

Effect of the Landscape on Functional and Spatial Connectivity in *Magnolia Cubensis* (Magnoliaceae) in Two Mountain Massifs of Cuba

Majela Hernández Rodríguez (✉ majelahr@gmail.com)

Facultad de Biología, Universidad de La Habana <https://orcid.org/0000-0002-1142-5296>

Ernesto Testé Lozano

Universidad de la Habana Jardín Botánico Nacional

Emily Veltjen

University of Ghent: Universiteit Gent

Jhonny Quintana Delgado

University of Havana: Universidad de la Habana

Alejandro Palmarola Bejerano

Universidad de la Habana Jardín Botánico Nacional

Marlyn Valdés de la Cruz

Universidad de la Habana Facultad de Biología

Pieter Asselman

Ghent University: Universiteit Gent

Isabel Larridon

Royal Botanic Gardens: Real Jardín Botánico

Marie-Stéphanie Samain

INECOL: Instituto de Ecología

Luis Roberto González Torres

The University of British Columbia

Research Article

Keywords: gene flow, genetic patterns, isolation by distance, least cost path, significant evolutionary unit, spatial autocorrelation

Posted Date: April 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-432614/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Landscape features impact gene flow and the spatial patterns of genetic variation between populations of a species. Because many Magnoliaceae species occur in fragmented and highly disturbed landscapes, the family is an excellent model for landscape genetic studies. This research focuses on the subspecies and localities of *Magnolia cubensis* and aims to: (1) compare the genetic diversity, (2) search for genetic patterns, (3) describe the functional connectivity and (4) assess the structural connectivity of the landscape. This study employs 21 microsatellite markers on two subspecies, complemented with landscape characteristics of the Guamuhaya and Sierra Maestra massifs in Cuba. *Magnolia cubensis* subsp. *acunae* does not have a well-defined spatial genetic pattern: there is no evidence of isolation by distance or spatial autocorrelation and the little genetic differentiation between the two localities does not reflect the characteristics of the landscape that separates them or the cost values to cross it. *Magnolia cubensis* subsp. *cubensis* presents evidence of isolation by distance and the autocorrelation analyses indicate that the approximate scale of the genetic neighborhood is between 35 and 40 km. There is a marked genetic structure with probability values that indicate the existence of three genetic groups. Large genetic differentiation was only found between Gran Piedra and the other localities, which reflects low gene flow. Our results support the recognition of these subspecies at the species level. We define one evolutionary significant unit in *Magnolia cubensis* subsp. *acunae* and two evolutionary significant units in *Magnolia cubensis* subsp. *cubensis*. These results must be combined with ecological, social and distribution data, in order to obtain a more accurate and realistic perspective of the conservation management strategies for these taxa.

Introduction

Landscape features impact gene flow and the spatial patterns of genetic variation between populations of the same species (Anderson 2010). With the emergence of landscape genetics as a discipline (Manel et al. 2003), the ability of genetic markers to quantify genetic diversity is being combined with spatial data (Scribner et al. 2005; Apodaca et al. 2012; Frantz et al. 2012). This combination is possible due to the existence of great variety of statistical methods which evaluate the role that spatial variables play in genetic diversity and population structure. From a conservationist perspective, landscape genetics enables identifying barriers that reduce gene flow or genetic diversity (Epps et al. 2005; Frantz et al. 2012), as well as predicting the effects of management measures on genetic variation and connectivity (Scribner et al. 2005; Apodaca et al. 2012).

Because many Magnoliaceae species occur in fragmented and highly disturbed landscapes, the family is an excellent model for such landscape genetic studies. In this plant family, 48% of the species are assessed as threatened (Rivers et al. 2016). Different molecular *Magnolia* case-studies have been conducted mainly focused on genetic diversity and population structure and their relationship with ecological, morphological, taxonomic and physiological aspects (e.g. Hirayama et al. 2007; Tamaki et al. 2018; Rico & Gutiérrez Becerril 2019; Chávez-Cortazar et al. 2020; Aldaba-Núñez et al. 2021). Most studies did not consider the landscape, and only a few studies have compared *Magnolia* genetic diversity with

the underlying spatial data (Kikuchi and Isagi 2002; Isagi et al. 2007; Hernández et al. 2020a). Although these studies are approximations to the description of genetic patterns in geographical space, they lack an integrative analysis between genetic variation patterns and the landscape or environment characteristics.

Within the Magnoliaceae family, *Magnolia cubensis* Urb. is a perfect case study to execute a full integrative landscape genetic analysis because we have been intensively surveying their populations for years and its full distribution is documented (Palmarola et al. 2018; Testé et al. 2019; Simón 2020). It is one of the six species endemic to Cuba (Palmarola et al. 2016; Greuter and Rankin 2017) that belong to the section *Talauma* subsect. *Cubenses* (Figlar & Nootboom 2004). Currently, the species is represented by two subspecies: *Magnoliacubensis* subsp. *cubensis* and *M. cubensis* subsp. *acunae* Imkhan. (Greuter and Rankin 2017). Although morphologically the two are very similar (Hernández, 2014), it has been suggested to move the subspecies to the species rank based on molecular and geographical arguments (Palmarola et al. in prep.; Veltjen et al. in review). *Magnolia cubensis* subsp. *acunae* is a local endemic from the Guamuhaya massif in central Cuba. *Magnolia cubensis* subsp. *cubensis* is restricted to the Sierra Maestra massif in the southeast of Cuba (González-Torres et al. 2016). Populations of both subspecies have been reduced mainly by overexploitation, illegal logging, loss of habitat and low natural regeneration, causing that both taxa have been assessed as Endangered (Cicuzza et al. 2007; González-Torres et al. 2016; Hernández et al. 2020b, c). Due to these problems, conservation actions have increased in recent years (Molina-Peregrín et al. 2014; Palmarola et al. 2015a, b; Testé 2015; Palmarola et al. 2018; Testé et al. 2019).

A landscape genetic study of the populations of the *Magnolia cubensis* subspecies is intriguing as, based on the *a priori* knowledge we have on their characteristics, the influence of the landscape on the genetic diversity cannot be predicted. On the one hand, the large size of their individuals (Palmarola et al. 2018), their longevity (Petit and Hampe 2006), the absence of clonal growth and seeds dispersed by birds (Vásquez et al. 2001) can mitigate the loss of genetic diversity and population structuring, even in the presence of a fragmented landscape or with elements that constitute barriers to gene flow (Honnay and Jacquemyn 2007). On the other hand, small population sizes, being forest specialists (González-Torres et al. 2016), potential self-incompatibility (Yasukawa et al. 1992) and pollination by insects (Thien 1974), can make these trees susceptible to the fragmentation of habitat and landscape patterns (Kolb and Diekmann 2005). Some authors such as Honnay and Jacquemyn (2007), Isagi et al. (2007) and Aparicio et al. (2012) found that in species with comparable characteristics, the genetic diversity and population structure are affected by the loss of habitat and isolation.

Using microsatellites markers, this integrative landscape genetic study aims to solve different research questions that can be structured in four main categories (1) **genetic diversity**: Do the subspecies and localities of *Magnolia cubensis* present differences in their genetic diversity? (2) **genetic structure**: Are the populations of *Magnolia cubensis* subsp. *acunae* and *M. cubensis* subsp. *cubensis* genetically structured? Are the genetic distances between localities explained by the geographical distances? Is the genotype of an individual in a locality dependent on the genotype of an individual in a nearby locality? (3)

functional connectivity: How much gene flow, expressed as genetic differentiation, is there between localities within a subspecies? How many evolutionary significant units can be defined for *Magnolia cubensis*? (4) **structural connectivity:** Do landscape characteristics of the Guamuhaya and Sierra Maestra massifs constitute barriers to gene flow? Do environmental variables contribute significantly to genetic patterns?

Material And Methods

Study site and sampling

This study was carried out in two localities in the Guamuhaya massif and five localities in the Sierra Maestra massif, Cuba (Fig. 1). These areas contain 90% of the known individuals of *M. cubensis* subsp. *acunae* and *M. cubensis* subsp. *cubensis* and present different IUCN management categories according to the National System of Protected Areas (SNAP) (2009). We collected leaf samples from 227 individuals and stored them in self-sealed bags with silica gel. The number of DNA samples per locality is shown in Table 1.

DNA extraction and genotypification

We extracted DNA from dried leaf tissue using a modified cetyltrimethylammonium bromide (CTAB) extraction protocol (Doyle and Doyle 1990), with MagAttract Suspension G solution mediated cleaning (Xin and Chen 2012). We genotyped individuals with 21 microsatellite or Simple Sequence Repeats (SSR) markers developed on four Neotropical *Magnolia* species (Veltjen et al. 2019). PCR conditions and primer labeling follow Veltjen et al. (2019). Fragment analyses were executed by Macrogen Inc. (Seoul, South Korea) and the results were analyzed in Geneious v.8.0.5 (<http://www.geneious.com>; Kearse et al. 2012), using the microsatellite plugin. Those individuals in whom more than 70% of markers could not be genotyped were excluded from the analyses. In the same way, the markers MA39_442 and MA42_296 for *M. cubensis* subsp. *acunae* were eliminated, because they did not amplify for more than 50% of the individuals. Therefore, the genetic characterization was carried out with 19 SSR markers for *M. cubensis* subsp. *acunae*, 21 SSR markers for *M. cubensis* subsp. *cubensis* and 19 overlapping SSR markers for the species.

Marker testing and genetic diversity

We calculated the deviation from Hardy-Weinberg proportions, linkage disequilibrium and the inbreeding coefficient (F_{IS}) for each *locus*, using 10,000 dememorization steps, 50 batches and 5,000 iterations per batch in Genepop v.4.3 (Rousset 2008). We calculated deviations of both the uncorrected (Waples 2015) and (sequential Bonferroni) corrected p-values to the nominal level of $\alpha = 0.05$ for both analyses.

Genetic diversity of each locality was quantified based on the following parameters in GeneAIEx v.6.5 (Peakall and Smouse 2012): number of alleles per *locus* (A), effective alleles (A_e), number of private alleles (A_p), observed heterozygosity (H_o), expected heterozygosity (H_e), inbreeding coefficient (F_{IS}) and

percentage of polymorphic *loci* (P). The allelic richness (A_R) and significance of F_{IS} was calculated with FSTAT v.2.9.3.2 (Goudet 1995).

Description of spatial genetic patterns

For each subspecies, the genetic distances between individuals and between localities were estimated from codominant genotypic distances (GD) and lineal genetic distances (LinGD), calculated according to Smouse and Peakall (1999), and the missing data were interpolated. From the coordinates of each individual and the elevation data, geographical distance matrices (GGD) and distance matrices in altitude (ALTD) were calculated, in kilometers, between individuals and between localities of the same subspecies. The matrices were obtained in GeneAEx v.6.5 by a modification of the Haversine formula (Sinnot 1984). Considering the spatial distribution of the individuals, the matrix of real geographic distances (GGD_R) was calculated as the hypotenuse of a triangle where the legs were the GGD and ALTD matrices. To calculate the distances between localities of the same subspecies, the coordinates of one individual (located in the center of each locality) were taken as representative of that locality. The subsequently mentioned structure analyses were carried out between subspecies and between localities within the subspecies. The genetic-spatial patterns (isolation by distance, spatial autocorrelation and functional connectivity) were analyzed for the two subspecies independently.

1) Genetic structure

We visualized genetic distances between subspecies, localities and individuals with a Principal Coordinates Analysis (PCoA) from the GD matrix in GeneAEx v.6.5. The groups defined *a priori* were the two subspecies, as well as the different localities sampled. The method used was the standardized covariance matrix and the scores corresponding to the principal coordinates 1 and 2 were plotted.

To estimate the probability of each individual belonging to a subspecies or locality, based on their genotype, we ran two data sets in Structure v.2.3 (Pritchard et al. 2000): grouped by subspecies and grouped by localities. The runs were performed under the following conditions: 10,000 Markov chain Monte Carlo replicates after an initial burn-in of 10,000, using correlated allelic frequencies and assuming the admixture model. To obtain probability values of the allocation of individuals to each genetic group K , we used five repetitions for each evaluated value of K , set to run from 1 to 8 for the species and *M. cubensis* subsp. *acunae*, and from 1 to 10 for *M. cubensis* subsp. *cubensis*. We determined the most probable number of groups from the value of ΔK obtained according to the method of Evanno et al. (2005). The results were analyzed in Structure Harvester Web v.0.6.94 (Earl and vonHoldt 2012). An individual was considered to be a member of a genetic group when its probability of belonging to that group was higher than 0.5.

2) Isolation by distance and spatial autocorrelation

To identify possible isolation by distance patterns between and within localities of the same subspecies, we performed Mantel correlation tests with 999 permutations between lineal genetic distances (LinGD)

and real geographical distances (GGD_R), whereby $\alpha < 0.01$ was considered statistically significant. To analyze whether the genotype of an individual is dependent on the genotype of another close individual, spatial autocorrelation analyses were performed according to Miller (2005). Taking into consideration the distance between the individuals, and that the individuals were not randomly distributed within a locality, we constructed distance classes that contain comparable pairs of individuals. We constructed 5 distance classes for *Magnolia cubensis* subsp. *acunae* and 10 for *Magnolia cubensis* subsp. *cubensis*. Bonferroni corrections were applied by the calculation of the V statistic (Miller 2005). The significance of V for each data set is evaluated via random allocation of individuals and genotypes among the set of sampled points on the landscape and calculated for 1,000 randomization replicates. The proportion of randomization replicates where random value of V was equal to the observed V provides an estimate of the p-value for each test. The results were visualized through correlograms between the values of the A_y statistic (average of the genetic distances between all pairs of individuals that are within a distance class) and the geographical distance classes. The spatial autocorrelation approaches are useful for defining the scale of genetic neighborhoods (Wright 1943) and the correlograms allowed us to elucidate when genetic and geographic distance are no longer correlated. The isolation by distance analyses were performed in GeneAEx v.6.5 and the spatial autocorrelation was realized in ALLELES IN SPACE (Miller 2005).

3) Functional connectivity

We estimated genetic differentiation between localities through pairwise comparisons of F_{ST} (Weir and Cockerham 1984), G_{ST} (Nei 1973; Nei and Chesser 1983) and D_{JOST} (Jost 2008) values calculated in R (R Core Team 2020 <https://www.R-project.org/>) using the fastDivPart function of the package diveRsity (Keenan et al. 2013). Interpretation of genetic differentiation followed the criteria of Hartl and Clark (1997).

Landscape data acquisition

The scenarios corresponding to the GuamuHaya and Sierra Maestra massifs were cut out from the vector map of the land surface of Cuba. Each of the massifs was analyzed separately for landscape data acquisition. For the description of the landscape, we used the vegetation map (Estrada et al. 2012) with a scale of 1: 100,000 referred to as the 'Vegetation' variable in further analyses. We used 6 other variables: the Digital Elevation Model (DEM), the index of the landform (Topography), the slope of the terrain (Slope), Temperature Seasonality (BIO4), Annual Precipitation (BIO12) and Precipitation Seasonality (BIO15), considering their contribution to the distribution models of the *Magnolia* subsect. *Cubenses* obtained by Simón (2020). The layers of the topographic and climatic variables were downloaded from WorldClim (<http://www.worldclim.org>). These variables are only available on a global scale at a resolution of 30 arc-seconds ($\sim 1 \text{ km}^2$), so it was necessary to trim all the variables to a standardized pixel size of 0.7883 km^2 . The processing was carried out in QGIS v.3.4.12 and Global Mapper v.15.0 (www.globalmapper.com).

Landscape genetics analyses

We used three analyses to correlate spatial genetic patterns and landscape characteristics: 1) Monmonier's Maximum Difference algorithm, 2) partial Mantel tests and 3) minimum cost models.

1) To identify barriers to gene flow we used Monmonier's Maximum Difference algorithm (Monmonier 2010). This procedure uses the Delaunay triangulation (Watson 1992; Brouns et al. 2003) to create, based on Nei's genetic distances (Nei et al. 1983), a network that connects all individuals and allows the detection of areas with significant changes in allele frequencies. The residuals of the genetic distances derived from the linear regression between all pairs of genetic and geographic distances were used to create the *M. cubensis* network. The number of barriers specified was 10. The analysis was performed in ALLELES IN SPACE and the visualization of the Delaunay triangulation on the maps of each mountain massif was performed in QGIS v.3.4.12.

2) To determine which environmental variables contribute significantly to genetic patterns, partial Mantel tests were used. We created environmental distances (BIO4D, BIO12D and BIO15D) from the pairwise differences between the individuals for each variable and combined environmental distances (EnvD) from the pairwise differences between the individuals taking into account all the environmental variables. Matrices of lineal genetic distances (LinGD), real geographical distances (GGD_R) and environmental distances (EnvD, BIO4D, BIO12D and BIO15D) were investigated for their correlation. For the creation of the distance matrices: EnvD, BIO4D, BIO12D and BIO15D, the values of the BIO4, BIO12 and BIO15 layers were extracted from the coordinates of each individual in QGIS v.3.4.12 using the Microsoft Excel add-in tool PopTools. Three groups of analyses were carried out: a) partial correlation between genetic (LinGD) and environmental distances (EnvD, BIO4D, BIO12D and BIO15D), in which the effect of geographical (GGD_R) distances was controlled, b) partial correlation between geographical (GGD_R) and environmental distances (EnvD, BIO4D, BIO12D and BIO15D), in which the effect of genetic distances (LinGD) was controlled, and c) partial correlation between genetic (LinGD) and geographic (GGD_R) distances, in which the effect of environmental distances (EnvD, BIO4D, BIO12D and BIO15D) was controlled. The tests were run in R using the vegan package (Oksanen et al. 2019), with 1,000 permutations and the Spearman's correlation coefficient.

3) To evaluate the structural connectivity of the landscape in each massif, minimum cost models (Least-cost-path or LCP) were used (Adriaensen et al. 2003). For the construction of the LCP models, cost or friction values were assigned to a set of environmental variables (Vegetation, BIO4, BIO12 and BIO15), as well as geographical ones (DEM, Slope and Topography). Friction maps were constructed in QGIS v.3.4.12 based on seven scenarios that explicitly consider the influence of landscape heterogeneity. For the first scenario the influence of the vegetation was tested. For this the vegetation map was reclassified to three classes, with low cost values for those plant formations where the species is distributed. For the rest of the scenarios the influence of the environmental and geographic variables was tested. For that, the values of each environmental variable (BIO4, BIO12 and BIO15) and each geographical variable (DEM, Topography, Slope) were extracted, for all the known individuals. The minimum and maximum values

and the 10th and 90th percentiles were calculated for each variable with the Microsoft Excel PopTools add-in tool. Each environmental and geographic layer was reclassified to three classes, where the highest cost (100) was assigned to cells that presented values less than the minimum, or greater than the maximum. The average cost (10) to the cells whose values were outside the 10th and 90th percentiles, but within the range, while the lowest cost values (1) were assigned to cells with values included in the interpercentile range (Online Resource 1).

Finally, a multivariate resistance surface was created by adding each reclassified variable (raster map with a cell size of 90 m²), representing 7 different gene flow scenarios. The least cost routes between localities were calculated under the seven hypothetical scenarios previously described. The analysis was carried out in QGIS v.3.4.12 through the Least-cost-path extension. To visualize the possible dispersal routes between localities, these results were represented on a raster map. Ultimately, the cost values obtained for each route were correlated with the genetic distances obtained from the gene flow estimators (F_{ST}), and the geographic distances between localities; using a Spearman correlation in Past v.3.14 (Hammer et al. 2001). This comparison was based on the logical assumption that the rate of gene flow between any two localities is inversely proportional to the cost of movement between them (i.e., if the cost is high, the gene flow must be low).

Results

Marker testing and genetic diversity

The Hardy-Weinberg equilibrium deviations for *Magnolia cubensis* show different results for the two subspecies. For *M. cubensis* subsp. *acunae*, no statistically significant differences were obtained for any of the markers. However, in *M. cubensis* subsp. *cubensis* 14 *loci* presented F_{IS} values significantly different from zero, in at least one locality. In Pico Turquino and Gran Piedra, more than 20% of the markers showed significant deviations, with 9 and 6 *loci*, respectively. Linkage analysis showed for *M. cubensis* subsp. *acunae* that 18 of 171 pairs of markers were found to be linked. In *M. cubensis* subsp. *cubensis*, significant differences were found for 7 pairs of the 210 combinations.

The **genetic diversity** measurements per taxon and locality of *Magnolia cubensis* can be found in Table 2. *Magnoliacubensis* subsp. *cubensis* showed higher values for the genetic diversity parameters compared to *M. cubensis* subsp. *acunae*. At the infraspecific level for *M. cubensis* subsp. *acunae*, Topes de Collantes showed the highest mean values for number of alleles, effective alleles, allelic richness and heterozygosity. Lomas de Banao presented the lowest values for all the diversity measures analyzed. Both localities showed a similar percentage of polymorphism, higher than 90%. For *M. cubensis* subsp. *cubensis*, La Bayamesa presented the highest values for the genetic diversity measures, while the minimum values for the different diversity measures were obtained by various locations: El Gigante A (5.10), Pico Caracas A_e (3.08) and A_R (4.317), and Pico Turquino H_o (0.37). In all the localities, a deficit of heterozygotes is observed, since the mean values of H_o were exceeded by H_e, with the largest difference in the individuals of Pico Turquino. For the inbreeding coefficient, the highest value was in Pico Turquino

and the lowest in La Bayamesa. Detailed information on the genetic diversity estimates per *locus* in *M. cubensis* subsp. *acunae* is shown in Online Resource 2 and for *M. cubensis* subsp. *cubensis* it is given in Quintana et al. (in review).

Genetic structure

The **principal coordinates analysis** of the three datasets (i.e. dataset 1: data of the two subspecies combined; and datasets 2 & 3: data of each subspecies separately) is presented in Fig. 2. When including the data of both subspecies, the PCoA showed well-defined groups corresponding to the two subspecies of *Magnolia cubensis* (Fig. 2a). At the infraspecific level for *M. cubensis* subsp. *acunae*, no clustering was observed for individuals from the Topes de Collantes and Lomas de Banao localities (Fig. 2b). In contrast, in *M. cubensis* subsp. *cubensis*, the PCoA contains five well-defined groups. The clusters containing the individuals of Pico Caracas, Pico Turquino, La Bayamesa and El Gigante present a slight overlap; while that of Gran Piedra does not overlap with the rest of the localities (Fig. 2c).

Barplots respecting the ΔK criterium to determine the optimal K and Structure Harvester output of the **structure analyses** of the three datasets are shown in Fig. 3 and Online Resource 3, respectively. The structure analyses including both subspecies result in $\Delta K = 2$, dividing the individuals in two genetic groups *cf.* the subspecies (Fig. 2a). There is a smaller, secondary peak at $K = 4$ whereby one genetic group corresponds to *Magnolia cubensis* subsp. *acunae* and the other three to *M. cubensis* subsp. *cubensis*. When analyzing the subspecies separately, the *M. cubensis* subsp. *acunae* dataset has a $\Delta K = 2$, yet, the mean $L(K)$ does not increase between 1 and 2; hence, the most likely true optimal K value for this dataset is 1. The barplots of $K = 2$ (Fig. 2b) show a mixed pattern that includes individuals from both locations in each genetic group. For *M. cubensis* subsp. *cubensis* $\Delta K = 3$, with a secondary peak at $\Delta K = 5$ (Online Resource 3). The genetic structure of *M. cubensis* subsp. *cubensis* for $K = 3$ showed the following composition: Pico Turquino shows individuals of two different genetic groups, while the localities Pico Caracas, La Bayamesa and El Gigante have individuals of only one genetic group. However, they are non-exclusive because 1) the group of Pico Caracas is present in Pico Turquino; and 2) La Bayamesa, El Gigante and the rest of the individuals of Pico Turquino belong to the same genetic group. Only Gran Piedra shows a genetic group exclusive of this locality.

Isolation by distance and spatial autocorrelation

No evidence of **isolation by distance** was found for *Magnolia cubensis* subsp. *acunae* as Mantel's correlation test between genetic and geographic distances showed low statistically non-significant correlation coefficients (Fig. 4a). For *M. cubensis* subsp. *cubensis*, the correlation coefficient was high ($r^2 = 0.946$) and statistically significant (Fig. 4b).

These results were supported by the **spatial autocorrelation analyses** (significant spatial autocorrelation is indicated in red in Fig. 4). For *Magnolia cubensis* subsp. *acunae* non-significant values were obtained in all distance classes. In contrast, significant and positive autocorrelation values were obtained in the first five distance classes for *M. cubensis* subsp. *cubensis* (up to 35 km). The transition from positive to

no correlation indicates the approximate scale of the genetic neighborhood. For this subspecies, the genetic distances increased with the geographic distance, as expected when there is isolation by distance.

Functional connectivity

The **genetic differentiation** between populations of two taxa of *Magnolia cubensis* is presented in Table 3. The pairwise D_{JOST} values range from 0.035 to 0.422, the pairwise F_{ST} values from 0.036 to 0.290 and the pairwise G_{ST} values from 0.021 to 0.170. For all the pairwise comparisons, the lowest values were found between localities of *M. cubensis* subsp. *acunae*. The highest values for allelic differentiation (D_{JOST}) corresponded to Gran Piedra - Turquino and for genetic fixation (F_{ST} and G_{ST}) to Gran Piedra – Pico Caracas.

Landscape genetics analyses

Monmonier's Maximum Difference algorithm showed that the highest rates of genetic change were located mainly between localities and not within them (Fig. 5). Of a maximum of 10 barriers (previously assigned to the algorithm), 3 were assigned for *Magnolia cubensis* subsp. *acunae* and 7 for *M. cubensis* subsp. *cubensis*. The geometry of the barriers detected, suggests that the main differentiation zones are found between the West and the East of each massif. The barriers never completely surround the localities, even for those more isolated localities. Even though the barriers to gene flow appear to extend along a latitudinal gradient for the two evaluated subspecies, in the case of *M. cubensis* subsp. *acunae* its limits do not cover the entire distance between localities. These results support the existence of barriers to gene flow that correspond in some cases with landscape features: the Agabama river valley in Guamuhaya and the Santiago de Cuba basin (also known as “Paso de la Barbacoa”) in the area between the Sierra Maestra and the Gran Piedra mountain range.

Mantel's partial correlations between geographic and genetic distances (controlling for the effect of environmental distance) support the Mantel test results. That is: a significant correlation was found only in *Magnolia cubensis* subsp. *cubensis*, in which the presence of isolation by distance was also detected. None of the correlations between genetic and environmental distances (controlling for the effect of geographic distance) were significant. However, statistically significant correlations were detected between geographic and environmental distances (controlling for the effect of genetic distance), for both subspecies (Table 4).

The **cost values** and the **least cost path** derived from the minimum cost models are showed in Table 5 and Fig. 6, respectively. Although in general the geographical distances, the cost distances and the fixation index (F_{ST}) between localities show positive and significant correlations, this relationship does not seem so evident when we compare between the subspecies. For example: for the Topes de Collantes – Lomas de Banao route (*Magnolia cubensis* subsp. *acunae*) the cost value is approximately 7,000; between the two locations there is a distance of 43 km and $F_{ST} = 0.04$. In contrast, the Pico Turquino – El

Gigante route (*M. cubensis* subsp. *cubensis*), with a comparable 41 km distance, showed much lower cost values (ca. 3,000), even when the fixation index is higher ($F_{ST} = 0.16$) (Table 5). In Guamuhaya, the least cost path between Topes de Collantes and Banao crosses the northern portion of the Agabama river valley (Fig. 6). In the Sierra Maestra, the cost path obtained between the different localities showed almost total overlap. For example, the Pico Caracas – El Gigante route implicitly includes the routes Pico Caracas – Turquino and Pico Caracas – La Bayamesa (Fig. 7). Taking into account that the similarity between the minimum cost paths is obtained independently, and the path obtained for the cost accumulated surface, the variables with the greatest contribution in the Guamuhaya massif are BIO4 and BIO15; while in the Sierra Maestra massif it is BIO12, BIO15 and MDE (Online Resource 4).

Discussion

1) Genetic diversity

The higher genetic diversity of *Magnolia cubensis* subsp. *cubensis* compared to *M. cubensis* subsp. *acunae* (Table 2) could be explained by the larger geographic extent of the population (Fig. 1b), the presence of more habitat heterogeneity (Capote and Berzaín 1984; Simón 2020), the habitat being conserved better (Palmarola et al. 2012; González-Torres et al. 2016), or a combination of the three. Gitzendanner and Soltis (2000) state that many fragmented habitats have lost the ability to support plant populations large enough to maintain a mutation-drift equilibrium. Therefore, the fragmentation of the habitat of *M. cubensis* subsp. *acunae* increases the risk that the population will reach a threshold below which it is not viable. Fortunately, its longevity guarantees that the time between generations is extended and therefore, the loss of alleles due to genetic drift is moderated (Honnay and Bossuyt 2005).

The mean values of allelic richness, observed and expected heterozygosity (Table 2) found for *M. cubensis* subsp. *acunae* exceeded that reported by Hernández et al. (2020a) (Topes de Collantes: $A_R = 4.10$ $H_o = 0.50$ $H_e = 0.54$; Lomas de Banao: $A_R = 2.93$ $H_o = 0.36$ $H_e = 0.41$), most likely because of the higher number of microsatellite markers used in the present study. The diversity measures were lower than the diversity values reported by Veltjen et al. (2019) in *M. cubensis* subsp. *cubensis* in Pico Turquino ($H_o = 0.60$ $H_e = 0.61$), but similar to the values of *M. cubensis* subsp. *acunae* in Topes de Collantes ($H_o = 0.60$ $H_e = 0.59$). These authors used a higher number of markers (30 loci), even when the number of samples used per subspecies was less (20 individuals). A similar behavior was reported by Hall and Beissinger (2014), who compared diversity values in several studies based on sample size and the number of markers used. Both results show that in population genetic studies, the cost-benefit relationship should guarantee an equilibrium between the number of markers and the number of samples.

When compared with other similar studies in non-Caribbean *Magnolia*, in general, the average values of the number of alleles per *locus*, allelic richness, observed and expected heterozygosity are lower to those reported for *Magnolia tomentosa* Thunb. (Setsuko et al. 2004), *M. stellata* (Siebold & Zucc.) Maxim. (Setsuko et al. 2007; Tamaki et al. 2008), *M. shiluensis* (Chun & Y.F.Wu) Figlar (as *Michelia shiluensis*

Chun & Y.F.Wu; Deng et al. 2020) and five *Magnolia* species of Veracruz, Mexico (Aldaba-Núñez et al. 2021); but higher than in *M. sieboldii* K. Koch (Kikuchi and Isagi 2002), *M. coriacea* (Hung T.Chang & B.L.Chen) Figlar (as *Michelia coriacea* Hung T.Chang & B.L.Chen; Zhao et al. 2012) and *M. wufengensis* L.Y.Ma & L.R.Wang (Wang et al. 2017).

When we compare the genetic diversity of the localities and their SNAP management category – which serves as a proxy for the degree of habitat disturbance, we do not retrieve that the habitats categorized to be in a higher degree of disturbance harbor less genetic diversity. In *Magnolia cubensis* subsp. *acunae*, the Topes de Collantes locality was the most diverse, while Lomas de Banao showed lower values for all the measures. However, Hernández et al. (2020a) found that, the Ecological Reserve Lomas de Banao has more conserved forests and less fragmentation than the Protected Natural Landscape Topes de Collantes, in correspondence with its management category (SNAP 2009). In *M. cubensis* subsp. *cubensis*, the most diverse locality was the National Park La Bayamesa and the least diverse were the Ecological Reserves Pico Caracas and El Gigante; yet all three have a similar degree of management category (SNAP 2009).

Our genetic diversity results may be determined by the number of individuals and the sample size of each locality. Poudel et al. (2014) demonstrated that the sample size can have a significant effect on the estimate of expected heterozygosity in microsatellite markers. Furthermore, Vergeer et al. (2003) state that the population size is positively correlated with the heterozygosity observed. However, most likely, the genetic diversity results are explained by demographic causes. On the one hand, Cuban magnolias were highly exploited for timber for many years (Carreras and Dechamps 1995; Álvarez 2003). On the other hand, low natural regeneration has been reported by Palmarola et al. (2018) in *M. cubensis* subsp. *acunae*, and by Molina-Pelegrín et al. (2014) and Testé et al. (2019) for *M. cubensis* subsp. *cubensis* in El Gigante and Gran Piedra, respectively. The four less genetically diverse localities (i.e. Lomas de Banao, El Gigante, Pico Caracas and Gran Piedra) hence may partly be explained by their demographic history.

2) Genetic structure

The PCoA analysis (Fig. 2a) and the structure analysis (Fig. 3a) showed two well-defined groups, corresponding to each subspecies. Currently, these taxa are considered subspecies (Imchanitzkaja 1991, 1993; Figlar and Nootboom 2012; Greuter and Rankin 2017). However, the large distance between their populations (ca. 400 km), morphological differentiation (Hernández 2014), ecological features (Simón 2020), phylogenetic evidence (Veltjen et al. in review), as well as our results, support the decision to consider them as different species.

In the analyses per subspecies, all spatial genetic patterns (genetic structure: Fig. 2b, c and Fig. 3b, c, isolation by distance and spatial autocorrelation: Fig. 4) show similar results. In *Magnolia cubensis* subsp. *acunae* there is no substructure, no isolation by distance and no spatial autocorrelation. In *M. cubensis* subsp. *cubensis*, there is substructure, with significant isolation by distance and significant values in the first five distance classes of the spatial autocorrelation analysis. The structure (Fig. 3c) and PCoA (Fig. 2c) analyses clearly separate the GRP (Gran Piedra) locality from the other three localities.

Next, the REC (Pico Caracas) locality is genetically detectable as a randomly mating population; however, there are migrants of this genetic cluster detectable in the TUR (Pico Turquino) populations. In the PCoA, REC is differentiating a little bit from the rest of the localities, but still overlaps.

3) Functional connectivity

The fixation index (F_{ST}) is a measure of the genetic divergence between subpopulations. Values close to zero, such as those found in *Magnolia cubensis* subsp. *acunae* (Table 3), indicate that there is similar heterozygosity in each locality compared to the value of the two localities together, so genetic differentiation is low (Hartl and Clark 1997). High values of this coefficient ($F_{ST} > 0.15$), such as those detected in *M. cubensis* subsp. *cubensis* (Table 3), show there has been less gene flow. However, between pairs of geographically contiguous localities such as Pico Caracas – Pico Turquino, Pico Turquino – La Bayamesa and La Bayamesa – El Gigante (Table 3), the genetic differentiation was moderate. If we compare our results with the F_{ST} values for Caribbean *Magnolia* (Veltjen et al. 2019), the F_{ST} values between some localities of *M. cubensis* subsp. *cubensis* (Pico Caracas – El Gigante: $F_{ST} = 0.23$, Gran Piedra and the rest: $0.20 < F_{ST} < 0.30$), are comparable to those found for sister species pairs more than for population pairs. In general, the genetic differentiation between localities supports the results of the genetic patterns. The locality of Pico Caracas presents values of genetic differentiation with Pico Turquino, similar to (or lower than) other pairs of localities that belong to the same genetic group (e.g., La Bayamesa – El Gigante). The results of genetic differentiation, together with those obtained in the Structure analysis and the PCoA, suggest to consider Gran Piedra as a separate genetic entity and not to appraise Pico Caracas as an independent genetic unit.

The clear genetic cluster in the Structure barplots for the Gran Piedra individuals (Fig. 3c), as well as their differentiation with the rest of the localities (Table 3), can be the cause of the large number of private allelic variants at this locality (Table 2), resulting from the geographical distance and isolation (Fig. 1b) or the diversification history of the Cuban magnolias. The geographical distance between localities can regulate the gene flow between them and has a direct effect on the existing genetic differentiation. The characteristic red-orange arilloid of the mature Magnoliaceae seeds (Judd et al. 2016) is easily visible and ingested by birds. As we do not find an indication of (recent) gene flow between Gran Piedra and the other localities, it could be that the distance of 88.49 km is either too far to overcome by the seed dispersers, which causes the barrier in gene flow; or it may be because of the landscape limiting the birds, even though they could overcome this distance.

In the study by Veltjen et al. (in review), the precise relationship among the two *Magnolia cubensis* subspecies and a third Cuban species from that subsection (*M. cristalensis* Bisse) was unclear, and the estimated divergence time of that clade was 5–2 million years ago. Hence, considering the simulations of Iturralde-Vinent (2006) and the work of Borhidi (1996), it is a possibility that Gran Piedra was the first Cuban locality to which *Magnolia* members of the subsection *Cubenses* migrated after which the rest of the Cuban localities were reached and diversification occurred. Our study shows clear genetic differences between the individuals of *M. cubensis* subsp. *cubensis* of Gran Piedra and those of the rest of the Sierra

Maestra mountain massif. This was not unexpected based on its complex taxonomic history. *Magnoliacubensis* subsp. *cubensis* is one of the Cuban magnolias that underwent numerous taxonomic changes, mainly at the infraspecific level. The current subspecies was treated by Imchanitzkaja (1991, 1993) as two different taxa: *Magnoliacubensis* subsp. *cubensis* with distribution in the Gran Piedra mountain range and *Magnolia cubensis* subsp. *turquinensis* Imkhan. from the Sierra Maestra mountain range. Recently, Greuter and Rankin (2017) synonymized *M. cubensis* subsp. *turquinensis* within *M. cubensis* subsp. *cubensis* based on the similarity between taxa in the leaf shape and leaf areole size.

Considering the different data (Fig. 2, 3, Table 2, 3), and according to the criteria of Ryder (1986) and Moritz (2002), we can define one evolutionary significant unit in *Magnolia cubensis* subsp. *acunae*, and two significant evolutionary units in *M. cubensis* subsp. *cubensis*: Pico Caracas – Turquino – La Bayamesa – El Gigante and Gran Piedra. Defining the conservation units within each subspecies makes it possible to preserve their evolutionary processes, previously not considered in conservation (Moritz 2002), and it is essential for the long-term survival of these threatened magnolias.

4) Structural connectivity

The coincidence between genetic and geographical barriers shown by Monmonier's algorithm, should not be taken as evidence of a relationship between gene flow patterns and the studied landscape features. This can be explained in a relatively simple way in *Magnolia cubensis* subsp. *cubensis*: if the landscape characteristics reflected the gene flow patterns, only large genetic differences would be found between Gran Piedra and the rest of the localities, due to the existence of Santiago de Cuba basin (CNNG 2000), which represents an area of differences in altitude, topography, vegetation and land use, between both mountain ranges (Sierra Maestra and Gran Piedra) (Fernández and Pérez 2008). However, the genetic fixation between Pico Caracas – El Gigante (both in the Sierra Maestra mountain range) was greater than between El Gigante – Gran Piedra (localized in different mountain ranges: the first in Sierra Maestra and the second one in Gran Piedra).

From the results of the partial Mantel tests, it can be inferred that the geographical, rather than the environmental, distances explain the genetic differences found between localities. However, statistically significant correlations were detected between geographic and environmental distances for the two taxa, by controlling for genetic effect. Environmental differences between the distribution areas of Cuban magnolias are reported by Simón (2020). The latter study found that the greatest niche similarity in the family did not correspond to the taxa that were phylogenetically closest, but instead to those that live in the same mountain range. However, this analysis was based on the comparison of climatic niche between species and did not include the possible microclimatic differences in the whole distribution range for each taxon.

The correlation between geographical and environmental distances can be explained by microclimatic differences. In general, the mountain systems where these magnolias live have higher precipitations and lower temperatures than those of the rest of the country (Durán 2012). In turn, a single mountain at a higher elevation corresponds to an increase in the average precipitation and a decrease in the average

temperature; therefore, individuals found at different elevations will also develop under different microclimatic conditions (Díaz 1988).

The differences in the relation of the geographical distances, the cost distances and the fixation index (F_{ST}) between localities for each subspecies, not necessarily indicate that they have different dispersal ability or seed dispersers. Their seeds are very similar (Judd et al. 2016), the possible dispersers in Cuba (Hernández et al. 2020a) are distributed over the whole island and are permanent residents (Garrido and Kirkconnell 2011). Hence, it is most likely that these differences are because the Least-cost-path models reflect the present environmental data, and not the environmental data from when the dispersal happened.

Although the Least-cost-path models may be related to a resistance to dispersal in each mountain massifs, it is necessary to clarify that the models were based on the habitat requirements of *Magnoliacubensis*. It is possible that the seed dispersers have similar requirements, in which case the model would represent the resistance to the movement of the disperser, and therefore, the gene flow of these species. Although the method has some limitations, the Least-cost-path models can be used for the establishment of biological corridors (Li et al. 2010; Alexander et al. 2016).

It should be noted that the analyses presented above (partial Mantel tests, Least-cost-path), attempted to evaluate the effect of landscape characteristics and some environmental variables on migration, dispersal and gene flow of *Magnoliacubensis*, not the possible interaction between the environment and the genetic variation of these taxa. Although some studies of landscape genetics in plants (Vandepitte et al. 2007; Schmidt et al. 2008) have focused on evaluating whether genetic diversity is related to soil type, elevation, humidity, temperature or vegetation, the type of marker used in the present study excludes this type of analysis. Neutral genetic diversity (such as that quantified by microsatellite markers) is not directly related to adaptive genetic variation and is only indirectly influenced by local environmental factors if the latter affect processes such as gene flow (Ouburg et al. 2010). To evaluate the direct effect of environmental factors on the genetic diversity of these magnolias, adaptive genetic diversity would have to be quantified through another type of molecular marker such as Single Nucleotide Polymorphism (SNPs). They make it possible to test which SNPs are in genes or close to genes under selection and then to correlate this to environmental variables.

Conclusions

In conclusion, the current landscape features do not clearly correlate with the patterns of genetic diversity; and the results appear to be taxon-specific. For *Magnoliacubensis* subsp. *acunae*, we do not find a well-defined spatial genetic pattern. The population structure does not correspond to the demographic groups or the topology of the localities, and the diversity values are not related to the degree of fragmentation of the habitat, but rather to the number of individuals. There is no evidence of isolation by distance or spatial autocorrelation. Furthermore, the little genetic differentiation between the two localities does not reflect the characteristics of the landscape that separates them or the cost values to cross it.

For *Magnoliacubensis* subsp. *cubensis*, we find evidence of isolation by distance and the autocorrelation analyses indicate that the approximate scale of the genetic neighborhood is between 35 and 40 km. In this subspecies, there is a marked genetic structure with probability values that indicate the existence of three genetic groups. However, for one of the defined genetic groups (El Gigante), the genetic differentiation with the other localities was moderate and migrants were detected. Large genetic differentiation was only found between Gran Piedra and the other localities, which reflects low gene flow. These results indicate that functional connectivity and landscape characteristics do not always correspond, yet that geographic distance does correlate with gene flow.

We propose to consider *Magnolia cubensis* subsp. *acunae* and *M. cubensis* subsp. *cubensis* as different species. Furthermore, our data suggest defining *M. cubensis* subsp. *acunae* as one evolutionarily significant unit and recognize two evolutionary significant units in *M. cubensis* subsp. *cubensis*: Pico Caracas – Turquino – La Bayamesa – El Gigante and Gran Piedra. We propose a Least-cost-path to connect the localities of *M. cubensis* subsp. *acunae* and those of *M. cubensis* subsp. *cubensis* in the Sierra Maestra mountain range.

The knowledge gained by this study must be combined with ecological, social and distribution data, to obtain a more accurate and realistic perspective of the conservation management strategies of these taxa. However, the examples in Cuba that complement management strategies with genetic or landscape data are practically non-existing. Therefore, in Cuba, conservation genetics still requires integration into conservation practice, at the level of legislation and management.

Declarations

Funding

This work was supported by the Whitley Fund for Nature, Franklinia Foundation and Planta! – Plantlife Conservation Society.

Conflicts of interest/Competing interests

Majela Hernández Rodríguez, Ernesto Testé, Emily Veltjen, Alejandro Palmarola Bejerano, Jhonny Quintana Delgado, Marlyn Valdés de la Cruz, Pieter Asselman, Isabel Larridon, Marie-Stéphanie Samain and Luis Roberto González Torres declare that they have no conflict of interests.

Availability of data and material

The datasets generated and analyzed during the current study are available from the corresponding author. However, most of the data generated during this study are included in this published article and its supplementary information files.

Code availability: Not applicable

Authors' contributions

Study design: Majela Hernández Rodríguez, Ernesto Testé Lozano, Luis Roberto González-Torres and Emily Veltjen. Field work and data collection: Majela Hernández Rodríguez, Ernesto Testé, Alejandro Palmarola Bejerano and Luis Roberto González Torres. Lab work: Majela Hernández Rodríguez, Emily Veltjen, Ernesto Testé, Pieter Asselman and Alejandro Palmarola Bejerano. Data analysis: Majela Hernández Rodríguez, Jhonny Quintana Delgado and Emily Veltjen. Writing and original draft preparation: Majela Hernández Rodríguez. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Additional declarations for articles in life science journals that report the results of studies involving humans and/or animals:

Ethics approval: Not applicable

Consent to participate: Not applicable

Consent for publication: Not applicable

Acknowledgments

This work was financially supported the Whitley Fund for Nature, Franklinia Foundation, Planta! – Plantlife Conservation Society, the Ghent University and the National Botanical Garden of Cuba. We would like to thank to the National Center of Protected Areas of Cuba for the support in the field trips for the collection of the samples. We are most grateful for all the people who helping us during field work: Norlys Albelo, José L. Moscoso, Alexei Rodríguez, Yenia Molina, Adonis Sosa, Daryl D. Cruz y Reinier Nuñez.

References

1. Adriaensen F, Chardon JP, De Blust G, Swinnen E, Villalba S, Gulinck H, Matthysen E (2003) The application of 'least-cost' modelling as a functional landscape model. *Landsc Urban Plan* 64(4):233–247
2. Aldaba-Núñez FA, Veltjen E, Martínez EM, Samain M-S (2021) Disentangling species delineation and guiding conservation of endangered Magnolias in Veracruz, Mexico. *Plants* 10(4):673
3. Alexander JL, Olimb SK, Bly KL, Restani M (2016) Use of least-cost path analysis to identify potential movement corridors of swift foxes in Montana. *J Mammal* 97(3):891–898
4. Álvarez A (2003) Estado de la diversidad biológica de los árboles y bosques en Cuba. Informe presentado en Taller Regional sobre los Recursos Genéticos Forestales de Centroamérica, Cuba y México
5. Anderson CD et al (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Mol Ecol* 19:3565–3575

6. Aparicio A, Hampe A, Fernández-Carrillo L, Albaladejo RG (2012) Fragmentation and comparative genetic structure of four mediterranean woody species: complex interactions between life history traits and the landscape context. *Divers Distrib* 18:226–235
7. Apodaca JJ, Rissler LJ, Godwin JC (2012) Population structure and gene flow in a heavily disturbed habitat: implications for the management of the imperilled Red Hills salamander (*Phaeognathus hubrichti*). *Conserv Genet* 13(4):913–923
8. Borhidi A (1996) *Phytogeography and Vegetation Ecology of Cuba*. Akademiai Kiado, Budapest
9. Brouns G, De Wulf A, Constaes D (2003) Delaunay triangulation algorithms useful for multibeam echosounding. *J Surv Eng* 129:79–84
10. Carreras R, Dechamps R (1995) Anatomía de la madera de 157 especies forestales que crecen en Cuba y sus usos tecnológicos, históricos y culturales. *Musee Royal de L’Afrique Centrale, Documentation Sciences Economiques Viol.* 9. Tervuren, Belgique
11. Capote RP, Berazaín R (1984) Clasificación de las formaciones vegetales de Cuba. *Jard. Bot. Nac-Univ. Habana.* 5(2)
12. Chavez-Cortazar A, Mata-Rosas M, Oyama K, Samain MS, Quesada M (2020) Induction of somatic embryogenesis and evaluation of genetic stability in regenerated plants of *Magnolia dealbata*. *Biol Plant* 64:224–233
13. Cicuzza D, Newton A, Oldfield S (2007) *The red list of Magnoliaceae*. Fauna and Flora International, Cambridge
14. CNNG (2000) *Diccionario geográfico de Cuba*. Oficina Nacional de Hidrografía y Geodesia, La Habana, Cuba
15. 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
16. Díaz LR (1988) *Nuevo Atlas Nacional de Cuba*. In: Gutiérrez GO, Sánchez EA, Hernández JR et al (eds) *Instituto Cubano de la Academia de ciencias de Cuba*. XXIV
17. Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
18. Durán I (2012) *Comportamiento de la precipitación en estaciones meteorológicas seleccionadas de Cuba*. Dissertation, Facultad de Geografía, Universidad de La Habana
19. Earl DA, vonHoldt BM (2012) Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4(2):359–361
20. Epps CW, Palsb Il PJ, Wehausen JD, Roderick GK, Ramey IIRR, McCullough DR (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecol Lett* 8:1029–1038
21. Estrada R, Martín G, Martínez P et al (2012) Mapa (BD-SIG) de vegetación natural y seminatural de Cuba v.1 sobre Landsat EMT 7 slc-off gap filled, circa 2011. *IV Congreso de Biodiversidad y Ecosistemas*

22. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
23. Figlar RB, Nootboom HP (2004) Notes on Magnoliaceae IV. *Blumea* 49(1):87–100
24. Frantz AC, Bertouille S, Eloy MC, Licoppe A, Chaumont F, Flamand MC (2012) Comparative landscape genetic analysis show a Belgian motorway to be a gene flow barrier for red deer (*Cervus elaphus*), but not wild boars (*Sus scrofa*). *Mol Ecol* 21(14):3445–3457
25. Garrido OH, Kirkconnell A (2011) *Aves de Cuba*. Cornell University Press, Ithaca
26. Fernández A, Pérez de los Reyes R (eds) (2008) *GeoCuba: evaluación del Medio Ambiente cubano*. Ministerio de Ciencia, Tecnología y Medio Ambiente
27. Gitzendanner MA, Soltis P (2000) Patterns of genetic variation in rare and widespread plant congeners. *Am J Bot* 87(6):783–792
28. González-Torres LR, Palmarola A, González Oliva L, Bécquer ER, Testé E, Barrios D (eds) (2016) *Lista roja de la flora de Cuba*. *Bissea* 10(NE 1):1-352
29. Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *J Hered* 86:485–486
30. Granado L (2015) *Estructura poblacional, distribución geográfica y conservación de Magnolia cubensis subsp. acunae* (Magnoliaceae). Dissertation, Facultad de Biología, Universidad de La Habana
31. Greuter W, Rankin R (2017) *Espermatófitos de Cuba - Inventario preliminar*. Botanischer Garten und Botanisches Museum Berlin Zentraleinrichtung der Freien Universität, Berlin
32. Hall LA, Beissinger SR (2014) A practical toolbox for design and analysis of landscape genetics studies. *Landsc Ecol* 29(9):1487–1504
33. Hammer O, Harper DAT, Ryan PD (2001) PAST: Paleontological Statistics software package for education and data analysis. *Palaeont Electr* 4(1):9
34. Hartl DL, Clark AG (1997) *Principles of Population Genetics*, 3rd edn. Sinauer, Sunderland
35. Hernández M (2014) Identificación de las subespecies de *Magnolia cubensis* (Magnoliaceae) mediante análisis digital de imágenes de las hojas. *Rev Cub de Cienc Bio* 3(2):53–60
36. Hernández M, Palmarola Bejerano A, Veltjen E, Asselman P, Testé E, Larridon I, González-Torres LR (2020a) Population structure and genetic diversity of *Magnolia cubensis* subsp. *acunae* (Magnoliaceae): effects of habitat fragmentation and implication for conservation. *Oryx* 54:1–9
37. Hernández M, Testé E, Albelo N, Valle O, Simón R, Palmarola Bejerano A (2020b) *Magnolia cubensis* subsp. *acunae* - Hoja de taxón. *Bissea* 14(NE 1)
38. Hernández M, Testé E, Simón R, Palmarola Bejerano A, González-Torres L, Molina, Y (2020c) *Magnolia cubensis* subsp. *cubensis* - Hoja de taxón. *Bissea* 14(NE 1)
39. Hirayama K, Ishida K, Setsuko S, Tomaru N (2007) Reduced seed production, inbreeding, and pollen shortage in a small population of a threatened tree. *Magnolia stellata* *Biol Conserv* 136:315–323

40. Honnay O, Bossuyt B (2005) Prolonged clonal growth: escape route or route to extinction? *Oikos* 108(2):427–432
41. Honnay O, Jacquemyn H (2007) Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conserv Biol* 21(3):823–831
42. Imchanickaja NN (1991) Rod *Magnolia* L. (Magnoliaceae) vo flore Kuby. *Novosti Sist Vysš Rast* 28:58–77
43. Imkhanitzkaja NN (1993) Genus *Talauma* A. L. Juss. (Magnoliaceae) in flora Cubae. *Novosti Sist Vysš Rast* 29:76–84
44. Isagi Y, Tateno R, Matsuki Y, Hirao A, Watanabe S, Shibata M (2007) Genetic and reproductive consequences of forest fragmentation for populations of *Magnolia obovata*. *Ecol Res* 22:382–389
45. Iturralde-Vinent MA (2006) Meso-Cenozoic Caribbean Paleogeography: Implications for the Historical Biogeography of the Region. *Int Geol Rev* 48:791–827
46. Jost L (2008) G(ST) and its relatives do not measure differentiation. *Mol Ecol* 17:4015–4026
47. Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2016) *Plant systematics: a phylogenetic approach*, 4th edn. Sinauer Associates, Sunderland
48. Kearse M, Moir R, Wilson A et al (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649
49. Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA, O'Hara RB (2013) diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol* 4:782–788
50. Kikuchi S, Isagi Y (2002) Microsatellite genetic variation in small and isolated populations of *Magnolia sieboldii* ssp. *japonica*. *Heredity* 88:313–321
51. Kolb A, Diekmann M (2005) Effects of life history traits on responses of plant species to forest fragmentation. *Conserv Biol* 19(3):929–938
52. Li H, Li D, Li T, Qiao Q, Yang J, Zhang H (2010) Application of least-cost path model to identify a giant panda dispersal corridor network after the Wenchuan earthquake—Case study of Wolong Nature Reserve in China. *Ecol Model* 221(6):944–952
53. Luijten S, Dierick HA, Oostermeijer JGB, Raijmann LEL, den Nijs JCM (2000) Population size, genetic variation, and reproductive success in a rapidly declining, self-incompatible perennial (*Arnica montana*) in The Netherlands. *Biol Conserv* 14:1776–1787
54. Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol* 18(4):189–197
55. Miller MP (2005) Alleles In Space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. *J Hered* 96(6):722–724
56. Molina Y, Santos-Chacón W, Sosa-López A, Arcia-Chávez M, Hechavarría-Kindelán O, Rosales-Rodríguez M (2014) Estructura poblacional de *Magnolia cubensis* Urb. subsp. *cubensis* en la Reserva

57. Monmonier MS (2010) Maximum-difference barriers: an alternative numerical regionalization method. *Geogr Anal* 5:245–261
58. Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst Biol* 51(2):238–254
59. Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70(12):3321–3323
60. Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. *Ann Hum Genet* 47(3):253–259
61. Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J Mol Evol* 19:153–170
62. Oksanen J, Blanchet FG, Friendly M et al (2019) *vegan: Community Ecology Package*. R package version 2.5. 4. 2019
63. Ouburg NJ et al (2010) An essay on the necessity and feasibility of conservation genomics. *Conserv Genet* 11:643–653
64. Palmarola A, González-Torres LR, Barrios D, Albelo N, León J (2012) Proyecto de Conservación integral del “mantequero” en Guamuhaya. *Bissea* 6(2):2
65. Palmarola A, González-Torres LR, Granado L, Hernández M (2015a) *Magnolia cubensis* subsp. *acunae*. *Bissea* 9(NE 4):405
66. Palmarola A, González-Torres LR, Hernández M, Molina Y, Gordillo M, Granado L, Cruz DD (2015b) *Magnolia cubensis* subsp. *cubensis*. *Bissea* 9(NE 4):406–407
67. Palmarola A, Romanov MS, Bobrov A, González-Torres yLR (2016) Las magnolias de Cuba: *Talauma* - problemática taxonómica y nomenclatural. *Rev. Jard. Bot. Nac-Univ. Habana*. 37
68. Palmarola A, Granado L, Testé E, Hernández M, Albelo N, González-Torres LR (2018) Estructura poblacional y distribución de *Magnolia cubensis* subsp. *acunae* (Magnoliaceae). *Rev. Jard. Bot. Nac-Univ Habana* 39:103–111
69. Peakall R, Smouse PE (2012) GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics* 28:2537–2539
70. Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annu Rev Ecol Syst* 37:187–214
71. Poudel RC, Moller M, Li DZ, Shah A, Gao LM (2014) Genetic diversity, demographical history and conservation aspects of the endangered yew tree *Taxus contorta* (syn. *Taxus fuana*) in Pakistan. *Tree Genet Genomes* 10:653–665
72. Core Team R (2020) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna
73. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959

74. Rico Y, Gutierrez Becerril BA (2019) Species delimitation and genetic structure of two endemic *Magnolia* species (section *Magnolia*; Magnoliaceae) in Mexico. *Genetica* 147:57–68
75. Rivers M, Beech E, Murphy L, Oldfield S (2016) The Red List of Magnoliaceae. Revised and extended. Botanic Gardens Conservation International, Richmond (UK)
76. Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Eco Res* 8:103–106
77. Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends Ecol Evol* 1:9–10
78. Schmidt T et al (2008) Effects of landscape structure on genetic diversity of *Geum urbanum* L. populations in agricultural landscapes. *Flora* 204:549–559
79. Scribner KT, Blanchong JA, Bruggeman DJ, Epperson BK, Lee CY, Pan YW, Luukkonen DR (2005) Geographical genetics: conceptual foundations and empirical applications of spatial genetic data in wildlife management. *J Wildl Manag* 69(4):1434–1453
80. Setsuko S, Ishida K, Tomaru N (2004) Size distribution and genetic structure in relation to clonal growth within a population of *Magnolia tomentosa* Thunb. (Magnoliaceae). *Mol Ecol* 13:2645–2653
81. Setsuko S, Ishida K, Ueno S, Tsumura Y, Tomaru N (2007) Population differentiation and gene flow within a metapopulation of a threatened tree, *Magnolia stellata* (Magnoliaceae). *Am J Bot* 94:128–136
82. Simón R (2020) Análisis de la distribución y el nicho ecológico de *Magnolia* sect. *Talauma* en Cuba. Master dissertation, Facultad de Biología, Universidad de La Habana
83. Sinnott RW (1984) Virtues of the Haversine. *SandT* 68(2):158
84. Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82:561–573
85. SNAP (Centro Nacional de Áreas Protegidas) (2009) Plan del Sistema Nacional de Áreas Protegidas 2009–2013. Centro Nacional de Áreas Protegidas, La Habana, Cuba
86. Tamaki I, Kawashima N, Setsuko S, Itaya A, Tomaru N (2018) Morphological and genetic divergence between two lineages of *Magnolia salicifolia* (Magnoliaceae) in Japan. *Biol J Linn Soc* 125(3):475–490
87. Tamaki I, Setsuko S, Tomaru N (2008) Genetic variation and differentiation in populations of a threatened tree, *Magnolia stellata*: factors influencing the level of within-population genetic variation. *Heredity* 100:415–423
88. Testé E, González-Oliva L, Márquez A (2015) Invasión actual y potencial del árbol tóxico *Rhus succedanea* (Anacardiaceae) en el Paisaje Natural Protegido Topes de Collantes. *Cuba Rev Jard Bot Nac-Univ Habana* 36:173–180
89. Testé E, Gordillo M, Palmarola A, Hernández M, González-Torres LR (2019) Estructura poblacional de *Magnolia cubensis* subsp. *cubensis* (Magnoliaceae) en el Paisaje Natural Protegido Gran Piedra. *Rev. Jard. Bot. Nac-Univ Habana* 40:19–21

90. Thien LB (1974) Floral biology of Magnolia. *Amer J Bot* 61(10):1037–1045
91. Vandepitte K et al (2007) Landscape genetics of the self-compatible forest herb *Geum urbanum*: effects of habitat age, fragmentation and local environment. *Mol Ecol* 16:4171–4179
92. Vásquez SG, Sánchez LR, Pineda MR (2001) Las semillas de la magnolia, 3. *Revista de Divulgación Científica y Tecnológica de la Universidad Veracruzana Volumen XXIII*
93. Veltjen E, Asselman P, Hernández M, Palmarola A, Testé E, González-Torres LR, Goetghebeur P, Larridon I, Samain M-S (2019) Genetic patterns in Neotropical Magnolias (Magnoliaceae) using *de novo* developed microsatellite markers. *Heredity* 122(4):485–500
94. Veltjen E, Testé E, Bejerano A, Asselman P, Hernández M, González-Torres LR, Chatrou LW, Goetghebeur P, Larridon I, Samain M-S (in review) The evolutionary history of the Caribbean Magnolias (Magnoliaceae): testing species delimitations and biogeographical hypotheses using molecular data. *Mol. Phylogenetics Evol*
95. Vergeer P, Rengeling R, Copal A, Ouborg NJ (2003) The interacting effects of genetic variation, habitat quality and population size on performance of *Succisa pratensis*. *J Eco* 91:18–26
96. Wang L, Xiao AH, Ma LY, Chen FJ, Sang ZY, Duan J (2017) Identification of *Magnolia wufengensis* (Magnoliaceae) cultivars using phenotypic traits, SSR and SRAP markers: Insights into breeding and conservation. *Genet. Mol. Res.* 2017: 16
97. Waples RS (2015) Testing for Hardy-Weinberg proportions: have we lost the plot? *J Hered* 106(1):1–19
98. Watson DF (1992) *Contouring: a guide to the analysis and display of spatial data*. Pergamon Press, New York
99. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evo* 38:1358–1370
100. Wright S (1943) Isolation by distance. *Genetics* 28:114–138
101. Xin Z, Chen J (2012) A high throughput DNA extraction method with high yield and quality. *Plant Methods* 8(26)
102. Zhao X, Ma Y, Sun W, Wen X, Milne R (2012) High genetic diversity and low differentiation of *Michelia coriacea* (Magnoliaceae), a critically endangered endemic in southeast Yunnan, China. *Int J Mol Sci* 13:4396–4411

Tables

Table 1 Demographic information of *Magnolia cubensis*, collected during expeditions between 2010-2020. The table includes acronyms for each locality (Pop column), the know individuals (IND), the number of DNA samples (N), the number of final samples (N_f) after to delete the individuals in whom more than 70% of markers could not be genotyped and between parenthesis the percentage (%) that represent the samples of the total know individuals per locality and taxa

Taxa	Locality	Pop	IND	N	N _F (%)
<i>Magnolia cubensis</i>	NPL 'Topes de Collantes'	TOP	368	39	36 (9.8)
subsp. <i>acunae</i>	ER 'Lomas de Banao'	BAN	70	9	9 (12.9)
		Total	438	48	45 (10.3)
<i>Magnolia cubensis</i>	ER 'Pico Caracas'	REC	50	36	36 (72.0)
subsp. <i>cubensis</i>	NP 'Pico Turquino'	TUR	804	46	39 (4.9)
	NP 'La Bayamesa'	BAY	319	50	47 (14.7)
	ER 'El Gigante'	GIG	48	17	16 (33.3)
	NPL 'Gran Piedra'	GRP	60	30	28 (46.7)
		Total	1282	179	166 (13.0)

NPL: Natural Protected Landscape; ER: Ecological Reserve; NP: National Park.

Table 2 Measures of genetic diversity per locality and taxa of *Magnolia cubensis*, estimated by simple sequence repeat markers. The number of SSR markers per taxon is given between parenthesis

Taxa	Pop	N	A (min-max)	A _e	A _R	H _o	H _e	HW	F _{IS}	P (%)
<i>Magnolia cubensis</i>	TOP	36	6.16 (1-11)	3.55	4.54	0.60	0.64	0	0.069	94.74
subsp. <i>acunae</i>	BAN	9	3.74 (1-6)	2.25	3.74	0.56	0.53	0	0.002	94.74
(19)	Total	45	4.72 (1-11)	2.80	4.46	0.56	0.58	-	0.031	94.74
<i>Magnolia cubensis</i>	REC	36	5.57 (1-17)	3.08	4.32	0.43	0.51	4	0.182	85.71
subsp. <i>cubensis</i>	TUR	39	6.52 (2-17)	3.43	5.18	0.37	0.57	9	0.364	100.00
(21)	BAY	47	6.76 (1-16)	4.14	5.24	0.50	0.56	3	0.110	90.48
	GIG	16	5.10 (1-13)	3.24	4.80	0.45	0.55	2	0.209	85.71
	GRP	28	5.86 (2-11)	3.16	4.82	0.47	0.57	6	0.207	100.00
	Total	166	5.96 (1-17)	3.41	6.46	0.44	0.55	-	0.249	92.38

Pop: locality acronyms according to Table 1. **N:** number of sampled individuals. **A:** average number of alleles per *locus*. **A_e:** effective alleles. **A_R:** allelic richness. **Ho:** observed heterozygosity. **He:** expected heterozygosity. **HW:** number of *loci* that deviated from Hardy–Weinberg proportions. **F_{IS}:** inbreeding coefficient calculated according to Weir & Cockerham (1984). **P:** percentage of polymorphic *loci*. *p < 0.005 Bonferroni corrected probability were considered statistically significant. Detailed information on the genetic diversity estimates per *locus* in *Magnolia cubensis* subsp. *cubensis* is given in Quintana *et al.* (in review).

Table 3 Genetic differentiation between localities of two taxa of *Magnolia cubensis*. The number of SSR markers per taxon is given between parentheses

Taxa	Locality						
	TOP	BAN	REC	TUR	BAY	GIG	GRP
<i>Magnolia cubensis</i> subsp. <i>acunae</i> (19)		0.036 (F _{ST})					
		0.021 (G _{ST})					
	BAN	0.035					
<i>Magnolia cubensis</i> subsp. <i>cubensis</i> (21)	REC		0.115 (F _{ST})		0.168 (F _{ST})	0.229 (F _{ST})	0.290 (F _{ST})
			0.065 (G _{ST})		0.095 (G _{ST})	0.128 (G _{ST})	0.170 (G _{ST})
	TUR	0.132			0.089 (F _{ST})	0.158 (F _{ST})	0.267 (F _{ST})
					0.050 (G _{ST})	0.087 (G _{ST})	0.153 (G _{ST})
	BAY	0.193	0.089			0.128 (F _{ST})	0.245 (F _{ST})
						0.069 (G _{ST})	0.141 (G _{ST})
	GIG	0.306	0.205	0.178			0.206 (F _{ST})
							0.117 (G _{ST})
	GRP	0.414	0.422	0.379	0.326		

Locality acronyms are according to Table 1. Pairwise fixation indices: F_{ST} (Weir and Cockerham, 1984) and G_{ST} (Nei, 1973; Nei and Chesser, 1983), are placed above the diagonal; and allelic differentiation: D_{JOST} (Jost, 2008) is placed below the diagonal.

Table 4 Partial Mantel test values (r^2) and their significance (p) for the correlation between lineal genetic distances, environmental distances and geographic distances (km) among localities of *Magnolia cubensis*. The number of SSR markers per taxon is given between parentheses

Taxa	Partial Mantel test	r_s	p
<i>Magnolia cubensis</i> subsp. <i>acunae</i> (19)	(LinGD x EnvD) GGD _R	-0,044	0,688
	(GGD _R x EnvD) LinGD	0,152	0,001*
	(LinGD x GGD _R) EnvD	0,017	0,428
	(LinGD x BIO4D) GGDR	0,052	0,266
	(GGD _R x BIO4D) LinGD	0,902	0,001*
	(LinGD x GGD _R) BIO4D	-0,049	0,699
	(LinGD x BIO12D) GGD _R	-0,055	0,733
	(GGD _R x BIO12D) LinGD	0,810	0,001*
	(LinGD x GGD _R) BIO12D	0,041	0,266
	(LinGD x BIO15D) GGD _R	-0,031	0,619
	(GGD _R x BIO15D) LinGD	0,502	0,001*
	(LinGD x GGD _R) BIO15D	0,009	0,375
<i>Magnolia cubensis</i> subsp. <i>cubensis</i> (21)	(LinGD x EnvD) GGD _R	-0,048	0,924
	(GGD _R x EnvD) LinGD	0,172	0,001*
	(LinGD x GGD _R) EnvD	0,752	0,001*
	(LinGD x BIO4D) GGD _R	-0,089	1,000
	(GGD _R x BIO4D) LinGD	0,355	0,001*
	(LinGD x GGD _R) BIO4D	0,736	0,001*
	(LinGD x BIO12D) GGD _R	-0,067	0,999
	(GGD _R x BIO12D) LinGD	0,278	0,001*
	(LinGD x GGD _R) BIO12D	0,744	0,001*
	(LinGD x BIO15D) GGD _R	-0,123	1,000
	(GGD _R x BIO15D) LinGD	0,275	0,001*
	(LinGD x GGD _R) BIO15D	0,755	0,001*

The partial Mantel tests are defined in the format (A x B) C whereby A and B are the variables that are being tested for partial correlation, while the variable C is controlled. LinGD: lineal genetic distances; EnvD: environmental distances; GGD_R: real geographic distances; BIO4D: Temperature Seasonality distance; BIO12D: Annual Precipitation distance; BIO15D: Precipitation Seasonality distance.

Table 5 Distance between localities, cost values derived from the Least-cost-path models and F_{ST} for all comparisons between localities in *Magnolia cubensis*. Spearman's correlation coefficients (r_s) and associated probability (p) are included for the correlations. The pair of localities are organized by geographic proximity. Locality acronyms according to Table 1

Species	Pair of localities	Distance (km)	Cost	F_{ST}
<i>Magnolia cubensis</i> subsp. <i>acunae</i>	TOP-BAN	43.29	7,517.88	0.04
<i>Magnolia cubensis</i> subsp. <i>cubensis</i>	TUR-BAY	14.67	499.66	0.09
	REC-TUR	18.93	1,418.73	0.12
	BAY-GIG	26.91	2,283.51	0.13
	REC-BAY	33.54	1888.40	0.17
	TUR-GIG	41.32	2783.17	0.16
	REC-GIG	60.24	4171.91	0.23
	GIG-GRP	88.49	27,980.31	0.21
	BAY-GRP	112.20	29,969.24	0.25
	TUR-GRP	126.60	30,468.90	0.27
	REC-GRP	144.83	31,857.64	0.29

Correlation geographic distance - cost: $r_s = 0.94$; $p < 0.0001$ *

Correlation cost - F_{ST} : $r_s = 0.74$; $p < 0.0001$ *

Correlation geographic distance - F_{ST} : $r_s = 0.81$; $p < 0.0001$ *

Figures

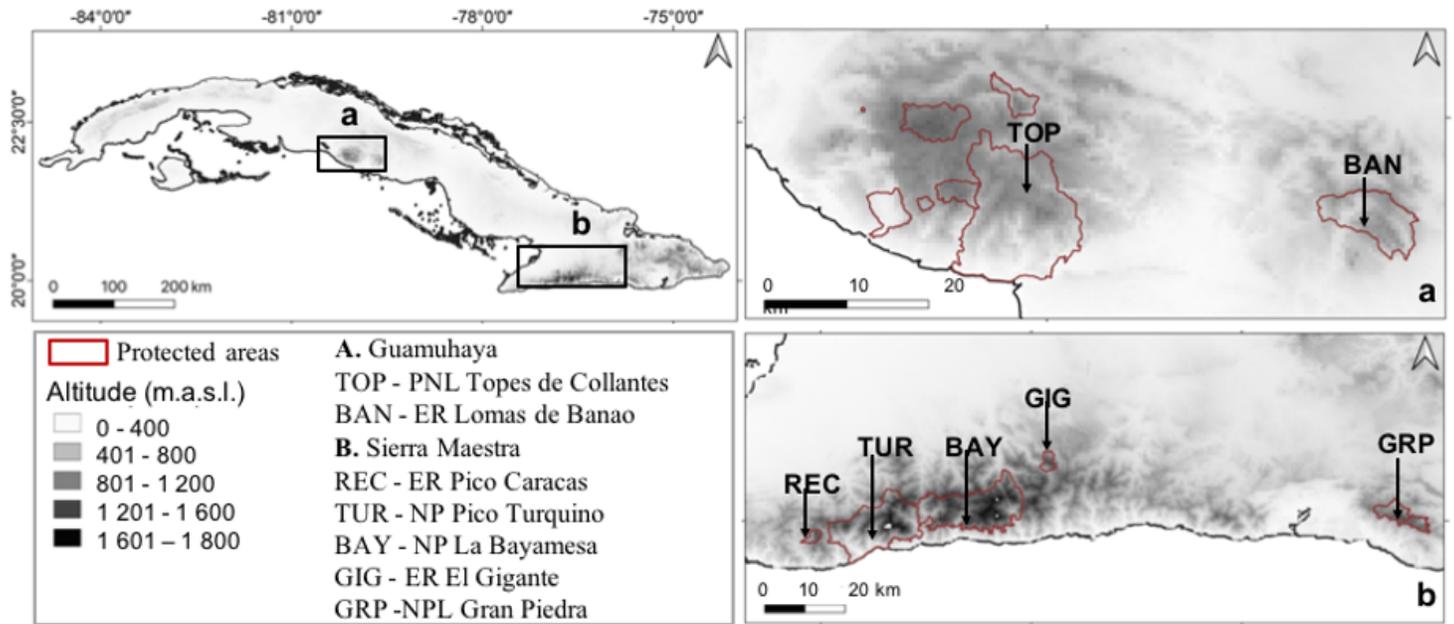


Figure 1

Study areas and sample localities of *Magnolia cubensis* in Guamuhaia and Sierra Maestra massifs, in Cuba. Locality acronyms correspond to those used in Table 1. NPL: Natural Protected Landscape; ER: Ecological Reserve; NP: National Park (according SNAP 2009); m.a.s.l.: meters above the sea level Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

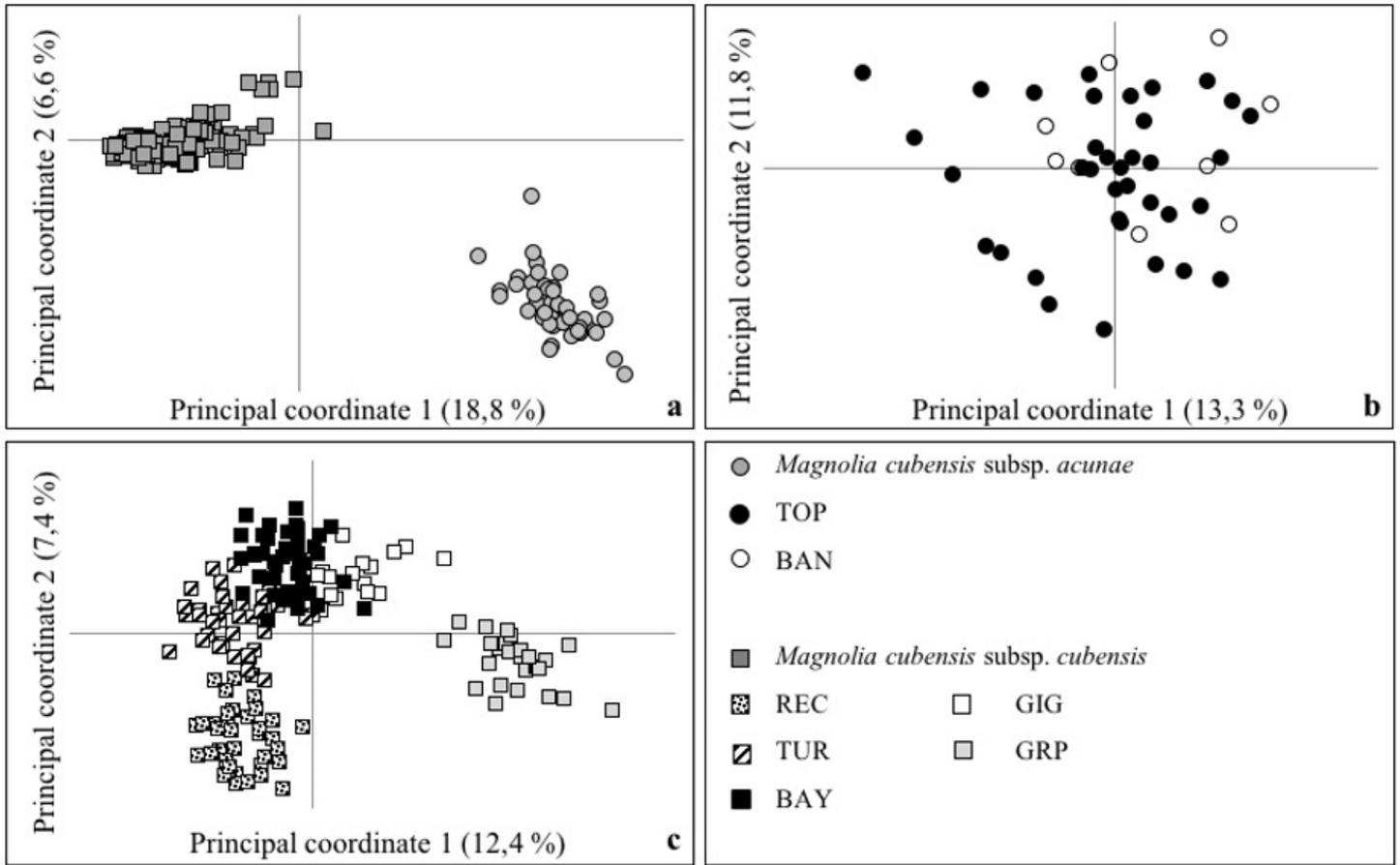


Figure 2

Principal coordinates analysis based on an SSR genetic distance matrices of *Magnolia cubensis* individuals. a. Analysis including both subspecies (19 SSR markers). b-c. Analysis per subspecies: *Magnolia cubensis* subsp. *acunae* (19 SSR markers) (b); *Magnolia cubensis* subsp. *cubensis* (21 SSR markers) (c). Locality acronyms are according to Table 1

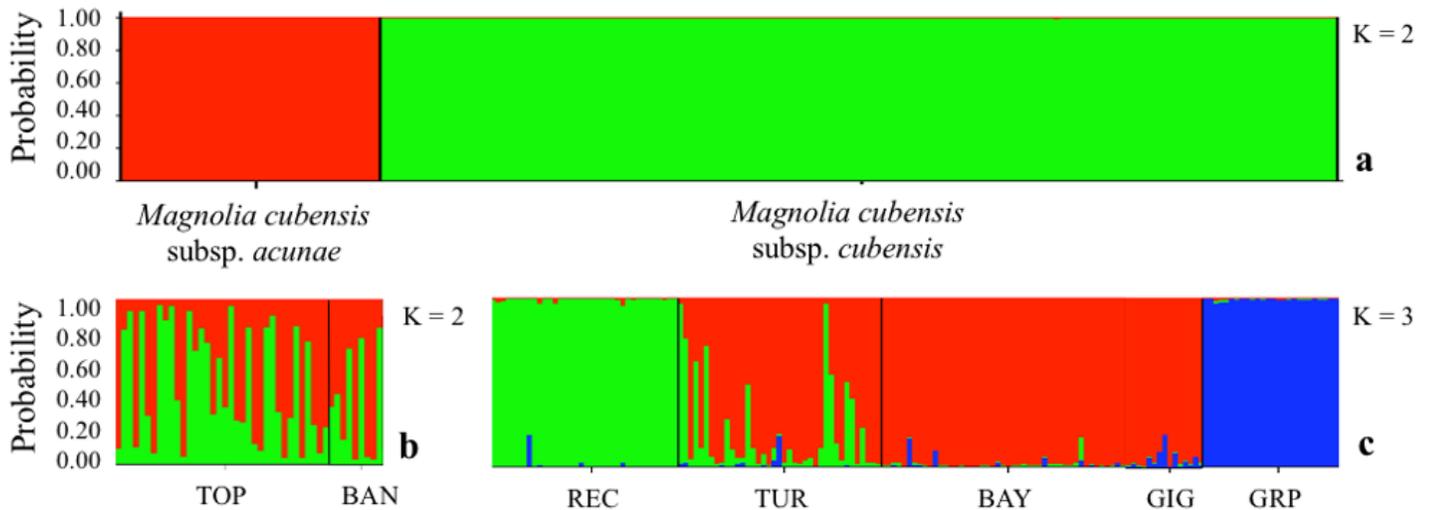


Figure 3

Structure barplots that illustrate the probability of genetic group allocation of *Magnolia cubensis* individuals, based on allelic frequencies of SSR data. a. Analysis including both subspecies. b-c. Analysis per subspecies: *Magnolia cubensis* subsp. *acunae* (b); *Magnolia cubensis* subsp. *cubensis* (c) (modified from Quintana et al. in review). Locality acronyms are according to Table 1

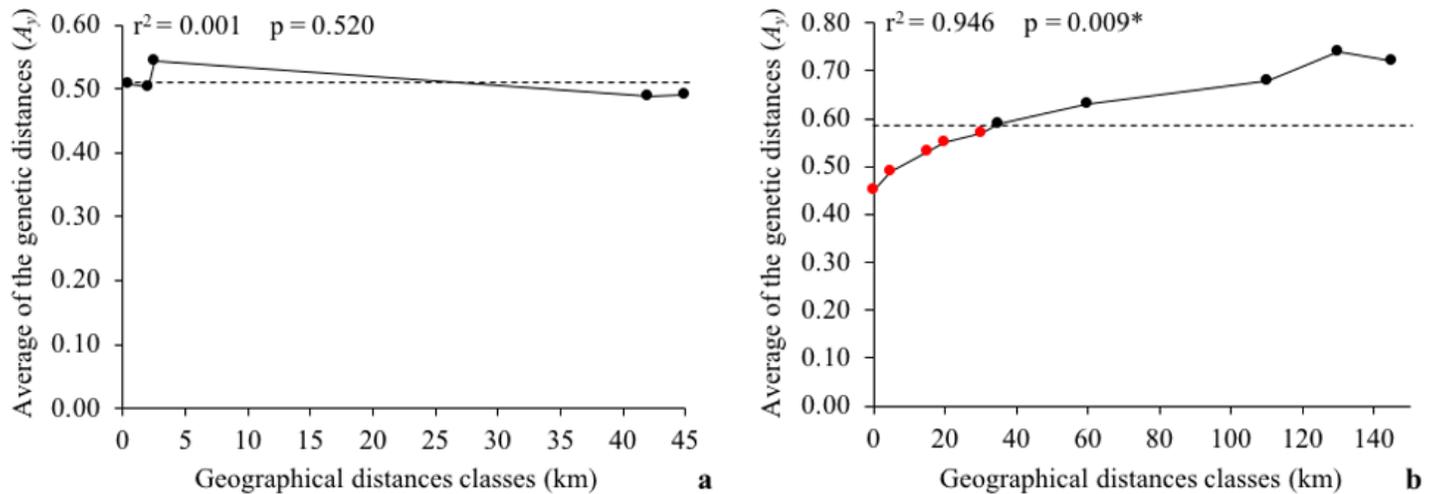


Figure 4

Spatial autocorrelation analyses for *Magnolia cubensis*. Spatial autocorrelation is visualized by the red datapoints that have a probability of observing a random value of $V \geq$ observed V by chance. Mantel test correlation coefficient (r^2) and their significance (p) for the correlation between lineal genetic distances and geographic distances (km). $*p < 0.01$ was considered statistically significant. a. *Magnolia cubensis* subsp. *acunae*. b. *Magnolia cubensis* subsp. *cubensis*

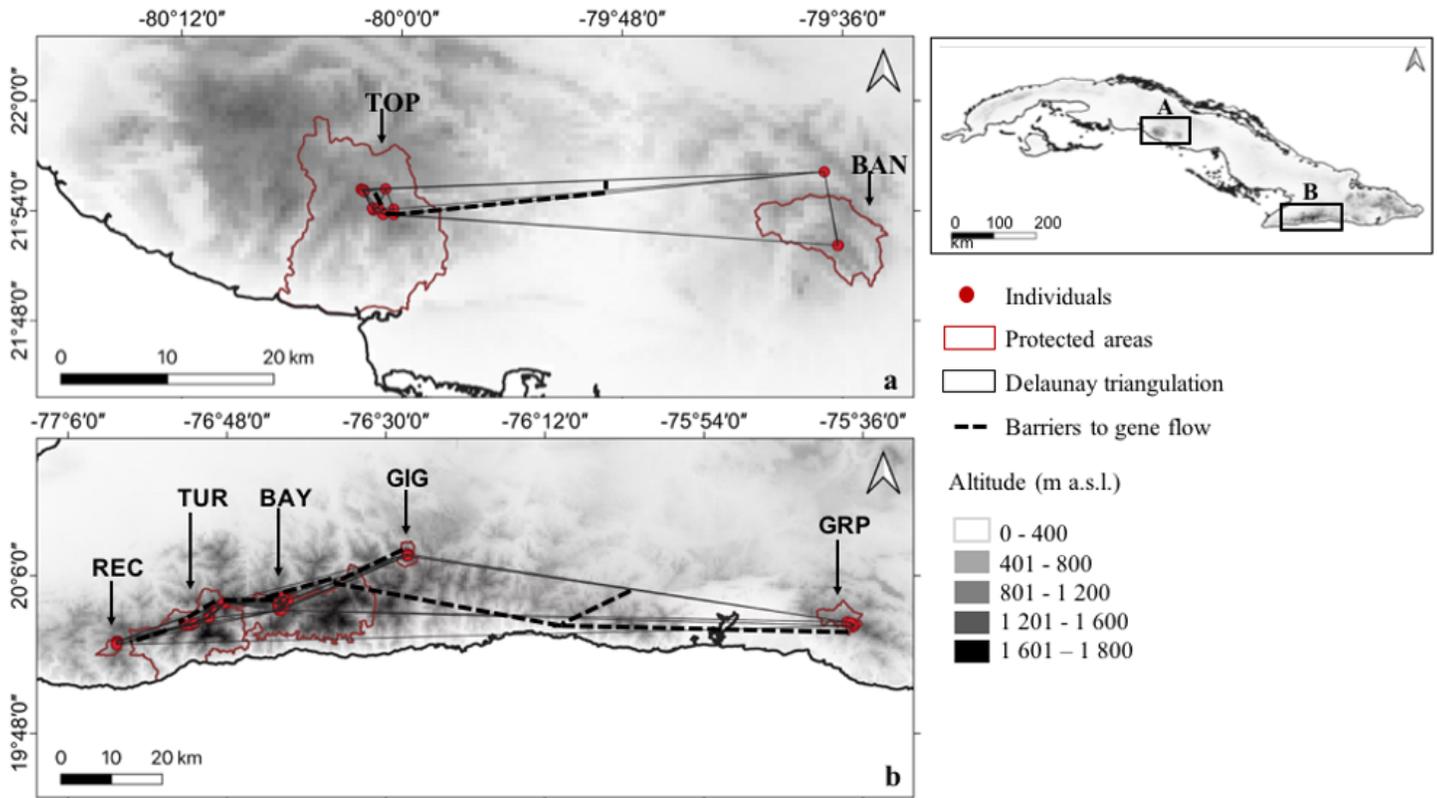


Figure 5

Barriers to gene flow between localities of *Magnolia cubensis* as detected by Monmonier's algorithm corresponding to the Guamuhaya (a) and Sierra Maestra (b) massifs, in Cuba. Light solid lines represent the Delaunay triangulation and heavy dashed lines indicate barriers to gene flow. Locality acronyms according to Table 1 Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

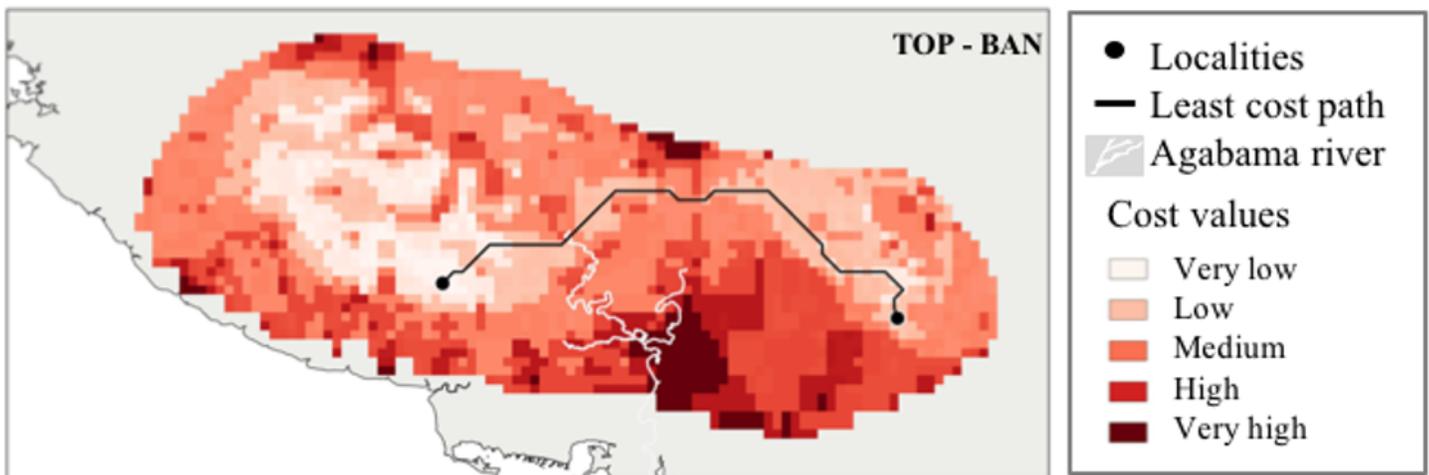


Figure 6

Connectivity between localities of *Magnolia cubensis* subsp. *acunae* in Guamuhaya, Cuba developed with Least-cost-path (LCP) analysis. Black lines depict Least-cost-paths. The resistance surface is shown as a grid gradient in function of the accumulated cost values, from light (very low cost) to dark red (very high cost); white circles are sampling sites. Locality acronyms are according to Table 1 Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

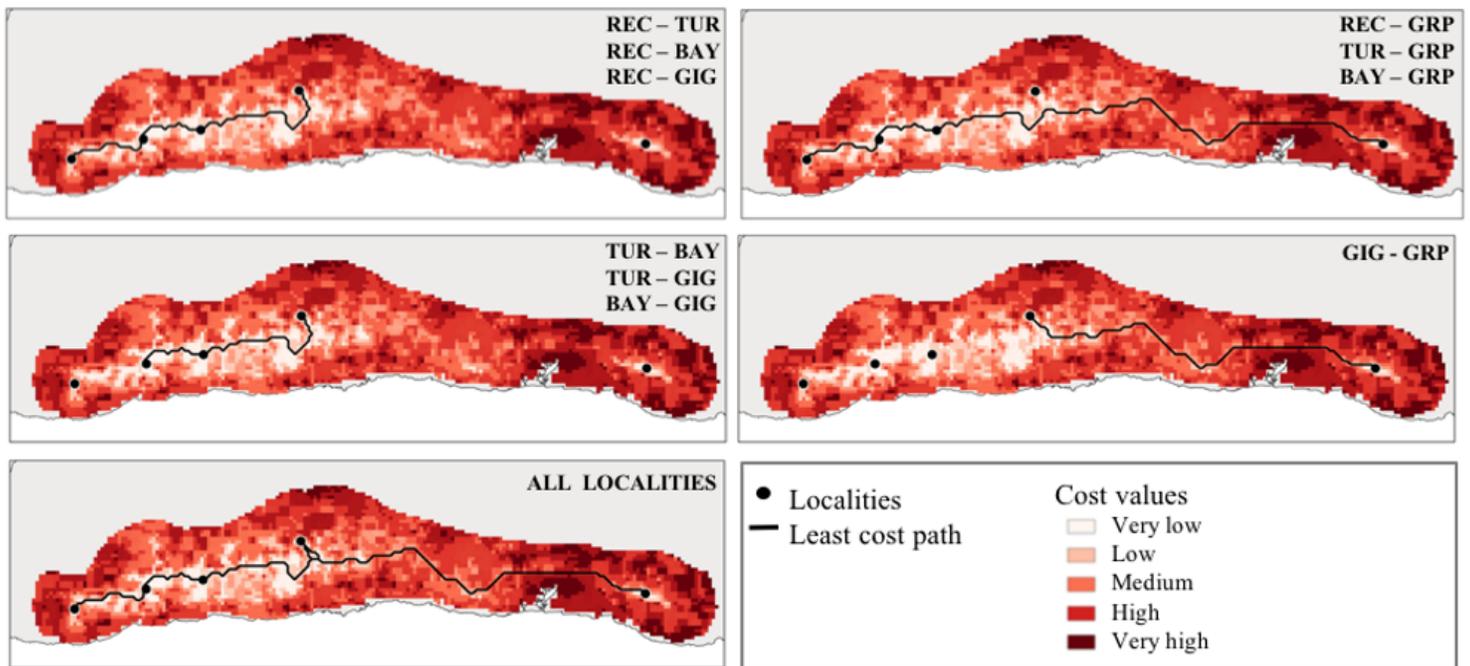


Figure 7

Connectivity between localities of *Magnolia cubensis* subsp. *cubensis* in Sierra Maestra, Cuba developed with Least-cost-path (LCP) analysis. Black lines depict Least-cost-paths. The resistance surface is shown as a grid gradient in function of the accumulated cost values, from light (very low cost) to dark red (very high cost); white circles are sampling sites. It is showed the Least-cost-paths for each pair of localities and between all localities. Locality acronyms according to Table 1 Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM1..docx](#)

- [ESM2.docx](#)
- [ESM3.docx](#)
- [ESM4.docx](#)