Effect of the landscape on functional and spatial connectivity in *Magnolia cubensis* (Magnoliaceae) in two mountain massifs of Cuba

Majela Hernández Rodríguez1 · Ernesto Testé2,3 · Emily Veltjen4,5 · Jhonny Quintana Delgado1 · Alejandro Palmarola Bejerano2 · Marlyn Valdés de la Cruz1 · Pieter Asselman4 · Isabel Larridon4,6 · Marie-Stéphanie Samain4,7 · Luis Roberto González Torres8,9

Received: 16 April 2021 / Accepted: 13 August 2021 / Published online: 20 September 2021 © The Author(s), under exclusive licence to Springer Nature B.V. 2021

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 provide an excellent model for landscape genetic studies. This research focuses on the subspecies and subpopulations of *Magnolia cubensis* and aims to: (1) compare the genetic diversity between the different subspecies and their subpopulations, (2) search for patterns between the spatial and genetic distance, (3) quantify the functional connectivity and (4) assess the structural connectivity of the landscape. This study employs 21 microsatellite markers to quantify the genetic diversity, complemented with seven landscape characteristics of the Guamuhaya and Sierra Maestra massifs in Cuba. *Magnolia cubensis* subsp. *acucae* 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 subpopulations 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 subpopulations, 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. *acucae* 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.

Keywords Gene flow · Genetic patterns · Isolation by distance · Least cost path · Significant evolutionary unit · Spatial autocorrelation
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 combined with spatial data (Scribner et al. 2005; Apodaca et al. 2012; Frantz et al. 2012). The existence of a great variety of statistical methods evaluate the role of 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 provides an excellent model for landscape genetic studies. In this plant family, 48% of species have been assessed as threatened (Rivera et al. 2016). Case studies using molecular data have been conducted on Magnolia species, 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 and 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 genetic diversity of Magnolia species with underlying spatial data (Kikuchi and Isagi 2002; Isagi et al. 2007; Tamaki et al. 2008; Tamaki et al. 2016; Hernández et al. 2020a). Although these studies provided approximations to the description of genetic patterns in geographical space, they lacked an integrative analysis between genetic variation patterns and the landscape or environmental characteristics.

Within the Magnoliaceae family, Magnolia cubensis Urb. provides a perfect case study to execute a full integrative landscape genetic analysis, because its subpopulations have been intensively surveyed for years which enabled documenting its full distribution (Palmarola et al. 2018; Testé et al. 2019; Simón 2020). It is one of five Magnolia species endemic to Cuba (Palmarola et al. 2016; Greuter and Rankin 2017) and belongs to the section Talauira subsection Cubenses (Figlar and Nooteboom 2004). Currently, the species is represented by two subspecies: Magnolia cubensis subsp. cubensis and M. cubensis subsp. acunae Imkhan. (Greuter and Rankin 2017). Although morphologically the two subspecies are very similar (Hernández 2014), it has been suggested to move the subspecies to the species rank based on molecular and geographical arguments (Simón 2020; Veltjens et al. in review). Magnolia cubensis subsp. acunae is a local endemic from the Guamuhaya massif in central Cuba, and 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 (González-Torres et al. 2016; Hernández et al. 2020b, c), causing that both taxa have been assessed as Endangered (Hernández et al. 2020b, c). In response, 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), which could greatly benefit from scientific data to guide future conservation management.

It is not straightforward to make general statements on the effect of habitat fragmentation on the genetic diversity of populations, because each species responds differently on habitat fragmentation, depending on its ecology and evolutionary history (Nazareno and Jump 2012). Both subspecies of Magnolia cubensis present characteristics that, in theory, could either mitigate the loss of genetic diversity in the presence of a fragmented landscape, or that could make them more susceptible to the landscape attributes. 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) could 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 (Honman and Jacquemyn 2007). Small population sizes, habitat specificity (González-Torres et al. 2016), potential self-incompatibility (Yasukawa et al. 1992) and pollination by insects (Thien 1974), could make these trees susceptible to the fragmentation of habitat and landscape patterns (Kolb and Diekmann 2005). Honman and Jacquemyn (2007), Isagi et al. (2007) and Aparicio et al. (2012) found that in species with comparable characteristics to Magnolia cubensis, the genetic diversity and population structure are affected by the loss of habitat and isolation.

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 subpopulations of Magnolia cubensis present differences in their genetic diversity? (2) spatial genetic patterns: Regardless of their spatial structure, what is the number of genetic groups? Are the genetic distances between subpopulations explained by the geographical distances? Is the genotype of an individual in a subpopulation dependent on the genotype of an individual in a nearby subpopulation? (3) functional connectivity: How much genetic differentiation is there between subpopulations 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 subpopulations in the Guamuhaya massif and five subpopulations in the Sierra Maestra massif, Cuba (Fig. 1). These areas contain 90% of the known individuals of *Magnolia cubensis* subsp. *acunae* and *M. cubensis* subsp. *cubensis* and have 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 subpopulation 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 followed 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.

Table 1 Demographic information of *Magnolia cubensis*, collected during expeditions between 2010 and 2020

| Taxa             | Subpopulation | SPop | IND | N  | NF (%) |
|------------------|---------------|------|-----|----|--------|
| *Magnolia cubensis* subsp. *acunae* | NPL ‘Topes de Collantes’ | TOP  | 368 | 39 | 36 (9.8) |
|                  | ER ‘Lomas de Banao’ | BAN  | 70  | 9  | 9 (12.9) |
|                  | Total         |      | 438 | 48 | 45 (10.3) |
| *Magnolia cubensis* subsp. *cubensis* | ER ‘Pico Caracas’ | REC  | 50  | 36 | 36 (72.0) |
|                  | 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) |

The table includes acronyms for each subpopulation (SPop column), the number of known individuals (IND), the number of DNA samples (N) and the final number of genotyped samples used in the three SSR datasets (NF). Between parenthesis the percentage of representation is given (%) i.e. the number of genotyped samples compared to the number of known individuals per subpopulation and taxon

*NPL* Natural protected landscape, *ER* Ecological Reserve, *NP* National Park

Fig. 1 Study areas and sample subpopulations of *Magnolia cubensis* in Guamuhaya and Sierra Maestra massifs, in Cuba. Subpopulation 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
acuana, 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<sub>IS</sub>) for each *locus*, using 10,000 dememorization steps, 50 batches and 5000 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 α = 0.05 for both analyses.

Genetic diversity of each subpopulation was quantified based on the following parameters in GeneAlEx v.6.5 (Peakall and Smouse 2012): number of alleles per *locus* (A), effective alleles (A<sub>e</sub>), number of private alleles (A<sub>p</sub>), observed heterozygosity (H<sub>O</sub>), expected heterozygosity (H<sub>E</sub>), inbreeding coefficient (F<sub>IS</sub>) and percentage of polymorphic *loci* (P). The allelic richness (A<sub>R</sub>) and significance of F<sub>IS</sub> were calculated with FSTAT v.2.9.3.2 (Goudet 1995).

**Spatial genetic patterns**

For each subspecies, the genetic distances between individuals and between subpopulations were estimated from codominant genotypic distances (GD) and linear 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 subpopulations of the same subspecies. The matrices were obtained in GeneAlEx v.6.5 by a modification of the Haversine formula (Simnot 1984). Considering the spatial distribution of the individuals, the matrix of real geographic distances (GGD<sub>R</sub>) was calculated as the hypotenuse of a triangle where the legs were the GGD and ALTD matrices. To calculate the distances between subpopulations of the same subspecies, the coordinates of one individual (located in the center of each subpopulation) were taken as representative of that subpopulation.

**Genetic structure**

We visualized genetic distances between subspecies, subpopulations and individuals with a Principal Coordinates Analysis (PCoA) from the GD matrix in GeneAlEx v.6.5. The PCoA is a multivariate technique that allows to find and plot the major patterns of genetic distances between samples, and the graphic results are intuitive and easily interpreted. The groups defined a priori were the two subspecies, as well as the different subpopulations 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 subpopulation, based on their genotype, we ran two data sets in Structure v.2.3 (Pritchard et al. 2000): grouped by subspecies and grouped by subpopulations. 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. *acuana*; 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.

**Isolation by distance**

To identify possible isolation by distance patterns between and within subpopulations of the each subspecies dataset, we performed Mantel correlation tests with 999 permutations between linear genetic distances (LinGD) and real geographical distances (GGD<sub>R</sub>), whereby α < 0.01 was considered statistically significant.

**Spatial autocorrelation**

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) on each subspecies dataset. Taking into consideration the distance between the individuals, and that the individuals were not randomly distributed within a subpopulation, we constructed distance classes that contain comparable pairs of individuals. We constructed 5 distance classes for *Magnolia cubensis* subsp. *acuana* 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 1000 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 Ay 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 GeneAlEx v.6.5 and the spatial autocorrelation was realized in ALLELES IN SPACE (Miller 2005).

Functional connectivity

We estimated genetic differentiation between subpopulations 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 ($F_{ST}$) followed the criteria of Hartl and Clark (1997): $< 0.05$ (Little), $0.05 – 0.15$ (Moderate), $0.15 – 0.25$ (Great) and $> 0.25$ (Very great).

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 subsection 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 (~ 1 km²), so it was necessary to trim all the variables to a standardized pixel size of 0.7883 km². The processing was carried out in QGIS v.3.4.12 and Global Mapper v.15.0 (www.globalmapper.com).

Structural connectivity: 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 linear genetic distances (LinGD), real geographical distances (GGDg) and environmental distances (EnvD, BIO4D, BIO12D and BIO15D) were investigated for their correlation. To create 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 PopTools (Hood 2010). Three types of analyses were carried out: (a) partial correlation between genetic (LinGD) and environmental (EnvD, BIO4D, BIO12D and BIO15D) distances, in which the effect of geographical (GGDg) distances was controlled, (b) partial correlation between geographical (GGDg) and environmental (EnvD, BIO4D, BIO12D and BIO15D) distances, in which the effect of genetic (LinGD) distances was controlled, and (c) partial correlation between genetic (LinGD) and geographic (GGDg) distances, in which the effect of environmental (EnvD, BIO4D, BIO12D and BIO15D) distances was controlled. The tests were run in R using the vegan package (Oksanen et al. 2019), with 1000 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 calculated (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 variables that explicitly consider the influence of landscape hétéroge-

 Springer
neity. For the first friction map the influence of the vegetation was tested. For this, the vegetation map was reclassified to three classes, with low cost values (1) for those plant formations where the species is distributed, average cost (10) was assigned to other natural plant formations and the highest cost (100) was assigned to costal vegetation. For the other six friction map the influence of the environmental and geographic variables was tested. 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 PopTools (Hood 2010). 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. An average cost (10) was set 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). A multivariate resistance surface was created by adding each reclassified layer (raster map with a cell size of 90 m²), representing seven different variables that could influence resistance to gene flow. The least cost routes between subpopulations were calculated under the assumption that the rate of gene flow between any two subpopulations 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 showed different results for the two subspecies. For M. cubensis subsp. acucae, no statistically significant differences were obtained for any of the markers. In M. cubensis subsp. cubensis 14 loci presented FIS values significantly different from zero, in at least one subpopulation. In Pico Turquino and Gran Piedra, more than 20% of the markers showed significant deviations, with 9 and 6 out of the 21 loci, respectively. Linkage analysis showed for M. cubensis subsp. acucae 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 subpopulation of Magnolia cubensis can be found in Table 2. Magnolia cubensis subsp. cubensis showed higher values for the genetic diversity parameters compared to M. cubensis subsp. acucae. At the infraspecific level for M. cubensis subsp. acucae, Topes de Collantes showed the highest

Table 2  Measures of genetic diversity per subpopulation and taxon of Magnolia cubensis, estimated by SSR markers

| Taxa                  | SPop | Np  | A (min–max) | Ae  | Ar  | Ho  | Hz  | HW  | FIS  | P (%) |
|-----------------------|------|-----|-------------|-----|-----|-----|-----|-----|------|------|
| Magnolia cubensis     |      |     |             |     |     |     |     |     |      |      |
| subsp. acucae (19)    | TOP  | 36  | 6.16 (1–11) | 3.55| 4.54| 0.60| 0.64| 0   | 0.069| 94.74|
|                       | BAN  | 9   | 3.74 (1–6)  | 2.25| 3.74| 0.56| 0.53| 0   | 0.002| 94.74|
|                       | Total| 45  | 4.72 (1–11)| 2.80| 4.46| 0.56| 0.58| –   | 0.031| 94.74|
| Magnolia cubensis     |      |     |             |     |     |     |     |     |      |      |
| subsp. cubensis (21)  | REC  | 36  | 5.57 (1–17)| 3.08| 4.32| 0.43| 0.51| 4   | 0.182| 85.71|
|                       | TUR  | 39  | 6.52 (2–17)| 3.43| 5.18| 0.37| 0.57| 9   | 0.364| 100.00|
|                       | 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|

The number of SSR markers per taxon is given between parentheses

SPop subpopulation acronyms according to Table 1. Np, number of sampled individuals. A, average number of alleles per locus. Ae, average number of effective alleles. Ar, average allelic richness. Ho, average observed heterozygosity. Hz, average expected heterozygosity. HW, number of loci that deviated from Hardy–Weinberg proportions. FIS inbreeding coefficient calculated according to Weir and Cockerham (1984). P, percentage of polymorphic loci. Detailed information on the genetic diversity estimates per locus in Magnolia cubensis subsp. cubensis is given in Quintana et al. (2021)
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 subpopulations 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 $H_e$ (0.37). In all the subpopulations, a deficit of heterozygotes is observed, since the mean values of $H_e$ 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. (2021).

### 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 two separate clusters 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 subpopulations (Fig. 2b). In contrast, the five subpopulations in *M. cubensis* subsp. *cubensis* were somewhat detectable in the PCoA as separate clusters. The clusters containing the individuals of Pico Caracas, Pico Turquino, La Bayamesa and El Gigante overlap; while that of Gran Piedra does not overlap with the rest of the subpopulations (Fig. 2c). Although the PCoA allowed to define groups in two of the three datasets, the two axes explain little variation in the data.

Barplots with the optimal K genetic groups following the $\Delta K$ criterium 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

---

**Fig. 2** Principal coordinates analysis based on SSR genetic distance matrices of *Magnolia cubensis* individuals. Subpopulation acronyms are according to Table 1. a Analysis including both subspecies (19 SSR markers). b Analysis of *Magnolia cubensis* subsp. *acunae* (19 SSR markers). c Analysis of *Magnolia cubensis* subsp. *cubensis* (21 SSR markers)
other three to *M. cubensis* subsp. *cubensis*. When analyzing the subspecies separately, the *M. cubensis* subsp. *acunae* dataset has a Δ*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* Δ*K* = 3, with a secondary peak at Δ*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 subpopulations 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 subpopulation.

### Isolation by distance

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 significant

### Table 1

| Analysis | Subspecies |
|----------|------------|
| a | Analysis including both subspecies. |
| b | Analysis of *Magnolia cubensis* subsp. *acunae*. |
| c | Analysis of *Magnolia cubensis* subsp. *cubensis* (modified from Quintana et al. 2021) |

### Fig. 3

Structure barplots that illustrate the probability of genetic group allocation of *Magnolia cubensis* individuals, based on allelic frequencies of SSR data. Subpopulation acronyms are according to Table 1. a Analysis including both subspecies. b Analysis of *Magnolia cubensis* subsp. *acunae*. c Analysis of *Magnolia cubensis* subsp. *cubensis* (modified from Quintana et al. 2021)

### Fig. 4

Spatial autocorrelation analyses for *Magnolia cubensis*. Dashed lines indicate the average pairwise genetic distance from the full data set. Spatial autocorrelation is visualized by the red datapoints that have a probability of observing a random value of V ≥ observed V by chance. Mantel test correlation coefficient (r²) and their significance (p) for the correlation between linear genetic distances and geographic distances (km). *p < 0.01 was considered statistically significant.

- a *Magnolia cubensis* subsp. *acunae*.
- b *Magnolia cubensis* subsp. *cubensis*.
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).

**Spatial autocorrelation**

These results were supported by the spatial autocorrelation analyses (significant spatial autocorrelation is indicated in red in Fig. 4). For *Magnolia cubensis* subsp. *acuace* 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* (between 35 and 40 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 subpopulations of *M. cubensis* subsp. *acuace*. 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.

### Table 3  Genetic differentiation between seven subpopulations of two taxa of *Magnolia cubensis*

| Taxa                  | Subpopulation | BAN | TUR | GRP |
|-----------------------|---------------|-----|-----|-----|
| *Magnolia cubensis*   | TOP           | 0.035 | 0.115 | 0.306 | 0.414 |
| *Magnolia cubensis*   | BAN           |       | 0.065 | 0.089 | 0.326 |
|                       | REC           |       |       | 0.089 |       |
|                       | TUR           |       |       |       | 0.379 |
|                       | BAY           |       |       |       | 0.326 |
|                       | GIG           |       |       |       |       |
|                       | GRP           |       |       |       |       |

The number of SSR markers per taxon is given between parentheses.

Subpopulation 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.

**Structural connectivity: Landscape genetics analyses**

Monmonier’s Maximum Difference algorithm showed that the highest rates of genetic change were located mainly between subpopulations and not within them (Fig. 5). Of a maximum of 10 barriers (previously assigned to the algorithm), three were assigned for *Magnolia cubensis* subsp. *acuace* and seven 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 subpopulations, even for those more isolated ones. 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. *acuace* its limits do not cover the entire distance between subpopulations. 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 shown in Table 5 and Fig. 6, respectively. Although in general the geographical distances, the cost distances and the fixation index \((F_{ST})\) between subpopulations 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 (\(M.\) cubensis subsp. acunae) the cost value is approximately 7000; 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. 3000), 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 subpopulations 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**

**Genetic diversity**

The higher genetic diversity of \(M.\) 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 Berazaín 1984; Simón 2020), the habitat being conserved better (Palmarola et al. 2012; González-Torres et al. 2016), or a combination of any of the three. Gitzendanner and Soltis (2000) stated 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).
Lomas de Banao: AR = 2.93, Ho = 0.36, He = 0.41), most likely in than the diversity values reported by Veltjen et al. (2019) in the present study. The diversity measures were lower because of the higher number of microsatellite markers used.

When we compare the genetic diversity of the subpopulations, cost values derived from the Least-cost-path models and FST values for all pairwise comparisons between subpopulations of Magnolia cubensis.

| Species                  | Pair of subpopulations | Distance (km) | Cost   | FST   |
|--------------------------|------------------------|---------------|--------|-------|
| Magnolia cubensis subsp. acucae | TOP-BAN                | 43.29          | 7517.88 | 0.04  |
| Magnolia cubensis subsp. acucae | TUR-BAY                | 14.67          | 499.66  | 0.09  |
| Magnolia cubensis subsp. acucae | REC-TUR                | 18.93          | 1418.73 | 0.12  |
| Magnolia cubensis subsp. acucae | BAY-GIG                | 26.91          | 2283.51 | 0.13  |
| Magnolia cubensis subsp. acucae | REC-BAY                | 33.54          | 1888.40 | 0.17  |
| Magnolia cubensis subsp. acucae | TUR-GIG                | 41.32          | 2783.17 | 0.16  |
| Magnolia cubensis subsp. acucae | REC-GIG                | 60.24          | 4171.91 | 0.23  |
| Magnolia cubensis subsp. acucae | GIG-GRP                | 88.49          | 27980.31| 0.21  |
| Magnolia cubensis subsp. acucae | BAY-GRP                | 112.20         | 29969.24| 0.25  |
| Magnolia cubensis subsp. acucae | TUR-GRP                | 126.60         | 30468.90| 0.27  |
| Magnolia cubensis subsp. acucae | REC-GRP                | 144.83         | 31857.64| 0.29  |

Spearman’s correlation coefficients (r_s) and associated probability (p) for each of the three variables. The pairs of subpopulations are sorted by geographic proximity (low–high). Subpopulation acronyms according to Table 1:

- Correlation geographic distance—cost: r_s = 0.94; p < 0.0001*
- Correlation cost—FST: r_s = 0.74; p < 0.0001*
- Correlation geographic distance—FST: r_s = 0.81; p < 0.0001*

The number of SSR markers per taxon is given between parentheses. The partial Mantel tests are defined in the format (A × B) C whereby A and B are the variables that are being tested for partial correlation, while the variable C is controlled.

**Table 5** Distance between subpopulations, cost values derived from the Least-cost-path models and FST values for all pairwise comparisons between subpopulations of Magnolia cubensis.

The mean values of allelic richness, observed and expected heterozygosity (Table 2) found for *M. cubensis* subsp. acucae exceeded that reported by Hernández et al. (2020a) (Tobres de Collantes: AR = 7.93 H_o = 0.60 H_e = 0.22; Lomas de Banao: AR = 7.93 H_o = 0.60 H_e = 0.41), and the number of samples 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. From the latter study and confirmed by our result, it is clear that 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. shihuisensis* (Chun & Y.F.Wu) Figlar (Deng et al. 2020) and five Magnolia species of Veracruz, Mexico (Aldabá-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 and 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 subpopulations 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. acucae, the Tobres de Collantes subpopulation was the most genetically 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 genetically diverse subpopulation was the National Park La Bayamesa and the least genetically 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 subpopulation. 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) stated 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. *acuanae*, and by
Molina-Pelegreń 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 subpopulations (i.e. Lomas de Banao, El Gigante, Pico Caracas and Gran Piedra) hence may partly be explained by their demographic history.

**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 Nooteboom 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. *acuæae* 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.

**Functional connectivity**

The fixation index (FST) is a measure of the genetic divergence between subpopulations. Values close to zero, such as those found in *Magnolia cubensis* subsp. *acuæae* (Table 3), indicate that there is similar heterozygosis in each subpopulation compared to the value of the two subpopulations together, so genetic differentiation is low (Hartl and Clark 1997). High values of this coefficient (FST > 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 subpopulations 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 La Bayamesa – El Gigante (Table 3), the genetic differentiation with Pico Turquino, similar to (or lower than) other pairs of subpopulations 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 barcharts for the Gran Piedra individuals (Fig. 3c), as well as their differentiation with the rest of the subpopulations (Table 3), can be the cause of the large number of private allelic variants at this subpopulation (Table 2), resulting from the geographical distance and isolation (Fig. 1b) or the diversification history of the Cuban magnolias. The geographical distance between subpopulations 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 subpopulations, 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 subpopulation to which *Magnolia* members of the subsection *Cubenses* migrated after which the rest of the Cuban subpopulations 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. *Magnolia cubensis* 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: *Magnolia cubensis* 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* with *M. cubensis* subsp. *cubensis* based on the similarity between taxa in the leaf shape and leaf areole size (Palmarola et al. 2015a, b).

Considering the different data (Figs. 2 and 3, Table 2 and 3), and according to the criteria of Ryder (1986) and Moritz (2002), we can define one evolutionary significant unit in *Magnolia cubensis* subsp. *acuæae*, 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.

**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 subpopulations, due to the existence of the 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 the 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 subpopulations. 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 subpopulations for each subspecies, not necessarily indicate that they have different dispersal ability or seed dispersers. Their seeds are very similar (Imchaitzka 1991) and 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 massif, it is necessary to clarify that the models were based on the habitat requirements of *Magnolia cubensis*. 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 (i.e. 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 *Magnolia cubensis*, 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 and conservation management guidelines**

In conclusion, the current landscape features do not clearly correlate with the patterns of genetic diversity; and the results are taxon-specific. For *Magnolia cubensis* subsp. *acuinae*, 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 subpopulations, 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 subpopulations does not reflect the characteristics of the landscape that separates them or the cost values to cross it. For Magnolia cubensis 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 from the other subpopulations was moderate and migrants were detected. Large genetic differentiation was only found between Gran Piedra and the other subpopulations, 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 confirm to consider Magnolia cubensis subsp. acucae and M. cubensis subsp. cubensis as different species. Furthermore, our data suggest defining M. cubensis subsp. acucae as one evolutionarily significant unit (ESU) and subsequently one conservation unit (CU). We recognize two ESUs and CUs in M. cubensis subsp. cubensis: Pico Caracas – Turquino – La Bayamesa – El Gigante and Gran Piedra. We propose a Least-cost-path to connect the subpopulations of M. cubensis subsp. acucae and the four subpopulations 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. In Cuba there are no examples of conservation actions that use the combination of genetic and landscape data: conservation genetics still requires integration into conservation practice, at the level of legislation and management.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10592-021-01395-6.

Acknowledgements This work was financially supported the Whitley Fund for Nature, Franklinia Foundation, 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 to all the people who have helped us during field work: Norlys Albelo, José L. Moscoso, Oliver Valle, Eldis Béquer, Alexei Rodríguez, Yenia Molina, Adonis Sosa, Daryl D. Cruz y Reinier Núñez.

Authors’ contributions Study design MHR, ETL, LRG-T and EV. Field work and data collection MHR, ET, APB and LRG-T. Lab work: MHR, EV, ET, PA and APB. Data analysis: MHR, JQD and EV. Writing and original draft preparation: MHR. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the Whitley Fund for Nature, Franklinia Foundation (Project Number 2019–03) and Plantlife Conservation Society.

Data Availability 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.

Declarations

Conflict of interest 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.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

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

References

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

Fernández A, Pérez de los Reyes R (eds) (2008) GeoCuba: evaluación del Medio Ambiente cubano. Ministerio de Ciencia, Tecnología y Medio Ambiente

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

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

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

Anderson CD et al (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. Mol Ecol 19:3565–3575

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

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
Waples RS (2015) Testing for Hardy-Weinberg proportions: have we lost the plot? J Hered 106(1):1–19
Watson DF (1992) Contouring: a guide to the analysis and display of spatial data. Pergamon Press, New York
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evo 38:1358–1370
Wright S (1943) Isolation by distance. Genetics 28:114–138
Xin Z, Chen J (2012) A high throughput DNA extraction method with high yield and quality. Plant Methods 8(26):1–7
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

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.