scientific reports

Check for updates

OPEN

ddRAD sequencing of the endangered species *Primula palinuri* Petagna reveals high levels of inter-population diversity

Paloma Perez-Bello[1,6]✉, Ilaria Panero[2,6]✉, Emanuele De Paoli[1], Valentino Casolo[1], Fabio Attorre[2], Vito Emanuele Cambria[2], Sandro Strumia[3], Annalisa Santangelo[4], Costantino Bonomi[5], Giuseppe Fabrini[2] & Fabio Marroni[1]✉

Population genetics studies provide information regarding genetic diversity and genetic structure of populations. This information can in turn be used to inform conservation strategies. In view of programmed conservation efforts, we aimed at investigating the genetic composition of *Primula palinuri* Petagna, an endemic vulnerable species in Southern Italy. Ninety-one individuals of *P. palinuri* from seven populations were investigated using double digest Restriction Associated DNA sequencing (ddRAD-seq) and a total of 99,014 Single nucleotide polymorphisms (SNPs) were identified. Low levels of heterozygosity (0.08–0.12) and high levels of kinship (0.20–0.30) and inbreeding (0.34–0.49) were detected in all study populations. Kinship computed irrespective of population was low, suggesting that the level of inter-population relatedness is low. Analysis of the genetic structure showed that the populations were differentiated from each other and formed three major clusters, broadly corresponding to the geographic sampling locations. We identified the SNPs mostly contributing to the separation in population clusters, and we identified a reduced set of 40 SNPs that can successfully discriminate between the three population clusters. In conclusion, our study shows that the sampled *P. palinuri* populations have low heterozygosity, low intra-population variability and high inter-population variability.

Keywords Endemic species, Population genetics, Genotyping, Admixture, Population structure

The Mediterranean basin is recognized as a biodiversity hotspot, hosting numerous endemic species and like many other regions globally, is experiencing significant biodiversity loss due to habitat degradation, climate change and anthropogenic pressures[1]. To counter these declines, biodiversity conservation efforts are increasingly urgent, with initiatives spanning habitat restoration, species reintroduction, and other conservation translocation actions[2].

Effective restoration and conservation strategies often benefit from complementary techniques such as assisted gene flow and translocation efforts, which help improve the adaptive capacity and resilience of vulnerable populations[3–5]. Several factors can influence translocation success, including habitat suitability[6] and the characteristics of introduced propagules such as number type and age[7]. Nevertheless, the role of genetic diversity in translocation remains underexplored. Several authors demonstrated that higher genetic diversity increases short-term survival in translocated populations, reinforcing the need to use genetically diverse plant material[7,8].

The European Union, through the LIFE funding program, co-financed the LIFE SEEDFORCE project to enhance the conservation status of 29 plant species listed in Annex II of the Council Directive 92/43/EEC (Habitats Directive), through targeted translocation actions (https://lifeseedforce.eu). According to the IV EU

[1]Department of Agricultural, Food, Environmental and Animal Sciences (Di4A), University of Udine, Via delle Scienze, Udine 206 - 33100, Italy. [2]Department of Environmental Biology, Sapienza University of Rome, Piazzale Aldo Moro, Roma, RM 5 - 00100, Italy. [3]Department of Environmental, Biological and Pharmaceutical Sciences and Technologies, University of Campania "Luigi Vanvitelli", Caserta, Italy. [4]Department of Biology, University of Naples Federico II, Naples, Italy. [5]Botany Section, Museo Delle Scienze (Muse), Trento, Italy. [6]Paloma Perez-Bello and Ilaria Panero contributed equally to this work. ✉email: palperbel@gmail.com; ilaria.panero@uniroma1.it; fabio.marroni@uniud.it

Nature Directives Reporting ex art. 17 (2013–2018)[9], these species currently have either an "Unfavourable-Inadequate" (19 species) or "Unfavourable-Bad" (10 species) conservation status. The project employs an integrated approach to mitigate or remove threats across 139 targeted sites. Over 25,000 reintroductions are planned, with 20,000 occurring exclusively in Italy.

Several of the 29 target species of LIFE SEEDFORCE are endemic to the Mediterranean Biogeographic Region including *Primula palinuri* Petagna (*Primulaceae*, an endangered species on the IUCN Red List)[10]. *P. palinuri* is the only species of the genus *Primula* living on the Mediterranean coasts[11]. It is endemic to a limited area of the southern Tyrrhenian coast, stretching from Cape Palinuro (Campania region) to Cape Scalea (Calabria region), and is preferentially found on cliff sites[12], usually in soil pockets inside rock fractures. Geological changes, coastal retreat, and Quaternary climatic oscillations (glacial-interglacial) have been fundamental in shaping the distribution of many species[13–16] and *P. palinuri* likely underwent a similar process, with current populations representing remnants of larger historical ones.

P. *palinuri* has a reproductive mechanism involving distyly, with populations having two distinct flower types: long-styled with short anthers and short-styled with long anthers. This setup promotes cross-pollination, reducing self-fertilization and enhancing genetic diversity[17].

To date, relatively few studies have assessed genetic diversity of *P. palinuri* species, usually genotyping a few loci, in the context of phylogenetic studies of the genus *Primula*[18–20]. With the aim of filling this knowledge gap, also in consideration of the habitat fragmentation and the status of endangered species, we set out to perform genome-wide population genetics analysis of *P. palinuri*. We used a genome-wide genotyping approach which involves the double digestion of DNA with restriction enzymes (ddRAD-seq) to investigate the genetic structure of the endangered *Primula palinuri* Petagna in its natural habitat along the coastline of Campania, Basilicata and Calabria. Our findings are important for understanding the genetic diversity within this species, providing valuable insights to inform conservation strategies that enhance population resilience and long-term persistence.

Materials and methods
Sample collection
A total of 96 individuals from *P. palinuri* were collected from seven sampling sites in Campania, Calabria and Basilicata regions in Italy, as reported in Table 1 and shown in Fig. 1. Each sampling site was considered as a population.

Fresh leaves were collected from randomly selected individuals at each selected sampling site, ensuring that the individuals were non-clonal (genet). The sampling was randomized, aiming to cover the entire population area. From each individual, 2gr of fresh material were collected, selecting young leaves and excluding senescent or pathogen-affected leaves. Each sample was placed in a breathable paper envelope sealed in a plastic bag filled with silica gel (Laboratorio Vetrerie Scientifiche SRL).

Samples were transported to the Germplasm Bank of the Botanical Garden of Rome (University of Rome, La Sapienza) and stored at 4 °C, with the silica gel replaced after 24 h. Once the samples were dried, they were sent for genetic analyses to the University of Udine.

Permission to collect vegetal material was obtained by Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA) and by the Italian Ministry environment and energy security. Voucher specimens for *P. palinuri* petagna are conserved in the Herbarium Horti Botanici Pisani (PI), collected by Gregorio Rigo in Campania (Italy) and stored with accession PI 022579 (https://pi.jacq.org/PI022579).

DNA extraction, library preparation and sequencing
DNA was extracted from leaves tissue using the kit NucleoSpin Plant II (Thermo Fisher Scientific). Sequencing libraries were prepared by IGA Technology services s.r.l (Udine, Italy), following the protocol for ddRAD sequencing (double digest Restriction Associated DNA sequencing)[21] with minor custom modifications. The procedure included DNA quantification with Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) and normalization to a uniform concentration followed by a double enzymatic digestion with SphI and EcoRI endonucleases (New England BioLabs).

Code	Location	Total individuals	Retained Individuals
LAM	Lambro river	8	8
FIU	Fiuzzi	12	12
ID	Dino Island	12	9
PC	Punta Caina	12	12
SGP	San Giovanni a Piro	12	12
CIM	Vaccuta	20	20
PPA	Palinuro Cape	20	18
ALL		96	91

Table 1. Summary of studied samples. Code: Three-letter code collection site. Location: collection site. Total individuals: number of individuals from which DNA was extracted and sequenced. Retained individuals: sequenced samples that passed quality control.

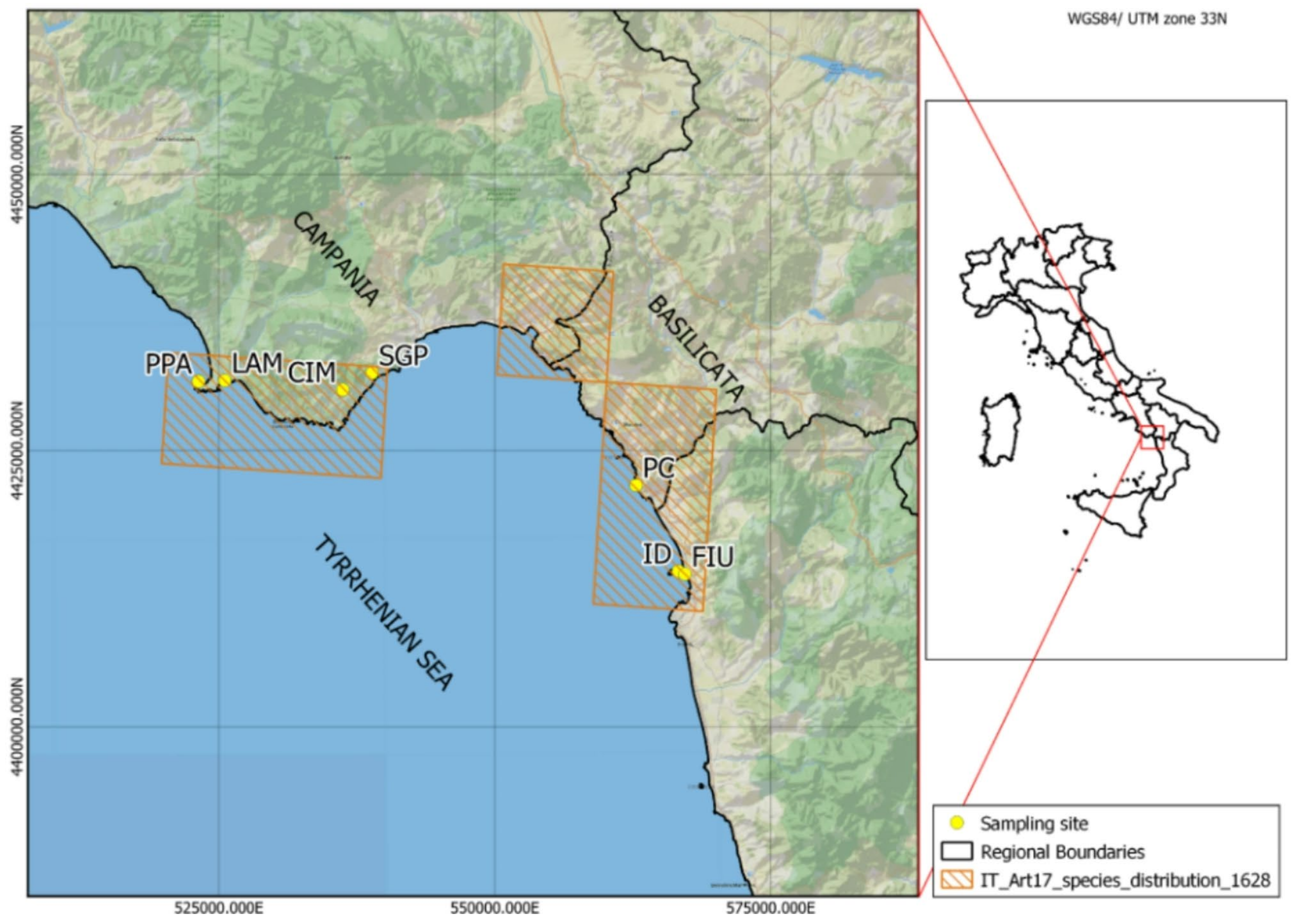

Fig. 1. Geographic location of the studied samples. Geographic locations correspond to LAM (Lambro river), FIU (Fiuzzi), ID (Dino' Island), PC (Punta Caina), SGP (San Giovanni a Piro), CIM (Vaccuta) and PPA (Palinuro Cape). The 10 × 10 km grid regions represent the official distribution of *P. palinuri* according to the Habitat Directive Report (2013–2018) https://reportingdirettivahabitat.isprambiente.it/downloads.

DNA fragments were purified using AMPureXP beads (Agencourt) followed by barcoded adapter ligation. Final libraries were checked with both Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) and with Bioanalyzer High sensitivity DNA assay (Agilent technologies, Santa Clara, CA).

Paired-end sequencing 2 × 150 bp was performed on a Novaseq 6000 instrument (Illumina, San Diego, CA).

De Novo DdRAD genotyping and variant calling

Multiplexed samples present within the same sequencing lane were demultiplexed based on their index sequence and sorted into separated files using the process_radtags package included in the Stacks v2.61 software[22] by IGA technology services s.r.l. (Udine, Italy). Prior to genetic analyses, five individuals with less than 250 K read pairs were discarded. As no reference genome is available for *Primula palinuri*, a catalog reference was created with the *de-novo* pipeline for pair end sequencing using default parameters. Quality control was performed following the approach of Magris and colleagues[23]. Briefly, only loci present in at least 75% of the individuals (-R 0.75) were retained for further analyses; this ensured that only SNPs with a high information content were analysed. In addition, only SNPs with an observed heterozygosity of 0.8 or lower (--max-obs-het 0.8) were included. This choice is motivated by the fact that extremely high levels of heterozygosity are not common in natural populations and may be caused by the presence of reads originated by regions with high sequence similarity to the genotyped stack. In addition, the parameter --fstats was included to enable SNP and haplotype-based F statistics.

To ensure high genotyping quality, polymorphisms with less than 5 reads of coverage were considered uninformative and were removed using vcftools[24] -minDP 5. Finally, only genotyped positions that were informative in at least 50% of the samples were retained for further analyses.

Population genetics analysis

The software Stacks v2.0[22] was used to compute the observed heterozygosity (Ho), the expected heterozygosity assuming Hardy-Weinberg equilibrium (He), nucleotide diversity (π), the number of private alleles and the pairwise F_{st}. Effective population size (Ne) was calculated for each population using NeEstimator v2[25], following the linkage disequilibrium (LD) random mating model with a minor allele frequency (MAF >= 0.2). Since

NeEstimator cannot read variant call format (vcf) file and requires GENEPOP format, vcf files were converted in GENEPOP using the R libraries vcfR[26] and adegenet[27].

The relatedness between individuals was assessed by calculating the maximum likelihood estimator of identity-by-descent (IBD) included in the R package SNPRelate[28]. Kinship and Inbreeding coefficients were calculated including SNPs with a minimum allele frequency of 5%, and a threshold of 0.2 for Linkage Disequilibrium pruning. To address robustness of population-specific estimates to varying sample size, down sampling was performed by randomly selecting 8 individuals per populations.

Principal component analysis (PCA) was performed using the R library SNPRelate[28]. A neighbor-joining (NJ) phylogenetic tree[29] was constructed using the nj function of the ape library.

To further investigate population structure, Admixture analysis[30] was performed. The cross-validation procedure[31] was used to determine the optimal number of populations K. For each individual, admixture computes the probability Q (from Q1 to QK) that the individual belongs to each of the K populations. maxQ was defined as the maximum probability Q across all K clusters. Individuals with a maxQ lower than 0.95 were considered admixed.

An additional principal component analysis was carried out using the R function prcomp to identify the SNPs most contributing to the variance explained by the PCA, i.e. those with the more extreme coordinates in the rotated reference system. Only SNPs without missing data were included in this analysis. Similarity of stacks with known proteins was performed using blastx[32] against the SwissProt protein database. Hits were limited to those with an e-value lower than 10^{-5}.

To identify SNPs showing the highest values of F_{st}, pairwise F_{st} estimation was performed between populations[33]. Outliers were arbitrarily defined as the top 5% SNPs in the pairwise F_{st} distribution. As before, stacks containing the selected outliers were analysed using blastx[32] against the SwissProt protein database.

A Mantel test[34] was used to evaluate the correlation between genetic distance (measured as pairwise F_{st}) and geographic distance using the R function mantel included in the R library vegan and setting the number of permutations to 9999.

The workflow used for the present work is summarized in Supplementary Fig. S1.

Results

Sample collection, sequencing summary and SNP discovery

The selected sites for the study encompassed all the representative areas where *P. palinuri* is found; this enabled to sample the full ecological and genetic diversity of the species. Figure 1 shows the locations of the populations used in the present study.

A total of 96 samples were sequenced, five of which failed quality control due to low number of read pairs (less than 250 K) and were discarded. The ninety-one samples retained for analysis had a total of 172,831,301 read pairs, with an average of about 1.9 M reads per individual, ranging from 322 K to 4.3 M reads. A mean coverage of 27.78x was obtained, with individual values that ranged from 9.81x to 52.92x. A total of 99,014 polymorphic positions satisfied quality control and were retained for analysis. The number of reads, mean coverage, mapped reads, maxQ and the observed heterozygosity for each individual are listed in Supplementary table S1. Summary population genetic statistics are reported in Table 2.

Genomic diversity and population structure

The boxplot represented in Fig. 2a shows the observed heterozygosity across the populations. LAM and PPA show the lowest heterozygosity, and ID and PC the highest. Figure 2b shows the distribution of kinship coefficients for all pairwise comparisons between individuals of the same population (colored boxplots) and for all pairwise comparisons regardless of their population (white boxplot). Overall, the kinship coefficient between individuals was relatively high, with the highest values in LAM and PPA and the lowest in CIM. The median kinship coefficient for all pairwise comparisons regardless of population of origin, is significantly lower than the within-population kinship coefficient. The inbreeding coefficients of the populations showed a similar pattern among populations, with the highest inbreeding for LAM and PPA and the lowest value for CIM (Fig. 2c). As shown in Supplementary Fig. S2, the down sampling analyses confirmed that differences in sample size across

Code	Coverage	He	Ho	π	Private alleles	Kinship	Inbreeding	Ne
LAM	33.3	0.08	0.08	1.37E-03	2621	0.30	0.49	38.3
FIU	20.4	0.10	0.10	1.62E-03	4734	0.28	0.46	114
ID	25.9	0.11	0.12	1.58E-03	5240	0.27	0.37	9.8
PC	27.2	0.11	0.11	1.69E-03	8942	0.27	0.42	11.5
SGP	28.4	0.10	0.10	1.62E-03	3680	0.25	0.39	8.6
CIM	30.2	0.12	0.11	1.56E-03	8357	0.20	0.34	22.5
PPA	27.8	0.08	0.08	1.29E-03	5047	0.30	0.46	2.8

Table 2. Summary of population analysis. Code: Three-letter code collection site. Coverage: mean coverage obtained per population; He and Ho refers to mean expected and observed heterozygosity respectively; π: nucleotide diversity; Private alleles: number of exclusive alleles per population; Kinship: mean kinship coefficient; Inbreeding: mean inbreeding coefficient. Ne: effective population size.

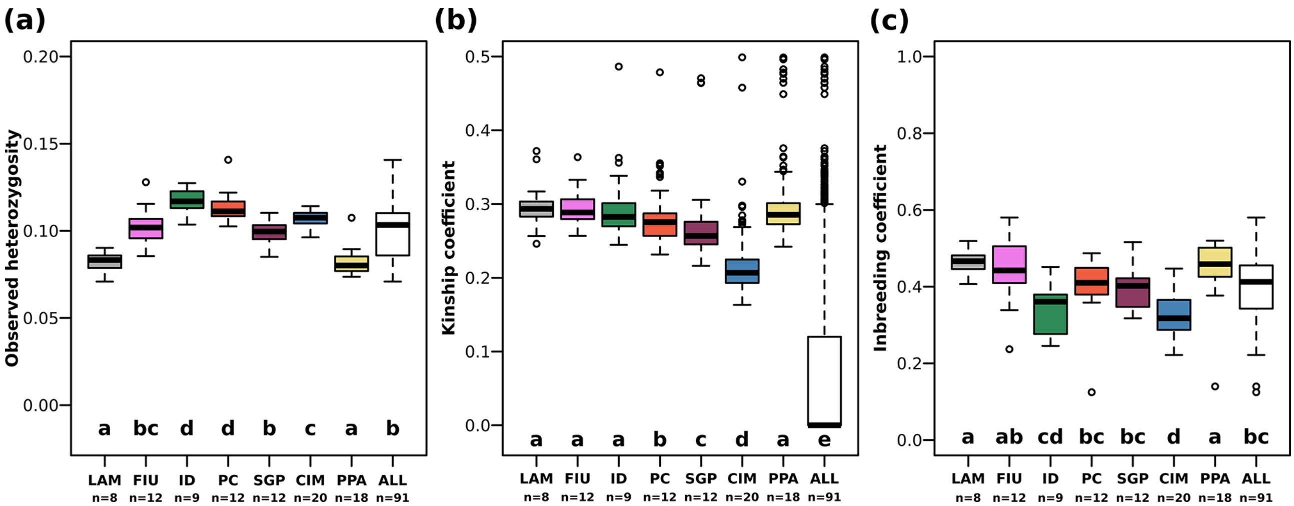

Fig. 2. Boxplots of population genetics parameters. (**a**) Observed heterozygosity for different populations. (**b**) Kinship coefficient and (**c**) Inbreeding coefficient. Population names are listed in Table 1. "ALL" denote statistics computed considering the whole dataset as a single population. Populations with the same letter are not significantly different from each other (Wilcoxon's test). Geographic locations correspond to LAM (Lambro river), FIU (Fiuzzi), ID (Dino' Island), PC (Punta Caina), SGP (San Giovanni a Piro), CIM (Vaccuta) and PPA (Palinuro Cape). "n" denotes the number of sequenced individuals per population.

populations do not compromise the reliability of population-specific estimates. The results for Ho, IBD, and kinship coefficients remained consistent across all replicates.

PCA analysis separated the sampled individuals in three clusters: cluster 1, including individuals from populations of FIU, ID and PC; cluster 2, including individuals from LAM and PPA and cluster 3, with individuals collected in SGP and CIM (Fig. 3a). For admixture analysis, cross-validation suggested an optimal number of 5 clusters (Supplementary Fig. S3). According to this interpretation (Supplementary Fig. S5), FIU and ID form a cluster separated from PC, another cluster is formed by LAM and PPA, and the remaining populations are separated. Although the cross-validation procedure does not suggest it as the most probable number of clusters, the use of K = 3 (Fig. 3b) results in the same three clusters identified by PCA. For completeness, results of admixture analyses for K ranging from 4 to 7 are reported in Supplementary Fig. S4, Fig. S5, Fig. S6 and Fig. S7 respectively.

The phylogenetic tree (Fig. 3c) also suggests that the studied populations broadly form the same three clusters identified by the PCA. However, it also suggests that within clusters, populations are clearly separated. A further PCA performed separately on cluster 1 (Supplementary Fig. S8), cluster 2 (Supplementary Fig. S9), and cluster 3 (Supplementary Fig. S10) confirmed the genetic separation between populations belonging to the same cluster.

Identification of SNPs most contributing to genetic differentiation

To identify the SNPs responsible for the differentiation between clusters, the SNPs with the most extreme coordinates in the rotated system were selected. In Fig. 4 we represent as red arrows the 40 SNPs with highest loadings in the first two principal components; such SNPs are responsible for differentiating the three clusters. The genotype obtained for each individual and the proportion for each population are available in Supplementary table S2. The blastx analysis revealed six hits among the top 40 SNPs, identifying sequences associated with transposon-related proteins, pathogenic effectors, structural plant proteins, and cell cycle regulators (Supplementary table S3).

An additional analysis aimed at SNPs differentiating the studied population is the F_{st} outlier analysis. We selected the stacks containing the top 5% SNPs showing the highest F_{st} and mapped them against known proteins using blastx[32]. Results are shown in Supplementary table S4. Out of 155 analysed stacks, 36 returned a blast hit and for 33 of them the best hit was the Retrovirus-related Pol polyprotein from transposon TNT 1–94 (protein ID P10978.1). The percentage of stacks containing such sequence (21%) is substantially higher than expected by chance; our resampling experiments revealed an average percentage of 4.4% (c.i 3.2-5.6%, data not shown).

Genetic and geographic distances are correlated, with a Mantel statistics correlation value of $r = 0.83$ and a p-value = 0.001. A scatterplot representing the genetic distance estimated by pairwise F_{st} as a function of geographical distance, together with Mantel statistics and the corresponding p-value, is reported in Supplementary Fig. S11.

Discussion

Geographical isolation and habitat fragmentation over time could lead to genetic differentiation between small populations, but the extent of these genetic differences remains largely unquantified[35,36]. Genetic variability in endangered plant species is a significant concern, given the need to avoid biodiversity loss. Therefore, molecular data are essential for conserving genetic diversity, preventing the loss of valuable alleles and to take action against

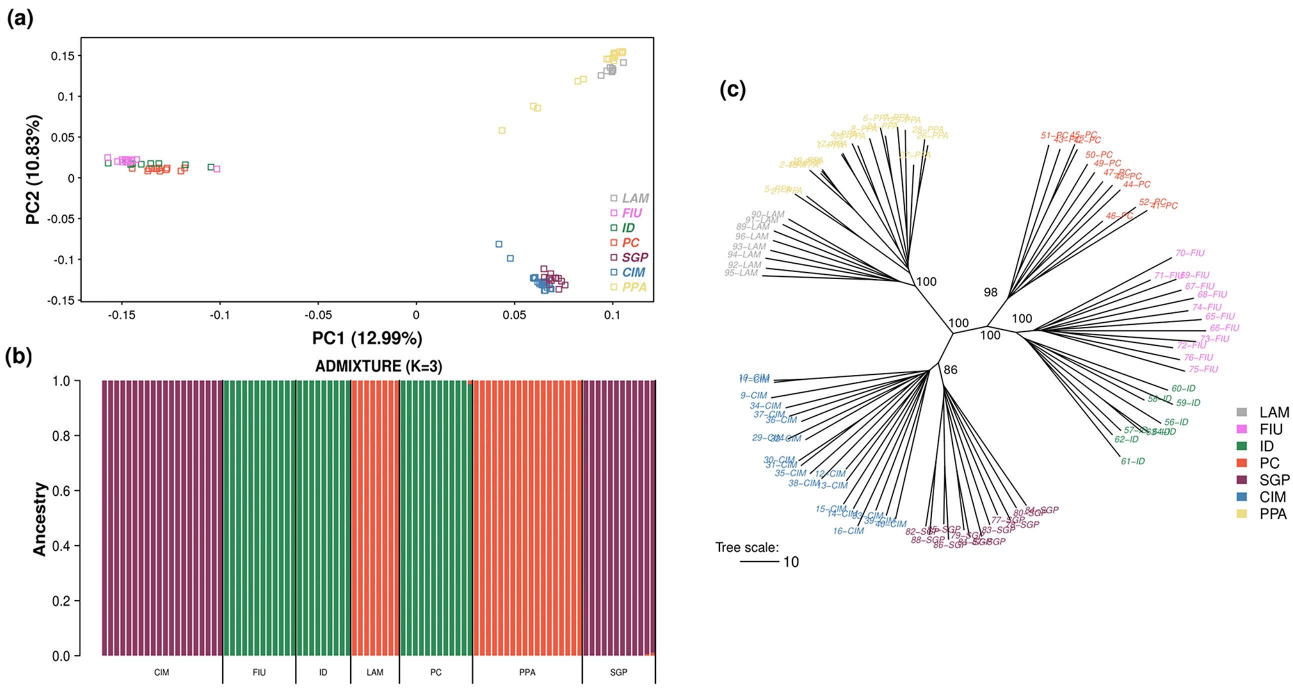

Fig. 3. PCA, admixture analysis and phylogenetic tree. (**a**) Principal component analysis (PCA) with colors representing each population and each individual depicted with a square. (**b**) Admixture analysis in the studied individuals of *Primula palinuri*. Each barplot represents an individual. The elevation of each colored bar represents the probability of assignment to the corresponding population when assuming the presence of three ancestral populations (K = 3). (**c**) Phylogenetic neighbor-joining unrooted tree based on Nei's distance and bootstrap = 100. Geographic locations correspond to LAM (Lambro river), FIU (Fiuzzi), ID (Dino' Island), PC (Punta Caina), SGP (San Giovanni a Piro), CIM (Vaccuta) and PPA (Palinuro Cape).

inbreeding depression[37–39]. In this study, we assessed the population structure and genetic diversity of seven *P. palinuri* populations using a large set of genome-wide SNPs generated through ddRAD sequencing.

The heterozygosity recorded in *P. palinuri* ranged from 0.08 in PPA to 0.12 in ID (Fig. 2a). While it is hard to give an absolute evaluation of heterozygosity levels, the heterozygosity recorded in the present study can be considered to be low. Similar values of heterozygosity have been observed in *Eokochia saxicola*[16], another endangered endemic species to coastal cliff of Southern Italy, living in Cilento on the same cliff of *P. palinuri* even if at different distance from sea level. Studies on biallelic RAPD markers of the congeneric *Primula veris*, revealed similar values of heterozygosity[40]. While the species *P. veris* is not considered endangered, authors reported a decline in the populations size in the study site (Kosovo)[40]. However, recent ddRAD-seq[10] analysis on *P. veris* in Estonia[41], revealed levels of heterozygosity substantially higher than those measured in *P. palinuri*.

Our results showed relatively high levels of intra-population relatedness (Fig. 2b) and inbreeding (Fig. 2c) in *P. palinuri*. The high levels of inbreeding and kinship are not justified by the population size, since the latest estimate suggested a population size of 18,500 individuals[10]. A possible explanation is that, in spite of a large population size, *P. palinuri* is divided into several small sub-populations, experiencing relatively small gene flow mostly between geographically close populations. This is in agreement with the observation that the average kinship is high within population and low between populations (Fig. 2b) and that the populations are clearly differentiated. The population differentiation observed in the present work shows that three major clusters can be defined based on the PCA analysis and on the phylogenetic tree (Fig. 2a and c respectively). The three clusters, with FIU, ID and PC forming cluster 1, LAM and PPA forming cluster 2, and SGP and CIM forming cluster 3, also reflects the geographic location of the sampling points (Fig. 1). This result is confirmed by Mantel's test, showing a significant correlation ($r = 0.83$ and p-value = 0.001) between genetic and geographic distance. Even at relatively short distances, genetic differentiation increases with distance, as could be expected under the isolation by distance scenario[42]. This is also in agreement with the observation that the three main clusters can be further and clearly separated in the seven populations. This is observed in the phylogenetic tree (Fig. 2c), but is also evident in the clear separation obtained in the admixture analysis with K = 7 (Supplementary Fig. S6) and in the PCA analysis performed in each of the three clusters separately (Supplementary Fig. S7, Fig. S8 and Fig. S9). All the results suggest a hierarchical structure of population, probably driven by geographic distance.

The exchange of genetic material with nearby populations may be driven by.

pollinator species such as *Anthophora plumipes*, *Bombus terrestris*, and *Macroglossum stellatarum* that have been identified as potential pollinators of *P. palinuri*[43]. The potential of these possible pollinators for *P. palinuri* may account for gene exchange between nearby populations identified in the three clusters. These species are generalists capable of covering large distances[44]. Moreover, male individuals of *Anthophora plumipes* could

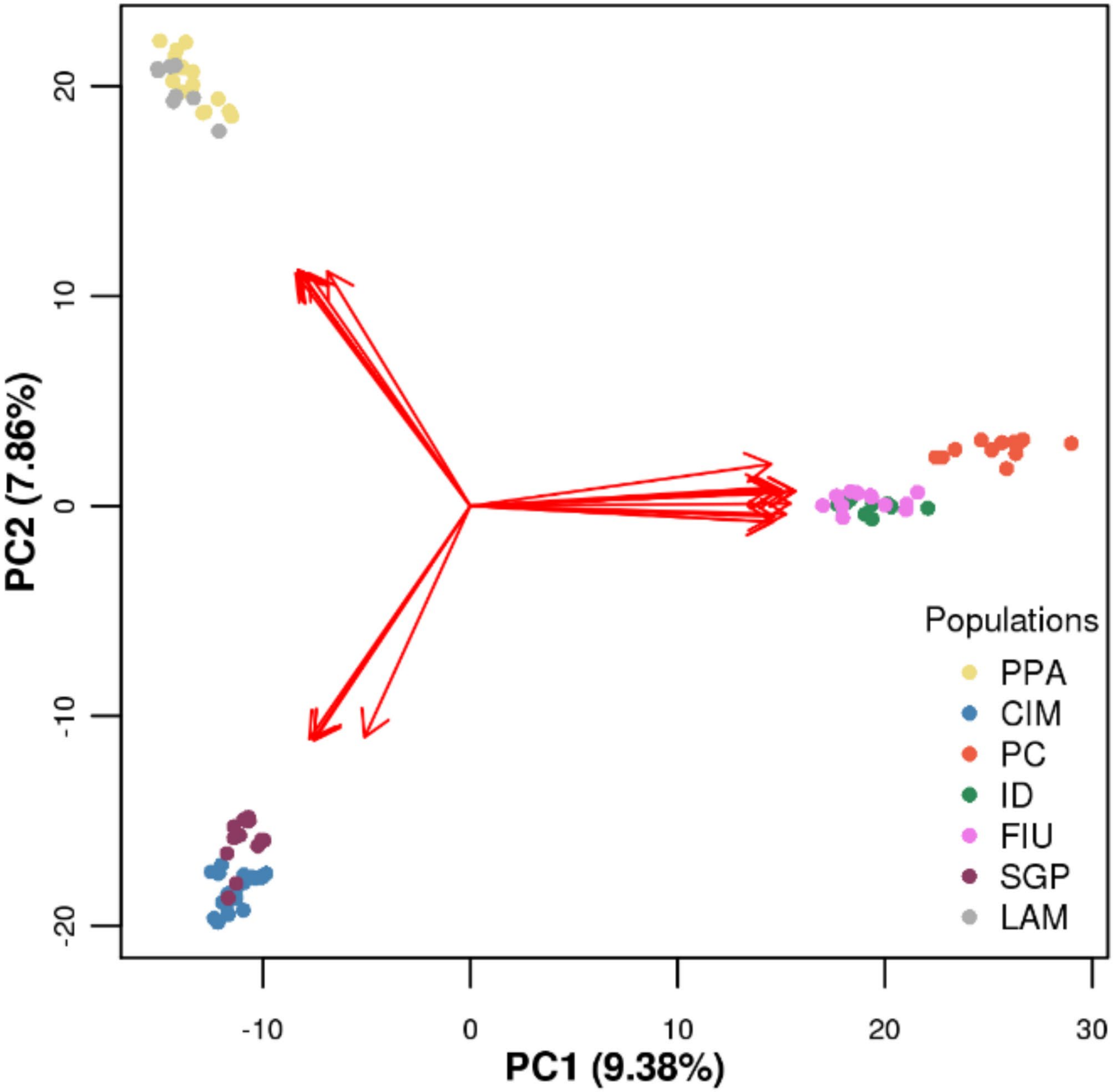

Fig. 4. Biplot with a subset of 40 SNPs. Subset of the 40 SNPs more contributing to the separation of population clusters. Colors correspond to different populations and SNPs are represented with red arrows. Geographic locations correspond to LAM (Lambro river), FIU (Fiuzzi), ID (Dino' Island), PC (Punta Caina), SGP (San Giovanni a Piro), CIM (Vaccuta) and PPA (Palinuro Cape).

promote cross-pollination by flying over greater distances[43], but the lack of intermediate populations, along with the high fragmentation of the *P. palinuri* range, may limit this process. This is further supported by Kaldra et al., who highlight that heterostylous species such as *P. veris*, which rely on pollen transfer between different floral morphs, are particularly vulnerable to fragmentation due to their dependence on specific pollinator interactions for maintaining genetic diversity[45].

The differentiation of the three clusters is also reflected by SNPs private to each population cluster. The identification of the SNPs that were responsible for discriminating the three clusters revealed that with only 40 SNPs, the three clusters can be clearly identified, suggesting that these SNPs can be used as diagnostic SNPs for membership of one of the three clusters. This may potentially help in designing cheap assays for assigning individuals to one group of population by typing only a few markers. Sequence similarity analysis revealed that only six of these 40 stacks had high similarity to know proteins. Of them, three had similarity to a transposable element (protein id: P10978.1).

Another widely used approach to identify SNPs responsible of population differentiation is the F_{st} outlier analysis. In our analysis we selected the 100 stacks harboring the SNPs with higher F_{st}. Again, we assessed sequence similarity, and we observed that 21 stacks had similarity to the same transposable element (protein id: P10978.1).

The frequency of transposable element can be simply due to the abundance of TE in *P. palinuri* genome, or may be caused by recent movements of TE, or by fast evolution of TE sequence. According to our resampling experiment, the average abundance of the TE in the genome is 4.4%, compared to 21% identified by F_{st} analysis. This may suggest that the presence of TE-related sequences in divergent stacks is not due to chance alone.

While defining a specific conservation strategy for *P. palinuri* was not the primary aim of this study, our findings underscore the importance of integrating genetic structure information into translocation planning[46]. In particular, the observed low intra-population heterozygosity and high inter-population differentiation suggest potential local adaptations that must be carefully considered[37]. This genetic structuring implies that sourcing individuals for translocation should prioritize genetic compatibility and adaptation to local conditions to avoid outbreeding depression.

The high levels of relatedness within populations compared to the low level of kinship between populations may suggest translocating admixture of individuals from different populations as has been suggested before[47–49] to increase the genetic diversity of the new population and reduce the risk of genetic impoverishment. On the other hand, the high level of genetic differentiation suggests that some populations may be particularly well adapted to local conditions. The possibility of outbreeding depression is real and, although to a limited extent, it has been observed in the congeneric species *P. veris*[46]. Consequently, it is crucial to adopt a balanced approach to avoid the introduction of too many different ecotypes that may not be adequately adapted to the new environment[37].

Our results substantially improve the understanding of the conservation status of *P. palinuri* in Italy. However, some limitations are present. Specifically, no reference genome is available for this species; thus, each stack has to be treated individually, and this limits the study of potentially adaptive regions composed by more stacks. In addition, due to the strong limitations in sample availability, the size of populations was small and uneven, although we showed that sample size had no major effect on several population genetic parameters estimation.

In conclusion, our results confirm the urgency of adopting conservation actions to preserve and increase genetic diversity in *P. palinuri* populations. While this study offers insights into population structure and genetic diversity, further research on translocation outcomes and admixture effects would support refined strategies for maintaining and enhancing the resilience of *P. palinuri* populations across its range.

Data availability

The sequencing reads used in this study are available in Sequence Read Archive (SRA), under the accession PRJNA1153737 (https://www.ncbi.nlm.nih.gov/sra/PRJNA1153737).

Code availability

Code availability Scripts and functions used in the present work are released on github under the MIT License: (https://github.com/genomeud/Seedforce/tree/main/Primula_palinuri).

Received: 22 November 2024; Accepted: 10 April 2025
Published online: 30 April 2025

References

1. Myers, N., Mittermeier, R., Mittermeier, C., Da Fonseca, G. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858 (2000).
2. IUCN/SSC. *Guidelines on the Use of Ex Situ Management for Species Conservation. Version 2.0.* (2014). www.iucn.org/about/work/programmes/species/publications/iucn_guidelines_and__policy__statements/
3. Zimmer, H. C., Auld, T. D., Cuneo, P., Offord, C. A. & Commander, L. E. Conservation translocation - An increasingly viable option for managing threatened plant species. *Aust J. Bot.* **67**, 501–509 (2019).
4. Gargiulo, R. et al. Conservation of the threatened species, pulsatilla vulgaris mill. (Pasqueflower), is aided by reproductive system and polyploidy. *J. Hered.* **110**, 618–628 (2019).
5. Van Rossum, F., Le Pajolec, S., Raspé, O. & Godé, C. Assessing population genetic status for designing plant translocations. *Front. Conserv. Sci.* **3**, (2022).
6. Bellis, J., Bourke, D., Maschinski, J., Heineman, K. & Dalrymple, S. Climate suitability as a predictor of conservation translocation failure. *Conserv. Biol.* **34**, 1473–1481 (2020).
7. Menges, E. S. Restoration demography and genetics of plants: When is a translocation successful? *Austral. J. Bot.* **56**, 187–196 (2008). https://doi.org/10.1071/BT07173
8. Schäfer, D., Vincent, H., Fischer, M. & Kempel, A. The importance of genetic diversity for the translocation of eight threatened plant species into the wild. *Glob Ecol. Conserv.* **24**, (2020).
9. Ercole, S. Monitoraggio e Stato Della biodiversita in Italia. *Ecoscienza* **2**, 310–335 (2022).
10. Gangale, C., Uzunov, D. & Cesca, G. Primula Palinuri. *IUCN Red List. Threatened Species*. **2013** https://doi.org/10.2305/IUCN.UK.2011 (2015).
11. Aronne, G., Buonanno, M. & De Micco, V. Reproducing under a warming climate: Long winter flowering and extended flower longevity in the only mediterranean and maritime primula. *Plant. Biol.* **17**, 535–544 (2015).
12. De Micco, V. & Aronne, G. Occurrence of morphological and anatomical adaptive traits in young and adult plants of the rare mediterranean cliff species Primula palinuri Petagna. *Sci. World J.* (2012). (2012).
13. Taberlet, P., Fumagalli, L., Wust-Saucy, A. G. & Cosson, J. F. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* **7**, 453–464 (1998).
14. Hewitt, G. M. Speciation, hybrid zones and phylogeography or seeing genes in space and time. *Mol. Ecol.* **10**, 537–549 (2001).
15. Heuertz, M. et al. Chloroplast DNA variation and postglacial recolonization of common Ash (Fraxinus excelsior L.) in Europe. *Mol. Ecol.* **13**, 3437–3452 (2004).
16. Strumia, S., Santangelo, A., Galise, T. R., Cozzolino, S. & Cafasso, D. Hopping or jumping on the Cliffs: The unusual phylogeographical and demographic structure of an extremely narrow endemic mediterranean plant. *Front. Plant. Sci.* **12**, (2021).

17. Aronne, G., Buonanno, M. & De Micco, V. Assessment of distyly syndrome in primula Palinuri Petagn. A rare species living on maritime vertical cliffs. *Plant Syst. Evol.* https://doi.org/10.1007/s00606-013-0931-6 (2013).
18. Crema, S. & Cristofolini, G. Filogenesi Di primula subsect. Euauricula Un Esempio Di origine ed evoluzione Dell Endemismo Alpino. *Annali Del. Museo Civico Di Rovereto*. **28**, 135–159 (2013).
19. Crema, S., Kadereit, J. W. & Cristofolini, G. Phylogenetic insights into primula sect. Auricula in the apennine Peninsula. *Flora Mediterranea*. **23**, 157–172 (2013).
20. Boucher, F. C., Casazza, G., Szövényi, P. & Conti, E. Sequence capture using RAD probes clarifies phylogenetic relationships and species boundaries in primula sect. Auricula. *Mol. Phylogenet Evol.* **104**, 60–72 (2016).
21. Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest RADseq: an inexpensive method for de Novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**, (2012).
22. Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. A. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140 (2013).
23. Magris, G. et al. ddRAD-seq reveals the genetic structure and detects signals of selection in Italian brown trout. *Genet. Sel. Evol.* **54**, (2022).
24. Danecek, P. et al. The variant call format and vcftools. *Bioinformatics* **27**, 2156–2158 (2011).
25. Do, C. et al. NeEstimator v2: Re-implementation of software for the Estimation of contemporary effective population size (Ne) from genetic data. *Mol. Ecol. Resour.* **14**, 209–214 (2014).
26. Knaus, B. J. & Grünwald, N. J. vcfr: a package to manipulate and visualize variant call format data in R. in *Mol. Ecol. Resour.* **17** 44–53 (2017).
27. Jombart, T. & Ahmed, I. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**, 3070–3071 (2011).
28. Zheng, X. et al. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012).
29. Nei, M. Analysis of gene diversity in subdivided populations. *PNAS* **70**, 3321–3323 (1973).
30. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based Estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
31. Fu, W. & Perry, P. O. Estimating the number of clusters using cross-validation. *J. Comput. Graphical Stat.* **29**, (2017).
32. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215** (3), 403–410 (1990).
33. Kitada, S., Nakamichi, R. & Kishino, H. Understanding population structure in an evolutionary context: Population-specific FST and pairwise FST. *G3: Genes, Genomes, Genetics* 11, (2021).
34. Diniz-Filho, F., Soares, J. A., Lima, J. S. & T. N. & Mantel test in population genetics. *Genet. Mol. Biol.* **4**, 475–485 (2013).
35. Waples, R. S., Punt, A. E. & Cope, J. M. Integrating genetic data into management of marine resources: How can we do it better? *Fish Fish.* **9**, 423–449 (2008).
36. Laikre, L., Nilsson, T., Primmer, C. R., Ryman, N. & Allendorf, F. W. Importance of genetics in the interpretation of favourable conservation status. *Conserv. Biol.* **23**, 1378–1381 (2009).
37. Mckay, J. K., Christian, C. E., Harrison, S. & Rice, K. J. 'How local is local?"-A review of practical and conceptual issues in the genetics of restoration. *Restor. Ecol.* **13**, 432–440 (2005).
38. Gentili, R. et al. Conservation genetics of two Island endemic Ribes spp. (Grossulariaceae) of Sardinia: survival or extinction? *Plant. Biol.* **17**, 1085–1094 (2015).
39. Gentili, R. et al. Genetic variability of the first-generation of Ribes Sardoum, a threatened relic plant requiring translocation measures. *Plant. Biosyst*. **153**, 1–4 (2019).
40. Berisha, N., Millaku, F., Gashi, B., Krasniqi, E. & Novak, J. Initial determination of DNA polymorphism of some primula veris L. populations from Kosovo and Austria. *Physiol. Mol. Biology Plants*. **21**, 117–122 (2015).
41. Reinula, I. et al. Beware of the impact of land use legacy on genetic connectivity: A case study of the long-lived perennial primula veris. *Biol. Conserv.* **292**, (2024).
42. Wright, S. *Isolation Distance Genet.* **28**, (1943).
43. Giovanetti, M. & Aronne, G. Insect visitors to the rare plant primula Palinuri Petagna (Primulaceae) in Southern Italy: A framework for future research. *J. Apic. Res.* **51**, 280–281 (2012).
44. Ne'eman, G., Shavit, O., Shaltiel, L. & Shmida, A. Foraging by male and female solitary bees with implications for pollination. *J. Insect Behav.* **19**, 383–401 (2006).
45. Kaldra, M. et al. Skewed morph ratios lead to lower genetic diversity of the heterostylous primula veris in fragmented grasslands. *Plant. Biol.* **25**, 703–714 (2023).
46. Nistelberger, H. M. et al. Genetic mixing in conservation translocations increases diversity of a keystone threatened species, bettongia lesueur. *Mol. Ecol.* https://doi.org/10.1111/mec.17119 (2023).
47. Van Rossum, F., Hardy, O. J., Le Pajolec, S. & Raspé, O. Genetic monitoring of translocated plant populations in practice. *Mol. Ecol.* **29**, 4040–4058 (2020).
48. Albrecht, M. A. & Edwards, C. E. Genetic monitoring to assess the success of restoring rare plant populations with mixed gene pools. *Mol. Ecol.***29** 4037–4039 (2020). https://doi.org/10.1111/mec.15658
49. Godefroid, S. et al. How successful are plant species reintroductions? *Biol. Conserv.* **144** 672–682 (2011). https://doi.org/10.1016/j.biocon.2010.10.003

Acknowledgements
The authors gratefully acknowledge Nicoletta Felice from the University of Udine for her substantial contribution to sample management and DNA extraction.

Author contributions
P.P and F.M designed and performed the bioinformatic analyses and interpreted the results.I.P and G.F designed and conducted sampling. I.P, E.P, S.S, A.S, V.C, C.B, VE.C and F.A contributed with results interpretation.P.P, I.P and F.M wrote the first draft of the manuscript.All authors read and approved the manuscript.

Funding
This work received funding from the LIFE SEEDFORCE program - Using SEED banks to restore and reinFORCE the endangered native plants of Italy under grant agreement LIFE20 NAT/IT/001468 (EC CINEA2021; LIFE SEEDFORCE 2022).

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-98334-z.

Correspondence and requests for materials should be addressed to P.P.-B., I.P. or F.M.

Reprints and permissions information is available at www.nature.com/reprints.

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