Canadian Science Publishing | Botany

Conservation assessment of a range-edge population of Trichophorum planifolium (Cyperaceae) reveals range-wide inbreeding and locally divergent environmental conditions

Journal:	*Botany*
Manuscript ID	cjb-2021-0195.R2
Manuscript Type:	Article
Date Submitted by the Author:	28-Jan-2022
Complete List of Authors:	Nowell, Victoria J.; Agriculture and Agri-Food Canada Ottawa Research and Development Centre, Botany; Carleton University, Department of Biology Wang, Song; Agriculture and Agri-Food Canada Ottawa Research and Development Centre, Botany Smith, Tyler; Agriculture and Agri-Food Canada Ottawa Research and Development Centre, Botany; Carleton University, Department of Biology
Keyword:	abundant center hypothesis, conservation genetics, microsatellites, Trichophorum planifolium, peripheral populations
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

Conservation assessment of a range-edge population of *Trichophorum planifolium* (Cyperaceae) reveals range-wide inbreeding and locally divergent environmental conditions

Victoria J. Nowell[1,2], Song Wang[1], and Tyler W. Smith[1,2,3]

[1]Agriculture and Agri-Food Canada, Ottawa Research and Development Centre, 960 Carling Avenue, Ottawa, Ontario K1A 0C6
[2]Carleton University, Department of Biology, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6
[3]Corresponding author: tyler@plantarum.ca, https://orcid.org/0000-0001-7683-2653

Abstract

Peripheral populations hold special conservation significance as repositories of genetic diversity, while they may also be at increased risk of extirpation. We collected genetic, ecological, and distributional data to evaluate an endangered range-edge population of *Trichophorum planifolium*. Our data show that, range-wide, *T. planifolium* has remarkably low within-population diversity, and most populations are dominated by a small number of unique genotypes. However, the extremely low observed and expected heterozygosity we document is comparable to results from other woodland sedges. The life history of this species results in high levels of inbreeding. The low diversity of the single remaining Canadian population is therefore not unusual for this species, and does not support suggestions that this population suffers from elevated inbreeding depression. However, it is geographically isolated, and its soil environment is unusual for the species. These conditions may have facilitated local adaptation of conservation value for the long-term persistence of this species. In addition, we identified populations within the core range of this species that are comparably isolated to the peripheral population, or that demonstrate greater genetic divergence. Consequently, while distinctions between core and peripheral populations are important for conservation management, understanding variation across the range of a species is key to effective protection.

Keywords

abundant center hypothesis; conservation genetics; microsatellites; *Trichophorum planifolium*; peripheral populations

Introduction

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) and climate change. 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 (Sagarin and Gaines 2002; Eckert et al.

2008). The lack of consistent range-edge genetic differentiation may be attributable in part to species-specific characteristics (Gibson et al. 2009). In plants, life-history traits (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 are 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 nor nectaries (Crins 1989; Iwanycki et al. 2010). Seeds mature in mid-summer and disperse in late July to August (Iwanycki et al. 2010). A specific mechanism for long-distance 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 maternal 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 evaluate 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. During sampling, we estimated the size of each population using four categories: fewer than 100, 100-500, 500-1000 and more than 1000 individuals. 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. Twenty samples per population were randomly selected for genetic analysis. For populations with fewer than 20 samples, all individuals were used.

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 employed eleven nuclear microsatellite markers that were previously developed to assess the population genetics of *T. planifolium* (Table 2, see Nowell et al. 2015 for details). 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 three 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 loci. We retained only those samples for which we could score at least eight 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 the 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.

To determine if population size had an influence on genetic diversity (eMLG) or divergence (F_D), we used a linear regression of these variables against population size. Population size categories were coded as values 1-4 (very small, small, medium and large) for this analysis.

Preliminary review of the data revealed excessive levels of homozygosity, and no locus was in Hardy-Weinburg equilibrium (data not shown). 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 ordination 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 visualized genetic variation among the 97 complete, distinct MLG using a Principal Components Analysis of allele frequencies for each MLG. We assessed spatial structure using spatial Principal Components Analysis (i.e., sPCA: Jombart et al. 2008; Thioulouse et al. 2018). This is a form of co-inertia analysis that identifies the spatial pattern present in population genetic data. The SSR data is incorporated as population allele frequencies.

Spatial relationships are defined by a neighbourhood. We defined the spatial neighbourhood of a population as all other populations within 650 km, with the strength of each link weighted inversely by geographic 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 used the `adegenet` package to complete this procedure (via function `spca`, Jombart 2008). 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 seven 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 the 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 niche 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 occurrence 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 occurrences near a point, with low values (near zero) assigned to isolated points, and increasing values for points with more and 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 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.

We quantified the degree of divergence of populations relative to the species as a whole for a number of variables (eMLG, F_D, Spatial Isolation, Climate Suitability) using the percentile of that population. Populations in the first/last 5 percentiles for a variable (i.e., having a value greater/less than 95% of all populations) were considered highly divergent for that variable. Populations in the top/bottom 20 percentiles were considered moderately divergent. 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 eight loci, and completed all 11 loci for 408 individuals (Table 1, Table 2). Forty-six alleles were detected (mean 4.2 per locus, range 2–9, Table 2), and a total of 115 unique 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.

Four hundred and sixty-five (92.1%) individuals were homozygous for all loci, 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 occurring in 16 of 18 samples from OHPR.

The mean genetic divergence, F_D, was 0.79 (range 0.59–0.9, Table 1, 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).

Estimated population size did not have a statistically significant relationship with either eMLG (regression adj. R^2 = 0.016, p = 0.24), or F_D (adj. R^2 = -0.03, p = 0.84) (Figure 3).

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. On the (non-spatial) PCA plot the Missouri samples form a discrete grouping; the single Maryland population is somewhat separated, and the Ontario samples cluster with the Missouri grouping (Figure 4). The other populations are intermixed in the center of the plot without obvious geographic structure. 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.

Soils

The cophenetic correlation of the UPGMA clustering of soil data was 0.94, indicating strong correspondence between the dendrogram structure and the Euclidean distances calculated from the soil variables. The optimal number of clusters identified by the Kelley-Gardner-Sutcliffe test was four (Figure 5). Twenty-two 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 3).

Geographic Isolation and Climate Suitability

Population isolation for the GBIF herbarium records ranged from 0.00716 to 0.113 (lower values indicate greater isolation) (Figure 6, Table 1). The Canadian population had an isolation index of 7.18e-3, putting it in the 1st percentile of the 266 records. Of the populations we sampled, KYLE and NYMP were also in the 1st percentile, with isolation index values of 7.19e-3 and 7.48e-3 respectively (Table 1). 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 6).

DISCUSSION

The most striking observation in our study is the extremely low genetic variation and high homozygosity present within populations (mean F_{IS} 0.91, H_O 0.01), and high level of differentiation among populations (F_{ST} = 0.80). These values are higher even than typical for selfing species (i.e., the mean F_{ST} for selfing species with similar life history traits is 0.42–0.57, following Hamrick and Godt 1996; Nybom 2004), and as a wind-pollinated species we expect *T. planifolium* is at least potentially an out-crossing species. Our data suggest *T. planifolium* has a mixed mating system, and the high levels of homozygosity suggest reproduction is dominated by inbreeding.

This supports Crins' (1989) hypothesis that the low stature of *T. planifolium*, and the complex terrain it occupies, limit outcrossing and leads to inbreeding. 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 (monoecious *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 meters (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 inbreeding and 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 and inbreeding. 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 (Kull and Oja 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).

As 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 result 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, representing both Group 1 and Group 2: *Trichophorum caespitosum*: Ho = 0.01 to 0.04 (Godt et al. 1996; densely caespitose as in Group 1), *Carex loliaceae*: Ho = 0.00 to 0.12 (Kull and Oja 2007; rhizomatous, Group 2) and *Carex magellanica* subsp. *irrigua*: Ho = 0.00 to 0.03 (Kull and Oja 2010; rhizomatous, Group 2).

Strong founder effects may also explain the apparent lack of relationship between population census size and genetic diversity (Figure 3). Small populations are more susceptible to the loss of alleles through genetic drift (Wright 1931), and so are expected to maintain lower levels of genetic diversity than larger populations. This pattern has been documented in many systems (Frankham 1996), including woodland sedges (Godt et al. 1996). However, if populations are founded by a small number of individuals, and dominated by selfing, as we propose is the case for *T. planifolium*, populations may never contain enough variation for drift to have a detectable impact. In such cases population size

may be less important than migration rate: the number of different source populations that have contributed individuals to the given population may be more influential than the total number of individuals present at a given time. We cannot rigorously assess this issue with our data, but it does suggest that further study of metapopulation dynamics may be fruitful in understanding genetic diversity in understory sedges like *T. planifolium*.

The main geographic structure in the genetic data reveals strong differentiation between Missouri and the eastern populations (Figure 4). This suggests current *T. planifolium* populations may have emerged from separate eastern and western glacial refugia (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 single remaining Canadian population shows somewhat reduced diversity (20th percentile for eMLG; Figure 2). We have argued above that population dynamics in this species is characterized by founder effects followed by inbreeding, which

is often associated with reduced fitness (Angeloni et al. 2011). However, this cost may be reduced in self-pollinating species (Holtsford and Ellstrand 1990; Busch 2005), as in some cases inbreeding may result in the 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 into patterns of colonization, reproductive strategies and gene flow. 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 (Table 1). 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 is not unusual for forest sedges, and as such does not suggest inbreeding depression is 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 (Koelling et al. 2011). These results also suggest the risk of outbreeding depression, 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). While we have focused on the status of the peripheral Canadian population, other populations are similarly isolated (e.g., KYLE, NYMP), or more divergent genetically (e.g., NYMP, VAHB). This indicates that conserving intraspecific diversity requires assessing the full range of a species. Prioritizing populations for protection based on their geographic location or the political jurisdiction they occur in risks overlooking important variation within the distribution of a species.

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, and Paul O'Hara, whose extensive field studies of *Trichophorum planifolium* in Canada helped set this project in motion. 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

The data and code used to complete the analyses in this paper are archived with Zenodo (Smith 2022).

References

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

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., and Anderson, R. P. 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., Luikart, G. H., and Aitken, S. N. 2013. Conservation and the genetics of populations, 2nd ed. Wiley.

Angeloni, F., Ouborg, N. J., and Leimu, R. 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 Kohn, J. R. 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. Oxford University Press, New York. Pp. 3–30.

Baskauf, C. J., Jinks, N. C., Mandel, J. R., and McCauley, D. E. 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., Tonkyn, D. W., O'Grady, J. J., and Frankham, R. 2002. Contribution of inbreeding to extinction risk in threatened species. Conservation Ecology **6**: 16.

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., Campbell, R. W., and Squires, K. A. 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., Milligan, B. G., and Strand, A. E. 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., Bininda-Emonds, O. R. P., Mace, G. M., and Wayne, R. K. 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 **103:** 57-60.

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

Crins, W. J., and Rettig, J. H. 2002. *Carex* sect. *Acrocystis* Dumortier. In: Flora of North America Editorial Committee ed., Flora of North America Volume 23: Magnoliophyta: Commelinidae (in part): Cyperaceae. Oxford University Press, New York. pp. 532–545.

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

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

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

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

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

Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. Conservation Biology 10(6): 1500–1508. doi:10.1046/j.1523-1739.1996.10061500.x.

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

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

GADM. 2022. GADM Maps and Data. Available from https://gadm.org/.

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

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

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

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

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

Hamrick, J. L., and Godt, M. J. W. 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., Kamm, U., and Gugerli, F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landscape Ecology **21**: 797–807.

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

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

Iwanycki, N., McGoey, B., and Harrison, K. 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., Reuter, H. I., Nelson, A., and Guevara, E. 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. (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24: 1403-1405. doi: 10.1093/bioinformatics/btn129

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

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

Kelley, L. A., Gardner, S. P., and Sutcliffe, M. J. 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., Hamrick, J. L., and Mauricio, R. 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 Mattick, J. S. 2008. Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nature Protocols **3**: 1452–1456.

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

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

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

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

Loveless, M. D., and Hamrick, J. L. 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 [eds.] 1993+. Flora of North America North of Mexico [Online]. 22+ vols. New York and Oxford. Vol. 3. http://beta.floranorthamerica.org/Hamamelis. [Accessed 20 February 2022]

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

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., and Anderson, R. P. 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., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., and Toxopeus, A. G. 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., Wang, S., and Smith, T. W. 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., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., and Wagner, H. 2019. vegan: Community ecology package. R package version 2.5-6.

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

Pritchard, J. K., Stephens, M., and Donnelly, P. 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.

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

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

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

Sagarin, R.D., and Gaines, S.D. 2002. The 'abundant centre' distribution: to what extent is it a biogeographical rule? Ecology Letters 5(1): 137–147. doi:10.1046/j.1461-0248.2002.00297.x.

Schell, C. M., and Waterway, M. J. 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. 2022. plantarum/trich: *Trichophorum planifolium* SSR data and analysis (v1.0.1). Zenodo. Available from https://doi.org/10.5281/zenodo.5850461.

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

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

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

Teixeira, J. C., and Huber, C. D. 2021. The inflated significance of neutral genetic diversity in conservation genetics. Proceedings of the National Academy of Sciences **118:** e2015096118.

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

Trabucco, A., and Zomer, R. 2019. Global aridity index and potential evapotranspiration (ET0) climate database v2. DOI: 10.6084/M9.FIGSHARE.7504448.V3.

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.

Wright, S. 1931. Evolution in Mendelian Populations. Genetics 16(2): 97–159.

Yakimowski, S. B., and Eckert, C. G. 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., Boyle, T., and Brown, T. 1996. The population genetic consequences of habitat fragmentation for plants. Trends in Ecology & Evolution **11**: 413–418.

Tables

Table 1: *Trichophorum planifolium* populations sampled. Pop Code: the population code, the first two letters indicate the US state or Canadian province. Lat and Lon: geographic coordinates. Pop Size: estimated population size, V Small: < 100 individuals; Small 100-500 individuals; Medium 500-1000 individuals; Large > 1000 individuals. 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: MLG 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. I: geographic isolation index, × 1000; lower values indicate more isolated populations. The Mean row contains the population means; the Total row contains values calculated for all samples as a single population. The Total value for the F_D column is the global F_{ST} value.

Pop Code	Lat	Lon	Pop Size	N	Ap	Ar	MLG	eMLG	Hs	Ho	F_{IS}	F_D	I
CTTP	41.371	-73.452	Medium	17(16)	0	1.00	1	1.00	0.00	0.00		0.89	64
KYLE	37.591	-83.782	V Small	5(2)	1	1.07	3		0.05	0.02	0.60	0.84	7
MABM	42.302	-72.530	Large	19(14)	1	1.12	5	3.63	0.07	0.00	0.93	0.81	50
MAMT	42.447	-72.539	Large	19(19)	0	1.06	3	2.05	0.03	0.03	0.13	0.85	39
MAST	42.059	-72.130	Small	20(20)	0	1.00	1	1.00	0.00	0.00		0.89	72

MDDM	39.568	-78.912	V Small	20(19)	0	1.32	7	4.53	0.18	0.02	0.87	0.76	11
MOBS	36.948	-90.993	Large	41(17)	2	1.31	17	6.53	0.16	0.01	0.93	0.70	34
MOBU	37.536	-91.212	Small	14(11)	1	1.45	6	5.34	0.25	0.01	0.97	0.60	54
MOHC	37.023	-92.168	Large	14(12)	0	1.11	4	3.43	0.06	0.01	0.90	0.83	40
MOOC	37.664	-90.900	Small	17(12)	0	1.03	3	2.18	0.02	0.01	0.65	0.86	61
MOSH	37.200	-91.326	Small	20(15)	0	1.14	5	4.21	0.07	0.01	0.84	0.78	50
NHER	42.764	-71.361	Small	19(14)	1	1.27	4	3.51	0.15	0.03	0.79	0.68	26
NJKP	40.761	-74.716	Medium	17(16)	0	1.14	3	2.18	0.07	0.01	0.93	0.82	63
NYMP	43.030	-77.574	Large	19(19)	0	1.02	2	1.53	0.01	0.00	1.00	0.90	7
NYTS	42.118	-73.497	Large	20(15)	0	1.29	3	2.89	0.17	0.00	1.00	0.67	38
OHBH	39.621	-82.405	Small	13(13)	0	1.00	1	1.00	0.00	0.00		0.80	18
OHHH	39.438	-82.531	Small	20(20)	0	1.20	7	5.32	0.11	0.00	0.96	0.71	19
OHPR	38.653	-83.164	V Small	18(17)	1	1.04	3	2.11	0.02	0.00	1.00	0.77	18
OHSC	39.058	-83.389	Medium	17(15)	0	1.00	1	1.00	0.00	0.00		0.87	18
ONRB	43.266	-79.921	Small	27(22)	1	1.01	2	1.37	0.01	0.00	1.00	0.87	7
PABE	41.044	-77.654	V Small	15(7)	0	1.11	4	3.31	0.06	0.00	1.00	0.82	12
PACR	40.370	-79.228	V Small	6(5)	0	1.06	2		0.04	0.00	1.00	0.82	10
PAGP	40.075	-76.909	Medium	18(13)	0	1.34	7	4.71	0.19	0.05	0.76	0.68	33

PASG	40.098	-76.948	V Small	9(8)	0	1.04	2		0.02	0.00	1.00	0.76	33
VAAC	37.804	-78.249	Medium	17(14)	0	1.09	3	2.59	0.05	0.00	1.00	0.81	13
VAHB	37.189	-80.419	Small	18(14)	0	1.00	1	1.00	0.00	0.00		0.90	29
VASC	36.820	-79.067	V Small	3(1)	0	1.09	2		0.09	0.00	1.00	0.75	14
WVFM	38.664	-79.240	Large	19(18)	2	1.30	9	6.08	0.16	0.01	0.97	0.67	27
WVMO	38.519	-79.371	Large	20(18)	1	1.30	7	4.92	0.16	0.00	1.00	0.72	26
WVSS	39.222	-77.820	V Small	4(2)	0	1.07	2		0.05	0.00	1.00	0.85	40
Mean				16.8(13.6)	0.37	1.13	7.6	3.3	0.08	0.01	0.89	0.79	
Total				505(408)			115	9.03	0.08	0.01	0.91	0.80	

Table 2: Microsatellite loci. See Nowell et al. 2015 for details.

Locus	Motif	Alleles	Size
TP142	AAT	3	136–148
TP152	AG	3	146–156
TP174	AAT	3	167–170
TP325	AG	9	151–177
TP326	AG	2	216–226
TP330	AAT	3	341–358
TP341	AGC	7	284–323
TP406	AAT	5	198–216
TP434	AAT	3	163–169
TP45	AAG	2	382–368
TP80B	AAT	6	208–226

640

Table 3: 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

645

Figure Captions

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. Lambert Conformal Conic projection with NAD83 datum, shapefiles from GADM (2022).

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

Figure 3: Genetic diversity (eMLG) and divergence (F_D) by population size. V small: < 100 individuals; Small: 100-500; Medium: 500-1000; Large: 1000+.

Figure 4: Spatial structure of genetic diversity in *Trichophorum planifolium*. A. Principal Coordinates Analysis of 97 unique microsatellite MLGs. The centroids 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 (MO, green squares), Maryland (MD, red circles) and Ontario (ON, blue diamonds). B. sPCA map. The color for each population indicates its position along the first axis of the sPCA ordination. Lambert Conformal Conic projection with NAD83 datum, shapefiles from GADM (2022).

Figure 5: Soil analysis of *T. planifolium* populations. A. UPGMA dendrogram, based on the Euclidean distance among sites, calculated from nine soil variables. The four clusters identified by the Kelley-Gardner-Sutcliffe test are indicated. B. 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. Lambert Conformal Conic projection with NAD83 datum, shapefiles from GADM (2022).

Figure 6: Geographic isolation and climate suitability of *T. planifolium* records. Population Isolation based on the density of occurrences around each GBIF record using the Utilization Distribution; Climate Suitability presents the fitted values for each record from a Maxent distribution model. Red lines indicate the values for the last remaining Canadian population.