Molecular Ecology (2005) **14**, 1603–1611 doi: 10.1111/j.1365-294X.2005.02494.x

# Genetic variation and evolution of *Polaskia chichipe* (Cactaceae) under domestication in the Tehuacán Valley, central Mexico

ADRIANA OTERO-ARNAIZ,*‡ ALEJANDRO CASAS,* JAMES L. HAMRICK† and JENNIFER CRUSE-SANDERS†§

*Centro de Investigaciones en Ecosistemas, Universidad Nacional Autónoma de México, Campus Morelia. Apdo. Postal 27-3 (Xangari) Morelia 58089, Michoacán, México, †Departments of Plant Biology and Genetics, 2502 Plant Sciences, University of Georgia, Athens, Georgia 30602, USA

**Abstract**

**_Polaskia chichipe_ is a columnar cactus under artificial selection in central Mexico because of its edible fruits. Our study explored the effect of human manipulation on levels and distribution of genetic variation in wild, silviculturally managed and cultivated sympatric populations. Total genetic variation, estimated in nine populations with five microsatellite loci, was $H_T = 0.658 \pm 0.026$ SE, which was mainly distributed within populations ($H_S = 0.646$) with low differentiation among them ($F_{ST} = 0.015$). Fixation index ($F_{IS}$) in all populations was positive, indicating a deficit of heterozygous individuals with respect to Hardy–Weinberg expectations. When populations were pooled by management type, the highest expected heterozygosity ($H_E = 0.631 \pm 0.031$ SE) and the lowest fixation index ($F_{IS} = 0.07$) were observed in wild populations, followed by cultivated populations ($H_E = 0.56 \pm 0.03$ SE, $F_{IS} = 0.14$), whereas the lowest variation was found in silviculturally managed populations ($H_E = 0.51 \pm 0.05$ SE, $F_{IS} = 0.17$). Low differentiation among populations under different management types ($F_{ST}$ 0.005, $P < 0.04$) was observed. A pattern of migration among neighbouring populations, suggested from isolation by distance ($r^2 = 0.314$, $P < 0.01$), may have contributed to homogenizing populations and counteracting the effects of artificial selection. _P. chichipe_, used and managed for at least 700 generations, shows morphological differentiation, changes in breeding system and seed germination patterns associated with human management, with only slight genetic differences detected by neutral markers.**

*Keywords*: columnar cacti, crop evolution, dinucleotide repeats, domestication, genetic resources, genetic structure, microsatellites

*Received 4 May 2004; revision received 3 September 2004; accepted 12 January 2005*

## Introduction

The semiarid Tehuacán-Cuicatlán Valley, central Mexico has been identified as a key location for the origins of agriculture and plant domestication in the New World (MacNeish 1967). People of this region domesticated a number of plant species including species of columnar cacti (Casas *et al*. 1999, 2002). Currently, people in the region gather fruit from 20 species of columnar cacti from wild populations. They also practice silvicultural management of wild populations of 12 species *in situ* that involves favouring individuals with useful phenotypes when natural vegetation is cleared for agriculture (Casas *et al*. 1999). Seven of these species are also cultivated in home gardens and terraces, mainly through vegetative propagation (Casas *et al*. 1999, 2001).

Both cultivation and silvicultural management may involve domestication processes (Casas *et al*. 1999) that can lead to phenotypic divergence between wild and manipulated populations (De Candolle 1882; Darwin 1883; Zohary

Correspondence: Adriana Otero-Arnaiz, ‡Present address: Department of Biological Sciences, Indiana University, South Bend IN 46634-7111, USA; Fax: 574-520-6589; E-mail: moteroar@iusb.edu. §Present address: Rancho Santa Ana Botanic Garden, 1500 N. College Ave., Claremont, CA 91711, USA.

1984; Harlan 1992). Plant domestication and human manipulation of populations occur at varying intensities and can result in a continuum of morphological, physiological, and genetic differentiation between wild and partially or fully domesticated variants. Studies of both incipient and advanced degrees of domestication contribute to an understanding of how the process of evolution under artificial selection operates. They also allow for evaluation of the impact of the current traditional management on genetic diversity. Columnar cacti in the Tehuacán Valley are ideal biological systems for studying ongoing domestication processes and genetic interactions through gene flow among wild and domesticated populations. Fruits of these plant species are the primary targets of artificial selection, and farmers select on phenotypic characters such as fruit size, pulp colour and flavor, peel thickness, and thorniness (Casas & Barbera 2002).

*Polaskia chichipe* (Glosselin) Backeberg is one of the columnar cacti species under silvicultural management and cultivation in the Tehuacán Valley (Casas *et al*. 1999). Recent studies (Carmona & Casas 2005) documented that artificial selection has led to morphological differentiation among populations under different forms of management. On average, fruits from cultivated populations are significantly larger with more pulp, and contain larger and more numerous seeds than those from silviculturally managed populations. On the other hand, fruits and seeds from silviculturally managed populations are larger than those from unmanaged wild populations (Carmona & Casas 2005). Furthermore, artificial selection favoured individuals that were able to produce seeds through self-pollination leading to higher selfing rates among managed plants than nonmanaged individuals. This pattern was evident in pollination experiments. A longer fruit production season, as well as faster and higher percentage of seedling germination, was also found in populations manipulated by humans (Otero-Arnaiz *et al*. 2003).

Populations under silvicultural management and cultivation are derived from wild populations and all coexist in the same geographical area. Artificial selection on silviculturally managed populations of *P. chichipe* involves tolerating, enhancing in abundance or propagating desirable individuals while clearing nondesirable individuals from agricultural fields and pastures. Propagation of desirable individuals is conducted by sowing seeds and sometimes by planting branch cuttings (Carmona & Casas 2005).

Genetic differentiation among *P. chichipe* populations was expected because previous experiments with crosses among and within populations demonstrated that fruit production was more successful when pollen was donated from the same population (Otero-Arnaiz *et al*. 2003). Additionally, flowering peaks occur with a difference of 1 month between wild and silvicultural populations and another month between silvicultural and cultivated populations resulting in partial temporal isolation among these populations (Otero-Arnaiz *et al*. 2003).

In other species of columnar cacti studied, genetic structure among populations, measured as $G_{ST}$, was usually low. The only exception being *Lophocoreus schotti*, which has a relatively high $G_{ST}$ among populations from separate geographical regions in Sonora and Baja California, Mexico (Hamrick *et al*. 2002). Given the known patterns of management, artificial selection, and reproductive biology of *P. chichipe*, as well as the patterns of genetic structure reported for other columnar cacti, we explored the following topics: (i) the effects of management practices on genetic diversity of this cactus species; (ii) the effect of migration in relation to geographical distance in the genetic structure of populations; (iii) the levels of inbreeding among populations under different management types. According to previous studies on reproductive biology in *P. chichipe* (Otero-Arnaiz *et al*. 2003), the level of inbreeding was predicted to be higher in manipulated populations than in wild populations.

This study used microsatellites to survey the genetic variation and structure of wild, silvicultural and cultivated populations of *P. chichipe*. The results are interpreted in the light of information derived from studies of morphological variation (Carmona & Casas 2005) and reproductive biology (Otero-Arnaiz *et al*. 2003) to improve understanding of the evolutionary consequences of human management of *P. chichipe* populations.

## Materials and methods

### *Populations, individuals and plant tissue sampled*

A total of nine populations (three wild, three silviculturally managed, and three cultivated) were sampled within and around the villages of San Luis Atolotitlán, Caltepec, and Metzontla, in the state of Puebla, Mexico (Fig. 1). Flower buds were collected from 30 individuals per population and maintained at –70 °C until DNA extraction. A method recommended for DNA extraction of cactus (De la Cruz *et al*. 1997), modified for miniprep, was used.

### *Microsatellite development and genotyping*

Five microsatellite loci were identified specifically for *Polaskia chichipe*. Three of them were described previously (Otero-Arnaiz *et al*. 2004), and two were developed *de novo* following the same procedure (Table 1). Polymerase chain reactions (PCRs) were performed using 7.5-μL reactions: 50–100 ng of DNA, 0.25 mM each dNTP, 2.5 mM $MgCl_2$, 0.5 μM of each primer (fluorescent with either FAM, HEX or TET), 0.25 unit of *Taq* DNA polymerase (Perkin-Elmer), and 1 × *Taq* buffer (Perkin-Elmer), 1 × *Taq* master buffer.


**Fig. 1** The Tehuacán Valley (study area). Sampled locations in the municipalities of Caltepec and Zapotitlan, Puebla, Mexico. Wild (W) and Silviculturally managed (M) populations. Cultivated populations are shown as San Luis, Atolotitlan (C1), Caltepec (C2), and Mentzontla (C3).

The reactions were denatured at 94 °C for 4 min, followed by 25 cycles of denaturing at 94 °C for 1 min, annealing at primer-specific temperatures (Table 1) for 1 min, extension at 72 °C for 2 min and a final extension at 72 °C for 4 min, using a GenAmp 9700 thermocycler (Applied Biosystems). PCR products were mixed with formamide and ROX-500 size standard (Applied Biosystems). Samples were denatured 2 min at 95 °C. Analysis of microsatellite loci was conducted using an ABI 377–96 DNA sequencer in Genescan mode to detect the labelled primers and internal size standard in a 4.5% denaturing polyacrylamide gel. Allele size and scoring was performed with GENOTYPER software (Applied Biosystems).

### *Mating patterns*

To estimate levels of selfed vs. outcrossed offspring among maternal individuals, fruits of seven to 14 (mean = 10.7) mother trees were sampled from a total of six populations (two cultivated, two silviculturally managed and two wild) in April 2001. Seeds were germinated in January 2002. Between 11 and 17 seedlings per mother tree per population (956 seedlings in total) were genotyped, and maternal genotype was directly assessed.

For each population, the mean outcrossing ($t$) and selfing ($s = 1 - t$) rates were estimated. A minimum estimate of apparent selfing caused by biparental inbreeding was calculated as the difference between the single-locus selfing rate and the multilocus selfing rate ($bf = tm - ts$). The proportion of plants with exclusively outcrossing progeny ($T$) was estimated as another measure of mating system. Expectation–maximization iteration was used to find maximum-likelihood estimations to infer allele frequencies in the gene pool and the proportion of progeny that are the result of outcrossing using the program MLTR 2.4 (Ritland 2002). Standard errors for the outcrossing rates were calculated from 1000 bootstrap using the family as the unit of resampling.

### *Genetic variation within and among populations and management types*

For each population of *P. chichipe*, genetic diversity was estimated by standard genetic parameters: observed alleles ($A_O$), observed heterozygosity ($H_O$), and expected heterozygosity ($H_E$) under the hypothesis of Hardy–Weinberg (HW) genotypic proportions (Nei 1978). Differences with respect to HW genotypic proportions were calculated by Wright's $F_{IS}$ (Weir & Cockerham 1984) and heterozygote deficiency was tested using Markov chain randomization (dememorization 1000, batches 100, iterations per batch 1000) to estimate significance test ($U$-test) using GENEPOP (Raymond & Rousset 1995), with Bonferroni corrections. Differences

**Table 1** *Polaskia chichipe* microsatellite loci, primer sequences, number of alleles ($n$), and range of allele size

| Primer name | Primer sequence (5′–3′) | $n$ | Size | Fluorescence | Annealing temperature |
|---|---|---|---|---|---|
| *Pchi9** | GTGGCCGAGAAAGAAGTTTG<br>AAAGGCCCAAATCATAAGCA | 6 | 208–218 | 6FAM | 60 °C |
| *Pchi21** | CGTTTAGCCCCTCTTTCTCC<br>GTTCCCAACTGACCGACAAC | 8 | 120–134 | HEX | 60 °C |
| *Pchi54** | CCTTGAGCTTTGACATTGAGA<br>GGAAGGTTTTCATTGGATGAG | 11 | 158–182 | HEX | 60 °C |
| *Pchi20* | GTGGCCGAGAAAGAAGTTTG<br>AAAGGCCCAAATCATAAGCA | 9 | 230–256 | 6FAM | 56 °C |
| *Pchi50* | CCTGGGCAAACCTCTGTTTA<br>GTTCCCAACTGACCGACAAC | 8 | 216–230 | TET | 56 °C |

*Primer designations as in Otero-Arnaiz *et al*. (2004).

in expected heterozygosity ($H_E$) and inbreeding coefficient ($F_{IS}$) among populations and management type were tested with a Wilcoxon rank test performed in SIMSTAT version 2.06 (Péladeau 1996). The number of exclusive alleles was determined by inspection of population allele distributions.

The Wright's fixation indices (Wright 1965) were estimated for all populations and pooled within and among management types using SPAGEDI 1.1 (Hardy & Vekemans 2002). Significance levels of the statistics were determined with 10 000 permutation tests. These parameters are estimates of the correlations of genes within individuals relative to populations ($F_{IS}$) and of individuals and populations relative to all populations or the species ($F_{IT}$ and $F_{ST}$, respectively). Additionally, an equivalent estimate of $F_{ST}$ ($R_{ST}$) based in allele size differences was calculated using the program RST22 (Goodman 1997). Because positive values of $F_{IS}$ indicate deviations from HW genotypic proportions that can result by self-pollination or mating among relatives, we expected more positive values in manipulated than in wild populations.

There is no consensus yet on whether $F_{ST}$- or $R_{ST}$-based approaches are better for handling microsatellite data. Both have disadvantages. While $R_{ST}$ overestimates differentiation if microsatellites mutate by large insertions or deletions (Di Rienzo *et al*. 1994), $F_{ST}$ estimators can underestimate differentiation if mutation is largely stepwise (Slatkin 1995). We compared $R_{ST}$ and $F_{ST}$ values using a simple testing procedure proposed by Hardy *et al*. (2003) based on allele size randomizations under the null hypothesis of no contribution of stepwise mutation to genetic differentiation $R_{ST} = F_{ST}$. The tests compare $R_{ST}$ computed after allele size permutation ($pR_{ST}$), which depends solely on allele identity/nonidentity with the value computed before randomization. The expectation is that $pR_{ST}$ should equal $F_{ST}$ computed on the same data. The test statistic for accepting or rejecting the null hypothesis is based on the proportion of $pR_{ST}$ values larger than the observed $R_{ST}$ (one-tailed test). Single locus and multilocus global $F_{ST}$ (Weir & Cockerham 1984) and $R_{ST}$ (Michalakis & Excoffier 1996) were estimated, as well as $pR_{ST}$ after 10 000 permutations with the program SPAGEDI (Hardy & Vekemans 2002).

### *Relationship between geographical and genetic distance*

Isolation by distance was tested using the method developed by Rousset (1997). This method uses estimates of $F_{ST}/1 - F_{ST}$ for pairs of populations, and is based on the expected increase of genetic differentiation between pairs of populations according to geographical distance. Paired $F_{ST}/1 - F_{ST}$ estimations for the nine populations were calculated with SPAGEDI 1.1 (Hardy & Vekemans 2002) and then a regression analysis of these estimates against the logarithm of the Euclidean distance was performed. Geographic distance among pairs of population was measured as the straight-line distance between the estimated population centres. The significance of isolation by distance was tested by a randomization process where populations were permuted among locations 10 000 times, which provided the null hypothesis of no correlation among geographical and genetic distances. *P* values were estimated as the proportion of this distribution falling above the observed regression slope.

In order to visualize the genetic relationship among populations, a matrix of genetic distances (Nei 1972) was calculated and a clustering dendrogram was constructed using the UPGMA method with the program TFPGA (Miller 1997). Confidence levels for the phenogram were constructed by bootstrapping the original data 10 000 times with replacement over all loci.

### *Reductions in effective population size*

Population bottlenecks can be estimated with a combination of microsatellite markers and novel statistical methods (Cornuet & Luikart 1996; Piry *et al*. 1999; Maudet *et al*. 2002). Populations that have experienced a recent reduction of their effective population size (between $2N_e$ and $4N_e$ generations, according to Piry *et al*. 1999) exhibit a correlative reduction of the allele numbers ($A_O$) and genetic diversity ($H_E$) at polymorphic loci. The number of alleles is reduced faster than genetic diversity, therefore in a recently bottlenecked population the observed genetic diversity is higher than the expected equilibrium genetic diversity ($H_{eq}$) computed from the observed number of alleles ($A_O$), under the assumption of mutation-drift equilibrium (Cornuet & Luikart 1996).

Detection of bottlenecks in *P. chichipe* populations was conducted using the two-phased model of mutation (TPM). Because few loci follow the stepwise-mutation model (SMM or strict single-step mutations), tests for bottleneck were performed with 95% of SMM and 5% of multiple-step mutation and a variance of 12 among multiple steps, as recommended by Piry *et al*. (1999) for microsatellite markers. A Wilcoxon sign rank test was performed using the BOTTLENECK software (Cornuet & Luikart 1996) to determine whether populations exhibited a significant number of loci with genetic diversity excess.

## Results

A total of 42 alleles were detected at the five microsatellite loci analysed. Number of alleles per locus ranged from six to 11, averaging 8.4 alleles per locus (Table 1).

### *Mating patterns*

Mating system parameters did not differ significantly among populations or management types (Table 2). Multilocus

**Table 2** Proportion of outcrossed progeny ($t_m$), maternal fixation index ($f$), biparental inbreeding ($bf$), and proportion of individuals with exclusively outcrossed progeny ($T$) for six populations of *Polaskia chichipe* under different types of management in Tehuacán Valley, Mexico

| Population | $N$ | $t_m$ | $f$ | $bf$ | $T$ |
|---|---|---|---|---|---|
| Wild | | | | | |
| W1 | 9/148 | 0.898 (0.06) | 0.000 (0.00) | 0.064 (0.04) | 0.22 |
| W2 | 11/187 | 0.949 (0.06) | 0.005 (0.05) | 0.060 (0.05) | 0.54 |
| Silvicultural | | | | | |
| M1 | 13/151 | 0.890 (0.07) | 0.008 (0.06) | 0.090 (0.06) | 0.42 |
| M3 | 7/110 | 0.823 (0.08) | 0.000 (0.04) | 0.005 (0.06) | 0.143 |
| Cultivated | | | | | |
| C1 | 14/251 | 0.890 (0.07) | 0.007 (0.07) | 0.102 (0.06) | 0.29 |
| C2 | 10/120 | 0.979 (0.04) | 0.060 (0.07) | 0.066 (0.05) | 0.22 |
| Average for management type | | | | | |
| Wild | 20/335 | 0.924 (0.06) | 0.003 (0.02) | 0.062 (0.04) | 0.40 |
| Silvicultural | 20/261 | 0.857 (0.08) | 0.004 (0.05) | 0.047 (0.06) | 0.32 |
| Cultivated | 24/371 | 0.935 (0.06) | 0.034 (0.07) | 0.084 (0.05) | 0.26 |

estimates of the proportion of outcrossing progeny ($t_m$) ranged from 0.82 to 0.98 and averaged 0.90 (± 0.09 SE). Biparental inbreeding ($bf$) ranged from 0.005 to 0.102 and averaged 0.064 (± 0.07 SE).

On average, cultivated populations had the highest estimate of $t_m$ (0.935 ± 0.06 SE) and silvicultural populations the lowest (0.857 ± 0.08 SE). Cultivated populations also had the highest estimates of biparental inbreeding ($bf$ = 0.084 ± 0.05 SE). The proportion of individuals with exclusively outcrossed offspring ranged among 0.14 and 0.55. On average, cultivated populations had the lowest proportion of outcrossed offspring (0.26), wild populations had the highest (0.40) and the silvicultural populations had an intermediate proportion (0.32).

### *Genetic variation within and between populations and management types*

Out of the 264 individual plants sampled, a total of 260 different genotypes were identified, indicating that at most 2% of the individuals sampled could be clones. Vegetative propagation, through planting branch cuttings, has been reported by local people (Carmona & Casas 2005). One identical genotype was observed in two individuals of the silvicultural population M3, and this genotype was also observed in one individual from the wild population W1. Two other individuals of the cultivated population C3 and one of population C1 shared an identical genotype. Similarly, populations M2 and W2 shared two identical genotypes.

All five loci had one common allele and one or more relatively rare alleles. At three loci the most common allele was the same in all nine populations. At *Pchi21* and *Pchi50* the most common allele was different in populations C1 and W3, respectively. Data on allele frequencies are available by request from the author for correspondence.

Levels of diversity ($H_E$) and allelic richness ($A_O$) did not differ significantly ($P < 0.05$) between populations. Genetic variation ($H_E$) across all populations varied from 0.559 ± 0.094 SE in population M3 to 0.726 ± 0.069 SE in population W3 (Table 3). Population M1 with two unique alleles had the highest mean observed number of alleles per locus (6.8). Whereas cultivated populations, each with a unique allele, had 5.2–6.4 alleles per locus. The average genetic variation ($H_E$) within populations was higher in wild populations (0.685 ± 0.16 SE) than in cultivated populations (0.660 ± 0.132 SE), and the lowest occurred in silviculturally managed populations (0.621 ± 0.248 SE). Values of $H_O$ in all populations were lower than $H_E$ values indicating a deficiency of heterozygous individuals with respect to HW expectations. This deficit was significant in three

| Population | Sample size | $A_O$ | $H_O$ | $H_E$ | $F_{IS}$ |
|---|---|---|---|---|---|
| W1 | 30 (30) | 6.2 ± 1.2 | 0.587 ± 0.08 | 0.673 ± 0.09 | 0.130* |
| W2 | 28.4 (27–29) | 5.8 ± 0.7 | 0.613 ± 0.08 | 0.635 ± 0.08 | 0.036 |
| W3 | 29.4 (28–30) | 5.8 ± 0.7 | 0.693 ± 0.06 | 0.726 ± 0.07 | 0.048 |
| M1 | 27.4 (24–30) | 6.8 ± 0.9 | 0.590 ± 0.09 | 0.668 ± 0.10 | 0.119 |
| M2 | 27.6 (25–29) | 4.6 ± 0.9 | 0.439 ± 0.10 | 0.586 ± 0.11 | 0.254* |
| M3 | 22 (18–26) | 4.4 ± 0.9 | 0.493 ± 0.08 | 0.559 ± 0.09 | 0.121 |
| C1 | 28.4 (24–30) | 6.4 ± 0.9 | 0.608 ± 0.07 | 0.651 ± 0.08 | 0.066 |
| C2 | 29 (27–30) | 6.2 ± 0.7 | 0.587 ± 0.05 | 0.674 ± 0.06 | 0.131 |
| C3 | 28 (28–30) | 5.2 ± 0.6 | 0.488 ± 0.07 | 0.631 ± 0.07 | 0.229* |
| Wild | 87.8 (85–89) | | 0.631 ± 0.03 | 0.683 ± 0.04 | |
| Silvicultural | 77 (70–85) | | 0.507 ± 0.05 | 0.621 ± 0.05 | |
| Cultivated | 85.4 (75–90) | | 0.560 ± 0.03 | 0.660 ± 0.04 | |

Exact test for heterozygote deficit using a Markov chain (dememorization 10 000, batches 100, iteration per batch 1000). *$P < 0.001$.

**Table 3** Genetic variation for nine *Polaskia chichipe* populations under different types of management in the Tehuacán Valley, Mexico. Populations are numbered according to the territory of the municipality that they belong: Atolotitlan (1), Caltepec (2) and Mentzontla (3) for wild (W), silvicultural (M) and cultivated populations (C). Mean sample size (and minimal and maximal values), mean observed number of alleles per locus ($A_O$), mean observed and expected heterozygosities ($H_O$ and $H_E$), with their SE and a measure of heterozygote deficiency ($F_{IS}$) are shown

**Table 4** Multilocus estimates of hierarchical *F*-statistics and their significance after 10 000 permutation tests between populations under the same management type, among different management types, and pooled within management type for all populations of *Polaskia chichipe* studied in the Tehuacán Valley, Mexico

| Management type | $F_{IS}$ | $F_{IT}$ | $F_{ST}$ | $R_{ST}$ |
|---|---|---|---|---|
| Wild | 0.071** | 0.080** | 0.009* | 0.001 |
| Silvicultural | 0.173*** | 0.190*** | 0.022** | 0.0189* |
| Cultivated | 0.142*** | 0.157*** | 0.017** | 0.0183* |
| Among different management types | 0.135*** | 0.140*** | 0.005* | |
| Polaskia chichipe | 0.126*** | 0.140*** | 0.015*** | 0.009* |

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

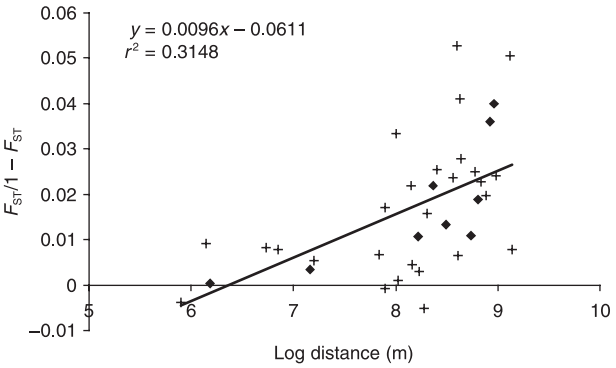


**Fig. 2** Relationship between pairwise comparisons of genetic and log geographical distance. Multilocus estimates of $F_{ST}/1 - F_{ST}$ are plotted against the Euclidian distances. Crosses (+) illustrate comparisons between populations under different management, and diamonds (◆) were used for comparisons between populations under the same management type.

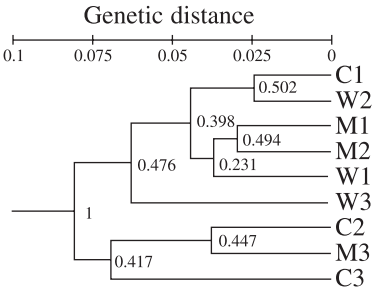


**Fig. 3** UPGMA phenogram based on Nei's (1972) genetic distances estimated for nine populations of *Polaskia chichipe* under different levels of management in the Tehuacán Valley, Mexico. Proportions of similar replicates after 10 000 permutations are shown for each grouping.

populations after Bonferroni corrections (Table 3). Although positive $F_{IS}$ values were observed at most of the loci in each population, not all were significant. In populations W1 and C3, $F_{IS}$ was significant at locus *Pchi50*. Significant values for $F_{IS}$ were also found at *Pchi21* and *Pchi50* in population M2. When populations under a similar management type were analysed together, inbreeding coefficients were significantly different from zero in silvicultural and cultivated populations after a randomization test (*U*-test). The $F_{IS}$ in silvicultural populations was higher than in cultivated populations, and the lowest $F_{IS}$ was observed in wild populations (Table 4). Differentiation among populations was low but significant ($F_{ST} = 0.015$; $P < 0.001$) with a significant contribution caused by management type ($F_{ST} = 0.005$; $P < 0.04$). The strongest differentiation among populations under similar management type was among silvicultural and cultivated populations ($F_{ST} = 0.022$ and 0.017, respectively), wild populations were the least different among themselves ($F_{ST} = 0.009$, $P < 0.04$). Differentiation based on allele sizes ($R_{ST}$) is shown in Table 4 for comparisons with other studies. However, $F_{ST}$ seems to be a better estimate of genetic differentiation in *Polaskia chichipe* because permutation test did not detect any $R_{ST}$ value significantly larger than $pR_{ST}$. When each population was analysed for genetic bottleneck under TPM with 95% SMM, none were significant after 5000 iterations ($P < 0.05$).

### *Relationship between geographical and genetic distance*

The most genetically similar populations were C1 and W2 ($F_{ST} = -0.005$, $R_{ST} = -0.001$) and the most different were M2 and C3 ($F_{ST} = 0.051$, $R_{ST} = 0.061$). There was a positive correlation between genetic differentiation ($F_{ST}/1 - F_{ST}$) and log geographical distance. This correlation was statistically significant, with a significant $b = 0.0096$ after 10 000 permutations of locations among populations ($P = 0.007$, Fig. 2). This result indicates restricted gene flow and a pattern of isolation by distance among *P. chichipe* populations (Rousset 1997). There was high sampling variance inherent to pairwise $F_{ST}/1 - F_{ST}$ ratios. Geographic distance among populations explained only 31.4% of the total genetic variation. The slope of the regression was 0.009, indicating an indirect estimation of a neighbourhood size of 104 individuals (from $b = 1/4N\pi\sigma^2$, see Rousset 1997). In Fig. 2, the regression coefficient for $F_{ST}/1 - F_{ST}$ ratios was higher between pairs of populations under the same management type ($r^2 = 0.57$, $P = 0.13$) with respect to the ratios between populations under different management type ($r^2 = 0.27$, $P = 0.02$).

Bootstrap analysis of the phenogram provides strong support ($P < 0.001$) for differences among populations according to their spatial distribution (Fig. 3). Two groups of populations were clustered together. Populations C3, M3 and C2, which occupy the western side of the study area, comprise a first group, and the rest of the populations occupying the eastern side of the area were clustered in a second group. In the second group, the most similar

populations were C1 and W2. Populations M1, M2 and W1 occurred at the centre of the sampling area and formed a subgroup. Population W3, also in the second group, was the population with the highest genetic diversity ($H_E = 0.73$).

## Discussion

The process of domestication is expected to result in significant reductions of genetic variation, as has been observed in most species with signs of advanced levels of domestication (Doebley 1992). Considering the patterns of management in *Polaskia chichipe* that involve tolerating or propagating desirable individuals while clearing non-desirable individuals, we expected to find a reduction of genetic variation in manipulated populations compared to wild populations. Such reduction was expected to be stronger in silviculturally managed populations because artificial selection involves the selective removal of part of a population when clearing vegetation. However, in cultivated populations, genetic diversity could be enhanced by people introducing new variants from local sources or other towns, as has been documented for *Stenocereus stellatus* by Casas *et al*. (2001). Results of our study of *P. chichipe* are not consistent with this pattern. Although the highest genetic variation was found in wild populations, with lower variation in silviculturally managed and cultivated populations, there were no significant differences in the levels of genetic variation among management types.

Low genetic differentiation among populations can result from gene flow that potentially slows the domestication process of *P. chichipe* by mixing up genotypes and preventing diversification caused by artificial selection. This, however, does not preclude morphological differences observed among populations under different management type (Carmona & Casas 2005) and artificial selection favouring a higher proportion of individuals with self-pollination (Otero-Arnaiz *et al*. 2003). Such low genetic differentiation could be caused by the sympatry of the populations. *P. chichipe* is a species restricted to a small region with volcanic soils and the primary flower visitors can potentially travel distances further than that among populations (0.4–9.4 km); *Apis melifera* may fly within a radius of nearly 2.5 km (Metcalf & Flint 1974) and hummingbirds may cover distances > 5.5 km /d (Arizmendi 2001). However, the pattern of foraging activity is more frequently restricted to short distances (Otero-Arnaiz *et al*. 2003) limiting migration to neighbouring populations.

Characteristics of the species, such as long-lived perennial habit and limited vegetative propagation, increase the time of fixation of desirable characters in a population through artificial selection, as has been postulated for ancient domesticated crops (Zohary & Hopf 2000). Within populations, artificial selection can be delayed because the cacti are slow to reach maturity. People tolerate or eliminate individual cacti based on their ability to distinguish the desirable characteristics in the plants which does not occur until after fruit production (probably more than 10 years after seed germination, according to Carmona & Casas 2005).

Genetic variation found in *P. chichipe* with five microsatellite loci is within the range of microsatellite variation reported for other species of long-lived plants with mixed-mating system (0.64–0.85 in some trees according to Byrne *et al*. 1996; Chase *et al*. 1996; Aldrich *et al*. 1998). An excess of homozygosity within populations can be due to at least two causes: inbreeding and population subdivision. Positive $F_{IS}$ values found in all populations indicate potential levels of inbreeding that are inconsistent with the mostly outcrossed mating system. However, mating with close relatives could be common, biparental inbreeding contributed significantly to selfing rates. An apparent deficiency of heterozygotes could result from Wahlund effects (Hartl & Clark 1997). If selfing occurs, all loci should have heterozygous deficiencies and similar $F_{IS}$, which was not the case with *P. chichipe*. Our data are not sufficient to resolve whether there was substructure within our populations that could result in a Wahlund effect. However, in a previous study of local substructure and gene flow in *P. chichipe*, we found evidence to support within population structure as a cause of heterozygote deficiency (Otero-Arnaiz 2004).

*Polaskia chichipe* is pollinated by several species of insects with foraging activity frequently restricted to short flights among flowers on one or few individuals. *P. chichipe* is also pollinated by hummingbirds, which may fly longer distances than bees but commonly fly short distances compared with bats, the primary pollinator of other species of columnar cacti (Otero-Arnaiz *et al*. 2003). The possibility to produce seeds by self-pollination in *P. chichipe* appears to favour inbreeding associated to the pattern of movement of pollinators among nearest neighbours, but both the selfing and inbreeding coefficient do not vary among populations. The proportion of individuals producing selfed progenies ($1 - T$, Table 2) varied among populations regardless of management type. The highest value was observed in M3 population. When values were averaged by management type, cultivated populations had the highest proportion of selfed progeny and wild populations the lowest. This result is consistent with our previous study on reproductive biology of *P. chichipe* in which we found that self-pollination treatments are successful in more individuals of cultivated and silvicultural populations than in wild populations (Otero-Arnaiz *et al*. 2003).

Although frequency of vegetative propagation was at most 2%, presence of clones in different populations suggests routes of recent dispersal by humans within populations (in M3 and C2) and between populations (C2-C1, W1 to M3, W3 to M2). Reproduction by seed seems to

be the primary form of propagation under natural conditions. This favours gene flow among neighbouring populations because fleshy fruits of *P. chichipe* are a source of food and water for a variety of birds and bats that act as seed dispersers. Migration between neighbouring populations, as suggested by the isolation-by-distance analysis, contributed to homogenization of genetic variation among nearby populations. Even if migration occurs preferentially among populations under the same management type, as was shown with crossing experiments (Otero-Arnaiz *et al*. 2003), neighbouring populations under different management types are less differentiated than more distant populations under the same management type because gene flow counteracts population subdivision on a local level. This observation is supported by the fact that distance explained the 57% of differentiation among populations under the same management type. When populations under different forms of management were analysed, distance explained only 27% of the variation. Other factors, such as artificial selection or inbreeding, might have contributed to differentiation of populations under different management type.

Bottlenecks associated with the domestication process in *P. chichipe* would be expected to be stronger under silvicultural management, intermediate in cultivated population and weak or nonexistent in wild populations. Changes resulting from silvicultural management in which some individuals are deliberately eliminated could cause significant reductions in population size. This type of management has the potential to create genetic bottlenecks in which effective population size ($N_e$) is reduced. The degree of severity of the potential bottleneck created by silvicultural management type will depend upon the genetic variation existing in the stand of plants maintained and it will be less severe if supplementary sources of pollen or seeds significantly contribute to the regeneration of the population. We did not detect recent bottleneck reductions in effective size of populations under any management type. Our results suggest that populations manipulated by humans have not suffered genetic bottlenecks as a result of the constant migration among populations. These results could also mean that bottlenecks occurred in a too distant past to be detected by our analyses, or that sampled population sizes were too small. Alternatively, the observed pattern could result from directional gene flow from wild to manipulated populations, avoiding effects of size reductions of manipulated populations. To test these hypotheses, it is necessary to estimate patterns and directionality of pollen and seed flow among populations under different management type.

Although *P. chichipe* has a long history of use and management that has resulted in morphological differentiation, changes in breeding system and seed germination patterns associated to human management, this is reflected only slightly in genetic differences among populations as detected by neutral markers. Migration and population history appear to have been strongly counteracting the effects of artificial selection within this species.

## Aknowledgements

We thank R. Petit, the subject editor, for all the helpful comments that improved the paper. The work was financially supported by the DGAPA, UNAM (IN 224799), SEMARNAT/CONACYT (Grant 2002–0544), and by the Vavilov Frankel Award 2001 given to A. O-A for research conducted at the University of Georgia. We also thank E. Pérez-Negrón, E. Arellano and S. Rangel for fieldwork assistance, E. González, F. Jiménez, M. Cruz and D. Lugo for their help with laboratory work, M. Mackenzie and D.E. Walker for their suggestions regarding microsatellite genotyping and scoring, and H. Ferreira for computer-work assistance. We thank D. Piñero, A. Valiente, L. Eguiarte, E. Alvarez, and M. Quesada for the review of the preliminary versions of the manuscript, as well as the generosity of the people of San Luis Atolotitlán, Caltepec and Metzontla, Puebla, Mexico.

## References

Aldrich PR, Hamrick JL, Chavarriaga P, Kochert G (1998) Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. *Molecular Ecology*, **7**, 933–944.

Arizmendi MC (2001) Multiple ecological interactions: nectar robbers and hummingbirds in a highland forest in Mexico. *Canadian Journal of Zoology*, **79**, 997–1006.

Byrne M, Marquez-Garcia MI, Uren T, Smith DS, Moran GF (1996) Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. *Australian Journal of Botany*, **44**, 331–341.

Carmona A, Casas A (2005) Management, phenotypic patterns and domestication of *Polaskia chichipe* (Cactaceae) in the Tehuacán Valley, central Mexico. *Journal of Arid Environments*, **60**, 115–132.

Casas A, Barbera G (2002) Mesoamerican domestication and diffusion. In: *Cacti: Biology and Uses* (ed. Nobel PS), pp. 143–162. California University Press, Berkeley, California.

Casas A, Caballero J, Valiente-Banuet A (1999) Use, management and domestication of columnar cacti in south-central Mexico: a historical perspective. *Journal of Ethnobiology*, **19**, 71–95.

Casas A, Valiente-Banuet A, Viveros JL (2001) Plant resources of the Tehuacán-Cuicatlán Valley, Mexico. *Economic Botany*, **55**, 129–166.

Casas A, Valiente-Banuet A, Caballero J (2002) Evolutionary trends in columnar cacti under domestication in south-central Mexico. In: *Columnar Cacti and their Mutualists: Evolution, Ecology, and Conservation* (eds Fleming TH, Valiente-Banuet A), pp. 137–163. University of Arizona Press, Tucson.

Chase M, Kesseki R, Bawa K (1996) Microsatellite markers for population and conservation genetics of tropical trees. *American Journal of Botany*, **83**, 51–57.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Darwin C (1883) *The Variation of Animals and Plants under Domestication*, 2nd edn. D. Appleton & Co., New York.

De Candolle A (1882) *Origin of Cultivated Plants* D. Appleton & Co., New York and London.

De la Cruz M, Ramírez MF, Hernández H (1997) DNA isolation and amplification from cacti. *Plant Molecular Biology Reporter*, **15**, 319–325.

Di Rienzo A, Peterson AC, Garza JC *et al*. (1994) Mutational processes at simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 3166–3170.

Doebley J (1992) Molecular systematics and crop evolution. In: *Molecular Systematic of Plants* (eds Soltis PS, Soltis DE, Doyle JJ), pp. 202–222. Chapman & Hall, New York.

Goodman SJ (1997) $R_{ST}$ CALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Molecular Ecology*, **6**, 881–885.

Hamrick JL, Nason JD, Fleming TH, Nassar JM (2002) Genetic diversity in columnar cacti. In: *Columnar Cacti and their Mutualists: Evolution, Ecology, and Conservation* (eds Fleming TH, Valiente-Banuet A), pp. 122–133. University of Arizona Press, Tucson.

Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyze spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hardy OJ, Charbonnel N, Fréville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, **163**, 1467–1482.

Harlan JR (1992) Origins and processes of domestication. In: *Grass Evolution and Domestication* (ed. Chapman GP), pp. 159–175. Cambridge University Press, Cambridge.

Hartl DL, Clark AG (1997) *Principles of Population Genetics*. Sinauer Associates, Sunderland, Massachusetts.

MacNeish RS (1967) A summary of subsistence. In: *The Prehistory of the Tehuacán Valley* (ed. Byres DS). University of Texas Press, Austin.

Maudet C, Miller C, Bassano B *et al*. (2002) Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [*Capra ibex* (*ibex*) ]. *Molecular Ecology*, **11**, 421–436.

Metcalf CL, Flint WP (1974) *Insectos Destructivos E Insectos Útiles*. Compañía Editorial Continental, Mexico City.

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, **142**, 1061–1064.

Miller MP (1997) *Tools for population genetic analyses (TFPGA) 1.3. A Windows program for the analysis of allozymes and molecular population genetic data*. Computer software distributed by author.

Nei M (1972) Genetic distance between populations. *American Naturalist*, **106**, 283–292.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Otero-Arnaiz A (2004) *Diferenciación t flujo génico entre poblaciones de* Polaskia chichipe *con distintos niveles de domesticación en zonas áridas del centro de México*. Tesis de Doctorado en Ciencias Biológicas, Universidad Nacional Autónoma de México.

Otero-Arnaiz A, Cruse-Sanders J, Casas A, Hamrick J (2004) Isolation and characterization of microsatellites in the endemic columnar cactus: *Polaskia chichipe* and cross-species amplification within the tribe Pachycereeae (Cactaceae). *Molecular Ecology Notes*, **4**, 265–267.

Otero-Arnaiz A, Casas A, Bartolo C, Perez-Negron E, Valiente-Banuet A (2003) Evolution of *Polaskia chichipe* (Cactaceae) under domestication in the Tehuacan Valley, central Mexico: reproductive biology. *American Journal of Botany*, **90**, 593–602.

Péladeau N (1996). SIMSTAT. Provalis Research, Montreal.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in effective population size from allele frequency data. *Journal of Heredity*, **90**, 502–503.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Ritland K (2002) Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity*, **88**, 221–228.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Slatkin M (1995) A measure of population subdivision based in microsatellite allele frequencies. *Genetics*, **139**, 157–162.

Weir B, Cockerham C (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wright S (1965) The interpretation of population structure by *F*-statistics with special regard to systems of mating. *Evolution*, **19**, 395–420.

Zohary D (1984) Modes of evolution in plants under domestication. In: *Plant Biosystematics* (ed. Grant WF), pp. 579–586. Academic Press Canada, Montreal.

Zohary D, Hopf M (2000) *Domestication of Plants in the Old World*. Oxford Science Publications, Oxford.

---

A. Otero-Arnaiz is a postdoctoral researcher in plant population genetics and reproductive systems. She is involved in a research team studying the evolution of pollen competition in Acacia species. She is also interested in the conservation and domestication of plants.

---