MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 374–389 doi: 10.1111/mec.12632

Diversification in the South American Pampas: the genetic and morphological variation of the widespread *Petunia axillaris* complex (Solanaceae)

CAROLINE TURCHETTO,* NELSON J. R. FAGUNDES,* ANA L. A. SEGATTO,* CRIS KUHLEMEIER,† VIVIANA G. SOLÍS NEFFA,‡ PABLO R. SPERANZA,§ SANDRO L.BONATTO¶ and LORETA B. FREITAS*

**Laboratory of Molecular Evolution, Department of Genetics, Universidade Federal do Rio Grande do Sul, PoBox 15053, 91501-970, Porto Alegre, RS, Brazil, †Institute of Plant Sciences, Altenbergrain 21 CH-3013, Bern, Switzerland, ‡Instituto de Botánica del Nordeste (UNNE-CONICET), Sargento Cabral 2131, Corrientes 3400, Argentina, §Facultad de Agronomía, Universidad de la República, Av. E. Garzón 780, Montevideo 12900, Uruguay, ¶Laboratory of Genomic and Molecular Biology, Pontifícia Universidade Católica do Rio Grande do Sul, Ipiranga 6681, Porto Alegre, RS 90610-001, Brazil*

Abstract

Understanding the spatiotemporal distribution of genetic variation and the ways in which this distribution is connected to the ecological context of natural populations is fundamental for understanding the nature and mode of intraspecific and, ultimately, interspecific differentiation. The *Petunia axillaris* complex is endemic to the grasslands of southern South America and includes three subspecies: *P. a. axillaris*, *P. a. parodii* and *P. a. subandina*. These subspecies are traditionally delimited based on both geography and floral morphology, although the latter is highly variable. Here, we determined the patterns of genetic (nuclear and cpDNA), morphological and ecological (bioclimatic) variation of a large number of *P. axillaris* populations and found that they are mostly coincident with subspecies delimitation. The nuclear data suggest that the subspecies are likely independent evolutionary units, and their morphological differences may be associated with local adaptations to diverse climatic and/or edaphic conditions and population isolation. The demographic dynamics over time estimated by skyline plot analyses showed different patterns for each subspecies in the last 100 000 years, which is compatible with a divergence time between 35 000 and 107 000 years ago between *P. a. axillaris* and *P. a. parodii*, as estimated with the IMa program. Coalescent simulation tests using Approximate Bayesian Computation do not support previous suggestions of extensive gene flow between *P. a. axillaris* and *P. a. parodii* in their contact zone.

Keywords: CAPS, diversification, floral morphology, grasslands, *Petunia axillaris*, plastid markers

Received 12 July 2012; revision received 20 November 2013; accepted 22 November 2013

Introduction

Understanding the spatiotemporal distribution of genetic variation and the ways in which this distribution is connected to the ecological context of natural populations is fundamental for understanding the nature and mode of intraspecific and, ultimately, interspecific differentiation (Hewitt 2001). Diversification due to ecological divergence may occur by natural selection that favours different alleles/allelic combinations in different environments or, in the case of flowering plants, due to differences in floral traits that attract different pollinators and result in pollinator-mediated selection (Armbruster 1991; Armbruster & Schwaegerle 1996; Cresswell 1998; Medel *et al.* 2003; Perez-Barrales

Correspondence: Loreta B. Freitas, Fax: 55 51 3308 9823; E-mail: loreta.freitas@ufrgs.br

et al. 2007; Chalcoff *et al.* 2008; Ordano *et al.* 2008; Cuartas-Dominguez & Medel 2010; Rosas-Guerrero *et al.* 2011). Alternatively, diversification may arise as a by-product of population isolation due to palaeoclimatic fluctuations (Knowles 2001; Arora *et al.* 2010; Chen *et al.* 2010; Harris & Taylor 2010; Dong-Rui *et al.* 2011) or other environmental factors in allopatry, whereby genetic isolation resulting from geographical factors eventually leads to the evolution of reproductive barriers and speciation. The spatial arrangement of genetically divergent populations is important in allopatric diversification, as it involves the geographical separation of a previously continuous metapopulation by geological or climatic causes or a founder event resulting from long-distance dispersal (Turelli *et al.* 2001; Gavrilets 2003). A lack of gene flow, drift and secondary adaptation to local environmental conditions among geographically separated populations can eventually result in morphologically and genetically distinct lineages. On the other hand, morphological variation may represent plastic genomes interacting with different environments without any relevant underlying genetic differentiation. This phenomenon is common in nature, for example, after recent adaptive radiation when ecologically specialized forms arise rapidly from a single ancestral form and occupy a variety of ecological niches (Gavrilets & Losos 2009).

The southeastern area of South America – including Uruguay, northern, central and eastern Argentina, southern Brazil, the southern half of Paraguay and southern Bolivia – is dominated by open vegetation of several different types. To the south, the Ventania and Tandilia Sierras delimit this area at 39°S, and the southern Brazilian plateau delimits the area to the north/northeast. To the west, the Andean Piedmont and the Sub-Andean and Pampean Sierras delimit the Pampas. In geomorphological and botanical terms, this is a complex region known as Pampas in Brazil, Uruguay and Southern Argentina and as Chaco in its central and northern distribution in Argentina and Bolivia (Fig. 1 and Fig. S1A, Supporting information). The Uruguayan and Brazilian Pampas exhibit a clear geological continuity with rocky outcrops dating from the Proterozoic to the Quaternary and resulting a great variety of soil types derived from basalt, granite, sandstone and silt (Grela 2004; Fregonezi *et al.* 2013). The climate is variable in this region and promotes, together with the different soil types, the occurrence of a plethora of ecological conditions (Cabrera & Willink 1973; Iriondo & García 1993). However, with the exception of a few major rivers such as the Uruguay and Negro, no obvious geographical barrier exists. Despite its wide area, little is known about the processes influencing the diversification of the plant species in the grassland

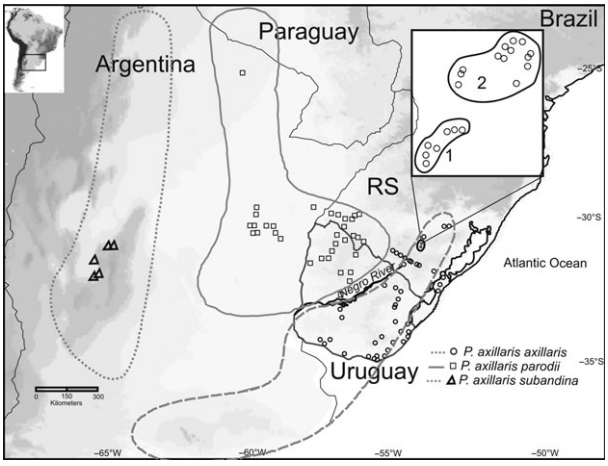

Fig. 1 Map showing the geographical distribution of each subspecies according to the literature (Ando 1996) and collection sites (see more details in Table S1, supporting information).

mosaics of southeastern South America (Speranza *et al.* 2007; Turchetto-Zolet *et al.* 2013).

Petunia axillaris (Lam.) Britton, Sterns & Poggenb (Solanaceae) has a wide distribution in the open vegetation area of southeastern South America (Fig. 1) and is the only *Petunia* species that has white flowers. *Petunia axillaris* has a corolla that ranges from 3 to 7 cm long (Fig. S2, Supporting information) (Stehmann *et al.* 2009) and is pollinated by hawkmoths (Ando *et al.* 1995; Venail *et al.* 2010; Klahre *et al.* 2011), although other insects are commonly found to be associated with the flowers (Kokubun *et al.* 1997; Tsukamoto *et al.* 2003). Three allopatric subspecies have been suggested (Ando 1996) for *P. axillaris* according to morphological differences in four floral characteristics observed in dried specimens: *P. a. axillaris*, *P. a. parodii* and *P. a. subandina* (Fig. 1). Two subspecies are found in the Pampas region but only in flat areas, and *P. a. subandina* is found exclusively in the uplands of the Sub-Andean Mountains and adjacent regions (Ando 1996). However, differences in flower morphology have been described for *P. a. axillaris* when Brazilian populations were compared to populations from Uruguay and Argentina (Kokubun *et al.* 2006). Additionally, an intermediate morphology between subspecies *P. a. axillaris* and *P. a. parodii* has been reported for individuals from populations found along the Negro River, which flows through central Uruguay and delimits the area of occurrence of *P. a. parodii* and *P. a. axillaris* (Kokubun *et al.* 1997). Kokubun *et al.* (1997) have suggested that these individuals with intermediate morphology might be hybrids between the two subspecies. Although the recognition in the *P. axillaris* subspecies has been based exclusively on these traits and the geographical distribution, it is unfortunate that morphological variations

in the *P. axillaris* complex have been studied primarily in greenhouse-grown plants (Ando *et al.* 1994, 1995; Ando 1996) or geographically limited populations (Kokubun *et al.* 1997). From a taxonomic and evolutionary perspective, it is fundamental to understand the extent of variation found in natural populations.

This species complex has not been previously analysed using genetic markers aside from the study of a few localized populations in studies focusing on other species (Kulcheski *et al.* 2006; Lorenz-Lemke *et al.* 2006; Fregonezi *et al.* 2013). Most evolutionary and phylogenetic studies in *Petunia* used chloroplast (cp) DNA sequence variability and have typically revealed little genetic differentiation and high haplotype sharing between the species (Ando *et al.* 2005; Kulcheski *et al.* 2006; Chen *et al.* 2007; Lorenz-Lemke *et al.* 2010).

In this study, we analysed the morphological, climate and molecular (both cpDNA and nuclear) data from a large number of populations of the three *P. axillaris* subspecies covering a large part of the geographical range of *P. a. axillaris* and *P. a. parodii*, including their contact zone. First, we determined and compared the genetic structure and morphological differentiation of *P. axillaris* populations and tested whether the obtained patterns correspond to the three described subspecies. We also performed a multivariate analysis on bioclimatic variables to visualize major climatic differences between the populations to determine whether these difference correlate with the subspecies delimitation. Comparing these results allowed us to test whether the subspecies previously delimited mainly by the morphological variation are real evolutionary units or are morphotypes determined by the different environments. Second, we tested previous suggestions that *P. a. axillaris* and *P. a. parodii* may hybridize in their contact zone by using coalescent simulations in an approximate Bayesian computation approach. Finally, we estimated the divergence time between *P. a. axillaris* and *P. a. parodii* and various demographic parameters such as effective population sizes and their fluctuation along time to better understand the evolutionary history of the species.

Material and methods

Sample collection

Natural populations of *P. axillaris* were visited during two spring seasons. Overall, 102 localities were sampled including 63 occurring in the *P. a. axillaris* range, 34 occurring in the *P. a. parodii* range and five occurring in the *P. a. subandina* range (see Fig. 1). We classified each population at the level of *P. axillaris* subspecies according to the geographical boundaries and morphological characteristics proposed previously (Ando *et al.* 1994, 1995; Ando 1996) (Table S1, Supporting information; this classification was used throughout the entire data analysis). For every population, the geographical coordinates were obtained using a global positioning system (GPS) unit, and an exsiccate was generated and deposited in the ICN Herbarium, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, or the BHCB Herbarium, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil. For some of the analyses, we pooled populations together into regional groups (Table S1, Supporting information) based on the geographical distances between sampling points and landscape features such as rivers and the geological context of the localities. This strategy was intended to provide an improved characterization of the genetic and morphologic variation at a broader scale, and we referred to these as regional level analyses. The geographical distribution of the three subspecies and localities sampled in this study are shown in Fig. 1. Detailed information on the sampled populations is provided in Table S1, Supporting information.

DNA extraction, cpDNA sequencing and analysis of cleaved amplified polymorphic sequence (CAPS) markers

We collected fresh leaves from 614 individuals and extracted DNA using the method of Roy *et al.* (1992). We used the polymerase chain reaction (PCR) to amplify two intergenic cpDNA spacers, *trnS-trnG* and *trnH-psbA,* using the universal primers described by Hamilton (1999) and Sang *et al.* (1997), respectively. These markers are among the fastest evolving cpDNA regions in the *Petunia* genus (Lorenz-Lemke *et al.* 2006, 2010). The PCR amplification was performed as described in Lorenz-Lemke *et al.* (2006), and the resulting products were purified according to the 20% polyethylene glycol (PEG) method (Dunn & Blattner 1987) and sequenced using a MEGABACE1000 system (GE Healthcare) following the manufacturer's instructions. The sequences were deposited in GenBank (Accession nos: DQ225609 to DQ225665, DQ225367 to DQ225423, JF917318 to JF918431, and JN097453) and manually aligned. The alignments were visually assessed using GENEDOC (Nicholas & Nicholas 1997). All insertion/deletion events (indels) that involved poly-T/A stretches were eliminated from the analyses because their mutational dynamics cannot be assessed (Aldrich *et al.* 1988).

To assess the genetic diversity in nuclear DNA, six codominant CAPS markers that have been used to study genetic diversity in several plant species (Tsumura *et al.* 1999; Barth *et al.* 2002; Kaundun & Matsumoto 2003; Huang *et al.* 2006) were analysed. The CAPS named HF1, EPF1, MYB75, MYB60, MYBX and PAL2B

were developed for *Petunia* (Venail *et al.* 2010) and are available online at http://www.botany.unibe.ch/deve/caps/index.html. These markers were evaluated using the PCR-restriction fragment length polymorphism (RFLP) approach, and the different cleavage patterns visually detected after a 2.5% agarose gel electrophoresis were considered different alleles. Given the small allele number, only a subsample of the 614 individuals was genotyped; this totalled 353 individuals randomly sampled within the populations (Tables S1 and S2, Supporting information). Table S3 in the supporting information provides additional information and the detailed protocols for this procedure. To verify whether the different band patterns detected by gel electrophoresis represent different DNA sequences, the PCR products resulting in each pattern were purified with PEG (Dunn & Blattner 1987) and sequenced using a MEGABACE 1000 system (GE HealthCare) following the manufacturer's specifications. All of the different restriction band patterns were confirmed as different alleles by sequencing and are available in GenBank under the Accession nos JF737840 to JF737855.

Morphological analysis

We examined the floral morphology of 790 individuals (1689 flowers, 79 populations; Tables S1 and S4, Supporting information). In the field, we used digital callipers to measure four flower traits: the length of the floral tube, the diameter of the corolla, the total length of the corolla and the distance between the anthers of the longest and medium-length stamens (Fig. S2, Supporting information). These are the most important traits for subspecies discrimination (Kokubun *et al.* 2006). To assess both the population and individual variability, we measured three flowers per individual and ten individuals per population whenever possible. Some populations were not measured (Table S1, Supporting information) because most of the measurements were performed during a second visit to the sites, when some of these populations could not be found because of intense drought. This situation was particularly critical for *P. a. subandina* for which only a single population could be measured.

We used a discriminant analysis of principal components (DAPC) to determine whether the subspecies could be distinguished by these four morphometric variables. The DAPC relies on data transformation using principal component analysis (PCA) as a prior step to discriminant analysis (DA), which maximizes the separation between groups. We also performed correspondence analyses (CA) using the regional population groups as units (using the within group average morphometric values).

To determine the location of the greatest morphological breaks and whether these coincide with the proposed geographical limits of the subspecies, a matrix of the squared Mahalanobis distance per individual using the flower traits was constructed, and its geographical significance was evaluated using the Monmonier algorithm as implemented in the software BARRIER 2.2 (Manni *et al.* 2004). The DAPC and CA were performed with the adegenet package for R (Jombart 2008), and Mahalanobis distance computations were performed using the STATISTICA 9.0 package (available at http://www.statsoft.com/).

Climate analyses

To test whether the morphological subspecies are discernible based on bioclimatic variables without their a priori designation, we performed a PCA using the occurrence data supplied by the samples used in the genetic analyses. The PCA was performed using the climate data (19 WorldClim bioclimatic layers for temperature and precipitation, see http://www.worldclim.org/bioclim) extracted through the program DIVA GIS (available at http://www.diva-gis.org/) and using variables that had been log-transformed in the STATISTICA program. The analysis was based on the correlation matrix (correlation as ss/N-1), which does not intensify the range of variation (Jolliffe 2000). The first two principal components were plotted, and the variation found for each point was compared to the description of the subspecies.

Genetic diversity and population structure using molecular data

For the cpDNA haplotypes, basic descriptive molecular diversity statistics were calculated in ARLEQUIN 3.5.1.2 as the haplotype and nucleotide diversity (Excoffier & Lischer 2010). The evolutionary relationship between the haplotypes was estimated by a median-joining network (Bandelt *et al.* 1999) calculated in the NETWORK 4.1.0.9 program (available at www.fluxux-engineering.com). The degree of genetic differentiation between subspecies was estimated using analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) and between subspecies and regional groups of populations by pairwise Φ_{ST}s, using ARLEQUIN.

For the CAPS markers, basic descriptive statistics such as diversity indices, observed (H_O) and expected (H_E) heterozygosities, test of Hardy–Weinberg equilibrium (HWE) and the between-subspecies pairwise F_{ST}s and AMOVA were calculated using ARLEQUIN. A Bayesian clustering method (Pritchard *et al.* 2000) implemented in the STRUCTURE 2.3.4 software was used to detect and test the presence of any genetic structure between the

samples using the admixture model (Pritchard *et al.* 2000; Falush *et al.* 2003) and correlated allelic frequencies (Falush *et al.* 2003) parameters but without any a priori information regarding population structure. This method assigns a probability that each individual belongs to each *K* ancestry cluster. We ran 10 replicates of the program from $K = 1$ to $K = 6$, allowing 250 000 Markov chain Monte Carlo (MCMC) steps as burn-in and 1 000 000 sampling steps. To choose the best value for *K*, we first used the method of Pritchard *et al.* (2000), which selects the *K* value returning the highest a posteriori probability. We then used the method of Evanno *et al.* (2005), which considers the second-order rate of change in the likelihood of the data given *K*, $\Pr(X \mid K)$. The individual ancestry coefficients were calculated by the average pairwise similarity of individual assignments across runs with Clumpp (Jakobsson & Rosenberg 2007) using the FullSearch method and weighted by the number of individuals in each population; Distruct (Rosenberg 2004) was used to plot the individual ancestry coefficients.

To better understand the structure of the populations based on the CAPS population variability but without the genetic assumption posed by the Bayesian structure method, we performed CA using the allelic frequency for each regional population group as described above for the morphological data. Finally, the CAPS pairwise F_{ST} values were used to estimate the main geographical barriers between the populations in the Monmonier algorithm implemented in Barriers again, as described above for the morphological analyses.

Test of hybridization scenarios using approximate Bayesian computation

One of our goals was to test the suggestions of hybridization between subspecies *axillaris* and *parodii* in their contact zone (Kokubun *et al.* 1997). However, it is difficult to distinguish gene flow from shared ancestral polymorphism in *Petunia* because the latter is common in *Petunia* (see Lorenz-Lemke *et al.* 2006, 2010). To examine these questions, we tested three scenarios using an approximate Bayesian computation (ABC) approach (see review in Bertorelle *et al.* 2010). The first scenario is divergence with no gene flow allowed (allopatric model). In the second scenario, gene flow may occur throughout the entire divergence process (parapatric model). The third scenario includes only recent gene flow after divergence (secondary contact model). These models were not tested over the entire distribution but over three selected population pairs from each subspecies, following an approach described previously (e.g. Surget-Groba & Kay 2013). The first pair (named Disjunct) was composed of allopatric populations from distant regions. This pair could be used as a control for the allopatric model because no gene flow was expected between them. The second pair (named Negro River) consisted of populations located in the region around the Negro River, the contact zone. The third pair (named Campanha) consisted of two populations in southern Brazil from the two subspecies for which some of our analyses suggested the presence of individuals of intermediate morphology or mixed genetic ancestry. Therefore, we tested three scenarios using three population sets each.

All the ABC analyses were performed using ABCTOOLBOX 1.0 (Wegmann *et al.* 2010) using both the cpDNA and CAPS data sets. The cpDNA data were modelled with the substitution rate as described above for the BSP analysis and no recombination. The CAPS data were modelled as single nucleotide polymorphism (SNP) data (biallelic marker), assuming no mutation. This is equivalent to assuming that every CAPS allele identical by state would also be identical by descent at some time in the past and provides a good approximation when the timescale of the process is much shorter than the timescale for which mutation becomes an important evolutionary force affecting allelic frequencies.

For the scenarios parameters, we defined prior distributions (Table S5, Supporting information) that were largely based on the results of the IMa analysis. For example, we used a uniform prior for divergence time distributed between 15 thousand years ago (kya) and 90 kya. We used a uniform prior for the time of secondary contact (between 3 and 5 kya) based on the latest geological events affecting sedimentary regions in the Quaternary (Ponce *et al.* 2011) and a uniform prior of migration rate (between 0.001 and 0.0001 effective migrant chromosomes per generation). Because migration is an important parameter in the scenarios we want to discriminate, there is a compromise between using low migration rates and the power to discriminate between scenarios (the lower the migration rates in a scenario with migration, the more this rate will resemble a truly allopatric scenario). These values were used because, while calibrating the models (see below), we observed that they could generate summary statistic values close to those observed in the original data and because they appeared reasonable compared with other studies.

We used a total of 19 summary statistics, with six extracted from the CAPS data (average number of alleles in each population, average H_E in each population, total H_E and F_{ST}) and 13 extracted from the cpDNA data (number of haplotypes in each population, number of segregating sites in each population, number of private segregating sites in each population, total number of segregating sites, Fu's F_S in each population, nucleotide diversity in each population, average

nucleotide diversity and Φ_{ST}). We performed an initial 10 000 simulations to evaluate whether the range of prior values was able to generate simulations close to the observed data and then performed 500 000 simulations for each scenario for the model selection procedure. The original summary statistics were transformed by partial least squares (PLS) after Box-Cox transformation according to Wegmann and Excoffier (2008), and the 1000 best simulations (those closest to the observed summary statistics) were retained for model selection. The Bayes factors (BFs) for pairwise model comparisons were estimated based on seven PLS components and interpreted according to Kass and Raftery (1995). It is important to note that these analyses focus on performing model (scenarios) selection by integrating over the uncertainty of the demographic parameters and not estimating the demographic parameters per se.

Demographic and IMa analysis with cpDNA data

To detect evidence for deviation of the neutral equilibrium model of evolution, we performed Tajima's *D* (1989) and Fu's F_S (1997) neutrality tests using the ARLEQUIN program. These tests were conducted for the species as a whole, for the subspecies and for the regional population groups. In addition, changes in population size over time for subspecies were estimated with a Bayesian skyline plot analysis (BSP, Drummond *et al.* 2005) performed in BEAST 1.6.1 (Drummond & Rambaut 2007). The priors for this analysis were a strict molecular clock model with a mean substitution rate of 2.8×10^{-9} and a standard deviation of 5.4×10^{-11} per site per year (Lorenz-Lemke *et al.* 2010) and an HKY (Hasegawa, Kishino, and Yano) nucleotide substitution model. MCMC was performed for 100 000 000 steps, sampling every 10 000 steps. The computation of the BSP and convergence checking were performed in TRACER 1.5 (available at http://beast.bio.ed.ac.uk/Tracer).

We used the isolation-with-migration model (Nielsen & Wakeley 2001; Hey & Nielsen 2004) implemented in the IMa program (Hey & Nielsen 2007) to estimate population demographic parameters for the two Pampean subspecies *P. a. axillaris* and *P. a. parodii*, especially their divergence time. This program uses a MCMC approach to estimate the posterior probability densities of demographic parameters including the effective population sizes, divergence time and migration rates. However, because we could not obtain stable estimates of the migration rates, we excluded the migration parameters and used a pure allopatric scenario. This choice was further justified by the results from the ABC analysis, which did not show support for divergence models including gene flow (see below). We used 50 Metropolis-coupled chains with a 20 000-step burn-in period, followed by 120 hours of Markov chain interactions based on the HKY nucleotide substitution model, the same substitution rate used for the BSP analysis and with the inheritance scalar set to 0.25. We assumed a value of 2 years for the generation time in all analyses, as based on our field observations during more than 10 years of collecting *Petunia* species.

Results

Morphological variation

Petunia axillaris exhibits wide morphological variation, and it was not uncommon to find individuals with a flower morphology that was inconsistent with its classification at the subspecific level based on the traditional geographical boundaries assumed for each subspecies according to Ando (1996). Considering the typical values for the diagnostic floral traits (Ando *et al.* 1995; Ando 1996), ≈64% of the individuals would be classified within the (geographically) expected subspecies and the remaining individuals would be classified within another subspecies (Table 1).

The first two axes of the multivariate DAPC (Fig. 2A) show that most of the individuals are organized in distinct clusters corresponding to the three subspecies, but most of the individuals of *P. a. subandina* overlap with one extreme of the distribution of *P. a. axillaris*. There is also some overlap between *P. a. axillaris* and *P. a. parodii* individuals, but not between the latter and *P. a. subandina*. Interestingly, overlapping between *P. a. axillaris* and *P. a. parodii* is mostly restricted to individuals from a set of populations located in the parapatric zone on the opposite margins of the Negro River, Uruguay (Fig. S3, Supporting information). Also interesting is that *P. a. axillaris* populations from its northern area (Brazil) present a wider within-population variation. A notable example of this is the Guaritas region in Serra do Sudeste, which encompasses populations sampled within a geographical region slightly larger than 30 km. These populations show a broader variation in flower morphology than do those of the entire distribution of *P. a. parodii*. Importantly, at the population level, the correspondence analysis of the regional population groups shows that they are clearly separated according to their subspecies (Fig. 2B), suggesting that these traits present high intrapopulation variability.

The Monmonier maximum difference algorithm revealed that, using the regional groups of populations as units, the first barrier separates the populations north and south of the Negro River (see Fig. 1) and continues in the northeastern direction towards Brazil, which corresponds to the proposed limits of *P. a. parodii* and *P. a. axillaris* (Ando 1996).

Table 1 Morphological classification of the *Petunia axillaris* complex. Proportions of individuals classified correctly or incorrectly according to the morphology expected for each subspecies in relation to the accepted geographical boundaries (see Fig. 1). The first column lists the subspecies based on geographical location and the first row provides the morphological identification

	N	subsp. *axillaris* (%)	subsp. *parodii* (%)	subsp. *subandina* (%)	Atypical (%)*
subsp. *axillaris*	538	292 (54)	14 (3)	215 (40)	17 (3)
subsp. *parodii*	237	0 (0)	198 (84)	32 (13)	7 (3)
subsp. *subandina*	15	0 (0)	0 (0)	15 (1)	0 (0)

*Atypical – individuals morphology classified as either subspecies. Correctly classified – 64%; % incorrectly classified – 36%.

Fundamental niche divergence

The PCA of the 19 bioclimatic variables revealed a high degree of niche divergence among the subspecies of *P. axillaris*. The first two principal components explained 73.6% of the variation, with the first component (PC1) accounting for 39.7% of the variation (Fig. 2C). In PC1, the variables related to precipitation explained 48% of the variation. The niche separation of *P. a. subandina* is evident, but *P. a. axillaris* and *P. a. parodii* were closer to each other in the PCA space with a small overlapping region between them. The populations associated with the overlap region are those in Uruguay around the Negro River, a location where the ranges of the two subspecies are close in proximity.

cpDNA genetic variability

The *trnH-psbA* and *trnS-trnG* intergenic cpDNA spacers were analysed in a single alignment totalling 1064 base pairs (bp). A total of 28 single-base substitutions (12 transitions and 17 transversions) and one inversion of 29 bp were found; the micro-inversion was coded as a single mutational event. Thirty-five haplotypes were found in the 614 samples analysed for the three subspecies. Table 2 shows the diversity indices for the cpDNA markers. *Petunia axillaris*, which had higher morphological variation, exhibited higher values for both haplotype and nucleotide diversity than did the other subspecies.

Most of the haplotypes in the network (Fig. 3) are separated by a single change. At a broad scale, the network is not structured geographically or in relation to the putative subspecies to which each individual belong, but some structure occurs in parts of the network. Another important pattern is that haplotype sharing between subspecies occurs mostly with central and frequent haplotypes, but the terminal haplotypes are not shared between subspecies. The sharing of haplotypes between populations (see Fig. S1B, Supporting information) was also unrelated to the level of admixture between the subspecies inferred using a Bayesian approach (see below). The extensive sharing of haplotypes between subspecies is also reflected in the AMOVA results, where only 6.9% of the total variation was found between subspecies. On the other hand, the portion of the variation found between populations *within* subspecies is high (74.3%), with the remaining 18.8% of the variation being distributed within populations.

CAPS genetic variability

We genotyped 354 individuals of the three subspecies for six CAPS markers (Table S1 and S2, Supporting information), and we identified new alleles in two of the six markers. New alleles were identified when new mutations promoted the gain or loss of a restriction site, which was identified as a novel band pattern after electrophoresis. All of these alleles were confirmed by sequencing.

Considering all of the CAPS loci, *P. a. axillaris* had the largest number of alleles and the highest H_E (Table 3). At the subspecies level, there were strong departures of HWE (Table 3), which may be explained by the general self-compatible (SC) status of *P. axillaris* throughout most of its range (Kokubun *et al.* 2006) because SC reduces heterozygosity due to inbreeding and genetic drift increases differentiation between populations. This also likely led to a Wahlund effect resulting from the pooling of the isolated populations in each subspecies.

AMOVA was used to investigate the overall distribution of genetic diversity between the three subspecies, revealing 19.6% of the genetic variation between the subspecies. Most of the genetic variation was found within populations (49.1%) and among populations within subspecies (31.3%). The pairwise F_{ST} values revealed genetic differentiation between the *P. axillaris* subspecies (0.174 *axillaris* vs. *parodii*; 0.256 *axillaris* vs. *subandina*; 0.433 *subandina* vs. *parodii*), with all the pairwise values being significant ($P < 0.001$).

In the STRUCTURE analysis of CAPS genotypes (Fig. 4), $K = 2$ was found to be the best number of genetic clusters by all methods (Pritchard *et al.* 2000; Evanno *et al.* 2005; Gao *et al.* 2007). No meaningful substructure within these two clusters could be seen in the $K > 2$ bar plots (see Fig. S4, Supporting information). In general, one of the clusters was related to the individuals

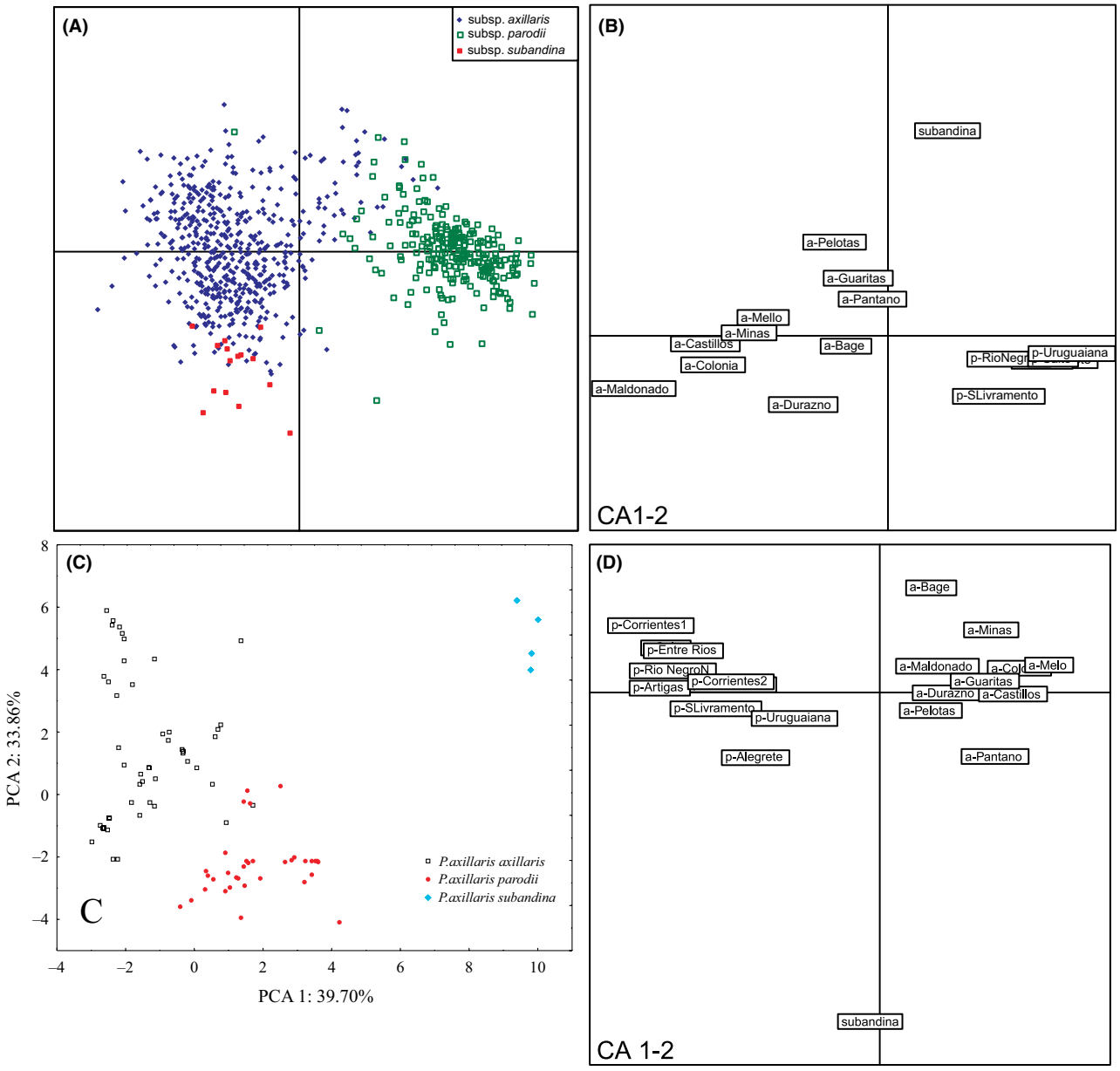

Fig. 2 (A) Discriminant analysis of principal components of the floral characters of 790 individuals from the three subspecies of *Petunia axillaris*. (B) Axis 1 and 2 of the correspondence analysis of 17 regional groups of populations based on the floral characters. 'a-' and 'p-' in the labels identify *P. a. axillaris* and *P. a. parodii* populations, respectively. (C) The extent of the ecological separation of *Petunia axillaris* for the three subspecies based on variations in climate and elevation. The points correspond to each population of *P. axillaris* studied. Climatic variables contained in the Bioclimate packed through the program DIVA GIS (available at: http://www.diva-gis.org/). (D) Axis 1 and 2 of the CA analysis of the allelic frequencies of six CAPS markers on 21 regional groups of populations. Labels as in (B).

identified as *P. a. axillaris*, and the other is more associated with the individuals classified as *P. a. parodii* (see Fig. 4). However, some individuals of both subspecies exhibited a significant proportion of their ancestry in both clusters. Two geographical regions appear to have a higher incidence of individuals showing 'dual ancestry', in agreement with the results obtained in the analysis of morphological traits. One of these regions is within the vicinity of the margins of the Negro River in Uruguay, which corresponds to the geographical distribution boundaries of the two Pampean subspecies. The second region is Bage in southern Brazil, a site at which only *P. a. axillaris* occurs; interestingly, the other subpopulations in this region had flower morphologies more similar to that of *P. a. parodii*. Individuals belonging to all the populations of *P. a. subandina* were either

Table 2 Diversity indices and neutrality tests for the cpDNA markers for each subspecies and for each group of the populations

Parameters	No. of plants	No. of haplotypes	Nucleotide diversity π % (SD)	Haplotype diversity h (SD)	Tajima's *D*	Fu's F_S
P. axillaris complex	614	35	0.22 (0.13)	0.70 (0.02)	−1.14	−17.58*
P. a. axillaris	359	17	0.21 (0.13)	0.77 (0.02)	−0.49	−2.60
P. a. parodii	230	23	0.20 (0.21)	0.55 (0.04)	−0.86	−9.43*
P. a. subandina	25	5	0.19 (0.12)	0.58 (0.09)	1.37	0.98
Durazno group, UY	12	3	0.08 (0.07)	0.62 (0.09)	−0.43	0.39
Colonia group, UY	28	1	0.00	0.00	—	—
Maldonado group, UY	50	4	0.06 (0.05)	0.48 (0.06)	−0.80	−0.49
Minas group, UY	19	2	0.15 (0.10)	0.51 (0.05)	2.17	4.24
Castillos group, UY	30	6	0.26 (0.16)	0.80 (0.04)	1.02	1.37
Melo group, UY	20	5	0.30 (0.18)	0.81 (0.04)	0.89	2.14
Pelotas group, BR	24	3	0.06 (0.05)	0.36 (0.11)	0.40	0.38
Bage group, BR	47	4	0.17 (0.11)	0.50 (0.07)	0.41	0.80
Guaritas group, BR	114	7	0.14 (0.10)	0.70 (0.03)	0.20	0.36
Pantano group, BR	15	2	0.15 (0.11)	0.53 (0.05)	2.18	3.99
Rio Negro N group, UY	22	6	0.35 (0.20)	0.83 (0.04)	1.62	1.78
Salto group, UY	30	6	0.32 (0.19)	0.72 (0.05)	1.53	2.23
Artigas group, UY	17	4	0.30 (0.18)	0.64 (0.07)	2.62	2.99
Slivramento group, BR	25	5	0.19 (0.12)	0.68 (0.08)	−0.82	1.02
Alegrete group, BR	24	2	0.01 (0.01)	0.08 (0.07)	−1.16	−1.03
Uruguaiana group, BR	34	3	0.05 (0.05)	0.21 (0.09)	−2.00*	0.22
Entre Rios group, AR	33	3	0.04 (0.04)	0.12 (0.07)	−2.18*	−0.20
Corrientes 1 group, AR	27	5	0.18 (0.12)	0.49 (0.12)	−0.15	0.56
Corrientes 2 group, AR	18	2	0.01 (0.02)	0.11 (0.10)	−1.16	−0.79
Formosa group, AR	6	2	0.03 (0.04)	0.33 (0.22)	−0.93	−0.01
Cordoba group, AR	25	5	0.19 (0.12)	0.58 (0.09)	1.37	0.98

*$P < 0.01$.

classified into cluster 1 or 2 or had fractions of their genome associated with both.

On the other hand, in the analysis at the population level, but not the individual level, the CA of the CAPS allele frequency clearly group the 21 regional population groups (Fig. 2D) in three distinct clusters that correspond to the three subspecies. Finally, the main barrier estimated by the Monmonier algorithm separates the populations north and south of the Negro River and continues towards the east in Rio Grande do Sul, Brazil (see Fig. 1). Again, this corresponds to the limits of the distribution of *P. a. parodii* and *P. a. axillaris* in Uruguay as proposed by Ando (1996), overlapping largely with the estimated barrier inferred using morphological data above.

Approximate Bayesian computation

Although the above results suggest that in general, the subspecies *P. a. axillaris* and *P. a. parodii* are differentiated on a morphological, environmental and genetic basis, some populations show evidence of individuals with mixed ancestry/morphology. We tested the three sets of population pairs using ABC model selection to determine which of the three scenarios (divergence with no gene flow, divergence with low but continuous gene flow or divergence with only recent gene flow: allopatric, parapatric and secondary contact models, respectively) better explains the cpDNA and CAPS genetic data. For the 'Disjunct' pair comparison, the allopatric model was, as expected, better supported than was either the parapatric or the recent contact model (BF = 308, BF = 146, respectively), with decisive support (Kass & Raftery 1995). Similarly, for both the 'Negro River' and 'Campanha' population pairs, the allopatric model was better supported than was either the parapatric or the recent contact model (BF = 388, BF = 2515, respectively, for the 'Negro River'; BF = 223, BF = 128, respectively, for the 'Campanha'). Therefore, these results indicate that despite previous suggestions of hybridization, there is no support to significant recent gene flow connecting these populations since the divergence of the subspecies.

Demographic parameters and isolation-with-migration model with cpDNA

Given our findings that the three subspecies are most likely independent evolutionary units, we used the isolation-with-migration demographic model in IMa to

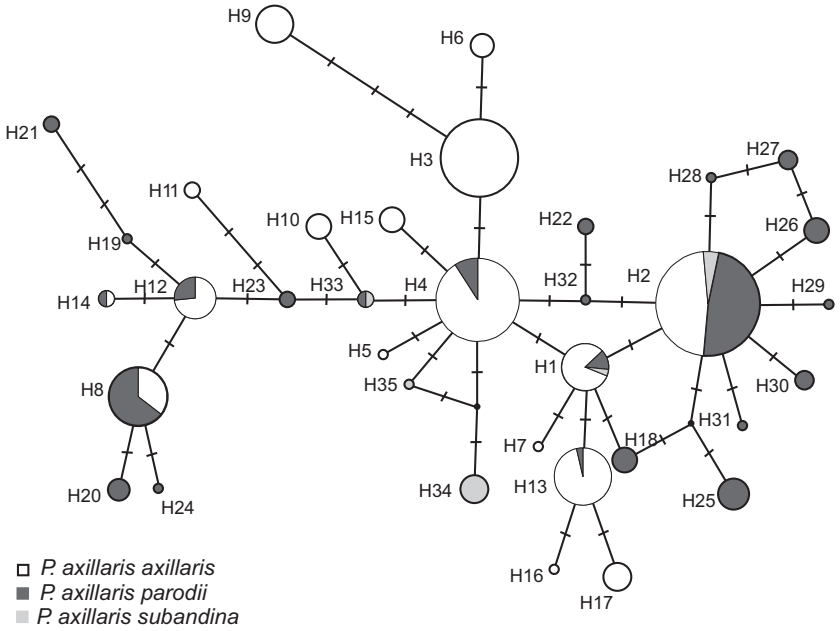

Fig. 3 Network showing the evolutionary relationships between 35 haplotypes of the populations of *Petunia axillaris* (combination of *trnS-trnG* and *trnH-psbA* sequencing). Each colour represents one of the three subspecies according to the legend. The area of the circles is proportional to the frequency of the haplotype. The perpendicular bars indicate the number of evolutionary steps separating the haplotypes.

Table 3 Absolute and relative frequencies of alleles and observed and expected heterozygosity per CAPS loci for each subspecies of *Petunia axillaris*

subsp.	Locus	N. alleles	Allele frequency (1; 2; 3; 4)	Obs. Het	Exp. Het
axillaris	HF1	3	0.76; 0.19; 0.05; 0.00	0.17*	0.38
	EPF1	2	0.03; 0.97; 0.00; 0.00	0.05	0.06
	Myb75	4	0.06; 0.33; 0.58; 0.03	0.24*	0.56
	Myb60	2	0.31; 0.69; 0.00; 0.00	0.23*	0.43
	Mybx	2	0.01; 0.99; 0.00; 0.00	0.02	0.02
	Pal2B	2	0.34; 0.66; 0.00; 0.00	0.28*	0.45
	Mean	2.5			
parodii	HF1	2	0.01; 0.99; 0.00; 0.00	0.03	0.03
	EPF1	2	0.03; 0.97; 0.00; 0.00	0.04	0.05
	Myb75	3	0.43; 0.30; 0.27; 0.00	0.15*	0.65
	Myb60	2	0.59; 0.41; 0.00; 0.00	0.04*	0.48
	Mybx	1	0.00; 1.00; 0.00; 0.00	0.00	0.00
	Pal2B	2	0.03; 0.97; 0.00; 0.00	0.06	0.06
	Mean	2			
subandina	HF1	3	0.17; 0.64; 0.19; 0.00	0.06*	0.50
	EPF1	2	0.33; 0.67; 0.00; 0.00	0.00*	0.47
	Myb75	2	0.12; 0.88; 0.00; 0.00	0.25	0.23
	Myb60	1	0.00; 1.00; 0.00; 0.00	0.00	0.00
	Mybx	1	0.00; 1.00; 0.00; 0.00	0.00	0.00
	Pal2B	1	0.00; 1.00; 0.00; 0.00	0.00	0.00
	Mean	1.7			

*$P < 0.01$.

make further inferences regarding the history of divergence between *P. a. axillaris* and *P. a. parodii*, the two subspecies with better sampling. Overall, there was evidence of large population sizes for both subspecies [with a mean estimate of ≈1 400 000 and ≈2 000 000 individuals for *P. a. axillaris* and *P. a. parodii*, respectively,

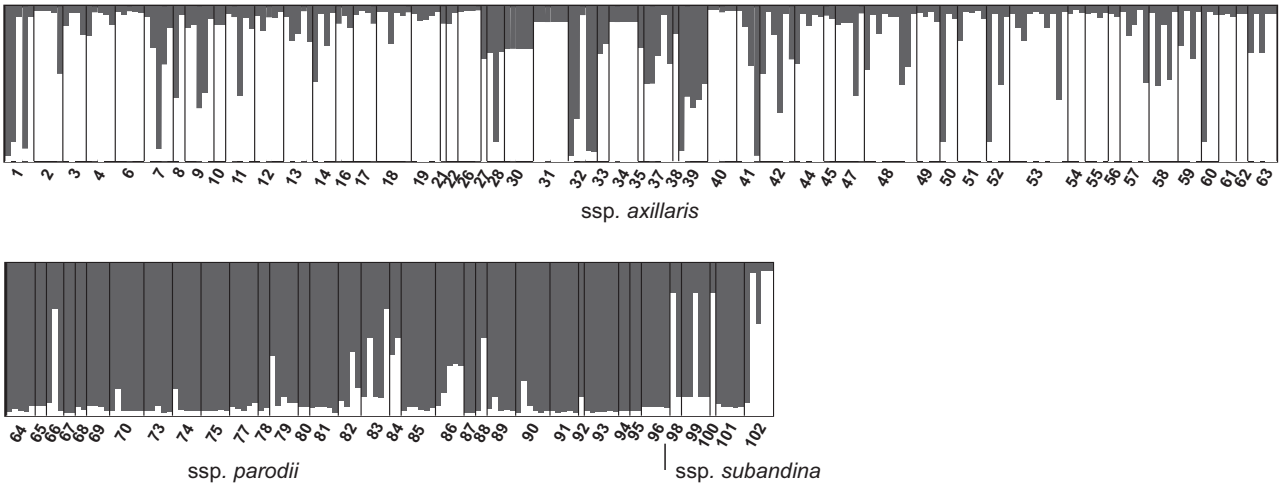

Fig. 4 Genomic constitution inferred by the program STRUCTURE based on six nuclear CAPS markers. Each population is represented as a vertical line partitioned into $K = 2$ coloured components representing membership in one of the genetic clusters. The population names are according to Table S1 (Supporting information).

and an estimated divergence time of ≈58 kya (95% credible interval between ≈35–107 kya)]. This finding is largely in agreement with a divergence in the late Pleistocene and before the Last Glacial Maximum.

With respect to the neutrality tests (Table 2), the only significant values of Fu's F_S were the negative values for *P. axillaris* as a whole and *P. a. parodii*, but this test was also negative, but not significant for *P. a. axillaris*. Tajima's *D* values were also negative (but not significant) for the above three taxa. On the other hand, the values for *P. a. subandina* were positive (but not significant) for both tests. These results suggest a history of population growth for *P. a. parodii* and *P. a. axillaris*, with a stronger suggestion in the former, and a possible population reduction for *P. a. subandina*. The Bayesian skyline plots (Fig. 5) mostly supported these conclusions, suggesting a period of population growth more intense approximately 100 kya for *P. a. parodii* (Fig. 5A), a signal of population decline since approximately 50 kya for *P. a. subandina* (Fig. 5B) and a moderate bottleneck beginning approximately 20 kya and an expansion since 10 kya for *P. a. axillaris* (Fig. 5C).

Discussion

Morphological and molecular diversity and differentiation between P. axillaris subspecies

We found wide variation in the four floral traits given that approximately 35% of individuals would be classified into a subspecies different from the one expected based on their geographical location (Table 1). A question we asked is whether, despite such a high rate of misassignment using standard univariate statistics, it is possible to find discontinuities in flower morphology using multivariate methods and whether these discontinuities would be consistent with the traditional geographical classification of *P. axillaris* subspecies. Using the multivariate DPCA method, the floral morphology clearly differentiated the populations of the three subspecies (Fig. 2B), but there was overlap between some individuals in a few populations (see below) (Fig. 2A and Fig. S3, Supporting information).

Notably, the structuring of the sampling sites according to their climatic variables is also consistent with the traditional distribution of *P. axillaris* subspecies (Fig. 2C) and their different morphologies. Moreover, Aoki and Hattori (1991) reported that the shape of the corolla in *Petunia* species is significantly associated not only with the growing temperature but also soil type; the distribution of *P. a. parodii* includes areas of soil originated mostly from basalt rocks, but *P. a. axillaris* occupies a region where soils derived from granites (or Quaternary sediments) are predominant (Bizzi *et al.* 2001). These results may suggest that differences in flower morphology between subspecies merely reflect environmental conditions instead of a meaningful evolutionary difference between them. For example, Solis-Neffa (2009) showed that the morphotypes of *Turnera sidoides pinnatifida*, which presents a distribution similar to the *P. axillaris* complex, live in different habitats. This author proposed that the differences observed between populations could be the consequence of either adaptation to different environments or a high phenotypic plasticity of the species. Here, the clear genetic distinction between the subspecies, especially with the CAPS markers (Fig. 2D), indicates that these morphological differences are related to population differentiation and

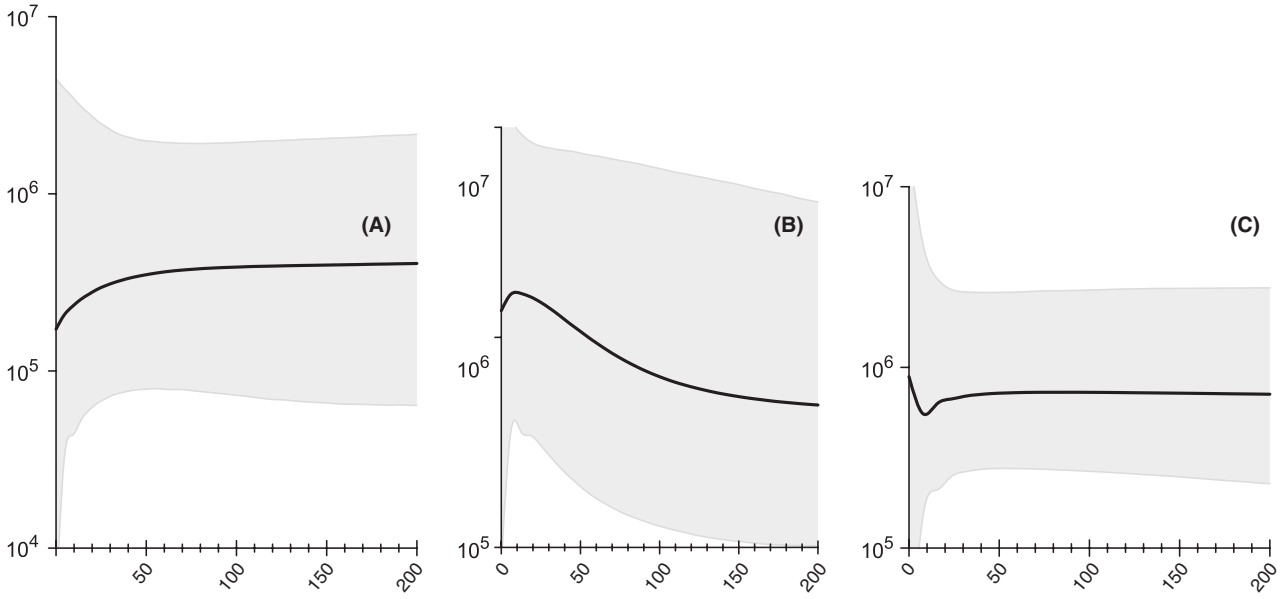

Fig. 5 Bayesian skyline plots showing the fluctuations in effective population size (N_e) over time 200 (thousands years ago). (A) *Petunia axillaris parodii* (B), *Petunia axillaris subandina and* (C) *Petunia axillaris axillaris*. Median estimate (dark line) and upper and lower interval (gray area).

genetic adaptation rather than simple phenotypic plasticity. It is widely accepted that ecological divergence due to habitat differences plays an important role in population differentiation (Foster *et al.* 2007; Lowry *et al.* 2008; Zheng & Ge 2010). Based on controlled experiments in the laboratory, Venail *et al.* (2010) proposed that the difference in floral tube length between these two subspecies indicates that they are visited by different hawkmoth species that have proboscides of different lengths, suggesting their differences in floral characters may additionally be explained as an adaptive response to the divergent selective pressures generated by pollinators. Although this process appears to contribute widely to plant diversification in this geographical region of South America (Fregonezi *et al.* 2013), studies of plant–pollinator interactions in the field are necessary to test this hypothesis. Therefore, the differences in the floral characteristics in the *P. axillaris* complex could be explained by climatic and edaphic factors (Fig. 2C) most likely coupled with the adaptive response to divergent selective pressures generated by other factors such as pollinator selection.

We also found that the morphological variation in this species is much higher than previously recognized (Ando *et al.* 1994; Ando 1996), particularly for *P. a. axillaris* as seen in the high dispersion of individuals found within some population groups in the DAPC axes (Fig. S3, Supporting information). For example, the populations found in the Guaritas region in Serra do Sudeste, Brazil, displayed the greatest morphological variation of the entire complex. This area is approximately 900 km^2 with elevations ranging between 200 and 500 m and several forms of physiognomic and vegetation diversity (Caporal & Boldrini 2007). One possible explanation for the high morphological diversity in this region could lie in the range of soil diversity in a small geographical area. However, *P. a. axillaris* and its sister species *P. exserta*, a species endemic to rock formations in the Serra do Sudeste region that has a long corolla tube similar to that of *P. a. parodii* (Lorenz-Lemke *et al.* 2006), seem to hybridize in this region. Therefore, considering that hybridization could occur in both directions (Lorenz-Lemke *et al.* 2006), such events may have also contributed to the wide morphological variation found in this region for *P. a. axillaris* individuals and their higher morphological similarities with *P. a. parodii*.

Although only few populations have been evaluated genetically and only one morphologically for *P. a. subandina*, this subspecies is clearly distinct from the other two subspecies based on morphology (Fig. 2B) and CAPS allele frequency (Fig. 2D), even though there is overlap at the individual level (e.g. Fig. 2A). Overall, these findings suggest that *P. a. subandina* also represents an independent evolutionary unit for this species, but additional sampling over a larger area of its distribution is needed.

Testing hypothesis of hybridization between P. a. axillaris and P. a. parodii

Our ABC-based model selection procedure using cpDNA and CAPS data clearly shows that the finding

of a high proportion of 'misassigned' individuals at the margins of the Negro River, Uruguay, and in the 'Campanha' region, Brazil, does not suggest extensive recent gene flow between the two subspecies as originally proposed at least for the Negro River area (Ando *et al.* 1994; Kokubun *et al.* 1997). Although our best-supported model (allopatric) does not represent the entire history of these two subspecies, it is more strongly supported than are the other models involving gene flow between the subspecies. Indeed, the *axillaris* and *parodii* subspecies behave like allopatric entities. Because the Negro River lies on the border of the *axillaris* and *parodii* subspecies distribution, the population in this region could represent a convergent status on the basis of ecological similarities, which is supported by the results of the PCA using climatic data (Fig. 2C).

On the other hand, cpDNA results seem to present a very different picture with high haplotype sharing between subspecies, a small proportion of total genetic diversity attributed to divergence between the subspecies (AMOVA results) and a network with a lack of structure in relation to the subspecies or geography (Fig. 3 and Fig. S1B, Supporting information). This lack of cpDNA differentiation between the subspecies could be explained by lineage sorting of ancestral polymorphisms due to recent divergence or by extensive recent gene flow. However, several lines of evidence support the recent divergence explanation. First, haplotype sharing between subspecies occurs mostly with frequent and central haplotypes, and the sharing of haplotypes between populations was unrelated to the level of admixture between the populations (as estimated by structure analyses). Moreover, the divergence between *P. a. axillaris* and *P. a. parodii* estimated with IMa is relatively recent, between 35 and 107 kya. Finally, the ABC rejected the scenario of recent gene flow, even between populations of *P. a. axillaris* and *P. a. parodii* in the contact zone. Actually, low differentiation with extensive cpDNA haplotype sharing is common in species of *Petunia* and its sister genus *Calibrachoa*, with most studies suggesting rapid morphological divergence (Ando *et al.* 2005; Kulcheski *et al.* 2006; Lorenz-Lemke *et al.* 2010; Fregonezi *et al.* 2013). Another line of evidence for the hypothesis of shared ancestral polymorphisms involves the life history characteristics of *Petunia* such as the maternal inheritance of cpDNA (Derepas & Dulieu 1992) and limited seed dispersal, which is restricted to short distances (Stehmann *et al.* 2009).

Divergence and demographic scenarios

The divergence time between *P. a. axillaris* and *P. a. parodii* between ≈35 and 107 kya found with the IMa analysis is relatively recent and compatible with their extensive sharing of cpDNA haplotypes (Fig. 3), but this divergence time may be greater than that between *P. a. axillaris* and its sister *P. exserta* (≈17 kya, A. Segatto, C. Turchetto, S.L. Bonatto, L.B. Freitas, unpublished data). *Petunia exserta* and *P. a. axillaris* are distinguished primarily by differences in habitat and floral syndromes. The former is hummingbird pollinated, and the latter is hawkmoth pollinated (Lorenz-Lemke *et al.* 2006).

The effective population size (N_e) and its dynamics in time as estimated by the BSP approach using the cpDNA data showed different patterns for each subspecies (Fig. 5), which are mostly concordant with the signals from the neutrality tests (Table 3). One important point is that these different demographic patterns support the evolutionary independence of the three subspecies because that would otherwise present similar histories. Note also that the median estimates of the present day N_e for *P. a. axillaris* and *P. a. parodii* in the BSPs, which are approximately $1E^6$, are broadly similar to the independent N_e estimates from IMa (that used both the cpDNA and CAPS data). Because these are the first genetic- or census-based population size estimates for this species, we could compare these values with only similar estimates for some species of *Petunia* from the highlands (Lorenz-Lemke *et al.* 2010) and one lowland species from the sister species *Calibrachoa* (Mäder *et al.* 2013). Interestingly, both of the latter studies also estimated N_es to be approximately $1E^6$, and these are large numbers as expected for herbaceous species with large distributions. The difference in the demographic history between *P. a. axillaris* and *P. a. parodii*, as seen in the BSPs (Fig. 5), begins approximately between 50 and 100 kya, which was the estimated divergence time between them. Because the subspecies are distributed in different geomorphological areas (Fig. 1 and Fig. S1A, Supporting information) and present clear niche (bioclimatic variables) divergence (Fig. 2C), they likely responded differently to the palaeoclimatic fluctuation over the last hundred thousand years. For example, *P. a. parodii* distribution extends well into the Chaco Plain, which is characterized by a highly seasonal climate that produces xeromorphic forest formations (Spichiger *et al.* 2004). The climate was drier and at least 5–7 °C cooler during glacial periods, and the Antarctic cold fronts were more frequent and more intense. These conditions could have allowed these fields to extend approximately 750 km to the north and reach latitudes of 28–20°S (Behling 2002), which could have allowed *P. a. parodii* to expand in range and size (Fig. 5A). Interestingly, a similar pattern of population growth was found in the highland *Petunia altiplana* (Lorenz-Lemke *et al.* 2010), which is also adapted to a colder climate. On the other hand, *P. a. axillaris* underwent a moderate

bottleneck around the LGM followed by a size recovery (Fig. 5C). A similar pattern of population increase following the end of the LGM was found in *Calibrachoa heterophylla*, a species that is also distributed (although much more restricted) in the eastern lowlands of the Pampas, a region that was much affected by the increase in temperature and humidity at the beginning of the Holocene (Mäder *et al.* 2013).

Acknowledgements

The authors thank AP Lorenz-Lemke and G Mäder for assistance in field collection and discussions. We also thank three anonymous reviewers and the Subject Editor for the comments and suggestions that improved this manuscript. This project was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), Programa de Pós Graduação em Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul (PPGBM-UFRGS), and the Swiss National Science Foundation NCCR 'Plant Survival' Program.

References

Aldrich J, Cherney BW, Merlin E, Christopherson L (1988) The role of insertions deletions in the evolution of the intergenic region between *psbA* and *trnH* in the chloroplast genome. *Current Genetics*, **14**, 137–146.

Ando T (1996) Distribution of *Petunia axillaris* (Solanaceae) and its new subspecies in Argentina and Bolivia. *Acta Phytotaxonomica et Geobotanica*, **47**, 19–30.

Ando T, Lida S, Kokubun H, Ueda Y, Marchesi E (1994) Distribution of intraspecific taxa of *Petunia axillaris* (Solanaceae) in Uruguay as revealed by discriminate analysis. *Acta Phytotaxonomica et Geobotanica*, **45**, 95–109.

Ando T, Lida S, Kokubun H, Ueda Y, Marchesi E (1995) Distribution of *Petunia axillaris* sensu lato in Uruguay as revealed by discriminate analysis of the live plants. *Journal of the Japanese Society for Horticultural Science*, **64**, 381–391.

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.

Aoki C, Hattori K (1991) Attempt to classify *Petunia* species on the basis of corolla shape. *Japanese Journal of Breeding*, **41**, 433–442.

Armbruster WS (1991) Multilevel analysis of morphometric data from natural plant-populations – insights into ontogenic, genetic, and selective correlations in dalechampia-scandens. *Evolution*, **45**, 1229–1244.

Armbruster WS, Schwaegerle KE (1996) Causes of covariation of phenotypic traits among populations. *Journal of Evolutionary Biology*, **9**, 261–276.

Arora N, Nater A, van Schaik CP *et al.* (2010) Effects of Pleistocene glaciations and rivers on the population structure of Bornean orangutans (Pongo pygmaeus). *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 21376–21381.

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Barth S, Melchinger AE, Lubberstedt T (2002) Genetic diversity in *Arabidopsis thaliana* L. Heynh. investigated by cleaved amplified polymorphic sequence (CAPS) and inter-simple sequence repeat (ISSR) markers. *Molecular Ecology*, **11**, 495–505.

Behling H (2002) South and southeast Brazilian grasslands during late Quaternary times: a synthesis. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **177**, 19–27.

Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, **19**, 2609–2625.

Bizzi LA, Schobbenhaus, Gonçalves JH *et al.* (2001) *Geology, Tectonics and Mineral Resources of Brazil: Geographic Information System (GIS) and Maps at the 1:2500000 Scale*. CPRM, Brasilia, 4 CD-Rom.

Cabrera AL, Willink A (1973) *Biogeografía de América Latina*. Secretaría General OEA, Washington, DC.

Caporal FJ, Boldrini IL (2007) Florística e fitossociologia de um campo manejado na Serra do Sudeste, Rio Grande do Sul. *Revista Brasileira de Biociências*, **5**, 37–44.

Chalcoff VR, Ezcurra C, Aizen MA (2008) Uncoupled geographical variation between leaves and flowers in a south-Andean Proteaceae. *Annals of Botany*, **102**, 79–91.

Chen S, Matsubara K, Omori T *et al.* (2007) Phylogenetic analysis of the genus *Petunia* (Solanaceae) based on the sequence of the *Hf1* gene. *Journal of Plant Research*, **120**, 385–397.

Chen W, Liu S, Liu Y *et al.* (2010) Phylogeography of the large white-bellied rat *Niviventer excelsior* suggests the influence of Pleistocene glaciations in the Hengduan Mountains. *Zoological Science*, **27**, 487–493.

Cresswell JE (1998) Stabilizing selection and the structural variability of flowers within species. *Annals of Botany*, **81**, 463–473.

Cuartas-Dominguez M, Medel R (2010) Pollinator-mediated selection and experimental manipulation of the flower phenotype in *Chloraea bletioides*. *Functional Ecology*, **24**, 1219–1227.

Derepas A, Dulieu H (1992) Inheritance of the capacity to transfer plastids by the pollen parent in *Petunia hybrida* Hort. *Journal of Heredity*, **83**, 6–10.

Dong-Rui J, Teng-Liang L, Liu-Yang W, Dang-Wei Z, Jian-Quan L (2011) Evolutionary history of an alpine shrub *Hippophae tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. *Biological Journal of the Linnean Society*, **102**, 37–50.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 1185–1192.

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

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 Resources*, **10**, 564–567.

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, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Foster SA, McKinnon GE, Steane DA, Potts BM, Vaillancourt RE (2007) Parallel evolution of dwarf ecotypes in the forest tree *Eucalyptus globulus*. *New Phytologist*, **175**, 370–380.

Fregonezi J, Turchetto C, Bonatto S, Freitas L (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.

Gao H, Williamson S, Bustamante CD (2007) A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics*, **176**, 1635–1651.

Gavrilets S (2003) Perspective: models of speciation: what have we learned in 40 years? *Evolution*, **57**, 2197–2215.

Gavrilets S, Losos JB (2009) Adaptive radiation: contrasting theory with data. *Science*, **323**, 732–737.

Grela IA (2004) *Geografía Florística de las Especies Arbóreas de Uruguay: Propuesta Para la Delimitación de Dendrofloras. Programa de Desarrollo de las Ciencias Básicas*, p. 103. Universidad de la República, Montevideo.

Hamilton MB (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular Ecology*, **8**, 521–523.

Harris LN, Taylor EB (2010) Pleistocene glaciations and contemporary genetic diversity in a Beringian fish, the broad whitefish, *Coregonus nasus* (Pallas): inferences from microsatellite DNA variation. *Journal of Evolutionary Biology*, **23**, 72–86.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences, USA*, **104**, 2785–2790.

Huang J, Li M, Zhang Y *et al.* (2006) A genetic diversity study of silkworm using cleaved amplified polymorphic sequence (CAPS) markers. *Biochemical Systematics and Ecology*, **34**, 868–874.

Iriondo MH, García NO (1993) Climatic variations in the Argentine plains during the last 18,000 years. *Palaeogeography, Palaeoclimatology, and Palaeoecology*, **101**, 209–220.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jolliffe IT (2000) *Principal Component Analysis*. pp. 487. Springer, New York.

Jombart T (2008) *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Kaundun SS, Matsumoto S (2003) Development of CAPS markers based on three key genes of the phenylpropanoid pathway in tea, *Camellia sinensis* (L.) O. Kuntze, and differentiation between *assamica* and *sinensis* varieties. *Theoretical and Applied Genetics*, **106**, 375–383.

Klahre U, Gurba A, Hermann K *et al.* (2011) Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. *Current Biology*, **21**, 730–739.

Knowles LL (2001) Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, **10**, 691–701.

Kokubun H, Ando T, Kohyama S *et al.* (1997) Distribution of intermediate forms of *Petunia axillaris* subsp. *axillaris* and subsp. *parodii* (Solanaceae) in Uruguay revealed by discriminate analysis. *Acta Phytotaxonomica et Geobotanica*, **48**, 173–185.

Kokubun H, Nakano M, Tsukamoto T *et al.* (2006) Distribution of self-compatible and self-incompatible populations of *Petunia axillaris* (Solanaceae) outside Uruguay. *Journal of Plant Research*, **119**, 419–430.

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

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.

Lorenz-Lemke AP, Togni PD, Mäder G *et al.* (2010) Diversification of plant species in a subtropical region of eastern South American highlands: a phylogeographic perspective on native *Petunia* (Solanaceae). *Molecular Ecology*, **19**, 5240–5251.

Lowry DB, Rockwood RC, Willis JH (2008) Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution*, **62**, 2196–2214.

Mäder G, Fregonezi JN, Lorenz-Lemke AP, Bonatto SL, Freitas LB (2013) Geological and climatic changes in Quaternary shaped the evolutionary history of *Calibrachoa heterophylla*, an endemic South-Atlantic species of petunia. *BMC Evolutionary Biology*, **13**, 178–190.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology*, **76**, 173–190.

Medel R, Botto-Mahan C, Kalin-Arroyo M (2003) Pollinator-mediated selection on the nectar guide phenotype in the Andean monkey flower, *Mimulus luteus*. *Ecology*, **84**, 1721–1732.

Nicholas KB, Nicholas KBJ (1997) GeneDoc: a Tool for Editing and Annotating Multiple Sequences Alignments. Available from www.psc.edu/biomed/genedoc.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Ordano M, Fornoni J, Boege K, Dominguez CA (2008) The adaptive value of phenotypic floral integration. *New Phytologist*, **179**, 1183–1192.

Perez-Barrales R, Arroyo J, Armbruster WS (2007) Differences in pollinator faunas may generate geographic differences in floral morphology and integration in *Narcissus papyraceus* (Alarcissiopapyraceris). *Oikos*, **116**, 1904–1918.

Ponce JF, Rabassa J, Coronato A, Borromei AM (2011) Palaeogeographical evolution of the Atlantic coast of Pampa and Patagonia from the last glacial maximum to the Middle Holocene. *Biological Journal of the Linnean Society*, **103**, 363–379.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Rosas-Guerrero V, Quesada M, Armbruster WS, Perez-Barrales R, Smith SD (2011) Influence of pollination specialization and breeding system on floral integration and phenotypic variation in ipomoea. *Evolution*, **65**, 350–364.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

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.

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

Solis-Neffa VG (2009) Geographic patterns of morphological variation in *Turnera sidoides* subsp. *pinnatifida* (Turneraceae). *Plant Systematics and Evolution*, **284**, 231–253.

Speranza PR, Seijo JG, Grela IA, Neffa VGS (2007) Chloroplast DNA variation in the *Turnera sidoides* L. complex (Turneraceae): biogeographical implications. *Journal of Biogeography*, **34**, 427–436.

Spichiger R, Calenge C, Bise B (2004) Geographical zonation in the Neotropics of tree species characteristic of the Paraguay-Paraná Basin. *Journal of Biogeography*, **31**, 1489–1501.

Stehmann JR, Lorenz-Lemke AP, Freitas LB, Semir J (2009) The genus Petunia. In: *Petunia: Evolutionary, Developmental and Physiological Genetics* (eds Gerats T, Strommer J). pp. 1–28, Springer, New York.

Surget-Groba Y, Kay KM (2013) Restricted gene flow within and between rapidly diverging Neotropical plant species. *Molecular Ecology*, **22**, 4931–4942.

Tajima F (1989) Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tsukamoto T, Ando T, Takahashi K *et al.* (2003) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by loss of pollen function. *Plant Physiology*, **131**, 1903–1912.

Tsumura Y, Tomaru N, Suyama Y, Bacchus S (1999) Genetic diversity and differentiation of *Taxodium* in the south-eastern United States using cleaved amplified polymorphic sequences. *Heredity*, **83**, 229–238.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C (2013) Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, **22**, 1193–1213.

Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends in Ecology and Evolution*, **16**, 330–343.

Venail J, Dell'Olivo A, Kuhlemeier C (2010) Speciation genes in the genus *Petunia*. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **365**, 461–468.

Wegmann D, Excoffier L (2008) Efficient approximate bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics*, **182**, 1207–1218.

Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, **11**, 116.

Zheng XM, Ge S (2010) Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology*, **19**, 2439–2454.

Collectively the group is interested in investigating evolutionary process and plant speciation.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 *Petunia axillaris* populations sampled for cpDNA, CAPS markers and morphology.

Table S2 Individual CAPS genotypes and their alleles per locus.

Table S3 CAPS markers chromosome location, primers sequences, annealing temperatures and restriction enzymes used per marker.

Table S4 Morphological measurements (mm) of all sampled individuals.

Table S5 Priors distributions for the demographic parameters in ABC analysis.

Fig. S1 (A) Extent of the Pampas region showing some geographical regions discussed in the text; (B) distribution of the haplotypes for the cpDNA markers.

Fig. S2 Photographs of the natural populations of *Petunia axillaris* analysed in this study: (A–D) *Petunia axillaris axillaris*; (E–H) *Petunia axillaris parodii*; and (I–L) *Petunia axillaris subandina*.

Fig. S3 Principal Components analysis of the characters of 790 individuals from 79 populations of *Petunia axillaris*. The ellipses represent the regional groups. The regional group showing the high morphological variability found for the Guaritas group.

Fig. S4 Genomic constitution inferred by the program STRUCTURE based on six nuclear CAPS markers.

Data S1 Data accessibility: alignment.

Data S2 Data accessibility.

Data S3 Data accessibility: climatic variables.