Conserv Genet (2018) 19:43–55
DOI 10.1007/s10592-017-0995-x

RESEARCH ARTICLE

Genetic structure, diversity, and hybridization in populations of the rare arctic relict *Euphrasia hudsoniana* (Orobanchaceae) and its invasive congener *Euphrasia stricta*

Katharine J. Zlonis[1,2] · Briana L. Gross[2]

Received: 25 October 2016 / Accepted: 4 July 2017 / Published online: 8 July 2017

Abstract Arctic relict populations, which persist in disjunct locations far south of a species' normal range, are at the frontline of climate change and may be especially susceptible to the negative impacts of climate warming. Further, these relict populations may face increasing contact with, or become outcompeted by, invasive species if the invasive taxa are spreading along with the warming climate. Relict populations are simultaneously of particular conservation importance due to their unique genetic make-up and potential for adaptations to warmer temperatures compared to populations at the core of the species range. In this study, we used genotyping-by-sequencing to study the population genetics of *Euphrasia hudsoniana*, a polyploid arctic disjunct of conservation concern, at the southern edge of its range along the northwestern shore of Lake Superior. In addition, we examined evidence for hybridization with its invasive congener, *E. stricta*. Overall, we found clear differentiation between the native and invasive species indicated by nearly all analyses. There was limited evidence for gene flow from the invasive into the native species, but patterns were consistent with more extensive gene flow in the opposite direction. Differentiation among native populations was low, yet two of the five populations fell into a separate, distinct group based on STRUCTURE analyses. Continued genetic monitoring of these populations will help elucidate whether hybridization with invasives is a burgeoning threat for this arctic relict.

Keywords Arctic disjunct · Climate change · Polyploid · Genotyping-by-sequencing · Invasive · Hybridization

Introduction

The flora of the arctic region has been shaped by dramatic climatic changes and associated shifts in the ranges of its component species. Arctic and alpine taxa have evolved and migrated with advancing and retreating ice-sheets and changes in sea level (Abbott and Brochmann 2003; Billings and Mooney 1968; Davis et al. 2005). Populations that recolonized the newly exposed land came from areas south of the ice-sheets or from refugia that persisted further north (e.g., Beringia; Nimis et al. 1998). Following such dramatic range shifts, relict populations have persisted in refugia with appropriate microclimates, even after the bulk of the species' climate envelope has shifted northward (Hampe and Jump 2011). These populations are often small, due to restricted habitat, and separated from the majority of their conspecifics by hundreds or thousands of miles. Relict populations are thus subject to genetic drift and an inevitable reduction in diversity over time (Ellstrand and Elam 1993; Bauert et al. 1998). At the same time, these populations potentially represent 'old' gene pools, having persisted in-place for much longer than populations on the leading edge of the arctic migration front. Thus, relict populations might contain genetic information and adaptations that are lacking from the core of the species range and are of special conservation concern (Hampe and Petit 2005).

Electronic supplementary material The online version of this article (doi:10.1007/s10592-017-0995-x) contains supplementary material, which is available to authorized users.

✉ Katharine J. Zlonis
wink0113@d.umn.edu

[1] Leech Lake Division of Resource Management, 15756 State 371 Northwest, Cass Lake, MN 56633, USA

[2] Department of Biology, University of Minnesota Duluth, 1035 Kirby Drive, Duluth, MN 55812, USA

Beyond their value as repositories of genetic information, communities of relict species can also serve as early indicators of the impact of climate change on plant distribution and community composition. Relict species populations may be especially vulnerable to the effects of a changing climate because nearby habitat does not match their climate tolerance, making migration through the surrounding landscape to the core habitat unlikely. Thus, any change in suitable conditions could result in the extirpation of relict populations. In addition, arctic species at the southern margin of their range may be particularly susceptible to declines (Doak and Morris 2010; Lesica and McCune 2004). Habitat for arctic relicts may be further deteriorated if invasive species become more prevalent as the climate warms and displace native species (Thuiller et al. 2007). These processes can be seen as extensions of the climate shifts that initially created these relict populations, except that current climate change is occurring on a more rapid scale than previous shifts. The response of these arctic relict populations to climate shifts can therefore inform our understanding of the future of plant population dynamics in the face of climate change (Hampe and Jump 2011).

Euphrasia hudsoniana (Hudson Bay eyebright) is an annual, low-arctic species that displays the classic pattern of a core distribution and relict populations far outside of that core area (Fig. 1a, Given and Soper 1981). The species is mainly distributed in the northeastern provinces of Canada, surrounding Hudson Bay. However, remnant populations still exist much further south along the northern shore of Lake Superior in Minnesota (USA) and Ontario (Canada), and on Isle Royale in Michigan (USA). Populations are variously referred to as *E. hudsoniana* var. *contracta* or *E. hudsoniana* var. *ramosior* (together hereafter referred to simply as *E. hudsoniana*). Although its phylogeography has not been studied, the simplest explanation for the current distribution of *E. hudsoniana* is that the plant was a component of communities that persisted south of the ice sheet that covered North America during the last glacial maximum. As the ice sheets retreated and species moved north to their current distributions, relict populations remained in areas with suitable climate. Lake Superior generates microhabitats that mimic arctic conditions in the form of cool summer temperatures and high humidity due to frequent fog. In addition, ice thrust along the rocky shoreline regularly removes boreal and temperate forest species that thrive nearby and would normally outcompete *E. hudsoniana* populations (Given and Soper 1981). These microhabitats also house associated arctic relicts, for example, *Primula mistassinica* and *Pinguicula vulgaris*, which also rely on the localized climate around Lake Superior for their survival. Lake Superior is warming particularly rapidly and temperatures on the north shore of Lake Superior are expected to increase 1.6–2.9 °C in the next 35 years (2.9–4.4 °C by 2090), putting these communities at significant risk (Huff and Thomas 2014).

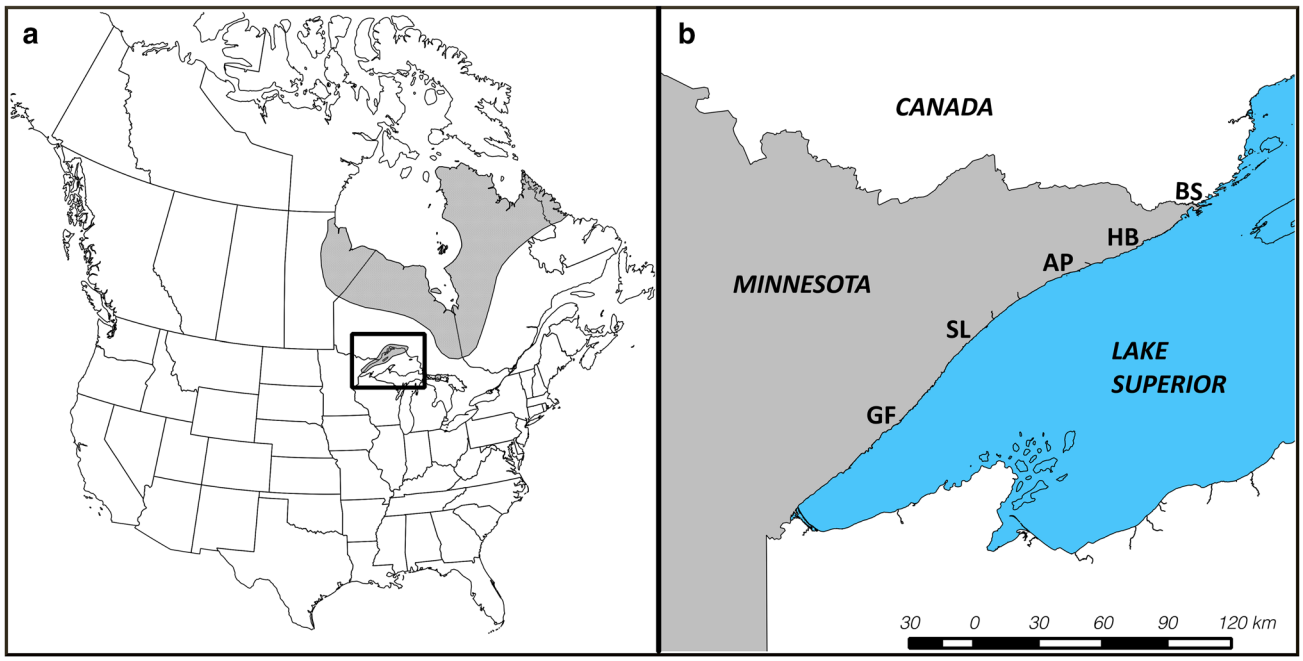

Fig. 1 Full (**a** *shaded area*, adapted from Given and Soper 1981 and Smith 1988) and disjunct range (**b**) of *Euphrasia hudsoniana*. Disjunct range along the northwestern shore of Lake Superior in Minnesota, USA is within the *bold rectangle* depicted in (**a**), and sampling locations are indicated with site name abbreviations

In addition to climate change, relict *E. hudsoniana* populations are threatened by competition and potential hybridization with an invasive congener. *Euphrasia stricta* is native to Europe and has been officially documented in the northeastern states and provinces of the US and Canada (USDA database), including Minnesota (MN DNR 2016). *E. stricta* is not restricted to arctic habitats, and is commonly observed in disturbed areas along trails, gravel roads, and openings in the forested regions of Minnesota (Reznicek et al. 2011; MN DNR 2016). The invasive is also found in close proximity to at least two relict populations of *E. hudsoniana* along the shore of Lake Superior, and hybridization between the native and invasive species has been suggested based on morphological examination of voucher specimens (L. Gerdes, personal communication).

The goal of this study was to characterize the population genetic structure of the relict populations of *E. hudsoniana* that occur along the northern shore of Lake Superior in Minnesota (USA). *Euphrasia* is a challenging species to work with in terms of genetic analysis; it has not been possible to germinate it in the greenhouse and it is a allotetraploid with no closely related reference genome species. However, it is still amenable to genotyping-by-sequencing (GBS) approaches, which can be used to generate SNP data regardless of ploidy level and without the requirement of reference genome for read assembly. Using the GBS approach, we examined the population genetic structure of five relict populations of *E. hudsoniana* and two populations of the invasive *E. stricta* that occur in close proximity with those populations. In particular, we sought to determine: (1) the extent of genetic diversity and differentiation among populations, (2) genetic relationships among populations in comparison with their physical locations along the shore and, (3) whether there is evidence for hybridization between the native and invasive *Euphrasia*.

Materials and methods

Study species

Euphrasia hudsoniana is an annual, tetraploid (G. Gusarova, personal communication), hemiparasitic herb native to North America (Fig. 1a). It is likely that *E. hudsoniana* is an allotetraploid, like other species in the genus (Gusarova et al. 2008), but this has not been investigated in detail for *E. hudsoniana*. Individuals of this species are small in stature; plants measured in 2014 at some overlapping sites as in this study were 6–8 cm tall, and had approximately 15 leaves that were 0.5 cm wide × 0.5 cm long, on average (Winkler and Etterson 2015). Its range is mainly in Canada, extending from the northern border of Quebec along the shores of Hudson Bay throughout the northern region of Ontario and west into Manitoba (Given and Soper 1981; Smith 1988). The southernmost extent of the species' range is in the northern US, where disjunct populations of *E. hudsoniana* are found along the north shore of Lake Superior in Minnesota and on Isle Royale in Michigan. It occupies gravelly habitats on the shores of lakes and streams, and on the shore of Lake Superior it grows in cracks in rocks immediately adjacent to the lake. As a small-flowered *Euphrasia* species (corolla length 4–6.5 mm), *E. hudsoniana* is thought to reproduce primarily via self-fertilization (French et al. 2005; Sell and Yeo 1970).

Euphrasia stricta is an annual allotetraploid (Gusarova et al. 2008) native to Europe (Downie and McNeill 1988). This species is spreading rapidly in northeastern Minnesota, and is nearly ubiquitous on disturbed forest roads, trails, and openings. *Euphrasia* taxonomy is complicated and some species can sometimes be difficult to distinguish in the field, but generally *E. stricta* is slightly larger and more robust than *E. hudsoniana*, has sharply dentate leaves with bristle tips, its leaves and stems are less hirsute, and it has larger corollas than *E. hudsoniana* (Reznicek et al. 2011; Sell and Yeo 1970). Work on *Euphrasia* taxonomy is ongoing at Flora of North America and we refer to the species currently recognized by the state of Minnesota, *E. hudsoniana* and *E. stricta*, throughout this paper.

We collected samples from populations of these species at five sites in Minnesota including Gooseberry Falls State Park (GF; in Castle Danger, MN), Sugarloaf Point Scientific and Natural Area (SL; in Schroeder, MN), Artist Point (AP; in Grand Marais, MN), Horseshoe Bay (HB; in Hovland, MN), and an unnamed site, referred hereafter as "Border Site," within 5 km of the Canadian border (BS) (Fig. 1b). *E. hudsoniana* is listed as a species of special concern in Minnesota thus we are unable to publish population coordinates; however, coordinates can be supplied in response to individual requests upon approval from the Minnesota DNR. Of these sites, *E. stricta* is found near two *E. hudsoniana* populations at AP and HB. At HB *E. stricta* grows along a gravel road leading to the cobble shore within approximately 50 m of the *E. hudsoniana* population. At AP, most identified *E. stricta* plants are growing within approximately 100 m of *E. hudsoniana* plants and there are several patches of plants at this site that appear phenotypically intermediate to the invasive and native species. One phenotypically intermediate plant (AP_04) was collected and labeled according to the species that it most resembled based on evaluation in the field (*E. hudsoniana*). Intermediate plants were not avoided in the collection strategy, but only one was collected in order to maintain even sampling across the rest of the population.

Sample collection, DNA extraction, and library preparation

To optimize the GBS (Elshire et al. 2011) restriction digest protocols for this study, we collected and processed a small set of leaf and stem tissues from *E. hudsonia*. These initial samples were collected from six individuals from a population in Two Harbors, MN and from two individuals from a population in Grand Marais, MN. Samples (10–80 mg) were stored on ice and in the refrigerator until being transported to the laboratory, frozen with liquid nitrogen, and ground. Genomic DNA was extracted using Qiagen Plant DNeasy isolation kits (Qiagen Inc., Valencia CA, USA). Samples were pooled to achieve a concentration of 50 ng/µl. Five hundred ng of DNA was digested with 5 U of *Hin*dIII restriction enzyme and run on a 0.8% agarose gel with uncut DNA and lambda *Hin*dIII standards. The pooled sample was sent to the Cornell University Institute of Biotechnology for test GBS library creation with two enzymes. The *Pst*I enzyme was selected for further sample processing.

Euphrasia hudsoniana tissue samples were collected haphazardly from five populations (Fig. 1b): 16 samples from GF, 15 samples from SL, 14 samples from AP, 12 samples from HB, and 14 samples from BS (total of 71 native samples). In addition, we collected invasive *E. stricta* tissue samples haphazardly from two sites: 12 samples each from AP and HB (total of 24 invasive samples). Samples were prepared and DNA was extracted as above from 3 to 80 mg samples. A vacuum concentrator was used, when necessary, to achieve a high enough concentration for GBS. Thirty µl of DNA for each sample was transferred to a 96-well plate; sample concentrations ranged from 15 to 76 ng/µl and resulting sample DNA mass ranged from 462 to 2278 ng. The minimum concentration recommended for GBS is 30 ng/µl, although samples with concentrations as low as 10 ng/µl can be processed. Eleven of our samples had concentrations lower than the recommended threshold, seven of which had concentrations between 29 and 30 ng/µl. Five samples with low concentrations were *E. stricta* individuals from HB (3) and AP (2), the remaining six were *E. hudsoniana* with one or two samples from each population. The lowest sample concentration submitted for our study was 15.41 ng/µl. Higher concentrations of DNA were difficult to obtain due to the small size and low biomass of *Euphrasia* plants.

Data generation and analysis

DNA samples were digested with the restriction enzyme *Pst*I and the resulting libraries were sequenced on the Illumina NextSeq 500 at the Cornell University Institute of Biotechnology Genomic Diversity Facility. Ninety-five samples and a blank control were sequenced. The non-reference pipeline UNEAK (Bradbury et al. 2007; Glaubitz et al. 2014; Lu et al. 2013), which was developed for SNP discovery and genotyping in non-model systems such as *Euphrasia* spp., was implemented in Tassel 3.0. The UNEAK pipeline is conservative, trimming reads to 64 base pairs, stacking identical reads into 'tags', and retaining only reciprocal tag pairs that differ for a single base pair in an effort to eliminate matches across paralogs. SNPs are identified using information from the whole dataset (not within a single individual), and are not required to meet a particular allelic dosage in order to be included in the SNP dataset. At the individual level, the SNPs called by the pipeline could represent a variety of allelic dosage levels (AABB, ABBB, or AAAB). Due to this fact, allelic dosage was treated as unknown in subsequent analyses. The UNEAK pipeline SNP calling procedure resulted in a total of 71,568 biallelic SNP loci. The resulting SNPs were subsequently filtered using VCFtools (v0.1.11, Danecek et al. 2011) and Tassel 5.0 (Bradbury et al. 2007) according to the following parameters: (1) loci with a mean coverage depth <3× were excluded, (2) loci with >20% missing data across individuals were excluded, (3) individuals with >10% missing data were excluded, and (4) loci with minor allele frequency <0.01 were excluded (McAllister and Miller 2016). Note that this stringent filtering exceeds the mean coverage depth of 1× used in the standard UNEAK analysis pipeline for population genetic analysis utilizing GBS markers for a species without a reference genome (Lu et al. 2013). SNPs were filtered separately for *E. hudsoniana* and *E. stricta* (resulting in 799 and 625 SNPs, respectively; online Table S1) then merged to ensure that locus quality was high within each species for separate analyses; hereafter referred to as the 'merged' dataset. This dataset retained the highest number of individuals in total, but filtering the species separately resulted in a small number of retained SNPs overlapping both species (196 SNPs). To ensure that the low number of overlapping SNPs did not affect results, we also completed the filtering for both species combined as a single group, resulting in 972 total SNPs; hereafter referred to as the 'filtered together' dataset. Most analyses were conducted using both datasets, and the dataset used for each analysis (either 'merged' or 'filtered together') is referenced in the figure or table legend. Analyses based on the 'filtered together' dataset are featured in the main text, with the analysis of the 'merged' dataset provided in the online Figs. S1-S3.

As a basic measure of population differentiation, Nei's D between populations and individuals was calculated in R (R Core Team 2015) based on the 'filtered together' dataset using package StAMPP (Statistical Analysis of Mixed Ploidy Populations; Pembleton et al. 2013). This package was designed specifically to calculate allele frequencies

and population genetic statistics for polyploid SNP data. StAMPP can calculate these statistics for unambiguous genotypes where allele dosage is known or, as in this case, for ambiguous genotypes treated as pseudodiploids that confer less precise, but still biologically meaningful results (Pembleton et al. 2013). AMOVAs were calculated based on Nei's distances using StAMPP to assess within and among population variation at multiple levels (Pembleton et al. 2013). Mantel tests were used to assess correlations between geographic and genetic distance (Nei's D) between native populations using the ade4 package in R (Dray and Dufour 2007). Neighbor-joining networks (Saitou and Nei 1987; Tamura et al. 2004) were constructed using 500 bootstrap replicates in MEGA7 (Kumar et al. 2016) for datasets including all individuals and for native individuals separately (note that bootstrap values for resulting trees are not presented in the main text, but can be found in supplementary figures). Multidimensional scaling (MDS) analyses were run for all individuals and native and invasive individuals separately with 5 dimensions using standard identity-by-state (IBS) distance matrices calculated in TASSEL 5.0 (Bradbury et al. 2007). Results for the first two principal coordinates were plotted in R.

Population structure and admixture was evaluated using STRUCTURE (Pritchard et al. 2000) for datasets including all individuals and for native individuals separately with a burn in length of 5000 and 50,000 MCMC repetitions. PLOIDY was set to 4 and RECESSIVE ALLELES was set to 1 to allow for a polyploid dataset of unknown allelic dosage. An admixture model was used assuming correlated allele frequencies. Twenty replicate runs were performed for each value of K = 1 through K = 6. Structure Harvester (Earl and vonHoldt 2012) was used to infer the best value of K from the replicate STRUCTURE runs using L(K), DeltaK, and Evanno methods (Evanno et al. 2005). The online interface of CLUMPAK (Kopelman et al. 2015) was used to identify and visualize individual assignment to clusters. Introgression between invasive and native samples was evaluated via a 'hybrid index' score calculated in INTROGRESS v1.22 using the 'filtered together' dataset (Gomert and Buerkle 2009, 2010). The *E. stricta* populations were used as one 'parental' group, the SL, BS, and GF *E. hudsoniana* populations as the other 'parental' group, and the HB and AP *E. hudsoniana* individuals were evaluated for introgression from the two 'parents'. The INTROGRESS package does not accommodate polyploid genotypes, so the data was converted to a dominant (presence/absence) format.

Genetic diversity statistics, consisting of observed and expected heterozygosity within populations, were calculated for the 'filtered together' dataset in GenoDive (Meirmans and Van Tienderen 2004). For expected heterozygosity (H_E) data was treated as tetraploid, coupled with the option to "Correct for unknown dosage of alleles". For observed heterozygosity (H_O), the correction option was not available, so results should be interpreted with caution. However, when measures of H_E were calculated without the correction, they were very similar to measures calculated with the correction applied (data not shown), implying that the results were not strongly influenced by the application of the correction. To ensure that coverage depth did not adversely affect heterozygote calls, we also analyzed 'filtered together' datasets with 5× and 10× minimum depth (with all other filtering parameters the same; filtering outcomes are shown in online Table S1).

Results

SNP calling and filtering

Illumina sequencing of 95 individuals resulted in 374,285,411 good barcoded reads, from which a total of 4,157,020 tags were identified, with a minimum of three reads per tag. From these sequence tags, 166,569 reciprocal tag pairs, or two-node networks, remained for SNP calling after removing errors. The total number of raw SNPs called from the UNEAK pipeline was 71,568. After the resulting SNP dataset was filtered separately by species, 80 individuals remained (60 native, 20 invasive) and 1,228 loci were retained (799 native with 49× average depth, 625 invasive with 85× average depth), with 196 of those loci overlapping both species—the 'merged' dataset (online Table S1). When SNPs of both species were filtered together, 70 individuals remained (62 native, 8 invasive) and 972 loci were retained with 72× average depth—the 'filtered together' dataset (online Table S1). Note that many invasive individuals were lost when SNPs were filtered together because SNPs might pass multiple filtering steps based on robust coverage and call rates in the native species (even if they were not present in most invasive individuals), causing invasives to be excluded when individuals with >10% missing data were filtered out. All SNPs retained for analysis had no more than 20% missing data across individuals (i.e., at least an 80% call rate). In both datasets, the final number of SNPs retained after filtering was much lower than the number of raw SNPs—a pattern that has been observed for other polyploid species (Clevenger and Ozias-Akins 2015). For *Euphrasia*, it is likely that the retention of SNPs might be improved with the development of a reference genome, but the challenge of calling SNPs in a non-model polyploid species from GBS data is well-demonstrated in this study. Overall, the results of all analyses were congruous regardless of whether SNPs were filtered separately or together. Thus, we present the results from the larger 'filtered together' dataset in the main text, with corresponding

results from the 'merged' dataset provided in the online supplementary figures (see online Figs. S1-S3). Any major differences in the results based on the two datasets are noted.

Analysis of native and invasive populations

Clear genetic divergence between native *E. hudsoniana* and invasive *E. stricta* was demonstrated by multiple analyses. MDS plots show that individuals identified as *E. hudsoniana* in the field cluster tightly together, while *E. stricta* individuals cluster into one tight group with additional individuals in a loose association (Fig. 2a; online Fig. S1a). The one exception to this pattern is individual AP_04, which was labeled as a native in the field but also considered to have an intermediate phenotype between the native and invasive species—it falls in an intermediate position between the native and invasive individuals.

Neighbor-joining networks corroborate the MDS analysis, with invasive individuals branching separately from native individuals, again with the exception of individual AP_04 (Fig. 3a; online Fig. S2a). Unlike all other native plants, AP_04 appears to share common ancestry with invasive plants from both AP and HB. However, when invasive plants were excluded from analysis, AP_04 appears to have common ancestry with other individuals from AP (Fig. 3b; online Fig. S2b). Finally, the neighbor-joining network including all individuals shows that invasive individuals from AP and HB group together, rather than forming reciprocally monophyletic groups based on population of origin.

STRUCTURE simulations based on ambiguous polyploid genotypes can lead to poor estimates of K (STRUCTURE manual), so the results of these simulations should be interpreted only in relation to additional evidence from other analyses. With this caveat noted, the STRUCTURE results do correspond to MDS and neighbor-joining results indicating that native and invasive populations are generally differentiated. When all native and invasive individuals are analyzed together, log-likelihood, rate of change of the likelihood distribution (Evanno method), and Delta K values (online Fig. S4a) support two (K = 2) distinct clusters (Fig. 4a; online Fig. S3a). The clusters identified by CLUMPAK separate native and invasive individuals into distinct populations, again with the exception of AP_04, which appears to have shared ancestry from both *Euphrasia* species. There also appears to be shared ancestry between invasive individuals from AP and HB and the native species, including AP_INV_04, AP_INV_08, and HB_INV_02 based on the 'filtered together' dataset, along with AP_INV_05, AP_INV_12, AP_INV_14, AP_INV_15, and HB_INV_09 based on the 'merged' dataset. These individuals also fell into intermediate positions on the MDS plots. In the INTROGRESS analysis, a hybrid

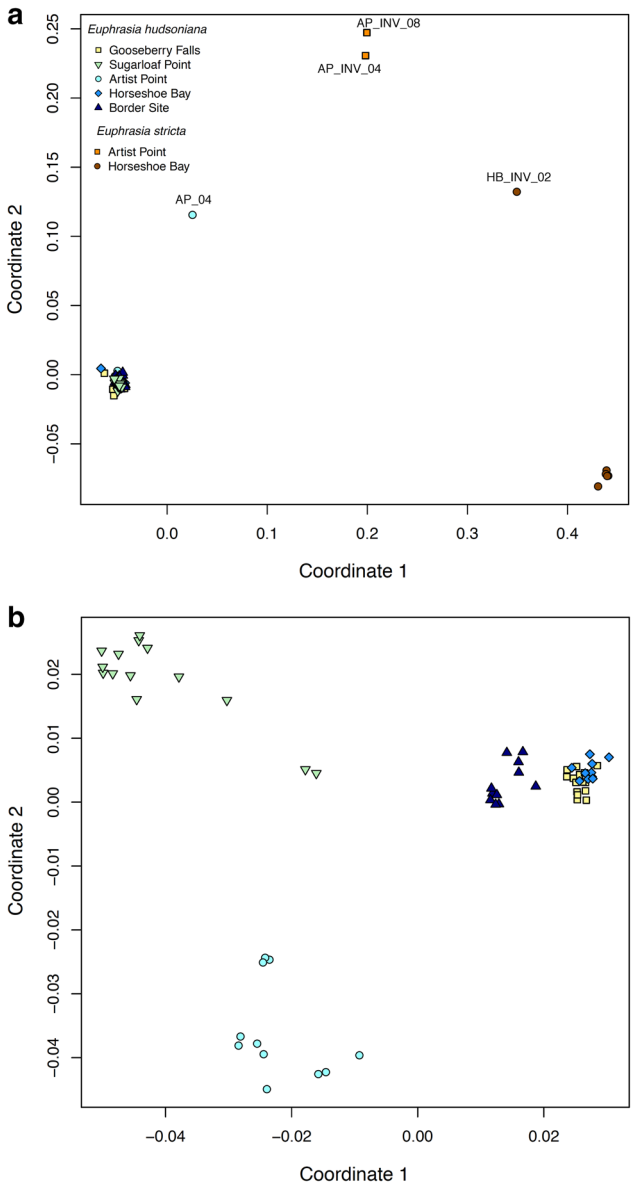

Fig. 2 MDS analysis of SNP data 'filtered together' (972 SNPs) for **a** *Euphrasia hudsoniana* and *E. stricta* individuals sampled from five and two populations respectively and **b** *E. hudsoniana* individuals only, excluding hybrid individual AP_04. The first two principal coordinates are plotted

index of 0 corresponds to a complete genomic assignment to *E. hudsoniana*, and a hybrid index of 1 corresponds to a complete genomic assignment to *E. stricta*. Out of the AP and HB individuals evaluated for introgression, AP_04 had the highest hybrid index of 0.21, while all other individuals from the AP and HB populations ranged from 0.003 to 0.015 (online Table S2).

The AMOVA analysis of Nei's distances showed that among-species differences (i.e., when *E. stricta* was considered one population and *E. hudsoniana* was considered

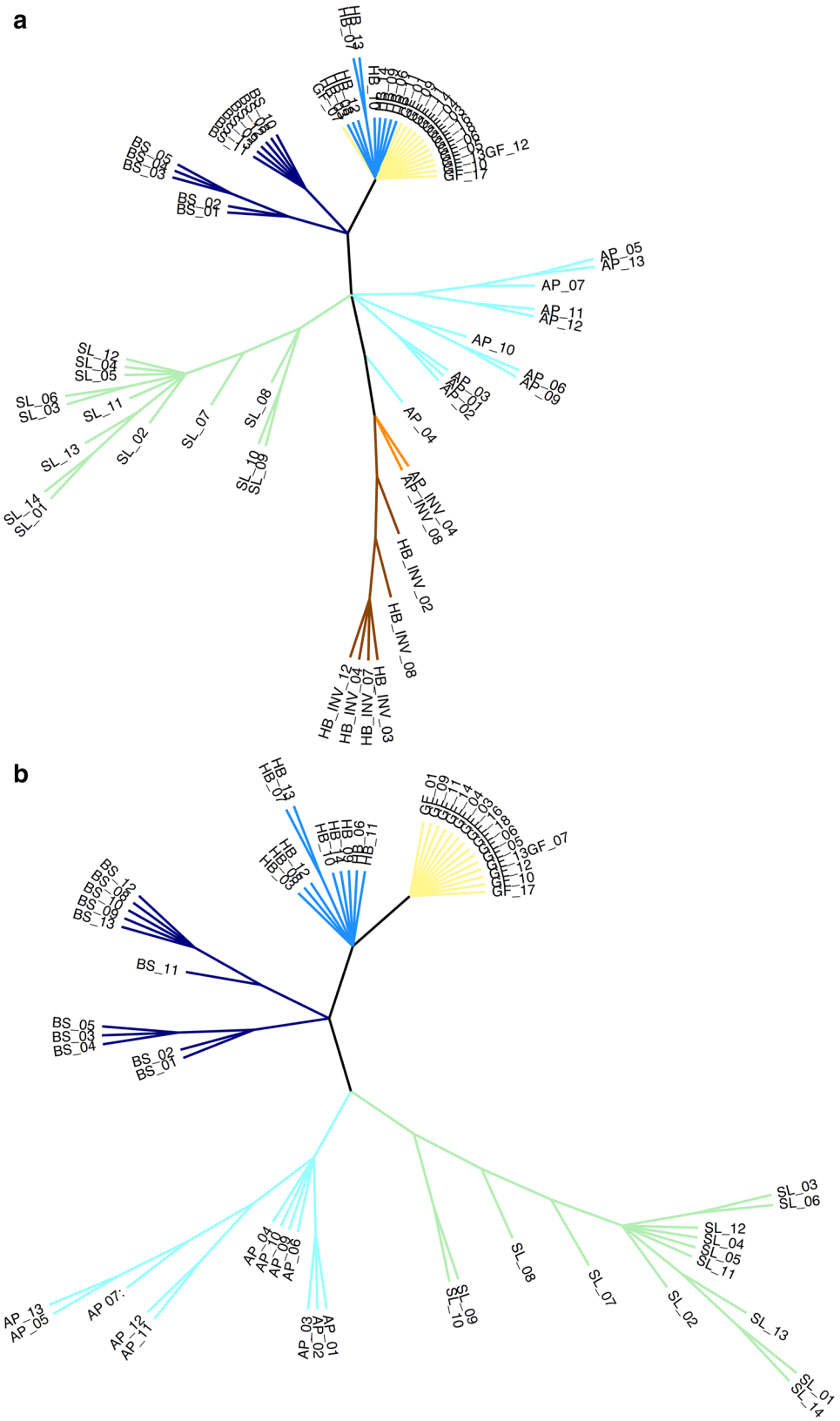

Fig. 3 Neighbor joining tree based on SNPs 'filtered together' (972 SNPs) for **a** *Euphrasia hudsoniana* and *E. stricta* individuals sampled from five and two populations respectively and **b** *E. hudsoniana* individuals only. See Fig S5 in supplementary figures for bootstrap values

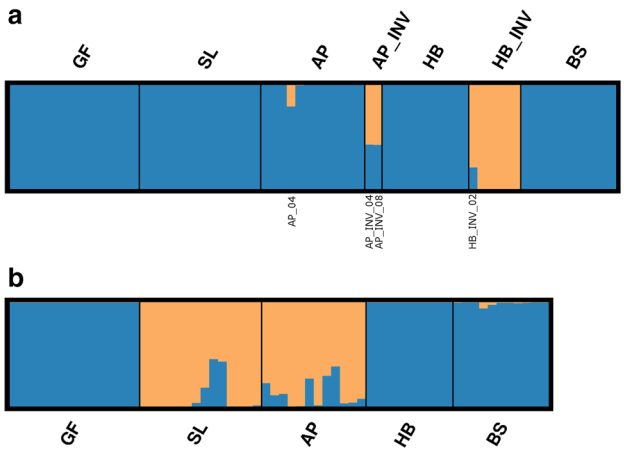

Fig. 4 STRUCTURE results based on SNP data 'filtered together' (972 SNPs) for best supported values of K for **a** *Euphrasia hudsoniana* and *E. stricta* individuals sampled from five and two populations respectively (K = 2) and **b** *E. hudsoniana* individuals (K = 2)

a second population) accounted for 96% of total variation. Among-population differences accounted for 95–96% of total variation when all populations were included, depending on whether individual AP_04 was included in analysis (Table 1). Expected and observed heterozygosity were calculated for a variety of coverage depths, and showed a gradual increase for all populations with higher minimum depth of coverage (Table 2), perhaps due to heterozygotes with low coverage being miscalled as homozygotes. This fact, along with the ability to specify unknown allelic dosage for H_E but not for H_O, requires that these estimates of diversity be interpreted with caution, and not considered comparable to values calculated for diploid populations with known allelic dosage. Instead, the values are useful for relative analyses, to compare H_O and H_E within each population, and compare the invasive to the native populations within each analysis. Overall, H_O was higher than H_E, and H_O and H_E were greatest for invasive populations (Table 2).

Analysis of native populations

Population structure among native individuals was subtle. In general, MDS results show that *E. hudsoniana* individuals cluster with other individuals collected from the same geographic location, but individuals originating from populations BS, HB, and GF group more closely together than to individuals from AP or SL (Fig. 2b; online Fig. S1b). Similarly, individuals group with other individuals collected from the same location on the neighbor-joining network based on the 'filtered together' dataset (Fig. 3b).

Table 1 AMOVA results for Nei's pairwise genetic distances of *Euphrasia hudsoniana* and *E. stricta* individuals from populations along the north shore of Lake Superior

	SSD	MSD	df	sigma2	% variation	p value
***E. hudsoniana* vs E. *stricta*—species comparison**						
Among population	0.04171	0.041710	1	0.002934	0.96	0
Among individuals	0.00895	0.000132	68	0.000132	0.04	
Total	0.05066	0.000734	69			
Population comparisons						
***Euphrasia hudsoniana* and *E. stricta* populations**						
Among population	0.04815	0.008025	6	0.000823	0.95	0
Among individuals	0.00251	0.000040	63	0.000040	0.05	
Total	0.05065	0.000734	69			
***Euphrasia hudsoniana* and *E. stricta* populations excluding AP_04**						
Among population	0.04822	0.008037	6	0.000837	0.96	0
Among individuals	0.00192	0.000031	62	0.000031	0.04	
Total	0.05015	0.000737	68			
***Euphrasia hudsoniana* populations**						
Among population	0.00107	0.000267	4	0.000020	0.49	0
Among individuals	0.00119	0.000021	57	0.000021	0.51	
Total	0.00226	0.000004	61			
***Euphrasia hudsoniana* populations excluding AP_04**						
Among population	0.00095	0.000238	4	0.000019	0.63	0
Among individuals	0.00061	0.000011	56	0.000011	0.37	
Total	0.00156	0.000061	60			

Based on the 'filtered together' dataset (972 SNPs). The 'Native vs. Non-native' comparison is between all *E. hudsoniana* individuals in one group, with all *E. stricta* individuals in the other group (no other population divisions). All other comparisons are conducted with populations intact within each species

Table 2 Genetic diversity statistics for *Euphrasia hudsoniana* and *E. stricta* populations including observed heterozygosity (H_O) and expected heterozygosity within populations (H_E) for 'filtered together' datasets filtered at 3×, 5×, and 10× mean site depths

Population	Genetic diversity					
	Mean site depth 3× 972 SNPs		Mean site depth 5× 896 SNPs		Mean site depth 10× 651 SNPs	
	H_O	H_E	H_O	H_E	H_O	H_E
BS	0.178	0.109	0.186	0.113	0.219	0.128
HB	0.175	0.103	0.183	0.107	0.218	0.122
HB_INV	0.191	0.163	0.207	0.243	–	–
AP	0.187	0.127	0.196	0.132	0.234	0.154
AP_INV	0.232	0.182	0.247	0.233	0.311	0.296
SL	0.186	0.112	0.194	0.115	0.231	0.13
GF	0.176	0.105	0.187	0.11	0.220	0.121
Average	0.189	0.129	0.200	0.150	0.239	0.159

No statistics were calculated for the HB_INV population at the 10 mean site depth because only one individual remained after filtering

Interestingly, the individuals from BS split into two different groups on the neighbor-joining tree according to two different geographic locations plants were collected from at that site, suggesting sub-population structure. The neighbor-joining network constructed using the 'merged' dataset was generally similar but without good population-level grouping, perhaps due to the small number of SNPs in the dataset (online Fig. S2b). In the STRUCTURE analysis of native populations, two (K = 2) clusters of native individuals were best supported by simulations (online Fig. S4b), with AP and SL forming one group, and BS, HB, and GF forming the other group (Fig. 4b; online Fig. S3b). Individuals from SL and AP show shared ancestry with individuals from the BS/HB/GF group. The AMOVA analysis of Nei's distances showed that among-population differences within *E. hudsoniana* accounted for 37–49% of total variation, depending on whether individual AP_04 was included in analysis (Table 1). There was no significant correlation between geographic and genetic distance ($r = -0.26$, $P = 0.78$). Values of H_O and H_E were similar across native populations, with a slight trend towards higher values of H_O for the SL and AP populations (Table 2).

Discussion

How do relict populations evolve in isolation from the populations in the home range of the species? How do they respond to evolutionary pressures of a changing climate and contact with invasive congeners? Answering these questions requires both genetic and phenotypic examination of relict populations across time and space. Here, we use the former approach to understand the current genetic status of relict populations of *E. hudsoniana* along the northern shore of Lake Superior in Minnesota. In particular, this study is designed to evaluate the extent of diversity and differentiation among populations of the native species, and to determine whether hybridization that has been suggested based morphological examination of herbarium specimens is detectable at the genetic level.

Population relationships and genetic diversity

Relict populations of *E. hudsoniana* are isolated not only from the home range of the species, but also from each other. Populations of these species occur sporadically on rocky habitats on the shore of Lake Superior, and do not form contiguous populations along the coast (Smith 1988). Many populations are endangered by human traffic because the rocky outcrops are frequented by tourists—it is therefore likely that some populations have already been destroyed by human activity, leaving the remaining populations even more isolated. The species is annual and does not reproduce clonally, and gene flow between populations is most likely to occur via insect-mediated pollen movement or water-based seed transport along the shore. In some places, seed transport is also likely mediated by humans on muddy shoes and clothing due to frequent human visits to scenic areas of Lake Superior. The level of connectivity of these populations is unknown, as is their relationship to each other along the coastline.

Small, isolated populations, such as those of *E. hudsoniana*, often experience a relative deficit of heterozygosity (Bauert et al. 1998); however, this wasn't observed in our study (Table 2). One possible explanation for excess heterozygosity in these populations is fixed heterozygosity in the tetraploid genome of this species (Meirmans and Van Tienderen 2013). Many arctic plants are polyploid, and the duplicated genome is hypothesized to be a safeguard against loss of genetic diversity in harsh habitats and isolated populations (Brochmann and Steen 1999; Abbott and Brochmann 2003). Thus, the fixed

heterozygosity could benefit *E. hudsoniana* in Minnesota, allowing the species to persist in small isolated populations while maintaining diversity. Despite this fixed heterozygosity, it is clear that *E. hudsoniana* populations are not as diverse as the introduced *E. stricta*; measures of observed and expected heterozygosity are generally higher in the invasive species than in the native species, and the genetic distance among individuals is higher for *E. stricta* than for *E. hudsoniana*. Although no one has conducted a population genetic survey of introduced *E. stricta*, most introduced populations are subject to genetic bottlenecks, so it is interesting that the diversity of this invasive exceeds that of the native species.

This study reveals clear differentiation between the native and invasive species, and also subtle structure among the native populations. The two invasive populations showed differentiation from one another, but there was still interdigitation of populations across analyses. That is, individuals from AP_INV were frequently most closely related to individuals from HB_INV, and vice-versa; this was particularly evident in the neighbor-joining networks and MDS plots. This pattern may indicate multiple introductions of *E. stricta* to the shore of Lake Superior, or a larger population of *E. stricta* within Minnesota that consists of multiple genetically distinct groups, and warrants further study.

The native species populations fell into two main groups, but the groups did not correspond to the relative geographic locations along the coast. The most southern population, GF, grouped with the two most northern populations, HB and BS, while SL and AP formed a separate group. This discordance between the locations of the populations and their geographical location is also evidenced by the non-significant Mantel test. The close relationship between SL and AP might be due to seed dispersal, potentially by humans, as both SL and AP are easily accessible and frequently visited by tourists. Note that while the population differentiation between the SL/AP and the HB/BS/GF groups appeared strong based on some analyses, levels of differentiation among *E. hudsoniana* populations were lower compared to the differentiation between native individuals based on the MDS (Fig. 2a). Overall, our results clearly show that populations could be differentiated from each other (and, at BS, population substructure could also be identified; Fig. 3b), suggesting that continued conservation efforts to preserve all relict population groups are necessary to maintain existing differentiation and diversity. Nonetheless, the *E. hudsoniana* populations were still similar enough to group together within the species when compared to another species of *Euphrasia*. Thus, despite the small population sizes and potential for genetic drift in isolation, *E. hudsoniana* populations still appear coherent at the species level.

Hybridization and conservation implications

Hybridization between the native *E. hudsoniana* and the invasive congener *E. stricta* has been suggested based on morphological examination of herbarium specimens. While hybridization can play a creative role in evolution (Rieseberg 1997; Gross and Rieseberg 2005), it can also pose a danger to a rare species if they are 'genetically swamped' by a more common species (Rhymer and Simberloff 1996; Todesco et al. 2016). This is an issue of serious concern for *E. hudsoniana* because contact with *E. stricta* is novel (on an evolutionary time-scale), and may increase through time.

Our sampling suggests that interspecific hybridization is not occurring at a high level for the surveyed native populations; only one hybrid was detected out of all analyzed native samples (AP_04) and all other individuals appeared to be pure *E. hudsoniana*. Conversely, most invasive individuals from AP and a few from HB resemble native plants in their genetic make-up, due to either shared ancestry or hybridization. This suggests that very little gene flow has occurred from *E. stricta* to *E. hudsoniana* where these species are coming into contact, but that gene flow is happening more frequently from *E. hudsoniana* to *E. stricta*. While seemingly counterintuitive, these results correspond to the predictions from simulations and other studies that when invading species come into contact with native populations, gene flow typically occurs from the native to the non-native (Currat et al. 2008; Owens et al. 2016). We note that the gene flow from *E. hudsoniana* into *E. stricta* might be responsible for the relatively high diversity in the invasive species; further study of the *E. stricta* populations in the invasive range are required to evaluate this possibility.

These results indicate that the process of hybridization has not had a major impact on native populations at this point. The date of the invasion of Minnesota by *E. stricta* is unknown, nor is it known how long the populations of *E. stricta* have been present along the shore of Lake Superior. If contact between *E. stricta* and *E. hudsoniana* is recent, it is possible that this study is documenting the 'leading edge' of the process of hybridization between the native and invasive *Euphrasia* species. It is important to note that the admixed native individual (AP_04) was initially identified as having an intermediate morphology during sample collection, and was a part of a group of individuals with intermediate phenotypes. These individuals occur in traditional *E. hudsoniana* habitat, but are larger and more robust than average to the point where they begin to resemble *E. stricta*. Our observation of *E. hudsoniana* populations since the collection of samples for this study do show an increase in the number of relatively large plants that appear intermediate or resemble *E. stricta*. While we cannot exclude the

possibility that the phenotypes in these populations are the result of phenotypic plasticity, the current genetic analysis suggests that the increase in the number of large or intermediate *E. hudsoniana* individuals may be due to an increased frequency of hybridization. Alternatively, gene flow from *E. stricta* to *E. hudsoniana* may be prevented by a reproductive isolating barrier. For example, although many *Euphrasia* spp. readily hybridize (Yeo 1968; Liebst and Schneller 2005; Liebst 2008), *E. hudsoniana* and *E. stricta* potentially have contrasting breeding systems in which the small flowered *E. hudsoniana* may primarily self-fertilize (French et al. 2005; Liebst and Schneller 2005), while *E. stricta* primarily outcrosses via insect pollination (Hegland and Totland 2012). If this or another reproductive barrier between these species exists, the number of hybrids may not change in the future. In either case, a genetic study conducted in 5–10 years would be useful to shed light on the dynamics of genetic exchange (or lack thereof) between the species.

In terms of conservation implications, this study makes it clear that removal of *E. stricta* populations from the shore of Lake Superior in the near future would still lead to the preservation of *E. hudsoniana* relict populations as a genetically distinct group, free from genetic material from the invasive species. Moreover, our work provides preliminary evidence that an intermediate phenotype is a good guide to the underlying genotype of a *Euphrasia* individual in terms of assessing introgression status, and the genetic invasion is not cryptic at the phenotypic level as it is in some taxa (Johnson et al. 2016; Lucek 2016; Holsbeek et al. 2008; Saltonstall 2002). While this requires further confirmation, the match between phenotype and genotype has the potential to make removal of hybrid individuals straightforward, provided that introgression does not reach an advanced state where most individuals are advanced generation hybrids.

Conclusion

This study focused on five native populations of *E. hudsoniana* at the disjunct southern margin of their range in North America and represents the first population genetic analysis of this species. Our results provide a baseline of genetic information to inform future studies and conservation of this species. For example, closer examination of the genetics of populations at AP and SL compared to GF, HB, and BS may be warranted to determine if these populations comprise a separate, cryptic species. Additional sampling across time and space will continue to provide valuable insights into the dynamics of *E. hudsoniana* populations across the species' range and relative potential effects of climate change within the species' core and disjunct ranges. Continued genetic monitoring of native and invasive populations, particularly at AP and HB, may allow us to view how the genetic effects of hybridization play out in this system.

Overall, we show here that *E. hudsoniana* and *E. stricta* populations in close proximity along the shore are distinct and that gene flow into the native species is currently limited. Future monitoring efforts would be beneficial to determine whether *E. stricta* is increasingly coming into contact with *E. hudsoniana* and whether hybridization is becoming more common. If efforts to remove coastal populations of *E. stricta* are made in the near future, our study suggests that the *E. hudsoniana* populations will not retain a strong genetic signature of hybridization. The relative genetic and phenotypic similarity (Winkler and Etterson 2015) among native *E. hudsoniana* populations may suggest that they are all at similar risk of negative effects due to climate change. *E. hudsoniana* is a component of the unique assemblage of rare arctic species in Minnesota that are relicts of the fascinating geologic history of this region. Conservation of this species in the face of increasing habitat loss and climate change will depend on an understanding of the dynamics underlying their populations, which necessitates continued monitoring and research.

Acknowledgements We acknowledge L. Gerdes and G. Gusarova for sharing their expertise on *Euphrasia* species, M Cleveland, A McCormick, the Grand Portage Band of Lake Superior Chippewa, the Minnesota Department of Natural Resources, and United States Forest Service for facilitating or providing permission to collect samples, J. Manolis for assistance collecting samples, A. Boser, J. Horky, A. Schumann, and M. Wedger for lab support, the Gross, Strasburg, and Etterson labs, and two anonymous reviewers for helpful comments on this manuscript. This work was funded in part by the Coastal Zone Management Act of 1972, as amended, administered by the Office for Coastal Management, National Oceanic and Atmospheric Administration under Award NA14NOS4190055 provided to the Minnesota Department of Natural Resources for Minnesota's Lake Superior Coastal Program.

References

Abbott RJ, Brochmann C (2003) History and evolution of the arctic flora: in the footsteps of Eric Hultén. Mol Ecol 12:299–313

Bauert MR, KÄLin M, Baltisberger M, Edwards PJ (1998) No genetic variation detected within isolated relict populations of *Saxifraga cernua* in the Alps using RAPD markers. Mol Ecol 7:1519–1527

Billings WD, Mooney HA (1968) The ecology of arctic and alpine plants. Biol Rev 43:481–529

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Brochmann C, Steen SW (1999) Sex and genes in the flora of Svalbard—implications for conservation biology and climate change. Det Norske Videnskaps-Akademi I Matematisk-Naturvitenskapelig Klasse Skrifter, Ny Serie 38:33–72

Clevenger JP, Ozias-Akins P (2015) SWEEP: a tool for filtering high-quality SNPs in polyploid crops. G3 5:1797–1803. doi:10.1534/g3.115.019703

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. Evol Int J Org Evol 62:1908–1920

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158. doi:10.1093/bioinformatics/btr330

Davis MB, Shaw RG, Etterson JR (2005) Evolutionary responses to changing climate. Ecology 86:1704–1714

Doak DF, Morris WF (2010) Demographic compensation and tipping points in climate-induced range shifts. Nature 467:959–962

Downie SR, McNeill J (1988) Description and distribution of *Euphrasia stricta* in North America. Rhodora 90:223–231

Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22:1–20

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. doi:10.1007/s12686-011-9548-7

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. Annu Rev Ecol Evol Syst 24:217–242

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE. doi:10.1371/journal.pone.0019379

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. doi:10.1111/j.1365-294X.2005.02553.x

French GC, Ennos RA, Silverside AJ, Hollingsworth PM (2005) The relationship between flower size, inbreeding coefficient and inferred selfing rate in British *Euphrasia* species. Heredity 94:44–51

Given DR, Soper JH (1981) The artic-alpine element of the vascular flora at lake superior. National Museum of Natural Resources Publications in Botany, No. 10

Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9:e90346

Gomert Z, Buerkle CA (2009) A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. Mol Ecol 18:1207–1224

Gomert Z, Buerkle CA (2010) *introgress*: a software package for mapping components of isolation in hybrids. Mol Ecol Resour 10:378–384

Gross BL, Rieseberg LH (2005) The ecological genetics of homoploid hybrid speciation. J Hered 96:241–252

Gusarova G, Popp M, Vitek E, Brochmann C (2008) Molecular phylogeny and biogeography of the bipolar *Euphrasia* (Orobanchaceae): recent radiations in an old genus. Mol Phylogenet Evol 48:444–460

Hampe A, Jump AS (2011) Climate relicts: past, present, future. Annu Rev Ecol Evol Syst 42:313–333

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8:461–467

Hegland SJ, Totland Ø (2012) Interactions for pollinator visitation and their consequences for reproduction in a plant community. Acta Oecol 43:95–103

Holsbeek G, Mergeay J, Hotz H, Plotner J, Volckaert FA, De Meester L (2008) A cryptic invasion within an invasion and widespread introgression in the European water frog complex: consequences of uncontrolled commercial trade and weak international legislation. Mol Ecol 17:5023–5035

Huff A, Thomas A (2014) Lake Superior climate change impacts and adaptation. Prepared for the Lake Superior Lakewide Action and Management Plan – Superior Work Group

Johnson MG, Lang K, Manos P, Golet GH, Schierenbeck KA (2016) Evidence for genetic erosion of a California native tree, *Platanus racemosa*, via recent, ongoing introgressive hybridization with an introduced ornamental species. Conserv Genet 17:593–602

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol Ecol Resour 15:1179–1191

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol. doi:10.1093/molbev/msw054

Lesica P, McCune B (2004) Decline of arctic-alpine plants at the southern margin of their range following a decade of climatic warming. J Veg Sci 15:679–690

Liebst B (2008) Do they really hybridize? A field study in artificially established mixed populations of *Euphrasia minima* and *E. salisburgensis* (Orobanchaceae) in the Swiss Alps. Plant Syst Evol 273:179–189

Liebst B, Schneller J (2005) How selfing and intra-and interspecific crossing influence seed set, morphology and ploidy level in *Euphrasia*: an experimental study of species occurring in the Alps of Switzerland. Plant Syst Evol 255:193–214

Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, Buckler ES, Costich DE (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. PLoS Genet 9:e1003215

Lucek K (2016) Cryptic invasion drives phenotypic changes in central European threespine stickleback. Conserv Genet 17:993–999

McAllister CA, Miller AJ (2016) Single nucleotide polymorphism discovery via genotyping by sequencing to assess population genetic structure and recurrent polyploidization in *Andropogon gerardii*. Am J Bot 103:1314–1325

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

Meirmans PG, Van Tienderen PH (2013) The effects of inheritance in tetraploids on genetic diversity and population divergence. Heredity 110:131–137

MN DNR (2016) *Euphrasia hudsoniana* var. *ramosior*. Minnesota's endangered, threatened, and special concern species: rare species guide. http://www.dnr.state.mn.us/rsg/profile.html?action=elementDetailandselectedElement=PDSCR0P053

Nimis P, Malyshev L, Bolognini G, Friesen N (1998) A multivariate phytogeographic analysis of plant diversity in the Putorana Plateau (N. Siberia). Opera Bot 136:1–72

Owens GL, Baute GJ, Rieseberg LH (2016) Revisiting a classic case of introgression: hybridization and gene flow in Californian sunflowers. Mol Ecol 25:2630–2643

Pembleton LW, Cogan NO, Forster JW (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol Ecol Resour 13:946–952

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

R Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/

Reznicek AA, Voss EG, Walters BS (2011) Michigan Flora Online. University of Michigan. Web. August 11, 2016. http://michiganflora.net/species.aspx?id=1865

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. Annu Rev Ecol Evol Syst 27:83–109

Rieseberg LH (1997) Hybrid origins of plant species. Annu Rev Ecol Evol Syst 28:359–389

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Saltonstall K (2002) Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. Proc Natl Acad Sci USA 99:2445–2449

Sell PD, Yeo PF (1970) A revision of the North American species of *Euphrasia* L. (Scrophulariaceae). Bot J Linn Soc 63:189–234

Smith W (1988) Vascular plants, Chapter 1. In: Coffin B, Pfannmuller L (eds) Minnesota's endangered flora and fauna. University of Minnesota Press, Minneapolis, p 185

Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci USA 101:11030–11035

Thuiller W, Richardson DM, Midgley GF (2007) Will climate change promote alien plant invasions? In: Nentwig W (ed) Biological invasions. Springer, New York, pp 197–211

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

Winkler KJ, Etterson JR (2015) Natural selection in populations of *Euphrasia hudsoniana* var. *ramosior* along the north shore of Lake Superior. Technical report for Minnesota's Lake Superior Coastal Program

Yeo PF (1968) The evolutionary significance of the speciation of *Euphrasia* in Europe. Evolution 22:736–747. doi:10.2307/2406899