Conservation implications for the Iberian narrowly endemic *Androsace cantabrica* (Primulaceae) using population genomics with target capture sequence data

Jungle Ke Liang[1], Amelia Shepherd-Clowes[1], Pablo Tejero Ibarra[1], and Juan Viruel[1]

[1]Affiliation not available

December 06, 2024

Jungle Ke Liang[1,2], Amelia Shepherd-Clowes[1], Pablo Tejero Ibarra[3,4]*, Juan Viruel[1]*+

[1]Royal Botanic Gardens, Kew, Richmond, TW9 3DS, England

[2]Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, PO Box 94240, 1090GE, Amsterdam, The Netherlands

[3]Instituto Pirenaico de Ecología (IPE-CSIC), Av. Nuestra Señora de la Victoria 16, 22700 Jaca, Huesca, Spain

[4]Aranzadi Zientzia Elkartea, Zorroagagaina 11, 20014 Donostia, Spain

* Both authors contributed as senior authors to this publication.

+ Author for correspondence (j.viruel@kew.org, juanviruel@gmail.com)

Jungle Ke Liang: k.liang@kew.org

Amelia Shepherd-Clowes: A.Shepherd-Clowes@kew.org

Pablo Tejero Ibarra: ptibarra@aranzadi.eus

ABSTRACT

Androsace cantabrica (Losa & P. Monts.) Kress is a narrow endemic polyploid restricted tofew northern Iberian mountains, potentially threatened by global warming and human activities. However, *A. cantabrica* is taxonomically not accepted and is considered a synonym of *A. adfinis*. We investigated six *A. cantabrica* populations, which cover its entire distribution range, together with related taxa and used Angiosperms353 target capture sequence data at two scales: (1) applying phylogenomic approaches to resolve species-level taxonomic conflicts and (2) conducting population genomic analysis to provide conservation recommendations.

We optimized the use of the universal Angiosperms353 target capture bait set in conservation genetics for the first time, demonstrating its versatility in being resolutive at phylogenetic and population genetics scales. Based on nuclear and plastid data, phylogenetic results resolved *A. cantabrica* as an independent clade from *A. adfinis*, thus supporting it as a distinct species. Phylogenetic incongruence between plastid

and nuclear data sheds new light on the origin of *A. cantabrica*, sister to *A. adfinis* with plastid sequence data and sister to a clade formed by *A. pyrenaica*, *A. laggeri*, *A. halleri* and *A. rioxana* using nuclear genes. Intronic regions were used to call SNPs and calculate population genetics parameters. Population genetic structure results divide the six populations into two conservation management units with a clear geographic separation (east and west) although low genetic differentiation ($F_{ST} = 0.05$) between them. We estimated the distribution range, population size and threats of *A. cantabrica* and, under the IUCN Red List criteria B1ab(ii,iii) + 2ab(ii,iii), classified it as Vulnerable (VU). We recommend translocations for the western genetic group due to its lower genetic diversity ($H_E = 0.143$) to increase effective population size and habitat threat management for the eastern genetic group ($H_E = 0.184$). We conducted ex-situ conservation collected and identified a potential micro-reserve.

Keywords: alpine ecosystem, Angiosperms353, conservation genetics, global warming, phylogenomics, population genomics, threatened species.

1 INTRODUCTION

Alpine environments, dominated by perennial herbs, face severe impacts from global climate change (Seddon *et al.* , 2016). Global warming has led to significant changes, such as the encroachment of woody subalpine plants, narrowing of alpine ecosystems (Capers & Stone, 2011) or the increase of diversity in European summits (Steinbauer *et al.* , 2018). While the loss of alpine habitat could be compensated by glacier retreatment (Whittaker, 1993; Losapio *et al.* , 2021), the snowline has been wholly lost in the lower and southern mountain regions where alpine plants are currently refuged in their ridges and peaks facing potential local extinction (Rumpf *et al.* , 2022). Global warming also affects alpine plant reproduction, including plant-pollinator interactions (Inouye, 2020) and seed germination (Mondoni *et al.* , 2012). Moreover, human disturbance in mountainous areas, such as civil infrastructures or recreation resources, leads to alpine habitat fragmentation and quality decline (Winkler, 2020; Chardon *et al.* , 2023). Therefore, alpine ecosystems and their species are considered vulnerable to environmental changes (Schwager & Berg, 2019).

By 2100, 36-55% of the alpine species in European mountains are predicted to lose more than 80% of their habitats (Inouye, 2020). However, limited information is available regarding the current conservation status of many European alpine species, such as those included in the *Androsace* L. section *Aretia* (L.). Section*Aretia* includes narrow endemics with low dispersal ability (Anderberg & Kelso, 1996), with 34 recognized species (Boucher *et al.* , 2021) mainly distributed in the "European Alpine System" (Ozenda, 1995). Only a handful of *Aretia* species have undergone threat assessments (Fasciani & Pace, 2015; Eustacchio *et al.* , 2023). In Spain, *Androsace cantabrica* (Losa & P. Monts.) Kress has been included in the list of priority species for conservation (Moreno Saiz*et al.* , 2008).

Androsace cantabrica is an endemic species to the central region of the North Iberian Cordillera Cantabrica (Fig. 1A; Kress, 1997). It is a perennial, monoecious, and allogamous plant with small, densely clustered rosettes. The stem is usually less than 5cm long, and the flower corolla is deep pink (Figs. 1B and C; Kress, 1997).*Androsace cantabrica* is known to occur in seven localities; however, population size estimates are only known in four of them (Fig. 1A, in red), with less than 6,000 individuals estimated across 20 1x1km UTM quadrats (Baudet *et al.* (2004). This species is found on siliceous or acidic substrates in mountainous areas above 2000m, typically in ridges, and often associated with low shrubs or pastures (Tejero *et al.* , 2022). The central distribution core is centred around the Tres Mares area, partially overlapping with the Alto Campoo ski resort. All population sites have traditionally been subjected to controlled burning to promote pasture development. Global warming will likely affect its reproductive output, like Tejero *et al.* (2022) observed lower germination rates in experiments with warmer temperatures. Baudet *et al.* (2004) proposed to categorise*A. cantabrica* as "Endangered" in the Spanish Red List, later confirmed by Moreno Saiz *et al.* (2008). However, *A. cantabrica* is not an accepted species name (https://powo.science.kew.org/, accessed 14th July 2024); instead, it is treated as a synonym of *A. adfinis* subsp. *adfinis* Biroli. This taxonomic uncertainty has conservation implications because the European conservation framework does

not consider it a species, thus not categorised as threatened. Scientific evidence is urgently needed to resolve this taxonomic conflict as a first step to provide effective conservation (Godfray *et al.* , 2004).

Molecular phylogenetics has emerged as a crucial tool for addressing taxonomic challenges (de Queiroz & Gauthier, 1992), and its resolution has greatly enhanced with the advent of high-throughput sequencing methodologies (Campos *et al.* , 2023). However, attempts to clarify species boundaries in this complex of *Androsace* species were mostly based on traditional molecular approaches. Previous morphological studies suggested that *A. cantabrica* was closely related to*A. laggeri* A.Huet (Kress, 1997). Genetic data using different molecular markers, such as the plastid *trn* L-F region and the internal transcribed spacer (ITS, Schneeweiss *et al.* , 2004), amplified fragment length polymorphism (AFLPs, Dixon *et al.* , 2008), and double digest restriction-site associated DNA (ddRAD-seq, Boucher *et al.* , 2021), proposed that *A. cantabrica* and*A. adfinis* are sister taxa, forming a *cantabrica-adfinis* clade sister to a clade formed by *A. halleri* L. and *A. laggeri* (/halleri clade hereafter), but with low bootstrap support. However, a more recent phylogenetic tree reconstructed with full plastome sequences suggests that the clade formed by *A. cantabrica* and *A. adfinis* is not sister to the /halleri clade (Smyčka *et al.,* 2022). Therefore, the evolutionary relationships among the Iberian *Androsace* section *Aretia* taxa and *A. adfinis* remain uncertain and still need to be resolved.

Targeted sequencing using the universal Angiosperms353 probe set can generate hundreds of homologous low-copy nuclear loci sequences, establishing it as a powerful tool in plant evolutionary studies (Johnson *et al.* , 2019). This approach is cost-effective and allows the use of herbarium materials in phylogenomic analysis (Brewer*et al.* , 2019). Nuclear genes can yield a distinct phylogenetic topology compared to plastid genes (Stubbs *et al.* , 2023). By combining these two genomic sources, researchers can explore reticulate evolution and potential hybrid origins more effectively (Vriesendorp & Bakker, 2005). However, most previous studies using Angiosperms353 data have primarily focused on clade boundaries at the genus, family, and order levels (e.g., *Nepenthes* (Nepenthaceae), Murphy *et al.* , 2020; Gentianales, Antonelli *et al.* , 2021; Primulaceae, Larson *et al.* , 2023), with few addressing species-level taxonomic conflicts (e.g., Campos *et al.* , 2023). In addition to its application in phylogenomics, Angiosperms353 data can be utilised in population genetic studies (Slimp *et al.* , 2021), which is invaluable for designing effective conservation plans for threatened species (Liu & Zhao, 1999; Xiong *et al.* , 2024). Compared to RAD-Seq (Davey & Blaxter, 2010), Angiosperms353 offers a more cost-effective alternative with reduced missing data, and it can be used in plants with different genome sizes (Slimp *et al.* , 2021). However, to our knowledge, the application of Angiosperms353 in practical conservation genetics has yet to be reported.

Our research objectives are threefold: to clarify the taxonomic status of *A. cantabrica* using nuclear Angiosperms353 loci and plastid data; to evaluate its threatened status and IUCN category; and to provide conservation recommendations for *A. cantabrica* based on population genetics analysis. This approach will also enable us to evaluate the effectiveness of Angiosperms353 in conservation genetics research.

2 MATERIALS AND METHODS

2.1 Plant material

In the summer of 2020, *A. cantabrica* populations were sampled from six locations that collectively represent its distribution range, as illustrated in Fig. 2A. Within these populations, its distribution is often fragmented, resulting in multiple subpopulations, notably in Tres Meras (TM). A total of thirty-five individuals were collected, with six individuals sampled from each population, except for Hoya Continua (HC), where only one individual was available for analysis. Three additional individuals were sampled from different subpopulations within TM, along with one more from a separate Valdecebollas (VB) subpopulation. Related taxa of *A. cantabrica* from the northern Iberian Peninsula were also sampled, including six individuals from an *A. halleri* subsp. *nuria* Schönsw. & Schneew. population (Fig. 2B; Suppl. Table 1). Fresh leaf tissue samples were dried in silica gel, and specimen vouchers were preserved at the JACA Herbarium (Suppl. Table 1). Specimens from the four Alps taxa were obtained from the Kew Herbarium (Suppl. Fig. 1), with additional sequencing data sourced for related taxa and *Primula matthioli* (L.) V.A.Richt. which was selected as an

outgroup (Suppl. Table 2). Field sampling also involved recording population sizes, key reproductive traits, and the conservation status of each population. We further reassessed the IUCN category following the standards set by the IUCN Species Survival Commission (2012).

2.2 Molecular methods to generate Angiosperms353 sequence data

Total DNA was isolated using a modified CTAB protocol (Doyle & Doyle, 1987). Genomic libraries were constructed as optimized in (Viruel*et al.* , 2019) using half volumes of the NEBNext® UltraTM II DNA Library Prep Kit for Illumina® (New England Biolabs, Ipswich, MA, United States), purified using AMPure XP magnetic beads and multiplexed with NEBNext® Multiplex Oligos for Illumina® (Dual Index Primer Sets I and II). Equimolar pools containing twelve genomic libraries were enriched with half-reactions of the Angiosperms353 probe kit (Johnson *et al.* , 2019; Baker *et al.* , 2022) following myBaits®kit manual v5.03 (Arbor Biosciences). DNA concentrations were calculated using a QuantusTM fluorometer (Promega Corp.), and an Agilent 4200 TapeStation (Agilent Technologies, Santa Clara, CA, United States) was used to assess fragment length. Sequencing was performed on a HiSeq (Illumina, Inc.) by Macrogen (Seoul, South Korea), producing 150 bp paired-end reads.

2.3 Quality filtering of FASTQ raw data

The raw sequencing files were checked for quality using FastQC (Andrews, 2010) and MultiQC (Ewels *et al.* , 2016), then trimmed using Trimmomatic (Bolger *et al.* , 2014) to remove adapters and reads with low quality (LEADING:30 TRAILING:30). Paired reads were used as input in HybPiper (Johnson *et al.* , 2016) and the "mega353" target file (McLay *et al.* , 2021) was used to recover Angiosperms353 loci sequences. Reads were mapped to the mega353 reference using BWA (Li & Durbin, 2009) and were then assembled *de novo* using SPAdes (Bankevich *et al.* , 2012). Exon, intron and supercontig sequences were recovered using Exonerate (Slater & Birney, 2005). We excluded genes flagged with paralog warnings by Hybpiper and genes that were not recovered in at least 75% of samples.

We extracted protein-coding and intergenic sequences from the complete plastid genome of *Androsace mariae* Kanitz (GenBank: MT732944) and removed duplicates and sequences shorter than 200bp, resulting in a plastome reference of 125 plastid fragments. This reference was then used to recover plastid sequences with HybPiper, as described above.

2.4 Estimation of ploidy

We estimated the ploidy levels of *Androsace* samples using nQuire (Weiß *et al.* , 2018) following the approach by Viruel *et al.* (2019). To prepare the reference file for nQuire, we extracted the longest exon recovered per gene using bioawk (available at https://github.com/lh3/bioawk) and excluded any genes that received paralog warnings from the initial reference. We then evaluated ploidy by analyzing the delta log-likelihood (Δ log$L$ values produced by nQuire across three models —diploid, triploid, and tetraploid— to identify the best-supported ploidy level.

2.5 Phylogenomic analysis

We reconstructed phylogenetic trees using one representative sample from each of the six *A. cantabrica* populations and all other taxa. Some genome skimming data available online (Suppl. Table 2) were added for reconstructing the plastid phylogeny.

Loci sequences were aligned with MAFFT (–auto; Katoh & Standley, 2013), and then the alignments were trimmed with trimAl (-automated1; Capella-Gutiérrez *et al.* , 2009). For both Angiosperms353 loci and the 125 recovered plastid fragments, we inferred nuclear and plastid phylogenetic trees using both coalescent and concatenated maximum likelihood (ML) approaches. In the coalescent approach, we inferred single-locus phylogenetic trees from each trimmed alignment using IQ-TREE (Minh *et al.* , 2020) with 1000 ultrafast bootstrap replicates (-bb 1000; Hoang *et al.* , 2018), and branches with less than 10% bootstrap support were collapsed with Newick utilities (Junier & Zdobnov, 2010). We then used ASTRAL-III (Zhang *et al.* ,

2018) to infer the species tree (hereafter, ASTRAL tree), applying the "-t 3" flag to annotate local posterior probabilities (LPP) for each node.

In the concatenated ML approach, all trimmed alignments were concatenated with FASconCAT-G (Kück & Meusemann, 2010). The best-fit model inferred by IQ-TREE (-m MFP) was applied in RAxML-NG (Kozlov*et al.* , 2019) to infer the species tree using the concatenated partitioned matrix with 1000 bootstrap replicates (–tree pars{20} –bs-trees 1000; hereafter, RAxML tree). Additionally, we implemented a greedy strategy (Lanfear *et al.* , 2012) with the relaxed hierarchical clustering algorithm (Lanfear *et al.* , 2014) to select the best partition model, which was applied in IQ-TREE to infer the species tree with 1000 SH-like approximate likelihood ratio test replicates (-alrt 1000; Guindon *et al.* , 2010) and 1000 ultrafast bootstrap replicates (hereafter, IQ-partition tree). We visualized the phylogenetic trees using Dendroscope (Huson & Scornavacca, 2012) and FigTree (available at https://github.com/rambaut/figtree).

To investigate potential phylogenetic conflicts and signs of reticulate evolution, we used SplitsTree4 (Huson & Bryant, 2006) to create a split network based on the Neighbor-Joining algorithm with the Angiosperms353 data. In the resulting network, we masked specific samples to retain only those within the *A. cantabrica, A. adfinis* and /halleri clades for focused analysis.

2.6 Variant calling and filtering

To compare the population genetic results between the threatened*A. cantabrica* and the non-threatened *A. halleri* , we performed variant calling and population genetic analyses using 35*A. cantabrica* samples from six populations and six samples of*A. halleri* subsp. *nuria* from a single population. We followed the pipelines and scripts provided by https://github.com/lindsawi/HybSeq-SNP-ExtractionSlimp *et al.* (2021, available at) with some modifications. In their pipeline, Slimp *et al.* (2021) used supercontig sequences, demonstrating that most genetic variation occurred in flanking non-coding regions, which tend to accumulate mutations quickly due to limited functional constraints (Palumbi, 1996). We used sequences from supercontig and intron regions separately for comparative analyses. We prepared a reference file for supercontigs and introns using the same approach described above to generate the nQuire reference, in this case, selecting each gene's longest supercontig and intron sequence. Additionally, we excluded any genes flagged by HybPiper for paralogy warnings (Bryc *et al.* , 2013).

To obtain single-nucleotide polymorphisms (SNPs) data, we used the framework developed by DePristo *et al.* (2011) in GATK (McKenna*et al.* , 2010). We combined aligned and unaligned reads to the reference, removed duplicate sequences, and performed genotype calling collectively for all samples after generating preliminary variants individually for each sample (Poplin *et al.* , 2018) in a Variant Call Format (VCF) file. The filtering conditions we conducted on the initial VCF file included using a "hard filter" (QD $<$ 5.0 || FS $>$ 60.0 || MQ $<$ 40.0 || MQRankSum $<$ -12.5 || ReadPosRankSum $<$ -8.0), removing indels and SNPs with missing data in GATK, and removing linked SNPs in PLINK (Chang*et al.* , 2015). We conducted a Base Quality Score Recalibration in GATK and repeated the variant calling step. To address the potential effects of polyploidy, which can artificially increase heterozygosity and allelic richness (Hokanson & Hancock, 1998), it is essential to filter fixed heterozygotes in SNP datasets in polyploid species (e.g., Douglas et al., 2015; Cornille et al., 2016; Blischak et al., 2018; Pavan et al., 2020). We removed loci with observed heterozygosity (H_O) $>$ 0.5 from *A. cantabrica* (Appendix S1) data using the R package "VCFR" (Knaus & Grünwald, 2017). We established this filter by comparing heterozygosity and inbreeding coefficient results for the diploid *A. halleri* to those obtained for the tetraploid *A. cantabrica* (see Appendix S1 and Results). The unfiltered data were retained for comparative studies.

2.8 Population genetic indicators for conservation recommendations

Following the framework proposed by Ottewell *et al.* (2016) for conservation planning, we calculated three population genetic indicators: genetic differentiation (F_{ST}), genetic diversity (observed and expected heterozygosity,H_O and H_E) and inbreeding coefficient (F_{IS}). To identify conservation management units and set conservation priorities, we analyzed the population genetic structure as outlined by Fraser & Bernatchez (2001) using two primary approaches: (1) Principal Coordinate Analysis (PCoA): We generated genetic

distance-based PCoA plots in GeneAlEx (Peakall & Smouse, 2012) to visualize genetic relationships among populations; (2) Clustering Analysis: We inferred the optimal number of genetic clusters (K) in STRUCTURE (Pritchard *et al.* , 2000) using 10,000 burnin and 100,000 MCMC generations, with ten replicates per K value, testing up to K equal to the number of populations plus two. The most likely K was determined following Evanno *et al.* (2005) approach as implemented in Structure Harvester (Earl & vonHoldt, 2012), and the results were visualized with StructuRly (Criscuolo & Angelini, 2020). File format conversions between software were conducted in PGDSpider (Lischer & Excoffier, 2012).

3 RESULTS

3.1 Sequence data and recovery

On average, 32.98% of reads were mapped to target regions, ranging from 14.88% to 66.83%. The sequence length recovery rate, relative to gene lengths in the "mega353" target file, averaged 80.56%, ranging from 25.55% to 86.52% (Suppl. Table 3). Sample 21B61 was removed from analysis due to the absence of gene recovery. Sample Z4 had a notably low recovery rate of 25.55%, the lowest among *A. cantabrica* samples (Suppl. Table 3). This sample was temporarily retained to assess the impact of low-recovery samples and variation in recovery rates on variant calling and population genetic analyses. The average plastid sequence length recovery rate was 92.03%, ranging from 60.69% to 98.64%, excluding herbarium samples (Suppl. Table 4).

3.2 Phylogenetic trees and network

The phylogenetic trees reconstructed with the Angiosperms353 loci (hereafter, Angiosperms353 trees) show strong support, with bootstrap values above 90% for most inter-species nodes (Figs. 3A and Suppl. Fig. 2). The IQ-partition and RAxML trees are largely congruent, with*A. cantabrica* resolved as a sister to the /halleri clade (Fig. 3A). In the ASTRAL tree, a phylogenetic conflict was observed regarding the placement of *A. rioxana* A.Segura (Suppl. Fig. 2); here,*A. rioxana* was resolved as a sister to the /halleri and *A. cantabrica* clades rather than being embedded within the /halleri clade, as seen in other phylogenetic trees (Fig. 3). In the plastid tree,*A. cantabrica* is resolved as a sister to *A. adfinis* , with these two forming a clade that is sister to *A. alpina* (L.) Lam.. The /halleri clade is resolved as a sister to the clade formed by*A. alpina* , *A. adfinis* , and *A. cantabrica* (Fig. 3B).

Our ploidy estimation analysis predicts that *A. rioxana* , *A. cantabrica* and *A. adfinis* subsp. *brigantiaca* (Jord. & Fourr.) Kress are tetraploid, whereas all other species tested are estimated as diploid (Suppl. Table 5). The split network analysis reveals indicated gene flow and reticulate evolution in the formation of these species, with *A. cantabrica* emerging as a relatively independent evolutionary branch related to *A. adfinis* and the /halleri clade, yet showing more ancestral gene flow with the /halleri species (Fig. 3C).

3.3 Variant calling and filtering

Initial filtering steps produced several SNP datasets, including a "hard filter" dataset (SNPs_HF), a dataset excluding SNPs with missing data (SNPs_HF_NMD), and a set excluding linked SNPs (SNPs_HF_NMD_NLD) (Table 1). Removing the low-recovery sample Z4 from the analysis altered the STRUCTURE results, reducing the optimal number of clusters from$K$ =5 (with Z4) to K =2 (without Z4; (all comparative analysis results for this section can be found in Appendix S1). The population structure position of Z4 diverged from its native population. Consequently, we opted to discard Z4 in subsequent analyses.

Despite the high SNP count from supercontigs (34,138 SNPs after linkage disequilibrium filtering), no significant difference in population genetics parameters (F_{ST},H_E and F_{IS}) was observed between datasets derived from supercontigs and introns (Table 1, Appendix S1). Therefore, intron-derived SNPs were selected for final analyses to ensure robustness and compatibility across related studies.

Table 1. Number of variants produced by different filtering steps (Appendix S1).

Condition	N	SNPs+ indels	Only SNPs

Table 1. Number of variants produced by different filtering steps (Appendix S1).

A. cantabrica supercontig (with Z4)	34	62132	49523
A. cantabrica supercontig (without Z4)	33	61987	49412
A. cantabrica intron	33	8760	6964
A. halleri supercontig	6	25253	20827
A. halleri intron	6	2759	2253

3.4 Population genetic analysis

Using two SNP datasets, one with 869 SNPs and one filtered to retain only those with H_O [?] 0.5 (462 SNPs), we observed consistent population structure results in PCoA and STRUCTURE analyses, except for the position of individual TM_6 in the PCoA plot under 869 SNPs (Fig. 4). The population division separated western (Group W: LL, EP, HC) and eastern populations (Group E: TM, VB, CD) (Fig. 4). Although the most likely number of K estimated was 3 for the 869 SNPs matrix, both groups maintained similar genetic structure.F_{ST} values indicate low differentiation between groups (i.e., $F_{ST} < 0.15$), aligning with limited recent gene flow observed between them (Table 2).

Table 2. Genetic differentiation (F_{ST}) values between five location populations and two groups, with the lower triangle bei

Populations
TM
VB
CD
LL
EP
Group W
Group E

Table 3. Sample size (N), observed heterozygosity (H_O), expected heterozygosity (H_E) and inbreeding coefficients (F_{IS}) va

Populations
A. halleri
A. cantabrica
TM
VB
CD
LL
EP
Mean
Group E
Group W
Mean

Using *A. cantabrica* intron-derived data, we observedF_{IS} values (> -0.5) when SNPs with$H_O > 0.5$ were not removed (Table 3). However,F_{IS} values became relatively low negative after filtering. In both cases, Group E exhibits relatively higherH_E values than Group W, with the TM population having the highest H_E value. The genetic diversity of the TM population remains lower than that of the non-threatened and diploid *A. halleri* (Table 3).

3.5 Reproductive biology, distribution range and population size of *A. cantabrica*

Both pistils and stamens of *A. cantabrica* are situated within the corolla tube (Fig. 5A), which makes wind pollination highly improbable. Initially, the corolla exhibits a prominent yellow ring within the throat, which fades after pollination along with the pink of the corolla, serving as a signal to potential pollinators (Fig. 5B). This signalling contrasts with other *Androsace* species, where the colour change post-pollination is typically more pronounced transitioning to a deeper hue (Fig. 5C). *Androsace cantabrica* generally attracts various Hymenoptera and Diptera pollinators (Fig. 5D) and shows notable variation in peduncle length due to growth conditions: individuals growing amidst taller weeds and shrubs develop longer peduncles (Figs. 5E and F).

Table 4 Estimated population sizes and threat factors of *A. cantabrica* populations.

Location
TM
VB
CD
LL
HC
EP

Shrubs and heath are colonizing mountain peaks and ridges. + Ski resorts and hiking trails. # This number corresponds to

Additionally, our findings indicate a relatively low reproductive yield, as *A. cantabrica* only produces 3-4 seeds per flower (Fig. 5G), with notably larger seeds than most other *Androsace* species (Fig. 5H). These traits and restricted habitat distribution above 2,000 m highlight the species' adaptation to its alpine environment and vulnerability to shifts in habitat quality and availability.

As detailed in Table 4, our population size estimates indicate around 100 individuals in the Peña Prieta subpopulation and approximately 2,500 in Tres Mares (TM), where we observed a decline in numbers likely due to pressures from nearby ski resorts and hiking trails. Other populations, including those in Hoya Continua (HC) and other small subpopulations, collectively estimate fewer than 6,000 individuals. Key threats to*A. cantabrica* include habitat encroachment from shrub expansion, particularly in areas that were once controlled by herbivory or fire, which are now dominated by taller vegetation (Figs. 6A-C). These encroaching shrubs — namely *Vaccinium uliginosum* L.,*Juniperus communis* var. *saxatilis* Pall., and *Calluna vulgaris* (L.) Hull (Figs. 6D–F) — create competitive pressures that limit the available light and space for *A. cantabrica* . Consequently, most individuals are observed growing along the margins of these shrubs.

3.6 Conservation status and recommendations

Based on our distribution range and population size results (Section 3.5), we propose that *A. cantabrica* be classified as Vulnerable (VU) under the IUCN Red List criteria: B1ab(ii,iii) + 2ab(ii,iii). This categorization is supported by an estimated extent of occurrence (EOO) of less than 20,000 km^2 (B1); an area of occupancy (AOO) of less than 2,000 km^2 (B2); fewer than 10 locations (a), and ongoing declines in both area of occupancy (b(ii)) and habitat extent and quality (b(iii)).

Considering the population genetic structure analysis results, we compared the genetic diversity values between the two genetic groups and observed a higher genetic diversity in Group E. Conservation measures for Group E are recommended to primarily focus on managing ecological and demographic threats at the species level to maintain population size and gene flow. For Group W, emphasis on in-situ diversity restoration and facilitating translocations of individuals to increase population size, especially for the small-size (individuals < 100) EP population, are recommended.

4 DISCUSSION

Our study reveals that *Androsace cantabrica* represents a genetically distinct lineage within the /halleri clade, underscoring its unique evolutionary history and conservation significance. Our study exemplifies the use

of Angiosperms353 target capture data in species delimitation using phylogenomic analysis, and we establish that *A. cantabrica* maintains high phylogenetic differentiation from closely related species and clades, supporting its taxonomic independence and validating its conservation priority. The population genetic diversity identified through our analysis provides a foundation for actionable conservation strategies, especially in light of anthropogenic pressures and climate change.

4.1 Species boundary delimitation of *Androsace cantabrica* using phylogenomics

Clarifying the taxonomic status of *A. cantabrica* has both scientific and conservation implications, as distinguishing it from closely related taxa is crucial for understanding its threatened status and prioritizing conservation measures (Godfray et al., 2004; Ottewell et al., 2016; Kress et al., 2017). We implemented a phylogenomic approach to resolve whether *A. cantabrica* is a valid species and, therefore, threatened or if it is part of a wider and non-threatened taxon (*A. adfinis* subsp. *adfinis*). Phylogenetic reconstructions based on hundreds of nuclear loci and plastid sequences obtained from Angiosperms353 target capture data consistently resolved*A. cantabrica* as a monophyletic clade, confirming its phylogenetic distinctiveness in alignment with previous morphological (Kress, 1997) and karyotypic data (Kress, 1984). Our phylogenomic approach improved our understanding of *A. cantabrica* and allowed us to investigate its distribution, genetic diversity and evolutionary origin and supported its categorization as a valid species in need of conservation.

The placement of *A. cantabrica* as sister to the /halleri clade in the nuclear Angiosperms353 phylogenetic trees, rather than to *A. adfinis,* aligns with the morphological groupings of *Androsace* sect. *Aretia* as per Smith & Lowe (1997). Interestingly, while nuclear Angiosperms353 data positioned *A. cantabrica* as a sister to the /halleri clade, plastid data and previous studies suggested a closer phylogenetic relation to *A. adfinis* . This topological incongruence may reflect a divergence in evolutionary paths between the plastid and nuclear genomes, a phenomenon observed across many plant species (e.g., Galbany-Casals et al., 2014; Viruel et al., 2018 Favre et al., 2022; Liu et al., 2023).

The evolutionary origin and taxonomic identity of *A. cantabrica* have been subjects of significant debate. Initially, it was hypothesized that *A. cantabrica* was an allopolyploid species resulting from hybridization between *A. laggeri* ($2n = 38$, localized in the central Pyrenees) and *A. halleri* ($2n = 38$, distributed in the Cantabrian Mountains, Pyrenees, Massif Central, and Vosges; Kress, 1984). However, this hypothesis was challenged by Dixon et al. (2008), who used amplified fragment length polymorphism (AFLP) data to refute a close relationship between *A. cantabrica* and either*A. halleri* or *A. laggeri* . Instead, they suggested that*A. cantabrica* was an autopolyploid related to the southwestern Alps' *A. adfinis s.l.* , which includes *A. adfinis* subsp.*adfinis* , *A. adfinis* subsp. *puberula* (Jord. & Fourr.) Kress, and *A. adfinis* subsp. *brigantiaca* . Our data indicate that *A. cantabrica* is a polyploid and likely has a complex evolutionary history. While Dixon et al. (2008) proposed an autopolyploid origin closely related to *A. adfinis* , our analysis suggests a different scenario, albeit without conclusively identifying the exact polyploidy type. Although paralogous genes were limited in our analysis, which included *A. cantabrica* samples, definitive proof of an autopolyploid origin is absent. For example, *A. adfinis* subsp. *brigantiaca* is suspected of being a recent hybrid (Boucher et al., 2016) with tetraploid features, yet only exhibited six paralogous genes in HybPiper analyses. Possible scenarios for *A. cantabrica* origin include an ancient homoploid hybridization event involving the ancestors of *A. cantabrica* and the /halleri clade, followed by local polyploidization, or an initial allopolyploidization. In any case, given its phylogenetic, morphological, and karyological uniqueness, *A. cantabrica* should be considered a valid species with an evolutionary trajectory shaped by rapid speciation, introgression, and possibly hybridization in alpine environments (Hibbins et al., 2020; Smyčka et al., 2022).

4.2 Conservation status and strategies for *Androsace cantabrica*

Our results support classifying *Androsace cantabrica* as Vulnerable (VU) based on the IUCN Red List framework, proposed under criteria B1ab(ii,iii) + 2ab(ii,iii), which accounts for its restricted distribution, ongoing declines in area of occupancy, and quality of habitat. Our research indicates that the estimated extent of occurrence (EOO) is below 20,000 km^2, with an area of occupancy (AOO) under 2,000 km^2, thus meeting the spatial thresholds for Vulnerable status. This is similar to the reasoning to categorise

A. hemisphaerica Ludlow as Endangered due to a very limited distribution range (i.e., EOO of 1,008 km2; Bhutan Endemic Flowering Plants Workshop, 2017). Furthermore,*A. cantabrica* populations are restricted to fewer than ten isolated locations, each experiencing habitat encroachment from shrub expansion and ongoing degradation due to human activities, consistent with criteria B1ab(ii,iii) + 2ab(ii,iii). One notable limitation of our IUCN assessment is the potential underestimation of population size due to the species' association with dense shrub margins, making locating individuals challenging. The ongoing shrub expansion reduces the visibility of A. cantabrica and exacerbates competition for light and space, threatening population stability across its range. Additionally, while the current distribution data meets IUCN's "Vulnerable" criteria, further decline in shrub-controlled habitats could eventually lead to "Endangered" status.

Our revised population estimates (see Table 4) underscore the challenges of accurately assessing population sizes, as we observed potential discrepancies with past reports, especially in highly disturbed areas like Tres Mares (TM). For instance, TM showed a marked decline in population size compared to previous counts, likely due to human disturbances from ski resort expansions, hiking trail use, and trampling. Such disturbances, combined with the environmental pressures from global warming, drive shrub expansion into alpine zones. In the TM region, this shrub encroachment and diminished herbivory and fire threaten the remaining open meadow habitats essential for *A. cantabrica* survival. These disturbances significantly affect the species' distribution and resilience, as reflected in the observed population declines and limited recruitment in this location.

Our results demonstrate the strengths of Angiosperms353 as an effective tool for refining the conservation status of polyploid species like*A. cantabrica* . The population genetic analysis divides *A. cantabrica* populations into two genetic conservation units: Group W, comprising western populations with lower genetic diversity, and Group E, which includes eastern populations showing relatively higher genetic diversity (H_E values as high as 0.184). While the genetic data confirms low inbreeding and healthy population structure across regions, they also reveal low genetic diversity within smaller populations (Group W), suggesting that some populations' habitat fragmentation and small sizes could increase genetic vulnerability over time. For Group W, in-situ conservation measures should focus on bolstering genetic diversity by translocating individuals from genetically diverse populations within this group. The Peña Prieta population has an estimated population of approximately 1,600 individuals fragmented into smaller subpopulations, such as the LL population near Peña Prieta, with approximately 100 individuals. Future work should involve gathering complementary genetic data from HC and Peña Prieta localities to identify potential donor populations with higher genetic diversity within the same genetic group and conducting translocations to strengthen the population size and genetic diversity in the EP population. Although the TM population exhibits the highestH_E, translocating individuals from TM (i.e., Group E) to Group W is not advisable due to the potential risk of outbreeding depression (Lynch, 1991). Avoiding translocations between genetically differentiated populations is crucial without experiments investigating the risk of outbreeding depression, as it can reduce fitness through adverse breeding effects (Liu & Zhao, 1999). After any translocation efforts, if necessary, establishing a monitoring program to track fruiting rates, seed setting, and seedling survival will be essential for assessing population health and adaptability (Liu & Zhao, 1999). Despite the relatively high H_E of the TM population, it is still lower than that of other non-threatened taxa (e.g., *A. halleri* subsp. *nuria*), emphasizing the importance of mitigating anthropogenic disturbances in the TM area to preserve its genetic diversity. As an overall recommendation, conservation practices should focus on the genetic group W by reducing threats where appropriate and feasible, for example, reducing shrub competition to improve habitat suitability or collaborating with ski resorts near Peña Prieta to develop conservation and sustainable practices that mitigate human impact on surrounding habitats.

Conversely, for Group E, conservation efforts should emphasize mitigating ecological threats by managing shrub encroachment to maintain habitat openness essential for *A. cantabrica* 's survival. Shrub encroachment poses a significant threat to the TM area, with impacts particularly severe at higher altitudes where open habitats are more vulnerable to invasive shrub growth. This underscores the need for targeted interventions, such as controlled burning or grazing, where shrubs colonize high-altitude habitats and limit suitable growing spaces for *A. cantabrica* . Additionally, we recommend long-term monitoring of all known populations

alongside efforts to locate and characterize additional populations.

Ex-situ conservation approaches are efficient for the long-term conservation of threatened species (e.g., Schoen and Brown, 2001; Wambugu et al., 2023). In 2021, more than 2,000 seeds were collected as part of the PRIOCONEX project (https://sites.google.com/aranzadi.eus/prioconex) to be stored at the Seed Bank in Gipuzkoa, Spain (Accession number 52/2020). Seeds from more than 50 mother plants were collected from the TM area to preserve most of its genetic diversity, and several morphometric measurements and germination protocols were conducted (Tejero et al., 2022). Ex-situ conservation of seeds from the western group is also recommended, but due to the scarcity of the species in the area, this task might be demanding in time and prospection.

Micro-reserves might be very efficient for conservation of cryptic populations of endangered plant species without a strong impact in local land use (Laguna, 2000; Médail et al., 2021). We recommend the creation of a micro-reserve in a specific locality in TM area (ETRS89 UN 86101 65592; 2058 MASL) which hosts the most conspicuous and dense known population and with the highest heterozigosity values. Additionally, it seems valuable to initiate high-altitude reintroductions given the projected habitat loss by 2070 (MITECO, 2011). Such measures align with PRIOCONEX, which focuses on ex-situ conservation in response to climate impacts on alpine habitats (Yuste et al., 2021). Notably, *A. cantabrica* produces larger seeds than most *Androsace* species, likely indicative of a K-selected reproductive strategy to adapt to harsh environmental conditions (Sam, 2013). However, seed production is limited (3–4 seeds per flower), and *A. cantabrica* exhibits cold-dependent germination (Tejero et al., 2022), complicating in-situ population expansion. Further research to investigate the effects of environmental changes on germination and the survivability of seedlings will provide essential knowledge for the long-term conservation of this species.

4.3 Practical considerations of using Angiosperms353 in conservation genetics

Our results exemplify the potential of using target capture sequencing with the universal bait panel Angiosperms353 for population genetic studies (Slimp et al., 2021), offering high data quality and valuable cost efficiency for population-level analyses. Angiosperms353 target capture has broad applications for conservation genetics, effectively capturing intraspecific variation within populations and supporting conservation genomics for rare or threatened taxa. The ability to integrate herbarium samples makes it particularly suited for conservation genetics (Slimp et al., 2021). Although Phang et al. (2023) found that population structure analysis using Angiosperms353 yielded limited resolution within species, our results demonstrate a higher resolution, likely because we did not combine multiple species when calling SNPs.

Based on SNPs data extracted from intronic regions, we developed specific recommendations: 1) inspecting samples with low target sequence recovery, as generated in HybPiper, which could serve as indicators to assess subsequent SNPs recovery and/or errors in population genetic structure results (Yi & Latch, 2022), although the levels of missing data are expected to be lower than in other approaches such as RADseq (Slimp et al., 2021); 2) prioritization of intronic SNPs for population genetics due to their higher mutation rate, although this could vary in other species; 3) exclusion of paralog genes (Bryc et al., 2013), and 4) removing highly heterozygous SNPs ($H_O > 0.5$) to enhance data accuracy in polyploid species.

5. Conclusions

Our study confirms the taxonomic and phylogenetic distinctiveness of*Androsace cantabrica* and emphasizes the utility of Angiosperms353 target capture data in resolving species-level conflicts within complex plant groups. We demonstrated that *A. cantabrica* is a distinct species requiring conservation action, counter to previous hypotheses suggesting a close affiliation with *A. adfinis* subspecies. The contrasting results between nuclear and plastid phylogenies highlight the complex evolutionary history of *A. cantabrica* and related taxa, underlining the need for integrated molecular approaches to untangle rapid radiations and reticulate evolution. Future research should further investigate the polyploid origin of *A. cantabrica* and monitor its genetic structure and diversity in the face of ongoing climate change. Long-term conservation planning, including habitat management, controlled translocations, and ex-situ conservation, will be vital to prevent genetic erosion and habitat loss for this Vulnerable alpine species.

The Angiosperms353 target capture approach proved effective for population-level conservation genetics, even in a polyploid species like*A. cantabrica* . Moreover, we advocate for the adoption of Angiosperms353 in similar conservation genetics studies, given its cost-effectiveness, sample efficiency, and the potential to incorporate herbarium samples whilst enabling comparative studies between species based on population genetic metrics calculated using the same set of molecular markers.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA ARCHIVING STATEMENT

Data for this study are available at: to be completed after the manuscript is accepted for publication.

LITERATURE CITED

Anderberg AA, Kelso S . 1996 . Phylogenetic implications of endosperm cell wall morphology in Douglasia, *Androsace* , and*Vitaliana* (Primulaceae). *Nordic Journal of Botany* 16: 481–486.**Andrews S . 2010** . Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Available at https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [accessed 07 July 2023].**Antonelli A, Clarkson JJ, Kainulainen K, Maurin O, Brewer GE, Davis AP, Epitawalage N, Goyder DJ, Livshultz T, Persson C, Pokorny L, Straub SCK, Struwe L, Zuntini AR, Forest F, Baker WJ . 2021** . Settling a family feud: a high-level phylogenomic framework for the Gentianales based on 353 nuclear genes and partial plastomes. *American Journal of Botany* 108: 1143–1165.**Baker WJ, Bailey P, Barber V, Barker A, Bellot S, Bishop D, Botigué LR, Brewer G, Carruthers T, Clarkson JJ, Cook J, Cowan RS, Dodsworth S, Epitawalage N, Françoso E, Gallego B, Johnson MG, Kim JT, Leempoel K, Maurin O, Mcginnie C, Pokorny L, Roy S, Stone M, Toledo E, Wickett NJ, Zuntini AR, Eiserhardt WL, Kersey PJ, Leitch IJ, Forest F . 2022** . A Comprehensive Phylogenomic Platform for Exploring the Angiosperm Tree of Life. *Systematic Biology* 71: 301–319.**Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA .2012** . SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* 19: 455–477.**Baudet A, Blanca G, Güemes J, Moreno Saiz J, Ortiz S . 2004** . *Androsace cantabrica* (Losa & P. Monts.) Kress. *Atlas y Libro Rojo de la Flora Vascular de España* . Madrid, 580–581.**Blischak PD, Kubatko LS, Wolfe AD. 2018.** SNP genotyping and parameter estimation in polyploids using low-coverage sequencing data. Bioinformatics 34: 407–415.**Bolger AM, Lohse M, Usadel B . 2014** . Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.**Boucher FC, Dentant C, Ibanez S, Capblancq T, Boleda M, Boulangeat L, Smyčka J, Roquet C, Lavergne S . 2021** . Discovery of cryptic plant diversity on the rooftops of the Alps. *Scientific Reports* 11: 11128.**Boucher FC, Zimmermann NE, Conti E . 2016** . Allopatric speciation with little niche divergence is common among alpine Primulaceae. *Journal of Biogeography* 43: 591–602.**Brewer GE, Clarkson JJ, Maurin O, Zuntini AR, Barber V, Bellot S, Biggs N, Cowan RS, Davies NMJ, Dodsworth S, Edwards SL, Eiserhardt WL, Epitawalage N, Frisby S, Grall A, Kersey PJ, Pokorny L, Leitch IJ, Forest F, Baker WJ . 2019** . Factors Affecting Targeted Sequencing of 353 Nuclear Genes From Herbarium Specimens Spanning the Diversity of Angiosperms. *Frontiers in Plant Science* 10.**Bryc K, Patterson N, Reich D . 2013** . A Novel Approach to Estimating Heterozygosity from Low-Coverage Genome Sequence.*Genetics* 195: 553–561.**Bhutan Endemic Flowering Plants Workshop. 2017.** *Androsace hemisphaerica* . The IUCN Red List of Threatened Species 2017: e.T83595192A84447331. https://dx.doi.org/10.2305/IUCN.UK.2017-3.RLTS.T83595192A84447331.en. Accessed on 30 October 2024.**Campos M, Kelley E, Gravendeel B, Médail F, Maarten Christenhusz JM, Fay MF, Catalán P, Leitch IJ, Forest F, Wilkin P, Viruel J . 2023** . Genomic, spatial and morphometric data for discrimination of four species in the Mediterranean Tamus clade of yams (Dioscorea, Dioscoreaceae). *Annals of Botany* 131: 635–654.**Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T .2009** . trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinfor-*

matics 25: 1972–1973.**Capers RS, Stone AD . 2011** . After 33 Years, Trees More Frequent and Shrubs More Abundant in Northeast U.S. Alpine Community. *Arctic, Antarctic, and Alpine Research* 43: 495–502.**Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ . 2015** . Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 4: s13742-015-0047–8.**Chardon NI, Stone P, Hilbert C, Maclachlan T, Ragsdale B, Zhao A, Goodwin K, Collins CG, Hewitt N, Elphinstone C .2023** . Species-Specific Responses to Human Trampling Indicate Alpine Plant Size Is More Sensitive than Reproduction to Disturbance.*Plants* 12: 3040.**Coomes DA, Grubb PJ . 2003** . Colonization, tolerance, competition and seed-size variation within functional groups. *Trends in Ecology & Evolution* 18: 283–291.**Cornille A, Salcedo A, Kryvokhyzha D, Glémin S, Holm K, Wright SI, Lascoux M. 2015.** Genomic signature of successful colonization of Eurasia by the allopolyploid shepherd's purse (*Capsella bursa-pastoris*). *Molecular Ecology* 25: 616-629.**Criscuolo NG, Angelini C . 2020** . StructuRly: A novel shiny app to produce comprehensive, detailed and interactive plots for population genetic analysis. *PLOS ONE* 15: e0229330.**Davey JW, Blaxter ML . 2010** . RADSeq: next-generation population genetics. *Briefings in Functional Genomics* 9: 416–423.**DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ . 2011** . A framework for variation discovery and genotyping using next-generation DNA sequencing data.*Nature Genetics* 43: 491–498.**Dixon CJ, Schönswetter P, Schneeweiss GM . 2007** . Traces of ancient range shifts in a mountain plant group (*Androsace halleri* complex, Primulaceae).*Molecular Ecology* 16: 3890–3901.**Dixon CJ, Schonswetter P, Schneeweiss GM . 2008** . Morphological and Geographical Evidence are Misleading with Respect to the Phylogenetic Position and Origin of the Narrow Endemic Polyploid *Androsace cantabrica* (Primulaceae). *Systematic botany* 33: 384–389.**Douglas G, Gos G, Steige K, Salcedo A, Holm K, Josephs EB, Arunkumar R, J. Ågren A, Hazzouri KH, Wei Wang, Platts AE, Williamson RJ, Neuffer B, Lascoux M, Slotte T, Wright SI. 2015** . Hybrid origins and the earliest stages of diploidization in the highly successful recent polyploid *Capsella bursa-pastoris* . *PNAS* 12: 2806–2811.**Doyle JJ, Doyle JL (Eds.) . 1987** . A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical bulletin* : 11–15.**Earl DA, vonHoldt BM . 2012** . STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.**Eustacchio E, Bonelli M, Beretta M, Monti I, Gobbi M, Casartelli M, Caccianiga M . 2023** . Pollen and floral morphology of *Androsace* brevis (Hegetschw.) Ces. (Primulaceae), a vulnerable narrow endemic plant of the Southern European Alps.*Flora* 301: 152256.**Evanno G, Regnaut S, Goudet J .2005** . Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14: 2611–2620.**Ewels P, Magnusson M, Lundin S, Käller M .2016** . MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32: 3047–3048.**Fasciani P, Pace L . 2015** . Conservation of Endangered Species: *Androsace mathildae* Levier (*Primulaceae*) in Central Italy. *American Journal of Plant Sciences* 06: 3175.**Favre A, Paule J, Ebersbach J .2022** . Incongruences between nuclear and plastid phylogenies challenge the identification of correlates of diversification in Gentiana in the European Alpine System. *Alpine Botany* 132: 29–50.**Fraser DJ, Bernatchez L . 2001** . Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology* 10: 2741–2752.**Frattaroli AR, Di Martino L, Di Cecco V, Catoni R, Varoni L, Di Santo M, Gratani L . 2013** . Seed germination capability of four endemic species in the Central Apennines (Italy): relationships with seed size. *LAZAROA* 34: 43–53.**Galbany-Casals M, Unwin M, Garcia-Jacas N, Smissen RD, Susanna A, Bayer RJ . 2014** . Phylogenetic relationships in Helichrysum (Compositae: Gnaphalieae) and related genera: Incongruence between nuclear and plastid phylogenies, biogeographic and morphological patterns, and implications for generic delimitation. *TAXON* 63: 608–624.**Gérard M, Vanderplanck M, Wood T, Michez D .2020** . Global warming and plant–pollinator mismatches.*Emerging Topics in Life Sciences* 4: 77–86.**Godfray HCJ, Knapp S, Mace GM . 2004** . The role of taxonomy in species conservation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359: 711–719.**Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O .2010** . New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syste-*

matic Biology 59: 307–321.**Hibbins MS, Gibson MJ, Hahn MW .2020** . Determining the probability of hemiplasy in the presence of incomplete lineage sorting and introgression (A Rokas and PJ Wittkopp, Eds.). *eLife* 9: e63753.**Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS . 2018** . UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* 35: 518–522.**Hokanson K, Hancock J . 1998** . Levels of allozymic diversity in diploid and tetraploid Vaccinium sect. Cyanococcus (blueberries). *Canadian Journal of Plant Science* 78: 327–332.**Huson DH, Bryant D . 2006** . Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254–267.**Huson DH, Scornavacca C .2012** . Dendroscope 3: An Interactive Tool for Rooted Phylogenetic Trees and Networks. *Systematic Biology* 61: 1061–1067.**Inouye DW . 2020** . Effects of climate change on alpine plants and their pollinators. *Annals of the New York Academy of Sciences* 1469: 26–37.**Ishihama F, Ueno S, Tsumura Y, Washitani I . 2005** . Gene flow and inbreeding depression inferred from fine-scale genetic structure in an endangered heterostylous perennial, Primula sieboldii. *Molecular Ecology* 14: 983–990.**IUCN Species Survival Commission . 2012** .*IUCN Red List Categories and Criteria: Version 3.1. Second edition.* IUCN, Gland, Switzerland.**Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC, Wickett NJ .2016** . HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* 4: 1600016.**Johnson MG, Pokorny L, Dodsworth S, Botigué LR, Cowan RS, Devault A, Eiserhardt WL, Epitawalage N, Forest F, Kim JT, Leebens-Mack JH, Leitch IJ, Maurin O, Soltis DE, Soltis PS, Wong GK shu, Baker WJ, Wickett NJ . 2019** . A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using k-Medoids Clustering. *Systematic Biology* 68: 594–606.**Junier T, Zdobnov EM . 2010** . The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. *Bioinformatics* 26: 1669–1670.**Katoh K, Standley DM . 2013** . MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30: 772–780.**Knaus BJ, Grünwald NJ .2017** . vcfr: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources* 17: 44–53.**Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A .2019** . RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35: 4453–4455.**Kress A . 1984** . *Primulaceen-Studien: Chromosomenzählungen an verschiedenen Primulaceen. Teil A, Androsace. 3* . A. und I. Kress.**Kress A . 1997** . *Androsace .Flora Iberica V* . Madrid: Real Jardín Botánico, CSIC., 22–40.**Kück P, Meusemann K . 2010** . FASconCAT: Convenient handling of data matrices. *Molecular Phylogenetics and Evolution* 56: 1115–1118.**Lanfear R, Calcott B, Ho SYW, Guindon S . 2012** . PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses.*Molecular Biology and Evolution* 29: 1695–1701.**Laguna E. 2000** . The Micro-Reserves as a Tool for Conservation of Threatened Plants in Europe. Nature and Environment Series nº 121. Council of Europe. ISBN: 92-871-4664-0**Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A . 2014** . Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology* 14: 82.**Larson DA, Chanderbali AS, Maurin O, Gonçalves DJP, Dick CW, Soltis DE, Soltis PS, Fritsch PW, Clarkson JJ, Grall A, Davies NMJ, Larridon I, Kikuchi IABS, Forest F, Baker WJ, Smith SA, Utteridge TMA .2023** . The phylogeny and global biogeography of Primulaceae based on high-throughput DNA sequence data. *Molecular Phylogenetics and Evolution* 182: 107702.**Li H, Durbin R .2009** . Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* 25: 1754–1760.**Lischer HEL, Excoffier L . 2012** . PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28: 298–299.**Liu TJ, Zhang SY, Wei L, Lin W, Yan HF, Hao G, Ge XJ . 2023** . Plastome evolution and phylogenomic insights into the evolution of Lysimachia (Primulaceae: Myrsinoideae). *BMC Plant Biology* 23: 359.**Liu Z, Zhao G . 1999** . Population genetics and its implications for conservation of rare and endangered plants.*Biodiversity Science* 07: 340–346.**Losapio G, Cerabolini BEL, Maffioletti C, Tampucci D, Gobbi M, Caccianiga M . 2021** . The Consequences of Glacier Retreat Are Uneven Between Plant Species.*Frontiers in Ecology and Evolution* 8.**Lynch M .1991** . The Genetic Interpretation of Inbreeding Depression and Outbreeding Depression. *Evolution* 45: 622–629.**McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA . 2010** . The Genome Analysis Toolkit: A MapReduce framework for

analyzing next-generation DNA sequencing data. *Genome Research* 20: 1297–1303.**McLay TGB, Birch JL, Gunn BF, Ning W, Tate JA, Nauheimer L, Joyce EM, Simpson L, Schmidt-Lebuhn AN, Baker WJ, Forest F, Jackson CJ . 2021** . New targets acquired: Improving locus recovery from the Angiosperms353 probe set. *Applications in Plant Sciences* 9.M**édail F, Diadema K, Pouget M, Baumel A. 2021** . Identification of plant micro-reserves using conservation units and population vulnerability: the case of an endangered endemic snowflake (acis nicaeensis) in the mediterranean basin hotspot. J Nat Conserv 61:125980. **Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R . 2020** . IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* 37: 1530–1534.**MITECO . 2011** .*Androsace cantabrica* . Available at https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/Androsace_cantabrica_tcm30-200325.pdf.**Mondoni A, Rossi G, Orsenigo S, Probert RJ . 2012** . Climate warming could shift the timing of seed germination in alpine plants.*Annals of Botany* 110: 155–164.**Moreno Saiz J, Algarra J, P. B, Carbajal R, Domínguez Lozano F, Marrero Gómez MV, Acedo C, Acevedo A, E A, Aldezabal A, Amich F, Aymerich P, Baudet A, G B, G B, Blanché C, Bueno A, Cabezudo B, Crespo M, Wildpret W . 2008** . *Lista Roja 2008 de la flora vascular española* .**Morozowska M, Czarna A, Kujawa M, Jagodzinski AM . 2011** . Seed morphology and endosperm structure of selected species of Primulaceae, Myrsinaceae, and Theophrastaceae and their systematic importance. *Plant Systematics and Evolution* 291: 159–172.**Moza MK, Bhatnagar AK .2007** . Plant reproductive biology studies crucial for conservation. *Current Science* 92: 1207.**Murphy B, Forest F, Barraclough T, Rosindell J, Bellot S, Cowan R, Golos M, Jebb M, Cheek M . 2020** . A phylogenomic analysis of Nepenthes (Nepenthaceae).*Molecular Phylogenetics and Evolution* 144: 106668.**Ottewell KM, Bickerton DC, Byrne M, Lowe AJ .2016** . Bridging the gap: a genetic assessment framework for population-level threatened plant conservation prioritization and decision-making. *Diversity and Distributions* 22: 174–188.**Ozenda P . 1995** . L'endémisme au niveau de l'ensemble du Système alpin. *Acta Botanica Gallica* 142: 753–762.**Palumbi SR . 1996** . The polymerase chain reaction. *Molecular systematics* : 205–247.**Pavan S, Delvento C, Ricciardi L, Lotti C, Ciani E, D'Agostino N. 2020.** Recommendations for Choosing the Genotyping Method and Best Practices for Quality Control in Crop Genome-Wide Association Studies.*Frontiers in Genetics* 11: 447.**Peakall R, Smouse PE .2012** . GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update.*Bioinformatics* 28: 2537–2539.**Phang A, Pezzini FF, Burslem DFRP, Khew GS, Middleton DJ, Ruhsam M, Wilkie P . 2023** . Target capture sequencing for phylogenomic and population studies in the Southeast Asian genus Palaquium (Sapotaceae). *Botanical Journal of the Linnean Society* : boad022.**Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Auwera GAV der, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, Shakir K, Thibault J, Chandran S, Whelan C, Lek M, Gabriel S, Daly MJ, Neale B, MacArthur DG, Banks E .2018** . Scaling accurate genetic variant discovery to tens of thousands of samples. : 201178.**Pritchard JK, Stephens M, Donnelly P . 2000** . Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155: 945–959.**de Queiroz K, Gauthier J . 1992** . Phylogenetic Taxonomy.*Annual Review of Ecology and Systematics* 23: 449–480.**Roy T, Cole LW, Chang TH, Lindqvist C . 2015** . Untangling reticulate evolutionary relationships among New World and Hawaiian mints (Stachydeae, Lamiaceae). *Molecular Phylogenetics and Evolution* 89: 46–62.**Rumpf SB, Gravey M, Brönnimann O, Luoto M, Cianfrani C, Mariethoz G, Guisan A . 2022** . From white to green: Snow cover loss and increased vegetation productivity in the European Alps.*Science* 376: 1119–1122.**Sam MS . 2013** . K-STRATEGY. Available at https://psychologydictionary.org/k-strategy/ [accessed 16 July 2023].**Schneeweiss GM, Schönswetter P, Kelso S, Niklfeld H . 2004** . Complex Biogeographic Patterns in*Androsace* (Primulaceae) and Related Genera: Evidence from Phylogenetic Analyses of Nuclear Internal Transcribed Spacer and Plastid trnL-F Sequences. *Systematic Biology* 53: 856–876.**Schoen DJ, Brown AHD. 2001.** The Conservation of Wild Plant Species in Seed Banks: Attention to both taxonomic coverage and population biology will improve the role of seed banks as conservation tools. *BioScience* 51: 960–966.**Schwager P, Berg C . 2019** . Global warming threatens conservation status of alpine EU habitat types in the European Eastern Alps. *Regional Environmental Change* 19: 2411–2421.**Seddon AWR, Macias-Fauria M, Long PR, Benz D, Willis KJ . 2016** . Sensitivity of global terrestrial ecosystems to climate variability. *Nature* 531: 229–232.**Slater GSC,**

Birney E . 2005 . Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6: 31.**Slimp M, Williams LD, Hale H, Johnson MG . 2021** . On the potential of Angiosperms353 for population genomic studies.*Applications in Plant Sciences* 9.**Smith G, Lowe D .1997** . *The Genus Androsace* . Avon Bank, Pershore: Alpine Garden Society.**Smith SA, Moore MJ, Brown JW, Yang Y .2015** . Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evolutionary Biology* 15: 150.**Smyčka J, Roquet C, Boleda M, Alberti A, Boyer F, Douzet R, Perrier C, Rome M, Valay JG, Denoeud F, Šemberová K, Zimmermann NE, Thuiller W, Wincker P, Alsos IG, Coissac E, Lavergne S . 2022** . Tempo and drivers of plant diversification in the European mountain system. *Nature Communications* 13: 2750.**Steinbauer MJ, Grytnes JA, Jurasinski G, Kulonen A, Lenoir J, Pauli H, Rixen C, Winkler M, Bardy-Durchhalter M, Barni E, Bjorkman AD, Breiner FT, Burg S, Czortek P, Dawes MA, Delimat A, Dullinger S, Erschbamer B, Felde VA, Fernández-Arberas O, Fossheim KF, Gómez-García D, Georges D, Grindrud ET, Haider S, Haugum SV, Henriksen H, Herreros MJ, Jaroszewicz B, Jaroszynska F, Kanka R, Kapfer J, Klanderud K, Kühn I, Lamprecht A, Matteodo M, di Cella UM, Normand S, Odland A, Olsen SL, Palacio S, Petey M, Piscová V, Sedlakova B, Steinbauer K, Stöckli V, Svenning JC, Teppa G, Theurillat JP, Vittoz P, Woodin SJ, Zimmermann NE, Wipf S . 2018** . Accelerated increase in plant species richness on mountain summits is linked to warming. *Nature* 556: 231–234.**Stöcklin J . 1992** . Umwelt, Morphologie und Wachstumsmuster klonaler Pflanzen : eine Ubersicht. *Botanica Helvetica* 102: 3–21.**Stubbs RL, Theodoridis S, Mora-Carrera E, Keller B, Yousefi N, Potente G, Leveille-Bourret E, Celep F, Kochjarova J, Tedoradze G, Eaton DAR, Conti E . 2023** . Whole-genome analyses disentangle reticulate evolution of primroses in a biodiversity hotspot. *New Phytologist* 237: 656–671.**Taylor SA, Larson EL . 2019** . Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nature Ecology & Evolution* 3: 170–177.**Tejero P, Otamendi M, Arrieta M, Etxeberria M, Escrig AA, Hermosilla-Lorenzo B, Navarro L, Martinez-Ortega M, Calvo-Yuste J, Malvar-Ferreras T, Ezquerra V, Ruiz A, Urquiola L, Martinez N, Asiain M, Aguinako I, Zabala J, Vicente-Ferrandez J, Palacio S, Viruel J, Pokorny L, Rincon M, Villaverde T, Garmendia-Altuna J . 2022** . Informe Cientifico-Tecnico del proyecto PRIOCONEX. DIGITAL.CSI. https://doi.org/10.20350/digitalCSIC/14688.**Viruel J, Forest F, Paun O, Chase MW, Devey D, Sousa Couto R, Segarra-Moragues JG, Catalan P, Wilkin P. 2018.** A nuclear *Xdh* phylogenetic analysis of yams (*Dioscorea* : Dioscoreaceae) congruent with plastid trees reveals a new Neotropical lineage. *Botanical Journal of the Linnean Society* 187: 232-246.**Viruel J, Conejero M, Hidalgo O, Pokorny L, Powell RF, Forest F, Kantar MB, Soto Gomez M, Graham SW, Gravendeel B, Wilkin P, Leitch IJ . 2019** . A Target Capture-Based Method to Estimate Ploidy From Herbarium Specimens. *Frontiers in Plant Science* 10.**Vriesendorp B, Bakker FT . 2005** . Reconstructing Patterns of Reticulate Evolution in Angiosperms: What Can We Do? *Taxon* 54: 593–604.**Weiss CL, Pais M, Cano LM, Kamoun S, Burbano HA . 2018** . nQuire: a statistical framework for ploidy estimation using next generation sequencing. *BMC Bioinformatics* 19: 122.**Whittaker RJ . 1993** . Plant Population Patterns in a Glacier Foreland Succession: Pioneer Herbs and Later-Colonizing Shrubs. *Ecography* 16: 117–136.**Winkler DE .2020** . Contemporary Human Impacts on Alpine Ecosystems: The Direct and Indirect Effects of Human-Induced Climate Change and Land Use. In: Goldstein MI, In: DellaSala DA, eds. *Encyclopedia of the World's Biomes* . Oxford: Elsevier, 574–580.**Wambugu PW, Nyamongo DO, Kirwa EC. 2023.** Role of Seed Banks in Supporting Ecosystem and Biodiversity Conservation and Restoration. *Diversity* 15:896. https://doi.org/10.3390/d15080896**Xiong Y, Zhao Y, He Y, Zhao L, Hu H, Barrett RL, Chen Z, Lu L . 2024** . Current progress and future prospects for understanding genetic diversity of seed plants in China. *Biological Diversity* .**Yi X, Latch EK .2022** . Nonrandom missing data can bias Principal Component Analysis inference of population genetic structure. *Molecular Ecology Resources* 22: 602–611.**Yuste JC, Arrieta M, Maddi Otamendi, Agusti Agut, Brais Hermosilla, M. Montserrat Martinez-Ortega, Navarro L, Palacio S, Viruel J, Naroa Martinez, Zulaika J, Villagrasa E, Imano Aguinaco, Asiain M, Ferrandez JV, Etxeberria M, Ezquerra V, Garmendia J, Ibarra PT . 2021** . First results from PRIOCONEX project: Converging climate change and *ex situ* conservation.**Zhang C, Rabiee M, Sayyari E, Mirarab S .2018** . ASTRAL-III: polynomial time

species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19: 153.

FIGURE CAPTIONS

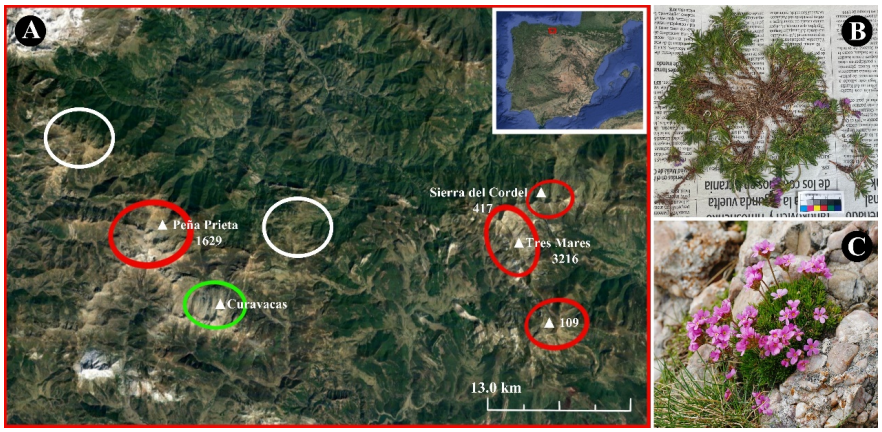

Figure 1. (A) Map of currently known *A. cantabrica* locations. Red circles indicate population sites with census estimates from Baudet *et al.* (2004). Baudet *et al.* (2004) did not detect *A. cantabrica* in the Curavacas region, marked by the green circle. White circles show marginal populations described more recently. (B) *A. cantabrica* specimen close-up. (C) *A. cantabrica* in its natural habitat.

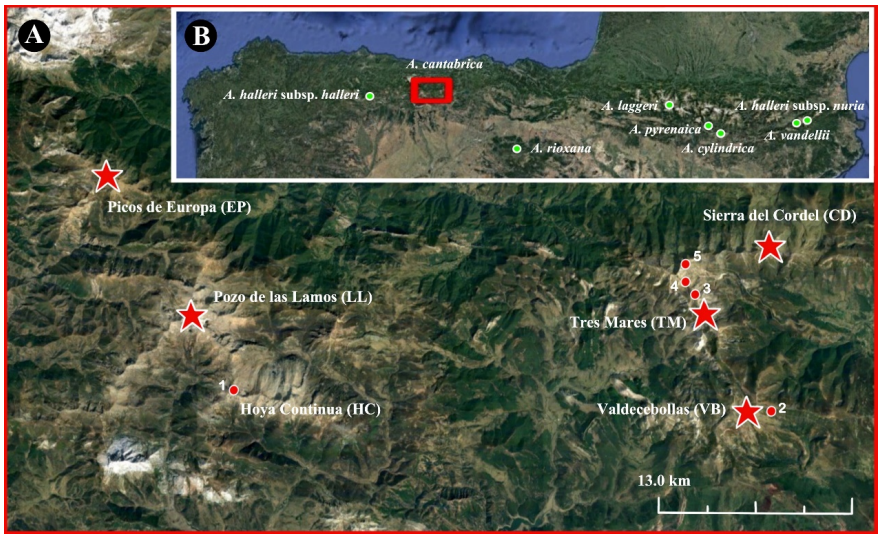

Figure 2. Sampling sites of *A. cantabrica* and related taxa. (A) Six sampling locations for *A. cantabrica* . Pentagrams represent six individuals of a subpopulation sampled, and dots represent a single individual sampled. (B) Sampling sites for related taxa of*A. cantabrica* in the northern Iberian Peninsula.

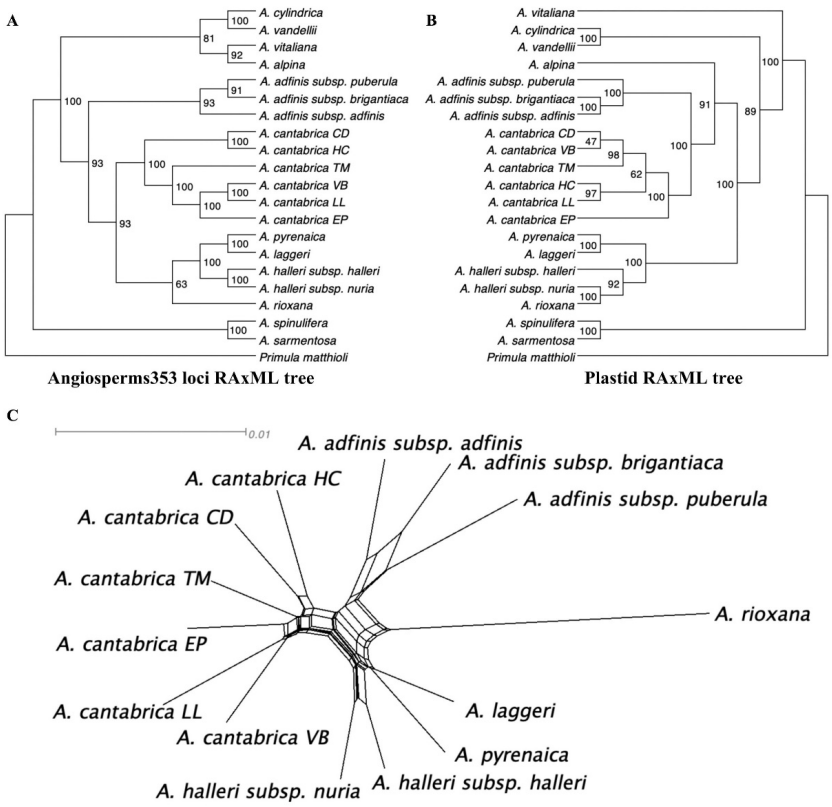

Figure 3. Phylogenetic tree and network split plots. (A) Angiosperms353 loci RAxML tree with node values indicating support from 1000 bootstrap replicates. (B) RAxML tree of 125 plastid fragments derived from Angiosperms353 off-target data. *A. adfinis* and*A. alpina* plastid sequences are derived from the online genome skimming data. (C) Angiosperms353 loci phylogenetic network split-plot.

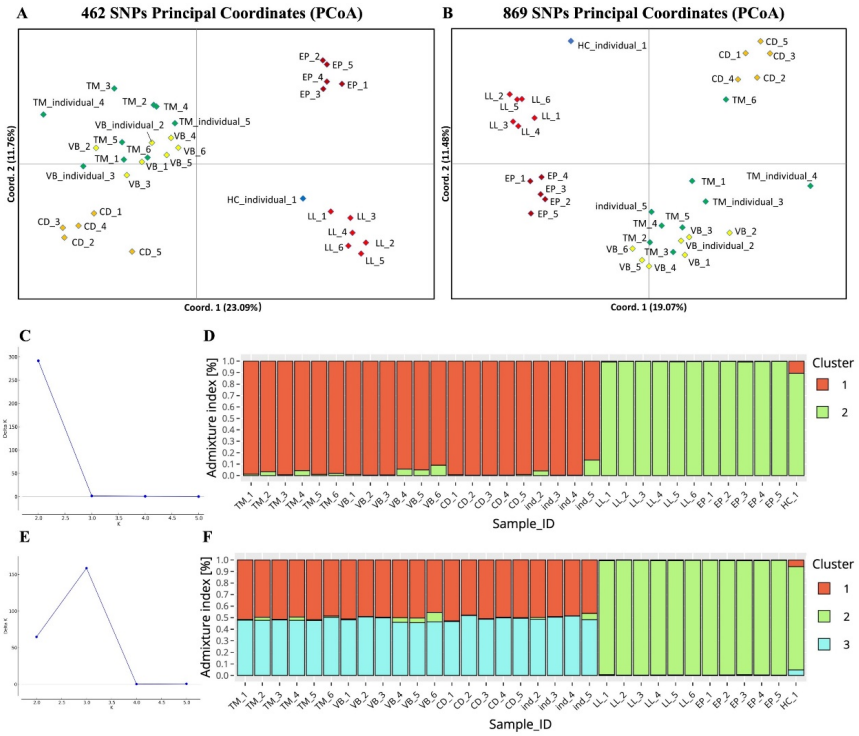

Figure 4. Results of population structure analyses using 462 SNPs (ACD) and 869 SNPs (BEF), respectively. (AB) PCoA plots, generated in GeneAlEx. (CE) DeltaK plots, generated in Structure Harvester from the STRUCTURE outputs, with the corresponding K values at their peaks on the plots representing the optimal cluster status for population structure (Evanno *et al.* , 2005). (DF) Population structure plots of all individuals are generated in StructuRly.

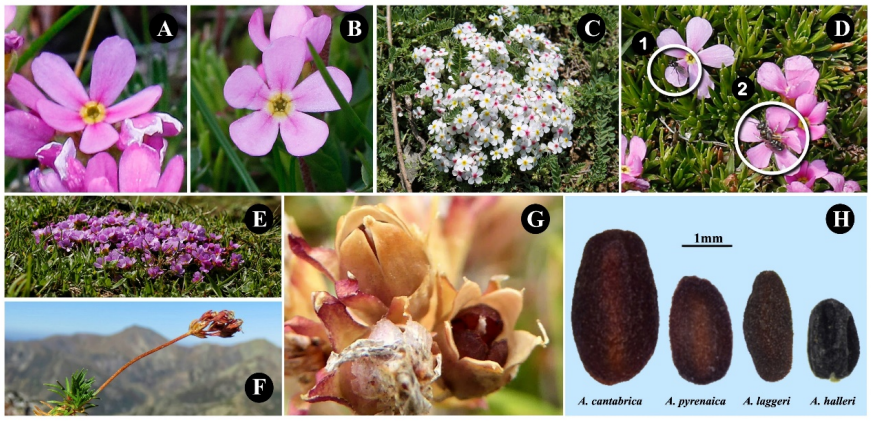

Figure 5 . Information on the reproductive biology of *A. cantabrica* . (A) *A. cantabrica* 's corolla characteristics at the beginning of flowering. (B) *A. cantabrica* 's corolla characteristics after pollination. (C) *A. ovczinnikovii* Schischk. &Bobrov, an example of the corolla throat turning from yellow to red after pollination in the genus *Androsace* . (D) *A. cantabrica* 's primary pollinators: 1. Diptera insects 2. Hymenoptera insects. (E) *A. cantabrica* grows in low, weedy meadows and has relatively short peduncles. (F) *A. cantabrica* growing in taller weeds or shrubs has longer peduncles, which could exceed 10cm. (G)*A.*

cantabrica 's dehiscent capsule contains 3-4 seeds. (H) Seed size comparison of *A. cantabrica* , *A. pyrenaica* ,*A. laggeri* and *A. halleri* .

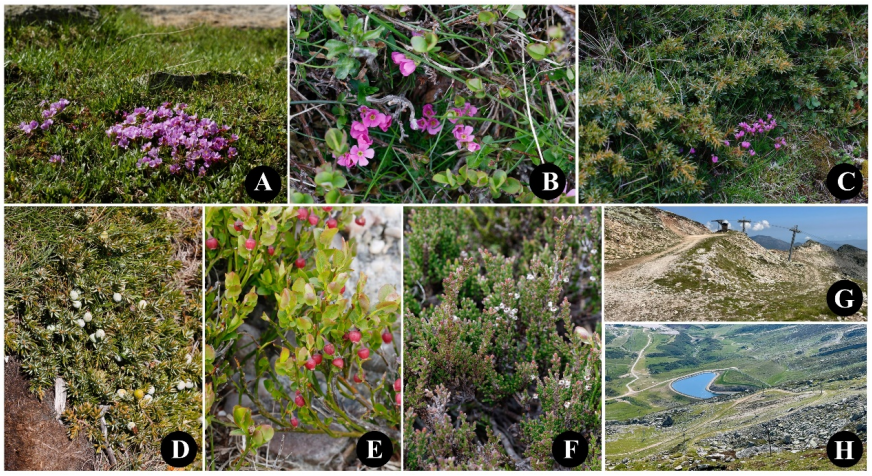

Figure 6. Habitat information for *A. cantabrica* . Three types of habitat: i) open alpine meadows around rocks (A), ii) within shorter shrubs (B), and iii) on the edge of taller and denser shrubs (C). Three main shrub species dominate the landscape: i) *Vaccinium uliginosum* L. (D), ii) *Juniperus communis* var. *saxatilis* Pall. (E) and *Calluna vulgaris* (L.) Hull (F). Anthropogenic disturbance in the Tres Meras (TM) area: construction of ski resorts and hiking trails (G-H).