Systematics and Biodiversity (2015), **13(6)**: 571–580

Taylor & Francis
Taylor & Francis Group

Research Article

Genetic diversity of *Viola cazorlensis* Gand., an endemic species of Mediterranean dolomitic habitats: implications for conservation

JOSE LUIS CÁNOVAS[1], JUAN FRANCISCO JIMÉNEZ[1], JUAN FRANCISCO MOTA [2] & PEDRO SÁNCHEZ GÓMEZ[1]

[1]Departamento de Biología Vegetal (Botánica), Universidad de Murcia, Campus de Espinardo s/n, E-30100, Murcia, Spain
[2]Departamento de Biología y Geología, Universidad de Almería, Carretera Sacramento s/n CITE-2B, E-04120 La Cañada de San Urbano (Almería), Spain

(Received 21 April 2015; accepted 20 July 2015)

Viola cazorlensis is a South Iberian endemic (Spain), which is protected at European level, nationally as well as regionally, and considered vulnerable (VU) in accordance with IUCN criteria. This study has researched the genetic variability of seven representatives of *V. cazorlensis* populations, both inter- and intra-populationally, through the use of ISSR molecular markers and the sequencing of plastidial intergenic spacers. Results obtained from the ISSR markers indicate that these *V. cazorlensis* populations are not genetically impoverished, and that no clear genetic structure pattern exists. From the sequencing of plastidial intergenic spacers, the presence of different haplotypes has been observed, which becomes more evident as geographic distance among populations increases. Furthermore, there is a certain gene flow among them, more effective at a nuclear level, which could be mediated by *Macroglossum stellatarum* pollinator. The results obtained lead to the discussion of a series of conservation measures.

Key words: allogamy, dolomite endemic, gene flow, genetic diversity, ISSR markers, plastidial markers

Introduction

Viola cazorlensis Gand. is a South Iberian endemic (Spain), whose main populations are located in the mountains of Cazorla, Segura and Las Villas (Jaén), Sierra de Castril (Granada); more restricted ones are to be found in Sierra Mágina (Jaén), Sierra de Alcaraz (Albacete), and Sierra de Mojantes (Murcia). Herrera and Bazaga (2010) and Mota et al. (2008) consider this species as a strict habitat specialist, always on either cataclastic dolomites (rocks formed by fracturing and comminution) or on severely exposed (windswept) xeric dolomitic rocky soil. Its chromosome number is 2n = 20, independently of the population sampled to determine it (Leal, Ortiz, & Pajarón, 1980; Luque, 1981; Merxmüller & Lippert, 1977). Although its flowers are self-compatible, autogamous pollination rarely occurs, so it is considered an allogamous species (Blanca et al., 1999).

As regards its conservation and legal protection status, *Viola cazorlensis* is considered an endangered species within the category of 'vulnerable': VU A2acd (Moreno, 2008), which indicates that a reduction in population size has been verified over the past few years. This reduction is due to a series of threat factors related to the increase in number of wild herbivorous populations (Spanish Ibex and other ungulates), domestic herds (sheep and goats), and to other human activities. The latter include tourism and the development of infrastructures (roads, paths, etc.). As to legislation, this species is protected at European level, as well as nationally and regionally. As far as this regional aspect is concerned, this species is protected in the three Autonomous Communities where it is present (Andalucía, Castilla-La Mancha and Murcia) (Sánchez-Gómez, Carrión, Hernández, & Guerra, 2002). Moreover, all of its populations are located in areas protected under different legal figures, the majority in regional National Parks and, at European level, in sites of Community Importance (SCI) as well as in Special Bird Protection Areas (SBPA), within the Nature 2000 Network frame.

Several authors (Allendorf & Luikart, 2007; Frankham, Ballou, & Briscoe, 2009; Hedrick, 2005) have noted that it is necessary to know the different genetic features of species for the elaboration of effective and efficient

Correspondence to: Juan Francisco Jiménez. E-mail: fjimenez@um.es

ISSN 1477-2000 print / 1478-0933 online

http://dx.doi.org/10.1080/14772000.2015.1079275

conservation plans. Loss of genetic diversity might be related to a decrease in population viability, as this could have an impact on the ability of individuals to face biotic or abiotic changes (Booy, Hendriks, Smulders, van Groenendael, & Vosman, 2000; Fischer & Matthies, 1998; Hedrick, 2001). This might result in a reduction or even in the disappearance of populations, as well as in the appearance of such phenomena as endogamy and genetic drift. Along this line, different studies have been carried out on *Viola cazorlensis* Andalusian populations, which dealt with the species' response before herbivorism, together with the relationship between genetic (and epigenetic) and floral variability as well as fecundity (Herrera & Bazaga, 2008, 2009, 2010, 2011).

In order to clarify inter- and intra-population variability in *Viola cazorlensis*, this study focused on two types of molecular markers: Inter Simple Sequence Repeat (ISSR; Zietkiewicz, Rafalski, & Labuda, 1994), and plastidial DNA sequencing. The main aim of this study is to quantify *V. cazorlensis* values and the distribution of genetic variability at both inter- and intra-population levels, and therefore, to offer the most suitable measures for the conservation of the species.

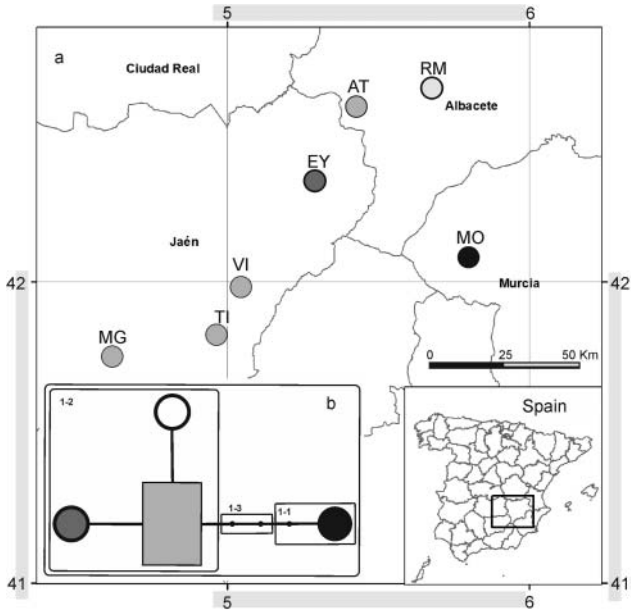

Fig. 1. a. Map of *Viola cazorlensis* distribution. Observed haplotypes have been placed over geographic location of populations. b. CpDNA haplotype network among four haplotypes obtained from 35 sampled individuals belonging to seven populations. Nested clade design was performed with ANECA v 1.2. Codes of populations are listed in Table 1.

Materials and methods

Sampling

Seven populations of *Viola cazorlensis* were sampled in 2011, covering the entire range of the species, but focusing on Sierra de Cazorla (Jaén), which is the principal nucleus of the species (Table 1 and Fig. 1). Typically, leaf material of 30−32 individuals per population was sampled, but in the Río Mundo population only 21 individuals could be sampled. To avoid DNA degradation, plant material was packed in zip-lock bags with silica gel until DNA extraction (Chase & Hills, 1991; Sytsma, Givnish, Smith, & Hahn, 1993).

DNA extraction and ISSR amplification

Total genomic DNA was extracted from silica gel dried leaf material following the 2 × Cetyl trimethyl ammonium bromide (CTAB) method (Doyle & Doyle, 1987), and cooled at −20°C until analysis.

ISSR reactions were performed in 25 μl, containing 10 mM Tris-HCl (pH 8.8, 25°C), 50 mM KCl, 2.5 mM $MgCl_2$, 200 mM each of dATP, dCTP, dGTP and dTTP, 15 ng primer, approximately 25 ng genomic DNA and one unit of Taq polymerase. A control, containing all the components except genomic DNA, was included in each set of

Table 1. *Viola cazorlensis* populations from which material was sampled. Location is displayed in 1 km UTM coordinates. N, Number of individuals sampled. P, Percentage of polymorphic loci. H_0, Intra-population genetic diversity. H_s, Genetic diversity within species. H_t, Total genetic diversity. G_{ST}, Genetic distance among populations.

Pop code	Pop name (Province)	Location (UTM coord)	N	P	H_0	H_s	H_t	G_{ST}
RM	Río Mundo (Albacete)	30S 568 4265	21	72	0.210			
AT	Arroyo del Tejo (Albacete)	30S 543 4259	32	85	0.236			
TI	Tíscar (Jaén)	30S 497 4183	30	82	0.230			
EY	El Yelmo (Jaén)	30S 530 4235	31	75	0.215			
MO	Mojantes (Murcia)	30S 580 4209	31	61	0.168			
MG	Sierra Mágina (Jaén)	30S 459 4172	32	63	0.177			
VI	Viñuelas (Jaén)	30S 505 4199	32	72	0.190			
TOTAL						0.215	0.257	0.104

reactions to confirm that no contamination had occurred. Amplifications were carried out in an Eppendorf Thermocycler under the following conditions: an initial cycle at 94°C for 2 min; 35 cycles of 30 s at 94°C, 30 s at different optimal annealing temperatures, 1 min at 72°C; a final cycle at 72°C for 10 min. The resulting reactions were analysed by electrophoresis on 1.5% agarose gel stained with ethidium bromide. Duplicate amplifications were conducted for each sample to ensure reproducible results and to minimize errors. Bands that could not be reproduced in any assays were not considered for further analysis. A total of 20 primers were initially examined for variability. Five of these primers (see online supplemental material, which is available from the article's Taylor & Francis Online page at http://dx.doi.org/10.1080/14772000.2015.1079275) showed variability between samples.

Gel images were captured with the Kodak Gel Logic System, and fragment sizes were determined by comparison to Hyperladder 50 bp (Bioline) using the Kodak 1D 3.6 software. The presence or absence of each ISSR fragment was treated as a binary character (coded 1 and 0, respectively) and used to construct the original data matrix. Bands showing the same gel mobility were assumed to be homologous. Following suggestions by Grosberg, Levitan, and Cameron (1996), no attempts were made to code for band intensity. DNA bands showing quantitative variation in brightness were scored as present, regardless of their intensities, and as absent if they were undetectable.

cpDNA sequencing

Two plastid regions (*trn*T-*trn*L and *trn*L-t*rnF* intergenic spacers) were amplified in five individuals from each population included in the ISSR analysis. Amplification reactions were conducted in 50 μl volumes containing approximately 20 ng of genomic DNA, 0.2 mM of each dNTP, 2.5 mM $MgCl_2$, 2 units of Taq Polymerase (Biotools), the buffer provided by the manufacturer and the primer combinations trnA-trnB for *trn*T-*trn*L region, and trnL-trnF for *trn*L-*trn*F intergenic spacer (Taberlet, Gielly, Patou, & Bouvet, 1991) at a final concentration of 0.4 μM. Reactions were performed in an Eppendorf Mastercycler using the following program: an initial cycle at 94°C for 3 min; 35 cycles of 30 s at 94°C, 30 s at 52°C, and 1 min at 72°C. A final cycle at 72°C for 8 min was included to terminate amplification products. Finally, 2 μl of the amplification products were visualized on 1.5% agarose gel and successful amplifications were cleaned with the GenElute PCR clean-up kit (Sigma-Aldrich, Madrid). For sequencing, purified PCR products were reacted with BigDye terminator cycle sequencing ready reaction (Perkin-Elmer, Applied Biosystems, Madrid) using amplification primers. For each product, both strands were sequenced.

Analysis of ISSR data

A number of parameters related to gene diversity have been calculated from the binary matrix, with the aid of POPGENE 1.3 software (Yeh, Yang, Boyle, Ye, & Mao, 1997), which allowed for the comparison of gene diversity values at inter- and intra-population levels. The parameters estimated were percentage of polymorphic loci, H_0 (intra-population genetic diversity), H_s (the diversity within species), and H_t (Total gene diversity; Nei, 1973). Genetic distance among populations (G_{ST}) was computed following Nei (1977).

On the other hand, the Analysis of Molecular Variance (AMOVA) indicates how genetic variation spreads over diverse hierarchical levels: intra-population, inter-population, and among five population groups created according to their geographic location. It also allowed for the obtainment of an estimated Wright Fixation Index (F_{ST}). Said groups were: (1) MO (Mojantes) population, (2) RM (Río Mundo) population, (3) AT (Arroyo del Tejo) population, (4) Populations located in Sierra Cazorla (TI, Tíscar; EY, El Yelmo; VI, Viñuelas), (5) and MG (Sierra Mágina). Mantel Tests correlated genetic diversity distribution with geographic distance between populations. Both analyses were performed with ARLEQUIN 3.5 software (Excoffier & Lischer, 2010). Cluster analyses were performed with PC-Ord 6 (McCune & Mefford, 2011). This sort of analysis requires selecting not only a linkage strategy, but also a measure of the distance or dissimilarity between samples (McCune & Grace, 2002; Peck, 2010). The most common among the latter are nearest neighbour (or single linkage), furthest neighbour (or complete linkage), UPGMA (Unweighted Pair Group Method with Arithmetic Mean) and Ward's method.

While the two first ones are antagonistic, UPGMA offers intermediate solutions, which explains why it is so widely implemented. Ward's method is also quite widespread in that it is more discriminative when determining grouping levels. As this last procedure is only compatible with Euclidean distance, said measure was used to generate distance semi-matrices from the binary data gross matrix (presence-absence of ISSR bands). Cluster analysis was applied both to the 209 individuals separately as well as to the said individuals grouped in each one of the seven populations under consideration. PAST 2.17c (Hammer, Harper, & Ryan, 2001) was implemented in order to calculate bootstrap support for each branch in the dendrograms. Only identical dendrograms and those offering scarce additional information have been excluded when displaying results. PC-Ord 6 was also used to show these seven populations in a multidimensional space, through a Principal Coordinates Analysis (PCO).

A Bayesian model-based analysis was performed to infer population structure with Structure version 2.2 (Falush, Stephens, & Pritchard, 2007; Pritchard,

Stephens, & Donnelly, 2000). The F model, based on an admixture ancestry model with correlated allele frequencies, was imposed to estimate the posterior probabilities [LnP(D)] of K groups (Pritchard & Wen, 2004) and the individual percentages of membership assigned to them according to their molecular multilocus profiles (Falush, Stephens, & Pritchard, 2003, 2007). Probabilities for a range of K were examined starting from 1 to the number of sampled populations plus one (K = 1−8), using a burn-in period and run length of the Markov chain Monte Carlo (MCMC) of 10^5 and 10^6 iterations respectively, and this was replicated 20 times. The results were uploaded into Structure Harvester (Earl & von Holdt, 2012, available at http://taylor0.biology.ucla.edu/struct_harvest/), which estimates the most likely K value (ΔK), following Evanno, Regnaut, and Goudet (2005) method. We used CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) to make a consensus of the results from the independent runs for the optimal K. For the consensus the Greedy option with random input order and 100,000 repeats was used. The consensus was visualized in DISTRUCT 1.1 (Rosenberg, 2004).

Analysis of cpDNA data

For each individual and sequenced DNA region, forward (5′−3′) and reverse (3′−5′) sequences were checked for inaccurate base calling using Chromas Lite 2.01 (Technelysium Pty Ltd, Shannon, Ireland). Consensus sequences were aligned using CLUSTALX (Thompson, Gibson, Plewniak, Jeanmougin, & Higgins, 1997). Bioedit (Hall, 1999) was used for minor manual alignment adjustments.

TCS v.1.21 (Clement, Posada, & Crandall, 2000) was used for a statistical parsimony network approach. In this analysis, each insertion/deletion (indel) was considered as a single mutation event, and therefore coded as substitutions (A or T) in the final alignment (Martínez-Nieto, Segarra-Moragues, Merlo, Martínez-Hernández, & Mota, 2013; Qiu, Guan, Fu, & Comes, 2009). The resulting haplotype network was nested into hierarchical clades using the automated implementation of Nested Clade Phylogeographic Analysis (NCPA, Templeton, Routman, & Phillips, 1995) provided by the program ANECA v.1.2 (Panchal, 2007). ANECA software implements both TCS v.1.21 and GEODIS v.2.5 (Posada, Crandall, & Templeton, 2000) and an inference key (Zhang, Tan, & Sota, 2006) which automates the inference process, providing a framework for replicating analyses in an objective way. Although NCPA has been criticized by some authors (Knowles & Maddison, 2002) it is a powerful method for reconstructing the phylogeography of a species where prior information is lacking (Templeton, 2004).

Results

ISSR molecular markers

Once 209 individuals were analysed using the 5 ISSR primers, a total of 100 bands was obtained, of which 98% were polymorphic. ISSR fragment sizes varied between 200 and 1600 bp. No population-exclusive phenotypes were found, which could be an indication of active or recent contact among the populations.

The percentage of polymorphic *loci* ranged from 61% of MO population to 85% of AT population (Table 1). Genetic diversity values obtained from Nei indexes (H_0) (Table 1) go from those shown by AT population (H_0 = 0.236) to those shown by MO population (H_0 = 0.168), while average gene diversity within populations (H_s) was 0.215, and the total gene diversity (H_t) was 0.257. These values indicate that populations were not genetically impoverished, and that there were no differences among them. AMOVA pointed in the same direction, and showed that the major part of the variations was distributed within populations (86.61%). The Fixation Index obtained through POPGENE (G_{ST}) offered a value of 0.104, whereas that obtained from AMOVA (F_{ST}) was 0.134. These values indicate that there was low to moderate genetic differentiation among populations. When the (F_{ST}) index was applied to pairs of populations (Table 2), it was observed that some populations − such as RM in relation with VI (F_{ST} = 0.232) or with MG (F_{ST} = 0.212) − were beginning to show a high genetic differentiation. Adversely, these differences were almost inexistent among the three populations within one group: AT, TI and EY. The Mantel test (r = 0.338; P = 0.075) revealed low and no significant correlation between genetic and geographic distances.

Cluster analysis of the 209 individuals offered no conclusive results. Even though a tendency for some populations to present their individuals in well-defined groups was observed throughout all analyses performed, it was also true that bootstrap values never surpassed 5%, not even in the case of better-defined population groups. Nonetheless, in these groups individuals belonging to one population appeared to gather in the same tree branch, with

Table 2. Fixation Indexes (F_{ST}) between *Viola cazorlensis* population pairs.

Pop	RM	AT	TI	EY	MO	MG	VI
RM	0						
AT	0.141	0					
TI	0.159	0.054	0				
EY	0.153	0.074	0.067	0			
MO	0.172	0.119	0.136	0.074	0		
MG	0.212	0.137	0.149	0.132	0.174	0	
VI	0.232	0.102	0.134	0.160	0.173	0.124	0

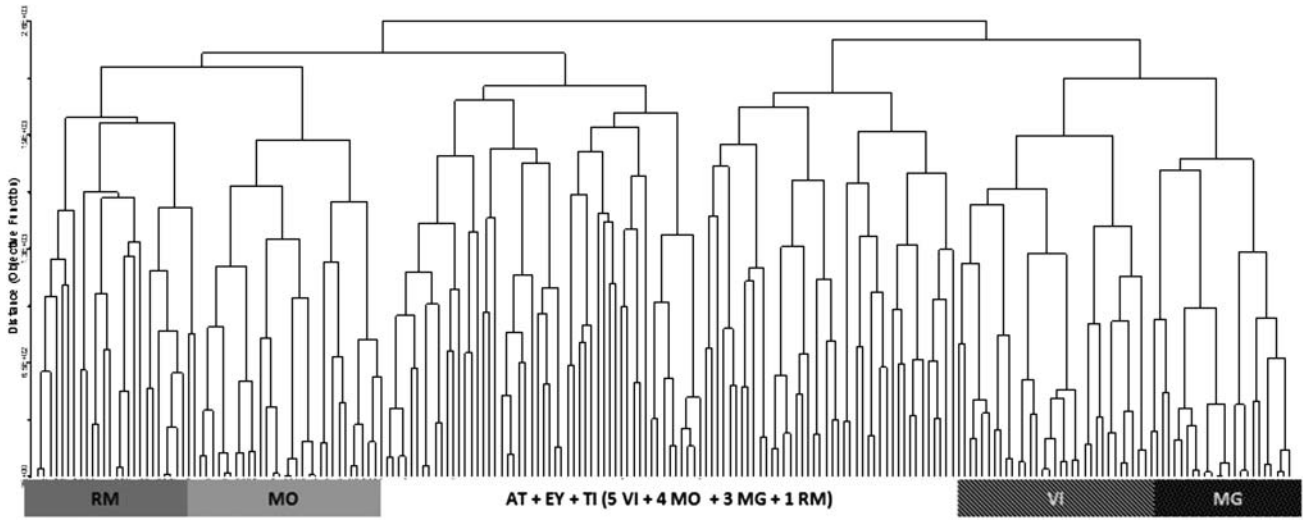

Fig. 2. Dendrogram constructed by Ward's method and based on ISSR data showing relationships among 209 individuals analysed (for abbreviations see Table 1).

hardly any intercalation of other populations' individuals. The dendrogram obtained from Ward's method (Fig. 2) could very well illustrate this point. Individuals were grouped in four main clusters. One of these clusters grouped most individuals in RM and MO populations in two close branches. The same situation happened for those individuals belonging to VI and MG populations. However, the other two clusters included a mixture of individuals from the remaining populations (AT, EY and TI) and, moreover, all the branches showed low statistical support.

As far as populations' dendrograms are concerned, all of the strategies showed the same result (Fig. 3), with the exception of the furthest neighbour tree; thus, a narrow relationship between EY and TI populations was reflected. Said populations also appeared as linked to RM and AT. These four populations integrated in one of the two main branches in the dendrograms; MG, MO and VI populations were located in the other. The latter seemed to be related to the RM, AT, EY and TI group in the case of the furthest neighbour. In the vast majority of cases, bootstrapping offered $>$ 50% values. PCO analysis for the populations (Fig. 4) also validated the results obtained from the cluster analysis.

Bayesian analysis of ISSR markers, carried out with STRUCTURE, showed the highest value for $K = 6$. As it can be noted, four of the populations (RM, MO, MG and VI) adjust relatively well to four of the groups established by the program. Residual populations (AT, TI and EY) showed a mixed proportion of membership to the remnant two clusters (Fig. 5).

Plastidial intergenic spacers

Alignment of 35 individuals for the two plastid regions yielded 1285 nucleotide sites (379, trnT-L; 906, trnL-F), of which 1279 were constant and six variable and parsimony-informative as well.

Four haplotypes were identified, three of them were exclusive for single populations (MO, EY and RM), and the other was shared by the rest of populations (Table 3). The haplotype from MO was linked to the most frequent haplotype by three missing haplotypes. Haplotypes from

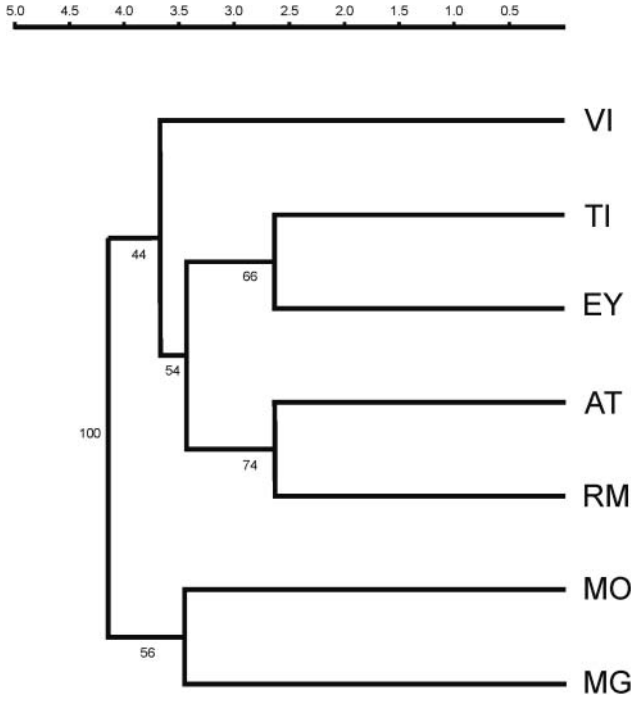

Fig. 3. Dendrogram UPGMA based on dissimilarity values showing the relationships between *Viola cazorlensis* populations. Numbers indicate bootstrap values obtained. Codes of populations are listed in Table 1.

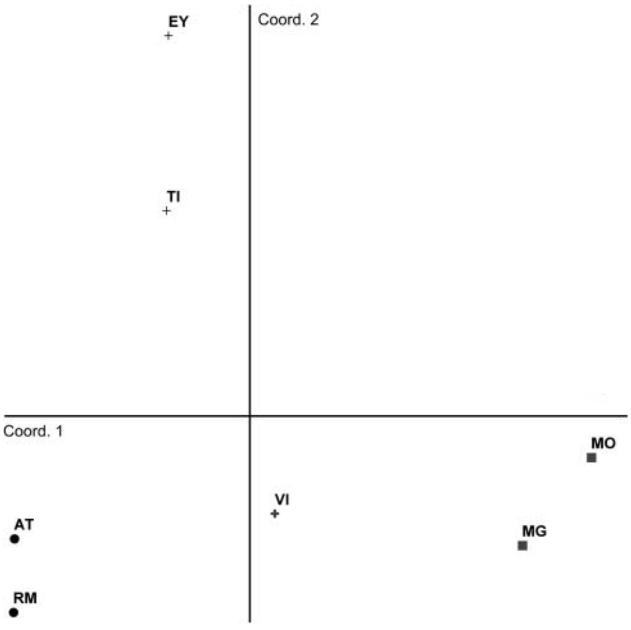

Fig. 4. Scatter plot of the seven sampled populations based on two first components of the PCO. The names of the localities are abbreviated according with the Table 1.

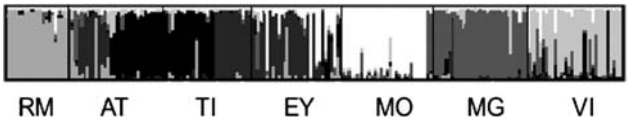

Fig. 5. Plot showing the membership of individuals for six predefined clusters obtained with STRUCTURE.

EY and RM were related to this frequent haplotype. The hierarchical nested design of the haplotype network identified three clades (Fig 1*b*). The 1-1 clade included the haplotype from MO and a missing haplotype. The 1-2 clade comprised the more frequent haplotype and the haplotypes from EY and RM that were connected to the former by a mutational step. The 1-3 clade was formed by two missing haplotypes, and it was in an intermediate position between 1-1 and 1-3 clades. NCPA suggested that no geographic structure of haplotypes occurred for clades 1-1 and 1-3, however it showed allopatric fragmentation for the 1-2 clade, as could be expected in narrow and isolated populations.

Table 3. Position of different mutations found for *Viola cazorlensis trn*T-F region.

		*trn*T-L			*trn*L-F		
Pop	Haplotype	267	466	467	709	759	980
RM	1	G	-	-	A	G	G
AT	2	G	A	-	A	G	G
TI	2	G	A	-	A	G	G
EY	3	G	A	-	-	G	G
MO	4	A	A	A	A	A	A
MG	2	G	A	-	A	G	G
VI	2	G	A	-	A	G	G

Discussion

Intra-population genetic diversity

The study of genetic diversity levels in plants can oftentimes be regarded as a complicated issue, for it can be affected not only by intrinsic factors related to the biology of the species, but also by extrinsic ones (Falk & Holsinger, 1991). Numerous studies have noted that those species with a more restricted geographic distribution usually show lower genetic diversity than more widespread congeners (Hamrick & Godt, 1990; Nybom, 2004; Piñeiro, Fuertes, Menezes, & Nieto, 2009), although this is not always so (Cole, 2003; Gitzendanner & Soltis, 2000). *Viola cazorlensis* exemplifies a species with a higher genetic diversity than might be expected. High genetic diversity maintained in rare plants is attributable to a number of factors (Zawko, Krauss, Dixon, & Sivasithamparam, 2001), such as recent reduction of population size plus insufficient time for isolation, or extensive, recurrent gene flow (Chiang et al., 2006; Maguire & Sedgley, 1997). In addition, different reproductive systems can also condition genetic diversity in a species. Predominantly outcrossing species have higher levels of variability within populations than selfing species (Hamrick & Godt, 1990; Loveless & Hamrick, 1984). Although flowers of *V. cazorlensis* are self-compatible and may occasionally produce fruits in absence of pollinators, the activity of these is essential for fruit set (Herrera, 1993). The value of intra-population genetic diversity found (H_s = 0.215) is very close to that reported by Nybom (2004) for endemic species studied with dominant markers. Comparisons with other studies are difficult since genetic diversity depends on numerous factors, such as life history, breeding system, growth life forms, geographic range and even the molecular marker used (Nybom, 2004; Powell et al., 1996). In spite of these complications, if we compare the results of studies for chasmophyte plants using dominant markers, it appears that the genetic diversity of *V. cazorlensis* is slightly higher than *Antirrhinum microphyllum* (obligate outcrosser, Torres, Iriondo, & Pérez, 2003) or far higher than *Antirrhinum subbaeticum* (selfing species, Jiménez, Sánchez-Gómez, Güemes, Werner, & Rosselló, 2002). In the case of special edaphic habitats, the value found for *V. cazorlensis* does not significantly differ from that of several gypsophile or subgypsophile plants (Jiménez & Sánchez-Gómez, 2012; Martínez-Nieto et al., 2013); it is even higher (Hickerson & Wolf, 1998; Pérez-Collazos & Catalán, 2008; Pérez-Collazos, Sánchez-Gómez, Jiménez, & Catalán, 2009).

Apart from this, genetic diversity values obtained are in line with Herrera and Bazaga (2008), which indicates that populations are not genetically impoverished although they show in fragmented groups and in relatively small sizes.

Inter-population genetic structure

The normal consequences of population fragmentation are a reduction in intra-population genetic diversity and increasing of inter-population differentiation (Young, Boyle, & Brown, 1996). It has generally been observed in outcrossing species that the degree of genetic differentiation between populations is smaller than in selfing species (Hamrick & Godt, 1996). Furthermore, to explain the genetic structure, emphasis has generally been placed on the importance of genetic interchange between populations, but if gene flow ceases, shared common ancestry and similar selective regimes could determine the genetic similarities between populations (Schaal, Hayworth, Olsen, Rauscher, & Smith, 1998). With two or three exceptions, *Viola cazorlensis* populations exhibit moderate levels of pairwise genetic differentiation. These results are in contrast with those obtained by Herrera and Bazaga (2008), but their study was restricted to Cazorla populations. Populations might have fragmented relatively recently and did not have enough time to differentiate; furthermore, the perennial nature of this species may have played a part in this low differentiation rate.

As regards the genetic structure of *Viola cazorlensis*, according to the analyses performed, there is not a strong geographic structure. The majority of analyses show only a slight difference among populations and point at the fact that both MG and MO populations are the most deviant, perhaps as a consequence of their lower genetic diversity. Dendrograms obtained for the seven populations (Fig. 3) reflect a population structure based on an average statistical support (bootstrapping), in which populations from Sierra de Cazorla, at large, are gathered in one group, and so-called peripheral ones (MG and MO) find themselves in another one. Nevertheless, it is true that Cazorla populations do not show a geographic structure and that in some analyses, VI population (Fig. 2; 209 individual cluster) is linked to MO and MG. This last group might result, as stated before, from markers for these populations (i.e. those not representing autapomorphia or which are synapomorphic for the seven populations). The absence of Sierra de Cazorla intermediate population sampling could be another possible explanation for this, even though Herrera and Bazaga (2008), in a study carried out with AFLP markers, noted that there was no genetic structure pattern for *V. cazorlensis* populations. According to these authors, all these populations would behave as a continuum and would be subject to an important gene flow. This flow might be mediated by *Macroglossum stellatarum*, a migratory hummingbird hawk-moth able to fly long distances (Pittaway, 1997−2008). No data exist to test the maximum dispersal distance for *M. stellatarum*, but the observations by Martins and Johnson (2007) suggest that hawk-moths can disperse quite widely (see also Stockhouse, 1973). According to Herrera and Bazaga (2008), this pollinator is well-represented in populations from Jaén, thus, it is only logical to conclude that it might also be so in the remaining *V. cazorlensis* area, and that it is, to a great extent, in charge of gene flow between these populations. *Macroglossum stellatarum*, could contribute to maintaining both intra-population genetic diversity levels, and could favour the poor geographic structure of the populations. Similar results have been found in *Oenothera harringtonii*, pollinated primarily by hawk-moths. Hawk-moths can travel up to 20 miles in just one night, and may therefore contribute significantly to long-distance gene flow among populations (Stockhouse, 1973). According to Skogen (personal communication), as regards *O. harringtonii*, long-distance pollination has widespread implications ranging from limiting population divergence, accelerating the spread of adaptive traits, disrupting gene complexes, and maintaining species cohesion.

Although *Macroglossum stellatarum*'s migratory routes that might be favouring contact between specific populations and the isolation of others are still unknown, spatial separation could account for lower degrees of genetic diversity in MG and MO peripheral populations, as diversity of most taxa decreases with distance from natural habitat (Martins & Johnson, 2007).

However, focusing on the information provided by plastidial sequences it is clear that higher F_{ST} values in population pairs do not correspond to the best differentiated haplotypes. This sort of incongruence between nuclear variation distribution and plastidial distribution has been explained for species with an anemogamous reproductive system, in which pollen can travel far (Stehlik, 2002). In spite of the fact that *Viola cazorlensis* is an entomogamous species, the presence of four haplotypes indicates that *Macroglossum stellatarum* could be really effective when keeping the gene flow between populations, but only at a nuclear level through pollen transport. At a plastidial level, this gene flow is not self-evident, possibly due to the absence of spreading systems for seeds at long distances (Herrera & Bazaga, 2008). The analysis of the different haplotypes indicates that MO population, the furthest geographically, is the most dissimilar of all the populations. These data might be related as all the populations, except for MO and MG, find themselves located in a continuous mountain range, oriented towards the southeast-northeast. Therefore, a pollinator's transit could be facilitated and the existence of favourable intermediate habitats where *V. cazorlensis* once existed cannot be discarded. Yet, MO population's geographic location, a relatively isolated mountain chain, might have been

impacting the gene flow with other populations for a longer period, thus favouring its differentiation. This could also be the case of MG, although trying to prove this theory is correct would require further research.

Implications on conservation

The ultimate goals of conservation are to ensure sustainable survival of populations and to preserve their evolutionary potential. The estimates of genetic diversity and genetic differentiation provide a basis for implementing efficient and practical conservation programmes for *Viola cazorlensis*. This species lives in narrow and fragmented populations which may lead to the loss of genetic diversity (Ellstrand & Elam, 1993; Young et al., 1996). Nevertheless, in the present study, genetic variability distribution analysis, as estimated by ISSRs, suggests that 86.4% of the total genetic variation is still harboured within populations. What is more, the number of individuals in any of the populations sample exceeds 500, the estimated minimum effective to avoid phenomena deriving from inbreeding depression or genetic drift (Franklin, 1980). This leads to the next conclusion: the species is not endangered yet, and, for conservation purposes, it would probably be sufficient to maintain the populations located across the whole distribution range, to ensure continued representation of the total genetic diversity.

Should a population lose members, the reintroduction of individuals from near populations might sustain its viability. Nevertheless, MO population presents the lowest genetic viability values and the most singular haplotype, which make it more susceptible to suffering the effects of genetic drift. Thus, it would be advisable to gather germplasm and to store it in suitable germplasm banks, as well as to perform a close follow-up of this population so as to avoid a decrease in the viability of its individuals and subsequent local extinction. According to the haplotype distribution, seeds from RM, EY, MG populations and some from the Cazorla nucleus should be collected, always taking into account that the number of individuals is limited in populations. These guidelines must be considered in the conservation policies for *Viola cazorlensis*. Similar recommendations have been suggested for other species linked to dolomite outcrops (Salmerón-Sánchez, Martínez-Nieto et al., 2014; Salmerón-Sánchez, Merlo et al., 2014). The high number of endemic and endangered species that gather in these habitats (Mota et al., 2008) require further research of this kind, where it would be specially interesting to go deeper into their behaviour as meta-populations.

Acknowledgements

We would like to thank Antonio Catalán Hernández, José Antonio López Donate and Eduardo Picazo from the Consejería de Agricultura de Castilla-La Mancha, Juan Bautista Vera Pérez from the University of Murcia, and Miguel Ángel Carrión Vilches from the Consejería de Presidencia de la Región de Murcia for their collaboration in data and samples gathering.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the Dirección General de Universidades PEPLAN PROJECT framework (Subproject 3: Molecular Markers in the Conservation and Management of Murcia's Flora).

Supplemental data

Supplemental data for this article can be accessed here.

ORCID

Juan Francisco Mota http://orcid.org/0000-0001-5754-279X

References

Allendorf, F. W., & Luikart, G. (2007). *Conservation and the genetics of populations*. Oxford: Wiley-Blackwell.

Blanca, G., Cabezudo, B., Hernández-Bermejo, J. E., Herrera, C. M., Molero Mesa, J., Muñoz, J., & Valdés, B. (1999). *Libro Rojo de la Flora Silvestre Amenazada de Andalucía. Tomo I: Amenazadas*. Sevilla: Consejería de Medio Ambiente, Junta de Andalucía.

Booy, G., Hendriks, R. J. J., Smulders, M. J. M., van Groenendael, J. M., & Vosman, B. (2000). Genetic diversity and the survival of populations. *Plant Biology, 2*, 379–395.

Chase, M. W., & Hills, H. H. (1991). Silica gel: An ideal material for field preservation of leaf samples for DNA studies. *Taxon, 40*, 215–220.

Chiang, Y. C., Hung, K. H., Schaal, B. A., Ge, X. J., Hsu, T. W., & Chaing, T. Y. (2006). Contrasting phylogeographical patterns between mainland and island taxa of the *Pinus luchuensis* complex. *Molecular Ecology, 15*, 765–779.

Clement, M., Posada, D., & Crandall, K. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology, 9*, 1657–1659.

Cole, C. T. (2003). Genetic variation in rare and common plants. *Annual Review of Ecology, Evolution, and Systematics, 34*, 213–237.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin, 19*, 11–15.

Earl, D. A., & von Holdt, B. M. (2012). Structure Harvester: A website and program for visualizing STRUCTURE output

and implementing the Evanno method. *Conservation Genetics Resources, 4*, 359−361.
Ellstrand, N. C., & Elam, D. R. (1993). Population genetic consequences of small population size: Implications for plant conservation. *Annual Review of Ecology, Evolution, and Systematics, 24*, 217−242.
Evanno, G., Regnaut, S,., & Goudet, J. (2005). Detecting the number of cluster of individuals using the software Structure: A simulation study. *Molecular Ecology, 14*, 2611−2620.
Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perfom population genetics analyses under Linux and Windows. *Molecular Ecology Resources, 10*, 564−567.
Falk, D. A., & Holsinger, K. E. (1991). *Genetics and conservation of rare plants*. New York: Oxford University Press.
Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics, 164*, 1567−1587.
Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes, 7*, 574−578.
Fischer, M., & Matthies, D. (1998). RAPD variation in relation to population size and plant fitness in the rare *Gentianella germanica* (*Gentianaceae*). *American Journal of Botany, 85*, 811−819.
Frankham, R., Ballou, J. D., & Briscoe, D. A. (2009). *Introduction to conservation genetics* (2nd ed.). Cambridge: Cambridge University Press.
Franklin, I. R. (1980). Evolutionary change in small populations. In M. E. Soulé, & B. Wilcox (Eds.), *Conservation biology: An evolutionary-ecological perspective* (pp. 135−149). Sunderland: Sinauer Associates.
Gitzendanner, M. A., & Soltis, P. (2000). Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany, 87*, 783−792.
Grosberg, R. K., Levitan, D. R., & Cameron, B. (1996). Characterization of genetic structure and genealogies using RAPD-PCR markers: A random primer for the novice and nervous. In J. Ferraris, & S. R. Palumbi (Eds.), *Molecular zoology. Advances, strategies, and protocols* (pp. 67−100). New York: Wiley-Liss.
Hall, T. A. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Window 95/98/NT. *Nucleic Acids Symposium Series, 41*, 95−98.
Hammer, Ø., Harper, D. A. T., & Ryan, P. D.(2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica, 4*, 1−9.
Hamrick, J. L., & Godt, M. J. W. (1990). Allozyme diversity in plants species. In A. H. D. Brown, M. T. Clegg, A. L. Kahler, & B. S. Weir (Eds.), *Plant population genetics, breeding and germplasm resources* (pp. 43−63). Sunderland, MA: Sinauer Associates.
Hamrick, J. L., & Godt, M. J. W. (1996). Conservation genetics of endemic plant species. In J. C. Avise & J. L. Hamrick (Eds.), *Conservation genetics: Case histories from nature* (pp. 281−304). New York: Chapman and Hall.
Hedrick, P. W. (2001). Conservation genetics: Where are we now? *Trends in Ecology & Evolution, 16*, 629−636.
Hedrick, P. W. (2005). *Genetics of populations*. Sudbury: Jones & Bartlett Learning.
Herrera, C. M. (1993). Selection of floral morphology and environmental determinants of fecundity in a hawk-moth pollinated violet. *Ecological Monographs, 63*, 251−275.
Herrera, C. M., & Bazaga, P. (2008). Population-genomic approach reveals adaptive floral divergence in discrete populations of a hawk moth-pollinated violet. *Molecular Ecology, 17*, 5378−5390.
Herrera, C. M., & Bazaga, P. (2009). Quantifying the genetic component of phenotypic variation in unpedigreed wild plants: Tailoring genomic scan for within-population use. *Molecular Ecology, 18*, 2602−2614.
Herrera, C. M., & Bazaga, P. (2010). Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytologist, 187*, 867−876.
Herrera, C. M., & Bazaga, P. (2011). Untangling individual variation in natural populations: Ecological, genetic and epigenetic correlates of long-term inequality in herbivory. *Molecular Ecology, 20*, 1675−1688.
Hickerson, L. L., & Wolf, P. G. (1998). Population genetic structure of *Arctomecon californica* Torrey & Frémont (Papaveraceae) in fragmented and unfragmented habitat. *Plant Species Biology, 13*, 21−33.
Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics, 23*, 1801−1806.
Jiménez, J. F., & Sánchez-Gómez, P. (2012). Molecular taxonomy and genetic diversity of *Moricandia moricandioides* subsp. *pseudofoetida* compared to wild relatives. *Plant Biosystems, 146*, 99−105.
Jiménez, J. F., Sánchez-Gómez, P., Güemes, J., Werner, O., & Rosselló, J. A. (2002). Genetic variability in a narrow endemic snapdragon (*Antirrhinum subbaeticum*, *Scrophulariaceae*) using RAPD markers. *Heredity, 89*, 387−393.
Knowles, L. L., & Maddison, W. P. (2002). Statistical phylogeography. *Molecular Ecology, 11*, 2623−2635.
Leal, J., Ortiz, A., & Pajarón, S. (1980). Números cromosómicos para la Flora Española 155−161. *Lagascalia, 9*, 269−272.
Loveless, M. D., & Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology, Evolution, and Systematics, 15*, 65−95.
Luque, T. (1981). Números cromosómicos para la Flora Española 210-213. *Lagascalia, 10*, 236−237.
Maguire, T. L., & Sedgley, M. (1997). Genetic diversity in *Banksia* and *Dryandra* (Proteaceae) with emphasis on *Banksia cuneata*, a rare and endangered specs. *Heredity, 79*, 394−401.
Martínez-Nieto, M. I., Segarra-Moragues, J. G., Merlo, M. E., Martínez-Hernández, F., & Mota, J. F. (2013). Genetic diversity, genetic structure and phylogeography of the Iberian endemic Gypsophila struthium (Caryophyllaceae) as revealed by AFLP and plastid DNA sequences: Connecting habitat fragmentation and diversification. *Botanical Journal of the Linnean Society, 173*, 654−675.
Martins, D. J., & Johnson, D. S. (2007). Hawkmoth pollination of aerangoid orchids in Kenya, with special reference to nectar sugar concentration gradients in the floral spurs. *American Journal of Botany, 94*, 650−659.
McCune, B., & Grace, J. B. (2002). *Analysis of ecological communities*. Gleneden Beach, OR: MjM Software.
McCune, B., & Mefford, M. J. (2011). *PC-ORD. Multivariate analysis of ecological data*. Version 6. Gleneden Beach, OR: MjM Software.
Merxmüller, H., & Lippert, W. (1977). Veilchenstudien, V-VII. *Mitteilungen der Botanischen Staatssammlung München, 10*, 236−237.
Mota, J. F., Medina-Cazorla, J. M., Navarro, F. B., Pérez-García, F. J., Pérez-Latorre, A., Sánchez-Gómez, P., … Merlo, M.

E. (2008). Dolomite flora of the Baetic Ranges glades (South Spain). *Flora, 203*, 359−375.

Moreno, J. C. (2008).(coord.). *Lista Roja 2008 de la flora vascular española*. Madrid: Dirección General de Medio Natural y Política Forestal (Ministerio de Medio Ambiente, y Medio Rural y Marino), y Sociedad Española de Biología de la Conservación de Plantas.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA, 70*, 3321−3323.

Nei, M. (1977). F-statistics and analysis of gene diversity in subdivided populations. *Annals of Human Genetics, 41*, 225−233.

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology, 13*, 1143−1155.

Panchal, M. (2007). The automation of nested clade phylogeographic analysis. *Bioinformatics, 23*, 509−510.

Peck, J. E. (2010). *Multivariate analysis for community ecologists: Step-by-step using PC-ORD*. Gleneden Beach, OR: MjM Software.

Pérez-Collazos, E., & Catalán, P. (2008). Conservation genetics of the endangered Iberian steppe plant *Ferula loscosii* (Apiaceae). *Plant Biology, 10*, 492−501.

Pérez-Collazos, E., Sánchez-Gómez, P., Jiménez, J. F., & Catalán, P. (2009). The phylogeographical history of the Iberian steppe plant *Ferula loscosii* (Apiaceae): A test of the abundant-centre hypothesis. *Molecular Ecology, 18*, 848−861.

Piñeiro, R., Fuertes, J., Menezes, M., & Nieto, G. (2009). Low genetic diversity in the rare Madeiran endemic *Armeria maderensis* (*Plumbaginaceae*). *Folia Geobotanica, 44*, 65−81.

Pittaway, A. R. (1997-2008). Sphingidae of the Western Palaearctic. Available from http://tpittaway.tripod.com/sphinx/list.htm (accessed 4 September 2015)

Posada, D., Crandall, K. A., & Templeton, A. R. (2000). GeoDis: A program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology, 9*, 487−488.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., & Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding, 2*, 225−238.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*, 945−959.

Pritchard, J. K., & Wen, W. (2004). *Documentation for structure software: Version 2*. Chicago: University of Chicago.

Qiu, Y., Guan, B., Fu, C., & Comes, H. P. (2009). Did glacials and/or interglacials promote allopatric incipient speciation in East Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dysosma versipellis*. *Molecular Phylogenetics and Evolution, 51*, 281−293.

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes, 4*, 137−138.

Salmerón-Sánchez, E., Martínez-Nieto, M. I., Martínez-Hernández, F., Garrido-Becerra, J. A., Mendoza-Fernández, A. J., Gil de Carrasco, … Mota, J. F. (2014). Ecology, genetic diversity and phylogeography of the Iberian endemic plant *Jurinea pinnata* (Lag.) DC.(Compositae) on two special edaphic substrates: Dolomite and gypsum. *Plant Soil, 374*, 233−250

Salmerón-Sánchez, E., Merlo, M. E., Medina-Cazorla, J. M., Pérez-García, F. J., Martínez-Hernández, F., Garrido-Becerra, J. A., … Mota, J. F. (2014). Variability, genetic structure and phylogeography of the dolomitophilous species *Convolvulus boissieri* Steud. (Convolvulaceae) in the Betic ranges, inferred from AFLPs, cpDNA and ITS sequences. *Botanical Journal of the Linnean Society, 176*, 506−523.

Sánchez-Gómez, P., Carrión, M. A., Hernández, A., & Guerra, J. (2002). *Libro Rojo de la flora silvestre protegida de la Región de Murcia. Tomo II*. Murcia: Consejería de Agricultura, Agua y Medio Ambiente, Dirección General del Medio Natural.

Schaal, B. A., Hayworth, D. A., Olsen, K. M., Rauscher, J. T., & Smith, W. A. (1998). Phylogeographic studies in plants: Problems and prospects. *Molecular Ecology, 7*, 465−474.

Stehlik, I. (2002). Glacial history of the alpine herb *Rumex nivalis* (Polygonaceae): A comparison of common phylogeographic methods with nested clade analysis. *American Journal of Botany, 89*, 2007−2016.

Stockhouse, R. E. (1973). Biosystematic studies of Oenothera L. subgenus Pachylophus. Ph.D. dissertation. Fort Collins, Colorado, USA: Colorado State University.

Sytsma, K. J., Givnish, T. J., Smith, J. F., & Hahn, W. J. (1993). Collection and storage of land plant samples for macromolecular comparisons. *Methods in Enzymology, 224*, 23−37.

Taberlet, P., Gielly, L., Patou, G., & Bouvet, J. (1991). Universal primers for amplification of three noncoding regions of chloroplast DNA. *Plant Molecular Biology, 17*, 1105−1109.

Templeton, A. R. (2004). Statistical phylogeography: Methods of evaluating and minimizing inference errors. *Molecular Ecology, 13*, 789−809.

Templeton, A. R., Routman, E., & Phillips, C. A. (1995). Separating population structure from population history − A cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambyastoma tigrinum*. *Genetics, 140*, 767−782.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research, 24*, 4876−4882.

Torres, E., Iriondo, J. M., & Pérez, C. (2003). Genetic structure of an endangered plant, *Antirrhinum microphyllum* (*Scrophulariaceae*): Allozyme and RAPD analysis. *American Journal of Botany, 90*, 85−92.

Yeh, F. C., Yang, R. C., Boyle, T., Ye, Z. H., & Mao, J. X. (1997). *POPGENE, the user-friendly hareware for population genetic analysis*. Edmonton: Molecular Biology and Biotecnology Centre, University of Alberta.

Young, A., Boyle, T., & Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution, 11*, 413−418.

Zawko, G., Krauss, S. L., Dixon, K. W., & Sivasithamparam, K. (2001). Conservation genetics of the rare and endangered *Leucopogon obtetus* (Ericaceae). *Molecular Ecology, 10*, 2389−2396.

Zhang, A. B., Tan, S., & Sota, T. (2006). Autoinfer 1.0: A computer program to infer biogeographical events automatically. *Molecular Ecology Notes, 6*, 597−599.

Zietkiewicz, E., Rafalski, A., & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics, 20*, 176−183.

Associate Editor: Nadia Bystriakova