap percent.

The amplification mix (2 µl) contained 50 ng/µl genomic DNA; 1.6 µl of MgCl₂ (25 mM), 0.7 µl of oligonucleotide (primer forward), 0.7 µl of oligonucleotide (primer reverse), (1 µM), 0.3 µl of dNTPs (10 mM), 2.4 µl of buffer (5×) and 0.5 µl of dH₂O and 0.3 µl (1.5U) of TaqDNA polymerase (Promega®). Amplification was carried out in a Thermo Scientific® Artik endpoint thermal cycler; the PCR program consisted of pre-denaturation at 95 °C for three minutes, followed by 30 cycles, with denaturation at 95 °C for 30 seconds, followed by the alignment stage for 30 seconds, 72 °C for 90 seconds (extension stage), and finally a temporal extension at 72 °C for six minutes. The alignment temperature was specific for each micro-satellite: 54.6 °C MA39_159 and MA40_282; 64°C MA39_182; 59.8°C MA39_185; 58.6 °C MA39_236; 62°C MA39_342; 56.2°C MA39_348; 59.2°C MA41_076; 54°C MA41_373; 56.1°C MA42_028; 59.2 °C MA42_231; and 57.2 °C MA42_471.

The PCR products were visualized in 15 % acrylamide gels by means of electrophoresis at 90 volts for 90 minutes, and stained with ethidium bromide at a concentration of 10 mg/ml, then visualized with ultraviolet light. Once the electropherograms were obtained, the size of each of the fragments was identified with the Gel Analyzer 2010a program (Lazar and Lazar, 2021). Subsequently, the genotypes of each individual were identified, to build a matrix for estimated the data quality.

Data quality

The data obtained were analyzed in the Micro-Checker version 2.2.3 program (Van Oosterhout et al., 2004) in order to detect null alleles using 10,000 randomizations and a confidence interval of 95 %. For estimated the deviations from the Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) with Arlequin version 3.5.2.2 (Excoffier and Lischer, 2010). The test of HWE for each SSR and each locality, using 100,000 steps in a Markov chain and 100,000 dememorization steps. The test linkage disequilibrium (LD) between all pairs of loci in each locality was estimated with 10,000 permutations and two initial conditions in an expectation-maximization (EM) algorithm with a significance level of 0.05.

Statistical analysis

The genetic diversity values estimated for the seven localities of *M. mexicana* with the program GenAlEx (Genetic Analysis in Excel) version 6.3 (Peakall and Smouse, 2006) were: number of alleles per locus (Na), effective number of alleles per locus (Ne), observed (Ho) and expected (He) heterozygosity, Shannon index (I), Hardy-Weinberg equilibrium deviation (chi-square), and fixation index (Fis; Wright, 1922). In addition, the Nei genetic distances (Nei et al., 1983) were estimated with the GenAlEx program, and with the distance matrix obtained, a dendrogram was generated by means of the neighbor-joining algorithm in the POPTREE2 program (Takezaki et al. 2009). The node significances of tree were evaluated using bootstrap probabilities based on 1000 replicates.

The genetic structure of the *M. mexicana* populations was estimated by means of molecular analysis of variance (AMOVA) with 10,000 permutations to obtain genetic fixation indices and their significance, that is, global index of genetic

differentiation between populations (Fst), index of inbreeding within populations (Fis) and index of total inbreeding (Fit), to quantify the distribution of variability within and between localities (Excoffier and Lischer, 2010). In addition, an assignation analysis was carried out with STRUCTURE version 2.3.4 (Pritchard et al., 2000), with a systematic Bayesian clustering approach, applying the Markov chain Monte Carlo (MCMC) algorithm with correlated allele frequencies and under a mixture model. For each cluster, we assumed *K* values ranging from 2 to 20, with 10 independent MCMC simulation runs were performed, each consisting of 150,000 replicates of MCMC chains were carried out, after an initial period of 150,000 permutations. Then, 100 new repetitions of the MCMC method were run for the best *K* value. The optimal value of *K* was estimated in two ways, the method of Pritchard et al. (2000), which favors the model with the maximum value of $\ln \Pr(X|K)$ and the ΔK method (Evanno et al., 2005), which selects the model with the maximum second-order rate of change of $\ln \Pr(X|K)$. The results were visualized in the STRUCTURE SELECTOR web server (Li and Liu, 2018).

To contrast the results of the AMOVA, the cluster analysis with the neighbor-joining technique (genetic distance), and the Bayesian analysis (probability of belonging to a group by re-sampling), a principal component and discriminant analysis (DAPC) was carried out, to identify and describe groups of genetically related individuals. With principal component analysis (PCA) it is possible to reduce the number of original variables to a smaller set of independent variables (called principal components, PC), and identify the characters that contribute to explaining the highest percentage of genetic variation (Gutiérrez-Lozano et al., 2021); with discriminant function analysis (DFA) it was possible to examine the patterns of variation and identify the genetic variables that maximized the differences between populations (McCune and Grace, 2002). The DAPC is a standard technique free of assumptions about HW equilibrium or linkage disequilibrium and is based on the same classical statistical model as AMOVA, to estimate the differentiation between groups by means of a multivariate distance (Jombart et al., 2010). The analysis was carried out with the STATISTICA program, version 7 (Hilbe, 2007).

The decrease or expansion in the population sizes was estimated with a bottleneck analysis (Cournuet and Luikart, 1996), considering both the stepwise mutation model (SMM) and the infinite island model (IAM), using the standardized distribution test and Wilcoxon nonparametric test for heterozygosity by population. A paired Mantel test was carried out between the Nei genetic distances and the geographic distances of the *M. mexicana* populations to estimate the association between genetic structure *versus* isolation by distance. The collection of geographic data was obtained from distances in a straight line in km, from one location to another with the help of Google Earth Application.

Results

Data quality

No evidence for scoring errors due to stuttering or allele drop-out in any of the loci was detected by Microchecker software. However, null alleles were present in one of 12 loci amplified, but only in the La Tuza locality with frequencies between 5 % and 8 % should have only minor effects on classical estimates of population differentiation (van Oosterhout et al., 2004), they were subsequently maintained in all analysis. The test for HWE (12 locus-population combinations) and LD pattern (all pairs of loci in each locality), did not suggest a systematic bias across locus-population combinations and pairs of loci ($p = 0.05$).

Genetic diversity

The *Magnolia mexicana* populations showed variation in the genetic diversity indices (in terms of N_a , N_e , H_e , H_o , I , and F_{is}), as indicated in Table 4. The number of alleles per locus (N_a) per population ranged between 2.97 (La Tuza) and 5.75 (Zapotla), with an average of 4.33. The average number of effective alleles (N_e) in all populations was 3.44, with a range of variation from 2.48 (La Tuza) to 4.44 (Zapotla). Expected heterozygosity (H_e) and observed heterozygosity (H_o) ranged from 0.46 (La Tuza) to 0.70 (Zapotla), and from 0.25 (La Tuza) to 0.41 (Vista Hermosa), with averages of 0.63 and 0.33, respectively. The average value of the Shannon index (I) was 1.21, with a range of 0.82 (La Tuza) to 1.14 (Zapotla). The average fixation index (F_{is}) was 0.52, ranging from 0.41 (Vista Hermosa) to 0.59 (Chocamán). None of the populations meet the assumption required for Hardy-Weinberg equilibrium (HWE).

Genetic structure

The AMOVA analysis showed that 49 % of the total molecular variation was explained by differences between individuals, while 36 % variance was found within individuals, and only 15% variance among populations. All the localities analyzed presented statistically significant differences between each other ($p < 0.001$; Table 3).

In the neighbor-joining cluster tree, the seven localities can be divided into two main groups. Group one is the Chocamán locality, and group two includes the rest of the localities, but Las Margaritas locality was the most divergent, so it can be distinguished as subgroup 1 within group 2. The other populations in group 2 formed subgroup 2 within the group, so they are more similar to each other. The Bayesian clustering analysis shows that the optimal number of genetic groups was six, according to the maximum value of ΔK . All the groups includes a mixture of the six genetic groups, however, in three localities a higher proportion of a genetic group is visualized (La Tuza, Vista Hermosa y Chocamán; Figure 3).

The DAPC results indicated that the first two factors explained 90.83 % of the total variation in the data (Figure 4). The values obtained with the F test show that there are statistically significant differences between the majority of the analyzed populations. In the ordination diagram the points that represent the genetic characteristics of the individuals of the Chocamán population (F , 31.41; $p < 0.0001$) can be seen to be

separated from the group of points that include to the rest of the localities which make up a more heterogeneous group. The DAPC results, which represent the natural distribution of genetic variability of *M. mexicana* localities, were consistent with the results of the Bayesian analysis.

The infinite island model (IAM) estimated that there are recent bottlenecks in all localities; in the stepwise mutation model (SMM), only the localities of El Cajón, Duraznillo and Zapotla did not present a reduction in population size (Table 4). Additionally, the Mantel correlation test had a low correlation coefficient ($r = 0.2085$), considered statistically non-significant ($p = 0.81$), indicating no effect of distance on genetic variation.

Discussion

Different approaches to the study of the *Magnolia mexicana* genetic diversity and structure can be made; either by measuring neutral molecular variants, such as RAPD (Medrano et al., 2017), ISSR (Arteaga-Ríos et al., 2020), or codominant and highly polymorphic markers such as SSRs (Aldaba-Núñez et al., 2021). The importance of estimating variability through SSR markers lies in the possibility of obtaining basic information on genetic diversity and structure (distribution of genetic variation in populations), useful for the implementation of conservation programs for plant species (Aldaba-Núñez et al., 2021; Chavéz-Cortázar et al., 2021; Claerhout et al., 2023; Hernández et al., 2020; Tamaki et al., 2016; von Kohn et al., 2018; Veltjen et al., 2018).

The estimated genetic variability values in *M. mexicana* was similar ($H_e = 0.46$ -0.70, arithmetic mean: 0.63), to those found in other investigations in populations of the same species of the section *Talauma* in Mexico ($H_e = 0.60$, Aldaba-Núñez et al., 2021), in other species of neotropical magnolias ($H_e = 0.458$ -0.712; Veltjen et al., 2018), in two species of Chinese magnolias ($H_e = 0.521$ -0.718, Deng et al., 2020; Wang et al., 2019), and in *Magnolia kobus* from Japan ($H_e = 0.723$, Tamaki et al., 2016). In *M. acuminata* from the USA and Canada, the H_e values were higher, between 0.828 and 0.899 (Budd et al., 2015). The moderate differences in the values of genetic diversity estimated in these studies may be related to various factors, such as life history, degree of habitat fragmentation, anthropogenic pressures, or geographical characteristics of the distribution area, among others. With habitat fragmentation, the sizes of the populations are reduced and the distance between them is increased, which translates into lower gene flow (increase in inbreeding, decrease in heterozygosity), and increased genetic divergence between populations (Tamaki et al., 2016).

Table 3

Analysis of molecular variance (AMOVA) for seven localities of *Magnolia mexicana*.

Source of variance	Degrees of freedom	Sum of squares	Variance components	Total variance (%)	P value
Between populations (F_{st})	6	122.972	0.791	15%	<0.001
Between individuals (F_{is})	53	379.061	2.609	49%	<0.001
Within individuals (F_{it})	60	116	1.933	36%	<0.001

F_{st} = Fixation index, F_{is} = Inbreeding index, F_{it} = Total inbreeding index

Table 4

Genetic diversity values in evaluated *Magnolia mexicana* populations. N: number of individuals, Na: average alleles per locus, Ne: effective number of alleles per locus, I: Shannon index, Ho: observed heterozygosity, He: expected heterozygosity, Fis: Fixation index.

Population	N	Na	Ne	I	Ho	He	Fis	HWE
El Cajón	9	5.00 ± 0.57	3.78 ± 0.38	1.38 ± 0.11	0.31 ± 0.10	0.70 ± 0.03	0.58 ± 0.14	$\chi^2 = 27.54$, $p = 0.001$
La Tuza	5	2.91 ± 0.49	2.48 ± 0.41	0.82 ± 0.17	0.25 ± 0.10	0.46 ± 0.08	0.48 ± 0.15	$\chi^2 = 10.26$, $p = 0.025$
Las Margaritas	7	4.41 ± 0.60	3.64 ± 0.51	1.27 ± 0.13	0.31 ± 0.10	0.66 ± 0.04	0.59 ± 0.13	$\chi^2 = 17.72$, $p = 0.012$
Duraznillo	10	4.83 ± 0.53	3.57 ± 0.38	1.32 ± 0.11	0.36 ± 0.07	0.68 ± 0.03	0.48 ± 0.10	$\chi^2 = 21.47$, $p = 0.006$
Vista Hermosa	7	3.58 ± 0.33	2.88 ± 0.21	1.10 ± 0.08	0.41 ± 0.12	0.62 ± 0.03	0.41 ± 0.17	$\chi^2 = 10.22$, $p = 0.026$
Chocamán	8	3.83 ± 0.53	3.26 ± 0.44	1.14 ± 0.15	0.28 ± 0.10	0.61 ± 0.06	0.59 ± 0.14	$\chi^2 = 13.88$, $p = 0.026$
Zapotla	14	5.75 ± 0.78	4.44 ± 0.57	1.47 ± 0.16	0.40 ± 0.11	0.70 ± 0.05	0.50 ± 0.13	$\chi^2 = 35.91$, $p = 0.001$

Table 5

Probability obtained by bottleneck analysis. IAM Infinite island model, SMM Stepwise mutation model. The Wilcoxon test shows only the right tail, since the left tail was not statistically significant in four population.

Population	Z-test		Wilcoxon	
	IAM	SMM	IAM	SMM
El Cajón	0.00024	0.03198	0.00048	0.06396*
La Tuza	0.00488	0.02441	0.00976	0.04882
Las Margaritas	0.00061	0.01709	0.00122	0.03418
Duraznillo	0.00036	0.06469*	0.00073	0.12393*
Vista Hermosa	0.00001	0.00061	0.00024	0.00122
Chocamán	0.00048	0.00170	0.00097	0.00341
Zapotla	0.00036	0.05493*	0.00036	0.10986*

* Non-significant result.

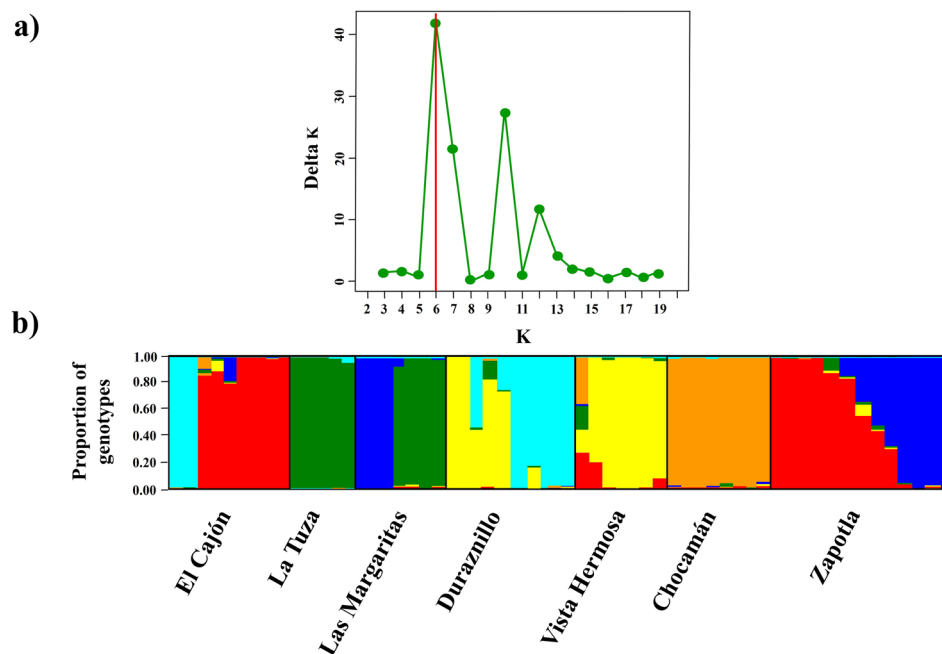


Figure 3
a) Estimated number of genetic groups (K) derived from clustering analysis from STRUCTURE. Delta K was calculated using the method described by Evanno et al. (2005); b) Genetic clusters obtained with K=6.

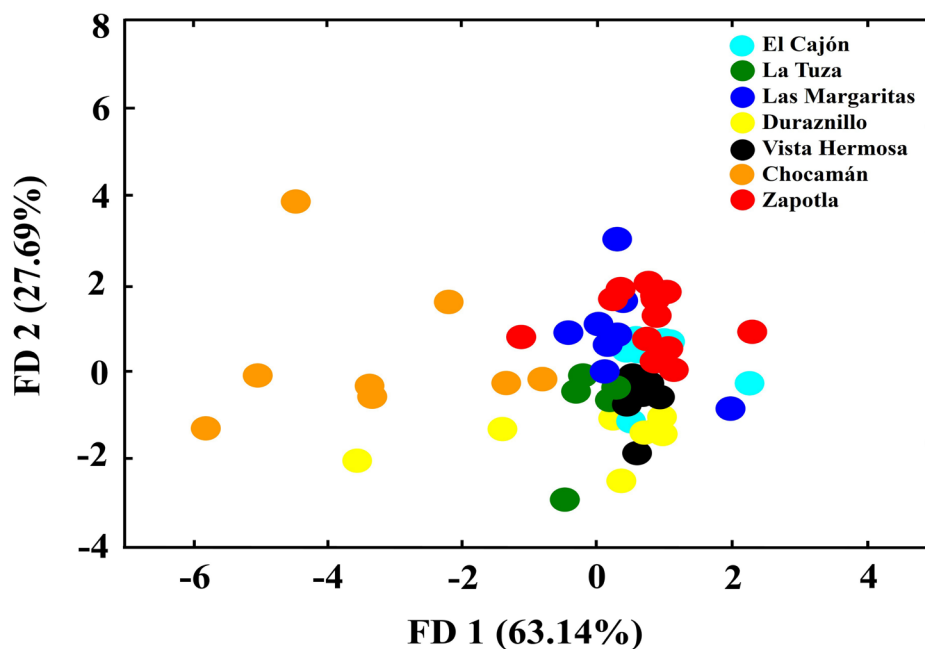


Figure 4
Discriminant analysis of principal components (DAPC) among *Magnolia mexicana* localities.

The values of the genetic diversity index in *M. mexicana* were relatively high in all localities (0.82-1.47), compared to those estimated in other populations of the same species ($I = 0.41-0.47$), and in different species of the genus, considered in extinction risk, such as *M. officinalis* ($I = 0.286$; Yu et al., 2011); *M. ashei* ($I = 0.172-0.617$; von Koh et al., 2018), and *M. pacifica* ($I = 0.39$; Muñoz-Castro et al., 2020). However, the differences in the values of genetic diversity could be due to different causes, related to aspects inherent to the evolutionary history and current conditions of the populations of each species (as mentioned in the previous paragraph), and/or to the methods used in each study (molecular markers, sample size, among others). The Shannon index values, for example, are influenced by the sample size and the presence of loci with a high number of alleles (Konopiński, 2020). In accordance with the above, the localities with the largest number of individuals (Duraznillo, El Cajón and Zapotla) presented the highest values of genetic diversity. Although these results should be interpreted with caution, because the gene pool could be overestimated due to the number of individuals, the H_e and I values obtained suggest that the current decline in *M. mexicana* populations could be more due to causes derived from the human disturbance, than to inbreeding or genetic drift (Deng et al., 2020).

The deviation from Hardy-Weinberg equilibrium (HWE) in most of the analyzed loci and the intermediate values of inbreeding and genetic diversity in the analyzed population, could be related to the high degree of fragmentation and disturbance of the tropical forest, habitat of the *M. mexicana* populations. Aldaba-Núñez et al. (2021) obtained similar results in some populations of the same species in the state of Veracruz, Mexico. It is probable that gene flow and diversity in the *M. mexicana* populations will continue to decrease considerably because there are no management and conservation programs for the species, nor for the ecosystem where it grows. In addition, the phenomenon of genetic drift can increase differentiation and/or cause the extinction of populations because the alleles that are lost or fixed in each of them will be different (Groom et al., 2005; Menges, 2000).

The genetic differentiation coefficient between *M. mexicana* populations was intermediate or moderate value ($F_{st} = 0.245$), according to the criteria proposed by Wright (1965). However, if genetic variability continues to decline and inbreeding increases, the species' responsiveness to changes in environmental conditions will decrease, putting its viability at risk (Frankham et al., 2015).

The AMOVA results indicate that most of the variation (>49 %) was explained by differences among individuals, and secondarily within individuals (36 %). This pattern of genetic differentiation is consistent with that observed in other Magnoliaceae species (Chávez-Cortázar et al., 2021; García-Montes et al., 2022; Medrano-Hernández et al., 2016; Newton et al., 2008).

In the present study, the clustering analysis divided the populations into two main groups, and the Bayesian analysis categorized them into three groups with different combinations of genetic information. The results of the DAPC were similar to those obtained the neighbor-joining technique, and

compatible with those obtained with the in the Bayesian analysis, so it is advisable to use a variety of different analysis techniques to obtain complementary information (García-Montes et al., 2022). With the information analysis techniques used, two independent populations were recognized, the first from the Chocamán locality, made up of only eight individuals close to each other (probably related), who grow in similar environmental conditions in orchards near houses. The Zapotla locality (Zongolica, Veracruz) grows in a slightly disturbed forest, so the density of adult individuals is higher than in the other populations. However, the extraction of fruits for sale is excessive and drastically affects the rate of natural regeneration; consequently, it is likely that in the near future the Zapotla population will disappear locally (Frankman et al., 2015; Gutiérrez-Lozano et al., 2021).

The infinite allele model results indicate that some of *M. mexicana* populations present recent bottlenecks, probably as a consequence of the low numbers of individuals, low genetic diversity and high inbreeding (Frankman et al., 2015). The exception were the El Cajón, Duraznillo y Zapotla localities, since the distribution of allelic frequencies indicates that it has not suffered recent bottlenecks. The step mutation model, which is considered more appropriate for the analysis of results with microsatellite markers (Putman and Carbone, 2014), also detect evidence of bottlenecks in three localities. In any case, it is probable that the number of individuals in the El Cajón and Zapotla localities will decrease in the short term due to the strong pressure exerted by the inhabitants through the extraction of flowers for medicinal purposes and food (Gutiérrez-Lozano et al., 2021). In the particular case of the locality Duraznillo due to the change in land use resulting from agriculture and livestock.

The Mantel test results showed that genetic divergence between *M. mexicana* populations is not related to geographic distance, which has also been observed in other Magnoliaceae species, such as the *M. pacifica* complex in Mexico (Muñoz-Castro et al., 2020) and in *Houpoëa officinalis* in China (Li et al., 2018). It is possible that other factors with local influence, such as crop and livestock agriculture, and trade in the flowers and fruit of *M. mexicana* in El Cajón, Las Margaritas, La Tuza, and Duraznillo have a greater influence on population genetic variation than isolation by distance.

The low values of genetic diversity in some of the populations (Chocamán, El Cajón y La Tuza) and the high genetic differentiation of the Chocamán and Zapotla populations could be a reflection of recent bottlenecks caused by the low density of individuals and by the mass extraction of flowers and fruits for edible and medicinal use (Song and Liu, 2019; Deng et al., 2020), but the influence of population isolation due to habitat fragmentation and disturbance should also not be underestimated. To maintain the regeneration capacity of native populations and reduce the extinction risk of *M. mexicana*, it is important to develop effective in situ conservation strategies and programs in the short term to reduce the impact of human activities on their habitat (Aldaba-Núñez et al., 2021; Chávez-Cortázar et al., 2021; Claerhout et al., 2023; Hernández et al., 2020; Tamaki et al., 2016; von Kohn et al., 2018; Veltjen et al.,

2018). With these specific actions, the probability of seedling survival, recruitment of new individuals, and the sizes of *M. mexicana* populations could be increased through natural regeneration as has happened in other magnolia species that were at risk of extinction (Song and Liu, 2019). Estimates of the distribution of the genetic variation of the *M. mexicana* populations through SSR markers provided basic and specific information which can be used in the implementation of a management and conservation program for the species.

Acknowledgments

This work was part of the CONACyT research project, CB-2016/284484 "Effect of climate change on relict populations of trees; integrating dendrochronology and population genetics." The authors thank Mr. Arnulfo, Francisco, Guillermo (biologist), Lorenzo (Puebla), Diego and Joaquin (Veracruz), for their kindness and support during the field work, and PhD candidate José Luis Reyes Ortiz for preparing the study area map and help in the field.

References

- Agarwal M, Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports* 27:617-31. <https://doi.org/10.1007/s00299-008-0507-z>
- Aldaba-Núñez FA, Veltjen E, Samain MS, Martínez-Salas E (2021) Disentangling species delineation and guiding conservation of endangered Magnolias in Veracruz, México. *Plants* 10:673. <https://doi.org/10.3390/plants10040673>
- Arteaga-Ríos LD, Mejía-Carranza J, Piña-Escutia JL, González-Díaz JG, Rivera-Colín A (2020) Comparación molecular y morfológica entre ejemplares de *Magnolia mexicana* "Yoloxóchitl" (Magnoliaceae) del estado de México y Veracruz. *Polibotánica* 49:107-124.
- Budd C, Zimmer E, Freeland JR (2015) Conservation genetics of *Magnolia acuminata*, an endangered species in Canada: can genetic diversity be maintained in fragmented, peripheral populations? *Conservation Genetics* 16(6):1359-1373. <https://doi.org/10.1007/s10592-015-0746-9>
- Chávez-Cortázar A, Oyama K, Ochoa-Zavala M, Mata-Rosas M, Veltjen E, Samain ME, Quesada M (2021) Conservation genetics of relict tropical species of *Magnolia* (section Macrophylla). *Conservation Genetics* 22:259-273. <https://doi.org/10.1007/s10592-021-01334-5>
- Claerhout T, Strubbe D, Castillo-Torres R, Samain MS, Chatrou LW, Veltjen E (2023) An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola. *Biodiversity and Conservation* 32:1205-1231. <https://doi.org/10.1007/s10531-023-02546-6>
- Cournuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001-2014. <https://doi.org/10.1093/genetics/144.4.2001>
- Deng Y, Liu T, Xie Y, Wei Y, Xie Z, Shi Y, Deng X (2020) High genetic diversity and low differentiation in *Michelia shiluensis*, an endangered *Magnolia* species in South China. *Forests* 11(4):469. <https://doi.org/10.3390/f11040469>
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh tissue. *Phytochemical Bulletin* 19(1):11-15.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14(8):2611-2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10(3):564-567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Figlar RB (2006). Taxonomy topics—a new classification for *Magnolia*. In: Royal Horticultural Society ed. *Rhododendrons with Camellias and Magnolias*. Cornwall: MPGBooks Ltd. 69-82.
- Frankham R, Ballou JD, Briscoe DA (2015) Introduction to conservation genetics, 2nd edn. Cambridge.
- García-Montes MA, Cibrián-Jaramillo A, Reyes-Zepeda F, Chacón-Hernández JC, Vanoye-Eligio V, Sánchez-González A, Octavio-Aguilar P (2022) *Magnolia tamaulipana*: genetic evaluation shows high vulnerability in a narrow distribution. *International Journal of Plant Sciences* 183(2). <https://doi.org/10.1086/717817>
- Groom MJ, Meffe GK, Carroll CR (2005) Principles of Conservation Biology. Sunderland Massachusetts, Sinauer Associated.
- Gutiérrez-Lozano M, Reyes JL, Octavio-Aguilar P, Galván-Hernández DM, Sánchez-González A, Vázquez-García JA (2021) Variación en la morfología foliar, floral y frutal de *Magnolia mexicana* (DC.) G. Don (Sección Talauma, Magnoliaceae) en México. *Botanical Sciences* 99(4):955-975. <https://doi.org/10.17129/botsci.2800>
- Hernández M, Palmarola A, Veltjen E, Asselman P, Teste E, Larridon I, Samain MS, González-Torres R (2020) Population structure and genetic diversity of *Magnolia cubensis* subsp. *acunae* (Magnoliaceae): effects of habitat fragmentation and conservation. *Oryx* 54(4):451-459. <https://doi.org/10.1017/S003060531900053X>
- Hilbe JM (2007) STATISTICA 7: An Overview. *Am Stat* 61:91-94. <https://doi.org/10.1198/000313007X172998>
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* 11(1):1-15. <https://doi.org/10.1186/1471-2156-11-94>
- Konopiński MK (2020) Shannon diversity index: a call to replace the original Shannon's formula with unbiased estimator in the population genetics studies. *PeerJ* 8:e9391. <https://doi.org/10.7717/peerj.9391>
- Lazar I, Lazar E (2021) GelAnalyzer 19.1. Available at <<https://www.gelanalyzer.com>> [cited 11.03.2022].
- Li H, Yang X, Yang Z (2018) Genetic diversity, population genetic structure and protection strategies for *Houpoëa officinalis* (Magnoliaceae), an Endangered Chinese Medical Plant. *Journal of Plant Biology* 61:159-168. <https://doi.org/10.1007/s12374-017-0373-8>
- Li Y, Liu JX (2018) Structure selector. A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources* 18:176-177. <https://doi.org/10.1111/1755-0998.12719>
- Linsky J, Crowley D, Beckman Bruns E, Coffey EED (2022) Global Conservation Gap Analysis of *Magnolia*. Atlanta Botanical Garden. Available at <<https://www.globalconservationconsortia.org/resources/global-conservation-gap-analysis-of-magnolia/>> [cited 11.03.2022].
- McCune B, Grace JB (2002) Analysis of ecological communities. MjM Software Design: Gleneden Beach, Oregon.
- Medrano-Hernández J, Rodríguez JJ, Reyes-Trejo B, Peña-Ortega G (2017) Molecular characterization using ISSR primers of *Magnolia mexicana* DC. from two regions in Zongolica, Veracruz, Mexico. *Revista Chapingo Serie Ciencias Forestales y del Ambiente* 23(3):427-436. <https://doi.org/10.5154/rchscfa.2017.03.019>
- Menges ES (2000) Applications of population viability analyses in plant conservation. *Ecological Bulletins* 48:73-84. <https://www.jstor.org/stable/20113249>
- Muñiz-Castro MA, Vázquez-García JA, Castro-Felix P, Carranza-Aranda AS, Santerre A (2020) Population genetics, species boundaries, and conservation in the *Magnolia pacifica* species complex along a continentality and moisture gradient in Western Mexico. *Botanical Sciences* 98(3):500-516. <https://doi.org/10.17129/botsci.2551>
- Nei M, Tajima F, Tatenno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* 19:153-170. <https://doi.org/10.1007/BF02300753>
- Newton AC, Gow J, Robertson A, Williams-Linera G, Ramírez-Marcial N, González-Espinosa M, Allnut TR, Ennos R (2008) Genetic variation in two rare endemic Mexican trees, *Magnolia sharpie* and *Magnolia schiedeana*. *Silvae Genetica* 57:348-356. <https://doi.org/10.1515/sg-2008-0051>

- Peackall R, Smouse PE (2006) GenAlex 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6(1):288-295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Pennington T, Sarukhán J (1998) *Arboles tropicales de México. Manual para la identificación de las principales especies*. Universidad Nacional Autónoma de México y Fondo de Cultura Económica, México, 521 p, ISBN 9681678559.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:94-959. <https://doi.org/10.1093/genetics/155.2.945>
- Putman AJ, Carbone I (2014) Challenges in analysis and interpretation of micro-satellite data for population genetic studies. *Ecology and Evolution* 4:4399-4428. <https://doi.org/10.1002/ece3.1305>
- Quintana-Delgado J, Hernández-Rodríguez M, Teste E, Palmarola A, Veltjen E, Valdes M, Asselman P, Larridon I, Samain ME, González-Torres LR (2022) Estructura y diversidad genética de *Magnolia cubensis* subsp. *cubensis* (Magnoliaceae). *Revista Jardín Botánico Nacional Universidad Habana* 43:33-46.
- Rivers MC, Martínez-Salas S, Samain MS (2016). *Magnolia mexicana*. The IUCN Red List of Threatened Species 2016: e.T193977A2293105. Available at <<https://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T193977A2293105.en.>> [cited 11.03.2022].
- Ruiz-Jiménez CA, Téllez-Valdés O, Luna-Vega I (2012) Clasificación de los bosques mesófilos de montaña de México; afinidades de la flora. *Revista Mexicana de Biodiversidad* 83:1110-1144. <https://doi.org/10.7550/rmb.29383>
- SEMARNAT (2010) Norma Oficial Mexicana. NOM-059-SEMARNAT-2010, Protección ambiental-Especies nativas de México de flora y fauna silvestres-Categorías de riesgo y especificaciones para su inclusión, exclusión o cambio-Lista de especies en riesgo. Diario Oficial de la Federación. 2da Sección, 30 de diciembre de 2010. Available at <<https://www.dof.gob.mx/normasOficiales/4254/semarnat/semarnat.htm>> [cited 11.03.2022].
- Song C, Liu H (2019) Habitat differentiation and conservation gap of *Magnolia biondii*, *M. denudata*, and *M. sprengeri* in China. *PeerJ* 6:e6126. <https://doi.org/10.7717/peerj.6126>
- Tamaki I, Setsuko S, Tomaru N (2016) Genetic diversity and structure of remnant *Magnolia stellata* populations affected by anthropogenic pressures and a conservation strategy for maintaining their current genetic diversity. *Conservation Genetics* 17:715-725. <https://doi.org/10.1007/s10592-016-0817-6>
- Takeshi N, Nei M, Tomaru K (2009) POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Molecular Biology Evolution* 27:747-752. <https://doi.org/10.1093/molbev/msp312>
- UICN (2012) The IUCN Red List of Threatened Species. Version 2021-3. Available at <<https://www.iucnredlist.org>> [cited 11.03.2022].
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4(3):535-538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
- Vázquez-García JA, Muñiz-Castro MÁ, Arroyo F, Pérez AJ, Serna M, Cuevas-Guzmán R, Gurrola-Díaz CM (2013) Novelty in neotropical *Magnolia* and an addendum proposal to the IUCN Red List of Magnoliaceae. In: Vázquez-García JA, Muñiz-Castro MÁ, Salcedo Pérez E, Hernández Álvarez E, Escoto García T, Díaz Echavarría N (eds) *Recursos Forestales en el Occidente de México: Diversidad, Manejo, Producción, Aprovechamiento y Conservación*. Guadalajara: Universidad de Guadalajara, pp 461-496, ISBN 9786078072736
- Veltjen E, Asselman P, Hernández M, Palmarola A, Testé E, González LR, Goetghebeur P, Larridon I, Samain MS (2018) Genetic patterns in Neotropical magnolias (Magnoliaceae) using de novo developed microsatellite markers. *Heredity* 122:485-500. <https://doi.org/10.1038/s41437-018-0151-5>
- Von Kohn C, Conrad K, Pooler M, Kramer M (2018) Genetic diversity of *Magnolia ashei* characterized by SSR markers. *Conservation Genetics* 19:923-936. <https://doi.org/10.1007/s10592-018-1065-8>
- Waizel-Bucay J (2002) Uso tradicional e investigación científica de *Talauma mexicana* (D.C.) Don., o flor del corazón. *Revista Mexicana de Cardiología* 13:31-38.
- Wang X, Chen W, Luo J, Yao Z, Yu Q, Wang Y, et al., (2019) Development of EST-SSR markers and their application in an analysis of the genetic diversity of the endangered species *Magnolia sinostellata*. *Molecular Genetics and Genomics* 294:135-147. <https://doi.org/10.1007/s00438-018-1493-7>
- Wang YB, Liu BB, Nie ZL, Chen HF, Chen FJ, Figlar RB, Wen J (2020) Major clade and a revised classification of *Magnolia* and Magnoliaceae based on whole plastid genome sequences via genome skimming. *Journal of Systematics and Evolution* 58(5):673-695. <https://doi.org/10.1111/jse.12588>
- Wright S (1965) The interpretation of population structure by F-Statistics with special regard to systems of mating. *Evolution* 19:395-420. <https://doi.org/10.1111/j.1558-5646.1965.tb01731.x>