Landscape Genomics of the Tussock Cottongrass (Eriophorum vaginatum) and the Dwarf Birch (Betula nana) In North Central Alaska

Elizabeth Stunz

University of Texas at El Paso

Follow this and additional works at: https://scholarworks.utep.edu/open_etd

Part of the Biology Commons, Ecology and Evolutionary Biology Commons, and the Genetics Commons

Recommended Citation

Stunz, Elizabeth, "Landscape Genomics of the Tussock Cottongrass (Eriophorum vaginatum) and the Dwarf Birch (Betula nana) In North Central Alaska" (2022). Open Access Theses & Dissertations. 3734. https://scholarworks.utep.edu/open_etd/3734

This is brought to you for free and open access by ScholarWorks@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of ScholarWorks@UTEP. For more information, please contact lweber@utep.edu.
LANDSCAPE GENOMICS OF THE TUSSOCK COTTONGRASS (*ERIOPHORUM VAGINATUM*) AND THE DWARF BIRCH (*BETULA NANA*) IN NORTH CENTRAL ALASKA

ELIZABETH ANNE STUNZ

Doctoral Program in Ecology and Evolutionary Biology

APPROVED:

___________________________________________
Michael Moody, Ph.D., Chair

___________________________________________
Vanessa Lougheed, Ph.D.

___________________________________________
Ming-Ying Leung, Ph.D.

___________________________________________
Ned Fetcher, Ph.D.

___________________________________________
Stephen L. Crites, Jr., Ph.D.
Dean of the Graduate School
Dedication

I dedicate my dissertation to my partner, family, and close friends who have always been there for me. I also dedicate my dissertation to my family members and friends who have passed, including my dearest Imogen.
LANDSCAPE GENOMICS OF THE TUSSOCK COTTONGRASS (*ERIOPHORUM VAGINATUM*) AND THE DWARF BIRCH (*BETULA NANA*) IN NORTH CENTRAL ALASKA

by

ELIZABETH ANNE STUNZ, B.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Biological Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

December 2022
Acknowledgements

First and foremost, I would like to recognize and express my gratitude to my PhD advisor, Dr. Michael Moody. I am thankful for his critical eye, patience, and for giving me and my nontraditional education a chance. He’s an impressive researcher and writer, and has taught me to consider every aspect of research from all angles and to never settle for a seemingly easy answer. He’s showed me the benefit of working and problem solving independently, as well as a part of a team. Thank you for the conversation, inspiration, and invaluable botany and plant systematics skills. I will always appreciate the discussion of literature, both scientific and fiction.

Thanks to Stephen Escarzaga for his unending support and willingness to share his GIS skills. I would also like to extend my sincere thanks to my committee members, thank you for the support and guidance. Thank you, Dr. Jon Mohl, for the bioinformatics guidance and working with me to build my computer scientist skills. I will always be thankful for the thoughtful discussions and your patience. Thank you, Dr. Phil Lavretsky, for the ddRAD guidance, the access and use of your lab, and the conversation.

Thank you to Bob Muscarella and Joaquin Ortego for the demographic modeling advice, your guidance and dialogue is very much appreciated. Thank you to Austin Frisbey for field assistance and Tom Parker for his contribution to *Eriophorum vaginatum* collections. Thanks to Gus Shaver for invaluable help in the field, providing location information for collection sites, and guidance.

Finally, I could not have finished this PhD without the support of my amazing family and friends. Thank you for always having my back and sticking with me through it all. I can never express enough gratitude for your love, patience, and generosity.
Abstract

Global climate change has resulted in geographic range shifts of flora and fauna at a global scale. Extreme environments, like the Arctic, are seeing some of the most pronounced changes. This region covers 14% of the Earth’s land area, and while many arctic species are widespread, understanding ecotypic variation at the genomic level will be important for elucidating how range shifts will affect ecological processes. Increase in shrub cover is a major effect of ongoing climate change in arctic tundra ecosystems. The relative increases in abundance and cover of shrub species such as birch, willow, and alder (Betula, Salix, and Alnus spp.) are predicted to modify ecological communities by altering ecosystem processes and outcompeting other arctic plant species, such as the tussock cottongrass (Eriophorum vaginatum L.). Eriophorum vaginatum is a foundation species of the moist acidic tundra, whose potential decline due to competition from shrubs may affect ecosystem stability in the Arctic. Here, I examine the genomic population structure, local adaption, and genotype-environment associations of two widespread arctic plant species, the tussock cottongrass (Eriophorum vaginatum) and the dwarf birch (Betula nana L.) using thousands of genomic markers obtained from double-digest Restriction-site Associated DNA sequencing (ddRAD-seq). I then compare environmental niche models for both species from the Last Glacial Maximum (LGM) to the year 2070 to examine the potential of range expansion and persistence of each species in a warming arctic.

In Chapter 1, genetic variation was identified in 273 individuals of E. vaginatum from 17 sites along a latitudinal gradient in north central Alaska. These 17 sites were selected due to their inclusion in 30+ years of ecological research as well as their location within a region that was part of the Beringian refugium. A genome-wide SNP dataset was used to investigate population structure, genomic diversity, genotype by environment association and environmental niche
modeling. A comprehensive dataset of 3,879 loci and 10,734 SNPs was used to conduct genotype by environment association analyses and revealed environmentally-associated variation. A neutral dataset of 2,776 loci represented by a single SNP was used to conduct population structure, genomic diversity and landscape resistance analyses across the sampled range of *E. vaginatum*, and supported multiple genetic clusters across sites, including a genetic break between populations north and south of treeline. Gene flow, landscape resistance, and genotype-environment association analyses all supported the influence of subrefugial isolation, contemporary isolation by resistance, and adaptation on current population structure. Using genotype-environment association analyses, 45 candidate loci were identified, with most identified genes related to abiotic stress. Our results supported a hypothesis of limited gene flow related to both spatial and environmental factors for *E. vaginatum*. These results, in combination with life history traits, suggest a limited range expansion of southern ecotypes northward as the tundra warms. These results also have implications for northern ecotypes, as lower competitive attributes may put this foundational species at a disadvantage as the tundra warms and shrub cover increases.

In Chapter 2, I used a genome-wide SNP dataset to investigate population structure, genomic variation, and local adaptation of 109 *B. nana* individuals from 9 sites along a latitudinal gradient in north central Alaska. These sites were chosen to overlap with those sampled for *E. vaginatum* in Chapter 1 to allow for a comparison of population structure, genomic variation, and adaptation of the two species in the same region. A neutral SNP dataset of 1,039 loci (each represented by a single SNP) was used to demonstrate two genetic clusters, one composed of individuals from the No Name (NN) site and the other composed of individuals from all other sites. The general lack of population structure and absence of allelic variation related to environment along the cline of the latitudinal gradient was likely due to high co-ancestry,
incomplete lineage sorting of a relatively continuous population with recent isolation, or previously disjunct populations reconnected by contemporary widespread gene flow. The low levels of co-ancestry of NN with the other sites in addition to the high number of private alleles identified for NN may indicate the presence of *B. glandulosa*, or an admixed variant between *B. nana* and *B. glandulosa* at this site. The lack of structure related to environment, treeline, and geography suggests that *B. nana* did not share a similar evolutionary history with *E. vaginatum* along the same latitudinal gradient. The increased prevalence of *Betula* pollen and macrofossils in the region during warming fluxes during the early and mid-Holocene, and generally higher levels of *Betula* pollen south of the Brooks Range further suggest that while *Betula*, and potentially *B. nana*, was present on the north side of Brooks Range during the LGM, the genus was likely not prevalent. Post-glacial expansion of southern *B. nana* populations northward could also lead to high levels of co-ancestry between populations north and south of the Brooks Range. The lack of genetic structure and genetic signature related to environmental variation could indicate a higher tolerance for environmental shifts (plasticity) across the range, which could facilitate a competitive advantage for genotypes under climate change.

In Chapter 3, I conducted a literature review of graminoid and deciduous shrub distribution patterns in the Arctic and used environmental niche modeling to investigate temporal variation of suitable habitat for *E. vaginatum* and *B. nana* in Alaska. Environmental niche models (ENMs) demonstrated small areas of moderately suitable habitat for *B. nana* during the LGM, with general increases in suitable habitat area through the Mid-Holocene and the present, primarily in southern and west central Alaska. For *E. vaginatum*, highly suitable habitat decreased in cover from the LGM to the Mid-Holocene, and both moderately and highly suitable habitat continued to decrease in the present, with a suitable habitat shift northward since the LGM. Importantly, and as supported
by the literature review, modeling actual species distributions is complex, and the incorporation of population-level and ecological community factors would improve predictions of realistic expansion and population persistence under climate change. While shrubs are expected to increase in height and density, and landscape resistance could hinder range expansion of graminoids as the Arctic warms, fine-scale environmental variation, nutrient availability, dispersal, gene flow, genetic differentiation, local adaptation, competition, and community structure will also shape both species distributions. While *B. nana* will likely expand rapidly in areas of highly suitable habitat in Alaska, as supported by a lack of local adaptation and narrow environmental tolerances (Chapter 2), moist and warm sites will likely see the greatest increases in density. While *E. vaginatum* may be able to dominate in colder and/or wetter sites, especially as deeper rooting will allow access to nutrients in thawed permafrost layers, *E. vaginatum* extent is not expected to increase in the taiga, where landscape resistance is high, treeline reduces gene flow between populations, and competition with shrubs will likely limit population expansion in the warming arctic tundra.

My dissertation work utilizes genome-wide SNP datasets to investigate the evolutionary potential of two arctic plant species under climate change. Using landscape-level genomic analyses, population structure, genotype-environment associations, and gene flow barriers were well-supported for *E. vaginatum* and lacking for *B. nana*, highlighting disparate evolutionary histories and trajectories for these species in north central Alaska. These findings can inform conservation efforts to promote persistence and genetic variation of either or both species under climate change, given the conservation objective.
# Table of Contents

Dedication........................................................................................................................................ iii
Acknowledgements........................................................................................................................... v
Abstract............................................................................................................................................... vi
Table of Contents ............................................................................................................................... x
List of Tables ......................................................................................................................................... xiii
List of Figures ......................................................................................................................................... xiv

Chapter 1: Landscape genomics provides evidence of ecotypic adaptation and a barrier to gene flow at treeline for the arctic foundation species *Eriophorum vaginatum* .......... 1
  Abstract ............................................................................................................................................... 2
  Introduction ......................................................................................................................................... 3
  Materials and Methods ..................................................................................................................... 7
    *Study area* ................................................................................................................................. 7
    *Sample collection* ...................................................................................................................... 8
    *DNA extraction and ddRAD-seq library preparation* ............................................................... 8
    *SNP identification and genotyping* ............................................................................................ 10
    *Patterns of genomic diversity* .................................................................................................... 12
    *Population structure* .................................................................................................................. 12
    *Demographic and environmental niche modeling* ................................................................. 13
    *Genotype-environment association* ......................................................................................... 16

Results ............................................................................................................................................... 18
  *Genomic sequence data and genomic diversity* ........................................................................ 18
  *Population structure* .................................................................................................................... 19
  *Demographic and environmental niche modeling* ................................................................. 20
  *Genotype-environment association* ......................................................................................... 21

Discussion .......................................................................................................................................... 22
  *Landscape genomics of an arctic foundation species* ............................................................... 22
  *Environmental heterogeneity explains allelic turnover across a latitudinal cline* .......... 27

Data Availability Statement ............................................................................................................... 30
### Contents

- **Conflict of Interest** .......................................................................................................................... 30
- **Author Contributions** ......................................................................................................................... 30
- **Funding** .................................................................................................................................................. 31
- **Acknowledgments** ................................................................................................................................ 31
- **References** ............................................................................................................................................. 31
- **Supplementary Material** ....................................................................................................................... 51
- **Tables and Figures** ................................................................................................................................. 52

**Chapter 2: Population structure, demographics, and local adaptation of the arctic dwarf birch**

*Betula nana* .......................................................... 62

- **Abstract** .................................................................................................................................................. 63
- **Introduction** ............................................................................................................................................. 65
- **Materials and Methods** .......................................................................................................................... 72
  - *Study area* ............................................................................................................................................. 72
  - *Sample collection* ................................................................................................................................. 74
  - *DNA extraction and ddRAD-seq library preparation* ........................................................................... 75
  - *SNP identification and genotyping* ........................................................................................................ 77
  - *Patterns of genomic diversity* .............................................................................................................. 79
  - *Population structure* ............................................................................................................................ 79
  - *Genotype-environment association* ..................................................................................................... 81
- **Results** ................................................................................................................................................... 82
  - *Genomic sequence data and genomic diversity* .................................................................................. 82
  - *Population structure* ............................................................................................................................ 84
  - *Genotype-environment association* ..................................................................................................... 88
- **Discussion** ............................................................................................................................................... 88
- **References** ............................................................................................................................................ 94

**Chapter 3: Comparative ENM modeling of arctic dwarf birch**

*Betula nana* and the arctic tussock cottongrass *Eriophorum vaginatum* in Alaska in the context of a review of overall graminoid and deciduous shrub distribution patterns in the Arctic .............................................. 108

- **Abstract** .................................................................................................................................................. 109
- **Introduction and Review** .......................................................................................................................... 111
  - *Distribution of genetic diversity in the Arctic* ..................................................................................... 115
  - *Population structure in the Arctic* ........................................................................................................ 118
List of Tables

**Table 1.** Collection sites, GPS coordinates, elevation (meters), and vegetation type of *E. vaginatum* in north central Alaska. MAT = Moist Acidic Tundra. Genetic cluster assignment based on STRUCTURE analysis at *K* = 3 (**Figure 4**). ................................................................. 52

**Table 2.** Genetic diversity summary and demographic statistics for the neutral data set of the 17 *Eriophorum vaginatum* sites sampled in north central Alaska. Sites are ordered geographically from South to North and site abbreviations follow **Table 1**. Sample size (*N*), allelic richness (*A*), private alleles (*P*), observed heterozygosity (*H*), expected heterozygosity (*H*), inbreeding coefficient (*F*), and effective population size (*N*). ................................................................. 53

**Table 3.** Maximum likelihood of population effects (MLPE) models relating pairwise *F* to pairwise distance matrices of isolation by resistance (IBR), isolation by environment (IBE), and isolation by distance (IBD) and ranked by AICc. *R* = marginal *R* approximation of mixed model fixed effects. ................................................................. 53

**Table 4.** Annotated *E. vaginatum* candidate genes with a percentage identity match of at least 80.0 and an E-value threshold of at least 1 X 10^-4 and RDA *R* value ≥ 0.8. RDA Predictors had the highest *R* values for that association with the given gene. 2’ RDA Predictors had *R* values > 0.7 and are listed from highest to lowest. E-value and Similarity % for gene ID of each locus. MF = Molecular Function, and BP = Biological Process. *Genes with stress response association. ................................................................. 53

**Table 2.1.** Collection sites, GPS coordinates, elevation (meters) and vegetation type of *B. nana* in north central Alaska. MAT = Moist Acidic Tundra. ............................................................................. 74

**Table 2.2.** Annotated *B. nana* outlier genes with a percentage identity match of at least 80.0 and an E-value threshold of at least 1 X 10^-4. ............................................................................. 83

**Table 2.3.** Genetic diversity summary and demographic statistics for the neutral data set (after clone removal) of the 9 *Betula nana* populations and 8 overlapping *Eriophorum vaginatum* populations sampled in north central Alaska. Populations are ordered geographically from South to North and site abbreviations follow **Table 2.1**. Sample size (*N*), allelic richness (*A*), private alleles (*P*), observed heterozygosity (*H*), expected heterozygosity (*H*), inbreeding coefficient (*F*), and effective population size (*N*). ............................................................................. 84
List of Figures

**Figure 1.** Map of *Eriophorum vaginatum* sampling locations and modeled maximum Pleistocene glacial extent (Kaufman and Manley, 2004; Kaufman et al., 2011) along a latitudinal gradient in north central Alaska. Blue stars designate reciprocal transplant gardens (Shaver et al. 1986; Bennington et al. 2012). Treeline is indicated by the dashed black line and The Continental Divide is indicated by the burnt orange line. The inset shows the extent of the Beringian region, outlined with a dashed yellow line. The 17 collection site abbreviations are as for Table 1. 56

**Figure 2.** Images showing (A) the habitat of the Coldfoot sampling location south of treeline, (B) a mature *Eriophorum vaginatum* tussock at Coldfoot, and (C) the habitat of the Prudhoe Bay sampling location in north central Alaska. All photos by E. Stunz. 57

**Figure 3.** Maps of the study area with (A) mean temperature of the driest quarter (tdq) and (B) precipitation seasonality (prs) WorldClim underlying data layers using ARCMAP v.10.7.1 (ESRI, 2011). Note that *divMigrate* migration values ≥ 0.80 and direction indicated with arrows are also provided in 2A. Site abbreviations as for Table 1. 58

**Figure 4.** Population structure results from the neutral SNP data set. (A) Scatter plot from the discriminant analysis of principal components (DAPC) with 80 principal components retained on the first two Discriminant Analysis axes showing the differentiation between the three groups and inertia ellipses. Each color represents a cluster as identified with the Bayesian Information Criterion (BIC). Red dots represent individuals from the Eagle Creek (EC) population, yellow dots represent individuals from the population south of treeline except for EC (South), and blue dots represent individuals from populations north of treeline (North). (B) Bar graph of *structure* results for population structure analysis. Each vertical bar represents an individual, and colors show the proportion of ancestry assigned to each of the three clusters (K = 2 and K = 3), as inferred from ΔK values. *Eriophorum vaginatum* populations are ordered from south to north location along the latitudinal gradient in north central Alaska. See Table 1 for collection site abbreviations. 59

**Figure 5.** MaxEnt environmental niche model (ENM) maps of Alaska, U.S.A. adapted to depict *Eriophorum vaginatum* habitat suitability during the Last Glacial Maximum (LGM), Mid-Holocene and present. Modeled LGM layers were derived from the Community Climate System Model (CCSM) (Braconnot et al., 2007). Current and LGM layer habitat suitability bin values were averaged to create the Mid-Holocene climate raster. 60

**Figure 6.** RDA plot demonstrating predictor associations as related to ecotype for the comprehensive SNP data set. iso = Isothermality, tdq = Mean temperature of the driest quarter (Feb, Mar, Apr), twq = Mean temperature of the warmest quarter (Jun, Jul, Aug), prd = Precipitation of the driest month (Apr), prs = Precipitation seasonality. 61

**Figure 2.1.** Map of *Betula nana* sampling locations and modeled maximum Pleistocene glacial extent (Kaufman and Manley, 2004; Kaufman et al., 2011) along a latitudinal gradient in north central Alaska. Treeline is indicated by the dashed black line. The 9 collection site abbreviations are as for Table 2.1. 73

**Figure 2.2.** *fineRADstructure* co-ancestry matrix. Co-ancestry pairwise coefficients are color-coded from low (yellow) to high (black) and individuals are clustered according to the pairwise matrix of co-ancestry coefficients in the dendrogram. Site abbreviations follow Table 2.1. The No Name (NN) group is distinguished with a light blue branch in the dendrogram and bars alongside NN individuals in the matrix. 85
Figure 2.3. Population structure results from the neutral SNP data set. (A) Scatter plot from the discriminant analysis of principal components (DAPC) with 40 principal components retained on the first two Discriminant Analysis axes showing the differentiation between the 9 populations and inertia ellipses. Each color represents a sampled site. Blue dots represent individuals from the Eagle Creek (EC) population, purple dots represent individuals from No Name (NN) population, lavender dots represent individuals from the Gobbler’s Knob (GO) population, gold dots represent individuals from the Coldfoot (CF) population, yellow dots represent individuals from the Timberline (TB) population, light orange dots represent individuals from the Atigun Camp (AT) population, orange dots represent individuals from the Toolik Lake (TL) population, dark orange dots represent individuals from the Sagwon River (SR) population and red dots represent individuals from the Sagwon (SG) population. (B) Bar graph of STRUCTURE results for population structure analysis. Each vertical bar represents an individual, and colors show the proportion of ancestry assigned to each of the two clusters ($K = 2$), as inferred from $\Delta K$ values. *Betula nana* populations are ordered from south to north location along the latitudinal gradient in north central Alaska.

Figure 3.1. MaxEnt environmental niche model (ENM) maps of Alaska, U.S.A. adapted to depict (A) *Betula nana* and (B) *Eriophorum vaginatum* habitat suitability during the Last Glacial Maximum (LGM), Mid-Holocene, ‘present’ (averages from 1970 to 2000 climate records), and future (2070). Modeled LGM and future layers were derived from the Community Climate System Model (CCSM) (Braconnot et al., 2007), with future layers derived from the Representative Concentration Pathway (RCP) 8.5 model. Current and LGM layer habitat suitability bin values were averaged to create the Mid-Holocene climate raster.
Chapter 1: Landscape genomics provides evidence of ecotypic adaptation and a barrier to
gene flow at treeline for the arctic foundation species *Eriophorum vaginatum*

Elizabeth Stunz¹*, Ned Fetcher², Philip Lavretsky¹, Jonathon E. Mohl³, Jianwu Tang⁴, 
Michael L. Moody⁴*

¹University of Texas at El Paso, Department of Biological Sciences, El Paso, TX, United States
²Wilkes University, Institute for Environmental Science and Sustainability, Wilkes-Barre, PA, United States
³University of Texas at El Paso, Department of Mathematical Sciences and Border Biomedical Research Center, El Paso, TX, United States
⁴Marine Biological Laboratory, The Ecosystems Center, Woods Hole, MA, United States

* Correspondence:
Elizabeth Stunz
estunz@utep.edu

Michael L. Moody
mlmoody@utep.edu

Keywords: Arctic¹, climate change², *Eriophorum vaginatum*³, landscape genomics⁴, environmental niche modelings⁵, genotype-environment association analyses⁶, refugia⁷

Number of words: 8,226

Number of figures: 6; 1 supplemental

Number of tables: 4; 7 supplemental
ABSTRACT

Global climate change has resulted in geographic range shifts of flora and fauna at a global scale. Extreme environments, like the Arctic, are seeing some of the most pronounced changes. This region covers 14% of the Earth’s land area, and while many arctic species are widespread, understanding ecotypic variation at the genomic level will be important for elucidating how range shifts will affect ecological processes. Tussock cottongrass (*Eriophorum vaginatum* L.) is a foundation species of the moist acidic tundra, whose potential decline due to competition from shrubs may affect ecosystem stability in the Arctic. We used double-digest Restriction Site-Associated DNA sequencing to identify genomic variation in 273 individuals of *E. vaginatum* from 17 sites along a latitudinal gradient in north central Alaska. These sites have been part of 30+ years of ecological research and are inclusive of a region that was part of the Beringian refugium. The data analyses included genomic population structure, demographic models, and genotype by environment association. Genome-wide SNP investigation revealed environmentally-associated variation and population structure across the sampled range of *E. vaginatum*, including a genetic break between populations north and south of treeline. This structure is likely the result of subrefugial isolation, contemporary isolation by resistance, and adaptation. Forty-five candidate loci were identified with genotype-environment association analyses, with most identified genes related to abiotic stress. Our results support a hypothesis of limited gene flow based on spatial and environmental factors for *E. vaginatum*, which in combination with life history traits could limit range expansion of southern ecotypes northward as the tundra warms. This has implications for lower competitive attributes of northern plants of this foundation species likely resulting in changes in ecosystem productivity.
INTRODUCTION

Investigating the adaptive constraints of plants is critical to better determine how different species and communities may respond to an altered climate (Shaver et al., 2000) and how best to predict ecological community shifts in the future (Ikeda et al., 2017b, Smith et al., 2019). Ecological communities are transforming and species distributions are shifting in response to a changing climate. Such changes are more evident in environments with a steep transition between ecosystems, or in biomes experiencing extreme effects of global climate change (Chen et al., 2011; Felde et al., 2012). Among these, arctic and alpine environments are areas experiencing some of the most pronounced effects of climate change, with a temperature increase of more than 0.5°C per decade over the past 40 years in the arctic tundra alone (Serreze et al., 2000; Sturm et al., 2005; Chandler et al., 2015) and an 11°C increase projected by 2100 (IPCC, 2014). Changes in climate have already led to plant community shifts (Villarreal et al., 2012; Hollister et al., 2015), including northerly shifts in treeline (Harsch et al., 2009) and increased shrub density (Tape et al., 2006). These shifts are also important for foundation species in which ecotypic variation influences gross primary productivity (GPP) and ecosystem response to climate change (Curasi et al., 2019). This has implications for the carbon cycle, as the soil and vegetation in the Arctic store large amounts of carbon (Schaefer et al. 2014). Determining population structure, or lack thereof, across species’ range distributions and understanding whether environmental variables are responsible for such patterns are both essential to model potential responses to ongoing climate change (Smith et al., 2019; Reiskind et al., 2021).

Among geological events with lasting effects on populations, glaciation events often serve as a major factor leading to genetic divergence (Ferris et al., 1993; Soltis et al., 1997). In fact, comparative analyses of spatial genetic structure and diversity of multiple circumpolar arctic plant
species identify Pleistocene glaciation, the resulting physical barriers, and glacial refugia as responsible for shaping patterns across the Arctic (Alsos et al., 2012; Eidesen et al., 2013). The Beringian region, which remained mostly unglaciated throughout the Pleistocene glacial-interglacial cycles, is known to have served as an important refugium for multiple arctic plant and animal species, and therefore served as sources for post-glacial expansions of these populations (Abbott and Brochmann, 2003; Alsos et al., 2005; Brubaker et al., 2005). Identifying refugium origin for the contemporary distribution of arctic taxa has been an emphasis of circumpolar population genetic studies (Alsos et al., 2005; 2007), whereas evidence for differently adapted genotypes advancing from refugia has only begun to be explored (Ikeda and Setoguchi, 2017; Napier et al., 2019; Wang et al., 2021).

Refugia like the Beringian region had a heterogeneous landscape (Billings, 1992; Lapointe et al., 2017), and local adaptations within these landscapes could have a comparably strong effect on the post-glaciation distribution of genotypes. During the Last Glacial Maximum (LGM; ~20 thousand years before present (kyr BP)), tundra-steppe in the north and spruce woodlands in the south of the Beringian refugium were isolated by glaciation over the Brooks Range (Billings, 1992; Lapointe et al., 2017), forming potential subrefugia. Post-glaciation, these habitats are largely differentiated by treeline extending along the interface of the taiga-tundra biomes, which functions as an ecological barrier (Chapin et al., 1996) generally delimited by permafrost depth, soil availability, growing season temperature, and reduced albedo (Billings, 1973; Chapin et al., 1996). Treeline can also pose a physical barrier for wind-pollinated or wind-dispersed plants, both of which are dominant among arctic taxa (Chapin et al., 1996; Dahl, 1963). Therefore, there is both older as well as more contemporaneous landscape heterogeneity in the Arctic that could lead to adaptive and physical barriers through time.
Landscape-scale genetic constraints for arctic plant species have been proposed, especially for ecotypic specialization and adaptational lag (Bennington et al., 2012; Mazer et al., 2013; McGraw et al., 2015). Local adaptation across a species’ range can lead to differences in the thermal optima or climatic niches of populations, resulting in ecotypes with narrower environmental tolerances if adaptation is strong (Forester et al., 2016; Peterson et al., 2018). This is pertinent to the Alaskan Arctic where there has been an environmental cline that has remained relatively stable over the last ~6,000 years (Billings, 1992). Lag in dispersal and establishment can hamper plant ecotypes from adjusting their ranges to track and remain in climate optima, resulting in reduced fitness of local genotypes, or adaptational lag. Here we use tussock cottongrass (*Eriophorum vaginatum* L.; Cyperaceae), a sedge that exemplifies arctic plant distribution across tundra and taiga biomes and occurs throughout the Beringian region, to investigate landscape genomic patterns.

*Eriophorum vaginatum* is a wind-pollinated and wind-dispersed sedge, and a foundation species of the moist acidic tundra (MAT). It has a circumarctic and circumboreal distribution and is a common component of taiga forest in muskeg and bogs (Wein, 1973; Fetcher and Shaver, 1982, 1983; Brown and Kreig, 1983). As a foundation species, *E. vaginatum* plays a strong role in structuring the ecological network where it occurs. In tundra sites, *E. vaginatum* can account for up to one-third of ecosystem productivity (Chapin and Shaver, 1985), and is prevalent throughout Alaska, northern Canada, and northern Russia. Because tussocks are densely distributed and can persist over 100 years, new recruitment is likely uncommon (McGraw and Shaver, 1982; Gartner et al., 1983). This is pertinent under climate change, as the climate optima for ecotypes of tussock cottongrass was displaced ~140 km northwards between 1993 and 2010 in Alaska (McGraw et al., 2015).
Importantly, over 30 years of reciprocal transplant studies have uncovered measurable phenotypic variation of *E. vaginatum* across an arctic latitudinal gradient within the geographic range of the eastern Beringian refugium (Figure 1). These include: (1) Tussocks transplanted back into their home sites showing home site advantage in flowering rates and survival; (2) Some tussock adaptations correlated with latitude of population origin in light-saturated photosynthetic rate and stomatal density; and (3) Tussock adaptations related to leaf phenology and plastic responses correlated with the site of origin occurring north or south of the treeline, which signifies the arctic tundra/taiga ecosystem shift (Fetcher and Shaver, 1990; Bennington et al. 2012; Parker et al., 2017, 2021; Peterson et al., 2012; Souther et al., 2014). These adaptations have led to the hypothesis that across this latitudinal gradient there are genetic constraints related to ecotype site of origin, and ecological adaptations should be present at multiple scales that will be manifest through landscape genomic analyses.

Here we use thousands of double-digest Restriction Site-Associated DNA sequencing (ddRAD-seq) single nucleotide polymorphisms (SNPs) to investigate broad patterns of population structure, gene flow, and associations with landscape variables in *E. vaginatum* populations along the latitudinal gradient in the north central Alaskan Arctic that is the site of the long-term ecological studies cited above. We address the following questions: (1) Is population structure delimited at the ecosystem level among these populations? (2) Is there evidence of genetic structure linked to adaptation for latitude of origin? We hypothesize that the Brooks Range glaciation within the Beringian refugium effectively isolated populations to the north and south and that this resulted in the confinement of *E. vaginatum* to putative subrefugia as glaciers expanded before the LGM. When separated for a sufficient amount of time, genetic drift and/or divergent selection should have altered the genomes of the isolated *E. vaginatum* populations, and population structure
analyses should demarcate these isolated populations as unique. More recently, treeline may be expected to pose an important barrier to gene flow for contemporary populations of *E. vaginatum*, as the treeline ecotone represents the current boundary between glacially isolated habitats. Finally, we hypothesize that genetic correlation associated with environmental predictors will be uncovered along the stable arctic latitudinal cline corresponding to local niche dynamics that have resulted in the ecotypic variation identified through 30+ years of reciprocal transplant garden studies.

**Materials and Methods**

**Study area**

The study area consisted of a latitudinal gradient covering ~426 km in north central Alaska (*Figure 1; Table 1*), from north of Fairbanks (65.433°, -145.512°) to Prudhoe Bay (70.327°, -149.065°). The Continental Divide at the crest of the Brooks Range divides the region into south slope and north slope components, roughly at the intersection of two climatic regions, the Arctic and the interior (Haugen, 1982; Brown and Kreig, 1983). Beginning just north of Fairbanks, taiga vegetation is dominant, with forest and lowland treeless bogs comprising much of the interior. Alpine or tussock tundra is found once elevations exceed ~700 m. Treeline occurs on the south slope of the Brooks Range (*Figure 1*), while tundra plant communities are found north of treeline. Permafrost is continuous north of treeline and discontinuous in most of the interior, where soil parent materials, slope angle and aspect, drainage, and vegetation often indicate permafrost presence (Brown and Kreig, 1983). In general, the interior is warmer and annual rainfall declines north of the Brooks Range towards the coastal plain (*Supplementary Table S1*).
**Sample collection**

During the summers of 2015, 2016, and 2017, leaf samples were collected from *E. vaginatum* from 16-18 individuals at each of 17 sites (273 accessions) along the latitudinal gradient following the Dalton Highway (Figures 1, 2). We sampled away from reciprocal transplant gardens and used comparably large individuals to avoid the potential of using plants that would have germinated since the transplant gardens were established. Logistical access to this region is limited and cost prohibitive outside of the Dalton Highway as it can only be accessed by helicopter, which was needed for two coastal plain collections. Latitudinal distance between populations was ≤ 0.75° with denser sampling near treeline and at the southern end of the range. Individuals were sampled at least 50 m from the roadside and at least 10 m apart to minimize bias based on same seed parentage among the tussocks. Sampled leaves were dried and stored in silica gel. Sites north of treeline are classified as MAT (Britton, 1966; Wein and Bliss, 1974; Shaver et al., 1986), while most sites south of treeline are muskeg or tussock bog unless above 700 m (Table 1) (Cleve and Dyrness, 1983; Kummerow, 1983; Shaver et al., 1986). The sampling strategy was designed to include sites that overlapped with gardens used in long-term ecological studies (Shaver et al., 1986).

**DNA extraction and ddRAD-seq library preparation**

Genomic DNA was extracted from 50 mg of dried leaf tissue for 16 samples per site using a modified CTAB method (Doyle and Doyle, 1987). DNA concentrations were quantified using the Qubit dsDNA BR Assay Kit (Invitrogen) and Qubit 3.0 Fluorometer (Thermo Fisher Scientific). DNA samples with a minimum concentration of 0.02 µg/µL were included for
preparation of double-digest Restriction Site-Associated DNA (ddRAD) -seq libraries, which generally followed DaCosta and Sorenson (2014) and Hernández et al., (2021).

To assemble ddRAD libraries, 43 µL of each ~25 ng/µL DNA sample was digested with 20 U of EcoRI and 20 U of MspI restriction enzymes (New England Biolabs; cut sites: 5’-GAATTC-3’; 5’-CCGG-3’) and 5 µL of 10× NEBuffer 4 (New England Biolabs) in a thermocycler at 37°C for 30 min followed by enzyme deactivation at 65°C for 20 min. Custom barcodes and indices were ligated to digested DNA, with the addition of 50 nM of custom EcoRI and MspI adapters (containing individual barcodes), 2 µL of 10× NEBuffer (New England Biolabs), 0.6 µL of rATP (Promega), 0.4 µL of ddH₂O, and 1 µL of T4 DNA ligase (New England Biolabs) to each 50 µL sample of digested DNA. Ligation was completed in the thermocycler (22°C for 30 min), followed by enzyme deactivation (65°C for 20 min). Adapter-ligated DNA fragments were then double-side size selected with a 0.8× SPRI bead clean up. Right-sided selection for large fragments was done with AMPure XP beads (Beckman Coulter, Inc.) added at 0.55× volume of the starting ligated DNA solution, and the supernatant transferred to new tubes. Subsequently, a left-sided size selection was done with 0.25× volume AMPure XP beads added to the supernatant. The supernatant was then discarded and the beads were washed twice with 80% ethanol and air-dried. DNA was re-suspended with 25 µL ddH₂O and eluted for a minimum of 30 minutes.

To amplify size-selected DNA fragments, PCR was performed with a solution of 15 µL template DNA, 30 µL of Phusion High-Fidelity PCR Master Mix (Thermo Scientific), 9 µL of ddH₂O and 3 µL of each 10 µM primer specific to sequences at the ends of each custom barcode and index (see DaCosta & Sorenson, 2014). PCR was performed in a thermocycler: 98°C for 30 s; 22 cycles at 98°C for 10 s, 60°C for 30 s, 72°C for 40 s and 72°C for 5 min. Amplified DNA fragments across samples were cleaned using a 1.8× AMPure XP bead clean-up protocol.
(Beckman Coulter, Inc.). After adding beads and discarding the supernatant, beads were washed twice with 80% ethanol and re-suspended in 40 µL ddH$_2$O. Finally, cleaned PCR products were quantified with the Qubit dsDNA BR Assay Kit (Invitrogen) and Qubit 3.0 Fluorometer (Thermo Fisher Scientific). Additionally, samples were visualized with gel electrophoresis to ensure that all were similarly size-selected. Equimolar concentrations of samples with unique barcode combinations were pooled. Multiplexed libraries were sequenced at the Center for Genome Research and Biocomputing at Oregon State University on an Illumina HiSeq 3000.

**SNP identification and genotyping**

Files of raw Illumina read sequence data were de-multiplexed and analyzed via the *de novo* assembly method in **STACKS** 2.41 (Catchen et al., 2011, 2013), and using in-house Python scripts. First, the *process_radtags* program in **STACKS** was used to filter out reads with a Phred Quality score < 33, remove reads without intact radtags, and trim all reads to 145 bp. The sequences were then processed in the program *ustacks* to align the short-read sequences into ‘stacks.’ Stacks were compared using a maximum likelihood framework (Hohenlohe et al., 2010) to identify loci and detect single nucleotide polymorphisms (SNPs). The minimum depth of coverage for stack creation (m) was set to three (m 3; default), The maximum distance (in nucleotides) between each stack (M) was set to four (M 4) and all other parameters were set to default values. The *cstacks* program was used to build a catalog of consensus loci based on matching sets of reads built in *ustacks*. The number of mismatches permitted between sampled loci when building a catalog (n) was set to four (n 4). Parameter choice was guided by optimization methods of Paris et al. (2017) by assessing the number of loci and SNPs retained across 80% of individuals from each site (r80).
for values 1-9 for M, assuming M = n, and fixing m to 3 (Paris et al., 2017; Rochette and Catchen, 2017).

Sets of putative loci were then compared to the catalog of loci with the sstacks program. Data was arranged by locus, instead of by sample, with the tsv2bam program. The ddRAD data was analyzed locus by locus across all individuals to genotype individuals for each SNP with the gstacks program. The data set was then pruned for minor alleles occurring in less than 1% of reads, as rare genotypes are likely the result of sequencing error (Tabangin et al., 2009). The STACKS populations program was run to retain only loci found in all sampled populations with ≤ 20% missing data. Two sets of SNPs were created: one single SNP per locus “stack” set containing putatively neutral markers (or the neutral dataset) and a set containing all SNPs (or the comprehensive set) used for RDA analysis.

Loci under balancing or directional selection were identified with BayeScan v.2.1 (Foll and Gaggiotti, 2008) via a Bayesian $F_{ST}$ outlier test with default values, 20 pilot runs of 5,000 iterations were implemented, followed by 100,000 iterations, sampled every 10 iterations, with a 50,000 iteration burn-in. Outputted iterations were increased to ensure convergence of the MCMC chain (confirmed by visual assessment). In addition, prior odds settings were increased to 500 for the comprehensive dataset, as recommended for candidate loci identification in large datasets (Foll, 2010), and held at 100 for the neutral dataset. Candidate outlier loci were identified as those with a $q$-value < 0.05 (or a false discovery rate [FDR] of 5%). Additionally, SNPs deviating from Hardy-Weinberg Equilibrium (HWE; p ≤ 0.001) were identified with PLINK$^{1}$ v.1.07 (Purcell et al., 2007). Outlier loci and SNPs deviating from HWE were removed to create the neutral SNP data set used for estimating genetic diversity, demographic statistics, and population structure.

---

$^{1}$ http://pngu.mgh.harvard.edu/purcell/plink/
Patterns of genomic diversity

Heterozygosity (observed ($H_o$) and expected ($H_e$)), allelic richness ($Ar$) and the inbreeding coefficient ($F_{IS}$) were estimated across all loci and for each population with the \textit{divBasic} function in the \texttt{diveRsity} \texttt{R} package v.1.9.90 (Keenan et al., 2013). Private alleles were estimated with the \texttt{R} \textit{PopGenReport} package v.3.0.4 (Adamack and Gruber, 2014; Gruber and Adamack, 2015). Effective population size ($Ne$) was estimated with NeEstimator v.2.1 (Waples and Do, 2008; Do et al., 2014) using the bias-corrected linkage disequilibrium method (Waples, 2006), random mating, and an allele frequency threshold of $\geq 0.02$ for $Ne_{LD}$ calculation.

Population structure

Pairwise estimates of Weir and Cockerham’s (1984) unbiased $F_{ST}$ were calculated between populations using \texttt{GENODIVE} 2.0b2.5 (Meirmans and van Tienderen, 2004). The neutral dataset was processed in \texttt{STRUCTURE} v.2.3.4 (Pritchard et al., 2000; Falush et al., 2003), which uses a Bayesian clustering algorithm to assign individuals to genetic clusters ($K$). The analysis was comprised of 20,000 burn-in iterations followed by 50,000 replicates of each population value ($K = 1-10$), and each run was conducted 10 times (Schweizer et al., 2016). To determine the optimal value of $K$, the $\Delta K$ statistic (Evanno et al., 2005) was evaluated using \texttt{STRUCTURE HARVESTER} v.0.6.94 (Earl and vonHoldt, 2012). The greedy method and 1,000 random permutations were used in \texttt{CLUMPP} v.1.1.2 to account for variation in cluster assignment across \texttt{STRUCTURE} runs. Bar charts displaying the proportion of cluster membership for each individual were created and modified with \texttt{DISTRUCT} v.1.1 (Rosenberg, 2004). Population structure was also investigated using Discriminant Analysis of Principal Components (DAPC) and Bayesian clustering in the \texttt{R} package \texttt{adegenet} v.2.0.1.
DAPC is useful to avoid a priori assignment and corroborate results of \textit{structure} as it provides a non-model-based method to estimate cluster assignment of individuals. The \textit{find.clusters} function was run in \textit{adegenet} to transform allele frequencies and determine the optimal number of clusters ($K = 1$-10) via \textit{k}-means clustering of principal components and the Bayesian Information Criterion (BIC) (Jombart et al., 2010).

Hierarchical partitioning of genetic variance was evaluated with an analysis of molecular variance (AMOVA) (Excoffier et al., 1992) using the \textit{poppr.amova} function with 99 permutations in the \textit{poppr} package v.2.8.0 (Kamvar et al., 2014, 2015). To determine gene flow patterns in our study area, the \textit{divMigrate} function of the \textit{diveRsity} package v.1.9.90 (Keenan et al., 2013) was used to estimate asymmetric gene flow in relation to contemporary levels of genetic diversity between populations. Nei’s $G_{ST}$ method was used to calculate values of relative directional migration between sampled sites and investigate source and sink dynamics. To test whether migration between populations was significantly asymmetrical, 1,000 bootstrap replicates were performed to calculate 95% confidence intervals (Sundqvist et al., 2016).

\textbf{Demographic and environmental niche modeling}

An environmental niche model (ENM) was used to create environmental suitability maps, to test various demographic models, and predict the geographical distribution of \textit{E. vaginatum} suitable habitats past and present within the range of the Beringian refugium. To build an ENM, the species distribution was modeled for the present and projected into climate scenarios of the mid-Holocene (~6 kyr BP), and the LGM (~20 kyr BP). The initial ENM was built with 19 bioclimatic variables obtained from the WorldClim 2.0 Bioclimatic database (Fick and Hijmans, 2017) and derived from climatic records from 1970 to 2000 with the maximum entropy algorithm
of MaxEnt v.3.4.3 (Phillips et al., 2006; Phillips and Dudík, 2008). MaxEnt uses species occurrence data and predictor variables to predict probability distribution of a species. WorldClim data were downloaded at 30 arc seconds spatial resolution. *Eriophorum vaginatum* species occurrence data were obtained from the Alaska Vegetation Plots Database (Nawrocki et al., 2020), adding 1,228 occurrences to the 17 sampled sites used in this study, for a total of 1,245 unique species occurrence records.

*ENMEVAL* v.2.0.2 (Muscarella et al., 2014) was used to determine optimal parameters, feature class (FC), and regularization multiplier (RM) settings for MaxEnt. A total of 64 models were created using eight FC combinations (L, LQ, LQP, H, T, LQH, LQHP, LQHPT in which L = linear, Q = quadratic, H = hinge, P = product and T = threshold) (Muscarella et al., 2014) and eight RMs (0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0) (Manzoor et al., 2020). The “block” method was implemented to partition data into calibration and evaluation datasets in order to evaluate model performance (Muscarella et al., 2014; González-Serna et al., 2019). The model with the lowest Akaike Information Criterion corrected for small sample size (AICc; Burnham and Anderson, 2002) was identified as the optimal model and implemented for the final run in MaxEnt. To obtain a subset of uncorrelated environmental variables, a Pearson correlation coefficient cutoff ($r \geq 0.75$ and $r \leq -0.75$) was applied to remove variables for the final MaxEnt run (Dormann et al., 2013; Manzoor et al., 2020). Modeled LGM layers for retained environmental variables were derived from the Community Climate System Model (CCSM) (Braconnot et al., 2007) and included as a projection layer in MaxEnt. *ARCMap* v.10.7.1 (ESRI, 2011) was used to classify cell values (with a continuous range of 0 to 1) of current and LGM rasters into 20 equal interval habitat suitability bins (González-Serna et al., 2019) to reduce file size and speed up processing. A Mid-Holocene climate raster was created by averaging current and LGM layer habitat suitability bin values.
Respective output raster files for current environmental conditions were obtained for use in 
CIRCUISTSCAPE (McRae, 2006).

To examine the landscape resistance effects of different vegetation classes and canopy 
cover on gene flow, 30 m categorical land cover data were obtained from the National Aeronautics 
and Space Administration Arctic-Boreal Vulnerability Experiment (NASA ABoVE) ABoVE: 
Landsat-derived Annual Dominant Land Cover Across ABoVE Core Domain, 1984-2014 dataset 
(Wang et al., 2019). The 15-class system used for this data was condensed into four categories: 
forest, woodland, shrubland, and other. Forest included Evergreen Forest, Deciduous Forest and 
Mixed Forest classes with woody vegetation > 3 m tall and > 60% canopy coverage. The woodland 
category also included vegetation > 3 m tall, but with 30-60% canopy coverage. Shrubland 
comprised Low Shrub, Tall Shrub and Open Shrub classes with woody vegetation between 5 cm 
and 3 m tall and with 30-60% canopy coverage. Other included Herbaceous, Tussock Tundra, 
Sparsely Vegetated, Fen, Bog, Shallow/littoral, Barren, and Water classes, in addition to areas with 
missing data, which are likely covered by snow or ice. The dataset included 31 raster bands of land 
cover data, each corresponding to a year during 1984-2014 of land cover classification. To 
determine and implement the most frequent classification during this time series, a per pixel 
summarization was performed across the bands. Using ArcMap, the original 30 m resolution 
dataset was resampled to 60 m to decrease processing requirements during runs in CIRCUISTSCAPE.

A series of models were designed to examine the effects of isolation by distance (IBD), 
resistance (IBR), and environment (IBE) on genetic distance and run in CIRCUISTSCAPE following Van 
Strien et al., (2012) and Emel et al., (2020). For wind-pollinated and wind-dispersed E. vaginatum, 
resistance due to forest land cover in the Arctic landscape is expected to be a primary physical 
factor limiting gene flow. As the extent to which land cover types affect gene flow is unknown,
we tested a series of resistance surfaces for Forest, Woodland, and Shrubland categories, assigning combinations of low resistance (2 or 5), medium resistance (10, 20, or 50), and high resistance (100 or 500). Landscape classes in the Other category were assigned a resistance of 1, or lack of resistance. An ENM model based on habitat suitability using MaxEnt output was used to examine IBE. To create a null model, all cells were assigned a resistance cost value of 1, equivalent to an IBD scenario. Pairwise landscape resistance matrices were calculated between sites for 11 different models.

To determine which model best reflected gene flow across sites, we used maximum likelihood of population effects (MLPE) models to test the IBD, IBE, or IBR models with pairwise $F_{ST}$ genetic distances as response variables and pairwise resistance matrices as explanatory variables. MLPE models implement a type of linear regression on distance matrices while accounting for random effects of pairwise data (Clarke et al., 2002). MLPE models were utilized with modified lmer models in the R package lme4 v.1.1.27.1 (Bates et al., 2015). Best supported models were identified using AICc. The r.squaredGLMM function in the MUMIN R package v.1.43.17 (Barton, 2018) was used to calculate marginal $R^2$ values and provide a measure of goodness-of-fit.

Genotype-environment association

As geographic and landscape variables often influence patterns of genomic variation and spatial distribution of plants (Manel et al., 2003; Sork et al., 2013; Lind et al., 2017), genotype-environment association (GEA) methods were implemented to investigate these associations using the comprehensive data set with missing data imputed based on averaged allele frequencies for each site. Nineteen environmental predictors that are important for ecological models in the Arctic
(Pearson et al., 2013) were obtained from the WorldClim 2.0 Bioclimatic database (Fick and Hijmans, 2017). We used averages from 1970 to 2000 for GPS locations of sampling sites to distinguish climatic attributes of populations. Global multi-resolution terrain elevation data (GTMED2010; Danielson and Gesch, 2011) was obtained from Google Earth Engine (Gorelick et al., 2017) based on GPS coordinates. In addition, Moran’s eigenvector maps (MEMs) were calculated based on linear distances among sites with the \texttt{adespatial v.0.0-7} (Dray et al., 2016) and \texttt{spdep v.0.6-9} (Bivand et al., 2013) packages. MEMs were included as predictor variables if their Moran’s I value was associated with significant (100 permutations, $p < 0.05$) allele frequency variance.

Due to high correlation ($R^2 > 0.7$) of predictor variables (see Results), a principal components analysis (PCA) was conducted with the \texttt{prcomp} function of the \texttt{stats} package v.3.6.0 (R Core Team, 2013) to estimate the influence of predictors on principal component (PC) axes and to reduce variables. Two environmental PCAs were run that included temperature and precipitation, which were grouped into two subsets of 11 temperature variables and 8 precipitation variables (\textbf{Supplementary Table S2}). The first two axes were retained for both PCAs, following the Kaiser-Guttman criterion (Guttman, 1954). If several predictor variables strongly influence PC1 and PC2 axes in both PCAs, which can make interpretations of results difficult if environmental variables are summarized (Rellstab et al., 2015), an approach often taken in PCA (Brauer et al., 2018; Wellband et al., 2019). Predictor variables with $R^2 > 0.75$ and $< -0.75$ (Schweizer et al., 2016; Forester et al., 2018) and VIFs $> 20$ were removed (Ter Braak, 1988; Pienitz et al., 1995).

The \texttt{vegan} v.2.5.6 (Oksanen et al., 2013) and \texttt{psych} v.1.8.12 (Revelle, 2014) packages were used to run a redundancy analysis (RDA) to investigate co-variation of alleles in response to
environmental predictors. Although candidate outlier loci analyses using ddRAD are incomplete as they only sample a fraction of the genome (Lowry et al., 2017), they can be informative for identifying some genes that may have a role in adaptation. The relatively small genome of *E. vaginatum* (1C ≈ 0.4 pg; Rewers et al., 2012) also provides for a higher probability of capturing candidate outlier loci. Site-based allele frequencies used for RDA were computed with *adegenet* v.2.0.1. The RDA was then run using the reduced predictor variable data set (n = 7; see Results). Significance of the final models and constrained axes were identified with 999 permutations and a *p*-value of 0.05 (Forester et al., 2018). SNPs that loaded ±2 SD from the mean loading of significant RDA axes were recognized as candidate outlier loci to retain as many potential candidate loci as possible and to discern between neutral loci that share similar spatial signatures to outliers (Forester et al., 2018). The strongest correlations between candidate SNPs and variables were then identified based on the highest correlation coefficients (Forester et al., 2018). Candidate loci found with RDA were investigated using the *blastn* function in *BLAST* (Altschul et al., 1990), and queried against the NCBI non-redundant nucleotide database and the transcriptome of *E. vaginatum* (Mohl et al., 2020). Subsequently, candidate loci related to local adaptation were annotated with the transcriptome to find the associated Gene Ontology term for each gene based on a percentage identity match of at least 80.0 and an *E*-value threshold of at least $1 \times 10^{-4}$.

**RESULTS**

*Genomic sequence data and genomic diversity*

After removal of low-quality reads and reads without radtags, a total of 546,642,395 single-end sequences were retained for 273 individuals, with an average of 2,002,353 sequencing reads per individual. All samples had a coverage depth $> 13 \times$. The comprehensive SNP data set
contained 3,879 loci and 10,734 SNPs. Twenty-one outlier loci, each with a single SNP, were identified using BayeScan, and these were removed to create a putatively neutral dataset. The neutral data set was comprised of 2,776 loci, each represented by a single SNP, after filtering out SNPs deviating from HWE ($p \leq 0.001$).

For the neutral SNP data set, global population estimates revealed a mean $F_{IS} = -0.003$ (SE $= \pm 0.001$). $H_o$ (mean $= 0.173$; SE $= \pm 0.0008$) and $H_e$ (mean $= 0.173$; SE $= \pm 0.0008$) varied little among sites (Table 2). Allelic richness (Ar) ranged from 1.590 to 1.667 among sites, with one private allele (in EC) identified (Table 2). $N_e$ varied from 39.9 to $\infty$ across sites, but was generally high except near treeline and for isolated sites in the south.

Population structure

On average, we attained a global $F_{ST}$ of 0.020, with pairwise values generally being lower among sites above treeline (0.002-0.010; mean = 0.006) than below treeline (excluding EC; 0.004-0.026; mean = 0.014). Eagle Creek had relatively high $F_{ST}$ values in all pairwise comparisons (0.034 to 0.060; mean = 0.047). While all pairwise comparisons were significant ($p < 0.01$) between sites below treeline, this was not true between some sites above treeline (Supplementary Table S3). AMOVA results demonstrated that a high percentage of neutral genetic variation occurred within sites (94.9%), however, variation was significant among all comparisons (Supplementary Table S4).

Highest $divMigrate G_{ST}$ migration values ($\geq 0.80$) were found between sites north of treeline (CH, AT, TL, AN, SG, CP, PB) and between CC, EL, NN, GO, and CF south of treeline (Figure 3; Supplementary Table S5), with a break at treeline. Migration values were lower for sites sampled adjacent to the treeline boundary and among more isolated higher elevation sites.
(EC, NC, and VM) south of treeline. Migration values from SG into other northern sites were consistently higher than from other sites (Figure 3; Supplementary Table S5). Asymmetric gene flow was found between most sites, but values were not significant.

The best $\Delta K$ value resulting from STRUCTURE analyses was $K = 2$ (Supplementary Figure S1A). However, the BIC score for $K = 3$ from DAPC analyses was optimal (Supplementary Figure S1B). Populations sampled north or south of treeline formed clusters at both $K = 2$ or $K = 3$ for DAPC and STRUCTURE analyses (Figure 4). At $K = 3$, EC was identified as a unique cluster in the DAPC scatter plot and in barplots of STRUCTURE results. $\Delta K$ methods, in particular, have been shown to bias towards $K = 2$ for populations with subtle population structure and in simulated data when $K = 1$ or $K = 3$ were supported for simulations (Cullingham et al., 2020). In this case, EC was also differentiated by high $F_{ST}$ values in pairwise comparisons and low $G_{ST}$ based migration values compared to other populations. Populations on either side of treeline (ST and TB) showed evidence of admixture based on both STRUCTURE and DAPC results (Figure 4).

Demographic and environmental niche modeling

The contemporary species distribution model and climate scenario projections from the LGM (~20 kyr BP) identified suitable habitat throughout the Beringian refugium for *E. vaginatum* (Figure 5) outside of the glaciated regions within the refugium (notably including north and south of the Brooks Range). Projection from the mid-Holocene (~6 kyr BP) shows suitable habitat extended throughout most of northern Alaska and was continuous across the formerly glaciated Brooks Range. MLPE models incorporating increased resistance for forest, woodland, and shrubland supported IBR as an important factor in shaping neutral genetic structure across sites.
The best MLPE models favored higher forest resistance compared to other categories, and higher resistance from woodland than shrubland was also important. The best IBR models explained ~21% of the genetic variation, which was ~17% more than the variation explained by the null IBD model (Table 3). Including IBE (ENM) did not improve the best IBR models, but IBE alone and models incorporating IBE also outperformed the IBD model in all measures.

**Genotype-environment association**

Two environmental PCAs were run: 1) temperature and 2) precipitation. PC1 and PC2 explained 90.5% of the total variation of the temperature PCA and 96.0% of the precipitation PCA. All predictors strongly influenced both PC1 and PC2, so environmental predictors were not grouped into summarized environmental variables. The retained variables explaining significant allelic variation for the RDA analysis were: 1) isothermality (iso; annual mean diurnal range/annual temperature range), 2) mean temperature of the driest quarter (tdq; Feb - Apr), 3) mean temperature of the warmest quarter (twq; Jun - Aug), 4) precipitation of the driest month (pdm; Apr), 5) precipitation seasonality (prs; coefficient of variation estimated from the standard deviation of monthly precipitation estimates), 6) MEM1 and 7) MEM2 (Supplementary Table S6).

The RDA model was significant ($p = 0.001$), and predictor variables explained 13.2% of the total genetic variance. The RDA1 discriminant axis was significant ($p = 0.001$), and explained 30.1% of the constrained variation. The RDA2 axis was not significant ($p = 0.070$). There were 165 candidate SNPs identified and detected on the RDA1 axis and 162 candidate SNPs with high correlations ($R^2 \geq 0.7$) to predictor variables, which were retained to further investigate local adaptation (Table 4, Supplementary Table S7). Of the candidate SNPs with correlations $R^2 \geq 0.7$
to predictor variables, most were correlated with the MEM1 variable (141 SNPs). The remaining candidate SNPs were correlated with tdq (11 SNPs) and twq (10 SNPs). Gene Ontology annotations were found for 45 candidate loci after a search against the *E. vaginatum* transcriptome (Mohl et al., 2020, Table 4, Supplementary Table S7).

**DISCUSSION**

*Landscape genomics of an arctic foundation species*

Three genetic clusters were identified for sampled *E. vaginatum* (Figure 4; EC, North, and South) across a latitudinal gradient encompassing two ecosystems in the north central Alaskan Arctic. Most populations genetically clustered by geographical locations north or south of treeline (Figures 1, 4; Table 1) with the exception of treeline adjacent populations (ST and TB), which had a near equal genetic assignment to both north or south genetic clusters, and EC, which formed its own genetic cluster. Importantly, population structure results support the presence of glacial barriers to gene flow for *E. vaginatum*, and corroborates long-term studies finding adaptations shared among plants north versus south of the Brooks Range (Fetcher and Shaver, 1990; Bennington et al., 2012).

Post-glaciation, the Brooks Range could have provided an allopatric barrier for *E. vaginatum* colonization. However, our landscape-level sampling suggests the gene flow barrier is at treeline, well below the summit of the range. The Continental Divide defines a north and south slope through the Brooks Range and populations above treeline from the south slope of the Brooks Range (i.e., CH) clusters with north slope populations while south slope treeline adjacent populations ST (south of treeline) and TB (north of treeline) show evidence of an admixture zone only at treeline. While treeline in the Arctic signifies a contemporary ecosystem change, the arctic
flora was also shaped in part by patterns of glaciation and distribution of refugia. The Beringian refugium was an important source for post-glaciation vegetation, but in Alaska it was divided by glaciation over the Brooks Range (Figure 1) that has likely affected the genetic structure of the Alaskan arctic flora. Here, we propose a hypothesis for the contemporary genetic structure of *E. vaginatum* that could have implications for the adaptations found in long-term studies of plants in our transect. Two subrefugia were present north and south of the Brooks Range glaciation (Figure 5), with each population accumulating genetic variation in allopatry via genetic drift and environmental adaptation; this was followed by post-glaciation advancement. Both geophysical attributes and climatic events in this region of Alaska support the idea of two centers of origin for *E. vaginatum*. The Brooks Range remained glaciated throughout the Pleistocene glacial cycles (Figure 1; Kaufman and Manley, 2004; Kaufman et al., 2011), creating a formidable barrier to gene flow between northern and southern regions of the eastern Beringian refugium. Pollen of *E. vaginatum* has been found in Yedoma sediment samples, indicating its presence in both regions during paleoclimate fluctuations encompassing the LGM (Schirrmeister et al., 2016; Lapointe et al., 2017). Likewise, pollen profiles in central and northern Alaska during the last ~36 kyr suggest a flora in the region similar to that found today, with dominance of graminoids in the north and spruce woodlands in the south (Billings, 1992; Finkenbinder et al., 2014; Schirrmeister et al., 2016; Lapointe et al., 2017). Demographic analyses suggest that at the glacial maximum environmental conditions in the two regions were suitable for *E. vaginatum* as well (Figure 5). As such, *E. vaginatum* (along with the rest of the flora) had the opportunity to adapt to contrasting abiotic and biotic factors reflecting the contemporary ecosystems that now have a transition zone at the treeline ecotone.
Adaptive variation is supported by GEA analyses that indicate environmental predictors (Figure 6) contribute to allelic turnover with a shift related to changes at treeline (Supplementary Table S6), corroborating evidence from long-term ecological studies (e.g. Fetcher and Shaver, 1990; Bennington et al., 2012). Genetic structure would also be reinforced through neutral variation accumulated during the period of allopatry caused by glacial isolation surrounding the LGM. IBR based on vegetation cover, especially denser forest, also has a large effect on contemporary gene flow and probably has a large effect on the genetic structure we found here. A particularly important consequence of IBR due to forest density is its role in preventing future gene flow north, as potentially better adapted ecotypes from the south will have restrained capacity to migrate north. In summary, both geographic isolation and adaptation influenced the genetic structure that is now delimited at treeline, as supported by demographic modeling analyses, which identified that both IBR and IBE had a more significant role in restricting gene flow than IBD (Table 3).

Migration patterns also support bidirectional gene flow from the proposed subrefugia. The divMigrate analyses indicate gene flow to the southernmost site above treeline (i.e., CH) on the south slope of the Brooks Range coming from SG, which is located just beyond the northern extent of the LGM (Figures 1, 3). Between June and July, when *E. vaginatum* flowers and fruits, prevailing daytime winds are from the north in this region (Zhang et al., 2016; Environmental Data Center Team, 2022) providing a mechanism for gene flow over the Continental Divide. SG also is the northern population with highest gene flow to all other northern sites, indicating its likely importance as a source population post-glaciation. It is also clear that connectivity and gene flow is high throughout the north slope and coastal plain (Supplementary Table S3), probably due to the lack of the primary IBR variable of forest cover. These results suggest that this region could to
some extent be treated as a large population with high genetic connectivity. The highest migration values among South populations are from sites just south of treeline, including GO, the northernmost site beyond the southern glacial boundary at the LGM. Thus, potential source populations on either side of the Brooks Range glaciation likely had a role in colonizing to treeline where gene flow was restricted. Further, logistically challenging sampling along a longitudinal transect above treeline and along the south slope of the Brooks Range would further verify the geographic extent of these results. The denser forest cover and more disjunct *E. vaginatum* population along the southern latitudinal gradient would lead to the lower genetic connectivity of these populations. At the treeline ecotone migration values were low and the two sites sampled at the treeline boundary (i.e., ST and TB) had similar migration values from North and South clusters (*Supplementary Table S5*). If gene flow proceeded from subrefugia rather than an ancestral gene pool, then the admixed samples at treeline identified in population structure analyses (*Figure 4*) signify an admixture zone along the treeline ecotone, which further supports treeline as a gene flow barrier. Our modeling of IBR suggests that land-cover resistance continues to have a strong effect on gene flow.

Alternatively, the environmental shift accompanying the moving treeline could have led to rapid evolution during post-glacial expansion if seed banks from widespread species like *E. vaginatum*, that innately carry high allelic diversity, are already present in regions with retreating glaciers (Alsos et al., 2007); and thus, provide the standing genetic variation for rapid adaptive change (Hermisson and Pennings, 2005; Dlugosch et al., 2015). Although such a scenario has been hypothesized for *E. vaginatum* in alpine environments (Walker et al., 2019), the known geological and vegetation history of the region, along with *divMigrate* gene flow patterns (*Figure 3*) and evidence of population structure (*Figure 4*) identified here makes this alternative unlikely.
Much of the Arctic has been revegetated relatively recently from refugia, which have typically been represented as a genetic pool without considering adaptive variation within each refugium (Tremblay and Schoen, 1999; Abbott et al., 2000; Alsos et al., 2005, 2007). If *E. vaginatum* from northern and southern Beringian subrefugia represent adaptive ecotypes, as suggested by the long-term ecological studies (Bennington et al., 2012) and molecular data (Mohl et al., 2020; and these results), that colonized formerly glaciated regions, then we would expect them to have more likely colonized ecosystems to which they are best adapted. For example, *E. vaginatum* populations from north of treeline have lower phenotypic plasticity (Fetcher and Shaver, 1990; Parker et al., 2017) and are less responsive in gross primary productivity (Curasi et al., 2019) compared to those south of treeline, which could affect persistence of each ecotype as the Arctic warms. Similarly, we provide evidence of allelic turnover in GEA analyses for populations of *E. vaginatum*, suggesting that populations in these two regions are not simply genetically differentiated due to genetic drift in isolation, but have adapted to unique genetic niche space. Together, the data demonstrates the importance of determining the extent and cause for genetic divergence between populations, which is particularly critical when developing species distribution models for climate change (Massatti and Knowles, 2016; Ortego and Knowles, 2020) that include foundation species such as *E. vaginatum*. Attempts to verify the distribution of ecotypes and the adaptive potential of *E. vaginatum* will require expanding landscape genomic studies across entire Arctic ecosystems including refugial regions, and building on the ground-breaking genomic biodiversity work of Wang et al. (2021).

Finally, the site EC, occurring in the southern extent of the sampling range, was identified as genetically distinct (Figure 4), which was unexpected given the similarity in ecological attributes of EC and other populations south of treeline (Fetcher and Shaver, 1990; Bennington et
Due to higher landscape resistance south of treeline, populations of *E. vaginatum* are more sparsely distributed with lower genetic connectivity (Figure 3, Supplementary Table S5). The EC site is at the southern margin of the modeled LGM extent of the Beringian refugium (Figure 1; Kaufman and Manley, 2004; Kaufman et al., 2011). Given the population disjunction and potential for isolation due to glaciation patterns, genetic differentiation of EC could be due to bottlenecking or a founder event as supported by this population having the lowest allelic richness and low observed heterozygosity, as compared to the other sampled sites (Table 2). However, effective population size for EC was relatively high (*Ne* = 5923.3) and it is not uniquely isolated among southern populations. Thus, we alternatively posit that EC may represent a population adapted to alternative niche space as supported by GEA analyses that uncovered association between allelic turnover and precipitation variables (iso and prs) for this population relative to others (Figures 3B, 6, Table 4). While our GEA analyses didn’t uncover specific genes correlated to iso and prs, Mohl et al. (2020) found that EC had differential expression of abiotic stress genes not found in other populations along the latitudinal distribution. Future work will benefit from examining southern populations with similar precipitation patterns to further elucidate the importance of adaptive variation for this and other similar populations that may have been uniquely isolated during glacial periods.

**Environmental heterogeneity explains allelic turnover across a latitudinal cline**

The Arctic covers roughly 7% of the Earth’s surface area or 14% of its land area, and its vegetation history has been shaped by recurring patterns of glaciation, geological landscape structure, and environmental shifts in attributes such as permafrost depth (Billings, 1973; Meyerhoff and Meyerhoff, 1973). Likewise, the transitions associated with treeline in the Arctic
aren’t necessarily abrupt; instead there is a cline that could require adaptations that can range from local specificity to wide-ranging plasticity (Fetcher and Shaver, 1990; Bennington et al., 2012; Curasi et al., 2019). This has been exemplified in long-term reciprocal transplant studies that have shown adaptive differences in *E. vaginatum* are complex and related not only to position north or south of treeline, but also to home site and latitude (Shaver et al., 1986; Fetcher and Shaver, 1990; Bennington et al., 2012; Curasi et al., 2019; Mohl et al., 2020). As expected, a genetic signature was recovered related to environment along the cline of the latitudinal gradient, not just at treeline. Evidence for environmental association with allele frequency turnover was found for all predictors in the RDA: MEM1, temperature, and precipitation (Figure 6). RDA1 was correlated with most environmental variables considered and corresponded with the distribution of populations along the latitudinal gradient.

The primary temperature predictors (tdq and twq) relate to spring and summer conditions, which encompass the limited growing season of *E. vaginatum* in the Arctic and show a clinal change in both mean and variance along the latitudinal gradient (Supplementary Table S1). The short growing season for most plants in the Arctic begins in late spring following snowmelt, as air temperatures and photosynthetic rates increase. Photosynthetic rate then remains relatively high from mid-June to mid-August (Defoliart et al., 1988). Consequently, late spring and summer temperature differences are likely critical for vegetative phenology of *E. vaginatum* (Siegenthaler et al., 2013; Parker et al., 2017, 2021). Similar physiological adaptations have been observed for other arctic plant species, especially graminoids, forbs, and deciduous shrubs (Chapin, 1987; Chapin and Shaver, 1996); thus, selective pressures could lead to similar adaptation for co-occurring widespread arctic species.
Several potentially adaptive SNPs (162 candidate SNPs and 131 loci with high correlations to predictor variables; $R^2 \geq 0.7$) were identified with GEA methods. MEM1, tdq, and twq, in succession, were the predictors for the 45 loci identified to gene and function (Table 4). MEM variables can represent unmeasured environmental predictors, have a spatial component (Manel et al., 2010; Fitzpatrick and Keller, 2015), and provide important predictive value for allele frequency turnover (Martins et al., 2018; Gibson and Moyle, 2020) included among other arctic-alpine plants (Manel et al., 2012; Bothwell et al., 2013). These three variables most frequently correlated with candidate loci that have roles in abiotic stress response (Table 4). These include transcription factors belonging to gene families established in stress response pathways (e.g. LRR, SIZ1, and EIL; Yeh et al., 2012; Liu et al., 2020; Poór et al., 2021). The importance of these genes is predictable due to the need of plants to adjust to extreme fluctuations in temperature in the Low Arctic throughout the growing season (see Supplementary Table S1) and should be the starting point in investigating genes attributing to ecotype adaptation.

In summary, genomic structure and differentiation were identified across 17 *E. vaginatum* sites along a ~426 km latitudinal gradient in north central Alaska, which was within the boundary of the Beringian refugium. A major genetic boundary was found at the treeline ecotone. Strong evidence of IBR was found across the study region, highlighting the influence of vegetation type and density on *E. vaginatum* neutral genetic structure. Importantly, IBR due to forest cover can signify a major obstacle for gene flow of potentially better-adapted genotypes northward with rapid climate change. GEA results support ecotypic adaptations elucidated over decades of ecological studies for this arctic foundation species and suggest potential important environmental variables along with candidate loci, many of which are associated to abiotic stress gene pathways. These results are consequential for increasing predictive accuracy of distribution changes under climate
change. An understanding of adaptive variation should be incorporated into developing hybrid environmental niche and species distribution models (Ikeda et al., 2017a; Razgour et al., 2019; Waldvogel et al., 2020). The combination of broader sampling and genomic studies of other foundational arctic plants will increase understanding of local adaptation, gene flow, and environmental associations crucial for determining the long-term consequences of climate change in the arctic flora. We conclude that accounting for ecotypic physiology, gene flow, local adaptation and gene expression of foundation species under changing climates will lead to a greater understanding of response from the level of the individual to the ecosystem.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession numbers can be found below at: https://www.ncbi.nlm.nih.gov/, BioProject PRJNA803172, accession numbers SAMN25639742-SAMN25640014. The STACKS pipeline Python script, Redundancy Analysis R code, and ENMEVAL R code can be found at https://github.com/estunz/EvLG2022.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

MLM, NF, and JT designed the study. MLM, ES, and NF organized and performed field collections. ES and MLM wrote the manuscript. ES performed lab research and analyzed the data.
JEM contributed bioinformatic analytical tools and training and assisted with data analyses. PL provided ddRAD data collection techniques, lab resources, and assistance. All authors reviewed and approved the final manuscript.

FUNDING

This research was made possible by funding provided by NSF/PLR-1417645 to MLM. The Botanical Society of America Graduate Student Research Award and the Dodson Research Grant from the Graduate School of the University of Texas at El Paso provided assistance to ES. The grant 5U54MD007592 from the National Institute on Minority Health and Health Disparities (NIMHD), a component of the National Institutes of Health (NIH) provided bioinformatics resources and support of JM.

ACKNOWLEDGMENTS

Thanks to Stephen Escarzaga for contributing GIS skills, and Bob Muscarella and Joaquín Ortego for demographic modeling advice. Thank you to Austin Frisbey for field assistance and Tom Parker for collections. We are thankful to Gus Shaver for invaluable help in the field, providing location information for collection sites and critical discussion of the research. We would also like to acknowledge logistic support from Toolik Field Station and Arctic LTER (NSF/PLR-1637459).

REFERENCES


Dahl, E. (1963) “Plant migrations across the North Atlantic Ocean and their importance for the

35


both dominant and codominant markers: a Bayesian perspective. *Genetics* 180, 977-993. doi: 10.1534/genetics.108.092221


Late Quaternary paleoenvironmental records from the Chatanika River valley near Fairbanks (Alaska). *Quat. Sci. Rev.* 147, 259-278. doi: 10.1016/j.quascirev.2016.02.009


Ter Braak, C. J. F. (1988). CANOCO - a FORTRAN program for canonical community ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis (version 2.1). (No. LWA-88-02) MLV.

doi: 10.1046/j.1365-294x.1999.00680.x


Villarreal, S., Hollister, R. D., Johnson, D. R., Lara, M. J., Webber, P. J., and Tweedie, C. E.
015508. doi: 10.1088/1748-9326/7/1/015508

Waldvogel, A. M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R.,
et al. (2020). Evolutionary genomics can improve prediction of species’ responses to

(2019). Plastic and genetic responses of a common sedge to warming have contrasting

ABoVE: Landsat-derived Annual Dominant Land Cover Across ABoVE Core Domain,

Late Quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature*
600, 86-92. doi: 10.1038/s41586-021-04016-x

Waples, R. (2006). A bias correction for estimates of effective population size based on linkage

doi: 10.1111/j.1755-0998.2007.02061.x


doi: 10.1111/mec.14965


doi: 10.1016/j.plantsci.2012.06.004


**Supplementary Material**

The Supplementary material for this article is available online

at: https://www.frontiersin.org/articles/10.3389/fpls.2022.860439/full#supplementary-material
**Tables and Figures**

Table 1. Collection sites, GPS coordinates, elevation (meters), and vegetation type of *E. vaginatum* in north central Alaska. MAT = Moist Acidic Tundra. Genetic cluster assignment based on STRUCTURE analysis at $K = 3$ (Figure 4).

<table>
<thead>
<tr>
<th>Site</th>
<th>Latitude, Longitude</th>
<th>Elevation (m)</th>
<th>Vegetation Type</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eagle Creek (EC)</td>
<td>65.4332°, -145.5118°</td>
<td>771</td>
<td>MAT</td>
<td>EC</td>
</tr>
<tr>
<td>Nome Creek (NC)</td>
<td>65.3646°, -147.0406°</td>
<td>511</td>
<td>Muskeg</td>
<td>South</td>
</tr>
<tr>
<td>Victoria Mountain (VM)</td>
<td>65.7832°, -147.0681°</td>
<td>960</td>
<td>MAT</td>
<td>South</td>
</tr>
<tr>
<td>Colorado Creek (CC)</td>
<td>65.4705°, -148.2666°</td>
<td>192</td>
<td>Muskeg</td>
<td>South</td>
</tr>
<tr>
<td>Elliott Highway (EL)</td>
<td>65.3081°, -149.1230°</td>
<td>720</td>
<td>MAT</td>
<td>South</td>
</tr>
<tr>
<td>No Name Creek (NN)</td>
<td>66.1171°, -150.1676°</td>
<td>167</td>
<td>Tussock bog</td>
<td>South</td>
</tr>
<tr>
<td>Gobbler’s Knob (GO)</td>
<td>66.7459°, -150.6862°</td>
<td>520</td>
<td>Muskeg</td>
<td>South</td>
</tr>
<tr>
<td>Coldfoot (CF)</td>
<td>67.2631°, -150.1591°</td>
<td>321</td>
<td>Muskeg</td>
<td>South</td>
</tr>
<tr>
<td>South of Timberline (ST)</td>
<td>68.0006°, -149.7469°</td>
<td>702</td>
<td>MAT</td>
<td>South/North</td>
</tr>
<tr>
<td>Atigun Camp (AT)</td>
<td>68.0300°, -149.6737°</td>
<td>760</td>
<td>MAT</td>
<td>North/South</td>
</tr>
<tr>
<td>Chandalar (CH)</td>
<td>68.0518°, -149.6115°</td>
<td>968</td>
<td>MAT</td>
<td>North</td>
</tr>
<tr>
<td>Toolik Lake (TL)</td>
<td>68.6292°, -149.5778°</td>
<td>758</td>
<td>MAT</td>
<td>North</td>
</tr>
<tr>
<td>Anaktuvuk (AN)</td>
<td>68.9945°, -150.2871°</td>
<td>341</td>
<td>MAT</td>
<td>North</td>
</tr>
<tr>
<td>Sagwon (SG)</td>
<td>69.4244°, -148.6976°</td>
<td>299</td>
<td>MAT</td>
<td>North</td>
</tr>
<tr>
<td>Coastal Plain (CP)</td>
<td>69.7704°, -150.3128°</td>
<td>173</td>
<td>MAT</td>
<td>North</td>
</tr>
<tr>
<td>Prudhoe Bay (PB)</td>
<td>70.3270°, -149.0645°</td>
<td>8</td>
<td>MAT</td>
<td>North</td>
</tr>
</tbody>
</table>
Table 2. Genetic diversity summary and demographic statistics for the neutral data set of the 17 Eriophorum vaginatum sites sampled in north central Alaska. Sites are ordered geographically from South to North and site abbreviations follow Table 1. Sample size (N), allelic richness (Ar), private alleles (Pa), observed heterozygosity (Ho), expected heterozygosity (He), inbreeding coefficient (FIS) and effective population size (Ne).

<table>
<thead>
<tr>
<th>Site</th>
<th>N</th>
<th>Ar</th>
<th>Pa</th>
<th>Ho</th>
<th>He</th>
<th>FIS</th>
<th>Ne</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC</td>
<td>16</td>
<td>1.590</td>
<td>1</td>
<td>0.171</td>
<td>0.167</td>
<td>-0.025</td>
<td>5923.3</td>
</tr>
<tr>
<td>NC</td>
<td>15</td>
<td>1.634</td>
<td>0</td>
<td>0.174</td>
<td>0.173</td>
<td>0.002</td>
<td>202.0</td>
</tr>
<tr>
<td>VM</td>
<td>15</td>
<td>1.627</td>
<td>0</td>
<td>0.171</td>
<td>0.171</td>
<td>0.004</td>
<td>39.9</td>
</tr>
<tr>
<td>CC</td>
<td>16</td>
<td>1.645</td>
<td>0</td>
<td>0.175</td>
<td>0.174</td>
<td>-0.003</td>
<td>60.9</td>
</tr>
<tr>
<td>EL</td>
<td>16</td>
<td>1.658</td>
<td>0</td>
<td>0.178</td>
<td>0.176</td>
<td>-0.014</td>
<td>∞</td>
</tr>
<tr>
<td>NN</td>
<td>16</td>
<td>1.661</td>
<td>0</td>
<td>0.177</td>
<td>0.178</td>
<td>0.004</td>
<td>∞</td>
</tr>
<tr>
<td>GO</td>
<td>16</td>
<td>1.663</td>
<td>0</td>
<td>0.173</td>
<td>0.177</td>
<td>0.023</td>
<td>∞</td>
</tr>
<tr>
<td>CF</td>
<td>18</td>
<td>1.667</td>
<td>0</td>
<td>0.184</td>
<td>0.177</td>
<td>-0.036</td>
<td>∞</td>
</tr>
<tr>
<td>ST</td>
<td>16</td>
<td>1.645</td>
<td>0</td>
<td>0.170</td>
<td>0.175</td>
<td>0.029</td>
<td>186.4</td>
</tr>
<tr>
<td>TB</td>
<td>16</td>
<td>1.632</td>
<td>0</td>
<td>0.166</td>
<td>0.171</td>
<td>0.028</td>
<td>221.4</td>
</tr>
<tr>
<td>CH</td>
<td>16</td>
<td>1.643</td>
<td>0</td>
<td>0.179</td>
<td>0.175</td>
<td>-0.029</td>
<td>314.5</td>
</tr>
<tr>
<td>AT</td>
<td>16</td>
<td>1.634</td>
<td>0</td>
<td>0.171</td>
<td>0.172</td>
<td>0.011</td>
<td>∞</td>
</tr>
<tr>
<td>TL</td>
<td>16</td>
<td>1.627</td>
<td>0</td>
<td>0.171</td>
<td>0.169</td>
<td>-0.011</td>
<td>∞</td>
</tr>
<tr>
<td>AN</td>
<td>16</td>
<td>1.627</td>
<td>0</td>
<td>0.167</td>
<td>0.169</td>
<td>0.017</td>
<td>∞</td>
</tr>
<tr>
<td>SG</td>
<td>17</td>
<td>1.645</td>
<td>0</td>
<td>0.174</td>
<td>0.173</td>
<td>-0.006</td>
<td>∞</td>
</tr>
<tr>
<td>CP</td>
<td>16</td>
<td>1.656</td>
<td>0</td>
<td>0.172</td>
<td>0.175</td>
<td>0.019</td>
<td>∞</td>
</tr>
<tr>
<td>PB</td>
<td>16</td>
<td>1.653</td>
<td>0</td>
<td>0.173</td>
<td>0.175</td>
<td>0.013</td>
<td>∞</td>
</tr>
</tbody>
</table>

Table 3. Maximum likelihood of population effects (MLPE) models relating pairwise FST to pairwise distance matrices of isolation by resistance (IBR), isolation by environment (IBE), and isolation by distance (IBD) and ranked by AICc. R2m = marginal R2 approximation of mixed model fixed effects.

<table>
<thead>
<tr>
<th>Resistance values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>IBR</td>
</tr>
<tr>
<td>IBR</td>
</tr>
<tr>
<td>IBR</td>
</tr>
<tr>
<td>IBR</td>
</tr>
<tr>
<td>IBR+IBE</td>
</tr>
<tr>
<td>IBR</td>
</tr>
<tr>
<td>IBR+IBE</td>
</tr>
<tr>
<td>IBR+IBE</td>
</tr>
<tr>
<td>IBE</td>
</tr>
<tr>
<td>IBD</td>
</tr>
</tbody>
</table>
Table 4. Annotated *E. vaginatum* candidate genes with a percentage identity match of at least 80.0 and an E-value threshold of at least $1 \times 10^{-4}$ and RDA $R^2$ value $\geq 0.8$. RDA Predictors had the highest $R^2$ values for that association with the given gene. 2' RDA Predictors had $R^2$ values $> 0.7$ and are listed from highest to lowest. E-value and Similarity % for gene ID of each locus. MF = Molecular Function, and BP = Biological Process. *Genes with stress response association.

<table>
<thead>
<tr>
<th>Locus ID</th>
<th>Gene</th>
<th>Protein</th>
<th>GO Functional Term</th>
<th>RDA $R^2$</th>
<th>RDA Predictor</th>
<th>2' RDA Predictor</th>
<th>E-value</th>
<th>Similarity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>42946</td>
<td>EMB8</td>
<td>Embryogenesis-associated EMB8</td>
<td>MF</td>
<td>0.933</td>
<td>MEM1</td>
<td>N/A</td>
<td>1.00E-70</td>
<td>100</td>
</tr>
<tr>
<td>68081</td>
<td>HIPP26*</td>
<td>Heavy metal-associated isoprenylated plant protein 26-like</td>
<td>MF, BP</td>
<td>0.911</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>99</td>
</tr>
<tr>
<td>2666657</td>
<td>VPS37</td>
<td>Vacuolar-sorting-associated 37 homolog 1-like</td>
<td>MF, BP</td>
<td>0.907</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>27083</td>
<td>LRR*</td>
<td>Probable LRR receptor-like serine/threonine-protein kinase</td>
<td>MF</td>
<td>0.906</td>
<td>tdq</td>
<td>MEM1</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>53674</td>
<td>TMEM33</td>
<td>Transmembrane 33 homolog</td>
<td>BP</td>
<td>0.898</td>
<td>MEM1</td>
<td>N/A</td>
<td>1.00E-21</td>
<td>91</td>
</tr>
<tr>
<td>57279</td>
<td>pKIW1502</td>
<td>fruit protein pKIW1502-like</td>
<td>MF</td>
<td>0.897</td>
<td>MEM1</td>
<td>N/A</td>
<td>2.00E-69</td>
<td>99</td>
</tr>
<tr>
<td>61802</td>
<td>CEK2*</td>
<td>Probable choline kinase 2</td>
<td>MF</td>
<td>0.893</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>47488</td>
<td>ASAT</td>
<td>Putative aspartate aminotransferase</td>
<td>MF, BP</td>
<td>0.891</td>
<td>tdq</td>
<td>MEM1</td>
<td>3.00E-77</td>
<td>99</td>
</tr>
<tr>
<td>1469</td>
<td>ADH2*</td>
<td>Alcohol dehydrogenase class-3</td>
<td>MF, BP</td>
<td>0.874</td>
<td>MEM1</td>
<td>N/A</td>
<td>8.00E-68</td>
<td>99</td>
</tr>
<tr>
<td>3039887</td>
<td>GRF2</td>
<td>14-3-3-like protein GF14 omega</td>
<td>BP</td>
<td>0.871</td>
<td>MEM1</td>
<td>N/A</td>
<td>1.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>63730</td>
<td>At2g45590*</td>
<td>Receptor-like serine threonine-kinase At2g45590</td>
<td>MF</td>
<td>0.865</td>
<td>MEM1</td>
<td>N/A</td>
<td>2.00E-69</td>
<td>99</td>
</tr>
<tr>
<td>11498</td>
<td>CBP</td>
<td>Serine carboxypeptidase</td>
<td>MF</td>
<td>0.847</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>33151</td>
<td>SIZ1*</td>
<td>E3 SUMO-ligase SIZ1 isoform X1</td>
<td>MF</td>
<td>0.846</td>
<td>twq</td>
<td>MEM1</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>38057</td>
<td>RPC3</td>
<td>DNA-directed RNA polymerase III subunit RPC3</td>
<td>MF, BP</td>
<td>0.845</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>1151144</td>
<td>POLD2</td>
<td>DNA polymerase delta small subunit</td>
<td>MF, BP</td>
<td>0.841</td>
<td>MEM1</td>
<td>tdq</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Gene Symbol</td>
<td>Description</td>
<td>MF, BP</td>
<td>MEM1</td>
<td>N/A</td>
<td>p-Value</td>
<td>FDR</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>-------------</td>
<td>--------</td>
<td>------</td>
<td>-----</td>
<td>---------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>38483</td>
<td>PISD*</td>
<td>Phosphatidylserine decarboxylase proenzyme 1, mitochondrial-like</td>
<td>MF, BP</td>
<td>0.837</td>
<td>MEM1</td>
<td>N/A</td>
<td>2.00E-69</td>
<td>99</td>
</tr>
<tr>
<td>4094</td>
<td>EIL1A*</td>
<td>Ethylene Insensitive 3-like 1</td>
<td>MF, BP</td>
<td>0.829</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>26505</td>
<td>FC2*</td>
<td>Ferrochelatase-2, chloroplastic</td>
<td>MF, BP</td>
<td>0.827</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>45291</td>
<td>ABCB*</td>
<td>ABC transporter B family member 13-like isoform</td>
<td>MF</td>
<td>0.823</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>4623</td>
<td>VIP6</td>
<td>CTR9 homolog</td>
<td>MF, BP</td>
<td>0.823</td>
<td>MEM1</td>
<td>N/A</td>
<td>1.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>47238</td>
<td>AHK*</td>
<td>Histidine kinase</td>
<td>MF, BP</td>
<td>0.821</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>1847</td>
<td>CRR7</td>
<td>Chlororespiratory Reduction 7, chloroplastic</td>
<td>MF, BP</td>
<td>0.811</td>
<td>MEM1</td>
<td>N/A</td>
<td>3.00E-07</td>
<td>97</td>
</tr>
<tr>
<td>69795</td>
<td>ZBED1</td>
<td>BED zinc finger, hAT family dimerization domain isoform 1</td>
<td>MF, BP</td>
<td>0.811</td>
<td>MEM1</td>
<td>N/A</td>
<td>3.00E-27</td>
<td>92</td>
</tr>
<tr>
<td>45293</td>
<td>UBP27*</td>
<td>Ubiquitin carboxyl-terminal hydrolase 27 isoform</td>
<td>MF, BP</td>
<td>0.810</td>
<td>MEM1</td>
<td>N/A</td>
<td>4.00E-71</td>
<td>100</td>
</tr>
<tr>
<td>70778</td>
<td>ACO1</td>
<td>1-aminocyclopropane-1-carboxylate oxidase homolog 1-like</td>
<td>MF</td>
<td>0.808</td>
<td>MEM1</td>
<td>N/A</td>
<td>2.00E-68</td>
<td>98</td>
</tr>
</tbody>
</table>
Figure 1. Map of *Eriophorum vaginatum* sampling locations and modeled maximum Pleistocene glacial extent (Kaufman and Manley, 2004; Kaufman et al., 2011) along a latitudinal gradient in north central Alaska. Blue stars designate reciprocal transplant gardens (Shaver et al. 1986; Bennington et al. 2012). Treeline is indicated by the dashed black line and The Continental Divide is indicated by the burnt orange line. The inset shows the extent of the Beringian region, outlined with a dashed yellow line. The 17 collection site abbreviations are as for Table 1.
Figure 2. Images showing (A) the habitat of the Coldfoot sampling location south of treeline, (B) a mature *Eriophorum vaginatum* tussock at Coldfoot, and (C) the habitat of the Prudhoe Bay sampling location in north central Alaska. All photos by E. Stunz.
Figure 3. Maps of the study area with (A) mean temperature of the driest quarter (tdq) and (B) precipitation seasonality (prs) WorldClim underlying data layers using ArcMap v.10.7.1 (ESRI, 2011). Note that divMigrate migration values ≥ 0.80 and direction indicated with arrows are also provided in 2A. Site abbreviations as for Table 1.
Figure 4. Population structure results from the neutral SNP data set. (A) Scatter plot from the discriminant analysis of principal components (DAPC) with 80 principal components retained on the first two Discriminant Analysis axes showing the differentiation between the three groups and inertia ellipses. Each color represents a cluster as identified with the Bayesian Information Criterion (BIC). Red dots represent individuals from the Eagle Creek (EC) population, yellow dots represent individuals from the population south of treeline except for EC (South), and blue dots represent individuals from populations north of treeline (North). (B) Bar graph of STRUCTURE results for population structure analysis. Each vertical bar represents an individual, and colors show the proportion of ancestry assigned to each of the three clusters ($K = 2$ and $K = 3$), as inferred from $\Delta K$ values. *Eriophorum vaginatum* populations are ordered from south to north location along the latitudinal gradient in north central Alaska. See Table 1 for collection site abbreviations.
Figure 5. MaxEnt environmental niche model (ENM) maps of Alaska, U.S.A. adapted to depict *Eriophorum vaginatum* habitat suitability during the Last Glacial Maximum (LGM), Mid-Holocene and present. Modeled LGM layers were derived from the Community Climate System Model (CCSM) (Braconnot et al., 2007). Current and LGM layer habitat suitability bin values were averaged to create the Mid-Holocene climate raster.
Figure 6. RDA plot demonstrating predictor associations as related to ecotype for the comprehensive SNP data set. iso = Isothermality, tdq = Mean temperature of the driest quarter (Feb, Mar, Apr), twq = Mean temperature of the warmest quarter (Jun, Jul, Aug), prd = Precipitation of the driest month (Apr), prs = Precipitation seasonality.
Chapter 2: Population structure, demographics, and local adaptation of the arctic dwarf birch (*Betula nana*)
ABSTRACT

Increase in shrub cover is a major effect of ongoing climate change in arctic tundra ecosystems. Here, I focus on the widespread and deciduous shrub, *Betula nana*, to better understand its spatial distribution of genetic diversity and local adaptation of populations. I used double-digest Restriction Site-Associated DNA sequencing to identify genomic variation in 109 individuals of *B. nana* from 9 sites along a latitudinal gradient in north central Alaska. These sites were chosen to overlap with those sampled for *E. vaginatum* in Chapter 1 to allow for a comparison of population structure, genomic variation, and adaptation of the two species in the same region. A neutral SNP dataset of 1,039 loci (each represented by a single SNP) was used to demonstrate two genetic clusters, one composed of individuals from the No Name (NN) site and the other composed of individuals from all other sites. The general lack of population structure and absence of allelic variation related to environment along the cline of the latitudinal gradient was likely due to high co-ancestry, incomplete lineage sorting of a relatively continuous population with recent isolation, or previously disjunct populations reconnected by contemporary widespread gene flow. The low levels of co-ancestry of NN with the other sites in addition to the high number of private alleles identified for NN may indicate the presence of *B. glandulosa*, or an admixed variant between *B. nana* and *B. glandulosa* at this site. The lack of structure related to environment, treeline, and geography suggests that *B. nana* did not share a similar evolutionary history with *E. vaginatum* along the same latitudinal gradient. The increased prevalence of *Betula* pollen and macrofossils in the region during warming fluxes of the early and mid-Holocene, and generally higher levels of *Betula* pollen south of the Brooks Range further suggest that while *Betula*, and potentially *B. nana*, was present on the north side of Brooks Range during the Last Glacial Maximum, the genus was likely not prevalent. Post-glacial expansion of southern *B. nana*
populations northward could also lead to high levels of co-ancestry between populations north and south of the Brooks Range. The lack of genetic structure and genetic signature related to environmental variation could indicate a higher tolerance for environmental shifts (plasticity) across the range, which could facilitate a competitive advantage for genotypes under climate change.
INTRODUCTION

Increase in shrub cover is a major effect of ongoing climate change in arctic tundra ecosystems. Relative increases in shrub abundance and cover of species such as birch, willow, and alder (Betula, Salix, and Alnus spp.) are expected to alter ecological communities by modifying biodiversity and affecting ecosystem processes (Myers-Smith et al., 2011). In northern Alaska, repeat aerial photography between 1948 and 2001 demonstrated a considerable increase in deciduous shrub cover, especially along hill slopes and in river valleys (Tape et al., 2006). In riparian areas, birch, willow, and Siberian alder (Alnus viridis) are becoming increasingly dominant due to more available nutrients, a longer growing season, and facilitated dispersal of seeds (Tape et al., 2006; Naito and Cairns, 2011; Berner et al., 2018). In a study mapping plant and shrub above-ground biomass (AGB) on the North Slope of Alaska, shrub dominance was highest where average summer temperatures were warmest (specifically in the Brooks Range Foothills), and shrub AGB was 4× greater between the coldest and warmest areas of the North Slope (Berner et al., 2018). While low temperatures that limit reproduction partially explain the sparser abundance of shrub taxa in the northern extent of their ranges in the Low Arctic, shrub species like the dwarf birch (Betula nana L.; Betulaceae) and Salix spp. (Salicaceae) have been shown to respond quickly and outcompete other arctic plant species under conditions of increased air temperature and nutrient availability (Tape et al., 2006; Myers-Smith et al., 2011). DeMarco et al. (2014) found that after 18 years of warming and nutrient addition in deciduous shrub tundra of the Alaskan Arctic, B. nana and Salix spp. dominated all treatment plots while evergreen shrub, forb, and graminoid diversity declined.
As the abundance of shrubs increases in the Alaskan arctic tundra, alteration of ecosystem processes such as nutrient exchange, hydrologic dynamics, energy fluxes, and carbon (C) cycling is anticipated (Myers-Smith et al., 2011). The expansion of shrubs into tussock tundra will likely increase the amount of C storage in plants due to the high C to nitrogen (N) ratio and high lignin content in their woody tissue (Hobbie, 1996; Weintraub and Schimel, 2005). Soil nutrient cycling and availability will likely be affected as increased snowpack below shrubs insulates the soil and increases microbial activity during the winter, thereby increasing C mineralization rates and loss of soil C (Weintraub and Schimel, 2005). In a two-year comparative study of root and litter decomposition in Alaskan moist acidic tundra, *B. nana* leaf litter was found to decompose faster than three other tundra plant species (*Eriophorum vaginatum*, *Vaccinium vitis-idaea*, and *Rhododendron tomentosum*) (McLaren et al., 2017). As shrubs expand, leaf litter and fine root increases will likely result in increased turnover of the quickly-decomposing fine root and leaf carbon pool (McLaren et al., 2017), which may ultimately determine whether an increase of net ecosystem C loss or gain occurs (Myers-Smith et al., 2011; Weintraub and Schimel, 2005).

The potential for range expansion of a species is important under climate change scenarios and can depend on traits such as dispersal ability and colonization history. Plants with high dispersal ability are more likely to retain genetic diversity due to high rates of gene flow between populations. Plant species with light, wind-dispersed seeds, like birch, willow, and alder, are expected to have low genetic differentiation and weak genetic structure (Belton et al., 2021). A comparison of chloroplast DNA (cpDNA) from English and Scottish populations of black alder (*Alnus glutinosa*; Cubry et al., 2015) and birch (*Betula pubescens* and *B. pendula*; Belton et al., 2021) demonstrated low genetic differentiation (GST = 0.268 to 0.38) between sampled populations within each species. Populations of widespread, wind-pollinated, and wind-dispersed
Woody plants tend to have low genetic differentiation and high genetic diversity, as genetic drift or population extinction is less likely to remove novel alleles (Hamrick and Godt, 1996). For example, although half of the range of dwarf willow (Salix herbacea) in Europe was estimated to have been lost, genetic diversity was reduced by only 5% because of the species’ high dispersal ability and wide-scale recolonization after glaciations of the Last Glacial Maximum (LGM; 20,000 years before present (kyr BP)) (Alsos et al., 2009, 2012). Alsos et al. (2015) found evidence for multiple events of long-distance dispersal (LDD) between B. nana populations occurring 280 to >3,000 km away, suggesting LDD is frequent in arctic B. nana populations (Alsos et al., 2015). Genetic diversity of B. nana populations has also been found to be higher than expected for a long-lived shrub with a propensity for clonality. On the Taymyr Peninsula in the Siberian forest-tundra ecotone, the weak geographical structure and higher genetic diversity found for B. nana populations as compared to other shrubs (Alnus alnobetula and Salix spp.) suggested more efficient seed dispersal and recruitment from multiple populations (Meucci et al., 2021).

Shrub population fragmentation and geographic range expansion can be highly variable across short horizontal distances due to landscape heterogeneity, environmental characteristics across multiple scales, and species-specific characteristics. In the absence of gene flow, reduction of population size and/or population isolation can partition genetic diversity and beneficial alleles can ultimately be lost due to random processes such as genetic drift. For instance, in a study of age structure of Salix spp. stands in southwest Yukon, evidence of a potential widespread expansion was found for two of the six sampled sites. While a high proportion of older individuals could indicate a previous rapid expansion at one site, older age classes were absent at the other site, which could indicate recent establishment or increased turnover (Danby and Hik, 2007). Other factors, such as the type of forest-tundra ecotone (such as open-forest or woodland), vegetation
density, permafrost presence, and low temperature can influence shrub expansion and establishment (Danby and Hik, 2007) as well. Reduction of seedling and sapling growth rates, viable seed production, and seedling recruitment have also been found to contribute to limited reproduction and growth of *B. glandulosa* (Hermanutz et al., 1989), *Alnus viridis* subsp. *fruticosa*, and *Salix* spp. at the northern edge of their ranges (Danby and Hik, 2007; Myers-Smith et al., 2011). Here, I focus on the widespread and deciduous shrub, *Betula nana*, to better understand its spatial distribution of genetic diversity and local adaptation of populations in north central Alaska.

*Betula nana* is a monoecious, wind-pollinated, and wind-dispersed circumpolar and circumboreal deciduous shrub. *Betula nana* individuals can live up to 147 years (Miller, 1975; De Groot et al., 1997; Eidesen et al., 2015). The dwarf birch has the propensity for clonal growth (via ramets), which has been hypothesized as its dominant reproductive strategy (De Groot et al., 1997). *Betula nana*’s production of two shoot types (both long and short) gives the shrub a competitive advantage by allocating more energy to aboveground biomass production (Bret-Harte et al., 2001). This allocation strategy can lead to dense canopies (Deslippe and Simard, 2011), which reduce growth of other plants by restricting light availability (Bret-Harte et al., 2001; Myers-Smith et al., 2011). These results suggest that *B. nana* could have a competitive advantage in current tundra ecosystems as permafrost thaws and more soil nutrients become available under climate change (Bret-Harte et al., 2001; Sturm et al., 2001; Hinzman et al., 2005; Tape et al., 2006).

Two subspecies, *B. nana* subsp. *nana* and *B. nana* subsp. *exilis*, comprise *B. nana* (Elven et al., 2011). The subspecies are separated morphologically and well-distinguished by molecular data (Elven et al., 2011; Eidesen et al., 2015). The two subspecies are mostly allopatric. *Betula nana* subsp. *nana* occurs in Europe, Greenland, and western Asia, while *B. nana* subsp. *exilis* is found in central and eastern Asia, from Alaska to northern Canada in North America, and in
Greenland (Hultén and Fries, 1986; De Groot et al., 1997; Elven et al., 2011). In north central Alaska, *B. nana* subsp. *exilis* has a less continuous distribution in the northern moist acidic tundra ecosystem and a more continuous distribution in the taiga ecosystem south of the Brooks Range (Tape et al., 2006; Elven et al., 2011; Eidesen et al., 2015). The shrub mostly occurs in low-nutrient acidic areas, but also in acidic barrens, muskegs, peat bogs, near stream banks, and on rocky slopes and subalpine summits in arctic tundra and taiga ecosystems (Graglia et al., 2001; Torp et al., 2010; Stark et al., 2015).

The distribution of *B. nana* overlaps with *B. glandulosa*, the shrub birch, across Alaska and northern Canada in North America and into Greenland (De Groot et al., 1997). Although Wein and Bliss (1974) described a shift in Alaska from *B. nana* subsp. *exilis* to *B. glandulosa* in moist tussock tundra where black spruce (*Picea mariana*) occurs, hybridization has been hypothesized to be frequent, to the extent that Hultén considered *B. glandulosa* and *B. nana* subsp. *exilis* to be completely introgressed in North America (Hultén, 1968, 1971). *B. nana* subsp. *exilis* is a prostrate shrub growing up to 1 m in height with short (0.5 to 1 cm long) leaf blades that are either glabrous or with a few glands once mature (De Groot et al., 1997). *B. glandulosa* can be either a prostrate or erect shrub ranging from 0.3 to 2.5 m in height (De Groot et al., 1997). Leaves range from 0.5 to 2 cm long, typically with glandular spots on the abaxial side (De Groot et al., 1997). Due to hybridization between the two (and with other *Betula* species), and in addition to similar morphological features, *B. nana* subsp. *exilis* and *B. glandulosa* are difficult to distinguish, especially where they co-occur (De Groot et al., 1997). As I focus on *B. nana* subsp. *exilis* in this chapter, it will be hereafter referred to as *B. nana*; *B. nana* subsp. *nana* will be designated with the subspecies epithet.
Weak population structure has often been found for arctic deciduous shrubs, primarily due to long-distance dispersal (Cortés et al., 2014; Alsos et al., 2015; Meucci et al., 2021), and genetic diversity can vary from low (Alsos et al., 2002, 2015; Cortés et al., 2014) to moderate/high (Meucci et al., 2021) across sites. Moderate differentiation has been identified across some B. nana populations (Alsos et al., 2002; Borrell et al., 2018). On Svalbard, Norway, relatively low genetic diversity and moderate genetic differentiation characterized B. nana subsp. nana populations (Alsos et al., 2002). Genetic diversity across populations was still considered noteworthy and was attributed to past sexual reproduction and long-lived clones derived from a large initial gene pool (Alsos et al., 2002). Moderate genetic differentiation found among Svalbard populations was explained by inbreeding and genetic drift, which is consistent with patterns seen for highly fragmented B. nana subsp. nana populations (Alsos et al., 2002). Borrell et al. (2018) compared fragmented, highly genetically differentiated populations in Britain and continuous, large populations in Scandinavia and found significantly higher allelic richness and expected heterozygosity across Scandinavia sites, as would be expected for more continuous populations in the Arctic as compared to the fragmented population distribution of small populations of Britain. (Borrell et al., 2018).

In Europe, population structure has been identified across B. nana subsp. nana populations (Jadwiszczak et al., 2017), suggesting local adaptation and potential ecotypic differentiation. Population structure analyses showed significant differentiation of Polish populations from the remaining populations, likely due to selection, genetic drift, and restricted gene flow leading to isolation and small population sizes (Jadwiszczak et al., 2017). Genetic diversity of relict (or populations near the southern, retreating edge) populations was higher on average than widespread, centrally located populations in Finland and Russia, likely due to retained genetic
variability from past sexual reproduction and/or hybridization (Alsos et al., 2002; Jadwiszczak et al., 2017). High genetic connectivity of north central European B. nana subsp. nana populations supports range expansion of B. nana subsp. nana northward, similarly to northern B. nana populations becoming more continuous in the Arctic (Tape et al., 2006; Naito and Cairns, 2011; Berner et al., 2018).

Betula nana often co-occurs with the tussock cottongrass (Eriophorum vaginatum L.; Cyperaceae), a foundation species of the moist acidic tundra (MAT) in north central Alaska. Eriophorum vaginatum has similar life history traits as B. nana, as it is also a wind-pollinated and wind-dispersed plant, but it is a graminoid as opposed to a deciduous shrub. Due to shared attributes of wind pollination and wind dispersal, we expect a similar population structure of B. nana populations as that found for E. vaginatum along the same latitudinal gradient in north central Alaska (Stunz et al., 2022). Pollen of Eriophorum and Betula in yedoma soil profiles north and south of the Brooks Range could indicate that Betula nana persisted in refugia north of the Brooks Range during the LGM, but the pollen may be from another Betula species (Livingstone, 1955; Naito and Cairns, 2011; Schirrmeister et al., 2016; Lapointe et al., 2017). Macrofossils identified to B. glandulosa and/or B. nana found south of the Brooks Range provide stronger support for B. nana in southern refugia during the LGM (Schirrmeister et al., 2016). Long-term research and genomic investigations of E. vaginatum now suggest differentiation due to older and contemporary gene flow barriers in addition to local adaptation. Given gene flow expectations and similar life history traits, we expect a similar structure for B. nana, assuming it shared a similar historic distribution with E. vaginatum. Notably, we expect the Brooks Range glaciation to have reinforced adaptations in subrefugia north and south of the Brooks Range. Furthermore, the contemporary gene flow barrier of treeline is expected to influence structure between populations north and south.
of treeline, as found for *E. vaginatum* (Stunz et al., 2022). Alternatively, there may be a lack of both population structure and local adaptation across *B. nana* sites, as has been found for multiple shrub species in arctic and taiga ecosystems (Cortés et al., 2014; Alsos et al., 2015; Meucci et al., 2021), due to high genetic connectivity across large populations.

Here, I investigate clonality, genomic diversity, adaptation, and gene flow to better understand the ability of *B. nana* to expand in density and distribution in a whole ecosystem setting. Using thousands of single nucleotide polymorphisms (SNPs), I will address the following questions: (1) Is there population structure among *B. nana* populations along the same latitudinal gradient in north central Alaska as found for *E. vaginatum* (Stunz et al., 2022)? (2) Is there evidence of genetic structure linked to adaptation for latitude of origin? (3) Does genetic diversity vary along the latitudinal gradient?

**MATERIALS AND METHODS**

**Study area**

The study area consisted of a latitudinal gradient covering ~324 km in north central Alaska (Figure 2.1; Table 2.1), from north of Fairbanks (65.433°, -145.512°) to Sagwon (69.424°, -148.698°). The Continental Divide at the crest of the Brooks Range divides the region into south slope and north slope components, roughly at the intersection of two climatic regions, the Arctic and the interior (Haugen, 1982; Brown and Kreig, 1983). Beginning just north of Fairbanks, taiga vegetation is dominant, with forest and lowland treeless bogs comprising much of the interior. Alpine or tussock tundra is found once elevations exceed ~700 m. Treeline occurs on the south slope of the Brooks Range (Figure 2.1), while tundra plant communities are found north of treeline. Permafrost is continuous north of treeline and discontinuous in most of the interior, where soil
Figure 2.1. Map of *Betula nana* sampling locations and modeled maximum Pleistocene glacial extent (Kaufman and Manley, 2004; Kaufman et al., 2011) along a latitudinal gradient in north central Alaska. Treeline is indicated by the dashed black line. The 9 collection site abbreviations are as for Table 2.1.
Table 2.1. Collection sites, GPS coordinates, elevation (meters) and vegetation type of *B. nana* in north central Alaska. MAT = Moist Acidic Tundra.

<table>
<thead>
<tr>
<th>Site</th>
<th>Latitude (N), Longitude (W)</th>
<th>Elevation (m)</th>
<th>Vegetation Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eagle Creek (EC)</td>
<td>65.4332°, -145.5118°</td>
<td>771</td>
<td>MAT</td>
</tr>
<tr>
<td>No Name Creek (NN)</td>
<td>66.1171°, -150.1676°</td>
<td>167</td>
<td>Tussock bog</td>
</tr>
<tr>
<td>Gobbler’s Knob (GO)</td>
<td>66.7459°, -150.6862°</td>
<td>520</td>
<td>Muskeg</td>
</tr>
<tr>
<td>Coldfoot (CF)</td>
<td>67.2631°, -150.1591°</td>
<td>321</td>
<td>Muskeg</td>
</tr>
<tr>
<td>Timberline (TB)</td>
<td>68.0300°, -149.6737°</td>
<td>760</td>
<td>MAT</td>
</tr>
<tr>
<td>Atigun Camp (AT)</td>
<td>68.1730°, -149.4392°</td>
<td>1,063</td>
<td>MAT</td>
</tr>
<tr>
<td>Toolik Lake (TL)</td>
<td>68.6292°, -149.5778°</td>
<td>758</td>
<td>MAT</td>
</tr>
<tr>
<td>Sagwon River (SR)</td>
<td>68.9075°, -148.8433°</td>
<td>400</td>
<td>MAT</td>
</tr>
<tr>
<td>Sagwon (SG)</td>
<td>69.4244°, -148.6976°</td>
<td>299</td>
<td>MAT</td>
</tr>
</tbody>
</table>

Parent materials, slope angle and aspect, drainage, and vegetation often indicate permafrost presence (Brown and Kreig, 1983). In general, the interior is warmer, and annual rainfall declines north of the Brooks Range towards the coastal plain.

**Sample collection**

During the summers of 2015 and 2017, leaf samples were collected from 13 *B. nana* individuals at each of the 9 sampled sites (109 accessions) along the latitudinal gradient following the Dalton Highway (Figure 2.1). Away from the Dalton Highway, logistical access to this region is limited and cost prohibitive, as it can only be accessed by helicopter. The latitudinal distance between populations was ≤ 0.75°. Individuals were sampled at least 50 m from the roadside and at least 10 m apart to minimize bias based on same seed parentage. Sampled leaves were dried and stored in silica gel. Sites north of treeline are classified as MAT (Britton, 1966; Wein and Bliss, 1974; Shaver et al., 1986), while most sites south of treeline are muskeg or tussock bog unless above 700 m (Table 2.1) (Cleve and Dyrness, 1983; Kummerow, 1983; Shaver et al., 1986). The sampling strategy was designed to include sites that overlapped with *E. vaginatum* sites from Stunz.
et al. (2022), but sampling only extended as far north as Sagwon as *B. nana* individuals were not found in the coastal region near Prudhoe Bay.

**DNA extraction and ddRAD-seq library preparation**

Genomic DNA was extracted from 25 mg of dried leaf tissue for 13 samples per site. Due to high concentrations of phenolic compounds identified in *B. nana* leaves (Graglia et al., 2001; Wang et al., 2013), a modified CTAB DNA extraction method was used to remove suspended secondary compounds in order to obtain high quality DNA. The Wang et al. (2013) method combines steps from three protocols (Doyle and Doyle, 1987; Cullings, 1992; Zeng et al., 2002), and added steps include the use of ice-cold TNE buffer to remove polysaccharides, keeping tubes on ice to reduce nuclease activity and potential DNA degradation, and adding Sarkosyl (sodium lauroyl sarcosinate) solution and Proteinase K to denature and degrade contaminating proteins (Hong et al., 1995; Collard et al., 2007; Sahu et al., 2012; Wang et al., 2013). DNA concentrations were quantified using the Qubit dsDNA BR Assay Kit (Invitrogen) and Qubit 3.0 Fluorometer (Thermo Fisher Scientific). DNA samples with a minimum concentration of 0.02 µg/µL were included for preparation of double-digest Restriction Site-Associated DNA (ddRAD) -seq libraries, which generally followed DaCosta and Sorenson (2014) and Stunz et al. (2022).

To assemble ddRAD libraries, 43 µL of each ~25 ng/µL DNA sample was digested with 20 U of EcoRI and 20 U of MspI restriction enzymes (New England Biolabs; cut sites: 5’-GAATTC-3’; 5’-CCGG-3’) and 5 µL of 10× NEBuffer 4 (New England Biolabs) in a thermocycler at 37°C for 30 min followed by enzyme deactivation at 65°C for 20 min. Custom barcodes and indices were ligated to digested DNA, with the addition of 50 nM of custom EcoRI and MspI adapters (containing individual barcodes), 2 µL of 10× NEBuffer (New England Biolabs), 0.6 µL
of rATP (Promega), 0.4 µL of ddH₂O, and 1 µL of T4 DNA ligase (New England Biolabs) to each 50 µL sample of digested DNA. Ligation was completed in the thermocycler (22°C for 30 min), followed by enzyme deactivation (65°C for 20 min). Adapter-ligated DNA fragments were then double-side size selected with a 0.8× SPRI bead clean up. Right-sided selection for large fragments was done with AMPure XP beads (Beckman Coulter, Inc.) added at 0.55× volume of the starting ligated DNA solution, and the supernatant was transferred to new tubes. Subsequently, a left-sided size selection was done with 0.25× volume AMPure XP beads added to the supernatant. The supernatant was then discarded, and the beads were washed twice with 80% ethanol and air-dried. DNA was re-suspended with 25 µL ddH₂O and eluted for a minimum of 30 minutes.

To amplify size-selected DNA fragments, PCR was performed with a solution of 15 µL template DNA, 30 µL of Phusion High-Fidelity PCR Master Mix (Thermo Scientific), 9 µL of ddH₂O, and 3 µL of each 10 µM primer specific to sequences at the ends of each custom barcode and index (see DaCosta and Sorenson, 2014). PCR was performed in a thermocycler: 98°C for 30 s; 22 cycles at 98°C for 10 s, 60°C for 30 s, 72°C for 40 s, and 72°C for 5 min. Amplified DNA fragments across samples were cleaned using a 1.8× AMPure XP bead clean-up protocol (Beckman Coulter, Inc.). After adding beads and discarding the supernatant, beads were washed twice with 80% ethanol and re-suspended in 40 µL ddH₂O. Finally, cleaned PCR products were quantified with the Qubit dsDNA BR Assay Kit (Invitrogen) and Qubit 3.0 Fluorometer (Thermo Fisher Scientific). Additionally, samples were visualized with gel electrophoresis to ensure that all were similarly size-selected. Equimolar concentrations of samples with unique barcode combinations were pooled. Multiplexed libraries were sequenced at the Center for Genome Research and Biocomputing at Oregon State University on an Illumina HiSeq 3000.
SNP identification and genotyping

Files of raw Illumina read sequence data were de-multiplexed and analyzed via the de novo assembly method in STACKS 2.41 (Catchen et al., 2011, 2013) and using in-house Python scripts. First, the quality of raw sequences was assessed before trimming with FASTQC (Andrews, 2010). The process_radtags program in STACKS was then used to remove reads without intact radtags and trim all reads to 144 bp. The quality of sequences was again assessed post-trimming with FASTQC to confirm removal of adapters and low-quality sequence data. The sequences were then processed in the program ustacks to align the short-read sequences into ‘stacks.’ Stacks were compared using a maximum likelihood framework (Hohenlohe et al., 2010) to identify loci and detect single nucleotide polymorphisms (SNPs). The minimum depth of coverage for stack creation (m) was set to three (m 3; default), The maximum distance (in nucleotides) between each stack (M) was set to four (M 4), and all other parameters were set to default values. The cstacks program was used to build a catalog of consensus loci based on matching sets of reads built in ustacks. The number of mismatches permitted between sampled loci when building a catalog (n) was set to four (n 4). Parameter choice was guided by the optimization methods of Paris et al. (2017) by assessing the number of loci and SNPs retained across 80% of individuals from each site (r80) for values 1-9 for M, assuming M = n, and fixing m to 3 (Paris et al., 2017; Rochette and Catchen, 2017).

Sets of putative loci were then compared to the catalog of loci with the sstacks program. Data was arranged by locus, instead of by sample, with the tsv2bam program. The ddRAD data was analyzed locus by locus across all individuals to genotype individuals for each SNP with the gstacks program. The data set was then pruned for minor alleles occurring in less than 1% of reads, as rare genotypes are likely the result of sequencing error (Tabangin et al., 2009). The STACKS
populations program was run to retain only loci found in all sampled populations with \( \leq 20\% \) missing data. Additionally, individuals with \( \geq 20\% \) missing loci were identified with the populations.log.distribs file from the STACKS population program output and subsequently removed.

Two sets of SNPs were created: one single SNP per locus “stack” set containing putatively neutral markers (or the neutral dataset), and a set containing all SNPs (or the comprehensive set) used for redundancy analysis (RDA). Loci under balancing or directional selection were identified with BayeScan v.2.1 (Foll and Gaggiotti, 2008) via a Bayesian \( F_{ST} \) outlier test with default values, 20 pilot runs of 5,000 iterations were implemented, followed by 100,000 iterations, sampled every 10 iterations, with a 50,000 iteration burn-in. Outputted iterations were increased to ensure convergence of the MCMC chain (confirmed by visual assessment). Candidate outlier loci were identified as those with a \( q \)-value < 0.05 (or a false discovery rate [FDR] of 5%). Additionally, SNPs deviating from Hardy-Weinberg Equilibrium (HWE; \( p \leq 0.001 \)) were identified with PLINK\(^2\) v.1.07 (Purcell et al., 2007). Outlier loci and SNPs deviating from HWE were removed, and one SNP per locus was retained to create the neutral SNP data set used for estimating genomic diversity, demographic statistics, and population structure.

Candidate loci found with BayeScan were investigated using the blastn function in BLAST (Altschul et al., 1990b) and queried against the NCBI non-redundant nucleotide database to find the associated Gene Ontology term for each gene based on a percentage identity match of at least 80.0 and an E-value threshold of at least \( 1 \times 10^{-4} \).

\(^2\) http://pngu.mgh.harvard.edu/purcell/plink/
Patterns of genomic diversity

Heterozygosity (observed ($H_o$) and expected ($H_e$)), allelic richness ($Ar$) and the inbreeding coefficient ($F_{IS}$) were estimated across all loci and for each population with the \textit{divBasic} function in the \textit{diveRsity} R package v.1.9.90 (Keenan et al., 2013). Private alleles were estimated with the R \textit{PopGenReport} package v.3.0.4 (Adamack and Gruber, 2014; Gruber and Adamack, 2015). Clonality of individuals was assessed with the Assign Clones analysis in GENODIVE 2.0b2.5 (Meirmans and Van Tienderen, 2004) using an infinite allele mutation model and a distance threshold of 0 to 4. The Corrected Nei’s diversity index, 999 permutations, and randomization of alleles across all populations were also used to investigate clonal structure. Effective population size ($Ne$) was estimated with NeEstimator v.2.1 (Waples and Do, 2008; Do et al., 2014) using the bias-corrected linkage disequilibrium method (Waples, 2006), random mating, and an allele frequency threshold of $\geq 0.02$ for $Ne_{LD}$ calculation.

Population structure

The neutral dataset was processed in STRUCTURE v.2.3.4 (Pritchard et al., 2000; Falush et al., 2003), which uses a Bayesian clustering algorithm to assign individuals to genetic clusters ($K$). The analysis was comprised of 20,000 burn-in iterations followed by 50,000 replicates of each population value ($K = 1-10$), and each run was conducted 10 times (Schweizer et al., 2016). To determine the optimal value of $K$, the $\Delta K$ statistic (Evanno et al., 2005) was evaluated using STRUCTURE HARVESTER v.0.6.94 (Earl and vonHoldt, 2012). The greedy method and 1,000 random permutations were used in CLUMPP v.1.1.2 to account for variation in cluster assignment across STRUCTURE runs. Bar charts displaying the proportion of cluster membership for each individual were created and modified with DISTRICT v.1.1 (Rosenberg, 2004). Population
structure was also investigated using Discriminant Analysis of Principal Components (DAPC) and Bayesian clustering in the R package adegenet v.2.0.1 (Jombart et al., 2010). DAPC is useful to avoid a priori assignment and corroborate results of STRUCTURE as it provides a non-model-based method to estimate cluster assignment of individuals. The find.clusters function was run in adegenet to transform allele frequencies and determine the optimal number of clusters ($K = 1$-$10$) via $k$-means clustering of principal components and the Bayesian Information Criterion (BIC) (Jombart et al., 2010).

To investigate the recent shared ancestry of $B.\ nana$ individuals, fineRADstructure (Malinsky et al., 2018) was run using the neutral dataset. fineRADstructure is particularly useful for investigating relationships between and within populations containing family groups and estimating ancestry sources of populations, methods that are unavailable in STRUCTURE (Malinsky et al., 2018). The inclusion of multiple SNPs at a locus further refines nearest neighbor assignment in fineRADstructure, which is opposed to typical single SNP population structure methods (Malinsky et al., 2018). In addition, all SNPs occurring at a locus are assumed to be in nearly complete linkage disequilibrium (LD), and therefore historical recombination is not considered (Malinsky et al., 2018).

Two methods, RADpainter and finestructure (Lawson et al., 2012), comprise the fineRADstructure package. Without providing a priori population assignment, RADpainter first creates a co-ancestry matrix by identifying identical or most similar haplotypes at each locus, and finestructure utilizes a Markov chain Monte Carlo (MCMC) algorithm to cluster individuals by most similar haplotype and infer population structure. After a burn-in of 100,000 steps and 100,000 MCMC iterations, a tree-building algorithm was run with default parameters. R scripts (fineradstructureplot.r and finestructurelibrary.r, available at
http://cichlid.gurdon.cam.ac.uk/fineRADstructure.html) were then used to visualize the results.

**Genotype-environment association**

As geographic and landscape variables often influence patterns of genomic variation and spatial distribution of plants (Manel et al., 2003; Sork et al., 2013; Lind et al., 2017), genotype-environment association (GEA) methods were implemented to investigate these associations using the comprehensive data set with missing data imputed based on averaged allele frequencies for each site. Nineteen environmental predictors that are important for ecological models in the Arctic (Pearson et al., 2013) were obtained from the WorldClim 2.0 Bioclimatic database (Fick and Hijmans, 2017). We used averages from 1970 to 2000 for GPS locations of sampling sites to distinguish climatic attributes of populations. Global multi-resolution terrain elevation data (GTMED2010; Danielson and Gesch, 2011) were obtained from Google Earth Engine (Gorelick et al., 2017) based on GPS coordinates. In addition, Moran’s eigenvector maps (MEMs) were calculated based on linear distances among sites with the R adespatial v.0.0-7 (Dray et al., 2016) and spdep v.0.6-9 (Bivand et al., 2013) packages. MEMs were included as predictor variables if their Moran’s I value was associated with significant (100 permutations, $p < 0.05$) allele frequency variance. Predictor variables with $R^2 > 0.75$ and $< -0.75$ (Schweizer et al., 2016; Forester et al., 2018) were removed.

The R vegan v.2.5.6 (Oksanen et al., 2013) and psych v.1.8.12 (Revelle, 2014) packages were used to run a redundancy analysis (RDA) to investigate co-variation of alleles in response to environmental predictors. Although candidate outlier loci analyses using ddRAD are incomplete as they only sample a fraction of the genome (Lowry et al., 2017), they can be informative for identifying some genes that may have a role in adaptation. The relatively small genome of *B. nana*
(1C ≈ 0.46 pg; Anamthawat-Jonsson et al., 2010) increases the probability of capturing candidate outlier loci. Site-based allele frequencies used for RDA were computed with adegenet v.2.0.1. The RDA was then run using the reduced predictor variable dataset (n = 7; see Results). The significance of the final models and constrained axes was identified with 999 permutations and a p-value of 0.05 (Forester et al., 2018).

**RESULTS**

**Genomic sequence data and genomic diversity**

Once low-quality reads and reads without radtags were removed, 212,912,575 single-end sequences were retained across 109 individuals, with an average of 1,953,326 reads/individual. All samples had a coverage depth > 12×. Clonality analyses with GENODIVE were inconclusive, as no clonal pairs were recognized with a distance threshold below four, meaning that at least four mutations between individuals were required to assign clonal pairs. Individuals with ≥ 20% missing loci were removed, resulting in a dataset of 101 individuals for fineRADstructure analysis. A comprehensive dataset (1,294 loci/3,817 SNPs) and a neutral, single SNP dataset (1,039 loci, each represented by one SNP) were created once SNPs deviating from HWE (p ≤ 0.001) were filtered using PLINK. Nine outlier loci were identified with BayeScan. Gene Ontology associations were found for three candidate loci using the blastn function in BLAST and searched against the NCBI non-redundant nucleotide database (Table 2.2). The nine outlier loci were removed to create the putatively neutral dataset. Using output from fineRADstructure, 10 clonal pairs and two family groups (one group of three individuals and another of five individuals) were identified using a > 13% co-ancestry cut-off. Only one individual from each clonal pair or family
group was retained, resulting in a neutral, single SNP dataset of 85 individuals and 1,039 loci/SNPs for genomic diversity statistics and subsequent population structure analyses.

**Table 2.2.** Annotated *B. nana* outlier genes with a percentage identity match of at least 80.0 and an E-value threshold of at least 1 X 10^{-4}.

<table>
<thead>
<tr>
<th>Locus ID</th>
<th>Gene ID</th>
<th>E-value</th>
<th>% Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>8270</td>
<td>Probable lysine-specific demethylase ELF6</td>
<td>2.00E-45</td>
<td>91%</td>
</tr>
<tr>
<td>25669</td>
<td>Putative cytochrome c biosynthesis ccmC-like mitochondrial protein</td>
<td>8.00E-64</td>
<td>99%</td>
</tr>
<tr>
<td>36794</td>
<td>Probable LRR receptor-like serine/threonine protein kinase At1g06840</td>
<td>2.00E-44</td>
<td>94%</td>
</tr>
</tbody>
</table>

A mean $F_{IS} = 0.001$ (SE = ± 0.005) was found for global population estimates of the *B. nana* neutral dataset, with relatively high $F_{IS}$ (0.228) in CF. Heterozygosity was generally low and varied little among sites, as a mean $H_o$ of 0.102 (SE = ± 0.002) and a mean $H_e$ of 0.103 (SE = ± 0.002) were found (Table 2.3). Allelic richness ($A_r$) ranged from 1.270 to 1.370, and many private alleles ($P_o$) were identified across sites. Notably, 111 private alleles were identified in NN (Table 2.3). $N_e$ ranged from 43.7 to $\infty$ across sites.
Table 2.3. Genetic diversity summary and demographic statistics for the neutral data set (after clone removal) of the nine *Betula nana* populations and eight overlapping *Eriophorum vaginatum* populations sampled in north central Alaska. Populations are ordered geographically from South to North, and site abbreviations follow Table 2.1. Sample size (*N*), allelic richness (*Ar*), private alleles (*Pa*), observed heterozygosity (*Ho*), expected heterozygosity (*He*), inbreeding coefficient (*FIS*), and effective population size (*Ne*).

<table>
<thead>
<tr>
<th>Site</th>
<th><em>N</em></th>
<th><em>Ar</em></th>
<th><em>Pa</em></th>
<th><em>Ho</em></th>
<th><em>He</em></th>
<th><em>FIS</em></th>
<th><em>Ne</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Betula nana</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC</td>
<td>13</td>
<td>1.340</td>
<td>17</td>
<td>0.096</td>
<td>0.105</td>
<td>0.078</td>
<td>∞</td>
</tr>
<tr>
<td>NN</td>
<td>7</td>
<td>1.310</td>
<td>111</td>
<td>0.121</td>
<td>0.102</td>
<td>-0.190</td>
<td>∞</td>
</tr>
<tr>
<td>GO</td>
<td>9</td>
<td>1.350</td>
<td>17</td>
<td>0.098</td>
<td>0.110</td>
<td>0.102</td>
<td>∞</td>
</tr>
<tr>
<td>CF</td>
<td>10</td>
<td>1.330</td>
<td>19</td>
<td>0.086</td>
<td>0.112</td>
<td>0.228</td>
<td>∞</td>
</tr>
<tr>
<td>TB</td>
<td>7</td>
<td>1.270</td>
<td>8</td>
<td>0.093</td>
<td>0.087</td>
<td>-0.058</td>
<td>43.7</td>
</tr>
<tr>
<td>AT</td>
<td>10</td>
<td>1.310</td>
<td>7</td>
<td>0.092</td>
<td>0.095</td>
<td>0.024</td>
<td>∞</td>
</tr>
<tr>
<td>TL</td>
<td>8</td>
<td>1.320</td>
<td>12</td>
<td>0.110</td>
<td>0.101</td>
<td>-0.098</td>
<td>∞</td>
</tr>
<tr>
<td>SR</td>
<td>9</td>
<td>1.370</td>
<td>22</td>
<td>0.117</td>
<td>0.114</td>
<td>-0.028</td>
<td>∞</td>
</tr>
<tr>
<td>SG</td>
<td>12</td>
<td>1.360</td>
<td>9</td>
<td>0.108</td>
<td>0.105</td>
<td>-0.037</td>
<td>∞</td>
</tr>
<tr>
<td><strong>Eriophorum vaginatum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC</td>
<td>16</td>
<td>1.590</td>
<td>1</td>
<td>0.171</td>
<td>0.167</td>
<td>-0.025</td>
<td>5923.3</td>
</tr>
<tr>
<td>NN</td>
<td>16</td>
<td>1.661</td>
<td>0</td>
<td>0.177</td>
<td>0.178</td>
<td>0.004</td>
<td>∞</td>
</tr>
<tr>
<td>GO</td>
<td>16</td>
<td>1.663</td>
<td>0</td>
<td>0.173</td>
<td>0.177</td>
<td>0.023</td>
<td>∞</td>
</tr>
<tr>
<td>CF</td>
<td>18</td>
<td>1.667</td>
<td>0</td>
<td>0.184</td>
<td>0.177</td>
<td>-0.036</td>
<td>∞</td>
</tr>
<tr>
<td>TB</td>
<td>16</td>
<td>1.632</td>
<td>0</td>
<td>0.166</td>
<td>0.171</td>
<td>0.028</td>
<td>221.4</td>
</tr>
<tr>
<td>AT</td>
<td>16</td>
<td>1.634</td>
<td>0</td>
<td>0.171</td>
<td>0.172</td>
<td>0.011</td>
<td>∞</td>
</tr>
<tr>
<td>TL</td>
<td>16</td>
<td>1.627</td>
<td>0</td>
<td>0.171</td>
<td>0.169</td>
<td>-0.011</td>
<td>∞</td>
</tr>
<tr>
<td>SG</td>
<td>17</td>
<td>1.645</td>
<td>0</td>
<td>0.174</td>
<td>0.173</td>
<td>-0.006</td>
<td>∞</td>
</tr>
</tbody>
</table>

Population structure

A highly supported split (100% branch support) between a distinct group of only NN individuals and a group of all remaining individuals was demonstrated with a fineRADstructure clustered co-ancestry matrix for the neutral dataset (Figure 2.2). A poorly supported split (57.4% branch support) occurred within the group of remaining sites (excluding NN), separating GO and SR into one group and TB and AT into another, with individuals from other sites interspersed in both groups. None of the *B. nana* populations formed a supported cluster, except for NN.
Figure 2.2. fineRADstructure co-ancestry matrix. Co-ancestry pairwise coefficients are color-coded from low (yellow) to high (black), and individuals are clustered according to the pairwise matrix of co-ancestry coefficients in the dendrogram. Site abbreviations follow Table 2.1. The No Name (NN) group is distinguished with a light blue branch in the dendrogram and bars alongside NN individuals in the matrix.
Individuals from various populations clustered together, delineating shared co-ancestry among individuals. Fine-scale structure of multiple family groups and clonal pairs were also evident and denoted by higher levels of co-ancestry (> 13%) with one another. Family groups and clonal pairs were found in all sites except EC, CF, and SG and always consisted of individuals from the same site. The most clonal pairs occurred in populations north of the Brooks Range TB (3), SR (3), and AT (2). One pair each was found at GO and TL. The largest family group consisted of five NN individuals, with a smaller family group consisting of three individuals in TL.

STRUCTURE output was not informative across the nine sites, likely due to the presence of family groups and recent co-ancestry across populations. $K = 2$ was the best supported $K$-value in the BIC score in DAPC analyses (Figure 2.3). Plotting DAPC results by site showed notable genetic distance of NN from the other sites, suggesting one cluster containing NN individuals and one cluster containing individuals from all other sites, though assignment on an individual basis did not support a unique cluster of only NN individuals.
Figure 2.3. Population structure results from the neutral SNP data set. (A) Scatter plot from the discriminant analysis of principal components (DAPC) with 40 principal components retained on the first two Discriminant Analysis axes showing the differentiation between the 9 populations and inertia ellipses. Each color represents a sampled site. Blue dots represent individuals from the Eagle Creek (EC) population, purple dots represent individuals from No Name (NN) population, lavender dots represent individuals from the Gobbler’s Knob (GO) population, gold dots represent individuals from the Coldfoot (CF) population, yellow dots represent individuals from the Timberline (TB) population, light orange dots represent individuals from the Atigun Camp (AT) population, orange dots represent individuals from the Toolik Lake (TL) population, dark orange dots represent individuals from the Sagwon River (SR) population, and red dots represent individuals from the Sagwon (SG) population. (B) Bar graph of STRUCTURE results for population structure analysis. Each vertical bar represents an individual, and colors show the proportion of ancestry assigned to each of the two clusters ($K = 2$), as inferred from $\Delta K$ values. *Betula nana* populations are ordered from south to north location along the latitudinal gradient in north central Alaska.
**Genotype-environment association**

Significant final models and constrained axes using 999 permutations and a \( p \)-value of 0.05 were not identified for the comprehensive dataset due to insignificant co-variation of allele frequencies and environmental predictors along the gradient.

**Discussion**

Genetic structure related to geography, treeline, or other environmental factors was not identified for *B. nana* at the landscape scale. Consequently, the hypothesis that this shrub species shared a similar evolutionary history with the dominant graminoid examined over the same geographic range was not supported. The only structure identified for *B. nana* across this latitudinal gradient in north central Alaska was supported by fineRADstructure and DAPC analyses and included two clusters, one comprised of all individuals from No Name Creek (NN), while the other comprised all other populations sampled. The fineRADstructure algorithm infers most recent shared ancestry among individuals, and the lack of population structure for the cluster of eight sites suggests high co-ancestry or incomplete lineage sorting of a relatively continuous population after recent isolation, or previously disjunct populations reconnected by contemporary widespread gene flow. These results differ from those found for *E. vaginatum* across the same latitudinal gradient, in which three genetic clusters (EC, south of treeline, and north of treeline) were found with a break at treeline supporting the presence of glacial gene flow barriers and the contemporary environmental gradient (Stunz et al., 2022).

Pollen of both *Eriophorum* and *Betula* has been found in yedoma soil profiles south and north of the Brooks Range during the last ~30,000 years, suggesting the region was a refugium for both during the LGM (Livingstone, 1955; Naito and Cairns, 2011; Schirrmeister et al., 2016;
Lapointe et al., 2017), however specific species aren’t identified. *Betula glandulosa/nana* macrofossils south of the Brooks Range support the presence of *B. glandulosa* and/or *B. nana* during the early to mid-Wisconsin glaciation (up until ~15 kya BP) (Schirrmeister et al., 2016). While graminoid tundra was thought to dominate Beringia during the Wisconsin glaciation, much of the vegetation had transitioned to shrubs (including *Betula*) by the Holocene (Ager and Phillips, 2008; Naito and Cairns, 2011). Yedoma soil profiles show *Betula* was present in ~30 kyr BP paleoenvironmental records, with increases in *Betula, Salix,* and *Populus* north of the Brooks Range ~12.5 kyr BP in response to warmer temperatures and a wetter climate (Lapointe et al., 2017). *Betula, Alnus,* and *Salix* pollen were also found in Alaska North Slope lake sediment core samples, suggesting northern refugia for these shrub lineages during the LGM (Livingstone, 1955; Naito and Cairns, 2011). These patterns are similar to what Brubaker et al. (2005) found for levels of *Betula* pollen across Beringia. While the amount of *Betula* pollen largely varied among sites, *Betula* pollen reached levels of > 10% in scattered sites across Beringia 21 to 20 kya BP and generally increased from ~18 to 9 kya BP, becoming abundant (> 40%) across much of central and southern Alaska by 15 kya BP. *Betula* pollen reached high levels (> 60%) at about a quarter of sites, primarily in central and southern Alaska, 14 to 9 kya BP (Brubaker et al., 2005). Warming periods during the early and mid-Holocene, leading to the retreat of most glacial masses by 10 kya BP, would have permitted further population expansion of *Betula* from refugial populations in Alaska (Hamilton, 1986; Kaufman and Manley, 2004; Mason and Bigelow, 2008). These results suggest that while *Betula,* and potentially *B. nana,* was present on the north side of the Brooks Range during the LGM, the genus was likely not prevalent, which is also supported by very limited suitable habitat modeled in the *B. nana* ENM (Ch. 3; Figure 3.1). If the abundance of *B. nana* was very low in the northern refugium compared to *E. vaginatum* during the LGM, with abundance in
the southern refugium increasing much earlier than in the north, the spread of *B. nana* may have been primarily from southern populations and differentiation due to allopatri, and would not be apparent through genetic structure as found for the populations studied here in comparison to *E. vaginatum*.

Given that *B. nana* population genomics does not correspond to that found for *E. vaginatum*, it is relevant to determine if the pattern we found is consistent with those described for other arctic shrubs. Arctic and alpine population genetic studies of wind-pollinated and wind-dispersed deciduous shrub species have found: (1a) a general lack of population structure and low genetic diversity (*S. herbacea* in the Swiss Alps; Cortés et al., 2014) or (1b) high genetic diversity (*B. nana* subsp. *nana* in Siberia; Meucci et al., 2021), in addition to (2) higher than expected genetic differentiation and moderate genetic diversity (*B. nana* subsp. *nana* on Svalbard; Alsos et al., 2002).

Genetic diversity for *B. nana* and *E. vaginatum* across the same eight sites (EC, NN, GO, CF, TB, AT, TL, SG) showed trends of lower heterozygosity and allelic richness, more private alleles, and increased inbreeding of *B. nana* as compared to *E. vaginatum* (Table 2.3). Observed heterozygosity of *B. nana* (mean = 0.101; SE = ±0.002) was also lower than expected heterozygosity (mean = 0.102; SE = ±0.002), while both observed (mean = 0.173; SE = ±0.001) and expected heterozygosity (mean = 0.173; SE = ±0.001) were nearly identical for *E. vaginatum*. Several clonal pairs and family groups of *B. nana* individuals were identified in multiple populations, with individuals with higher than average co-ancestry mostly in northern sites. Increased clonality north of treeline, or in arctic tundra, corresponds with patterns seen for *B. nana* populations elsewhere, where less viable seeds were produced and sexual reproduction rates were especially low (Alsos et al., 2002; Jadwiszczak et al., 2017).
Very low heterozygosity and inbreeding may reflect low rates of sexual reproduction, as found by paleobotanical analyses of Svalbard *B. nana* subsp. *nana* populations (Alsos et al., 2002). On Svalbard, *B. nana* subsp. *nana* was more widespread 6 to 8 kyr BP, followed by climatic cooling 4 to 2.5 kyr BP and a subsequent decrease in sexual reproduction leading to higher rates of inbreeding and population fragmentation resulting from genetic drift (Birks, 1991; Alsos et al., 2002). Reduced sexual reproduction of shrubs, constrained by periodic cooling and a propensity for clonality, would likely lead to more continuous populations and low levels of genetic variation (Alsos et al., 2002). In general, low heterozygosity and low allelic richness across *B. nana* populations (as compared to *E. vaginatum*) support more recent population expansion with high genetic connectivity and low genetic differentiation across sites. These patterns are similar to what was found for *S. herbacea* populations in the Swiss Alps by Cortés et al. (2014). Lack of genetic differentiation across *B. nana* populations may also be attributed to wind dispersal of seeds facilitating gene flow over long distances, thereby increasing genetic connectivity and homogenizing genetic diversity across populations due to infrequent outcrossing of long-lived clones and/or recent co-ancestry.

The low levels of co-ancestry of NN with the other sites may be due to lower rates of contemporary gene flow, the presence of *B. glandulosa*, or an admixed variant between *B. nana* and *B. glandulosa* (De Groot et al., 1997; Eidesen et al., 2015). Many private alleles found across sites may also be partly explained by the presence and/or hybridization with *B. glandulosa* as well. There has been taxonomic uncertainty about whether *B. glandulosa* should be considered a unique species from *B. nana* (Hultén, 1968; Fredskild, 1991). Plants collected at NN had different morphology than those at other sites, including larger and dense glands on twigs, but these
characteristics show high variation across the range. Leaf shape has been used to distinguish the species as well, but in general, leaf characteristics can be highly variable, at times matching those that have been considered common for each species within a single plant. Yet, our results show an inordinately high number of private alleles (111) compared to all other sampled populations, and fineRADstructure and DAPC both support differentiation of the population from the other sampled populations. The observed variation may be supporting evidence for differentiation of *B. glandulosa*. However, a much broader sampling for both morphology and molecular data will be required to better evaluate species limits. The intermediate characters seen in this population may also indicate admixture. If this is the case, it would be important to actively sample more populations with *B. glandulosa* morphology to be able to identify admixed individuals or populations from parent species populations. While further taxonomic investigations need to be conducted, these patterns of population structure are likely influenced by the presence of *B. glandulosa* and/or hybridity between *B. nana* and *B. glandulosa* (especially in NN), as the range of *B. glandulosa* overlaps with *B. nana* in Alaska and hybridization in *Betula* is common (De Groot et al., 1997; Eidesen et al., 2015). Extensive sharing of chloroplast DNA haplotypes among *Betula* species, including *B. nana* subsp. *nana*, *B. pendula*, and *B. pubescens*, have been explained primarily by frequent hybridization and introgression (Palme et al., 2004; Eidesen et al., 2015). Taxonomic confusion and difficulty in species identification of *Betula glandulosa* and *B. nana* have been well-documented due to plasticity, lack of species boundaries, or hybridization with each other and other *Betula* species in their range (De Groot et al., 1997).

As the Arctic warms, low seedling recruitment, narrower environmental tolerances, adaptational lag, and limited gene flow between *E. vaginatum* populations south and north of treeline may inhibit the ability of *E. vaginatum* to adjust its range (Mcgraw et al., 2015; Stunz et
al., 2022). In contrast, more continuous populations, a general lack of local adaptation, and developmental plasticity should facilitate range expansion of *B. nana*. Genotype-environment associations found for *E. vaginatum* populations (Stunz et al., 2022) indicate significant relationships between clinal allelic variation and environmental predictors across the latitudinal gradient and support adaptation of populations to unique genetic niches. Significant associations of environment and allele frequency turnover along the latitudinal gradient were not found for *B. nana* populations. The lack of genetic structure and niche specificity of genotypes could indicate higher tolerance for environmental shifts (plasticity) across the range. This could facilitate a competitive advantage as genotypes should be able to respond to shifting environmental conditions under climate change. Additionally, there were outlier loci (candidate genes for selection) associated with genes that have roles in abiotic stress response and sexual reproduction (Table 2.2). Specifically, lysine-specific demethylase ELF6 has a role in promotion of carpel cell elongation, which can increase the likelihood of sexual reproduction by increasing carpel growth (Keyzor et al., 2021). Association with the LRR gene family, which plays a role in stress response, was also found (Liu et al., 2020). The importance of these genes is not surprising as plants need to adjust to warming fluxes, and outcrossing of shrubs is expected to increase as the Arctic warms. Investigation of these genes and gene families would be useful to increase understanding of adaptation of *B. nana* in a warming arctic.

In conclusion, as permafrost thaws and the Arctic warms, *Betula nana* has the potential to alter dynamics of the MAT ecosystem if it becomes more prevalent and, along with other shrubs, displaces *E. vaginatum*. Increases in sexual reproduction, seed germination, and hybridization may also lead to increased genetic variability and population size as more soil nutrients become available, which may lead to further displacement of *E. vaginatum*. The low genetic variation
across *B. nana* sites could indicate a lack of ecotypic adaptations and support more plasticity in this species, which will likely confer a competitive advantage as it will be more readily adaptable to increasing temperatures. In addition, factors typically cited as restricting sexual reproduction in northern populations (i.e., cold temperatures and low nutrient availability) of arctic plants, especially shrubs, should not be as limiting (Douhovnikoff et al., 2010). Further analyses of local adaptation, paired with broader sampling, will be important to determine plasticity and lack of adaptive constraints that could lead to increased competitiveness and abundance of *B. nana* in a warming arctic.

**REFERENCES**


Small-scale patterns in snowmelt timing affect gene flow and the distribution of genetic
diversity in the alpine dwarf shrub *Salix herbacea*. *Heredity* 113, 233–239.

Phylogeography and population genetics of black alder (*Alnus glutinosa* (L .) Gaertn .) in

Cullings, K. W. (1992). Design and testing of a plant-specific PCR primer for ecological and


(GTMED2010).


Chapter 3: Comparative ENM modeling of arctic dwarf birch (*Betula nana*) and the arctic tussock cottongrass (*Eriophorum vaginatum*) in Alaska in the context of a review of overall graminoid and deciduous shrub distribution patterns in the Arctic
ABSTRACT

Improving understanding of the response of plant functional groups to a changing climate can inform how ecological communities may shift in the future. Here, I conduct a literature review of graminoid and deciduous shrub distribution patterns in the Arctic and use environmental niche modeling to investigate temporal variation of suitable habitat for *E. vaginatum* and *B. nana* in Alaska. Environmental niche models (ENMs) demonstrated small areas of moderately suitable habitat for *B. nana* during the LGM, with general increases in habitat suitability and suitable habitat area through the Mid-Holocene and the present, primarily in southern and west central Alaska. For *E. vaginatum*, highly suitable habitat decreased in cover from the LGM to the Mid-Holocene, and both moderately and highly suitable habitat continued to decrease in the present, with a suitable habitat shift northward since the LGM. Importantly, and as supported by the literature review, modeling species actual distributions is complex, and the incorporation of population-level and ecological community factors would improve predictions of realistic expansion and population persistence under climate change. While shrubs are expected to increase in height and density, and landscape resistance could hinder range expansion of graminoids as the Arctic warms, fine-scale environmental variation, nutrient availability, dispersal, gene flow, genetic differentiation, local adaptation, competition, and community structure will also shape both species’ distributions. While *B. nana* will likely expand rapidly in areas of highly suitable habitat in Alaska, as supported by a lack of local adaptation and narrow environmental tolerances (Chapter 2), moist and warm sites will likely see the greatest increases in density. While *E. vaginatum* may be able to dominate in colder and/or wetter sites, especially as deeper rooting will allow access to nutrients in thawed permafrost layers, *E. vaginatum* extent is not expected to increase in the taiga, where landscape resistance is high, treeline reduces gene flow between
populations, and competition with shrubs will likely limit population expansion in the warming arctic tundra.
**INTRODUCTION AND REVIEW**

Arctic flora populations occur in either arctic tundra or taiga ecosystems and are often characterized by low genetic diversity. Shrubby and herbaceous plants that predominated arctic forests and southern mountain ranges during the late Tertiary are believed to be the ancestors of present-day arctic plant species (Hultén, 1937; Murray, 1995). The Arctic supported continuous forest across Asia and America during the Tertiary and potentially into the early Pleistocene, as cooling cycles and continental glaciation shifted the arctic tundra to a treeless landscape (Repenning and Brouwers, 1992; Murray, 1995). The current arctic flora comprises approximately 2,218 vascular species (Elven et al., 2011), and among these, two key plant functional types, deciduous shrubs and graminoids, are well-represented across plant communities (Walker et al., 1989; Chapin et al., 1996). The influence of foundation species such as the graminoid *Eriophorum vaginatum* on gross primary productivity (GPP) of ecosystems in the Arctic has been well-documented (Souther et al., 2014; Parker et al., 2017, 2021; Curasi et al., 2019, 2022). If current climate change leads to community level shifts, such as increasing shrub density (Tape et al., 2006; Bjorkman et al., 2018; Wilcox et al., 2019), this could change ecosystem productivity.

Environmental factors can influence the distribution of graminoids and shrubs, such as differences in rooting depth and seasonal growth patterns. In the tundra, graminoids’ root growth continues during the growing season, and fine roots can extend to depths of 40 cm, often into soil horizons with recently thawed permafrost (Wang et al., 2016, 2018; Blume-Werry et al., 2019). In contrast, Wang et al. (2016) showed that at a Siberian tundra site, deciduous shrubs have high fine root biomass at the beginning of the growing season, as well as a shallower rooting pattern, with roots rarely extending deeper than 15 cm, allowing these species to take advantage of available soil nutrients in shallow soil layers immediately following snowmelt. In northern Sweden, Blume-
Werry et al. (2019) found that while dwarf shrubs such as *Vaccinium uliginosum* and *Betula nana* grew roots deeper in permafrost thaw treatments than in control treatments, they were unable to reach the thawed permafrost layer (Blume-Werry et al., 2019). In contrast, the graminoid *Eriophorum vaginatum* was able to track receding thaw and grow roots deeper (> 60 cm) into newly thawed permafrost later in the growing season (Blume-Werry et al., 2019). Increased root litter input also corresponded with deeper permafrost thaw, which has implications for increased C loss from root litter and exudates, and even permafrost C if root litter is deep enough. Graminoids like *E. vaginatum* with deciduous roots may stimulate microbial decomposition especially (Bennington et al., 2012; Blume-Werry et al., 2019). Differences in rooting depth of graminoids and shrubs supports niche differentiation in soil nutrient uptake, which could play a large role in influencing species distribution patterns at a landscape scale and will likely have long-term implications with climate change and deeper active layers.

Shrubs have been shown to influence active layer thickness in the arctic tundra, with more pronounced influence expected as shrubs increase in height, density, and abundance (Tape et al., 2006; Myers-Smith et al., 2011; Bjorkman et al., 2018; Wilcox et al., 2019). Different shrub species have also been shown to affect snowmelt timing in varying degrees in the Arctic. According to a study in the tundra of Northwest Territories, Canada, areas dominated by shrub birch (*Betula glandulosa*) became snow-free significantly earlier than areas dominated by alder (*Alnus alnobetula*) and other types of vegetation (Wilcox et al., 2019). Earlier snowmelt in birch-dominated areas resulted in a deeper active layer, regardless of snow depth or slope. Protrusion of birch branches through the snowpack, leading to lower albedo and less snow accumulation than tundra or areas dominated by alder, partially explained the earlier snowmelt (Wilcox et al., 2019). Snowfall is predicted to decrease across the Arctic under climate change (Walsh et al., 2020; Henry
et al., 2022), and increases in shrub prevalence, height, and density will likely result in more shrub protrusion through the snowpack, leading to earlier snowmelt and a deeper active layer (Wilcox et al., 2019). A deeper active layer in the tundra may permit deeper rooting and increases in aboveground cover of *E. vaginatum*, as found in a decade-long permafrost thaw experiment (Blume-Werry et al., 2019). While the greatest biomass increases were belowground, the abundance of *E. vaginatum* nearly doubled in deep thaw treatments, as compared to control plots, while deciduous shrub abundance did not differ significantly between treatments (Blume-Werry et al., 2019). These results suggest that while shrub height and cover are generally expected to increase as the Arctic tundra warms, microenvironmental variables such as soil moisture and site temperature will likely be important as the greatest increases in shrub cover were seen in warmer, moist sites (Chapin et al., 2005; Elmendorf et al., 2012; Henry et al., 2022). Graminoids may be able to dominate in colder and/or wetter sites, especially as deeper rooting will allow access to nutrients in thawed permafrost layers (Wang et al., 2016, 2018; Blume-Werry et al., 2019). Arctic sites in warmer, moist areas could see drastic changes to tundra plant communities as shrubs increase, including loss of bryophytes and lichens (Elmendorf et al., 2012; Henry et al., 2022).

While shrubs and graminoids have stark morphological and physiological differences, phenotypic plasticity has been documented for both functional groups (Bret-Harte et al., 2001; Deslippe and Simard, 2011; Kreyling et al., 2019; Dobbert et al., 2021). Phenotypic plasticity confers flexibility in growth strategy in response to environmental variation, which can be particularly advantageous in a changing climate (Bret-Harte et al., 2001). High adaptability and seasonal microenvironmental variation have been demonstrated for shrubs *Betula nana* and *Empetrum nigrum* ssp. *hermaphroditum* in the Low Arctic of Norway, suggesting phenotypic plasticity of these species (Dobbert et al., 2021). By monitoring radial stem growth of the shrubs
for five years, responsiveness to seasonal fluctuations in temperature and soil moisture availability was detected for both species. *Betula nana* showed increased growth during years with warmer winters, high snowfall, and early snowmelt and responded negatively to high summer temperatures with soil temperatures exceeding 10 °C (Dobbert et al., 2021). While phenotypic plasticity has been described more often for shrubs, it has also been documented in graminoids, such as for the grass *Arrhenatherum elatius*, after exposure to winter warming and frost stress treatments in sites across northern Europe (Kreyling et al., 2019). In general, *A. elatius* populations in warmer sites with fluctuating climates showed higher phenotypic plasticity. (Kreyling et al., 2019). These results are comparable to those found for *E. vaginatum* across a latitudinal gradient in north central Alaska, in which populations from warmer sites south of treeline had higher phenotypic plasticity (Fetcher and Shaver, 1990; Bennington et al., 2012; Parker et al., 2021).

How these plants will be able to persist or expand in the future arctic under climate change will, in a great part, depend on past and current propagule movement across the landscape. Some of the most common and widespread deciduous shrub (e.g., *Alnus* spp., *Betula* spp., and *Salix* spp.) and graminoid (e.g., *Carex* spp. and *Eriophorum vaginatum*) taxa in the Arctic are primarily wind-dispersed and wind-pollinated (though insect pollination prevails for some *Salix* spp.) (Bliss, 1971; Kevan, 1972; Argus, 1974). Wind pollination and dispersal are generally believed to promote widespread gene flow and low genetic differentiation. However, loss of habitat, low levels of seedling recruitment, and evolutionary forces (i.e., genetic drift and selection) with environmental isolation can limit connectivity and increase population genetic differentiation. Low seedling recruitment has been described for both graminoids and shrubs in the Arctic (e.g., *Eriophorum vaginatum* and *Betula nana*) (Chester and Shaver, 1982; Fetcher and Shaver, 1982; McGraw and Shaver, 1982). *Eriophorum vaginatum* and *B. nana* also have long lifespans (~187 years for *E.*
and up to ~147 years for *B. nana* (Miller, 1975) and reproduce vegetatively, which are important factors contributing to population persistence and can compensate for low seedling recruitment (De Witte and Stöcklin, 2010). Low recruitment can lead to lower genetic diversity of populations, as individuals with a low progeny survival rate will also have a lower chance of spreading their genes.

Here, I compare genetic diversity, genetic differentiation, population structure, ecotypic variation, and genotype-environment association findings for graminoid and deciduous shrub functional types in the scientific literature. Then, a comparison is made between one representative species for each group, the graminoid *Eriophorum vaginatum* and the deciduous shrub *Betula nana*, based on environmental niche models (ENMs). This comparison will include past, present, and future ENMs for the two species concluding with an assessment of how each group might persist in a warming arctic.

**Distribution of genetic diversity in the Arctic**

In general, arctic plants are expected to have higher genetic diversity in historically refugial populations (i.e., Beringia) and lower genetic diversity in more recently deglaciated regions due to founder effects, population bottlenecks, and genetic drift during and after postglacial range expansion (Hewitt, 1996; Abbott and Brochmann, 2003; Alsos et al., 2005; Brubaker et al., 2005). These expectations have been supported by population genetic analyses of shrubs, forbs, and graminoids in taiga and tundra ecosystems on a circumarctic scale. They have been attributed to similar glacial and post-glacial patterns of range shifts, demographic processes, and persistence in glacial refugia (Stenström et al., 2001; Taberlet et al., 2012; Eidesen et al., 2013). Wind-pollinated plant species with wide geographical ranges are expected to retain higher genetic diversity as
compared to animal-pollinated species, selfing species, and/or species with narrow geographical ranges (Hamrick and Godt, 1990).

**Graminoids.** Large populations of wind-pollinated and wind-dispersed plants are expected to have similar genetic diversity and low genetic differentiation due to long-distance dispersal events and high rates of gene flow across populations (Alsos et al., 2005, 2009, 2015). Extreme long-distance dispersal events of graminoids in the Arctic are common and have been extensively documented for *Carex* spp. (Jonsson et al., 1996; Stenström et al., 2001; Schönswetter et al., 2008). For instance, Jonsson et al. (1996) described the colonization of lava fields by *C. bigelowii* in Iceland, which was supported by phylogeographic analyses and the absence of local source populations. High levels of allelic variation and low genetic differentiation were also found for three *C. bigelowii* sites in Iceland, separated by up to 35 km, suggesting extensive gene flow over long distances (Jonsson et al., 1996). Relatively high genetic diversity within populations was found for *Carex* spp. populations in Siberia and Sweden (Stenström et al., 2001; Schönswetter et al., 2008), and genetic variation, both within populations and within taxa, was high for *C. bigelowii, C. lugens, C. ensifolia, and C. stans*, consistent with patterns seen for wind-pollinated and outcrossing taxa with wide geographical ranges (Hamrick and Godt, 1990; Stenström et al., 2001). Consistent with other graminoid populations, genetic diversity was similar, though relatively low, across 17 *E. vaginatum* sites along a latitudinal gradient in north central Alaska (Stunz et al., 2022). However, genetic structure was found between arctic ecosystems, while overall genetic differentiation was low (Stunz et al., 2022). Most studies show a homogenizing effect of gene flow across graminoid populations in the Arctic, leading to low genetic differentiation and the maintenance of similar genetic diversity across populations.
**Shrubs.** Long-distance dispersal events have also been frequently documented for deciduous shrubs in the Arctic. For example, Fredskild (1983) estimated *Alnus crispa* arrived on Greenland ~4 kyr BP via long-distance dispersal events from populations in Labrador, in eastern Canada. Multiple long-distance events have been described for *Salix herbacea* at a circumarctic and circumboreal scale. Alsos et al. (2009) investigated the genetic diversity and population structure of *S. herbacea* populations along an east-west Atlantic transect using amplified fragment length polymorphisms (AFLPs). Five groups related to geographic regions along the sampled longitudinal gradient were identified using population structure analyses, with a deep evolutionary split occurring in Greenland (Alsos et al., 2009). Genetic diversity in the eastern cluster (Europe and eastern Greenland) was generally higher than in the western cluster (western Greenland and eastern Canada), with extensive population connectivity supported by the relatively low number of private alleles in both clusters (Alsos et al., 2009). Fossil records, modeling, and genetic data supported a general northward postglacial colonization of *S. herbacea*, including colonization of Iceland and Greenland during the Holocene (Alsos et al., 2009). Yet, genetic connectivity isn’t always high across shrub populations in the Arctic, as found for *B. nana* populations on Svalbard (Alsos et al., 2002). While *B. nana* populations were less fragmented and larger than other thermophilous plant populations, genetic diversity was lower and genetic differentiation was higher among populations than the averages found for wind-dispersed, outcrossing, wind-pollinated species, including other woody perennials (Hamrick and Godt, 1990; Alsos et al., 2002). While these results support an overall trend of high population connectivity across deciduous shrub populations in the Arctic maintained by high levels of gene flow and long-distance dispersal, trends can vary depending on the species and sampled populations.
Relatively high genetic diversity is also expected within populations of species with high production of viable seeds and populations influenced by multiple dispersal events, both of which have been documented for shrub species *Salix herbacea* and *Betula nana* at multiple circumarctic sites (Alsos et al., 2007; Huebner and Bret-Harte, 2019). On the other hand, genetic evidence of multiple dispersal events has not been demonstrated for graminoid populations in the Arctic. Recent warming in the Arctic (e.g., ~16 to ~14 kyr BP (Bartlein et al., 1991; Brubaker et al., 2005), ~10 kya BP (Hamilton, 1986; Serreze et al., 2000; Kaufman and Manley, 2004; IPCC, 2014) has been cited as particularly important for increases in *Betula* spp. abundance and increased sexual reproduction (especially in the High Arctic and populations near the northern edge of the growing range), which could lead to domination of shrub species in arctic seedbanks (Huebner and Bret-Harte, 2019), increased dispersal events, higher rates of seedling recruitment, and increased genetic diversity. Establishment of seedlings with subsequent clonal growth could lead to increased shrub population establishment and/or higher genetic diversity in continuous populations (Alsos et al., 2007; Huebner and Bret-Harte, 2019), especially in microhabitats with warm summers and mesic soil conditions (Elmendorf et al., 2012; Cahoon et al., 2016; Ackerman et al., 2017; Bjorkman et al., 2018).

**Population structure in the Arctic**

Population structure of wind-pollinated and wind-dispersed shrub and graminoid species in the Arctic is expected to be weak due to long-distance pollination and dispersal events increasing population connectivity (Hamrick and Godt, 1990; Stenström et al., 2001; Alsos et al., 2015; Chau et al., 2019). In general, populations of long-distance dispersing plants are expected to be structured by geological processes and major physical and climatic barriers (Sanmartín and
Ronquist, 2004; Lee et al., 2017) at large time and spatial scales, while local dispersal, heterogeneous landscape, geographical, and environmental features can influence structure at a finer spatial scale (Manel et al., 2003; Eidesen et al., 2013; Chau et al., 2019). While landscape-level studies of population dynamics in the Arctic have not been examined at a broad scale, there has been some recent progress (Wang et al., 2021; Alsos et al., 2022).

**Graminoids.** Structure has been reported for multiple graminoid species across the Arctic, and is often attributed to geological and ecological barriers to gene flow. Genomic data (ddRAD SNP data) were used to investigate genomic diversity and identify population structure and adaptation of the species *Carex scirpoidea* across sites in the northern United States and Canada (Bard et al., 2021) and across North America and Europe (Westergaard et al., 2019), including two arctic sites. Bard et al. (2021) found evidence of genetic differentiation and adaptations related to ecosystem differences, as three groups were identified across sites corresponding to geographic regions along a longitudinal gradient. Additionally, a study examined genetic variation using AFLP markers for two graminoid species across continuous Arctic and non-Arctic regions (Birkeland et al., 2017). Population structure analyses delineated two groups in *C. capillaris*, which corresponded to two assumed subspecies, *C. capillaris* subsp. *fuscidula* and *C. capillaris* subsp. *capillaris*, as well as two groups for *Kobresia simpliciuscula*, corresponding to *K. simpliciuscula* subsp. *simpliciuscula* and subsp. *subholarctica* respectively (Birkeland et al., 2017). The split between *K. simpliciuscula* and *C. capillaris* subspecies was primarily associated with geographical and ecological barriers. The *K. simpliciuscula* subspecies were split between the sub-Arctic areas of Europe and arctic sites (Elven et al., 2011). Similarly, the split between *C. capillaris* subspecies was between arctic areas and non-arctic areas of Iceland, Greenland, and Canada and alpine and temperate areas of Europe (Elven et al., 2011). Ecological barriers,
primarily between arctic and alpine (or temperate) regions, were found to be especially important in supporting the split between *C. capillaris* subspecies, proving to have more of an influence than the geographic distance that separated populations by the Atlantic Ocean.

Ecological barriers can also be important for population structure at a finer sampling scale, as supported by a landscape genomics investigation of *Eriophorum vaginatum* (Stunz et al., 2022) in north central Alaska. Three populations/clusters were found across 17 *Eriophorum vaginatum* sites, and while one cluster was recognized at the southeast extreme of the sampling range at a higher elevation site, the remaining sites clustered into a population south of treeline or a population north of treeline in arctic tundra, supporting treeline as a major ecological and gene flow barrier for *E. vaginatum* in the region (Stunz et al., 2022). Landscape resistance, attributed primarily to forest cover, was also found to be important in explaining general disjunction and lower genetic connectivity of sites south of treeline (Stunz et al., 2022). In contrast to more structurally fragmented landscapes in temperate and alpine regions found at lower latitudes, landscape resistance tends to be lower in the arctic tundra, where large expanses of open landscapes and strong winds occur (Alsos et al., 2015). Lower landscape resistance in arctic tundra is also supported by high genetic connectivity of *E. vaginatum* sites occurring north of treeline in north central Alaska (Stunz et al., 2022)

**Shrubs.** Most population genetics investigations of shrubs in the Arctic have examined broad-scale patterns (Alsos et al., 2009, 2012) and have shown similar trends to graminoids (Birkeland et al., 2017; Westergaard et al., 2019). Only a few studies have examined landscape-scale population structure of shrubs in the Arctic. At the landscape level, weak population structure and relatively high genetic diversity were found for *Betula nana* on the Taymyr Peninsula in the Siberian forest-tundra ecotone (Meucci et al., 2021). Genetic diversity was higher for *B. nana*
populations than for other shrubs (*Alnus alnobetula* and *Salix* spp.) in the study region and was attributed to efficient seed dispersal and genetic connectivity between multiple populations (Meucci et al., 2021). Similar patterns were also found for three populations of *B. nana* on Svalbard using isozyme diversity markers (Alsos et al., 2002). However, for *B. nana* in north central Alaska (Chapter 2) across the same range as *E. vaginatum*, structure was not found except for one differentiated population that may indicate a taxonomic boundary (Stunz et al., 2022). Moderate genetic diversity and weak population structure suggest that while inbreeding is frequent, relatively high genetic diversity is maintained across *B. nana* populations. These results suggest that overall rates of pollen dispersal, subsequent pollination, and viable seed production and germination are higher and/or more successful for shrubs than graminoids, and ecological barriers may also play a stronger role in shaping graminoid population structure, given the limited landscape-level studies conducted so far in the Arctic. Vegetative reproduction and aggressive growth response of shrubs (De Groot et al., 1997; Bret-Harte et al., 2001; Myers-Smith et al., 2011) may also lead to larger, continuous shrub populations, facilitating pollen and seed dispersal within and across populations.

Generally, sexual reproduction and seedling recruitment is considered to be low for most arctic plants (Bliss, 1958), and vegetative reproduction is common for many species, including shrubs and graminoids, in arctic and alpine ecosystems (Shaver et al., 1979; Fetcher and Shaver, 1982; Jonsson et al., 1996; Douhovnikoff et al., 2010; Zhou et al., 2019). Yet, clonal growth of shrubs in the Arctic has been cited as especially important for retention of genetic diversity from past sexual reproduction and a means of dominating ecosystems via aggressive spread of stands and formation of dense canopies, especially under warming (Bret-Harte et al., 2001; Alsos et al., 2009; Swanson, 2015). For instance, sexual reproduction of *Salix herbacea* is considered rare across its growing range, and stands of the shrub mostly spread vegetatively (Alsos et al., 2009).
The moderate genetic diversity of *B. nana* Svalbard populations and *B. glandulosa* populations on Baffin Island (Canada) were both attributed to past sexual reproduction and the subsequent retention of genetic diversity by long-lived clones (Hermanutz et al., 1989; Alsos et al., 2002). However, clonality was not found to be common among *B. nana* populations in Alaska (Chapter 2). *Betula glandulosa* clonal populations on Baffin Island were also found to have low to no viable seed production and seedling recruitment, which was supported by low genotypic diversity within sites and low genetic differentiation across sites (Hermanutz et al., 1989). Cold temperatures and low nutrient availability have been cited as primary reasons for infrequent sexual reproduction in arctic plants, especially shrubs, and increasing documentation of sexual recruitment may be indicative of a transition to sexual reproduction as the Arctic warms (Douhovnikoff et al., 2010).

**Arctic landscape genetics, ecotypes, and genotype-environment associations**

Ecotypes evolve when natural selection acts on broadly distributed species, leading to locally-adapted populations across niches (Jain and Martins, 1979). Plant species may effectively cope with environmental variation at a landscape scale via phenotypic plasticity, the evolution of ecotypes, or a combination of both (Sultan, 1995; De Jong, 2005; Geng et al., 2007). Common garden and reciprocal transplant experiments have often been used to disentangle phenotypic plasticity and ecotypic variation in arctic and alpine plant populations (Clausen et al., 1940; Chapin and Chapin, 1981; Shaver et al., 1986; Fetcher and Shaver, 1990), with the more recent inclusion of genetic data to identify locally-adapted genotypes, or ecotypes (Geng et al., 2007; Stunz et al., 2022). While ecotypic variation can be advantageous in a stable environment, environmental change can lead to a mismatch between ecotype and optimal climate, or adaptational lag (Aitken et al., 2008; Mcgraw et al., 2015). Phenotypically plastic individuals, on the other hand, are defined
by their environmentally-induced variation and, therefore, are better equipped to cope with climatic instability (Sultan, 1995). Ecotypes have been described for multiple graminoid species in the Arctic (Chapin and Chapin, 1981; Bennington et al., 2012) and attributed to pressures of environmental selection (Curasi et al., 2019) and local adaptation (Stunz et al., 2022). In contrast, ecotypes in arctic shrub species are less expected due to generally high population connectivity and large clonal populations (Alsos et al., 2002; Chapter 2), often characterized by phenotypically plastic attributes (Bret-Harte et al., 2001; Deslippe and Simard, 2011; Dobbert et al., 2021).

**Graminoids.** Investigating ecotypes of graminoid species has been a focus of researchers for decades in the Arctic, notably in Alaska. For example, Shaver et al. (1979) examined the presence of ecotypes of *Carex aquatilis* across five ice-wedge polygons occurring in wet meadow tundra in Barrow, Alaska. Evidence of ecotypic differentiation was found for each of the five populations in terms of allocation, metabolism, and growth requirements. Chapin and Chapin (1981) further examined the presence of ecotypes of *C. aquatilis* at a local scale and along a latitudinal gradient extending from Barrow, Alaska, to north central Colorado. Five reciprocal transplant gardens were established, three in Alaska and two in Colorado, and while the short duration of the experiment (three years) likely did not allow for ecotypes to fully respond to novel environmental conditions, home-site advantage was found for individuals from arctic and muskeg sites in Alaska (Chapin and Chapin, 1981), supporting ecotypic differentiation across sites. In general, environmental differences and adaptive differences across sites supported differentiation of ecotypes of *Carex* spp. in Alaska.

Bard et al. (2021) and Westergaard et al. (2019) used phylogenomic and population genomic analyses to support ecotypic differentiation of *C. scirpoidea* populations. In an
investigation of two *C. scirpoidea* subspecies, subsp. *scirpoidea* and subsp. *convoluta*, in North America, evidence of restricted gene flow between the subspecies combined with morphological and ecological distinction of subsp. *convoluta* supported the reclassification of subsp. *convoluta* as a *C. scirpoidea* ecotype (Bard et al., 2021). Westergaard et al. (2019) identified three genetically differentiated *C. scirpoidea* populations in Norway, all of which could be considered ecotypes, as supported by local adaptation, strong population structure, and low levels of genetic diversity (Westergaard et al., 2019). These results are consistent with the results of ecological studies of Shaver et al. (1986), which have identified differences in ecotypes of *E. vaginatum* across a latitudinal gradient in north central Alaska.

Ecotypic differentiation of *E. vaginatum* in Alaska has been well-documented over the last four decades (Shaver et al., 1986; Fether and Shaver, 1990; Bennington et al., 2012; Peterson et al., 2012; Souther et al., 2014; Parker et al., 2017, 2021; Stunz et al., 2022). Reciprocal transplant gardens have been a key component of research highlighting ecotypic differentiation of *E. vaginatum*. For example, tussock cottongrass individuals that were transplanted back into their home site along a latitudinal gradient had generally higher survival rates, flower production, and biomass than plants from "away" sites (Bennington et al., 2012; Souther et al., 2014). Likewise, light-saturated photosynthetic rate and stomatal density were correlated with the latitude of population origin (Peterson et al., 2012; Souther et al., 2014). More broadly defined ecotype-specific differences were correlated with site of origin occurring either north or south of the Brooks Range in northern Alaska (Bennington et al., 2012; Fether & Shaver, 1990; Shaver et al., 1986). Reciprocal transplant experiments also showed plants from the south of the Brooks Range exhibited higher plasticity and survival when moved north, while the inverse occurred when plants from the north were moved south (Bennington et al., 2012; Peterson et al., 2012; Parker et al.,
Population genomic studies have shown evidence for adaptation across *E. vaginatum* sites in north central Alaska, as genetic breaks between ecotypes also corresponded to ecosystem differences between taiga and tundra biomes at treeline (Stunz et al., 2022). Genotype-environment associations (GEAs) can be used to elucidate local adaptation of ecotypes by investigating the clinal allelic variation of genotypes and environmental variables. To further examine variation in adaptation of ecotypes, Stunz et al. (2022) used GEAs to identify environmental predictors that influence allelic turnover and correspond to a shift between ecosystems at treeline.

**Shrubs:** Phenotypic plasticity is well-supported in arctic shrub species and has been cited as a primary driver of morphological variation across populations (Chapin et al., 1995, 1996; Bret-Harte et al., 2001; Deslippe and Simard, 2011; Dobbert et al., 2021). For example, the responsiveness of radial stem growth to changes in soil moisture and temperature supported phenotypic plasticity of shrubs *Betula nana* and *Emetrum nigrum ssp. hermaphroditum* in the Low Arctic of Norway (Dobbert et al., 2021). As shrub populations tend to be large and have high genetic connectivity in the Arctic, the influence of environmental predictors on clinal allelic variation isn’t expected to be as pronounced or significant. These expectations have been supported by results of spatial genotype associations of *Salix herbacea* in alpine sites. Cortés et al. (2014) assessed gene flow and genetic differentiation of *S. herbacea*, a widely distributed arctic and alpine shrub species, using SSR loci from 12 sites in either snow bed or ridge microhabitats across three mountains in the Swiss Alps. While genetic diversity at snow bed sites was significantly higher than at ridge sites, subpopulations were not genetically differentiated due to high levels of gene flow across sites (Cortés et al., 2014). Although the study did not take place in the Arctic, the geographic heterogeneity and microenvironmental variation are similar in the Swiss Alps alpine habitat and the taiga ecosystem in the Arctic (Cortés et al., 2014). While locally-
adapted populations might be expected in ecosystems with high habitat variability, the high connectivity of S. herbacea populations supports the weak influence of landscape resistance on gene flow of deciduous shrub populations in taiga ecosystems.

In the Arctic, Betula nana populations are generally large and unstructured, as supported by population genomics of B. nana in north central Alaska, in which treeline did not function as a barrier to gene flow between populations in taiga and tundra biomes, as opposed to findings for the graminoid E. vaginatum with a similar distribution (see Chapter 2). Either increased population connectivity or historical processes post-glaciation in the Arctic has resulted in a lack of strong genetic differentiation in B. nana as compared to E. vaginatum.

**Artic plant distributions under climate change**

A general consensus that elevated growing season temperature and increased soil nutrient availability promote the dominance of shrubs in tundra ecosystems is prevalent in arctic research (Shaver et al., 2001; Elmendorf et al., 2012; Heskel et al., 2014; Leffler et al., 2016). Significant increases in reproduction and growth were found for shrubs, as compared to graminoids, in a review of data from 46 studies using open top chambers (OTC) to investigate arctic and alpine plant responses to experimental warming (Fazlioglu and Wan, 2021). DeMarco et al. (2014) also found a general trend of deciduous shrub dominance, as Betula nana and Salix spp. dominated all treatment plots while evergreen shrub, forb, and graminoid diversity declined in an investigation of plant community response to 18 years of warming and nutrient addition at sites near Toolik Lake in northern Alaska.

Other studies have shown site-specific plant functional group responses to OTC warming treatments and even opposing responses within functional groups to warming treatments.
(Elmendorf et al., 2012). In a synthesis of 61 circumarctic tundra warming experiments, the greatest increases in deciduous shrub (and total shrub) abundance occurred in relatively warm areas with moist soil, while graminoid abundance increased the most at colder sites (Elmendorf et al., 2012). While an initial increase in shrub abundance occurred at cold and dry sites, the increase was not sustained, and abundance increases in dwarf shrubs were found to decrease over time with warming (Elmendorf et al., 2012). Graminoids did not show consistent trends in abundance over time, and while sedges had the greatest long-term increases at wet sites, grasses increased most at dry sites, and rushes were unresponsive to soil moisture conditions (or responded negatively at dry sites) (Elmendorf et al., 2012).

Many recent studies are veering away from generalizations of plant functional group response to warming, citing fine-scale environmental variation, nutrient availability, dispersal, immigration, genetic differentiation, and local adaptation all as important factors contributing to arctic ecosystem responses to warming (Natali et al., 2012; Ackerman et al., 2017; Bjorkman et al., 2017, 2018; Myers-Smith et al., 2019). While functional groups have been the common focus for investigations of plant traits in arctic plant communities, many researchers have found that most variation can’t be explained at the functional group level (Elmendorf et al., 2015; Bjorkman et al., 2017; Myers-Smith et al., 2019). Local adaptation and genetic differentiation across populations have been documented for many widespread arctic species (Alsos et al., 2012; Bjorkman et al., 2017; Stunz et al., 2022), and barriers to gene flow or other restraints leading to mismatches in optimal climate along spatial gradients can limit community response to climate change (Elmendorf et al., 2015; Mcgraw et al., 2015). Bjorkman et al. (2018) used nearly 30 years of vegetation survey data from warming treatments at 117 tundra sites and found that the strongest spatial relationships between plant traits (such as height, specific leaf area, and leaf nitrogen (N))
content) and temperature and soil moisture gradients were mostly explained by species turnover. Community height increased across all sites as taller, thermophilous plant species increased in the tundra, particularly in areas where soil moisture was high (Bjorkman et al., 2018).

*Eriophorum vaginatum* vs. *Betula nana* under climate change

Many warming and nutrient addition experiments have been conducted to better understand the potential for range expansion and persistence of the graminoid *Eriophorum vaginatum* and deciduous shrub *Betula nana* in a warming arctic, given their prominence in the ecosystem. Often, negligible effects or decreased growth and abundance of *E. vaginatum* have been found in response to warming (Arft et al., 1999; Leffler et al., 2016; Parker et al., 2017, 2021). Arft et al. (1999) investigated the growth response of graminoids, including *E. vaginatum*, across 13 alpine, Low Arctic, and High Arctic sites to warming over the course of four years, but results were inconsistent as growth response only significantly increased during the second year of the experiment. In a 20-year warming and increased snowpack experiment at Toolik Lake Field Station in northern Alaska, an increase in shrub dominance, primarily by *Salix pulchra*, and a decline in *E. vaginatum* cover was reported for all treatments (increased snowpack, increased temperature, and combined) (Leffler et al., 2016). Increased snow insulation, greater thaw depth, and deeper soil drainage were all cited as important factors leading to shrub dominance, as areas without run-off would likely respond differently and not experience similar shrub abundance increases (Leffler et al., 2016).

Yet, site-specific conditions and experiment duration can have an impact on growth response, even for sites in close proximity. During a two-year warming experiment at Toolik Lake Field Station, *E. vaginatum* dominated in aboveground biomass and net productivity increase in soil warming plots, or the use of snow fences to trap layers of snow during the winter, while a
positive warming effect was not found for *B. nana* (Natali et al., 2012). Soil warming plots that increase N mineralization during the winter are expected to support earlier leaf-out in early season plants due to higher levels of available N in the early growing season (Natali et al., 2012). In a three-year experiment at Innnavait Creek (< 20 km from Toolik Lake Field Station), *E. vaginatum* increased by 36% in aboveground biomass with warming and accelerated leaf growth by 10 days in a 10-day snowmelt advance treatment (Livensperger et al., 2016). Early-season advances in green-up and leaf production with warming were also found for *E. vaginatum* during one growing season near Toolik Lake, but total leaf production was not significantly higher for warmed tussocks (Sullivan and Welker, 2005). Similar insignificant shifts in earlier green-up and/or delayed senescence with warming have been reported for *E. vaginatum* at sites in north central Alaska (Parker et al., 2017, 2021; Livensperger et al., 2019).

Ecotype-specific differences within species have been demonstrated as well, suggesting that adaptations within specific environments can affect response to climate change. Parker et al. (2021) demonstrated plasticity of an *E. vaginatum* ecotype from a site south of treeline compared to plants from sites north of treeline in north central Alaska. Over two growing seasons, reciprocal transplant gardens were used to investigate the response of three ecotypes. Though no significant differences in growth rate were evident between ecotypes in control or warming treatments, green-up timing of the ecotype south of treeline was found to be more responsive to snowmelt than ecotypes north of treeline (Parker et al., 2021). The southern ecotype also had a longer growing season and later senescence date than northern ecotypes at all sites and gardens (Parker et al., 2017, 2021).

Response of *B. nana* to warming has generally resulted in abundance increases and the formation of dense canopies (Bret-Harte et al., 2001; Deslippe and Simard, 2011; Deslippe et al.,
In Alaska, increases in \textit{B. nana} abundance have been demonstrated in response to warming treatments (Bret-Harte et al., 2001; Sturm et al., 2001; Stark et al., 2015), primarily due to alteration of biomass allocation (i.e., switching from short to long shoot production) (Deslippe & Simard, 2011; Deslippe et al., 2012; DeMarco et al., 2014). \textit{Betula nana} also increased branching rate in nutrient addition experiments, showing rapid and consistent growth after an initial lag of two to three years (Bret-Harte et al., 2002, 2004). Experiments investigating senescence timing of \textit{B. nana} at Imnavait Creek, Alaska, found that early snowmelt, warming, and combined warming and early snowmelt treatments resulted in earlier leaf appearance and expansion (one to eight day advance) and a one to four day earlier senescence, though the alteration in phenology was not considered significant (Livensperger et al., 2016, 2019). \textit{Betula nana}-dominated (rather than graminoid-dominated) communities in Greenland were found to be stronger carbon (C) sinks during a warm and wet growing season due to increased gross ecosystem photosynthesis and reduced ecosystem respiration, as influenced by larger \textit{B. nana} canopies retaining snow and cooling soil in summer (Cahoon et al., 2016). Yet, \textit{B. nana} communities could become C sources rather than sinks in a warm and dry growing season, resulting in losses of early-season C uptake due to drought-related reductions in gross ecosystem photosynthesis and potential ecosystem respiration increases (Cahoon et al., 2016). Whether arctic plant communities will become net C sinks or sources is uncertain and varies widely across climate change scenarios. For example, averaged high emission modeled scenarios predict a loss of 90% of permafrost in shallow (< 3 m depth) soils by 2300, and net uptake of CO$_2$ by vegetation may not be able to offset such a large C release from permafrost (Bruhwiler et al., 2021). In addition, future projections often fail to incorporate processes such as rapid permafrost thaw, winter CO$_2$ emissions, and extreme winter weather events that will influence C uptake and growth of vegetation in the Arctic (Bruhwiler et
Pearson et al. (2013) estimated an increase of up to 52% in woody cover across the Arctic under multiple climate scenarios by 2050, and if precipitation is not able to offset the loss of moisture in warming soils, evapotranspiration and respiration of *B. nana* and other deciduous shrub communities may contribute significantly to net C emissions in the Arctic (Bruhwiler et al., 2021).

Using modeling to predict how plant functional groups will respond to a changing climate can inform how ecological communities may shift in the future. Future geographical distribution models can also influence conservation measures to promote species persistence and protect populations with high genetic variability (Alsos et al., 2009, 2012; Çoban et al., 2020). However, as discussed above, functional group modeling may be inappropriate because responses can vary across species and ecosystems. Examining species-level distribution for the most common species in the Arctic may be particularly informative because of their influence on ecosystem structuring, productivity, and response. In light of the review above, I examine past, present, and future ENM models of *E. vaginatum* in relation to *B. nana* in Alaska to investigate how patterns of past distribution may influence species persistence in the future.

**METHODS**

Environmental niche models (ENMs) were used to create environmental suitability maps, test various demographic models, and predict the geographical distribution of *B. nana* and *E. vaginatum* suitable habitats in the past, present, and future within the range of the Beringian refugium.

To build an ENM, the species distribution was modeled for the ‘present’ (derived from records from 1970 to 2000) and projected into climate scenarios of the LGM (~20 kyr BP) and
future (2070). A Mid-Holocene (~6 kyr BP) climate raster was created by averaging current and LGM layer habitat suitability cell values. The initial ENM was built with 19 bioclimatic variables obtained from the WorldClim 2.0 Bioclimatic database (Supplementary Table S2; Fick and Hijmans, 2017) and derived from climatic records from 1970 to 2000 with the maximum entropy algorithm of MaxEnt v.3.4.3 (Phillips et al., 2006; Phillips and Dudik, 2008). MaxEnt uses species occurrence data and predictor variables to predict the probability distribution of a species. WorldClim data were downloaded at 30 arc seconds spatial resolution. *Betula nana* species occurrence data were obtained from the Alaska Vegetation Plots Database (Nawrocki et al., 2020), adding 3,923 occurrences to the 9 sampled sites used in this study, for a total of 3,932 unique species occurrence records (see Chapter 1 for LGM, mid-Holocene, and present ENM specifics for *E. vaginatum*).

ENMEVAL v.2.0.2 (Muscarella et al., 2014) was used to determine optimal parameters, feature class (FC), and regularization multiplier (RM) settings for MaxEnt. A total of 64 models were created using 8 FC (L, LQ, LQP, H, T, LQH, LQHP, LQHPT in which L = linear, Q = quadratic, H = hinge, P = product, and T = threshold) combinations (Muscarella et al., 2014), and eight RMs (0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0) (Manzoor et al., 2020). The “block” method was implemented to partition data into calibration and evaluation datasets in order to evaluate model performance (Muscarella et al., 2014; González-Serna et al., 2019). The model with the lowest Akaike Information Criterion corrected for small sample size (AICc; Burnham and Anderson, 2004) was identified as the optimal model and implemented for the final run in MaxEnt. To obtain a subset of uncorrelated environmental variables, a Pearson correlation coefficient cutoff ($r \geq 0.75$ and $r \leq -0.75$) was applied to remove variables for the final MaxEnt run (Dormann et al., 2013; Manzoor et al., 2020). Modeled LGM layers for retained environmental variables were derived.
from the Community Climate System Model (CCSM) (Braconnot et al., 2007) and included as a projection layer in MaxEnt. ARCMAP v.10.7.1 (ESRI, 2011) was used to create a Mid-Holocene climate raster by averaging current and LGM layer habitat suitability cell values.

Future ENMs were created for both *B. nana* and *E. vaginatum* using modeled layers for retained environmental variables derived from the Community Climate System Model (CCSM) Representative Concentration Pathway (RCP) 8.5 model for 2070 (Braconnot et al., 2007) and included as a projection layer in MaxEnt. RCP models vary depending on future emission scenarios, and the RCP 8.5 model comprises a high greenhouse gas concentration scenario in which the emission of greenhouse gases increases over time (San José et al., 2016). ARCMAP was used to classify cell values (0-1 range) for ENM map creation, and binning and normalization were not used in order to increase precision of ENM output. Raw ENM raster layers for each model/projection and both *B. nana* and *E. vaginatum* were re-classified into discrete values with the following scheme: ENM values ranging from 0 to 0.5 were classified as "not suitable," while values ranging from 0.5 to 1 were classified as "suitable." To support the calculation of suitable area, each re-classified raster was re-projected from a geographic coordinate reference system into the Alaska Albers Equal Area Conic projection (EPSG: 3338). Area was then calculated for "suitable ENM value" pixels in each raster in square kilometers (km²).

**RESULTS AND DISCUSSION**

Two ENMs were identified as best-fit, due to identical AICc values, for *B. nana*. One utilized a linear, quadratic, hinge, product, and threshold FC with a 0.5 RM, and the other employed a threshold FC with a 0.5 RM. Retained variables and their contributions were also identical for both models, and modeled suitable habitat was indistinguishable between ENMs, so
only the former was used for ENM maps. Environmental variables with the highest permutation importance in the *B. nana* ENM were mean temperature of the warmest quarter (Jun, Jul, Aug), precipitation of the wettest month (Aug), and mean temperature of the wettest quarter (Jul, Aug, Sept), contributing 29%, 23.2%, and 17.2% respectively.

The most supported ENM for *E. vaginatum* utilized a linear, quadratic, hinge, product, and threshold FC with a 0.5 RM. Variables with the highest permutation importance in the *E. vaginatum* ENM were temperature seasonality, maximum temperature of the warmest month (Jul), and precipitation of the coldest quarter (Dec, Jan, Feb) contributing 30.3%, 22.2%, and 15.2% respectively.

The contemporary ENM and climate scenario projections from the LGM (20 kyr BP) estimated 1,050.98 km$^2$ of suitable habitat for *B. nana* during the LGM in Alaska. Suitable habitat was predicted to cover 1,966.11 km$^2$ during the Mid-Holocene (6 kyr BP) and 462,505.59 km$^2$ in the present (Figure 3.1). Highly suitable habitat was not identified for *B. nana* during the LGM, and patches of moderately suitable habitat were found primarily in western Alaska. Moderately suitable habitat drastically increased during (the time between the LGM to) the Mid-Holocene across all of Alaska, with larger continuous areas in southwestern and west central Alaska extending across the Brooks Range. Projection of suitable habitat continued to increase in area across the state, with the addition of large swathes of highly suitable habitat in southern and west central Alaska, during (or between the Mid-Holocene and) the present.
Figure 3.1. MaxEnt environmental niche model (ENM) maps of Alaska, USA, adapted to depict (A) *Betula nana* and (B) *Eriophorum vaginatum* habitat suitability during the Last Glacial Maximum (LGM), Mid-Holocene, and ‘present’ (averages from 1970 to 2000 climate records). Modeled LGM layers were derived from the Community Climate System Model (CCSM) (Braconnot et al., 2007). Current and LGM layer habitat suitability bin values were averaged to create the Mid-Holocene climate raster.

Suitable habitat for *E. vaginatum* in Alaska was estimated to cover 717,451.23 km$^2$ during the LGM, 404,803.15 km$^2$ during the Mid-Holocene, and 199,540.85 km$^2$ in the present. Suitable habitat, especially highly suitable habitat, decreased in cover from the LGM to the present for *E. vaginatum*, and there was a general shift northward since the LGM. Suitable habitat was identified throughout the Beringian refugium for *E. vaginatum* (Figure 3.1), outside of the glaciated regions within the refugium (notably including north and south of the Brooks Range). Projection from the
mid-Holocene shows suitable habitat extended throughout most of northern Alaska, as well as across portions of the formerly glaciated Brooks Range.

Using future ENMs to determine the extent of suitable habitat for both *E. vaginatum* and *B. nana* in 2070 under the RCP 8.5 scenario produced unrealistic results and will not be reported here or included in Figure 3.1. While *E. vaginatum* may be able to dominate in colder and/or wetter sites under climate change, especially as deeper layers of permafrost thaw and deeper rooting will allow access to nutrients in these layers (Wang et al., 2016, 2018; Blume-Werry et al., 2019), the incorporation of population-level and ecological community factors would improve predictions of realistic expansion and population persistence, as generally *E. vaginatum* extent is not expected to increase in taiga ecosystems, where landscape resistance is high, treeline reduces gene flow between populations south and north of treeline, and competition with shrubs will likely limit population expansion in the arctic tundra under climate change (Elmendorf et al., 2012; Cahoon et al., 2016; Ackerman et al., 2017; Bjorkman et al., 2018; Stunz et al., 2022). Future ENMs are especially difficult to assess as future models cannot be validated, and there’s a lack of input data and parameters to correctly constrain the model. Additionally, the output from the RCP 8.5 and RCP 4.5 scenarios didn’t vary much for suitable habitat for *B. nana* or *E. vaginatum* in 2070. Curasi et al. (2022) produced ecological niche models predicting a dramatic decrease in *E. vaginatum* density on the Alaskan North Slope by 2100. By including more variables, such as edaphic properties (i.e., soil moisture, bulk density, and slope) in addition to the WorldClim dataset, (similarly to the approach in Curasi et al. (2022)), in ENMs should decrease uncertainty and more closely resemble expectations for range shifts of both *B. nana* and *E. vaginatum* in the future.
On a broad scale, winter precipitation and mean summer temperature are predicted to increase under climate change in the Arctic (Leffler et al., 2016). Increased frequency of exceptionally warm and cold growing seasons in addition to extreme weather events, such as warm winters with both snow and rain, are also expected (Walsh et al., 2020; Henry et al., 2022). In a warming arctic, landscape heterogeneity and microclimate variables will also be important to consider for shifts in abundance on a finer scale. The advance of treeline will likely increase landscape resistance and reduce population connectivity for many arctic tundra plant species (Dial et al., 2022). Soil moisture, in conjunction with warmer temperatures, will likely play a factor in increased abundance and growth of deciduous shrubs in the Arctic (Elmendorf et al., 2012; Cahoon et al., 2016; Ackerman et al., 2017; Bjorkman et al., 2018). In a dendroecological investigation of *S. pulchra* on the North Slope of Alaska, increased growth was found for *S. pulchra* under increased June temperature in both upland tundra and riparian sites over the past 25 years, yet a model of decelerating growth better explained growth at the upland sites, with sustained growth supported for *S. pulchra* at the riparian sites (Ackerman et al., 2017). Cahoon et al. (2016) found that a warm and wet growing season resulted in more rapid phenological advancement, canopy development, and carbon gain in *B. nana*-dominated communities than in graminoid-dominated communities. Soil moisture was also found to be more important than temperature alone for explaining spatial variation in multiple plant functional traits such as specific leaf area and leaf N content (Bjorkman et al., 2018). In general, moist soil conditions and warmer temperatures will likely facilitate shrub expansion, but wetlands and drier areas will likely see limited increases in deciduous shrubs (Elmendorf et al., 2012; Cahoon et al., 2016; Ackerman et al., 2017; Bjorkman et al., 2018).
Future warming-driven changes in traits and associated ecosystem functions will probably depend on soil moisture conditions at a site, as warmer growing season temperatures paired with moist soil conditions have been found to be especially important for increased abundance and sustained growth of *B. nana* and other deciduous shrubs in the Arctic (Elmendorf et al., 2012; Cahoon et al., 2016; Ackerman et al., 2017; Bjorkman et al., 2018). The effect of temperature on shrub growth can be variable, as multiple long-term studies using experimental warming treatments showed a weak or absent growth response of *Betula nana* and *Salix* spp. in comparison to herbaceous species in alpine, Low Arctic, and High Arctic sites (Chapin et al., 1995, 1996; Arft et al., 1999). However, *B. nana* has an uncommon advantage over other deciduous shrubs in its ability to produce two shoot types (both short and long), and in multiple studies near Toolik Lake, *B. nana* dominated treatment plots containing major tussock-tundra species (various graminoids, mosses, evergreen shrubs, and deciduous shrubs) when experimental warming was paired with nutrient addition (Bret-Harte et al., 2001; Chapin et al., 1996).

These findings are in strong agreement with environmental variables contributing most to ENMs of *B. nana* in Alaska. Variables with the highest permutation importance in the *B. nana* ENMs were mean temperature of the warmest quarter (29%), precipitation of the wettest month (23.2%), and mean temperature of the wettest quarter (17.2%). Summer temperature and summer precipitation have also been shown to influence the establishment of *B. nana* ramets (large, rooted branches) in the Siberian tundra, as establishment was found to be more frequent in areas with high summer precipitation, and declines in establishment corresponded to decreases in summer precipitation (Li et al., 2016).

Summer temperature and precipitation were found to be the most important for predicting *B. nana* occurrence, while temperature seasonality (30.3%), maximum temperature of the warmest
month (22.2%), and precipitation of the coldest quarter (15.2%) had the highest permutation importance for predicting the distribution of *E. vaginatum* in Alaska. Temperature seasonality, or the measure of annual temperature variation, and precipitation during the winter may be important for culm and spike (or stem and inflorescence) formation, as they are formed during the prior growing season in the Arctic before growth ceases in the fall (Wein, 1973). Culms begin to elongate, or ‘green-up,’ during the following spring, when water becomes available and soil is typically still frozen (Bliss and Peterson, 1992). Spring and summer temperature differences are also important for *E. vaginatum* vegetative phenology (Siegenthaler et al., 2013; Parker et al., 2017, 2021), which was also supported by GEA findings, in which many potentially adaptive SNPs were correlated with temperature of the driest quarter and temperature of the warmest quarter variables (Stunz et al., 2022). In the Arctic, the short growing season for most plants begins in late spring following snowmelt, as air temperatures and photosynthetic rates increase, and photosynthetic rate then remains relatively high from mid-June to mid-August, during the limited growing season (Defoliart et al., 1988). Similar physiological adaptations are common for arctic plant species and have been described for graminoids, forbs, and deciduous shrubs (Chapin, 1987; Chapin and Shaver, 1996).

While seasonal patterns and climate play an important role in determining suitable habitat for a species, there are multiple additional factors that contribute to its realized niche. For instance, human activities and geographic characteristics can affect population structure and gene flow of the same species in different locations (Alsos et al., 2002; Borrell et al., 2018). Though weak population structure and low to moderate genetic diversity have been reported for *Betula nana* across the Arctic (Alsos et al., 2002; Meucci et al., 2021; Chapter 2), population fragmentation and genetic drift can impact populations of this deciduous shrub in varying degrees. Inbreeding,
loss of habitat, and evolutionary pressures such as genetic drift could all contribute to increased genetic differentiation, which is consistent with patterns seen for highly fragmented *B. nana* populations outside the Arctic, such as in the UK (Alsos et al., 2002; Borrell et al., 2018).

Environmental niche models are also limited as they do not incorporate population-level attributes, such as genetic diversity, ecotypic variation, and population structure, all of which can play an important role in influencing the actual suitable distribution of a species. As reported in Chapter 2, relatively low genetic diversity and weak population structure were found for nine *B. nana* populations along a latitudinal gradient in north central Alaska, with population structure likely due to the presence of *B. glandulosa* or hybridity between *B. nana* and *B. glandulosa*. Ecotypic variation was not detected with population structure analyses (Chapter 2), and phenotypic plasticity of *B. nana* in the Alaskan tundra has been well-supported (Bret-Harte et al., 2001; Deslippe et al., 2011), suggesting that *B. nana* could thrive in a warming arctic by responding plastically to varying environmental conditions. Phenotypic plasticity of *B. nana* also relates to increased competitiveness and will likely allow the shrub to dominate and influence community structure throughout much of the Arctic. If these population-level factors were able to be incorporated in the ENM, the predicted suitable habitat would more accurately reflect the actual suitable habitat area of *B. nana* in Alaska, which will likely generally increase from the present to 2070.

Environmental niche models for *E. vaginatum* showed a general decrease in suitable habitat area from the LGM to the present. Ecotypes of multiple graminoid species in the Arctic have been described in the scientific literature over the last ~40 years (Shaver et al., 1979; Chapin and Chapin, 1981; Stunz et al., 2022), supporting local adaptation across populations, and as the Arctic warms, could lead to a mismatch between ecotypes and climate optima (Aitken et al., 2008; Mcgraw et
Population genetic investigations of *E. vaginatum* at a finer scale have supported population structure as well, with gene flow barriers and landscape resistance limiting population connectivity (Stunz et al., 2022). Predictions based on this genetic data are in agreement with environmental niche modeling for *E. vaginatum* from the LGM to the present and suggests limited population expansion of *E. vaginatum* in Alaska in the future, especially where landscape resistance is high.

In conclusion, a general increase in shrub height and density is expected across much of the Arctic, and low levels of contemporary gene flow and landscape resistance could hinder range expansion of graminoids in a warming arctic, but it’s not so simple. While ENMs can be useful to determine what environmental variables contribute most to shaping a species’ distribution, fine-scale environmental variation, nutrient availability, dispersal, gene flow, genetic differentiation, local adaptation, competition, and community structure all contribute to the realized niche of a species and its potential to respond to climate change. In Alaska, *B. nana* will likely expand rapidly in areas of highly suitable habitat, as facilitated by large populations and a large gene pool, in addition to a lack of local adaptation and narrow environmental tolerances (Alsos et al., 2002; Chapter 2 results). *Betula nana* was predicted to lose 5 to 8% of genetic diversity by 2080 on a circumarctic scale due to frequent long-distance dispersal and the relatively uniform distribution of genetic diversity across its range (Alsos et al., 2012). Expansion of *E. vaginatum* will likely be limited due to local adaptation/narrower environmental tolerances, landscape resistance, and barriers to gene flow (McGraw and Shaver, 1982; Mcgrew et al., 2015; Mohl et al., 2020; Stunz et al., 2022), though graminoid abundance may increase in wetland and colder sites, especially if more plastic and responsive ecotypes (i.e., from south of treeline) are introduced via assisted migration (Elmendorf et al., 2012; Parker et al., 2021).
REFERENCES


Çoban, H. O., Örücü, Ö. K., and Arslan, E. S. (2020). MaxEnt Modeling for Predicting the Current and Future Potential Geographical Distribution of *Quercus libani* Olivier.


Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role, in divergence and


Hultén, E. (1937). *Outline of the history of arctic and boreal biota during the Quaternary period: their evolution during and after the glacial period as indicated by the equiformal progressive areas of present plant species.* Stockholm.


Climate change drives expansion of Antarctic ice-free habitat. *Nature* 547, 49–54.


**Appendix**

**SUPPLEMENTAL TABLES**

**Table S1.** Mean annual temperature and mean annual rainfall (obtained from the WorldClim 2.0 Bioclimatic database (Fick and Hijmans, 2017)). 17 site abbreviations as for Table 1.

<table>
<thead>
<tr>
<th>Site</th>
<th>Minimum-Maximum Mean Temperature Range (°C)</th>
<th>Mean Annual Temperature (°C)</th>
<th>Mean Annual Precipitation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eagle Creek (EC)</td>
<td>-28.30° - 18.40°</td>
<td>-4.94°</td>
<td>247</td>
</tr>
<tr>
<td>Nome Creek (NC)</td>
<td>-29.90° - 20.10°</td>
<td>-4.87°</td>
<td>319</td>
</tr>
<tr>
<td>Victoria Mountain (VM)</td>
<td>-25.20° - 17.30°</td>
<td>-4.55°</td>
<td>290</td>
</tr>
<tr>
<td>Colorado Creek (CC)</td>
<td>-30.30° - 20.90°</td>
<td>-4.09°</td>
<td>288</td>
</tr>
<tr>
<td>Elliott Highway (EL)</td>
<td>-25.30° - 17.80°</td>
<td>-3.98°</td>
<td>311</td>
</tr>
<tr>
<td>No Name Creek (NN)</td>
<td>-31.20° - 21.90°</td>
<td>-5.07°</td>
<td>334</td>
</tr>
<tr>
<td>Gobbler’s Knob (GO)</td>
<td>-27.70° - 18.70°</td>
<td>-4.98°</td>
<td>348</td>
</tr>
<tr>
<td>Coldfoot (CF)</td>
<td>-31.20° - 19.80°</td>
<td>-5.58°</td>
<td>297</td>
</tr>
<tr>
<td>South of Timberline (ST)</td>
<td>-27.70° - 15.10°</td>
<td>-6.97°</td>
<td>279</td>
</tr>
<tr>
<td>Timberline (TB)</td>
<td>-28.00° - 15.60°</td>
<td>-6.92°</td>
<td>282</td>
</tr>
<tr>
<td>Chandalar (CH)</td>
<td>-26.90° - 14.10°</td>
<td>-7.55°</td>
<td>288</td>
</tr>
<tr>
<td>Atigun Camp (AT)</td>
<td>-25.40° - 11.50°</td>
<td>-8.75°</td>
<td>279</td>
</tr>
<tr>
<td>Toolik Lake (TL)</td>
<td>-29.30° - 16.00°</td>
<td>-8.74°</td>
<td>239</td>
</tr>
<tr>
<td>Anaktuvuk (AN)</td>
<td>-32.90° - 18.10°</td>
<td>-9.72°</td>
<td>210</td>
</tr>
<tr>
<td>Sagwon (SG)</td>
<td>-34.00° - 17.70°</td>
<td>-10.08°</td>
<td>183</td>
</tr>
<tr>
<td>Coastal Plain (CP)</td>
<td>-33.10° - 15.80°</td>
<td>-10.37°</td>
<td>171</td>
</tr>
<tr>
<td>Prudhoe Bay (PB)</td>
<td>-31.60° - 11.80°</td>
<td>-11.35°</td>
<td>161</td>
</tr>
</tbody>
</table>
Table S2. Variable subsets of 19 temperature and precipitation variables (from the WorldClim 2.0 Bioclimatic database (Fick and Hijmans, 2017) used in Principal Components Analysis (PCA). Mean diurnal range = mean of monthly maximum temperature-minimum temperature, Isothermality = annual mean diurnal range/annual temperature range, Temperature seasonality (standard deviation * 100), Maximum temperature of the warmest month (Jul), Minimum temperature of the coldest month (Jan), Temperature annual range (maximum temperature of the warmest month (Jul)-minimum temperature of the coldest month (Jan), Mean temperature of the wettest quarter (Jul, Aug, Sept), Mean temperature of the driest quarter (Feb, Mar, Apr), Mean temperature of the warmest quarter (Jun, Jul, Aug), Mean temperature of the coldest quarter (Dec, Jan, Feb), Precipitation of the wettest month (Aug), Precipitation of the driest month (Apr), Precipitation seasonality = coefficient of variation estimated from the standard deviation of monthly precipitation estimates; Precipitation of the wettest quarter (Jul, Aug, Sept), Precipitation of the driest quarter (Feb, Mar, Apr), Precipitation of the warmest quarter (Jun, Jul, Aug), Precipitation of the coldest quarter (Dec, Jan, Feb).

<table>
<thead>
<tr>
<th>Temperature Variables</th>
<th>Precipitation Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual mean temperature (°C)</td>
<td>Annual precipitation (mm)</td>
</tr>
<tr>
<td>Mean diurnal range (°C)</td>
<td>Precipitation of the wettest month (mm)</td>
</tr>
<tr>
<td>Isothermality (%)</td>
<td>Precipitation of the driest month (mm)</td>
</tr>
<tr>
<td>Temperature seasonality (°C)</td>
<td>Precipitation seasonality (coefficient of variation)</td>
</tr>
<tr>
<td>Maximum temperature of the warmest month (°C)</td>
<td>Precipitation of the wettest quarter (mm)</td>
</tr>
<tr>
<td>Minimum temperature of the coldest month (°C)</td>
<td>Precipitation of the driest quarter (mm)</td>
</tr>
<tr>
<td>Temperature annual range (°C)</td>
<td>Precipitation of the warmest quarter (mm)</td>
</tr>
<tr>
<td>Mean temperature of the wettest quarter (°C)</td>
<td>Precipitation of the coldest quarter (mm)</td>
</tr>
<tr>
<td>Mean temperature of the driest quarter (°C)</td>
<td></td>
</tr>
<tr>
<td>Mean temperature of the warmest quarter (°C)</td>
<td></td>
</tr>
<tr>
<td>Mean temperature of the coldest quarter (°C)</td>
<td></td>
</tr>
</tbody>
</table>
### Table S3. Table of Pairwise $F_{ST}$ values for the neutral dataset. $F_{ST}$ values are on the lower diagonal and p-values are on the upper diagonal. * refers to significance at $\leq 0.01$.

<table>
<thead>
<tr>
<th></th>
<th>EC</th>
<th>NC</th>
<th>VM</th>
<th>CC</th>
<th>EL</th>
<th>NN</th>
<th>GO</th>
<th>CF</th>
<th>ST</th>
<th>TB</th>
<th>CH</th>
<th>AT</th>
<th>TL</th>
<th>AN</th>
<th>SG</th>
<th>CP</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>NC</td>
<td>0.034</td>
<td>0.001*</td>
<td>0.002*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>VM</td>
<td>0.039</td>
<td>0.012</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>CC</td>
<td>0.039</td>
<td>0.008</td>
<td>0.015</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>EL</td>
<td>0.039</td>
<td>0.008</td>
<td>0.013</td>
<td>0.008</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>NN</td>
<td>0.038</td>
<td>0.009</td>
<td>0.012</td>
<td>0.008</td>
<td>0.010</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>GO</td>
<td>0.041</td>
<td>0.012</td>
<td>0.015</td>
<td>0.011</td>
<td>0.011</td>
<td>0.005</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>ST</td>
<td>0.043</td>
<td>0.011</td>
<td>0.017</td>
<td>0.014</td>
<td>0.012</td>
<td>0.006</td>
<td>0.004</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>TB</td>
<td>0.050</td>
<td>0.021</td>
<td>0.025</td>
<td>0.023</td>
<td>0.021</td>
<td>0.016</td>
<td>0.015</td>
<td>0.015</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>CH</td>
<td>0.054</td>
<td>0.022</td>
<td>0.026</td>
<td>0.022</td>
<td>0.023</td>
<td>0.018</td>
<td>0.014</td>
<td>0.016</td>
<td>0.017</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>AT</td>
<td>0.054</td>
<td>0.025</td>
<td>0.029</td>
<td>0.025</td>
<td>0.025</td>
<td>0.021</td>
<td>0.022</td>
<td>0.019</td>
<td>0.018</td>
<td>0.016</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>TL</td>
<td>0.060</td>
<td>0.027</td>
<td>0.033</td>
<td>0.029</td>
<td>0.027</td>
<td>0.027</td>
<td>0.024</td>
<td>0.026</td>
<td>0.024</td>
<td>0.018</td>
<td>0.019</td>
<td>0.008</td>
<td>0.006</td>
<td>0.004*</td>
<td>0.0078</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>AN</td>
<td>0.054</td>
<td>0.026</td>
<td>0.031</td>
<td>0.028</td>
<td>0.024</td>
<td>0.023</td>
<td>0.023</td>
<td>0.019</td>
<td>0.019</td>
<td>0.015</td>
<td>0.008</td>
<td>0.003*</td>
<td>0.002</td>
<td>0.002</td>
<td>0.004</td>
<td>0.003*</td>
<td>0.001*</td>
</tr>
<tr>
<td>SG</td>
<td>0.052</td>
<td>0.025</td>
<td>0.029</td>
<td>0.023</td>
<td>0.027</td>
<td>0.021</td>
<td>0.024</td>
<td>0.022</td>
<td>0.017</td>
<td>0.016</td>
<td>0.008</td>
<td>0.003</td>
<td>0.002</td>
<td>0.004</td>
<td>0.003</td>
<td>0.001*</td>
<td>0.001*</td>
</tr>
<tr>
<td>CP</td>
<td>0.050</td>
<td>0.020</td>
<td>0.025</td>
<td>0.021</td>
<td>0.020</td>
<td>0.017</td>
<td>0.018</td>
<td>0.017</td>
<td>0.016</td>
<td>0.015</td>
<td>0.010</td>
<td>0.008</td>
<td>0.006</td>
<td>0.002</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>PB</td>
<td>0.045</td>
<td>0.019</td>
<td>0.022</td>
<td>0.020</td>
<td>0.020</td>
<td>0.015</td>
<td>0.015</td>
<td>0.016</td>
<td>0.015</td>
<td>0.015</td>
<td>0.010</td>
<td>0.006</td>
<td>0.009</td>
<td>0.006</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
</tr>
</tbody>
</table>
Table S4. Analysis of Molecular Variance (AMOVA) results for 17 *Eriophorum vaginatum* sites in north central Alaska for the neutral SNP data set.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>DF</th>
<th>% Variation</th>
<th>F-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Clusters</td>
<td>2</td>
<td>3.112</td>
<td>0.031</td>
<td>0.01</td>
</tr>
<tr>
<td>Between Sites Within Clusters</td>
<td>14</td>
<td>2.024</td>
<td>0.021</td>
<td>0.01</td>
</tr>
<tr>
<td>Within Sites</td>
<td>256</td>
<td>94.864</td>
<td>0.051</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Table S5. \textit{divMigrate} $G_{ST}$ based migration values. Values by row indicate migration for each site going into sites listed by column. Migration values $\geq 0.8$ are in bold.

<table>
<thead>
<tr>
<th></th>
<th>EC</th>
<th>NC</th>
<th>VM</th>
<th>CC</th>
<th>EL</th>
<th>NN</th>
<th>GO</th>
<th>CF</th>
<th>ST</th>
<th>TB</th>
<th>CH</th>
<th>AT</th>
<th>TL</th>
<th>AN</th>
<th>SG</th>
<th>CP</th>
<th>PB</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC</td>
<td>0.457</td>
<td>0.408</td>
<td>0.418</td>
<td>0.425</td>
<td>0.416</td>
<td>0.409</td>
<td>0.409</td>
<td>0.344</td>
<td>0.328</td>
<td>0.313</td>
<td>0.321</td>
<td>0.300</td>
<td>0.325</td>
<td>0.345</td>
<td>0.338</td>
<td>0.371</td>
<td></td>
</tr>
<tr>
<td>NC</td>
<td>0.484</td>
<td>0.646</td>
<td>0.792</td>
<td>0.771</td>
<td>0.762</td>
<td>0.712</td>
<td>0.745</td>
<td>0.556</td>
<td>0.548</td>
<td>0.503</td>
<td>0.517</td>
<td>0.491</td>
<td>0.537</td>
<td>0.538</td>
<td>0.586</td>
<td>0.577</td>
<td></td>
</tr>
<tr>
<td>VM</td>
<td>0.430</td>
<td>0.685</td>
<td>0.660</td>
<td>0.699</td>
<td>0.718</td>
<td>0.652</td>
<td>0.670</td>
<td>0.527</td>
<td>0.511</td>
<td>0.497</td>
<td>0.477</td>
<td>0.457</td>
<td>0.474</td>
<td>0.492</td>
<td>0.531</td>
<td>0.556</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>0.450</td>
<td>0.779</td>
<td>0.623</td>
<td>0.800</td>
<td>0.803</td>
<td>0.737</td>
<td>0.750</td>
<td>0.551</td>
<td>0.566</td>
<td>0.544</td>
<td>0.507</td>
<td>0.491</td>
<td>0.513</td>
<td>0.572</td>
<td>0.583</td>
<td>0.591</td>
<td></td>
</tr>
<tr>
<td>EL</td>
<td>0.454</td>
<td>0.766</td>
<td>0.656</td>
<td>0.807</td>
<td>0.765</td>
<td>0.733</td>
<td>0.764</td>
<td>0.563</td>
<td>0.538</td>
<td>0.524</td>
<td>0.540</td>
<td>0.499</td>
<td>0.528</td>
<td>0.516</td>
<td>0.590</td>
<td>0.588</td>
<td></td>
</tr>
<tr>
<td>NN</td>
<td>0.458</td>
<td>0.720</td>
<td>0.639</td>
<td>0.790</td>
<td>0.760</td>
<td>0.858</td>
<td>0.885</td>
<td>0.641</td>
<td>0.611</td>
<td>0.585</td>
<td>0.545</td>
<td>0.524</td>
<td>0.545</td>
<td>0.545</td>
<td>0.643</td>
<td>0.619</td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>0.431</td>
<td>0.653</td>
<td>0.623</td>
<td>0.733</td>
<td>0.715</td>
<td>0.816</td>
<td>0.932</td>
<td>0.661</td>
<td>0.645</td>
<td>0.565</td>
<td>0.569</td>
<td>0.513</td>
<td>0.555</td>
<td>0.561</td>
<td>0.626</td>
<td>0.648</td>
<td></td>
</tr>
<tr>
<td>CF</td>
<td>0.440</td>
<td>0.742</td>
<td>0.617</td>
<td>0.741</td>
<td>0.783</td>
<td>0.878</td>
<td>0.950</td>
<td>0.729</td>
<td>0.674</td>
<td>0.624</td>
<td>0.627</td>
<td>0.546</td>
<td>0.608</td>
<td>0.615</td>
<td>0.674</td>
<td>0.660</td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td>0.376</td>
<td>0.548</td>
<td>0.471</td>
<td>0.565</td>
<td>0.570</td>
<td>0.637</td>
<td>0.659</td>
<td>0.681</td>
<td>0.622</td>
<td>0.603</td>
<td>0.594</td>
<td>0.573</td>
<td>0.612</td>
<td>0.647</td>
<td>0.647</td>
<td>0.628</td>
<td></td>
</tr>
<tr>
<td>TB</td>
<td>0.356</td>
<td>0.542</td>
<td>0.505</td>
<td>0.577</td>
<td>0.581</td>
<td>0.631</td>
<td>0.700</td>
<td>0.719</td>
<td>0.654</td>
<td>0.674</td>
<td>0.594</td>
<td>0.656</td>
<td>0.689</td>
<td>0.694</td>
<td>0.662</td>
<td>0.662</td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>0.360</td>
<td>0.516</td>
<td>0.448</td>
<td>0.538</td>
<td>0.552</td>
<td>0.572</td>
<td>0.577</td>
<td>0.657</td>
<td>0.622</td>
<td>0.641</td>
<td>0.795</td>
<td>0.768</td>
<td>0.771</td>
<td>0.822</td>
<td>0.772</td>
<td>0.741</td>
<td></td>
</tr>
<tr>
<td>AT</td>
<td>0.359</td>
<td>0.514</td>
<td>0.455</td>
<td>0.520</td>
<td>0.527</td>
<td>0.560</td>
<td>0.552</td>
<td>0.613</td>
<td>0.612</td>
<td>0.660</td>
<td>0.798</td>
<td>0.827</td>
<td>0.834</td>
<td>0.972</td>
<td>0.810</td>
<td>0.837</td>
<td></td>
</tr>
<tr>
<td>TL</td>
<td>0.333</td>
<td>0.503</td>
<td>0.432</td>
<td>0.514</td>
<td>0.523</td>
<td>0.551</td>
<td>0.541</td>
<td>0.608</td>
<td>0.634</td>
<td>0.625</td>
<td>0.860</td>
<td>0.856</td>
<td>0.861</td>
<td>1.000</td>
<td>0.854</td>
<td>0.814</td>
<td></td>
</tr>
<tr>
<td>AN</td>
<td>0.360</td>
<td>0.558</td>
<td>0.485</td>
<td>0.550</td>
<td>0.588</td>
<td>0.590</td>
<td>0.598</td>
<td>0.652</td>
<td>0.680</td>
<td>0.693</td>
<td>0.875</td>
<td>0.894</td>
<td>0.859</td>
<td>0.891</td>
<td>0.825</td>
<td>0.834</td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>0.373</td>
<td>0.544</td>
<td>0.475</td>
<td>0.581</td>
<td>0.547</td>
<td>0.598</td>
<td>0.589</td>
<td>0.631</td>
<td>0.660</td>
<td>0.652</td>
<td>0.855</td>
<td>0.932</td>
<td>0.959</td>
<td>0.884</td>
<td>0.987</td>
<td>0.896</td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>0.384</td>
<td>0.560</td>
<td>0.498</td>
<td>0.584</td>
<td>0.608</td>
<td>0.632</td>
<td>0.619</td>
<td>0.666</td>
<td>0.635</td>
<td>0.628</td>
<td>0.745</td>
<td>0.781</td>
<td>0.764</td>
<td>0.795</td>
<td>0.953</td>
<td>0.916</td>
<td></td>
</tr>
<tr>
<td>PB</td>
<td>0.402</td>
<td>0.570</td>
<td>0.503</td>
<td>0.577</td>
<td>0.575</td>
<td>0.669</td>
<td>0.648</td>
<td>0.694</td>
<td>0.661</td>
<td>0.651</td>
<td>0.746</td>
<td>0.815</td>
<td>0.733</td>
<td>0.768</td>
<td>0.895</td>
<td>0.896</td>
<td></td>
</tr>
</tbody>
</table>
Table S6. Values for retained RDA variables ((all except MEM1 and MEM2 obtained from the WorldClim 2.0 Bioclimatic database (Fick and Hijmans, 2017). See Table 1 for 17 site abbreviations, iso = Isothermality, tdq = Mean temperature of driest quarter (Feb, Mar, Apr), twq = Mean temperature of warmest quarter (Jun, Jul, Aug), prd = Precipitation of driest month (Apr), prs = Precipitation seasonality.

<table>
<thead>
<tr>
<th>Site</th>
<th>iso</th>
<th>tdq (°C)</th>
<th>twq (°C)</th>
<th>prd (mm)</th>
<th>prs</th>
<th>MEM1</th>
<th>MEM2</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC</td>
<td>24.89</td>
<td>-13.00</td>
<td>11.37</td>
<td>6.00</td>
<td>73.46</td>
<td>-1.32</td>
<td>-0.72</td>
</tr>
<tr>
<td>NC</td>
<td>22.92</td>
<td>-13.70</td>
<td>12.62</td>
<td>10.00</td>
<td>57.89</td>
<td>-1.46</td>
<td>-0.71</td>
</tr>
<tr>
<td>VM</td>
<td>22.47</td>
<td>-12.98</td>
<td>11.57</td>
<td>7.00</td>
<td>59.66</td>
<td>-1.46</td>
<td>-0.70</td>
</tr>
<tr>
<td>CC</td>
<td>21.14</td>
<td>-12.75</td>
<td>14.08</td>
<td>8.00</td>
<td>58.04</td>
<td>-1.35</td>
<td>-0.30</td>
</tr>
<tr>
<td>EL</td>
<td>21.27</td>
<td>-11.92</td>
<td>12.18</td>
<td>8.00</td>
<td>68.07</td>
<td>-1.27</td>
<td>-0.10</td>
</tr>
<tr>
<td>NN</td>
<td>21.80</td>
<td>-15.13</td>
<td>14.55</td>
<td>10.00</td>
<td>61.60</td>
<td>-0.82</td>
<td>0.89</td>
</tr>
<tr>
<td>GO</td>
<td>19.38</td>
<td>-13.57</td>
<td>12.98</td>
<td>11.00</td>
<td>59.46</td>
<td>-0.10</td>
<td>1.59</td>
</tr>
<tr>
<td>CF</td>
<td>21.73</td>
<td>-14.32</td>
<td>12.77</td>
<td>9.00</td>
<td>61.63</td>
<td>0.08</td>
<td>1.55</td>
</tr>
<tr>
<td>ST</td>
<td>22.33</td>
<td>-15.23</td>
<td>9.38</td>
<td>7.00</td>
<td>66.91</td>
<td>0.54</td>
<td>1.27</td>
</tr>
<tr>
<td>TB</td>
<td>22.29</td>
<td>-15.25</td>
<td>9.55</td>
<td>10.00</td>
<td>62.59</td>
<td>0.54</td>
<td>1.27</td>
</tr>
<tr>
<td>CH</td>
<td>21.83</td>
<td>-16.03</td>
<td>8.60</td>
<td>9.00</td>
<td>62.11</td>
<td>0.76</td>
<td>0.88</td>
</tr>
<tr>
<td>AT</td>
<td>20.21</td>
<td>-16.98</td>
<td>6.18</td>
<td>10.00</td>
<td>62.47</td>
<td>0.92</td>
<td>0.05</td>
</tr>
<tr>
<td>TL</td>
<td>21.73</td>
<td>-19.75</td>
<td>9.93</td>
<td>7.00</td>
<td>64.45</td>
<td>1.06</td>
<td>-0.50</td>
</tr>
<tr>
<td>AN</td>
<td>21.45</td>
<td>-22.30</td>
<td>11.00</td>
<td>7.00</td>
<td>62.08</td>
<td>0.98</td>
<td>-0.59</td>
</tr>
<tr>
<td>SG</td>
<td>21.44</td>
<td>-26.58</td>
<td>10.67</td>
<td>6.00</td>
<td>60.53</td>
<td>1.00</td>
<td>-1.15</td>
</tr>
<tr>
<td>CP</td>
<td>19.61</td>
<td>-26.63</td>
<td>9.23</td>
<td>5.00</td>
<td>65.70</td>
<td>0.96</td>
<td>-1.36</td>
</tr>
<tr>
<td>PB</td>
<td>15.96</td>
<td>-27.42</td>
<td>6.60</td>
<td>4.00</td>
<td>65.10</td>
<td>0.94</td>
<td>-1.38</td>
</tr>
</tbody>
</table>
Table S7. Annotated *E. vaginatum* candidate genes with a percentage identity match of at least 80.0 and an E-value threshold of at least $1 \times 10^{-4}$ and RDA $R^2$ value between 0.7 and 0.79.

<table>
<thead>
<tr>
<th>Locus ID</th>
<th>Gene ID</th>
<th>RDA $R^2$</th>
<th>RDA Top Predictor</th>
</tr>
</thead>
<tbody>
<tr>
<td>42747</td>
<td>Ribonuclease H domain</td>
<td>0.795</td>
<td>MEM1</td>
</tr>
<tr>
<td>82292</td>
<td>Acyltransferase-like protein</td>
<td>0.794</td>
<td>MEM1</td>
</tr>
<tr>
<td>1523776</td>
<td>DEAD-box ATP-dependent RNA helicase</td>
<td>0.793</td>
<td>MEM1</td>
</tr>
<tr>
<td>66558</td>
<td>Cysteine-rich receptor-like protein kinase 10-like</td>
<td>0.791</td>
<td>MEM1</td>
</tr>
<tr>
<td>43633</td>
<td>Protein ENL-like</td>
<td>0.780</td>
<td>MEM1</td>
</tr>
<tr>
<td>74218</td>
<td>AAA ATPase At3g50940-like</td>
<td>0.778</td>
<td>MEM1</td>
</tr>
<tr>
<td>73664</td>
<td>Cellulose synthase-like protein G2-like</td>
<td>0.778</td>
<td>MEM1</td>
</tr>
<tr>
<td>70372</td>
<td>Acyl-binding domain-containing-4</td>
<td>0.771</td>
<td>tdq</td>
</tr>
<tr>
<td>1963599</td>
<td>Magnesium transporter NIPA2-like</td>
<td>0.769</td>
<td>MEM1</td>
</tr>
<tr>
<td>71022</td>
<td>Dynamin-related protein 5A</td>
<td>0.767</td>
<td>MEM1</td>
</tr>
<tr>
<td>47232</td>
<td>Zinc finger-like protein</td>
<td>0.764</td>
<td>MEM1</td>
</tr>
<tr>
<td>1721445</td>
<td>Zinc-dependent exopeptidases superfamily protein</td>
<td>0.756</td>
<td>twq</td>
</tr>
<tr>
<td>32318</td>
<td>UPF0496 protein 4</td>
<td>0.755</td>
<td>MEM1</td>
</tr>
<tr>
<td>45472</td>
<td>Pentatricopeptide repeat-containing mitochondrial-like</td>
<td>0.747</td>
<td>tdq</td>
</tr>
<tr>
<td>1405</td>
<td>Actin cytoskeleton-regulatory complex PAN1 isoform X1</td>
<td>0.744</td>
<td>MEM1</td>
</tr>
<tr>
<td>68114</td>
<td>Flowering time control protein FCA isoform X2</td>
<td>0.725</td>
<td>MEM1</td>
</tr>
<tr>
<td>32648</td>
<td>Heavy metal ATPase</td>
<td>0.714</td>
<td>MEM1</td>
</tr>
<tr>
<td>36463</td>
<td>ATP-dependent zinc metalloprotease FTSH chloroplastic</td>
<td>0.713</td>
<td>MEM1</td>
</tr>
</tbody>
</table>
SUPPLEMENTAL FIGURE

Figure S1. (A) $\Delta K$ values used to determine number of clusters ($K$) of STRUCTURE results (B) Bayesian Information Criterion (BIC) used to infer $K$ for Discriminant Analysis of Principal Com
Vita

Elizabeth Anne Stunz was born in Boise, Idaho and received a Bachelor of Arts degree in Creative Writing and French from the University of Idaho in 2009. In 2013 she returned to the University of Idaho to take prerequisites for a degree program in plant science, and in 2015 she was admitted to the Ecology and Evolutionary Biology Doctoral Program at the University of Texas at El Paso (UTEP). From 2015 to 2018 she worked as a research assistant in the Plant Evolution Lab, where she worked with Dr. Michael Moody to spearhead efforts to establish methodology for Restriction-site Associated DNA Sequencing (RAD-seq), and the optimization of plant DNA extraction protocols. During this time, she also conducted months of cumulative field research, both in the Arctic and in the Chihuahuan Desert, to collect plant tissue of various species. Additionally, Elizabeth trained and mentored several undergraduate and graduate students in multiple facets of lab and field work and worked as a Teaching Assistant for the UTEP Biology Department, including four semesters teaching the Botany lab and two semesters teaching the Plant Diversity & Systematics lab.

While at UTEP, she volunteered for multiple outreach events with 500 Women Scientists, UTEP Biodiversity Collections and as a Botanical Society of America (BSA) Planting Science mentor. She wrote multiple proposals and received numerous grant awards, including UTEP’s Dodson Grant (2017, 2019) and the BSA’s Graduate Student Research Grant (2017). Elizabeth presented her research at Evolution (2016, 2017, 2019) and Botany (2018, 2022) conferences and has published two peer-reviewed manuscripts, with two manuscripts in preparation. She plans to pursue a career as a plant conservation geneticist and can be contacted at lizstunz@gmail.com.

This dissertation was typed by Elizabeth Anne Stunz.