Population genetics of *Trichophorum planifolium* structured by inbreeding, not distribution limits

Victoria J. Nowell, Song Wang, and Tyler W. Smith

2021-02-11

Note

This is the source file for our manuscript. It contains all of the writing and analysis contained in our publication. However, final formatting, including revisions from co-authors and peer reviewers, were done in a separate Word MS word document. It's just too big a hassle to copy those changes back to this document. Consequently, there will be some differences between this file and the published document. In case of disagreement, the published manuscript is the canonical report we stand behind.

Abstract

Peripheral populations are understood to hold special conservation significance as repositories of divergent genetic diversity. They may also be at increased risk of extirpation, due to a combination of small size, isolation, and marginal habitat conditions. We collected genetic, ecological, and distributional data to evaluate these hypotheses for an endangered range-edge population of the woodland sedge *Trichophorum planifolium*. Our data show that, range-wide, *T. planifolium* has remarkably low within-population diversity (F_{ST} 0.8, 1% heterozygosity). Most populations are dominated by a small number of distinct multi-locus genotypes. Population differentiation for this species is much higher than other outcrossing, wind-pollinated species. However, the extremely low observed heterozygosity we document is comparable to results from other woodland sedges. This suggests that the life history of this group of plants results in high levels of inbreeding. In this context, the somewhat reduced diversity and limited divergence of the lone remaining Canadian population is not remarkable, and does not support suggestions that the population suffers from unusual levels of inbreeding. Despite being located on the geographic margin of the species' range, the climatic conditions at the remaining Canadian population are not extreme compared to the rest of the range in the northeastern US. However, it is geographically isolated, and the soil conditions at the location are unusual for the species. Cluster analysis shows that most populations across the range of the species share relatively homogeneous soil conditions. In constrast, the Canadian population forms a small outlying cluster with extreme nutrient values. Thus, while this outlying population does not appear to be divergent genetically, it's geographic isolation and divergent environmental conditions indicate it may hold important conservation value.

Keywords

abundant center hypothesis; conservation genetics; microsatellites; *Trichophorum planifolium*; peripheral populations; range limits; species distribution models

Introduction

No limit, aim for 8000 words.

In high-latitude countries, a high proportion of species identified as requiring legal protection (i.e., Endangered species) are actually geographically peripheral populations of species common at lower latitudes (Yakimowski and Eckert, 2007; Gibson et al., 2009). For example, in Canada, approximately 90% of listed Threatened or Endangered plant species are locally rare representatives of species that reach the northern extent of their range in Canada, but are abundant further south (Yakimowski and Eckert, 2007). Concern over the investment of conservation resources into species that are locally rare but globally common has led to debate over whether, or when, peripheral populations merit national concern (e.g., Hunter and Hutchinson, 1994; Hampe and Petit, 2005; Gibson et al., 2009).

Peripheral populations are typically small and geographically isolated due to increased habitat variability and reduced resource availability at the range margin (i.e., the abundant center hypothesis, Brown, 1984). Consequently, they may be at greater risk of extirpation due to genetic factors (i.e., inbreeding depression) as well as stochastic demographic, environmental and catastrophic factors (Soulé and Mills, 1998; Brook et al., 2002). These risks may be exacerbated by anthropogenic threats such as habitat loss and fragmentation (Young et al., 1996). Nevertheless, peripheral populations may be important sources of genetic diversity for the species (Safriel et al., 1994; Leppig and White, 2006; Eckert et al., 2008), harbouring unique genetic diversity and/or adaptations to conditions at the range edge (Lesica and Allendorf, 1995; Bunnell et al., 2004).

However, empirical studies attempting to quantify the genetic differentiation of peripheral plant populations have produced mixed results (Eckert et al., 2008). The lack of consistent range-edge genetic differentiation may be attributable in part to species-specific attributes (Gibson et al., 2009). In plants, life-history characteristics (i.e., breeding system and dispersal mechanisms) are a major determinant of population genetic structure, because they strongly impact the rates of gene flow (Loveless and Hamrick, 1984; Frankel et al., 1995). Knowledge of the spatial distribution of intraspecific genetic diversity can guide conservation efforts for peripheral populations, especially when the reproductive biology and ecology of the species is not well understood.

Lesica and Allendorf (1995) proposed a framework for evaluating the conservation value of peripheral populations, identifying population isolation and divergent environmental conditions as main criteria. In practice, neutral genetic markers, such as microsatellites, can be used to identify populations that have experienced restricted gene flow (i.e., are genetically isolated) (Crandall et al., 2000; Moritz, 2002). However, microsatellite markers cannot provide information on adaptive divergence of populations (Holderegger et al., 2006; Teixeira and Huber, 2021). Therefore, to capture variation in selection pressures across the species' range, environmental differences among populations should be assessed (Allendorf et al., 2013).

Trichophorum planifolium (Sprengel) Palla, commonly known as Few-flowered Club-rush or Bashful Bulrush, is a species that is rare at its northern limit in Canada, but is common further south. It is a perennial woodland species in the family Cyperaceae, typically occurring on mesic to dry, often rocky, slopes in association with hardwoods, especially Oaks (*Quercus* spp.) (Crins, 2002). The range of *T. planifolium* extends from Massachusetts west to Ontario and south to Virginia and Kentucky, with disjunct populations occurring in Missouri and southern Illinois (Figure 1).

The only extant locality in Canada is in southern Ontario, in the nature sanctuaries at Royal Botanical Gardens in Hamilton, Ontario. *Trichophorum planifolium* was previously known from one other location in Ontario, in Rouge Valley near Toronto, but no plants have been located at this site since 2005 (Smith and Rothfels, 2007). Due to its limited Canadian range and an apparent decline in population size, *T. planifolium* is listed as Endangered under the Canadian Species at Risk Act (SARA, COSEWIC, 2000).

Trichophorum planifolium is caespitose, growing in dense, low tufts (~ 40 cm high). Individual plants reach a maximum diameter of 20-30 cm, and lack the long rhizomes present in clonally spreading sedges such as *Carex pensylvanica* Lamarck (Crins and Rettig, 2002). It is wind-pollinated, with styles and stamens exerted at anthesis, and has neither a showy perianth or nectaries (Crins, 1989; Iwanycki et al., 2010). Seeds mature in mid-summer and disperse in late July to August (Iwanycki et al., 2010). A mechanism for seed dispersal has not been confirmed (Smith and Rothfels, 2007), but it has been suggested that the matting of leaves and stems during the fruiting period may limit most seed dispersal to within a meter of the parent plants, leading to the formation of colonies of closely related sibling groups (Crins, 1989). It has been suggested that low

recruitment observed at the Hamilton site may be due to limited intrapopulational genetic diversity (Smith and Rothfels, 2007).

The main objective of our research was to apply Lesica and Allendorf's (1995) framework to assess the conservation value of the last remaining peripheral population of *Trichophorum planifolium* in Canada. Specifically, we assess genetic diversity (richness and divergence) using neutral microsatellite markers, to evalute the hypotheses that this peripheral population has reduced richness and increased divergence relative to central populations. Additionally, we assess variation in abiotic conditions (climate niche and local soil properties) and spatial isolation, to test the hypotheses that peripheral populations are subject to divergent environmental conditions and reduced connectivity to core populations.

MATERIALS AND METHODS

Population Sampling

In spring 2014, we sampled 29 populations of *T. planifolium* across its geographic distribution (Figure 1, Table 1). An additional population sample was previously obtained in 2011 in Kentucky. At each site, we sampled individuals haphazardly, with at least 1 m between individuals, and a maximum distance of 100 m between the most distant samples. In total, 919 tissue samples were obtained, with a minimum of four and a maximum of 50 samples per population (mean = 31.7). Our protocol was modified for one very small population (PACR), where one sample was taken at a distance of 150 m to increase sample size to seven. A single herbarium voucher was collected per population, except for the locations in Ontario, New Hampshire, and West Virginia, where the species is Endangered and collection of whole plants is prohibited. Vouchers have been deposited at DAO (abbreviations follow Thiers, 2020).

Tissue samples were immediately placed on silica gel and stored at room temperature prior to molecular analysis. A random subsample of 20 samples per population were selected for genetic analysis. For populations with fewer than 20 samples, all individuals were used.

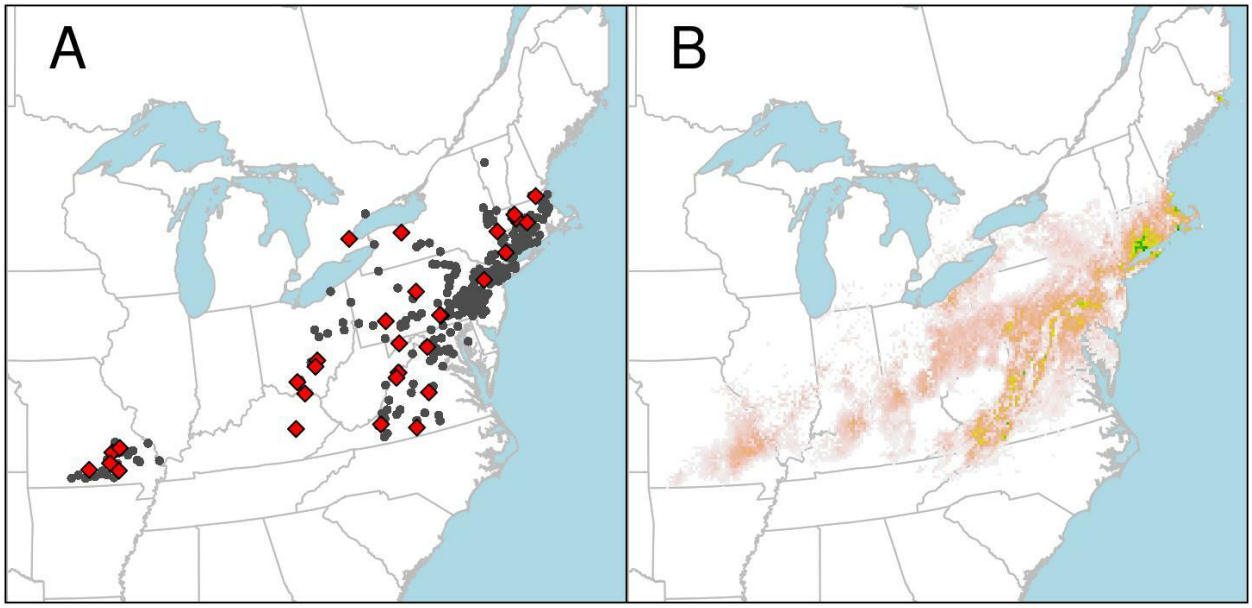

Figure 1: *Trichophorum planifolium* distribution. A. Populations Sampled: Grey points are GBIF records; red diamonds are populations sampled for the genetic survey and soil analysis. B. Maxent species distribution model. Highest suitability is indicated in green, declining through yellow, orange and grey.

Table 1: *Trichophorum planifolium* populations sampled. Population: the population code, the first two letters indicate the US state or Canadian province. Lat and Lon: geographic coordinates. N: total samples per population (samples with all 11 loci scored in parentheses). Ap: private alleles. Ar: mean allelic richness per locus. MLG: multilocus genotypes. eMLG: multilocus genotypes in rarefied samples of 10 individuals. Hs: gene diversity. Ho: observed heterozygosity. F_{IS}: population inbreeding coefficient. F_D : mean pairwise F_{ST} for the population when compared to all other populations. The value in the Total row for the F_D column is the global F_{ST} value.

population	Latitude	Longitude	N	Ap	Ar	mlg	emlg	Hs	Ho	Fis	Fst
CTTP	41.3714	-73.4527	17(16)	0	1.00	1	1.00	0.00	0.00		0.89
KYLE	37.5906	-83.7815	5(2)	1	1.07	3		0.05	0.02	0.60	0.84
MABM	42.3017	-72.5295	19(14)	1	1.12	5	3.63	0.07	0.00	0.93	0.81
MAMT	42.4471	-72.5389	19(19)	0	1.06	3	2.05	0.03	0.03	0.13	0.85
MAST	42.0585	-72.1300	20(20)	0	1.00	1	1.00	0.00	0.00		0.89
MDDM	39.5678	-78.9119	20(19)	0	1.32	7	4.53	0.18	0.02	0.87	0.76
MOBS	36.9475	-90.9931	41(17)	2	1.31	17	6.53	0.16	0.01	0.93	0.70
MOBU	37.5356	-91.2116	14(11)	1	1.45	6	5.34	0.25	0.01	0.97	0.60
MOHC	37.0226	-92.1677	14(12)	0	1.11	4	3.43	0.06	0.01	0.90	0.83
MOOC	37.6643	-90.8998	17(12)	0	1.03	3	2.18	0.02	0.01	0.65	0.86
MOSH	37.1996	-91.3258	20(15)	0	1.14	5	4.21	0.07	0.01	0.84	0.78
NHER	42.7643	-71.3613	19(14)	1	1.27	4	3.51	0.15	0.03	0.79	0.68
NJKP	40.7612	-74.7158	17(16)	0	1.14	3	2.18	0.07	0.01	0.93	0.82
NYMP	43.0301	-77.5737	19(19)	0	1.02	2	1.53	0.01	0.00	1.00	0.90
NYTS	42.1176	-73.4972	20(15)	0	1.29	3	2.89	0.17	0.00	1.00	0.67
OHBH	39.6205	-82.4048	13(13)	0	1.00	1	1.00	0.00	0.00		0.80
OHHH	39.4381	-82.5306	20(20)	0	1.20	7	5.32	0.11	0.00	0.96	0.71
OHPR	38.6525	-83.1637	18(17)	1	1.04	3	2.11	0.02	0.00	1.00	0.77
OHSC	39.0583	-83.3886	17(15)	0	1.00	1	1.00	0.00	0.00		0.87
ONRB	43.2656	-79.9210	27(22)	1	1.01	2	1.37	0.01	0.00	1.00	0.87
PABE	41.0438	-77.6540	15(7)	0	1.11	4	3.31	0.06	0.00	1.00	0.82
PACR	40.3703	-79.2275	6(5)	0	1.06	2		0.04	0.00	1.00	0.82
PAGP	40.0746	-76.9091	18(13)	0	1.34	7	4.71	0.19	0.05	0.76	0.68
PASG	40.0976	-76.9484	9(8)	0	1.04	2		0.02	0.00	1.00	0.76
VAAC	37.8035	-78.2493	17(14)	0	1.09	3	2.59	0.05	0.00	1.00	0.81
VAHB	37.1885	-80.4192	18(14)	0	1.00	1	1.00	0.00	0.00		0.90
VASC	36.8198	-79.0665	3(1)	0	1.09	2		0.09	0.00	1.00	0.75
WVFM	38.6642	-79.2404	19(18)	2	1.30	9	6.08	0.16	0.01	0.97	0.67
WVMO	38.5188	-79.3705	20(18)	1	1.30	7	4.92	0.16	0.00	1.00	0.72
WVSS	39.2222	-77.8200	4(2)	0	1.07	2		0.05	0.00	1.00	0.85
Total			505(408)			115	9.03	0.08	0.01	0.91	0.80

Genetic Analysis

For each individual, 20 mg of dried leaf tissue was ground with 3 mm-diameter stainless steel beads in a 2 mL Eppendorf tube at 30 Hz for 2 min using a TissueLyser II (Qiagen, Venlo, Limburg, Netherlands). Total genomic DNA was extracted from ground tissue using the Nucleospin Plant II Kit (Machery-Nagel, Düren, Germany). The quality and concentration of DNA obtained was verified using a Nanodrop 2000 UV-Vis Spectrophotometer (ThermoScientific, Waltham, Massachusetts, U.S.A.).

We developed eleven nuclear microsatellite markers (see Nowell et al., 2015 for details) to assess the population genetics of *T. planifolium*. Amplification of microsatellite fragments were performed in 8 µl reaction volumes containing 0.416 µL primer mix (10 mM; 10:1 untagged to tagged primer), 0.192 µL dye-labelled CAG Tag (10 mM; 6-FAM or VIC, Life Technologies, Carlsbad, California, U.S.A.), 0.24 µL DMSO, 4 µL 2X Phusion High-Fidelity Master Mix with HF Buffer (New England Biolabs, Ipswich, Massachusetts, U.S.A.), 2.152 µL ddH2O and 1 µL DNA (10 ng/µL) using a T-100 Thermal Cycler (Bio-Rad, Hercules, California, U.S.A.). Cycling conditions followed Touchdown-TD PCR (Korbie and Mattick, 2008) as in Nowell et al. (2015). PCR products were visualized on a 1% agarose gel stained with GelRed (Biotium, Hayward, California, U.S.A.) and viewed with a High Performance Transilluminator (UVP, Upland, California, U.S.A.) with a 100 bp DNA ladder (New England Biolabs) to confirm the presence and size of amplification products and absence of contamination prior to genotyping.

Amplification products were subsequently pooled into groups of four and visualized by capillary electrophoresis using a 3130xL Genetic Analyzer (Life Technologies) with the GeneScan 500 LIZ Size Standard (Life Technologies). Individual samples were genotyped using Genemapper v.5 software (Life Technologies) and verified with manual scoring.

Due to amplification failures, we were unable to score all samples for all 11 markers. We retained only those samples for which we could score at least 8 loci. We enumerated private alleles, calculated allelic richness (El Mousadik and Petit, 1996; Adamack and Gruber, 2014), gene diversity, observed heterozygosity and the fixation index F_{IS} (Nei, 1987; Goudet and Jombart, 2015) for each population. Tests for departures from Hardy-Weinberg equilibrium (HWE) were performed with the R package 'pegas' ver. 0.13 (Paradis, 2010).

We counted the number of distinct multi-locus genotypes (MLG) for each population, using the methods provided by Kamvar et al. (2015). When samples with incomplete genotypes could not be unambiguously assigned to a single MLG, we randomly assigned them to a matching MLG in same population, if present. To account for differences in sample size, we also calculated rarefied MLG counts for each population, as a standardized measure of genetic richness. The rarefaction estimates were calculated for subsamples of 10 individuals per population (referred to as eMLG hereafter, Oksanen et al., 2019). Populations with fewer than 10 invidivduals scored were excluded from rarefaction, and all analyses based on eMLG.

We also assessed genetic divergence, using an index based on population pairwise-F_{ST} values. We calculated our divergence index (F_D) as the mean of pairwise-F_{ST} values for a given population, when compared with all other populations.

Preliminary review of the data revealed excessive levels of homozygosity, and no locus was in Hardy-Weinburg equilibrium. Five populations had a single MLG shared by all individuals and no heterozygosity (Table 1). As discussed below, this indicates high levels of selfing and/or inbreeding among close relatives. As such, analyses based on assumptions of Hardy-Weinburg equilibrium, linkage equilibrium and panmictic populations, such as STRUCTURE (Pritchard et al., 2000), are inappropriate for our data. Accordingly, we used multivariate ordinations to visualize genetic structure among populations (Jombart et al., 2008). Similarly, we did not test for the presence of null alleles, as we are not aware of a test that can distinguish between excessive homozygosity due to the presence of null alleles, and excessive homozygosity to due to high levels of inbreeding (e.g., Brookfield, 1996).

We assessed spatial structure using spatial Principal Components Analysis (i.e., sPCA: Jombart et al., 2008; Thioulouse et al., 2018). We defined the spatial neighbourhood of a sample as all other samples within 650 km, with the strength of each link weighted inversely by distance (via function `spdep::dnearneigh` with the upper distance bound set to 650). We chose 650 km to ensure there were no disconnected subgroups (i.e., there was at least one link between the Missouri populations and the eastern populations). We tested

for statistically significant spatial autocorrelation using a randomization test (function `global.rtest` in R package `adegenet`).

Soil Sampling and Analysis

Soil samples, composed of at least 7 sub-samples, were collected at each population (except WVMO and VASC, where the substrate was too rocky for soil collection, and KYLE, for which tissue samples were collected several years prior to the development of sampling design implemented here), with sub-samples made to rooting depth (~10 cm) immediately adjacent to *T. planifolium* plants. Composite samples were submitted to the University of Massachusetts Soil Laboratory (American Samples) or the University of Guelph Soil Laboratory (the Canadian sample), and soil properties including pH, cation exchange capacity (CEC), organic matter (OM) and soil nutrients (Calcium, Magnesium, Manganese, Phosphorous, Potassium, Zinc) were quantified.

We used hierarchical clustering to visualize differences in soil conditions. Each variable was scaled to mean = 0, standard deviation = 1, and the Euclidean distance among sites was used to construct UPGMA cluster dendrograms (Legendre and Legendre, 2012). We validated the dendrogram structure using the cophenetic correlation coefficient, and identified the optimal number of clusters using the Kelley-Gardner-Sutcliffe penalty function (Kelley et al., 1996).

Geographic and Ecological Analysis

We used herbarium records from GBIF.org (2019) to characterize the geographic isolation and climatic distribution of *T. planifolium*. To control for collection bias (see Radosavljevic and Anderson, 2014), we used the `spThin` package (Aiello-Lammens et al., 2015) to create a thinned data set, in which retained samples were at least 10 km from all other samples.

We quantified the isolation of each occurence in the thinned GBIF data as the fitted value of a kernel density smoother fit to those data, the "Utilization Distribution" of Worton (1989), and implemented in the R package `adehabitat` (Calenge, 2006). This approach incorporates both distance and density of occurences near a point, with low values (near zero) assigned to isolated points, and increasing values for points with more & closer neighbours. There is no upper bound to this value, but records used in our study had values from 0.007 (most isolated) to 0.113 (least isolated; see results).

To quantify climate suitability for observations, we constructed a Species Distribution Model (SDM) using a Maxent (Phillips et al., 2017). We limited the study extent to 31N-50N degrees latitude, 66W-98W degrees longitude; this area was chosen to include a 500 km buffer around a minimum convex polygon containing all GBIF records for *T. planifolium*. Background training data was limited to level 1 Ecoregions (Omernik and Griffith, 2014) in which *T. planifolium* is present within this extent. The following environmental layers were considered: 19 bioclimatic variables (WorldClim: Fick and Hijmans, 2017, www.worldclim.org), aridity (Trabucco and Zomer, 2019), and terrain ruggedness. Terrain ruggedness was calculated from SRTM DEM data (Jarvis et al., 2008) with 3-arc second (approximately 90m) horizontal resolution. We used the Terrain Ruggedness Index in QGIS (QGIS.org, 2020). The resulting raster was upscaled to match the 30 arc-second (~1km) resolution of the other layers, using the "average" resampling method. To reduce collinearity among variables, We selected a subset of these layers such that the variance inflation factor was less than 10 (function `vif` in the R package `usdm`: Naimi et al., 2014). The final set of variables included aridity, WorldClim Bio2 (mean diurnal temperature range), Bio6 (minimum temperature of coldest month), Bio13 (precipitation of wettest month), and Terrain Ruggedness. Maxent model parameters were tuned with the `ENMevaluate` function from the package `ENMeval` (Muscarella et al., 2014). We used the fitted cumulative log-log values from the SDM as an index of climate suitability for each record.

All analyses were performed in R version 4.0 (R Core Team, 2020), unless otherwise specified.

RESULTS

Genetics

We scored a total of 505 individual genotypes (mean 16.8 per population) for at least 8 loci, and completed all 11 loci for 408 individuals (Table 1). 46 alleles were detected (mean 4.2 per locus, range 2–9), and a total of 115 unique multi-locus genotypes (MLG) were described. All but three MLG were confined to a single population. One MLG was shared by two Missouri populations; one was shared by two Ohio populations; and one MLG was found in three populations, one each in Connecticut, Maryland, and eastern New York.

465 (92.1%) individuals were homozygous for all alleles, with only 40 individuals heterozygous at one or more loci. Of the 25 populations with at least 10 individuals genotyped, the mean eMLG (i.e., genotypes per 10 individuals) was 3.1 (range 1–6.5, Figure 2). The Canadian population had an eMLG of 1.37, corresponding to the 20th percentile of the distribution.

The pair-wise F_{ST} values ranged from 0 to 1 (mean 0.79). A group of five populations shared mutual pairwise-F_{ST} values of exactly 1: CTTP, MAST, OHBH, OHSC, VAHB. This is a consequence of each of these populations being fixed for a different homozygous MLG. Ohio populations OHBH and OHPR had a pairwise F_{ST} of -0.01 (corrected for heterozygosity, Goudet and Jombart, 2015), as a consequence of the single MLG present in OHBH (all 13 samples), also occuring in 16 of 18 samples from OHPR.

The mean genetic divergence, F_D, was 0.79 (range 0.59–0.9, Figure 2). F_D for the Ontario population was 0.87, corresponding to the 80th percentile of the distribution. The least divergent populations (in increasing order) were MOBU, WVFM, NYTS, NHER and PAGP (F_D = 0.60 – 0.68). The most divergent populations (in decreasing order) were NYMP, VAHB, MAST, CTTP, OHSC, ONRB (F_{ST} 0.87–0.9).

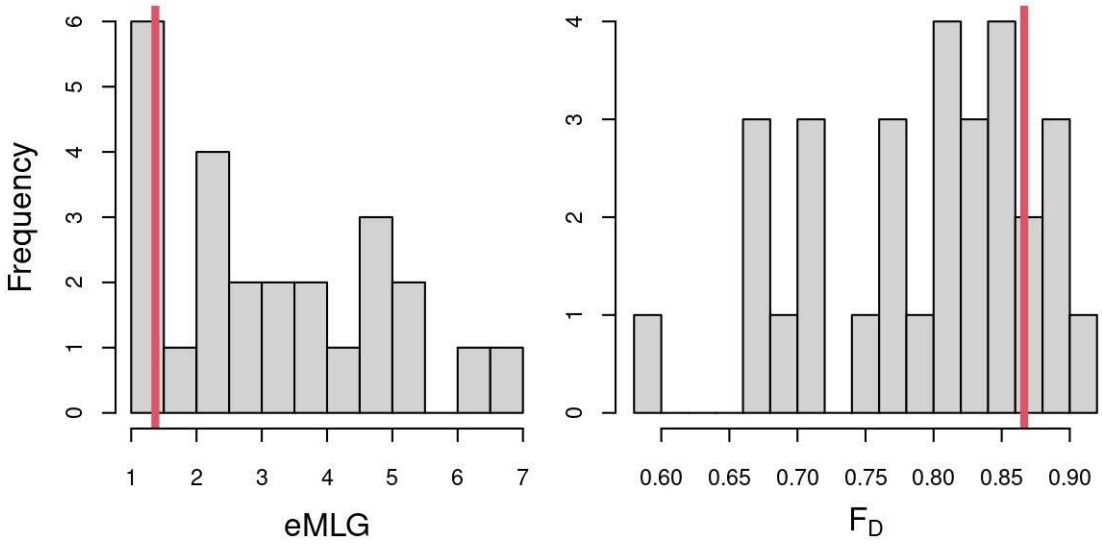

Figure 2: Distribution of rarified multi-locus genotypes (eMLG) and population mean pair-wise Fst values (F_D). The red line indicates the population in Ontario, Canada.

There was significant spatial structure in the genetic data, as determined by a Monte Carlo test of Morin's I (function `adespatial::global.rtest`, $p < 0.001$). This is largely due to the relative divergence of the Missouri populations (Figure 3). On the PCoA plot, the Missouri samples form a discrete grouping on the left side; the single Maryland population is somewhat separated in the upper right corner, and the Ontario samples cluster with the Missouri grouping on the lower left. The other populations are intermixed in the center of the plot. A similar pattern is shown in the sPCA map, which shows the Missouri samples have low values on sPCA axis 1, compared to higher values for the eastern populations. The populations in Ohio, in the center of the geographic range show a mix of 'eastern' and 'western' affinities.

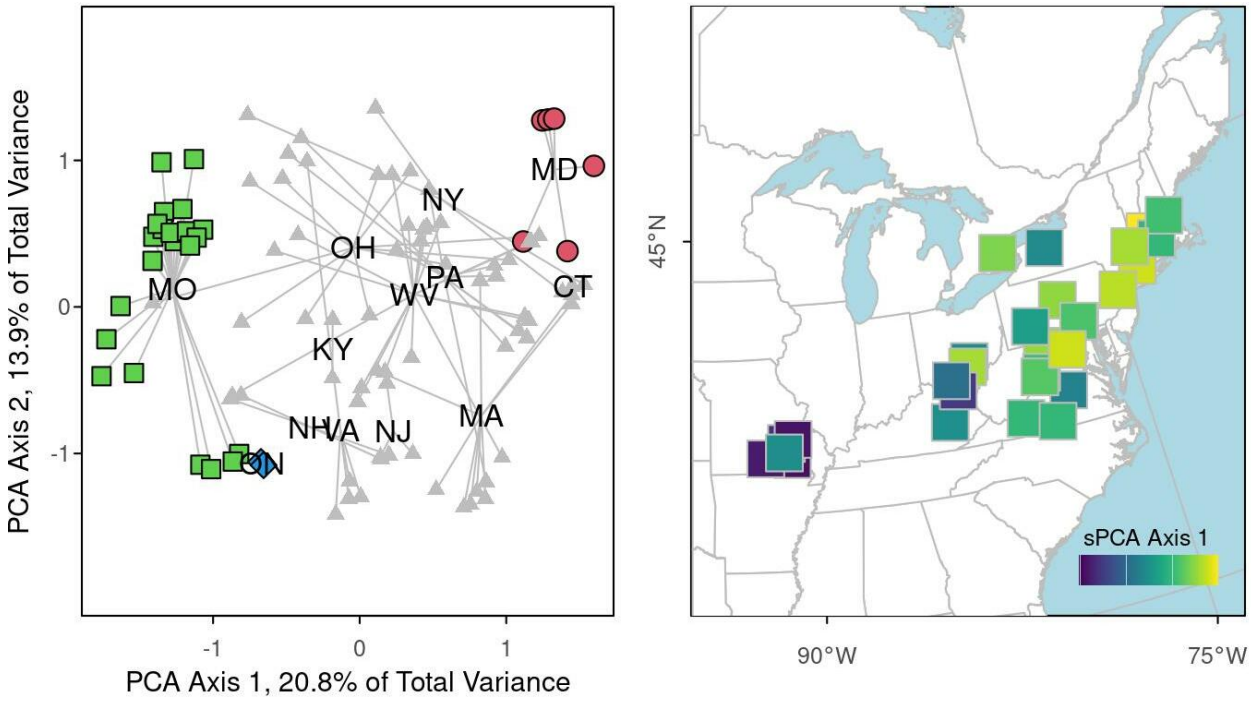

Figure 3: Spatial Structure of Genetic Diversity in *Trichophorum planifolium*. Left: Principal Coordinates Analysis of 97 unique microsatellite MLGs. The centroid of all samples from each state or province are labelled, connected to each of the unique MLGs for that jurisdiction. The strongest geographic structure is the clustering of Missouri samples on the left (MO, green squares), Maryland on the right (MD, red circles) and Ontario at the lower left (ON, blue diamonds). Right. sPCA map. The color for each population indicates its position along the first axis of the sPCA ordination.

Soils

The cophenetic correlation of the UPGMA clustering of soil data was 0.94, indicating strong correspondance between the dendrogram structure and the original distance values. The optimal number of clusters identified by the Kelley-Gardner-Sutcliffe test was four (Figure 4). 22 of 27 assessed populations belonged to the main cluster, with three outlying clusters containing one or two populations each. The Maryland population MDDM formed a cluster on its own, characterized by low pH, high organic matter, and high values of P, K, and Mn. Populations MABM and PAGP formed a cluster, characterized by high Ca and Mg. Populations ONRB and PACR formed the fourth cluster, characterized by high Ca and Zn and low Mn (Table 2).

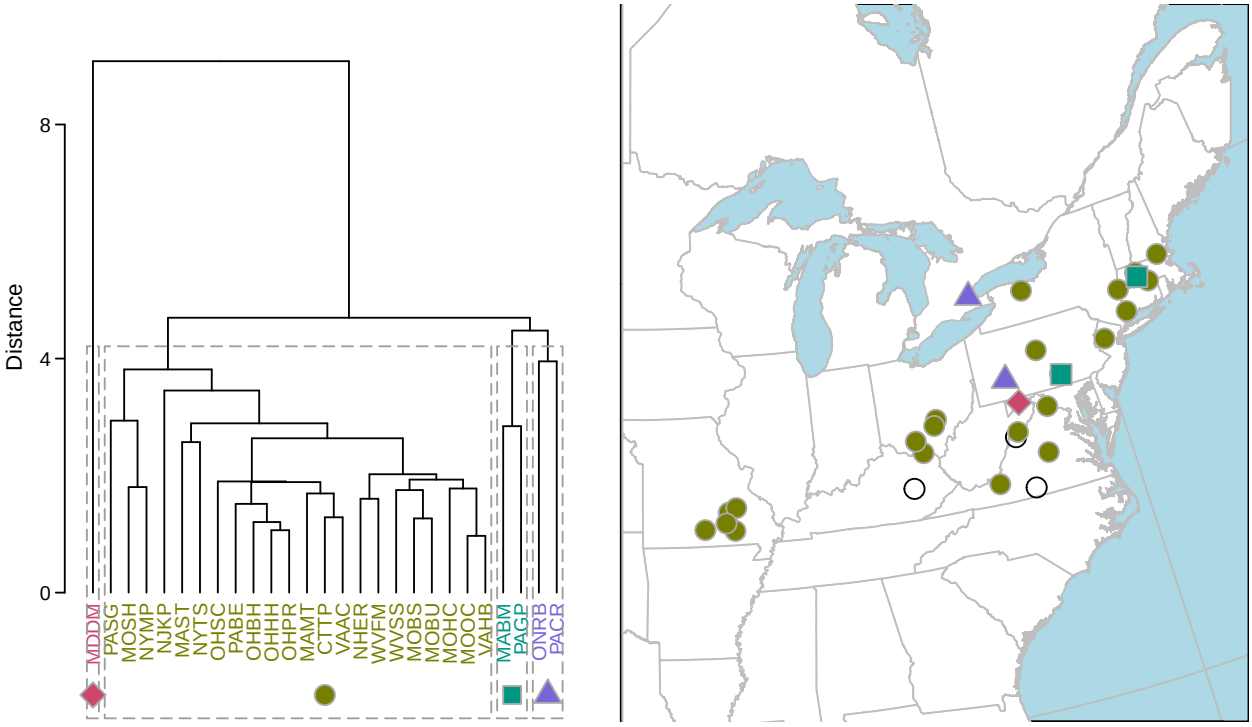

Figure 4: Soil Analysis of *T. planifolium* populations. Left: UPGMA dendrogram, based on 9 soil variables. The four clusters identified by the Kelley-Gardner-Sutcliffe test are indicated. Right: Geographic distribution of soil clusters. The plot symbols on the map indicate the soil cluster from the dendrogram; empty circles are populations without soil data.

Geographic Isolation and Climate Suitability

Population isolation for the herbarium records ranged from 0.00716 to 0.113 (lower values indicate greater isolation) (Figure **??**). The Canadian population had an isolation index of 0.00718, putting it in the 1st percentile of the 266 records. Climate suitability for each of the herbarium records ranged from 1.9e-5 to 4.5e-3; suitability at the Ontario population scored 3.6e-4, the 30th percentile of the total distribution (Figures 1 and **??**).

```
## pdf
##   2
```

DISCUSSION

The most striking observation in our study is the extremely low variation present within populations (mean F_{IS} 0.91, 1% heterozygosity), and high level of differentiation among populations (F_{ST} = 0.80). *Trichophorum planifolium* is apparently a wind-pollinated, outcrossing species, but these values are higher even than typical

Table 2: Soil analysis for *T. planifolium* populations. Values reported as mean ± standard deviation for each cluster. Note: the MDDM cluster had only one population, so there is no standard deviation to report.

	Main	MABM & PAGP	ONRB & PACR	MDDM
CEC (meq/100g)	16.0 ± 2.9	19.2 ± 1.8	12.5 ± 5.8	26.3
OM (%)	6.6 ± 2.9	12.2 ± 4.2	5.1 ± 0.5	20.3
pH	5.4 ± 0.5	5.7 ± 0.2	5.7 ± 0.4	4.2
Ca (ppm)	352.9 ± 275.0	1065.0 ± 200.8	1122.5 ± 321.7	341.0
K (ppm)	68.5 ± 15.5	106.5 ± 33.2	53.5 ± 19.1	160.0
Mg (ppm)	64.8 ± 46.0	186.5 ± 7.8	94.1 ± 79.3	44.0
Mn (ppm)	30.0 ± 13.8	50.2 ± 5.2	14.9 ± 5.6	271.8
P (ppm)	4.3 ± 2.0	4.0 ± 0.1	3.2 ± 3.4	15.8
Zn (ppm)	1.5 ± 0.7	3.7 ± 0.8	4.3 ± 0.2	2.8

for selfing species (i.e., the mean F_{ST} for selfing species with similar life history traits is 0.4–0.57, following Hamrick and Godt, 1996; Nybom, 2004).

This supports Crins' (1989) hypothesis that the low stature of *T. planifolium*, and the complex terrain it occupies, limits outcrossing. Crins also noted that *T. planifolium* is protandrous, presumably an adaptation to lessen the frequency of self-pollination. However, Friedman and Barrett (2009) showed that the spatial and temporal separation of male and female flowers in woodland sedges (monecious *Carex*) are largely ineffective at preventing geitonogamy. They studied seven *Carex* species, many of which co-occur with *T. planifolium* in woodland habitats (pers. obs) and have similar caespitose habit. They found only one species had an outcrossing rate greater than 0.1 (three species had no polymorphic allozymes, which may reflect a long history of inbreeding).

Lacking any obvious adaptations for long-distance seed dispersal, gene flow via seed dispersal is likely similarly limited in *T. planifolium*. It is well documented that plant seeds often travel only one or a few metres (Cain et al., 2000). Seed dispersal in *T. planifolium* occurs in late July and August, during which time the plants' leaves become matted. Crins (1989) suggested that this would facilitate deposition of the seeds immediately adjacent to parent plants (i.e., seeds are gravity dispersed), resulting in small spatial clumps of closely related tussocks. Like self-pollination, mating among closely related individuals limits recombination and leads to increased homogeneity of genotypes in the population (Loveless and Hamrick, 1984).

In combination, these life history characteristics may explain the low population genetic diversity documented here as the outcome of founder effects, followed by population expansion dominated by selfing. This is consistent with a general pattern noted in other sedges (*Carex*), which can be divided into two life history categories (Ford et al., 1998). "Group 1" are caespitose, with multiple hermaphroditic culms on each plant, with flowers in close proximity. This facilitates geitonogamy, resulting in low intra-population diversity, but high inter-population diversity (e.g., Waterway, 1990; Schell and Waterway, 1992; Whitkus, 1992 p. @KullOja_2007).

In contrast, "Group 2" *Carex* are rhizomatous, with more space between culms from the same individual, and more physical mixing among individuals. This leads to greater intra-population diversity, and less differentiation among populations (Ford et al., 1998).

A caespitose species incapable of rhizomatous spread, *Trichphorum planifolium* fits the description of *Carex* Group 1, aside from having bisexual flowers. This likely further exacerbates inbreeding, as it may engage in both autogamous and geitenogamous selfing. The low levels of heterozygosity we observed (Ho 0.00–0.03, Table 1) are comparable to values reported for other sedges: *Trichophorum caespitosum*: Ho = 0.01 to 0.04 (Godt et al., 1996), *Carex loliaceae*: Ho = 0.00 to 0.12 (Kull and Oja, 2007) and *Carex magellanica* subsp. *irrigua*: Ho = 0.00 to 0.03 (Kull and Oja, 2010).

The main geographic structure in the genetic data reveals a broad division between Missouri and the eastern populations (Figure 3). This suggests current *T. planifolium* populations may have emerged from glacial

refugia both in the Appalachian Mountains and on the Ozark plateau (Delcourt and Delcourt, 1991; Meyer, 1997; Soltis et al., 2006). However, if this is the case, it is interesting to note that populations representing both eastern and western genetic groups appear in close proximity in Ohio. The ordination analysis also indicates the Ontario population is genetically more similar to the Missouri populations than the geographically closer eastern group.

Gene flow via pollen or seed between Missouri in the west, and Ohio and Ontario to the east, is unlikely, given the long distances between these locations ($>$ 500 km). However, the SDM indicates pockets of suitable environment exist along the Ohio river in southern Indiana and northern Kentucky (Figure 1). This may have served as a post-glacial colonization route linking populations from the two regions; there may even be overlooked populations of *T. planifolium* in this area, as this species is easily missed.

The lone remaining Canadian population shows somewhat reduced diversity (20th percentile for eMLG; Figure 2). We have argued above that population dynamics in this species are characterized by bottlenecks followed by inbreeding. This is often associated with a reduction in fitness (Angeloni et al., 2011). However, this cost may be reduced in self-pollinating species (Holtsford and Ellstrand, 1990; Busch, 2005), and can be further mitigated if the inbreeding leads to purging of deleterious alleles (Crnokrak and Barrett, 2002).

We cannot address this issue directly with our data. Microsatellites are neutral markers, they do not provide direct evidence of adaptive diversity; they only provide insight into gene flow and reproductive strategies. We can say that *T. planifolium* populations with low diversity are relatively common and widespread (e.g., CT, MA, OH, VA), and several of them are quite large and apparently stable (personal observation). Furthermore, this pattern of limited intra-population genetic diversity is common in other woodland sedges. From this, we conclude that the reduced genetic diversity of the last Canadian population does not suggest inbreeding depression likely to be a threat to its persistence.

The Canadian population is moderately divergent genetically (80th percentile for F_D; Figure 2), and climatic conditions at that location are not extreme for the species (30th percentile for suitability). However, it is extremely isolated, and the soil conditions at the site are shared with only one other population. This provides indirect evidence that this population may be adapted to site-specific environmental conditions, indicating elevated conservation value (as per Lesica and Allendorf, 1995). Indeed, protection of many populations of *T. planifolium* would be necessary to conserve species-level genetic diversity as it is chiefly partitioned among, rather than within, populations (see Baskauf et al., 2014 or @KoellingEtAl_2011, for example). These results also suggest that risks of outbreeding depression are high with outcrossing, and future restorative efforts (i.e., augmentation) for the Canadian population should be avoided or approached with caution (Frankham et al., 2011). It should also be considered that mixing of populations with differing genetic backgrounds could decrease the conservation value of the Canadian population (Barrett and Kohn, 1991).

Acknowledgements

The authors thank the staff of Royal Botanical Gardens in Hamilton, Ontario, especially Natalie Iwanycki, who provided detailed reports on *T. planifolium* in Canada. Kasia Dadej (Agriculture and Agri-Food Canada) assisted with scoring the microsatellite data. The following people assisted with obtaining access to field sites: George Yaskievych, Rick Gardner, Tom Wieboldt, P. J. Harmon, Jim Vanderhorst, John Kunsman, Jeff Saarela, Steve Grund, Steve Glenn, Sara Cairns, Bryan Connolly, Robert Bertin, Karen Searcy, Rich Ring, Steven Daniel, Greg Iffrig, Brian Davidson, April Daniels, Dale Egbert, Lynn Prater, Nathan Jester, Rich McCarty, Dave Minney, Bill Stanely, Phil Hoffman, Glen Cobb, Levi Miller, Pearl's family, Bruce Metz, Fawn Kearns, Barbara Sargent, Tom Metzger, the Morris County Parks Commission, Becky Petro, Julie Richburg, Nancy Putnam, Jesse Jaycox, and Mark Quinn. Funding for this project was provided by the Ontario Ministry of Natural Resources Species at Risk Stewardship Fund (SAR-SF), the Environment and Climate Change Canada Interdepartmental Recovery Fund (IRF), and the Carleton University Bonneville Family Scholarship.

AUTHOR CONTRIBUTIONS

VJN and TWS conceived the study. VJN designed the field sampling, and VJN and TWS completed the sample collection. VJN and SW completed the lab experiments. VJN completed the analysis, with direction and supervision from TWS. VJN completed the first draft of the manuscript. TWS provided additional analyses and completed the final draft of the manuscript.

Data Availability

This manuscript and all associated data is available at the GitHub repository, https://github.com/plantarum/trich for inspection during peer review. This repository will be transfered to Zenodo for permanent, public archiving after completion of the peer review process.

Extra Figure

Added during revision

```
## pdf
##   2
```

References

Adamack, A. T., and B. Gruber. 2014. PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution* 5: 384–387.

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2015. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* 38: 541–545.

Allendorf, F. W., G. H. Luikart, and S. N. Aitken. 2013. Conservation and the genetics of populations, 2nd ed. Wiley.

Angeloni, F., N. J. Ouborg, and R. Leimu. 2011. Meta-analysis on the association of population size and life history with inbreeding depression in plants. *Biological Conservation* 144: 35–43.

Barrett, S. C. H., and J. R. Kohn. 1991. Genetic and evolutionary consequences of small population size in plants: implications for conservation. *In* D. A. Falk, and K. W. Holsinger [eds.], Genetics and Conservation of Rare Plants, 3–30. Oxford University Press, New York.

Baskauf, C. J., N. C. Jinks, J. R. Mandel, and D. E. McCauley. 2014. Population genetics of Braun's Rockcress (*Boechera perstellata*, Brassicaceae), an endangered plant with a disjunct distribution. *Journal of Heredity* 105: 265–275.

Brook, B. W., D. W. Tonkyn, J. J. O'Grady, and R. Frankham. 2002. Contribution of inbreeding to extinction risk in threatened species. *Conservation Ecology* 6.

Brookfield, J. F. Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* 5: 453–455.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. *The American Naturalist* 124: 255–279.

Bunnell, F. L., R. W. Campbell, and K. A. Squires. 2004. Conservation priorities for peripheral species: the example of British Columbia. *Canadian Journal of Forest Research* 34: 2240–2247.

Busch, J. W. 2005. Inbreeding depression in self-incompatible and self-compatible populations of *Leavenworthia alabamica*. *Heredity* 94: 159–165.

Cain, M. L., B. G. Milligan, and A. E. Strand. 2000. Long-distance seed dispersal in plant populations. *American Journal of Botany* 87: 1217–1227.

Calenge, C. 2006. The package 'adehabitat' for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling* 197: 516–519.

COSEWIC. 2000. COSEWIC assessment and status report on the bashful bulrush *Trichophorum planifolium* in Canada. Committee on the Status of Endangered Wildlife in Canada, Ottawa, Ontario, Canada.

Crandall, K. A., O. R. P. Bininda-Emonds, G. M. Mace, and R. K. Wayne. 2000. Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution* 15: 290–295.

Crins, W. J. 1989. Status of the few-flowered club-rush, *Scirpus verecundus* (Cyperaceae), in Canada. *Canadian Field-Naturalist*: 57–60.

Crins, W. J. 2002. *Trichophorum* Persoon. *In* Flora of North America Editorial Committee [ed.], Flora of North America north of Mexico, 28–31. Oxford University Press, New York.

Crins, W. J., and J. H. Rettig. 2002. *Carex* sect. *Acrocystis* Dumortier. *In* Flora of North America Editorial Committee [ed.], Flora of North America Volume 23: Magnoliophyta: Commelinidae (in part): Cyperaceae, 532–548. Oxford University Press, New York.

Crnokrak, P., and S. C. H. Barrett. 2002. Perspective: purging the genetic load: a review of the experimental evidence. *Evolution* 56: 2347–2358.

Delcourt, H. R., and P. Delcourt. 1991. Quaternary ecology: a paleoecological perspective. Springer Science & Business Media.

Eckert, C. G., K. E. Samis, and S. C. Lougheed. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology* 17: 1170–1188.

El Mousadik, A., and R. J. Petit. 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* 92: 832–839.

Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37: 4302–4315.

Ford, B. A., D. A. R. McQueen, R. F. C. Naczi, and A. A. Reznicek. 1998. Allozyme variation and genetic relationships among species in the *Carex willdenowii* complex (Cyperaceae). *American Journal of Botany* 85: 546–552.

Frankel, O. H., A. H. D. Brown, and J. J. Burdon. 1995. The conservation of plant biodiversity. Cambridge University Press.

Frankham, R., J. D. Ballou, M. D. B. Eldridge, R. C. Lacy, K. Ralls, M. R. Dudash, and C. B. Fenster. 2011. Predicting the probability of outbreeding depression. *Conservation Biology* 25: 465–475.

Friedman, J., and S. C. H. Barrett. 2009. The consequences of monoecy and protogyny for mating in wind-pollinated *Carex*. *New Phytologist* 181: 489–497.

GBIF.org. 2019. GBIF occurrence download. *https://doi.org/10.15468/dl.qoendw*. Website [accessed 17 December 2019].

Gibson, S. Y., R. C. Van Der Marel, and B. M. Starzomski. 2009. Climate change and conservation of leading-edge peripheral populations. *Conservation Biology* 23: 1369–1373.

Godt, M. J. W., B. R. Johnson, and J. L. Hamrick. 1996. Genetic diversity and population size in four rare southern Appalachian plant species. *Conservation Biology* 10: 796–805.

Goudet, J., and T. Jombart. 2015. hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22.

Hampe, A., and R. J. Petit. 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters* 8: 461–467.

Hamrick, J. L., and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 351: 1291–1298.

Holderegger, R., U. Kamm, and F. Gugerli. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology* 21: 797–807.

Holtsford, T. P., and N. C. Ellstrand. 1990. Inbreeding effects in *Clarkia tembloriensis* (Onagraceae) populations with different natural outcrossing rates. *Evolution* 44: 2031–2046.

Hunter, M. L., and A. Hutchinson. 1994. The virtues and shortcomings of parochialism: conserving species that are locally rare, but globally common. *Conservation Biology* 8: 1163–1165.

Iwanycki, N., B. McGoey, and K. Harrison. 2010. 2010 population assessment of the Few-flowered Club-rush (*Trichophorum planifolium*) at Cootes Paradise, Royal Botanical Gardens. Royal Botanical Gardens, Hamilton, Ontario.

Jarvis, A., H. I. Reuter, A. Nelson, and E. Guevara. 2008. Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT). *CGIAR-CSI SRTM – SRTM 90m DEM Digital Elevation Database*. Website http://srtm.csi.cgiar.org/ [accessed 31 July 2020].

Jombart, T., S. Devillard, A.-B. Dufour, and D. Pontier. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* 101: 92–103.

Kamvar, Z. N., J. C. Brooks, and N. J. Grünwald. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics* 6: 208.

Kelley, L. A., S. P. Gardner, and M. J. Sutcliffe. 1996. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. *Protein Engineering, Design and Selection* 9: 1063–1065.

Koelling, V. A., J. L. Hamrick, and R. Mauricio. 2011. Genetic diversity and structure in two species of *Leavenworthia* with self-incompatible and self-compatible populations. *Heredity* 106: 310–318.

Korbie, D. J., and J. S. Mattick. 2008. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols* 3: 1452–1456.

Kull, T., and T. Oja. 2010. Allozyme diversity and geographic variation among populations of the locally endangered taxon *Carex magellanica* subsp. *irrigua* (Cyperaceae). *Folia Geobotanica* 45: 323–336.

Kull, T., and T. Oja. 2007. Low allozyme variation in *Carex loliacea* (Cyperaceae), a declining woodland sedge. *Annales Botanici Fennici* 44: 267–275.

Legendre, P., and L. Legendre. 2012. Numerical ecology. 3rd Edition. Elsevier, New York.

Leppig, G., and J. W. White. 2006. Conservation of peripheral plant populations in California. *Madroño* 53: 264–274.

Lesica, P., and F. W. Allendorf. 1995. When are peripheral populations valuable for conservation? *Conservation Biology* 9: 753–760.

Loveless, M. D., and J. L. Hamrick. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* 15: 65–95.

Meyer, F. G. 1997. *Hamamelis* Linnaeus. *In* Flora of North America Editorial Committee [ed.], Flora of North America Volume 3: Magnoliophyta: Magnoliidae and Hamamelidae, Oxford University Press, New York.

Moritz, C. 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology* 51: 238–254.

Muscarella, R., P. J. Galante, M. Soley-Guardia, R. A. Boria, J. M. Kass, M. Uriarte, and R. P. Anderson. 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution* 5: 1198–1205.

Naimi, B., N. A. S. Hamm, T. A. Groen, A. K. Skidmore, and A. G. Toxopeus. 2014. Where is positional uncertainty a problem for species distribution modelling? *Ecography* 37: 191–203.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press.

Nowell, V. J., S. Wang, and T. W. Smith. 2015. Development and characterization of 11 microsatellite primers for the sedge *Trichophorum planifolium*(Cyperaceae). *Applications in Plant Sciences* 3: 1500050.

Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* 13: 1143–1155.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, et al. 2019. vegan: Community ecology package. R package version 2.5-6.

Omernik, J. M., and G. E. Griffith. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Environmental Management* 54: 1249–1266.

Paradis, E. 2010. pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics* 26: 419–420.

Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire, and M. E. Blair. 2017. Opening the black box: an open-source release of Maxent. *Ecography* 40: 887–893.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

QGIS.org. 2020. QGIS geographic information system. Open Source Geospatial Foundation Project.

Radosavljevic, A., and R. P. Anderson. 2014. Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of Biogeography* 41: 629–643.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Safriel, U. N., S. Volis, and S. Kark. 1994. Core and peripheral populations and global climate change. *Israel Journal of Plant Sciences* 42: 331–345.

Schell, C. M., and M. J. Waterway. 1992. Allozyme variation and the genetic structure of populations of the rare sedge *Carex misera* (Cyperaceae). *Plant Species Biology* 7: 141–150.

Smith, T. W., and C. J. Rothfels. 2007. Recovery strategy for Few-flowered Club-rush (*Trichophorum planifolium*) in Ontario. Ontario Ministry of Natural Resources, Peterborough, Ontario.

Soltis, D. E., A. B. Morris, J. S. McLACHLAN, P. S. Manos, and P. S. Soltis. 2006. Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* 15: 4261–4293.

Soulé, M. E., and L. S. Mills. 1998. No need to isolate genetics. *Science* 282: 1658–1659.

Teixeira, J. C., and C. D. Huber. 2021. The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences* 118.

Thiers, B. 2020. Index Herbariorum: a global directory of public herbaria and associated staff. New York Botanical Garden. Website http://sweetgum.nybg.org/science/ih/.

Thioulouse, J., S. Dray, A.-B. Dufour, A. Siberchicot, T. Jombart, and S. Pavoine. 2018. Multivariate analysis of ecological data with ade4. Springer New York, New York, NY.

Trabucco, A., and R. Zomer. 2019. Global aridity index and potential evapotranspiration (ET0) climate database v2. 1705236666 Bytes.

Waterway, M. J. 1990. Genetic differentiation and hybridization between *Carex gynodynama* and *C. mendocinensis* (Cyperaceae) in California. *American Journal of Botany* 77: 826–838.

Whitkus, R. 1992. Allozyme variation within the *Carex pachystachya* complex (Cyperaceae). *Systematic Botany* 17: 16–24.

Worton, B. J. 1989. Kernel Methods for Estimating the Utilization Distribution in Home-Range Studies. *Ecology* 70: 164–168.

Yakimowski, S. B., and C. G. Eckert. 2007. Threatened peripheral populations in context: geographical variation in population frequency and size and sexual reproduction in a clonal woody shrub. *Conservation Biology* 21: 811–822.

Young, A., T. Boyle, and T. Brown. 1996. The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution* 11: 413–418.