RESEARCH ARTICLE



Low genetic differentiation despite high fragmentation in the endemic serpentinophyte *Minuartia smejkalii* (*M. verna* agg., Caryophyllaceae) revealed by RADSeq SNP markers

Bojana Stojanova¹ · Maria Šurinová^{2,3} · Vojtěch Zeisek^{2,3} · Zuzana Münzbergová^{2,3} · Hana Pánková²

Received: 28 May 2019 / Accepted: 20 November 2019 / Published online: 2 January 2020 © Springer Nature B.V. 2020

Abstract

Minuartia smejkalii is an obligate serpentinophyte plant endemic to the Czech Republic. Since the 1960s, the species' habitat has undergone strong human-mediated fragmentation, resulting in extinction of some populations and dramatic size reduction of the remaining populations. Thus, contrary to the typically stable serpentine habitats, *M. smejkalii* habitats underwent a recent and severe decline, which can exacerbate the effects of fragmentation on population genetic structure. We examined the genetic structure of all known *M. smejkalii* populations and two populations of *M. corcontica* and *M. caespitosa*, which are closely related, using RADSeq. The results indicate low, but clear differentiation among the three species, thus supporting the status of *M. smejkalii* as an independent taxon, though more extensive analysis of the whole group is needed. We further show high genetic diversity within *M. smejkalii* populations, low to moderate among-populations differentiation, and moderate regional differentiation. This could be due to the outcrossing mating system of *M. smejkalii* promoting high levels of gene flow and historical factors (multiple founder events, a recent bottleneck and/or a genetic time lag). We finally demonstrate that 2–3% of the markers show differentiation patterns consistent with divergent selection, suggesting that some local adaptation might have occured in *M. smejkalii*. Based on our observations, but without any experimental testing for local adaptation, if a conservation action is to be carried out, we recommend strictly separating the material from the two regions, and if possible, separating the populations within a region.

Keywords Caryophyllaceae · SNP markers · Anonymous loci · Relict species · Edaphic isolation · Species conservation

Introduction

Securing species persistence under the threat of humanmediated habitat fragmentation and reduction is a major concern for conservation biologists (Fahrig 2003; Lindenmayer and Fischer 2013). Classical in situ conservation approaches

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10592-019-01239-4) contains supplementary material, which is available to authorized users.

Bojana Stojanova bojana.stojanova@gmail.com

- ¹ Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava, Czech Republic
- ² Institute of Botany, Academy of Sciences of the Czech Republic, Průhonice, Czech Republic
- ³ Department of Botany, Faculty of Science, Charles University, Prague, Czech Republic

consist of restoration of habitat and species diversity, maintaining or increasing population sizes, and reintroducing species in locations where they went extinct. Lately, conservation policies are increasingly focusing on the preservation of genetic diversity in threatened species (Rodríguez-Quilón et al. 2016). Because it reduces population size and connectivity, habitat fragmentation can erode genetic variation through increased genetic drift and inbreeding, and reduced gene flow (Young et al. 1996). In endangered species in particular, high genetic diversity is important to buffer the potential negative impacts of environmental change, increase the adaptive potential of populations, and reduce the risk of inbreeding depression (Mimura et al. 2017).

Endemic species are a priority category for conservation, since their restricted geographic distribution makes them particularly vulnerable to habitat fragmentation and reduction (Wolf 2001; Wolf and Harrison 2001). Maintaining high genetic diversity is crucial for the long-term conservation of endemic species, as they rarely, if at all, migrate towards new habitats when facing habitat degradation. A typical example of such a scenario is obligate serpentinophyte species, whose survival is bound to serpentine soils in nature. Serpentine soils have an ubiquitous but highly patchy distribution, covering less than 1% of the total dryland, which limits the possibility for migration of serpentine-adapted populations. These soils are characterized by high concentrations of toxic heavy metals (magnesium, nickel, cobalt, chromium), low levels of nutrients (potassium, phosphorous, nitrogen), and are often steep, rocky and subject to erosion. These conditions are frequently linked with low water retention. Thus, populations growing on serpentine soils experience strong selective pressures, as they need to adapt to the soil chemical composition, drought, and high temperatures (Linhart and Grant 1996; Brady et al. 2005). Due to the extreme conditions, the vegetation on serpentine soils is usually sparse, which results in a less competitive environment. As a consequence, serpentine-adapted populations, even when capable of growing on non-serpentine soils alone, are quickly extirpated because they cannot outcompete the non-serpentine species, likely because of a trade-off between stress tolerance and competitive abilities (Kruckeberg 1984; Harrison and Rajakaruna 2011).

The strong selective pressure and spatial isolation of serpentine soils are expected to result in low genetic diversity in serpentine species through bottlenecks, founder effects, or rapid elimination of poorly adapted genotypes (Whittaker 1954; Moore et al. 2013). Yet a remarkable number of serpentine species show high genetic variation within populations and small regional scales that are equal to or surpass genetic variation of their non-serpentine congeners (Ducousso et al. 1990; Linhart and Grant 1996; Mengoni et al. 2000; Mattner et al. 2002; Quintela-Sabarís et al. 2010; Moore et al. 2013). Possible explanations for this are gene flow from neighboring serpentine and non-serpentine habitats (if the species is not a strict serpentinophyte), multiple events of colonization in the past when serpentine soils acted as refugia for non-competitive genotypes, or the accumulation of new mutations and variants if the serpentine populations were isolated for a long time (Ducousso et al. 1990; Mengoni et al. 2000; Słomka et al. 2011). Typical serpentine habitats, although scattered, are usually stable over long time periods (Cheptou et al. 2017), which can help to maintain and build up genetic diversity of serpentine species.

A disjunct geographic range encompassing a wide edaphic niche ranging from calcareous and siliceous substrates to serpentines soils is well exemplified in the *Minuartia verna* agg. The *M. verna* agg. includes perennial taxa widely distributed in the Northern hemisphere (Meusel et al. 1965), but confined to various open habitats. The complex was widely distributed during the last glacial maxima, but in the Holocene its range retracted to so-called interglacial refugia on various extreme substrates, including serpentine soils, by the changing climatic conditions and increased competition from other species (Novák 1960). Being subjected to long-term geographic isolation and highly stressful serpentine soil conditions, the relict populations of M. verna agg. became genetically isolated through genetic drift, mutation and selection, resulting in several relict species (sensu biogeographic relicts, Lomolino et al. 2017) within the aggregate (Novák 1960; Mayer et al. 1994; Chytrý et al. 2017). Depending on the taxonomic concept, the M. verna agg. complex consists of between 3 and 15 species, often with contrasting edaphic preferences and geographic ranges. One of these is species is *Minuarita smejkalii*, an endemic to the Czech Republic and confined to small serpentine outcrops in highlands of the Central part of the Czech Republic (Dvořáková 1988). Two other species of the M. verna agg. complex are present in the Czech Republic: M. corcontica is an endemic occurring exclusively in the in Krkonoše Mountains (Dvořáková 2003), and M. caespitosa has a distribution which extends to Germany and Belgium (Kubát 2002). Up until now, the distinction among the three Minuartia species has only been made based on morphology (Dvořáková 1988, 2003), and no genetic comparisons have yet been made.

Being an obligate serpentinophyte, *M. smejkalii*'s natural occurrences are highly restricted by the patchy distribution of serpentine soils. Known habitats of *M. smejkalii* have been dramatically reduced in size and fragmented since the 1960–1970 by human actions. Thus, contrary to their typically fragmented but stable-over-time serpentine habitats, the habitat of *M. smejkalii* underwent a recent destruction followed by strong population decline. This could result in a severe reduction of genetic diversity through genetic drift, inbreeding, new and strong selective pressures, possibly leading to an extinction vortex (Cheptou et al. 2017).

In this study, we used codominant RADseq SNP markers to genotype individuals from all known localities of M. smejkalii, as well as two Czech populations of each M. caespitosa and M. corcontica. Our first aim was to disentangle the three congeneric species occurring within the Czech Republic based on the SNP markers, in order to contribute to the understanding of the taxonomic status of M. smejkalii and thus help conservation by clarifying its IUCN status. We then focused on the structuring of genetic diversity within the species of M. smejkalii and examined how it is partitioned on different geographic scales-between regions, among populations within a region, and among subpopulations. We also tested for the presence of outlier loci in our data set, which are putatively under selection, or at the very least population- or region-specific (private) alleles (Foll and Gaggiotti 2008). Altogether, these results are the first step into developing an efficient conservation strategy for M. smejkalii that maximizes its genetic diversity while avoiding deleterious effects of inbreeding or outbreeding depression.

Materials and methods

Population description and sampling

Minuartia smejkalii (Dvořáková) is a Czech-endemic species, with a self-compatible, but predominantly outcrossing mating system. M. smejkalii belongs to the species of priority European interest according to Habitats directive 92/43/EEC, Annex II. It is also included in the Convention on the Conservation of European Wildlife and Natural Habitats and the IUCN international Red list (Bilz 2011). According to the Czech laws, M. smejkalii is critically endangered (Grulich 2012). The distribution of M. sme*ikalii* has historically been limited to three regions, all occurring within an area of 500 km²: Borecká skalka Natural Monument, Hadce u Želivky National Natural Monument and Hadce u Hrnčíř Natural Monument (the latter two hereafter referred to as Želivka and Hrnčíře). Populations in the Borecká skalka region went extinct because of mining activities and the expansion of self-seeded trees in the 1960s. More than half of the available serpentine rocks suitable for the species in the Želivka region were destroyed by the construction of a motorway and a dam between 1965 and 1975 (Pešout 2001). In Hrnčíře, M. sme*jkali* occurred on several serpentine localities but they were destroyed by agricultural intensification in the 1960s. Currently, M. smejkalii in Hrnčíře survives only on one site which is an open area in a cultural forest.

In the Želivka region *M. smejkalii* occurs on rock platforms, crevices, and shallow serpentine soils, and fragmented grassland patches in open pine forests with sporadic vegetation (*Asplenion serpentinii* and *A. cuneifolii-Armerion serpentini* communities). There are six

Table 1 Populati characteristics geographically distinct populations in the Želivka region: Z1–5 which are all within 600 m from of each other and Z6, which is located about 2000 m to the east. Populations Z2, Z3 and Z4 were further divided in subpopulations, which correspond to distinctive rock platforms (Table 1; Fig. 1). The total population size in the Želivka region varied between 400 and 1200 in the last 13 years (Pánková, pers.obs). Populations on the border of the region are the smallest or have already gone extinct. In the Hrnčíře region, *M. smejkalii* grows in one population (marked as H) divided in two subpopulations—one in cultural pine forest openings with 21–97 individuals in the last 13 years, and a new subpopulation, naturally established in 2011 along a nearby forest path after tree cutting with 110–227 individuals (Table 1; Fig. 1).

Two other Minuartia species from the M. verna agg. complex occurring in the Czech Republic were also included in the study. Minuartia caespitosa has a broad distribution from eastern Belgium, through north-western Germany to northern Czech Republic (Hejný and Slavík 2003). Minuartia caespitosa occurs in the Czech Republic in a 2 km² area in Hradčanské stěny (Hradčany further in the text), where they occur in calcareous grasslands on small, unstable sand dunes with sporadic vegetation at the bottom of sandstone rocks (Sádlo et al. 2011). For the purpose of this study, two populations 0.4 km from each other in the region of Hradčany were sampled. Minuartia corcontica is endemic to the Czech Republic that at present grows in two populations in the Krkonoše National Park (Krkonoše further in the text), on pyroxen gneiss with a high mineral and sulphide content covering an area of approximately 0.06 km². Only one population in this area (divided in two subpopulations), is accessible for sampling and was thus sampled for this study (Fig. 1).

on genetics	Species	Region	Subpops	Рор	Ν	λ	Ia	% nonHW	H _{obs}	H _{exp}	F _{IS}
	M. corcontica	Krkonoše	1	MC	5	0.8	5	5.687	0.206	0.154	-0.335
			1	VC	5	0.8	4.07	6.299	0.221	0.173	-0.281
	M. caespitosa	Hradčany	1	SK1	5	0.8	4.63	5.074	0.214	0.174	-0.232
			1	SK2	4	0.75	2.85	5.774	0.213	0.174	-0.224
	M. smejkalii	Želivka	1	Z1	9	0.889	1.93	7.962	0.210	0.192	-0.095
			11	Z2	28	0.964	2.5	5.424	0.259	0.217	-0.193
			1	Z3	9	0.889	1.47	20.385	0.269	0.206	-0.305
			2	Z4	10	0.9	2.16	9.099	0.247	0.204	-0.213
			1	Z5	10	0.9	2.92	9.274	0.244	0.213	-0.146
			1	Z6	9	0.889	2.47	7.612	0.213	0.208	-0.027
		Hrnčíře	1	Н	14	0.929	10.88	12.073	0.255	0.210	-0.218
				Total	108	0.991					

N number of genotyped individuals, λ Simson index of the probability that two random genotypes in the population are different, *Ia* index of association, *%nonHW* percentage of the loci in the population that are not in Hardy–Weinberg equilibrium, H_{obs} observed heterozygosity, H_{exp} expected heterozygosity, F_{IS} fixation index





For genotyping we sampled vegetative tissue that consisted of a stem fragment with a few green leaves. The number of individuals sampled per population varied according to the species and population size and structure. In populations Z1, Z5 and Z6, for which a subpopulation structure was not defined, up to ten individuals per population were sampled. In populations that were further divided in subpopulations, three (Z2 with 11 subpopulations) to five (H, Z3, and Z4 with 2 subpopulations) individuals were sampled per subpopulation. For M. corcontica, five individuals per subpopulation, i.e. 10 samples in total were collected (Table 1). The number of samples was limited by the low population size and difficulty in accessing the species. We sampled the same number of plants for M. caespitosa, i.e. five individuals per population. While this species is more widespread, for the purpose of this study we only sampled the population that is geographically the closest to M. smejkalii. Altogether, this resulted in 119 Minuartia samples.

DNA extraction and RADSeq genotyping

Total genomic DNA of 119 individuals was extracted from dehydrated leaves using a modified cetyltrimethylammonium bromide (CTAB) method (Lodhi et al. 1994). Library preparation and sequencing were made by SNPsaurus, LLC, following a previously established protocol (Russello et al. 2015). Genomic DNA was converted into nextRAD genotyping-by-sequencing libraries. Genomic DNA was first fragmented with Nextera reagent (Illumina, Inc.), which also ligates short adapter sequences to the ends of the fragments. The Nextera reaction was scaled for fragmenting 10 ng of genomic DNA, although 20 ng of genomic DNA was used for input to compensate for the amount of degraded DNA in the samples and to increase fragment sizes. Fragmented DNA was then amplified for 26 cycles at 73°, with one of the primers matching the adapter and extending nine nucleotides into the genomic DNA with the selective sequence GTGTAG AGC. Thus, only fragments starting with a sequence that can be hybridized by the selective sequence of the primer were efficiently amplified. The indexes used and the PCR protocol are part of Nextera DNA library preparation kit (Illumina, Inc.). No technical replicates were included. The nextRAD libraries were sequenced on a HiSeq4000 with two lanes (60 and 59 individuals in each line). The nextRAD libraries were sequenced on a HiSeq4000 with two lanes of 150 bp reads (University of Oregon). The genotyping analysis used custom scripts (SNPsaurus, LLC) that trimmed the reads using bbduk (Bushnell 2014).

A de novo reference genome was created by collecting 10 million reads in total, evenly from the samples, and excluding loci that had a depth of coverage of fewer than five (insufficient coverage) or more than 1000 (derived from clonal amplification or repetitive elements; Morris et al. 2016). The remaining loci were then aligned to each other to identify allelic loci and collapse allelic haplotypes to a single locus. All reads were mapped to the reference with an alignment identity threshold of 90% using bbmap (Bushnell 2014). Genotype calling was done using Samtools and bcftools. The vcf was filtered to remove alleles with a population frequency of < 3% which are likely due to a genotyping error. Loci that were heterozygous in all samples or had more than 2 alleles per sample (suggesting collapsed paralogs) were removed. The absence of artifacts was confirmed by counting SNPs at each read nucleotide position and determining that SNP number did not increase with reduced base quality at the end of the read.

Statistical analyses

Because our RADSeq loci are anonymous, we tested for candidate loci under selection (outlier loci) using BayeScan 2.1 (Foll and Gaggiotti 2008). Twenty pilot runs of 5000 iterations were run, and the estimates were made from 5000 iterations with a thinning interval of 10, after a burn-in period of 10,000 iterations. We set a prior odds ratio of 10 (prior belief that a selection model is 1/10 as likely as a neutral model for a given SNP), which is considered as a strong evidence for selection (Foll and Gaggiotti 2008).

For the purposes of the evaluation of the genetic clustering of the individuals, we considered statistical techniques based on Bayesian clustering as very effective because they do not involve a priori hypotheses about sample clustering. We used Bayesian K-means clustering as described for Discriminant Analysis of Principal Components (DAPC, Jombart et al. 2010). It processes PC-scaled data by K-means clustering and maximizes the manifestation of the major pattern involved in the data by subsequent discriminant analysis (Jombart et al. 2010). The best clustering K was chosen according to step-wise decrease of BIC (the biggest decrease) with regard to absolute minimal value of the BIC, as discussed in Jombart et al. (2010).

Most computations were performed in R 3.4 (R Core Team 2014). We used packages ade4 (Dray and Dufour 2007), adegenet (Jombart 2008), APE (Paradis et al. 2004), pegas (Paradis 2010), Poppr (Kamvar et al. 2014) and vcfR (Knaus and Grünwald 2017). We calculated the distribution and diversity of multi locus genotypes (MLGs) within species (package poppr), F-statistics (package hierfstat), allelic richness (package poppr), index of association (Brown et al. 1980; Smith et al. 1993; Agapow and Burt 2001; package poppr), departure from Hardy–Weinberg equilibrium (HWE, Jombart 2008; package pegas), Principal Coordinate Analysis (PCoA, Dray et al. 2003; packages ade4 and adegenet), and K-means clustering (with 10,000,000 iterations and maximal K = 35, Jombart et al. 2010; package adegenet). We applied NeigborNet network analysis using Uncorrected_P

distance matrix in SplitsTree (Huson and Bryant 2005) to better display presumably tree-like relationships among samples.

Because F_{ST} estimates in small populations are dependent on the sample size, for the F_{ST} comparisons among populations of *M. smejkalii* we randomly sampled ten individuals in population Z2 which had 30 individuals in total. We did not estimate F_{ST} among the populations of different species, as they have been completely geographically isolated since the last glacial period, which effectively prevents gene flow among them.

Testing of how genetic diversity was structured between species, geographic regions, populations and subpopulations was conducted using analysis of molecular variance (AMOVA). Because the data set was not balanced, in particular with the number of subpopulations and populations differing among regions and species, we first examined how the variance was distributed in the entire data set (all populations), including all hierarchical levels, and then we ran a separate AMOVA for the *M. smejkalii* data only (populations Z1–Z6, H; hierarchical levels—region, population, subpopulation), for the Želivka populations only (populations Z1–Z6; hierarchical levels population, subpopulation), and for populations Z2 and Z3 each (subpopulation).

Finally, we tested for the occurrence of recent population bottlenecks using the software BOTTLENECK 1.2.02 (Piry et al. 1999), assuming an infinite allele model of mutation which is the most suitable for biallelic SNP markers (Cornuet and Luikart 1996; Kogura et al. 2011). Because sample sizes were smaller than the recommended sizes except for population Z2, we only analyzed population Z2 on its own, and we also analyzed all Želivka populations as a single population to increase the power of the analyses. To determine whether observed population heterozygosity was significantly greater than that expected under mutation-drift balance, we applied the sign test and the standardized difference test based on 10,000 iterations.

Results

SNP loci characteristics

Out of the 119 genotyped individuals, nine were eliminated because they had over 75% missing data—one from *M. caespitosa* (population SK2), two from Z2 and six from population H. There were 12,599 potential SNP sites (variants), distributed over 2800 short reads (contigs). The average depth of coverage was 67.15 (standard error 2.43, Online resource 3). For the subsequent analyses, only loci with <2.5% missing data (null alleles) were retained, resulting in 1 143 polymorphic SNP sites (loci).

Identification of putatively adaptive loci

Out of the 1143 analyzed loci, a negligible number were under divergent selection among species and between the two regions of *M. smejkalii* (six and three respectively). We detected 20 loci among populations of *M. smejkalii* (1.7% of the total number of analyzed loci), and 34 among subpopulations (3.2%, Online resource 2) under putative divergent selection. We thus recalculated F_{ST} without outlier loci for populations and subpopulations only.

Population genetics and demography

Genotypic richness was at its maximal value in all populations since each individual had a unique multilocus genotype. The index of association was between 1.5 and 3 in the Želivka populations, and 10.8 in population H, suggesting higher linkage disequilibrium in the latter. In the other species, the index of association was between 3 and 4.5 for both *M. corcontica* and *M. caespitosa* (Table 1). Testing for Hardy–Weinberg equilibrium within each population showed that between 5 and 20% of the loci were not in equilibrium. This was likely due to heterozygote excess which was observed for all populations (negative F_{IS}) to different extents.

Assuming an infinite allele mutation model, the number of loci exhibiting heterozygosity excess relative to the expected heterozygosity under mutation-drift balance was significantly higher (p value < 0.001 for all statistical tests and all analyzed samples), suggesting the occurrence of a recent population bottleneck. Note that because of the low sample sizes for *M. corcontica*, *M. caespitosa* and *M. smejkalii* in the Hrnčíře region, testing for bottlenecks could only be done in the Želivka region and in population Z2.

Among-species differentiation

According to the K-means clustering, the most plausible number of clusters was two when all species are analyzed together (Online resource 3). With two clusters, the DAPC results showed a clear separation between M. smejkalii and the other species, with all specimens being correctly assigned to their designated species (Fig. 2). A biologically meaningful structure was also observed with four clusters: M. concortica, M. caespitosa, and the regions Želivka and Hrnčíře, although the clustering was not supported by the step-wise selection method of K. The SpitsTree clustered all individuals of M. corcontica and M. caespitosa in two sister clades separated from all *M. smejkalii* individuals (Fig. 3). The AMOVA results showed that most of the genetic variation was found within populations (individual level, Fig. 4). About 20% of the genetic variation was explained by the species, and population and subpopulation explained a small



Fig. 2 Results of K-means clustering with two clusters. First line below the x-axis corresponds to population names, second to species



Fig. 3 SplitsTree network delimiting species (*M. corcontica*, *M. caespitosa* and *M. smejkalii*). Within populations of *M. smejkalii*, samples from Hrnčíře are nested within samples from Želivka. The

population from Hrnčíře is strongly differentiated from the Želivka population, especially in comparison with the other species (*M. caespitosa*, *M. corcontica*)

(2.7 and 1.8%, respectively) but significant amount of the genetic variation.

Among-region and population differentiation

Global F_{ST} among populations of *M. smejkalii* was 0.036. Pairwise F_{ST} between populations were quite similar for all populations, ranging from 0.027 to 0.052, including population H which comes from a different region than the populations Z1–Z6. Removing the loci putatively under selection from the estimates resulted in lower global F_{ST} (0.022) and overall lower pairwise F_{ST} among populations (Table 2).

According to the K-means clustering, the best number of clusters for *M. smejkalii* populations were analyzed was one—the BIC for K=2 was not supported by the step-wise selection for K (Online resource 3). The SplitsTree clustered individuals of the Hrnčíře population in one clade, nested in the Želivka clade (Fig. 3). The AMOVA results showed that almost 77% of the total genetic variation was found within subpopulations, 18% was between regions, 1.9% among

Fig. 4 Percentage of variation explained by the embedded hierarchical levels (species, region, population, subpopulation) in different AMOVA. All species-analysis of the entire data set, M. smejkalii-only M. smejkalii populations, Želivka-only the six populations from Želivka, Z2 and Z3—only the populations divided into subpopulations (11 and two subpopulations of populations Z2 and Z3 respectively). Significant variance components are marked with an asterisk



Table 2	Pairwise population
F _{ST} estir	nates for all loci (above
the diag	onal) and for putatively
neutral	loci only (below
diagona	1)

	Z1	Z2	Z3	Z4	Z5	Z6	Н
Z1		0.039	0.043	0.043	0.048	0.038	0.052
Z2	0.035		0.027	0.028	0.035	0.033	0.037
Z3	0.040	0.027		0.035	0.038	0.039	0.040
Z4	0.036	0.023	0.031		0.041	0.036	0.038
Z5	0.038	0.026	0.030	0.033		0.034	0.047
Z6	0.034	0.026	0.036	0.030	0.030		0.045
Н	0.038	0.028	0.029	0.030	0.034	0.032	

To avoid comparing samples of different size, in population Z2 ten individuals (out of 28) were randomly drawn and used for the estimates

subpopulations, whereas the population did not have significant effect on genetic diversity partitioning (Fig. 4).

Within region differentiation (Želivka)

Within the region of Želivka, the K-means clustering suggested that the best estimate of structure was obtained for K=1, and the SplitsTree analysis did not detect any clustering by population. According to the AMOVA results, population and subpopulation were significant and explained 14.5% and 2.1% of the total variation, respectively (Fig. 4).

Discussion

Patterns of among-species genetic variation

Present-day conservation policies require a reliable taxonomic identification of the endangered taxa, especially for establishing conservation priorities when available resources are limited (McNeely 2002; Wege et al. 2015). The taxonomic status of the Czech endemic species *M. smejkalii* is not fully resolved, with a recent taxonomic study classifying it as a synonym of the widespread *M. verna* ssp. *gerardii* (Dillenberger and Kadereit 2014). However, Dillenberger and Kadereit (2014) did not mention having any sample from the Czech Republic for their analysis and likely thus did not have *M. smejkalii* in their material at all. The best way for proper distinction between *Minuartia* species is the morphology of seeds (Hejný and Slavík 2003). Comparison of the seed morphology of *M. verna* ssp. *verna* (Germany 300–400 m above sea level), *M. verna* ssp. *gerardii* (Austrian Alps, more than 1200 m above sea level), and *M. smejkalii, M. caespitosa*, and *M. corcontica* from this study under electron microscope showed that each subspecies/species had different seed morphology, especially in its protuberance (Pánková and Machač unpubl).

Our clustering analyses clearly separated all M. smejkalii individuals from M. corcontica and M. caespitosa, the other two species which originated from M. verna agg. and occur in the Czech Republic, suggesting an independent taxonomic status of M. smejkalii. The exact taxonomic level of M. smejkalii (species, subspecies, ecotype) needs yet to be determined in a global treatment of the whole *M. verna* agg., which is the goal of an ongoing study which uses very broad taxonomic sampling of Minuartia in central Europe (Jindřich Chrtek, pers. com.). Although this expanded sampling might produce results which do not warrant its endemic status in the Czech Republic, the differentiation of M. smejkalii from the other Minuartia species of the M. verna agg. in the Czech Republic supports its treatment as a separate entity in conservation programs (Montalvo and Ellstrand 2000; Hufford and Mazer 2003; McKay et al. 2005).

Patterns of within species genetic variation

Minuartia smejkalii is an endemic taxon with fragmented distribution, reduced number of populations, small population size, growing in an unfavorable, serpentine soil environment, which can result in high genetic drift and inbreeding, and strong local selective pressure. This should cause reduced within population genetic diversity, and possibly increased among-population genetic variation (Hamrick and Godt 1989; Ellstrand and Elam 1993; Linhart and Grant 1996; Whitlock et al. 2016). Our AMOVA results showed however moderate regional differentiation, low to moderate population differentiation within the Zelivka region, and high genetic variation within populations. Similar results have also been observed for other rare, endangered, and endemic species (Ellstrand and Elam 1993). For instance, Petunia secreta, a Brazilian endemic with a very similar distribution to M. smejkalii, showed genetic differentiation between two regions about 20 km apart, but not within regions, where the populations were < 0.2 km apart (Turchetto et al. 2016).

Genetic differentiation at relatively small geographic scales, such as those between the regions Želivka and Hrnčíře in this study, is commonly observed in serpentine species (Linhart and Grant 1996; Mattner et al. 2002). This regional differentiation could be caused by the lack of connectivity among serpentine patches between regions resulting in isolation by distance, and/or because of differences in regional environmental conditions, resulting in contrasting selective pressures. The levels of regional differentiation in M. smejkalii according to FST estimates in this study were between 0.028 and 0.052 (depending on the populations and loci included in the estimates), which is much lower than those found for other serpentine Minuartia species: 0.186 for M. laricifolia ssp laricifolia, 0.1–0.279 for M. laricifolia ssp ophiolitica, 0.249 for M. biflora and 0.895 for the selfing M. rubella (Borgen 1999; Moore et al. 2013). This could suggest some gene flow between regions in the past despite the ridge that separates the Zelivka and Hrnčíře sites. The different F_{ST} estimates can also be due to the different types of markers used in different studies with contrasting polymorphism, genome coverage or dominance levels (AFLP, Moore et al. 2013; allozymes, Borgen 1999). However, F_{ST} estimates are also influenced by the sampling method, population size, history, or mating system, which makes it difficult to conclude on the extent to which the marker type influence the estimates (Mariette et al. 2002; Leipold et al. 2018).

Low levels of regional genetic differentiation however are not necessarily excluding the possibility of local adaptation. Mosca et al. (2012) showed that some conifer populations differentiated by <5% according to molecular markers, could still show strong patterns of local adaptation in quantitative genetic traits. In line with this, we detected between 2 and 3% (depending on the levels of analysis) of outlier loci putatively under divergent selection suggesting the occurrence of local adaptation on the population level (nested within region). Interpreting outlier loci as adaptive should however be made with caution, as outliers can result also from low recombination rates due to non-evolutionary processes (Shafer et al. 2015). To confirm their adaptive value, the outlier loci need to be either correlated with fitness related traits, or their function needs to be identified using a reference genome (Fuentes-Pardo and Ruzzante 2017) which is not yet available for M. smejkalii or any other Minuartia species.

Genetic differentiation among the populations in the Želivka region was very low when all *M. smejkalii* species were analyzed with AMOVA, and moderate when only populations from Želivka were analyzed. Therefore, most of the genetic variation in our system, regardless of the AMOVA hierarchy was within populations. Similar results have also been observed for the Alpine populations of *M. laricifolia* spp. *laricifolia* (Moore et al. 2013). Several serpentine soil

specialist species also showed relatively low among-population genetic differentiation and high within-population genetic variation (Mengoni et al. 2000; Mattner et al. 2002; Quintela-Sabarís et al. 2010; Babst-Kostecka et al. 2014). More generally, population genetic differentiation of rare species is sometimes comparable and even greater than that of congeneric widespread species (Gitzendanner and Soltis 2000; Zawko et al. 2001; Ellis et al. 2006; Turchetto et al. 2016; Lanes et al. 2018) especially in perennial plants (Cabrera-Toledo et al. 2008).

High within- and low among-population genetic diversity can be a consequence of an outcrossing mating system and long life cycle (Nybom and Bartish 2000). Outcrossing species have in general F_{ST} values of 0.2 or less, (Mattner et al. 2002) and high within population genetic variation, in particular in dioecious or self-incompatible species (Cabrera-Toledo et al. 2008). Although M. smejkalii is self-compatible (Stojanova, pers.obs.), it does not have any mechanisms facilitating self-pollination: flowers isolated from pollinators tend to produce smaller seed sets than open pollinated flowers (Pánková, pers. obs.). Furthermore, some plants grown in a greenhouse show protogynous flowering patterns—early season flowers are mostly female, and hermaphrodite flowers develop later during the season (Stojanova, pers. obs.). Altogether, these results suggest a highly outcrossing mating system, which could explain the observed levels of genetic variation.

Another possible explanation for the observed patterns of genetic diversity in M. smejkalii is the effect of historical evolutionary processes. For instance, independent colonization of different sites, with multiple founder events or gene flow between serpentine and (now extinct) non-serpentine populations may have contributed to historically high levels of genetic variation in some serpentine species (Moore et al. 2013). This scenario seems unlikely for *M. smejkalii*, as no other related species (i.e. descendants from *M. verna* agg.) has been recorded in the region. High levels of genetic variation can also be maintained by genetic time lags, especially in long-lived species that only recently encountered habitat fragmentation (Cabrera-Toledo et al. 2008; Epps and Keyghobadi 2015). The habitat fragmentation of M. smejkalii began in the 1960, and the species could have had as little as four generations since then. Indeed, plants recorded in 2001 are still alive and reproducing at the time of this study, 18 years later. In line with this, the observed heterozygosity was higher than the expected heterozygosity in all populations, consistent with occurrence of a recent bottleneck. This suggests that the populations of M. smejkalii have still not reached a genetic equilibrium after the disturbance of their habitat, and the high levels of genetic diversity are likely a transient state (Cornuet and Luikart 1996). Thus, without conservation efforts the populations genetic diversity might decline in the future.

Conservation recommendations

Despite recent habitat fragmentation and population size reduction, *M. smejkalii* still maintains high levels of genetic diversity within populations, moderate regional differentiation and low to moderate among-populations differentiation. Determining the exact causes for the observed genetic structure requires further observations and experimental work to characterize the breeding system of the species in more detail, experimentally test for local adaptations, and detect potential effects or absence of selective pressure. It is, however, plausible that the high levels of genetic diversity are only a transient state following a recent population disturbance, and will not be maintained indefinitely if no conservation actions are undertaken.

The ex situ conservation efforts of *M. smejkalii* should focus on maintaining the observed genetic diversity while increasing population size and territory. This requires choosing appropriate seed source to enhance population size for each region. Based solely on the results of this study, it is difficult to unequivocally recommend a single conservation strategy.

Given the genetic differentiation between regions, we do not recommend mixing individuals from the regions of Želivka and Hrnčíře. Within a region, low among-population differentiation is taken as a justification for the use of transplants regardless of their population of origin, since they will not erode the local genetic structure of the extant populations (Cabrera-Toledo et al. 2008). However, patterns of molecular diversity are not necessarily related to adaptive traits (Reed and Frankham 2001; Kramer and Havens 2009). In line with this, we detected several outlier loci which could be compatible with patterns of divergent selection. Based on these observations, one might recommend a conservation strategy using exclusively material coming from the same population. A counter argument to this is the mating system of M. smejkalii, which is likely outcrossing. If this is indeed true, using only within-population/within-region transplants could increase the risk of inbreeding depression through biparental inbreeding. Experimental assays are currently being conducted to test for the existence of inbreeding and outbreeding depression in M. smejkalii, which will provide the necessary input for a reliable and efficient conservation strategy. In the meanwhile, if any conservation action is to be carried out, we recommend the use of individuals from the same region chosen to maximize the genetic diversity within populations.

Acknowledgements The study was supported by Project LIFE for Minuartia (Grant No. LIFE15NAT/CZ/000818), which is implemented with the financial contribution of the European Union, the Life program and with the financial contribution of the Ministry of Environment of the Czech Republic and party by and the long-term research development Project RVO 67985939. Data availability Genomic data is available at BioSample, ID: SUB5620362.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Agapow P-M, Burt A (2001) Indices of multilocus linkage disequilibrium. Mol Ecol Notes 1:101–102
- Babst-Kostecka AA, Parisod C, Godé C, Vollenweider P, Pauwels M (2014) Patterns of genetic divergence among populations of the pseudometallophyte *Biscutella laevigata* from southern Poland. Plant Soil 383:245–256
- Bilz M (2011) The IUCN Red List of Threatened Species. Minuartia smejkalii. The IUCN Red List of Threatened Species 2011
- Borgen L (1999) Genetic variation in *Minuartia* (Caryophyllaceae) in Svalbard. Nord J Bot 19:179–192
- Brady KU, Kruckeberg AR, Bradshaw HD Jr (2005) Evolutionary ecology of plant adaptation to serpentine soils. Annu Rev Ecol Evol Syst 36:243–266
- Brown AHD, Feldman MW, Nevo E (1980) Multilocus structure of natural populations of *Hordeum spontaneum*. Genetics 96:523–536
 Bushnell B (2014) BBMap: a fast, accurate, splice-aware aligner
- Cabrera-Toledo D, Gonzalez-Astorga J, Vovides AP (2008) Heterozygote excess in ancient populations of the critically endangered *Dioon caputoi* (Zamiaceae, Cycadales) from central Mexico. Bot J Linn Soc 158:436–447
- Cheptou P-O, Hargreaves AL, Bonte D, Jacquemyn H (2017) Adaptation to fragmentation: evolutionary dynamics driven by human influences. Philos Trans R Soc B 372:20160037
- Chytrý M, Danihelka J, Kaplan Z, Pyšek P (2017) Flora and vegetation of the Czech Republic. Springer, Cham
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014
- Dillenberger MS, Kadereit JW (2014) Maximum polyphyly: multiple origins and delimitation with plesiomorphic characters require a new circumscription of Minuartia (Caryophyllaceae). Taxon 63:64–88
- Dray S, Dufour A-B (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22:1–20
- Dray S, Chessel D, Thioulouse J (2003) Co-inertia analysis and the linking of ecological data tables. Ecology 84:3078–3089
- Ducousso A, Petit D, Valero M, Vernet P (1990) Genetic variation between and within populations of a perennial grass: *Arrhenatherum elatius*. Heredity 65:179
- Dvořáková M (1988) Minuartia smejkalii, eine neue Art aus der Minuartia gerardii-Gruppe (Caryophyllaceae). Preslia 60:1–9
- Dvořáková M (2003) Minuartia pauciflora, a Carpathian endemic species of the M-verna group. Preslia 75:349–356
- Ellis JR, Pashley CH, Burke J, McCauley DE (2006) High genetic diversity in a rare and endangered sunflower as compared to a common congener. Mol Ecol 15:2345–2355
- Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. Annu Rev Ecol Syst 24:217–242
- Epps CW, Keyghobadi N (2015) Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. Mol Ecol 24:6021–6040

- Fahrig L (2003) Effects of habitat fragmentation on biodiversity. Annu Rev Ecol Evol Syst 34:487–515
- Foll M, Gaggiotti OE (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180:977–993
- Fuentes-Pardo AP, Ruzzante DE (2017) Whole-genome sequencing approaches for conservation biology: advantages, limitations and practical recommendations. Mol Ecol 26:5369–5406
- Gitzendanner MA, Soltis PS (2000) Patterns of genetic variation in rare and widespread plant congeners. Am J Bot 87:783–792
- Grulich V (2012) Red List of vascular plants of the Czech Republic. Preslia 84:631–645
- Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding and genetic resources. Sinauer Associates, Sunderland, pp 43–63
- Harrison S, Rajakaruna N (2011) Serpentine: the evolution and ecology of a model system. University of California Press, Berkeley
- Hejný S, Slavík B (2003) Květena ČR, díl 2. Academia, Praha
- Hufford KM, Mazer SJ (2003) Plant ecotypes: genetic differentiation in the age of ecological restoration. Trends Ecol Evol 18:147–155
- Huson DH, Bryant D (2005) Estimating phylogenetic trees and networks using SplitsTree 4. Manuscr. Prep. Softw. www.split.org. Accessed 22 Dec 2019
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94
- 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
- Knaus BJ, Grünwald NJ (2017) VCFR: a package to manipulate and visualize variant call format data in R. Mol Ecol Resour 17:44–53
- Kogura Y, Seeb JE, Azuma N, Kudo H, Abe S, Kaeriyama M (2011) The genetic population structure of lacustrine sockeye salmon, *Oncorhynchus nerka*, in Japan as the endangered species. Environ Biol Fishes 92:539–550
- Kramer AT, Havens K (2009) Plant conservation genetics in a changing world. Trends Plant Sci 14:599–607
- Kruckeberg AR (1984) California serpentines: flora, vegetation, geology, soils, and management problems. University of California Press, Berkeley
- Kubát K (2002) Klíč ke květeně České republiky. Acad, Praha
- Lanes ÉC, Pope NS, Alves R, Carvalho Filho NM, Giannini TC, Giulietti AM, Imperatriz-Fonseca VL, Monteiro W, Oliveira G, Silva AR (2018) Landscape genomic conservation assessment of a narrow-endemic and a widespread morning glory from Amazonian savannas. Front Plant Sci 9:532
- Leipold M, Tausch S, Hirtreiter M, Poschlod P, Reisch C (2018) Sampling for conservation genetics: how many loci and individuals are needed to determine the genetic diversity of plant populations using AFLP? Conserv Genet Resour. https://doi.org/10.1007/ s12686-018-1069-1
- Lindenmayer DB, Fischer J (2013) Habitat fragmentation and landscape change: an ecological and conservation synthesis. Island Press, Washington, DC
- Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. Annu Rev Ecol Syst 27:237–277
- Lodhi MA, Ye G-N, Weeden NF, Reisch BI (1994) A simple and efficient method for DNA extraction from grapevine cultivars and Vitis species. Plant Mol Biol Report 12:6–13
- Lomolino MV, Riddle BR, Whittaker RJ (2017) Biogeography: biological diversity across space and time. Sinauer Associates Inc., Sunderland

- Mariette S, Le Corre V, Austerlitz F, Kremer A (2002) Sampling within the genome for measuring within-population diversity: trade-offs between markers. Mol Ecol 11:1145–1156
- Mattner J, Zawko G, Rossetto M, Krauss SL, Dixon KW, Sivasithamparam K (2002) Conservation genetics and implications for restoration of *Hemigenia exilis* (Lamiaceae), a serpentine endemic from Western Australia. Biol Conserv 107:37–45
- Mayer MS, Soltis PS, Soltis DE (1994) The evolution of the *Streptanthus glandulosus* complex (Cruciferae): genetic divergence and gene flow in serpentine endemics. Am J Bot 81:1288–1299
- McKay JK, Christian CE, Harrison S, Rice KJ (2005) "How local is local?"—a review of practical and conceptual issues in the genetics of restoration. Restor Ecol 13:432–440
- McNeely JA (2002) The role of taxonomy in conserving biodiversity. J Nat Conserv 10:145–153
- Mengoni A, Gonnelli C, Galardi F, Gabbrielli R, Bazzicalupo M (2000) Genetic diversity and heavy metal tolerance in populations of *Silene paradoxa* L. (Caryophyllaceae): a random amplified polymorphic DNA analysis. Mol Ecol 9:1319–1324
- Meusel H, Jäger E, Rauschert J, Weinert E (1965) Vergleichende Chorologie der zentraleuropäischen Flora. Jena
- Mimura M, Yahara T, Faith DP, Vázquez-Domínguez E, Colautti RI, Araki H, Javadi F, Núñez-Farfán J, Mori AS, Zhou S (2017) Understanding and monitoring the consequences of human impacts on intraspecific variation. Evol Appl 10:121–139
- Montalvo AM, Ellstrand NC (2000) Transplantation of the Subshrub Lotus scoparius: testing the home-site advantage hypothesis. Conserv Biol 14:1034–1045
- Moore AJ, Merges D, Kadereit JW (2013) The origin of the serpentine endemic *Minuartia laricifolia* subsp. *ophiolitica* by vicariance and competitive exclusion. Mol Ecol 22:2218–2231
- Morris GP, Hu Z, Grabowski PP, Borevitz JO, de Graaff M-A, Miller RM, Jastrow JD (2016) Genotypic diversity effects on biomass production in native perennial bioenergy cropping systems. Gcb Bioenergy 8:1000–1014
- Mosca E, Eckert AJ, Di Pierro EA, Rocchini D, La Porta N, Belletti P, Neale DB (2012) The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. Mol Ecol 21:5530–5545
- Novák FA (1960) Fylogenese serpentinových typů. [Phylogeny of serpentine types.]. Preslia 32:1–8
- Nybom H, Bartish IV (2000) Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. Perspect Plant Ecol Evol Syst 3:93–114
- Paradis E (2010) pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 26:419–420
- Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290
- Pešout P (2001) Final report: the rescue of Dolnokralovice serpentines and part of Vltava slopes in cooperation with owners. ČSOP, Vlašim
- Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered 90:502–503
- Quintela-Sabarís C, Vendramin GG, Castro-Fernández D, Fraga MI (2010) Chloroplast microsatellites reveal that metallicolous populations of the Mediterranean shrub *Cistus ladanifer* L have multiple origins. Plant Soil 334:161–174

- R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. Evolution 55:1095–1103
- Rodríguez-Quilón I, Santos-del-Blanco L, Serra-Varela MJ, Koskela J, González-Martínez SC, Alía R (2016) Capturing neutral and adaptive genetic diversity for conservation in a highly structured tree species. Ecol Appl 26:2254–2266
- Russello MA, Waterhouse MD, Etter PD, Johnson EA (2015) From promise to practice: pairing non-invasive sampling with genomics in conservation. PeerJ 3:e1106
- Sádlo J, Petřík P, Boublík K, Rychtařík P, Šímová I (2011) Diverzita rostlinstva Hradčanských stěn (Dokesko) a její příčiny [Habitats, vegetation and flora of the Hradčanské stěny rocks (Doksy region, northern Bohemia): causes of diversity]. Zpr Čes Bot Společ 46:17–38
- Shafer AB, Wolf JB, Alves PC, Bergström L, Bruford MW, Brännström I, Colling G, Dalén L, De Meester L, Ekblom R (2015) Genomics and the challenging translation into conservation practice. Trends Ecol Evol 30:78–87
- Słomka A, Sutkowska A, Szczepaniak M, Malec P, Mitka J, Kuta E (2011) Increased genetic diversity of Viola tricolor L. (Violaceae) in metal-polluted environments. Chemosphere 83:435–442
- Smith JM, Smith NH, O'Rourke M, Spratt BG (1993) How clonal are bacteria? Proc Natl Acad Sci 90:4384–4388
- Turchetto C, Segatto ALA, M\u00e4der G, Rodrigues DM, Bonatto SL, Freitas LB (2016) High levels of genetic diversity and population structure in an endemic and rare species: implications for conservation. AoB Plants. https://doi.org/10.1093/aobpla/plw002
- Wege JA, Thiele KR, Shepherd KA, Butcher R, Macfarlane TD, Coates DJ (2015) Strategic taxonomy in a biodiverse landscape: a novel approach to maximizing conservation outcomes for rare and poorly known flora. Biodivers Conserv 24:17–32
- Whitlock R, Hipperson H, Thompson DBA, Butlin RK, Burke T (2016) Consequences of in situ strategies for the conservation of plant genetic diversity. Biol Conserv 203:134–142
- Whittaker RH (1954) The ecology of serpentine soils. Ecology 35:258–288
- Wolf A (2001) Conservation of endemic plants in serpentine landscapes. Biol Conserv 100:35–44
- Wolf AT, Harrison SP (2001) Effects of habitat size and patch isolation on reproductive success of the serpentine morning glory. Conserv Biol 15:111–121
- Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. Trends Ecol Evol 11:413–418
- Zawko G, Krauss SL, Dixon KW, Sivasithamparam K (2001) Conservation genetics of the rare and endangered *Leucopogon obtectus* (Ericaceae). Mol Ecol 10:2389–2396

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