RESTORATION ECOLOGY
The Journal of the Society for Ecological Restoration
SER SOCIETY FOR ECOLOGICAL RESTORATION

RESEARCH ARTICLE

Assessing population genomic structure and polyploidy: a crucial step for native plant restoration

Bryce A. Richardson[1,2], Rob Massatti[3], Nurul Islam-Faridi[4], Skylar Johnson[5], Francis F. Kilkenny[6]

Establishing an effective restoration program requires baseline genetic information to make sound decisions for seed increase and transfer. For many plants this information is lacking, especially among native forbs that are critical for pollinator health. *Erigeron speciosus* is a widespread, perennial forb occupying montane environments in the western United States and Canada. This species is important in fostering pollinator diversity. Our study examines the population genetic patterns across the species range using reduced-representation sequencing and surveys for genome duplication using flow cytometry and cytology. These genomic tools provide critical information for seed increase and seed transfer, necessary for restoration programs. Population genetic differentiation (F_{ST}) average was 0.13 and ranged from 0.05 to 0.24 among 23 collection sites. Model-based Bayesian clustering supported a model with collection sites grouped into two populations, occupying distinct geographic regions of this species range. A genetic distance-based neighbor-joining tree also supported this division. Flow cytometry of 53 samples from 17 populations had 2C values that ranged from 1.7 to 3.6 pg with a mean 2C value of 2.3 pg. Putative triploids were found in two individuals from one collection site. The spatial distribution of genetic structure supports regionally based taxonomic descriptions of two varieties: *speciosus* in the North and *macranthus* in the South. This assessment of genetic structure and genome duplication describes an effective approach in developing baseline genetic information for restoration species, especially those species that may harbor complex taxonomy and polyploidy.

Key words: Asteraceae, genetic diversity, polyploidy, population genetic structure, restoration, whole-genome duplication

Implications for Practice

- Population genomic approaches can be applied to a variety of species and inform the development of restoration programs. As illustrated in this study, distinct genetic groups (i.e. populations) warrant separate seed increase programs.
- Selection of seed sources within genetic groups should also be chosen with prudence given the strong association between genetic and geographic distance or within groups that have broad environmental gradients.
- A survey of cytological variation among seed collections destined for a species restoration program is a necessary step for seed increase. Cytotype screening would prevent combining mixed polyploid individuals that may reduce seed yield in cultivation due to genomic incompatibilities.

Introduction

As efforts to restore degraded landscapes increase in scope and intensity worldwide (Mills et al. 2020), an important strategy to foster ecosystem health has been to diversify restoration seed mixes to include native forbs. Forbs provide critical resources to support pollinator networks (e.g. Dumroese et al. 2016; Burkle & Runyon 2019) and regional avian fauna (e.g. Dumroese et al. 2015), as well as increase plant community resiliency by diversifying the representation of functional traits (Isbell et al. 2011; Funk 2021). Despite their utility, several challenges hinder the increased use of forbs in restoration. In particular, forbs utilize a broad array of life history strategies, e.g. breeding systems and pollination mechanisms, that make predicting range wide patterns of genetic diversity and differentiation difficult, and therefore using and developing genetically appropriate restoration materials impractical (e.g. Massatti et al. 2020). Moreover, and similar to other groups of plants utilized in restoration like graminoids, many widespread forbs of restoration interest display spatially explicit patterns of morphological variation that have been used to inform taxonomy at the infraspecific level; this variation may reflect a variety of intraspecific relationships along an

Author contributions: BAR, RM, FFK conceived the experiment; FFK, BAR coordinated tissue collections; BAR, NI-F, RM, SJ analyzed the data; BAR wrote the manuscript; all authors contributed to manuscript revision.

[1]Rocky Mountain Research Station, USDA Forest Service, Moscow, ID 83843, U.S.A.
[2]Address correspondence to B. A. Richardson, email brichardson02@fs.fed.us
[3]Southwest Biological Center, US Geological Survey, Flagstaff, AZ 86001, U.S.A.
[4]Southern Research Station, USDA Forest Service, College Station, TX 77843, U.S.A.
[5]Independent Researcher, Pullman, Washington 99163, U.S.A.
[6]Rocky Mountain Research Station, USDA Forest Service, Boise, ID 83702, U.S.A.

doi: 10.1111/rec.13740
Supporting information at:
http://onlinelibrary.wiley.com/doi/10.1111/rec.13740/suppinfo

evolutionary continuum (Coates et al. 2018). Additional complexity may result from genome duplication (i.e. polyploidy; Kramer et al. 2018) and hybridization, which is often difficult to assess without appropriate data.

Given the biological complexities inherent across species, understanding species-specific evolutionary processes influencing differentiation and diversification is an essential prerequisite to developing restoration seed mixes (Breed et al. 2019). Fortunately, genomic sequencing data can provide a broad foundation for implementing informed seed collection, seed increase, and dissemination of seeds or seedlings (i.e. seed transfer). For example, population-level genetic diversity and structure data can inform seeding efforts (e.g. Durka et al. 2017) to avoid inbreeding or outbreeding depression (Fenster & Galloway 2000; Hufford et al. 2012). Inbreeding or outbreeding depression may reduce plant performance in the early generations after a restoration treatment when performance of native plant materials (NPMs) is critical to, e.g. mitigate the influence of invasive species. In addition, these data can guide seed collecting efforts by illuminating how adaptive genetic variation correlates to environmental gradients (Shryock et al. 2015; Faske et al. 2021) and by providing baseline information on how species are demographically connected (e.g. through migration) across their ranges (Massatti et al. 2018); these biological characteristics of species are critical to consider as seed sources are selected for development into new NPMs using different strategies (e.g. the various mix and match approaches described in Bucharova et al. 2019). Furthermore, flow cytometry and cytogeography approaches can assess the frequency and geographic patterns of polyploidy (McArthur & Sanderson 1999; Laport et al. 2012), which may affect performance across a species' environmental distribution (e.g. Ramsey 2011; Zaiats et al. 2020; Richardson et al. 2021) and intraspecific compatibility (Kramer et al. 2018). Finally, population genetic studies can define or add support to describe intraspecific variation or taxonomic treatments (Richardson & Meyer 2012; Widener & Fant 2018) and elucidate interspecific hybridization (Winkler & Massatti 2020; Ruiz et al. 2021). These approaches are transferable across species with little modification and can be completed rapidly (i.e. 1–2 years) to provide information on seed sourcing for seed increase or design and interpretation of common garden studies.

Herein, we analyze genomic patterns of *Erigeron speciosus* DC. (Aspen fleabane; Asteraceae), a cross-pollinating native forb with spatially structured morphological variation that has been used to describe taxonomic varieties. As such, we enhance available information about a future restoration species that exhibits a particular, though not uncommon, suite of biological traits. We demonstrate a workflow that combines population genomics and cytogeographic approach to inform downstream management efforts related to seed collecting, seed transfer, and NPM development. Specifically, we generate baseline genetic data using next-generation sequencing, flow cytometry and microscopy to survey genetic diversity, population genetic structure, and assess cytogeographic variation. We illustrate the restoration ramifications of these patterns in the context of management across western North America, and we set results in the broader context of restoration as it pertains to international efforts and the UN Decade on Ecological Restoration.

Methods

Study Species

Erigeron is a speciose genus that likely originated in North America (Noyes 2000). Western North America is home to a diverse array of both localized and widespread *Erigeron* species (Nesom 2006), including *Erigeron speciosus*, an ecologically important, perennial forb occupying forested and nonforested montane communities from the U.S.–Mexico border to southern Canada (Gucker & Shaw 2018). *Erigeron speciosus* has been identified as a key restoration species based on its ecological breadth across shrubland, forested, and alpine communities, the ability for an individual to survive for more than a decade (Inouye 2008), its prolonged flowering period in mid-summer, and pollinator visitation by numerous insect families (Burkle & Irwin 2009).

Different taxonomic treatments have been proposed for *E. speciosus* and sister taxa based primarily on hairs and glandularity of phyllaries, leaves, and stems. Cronquist et al. (1994) recognized two varieties: *speciosus* and *macranthus* (Nutt.) Cronquist. These two proposed varieties have overlapping distributions, but var. *speciosus* is more prevalent in the Northwest (Idaho, western Montana, and Washington) and var. *macranthus* is found primarily in the South (Cronquist et al. 1994). Alternatively, Nesom (2006) suggests *E. speciosus* is a cohesive, albeit variable, taxonomic unit and where distinguishing morphological characters can intergrade with the sympatric *E. subtrinervis* Rydb., especially in the northern portions of their overlapping range (i.e. Idaho, Montana, and Wyoming). While some species within *Erigeron* are known for apomixis that is derived from auto or allopolyploidy (Noyes & Rieseberg 2000; Noyes 2006), polyploidy is unknown in *E. speciosus* and *E. subtrinervis*, which both have a base chromosome number $n = 9$ (Jones & Young 1983).

Sample Collection

Leaf tissue was collected from 23 sites of *E. speciosus* (Fig. 1) across steep environmental gradients in the central and northern Rocky Mountain region (Fig. S1). These collections primarily reflected USDA Forest Service Region 4's interest in developing this species for restoration, and therefore sampling was focused within this region. In addition, one population of *E. formosissimus* Greene was included as an outgroup. Individual samples were placed in envelopes with silica in the field, and then dried under vacuum at room temperature. Dried tissue was ground with a bead mill and DNA was extracted following the DNeasy plant extraction kit protocol (Qiagen, Germantown, MD, U.S.A.). DNA yields were standardized to 3 ng/μL prior to library preparation. While the intent was to have $n \geq 10$ for each population, during this process, some individuals were omitted due to poor DNA yields or amplification, leading to a collection site sample size ranging between 2 and 14 individuals (Table 1).

DNA Preparation and Sequencing

Genomic DNA samples, including outgroup specimens, were individually barcoded and processed into libraries using a restriction fragment-based procedure (Peterson et al. 2012)

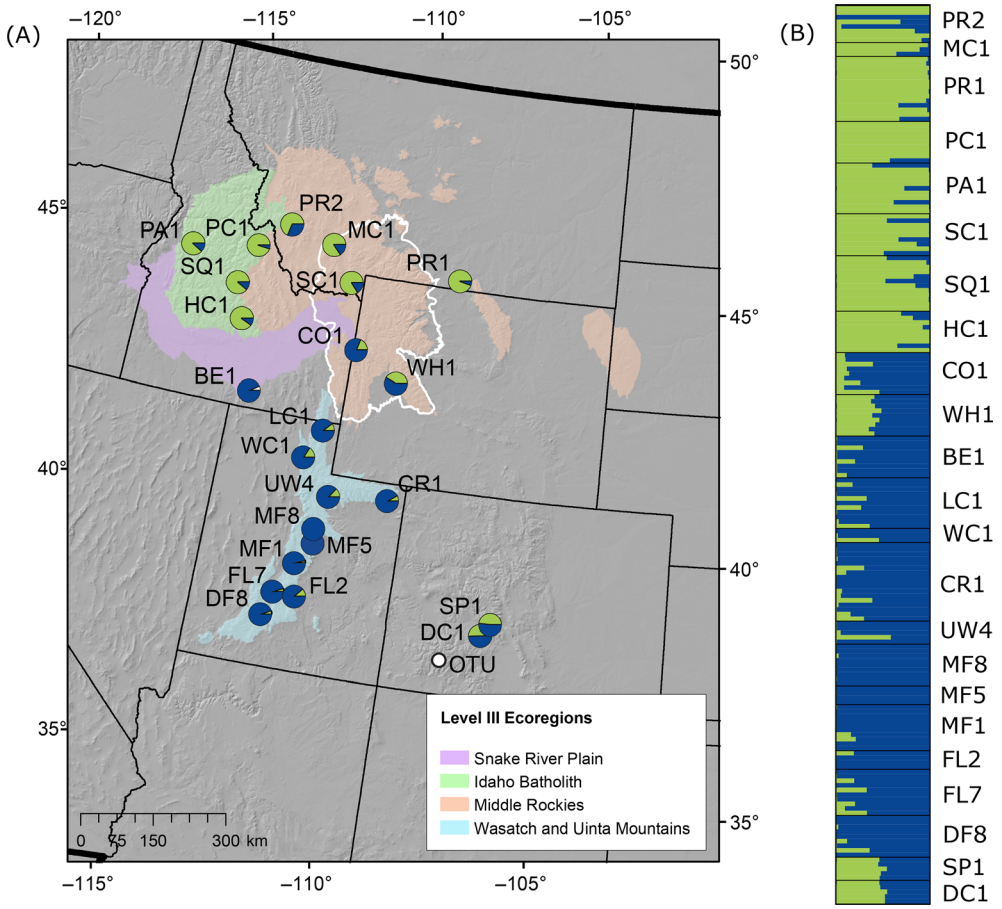

Figure 1. (A) Mapped locations of *Erigeron speciosus* collection sites. The pie charts show the location and STRUCTURE assignment probabilities for each collection site based on a two-population model ($K = 2$). The colored polygons indicate four Omernik level III ecoregions (Idaho Batholith, Middle Rockies, Wasatch, and Uinta Mountains). White outline shows the Greater Yellowstone Region boundary. The outgroup taxonomic unit (OTU) is *E. formosissimus*. (B) Barplot of individual sample STRUCTURE assignment probabilities and horizontal black lines separate collection sites.

modified to utilize separate indexing reads. DNA was doubly digested with EcoRI and MspI restriction enzymes, followed by the ligation of Illumina adaptor sequences and barcodes. Each sampled individual was barcoded using unique combinations of forward and reverse indexes. Ligation products were pooled and amplified using 12 cycles of polymerase chain reaction. A Pippin Prep (Sage Science, Beverly, MA, U.S.A.) was used to size select amplicons from 400 to 600 base pairs. Libraries were sequenced on a HiSeq 4000 (Illumina, San Diego, CA, U.S.A.) at the University of Oregon's Genomics and Cell Characterization Core Facility to generate single-end 100 base-pair reads.

Raw sequence data were demultiplexed using custom scripts and "fastq-multx" in ea-utils (Aronesty 2011). The remaining data processing and read assembly steps were accomplished using STACKS v2.53 (Catchen et al. 2013) and followed the r80 protocol detailed in Rochette and Catchen (2017). Within STACKS, the process_radtags script was used to exclude raw reads containing more than four low-quality sites and adapter contamination. Parameters affecting the assembly were assessed based on how parameter combinations affected r80 loci (i.e. those found in 80% of samples or more); the optimal parameter set associated with the plateau of the number of r80 loci was selected (Fig. S2; Paris et al. 2017; Rochette & Catchen 2017). Values used in the final assembly were read to initiate a new putative allele (-m in ustacks) = 3; number of mismatches allowed between the two alleles of a heterozygote sample (-M in ustacks) = 6; mismatches allowed between any two alleles of the population (-n in cstacks) = 6. The populations program in STACKS (Catchen et al. 2013) was executed to generate the datasets used in analyses under the following settings: minimum number of populations a locus must be present in to process a locus (-p) = 2; minimum minor allele frequency required to process a nucleotide site at a locus (--min_maf) = 0.02; maximum observed heterozygosity required to process a nucleotide site at a locus (--max_obs_het) = 0.7; minimum percentage of individuals in a population required to process a locus for that populations (-*r*) = 0.5; and the maximum value to keep an F_{ST} measurement (--*p*-value cutoff) = 0.05.

Table 1. *Erigeron speciosus* genomic data processing information for each collection site. *K* is the STRUCTURE assignment to either the North or South population. *N* is the number of individuals included per locality (excludes individuals with few raw reads), utilized reads is reads used for data analyses after filtering, coverage is average coverage per locus. Genetic diversity statistics per sampling site include *p* (variance)—mean frequency of most frequent allele per locus; H_E (variance)—mean expected heterozygosity; π (variance)—mean nucleotide diversity; and F_{IS} (variance)—mean inbreeding coefficient.

Site ID	*K*	*N*	*Latitude*	*Longitude*	*Elevation (m)*	*Utilized Reads*	*Coverage*	*p*	H_E	π	F_{IS}
HC1	North	9	43.70	−114.19	2,144	716,735	17.7	0.885 (0.022)	0.159 (0.034)	0.172 (0.040)	0.022 (0.084)
SQ1	North	10	44.38	−114.50	1,961	750,167	18.1	0.879 (0.022)	0.168 (0.033)	0.182 (0.039)	0.041 (0.092)
SC1	North	8	44.80	−111.41	2,176	774,392	18.4	0.877 (0.022)	0.172 (0.033)	0.188 (0.039)	0.039 (0.089)
PA1	North	11	44.94	−115.94	1,666	709,976	17.2	0.882 (0.021)	0.166 (0.033)	0.178 (0.037)	0.048 (0.092)
PR1	North	14	45.17	−108.43	2,198	890,751	18.1	0.875 (0.022)	0.176 (0.032)	0.186 (0.036)	0.089 (0.106)
PC1	North	9	45.19	−114.15	1,979	790,623	19.2	0.888 (0.022)	0.155 (0.034)	0.168 (0.040)	0.054 (0.090)
MC1	North	2	45.48	−112.08	1,975	691,492	16.3	0.887 (0.028)	0.144 (0.039)	0.189 (0.068)	0.002 (0.058)
PR2	North	8	45.72	−113.34	1,908	770,131	17.9	0.889 (0.024)	0.150 (0.035)	0.167 (0.044)	0.024 (0.106)
DF8	South	9	38.02	−112.20	3,125	552,757	16.5	0.898 (0.021)	0.142 (0.031)	0.155 (0.037)	0.010 (0.073)
DC1	South	5	38.25	−106.67	2,825	595,369	16.5	0.897 (0.023)	0.139 (0.035)	0.162 (0.048)	0.050 (0.086)
FL7	South	10	38.48	−111.45	2,803	721,261	17.9	0.887 (0.021)	0.159 (0.032)	0.172 (0.038)	0.028 (0.087)
FL2	South	4	38.50	−112.01	2,832	691,678	16.7	0.898 (0.026)	0.130 (0.037)	0.162 (0.058)	0.025 (0.075)
SP1	South	5	38.50	−106.45	2,931	652,278	16.9	0.890 (0.024)	0.148 (0.036)	0.173 (0.049)	0.041 (0.086)
MF1	South	10	39.13	−111.60	2,389	807,951	17.6	0.888 (0.021)	0.156 (0.033)	0.169 (0.039)	0.072 (0.103)
MF5	South	4	39.58	−111.22	2,929	782,588	17.7	0.900 (0.027)	0.127 (0.037)	0.158 (0.058)	0.041 (0.082)
MF8	South	9	39.87	−111.27	2,738	775,462	17.3	0.904 (0.021)	0.132 (0.032)	0.143 (0.038)	0.038 (0.103)
UW4	South	5	40.54	−111.03	2,642	640,488	17	0.889 (0.023)	0.152 (0.035)	0.178 (0.048)	0.049 (0.088)
CR1	South	12	40.65	−109.50	2,465	767,854	18	0.885 (0.021)	0.161 (0.032)	0.172 (0.036)	0.049 (0.094)
WC1	South	3	41.23	−111.85	1,843	717,671	17.1	0.884 (0.028)	0.149 (0.039)	0.192 (0.065)	0.017 (0.068)
LC1	South	11	41.82	−111.47	2,312	731,874	17.2	0.881 (0.021)	0.167 (0.032)	0.179 (0.037)	0.050 (0.094)
BE1	South	9	42.32	−113.60	2,274	722,203	17.4	0.897 (0.021)	0.142 (0.032)	0.155 (0.038)	0.019 (0.077)
WH1	South	9	42.98	−109.76	2,604	772,952	17.9	0.876 (0.022)	0.174 (0.033)	0.189 (0.039)	0.057 (0.098)
CO1	South	9	43.50	−110.97	2,422	742,387	17.6	0.884 (0.021)	0.163 (0.032)	0.178 (0.039)	0.046 (0.089)
	North							0.872 (0.018)	0.187 (0.027)	0.189 (0.028)	
	South							0.869 (0.017)	0.193 (0.024)	0.194 (0.025)	

Genetic Analyses

We employed three methods to assess population structure in *E. speciosus*. First, Bayesian clustering was implemented in STRUCTURE v2.3.4 (Falush et al. 2003). STRUCTURE was run across K (population) values ranging from 1 to 9 without assigning population membership a priori. Twenty independent runs per K were conducted, each with 150,000 burn-in and 250,000 Markov chain Monte Carlo iterations, using an admixture model with correlated allele frequencies. STRUCTURE HARVESTER v0.6.94 (Earl & von Holdt 2012) and DISTRUCT v1.1 (Rosenberg 2004) were used to visualize results, and the most probable K was chosen based on the second order rate change of log probability scores (Evanno et al. 2005). Second, population genetic differentiation was visualized using Nei genetic distance (Nei 1978) in a unweighted pair group method with arithmetic mean (UPGMA) dendrogram using the *aboot* function in POPPRV2.8.6 (Kamvar et al. 2014) with 100× bootstrap resampling. Third, principal components analysis (PCA) was performed on the dataset using the *dudi.pca* function in ADGENET 2.17 (Jombart & Ahmed 2011). To further assess spatial and genetic patterns, correlations were calculated (r values) between collection site principal component (PC) 1 and 2 eigenvector values and geographic coordinates. Separate analyses were run for the range-wide collection and within populations determined by STRUCTURE analysis. Isolation by distance (IBD) was assessed with a Mantel test in ADE v1.7-15 (Dray & Dufour 2007) using geographic distance (km) and Nei's genetic distance. In addition, collection site pairwise population differentiation (F_{ST}), expected heterozygosity (H_{exp}), and nucleotide diversity (π) were calculated in STACKS. These analyses were conducted in R statistical software v4.05 (R Core Team 2021) and scripts can be found at the github repository (https://github.com/brichardsonfs/Erigeron_speciosus_seq).

To determine if the North and South *E. speciosus* populations (see Results section) are distributed across distinct subsets of environmental space within the study region, we characterized regional-scale climatic gradients using the 19 bioclimatically informative variables from WorldClim v2 (Fick & Hijmans 2017). We reduced the 19 variables to three PCs using "princomp" in R, as implemented through "rasterPCA" in rstoolbox (Leutner et al. 2019). PC values were extracted for three sets of points including: (1) collection sites; (2) 10,000 randomly selected background across the geographic extent of the study area; and (3) vetted *E. speciosus* herbarium records accessioned since 1995 and available on the SEINet data portal (https://swbiodiversity.org/seinet).

Flow Cytometry and Cytological Analyses

Two approaches were used to ascertain genome duplication. First, flow cytometry was conducted on a subset of fresh, field collected leaf tissue ($n = 44$) when available. Dried leaf tissue did not produce an adequate signal on the flow cytometer. In addition, since there was available seed from the CR1 population, a more intensive sampling of this population was conducted with seedlings grown from collection site bulked seed ($n = 9$; Table S1). Flow cytometry (Cyflow Space, Partec GmbH, Münster, Germany) was conducted using approximately 0.3 cm^2 of chopped fresh leaf tissue of the standard and the experimental sample. Leaf tissue was prepared using the protocol provided by the Cystain UV Precise staining buffer (Sysmex, Inc.) and a previously published protocol (Richardson et al. 2012) using an internal standard, *Acer negundo*, 2C value = 1.07 pg. Measurements were made with the UV laser at 375 nm. Genome size was calculated using the equation from Dolezel et al. (2007). Second, cytological analysis was conducted on the CR1 population from excised root tips (1 cm) of greenhouse-grown seedlings to determine the ploidy level of five of the nine samples whose genome sizes were estimated using flow cytometry (Table S1). Root tips were pretreated with 2.5 mM (w/v) 8-hydroxyquinoline, placed in the dark for 3 hours at room temperature, rinsed with ddH_2O for 2 minutes, then fixed in 3:1 fixative (95% ethanol:glacial acetic acid). Chromosome preparations were conducted within a week of the fixative treatment. Fixed root tips were rinsed with ddH_2O to remove fixative for 30 minutes, then hydrolyzed mildly with 0.2 N HCl at 60°C for 10 minutes and then 10 minutes at RT, rinsed with ddH2O twice 10 minutes each followed by 0.01M cold 4°C citrate buffer (pH 4.5). The enzyme mixture consisted of 2% cellulase RS (w/v), 1% macerozyme R10 (w/v) (Yakult Pharmaceutical Ind. Co., LTD, Japan), 2% pectolyase Y23 (w/v) (Kyowa Chemical Products, Co., Ltd, Japan), 30% cellulase (v/v), C2730, 30% pectinase (v/v), P2611, Sigma-Aldrich, St. Louis, MO, U.S.A., and 40% of 0.01M citrate buffer (pH 4.5). The enzyme digestion time varied from 12 to 25 minutes depending on the thickness of root-tips. Enzyme solution was carefully removed without disturbing the tips and samples were washed three times with fresh buffer. Chromosome spreads were prepared as described by Jewell and Islam-Faridi (1994). Chromosome spreads were stained with DAPI (4',6-diamidino-2-phenylindole; Vectashield, H-1200, Vector Laboratories, U.S.A.), and images were viewed under a 63× plan apochromatic oil-immersion objective with UV-filter (Chroma Technology, Bellows Falls, VT, U.S.A.) equipped with an epi-fluorescence microscope (AxioImager M2, Carl Zeiss, Inc., Germany). Digital images were captured with a Cool Cube 1 (metaSystem Group, Inc., Boston, MA, U.S.A.) charge-coupled device camera. Images were preprocessed with ISIS v5.1 (MetaSystems Group, Inc., and then processed further with Adobe Photoshop CC2021 (Adobe Systems, Inc., New York, NY, U.S.A.) to obtain chromosome counts.

Results

The filtered *Erigeron speciosus* sequence dataset included 185 samples from 23 collection sites and 175,419 SNPs at 50,034 loci. Collection site H_{exp} ranged from 0.13 to 0.18 (Table 1). Genetic structure was evident among some collection sites, with pairwise F_{ST} ranging from 0.05 to 0.24 (mean of 0.13; Table S2). STRUCTURE analyses supported grouping collection sites into two populations ($K = 2$), based on maximized second-order rate change (Fig. S3). The two populations were geographically distributed into a northern population

comprising central Idaho and Montana and a southern population comprising southeast Idaho, Utah, and Colorado (Fig. 1A), hereafter referred to as North and South populations. The STRUCTURE assignments were supported by genetic distance measures showing the divergence between these clades (Fig. 2). The exception to the divergent assignment probabilities between these two populations were eastern collection sites in Wyoming and Colorado (WH1, SP1, and DC1) that had more admixture. Except for WH1, these collection sites were rooted between the northern and southern clades (Fig. 2), and assignment probabilities indicated admixed individuals (Fig. 1B).

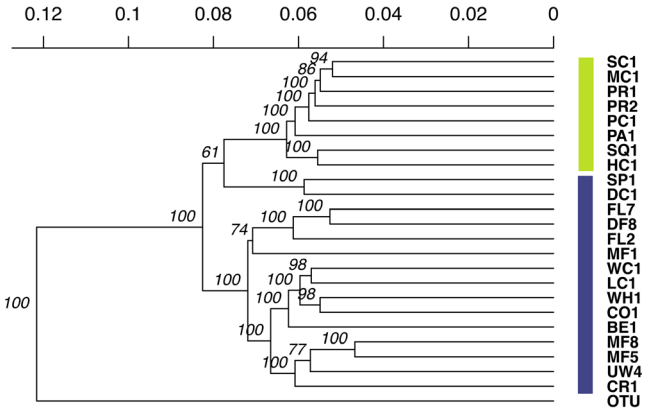

Figure 2. Unweighted pair group method with arithmetic mean (UPGMA) dendrogram of *Erigeron speciosus* collection site. The scale at the top shows Nei's genetic distance and colored bars indicate majority assignment to one of two populations, based on STRUCTURE analyses (Fig. 1). Values adjacent to the nodes are bootstrap support. The outgroup taxonomic unit (OTU) is *E. formosissimus*.

WH1 placement within the southern group is likely due to higher likelihood values for the South population, compared to DC1 and SP1 (Fig. 2). These relationships among sites and populations are also reflected in the PCA ordination (Fig. S4).

Overall, genetic differentiation and geographic distance was found to be correlated ($r^2 = 0.5$, $p < 0.0001$). This was also confirmed with a Mantel test ($p = 0.001$), supporting IBD. A closer examination of collection sites within North and South populations showed that IBD was significant in the South population ($p < 0.0001$) and nonsignificant in the North population ($p = 0.12$) (Fig. S5). The genetic variation that partitions the North and South populations is assorted among PC1 (Fig. S4), whereas association with latitude is assorted along PC2 ($r = -0.87$, $p = 0.0001$; Figs. 3A & S4). This North–South pattern in IBD was also reflected in collection site nucleotide diversity (π) in the South. A positive correlation was found between π and latitude among South collection sites ($r = 0.67$, $p = 0.006$; Fig. 3B), but no significant association was found between π and latitude for the North population. Average π values of collection sites above 40° latitude were 0.178 (SD = 0.01), whereas average values below 40° latitude were 0.162 (SD = 0.01) (Table 1).

Average genome size (2C value) from flow cytometry was 2.3 pg and ranged from 1.7 to 3.6 pg based on 53 field-collected samples. Two samples were approximately 1.5× the mean genome size, 3.1 and 3.6 pg, suggesting triploidy (Table S1). These field-collected specimens could not be confirmed with cytological analysis because of the availability of root and/or apical shoot samples. Cytological analysis of root tips from seven greenhouse-grown seedlings from one population (CR1) showed C values between 1.8 and 2.7 pg had $2n = 18$ chromosomes, and heterochromatic band variation was observed for some samples (Table S1; Fig. S6).

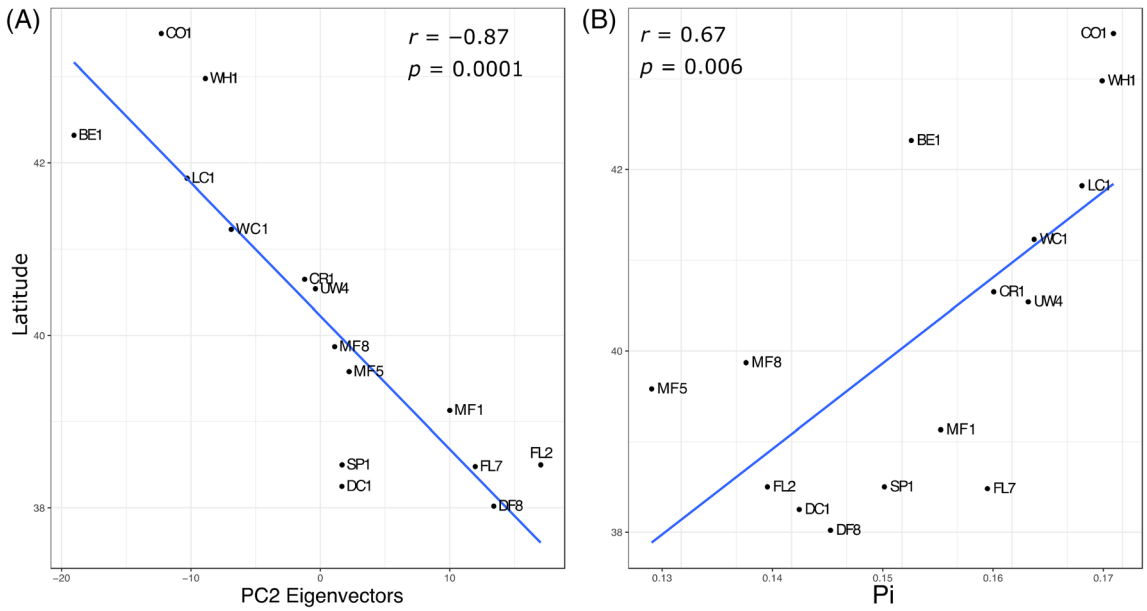

Figure 3. The linear relationship between (A) latitude and PC2 eigenvector values and (B) latitude and Pi, nucleotide diversity, of the 15 collection sites within the South population.

The majority of environmental variation across the study area was explained by the first three PC axes (PC1–52.4%, PC2–14.4%, and PC3–9.6%). PC1 values were strongly negatively correlated with precipitation variables (i.e. higher PC1 values represent drier climates), including annual precipitation (BIO12, $r = -0.88$), precipitation of the wettest month (BIO13, −0.74), precipitation of the driest month (BIO14, −0.77), precipitation of the wettest quarter (BIO16, −0.76), precipitation of the driest quarter (BIO17, −0.83), and precipitation of the coldest quarter (BIO19, −0.83). Alternatively, PC2 and PC3 were correlated with temperature variables, including positive correlations with minimum temperature of the coldest month for both (BIO16, 0.74 and 0.7, respectively), and negative correlations with mean diurnal range (BIO2, −0.82) and temperature annual range (BIO7, −0.72) for PC3 only. *Erigeron speciosus* is distributed across a wide breadth of environmental space within the study region and collection sites are representative of that space (Fig. S1).

Discussion

Population genetic analyses of *Erigeron speciosus* support two populations: a North population that occurs in central Idaho and western Montana and a South population that occurs in Utah and southeast Idaho. Collection sites in Wyoming and Colorado (e.g. WH1 and SP1) suggest that admixture between these populations is more common farther east in the species range. Genetic differentiation among collection sites ranged from low to high (mean $F_{ST} = 0.13$, range 0.05–0.24) with higher collection site pairwise F_{ST} values found between the North and South populations. These F_{ST} values may reflect this species' life history traits, specifically insect-vectored pollination, and herbaceous growth, which has been shown to increase population genetic structure (Duminil et al. 2007; Gamba & Muchhala 2020). Another factor that may affect population genetic structure is discontinuous distribution. Since *E. speciosus* is restricted to montane environments, arid lower elevation sagebrush ecosystems could act as a barrier to gene flow. The distinct North and South populations coincide with the Snake River Plain, a large (approximately 130 km), lower elevation gap in this species range. During the late Pleistocene, Lake Bonneville, which covered much of northern Utah and eastern Nevada, may have also acted as a barrier to gene flow. Corridors for introgression may be limited to montane habitats further east in Wyoming and Colorado. However, historically these routes (e.g. the Greater Yellowstone region in northwestern Wyoming) likely limited geneflow during the late Pleistocene, since this region was extensively glaciated (Licciardi & Pierce 2018). Thus, admixed populations occurring in western Wyoming (i.e. populations CO1 and WH1) could be a more recent occurrence during Holocene warming as North and South populations expanded into western Wyoming. The locality of the North–South introgression zone near Yellowstone, is similar to those reported between two aspen populations (*Populus tremuloides*; Callahan et al. 2013). Moreover, the area from Yellowstone south into western Wyoming and Colorado is shown to be a contact/hybrid zone for several plant species (Swenson & Howard 2005).

Spatial patterns of variation within *E. speciosus* illustrate the importance of using neutral genomic loci to guide the development of management programs for species used in restoration. Without such knowledge, managers across the Intermountain West typically rely upon Omernik Level III ecoregions (Omernik & Griffith 2014) to guide/constrain regional seed transfer and NPMs development strategies (Bower et al. 2014; Shryock et al. 2017). If used for this species, the Middle Rockies ecoregion, which incorporates montane regions both north and south of the Snake River Plain, would facilitate the mixture of genotypes from the North and South populations. In effect, such a management decision would degrade natural patterns of genetic diversity formed by long-term historical and evolutionary processes, which are recognized as an important component of biodiversity (CBD 1992). Complex patterns of local adaptation are expected in widespread species with large population sizes that display regional population structure across broad environmental gradients (Richardson et al. 2012; Shryock et al. 2015). Unknowingly mixing North and South seed sources into a seed increase program or seed transfer across these boundaries could lead to deleterious processes such as outbreeding depression, and potentially affect long-term restoration goals (Hufford et al. 2012).

IBD was significant for collection sites in the South population, but not in the North. Two factors that could be contributing to this difference are physiography and barriers to postglacial colonization. In the South population, especially in Utah, physiography of north–south trending mountain ranges likely provides corridors for propagule movement. However, east–west movement is likely impeded by basins too arid to support this species and long-distance pollination, as few collection locations were found associated with low levels of precipitation. In the North population, the basin and range physiography are not as pronounced, and cooler climates provide *E. speciosus* with lower elevation habitat and probably a more contiguous range, resulting in increased gene flow and less isolation. Holocene warming likely resulted in expansion of habitat in the North, while the South may have experienced a range contraction and more isolated populations at higher elevations. This hypothesis appears to be supported by the association between genetic distance and latitude, population nucleotide diversity, π, has a positive association with latitude in the South population. This diversity-latitude pattern could be caused by (1) increasing isolation and smaller populations in lower latitudes leading to decreasing nucleotide diversity due to evolutionary processes like genetic drift, (2) increasing nucleotide diversity at the northern edge of the South population due to introgression with the North population (e.g. admixture found in sites CO1 and WH1), or (3) a combination of both. Many temperate plant species show patterns of increasing genetic diversity with latitude through the mid-latitudes of the species range, followed by a decline in diversity at higher latitudes near the leading edge of range expansion (Petit et al. 2003). However, this pattern is not supported for many tree species in Eastern North America. Here, higher genetic diversity can extend beyond the mid-latitudes and into high latitude range limits (Lumibao et al. 2017). In *E. speciosus*, we do not observe declining genetic

diversity with increasing latitude at the northern range, supporting the patterns found in Eastern North America trees. It is important to note, however, that northern range limit of this species extends at least 6°N in latitude (approximately 500 km) beyond our sampling. Declining genetic diversity in the last 500 km of the species northern range could be possible.

Based on flow cytometry from a subset of field collected samples and greenhouse-grown seedlings, *E. speciosus* appears to be predominantly diploid with an average genome size (2C value) of 2.3 pg. The 2C values are similar to previous reports of *Erigeron* species ranging between 1.8 pg in *E. philadelphicus* to 3.1 pg in *E. alpinus* (Bai et al. 2012; Garcia et al. 2013). Triploidy is suspected in two of 53 flow cytometry samples. These samples are 3.1 and 3.6 pg, which is roughly 1.5× the average genome size. Diploid ($2n = 18$) intraspecific genome size, confirmed by cytological analysis, varied in *E. speciosus* from 1.8 to 2.7 pg. This is a 1.6× difference genome size between individuals from the same collection site (CR1). There could be different factors affecting the genomes size in this species. Genome size variation can be caused by transposable elements (Hawkins et al. 2006), satellite repetitive elements (Bitonti et al. 1999; Thakur et al. 2021), telomeric repeats (Fajkus et al. 2005), or supernumerary chromosomes (Jones & Houben 2003). Variable heterochromatic bands were observed in this species, but the specific causes of this variation mentioned above remain unknown. Some instrumental error could also be attributed to flow cytometry. Plant secondary metabolites, particularly phenolics, can affect fluorophore binding and fluorescence in propidium iodide staining (Price et al. 2000; Greilhuber 2005), leading to a degraded signal. However, less is known on how plant compounds interact with DAPI stain used in this study.

Population genomics is increasingly being used to help guide restoration seed-sourcing decisions (Durka et al. 2017). As demonstrated here, a population genomics approach paired with flow cytometry is especially useful for native forb species that routinely display taxonomical complexity and harbor polyploidy (Kramer et al. 2018). Compared to traditional molecular genetic approaches, genome-wide loci allow for better resolution of demographic and biogeographic processes (Ellegren 2014), which can lead to more informed seed-sourcing decisions (Breed et al. 2019; Massatti et al. 2020). In the western United States, population genomic information is being developed for an expanding set of restoration species (e.g. Shryock et al. 2017; Massatti et al. 2018). We advocate that this approach is complementary and a precursor to common garden studies. Information from studies like this one can be useful in designing and interpreting trait-based experiments that address adaptive variation. Studies that pair neutral genetic structure, cytogenetics, and trait-based variation provide a set of comprehensive information for restoration species. Such progress has been done in some restoration species, e.g. *Pseudoroegneria spicata* (St. Clair et al. 2013; Massatti et al. 2018) and *Artemisia tridentata* (Richardson et al. 2012; Richardson & Chaney 2018). These datasets can lead to genetic-environment analyses that synthesize genomic, trait, and environmental data, providing a comprehensive assessment of adaptive genetic variation and how the environment shapes these patterns (Temunović et al. 2020; Faske et al. 2021). As the set of species with genetic information grows, restoration programs will benefit by having more tools to meet restoration targets, such as plant communities with genetically appropriate restoration materials that ultimately support pollinator health relationships within ecosystems (Bucharova et al. 2021).

The population genetic structure could be compatible with either of the two proposed taxonomic treatments: two regional varieties *macranthus* and *speciosus* (Cronquist et al. 1994) or introgression of *E. speciosus* and *E. subtrinervis* (Nesom 2006). In the former taxonomic treatment, the North and South populations appear to coincide with the general geographic distribution of the two described varieties: *speciosus*, found in the Northwest (Washington, Oregon, northern Idaho, and western Montana), and *macranthus*, found further south and east (Cronquist et al. 1994). According to the latter taxonomic treatment, the North and South populations would reflect varied hybridization with *E. subtrinervis*. However, under the hypothesis of hybridization with *E. subtrinervis*, we would expect a third genetic group to emerge in the STRUCTURE analyses. Further study using herbarium samples of both species is needed to resolve the taxonomic treatments.

Both genetic structure and diversity within *E. speciosus* have a strong spatial component that can inform the development of a restoration program. Given the clear segregation of two populations (North and South), separate seed increase and seed transfer programs for each region will best support observed patterns of genetic diversity (e.g. evolutionary units potentially best represented by varieties; Coates et al. 2018). Management programs armed with this knowledge may avoid entangling divergent evolutionary processes that have the capacity to influence restoration success (e.g. outbreeding depression; Frankham et al. 2011). Population structure in many organisms of temperate regions have been highly influenced by the last glacial period and therefore long time frames (Richardson & Meyer 2012; Massatti & Knowles 2020); such long-term processes may influence reproductive compatibility of individuals in early generations when plant performance is critical for promoting restoration success (Hufford et al. 2012). We recommend careful consideration when assembling sources to develop NPMs (Bucharova et al. 2019) or transferring seeds within North and South populations, due to the distribution of populations across broad environmental space (McKay et al. 2005). Moreover, different strategies for developing seed increase lots could be employed among populations. Given the stronger IBD found within the South, stricter guidelines for choosing pooled seed collections, whereas the North population could have broader collection requirement. Prior to seed increase, we also recommend screening seed sources for polyploidy; although polyploidy appears to be infrequent in *E. speciosus*, introduction of mixed cytotypes into a seed increase program could increase incompatibility, lowering seed yield or introduce infertility in the offspring (Kramer et al. 2018). Since polyploidy is infrequent in this species, triploids could be excluded from the seed increase program. However, for species that have a higher frequent polyploidy or exhibit niche-specific patterns, different

strategies and potentially separate seed increase programs need to be considered to fulfill restoration needs. Ultimately, implementation of a genetically informed management program will help improve the success of restoration outcomes and species interactions within natural communities, and maintain patterns of evolutionary history across the species' range.

Acknowledgments

The authors thank R. Lehman, K. Memmott, D. Netz, J. Proctor, A. Tonnesen, and their field crews for assistance in field collections. Thanks also to T. Tobiasson, J. Irwin, and L. Grossfurthner for taxonomic and laboratory assistance; M. Mahalovich for supplying seeds; and E. Milano and two anonymous reviewers for revising the manuscript. Funding was provided by the USDA Forest Service, Region 4. This research was also supported in part by the Great Basin Native Plant Project, USDA Forest Service, Rocky Mountain Research Station, and USDOI Bureau of Land Management. The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

LITERATURE CITED

Aronesty E (2011) ea-utils: command-line tools for processing biological sequencing data. https://github.com/ExpressionAnalysis/ea-utils

Bai C, Alverson WS, Follansbee A, Waller DM (2012) New reports of nuclear DNA content for 407 U.S. plant species. Annals of Botany 110:1623–1629. https://doi.org/10.1093/aob/mcs222

Bitonti MB, Cozza R, Chiappetta A, Contento A, Minelli S, Ceccarelli M, Gelati MT, Maggini F, Baldoni L, Cionini PG (1999) Amount and organization of the heterochromatin in *Olea europaea* and related species. Heredity 83:188–195. https://doi.org/10.1046/j.1365-2540.1999.00564.x

Bower AD, St. Clair JB, Erickson V (2014) Generalized provisional seed zones for native plants. Ecological Applications 24:913–919. https://doi.org/10.1890/13-0285.1

Breed MF, Hamilton PA, Bluth C, Byrne M, Gaget V, Gillie NJC, et al. (2019) The potential of genomics for restoring ecosystems and biodiversity. Nature Reviews 20:615–628. https://doi.org/10.1038/s41576-019-0152-0

Bucharova A, Bossdorf O, Hölzel N, Kollmann J, Prasse R, Durka W (2019) Mix and match: regional admixture provenancing strikes a balance among different seed-sourcing strategies for ecological restoration. Conservation Genetics 20:7–17. https://doi.org/10.1007/s10592-018-1067-6

Bucharova A, Lampei C, Conrady M, May E, Matheja J, Meyer M, Ott D (2021) Plant provenance affects pollinator network: implications for ecological restoration. Journal of Applied Ecology 1–11. https://doi.org/10.1111/1365-2664.13866

Burkle L, Irwin R (2009) The importance of interannual variation and bottom–up nitrogen enrichment for plant–pollinator networks. Oikos 118:1816–1829. https://doi.org/10.1111/j.1600-0706.2009.17740.x

Burkle LA, Runyon JB (2019) Floral volatiles structure plant–pollinator interactions in a diverse community across the growing season. Functional Ecology 33:2116–2129. https://doi.org/10.1111/1365-2435.13424

Callahan CM, Rowe CA, Ryel RJ, Shaw JD, Madritch MD, Mock KE (2013) Continental-scale assessment of genetic diversity and population structure in quaking aspen (*Populus tremuloides*). Journal of Biogeography 40:1780–1791. https://doi.org/10.1111/jbi.12115

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. Molecular Ecology 22:3124–3140. https://doi.org/10.1111/mec.12354

CBD (1992) Convention on biological diversity. United Nations. Earthscan Publications Ltd, London and Sterling, VA. https://www.cbd.int/doc/legal/cbd-en.pdf

Coates DJ, Byrne M, Moritz C (2018) Genetic diversity and conservation units: dealing with the species-population continuum in the age of genomics. Frontiers in Ecology and Evolution 6:165. https://doi.org/10.3389/fevo.2018.00165

Cronquist A, Holmgren AH, Holmgren NH, James R (1994) Intermountain flora, Vol. 5: Asterales. Intermountain Flora: Vascular Plants of the Intermountain West, USA. New York Botanical Garden, Bronx, New York

Dolezel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. Nature Protocols 2:2233–2244. https://doi.org/10.1038/nprot.2007.310

Dray S, Dufour A (2007) The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software 22:1–20. https://doi.org/10.18637/jss.v022.i04

Duminil J, Fineschi S, Hampe A, Jordano P, Salvini D, Vendramin GG, Petit RJ (2007) Can population genetic structure be predicted from life-history traits? The American Naturalist 169:662–672. https://doi.org/10.1086/513490

Dumroese RK, Luna T, Richardson BA, Kilkenny FF, Runyon JB (2015) Conserving and restoring habitat for Greater Sage-Grouse and other sagebrush-obligate wildlife: the crucial link of forbs and sagebrush diversity. Native Plants Journal 16:276–299. https://doi.org/10.3368/npj.16.3.276

Dumroese RK, Luna T, Pinto JR, Landis TD (2016) Forbs: Foundation for restoration of monarch butterflies, other pollinators, and Greater Sage-Grouse in the western United States. Natural Areas Journal 36:499–511. https://doi.org/10.3375/043.036.0415

Durka W, Michalski SG, Berendzen KW, Bossdorf O, Bucharova A, Hermann J-M, Hölzel N, Kollmann J (2017) Genetic differentiation within multiple common grassland plants supports seed transfer zones for ecological restoration. Journal of Applied Ecology 54:116–126. https://doi.org/10.1111/1365-2664.12636

Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359–361. https://doi.org/10.1007/s12686-011-9548-7

Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. Trends in Ecology & Evolution 29:51–63. https://doi.org/10.1016/j.tree.2013.09.008

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fajkus J, Sýkorová E, Leitch AR (2005) Telomeres in evolution and evolution of telomeres. Chromosome Research 13:469–479. https://doi.org/10.1007/s10577-005-0997-2

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587. https://doi.org/10.1093/genetics/164.4.1567

Faske TM, Agneray AC, Jahner JP, Sheta LM, Leger EA, Parchman TL (2021) Genomic and common garden approaches yield complementary results for quantifying environmental drivers of local adaptation in rubber rabbitbrush, a foundational Great Basin shrub. Evolutionary Applications 14:2881–2900. https://doi.org/10.1111/eva.13323

Fenster CB, Galloway LF (2000) Inbreeding and outbreeding depression in natural populations of *Chamaecrista fasciculata* (Fabaceae). Conservation Biology 14:1406–1412. https://doi.org/10.1046/j.1523-1739.2000.99234.x

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37:4302–4315. https://doi.org/10.1002/joc.5086

Frankham R, Ballou JD, Eldridge MD, Lacy RC, Ralls K, Dudash MR, Fenster CB (2011) Predicting the probability of outbreeding depression. Conservation Biology 25:465–475. https://doi.org/10.1111/j.1523-1739.2011.01662.x

Funk JL (2021) Revising the trait-based filtering framework to include interacting filters: lessons from grassland restoration. Journal of Ecology 109:3466–3472. https://doi.org/10.1111/1365-2745.13763

Gamba D, Muchhala N (2020) Global patterns of population genetic differentiation in seed plants. Molecular Ecology 29:3413–3428. https://doi.org/10.1111/mec.15575

Garcia S, Hidalgo O, Jakovljević I, Siljak-Yakovlev S, Vigo J, Garnatje T, Vallès J (2013) New data on genome size in 128 Asteraceae species and subspecies, with first assessments for 40 genera, 3 tribes and 2 subfamilies. Plant Biosystems 147:1219–1227. https://doi.org/10.1080/11263504.2013.863811

Greilhuber J (2005) Intraspecific variation in genome size in angiosperms: identifying its existence. Annals of Botany 95:91–98. https://doi.org/10.1093/aob/mci004

Gucker CL, Shaw NL (2018) Aspen fleabane (*Erigeron speciosus*). In: Gucker CL, Shaw NL (eds) Western forbs: biology, ecology, and use in restoration. Great Basin Fire Science Exchange, Reno, Nevada. http://greatbasinfirescience.org/western-forbs-restoration

Hawkins JS, Kim H, Nason JD, Wing RA, Wendel JF (2006) Differential lineage-specific amplification of transposable elements is responsible for genome size variation in *Gossypium*. Genome Research 16:1252–1261. https://doi.org/10.1101/gr.5282906

Hufford KM, Krauss SL, Veneklaas EJ (2012) Inbreeding and outbreeding depression in *Stylidium hispidum*: implications for mixing seed sources for ecological restoration. Ecology and Evolution 2:2262–2273. https://doi.org/10.1002/ece3.302

Inouye DW (2008) Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. Ecology 89:353–362. https://doi.org/10.1890/06-2128.1

Isbell F, Calcagno V, Hector A, Connolly J, Harpole WS, Reich PB, et al. (2011) High plant diversity is needed to maintain ecosystem services. Nature 477:199–202. https://doi.org/10.1038/nature10282

Jewell DC, Islam-Faridi N (1994) A technique for somatic chromosome preparation and C-banding of maize. Pages 484–493. In: The maize handbook. Vol 1994. Springer, New York

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jones N, Houben A (2003) B chromosomes in plants: escapees from the A chromosome genome? Trends in Plant Science 8:417–423. https://doi.org/10.1016/S1360-1385(03)00187-0

Jones AG, Young DA (1983) Generic concepts of Aster (Asteraceae): a comparison of cladistic, phenetic, and cytological approaches. Systematic Botany 8:71. https://doi.org/10.2307/2418564

Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281. https://doi.org/10.7717/peerj.281

Kramer AT, Wood TE, Frischie S, Havens K (2018) Considering ploidy when producing and using mixed-source native plant materials for restoration. Restoration Ecology 26:13–19. https://doi.org/10.1111/rec.12636

Laport RG, Minckley RL, Ramsey J (2012) Phylogeny and cytogeography of the North American Creosote Bush (*Larrea tridentata*, Zygophyllaceae). Systematic Botany 37:153–164. https://doi.org/10.1600/036364412X616738

Leutner B, Horning N, Schwalb-Willmann J, Hijmans RJ (2019) RStoolbox: tools for remote sensing data analysis. R package version 0.2.6

Licciardi JM, Pierce KL (2018) History and dynamics of the Greater Yellowstone Glacial System during the last two glaciations. Quaternary Science Reviews 200:1–33. https://doi.org/10.1016/j.quascirev.2018.08.027

Lumibao CY, Hoban SM, McLachlan J (2017) Ice ages leave genetic diversity "hotspots" in Europe but not in eastern North America. Ecology Letters 20:1459–1468. https://doi.org/10.1111/ele.12853

Massatti R, Knowles LL (2020) The historical context of contemporary climatic adaptation: a case study in the climatically dynamic and environmentally complex southwestern United States. Ecography 43:735–746. https://doi.org/10.1111/ecog.04840

Massatti R, Prendeville HR, Larson S, Richardson BA, Waldron B, Kilkenny FF (2018) Population history provides foundational knowledge for utilizing and developing native plant restoration materials. Evolutionary Applications 11:2025–2039. https://doi.org/10.1111/eva.12704

Massatti R, Shriver RK, Winkler DE, Richardson BA, Bradford JB (2020) Assessment of population genetics and climatic variability can refine climate-informed seed transfer guidelines. Restoration Ecology 28:485–493. https://doi.org/10.1111/rec.13142

McArthur ED, Sanderson SC (1999) Cytogeography and chromosome evolution of subgenus Tridentatae of *Artemisia* (Asteraceae). American Journal of Botany 86:1754–1775. https://doi.org/10.2307/2656673

McKay JK, Christian CE, Harrison S, Rice KJ (2005) "How local is local?"—a review of practical and conceptual issues in the genetics of restoration. Restoration Ecology 13:432–440. https://doi.org/10.1111/j.1526-100X.2005.00058.x

Mills A, Christophersen T, Wilkie ML, Mansur E (2020) The United Nations Decade on ecosystem restoration: catalysing a global movement. Unasylva 252:119–126

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590. https://doi.org/10.1093/genetics/89.3.583

Nesom GL (2006) *Erigeron*. Pages 256–348. In: Flora of North America Editorial Committee (ed) Flora of North America North of Mexico. Volume 20 Magnoliophyta: Asteridae, part 8: Asteraceae, part 3 Asterales, part 3 Aster order. Oxford University Press, New York

Noyes RD (2000) Biogeographical and evolutionary insights on *Erigeron* and allies (Asteraceae) from ITS sequence data. Plant Systematics and Evolution 220:93–114. https://doi.org/10.1007/BF00985373

Noyes RD (2006) Intraspecific nuclear ribosomal DNA divergence and reticulation in sexual diploid *Erigeron strigosus* (Asteraceae). American Journal of Botany 93:470–479. https://doi.org/10.3732/ajb.93.3.470

Noyes RD, Rieseberg LH (2000) Two independent loci control agamospermy (apomixis) in the triploid flowering plant *Erigeron annuus*. Genetics 155:379–390. https://doi.org/10.1093/genetics/155.1.379

Omernik JM, Griffith GE (2014) Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental Management 54:1249–1266. https://doi.org/10.1007/s00267-014-0364-1

Paris JR, Stevens JR, Catchen JM (2017) Lost in parameter space: a road map for stacks. Methods in Ecology and Evolution 8:1360–1373. https://doi.org/10.1111/2041-210X.12775

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7:e37135. https://doi.org/10.1371/journal.pone.0037135

Petit RJ, Aguinagalde I, de Beaulieu J-L, Bittkau C, Brewer S, Cheddadi R, et al. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300:1563–1565. https://doi.org/10.1126/science.1083264

Price HJ, Hodnett G, Johnston JS (2000) Sunflower (*Helianthus annuus*) leaves contain compounds that reduce nuclear propidium iodide fluorescence. Annals of Botany 86:929–934. https://doi.org/10.1006/anbo.2000.1255

R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ramsey J (2011) Polyploidy and ecological adaptation in wild yarrow. Proceedings of the National Academy of Sciences of the United States of America 108:7096–7101. https://doi.org/10.1073/pnas.1016631108

Richardson BA, Chaney L (2018) Climate-based seed transfer of a widespread shrub: population shifts, restoration strategies, and the trailing edge. Ecological Applications 28:2165–2174. https://doi.org/10.1002/eap.1804

Richardson BA, Meyer SE (2012) Paleoclimate effects and geographic barriers shape regional population genetic structure of blackbrush (*Coleogyne*

ramosissima: Rosaceae). Botany 90:293–299. https://doi.org/10.1139/b2012-002

Richardson BA, Page JT, Bajgain P, Sanderson SC, Udall JA (2012) Deep sequencing of amplicons reveals widespread intraspecific hybridization and multiple origins of polyploidy in big sagebrush (*Artemisia tridentata*; Asteraceae). American Journal of Botany 99:1962–1975. https://doi.org/10.3732/ajb.1200373

Richardson BA, Germino MJ, Warwell MV, Buerki S (2021) The role of genome duplication in big sagebrush growth and fecundity. American Journal of Botany 108:1405–1416. https://doi.org/10.1002/AJB2.1714

Rochette NC, Catchen JM (2017) Deriving genotypes from RAD-seq short-read data using Stacks. Nature Publishing Group 12:2640–2659. https://doi.org/10.1038/nprot.2017.123

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. Molecular Ecology Notes 4:137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Ruiz DC, Machío IV, Nieto AH, Feliner GN (2021) Hybridization and cryptic speciation in the Iberian endemic plant genus *Phalacrocarpum* (Asteraceae-Anthemideae). Molecular Phylogenetics and Evolution 156:107024. https://doi.org/10.1016/j.ympev.2020.107024

Shryock DF, Havrilla CA, DeFalco LA, Esque TC, Custer NA, Wood TE (2015) Landscape genomics of Sphaeralcea ambigua in the Mojave Desert: a multivariate, spatially-explicit approach to guide ecological restoration. Conservation Genetics 16:1303–1317. https://doi.org/10.1007/s10592-015-0741-1

Shryock DF, Havrilla CA, DeFalco LA, Esque TC, Custer NA, Wood TE (2017) Landscape genetic approaches to guide native plant restoration in the Mojave Desert. Ecological Applications 27:429–445. https://doi.org/10.1002/eap.1447

St. Clair JB, Kilkenny FF, Johnson RC, Shaw NL, Weaver G (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. https://doi.org/10.1111/eva.12077

Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. The American Naturalist 166:581–591. https://doi.org/10.1086/491688

Temunović M, Garnier-Géré P, Morić M, Franjić J, Ivanković M, Bogdan S, Hampe A (2020) Candidate gene single nucleotide polymorphism variation in floodplain populations of pedunculate oak (*Quercus robur* L.) near the species' southern range margin: weak differentiation yet distinct associations with water availability. Molecular Ecology 29:2359–2378. https://doi.org/10.1111/mec.15492

Thakur J, Packiaraj J, Henikoff S (2021) Sequence, chromatin and evolution of satellite DNA. International Journal of Molecular Sciences 22:4309. https://doi.org/10.3390/ijms22094309

Widener L, Fant JB (2018) Genetic differentiation and diversity of two sympatric subspecies of *Castilleja affinis*: a comparison between the endangered serpentine endemic (spp. *neglecta*) and its widespread congener (ssp. *affinis*). Conservation Genetics 19:365–381. https://doi.org/10.1007/s10592-017-1009-8

Winkler DE, Massatti R (2020) Unexpected hybridization reveals the utility of genetics in native plant restoration. Restoration Ecology 28:1047–1052. https://doi.org/10.1111/rec.13189

Zaiats A, Lazarus B, Germino MJ, Serpe MD, Richardson BA, Buerki S, Caughlin TT (2020) Intraspecific variation in surface water uptake in a perennial desert shrub. Functional Ecology 34:1170–1179. https://doi.org/10.1111/1365-2435.13546

Supporting Information

The following information may be found in the online version of this article:

Table S1. *Erigeron speciosus* 2C genome size in picograms (pg).
Table S2. The upper diagonal shows collection site pairwise F_{ST} values for each collection site.
Figure S1. (A) Mapped locations of *Erigeron speciosus* collection sites (squares).
Figure S2. (A) The number of loci shared across 80% of individuals (r80), and (B) the distribution of the number of SNPs per locus across various *M* and *n* filtering parameter values in *STACKS*.
Figure S3. Second-order rate change of log probability (Delta K) for population models (K) 2 to 8.
Figure S4. Principle coordinates plot of *E. speciosus* SNP variation.
Figure S5. Relationship between genetic distance and geographic distance (km).
Figure S6. Different stages of somatic chromosome spreads of *E. speciosus*.

Coordinating Editor: Paul Nevill

Received: 21 December, 2021; First decision: 28 February, 2022; Revised: 23 May, 2022; Accepted: 24 May, 2022