SMALL-SCALE INVASION DYNAMICS OF A NATIVE TREE, *JUNIPERUS VIRGINIANA*, IN OHIO

A thesis submitted

to Kent State University in partial

fulfillment of the requirements for the

degree of Master of Science

by

Hannah M. Hartman

December, 2023

Thesis written by

Hannah M. Hartman

B.S., Seton Hill University, 2018

M.S., Kent State University, 2023

Approved by

Oscar J. Rocha, Ph.D. _____, Advisor

Heather Caldwell, Ph.D. _____, Chair, Department of Biological Sciences

Mandy Munro-Stasiuk, Ph.D. _____, Dean, College of Arts and Sciences

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to the greatest advisor I could ask for, Dr. Oscar Rocha. Your dedication to your students, humor, passion for science, and patience were crucial to my development as a student and researcher. Thank you for welcoming me into your lab, letting me watch your dog, and for always helping with field work. I am a better student, scientist, and educator because I worked with you, and I will always be grateful for that.

I would like to thank my brothers and my chosen family of friends for your unwavering support in all aspects of my life, for lifting me up in tough times, and for celebrating with me in happy times. My family wouldn't be complete without my cats, who supported me immensely every day. To my family, friends, and cats, thank you for your love and care.

Thank you to my committee members, Dr. David Ward and Dr. Sangeet Lamichhaney. Thank you for your feedback and for encouraging me to push myself in my research. I enjoyed not only our meetings, but dinners, trips, and conversations with you as well. Thank you for your wisdom and for enriching my experience at Kent State. I would also like to thank all of those who collaborated with me on my research, especially those on the ERC project and Dr. Andrea Case, who inspired me to delve into bioinformatics. Thank you all for challenging me as a researcher and a student.

Additionally, I am grateful for the support of the Department of Biological Sciences at Kent State. Specifically, I would like to thank Susan Kieklak, Michelle Kane-Sutton, and Lajuane Mathis for their administrative support and kindness in answering my many questions. I am thankful for the support of Shelley Jurkiewicz and all of the Teaching Assistants I had the

pleasure of working with at Kent State. It was a joy to serve as a TA at Kent State and my teaching skills improved greatly with the support of these individuals.

Finally, I would like to thank those who supported my research financially. Thank you to The National Science Foundation (DEB 2214146), The Graduate Student Senate at Kent State University, and the Art and Margaret Herrick Aquatic Ecology Research Facility Student Research Grant for funding this work.

CHAPTER I

INTRODUCTION

In this thesis, I use various genetic diversity measures to explore the patterns that drive the colonization of new habitats by the native invader, *Juniperus virginiana* (eastern redcedar). Woody plants like eastern redcedar (ERC) threaten native grassland species by altering ecosystems' species diversity, nutrient cycling, soil. chemistry, natural disturbance regimes and hydrology (Eldridge et al. 2011, Ratajczak et al. 2012, Donovan et al. 2018, Ward 2020). Dispersal patterns affect the genetic diversity and structure of plant populations, which drive plant evolution and the ways in which a population responds to selection pressures (Sork and Smouse 2006, Chybicki and Oleksa 2018). By studying dispersal patterns, I aim to provide a more detailed understanding of how newly established populations of ERC may persist over time.

In this thesis, I seek to answer the question: is ERC's pattern of range expansion driven by diffusion from the edge of the range or by long-distance dispersal events? By answering this question, I may be able to provide suggestions for improvement of monitoring and management of grasslands in response to encroaching ERC populations.

Thesis Structure. I have constructed this thesis in the following way: Chapter I is an introduction to the issue of woody encroachment by native invaders like *Juniperus virginiana* with a brief overview of the plant itself and of my study.

In Chapter II, (*Temporal changes in genetic diversity reveal small-scale invasion dynamics of* Juniperus virginiana *L. var.* virginiana *in the Lakeside Daisy State Nature Preserve in Ohio*), I present the results of a genetic analysis of a small population of ERC in northern Ohio. I examined the levels of genetic diversity and genetic structure of a recently-established population of ERC in the Lakeside Daisy State Nature Preserve (LDSNP). This study showed that ERC encroachment of LDSNP resulted from multiple and reiterated gene flow events from the edge of the range, mainly through animal-mediated seed dispersal. Furthermore, the identification of multiple distinct groups of genetically similar individuals inspired research into microclimates at LDSNP and how it may filter ERC establishment in the preserve. A portion of this chapter has been accepted for publication with *Invasive Plant Science and Management*.

In Chapter III, (*Microclimate may affect genetic diversity of* Juniperus virginiana *L. var.* virginiana *in the Lakeside Daisy State Nature Preserve in Ohio*), I examine the possible relationship between genetic diversity and microclimate diversity at LDSNP. Climatic conditions are known to limit the range of a species and can impact the genetic diversity of an invasive plant population. LDSNP is a unique preserve because in some areas its soil is very shallow and excessively drained, but in other areas glacial till and lake sediments have accumulated on top of the bedrock leading to deeper soil that frequently floods. The differences between soil types are fine-scale and may cause small-scale environmental filtering of incoming ERC. I studied the correlation between genetic diversity and geographic location and soil type at LDSNP by observing differentially expressed loci in trees living in different microclimates. A portion of this chapter has been accepted for publication with *Invasive Plant Science and Management.*

Finally, Chapter IV highlights the general conclusions of the results presented in this thesis and suggests future possibilities for further research.

CHAPTER II

TEMPORAL CHANGES IN GENETIC DIVERSITY REVEAL SMALL-SCALE INVASION DYNAMICS OF *JUNIPERUS VIRGINIANA* L. VAR. *VIRGINIANA* IN THE LAKESIDE DAISY STATE NATURE PRESERVE IN OHIO

A portion of this work is published in

Hartman, H. M. & Rocha, O. J. (2023). Temporal changes in genetic diversity reveal small-scale invasion dynamics of the eastern Redcedar (*Juniperus virginiana* L. var. *virginiana*) in the Lakeside Daisy State Nature Preserve in Ohio. *Invasive Plant Science and Management*, 1–32. https://doi.org/10.1017/inp.2023.23

Author Contributions: HMH and OJR designed the study, collected the data, performed data and statistical analyses, and wrote the manuscript.

INTRODUCTION

There is ample evidence that native and nonnative species invasions have similar impacts on biodiversity (Sagoff 1999, Simberloff et al. 2012, Schnelle 2019, Yazlik and Ambarli 2022). While most invasive species are not native, there are native species that have become too abundant and have been implicated as drivers of recent extinctions (Blackburn et al. 2019). Some native invaders became problematic after intracontinental movements outside their native range into new suitable habitats. However, other species expand their range from their edges and become increasingly dominant because of their prolific fruit production, seed dispersal by birds

and mammals, anthropogenic disturbances, overgrazing pastures, and tolerance of environmental extremes (Schnelle 2019, Aththanayaka et al. 2023). Range expansion by woody native plants has been documented globally (Gettys and Schnelle 2018, Schnelle 2019, Negi et al. 2021, Ward et al. 2022).

Woody encroachment by native species is an important issue in grasslands globally resulting from overgrazing by cattle, anthropogenic interventions, and reduced fire frequency (Briggs et al. 2002, Eldridge et al. 2011, Ratajczak et al. 2012, Ward 2020). Encroachment of woody plants into grasslands reduces the species diversity and changes the quality and quantity of light reaching the understory (Ratajczak et al. 2012). In addition, woody encroachment modifies nutrient cycling, ecosystem productivity, soil chemistry, natural disturbance regimes, and local hydrology (Eldridge et al. 2011, Ratajczak et al. 2012, Donovan et al. 2018, Ward 2020). Climate change may also exacerbate the problem as the increase in global CO_2 levels may benefit C3 woody species over C4 grasses, allowing for further grassland transition to shrubland (Tunnell et al. 2004, Eldridge et al. 2011, Schnelle 2019).

Eastern redcedar is a widespread native tree in North America (Ward 2021). ERC is currently expanding its range out of its traditional habitat and into open areas like grasslands (Vasiliauskas and Aarssen 1992, Briggs et al. 2002, Ward 2021). ERC is commonly used as a windbreak surrounding agricultural fields and can aggressively colonize neighboring grasslands, rapidly transforming them into closed-canopy ERC forests (Briggs et al. 2002, Donovan et al. 2018). Although ERC's historical niche is limestone soils and cliffsides, it can survive in many environments and thrives in xeric environments where competition with other plants is reduced (Lawton and Cothran 2000, Ward 2020, Sangüesa-Barreda et al. 2021). In addition, as an animal-dispersed species, ERC seeds can be dispersed away from the mother tree, allowing

further movement into new areas (Holthuijzen and Sharik 1985, Horncastle et al. 2004). Like other successful invaders, ERC trees become sexually mature early in their lives; female ERC trees produce many seed-containing fleshy cones when they are ten-years-old and males develop pollen-producing cones when they are six-years-old, leading to rapid reproduction and dispersal (Van Haverbeake and Read 1976, Aronson et al. 2007, Wickert et al. 2017, K Shvach, personal communication). Moreover, increasing severe droughts brought on by climate change could give ERC an even more decisive competitive advantage over native grassland plant species (Kaskie et al. 2019).

One area affected by ERC invasion is Lakeside Daisy State Nature Preserve. Located in Ohio on the Marblehead Peninsula on the shore of Lake Erie, LDSNP is a unique preserve, including extensive wetlands and a small prairie traditionally dominated by short grasses and wildflowers (Figure 1). This preserve was likely historically maintained by its xeric environment and shallow soil, limiting competition from other plant species for the native flora. By observing historical satellite imagery, clumps of established trees can be observed in the 1990s, just shortly after the preserve's founding in 1988 (Ohio Department of Natural Resources). Possible reasons for the establishment of ERC trees at LDSNP could include urbanization of the peninsula and subsequent landscaping of privately-owned land using ERC or anthropogenic disturbance caused by limestone quarrying at the site. Additionally, ERC trees populate the sides of highways and hills surrounding LDSNP and could be dispersing into the preserve from these areas. ERC's affinity for xeric environments where competition with other plants is low and shade is minimal has allowed this population to rapidly colonize the prairie at LDSNP following its introduction (Lawton and Cothran 2000, Ward 2020, Sangüesa-Barreda et al. 2021).

Grasslands, like LDSNP, are considered among the most endangered ecosystems globally (Leis et al. 2017, Kaskie et al. 2019). Changes in plant community composition linked to woody encroachment can modify the belowground biomass of these ecosystems, reducing native grasses' soil-stabilizing effects (Watson et al. 2019). In semiarid environments like grasslands, soil stabilization is critical to prevent erosion and retain the little available water (Eldridge et al. 2011, Kaur et al. 2020, Knapp et al. 2020). Additionally, the fragmentation of grasslands resulting from anthropogenic influence and woody species encroachment can result in biodiversity losses (Leis et al. 2017). Up to 330 million hectares of grasslands were experiencing shrub encroachment in 2011, negatively affecting many of the economic and ecological resources that grasslands possess (Eldridge et al. 2011).

Two theoretical models describe how invasions typically occur: outward diffusion from the edge of the range and long-distance dispersal followed by local expansion (Auld and Coote 1980, Campbell and Dooley 1992, Gorchov et al. 2014). Diffusion of plant species is characterized by seed and pollen traveling short distances from the edge of the range and may result in local adaptation due to isolated reproduction in different areas (Auld and Coote 1980, Campbell and Dooley 1992, Keller et al. 2017). On the other hand, long-distance dispersal may have a homogenizing effect on genetic diversity because it allows for sharing of genetic material through pollen and seeds between geographically separate areas of the range (Campbell and Dooley 1992, Barriball et al. 2015). Additionally, long-distance dispersal may lead to an increased rate of spread as new adult trees in an area can facilitate the deposition of more seeds by providing perches for birds, the main dispersers of some woody encroachers (Holthuijzen and Sharik 1985, Higgins et al. 2000, Horncastle et al. 2004). If seeds are transported successfully over long distances, range expansion can occur even more rapidly for the species (Moody and

Mack 1988). It is, therefore, essential to identify patterns of a range expansion in order to model and predict them more accurately. Better prediction of spread rate and pattern can give managers the ability to target management efforts on new colonies before they expand further (Moody and Mack 1988, Higgins et al. 2000, Gorchov et al. 2014).

Using neutral genetic markers provides one way to infer the historical seed dispersal rates and patterns (Hamrick and Trapnell 2011). Dispersal patterns of seeds and pollen affect the genetic diversity and structure of plant populations, which is a driver of plant evolution and determines the population's response to selection pressures (Sork and Smouse 2006, Chybicki and Oleksa 2018). Therefore, studying the changes in the genetic structure over time can provide a more detailed understanding of how a newly established population of ERC persists over time (Céspedes et al. 2003, Roser et al. 2017). I suggest that higher genetic diversity allows the population to respond to environmental variation and may affect the success of ERC trees in different environmental conditions (Gonzales et al. 2009, Fuchs et al. 2013). I used genetic similarity among ERC trees over a temporal scale in the Lakeside Daisy State Nature Preserve to address the question: is ERC's pattern of range expansion driven by diffusion from the edge of the range or by long-distance dispersal events? Answering this question will help us estimate patterns of dispersal and colonization of ERC and inform a model of ERC's range expansion in grasslands and other open areas. Such models can help determine where to prioritize control measures and proactively manage expanding populations (Donovan et al. 2018, Kaskie et al. 2019).

I used genetic similarity among ERC trees over a temporal scale in the Lakeside Daisy State Nature Preserve to address the question: is ERC's pattern of range expansion driven by diffusion from the edge of the range or by long-distance dispersal events? Answering this

question will help us estimate patterns of dispersal and colonization of ERC and inform a model of ERC's range expansion in grasslands and other open areas. Such models can help determine where to prioritize control measures and proactively manage expanding populations (Donovan et al. 2018, Kaskie et al. 2019).

METHODS

Study site. The Lakeside Daisy State Nature Preserve is at the eastern end of the Marblehead Peninsula on Lake Erie in Ohio (41.53 °N, 82.73 °W) (Figure 1) (ESRI 2011). The 55-hectare preserve was established to protect the only naturally-occurring population of the federally threatened lakeside daisy (*Tetraneuris herbacea*) in the United States and includes an old limestone quarry of the Marblehead geological series (Ohio Department of Natural Resources). This study focuses on the 8.9-hectare parcel of prairie habitat with a high presence of ERC (Figure 1i), as much of the rest of the 55-hectare preserve (Figure 1ii) is made up of wetlands where ERC does not have a dominant presence. ERC presence in this parcel of prairie habitat threatens the native flora of LDSNP, including the federally threatened lakeside daisy. Mean annual precipitation ranges from 686 to 914 mm, and mean annual air temperature ranges from 7 to 11 ℃.

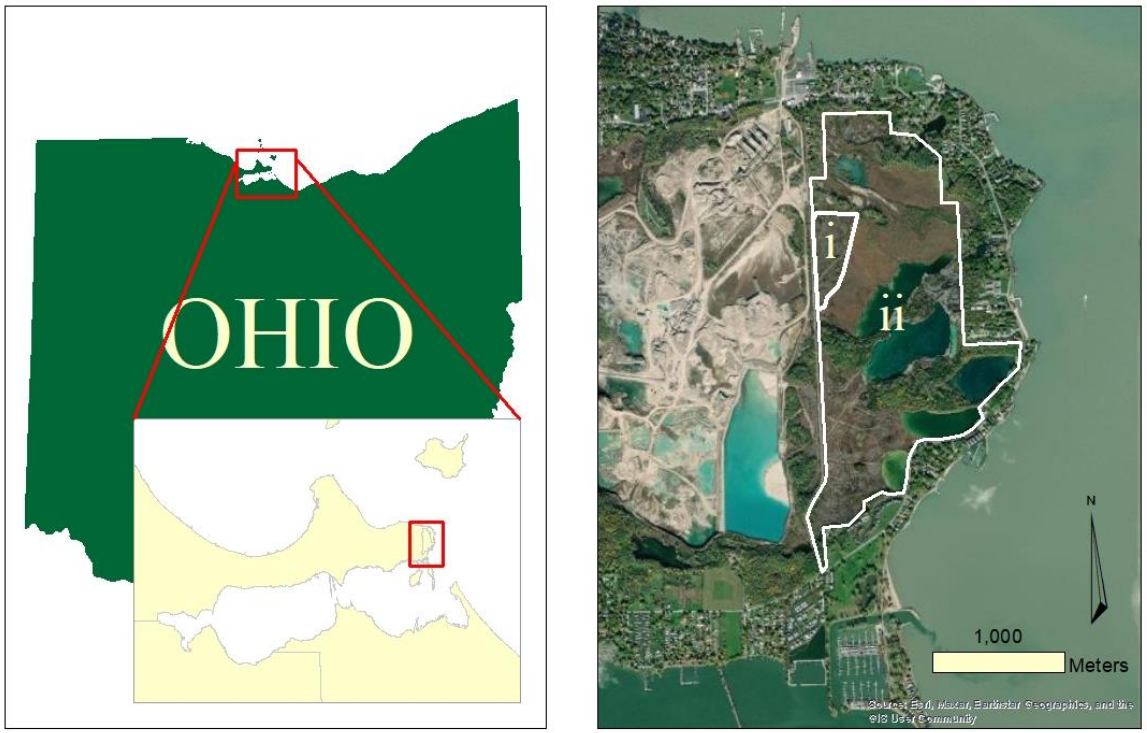

Figure 1. Location of Lakeside Daisy State Nature Preserve (LDSNP) on Marblehead Peninsula in northern Ohio. The sampling area (i), where the population of encroaching ERC is located, is approximately 8.9 hectares out of the entire (ii) 55-hectare preserve. Map created using ArcMap (ESRI 2011).

Sampling protocol. Leaf samples from 189 ERC trees from Lakeside Daisy State Nature Preserve were collected and stored in resealable plastic bags in a -20 °C freezer until DNA extraction. I sampled 170 ERC trees using a grid-like pattern from the north to the south end of the 8.9-hectare sampling area. Eleven sampling transects run from east to west every 50 m, and I collected leaf tissue from the closest tree every 10 m along each transect. The transect length varied as the transect stopped when it reached the wetland where ERC trees were absent. Overall, the sampled area had a perimeter of 1389 m, with transect two being the longest (24 trees sampled). Tree height, coordinates, presence or absence of female cones, and diameter at

breast height (DBH) for tall trees or base diameter for short trees were noted. The youngest tree found bearing female cones, indicating sexual maturity, was determined to be six years old.

Additionally, nine large trees found in a clump in one portion of the 8.9-hectare sampled area and ten large trees along the roadside on the edge of the sampling area were targeted as candidates for founding members of the population at LDSNP. These trees were sampled outside the grid sampling scheme, but they were sampled in the same way and included in the data analysis together with the rest of the trees sampled. LDSNP is located in the less-developed eastern end of the Marblehead Peninsula and is surrounded by a town with private residences (Figure 1). Large trees can be seen in the town of Marblehead, on other islands close to LDSNP, and along the highways in northeast Ohio, but these were located on privately-owned property and therefore, could not be sampled.

Wood core samples were collected using an increment borer from 45 trees where a DBH measurement was taken; meanwhile, stem cross sections were collected from 50 trees that had a base diameter taken. Cores and cross-sections were frozen in a -20 °C freezer until processed. The annual rings in each tree core or stem cross-section were counted under a dissecting microscope to determine the age of each tree. A core or cross section could not be obtained for all trees as some trees were missing the tags that were placed during the initial sampling and could not be reliably located using GPS coordinates. Linear relationships between stem diameter and age of the tree, based on growth ring count, was used to estimate the age of those trees for which a core was not collected (41 tall trees) or stem cross section (53 small trees). ERC growth rates are highly variable across sexes and microsite conditions, especially once trees are sexually mature (Quinn and Meiners 2004), so this estimation may have resulted in miscalculation of the age of trees and may have contributed to the disproportionate sample sizes of each age group.

The trees were grouped into five age groups of 10-year increments rather than on a continuous age scale to help account for any error resulting from age estimation.

DNA extraction and microsatellite analysis. Total genomic DNA was extracted from the leaf tissue of each ERC tree using a modification of the CTAB protocol described by Doyle and Doyle (1987). Subsequently, all DNA samples were diluted to approximately 10 ng/µl for polymerase chain reaction (PCR) amplification. PCR was performed using eight microsatellite markers, seven of these were developed in our lab for genetic analysis of ERC and one was developed for genetic analysis of *Juniperus communis* (Michalczyk et al. 2006) (Table 1). PCR was performed in a final volume of 10 µl, containing approximately 15 ng of genomic DNA, 10 mM tris buffer with KCl, pH 8.8, 1.88 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 µM of each primer, and 1 unit of *Taq* polymerase using a PTC-200 thermal cycler (MJ Research, Watertown, MA, USA). The PCR was conducted using a touchdown annealing approach to improve the specificity of primer binding to the target DNA. Annealing temperature decreased by 1 °C every three cycles from 60 to 56 °C during the first 15 cycles then annealing temperature remained at 55 °C for the remaining 30 cycles. The thermocycling profile consisted of initial denaturation at 94 °C for 2 min, followed by 15 cycles of 94 °C for 30 s, 60–56 °C for 30 s, 72 °C for 30 s, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s, a final extension of 72 °C for 3 min, and a holding temperature of 5 °C. All forward primers were labeled using the fluorescent dyes FAM or TET. The PCR product size was determined using an ABI 3730 DNA sequencer (Applied Biosystems, Waltham, MA) at MC Lab (San Francisco, CA). The program Peak Scanner was used to determine the multilocus genotype of all trees.

Double-digest restriction site associated DNA sequencing (ddRADseq) analysis. Eighty-four trees were chosen for double-digest restriction site associated DNA sequencing (ddRADseq)

analysis. These trees were chosen because they were able to be aged definitively. Used in this analysis were five individuals in the 0-9 age group, 41 individuals in the 10-19 age group, 27 in the 20-29 age group, and 11 individuals older than 30 years. These age groups will be referred to as LD1, LD2, LD3, and LD4, respectively. Undiluted total genomic DNA was sent to Admera Health Biopharma Services (South Plainfield, NJ) for library preparation and sequencing using a ddRADseq protocol. Once ddRADseq was performed, the samples were analyzed using open source software on the Ohio Supercomputer (1987) web interface. FastQC (version 0.11.5, Andrews 2010) was used to analyze the quality of reads from the samples sent for sequencing. This program was used to visualize important information about sample files including sequence quality, per sequence GC content, per base sequence and N content, sequence lengths, sequence duplication, and overrepresented sequences. After analyzing using FastQC, the program cutadapt (version 4.4, Martin 2011) was used to trim adapters sequences used in ddRADseq from the ends of the fragments.

The Stacks (version 2.62 Catchen et al. 2013) pipeline was used to create a SNP dataset from the trimmed files to be used for further analysis. This dataset was filtered using limits of minimum coverage at 7x and maximum coverage at 30x and missingness of 20% using VCFtools (version 0.1.16, Danecek et al. 2011). After filtering, per-site nucleotide divergency and nucleotide diversity in windows of 10,000 were calculated for each age group using VCFtools.

Temporal variation in genetic diversity. The program GenAlEx (version 6.4, Peakall and Smouse 2006) was used to determine the levels of genetic variation found within and among ERC trees from different age categories. This program was used to calculate common indicators of genetic diversity such as the number of alleles per locus (Na), effective number of alleles

(Ne), number of unique alleles in each age group, and observed and unbiased expected heterozygosity (Ho and He, respectively). Ne indicates the number of equally frequent alleles that would take to achieve the same He in each age group and unique alleles to an age group are present only in that age group. Deviations from Hardy–Weinberg equilibrium and linkage disequilibrium were determined using Genepop (version 4.7, Raymond and Rousset 1995, Rousset 2008).

I also used GenAlEx to estimate Wright's F coefficients and to conduct an Analysis of Molecular Variance (AMOVA) to assess the distribution of genetic diversity among ERC established each decade to determine levels of differentiation between age groups (Peakall and Smouse 2006). Because genetic diversity estimates are positively correlated with sample size, I obtained thirty random subsamples of ten individuals for age groups 10-19, 20-29, and 30-39 to compare to the 0-9 and 40 and greater (40+) age groups in order to account for the disparity in sample size among age groups in this study (Ward and Jasieniuk 2009). After random subsampling, I calculated Na, Ne, the number of unique alleles, Ho, He, and pairwise genetic differentiation index (Fst) among all age groups. In addition, I conducted one-way analyses of variance (ANOVA) using the thirty random samples taken from age groups 10-19, 20-29, and 30-39 to determine if there were significant differences in the means of Na, Ne, and Ho among groups. I use pairwise *post hoc* Tukey tests among age groups for these three diversity indicators. Additionally, pairwise Fst values were calculated from the SNP dataset using VCFtools. I also used GenAlEx to conduct Principal Coordinate Analysis (PCoA) to visualize the genetic relationships among trees of different age groups.

Genetic Structure. The genetic structure of ERC was analyzed with the use of Bayesian model-based clustering with the program STRUCTURE (version 2.3.4, Pritchard et al. 2000, Hubisz et

al. 2009). This analysis was conducted to determine if there are changes in the genetic structure among age groups. This program predicts the most likely number of subpopulation clusters for the populations sampled and recalculates F-statistics. All STRUCTURE runs used a burn-in length of 100,000 followed by 1,000,000 Markov chain Monte Carlo repetitions. A graphic representation of the results generated by STRUCTURE was presented using the program DISTRUCT (version 1.1, Rosenberg 2003). Options were selected to allow admixture, assume independence among loci and ignore population affiliations when defining clusters. To determine the likeliest number of subpopulation clusters (K), I followed the methodology proposed by Evanno et al. (2005). All probable K values were run 20 times to obtain delta K, which is an ad hoc measure based upon the second order rate of change of the likelihood function with respect to each K value (Evanno et al. 2005). According to this procedure, the modal value of the delta K can be used as an indicator of the number of ancestral population clusters in the area. The program STRUCTURE HARVESTER (version 6.0, Earl and von Holdt 2012) was used for calculating parameters of Evanno et al. (2005).

An ADMIXTURE plot was created using the SNP dataset to accompany the STRUCTURE plot created using microsatellite data. ADMIXTURE (version 1.3, Alexander et al. 2009) is a clustering software that is similar to STRUCTURE and produces a similar plot that estimates the ancestry of individuals using large SNP datasets. First, PLINK (version 1.90, Purcell 2020) was used to conduct linkage pruning (Purcell et al. 2007). The resulting files were used to identify the value of K that resulted in the lowest cross-validation (CV) error using ADMIXTURE. This value represents the most likely number of ancestral clusters that make up the population. The appropriate value of K was used to create the final ADMIXTURE plot using R (version 4.2.1 R Core Team 2022, Wickham 2016). Additionally, a principal component

analysis (PCA) plot based on the SNP dataset was created in R using output from PLINK. An additional PCA was created with the same method after the removal of three outliers from the previous PCA.

RESULTS

Levels of genetic diversity. The results of this study showed a total of 63 alleles among the 189 ERC trees at LDSNP examined using eight microsatellite loci. After random subsampling, average Na values ranged from 2.125 ± 0.227 in the 40+-year-old group to 3.250 ± 0.453 in the 0- to 9-year-old group (Figure 2, Table 2). Ne values showed a similar trend, with averages ranging from 1.671 ± 0.153 in the 40+-year-old group to 2.079 ± 0.221 in the 0- to 9-year-old group (Figure 2, Table 2). There is a significant overall effect of age group on Na, Ne, and Ho (MANOVA, Wilks Lambda = 0.1188, $F_{8,288} = 68.46$, $P < 0.0001$). ANOVA and Tukey honest significance difference (HSD) tests revealed significant differences in mean values of Na and Ne between different age groups of ERC (Na: $F_{4,145} = 197.515$; $P < 0.0001$) (Ne: $F_{4,145} = 46.538$; $P < 0.0001$) (Figure 2). My findings indicate that, considering all trees with no subsampling, the average number of unique alleles per tree increased from 0 in the oldest group to 0.44 in the youngest group (Table 3). When random subsampling was conducted in age groups with over 10 individuals, I found that the number of unique alleles in each age group increased from 0.033 unique alleles in the 40+-year-old group to 3.667 unique alleles in the 10- to 19-year old group (Table 2). Although there were no unique alleles in the 40+-year-old age group when considering all trees in the dataset, the random subsampling of the dataset revealed that there is there a low likelihood of finding a unique allele among the oldest group of trees (unique allele = 0.033 ± 0.033) (Table 2; Table 3). Ne values increased only slightly among all ages, but the difference

between the mean Ne values of the youngest group and the older groups was significant (Figure 2).

Our analysis also showed that two of the eight loci are not in Hardy-Weinberg Equilibrium in the sample as a whole, namely, JV4 and JV10 (Hardy-Weinberg exact test p-values 0.0014 ± 0.0014 and 0.0095 ± 0.0050, respectively). Additionally, there are indications of linkage disequilibrium among loci JV4 and JV7 (p-values 0.0256 ± 0.0064) and loci JV11 and JV8 (p-values 0.0460 ± 0.0127). The pairs of loci showing linkage disequilibrium were inconsistent among age groups.

Using ddRADseq, a total of 2,178,467 SNPs were identified in the dataset, which was reduced to 487,866 sites after filtering for coverage and missingness. The nucleotide divergency per site ranged from 0.138 in LD1 to 0.144 in LD3. The number of variants per window of 10,000 was 1.959 for LD1, 4.252 for LD2, 3.842 for LD3, and 2.804 for LD4.

Genetic Structure. The AMOVA revealed that only 1% molecular variance was among age groups. My analysis also showed that genetic differentiation among age groups increased slightly as the age difference between age groups increased (Table 5). I found that observed heterozygosity was highest in the youngest age group (Ho = 0.581 ± 0.119) and lowest in the oldest age group (Ho = 0.458 ± 0.107) (Table 2). My analysis using ANOVA and Tukey HSD *post hoc* test revealed significantly different mean Ho values for the 0-9, 10-19, 20-29, and 40+age groups, but the 30-39 age group was not significantly different from either the 20-29 or 40+age groups ($F_{4,145} = 60.117$; $P < 0.0001$). Overall, I saw an excess in heterozygosity and a low level of differentiation among age groups (Fst = 0.037 ± 0.009) (Table 2).

I used the Bayesian model-cluster program STRUCTURE to describe the genetic structure among age groups. This analysis showed that the genetic structure could be best described by four genetically distinct clusters (K = 4) (Figure 3A). Additionally, this analysis showed that all four clusters are present in all age groups and all clusters showed a similar contribution to the genetic diversity of all age groups (Figure 3B). A principal coordinate analysis shows two distinct groups of trees when the first two axes are plotted (Figure 4). It is worth noting that individuals from all age groups are represented in both groups in the principal coordinate analysis in similar proportions (Figure 4). Principal coordinates 1 and 2 explained 16.8% and 11.3% of the variation, respectively.

I also used the cluster program ADMIXTURE to create a more detailed ancestry plot using the SNP dataset. This analysis showed that the genetic structure of this population could best be described by three genetically distinct clusters (K = 3) (Figure 5A). This analysis showed that all three ancestral lineages were present in all age groups, but that they were present in different percentages of individuals over time (Figure 5B). A PCA showed that many individuals in this population are close in genetic distance, but three individuals are further away from the larger group (Figure 6A). Notably, the distant individuals are from different age groups from one another. Once the outliers were removed, it is clear that the LD1 and LD4 groups are mainly in close proximity to one another while the LD2 and LD3 groups have a wider range of genetic distances between individuals (Figure 6B). Principal components 1 and 2 explained 6.5% and 6.35% of the variation, respectively.

Table 1. Locus name, oligonucleotide primer sequences, repeat motifs, and PCR product size range each of the seven microsatellite loci for genetic analysis. Seven primer sequences were developed for *Juniperus virginiana* and one microsatellite locus developed for *Juniperus communis.*

Locus ID	FORWARD PRIMER (5'-3')	REVERSE PRIMER (5'-3')	Repeated Motif	Product size range (bp)
JV1	AATGTTCGATCCATTAAAGAGG	TTATAGCATTGGCTGCATTTAG	(ATA)5	380-430
JV2	AGTCTAATTTTGGGCATGATAG	GTTGGCTAAATCTTCCCTGTT	(TAT)5	270-310
JV3	CACTCAACCTAGTCAAGATCCA	TCAATTTGAAGAAACATGACTG	(TTA)5	420-470
JV4	TCACCATTTCTCGAGGTATTAG	ACTCCACCAATAGACTAGGAGC	(TAT)6	380-430
JV5	TCTTGTGCAACATACTTCCTTC	AAGAAGTTGAAAGCTAAGTGGG	(CTT)5	380-430
JV6	TTTCCAGGACTTGTTGTCATAG	CATGTTACACCTACCATTCCAC	(CTT)7	150-210
JV7	ACACCTCAGAAAATGGAATGAC	TTAGCCACTAGTTCCAATGATG	(ATT)7	120-170
JC1	TGTGTTTATTCTCCCCATCT	CCCCCAGTTATTCTAAACATT	(CA)20	121-147

Table 2. Sample size after subsampling (N), number of alleles (Na), number of effective alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), number of unique alleles, and fixation index (F) across all loci for each age group.

Age Group	N	Na	Ne	Ho	He	Unique Allele	F
0-9	9	3.250 (±0.453)	2.079 (±0.221)	0.581 (±0.119)	0.467 (±0.075)	3.467 (±0.202)	-0.241 (±0.129)
10-19	10	3.188 (±0.439)	1.833 (±0.214)	0.484 (±0.143)	0.405 (±0.062)	3.667 (±0.273)	-0.155 (±0.082)
20-29	10	3.104 (±0.475)	1.811 (±0.189)	0.472 (±0.098)	0.413 (±0.063)	3.467 (±0.243)	-0.119 (±0.111)
30-39	10	2.879 (±0.369)	1.772 (±0.151)	0.527 (±0.097)	0.401 (±0.057)	2.833 (±0.254)	-0.273 (±0.095)
40+	5	2.125 (±0.227)	1.671 (±0.153)	0.458 (±0.107)	0.361 (±0.066)	0.033 (±0.033)	-0.245 (±0.117)
All trees	189	4.075 (±0.406)	1.810 (±0.069)	0.493 (±0.040)	0.414 (±0.024)	0.7000 (±0.287)	-0.163 (±0.046)

Table 3. Number of unique alleles present in each locus and age group of ERC. Number of unique alleles and total number of unique alleles per age group were calculated considering all individuals in each age group.

Age Group	N	JV1	JV2	JV3	JV4	JV5	JV6	JV7	JC1	Total
0-9	9	1	0	0	0	0	0	2	1	4
10-19	95	0	1	0	4	1	1	6	2	13
20-29	48	0	1	1	0	0	1	1	0	4
30-39	32	0	1	0	0	2	0	2	0	5
40+	5	0	0	0	0	0	0	0	0	0

Table 4. List of unique allele size (bp) found in each age group of ERC. All individuals in each age group were considered for analysis.

Locus	0-9	10-19	20-29	30-39
JV1	403	0	0	0
JV2	0	270	294	285
JV3	0	0	403	0
JV4	0	367 373 376 409	0	0
JV5	0	377	0	374 380
JV6	0	141	156	0
JV7	110 167	113 116 143 158 161 182	134	149 173
JC1	119	123 129	0	0

Table 5. Pairwise Fst values demonstrating similarity between pairs of age groups of ERC.

Age group	0-9	10-19	20-29	30-39	40+
0-9	0.000				
10-19	0.035	0.000			
20-29	0.034	0.024	0.000		
30-39	0.042	0.024	0.020	0.000	
40+	0.037	0.026	0.024	0.024	0.000

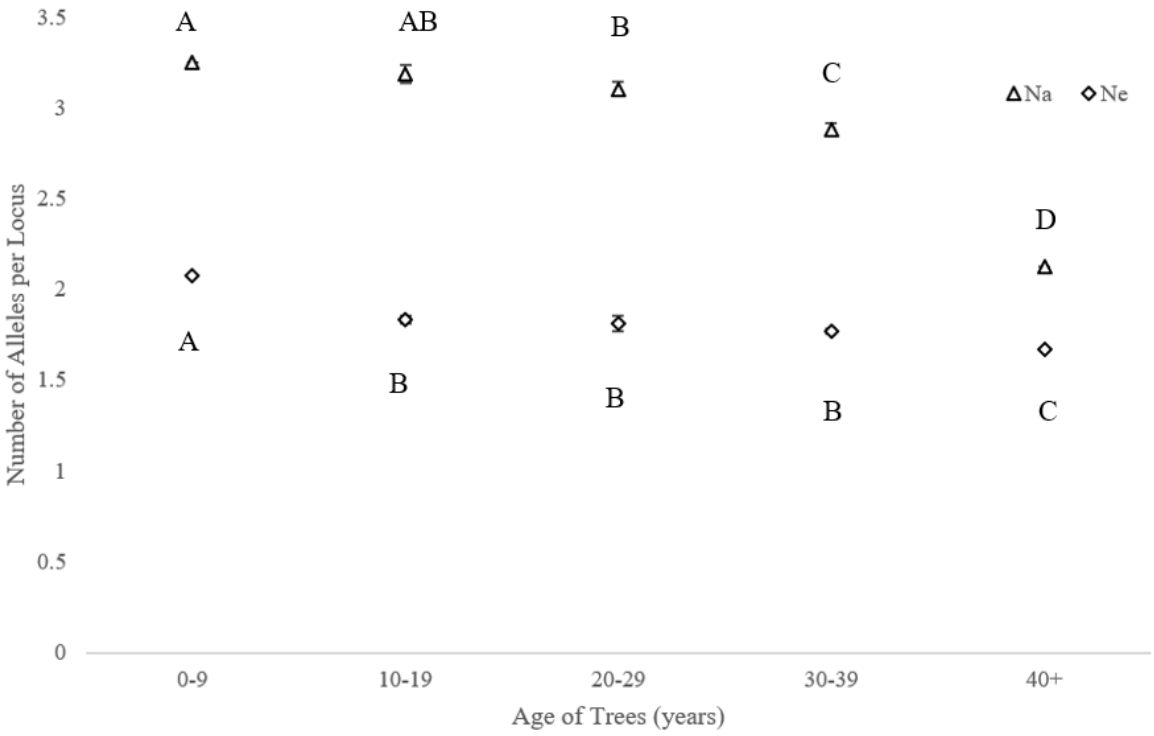

Figure 2. Average number of alleles (Na) and average number of effective alleles (Ne) for various age groups of ERC. Na and Ne values were calculated per locus from 30 random subsamples of age groups 10-19, 20-29, and 30-39. Values for Na and Ne for age groups 0-9 and 40+ were calculated using all individuals (N = 9 and N = 5, respectively). Sample size disparity among age groups may impact Na and Ne values. Error bars correspond to standard errors resulting from ANOVA. Letters indicate significantly different means determined using Tukey HSD *post hoc* pairwise comparisons.

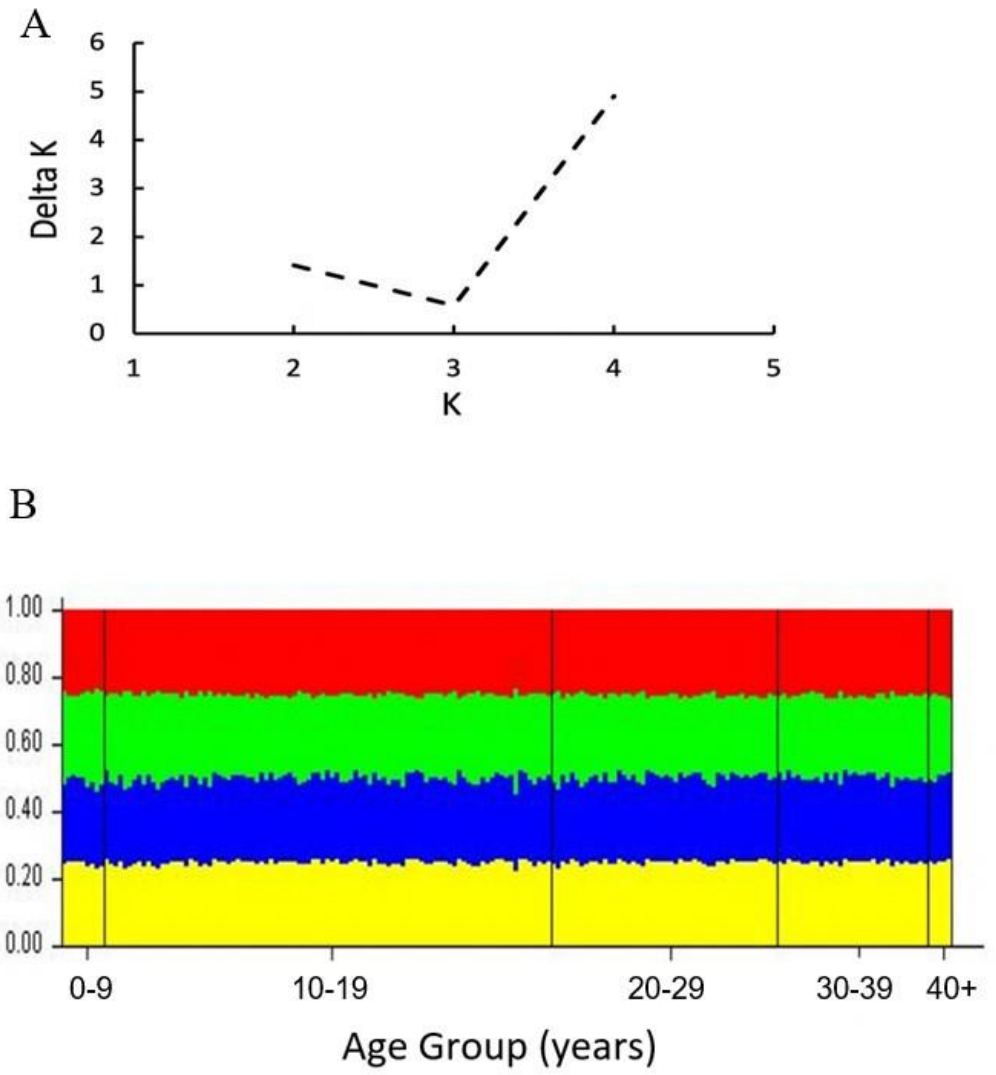

Figure 3. STRUCTURE plot of ERC age groups at LDSNP. STRUCTURE estimates of cryptic population structure of the ERC trees in the Lakeside Daisy State Nature Preserve. (A) Calculation of the second order rate of change (Delta K), determined by the modal peak. The modal peak for natural populations is at K=4. (B) STRUCTURE plot of ancestral subpopulations from the natural populations, with different colors representing the four population clusters and each line on the x-axis representing a single individual arranged in age groups, with the percent of its genome identified by the y-axis (Cluster 1: yellow, Cluster 2: blue, Cluster 3: green, Cluster 4: red).

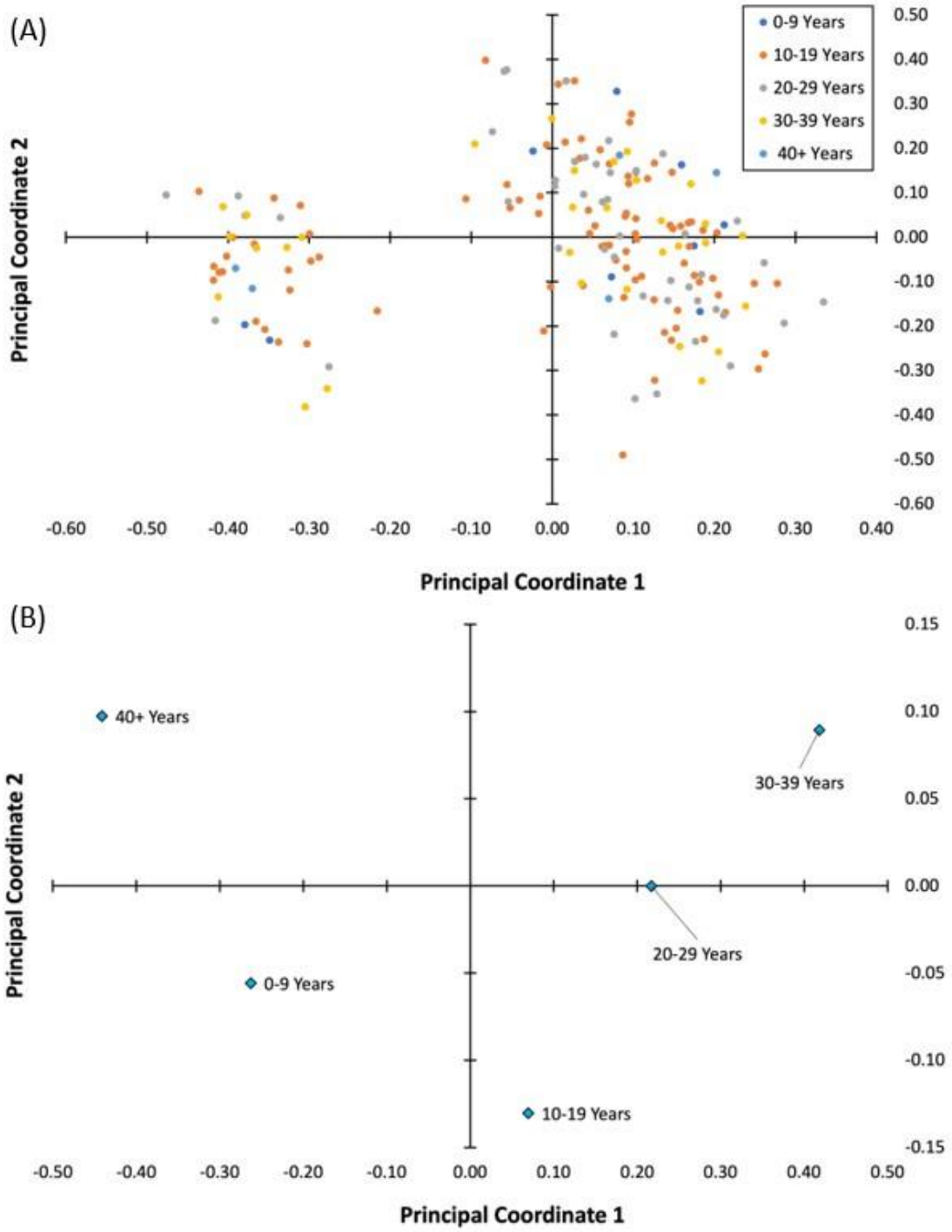

Figure 4. The Principal Coordinate Analysis (PCoA) of ERC age groups at LDSNP. (A) individual trees and (B) age groups at LDSNP plotted based on eight polymorphic microsatellite markers.

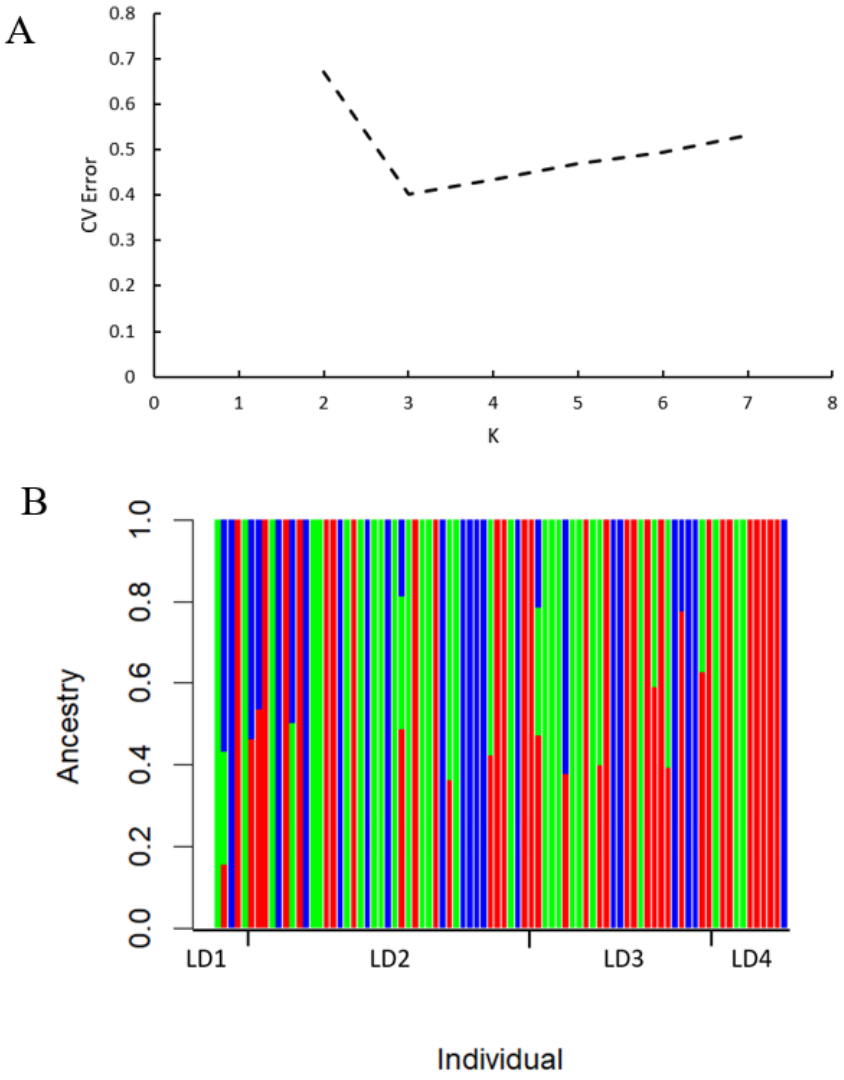

Figure 5. ADMIXTURE plot of ERC age groups at LDSNP. ADMIXTURE estimates the population ancestry of ERC trees at Lakeside Daisy State Nature Preserve based on SNP data. (A) Calculation of cross-validation error resulting from various K values with the lowest cv error being for K=3. (B) ADMIXTURE plot of ancestral subpopulations from the natural populations, with different colors representing the three population clusters identified using SNP data. Each line on the x-axis represents a single individual arranged in age groups, with the percent of its genome identified by the y-axis (Cluster 1: blue, Cluster 2: red, Cluster 3: green).

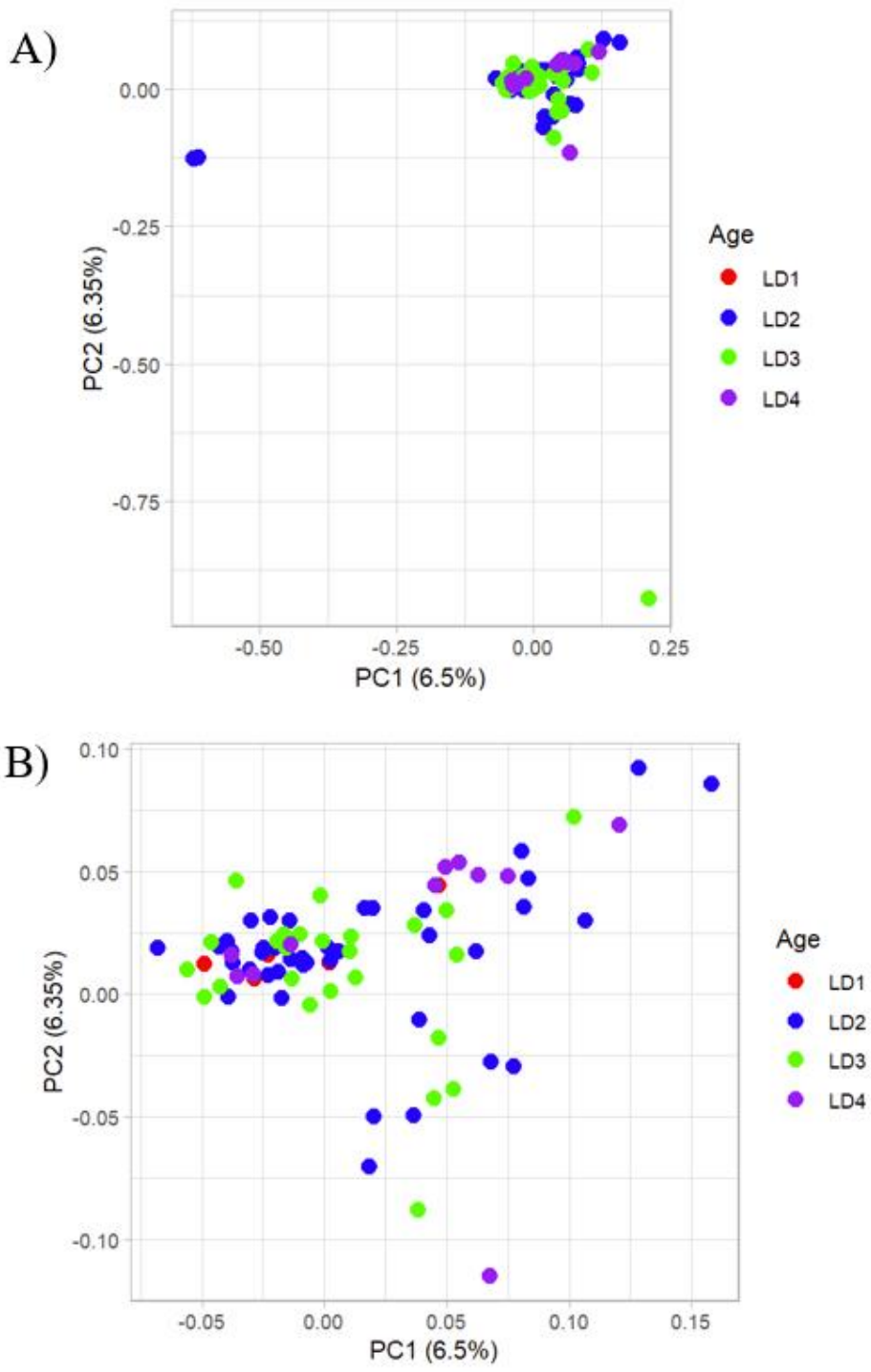

Figure 6. The Principal Component Analysis (PCA) of ERC age groups at LDSNP. (A) Individual trees separated into age groups and (B) individual trees separated into age groups with outliers removed. Created using SNP data.

DISCUSSION

Levels of genetic diversity. The results of this study showed that the genetic diversity of a population of ERC in Lakeside Daisy State Nature Preserve is increasing over time. Some genetic indicators such as Na increased greatly in younger age groups compared to older ones and others like Ne increased only slightly among all ages (Table 2). Although some trends were not as pronounced as others, Na, Ne, and Ho all showed a significant overall increase in younger age groups when measuring with microsatellite data. These findings suggest that, although there is a higher number of unique alleles found in younger populations, the new alleles are in low frequency within all populations, suggesting the arrival of new variants over time (Table 2).

While the microsatellite data shows compelling trends, when measuring genetic diversity using SNP data, no clear trend can be seen. The nucleotide divergency for each of the age groups did not show a clear trend of increasing or decreasing over time, but rather showed a trend of increasing nucleotide divergency with sample size. Additionally, the number of variants per window that were found also correlates closely with sample size for each age group but show no other trend. It is well known that sample size is correlated with increased genetic diversity so these values cannot be used to support or refute these conclusions until further analysis is conducted (Ward and Jasieniuk 2009).

Eight loci in this study were not in Hardy-Weinberg Equilibrium. Departure from Hardy Weinberg Equilibrium (HWE) may occur due to various causes, including purifying selection, inbreeding, population substructure, copy number variation, or genotyping error. For example, deviations from Hardy-Weinberg equilibrium are likely to occur in expanding populations and structured populations such as ERC in the Lakeside Daisy State Nature Preserve (Chen et al.

2017, Meisner and Albrechtsen 2019). Moreover, genotype frequency deviation in finite random-mating populations may also result from the difference between the gene frequencies of male and female gametes, which is determined by two independent causes: the gene frequency difference between male and female parents and the sampling error due to the finite number of offspring (Wang 1996). Given that this population was recently established and that seeds from nearby trees continue to arrive at LDSNP, it is reasonable to expect that the population will not be in equilibrium. Additionally, there were indications of linkage disequilibrium in the population of ERC at LDSNP. Disequilibrium in expanding populations can also result from the arrival of individuals from different provenances or regions with unusual genomic patterns, which can distort the inference of the genetic structure of the population (Liu et al. 2022).

The results of this study help to elucidate the historical patterns of encroaching of ERC in the LDSNP and provide insight into how dispersal and genetic admixture can contribute to its range expansion. These findings show that despite the size of the study area, there is considerable genetic diversity in the ERC population. Furthermore, these findings suggest that genetic diversity has continued to build over the last four to five decades. Because genetic diversity is essential for population survival and adaptation to changing environments, these findings suggest that increasing genetic diversity during ERC invasion of open areas must be a critical driver for its range expansion.

Ward et al. (2008) reviewed published studies examined that the changes in allelic diversity of alien species expanding in their new range. The invasion processes of non-native species imply an initial founder effect into a novel range; therefore, the new populations typically have lower diversity relative to populations in their native range (Ward et al. 2008, Zimmermann et al. 2010, Uller and Leimu 2011, Zhao et al. 2013). Moreover, environmental

filtering also causes diversity loss in alien species after arriving at their new location, further lowering genetic diversity (Dar et al. 2020, Dormontt et al. 2014, Vyšniauskienė et al. 2011). In contrast, native species that become invasive do not always experience loss of genetic diversity due to environmental filtering (Negi et al. 2021). Native invasive species, like ERC, have more sources of genetic diversity close to their invasive range than most non-native invasive species and are therefore more likely to experience intraspecific hybridization with other ecotypes, which allows the species to cope with stresses resulting from invading new habitat (Castillo et al. 2021, Glisson and Larkin 2021, Negi et al. 2021).

Some researchers question the role that environmental filtering alone plays in the success of a species in a given environment, citing competition as a limiting factor in a population's success (Cadotte & Tucker, 2017). Indeed, ERC is highly limited by competition, specifically by other trees through shading (Ward, 2020; Hamati et al. 2023). However, when not competing for light, ERC is highly successful in outcompeting other plants. Therefore, when considering the community composition of the prairie areas experiencing ERC encroachment, the differences in genetic diversity noted here could, in fact, be due to environmental filtering.

Changes in the genetic diversity of native species increasing their range by invading new environments have received little attention. These results show that changes in genetic diversity can occur within a short time, with the average number of alleles per locus increasing by 65% in 3 decades (Table 2, Figure 2). Excoffier et al. (2009) examined the genetic consequences of range expansion. They proposed that range expansions could promote the surfing of rare variants into newly occupied territories. These findings revealed the accumulation of unique low-frequency alleles in the LDSNP as time progressed (Table 2, Table 3, Table 4). As expected for native species experiencing range expansion, my findings indicate that the newly established

population was not isolated from other nearby populations, which explains the arrival of additional genetic diversity in the population. I observed nearby populations of ERC on the Marblehead Peninsula, with some under 1 km away, and on surrounding islands including Kelley's Island, approximately 8 km away.

Genetic Structure. The results of this study showed that Ho increased significantly in younger populations of ERC and that there was an excess of heterozygosity in the population of ERC at LDSNP. While this trend is interesting, it is possible that the smaller sample size of the oldest group contributed to the lower observed heterozygosity value (Ward and Jasieniuk 2009). On the other hand, individuals in the older age groups may have died off over time; therefore, there may be an underrepresentation of genetic diversity for that age group. An excess in heterozygosity in a population indicates little biparental inbreeding is occurring in a population. Accompanied by the fact that ERC is a dioecious species and does not engage in selfing as a monecious species would, it is likely that there is an influx of new lineages of ERC at LDSNP (Willi et al. 2020).

Using the STRUCTURE analysis, I found that the population at LDSNP was best described with four genetically distinct clusters, but also that they were all similarly represented in all age groups in the preserve (Figure 3). This finding suggests that all individuals originated from previously admixed populations of the four clusters. Despite the similar contribution of the four lineages to all age groups in the preserve, it appears that there are two distinct clusters of individuals that are closer in genetic distance to one another, identified by the principal coordinate analysis (Figure 4).

Genetic diversity in invasive plant populations accumulates through multiple introductions, gene flow, mutation, and hybridization among plants of different origins (Gaskin

and Schaal 2002, Espeland 2013, Jeschke and Heger 2018, Carr et al. 2019). However, the results of the STRUCTURE analysis show that all individuals at LDSNP contain the same admixture of four genetic clusters, suggesting that all individuals are derived from the same admixed source population. In addition, the high heterozygosity and the lack of inbreeding maintains the four genetically distinct clusters. As ERC is a dioecious wind-pollinated and animal-dispersed species it is possible and likely that there is continued gene flow between source populations.

When conducting similar analyses with a more extensive dataset of SNPs rather than eight microsatellite regions, different results were obtained that tell a similar story. The ADMIXTURE plot shows that the population at LDSNP is best described with three genetically distinct clusters and that these are all represented in each of the age groups in the preserve (Figure 5). However, this more detailed chart reveals that the older individuals in the population have a less admixed ancestry than the younger individuals. This supports the idea that all lineages have been present in the preserve since the establishment of ERC at LDSNP. However, the existence of only single-ancestry individuals in the oldest age group and admixture seen in younger individuals is likely a result of mating among trees in the preserve rather than mixing of clusters outside of the preserve. Individuals that are not admixed are present in some younger populations, namely LD3 and LD2, indicating that the continual input from the source population is likely.

The principal component analysis shows the majority of individuals at LDSNP are relatively close in genetic distance to one another (Figure 6A). Two individuals in the LD2 age group and one individual in the LD3 age group have greater genetic distance from the others. Interestingly, these individuals are not part of the oldest age group, which may suggest instances

of new colonization events in the preserve after the initial establishment of ERC. Upon closer inspection of the main group of individuals in the PCA, a nonrandom distribution of the LD1 and LD4 individuals while there is a wider range of variability in the LD2 and LD3 groups. It is possible that this is a result of the larger sample size in the LD2 and LD3 groups, but could also be indicative of increasing genetic variability in these groups.

Seed dispersal outside of a local area has been documented for numerous invasive species (Horncastle et al. 2004, Nathan et al. 2008, Martinod and Gorchov 2017). This type of seed dispersal is especially common among species with fruiting phenologies that mature their fruits in the late fall (Barriball et al. 2015, Bartowitz and Orrock 2016, McNeish and McEwan 2016). ERC trees initiate fruit maturation in September and bear mature fruits through the winter months in LDSNP (Shvach, personal communication). These fruits are available to be dispersed by a diverse array of birds and mammals over the winter when resources are scarce (Horncastle et al. 2004). Avian dispersers of ERC can disperse seeds short and intermediate distances from the source and are primarily resident or nomadic birds, traveling with no regular pattern following resource availability (Holthuijzen and Sharik 1985, Horncastle et al. 2004, Shvach, personal communication). Additionally, small mammals are likely to carry seeds short distances (<100 m) while medium-sized mammals may spread seeds longer distances (1 km or more) (Horncastle et al. 2004). ERC can be distributed by humans as well through use of the tree for landscaping and windbreaks and are often distributed very long distances from the seed source, resulting in long-distance dispersal (Donovan et al. 2018). Therefore, the behavior of ERC's seed dispersers could result in the colonization of LDSNP by ERC seeds from within LDSNP and from surrounding populations from varying distances away. This seed-dispersal pattern likely

explains ERC's rapid encroachment of LDSNP and the increase in allele diversity and heterozygosity in the younger age groups.

Once introduced to a new suitable habitat, ERC can quickly dominate through its strong competitive abilities, similar to many non-native invasive plants (Briggs et al. 2002, Donovan et al. 2018). ERC thrives in environments where competition with other plants for sunlight is low and can tolerate many moisture levels, especially in xeric environments (Ward 2020, Hamati 2022, Hamati et al. 2023). Additionally, birds are more likely to deposit seeds near perching sites, making previously colonized open areas more likely to be sites of future seed deposition than undisturbed open areas (Holthuijzen and Sharik 1985, Higgins et al. 2000, Horncastle et al. 2004). When long-distance dispersal occurs, ERC becomes established in a new area, and the acceleration of invasion quickly increases (Moody and Mack 1988). Therefore, it is imperative to prevent long-distance dispersal and establishment of ERC into undisturbed grasslands to maintain the integrity of the ecosystem and prevent further invasion.

CONCLUSIONS

Overall, I found that the invasion pattern of the native ERC has similarities and differences from that of non-native invasive plants. First, contrary to most non-native invasive plants, there has been an increase in allelic diversity in the ERC population at the Lakeside Daisy State Nature Preserve in just a few decades. This finding follows the predictions of Excoffier et al. (2009); the new alleles are rare (low frequency). Moreover, the results of the STRUCTURE analysis and trends in genetic diversity also suggest sustained gene flow from small clusters of trees neighboring LDSNP originating from previously admixed source populations now found along highways, farmland, and in yards in and around Marblehead. Because ERC is a dioecious

tree, outcrossing promotes admixture among clusters, increasing the range of environmental conditions where it can thrive. Hybridization among different ecotypes is often a condition that facilitates invasion of both non-native and native invaders (Vilà et al. 2000, Ellstrand 2009, Castillo et al. 2021).

Finally, while a clear advancing front could not be identified, the level of admixture observed in all ages in my analysis suggests that diffusion from newly established foci trees through seeds dispersed at short and intermediate distances drives ERC's range expansion. Additionally, anthropogenic distribution resulting from using ERC for landscaping can result in the long-distance dispersal of trees from different origins, which in turn facilitates hybridization and further invasion of grasslands and open areas. The new allelic variants seen in the younger individuals in my data set are present at low frequencies and are likely to result from introgression with external sources in this admixed population through pollen and seed dispersal from neighboring populations. Specifically, in the case of LDSNP, ERC has likely arrived through a combination of animal- and human-induced means from neighboring populations and has subsequently disseminated through the preserve's prairie from those introductions. At LDSNP, ERC is quickly dominating the dry, open areas of the preserve but is being outcompeted in wetter areas.

It is imperative that managers control the spread of ERC to new areas to prevent the establishment of these fast-growing populations. In order to prevent the establishment of new stands of ERC most effectively, managers should eliminate satellite populations before individual trees are able to reach sexual maturity. Additionally, humans have the responsibility to prevent long-distance dispersal of invasive plants through anthropogenic means such as landscaping. By preventing the long-distance dispersal of ERC, managers can limit the amount

of intraspecific hybridization occurring among different ecotypes with the goal of limiting genetic diversity and potential for evolution of beneficial traits in invasive plants.

CHAPTER III

MICROHABITAT DIFFERENCES AFFECT THE SMALL-SCALE SPATIAL DISTRIBUTION OF GENETIC DIVERSITY OF THE EASTERN REDCEDAR (*JUNIPERUS VIRGINIANA* L. VAR. *VIRGINIANA*)

A portion of this work is published in

Hartman, H. M. & Rocha, O. J. (2023). Temporal changes in genetic diversity reveal small-scale invasion dynamics of the eastern Redcedar (*Juniperus virginiana* L. var. *virginiana*) in the Lakeside Daisy State Nature Preserve in Ohio. *Invasive Plant Science and Management*, 1–32. https://doi.org/10.1017/inp.2023.23

Author Contributions: HMH and OJR designed the study, collected the data, performed data and statistical analyses, and wrote the manuscript.

INTRODUCTION

Woody encroachment is an important issue in grasslands worldwide, as encroaching native woody species like *Juniperus virginiana* (eastern redcedar) threaten native flora in grasslands (Briggs et al. 2002, Eldridge et al. 2011, Ratajczak et al. 2012, Ward 2020). Woody plants like the eastern redcedar (ERC) reduce species diversity in grasslands, change the quality and quantity of light reaching the understory, and affect ecosystem functions when encroaching into grasslands (Ratajczak et al. 2012). Additionally, the fragmentation of grasslands resulting

from anthropogenic influence and woody species encroachment can result in biodiversity losses (Leis et al. 2017). Up to 330 million hectares of grasslands were experiencing shrub encroachment in 2011, negatively affecting many of The Great Plains' economic and ecological resources (Eldridge et al. 2011).

Although ERC's historical niche is limestone soils and cliffsides, it can survive in many environments and thrives in xeric environments where competition with other plants is reduced (Lawton and Cothran 2000, Ward 2020, Sangüesa-Barreda et al. 2021, Hamati 2022, Hamati et al. 2023). In addition, as an animal-dispersed species, ERC seeds can be dispersed away from the mother tree, resulting in movement into new areas (Holthuijzen and Sharik 1985, Horncastle et al. 2004). Moreover, increasing severe droughts brought on by climate change could give ERC an even more decisive competitive advantage over native grassland plant species (Kaskie et al. 2019).

Many invasive plants experience a founder effect when introduced into a novel range, causing populations to have lower genetic diversity relative to native plants in their native range (Ward et al. 2008, Zimmermann, et al. 2010, Uller and Leimu 2011, Zhao et al. 2013). Additionally, environmental filtering causes diversity loss after populations arrive at their new location, often resulting in local adaptations in environmentally challenging sites (Vyšniauskienė et al. 2011, Dormontt et al. 2014, Dar et al. 2020). In contrast, pre-adapted characteristics, such as phenotypic plasticity, can lead to aggressive range expansion by colonizing non-native species without diversity loss (Ward et al. 2008). Changes in the genetic diversity of non-native plants invading new ranges has been widely studied, but genetic changes experienced by native species expanding their range by encroaching into new environments, especially those with challenging

environmental conditions, is an understudied topic (Ward 2008, Zhao et al. 2013, Yazlik & Ambarli, 2022).

There is ample evidence that closely adjacent populations show morphological and genetic differences across sharp boundaries (Antonovics and Bradshaw 1970, Antonovics 1972, Caisse and Antonovics 1978). Early evidence shows differences between tolerant and non-tolerant populations of *Anthoxanthum odoratum*, sweet vernal grass, growing over the boundaries of mine tip and pasture soils (Antonovics and Bradshaw 1970). They showed that metal tolerant plants differ both physiologically and morphologically from normal plants in several features other than tolerance. Moreover, other studies indicate the divergence among adjacent populations results from intense disruptive selection and ultimately lead to the establishment of reproductive isolation (McNeilly and Antonovics 1968, Caisse and Antonovics 1978, Antonovics 2006). More recent studies revealed similar effects of fine scale environmental variation on the genetic structure of herbaceous and woody plants (Triest et al. 2014, Mosca et al. 2018, Mizuki et al. 2010).

Using neutral genetic markers provides one way to infer the historical seed dispersal rates and patterns (Hamrick and Trapnell 2011). Additionally, using nucleotide polymorphisms (SNPs) provides a way to infer patterns of seed dispersal using neutral and non-neutral genetic markers. Dispersal patterns of seeds and pollen affect the genetic diversity of plant populations, which is a driver of plant evolution and determines the population's response to selection pressures (Sork and Smouse 2006, Chybicki and Oleksa 2018). Higher genetic diversity allows a population to better respond to environmental variation and may affect the success of ERC trees in different conditions (Gonzales et al. 2009, Fuchs et al. 2013). I used genetic similarity among ERC trees over a spatial scale in the Lakeside Daisy State Nature Preserve to elucidate how

environmental filtering may affect the distribution of genetic diversity in an encroaching population of ERC. Answering this question will help us understand the patterns of dispersal and colonization of ERC.

METHODS

Study site. The Lakeside Daisy State Nature Preserve is at the eastern end of the Marblehead Peninsula on Lake Erie in Ohio (41.53°N, 82.73°W). The 55-hectare preserve was established to protect one of the few naturally-occurring populations of the federally threatened lakeside daisy (*Tetraneuris herbacea*) and includes an old limestone quarry of the Marblehead geological series (Ohio Department of Natural Resources). A bedrock of limestone and dolomite underlies this region. Soils in the preserve are on reefs on the lake plains and are very shallow and excessively drained. In some parts of the formation, glacial till and lake sediments accumulate on top of the bedrock, forming poorly drained soils and extensive wetlands. In addition, the preserve includes patches of poorly drained soils on a smaller scale. This study focuses on the 8.9-hectare parcel of prairie habitat with a high presence of ERC (Figure 7i), as much of the rest of the 55-hectare preserve (Figure 7ii) is made up of wetlands where ERC does not have a dominant presence. ERC presence in this parcel of prairie habitat threatens the native flora of LDSNP, including the federally threatened lakeside daisy. Mean annual precipitation ranges from 686 to 914 mm, and mean annual air temperature ranges from 7 to 11°C.

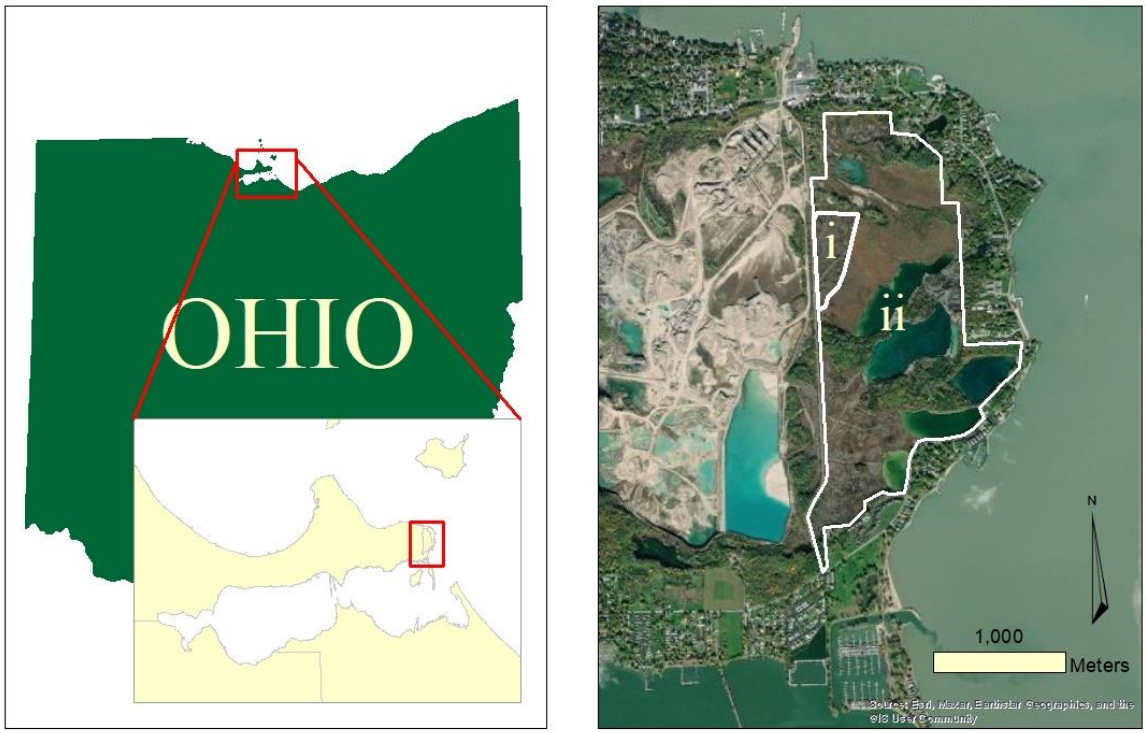

Figure 7. Location of Lakeside Daisy State Nature Preserve (LDSNP) on Marblehead Peninsula in northern Ohio. The sampling area (i), where the population of encroaching ERC is located, is approximately 8.9 hectares out of the entire (ii) 55-hectare preserve. Map created using ArcMap (ESRI 2011).

Sampling protocol. Leaf samples from 189 ERC trees from Lakeside Daisy State Nature Preserve were collected and stored in resealable plastic bags in a -20 °C freezer until DNA extraction. I sampled 170 ERC trees using a grid-like pattern from the north to the south end of the 8.9-hectare sampling area. Eleven sampling transects run from east to west every 50 m, and I collected leaf tissue from the closest tree every 10 m along each transect. The transect length varied as the transect stopped when it reached the wetland where ERC trees were absent. Overall, the sampled area had a perimeter of 1389 m, with transect two being the longest (24 trees sampled). Tree height, coordinates, presence or absence of female cones, and diameter at

breast height (DBH) for tall trees or base diameter for short trees were noted. The youngest tree found bearing female cones, indicating sexual maturity, was determined to be six-years-old.

Additionally, nine large trees found in a clump in one portion of the 8.9-hectare sampled area and ten large trees along the roadside on the edge of the sampling area were targeted as candidates for founding members of the population at LDSNP. These trees were sampled outside the grid sampling scheme, but they were sampled in the same way and included in this data analysis together with the rest of the trees sampled. LDSNP is located in the less-developed eastern end of the Marblehead Peninsula and is surrounded by a town with private residences (Figure 7). Large trees can be seen in the town of Marblehead, on other islands close to LDSNP, and along the highways in northeast Ohio, but these were located on privately-owned property and therefore, could not be sampled.

DNA extraction and microsatellite analysis. Total genomic DNA was extracted from the leaf tissue of each ERC tree using a modification of the CTAB protocol described by Doyle and Doyle (1987). Subsequently, all DNA samples were diluted to approximately 10 ng/µl for polymerase chain reaction (PCR) amplification. PCR was performed using eight microsatellite markers, seven of these were developed in our lab for genetic analysis of ERC and one was developed for genetic analysis of *Juniperus communis* (Michalczyk et al. 2006) (Table 6). PCR was performed in a final volume of 10 µl, containing approximately 15 ng of genomic DNA, 10 mM tris buffer with KCl, pH 8.8, 1.88 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 µM of each primer, and 1 unit of *Taq* polymerase using a PTC-200 thermal cycler (MJ Research, Watertown, MA, USA). The PCR was conducted using a touchdown annealing approach to improve the specificity of primer binding to the target DNA. Annealing temperature decreased by 1 °C every three cycles from 60 to 56 °C during the first 15 cycles then annealing temperature remained at 55 °C for the

remaining 30 cycles. The thermocycling profile consisted of initial denaturation at 94 °C for 2 min, followed by 15 cycles of 94 °C for 30 s, 60–56 °C for 30 s, 72 °C for 30 s, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 30 s, a final extension of 72 °C for 3 min, and a holding temperature of 5 °C. All forward primers were labeled using the fluorescent dyes FAM or TET. The PCR product size was determined using an ABI 3730 DNA sequencer (Applied Biosystems, Waltham, MA) at MC Lab (San Francisco, CA). The program Peak Scanner was used to determine the multilocus genotype of all trees.

Double-digest restriction site associated DNA sequencing (ddRADseq) analysis. Eighty-four trees were chosen for double-digest restriction site associated DNA sequencing (ddRADseq) analysis. These trees were chosen because they were able to be aged definitively using stem cross sections or cores. Undiluted total genomic DNA was sent to Admera Health Biopharma Services (South Plainfield, NJ) for library preparation and sequencing using a ddRADseq protocol. Once ddRADseq was performed, the samples were analyzed using open source software on the Ohio Supercomputer (1987) web interface. FastQC (version 0.11.5, Andrews 2010) was used to analyze the quality of reads from the samples sent for sequencing. This program was used to visualize important information about sample files including sequence quality, per sequence GC content, per base sequence and N content, sequence lengths, sequence duplication, and overrepresented sequences. After analyzing using FastQC, the program cutadapt (version 4.4, Martin 2011) was used to trim adapters sequences used in ddRADseq from the ends of the fragments.

The Stacks (version 2.62 Catchen et al. 2013) pipeline was used to create a SNP dataset from the trimmed files to be used for further analysis. For this pipeline, the samples were separated into groups based on approximate soil depth in their microhabitat. The soil was not

measured, but each tree was placed into a category of Shallow, Medium, or Deep soil based on visual observation. There were 26 individuals in the deep soil group, 36 individuals in the medium soil group, 22 individuals in the shallow soil group. The dataset was filtered using limits of minimum coverage at 7x and maximum coverage at 30x and missingness of 20% using VCFtools (version 0.1.16, Danecek et al. 2011). After filtering, per-site nucleotide divergency and nucleotide diversity in windows of 10,000 bp were calculated for each soil group using VCFtools.

Variation in genetic diversity in space. The program GenAlEx (version 6.4, Peakall and Smouse 2006) was used to determine the levels of genetic variation found within and among ERC trees from different age categories and locations within LDSNP. Based on the eight polymorphic microsatellite loci with a total of 63 alleles, I conducted a spatial genetic autocorrelation analysis using GenAlEx. GenAlEx calculates the genetic spatial correlation (*r*) over all age classes (*rc*), where the individual components of the numerator and denominator of r at a given distance class are summed across age groups (Smouse and Peakall 1999). This calculation is a better estimator than a simple arithmetic average across age classes because each pairwise comparison in each population contributes equally to the estimation of *rc*.

Additionally, I used GenAlEx to conduct Principal Coordinate Analysis (PCoA) to visualize the genetic relationships among trees of different age groups. The two groups of individual trees in the two groups of trees identified in the PCoA were plotted where they exist in space at LDSNP using their global coordinates in ArcMap to visually see where the trees in each group were established at LDSNP (version 10.8.2, ESRI 2011).

Genetic Structure. An ADMIXTURE plot was created using the SNP dataset to accompany the STRUCTURE plot created using microsatellite data. ADMIXTURE (version 1.3, Alexander et al. 2009) is a clustering software that is similar to STRUCTURE and produces a similar plot that estimates the ancestry of individuals using large SNP datasets. First, PLINK (version 1.90, Purcell 2020) was used to conduct linkage pruning (Purcell et al. 2007). The resulting files were used to identify the value of K that resulted in the lowest cross-validation (CV) error using ADMIXTURE. This value represents the most likely number of ancestral clusters that make up the population. The appropriate value of K was used to create the final ADMIXTURE plot using R (version 4.2.1 R Core Team 2022, Wickham 2016). Additionally, a principal component analysis (PCA) plot based on the SNP dataset was created in R using output from PLINK. An additional PCA was created with the same method after the removal of three outliers from the previous PCA.

RESULTS

The genetic spatial autocorrelation analysis revealed a consistent pattern of significant positive local genetic structure among closely-spaced trees (Figure 8). The correlogram shows the correlation for ERC trees as a function of the distance between trees. I found positive correlation values for trees in the first two distance classes (less than 110 m). For larger distance classes, the spatial autocorrelation is negative or close to zero, indicating little genetic structure. However, in two intermediate distance classes (classes six and seven, between 325 and 380 m) there is also positive genetic structure among trees.

A principal coordinate analysis (PCoA) created using microsatellite data shows two distinct groups of trees when the first two axes are plotted (Figure 9). Principal coordinates one and two explain 16.8% and 11.3% of the variation, respectively. When these two groups of trees

are mapped using their physical distribution on the preserve, a nonrandom distribution emerges (Figure 10). Group A identified with the PCoA exists at the north and south ends of the preserve while Group B has a more homogeneous distribution across the preserve (Figure 10).

Using ddRADseq, a total of 2,367,709 SNPs were identified across all individuals in the LDSNP population; the number of SNPs for this population was reduced to 487,977 sites after filtering for coverage and missingness. When analyzing using soil groups, the average nucleotide divergency per site was approximately 0.144 for all soil groups. However, the average number of variants per window of 10,000 bp was 3.624 for the shallow soil group, 4.152 for the medium soil group, and 3.789 for the deep soil group.

A principal component analysis (PCA) created using SNP data showed a nonrandom distribution of genetic distance between individuals (Figure 11A). When the outlier samples from the PCA are removed, the correlation between genetic variability and soil depth is shown more clearly (Figure 11B). Principal components one and two explain 6.5% and 6.35% of the variation, respectively. Additionally, an ADMIXTURE plot shows that all ancestral clusters are present in each soil group, but not in the same ratio across soil groups (Figure 12). The ADMIXTURE plot shows 9.09% of individuals in the shallow group, 25.0% in the medium group, and 15.4% in the deep have ancestry from multiple clusters.

Table 6. Locus name, oligonucleotide primer sequences, repeat motifs, and PCR product size range each of the seven microsatellite loci for genetic analysis. Seven primer sequences were developed for *Juniperus virginiana* and one microsatellite locus developed for *Juniperus communis.*

Locus ID	FORWARD PRIMER (5'-3')	REVERSE PRIMER (5'-3')	Repeated Motif	Product size range (bp)
JV1	AATGTTCGATCCATTAAAGAGG	TTATAGCATTGGCTGCATTTAG	(ATA)5	380-430
JV2	AGTCTAATTTTGGGCATGATAG	GTTGGCTAAATCTTCCCTGTT	(TAT)5	270-310
JV3	CACTCAACCTAGTCAAGATCCA	TCAATTTGAAGAAACATGACTG	(TTA)5	420-470
JV4	TCACCATTTCTCGAGGTATTAG	ACTCCACCAATAGACTAGGAGC	(TAT)6	380-430
JV5	TCTTGTGCAACATACTTCCTTC	AAGAAGTTGAAAGCTAAGTGGG	(CTT)5	380-430
JV6	TTTCCAGGACTTGTTGTCATAG	CATGTTACACCTACCATTCCAC	(CTT)7	150-210
JV7	ACACCTCAGAAAATGGAATGAC	TTAGCCACTAGTTCCAATGATG	(ATT)7	120-170
JC1	TGTGTTTATTCTCCCCATCT	CCCCCAGTTATTCTAAACATT	(CA)20	121-147

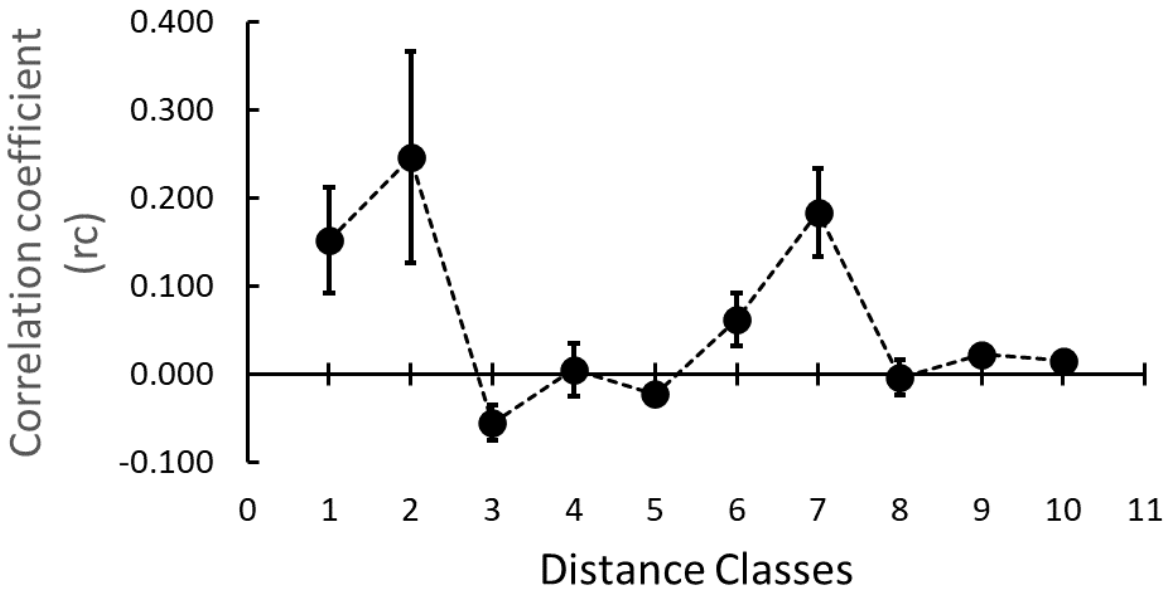

Figure 8. Genetic spatial autocorrelation of ERC at LDSNP. This spatial autocorrelation shows similarity of trees over distance classes at Lakeside Daisy State Nature Preserve based on genetic distance for each individual locus and all loci combined across all age groups. Error bars indicate standard errors.

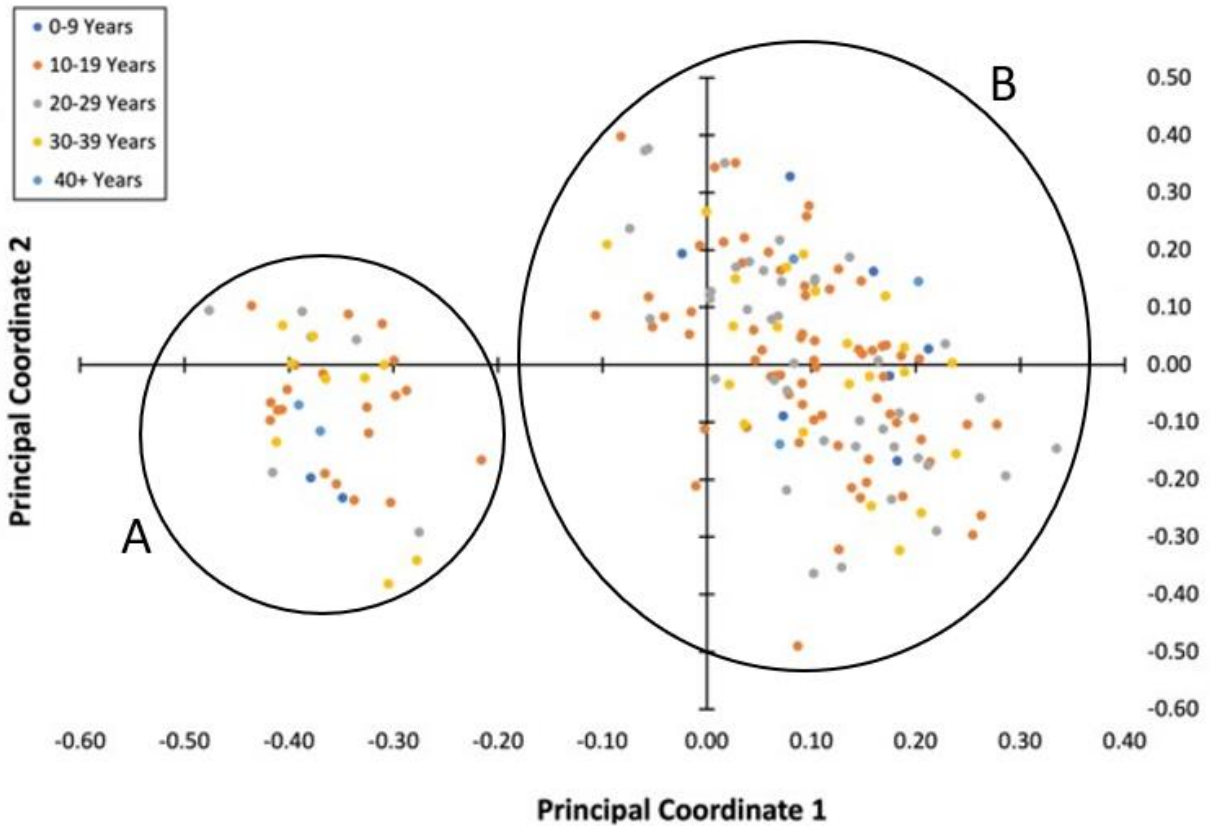

Figure 9. The Principal Coordinate Analysis (PCoA) of ERC at LDSNP with smaller groups identified. Individual trees at LDSNP plotted based on eight polymorphic microsatellite markers. Groups of trees used for further spatial analysis are marked on plot as A and B.

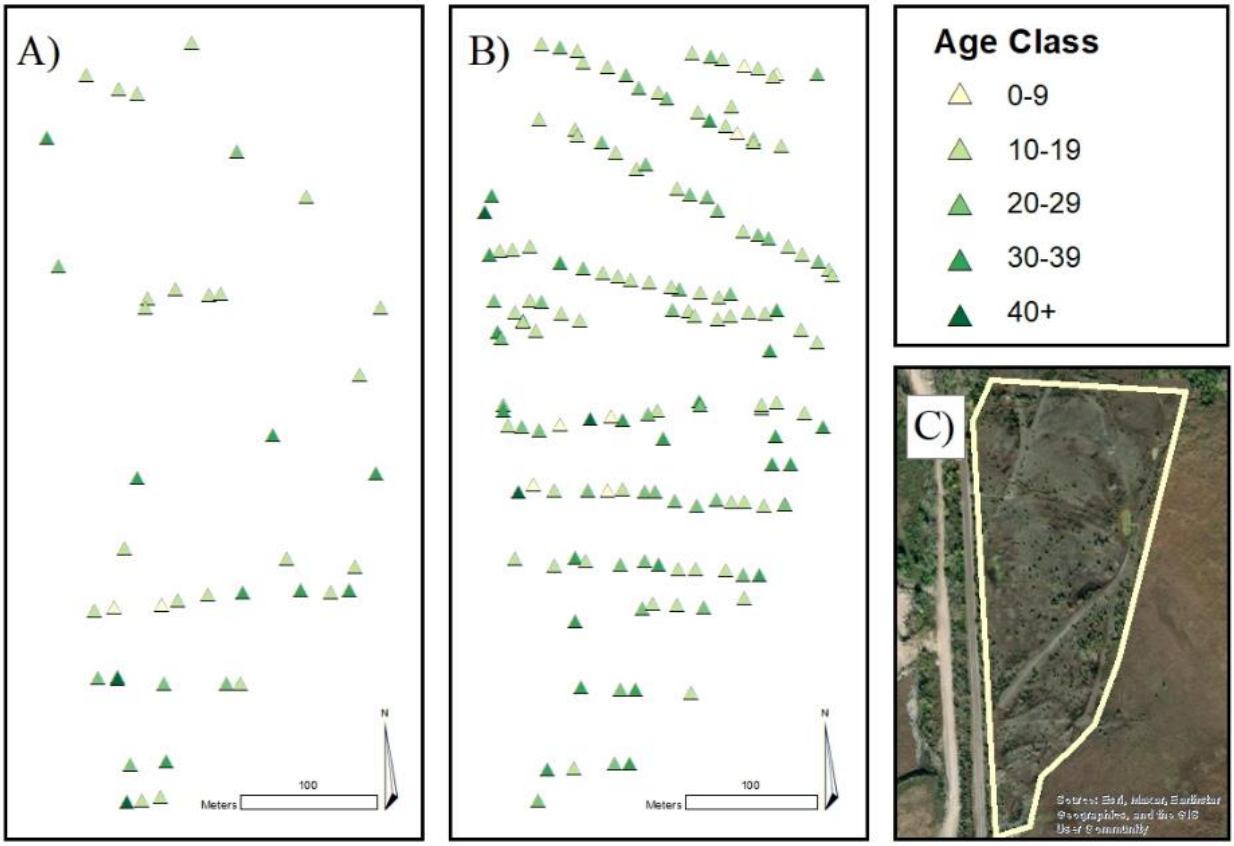

Figure 10. Spatial distribution of ERC sampled at LDSNP with ages. Trees sampled for this study in group A and B identified in the PCoA plotted in space. These trees are located within the 8.9-hectare parcel of land where sampling was conducted (C). Map created using ArcMap (ESRI, 2011).

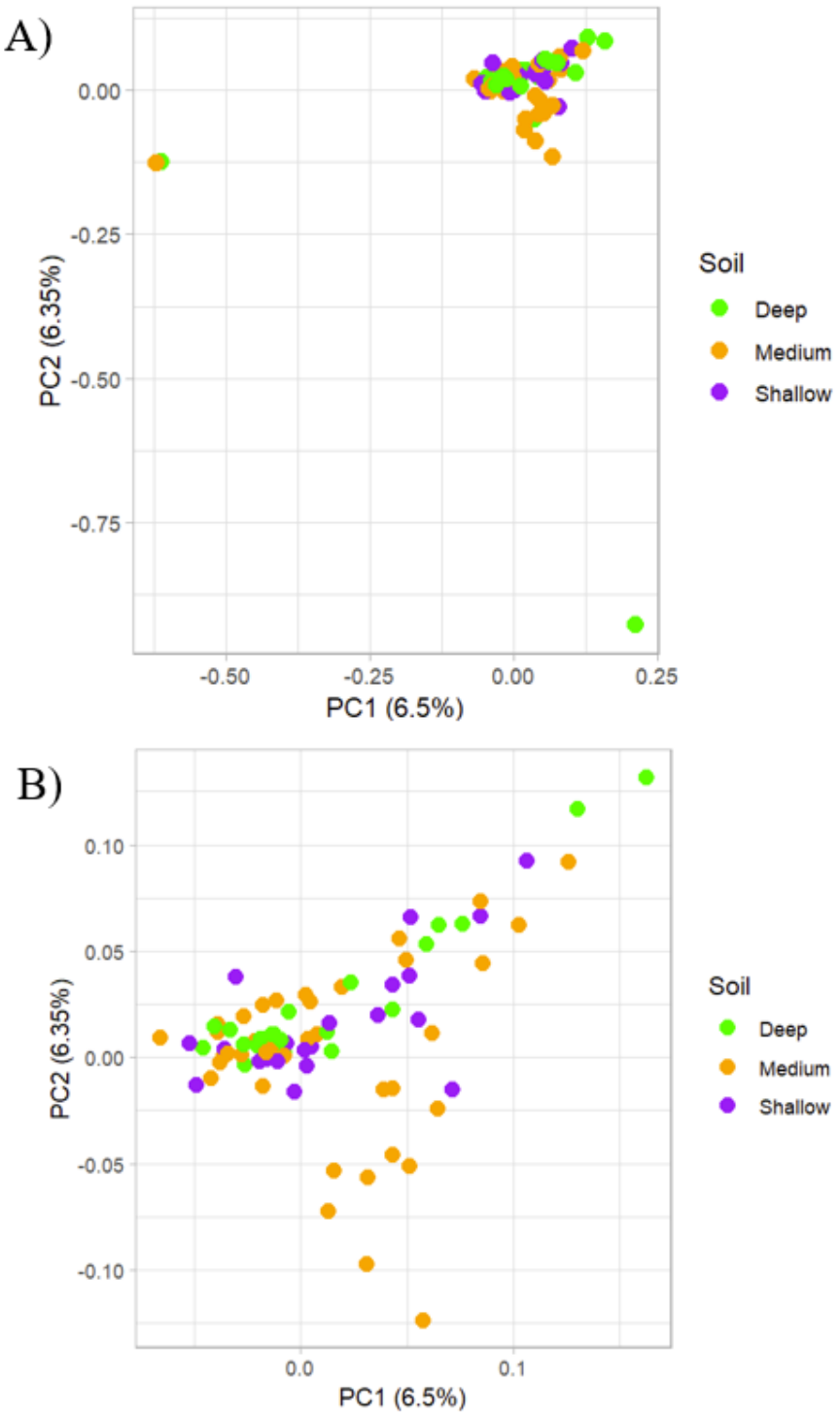

Figure 11. Principal Component Analysis (PCA) of ERC in different soil types at LDSNP. (A) Individual trees separated into soil groups and (B) individual trees separated into soil groups with outliers removed. Created using SNP dataset.

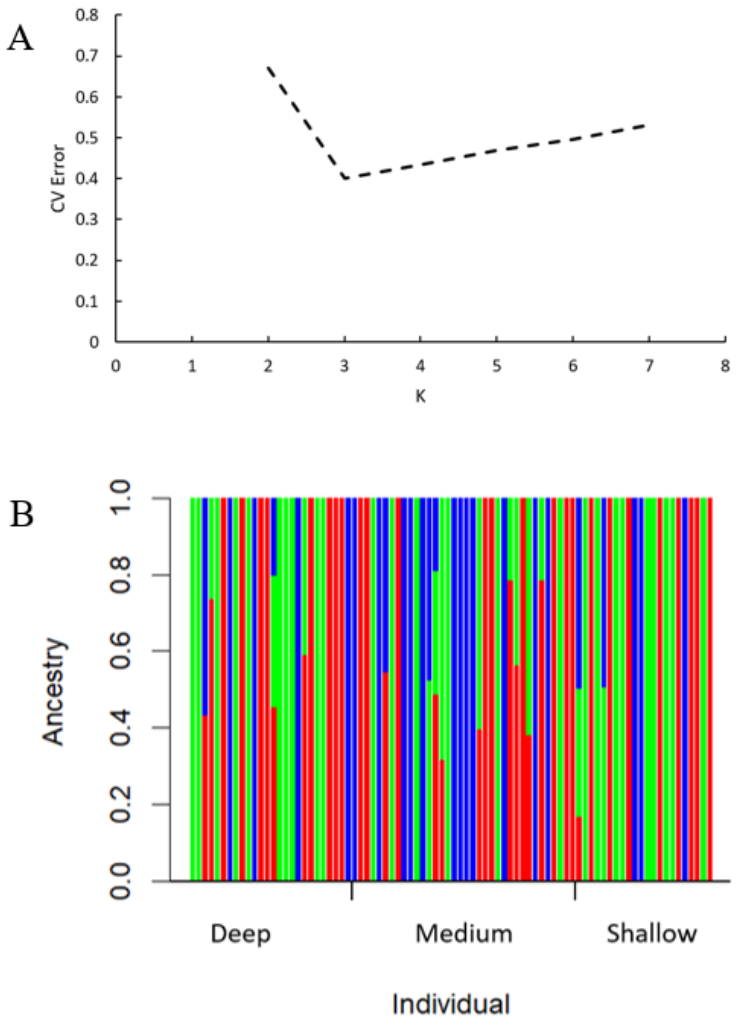

Figure 12. ADMIXTURE plot of ERC in different soil types at LDSNP. ADMIXTURE estimates the population ancestry of ERC trees at LDSNP based on SNP data. (A) Calculation of cross-validation error resulting from various K values with the lowest cv error being for K=3. (B) ADMIXTURE plot of ancestral subpopulations from the natural populations, with different colors representing the three population clusters identified using SNP data. Each line on the x-axis represents a single individual arranged into groups based on soil depth in the individual's microsite, with the percentage of its genome identified by the y-axis (Cluster 1: blue, Cluster 2: red, Cluster 3: green).

DISCUSSION

The results of this study revealed interesting trends in the distribution of genetic diversity of ERC across space and microhabitat conditions at LDSNP. I have found evidence that ERC trees that are closer in spatial distance are also closer in genetic distance (Figure 8). Additionally, my results show that there is a positive correlation in genetic similarity in trees that are at intermediate distances (325-380 m) away from one another (Figure 8). This result was not expected and led to further investigation into what could be causing this pattern.

Two distinct groups of individuals that are closer in genetic distance were identified using the principal coordinate analysis (Figure 9). While one group corresponds to trees widely distributed across the studied area of the preserve (Figure 10B), the second had a more limited distribution. This second group primarily occurs in the southern and northern parts of the prairie area (Figure 10A), corresponding to areas with deeper soils at both ends of the preserve. Meanwhile, trees in the larger group are most abundant in the center of the preserve (Figure 10B), where the soils on top of the limestone rock are thin. It is possible that the positive correlation of trees that are 325-380 m away from one another could be the result of the patchy distribution of the deep soils in the preserve, many of which are intermediate distances from one another (Figure 8; Figure 10C).

Other studies have shown environmentally challenging sites are conducive to the formation of distinct subpopulations (Espeland 2013, Gilbert et al. 2017, Tognetti et al. 2019). Local adaptations may result from the contrasting soil types present in LDSNP; most of the preserve includes very shallow and excessively drained soils on limestone bedrock creating xeric conditions that limit plant growth. However, sediments accumulate in some parts of the preserve forming deeper soils with high holding capacity and seasonal flooding. Intraspecific

hybridization among individuals capable of surviving under more stressful microhabitat conditions may contribute to the generation of successful novel genotypes as reported for other species (Gaskin and Schaal 2002, Espeland 2013, Pfennig et al. 2016).

When comparing trees based on their soil type, patterns emerged in genetic diversity for trees living in different soil microhabitats. While the patterns are not clear due to sample size discrepancies in soil groups, it was notable that the nucleotide divergency per site and the number of variants per window were highest for the medium soil group, followed by the shallow soil group and the deep soil group. Trees in areas of high environmental fluctuation appeared to have more genetic diversity which allows them to respond to the changing environment around them (LaForgia et al. 2020). The medium soil environment at LDSNP fluctuates between flooding and drought throughout the year, a microhabitat that could prove challenging for most plant life.

The results of the PCA show a nonrandom distribution of trees bearing different genotypes among soil types when plotted using genetic distance (Figure 11). The PCA shows that the shallow soil group and the deep soil group occupy a narrow area of the PCA while the medium soil group has a wider distribution on the PCA (Figure 11B). This could indicate that there is greater genetic diversity in the medium soil group when compared to the deep and shallow soil groups. All ancestral clusters are represented in each soil group in the ADMIXTURE plot, but they are not present in the same ratio for all soil groups (Figure 12). The shallow and deep soil groups show a larger proportion of individuals with ancestry stemming from a single cluster when compared to the medium soil group, where there is a greater ratio of individuals with ancestry from multiple clusters. The results of the ADMIXTURE plot further support that medium soil group has greater genetic diversity than the other two soil groups.

The results of this study help to elucidate the ways in which different environmental conditions may affect the genetic diversity of this population of ERC. Many studies have shown that environmental selection pressure differences between sites can lead to genetic differences in plants despite the existence of gene flow between populations (McNeilly and Antonovics 1968, Pannell & Fields 2013, Mosca et al 2018). Furthermore, these findings suggest that at LDSNP, genetic diversity in ERC is greater in areas with greater flux in environmental conditions and lower in areas with more specific environmental pressures like shallow soil and high shade. Similarly, Willi et al. (2020) found increased adaptation potential in populations with greater admixture. In this study, the medium soil group had the greatest admixture rate between ancestral clusters and existed in the section of the preserve with the most flux in environmental conditions, a microhabitat that may require more adaptation potential for a population to persist over time. Population genetic structure drives evolutionary dynamics in plant populations and in populations with high environmental heterogeneity, the differential selection among habitats can lead to substantial population differentiation (Pannel and Fields, 2013).

In summary, these results suggest that microhabitat conditions are filtering individuals according to their genotype. Individuals with a broader range of genetic diversity appear to establish and grow in medium soils that experience high flux in water retention. Meanwhile, trees in the deep and shallow soils appear to be part of distinct subpopulations based on their unique features, allowing them to establish in more xeric or mesic conditions. Specifically, trees in the shallow soils in the middle of the preserve experience drought most of the year, a condition that is not tolerated by most plant species. Therefore, this dry part of the preserve results in little competition from other plants. ERC thrives in dry areas with little shade and

competition, so this may not be as environmentally challenging for ERC as it would be for other plants (Ward, 2020).

This pattern of the distribution of genetic diversity could be why the genetic diversity measures of the shallow soil group of plants appear to be lower than the other two soil types. On the other hand, the deeper soils at LDSNP are located at either end of the preserve and have much more competition from other plants, specifically tall trees that cast shade on the plants around them. These areas do not typically flood, but they would be challenging for ERC due to its affinity for high sun exposure (Ward, 2020). The medium soil areas provide mainly open access to high sun environments, with the main competitor being grasses and other small plants, but flood heavily throughout the year. This microhabitat condition could be where ERC seedlings from many different ancestries may get established because of the deeper soils and high sun; however, seedlings may not persist here over time due to extensive periods of flooding.

CONCLUSIONS

In this study, ERC individuals at LDSNP are genetically more similar to one another when established in the same microhabitat compared to individuals in other microhabitats. My findings suggest that trees bearing higher genetic diversity in environmentally challenging sites allow the ERC population to persist in these sites over time. Increased genetic diversity due to hybridization among trees from different origins (ecotypes) could allow ERC to continue invading prairie areas outside of its native range. When competing with only grasses and other small plants, ERC has persisted over three decades at LDSNP in various soil moisture conditions.

CHAPTER IV

CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions. I found an increase in allelic diversity after just four decades of establishment at Lakeside Daisy State Nature Preserve (LDSNP) rather than the genetic bottleneck that would be expected for non-native plants (Dormontt et al. 2014). This increase in genetic diversity is likely due to the abundant sources of ERC seed in Ohio that can be transported various distances by animals and humans. The admixture of many sources can increase the likelihood of hybridization between ecotypes, which may increase the range of environmental conditions where ERC can persist over time by increasing the diversity of alleles in a population (Vilà et al. 2000, Ellstrand 2009, Castillo et al. 2021). Additionally, I found evidence that ERC in similar microclimates may be more genetically similar to one another when compared to those residing in other microclimates and that microclimates that fluctuate more may house a greater diversity of ERC.

Anthropogenic movement of ERC seeds for landscaping is a large contributor to the deposition of different ecotypes to new areas and should be avoided by land managers (Schnelle 2019). My data indicate that encroachment results primarily from local seed production and intermediate-distance seed dispersal by animals with occasional arrival of propagules via long-distance dispersal facilitated by humans. It is imperative that the spread of ERC to new areas is prevented before these fast-growing populations can be established. Therefore, adequate

management requires consistent monitoring of grasslands to detect and eradicate newly established foci before initiating seed production with a particular focus on grasslands near other populations of ERC (Horncastle et al. 2004). Since the earliest recorded cone-producing female tree in this study was six years old, managers could survey grasslands and eradicate ERC every five years to prevent the establishment of new ERC stands most efficiently.

Future Directions. ERC is not a model organism and therefore, using a genomic study to its full potential was difficult. The use of open-source software like STACKS has made the study of ERC's genome possible in the capacity included in this thesis, but the establishment of a recognized genome would aid in the further study of the ERC genome. Additionally, there are a multitude of genomic studies that can be conducted to further explore the ERC genome including comparisons between loci and environmental conditions. Further study into the effect of environmental conditions on ERC may yield interesting information about adaptations of ERC in different microhabitats.

REFERENCES

Alexander, D. H, Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19, 1655–1664.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Antonovics, J. (1972). Population dynamics of the grass *Anthoxanthum odoratum* on a zinc mine. *Journal of Ecology*, 60, 351-365. https://doi.org/10.2307/2258352

Antonovics, J. (2006). Evolution in closely adjacent plant populations x: Long-term persistence of prereproductive isolation at a mine boundary. *Heredity*, 97, 33–37. https://doi.org/10.1038/sj.hdy.6800835

Antonovics, J., & Bradshaw, A. D. (1970). Evolution in closely adjacent plant populations viii. Clinal patterns at a mine boundary. *Heredity*, 25, 349–362. https://doi.org/10.1038/hdy.1970.36

Aronson, M. F. J., Handel, S. N., & Clemants, S. E. (2007). Fruit type, life form and origin determine the success of woody plant invaders in an urban landscape. *Biological Invasions*, 9, 465-475.

Aththanayaka, C. P., Siyasinghe, D. P., Prakash, S. L., Bloch, C. P., & Surasinghe, T. D. (2023). Native and exotic plant invasions vary across habitat types and anthropogenic disturbances in a tourism-heavy protected area. *Biological Invasions*, 25, 411-429.

Auld, B. A. & Coote, B. G. (1980). A model of a spreading plant population. *Oikos*, 34, 287-292.

Barriball, K., McNutt, E. J., Gorchov, D. L., & Rocha, O. J. (2015). Inferring invasion patterns of *Lonicera maackii* (Rupr) Herder (Caprifoliaceae) from the genetic structure of 41 naturalized populations in a recently invaded area. *Biological Invasions*, 17, 2387–2402.

Bartowitz, K. J. & Orrock, J. L. (2016). Invasive exotic shrub (*Rhamnus cathartica*) alters the timing and magnitude of post-dispersal seed predation of native and exotic species. *Journal of Vegetation Science*, 27, 789-799.

Blackburn, T. M., Bellard, C., & Ricciardi, A. (2019). Alien versus native species as drivers of recent extinctions. *Frontiers in Ecology and the Environment*, 17, 203-207.

Briggs, J. M., Hoch, G. A., & Johnson, L. C. (2002). Assessing the rate, mechanisms, and consequences of the conversion of tallgrass prairie to *Juniperus virginiana* forest. *Ecosystems*, 5, 578–586.

Cadotte, M. W., & Tucker, C. M. (2017). Should environmental filtering be abandoned? Trends in Ecology and Evolution, 32, 429–437. http://dx.doi.org/10.1016/j.tree.2017.03.004.

Caisse, M., & Antonovics, J. (1978). Evolution in closely adjacent plant populations. *Heredity*, 40, 371–384. https://doi.org/10.1038/hdy.1978.44

Campbell, D. R. & Dooley, J. L. (1992). The spatial scale of genetic differentiation in a hummingbird-pollinated plant: comparison with models of isolation by distance. *American Naturalist*, 139, 735-748.

Carr, A. N., Hooper, D. U., & Dukes, J. S. (2019). Long-term propagule pressure overwhelms initial community determination of invader success. *Ecosphere*, 10, e02826.

Castillo, M. L., Schaffner, U., van Wilgen, B. W., Manuel Montaño, N., Bustamante, R. O., Cosacov, A., Mathese, M. J., & Le Roux, J. J. (2021). Genetic insights into the globally invasive and taxonomically problematic tree genus *Prosopis*. *AoB Plants*, 13, 1-13.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22, 3124-3140.

Céspedes, M., Gutierrez, M. V., Holbrook, N. M., & Rocha, O. J. (2003). Restoration of genetic diversity in the dry forest tree *Swietenia macrophylla* (Meliaceae) after pasture abandonment in Costa Rica. *Molecular Ecology*, 12, 3201–3212.

Chen, B., Cole, J. W., & Grond-Ginsbach, C. (2017). Departure from Hardy Weinberg Equilibrium and Genotyping Error. *Frontiers in Genetics*, 8, 1-6.

Chybicki, I. J. & Oleksa, A. (2018). Seed and pollen gene dispersal in *Taxus baccata*, a dioecious conifer in the face of strong population fragmentation. *Annals of Botany*, 122, 409–421.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dar, T., Bhat, B. A., Khuroo, A. A., Verna, S., & Islam, S. U. (2020). Genetic diversity and population structure of an invasive plant species differ in two non-native regions with differing climate invasion success. *Nordic Journal of Botany*, 38, 1-9.

Donovan, V. M., Burnett, J. L., Bielski, C. H., Birgé, H. E., Bevans, R., Twidwell, D., & Allen, C. R. (2018). Social–ecological landscape patterns predict woody encroachment from native tree plantings in a temperate grassland. *Ecology and Evolution*, 8, 9624–9632.

Dormontt, E. E., Gardner, M. G., Breed, M. F., Rodger, J. G., Prentis, P. J., & Lowe, A. J. (2014). Genetic bottlenecks in time and space: Reconstructing invasions from contemporary and historical collections. *PLoS One*, 9, 1-11.

Doyle, J. J. & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19, 11-15.

Earl, D. A. & von Holdt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources*, 4, 359-361

Eldridge, D. J., Bowker, M. A., Maestre, F. T., Roger, E., Reynolds, J. F., & Whitford, W. G. (2011). Impacts of shrub encroachment on ecosystem structure and functioning: Towards a global synthesis. *Ecology Letters*, 14, 709-722

Ellstrand, N. C. (2009). Evolution of invasiveness in plants following hybridization. *Biological Invasions*, 11, 1089-1091.

Espeland, E. K. (2013). Predicting the dynamics of local adaptation in invasive species. *Journal of Arid Land*, 5, 268-274

ESRI. (2011). ArcGIS Desktop: version 10.8.2. Redlands CA: Environmental Systems Research Institute.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611-2620.

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, 40, 481-501.

Fuchs, E. J., Robles, T., & Hamrick, J. L. (2013). Spatial distribution of *Guaiacum sanctum* (Zygophyllaceae) seedlings and saplings relative to canopy cover in Palo Verde National Park, Costa Rica. *Revista De Biología Tropical*, 61, 1521-1533.

Gaskin, J. F. & Schaal, B. A. (2002). Hybrid *Tamarix* widespread in U.S. invasion and undetected in native Asian range. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 11256-11259.

Gettys, L. A. & Schnelle, M. A. (2018). The Natives Are Restless: Proceedings from the ASHS Invasive Plants Research Interest Group 2017 and 2018 Workshops. *HortTechnology*, 29, 1.

Glisson, W. J. & Larkin, D. J. (2021). Hybrid watermilfoil (*Myriophyllum spicatum* X *Myriophyllum sibiricum*) exhibits traits associated with greater invasiveness than its introduced and native parental taxa. *Biological Invasions*, 23, 2417-2433

Gonzales, E., Hamrick, J. L., Smouse, P. E., Trapnell, D. W., & Peakall, R. (2009). The impact of landscape disturbance on spatial genetic structure in the Guanacaste tree, *Enterolobium cyclocarpum* (Fabaceae). *Journal of Heredity*, 101, 133–143.

Gorchov, D. L., Castellano, S. M., & Noe, D. A. (2014). Long-distance dispersal and diffusion in the invasion of *Lonicera maackii*. *Invasive Plant Science and Management*, 7, 464–472

Hamati, S. (2022). Ecophysiology of *Juniperus virginiana* encroachment in Ohio. Ph.D dissertation. Kent, OH: Kent State University. 244 p.

Hamati, S., Medeiros, J. S., & Ward, D. (2023). Effects of post oak (*Quercus stellata*) and smooth brome (*Bromus inermis*) competition on water uptake and root partitioning of eastern redcedar (*Juniperus virginiana*). *PLoS One*, 18, e0280100. https://doi.org/10.1371/journal.pone.0280100.

Hamrick, J. L. & Trapnell, D. W. (2011). Using population genetic analyses to understand seed dispersal patterns. *Acta Ecologica Sinica*, 37, 641–649.

Higgins, S. I., Richardson, D. M., & Cowling, R. M. (2000). Using a dynamic landscape model for planning the management of alien plant invasions. *Ecological Applications*, 10, 1833-1848.

Holthuijzen, A. M. & Sharik, T. L. (1985). The avian seed dispersal system of eastern redcedar (*Juniperus virginiana*). *Canadian Journal of Botany*, 63, 1508–1515.

Horncastle, V. J., Hellgren, E. C., Mayer, P. M., Engle, D. M., & Leslie, D. M. (2004). Differential consumption of eastern redcedar (*Juniperus virginiana*) by avian and mammalian guilds: Implications for tree invasion. *American Midland Naturalist*, 152, 255-267.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9, 1322-1332.

Jeschke, J. M. & Heger, T., eds (2018). Invasion Biology: Hypotheses and Evidence. Wallingford, Oxfordshire, UK: CAB International. Pp 147-153.

Kaur, R., Joshi, O., & Will, R. E. (2020). The ecological and economic determinants of eastern redcedar (*Juniperus virginiana*) encroachment in grassland and forested ecosystems: A case study from Oklahoma. *Journal of Environmental Management*, 254, 1-8.

Kaskie, K. D., Wimberly, M. C., & Bauman, P. J. (2019). Rapid assessment of juniper distribution in prairie landscapes of the Northern Great Plains. *International Journal of Applied Earth Observation and Geoinformation*, 83, 1–9.

Keller, S. R., Chhatre, V. E., & Fitzpatrick, M. C. (2017). Influence of range position on locally adaptive gene–environment associations in *Populus* flowering time genes. *Journal of Heredity*, 109, 47–58.

Knapp, A. K., Chen, A., Griffin-Nolan, R. J., Baur, L. E., Carroll, C. J. W., Gray, J. E., Hoffman, A. M., Li, X., Post, A. K., Slette, I. J., Collins, S. L., Luo, Y., & Smith, M. D. (2020). Resolving the Dust Bowl paradox of grassland responses to extreme drought. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 22249-22255.

LaForgia, M. L., Harrison, S. P., Latimer A. M. (2020). Invasive species interact with climatic variability to reduce success of natives. *Ecology*, 101, e03022. https://doi.org/10.1002/ecy.3022

Lawton, R. O. & Cothran, P. (2000). Factors influencing reproductive activity of *Juniperus virginiana* in the Tennessee Valley. *Journal of the Torrey Botanical Society*, 127, 271-279.

Leis, S. A., Blocksome, C. E., Twidwell, D., Fuhlendorf, S. D., Briggs, J. M., & Sanders L. D. (2017). Juniper invasions in grasslands: Research needs and intervention strategies. *Rangelands*, 39, 64–72.

Liu, Q., Wu, D., & Wang, C. (2022). Identification of genomic regions distorting population structure inference in diverse continental groups. *Quantitative Biology*, 10, 287-298.

Martin, M. (2011). CUTADAPT removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17, 10. https://doi.org/10.14806/ej.17.1.200

Martinod, K. L. & Gorchov, D. L. (2017). White-tailed deer browse on an invasive shrub with extended leaf phenology meets assumptions of an apparent competition hypothesis. *AoB Plants*, 9, 1-14.

McNeilly, T., & Antonovics, J. (1968). Evolution in closely adjacent plant populations IV. Barriers to gene flow. *Heredity*, 23, 205–218. https://doi.org/10.1038/hdy.1968.29

McNeish, R. E. & McEwan, R. W. (2016). A review on the invasion ecology of Amur honeysuckle (*Lonicera maackii*, Caprifoliaceae) a case study of ecological impacts at multiple scales. *Journal of the Torrey Botanical Society*, 143, 367-385.

Meisner, J. & Albrechtsen, A. (2019). Testing for Hardy-Weinberg equilibrium in structured populations using genotype or low-depth next generation sequencing data. *Molecular Ecology Resources*, 19, 1144-1152.

Michalczyk, I. M., Sebastiani, F., Buonamici, A., Cremer, E., Mengel, C., Ziegenhagen, B., & Vendramin, G. G. (2006). Characterization of highly polymorphic nuclear microsatellite loci in *Juniperus communis*. L. *Molecular Ecology Notes*, 6, 346–348.

Mizuki, I., Ishida, K., Tani, N., & Tsumura, Y. (2010). Fine-scale spatial structure of Genets and sexes in the dioecious plant *Dioscorea japonica*, which disperses by both bulbils and seeds. *Evolutionary Ecology*, 24, 1399–1415. https://doi.org/10.1007/s10682-010-9396-z

Moody, M. E. & Mack, R. N. (1988). Controlling the spread of plant invasions: The importance of nascent foci. *Journal of Applied Ecology*, 25, 1009-1021.

Mosca, E., Di Pierro, E. A., Budde, K. B., Neale, D. B., & González-Martínez, S. C. (2018). Environmental effects on fine-scale spatial genetic structure in four alpine keystone forest tree species. *Molecular Ecology*, 27, 647–658. https://doi.org/10.1111/mec.14469

Nathan, R., Schurr, F. M., Spiegel, O., Steinitz, O., Trakhtenbrot, A., & Tsoar, A. (2008). Mechanisms of long-distance seed dispersal. *Trends in Ecology & Evolution*, 23, 638-647.

Negi, V. S., Maletha, A., Pathak, R., & Maikhuri, R. K. (2021). Expansion of a native species and its impacts on alpine ecosystems, Indian Himalaya. *Biologia*, 76, 889-899.

Ohio Department of Natural Resources. Lakeside Daisy State Nature Preserve. ohiodnr.gov/wps/portal/gov/odnr/go-and-do/plan-a-visit/find-a-property/lakeside-daisy-state-nature-preserve. Accessed: January 7, 2022

Ohio Supercomputer Center. 1987. Ohio Supercomputer Center. Columbus OH: Ohio Supercomputer Center. http://osc.edu/ark:/19495/f5s1ph73.

Pannell, J. R., & Fields, P. D. (2013). Evolution in subdivided plant populations: Concepts, recent advances and future directions. *New Phytologist*, 201, 417–432. https://doi.org/10.1111/nph.12495

Peakall, R. & Smouse, P. E. (2006). Genalex 6: Genetic Analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288-295.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J. & Sham, P. C. (2007). PLINK: A toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, 81, 559-575.

Purcell, S. PLINK (version 1.90). (2020). http://pngu.mgh.harvard.edu/purcell/plink/.

Quinn, J. A. & Meiners, S. T. (2004). Growth rates, survivorship, and sex ratios of *Juniperus virginiana* on the New Jersey Piedmont from 1963 to 2000. *Journal Torrey Botanical Society*, 131, 187-194.

R Core Team (2022). R: version 4.2.1. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Ratajczak, Z., Nippert, J. B., & Collins, S. L. (2012). Woody encroachment decreases diversity across North American grasslands and savannas. *Ecology*, 93, 697-703.

Raymond, M & Rousset, F. (1995). GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248-249.

Rosenberg, N. A. (2003). District: A program for the graphical display of population structure. *Molecular Ecology Notes*, 4, 137-138.

Roser, L. G., Ferreyra, L. I., Saidman, B. O., & Vilardi, J. C. (2017). Ecogenetics: An R package for the management and exploratory analysis of spatial data in landscape genetics. *Molecular Ecology Resources*, 17, 241-250.

Rousset, F. (2008). Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, 8, 103-106.

Sagoff, M. (1999). What's wrong with exotic species? *Report from the Institute for Philosophy & Public Policy*, 19, 16-23.

Sangüesa-Barreda, G., García-Cervigón, A. I., García-Hidalgo, M., Rozas, V., Martín-Esquive, J. L., Martín-Carbaja, J., Martínez, R., & Olano, J. M. (2022). Vertical cliffs harbour millennia-old junipers in the Canary Islands. *Ecology*, e3633.

Schnelle, M. A. (2019). Native woody plants of the southern United States with weedy or invasive tendencies: A review of common offenders. *HortTechnology*, 29, 567-570.

Simberloff, D., Souza, L., Nunez, M. A., Barrios-Garcia, N., & Bunn, W. (2012). The natives are restless, but not often and mostly when disturbed. *Ecology*, 93, 598-607.

Smouse, P. E. & Peakall, R. (1999). Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity*, 82, 561-573

Sork, V. L. & Smouse, P. E. (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology*, 21, 821-836.

Triest, L., Sierens, T., & Terer, T. (2013). Diversity and fine-scale spatial genetic structure of *Cyperus papyrus* populations in Lake Naivasha (Kenya) using microsatellite markers. *Hydrobiologia*, 737, 131–144. https://doi.org/10.1007/s10750-013-1584-8

Tunnell, S. J., Stubbendieck, J., Huddle, J., Brollier, J. (2004). Seed dynamics of eastern redcedar in the mixed-grass prairie. *Great Plains Research*, 14, 129-142.

Uller, T. & Leimu, R. (2011). Founder events predict changes in genetic diversity during human-mediated range expansions. *Global Change Biology*, 17, 3478-3485.

Van Haverbeke, D. F. & Read, R. A. (1976). Genetics of eastern redcedar. *U.S. Forest Service, no. 32*.

Vasiliauskas, S. A., Aarssen, L. W. (1992). Sex ratio and neighbor effects in monospecific stands of *Juniperus virginiana*. *Ecology*, 73, 622-632.

Vilà, M., Weber, E., Antonio, C. M. D. (2000). Conservation implications of invasion by plant hybridization. *Biological Invasions*, 2, 207-217.

Vyšniauskienė, R., Rančelienė, V., Žvingila, D., & Patamsytė, J. (2011). Genetic diversity of invasive alien species *Lupinus polyphyllus* populations in Lithuania. *Žemdirbystė Agriculture*, 98, 383-390.

Wang, J. (1996). Deviation from Hardy-Weinberg proportions in finite populations. *Genetic Research*, 68, 249-257.

Ward, D. (2020). Shade is the most important factor limiting growth of a woody range expander. *PLoS ONE*, 15, 1-20.

Ward, D. (2021). Shade affects fine-root morphology in range-encroaching eastern redcedars (*Juniperus virginiana*) more than competition, soil fertility and pH. *Pedobiologia*, 84, 1-8.

Ward, D., Pillay, T., Mbongwa, S., Kirkman, K., Hansen, E., & Van Achterbergh, M. (2022). Reinvasion of native invasive trees after a tree-thinning experiment in an African savanna. *Rangeland Ecology and Management*, 81, 69-77.

Ward, S. M., Gaskin, J. F., & Wilson, L. M. (2008). Ecological genetics of plant invasion: What do we know? *Invasive Plant Science and Management*, 1, 98-109.

Ward, S. M. & Jasieniuk, M. (2009). Review: Sampling weedy and invasive plant populations for genetic diversity analysis. *Weed Science,* 57, 593-602.

Watson, P. A., Alexander, H. D., & Moczygemba, J. D. (2019). Coastal prairie recovery in response to shrub removal method and degree of shrub encroachment. *Rangeland Ecology and Management*, 72, 275-282.

Wickert, K. L., O'Neal, E. S., Davis, D. D., & Kasson, M. T. (2017). Seed production, viability, and reproductive limits of the invasive *Ailanthus altissima* (Tree-of-Heaven) within invaded environments. *Forests*, 8, 226.

Wickham H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag.

Willi, Y., Fracassetti, M., Bachmann, O., & Van Buskirk, J. (2020). Demographic processes linked to genetic diversity and positive selection across a species' range. *Plant Communications,* 1, 1-13. https://doi.org/10.1016/j.xplc.2020.100111

Yazlik, A. & Ambarli, D. (2022). Do non-native and dominant native species carry a similar risk of invasiveness? A case study for plants in Turkey. *NeoBiota,* 76, 53-72.

Zhao, J., Solís-Montero, L., Lou, A., & Vallejo-Marín, M. (2013). Population structure and genetic diversity of native and invasive populations of *Solanum rostratum* (Solanaceae). *PLoS One*, 8, 1-9.

Zimmermann, H., Ritz, C. M., Hirsch, H., Renison, D., Wesche, K., & Hensen, I. (2010). Highly reduced genetic diversity of *Rosa rubiginosa* L. populations in the invasive range. *International Journal of Plant Science*, 171, 435-446.