

Genetic Diversity of Twelve Switchgrass Populations Using Molecular and Morphological Markers

Laura M. Cortese · Joshua Honig · Chris Miller · Stacy A. Bonos

© Springer Science+Business Media, LLC. 2010

Abstract Switchgrass (*Panicum virgatum* L.) is a warm season, C₄ perennial grass native to most of North America with numerous applications, including use as a bioenergy feedstock species. To date, no studies on genetic diversity in switchgrass have been conducted that use both molecular and morphological markers. The objectives of this study were to assess genetic diversity and determine differences among and between 12 switchgrass populations grown in New Jersey by examining both morphological and molecular characteristics, and to determine whether morphological, molecular, and/or combined data sets can detect ecotype and/or geographical differences at the population level. Twelve plants from each population were characterized with 16 switchgrass expressed sequence tag-simple sequence repeat markers (EST-SSRs) and seven morphological characters. Data was analyzed using GenAIEx and Unweighted Pair-Group Method of Averages (UPGMA) cluster analysis. Most (64%) of the molecular variation in switchgrass populations exists among individuals within populations, with lesser amounts between populations (36%). Upland and lowland populations were distinguished in all three data sets. Some eastern US and midwestern US populations were distinct in all three data sets. Similarities were observed between all three data sets indicating molecular markers may be useful for identifying morpho-

logical differences or other adaptive traits. The combined data set was the most useful in differentiating populations based on geography and found separation between mid-western and eastern upland populations. The results indicate that the combination of morphological and molecular markers may be useful in future applications such as genetic diversity studies, plant variety protection, cultivar identification, and/or identifying geographic origin.

Keywords Ecotype · Genetic diversity · Lowland · Switchgrass · Upland

Abbreviations

AFLP	Amplified fragment length polymorphism
AMOVA	Analysis of molecular variance
EST-SSR	Expressed sequence tag-simple sequence repeats
PCA	Principal component analysis
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
UPGMA	Unweighted pair-group method of averages

Introduction

Switchgrass (*Panicum virgatum* L.) is a warm season, C₄ perennial grass native to most of North America. It has numerous applications including hay, forage, habitat restoration, and erosion control as well as a component in seeded native grass mixtures and buffer strips [32]. Switchgrass has also been selected by the U.S. Department of Energy (US-DOE) as a model bioenergy feedstock species. Switchgrass is an outcrossing polyploid that has been classified into upland and lowland ecotypes based on morphology and habitat preference [7]. Upland ecotypes are commonly

L. M. Cortese (✉) · J. Honig · S. A. Bonos
Department of Plant Biology and Pathology,
School of Environmental and Biological Sciences,
Rutgers University, Foran Hall, 59 Dudley Road,
New Brunswick, NJ 08901, USA
e-mail: lauracor@eden.rutgers.edu

C. Miller
USDA-NRCS, Cape May Plant Materials Center,
1536 Route 9 North,
Cape May Court House, NJ 08210, USA

octaploids ($2n = 8x = 72$) and occasionally hexaploids ($2n = 6x = 54$) or tetraploids ($2n = 4x = 36$) and are fine-stemmed with various amounts of pubescence on the leaf blades, semi-decumbent, and broad based, with heights of 92 to 152 cm and adapted to drier habitats. The lowland ecotypes are typically tetraploid ($2n = 4x = 36$) and are coarse stemmed, erect, glabrous, more robust, and found in bunches in wetter sites with heights of 61 to 305 cm [20, 26].

The US-DOE launched a research effort beginning in the early 1990s on breeding, culture, and physiology of switchgrass for use as a biofuel because it demonstrates high productivity across a wide geographic range, suitability for marginal land, low water and nutrient requirements, as well as positive environmental benefits [42]. Breeding programs are dependent on genetic variation for the development of improved cultivars. Therefore, the knowledge of genetic diversity is pertinent to improving overall plant characteristics which will allow for a systematic sampling of germplasm for breeding and conservation purposes [25, 41]. Significant genetic diversity has been observed in switchgrass [3, 4, 20, 33, 34].

Previous genetic diversity studies in switchgrass utilized either molecular [4, 16, 20, 33, 34] or morphological markers [3, 5]. Random amplified polymorphic DNA (RAPD) markers were used to analyze genetic relatedness in switchgrass and separate populations into upland and lowland ecotypes [16]. Additionally, Hultquist et al. [20] identified a chloroplast DNA (cpDNA) restriction fragment length polymorphism (RFLP) associated with the upland–lowland ecotype classification, and was able to designate two cytotypes, U (upland) and L (lowland), in switchgrass. Both studies found differences between upland and lowland populations, but no variation associated with ploidy level or individual ecotypes. RFLPs as well as sequencing of the chloroplast *trnL* intron were also shown to be useful in distinguishing between upland and lowland populations in switchgrass [33]. RAPDs were also used to identify structural patterns and spatial variation of switchgrass populations from the northern and central US. However, switchgrass cultivars could not be distinguished from prairie remnant populations, and little to no marker variation was associated with geographic zones (Casler et al. [4, 5]). Expressed sequence tag-simple sequence repeats (EST-SSR) and flow cytometry were also used to determine genetic variability within and among 31 switchgrass populations to determine relationships between ploidy level and genetic variation [34]. Upland and lowland ecotypes were differentiated and there was some evidence of clustering of populations based on geographic location but no association with ploidy level was identified [34]. Morphological variables have also been utilized to characterize phenotypic variability among switchgrass ecotypes. In a study comparing prairie remnant collections in the northern US with

switchgrass cultivars, morphological markers were useful in distinguishing between upland and lowland ecotypes and identified a high degree of phenotypic variability between populations collected from sites within close proximity [3]. No studies to date on genetic diversity in switchgrass have been conducted that use both molecular and morphological markers. The utility of using both molecular and morphological markers has been demonstrated in other species, including cotton (*Gossypium hirsutum* L.) [46], common bean (*Phaseolus vulgaris* L.) [8, 15], peanut (*Arachis hypogea* L.) [10], and *Cucurbita pepo* L. [11, 12]. For example, in *C. pepo*, many landraces cannot be assigned to a given known morphotype; therefore, characterization based on the use of both molecular and morphological markers is essential for elucidating the genetic relationships of ecotypes within this species [11]. In general, the use of both molecular and morphological markers is recommended because each data set provides complementary information with greater power of resolution in genetic diversity analyses [15, 29]. The use of both morphological and molecular markers classify genotypes better than employing only continuous phenotypic variables or only discrete phenotypic variables when assessing genetic diversity [12] and phylogenetic relationships [29]. Both molecular and morphological markers are also valuable for the identification of distinct populations or genotypes for conservation, optimum sites for germplasm collection, and ongoing changes in the pattern of diversity over time. Additionally, morphological and molecular markers are useful for the evaluation and utilization of genetic resources, the study of diversity of pre-breeding and breeding germplasm, and for the protection of the breeder's intellectual property rights [12, 37].

Furthermore, although a significant amount of genetic diversity exists within switchgrass, little research has been conducted on the level of genetic diversity and local adaptation among different ecotypes of switchgrass currently recommended for habitat restoration and biofuel production in the northeast region of the US. The objectives of this study were to assess genetic diversity and determine molecular and morphological differences within and between 12 different switchgrass populations grown in New Jersey by examining both morphological and molecular characteristics, and to determine whether morphological, molecular, and/or combined data sets can detect ecotype and/or geographical differences between these populations.

Materials and Methods

Plant Material

Switchgrass seed from 12 populations was obtained from various sources. Brooklyn, 'Carthage' [30], 'High Tide'

(Miller et al. [31]; <http://plant-materials.nrcs.usda.gov/NJPMC/releases.html>), ‘Shelter’, and Timber germplasm sources were obtained from the Natural Resources Conservation Service—United States Department of Agriculture Plant Materials Center in Cape May NJ and represented eastern ecotypes. All of the additional germplasm sources [Argentina, ‘Caddo’ [17], ‘Kanlow’, ‘Pathfinder’ [38], ‘Shawnee’ [49], ‘Sunburst’ [1], and Turkey] were obtained from the Plant Introduction (PI) collection curated by the Germplasm Resources Information Network (GRIN) and included standard cultivars developed in the Midwest and other germplasm sources from other countries (Table 1). Kanlow represented a lowland ecotype, while Brooklyn, Caddo, Carthage, Pathfinder, Shawnee, Shelter, and Sunburst, represented upland ecotypes [18, 20, 21, 28]. Carthage, High Tide, Argentina, and Turkey have not yet been classified as upland or lowland ecotypes but have morphologies consistent with the upland designation, while Timber has morphology consistent with the lowland designation (Table 1).

Seed of each population was germinated in Pro-Mix HP (K.C. Shafer, York, PA) in 30.5×38.1 cm flats. Individual plants were transplanted to 48-celled flats and held under greenhouse conditions for approximately 8 weeks. Plants were transplanted to a spaced-plant nursery in the spring of 2005 at the Rutgers University Plant Biology Research and Extension Farm at Adelphia, NJ. Individual plants were spaced 0.9 M apart with 12 plants per row. Four rows (or 48 genotypes) of each population were planted together and were spaced 0.9 M apart. Populations were not replicated in space; therefore, valid hypothesis tests regarding cultivar differences for phenotypic traits cannot be made. Morphological measurements were taken on 12 random individuals from each of the 12 different switchgrass populations in 2005 and 2006. Only a total of 139 individual plants were

included in the morphological analysis due to death of some individual plants between 2005 and 2006. Measurements included plant height, panicle length, flag leaf height, length and width, heading date (when panicles first became visible), and anthesis date (50% flowering). Plant height was measured from the soil surface to the average height of the majority of the panicles. Panicle length was measured from the bottom node to the tip of the panicle. Flag leaf height was measured as the distance between the soil surface and the collar of the flag leaf. Flag leaf width was measured at the widest point of the flag leaf. Flag leaf length was measured from the collar to the tip of the leaf blade. Plant height, heading date, and anthesis date were measured on an individual plant basis to obtain one measurement per plant. The remaining measurements were taken from three panicles from each of the 12 plants per population and averaged to obtain one measurement per plant. All measurements were averaged over both years, and averages for each plant were used in subsequent analyses.

DNA Extraction Leaf tissue was also collected from 12 individuals from each population for molecular marker analysis. Leaf tissue was ground in liquid nitrogen, and DNA was isolated using the Sigma® GenElute™ Plant Genomic DNA Miniprep kit (Sigma-Aldrich Co., St. Louis, MO) with the following modifications. Microcentrifuge tubes containing ground plant samples were kept on ice prior to incubation at 65°C. Five microliter RNaseA (Qiagen Inc., Valencia, CA) was added to microcentrifuge tubes prior to the addition of Lysis Solution [Part A and Part B]. Binding Solution was added to flow-through liquid and was mixed by pipetting. In addition, pre-warmed Elution Solution was incubated for 5 min at room temperature (15–25°C) before centrifuging at maximum speed for 1 min.

Table 1 Twelve switchgrass populations evaluated in NJ for morphological and molecular markers and their ecotype designation and origin

Population	Ecotype	GRIN Accession ID	Origin
Argentina	Upland ^a	PI 337553	Rafaela Experiment Station, Santa Fe, Argentina
Brooklyn	Upland		Brooklyn, NY
Caddo	Upland	PI 476297	Stillwater, OK
Carthage	Upland ^a	PI 421138	Carthage, NC
High Tide	Upland ^a		Chesapeake Bay area, Perryville, MD
Kanlow	Lowland	PI 421521	Wetumka, OK
Pathfinder	Upland	PI 642192	Domestic collections from NE and KS
Shawnee	Upland	PI 591824	Shawnee National Forest, IL; from 1 cycle of selection of Cave-in-Rock, selected in Nebraska
Shelter	Upland	PI 430240	Saint Mary’s, WV
Sunburst	Upland	PI 598136	Near Yankton, SD
Timber	Lowland ^a		NC
Turkey	Upland ^a	PI 204907	Ankara, Turkey

^a These populations have not been characterized but exhibit characteristics very similar to the ecotype designation written

PCR Reaction and Genotyping Publicly available microsatellite (SSR) markers derived from expressed sequence tag (EST) sequences from switchgrass were utilized for the molecular marker analysis [47]. Thirty-two SSR primer pairs (Integrated DNA Technologies, Coralville, IA) were tested for polymorphism on the 12 individuals from each population totaling 144 individual plant samples. Each polymerase chain reaction contained 10× Ramp-Taq Buffer (160 mM (NH₄)₂SO₄, 670 mM Tris–HCl pH 8.3, 0.1% Tween-20), 2 mM MgCl₂ (Denville Scientific Inc., Metuchen, NJ), 0.25 mM each dNTP, 2 μL of 500 pM/μL reverse primer, 1.5 μL of 500 pM/μL forward primer (the 5' end of the forward primer was fluorescently labeled with one of four possible dyes—6-FAM™, NED™, PET™, or VIC®; Applied Biosystems, Foster City, CA), 0.5 units of Ramp-Taq™ DNA Polymerase, and 25 ng of template DNA in a total volume of 12.5 μL. PCR was performed using Applied Biosystems GeneAmp® PCR System 9700 thermocyclers with the following profile: initial denaturation at 94°C for 7 min; 30 cycles of 94°C for 30 s, 56°C for 45 s, and 72°C for 45 s; 8 cycles of 94°C for 30 s, 53°C for 45 s, and 72°C for 45 s; final extension at 72°C for 10 min. Four polymerase chain reaction products labeled with different fluorescent dyes were pooled (1 μL of each reaction) and the volume was brought to 10 μL with sterile water. One microliter of the pooled mixture was combined with 9 μL Hi-Di Formamide (Applied Biosystems, Foster City, CA) and 1 μL of the Genescan-500 ROX Size Standard (Applied Biosystems Foster City, CA). The samples were heated at 95°C for 5 min and quenched on ice. Raw data was generated on the ABI 3130 genetic analyzer and genotypes were scored using GeneMapper Version 3.7 software (Applied Biosystems, Foster City, CA).

Morphological Data Analysis Population means of plant measurements over both years including plant height, panicle height, flag leaf height, length, and width, heading date, and anthesis date from 139 individuals as well as phenotypic standard deviations of each mean value were determined using the Proc Means procedure in SAS Version 9.1 (SAS Institute [43]). Given the experimental design utilized in this experiment, no statistical hypothesis tests regarding population differences for phenotypic traits, actual or implied, were used. Population means were used only as a descriptive tool without any statistical inferences. The 12 population means were then subjected to principal component analysis (PCA) using the Proc Princomp procedure in SAS (SAS Institute [43]). Principal components were used as input variables for a cluster analysis using the unweighted pair-group method of averages (UPGMA) to generate a dendrogram using the Proc Cluster procedure in SAS Version 9.1 (SAS Institute [43]).

Molecular Data Analysis Statistical procedures for genetic analysis were developed for diploid organisms and are not sufficient for analysis of organisms with higher ploidy levels that contain more than two alleles at a given locus. Switchgrass is primarily tetraploid and octaploid and can have four or eight possible alleles at a given locus. Therefore, the polymorphic SSR bands for each individual were scored individually for presence or absence [16, 24, 33]. This resulted in a data set of 1's and 0's for 103 alleles (141 individuals from 12 populations). Only 141 individuals were included in the molecular analysis due to poor amplification of three individual genotypes. Nei's genetic distance matrix was calculated from presence/absence data from 141 individuals according to Nei [35] using GenAlEx Version 6.2 [40]. The genetic distance matrix was used as input for a cluster analysis using the unweighted pair-group method of averages (UPGMA) to generate a dendrogram using the Proc Cluster procedure using Ward's minimum-variance criteria in SAS Version 9.1 (SAS Institute [43]). Ward's method utilizes an analysis of variance approach for evaluation of distances between clusters and attempts to minimize the sum of squares (SS) of any two clusters that could be formed [50]. Genetic distances were subjected to an analysis of molecular variance (AMOVA) using GenAlEx 6.2 based on 999 permutations [40]. AMOVA allows for a partitioning of molecular variance within and among populations and tests the significance of partitioned variance components using permutational testing procedures [9]. ΦPT values, analogous to Fst when data are haploid or binary, are calculated in an AMOVA and represent the proportion of the total variance that is partitioned between populations [9, 40].

Combined Data Analysis Plants that had both morphological and molecular marker data were included in a third data set. This combined data set contained 110 characters, the seven morphological measurements from the morphological marker data set and the 103 SSR alleles from the molecular marker data set, from 119 total individuals from 12 populations, with eight to 12 individuals per population. Only 119 individuals had both morphological and molecular data due to plant death and poor amplification and were thus included in the combined analysis. Means of morphological measurements for each population, used only as a descriptive tool without any statistical inferences, were computed using the Proc Means procedure in SAS, and were subjected to a PCA using the Proc Princomp analysis in SAS to generate seven principal components. Nei's genetic distance matrix (Table 2) was calculated from the molecular presence/absence data according to Nei [35] using GenAlEx Version 6.2 [40]. The genetic distance matrix was then subjected to PCA to generate seven principal components. The seven principal components from the morphological data and the seven from the molecular

Table 2 Pairwise population matrix of Nei's genetic distance for 12 switchgrass populations. Individuals from each population were evaluated for presence or absence of 103 alleles amplified by 16 EST-SSR primer pairs

	CAD	SHA	AR	TU	SU	KA	SHE	HT	PA	BR	TI
Shawnee (SHA)	0.028										
Argentina (AR)	0.035	0.042									
Turkey (TU)	0.072	0.064	0.083								
Sunburst (SU)	0.062	0.050	0.067	0.078							
Kanlow (KA)	0.197	0.225	0.233	0.218	0.236						
Shelter (SHE)	0.101	0.105	0.118	0.071	0.112	0.213					
High Tide (HT)	0.123	0.129	0.150	0.147	0.163	0.170	0.135				
Pathfinder (PA)	0.047	0.052	0.069	0.103	0.083	0.209	0.117	0.129			
Brooklyn (BR)	0.128	0.158	0.168	0.162	0.169	0.151	0.142	0.065	0.134		
Timber (TI)	0.169	0.193	0.192	0.189	0.204	0.073	0.188	0.113	0.180	0.110	
Carthage (CAR)	0.073	0.091	0.104	0.097	0.117	0.193	0.104	0.110	0.062	0.106	0.173

The distance values were generated based on the similarity index between populations

marker data were combined into one data set of 14 principal components, comprising the third combined data set. This data set was then used as input for a cluster analysis using the unweighted pair-group method of averages (UPGMA) to generate a dendrogram using the Proc Cluster procedure in SAS Version 9.1 (SAS Institute [43]).

Results and Discussion

Morphological Characterization

The UPGMA analysis of morphological measurements (Fig. 1) corresponded well to the ecotype origins (Table 1) for the populations evaluated with a few exceptions. The UPGMA analysis resulted in three distinct clades with the lowland types (Kanlow and Timber) forming a separate clade from the upland types (Fig. 1). Within the upland group, Carthage formed a separate clade distinct from the other upland ecotypes. This is not surprising since Carthage is the only upland ecotype collected from North Carolina and is visibly distinct from other upland ecotypes evaluated (Table 3). The remaining upland populations consisting of Argentina, Caddo, Brooklyn, Pathfinder, Turkey, Shelter, Sunburst, High Tide, and Shawnee formed the third clade. These populations also grouped together in a Principal Component Analysis (data not shown). Of interest was the inclusion of the ecotypes from Argentina and Turkey in the upland clade. Similar results were observed for the Argentina and Turkey germplasm sources by Narasimhamoorthy et al. [34]. It is possible that these were introductions from North America [13]. These results indicate that the morphological measurements utilized in the cluster analysis were effective in distinguishing between upland and lowland switchgrass

ecotypes which is similar to previous studies of morphological characterization of switchgrass ecotypes [3].

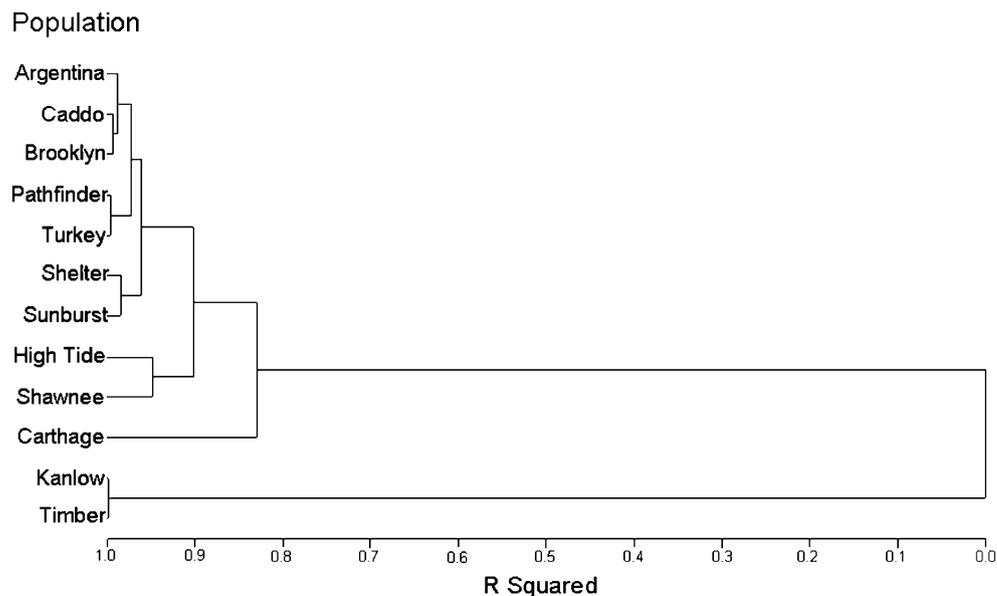
Molecular Characterization

The 32 EST-derived SSR primer pairs obtained from Tobias et al. [47] were tested on the 12 individual plants from each population. Sixteen of the 32 primer pairs amplified a polymorphic SSR locus and were highly informative among the switchgrass populations evaluated. These 16 were chosen for genetic analysis. The 16 polymorphic loci identified between two and 12 alleles with an average of 6.5 alleles per locus in 12 populations of switchgrass. The 16 SSR loci amplified a total of 103 alleles that were used for genetic analysis.

Genetic Diversity among Switchgrass Populations

Similarity coefficients of Nei's genetic distance between switchgrass populations ranged from 0.028 to 0.236 (Table 2). The highest degree of similarity was observed between Caddo and Shawnee, two upland populations, while the lowest degree of similarity (most diversity) was observed between Sunburst and Kanlow, an upland and a lowland population, respectively. Some ecotypes evaluated in this study were shown to be very closely related and shared a high degree of genetic similarity, while other ecotypes were quite diverse. Pairwise Jaccard genetic distances based on a dissimilarity index among three switchgrass populations evaluated with RFLP markers ranged from 0.70 to 0.82 between genotypes [33]. These findings were similar to those reported in this paper, which when converted to a dissimilarity matrix (1-similarity coefficient; data not shown) range from 0.76 to 0.97. It is possible that the values reported here indicated a greater

Fig. 1 Cluster analysis using UPGMA of morphological measurements of seven characters averaged over two years in 12 switchgrass populations. Distances between clusters are expressed in R^2 values, where high R^2 values indicate more similarity



degree of similarity between populations due to the fact that more ecotypes were analyzed and a different marker system was used. Similarity coefficients among 14 switchgrass populations evaluated with RAPD markers ranged from 0.53 to 0.78 [16], and indicated a greater degree of diversity between the populations and differed considerably from what was shown here or by Missaoui et al. [33].

The difference in similarity coefficients between this study and Missaoui et al. [33] and that of Gunter et al. [16] could be due to the fact that different populations were evaluated in each study and different marker systems were

utilized. In a comparison of RAPD, RFLP, AFLP, and SSR markers for utility in diversity studies in maize, it was shown that AFLPs and RFLPs were the most highly correlated marker systems for genetic distances, closely followed by SSRs and RFLPs. SSRs and RAPDs were the least highly correlated of the markers [14]. The study indicated that AFLPs, RFLPs, and SSRs all showed similar degrees of genetic diversity, while RAPDs produced results inconsistent with those of the aforementioned markers. Additionally, EST-SSRs and RFLPs are both highly conserved and therefore may result in the generation of

Table 3 Means of morphological measurements of 12 switchgrass populations grown in NJ in 2005 and 2006

Population	Ecotype	Plant height (cm)	Panicle length (cm)	Flag leaf height (cm)	Flag leaf length (cm)	Flag leaf width (cm)	Heading date ^a (Julian)	Anthesis date ^b (Julian)
Argentina	Upland ^c	129±10 ^d	44±4	78±11	39±4	1.0±0.1	194±2	235±4
Brooklyn	Upland ^c	126±8	44±8	71±8	43±6	1.0±0.2	186±6	232±5
Caddo	Upland	129±15	46±7	79±13	41±4	0.9±0.2	185±5	230±3
Carthage	Upland ^c	156±14	67±8	92±15	50±3	1.1±0.1	195±8	234±3
High Tide	Upland ^c	149±8	48±10	95±6	47±8	1.1±0.2	197±6	238±6
Kanlow	Lowland	198±11	58±6	140±9	52±4	1.3±0.2	208±6	243±5
Pathfinder	Upland	137±11	49±8	86±12	44±4	1.0±0.1	190±4	234±5
Shawnee	Upland	133±8	45±10	98±21	40±4	0.9±0.1	190±5	235±6
Shelter	Upland	129±11	37±4	79±9	40±5	1.0±0.1	180±6	233±2
Sunburst	Upland	137±10	37±7	84±13	36±7	0.9±0.1	184±7	237±6
Timber	Lowland ^c	198±11	56±8	138±12	54±4	1.2±0.1	203±4	243±6
Turkey	Upland ^c	137±10	44±6	82±6	41±7	0.9±0.1	186±2	232±8

Measurements were averaged over both years

^a Heading date was recorded when panicles were first visible

^b Anthesis date was recorded when about 50% of flowers were open

^c These populations have not been characterized but exhibit characteristics very similar to the ecotype designation written

^d Phenotypic standard deviation of each mean value. Statistical comparisons between populations are not valid due to lack of spatial replication

smaller genetic distances than a marker system such as RAPDs which are not well conserved.

UPGMA analysis of EST-SSR marker variation corresponded well to collection sites (Table 1) and to the morphological analysis (Fig. 1) with a few exceptions. The UPGMA analysis of the marker data resulted in four distinct clades (Fig. 2). In this analysis, Kanlow and Timber, the two lowland ecotypes, formed their own group. High Tide and Brooklyn, both collected from eastern US (Table 1) formed a distinct group. The other two clades made up the rest of the upland ecotypes and were more similar to each other than to the two previously mentioned clades (Kanlow and Timber and High Tide and Brooklyn). The smaller of the two clades contained Turkey, Shelter and Carthage while the large clade contained Caddo, Shawnee, Argentina, Pathfinder, and Sunburst. These two clades are similar to the results of the morphological analysis for the upland ecotypes, except that Carthage is included with the other upland ecotypes. It is interesting to note that the UPGMA analysis grouped some of the upland ecotypes by geographic region. High Tide was collected from MD and Brooklyn was collected from NY. Additionally, Shelter was collected from WV while Carthage from NC. This is an indication that ecotypes collected from close geographic regions may share some common alleles.

These results indicate that EST-SSR markers were effective in distinguishing between switchgrass populations. The similarity between morphological and molecular marker analysis indicates the genic SSR markers may be highlighting expressed traits with adaptive significance. EST-SSRs are derived from transcribed genes and often characterize functionally relevant polymorphisms. They have been shown to have roles in

gene expression, regulation of DNA recombination, transcription and translation, as well as putative roles in providing an adaptive advantage [6, 27, 48]. The results shown here reveal that EST-SSR markers were useful in differentiating between closely related germplasm sources and could be used to supplement morphological and agronomic data used for plant variety protection and/or cultivar identification.

AMOVA

Results of the AMOVA (Table 4) indicated that most (64%) of the molecular variation in switchgrass populations exists among individuals within populations, with lesser amounts among populations (36%). Permutation tests (based on 999 permutations) suggest that the overall Φ_{PT} was significant ($\Phi_{PT}=0.37, P=0.001$; Table 4), which indicates the differences among ecotypes are significant. Similar results were observed in switchgrass [3, 4, 16, 33, 34] as well as other outcrossing species including buffalograss [*Buchloe dactyloides* (Nutt.) Engelm.] [19, 39], *Physaria bellii* G. A. Mulligan [22], perennial ryegrass (*Lolium perenne* L.) [23], and creeping bentgrass (*Agrostis stolonifera* L.) [24].

Combined Analysis

UPGMA analysis of the combined principal components revealed a cluster pattern somewhat similar to that of both the morphological and molecular cladograms. The combined analysis resulted in three distinct clades, with the lowland types (Kanlow and Timber) forming a separate clade from the upland types (Fig. 3). This is similar to the results seen in the morphological (Fig. 1) and the molecular

Fig. 2 An UPGMA dendrogram analysis of 12 switchgrass populations using 103 alleles from 16 EST-SSR primer pairs. Distances between clusters are expressed in R^2 values, where high R^2 values indicate more similarity

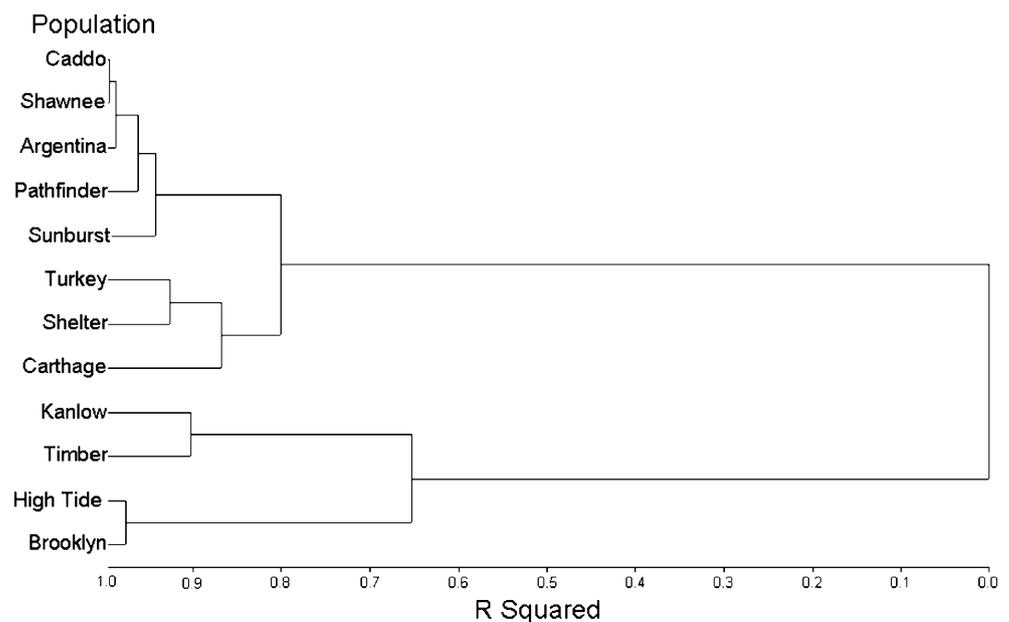


Table 4 Analysis of molecular variance (AMOVA) for 141 individuals from 12 switchgrass populations based on 16 switchgrass EST-SSR markers

Source of variation	Degrees of freedom	Sum of squares	Means square	Est. Var.	% variation	<i>P</i> value
Among populations	11	675.637	61.422	4.554	36%	0.001
Within populations	129	1022.377	7.925	7.925	64%	0.001
Total	140	1698.014		12.479	100%	0.001
$\Phi_{PT}=0.365$						0.001

(Fig. 2) analyses. Within the upland groups, Carthage and High Tide formed a second clade which differs from both the morphological cluster analysis and the molecular cluster analysis (Figs. 1 and 2). Carthage originated in NC while High Tide was collected in MD, the two sites being in relatively close geographic proximity to one another. The remaining upland populations, Argentina, Pathfinder, Sunburst, Turkey, Shawnee, Caddo, Brooklyn, and Shelter, comprised the third clade, similar to the morphological cluster analysis (Fig. 1). The cluster analysis of the combined data set most closely resembles that of the morphological data set with one major exception. The combined data included High Tide in a clade with Carthage, whereas in the morphological analysis High Tide clustered with all the other upland populations and Carthage formed its own clade.

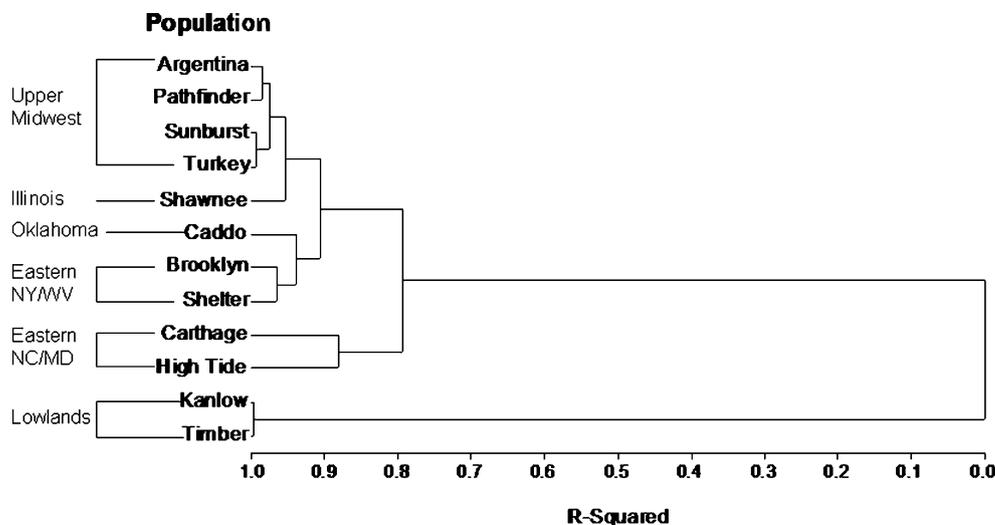
The cluster analysis of the combined data set fit most closely to geographic origin compared to all other analyses if it is assumed that the populations Argentina and Turkey are introductions of ecotypes from the upper Midwest. The upper Midwest ecotypes, i.e., Sunburst, Pathfinder, and Shawnee clustered with the majority of the upland populations. Caddo, originating from OK, also grouped with the majority of the upland populations. In a cluster analysis of SSR markers in 31 switchgrass accessions,

upland ecotypes Caddo, Argentina, Turkey, Sunburst, and Shawnee were all included in the same cluster [34].

The eastern upland ecotypes were split into two groups: Brooklyn and Shelter, and Carthage and High Tide. Brooklyn and Shelter are lower growing than Carthage and High Tide (personal observation) and may be the reason for the similarities observed. The lowlands formed their own clade irrespective of geographic origin, which was similar to the other analyses reported in this experiment and others [16, 33, 34]. Although others have reported genetic differences among Midwest ecotypes [4, 34], this is the first report of clear separation between midwestern and eastern ecotypes of switchgrass.

Although other types of analyses with different assumptions may reveal different genetic relationships, there are assumptions made with all types of genetic analyses. For example, UPGMA assumes a constant rate of nucleotide substitutions over time [36], while maximum likelihood (ML) makes assumptions about the ratio of the transition rate to the transversion rate and GC content [45]. Several other genetic diversity studies have utilized both UPGMA and neighbor-joining (NJ) analyses and have found similar results from both analyses [2, 44, 51, 52]. Although other methods may provide insight and further discriminate genetic similarities and differences between populations, a

Fig. 3 Cluster analysis using UPGMA of 12 switchgrass populations using 14 principal components, seven morphological and seven molecular components. Distances between clusters are expressed in R^2 values, where high R^2 values indicate more similarity



principal component analysis supported the UPGMA results reported here (data not shown).

Conclusions

This is the first study to combine morphological and molecular markers to evaluate genetic diversity in switchgrass. It is also the first report to evaluate the genetic diversity of eastern switchgrass ecotypes compared to midwestern ecotypes. The combined data set corresponded best to ecotype origin and did separate upland and lowland ecotypes as well as some midwestern and eastern upland ecotypes. The distinctness between ecotypes reported here provides further evidence that regional breeding programs may be necessary to optimize genetic diversity in a given area as well as performance and biomass yield potential. The combination of morphological and EST-SSR markers were useful in differentiating between closely related germplasm sources and could have applications in identifying unknown origins of germplasm sources, plant variety protection, and/or cultivar identification.

Acknowledgments The authors would like to thank the anonymous reviewers for their comments and suggestions and Drs. Mike Casler, Peter Smouse, and Lena Struwe for their time and advice on data analysis. We are grateful to Eric Weibel, Rob Shortell, PhD, Matt Koch, Christine Kubik, and Matt Mattia for assistance with data collection. This research was supported by the Rutgers Center for Turfgrass Science.

References

- Boe A, Ross JG (1998) Registration of 'Sunburst' switchgrass. *Crop Sci* 38:540
- Budak H, Shearman RC, Parmakasiz I, Gaussoin RE, Riordan TP, Dweikat I (2004) Molecular characterization of Buffalograss germplasm using sequence-related amplified polymorphism markers. *Theor Appl Genet* 108:328–334
- Casler MD (2005) Ecotypic variation among switchgrass populations from the Northern USA. *Crop Sci* 45:388–398
- Casler MD, Stendal CA, Kapich L, Vogel KP (2007) Genetic diversity, plant adaptation regions, and gene pools for switchgrass. *Crop Sci* 47:2261–2273
- Casler MD, Vogel KP, Taliaferro CM, Ehlke NJ, Berdahl JD, Brummer EC et al (2007) Latitudinal and longitudinal adaptation of switchgrass populations. *Crop Sci* 47:2249–2260
- Chistiakov DA, Hellemans B, Volckaert FAM (2006) Microsatellites and their genomic distribution, evolution, function, and applications: a review with special reference to fish genetics. *Aquaculture* 255:1–29
- Das MK, Fuentes RG, Taliaferro CM (2004) Genetic variability and trait relationships in switchgrass. *Crop Sci* 44:443–448
- Duran LA, Blair MW, Giraldo MC, Macchiavelli R, Prophete E, Nin JC et al (2005) Morphological and molecular characterization of common bean landraces and cultivars from the Caribbean. *Crop Sci* 45:1320–1328
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Ferguson ME, Bramel PJ, Chandra S (2004) Gene diversity among botanical varieties in peanut (*Arachis hypogaea* L.). *Crop Sci* 44:1847–1854
- Ferriol M, Pico B, Nuez F (2003) Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. *Theor Appl Genet* 107:271–282
- Franco J, Crossa J, Ribaut JM, Betran J, Warburton ML, Khairallah M (2001) A method for combining molecular markers and phenotypic attributes for classifying plant genotypes. *Theor Appl Genet* 103:944–952
- Freckmann RW, Lelongford MG (1993) *Panicum*. In: Flora of North America Editorial Committee (eds) Magnoliophyta:Comelinidae (in part):Poaceae, part 2. Flora of North America North of Mexico, vol 25. New York and Oxford, pp 450–488
- Garcia AAF, Benchimol LL, Barbosa AMM, Geraldi IO, Souza CL, de Souza AP (2004) Comparison of RAPD, RFLP, AFLP, and SSR markers for diversity studies in tropical maize inbred lines. *Genet Mol Biol* 27:579–588
- Gomez OJ, Blair MW, Frankow-Lindberg BE, Gullberg U (2004) Molecular and phenotypic diversity of common bean landraces from Nicaragua. *Crop Sci* 44:1412–1418
- Gunter LE, Tuskan GA, Wullschlegel SD (1996) Diversity among populations of switchgrass based on RAPD markers. *Crop Sci* 36:1017–1022
- Hein MA (1958) Registration of Caddo switchgrass. *Agron J* 50:399
- Hopkins AA, Taliaferro CM, Murphy CD, Christian D (1996) Chromosome number and nuclear DNA content of several switchgrass populations. *Crop Sci* 36:1192–1195
- Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalograss (*Buchloe dactyloides* (Nutt.) Englem). *Theor Appl Genet* 86:927–934
- Hultquist SJ, Vogel KP, Lee DJ, Arumuganathan K, Kaeppler S (1996) Chloroplast DNA and nuclear DNA content variations among cultivars of switchgrass, *Panicum virgatum* L. *Crop Sci* 36:1049–1052
- Hultquist SJ, Vogel KP, Lee DJ, Arumuganathan K, Kaeppler S (1997) DNA content and chloroplast DNA polymorphisms among switchgrasses from remnant midwestern prairies. *Crop Sci* 37:595–598
- Kothera L, Richards CM, Carney SE (2007) Genetic diversity and structure in the rare Colorado endemic plant *Physaria bellii* Mulligan (Brassicaceae). *Conserv Genet* 8:1043–1050
- Kubik C, Sawkins M, Meyer WA, Gaut BS (2001) Genetic diversity in seven perennial ryegrass (*Lolium perenne* L.) cultivars based on SSR markers. *Crop Sci* 41:1565–1572
- Kubik C, Honig JA, Meyer WA, Bonos SA (2009) Genetic diversity of creeping bentgrass cultivars using SSR markers. *Int Turfgrass Soc Res J* 11:533–547
- Kumar LS (1999) DNA markers in plant improvement: an overview. *Biotechnol Adv* 17:143–182
- Lewandowski I, Scurlock JMO, Lindvall E, Christou M (2003) The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. *Biomass and Bioenerg* 25:335–361
- Li Y, Korol AB, Fazima T, Nevo E (2004) Microsatellites within genes: structure, function, and evolution. *Mol Biol Evol* 21:991–1007
- Lu K, Kaeppler SW, Vogel K, Arumuganathan K, Lee DJ (1998) Nuclear DNA content and chromosome numbers in switchgrass. *Gt Plains Res* 8:269–280
- Marvaldi AE, Sequeira AS, O'Brien CW, Farrell BD (2002) Molecular and morphological phylogenetics of Weevils (Coleoptera, Curculionioidea): do niche shifts accompany diversification? *Syst Biol* 51:761–785

30. Miller C, Skaradek B, Bonos, SA (2006) Notice of release of 'Carthage' switchgrass. USDA-NRCS Somerset, NJ and the New Jersey Agricultural Experiment Station. August 2006:1–10
31. Miller C, Skaradek B, Bonos, SA (2007) Notice of release of 'High Tide' switchgrass. USDA-NRCS Cape May Courthouse, NJ and the New Jersey Agricultural Experiment Station. August 2007:1–9
32. Missaoui AM, Fasoula VA, Bouton JH (2005) The effect of low plant density on response to selection for biomass production in switchgrass. *Euphytica* 142:1–12
33. Missaoui AM, Paterson AH, Bouton JH (2006) Molecular markers for the classification of switchgrass (*Panicum virgatum* L.) germplasm and to assess genetic diversity in three synthetic switchgrass populations. *Genet Resour and Crop Evol* 53:1291–1302
34. Narasimhamoorthy B, Saha MC, Swaller T, Bouton JH (2008) Genetic diversity in switchgrass collections assessed by EST-SSR markers. *Bioenerg Res* 1:136–146
35. Nei M (1972) Genetic distance between populations. *Am Nat* 106:283–392
36. Nei M, Jin L (1989) Variances of the average numbers of nucleotide substitutions within and between populations. *Mol Biol Evol* 6:290–300
37. Newbury HJ, Ford-Lloyd BV (1997) Estimation of genetic diversity. In: Maxted N et al (eds) *Plant genetic conservation: the in situ approach*. Chapman and Hall, London, pp 192–206
38. Newell LC (1968) Registration of Pathfinder Switchgrass. *Crop Sci* 8:516
39. Peakall R, Smouse PE, Huff DR (1995) Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloe dactyloides*. *Mol Ecol* 4:135–147
40. Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
41. Rajasekar S, Fei S, Christians NE (2006) Analysis of genetic diversity in rough bluegrass determined by RAPD markers. *Crop Sci* 46:162–167
42. Sanderson MA, Reed RL, McLaughlin SB, Wullschleger SD, Conger BV, Parrish DJ et al (1996) Switchgrass as a sustainable bioenergy crop. *Bioresour Technol* 56:83–93
43. SAS Institute Inc (2004) SAS OnlineDoc® 9.1.3. SAS Institute, Cary, NC
44. Sehgal D, Rajpal VR, Raina SN, Sasanuma T, Sasakuma T (2009) Assaying polymorphism at DNA level for genetic diversity diagnostics of the safflower (*Carthamus tinctoris* L.) world germplasm resources. *Genetica* 135:457–470
45. Tatenno Y, Takezaki N, Nei M (1994) Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum parsimony methods when substitution rates vary with site. *Mol Biol Evol* 11:261–277
46. Tatineni V, Cantrell RG, Davis DD (1996) Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. *Crop Sci* 36:186–192
47. Tobias CM, Hayden DM, Twigg P, Sarath G (2006) Genic microsatellite markers derived from EST sequences of switchgrass (*Panicum virgatum* L.). *Mol Ecol Notes* 6:185–187
48. Varshney RK, Graner A, Sorrels ME (2005) Genic microsatellite markers in plants: features and applications. *Trends in Biotechnol* 23:48–55
49. Vogel KP, Hopkins AA, Moore KJ, Johnson KD, Carlson IT (1996) Registration of 'Shawnee' switchgrass. *Crop Sci* 36:1713
50. Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58:236–244
51. Yokoyama K, Biswas SK, Miyaji M, Nishimura K (2000) Identification and phylogenetic relationship of the most common pathogenic *Candida* species inferred from mitochondrial cytochrome b gene sequences. *J Clin Microbiol* 38:4503–4510
52. Zeller KA, Jurgenson JE, El-Assiuty EM, Leslie JF (2000) Isozyme and amplified fragment length polymorphisms from *Cephalosporium maydis* in Egypt. *Phytoparasitica* 28:121–130