Genetic status of the endangered plant species *Gladiolus palustris* in the western part of its distribution area

Laura Daco^{1,2,3} · Tiphaine Maurice^{1,4} · Serge Muller^{4,5} · Julie Rossa^{1,4} · Guy Colling¹

Received: 21 November 2018 / Accepted: 9 August 2019 / Published online: 24 August 2019
© The Author(s) 2019

Abstract

Many endangered plant species suffer from the effects of ongoing fragmentation of their populations leading to a loss of genetic diversity through genetic drift. In addition, populations of rare plants may also be affected by hybridization with other plant species. We studied the genetic population structure of populations of the endangered *Gladiolus palustris* in the western limit of its distribution area using AFLP markers. In addition, we clarified the taxonomic status of populations where hybridization with the closely related *G. imbricatus* was suspected based on morphology by sequencing the ITS region of the nuclear ribosomal DNA as well as two regions of the chloroplast DNA. Our analysis of the genetic population structure of *G. palustris* revealed a strong differentiation among geographical regions, which was much higher than the differentiation among populations within regions. Most populations retained a considerable amount of genetic variation. To counteract the future loss of genetic diversity through drift, we recommend using the largest populations per region as seed source to increase genetic diversity in genetic depauperate populations of the same region. Our genetic analyses indicated that some French populations of *G. palustris* are of hybrid origin. These hybridization events are likely to be ancient as these populations are very isolated and some parent taxa went regionally extinct. As these hybrid populations may elucidate post-glacial distribution patterns of related *Gladiolus* taxa, we advocate that they deserve the same conservation efforts as purebred ones.

Keywords Conservation · Endangered · Genetic diversity · *Gladiolus palustris* · Plant species hybridization

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10592-019-01213-0>) contains supplementary material, which is available to authorized users.

✉ Laura Daco
ldaco@mnhn.lu

- ¹ Population Biology and Evolution, Musée national d'histoire naturelle, 25 rue Münster, 2160 Luxembourg, Luxembourg
- ² Institut de Botanique, Université de Strasbourg, 28 rue Goethe, 67000 Strasbourg, France
- ³ Fondation faune-flore, 24 rue Münster, 2160 Luxembourg, Luxembourg
- ⁴ Laboratoire Interdisciplinaire des Environnements Continentaux (LIEC), Université de Lorraine, CNRS, UMR 7360, rue du Général Delestraint, 57070 Metz, France
- ⁵ Muséum national d'Histoire naturelle, UMR 7205 ISYEB, CNRS, UPMC, EPHE, Sorbonne Universités, CP 39, 16 rue Buffon, 75005 Paris, France

Introduction

Many endangered plant species are considered to suffer from the effects of ongoing habitat destruction, which has often led to an increased isolation and fragmentation of their populations. As a consequence, populations of rare plants often suffer from a loss of genetic diversity due to the loss of alleles through genetic drift (Young et al. 1996; Frankham and Wilcken 2006; Yuan et al. 2012). Genetic differentiation among populations and the loss of genetic diversity is expected to increase with time since fragmentation (Coates 1988; Gitzendanner and Soltis 2000; Zawko et al. 2001). In Europe, nutrient-poor wet grasslands belong to the most threatened and fragmented habitat types (van Duren et al. 1997; Bakker and Berendse 1999; Oelmann et al. 2009). Many of these grasslands have been drained and fertilized to increase agricultural productivity or have been abandoned. These changes in land use practices have led to the decline and genetic impoverishment of many plant species (Bakker and Berendse 1999; Vergeer et al. 2003; Wesche et al. 2012). Among the plant species that have become rare

through the destruction and fragmentation of wet grasslands is the sword-lily species *Gladiolus palustris* GAUDIN, a Western and Central European species which is declining all across Europe and is considered to be threatened in many countries (Bilz et al. 2011; UICN France et al. 2012). *G. palustris* is legally protected in the European Union (Annex II Habitats Directive 92/43/EEC) and in Switzerland. The closely related *Gladiolus imbricatus* L., considered to be less threatened, has a larger distribution area extending from Central to Eastern Europe (Fig. 1). Morphological distinction of both taxa in the field is difficult without examining the corm, which is often not possible for conservation reasons. At the western limit of its distribution area in France *Gladiolus palustris* occurs in very isolated populations and some populations have been identified morphologically as *G. imbricatus* but their taxonomic status is unclear (Tison and Girod 2014).

Closely related congeneric taxa are likely to hybridize as they often have incomplete reproductive isolation (Mallet 2005, 2008; Soltis and Soltis 2009; Soltis 2013). When the ecological niches of closely related species are similar, they may coexist at certain parts of their distribution area and hybridization can occur in these contact zones especially when both taxa have the same ploidy level (Chapman and Abbott 2010) as it is the case for *G. palustris* and *G. imbricatus* ($2n=4x=60$; Cantor and Tolety 2011). Hybridization and polyploidy strongly influenced the evolution of European *Gladiolus* species (van Raamsdonk and de Vries 1989; Cantor and Tolety 2011; Tison and Girod 2014). It has been recently confirmed by molecular and

morphological analyses that *G. palustris* and *G. imbricatus* naturally hybridize to form the nothospecies *G. × sulistrovicus* (Cieślak et al. 2014; Szczepaniak et al. 2016) in the eastern part of the distribution area of *G. palustris* where the distributions of both taxa partially overlap. Molecular data showed that in the case of *G. × sulistrovicus* hybridization is unidirectional with *G. palustris* as the maternal species and *G. imbricatus* as the pollen donor (Szczepaniak et al. 2016).

Although interspecific hybridization is considered to be an important force of plant evolution (Soltis and Soltis 2009), it also can have detrimental effects on endangered plant taxa through genetic swamping and hybrid backcrossing with one of the parental taxa (Allendorf et al. 2001; Todesco et al. 2016). The discrimination between cryptic hybrids (Gaskin and Kazmer 2009) and purebred taxa has important implications for the delineation of conservation units (Allendorf et al. 2001). The clarification of the taxonomic status of populations of endangered species is important for appropriate conservation measures (Smith and Waterway 2008; Nierbauer et al. 2017). Although hybrid populations may have a certain conservation interest, protective laws often ignore hybrids. It has been argued that hybrid populations are not worth conserving because hybridization events may be due to anthropogenic influences such as the introduction of plants and animals, habitat fragmentation and habitat modification (Allendorf et al. 2001). The use of molecular methods can help to clarify the taxonomic status of populations of closely related rare plant species (Turchetto et al. 2015; Segatto et al. 2017).

Fig. 1 Distribution map and study sites of *G. palustris* and *G. imbricatus* across Central and Western Europe. For population names, see Table 1, regions are in between brackets. Gray circles represent *G. palustris* populations and black circles represent *G. imbricatus* populations. Taxonomic identification was based on morphology. Populations with evidence for hybridization based on genetic analyses are marked with asterisk. Map created with SimpleMappr modified from Meusel et al. (1965) and Szczepaniak et al. (2016)

We studied the genetic population structure of the endangered *G. palustris* in the western part of its distribution area in France and Switzerland using AFLP markers to assess the effects of habitat fragmentation. We wanted to investigate the genetic diversity and conservation status of extant populations as a basis for more appropriate management measures. We furthermore investigated some isolated populations with unclear taxonomic status at the western limit of the distribution area of *G. palustris*. We therefore sequenced the ITS region of the nuclear ribosomal DNA as well as two regions of the chloroplast DNA for a large number of *G. palustris* populations in Western and Central Europe and compared them to populations of the closely related *G. imbricatus* taxa occurring in the same area.

We asked the following questions: (i) does habitat fragmentation affect the genetic structure and diversity of *G. palustris* populations? (ii) are the populations of *G. palustris* at the western distribution limit genetically distinct from core populations in Central Europe and what is their taxonomic status?

Materials and methods

AFLP study of *G. palustris*

Thirteen populations of *G. palustris* located in the western part of the distribution area in France and Switzerland were used to conduct an AFLP study (Table 1). The studied populations represent a large part of the extant populations in the area. At each study site, leaf samples were collected from individual plants along a 20 m long transect. In most populations, leaves from 14 to 16 individuals were sampled (Table 1). The leaf material was placed immediately in separate paper bags and conserved in silica gel for storage until DNA extraction. The minimum distance among sampled plants was one meter in order to avoid sampling clones. Population sizes were estimated as the number of flowering individuals (Table 1).

The AFLP reactions were performed according to the AFLP® Core Reagent Kit (Invitrogen™). Approximately 100 ng of genomic DNA was digested at 37 °C for 2 h with 0.8 µl of the restriction enzymes EcoRI and MseI (1.25 U/µl each) and 2 µl of 5× Reaction Buffer in a final volume of 10 µl. Endonucleases were then inactivated at 70 °C during 15 min.

Adaptor ligation was achieved by adding 9.6 µl of Adapter/Ligation Solution and 0.4 µl of T4 DNA ligase (1 U/µl) to the previous reaction and by incubating the mix for 2 h at 20 °C. After diluting the ligation solution to 2:5, 2 µl were used to perform pre-amplification by adding it to 9.4 µl of AFLP® Pre-amp Primer Mix I (Invitrogen™), 1.2 µl of 10× PCR buffer plus Mg and 0.2 µl of Taq DNA polymerase

(5 U/µl, Thermo Scientific). Polymerase chain reaction was performed with 20 cycles at 94 °C for 30 s, 56 °C for 1 min and 72 °C for 1 min. The solution was then diluted with 115 µl of ddH₂O.

Two out of twelve primer combinations with distinct polymorphic loci were selected for the selective amplification: E-AAG/M-CAC and E-AAC/M-CAC. Amplifications were performed using 5 µl of diluted pre-amplification reaction added to 0.4 µl of dNTPs (10 mM), 2 µl of 10X PCR buffer with (NH₄)₂SO₄, 1.2 µl of MgCl₂ (25 mM), 0.16 µl of Taq polymerase (5U/µl, Thermo Scientific), 1 µl of EcoRI primers (1 µM) and 1 µl MseI primer (5 µM) in a total volume of 20 µl. Amplifications were programmed for 1 cycle at 94 °C for 2 min, 10 cycles of 20 s at 94 °C, 30 s at 66 °C and 2 min at 72 °C, where the annealing temperature was reduced by 1 °C each cycle. Amplification continued for 20 more cycles with the same conditions except for a 56 °C annealing temperature. The final extension step was conducted at 60 °C for 30 min.

Capillary electrophoresis of all samples was performed with 5 µl of a 1:10 dilution of the selective amplification products of AFLP and 1 µl of a 1:4 dilution of ET550-ROX Size Standard (GE Healthcare) on a MegaBACE 500 DNA Analysis System (GE Healthcare) after a denaturation step of 95 °C for 2 min.

Statistical analysis of AFLP data

The fragments amplified by AFLP primers were scored using MegaBACE Fragment profiler v1.2 (GE Healthcare) as either present (1) or absent (0). Fragments with lengths between 60 and 500 base pairs were included in the analysis. Due to their geographic proximity, data from the two populations from Perrignier were pooled together for further statistical analysis resulting in a total of twelve populations. To estimate the error rate of the AFLP analysis we made three replicates from four randomly chosen individuals from four populations. The mean error rate per sample was calculated as the number of errors divided by the total number of phenotypic comparisons within replicated samples (see Paun and Schönschwetter 2012).

Genetic diversity within populations was estimated as Nei's expected gene diversity (*He*) within a population that averages expected heterozygosity of the marker loci (Nei 1987) using the program GenoDive which allows to correct for unknown dosage of alleles in polyploids (Meirmans and Van Tienderen 2004). The fact that *G. palustris* is a tetraploid may affect the estimates of genetic diversity, as polyploids often maintain greater genetic diversity (Meirmans et al. 2018) and the dosage of alleles is not known (Dufresne et al. 2014). However, due to the sheer abundance of anonymous nuclear markers scattered over the entire genome, dominant markers such as AFLP are considered to be a

Table 1 Location of the 32 studied populations of *Gladiolus palustris* and *G. imbricatus* in Western and Central Europe

Country—region	Population	Abbr.	Year	n_{AFLP}	N	Lat (°N)	Long (°E)	Method	He	Haplotype
<i>Gladiolus palustris</i>										
Czech Republic—South Moravia	Hodonín	Hod	2007			48.895	17.097	2,3		A
France—Ain	Mont d'Ain	MtA	2006	15	270	46.138	5.602	1,2,3	0.264	A
France—Ain	Nantua Nord	Nn	2006	14	300	46.163	5.616	1,2,3	0.266	A
France—Ain	Oyonnax	Oy	2006	16	200	46.241	5.707	1,2,3	0.235	A
France—Bas-Rhin	Herbsheim	He	2006	40	640	48.351	7.656	1,2,3	0.198	A
France—Bas-Rhin	Osthouse	Os	2006	19	100	48.379	7.666	1,2,3	0.178	A
France—Haute-Savoie	Les Rosses	Ro	2006	3	10	46.197	6.281	1,3	0.205	A
France—Haute-Savoie	Margencel	Ma	2006	16	40	46.325	6.424	1,2,3	0.231	A
France—Haute-Savoie	Perrignier	Pe	2006	16	270	46.310	6.420	1,2,3	0.218	A
France—Haute-Savoie	Perrignier petite	Pp	2006	10	^a	46.310	6.420	1,2,3	^a	A
France—Haute-Savoie	Reuland	Re	2006	15	6300	46.316	6.403	1,2,3	0.234	A
France—Jura	Légna	Le	2006	16	4000	46.427	5.622	1,2,3	0.258	A
France—Jura	Ranchette	Ra	2006	16	850	46.355	5.823	1,2,3	0.213	A
France—Territoire de Belfort	Felon ^b	Fel	2007			47.739	6.938	2,3		B
Germany—Bavaria	Ried	Ri	2011			49.672	9.142	2,3		A
Hungary—Csongrád	Zákányszék	Za	2008			46.301	19.867	2,3		A
Italy—Lombardy	Como1	CoI	2010			46.111	9.255	2,3		A
Italy—Lombardy	Como2	CoII	2010			45.988	9.226	2,3		A
Italy—Lombardy	Como3	CoIII	2010			46.067	9.329	2,3		A
Italy—Piedmont	Ivrea	Iv	2010			45.481	7.907	2,3		A
Italy—Piedmont	Turin3	TuIII	2010			45.258	7.814	2,3		A
Italy—Trentino	Rovereto	Rov	2010			45.856	11.034	2,3		A
Switzerland—Haute-Savoie ^c	Faverges	Fa	2006	16	200	46.224	6.284	2,3	0.239	A
<i>Gladiolus imbricatus</i>										
Czech Republic—Olomouc	Krahulci L	KraL	2007			49.799	17.377	2,3		E
Czech Republic—Olomouc	Krahulci R	KraR	2007			49.800	17.379	2,3		A
Czech Republic—Olomouc	Miroslav	Mir	2007			50.217	17.132	2,3		D
Czech Republic—Pardubice	Rohozná	Roh	2007			49.810	15.812	2,3		E
France—Drôme	Lus-La-Croix-Haute	LCH	2008			44.670	5.700	2,3		C
France—Puy-de-Dôme	Lezoux	Lez	2008			45.852	3.375	2,3		D
Germany—Saxony	Dauban	Dau	2008			51.290	14.614	2,3		D
Germany—Saxony	Dubrauke	Dub	2008			51.236	14.620	2,3		D
Italy—Piedmont	Turin2	TuII	2010			45.288	7.792	2,3		D

Abbr. population abbreviation, year of collection, n_{AFLP} sample size for the AFLP study, N population size. Applied methods: 1 AFLP analysis, 2 ITS sequencing, 3 cpDNA sequencing. He Nei's expected heterozygosity. Haplotypes of the combined cpDNA consensus sequences

^aPopulation Perrignier petite was pooled together with population Perrignier for the AFLP analysis (see text)

^bPopulation Felon was mentioned as *G. imbricatus* in Tison and Girod 2014

^cDue to geographic proximity, the Swiss population Faverges was integrated into the Haute-Savoie region

powerful source of information to study population genetic structure even in polyploids (see Dufresne et al. 2014). In GenoDive, dominant data such as AFLP are coded as haploids setting the maximum ploidy level to one. To test for the relationship between population size and genetic diversity we correlated population size (log-transformed) and Nei's gene diversity.

The genetic structure of *G. palustris* at the landscape level was studied using a Bayesian clustering method to

infer population structure and assign individuals to population groups, as implemented in STRUCTURE version 2.3.4 (Pritchard et al. 2000) which allows the analysis of dominant data (Falush et al. 2007). We used a model of no population admixture and a model of admixture for the ancestry of the individuals in two different runs without prior information about the regional membership of the populations and assumed that the allele frequencies are correlated within populations. We conducted a series of ten independent runs

for each value of K (the number of clusters) between one and twelve to quantify the amount of variation of the likelihood of each K . We found that a length of the burn-in of 100,000 and Markov chain Monte Carlo (MCMC) of 500,000 was sufficient. Longer burn-in or MCMC did not significantly change the results.

The model choice criterion implemented in STRUCTURE to detect the K most appropriate to describe the data is an estimate of the posterior probability of the data for a given K , $\Pr(X|K)$ (Pritchard et al. 2000). We used an ad hoc quantity based on the second order rate of change of the likelihood function (DK) with respect to K (Evanno et al. 2005) as implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2012) as an indicator of the signal detected by STRUCTURE.

Finally, the ten runs of the STRUCTURE simulations using the model of no population admixture with the most appropriate K were aligned using the Greedy option in CLUMPP V1.1.2 (Jakobsson and Rosenberg 2007). Convergence of the ten replicate runs for the most appropriate K was high as they produced very similar clustering results as shown by the pairwise G' (similarity function) values (> 0.99) for each pair of permuted runs in CLUMPP. The mean membership coefficients were represented as a bar graph using DISTRUCT 1.1 (Rosenberg 2004).

We also conducted a discriminant analysis of principal components (DAPC) with the R package Adegenet version 2.1.1 (Jombart 2008) as DAPC is not based on pre-defined population genetic models and makes no assumptions about Hardy–Weinberg equilibrium, which is potentially an advantage in the analysis of polyploids (see Turchetto et al. 2015; Meirmans et al. 2018). We first set an a priori group number of ten in the DAPC analysis and in a second step reduced the group number to four based on the results of the first analysis.

To represent the overall genetic relationship among individuals a neighbor-net network (Bryant and Moulton 2004) was constructed using SplitsTree version 4.14.6 (Huson and Bryant 2006). We used uncorrected P method to calculate pairwise distances among individuals (Nei and Kumar 2000).

A hierarchical analysis of molecular variance (AMOVA) was used to partition the genetic variability among geographic regions, populations within regions and individuals as implemented in GenAlEx (Excoffier et al. 1992; Peakall and Smouse 2006, 2012; Meirmans 2006). The variance components from the analysis were used to estimate Φ -statistics, which are similar to F -statistics (Excoffier et al. 1992). We tested the association between pairwise Φ_{st} genetic distances among all pairs of populations and geographic distances using a Mantel test, as implemented in GenAlEx (1000 permutations). We estimated the population-specific F_{ST} values using BAYESCAN 2.1 (Foll and

Gaggiotti 2008) with the default settings and used a linear regression to test if there was a relationship between the population-specific F_{ST} values and measures of genetic diversity (H_e) of populations. We expect a strong relationship if genetic differentiation is affected by genetic drift (both current and historical), reflecting a migration–drift disequilibrium (Cox et al. 2011).

ITS and cpDNA study of *G. palustris* and *G. imbricatus*

We collected plant material from 23 populations of *G. palustris* and nine populations of *G. imbricatus* across France, Switzerland, Italy, Hungary, the Czech Republic and Germany (Fig. 1, Table 1). Taxonomic identification in the field was based on morphology and was done in collaboration with local botanists who also helped with the sampling of the plant material. At each study site, leaf samples were collected from two to 42 randomly selected individual plants. To include *Gladiolus communis* subsp. *byzantinus* as a phylogenetic outgroup, seeds from the Andalusian Seed Bank (Accession number BGVA 13452) originating from Mallorca were germinated and cultivated to the seedling stage. Leaf material of five seedlings was sampled and preserved in silica gel until DNA extraction. Genomic DNA was extracted using a DNeasy Plant Mini Kit (Qiagen) starting from approximately 100 mg of fresh leaf material or 10 mg of dried material.

ITS and cpDNA sequencing

Nuclear ITS region

PCR amplification and sequencing of the ITS region was implemented with ITS4 and ITS5 primers (White et al. 1990). The PCR protocol was conducted using approximately 60 ng of genomic DNA, 2.5 μ l of deoxyribonucleotide triphosphates (dNTPs) mix (2 mM), 1 $\times$ Pfu buffer with $MgSO_4$, 0.625 U of native Pfu DNA polymerase (Thermo Scientific®) and 1.75 μ l of each primer (10 μ M), in a total volume of 25 μ l. Cycling conditions consisted of an initial denaturation at 95 °C for 3 min, followed by 35 cycles with denaturation at 95 °C for 1 min, an annealing at 50.8 °C for 1 min and 2 min of elongation at 72 °C, and a final elongation at 72 °C for 7 min. All PCR products were visualized by electrophoresis on 1.5% agarose gels stained with SYBR® Safe DNA Gel Stain (Invitrogen) to check for amplification. The amplified samples were purified using the QIAquick PCR Purification Kit Protocol (Qiagen).

Due to fungal contamination in the populations Nantua Nord and Les Rosses, an annealing temperature gradient was used to maximize PCR product yield to extract the desired PCR product after gel migration. To generate sequences for the samples from population Nantua Nord the genomic DNA

was amplified with an annealing temperature of 52.1 °C in a 50 µl reaction, migrated on a 1.5% agarose gel, purified with a GeneJET™ Gel Extraction Kit (Thermo Fisher Scientific) and then amplified a second time using the same conditions. For population Les Rosses we were not able to generate a clean ITS sequence.

Purified PCR products were sequenced with the dideoxy chain termination method using the Big-Dye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) with 2 µl of purified PCR product, 0.5 µl of Ready Reaction Premix (2.5×), 2 µl of sequencing buffer (5×) and 0.5 µl of primer (5 µM) in a total of 10 µl. Cycling conditions consisted of 1 min at 96 °C and 25 cycles of 10 s at 96 °C, 5 s at 50 °C and 4 min at 60 °C. The samples were precipitated with ethanol and EDTA according to the same kit, resuspended in 10 µl of Hi-Di™ Formamide and run on a 3730xl DNA Analyzer (Applied Biosystems).

cpDNA markers trnQ(UUG)-rpS16x1 and trnV(UAC)x2-ndhC

PCR amplification and sequencing of the cpDNA was carried out using the two primers trnQ(UUG)-rpS16x1 and trnV(UAC)x2-ndhC (Shaw et al. 2007). The PCR protocol was identical to the one used for the ITS amplification except that we used an annealing temperature of 55 °C and a total volume of 50 µl. The amplified samples were purified using the QIAquick Gel Extraction Kit (Qiagen). The purified products were then sequenced using the DYEnamic ET Dye Terminator Cycle Sequencing Kit for MegaBACE DNA Analysis Systems (GE Healthcare). To approximately 20 ng of purified PCR product, we added 1 µl of primer (5 µM) and 4 µl of Sequencing Reagent Premix in a total of 10 µl. Cycling conditions were 25 cycles of 95 °C for 20 s, 50 °C for 15 s and 60 °C for 1 min. Precipitation of the cycle sequencing products was achieved with sodium acetate/EDTA buffer and ethanol according to the DYEnamic ET Dye Terminator Cycle Sequencing Kit for MegaBACE DNA Analysis Systems (GE Healthcare). The samples were then resuspended in 10 µl of MegaBACE loading solution (GE Healthcare) and run on a MegaBACE 500 DNA Analysis System (GE Healthcare). Due to technical problems with the MegaBACE sequencer during the project, part of the samples had to be sent to Macrogen (Republic of Korea) for sequencing on a 3730xl DNA Analyzer (ABI).

Sequence editing and alignment

Between two and five randomly chosen samples per population were sequenced for the ITS region and the cpDNA markers trnQ(UUG)-rpS16x1 and trnV(UAC)x2-ndhC to check for intrapopulation variation and to build robust consensus sequences. The raw sequencing ITS data were imported into Geneious 11.1.5 (<https://www.geneious.com>,

Kearse et al. 2012) and assembled by population and both forward and reverse sequences were used for the assembly. The alignment of the consensus sequences of the populations was performed with MUSCLE (Edgar 2004). Raw data from the two cpDNA regions were analyzed with CodonCode Aligner 4.0.2 (CodonCode Corporation, www.codoncode.com). An assembly of the data was done for each primer pair for each sample, checked and edited if necessary. The consensus sequences obtained for each population with the two primer pairs were combined into a single cpDNA consensus sequence per population. These consensus sequences were imported into the Geneious software to be aligned with the MUSCLE alignment tool. Obvious alignment errors were manually edited.

Statistical analysis of sequence data

A maximum likelihood (ML) analysis was performed using the PHYML algorithm (Guindon et al. 2010) implemented in Geneious. The model of nucleotide substitution used was determined with jModeltest 2.1.5 (Darriba et al. 2012). We chose to use the substitution model GTR + G for the ITS data set and the GTR model for the cpDNA data set. The number of bootstraps was set to 1000 and the gamma distribution parameter for the ITS data set was fixed to 1.134 according to the jModeltest results.

We conducted Bayesian analyses of phylogeny in MrBayes 3.2.2 (Ronquist et al. 2012) for both the combined cpDNA and the ITS after simple indel coding with SeqState 1.4.1 (Müller 2005). The models were the same as for the ML analysis. Four Monte Carlo Markov chains were run in two runs of 1,000,000 generations each. The burn-in discarded the first 25% of the trees and every 1000 generations Markov chains were sampled. A consensus tree was built after the average standard deviation of split frequencies was below 0.01 and posterior probabilities of the clades were computed. Dendroscope 3.5.9 (Huson and Scornavacca 2012) was used to produce tanglegrams.

To construct a phylogenetic network of the cpDNA haplotypes, we used PopART v1.7 (Leigh and Bryant 2015) with the minimum spanning network algorithm (Bandelt et al. 1999).

Results

AFLP analysis of *Gladiolus palustris* populations

Genetic diversity within populations

Using two primer combinations, 130 AFLP bands were scored with no private bands specific to a *G. palustris* population. All 212 individuals had a unique multilocus

genotype. The estimated mean error rate per sample was 3.6%. Error rates for AFLP below 5% are considered acceptable (see Bonin et al. 2004). The expected heterozygosity, as measured by Nei's gene diversity (H_e), was similar in all populations (range H_e : 0.18–0.27; Table 1). Mean genetic diversity varied among geographic regions (ANOVA, $F_{3,8}=5.728$, $P=0.02$) and was significantly lower in the Alsace region (0.19) in comparison to the Ain region (0.26; Post-hoc Tukey test). We found no significant relationship between genetic diversity as measured by Nei's gene diversity and population size ($r=0.33$, $P=0.30$).

Population genetic structure

Using the modal value of ΔK rather than the maximum value of $L(K)$ allowed us to identify with STRUCTURE several groups corresponding to the uppermost hierarchical level of partitioning among populations. In the model with no admixture, the highest modal value of ΔK was at $K=2$ which separated the two Bas-Rhin populations from all other populations indicating that the *G. palustris* populations of the Rhine Valley are genetically distinct. The second highest modal value of ΔK was at $K=4$ corresponding to the number of geographic regions (Fig. 2, Appendix Fig. S1a). In the STRUCTURE model with admixture, the highest modal was also at $K=2$ corresponding to the uppermost level of structure already detected by the model with no admixture (Appendix Fig. S1b). An examination of the standard deviation estimate of $L(K)$ showed that its value strongly increased after K values higher than four indicating that the most likely number of groups would be four. In a third model, we run STRUCTURE on a reduced dataset in a model with admixture without the group containing the two Bas-Rhin populations to detect potential subdivisions of the second group containing the other populations. This

analysis confirmed the subdivision of the second group into the same clusters as the analysis with no admixture (results not shown). The DAPC analysis confirmed that $K=4$ corresponded to the most likely number of clusters (Fig. 3). Membership probabilities of individuals to the four groups in the DAPC analysis were nearly identical to those of the STRUCTURE analysis with no admixture (results not shown).

There was a nearly complete correspondence between the four clusters identified by STRUCTURE and the four geographic regions (Table 1, Fig. 2). A first cluster corresponded to the two populations of the French Jura region and the only Swiss population from the Haute-Savoie region. The second cluster corresponded to the four French populations of the Haute-Savoie region and the third cluster to the three populations of the Ain region. The fourth cluster corresponded to the two Bas-Rhin populations situated in the Rhine Valley.

The neighbor-net network revealed a clustering pattern very similar to the clusters identified by STRUCTURE (Figs. 2, 4). The separation of the two Bas-Rhin populations was highly supported by bootstrapping (842 out of 1000 replicates). In contrast to the STRUCTURE groups, the two populations of the Jura region were clearly separated in the neighbor-net network (Fig. 4). Population Ranchette was more related to the Swiss and Haute-Savoie region whereas population Légna was related to the Ain region. The Swiss population was placed at an intermediate position between the Haute-Savoie region and the Ranchette population of the Jura region.

The estimate of overall F_{ST} obtained by the AMOVA was high ($\Phi_{st}=0.46$). There was a strong differentiation among regions ($\Phi_{st}=0.33$) and a lower differentiation among populations within regions ($\Phi_{st}=0.14$). The variation among individuals within populations accounted for another 54%. Nei's gene diversity of a

Fig. 2 Estimated population structure of *G. palustris* inferred by a Markov chain Monte Carlo Bayesian clustering method (STRUCTURE) of AFLP data using a model of no population admixture. Each individual is represented by a vertical line, which is partitioned in a maximum of $K=4$ differently colored segments that represent the

individual's estimated membership fractions in four clusters. Vertical black lines separate the twelve populations. Runs of ten simulations were aligned using CLUMPP (see text for "Results"). For population names see Table 1. Geographical origin of individuals is indicated. (Color figure online)

Fig. 3 Estimated population structure of *G. palustris* inferred by discriminant analysis of principal components (DACP) of AFLP data from twelve populations. Posterior group assignments and geographi-

cal origin of individuals: Cluster 1=Jura and Switzerland; Cluster 2=Haute-Savoie; Cluster 3=Ain; Cluster 4=Bas-Rhin. Eigenvalues of the three retained discriminant functions are indicated

Fig. 4 Neighbor-net network (Bryant and Moulton 2004) for AFLP data of 212 *Gladiolus palustris* individuals inferred using SplitsTree (Huson and Bryant 2006). Geographical origin of individuals is indicated

population decreased with its specific F_{ST} value (Fig. 5; $r = -0.90$, $P < 0.001$) indicating that populations were not in migration–drift equilibrium. Overall genetic differentiation between all twelve populations (pairwise Φ_{st}) was related to their geographic distance (Mantel test, $r = 0.84$, $P < 0.001$; Fig. 6) indicating isolation by distance (IBD).

ITS and cpDNA study of *G. palustris* and *G. imbricatus*

Sequence and nucleotide variations

We obtained 32 assembled ITS consensus sequences (Genbank accession numbers MK005888–MK005919) and 33 combined consensus sequences from the two cpDNA markers (Genbank accession numbers MK014501–MK014566). We found no evidence of intrapopulation variation for the three markers, except in two populations (KraL and Roh) where part of the sampled individuals differed only by indels in their trnV(UAC)x2-ndhC region. In these cases, we randomly chose one of the two haplotypes as the populational consensus sequences. Intra-genomic polymorphic sites in ITS sequences were marked as ambiguities. The data matrix used for the phylogenetic analysis containing the aligned consensus sequences was 723 bp long for the ITS region and 1015 bp for the combined cpDNA regions. Without considering the outgroup, 89 bp sites were variable for the ITS region and 43 bp sites were variable for the combined cpDNA regions, adding up to 12.8% and 4.2% of variable sites, respectively.

Fig. 5 Relation between genetic diversity (H_e) and the population-specific F_{ST} values in *Gladiolus palustris*. Population Les Rosses, with only three samples, was excluded from the analysis. $r = -0.90$, $P < 0.001$

Fig. 6 Relation between genetic distances (pairwise Φ_{st}) and geographic distances for twelve sampled populations of *Gladiolus palustris*. Note log-scale for geographic distances. $r = 0.84$, $P < 0.001$ (Mantel test)

Phylogenetic and phylogeographic analyses

Phylogenetic trees resulting from the ML and Bayesian analysis of cpDNA and ITS data were presented in a

Fig. 7 Tanglegram of the phylogenetic trees based on **a** cpDNA and **b** ITS markers. Colors represent: green=outgroup, blue=*Gladiolus palustris* clade, red=*G. imbricatus* clade, light blue=outliers. Numbers near the nodes represent the posterior probabilities of the MrBayes tree. The bootstrap values from the maximum likelihood (ML) analysis are represented in between brackets behind the posterior probabilities. Asterisk collapsed branch in the ML tree. GP *G. palustris*, GI *G. imbricatus*, for population names see Table 1. (Color figure online)

combined tanglegram. They showed generally the same branching, with only slight changes in topology and branch support values (Fig. 7). Both the cpDNA and the ITS analysis indicated that there were two main clades regrouping the great majority of the *Gladiolus palustris* and *G. imbricatus* populations, respectively. The ITS trees allowed us to further subdivide the main *G. palustris* clade into two distinct subgroups (ML bootstrap = 99.6% and posterior probability PP > 0.99). The first subgroup contained the French populations of *G. palustris* from the Bas-Rhin region (Herbsheim, Osthause), and *G. palustris* populations from Italy (TurinIII, ComoI, ComoII, ComoIII, Rovereto), from Hungary (Zákányszék) and Germany (Ried). The second subgroup included all French populations of *G. palustris* from the adjacent regions of Jura (Légna, Ranchette), Ain (Mont d'Ain, Nantua Nord, Oyonnax) and Haute-Savoie (Margencel, Perrignier, Perrignier petite, Reuland), the Swiss population (Faverges) and one population from Italy (Ivrea). The second main clade regrouped *G. imbricatus* populations from Germany (Dauban, Dubraucke), Italy (TurinII) and the Czech Republic (Miroslav, Rohozná, Krahulci L). In the cpDNA tree the Czech *G. imbricatus* populations of Rohozná and Krahulci L were separated from the main *G. imbricatus* group (ML bootstrap = 89.3% and PP > 0.99) (Fig. 7).

Some populations of both taxa were classified into different clades in the cpDNA and ITS trees. The Czech population of Krahulci R, determined morphologically as *G. imbricatus*, was placed into the *G. palustris* clade according to its cpDNA sequences and in the *G. imbricatus* clade in the ITS tree. The Czech *G. palustris* population of Hodonin was included in the *G. palustris* clade according to the cpDNA data, but the ITS sequence data revealed a position within the *G. imbricatus* clade. Population Lus-La-Croix-Haute, determined morphologically as *G. imbricatus*, was placed in the *G. palustris* clade in the ITS tree (Fig. 7) where it had an outlier position (ML bootstrap = 90.9%, PP = 0.91). In the cpDNA tree the same population was placed outside the two main *G. palustris* and *G. imbricatus* clades (ML bootstrap = 81.1, PP > 0.99). A closer examination of the cpDNA sequences showed that the haplotype of population Lus-La-Croix-Haute differed both from the *G. palustris* and *G. imbricatus* haplotypes by several indels (Table 2). The second putative French *G. imbricatus* population Lezoux grouped together with the *G. imbricatus* population from Lus-La-Croix-Haute in the ITS trees but was placed inside the *G. imbricatus* clade according to its cpDNA sequences. The *G. palustris* population Felon was placed within the *G. palustris* ITS clade but had a particular outgroup position in the cpDNA trees (Fig. 7). A comparison of the cpDNA sequences from Felon with the outgroup species *G. communis* showed a 100% identity.

Haplotype network

The minimum spanning network calculated with the cpDNA sequences of all *G. palustris* and *G. imbricatus* populations revealed five different haplotypes (Fig. 8, Table 1). For *G. imbricatus* the minimum spanning haplotype network confirmed the phylogenetic subdivision of the *G. imbricatus* cpDNA tree with the Czech populations of Rohozná and Krahulci L sharing the same haplotype. The position of the *G. palustris* population Felon and the *G. imbricatus* population Lus-La-Croix-Haute was intermediate between the *G. palustris* and *G. imbricatus* haplotype clades. Both populations were separated from these clades by three, respectively, five mutational steps. The haplotype of population Felon was identical to the haplotype of the outgroup species *G. communis* (not shown in Fig. 8), whereas the haplotype of Lus-La-Croix-Haute differed from extant *G. palustris* or *G. imbricatus* haplotypes (Table 2).

Discussion

Genetic diversity and population structure of *G. palustris*

The overall genetic diversity found in the populations of *G. palustris* was comparable to that of other *Gladiolus* species like *G. hybridus* (Chaudhary et al. 2018) and other long-lived perennial plant species (Nybohm 2004). The fact that *G. palustris* is polyploid could contribute to the conservation of genetic diversity even in fragmented and isolated populations (see Doyle and Sherman-Broyles 2017; Meirmans et al. 2018). Other rare Central European plants such as *Saxifraga rosacea* subsp. *sponhemica* (Walisch et al. 2014), *Arnica montana* (Maurice et al. 2016) and *Iris pumila* (Dembicz et al. 2018) are also known to have conserved high levels of genetic diversity despite strong habitat fragmentation. The longevity of these plants and the long-term stability of their habitats could explain the maintenance of genetic diversity (Young et al. 1996; Tang et al. 2010). Many studies have found a reduction of genetic diversity in small and fragmented populations through genetic drift (Leimu et al. 2006). In *G. palustris* we found no relationship between the genetic diversity of the populations and their size. The relationship between genetic diversity and population-specific F_{ST} -values indicated a migration-drift disequilibrium suggesting that the population structure of *G. palustris* has been modified by genetic drift (Whitlock and McCauley 1999; Cox et al. 2011; Walisch et al. 2014).

In *G. palustris* the strong correlation between genetic and geographic distances indicated an isolation by distance pattern (IBD) where there is higher gene flow among geographically close populations. As the current distribution

Table 2 Variable sites and haplotypes of the combined cpDNA consensus sequences of *Gladiolus palustris* (23 individuals) and *G. imbricatus* (nine individuals)

Population	Variable sites	Haplotype
GP_Ri	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	A
GP_Re	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_CoI	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GI_KraR	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_CoIII	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_CoII	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Fa	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_He	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Hod	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Le	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Ro	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Os	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Rov	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_TuIII	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Za	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Iv	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Ma	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_MtA	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Nn	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Oy	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Pe	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Pp	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Ra	ATCCAACCCCAAAAGAAACTAT-----CATAAAT-C	
GP_Fel	CTACAACCCCAAAAGAAACTA-GTTCTCATCAATAT	B
GI_LCH	CGACAACCCCAAAAGAAACTA-GTTCT-A-----C	C
GI_Dau	CTA-----AGT-GTTCTCCTAAAT-C	D
GI_Dub	CTA-----AGT-GTTCTCCTAAAT-C	
GI_Mir	CTA-----AGT-GTTCTCCTAAAT-C	
GI_Lez	CTA-----AGT-GTTCTCCTAAAT-C	
GI_TuII	CTA-----AGT-GTTCTCCTAAAT-C	
GI_KrahL	CTA-----ATA-GTTCTCCTAAAT-C	E
GI_Roh	CTA-----ATA-GTTCTCCTAAAT-C	

Number of segregating sites: 8; number of parsimony-informative sites: 6. GP=*G. palustris*, GI=*G. imbricatus*, for population names see Table 1

is highly disjunct and most extant populations are highly isolated, current gene flow among populations is probably very low or absent. It is likely that the observed IBD pattern

may reflect historical gene flow when suitable habitats of *G. palustris* were much more common. The conservation of the historical pattern may be explained by the fact that

Fig. 8 Minimum spanning network of cpDNA sequences of 23 *Gladiolus palustris* and nine *G. imbricatus* populations. The number of mutations among the different haplotypes are marked as hatch marks. GP = *G. palustris*, GI = *G. imbricatus*, for population names see Table 1

the habitat fragmentation is quite recent and mainly due to the intensification of agricultural practices during the last century. Moreover, the fact that *G. palustris* is a long-lived geophyte may slow down the effects of genetic drift (Aguilar et al. 2008) and thus contribute to the conservation of the historical pattern.

The results of the STRUCTURE analysis suggest that the genetic population structure of *G. palustris* in the studied area is hierarchical and consists of two main clusters substructured into regional groups of populations. The distinctiveness of the low-altitude Rhine Valley populations is remarkable and was confirmed by the neighbor-net network and the DAPC analysis. It is likely that these isolated populations are the remnants of a much larger population of the Alsatian Ried, which may be locally adapted to the warmer climatic conditions of the Rhine Valley. Common garden experiments and reciprocal transplant experiments would be necessary to confirm this hypothesis. The AMOVA revealed that a large part of the molecular variation of the *G. palustris* populations is due to differences among geographical regions indicating that the regional clusters have been isolated for a long time, probably since the post-glacial period. An additional explanation for the strong differentiation among populations could be the breeding system of *G. palustris* (facultative autogamous; Maurice, unpublished) as facultative selfing reduces gene flow and increases genetic differentiation among populations (Mable and Adam 2007).

Hybridization and phylogeography

Interspecific hybrids between *G. palustris* and *G. imbricatus* are difficult to detect in the field as they are either morphologically similar to *G. imbricatus* or morphologically intermediate between the two parental taxa (Szczepaniak et al. 2016). The results of the cpDNA and ITS analyses

confirmed that *G. palustris* and *G. imbricatus* are genetically distinct and represent two different clades (see Szczepaniak et al. 2016). Our genetic analyses also showed that the three French populations Lezoux, Lus-la-Croix-Haute and Felon at the western limit of the distribution area of *G. palustris*, considered being *G. imbricatus* (Tison and Girod 2014), are in fact of hybrid origin. Our cpDNA and ITS data suggest that population Lezoux is a hybrid taxon between *G. imbricatus* as maternal species and *G. palustris* as a pollen donor. This indicates that hybridization events between both taxa are not only unidirectional as reported in Szczepaniak et al. (2016) for the nothospecies *G. × sulistrovicus* with *G. palustris* as the maternal species and *G. imbricatus* as the pollen donor. In population Lus-La-Croix-Haute the cpDNA data suggest a hybridization event with an unknown *Gladiolus* lineage acting as maternal species. The separation of both Lezoux and Lus-La-Croix-Haute from the two larger ITS subgroups of *G. palustris* suggest that these hybridization events may be ancient. The lineage of *G. palustris* involved in these hybridization events is likely to be nowadays extinct, especially in the case of population Lezoux that occurs beyond the western range limit of extant *G. palustris* populations. Moreover, as there are no extant populations of *G. imbricatus* in this region, recent hybridization events cannot have taken place. Our results suggest that in the post-glacial period, the area of *G. imbricatus* may have extended further to the West and that population Lezoux may reflect an ancient hybridization event between *G. imbricatus* and *G. palustris*. This hybridization event at the western range limit of *G. palustris* in France may be qualified as genetic swamping where parental lineages are replaced by hybrids (see Todesco et al. 2016).

The isolated population Felon was morphologically less clearly related to *G. imbricatus* (see Tison and Girod 2014) and was determined as *G. palustris* by the field collector in

our study (Table 1). Our cpDNA analysis suggests that population Felon may be an interspecific hybrid between *G. communis* as maternal species and *G. palustris* as a pollen donor. Population Felon is situated in the Burgundian gate, which is a flat saddle about 30 km wide at an altitude of about 400 m between the Vosges and the Jura mountains. It connects the Rhine Valley and the foothills of the Saône Valley. During the post-glacial period, numerous plant and animal species used this gateway on their way north (Wassmer et al. 1994; Sternberg 1998; Granoszewski et al. 2004; Scholler et al. 2012; Westrich and Büllès 2016). To our knowledge, there is currently no extant *G. communis* population in the Felon region, but a population determined as *G. communis* occurs some 35 km away near Mulhouse in France (Hoff 2018). Population Felon may thus reflect an ancient hybridization event during the post-glacial period between *G. communis* and *G. palustris*.

Our genetic analyses further revealed that two Czech populations are also of hybrid origin and may correspond to the recently described nothospecies *G. x sulistrovicus* (Szczeplaniak et al. 2016). These two hybrid populations are situated at the eastern limit of the main distribution area of *G. palustris* in Central Europe where *G. palustris* and *G. imbricatus* are partially sympatric. In this area, interspecific hybridization between the two closely related *Gladiolus* species is facilitated due to their occurrence in the same habitat and their overlapping flowering periods (Szczeplaniak et al. 2016).

These results suggest a different hybridization asymmetry at the western than at the eastern range limit of the distribution area of *G. palustris* in Western and Central Europe, respectively. At the western range limit in France, we found indications for nuclear introgression from the ‘common’ *G. palustris* into the ‘rare’ *G. imbricatus* that became locally extinct, whereas at the eastern range limit in Czech Republic there were signs of nuclear introgression from *G. imbricatus* into the locally ‘rare’ *G. palustris*. However, as the distribution area of *G. imbricatus* in the East was only sampled with a few populations, we cannot distinguish between the influence of gene flow or ancestral polymorphisms in the incongruence of the cpDNA and ITS dataset, especially as polyploids may retain ancestral polymorphisms much longer than diploids (see Slotte et al. 2008). Our findings are in line with expectations of unidirectional ‘pollen swamping’ of rare species by more abundant congeners (Ellstrand and Elam 1993; Levin et al. 1996; Beatty et al. 2010) and of extinction through genetic swamping (Todesco et al. 2016). Hybridization has been shown to present an extinction risk for other rare plant species like *Nuphar pumila* (by the common *N. lutea*; Arrigo et al. 2016) and the endangered *Castilleja levisecta* (by the common *C. hispida*; Sandlin et al. 2018 unpublished). However, there are also examples where the gene flow is unidirectional from the rare species into the

common conspecific as in the Mediterranean *Lotus fulgurans* and *L. dorycnium* (Conesa et al. 2010).

Conservation implications

The analysis of the genetic population structure of *G. palustris* revealed a strong differentiation among geographical regions indicating that current gene flow is very low. This implies that the conservation of populations from each regional cluster is important to preserve genetic diversity of this legally protected taxon. Although most populations within regions retained a considerable amount of genetic variation, genetic diversity is likely to decrease in small populations of *G. palustris* due to genetic drift as the extant populations are not in a migration–drift equilibrium.

To counteract the future loss of genetic diversity through drift, artificial gene flow has been advocated (Keller and Waller 2002; Hamilton et al. 2017; Van Rossum and Raspé 2018) to increase population genetic diversity and evolutionary resilience. The moderate level of population differentiation within regions suggests that the largest populations per region could be used as seed source to increase genetic diversity in genetic depauperate populations of the same region, especially in small populations. However, as we cannot exclude the risk of outbreeding depression, we suggest conducting first within and between-population crossing experiments in a common garden experiment and measure the vigor and fitness of the offsprings. Introducing seed material instead of young plants would allow for selection of maladapted genotypes before establishment and flowering (Willi et al. 2007; Maschinski et al. 2013; Van Rossum and Raspé 2018). As the isolated Bas-Rhin populations in the Alsatian Ried occur at a lower altitude, we suggest investigating in a further study if these lowland populations are locally adapted to warmer climatic conditions in the Rhine Valley. This would further enhance their conservation interest especially in the context of climate change.

Our phylogenetic analysis permitted to clarify the effects of ancient and recent hybridization events on the conservation of *G. palustris*. It showed that none of the presumed *G. imbricatus* populations in France correspond to a purebred *G. palustris* or *G. imbricatus* species. Our results indicate that certain hybrid populations at the western range limit of *G. palustris* may be products of ancient hybridization events where one of the parent species went regionally extinct. Because these populations seem to be self-sustainable without the presence of the parent taxa, we advocate that efforts should be made to conserve the French hybrid populations due to their biogeographic interest. In contrast, the recently discovered hybrid populations at the eastern range limit of *G. palustris* in Central Europe confirm that *G. palustris* and *G. imbricatus* are closely related and are still able to hybridize if they co-occur (Szczeplaniak et al. 2016). It is however

unclear if this recent hybrid can exist without the parent species as it had been shown that the hybrid has a reduced seed set (see Szczepaniak et al. 2016). The nothospecies *G. x sulistrovicus* should therefore be monitored over several generations to evaluate its demographic and genetic dynamics.

Acknowledgements Financial support by the Musée national d'histoire naturelle Luxembourg is greatly acknowledged. We thank Philippe Antonetti, Eszter Aradi, Eric Brugel, Géraldine Garnier, Gaëlle Grandet, Christophe Hennequin, Gilberto Parolo, Markus Peintinger and Frank Richter for providing plant samples and assistance during fieldwork and Diethart Matthies and two anonymous referees for making valuable comments on earlier versions of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* 17:5177–5188. <https://doi.org/10.1111/j.1365-294X.2008.03971.x>
- Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends Ecol Evol* 16:613–622. [https://doi.org/10.1016/S0169-5347\(01\)02290-X](https://doi.org/10.1016/S0169-5347(01)02290-X)
- Arrigo N, Bétrisey S, Graf L, Bilat J, Gerber E, Kozłowski G (2016) Hybridization as a threat in climate relict *Nuphar pumila* (Nymphaeaceae). *Biodivers Conserv* 25:1863–1877. <https://doi.org/10.1007/s10531-016-1165-z>
- Bakker JP, Berendse F (1999) Constraints in the restoration of ecological diversity in grassland and heathland communities. *Trends Ecol Evol* 14:63–68. [https://doi.org/10.1016/S0169-5347\(98\)01544-4](https://doi.org/10.1016/S0169-5347(98)01544-4)
- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036>
- Beatty GE, Philipp M, Provan J (2010) Unidirectional hybridization at a species' range boundary: implications for habitat tracking. *Divers Distrib* 16:1–9. <https://doi.org/10.1111/j.1472-4642.2009.00616.x>
- Bilz M, Kell SP, Maxted N, Lansdown RV (2011) European red list of vascular plants. Publications Office of the European Union, Luxembourg
- Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Mol Ecol* 13:3261–3273. <https://doi.org/10.1111/j.1365-294X.2004.02346.x>
- Bryant D, Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol* 21:255–265. <https://doi.org/10.1093/molbev/msh018>
- Cantor M, Tolety J (2011) *Gladiolus*. In: Kole C (ed) Wild crop relatives: genomic and breeding resources. Springer, Berlin, pp 133–159
- Chapman MA, Abbott RJ (2010) Introgression of fitness genes across a ploidy barrier. *New Phytol* 186:63–71. <https://doi.org/10.1111/j.1469-8137.2009.03091.x>
- Chaudhary V, Kumar M, Sharma S, Kumar N, Kumar V, Yadav HK, Sharma S, Sirohi U (2018) Assessment of genetic diversity and population structure in gladiolus (*Gladiolus hybridus* Hort.) by ISSR markers. *Physiol Mol Biol Plants* 24:493–501. <https://doi.org/10.1007/s12298-018-0519-2>
- Cieślak E, Szczepaniak M, Kamiński R, Heise W (2014) Stan zachowania krytycznie zagrożonego gatunku *Gladiolus paluster* (Iridaceae) w Polsce—analiza zmienności genetycznej osobników w uprawie Ogrodu Botanicznego Uniwersytetu Wrocławskiego w kontekście prowadzonych działań ochronnych. *Fragm Florist Geobot Pol* 21:49–66
- Coates D (1988) Genetic diversity and population genetic structure in the rare chattering grass wattle, *Acacia anomala* court. *Aust J Bot* 36:273. <https://doi.org/10.1071/BT9880273>
- Conesa MA, Mus M, Rosselló JA (2010) Who threatens who? Natural hybridization between *Lotus dorycnium* and the island endemic *Lotus fulgurans* (Fabaceae). *Biol J Linn Soc* 101:1–12. <https://doi.org/10.1111/j.1095-8312.2010.01456.x>
- Cox K, Broeck AV, Calster HV, Mergeay J (2011) Temperature-related natural selection in a wind-pollinated tree across regional and continental scales. *Mol Ecol* 20:2724–2738. <https://doi.org/10.1111/j.1365-294X.2011.05137.x>
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772. <https://doi.org/10.1038/nmeth.2109>
- Dembicz I, Szczeparska L, Moysiienko II, Wódkiewicz M (2018) High genetic diversity in fragmented *Iris pumila* L. populations in Ukrainian steppe enclaves. *Basic Appl Ecol* 28:37–47. <https://doi.org/10.1016/j.baae.2018.02.009>
- Doyle JJ, Sherman-Broyles S (2017) Double trouble: taxonomy and definitions of polyploidy. *New Phytol* 213:487–493
- Dufresne F, Stift M, Vergilino R, Mable BK (2014) Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Mol Ecol* 23:40–69. <https://doi.org/10.1111/mec.12581>
- 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>
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annu Rev Ecol Syst* 24:217–242. <https://doi.org/10.1146/annurev.es.24.110193.001245>
- 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, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- 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>
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180:977–993. <https://doi.org/10.1534/genetics.108.092221>
- Frankham R, Wilcken J (2006) Does inbreeding distort sex-ratios? *Conserv Genet* 7:879–893. <https://doi.org/10.1007/s10592-006-9129-6>

- Gaskin JF, Kazmer DJ (2009) Introgression between invasive saltcedars (*Tamarix chinensis* and *T. ramosissima*) in the USA. *Biol Invasions* 11:1121–1130. <https://doi.org/10.1007/s10530-008-9384-1>
- Gitzendanner MA, Soltis PS (2000) Patterns of genetic variation in rare and widespread plant congeners. *Am J Bot* 87:783. <https://doi.org/10.2307/2656886>
- Granoszewski W, Nita M, Nalepka D (2004) *Vitis vinifera* L. subsp. *sylvestris* (C. C. Gmelin) Hegi—wild grape-vine. In: Ralska-Jasiewiczowa M, Latalowa M, Wasylkowska K, Tobolski K, Madeyska E, Wright HE, Turner C (eds) Late glacial and holocene history of vegetation in Poland based on isopollen maps. Polish Academy of Sciences, Szafer Institute of Botany, Kraków, pp 245–249
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321. <https://doi.org/10.1093/sysbio/syq010>
- Hamilton JA, Royauté R, Wright JW, Hodgskiss P, Ledig FT (2017) Genetic conservation and management of the California endemic, Torrey pine (*Pinus torreyana* Parry): implications of genetic rescue in a genetically depauperate species. *Ecol Evol* 7:7370–7381. <https://doi.org/10.1002/eece3.3306>
- Hoff M (2018) Données de la flore d'Alsace (SBA). UMS PatriNat (AFB-CNRS-MNHN), Paris. Occurrence dataset <https://doi.org/10.15468/1c9lfo> accessed via GBIF.org on 2018-10-01. <https://www.gbif.org/occurrence/1867669282>
- Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23:254–267. <https://doi.org/10.1093/molbev/msj030>
- Huson DH, Scornavacca C (2012) Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst Biol* 61:1061–1067. <https://doi.org/10.1093/sysbio/sys062>
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- 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>
- Kearse R, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- 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](https://doi.org/10.1016/S0169-5347(02)02489-8)
- Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network construction. *Methods Ecol Evol* 6:1110–1116. <https://doi.org/10.1111/2041-210X.12410>
- Leimu R, Mutikainen P, Koricheva J, Fischer M (2006) How general are positive relationships between plant population size, fitness and genetic variation? *J Ecol* 94:942–952. <https://doi.org/10.1111/j.1365-2745.2006.01150.x>
- Levin DA, Francisco-Ortega J, Jansen RK (1996) Hybridization and the extinction of rare plant species. *Conserv Biol* 10:10–16. <https://doi.org/10.1046/j.1523-1739.1996.10010010.x>
- Mable BK, Adam A (2007) Patterns of genetic diversity in outcrossing and selfing populations of *Arabidopsis lyrata*. *Mol Ecol* 16:3565–3580. <https://doi.org/10.1111/j.1365-294X.2007.03416.x>
- Mallet J (2005) Hybridization as an invasion of the genome. *Trends Ecol Evol* 20:229–237. <https://doi.org/10.1016/j.tree.2005.02.010>
- Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philos Trans R Soc B* 363:2971–2986. <https://doi.org/10.1098/rstb.2008.0081>
- Maschinski J, Wright SJ, Koptur S, Pinto-Torres EC (2013) When is local the best paradigm? Breeding history influences conservation reintroduction survival and population trajectories in times of extreme climate events. *Biol Conserv* 159:277–284. <https://doi.org/10.1016/j.biocon.2012.10.022>
- 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>
- Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60:2399–2402. <https://doi.org/10.1111/j.0014-3820.2006.tb01874.x>
- Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* 4:792–794. <https://doi.org/10.1111/j.1471-8286.2004.00770.x>
- Meirmans PG, Liu S, van Tienderen PH (2018) The analysis of polyploid genetic data. *J Hered* 109:283–296. <https://doi.org/10.1093/jhered/esy006>
- Meusel H, Jäger EJ, Rauschert S, Weinert E (eds) (1965) Vergleichende Chorologie der zentraleuropäischen Flora: Karten
- Müller K (2005) SeqState: primer design and sequence statistics for phylogenetic DNA datasets. *Applied Bioinformatics* 4:65–69. <https://doi.org/10.2165/00822942-200504010-00008>
- Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, p 512
- Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, Oxford, p 333
- Nierbauer KU, Paule J, Zizka G (2017) Heteroploid reticulate evolution and taxonomic status of an endemic species with bicentric geographical distribution. *AoB Plants* 9:plx002. <https://doi.org/10.1093/aobpla/plx002>
- Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol Ecol* 13:1143–1155. <https://doi.org/10.1111/j.1365-294X.2004.02141.x>
- Oelmann Y, Broll G, Hölzel N, Kleinebecker T, Vogel A, Schwartz P (2009) Nutrient impoverishment and limitation of productivity after 20 years of conservation management in wet grasslands of north-western Germany. *Biol Conserv* 142:2941–2948. <https://doi.org/10.1016/j.biocon.2009.07.021>
- Paun O, Schönschetter P (2012) Amplified fragment length polymorphism (AFLP)—an invaluable fingerprinting technique for genomic, transcriptomic and epigenetic studies. *Methods Mol Biol* 862:75–87. https://doi.org/10.1007/978-1-61779-609-8_7
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542. <https://doi.org/10.1093/sysbio/sys029>
- Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* 4:137–138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Scholler M, Hemm V, Lutz M (2012) *Erysiphe platani*: monitoring of an epidemic spread in Germany and molecular characterization based on rDNA sequence data. *Andrias* 19:263–272

- Segatto ALA, Reck-Kortmann M, Turchetto C, Freitas LB (2017) Multiple markers, niche modelling, and bioregions analyses to evaluate the genetic diversity of a plant species complex. *BMC Evol Biol* 17:234. <https://doi.org/10.1186/s12862-017-1084-y>
- Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am J Bot* 94:275–288. <https://doi.org/10.3732/ajb.94.3.275>
- Slotte T, Huang H, Lascoux M, Ceplitis A (2008) Polyploid speciation did not confer instant reproductive isolation in *Capsella* (Brassicaceae). *Mol Biol Evol* 25:1472–1481. <https://doi.org/10.1093/molbev/msn092>
- Smith TW, Waterway MJ (2008) Evaluating the taxonomic status of the globally rare *Carex roanensis* and allied species using morphology and amplified fragment length polymorphisms. *Syst Bot* 33:525–535. <https://doi.org/10.1600/036364408785679824>
- Soltis PS (2013) Hybridization, speciation and novelty. *J Evol Biol* 26:291–293. <https://doi.org/10.1111/jeb.12095>
- Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. *Annu Rev Plant Biol* 60:561–588. <https://doi.org/10.1146/annurev.arplant.043008.092039>
- Sternberg K (1998) The postglacial colonization of Central Europe by dragonflies, with special reference to Southwestern Germany (Insecta, Odonata). *J Biogeogr* 25:319–337
- Szczepaniak M, Kamiński R, Kuta E, Słomka A, Heise W, Cieślak E (2016) Natural hybridization between *Gladiolus palustris* and *G. imbricatus* inferred from morphological, molecular and reproductive evidence. *Preslia* 88:137–161
- Tang L, Xiao Y, Li L, Guo Q, Bian Y (2010) Analysis of genetic diversity among Chinese *Auricularia auricula* cultivars using combined ISSR and SRAP markers. *Curr Microbiol* 61:132–140. <https://doi.org/10.1007/s00284-010-9587-4>
- Tison J-M, Girod C (2014) Synopsis du genre *Gladiolus* L. (Iridaceae) en France. *Bot J Bot Soc France* 68:69–75
- Todesco M, Pascual MA, Owens GL, Ostevik KL, Moyers BT, Hübner S, Heredia SM, Hahn MA, Caseys C, Bock DG, Rieseberg LH (2016) Hybridization and extinction. *Evol Appl* 9:892–908. <https://doi.org/10.1111/eva.12367>
- Turchetto C, Segatto ALA, Beduschi J, Bonatto SL, Freitas LB (2015) Genetic differentiation and hybrid identification using microsatellite markers in closely related wild species. *AoB Plants* 7:plv084. <https://doi.org/10.1093/aobpla/plv084>
- UICN France, FCBN, MNHN (2012). La Liste rouge des espèces menacées en France—Chapitre Flore vasculaire de France métropolitaine: premiers résultats pour 1 000 espèces, sous-espèces et variétés. Dossier électronique
- Van Duren IC, Boeye D, Grootjans AP (1997) Nutrient limitations in an extant and drained poor fen: implications for restoration. *Plant Ecol* 133:91–100. <https://doi.org/10.1023/A:1009728007279>
- van Raamsdonk LWD, de Vries T (1989) Biosystematic studies in European species of *Gladiolus* (Iridaceae). *Plant Syst Evol* 165:189–198. <https://doi.org/10.1007/BF00936001>
- 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>
- Vergeer P, Rengelink R, Copal A, Ouborg NJ (2003) The interacting effects of genetic variation, habitat quality and population size on performance of *Succisa pratensis*. *J Ecol* 91:18–26. <https://doi.org/10.1046/j.1365-2745.2003.00736.x>
- Walisch TJ, Matthies D, Hermant S, Colling G (2014) Genetic structure of *Saxifraga rosacea* subsp. *sponhemica*, a rare endemic rock plant of Central Europe. *Plant Syst Evol* 301:251–263. <https://doi.org/10.1007/s00606-014-1070-4>
- Wassmer T, Himmelsbach W, Himmelsbach R (1994) Dungbewohnende Blatthornkäfer (Scarabaeoidea) und Wasserkäfer (Hydrophilidae) aus dem Hessental bei Schelingen im Kaiserstuhl. *Mitt Bad Landesver Naturkd Naturschutz*, NF 16:75–83
- Wesche K, Krause B, Culmsee H, Leuschner C (2012) Fifty years of change in Central European grassland vegetation: large losses in species richness and animal-pollinated plants. *Biol Conserv* 150:76–85. <https://doi.org/10.1016/j.biocon.2012.02.015>
- Westrich P, Bülls J (2016) *Epeolus fallax*, ein Brutparasit von *Colletes hederæ* und eine für Deutschland neue Bienenart (Hymenoptera, Apidae). *Eucera, Beiträge zur Apidologie* 10:15–26
- White TJ, Bruns T, Lee S, Taylor JL (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods and applications. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols. A guide to methods and applications. Academic Press, San Diego, pp 315–322
- Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity* 82:117–125
- Willi Y, van Kleunen M, Dietrich S, Fischer M (2007) Genetic rescue persists beyond first-generation outbreeding in small populations of a rare plant. *Proc R Soc B* 274:2357–2364. <https://doi.org/10.1098/rspb.2007.0768>
- Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends Ecol Evol* 11:413–418. [https://doi.org/10.1016/0169-5347\(96\)10045-8](https://doi.org/10.1016/0169-5347(96)10045-8)
- Yuan J, Cheng F, Zhou S (2012) Genetic structure of the tree peony (*Paeonia rockii*) and the Qinling Mountains as a geographic barrier driving the fragmentation of a large population. *PLoS ONE* 7:e34955. <https://doi.org/10.1371/journal.pone.0034955>
- Zawko G, Krauss SL, Dixon KW, Sivasithamparam K (2001) Conservation genetics of the rare and endangered *Leucopogon obtectus* (Ericaceae). *Mol Ecol* 10:2389–2396. <https://doi.org/10.1046/j.0962-1083.2001.01378.x>

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