



# Molecular evidence of species- and subspecies-level distinctions in the rare *Orchis patens* s.l. and implications for conservation

Jacopo Calevo<sup>1,2</sup> · Roberta Gargiulo<sup>2</sup> · Leif Bersweden<sup>2,3</sup> · Juan Viruel<sup>2</sup> · Cristina González-Montelongo<sup>4</sup> · Khellaf Rebbas<sup>5</sup> · Lamia Boutabia<sup>6</sup> · Michael F. Fay<sup>2,7</sup>

Received: 26 October 2020 / Revised: 8 February 2021 / Accepted: 11 February 2021 /  
Published online: 25 February 2021  
© The Author(s) 2021

## Abstract

Characterizing genetic diversity and structure of populations is essential for the effective conservation of threatened species. *Orchis patens* sensu lato is a narrowly distributed tetraploid species with a disjunct distribution (i.e., Northern Italy, North Africa and the Canary Islands), which is facing a severe decline. In this study, we evaluated levels of genetic diversity and population structuring using 12 new nuclear microsatellite markers. Our analyses of genetic differentiation based on multiple approaches (Structure analysis, PCA analysis, and *F*-statistics using the ploidy-independent *Rho*-index) showed that gene flow is low across the range of *O. patens* s.l., particularly in the Canary Islands. Clear differences in allele frequencies between Italy, Algeria and the Canary Islands underlie the genetic differentiation retrieved. Our study provides support for the recognition of *O. canariensis* as a sister species to *O. patens* and the separation of the Italian populations as a new subspecies of *O. patens*. Despite the high heterozygosity values found in all populations (ranging from 0.4 to 0.7), compatible with the tetraploid status of the species, small population sizes and reduced gene flow will be likely detrimental for the different populations in the long term, and we recommend immediate conservation actions to counteract further fragmentation and population decline.

**Keywords** Disjunct distribution · Orchids · Population genetics · Simple Sequence Repeats (SSRs)

---

Communicated by Daniel Sanchez Mata.

---

✉ Jacopo Calevo  
jacopo.calevo@unito.it

Extended author information available on the last page of the article

## Introduction

As we are facing the sixth age of extinction (Swarts and Dixon 2009; Barnosky et al. 2011), an increasing number of species are threatened by decreased population sizes, habitat degradation and habitat fragmentation. Endangered species are in most cases characterised by small, fragmented and declining populations (Brooks et al. 2002). Such limited groups of individuals are at high risk of extinction owing to decreased genetic diversity and the growing effect of biological stochasticity (Fahrig 2003; Henle et al. 2004). There is strong evidence, indeed, that genetic factors, such as low genetic diversity and consequently reduced reproductive fitness, are driving threatened taxa to extinction (Spielman et al. 2004; Vellend and Geber 2005; Garner et al. 2020). Preserving genetic diversity is therefore one of the essential requirements to ensure population survival (Hoban et al. 2020; Laikre et al. 2020), especially under anthropogenic and environmental pressures, including climate change (Hewitt 2004; Jump et al. 2009). Rare species with low genetic diversity are more vulnerable to environmental and biological stresses compared to species with higher genetic diversity (Leimu et al. 2006; Honnay and Jacquemyn 2007; Aguilar et al. 2008), and small, isolated and declining populations of rare species, under climate change, may also become maladapted (Borrell et al. 2020). Habitat fragmentation or the complete loss of the habitat are major reasons for depletion of genetic variation in plants (Young et al. 1996; Frankham et al. 2002). Generally, species with narrow distribution ranges or small and/or isolated populations are more affected by loss of their natural habitat, and they are more susceptible to extinction due to depletion of genetic diversity than those species that are widely distributed and have relatively large, non-fragmented populations (Hamrick and Godt 1996; Frankham et al. 2002; Pandey et al. 2015), although life-history traits may play an important role in buffering genetic diversity loss (Honnay and Jacquemyn 2007).

Spatial distribution of genetic variation within a species may offer useful insights into fundamental ecological and environmental processes and information important for identifying distinct genetic groups (if present) across its range (Allendorf and Luikart 2007; Mkare et al. 2017; Gargiulo et al. 2019a). The delineation of conservation units for an endangered species is critical not only for their long-term survival, but also for directing the prioritization of conservation efforts (Petit et al. 1998; Schwartz et al. 2007; Volkmann et al. 2014; Mkare et al. 2017; Médail and Baumel 2018). Characterizing genetic diversity and structure of populations is useful in informing assisted gene flow (Borrell et al. 2020), or to identify high priorities for effective conservation of threatened species (Fay 2018; Väli et al. 2019). This is even more urgent for orchids, as approximately half of the 1641 currently assessed in the IUCN Red List are in the categories Vulnerable, Endangered and Critically Endangered (Fay 2020; IUCN, 2020).

The genus *Orchis*, which includes about 20–25 species, is divided into two subgenera, *Orchis* and *Masculae*, with the first commonly called the “anthropomorphic group” (Kretzschmar et al. 2007). The topology of Bateman et al. (2003) shows the non-anthropomorphic *Orchis* species are well-supported as monophyletic. Cytological examinations of different *Orchis* taxa have revealed a main chromosome number of  $2n=42$  (e.g. Mrkvicka 1992; Aedo and Herrero 2005), but exceptions with aberrant values have been found in section *Robustocalcare* ( $2n=40$ ; Hautzinger 1978), which includes the only tetraploid species in the genus, *O. patens*. However, more recent counts indicated tetraploid *O. patens* with  $2n=84$  (Pridgeon et al. 2001; Bernardos et al. 2006).

*Orchis patens* sensu lato is a species with a disjunct geographic range, with *O. patens* subsp. *patens* occurring in Italy (Liguria region) and North Africa (Algeria and Tunisia),

and *O. patens* subsp. *canariensis* occurring in the Canary Islands (Orsenigo et al. 2016). Several authors have considered *O. canariensis* (syn. *O. patens* subsp. *canariensis*) and *O. patens* (syn. *O. patens* subsp. *patens*) as sister species (e.g., Bateman et al. 2003; Bernardos et al. 2006), unlike WCSP (2020), in which *O. canariensis* is considered as the basionym of *O. patens* subsp. *canariensis*. Although not resolving relationships in *O. patens* *s.l.*, ITS-based phylogenetic analyses showed that *O. patens* *s.l.* is the sister group of *O. spitzelii* (Bateman et al. 2003), a species predominantly found in Mediterranean regions, including northern Africa and western Asia, with outlying populations on the island of Gotland. A recent study on seed micromorphology of *O. patens* subsp. *patens*, however, identified some similarities with the *O. mascula* group (Calevo et al. 2017), a group of taxa distributed in Europe, north-western Africa and western Asia, thus opening up the hypothesis of an allotetraploid origin. *Orchis patens* is assessed in the Italian (Orsenigo et al. 2016) and European IUCN Red Lists (Rankou 2011) as Endangered (EN), and in the Mediterranean IUCN Red List (Calevo et al. 2018) as Vulnerable (VU), but all the assessments agree that, given the clear morphological differences between subspecies and the lack of studies investigating genetic diversity, clarifying taxonomic uncertainties at species and subspecies levels should be a priority for the conservation of *O. patens* *s.l.*

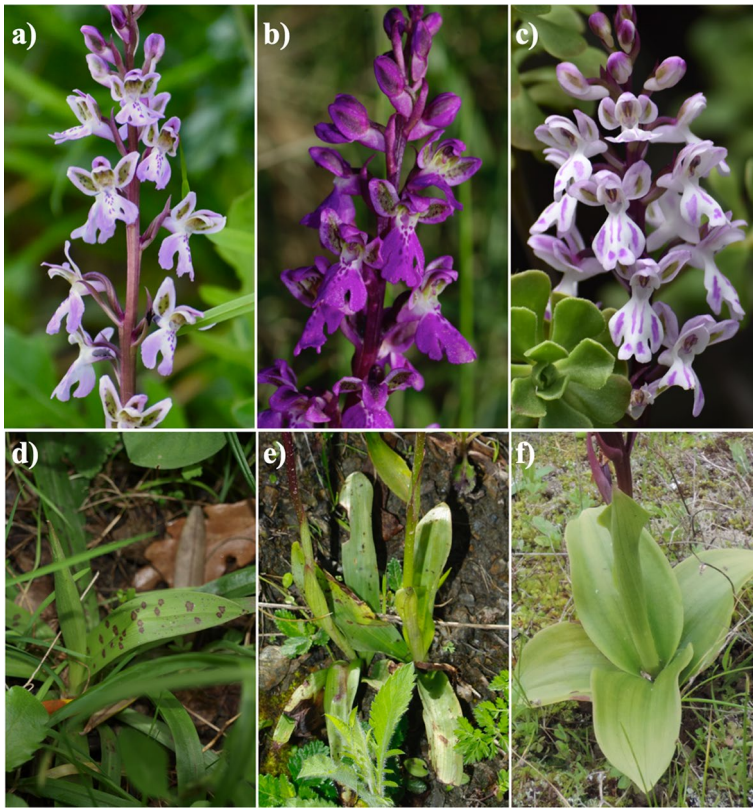
Moreover, although *O. patens* is reported as the only tetraploid species of the genus (Mrkvicka 1992; Pridgeon et al. 2001), it is not clear whether different cytotypes (ploidy levels) occur (Bernardos et al. 2006) and codominant markers, especially microsatellites, may provide indirect information through the allele counts observed (Besnard and Baali-Cherif 2009; Gompert and Mock 2017). In this study, we designed new nuclear microsatellite markers to elucidate genetic structure and diversity in *O. patens* *s.l.*, which will aid in the development of appropriate conservation strategies. In particular, our aims were to i) ascertain the occurrence of gene flow among populations, with a focus on indirectly detecting differences in ploidy; ii) evaluate the infraspecific delimitation and taxonomic rank of *O. patens* subsp. *patens* and *O. patens* subsp. *canariensis*; iii) verify the potential use and informativity of the markers herein developed in other congeneric species.

We refer to Canarian populations as *O. canariensis* and to African and Italian populations as *O. patens* *s.s.* from here on.

## Materials and methods

### Plant material

We collected leaf samples from populations of both subspecies, representative of the distribution range of the species. Samples of *O. patens* *s.s.* were collected in north-western Italy (in figures referred to as OP samples, Fig. 1a and d), in four localities of the Liguria region (16 samples from Romaggi, 11 samples from Breccanecca, 12 samples from Portofino and six samples from Capreno), and in two localities of Algeria (Fig. 1b and e): ten samples from Ait Zikki, in the Bejaia region (in figures referred to as OPAZ samples), and ten samples from Ain Talhi, in the Souk-Ahras region (in figures referred to as OPSA samples), carefully avoiding natural hybrids. Samples of *O. canariensis* (Fig. 1c and f) were collected from Tenerife (ten samples, in figures referred to as TN), Gran Canaria (ten samples, in figures to referred as GC) and La Gomera (nine samples, in figures referred to as LG). Collection of plant material was performed after requesting permission from the competent authorities and respecting regional rules.



**Fig. 1** Inflorescences and leaves of *Orchis patens* s.s. from Italy (**a**, **d**) and Algeria (**b**, **e**), and *O. canariensis* from Gran Canaria (**c**) and La Gomera (**f**). It is possible to notice differences in morphology and colours. In particular, leaves in *O. canariensis* (**f**) are larger and more rounded than those lanceolate of the subspecies *patens* (**d**, **e**). Furthermore, leaves in the Italian populations (**d**) present pale spots that are instead absent in the African populations (**e**), the one described by Desfontain (1799). Pictures by Jacopo Calevo (**a**, **c**, **d**), Khellaf Rebbas (**b**, **e**) and Ángel B. Fernandez (**f**)

## DNA extraction and Illumina sequencing

Total genomic DNA was extracted from silica-dried leaf tissues using the DNeasy Plant Mini Kit (QIAGEN) according to the manufacturer's instructions, after disruption with TissueLyser. The quantity and quality of DNA samples were assessed by spectrophotometry (ND-1000 Spectrophotometer NanoDropH; Thermo Scientific, Wilmington, Germany) and by electrophoresis on a 1% agarose gel.

A genome-skimming approach (Straub et al. 2012; Dodsworth 2015) was used for the detection of microsatellites (Viruel et al. 2018; Landoni et al. 2020). Genomic libraries were prepared using NEBNext® UltraTM II DNA Library Prep Kit for Illumina® (New England Biolabs, Ipswich, MA, United States) with AMPure XP magnetic beads for size selection (300–350 bp) and NEBNext® Multiplex Oligos for Illumina® (Dual Index Primer Sets I and II) as barcodes for simultaneous sequencing (Viruel et al. 2019). Library quality was evaluated using a Quantus™ fluorometer (Promega Corp.) and an Agilent 4200 TapeStation (Agilent Technologies, Santa Clara, CA, United States). Multiplexed

libraries were then sequenced on an Illumina MiSeq (Illumina, Inc.) lane. Trimmomatic v0.35 (Bolger et al. 2014) was used to remove low-quality reads and adapter sequences flagged by FastQC v0.11.7 (Andrews 2010).

## Microsatellite design and analysis

Microsatellites were identified and primers designed with MSATCOMMANDER 1.0.8 (Faircloth 2008); duplicate reads and those containing more than one SSR array were manually checked and removed. Primers fulfilling the following criteria were selected: optimal size 18–22 bp, not directly flanking the SSR motif, lacking ambiguous bases, low self- and pair-product complementary parameters, amplicon expected size < 300 bp, and melting temperature ( $T_m$ ) difference < 1.5 °C (Viruel et al. 2018).

Preliminary screening was carried out for 40 nuclear primer pairs on a 2% agarose gel. After testing and optimization of primers with a subset of samples, the final set of markers consisted of 12 microsatellite primer pairs (Table 1); other primer pairs were discarded due to issues with reproducibility and PCR product yield, possibly resulting from the tetraploid status of the species. Representative sequences for the microsatellite sequences were deposited in GenBank (accession numbers MT799099–MT799110).

PCR mix was prepared by adding 6  $\mu$ L of (2 $\times$ ) DreamTaq PCR MasterMix (ThermoFisher Scientific), 0.5  $\mu$ L of 0.4% (w/v) bovine serum albumin (BSA), 0.5  $\mu$ L of each primer (10  $\mu$ M), 1  $\mu$ L of template DNA (ca. 10 ng/ $\mu$ L), and H<sub>2</sub>O up to a final volume of 10  $\mu$ L. Amplifications were performed with the following PCR conditions: 4 min of initial denaturation at 94 °C followed by 25 to 35 cycles (see Table 1 for primer details) of 30 secs of denaturation at 94 °C, 30 secs of annealing (see Table 1 for temperature details), 45 secs of elongation at 72 °C and an extra extension of 10 min at 72 °C, with the exception for primer pairs Opat\_27 for which it was necessary to perform a touch-down PCR by decreasing annealing temperature of 0.2 °C at each cycle (see Table 1). One  $\mu$ L of PCR product was added to 10  $\mu$ L HiDi™ formamide (Applied Biosystems) and 0.15  $\mu$ L 500 ROX Size Standard (Applied Biosystems) and capillary electrophoresis was run on an ABI3730 DNA Analyzer (Applied Biosystems).

Fragment sizing was performed in GeneMapper v.5 (Applied Biosystems) and peaks were visually checked before assigning final allele sizes. We inferred the minimum chromosome number from the maximum number of peaks detected; this procedure has been shown to give results consistent to flow cytometry (Besnard and Baali-Cherif 2009; Stark et al. 2011; Nemorin et al. 2013; Teixeira et al. 2014; Donkpegan et al. 2015; Gompert and Mock 2017). We used the R package *polysat* v1.7 (Clark and Jasieniuk 2011; Clark and Schreier 2017) to explore segregation patterns in our microsatellite loci. Understanding whether loci are polysomic or are inherited as separate subgenomes can give some clues about the origin of polyploidy in *O. patens* and, most importantly, can inform about the correct treatment of the dataset for subsequent analyses (Dufresne et al. 2014; Hardy 2016; Meirmans et al. 2018; Meirmans and Liu 2018). To avoid the bias induced by population structure, we kept individuals from Italy, Algeria and the Canary Islands as three separate populations before performing the analyses in *polysat*; we used the function *processDataSetAll* with default parameters and excluding the locus Opat\_38 (see results).

To check the transferability of the primers hereby developed, we also tested them in five other ‘non-anthropomorphic’ *Orchis* taxa, *O. provincialis*, *O. mascula* subsp. *mascula*, *O. mascula* subsp. *ichnusae*, *O. spitzelii*, and *O.  $\times$ fallax* (*O. patens*  $\times$  *O. provincialis*), and four anthropomorphic species, *O. anthropophora*, *O. simia*, *O. militaris* and *O. purpurea*.

**Table 1** Primer pairs designed for nuclear microsatellite loci in *Orchis patens s.l*

Primers ID	Msat motif	Primer sequences	Expected length (bp)	Length (bp)	$T_a$ (°C)	PCR Cycles
Opat_11	(ACT) <sub>13</sub>	(FAM) GGTGAAATGAGT CGGAAGTC ACCGCACTGCAAATAAAT CC	221	81–239	55	30
Opat_13	(CCG) <sub>9</sub>	(JOE) GAACCGTACCGAGTG AGAC CTGAGCTATGAACCAGTC CC	150	133–147	55	30
Opat_14	(ACAT) <sub>10</sub>	(JOE) GGGTAGGGCAGCTAT AAAGG AGAGAAGGTGGAAGAAGG TG	235	209–239	55	35
Opat_20	(AC) <sub>17</sub>	(FAM) CTAACGCGGTCC AATTCTTC GGTACTATCGGGTTGGCC	112	91–116	59	30
Opat_27	(CT) <sub>14</sub>	(FAM) GTTCACAACATTGCG AGAGG TCTTGGGAGGTTACAGCT TC	173	129–183	60–0.2	30
Opat_29	(CTT) <sub>11</sub>	(JOE) ATGGTTCCTCTGGTT CTTCG CAATAGCCTTCTTTGCCGTC	130	110–139	53	25
Opat_30	(AG) <sub>13</sub>	(FAM) TGTCACCACGGA ACAGTTAC CTCTCCAAACTCCTCCAA CC	173	152–218	53	30
Opat_32	(CT) <sub>20</sub>	(FAM) CCATCACATTGACTG CATCTG GAATTCACGTCTCTGGGCA	197	168–204	53	25
Opat_34	(CT) <sub>14</sub>	(FAM) AGGCAGTCTGA TTCGATATG GAACACGACATTGCCCAGA	174	153–193	52	30
Opat_36	(CT) <sub>8</sub>	(JOE) TATCCACTGCTCGCT ATGTG CACATCAGCTTCGATGGA GA	134	127–177	53	30
Opat_38	(CT) <sub>10</sub>	(JOE) TGCAACAACCACTTT CTCAC CTAGCACAAGGATGCTGA GT	115	101–103	53	25
Opat_39	(GA) <sub>15</sub>	(FAM) GGAAGTTAGTCCCTC CGAAC CTCCAAGTCAATTCGCA TCT	147	138–236	53	30

GenBank accession numbers: MT799099–MT799110.  $T_a$ , annealing temperature

## Genetic diversity and structure

As we did not have prior information about the extent of gene flow in *O. patens s.l.*, we conducted the analyses both by grouping individuals based on geographical regions (i.e., Italy, Algeria and the Canary Islands) and on sampling localities (i.e., nine sampling localities across the three regions).

Standard population genetic parameters such as mean number of alleles ( $Num$ ), effective number of alleles ( $Eff\_num$ ), observed heterozygosity ( $H_O$ ), gene diversity ( $H_S$ ) and inbreeding coefficient  $G_{IS}$  based on the 12 nuclear microsatellites were estimated using the program GenoDive v.3 (Meirmans and Van Tienderen 2004; Meirmans 2020). In addition, we performed a hierarchical analysis of molecular variance (AMOVA; Meirmans 2012) based on the  $Rho$ -statistics ( $\rho$ -statistics), which is ploidy-independent (Ronfort et al. 1998). All analyses in GenoDive were performed by applying the correction for unknown dosage of the alleles. Test for deviations from Hardy–Weinberg proportions was not attempted since it was not possible to record allelic dosage information for all the loci (Meirmans et al. 2018; Gargiulo et al. 2019b). We used the programme SPAGeDI v.1.5 (Hardy and Vekemans 2002) to produce a matrix of pairwise  $Rho$  values testing significance with 20,000 permutations; in addition, we estimated the selfing rate per locality using the standardized identity disequilibrium (Hardy 2016). Monomorphic loci were excluded from the computations.

We further analyzed genetic structure using different approaches. In *polysat*, we constructed a neighbor-joining tree based on a square dissimilarity matrix and we performed a principal component analysis (PCA) based on Bruvo distances (Bruvo et al. 2004). The model-based clustering procedure implemented in Structure v2.3.4 (Pritchard et al., 2000) was used to evaluate the number of different genetic clusters,  $K$ . We applied the Recessive Alleles model (Falush et al. 2007) to handle missing allelic dosage information. Preliminary analyses were conducted both with the Admixture model and the Non-Admixture model, testing  $1 < K < 8$ . The final analysis was conducted with  $K$  ranging from 1 to 8, with a  $5 \times 10^5$  burn-in period,  $5 \times 10^5$  MCMC runs, and ten iterations per each  $K$  value, using the Non-Admixture model for the total dataset, and the Admixture model for subsets based on sampling locations and infraspecific delimitations. The most likely value of  $K$  was evaluated using  $\Delta K$  (Evanno et al. 2005) and the estimators introduced by Puechmaile (2016) for uneven datasets (MedMedK, MedMeanK, MaxMedK and MaxMeanK) implemented in STRUCTURESELECTOR (Li and Liu 2018). Structure results were summarized with CLUMPAK (Kopelman et al. 2015). Simulations studies have shown that Structure is particularly robust compared to other clustering methods when dosage information is incomplete, even in the presence of mixed ploidies (Stift et al. 2019).

## Results

### Microsatellite design and analysis

Selected primer pairs, microsatellite motifs, and size ranges are reported in Table 1. The 12 newly designed nuclear microsatellite primer pairs amplified all samples and the resulting loci were variable with some exceptions. In particular, locus Opat\_38 was not variable for the Italian *O. patens* but showed at least two alleles (Table 2) in the

**Table 2** Locus-based estimates of genetic diversity for the *Orchis patens* s. l. populations

Loci ID (populations)	<i>N</i>	<i>Num</i>	<i>Eff_num</i>	<i>H<sub>O</sub></i>	<i>H<sub>S</sub></i>	<i>G<sub>IS</sub></i>	<i>Peaks</i>
<i>Italy</i>							
	45						
Opat_11		9	3.978	0.859	0.754	- 0.139	4
Opat_13		3	2.118	0.683	0.532	- 0.285	3
Opat_14		5	3.354	0.852	0.707	- 0.205	4
Opat_20		4	1.571	0.280	0.371	0.246	2
Opat_27		6	2.319	0.689	0.573	- 0.201	4
Opat_29		5	2.123	0.313	0.540	0.421	3
Opat_30		6	4.601	0.893	0.788	- 0.132	4
Opat_32		9	5.488	0.854	0.824	- 0.036	4
Opat_34		8	4.051	0.837	0.759	- 0.103	3
Opat_36		6	2.234	0.526	0.559	0.059	3
Opat_38*		1	1.000	0.000	0.000	-	1
Opat_39*		8	2.896	0.820	0.659	- 0.245	3
Mean (excluding monomorphic)		6.273	3.157	0.691	0.642	- 0.056	
<i>Algeria</i>							
	20						
Opat_11		11	4.418	0.900	0.786	- 0.145	3
Opat_13		5	2.463	0.721	0.605	- 0.192	3
Opat_14		7	3.527	0.842	0.728	- 0.156	3
Opat_20		6	3.867	0.258	0.788	0.672	2
Opat_27		5	1.317	0.213	0.251	0.154	4
Opat_29		7	4.449	0.463	0.809	0.428	2
Opat_30		10	5.758	0.792	0.844	0.062	4
Opat_32		10	4.894	0.925	0.808	- 0.145	3
Opat_34		11	6.617	0.771	0.869	0.113	4
Opat_36		3	1.311	0.162	0.251	0.352	2
Opat_38*		2	1.487	0.000	0.364	1.000	1
Opat_39*		9	4.090	0.875	0.768	- 0.140	3
Mean		7.167	3.683	0.577	0.656	0.121	
<i>Canary Islands</i>							
	29						
Opat_11		5	2.995	0.848	0.673	- 0.260	3
Opat_13		2	1.075	0.066	0.073	0.090	2
Opat_14		4	2.145	0.718	0.540	- 0.331	3
Opat_20		13	6.497	0.480	0.872	0.450	2
Opat_27		15	10.416	0.621	0.925	0.329	4
Opat_29		8	3.487	0.422	0.736	0.426	2
Opat_30		22	4.984	0.937	0.808	- 0.160	3
Opat_32		8	5.098	0.330	0.835	0.604	2
Opat_34		11	6.055	0.491	0.859	0.428	2
Opat_36		2	1.231	0.000	0.201	1.000	1
Opat_39*		1	1.000	0.000	0.000	-	1
Mean (excluding monomorphic)		9	4.398	0.491	0.716	0.257	

*N* number of individuals studied per population, *Num* average number of alleles per locus, *Eff\_num* the effective number of alleles, *H<sub>O</sub>* observed heterozygosity, *H<sub>S</sub>* gene diversity, *G<sub>IS</sub>* inbreeding coefficient. Peaks: maximum number of peaks observed in individual electropherograms

\*Notice that *Opat\_38* was monomorphic in the Italian populations of *O. patens* and polymorphic in Algeria. No amplification product was detected in the samples from the Canary Islands. *Opat\_39* was monomorphic



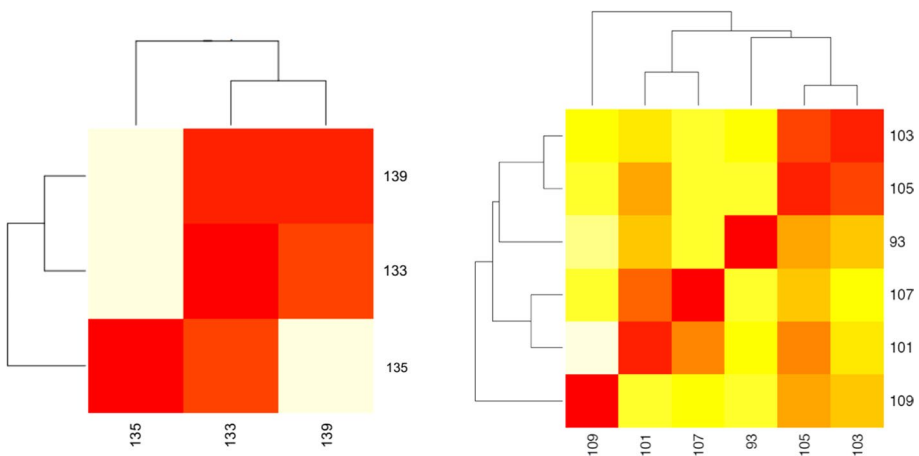
**Table 2** (continued)

in the samples from the Canary Islands, but polymorphic in *O. patens* from Italy and Algeria. All mean values of genetic diversity indices were computed after the exclusion of *Opat\_38* and *Opat\_39*

Algerian populations, and it did not show any amplification product in the samples from the Canary Islands. Locus *Opat\_39*, with multiple alleles in *O. patens* from Italy and Algeria (Table 2), was not variable for samples from the Canary Islands.

The number of alleles as inferred by the maximum number of peaks in the electropherograms was four in all populations. Every genotyped individual had at least one locus displaying three peaks, except GC8 in the Canary Islands, which exhibited two peaks at every locus (Table 2). Therefore, we did not find strong evidence of populations with mixed ploidy. The analysis in *polysat* showed that polysomic inheritance prevailed at all loci, with one exception represented by locus *Opat\_36* in the Algerian samples (Fig. 2). Consequently, data were treated as for autopolyploid species in the subsequent analyses.

To verify the transferability of these microsatellite loci, we tested the primers in other *Orchis* taxa (Table 3). *Opat\_27*, *Opat\_32*, *Opat\_34*, *Opat\_36* and *Opat\_38* did not amplify *O. provincialis* samples (with locus *Opat\_38* being variable only in Algerian *O. patens*). All the primers were transferable to *O. ×fallax*, the natural hybrid between *O. patens* and *O. provincialis*, with the only exception of *Opat\_13* and *Opat\_38* that showed no variability (they were not variable in *O. provincialis* and *O. patens*, respectively). Six out of 12 primer pairs successfully amplified and showed variability also in *O. mascula* subsp. *mascula* and *O. mascula* subsp. *ichnusae* samples. Three primer pairs amplified and showed variability also in *O. spitzelii*. Furthermore, *Opat\_13* amplified and was variable also in four ‘anthropomorphic’ species: *O. anthropophora*, *O. simia*, *O. militaris* and *O. purpurea* (Bersweden et al. in rev.).



**Fig. 2** Examples of heatmaps of the P-values for correlations between alleles in Algerian populations; the colour gradient represents the range of P-values (red: 0; light yellow: 1). Left: locus *Opat\_36* showing a disomic pattern in the Algerian individuals. Right: locus *Opat\_20* showing a polysomic pattern

**Table 3** Result of cross-species transferability in *Orchis* taxa using the 12 microsatellite primers developed from *Orchis patens s.l.*

Species	Opat markers											
	11	13	14	20	27	29	30	32	34	36	38	39
<i>O. patens</i> (Italy)	X	X	X	X	X	X	X	X	X	X	O	X
<i>O. patens</i> (Algeria)	X	X	X	X	X	X	X	X	X	X	X	X
<i>O. canariensis</i>	X	X	X	X	X	X	X	X	X	X	-	O
<i>O. provincialis</i>	O	O	X	X	-	O	O	-	-	-	-	X
<i>O. mascula</i>	X	X	X	NA	NA	X	O	X	NA	X	NA	O
<i>O. mascula</i> subsp. <i>ichnusae</i>	X	X	X	NA	NA	X	O	X	NA	X	NA	O
<i>O. spitzelii</i>	NA	NA	NA	NA	NA	NA	NA	X	NA	X	NA	X
<i>O. ×fallax</i>	X	O	X	X	X	X	X	X	X	X	O	X

X-markers successfully amplified a variable PCR product; O-markers successfully amplified but the product was not variable;—markers that did not show any amplification pattern; NA untested

## Genetic diversity and structure

Genetic diversity parameters for each population and locality are summarized in Tables 2 and 4. The results indicate that genetic diversity indices are similar across all populations and localities, except Tenerife, which shows slightly lower heterozygosities and number of alleles (Table 4). The number of effective alleles ( $Eff\_Num$ ) per population varied between 3.2 (in Italy, averaged over 11 polymorphic loci) and 4.4 (in the Canary Islands, averaged over ten polymorphic loci). The lowest  $H_o$  was found in the Canary Islands populations (0.49, reflecting the low  $H_o$  observed in Tenerife; Table 4) and the highest in the Italian populations (0.69);  $H_s$  ranged between 0.64 in Italy and 0.72 in the Canary Islands (Table 2). Negative values of  $G_{IS}$  were found in all populations, possibly reflecting the tetraploid status of the species; however, locus-specific high  $G_{IS}$  were found for Opat\_20 and Opat\_29, possibly indicating null alleles or drop-out (i.e., alleles that do not produce amplification by PCR). The highest selfing rate, as

**Table 4** Genetic diversity indices as estimated from samples of *Orchis patens s. l.* collected in each locality

Locality	Sample Size	$Num$	$Eff\_num$	$H_o$	$H_s$	$G_{IS}$	Selfing Rate (SE)*
Portofino (Italy)	12	5.1	3.085	0.653	0.622	-0.049	0.072 (0.159)
Breccanecca (Italy)	11	4.4	3.196	0.706	0.648	-0.09	0
Romaggi (Italy)	16	4.6	2.942	0.675	0.639	-0.056	0.098 (0.139)
Capreno (Italy)	6	4.3	2.990	0.689	0.666	-0.035	0 (0.157)
Bejaia (Algeria)	10	6	3.497	0.653	0.69	0.054	0.288 (0.187)
Souk-Ahras (Algeria)	10	5.1	3.149	0.557	0.634	0.122	0.035 (0.238)
Tenerife (Canary Islands)	10	2.8	2.070	0.44	0.439	-0.003	0 (0.367)
Gran Canaria (Canary Islands)	10	4.7	2.921	0.503	0.62	0.188	0
La Gomera (Canary Islands)	9	6.8	3.893	0.535	0.66	0.189	0.160 (0.223)

Monomorphic loci were excluded from the computations

\*Selfing rate based on the standardised identity disequilibrium, as calculated in SPAGeDI, using allelic phenotypes. SE: standard error based on jackknife over loci

computed in SPAGeDI was found in the Algerian locality OPAZ; absence of selfing was estimated in Italian and Canarian localities (Table 4).

The hierarchical AMOVA (Table 5) showed most of the variation at the individual level, which is expected when using polymorphic (and polyploid) microsatellite loci. Moreover, differentiation among groups was the highest for the group formed respectively by Algerian/Italian populations and the Canary Islands populations (*Rho\_ct*: 0.421) and the lowest when one of the groups included Italian and Canarian populations (*Rho\_ct*: 0.162). In parallel, the matrix of pairwise *Rho*-values indicated strong differentiation among populations in Italy, Algeria and the Canary Islands and also among the individuals collected in different islands of the archipelago (Fig. 3). When performing AMOVA on the Italian subpopulations, for which we had more localities and individuals (i.e., 45), the results (Table 6) pointed to higher levels of variation within (89.8%) rather than between the subpopulations (10.2%).

The neighbor-joining tree based on the square matrix of genotype dissimilarities obtained with *polysat* indicated three clear clusters (Fig. 4). In particular, the cluster including all the samples from the Canary Islands (light blue) is divergent from the other two. The same pattern was observed in the principal component analysis (PCA) based on Bruvo distances (Fig. 5).

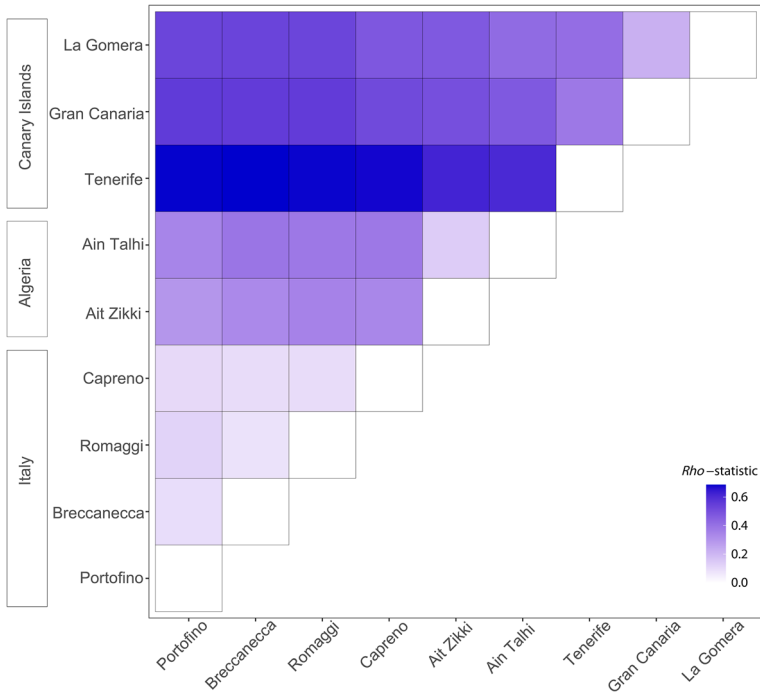
DeltaK values computed on the output of the Structure analysis indicated a strong signal for  $K=2$  (Supplementary information Fig. S1), whereas the other estimators (MedMedK, MedMeanK, MaxMedK and MaxMeanK) indicated three clusters (Supplementary information Fig. S2). The clusters underlined, once again, a strong genetic differentiation between *O. patens s.s.* and *O. canariensis*) (Fig. 6).

The analysis conducted on the Italian and Algerian populations identified four potential clusters (Supplementary information Fig. S3), whereas the analysis conducted on the samples from the Canary Islands identified three clusters (Supplementary information Fig. S6). In both cases, it was possible to observe an important division driven by regionality (Fig. 7) which, in the Canary Islands, are consistent with the three islands

**Table 5** Hierarchical analyses of molecular variance (AMOVA) of *Orchis patens s. l.*, using a ploidy-independent Infinite Allele Model (*Rho*-statistics)

Groups	Source of Variation	% Vari	F-stat	F-value
<i>OP</i>	Within locality	0.468	<i>Rho_st</i>	0.532
<i>AL</i>	Among localities within group	0.119	<i>Rho_sc</i>	0.203***
<i>CI</i>	Among groups	0.413	<i>Rho_ct</i>	0.413***
<i>OP-AL</i>	Within locality	0.415	<i>Rho_st</i>	0.585
<i>CI</i>	Among localities within group	0.163	<i>Rho_sc</i>	0.282***
	Among groups	0.421	<i>Rho_ct</i>	0.421***
<i>AL-CI</i>	Within locality	0.477	<i>Rho_st</i>	0.523
<i>OP</i>	Among localities within group	0.288	<i>Rho_sc</i>	0.376***
	Among groups	0.235	<i>Rho_ct</i>	0.235
<i>OP-CI</i>	Within locality	0.482	<i>Rho_st</i>	0.518
<i>AL</i>	Among localities within group	0.357	<i>Rho_sc</i>	0.425***
	Among groups	0.162	<i>Rho_ct</i>	0.162

*OP*-*O. patens*, Italy; *AL*-*O. patens* Algeria; *CI*-*O. canariensis*, Canary Islands; % Vari.-Percentage of variation, 9999 permutations; *Rho\_st* proportion of variation within locations; *Rho\_sc* proportion of variation among locations within groups; *Rho\_ct* proportion of variation among groups; \*\*\*  $p < 0.001$ . Monomorphic loci were excluded from the computation



**Fig. 3** Heatmap of the pairwise differentiation among *Orchis patens s.l.* in different sampling localities, using the Rho-statistic as calculated in SPAGeDI on the 12 microsatellite loci

**Table 6** Analysis of molecular variance (AMOVA) on the Italian subpopulations (45 individuals, four subpopulations) by using a ploidy independent Infinite Allele Model

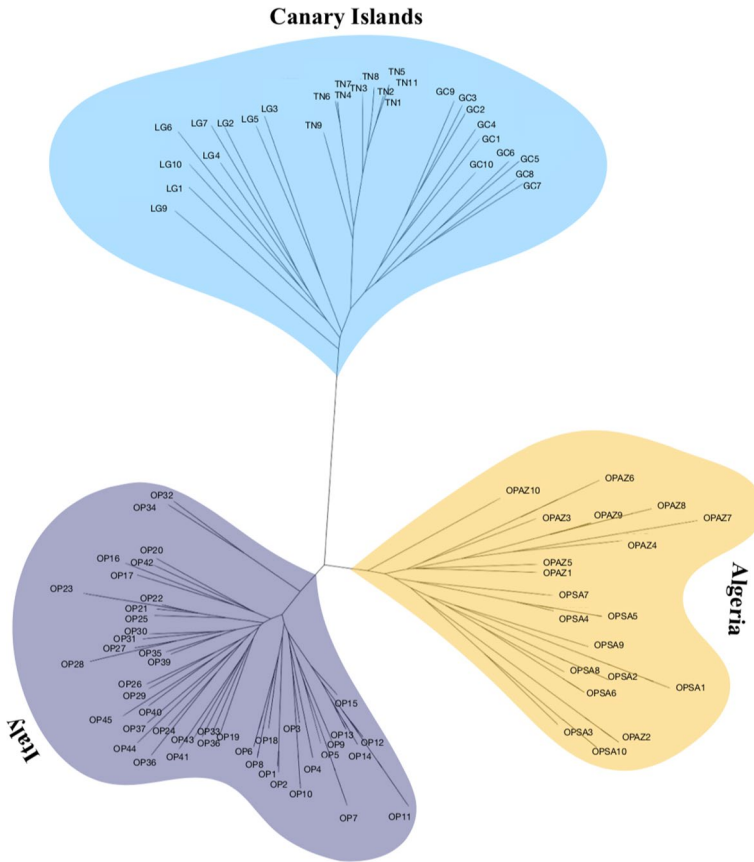
Source of variation	SSD	d.f	MS	Variance components	%Var	F-stat	F-value
Within population	73.023	41	1.781	1.781	0.898	–	–
Among population	11.926	3	3.975	0.202	0.102	Rho_st	0.102***

\*\*\*p < 0.001

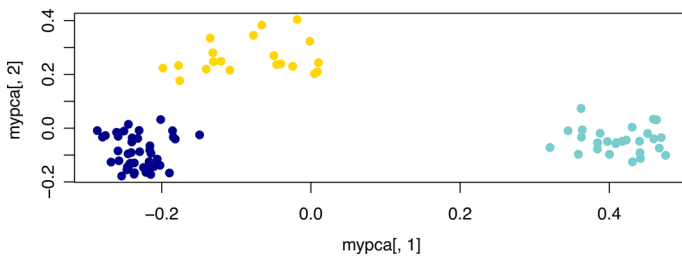
where samples were collected. For each of the Algerian and the Italian subsets, two clusters were identified (Supplementary information Figs. S4–S5) (Fig. 7).

### Discussion

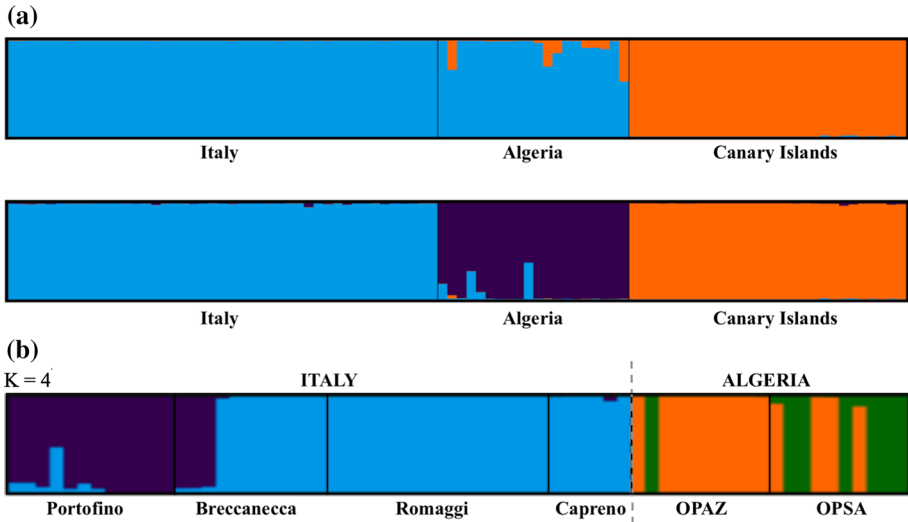
Preserving genetic variation is imperative to ensure the existence of any species, especially under anthropogenic and environmental stresses, including climate change (Hewitt 2004; Jump et al. 2009). For this reason, rare species with a narrow distribution or a fragmented habitat, and therefore potentially low genetic diversity, are a priority for studies aimed at investigating genetic diversity to establish appropriate conservation strategies. Orchidaceae



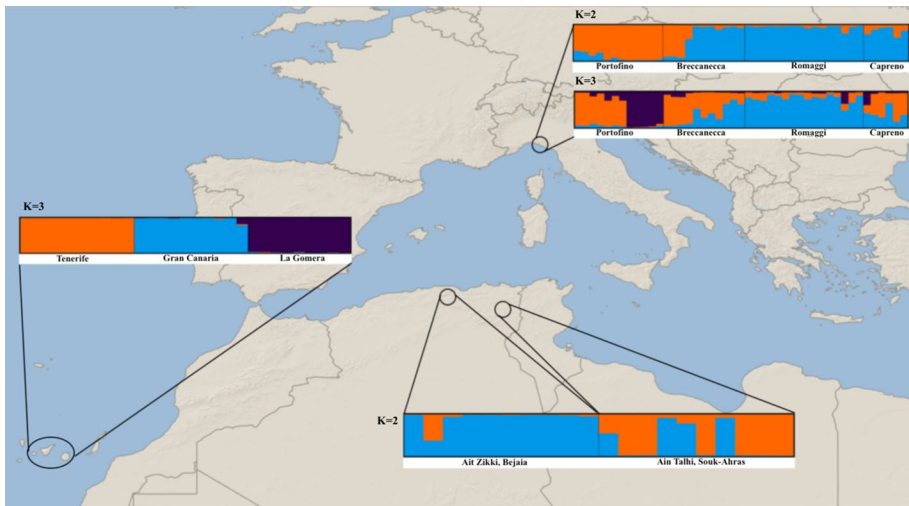
**Fig. 4** Neighbor-joining tree of *Orchis patens s.l.* based on a square dissimilarity matrix based on the 12 microsatellite loci. There are three clear clusters consistent with sampling countries. A “canariensis” cluster (TN, GC and LG samples; light blue), an Italian cluster (OP samples; blue) and an Algerian cluster (OPAZ and OPSA samples; gold)



**Fig. 5** Principal component analysis (PCA) of *Orchis patens s.l.* based on Bruvo distances among individuals, based on the 12 microsatellite loci. The three major groups are consistent with the Neighbor-joining tree, but it is possible to observe further diversification within populations. In particular, African samples (gold) seem to cluster in two minor groups (samples were collected in two different locations in Algeria). Canary Islands—light blue; Italy—blue; Algeria—gold)



**Fig. 6** **a** Structure clustering results for 94 samples of *Orchis patens s.l.* (45 samples from Italy, 20 from Algeria and 29 samples from Canary Islands) based on the variation at 12 microsatellite loci. Barplots are reported for K=2 (top) and K=3 (bottom), as determined using the Evanno method (Fig. S1) and the Puechmaille's estimators (Fig. S2); **b** Structure barplots of genetic differentiation in *O. patens s.s.* for all the Italian and Algerian sampling localities included in the present study



**Fig. 7** Structure barplots of genetic differentiation in *Orchis patens s.l.* for all the sampling localities included in the present study, based on the variation at 12 microsatellite loci. Each barplot is based on the most likely K value retrieved using the Evanno method and the Puechmaille's estimators (see Supplementary Figures)

are characterized by a diverse range of reproductive strategies (Cozzolino and Widmer 2005; Duffy and Johnson 2017; Schatz et al. 2020), geographic distribution and life histories that are reflected in a variety of patterns in the population genetic structure of different

species (Swarts et al. 2009). Clear species delimitations are important for orchids because many of the currently recognized species are of conservation concern and threats continue to increase due to habitat degradation (Cozzolino et al. 2003; Fay 2018). In this study we used 12 newly developed nDNA markers to determine the level of genetic differentiation among populations of the *O. patens* polyploid group, with a view to resolving taxonomic uncertainties for conservation purposes (Rankou 2011; Swarts et al. 2014; Orsenigo et al. 2016; Calevo et al. 2018; Médail and Baumele 2018).

## Genetic structure and taxonomic delimitation

Our analyses of genetic differentiation based on different approaches revealed the same clear clustering of samples (Figs. 3, 4, 5, 6), suggesting that the microsatellite data presented here carry a clear signal of population subdivision. In particular, samples from the Canary Islands are included in a genetic cluster divergent from the Italian and Algerian populations and are, in turn, characterized by genetic subdivision depending on the island of provenance (Figs. 3, 7; Table 5). The failure to amplify a product using Opat\_38 in the Canary Island populations further implies genetic divergence between *O. canariensis* and *O. patens* s.s. This locus was successfully amplified in the samples from Italy and Algeria. Even though both the Italian and Canarian plants were recently found to share the same mycobiont showing a high level of phylogenetic conservatism in their symbiotic interactions (Calevo et al. 2020), our results support the separation of *O. canariensis* as a distinct (probably sister) species, as initially described from Tenerife by Lindley (*Orchis canariensis*, Lindley 1835) based on morphological characters and as recognised by some authors (e.g., Aceto et al. 1999; Bateman et al. 2003) based on phylogenetic analyses. Analysis of the genetic structure based on microsatellites has proved to be informative in revealing cryptic infraspecific structure and subspecies delimitation, for example in the genera *Greenwayodendron* (Lissambou et al. 2019) and *Lophira* (Ewédjè et al. 2020).

The genetic structure of *O. canariensis*, as characterized by the number of subpopulations and the frequencies of different genetic variants (alleles) in each subpopulation (Chakraborty 1993), might be indicative of a high degree of genetic isolation among the subpopulations on the different islands. The strong clustering of Canarian populations (that mirrors the island of provenance) has been observed in other endemic plants of the islands such as *Lotus sessilifolius* (Yang et al. 2018) and *Bethencourtia* spp. (Rodríguez-Rodríguez et al. 2018), and the genetic differentiation among island populations may reflect the very beginning of speciation as proposed for the endemic *Kleinia neriifolia* by Sun and Vargas-Mendoza (2017). This, as suggested by the latter authors, might be expected as the ocean acts as a natural barrier among the islands.

At the same time, there is also some evidence of differentiation between the Italian and Algerian populations of *O. patens* s.s. (Figs. 3, 4, 5, 6, 7), which provides a genetic basis for the treatment of the Italian populations as a subspecies of *Orchis patens*. As mentioned for *O. canariensis*, microsatellite amplifications revealed a subpopulation-specific pattern (Tables 2 and 6) even among samples from Italy and Algeria, which may suggest that mutations in the priming sites are consistent with intraspecific divergence. This may be promoted by the tetraploid status of *O. patens*, in which the two chromosome sets undergo a range of somatic mutations and diverge if gene flow is reduced (Moody et al. 1993; Meir-mans and Liu 2018). Morphologically, the Italian populations are characterized by slightly spotted leaves (Fig. 1d), whereas spots are absent in African individuals (Fig. 1e) and were not mentioned in the first description of the species from Algeria (Desfontaines 1799).

Further morphological studies should be focused on detecting morphometric differences among the subspecies based on their entire phenotype, in order to strengthen their taxonomic circumscription and to collect more data about functional differences towards their conservation.

### Influence of polyploidy on current genetic diversity patterns

Although populations of *O. patens* have experienced natural and anthropogenic fragmentation, and fruit set is low and affected by hybridization with the sympatric species *O. provincialis*, the genetic structure found here reflects historical gene flow, especially between Italy and Algeria. *Orchis patens* may have been relatively common in the past within its narrow range and therefore the genetic effects of recent fragmentation and reduction in population number may yet to be detected. In polyploid species, loss of genetic diversity by drift is slowed down because of the availability of multiple allele copies compared to diploid species (Moody et al. 1993; Meirmans and Liu 2018). In addition, the effect of gene flow is amplified as more gene copies are involved. Consequently, the effects of fragmentation and genetic erosion may be not evident until more generations have elapsed. The levels of heterozygosity we found in *O. patens* are compatible with that of other polyploid species (Gargiulo et al. 2019b; Hagl et al. 2020), and in particular tetraploid orchid populations, such as *Gymnadenia conopsea* ( $H_O=0.77$ ; Gustafsson and Thorén 2001) and the *Dactylorhiza majalis* complex ( $H_O=0.77$  averaged across tetraploids; Balao et al. 2016). It is also possible that the current distribution of *O. patens* reflects long-distance dispersal from one of the three areas towards the other two, although our results do not allow us to speculate further on the most likely explanation.

Regarding the origin of *O. patens* s.l., the group has been considered sister of the *O. spitzelii* group based on an analysis of the ITS region (Bateman et al. 2003). These similarities have led to the hypothesis that *O. patens* could have been derived from the latter species as an autotetraploid, but a recent study based on seed micromorphology as a taxonomic tool also highlighted similarities with the *O. mascula* group (Calevo et al. 2017), suggesting the possibility of an allopolyploid origin. Our microsatellite analysis did not recover a segregation pattern typical of allopolyploid species (allopolyploidy is generally associated with disomic inheritance, at least for some loci). However, current segregation patterns are not always informative of past events, as the inheritance may shift from disomic to polysomic and vice-versa (Stift et al. 2008). However, more analyses are needed to better understand the origin of *O. patens* and the potential taxa involved in the allopolyploidization (Ramsey and Schemske 1998; Soltis et al. 2010; Barker et al. 2015).

### Conclusions and implications for conservation

Combining molecular evidence from our newly developed nuclear microsatellite markers with the geographical distribution of the studied populations, we propose the recognition of *O. patens* subsp. *canariensis* as a separate (possibly sister) species, i.e., *O. canariensis*, and the treatment of the Italian populations of *O. patens* as a distinct subspecies. Moreover, the primer set for the 12 microsatellite loci with transferability potential represents a useful tool for interspecific research in the genus *Orchis*. Our study of species delimitation within the *O. patens* s.l. complex can offer guidelines for further morphological and genetic investigations and conservation studies. Therefore, we recommend the following actions:



- *Orchis canariensis* should be assessed for the IUCN red list;
- the Mediterranean assessment of *O. patens* (Calevo et al. 2018) should be revised taking into account the taxonomic delimitation suggested here (which will probably lead to a higher category of threat), the evidence of restricted gene flow, narrow range and low number of individuals (Orsenigo et al. 2016);
- given that the availability of the main symbiotic fungus of *O. patens* s.l. at global scale seems not to be a limiting factor for their distribution (Calevo et al. 2020), conservation actions aimed at preventing further habitat fragmentation should be taken. In particular, local authorities and agencies in Algeria, Tunisia (where the natural population is already potentially extinct), Italy and the Canary Islands should institute measures to protect remaining individuals from damage and illicit collection.
- Further non-destructive sampling, even if difficult due to the rarity and low density of the species, may be required to better investigate the level of inbreeding and its contribution to the reproduction of the populations. Once reproduction mechanisms are clarified, strategies which facilitate pollination should be taken into account, also considering the low fruit-set of the species.

**Supplementary Information** The online version of this article (<https://doi.org/10.1007/s10531-021-02142-6>) contains supplementary material, which is available to authorized users.

**Acknowledgements** Authors are truly grateful to Robyn Cowan, László Csiba and Gabriele Rizzuto for technical support, and to Dr. Pierluigi Cortis, Dr. Errol Vela, Dr. Selim Amine, Dr. Giampiero Ciaschetti-National Park of Majella, and Dr. Ángel Fernández López-National Park of Garajonay for logistic and sampling support. We thank Prof. Richard Bateman and one anonymous reviewer for providing insightful comments which improved the quality of our manuscript.

**Authors' contributions** JC, CGM, KR and LBou. collected samples; JC, MFF, RG and JV designed research; JC and LBers. performed research; JC and RG analyzed data, JC wrote the paper. All authors reviewed and approved the final version of the manuscript.

**Funding** This research was conducted within the framework of the PhD program of Jacopo Calevo funded by the Italian Ministry of Education and Research (MIUR).

**Availability of data and material** Representative sequences were deposited in GenBank (Accession Numbers MT799099-MT799110).

## Compliance with ethical standards

**Conflicts of interest** The authors declare that they have no conflict of interest.

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

## References

- Aceto S, Caputo P, Cozzolino S, Gaudio L, Moretti A (1999) Phylogeny and evolution of *Orchis* and allied genera based on ITS DNA variation: morphological gaps and molecular continuity. *Mol Phylogenet Evol* 13:67–76
- Aedo C, Herrero A (2005) *Flora Iberica*. Vol. 21: Smilacaceae - Orchidaceae. Madrid
- Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* 17:5177–5188
- Allendorf FW, Luikart G (2007) Conservation and the genetics of populations. Blackwell Publishing, Malden
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Balao F, Tannhäuser M, Lorenzo M, Hedrén M, Paun O (2016) Genetic differentiation and admixture between sibling allopolyploids in the *Dactylorhiza majalis* complex. *Heredity* 116:351–361. <https://doi.org/10.1038/hdy.2015.98>
- Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA (2016) On the relative abundance of autopolyploids and allopolyploids. *New Phytol* 210:391–398
- Barnosky A, Matzke N, Tomiya S et al (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471:51–57. <https://doi.org/10.1038/nature09678>
- Bateman RM, Hollingworth PM, Preston J, Yibo L, Pridgeon AM, Chase MW (2003) Molecular phylogenetics and evolution of Orchidinae and selected Habenariinae (Orchidaceae). *Bot J Linn Soc* 142:1–40
- Bernardos S, García-Barriuso M, León-Arencibia MC, Reyes-Betancort A, González-González R, Padrón M, Amich F (2006) A cytotoxic study of three endemic orchids of the Canary Islands. *Ann Bot Fennici* 43:161–166
- Besnard G, Baali-Cherif D (2009) Coexistence of diploids and triploids in a Saharan relict olive: evidence from nuclear microsatellite and flow cytometry analyses. *C R Biol* 332:1115–1120
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina Sequence Data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Borrell JS, Zohren J, Nichols RA, Buggs RJ (2020) Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow. *Evol Appl* 13(1):161–175. <https://doi.org/10.1111/eva.12883>
- Brooks TM, Mittermeier RA, Mittermeier CG, Da Fonseca GAB, Rylands AB, Konstant WR, Flick P, Pilgrim J, Oldfield S, Magin G, Hilton-Taylor C (2002) Habitat loss and extinction in the hotspots of biodiversity. *Conserv Biol* 16:909–923
- Bruvo R, Michiels NK, D'Souza TG, Schulenburg H (2004) A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Mol Ecol* 13:2101–2106
- Calevo J, Giovannini A, Cornara L, Peccenini S, Monroy F (2017) *Orchis patens* Desf.: seed morphology of an endangered Mediterranean orchid. *Plant Biosyst* 151:770–774. <https://doi.org/10.1080/11263504.2017.1297335>
- Calevo J, Montagnani C, Véla E (2018) *Orchis patens*. The IUCN Red List of Threatened Species 2018:e.T175961A84665334. <https://doi.org/10.2305/IUCN.UK.2018-2.RLTS.T175961A84665334.en>
- Calevo J, Voyron S, Ercole E, Girlanda M (2020) Is the distribution of two rare *Orchis* Sister species limited by their main mycobiont? *Diversity* 12:262. <https://doi.org/10.3390/d12070262>
- Chakraborty R (1993) Analysis of genetic structure of populations: meaning, methods, and implications. In: Majumder PP (eds) *Human population genetics*. Springer, Boston
- Chase MW, Cameron KM, Barrett RL, Freudenstein JV (2003) DNA data and Orchidaceae systematics: a new phylogenetic classification. In: Dixon KW, Kell SP, Barrett RL, Cribb PJ (eds) *Orchid conservation*. Natural History Publications, Kota Kinabalu, pp 69–89
- Chase MW, Cameron KM, Freudenstein JV, Pridgeon AM, Salazar G, van den Berg C, Schuiteman A (2015) An updated classification of Orchidaceae. *Bot J Linn Soc* 177(2):151–174. <https://doi.org/10.1111/boj.12234>
- Clark LV, Jasieniuk M (2011) POLYSAT: an R package for polyploid microsatellite analysis. *Mol Ecol Resour* 11:562–566
- Clark LV, Schreier AD (2017) Resolving microsatellite genotype ambiguity in populations of allopolyploid and diploidized autopolyploid organisms using negative correlations between allelic variables. *Mol Ecol Resour* 17:1090–1103

- Cozzolino S, Noce ME, Musacchio A, Widmer A (2003) Variation at a chloroplast minisatellite locus reveals the signature of habitat fragmentation and genetic bottlenecks in the rare orchid *Anacamptis palustris* (Orchidaceae). *Am J Bot* 90:1681–1687. <https://doi.org/10.3732/ajb.90.12.1681>
- Cozzolino S, Widmer A (2005) Orchid diversity: an evolutionary consequence of deception? *Trends Ecol Evol* 20:487–494
- Desfontaines RL (1799) *Flora Atlantica, sive Historia plantarum, quae in Atlante, agro Tunetano et Algeriensi crescunt*, Parigi, L.G. Desgranges
- Dodsworth S (2015) Genome skimming for next-generation biodiversity analysis. *Trends Plant Sci* 20:525–527
- Donkpegan AS, Doucet JL, Dainou K, Hardy OJ (2015) Microsatellite development and flow cytometry in the African tree genus *Azelia* (Fabaceae, Caesalpinioideae) reveal a polyploid complex. *Appl Plant Sci* 3(1):1400097
- Duffy KJ, Johnson SD (2017) Specialized mutualisms may constrain the geographical distribution of flowering plants. *Proc R Soc B* 284:20171841
- Dufresne F, Stift M, Vergilino R, Mable BK (2014) Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Mol Ecol* 23:40–69
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Ewédjè E-EBK, Jansen S, Koffi GK, Staquet A, Pineiro R, Esaba RA, Obiang NLE, Dainou K, Biwolé AB, Doucet JL, Hardy OJ (2020) Species delimitation in the African tree genus *Lophira* (Ochnaceae) reveals cryptic genetic variation. *Conserv Genet*. <https://doi.org/10.1007/s10592-020-01265-7>
- Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Ann Rev Ecol Evol Syst* 34:487–515
- Faircloth BC (2008) MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol Ecol Resour* 8:92–94
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 7:574–578
- Fay MF (2018) Orchid conservation: how can we meet the challenges in the twenty-first century? *Bot Stud* 59:16
- Fay MF (2020) Orchids on the IUCN Global Red List. *Orchid Specialist Group Newsletter* 2020 (2):2–4. [https://www.iucn.org/sites/dev/files/content/documents/osg\\_newsletter\\_sept\\_final.pdf](https://www.iucn.org/sites/dev/files/content/documents/osg_newsletter_sept_final.pdf)
- Fay MF, Gargiulo R, Viruel J (2019) The present and future for population genetics, species boundaries, biogeography and conservation. *Bot J Linn Soc* 191:299–304
- Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to conservation genetics*. Cambridge University Press, Cambridge, UK
- Gargiulo R, Pironon S, Zheleznaya E, Sanchez MD, Balázs ZR, Podar D, Wilkinson T, Jäkäläniemi A, Kull T, Väre H, Fay MF (2019a) Phylogeography and post-glacial dynamics in the clonal-sexual orchid *Cypripedium calceolus* L. *J Biogeogr* 46:526–538. <https://doi.org/10.1111/jbi.13528>
- Gargiulo R, Worswick G, Arnold C, Pike LJ, Cowan RS, Hardwick KA, Chapman T, Fay MF (2019) Conservation of the threatened species, *Pulsatilla vulgaris* Mill. (Pasqueflower), is aided by reproductive system and polyploidy. *J Hered* 110(5):618–628. <https://doi.org/10.1093/jhered/esz035>
- Garner BA, Hoban S, Luikart G (2020) IUCN Red List and the value of integrating genetics. *Conserv Genet* 21:795–801. <https://doi.org/10.1007/s10592-020-01301-6>
- Gompert Z, Mock KE (2017) Detection of individual ploidy levels with genotyping-by-sequencing (GBS) analysis. *Mol Ecol Resour* 17:1156–1167
- Gustafsson S, Thorén PA (2001) Microsatellite loci in *Gymnadenia conopsea*, the fragrant orchid. *Mol Ecol Notes* 1:81–82
- Hagl PA, Gargiulo R, Fay MF, Solofondranohatra C, Salmona J, Suescun U et al (2020) Geographical structure of genetic diversity in *Loudetia simplex* (Poaceae) in Madagascar and South Africa. *Bot J Linn Soc*. <https://doi.org/10.1093/botlinnean/boaa098>
- Hamrick JL, Godt MJW (1996) Conservation genetics of endemic plant species. In: Avise JC, Hamrick JL (eds) *Conservation genetics*. Chapman and Hall, New York, pp 281–304
- Hardy OJ (2016) Population genetics of autopolyploids under a mixed mating model and the estimation of selfing rate. *Mol Ecol Resour* 16(1):103–117
- Hautzinger L (1978) Genus “*Orchis*” L. (Orchidaceae): Sectio “*Robustocalcare*” Hautzinger. *Ann Nat Mus Wien* 81:31–73
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* 2:618–620

- Henle K, Lindenmayer DB, Margules CR, Saunders DA, Wissel C (2004) Species survival in fragmented landscapes: where are we now? *Biodivers Conserv* 13:1–8
- Hewitt GM (2004) Genetic consequences of climatic changes in the Quaternary. *Philos Trans R Soc Lond B* 359:183–195
- Hoban S, Bruford M, Jackson JDU, Lopes-Fernandes M, Heuertz M, Hohenlohe PA et al (2020) Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. *Biol Conserv* 248:108654
- Honnay O, Jacquemyn H (2007) Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conserv Biol* 21:823–831
- IUCN (2020) The IUCN Red List of threatened species. Available at: <https://www.iucnredlist.org/>
- Jump AS, Marchant R, Penuelas J (2009) Environmental change and the option value of genetic diversity. *Trends Plant Sci* 14:51–58
- Kretzschmar H, Eccarius W, Dietrich H (2007) The Orchid Genera *Anacamptis*, *Orchis*, *Neotinea*. EchinoMedia, Bürgel
- Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour* 15(5):1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Laikre L, Hoban S, Bruford MW, Segelbacher G, Allendorf FW, Gajardo G et al (2020) Post-2020 goals overlook genetic diversity. *Science* 367:1083–1085
- Landoni B, Viruel J, Gómez R, Allaby RG, Brennan AC, Picó FX, Pérez-Barrales R (2020) Microsatellite marker development in the crop wild relative *Linum bienne* using genome skimming. *App Plant Sci* 8(5):e11349
- Leimu R, Mutikainen P, Koricheva J, Fischer M (2006) How general are positive relationships between plant population size, fitness, and genetic variation? *J Ecol* 94:942–952
- Li YL, Liu JX (2018) StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol Ecol Resour* 18:176–177. <https://doi.org/10.1111/1755-0998.12719>
- Lindley J (1835) Genera and species of orchidaceous plants. 263 London, Ridgways Picadilly
- Lissambou BJ, Couvreur TL, Atteke C et al (2019) Species delimitation in the genus *Greenwayodendron* based on morphological and genetic markers reveals new species. *Taxon* 68(3):442–454
- Médail F, Baumel A (2018) Using phylogeography to define conservation priorities: the case of narrow endemic plants in the Mediterranean Basin hotspot. *Biol Conserv* 224:258–266
- Meirmans PG (2012) AMOVA-based clustering of population genetic data. *J Hered* 103(5):744–750
- Meirmans PG (2020) genodive version 3.0: easy-to-use software for the analysis of genetic data of diploids and polyploids. *Mol Ecol Resour* 20:1126–1131
- Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* 4:792–794
- Meirmans PG, Liu S (2018) Analysis of Molecular Variance (AMOVA) for autopolyploids. *Front Ecol Evol* 6:66
- Meirmans PG, Liu S, Van Tienderen PH (2018) The analysis of polyploid genetic data. *J Hered* 109:283–296
- Mkare TK, van Vuuren BJ, Teske PR (2017) Conservation implications of significant population differentiation in an endangered estuarine seahorse. *Biodivers Conserv* 26:1275–1293. <https://doi.org/10.1007/s10531-017-1300-5>
- Mrkvicka A (1992) Liste der Chromosomenzahlen europäischer Orchideen. *Mitteilungsbl Arbeitskreis Heimische Orchid Baden-Württemberg* 24(1):669–678
- Moody ME, Mueller LD, Soltis DE (1993) Genetic variation and random drift in autotetraploid populations. *Genetics* 134:649–657
- Nemorin A, David J, Maledon E et al (2013) Microsatellite and flow cytometry analysis to help understand the origin of *Dioscorea alata* polyploids. *Ann Bot* 112:811–819
- Orsenigo S, Bacchetta G, Calevo J, Castello M, Cogoni D, Gennai M et al (2016) Global and regional IUCN red list assessments: 1. *Italian Bot* 1:61–85. <https://doi.org/10.3897/italianbotanist.1.8647>
- Pandey M, Richards M, Sharma J (2015) Microsatellite-based genetic diversity patterns in disjunct populations of a rare orchid. *Genetica* 143:693–704. <https://doi.org/10.1007/s10709-015-9867-9>
- Petit RJ, Mousadik AE, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conserv Biol* 12:844–855
- Pridgeon AM, Cribb PJ, Chase MW, Rasmussen FN (eds.) (2001) *Genera Orchidacearum*, vol. 2. Oxford University Press, Oxford

- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Puechmaille SJ (2016) The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Mol Ecol Resour* 16:608–627. <https://doi.org/10.1111/1755-0998.12512>
- Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annu Rev Ecol Evol Syst* 29:467–501
- Rankou H (2011) *Orchis patens*. The IUCN Red List of Threatened Species
- Rodríguez-Rodríguez P, Pérez de Paz PL, Sosa PA (2018) Species delimitation and conservation genetics of the Canarian endemic *Bethencourtia* (Asteraceae). *Genetica* 146:199–210. <https://doi.org/10.1007/s10709-018-0013-3>
- Ronfort JL, Jenczewski E, Bataillon T, Rousset F (1998) Analysis of population structure in autotetraploid species. *Genetics* 150:921–930
- Schatz B, Genoud D, Claessens J, Kleynen J (2020) Orchid-pollinator network in Euro-Mediterranean region: what we know, what we think we know, and what remains to be done. *Acta Oecol*. <https://doi.org/10.1016/j.actao.2020.103605>
- Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends Ecol Evol* 22:25–33
- Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *Proc Natl Acad Sci USA* 101:15261–15264
- Soltis DE, Buggs RJA, Doyle JJ, Soltis PS (2010) What we still don't know about polyploidy. *Taxon* 59:1387–1403
- Stark C, Michalski SG, Babik W, Winterfeld G, Durka W (2011) Strong genetic differentiation between *Gymnadenia conopsea* and *G. densiflora* despite morphological similarity. *Plant Syst Evol* 293:213–226. <https://doi.org/10.1007/s00606-011-0439-x>
- Stift M, Berenos C, Kuperus P, Van Tienderen PH (2008) Segregation models for disomic, tetrasomic and intermediate inheritance in tetraploids: a general procedure applied to Rorippa (yellow cress) microsatellite data. *Genetics* 179:2113–2123
- Stift M, Kolář F, Meirmans PG (2019) Structure is more robust than other clustering methods in simulated mixed-ploidy populations. *Heredity* 123:429–441. <https://doi.org/10.1038/s41437-019-0247-6>
- Straub SC, Parks M, Weitemier K et al (2012) Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *Am J Bot* 99(2):349–364
- Sun Y, Vargas-Mendoza CF (2017) Population structure, genetic diversity, and evolutionary history of *Kleinia nerifolia* (Asteraceae) on the Canary Islands. *Front Plant Sci* 8:1180. <https://doi.org/10.3389/fpls.2017.01180>
- Swarts ND, Dixon KW (2009) Terrestrial orchid conservation in the age of extinction. *Ann Bot* 104(3):543–556. <https://doi.org/10.1093/aob/mcp025>
- Swarts ND, Sinclair EA, Krauss SL, Dixon KW (2009) Genetic diversity in fragmented populations of the critically endangered spider orchid *Caladenia huegelii*: implications for conservation. *Conserv Genet* 10:1199–2120
- Swarts ND, Clements MA, Bower CC, Miller JT (2014) Defining conservation units in a complex of morphologically similar, sexually deceptive, highly endangered orchids. *Biol Conserv* 174:55–64
- Teixeira H, Rodriguez-Echeverría S, Nabais C (2014) Genetic Diversity and Differentiation of *Juniperus thurifera* in Spain and Morocco as Determined by SSR. *PLoS ONE* 9(2):e88996
- Väli Ü, Dombrovski V, Dzmitranok M, Maciorowski G, Meyburg BU (2019) High genetic diversity and low differentiation retained in the European fragmented and declining Greater Spotted Eagle (*Clanga clanga*) population. *Sci Rep* 9:3064. <https://doi.org/10.1038/s41598-019-39187-1>
- Vellend M, Geber MA (2005) Connections between species diversity and genetic diversity. *Ecol Lett* 8:767–781
- Viruel J, Haguenaer A, Juin M et al (2018) Advances in genotyping microsatellite markers through sequencing and consequences of scoring methods for *Ceratonia siliqua* (Leguminosae). *Appl Plant Sci* 6(12):e1201. <https://doi.org/10.1002/aps3.1201>
- Viruel J, Conejero M, Hidalgo O et al (2019) A Target capture-based method to estimate ploidy from herbarium specimens. *Front Plant Sci* 10:937. <https://doi.org/10.3389/fpls.2019.00937>. PMID:31396248; PMCID:PMC6667659
- Volkman L, Martyn I, Moulton V, Spillner A, Mooers AO (2014) Prioritizing populations for conservation using phylogenetic networks. *PLoS ONE* 9(2):e88945. <https://doi.org/10.1371/journal.pone.0088945>
- WCSP (2020) World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; <http://wmsp.science.kew.org/> Retrieved on 13 March 2020

- Yang JY, Ojeda DI, Santos-Guerra A, Molina RJ, Caujapé-Castells J, Cronk Q (2018) Population differentiation in relation to conservation: nuclear microsatellite variation in the Canary Island endemic *Lotus sessilifolius* (Fabaceae). *Conserv Genet Resour* 10(2):219–227
- Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends Ecol Evol* 11:413–418

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

## Authors and Affiliations

Jacopo Calevo<sup>1,2</sup>  · Roberta Gargiulo<sup>2</sup>  · Leif Bersweden<sup>2,3</sup>  · Juan Viruel<sup>2</sup>  ·  
Cristina González-Montelongo<sup>4</sup>  · Khellaf Rebbas<sup>5</sup>  · Lamia Boutabia<sup>6</sup>  ·  
Michael F. Fay<sup>2,7</sup> 

<sup>1</sup> Department of Life Sciences and Systems Biology, University of Turin, Viale Mattioli 25, 10125 Turin, Italy

<sup>2</sup> Royal Botanic Gardens, Kew, Richmond TW9 3DS, UK

<sup>3</sup> Queen Mary University of London, London E1 4NS, UK

<sup>4</sup> Department of Botany, Ecology and Plant Physiology, Universidad de La Laguna, Canary Islands, 456, 38200 Santa Cruz de Tenerife, Spain

<sup>5</sup> Département de Biologie, University Mohamed Boudiaf de M'sila, M'sila, Algeria

<sup>6</sup> Department of Agronomy Sciences, Faculty of Natural and Life Sciences, Chadli Bendjedid University, El Tarf, P.O. Box 73, 36000 El Tarf, Algeria

<sup>7</sup> University of Western Australia, Crawley, WA 6009, Australia

## Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

[onlineservice@springernature.com](mailto:onlineservice@springernature.com)