Winter 2021

# Genetic structure of Pseudoroegneria spicata in the Northern Palouse Prairie and Channeled Scablands

Ethan Timothy Bean
*Eastern Washington University*

# GENETIC STRUCTURE OF *PSEUDOROEGNERIA SPICATA* IN THE NORTHERN PALOUSE PRAIRIE AND CHANNELED SCABLANDS

---

A Thesis

Presented to

Eastern Washington University

Cheney, Washington

---

In Partial Fulfillment of the Requirements

For the Degree

Master of Science in Biology

---

By

Ethan Timothy Bean

Winter 2021

THESIS OF ETHAN TIMOTHY BEAN APPROVED BY

_____ DATE _____
Rebecca Brown

_____ DATE _____
Paul Spruell

_____ DATE _____
Jessica Allen

_____ DATE _____
Brian Buchanan

_____ DATE _____
Steve Larson

# ABSTRACT

GENETIC STRUCTURE OF *PSEUDOROEGNERIA SPICATA* IN THE NORTHERN PALOUSE PRAIRIE AND CHANNELED SCABLANDS

by

Ethan Timothy Bean

Winter 2021

Establishing genetically diverse communities that can adapt to dynamic selective pressures is crucial in ecological restoration. However, the genetic structure of native plant species used for restoration is often poorly understood. *Pseudoroegneria spicata,* is a keystone species of the Inland Northwestern US that has become a staple in commercially available restoration seed mixes. It is abundant in remnant prairies of the endangered Palouse Prairie ecoregion, which is characterized by rolling hills of deep loess soil. Less than 1% of native Palouse Prairie is left, due to agricultural conversion, with remnants highly fragmented and isolated. *P. spicata* is also common in nearby Channeled Scabland habitats formed by flooding during the last glacial advance and characterized by relatively shallow soils. Previous studies of the genetic structure of *P. spicata* have not included populations from the Northern Palouse and Channeled Scablands. My goal was to assess the genetic structure of *P. spicata* across Palouse Prairie remnants and Channeled Scabland habitats in Eastern Washington, and determine relatedness to commercial seed. I hypothesized that habitat differences between the Palouse and Channeled Scablands would lead to genetic differentiation between *P. spicata* populations. I also expected that local source-identified commercial plant

material would be more closely related to plants from our sampling locations than plant material collected from a larger nearby area and pooled. Plant DNA samples were collected from six locations in the Channeled Scablands and Palouse Prairie and two commercial sources. I calculated inbreeding coefficient, conducted principal component analyses, and used Bayesian cluster analysis to test for inbreeding, differentiation of populations, and relatedness of native propagule sources. *Pseudoroegneria spicata* differentiated along a north-south latitudinal gradient instead of between different habitat types. Commercially sourced seeds from a certified source-identified provenance were more closely related to northern sites, while seeds sourced from the larger region were more similar to southern sites. There were no signs of inbreeding in commercial seed sources, however I found possible evidence of hybridization at a native seed nursery. Samples from one of the larger Palouse remnants had a lower inbreeding coefficient than the rest of the sites, highlighting the importance of its continued preservation.

# ACKNOWLEDGEMENTS

I would like to thank Rebecca Brown for her input on project development and her in-depth reviews of this thesis. This project would not have been possible without Steve Larson's assistance with experimental design, sequencing, data preparation, and data analysis. I would like to extend thanks to Jessica Allen for her assistance with experimental design, and hours of help with analyzing and interpreting data. Thank you to Paul Spruell, Krisztian Magori, and Rob Massatti for their help with data analysis. Field collecting would not have been possible without the input and advice of Robin O'Quinn. Brian Buchannan was very helpful with his assistance with work in ArcGIS. Lin Johnson's was for my DNA extraction and library preparation. Permission to access land from the Washington Department of Natural Resources, Steptoe Butte State Park, the Bureau of Land Management, and the US Fish and Wildlife Service made this project possible. I would like to thank the Restoration Ecology Lab at EWU for their time and support with listening to practice talks, reading manuscripts, and giving much-needed input on my project along the way. Thank you to the staff and faculty at EWU for being understanding and supportive of my National Guard career which has frequently pulled me away from school and prolonged this process. Finally, I would like to thank Katy Bean for her field assistance, draft editing, and general support during my academic endeavors.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

## Introduction

Establishing genetically diverse communities that can adapt to dynamic selective pressures is a challenge that must be addressed in ecological restoration (Kilkenny et al. 2013). Many restoration projects use native seed mixes with local propagule sources for each species in the mix (Aavik et al. 2011). This creates a restoration site with similar appearance and ecosystem functions to remnant sites that have remained mostly untouched since human colonization (Larson et al. 2011). The genetic structure of the native plants in the mix, and the local population area must be defined to best source each propagule in the seed mix (Aavik et al. 2011). Once the local population area is defined, propagule sources for each species in the mix can be identified.

Restoration sites with locally adapted genetically diverse plants are best suited to provide ecosystem functions, support native organisms, and persist within local climate and soil types (Johnson and Agarwal 2005, Aavik et al. 2011). Using genetically diverse seeds in restoration can produce larger, more robust plants in the site (Bischoff et al. 2010). Restored populations deficient in genetic diversity may to struggle to adapt to selective pressures caused by changing climates (Basey et al. 2015).

Local adaptation of plant material is equally as important as genetic diversity, this is because phenological traits such as germination and flower timing evolve to be best suited to local climates (Clausen et al. 1941, Bischoff et al. 2010). Identification of rare genotypes contained within fragmented or ecologically unique habitats should be prioritized because they may become important to the species' ability to adapt to changing conditions (Dolan et al. 2008, Rodriguez-Quilon et al. 2016). Parameters such as soil type, climate, climate change predictions, and genetic analyses have all been used

to identify appropriate propagule sources for use in restoration (Kilkenny et al. 2013, St Clair et al. 2013 Gibson et al. 2019). Genetic analyses have also been used; This allows identification of local genotypes, reducing the risk of introducing maladapted genotypes which can cause cause outbreeding depression (Crémieux et al. 2010).

Sufficient genetic diversity within seed mixes is important because it can provide habitat structure for native organisms. Microhabitats created by different genotypes within a single population of plants can support a range of native organisms (Johnson and Agarwal 2005). In addition to supporting local native species, plants propagated from a provenance (the location from which propagules are sourced or originated) other than what they are currently adapted to may not survive with disease and pest pressures in the new location (Schoen and Brown 2001). Therefore, identification and use of local propagule sources ensures the best performance of the restored community (Bischoff et al. 2010, Aavik et al. 2011).

Native plant material used in restoration projects is often purchased from commercial production facilities, and there are several practices these facilities should follow to ensure production of local and genetically diverse provenances. Some efficiency practices can promote maladapted genotypes or low genetic diversity in commercial seed, creating outbreeding depression and founder effects in the restoration site (Fant et al. 2008, Aavik et al. 2011). For example, accidental selection may occur in for plants that are larger, flower sooner, and produce more seeds than their wild relatives; however, this can come at a trade off with the plant's viability in the wild (Schröder and Prasse 2013b). Implementing measures to prevent loss of seeds in processing and minimize accidental hybridization between ecotypes can help ensure production of local

genetically diverse propagules (Basey et al. 2015). Additionally, native plant nurseries should regularly refresh their seed provenances with seed sourced from wild sites to avoid alteration of life history traits within propagules over time (Schröder and Prasse 2013a).

One widespread native grass that is used in restoration projects across western North America is *Pseudoroegneria spicata* (Pursh) Á. Löve (bluebunch wheatgrass). *P. spicata* is a native cool-season bunch grass that is valuable in habitat restoration because it is drought tolerant and resistant to invasion promoting early survival and establishment (Ogle et al. 2010, Davies and Johnson 2017). Because of its wide use and range, there are many native plant nurseries that sell *P. spicata* from a variety of provenances (personal observation). *P. spicata* rarely reproduces rhizomatously and is highly self-sterile making pollen and seeds the prominent method of gene dispersal (Keller 1948, Jensen et al. 1990, Hitchcock et al. 2018).

Studies of the genetic structure of *P. spicata* populations can be used to choose propagules sources for restoration, however this level of understanding is not available for all regions. Seed transfer zones have been defined for *P. spicata* in North America using common gardens and climate data, amplified fragment length polymorphisms (AFLPs), and Microsatellites (Larson et al. 2004, Fu and Thompson 2006, St. Clair et al. 2013). More recently genotyping by sequencing (GBS) identified five *P. spicata* genetic clusters across Utah, Nevada, Idaho, Oregon, and Washington (Massatti et al. 2018). One region where understanding of population genetic structure for *P. spicata* is lacking is the northern edge of Palouse Prairie Ecoregion and the Channeled Scablands in the Inland Northwestern United States.

The Palouse Prairie was formerly an abundant ecosystem that extended 16,000 $km^2$, but it is now limited to less than one percent of its original range (Noss et al. 1995, Sanchez-de Leon and Johnson-Maynard 2013). Fragmentation caused by land use conversion in this habitat may have genetic implications for local plant communities. The Palouse Prairie is characterized by deep, wind-blown, loess soils that accumulated during the Pleistocene atop thick basalt bedrock (Bryan 1926). After the last glacial advance, catastrophic floods with high volumes of water carved large areas of loess soil down to bedrock, giving them a thin profile and creating the Channeled Scablands (Bretz 1925). These Channeled Scablands weave through what was the northern extent of the Palouse Prairie leaving islands of loess soil surrounded by thin rocky soil over basalt (Bretz 1925). While islands of loess soil in and around the Channeled Scablands are not part of the EPA level IV Palouse Hills ecoregion, they are included in broader definitions of Palouse Prairie (Bowlick et al. 2015), have similar loess soil and plant communities, and were considered Palouse Prairie for this study (Figure 1, U.S. EPA 2013). In areas where deep loess deposits remain, soils retain moisture throughout the dry season which has led to their use for cultivation (Bryan 1926).

Conversion of the Palouse Prairie to agricultural crop land has led to ecological problems leading to a need for restoration. Starting in the 1870s, European-Americans began to colonize the area and ultimately converted 99% of the prairie to agricultural crops (Noss et al. 1995, Duffin, 2007). Habitat fragmentation caused by conversion of large areas to crop fields has relegated the remaining remnants of Palouse Prairie to isolated small islands of habitat that were inconvenient for plowing. Depending on the size and distance between fragmented local prairie habitats, plant populations living

within them may be reproductively isolated and contain rare genes. Moreover, extremely small and isolated habitats may suffer from inbreeding depression and genetic drift (Ellstrand and Elam 1993).

To aid recovery of the Palouse Ecosystem, Eastern Washington University (EWU) will be restoring 120 acres of wheat field to prairie habitat on its campus in Cheney, WA. Cheney is near the northern edge of the Palouse Ecoregion, where prairie and Channeled Scablands are found in close proximity. *Pseudoroegneria spicata* will be a foundational species used in the seed mix at EWU for this project and there is a need for a better understanding of local genetic structure to select an appropriate propagule source.

The objectives of this study were 1) to assess local populations of *P. spicata* for inbreeding, 2) to characterize the genetic structure of *P. spicata* in the Northern Palouse Prairie and Channeled Scablands of Eastern Washington, and 3) to determine a genetically appropriate commercial propagule source for EWU's restoration project. I hypothesized that due to fragmentation caused by conversion to agricultural land, *P. spicata* populations in the Northern Palouse and Channeled Scablands would have signs of inbreeding. I hypothesized that Palouse sites and Scabland sites would be genetically distinct due to adaptation to different habitats within the sites. Finally, I expected commercially available *P. spicata* seed that was certified by the Washington State Department of Agriculture (WSDA) as source-identified certified to be more genetically similar to local populations than Forest Service seed material collected from a large area and pooled.

## Methods

### *Study Area*

The study area for this project was the Northern Palouse and Channeled Scablands in the state of Washington. Sampling locations fell within Spokane, Grant, Jefferson, and Lincoln Counties. The overall average annual temperature is 8.44°C and the average annual precipitation is 414.94 mm (Table 1, Prism 2021). EWU's Restoration site is located within the town of Cheney, WA. This is located on an island of loess soil that was formerly part of the Palouse Prairie ecotype, surrounded by the Channeled Scablands in northeastern Washington.

### *Study Design*

Twelve plant tissue samples were collected from two sites in the Northern Palouse Prairie (Unnamed DNR land near Medical Lake, and Steptoe Butte State Park) and four sites in the Channeled Scablands (Unnamed DNR land on the Palouse River, Turnbull National Wildlife Refuge, Marcellus Shrub Steppe Natural Area Preserve, and BLM land near Hawk Creek) of Eastern Washington. Four sites were located on thin soil characteristic of the Channeled Scablands and two sites were located on deep loess soil characteristic of the Palouse Prairie. Sites were located across a range of latitudes, average temperatures, and average precipitation to cover locations in the two habitat types (Table 1). All sites were within 17.9 to 103 km of each other (Figure 1, Table 2). All sampling sites were public land, and permits were obtained where necessary. In addition to sampling sites, all tissue samples were collected from all plants that germinated in the greenhouse and survived until collection.

**Table 1.** Climate Data for Sampling Localities. Site names are as follows: Department of Natural Resources land on the Palouse River (DNR), BLM land near Davenport (Hawk Creek), Turnbull National Wildlife Refuge (Turnbull NWR), Marcellus Shrub Steppe Natural Area and Preserve (Marcellus NAP), Department of Natural Resources land near Medical Lake (Medical Lake), Steptoe Butte State Park (Steptoe Butte).

| Site Name | Annual Precipitation (mm) | Annual temperature (C) | Ecotype | Property type |
|---|---|---|---|---|
| DNR | 450.96 | 9.28 | Channeled Scabland | State |
| Hawk Creek | 343.21 | 8.13 | Channeled Scabland | Federal |
| Turnbull NWR | 454.80 | 8.31 | Channeled Scabland | Federal |
| Marcellus NAP | 299.07 | 8.94 | Channeled Scabland | State |
| Medical Lake | 411.33 | 8.01 | Palouse | State |
| Steptoe Butte | 530.24 | 7.95 | Palouse | State |

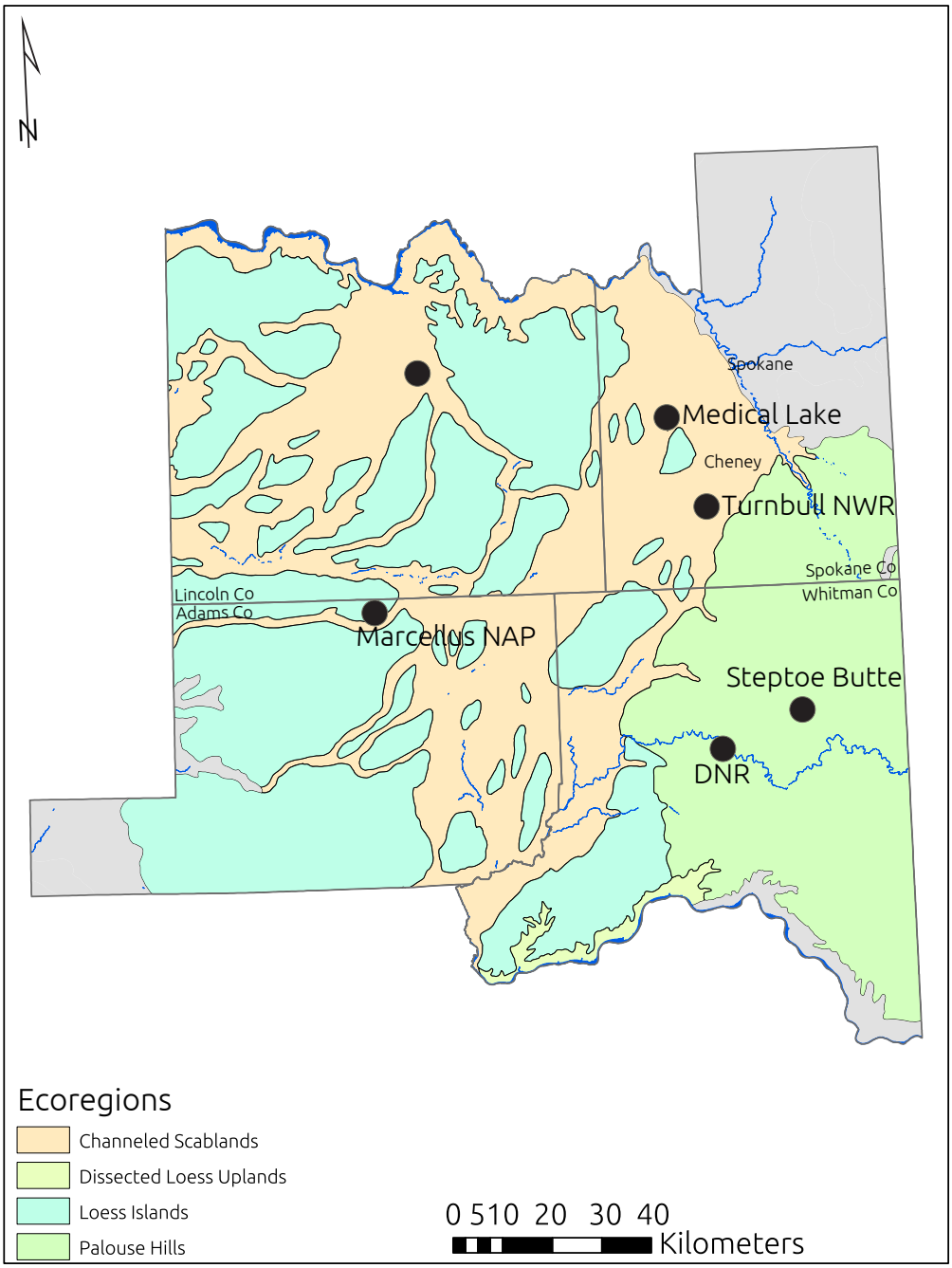


**Figure 1.** Sample site locations with EPA level IV ecoregions. Note: Although DNR is located in the Palouse Hills ecoregion, its location along the Palouse River has soil characteristics more similar to the Channeled Scablands and it was considered as a Scabland site. The Medical Lake site is also located on a Loess Island site that is not visible on this map (U.S. EPA 2013).

In addition to sampling on public land, seeds were obtained from BFI Native Seeds (BFI), and from the US Forest Service Coeur d'Alene Nursery (FS). The BFI seeds were collected from the same location as the Hawk Creek sample site, and then grown in seed increase lots. These seeds were a third-generation seed increase lot originally collected in 2010 (J. Benson Personal Communication May 28, 2019, J. Benson Personal communication 2010). BFI's Hawk Creek Provenance seeds are source-identified certified from the WSDA for genetic purity, cross-pollination, and weed contamination – seed production fields are inspected twice per growing season for this certification (J. Benson Personal Communication February 25, 2021). US Forest Service Region 1 Zone 1A seeds were collected from a much larger area and represent what the US Forest Service considers the seed provenance for the region including EWU's restoration site in Cheney, Washington (N. Robertson, Personal Communication 2019). BFI is contracted to grow seed increase lots for the US Forest Service controlling for genetic artifacts that may have appeared as a result of different nursery practices. The US Forest Service seeds were a second-generation seed increase lot (J. Benson Personal Communication May 28, 2019). Seeds from BFI and the Forest Service were planted in late April in a greenhouse at Eastern Washington University, with one seed per container using Sta-Green Potting Mix Plus Fertilizer and watered often enough to keep the soil moist.

*Plant Tissue Sampling*

To test for differentiation between plants growing in Palouse Prairie compared to Channeled Scablands, plant tissue samples were collected from six locations across both habitats. Two locations had deep loess soils characteristic of the Palouse Prairie ecotype,

and four locations had thin rocky soil that is characteristic of the Channeled Scablands. In all sampling locations plants were randomly selected by tossing an object.

To determine relatedness between local *P. spicata* populations and commercial seed, plants were germinated from two nurseries that provide a provenance for this region, the DNA from nursery sources was compared to the DNA from the six sampling locations in the Northern Palouse and Channeled Scablands. Additionally, plant tissue was sampled from the site where BFI Native Seeds collects for their Hawk Creek source-identified seed (J. Benson personal communication July 5, 2019) to compare relatedness after seed production generations.

*DNA Extraction, Library Preparation, and Genotyping by Sequencing*

Approximately 24 mg of dry leaf tissue from each plant was spun on a Retsch mixer mill 300 (Retsch GmbH, Haan, Germany) with Ballcone steel shot 1/8 inch at 28 Hz for 4 minutes to lyse plant cells. DNA was extracted using a MagMax™ Plant DNA kit and a KingFisher™ Flex Purification System (ThermoFisher, Waltham, MA) per manufacturer's protocol. DNA concentrations were normalized to 12ng/μL using an epMotion P5073 (Eppendorf, Hamburg, Germany) and then restricted with PstI-HF and Msp1 at 37°C for 2 hours, then 65 °C for 20 minutes, barcoded with Illumina barcodes (Illumina, San Diego, CA), amplified with a Simpliamp TC (ThermoFisher, Waltham, MA), multiplexed, and sequenced with a HiSeq 2500 (Illumina, San Diego, CA) by Utah State University in Logan, UT.

**Table 2.** Pairwise geographic distances between all sampling sites

| Distance between Sample Sites (km) | | | | | |
|---|---|---|---|---|---|
| | Steptoe Butte | DNR | Hawk Creek | Turnbull NWR | Marcellus NAP |
| DNR | 17.9 | | | | |
| Hawk Creek | 103 | 98.9 | | | |
| Turnbull NWR | 45.1 | 49.5 | 66.2 | | |
| Marcellus NAP | 86.3 | 70.3 | 48.7 | 67.8 | |
| Medical Lake | 64.8 | 68.5 | 51.2 | 20.4 | 68.7 |

*Filtering data*

Sequenced DNA was demultiplexed and aligned to a rough *P. spicata* genome by S. Larson at the Forage and Range Research Station in Logan, Utah. There were 25010 genetic markers per plant, a minimum allele frequency of 0.05, at least 5 read counts per marker per plant, and less than 30% of the data were missing. The genotypic data were converted into a genind object using the "adegenet" package in R (Jombart 2008, Jombart and Ahmed 2011,). To make the data set more manageable and to pseudo correct for linkage disequilibrium, the genomic data were filtered. All loci with missing data were removed, then the dataset was filtered down to only informative loci with the "poppr" package in R (Kamvar et al. 2014, Kamvar et al. 2015, R Core Team 2020). The resulting dataset contained 2648 genetic markers for all 92 plant samples. Basic statistics were calculated using the "basic.stats" function in the "hierfstat" package (Goudet and Jomabart 2020).

*Relatedness of Commercial seed and inbreeding*

To determine which commercial seed source was more closely related to sampling sites Nei's Genetic Distance (Nei's D) was calculated between populations using the R package "hierfstat" (Goudet and Jombart 2020, R Core Team 2020). Then an unweighted pair group method with arithmetic mean (UPGMA) dendrogram was drawn using Nei's D to visualize genetic differences between populations using the "ape" package in R (Paradis and Schliep 2019, R Core Team 2020). To determine inbreeding within subpopulations, Inbreeding Coefficient ($F_{IS}$) was calculated with 100 bootstraps and a 95% confidence interval using "hierfstat" in R (Goudet and Jomabart 2020, R Core Team 2020).

*Differentiation based on climate*

To test whether there were patterns of differentiation with climate, thirty-year norm data were downloaded from the PRISM database for both precipitation and average temperature (PRISM 2021, Table 1). These data were loaded into QGIS Version 3.16.3-Hannover and the data were extracted for each sampling location (QGIS Development Team 2020). Distance matrices were created for all physical sampling locations (excluding BFI and Forest Service plant material). Mantel tests with 999 permutations were run in the "ape" package in R to test for correlations between $F_{ST}$ and precipitation or mean temperature (Paradis and Schliep 2019).

*Population Structure in the Northern Palouse and Channeled Scablands*

Multiple methods were used to test for population genetic structure within all the sampled sites; a principal component analysis (PCA) with a mantel test for differentiation over geographic distance, a spatial principal component analysis (sPCA) which specifically testes for differentiation across geographic space, and Bayesian cluster analysis in STRUCTURE version 2.3.4 (Pritchard et al. 2000, Falush et al. 2003).

A PCA was used to investigate for genetic differences between populations of *P. spicata* in the Northern Palouse and Channeled Scablands. The filtered dataset was read into "R" (R Core Team 2020) with the read.structure function in the "adegenet" package (Jombart 2008, Jombart and Ahmed 2011). Then the "dudi.pca" function was used to analyze the filtered dataset for differences between populations. A mantel test was conducted to investigate for spatial patterns in the data. After the mantel test, the sPCA function was used in the adegenet package to analyze for both spatial and genetic structure. The geographic locations were read into R for all sampling locations (R Core

Team 2020). The plants that were grown from seed in the greenhouse were arbitrarily assigned one of two locations that belong to BFI Native Seeds for use with the sPCA. R did not recognize some of the sampling locations as different from each other. In order to use them for my analyses, I used the jitter function with a factor of 1.01 in the R base package to add some noise to the geographic locations. A similar technique has been used for plants grown in a greenhouse that only had a general source location for the seed (Massatti el al. 2018). Eigenvalues in the sPCA differ from those in normal PCA in that they represent both genetic diversity between sites, and geographic location within the area sampled (Jombart 2015). After running the sPCA on the dataset the "colorplot" function was used to visualize genetic differentiation across geographic space for *P. spicata*.

STRUCTURE version 2.3.4 was used to assign individuals to populations (Pritchard et al. 2000, Falush et al. 2003). One benefit of using STRUCTURE is that it attempts to account for linkage disequilibrium by grouping populations so they are not in disequilibrium (Pritchard et al. 2000). The filtered dataset was converted to a STRUCTURE formatted file with PGDSpider version 2.1.1.5 (Lischer and Excoffier 2012). The dataset was then read into STRUCTURE and run for 11 runs per K with correlated allele frequencies and the admixture model (Pritchard et al. 2000, Falush et al. 2003). Each run had 100,000 Burn-ins and 100,000 Markov chain Monte Carlos. The results of the runs were analyzed for best K using the Evanno method in CLUMPAK and STRUCTURE HARVESTER; then DISTRUCT was used to visualize the results (Rosenburg 2004, Evanno et al. 2005, Earl and vonHoldt 2012, Kopelman et al. 2015).

## Results

### *Genetic Distance, Inbreeding, and Local Adaptation*

Slight differentiation was observed between populations of *P. spicata* in the northern Palouse Prairie and Channeled Scablands of Washington ($F_{ST}$ 0.1441). All $F_{IS}$ values were negative indicating that they are not inbred, however Steptoe Butte samples had the lowest inbreeding coefficient, and the BFI samples had the highest (Figure 2). There are two distinct genetic groups within all samples, one contains Forest Service propagules, Steptoe Butte, and DNR; the other group contains BFI propagules with Medical Lake, Turnbull NWR, and Marcellus NAP, and Hawk Creek. The two sampling locations with the least genetic distance were Hawk Creek and Medical Lake (Nei's D = 0.0085), and the Two sampling locations with the most genetic distance were Forest Service and Medical Lake (Nei's D = 0.3229, Figure 3). No correlations were observed between temperature and genetic distance ($p=0.203$), or precipitation and genetic distance ($p=0.863$).

### *Population Structure of P. spicata in the Northern Palouse and Channeled Scablands*

The first principal component in the PCA explained nearly four times more variation than any other principal component (Figure 4). There was a significant genetic pattern associated with geographic location for *P. spicata* in the Northern Palouse and Channeled Scablands in Washington (Mantel test $p \leq 0.001$). Two genetic groups were identified using sPCA, one in the North and one in the south (Figure 5). There was more variation between sampling sites than within sites (Figure 6). There were no signs of differentiation between Palouse and Scabland sites, instead genetic differentiation is along a latitudinal gradient from south to north (Figure 6).

The best K value was chosen by ΔK for the STRUCTURE results in accordance with Evanno et al. (2005, Figure 7). Of the two commercial propagule sources tested, BFI plant material is more closely related to Medical Lake and Turnbull NWR, which are within 10 km to the North and South of the restoration site, respectively (Figure 8). There are two different genetic clusters of *P. spicata* in the Northern Palouse and Channeled Scablands (Figure 8). The first cluster includes the Northern Populations with seeds from BFI, and the second cluster includes the Southern Populations with seeds from the Forest Service (Figure 8).

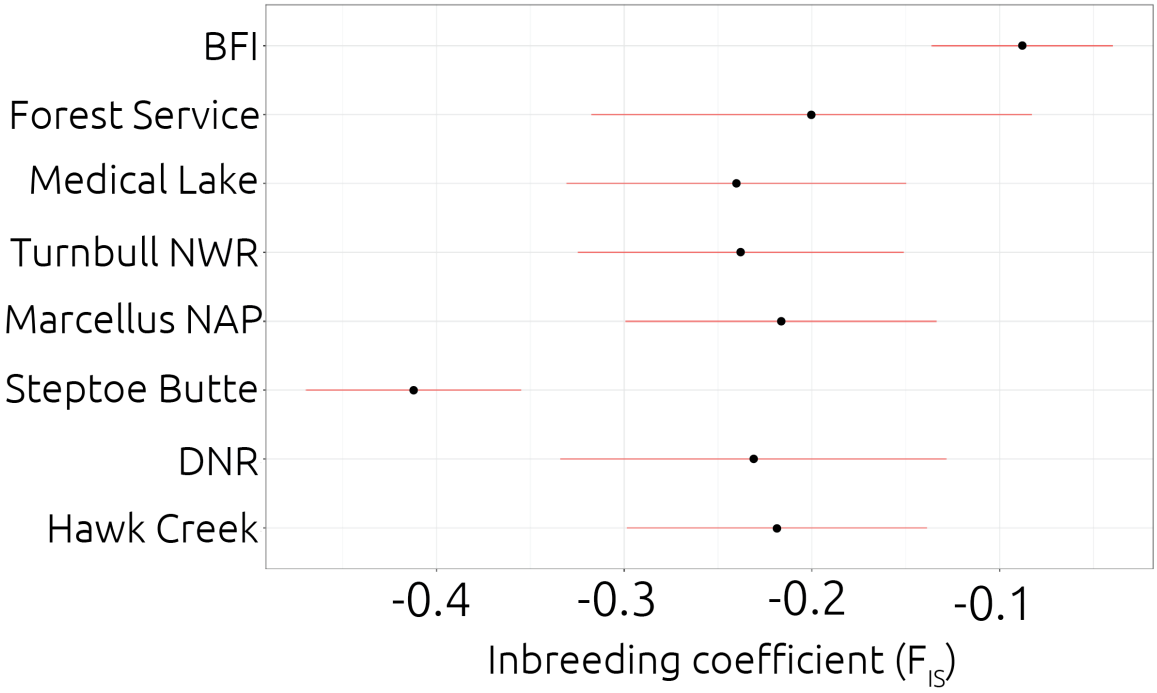


**Figure 2.** Bootstrapped inbreeding coefficient ($F_{IS}$) for each population with 95% confidence intervals. More negative values indicate less fixation of alleles within the population. Negative numbers indicate a population without inbreeding, a positive $F_{IS}$ would represent an inbred population.

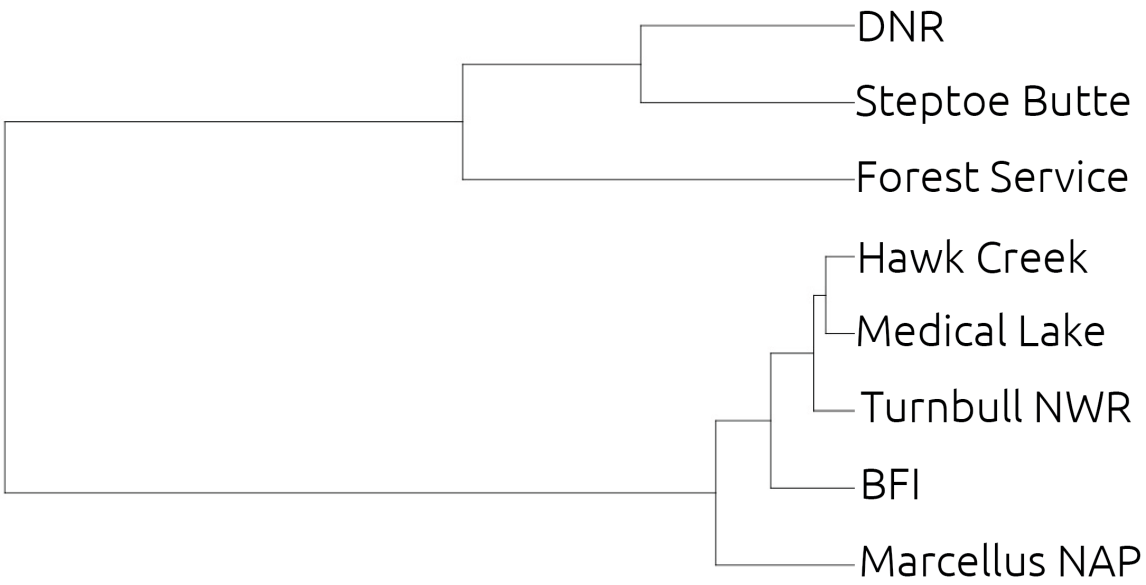


**FIGURE 3**. UPGMA dendrogram using Nei's Genetic Distance between sampling sites. The length of branches indicates the similarity between sites; sites with short branches between them are more closely related than sites with longer branches.

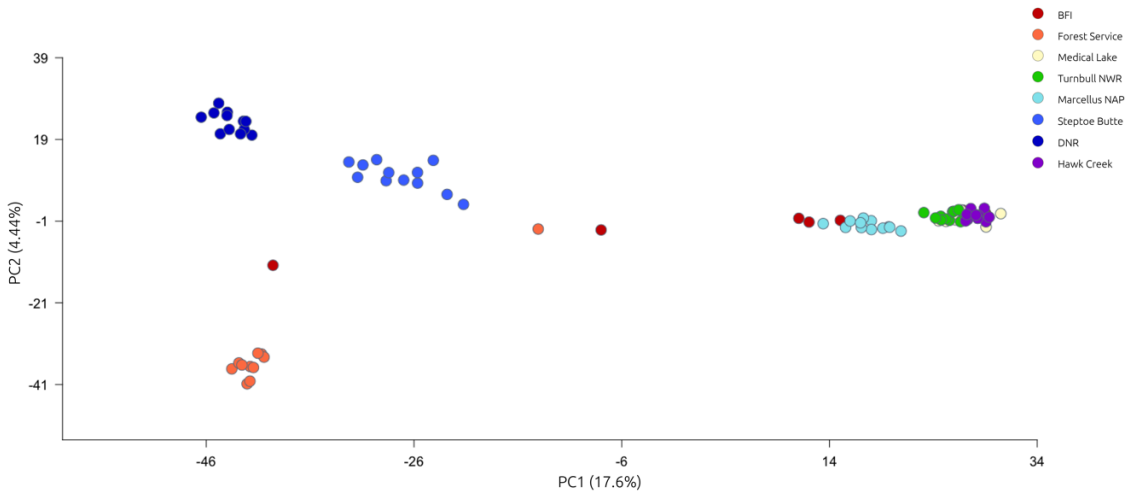


**Figure 4.** PC1 vs PC2 from PCA. Samples are color-coded by sampling location. The percentage of variation explained by each principal component is in parentheses.

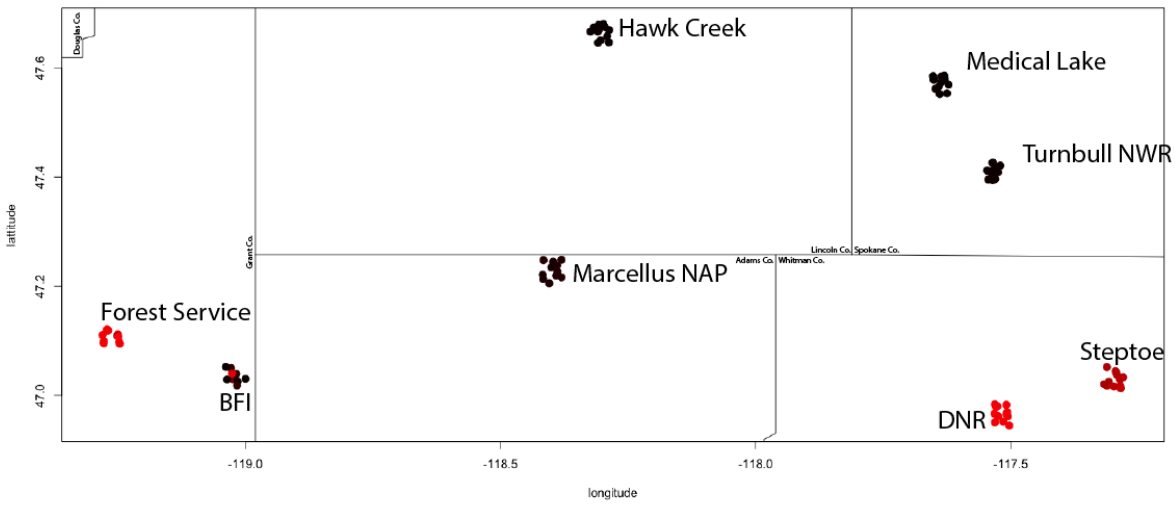


**Figure 5.** Color-coded plot of Spatial Principal Component Analysis. The first PC is color-coded in red. Note: because actual collection locations for BFI and Forest Service propagules were unknown, Forest Service and BFI were arbitrarily given locations for BFI Native Seeds who produces them. BFI provenance is collected from within the HAWK Creek Site (J. Benson Personal Communication July 5, 2019) and the Forest Service provenance is collected from a large region and pooled (N. Robertson personal communication 2019).

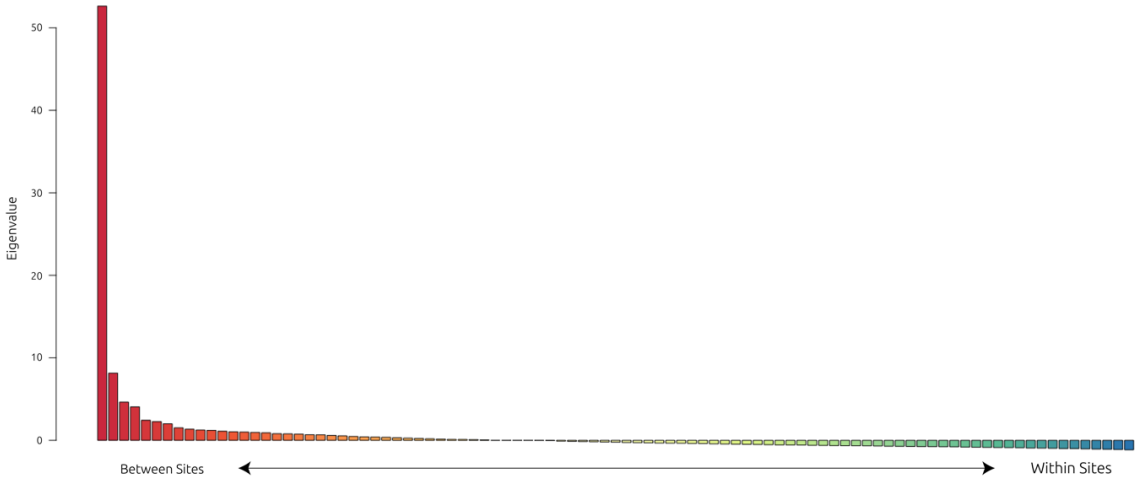


**Figure 6.** sPCA eigenvalues. Positive eigenvalues represent variation between sites and negative eigenvalues represent variation within sites.

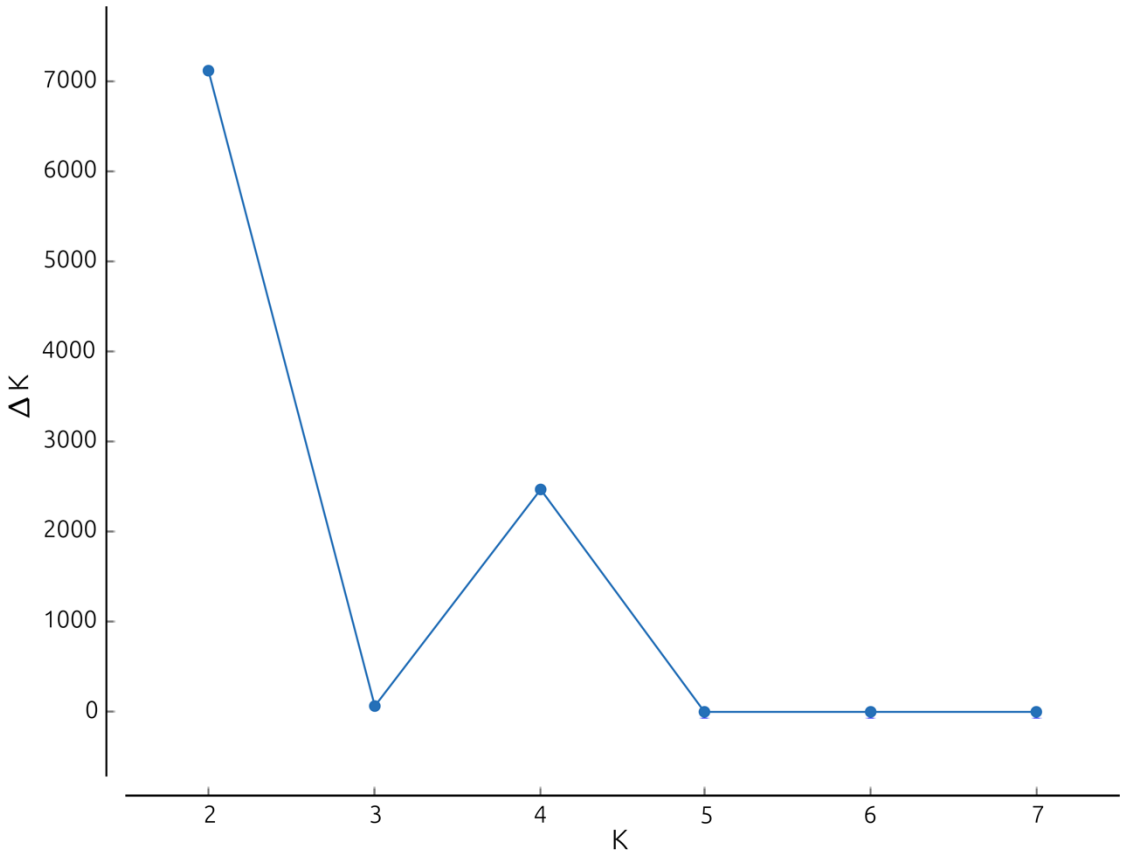


**Figure 7.** ΔK for each number of genetic clusters in STRUCTURE. K represents the number of genetic clusters identified by STRUCTURE. ΔK is based on the second order rate of change of the log probability of the correct number of genetic clusters, the value with the highest peak for ΔK represents the most likely number of clusters. (Evanno et al. 2005).

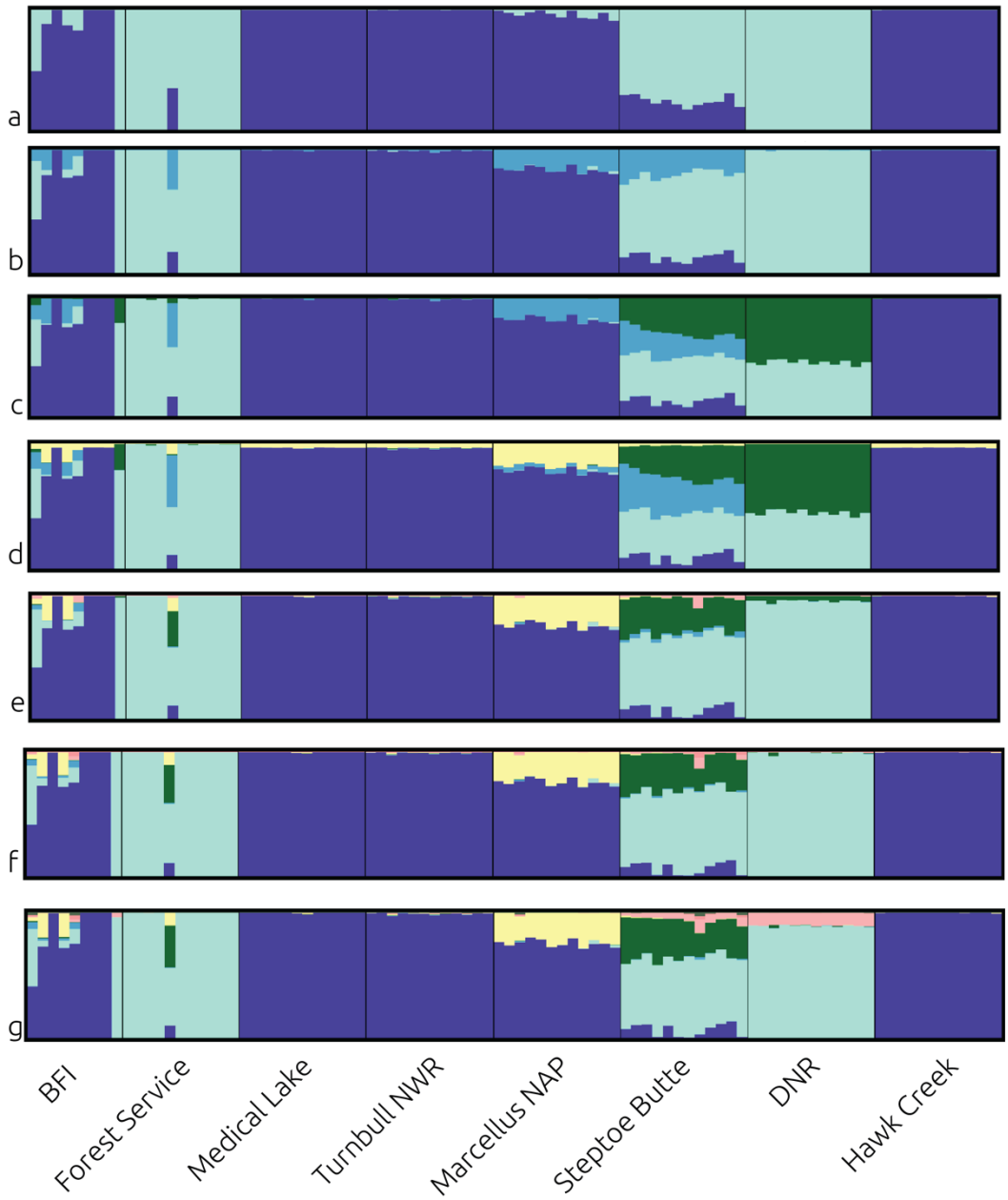


**Figure 8.** DISTRUCT print out for K=2 (a), K=3 (b), K=4 (c), K=5 (d), K=6 (e), K=7(f), and K=8 (g) runs in STRUCTURE. Colors represent different genetic clusters. Each column represents one individual, and each block surrounded by the black line represents a sample location/source.

## Discussion

My study identified two genetic clusters that varied along a latitudinal gradient from south to north and highlights the previously unstudied Northern Palouse and Channeled Scablands and did not support my hypothesis that populations would be differentiation with habitat (Figure 5). This study included two previously unstudied commercial provenance sources for *P. spicata* and investigated an area of the United States that to my knowledge has remained unstudied. My hypothesis that populations of *P. spicata* in the Northern Palouse and Channeled Scablands would have signs of inbreeding was not supported (Figure 2.). Plant material with a source-identified certification from the WSDA was more closely related to the northern sites, and Forest Service plant material from a large area was more closely related to southern sites, this did not support my hypothesis that source identified seeds (BFI) would be more closely related to sampling sites, instead BFI seeds were more appropriate for southern sites, and Forest Service Region 1 Zone 1 A for southern sites (Figure 3, Figure 5).

### *Relatedness of Sample Locations and Seed Sources*

There were two different genetic groups within plants sampled for this study. The southern group would likely benefit from being planted with seed from the Forest Service, and the Northern group, which surrounds EWU's Restoration site is more closely related to BFI's Hawk Creek source-identified seed. Massatti et al. (2018) found differentiation from south to north and suggested that populations differentiated genetically as they followed new habitable area during the last glacial retreat. Perhaps the two genetic clusters identified in this study are part of those identified by Massatti et al. (2018), or additional clusters; a meta analyses with these data would help to determine

this. At this small scale, gene flow may occur over a great distance through pollen and seed dispersal, preventing genetic differentiation between Channeled Scabland and Palouse habitats. Although there are commercial germplasm sources for *P. spicata* that have been investigated in relation to geographic range, they are genetically most similar to the area where they were originally sourced (Massatti et al. 2018). Adding these two sources and understanding their relationships to nearby populations will allow land managers to choose appropriate seed sources. Interestingly, although plants grown from BFI's Hawk Creek source-identified seed provenance are more similar to those sampled from Hawk Creek than Forest service plants, the Hawk Creek site is more closely related to Medical Lake and Turnbull NWR than it is to BFI propagules (Figure 3).

The genetic difference between seeds from BFI Native Seeds' Hawk Creek provenance and that of the Hawk Creek site deserves further examination. It is apparent in the dendrogram, STRUCTURE results, and inbreeding coefficient that the third-generation seed-increase lot from BFI's Hawk Creek provenance is slightly less genetically diverse than Hawk Creek samples and seems to have a different genetic makeup (Figure2, Figure 3, Figure 8). Three potential sources of genetic differences are, crosspollination from natural populations nearby, crosspollination from neighboring seed-increase fields, or insufficient sample size at the Hawk Creek site. It is also possible that the one individual that belongs entirely to a different genetic group from the BFI samples was mixed in during processing at the nursery. BFI Native seeds separated their seed production fields by at least 50 meters in accordance with the guidelines from the Washington State Department of Agriculture (J. Benson personal communication April 17, 2019) however at least one USDA native plant production facility separates their

grasses by 300 to 365 meters, which might be more realistic for *P. spicata* (Darris 2005). Without sampling from natural communities near BFI's seed-increase fields it is difficult to suggest which cause of hybridization is more likely, however it seems likely that the short separation between fields required by the WSDA might contribute. If nearby natural areas contain plants that are pollinating seed increase fields, the only way to avoid this would be to grow seed increase in an enclosed area, which would not be realistic for production. It is interesting to note that the Forest Service plants largely represent one gene pool (Figure 8). This might be an indicator that Forest Service plants grown at BFI are generally upwind of potential hybridization sources, so they have more genetic purity. The differences in inbreeding coefficient between BFI and Hawk Creek are likely the result of the three seed-increase generation distance from the natural population they were picked from and might not be noticeable in the first or second generation similar to FS plant material. During each generation of seed increase there are a plethora of opportunities to lose genetic diversity (Meyer and Monsen 1992, Basey et al. 2015). Although the cause of differentiation of BFI's Hawk Creek Provenance from the collection site is unknown, a seed lot that is only the first-generation seed would probably be a better genetic representation of the local gene pool.

*Management Implications*

My findings that there are two major genetic clusters in my sampling area suggest that seeds for a restoration in the Northern Palouse can be sourced from a nearby population that is at a similar latitude on either Palouse or Channeled Scabland habitat. Additionally, Steptoe Butte has the most within population variation suggesting that it might be a refuge for genetic diversity for the region.

Of all sampling locations, Steptoe Butte State park contains the most within population variation (Figure 2). Due to the amount of genetic diversity within this site, its conservation could help it continue to be a pool of genetic diversity for the region and make it useful for future restoration and rehabilitation projects. Additional population genetics studies that include Steptoe Butte State Park for other species and comparing management between it and other natural areas in the region would help identify 1. If Steptoe Butte is a refuge of genetic diversity for multiple native species within the region, and 2. What management practices if any could have led to the higher level of genetic diversity compared to other sites.

The identification of potential hybridization of a local propagule source at BFI native seeds points to a need for changed practices at native seed nurseries. Accidental use of native plant hybrids in restoration or remediation can introduce novel genetic material to an area beyond its natural range (Winkler and Massatti 2020). Although unintentional hybrid availability from nurseries has been observed there is rarely evidence that it may have happened at the nursery (Basey et al. 2015, Winkler and Massatti 2020). Recognition of practices at native seed nurseries that increase the chance of hybridization and mitigation of these practices would help keep local native plant provenances pure. Because the Hawk Creek provenance from BFI is produced under the guidelines of the WSDA source identified certification program this suggests that WSDA might need to reevaluate the guidelines of the program for predominantly outcrossing species such as *P. spicata.*

Although the BFI seed source is the least genetically diverse of the two commercial sources, it must be noted that it is not inbred and would still help establish a

genetically diverse, restored population at EWU's prairie restoration. Three seed-increase generations are near the limit before BFI refreshes their seed from a natural population (J. Benson Personal Communication May 28, 2019). Moreover, because Forest Service seeds are collected from a larger area, they would likely contain more genetic diversity than the initial starting Hawk Creek seed, making it difficult to compare the two sources. Studies on pollen dispersal for *P. spicata* would elucidate the possible need to separate seed production fields from each other by a greater distance to reduce the risk of accidental hybridization, however it would be nearly impossible prevent stochastic events that could lead to mixing of seed provenances.

Finally, this study highlights a need for continued research on the genetic structure and makeup of native plant materials used for restoration. More research is needed on management practices to improve the genetic purity of commercially grown restoration seed, and prevent the spread and novel genetic material beyond their native range. Others have already identified hybrids growing beyond their native range and suggest the use of genetic analyses to reduce the risk of transferring hybrids beyond their range (Winkler and Massatti 2020). Additionally, genetic analyses of other native plant species growing at Steptoe Butte and other large Palouse Prairie remnant sites would help justify continued preservation of these sites, and establish their potential as genetically diverse propagule sources for restoration.

LITERATURE CITED

Aavik, T., P. J. Edwards, R. Holderegger, R. Graf, and R. Billeter. 2011. Genetic consequences of using seed mixtures in restoration: A case study of a wetland plant *Lychnis flos-cuculi*. Biological Conservation 145:195-204.

Basey, A. C., J. B. Fant, and A. T. Kramer. 2015. Producing native plant materials for restoration: 10 rules to collect and maintain genetic diversity. Native Plants 16:37-52.

Bischoff, A., T. Steinger, and H. Müller-Schärer. 2010. The importance of plant provenance and genotypic diversity of seed material used for ecological restoration. Restoration Ecology 18:338-348.

Bowlick, F., K. Andermatt, C. Besmar, J. Erbe, R. Lopez III, J. White, D. W. Goldberg, and K. M. Carroll. 2015. Defining the Palouse: Using overlap analysis to delineate an informal region. Association of the Pacific Coast Geographers. 77:40-51.

Bretz, J. H. 1925. The Spokane flood beyond the channeled scablands. The Journal of Geology 33:97-115.

Bryan, K. 1926. The "Palouse Soil" problem with an account of elephant remains in wind-borne soil on the Columbia Plateau of Washington. US Geological Survey Bulletin 790-B:41-46.

Clausen, J., D. D. Keck, and W. M. Hiesey. 1941. Regional differentiation in plant species. The American Naturalist 75:231-250.

Crémieux, L. A. Bischoff, H. Müller-Schärer, and T. Steinger. 2010. Gene flow from foreign provenances into local plant populations: Fitness consequences and implications for biodiversity restoration. American Journal of Botany 97:94-100.

Darris, D. C. 2005. Seed production and establishment of Western Oregon Native Grasses. USDA Forest Service Proceedings. RMRS-P-35: 119-128.

Davies, K. W., and D. D. Johnson. 2017. Established perennial vegetation provides high resistance to reinvasion by exotic annual grasses. Rangeland Ecology and Management 70:748-754.

Dolan R. W., D. L. Marr, and A. Schnabel. 2008. Capturing genetic variation during ecological restorations: An example from Kankakee Sands in Indiana. Restoration Ecology 16:386-396.

Duffin, A. P. 2007. Plowed Under: Agriculture and environment in the Palouse. Seattle: University of Washington Press.

Earl, D. A., and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361.

Ellstrand, N. C., and D. R. Elam. 1993. Population genetic consequences of small population size: Implications for plant conservation. Annual Review of Ecology and Systematics 24:217-242.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology 14:2611-2620.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics Society of America 164:1567-1587.

Fant, J. B., R. M. Holmstrom, E. Sirkin, J. R. Etterson, and S. Masi. 2008. Genetic structure of threatened native populations and propagules used for restoration in a clonal species, American beachgrass (*Ammophila breviligulata* Fern.). Restoration Ecology 16:594-603.

Fu, Y., and D. Thompson. 2006. Genetic diversity of bluebunch wheatgrass (*Pseudoroegneria spicata*) in the Thompson River valley of British Columbia. Canadian Journal of Botany 84:1122-1128.

Gibson, A., C. R. Nelson, S. Rinehart, V. Archer, and A. Eramian. 2019. Importance of considering soils in seed transfer zone development: Evidence from a study of the native *Bromus marginatus.* Ecological Applications 29: e01835.

Goudet, J., and T. Jombart. 2020. hierfstat: Estimation and tests of hierarchical F-Statistics. R package version 0.5-7. https://CRAN.R-project.org/package=hierfstat

Hitchcock, C. L., and A. Cronquist. 2018. Flora of the Pacific Northwest: an illustrated manual. University of Washington Press, Seattle.

Jensen, K. B., Y. F. Zhang, and D. R. Dewey. 1990. Mode of pollination of perennial species of the *Tritaceae* in relation the genomically defined genera. Canadian Journal of Plant Science 70:215-225.

Johnson, M. T. J., and A. A. Agrawal. 2005. Plant genotype and environment interact to shape a diverse arthropod community on evening primrose (*Oenothera biennis*). Ecology. 86:874-885.

Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.

Jombart, T. 2015. A Tutorial for the spatial Analysis of Principle Components (sPCA) using *adegenet* 2.0.0. Imperial College London 1-50

Jombart, T. and Ahmed I. 2011. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070-3071.

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

Kamvar Z. N., J. F. Tabima, N. J. Grünwald. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2:e281.

Keller, W. 1948. Interpretation of self-fertility in grasses by frequency distributions. Journal of the American Society of Agronomy 40:894-900.

Kilkenny, F., B. St. Clair, and M. Horning. 2013. Climate change and the future of seed transfer zones. USDA Forest Service Proceedings. RMRS-P-69:87-89.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources 15:1179-1191.

Larson, D. L., J.B. Bright, P. Drobney, J. L. Larson, N. Palaia, P. A. Rabie, S. Vacek, and D. Wells. 2011. Effects of planting method and seed mix richness on the early stages of tallgrass prairie restoration. Biological Conservation 144:3127-3139.

Larson, S. R., T. A. Jones, and K. B. Jensen. 2004. Population structure in *Pseudoroegneria spicata* (*Poaceae: Triticeae*) modeled by Bayesian clustering of AFLP genotypes. American Journal of Botany 91:1789-1801.

Lischer, H. E. L., and L. Excoffier. (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28:298–299.

Massatti, R., H. R. Prendeville, S. Larson, B. A. Richardson, B. Waldron, and F. Kilkenny. 2018. Population history provides foundational knowledge for utilizing and developing native plant restoration materials. Evolutionary Applications 11:2025-2039.

Meyer, S. E., and S. B. Monsen. 1992. Genetic considerations in propagating native shrubs, forbs, and grasses from seed. Western Forest Nursery Association Meeting. 1-8.

Noss, R., E. Laroe, and M. Scott. 1995. Endangered ecosystems of the United States: A preliminary assessment of loss and degradation. US National Biological Service: Report nr 28.

Ogle, D., L. St. John, and T.A. Jones. 2010. Plant guide for bluebunch wheatgrass (*Pseudoroegneria spicata*). USDA Natural Resources Conservation Service.

Paradis E., and K. Schliep. 2019. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526-528.

PRISM Climate Group. 2021. Oregon State University. https://prism.oregonstate.edu/normals/

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

QGIS Development Team. 2020. QGIS Geographic Information System. Open Source Geospatial Foundation Project.

R Core Team 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rodriguez-Quilon, I., L. Santos-del-Blanco, M. Jesus Serra-Varela, J. Koskela, S. C. Gonzalez-Martinez, and R. Alia. 2016. Capturing neutral and adaptive genetic diversity for conservation in a highly structured tree species. Ecological Applications 26:2254-2266.

Rosenberg, N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4:137-138.

Sanchez-de Leon, Y., and J. Johnson-Maynard. 2013. Ecosystem carbon storage and cycling in restored and native grasslands of the Palouse region. Soil Science Society of America Journal 77:929-940.

Schoen, D. J., and A. H. D. Brown. 2001. The conservation or wild plant species in seed banks. BioScience 51:960-966.

Schröder, R., and R. Prasse. 2013. Cultivation and hybridization alter the germination behavior of native plants used in revegetation and restoration. Restoration Ecology 21:793-800.

Schröder, R., and R. Prasse. 2013. From nursery into nature: A study on performance of cultivated varieties of native plants used in re-vegetation, their wild relatives and evolving wild x cultivar hybrids. Ecological Engineering 60:428-437.

St. Clair, J., F. F. Kilkenny, R. C. Johnson, N. L. Shaw, and G Weaver. 2013. Genetic variation in adaptive traits and seed transfer zones for *Pseudoroegneria spicata* (bluebunch wheatgrass) in the northwestern United States. Evolutionary Applications 6:933-948.

U.S. Environmental Protection Agency. 2013. Level III and IV ecoregions of the continental United States: Corvallis Oregon, U.S. EPA National Health and Environmental Effects Research Laboratory

Winkler, D. E., and R. Massatti. 2020. Unexpected hybridization reveals the utility of genetics in native plant restoration. Restoration Ecology 28:1047-1052.

***Curriculum Vitae***

Author: Ethan Timothy Bean

Birth: September 10, 1988. Lansing, Michigan

EDUCATION:

**2012-2014** University of Washington, Tacoma
Declared Major: Environmental Studies

**2016** B.S., Biology, Washington State University
*Magna Cum Laude*

HONORS:

**2012-2013** University of Washington Certificate of High Scholarship

**2014-2016** Washington State University President's Honor Roll

**2014-2015** Washington State University Transfer Achievement Award

**Apr 2015** Nominated by WSU faculty and participated in Governor Inslee's Statewide Virtual Climate Conversation

**Dec 2015** Nominated by WSU faculty and participated in Governor Inslee's Statewide Virtual Climate Conversation

**2018-2020** Graduate Service Appointment – Eastern Washington University

**VOLUNTEER WORK:**

**PROFESSIONAL EXPERIENCE**

**July 2006 - March 2009** United States Army – Blackhawk maintainer and crew chief – Performed maintenance on Blackhawk helicopters, deployed to the Nineveh Province of Iraq as a Blackhawk maintainer and crew chief in support of Operation Iraqi Freedom, performed maintenance and flew as a member of a 4-person crew

**March 2009 - July 2012** United States Army Special Operations Aviation Regiment – Crew Chief – Operated as a member of a 4-5 person Blackhawk crew on multiple deployments worldwide, in charge of all scheduled and unscheduled maintenance for an assigned Blackhawk helicopter, keeping it flyable and ready to deploy with short notice, during flights, responsible for the safety of other crewmembers and passengers

**2015 - 2016** Volunteer, Mt. St. Helens Disturbance Ecology research – data entry, research site set up/take-down, assisted with maintenance of research plots

**Summer 2016** Volunteer, WSU Vancouver Soils Lab – gained proficiency with instruments such as CO2 Gas Analyzer, TOC analyzer and Data Loggers, Assembled an irrigation chamber for soil experiments

**Spring 2017** Volunteer, Coeur d'Alene Forest Service Nursery – transplanted native plants, sowed several species of native plants from seed and cuttings using machines and by hand, made cuttings of native species to start new plants, grafted trees for pathogen resistance and nursery stock

**June 2017 - May 2018** Coeur d'Alene Forest Service Nursery – Biological Science Aid – Collected data for a long-term project looking for families of five-needle pine resistant to the fungal pathogen, white pine blister rust, sewed, thinned, weeded, and transplanted native conifers for use in the blister rust program and various Forest Service orchards, collected data for a spore dispersal study by identifying spores under a microscope and getting accurate counts for calculation of dispersal densities on National Forest land, grafted treed for pathogen resistance, assisted with planting and transplanting of native trees, forbs, and grasses for production for various projects

**July 2017 - Present** Bean and Pie – Founder and Co-owner – Created original recipes to start the business, provide assistance with marketing, management decisions, and guiding the direction of the company

**February 2018 - Present** Army National Guard – Infantry Officer – In charge of training, admin, and planning operations for a group of 20 to 40 individuals, Laser and Radiation Safety Officer for 1-161 Infantry Regiment.

**September 2018 - April 2020** Eastern Washington University – Teaching Assistant – Assist professors with set-up, instruction, grading and grading discussion sections, available to students during office hours to answer additional questions and provide feedback on assignments

**January 2020 - Present** Eastern Washington University – Prairie Restoration Plan – Author on genetics section of restoration plan for a 155-acre restoration project