

Research Article

Advances towards a Marker-Assisted Selection Breeding Program in Prairie Cordgrass, a Biomass Crop

K. R. Gedye, J. L. Gonzalez-Hernandez, V. Owens, and A. Boe

Department of Plant Sciences, South Dakota State University, Brookings, SD 57007, USA

Correspondence should be addressed to J. L. Gonzalez-Hernandez, jose.gonzalez@sdstate.edu

Received 3 August 2012; Accepted 29 October 2012

Academic Editor: Shizhong Xu

Copyright © 2012 K. R. Gedye et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prairie cordgrass (*Spartina pectinata* Bosc ex Link) is an indigenous, perennial grass of North America that is being developed into a cellulosic biomass crop suitable for biofuel production. Limited research has been performed into the breeding of prairie cordgrass; this research details an initial investigation into the development of a breeding program for this species. Genomic libraries enriched for four simple sequence repeat (SSR) motifs were developed, 25 clones from each library were sequenced, identifying 70 SSR regions, and primers were developed for these regions, 35 of which were amplified under standard PCR conditions. These SSR markers were used to validate the crossing methodology of prairie cordgrass and it was found that crosses between two plants occurred without the need for emasculation. The successful cross between two clones of prairie cordgrass indicates that this species is not self-incompatible. The results from this research will be used to instigate the production of a molecular map of prairie cordgrass which can be used to incorporate marker-assisted selection (MAS) protocols into a breeding program to improve this species for cellulosic biomass production.

1. Introduction

Recent world issues associated with fuel consumption and supply have turned attention towards biofuel production, especially cellulosic biofuel. Perennial grasses provide an optimal source of cellulosic biomass due to their high yield potential. Prairie cordgrass (*Spartina pectinata* Bosc ex Link) is a perennial indigenous grass of North America and can be found as a native from Texas to near the Arctic Circle [1]. Ongoing studies on prairie cordgrass in comparison with switchgrass (*Panicum virgatum* L.) indicate that prairie cordgrass could produce more biomass than switchgrass [2]. Furthermore, results from the comparison of prairie cordgrass and switchgrass performed by Boe and Lee in 2007 [2] indicated that prairie cordgrass has a wider environmental amplitude and is adapted to poorly drained wet areas which can have high salinity and be poorly aerated, regions not suitable for the production of conventional crops such as maize (*Zea mays*) [2, 3]. These results are indicative of the potential of prairie cordgrass as a source of biomass for

cellulosic biofuel production. A research program at South Dakota State University (SDSU) is underway to develop native prairie cordgrass into a viable cellulosic biomass crop. The development of a new crop species requires a multidisciplinary approach; examining and validating each step before commercialization can occur. These steps include, but are not limited to, the assembling of a germplasm collection, the development of an accelerated breeding program, and the optimization of cultivation practices. A fundamental step to accelerate the breeding of prairie cordgrass is to determine and optimize a crossing protocol. prairie cordgrass has been intrinsically believed to be a protogynous outcrossing species, based on the mode of reproduction of its maritime relative smooth cordgrass (*Spartina alterniflora*) [4, 5]. From work performed upon other members of the *Spartina* genus, it has been conjectured that prairie cordgrass will have a similar method of reproduction. In the majority of other Gramineae species, breeding is performed via the initial emasculation of the floret to ensure that only cross

pollination can occur. Prairie cordgrass has an inflorescence composed of between 0 and 31 short paracladia and 11–13 long paracladia [6], bearing a total of 10–80 fertile spikelets, of single florets [7]. Physical emasculation of prairie cordgrass is essentially impractical, but this technique may not be necessary as with protogyny and ascertaining the appropriate timing, directed cross pollination is feasible. Self pollination has not been identified previously in prairie cordgrass and assumptions have been made that prairie cordgrass may be self incompatible, if this is the case then the classic mapping technique of producing recombinant inbred lines will be fundamentally impossible and alternate strategies will need to be utilized.

An initial stage of the prairie cordgrass project is to develop a molecular map of the species. Only limited molecular analyses have been performed upon prairie cordgrass, notably in contrast to its maritime relative *S. alterniflora*. In the National Centre for Biotechnology Information (NCBI), only 57 prairie cordgrass sequences have been deposited (predominantly regions of the nuclear and organelle genomes utilized for diversity analysis, that is, *Waxy* (AY508655 and AF372461), ITS (AF019843, AJ489796, and EF153082) and the chloroplast *trnL-trnF* intergenic spacer region (EF137568, EU056305, and AF372625)). A recent publication on the analysis of the transcriptome of prairie cordgrass has increased the available knowledge on the genome of the species [8]. Constructed from the analysis of the expression sequence tags (ESTs) and identified from the transcriptome of prairie cordgrass, a total of 26,302 contigs and 71,103 singletons were assembled with all sequence information available as supplemental data [8]. Additional molecular analysis of genetic diversity of prairie cordgrass in natural populations in Minnesota with amplified fragment length polymorphism (AFLP) has been performed [9].

An optimal marker system for the development of a molecular map of prairie cordgrass are SSRs or microsatellites. Microsatellites are highly utilized molecular markers and have been developed for the majority of agronomically and economically important crop species [10]; their efficacy arises predominantly due to their reproducibility, codominant inheritance, and abundance [11]. Modern techniques allow the rapid identification of microsatellite regions in species which have limited sequence information, and by using proprietary genomic library screening techniques, microsatellites have been developed for numerous non-model organisms, encompassing insects [12], birds [13], fish [14], mammals [15], and plants, including other species in the *Spartina* genera [16]. Previous studies on other members of the *Spartina* genera have developed 35 markers for microsatellite loci in *S. alterniflora* [16, 17]. Of these *S. alterniflora* markers, three have been found to amplify in prairie cordgrass [16]. SSRs have been used extensively in numerous crop species, in wheat (*Triticum aestivum* L.), and another Poaceae species; SSRs have been used to enhance molecular maps begun with restriction fragment length polymorphism (RFLP) and to identify genes of interest [18]. Furthermore, Wilde et al. in 2007 [19] found that the incorporation of MAS with SSR markers into a traditional

breeding program can result in a substantial increase in the incorporation of *Fusarium* head blight resistance, a quantitative trait, in wheat with the shortest possible time. The results of Wilde et al. [19] are indicative of the potential of MAS for the breeding of prairie cordgrass.

This paper details the first investigation into initial breeding and crossing of prairie cordgrass, specifically the production of F₁ individuals and the validation of a crossing protocol. Furthermore, the characterization of 35 microsatellite loci in prairie cordgrass from genomic DNA is discussed. The validation of the microsatellite loci occurred in a prairie cordgrass germplasm collection and a reciprocal cross as an initial step in the development of a molecular map, which can be further utilized during marker-assisted selection of lines with traits desirable for the improvement of prairie cordgrass as a biofuel crop.

2. Materials and Methods

2.1. Genomic Library Construction. Genomic DNA was extracted from four lines of prairie cordgrass using the method as described by Karakousis and Langridge in 2003 [20], with minor modifications. A mixture of the four random prairie cordgrass clones DNA, totaling > 100 µg, was sent to Genetic Identification Services (GIS) (<http://www.genetic-id-services.com/>) for the development of genomic libraries. The genomic libraries were enriched for the four simple sequence repeat (SSR) motifs (CA)_n, (GA)_n, (AAG)_n, and (CAG)_n. Subsamples of 100 clones were sequenced and primers were designed to flank the SSR regions using DesignerPCR, version 1.03 (Research Genetics, Inc.) and PRIMER 3 [21]. The SSR regions were classified according to Jones et al. [22] as being pure repeats (i.e., [N₁N₂]_X), compound repeats (i.e., [N₁N₂]_X[N₃N₄]_Y), and interrupted repeats (i.e., [N₁N₂]_XN₃[N₄N₅]_Y).

2.2. Plant Material. Wild germplasm has been collected from throughout the mid-western states of the United States, creating a core germplasm collection to be utilized in prairie cordgrass breeding (unpublished data). A geographically diverse sample of sixteen plants, collected from South Dakota, North Dakota, Minnesota, and Iowa, was characterized using the identified SSR loci. A sample from the closely related species *S. spartinae* was also included to examine cross species amplification. The germplasm collection was grown in standard greenhouse conditions.

Crosses between and amongst these genotypes were performed. Crossing between two plants of prairie cordgrass was performed in the following manner; two inflorescence at the appropriate stage of development were placed inside a crossing bag for seven days, producing F₁ plants (Figure 1). A reciprocal cross between two genotypes (designated RR2 and RR21) from the Red River morphotype was produced. The reciprocal cross produced 45 F₁ plants from the RR21 × RR2 directed cross and 49 F₁ plants from the RR2 × RR21 directed cross, a total of 94 plants. Two genotypes identified as clones (designated SP1.1B and SP1.2A) from the SSR evaluation were crossed via the same technique, producing a total of



FIGURE 1: Two prairie cordgrass heads undergoing crossing, contained within a single bag.

46 F_1 individuals; a total of 20 F_1 plants were produced when SP1.1B was used as the maternal parent and 26 F_1 plants when SP1.2A was used as the maternal parent. All F_1 plants and their respective parents were grown in standard greenhouse conditions.

2.3. DNA Extraction and Evaluation of SSR Primers. DNA was extracted from all examined plants using the method as described by Karakousis and Langridge in 2003 [20], with minor modifications. Evaluation of the primers was performed in a PCR reaction consisting of 2 U *Taq* DNA polymerase (Promega GoTaq), 1 × PCR Buffer (Promega GoTaq), 2 mM $MgCl_2$, 0.6 mM dNTPs, and 0.6 mM of each primer, with a final volume of 20 μ L [23]. PCR reactions were performed in a BioRad MyCycler (BioRad, Hercules, CA). Initially a gradient of annealing temperatures was used to determine optimal T_a , as these primers were designed to be utilized in a prairie cordgrass breeding program, a robust and high-throughput protocol was developed. Specific thermocycling conditions for the primers used in this study were as follows: an initial denaturation at 94°C for 5 minutes, followed by 35 cycles of 94°C for 1 minute, 53 or 55°C for 1 minute, 72°C for 1 minute, followed by an extension step of 72°C for 10 minutes, and a 10°C hold. PCR product was visualized on 8% nondenaturing PAGE; bands were scored and sized by comparison to a 100 bp ladder using AlphaEaseFC Software (Alpha Innotech, San Leandro, CA).

2.4. SSR Markers Analysis. Polymorphic bands in the two examined populations described above were scored on a presence (1) or absence (0) basis. An estimation of genetic distance was calculated with Nei and Li's algorithm [24] and the resulting matrix was clustered with unweighted pair group method with arithmetic mean (UPGMA) [25]. Analysis was performed using multivariate statistics package (MVSP) [26].

3. Results

3.1. Germplasm and Crossing. The crossing of the germplasm collection was successful with the production of a total of 14,813 putatively viable seeds from 110 crosses, performed in lines and collected from 48 distinct collection locations (Table 1). Numerous nonviable seeds were produced from each cross. A seed was determined to be viable if the actual seed (endosperm and embryo) was visible within the glumes (Figure 2). The number of viable seeds produced per head varied from 6 to 642, with an average of 134 seeds per cross. The crosses can be grouped into two, dependent upon the parents, those that were outcrossed and those that were selfed. Crosses that were designated “out” were between genotypes that were geographically diverse, while crosses designated “self” were produced from members of the population collected at the same location. The average number of seeds produced from outcrossed individuals was 143, with a range of 12 to 250, while selfed individuals produced an average of 112 seeds per cross, with a range of 6 to 642 (Figure 3).

3.2. Microsatellite Characterization. Of the total 100 genomic clones sequenced, 70 contained microsatellite motifs, 26 from the $(CA)_n$ enriched library, 25 from the $(GA)_n$ enriched library, 3 from the $(AAG)_n$ enriched library, and 16 from the $(CAG)_n$ enriched library. All 70 loci were examined and 35 were amplified with the standardized conditions (Table 2). The repeat SSR structure that occurred most frequently was the pure repeat (27), followed by the compound repeat (7) and only one interrupted repeat sequence was found (Table 2).

Only two primers produced monomorphic profiles from the analysis of the sixteen lines, SPSD004 from the $(GA)_n$ library and SPSD048 from the $(AAG)_n$ library. The remaining primers produced between 1 and 12 scorable bands. Of the examined primers, 11 were amplified between 1 and 7 bands in *S. spartinae* (Table 2).

All sequences were examined using BLASTn [27] on the National Center for Biotechnology Information (NCBI) server for similarities to other recorded sequences. Only two of the prairie cordgrass sequences displayed homology ($e \leq e^{-20}$) to recorded sequences, SPSD003 to six sequences from rice, wheat, and field mustard (*Brassica rapa*) (accession numbers: AP008214.1, AP005495.3, NM_001067901.1, AK111788.1, EU660901.1, and AC189183.2) and SPSD050 to five sequences all from rice (accession numbers: CR855236.1, AP008210.1, AL606607.3, NM_001059819.1, and AK108706.1) [25].

3.3. Crossing Population. A total of 94 individuals from the RR2 × RR21 and the RR21 × RR2 crosses were examined with 35 SSR primers. The amplified bands were scored based on presence or absence, the resulting data was analyzed with Nei and Li's coefficient producing a similarity matrix (data not shown), and the data was clustered with UPGMA producing a dendrogram (Figure 4). The results of the dendrogram indicate that each individual was a cross

TABLE 1: Summary of crosses performed amongst the core prairie cordgrass germplasm collection. The type of cross is designated “Out” for crosses between geographically diverse members and “Self” for members of the population collected at the same location.

Female parent	Male parent	Number of crosses	Type	Total no. of seeds
SP1	SP21	2	Out	126
SP1	SP71	1	Out	71
SP21	SP30	1	Out	51
SP21	SP31	6	Out	484
SP21	SP4	2	Out	248
SP21	SP42	1	Out	78
SP21	SP66	2	Out	233
SP21	SP7	1	Out	82
SP24	SP30	1	Out	180
SP29	SP22	1	Out	146
SP29	SP30	1	Out	42
SP3	SP77	1	Out	282
SP30	SP31	1	Out	164
SP30	SP52	1	Out	246
SP31	SP21	4	Out	408
SP31	SP30	1	Out	28
SP31	SP32	1	Out	38
SP31	SP46	3	Out	211
SP31	SP5	1	Out	81
SP31	SP66	1	Out	46
SP31	SP79	1	Out	198
SP32	SP21	1	Out	173
SP32	SP31	2	Out	310
SP32	SP45	1	Out	296
SP32	SP66	1	Out	240
SP32	SP67	1	Out	193
SP32	SP78	1	Out	94
SP35	SP78	1	Out	279
SP41	SP42	1	Out	6
SP41	SP45	1	Out	86
SP44	SP54	1	Out	312
SP44	SP59	1	Out	274
SP45	SP31	1	Out	282
SP45	SP49	1	Out	225
SP45	SP67	1	Out	138
SP49	SP40	1	Out	203
SP49	SP45	1	Out	176
SP49	SP66	2	Out	237
SP5	SP31	1	Out	189
SP5	SP78	1	Out	407
SP53	SP67	2	Out	43
SP54	SP44	1	Out	145
SP54	SP59	2	Out	81
SP56	SP58	1	Out	235
SP65	SP57	1	Out	70
SP66	SP31	1	Out	93
SP67	SP32	1	Out	99

TABLE 1: Continued.

Female parent	Male parent	Number of crosses	Type	Total no. of seeds
SP67	SP45	2	Out	583
SP67	SP53	1	Out	185
SP7	SP46	1	Out	160
SP72	SP1	1	Out	125
SP72	SP30	2	Out	459
SP75	SP13	1	Out	59
SP77	SP72	1	Out	163
SP77	SP79	1	Out	184
SP78	SP35	1	Out	94
SP78	SP79	1	Out	113
SP82	SP14	1	Out	642
SP21	SP21	13	Self	1470
SP22	SP22	1	Self	110
SP31	SP31	13	Self	1377
SP40	SP40	1	Self	126
SP41	SP41	2	Self	393
SP53	SP53	1	Self	12
SP74	SP74	2	Self	229

between the two parents. When the population derived from the cross between the two plants (SP1.1B and SP1.2A) and determined to be clones was examined, bands present in both parents were found to segregate in the progeny (data not shown), indicative of sexual recombination; this result provides validation that successful crossing occurred.

4. Discussion

The primary requirement of any breeding program is to ensure that accurate crosses are made; in many other members of the Poaceae this is achieved by physical or chemical emasculation. The prolific numbers of flowers per head in addition to the small size of the flowers make physical emasculation unfeasible. Furthermore, due to limited knowledge about the nature of the fertility of this species, chemical emasculation has not been developed for prairie cordgrass. The results from the SSR analysis indicate that utilizing the inherent protogyny of prairie cordgrass allows successful crossing between two individuals without the need for emasculation, confirming the validity of the breeding methodology used. The presence of individuals in the F₁ mapping populations which show limited genetic dissimilarity from the parents could be evidence of selfing; further investigations are required. The presence of these potential selfed individuals indicates that future breeding and/or mapping populations should be examined with the molecular markers devised in this research to remove suspect individuals. Subsequently, the successful crossing between two clones is indicative that prairie cordgrass may not be self-incompatible and that it may be possible to develop in this species conventional mapping populations, such as recombinant inbred lines. Further studies into self-compatibility



(a)



(b)

FIGURE 2: Prairie cordgrass seeds photographs taken over a light box. Illuminating prairie cordgrass with light box is used to distinguish the presence of a developed embryo in the seeds. (a) Non-viable seeds. (b) Viable seeds.

with investigations into potential apomictic prairie cordgrass plants are underway. The number of seeds observed in this research appears to be larger than what was previously described by Clayton et al. [7]. The variation between the two studies can be attributed to both environmental and genetic variations. Genetic variation in seed set in prairie cordgrass, although at this stage not quantified, is demonstrated by the range in viable seed set observed in this research.

The amplification of *S. spartinae* with the SSR primers developed in this research are indicative of the potential colinearity amongst the genomes of *Spartina* spp. and other grass species; this colinearity will allow easier identification and characterization of genes. The colinearity between prairie cordgrass and other Poaceae is currently utilized

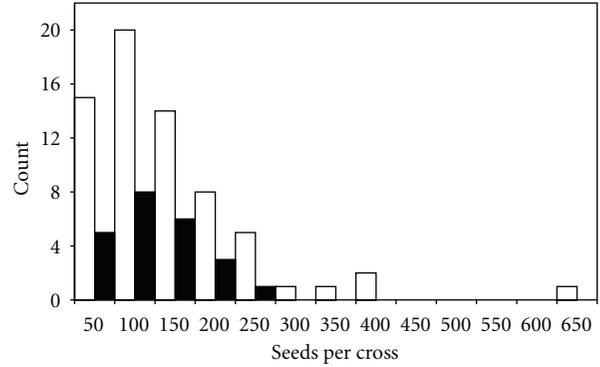


FIGURE 3: Summary graph of results of crosses between members of the core germplasm collection of prairie cordgrass. White bars indicate self crosses and black bars indicate out crosses.

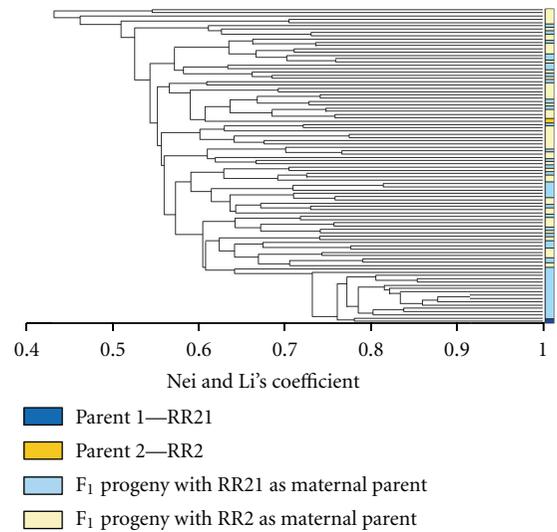


FIGURE 4: Unrooted dendrogram clustered with UPGMA of the genetic association of the F₁ progeny from a cross between prairie cordgrass lines RR21 and RR2, from the analysis of SSR regions with Nei and Li's Coefficient of genetic similarity.

to examine genes identified in related species in prairie cordgrass, specifically genes utilized to examine phylogeny (i.e., the *waxy* gene for granule bound starch synthase).

The prevalence and distribution of SSR regions across plant genomes are extremely variable. Variation in SSRs is not limited to their location, but also their motif, putative function, abundance, and repeat number [28]. The results of this analysis indicate that the two dinucleotide repeats (CA)_n and (GA)_n are more prevalent in prairie cordgrass (31% and 46%, resp.), than the two trinucleotide repeats (AAG)_n and (CAG)_n (6% and 17%, resp.). The prevalence of the dinucleotide motif in prairie cordgrass is similar to what was observed in the characterization of SSRs in other *Spartina* sp., where Blum et al. [16] found 82% of isolated SSRs contained dinucleotide repeats and Sloop et al. [17] found 71% containing similar motifs. In all three

TABLE 2: Summary and description of primer sequences designed for SSR regions identified in prairie cordgrass and their amplification products.

Locus	Library	Primer sequence (5'3')	SSR motif	Expected size (bp)	Size range (bp)	T_a (°C)	GenBank ID
SPSD001	(CA) _n	F: GCTGCTCCTCTTCTCTGTCT R: ACGGCACACTTAGTTTCTG	(CA) ₁₃ Perfect	180	212–386	55	GQ354531
SPSD002	(CA) _n	F: GCACTGTTTGGTGATGCTC R: CTGACGCAAGGTTGATGAG	(CA) ₁₃ Perfect	249	100–655	55	GQ354532
SPSD009	(CA) _n	F: ATGGTTTCACAAGTCGGAAGT R: CAGGGCTGCCTACAAGATG	(GT) ₁₅ Perfect	166	215–432	55	GQ354533
SPSD010	(CA) _n	F: AACCAAAAGGATAGACCCTA R: ACGAAATATGTGACCGATAC	(CA) ₂₈ Perfect	168	127–310	53	GQ354534
SPSD011	(CA) _n	F: CTAACGTATGTCGTTTCATGTGG R: AAGGCGATTTTAAGAGGCTAAG	(GT) ₁₁ Perfect	135	147–251	53	GQ354535
SPSD013	(CA) _n	F: GGGATGCTTTGTAGATAAGAAA R: TCTTCCTCTTACTCTGTCCACC	(CA) ₄₁ Perfect	157	85–173	55	GQ354536
SPSD016	(CA) _n	F: CCGACTACGAGCCACATT R: GTTCCACACATACGAAGGAGA	(CA) ₁₄ Perfect	155	156–268	53	GQ354537
SPSD019	(CA) _n	F: CCTGCTTACTCTTACTCCGTC R: ACCCTTTTTTCTTTTGGTCTC	(CA) ₁₃ (GA) ₇ Compound	134	220–365	53	GQ354538
SPSD020	(CA) _n	F: ATGAGACGATAGCAGGATGAC R: AGCAGATTACGATTCAGATGG	(CA) ₂₃ Perfect	178	146–288	55	GQ354539
SPSD025	(CA) _n	F: CAGTCCATGCAACTCAGAAGTA R: AACCTGATAGAAGTGGTCATGC	(CT) ₁₀ (CA) ₃₈ Compound	269	204–748	55	GQ354540
SPSD026	(CA) _n	F: GTGGAATCAACAACACCAGA R: GTCGCTTTAGCCCGTAAG	(GA) ₁₃ (GT) ₂₀ Compound	209	196–539	55	GQ354541
SPSD003	(GA) _n	F: ATGGAAACTGTCTGAACTGAC R: AGCAATAACCACAGAAGAGACC	(CT) ₂₈ Perfect	294	231–263	55	GQ354542
SPSD027	(GA) _n	F: TCAAACAATGGCGGAGAAG R: CTGGCTCCACCTCTTTGG	(CT) ₁₇ Perfect	214	188–224	55	GQ354543
SPSD004	(GA) _n	F: GTTGCTCGGTTCCAGTTG R: CGCCACACAAAAGTAGCC	(GA) ₂₂ Perfect	178	133 & 164	55	GQ354544
SPSD031	(GA) _n	F: TCGCACTTTTGATTCTCTTTAC R: TGGATGGATTAGGTTACTGTTG	(CT) ₂₀ Perfect	188	155–346	53	GQ354545
SPSD032	(GA) _n	F: CTCTCGCCCATGCTACTTA R: CCATTGCTATGTTGTTTGAGC	(CT) ₁₃ Perfect	196	146–193	53	GQ354546
SPSD034	(GA) _n	F: CAGGTCTACGGAGGTCACTAC R: TCAAAAAGAAGAGCACATACACA	(CT) ₁₂ Perfect	160	144–306	53	GQ354564
SPSD036	(GA) _n	F: TTCACCACACCACTTTATCC R: GGAAGCAACAAACATTGATG	(CT) ₂₉ Perfect	271	222–260	53	GQ354547
SPSD039	(GA) _n	F: CTTTCAGATAGTCCACTGATC R: AGCAATAACTGTGCATACCTCT	(GA) ₁₂ Perfect	206	193–448	55	GQ354548
SPSD040	(GA) _n	F: AATCGAAGTAGCAGACACCAAC R: CATGCGTTTTTCACTCATGTAG	(GA) ₁₁ Perfect	214	190–416	55	GQ354549
SPSD041	(GA) _n	F: CCCAACGATGATTCTCTTTG R: TCACGGTAACACGATTAGTCC	(GA) ₂₆ Perfect	135	104–180	53	GQ354550
SPSD042	(GA) _n	F: ACCTCCCACTCGTTGCTAC R: GCCATTGCTCTGTTGTTTG	(CT) ₁₉ Perfect	211	131–257	53	GQ354551
SPSD043	(GA) _n	F: GTTCAAATGCGAACAAATCAG R: ATTCGATCTCACATGCAACAC	(GA) ₃₀ Perfect	277	231–252	53	GQ354552

TABLE 2: Continued.

Locus	Library	Primer sequence (5'3')	SSR motif	Expected size (bp)	Size range (bp)	T_a (°C)	GenBank ID
SPSD044	(GA) _n	F: AGCTATATGACCCGAGACTGTG R: GGAATGGTCCCATACTTAATCC	(GA) ₂₀ Perfect	275	254–276	55	GQ354553
SPSD045	(GA) _n	F: AACGGAGGAAGTAATAAATAGC R: AGCACACACTAGCAAGGAC	(CT) ₁₆ Perfect	241	214–426	53	GQ354554
SPSD046	(GA) _n	F: CAGGTTTATCAGTGAAGACATC R: GAGGTTCTTAAAGGAACATAGC	(TA) ₉ (GA) ₁₂ Compound	278	57–426	53	GQ354555
SPSD047	(GA) _n	F: CCACCTTCCTTGATACAC R: CCACAACCTACCACCTC	(CT) ₁₇ Perfect	194	156–326	53	GQ354556
SPSD005	(AAG) _n	F: TGAACCAACATAACCTACCTG R: CCACACTAAACCGAACTTG	(TTC) ₁₉ Perfect	297	192–452	55	GQ354557
SPSD048	(AAG) _n	F: AAGGGCATAGTTTCAACCAAG R: CTTTTGCTTGTTTCATCAACATG	(AAG) ₁₀ Perfect	197	88	55	GQ354565
SPSD007	(CAG) _n	F: AATCCTTCGCCTATCCTACAC R: TTCACACAGCAGACAGAAGT	(CTG) ₈ (CT) ₉ Compound	168	146–519	55	GQ354558
SPSD008	(CAG) _n	F: GCAAGAACAGACTCAAGAGC R: CTGCTGCTGAAGTAAAAGTTG	(CAG) ₉ (CAA) ₄ (CAG) ₄ (CAA) ₂ Compound	276	144–309	55	GQ354559
SPSD049	(CAG) _n	F: TGGATTGTTTCTGATACTCCA R: CCATAAATTGCTGCATTATTCC	(TTG) ₅ (CTG) ₅ Compound	299	303–667	53	GQ354560
SPSD050	(CAG) _n	F: GAAGCAGAAAACACAGTATTGC R: TTGCTGGAATTTAACCTATCTG	(CTG) ₅ Perfect	225	225–505	53	GQ354561
SPSD053	(CAG) _n	F: ACGCCTTCTTCACTCCAAC R: GCCACCAGTTTTCATCACC	(CTG) ₈ Perfect	208	166–316	53	GQ354562
SPSD056	(CAG) _n	F: GTTCTCCAAAGTCTCCTCCT R: ATCTTTACCTTCCTTCTGGG	(TCC) ₆ TTC(TCC) ₂ Interrupted	197	79–230	53	GQ354563

studies the di- and trinucleotide repeats occurred as perfect, compound, and interrupted motifs. Based on the results found in this research, the genomic libraries enriched with the dinucleotide repeats have been extensively sequenced; the resulting sequence information will be screened to isolate additional SSR regions, primers will be designed, and the resulting markers will be used to develop a molecular map of prairie cordgrass. The molecular map will then be used to find linkage between SSR markers and traits of interest allowing future MAS to be performed.

Acknowledgments

The authors want to thank S. G. Dwire for his technical assistance. The research described in this paper was funded in part by grants from the North Central Sun Grant Initiative and from the USDA Feedstock Genomics Program to J. L. Gonzalez-Hernandez, A. Boe, and V. Owens and additional support from the South Dakota Agricultural Experimental Station. The authors acknowledge the use of the SDSU-FGCF supported in part by NSF/EPSCoR Grant no. 0091948 and by the State of South Dakota.

References

- [1] USDA-NRCS, *Plants Database*, USDA-NRCS, Urbandale, Iowa, USA, 2008.
- [2] A. Boe and D. K. Lee, "Genetic variation for biomass production in prairie cordgrass and switchgrass," *Crop Science*, vol. 47, no. 3, pp. 929–934, 2007.
- [3] J. L. Gonzalez-Hernandez, G. Sarath, J. M. Stein, V. Owens, K. Gedye, and A. Boe, "A multiple species approach to biomass production from native herbaceous perennial feedstocks," *In Vitro Cellular and Developmental Biology-Plant*, vol. 45, no. 3, pp. 267–281, 2009.
- [4] X. Fang, P. K. Subudhi, B. C. Venuto, and S. A. Harrison, "Mode of pollination, pollen germination, and seed set in smooth cordgrass (*Spartina alterniflora*, Poaceae)," *International Journal of Plant Sciences*, vol. 165, no. 3, pp. 395–401, 2004.
- [5] X. Fang, P. K. Subudhi, B. C. Venuto, S. A. Harrison, and A. B. Ryan, "Influence of flowering phenology on seed production in smooth cordgrass (*Spartina alterniflora* Loisel.)," *Aquatic Botany*, vol. 80, no. 2, pp. 139–151, 2004.
- [6] V. G. Kern, N. J. Guarise, and A. C. Vegetti, "Inflorescence structure in species of *Spartina* Schreb. (Poaceae: Chloridoideae: Cynodonteae)," *Plant Systematics and Evolution*, vol. 273, no. 1–2, pp. 51–61, 2008.
- [7] W. D. Clayton, M. S. Vorontsova, K. T. Harman, and H.

- Williamson, *GrassBase-The Online World Grass Flora*, 2006.
- [8] K. Gedye, J. Gonzalez-Hernandez, Y. Ban et al., "Investigation of the transcriptome of Prairie Cord grass, a new cellulosic biomass crop," *The Plant Genome Journal*, vol. 3, no. 2, pp. 69–80, 2010.
- [9] K. M. Moncada, N. J. Ehlke, G. J. Muehlbauer, C. C. Sheaffer, D. L. Wyse, and L. R. DeHaan, "Genetic variation in three native plant species across the State of Minnesota," *Crop Science*, vol. 47, no. 6, pp. 2379–2389, 2007.
- [10] L. Cardle, L. Ramsay, D. Milbourne, M. Macaulay, D. Marshall, and R. Waugh, "Computational and experimental characterization of physically clustered simple sequence repeats in plants," *Genetics*, vol. 156, no. 2, pp. 847–854, 2000.
- [11] R. K. Varshney, A. Graner, and M. E. Sorrells, "Genic microsatellite markers in plants: features and applications," *Trends in Biotechnology*, vol. 23, no. 1, pp. 48–55, 2005.
- [12] A. E. Van't Hof, B. J. Zwaan, I. J. Saccheri, D. Daly, A. N. M. Bot, and P. M. Brakefield, "Characterization of 28 microsatellite loci for the butterfly *Bicyclus anynana*," *Molecular Ecology Notes*, vol. 5, no. 1, pp. 169–172, 2005.
- [13] V. Saladin, D. Bonfils, T. Binz, and H. Richner, "Isolation and characterization of 16 microsatellite loci in the Great Tit *Parus major*," *Molecular Ecology Notes*, vol. 3, no. 4, pp. 520–522, 2003.
- [14] T. L. King, M. S. Eackles, and B. H. Letcher, "Microsatellite DNA markers for the study of Atlantic salmon (*Salmo salar*) kinship, population structure, and mixed-fishery analyses," *Molecular Ecology Notes*, vol. 5, no. 1, pp. 130–132, 2005.
- [15] M. M. Peacock, V. S. Kirchoff, and S. J. Merideth, "Identification and characterization of nine polymorphic microsatellite loci in the North American pika, *Ochotona princeps*," *Molecular Ecology Notes*, vol. 2, no. 3, pp. 360–362, 2002.
- [16] M. J. Blum, C. M. Sloop, D. R. Ayres, and D. R. Strong, "Characterization of microsatellite loci in *Spartina* species (Poaceae)," *Molecular Ecology Notes*, vol. 4, no. 1, pp. 39–42, 2004.
- [17] C. M. Sloop, H. G. McGray, M. J. Blum, and D. R. Strong, "Characterization of 24 additional microsatellite loci in *Spartina* species (Poaceae)," *Conservation Genetics*, vol. 6, no. 6, pp. 1049–1052, 2005.
- [18] S. Landjeva, V. Korzun, and A. Börner, "Molecular markers: actual and potential contributions to wheat genome characterization and breeding," *Euphytica*, vol. 156, no. 3, pp. 271–296, 2007.
- [19] F. Wilde, V. Korzun, E. Ebmeyer, H. H. Geiger, and T. Miedaner, "Comparison of phenotypic and marker-based selection for *Fusarium* head blight resistance and DON content in spring wheat," *Molecular Breeding*, vol. 19, no. 4, pp. 357–370, 2007.
- [20] A. Karakousis and P. Langridge, "High-throughput plant DNA extraction method for marker analysis," *Plant Molecular Biology Reporter*, vol. 21, no. 1, 2003.
- [21] S. Rozen and H. Skaletsky, "Primer3 on the WWW for general users and for biologist programmers," *Methods in Molecular Biology*, vol. 132, pp. 365–386, 2000.
- [22] E. S. Jones, M. P. Dupal, R. Kölliker, M. C. Drayton, and J. W. Forster, "Development and characterization of simple sequence repeat (SSR) markers for perennial ryegrass (*Lolium perenne* L.)," *Theoretical and Applied Genetics*, vol. 102, no. 2–3, pp. 405–415, 2001.
- [23] K. Gedye, J. Gonzalez-Hernandez, R. Schuelke, V. Owens, and A. Boe, "Development of SSR markers from genomic DNA and ESTs in prairie cordgrass, a cellulosic biomass crop," in *Proceedings of the Plant & Animal Genomes XIX Conference*, Town & Country Convention Center, San Diego, Calif, USA, January 2011.
- [24] M. Nei and W. H. Li, "Mathematical model for studying genetic variation in terms of restriction endonucleases," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 76, no. 10, pp. 5269–5273, 1979.
- [25] P. H. A. Sneath and R. R. Sokal, *Numerical Taxonomy*, W.H. Freeman and Company, San Francisco, Calif, USA, 1973.
- [26] W. L. Kovach, *MVSP-A Multivariate Statistical Package For Windows*, Kovach Computing Services: Pentraeth, Walesm, UK, 1998.
- [27] Z. Zhang, S. Schwartz, L. Wagner, and W. Miller, "A greedy algorithm for aligning DNA sequences," *Journal of Computational Biology*, vol. 7, no. 1-2, pp. 203–214, 2000.
- [28] Y.-C. Li, A. B. Korol, T. Fahima, A. Beiles, and E. Nevo, "Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review," *Molecular Ecology*, vol. 11, no. 12, pp. 2453–2465, 2002.