Conservation Genetics (2025) 26:913–927
https://doi.org/10.1007/s10592-025-01712-3

RESEARCH

Conservation genetics of the wild cotton *Gossypium turneri*: a species threatened by tourism-driven coastal development in Sonora, Mexico

Karla Fabiola Yescas-Romo[1,2] · Fernando Razo-Mendivil[3] · Corina Hayano-Kanashiro[2] · Francisco Molina-Freaner[4]

Received: 7 August 2024 / Accepted: 11 June 2025 / Published online: 2 July 2025
© The Author(s) 2025

Abstract

Crop wild relatives (CWR) are an important class of endangered wild species that are increasingly declining due to human-induced factors that also threaten other wild plants. This class of species has not received sufficient attention in conservation strategies and remains poorly studied. This holds true for the endangered wild cotton *Gossypium turneri*, which is endemic to the state of Sonora, Mexico. This species faces threats from disturbance associated with tourism-driven coastal development and requires urgent conservation actions. This study aimed to assess the genetic diversity and population structure of *G. turneri* using genotyping by sequencing (GBS) to offer insights into its conservation. Clustering and genetic structure analysis revealed three distinct genetic groups and potential populations. Observed (H_O) and expected (H_E) heterozygosity, along with genetic differentiation (F_{ST}) values, indicated moderate genetic diversity and differentiation. Effective population sizes (N_e) varied from 18 to 32, whereas census population sizes (N_c) ranged from 62 to 883 adult plants. Among the three inferred populations, the southernmost population exhibited the highest number of private alleles and allelic differences, making it a conservation priority due to its potential reservoir of genetic variation. Recommendations include seed collection for ex-situ conservation in Botanical Gardens for the three populations. For in-situ conservation, different measures should be implemented according to the level of human disturbance affecting each population. These include restricting further coastal development in less disturbed areas and promoting *G. turneri* as an ornamental plant in the most impacted population to enhance its survival.

Keywords GBS · SNPs · Genetic structure · Genetic diversity · Effective and census population size

Introduction

Wild plant populations face increasing threats from diverse human activities, leading to a significant number of endangered plants listed on the current International Union for Conservation of Nature (IUCN) Red List (IUCN 2023). Human-induced causes of endangerment include destruction, degradation or fragmentation of habitats, overexploitation, overgrazing by introduced ungulates, pollution, and displacement by invasive non-native species (DiPaolo and Villella 2022). While plants may be rare due to natural causes, the convergence of multiple factors, both natural and anthropogenic, often exacerbates the threat, consequently resulting in a severe decline of species (DiPaolo and Villella 2022). Among the human-induced causes of endangerment, tourism-related activities in coastal ecosystems have been listed as a significant threatening agent (Rankin et al. 2015; Simone et al. 2018).

✉ Corina Hayano-Kanashiro
angela.hayano@unison.mx

✉ Francisco Molina-Freaner
freaner@unam.mx

[1] Posgrado en Biociencias, Universidad de Sonora, Hermosillo, Sonora, México

[2] Departamento de Investigaciones Científicas y Tecnológicas (DICTUS), Universidad de Sonora, Hermosillo, Sonora, México

[3] Departamento de Ingeniería Genética, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV), Irapuato, Guanajuato, México

[4] Departamento de Ecología de la Biodiversidad, Instituto de Ecología, Universidad Nacional Autónoma de México, Hermosillo, Sonora, México

Although tourism can be utilized as a tool for conserving biodiversity (Buckley and Pascual 2024), often tourism-driven coastal development exerts significant pressure on plant species from coastal ecosystems (Simone et al. 2018; Samad et al. 2021). The construction of hotels, resorts, marinas, and other infrastructure can result in habitat loss and fragmentation, impacting vegetation that provides crucial ecosystem services such as mangroves (Brenner et al. 2018) as well as affecting sensitive endemic species (Samad et al. 2021). Uncontrolled tourism development and urban sprawl contribute to environmental degradation (Navarro 2019), leading to the disappearance of populations of endangered endemic species (Médail et al. 2021) and facilitating the introduction of invasive non-native species (Anderson et al. 2015). Consequently, tourism-related activities increase the risk of extinction of sensitive endemic plants from coastal habitats (Sulis et al. 2020; Médail et al. 2021).

Genetic diversity is crucial for species persistence as it helps to maintain the capacity to respond to diseases, pests, climate change, and other stressors (Chung et al. 2023). Species with higher genetic diversity are more likely to adapt and survive in the face of environmental changes than species with lower genetic diversity (Chung et al. 2023). Consequently, species with low genetic diversity have a reduced adaptive capacity and become more vulnerable to extinction under intense environmental changes (Azman et al. 2020; Li et al. 2021). Therefore, assessing genetic diversity is a critical objective for the development of conservation strategies and management plans, particularly for endangered species (Dorsey et al. 2023; Li et al. 2021), where the identification of conservation units or populations with high genetic diversity is of paramount importance for their preservation and management (Wu et al. 2023).

Crop wild relatives (CWR) are an important class of endangered wild species, with relevance to food security, that are increasingly declining due to human-induced factors that also threaten other wild plants (von Wettberg et al. 2022). For example, according to the IUCN criteria, 35% of the 224 evaluated CWR from Mesoamerica are currently threatened with extinction (Goettsch et al. 2021). Tourism development emerges as the primary threat driver for 46% of CWR, particularly in the genus *Gossypium* (Goettsch et al. 2021). Despite their importance, CWR are not usually a priority in conservation strategies (Petitpierre et al. 2023) and have been poorly studied, as is the case of *Gossypium turneri*, a Critically Endangered taxon that requires urgent conservation actions in Mexico (Goettsch et al. 2021).

G. turneri is a wild cotton species endemic to the state of Sonora, exclusively found within a limited area along the coastal region north of San Carlos in the municipality of Guaymas in northwestern Mexico (Wegier et al. 2018; Yescas-Romo et al. 2024). San Carlos is a major tourist destination in Sonora (Enríquez-Acosta and Avila-Martínez 2017), known for its natural attractions (Cisneros-Montemayor et al. 2020; Rivera-Martínez 2020; Quijada-Siu 2015). Established in 1964, San Carlos comprises several localities, including San Carlos Nuevo Guaymas, El Renacimiento, Marina Real, Villas de California, and La Manga, with a combined population of approximately 2,508 Mexican and foreign residents, primarily from the US and Canada, according to the 2020 census (Enríquez-Acosta and Avila-Martínez 2017; Chávez-Valdez and Enríquez-Acosta 2022). The area offers a tourist capacity of 1,950 rooms and annually attracts around 900,000 national and 300,000 foreign tourists (Chávez-Valdez and Enríquez-Acosta 2022). Most of the urban area comprises second residences of Mexican nationals (for weekends and vacations) and North American retirees (Enríquez-Acosta and Avila-Martínez 2017). The tourist infrastructure, including hotels, condominiums, marinas, and restaurants, is spread along 13 km of the coast (K. Yescas-Romo, personal observation). While coastal development has not been explosive, tourism infrastructure has encroached upon coastal habitats (Enríquez-Acosta and Avila-Martínez 2017). Consequently, *G. turneri* has been adversely affected by tourism-driven coastal development, leading to habitat degradation, fragmentation, and a decline in its distribution area (Wegier et al. 2018; Yescas-Romo et al. 2024), thereby increasing its risk of extinction. Given the increasing disturbances associated with coastal development in its distribution area and its potential as a genetic resource, urgent conservation actions are needed for the preservation of *G. turneri*. This study analyzes the genetic diversity and population structure of the wild cotton *G. turneri* using genotyping by sequencing (GBS) to provide valuable guidelines for its conservation. The objectives include: (a) determining the number of populations based on cluster and genetic structure analysis, (b) analyzing genetic diversity in the inferred populations, (c) identifying populations with greater genetic diversity and potential conservation units, and (d) proposing conservation strategies for the species.

Materials and methods

Study species and sample collection

G. turneri is a diploid species with 26 chromosomes (2n = 2x = 26; Yescas-Romo et al. 2021). This wild cotton is a shrub about one meter in height with hermaphrodite flowers (Fryxell 1978, 1988; Ulloa 2014) (Fig. 1). It is a self-compatible species capable of autonomous self-pollination through style curvature under pollinator scarcity. However, most flowers show approach herkogamy, suggesting that



Fig. 1 Natural habitat of *G. turneri*: (**a**) Individual of *G. turneri* (indicated with white arrow) in the coastal area of San Carlos, Guaymas, (**b**) flower of *G. turneri*, (**c**) mature fruit of *G. turneri* and (**d**) leaf of *G. turneri*

the species is predominantly outcrossed (Yescas-Romo et al. 2024). Flowers last one day and are visited by a diverse group of insects, including butterflies, moths, bees, and ants (Yescas-Romo et al. 2024). Seeds are on average 8.80 mm by 4.83 mm (length x width; Yescas-Romo et al. 2021), and although seed dispersal has not been properly studied, our observations suggest that seeds fall into the ground (gravity) and once in the soil, water during rainfall events may transport the seeds downstream (K. Yescas-Romo, personal observation).

G. turneri is currently classified as endangered under the Mexican legislation (DOF 2019) and critically endangered in the IUCN Red List (Wegier et al. 2018). As previously stated, its natural distribution is limited to the coastal region of the municipality of Guaymas, Sonora. Within its range, the species has a discontinuous distribution forming spatially discrete areas north of San Carlos, where samples were collected in three accessible sites throughout its distribution: El Mirador, La Manga, and Roca El Choyudo (Fig. 2). Due to its critical conservation status and to deter illegal collection, specific coordinates of the collection sites are withheld. In each site, very young new shoots were collected from 30 individuals, resulting in a total of 90 individuals in 2022. To minimize the likelihood of sampling closely related individuals, plants were chosen haphazardly, ensuring a minimum distance of 15 m between them while aiming to cover a significant portion of the population's range. Each plant was labeled and georeferenced using the eTrex® 32x GPS device (GARMIN). The gathered young shoots were immediately placed in a cold icebox and transported for storage at −80 °C until genomic DNA extraction.

Genomic DNA extraction

Genomic DNA extraction from young shoots was performed using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany), with modifications described in the supplementary information (Protocol S1). The quality and quantity of genomic DNA were assessed using the NanoDropTM Lite spectrophotometer (Thermo Scientific™), while the integrity of genomic DNA was visualized through gel electrophoresis on a 1% agarose gel.

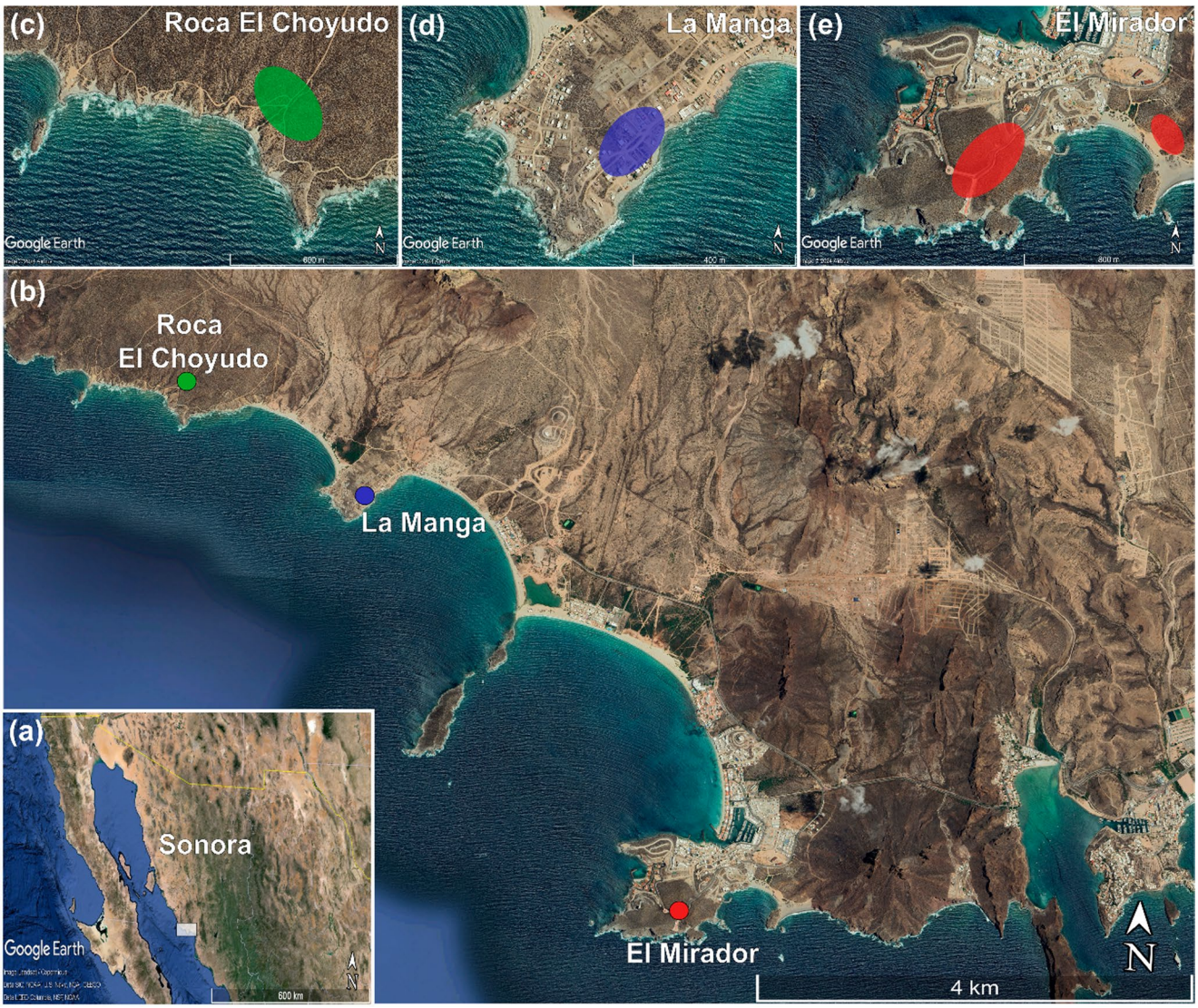


Fig. 2 Collection sites of *G. turneri* in Sonora, Mexico: (**a**) Study area (white box) in the state of Sonora, (**b**) study area indicating the three collection sites (El Mirador, La Manga and Roca El Choyudo), **c**) sampling areas at Roca El Choyudo (in green), **d**) La Manga (in blue) and **e**) El Mirador (in red)

GBS library preparation and sequencing

The genomic DNA of the samples was dispatched to the University of Wisconsin-Biotechnology Center DNA Sequencing Facility (https://biotech.wisc.edu/) for GBS library preparation and sequencing. During the GBS library construction process, genomic DNA underwent digestion using the methylation-sensitive ApeK1 restriction enzyme (Elshire et al. 2011). Subsequently, a unique barcode with an adapter sequence was ligated to each genomic DNA sample before pooling. Sequencing of the GBS libraries was conducted on an Illumina NovaSeq600 platform (Illumina, Inc.), generating 2 × 150-bp paired-end reads.

Sequence analysis and genetic variant calling

Raw sequencing reads were de-multiplexed based on individual and multiplexing indexes using Cutadapt v. 3.5 (Martin 2011) to split sequencing reads per sample. De-multiplexed raw and trimmed reads underwent quality control analysis using FastQC v. 0.11.9. (Andrews 2019). Subsequently, the raw reads were processed using Trimmomatic v. 0.39 (Bolger et al. 2014) to remove adapter sequences and reads with a final length less than 50 bp. The trimmed reads were aligned to the *G. turneri* reference genome (Udall et al. 2019; GenBank accession: GCA_008044935.1) using BWA-mem2 v. 2.2.0 (Md et al. 2019) with default parameters. The resulting alignment files were converted to binary

alignment map (BAM) format, sorted, and indexed using SAMtools v. 1.17 (Danecek et al. 2021).

Genetic variant calling was performed using Platypus v. 0.8.1.1 (Rimmer et al. 2014). Identified genetic variants were filtered using VCFtools v. 0.1.15 (Danecek et al. 2011) to retain biallelic sites with a minimum quality value of 30, a minor allele frequency of 0.01, and a limit of 5% missing data. Single nucleotide polymorphisms (SNPs) were retained, while insertions and deletions (InDels) were removed. Additionally, loci with missing data exceeding 10% in each population were excluded from further analysis.

Genetic structure and differentiation

SNPs identified in the genetic variant calling were analyzed to assess the genetic structure and differentiation of the populations. To estimate the potential number of populations, individual ancestry proportions were calculated using ADMIXTURE v. 1.3.0 (Alexander and Lange 2011) by varying the number of clusters (K-values) from 1 to 5, with predetermined parameters. The optimal K-value was determined through cross-validation (CV) error testing (Alexander and Lange 2011). Subsequently, individual ancestry values were visualized using the R package ggplot2 v. 3.4.4 (Wickham 2016) in R Core Team v. 4.3.3 (R Core Team 2024). In addition, a principal component analysis (PCA) was conducted using the R package ade4 v. 1.17–22 (Dray and Dufour 2007) to delineate the genetic relationships between individuals, genetic clustering, and differentiation among individuals and potential populations.

After determining the number of populations through individual ancestry proportion and PCA, an analysis of molecular variance (AMOVA) was performed and global F_{ST} and pairwise F_{ST} values were calculated to assess the level of genetic differentiation within and among populations. R Core Team v. 4.3.3 (R Core Team 2024) was utilized for these analyses. For the AMOVA, the R package poppr v. 2.9.5 (Kamvar et al. 2014) was used, and the significance of variance components was assessed using the randtest function from the ade4 v. 1.17–22 package (Dray and Dufour 2007). Global F_{ST} and pairwise F_{ST} values were calculated according to Weir and Cockerham (1984) using R package diveRsity v. 1.9.90 (Keenan et al. 2013) with 1,000 bootstraps for the calculation of 95% confidence intervals (CI). This comprehensive analysis provided insights into the genetic differentiation levels within and among populations, contributing to a deeper understanding of the population structure of *G. turneri*.

To assess whether genetic differentiation among populations could be attributed to isolation by distance (IBD), a Mantel test was conducted using the ade4 v. 1.17–22 package (Dray and Dufour 2007) in R Core Team v. 4.3.3 (R Core Team 2024). The test utilized 9,999 permutations to determine significance. Genetic distances between individuals were computed using Nei's method (Nei 1972) within the R package StAMPP v. 1.6.3 (Pembleton et al. 2013). Geographical distances were calculated based on the geographic coordinates of the individuals using GenAlEx v. 6.5 (Peakall and Smouse 2012).

Genetic diversity

To assess genetic diversity, SNPs were analyzed, and basic genetic diversity statistics were computed in R Core Team v. 4.3.3 (R Core Team 2024), both globally (across all samples) and within differentiated populations identified by different analyses of genetic structure to identify potential conservation units. The number of effective alleles per locus (N_{EA}) and allelic richness (A_R) were calculated using R packages dartRverse v. 1.0.2 (Gruber et al. 2018) and hierfstat v. 0.5–11 (Goudet 2005), respectively. Genetic diversity statistics, such as observed heterozygosity (H_O) and expected heterozygosity (H_E), were estimated using dartR v. 2.9.8 package (Gruber et al. 2018). To test for differences in genetic diversity statistics, t-test after testing the Bartlett's test of homogeneity of variances (Bartlett 1937) was applied to H_O and H_E using R package stats v. 4.3.3 (R Core Team 2024). The number of loci deviating from the Hardy-Weinberg equilibrium was determined using Guo and Thompson's test (1992) with 1,000 permutations with the packages pegas v. 1.3 (Paradis 2010). Additionally, nucleotide diversity (π) was estimated using TASSEL v. 5.0 (Bradbury et al. 2007).

Furthermore, for every differentiated population, the number of private alleles (P_A) was determined using the poppr v. 2.9.5 package (Kamvar et al. 2014). Additionally, the inbreeding coefficient (F_{IS}) in every differentiated population was estimated using the R package dartR v. 2.9.8 (Gruber et al. 2018) with 10,000 bootstraps for the calculation of 95% CI. Global F_{IS} was calculated according to Weir and Cockerham (1984) using R package diveRsity v. 1.9.90 (Keenan et al. 2013) with 1,000 bootstraps for the calculation of 95% CI. This comprehensive analysis provides a detailed understanding of the genetic diversity of *G. turneri*, which is crucial for conservation efforts.

Census and effective population size

The three studied populations are distributed along a coastal strip of 400–500 m parallel to the shoreline. In 2022, the distribution of each population was surveyed to determine the extent of its range. Once the boundaries were identified, a team of two people conducted a systematic census in the three studied populations. Due to differences in human disturbance, census methods varied slightly among

populations. In the least disturbed population (Roca El Choyudo), the systematic census was conducted by walking along a transect parallel to the shore. Along this transect, adult individuals (plants > 0.5 m in height) were counted on both sides of the line until reaching the population's limits, where no further plants were detected. The team of two people then moved inland to a new parallel transect and continued counting reproductive plants using tally counters. This process was repeated until the inland boundary of the population was reached. In the most human-impacted population (La Manga), the census was conducted among vacant lots, home gardens and along the streets of the village, carefully counting all adult plants. For El Mirador, the census followed the same method as Roca El Choyudo but was restricted to wild areas unaffected by human infrastructure. The total number of adult plants recorded at each site was used to estimate the census population size (N_c).

The contemporary effective population size (N_e) of each population was estimated using NeEstimator v.2.0.1 (Do et al. 2014). The analysis employed the bias-corrected method based on linkage disequilibrium (Waples 2006), assuming random mating, with a minimum allele frequency cutoff of 0.05. A 95% CI was obtained through jackknifing over individuals. The N_e/N_a ratio was then calculated for each population.

Divergence history

Based on the genetic structure analysis and the geographical distribution of the sampled populations, the divergence pattern among populations was estimated using the Approximate Bayesian Computation (ABC) approach in DIYABC Random Forest v.1.0 (Collin et al. 2021). The analysis evaluated six divergence scenarios (Fig. S1). Admixture events were not considered, because admixture analysis provided little support for their occurrence (see Results section). However, bottleneck events were incorporated into the analysis. For each scenario, 5,000 simulated training datasets were generated using 20,000 SNPs, with prior parameter distributions assumed to be uniform. Before bottleneck events, priors for both effective population sizes and divergence times (in number of generations) were set to default values, ranging from 10 to 10,000. For bottleneck events, priors for effective population sizes were specified to range from 10 to 50 individuals, with event durations varying between 10 and 100 generations. The most probable divergence scenario and its posterior probability were estimated using 2,000 random forest trees. The scenario receiving the highest number of votes in the analysis was considered the most probable.

Minimum sample size for conservation ex-situ

The conservation of this wild cotton species is critical due to the significant threats it faces. To develop effective ex-situ conservation strategies, it is essential to preserve its genetic diversity. To determine the minimum number of individuals required to adequately represent the genetic diversity of *G. turneri*, we applied the methodology proposed by Hoban et al. (2020). Genetic diversity data obtained through GBS from the three natural populations were resampled without replacement to generate pseudo-observed datasets, each representing different sample sizes for potential ex-situ conservation programs. Resampling was conducted in increments of five individuals within each of the previously identified populations. Genetic diversity in each dataset was assessed by calculating expected heterozygosity (H_E), which was then compared to the H_E values of the respective populations. This analysis was performed using the R package adegenet v. 2.1.10 (Jombart 2008). To evaluate the representativeness of the genetic diversity captured, each resampling and H_E estimation was assessed using 100 bootstrap replicates (Monteiro et al. 2022).

Results

Sequencing and SNPs calling

For the analysis of genetic diversity and population structure to identify potential conservation units, 90 individuals of *G. turneri* were sequenced using the GBS approach. A total of 1,518,231,980 paired-end raw reads were generated with an average of 16,583,296 reads per sample (ranged from 12,474,854 to 124,895,83) and an average quality score (Phred quality score) greater than 30. Following the removal of adapter sequences and read filtering, most samples yielded between 10,483,617 and 16,043,059 high-quality reads. However, one sample produced an exceptionally high number of reads, totaling 102,679,597, and was consequently excluded from further analysis. Therefore, a total of 89 samples of *G. turneri* were included in subsequent analyses. From this dataset, 820,901 genetic variants were called. After filtering, the final dataset comprised 89 individuals of *G. turneri* and 46,770 high-quality SNPs with 0.47% of missing data.

Genetic structure and differentiation

The analysis of individual ancestry proportion in ADMIXTURE v. 1.3.0 indicated that the best model fit for the samples was K = 3 with a CV score of 0.425 (Fig. S2). This model grouped individuals of *G. turneri* into three genetic

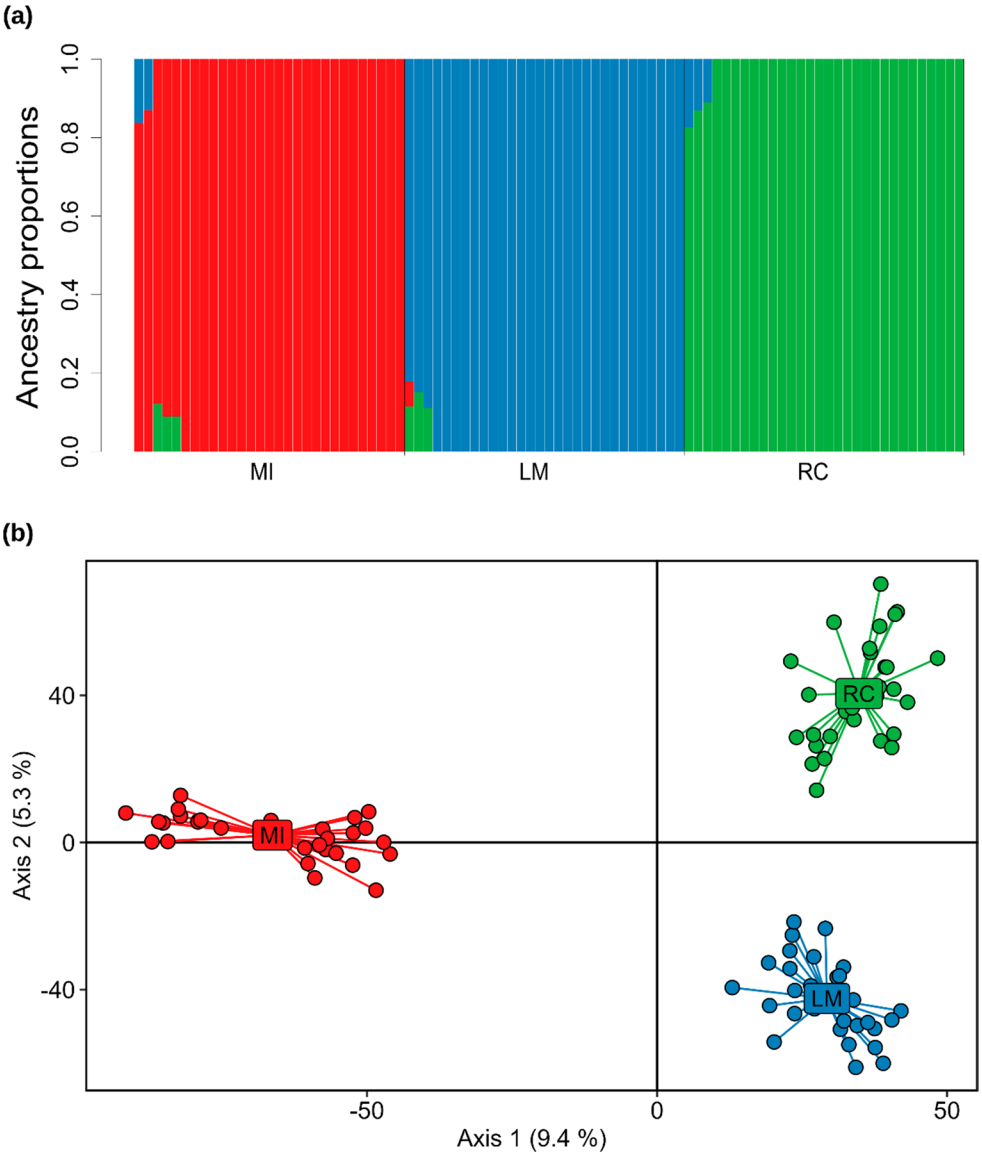


Fig. 3 Genetic differentiation of *G. turneri*: (**a**) individual ancestry proportion of plants for the most likely number of genetic clusters (K = 3) inferred in ADMIXTURE v. 1.3.0 and (**b**) principal component analysis (PCA) of *G. turneri* where each dot represents an individual. For both plots, colors indicate their respective collection sites: El Mirador (red), La Manga (blue), and Roca El Choyudo (green)

Table 1 Analysis of molecular variance of individuals and populations of *G. turneri*

Results	Degree of freedom	Sum of squares	Mean squares
Among populations	2	75,119.49	37,559.74
Between samples within population	86	457,526.36	5,320.07
Within samples	89	523,084.06	5,877.34
Total	177	1,055,729.92	5,964.57
Components of covariance	**Sigma**	**Percentage (%)**	**p-value**
Variation among populations	543.43	8.84	0.0009
Variation between samples within population	−278.63	−4.53	0.9730
Variation within samples	5,877.34	95.68	0.2507
Total variation	6,142.14	100	

clusters corresponding to their respective collection sites: El Mirador, La Manga, and Roca El Choyudo (Fig. 3a). Likewise, the PCA also grouped individuals of *G. turneri* according to their collection sites with no overlap among the groups (Fig. 3b). The PCA revealed three distinct genetic clusters, consistent with the analysis of individual ancestry proportions using K = 3 in ADMIXTURE v. 1.3.0.

Considering the inferred clusters as potential populations, the AMOVA showed significant genetic differentiation among populations (p-value = 0.0009) accounting for 8.84% of the total variation (Table 1). Similarly, global F_{ST} indicated moderate genetic differentiation among populations of *G. turneri* (F_{ST} = 0.088, 95% CI: 0.0873–0.0895). Pairwise F_{ST} values between El Mirador – La Manga and El Mirador – Roca El Choyudo were 0.099 (95% CI: 0.098–0.101) and 0.103 (95% CI: 0.102–0.105) respectively indicating a genetic differentiation between El Mirador and the

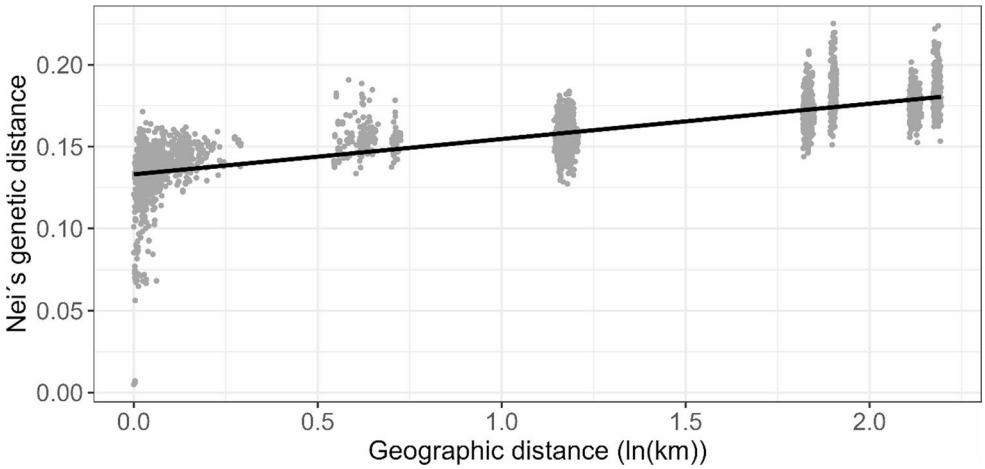


Fig. 4 Relationship between genetic and geographic distance among individuals of *G. turneri*

Table 2 Genetic diversity indexes estimate (mean ± standard error) in *G. turneri* for all samples and each population differentiated. Standard errors less than 0.00095 are indicated with a $^{+}$. For the FIS, the mean and the 95% CI were reported

Diversity index	Populations			Global (all samples)
	El Mirador	La Manga	Roca El Choyudo	
N	29	30	30	89
N_{EA}	1.839 ± 0.001	1.802 ± 0.001	1.829 ± 0.001	2.000 ± 0.000
A_R	1.830 ± 0.001	1.791 ± 0.001	1.818 ± 0.001	1.999 ± 0.000^{+}
π	0.243 ± 0.001	0.235 ± 0.001	0.239 ± 0.001	0.225 ± 0.001
H_O	0.244 ± 0.001	0.251 ± 0.001	0.256 ± 0.001	0.251 ± 0.000^{+}
H_E	0.238 ± 0.000^{+}	0.231 ± 0.000^{+}	0.235 ± 0.000^{+}	0.253 ± 0.000^{+}
F_{IS}	0.009 (0.061, 0.302)	−0.041 (−0.079, 0.009)	−0.041 (−0.107, 0.059)	−0.050 (−0.053, −0.047)

N: number of samples, N_{EA}: number of effective alleles per locus, A_R: allelic richness, π: nucleotide diversity, H_O: observed heterozygosity, H_E: expected heterozygosity, F_{IS}: inbreeding coefficient

rest of the populations. In contrast, the pairwise F_{ST} between La Manga – Roca El Choyudo was 0.061 (95% CI: 0.059–0.062). Despite the lower F_{ST} value between La Manga and Roca El Choyudo compared to El Mirador, the populations can still be differentiated according to the analysis of individual ancestry proportions (Fig. 3a) and the PCA (Fig. 3b). The Mantel test for IBD revealed a significant relationship between the genetic distances and geographical distances of the samples ($r = 0.791$, p-value = 0.0004) (Fig. 4), suggesting that greater geographical distances between samples correspond to greater genetic differentiation between them (Fig. 4).

The results from the analysis of individual ancestry proportions, PCA, AMOVA, and pairwise F_{ST} consistently demonstrated genetic clustering and differentiation among individuals and potential populations of *G. turneri*. Despite the species' limited distribution, individuals were consistently grouped into three genetic clusters or potential populations, which exhibited differentiation from one another. This alignment across multiple analyses underscores the robustness of the genetic structure observed in *G. turneri*.

Genetic diversity

To understand genetic diversity and identify potential conservation units, genetic diversity was analyzed globally (across all samples) and for each differentiated population. Across all samples, 2.000 ± 0.000 effective alleles were detected, with an A_R and π of 1.999 ± 0.000 and 0.225 ± 0.001, respectively. Among populations, N_{EA} varied between 1.802 ± 0.001 and 1.839 ± 0.001, with El Mirador exhibiting the greater number of effective alleles (Table 2). Additionally, El Mirador had the highest number of P_A, with 17,303 unique alleles, followed by Roca El Choyudo (9,910 unique alleles) and La Manga (5,429 unique alleles). The A_R and π varied from 1.791 ± 0.001 to 1.830 ± 0.001 and 0.235 ± 0.001 to 0.243 ± 0.001, respectively, among the populations (Table 2), with El Mirador seemingly harboring a greater number of different alleles and nucleotide polymorphisms.

Globally, across all samples, 7,380 loci (15.779%) were detected as significantly out of Hardy-Weinberg equilibrium. While the rest of the detected loci (> 84%) are in Hardy-Weinberg equilibrium. Besides, each population also exhibited loci out of Hardy-Weinberg equilibrium: El Mirador with 3,533 loci (7.553%), La Manga with 2,561 loci (5.475%), and Roca El Choyudo with 2,768 loci (5.918%).

The H_O for 89 individuals of *G. turneri* was slightly lower than H_E (Table 2), but not significantly after the t-test (p-value = 0.0798), indicating that, globally, the number of heterozygous individuals observed is the same as expected under Hardy-Weinberg equilibrium. However, in each population, the H_O was significantly higher than the H_E (Table 2) after t-test (p-values < 0.0001), suggesting an excess of heterozygous individuals. Among the populations, Roca El Choyudo showed the highest H_O, followed by La Manga and El Mirador (Table 2).

For all samples, the global FIS was −0.050 (Table 2). Additionally, among populations, the Fis values varied from 0.009 to −0.041 (Table 2). The FIS values were globally close to zero in all populations, indicating little evidence of inbreeding.

Census and effective population size

Census population size (N_c) varied from 62 to 883 individuals among populations, with Roca El Choyudo being the largest, followed by El Mirador and La Manga (Table 3). The effective population size (N_e), estimated using the linkage disequilibrium method, was consistently below 50 in all populations, with confidence intervals that did not overlap with zero (Table 3). The N_e/N_c ratio ranged from 0.03 to 0.52 among populations (Table 3).

Divergence history

The demographic analysis performed using DIYABC Random Forest v.1.0 identified scenario 5 (Fig. 5) as the most likely divergence scenario, receiving 658 votes and a posterior probability of 0.468. In this scenario, El Mirador (MI) represents the ancestral lineage, followed by an early divergence of Roca El Choyudo (RC) at time t2 (posterior distribution: 10 to 10,000 generations) and a later divergence of La Manga (LM) at time t1 (10 to 10,000 generations). After the divergence of LM, all three populations underwent a recent bottleneck starting at time t1–db, lasting between 10 and 100 generations. During these bottlenecks, effective population sizes were reduced to between 10 and 50 individuals (Fig. 5).

Table 3 Census and effective population size of *G. turneri*

Parameter	Populations		
	El Mirador	La Manga	Roca El Choyudo
N_e (95% CI)	18.7 (8.2, 65.7)	32.7 (21.8, 56.1)	27.0 (13.3, 89.4)
N_c	287	62	883
N_e/N_c	0.06	0.52	0.03

N_e: effective population size, N_c: census population size

Minimum sample size for conservation ex-situ

Simulations of expected heterozygosity (H_E) across different sample sizes ($n = 5$, 10, 15, 20, 25, 29, or 30) showed a consistent pattern across all three populations (Fig. 6). Genetic diversity in the resampled datasets reached values comparable to those of the original dataset (current diversity) when $p > 0.05$ (Fig. 6). In all three populations, a sample of 15 individuals was sufficient to capture the original genetic diversity. Therefore, collecting seeds from a random sample of 15 plants per population would be adequate for establishing an ex-situ collection of living plants while preserving the existing genetic variation.

Discussion

Levels and patterns of genetic diversity as well as differentiation were explored in *G. turneri* using genomic data from GBS to propose conservation strategies. Due to its limited distribution, relatively low N_e and increasing human disturbance in this area, it is necessary to identify potential conservation units or populations with higher genetic diversity to propose conservation plans. The results of this study have identified a population with higher genetic diversity as a conservation priority.

Genetic differentiation

Despite its small and restricted distribution, ancestry individual proportions and PCA analyses consistently grouped all sampled individuals into three genetic clusters, which can be considered as populations. Levels of differentiation among populations (F_{ST}) were moderate (0.088), with variation among populations from 0.061 to 0.102 (Wright 1965; Cai et al. 2021). These results are supported by the AMOVA analysis, which showed a significant genetic differentiation among populations with 8.84% of the total variation.

The levels of differentiation (F_{ST}) fall within the range previously reported for CWR (Table S1). Closer populations such as La Manga and Roca El Choyudo (≈2.2 km apart) showed lower values of differentiation than pairs of populations located at greater distances such as El Mirador and Roca El Choyudo (≈7.4 km apart), suggesting that gene flow could play an important role. The observed differentiation is congruent with what is known of major flower visitors (Yescas-Romo et al. 2024). Insects such as solitary, social bees and butterflies are known to transport pollen over limited distances (Gathmann and Tscharntke 2002; Smith et al. 2017; Webb and Bawa 1983), resulting in larger genetic differentiation among populations compared to plant species with highly mobile pollinators (Kramer et

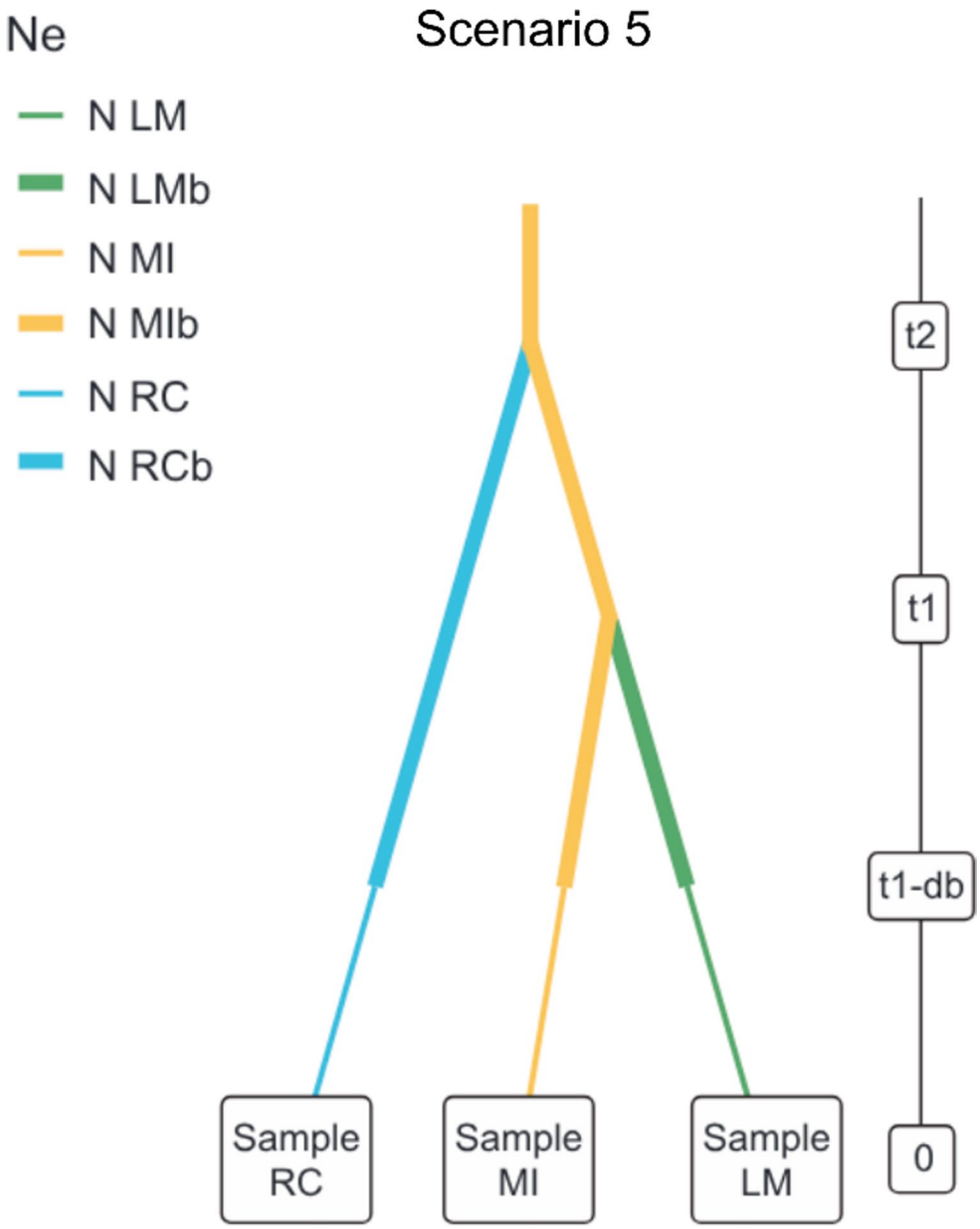


Fig. 5 Most likely divergence scenario for populations of *G. turneri* inferred using Approximate Bayesian Computation in DIYABC Random Forest v1.0. Roca El Choyudo (RC) and La Manga (LM) diverged from El Mirador (MI) at times t2 and t1, respectively, with both divergence times estimated to have occurred between 10 and 10,000 generations. Branch widths represent effective population sizes: wide branches indicate pre-bottleneck sizes, while narrow branches correspond to a bottleneck event beginning at time t1–db and lasting 10–100 generations. During the bottleneck, effective population sizes declined to between 10 and 50 individuals

al. 2011). Given that bees and butterflies are the most common flower visitors of *G. turneri*, it is not surprising that a pattern of isolation by distance among plants (Fig. 6) was detected, like other rare and endangered species distributed over a restricted range (Cohen and Griffen 2023). This pattern is also congruent with our observations that seed dispersal is by gravity and once in the soil, by water transport downstream (K. Yescas-Romo, personal observation).

The demographic analysis supported the divergence scenario in which El Mirador represent the ancestral lineage from which Roca El Choyudo and La Manga diverged (Fig. 5). This scenario aligns with the observed genetic data, as El Mirador retains a higher number of unique alleles and nucleotide polymorphisms. Additionally, the most supported scenario indicates that all three populations experienced recent bottlenecks, which likely influenced their levels of genetic diversity.

It is noteworthy that the ancestry individual proportions indicated a CV error of 0.428 for the next best model fit for the samples after K = 3, which was identified as K = 4. In this scenario, a new isolated patch (locally known as Piedras Pintas) in El Mirador was identified as a genetic cluster (Fig. S3). This genetic cluster identified by K = 4 was previously part of the continuous distribution of *G. turneri* in the El Mirador population. However, the construction of new houses in recent years has fragmented the distribution of this wild cotton in the population (Fig. S4). Although K = 4 is not the best fit for all samples, this model identifies Piedras

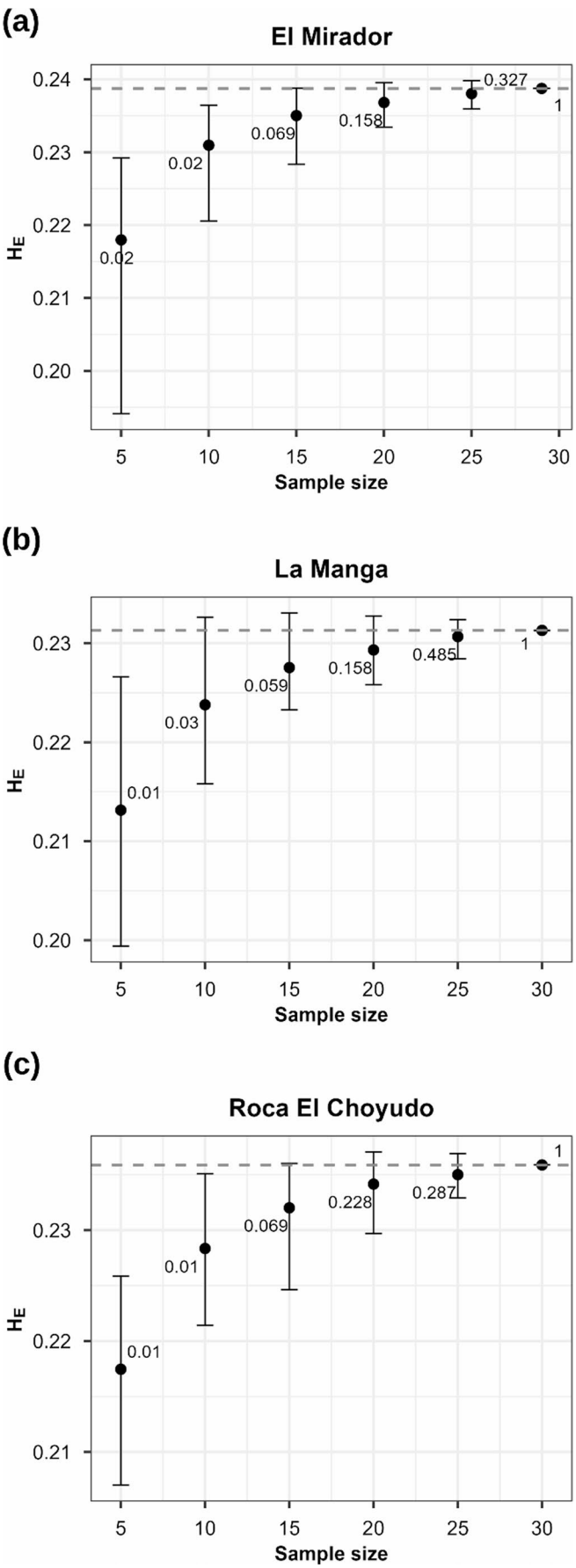


Fig. 6 Estimated expected heterozygosity (H_E) values for the resampled datasets using different sample sizes for populations of *G. turneri*: (a) El Mirador, (b) La Manga and (c) Roca El Choyudo. Points represent the average H_E, calculated using 100 bootstraps for each sample size. Error bars indicate the 95% CI and dashed lines represent the H_E values for each population. For each sample size in the populations, the corresponding p-values are displayed within the plot

Pintas as a genetic cluster and potentially a new population if fragmentation continues in this population (Fig. S3 and S4). This example highlights how tourism-driven coastal development in this area is causing fragmentation of the populations of this wild cotton.

Genetic diversity

The results regarding observed and expected heterozygosity both globally (H_O = 0.251, H_E = 0.253) and for populations (H_O = 0.244–0.256, H_E = 0.231–0.238) indicate that *G. turneri* has moderate levels of genetic diversity. While previous studies have investigated the genetic diversity in other wild cottons (Kuester and Nason 2012; de Menezes et al. 2014), these analyses were based on different markers, making direct comparisons challenging. However, a comparison with other CRW using reduced representation approaches shows that the results of the present study fall within the range that has been previously recorded (Table S1). Even so, *G. turneri* showed higher levels of genetic diversity than in pumpkin species such as *Curcubita argyrosperma* and *C. pepo* (Table S1) and lower levels compared to *Coffea canephora* (Table S1).

The analysis of F_{IS} indicates minimal evidence of inbreeding in *G. turneri*. Both the global F_{IS} value (−0.05) and the values observed within populations (−0.041 to 0.009) are within the range of values previously reported in CWR (Table S1) and are consistent with the known pollination biology of *G. turneri* (Yescas-Romo et al. 2024). Despite being self-compatible and capable of autonomous self-pollination, the F_{IS} values suggest little or no inbreeding. This is likely due to the approach herkogamy observed in most flowers, where stigmas are positioned above anthers (Yescas-Romo et al. 2024) and in consequence, outcrossing is likely the most common mating event. Bees and butterflies have been documented as the most common legitimate visitors (Yescas-Romo et al. 2024) and most likely they are depositing outcross pollen on stigmas during visits. Hence, the observed F_{IS} values are most likely a reflection of the mating system of *G. turneri*. However, it is necessary to estimate the outcrossing rates to confirm if this species is predominantly an outcrosser.

The estimated values of the contemporary effective population sizes (N_e) were <50 in the three populations.

Although N_e values for *G. turneri* populations (Table 3) are around the median reported for plants (Clarke et al. 2024), all three populations fall below the recommended threshold of 50 to prevent short-term inbreeding (Hoban et al. 2024). This suggests that the studied populations may be at significant inbreeding risk (Hoban et al. 2024). Census population sizes ranged from 62 to 883 adult plants (Table 3). Over the past five years, we have observed the removal of dozens of plants at La Manga and increasing fragmentation at El Mirador. These disturbances are likely to impact population dynamics, making it urgent to conduct demographic studies on both populations to assess their status and conservation needs.

Conservation

For this endangered CWR, this study reveals moderate levels of genetic diversity and differentiation, with minimal evidence of inbreeding. Moreover, the analysis of ancestry proportions and PCA consistently clustered individuals into three genetic groups that can be considered as populations and potential conservation units (Alves et al. 2023). Data on P_A, A_R, and π indicate that El Mirador harbors a higher number of unique and different alleles as well as nucleotide polymorphisms, highlighting its significance as a conservation priority. However, this population has recently experienced fragmentation due to coastal house construction (Fig. S4), leading to the elimination of numerous plants (K. Yescas-Romo, personal observation). Urgent seed collection is therefore recommended for the three populations for ex-situ conservation in a Botanical Garden (Cibrian-Jaramillo et al. 2013), prioritizing El Mirador due to its potential as the largest reservoir of genetic variation of *G. turneri*. Our analysis suggests that at least seeds of 15 random individuals should be collected per population to capture the genetic diversity of *G. turneri*. Additionally, we recommend maintaining each population isolated from each other as living plants in a Botanical Garden within the Sonoran Desert to ensure they experience similar abiotic and biotic conditions as in their natural habitat. Given the abundance and diversity of native bees in the region, these plants could be pollinated by bees and butterflies, as observed in natural populations (Yescas-Romo et al. 2024), rather than relying solely on self-pollination.

Unfortunately, the in-situ conservation prospects for the La Manga population of *G. turneri* are bleak due to its location within a small human settlement (Fig. S5). This population, being the smallest ones (N_c = 62), faces ongoing elimination of individuals due to activities such as house and road construction (K. Yescas-Romo, personal observation). However, it is worth noting that within this settlement, there has been a genuine interest among residents in conservation efforts. Therefore, implementing a program to encourage the adoption of this species as an ornamental plant in house gardens is proposed, considering that people find flowers quite attractive (Fig. 1). Although this proposal has potentially negative consequences, such as a reduction in H_O and modifications in traits associated with adaptation to water stress, it is probably the only viable option. Implementation of this proposal could begin with community involvement, especially through the participation of teachers at the local primary school, where children can learn how to propagate and care for the species.

In contrast to La Manga, the in-situ conservation outlook for El Mirador and particularly Roca El Choyudo is more promising, as until now this last population remains unaffected by human settlements (Fig. S6). For El Mirador, we propose a ¨micro-reserve¨ (sensu Laguna et al. 2004), surrounding the public space of this tourism attraction, restricting any development in this area. For Roca El Choyudo, we propose restricting coastal development. Along its distribution, *G. turneri* shares habitat with other plant species listed in the Mexican Norm of species in danger of extinction (e.g., *Peniocereus striatus*) (DOF 2019). Although this area was within a reserve (Cajón del diablo or Devil´s drawer) decreed in 1937 (SEMARNAP 2000), this was just a paper reserve as it was never implemented (Quijada-Siu 2015). Therefore, we propose strictly restricting coastal development in this area by enforcing the rules governing land use changes in Mexico (DOF 2000).

Tourism-driven coastal development often prioritizes short-term economic gains while generating a range of negative social and environmental impacts (Murray 2007). Tourism relies on natural resources and ecosystem services that are also essential for local communities (Williams and Ponsford 2009). The construction of hotels and other infrastructure can lead to habitat loss and fragmentation, compromising ecosystem services and threatening sensitive endemic species (Samad et al. 2021; Médail et al. 2021), such as *G. turneri*. Our study highlights how coastal development north of San Carlos is endangering the persistence of this endemic species. To mitigate environmental impacts, future tourism-related development in this region should integrate ecological knowledge into planning and decision-making processes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10592-025-01712-3.

Acknowledgements K. Yescas-Romo would like to thank the Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT-Mexico) for the PhD fellowship No. 774907. We are grateful to José Fulgencio Martínez Rodríguez and Karen Yoselin Jimenez Cedillo for their great assistance during fieldwork and Tania Molina for Fig 5 and Fig.S1. Special thanks to Carolina Aguirre for her invaluable assistance in DNA extractions and Dr. Luis Fernando Garcia Ortega for advice on sequencing.

Author contributions Sample collections were performed by K. F. Y. R. and F. M. F.; DNA extraction by K. F. Y. R. and data analysis by K. F. Y. R. and F. R. M. Interpretation of data/results and writing of the first draft of the manuscript by K. F. Y. R. and F. M. F. Project design and leadership by C. H. K. and F. M. F. All authors commented on previous versions of the manuscript and all authors read and approved the final manuscript.

Funding This study was funded by the operating funds from UNAM´s Institute of Ecology and project number 2203006 by the Interdisciplinary Faculty of Biological and Health Sciences of the Universidad de Sonora.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alexander DH, Lange K (2011) Enhancements to the ADMIXTURE algorithm for individual ancestry Estimation. BMC Bioinformatics 12:1–6. https://doi.org/10.1186/1471-2105-12-246

Alves F, Banks SC, Edworthy M, Stojanovic D, Langmore NE, Heinsohn R (2023) Using conservation genetics to prioritise management options for an endangered Songbird. Heredity 130:289–301. https://doi.org/10.1038/s41437-023-00609-6

Anderson LG, Rocliffe S, Haddaway NR, Dunn AM (2015) The role of tourism and recreation in the spread of non-native species: a systematic review and meta-analysis. PLoS ONE 10:1–15. https://doi.org/10.1371/journal.pone.0140833

Andrews S (2019) FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Azman A, Ng KKS, Ng CH et al (2020) Low genetic diversity indicating the threatened status of *Rhizophora apiculata* (Rhizophoraceae) in malaysia: declined evolution Meets habitat destruction. Sci Rep 10:1–12. https://doi.org/10.1038/s41598-020-76092-4

Bartlett MS (1937) Properties of sufficiency and statistical tests. Proc R Soc London Ser A 160:268–282. https://doi.org/10.1098/rspa.1937.0109

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Brenner L, Engelbauer M, Job H (2018) Mitigating tourism-driven impacts on mangroves in Cancún and the Riviera may, mexico: an evaluation of conservation policy strategies and environmental planning instruments. J Coastal Conserv 22:755–767. https://doi.org/10.1007/s11852-018-0606-0

Buckley R, Pascual U (2024) Role of tourism. In: Scheiner SM (ed) Encyclopedia of biodiversity, 3rd edn. Elsevier, Amsterdam, pp 562–566

Cai C, Xiao J, Ci X, Conran JG, Li J (2021) Genetic diversity of *Horsfieldia tetratepala* (Myristicaceae), an endangered plant species with extremely small populations to china: implications for its conservation. Plant Syst Evol 307:1–12. https://doi.org/10.1007/s00606-021-01774-z

Chávez-Valdez A, Enríquez-Acosta JA (2022) Agentes económicos y Su Entorno competitivo Del turismo En San Carlos Nuevo guaymas, México. Revista Propuestas Para El Desarrollo 6:1–18

Chung MY, Merilä J, Li J, Mao K, López-Pujol J, Tsumura Y, Chung MG (2023) Neutral and adaptive genetic diversity in plants: an overview. Front Ecol Evol 11:1–14. https://doi.org/10.3389/fevo.2023.1116814

Cibrian-Jaramillo A, Hird A, Oleas N, Ma H, Meerow AW, Francisco-Ortega J, Griffith MP (2013) What is the conservation value of a plant in a botanic garden? Using indicators to improve management of ex situ collections. Bot Rev 79:559–577. https://doi.org/10.1007/s12229-013-9120-0

Cisneros-Montemayor AM, Townsel A, Gonzales CM, Haas AR, Navarro-Holm EE, Salorio-Zuñiga T, Johnson AF (2020) Nature-based marine tourism in the Gulf of California and Baja California peninsula: economic benefits and key species. Nat Resource Forum 44:111–128. https://doi.org/10.1111/1477-8947.12193

Clarke SH, Lawrence ER, Matte JM, Gallagher BK, Salisbury SJ, Michaelides SN, Koumrouyan R, Ruzzante DE, Grant JWA, Fraser DJ (2024) Global assessment of effective population sizes: consistent taxonomic differences in meeting the 50/500 rule. Mol Ecol 33:1–14. https://doi.org/10.1111/mec.17353

Cohen JI, Griffen LR (2023) Conservation genetics of *Phlox hirsuta*, a serpentine endemic. Conserv Genet 24:137–152. https://doi.org/10.1007/s10592-022-01478-y

Collin FD, Durif G, Raynal L, Lombaert E, Gautier M, Vitalis R, Marin JM, Estoup A (2021) Extending approximate bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC random forest. Mol Ecol Resour 21:2598–2613. https://doi.org/10.1111/1755-0998.13413

Danecek P, Auton A, Abecasis G et al (2011) The variant call format and vcftools. Bioinformatics 27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek P, Bonfield JK, Liddle J et al (2021) Twelve years of samtools and BCFtools. GigaScience. https://doi.org/10.1093/gigascience/giab008

de Menezes IPP, Gaiotto FA, Hoffmann LV, Ciampi AY, Barroso PAV (2014) Genetic diversity and structure of natural populations of *Gossypium mustelinum*, a wild relative of cotton, in the basin of the de Contas river in bahia, Brazil. Genetica 142:99–108. https://doi.org/10.1007/s10709-014-9757-6

Diario Oficial de la Federación (DOF) (2000) Reglamento de la Ley General del Equilibrio Ecológico y la Protección al Ambiente en Materia de Areas Naturales Protegidas. https://www.dof.gob.mx/nota_detalle.php?codigo=2065550&fecha=30/11/2000#gsc.tab=0. Accessed 12 May 2024

Diario Oficial de la Federación (DOF) (2019) Anexo Normativo III. Lista de especies en riesgo de la Norma Oficial Mexicana NOM-059-SEMARNAT-2010. https://www.dof.gob.mx/nota_detalle.

php?codigo=5578808&fecha=14/11/2019#gsc.tab=0. Accessed 29 Apr 2024
DiPaolo DA, Villela JC (2022) Imperiled plants: introduction. In: DellaSala DA, MI Goldstein MI (eds) Imperiled: the encyclopedia of conservation. Elsevier, Amsterdam, pp 631–635
Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NeEstimator v2: re-implementation of software for the Estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour 14:209–214. https://doi.org/10.1111/1755-0998.12157
Dorsey BL, Salas-Morales SH, Gregory TJ (2023) Conservation genomics of *Dioon holmgrenii* (Zamiaceae) reveals a history of range expansion, fragmentation, and isolation of populations. Conserv Genet 25:335–355. https://doi.org/10.1007/s10592-023-01569-4
Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22:1–20. https://doi.org/10.18637/jss.v022.i04
Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE. https://doi.org/10.1371/journal.pone.0019379
Enríquez-Acosta JA, Ávila-Martínez L (2017) Impacto Del turismo residencial y Sol y Playa En Dos Lugares turísticos de Sonora. In: Enríquez-Acosta JA, Guillén-Lúgigo M, Valenzuela B (eds) (Coord) Patrimonio y turismo. Universidad de Sonora, Hermosillo, pp 16–36
Fryxell PA (1978) *Gossypium turneri* (Malvaceae), a new species from Sonora. Mexico Madroño 25:155–159
Fryxell PA (1988) Systematic botany monographs. Malvaceae of Mexico. The American Society of Plant Taxonomists, Mexico
Gathmann A, Tscharntke T (2002) Foraging ranges of solitary bees. J Anim Ecol 71:757–764. https://doi.org/10.1046/j.1365-2656.2002.00641.x
Goettsch B, Urquiza-Haas T, Koleff P et al (2021) Extinction risk of Mesoamerican crop wild relatives. Plants People Planet 3:775–795. https://doi.org/10.1002/ppp3.10225
Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-statistics. Mol Ecol Notes 5:184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x
Gruber B, Unmack PJ, Berry OF, Georges A (2018) Dartr: an r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Mol Ecol Resour 18:691–699. https://doi.org/10.1111/1755-0998.12745
Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48:361–372
Hoban S, Callicrate T, Clark J, Deans S, Dosmann M, Fant J, Gailing O, Havens K, Hipp AL, Kadav P, Kramer AT, Lobdell M, Magellan T, Meerow AW, Meyer A, Pooler M, Sanchez V, Spence E, Thompson P, Toppila T, Walsh S, Westwood M, Wood J, Griffith MP (2020) Taxonomic similarity does not predict necessary sample size for ex situ conservation: a comparison among five genera. Proc R Soc B 287:1–9. https://doi.org/10.1098/rspb.2020.0102
Hoban S, da Silva JM, Hughes A, Hunter ME, Kalamujić Stroil B, Laikre L, Mastretta-Yanes A, Millette K, Paz-Vinas I, Ruiz Bustos L, Shaw RE, Vernesi C, Coalition for Conservation Genetics (2024) Too simple, too complex, or just right? Advantages, challenges, and guidance for indicators of genetic diversity. Bioscience 74:269–280. https://doi.org/10.1093/biosci/biae006
IUCN (2023) The IUCN Red List of Threatened Species. Version 2023-1. https://www.iucnredlist.org. Accessed 29 Apr 2024
Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:403–1405. https://doi.org/10.1093/bioinformatics/btn129
Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. https://doi.org/10.7717/peerj.281
Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013) DiveRsity: an R package for the Estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 4:782–788. https://doi.org/10.1111/2041-210X.12067
Kramer AT, Fant JB, Ashley MV (2011) Influences of landscape and pollinators on population genetic structure: Examples from three *Penstemon* (Plantaginaceae) species in the Great Basin. Am J of Bot 98:109–121. https://doi.org/10.3732/ajb.1000229
Kuester AP, Nason JD (2012) Microsatellite loci for *Gossypium davidsonii* (Malvaceae) and other D-genome. Sonoran desert endemic cotton species. Am J Bot 99:e91–e93. https://doi.org/10.3732/ajb.1100421
Laguna E, Deltoro VI, Pérez-Botella J, Pérez-Rovira P, Serra L, Olivares A, Fabregat C (2004) The role of small reserves in plant conservation in a region of high diversity in Eastern Spain. Biol Conserv 119:421–426. https://doi.org/10.1016/j.biocon.2004.01.001
Li S, Li D, Zhang L, Shang W, Qin B, Jiang Y (2021) High levels of genetic diversity and an absence of genetic structure among breeding populations of the endangered Rufous-backed bunting in inner mongolia, china: implications for conservation. Avian Res 12:1–11. https://doi.org/10.1186/s40657-020-00236-3
Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200
Md V, Misra S, Li H, Aluru S (2019) Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In: Proceedings of 2019 IEEE international parallel & distributed processing symposium IPDPS 2019, pp 314–324. https://doi.org/10.1109/IPDPS.2019.00041
Médail F, Diadema K, Pouget M, Baumel A (2021) Identification of plant micro-reserves using conservation units and population vulnerability: the case of an endangered endemic snowflake (*Acis nicaeensis*) in the mediterranean basin hotspot. J Nat Conserv 61:1–8. https://doi.org/10.1016/j.jnc.2021.125980
Monteiro WP, Dalapicolla J, Carvalho CS, Costa Veiga J, Vasconcelos S, Ramos SJ, Gastauer M, Jaffé R, Caldeira CF (2022) Genetic diversity and structure of an endangered medicinal plant species (*Pilocarpus microphyllus*) in Eastern amazon: implications for conservation. Conserv Genet 23:745–758. https://doi.org/10.1007/s10592-022-01454-6
Murray G (2007) Constructing paradise: the impacts of big tourism in the Mexican coastal zone. Coastal Manage 35:339–355. https://doi.org/10.1080/08920750601169600
Navarro N (2019) Community perceptions of tourism impacts on coastal protected areas. J Mar Sci Eng 7:1–18. https://doi.org/10.3390/jmse7080274
Nei M (1972) Genetic distance between populations. Am Nat 106:283–292. https://doi.org/10.1086/282771
Paradis E (2010) Pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 26:419–420. https://doi.org/10.1093/bioinformatics/btp696
Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460
Pembleton L, Cogan N, Forster J (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol Ecol Resour 13:946–952. https://doi.org/10.1111/1755-0998.12129

Petitpierre B, Boserup J, Möhl A, Rometsch S, Aubry S (2023) Importance of agriculture for crop wild relatives conservation in Switzerland. Global Ecol Conserv 46:1–12. https://doi.org/10.1016/j.gecco.2023.e02588

Quijada-Siu F (2015) Gestión para la conservación ecológica del territorio: el caso del turismo de aventura en el Cañón del Nacapule, San Carlos, Sonora. Dissertation, El Colegio de Sonora

R Core Team (2024) R: A Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rankin BL, Ballantyne M, Pickering CM (2015) Tourism and recreation listed as a threat for a wide diversity of vascular plants: A continental scale review. J Environ Manage 154:293–298. https://doi.org/10.1016/j.jenvman.2014.10.035

Rimmer A, Phan H, Mathieson I et al (2014) Integrating mapping-, assembly-and haplotype-based approaches for calling variants in clinical sequencing applications. Nat Genet 46:912–918. https://doi.org/10.1038/ng.3036

Rivera-Martínez CM (2020) Capital natural y turismo náutico en San Carlos, Sonora, México. Dissertation, Centro de Investigaciones Biológicas del Noroeste

Samad FA, Medail F, Baumel A, Yakovlev SS, Roukoz A, Tohme G et al (2021) New insights on the conservation status of the endangered coastal endemic plant *Astragalus Berytheus* (Fabaceae) in Lebanon. Oryx 55:519–521. https://doi.org/10.1017/S0030605320000149

Secretaría de Medio Ambiente, Recursos Naturales y Pesca (SEMARNAP) (2000) Áreas naturales protegidas de México Con decretos federales (1899–2000). Instituto Nacional de Ecología-SEMARNAP and Programa de Naciones Unidas para el Desarrollo (PNUD), Mexico

Simone O, Chiara M, Guiseppe F et al (2018) Red listing plants under full National responsibility: extinction risk and threats in the vascular flora endemic to Italy. Biol Conserv 224:213–222. https://doi.org/10.1016/j.biocon.2018.05.030

Smith JP, Heard TA, Beekman M, Gloag R (2017) Flight range of the Australian stingless bee *Tetragonula carbonaria* (Hymenoptera: Apidae). Austral Entomol 56:50–53. https://doi.org/10.1111/aen.12206

Sulis E, Bacchetta G, Cogoni D, Gargano D, Fenu G (2020) Assessing the global conservation status of the rock Rose *Helianthemum caput-felis*. Oryx 54:197–205. https://doi.org/10.1017/S0030605318001424

Udall JA, Long E, Hanson C et al (2019) De novo genome sequence assemblies of *Gossypium raimondii* and *Gossypium turneri*. G3: Genes, Genomes, Gene 9:3079–3085. https://doi.org/10.1534/g3.119.400392

Ulloa M (2014) The diploid D genome cottons (*Gossypium* spp.) of the new world. In: Abdurakhmonov IY (ed) World cotton germplasm resources. InTech, Rijeka, pp 203–212

von Wettberg E, Toker C, Öskan H, Smýkal P (2022) Endangered wild crop relatives of the fertile crescent. In: Dellasala DA, Goldstein MI (eds) Imperiled: the encyclopedia of conservation. Elsevier, Amsterdam, pp 673–682

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conserv Genet 7:167–184. https://doi.org/10.1007/s10592-005-9100-y

Webb CJ, Bawa KS (1983) Pollen dispersal by hummingbirds and butterflies: a comparative study of two lowland tropical plants. Evolution 37:1258–1270. https://doi.org/10.1111/j.1558-5646.1983.tb00242.x

Wegier A, Alavez V, Vega M (2018) *Gossypium turneri*. The IUCN Red List of Threatened Species 2018. https://doi.org/10.2305/IUCN.UK.2018-1.RLTS.T71775005A71775008.en. Accessed 29 April 2024

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-, New York

Williams PW, Ponsford IF (2009) Confronting tourism's environmental paradox: transitioning for sustainable tourism. Futures 41:396–404. https://doi.org/10.1016/j.futures.2008.11.019

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395–420. https://doi.org/10.1111/j.1558-5646.1965.tb01731.x

Wu Q, Dong S, Zhao Y et al (2023) Genetic diversity, population genetic structure and gene flow in the rare and endangered wild plant *Cypripedium Macranthos* revealed by genotyping-by-sequencing. BMC Plant Biol 23:1–13. https://doi.org/10.1186/s12870-023-04212-z

Yescas-Romo KF, Molina-Freaner FE, Ramírez-Godina F, Reyes-Valdés MH, Martínez O, Hayano-Kanashiro C (2021) Seed morphology and cytogenetic characterization of an endangered wild cotton, *Gossypium turneri* fryxell, from Mexico. Genet Resour Crop Evol 68:2119–2129. https://doi.org/10.1007/s10722-021-01124-5

Yescas-Romo KF, Hayano-Kanashiro C, Franer-Molina F (2024) Pollination biology of *Gossypium turneri*: autonomous selfing provides reproductive assurance in an endangered wild cotton from sonora, Mexico. Plant Species Biol 39:260–274. https://doi.org/10.1111/1442-1984.12474

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.