American Journal of Botany 91(6): 797–803. 2004.

CORRELATION BETWEEN MOLECULAR MARKERS AND ADAPTIVELY SIGNIFICANT GENETIC VARIATION IN *BROMUS TECTORUM* (POACEAE), AN INBREEDING ANNUAL GRASS[1]

ALISA P. RAMAKRISHNAN,[2,3] SUSAN E. MEYER,[2,4] JENNIFER WATERS,[3] MIKEL R. STEVENS,[3] CRAIG E. COLEMAN,[3] AND DANIEL J. FAIRBANKS[3]

[2]USDA Forest Service, Rocky Mountain Research Station, Shrub Sciences Laboratory, Provo, Utah 84606 USA; and
[3]Plant and Animal Sciences Department, Brigham Young University, Provo, Utah 84602 USA

Single sequence repeat (SSR) and amplified fragment length polymorphic (AFLP) molecular marker genotypes in cheatgrass (*Bromus tectorum*) were compared to published data on phenotypic variation in seed dormancy, vernalization requirement, and resistance to the pathogen *Ustilago bullata*. Several features of cheatgrass facilitated this study: it is a recent invader in the western United States, has considerable phenotypic polymorphism, and is an obligate self-pollinator. Forty self-pollinating lines from four populations common to the three phenotypic data sets were analyzed for molecular genetic variation using seven SSR loci and 31 AFLP loci. We examined correlations between distance matrices using the Mantel test for each pair of studies. The two molecular data sets were significantly correlated ($r = 0.636$). The AFLP markers often distinguished among several lines with identical SSR genotypes. The AFLP data were also significantly correlated with the phenotypic data (r values from 0.4640 to 0.5658), but the SSR data were much more highly correlated (r values from 0.677 to 0.844). The difference between molecular marker systems was especially notable when an outlier population from Potosi Pass, Nevada, was excluded from the analysis. These results suggest that SSR markers may be good surrogates for phenotypic traits in population genetic studies of strongly inbreeding species such as cheatgrass.

Key words: AFLP; *Bromus tectorum*; cheatgrass; genetic variation; microsatellite; molecular marker; phenotypic variation; SSR.

Phenotypic traits under genetic control often vary within and among populations of a species. Some phenotypic variation, such as pathogen resistance or vernalization response, can be difficult to measure in a population genetic context because of the very large experiments that would be required. Providing there is association with the trait in question, molecular genetic markers can be used as surrogates for phenotypes in population genetic studies (Virk et al., 1996; Pillen et al., 2000). Unfortunately, identifying linked molecular markers for genes governing phenotypic traits can be difficult in a chasmogamous plant species, and measures of population variation are dependent on the relatedness between the parents used for the mapping population (Ford-Lloyd et al., 2001). In addition, mapped markers do not exist for most non-crop species, and it can be impractical to use molecular markers for the study of adaptively significant variation on a population level in such species (Virk et al., 2000).

Cheatgrass, or downy brome (*Bromus tectorum* L.), is a recent invader in western North America (ca.1890, Mack, 1981). It is an obligate self-pollinator (McKone, 1985), an annual, and has high fecundity under greenhouse conditions. Because of these characteristics, as well as its negative ecological impact in the western United States, it has been the subject of genetic research at both phenotypic and molecular levels. Genetically controlled phenotypic variation has been measured experimentally in flowering phenology (Rice and Mack, 1991; Meyer et al., in press), seed dormancy (Meyer et al., 1997; Meyer and Allen, 1999), and resistance to head smut (*Ustilago bullata* Berk., Meyer et al., 2001). Most of the previous molecular work has dealt with population-level isozyme polymorphism (Novak and Mack, 2001). Because all traits, both molecular and phenotypic, are inherited as a unit in this self-pollinating species (i.e., linked, Golding and Strobeck, 1980), high correlations among phenotypic and molecular traits would be expected.

Isozyme polymorphism is quite low in cheatgrass, prompting the suggestion that the success of cheatgrass may be due largely to phenotypic plasticity (Novak et al., 1991). Isozyme loci are inherently less polymorphic than PCR (polymerase chain reaction)-based marker loci due in part to selection pressure on gene products (O'Hanlon et al., 2000); thus, isozyme loci may underestimate the genetic variation present within and among populations. To directly associate molecular marker variation with phenotypic variation in cheatgrass, highly polymorphic markers may be more useful than isozymes. Several polymorphic simple sequence repeat markers (SSRs or microsatellites) have recently been developed for cheatgrass (Ramakrishnan et al., 2002). SSRs are areas of DNA characterized by di-, tri-, or tetra-nucleotide repeats. They are among the most polymorphic PCR-based markers with rates of mutation around 10^{-3} mutations per generation per locus (Udapa and Baum, 2001; Thuillet et al., 2002). Development of SSR markers is expensive and time intensive, because each locus must be characterized individually. However, once developed, SSRs may be easily and rapidly assayed, especially if several loci can be amplified in one PCR reaction (multiplexed). The DNA used for SSRs does not have to be of high quality because the DNA fragments amplified are small. Also of benefit to population genetic studies is the codominant nature of SSRs, allowing direct detection of heterozygotes.

[1] Manuscript received 14 August 2003; revision accepted 10 February 2004.

The authors thank the United States Department of Agriculture; Cooperative State Research, Extension, and Education Service; National Research Initiative, Plant Pathology Program (Grant # 98-35303-6957) for support of this research.

[4] E-mail: smeyer@fs.fed.us.

TABLE 1. Information about the four populations of *Bromus tectorum* from Utah (UT) and Nevada (NV) used in this study (from Meyer and Allen, 1999).

Population	Relative population size	Elev. (m)	Climate	
			Mean annual precip. (cm)	Mean temp. January/July (°C)
Potosi Pass, NV (Mojave Desert)	Medium (ca. 500 ha)	1500	25.4	1.7/26.5
Whiterocks, UT (Intermountain Desert)	Very large (>100 000 ha)	1450	17.7	−2.3/25.8
Hobble Creek, UT (Intermountain Foothill)	Large (ca. 10 000 ha)	1530	40.2	−2.1/24.8
Strawberry, UT (Intermountain Montane)	Medium (ca. 500 ha)	2400	56.3	−7.8/16.1

Amplified fragment length polymorphism (AFLP) is another commonly used PCR-based marker system (Pejic et al., 1998; Maguire et al., 2002) developed by Vos et al. (1995). It uses variation in restriction sites and length polymorphisms to estimate genetic variation across an entire genome, and as a result, the markers are often quite polymorphic. It has the advantage of revealing variation in many loci at once. It requires little development time, but is somewhat expensive (O'Hanlon et al., 2000). A potential drawback to AFLP markers is that codominance is difficult to detect. They also require highly purified DNA to prevent incomplete digestion by restriction enzymes.

In this study, we explore the relationships among different genetic traits in cheatgrass. First, we compare information content of AFLP markers and previously characterized SSRs. Second, we examine the correlation between molecular markers and genetically based phenotypic variation in seed dormancy, vernalization requirement, and resistance to *Ustilago bullata*. Our goal was to determine whether the patterns of phenotypic variation we observed in the earlier studies were the manifestation of a relatively small number of genotypes (inbreeding lines) represented by multiple individuals that could be characterized by their molecular marker fingerprints.

MATERIALS AND METHODS

Experimental design—We based our design on three of our published studies of genetically based phenotypic variation in cheatgrass (Meyer and Allen, 1999; Meyer et al., 2001, in press). Six to 10 inbreeding lines from each of four populations were included in these studies. The parental individuals were originally collected from three populations in Utah (UT) and one in southern Nevada (NV) for the study described in Meyer and Allen (1999; Table 1). We analyzed these 40 lines for six SSR loci and 31 polymorphic AFLP markers.

AFLP study—We grew one plant of each inbreeding line from seed in a greenhouse, collected leaves from plants that were about 10 cm tall into microcentrifuge tubes and froze them at −80°C prior to extraction. To extract DNA for the AFLPs, we ground the leaves in liquid nitrogen and used the Dneasy Plant Mini Kit (Qiagen Inc., Valencia, California, USA) to complete the extraction. We performed all steps of the AFLP analysis essentially according to Vos et al. (1995), with minor variations as follows. Following digestion, we incubated samples for 15 min at 70°C to stop the digestion. We used a final concentration of 10 mM ATP in the ligation reaction, which was carried out at 22°C overnight, then diluted the restriction/ligation product 1 : 9 with TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). We preamplified the DNA and diluted it 1 : 30 with TE before using it in the selective amplification. We used the following selective bases on the standard Vos et al. (1995) primer combinations: E-AA/M-CA, E-AA/M-CC, E-AT/M-CT, E-AT/M-CC, E-AT/M-CG, E-AC/M-CA, E-AC/M-CT, E-AC/M-CC, E-AC/M-CG, E-AG/M-CA, E-AG/M-CT, E-AG/M-CC, E-AG/M-CG.

We electrophoresed samples on a 5% denaturing polyacrylamide gel and visualized them using Kodak Imaging Screen K (Bio-Rad, Hercules, California, USA), which was exposed to the dried gel 6–12 h and visualized using the Molecular Imager FX (Bio-Rad). We scored the resulting images for ± amplification using the AFLP-Quantar 1.0 (Key Gene, Netherlands) analysis system.

SSR study—Sample tissue was collected as described earlier. We extracted DNA for the SSRs using a modified CTAB (cetyltrimethylammonium bromide) method (Bult et al., 1992). We then amplified the previously described SSR loci BT03, BT04, BT05, BT12, BT26, BT30, and BT33 with fluorescently labeled primers (Ramakrishnan et al., 2002). BT05, BT26, BT30 and BT33 amplified in a single PCR. BT03, BT04 and BT12 were amplified in separate reactions. We visualized PCR products on a Perkin-Elmer ABI 377 automated DNA sequencer (PE Applied Biosystems, Foster City, California, USA), run by the Brigham Young University DNA Sequencing Center, and analyzed the data using Genotyper (PE Applied Biosystems).

Phenotypic studies—Meyer and Allen (1999) measured variation in levels of primary dormancy of recently harvested seeds under five temperature regimes. Germination percentage under each regime was considered as a separate trait. The study included 10 randomly selected inbreeding lines from each of six populations. To compare phenotypic data with molecular data, we used the four populations common to all three phenotypic studies (Table 1).

Meyer et al. (in press) measured responses to vernalization treatment (chilling) as imbibed seeds and as seedlings, using progeny of the same 40 inbreeding lines as the previous study. Vernalization periods ranged from 0–14 weeks, but only the 0-, 2-, and 4-wk data are included in this analysis because response to longer vernalization treatments did not vary significantly. The response variables included proportion of plants flowering, weeks to initiation of flowering, and seed yield per plant. Magnitude of each response variable following each vernalization treatment was considered as a trait, for a total of 18 traits (seeds vs. seedlings, three vernalization periods, three response variables).

TABLE 2. Partitioning of variance within and among populations of *Bromus tectorum* from Utah and Nevada for two types of molecular marker data (SSR and AFLP), and variance accounted for by the first two coordinates from PCoA performed on each molecular marker data set. Percentage of total variance from AMOVA is indicated in parentheses.

Molecular marker	Among populations	Among lines within populations	PCoA-Variance percentage
SSR data (7 loci)	2.99 (83.3%)	0.60 (16.7%)	PCo1 = 55.8% PCo2 = 19.0%
AFLP data (31 loci)	4.15 (55.9%)	3.28 (44.2%)	PCo1 = 30.4% PCo2 = 16.1%

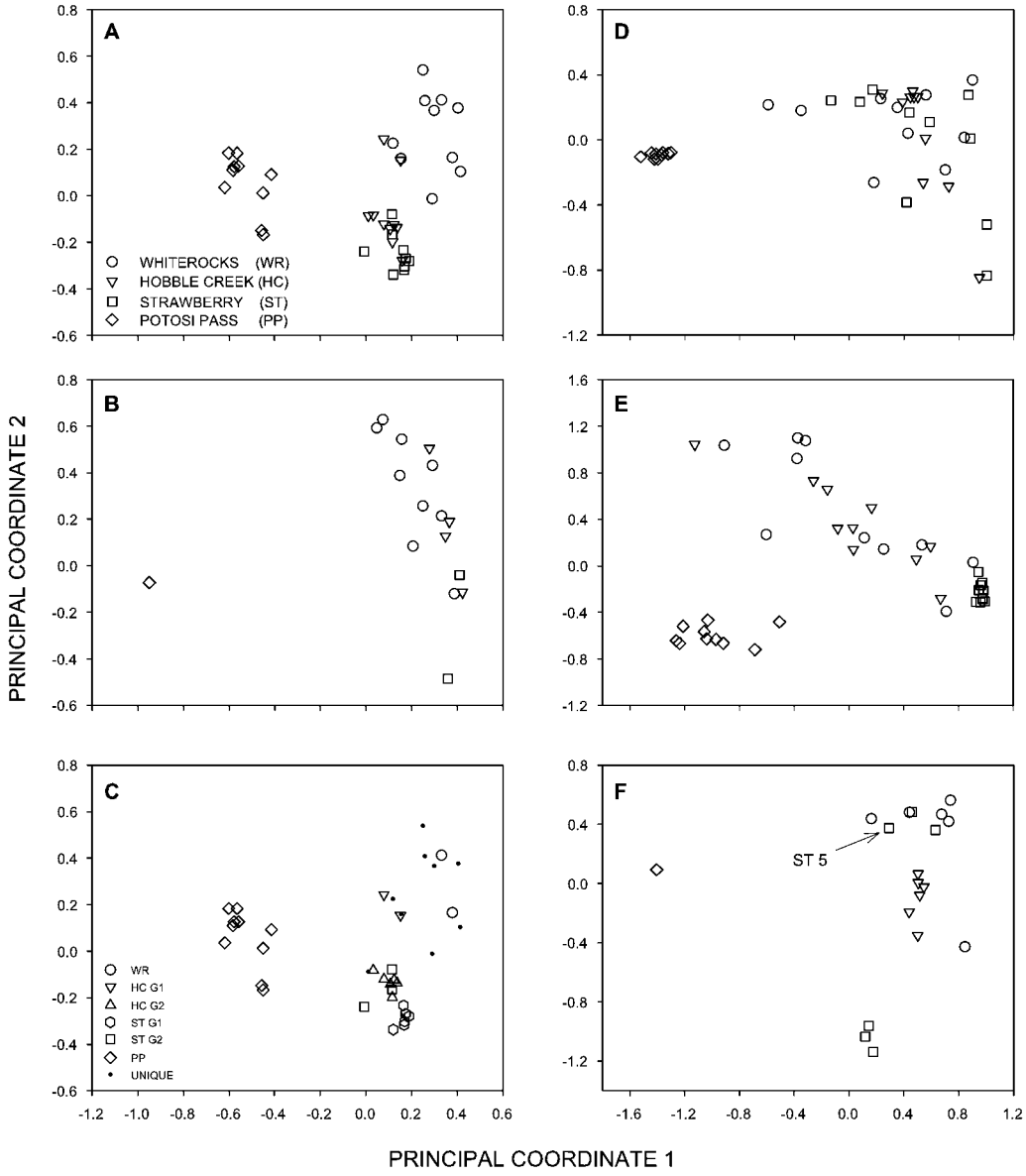

Fig. 1. Principal coordinates analyses of molecular and phenotypic data: (A) AFLP data (B) SSR data (C) AFLP data, with each symbol representing a different SSR genotype except for the unique genotypes, which all have the same symbol (D) vernalization data (E) seed dormancy data (F) smut resistance data.

Meyer et al. (2001) examined variation in resistance to the head smut pathogen *Ustilago bullata*. In this exploratory study, seeds from six of the 10 lines from each of the four populations in Table 1 were inoculated with bulk collections of *U. bullata* teliospores from both the co-occurring pathogen population and the other three populations. Infection percentage following inoculation with a specific pathogen population was considered as a trait, for a total of four traits. Plants were vernalized and allowed to bolt, and the percentage of smutted individuals was scored.

Data analysis—We used different similarity coefficients to construct distance matrices for each data set depending on the nature of the data. For the SSR data, we used a shared band similarity coefficient (Lynch, 1990) because we detected no null alleles. For the AFLP data, we used Dice's similarity coefficient (Dice, 1945; Nei and Li, 1979) because it weights band presence higher than band absence. For the phenotypic data, we re-analyzed the data from the original papers. We first standardized all values, then calculated the average taxonomic distance

$$\sum_{ij} = \sqrt{\sum_{k=1}^{n} \frac{1}{n}(x_{ki} - x_{kj})^2}$$

where i and j are means for individual lines, k is a value for a trait, and n is the number of traits for a comparison (Sneath and Sokal, 1973).

Using the distance matrices, we carried out principal coordinates analysis (PCoA) for each data set. Principal coordinates analysis is similar to principal components analyses, but it can extract scores from a broader range of data inputs such as a distance matrix (Quinn and Keough, 2002).

To test for significant correlation between pairs of distance matrices, we calculated a Mantel correlation statistic (Mantel, 1967). Because the Potosi Pass population is somewhat of an outlier in all data sets, we ran the Mantel test both with and without that population. This allows us to see how much correlation is present among populations that are more similar. Because the smut resistance data set has a subset of the lines from the other experiments, we compared resistance distance matrices to corresponding subsets of the distance matrices from the other data sets. The distance matrices, PCoAs, and

TABLE 3. Allele sizes at each of seven SSR loci for *Bromus tectorum* lines included in the study. The inbreeding lines are labeled by number and by population (WR = Whiterocks, Utah, HC = Hobble Creek, Utah, ST = Strawberry, Utah, and PP = Potosi Pass, Nevada).

Inbreeding line	SSR locus						
	BT03	BT04	BT05	BT12	BT26	BT30	BT33
WR 1	114	110	167	263	158	114	221
WR 2,8	114	110	175	263	156	114	221
WR 3	114	110	173	263	152	114	221
WR 4	114	114	169	252	156	114	221
WR 5	114	114	170	252	152	114	221
WR 6	114	114	169	263	156	114	221
WR 7	114	114	169	263	152	114	224
WR 9	114	110	169	263	156	114	221
WR 10	114	114	167	263	152	114	221
HC 1,2,3,6,8,9	114	114	173	252	152	114	221
HC 4,5	114	110	169	263	152	114	221
HC 7	114	110	173	252	152	114	221
HC 10	114	114	173	263	152	114	221
ST 1,3,4,5,8,9,10	114	114	178	252	152	116	221
ST 2,6,7	114	114	169	252	152	114	221
PP 1–10	112	108	171	261	156	118	223

Mantel tests were computed using NTSYSpc (Exeter Software, Setauket, New York, USA).

We graphed results from the PCoAs by plotting the scores on the first two principal coordinates and using a different symbol for each population. To better illustrate correlation between the AFLP and SSR data sets, we drew a second graph of the AFLP PCoA using a different symbol for each SSR genotype.

For all molecular data, we conducted analyses of molecular variance (AMOVAs, Excoffier et al., 1992) using Arlequin (Schneider et al., 2000), allowing 20% missing data for the AFLP AMOVA in order to exclude as few data as possible. We also calculated gene diversity values for each locus estimated as $(1 - \Sigma p_i^2)(2n/2n - 1)$ where n is the number of individuals in the population and p_i is the frequency of the i^{th} allele in the population (Nei, 1987). A value of 1.00 at a locus indicates that every individual has a different allele, while a value of zero indicates that the locus is fixed for a single allele, i.e., the same allele is present in every individual. Because all inbreeding lines were homozygous at every locus, we used one allele per locus per individual in our calculations and n instead of $2n$. We averaged the gene diversity over loci within populations and also calculated total gene diversity, treating the whole data set as one population. This enabled us to compare diversity present in each population to overall genetic diversity.

For the phenotypic data, we computed analyses of variance (ANOVAs) using SAS (version 8, Cary, North Carolina, USA) to determine the distribution of variance among components. Because of replication in the phenotypic data sets, we were able to calculate a variance component for within-line (non-genetic) variation as well as among populations and among lines within populations. Gene diversity statistics are not calculated for the phenotypic data because the molecular nature of the data is unknown.

RESULTS

Molecular markers—AFLP analysis yielded 1200 total bands and 31 polymorphic markers (2.3% polymorphism). The range of marker weight was 65–200 kb. Nine of the 31 polymorphic markers were fixed for either presence or absence in at least two populations. The percentage of total variance explained by within-population variance was 44%, with 56% of the total variance among populations (Table 2). The first two coordinates from the PCoA account for 46.5% of the total variation. Scores on these two PCo's group Potosi Pass separately from the other populations (Fig. 1A). Among the other three populations, Whiterocks and Strawberry are clearly separated, while Hobble Creek overlaps both Strawberry and Whiterocks.

TABLE 4. Gene diversity values for AFLPs and SSRs averaged across loci, calculated both for each population of *Bromus tectorum* from Utah (UT) and Nevada (NV) and for the whole data set pooling populations. A value of zero indicates fixation, i.e., all individuals have the same allele, while a value of one indicates that each individual has a different allele. Range of values for individual loci for each population are given in parentheses.

Population	SSRs	AFLPs
Whiterocks UT	0.368 (0–0.8222)	0.359 (0–0.571)
Hobble Creek UT	0.184 (0–0.4667)	0.287 (0–0.536)
Strawberry UT	0.133 (0–0.4667)	0.149 (0–0.467)
Potosi Pass NV	0	0.129 (0–0.500)
Total	0.8088 (0.7692–0.8282)	0.415 (0.229–0.517)

Of nine SSR markers originally described by Ramakrishnan et al. (2002), two were monomorphic (about 77% polymorphism). In the seven remaining loci, 24 total alleles were amplified, for an average of 3.43 alleles per locus. Six of the seven loci were fixed for different alleles in at least two populations (Table 3). Probably due in part to this fixation, 83% of the variance was partitioned among populations and 17% was partitioned within populations (Table 2). The variance percentages are somewhat misleading when compared to the AFLP data, as among-population variance components were similar in absolute magnitude for both markers, whereas the within-population component was much higher for AFLPs.

In the SSR PCoA, the first two coordinates account for 62.5% of the variation (Table 2 and Fig. 1B). The first coordinate separates Potosi Pass, NV from the other populations, largely because it has unique alleles at six loci (Table 3). As in the AFLP data, Whiterocks shows more overlap with Hobble Creek than with Strawberry.

When we label SSR genotypes on the AFLP PCoA (Fig. 1C), it becomes clear why the PCoA scattergrams for AFLPs and SSRs are similar. For each SSR genotype, there is often a cloud of AFLP genotypes. Though the AFLP data differentiates among more genotypes than the SSR data, similar AFLP genotypes often have the same SSR genotype.

The distribution of SSR genotypes among populations had an interesting pattern. All 10 inbreeding lines from the Potosi Pass population shared a common SSR genotype with unique alleles at six of seven loci, while the 10 inbreeding lines from Strawberry belonged to one of two SSR genotypes (Table 3). Six of the 10 inbreeding lines at Hobble Creek belonged to a single SSR genotype, while the remaining four lines represented three genotypes. The 10 Whiterocks inbreeding lines represented nine SSR genotypes. In this data set, none of the 16 SSR genotypes were found in more than one population.

The SSR data had higher total gene diversity than the AFLP data (Table 4). This can be attributed to the presence of multiple alleles in SSR data vs. two possible states in AFLP data. Populations followed the same rank order in terms of within-population mean gene diversity for both marker types, with Whiterocks showing the highest values and Potosi Pass showing the lowest. The gene diversity value for Potosi Pass based on SSRs was zero because all 10 individuals had identical genotypes. Mean gene diversity calculated on the pooled data set (all four populations combined) was higher than gene di-

TABLE 5. Partitioning of variance among populations, among lines within populations and within lines from ANOVA for three *Bromus tectorum* genecological experiments: seed dormancy regulation (Meyer and Allen, 1999), head smut resistance (Meyer et al., 2001) and vernalization requirement (Meyer et al., in press). Mean percentage of variance accounted for and range of variance percentage accounted for are given for each suite of traits. PCoA variance percentages are also included for the first two coordinates, used in Fig. 1.

Phenotypic trait	Among populations	Among lines within populations	Within lines	PCoA Variance percentage
Seed dormancy (5 regimes)	96.0 (93–99)	3.1 (0.7–5.4)	0.9 (0.2–1.3)	PCo1 = 65.0% PCo2 = 29.5%
Smut resistance (4 traits)	95.2 (86–98)	4.4 (1–13)	0.4 (<1)	PCo1 = 64.8% PCo2 = 30.2%
Vernalization (6 regimes)				PCo1 = 76.6% PCo2 = 8.5%
Proportion flowering	86.3 (47–99)	8.3 (0–26)	5.4 (0–26)	
Weeks to flowering	98.6 (97–100)	0.9 (0–2)	0.5 (<1)	
Seed yield/plant	95.0 (87–99)	3.4 (1–9)	1.6 (1–7)	

versity for any single population, due to the fixation of many loci for single alleles within populations.

Phenotypic traits—All traits in the vernalization experiment had over 86% of the variance accounted for among populations (Table 5). The proportion of flowering plants had the highest among-line variance proportion at 8.3%. All traits had some within-line variance, with the proportion of flowering plants having the highest variance at 5.4%. The first two coordinates on the PCoA accounted for 85% of the variation, with 77% for the first coordinate (Table 5). Potosi Pass is again very distinct, while the other populations are less so (Fig. 1D). Because so much of the variation is accounted for by the first coordinate, it is more important to observe scores for the first coordinate than for the second when making comparisons among lines.

Most of the variance in the seed dormancy data set is partitioned among populations, with only 3.1% partitioned among lines within populations (Table 5). There was some non-genetic variation, indicated by the within-line variance (0.9% of the total). The first two coordinates on the PCoA accounted for 95% of the variation (Table 5). Potosi Pass and Strawberry are separated from the other populations, but not much differentiation is evident between Hobble Creek and Whiterocks (Fig. 1E).

Similar to the dormancy data, the *U. bullata* resistance data display high variance among populations and low within-population variance (Table 5). There was some variance within lines. The first two coordinates on the PCoA account for 95% of the variation. Potosi Pass is again grouped apart from the other populations (Fig. 1F). Strawberry is separated into two distinct groups, one grouping with Whiterocks and one being unique. Hobble Creek lines are all grouped together, separate from the other populations. There are fewer data points in the resistance data set because a subset of lines from the other studies was included (Meyer et al., 2001).

Mantel correlations—All distance matrices calculated using data from all four populations were significantly correlated with each other ($P < 0.0001$, Table 6). The correlation between SSRs and head smut resistance had the highest r value, and the correlation between AFLPs and dormancy had the lowest, though it was still highly significant. Among Utah populations alone, most matrix pairs were still significantly correlated ($P < 0.05$) with the exception of AFLP with vernalization, AFLP with resistance, and the correlations among the phenotypic data. The highest correlation among Utah populations was between SSRs and AFLPs ($P < 0.0001$). When SSR genotypes are taken into account and graphed on the AFLP PCoA, it is clear that for a given SSR genotype, there are often several associated AFLP genotypes (Fig. 1C). This helps explain the high Mantel r values, even though there is more scatter in the AFLP graph than the SSR graph.

DISCUSSION

Most studies comparing AFLPs with SSRs deal with outcrossing organisms in either crop species or native populations (Pejic et al., 1998; Mariette et al., 2001; Maguire et al., 2002). Introduced species often have lower levels of variation than native species (Hamrick et al., 1979), and self-pollinators often have lower levels of variation than outcrossing species (Hamrick et al., 1991). Cheatgrass is an obligate self-pollinator, so gene flow is restricted to seed dispersal, and polymorphism is low, as evidenced by the 2.3% polymorphic bands in the AFLP data (also Novak et al., 1991).

Despite the low levels of polymorphism, there are several differences between the SSR and AFLP data. The SSR total gene diversity is higher than that calculated with AFLPs, but AFLPs detected diversity in Potosi Pass where SSRs detected none (Table 4). One reason the SSR total gene diversity is higher involves differences in genomic sampling: AFLPs are treated as biallelic loci, whereas SSRs are usually multiallelic. Because of this, estimates of diversity are not always well correlated and are highly dependent on the number of loci used, population sizes, and migration patterns (Mariette et al., 2002).

Though the AFLP markers often distinguished several genotypes within a single SSR genotype, the distance matrices for the two types of markers were highly correlated. This correlation can be attributed to the observation that although there is less variation in the SSR data, it rarely contradicts the AFLP data. There are instances in which two lines with very similar AFLP genotypes have different SSR genotypes, but overall, correlation between the two types of markers is high.

It is not prudent at this time to directly compare diversity measures from our data set to studies with other species, because such studies usually sample a larger number of individuals in a population (Lerceteau and Szmidt, 1999; Teulat et al., 2000; Green et al., 2001). We were unable to use more samples in this particular study due to the difficulty of measuring the phenotypic traits in question on a large scale. Stein-

TABLE 6. Mantel correlation values (r), with P values shown in parentheses. The lower diagonal is the correlation between complete data sets, and the upper diagonal is the correlation between measurements of variation excluding Potosi Pass, Nevada from the calculations.

	SSR	AFLP	Dormancy	Vernalization	Smut resistance
SSR		0.4636 (0.0001)	0.3565 (0.0004)	0.2009 (0.029)	0.2687 (0.0233)
AFLP	0.6365 (0.0001)		0.2300 (0.0162)	0.1648 (0.0926)	−0.095[1] (0.732)
Dormancy	0.6770 (0.0001)	0.4640 (0.0001)		−0.1229[1] (0.948)	0.0920[1] (0.143)
Vernalization	0.8327 (0.0001)	0.5197 (0.0001)	0.5325 (0.0001)		0.0368[1] (0.3518)
Smut resistance	0.8437 (0.0001)	0.5658 (0.0001)	0.6303 (0.0001)	0.7837 (0.0001)	

[1] Not significant at the $P < 0.1$ level.

ger et al. (2002) used a similar sampling scale when they compared various morphological traits to AFLP data. The correlations evident in our study may assist evolutionary interpretation of a larger survey of molecular genetic variation within and among populations of cheatgrass.

There are practical considerations to take into account when deciding which marker system to use in a comprehensive population-level analysis. The DNA for SSR markers does not have to be of high purity, as it does for AFLP markers, so analyzing field-collected tissue is easier. However, because AFLPs amplify more loci in a single reaction and have higher within-population diversity, they may also be a good choice for studying population-level diversity.

High correlations between isozymes and morphological traits have been observed in other obligately self-pollinating species (reviewed in Hamrick, 1989; Knapp and Rice, 1998). This association between neutral markers and genetically controlled phenotypic traits is often termed hitchhiking (Hedrick and Holden, 1979). In our data, the strong and consistent separation of the warm desert Potosi Pass population from the other three populations regardless of the measure used is the best evidence for hitchhiking in cheatgrass, but it is also apparent within and among the other populations.

Major differences in the levels of correlation depend on the molecular marker: the SSRs have higher Mantel r values than the AFLPs when compared with the phenotypic data. This may be attributed in part to the fact that the SSR variance structure is more similar to the phenotypic data than the AFLP data, with most of the variance distributed among populations. The low within-population variation of the SSR data (as seen in the PCoA) may allow better correlation with less precise data (such as phenotypic data), especially if both data sets have low within-population variation. The AFLP data set has such high within-population variation that it is more likely to conflict with other data sets that display most of their variation among populations, especially if those data sets contain some level of non-genetic information, as is possible in phenotypic data sets. In the data sets we used, it appears that the more similar the scale of variation, the higher the correlation values. Because the AFLPs measure genetic variation on a finer scale than the phenotypic data, the correlation values decrease.

In one striking case in the smut resistance data, some inbreeding lines were different phenotypically, yet had very similar genotypes. In this data set, lines are well clustered and have high correlation with SSR data. An exception to this is ST05, labeled on the resistance PCoA (Fig. 1F). This line is the same genotype as the three Strawberry lines toward the bottom of the graph, but it has a different resistance phenotype. Such a discrepancy is more difficult to explain when dealing with an obligate self-pollinator than when outcrossing plays a large role in generation and maintenance of genetic variation. One possibility is that smut resistance mutates at a higher rate than either of the marker systems we used (Botella et al., 1998). This would allow high similarities between lines as measured with molecular markers, while the resistance data set could show less similarity.

The molecular genetic variation we observed supports earlier conclusions based on phenotypic and isozyme data that there is significant genetic variation among cheatgrass populations in the western United States (Novak et al., 1991; Rice and Mack, 1991; Meyer et al., 1997, 2001, in press; Meyer and Allen, 1999). The high correlation between molecular markers and phenotypic traits is the best evidence yet that cheatgrass populations in the western United States are composed of many pre-adapted genotypes rather than a few ''general purpose'' genotypes. The close concordance between SSR genotype and phenotypic traits measured in our studies lends credence to the idea that SSR marker genotypes could indeed be used as surrogates for measured phenotypic traits in large-scale population genetic studies in which measurement of phenotype on a plant by plant basis is not possible.

Novak and Mack (2001) found evidence for seven distinct introduction events in North America using isozyme data. The high genetic variation in our study indicates that there may have been additional introductions, and that introduction events may have consisted of many similar inbreeding lines that were not identical, but that had already become distinct from each other, at least in terms of their AFLP genotypes, in their native locations before introduction to the United States. Because of the correlated molecular and phenotypic trait variation observed in this study, we conclude that the current distribution of both phenotypic and molecular marker variation is probably largely the result of selection on genetic diversity present in original founder populations rather than on genetic diversity generated through mutation following an introduction event.

LITERATURE CITED

BOTELLA, M. A., J. E. PARKER, L. N. FROST, P. D. BITTNER-EDDY, J. L. BEYNON, M. J. DANIELS, E. B. HOLUB, AND J. D. G. JONES. 1998. Three genes of the Arabidopsis *RPP1* complex resistance locus recognize distinct *Peronospora parasitica* determinants. *Plant Cell* 10: 1847–1860.

BULT, C., M. KALLERSO, AND S. YOUNGBAE. 1992. Amplification and sequencing of 16/18S rDNA from gel-purified total plant DNA. *Plant Molecular Biology Reporter-ISPMB* 10: 273–284.

DICE, L. R. 1945. Measures of the amount of ecologic association between species. *Ecology* 26: 297–302.

EXCOFFIER, L., P. SMOUSE, AND J. QUATTRO. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.

FORD-LLOYD, B. V., H. J. NEWBURY, M. T. JACKSON, AND P. S. VIRK. 2001. Genetic basis for co-adaptive gene complexes in rice (*Oryza sativa* L.) landraces. *Heredity* 87: 530–536.

GOLDING, G. B., AND C. STROBECK. 1980. Linkage disequilibrium in a finite population that is partially selfing. *Genetics* 94: 777–789.

GREEN, J. M., J. H. A. BARKER, E. J. P. MARSHALL, R. J. FROUD-WILLIAMS, N. C. B. PETERS, G. M. ARNOLD, K. DAWSON, AND A. KARP. 2001. Microsatellite analysis of the inbreeding grass weed Barren Brome (*Anisantha sterilis*) reveals genetic diversity at the within- and between-farm scales. *Molecular Ecology* 10: 1035–1045.

HEDRICK, P. W., AND L. HOLDEN. 1979. Hitch-hiking: an alternative to coadaptation for the barley and slender wild oat examples. *Heredity* 43: 79–86.

HAMRICK, J. L. 1989. Isozymes and the analysis of genetic structure in plant populations. *In* D. E. Soltis and P. S. Soltis [eds.], Isozymes in plant biology, vol. 4, 87–105. Dioscorides Press, Portland, Oregon, USA.

HAMRICK, J. L., M. J. W. GODT, D. A. MURAWSKI, AND M. D. LOVELESS. 1991. Correlations between species traits and allozyme diversity: implications for conservation biology. *In* D. A. Falk and K. E. Holsinger [eds.], Genetics and conservation of rare plants. Oxford University Press, New York, New York, USA.

HAMRICK, J. L., Y. B. LINHART, AND J. B. MITTON. 1979. Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. *Annual Review of Ecology and Systematics* 10: 173–200.

KNAPP, E. E., AND K. J. RICE. 1998. Comparison of isozymes and quantitative traits for evaluating patterns of genetic variation in purple needlegrass (*Nassella pulchra*). *Conservation Biology* 12: 1031–1041.

LERCETEAU, E., AND A. E. SZMIDT. 1999. Properties of AFLP markers in inheritance and genetic diversity studies of *Pinus sylvestris* L. *Heredity* 82: 252–260.

LYNCH, M. 1990. The similarity index and DNA fingerprinting. *Molecular Biology and Evolution* 7: 478–484.

MACK, R. N. 1981. Invasion of *Bromus tectorum* L. into western North America: an ecological chronicle. *Agro-Ecosystems* 7: 145–165.

MAGUIRE, T. L., R. PEAKALL, AND P. SAENGER. 2002. Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by AFLPs and SSRs. *Theoretical and Applied Genetics* 104: 388–398.

MANTEL, N. A. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27: 209–220.

MARIETTE, S., D. CHAGNE, C. LEZIER, P. PASTUSKA, A. RAFFIN, C. PLOMION, AND A. KREMER. 2001. Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity* 86: 469–479.

MARIETTE, S., V. LE CORRE, F. AUSTERLITZ, AND A. KREMER. 2002. Sampling within the genome for measuring within-population diversity: trade-offs between markers. *Molecular Ecology* 11: 1145–1156.

MCKONE, M. K. 1985. Reproductive biology of several bromegrasses (*Bromus*): breeding system, pattern of fruit maturation, and seed set. *American Journal of Botany* 72: 1334–1339.

MEYER, S. E., AND P. S. ALLEN. 1999. Ecological genetics of seed germination regulation in *Bromus tectorum* L. I. Phenotypic variance among and within populations. *Oecologia* 120: 27–34.

MEYER, S. E., P. S. ALLEN, AND J. BECKSTEAD. 1997. Seed germination regulation in *Bromus tectorum* (Poaceae) and its ecological significance. *Oikos* 78: 475–485.

MEYER, S. E., D. L. NELSON, AND S. CARLSON. In press. Ecological genetics of vernalization response in *Bromus tectorum* L. (Poaceae). *Annals of Botany*.

MEYER, S. E., D. L. NELSON, AND S. CLEMENT. 2001. Evidence for resistance polymorphism in the *Bromus tectorum–Ustilago bullata* pathosystem: implications for biocontrol. *Canadian Journal of Plant Pathology* 23: 19–27.

NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, New York, USA.

NEI, M., AND W. H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA* 76: 5269–5273.

NOVAK, S. J., AND R. N. MACK. 2001. Tracing plant introduction and spread: genetic evidence from *Bromus tectorum* (cheatgrass). *BioScience* 51: 114–122.

NOVAK, S. J., R. N. MACK, AND D. E. SOLTIS. 1991. Genetic variation in *Bromus tectorum* (Poaceae): population differentiation in its North American range. *American Journal Botany* 78: 1150–1161.

O'HANLON, P. C., R. PEAKALL, AND D. T. BRIESE. 2000. A review of new PCR-based genetic markers and their utility to weed ecology. *Weed Research* 40: 239–254.

PEJIC, I., P. AJMONE-MARSAN, M. MORGANTE, V. KOZUMPLICK, P. CASTIGLIONI, G. TARAMINO, AND M. MOTTO. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. *Theoretical and Applied Genetics* 97: 1248–1255.

PILLEN, K., A. BINDER, B. KREUZKAM, L. RAMSAY, R. WAUGH, J. FORSTER, AND J. LEON. 2000. Mapping new EMBL-derived barley microsatellites and their use in differentiating German barley cultivars. *Theoretical and Applied Genetics* 101: 652–660.

QUINN, G. P., AND M. J. KEOUGH. 2002. Experimental design and data analysis for biologists. Cambridge University Press, New York, New York, USA.

RAMAKRISHNAN, A. P., C. E. COLEMAN, S. E. MEYER, AND D. J. FAIRBANKS. 2002. Microsatellite markers for *Bromus tectorum* (cheatgrass). *Molecular Ecology Notes* 2: 22–23.

RICE, K. J., AND R. N. MACK. 1991. Ecological genetics of *Bromus tectorum* I. A hierarchical analysis of phenotypic variation. *Oecologia* 88: 77–83.

SCHNEIDER, S., D. ROESSLI, AND L. EXCOFFIER. 2000. Arlequin: a software for population genetics data analysis, version 2.0. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Geneva, Switzerland.

SNEATH, P. H. A., AND R. R. SOKAL. 1973. Numerical taxonomy: the principles and practice of numerical classification. W. H. Freeman, San Francisco, California, USA.

STEINGER, T., P. HALDIMANN, K. A. LEISS, AND H. MULLER-SCHARER. 2002. Does natural selection promote population divergence? A comparative analysis of population structure using amplified fragment length polymorphism markers and quantitative traits. *Molecular Ecology* 11: 2583–2590.

TEULAT, B., C. ALDAM, R. TREHIN, P. LEBRUN, J. H. A. BARKER, G. M. ARNOLD, A. KARP, L. BAUDOUIN, AND F. ROGNON. 2000. An analysis of genetic diversity in coconut (*Cocos nucifera*) populations from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. *Theoretical and Applied Genetics* 100: 764–771.

THUILLET, A., D. BRU, J. DAVID, P. ROUMET, S. SANTONI, P. SOURDILLE, AND T. BATAILLON. 2002. Direct estimation of mutation rate for 10 microsatellite loci in durum wheat, *Triticum turgidum* (L.) Thell. ssp. *durum* desf. *Molecular Biology and Evolution* 19: 122–125.

UDAPA, S. M., AND M. BAUM. 2001. High mutation rate and mutational bias at $(TAA)_n$ microsatellite loci in chickpea (*Cicer arietinum* L.). *Molecular Genetics and Genomics* 265: 1097–1103.

VIRK, P., B. V. FORD-LLOYD, M. T. JACKSON, AND H. J. NEWBURY. 1996. Predicting quantitative variation within rice germplasm using molecular markers. *Heredity* 76: 296–304.

VIRK, P., H. J. NEWBURY, M. T. JACKSON, AND B. V. FORD-LLOYD. 2000. Are mapped markers more useful for assessing genetic diversity? *Theoretical and Applied Genetics* 100: 607–613.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23: 4407–4414.