Inferring population structure and genetic diversity of the invasive alien Nootka lupin in Iceland

Jakub Skorupski,1,2 Magdalena Szenejko,1,2 Martyna Gruba-Tabaka,1,2 Przemysław Śmietana1,2 & Remigiusz Panicz3

1Institute of Marine and Environmental Sciences, University of Szczecin, Szczecin, Poland; 2Molecular Biology and Biotechnology Centre, University of Szczecin, Szczecin, Poland; 3Faculty of Food Sciences and Fisheries, Department of Meat Technology, West Pomeranian University of Technology in Szczecin, Szczecin, Poland

Abstract

Polar and subpolar regions are known for their particular vulnerability and sensitivity to the detrimental effects of non-indigenous species, which is well exemplified by the Nootka lupin (Lupinus nootkatensis) spread in Iceland. Since understanding the population and ecological genetics of invasive alien species offers hope for countering harmful biological invasions, the objective of the present study was to investigate interspecific variation in L. nootkatensis in Iceland in relation to a native population in Alaska. Moreover, we aimed to assess whether internal transcribed spacer 2 (ITS2) has sufficient phylogenetic applicability for a large-scale screening of the genetic diversity of a non-indigenous population of this species. This study, which is the first attempt to investigate the genetic diversity of the Nootka lupin in Iceland, included plant samples from eight locations in Iceland and one in Alaska. The analyses included genotyping by sequencing of the 417-nucleotide fragment of the 5.8S ribosomal RNA, ITS2 and part of the large subunit ribosomal RNA (GenBank MT026578-MT026580, MT077004). The main findings showed the presence of five previously unexplained single-nucleotide polymorphisms (SNPs); however, their discriminatory power for Icelandic populations was relatively low, since polymorphism information content (PIC) values ranged from 0.0182 to 0.0526, with average heterozygosity 0.0296. Concomitantly, analysis of multilocus genotypes (MLG) revealed sufficient differences in MLGs variants and their frequency to form genotypic patterns unique for Alaskan and Icelandic populations, revealing an internal genetic structure of the studied group. The proposed SNP panel needs to be supplemented with other nuclear and organellar markers.

Introduction

Invasive alien species pose serious ecological and economic problems (Thuiller et al. 2008; Skorupski et al. 2017). The harm to biodiversity from biological invasions results from the fact that they exert an array of unpredictable effects—both on individual species as well as on biocenotic systems (Richardson et al. 2000; Gallardo et al. 2019). This fact, in conjunction with the specific features of polar and subpolar ecosystems, such as relatively low diversity of native species, low temperatures that reduce the effectiveness of ecological homeostatic mechanisms, simplified trophic terrestrial networks, high rate of temperature increase due to climate change and increasing human activities and isolating geographical barriers, clearly indicates their...
particular vulnerability and sensitivity to the detrimen-
tal effects of a non-indigenous species spread (Xu et al.
2013; Bennett et al. 2015; Stefansson et al. 2016; CAFF
& PAME 2017). A good example illustrating this prob-
lem is the Nootka lupin (Lupinus nootkatensis Donn ex
Sims, 1810), native to Alaska and north-east Canada
and intentionally introduced to Iceland in 1885
(Schierbeck 1886). In the mid-20th century, the Nootka
lupin became naturalized on the island and is now
widely spread across the country, posing serious ecolog-
ical threats to native flora and fauna by displacing cold-
adapted native plant species, and negatively impacting
native pollinator communities (Bjarnason 1981; Wąsowicz
et al. 2013; Willow et al. 2017; Vetter et al.
2018; Guðjohnsen & Magnússon 2019).

Population and ecological genetics research offers
hope for counteracting harmful biological invasions
(Ward et al. 2008; Lawson Handley et al. 2011).
Knowledge of the genetic structure of non-indigenous
plant populations is helpful in determining their inva-
siveness potential, tracking the sources and routes of
invasions; forecasting their scale and assessing the genetic
consequences of invasions (Ward et al. 2008; Lawson
Handley et al. 2011; Harvey-Samuel et al. 2017; Wang
et al. 2017).

In this study, we tested part of the 18S–26S nrDNA
region, containing a fragment of the 5.8S ribosomal
RNA, ITS2 and part of the large subunit ribosomal RNA
(hereinafter referred to as the ITS2 region), as an
inter-population genetic variation marker of invasive
populations of the Nootka lupin in Iceland. This nuclear
DNA region has been proven to be a useful phylogenetic
marker in plants, helpful in resolving evolutionary rela-
tionships and taxonomic problems (Chen et al. 2010;
Han et al. 2013; Mishra et al. 2016). High intraspecific,
discriminatory ability (enough to distinguish even closely
related species), short length, high efficiency for PCR
amplification, high copy number of rRNA genes flanked
by well-conserved rRNA genes, and the fact that ITS2 is
not expressed, indicate its potential as a standard DNA
barcode available for seed plants identification (Alvarez
& Wendel 2003; Chase et al. 2007; Mäder et al. 2009; Yao
et al. 2010; Li et al. 2015; Mishra et al. 2016). Many authors reported its
utility for evolutionary studies in various plant groups
(Alvarez & Wendel 2003; Sonnante et al. 2003; Nieto-
Feliner & Rosselló 2007; Hughes et al. 2006; Chen et
al. 2010; Feng et al. 2016; Wang et al. 2016; Qin et al.
2017; Zhao et al. 2018; Duan et al. 2019), including represen-
tatives of the genus Lupinus (Käss & Wink 1997; Ainouche
et al. 2004; Eastwood et al. 2008; Mäder et al. 2009). At
the same time, the degree of sequence variation of the
ITSs region at the population level is, in the case of some
species, high enough to assess inter-population genetic
diversity (Yuan & Küpfer 1995; Desfeux & Lejeune
1996; Kollipara et al. 1997; Ainouche & Bayer 1999; Mäder
et al. 2009). This also applies to some lupin species, but
no data are available for L. nootkatensis in this respect
(Mäder et al. 2009).

Resolving the complex history of Nootka lupin, an
invasive species translocated to Iceland from its native
Alaska poses serious difficulties, and preliminary popula-
tion studies are needed urgently. Therefore, the main aim
of the study was to investigate interspecific genetic varia-
tion in L. nootkatensis in Iceland in comparison to plants
from Alaska. Moreover, we aimed to assess whether ITS2
has sufficient phylogenetic applicability for a large-scale
screening of the genetic diversity of non-indigenous pop-
ulations of Nootka lupin. To our knowledge this is the
first study to investigate the population structure of L.
nootkatensis in Iceland.

Material and methods

DNA was extracted from the dried leaves of 54 specimens
of Nootka lupin, from eight locations in Iceland and one
in Alaska, USA (as an out-group), collected in the sum-
mers of 2017 and 2019 (Fig. 1, Table 1). Total DNA was
isolated using the Syngen Plant DNA Mini Kit from
Syngen Biotech, following the manufacturer's protocol.
DNA concentration was measured by spectrophotometric

To amplify the target sequence a standard PCR was
used, with the following primers, which were designed
based on the previously recognized GenBank sequences:
MG236533.1 sequence of the Nootka lupin: forward
primer 5’-CCGTGAACCATCGAGTCTTT-3’ and reverse
primer 5’-ATTCCTCATGTTGGGCTTTTC-3’. Polymerase
chain reaction amplification was in a 20 µl volume con-
taining 3 µl (30 ng/µl) of template DNA, 2X NZT Taq PCR
Kit from EURx, 0.55 µM of each primer and deionized
water. The PCR reactions were performed in a T100TM
Thermal Cycler (Bio-Rad). The amplification profile
consisted of initial denaturation at 95°C for 5 min, followed
by 35 cycles of denaturation at 96°C for 10 s, annealing at
50°C for 10 s and elongation at 72°C for 30 s, and final
extension at 72°C for 5 min.

Amplified products were tested in a 1.5% agarose gel
for electrophoresis with SimplySafe™ (EURx) added in
TAE buffer. To visualize and document the results, the Gel
DocTM XR+ system was used (Bio-Rad). Polymerase
chain reaction products were subjected to Sanger
sequencing, carried out from both 5’ and 3’ ends (3730xl
DNA Analyzer from Applied Biosystems).

In order to detect variable sites obtained sequences
the products were subjected to a multiple alignment
by M-Coffee software (Wallace et al. 2006). Genetic diversity within the ITS2 region sequence was assessed on the basis of $\pi$ (Nei 1987), $k$ (Tajima 1983) and $\theta$, reflecting the expected resources of neutral variation in the population (Nei 1987), using DnaSP version 6 software (Rozas et al. 2017). To assess their conclusiveness and informativeness as a potential intraspecific genetic variation marker, $n_e$ (Kimura & Crow 1964), $H_{obs}$, PIC (Botstein et al. 1980) and $I$ (Shannon & Weaver 1949; Lewontin 1972) were calculated in POPGENE version 1.32 (Yeh & Boyle 1997) and Gene-Calc (Bińkowski & Miks 2018) software for each detected variable site.

The cumulative informative (discriminatory at the intraspecies level) value of all identified SNPs was estimated utilizing Haplotype Analysis version 1.05 (El Mousadik & Petit 1996) based on the analysis of MLGs, in which the number of expected and observed genotypes, the number of private genotypes, $R_e$ (El Mousadik & Petit 1996), $H_s$ (Finkeldey & Murillo 1999), $D_{st}$ (Finkeldey & Murillo 1999), $H_r$ (Finkeldey 1994) and $G_{st}$ (Nei 1973), were used.

Using the POPGENE version 1.32 package (Yeh & Boyle 1997), the internal genetic structure of the studied group was checked by calculating Wright’s $F$-statistics (for all variable sites), including $F_{is}$, $F_{st}$ and $F_{it}$ (Wright 1978; Hartl & Clark 1989). Gene flow between populations was estimated based on the $N_{m}$ indicator, estimated from $F_{st} = 0.25(1 - F_{st})/F_{st}$ (Wright 1931). Additionally, Tajima’s neutrality test and the Ewens-Watterson homozygosity test for each segregating site were performed to detect the possible effects of selection on inter-population allele distribution (Tajima 1989; Hedrick 2011). The first test was performed in DnaSP version 6 software, while the second in POPGENE version 1.32 software. LD and Hardy-Weinberg equilibrium were tested using Ohta’s two-locus analysis for subdivided populations (Ohta 1982) and POPGENE version 1.32 software. POPGENE version 1.32 was also used to test the internal genetic subdivision of the studied group. To illustrate genetic relations among populations, a dendrogram based on Nei’s (1972) genetic distance, calculated in POPGENE version 1.32 software, was constructed by the UPGMA method in MEGA version 10.1.7 software (Kumar et al. 2018).
A chi-squared statistics test was used to assess the significance of Hardy-Weinberg disequilibrium, as well as results of the Ohta’s test (POPGENE version 1.32). Differences in genotypes frequencies calculated for different populations were tested for statistical significance by the method developed by Zar (2009), using Student’s t-test.

Results and discussion

A 417 bp gDNA sequence of the Nootka lupin ITS2 region was submitted to GenBank under accession number MT026578-MT026580 and MT077004. Within this nucleotide sequence five variable sites were detected: four transitions (10C>T, 15C>T, 88C>T, 332G>A) and one transversion (167C>A; nucleotide numbering according to the sequence MT077004). The proportion of polymorphic sites is equal to 0.012. There are no other studies describing polymorphisms in the 18S–26S nrDNA sequence in the Nootka lupin. However, Aïnouche & Bayer (1999) reported 37 variable sites in the ITS2 for 44 Lupinus taxa, of which 17 are potentially informative. The same authors identified seven variable sites, including four that were potentially informative, in the 5.8 cistron in 25 Lupinus taxa. Eastwood et al. (2008) reported 169 variable sites, including 131 parsimony informative sites, in the 5.8S subunit and flanking internal transcribed spacers ITS1 and ITS2 in Lupinus genus. The obtained values of indexes measuring the DNA polymorphism—π = 0.0004, k = 0.1828, θ = 0.0023—indicate an excess of low-frequency polymorphisms and low genetic diversity in the studied group (Tajima 1989; Goodall-Copestake et al. 2012).

Tajima’s D value is equal to −1.6986, which may suggest either recent population expansion or purifying selection (Tajima 1989; Nei & Kumar 2000), although without statistical significance (0.10 > p > 0.05). The Ewens-Watterson test for neutrality of individual segregating sites showed that the F value (sum of square of allelic frequency) in no case exceeded the lower and upper limit of the 95% confidence region, the expected F value in the five variable sites that were analysed (Fig. 2).

Such results indicate that the Hardy-Weinberg homozygosity in a given sample is consistent with the equilibrium homozygosity under neutral theory, and there is no differential selection against the allelic variants (Watterson 1977; Hedrick 2011). These findings are confirmed by results of the Ohta’s two-locus analysis (Ohta 1982). The obtained average values for the components of the total variance of di locus LD (D12 = 0.0002, D12 = 0.0139, D12 = 0.0143, D12 = 0.0002) meet the condition D12>D12 and D12>D12 indicating that the relationships among variable sites result from limited migration and genetic drift and not from epistatic natural selection. Only polymorphic site 88C>T exhibited Hardy-Weinberg disequilibrium (−2 = 38.5714, df = 8, p < 0.0001). All studied groups were in the multilocus Hardy-Weinberg equilibrium, as no statistically significant departures from the Hardy-Weinberg rule were found.

The values of basic indicators characterizing an inter-population differentiating power of the segregating sites detected within the ITS2 region are summarized in the Table 2. The obtained average value of observed heterogeneity, PIC and I, equal to 0.0296, 0.0355 and 0.0902, respectively, indicate the low usability of the identified single SNPs as potential genetic markers and the low variety of their alleles in the studied group (Hildebrand et al. 1992). According to a commonly accepted interpretation, the five identified segregating sites can be categorized as a single nucleotide polymorphism rather than mutations, as their incidence exceeds the 1% threshold (Aggrey & Okimoto 2003) and equals 6% for 10C>T, 4% for 88C>T and 167C>A, and 2% for 15C>T and 332G>A in the studied group. Accordingly, data on a single variable site and combined five variable sites showed the highest DNA polymorphism at 10C>T, while the lowest polymorphism was found in the case of 15C>T and 332G>A. Nevertheless, all of them separately are far from the critical limit of heterozygosity determined to 70%, above which the marker is considered highly polymorphic (Ott 1992). Values of θ, are, in the case of each variable site, close to 1, which is the lowest possible value for diallelic loci. Such a situation occurs when one allele dominates in terms of frequency over the other that is very rare (Weir 1990).
As SNPs are diallelic markers, PIC value of a single SNP cannot exceed 0.5 (Kruglyak 1997; Kawuki et al. 2009). It is therefore recommended that several SNP sites are considered simultaneously (SNP panels, MLGs, haplotypes), since the cumulative level of polymorphism is usually increased (Kawuki et al. 2009). This has been proven for many plant species, including rye (Secale cereale L.), grapevine (Vitis vinifera L.), maize (Zea mays L.) and cassava (Manihot esculenta Crantz) (Ching et al. 2002; Salmaso et al. 2004; Hamblin et al. 2007; Varshney et al. 2007; Kawuki et al. 2009). For each variable site two alleles were detected in the present study, while the number of identified monolocus genotypes varies from two, in case of 10C>T, 15C>T, 167C>A and 332G>A, to 3, in case of 88C>T. Genotypes and alleles frequencies are summarized in Table 3.

Multilocus analysis revealed six observed MLGs, while the number of expected MLGs is 48. Populations 1, 2, 4, 6, 7 and 8 are monogenous and share the same ITS2 region genotype (GNTP1), while populations 3, 5 and 9 are polygenous and are characterized by four private genotypes (Table 4). The most common genotype is GNTP1; genotypes GNTP2 and GNTP5 are nearly 24 times less frequent; while GNTP3, GNTP4 and GNTP6 are 47 times less frequent. The overall picture of inter-population genotype differences is demonstrated by genotype frequency patterns of all the identified variable sites for specific populations. This indicates that genotypes other than GNTP1 constitute as much as 75% of the genotypes found in the Alaskan population, 30% in Icelandic populations 3 and 11% in Icelandic population 5. The average frequency of these genotypes for the Icelandic populations is 8%. The overall difference between genotype frequency patterns between plants from Alaska and plants from Iceland (grouped together) is statistically significant (t = 4.622, df = 5, p = 0.0057). Pairwise comparisons reveal statistically significant differences between populations 1 and 3 (t = 3.289, df = 10, p = 0.008), 1 and 9 (t = 4.622, df = 5, p = 0.0057), 2 and 3 (t = 3.289, df = 10, p = 0.008), 2 and 9 (t = 4.622, df = 5, p = 0.0057), 3 and 4 (t = 3.289, df = 10, p = 0.008), 3 and 6 (t = 3.289, df = 10, p = 0.008), 3 and 7 (t = 3.463, df = 10, p = 0.006), 3 and 8 (t = 3.463, df = 10, p = 0.006),

<table>
<thead>
<tr>
<th>Variable site</th>
<th>n</th>
<th>PIC</th>
<th>Hobs</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>10C&gt;T</td>
<td>1.0571</td>
<td>0.0526</td>
<td>0.0556</td>
<td>0.1269</td>
</tr>
<tr>
<td>15C&gt;T</td>
<td>1.0187</td>
<td>0.0182</td>
<td>0.0185</td>
<td>0.0526</td>
</tr>
<tr>
<td>88C&gt;T</td>
<td>1.0571</td>
<td>0.0526</td>
<td>0.0185</td>
<td>0.1269</td>
</tr>
<tr>
<td>167C&gt;A</td>
<td>1.0377</td>
<td>0.0357</td>
<td>0.0370</td>
<td>0.0922</td>
</tr>
<tr>
<td>332G&gt;A</td>
<td>1.0187</td>
<td>0.0182</td>
<td>0.0185</td>
<td>0.0526</td>
</tr>
<tr>
<td>Mean</td>
<td>1.0379</td>
<td>0.0355</td>
<td>0.0296</td>
<td>0.0902</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.0192</td>
<td>0.0172</td>
<td>0.0166</td>
<td>0.0372</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotype</th>
<th>10C&gt;T</th>
<th>15C&gt;T</th>
<th>88C&gt;T</th>
<th>167C&gt;A</th>
<th>332G&gt;A</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>0.9444</td>
<td>0.9815</td>
<td>0.9630</td>
<td>0.9630</td>
<td>–</td>
</tr>
<tr>
<td>CT</td>
<td>0.0556</td>
<td>0.0185</td>
<td>0.0185</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TT</td>
<td>–</td>
<td>–</td>
<td>0.0185</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CA</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.0370</td>
<td>–</td>
</tr>
<tr>
<td>GG</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.9815</td>
</tr>
<tr>
<td>GA</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.0185</td>
</tr>
<tr>
<td>Allele</td>
<td>0.9722</td>
<td>0.0278</td>
<td>0.9907</td>
<td>0.0093</td>
<td>0.9722</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>1-8</th>
<th>9</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNTP1</td>
<td>5'-CC-CC-CC-GG-3'</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.7000</td>
<td>1.0000</td>
<td>0.8889</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9200</td>
<td>0.2500</td>
</tr>
<tr>
<td>GNTP2</td>
<td>5'-CC-CC-CC-GG-GG-3'</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.2000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0400</td>
<td>0.0000</td>
</tr>
<tr>
<td>GNTP3</td>
<td>5'-CT-CC-CC-GG-3'</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>GNTP4</td>
<td>5'-CC-CC-GG-GG-3'</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>GNTP5</td>
<td>5'-CT-CC-CC-GG-GG-3'</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0200</td>
<td>0.2500</td>
</tr>
<tr>
<td>GNTP6</td>
<td>5'-CT-CC-CC-GG-GG-3'</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1111</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0200</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Citation: Polar Research 2020, 40, 4536, http://dx.doi.org/10.33265/polar.v40.4536

(page number not for citation purpose)
Genetic differentiation is illustrated by compounds of Nei’s (1973) genetic statistics, including $H_T$, $H_S$, $D_S$, and $G_{ST}$, calculated for each variable site and averaged for individual populations. Our results, summarized in Table 5, indicate that 27.2% ($G_{ST}$) of the total genetic diversity in the Nootka lupin, measured by identified ITS2 region SNPs, came from genetic diversity between populations, while 72.8% came from genetic diversity among plants within populations. For all grouped Icelandic populations these values were 13.4% and 86.6%, respectively, while for the Alaskan population they were 28.7% and 71.3%, respectively. These values are indicative of low genetic differentiation between the sampled populations. This is supported by the values of the total genetic diversity distributed among populations ($D_{AS}$), which range from 0.0016 to 0.0116 for the populations from Iceland (0.0216 for grouped non-indigenous populations) and is 0.0224 for the population from Alaska. Population differentiation may be, in this case, reduced by their isolation from each other and resulting limited migration (White et al. 2007). On the other hand, differences between populations from Iceland are much less pronounced than differences between all analysed non-indigenous populations taken together and the population of Alaska. This, as further indicated by the genotypic richness, expressing the number of genotypes found in a population corrected for sample size, is over 9.5 times greater for the indigenous Alaskan population than for the grouped populations sampled in Iceland. Similarly, the mean genetic distance between individuals from Alaska is over 2.9 greater than that between plants from Iceland.

A somewhat similar pattern of intra- and inter-population genetic diversity was found in the white lupin (*Lupinus albus* L.), for which 92% of allelic diversity was attributed to individuals within populations, while allelic diversity distributed among populations amounts to only 8% (Atnaf et al. 2017). Our results, supported by the abovementioned observation, have important implications for further genomic analyses, including studies involving high-throughput sequencing and genomic selection.

The average genetic distance between the plants ($D$) was 0.2, with the lowest values in the completely monomorphic (in terms of ITS2 region sequence) populations 1, 2, 4, 6, 7 and 8. The maximum genetic distance between individuals was registered in Eyrrarbakki ($D = 0.7333$). Interestingly, value of the mean genetic distance between individuals of the Icelandic invasive populations (0.2279) is almost equal to that obtained by Vyšniauskienė et al. (2011) for non-indigenous populations of the large-leaved lupin (*Lupinus polyphyllus* Lindl.) in Lithuania (0.272), calculated for RAPD polymorphism in 192 plants.

The mean value of the fixation index, measuring the difference between the expected and observed heterozygosity of populations compared to the total analysed group is 0.1522, which indicates a moderately high level of genetic differentiation among populations and structuring of the studied group (Wright 1978). The value of $F_{ST} = 0.0366$ suggests a slight excess of homozygotes within populations, while $F_{IS} = 0.1832$ indicates the same for the entire studied group. Both indicators show a moderate departure from the Hardy-Weinberg equilibrium across all populations and within all studied populations taken together (Balloux et al. 2003). Results of the Wright’s $F$-statistics are summarized in Table 6. Increased frequency of homozygotes, at the level of a single population and at the level of the whole studied group, is regarded as evidence for subdivision of the latter (Wahlund 1928). However, the deviation from the Hardy-Weinberg equilibrium described above is small enough (deviation at an

### Table 5 Results of genetic differentiation analysis of MLGs for the ITS2 region in *L. nootkatensis*.

<table>
<thead>
<tr>
<th>Population</th>
<th>$A_i$</th>
<th>$D_i$</th>
<th>$H_S$</th>
<th>$D_{ST}$</th>
<th>$H_T$</th>
<th>$G_{ST}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0040</td>
<td>0.0040</td>
<td>1.0000</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0040</td>
<td>0.0040</td>
<td>1.0000</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.7333</td>
<td>0.0852</td>
<td>0.0116</td>
<td>0.0968</td>
<td>0.1201</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0071</td>
<td>0.0071</td>
<td>1.0000</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.4000</td>
<td>0.0329</td>
<td>0.0064</td>
<td>0.0394</td>
<td>0.1638</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0063</td>
<td>0.0063</td>
<td>1.0000</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0016</td>
<td>0.0016</td>
<td>1.0000</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0016</td>
<td>0.0016</td>
<td>1.0000</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>0.6667</td>
<td>0.0556</td>
<td>0.0224</td>
<td>0.0779</td>
<td>0.2871</td>
</tr>
<tr>
<td>Mean</td>
<td>1.6667</td>
<td>0.2000</td>
<td></td>
<td>0.2724</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Number of different genotypes detected in each population. *Mean genetic distance between individuals. *Grouped Icelandic populations.

### Table 6 Summary of the F-statistics at all variable sites.

<table>
<thead>
<tr>
<th>Variable site</th>
<th>$F_{ST}$</th>
<th>$F_{IS}$</th>
<th>$F_{IR}$</th>
<th>$N_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−0.1013</td>
<td>−0.0263</td>
<td>0.0681</td>
<td>3.4195</td>
</tr>
<tr>
<td>2</td>
<td>−0.0588</td>
<td>−0.0062</td>
<td>0.0497</td>
<td>4.7812</td>
</tr>
<tr>
<td>3</td>
<td>0.4667</td>
<td>0.6522</td>
<td>0.3478</td>
<td>0.4688</td>
</tr>
<tr>
<td>4</td>
<td>−0.1111</td>
<td>−0.0112</td>
<td>0.0899</td>
<td>2.5313</td>
</tr>
<tr>
<td>5</td>
<td>−0.0588</td>
<td>−0.0062</td>
<td>0.0497</td>
<td>4.7812</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0366</td>
<td>0.1832</td>
<td>0.1522</td>
<td>1.3927</td>
</tr>
</tbody>
</table>

Citation: Polar Research 2020, 40, 4536, http://dx.doi.org/10.33265/polar.v40.i4536
absolute value of 3.4% for the inbreeding coefficient, and 18.3% for the overall fixation index) that the studied populations stay close to what is expected under panmixia, or high rate of clonal reproduction and self-pollination (Balloux et al. 2003). Our results are therefore consistent with the findings of Baldursson (1995), who pointed out that *L. nootkatensis* in Iceland depends on cross-pollination to a lesser extent than self-fertilization, which accounts for 70% of all pollination cases.

The number of migrant gametes among populations per generation (*N_m*) is moderately low, amounting to 1.3927. As there are obvious geographical barriers between the Alaskan and Icelandic populations (and assuming that the significance of modern introductions is negligible), there is much more conclusiveness in the gene flow rate between non-indigenous populations, which may reflect their inter-population genetic differentiation due to dispersal on newly occupied territories. Icelandic populations are characterized by an average value of gene flow equal to 1.6231, while the Alaskan figure is 0.6441. It has been shown that restricted dispersal is expressed by lower estimates of *N_m* and higher estimates of *F_ST* compared with wide dispersal (Bohonak et al. 1998). Whitlock & McCauley (1999) demonstrated that, since gene flow depends on the effective reproduction of migrants in the new location, the movement of genes does not necessarily have to be reflected by direct measures of dispersal. This may be particularly important in the case of invasive alien species, and for Icelandic populations it may be additionally explained by the planned and large-scale (e.g., aerial seeding by airplanes) use of Nootka lupin homogeneous seeds for revegetation and protection against soil erosion, carried out from the 1960s by the Soil Conservation Service of Iceland, the Iceland Forestry Service and the Icelandic Road and Coastal Administration (Thorsson & Hlidberg 1997; Olgeirsson 2007; Benediktsson 2015).

The UPGMA dendrogram (Fig. 3), constructed on the basis of Nei’s genetic distance values (summarized in Table 7), reveals the existence of four distinct clusters not
fully reflecting geographical proximity, which is usually a key factor influencing the genetic relatedness of populations (Wright 1943). The first cluster groups populations 1, 2, 4, 6, 7 and 8 (cluster I), the second groups populations 3 (cluster II), the third groups populations 5 (cluster III) and the fourth groups populations 9 (cluster IV). The last one represents plants from Alaska and, as expected, is of an out-group character, genetically most distant to all others. Post-introduction changes in genetic variability in invasive plant species, such as genetic drift, founder effects and responses to novel selection pressures, create differences between individuals in the natural range of the species and those introduced from outside that natural range, and also lead to changes in the genetic structure of introduced populations (Stout et al. 2015). It can be expressed by lower values of genetic diversity indices ($H_t$ and $H_e$ in this study) in these populations compared to natural ones, as shown by Wilson et al. (2009), Jogesh et al. (2015), Estoup et al. (2016) and Smith et al. (2020).

Clear differences between the Alaskan and Icelandic plants (populations 1–8 considered together) concern genotype frequency patterns, indicators of genetic differentiation analysis of MLGs, as well as Wright’s $F$-statistics. Differences between populations from Iceland are much less expressed and account for only 5.6% of the genetic distance between plants from Alaska and plants from Iceland. Complete homogeneity and genetic identity with respect to polymorphic sites within the ITS2 region of plants in cluster I can be explained by the fact that the locations of most of its populations (populations 2, 7 and 8 in land reclamation areas and populations 4 and 6 along roadsides) coincide with areas where planned sowing was carried out over the years by the Icelandic state (Olgeirsson 2007; Guðjohnsen & Magnússon 2019). It can be expected that large-scale use of homogeneous seed material results in founder effects (Ward et al. 2008). In turn, the reasons for genetic distinctness of cluster II and III can be seen in the location of sites occupied by populations 3 and 5. In both cases, of key importance can be the vicinity of urban centres (the Greater Reykjavik area in the case of population 3 and Stykkishólmur in the case of population 5), and the use of a variety of seed material by individual owners of home gardens planting them with the Nootka lupin for decorative purposes (Benediktsson 2015).

**Conclusion**

The main finding of the present study is detection of five previously unknown single nucleotide polymorphisms in the ribosomal DNA spacer 2 regions in the Nootka lupin. Their usefulness to detect genetic variation within non-indigenous populations in Iceland is very limited (measured for individual SNP and SNP haplotypes) because of the low discriminatory power of the identified SNPs at an inter-population level and the high genetic homogeneity of Icelandic populations. At the same time, the genetic distance between plants from Alaska and plants from Iceland is sufficient to identify the non-indigenous population. This is best illustrated by the average genetic distance, which is 35.8 times greater between the tested Alaskan population and the grouped Icelandic populations than between the Icelandic populations; also by the differences in MLG variants and their frequency, forming genotypic patterns unique for each indicated group.

This study is the first attempt to investigate the genetic diversity of *L. nootkatensis* in Iceland. To establish a molecular tool that can screen the genetic diversity of non-indigenous populations of this species on a large scale, the SNP panel should be supplemented with polymorphisms in the sequences of other genome regions. The discriminatory power of combined nuclear and organellar markers over markers based on an individual genomic region has been proven for many species (Gamache et al. 2003; Heuertz et al. 2004; Tollefsrud et al. 2009). Further studies are needed to develop a multilocus marker of inter-population, genetic variation

---

**Table 7** Pairwise population matrix of Nei’s genetic distance.

<table>
<thead>
<tr>
<th>Population</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>–</td>
<td>0.0000</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0000</td>
<td>–</td>
<td>0.0022</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0022</td>
<td>0.0000</td>
<td>–</td>
<td>0.0022</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0034</td>
<td>0.0014</td>
<td>0.0000</td>
<td>0.0014</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.0014</td>
<td>0.0014</td>
<td>0.0034</td>
<td>0.0014</td>
<td>0.0000</td>
<td>0.0014</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0022</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0014</td>
<td>0.0000</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0022</td>
<td>0.0000</td>
<td>0.0014</td>
<td>0.0014</td>
<td>0.0000</td>
<td>0.0000</td>
<td>–</td>
</tr>
<tr>
<td>8</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0022</td>
<td>0.0000</td>
<td>0.0014</td>
<td>0.0014</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0014</td>
</tr>
<tr>
<td>9</td>
<td>0.0314</td>
<td>0.0314</td>
<td>0.0341</td>
<td>0.0314</td>
<td>0.0334</td>
<td>0.0314</td>
<td>0.0314</td>
<td>0.0314</td>
<td>0.0314</td>
</tr>
</tbody>
</table>
of invasive populations of the Nootka lupin. The ongoing transition from genetics to genomics increases the popularity of genome-wide and reduced-representation techniques, such as a restriction site-associated DNA sequencing (RAD-seq) or high-throughput sequencing, allowing researchers to uncover a finer population structure than microsatellites do, with a smaller sample size (Jeffries et al. 2016; Stronen et al. 2019). A genomic approach has recently been used successfully to study the genetic structure of white lupin and narrow-leaved lupin (Lupinus angustifolius L.; Książkiewicz et al. 2017; Zhou et al. 2018; Hufnagel et al. 2020).

Despite the low value of the ITS2 region’s SNPs as inter-population genetic variation markers of invasive populations of the Nootka lupin in Iceland, they can help to explain the evolutionary processes to which this nrDNA region is subject to following introduction. Spatio-temporal demographic dynamics of L. nootkatensis, initiated by recent human-caused range changes, also may result in generating new patterns of genetic variation within and between populations from a plant’s native range and from invaded areas. Research results indicate the possible influence of differentiation due to isolation by geographic distance (limited migration), genetic drift and recent population expansion shaping the ITS2 region polymorphism in the new environment. This, in turn, leads to the conclusion that further investigation is necessary to identify a possible barcode overlapping in the case of the considered sequence in the Lupinus genus.

Acknowledgements

The authors wish to thank to Mr Sławomir Keszka, Ms Bożena Kościk and Mr Henryk Kościk who provided samples from Alaska, and the anonymous reviewers whose comments and suggestions helped improve this manuscript.

Disclosure statement

The authors report no conflict of interests.

Funding

Sampling in Iceland was possible thanks to the project Exchange of Knowledge, Experiences and Best Practices in Study and Control of the Invasive Alien Species Populations in Iceland and Poland, financed by the Financial Mechanism of the European Economic Area (EEA) 2009–2014 and the National Fund for Environmental Protection and Water Management (Poland).

References


Citation: Polar Research 2020, 40, 4536, http://dx.doi.org/10.33265/polar.v40.4536 (page number not for citation purpose)