Characterization of the USDA *Poa pratensis* collection using RAPD markers and agronomic descriptors

R.C. Johnson[1,*], W.J. Johnston[2], C.T. Golob[2], M.C. Nelson[3] and R.J. Soreng[4]

[1]*USDA, Western Regional Plant Introduction Station, Washington State University, Pullman, WA 99164-6402, USA;* [2]*Crop and Soils Sciences, Washington State University, Pullman, WA 99164-6420, USA;* [3]*US Department of Golf Association, Box 5844, Twin Falls, ID 83303, USA;* [4]*Botany Department, NMNH, Smithsonian Institution, Washington, D.C 20560-0166, USA;* *Author for correspondence (e-mail: rcjohnson@wsu.edu; phone: 509-335-3771; fax: 509-335-6654)*

Abstract

Characterization of germplasm collections is critical to assess collection diversity and enhance utilization. A *Poa pratensis* L. germplasm collection of 228 accessions representing 26 countries, along with 17 commercial check cultivars, was characterized using 86 random amplified polymorphic DNA (RAPD) markers and 17 agronomic descriptors. The Dice similarity coefficient used for RAPD data ranged from 0.56 to 0.95 and average Euclidean distance used for agronomic data ranged from 0.28 to 2.52. No two accessions had a similarity of one or a distance of zero, showing there were no duplicate entries. Cluster analysis of RAPD data using the unweighted pair-group method using arithmetic averages (UPGMA) revealed 11 accessions with particularly low similarity values. These were subsequently found to be misidentified *Poa* species (one each of *P. alpina, P. compressa, P. glauca, P. urssulensis* and seven *P. trivialis*). For RAPD data, 62% of the entries were in one large cluster with 46 additional clusters containing one to 13 accessions. For agronomic data, 89% of the entries were in four main clusters. This clustering pattern for RAPD and agronomic data suggested unique genotypes were generally under represented in the collection. The agronomic-based clusters showed some broad separation by accession origin, but in general, origin did not correspond closely with the clustering pattern. The correlation between the RAPD and agronomic-based distance matrices, excluding misidentified accessions, was highly significant (P<0.01) (n = 234, r = −0.14). However, the correlation represented a relatively small fraction of the total variation, indicating that both molecular and agronomic characterizations were needed to assess overall diversity.

Introduction

The National Plant Germplasm System (NPGS), through the United States Department of Agriculture (USDA), maintains extensive collections of plant genetic resources available for distribution to scientists worldwide. The Western Regional Plant Introduction Station (WRPIS), Pullman WA, is a part of the NPGS network of germplasm repositories. The *Poa* genus is among the forage and turfgrass collections at the WRPIS and is currently represented by more than 50 taxa and 750 accessions. The major species in this collection is *P. pratensis* L. (Kentucky bluegrass) currently with 348 accessions. *P. pratensis* is used extensively as a forage, and as turfgrass in lawns, golf courses, parks, and sports fields (Bashaw and Funk 1987). It is also used for stabilizing eroded and disturbed soils and improving soil structure and fertility. *P. pratensis* is thought to have originated in Eurasia and is widely adapted and cultivated in temperate climates throughout the world (Carrier and Bort 1916; Soreng 1990).

P. pratensis is a facultative apomictic species. Apomictic reproduction dominates, so progeny are

most often genetically identical to the seed bearing parent plant. In some cases sexual reproduction does occur and is a source of introgression of new genetic combinations (Åkerberg 1939; Clausen 1961; Huff and Bara 1993).

Molecular and agronomic characterization of germplasm collections are needed for assessing diversity, revealing duplication and misidentified accessions, and for identifying acquisition needs. Utilization of the USDA *P. pratensis* collection will also be enhanced as information concerning accessions becomes available to germplasm users. Characterization can be completed using traditional agronomic evaluations and also with molecular markers. There are advantages and disadvantages to both methods of characterization. Agronomic characters may be immediately useful to identify and select desirable genes or genotypes, but are subject to genotype by environment interactions, which are not expected with molecular-based evaluations. An assessment of the diversity of newly acquired accessions may be simplified by the use of molecular markers. The genetic distances of new accessions, or those from other collections, could be compared to those of existing accessions without conducting extensive field evaluations. But unless linked to specific traits, the random genetic variation would not be immediately useful for selection and breeding.

Among the molecular marker systems available, RAPD (random amplified polymorphic DNA) markers provide a relatively simple and easy PCR (polymerase chain reaction) based technique that can be applied to the large number of accessions often needed to characterize germplasm collections (Williams et al. 1990). There are some inherent difficulties with RAPDs associated with their reproducibility, dominant nature, and uncertain marker homology (Weising et al. 1995). Huff and Bara (1993) found silver-stained RAPD (ssRAPD) markers more reliable than RAPD analysis using agarose gels and ethidium bromide staining in *P. pratensis*. However, the agarose gel and ethidium bromide staining system was used successfully by Sweeney and Danneberger (1995) to characterize *P. annua* L. populations. And RAPDs have been successfully used for assessing genetic diversity and relatedness in a wide range of grass species including rice (*Oryza sativa* L.) (Virk et al. 1995), emmer wheat [*Triticum turgidum* L. subsp. *dicoccoides* (Körn. ex Asch. & Graebn.) Thell.] (Fahima et al. 1999), barley (*Hordeum vulgare* L.). (Strelchenko et al. 1999), *Lolium* species (Hayward et

al. 1994; Huff 1997), and the *Lolium-Festuca* complex (Šiffelová et al. 1997).

Johnston et al. (1997) completed phenotypic evaluations for agronomic factors and developed a core collection for the USDA *P. pratensis* collection. However, the clustering pattern and its relationship to accession origin, along with comparisons to molecular diversity analysis, were not completed. The objective of this study was to use RAPD markers to characterize the USDA *P. pratensis* collection and compare diversity based on RAPDs with diversity based on agronomic evaluations. The extent of duplication, and associations between germplasm origins and diversity patterns in the collection, will also be assessed along with recommendations for germplasm acquisition.

Materials and Methods

Characterization of the USDA collection of *P. pratensis* was initiated in 1994 on 228 accessions (at the time all accessions in the collection with ≥70% germination), along with 17 commercial cultivars representing morphological diverse turfgrass material. In May 1994, seeds were planted in 1 m rows in three randomized complete blocks at the Turfgrass Research Center, Washington State University, Pullman, WA. Rows were spaced 0.3 m apart within a 28- by 28-m plot area. Additional details of the experimental design can be found in Johnston et al. (1997). The 228 accessions represented 26 countries with the largest set of material collected in Alaska (86 accessions). The rest of North America was represented by 11 accessions. Northwest and Central Europe were represented by 42 accessions, with 33 from Southwest Asia, 24 from Central Asia, 12 from Scandinavia, eight from Russia, six from Spain, four from Japan, one from Morocco, and one from South Africa.

In May 1995, samples of fresh tissue were taken from 10 plants in each of the 735 plots (245 entries times three replications). Each sample consisted of a segment of un-emerged leaves approximately 3 mm long taken from just above the stem nodes. The 10 samples from each plot were combined in a 1.5-m microfuge tube and placed on ice. Fifty μL of TE extraction buffer [100 mM Tris-HCL (pH 8.0), 1.4 M NaCl, 20 mM EDTA, 2% hexadecyltrimethyl ammonium bromide (CTAB), and 0.4% 2-mercaptoethanol] was added to each tube and fresh tissue ground with a microfuge tube pestle mounted onto a

variable speed hand motor (Dremel model 732, Dremel, Racine WI, USA). After about 10 s of grinding, another 450 μL of buffer was added to the tube and the sample was ground for a few more seconds to suspend the sample in the buffer. The sample was again placed on ice until all samples for a given day (40 to 60) were ground. Samples were then incubated in a 65°C water bath for 1 to 2 h and extraction completed with chloroform:isoamyl alcohol (24:1) as outlined by Murray and Thompson (1980).

A PCR amplification protocol similar to that of Williams et al. (1990) was used. Initially, a random set of 12 accessions was used to screen 120, 10-mer primers obtained from Genosys Biotechnologies (The Woodlands TX, USA). The reaction volume of 25 μL contained 10 mM Tris-HCL (pH 8.3), 50 mM $MgCl_2$, 0.2 mM each dATP, dGTP, dCTP, and dTTP, 0.2 μL primer, and 1 unit *Taq* polymerase (Promega, Madison WI, USA). Approximately 2 ng DNA was used as template. Amplification took place in a thermal cycler (Perkin-Elmer, Norwalk CT, USA) for 40 cycles. Each cycle consisted of denaturation for 20 s at 94°C, 1 min annealing at 36°C, followed by a 3 min rise to 72°C, and primer elongation for 1 min at 72°C. The 40 cycles were followed by an 8 min extension period at 72°C. Amplification products were resolved by electrophoresis on 2% agarose gels (1% Seakem agarose, 1% NuSieve agarose) from FMC Bioproducts (Rockland ME, USA). PCR product was stained with ethidium bromide and visualized by UV-fluorescence.

From the 120 primers screened using 12 accessions, six primers were selected and used with DNA from the 735 plots described above. The selected primer sequences are given in Table 1. Eighty-six bands that reproduced across replications were scored as present (1) or absent (0). It was assumed that bands of the same molecular weight in different samples were identical.

Data for 17 agronomic characteristics were collected in the fall of 1994 and the spring and summer of 1995 as described by Johnston et al. (1997). These descriptors were adapted from the standard list developed by the U.S. Forage and Turf Grass Germplasm Committee. The descriptor list and agronomic data for each accession are available through the Germplasm Resources Information Network (GRIN) database (http://www.ars-grin.gov/npgs/). A brief description of these factors and when they were evaluated is given below.

Emergence (days from planting) was recorded in June 1994; emergence was considered successful when approximately 50% of a plot had emerged seedlings. MSMA (monosodium acid methanearsonate) phytotoxicity and powdery mildew infection (incited by *Erysiphe graminis* D.C.) were rated on a 0 to 3 scale with zero equaling no MSMA toxicity or no infection. (The MSMA herbicide was applied on 21 June 1994 to control a pervasive stand of witchgrass, *Panicum capillare* L.). The MSMA ratings were completed on 25 July 1994 and the powdery mildew infection rated on 15 May 1995.

Eight factors were evaluated on a 1 to 9 rating scale. These were leaf texture (9 = finest), leaf color (9 = darkest green), plant uniformity (9 = most), leaf habit (9 = most upright), dwarf character (9 = most), biomass (9 = most), turf potential (9 = best overall), and spring green-up (9 = earliest). Turf potential was a subjective rating intended to assess the potential for overall turfgrass quality based on visual evaluation. The eight rated factors were all evaluated in October 1994 expect spring green-up, which was evaluated in March 1995. Mean canopy height was measured on 12 October 1994. Heading date, 50% anthesis, harvest date, height at harvest, and seed yield were determined in 1995. Heading date was defined as when more than half the culms had emerged with visible heads. Harvest coincided with seed maturity, which is associated with the start of seed shattering (Morrison and Law 1978). Harvesting was completed by hand and plant material threshed, air cleaned, and weighed for yield.

Table 1. Primers, sequences and summary of RAPD markers used to characterize *Poa pratensis* accessions.

Primers	Sequence 5' to 3'	No. of markers	Size (bp) min–max	No. polymorphic markers	%polymorphic markers
4-70-74 1	GGACCGCTAG	13	290–1500	9	69
4-60-33 2	CTGCGATACC	14	220–2100	12	86
1-60-10 3	GCAGACTGAG	13	295–2000	8	62
CS-53 4	GCCTCATACC	16	310–1600	15	94
2-70-59 5	GCTCTCACCG	14	280–1320	14	100
3-70-67 6	AGCCTGACGC	16	420–1300	16	100
	total	86		74	mean = 85.2

352

Cluster analysis using the UPGMA (unweighted pair-group method using arithmetic averages) method was completed as described by Romesburg (1990). All distance matrix calculations and UPGMA clustering was completed using NTSYS-pc, version 2 (Exeter Software, Setauket, NY). For the RAPD data, the Dice coefficient (also known as the Sorenson coefficient) was used prior to UPGMA clustering. For the 17 agronomic characters, the data were standardized to have a mean of zero and a variance of one, average Euclidean distance calculated, and UPGMA cluster analysis completed. The average Euclidean distance is a dissimilarity coefficient; that is, the larger the value the greater distance between pairs of accessions. The Dice coefficient is a similarity coefficient; the larger the value the greater the similarity (the smaller the distance) between a given pair of accessions.

Using the COPH program in NTSYS-pc, the hierarchal system of clusters was used to produce a cophenetic matrix. The original distance matrices for RAPD data and for the average Euclidean distance was correlated with the cophenetic distance matrix using the MXCOMP program to give a subjective test of the goodness of fit for the cluster analysis.

For RAPD data, a principal coordinates analysis was completed as described by Gower (1966) using the D-CENTER and EIGEN programs in NTSYS-pc. For the agronomic data, SAS PROC PRINCOMP (SAS Institute 1985) was used to complete a principal components analysis. The principal component means and standard deviations were also calculated for clusters derived from the cluster analysis of the agronomic data. Correlation of the agronomic and RAPD-based distance matrices was completed using the MXCOMP program in NTSYS-pc. A Mantel test (Manly 1986) resulting from 1000 permutations was completed to determine the significance level of the correlation coefficient between the RAPD-based and agronomic-based distance matrices.

Results and Discussion

From the 120 primers evaluated, six primers were identified for use in characterizing the USDA *P. pratensis* collection. Those six were selected because they yielded a reasonably high number (86) of reproducible markers that were easily resolved by gel electrophoresis (Table 1). Additional primers fitting this criteria could also have been selected, but 86

markers were considered sufficient to give a high probability of distinguishing accessions (Virk et al. 1995). Among the primers, the percentage of polymorphic markers ranged from 62 to 100%, showing a high degree of molecular variation. The widest range in marker size was 1880 bp for primer 4-60-33 and the narrowest was 880 bp for primer 3-70-67 (Table 1). The number of reproducible markers per primer ranged from 13 to 16 with the percentage of polymorphic markers ranging from 62 to 100% (Table 1). All primers produced some weak and unreliable fragments. These were not scored but indicated that it was essential to repeat reactions or complete PCR on replications, as we have done, to ensure reliable results.

The correlation coefficient between the data matrix and the cophenetic matrix for RAPD data was 0.84. A correlation of greater than 0.80 is considered high enough to indicate that the clustering dendrogram gave a good representation of the original matrix (Romesburg 1990). The UPGMA hierarchal clustering program resulted in a range of Dice coefficient values from 0.56 to 0.95. There were no two accessions with a similarity of one, showing that the collection had no genetic duplicates.

The clustering pattern for the RAPD data showed a degree of 'chaining' (Romesburg 1990) as the clusters tended to become progressively larger as the similarity coefficient increased (Figure 1). As a result, it was not possible to find a set of clusters with a roughly equal number of accessions. Even when the dendrogram was cut over a range of similarity values, there tended to be a relatively large number of individual accessions forming single clusters, and then a single cluster with a large number of accessions. To illustrate this the dendrogram was cut at a similarity of 0.80 (Figure 1). That value was 40% of the distance from the maximum similarity of 0.95 to the minimum of 0.56, and was chosen because it appeared to give the best clustering balance possible. Cutting the dendrogram at a higher similarity value gave more and more single clusters and cutting at a lower similarity tended to give only a few clusters with a single large group containing the vast majority of accessions. Cutting the dendrogram at 0.80 resulted in 47 clusters of which 26 clusters were represented by a single accession or cultivar. Eleven clusters had two entries, four had three entries, three had four entries, one had eight, one had 13, and one had 152 entries (Table 2). This showed that even though there was no duplication in the collection, there was a large group of

accessions with relatively high similarity, suggesting limited diversity for a large portion of the collection.

Clusters 1 through 12 formed a set of highly dissimilar accessions (Figure 1). Field observation showed that the *P. pratensis* collection contained suspected misidentified accessions. Because of the low similarity compared with the rest of the collection, these were either misidentified or highly unusual *P. pratensis* accessions. Plants of these accessions were subsequently grown and mounts made for taxonomic identification. Seven of the 12 accessions were identified as *P. trivialis* L.. Of these, two were identified as *P. trivialis* subsp. *silvicola* (Guss.) Lindb. f. It is possible that all the *P. trivialis* are subspecies *silvicola* since all seven accessions showed lemma with pubescent marginal veins consistent with the subspecies. However, bulbous rhizomes must be observed to be certain, and insufficient rhizomes were available on five of the mounts to make an unequivocal determination to the subspecies level. Cluster 12 was tentatively identified as the Siberian species *P. urssulensis* Trin., and this is consistent with its Altai mountain origin. From this set of 12, the accession forming cluster 2 (PI 505896) was identified as *P. pratensis*, making it highly unique based on RAPD analysis (Table 2). Thus, the RAPD analysis was successful in distinguishing misidentified accessions within the USDA *P. pratensis* collection. In related studies, RAPDs were used by Martin et al. (1997) to identify *Oryza* species and misidentified *Capsicum* accessions were revealed in by RAPD analysis in a study by Rodriguez et al. (1999).
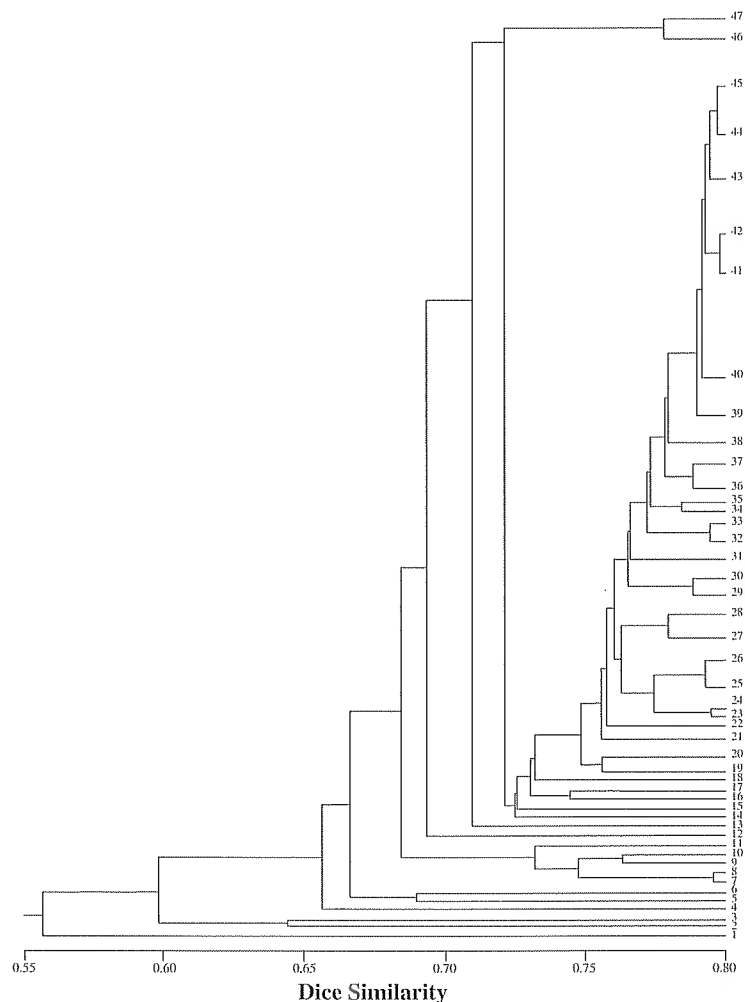

Figure 1. Dendrogram of RAPD data based on 86 markers and cut at 40% of the maximum to minimum range of the Dice similarity coefficient (0.80). The numbers refer to the resulting clusters described in Table 2.

Table 2. Number of accessions, origin, commercial cultivar checks, and species (if known to be other than *Poa pratensis*) for UPGMA clusters of RAPD data clusters cut at 40% of the distance from the maximum to the minimum Dice similarity.

Cluster	Accessions	Origin, species and cultivar
1	1	Morocco, *P. trivialis*
2	1	Russia, donated from the Vavilov Institute (VIR)
3	1	Iran, *P. trivialis*
4	1	Caucasus, *P. compressa*
5	1	Canada, *P. glauca* ssp. *glauca*
6	1	Switzerland, *P. alpina*
7	1	Iran, *P. trivialis* ssp. *sylvicola*
8	1	Iran, *P. trivialis*, some *P. pratensis* mix
9	1	Iran, *P. trivialis* ssp. *sylvicola*
10	1	Iran, *P. trivialis*
11	1	Afghanistan, *P. trivialis*
12	1	Russia, *P. urssulensis*
13	1	Czech Republic
14	1	Alaska
15	1	USA (other than Alaska)
16	1	USA (other than Alaska)
17	1	Alaska
18	1	Alaska
19	1	Iran
20	2	Norway (1) Netherlands (1)
21	2	Hungary (1) Denmark (1)
22	1	Alaska
23	1	Alaska
24	1	India
25	3	Alaska (2); cultivar Nublue
26	3	Alaska (1) Russia (1) Spain (1)
27	3	Iran (1) Russia (2)
28	2	Turkey (1) Russia (2)
29	2	Alaska (2)
30	2	Alaska (1) Kazakhstan (1)
31	2	Alaska (2)
32	2	Alaska (1) Kazakhstan (1)
33	2	S. Africa (1) Spain (1)
34	1	cultivar Mystic
35	1	cultivar Midnight
36	2	Alaska (1) Sweden (1)
37	3	Alaska (2) Turkey (1)
38	2	Alaska (2)
39	4	Alaska (1) Turkey (1); cultivars Dawn and Park
40	13	Alaska (11) Canada (1) Japan (1)
41	2	Alaska (1) Russia (1)
42	8	Alaska (1) Hungary (1) Iran (3) Turkey (1) Russia (1)
43	4	Alaska (1) Bulgaria (1) Kazakhstan (1) Sweden (1)
44	4	Kazakhstan (1) Netherlands (1) Russia (1) Spain (1)
45	152	Alaska (49), Russia (11), Iran (10), Poland (9), Turkey (7) Kazakhstan (6), Canada (5) Netherlands (5), and 50 others with less than 3 per country; cultivars Julia, Baron, Eclipse, Ikone, Baritia, Victa, Monopoly, Coventry, Banff, Plush, Kenblue, and Washington
46	1	Canada
47	1	Switzerland (questionable origin)

Although the cultivars Mystic and Midnight formed single clusters 34 and 35, commercial cultivars generally did not cluster together as separate groups but were integrated into clusters with the germplasm accessions (Table 2). A total of 71% of the commercial checks, 12 of the 17, were in cluster 45. That cluster, with 152 entries, represented 62% of all entries. Thus, there was some concentration of checks in cluster 45 but not enough to conclude a disproportionate amount. A number of the checks had relatively

high similarity with accessions. For example, the similarity between the cultivar Washington and PI 251166 from Bosnia was 0.92, and the similarity between the cultivar Banff and PI 311085 from Sweden was 0.91.

In most cases, country of origin was only marginally associated with the clustering pattern, if at all (Table 2). For example, the two most closely related accessions had a similarity of 0.95 but were none the less geographically distant. They were PI 204490 from Turkey and PI 314733 from Kazakhstan. The next most closely related accessions with a similarity slightly less than 0.95 were PI 251701, originally from Serbia and donated through Denmark, and PI 349174 from Alaska.

The principal coordinates analysis for the RAPD data showed that the eigenvalues for the first three principal components explained only 15% of the total variation, and 17 principal components were needed before more than 50% of the variance was explained. In other words, data reduction into two or three principal coordinates for easy illustration of the main features of the data was not possible. If this were the case it would suggest a good deal of correlation among the factors in the analysis (Manly 1986). But since it was not the case, it appeared that most markers provided a substantial amount of independent information concerning the genetic structure of the *P. pratensis* accessions and cultivars.

For agronomic data, the dendrogram based on average Euclidian distance did not represent the original agronomic distance matrix quite as well as the dendrogram based on RAPD data and the Dice coefficient. The cophenetic correlation was 0.78 compared to 0.84 for the RAPD data. However, it was high enough to assume that the dendrogram was not a greatly distorted representation of the distance matrix (Romesburg 1990). The range for average Euclidian distance was from a maximum of 2.52 to a minimum of 0.28. Thus, as with the RAPD data, there were no two accessions or cultivars that could be considered duplicates.

When the agronomic-based dendrogram was cut at the same relative point as the dendrogram based on RAPD data; that is, 40% of the range from minimum to the maximum distance values, 19 clusters resulted (Figure 2). Of these, 42% (8 of 19) of the clusters were represented by only one entry (Table 3). That fraction of single entry clusters was somewhat lower than those from the RAPD clustering, which had 55% (26 of 47) of the clusters represented by a single

entry. For the RAPD results, each misidentified accession was represented by a single cluster (Table 2), but for agronomic data, the 11 misidentified accessions were associated with only six clusters (Table 3). Six of the seven *P. trivialis* accessions were from Iran and Afghanistan and all six were in cluster 19. The remaining *P. trivialis* accession formed cluster 4, and was from Morocco. Clusters 1 through 3 and cluster 13 represented the other four misidentified species (Table 3). Thus, both RAPD and agronomic data distinguished the misidentified species. In the RAPD analysis, the *P. pratensis* accession PI 505896 was highly unique and formed cluster 2 (Figure 1 and Table 2). In the agronomic data, however, this accession was part of cluster 16 and thus was not highly unique compared with other *P. pratensis* accessions.

As expected, the cluster analysis organized the accessions into diverse groups based on differing agronomic characteristics (Table 4). Compared to clusters composed of *P. pratensis* accessions, the clusters of the misidentified species often showed extremes. This was exemplified by the slow emergence, fine leaf texture, dwarf character, late harvest, short height and low biomass of *P. glauca* compared to most *P. pratensis* clusters. Nevertheless, the diversity among the *P. pratensis* clusters was great enough that cluster means near those of the misidentified species could often be found. For example, *P. pratensis* in cluster 10 was nearly as slow to emerge as *P. glauca* in cluster 3, and the difference between the rapid emergence of *P. pratensis* in cluster 7 and *P. trivialis* in cluster 19 was not significant (Table 4). The *P. trivialis* accession forming cluster 4 contrasted with those in cluster 19 in that it had fine leaf texture, less uniformity, high MSMA burn, and shorter height.

Among the *P. pratensis* accessions, cluster 7 also had many attributes associated with high turf quality; that is, early green-up, dark green color, and high turf potential ratings (Table 4). Cluster 7 also had seed yield near the experiment wide value of 56.4 g m^{-2}, showing reasonably high yield with good turf attributes, yet none of the accessions were from the set of check cultivars, and accessions from cluster 7 originated from diverse origins (Table 3).

Four major agronomic clusters (16, 15, 14 and 6) were identified (Table 3). When the eleven misidentified accessions were excluded, those clusters included 209 or 89% of the total entries. The largest number of entries was in cluster 16 with 85 (Table 3), which was dominated by accessions from Eurasia. An examination of some key agronomic factors showed
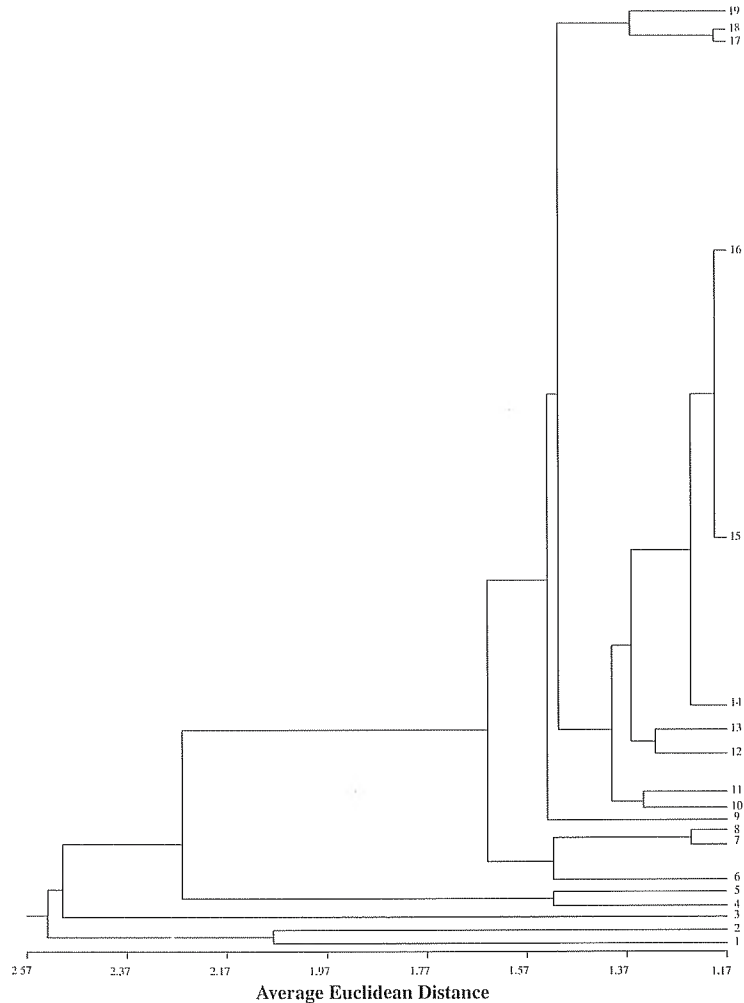
Figure 2. Dendrogram of agronomic data based on 17 descriptors and cut at 40% of the minimum to maximum range of the average Euclidean distance coefficient (1.17). The numbers refer to the resulting clusters described in Table 3.

cluster 16 was typified by accessions with generally lower leaf texture, leaf color, and turf potential ratings, especially compared to clusters 6 and 7. But seed production of cluster 16 was near the experiential wide mean of 56.4 g m^{-2}(Table 4). Despite the relatively low turf characteristics, the cluster included nine commercial cultivars, but only Dawn, Eclipse, and Julia had turf potential ratings greater than 6.0, which we considered acceptable for commercial use.

Cluster 15 had 56 entries and was also dominated by material from Eurasia. Accessions were typically early in heading and with high seed production (Table 4). Within this cluster there were accessions with both above average seed production and high turf quality ratings, which is a desirable combination for turfgrass cultivars. Although some accessions in cluster 15

showed potential for cultivar development, there were only two commercial cultivars in the cluster (Kenblue and Washington) (Table 3).

Clusters 14 and 6 were dominated by accessions from Alaska. Cluster 14 had 89% (42 of 47) of its entries from Alaska and cluster 6 had 100% (all 21) entries from Alaska. These two clusters represented 76% (63 of 83) of the total Alaska entries. However, clusters 14 and 6 were from different major branches in the dendrogram (Figure 2). The accessions in cluster 14 were actually more closely related to accessions in clusters 15 and 16, with mostly Eurasian origin, than those accessions in cluster 6. Cluster 6 had significantly darker color, more upright leaf habit, higher turf potential, was more dwarf, headed later, had later harvest, shorter harvest height, less biomass

Table 3. Number of accessions, origin, cultivar checks, and species (if other than *Poa pratensis*) for UPGMA clusters based on agronomic data cut at 40% of the distance from the minimum to the maximum average Euclidean distance.

Cluster 1	Accessions 1	Origin, species and cultivar Russia, *P. urssulensis*
2	1	Caucasus, *P. compressa*
3	1	Canada, *P. glauca* ssp. *glauca*
4	1	Morocco, *P. trivialis*
5	2	Alaska (2)
6	21	Alaska (19); cultivars Bartitia and Midnight
7	5	Alaska (2) Netherlands (1) Norway (1) Russia (1)
8	1	Netherlands
9	1	Alaska
1	1	Spain
11	6	Bosnia (1) Turkey (1); cultivars Banff, Barron, Victa
12	5	Alaska (2) Kazakhstan (2) Iran (1)
13	1	Switzerland, *P. alpina*
14	47	Alaska (42) Russia (1) Netherlands (1) Canada (1) Hungary (1) Iran (1)
15	56	Iran (10) Russia (7), Alaska (6), Kazakhstan (5), and 28 others represented by countries with one to three accessions each; cultivars Washington Kenblue
16	85	Alaska (12), Poland (7), Turkey (7), Sweden (6), Russia (4), and 49 others represented by countries with one to three accessions each; cultivars Plush, Park, Mystic, Nublue, Julia, Ikone, Coventry, Dawn, and Eclipse
17	1	cultivar Monopoly
18	2	Denmark (1) Hungary (1)
19	7	Iran (5) Afghanistan (1) (includes three *P. trivialis*, two *P. trivialis* ssp. *sylvicola*, with one *P. trivialis* with some *P. pratensis* contamination); Switzerland (1) (questionable origin)

and less yield then the other major cluster 14, 15 and 16 (Table 4). It included the cultivars Midnight and Bartitia, which had the highest color ratings of any of the commercial checks. Cluster 6 appeared to be a source of germplasm with high turf quality but without the desirable characteristic of average to high seed yield.

For the agronomic data, the first three principal components explained 62% of the total variation. The first principal component explained 31%, the second 16%, and the third 15% of the total variation. This results from of a degree of correlation among the agronomic factors (Manly 1986) as was observed previously by Johnston et al. (1997). Plots of the first two principal components showed the general separation of the clusters (Figure 3). The misidentified species from clusters 1, 2 and 3 were distinct from other clusters. For the four major agronomic clusters (6, 14, 15 and 16), the 95% confidence intervals for the first two principal components showed that they were easily distinguished (Figure 3). But cluster 19, containing the majority of the *P. trivialis* accessions, was very similar to cluster 16 when plotted in two dimensions. Unlike other clusters, the differences between clusters 19 and 16 were not observable until principal components 3 to 17 were examined.

Selection pressure provided by different environ-ments should in time result in natural selection and a differentiation of certain characteristics in a population (Frankel et al. 1995). The extent to which this occurs depends on the reproductive biology of the species, population genetic variation, and selection pressure. Thus, geographic origin of accessions might be expected to represent different macro-environments and thus explain much of the variation among populations from different regions. However, the cluster groupings were generally not strongly associated with particular geographic origins, especially for the RAPD data (Table 2).

There are a number of potential reasons why origin was not strongly associated with variation in clustering patterns in this study. First, country was the basis of most origin data. This is obviously a very broad geographic organizer and numerous environments could be expected within an origin. If well defined ecological and climatic zones could have been identified then the chances of relating these areas to the clustering pattern would have increased (Greene et al. 1999). For example, various climatic and soil factors were correlated with RAPD marker variation in wild emmer [*Triticum dicoccoides* (Korn. ex Asch. & Graebn.) Schweinf.] populations (Fahima et al. 1999). The demarcation of climate zones is becoming increasingly feasible as accurate longitude and latitude

Table 4. Mean values (when n > 1) for accessions within agronomic clusters in the USDA *Poa pratensis* L. collection. Species found misidentified and mistakenly placed in the *P. pratensis* collection are *P. urssulensis* (cluster 1), *Poa compressa* (cluster 2), *Poa glauca* (cluster 3), *Poa trivialis* (cluster 4 and 19), and *Poa alpina* (cluster 13).

Cluster	n	Emergence— days	Green- up[a]	Leaf texture[a]	Leaf color[a]	Turf potential[a]	Leaf habit[a]	Dwarf character[a]	Plant uniformity[a]	Canopy height— cm	MSMA burn[b]	Powdery mildew[b]	Heading date[c]	Date 50% anthesis[c]	Harvest date[c]	Harvest height— cm	Biomass[a]	Seed yield— g m^{-2}
1	1	10.0	5.33	5.67	6.00	6.20	6.67	7.33	4.33	11.3	0.670	1.00	145	176	200	60	3.33	20.5
2	1	11.3	4.33	4.33	4.33	3.67	4.00	2.00	8.00	17.0	0.000	0.00	145	175	222	55	6.67	46.8
3	1	18.0	7.00	8.67	5.67	6.43	7.00	9.00	3.00	3.7	0.000	0.00	115	142	208	42	1.67	15.3
4	1	10.0	5.00	6.50	4.50	4.85	4.50	4.33	3.67	12.7	2.770	0.00	138	151	177	65	4.00	19.5
5	2	16.5	4.67	5.75	5.50	5.43	5.25	5.92	2.17	6.9	2.050	0.00	120	147	190	73	2.42	24.4
6	21	16.1	4.14	6.28	6.59	6.44	6.26	6.58	5.17	12.4	0.142	0.57	129	148	191	46	4.14	23.6
7	5	12.1	7.57	6.31	7.75	7.34	6.93	6.60	6.57	12.6	0.400	0.73	121	147	184	77	4.70	53.6
8	1	14.7	5.67	7.77	6.10	7.00	6.67	6.00	7.33	18.3	0.330	2.00	118	145	177	66	5.00	27.3
9	1	16.0	6.00	4.33	4.33	4.57	5.00	2.33	2.33	19.3	0.670	1.67	115	140	181	89	4.00	13.9
10	1	17.0	5.67	4.00	7.00	6.00	6.67	4.33	6.00	26.0	0.000	0.00	135	150	180	80	5.33	77.5
11	6	13.3	4.45	5.63	6.28	5.91	5.72	4.06	7.06	20.2	0.000	2.57	137	153	182	76	6.33	42.8
12	5	13.1	6.40	4.33	5.13	4.05	3.53	2.53	5.60	14.4	0.354	0.27	118	148	177	98	5.07	100.8
13	1	10.0	5.00	3.67	5.43	4.33	4.00	5.00	7.67	16.3	0.000	0.00	118	145	174	46	5.00	76.8
14	47	15.8	5.44	5.75	5.84	5.37	4.73	4.12	4.89	13.7	0.206	0.11	118	144	181	70	4.61	45.2
15	56	13.8	6.37	5.97	5.75	5.70	5.41	4.49	6.40	18.0	0.035	0.72	119	148	177	91	5.44	87.8
16	85	14.6	5.50	5.44	5.33	4.90	4.38	2.99	6.15	19.1	0.058	1.53	124	150	181	85	5.95	53.1
17	1	12.7	4.67	4.00	4.33	3.77	4.00	1.67	8.00	22.3	0.000	0.00	138	158	182	79	7.67	87.2
18	2	16.0	4.17	3.84	3.83	3.50	3.50	1.83	5.17	21.8	0.165	0.50	137	153	187	81	5.17	33.4
19	7	10.2	5.57	4.81	4.26	4.57	5.12	4.17	5.52	20.7	0.619	0.05	138	151	175	87	5.43	34.5
LSD(0.05)		3.33	1.72	1.51	1.28	1.12	0.23	2.27	1.92	5.83	0.619	1.26	10.7	5.4	7.4	20.6	1.44	44.0

[a] Rated from 1 to 9: 9 = finest texture, darkest green leaf color, most plant uniformity, most upright leaf habit, most dwarf character, most biomass, highest turf potential, and earliest spring green-up.

[b] Rated from 0 to 3: 0 = no MSMA toxicity and no powdery mildew.

[c] Days from January 1, 1995.

[d] LSD values require multiplication by $[(n_1 + n_2)/n_1 n_2]^{1/2}$ where n_1 and n_2 are the number of accessions in the two cluster to be compared.
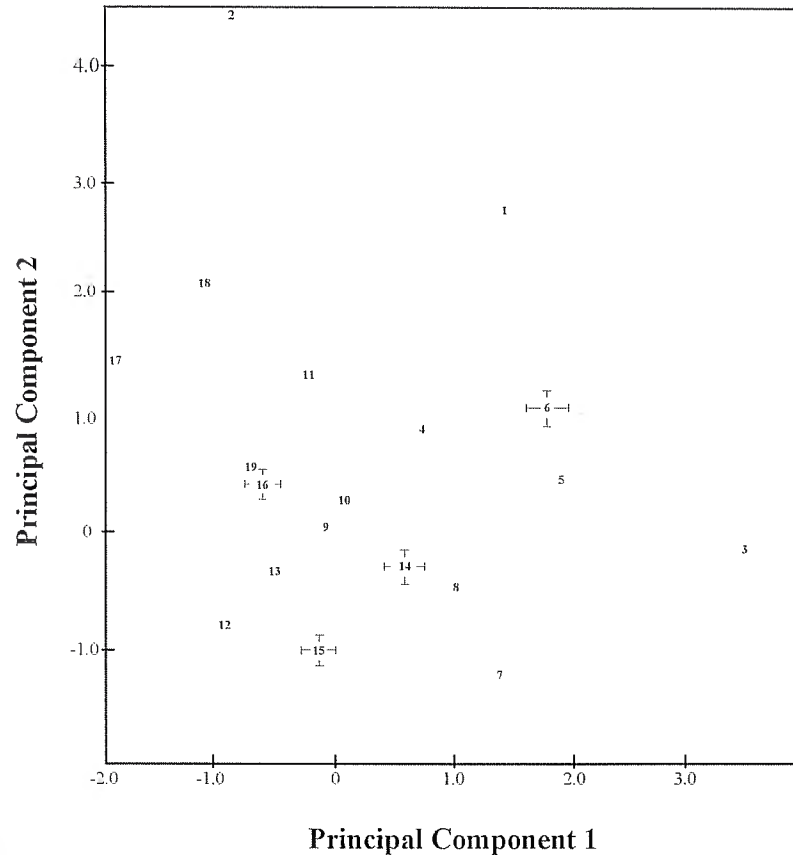
Figure 3. Plot of the first two principal components for each cluster derived from agronomic data based on 17 descriptors. For the four major clusters, which contained 89% of the *P. pratensis* accessions, a 95% confidence interval is shown.

coordinates are routinely obtained during field collection. For most accessions in this study, however, this type of data was not available.

Multiple transfer and introduction of material can also compromise origin data. For example, there were two sets of accessions in the collection, other than the commercial checks, that had the same cultivar names and as such were expected to be duplicates. In GRIN, these accessions were PI 266209 and PI 303056 designated as 'Prato', and accessions PI 303054 and PI 311085 designated as 'Golf'. For the pair of accessions designated as Prato the RAPD similarity was 0.80 and for the accessions designated as Golf the similarity was 0.87. Although these similarity values are relatively high there were many entries with much closer similarity values. Our expectation was that these pairs would be highly similar, but this was not the case. This lack of similarity was also observed in the agronomic data. The average Euclidian distance,

the reverse of similarity, was 1.30 between Prato accessions and 0.89 between Golf accessions. Thus, in both RAPD and morphological data, neither the set of Prato nor Golf accessions was highly similar. And in both RAPD and agronomic data, the two Prato accessions were less similar than the two Golf accessions. Thus, it cannot be assumed that accessions with the same identifying name, but entering a gene bank at different times and from different sources, are duplicates. Prato PI 266209 was originally introduced from the Netherlands in 1960, but later reintroduced as part of a seed collection donated by the Seed Exchange Office from Rome, Italy in 1964 and given PI 303056. Golf was originally introduced as part of the material coming from Italy in 1964 along with Prato, and then reintroduced in 1966 as part of a collection from Rumania at the Brasov Experiment Station. Although the reasons for the lack of duplication observed in these cases cannot be determined,

possibilities include misidentification, seed mixes, or selection processes, either natural or through human activity.

Aside from the need for accurate origin and geographic information, the cosmopolitan nature of *P. pratensis* could frustrate attempts to relate geography to clustering pattens. Most agree that the genus *Poa* originated in Eurasia (Hartley 1961; Soreng 1990). Although Carrier and Bort (1916) believe *P. pratensis* was introduced to the Americas, some suggest it may have been native to North America (Fernald 1950). Regardless of the origin, its wide adaptation has lead to a circumglobal distribution, spreading by natural means and by human activity to different temperate environments. Natural and human activity in Eurasia and the Americas is no doubt important in its distribution and its spread. For example, *P. pratensis* has been planted along many roads in Alaska, and roadsides were common collection sites for the Alaska accessions in this study. The process of natural spread and human transfer is ongoing, and this potential for introduction could have easily lead to variation among accessions within regions of origin and collection site areas.

Hybridization could have also lead to genetic variation within a region. Usually more than 90% of *P. pratensis* progeny are genetically identical through apomixis (Åkerberg 1939). But as a facultative apomict, sexual reproduction is a mechanism for introgression of new genetic variation and fixation of adaptive genes. Associated with this facultative apomixis, the chromosomes number in *P. pratensis* is highly variable (2n = 28 to ± 124) (Åkerberg 1942; Huff and Bara 1993). Sexual reproduction most often leads to aberrant individuals that are less fit than progeny resulting from apomixis, but some genotypes from sexual reproduction can show high fitness and high apomictic reproductive capacity (Clausen 1961).

In this study, high diversity was often noted in material collected within small geographic areas. For example, cluster 14 was rich in accessions from the Kenai, Alaska area (38%), and cluster 6 was rich in accessions from the Hope, Alaska area (48%). Both these locations are on the Kenai peninsula along Cook Inlet. Even though the clusters were agronomically distinct (Table 4), there were some accessions from Kenai in cluster 6 and some accessions from Hope in cluster 14. It is perhaps more logical to expect considerable variation within small geographic areas rather than homogeneity, especially as the processes of introduction of genotypes and hybridization are ongoing.

When this study was initiated there were approximately 286 accessions in the *P. pratensis* collection of which 228 were available for evaluation. There are now a total of 62 additional accessions listed as 'active' in GRIN for total of 348. Of the new accessions, the majority are cultivars originating from the U.S. Recent germplasm collections made in Mongolia, Ukraine, Turkey and China are in the process of seed regeneration and will be available soon. However, given the importance of *P. pratensis* for forage and turfgrass applications, the collection seems to be generally under represented by many key regions. Other than Alaska, most countries are represented in the collection by 10 or fewer accessions, indicating a general lack of a concentrated collection effort to date. Although the geographic origin of the genus *Poa* is difficult to determine, centers of species differentiation were identified by Hartley (1961) as the mountainous regions from northen Pakistan, India, and Burma to Kazakhstan and Mongolia. Given that only 29 accessions of *P. pratensis* are currently listed in GRIN as originating from Central Asia, additional germplasm collection appears to be needed. Other than Alaska, there is also a general lack of representation of wild or naturalized accessions from the United States, and also from Europe.

Using all 245 accessions and check cultivars, the distance matrix correlation coefficient between values of average Euclidean distance (agronomic data) and the Dice similarity coefficient (RAPD data) (r = − 0.32) was highly significant (P < 0.01). (The negative r value results when a distance matrix and a similarity matrix are correlated.) Although considerable variation was not explained by the correlation, the agronomic data had a degree of correspondence with the RAPD data. When the misidentified accessions were removed from the matrix correlation (n = 234), the correlation coefficient was reduced from −0.32 to −0.14. Even though the −0.14 value was still highly significant according the Mantel test, it showed that a good deal of the association between the RAPD data and the agronomic data could be attributed to the presence of species other than *P. pratensis*.

This study showed that RAPDs were useful in identifying misidentified species and in estimating diversity among accessions and cultivars of *P. pratensis*. The majority of accessions fell into one cluster at a Dice similarity of 0.8, suggesting that unique genotypes in the collection are generally under represented. This was also suggested by the agronomic data as 89% of the *P. pratensis* accessions fell into just four major clusters. Although the correlation between

the RAPD and agronomic data matrices was highly significant, it was relatively weak, showing that both molecular and agronomic characterization was needed for a comprehensive assessment of *P. pratensis* diversity.

Acknowledgements

References

Åkerberg E. 1939. Apomictic and sexual seed formation in *Poa pratensis*. Hereditas 25: 359–370.

Åkerberg E. 1942. Cytogentic studies in *Poa pratensis* and its hybrid with *Poa alpina*. Hereditas 28: 1–125.

Bashaw E.C. and Funk C.R. 1987. Apomictic grasses. In: Fehr W.R. (ed.), Principles of Cultivar Development Vol. 2. Macmillan Publishers, New York, NY Crop Species. Crop Species., pp. 40–82.

Carrier L. and Bort L.S. 1916. The history of Kentucky bluegrass and white clover in the United States. J. Amer. Soc. Agron. 8: 256–266.

Clausen J. 1961. Introgression facilitated by apomixis in polyploid Poas. Euphytica 10: 87–94.

Fahima T., Sun G.L., Beharav A., Krugman T., Beile A. and Nevo E. 1999. RAPD polymorphism of wild emmer wheat populations, *Triticum dicoccoides*, in Israel. Theoret. Appl. Genetics 98: 434–447.

Frankel O.H., Brown A.D.H. and Burdon J.J. 1995. The conservation of plant biodiversity. Cambridge Univ. Press, Cambridge, UK.

Fernald M.L. 1950. Gray's manual of Botany. 8th edn. American Book C., New York.

Gower J.C. 1966. Some properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325–338.

Greene S.L., Hart T.C. and Afonin A. 1999. Using geographic information to acquire wild crop germplasm for *ex situ* collections: II. Post-collection analysis. Crop Sci. 39: 843–849.

Hartley W. 1961. Studies on the origin, evolution, and distribution of the Gramineae IV. The genus *Poa* L. Aust. J. of Bot. 9: 156–161.

Hayward M.D., Mcadam N.J., Jones J.G., Evans C., Evans G.M., Forster J.W. et al. 1994. Genetic markers and the selection of quantitative traits in forage grasses. Euphytica 77: 269–275.

Huff D.R. 1997. RAPD characterization of heterogenous perennial ryegrass cultivars. Crop Science 37: 557–564.

Huff D.R. and Bara J.M. 1993. Determining genetic origins of aberrant progeny from apomictic Kentucky bluegrass using a combination of flow cytometry and silver stained RAPD markers. Theoret. Appl. Genetics 87: 201–208.

Johnston W.J., Nelson M.C., Johnson R.C. and Golob C.T. 1997. Phenotypic evaluation of *Poa pratensis* L.: USDA/ARS Plant introduction germplasm collection. Int. Turfgrass Soc. Res. J. 8: 305–311.

Manly B.F.J. 1986. Multivariate statistical methods: A primer. Chapman and Hall, London, UK.

Martin C., Juliano A., Newbury H.J., Lu B.-R., Jackson M.T. and Ford-Lloyd B.V. 1997. The use of RAPD markers to facilitate the identification of *Oryza* species. Genet. Resour. Crop Evol. 44: 175–183.

Morrison K.J. and Law A.L. 1978. Grass seed production. Washington State University, Pullman, WA, USA, Agricultural Extension Bulletin 704.

Murray M.G. and Thompson W.F. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8: 4321–4324.

Rodriguez J.M., Berke T., Engle L. and Nienhuis J. 1999. Variation among and within *Capsicum* species revealed by RAPD markers. Theor. Appl. Genet. 99: 147–156.

Romesburg H.C. 1990. Cluster analysis for researchers. Krieger Publishing, Malabar, FL, USA.

SAS Institute 1985. User's Guide: Statistics. 5th edn. SAS Institute, Cary, NC.

Šiffelová G., Pavelková M., Klabouchová A., Wiesner I. and Našinec V. 1997. Computer-aided fingerprinting of accessions from the ryegrass-fescue complex. Journ. of Agric. Sci., Cambridge 129: 257–265.

Soreng R.J. 1990. Chloroplast-DNA Phylogenetics and biogeography in a reticulating group: Study in *Poa* (Poaceae). Amer. J. of Bot. 77: 1383–1400.

Strelchenko P., Kovalyova O. and Kazutoshi O. 1999. Genetic differentiation and geographical distribution of barley germplasm based on RAPD markers. Genet. Resour. Crop Evol. 46: 193–205.

Sweeney P.M. and Danneberger T.K. 1995. RAPD characterization of *Poa annua* L. populations in golf course greens and fairways. Crop Sci. 35: 1676–1680.

Virk P.S., Newbury H.J., Jackson M.T. and Ford-Lloyd B.V. 1995. The identification and duplicate accessions within a rice germplasm collection using RAPD analysis. Theor. Appl. Genet. 90: 1049–1055.

Weising K., Nybom H., Wolff K. and Meyer W. 1995. DNA fingerprinting in Plants and Fungi. CRC Press, Boca Raton, FL, USA.

Williams J.G.K., Kubelik A.R., Livak K.J., Rafalski J.A. and Tingey S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531–6532.