Conservation Genetics (2024) 25:227–243
https://doi.org/10.1007/s10592-023-01554-x

RESEARCH ARTICLE

In situ conservation of *Heloscidium nodiflorum*: a crop wild relative of celery in Germany

Inga Mewis[1] · Nadine Bernhardt[2] · Maria Bönisch[2] · Jonathan Brassac[2] · Annette Naumann[1] · Matthias Zander[1] · Uta Schirmak[3] · Marion Nachtigall[3] · Christian Ulrichs[1]

Received: 14 November 2022 / Accepted: 24 July 2023 / Published online: 16 September 2023
© The Author(s) 2023

Abstract

Crop wild relatives (CWR) represent valuable sources of traits that can enhance the abiotic and biotic stress tolerance of crops. Due to climate change and expansion of land use, they may suffer from population decline, which increases the risk of losing these resources. The in situ conservation of CWR is receiving growing attention to preserve their intraspecific diversity. In contrast to the safeguarding of seeds ex situ, the preservation of taxa in their natural habitat facilitates the development of new adaptations to changing environmental conditions. However, studies aiming at the establishment of their in situ conservation are rare. Therefore, we identified and genotyped *Heloscidium nodiflorum* occurrences in Germany. *Heloscidium nodiflorum* is an endangered CWR of cultivated celery (*Apium graveolens*). For 20 occurrences the number of individuals, the number of sexually reproducing individuals, and potential threats were recorded. Genetic differences were examined using 17 newly developed simple sequence repeats (SSR). Population genetic analyses, like the inference of the compositional genetic differentiation, of population structure, and a Discriminant Analysis of Principal Components were conducted. The analyses revealed that most occurrences were genetically distinct. Geographically closely located populations are often also genetically more similar. Several populations revealed a composite genome indicative of hybridizations between distinct populations. Populations that experienced fixation were also identified. Based on the genetic diversity, the location in different ecogeographic units, and the support of local stakeholders, we recommend the establishment of 15 genetic reserves for *H. nodiflorum* in Germany.

Keywords Wild celery · In situ conservation · Plant genetic resources · Simple sequence repeats · Genetic diversity · Genetic reserves

Introduction

Numerous examples of the transfer of beneficial traits from a crop wild relative (CWR) into the related crop through introgression (Gong et al. 2022) underpin the need for the preservation of genetic resources for plant breeding (Dwivedi et al. 2009; Kole 2011). A significant number of CWR are endangered due to genetic erosion (Bilz et al. 2011) which could compromise breeding progress today and in the future (Wehling et al. 2017). Genetic erosion was mostly studied with a focus on landraces (Khoury et al. 2022), particularly, on annual cereal crops such as maize, wheat, rice, barley, and sorghum. In contrast, the loss of genetic diversity in CWR is not well known.

The in situ conservation of CWR was recognized as an important approach to safeguarding genetic resources. For example, the parties of the International Treaty on Plant

Inga Mewis and Nadine Bernhardt have contributed equally to this work.

✉ Inga Mewis
inga@entomology.de

✉ Nadine Bernhardt
nadine.bernhardt@julius-kuehn.de

[1] Division Urban Plant Ecophysiology, Humboldt-Universität zu Berlin, Lenzeallee 55-57, Berlin, Germany

[2] Julius Kühn-Institut (JKI) – Federal Research Centre for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Erwin-Baur-Str. 27, 06484 Quedlinburg, Germany

[3] Julius Kühn-Institut (JKI) – Federal Research Centre for Cultivated Plants, Institute for Breeding Research on Agricultural Crops, Erwin-Baur-Str. 27, 06484 Quedlinburg, Germany

Genetic Resources for Food and Agriculture (ITPGRFA) agreed to conserve CWR in situ and to re-establish viable populations in protected areas. This shall be achieved by supporting local communities engaged in the conservation of biological diversity (Art. 5 (d)) and by improving the sustainable use of plant genetic resources for food and agriculture (Art. 5 (e), Food and Agriculture Organisation 2001). However, the technique is globally poorly developed and receives little consideration (Vincent et al. 2022; Wambugu and Henry 2022). In Europe, the Natura 2000 protected area network was initiated by the nature conservation sector (European Council 1992). Maxted et al. (2015) suggested using this network to manage the genetic diversity of CWR in situ. In Germany, national programs aim for the implementation of international agreements. Consequently, i.e. the "Nationale Fachprogramm zur Erhaltung und nachhaltigen Nutzung pflanzengenetischer Ressourcen landwirtschaftlicher und gartenbaulicher Kulturpflanzen" of Germany (BLE 2012) advocates the protection of CWR in their natural habitat.

According to a proposal by Maxted et al. (1997) a genetic reserve conservation technique was defined as "the location, management, and monitoring of genetic diversity in natural wild populations within defined areas designated for active, long-term conservation". The goal is the conservation of a set of evolving populations that represent the intraspecific diversity of a target taxon. This technique also covers the ex situ conservation of germplasm as a backup copy, which facilitates access to the germplasm for research and breeding. Finally, the locations nominated for the establishment of genetic reserves will be managed by a network of institutions and people (Frese et al. 2018a; Maxted et al. 2015).

The financial resources for the in situ conservation of CWR are very limited. Currently, CWR generally profit only indirectly from conservation actions, if they are located within nature reserves. Management measures addressing them directly usually rely on local stakeholders that are willing to support this. Typically, they do not receive compensation for their expenses (Bönisch and Thormann 2021). Therefore, the selection of the most suited genetic reserve candidates is crucial. Genetic markers are a powerful resource of information to aid these decisions (Gradl et al. 2022).

When choosing sites for in situ conservation, the selection generally leaned on the minimum quality standards according to Iriondo et al. (2012). Briefly, the selected occurrences altogether should represent the intraspecific genetic diversity of the species under consideration. Therefore, genetic characteristics are considered in the selection and genetic reserve candidates should represent different natural habitats and ecogeographic units (EGU) to cover local adaptations. Occurrences should have a good and stable conservation status. Therefore, large and vital populations are selected when no acute threats can be identified. Further, the current management measures should allow the continued existence of the population for a foreseeable future. Sites located in existing protected areas are selected preferentially since long-term conservation conditions can be assumed. The local stakeholders to be involved in the establishment of the genetic reserves and the subsequent management should be willing to cooperate. These are in particular landowners and users as well as responsible nature conservation authorities. In detail, the considerations for the selection of genetic reserves are presented in Bönisch and Frese (2020).

The present work was part of a model project aiming at the implementation of Art. 5 of the ITPGRFA in Germany. To this end, four wild celery species native to Germany (i.e. *Apium graveolens* L., *Helosciadium inundatum* (L.) W. D. J. Koch, *H. nodiflorum* (L.) W. D. J. Koch and *H. repens* (Jacq.) W. D. J. have been chosen to explore the genetic reserve conservation technique (Bönisch et al. 2020; Bönisch and Thormann 2021; Herden et al. 2020). Here we show the results of the model taxon *H. nodiflorum.*

Helosciadium nodiflorum (L.) W. D. J. Koch (Syn. *Apium nodiflorum* L.) is a perennial species of the Apiaceae family. The taxon is insect-pollinated (Knuth 1898) but also self-compatible. Moreover, it can reproduce vegetatively via stolons (Bundesamt für Naturschutz 2017). It has its main distribution area in southwestern Europe. Its range extends in the northeast to the German federal states Rhineland-Palatinate, Hesse, Baden-Württemberg, North Rhine-Westphalia, and Saarland. The species occurs in small flowing waters or embankment areas with sparse vegetation. In addition to shading, competition from more vigorous plant species, flood events or overly intensive watercourse maintenance have a negative effect on populations. *Helosciadium nodiflorum* has been selected for this study as it has several interesting features. The taxon is (i) an endangered species (Metzing et al. 2018), (ii) shows resistance to the leafminer, *Lyriomiza trifolii* Burgess (Trumble et al. 1990, 1998), (iii) is a good accumulator of the elements Cd and Cu, and hence, could be used to clean contaminated water in a cost-efficient way (Kara 2014), (iv) is an edible but forgotten CWR (Guijarro-Real et al. 2019), and (v) contains a range of still largely unexplored secondary plant compounds (Guijarro-Real et al. 2015, 2019; Maggi et al. 2019).

Here we present an analysis of the genetic diversity of *H. nodiflorum* in Germany. In 2015 and 2016, individual numbers and threats of the occurrences were documented. Census data were recorded to assess potential fluctuations in population size in-between years. We developed simple sequence repeats (SSR) that were used to genotype 20 *H. nodiflorum* occurrences. Though, genotyping-by-sequencing (GBS) and related methods (Elshire et al. 2011; Poland et al. 2012) are more and more frequently used to address these topics (Leo et al. 2022; Stoll et al. 2020; Žerdoner

Čalasan et al. 2021). The use of SSR is well established to study population genetics and to select occurrences for conservation (Fievet et al. 2007; Frese et al. 2018b, 2019; Herden et al. 2020; Kohn et al. 2018). Indeeed, due to their ubiquitous occurrence in nuclear and organelle genomes, their codominant inheritance, and their high mutation rate, SSR remain a valuable tool. Particularly the high mutation rate leads to high levels of polymorphisms that can be used to assess the genetic similarity of closely related taxa and even between individuals (Guichoux et al. 2011; Hodel et al. 2016; Šarhanová et al. 2018).

In this study, a set of population genetic analyses were performed to assess (1) if the genetic diversity within and between occurrences is different, and to identify (2) what are potential reasons for the differences (e.g. recent and historical events shaping the genetic diversity). (3) We evaluate the data to see if populations can be delimitated. (4) Finally, genetic and census data were combined to identify the best candidates to establish genetic reserves. The data will provide a baseline for future monitoring of the genetic diversity of the taxon.

Materials and methods

Identification of sites, mapping, and sampling

Ecogeographic diversity can be used as a proxy for genetic differences between populations (Parra-Quijano et al. 2012). Therefore, sites located in different ecogeographic units (EGU) can be chosen to increase the chance of capturing a wide range of intraspecific genetic diversity. In this study, EGU were based on the main units as described in Meynen and Schmithüsen (1953) as they likely represent regions with specific abiotic and biotic characteristics. Regarding the selection of occurrences, particular value was given to the distribution of occurrence in different EGU, a high population abundance, and vitality indicating a good conservation state, as well as an interest of local stakeholders in the in situ conservation of the target species (Bönisch and Thormann 2021; Frese et al. 2018b).

In Germany, *H. nodiflorum* sites are mainly located in Rhineland-Palatinate and sporadically in Hesse, North Rhine-Westphalia, Baden-Württemberg, and Saarland. Information on *H. nodiflorum* occurrences provided by nature conservation authorities of the federal states were compiled and harmonized. About 250 sites had been reported, out of which 58 were considered potential genetic reserves. A revisiting of these sites in 2015 showed that only 43 still exist. The 43 sites (Fig. 1) were assessed to determine which ones were most appropriate for the sampling of leaf material for genetic analysis. Finally, leaf material has been sampled at 22 sites (Table 1) from 30 randomly chosen individuals. Leaves were collected, placed in filter bags, and silica-dried to conserve the DNA. Samples were taken from three occurrences in 2015 and from 19 occurrences in 2016. On-site monitoring was performed, e.g. pictures and geographic coordinates were taken to document the habitat. Furthermore, environmental conditions, including the risk of gradual degradation (loss of occurrences), have been recorded (Table S1). Seeds of some of the occurrences were collected in 2018 and deposited at the Genbank WEL (https://www.genbank-wel.uni-osnabrueck.de/index.php/de/).

SSR marker development

Genomic DNA was prepared from 8 mg of dried and ground leaf material using the InnuPREP Plant DNA Kit (Analytik Jena AG, Jena, Germany) according to the manufacturer's instructions with some modifications. From the manual of the manufacturer "Protocol 1" was selected for DNA isolation using Lysis Solution SLS. The solution was directly added to the plant samples and vortexed for 5 s, twisted, and then incubated for 30 min at 65 °C. The optional step to remove RNA was omitted. Because no SSR markers for genetic analysis of *H. nodiflorum* were generated previously, the company TraitGenetics GmbH (Gatersleben, Germany) was assigned to generate SSR markers using *H. nodiflorum* material from Kew Garden (accession number 200497). Using partial sequencing and subsequent marker analysis, TraitGenetics GmbH revealed 33 functional markers. For identifying informative genetic markers, leaves were used from ten individuals for each of three occurrences, i.e. ALB, RI1, and KNI (Table 1). The screening revealed a set of 17 polymorphic SSR markers (Table 2). All SSR markers are dinucleotides and the nucleotide motives are repeated up to 23 times.

In total, DNA was extracted from 660 individuals (30 specimens of each of the 22 occurrences). The DNA quality was assessed using a NanoDrop (Thermo Scientific) and gel electrophoresis. SSR markers were amplified in a total volume of 20 µl. The PCR mix consisted of 2 µl DNA extract (around 20–100 ng template DNA), 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.25 µM of each primer, and 0.5 U GoTaq DNA polymerase (Promega, Wisconsin, USA). The following PCR program was performed using a MyCycler™ Thermal Cycler System (Bio-Rad, München, Germany): 2 min 94 °C, 35 cycles of 1 min at 94 °C, 1 min 54 °C, and 2 min at 72 °C, followed by a final 5 min at 72 °C. Microsatellite analysis was conducted via capillary electrophoresis (ABI 3130xl Genetic Analyzer, Applied Biosystems, USA). Fragment sizes were determined using the GeneMapper software version 4.1 and recorded in a database developed by Applied Biosystems. The PCR of a sample was repeated once in case an expected fragment was absent. If the fragment remained absent, the allele was documented as a null allele. The

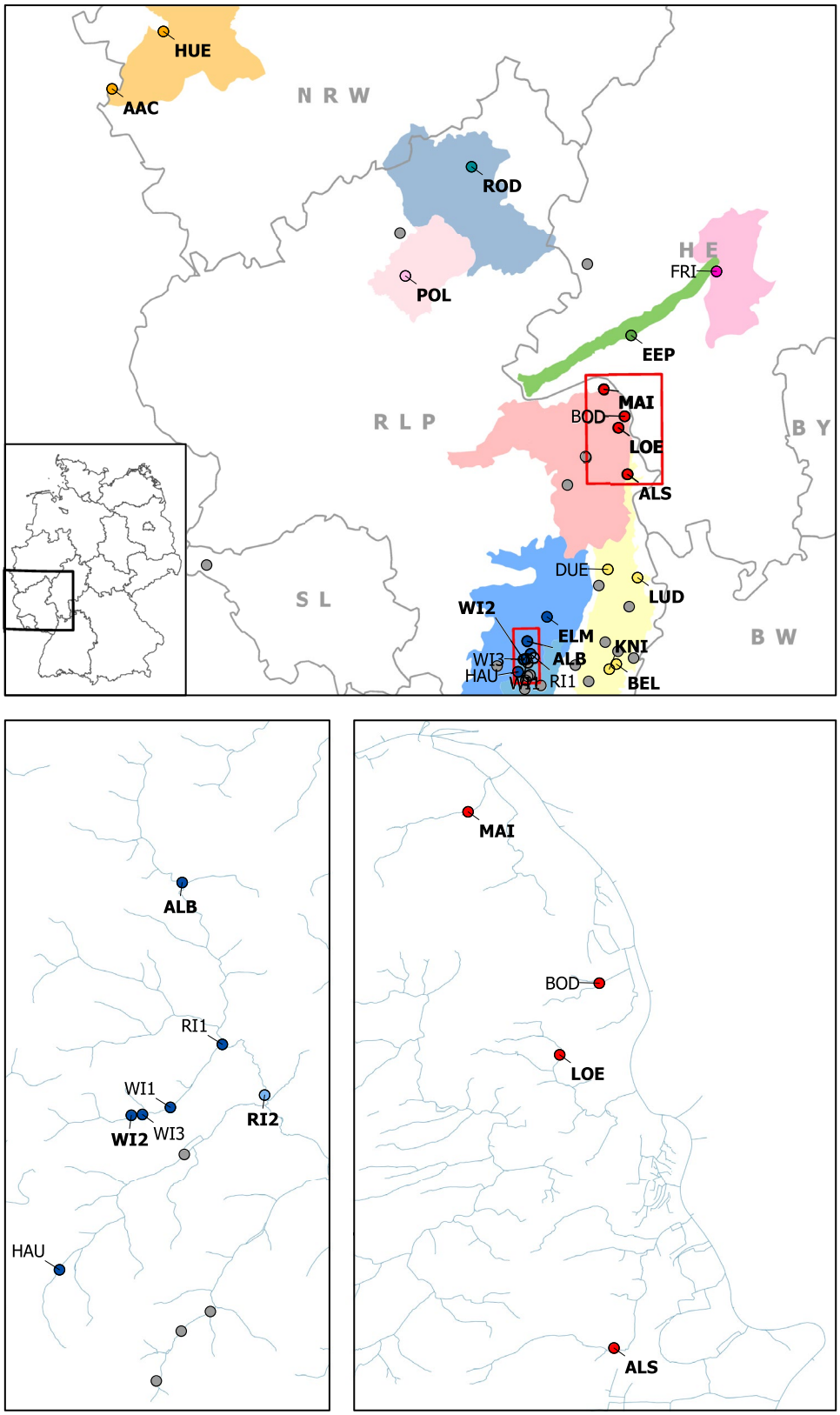


Fig. 1 Location of confirmed *H. nodiflorum* occurrences in Hesse (HE), North Rhine-Westphalia (NRW), Rhineland-Palatinate (RLP) and Saarland (SL) (top). Enlarged representation of the occurrences from Pfälzerwald and Dahner Felsenland as well as from Alzeyer Hügelland (bottom). Occurrences that were sampled are colored, others are grey. The color coding indicates the ecogeographic units (EGU, Table 1). Relevant EGU are shown as light colored areas. Occurrences nominated as genetic reserve are written in bold. Maps were created in ArcGISPro version 2.9.3. Boarders of the federal states were taken from © GeoBasis-DE/BKG 2007 and the water network from the Wasserwirtschaftsverwaltung Rheinland-Pfalz, https://www.geoportal.rlp.de/spatial-objects/343; accessed in June 2022)

Table 1 Geographic origin of the *H. nodiflorum* occurrences considered in the study

ID	Federal state	Municipality	First mention (year)	Ecogeographic unit
EPP	HE	Eppstein	1996	Hoher Taunus
FRI	HE	Friedberg (Hesse)	1990	Wetterau
AAC	NRW	Aachen	1991	Jülicher Börde
HUE	NRW	Hückelhoven	1985	Jülicher Börde
ALS	RLP	Alsheim	2010	Alzeyer Hügelland
BOD	RLP	Bodenheim	2012	Alzeyer Hügelland
LOE	RLP	Lörzweiler	2012	Alzeyer Hügelland
MAI	RLP	Mainz	1991	Alzeyer Hügelland
RI2	RLP	Rinnthal	1998	Dahner Felsenland
POL	RLP	Polch	2013	Mittelrheinisches Becken
ROD	RLP	Rodenbach bei Puderbach	1993	Montabaurer Westerwald
ALB	RLP	Albersweiler	1998	Pfälzerwald
ELM	RLP	Elmstein	2012	Pfälzerwald
HAU	RLP	Hauenstein	1998	Pfälzerwald
RI1	RLP	Rinnthal	1998	Pfälzerwald
WI1	RLP	Wilgartswiesen	1998	Pfälzerwald
WI2	RLP	Wilgartswiesen	1998	Pfälzerwald
WI3	RLP	Wilgartswiesen	1999	Pfälzerwald
BEL	RLP	Bellheim	1999	Vorderpfälzer Tiefland
DUE	RLP	Bad Dürkheim	2014	Vorderpfälzer Tiefland
KNI	RLP	Knittelsheim	2002	Vorderpfälzer Tiefland
LUD	RLP	Ludwigshafen am Rhein	2006	Vorderpfälzer Tiefland

Short identifier (ID), federal state (i.e. Hesse (HE), North Rhine-Westphalia (NRW), Rhineland-Palatinate (RLP), municipality, the year the occurrence was first mentioned and the ecogeopgraphic unit according to Meynen and Schmithüsen (1953) are given

Table 2 Overview of the selected SSR markers for genetic analyses in *H. nodiflorum* developed by TraitGenetics GmbH

Locus	Forward primer	Reverse primer	Repeat motive	Allele size range (nt)	Number of alleles
ANM0041	CCCTCTCGGCCAGAATCTT	TTGTCATAAATAAGTTGTCGCAAA	$GA_{(17)}$	114–129	2
ANM0052	CCTCGTAGCTTGTGCCTGTA	TCTTCATCGCTCCTGGATTC	$GA_{(11)}$	130–149	2
ANM0054	AAATAGACAGGCCCTGCAAA	TGTGTGTTGTATGTACGTATCAATGT	$CT_{(14)}$	98–103	3
ANM0056	CAGTTGGTGTGCACAACTAAAGAT	AAACACAGAATGGAGTTGGG	$CT_{(17)}$	157–173	5
ANM0057	AATATTATTGATTGGAGTGCGTTT	TGAGGTTGTAATAGGCTATCATCAGT	$CA_{(19)}$	138–144	2
ANM0058	TCTGAATAGTGTCTATATTGCGGG	CCTGCTCTAACAGAAGCTGAAA	$GA_{(15)}$	141–164	7
ANM0059	CATACAACAGCCATCCATCG	GAGAGTACCAGAAGGGATGAAA	$CT_{(15)}$	140–152	4
ANM0060	CAGCAAACTCCGGTAGGTGT	ATGTTCTCGTAATCGGCACC	$CT_{(13)}$	164–176	5
ANM0061	AAGTATACGCGGCAACTGCT	TGTTTACGCATACTTGTGGTGA	$GA_{(28)}$	132–152	3
ANM0062	TTGTTGCGCAAGACAAAGC	GGGATATGCATACATAAGTCTTGA	$GA_{(10)}$	156–167	4
ANM0063	CCTCATGTCACACTTGCTTCA	GATCTCTTCCGGAAACTCCA	$CT_{(10)}$	165–171	2
ANM0069	TGATTACCTTCGAATTGTTACCC	TGTCATCTGAATTATCTTTGGGAA	$AG_{(18)}$	166–168	2
ANM0071	GCATAGCTCCATATATTGCCG	TGAAGCTTCTCAGGCAAAGA	$CT_{(15)}$	167–175	3
ANM0072	ACAACTTCGCCTTGATACGC	TGTTTCTCATAATACTCAATTCTCTGA	$AC_{(17)}$	164–182	4
ANM0075	TTTGTTGCCTGCATGATTC	CAATGGGAGTCTCTCTCGGT	$GT_{(12)}$	183–189	4
ANM0076	GAGAACAGCAGATCGTTAAGCA	TGAAAGCGAGAAGGAGATGAG	$GA_{(13)}$	178–194	4
ANM0079	AAGCCACATAGCAAACCTGC	CGTGCAAAGTTGTGGTGTCT	$GA_{(14)}$	193–230	4

The markers were initially tested on 30 individuals of three *H. nodiflorum* populations

occurrences FRI and DUE (Table 1) were excluded from the analysis due to a limited number of successful amplification products.

SSR analysis

The fragment sizes of the remaining 600 individuals were determined and statistically analyzed. Eighty-six individuals having null alleles were removed. Hence, the data set used for subsequent analysis consisted of 514 individuals from 20 occurrences. The data sets with and without null alleles set are freely available from the OpenAgrar Repositorium: https://www.openagrar.de/receive/openagrar_mods_00073465. Descriptive genetic parameters (Table 3) were calculated using the SAS procedure ProcALLELE (version 9.4, SAS Institute, Heidelberg Germany). In case of a one nucleotide deviation from a marker-specific frequent fragment size, a manual binning procedure was performed to avoid overestimating the genetic diversity (Amos et al. 2007).

The compositional genetic differentiation (Δ_j) on the gene pool level was calculated using DifferInt (Gillet 2013) to identify two kinds of occurrences that are of special interest for conservation measures: (1) A high value of Δ_j shows that the genetic composition of an occurrence is deviating from the rest of the occurrences (the complement). (2) The lowest values of Δ_j indicate that an occurrence represents the composition of its complement the best. Diagrams were generated for visualization and colored using the software WinFIG (http://www.winfig.com).

The following genetic analysis were performed in R version 4.1.0 (R Core Team 2020). The number of multi-locus genotypes (MLG) and the expected number of MLG at the lowest common sample size (eMLG) were inferred using the R package *poppr* version 2.9.3 (Kamvar et al. 2014, 2015). The number of private alleles was inferred using the function *private_alleles*. The allelic richness was inferred using the *allelic.rich* function of the library *PopGenReport* (Mousadik and Petit 1996).

The degree of fixation within an occurrence was evaluated by the fixation index F_{IS} which is a measure of the intra-population underrepresentation of heterozygotes (Nei and Chesser 1983; Meirmans and Hedrick 2011). F_{IS} was inferred using FSTAT version 2.9.4 (Goudet 1995). Randomizations were conducted to test if the values are significantly different from zero. Expected and observed heterozygosity was estimated using the functions *Hs* and *Ho* of the R package *adegenet* version 2.1.5 (Jombart 2008; Jombart and Ahmed 2011).

Contrasting the expected and observed heterozygosities of each population, recent demographic changes can be inferred. Bottleneck version 1.2.02 (Piry et al. 1999) was used to test if the populations experienced bottlenecks. The tests were performed with 1000 iterations under the conservative one-step stepwise mutation model (SMM; Ohta and Kimura 1973) and the two-phase model (TPM; Di Rienzo et al. 1994) with 95% single-step mutations and a variance of 12 for the multiple steps as recommended for most SSR data (Piry et al. 1999). A Wilcoxon signed-rank test was used to detect populations impacted by a recent bottleneck.

Table 3 Overview on the number of alleles (including null alleles), the polymorphic information content (PIC), observed and expected heterozygosity (H_o, H_e) for each SSR marker in the data set consisting of 514 individuals

Locus	Number of alleles	Polymorphic information content (PIC)	Observed heterozygosity (Ho)	Expected heterozygosity (He)
ANM0041	6	0.370	0.202	0.392
ANM0052	4	0.558	0.302	0.630
ANM0054	5	0.529	0.253	0.606
ANM0056	11	0.829	0.543	0.848
ANM0057	6	0.444	0.123	0.474
ANM0058	12	0.758	0.260	0.788
ANM0059	6	0.594	0.218	0.625
ANM0060	13	0.837	0.407	0.852
ANM0061	9	0.623	0.232	0.854
ANM0062	7	0.690	0.142	0.732
ANM0063	5	0.270	0.043	0.308
ANM0069	8	0.679	0.190	0.678
ANM0071	6	0.491	0.300	0.566
ANM0072	6	0.328	0.172	0.370
ANM0075	7	0.679	0.140	0.723
ANM0076	10	0.564	0.60	0.644
ANM0079	17	0.676	0.287	0.691

A Discriminant Analysis of Principal Components (DAPC) was carried out to identify and describe genetic clusters with *adegenet*. The DAPC applies a multivariate statistical method in which the variance in the data set is divided into a between-group and within-group component, which aims at maximizing the discrimination between groups. The analysis was done using the function *dapc* and retaining 35 principle components (PCs) and three discriminant functions. The functions *scatter.dapc* and the R packages *ggplot2* version 3.4.2. (Wickham 2016) and *ggpubr* version 0.6.0 (Kassambara 2023) was used for visualization.

Isolation-by-distance was tested using the Mantel test between a matrix of genetic distances and a matrix of geographic distances using the *mantel.randtest* function of the R library *ade4* (Dray and Dufour 2007). The genetic distance was inferred using function *dist.genpop* and Edwards'distance. The genetic and geographic distances were plotted to be able to distinguish continuous clines from distant patches. For enhanced visualization, the local density was measured using the function *kde2d* (i.e. 2-deminensial kernel density estimation) of the R package *MASS* version 7.3–58.2 (Venables and Ripley 2002).

Population genetic structure was estimated and visualized using the R package *tess3r* version 1.1.0 (Caye et al. 2016; Caye and François 2016). *Tess3r* makes use of both genotypic and geographic data to estimate population genetic structure. The dataset was investigated assuming up to 20 clusters (K = 1–20) with ten iterations each. The number of clusters that explains the data the best was inferred based on cross-validation. The Q-matrix was visualized for all individuals and different values of K using the *barplot* function. The *plot* function was used to interpolate the values of the Q-matrix on a geographic map.

The data were analyzed with BayessAss 3.0.5 (BA3, Wilson and Rannala 2003) to estimate recent migration rates between population pairs using Markov Chain Monte Carlo (MCMC). BayessAss estimates rates of recent gene flow using multilocus genotype data. First, the structure file was converted to the immanc format using two perl scripts (onelinestr2str.pl and stacksStr2immanc.pl from https://github.com/stevemussmann/file_converters/). The BA3MSAT analysis was run for 10 million iterations, of which one million iterations were discarded as burn-in and sampling every 1000 iterations. The analysis was repeated 10 times setting the mixing parameters -a 0.3, -f 0.6, and -m 0.2 for allele frequencies, inbreeding coefficients and migration rates, respectively to obtain acceptance rates of those parameters within the optimal range of 20%–60% (cf. BA3 manual). The traces were inspected in *Tracer* version 1.7 (Rambaut et al. 2018) to detect any convergence issues. Following Meirmans (2014), the Bayesian deviance (Faubet et al. 2007) was calculated to identify the run with the best fit.

Results

Number of individuals

The years of the first mention of the occurrences (Table 1) vary between 1985 and 2014. All occurrences persisted from 2015 to 2016. The estimated number of individuals was the lowest for BOD, RI1, and ROD in 2015 and 2016 with less than 100 specimens per location. The most populated locations for 2015 and 2016 were BEL, MAI, and KNI with at least 1000 individuals in both years. Between 2015 and 2016 only six occurrences remained stable, a majority (11) increased by 1.6 to 37.5 times and for five locations the population size had been reduced by half or more (Table S1a). ALB, ELM, and ROD were the only occurrences that showed no sexually reproducing individuals (Tables 1, and S1a).

SSR marker development

The genetic differences within and between the 20 occurrences of *H. nodiflorum* in Germany were assessed using 17 newly developed SSR (Table 2). The median for the number of distinct alleles per SSR locus is seven. Marker ANM00052 is with four alleles the locus with the lowest number of alleles. In contrast, with 17 alleles, marker ANM0079 is the most diverse (Table 3). The polymorphic information content (PIC) varied between 0.270 (marker ANM0063) and 0.837 (marker ANM0060). High PIC values reveal the suitability of the respective markers for genetic analyses. The observed heterozygosity (H_o) ranges from 0.043 (ANM0063) to 0.543 (ANM0056) and the expected (He) heterozygosity from 0.308 (ANM0063) to 0.852 (ANM0056).

Population genetic analyses

The data set without null alleles consisted of 514 individuals. In total, 115 distinct alleles and 322 multi-locus genotypes (MLG) were detected (Fig. S1, Table 4). The sites EPP and RI2 showed only three MLG. For both sites, one MLG was dominating as it was shared by 25 and 27 individuals, respectively. In contrast, all individuals of LUD, LOE, and KNI carried different MLG. The other occurrences had one MLG shared by a few individuals while most were unique to single individuals (e.g. AAC, POL, and ELM). Only one MLG was shared between two occurrences (i.e. WI1 and WI3, Table S1b).

The occurrences with the highest number of private alleles were ALS, HUE, and MAI, with 26, 24, and 22 private alleles, respectively. EPP, AAC, LOE, RI2, HAU, WI1,

Table 4 Overview of individual counts and genetic analyses

Occurrence	Individuals (2015)	Individuals (2016)	Sexually reproducing individuals (2016)	N	MLG	eMLG	Private alleles	Mean allelic richness	F_{IS}	H_{obs}	H_{exp}	Δj	Additional arguments for and against the nomination as genetic reserves
EPP[a]	120	300	100	28	3	2.26	0	1.37	**-0.714**	0.189	0.110	0.575	sole occurrence in EGU on the edge of the species range, vital, phenotypic variability
FRI	90	200	50										low number of individuals, no information on genetics
AAC[a]	8	300	250	29	14	8.54	0	2.02	**0.172**	0.178	0.216	0.536	occurrence on the edge of the species range, vital, local stakeholders are willing to support it
HUE[a]	8	300	250	20	19	13.52	24	2.05	0.016	0.250	0.254	0.532	occurrence on the edge of the species range
ALS[a]	500	1000	1000	29	16	9.46	26	1.78	**-0.137**	0.318	0.280	0.589	large, stable and vital occurrence, not endagered
BOD	20	80	80	28	10	6.89	6	1.87	0.029	0.261	0.268	0.505	low number of individuals, low genetic diversity
LOE[a]	500	1000	1000	23	23	14.00	0	3.04	**-0.202**	0.537	0.447	0.290	large occurrence, represents the genetic composition of the other occurrences the best
MAI[a]	1000	1000	1000	28	26	13.52	22	1.84	0.051	0.246	0.259	0.532	large, stable and vital occurrence, not endagered, site management exists
RI2	80	100	50	30	3	2.19	0	1.12	**-0.865**	0.110	0.059	0.463	low number of individuals, low genetic diverstiy and signs of inbreeding
POL[a]	500	1000	1000	28	21	11.99	9	2.44	**0.092**	0.328	0.361	0.483	sole occurrence in EGU, large and vital occurrence in FFH protected area
ROD[a]	100	100	0	17	10	8.90	2	1.38	0.169	0.073	0.087	0.675	occurrence with highest compositional genetic differentiation, not endangered
ALB[a]	1500	500	0	20	12	9.36	1	1.89	**-0.513**	0.406	0.268	0.530	stable and vital occurrence in an undisturbed stream valley, vital, phenotypic variability
ELM[a]	200	200	0	29	21	11.64	2	2.03	-0.009	0.174	0.173	0.601	sole occurrence in spring area, stable and vital, low threat
HAU	500	800	200	28	18	10.40	0	1.99	**0.112**	0.292	0.329	0.449	genetically similar to LUD, that showed greater genetic diversity and appeared more vital
RI1[a]	1500	500	500	18	15	12.22	8	2.44	**0.291**	0.219	0.309	0.402	near-natural, stable and vital occurrence located in FFH protected area
WI1	600	300	50	30	18	10.35	0	1.62	**-0.378**	0.286	0.208	0.449	occurrences genetically identical with other occurrences of the same population
WI2[a]	1500	500	500	29	12	8.07	0	1.52	**-0.179**	0.233	0.198	0.449	sole occurrence in a pond, vital
WI3	1000	500	500	30	18	9.91	0	1.87	-0.080	0.253	0.234	0.431	occurrences genetically identical with other occurrences of the same population
BEL[a]	1000	1000	1000	26	20	11.74	0	2.35	0.061	0.333	0.354	0.450	large and stable occurrence in FFH protected area, site management ensured
DUE	100	300	0										low number of individuals, no information on genetics
KNI[a]	1200	1000	1000	14	14	14.00	4	2.99	**0.097**	0.403	0.447	0.323	large and vital occurrence, located in a village and therefore suitable for awareness raising
LUD[a]	500	1000	1000	30	30	14.00	15	1.92	**0.104**	0.249	0.278	0.493	large and stable occurrence, not endangered, site management ensured
Sum				514	323[b]								

The color coding of the occurrences reflects the ecogeographic units (Table 1). The number of individuals counted in 2015 and 2016, as well as the number of sexually reproducing individuals in 2016 is giving. N indicates the number of individuals for which SSR were successfully amplified. The number of multi-locus-genotypes (MLG), the expected number of MLG at the lowest common sample size (eMLG), the number of private alleles, the mean allelic richness, the fixation coefficient F_{IS}, the expected and observed heterozygosity, and the compositional genetic differentiation Δ_j is shown for each occurrence. The last column summarizes the arguments for and against the selection of genetic reserve candidates. F_{IS} values in bold are significantly different from zero (Table S1)

[a]Sites selected as genetic reserve candidates

[b]One MLG is shared between WI1 and WI3

WI2, WI3, and BEL showed no private alleles (Tables 4, and S1c). The mean allelic richness was the highest for LOE and KNI with 3.04 and 2.99, respectively. It was the lowest for RI2 (i.e. 1.12) and EPP (i.e. 1.37).

The fixation index F_{IS} ranged from 0.291 (RI1) to − 0.865 (RI2). An excess of homozygotes (i.e. a positive value significantly different from zero) was obtained for six occurrences (i.e. AAC, POL, HAU, RI1, KNI, and LUD). An excess of heterozygotes (i.e. a negative value significantly different from zero) was found for seven occurrences. The comparison of all values of heterozygosity (Table 4), revealed higher values for the observed heterozygosity (H_{obs}) for LOE, ALB, and KNI (i.e. 0.537, 0.406, and 0.403). Lower values of observed heterozygosity were obtained for EPP, RI2, and ROD (i.e. 0.189, 0.11, and 0.073). The largest difference between observed and expected heterozygosity (H_{exp}) was found for ALB ($H_{obs}-H_{exp}=0.138$, Table S1).

Most populations did not seem to experience a recent bottleneck except three populations (Table S2). BOD and ALS showed signatures of bottlenecks under both models with the one-tailed Wilcoxon rank test for heterozygosity excess ($P=0.0015$ and $P=0.0024$ for BOD and $P=0.0002$ and $P=0.0002$ for ALS under TPM and SMM, respectively). Though less intense, HAU also experienced a recent bottleneck ($P=0.0164$ and $P=0.0164$ under TPM and SMM, respectively). The analysis also detected signature of population increase in ELM under both models with the one-tailed Wilcoxon rank test for heterozygosity deficiency ($P=0.0009$ and $P=0.0003$ under TPM and SMM, respectively).

The mean complementary genetic differentiation Δ_{SD} amounts to 0.488 (Fig. 2). The genetic composition of ROD ($\Delta_j=0.675$) deviates most from its complement. In descending order, also ELM, ALS and EPP, and others differ from the complement. KNI ($\Delta_j=0.323$) and LOE ($\Delta_j=0.290$, Table 4) represent the genetic composition of the pooled remaining occurrences best. The sector radius (Fig. 2) of two neighboring occurrences differs at most by 7.9%. This means that populations differ moderately in their contribution to genetic differentiation.

A DAPC was performed to analyze the genetic structure. The first axis of the scatterplot (Fig. 3) describes 32.7% and the second 16.42% of the variation. The scatterplot reflects somewhat the geographic origin (Fig. 1). EPP and ROD appear genetically most distinct as they form independent clusters apart from the other populations. A cluster is also formed by RI1, RI2, WI1, WI2, and WI3. ALB and HAU from Pfälzerwald overlap with BEL, LUD, and KNI from Vorderpfälzer Tiefland. In the scatterplot, the clusters of BEL and LUD are partially overlapping, while the individuals of KNI form a dispersed and slightly more distinct group. ELM (Pfälzerwald), AAC (Jülicher Börde), POL (Mittelrheinisches Becken), and HUE (Jülicher Börde) form distinct clusters, with the latter appearing close to individuals of the Alzeyer Hügeland. The grouping of the individuals from the Alzeyer Hügelland reflects the geographic proximity: MAI individuals are closer and partially overlapping with individuals from BOD, which group closer to LOE. LOE individuals group closer to ALS. Occurrences overlapping

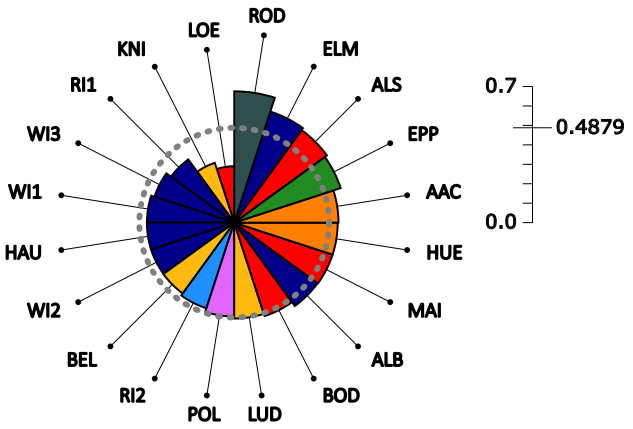


Fig. 2 Diagram of the differentiation between the *H. nodiflorum* occurrences at the gene-pool level. Each sector refers to one occurrence. The grey dashed circle marks the weighted mean of the sector radii (Δ_{SD}) of the pooled occurrences, which refers to the mean genetic differentiation (i.e. the complement). Occurrences are ordered by decreasing sector radius from the most unique to the most representative occurrence. An occurrence sharing no genetic types with the complement attains the maximal radius of $\Delta_{SD(j)}=1$, while an occurrence representing the complement perfectly shows $\Delta_{SD(j)}=0$. The sector angle is equal for the relative size of this occurrence among all occurrences. The color-coding reflects the ecogeographic units (Table 1)

in the scatterplot show admixture in the plot of posterior membership probabilities (Fig. S2).

A significant spatial structure has been detected by the Mantel test (Fig. S3) as the observed value of the correlation between the distance matrices (i.e. 0.42) lays outside of the reference distribution (i.e. under the absence of spatial structure). The scatterplot shows individuals with little genetic and geographic distance (i.e. the occurrences from Pfälzerwald), a smaller (i.e. the occurrences from Jülicher Börde), and a major cloud that are separated from each other. Hence, the genetic distance does not increase linearly with geographic distance (Fig. S4, Table S1d).

Moreover, ancestry coefficients were inferred that represent the proportions of an individual genome that may originate from multiple ancestral gene pools (Frichot et al. 2014). Cross-validation did not exhibit a clear plateau, but a steep slope until K = 7 followed by a slightly more gentle slope until K = 15 where the cross-validation score reaches a plateau and the standard deviations increase (Fig. S5). This is suggesting a coarse population structure with seven main clusters and isolation-by-distance as a factor for genetic differentiation. Smaller values of K are known to show superior estimates, therefore we will focus on K = 7. Ancestry coefficients are shown for different values of K, i.e. for K = 3–11 (Figs. 4 and S6). The seven main clusters are (1) EPP, (2) AAC/HUE, (3) ELM, (4) LUD/BEL, (5) MAI, (6) ALS, and (7) RI1/WI1/RI2/WI3/WI2 (Fig. 4). Other populations are showing signs of admixture from those seven clusters. For example, with K = 7, ROD shows admixture from MAI (ca. 50%), AAC/HUE (ca. 30%), and EPP (ca. 20%), however with K = 11, ROD is assigned to its own cluster (Fig. 4). With K = 7, MAI, BOD, and LOE belong to the same genetic cluster (5), however BOD and especially LOE show signs of admixture from LUD or more likely BEL as suggested by the analysis with K = 8. In contrast to other occurrences of Pfälzerwald, ALB belongs with ca. 50% to cluster (7) but shows some signal of admixture involving the clusters (4), (5), and (6) with ca. 20% each. BEL, HAU, KNI, and LUD share the same cluster (4), however with varying amount of admixture mainly associated with cluster (7). POL shows a strong signal of admixture mainly involving clusters (3) and (5). For K = 3 there is a north–south gradient for the genetic clusters (Fig. S6). The higher number of K, the more genetic clusters are formed in southwestern Rhineland-Palatinate (Fig. 1 and S5). The results are concordant with the DAPC analysis (Fig. S2).

The inference of recent migration rates (m) revealed that most populations have low migrant proportions with the exception of three populations ($m > 0.1$). The mean posterior probability for the migration rates among all populations are shown in Table S3. WI1 has the highest proportion of migrants originating from WI2 ($m = 0.20$). A high mean migration rate was also identified for WI1 to WI3 ($m = 0.13$). Moreover, ca. 17% individuals from LOE were identified as recent migrants coming from HAU. Migration events were always unidirectional.

Discussion

Number of individuals and condition of the occurrences

From 250 occurrences of *H. nodiflorum* that were previously reported, we selected 58 as potential candidates for our study. Only 43 of them could be confirmed during field trips and 20 were successfully genotyped. Some of these occurrences have been known for more than 30 years (Table 1). The fluctuations in individual counts can be a result of e.g., natural variations, succession, flooding, changes in land use, and habitat management. A methodological bias is considered low, as counting was conducted by the same botanists in 2015 and 2016, but for ELM. However, a reason for differences in counts of (sexually reproducing) individuals could lay in different dates when the sites were visited, e.g. LUD was first visited in August 2015 and again in June 2016 (Table S1a). Hence, the developmental status of *H. nodiflorum* may have differed for the recordings at different times. Moreover, clonal propagation may hamper the recognition of individual specimens. Regular monitoring is necessary to be sure about the evolution of population size and to assess

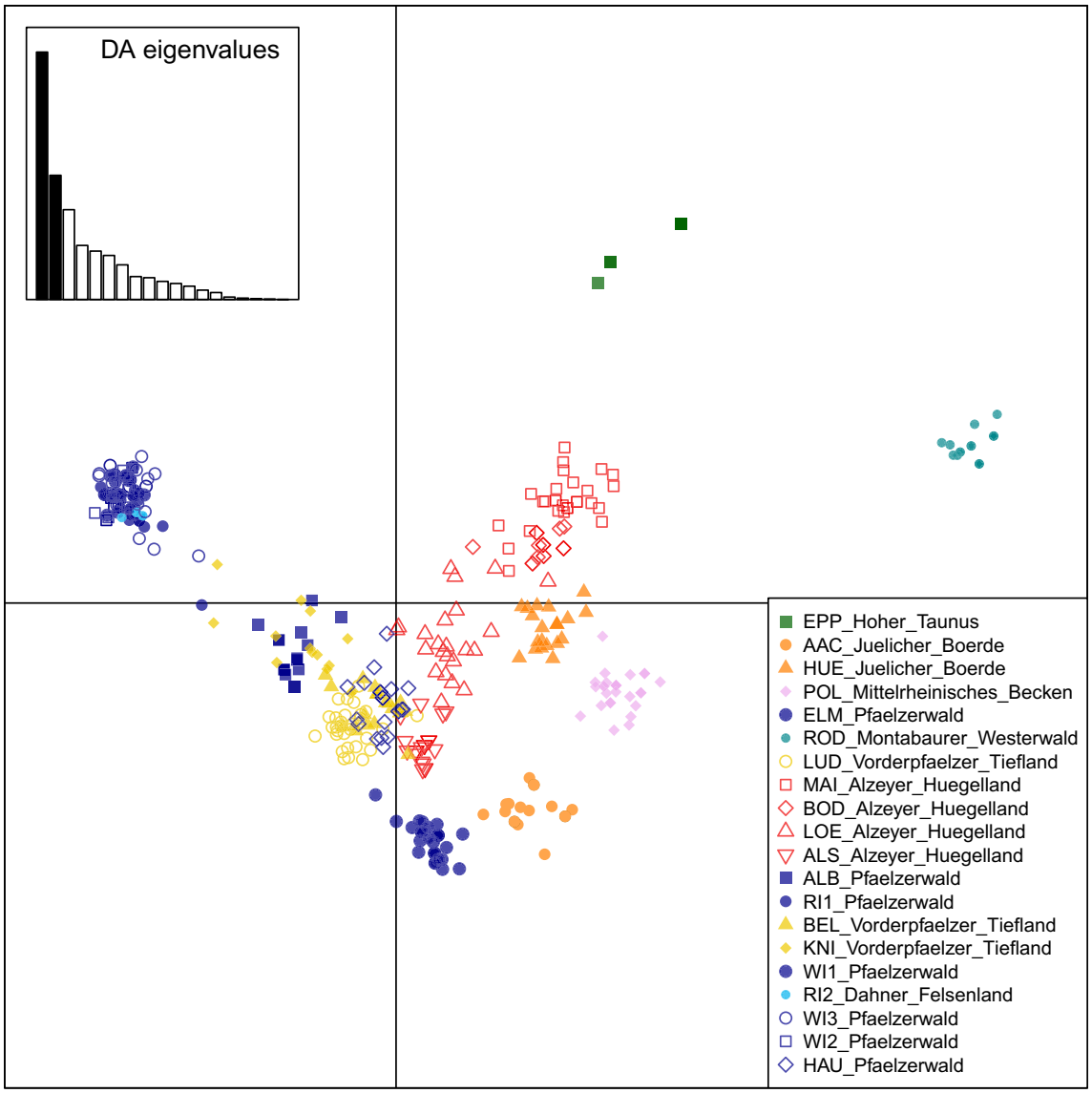


Fig. 3 Scatterplot showing the result of the Discriminant Analysis of Principal Components (DAPC) along the first two principal components. The plot displays the genetic structure of the 20 occurrences of *H. nodiflorum*. Each points refers to an individual. Samples collected at the same sites are displayed with the same symbol. Colors refer to the ecogeographic origin of the samples (Table 1). Eigenvalues of the analysis are displayed in inset

the impact of recent threats/management practices. Botanists noted that especially the progressing succession, e.g. an increase of reeds and woody plants (Table S1a) is a risk to the persistence of the occurrences. When the sites were visited again in 2018, the Spanish slug (*Arion vulgaris*) was identified as a major threat to *H. nodiflorum*. Flooding may represent threats as it can wash away entire occurrences. However, the disturbance through flooding may also create new open spaces that *H. nodiflorum* can colonize when competitive plants were detached.

Nomination of genetic reserves for *H. nodiflorum*

The genetic reserve conservation technique combines habitat protection, ex situ conservation, and the utilization of genetic resources for research and plant breeding (Maxted et al. 2015). In total, 15 sites (i.e. EPP, AAC, HUE, ALS, LOE, MAI, POL, ROD, ALB, ELM, RI1, WI2, BEL KNI, and LUD, Fig. 1, Table 4), were nominated as candidates for the establishment of genetic reserves. The nomination was based on the integration of the different data. Highest value was given to genetic distinctiveness (i.e. results of the DAPC, ancestry coefficients, allelic richness, F_{IS}, etc.) followed by population size and geographical aspects (location



Fig. 4 Barplot representation of ancestry coefficients for K = 3 till K = 11

in an EGU / distance to other occurrences). Moreover, additional knowledge (e.g. the location in a protected area or an already established suitable management has also been considered (Tables 4 and S1). The sites are located in six distinct EGU. In the following, for each EGU, the arguments to nominate a genetic reserve are presented.

Jülicher Börde (AAC and HUE)

The occurrences AAC and HUE are the most northwestern (Fig. 1). The number of individuals was low for both sites but increased from 2015 to 2016. Most individuals were sexually reproducing (Table 4). The DAPC returned distinct clusters for each occurrence (Fig. 3). However, individuals were assigned to the same ancestral genetic clusters (Figs. 4, and S5). This may indicate that both occurrences share a common ancestry, but they are genetically diverging. Differences between both occurrences are also visible from a higher number of MLG, and 24 private alleles for HUE, while no private alleles were found for AAC (Table 4). The significant positive F_{IS}-value indicates that AAC may have been impacted by fixation. Interestingly, HUE experienced a population expansion from 2015 to 2016 which continued in 2018 (data not shown). This is in contrast to AAC where the number of individuals stagnated. This may reflect a higher fitness of HUE compared to AAC. As the occurrences are somewhat genetically distinct and geographically remote, we recommend that both should be considered genetic reserves. In case of limited resources, HUE should be treated with priority. However, concerning AAC local stakeholders have already been identified that are willing to support the establishment of a genetic reserve.

Mittelrheinisches Becken and Montabaurer Westerwald (POL and ROD)

The occurrences POL and ROD are located in the north of Rhineland-Palatinate in different EGU. In the DAPC, POL and ROD form distinct clusters. ROD is the cluster most distinct from other *H. nodiflorum* locations (Figs. 2 and 3). ROD and POL are related but show traces of admixture with different genetic clusters: ROD shares strong genetic signatures with EPP, MAI/BOD/LOE, and AAC/HUE or their progenitor. POL shares a genetic signal with ELM and MAI/BOD/LOE (Fig. 4). Although the number of individuals and sexually reproducing individuals was high for POL (Table S1a), fixation was detected for this site (Table 4). POL also showed considerably more MLG, private alleles, and has a high allelic richness (Table 4).

No flowering individuals were recorded for ROD, which could indicate a trend toward clonal propagation, although the fixation index was not found significant (Table 4). The site was visited on June 27, 2016. At this time, flowering individuals had already been found for most of the occurrences (Table S1). As the populations show distinct genetic patterns and originate from different EGU, they should both be candidates for the establishment of genetic reserves.

Hoher Tanus (EPP)

EPP is the only genotyped occurrence from Hoher Taunus. Individual estimates were low, but one-third were flowering and the occurrences showed an excess of heterozygotes. Only three MLG and no private alleles have been detected, the allelic richness was low (Fig. S1, Table 4). EPP is one of the samples that deviate most from the complement (Fig. 2). It forms a separate cluster in the DAPC and belongs to a single ancestral genetic cluster in the analysis of ancestry proportions (Fig. 4). It does not seem to have experienced gene flow. However, EPP shares genetic signatures with AAC/HUE (i.e. K = 3–5) and ROD (Figs. 4 and S6), which could indicate the involvement of EPP as donor in hybridizations. EPP should become a genetic reserve as it is genetically distinct and the only occurrence suitable as a genetic reserve in this EGU.

Alzeyer Hügelland (MAI, BOD, LOE, and ALS)

Across the occurrences MAI, BOD, LOE, and ALS of the Alzeyer Hügelland, all individuals were sexually reproducing with moderately large populations, but for BOD. *Heloscciadium nodiflorum* is known to occur at BOD only since 2012. Most individuals of BOD appear to be admixted with MAI to a low degree (Fig. S2). This, in combination with the low genetic diversity of BOD (Fig. S1, Table 4), may be the result of a founder effect, where only a few individuals of MAI (Fig. S2) with a limited subset of alleles have recently started to populate this site. A founder effect in BOD is also in accordance with the signature of a recent and strong bottleneck (Table S2). Therefore, we do not nominate BOD as a genetic reserve. Although ALS was also impacted by a recent bottleneck, it should be nominated a genetic reserve candidate as it belongs to an independent genetic cluster when K > 4 (Figs. 4 and S5). It also had the highest number of private alleles. The ancestry coefficients support a common ancestry of MAI, BOD, and LOE as the occurrences belong to a single ancestral genetic cluster until K = 8. MAI had the highest number of MLG (26) and 22 private alleles. The mean allelic richness is the highest for LOE. Moreover, LOE is the occurrence with the smallest value for the compositional genetic differentiation (Δ_j), which means that it is reflecting the complement of the other 19 occurrences the best (Fig. 2). A reason for this could also be that LOE recently received migrants from HAU (Table S3 and Fig. S2). ALS and LOE show an excess of heterozygotes. This could be explained by the recent gene flow detected by the BayesAss analysis. ALS, MAI, and LOE should be chosen for the establishment of genetic reserves.

Vorderpfälzer Tiefland (BEL, KNI and LUD)

All genotyped occurrences of the Vorderpfälzer Tiefland seem vital regarding the number of individuals (Table 4). LUD is the occurrence with the highest number of MLG among all occurrences of this study and had the most private alleles of this EGU. KNI has the second highest allelic richness of the study. However, fixation was detected for this site. BEL, KNI, LUD, and HAU show related genetic signatures until K = 7 (Figs. 4 and S2). For K > 7, LUD was assigned to a genetic cluster that is different from BEL and KNI. The latter showed varying traces of hybridization/introgression. BEL shows signals of admixture with ALB (Fig. 3, K = 8–9, Fig. 4). Due to distinct genetic characteristics, BEL, KNI, and LUD should become genetic reserves. Moreover, as KNI is the only occurrence located in a village, a genetic reserve in this location would raise public awareness.

Pfälzerwald and Dahner Felsenland (ALB, ELM, HAU, WI1, WI2, WI3, RI1, and RI2)

The occurrences WI1, WI2, WI3, RI1 are located in close proximity along the same streamlet Freischbach. The streamlet flows into the small stream Wellbach, on whose lower course RI2 is located. Genetic analyses suggest that W1, W2, W3 and RI2 belong to the same metapopulation (Figs. 3, 4 and S5) with ongoing gene flow through migrating individuals (Table S3). RI2, WI1 and WI2 showed an excess of heterozygotes, which can be an indicator that the

population is vital. We recommend the establishment of a genetic reserve for WI2. It shows the highest individual numbers of the occurrences and is the only occurrence in a pond.

RI1 is somewhat genetically distinct from the other occurrences located at the Freischbach. Although, the fixation index points to a negative impact due to inbreeding and/or genetic drift, we argue that the high number of private alleles, the high allelic richness, and the location of RI1 in a Natura 2000 protected area, are suitable arguments to nominate RI1 as a genetic reserve. Although RI2 is the only occurrences of the Dahner Felsenland, we did not nominate it as a genetic reserve candidate. RI2 does not show genetically distinct features. Moreover, the census revealed very low individual numbers. Hence, a long persistence of the occurrence may be less likely, which violates a criterion for the selection of genetic reserves (Iriondo et al. 2012).

The occurrences further north (ELM, ALB) and south (HAU) of the Freischbach show a moderate number of MLG, but no or one to two private alleles (Table 4). For ALB and ELM no sexually reproducing individuals were found. As HAU is genetically similar to LUD (Figs. 3, 4, and S5) and fixation was detected, it is not considered a genetic reserve candidate. Moreover, HAU seems to have experienced a recent bottleneck (Table S2). ELM seems genetically distinct (Figs. 2, 3, 4, and S5). ELM is the only population that showed a significant heterozygosity deficiency indicative of population expansion (Table S2). There are signals of introgression between ALB and BEL (i.e. K = 8 and K = 9, Fig. 4). Especially ELM and ALB are good genetic reserve candidates as they are stable populations with special habitat conditions: ELM is located in a spring area and ALB in an undisturbed small stream.

Importance of genetic analysis for the selection of genetic reserves

The geographical distribution of genetic diversity that is revealed by genetic methods, is a result of past and ongoing events (Stoll et al. 2020). Generally, population differentiation can be caused by gene flow, genetic drift, or local adaptation (Shah et al. 2020; Wambugu and Henry 2022). Gene flow can, for example, occur via intraspecific hybridization (Stebbins 1950), which is defined as a "successful mating between individuals from well-differentiated populations originally isolated from one another and consequently with different gene frequencies" (Culley and Hardiman 2009). Therefore, hybridization can reshape the genetic composition of populations and create novel combinations of genes that allow adaptation to changing environments (Anderson et al. 2006; Anderson and Stebbins 1954; Arnold 1997; Stebbins 1950). Yet, there is no agreement on robust indicators modeling patterns of genetic diversity in time and space. However, allelic diversity has been proposed as a suitable candidate (Pereira et al. 2013; Schmeller et al. 2018; Wambugu and Henry 2022).

In our study, genotyping was essential to decide which occurrences are well-suited to establish genetic reserves. The genetic data were used to evaluate hypotheses on past and ongoing forces acting on the populations: The SSR data revealed that the genetic patterns are in part a result of isolation-by-distance (Figs. S3, S4, and S5). The pattern in the cross-entropy validation score (Fig. S4) suggests a coarse population structure with seven clusters but additional isolation-by-distance as a factor for genetic differentiation. The estimation of population genetic structure allowed us to identify populations with unique genetic patterns that may be a result of local adaptations. Hence, these occurrences should receive priority for in situ conservation. This also includes several populations that showed an exclusive signal of past gene flow via hybridizations combining genetic signals of the different ancestral genetic clusters (Figs. 4, and S2). Moreover, we identified occurrences that belonged to the same metapopulation that are connected via recent migration events (Table S3). Hence, it is sufficient if, initially, only one or a few of these occurrences are nominated as genetic reserves.

We identified seven populations that showed an excess of heterozygotes. This could be a temporary pattern as the result of the mixing of two previously isolated populations. However, it could indicate a progressive selection against deleterious recessive alleles, asexual reproduction (Ruiz Mondragon et al. 2022) or self-incompatibility (Stoeckel et al. 2006). Furthermore, six populations that experienced fixation were identified. Often, these populations were small in population size and exposed to threats. Small populations tend to inbreeding via mating of relatives. However, they also have a tendency to fix alleles via drift, which results also in among-population inbreeding characterized by a reduced fitness (Keller and Waller 2002). It has been shown that plant populations that experience inbreeding show negative effects related to seed set, germination, survival, and stress resistance. However, it depends on the degree and the genetic basis of inbreeding depression if the risk of population extinction is increased (Charlesworth and Charlesworth 1999; Keller and Waller 2002; Mable 2019). If the genetic diversity of such populations is distinct, it may still be worthwhile to nominate them as genetic reserves. It has been shown that a change in management can result in the recovery of such populations. For example, an occurrence of *Apium graveolens* that was believed to be extinct in the year 2004 could recover within approximately 15 years after the management of the site was changed (Hannappel and Bär 2021). The precondition for this is that stakeholders willingly support the occurrences. The genetic data can then present the baseline for genetic monitoring so that changes

in allele frequencies can be detected. If a loss in allele frequencies is recorded the management of the sites may be adjusted. Genetic monitoring may be able to reveal informative results every 10–20 years. For example, a loss in genetic allelic diversity was detected after 28 years in wheat wild relatives (Nevo et al. 2012).

Once established, monitoring within the genetic reserves should be conducted regularly (Wambugu and Henry 2022). It is needed to be sure about trends in population size. In the case of *H. nodiflorum*, this monitoring is needed to assess if the occurrences persisted the severe flooding in North Rhine-Westphalia and Rhineland-Palatinate from July 2021. Raging rivers could have carried away individuals or entire occurrences. It is unclear whether they were reestablished in other places.

Due to financial and time constraints, genetic analysis may not always be possible. Then, the knowledge about different EGU, the number of individuals at the site, and a trend in population size if known can be used as a proxy (Parra-Quijano et al. 2012) to choose the best genetic reserve candidates.

Conclusion

In times of climate change and intensification in land use, the protection of genetic resources is more important than ever. Traits held by CWR that can make our crops more robust against different kinds of stress need to be conserved. The choice of CWR populations that should be given priority for the establishment of genetic reserves is crucial. Whenever possible, the decision should be based on genetic data to ensure that the intraspecific diversity of the taxon under consideration is covered.

We have genotyped 20 occurrences of *H. nodiflorum* from Germany using 17 newly developed SSR. The SSR have proven to be informative to compare the genetic diversity of the occurrences. Most occurrences were genetically distinct. Isolation-by-distance was found to explain in part the detected genetic variations. Moreover, several populations showed traces of past hybridization and recent migration events, while others had signatures of recent bottlenecks. We identified occurrences that belong to the same metapopulation and we detected populations that experienced fixation.

Finally, based on the results from the genetic analyses, the size of the occurrences, the geographic location and the support of local stakeholders regarding the management of the occurrences, we nominate 15 occurrences for the establishment of genetic reserves for *H. nodiflorum*. The genetic reserves shall become part of the Wild Celery Network (https://netzwerk-wildsellerie.julius-kuehn.de/).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10592-023-01554-x.

Acknowledgements We thank the federal environmental agencies in Germany for providing records on *H. nodiflorum* observations and several botanists for gathering valuable information in the field. KEW Garden is acknowledged for accessions 200497 from Jordan. We are grateful to Lothar Frese who initiated and shaped the project and to anonymous reviewers for critical comments on earlier versions of this manuscript.

Author contributions IM and NB wrote the initial version of the manuscript. NB conducted the SSR data analysis supported by JB, MN, IM, AN, and US. MB coordinated the study and organized the fieldwork together with MZ. CU supervised the study. All authors approved the final version of the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by funds from the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Office for Agriculture and Food (BLE), grant number 2814BM110-112.

Data availability The SSR data sets with and without null alleles set are freely available from the OpenAgrar: https://www.openagrar.de/receive/openagrar_mods_00073465.

Declarations

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

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

Amos W, Hoffmann JI, Frodsham A, Zhang L, Best S, Hill AVS (2007) Automated binning of microsatellite alleles: problems and solutions. Mol Ecol Notes 7:10–14. https://doi.org/10.1111/j.1471-8286.2006.01560.x

Anderson E, Stebbins GL (1954) Hybridization as an evolutionary stimulus. Evolution 8:378. https://doi.org/10.2307/2405784

Anderson NO, Galatowitsch SM, Gomez N (2006) Selection strategies to reduce invasive potential in introduced plants. Euphytica 148:203–216. https://doi.org/10.1007/s10681-006-5951-7

Arnold ML (1997) Natural hybridization and evolution. Oxford University Press on, New York

Bilz M, Kell S, Maxted N, Lansdown RV (eds) (2011) European red list of vascular plants. Publications Office of the European Union

BLE (2012) Nationale Fachprogramm zur Erhaltung und nachhaltigen Nutzung pflanzengenetischer Ressourcen landwirtschaftlicher und

gartenbaulicher Kulturpflanzen. https://genres.de/fachportale/kultur-und-wildpflanzen/nationales-fachprogramm/
Bönisch M, Frese L (2020) Designation of genetic reserves for wild celery species in Germany. Crop Wild Relative 12:4–7
Bönisch M, Nachtigall M, Bernhardt N, Herden T, Zander, M. Friesen N, Frese L (2020) Genetische Erhaltungsgebiete für Wildsellerie-Arten (*Apium* und *Helosciadium*) als Bestandteil eines Netzwerkes genetischer Erhaltungsgebiete in Deutschland—Abschlussbericht. https://service.ble.de/ptdb/index2.php?detail_id=45943&ssk=673b20eafa&site_key=141&stichw=sellerie&zeilenzahl_zaehler=7#newContent
Bönisch M, Thormann I (2021) Wildpflanzen für Ernährung und Landwirtschaft - Sicherung in genetischen Erhaltungsgebieten. Natur Und Landschaftsplanung 96:461–467
Bundesamt für Naturschutz (2017) FloraWeb—Daten und Informationen zu Wildpflanzen und zur Vegetation Deutschlands Accessed 16 Feb 2017
Caye K, François O (2016) tess3r: inference of spatial population genetic structure. R package version 1.1.0. https://bcm-uga.github.io/TESS3_encho_sen/index.html
Caye K, Deist TM, Martins H, Michel O, François O (2016) TESS3: fast inference of spatial population structure and genome scans for selection. Mol Ecol Resour 16:540–548. https://doi.org/10.1111/1755-0998.12471
Charlesworth B, Charlesworth D (1999) The genetic basis of inbreeding depression. Genet Res 74:329–340. https://doi.org/10.1017/S0016672399004152
Culley TM, Hardiman NA (2009) The role of intraspecific hybridization in the evolution of invasiveness: a case study of the ornamental pear tree *Pyrus calleryana*. Biol Invasions 11:1107–1119. https://doi.org/10.1007/s10530-008-9386-z
Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. PNAS 91:3166–3170. https://doi.org/10.1073/pnas.91.8.3166
Dray S, Dufour A-B (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Soft 22:1–20. https://doi.org/10.1837/jss.v022.i04
Dwivedi SL, Upadhyaya HD, Stalker HT, Blair MW, Bertioli DJ, Nielen S, Ortiz R (2009) Plant Bred Rev 30
El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree *Argania spinosa* (L.) Skeels endemic to Morocco. Theor Appl Genet 92:832–839
Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:e19379. https://doi.org/10.1371/journal.pone.0019379
European Council (1992) European Council Directive 92/43/EEC of 21 May 1992 on the Conservation of Natural Habitats and of Wild Fauna and Flora: Accessed 14 Sep 2022 https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31992L0043&from=EN
Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. Mol Ecol 16:1149–1166. https://doi.org/10.1111/j.1365-294X.2007.03218.x
Fievet V, Touzet P, Arnaud J-F, Cuguen J (2007) Spatial analysis of nuclear and cytoplasmic DNA diversity in wild sea beet (*Beta vulgaris* ssp. *maritima*) populations: do marine currents shape the genetic structure? Mol Ecol 16:1847–1864. https://doi.org/10.1111/j.1365-294X.2006.03208.x
Food and Agriculture Organisation (2001) International Treaty on Plant Genetic Resources for Food and Agriculture: Accessed 16 Feb 2017 http://www.planttreaty.org
Frese L, Bönisch M, Herden T, Zander M, Friesen N (2018a) In-situ-Erhaltung von Wildselleriearten. Ein Beispiel für genetische Ressourcen von Wildlebenden Verwandten von Kulturarten (WVK-Arten). Naturschutz Und Landschaftsplanung 50:155
Frese L, Bönisch M, Nachtigall M, Schirmak U (2018b) Patterns of genetic diversity and implications for in situ conservation of wild celery (*Apium graveolens* L. ssp. *graveolens*). Agriculture 8:129. https://doi.org/10.3390/agriculture8090129
Frese L, Nachtigall M, Iriondo JM, Rubio Teso ML, Duarte MC, Pinheiro de Carvalho MÂA (2019) Genetic diversity and differentiation in *Patellifolia* (Amaranthaceae) in the Macaronesian archipelagos and the Iberian Peninsula and implications for genetic conservation programmes. Genet Resour Crop Evol 66:225–241. https://doi.org/10.1007/s10722-018-0708-4
Frichot E, Mathieu F, Trouillon T, Bouchard G, François O (2014) Fast and efficient estimation of individual ancestry coefficients. Genetics 196:973–983. https://doi.org/10.1534/genetics.113.160572
Gillet EM (2013) DifferInt: compositional differentiation among populations at three levels of genetic integration. Mol Ecol Resour 13:953–964. https://doi.org/10.1111/1755-0998.12145
Gong H, Rehman F, Ma Y, Zeng S, Yang T, Huang J, Li Z, Wu D, Wang Y (2022) Germplasm resources and strategy for genetic breeding of *Lycium* species: a review. Front Plant Sci 13:802936. https://doi.org/10.3389/fpls.2022.802936
Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J. Hered., 86(6), 485–486.
Gradl E, Lehmair TA, Poschlod P, Reisch C (2022) Building up a network of genetic conservation areas—A comprehensive approach to select target sites for the preservation of genetic variation in wild plant species. Biol Conserv. https://doi.org/10.1016/j.biocon.2022.109748
Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ (2011) Current trends in microsatellite genotyping. Mol Ecol Resour 11:591–611. https://doi.org/10.1111/j.1755-0998.2011.03014.x
Guijarro-Real C, Ribes-Moya AM, Fita AM, Prohens J, Rodriguez-Burruezo A (2015) Volatile fraction of fool's watercress (*Apium nodiflorum*) as a new spice herb and ingredient for salads. Bull UASVM Hortic. https://doi.org/10.15835/buasvmcn-hort:11645
Guijarro-Real C, Prohens J, Rodriguez-Burruezo A, Adalid-Martínez AM, López-Gresa MP, Fita A (2019) Wild edible fool's watercress, a potential crop with high nutraceutical properties. PeerJ 7:e6296. https://doi.org/10.7717/peerj.6296
Hannappel A, Bär K (2021) Erhaltungsgebiete für Wildsellerie. In: Magistrat der Stadt Nidda (Ed.), Salzquellen und Salzpflanzen, Nidda, Bad Salzhausen
Herden T, Bönisch M, Friesen N (2020) Genetic diversity of *Helosciadium repens* (Jacq.) W.D.J. Koch (*Apiaceae*) in Germany a crop wild relative of celery. Ecol Evol 10:875–890. https://doi.org/10.1002/ece3.5947
Hodel RGJ, Segovia-Salcedo MC, Landis JB, Crowl AA, Sun M, Liu X, Gitzendanner MA, Douglas NA, Germain-Aubrey CC, Chen S, Soltis DE, Soltis PS (2016) The report of my death was an exaggeration: a review for researchers using microsatellites in the 21st century. Appl Plant Sci 4:1600025. https://doi.org/10.3732/apps.1600025
Iriondo JM, Maxted N, Kell S, Ford-Lloyd BV, Lara-Romero C, Labokas J, Magos Brehm J (2012) Quality standards for genetic reserve conservation of crop wild relatives. In: Maxted N, Dulloo ME, Ford-Lloyd BV, Frese L, Iriondo J, Miguel AA, de Carvalho P (eds) Agrobiodiversity conservation securing the diversity of crop wild relatives and landraces. CABI, Wallingford, pp 72–77
Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart T, Ahmed I (2011) Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281. https://doi.org/10.7717/peerj.281

Kamvar ZN, Brooks JC, Grünwald NJ (2015) Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Front Genet 6:208. https://doi.org/10.3389/fgene.2015.00208

Kara Y (2014) The uptake of nickel and copper from water by *Apium nodifolium* (L.) Lag. Biosci Res 11:20–24

Kassambara A (2023) Ggpubr:"ggplot2" based publication ready plots: R package version 0.6.0, https://rpkgs.datanovia.com/ggpubr/.

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. Trends Ecol Evol 17:230–241. https://doi.org/10.1016/S0169-5347(02)02489-8

Khoury CK, Brush S, Costich DE, Curry HA, de Haan S, Engels JMM, Guarino L, Hoban S, Mercer KL, Miller AJ, Nabhan GP, Perales HR, Richards C, Riggins C, Thormann I (2022) Crop genetic erosion: understanding and responding to loss of crop diversity. New Phytol 233:84–118. https://doi.org/10.1111/nph.17733

Knuth P (1898) Handbuch der Blütenbiologie, Leipzig

Kole C (ed) (2011) Wild crop relatives: genomic and breeding resources: cereals. Springer, Berlin

Leo J, Bengtsson T, Carlsson AS, Brassac J, von Bothmer R (2022) Population structure and phylogeography of *Elymus mutabilis* and its genetic relationships with *E. transbaicalensis* (*Poaceae*). Nord J Bot. https://doi.org/10.1111/njb.03520

Mable BK (2019) Conservation of adaptive potential and functional diversity: integrating old and new approaches. Conserv Genet 20:89–100. https://doi.org/10.1007/s10592-018-1129-9

Maggi F, Giuliani C, Fico G, Ricciutelli M, Bramucci M, Quassinti L, Petrelli D, Vitali LA, Cianfaglione K, Tirillini B, Sut S, Dall'Acqua S (2019) Secondary metabolites, secretory structures and biological activity of water celery (*Apium nodiflorum* (L.) Lag.) growing in central Italy. Plant Biosyst 153:325–335. https://doi.org/10.1080/11263504.2018.1492992

Maxted N, Hawkes JG, Ford-Lloyd BV, Williams JT (1997) A practical model for in situ genetic conservation. In: Maxted N, Ford-Lloyd BV, Hawkes JG (eds) Plant genetic conservation: the in situ approach. Kluwer Academic Publishers, Dordrecht, pp 339–364

Maxted N, Avagyan A, Frese L, Iriondo J, Brehm JM, Singer A, Kell S (2015) ECPGR Concept for in situ conservation of crop wild relatives in Europe. http://www.cropwildrelatives.org/fileadmin/templates/cropwildrelatives.org/upload/documents/portal_content/ecpgr_concept_for_in__situ_conservation_of_cwr_in_europe_maxtedetal_2015.pdf

Meirmans PG (2014) Nonconvergence in Bayesian estimation of migration rates. Mol Ecol Res 14:726–733. https://doi.org/10.1111/1755-0998.12216

Meirmans PG, Hedrick PW (2011) Assessing population structure: F(ST) and related measures. Mol Ecol Resour 11:5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x

Metzing D, Garve E, Matzke-Hajek G (2018) Rote Liste und Gesamtartenliste der Farn-und Blütenpflanzen (Trachaeophyta) Deutschlands. Naturschutz Und Biologische Vielfalt 70(7):13–358

Meynen E, Schmithüsen J (1953) Handbuch der Naturräumlichen Gliederung Deutschlands/Unter Mitwirkung des Zentralausschusses für Deutsche Landeskunde

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Ann Hum Genet 47:253–259. https://doi.org/10.1111/j.1469-1809.1983.tb00993.x

Nevo E, Fu Y-B, Pavlicek T, Khalifa S, Tavasi M, Beiles A (2012) Evolution of wild cereals during 28 years of global warming in Israel. PNAS 109:3412–3415. https://doi.org/10.1073/pnas.1121411109

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet Res 22:201–204. https://doi.org/10.1017/S0016672300012994

Parra-Quijano M, Iriondo JM, Frese L, Torres E (2012) In situ conservation of crop wild relatives: a strategy for identifying priority genetic reserves sites. In: Maxted N, Dulloo ME, Ford-Lloyd BV, Frese L, Iriondo J, Miguel AA, de Carvalho P (eds) Agrobiodiversity conservation securing the diversity of crop wild relatives and landraces. CABI, Wallingford, pp 20–28

Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J, Gregory RD, Heip C, Höft R, Hurtt G, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M (2013) Ecology. Essential biodiversity variables. Science 339:277–278. https://doi.org/10.1126/science.1229931

Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a program for detecting recent effective population size reductions from allele data frequencies. J Hered 90(4):502–503

Poland JA, Brown PJ, Sorrells ME, Jannink J-L (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7:e32253. https://doi.org/10.1371/journal.pone.0032253

R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA (2018) Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol 67:901–904. https://doi.org/10.1093/sysbio/syy032

Ruiz Mondragon KY, Aguirre-Planter E, Gasca-Pineda J, Klimova A, Trejo-Salazar R-E, Reyes Guerra MA, Medellin RA, Piñero D, Lira R, Eguiarte LE (2022) Conservation genomics of *Agave tequilana* Weber var. azul: low genetic differentiation and heterozygote excess in the tequila agave from Jalisco Mexico. Peer J 10:e14398. https://doi.org/10.7717/peerj.14398

Šarhanová P, Pfanzelt S, Brandt R, Himmelbach A, Blattner FR (2018) SSR-seq: Genotyping of microsatellites using nextgeneration sequencing reveals higher level of polymorphism as compared to traditional fragment size scoring. Ecol. Evol. 8:10817–10833

Schmeller DS, Weatherdon LV, Loyau A, Bondeau A, Brotons L, Brummitt N, Geijzendorffer IR, Haase P, Kuemmerlen M, Martin CS, Mihoub J-B, Rocchini D, Saarenmaa H, Stoll S, Regan EC (2018) A suite of essential biodiversity variables for detecting critical biodiversity change. Biol Rev 93:55–71. https://doi.org/10.1111/brv.12332

Shah N, Wakabayashi T, Kawamura Y, Skovbjerg CK, Wang M-Z, Mustamin Y, Isomura Y, Gupta V, Jin H, Mun T, Sandal N, Azuma F, Fukai E, Seren Ü, Kusakabe S, Kikuchi Y, Nitanda S, Kumaki T, Hashiguchi M, Tanaka H, Hayashi A, Sønderkær M, Nielsen KL, Schneeberger K, Vilhjalmsson B, Akashi R, Stougaard J, Sato S, Schierup MH, Andersen SU (2020) Extreme genetic signatures of local adaptation during *Lotus japonicus* colonization of Japan. Nat Commun 11:253. https://doi.org/10.1038/s41467-019-14213-y

Stebbins GL (ed) (1950) Variation and evolution in plants. Columbia University Press, New York

Stoeckel S, Grange J, Fernández-Manjarres JF, Bilger I, Frascaria-Lacoste N, Mariette S (2006) Heterozygote excess in a self-incompatible and partially clonal forest tree species—*Prunus avium* L. Mol Ecol 15:2109–2118. https://doi.org/10.1111/j.1365-294X.2006.02926.x

Stoll A, Harpke D, Schütte C, Jimenez L, Letelier L, Blattner FR, Quandt D (2020) Landscape genetics of the endangered Atacama Desert shrub *Balsamocarpon brevifolium* in the context of habitat

fragmentation. Glob Planet Change 184:103059. https://doi.org/10.1016/j.gloplacha.2019.103059
Trumble JT, Dercks W, Quiros CF, Beier RC (1990) Host plant resistance and linear furanocoumarin content of *Apium* accessions. J Econ Entomol 83:519–525. https://doi.org/10.1093/jee/83.2.519
Trumble JT, Diawara MM, Quiros CF (1998) Breeding resistance in *Apium graveolens* to *Liriomyza trifolii*: antibioss and linear furanocoumarin content. Acta Horitc. https://doi.org/10.17660/ActaHortic.1998.513.2
Venables WN, Ripley BD (2002) Statistics and computing: modern applied statistics with S: Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York.
Vincent H, Hole D, Maxted N (2022) Congruence between global crop wild relative hotspots and biodiversity hotspots. Biol Conserv 265:109432. https://doi.org/10.1016/j.biocon.2021.109432
von Kohn C, Conrad K, Kramer M, Pooler M (2018) Genetic diversity of *Magnolia ashei* characterized by SSR markers. Conserv Genet 19:923–936. https://doi.org/10.1007/s10592-018-1065-8
Wambugu PW, Henry R (2022) Supporting in situ conservation of the genetic diversity of crop wild relatives using genomic technologies. Mol Ecol 31:2207–2222. https://doi.org/10.1111/mec.16402
Wehling P, Scholz M, Ruge-Wehling B, Hackauf B, Frese L (2017) Anpassung landwirtschaftlicher Kulturarten an den Klimawandel—Optionen aus Sicht der Züchtungsforschung. JFK 69:47–50. https://doi.org/10.5073/JfK.2017.02.04
Wickham H (2016) Ggplot2: elegant graphics for data analysis. Springer, New York
Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. Genetics 163:1177–1191. https://doi.org/10.1093/genetics/163.3.1177
Žerdoner Čalasan A, Hurka H, German DA, Pfanzelt S, Blattner FR, Seidl A, Neuffer B (2021) Pleistocene dynamics of the Eurasian steppe as a driving force of evolution: phylogenetic history of the genus *Capsella* (*Brassicaceae*). Ecol Evol 11:12697–12713. https://doi.org/10.1002/ece3.8015

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