*Acta Botanica Brasilica* 33(4): 664-671. October-December 2019.
doi: 10.1590/0102-33062019abb0080

# Genetic diversity in *Calibrachoa pygmaea* (Solanaceae): A hawkmoth-pollinated nightshade from the Pampas

**Geraldo Mäder[1,2], Alice Backes[1], Maikel Reck-Kortmann[1] and Loreta B. Freitas[1,2*]**

Received: March 12, 2019
Accepted: May 21, 2019

## ABSTRACT

*Calibrachoa pygmaea* is a unique species of *Calibrachoa*, especially concerning its flower morphology and the environment where it occurs. The species is self-incompatible and is narrowly distributed in wet and flooded fields of the Pampas region. We characterize the genetic diversity of the species based on traditional plastid markers and newly developed nuclear microsatellites to identify drivers that guide its evolution. Our results identified markers that are informative and useful for studying the population structure of *C. pygmaea*, as well as that of other species of *Calibrachoa*. Both marker sets were congruent in developing conclusions regarding the evolutionary scenario of *C. pygmaea*, and revealed that the genetic variability and population structure of the species could be explained by common allele fixation or shared ancestral polymorphism, while its diversification can be attributed mainly to the species' dispersal ability and certain ecological features.

**Keywords**: cross-amplification, gene flow, genetic diversity, plastid haplotypes, population genetics, SSR markers, wild petunias

## Introduction

*Calibrachoa* (Solanaceae), related to the worldwide known genus *Petunia* (Ando *et al.* 2005; Kulcheski *et al.* 2006), contains approximately 26 species divided into two subgenera, *Calibrachoa* subg. *Calibrachoa* and *Calibrachoa* subg. *Stimomphis* (Fregonezi *et al.* 2012). *Petunia* and *Calibrachoa* differ in chromosome number with diploid number 2n=14 and 2n=18, respectively. *Calibrachoa* species are distributed in southern South America, mainly in Brazil, Argentina, and Uruguay, between parallels 18° and 37°S (Greppi *et al.* 2013). The two species of *Calibrachoa* subg. *Calibrachoa* (*C. parviflora* and *C. pygmaea*) are herbaceous with an annual life cycle, whereas the other *Calibrachoa* species are usually small perennial shrubs (Fregonezi *et al.* 2013). All species maintain intercrossing capacity, except inter-subgenera (Watanabe *et al.* 1997). Species in the genus are used for breeding programs as pot plants and landscaping since 1990 (Rice 1997), primarily because of the vast amount of flowers produced per plant and enormous color variety (Kanaya *et al.* 2010).

*Calibrachoa pygmaea* occurs predominantly in the Pampean province, specifically in open and wet fields from southernmost Brazil, northwest Uruguay, and northeast Argentina, distributed in small patches of just a few individuals (Fig. 1A). This species belongs to *Calibrachoa* subg. *Calibrachoa* and its unique, in the genus, white hypocrateriform corolla (Fig. 1B, C) combined with other floral attributes are indicatives of hawkmoth-pollination (Fregonezi *et al.* 2012). Its strongly scented flowers open at dusk and keep closed during the day, with a change in

**1** Laboratório de Evolução Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul, 91501-970, Porto Alegre, RS, Brazil
**2** Programa de Pós-Graduação em Botânica, Universidade Federal do Rio Grande do Sul, 90509-900, Porto Alegre, RS, Brazil

* Corresponding author: loreta.freitas@ufrgs.br

the corolla color from white to pale yellow (AL Cazé unpubl. res.). It is self-incompatible and it has been suggested a robust and efficient interaction with its pollinator in nature (Tsukamoto *et al.* 2002).

Since mating and pollination systems have a strong influence on genetic diversity and population structure (Ghazoul 2005; Blambert *et al.* 2016), one can expect that species with small and fragmented populations would have high genetic structure and low gene flow among populations (Ellstrand & Elam 1993). Our main objective here was to test the above predictions using *C. pygmaea* as a model. To do this, we employed classical plastid markers and developed new nuclear microsatellite markers seeking to describe the genetic diversity and population structure of this species.

## Materials and methods

### Plant material and collection sites

We collected 73 individuals from three populations of *Calibrachoa pygmaea* (R.E. Fr.) Wijsman (Tab. 1) located in Pampas biome in Rio Grande do Sul State, Brazil (Fig. 1A). These three populations are more than 50 km distant from each other. We also included one individual of *C. pygmaea* collected at geographic coordinates 30°38'75''S 56°45'14''W, deposited voucher BHCB 79881 at the Universidade Federal de Minas Gerais herbarium. This individual was used to prepare genomic libraries for microsatellite development.

**Table 1.** Sampling information. RS – Rio Grande do Sul Brazilian state; BHCB – Herbarium of Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil; N – number of individuals.

| Population Code | Municipality | Geographical Coordinates | Voucher | N |
|---|---|---|---|---|
| Pop 1 | Uruguaiana, RS | 30°01'10"S 56°24'51"W | BHCB 102105 | 25 |
| Pop 2 | Alegrete, RS | 29°47'17"S 55°43'35"W | BHCB 102099 | 23 |
| Pop 3 | Santana do Livramento, RS | 30°32'58"S 56°07'06"W | BHCB 156811 | 25 |

We selected these populations because they are distributed in, or very close to the region where the typus and lectotypus of the species were collected (Stehmann 1999). All collected individuals presented the canonical morphology of the species. The few other populations found were too small to be used.

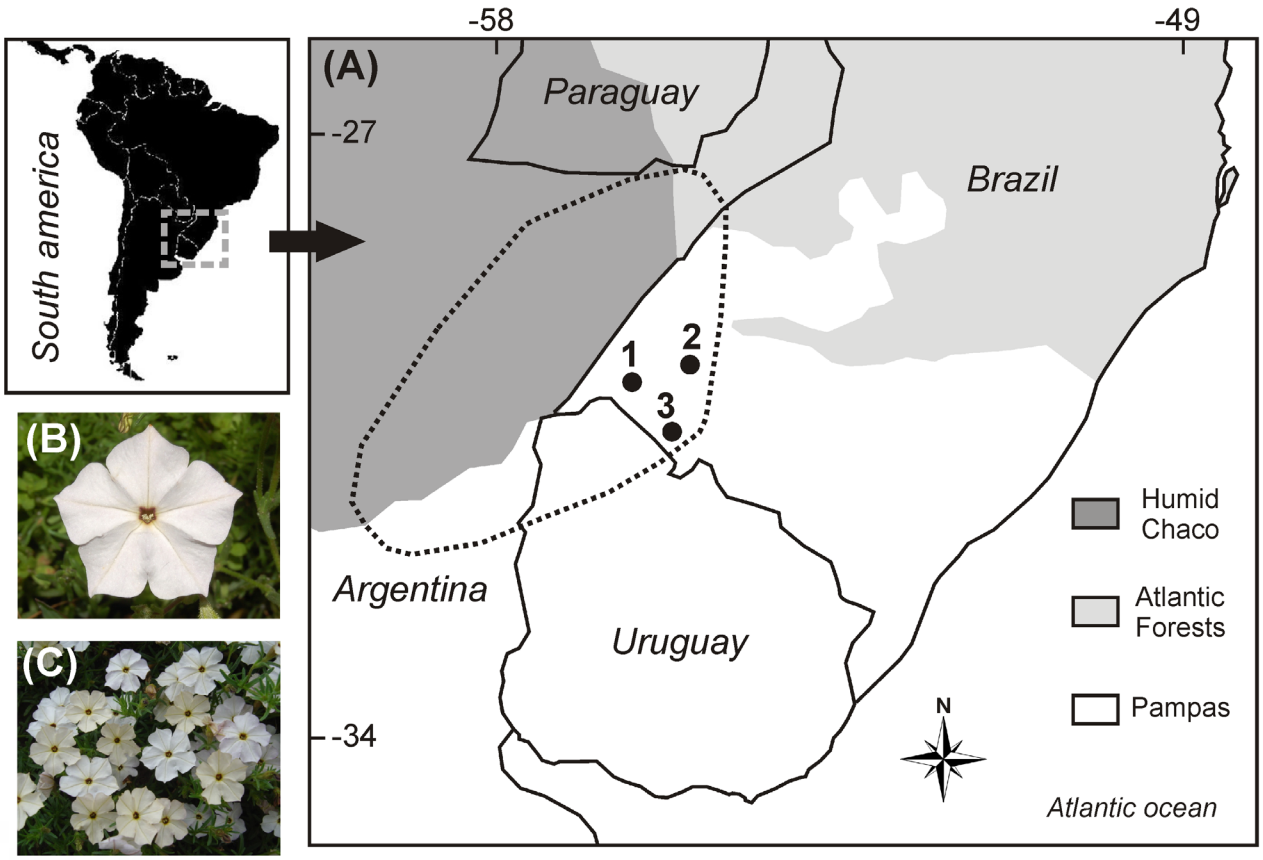


**Figure 1.** Analyzed *Calibrachoa pygmaea:* **A.** Collection sites; **B.** Flower frontal view; **C.** Individual overview highlighting massive flowering and change in flower color throughout the day. The dashed line corresponds to the geographical distribution of species according to herbarium records.

## DNA extraction

Four or five mature and young leaves were collected per individual, preserved in silica gel and subsequently powdered in liquid nitrogen. Genomic DNA was extracted with a cetyltrimethyl ammonium bromide (CTAB; Sigma-Aldrich, St. Louis, USA)-based protocol (Roy *et al.* 1992). DNA quality and quantity were evaluated by measuring the absorbance at 260 and 280 nm on a Nanodrop spectrophotometer (Thermo Scientific Corp., San Jose, USA). DNA was diluted to uniform concentration and stored at -20 °C until use.

## Microsatellite development and genotyping

An enriched library methodology was applied to isolate specific repeat motifs according to the protocols of Kriedt *et al.* (2011) and Silva-Arias *et al.* (2015) and design primers for successful loci. We also evaluated the transferability of these microsatellites loci to 13 *Calibrachoa* and two *Petunia* species (Tab. S1 in supplementary material), testing in three individuals per species the same primers and same conditions employed for *C. pygmaea*.

PCR amplifications were conducted in a final volume reaction of 10 μL containing 10 ng of genomic DNA as a template, 200 μM of each dNTP (Invitrogen, Carlsbad, USA), 1.7 pmol of fluorescently labelled M13(-21) primer, 3.5 pmol of reverse primer, 0.35 pmol of forward primer with a 5′-M13(-21) tail, 2.0 mM $MgCl_2$ (Invitrogen), 0.5 U of Platinum Taq DNA polymerase (Invitrogen), and 1× platinum Taq reaction buffer (Invitrogen). The PCR conditions consisted of an initial denaturation of 3 min at 94 °C; 30-35 cycles 20 s at 94 °C, 45 s at 55–58 °C, and 1 min at 72 °C; and a final 10-min extension cycle at 72 °C. The primer sequences, repeat motifs, respective annealing temperatures, and fluorescent dyes used in each fragment are shown in Table S2 in supplementary material**.**

The amplified products were visualized on a 2.5 % ultra-resolution agarose gel (Invitrogen) stained with 2 μL 0.001 % GelRedTM (Biotium, Inc., Hayward, USA) to check quality and, posteriorly, denatured and size-fractionated using capillary electrophoresis on an ABI 3100 genetic analyzer (Thermo Fisher Scientific Co., Waltham, USA) with a LIZ (500) molecular-size standard (Thermo Fisher Scientific Co.). GeneMarker 1.97 software (Softgenetics LLC, State College, PA, USA) was employed to determine the alleles. Additionally, all alleles were visually verified and scored.

## Plastid DNA amplification and sequencing

The plastid sequences (cpDNA) of intergenic spacers *rpl32-trnL, trnH-psbA,* and *trnS-trnG* were amplified and sequenced for individuals of the three *C. pygmaea* population (10 from population 1, 13 from population 2, and 12 from population 3) using universal primers (Shaw *et al.* 2007; Hamilton 1999; Sang *et al.* 1997, respectively). The PCR reactions were performed in 25-μL final volumes comprised of 1 U Taq polymerase (Invitrogen), 1x Taq polymerase buffer (Invitrogen), 0.2 mM each dNTP (Invitrogen), 2 mM MgCl2, 0.2 μM of each primer, and 30 ng of genomic DNA. The amplification conditions used were previously described (Lorenz-Lemke *et al.* (2006) for *trnH-psbA* and *trnS-trnG,* and Shaw *et al.* (2007) for *rpl32-trnL*). The PCR products were checked by horizontal electrophoresis in 1 % agarose (Invitrogen) and gel dyed with GelRedTM (Biotium), then purified with polyethylene glycol 3550 (Sigma-Aldrich) 20 % (Dunn & Blattner 1987). Sequencing was performed with an ABI 3730XL sequencer (Thermo Fisher Scientific Co.).

## Nuclear microsatellite genetic diversity

We used Fstat 2.9.3.2 software (Goudet 1995) to estimate summary statistics such as allelic frequencies, the number of alleles per locus (A), gene diversity (GD), allelic richness (AR), and inbreeding coefficient (*FIS*) for each locus for *C. pygmaea*. As an evaluation of the information context of the loci, Cervus 3.0.3 (Marshall *et al.* 1998; Kalinowski *et al.* 2007) was used to estimate the percentage of null alleles, the polymorphic index content (PIC), the levels of observed (Ho) and expected (He) heterozygosity, and significant deviations from the Hardy–Weinberg equilibrium (HWE).

We performed locus-by-locus analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) and estimated the overall and the pairwise *FST*s to assess the partition of the genetic variation among populations using Arlequin 3.5 (Excoffier & Lischer 2010) with $10^4$ permutations.

Moreover, to investigate the genetic dissimilarity among *C. pygmaea* individuals, discriminant analysis of principal components (DAPC; Jombart *et al.* 2010) was carried out in Adegenet 2.1.1 (Jombart 2008) without prior information on the individuals' spatial origin. The best number of genetic groups was assessed through *K*-means clustering according to the Bayesian information criterion (BIC). The number of Principal Components (PCs) used was estimated during the calculation process.

The genetic structure of *C. pygmaea* was also examined with the Structure 2.3 package (Pritchard *et al.* 2000) run without sampling location prior information and with the admixture model with correlated allele frequencies (Falush *et al.* 2003). The number of clusters (*K*) was evaluated from 1 to 6, with 10 independent runs per *K*-value. Each run was conducted using 2.5 x 105 burn-in periods and 106 repetitions after burn-in. The best *K*-value was identified from the maximum value of Δ*K* (Evanno *et al.* 2005) as implemented in Structure Harvester 0.6 (Earl & Holdt 2012). Pophelper 10.10 (http://pophelper.com/) was used to summarize the output from the Structure runs and generate bar plots.

We used a distance matrix based on shared alleles among individuals and collection sites to depict relationships among all of the *C. pygmaea* individuals using Msa 4.05 software (Dieringer & Schlötterer 2003). We employed Phylip software (http://evolution.genetics.washington.edu/phylip.html) to construct an unweighted neighbor-joining tree (N-J; Saitou & Nei 1987) based on the matrix of shared alleles.

### Plastid genetic diversity

For each plastid marker, both forward and reverse strands were assembled via ChromasPro 1.7.5 (Technelysium Pty Ltd, S. Brisbane, Australia) and the final sequences deposited at GenBank (MK619694-MK619791; https://www.ncbi.nlm.nih.gov/genbank/). DNA sequences were aligned through Mega7 software (Kumar *et al*. 2016) and the Clustal W algorithm, with manual editing when necessary. Haplotype (*h*) and nucleotide ($\pi$) diversity indexes, the number of variable sites, and AMOVA were calculated using Arlequin. The three plastid intergenic spacers were concatenated and treated as a single sequence across all analyses. The numbers of variable and informative sites in the alignment were obtained in Mega. Haplotypes were identified using DNAsp 5.10.01 (Rozas *et al.* 2003) and evolutionary relationships among them were estimated based on the median joining network in Network 5 (http://www.fluxus-engineering.com). Tajima's *D* (Tajima 1989) and Fu's *FS* (Fu 1997) neutrality tests were performed to detect evidence for deviation from a neutral equilibrium model of evolution using Arlequin. The program Alleles in Space 1.0 (Miller 2005) was also applied to associate the genetic and geographic distances between the three collection sites according to Mantel's test (Mantel 1967). For this analysis, *log*-transformed geographic distance was utilized to compare the three populations.

## Results

### Microsatellite loci development and transferability

Thirteen loci with a bright single band per allele were identified and used to genotype the 73 individuals from three populations of *C. pygmaea*. All loci exhibited polymorphisms in this species, and all individuals presented one or two alleles (consistent with the diploid condition of *C. pygmaea*) with the expected sizes based on clone sequences.

Cross-amplification was positive for 10 loci (Tab. S1 in supplementary material) under the same PCR conditions used for *C. pygmaea*, indicating that the developed markers can be useful to study other *Calibrachoa* species. Only Cpy 15, Cpy 29, and Cpy 57 loci were amplified across all *Calibrachoa* species; Cpy 29 was able to amplify *P. axillaris* individuals and Cpy 29, Cpy 57, and Cpy 80 were positive in cases of *P. integrifolia*. The highest number of loci was amplified in *C. parviflora* (nine).

### Genetic diversity and population structure in ***C. pygmaea*** based on nuclear markers

The number of alleles per locus varied from three to 32, with an average of 16 for *C. pygmaea* species; the percentage of null alleles was less than 1% for all loci and populations; all loci were informative with generally high PIC values, and the three populations presented private alleles (Tab. 2). Further, all variability indices were similar among populations. The observed ($H_o$) and expected (He) heterozygosity ranged from 0.1 to 0.8 and 0.3 to 1.0, with averages of 0.5 and 0.8, respectively among populations. Inbreeding coefficients (*FIS*) ranged from 0.1 to 0.7 across loci for the species and the majority of these values were significant at $\alpha = 0.05$ (Tab. 3). The loci Cpy 57, Cpy 62, Cpy 80, Cpy 89, Cpy 102, and Cpy

**Table 2.** Genetic information for 13 microsatellite loci in *Calibrachoa pygmaea*. A – number of alleles; AR – allelic richness; PIC – polymorphic information content; GD – genetic diversity; Nul – null alleles (%); Priv – number of private alleles; species – total or mean of three populations.

| Locus | Pop1 | | | | | | Pop2 | | | | | | Pop3 | | | | | | Species | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | AR | PIC | GD | Nul | Priv | A | AR | PIC | GD | Nul | Priv | A | AR | PIC | GD | Nul | Priv | A | AR | PIC | GD |
| Cpy02 | 5 | 4.7 | 0.4 | 0.49 | 0.32 | 2 | 4 | 4.0 | 0.6 | 0.70 | 0.37 | 0 | 7 | 6.6 | 0.6 | 0.7 | 0.54 | 3 | 9 | 8.9 | 0.6 | 0.65 |
| Cpy06 | 5 | 4.8 | 0.5 | 0.54 | 0.10 | 0 | 6 | 5.9 | 0.6 | 0.63 | 0.31 | 0 | 8 | 7.5 | 0.6 | 0.6 | 0 | 2 | 8 | 8.0 | 0.6 | 0.60 |
| Cpy15 | 7 | 6.6 | 0.6 | 0.65 | 0.07 | 1 | 6 | 5.9 | 0.6 | 0.71 | 0.07 | 4 | 7 | 6.7 | 0.6 | 0.7 | 0.30 | 3 | 8 | 7.9 | 0.6 | 0.68 |
| Cpy24 | 10 | 9.3 | 0.6 | 0.69 | 0.36 | 6 | 7 | 6.8 | 0.5 | 0.57 | 0.72 | 2 | 10 | 9.3 | 0.7 | 0.7 | 0.34 | 5 | 13 | 12.8 | 0.7 | 0.67 |
| Cpy29 | 2 | 2.0 | 0.3 | 0.35 | 0.26 | 2 | 2 | 2.0 | 0.3 | 0.38 | 0.25 | 0 | 3 | 2.8 | 0.3 | 0.3 | 0.43 | 0 | 3 | 2.9 | 0.3 | 0.34 |
| Cpy54 | 5 | 4.8 | 0.6 | 0.67 | 0.08 | 2 | 6 | 5.8 | 0.7 | 0.76 | 0.04 | 1 | 8 | 7.6 | 0.7 | 0.8 | 0 | 1 | 8 | 7.9 | 0.7 | 0.74 |
| Cpy57 | 16 | 15.2 | 0.9 | 0.91 | 0.34 | 0 | 15 | 14.6 | 0.9 | 0.93 | 0.19 | 1 | 15 | 14.0 | 0.9 | 0.9 | 0.30 | 2 | 25 | 24.8 | 0.9 | 0.94 |
| Cpy58 | 9 | 8.5 | 0.8 | 0.82 | 0.08 | 0 | 8 | 7.9 | 0.7 | 0.76 | 0.07 | 1 | 9 | 8.5 | 0.8 | 0.8 | 0.19 | 3 | 13 | 12.8 | 0.8 | 0.83 |
| Cpy62 | 18 | 17.1 | 0.9 | 0.94 | 0.12 | 5 | 16 | 15.5 | 0.9 | 0.92 | 0.31 | 4 | 15 | 14.0 | 0.8 | 0.9 | 0.33 | 2 | 26 | 25.7 | 0.9 | 0.92 |
| Cpy80 | 19 | 17.8 | 0.9 | 0.95 | 0.36 | 2 | 10 | 9.9 | 0.8 | 0.89 | 0.74 | 1 | 13 | 13.0 | 0.9 | 0.9 | 0.72 | 3 | 22 | 22.0 | 0.9 | 0.93 |
| Cpy89 | 14 | 13.0 | 0.9 | 0.90 | 0.09 | 4 | 14 | 13.6 | 0.9 | 0.92 | 0.07 | 5 | 11 | 10.6 | 0.9 | 0.9 | 0.04 | 1 | 19 | 18.6 | 0.9 | 0.91 |
| Cpy122 | 15 | 14.6 | 0.9 | 0.92 | 0.27 | 7 | 17 | 16.5 | 0.9 | 0.94 | 0.08 | 2 | 16 | 15.4 | 0.9 | 0.9 | 0.39 | 2 | 24 | 23.8 | 0.9 | 0.94 |
| Cpy144 | 24 | 22.0 | 0.9 | 0.97 | 0.19 | 4 | 17 | 16.4 | 0.9 | 0.93 | 0.20 | 2 | 20 | 18.6 | 0.9 | 1.0 | 0.16 | 1 | 32 | 31.5 | 1.0 | 0.96 |
| Mean | 11.5 | 10.8 | 0.7 | 0.75 | - | 2.7 | 9.8 | 9.6 | 0.7 | 0.77 | - | 1.8 | 10.9 | 10.4 | 0.7 | 0.8 | - | 2.2 | 16.2 | 16.0 | 0.8 | 0.78 |

144 deviated from HWE for all populations (P < 0.05) after Bonferroni's correction, whereas only Cpy 58 was in HWE equilibrium across the three populations. All loci that deviated from HWE presented heterozygote deficits that suggested high levels of inbreeding, just as confirmed through *FIS* values.

The AMOVA among the *C. pygmaea* populations uncovered that higher molecular variance existed among individuals within populations (~97 %), with a low but significant *FST* (*FST* = 0.023, P < 0.001). Based on pairwise *F*-statistics, equally low and significant *FST* values were observed with more considerable distances obtained between populations 1 and 2 (*FST* = 0.03).

DAPC analysis showed two main groups based on *K*-means clustering. The two groups of individuals were quite different in the scatter plot (Fig. 2A). Despite the clear separation between groups, they encompassed similar numbers of individuals from the three populations, though not indicating population structure (Fig. 2B) and confirming the AMOVA results.

**Table 3**. Genetic diversity for 13 microsatellites loci developed for *Calibrachoa pygmaea*. He - expected heterozygosity; Ho - observed heterozygosity; FIS - inbreeding coefficient. In bold, significant values at α = 0.05.

| Locus | Pop1 | | | Pop2 | | | Pop3 | | | Species | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Ho | He | FIS | Ho | He | FIS | Ho | He | FIS | Ho | He | FIS |
| Cpy02 | 0.2 | 0.5 | 0.5 | 0.3 | 0.7 | 0.5 | 0.2 | 0.7 | 0.7 | 0.3 | 0.6 | 0.6 |
| Cpy06 | 0.6 | 0.5 | -0.1 | 0.3 | 0.6 | 0.5 | 0.6 | 0.6 | 0.0 | 0.5 | 0.6 | 0.1 |
| Cpy15 | 0.6 | 0.6 | 0.1 | 0.6 | 0.7 | 0.2 | 0.4 | 0.7 | 0.5 | 0.5 | 0.7 | 0.3 |
| Cpy24 | 0.3 | 0.7 | 0.5 | 0.1 | 0.6 | 0.9 | 0.4 | 0.7 | 0.5 | 0.3 | 0.7 | 0.6 |
| Cpy29 | 0.2 | 0.3 | 0.4 | 0.2 | 0.4 | 0.4 | 0.1 | 0.3 | 0.6 | 0.2 | 0.3 | 0.5 |
| Cpy54 | 0.6 | 0.7 | 0.2 | 0.7 | 0.8 | 0.1 | 0.8 | 0.8 | 0.0 | 0.7 | 0.7 | 0.1 |
| Cpy57 | 0.4 | 0.9 | 0.5 | 0.6 | 0.9 | 0.3 | 0.5 | 0.9 | 0.5 | 0.5 | 0.9 | 0.5 |
| Cpy58 | 0.7 | 0.8 | 0.2 | 0.7 | 0.8 | 0.1 | 0.6 | 0.8 | 0.3 | 0.6 | 0.8 | 0.2 |
| Cpy62 | 0.7 | 0.9 | 0.2 | 0.5 | 0.9 | 0.5 | 0.4 | 0.9 | 0.5 | 0.5 | 0.9 | 0.4 |
| Cpy80 | 0.4 | 0.9 | 0.5 | 0.1 | 0.9 | 0.9 | 0.1 | 0.9 | 0.8 | 0.2 | 0.9 | 0.7 |
| Cpy89 | 0.7 | 0.9 | 0.2 | 0.8 | 0.9 | 0.2 | 0.8 | 0.9 | 0.1 | 0.8 | 0.9 | 0.2 |
| Cpy102 | 0.5 | 0.9 | 0.4 | 0.8 | 0.9 | 0.2 | 0.4 | 0.9 | 0.6 | 0.6 | 0.9 | 0.4 |
| Cpy144 | 0.6 | 1.0 | 0.3 | 0.6 | 0.9 | 0.4 | 0.7 | 0.9 | 0.3 | 0.6 | 1.0 | 0.3 |
| Mean | 0.5 | 0.7 | 0.3 | 0.5 | 0.8 | 0.4 | 0.5 | 0.8 | 0.4 | 0.5 | 0.8 | 0.4 |

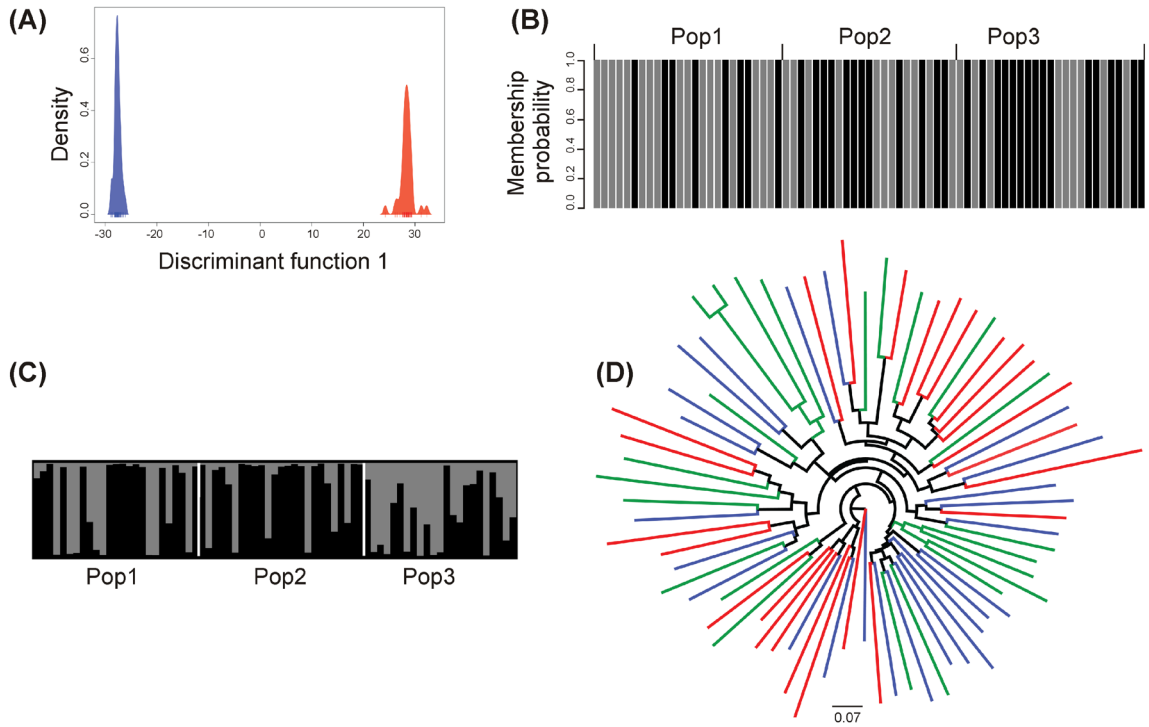


**Figure 2.** Genetic structure based on microsatellite markers: **A.** DAPC discrimination analysis; **B.** Membership probability for individuals from three populations; **C.** Genetic components as established in the Structure analysis; **D.** Evolutionary relationships for individuals from the three populations as inferred by N-J. In B and C, each vertical bar represents one individual and different colors correspond to each genetic component. In D, each color represents different populations (red – population 1; blue – population 2; and green – population 3).

Following the method of Evanno *et al.* (2005), the model-based clustering implemented in Structure found two distinct main genetic clusters (best K = 2), with both genetic components distributed in the three populations and the majority of individuals being purebred (Fig. 2C). The evolutionary relationships among individuals observed through N-J revealed there not to be a preferential association among individuals based on their collection sites (Fig. 2D).

### *Genetic diversity and population structure in* ***C. pygmaea*** *based on cpDNA*

The three intergenic regions obtained from 35 individuals of *C. pygmaea* produced a concatenated alignment of 2,011 base pairs (bp) long (955 bp in *rpl32-trnL;* 376 bp in *trnH-psbA;* and 680 bp in *trnS-trnG*). Across all individuals and populations, these regions presented 11 polymorphic sites (six transitions and five transversions), leading to 13 haplotypes. The evolutionary relationships among the 13 haplotypes (Fig. 3) did not exhibit a spatial correlation based on different collection sites. Each population presented exclusive haplotypes, and the highest number of haplotypes was observed in population 1. Haplotype H1 was the most frequently observed (46 %) and was present in all populations. Haplotype diversity considering all populations was $h$ = 0.86 and nucleotide diversity was $\pi$ = 0.07 % (±0.05 %). Concerning to population structure based on cpDNA, the AMOVA revealed that ~64 % of the total genetic variation was presented within populations. The fixation index was high and significant (*FST* = 0.36; P < 0.05). The Mantel test indicated a non-significant relationship between genetic and geographic distances (r = 0.0745; P = 0.02). Furthermore, the neutrality tests were negative and significant, (*D* = -1.74; P = 0.02; *FS* = -8.71; P = 0.0), suggesting the species did not follow the neutral evolutionary model.

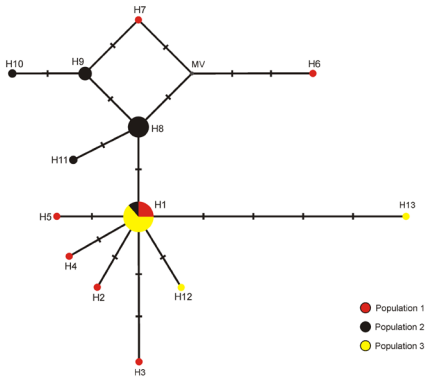


**Figure 3.** Evolutionary relationships among plastid haplotypes of *Calibrachoa pygmaea* obtained by the median-joining network approach. Colors identify populations as indicated in the graphic. Circle area is proportional to haplotype frequency and perpendicular bars indicate evolutionary steps between haplotypes. MV – median vector.

## Discussion

Herein, we described novel nuclear markers and plastid sequence to study the genetic diversity of *Calibrachoa* species and characterized genetic variability and population structure of *C. pygmaea*, a species narrowly distributed in the humid Pampas. Our main goal was to understand the processes conducting this species evolution.

*Calibrachoa pygmaea* is unique within the genus because of its flower morphology, pollination syndrome, ecological issues, and physiological aspects. This species displays a moth-pollination syndrome defined by a white corolla with a long and thin tube as well as strong scent emission at dusk (Fregonezi *et al.* 2012). Contrarily to the remaining species in *Calibrachoa*, *C. pygmaea* is an annual herb and features reduced DNA content (Mishiba *et al.* 2000).

Floral morphology and corolla color are key traits in pollinator attraction (Fenster *et al.* 2004; Schiestl & Johnson 2013), which is fundamental to species diversification and population structure (Bradshaw & Schemske 2003). Self-incompatible species, such as *C. pygmaea* (Tsukamoto *et al.* 2002), usually requires pollen vectors and mates (Charlesworth 2006), and fragmented or restricted distribution can limit partner availability (Yin *et al.* 2016).

An efficient evolutionary strategy in specialized pollination systems consists of maximizing the efficiency of specific pollinators that are successful in depositing pollen grains on stigmas and increasing the visitation frequency of conspecific flowers (Reynolds *et al.* 2009). In animal-pollinated species, the distances over which the pollen is dispersed are dependent on various ecological factors, such as the spatial distribution of population, density, and flowering phenology that may affect the foraging behavior of pollinators (Barrett 2003; Ghazoul 2005). In mass- and synchronized-flowering species, the pollen dispersal distance may be short because of the high proportion of pollination among neighbors (White & Boshier 2000). Furthermore, species with limited dispersal may have diminished genetic neighborhoods, implying an increase in mating between relatives (Carrillo-Angeles *et al.* 2011; Turchetto *et al.* 2015). The dynamics of gene flow among populations is conditioned upon several factors, such as species dispersal ability, the geographic distance among individuals and populations, landscape features, and ecological factors that facilitate or constrain gene movement (Cushman *et al.* 2016). The fruits of *C. pygmaea*, as all species of the *Calibrachoa*, are dried capsules that contain dozens of small seeds, with no dispersal mechanisms that fall close to the mother plant. The populations are, in general, small, with less than 30 individuals distributed in small patches, which blossom simultaneously and abundantly during the spring (October to December). As with all *Calibrachoa* species, *C. pygmaea* presents a vast number of flowers per plant (Kanaya *et al.* 2010).

In this work, we showed the highest component of genetic diversity of *C. pygmaea* is within populations based on both marker types, as seen through AMOVA results. This species possesses high genetic variability compared to other *Calibrachoa* species (*i.e.*, John *et al.* 2019), as observed by the number and haplotype diversity, as well as the nuclear marker diversity indices, despite also displaying elevated levels of inbreeding – suggested by the high and significant *F*IS values. Clearly dependent on pollinators, this species' genetic structure and diversity may be affected by crossings between related individuals more than gene flow between populations. Environmental factors can explain this due to the discontinuity of humid areas along the Pampas (Havrylenko *et al.* 2016), and perhaps the pollinator biology, as it is a very small hawkmoth and so it is incapable of flying long distances. Barochoric seed dispersal in this species (Pijl 1982) can also increase the crosses between relatives as was described for *Petunia* species (Turchetto *et al.* 2015; Rodrigues *et al.*, unpublished data). Therefore, inbreeding, small genetic structure, and limited seed dispersal may decrease individual heterozygosity, leading to the high observed values of *F*IS. Moreover, according to the microsatellite data, all populations share the same genetic matrix as suggested by Structure, DAPC, and N-J results. These data can be better explained by common allele fixation during population ancestor expansion (Excoffier & Ray 2008) or ancestral polymorphism sharing after population divergence (Schaal & Olsen 2000) instead of long-distance gene flow – this was also supported by the result of the Mantel test in plastid markers.

We noted high transferability in the microsatellites developed here, especially for *C. parviflora*, but also for species in the *Calibrachoa* subg. *Stimomphis*. Microsatellite flanking regions are conserved among closely related species, especially those recently diverged taxa (Moodley *et al.* 2015) as in the case of the *Calibrachoa* species (Fregonezi *et al.* 2012; 2013). In this way, microsatellites are ideal for testing hypotheses related to genetic segregation at fine spatiotemporal scales and are powerful tools in population and evolutionary genetics for all *Calibrachoa* species.

## Acknowledgements

This project was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). G.M. was supported by the PNPD-CAPES/PPG Botânica UFRGS and A.B. received a scholarship from the National Institutes for Science and Technology (INCT-CNPq) in Ecology, Evolution, and Biodiversity Conservation.

## References

Ando T, Kokubun H, Watanabe H, *et al.* 2005. Phylogenetic analysis of *Petunia* sensu Jussieu (Solanaceae) using chloroplast DNA RFLP. Annals of Botany 96: 289-297.

Barrett SCH. 2003. Mating strategies in flowering plants: the outcrossing – selfing paradigm and beyond. Philosophical Transactions of the Royal Society B: Biological Sciences 358: 991-1004.

Blambert L, Mallet B, Humeau L, Pailler T. 2016. Reproductive patterns, genetic diversity and in-mating depression in two closely related *Jumellea* species with contrasting patterns of commonness and distribution. Annals of Botany 118: 93-103.

Bradshaw HD, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. Nature 426: 176-178.

Carrillo-Angeles IG, Mandujano MC, Golubov J. 2011. Influences of the genetic neighborhood on ramet reproductive success in a clonal desert cactus. Population Ecology 53: 449-458.

Charlesworth D. 2006. Evolution of plant breeding systems. Current Biology 16: R726-R735.

Cushman SA, McRae HB, McGarigal K. 2016. Basics of landscape ecology: An introduction to landscapes and population processes for landscape geneticists. In: Balkenhol N, Cushman SA, McRae BH, McGarigal K. (eds.) Landscape genetics: concepts, methods, applications. Chichester, John Wiley & Sons Ltd. p. 9-34.

Dieringer D, Schlötterer C. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Molecular Ecology Notes 3: 167-169.

Dunn IS, Blattner FR. 1987. Charons 36 to 40: Multi-enzyme, high capacity, recombination deficient replacement vectors with polylinkers and polystuffers. Nucleic Acids Research 15: 2677-2698.

Earl EA, Holdt BM. 2012. Structure Harvester: a website and program for visualizing structure output and implementing the Evanno method. Conservation Genetics Resources 4: 359-361.

Ellstrand NC, Elam DR. 1993. Population genetic consequences of small population size: implications for plant conservation. Annual Review of Ecology and Systematics 24: 217-242.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611-2620.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Research 10: 564-567.

Excoffier L, Ray N. 2008. Surfing during population expansions promotes genetic revolutions and structuration. Trends in Ecology and Evolution 23: 347-351.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes - Application to human mitochondrial-DNA restriction data. Genetics 131: 479-491.

Falush D, Tephens MS, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination syndromes and floral specialization. Annual Review of Ecology, Evolution and Systematics 31: 375-403.

Fregonezi JN, Freitas LB, Bonatto SL, Semir J, Stehmann JR. 2012. Infrageneric classification of *Calibrachoa* (Solanaceae) based on morphological and molecular evidence. Taxon 61: 120-130.

Fregonezi JN, Turchetto C, Bonatto SL, Freitas LB. 2013. Biogeographic history and diversification of *Petunia* and *Calibrachoa* (Solanaceae) in the Neotropical Pampas grassland. Botanical Journal of the Linnean Society 171: 140-153.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915-925.

Ghazoul J. 2005. Pollen and seed dispersal among dispersed plants. Biological Reviews 80: 413-443.

Goudet J. 1995. FSTAT version 1.2: a computer program to calculate F-statistics. Journal of Heredity 86: 485-486.

Greppi JA, Hagiwara JC, Stehmann JR. 2013. Novelties in *Calibrachoa* (Solanaceae) and taxonomic notes on the genus for Argentina. Darwiniana 1: 173-187.

Hamilton MB. 1999. Four primers pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Molecular Ecology 8: 513-525.

Havrylenko SB, Bodoque JM, Srinivasan R, Zucarelli GV, Mercuri P. 2016. Assessment of the soil water content in the Pampas region using SWAT. Catena 137: 298-309.

John ALW, Mäder G, Fregonezi JN, Freitas LB. 2019. Genetic diversity and population structure of naturally rare *Calibrachoa* species with small distribution in southern Brazil. Genetics and Molecular Biology 42: 108-119.

Jombart T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403-1405.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11: 94. doi: 10.1186/1471-2156-11-94.

Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16: 1099-1106.

Kanaya T, Watanabe H, Kokubun H, *et al*. 2010. Current status of commercial *Calibrachoa* cultivars as assessed by morphology and other traits. Scientia Horticulturae 123: 488-495.

Kriedt RA, Ramos-Fregonezi AMC, Beheregaray LB, Bonatto SL, Freitas LB. 2011. Isolation, characterization and cross-amplification of microsatellite markers for *Petunia integrifolia* (Solanaceae) complex. American Journal of Botany 98: e277-e279.

Kulcheski FR, Muschner VC, Lorenz-Lemke AP, *et al*. 2006. Molecular phylogenetic analysis of *Petunia* Juss. (Solanaceae). Genetica 126: 3-14.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33: 1870-1874.

Lorenz-Lemke AP, Mäder G, Muschner VC, *et al*. 2006. Diversity and natural hybridization in a highly endemic species of *Petunia* (Solanaceae): a molecular and ecological analysis. Molecular Ecology 15: 4487-4497.

Mantel N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209-220.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7: 639-655.

Miller MP. 2005. Alleles In Space: computer software for the joint analysis of inter-individual spatial and genetic information. Journal of Heredity 96: 722-724.

Mishiba KI, Ando T, Mii M, *et al*. 2000. Nuclear DNA content as an index character discriminating taxa in the genus *Petunia* sensu Jussieu (Solanaceae). Annals of Botany 85: 665-673.

Moodley Y, Masello JF, Cole TL, *et al*. 2015. Evolutionary factors affecting the cross-species utility of newly developed microsatellite markers in seabirds. Molecular Ecology Resources 15: 1046-1058.

Pijl L. 1982. Principles of dispersal in higher plants. New York, Springer-Verlag.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Reynolds RJ, Westbrook MJ, Rohde AS, Cridland JM, Fenster CB, Dudash MR. 2009. Pollinator specialization and pollination syndromes of three related North American *Silene*. Ecology 90: 2077-2087.

Rice G. 1997. Petunias: a garden paradigm. Garden 122: 390-393.

Roy A, Frascaria N, MacKay J, Bousquet J. 1992. Segregating random amplified polymorphic DNAs (RAPDs) in *Betula alleghaniensis*. Theoretical and Applied Genetics 85: 173-180.

Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497.

Saitou N, Nei N. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4: 406-425.

Sang T, Crawford DJ, Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). American Journal of Botany 84: 1120-1136.

Schaal BA, Olsen KM. 2000. Gene genealogies and population variation in plants. Proceedings of the National Academy of Sciences of the United States of America 97: 7024-7029.

Schiestl FP, Johnson SD. 2013. Pollinator-mediated evolution of floral signals. Trends in Ecology and Evolution 28: 307-315.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. American Journal of Botany 94: 275-288.

Silva-Arias GA, Mäder G, Bonatto SL, Freitas LB. 2015. Novel microsatellites for *Calibrachoa heterophylla* (Solanaceae) endemic to the South Atlantic Coastal Plain of South America. Applications in Plant Sciences 3: 1500021. doi: 10.3732/apps.1500021.

Stehmann JR. 1999. Estudos taxonômicos da tribo Nicotianeae G. Don (Solanaceae): revisão de *Petunia* Jussieu, das espécies brasileiras de *Calibrachoa* La Llave & Lexarza e o estabelecimento do novo gênero *Petuniopsis* Stehmann & Semir. PhD Thesis, Universidade Estadual de Campinas, Campinas.

Tajima F. 1989. Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.

Tsukamoto T, Ando T, Watanabe H, *et al*. 2002. Differentiation in the status of self-incompatibility among *Calibrachoa* species (Solanaceae). Journal of Plant Research 115: 185-193.

Turchetto C, Lima JS, Rodrigues DM, Bonatto SL, Freitas LB. 2015. Pollen dispersal and breeding structure in a hawkmoth-pollinated Pampa grasslands species *Petunia axillaris* (Solanaceae). Annals of Botany 115: 939-948.

Watanabe H, Ando T, Iida S, *et al*. 1997. Cross-compatibility of *Petunia pubescens* and *P. pygmaea* with native taxa of *Petunia*. Journal of Japanese Society of Horticultural Science 66: 607-612.

White GM, Boshier DH. 2000. Fragmentation in Central American dry forests: genetic impacts on *Swietenia humilis* (Meliaceae). In: Young AG, Clarke GM. (eds.) Genetics, demography and viability of fragmented populations. Cambridge, Cambridge University Press. p. 293-311.

Yin G, Barrett SCH, Luo Y-B, Bai W-N. 2016. Seasonal variation in the mating system of a selfing annual with large floral displays. Annals of Botany 117: 391-400.