

# Phenotypic and Genetic Characterization of Western Prairie Clover Collections From the Western United States

Kishor Bhattarai,<sup>1</sup> B. Shaun Bushman,<sup>2</sup> Douglas A. Johnson,<sup>3</sup> and John G. Carman<sup>4</sup>

Authors are <sup>1</sup>Graduate Student and <sup>4</sup>Professor, Department of Plants, Soil, and Climate, Utah State University, Logan, UT 84322-4820, USA; and <sup>2</sup>Research Geneticist and <sup>3</sup>Plant Physiologist, US Department of Agriculture–Agricultural Research Service (USDA-ARS) Forage and Range Research Lab, Utah State University, Logan, UT 84322-6300, USA.

## Abstract

Few North American legumes are available for rangeland revegetation in the semiarid western United States. Western prairie clover (*Dalea ornata* [Douglas ex Hook.] Eaton & J. Wright) is a perennial legume with desirable forage characteristics and is distributed in the northern Great Basin, Snake River Basin, and southern Columbia Plateau. Understanding the genetic and ecotypic variability of this species is a prerequisite for developing populations suitable for revegetation purposes. To address this need, we established two common-garden plots of western prairie clover from 22 sites in Idaho, Oregon, and Washington. Significant variation was detected among the collections for all traits measured. Among the measured traits, flowering date was correlated with collection-site temperature and elevation. Population structure estimates from 474 amplified-fragment length polymorphism markers resulted in two distinct, genetically differentiated groups and a third admixed group, and flowering date played a significant role in discriminating those genetic-based groupings of collections. Positive correlations were observed between phenotypic and genetic distance matrices ( $r = 0.33$ ,  $P = 0.005$ ), phenotypic and geographic distance matrices ( $r = 0.35$ ,  $P = 0.002$ ), and genetic and geographic distance matrices ( $r = 0.31$ ,  $P = 0.009$ ). Based on these results, we recommend that two germplasm sources of western prairie clover be developed for use across the collection area, one from the Deschutes River region and the other encompassing Idaho, Washington, and eastern Oregon collection sites.

## Resumen

Existen pocas leguminosas disponibles en los Estados Unidos de Norteamérica para la revegetación en las zonas semiáridas del oeste Estados Unidos. El trébol de la pradera occidental (*Dalea ornata* [Douglas ex Hook.] Eaton & J. Wright) es una leguminosa perenne con características de un forraje deseable y se distribuye en el norte de la cuenca del Snake River, y en la meseta Columbia en el noroeste de los Estados Unidos. La comprensión de la variabilidad genética y ecotípica de esta especie es un requisito para el desarrollo de las poblaciones adecuadas para fines de revegetación. Para enfrentar esta necesidad, establecimos dos parcelas de jardín común de trébol de pradera occidental provenientes de 22 sitios de Idaho, Oregón, y Washington. Se detectó una variación significativa entre las colecciones para todos los rasgos medidos. Entre los rasgos medidos, la fecha de floración se correlacionó con la temperatura del sitio de la colección y la elevación. La estructura de la población se estimó de 474 marcadores de polimorfismo de la longitud del fragmento amplificado resultaron en dos grupos distintos genéticamente diferenciados y un tercer grupo mezclado. La fecha de floración desempeñó un papel importante en separar esas agrupaciones basándose en las colecciones genéticas. Se observaron correlaciones positivas entre los fenotipos y las matrices de distancia genética ( $r = 0.33$ ,  $P = 0.005$ ), fenotipo y las matrices de distancias geográficas ( $r = 0.35$ ,  $P = 0.002$ ), y genética y las matrices de geográficas ( $r = 0.31$ ,  $P = 0.009$ ). A partir de estos resultados recomendamos que dos fuentes de germoplasma del trébol de pradera occidental se desarrollen para su uso en toda la zona de colección, uno de la región del río Deschutes y otros que abarca de Idaho, Washington y sitios orientales colectados en Oregón.

**Key Words:** amplified fragment length polymorphism (AFLP), conservation, *Dalea*, Great Basin, legume, population structure, revegetation

## INTRODUCTION

The western United States has some of the highest numbers of endemic and imperiled species in the United States (Stein et al. 2000; Nachlinger et al. 2001). Use of a diversity of species in

rangeland revegetation programs can help minimize weed invasion, because they occupy available niches that could otherwise be colonized by invasive weeds (Pfisterer et al. 2004; Walker and Shaw 2005; Brown et al. 2008). Legumes are of particular interest because they provide biologically fixed nitrogen to associated species, increase plant production, enhance forage quality, and provide food sources for herbivores and pollinators (Cherney and Allen 1995; Madison and Robel 2001; Aydin and Uzun 2005; Walker and Shaw 2005). In addition, legumes can maintain and restore natural successional trajectories (Richards et al. 1998).

High seed cost and limited seed availability of western North American legumes have restricted their use in revegetation/restoration programs (Walker and Shaw 2005). Seed of

This research was partly funded by Great Basin Native Plant Selection and Increase Project through the Department of the Interior/Bureau of Land Management Great Basin Restoration Initiative and the USDA Forest Service Rocky Mountain Research Station.

Mention of a proprietary product does not constitute a guarantee or warranty of the product by USDA, Utah State University, or the authors and does not imply its approval to the exclusion of other products.

Correspondence: B. Shaun Bushman, USDA-ARS Forage and Range Research Laboratory, Utah State University, Logan, UT 84322-6300, USA. Email: Shaun.Bushman@ars.usda.gov

Manuscript received 13 January 2010; manuscript accepted 17 August 2010.

wildland-collected legumes is expensive, often several-fold higher than agronomically grown seed. Wildland-collected seed is of variable quality and might not be available when needed. Also, these legumes can be over-collected from specific wildland sites, thus disrupting the local population. Although seeds of western North American legumes are considerably more costly than grass seed, public land managers are interested in utilizing these legumes for reseeding and restoration because of their unique role in the ecosystem. Commercial seed production of these legumes could make seed more readily available at a lower cost for rangeland revegetation/restoration programs in the semiarid western United States.

*Dalea* is a widespread genus of the legume family (Fabaceae) comprising 62 species of prairie clovers in North America (US Department of Agriculture, Natural Resources Conservation Service [USDA, NRCS] 2009). *Dalea* is separated from other genera in the Amorphae tribe of Fabaceae based on a base chromosome number ( $x$ ) of 7 (rarely 8) and two collateral ovules (Barneby 1977). Western prairie clover (*Dalea ornata* [Douglas ex Hook.] Eaton & J. Wright) is an insect-pollinated legume that naturally is distributed throughout the northern Great Basin, southern Columbia River Plateau, and Snake River Plain (USDA, NRCS 2009). Its distribution, lack of toxicity to herbivores, and relatively upright growth habit make western prairie clover a candidate for commercial seed production.

In reseeding disturbed rangelands, there can be the potential for introducing genes from different populations into local gene pools. This potentially could reduce population fitness by increasing the frequency of maladapted genes or cytological differences in local populations or disrupting coadapted gene complexes (McKay et al. 2005). For outcrossing species that are sensitive to inbreeding depression, it can be important to increase genetic diversity in depauperate populations to increase population fitness (Ouborg et al. 2006) or to provide greater opportunities for adaptation to global change (Rice and Emery 2003). Therefore, identifying the amount and distribution of genetic variation in naturally occurring populations is important to make informed management decisions related to determination of conservation units.

Determining conservation units is not always straightforward, and includes analyses that infer short- and long-term evolutionary potential. The presence of local adaptation is suggested by growing collections in common gardens and identifying traits that are correlated to environmental variables at the collection sites (St. Clair et al. 2005). The quantitative traits correlated to collection site environments are indicative of the phenotypic variation necessary to respond to short-term environmental changes (McKay and Latta 2002), and this variation is important to fitness and adaptation. Long-term evolutionary potential is suggested by the use of molecular markers to infer population structure (McKay and Latta 2002). The resulting marker-based population structures identify gene flow barriers that can enable allelic differentiation, and can be correlated with higher differentiation for quantitative traits (Merila and Crnokrak 2001). Genetic structures, if detected by a sufficient number of markers and in a sufficient number of populations, indicate underlying genetic differences of groups of collections that can be used to preserve evolutionary potential. Genetic structures have been identified for several Great Basin and other semiarid

western United States species (Larson et al. 2004; Bushman et al. 2007; Jones et al. 2008; Philips et al. 2008).

Amplified fragment length polymorphism (AFLP) is a genetic fingerprinting technique useful for determining genetic diversity and population structure in various plant species (Meudt and Clarke 2007). The AFLP markers do not require extensive genetic information on the species of interest when generated, are largely considered neutral, and are robust in their ability to differentiate populations (Schlötterer 2004). Although AFLP markers are used to assess genetic diversity and define population structures in plants, few AFLP-based studies, or other DNA marker analyses, have been conducted on rangeland legumes.

The objectives of this study were to evaluate phenotypic and molecular marker diversity of 22 wildland collections of western prairie clover. Our hypothesis was that genetic population structure(s) exist within western prairie clover collections and that variation of some of the phenotypic characters is associated with the genetic variation. We assessed forage production and quality, flowering date, inflorescence weight, and growth habit in two common-garden settings, and genetic diversity among and within collections using AFLP markers. This information is useful in determining optimum release strategies of this species for commercial seed production.

## MATERIALS AND METHODS

### Seed Collections

Based on available literature and herbarium specimen data, 22 seed collections of western prairie clover were obtained from Washington, Idaho, and Oregon during the summer of 2005 (Table 1). Collection site data obtained at each collection included elevation, latitude, longitude (Table 1), and associated plant species. The collection sites ranged in elevation from 110 m to 1 163 m above sea level. Mean monthly temperature and precipitation data (average from 1961 to 1990) for each collection site were obtained from the Moscow Forestry Sciences Laboratory (2009). A 5–15-g section of insecticide strip (active ingredient Dochlorvos [2, 2-dichlorovinyl dimethyl phosphate] 18.6% and related compounds 1.4%, Hotshot No-Pest strip; United Industries Corp, St Louis, MO) was placed in each bag of collected seed to reduce seed weevil damage. Seeds were air-dried in a greenhouse, threshed in a Wintersteiger seed thresher (Model LD180; Wintersteiger, Des Moines, IA), and cleaned with sieves and a seed blower. Cleaned seeds were stored in a dark room maintained at 3°C with a relative humidity of 20–25%.

### Common-Garden Studies

Seeds were germinated in plastic boxes with moistened blotter paper at room temperature. After germination, seedlings were transplanted into Q-Plugs (International Horticulture Technologies, Hollister, CA), and placed within Ray Leach stubby cone-tainers (Stuewe and Sons, Corvallis, OR). The seedlings were grown in a greenhouse at the USDA–Agricultural Research Service Forage and Range Research Laboratory, Logan, Utah, under a 30/15°C day/night temperature regime. A single seedling was grown in each container. Adequate fertilization and watering was provided to the plants for 90 d

**Table 1.** Twenty-two western prairie clover collections with their collection site information, number of individuals sampled for genetic assays, percent polymorphic loci, and expected heterozygosity. Purple prairie clover (Dp) was used as a check and outgroup.

Collection ID	County and state	Latitude	Longitude	Elevation (m)	Temperature (°C)	No. of individuals	Polymorphic loci (%)	Expected heterozygosity
1	Wallowa, OR	45°41'N	116°48'W	631	10.5	8	34.4	0.12
2	Umatilla, OR	45°55'N	119°07'W	136	12.0	8	34.2	0.14
3	Benton, WA	45°56'N	119°20'W	131	11.9	8	36.7	0.14
4	Malheur, OR	43°24'N	117°07'W	1163	8.2	8	38.0	0.14
5	Benton, WA	46°21'N	119°21'W	148	11.8	8	41.6	0.15
6	Walla Walla, WA	46°04'N	118°54'W	110	12.3	8	40.7	0.16
7	Franklin, WA	46°17'N	119°11'W	160	11.8	8	36.9	0.14
8	Malheur, OR	42°47'N	117°43'W	1122	9.5	8	37.3	0.12
9	Sherman, OR	45°17'N	121°01'W	258	11.2	8	35.0	0.13
11	Elmore, ID	42°53'N	115°07'W	884	10.5	8	34.2	0.12
12	Owyhee, ID	42°44'N	115°54'W	945	10.4	8	34.2	0.12
13	Owyhee, ID	42°45'N	115°43'W	815	11.0	8	32.9	0.11
14	Jefferson, OR	44°43'N	121°02'W	696	9.6	8	32.1	0.10
15	Canyon, ID	43°19'N	116°35'W	713	11.5	8	33.5	0.12
16	Crook, OR	44°08'N	120°48'W	948	8.4	6	35.4	0.12
17	Wheeler, OR	44°37'N	120°07'W	1001	8.3	8	44.9	0.16
19	Asotin, WA	46°04'N	116°59'W	259	12.5	7	44.1	0.16
20	Wheeler, OR	44°37'N	120°20'W	699	10.0	8	45.1	0.17
21	Elmore, ID	43°02'N	115°09'W	960	9.9	8	30.8	0.09
22	Wheeler, OR	44°54'N	120°24'W	465	11.2	8	41.6	0.15
23	Jefferson, OR	44°34'N	121°02'W	941	8.2	7	28.5	0.09
25	Crook, OR	44°03'N	120°44'W	1038	8.0	8	29.7	0.10
Dp	—	—	—	—	—	8	40.1	0.15

before transplanting to field plots at Millville (lat 41°39'N, long 111°48'W, elevation 1350 m above sea level) and Hyde Park (lat 41°47'N, long 111°48'W, elevation 1380 m above sea level) locations in northern Utah. Soil at the Millville location is a Millville silt loam (coarse-loamy over sandy or sand-skeletal, mixed, superactive mesic, Calcic Haploxerolls). Soil at the Hyde Park location is a McMurdie silt loam (fine, montmorillonitic, mesic, Calcic Pachic Argixerolls). The experimental design at each of the common-garden locations was a randomized complete block with eight replications, and collections were assigned randomly within each replication. Each replication included five individual plants of each collection, for a total of 40 plants of each collection at each location. Plants were spaced 0.5 m apart within rows and between rows. Plants of commercially available purple prairie clover (*Dalea purpurea* Vent.) were included as checks at both locations. Seed for purple prairie clover came from transplants originated from a 16-ha commercial production field, started from a 0.5-kg pool of seed harvested from five prairie-remnant sites in southern Wisconsin (Oak Prairie Farm, Pardeeville, WI). Plots were planted in May 2006 and were routinely weeded and watered during the establishment year.

During the second and third year after establishment (2007 and 2008), collections at both locations were evaluated for dry matter yield (DMY) and number of inflorescences. For the DMY harvest at Hyde Park, plants were harvested at approximately 50% bloom (19 to 25 June) according to Miller (1984), and then after the first frost in October (17 to 25 October). At Millville, DMY was obtained after mature inflorescence harvest (26 July 2007, 4 August 2008), and then

again after the first frost in October (17 to 26 October). The summer and fall DMY harvest values were combined in each year at each location. The number of inflorescences was determined at 50% bloom at Hyde Park and immediately before inflorescence harvest at Millville in both years.

To minimize bias between early- and late-flowering collections, inflorescences were harvested sequentially from late June to late July each year at the Millville location. Because western prairie clover has racemose inflorescences and its mature seeds disperse from the inflorescence easily, plants were checked daily and the mature seed portions were collected until the entire inflorescence was harvested. The inflorescences were bulked by replication, dried in a greenhouse setting, and weighed. Flowering date was measured at Millville location over both years. The flowering date was recorded as the number of days from 1 January when the first flower emerged on each plant, and was averaged across plants for each collection on a replication basis.

Additional measurements of plant height, foliage diameter, crude protein (CP), acid detergent fiber (ADF), and neutral detergent fiber (NDF) were taken at the Hyde Park location. Inflorescence weight was not measured at Hyde Park because samples were harvested for DMY at 50% of the field bloom. Plant height and foliage diameter were obtained immediately prior to DMY harvest. Values of NDF, ADF, and CP were determined from the first DMY harvest from Hyde Park in 2007. Samples were ground (Cyclotec 1093 Sample Mill; Cyclotec, Hoganas, Sweden) to pass through a 1-mm diameter screen. Ground samples were analyzed for CP using total combustion procedures (LECO TruSpec C/N analyzer; LECO

Corp, St Joseph, MI). ADF and NDF were analyzed at Agri Analysis Inc (Leola, PA) following procedures of Association of Official Analytical Chemists (1990) and Mertens (2002), respectively, as adapted to the Ankom A200 filter bag technique.

Analysis of variance was conducted using the MIXED procedure of SAS (SAS Institute 2004) with collection considered a fixed factor and replication a random factor. Data were checked for normality and homoscedasticity before analysis, and subsequently transformed when needed. Significance was determined at  $\alpha = 5\%$ , and mean separations were conducted using the Tukey test. Pearson correlation coefficients and Spearman's rank coefficients were calculated using the CORR procedure of SAS.

Relationships among phenotypic traits and environmental and topographic variables at each collection site were evaluated by canonical correlation analysis (CCA) using the CANCELL procedure of SAS. Forage quality measurements were excluded from canonical correlation analysis, principal component analysis (PCA), and Mantel's test, because of their low variation among the collections, which reduced the power of discrimination for other phenotypic traits. Additionally, separate CCA, PCA, and Mantel's analyses were conducted on forage quality traits, and no significance was found. Therefore forage quality traits were not included in these analyses. The CCA procedure generated pairs of linear variable combinations with the highest correlations among a set of dependent (phenotypic) and independent (environmental and topographic) variables, which are termed canonical variables (Mardia et al. 1979).

PCA was conducted on the correlation matrix of phenotypic data using the FACTOR procedure of SAS. Significance of principal component axes was considered for components with eigenvalues greater than one, and an additional significance test was conducted with parallel analysis (PA; Horn 1965) using "paran" package in R (Dinno 2009). Significance for the grouping of collections in the principal component scatter plot was obtained using multiresponse permutation procedure (mrpp) using "vegan" package in R (Dixon 2003).

### Genetic Diversity and Population Structure

Plant tissues were collected from the shoot apical regions of eight individual plants from each collection prior to field transplanting. An additional eight plants of purple prairie clover were sampled for use as an outgroup. Tissues were lyophilized and stored at  $-20^{\circ}\text{C}$ . About 17 mg of dried plant tissue was used for each sample, and DNA was extracted using the Qiagen DNeasy 96-well procedure (Qiagen, Valencia, CA) following the manufacturer's protocol. The concentration and quality of genomic DNA were tested spectrophotometrically and by gel electrophoresis. Genomic DNA concentration was adjusted to approximately  $30 \text{ ng} \cdot \mu\text{L}^{-1}$ .

The AFLP procedure followed the method published in Vos et al. (1995), but was modified by resolving AFLP products on an ABI3730 capillary instrument (Life Technologies, Foster City, CA). Five selective primer pairs were chosen for analysis: E.AGG/M.CAT, E.AGA/M.CTG, E.AGC/M.CTA, E.AGT/M.CAC, and E.ACG/M.CTG. Amplified products were analyzed with Genescan (ABI, Foster City, CA) at the Center for Integrated Biosystems (Utah State University, Logan, UT). The

AFLP bands were scored visually for presence or absence using Genographer software (Benham 2001).

The percentage of polymorphic loci and within-collection expected heterozygosity ( $H_E$ ) was estimated in AFLP-SURV v1.0 (Vekemans et al. 2002) using the binary AFLP data. A matrix of pairwise genetic distance ( $F_{st}$ ), also obtained using AFLP-SURV, was used to construct a neighbor-joining (NJ) tree with Phylip software v3.69 (Felsenstein 2009). To assess model strength, bootstrap analysis was conducted using 1 000 permutations. Bayesian clustering analysis was conducted to assess population structures using STRUCTURE v2.2 (Falush et al. 2007), without advanced assignment of collections into populations. The raw binary data were analyzed with the RECESSIVE ALLELES option and the admixture model. Probabilities of  $K = 1$  through  $K = 9$  groupings were tested with three replications for each level of  $K$ . The Markov chain Monte Carlo (MCMC) procedure was used with 30 000 burn-in and 300 000 MCMC steps after burn-in to determine the probability of each structure model. The average estimated log probability of the data was plotted against the  $K$  values to observe structural fit of the data. Differences between the  $K$  and  $K-1$  average log probability values also were plotted to show changes in model fit, wherein a small change is an indication of the optimum number of robust population structures.

Important groups found in the STRUCTURE program and NJ tree were further tested with hierarchical Analysis of Molecular Variance (AMOVA), using Arlequin v3.1 software (Excoffier et al. 2005). First, a pairwise similarity matrix was generated among collections by estimating Dice's coefficient (Dice 1945) from the binary data of AFLP. Dice's coefficient was preferred over the simple matching coefficient because it estimates similarity based only on the band presence in at least one of the two individuals and thus reduces the bias in similarity estimates due to band absence in both individuals (Wong et al. 2001). The Dice's distance matrix was then used as an input file for the hierarchical AMOVA.

Mantel's  $Z$  statistic (Mantel 1967) was used to detect relationships among the genetic, phenotypic, and site characteristic distance matrices using the MXCOMP procedure of NTSYSpc (Rohlf 1998). For the genetic distance matrix, an  $F_{st}$  matrix obtained from AMOVA was used. A phenotypic matrix was constructed by using all phenotypic traits except forage quality traits. Collection-site characteristics (geographic distance, elevation, temperature, and precipitation) were each used to create site matrices. The ability of phenotypic traits to discriminate groups of collections identified by STRUCTURE or PCA was assessed by first using the STEPDISC procedure of SAS to determine the appropriate variables. Then the DISCRIM procedure of SAS was used to test significance of the appropriate discriminatory variables.

## RESULTS

### Common Gardens

The interaction between collections and years for DMY was not significant for either common-garden location (Hyde Park:  $F_{22,307} = 0.37$ ,  $P = 0.99$ ; Millville:  $F_{22,307} = 0.96$ ,  $P = 0.47$ ). Therefore, DMY was combined across years within each location. The DMY was not combined across locations because

**Table 2.** Means, standard deviations (SD), and F statistics of traits in western prairie clover collections measured at Hyde Park, Millville, and combined across both locations.

Trait	Mean	SD	F statistic
Hyde Park			
Dry matter yield, g · plot <sup>-1</sup>	165.0	48.3	F <sub>22,322</sub> = 11.08*** <sup>1</sup>
Plant height, cm	32.7	3.2	F <sub>22,321</sub> = 8.58***
Number of inflorescences	36.2	13.1	F <sub>22,322</sub> = 12.13***
Foliage diameter, cm	56.5	6.7	F <sub>22,321</sub> = 10.81***
Acid detergent fiber, %	34.0	1.5	F <sub>22,67</sub> = 2.53**
Neutral detergent fiber, %	43.1	2.4	F <sub>22,67</sub> = 3.91***
Crude protein, %	17.8	0.7	F <sub>22,63.1</sub> = 2.24**
Millville			
Dry matter yield, g · plot <sup>-1</sup>	113.5	33.4	F <sub>22,322</sub> = 15.83***
Inflorescence weight, g · plot <sup>-1</sup>	66.8	25.8	F <sub>22,318</sub> = 15.76***
Number of inflorescences	61.0	17.4	F <sub>22,322</sub> = 12.30***
Flowering date <sup>2</sup>	172.8	4.4	F <sub>22,318</sub> = 17.72***

<sup>1</sup>\*\*\* indicates  $P < 0.01$ ; \*\*\*,  $P < 0.001$ .

<sup>2</sup>Flowering date starts from 1 January.

of a significant interaction between locations and collections ( $F_{22,667} = 2.88$ ,  $P < 0.0001$ ). Collections differed significantly for DMY at both locations (Table 2), ranging from 85 g · plot<sup>-1</sup> to 242 g · plot<sup>-1</sup> at Hyde Park and 28 g · plot<sup>-1</sup> to 188 g · plot<sup>-1</sup> at Millville (Table 3). Collections Do9 and Do19 produced the greatest DMY at Hyde Park and Millville locations, respectively, but they were not significantly different from many other collections (Table 3). The 12 collections of western prairie clover with the highest average DMY values at Hyde Park were significantly higher in DMY than purple prairie clover (82 g · plot<sup>-1</sup>); however, at Millville they did not differ from purple prairie clover (135 g · plot<sup>-1</sup>). Collections Do1 and Do12 exhibited the lowest average DMY at both locations.

At Hyde Park, the collection by year interaction was not significant for number of inflorescences ( $F_{22,300} = 0.81$ ,  $P = 0.72$ ); hence, these data were combined across years. At Millville, a significant interaction was observed between collections and years for number of inflorescences ( $F_{22,307} = 2.24$ ,  $P = 0.0014$ ). However, rankings among collections were consistent as indicated by the Spearman rank coefficients ( $r = 0.63$ ,  $P = 0.0013$ ), and did not change among the highest and lowest collections. Therefore, we also combined the data across years for number of inflorescences. At both locations, number of inflorescences varied significantly (Table 2) with

**Table 3.** Means and differences for dry matter yield (DMY), number of inflorescences (No. of infl.), plant height, foliage diameter, acid detergent fiber (ADF), neutral detergent fiber (NDF), crude protein (CP), inflorescence weight (Infl. weight), and flowering date in western prairie clover collections, and purple prairie clover (Dp) at Hyde Park and Millville locations. Values followed by the same letter within a column are not significantly different at  $P = 0.05$ .

Collection ID	Hyde Park							Millville			
	DMY (g)	No. of infl.	Plant height (cm)	Foliage diameter (cm)	ADF (%)	NDF (%)	CP (%)	DMY (g)	No. of infl.	Infl. weight (g)	Flowering date <sup>1</sup>
1	85.7 g	19.4 fgh	26.3 f	64.2 abc	35.6 ab	39.3 d	17.0 b	27.9 g	20.8 h	20.0 i	173.8 abcdef
2	180.1 abcd	32.6 cdef	30.5 def	51.2F gh	33.3 ab	41.1 abcd	17.6 ab	108.8 cde	55.4 bcdef	46.3 defgh	176.6 ab
3	180.9 abcd	41.5 abcd	34.0 abcde	56.2 bcdefg	33.4 ab	41.7 ab	16.8 ab	120.2 bcde	67.8 abcde	56.0 cdefg	173.8 abcdef
4	187.5 abc	56.8 ab	38.4 ab	57.5 bcdefg	34.8 a	47.0 ab	17.7 ab	113.0 bcde	81.5 ab	86.1 abc	169.2 fgh
5	174.0 abcd	38.3 abcde	33.8 abcde	52.3 efgh	33.6 ab	41.5 abcd	18.1 ab	98.3 de	51.8 cdef	45.1 efgh	177.1 ab
6	186.2 abc	29.6 cdef	31.3 cdef	53.8 defgh	31.9 b	42.1 abcd	18.1 ab	154.3 abc	66.8 abcde	66.1 bcdefg	177.9 a
7	164.7 abcdef	29.2 cdef	31.6 cdef	53.5 defgh	33.0 ab	42.2 abcd	19.2 a	130.9 bcd	48.7 cdefg	44.6 efgh	174.9 abcde
8	115.6 defg	39.7 bcde	33.1 bcde	49.2 gh	35.2 ab	49.5 a	18.4 ab	91.5 def	60.0 abcdef	52.2 defgh	170.6 defgh
9	242.3 a	43.0 abcd	34.5 abcde	65.7 ab	35.6 ab	47.0 abc	16.7 b	156.4 ab	73.6 abc	114.2 a	170.4 efgh
11	98.0 fg	27.6 defg	33.1 bcde	50.7 gh	33.7 ab	41.7 abcd	18.3 ab	78.9 ef	39.6 fgh	36.8 ghi	175.8 abc
12	84.9 g	12.4 gh	33.0 bcde	45.1 h	32.8 ab	40.9 cd	16.9 ab	55.6 fg	27.3 gh	27.3 hi	176.8 ab
13	123.1 cdefg	20.9 efgh	36.7 abc	55.7 cdefg	34.2 ab	43.6 abcd	18.3 ab	98.7 de	41.5 efgh	43.4 fgh	176.6 ab
14	203.4 ab	59.5 a	30.4 def	58.8 abcdefg	33.5 ab	43.9 abcd	18.3 ab	113.7 bcde	68.4 abcd	77.4 abcde	167.4 ghi
15	132.3 cdefg	33.6 cdef	34.2 abcde	55.6 cdefg	34.3 ab	45.4 ab	17.1 ab	92.7 def	43.9 defg	44.2 efgh	175.4 abcde
16	195.6 abc	44.6 abcd	32.0C de	61.0 abcdef	34.4 ab	45.0 abcd	17.8 ab	114.6 bcde	74.3 abc	74.1 abcde	165.3 hi
17	170.5 abcde	39.8 abcde	26.2 f	61.7 abcde	34.6 ab	43.1 abcd	18.6 ab	115.2 bcde	80.2 ab	94.0 ab	169.5 fgh
19	226.9 ab	30.8 cdef	39.5 a	63.3 abcd	35.1 ab	43.6 abcd	17.3 ab	187.6 a	67.6 abcde	82.8 bcdef	175.8 abcd
20	210.8 ab	37.8 bcde	29.5 ef	62.2 abcd	36.0 ab	42.3 abcd	17.7 ab	114.6 bcde	66.6 abcde	99.4 ab	171.2 cdefg
21	155.1 bcdef	37.1 bcde	35.6 abcd	57.2 bcdefg	30.0 b	40.8 bcd	17.2 ab	130.8 bcd	64.0 abcdef	70.9 bcdefg	172.6 bcdefg
22	240.1 a	44.5 abcd	31.8 cde	67.8 a	37.1 a	44.1 abcd	17.8 ab	147.3 abc	86.2 a	101.8 ab	175.1 abcde
23	166.7 abcdef	55.7 ab	31.6 cdef	51.2 fgh	34.2 ab	42.0 bcd	18.7 ab	114.6 bcde	83.1 a	75.7 abcde	162.7 i
25	188.4 abc	48.7 abc	31.7 cde	63.5 abc	32.0 ab	42.6 bcd	17.2 ab	110.5 bcde	69.9 abcd	83.4 abcd	167.5 ghi
Dp	82.4 efg	9.1 h	33.3 abcde	41.5 h	32.8 ab	42.1 abcd	18.7 ab	134.8 abcde	63.2 abcdef	95.1 abc	179.5 a

<sup>1</sup>Flowering date is measured from 1 January.

**Table 4.** Pearson correlation coefficients (*r*) for dry matter yield (DMY), inflorescence weight (Infl. weight), number of inflorescences (No. of infl.), flowering date, plant height, foliage diameter, acid detergent fiber (ADF), neutral detergent fiber (NDF), and crude protein (CP) of 22 western prairie clover collections.

	DMY (Hyde Park)	DMY (Millville)	Infl. weight	No. of infl. (Millville)	No. of infl. (Hyde Park)	Flowering date	Plant height	Foliage diameter	ADF	NDF
DMY (Millville)	0.84*** <sup>1</sup>	—	—	—	—	—	—	—	—	—
Infl. weight	0.85***	0.71***	—	—	—	—	—	—	—	—
No. of infl. (Millville)	0.80***	0.71***	0.88***	—	—	—	—	—	—	—
No. of infl. (Hyde Park)	0.61**	0.37 ns	0.66***	0.82***	—	—	—	—	—	—
Flowering date	-0.26 ns	-0.05 ns	-0.50*	-0.58**	-0.76***	—	—	—	—	—
Plant height	0.15 ns	0.4 ns	0.07 ns	0.09 ns	0.04 ns	0.19 ns	—	—	—	—
Foliage diameter	0.62**	0.38 ns	0.66***	0.42 ns	0.30 ns	-0.25 ns	-0.12 ns	—	—	—
ADF	0.24 ns	0.00 ns	0.30 ns	0.17 ns	0.11 ns	-0.11 ns	-0.11 ns	0.45*	—	—
NDF	0.27 ns	0.28 ns	0.41 ns	0.41 ns	0.43*	-0.32 ns	0.37 ns	0.12 ns	0.42 ns	—
CP	-0.06 ns	0.07 ns	0.04 ns	0.16 ns	0.17 ns	-0.16 ns	0.22 ns	-0.30 ns	0.01 ns	0.08 ns

<sup>1</sup>ns indicates not significant; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ .

values ranging from 9 to 86 inflorescences per plant (Table 3). Collections with the highest number of inflorescences were Do14 (Hyde Park) and Do22 (Millville), both of which originated in central Oregon (Table 1). Purple prairie clover had the fewest inflorescences (9) at Hyde Park, and an intermediate number of inflorescences (63) at Millville. Collections Do1 and Do12 had the lowest numbers of inflorescences among the collections.

For inflorescence weight at Millville, the collection by year interaction was not significant ( $F_{22,305} = 1.55$ ,  $P = 0.06$ ) and the trait was combined across years. Collections differed significantly for inflorescence weight ( $F_{22,318} = 15.76$ ,  $P < 0.0001$ ) (Table 2) with collections from central Oregon generally having the highest values (Table 3). Collection Do9 had the greatest inflorescence weight ( $114 \text{ g} \cdot \text{plot}^{-1}$ ), yet did not differ from 10 other collections or from purple prairie clover (Table 3). Collection Do1 had the lowest inflorescence weight ( $20 \text{ g} \cdot \text{plot}^{-1}$ ); however, it was not significantly different from Do12 and Do11. Flowering dates ranged from 163 d (Do23) to 178 d (Do6) among Do collections, with average flowering date occurring about 21 June. Collection Do23 exhibited the earliest flowering but did not differ from Do16, Do14, and Do25. All of these collections were from the Deschutes River watershed in central Oregon. The latest flowering collection was Do6, which did not differ from 11 other collections and purple prairie clover. Purple prairie clover was later (180 d) in flowering than all western prairie collections (Table 3).

The collection by year interaction was not significant for plant height ( $F_{22,300} = 1.05$ ,  $P = 0.40$ ) or for foliage diameter ( $F_{22,300} = 0.72$ ,  $P = 0.82$ ) at the Hyde Park location. Therefore, these traits were combined across years. Significant variation was detected among the collections for plant height and foliage diameter (Table 2). Plant height ranged from 26 cm to 40 cm with Do19 having the greatest height, though it did not differ from seven other collections or from purple prairie clover (33 cm; Table 3). Collection Do17 was the shortest, but it did not differ from seven other collections. Foliage diameter ranged from 42 cm to 68 cm for the collections. Collection Do22 had the greatest foliage diameter, but it did not differ from eight other collections. Collection Do12 had the least foliage diameter (45 cm) among collections, but it did not differ from

seven other collections and purple prairie clover. The purple prairie clover check had a smaller foliage diameter (42 cm) than all the western prairie clover collections.

Differences were detected for ADF, NDF, and CP among the collections (Table 2). The ADF ranged from 30.0% to 37.1%, NDF ranged from 39.3% to 49.5%, and CP ranged from 16.7% to 19.2% (Table 3). Collection Do24 exhibited the greatest ADF, which differed from Do21 and Do6 but not from 19 other collections or from purple prairie clover. Collection Do8 had the greatest NDF, but it differed significantly only from five other collections (Table 3). Similarly, Do7 exhibited the greatest CP, which differed only from Do1 and Do9.

### Correlations of Phenotypic Characters

A high positive correlation ( $r = 0.84$ ,  $P < 0.0001$ ) was observed for DMY (Table 4), and number of inflorescences ( $r = 0.82$ ,  $P < 0.0001$ ) between collections at Hyde Park and Millville. The DMY at Hyde Park also was correlated with inflorescence weight ( $r = 0.85$ ,  $P < 0.0001$ ), numbers of inflorescences at both locations (Hyde Park:  $r = 0.61$ ,  $P = 0.0027$ ; Millville:  $r = 0.80$ ,  $P < 0.0001$ ), and foliage diameter ( $r = 0.62$ ,  $P = 0.0023$ ). The DMY at Millville also was correlated with inflorescence weight ( $r = 0.71$ ,  $P = 0.0002$ ) and numbers of inflorescences at Millville ( $r = 0.71$ ,  $P = 0.0002$ ). Flowering date was negatively correlated with number of inflorescences (Hyde Park:  $r = -0.76$ ,  $P < 0.0001$ ; Millville:  $r = -0.58$ ,  $P = 0.005$ ) and inflorescence weight (Millville:  $r = -0.50$ ,  $P = 0.018$ ), which indicated that earlier flowering plants produced more flowers and greater inflorescence weights. Among forage quality traits, NDF showed a moderate correlation to the number of inflorescences (Hyde Park:  $r = -0.43$ ,  $P = 0.046$ ), and ADF showed a moderate correlation to foliage diameter ( $r = 0.45$ ,  $P = 0.036$ ). No correlations were detected among ADF, NDF, and CP (Table 4).

Associations between phenotypic traits and environmental and topographic characteristics of the collection sites were assessed using canonical correlation analysis to identify traits with possible local adaptive significance. The first and second pairs of canonical variables contained 67% and 27% of the variation in the overall dataset, respectively, and only these two pairs of canonical variables exhibited significant correlations

**Table 5.** Correlations between original variables and canonical variables for traits. The phenotypic datasets (PHEN1 and PHEN2) are canonical variables obtained from phenotypic traits, and environmental and topographic datasets (CLIM1 and CLIM2) are canonical variables obtained from environmental and topographic variables.

Original variables	PHEN1	PHEN2
Dry matter yield, Millville	-0.27 ns <sup>1</sup>	-0.39 ns
Dry matter yield, Hyde Park	-0.07 ns	-0.69***
Inflorescence weight	0.34 ns	-0.42 ns
No. of inflorescences, Hyde Park	0.53*	-0.55**
No. of inflorescences, Millville	0.38 ns	-0.54**
Plant height	-0.27 ns	0.39 ns
Foliage diameter	0.17 ns	-0.30 ns
Flowering date	-0.87***	0.38 ns
	CLIM1	CLIM2
Precipitation	0.45*	0.05 ns
Temperature	-0.99***	0.02 ns
Elevation	0.86***	0.49*

<sup>1</sup>ns indicates not significant; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ .

( $P < 0.0001$  and  $P = 0.0049$ , respectively). The first canonical variable for the phenotypic data set (PHEN1) primarily represented flowering date ( $r = -0.87$ ,  $P < 0.0001$ ). The first canonical variable for the environmental and topographic data set (CLIM1) was represented primarily by temperature ( $r = -0.99$ ,  $P < 0.0001$ ) and elevation ( $r = 0.86$ ,  $P < 0.0001$ ; Table 5). There was a high negative correlation between PHEN1 and temperature ( $r = -0.95$ ,  $P < 0.0001$ ) and a positive correlation between PHEN1 and elevation ( $r = 0.82$ ,  $P < 0.001$ ; data not shown), indicating that flowering date was positively correlated with temperature and negatively correlated with elevation. The second canonical variable for the phenotypic data set (PHEN2) was correlated with DMY at Hyde Park ( $r = -0.69$ ,  $P = 0.0004$ ) and the number of inflorescences at both locations (Hyde Park:  $r = -0.55$ ,  $P = 0.008$ ; Millville:  $r = -0.54$ ,  $P = 0.009$ ; Table 5). The second canonical variable for the environmental and topographic dataset (CLIM2) correlated with elevation ( $r = 0.49$ ,  $P = 0.021$ ) but not with precipitation or temperature (Table 5). Thus, the phenotypic character with the greatest correlation to collection site environment was flowering date, which responded to elevation and temperature.

### Principal Component Analysis of Phenotypic Characters

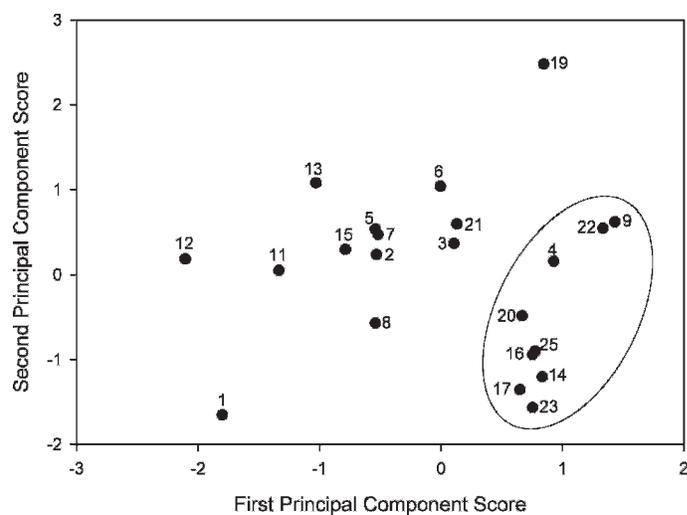
Eigenvalues of the first three principal components were greater than one and described 89% of the total variation. The first principal component was significant using parallel analysis (Table 6). The first principal component described 57% of the phenotypic variation, the second described 19%, and the third described 13%. Correlations for the first principal component were significant for DMY (Hyde Park:  $r = 0.91$ ,  $P < 0.0001$ ; Millville:  $r = 0.76$ ,  $P < 0.0001$ ), inflorescence weight ( $r = 0.95$ ,  $P < 0.0001$ ), number of inflorescences (Hyde Park:  $r = 0.80$ ,  $P < 0.0001$ ; Millville:  $r = 0.94$ ,  $P < 0.0001$ ), foliage diameter ( $r = 0.63$ ,  $P < 0.0018$ ), and flowering date ( $r = -0.57$ ,  $P = 0.0057$ ; Table 6). Correlations for the second principal component were significant for DMY at Millville ( $r = 0.56$ ,  $P = 0.0065$ ), plant height ( $r = 0.74$ ,  $P < 0.0001$ ), and flowering

**Table 6.** Phenotypic traits used in principal component analysis with their respective correlations in the first three principal component (PC) axes. Eigenvalues, adjusted eigenvalues from parallel analysis (PA), and the cumulative proportion of the total variation of the measured trait data set were given for each principal component axis.

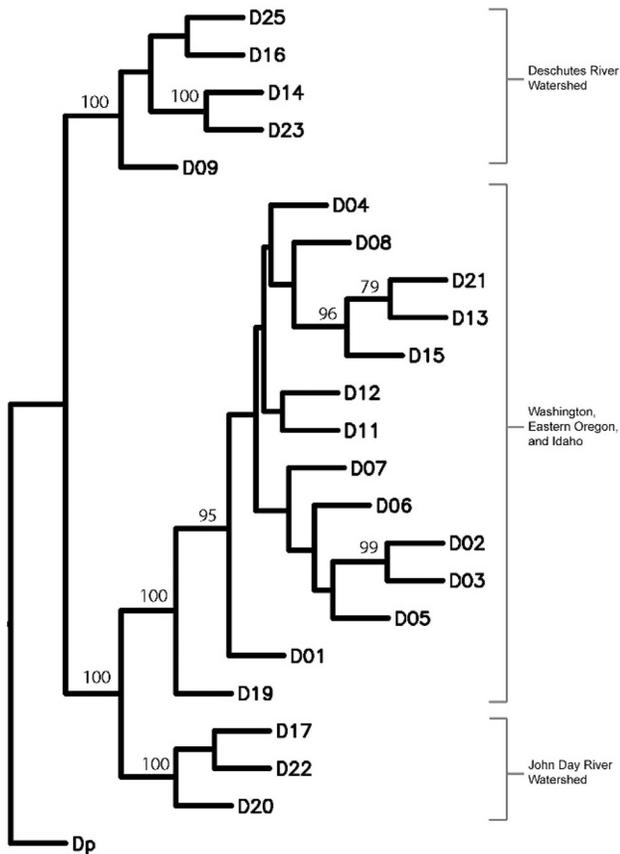
Traits	PC 1	PC 2	PC 3
Dry matter yield, Hyde Park	0.91*** <sup>1</sup>	0.24 ns	0.18 ns
Dry matter yield, Millville	0.76***	0.56**	0.02 ns
Inflorescence weight	0.95***	0.01 ns	0.13 ns
No. of inflorescences, Hyde Park	0.80***	-0.35 ns	-0.37 ns
No. of inflorescences, Millville	0.94***	-0.05 ns	-0.16 ns
Plant height	0.12 ns	0.74***	-0.53*
Foliage diameter	0.63**	-0.04 ns	0.64**
Flowering date	-0.57**	0.69***	0.32 ns
Eigenvalue	4.56	1.52	1.01
Adjusted eigenvalues (PA)	3.27	0.77	0.59
Cumulative proportion	0.57	0.76	0.89

<sup>1</sup>ns indicates not significant; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ .

date ( $r = 0.69$ ,  $P = 0.0004$ ), and the factor loading for the third principal component was significant for foliage diameter ( $r = 0.64$ ,  $P = 0.0014$ ; Table 6). The scatter plot of the collections based on the first and second principal component scores showed a clustering of collections with two outlying collections (Fig. 1). Collections originating from the Deschutes River and John Day River watersheds in central Oregon, as well as Do4 from eastern Oregon, exhibited significant grouping compared to the remaining collections (Fig. 1, circled), with an observed delta value of 1.307 and an expected delta value of 1.784 ( $P = 0.001$ ). The Do19 collection had the highest loadings for both principal components (i.e., high values for both sets of phenotypic traits), whereas Do1 had the lowest loadings for both components. Remaining collections from southeastern Washington, eastern Oregon, and Idaho were a large heterogeneous group.



**Figure 1.** Scatter plots of the first principal component score with the second principal component score obtained from the western prairie clover phenotypic data. The circle indicates a significant group designation from principal component analysis.

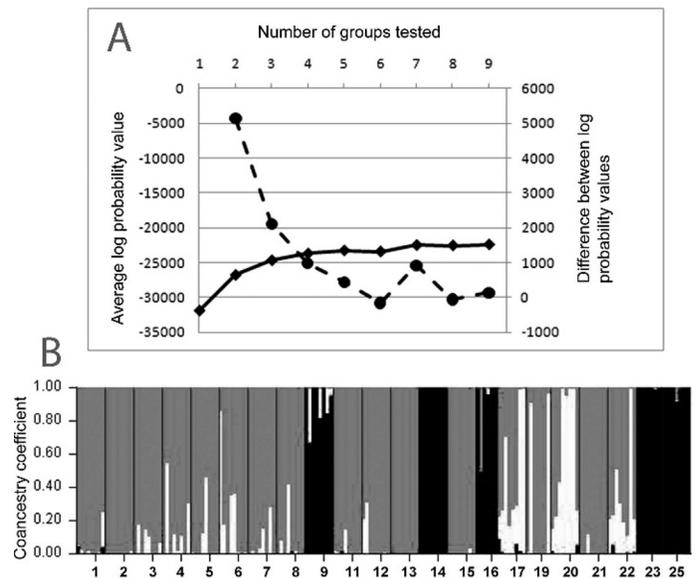


**Figure 2.** A consensus neighbor-joining dendrogram constructed using an  $F_{st}$  distance matrix of 22 western prairie clover collections, with associated bootstrap values (%). Only bootstrap values equal to or higher than 75% are shown.

### Genetic Diversity and Population Structure

Genetic diversity and population structure for the western prairie clover collections were determined with 474 AFLP bands. The percentage of polymorphic loci ranged between 29% and 45% with a mean of 36%, and the expected heterozygosity ranged from 0.09 to 0.17 with a mean of 0.13 (Table 1). AMOVA apportioned 31% of the variation among collections and 69% of variation within collections. Based on  $F_{st}$  distances, an NJ tree was used to investigate genetic structures, which showed two broad groups. The first group comprised five collections from the Deschutes River watershed (Fig. 2). The second group consisted of collections from geographically diverse areas in Idaho, Washington, and Oregon. Within the larger second group, three collections from the John Day River watershed, collection Do19, and collection Do1 each formed subgroups with bootstrap support. Additionally, several two-group subgroups of geographically approximate collections with high bootstrap support were detected.

Bayesian clustering was used to test for population structures, wherein one to nine groups were tested. The increase in average log probability values was minimal after the third group test, and the rate of change in the log probability values decreased sharply after that. These data indicated that little



**Figure 3.** **A.** The average log probability values of Bayesian clustering (solid line), and the differences between average log probability values (dashed line) of successive groups, testing groups 1–9. **B.** Bar plot showing Bayesian clustering results, where western prairie clover collections were fit into three a priori groups. Collection numbers correspond to those in Table 1 and smaller bars within collections represent the individual plants.

improvement in the model was obtained by including more than three groups (Fig. 3A). Five collections from the Deschutes River watershed in central Oregon constituted one genetically differentiated group (Fig. 3B). Fourteen collections from Idaho, Washington, and Oregon formed a second group. The Do17, Do20, and Do22 collections from the John Day River watershed were detected as a third group, but these three collections showed substantial admixture in that a large proportion of their AFLP markers were shared with the second group (Fig. 3B).

To test the significance of the three groups identified by Bayesian clustering, hierarchical AMOVA was used. Thirty-one percent of the total genetic variation was apportioned among the three groups, 10% ( $P < 0.001$ ) of the variation was apportioned among collections within groups, and the remaining 59% ( $P < 0.001$ ) was apportioned within collections. In comparison, when only two major groups, the Deschutes River watershed and the remaining collections (now including the John Day collections), were tested in hierarchical AMOVA, 34% ( $P < 0.001$ ) of the total genetic variation was apportioned among the two groups and 12% ( $P < 0.001$ ) among collections within groups. The slight difference of 3% of the variation apportioned among groups when comparing three-group and two-group scenarios suggests that the separation of the collections from the John Day River watershed as a population structure was not strong.

To determine if the marker-based genetic distances were correlated to the phenotypic, environmental, and geographic matrices, Mantel's correlation test statistic ( $Z$ ) was applied. Significance was found for the correlation between the phenotypic and genetic distance matrices ( $r = 0.33$ ,  $P = 0.005$ ), the phenotypic and geographic distance matrices ( $r = 0.35$ ,

$P = 0.002$ ), and the genetic and geographic distance matrices ( $r = 0.31$ ,  $P = 0.009$ ). Unlike some of the phenotypic characters, the genetic distance matrix did not correlate with precipitation or elevation differences (data not shown). However, the genetic distance matrix was correlated to mean annual temperature differences ( $r = 0.28$ ,  $P = 0.004$ ).

Because the genetic matrix was correlated with the phenotypic distance matrix, individual phenotypic characters were tested for their ability to discriminate the groups of collections identified by genetically differentiated groups (Fig. 3B), and the phenotypic group (Fig. 1). When the three genetically differentiated groups were tested (Deschutes River watershed collections, John Day River watershed collections, and remaining collections from Idaho, Washington, and eastern Oregon; Fig. 3B), flowering date, inflorescence weight, and plant height discriminated the three groups with 100% accuracy, with a Wilk's Lambda value of  $F_{6,34} = 14.30$  ( $P < 0.0001$ ). When only two genetically differentiated groups were compared (Deschutes River group and all the other collections including John Day), flowering date alone was retained in the stepwise regression model and discriminated with 100% accuracy ( $F_{1,20} = 29.33$ ,  $P < 0.0001$ ). When the phenotypic-based group of Deschutes and John Day River collections (with collection Do4) was compared against the Idaho, Washington, and eastern Oregon collections (Fig. 1), inflorescence weight, flowering date, and dry matter yields at both locations were identified as discriminating variables with 100% accuracy ( $F_{4,17} = 46.37$ ,  $P < 0.0001$ ). Finally, when comparing just the three John Day River collections against those in Idaho, Washington, and eastern Oregon, inflorescence weight, plant height, and the number of inflorescences at the Millville location, but not flowering date, were discriminatory variables with 100% accuracy ( $F_{3,13} = 28.67$ ,  $P < 0.0001$ ).

## DISCUSSION

Legumes for rangeland revegetation should ideally 1) have a broad distribution to minimize unanticipated crossing with related species; 2) be adapted to and compatible in the range of sites in which they will be planted; 3) be amenable to agronomic seed production; and 4) preferably be safe for grazing animals. Western prairie clover meets those requirements. It has a distribution in central and eastern Oregon, southeastern Washington, western Idaho, and several northern counties in Nevada and California (USDA, NRCS 2009); this precludes the risk of unanticipated crossing in those areas because hybridization with related species is likely already to have occurred naturally. Plant collections often are based on existing herbarium records, and require at least 100 individual plants at a site for a collection to be made. The collections we made of western prairie clover constituted many, but not all, of the counties reported to contain this species, and thus the phenotypic and genetic assessments reported herein might be limited to the geographical regions from which collections were obtained. As future populations of this species are reported, it would be interesting to extend phenotypic and genetic assessments to those populations.

The presence of western prairie clover from a range of elevations, ecoregions, temperature gradients, and precipitation regimes suggests broad adaptability, and phenotypic assess-

ments in the common-garden plots found significant variation for all characters (Tables 2, 3). The number of inflorescences, inflorescence weight, and plant height were used as general indicators of the potential for seed production and mechanical harvest. For all three characters, some western prairie clover collections had values greater than purple prairie clover, a related species that is commercially grown in the midwestern United States.

Dry matter yield and forage quality are important traits for livestock and wildlife grazing. Western prairie clover collections varied widely for DMY ( $28 \text{ g} \cdot \text{plot}^{-1}$  to  $188 \text{ g} \cdot \text{plot}^{-1}$  at Millville and  $85 \text{ g} \cdot \text{plot}^{-1}$  to  $242 \text{ g} \cdot \text{plot}^{-1}$  at Hyde Park), and on average greater DMY was obtained at Hyde Park than Millville ( $165 \text{ g} \cdot \text{plot}^{-1}$  vs.  $114 \text{ g} \cdot \text{plot}^{-1}$ , respectively). Dry matter yields were not associated with forage quality traits, suggesting that selection for high DMY would not necessarily lead to reduced forage quality, although some studies have shown otherwise (White and Wight 1984; van der Wal et al. 2006; Bhattarai et al. 2008). Despite significant differences among collections for forage quality traits, the differences were relatively small (Table 3). On average, western prairie clover exhibited promising forage quality traits when compared to other legumes. Western prairie clover exhibited greater forage quality compared to basalt milkvetch (*Astragalus filipes* Torr. ex A. Gray), a western North American legume evaluated at similar location (Bhattarai et al. 2008). In addition, purple prairie clover exhibited higher forage quality over 14 other legumes that were evaluated in Missouri, in the central United States (McGraw et al. 2004), and our study showed western prairie clover has comparable forage quality value with the purple prairie clover (Table 3).

Strong positive correlations among DMY, inflorescence weight, and number of inflorescences (Table 4) suggests that selection for high DMY likely would lead to greater seed yields, and that the number of inflorescences could be used to indirectly select for seed yield. Similar results were reported for *Astragalus filipes* (basalt milkvetch), another legume from the US Intermountain West (Bhattarai et al. 2008). Collections with the greatest inflorescence weight and DMY were Do9, Do22, and Do19, yet these were not significantly different than half of the other collections (Table 3). Interestingly, these three collections each represented one of the three different genetically differentiated groups: Do9 from the Deschutes River watershed, Do22 from the John Day River watershed, and Do19 from the remaining collections (Fig. 3B). These collections could be excellent contributions for germplasm release or development of breeding populations (Jones 2009). It would be interesting, therefore, to expand locality tests within geographic boundaries of those genetically differentiated groups.

Both NJ dendrogram and Bayesian clustering analyses separated the collections from the Deschutes River watershed as a homogeneous and distinct group (Figs. 2 and 3B). This homogeneity was detected at the two-group ( $K = 2$ ) test in the STRUCTURE program and remained consistent through successive tests (Fig. 3B and data not shown). Thus, the Deschutes River watershed collections comprise a genetically differentiated group of collections. Among the remaining collections, both Bayesian clustering and the NJ dendrogram identified significant groupings of the John Day River watershed collections (Figs. 2 and 3B). The John Day collections,

however, showed shared coancestry (genetic admixture) with the group of collections from Washington, eastern Oregon, and Idaho. This admixture also remained consistent throughout all further Bayesian clustering tests, and indicates that the collections from the John Day River watershed are genetically more similar to collections from more distant sites in this study than to those from the nearby Deschutes River watershed. For revegetation purposes, the presence of admixture and small (albeit significant) differences in molecular variance might not necessitate treating the John Day collections as a separate genetically differentiated group.

Although collections from the Deschutes and John Day River watersheds are genetically differentiated, they are in close geographic proximity and no apparent seed barrier exists between them. Major differences are not apparent in the climate or soils of the two watersheds (P. Pedone, personal communication, October 2009), and PCA of phenotypic measurements was not able to separate the two groups (Fig. 1). The causes of genetic differentiation among such groupings often cannot be determined succinctly, but possibilities include genetic drift, migration, selection/adaptation, and range expansion. The AFLP markers can represent the cumulative effects of such processes and show evidence of gene-flow barriers. Because molecular markers grouped the John Day River collections either separately or with distant collections in Idaho, eastern Oregon, and Washington (Fig. 2), and phenotypic measurements group the John Day River collections with nearby Deschutes River collections (Fig. 1), further studies are needed to determine the placement of John Day River watershed collections into a conservation unit.

One possible explanation for the genetic differences between the John Day River and Deschutes River watershed collections might be a differential historical disturbance of the Deschutes River watershed, which might have reduced its population size. This hypothesis is supported by the trend of relatively low expected heterozygosity ( $H_E$ ) in collections from the Deschutes River watershed (average < 11%) compared to collections from the John Day River watershed (average = 16%; Table 1). Additionally, the homogeneous grouping of individuals (Fig. 3B) and north-south distribution along the Deschutes River suggests that these collections might have originated from the same ancestral population. The Deschutes River was a popular route for Native Americans to travel to and from the Columbia River (Ray et al. 1938), and seed migration might have occurred along this corridor.

Traits with adaptive significance can be correlated with environmental variables of the collection sites because environmental variables partly determine the selective factors that will give rise to parallel variation in traits (Endler 1986). In our study, flowering date was positively correlated with temperature ( $r = 0.83$ ,  $P < 0.0001$ ) and negatively correlated with elevation ( $r = -0.55$ ,  $P = 0.0079$ ). Because high-elevation sites typically are colder than low-elevation sites, plants from high elevations generally grow and flower at relatively lower ambient temperatures than plants from lower elevations. Consequently, when grown in a common garden, western prairie clover collections from high-elevation, low-temperature sites had an earlier flowering date than those from low-elevation sites.

In this study a correlation ( $r = 0.33$ ,  $P = 0.005$ ) between genetic and phenotypic distance matrices suggests that some similar evolutionary processes might have been involved in the

diversification of the measured phenotypic traits and neutral genetic markers. Flowering date, inflorescence weight, and plant height accurately discriminated the three genetically differentiated groups among the collections, whereas flowering date was the sole discriminatory variable of the two genetically differentiated groups. Deschutes River collections exhibited a trend of difference for flowering date and plant height with John Day River collections, but not for other phenotypic traits (Table 3). Flowering date was not a significant variable for discriminating the John Day River collections from those in eastern Oregon, Washington, and Idaho. Flowering date is considered a life history trait (Crnokrak and Roff 1995), which has been shown to generally reflect neutral genetic diversity more than morphological traits (Merila and Crnokrak 2001). Additionally, flowering date also was the phenotypic character most highly correlated with environmental and topographic characteristics of the collection sites in this study. As a result, flowering date likely will play a substantial role in defining both genetic structure and local adaptation of western prairie clover.

## IMPLICATIONS

Western prairie clover is a nontoxic perennial legume native to the northern Great Basin, Snake River Basin, and southern Columbia Plateau in the western United States. Distribution and occurrence of western prairie clover in a variety of habitats makes it a potential candidate for use in rangeland revegetation programs. Based on genetic differences and phenotypic traits with possible adaptive significance, we recommend that separate seed sources (conservation units) be developed and locality tested from each of the two groups of collections described herein: one from the Deschutes River watershed and one from southeastern Washington, eastern Oregon, and Idaho. Further studies are needed to determine the placement of John Day River watershed collections into an optimal conservation unit.

## ACKNOWLEDGMENTS

We would like to thank Kevin Connors and Lisa Michaels for their help in both the field and laboratory. The authors also acknowledge graduate student grant funds awarded to Kishor Bhattarai from the Center for Integrated Biosystems, Utah State University.

## LITERATURE CITED

- ASSOCIATION OF OFFICIAL ANALYTICAL CHEMISTS. 1990. Official methods of analysis. Arlington, VA, USA: Association of Official Analytical Chemists. 1141 p.
- AYDIN, I., AND F. UZUN. 2005. Nitrogen and phosphorus fertilization of rangelands affects yield, forage quality and the botanical composition. *European Journal of Agronomy* 23:8–14.
- BARNEY, R. C. 1977. *Dalea imagines*. *Memoirs of the New York Botanical Garden* 27:1–892.
- BENHAM, J. J. 2001. Genographer—AFLP/Microsatellite Software. v. 1.6.0. Available at: <http://hordeum.oscs.montana.edu/genographer/>. Accessed 14 October 2007.
- BHATTARAI, K., D. A. JOHNSON, T. A. JONES, K. J. CONNORS, AND D. R. GARDNER. 2008. Physiological and morphological characterization of basalt milkvetch (*Astragalus filipes*): basis for plant improvement. *Rangeland Ecology and Management* 61:444–455.

- BROWN, C. S., V. J. ANDERSON, V. P. CLAASSEN, M. E. STANNARD, L. M. WILSON, S. Y. ATKINSON, J. E. BROMBERG, T. A. GRANT III, AND M. D. MUNIS. 2008. Restoration ecology and invasive plants in the semiarid West. *Invasive Plant Science Management* 1:399–413.
- BUSHMAN, B. S., S. L. LARSON, M. D. PEEL, AND M. E. PFRENDER. 2007. Population structure and genetic diversity in North American *Hedysarum boreale* Nutt. *Crop Science* 47:1281–1286.
- CHERNEY, J. H., AND V. G. ALLEN. 1995. Forages in a livestock system. In: R. F. Barnes, D. A. Miller, and C. J. Nelson [EDS.]. Forages: the science of grassland agriculture. Volume I. Ames, IA, USA: Iowa State University Press. p. 175–188.
- CRNOKRAK, P., AND D. A. ROFF. 1995. Dominance variance: associations with selection and fitness. *Heredity* 75:530–540.
- DICE, L. R. 1945. Measures of the amount of ecologic association between species. *Ecology* 26:297–302.
- DINNO, A. 2009. Exploring the sensitivity of Horn's parallel analysis to the distributional form of simulated data. *Multivariate Behavioral Research* 44:362–388.
- DIXON, P. 2003. Vegan, a package of R functions for community ecology. *Journal of Vegetation Science* 14:927–930.
- ENDLER, J. A. 1986. Natural selection in the wild. Princeton, NJ, USA: Princeton University Press. 358 p.
- EXCOFFIER, L., G. LAVAL, AND S. SCHNEIDER. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1:47–50.
- FALUSH, D., M. STEPHENS, AND J. K. PRITCHARD. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* 7:574–578.
- FELSENSTEIN, J. 2009. PHYLIP (Phylogeny Inference Package). Version 3.69. Seattle, WA, USA: University of Washington, Department of Genome Sciences (distributed by author).
- HORN, J. L. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika* 30:179–185.
- JONES, T. A. 2009. Conservation biology and plant breeding: special considerations for development of native plant materials for development of native plant materials for use in restoration. *Ecological Restoration* 27:8–11.
- JONES, T. A., S. R. LARSON, AND B. L. WILSON. 2008. Genetic differentiation and hybridization among *Festuca idahoensis*, *F. roemerii*, and *F. ovina* detected from AFLP, ITS, and chloroplast DNA. *Botany* 86:422–434.
- 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.
- MADISON, L. A., AND R. J. ROBEL. 2001. Energy characteristics and consumption of several seeds recommended for northern bobwhite food plantings. *Wildlife Society Bulletin* 29:1219–1227.
- MANTEL, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27:209–220.
- MARDIA, K. V., J. T. KENT, AND J. M. BIBBY. 1979. Multivariate analysis. New York, NY, USA: Academic. 518 p.
- MCGRAW, R. L., F. W. SHOCKLEY, J. F. THOMPSON, AND C. A. ROBERTS. 2004. Evaluation of native legume species for forage yield, quality, and seed production. *Native plants Journal* 5:152–159.
- McKAY, J. K., C. E. CHRISTIAN, S. P. HARRISON, AND K. J. RICE. 2005. "How local is local?" A review of practical and conceptual issues in the genetics of restoration. *Restoration Ecology* 1:432–440.
- McKAY, J. K., AND G. LATTA. 2002. Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution* 17:185–291.
- MERILA, J., AND P. CRNOKRAK. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology* 14:892–903.
- MERTENS, D. R. 2002. Gravimetric determination of amylase-treated neutral detergent fiber in feeds with refluxing in beakers or crucibles: collaborative study. *Journal of AOAC International* 85:1217–1240.
- MEUDT, H. M., AND A. C. CLARKE. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science* 12: 106–117.
- MILLER, D. A. 1984. Forage crops. New York, NY, USA: McGraw-Hill. 530 p.
- MOSCOW FORESTRY SCIENCES LABORATORY. 2009. Custom climate data requests. Available at: <http://forest.moscowfsl.wsu.edu/climate/customData/>. Assessed 6 October 2009.
- NACHLINGER, J., K. SOCHI, P. COMER, G. KITTEL, AND D. DORFMAN. 2001. Great Basin: an ecoregion-based conservation blueprint. Reno, NV, USA: The Nature Conservancy. 160 p.
- OUBORG, N. J., P. VERGEER, AND C. MIX. 2006. The rough edges of the conservation genetics paradigm for plants. *Journal of Ecology* 94:1233–1248.
- PRISTERER, A. B., J. JOSHI, B. SCHMID, AND M. FISCHER. 2004. Rapid decay of diversity–productivity relationships after invasion of experimental plant communities. *Basic and Applied Ecology* 5:5–14.
- PHILLIPS, N. C., S. R. LARSON, AND D. T. DROST. 2008. Detection of genetic variation in wild populations of three *Allium* species using amplified fragment length polymorphisms. *HortScience* 43:637–643.
- RAY, V. F., G. P. MURDOCK, B. BLYTH, O. C. STEWART, J. HARRIS, E. A. HOEBEL, AND D. B. SHIMKIN. 1938. Tribal distribution in eastern Oregon and adjacent regions. *American Anthropologists* 40:384–415.
- RICE, K. J., AND N. C. EMERY. 2003. Managing microevolution: restoration in the face of global change. *Frontiers in Ecology and the Environment* 1:469–478.
- RICHARDS, R. T., J. C. CHAMBERS, AND C. ROSS. 1998. Use of native plants on federal lands: policy and practice. *Journal of Range Management* 51:625–632.
- ROHLF, F. J. 1998. NTSYSpc numerical taxonomy and multivariate analysis system. Version 2.0. Setauket, NY, USA: Exeter Software. 31 p.
- SAS INSTITUTE. 2004. SAS/STAT 9.1 user's guide. Cary, NC, USA: SAS Institute, Inc. 5136 p.
- SCHLÖTTERER, C. 2004. The evolution of molecular markers—just a matter of fashion? *Nature Reviews Genetics* 5:63–69.
- ST. CLAIR, J. B., N. L. MANDEL, AND K. W. VANCE-BORLAND. 2005. Geneecology of Douglas fir in western Oregon and Washington. *Annals of Botany* 96:1199–1214.
- STEIN, B., L. S. KUTNER, AND J. S. ADAMS. 2000. Precious heritage: the status of biodiversity in the United States. New York, NY, USA: Oxford University Press. 399 p.
- [USDA, NRCS] US DEPARTMENT OF AGRICULTURE, NATURAL RESOURCES CONSERVATION SERVICE. 2009. The Plants Database. Baton Rouge, LA, USA: National Plant Data Center. Available at: <http://plants.usda.gov/java/profile?symbol=DALEA>. Accessed 30 November 2009.
- VAN DER WAL, R. S., C. VAN LIESHOUT, C. DORMANN, R. LANGVATN, AND S. D. ALBON. 2000. Trading forage quality for quantity? Plant phenology and patch choice by Svalbard reindeer. *Oecologia* 123:108–115.
- VEKEMANS, X., T. BEAUWENS, M. LEMAIER, AND I. ROLDAN-RUIZ. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology* 11:139–151.
- VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, AND M. KUIPER. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23:4407–4414.
- WALKER, S. C., AND N. L. SHAW. 2005. Current and potential use of broadleaf herbs for reestablishing native communities. In: N. L. Shaw, M. Pellant, and S. B. Monsen [EDS.]. Proceedings, Sage Grouse Habitat Restoration Symposium. Fort Collins, CO, USA: USDA Forest Service, Rocky Mountain Research Station. Proceedings RMRS-P-38. p. 46–61.
- WHITE, L. M., AND J. R. WIGHT. 1984. Forage yield and quality of dryland grasses and legumes. *Journal of Range Management* 37:233–236.
- WONG, A., M. R. FORBES, AND M. L. SMITH. 2001. Characterization of AFLP markers in damselflies: prevalence of codominant markers and implications for population genetic applications. *Genome* 44:677–684.