



Genetic survey extension of the threatened Iberian *Arnica montana* L. revealed the presence of divergent plastid lineages and highly structured populations in northern Spain

Carmen Bouza · Iria Lorenzo · Manuel A. Rodríguez-Gutián ·
Adrián Casanova · Carlos Real · Rosa Romero · Manuel Vera

Received: 8 September 2022 / Accepted: 22 December 2022 / Published online: 3 January 2023
© The Author(s) 2023

Abstract Iberian populations of *Arnica montana* L. (Asteraceae) represent a valuable resource both for conservation and pharmaceutical industry. Previous genetic analyses pointed out the presence of different genetic groups, but a wide region is still genetically unexplored. In order to fill this scientific gap, the present study analysed a wider sampling area along the northern Iberian Peninsula. Nuclear (i.e. microsatellite loci) and plastid DNA (cpDNA) molecular markers were used to assess the levels of genetic diversity and the population structure in 16 locations, eight analysed for the first time in the present study and eight representative of the different genetic groups previously identified. The two divergent cpDNA

groups previously described were found, but their distribution was extended and refined. Thus, one of the groups (suggested as ancestral) was predominantly distributed in adjacent zones of the Cantabrian coasts while the other (more related to Central-European populations) was predominant in inner Cantabrian regions and Pyrenees. Genetic diversity with microsatellite loci ($H_e=0.280$) was in accordance with the figures previously described, with a high level of population differentiation ($F_{ST}>0.500$) identifying the presence of up to five population genetic units. Genetic and geographical distances were not related (no isolation-by-distance pattern identified), suggesting an important effect of genetic drift. Finally, due to the conservation and evolutionary interest of the populations analysed, different management actions useful for the maintenance of wild *A. montana* resources are provided.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10722-022-01527-y>.

C. Bouza · I. Lorenzo · A. Casanova · M. Vera (✉)
Departamento de Zoología, Genética y Antropología Física. Facultad de Veterinaria, Universidad de Santiago de Compostela. Campus de Lugo, 27002 Lugo, Spain
e-mail: manuel.vera@usc.es

M. A. Rodríguez-Gutián · R. Romero
Departamento de Producción Vegetal, Escuela Politécnica Superior de Ingeniería, Universidad de Santiago de Compostela. Campus de Lugo, 27002 Lugo, Spain

C. Real
Departamento de Biología Funcional, Escuela Politécnica Superior de Ingeniería, Universidad de Santiago de Compostela. Campus de Lugo, 27002 Lugo, Spain

Keywords *Arnica montana* · Conservation · Genetic diversity · Management · Molecular markers · Population structure

Introduction

Information about the levels of genetic diversity and its landscape distribution is a keystone for the development of strategies that allow the conservation and management of these genetic resources (Frankham et al. 2002). This is particularly important for plant

species that are sessile with a limited biological dispersion range (Beckman et al. 2018), being more prone to the effect of adverse biotic and abiotic stresses (Atkinson and Urwin 2012; Leitao et al. 2020). Plant species are highly affected by human-mediated disturbances such as habitat fragmentation, changes in the land use, overexploitation, pollution, wildfires and the ongoing climate changes (Hautekeete et al. 2015). These disturbances can reduce the population sizes and the connectivity among populations, inducing undesirable effects of genetic drift and inbreeding (Ouborg et al. 2006). In addition to these recent processes, historical ones (e.g. climatic oscillations that occurred in the Pleistocene) also have an important contribution to the distribution of the genetic diversity (Birks 2019; Maroso et al. 2021).

Arnica montana L. (Asteraceae) is a perennial herb that is widely distributed across Europe (from Portugal to Scandinavia and the Baltic states; Ferguson 1976). This highly self-incompatible species (Luijten et al. 1996) undergoes sexual (mediated by different insects; see Kahmen and Poschlod 2000) and vegetative reproduction, with a flowering period in late spring and early summer in the Northern Hemisphere. There are no persistent seed banks, with germination taking place in autumn or in the next spring after seed dispersal (Luijten et al. 2000). *Arnica montana* is in high demand in the pharmaceutical industry due to the medicinal property of its secondary metabolites (basically sesquiterpene lactones) which are concentrated mainly in the flowers (Kriplani et al. 2017). Together with their anti-inflammatory action (Siedle et al. 2004), sesquiterpene lactones have many interesting activities that protect the plant against dangerous environmental stressors (e.g. UV radiation and extreme temperature) and herbivory (Zidorn 2010). Floral head recollection is made by hand, with Spain and Romania the main suppliers for the pharmaceutical industry (Mardari et al. 2019; Vera et al. 2020). This practice, carried out basically with no legislation, has diminished wild populations in different regions of Europe. Many European countries have included *A. montana* in their red list, developing specific laws for its protection (Allen et al. 2014). The species is globally catalogued by the International Union for Conservation of Nature (IUCN) as “Least Concern”; the IUCN also points out the decrease in wild populations in different countries and the recommends monitoring their resources (Falniowski

et al. 2013). Main threats to the species are related to changes in the land use, which reduce the size of its habitat, and uncontrolled exploitation (Falniowski et al. 2013; Moreno et al. 2019). In Spain, *A. montana* is listed in the Atlas of the Threatened Vascular Flora (Moreno et al. 2019), and it is also catalogued by the European Union in the Annex V of the CD 92/43/EEC, as a “plant species of community interest”, with suggestions to manage the exploitation of its wild resources.

Initial morphological analyses carried out in the 1940s identified the presence of two subspecies, *A. montana* subsp. *montana* and *A. montana* subsp. *atlantica* (Bolos y Vayreda 1946; see also Villar et al. 2019). The most common subspecies, which is widely distributed in European regions with continental climate and an altitude range of 0–3000 m a.s.l., is *A. montana* subsp. *montana*. On the other hand, *A. montana* subsp. *atlantica* is mainly distributed in Northern parts of the Iberian Peninsula and Southwestern France under oceanic climate conditions and an altitude of 0–400 m a.s.l. (maximum 1000 m) (Vera et al. 2014). However, overlapping biometric values used to differentiate both subspecies make it difficult to allocate morphologically the specimens of each subspecies in regions where both species cohabit, such as northern Spain (Romero et al. 2011).

Biochemical analyses of some Spanish populations from Galicia (NW Spain) identified two different biochemical groups (i.e. chemotypes) based on the abundance of different types of sesquiterpene lactones (SLs) in the flower extracts (Perry et al. 2009). Helenalin esters (H) are predominant in the Central-European chemotype and are associated with *A. montana* subsp. *montana*, while dihidrohelenalin esters (DH) are predominant in the Iberian chemotype found in Galicia and are associated with *A. montana* subsp. *atlantica* (Perry et al. 2009; Schmiderer et al. 2018). These SLs are the main metabolites that confer anti-inflammatory properties to the *A. montana* (Siedle et al. 2004). The anti-inflammatory action of H esters is greater than that of DH esters, but DH esters are less allergenic and, therefore, more valued in the pharmaceutical industry (Lass et al. 2008; Vera et al. 2014).

Despite its economic benefits, the number of genetic studies focused on the species are still quite scarce, and most of these scarce studies concentrate on the Central-Western European regions. These

studies have identified low-to-moderate genetic diversity and important degrees of differentiation among populations under low levels of gene flow (Luijten et al. 2000; Duwe et al. 2017; Van Rossum and Raspé 2018) and selective processes associated with altitude (Maurice et al. 2016). Genetic studies in the Iberian Peninsula have been focused on Galician populations (NW Spain) because these populations are the most conserved and variable (Vera et al. 2020). Using plastid DNA (cpDNA) sequencing, two different genetic groups concordant with the two previous chemotypes identified in the region have been described (cpDNA-group 1 and cpDNA-group 2 related to DH and H groups, respectively; see Vera et al. 2014). Despite morphological limitations for subspecies identification in Galicia, these two plastid groups have also been suggested, as divergent lineages, to be possibly related to *A. montana* subsp. *atlantica* (cpDNA-group 1) and *A. montana* subsp. *montana* (cpDNA-group 2) (see Vera et al. 2020). Schmiderer et al. (2018) compared Galician populations, composed of *A. montana* subsp. *atlantica* specimens, with Central-European populations composed of *A. montana* subsp. *montana*. These authors found a high genetic differentiation between Galician and European populations, which was justified by the presence of the two subspecies. Unfortunately, this study did not include all the diversity present in Galicia, analysing just a restricted area where only cpDNA-group 1 was present. The most exhaustive study carried out to date (~300 individuals distributed in 27 Galician locations) using cpDNA sequencing and microsatellite loci, identified the two cpDNA groups previously described (with cpDNA-group 2 restricted to a small area in the O Courel Mountains), subdivided into four highly differentiated population genetic units ($F_{ST}=0.441$) due to genetic drift, although local adaptations were also suggested (Vera et al. 2020). However, this study did not include other Northern Iberian regions, such as the Cantabrian Mountains and Pyrenees, where the species has been noticed, which would have improved knowledge on the distribution of the different cpDNA lineages (two or more) and the different demographic and/or evolutionary events driving it (i.e. glacial and interglacial periods during the Pleistocene which led to northern and/or southern expansions of animals and plants, see Hewitt 2000). Hence, a

wider sampling scenario across the North of the Iberian Peninsula is necessary to understand the distribution of the genetic groups identified.

Therefore, the goal of the present research is to analyse the genetic diversity levels and population structure of *A. montana* across the North of the Iberian Peninsula, including the Cantabrian Mountains and the Pyrenees. A battery of different molecular markers has been used (i.e. cpDNA sequencing and microsatellite genotyping) to compare the results obtained with those previously found in Galicia to understand the current extension of the different population units previously identified by Vera et al. (2020) and their possible relationship with European populations. This wider scenario gives a more informative picture of the species in its southern natural range which allows the proposal of more refined management and conservation actions for the species in the studied area.

Materials and methods

Locations

One hundred and four specimens of *A. montana* were gathered from eight locations (maximum linear distance between locations=752 km), covering different habitats (i.e. heathlands, peatlands and haymeadows) and at different altitudes, across the Northern Iberian Peninsula in June–July (i.e. flowering period) of 2021 (Fig. 1, Table 1). All these individuals were stored at the LUGO Herbarium, University of Santiago de Compostela. All sampled individuals belonged to distant plants (above 1 m between plants) within each location to avoid clonality related to the vegetative reproduction of the species (Kahmen and Poschlod 2000). The sampling size (N) in each location depended on the abundance of flowering plants, legislation imposed by autonomous administrations involved in the management of the species (due to its inclusion by the European Union in the Annex V of the CD 92/43/EEC), and, in some cases, availability after plant collection carried out by professional workers. Hence, all fieldwork had the proper administrative authorization. Moreover, eight locations from Galicia (NW Iberian Peninsula, 90 individuals; see Table 1), representing the four main genetic groups (two locations for each group) identified by Vera et al.

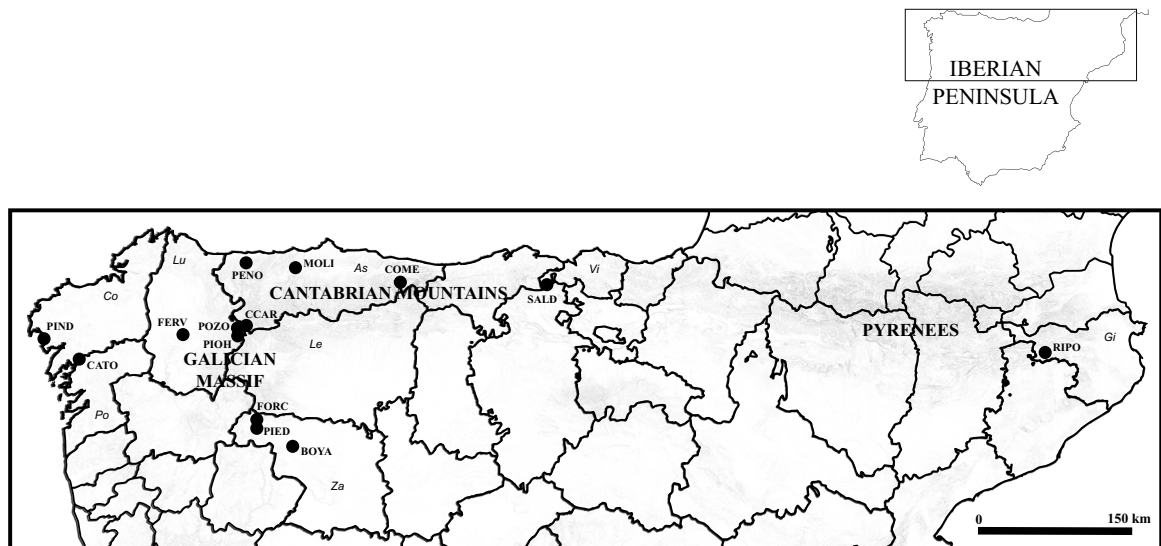


Fig. 1 Geographic distribution within northern Iberian Peninsula of the 16 *Arnica montana* studied locations. Elevation is represented in grey gradient from white (representing sea level) to dark grey (for higher elevations). Main mountain systems and regions (As, Asturias; Za, Zamora province; Vi,

Vitoria province; Gi, Girona province; Le, León province; Lu, Lugo province; Co, A Coruña province; Po, Pontevedra province) are indicated in capital grey and italic letters, respectively. Location codes are shown on Table 1

Table 1 Locations of *Arnica montana* analysed in the present study. Location (LOC), code (COD), number of individuals analysed (N), UTM coordinates, altitude (ALT) and habitat (HAB) are shown

LOC	COD	N	UTMX	UTMY	ALT	HAB
As: National Park Picos de Europa	COME	15	338601	4794446	840	Peatlands
As: La Molina	MOLI	9	716463	4806654	655	Peatlands
As: Penouta	PENO	4	675861	4913655	870	Heathlands
Za: Boya	BOYA	14	719797	4644983	865	Heathlands
Za: Piedras Negras	PIED	15	687429	4664842	1045	Peatlands
Za: Forcadura	FORC	17	687782	4671651	1695	Heathlands
Vi: Salduero	SALD	15	468592	4777308	1002	Peatlands
Gi: Ripollés	RIPO	15	433155	4672370	700	Heathlands
Co: Dumbría, Monte Pindo*	PIND	16	491000	4748000	535	Peatlands
Po: Catoira, Oeste, Brañas de Tixó*	CATO	4	522000	4725000	2	Peatlands
Lu: O Corgo, A Fervenza*	FERV	11	620000	4749000	352	Hay-meadows
Lu: Cervantes, Donís, O Piorno*	PIOH	10	674000	4747000	1110	Hay-meadows
Le: Candín, Balouta, Catro Caballos*	CCAR	10	678000	4752000	1410	Heathlands
Lu: Navia de Suarna, Rao, Marco do Pozo*	POZO	14	677000	4750000	1495	Heathlands
Lu: Folgoso do Courel, Alto do Couto*	COUT	10	656000	4719000	1330	Heathlands
Lu: Folgoso do Courel, Moreda, Pico Formigueiros*	PFOR	16	655000	4718000	1635	Heathlands

*locations previously analysed in Vera et al. (2020)

As, Asturias; Za, Zamora province; Vi, Vitoria province; Gi, Girona province; Le, León province; Lu, Lugo province; Co, A Coruña province; Po, Pontevedra province.

(2020), were included into the analyses to compare both genetic diversity levels and population differentiation in order to understand the proper distribution of these different genetic groups along the North of the Iberian Peninsula. Hence, the total number of individuals included was 194 (maximum linear distance between locations = 934 km).

Molecular analyses

Genomic DNA extraction was individually carried out with the E.Z.N.A. Plant DNA DS Mini Kit (OMEGA bio-tek) following the manufacturer's instructions using dried leaf (~40 mg) as target tissue.

cpDNA sequencing of two non-coding genic markers (i.e. *rps16* intron and *ycf4-cemA* spacer), in five individuals per location selected by their different microsatellite multilocus genotype or in all the individuals available within the location when $N < 5$, were obtained following the methodology of Vera et al. (2014).

Three different PCR reactions developed by Vera et al. (2020) were used to amplify fourteen microsatellite loci previously described (Duwe et al. 2015) as follows: Multiplex 1 composed of Arm01, Arm03, Arm05, Arm06 and Arm07; Multiplex 2 composed of Arm04, Arm08, Arm09, Arm11 and Armo03; and finally Multiplex 3 composed of Arm02, Arm10, Armo01 and Armo02. Amplified products were examined and genotyped using an ABI PRISM® 3730xl sequencer (Applied Biosystems) and the software GeneMapper 6.0 (Applied Biosystems).

Genetic diversity and population structure in microsatellites

The software MICROCHECKER v.2.3.3 (Van Oosterhout et al. 2004) was used to check genotyping errors such as null alleles, stuttering and allele dropout in the microsatellites analysed. Estimates of genetic diversity values within locations (i.e. average number of alleles per locus (N_a), allelic richness (A_r), observed (H_o) and expected heterozygosity (H_e)) were calculated with FSTAT 2.9.3 (Goudet 2001). Conformance to Hardy–Weinberg (HW) expectations per locus and location was tested using GENEPOP 4.0 (Rousset 2008). The software Genalex 6.5 (Peakall and Smouse 2012) was used to calculate the number of multilocus genotypes (MLG) present in

each location, which was necessary to estimate the proportion of clones in each location as $1\text{-MLG}/N$, where N was the sampling size. Estimates of effective population size (N_e) were computed according to the linkage disequilibrium (LD) method available in the program Ne Estimator 2.01 (Do et al. 2014).

Coefficients of population differentiation (global and pairwise F_{ST} values) among the locations were calculated using FSTAT and their significance was tested with 10,000 permutations. The number of genetically homogenous population units (K) present in the dataset was inferred by the Bayesian clustering method used by STRUCTURE v. 2.3.4 (Falush et al. 2007) for $K=1\text{--}17$ (number of locations + 1) with five replicates for each K tested to check reproducibility. Two hundred thousand Monte Carlo replicates after a burn-in period of 50,000 steps was applied in these runs. The online tool STRUCTURE HARVESTER v. 0.6.94 (Earl and vonHoldt 2012) was used to detect the most likely K following the methodology described by Evanno et al. (2005). The software POPHELPER (Francis 2017) was used to summarize the different outputs and to create plot representations.

A discriminant analysis of principal components (DAPC) was also carried out using the software ADEGENET 2.0.0 (Jombart and Ahmed 2011) on the R platform (R Core Team 2018). ADEGENET transforms genotype using a Principal Component Analysis (PCA). The number of Principal Components (PC) and discriminant functions (DF) retained in the analysis allowed explaining variance percentage above 90%. Analysis of molecular variance (AMOVA) was performed to assess the distribution of genetic variation within (F_{SC}) and among (F_{CT}) location groups using the software ARLEQUIN v. 3.5.1.3 (Excoffier et al. 2005), with locations grouped following the population units suggested by STRUCTURE. The significance was evaluated with 10,000 permutations. Finally, Isolation-By-Distance (IBD) among locations was assessed, comparing the genetic distances (calculated as $F_{ST}/(1-F_{ST})$, see Rousset 1997) and geographic distances using a Mantel test implemented in the NTSYS software (Rohlf 1993) with 10,000 permutations. The distance matrix tool implemented in the software QGIS 3.10 (QGIS 2017; <http://qgis.osgeo.org>) was used to calculate the linear distance between pairs of locations taken as their geographic distances.

Results

cpDNA variability

The same two haplotypes previously identified in Galicia were detected for each cpDNA marker (Hap1 and Hap2 for *rps16* intron; and Hap1 and Hap2 for *ycf4-cemA* spacer; GenBank Accession Numbers: KF679679, KF679680, KF679681 and KF679682, respectively; see Vera et al. 2014). When the two markers were combined, the three expected combined haplotypes (cHap) were identified (i.e. cHap1, cHap2 and cHap3; Vera et al. 2014, 2020), where cHap1 was associated with the cpDNA-group 1 and cHap2 and cHap3 were associated with cpDNA-group 2. Basically, locations closer to the Cantabrian coast (COME, MOLI, PENO, PIED and SALD) belonged to the cpDNA-group 1 (as detected in six Galician locations studied), while inner and Pyrenean locations (BOYA, FORC and RIPO) belonged to the cpDNA-group 2

(as found in the other two Galician samples, COUT and PFOR) (Table 2, Fig. 1).

Genetic diversity within locations using microsatellite loci

The microsatellite Arm01, Arm03 and Arm04 showed a high percentage of missing genotypes (>25%) and were consequently removed from the subsequent studies. Thus, eleven microsatellites were retained (>80% of individuals genotyped), which showed no evidence of scoring error due to null alleles, stuttering and/or large allelic dropout except for Arm05, Arm06, Arm08 (in RIPO) and Arm09 (in FORC) where presence of null alleles were suggested. However, because no bias was expected for global genetic estimators and null alleles affected only very few specific locations, the 11 loci were finally used in the population analyses. Clonality in the locations was low (Table 2) except for PENO where the four sampled individuals shared the same multilocus

Table 2 Genetic variation estimates in the 16 locations included. For cpDNA, number of individuals analysed (Ncp), combined haplotype for the two plastid markers following Vera et al. (2014) and cpDNA group are shown. For microsatellite loci, number of individuals analysed (N), number of different multilocus genotypes (MLG), proportion of clones (clones),

mean observed (H_o) and expected (H_e) heterozygosity, mean number of alleles per locus (N_a), mean allelic richness (Ar), conformance to Hardy–Weinberg Equilibrium (P_{HWE}) and effective population size (N_e , 95% Confidence Intervals between brackets) are shown

cpDNA				Microsatellite loci								
Location	Ncp	Combined (c)	cpDNA group	N	MLG	clones	H_o	H_e	N_a	Ar	P_{HWE}	N_e
COME	5	cHap1	1	15	10	0.333	0.227	0.240	2.4	1.8	0.005	3.2 [1.6–18.1]
MOLI	5	cHap1	1	9	7	0.222	0.367	0.300	2.5	2.0	0.593	2.2 [1.0–14.2]
PENO	4	cHap1	1	4	1	0.750	0.364	0.182	1.4	1.4	0.320	∞ [∞ – ∞]
BOYA	5	cHap2	2	14	14	0.000	0.299	0.308	2.3	1.9	0.953	10.8 [2.0– ∞]
PIED	5	cHap1	1	15	10	0.333	0.085	0.128	1.7	1.4	0.016	0.5 [0.1–2.0]
FORC	5	cHap2	2	17	17	0.000	0.239	0.281	2.3	1.9	0.032	49.6 [3.1– ∞]
SALD	5	cHap1	1	15	14	0.067	0.278	0.268	2.2	1.9	0.127	49.4 [3.1– ∞]
RIPO	5	cHap3	2	15	15	0.000	0.444	0.535	3.5	2.7	0.036	44.3 [10.0– ∞]
PIND*	3	cHap1	1	16	16	0.000	0.159	0.209	1.9	1.5	0.271	318.9 [2.7– ∞]
CATO*	3	cHap1	1	4	3	0.250	0.273	0.356	2.0	1.9	0.535	∞ [0.4– ∞]
FERV*	3	cHap1	1	11	9	0.182	0.268	0.233	1.7	1.5	0.245	∞ [1.2– ∞]
PIOH*	3	cHap1	1	10	10	0.000	0.277	0.259	2.0	1.7	0.741	17.1 [1.2– ∞]
CCAR*	3	cHap1	1	10	10	0.000	0.219	0.247	2.1	1.6	0.355	1.1 [0.5–2.5]
POZO*	3	cHap1	1	14	14	0.000	0.173	0.191	1.6	1.4	0.967	∞ [1.7– ∞]
COUT*	8	cHap2	2	10	10	0.000	0.339	0.409	2.5	2.0	0.023	∞ [6.2– ∞]
PFOR*	7	cHap2 (4), cHap3 (3)	2	15	15	0.000	0.286	0.300	2.3	1.8	0.192	∞ [∞ – ∞]

*data from Vera et al. (2020)

genotype. Therefore, due to the low number of individuals in some locations and their representation of the current populations in the different locations, all individuals were retained for the analyses that followed.

The mean number of alleles (N_a) per location ranged from 1.4 in PENO to 3.5 in RIPO (mean across locations \pm standard deviation = 2.3 ± 0.6), the mean allelic richness (A_r) from 1.4 in PENO to 2.7 in RIPO (1.9 ± 0.4), the mean observed heterozygosity (H_o) from 0.085 in PIED to 0.444 in RIPO (0.288 ± 0.109), and the mean expected heterozygosity (H_e) from 0.128 in PIED to 0.535 in RIPO (0.280 ± 0.120) (Table 2). All diversity values for the Galician locations analysed by Vera et al. (2020) were included in the ranges above described. Four locations (COME, PIED, FORC and RIPO) did not conform to Hardy–Weinberg (HW) expectations ($P < 0.05$) and the test for PENO should be taken with caution because maybe all individuals were probably clones (and the HWE test would not make sense). All the populations conformed to HW expectations after sequential Bonferroni correction ($P = 0.0045$). Effective population size (N_e) estimates were basically low, with some locations having values close to 50 (Table 2). In PENO, an N_e infinite value was obtained, which was generated by a negative estimator, precluding the LD-based estimation of N_e in this location.

Population differentiation and structure patterns

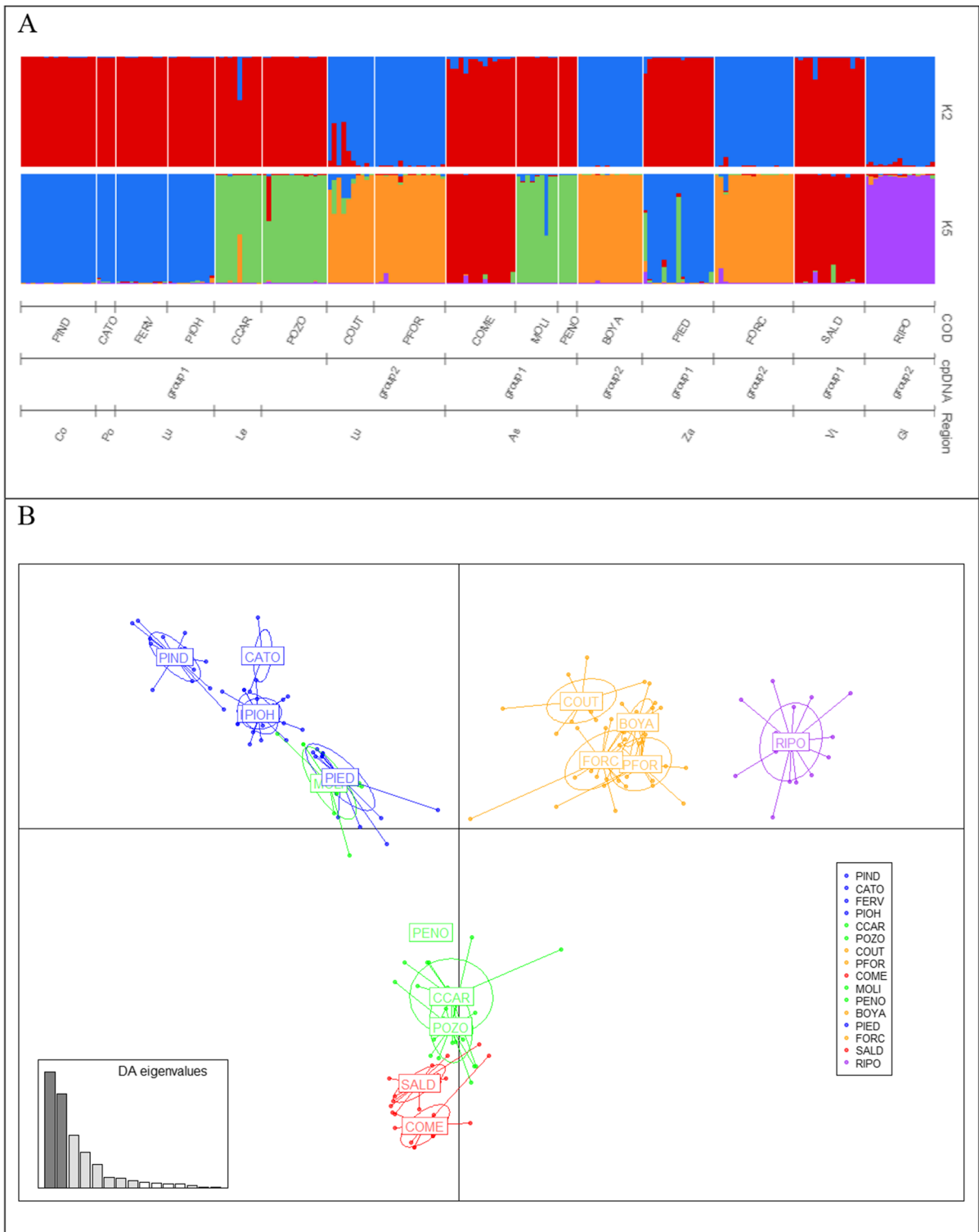
F_{ST} values between pairs of locations ranged from 0.126 to 0.757 (from 0.079 to 0.757 when Galician locations were included; see Supplementary Information 1). Global F_{ST} values were significant both for the new locations of the present study ($F_{ST} = 0.590$, $P < 0.001$) and when all previous and new populations were included ($F_{ST} = 0.545$, $P < 0.001$). The most likely K value while applying the ΔK approach proposed by Evanno et al. (2005) was 2, with also a peak for $K = 5$ in the ΔK plot given by STRUCTURE HARVESTER (Supplementary Information 2). The two groups for $K = 2$ were concordant with the results obtained with cpDNA (Fig. 2A). For $K = 5$ (Fig. 2A), RIPO created an own group (purple group), COME and SALD clustered together (red group), MOLLI and PENO were included in the group composed by CCAR and POZO (green group), and BOYA and

FORC were clustered with COUT and PFOR (orange group). Finally, PIED was clustered with PIND, CATO, FERV and PIOH (blue group, Fig. 2A). The DAPC plot confirmed basically the STRUCTURE results for $K = 5$, although MOLLI was clustered with the STRUCTURE blue group instead of the green one (Fig. 2B). AMOVA results with the five suggested groups by STRUCTURE assigned 51.72% of the global genetic variation to differences among groups ($F_{CT} = 0.517$, $P < 0.001$) and just 11.52% to variance among locations within group ($F_{SC} = 0.239$, $P < 0.001$), suggesting a good grouping. Finally, no IBD pattern was found when using neither the new eight locations of the present study ($r = -0.40381$, $P = 0.9249$) nor using the complete dataset (16 locations; $r = -0.11946$; $P = 0.8068$), suggesting a relevant effect of the genetic drift on the population genetic structure observed.

Discussion

Values of genetic diversity across the studied Northern Iberian region

Historical and recent processes shape the genetic variation and population differentiation present in natural populations (Maroso et al. 2021). Maintenance of genetic diversity can be a challenge for populations located at the limits of the species range because isolation and extinction/colonization events are frequent (Cires et al. 2011), but these populations can retain singular variants of high interest for conservation (Travis and Dytham 2004; Shay et al. 2021) and their human use should be rational and controlled. *Arnica montana* populations in the Iberian Peninsula represent the southern edge of the natural distribution of the species (Falniowski et al. 2013). Microsatellite genetic diversity values for the northern Iberian locations analysed for the first time in the present study (mean $H_e = 0.280$) were similar to those described in Galicia ($H_e = 0.249$, Schmiderer et al. 2018; $H_e = 0.245$, Vera et al. 2020) and lower than the figures found in European regions such as Germany and the Alps (mean $H_e = 0.512$; Duwe et al. 2017), Austria ($H_e = 0.657$, Schmiderer et al. 2018) and Belgium ($H_e = 0.463$, Van Rossum and Raspé 2018). As suggested by Vera et al. (2020), these results would agree with the “abundant centre” model proposed by Brown



◀**Fig. 2** STRUCTURE (A) and DAPC (B) plots obtained in the present study. For STRUCTURE analyses, $K=2$ and $K=5$ values are represented. Each vertical bar represents one individual, and the colour proportion for each bar represents the posterior probability of assignment of each individual to the different clusters (K). cpDNA group and region for each location are also included. For DAPC plot, the first and second discriminant factors (DF) are shown, including the weight of retained discriminant analysis (DA) eigenvalues on the left bottom box. Location codes are shown on Table 1

(1984) where there is a loss of abundance from the centre of species distribution to the limits of species distribution. This pattern has been described in other plant species (Cires et al. 2011; 2013) including Asteraceae members (Jump et al. 2003; McMinn et al. 2017).

Population differentiation and extension/distribution of Iberian genetic groups

The use of different molecular markers allows assessing the different temporal levels (i.e. historical and recent events) in order to explain the distribution of genetic variability among natural populations. The study of cpDNA markers provides information about phylogenetic relationships produced by past evolutionary processes such as post-glacial migration histories (Petit et al. 2005), while nuclear neutral microsatellite loci mainly provides information about recent demographic processes such as gene flow and genetic drift (Van Rossum and Raspé 2018). The present study identified the same haplotypes and cpDNA groups previously described in Galicia (Vera et al. 2014; 2020). Moreover, the wider scenario covered, extending the analyses along the Cantabrian Mountains and Pyrenees, provided more refined information about the distribution of the cpDNA groups. Thus, the cpDNA-group 1 (suggested by Vera et al. (2014) as ancestral) was mainly distributed close to the Cantabrian coast, while the cpDNA-group 2, which was more related to European haplotypes (Vera et al. 2014), was dominant in the inner region and the Pyrenees. The Cantabrian coast and adjacent mountains have been proposed as refugia during glacial periods for different plants such as the common beech (*Fagus sylvatica*) or the European yew (*Taxus baccata*) (Muñoz-Sobrino et al. 2009; Maroso et al. 2021). STRUCTURE results with microsatellite loci for $K=2$ were identical to those provided by

cpDNA markers. Moreover, $K=5$ was also identified as a probable clustering option following Evanno's methodology. In this case, the four previous population units identified in Galicia (Vera et al. 2020) were found, while a new one was associated with RIPO, the single location sampled at the Pyrenees. However, IBD patterns were not detected suggesting a relevant effect of genetic drift in the current population structure due basically to fragmentation (high F_{ST} values) and small effective size of locations. Due to previous biochemical and biometric analyses of the species (see Vera et al. 2020), the presence of adaptive variation cannot be ruled out, and further analyses at the genomic level using a higher number of molecular markers, such as Single Nucleotide Polymorphisms (SNPs) easily obtained using Genotyping-By-Sequencing (GBS) techniques developed thanks to the Next Generation Sequencing (NGS) methodologies (Casanova et al. 2021), are necessary.

Conservation recommendations

This study has confirmed the presence and singularity of the different genetic groups of *A. montana* in the North of the Iberian Peninsula, highlighting the interest in their conservation due to the peripheral situation of these populations within the natural distribution of the species. Following the genetic assessment framework for threatened plants developed by Ottewell et al. (2016), in situ recovery, recruitment increment and translocations would be suggested as management actions for conservation. Frankham (2022) has recently postulated the goals for restoration of genetic diversity in wild populations which would be applied by the Convention on Biological Diversity in the period 2030–2050. The two main goals are: i) the long-term retention of evolutionary potential and ii) connectivity enhancement among fragmented populations. Given the threatened status of the species, enforcement of legal figures for protection and regulated use of the genetic resources of this species is encouraged, as well as promotion of germplasm banks and habitat protection (Moreno et al. 2019), which may require interactions among different regional and local administrations in the North of Spain. To achieve the second goal, actions such as the increase of gene flow and genetic rescue have been suggested. Genetic rescue could be a feasible option because it focuses on the human-assisted movements

of individuals. However, crosses among breeders must be controlled to avoid outbreeding depression, although this effect seems not to be as harmful as expected (Frankham 2022). The effectiveness of translocations and genetic rescue in *A. montana* has been recently assessed in Belgian populations (Van Rossum et al. 2020). After two years, the success of translocations in the genetic restoration was established, with no signs of outbreeding depression in the first-generation hybrids. Preliminary analyses regarding the production of Galician *A. montana* in cultivated conditions have suggested a higher production (in terms of the number of flowering heads per plant, mean weight of flowering heads, and flowering head production/plant) in controlled crops compared to natural populations, maintaining their contents in sesquiterpene lactones (Romero et al. unpublished data). Undoubtedly, these crops would also help in the maintenance of wild populations used as germplasm sources and for ex situ conservation, but they would also be interesting for the pharmaceutical industry. Therefore, the identified genetic resources could be used to boost nearby depleted populations or be reintroduced in close regions where the species has disappeared (see Van Rossum et al. 2020), but after a genetic assessment and consideration of the possible local adaptations and adaptive variation.

Conclusion

The present study provides an extended and more refined information about the genetic diversity and structure of *A. montana* populations in the Northern Iberian Peninsula, which are the consequence of both historical and recent evolutionary processes. The presence of different genetic groups is undoubtable, and their preservation should be a priority given the threatened status of *A. montana* in the region. The genetic information provided may be used by the regional administration and policymakers to develop proper management and conservation actions for the reinforcement and recovery of wild populations, including in regions where the species has recently disappeared.

Acknowledgements Authors wish to thank C. Alonso, S. Sánchez-Darriba, L. Insua and S. Gómez for laboratory support and Dr. Lghetti and two anonymous reviewers for their

constructive comments which improved the initial version of this manuscript. The authors are grateful for the granting of the *Arnica montana* sampling permits to the regional administrations of the Principality of Asturias, Castilia and León, the Basque Country and Catalonia. They also thank Amparo Mora (P. Nac. Picos de Europa), Patricio Bariego, Amador Prieto and Roser Melero for their help in locating the Asturian, Basque, Zamoran and Catalan *arnica* populations studied, respectively. This study has been supported by Deputación de Lugo (Galicia, Spain) contract (Arnica studies, 2021/UC1362018-CP071).

Author Contributions Conceptualization: MV, CB, RR and MAR; field work: MAR and RR; formal analysis: IL, AC, CR, MV; funding acquisition: MV and RR; methodology: CB, IL, AC, MV; writing—original draft: MV; writing—review and editing: CB, IL, MAR, AC, CR, RR, MV. All authors have read and agreed to the published version of the manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was supported by Deputación de Lugo (Galicia, Spain) contract (Arnica studies, 2021/UC1362018-CP071).

Declarations

Conflict of interest The authors have no relevant financial or non-financial interest to disclose.

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

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

References

- Allen D, Bilz M, Leaman DJ, Miller RM, Timoshyna A, Window J (2014) European red list of medicinal plants. Publications Office of the European Union, Luxembourg
- Atkinson NJ, Urwin PE (2012) The interaction of plant biotic and abiotic stresses: from genes to the field. *J Exp Bot* 63:3523–3543. <https://doi.org/10.1093/jxb/ers100>
- Beckman NG, Bullock JM, Salguero-Gómez R (2018) High dispersal ability is related to fast life-history strategies. *J*

- Ecol 106:1349–1362. <https://doi.org/10.1111/1365-2745.12989>
- Birks HJB (2019) Contributions of quaternary botany to modern ecology and biogeography. *Plant Ecol Divers* 12:189–385. <https://doi.org/10.1080/17550874.2019.1646831>
- Bolos Y, Vayreda A (1946) El *Arnica montana* L. en la Península Ibérica. *Farmacognosia. Anales Del Inst José Celestino Mutis* 7:145–151
- Brown JH (1984) On the relationship between abundance and distribution of species. *Amer Nat* 124:255–279
- Casanova A, Maroso F, Blanco A, Hermida M, Rios N, Garcia G, Manuzzi A, Zane L, Verissimo A, Garcia-Marin JL, Bouza C, Vera M, Martinez P (2021) Low impact of different SNP panels from two building-loci pipelines on RAD-Seq population genomic metrics: case study on five diverse aquatic species. *BMC Genomics* 22:150. <https://doi.org/10.1186/s12864-021-07465-w>
- Cires E, Samain M-S, Goetghebeur P, Fernandez Prieto JA (2011) Genetic structure in peripheral Western European populations of the endangered species *Cochlearia pyrenaica* (Brassicaceae). *Plant Syst Evol* 297:75–85. <https://doi.org/10.1007/s00606-011-0500-9>
- Cires E, Perez R, Bueno A, Prieto JAF (2013) Genetic diversity in peripheral and central populations of the Cantabrian endemism *Genista legionensis* (Pau) M. Lainz (Fabaceae). *An Jard Bot Madrid* 70:91–96. <https://doi.org/10.3989/ajbm.2313>
- Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) NEESTIMATOR v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol Ecol Resour* 14:209–214. <https://doi.org/10.1111/1755-0998.12157>
- Duwe VK, Ismail SA, Buser A, Sossai E, Borsch T, Muller LAH (2015) Fourteen polymorphic microsatellite markers for the threatened *Arnica montana* (ASTERACEAE). *Appl Plant Sci* 3:1400091. <https://doi.org/10.3732/apps.1400091>
- Duwe VK, Muller LAH, Borsch T, Ismail SA (2017) Pervasive genetic differentiation among Central European populations of the threatened *Arnica montana* L. and genetic erosion at lower elevations. *Perspect Plant Ecol Evol Syst* 27:45–56. <https://doi.org/10.1016/j.ppees.2017.02.003>
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform* 1:47–50
- Falniowski A, Bazos I, Hodálová I, Lansdown R, Petrova A (2013) *Arnica montana*. The IUCN Red List Threat Species 2013: eT162327A5574104
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 7:574–578. <https://doi.org/10.1111/j.1471-8286.2007.01758.x>
- Ferguson IK (1976) *Arnica* L. In: Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA (eds) *Flora europaea*. Cambridge University Press, Cambridge, pp 189–190
- Francis RM (2017) POPHELPER: an R package and web app to analyse and visualize population structure. *Mol Ecol Resour* 17:27–32. <https://doi.org/10.1111/1755-0998.12509>
- Frankham R (2022) Evaluation of proposed genetic goals and targets for the convention on biological diversity. *Conserv Genet*. <https://doi.org/10.1007/s10592-022-01459-1>
- Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to conservation genetics*. Cambridge University Press, Cambridge
- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Accessed from <http://www.unil.ch/izea/software/fstat.html>.
- Hautekeete NC, Frachon L, Luczak C, Toussaint B, Van Landuyt W, Van Rossum F, Piquot Y (2015) Habitat type shapes long-term plant biodiversity budgets in two densely populated regions in north-western Europe. *Divers Distrib* 21:631–642. <https://doi.org/10.1111/ddi.12287>
- Hewitt GM (2000) The genetic legacy of the quaternary ice ages. *Nature* 405:907–913. <https://doi.org/10.1038/35016000>
- Jombart T, Ahmed I (2011) Adegnet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Jump AS, Woodward FI, Burke T (2003) *Cirsium* species show disparity in patterns of genetic variation at their range-edge, despite similar patterns of reproduction and isolation. *New Phytol* 160:359–370. <https://doi.org/10.1046/j.1469-8137.2003.00874.x>
- Kahmen S, Poschod P (2000) Population size, plant performance, and genetic variation in the rare plant *Arnica montana* L. in the Rhon. *Germany Basic Appl Ecol* 1:43–51. <https://doi.org/10.1078/1439-1791-00007>
- Kriplani P, Guarve K, Baghael US (2017) *Arnica montana* L.—a plant of healing: review. *J Pharm Pharmacol* 69:925–945. <https://doi.org/10.1111/jphp.12724>
- Lass C, Vocanson M, Wagner S, Schempp CM, Nicolas JF, Merfort I, Martin SF (2008) Anti-inflammatory and immune-regulatory mechanisms prevent contact hypersensitivity to *Arnica montana* L. *Exp Dermatol* 17:849–857. <https://doi.org/10.1111/j.1600-0625.2008.00717.x>
- Leitao ST, Araujo S, Rubiales D, Vaz Patto MC (2020) Abiotic and biotic stresses interaction in Fabaceae plants. Contributions from the grain legumes/soilborne vascular diseases/drought stress triangle. In: Hasanuzzaman M, Araujo S, Gill S (eds) *The plant family fabaceae*. Springer, Singapore, pp 237–260
- Luijten SH, Gerard J, Oostermeijer MJB, Van Leeuwen NC, den Nijs HCM (1996) Reproductive success and clonal genetic structure of the rare *Arnica montana* (Compositae) in the Netherlands. *Plant Syst Evol* 201:15–30
- Luijten SH, Dierick A, Oostermeijer JGB, Raijmann LEL, Den Nijs HCM (2000) Population size, genetic variation, and reproductive success in a rapidly declining, self-incompatible perennial (*Arnica montana*) in The Netherlands. *Conserv Biol* 14:1776–1787

- Mardari C, Birsan C, Stefanache C, Schiopu R, Grigoras V, Balaes T, Danila D, Tanase C (2019) Population structure and habitat characteristics of *Arnica montana* L. in the NE Carpathians (Romania). *Tuexenia* 39:401–421. <https://doi.org/10.14471/2019.39.012>
- Maroso F, Vera M, Ferreiro J, Mayol M, Riba M, Ramil-Rego P, Martínez P, Bouza C (2021) Genetic diversity and structure of *Taxus baccata* from the Cantabric-Atlantic area in northern Spain: a guide for conservation and management actions. *Forest Ecol Manag* 482:118844. <https://doi.org/10.1016/j.foreco.2020.118844>
- Maurice T, Matthies D, Muller S, Colling G (2016) Genetic structure of colline and montane populations of an endangered plant species. *AoB Plants* 8:plw057. <https://doi.org/10.1093/aobpla/plw057>
- McMinn RL, Russell FL, Beck JB (2017) Demographic structure and genetic variability throughout the distribution of Platte thistle (*Cirsium canescens* Asteraceae). *J Biogeogr* 44:375–385. <https://doi.org/10.1111/jbi.12879>
- Moreno JC, Iriondo JM, Martínez F, Martínez J and Salazar C (2019). Atlas y libro rojo de la flora vascular amenazada de España. Adenda 2017. Ministerio para la Transición Ecológica-Sociedad Española de Biología de la Conservación de Plantas. Madrid. 220 pp. Accessed from: https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/arnicamontana_tcm30-505849.pdf
- Muñoz-Sobrino C, Ramil-Rego P, Gomez-Orellana L, Ferreiro da Costa J, Díaz Varela RA (2009) Climatic and human effects on the post-glacial dynamics of *Fagus sylvatica* L. in NW Iberia. *Plant Ecol* 203:317–340. <https://doi.org/10.1007/s11258-008-9552-5>
- Ouborg NJ, Vergeer P, Mix C (2006) The rough edges of the conservation genetics paradigm for plants. *J Ecol* 94:1233–1248. <https://doi.org/10.1111/j.1365-2745.2006.01167.x>
- 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. *Divers Distrib* 22:174–188. <https://doi.org/10.1111/ddi.12387>
- Peakall R, Smouse PE (2012) GenA1Ex 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Perry NB, Burgess EJ, Rodriguez Guitian MA, Romero R, Lopez Mosquera E, Smallfield BM, Joyce NI, Littlejohn RP (2009) Sesquiterpene lactones in *Arnica montana*: helenalin and dihydrohelenalin chemotypes in Spain. *Planta Med* 75:660–666. <https://doi.org/10.1055/s-0029-1185362>
- Petit RJ, Dumnil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol Ecol* 14:689–701. <https://doi.org/10.1111/j.1365-294X.2004.02410.x>
- QGIS Development Team (2017) QGIS geographic information system. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>
- R Core Team (2018) R: a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. <https://www.R-project.org/>
- Rohlf F (1993) NTSYS-pc. Numerical taxonomy and multivariate analysis system, Version 2.1. Setauket, New York
- Romero R, Real C, Guitián MA, Barros MR, Rigueiro A, González-Hernández MP (2011) Estudio de la variabilidad biométrica de *Arnica montana* L. (Asteraceae) en el extremo occidental cantábrico (NW Ibérico), Actas del IX Col-loqui Internacional de Botànica Pirenaico-Cantàbrica: 379–388
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145:1219–1228
- Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8:103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Schmiderer C, Torres-Londono P, Lutz-Roeder A, Duwe VK, Novak J (2018) *Arnica montana* subsp *atlantica*: Really a subspecies? *Genet Resour Crop Evol* 65:1829–1838. <https://doi.org/10.1007/s10722-018-0653-2>
- Shay JE, Pennington LK, Montiel-Molina JAM, Toews D, Hendrickson B, Sexton J (2021) Rules of plant species ranges: applications for conservation strategies. *Front Ecol Evol* 9:700962. <https://doi.org/10.3389/fevo.2021.700962>
- Siedle B, Garcia-Pineros AJ, Murillo R, Schulte-Monting J, Castro V, Rungeler P, Klaas CA, Da Costa FB, Kisiel W, Merfort I (2004) Quantitative structure—activity relationship of sesquiterpene lactones as inhibitors of the transcription factor NF-kappa B. *J Med Chem* 47:6042–6054. <https://doi.org/10.1021/jm049937r>
- Travis JMJ, Dytham C (2004) A method for simulating patterns of habitat availability at static and dynamic range margins. *Oikos* 104:410–416
- Van Rossum F, Raspé O (2018) Contribution of genetics for implementing population translocation of the threatened *Arnica montana*. *Conserv Genet* 19:1185–1198. <https://doi.org/10.1007/s10592-018-1087-2>
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535–538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
- Van Rossum F, Hardy OJ, Le Pajolec S, Raspé O (2020) Genetic monitoring of translocated plant populations in practice. *Mol Ecol* 29:4040–4058. <https://doi.org/10.1111/mec.15550>
- Vera M, Romero R, Rodriguez-Guitian MA, Barros RM, Real C, Bouza C (2014) Phylogeography and genetic variability of the *Arnica montana* chemotypes in NW Iberian Peninsula. *Silvae Genet* 63:293–300. <https://doi.org/10.1515/sg-2014-0037>
- Vera M, Mora G, Rodriguez-Guitian MA, Blanco A, Casanova A, Real C, Romero R, Bouza C (2020) Living at the edge: population differentiation in endangered *Arnica montana* from NW Iberian Peninsula. *Plant Syst Evol* 306:44. <https://doi.org/10.1007/s00606-020-01673-9>

- Villar L, Ortiz S, Serrano S (2019) *Arnica*. In: Castroviejo, S. (Coord.) Flora Iberica, Vol. XVI (III). CSIC, Madrid, pp 2198–2203
- Zidorn C (2010) Altitudinal variation of secondary metabolites in flowering heads of the Asteraceae: trends and causes. *Phytochem Rev* 9:197–203. <https://doi.org/10.1007/s11101-009-9143-7>

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