

TECHNICAL NOTES

U.S. Department of Agriculture

Natural Resources Conservation Service

TN-PLANT MATERIALS-64

August 2001

Review of Paper “Mode of reproduction and amplified fragment length polymorphism variation in purple needlegrass (*Nassella pulchra*): utilization of natural germplasm sources”

Attached is a paper titled “Mode of reproduction and amplified fragment length polymorphism variation in purple needlegrass (*Nassella pulchra*): utilization of natural germplasm sources”. The paper was published in the British Journal of Molecular Ecology (2001) 10, 1165-1177. This paper was an USDA team effort involving the Agricultural Research Service, Natural Resources Conservation Service - Lockeford Plant Materials Center and San Francisco Urban Office.

This paper provides a source of information and background for personnel who are providing restoration and revegetation alternatives to landowners. It gives guidance on the genetic diversity of purple needlegrass and the result of distance from the seed source to the planting site.

Prepared by David A. Dyer, Manager, Lockeford Plant Materials Center,
Lockeford, California

CA-64-1

Mode of reproduction and amplified fragment length polymorphism variation in purple needlegrass (*Nassella pulchra*): utilization of natural germplasm sources

S. R. LARSON,* E. CARTIER,† C. L. McCracken* and D. DYER‡

*USDA -Agriculture Research Service, Forage and Range Research Laboratory, Utah State University, Logan, UT 84322-6300, USA,

†USDA-Natural Resource Conservation Service (NRCS), Livermore, California, USA, ‡USDA-NRCS, Lockeford Plant Materials Center, Lockeford, California, USA

Abstract

A dominant plant of the California grasslands, purple needlegrass [*Nassella pulchra* (Hitchc.) Barkworth] is an important revegetation species in its native range. The amplified fragment length polymorphism (AFLP) method was used to elucidate mode of reproduction and nucleotide variation among 11 natural populations and three selected natural germplasm releases of *N. pulchra*. A total of 12 co-dominant AFLPs, informative within eight populations, failed to reveal any heterozygous individuals, indicating very high selfing rates ($\bar{S}_H = 1$). Estimates of nucleotide diversity within populations ranged from 0 to 0.00069 (0.00035 average), whereas the total nucleotide divergence among populations ranged from 0.00107 to 0.00382 (0.00247 average). Measures of population differentiation (G_S) in terms of Shannon–Weaver diversity values and estimated nucleotide substitutions were 0.90 and 0.86, respectively. Although some of the sample populations contained a mixture of true breeding genotypes, most populations could be distinguished unambiguously. Moreover, geographical distance between the natural source populations was significantly correlated with genetic distance ($r = 0.60$) among the corresponding sample populations. Results indicate that inbreeding, combined with founder effects and/or selection, has contributed to the differentiation of *N. pulchra* populations. Foundation seed populations of the selected natural germplasm releases were genetically well defined and most similar to natural seed collected near the corresponding source populations. Thus, these commercial germplasm sources will be made practically available and useful for conservation plantings within the intended areas of utilization.

Keywords: amplified fragment length polymorphism, *Nassella pulchra*, self-fertilization

Received 17 August 2000; revision received 30 November 2000; accepted 11 December 2000

Introduction

Grasses are the most commonly used plant materials for primary soil stabilization and weed exclusion in many large- and small-scale revegetation efforts. Grasses used for revegetation have profound impacts evident by the ecological and agricultural importance of introduced species, such as crested wheatgrass (*Agropyron* spp.), that were successfully used over the past century to stabilize soils and productivity on highly degraded rangelands and abandoned croplands of western North America. The

release of commercial germplasm sources of native grass species, which will also be used in large-scale revegetation efforts, may have similar ecological and agricultural significance. Ecologists are concerned that the agronomic approach typically used for the production of these plant materials has been to select or breed cultivars with specific characteristics, which reduces genetic diversity (Roundy *et al.* 1997; Lesica & Allendorf 1999; Roundy 1999). Thus, many commercial sources of native plant materials may lack the genetic diversity needed to maintain adaptation in a natural environment. Moreover, nonlocal seed sources may not have adapted genomes and use of these plant materials may result in genetic pollution of natural populations (Roundy *et al.* 1997; Lesica & Allendorf 1999;

Correspondence: Steve Larson. Fax: 1 435 797 3075; E-mail: stlarson@cc.usu.edu

Roundy 1999). As the demand for native grass seed has grown, so has the demand for source-identified germplasm thought to represent naturally diverse and/or well-adapted ecotypes suitable for different sites or niches. Yet one of the most serious impediments to the implementation of native plant materials, particularly in many large-scale applications, has been seed availability (Richards *et al.* 1998). Although commercial growers can produce large quantities of high-quality grass seed, the feasibility of producing readily available, source-identified seed for many different sites is another important consideration. Locally collected seed or propagates from natural populations may be the best approach for many plant species. However, it is inherently difficult to anticipate the amount of grass seed needed in many large-scale fire rehabilitation efforts, much less collect or increase this seed from the locally native populations.

Believed to have been a dominant plant of pristine California grasslands, *Nassella pulchra* (Hitchc.) Barkworth is still the most common native bunchgrass in northern California and is also widespread in southern California (Bartolome 1981; Bartolome & Gemmill 1981; Heady *et al.* 1991). Closely related species, *N. lepida* (Hitchc.) Barkworth and *N. cernua* (Stebbins & Love) Barkworth, share similar distributions (primarily within California) and occasionally hybridize with *N. pulchra* (Love 1954). Evidence suggests that *N. pulchra* ($2n = 64$) contains a full complement of *N. lepida* ($2n = 34$) chromosomes, however, *N. cernua* ($2n = 70$) contains only 14 of the 17 chromosome pairs in *N. lepida* (Love 1954). Although cytogenetic relationships are not fully understood, these species behave as normal diploids and are thought to be allopolyploid (Stebbins & Love 1941; Love 1954). Like many Stipeae species, *N. pulchra* produces copious quantities of large viable seeds with sharp calli and twisting awns. These traits enable *N. pulchra* to rapidly colonize or increase abundance on many disturbed sites (Bartolome 1981; Bartolome & Gemmill 1981). These characteristics also make *N. pulchra* an important restoration tool within its native California range (Bartolome 1981; Dyer *et al.* 1996).

Quantitative genetic variation in morphology and phenology exist among populations of *N. pulchra*, which evidently reflects adaptation to different environments (Huntsinger *et al.* 1996; Knapp & Rice 1998). In a common garden evaluation of 10 diverse natural populations of *N. pulchra*, quantitative genetic variation in plant height, culm length, per cent heading, leaf length and width, seeds/culm, seed weight and awn weight was found to be associated with climatic differences and geographical distance (Knapp & Rice 1998). Based on these results, it was suggested that use of nonlocal seed sources, translocated across different regions of California, could result in poorly adapted plantings that may ultimately fail (Knapp & Rice 1998). Moreover, patterns of isozyme variability among

these 10 natural populations were correlated only with geographical distance and have limited usefulness in predicting the spatial scale over which *N. pulchra* seed should be translocated (Knapp & Rice 1998). Because measurements of quantitative trait variation are time- and labour-intensive, it was suggested that restorationists should use climatic information in making rapid decisions about appropriate spatial scales for translocating native grasses (Knapp & Rice 1998). The United States Department of Agriculture, Natural Resource Conservation Service, Plant Materials Center (PMC) at Lockeford, California has developed three natural germplasm sources of *N. pulchra* (e.g. LK115d, LK215e and LK315d). The identity of these natural *N. pulchra* germplasm sources reflect the Lockeford PMC as the releasing agency and the Major Land Resource Area (MLRA 15) for which these releases are best suited. Although MLRA boundaries were not intended to define seed transfer zones, these areas are characterized by particular patterns of soil type, climate, water resources and land use (USDA-SCS 1981). Approximately 50% of MLRA 15, Central California Coast Range, is native grass, shrub and brush rangeland. Another 35% is open woodland, also used for grazing. If plant cover is removed from the soils by fire, overgrazing, cultivation or logging, the hazard of erosion in this area is severe because of steep slopes and high-intensity rainfall (USDA-SCS 1981). Precipitation (typically 375–750 mm) in this area usually occurs during fall, winter or spring, but rarely in the summer. These natural germplasm sources (i.e. LK115d, LK215e and LK315d) were selected on the basis of flowering date, vigour, survival and growth from a replicated, common-garden evaluation of 32 accessions and released under Selected Class criteria of the Pre-Variety Germplasm Certification Standards (Association of Official Seed Certifying Agencies 1999; Young 1995). In order to protect the natural identity and diversity of these commercial germplasm sources, the foundation seed (used to establish commercial seed fields) was increased from seed collected directly from the original source populations. However, ecologists are also concerned that inadvertent directional selection may occur during cultivation (Knapp & Rice 1994). Moreover, determination of genetic identities for natural populations and commercial germplasm sources of native plant species is still a difficult problem of considerable importance to many public or private conservation groups and land managers.

The amplified fragment length polymorphism (AFLP) method (Vos *et al.* 1995) is a DNA fingerprinting technique that is highly informative in closely related germplasm (e.g. intraspecific) and can be used to estimate standard parameters of nucleotide sequence variation (π , D_{XY} and D_A) and differentiation among heterogeneous populations (G_S). The extent of DNA polymorphism within a heterogeneous population is often measured by nucleotide

Table 1 Identification of *Nassella pulchra* populations

County ID	(California)	Location	Latitude–Longitude
Tehama	Tehama	Red Bank Cr.	40°03'54"–122°36'00"
LK115d	Tehama	Hwy. 36 West of Bowman Rd.	40°17'00"–122°29'00"
Colusa	Colusa	Leesville	39°09'51"–122°23'39"
LK215e	Colusa	Walnut Valley Ranch, Lodoga	39°20'00"–122°30'00"
Calaveras	Calaveras	Hwy. 12 East of Evans Rd.	38°11'04"–120°50'34"
Contra Costa	Contra Costa	Mt. Diablo	37°52'21"–121°51'22"
Alameda	Alameda	Eden Canyon at I-580	37°42'00"–122°01'30"
LK315d	Alameda	Rancho Los Mochos BSA Camp	37°30'00"–121°32'00"
Marin	Marin	Lagunitas Dam	37°57'35"–122°36'36"
SF Yerba Buena	San Francisco	Yerba Buena Island	37°48'51"–122°21'41"
SF Bernal	San Francisco	Bernal Hill	37°44'36"–122°24'49"
SF St. Mary	San Francisco	St. Mary's Park	37°43'45"–122°25'25"
SF McLaren	San Francisco	John McLaren Park	37°43'08"–122°24'59"
SF Bayview	San Francisco	Bayview Hill	37°42'57"–122°23'47"
San Luis Obispo	San Luis Obispo	Waddell Ranch, Cayucos	35°27'04"–120°57'04"

diversity (π), which is defined as the average number of either nucleotide differences or substitutions per site for a group of DNA sequences (alleles) sampled (Nei & Li 1979; Nei & Miller 1990). Likewise, a standard measure of divergence between two heterogeneous populations is the net nucleotide divergence (D_A), where the average π within populations has been subtracted from D_{XY} , the total nucleotide divergence or average number of nucleotide differences per site between populations (Nei & Li 1979; Nei & Miller 1990). The apportionment of nucleotide sequence variation among (D_A/D_{XY}) populations is a standard measure of differentiation (G_S). Nei & Li (1979) developed a method for estimating π and D_{XY} using restriction endonucleases [restriction fragment length polymorphism (RFLP) method]. Variation within restriction/priming sites can also be used to estimate nucleotide sequence variation using the AFLP technique (Innan *et al.* 1999). However, a relatively small percentage (≈ 15 – 20%) of the AFLPs, analysed within genetic mapping populations of grasses, evidently arise from small insertion/deletions (indels) rather than nucleotide sequence variation at the restriction/priming sites (Becker *et al.* 1995; Bai *et al.* 1999; Larson *et al.* 2000b).

The primary objectives of this study were to determine mode of reproduction in *N. pulchra* and elucidate the genetic identities of LK115d, LK215e and LK315d relative to other natural germplasm sources collected within or near MLRA 15. Co-dominant AFLPs (i.e. indels) that vary within populations were identified and used to measure natural seedling rates (mode of reproduction). The remaining AFLP variation was used to estimate standard parameters of nucleotide sequence variation within and among populations. Differentiation among populations (G_S) was measured by the partitioning of AFLP variation in terms of Shannon–Weaver phenotypic diversity values (within and

among populations) and the proportion total nucleotide sequence variation (D_{XY}) partitioned among populations (D_A). The relationship between genetic distance and geographical distance was also investigated.

Materials and methods

Studies were initiated using foundation seed of LK115d, LK215e and LK315d and seed collected from each of 10 natural plants at 12 different sites (Table 1). Foundation seed of LK115d, LK215e and LK315d (lot numbered SFD-98-F2) was increased and certified under guidelines of the California Crop Improvement Association. Seed from the 12 natural populations was collected during April and May 1999. The spatial scale of these collection sites was approximately one ha or less. We attempted to germinate ≈ 150 seeds from each of the three foundation seed lots and 120 natural plants on moist blotter paper using a 2-week prechill at 5 °C, with the addition tetramethylthiuram disulfide as a fungicide. Seedlings were transplanted from germination boxes to individual cone soil containers maintained in an environmentally controlled greenhouse. Genomic DNA was extracted from ≈ 100 mg of young, green leaf tissue of each transplanted seedling, as described by Murray & Thompson (1980).

DNA fingerprinting was conducted using the AFLP technique according to the methods of Vos *et al.* (1995), except that *EcoRI* selective amplification primers were labelled with fluorescent 6-carboxy fluorescein (6-FAM) on the 5' nucleotide. The AFLP adapter sequences, pre-amplification primer sequences and selective amplification primer sequences are described in Table 2. The amplified fragments were fractionated and detected with an ABI373XL instrument (PE Applied Biosystems, Foster City, CA, USA) using 34 cm well-to-read polyacrylamide gels

Table 2 DNA sequences of amplified fragment length polymorphism (AFLP) primer and adapter sequences

Type	Name	Sequence
EcoRI adapter		5'-CTCGTAGACTGCGTACC-3'
		3'-CATCTGACGCATGGTTAA-5'
MseI adapter		5'-GACGATGAGTCCTGAG-3'
		3'-TACTCAGGACTCAT-5'
EcoRI + 1 primer	E01	5'-GACTGCGTACCAATTC-3'
EcoRI + 3 primers	E36	5'-GACTGCGTACCAATTCACC-3'
	E40	5'-GACTGCGTACCAATTCAGC-3'
	E41	5'-GACTGCGTACCAATTCAGG-3'
MseI + 1 primer	M02	5'-GATGAGTCCTGAGTAAC-3'
MseI + 3 primers	M47	5'-GATGAGTCCTGAGTAACAA-3'
	M48	5'-GATGAGTCCTGAGTAACAC-3'
	M49	5'-GATGAGTCCTGAGTAACAG-3'
	M50	5'-GATGAGTCCTGAGTAACAT-3'
	M60	5'-GATGAGTCCTGAGTAAC-3'
	M62	5'-GATGAGTCCTGAGTAAC-3'

formulated with 6.5% Long Ranger (FMC, Rockland, ME, USA), 7 M urea and 1× TBE running buffer. Each sample lane included the GS500-ROX (PE Applied Biosystems) internal lane size standards (labelled with fluorescent rhodamine X) ranging from 35 to 500 bp. Fluorescent signals of the 6-FAM-labelled AFLPs, between 35 and 500 bp, were identified using GENESCAN 3.1 software (PE Applied Biosystems). The GENESCAN sample (trace) files were subsequently analysed for the presence and absence of AFLP products, in ≈ 1 bp intervals, using Genographer (Benham *et al.* 1999). The degree of AFLP polymorphism and overall distribution of AFLP variation within and among populations was quantified using the Shannon–Weaver phenotypic diversity index as described for isozyme data (Chung *et al.* 1991; Dolan 1994; Knapp & Rice 1998) and AFLP data (Paul *et al.* 1997).

To help identify allelic AFLP fragments (co-dominant markers), phi correlation coefficients (Φ) were calculated among all pair-wise comparisons of all fragments over all primer pairs, within primer pairs and among comparisons of fragments differing by 3 bp (within each primer pair) using SAS Version 8 (SAS Institute Inc., Cary, NC, USA). Allelic fragments that might arise from indels would be expected to show phi coefficients (Φ) equal to -1 among homozygous genotypes or ≈ -0.3 among genotypes in Hardy–Weinberg equilibrium (in a two allele system). The genotypic scores for all fragments differing by 3 bp (within each primer pair), where $\Phi < 0$, were compared (Fig. 1) to identify co-dominant markers. These putative co-dominant markers were used to estimate selfing rates (\bar{S}_H), as described by Nordberg & Donnelly (1997) and Bergelson *et al.* (1998).

The computer program (Innan *et al.* 1999) used to estimate nucleotide variation based on F , the estimated

proportion of shared AFLP fragments between haploid genomes (with allelic fragments treated as shared characters), was kindly provided by Hideki Innan (University of Tokyo, Japan). For inbred genotypes, these values of F are the same as the average genetic similarity coefficients between individual plants, compared within or among populations, as described by Nei & Li (1979). This method of estimating nucleotide variation is based on the fact that each AFLP product represents a 16-bp sequence assay when using the *EcoRI* and *MseI* restriction enzymes, which have 6 and 4 bp recognition sequences, respectively, and three selective nucleotides on each of the two AFLP selective amplification primers. Therefore, each shared AFLP product indicates zero nucleotide differences over 16 bp, whereas polymorphisms reflect at least one nucleotide difference over 16 bp. The actual number of differences that contribute to each polymorphism is a function of F , which can be used to determine the overall number of nucleotide substitutions per site. These methods for estimating the number of nucleotide differences based on the proportion of shared AFLP products are good when GC content is near 0.5 and the when the number of differences is not too large (Innan *et al.* 1999). This assumption is tentatively acceptable for *Nassella pulchra* because the GC contents of wheat and maize, two other divergent Poaceae species, are very similar (45.5 and 46.0%, respectively) and close to 50% (Salinas *et al.* 1988; Montero *et al.* 1992). Allele frequencies, similarity coefficients (Nei & Li 1979), and proportions of shared fragments were calculated with SAS.

A UPGMA cluster analysis and scaled dendrogram were based on the proportion of shared bands (Nei & Li 1979) among individual genotypes (plants) developed using NTSYS-PC, version 2.02 (Exeter Software, Setauket, NY, USA). Bootstrap confidence levels (BCL) were recovered from the 50% majority-rule consensus tree of 1000 UPGMA searches (replicates), using PAUP* Version 4.0b4a (Sinauer Associates, Inc., Sunderland, MA, USA). Matrix correspondence between genetic similarity (i.e. proportion of shared fragments) and geographical distance or nucleotide divergences (D_{XY} or D_A) and geographical distances among populations was also tested (Mantel 1967) using NTSYS-PC.

Results

High germination rates and large vigorous seedlings were observed for the foundation seed lots of LK115d, LK215e and LK315d. However, some difficulties were encountered when germinating seed harvested from natural plants. The initial objective was to recover at least one seedling DNA sample from at least six natural plants at each of 12 collection sites (Table 1). However, only 86 of the 120 seed collections, harvested from 10 natural plants at each of 12 collection sites, produced at least one seedling. Only one

Table 3 Summary of amplified fragment length polymorphism (AFLP) genotyping for each *Nassella pulchra* population

Population	Number of plants	Average number of fragments per plant (range)	Number of different fragments	Polymorphic fragments (%)	Average Shannon–Weaver density (H) per AFLP primer pair
Tehama	6	406 (406)	406	0 (0.0)	0.00
LK115d	11	409.3 (408–410)	413	6 (1.5)	0.31
Calaveras	6	407.3 (407–408)	419	23 (5.5)	1.33
Alameda	6	408.5 (407–410)	420	23 (5.5)	1.17
LK315d	11	407.6 (406–410)	428	38 (8.9)	1.63
Colusa	6	404 (404)	404	0 (0.0)	0.00
LK215e	11	411 (411)	411	0 (0.0)	0.00
Contra Costa	6	407 (407)	424	32 (7.8)	0.40
SF St. Mary	6 (5*)	411 (411)	418	14 (3.3)	0.63
SF McLaren	5 (3*)	408 (407–409)	417	19 (4.6)	0.32
SF Bayview	6	406.8 (404–410)	419	22 (5.3)	1.29
SF Yerba Buena	6	412 (412)	412	0 (0.0)	0.00
Marin	6 (4*)	407.5 (406–410)	419	23 (5.5)	1.25
San Luis Obispo	6	403 (403)	403	0 (0.0)	0.00
Overall	98	407.6 (403–412)	513	178 (35.7)	5.77

*Actual number of natural plants sampled without replication among (identical) progeny genotypes.

fragments that varied within at least one population and 53 fragments that varied among, but not within, populations. The average Shannon–Weaver diversity (per AFLP primer pair) within populations (H_S) was 0.60, whereas the total AFLP diversity (H_T) was 5.77 (Table 3). Therefore, the proportion of total AFLP diversity among populations (G_S) was 0.90, which is considerably greater than the partitioning of *N. pulchra* isozyme variation ($G_S = 0.57$) using a similar Shannon–Weaver diversity index (Knapp & Rice 1998).

Small insertion/deletions (indels) were evident, particularly among comparisons of AFLPs 3 bp different (within primer pairs). The distributions of phi coefficients (Φ) among pair-wise comparisons of polymorphic AFLPs are almost identical when compared within and between primer pairs, except for the frequency of comparisons in which $\Phi = -1$ (Table 4). The percentage of comparisons in which $\Phi = -1$ is almost double within AFLP primer pairs

compared with between AFLP primer pairs (Table 4). Nearly half (10) of these comparisons, in which $\Phi = -1$, are among fragments 3 bp different (Table 4). Moreover, the percentage of comparisons in which $\Phi = -1$ is almost 30 times greater among comparisons of fragments 3 bp different (within AFLP primer pairs), than the number of comparisons in which $\Phi = -1$ between AFLP primer pairs (Table 4). Although 25% of the comparisons between AFLP primer pairs fall outside the 95% confidence interval (i.e. ≈ -0.20 to 0.20 with 98 plants) for zero association (Table 4), we do not recognize any molecular basis that would account for these correlations. We speculate that natural pedigree structure (relatedness within populations) effectively widens the 95% confidence interval for detecting significant associations between markers. In any case, we examined the possibility of indels among all pair-wise comparisons of fragments 3 bp different (within AFLP primer pairs) in which $\Phi < 0$ (Fig. 1). Although larger indels

Table 4 Distribution of phi coefficients (Φ) among pair-wise comparisons of polymorphic fragments as a measure of coupling phase associations (positive Φ) and repulsion phase associations (negative Φ)

Phi (Φ) coefficients	Fragments between AFLP primer pairs (%)	Fragments within AFLP primer pairs (%)	Fragments \leq 3 bp different within AFLP primer pairs (%)
= 1.00	85 (0.6)	9 (0.5)	0 (0.0)
0.80–1.00	51 (0.4)	10 (0.5)	1 (1.8)
0.60–0.80	141 (1.0)	14 (0.7)	0 (0.0)
0.40–0.60	370 (2.7)	45 (2.3)	3 (5.4)
0.20–0.40	1 170 (8.5)	170 (8.7)	6 (10.7)
0.00–0.20	4 767 (34.5)	676 (34.8)	6 (10.7)
–0.20–0.00	5 395 (39.1)	760 (39.1)	17 (30.4)
–0.40–0.20	1 239 (9.0)	165 (8.5)	7 (12.5)
–0.60–0.40	327 (2.4)	37 (1.9)	2 (3.6)
–0.80–0.60	142 (1.0)	20 (1.0)	3 (5.4)
–1.00–0.80	41 (0.3)	14 (0.7)	1 (1.8)
= – 1.00	82 (0.6)	23 (1.2)	10 (17.9)
Number of comparisons	13 810	1943	56

AFLP, amplified fragment length polymorphism.

may exist, small indels (i.e. 3 bp) evidently account for most of the unusual Φ coefficients observed within AFLP primer pairs (Table 4). In our interpretation of these data (Fig. 1), all indels show strong negative correlations (i.e. $\Phi = -1$). Moreover, 12 of these 15 co-dominant markers (Fig. 1) show compelling evidence of high selfing rates (i.e. $\bar{s}_H = 1$) within populations, including Calaveras, Contra Costa, Alameda, LK315d, Marin, SF St. Mary, SF McLaren and SF Bayview. Little or no variation of any sort was detected within the other populations. Because, seedlings of different wild plants (or plants grown together in the foundation seed fields of LK315d) were analysed, we are reasonably certain that these results reflect genetic variation among true-breeding genotypes (selfing mode of reproduction) rather than variation among colonizing plants that originated from reproductively isolated populations. Homozygous genotypes within LK315d are particularly interesting considering that the parents of these seedlings were grown under conditions that should have been highly conducive to cross-fertilization.

Although the size difference (3 bp) and frequency of putative allelic fragments (30/178) detected in this study are similar to that reported in AFLP studies of other grass (Becker *et al.* 1995; Bai *et al.* 1999; Larson *et al.* 2000b), it is possible that we failed to identify a substantial number of other indels. If we assume 30 fragments were allelic, then a total of 498 different, nonallelic sequences were analysed (i.e. 513 – 15 = 498). With an average size of 240 bp per fragment, this represents a sample of 119 520 bp. Virtually all of the indels within these sequences will produce polymorphic fragments. However, only 6.7% of the nucleotide variation within these sequences can be detected (i.e. 16 bp \times 298 fragments = 7968 bp total). This inference follows from the fact that restriction/priming

sites typically account for 16 bp per fragment, assuming one *EcoRI* site and one *MseI* site per fragment (Vos *et al.* 1995). Because five of these putative indels also showed evidence of nucleotide variation (Fig. 1), the actual number of polymorphisms detected was \approx 173 (i.e. 178 polymorphic fragments – 15 indels + 5 compound polymorphisms) ignoring convergent mutations. Therefore, \approx 8.7% of the polymorphisms were indels and the remaining 91.3% were nucleotide substitutions at the restriction/priming sites. Because only 6.7% of the nucleotide sequence variation can be detected using the AFLP method, described by Vos *et al.* (1995), we deduce that nucleotide substitutions were \approx 157-fold more common than indels. Bergelson *et al.* (1998) demonstrate that the frequency of nucleotide substitutions is sevenfold more common than indels in a survey of 5346 bp divided among coding and noncoding sequences of the *Adh*, *Dhs1* and *Gpa1* loci in *Arabidopsis thaliana*. Although it is possible that we failed to identify a substantial number of indels, our estimate of selfing rate (i.e. $\bar{s}_H = 1$) in *N. pulchra* is not likely to change in any case.

The average proportion of shared AFLP fragments (F) within and among populations was 0.995 and 0.953, respectively (Table 5). However, to estimate nucleotide sequence variation within and among populations, corrected values for the average proportion of shared AFLP fragments (F) were also computed by treating allelic fragments as shared characters (Table 5). The average proportion of shared AFLP fragments (F), corrected for indels, within and among populations was 0.996 and 0.962, respectively. As expected, values of F are higher when corrected for indels. The corresponding estimates of nucleotide variation within (π) and among populations (D_{XY}) were 0.00035 and 0.00247, respectively. Accordingly,

the net nucleotide divergence among populations (D_A) is 0.00212 and the proportion of total nucleotide variation among populations (G_S) is 0.86. Hence, the partitioning of nucleotide variation ($G_S = 0.86$) among these sample populations of *N. pulchra* is somewhat less than differentiation of among the Shannon–Weaver diversity values for AFLP variation ($G_S = 0.90$), but still much greater than the partitioning of Shannon–Weaver values ($G_S = 0.57$) for isozyme variation (Knapp & Rice 1998). Interestingly, estimates of nucleotide diversity ($\pi = 0.0004$), total nucleotide divergence ($D_{XY} = 0.0014$) and differentiation ($G_S = 0.64$) among populations of self-fertilizing *A. thaliana* (Bergelson *et al.* 1998) are very similar to our estimates for *N. pulchra*. Although population differentiation is certainly high in *N. pulchra*, estimates of nucleotide variation mam(Ber)18e.

oportioe

15d,

d

AFLPtechniquen

Fig. 1

Table 6 Estimated $\pi \times 1000$ (diagonal), $D_{XY} \times 1000$ (lower right) and $D_A \times 1000$ (upper right) within and between sample populations based on the estimated proportion of shared amplified fragment length polymorphisms (AFLPs) between haploid genomes (Innan *et al.* 1999)

	Teham.	LK115d	Colusa	LK215e	Calav.	Contra Costa	Alam.	LK315d	Marin	SF Yerba Buena	SF St. Mary	SF McLar	SF Bay	San Luis Obispo
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1)	0.00	0.85	2.23	2.43	1.66	1.95	1.98	1.83	2.64	2.95	3.03	2.54	2.81	3.62
(2)	1.45	1.20	1.31	1.76	0.48	1.09	1.19	0.97	2.17	2.35	2.10	2.00	2.01	2.82
(3)	2.23	1.91	0.00	1.45	1.53	1.63	1.59	1.44	2.57	2.82	2.70	2.27	2.28	3.29
(4)	2.43	2.36	1.45	0.00	1.98	2.47	2.31	1.83	3.04	2.89	2.77	2.60	2.54	3.82
(5)	1.91	1.33	1.78	2.23	0.50	1.31	1.34	1.25	2.32	2.37	2.25	1.96	2.03	3.04
(6)	2.10	1.84	1.78	2.62	1.71	0.31	1.51	1.28	2.09	2.28	2.41	2.12	1.87	3.27
(7)	2.23	2.04	1.84	2.56	1.84	1.91	0.50	1.38	2.26	2.31	2.52	2.09	2.09	3.10
(8)	2.17	1.91	1.78	2.17	1.84	1.78	1.97	0.69	2.30	2.28	2.42	2.26	2.20	3.35
(9)	2.89	3.02	2.82	3.29	2.82	2.49	2.76	2.89	0.50	1.40	1.21	0.98	1.18	3.30
(10)	2.95	2.95	2.82	2.89	2.62	2.43	2.56	2.62	1.65	0.00	1.14	1.11	1.17	3.55
(11)	3.15	2.82	2.82	2.89	2.62	2.69	2.89	2.89	1.58	1.26	0.25	0.73	0.99	3.17
(12)	2.76	2.82	2.49	2.82	2.43	2.49	2.56	2.82	1.45	1.33	1.07	0.44	0.89	3.13
(13)	3.09	2.89	2.56	2.82	2.56	2.30	2.62	2.82	1.71	1.45	1.39	1.39	0.56	2.87
(14)	3.62	3.42	3.29	3.82	3.29	3.42	3.35	3.69	3.55	3.55	3.29	3.35	3.15	0.00

Although geographical distance and genetic variation were significantly correlated, cluster analysis of the Calaveras, Marin and SF McLaren seedlings (Fig. 2) reveals mixtures of true-breeding genotypes that transcend and belie geographical origin. These observations may reflect seed dispersal from genetically different populations. Interestingly, the Contra Costa, Alameda and LK315 populations, from the inland side of San Francisco Bay, show a stronger affinity to other inland populations (i.e. Colusa, LK215e, Calaveras, LK115d and Tehama) that originate from collection sites much further away compared with populations collected from the nearby coastal side of San Francisco Bay (i.e. SF Yerba Buena, SF McLaren, SF St. Mary, SF Bayview and Marin). This putative division of coastal and inland populations, across San Francisco Bay was supported by the bootstrap confidence intervals (Fig. 2). Hierarchical analysis of D_{XY} values (Table 6) also indicates substantial differentiation ($G_S = 0.30$) between these two groups. Knapp & Rice (1998) assert that quantitative morphological variation among *N. pulchra* populations is primarily determined by climatic differences between the interior and coastal or maritime regions.

Discussion

Stebbins (1957) suggests that self-fertilization, in grasses and other plants, may be advantageous: (i) when conditions are unfavourable for crossing (such as periodic drought on the annual grasslands of California); (ii) for reproduction of one or two colonists following long-distance dispersal; and (iii) for reproduction of 'weedy' genotypes that are well-adapted to the disturbed niche, in which populations are constantly destroyed and quickly

built up again. Indeed, evidence suggests (Bartolome 1981; Bartolome & Gemmill 1981) that *Nassella pulchra* is an ephemeral species that has evolved to rapidly colonize or increase abundance on many disturbed sites. Results of this study indicate that self-fertilization in *N. pulchra* is another important adaptation that may be related to dispersal and colonization by this species.

Self-fertilization in *N. pulchra* has reduced variation within populations and increased apparent differences among populations. Estimates of nucleotide diversity within self-fertilizing populations of *N. pulchra* (0.00035) are almost two orders of magnitude less than estimates of nucleotide diversity within cross-fertilizing cultivars of *Pseudoroegneria spicata* (0.03400), also determined using the AFLP method (Larson *et al.* 2000a). Moreover, estimates of total nucleotide diversity within *N. pulchra* (0.00247) are also considerably less the estimates of nucleotide diversity in a multiple-origin polycross of 25 *P. spicata* populations (0.0366). Although no attempt was made to distinguish indels from nucleotide substitutions in *P. spicata* (Larson *et al.* 2000a), substantially more AFLP variation was observed in *P. spicata* than *N. pulchra*. Conversely, estimates of nucleotide differentiation within *N. pulchra* ($G_S = 0.86$) are more than one order of magnitude larger than differentiation between cultivars of *P. spicata* ($G_S = 0.07$). Large reductions in genetic variability within populations and increased differentiation among populations have also been observed in other self-fertilizing grasses (Price *et al.* 1984). Large values of G_S generally indicate restricted gene flow between populations (Bergelson *et al.* 1998), which in most plants occurs by pollen flow and/or seed dispersal. Although pollen flow is obviously reduced or eliminated in self-fertilizing species, seed dispersal of *N. pulchra* is

probably equal to or greater than seed dispersal of *P. spicata*. Selection and clonal expansion may also influence G_s by reducing variability within self-fertilizing populations (Bergelson *et al.* 1998). Clonal expansion of inbred founders may be expected to produce patches of homogeneous populations for weedy ephemeral species such as *Arabidopsis thaliana* (Bergelson *et al.* 1998). Relative to a cross-fertilizing population, selfing per se is only expected to reduce nucleotide variation by a factor of two when the selfing rate (\bar{S}_H) approaches one (Pollack 1987). Because recombination is reduced by self-fertilization, selection can also reduce the neutral genetic variability across large linkage blocks of DNA by way of hitchhiking effects or selective sweeps associated with favourable genes (Maynard Smith & Haigh 1974) or background selection against deleterious mutations (Charlesworth *et al.* 1993). Assuming normal rates of deleterious mutations, background selection may be expected to reduce the accumulation of neutral genetic variability within populations by another factor of 10 in a self-fertilizing population (Charlesworth *et al.* 1993). Therefore, a number of factors including reduced gene flow, selection and clonal expansion of inbred founders, may have reduced variation within self-fertilizing populations of *N. pulchra* and increased differentiation among these demes.

AFLP variation may reflect neutral genetic differences among populations of *N. pulchra*, not evident by isozyme analysis. In common garden studies, Knapp & Rice (1998) demonstrate that quantitative morphological variation among *N. pulchra* populations is correlated ($P = 0.0142$) with climatic differences across a relatively narrow range of longitudes, whereas isozyme variation is correlated ($P = 0.0036$) with geographical distance across a relatively wide range of latitudes. Based on these results, Knapp & Rice (1998) assert that poorly adapted plantings may result when using seed transferred between coastal and inland sites. Moreover, Knapp & Rice (1998) allege that isozymes (and perhaps molecular markers in general) have limited usefulness in predicting adaptive differences. However, the lack of congruence between morphological variation and isozyme variation (Knapp & Rice 1998) may reflect special limitations of the isozyme surveys that do not apply to methods of DNA fingerprinting that are potentially more informative. Isozymes are suitable for screening a relatively small number of neutral characters in a large number of plants (per population). However, estimates of genetic diversity and distance are affected far more severely by the number of loci sampled than by the number of individuals sampled (Gorman & Renzi 1979; Nei 1987). Even a few representative individuals may be sufficient to measure DNA variation if the number of loci examined is large. Indeed, the partitioning of Shannon–Weaver diversity values for AFLPs ($G_s = 0.90$), based on a relatively large number of loci in fewer individuals,

is considerably different from the partitioning of isozyme banding patterns ($G_s = 0.57$), even though the populations surveyed for isozyme variation originate from much wider range of latitudes. Although AFLP variation is also correlated with geographical distance, we observed a discontinuity of AFLP variation across a narrow region (San Francisco Bay and adjoining coastal mountains) separating coastal and inland areas that may be related to putative differences between the inland and coastal regions (Knapp & Rice 1998). Considering the self-fertilizing mode of reproduction and almost complete partitioning of all genetic variation among these populations of *N. pulchra*, it seems very likely that selection has somehow shaped AFLP variation in this species. Conversely, if genetic drift accounts entirely for the partitioning of AFLP variation among populations of *N. pulchra* then it is also likely that genetic drift had profound effects on the partitioning of quantitative morphological variation. We speculate that both selection and genetic drift influenced the partitioning of AFLP variation and quantitative morphological differences among natural populations of *N. pulchra*. Although adaptive variation within and among populations of *N. pulchra* may be important (Knapp & Rice 1998), there is not yet enough evidence to support the clear distinction or identification of two (i.e. inland and coastal) or more regionally distinct forms or ecotypes of *N. pulchra*. More extensive sampling of different sites is needed to detect genetic divergence among regionally distinct or ecologically distinct forms of *N. pulchra*. Hierarchical geographical analysis of genetic differences, within self-fertilizing species such as *N. pulchra*, might be efficiently performed by DNA fingerprinting a small number of plants from each of many different sites, perhaps even just one plant per site.

As expected, foundation seed lots of LK115d, LK215e and LK315d were genetically well defined, naturally diverse and most related to natural populations collected from nearby sites. Therefore, cultivated seed increases of LK115d, LK215e and LK315d probably did not have any inadvertent consequences on the genetic integrity of these commercial germplasm sources. Based on results of this study, we do not think that these plant materials are likely to cross with natural populations. However, the spatial scale of genetic identity among these Selected Class natural germplasm accessions and other natural populations is evidently very small. Moreover, the genetic identity of LK115d, LK215e and LK315d did not completely match that of natural populations collected from sites near the original source of these Selected Class germplasm releases. In fact, the AFLP method exposed genetic differences among virtually all of the natural populations, even within the city of San Francisco. Yet, one of the most ecologically important attributes of *N. pulchra* may be its ability to colonize disturbed sites (Bartolome 1981; Bartolome & Gemmill 1981). Moreover, much of the obvious variation

among populations of *N. pulchra* may be the result of inbreeding and genetic drift associated with clonal expansion of inbred founders. Of course, the value of genetic diversity and importance of germplasm conservation are widely recognized and greatly appreciated. To ensure planting success and adaptation, more than one commercial germplasm source (e.g. LK115d, LK215e and LK315d) could be used in the same planting. Ideally, native plant workers could develop and use genetically diverse synthetics (cross-fertilizing species) or composites (self-fertilizing species) for each region or ecotype, where in fact these intraspecific differences among regions or ecotypes may be significant. However, morphological variation would also be problematic for seed production, particularly for composites of true-breeding genotypes. In any case, the motive to use native species in large-scale revegetation efforts needs to be complemented by effective germplasm utilization, innovative agronomic techniques, and a viable industry that can efficiently produce a readily available source of high-quality seed for these species. Thus, commercial germplasm of native grass species can be made practically available and useful in many conservation plantings.

Acknowledgements

This work was supported by joint contributions of the USDA, the Department of Defense Strategic Environmental Research and Development Program CS1103 project, and the US Army BT25-EC-B09 project. We also acknowledge the assistance of Dr Eric Knapp, Dr Jeffery Steiner, and journal referees in providing helpful reviews of this manuscript.

References

- Association of Official Seed Certifying Agencies (1999) *Certification Handbook*. AOSCA, Meridian, ID.
- Bai G, Tefera H, Ayele M, Nguyen HT (1999) A genetic linkage map of *tef* [*Eragrostis tef* (Zucc.) Trotter] based on amplified fragment length polymorphism. *Theoretical and Applied Genetics*, **99**, 599–604.
- Bartolome JW (1981) *Stipa pulchra*, a survivor from the pristine prairie. *Fremontia*, **9**, 3–6.
- Bartolome JW, Gemmill B (1981) The ecological status of *Stipa pulchra* (Poaceae) in California. *Madroño*, **28**, 172–184.
- Becker J, Vos P, Kuiper M, Salamini F, Heun M (1995) Combined mapping of AFLP and RFLP markers in barley. *Molecular and General Genetics*, **249**, 65–73.
- Benham J, Jeung J-U, Jasieniuk M, Kanazin V, Blake T (1999) Genographer: a graphical tool for automated AFLP and microsatellite analysis. *Journal of Agricultural Genomics*, **4**, Article 4.
- Bergelson J, Stahl E, Dudeck S, Kreitman M (1998) Genetic variation within and among populations of *Arabidopsis thaliana*. *Genetics*, **118**, 1311–1323.
- Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics*, **134**, 1289–1303.
- Chung MG, Hamrick JL, Jones SB, Derda GS (1991) Isozyme variation within and among populations of *Hosta* (Liliaceae) in Korea. *Systematic Botany*, **16**, 667–684.
- Dolan RW (1994) Patterns of isozyme variation in relation to population size, isolation, and phylogeographic history in royal catchfly. *American Journal of Botany*, **81**, 965–972.
- Dyer AR, Fossum HC, Menke JW (1996) Emergence and survival of *Nassella pulchra* in a California grassland. *Madroño*, **43**, 316–333.
- Gorman GC, Renzi J (1979) Genetic distance and heterozygosity estimates in electrophoretic studies: effects of sample size. *Copeia*, 242–249.
- Heady HF, Bartolome JW, Pitt MD, Stroud MG, Savelle GD (1991) California grassland. In: *Natural Grasslands, Ecosystems of the World, Vol. 8A* (ed. Coupland RT), pp. 313–335. Elsevier, Amsterdam.
- Huntsinger L, McClaran MP, Dennis A, Bartolome JW (1996) Defoliation response and growth of *Nassella pulchra* (A. Hitchc.) Barkworth from serpentine and non-serpentine populations. *Madroño*, **43**, 46–57.
- Innan H, Terauchi R, Kahl G, Tajima F (1999) A method for estimating nucleotide diversity from AFLP data. *Genetics*, **151**, 1157–1164.
- Knapp EE, Rice KJ (1994) Starting from seed: genetic issues in using native grasses for restoration. *Restoration and Management Notes*, **12**, 40–45.
- Knapp EE, Rice KJ (1998) Comparison of isozymes and quantitative traits for evaluating patterns of genetic variation in purple needlegrass (*Nassella pulchra*). *Conservation Biology*, **12**, 1031–1041.
- Larson SR, Jones TA, McCracken CL, Hu Z-M, Palazzo AL (2000a) Genetic diversity of bluebunch wheatgrass cultivars and a multiple-origin polycross. *Crop Science*, **40**, 1142–1147.
- Larson SR, Rutger JN, Young KA, Raboy V (2000b) Isolation and genetic mapping of a non-lethal rice (*Oryza sativa* L.) *low phytic acid 1* mutation. *Crop Science*, **40**, 1397–1405.
- Lesica P, Allendorf FW (1999) Ecological genetics and the restoration of plant communities: mix or match? *Restoration Ecology*, **7**, 42–50.
- Love MR (1954) Interspecific hybridization in *Stipa* II. Hybrids of *S.cernua*, *S.lepidia*, and *S.pulchra*. *American Journal of Botany*, **41**, 107–110.
- Mantel NA (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.
- Maynard Smith J, Haigh J (1974) The hitch-hiking effect of a favourable gene. *Genetical Research*, **23**, 23–35.
- Montero LM, Filipinski J, Gil P, Capel J, MartPnez-Zapater JM, Salinas J (1992) The distribution of 5-methylcytosine in the nuclear genome of plants. *Nucleic Acids Research*, **20**, 3207–3210.
- Murray MG, Thompson WF (1980) Rapid isolation of high-molecular weight DNA. *Nucleic Acids Research*, **8**, 4321–4325.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the USA.*, **76**, 5269–5273.
- Nei M, Miller JC (1990) A simple method for estimating average number of nucleotide substitutions within and between populations from restriction site data. *Genetics*, **125**, 873–879.
- Nordberg M, Donnelly P (1997) The coalescent process with selfing. *Genetics*, **146**, 1185–1195.
- Paul S, Wachira FN, Powell W, Waugh R (1997) Diversity and genetic differentiation among populations of Indian and Kenyan

- tea (*Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. *Theoretical and Applied Genetics*, **94**, 255–263.
- Pollack E (1987) On the theory of partially inbreeding finite populations. I. Partial selfing. *Genetics*, **117**, 353–360.
- Price SC, Shumaker KM, Kahler AL, Allard RW, Hill JE (1984) Estimates of population differentiation obtained from enzyme polymorphisms and quantitative characters. *Journal of Heredity*, **75**, 141–142.
- Richards RT, Chambers JC, Ross C (1998) Use of native plants on federal lands: policy and practice. *Journal of Range Management*, **51**, 625–632.
- Roundy BA (1999) Lessons from historical rangeland revegetation for today's restoration. In: *Revegetation with Native Species: Proceedings* (comp. Holzworth LK, Ray RK), pp. 33–38. *RMRS-P-8*. USDA, Forest Service, Rocky Mountain Research Station, Ogden, UT.
- Roundy BA, Shaw NL, Booth DT (1997) Using native seeds on rangelands. In: *Proceedings: Using Seeds of Native Species on Rangelands* (comp. Shaw NL, Roundy BA), pp. 1–8. *Gen. Technical Report INT-GTR-372*. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, Ogden, UT.
- Salinas J, Matassi G, Montero LM, Bernardi G (1988) Compositional compartmentalization and compositional patterns in the nuclear genomes of plants. *Nucleic Acids Research*, **16**, 4269–4285.
- Stebbins GL (1957) Self fertilization and population variability in the higher plants. *American Naturalist*, **91**, 337–354.
- Stebbins GL, Love RM (1941) A cytological study of California forage grasses. *American Journal of Botany*, **28**, 371–382.
- USDA-SCS (1981) *Land Resource Regions and Major Land Resource Areas of the United States. Agriculture Handbook 296*. US Government Printing Office, Washington, D.C.
- Vos P, Hogers R, Bleeker M *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.
- Young SA (1995) Verification of germplasm origin and genetic status by seed certification agencies. In: *Proceedings: Wildland Shrub and Arid Land Restoration Symposium* (comp. Roundy BA, McArthur DE, Haley JS, Mann DK), pp. 293–295. *Gen. Technical Report INT-GTR-315*. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, Ogden, UT.

The USDA-NRCS Plant Materials Program aims to identify and develop a wide diversity of plant materials to solve environmental and conservation problems across all regions of the United States. The USDA-ARS Forage and Range Research Laboratory performs applied research related to many pasture and rangeland plants important to the western United States.
