Agroforest Syst
DOI 10.1007/s10457-014-9728-z

Genetic diversity analysis of switchgrass (*Panicum virgatum* L.) populations using microsatellites and chloroplast sequences

Madhugiri Nageswara-Rao · Micaha Hanson · Sujata Agarwal · C. Neal Stewart Jr. · Charles Kwit

Received: 20 March 2014 / Accepted: 9 July 2014

Abstract The agricultural landscape of the United States could soon be changed by planting of switchgrass (*Panicum virgatum* L.) cultivars to meet government-mandated targets for lignocellulosic bioenergy production and consumption. This alteration could affect the genetic structure of wild switchgrass populations, which are native to the eastern half of North America through cultivar introgression. In this study, PCR amplification of microsatellite fragments as well as chloroplast gene-specific markers were utilized to quantify the genetic diversity and structure of five native populations and three agronomic fields (hereafter 'populations') planted with switchgrass cultivars. Microsatellite polymorphism across all the switchgrass populations ranged from 91.4 to 100 %. Overall, natural switchgrass populations had significantly higher mean genetic diversity than agronomic switchgrass cultivars (0.262 ± 0.102 and 0.201 ± 0.082 respectively, t test $p < 0.008$). Natural switchgrass populations had significantly higher total genetic diversity within (H_S) and among (H_T) as compared to agronomic switchgrass cultivars. A clear separation of natural and agronomic switchgrass populations was noted using principal component analysis and STRUCTURE analysis. A grouping pattern similar to that obtained in the microsatellite study was observed when chloroplast nucleotide sequence variation was assessed. In the realm of bioenergy sustainability, our results highlight the need to consider the genetic structure of cultivars for bioenergy when they are grown in proximity to native switchgrass populations.

Keywords Agronomic cultivars · Biofuel · Chloroplast nucleotide sequencing · Conservation and restoration · Genetic variability · Natural populations

Electronic supplementary material The online version of this article (doi:10.1007/s10457-014-9728-z) contains supplementary material, which is available to authorized users.

M. Nageswara-Rao (✉) · M. Hanson · S. Agarwal · C. N. Stewart Jr. · C. Kwit
Department of Plant Sciences, University of Tennessee, 252 Ellington Plant Sciences, 2431 Joe Johnson Dr., Knoxville, TN 37996, USA
e-mail: mnrbhav@yahoo.com; mnrao@utk.edu

M. Nageswara-Rao
Department of Biology, New Mexico State University, P. O. Box 30001, MSC 3AF, Foster Hall, Las Cruces, NM 88003, USA

C. N. Stewart Jr.
BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

C. Kwit
Department of Forestry, Wildlife and Fisheries, University of Tennessee, 274 Ellington Plant Sciences, 2431 Joe Johnson Dr., Knoxville, TN 37996, USA

Introduction

Switchgrass (*Panicum virgatum* L., Poaceae), a warm season C_4 perennial grass is a common species

Published online: 22 July 2014

indigenous to tall grass prairies of eastern North America. It occurs from 15° to 55°N and east of 100°W, and is capable of growing in a variety of conditions, including on marginal lands (Moser and Vogel 1995; Vogel 2004; Wang et al. 2011). It is a predominantly self-incompatible, cross-pollinated polyploid species and is morphologically diverse (Talbert et al. 1983). It has a basic chromosome number of n = 9 and ploidy levels ranging from diploid (2n = 2x = 18) to dodecaploid (2n = 12x = 108) (Church 1940; Burton 1942; Nielsen 1944). Switchgrass has two widespread ecotypes: lowland and upland. The lowland ecotypes are taller than upland ecotypes and typically exhibit tetraploidy (2n = 4x = 36). The upland ecotypes are typically hexaploid (2n = 6x = 54) or octaploid (2n = 8x = 72), occupy northerly sites and have a spreading growth habit (Barnett and Carver 1967; Hopkins et al. 1996; Costich et al. 2010).

Switchgrass has received attention as a dedicated lignocellulosic biofuel crop to meet the United States government-mandated benchmarks for advanced biofuel production and consumption (McLaughlin and Kszos 2005; Wright and Turhollow 2010; Nageswara-Rao et al. 2013). The required mass-planting hypothesized to meet those benchmarks (McDonald et al. 2009) could lead to a significant modification of the landscape, especially in Mid-southern portion of the United States where the biomass yields of switchgrass are projected to be highest (Wullschleger et al. 2010). Not much information is available on the genetic structure of natural switchgrass populations in this region, and the introduction of agronomic or "non-native" cultivars may result in the alteration of current genetic structure of native wild conspecific populations through crop-to-wild introgression and/or seed dispersal (Kwit and Stewart 2012; Lewis 2013; Ridley et al. 2013). Hence, it is necessary to study the extant of genetic diversity of switchgrass for the conservation of its genetic resources before commercial plantings become widespread.

The objectives of this study were to utilize microsatellites and chloroplast DNA (cp-DNA) markers for the characterization of genetic diversity and structure of natural and agronomic switchgrass populations in a geographic area poised for landscape change incorporating extensive plantings for bioenergy. Microsatellites, also known as simple sequence repeat markers are highly polymorphic co-dominant DNA markers with a high degree of consistency and reproducibility (Morgante and Olivieri 1993). They are useful in population-level genetic diversity studies. They have been utilized for differentiating the genetic structure of populations, estimating the levels of gene flow, identifying cultivars/hybrids, and genetic imprinting (Clauss et al. 2002; Nageswara-Rao et al. 2008; Wang et al. 2011; Zalapa et al. 2011). Microsatellites have been utilized in switchgrass for analyzing the DNA polymorphisms in accessions largely obtained from the USDA Germplasm Resources Information Network (Narasimhamoorthy et al. 2008; Cortese et al. 2010; Zalapa et al. 2011), as well as assessing standing genetic diversity among natural switchgrass populations (Zhang et al. 2011). Markers from cp-DNA have also been used to determine the effect of fragmentation of populations, for assessing the distribution of genetic structure through the maternal lineage, and, in the case of switchgrass, for the classification of its ecotypes (Hultquist et al. 1996; Missaoui et al. 2006).

Materials and methods

Plant materials and DNA extraction

Total DNA was isolated from fresh leaves of individuals from five natural switchgrass populations and three fields planted with agronomic switchgrass cultivars (hereafter, all are simply referred to as "populations") in central and eastern Tennessee, USA (Supplementary Fig. 1). Between 12 and 14 individuals were sampled from each population except Morrison Meadow "East" and Rowe Gap, wherein nearly all existing individuals (n = 6) were sampled. To ensure sampling from different genets, the individuals sampled were >5 m apart. Seeds of *P. dichotomiflorum* and 'Cave-in-Rock' were obtained from USDA National Plant Germplasm System; DNA from leaf tissue of resultant seedlings was used for outgrouping purposes. Using DNeasy Plant Mini kit (Qiagen Inc., Valencia, CA, USA), total DNA was extracted from 200 mg of finely ground leaf tissue in a 2.0 ml micro-centrifuge tube. The extracted DNA was treated with 5 µl of RNaseA (10 mg/ml; Fisher Scientific, Pittsburgh, PA, USA). The DNA concentration was measured by a Nanodrop spectrophotometer ND2000 (Thermo Scientific, Waltham, MA, USA) as well as by electrophoresis using 1 % agarose

Table 1 Microsatellite primer pair sequence and genetic diversity parameters for switchgrass populations

Primer name	Sequence (5′ → 3′)	Repeat motif	Polymorphism (%)	Mean genetic diversity[a] (±SD[b])
SG4	F-GGAATTCAGAAGGCCCATTA	(TG)31	100	0.251 ± 0.085
	R-CCCGTCTCTGAAAAATCCAT			
SG6	F-TGCCTCTGAGTTGAATTTGC	(CA)8-(CA)18-(AG)8	95.8	0.256 ± 0.112
	R-GGAACCAAGGTGGTGAAACT			
SG7	F-AAGTTCAATCCCACCCTTTG	(GT)11	96.5	0.236 ± 0.077
	R-ACCAAACGGCAACTCACATA			
SG10	F-CTAGCCGATCTGTATCGCAA	(CA)16-(TA)7-(TA)11	97.9	0.264 ± 0.088
	R-GTTGGTGATCTGCCCCATA			
SG14	F-CCTTCCATTCAGATGGTTCC	(AC)42-(AC)14	100	0.261 ± 0.097
	R-TGTAGGTGGAGGGGTCTTTC			
SG15	F-TGCGTATTTATCCATGACCG	(GT)7-(GT)7-(GA)20	100	0.257 ± 0.089
	R-CATTTAGGGTTGGCGAGAAT			
SG16	F-CCCGTCTCTGAAAAATCCAT	(CA)29	96.7	0.236 ± 0.086
	R-GGAATTCAGAAGGCCCATTA			
SG17	F-CTACTAGGTGATTTCCCCGC	(TG)23	96.1	0.254 ± 0.094
	R-GCTCATGAATCCCTCCCTAA			
SG19	F-TAATCCACACACACACACGC	(CGA)5-(CG)6	96.7	0.252 ± 0.065
	R-AACAAGTTAAGCAGCAGCCA			
SG20	F-AGTACTGGCAGCAGCCTTTT	(TG)13	100	0.227 ± 0.064
	R-CCAAATAAAAGCGCCCTAAC			

[a] Shannon's information index, [b] standard deviation

gels with 1× TAE buffer (pH 8.0) followed by detection under UV light after staining with ethidium bromide.

Microsatellite PCR

PCR amplification of microsatellites was temperature cycled using a programmable thermal cycler (Eppendorf Mastercycler, Hamburg, Germany). In the present study, 30 polymorphic microsatellites primer pairs (Wang et al. 2011) were synthesized (Integrated DNA Technologies; www.idtdna.com) and tested by screening individuals (~ 20 samples randomly selected from various populations) from each switchgrass population. From this, 10 primer pairs, which generated unambiguous DNA fragments and consistent polymorphism, were chosen for subsequent analyses (Table 1). A total of 20 µl reaction mixture consisting of 10× reaction buffer (containing 500 mM KCl, 15 mM $MgCl_2$ and 100 mM Tris–HCl, pH 9.0), 200 µM dNTPs, 0.8 µM primer, 1 U *Taq* polymerase (Fisher Scientific, Pittsburgh, PA, USA) and 25–30 ng of template DNA was used for PCR amplification. Reactions were run at 95 °C for 5 min (initial denaturation) followed by 14 cycles of 94 °C for 20 s (denaturation), 58 °C for 1 min (annealing) and 72 °C for 30 s (extension). This was followed by additional 28 cycles of 94 °C for 20 s (denaturation), 55 °C for 1 min (annealing) and 72 °C for 30 s (extension). The final extension was carried out at 72 °C for 10 min. Separation of the amplified PCR products was carried out on 2.2 % agarose gels with 1× TAE buffer (pH 8.0) while the detection was under UV light after staining with ethidium bromide (Supplementary Fig. 2a). Microsatellite experiments were repeated thrice, and the consistency of banding patterns was confirmed. QIAxcel capillary electrophoresis system (Qiagen, Valencia, CA, USA) with a 25-bp DNA size marker (Qiagen, Frederick, MD, USA) was utilized to analyze the microsatellite PCR products (Supplementary Fig. 2b). BioCalculator™ software automatically documented and exported the data obtained. The software also produced a gel image and an electropherogram of the separated PCR products (Wang et al. 2009).

Chloroplast DNA PCR, gel purification and sequencing

Primers specific to the intergeneric region of *atp*I–*atp*H *ndh*A, *mat*K, *psb*J–*pet*A, *trn*CD and *trn*T(UGU)–*trn*L(UAA) (Taberlet et al. 1991; Sang et al. 1997; Shaw et al. 2007) were synthesized (Integrated DNA Technologies; www.idtdna.com) and utilized for the chloroplast gene PCR amplification of the switchgrass leaf DNA samples. DNA for the population comparison was prepared from pooled leaf samples of twelve plants each from the natural and agronomic populations tested (Lannér et al. 1997; Zeng et al. 2012). The contents in the PCR reaction mixture were as outlined above for microsatellites except 0.8 μM of the respective gene specific chloroplast primer pair was utilized (Table 2). The template DNA used in this experiment was 25–50 ng. A programmable thermal cycler (Eppendorf Mastercycler, Hamburg, Germany) was utilized for the PCR amplification. The initial denaturation was carried out at 94 °C for 3 min. This was followed by 39 cycles of 94 °C for 1 min (denaturation), 50 °C for 1 min (annealing) and 72 °C for 2 min (extension). The final extension was carried out at 72 °C for 20 min. Only for the primer pair specific to *psb*J-*pet*A region, annealing was carried out at 58 °C. PCR amplification products were separated on 1.6 % agarose gels with 1× TAE buffer (pH 8.0), detected under UV light after ethidium bromide staining, and the amplicons expected with respect to primer pairs used were confirmed (Table 2). The PCR products generated were extracted from the gel and gene-cleaned using a QIAquick PCR purification kit (Qiagen, Frederick, MD, USA). The purified PCR products were sequenced at the University of Tennessee Core Sequencing Facility.

Genetic data scoring and analysis

Microsatellite electropherograms were scored using binary coding with the presence of a PCR amplified product being scored as '1' and its absence as '0'. Estimation of genetic diversity parameters was carried out using POPGENE software (Yeh and Boyle 1997). The degree of polymorphism was assessed for each population by calculating the quantity of polymorphic amplification products. The frequency of the most recurrent microsatellite product had to be <95 % for it to be considered as polymorphic. Shannon's gene diversity index (calculated as $h = 1 - \Sigma\ Pi^2$, where 'Pi' is the frequency of the occurrence of the 'ith' amplified product over individuals within a population) was analyzed and subjected to *t* test analysis. The total genetic diversity from all the populations (H_T) and the mean diversity within each population (H_S) were calculated using POPGENE (Yeh and Boyle 1997). The total genetic diversity existing between the populations (G_{ST}) was estimated using the H_T and H_S values in the formula $G_{ST} = (H_T - H_S)/H_T$ (Nei 1973). To develop a dendrogram, we utilized cluster analysis following unweighted pair group method with arithmetic mean (UPGMA algorithm) using the squared Euclidean distance between individuals. To demonstrate the genetic relationships among individuals across each study site, we utilized an ordination generated by principal component analysis (PCA) using STATISTICA version 4.5 (Statsoft 1993; Nei 1972).

Genetic clustering of individuals within and among natural and agronomic switchgrass populations was performed using STRUCTURE (v.2.2), which employs a Bayesian clustering algorithm (Pritchard et al. 2000; Hadziabdic et al. 2010). Applying an admixture model, presuming that each sample depicted a part of its genome from each of *K* subpopulations and without any information on the origin of the population, we ran STRUCTURE using Markov chain Monte Carlo iterations having 10,000 generation runs and a burn-in period of 10,000 iterations. Using the ad hoc statistic ΔK, which is based on the second order rate of change in the log probability of the data between successive values of *K*, appropriate number of population clusters were identified (Evanno et al. 2005). To verify the reliability of the results obtained from independent Markov Chain results, an additional 1,000,000 generations and a burn-in period of 1,000,000 iterations were also performed. Within each population, the genetic similarity between all possible pairs of individuals (Sneath and Sokal 1972) was also computed.

The cp-DNA sequence comparison was analyzed using ClustalW Multiple Sequence Alignment (dot.imgen.bcm.tmc.edu) and a cluster analysis was performed following UPGMA neighbor-joining method (MAFFT sequence program). The pair-wise nucleotide similarity among natural as well as agronomic switchgrass cultivars were calculated using EMBOSS pair-wise alignment algorithms (http://www.ebi.ac.uk/Tools/psa/).

Table 2 Chloroplast primer pair sequences used for sequencing in switchgrass populations

Primer pairs	Sequence (5′ → 3′)	Gene loci name	Annealing temp. (°C)	Amplification length (bp)
SGCP21	F-CATTACAAATGCGATGCTCT	*trn*T(UGU)–*trn*L(UAA)	50	250–770
	R-TCTACCGATTTCGCCATATC			
SGCP22	F-ATAGGTACTGTARCTGGTATT	*psb*J–*pet*A	58	1,269
	R-AACAGTTTGAAAAGGTTCARTT			
SGCP23	F-TATTTACAAGCGGTATTCAAGCT	*atp*I–*atp*H *ndh*A	50	1,157
	R-CCAATCCAGCAGCAATAAC			
SGCP51	F-GGTTCAAGTCCCTCTATCCC	*trn*CD	50	500
	R-ATTTGAACTGGTGACACGAG			
SGCP52	F-ACTGTATCGCACTATGTATCA	*mat*K	50	1,800
	R-GAACTAGTCGGATGGAGTAG			

Results

Microsatellite genetic diversity

A total of ten microsatellite primer pairs were utilized in this switchgrass population genetic diversity study. These primers were found to be polymorphic for all the populations assessed in the present study. The percent polymorphism across each microsatellite primer pair ranged from 100 (in SG4, SG14, SG15 and SG20) to 95.8 (SG6, Table 1). The SG10 microsatellite primer pair exhibited the highest mean genetic diversity (0.264 ± 0.088), while the least mean genetic diversity was observed in SG20 microsatellite primer pair (0.227 ± 0.064). The mean number of alleles, the percent polymorphism and the mean genetic diversity observed across the switchgrass population studied are presented in Table 3. The agronomic switchgrass cultivar 'Alamo' from Vonore and Alcoa study sites had the highest mean number of alleles (15.0 and 13.7, respectively); in contrast, the natural switchgrass populations such as Morrison Meadow "East" and Rowe Gap had the least mean number of alleles (5.2 and 6.1, respectively). Microsatellite polymorphism ranged from 91.4 % in the agronomic cultivar 'Kanlow' from Vonore to 100 % in the natural population sampled from May Prairie. Morrison Meadow "East" and Rowe Gap had significantly higher genetic diversity as compared to agronomic switchgrass cultivars as well as to other natural switchgrass populations (Table 3). Overall, the natural switchgrass populations had significantly higher mean genetic diversity (0.262 ± 0.102) as compared to agronomic 'Alamo' and 'Kanlow' switchgrass cultivars (0.201 ± 0.082; Fig. 1). The total genetic diversity residing within (H_S) and among (H_T) populations were also significantly higher for the natural switchgrass populations as compared to agronomic switchgrass cultivars (Table 4). Further, pairwise population genetic similarity values among natural switchgrass populations are lower than the agronomic cultivars (0.59 vs. 0.76), Table 5.

Along with the evaluation of the genetic diversity, attempts were also made to assess the genetic structure of the natural and agronomic switchgrass populations. Principal component analysis (PCA) grouped the natural and agronomic switchgrass populations based on their geographic occurrence (Fig. 2). The STRUCTURE analysis revealed a clear clustering of populations (Fig. 3). An ad hoc quantity (ΔK), based on the rate of change of likelihood function, showed a clear peak at $K = 4$ (four population clusters). These clusters were more or less spatially contiguous and corresponded well with the PCA grouping of populations. The agronomic cultivars 'Alamo' and 'Kanlow' were grouped separately from the natural switchgrass populations Rowe Gap, Morrison Meadow "East", Morrison Meadow, Brockdell Road and May Prairie (Figs. 2, 3). While the microsatellite primer pairs used in this study clearly delineated the switchgrass populations based on the geographic region of occurrence, they were not able to differentiate among the switchgrass cultivars (Fig. 4). The dendrogram as well as STRUCTURE analysis revealed that the switchgrass cultivars 'Alamo' and 'Kanlow' were genetically more similar (Figs. 3, 4).

Table 3 Population study sites and genetic parameters for switchgrass populations

Study sites	Sample size	Mean number of alleles	Polymorphism (%)	Mean genetic diversity[#] (±SD[$])
Brockdell Road^ (Van Buren County)	14	12.5	97.6	0.207 ± 0.092^{b}
Morrison Meadow^ (Warren County)	12	13.5	95.6	0.197 ± 0.073^{b}
Morrison Meadow^ "East" (Warren County)	6	5.2	95.6	0.340 ± 0.128^{a}
May Prairie^ (Coffee County)	12	13.3	100	0.235 ± 0.097^{ab}
Rowe Gap^ (Franklin County)	6	6.1	96.7	0.334 ± 0.119^{a}
Vonore 'Alamo'* (Monore County)	13	15.0	98.9	0.195 ± 0.069^{b}
Vonore 'Kanlow'* (Monore County)	12	10.4	91.4	0.209 ± 0.103^{b}
Alcoa 'Alamo'* (Blount County)	13	13.7	97.8	0.199 ± 0.072^{b}

^ Natural populations. * Agronomic switchgrass cultivars. [#] Shannon's information index with different letters are significant at $p < 0.05$ (t test). [$] standard deviation

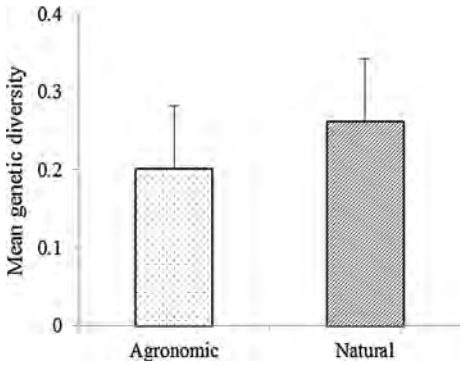

Fig. 1 Mean genetic diversity of natural and agronomic switchgrass populations (t test, $p < 0.008$)

Table 4 Mean genetic diversity and differentiation within and among switchgrass populations

	$H_S^{\dagger}$	$H_T^{\ddagger}$	$G_{ST}^{\S}$
Among natural populations	0.278 ± 0.040^{a}	0.208 ± 0.031^{a}	0.250
Among agronomic populations	0.221 ± 0.032^{b}	0.142 ± 0.016^{b}	0.357
Among all populations	0.274 ± 0.040	0.183 ± 0.021	0.332

The H_S, H_T, G_{ST} were calculated using POPGENE (Yeh and Boyle 1997)

[†] Mean genetic diversity within populations ($p < 0.0001$)

[‡] Total genetic diversity from all concerning populations ($p < 0.0001$)

[§] Total genetic diversity residing among populations (or population differentiation)

Chloroplast sequencing results

All the five chloroplast primer pairs were able to amplify clear PCR products from the switchgrass DNA. The PCR amplification product sizes ranged from 250 to 1,800 bp depending on the chloroplast primer used (Table 2, Supplementary Fig. 3a). To reveal the chloroplast nucleotide genome variation, the nucleotide sequence data was subjected to BLASTn analysis against the public GenBank database domain (http://www.ncbi.nlm.nih.gov/) using search and the amplified chloroplast DNA regions were re-confirmed. The in silico clustering analysis for all five chloroplast nucleotide sequences grouped 'Alamo' predominantly with the switchgrass chloroplast sequences and closely related monocot species (Supplementary Fig. 3b). The nucleotide sequences were aligned and the gene contigs were developed using VectorNTI software (Life Technologies, Grand island, NY, USA) (Supplementary Fig. 3c). To differentiate the agronomic switchgrass cultivars from the natural populations, samples of lowland (*P. dichotomiflorum*) and upland ('Cave-in-Rock') switchgrass were also included in the cp-DNA PCR and sequencing experiment. Following UPGMA (Sneath and Sokal 1972) as well as neighbor-joining method (involving 1,000 bootstrap values), a cladogram was generated. The pair-wise chloroplast nucleotide sequence similarity values are presented in Table 6. The tree generated was able to discriminate the switchgrass populations, and the clustering was approximated with their ploidy level variations

Table 5 Pairwise genetic similarity measures of switchgrass populations using microsatellites

	Alcoa 'Alamo'	Vonore 'Alamo'	Vonore 'Kanlow'	Brockdell Road	Morrison Meadow	Morrison Meadow "East"	May Prairie	Rowe Gap
Alcoa 'Alamo'	****	0.80	0.82	0.83	0.92	0.36	0.94	0.45
Vonore 'Alamo'		****	0.66	0.68	0.76	0.16	0.78	0.28
Vonore 'Kanlow'			****	0.73	0.83	0.26	0.84	0.36
Brockdell Road				****	0.80	0.38	0.82	0.39
Morrison Meadow					****	0.31	0.94	0.40
Morrison Meadow "East"						****	0.42	0.92
May Prairie							****	0.52
Rowe Gap								****

Genetic similarity measures were calculated using STATISTICA (Statsoft 1993; Sneath and Sokal 1972)

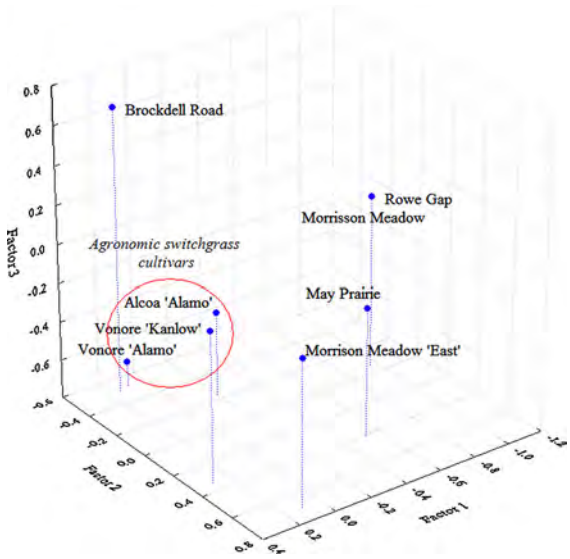

Fig. 2 Principal component analysis of switchgrass populations. Agronomic switchgrass populations are grouped inside the *red circle*. The first three axes of the PCA for the switchgrass populations were 30.26, 18.14, 15.22 % and explained a total variance of 63.62 %. (Color figure online)

(Fig. 5). Two major clusters were noted, wherein one of these contained the upland octaploid switchgrass ('Cave-in-Rock', outgroup), and the other contained lowland tetraploid cultivars (see Zalapa et al. 2011), the natural tetraploid switchgrass populations as well as tetraploid *P. dichotomiflorum*. Within the latter cluster, the agronomic cultivars ('Alamo' and 'Kanlow') clustered together, whereas the natural switchgrass populations formed a separate group.

Discussion

In this present study, the characterization of population genetic structure of natural and agronomic switchgrass populations using microsatellites and chloroplast molecular markers in Tennessee revealed that the native populations had higher genetic diversity than the agronomic cultivar populations. The levels of microsatellite primer pair polymorphism, the mean number of alleles observed and genetic diversity estimated in this study (Tables 1, 3) were on par with

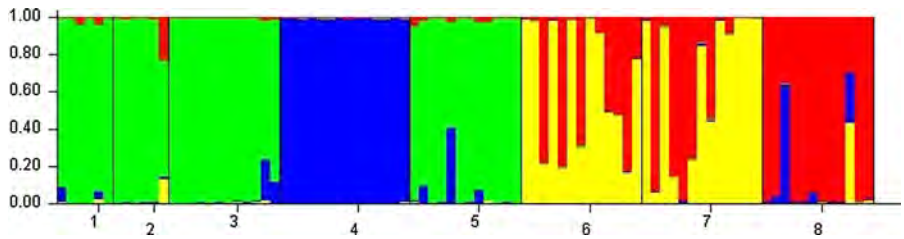

Fig. 3 Bayesian clustering analysis of natural (*1* Rowe Gap, *2* Morrison Meadow "East", *3* Morrison Meadow, *4* Brockdell Road, *5* May Prairie) and agronomic (*6* Alcoa 'Alamo', *7* Vonore 'Alamo', *8* Vonore 'Kanlow') switchgrass populations using STRUCTURE (v.2.2) program (Pritchard et al. 2000)

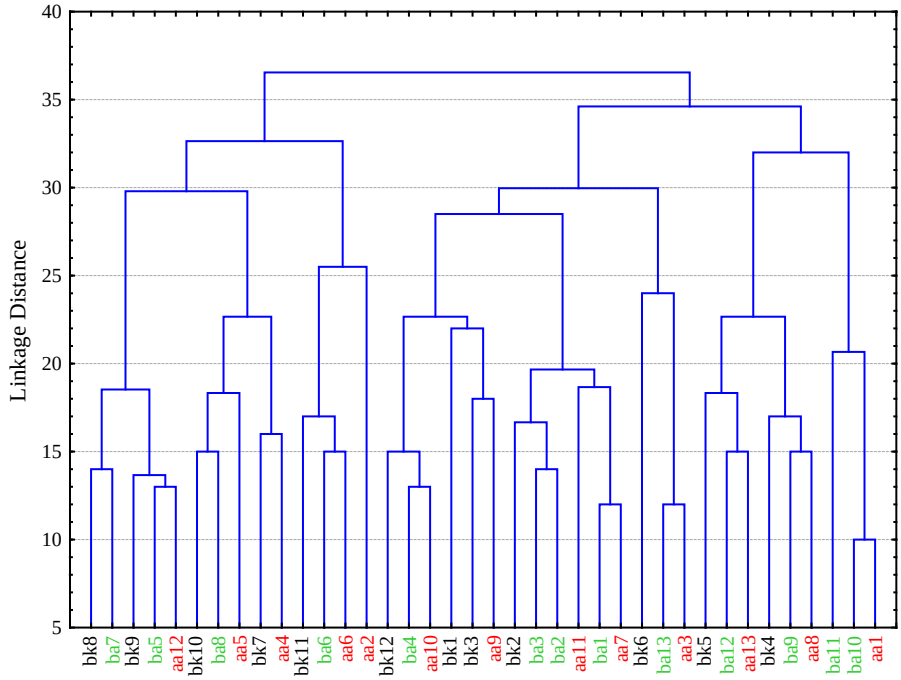

Fig. 4 Cluster analysis of agronomic switchgrass (aa, ba = 'Alamo' and bk = 'Kanlow') populations

Table 6 Pairwise similarity (in percentages) measures of switchgrass using chloroplast DNA nucleotide sequences

	'Alamo'	'Kanlow'	May Prairie	Brockdell Road	Morrison Meadow	*P. dichotomiflorum*	'Cave-in-Rock'
'Alamo'	****	98.2	96.7	98.3	98.3	95.9	94.5
'Kanlow'		****	98.0	99.7	99.6	97.4	95.6
May Prairie			****	98.2	98.3	98.5	96.9
Brockdell Road				****	99.6	97.4	95.8
Morrison Meadow					****	97.3	95.9
P. dichotomiflorum						****	96.2
'Cave-in-Rock'							****

Genetic similarity measures were calculated using EMBOSS pairwise alignment algorithms

the other switchgrass population genetic studies reported in the literature (Narasimhamoorthy et al. 2008; Huang et al. 2011; Zalapa et al. 2011; Zhang et al. 2011). The genetic diversity measured in this study was also similar to that noted for other self-incompatible, perennial grass species such as *Lolium perenne* L. (Hu et al. 2011), *Axonopus fissifolius* (Raddi) Kuhlm (Wang et al. 2010) and *Andropogon gerardii* Vitman (Selbo and Snow 2005). The genetic diversity of the natural switchgrass populations in Morrison Meadow "East" and Rowe Gap was significantly higher than the agronomic switchgrass cultivars (Table 3; Fig. 1). These two natural populations are small, fragmented, and also contained the least number of switchgrass individuals. Fragmented populations such as these have been shown to have distinct ecological and genetic structures (Gascon et al. 2000; Laurance 2000; Pimm and Raven 2000; Young and Boyle 2000; Uma Shaanker et al. 2001; Ravikanth et al. 2002; Honnay et al. 2005; Nageswara-Rao et al. 2007). Restricted gene flow among such fragments and their gene pool could facilitate population divergence (Hamrick and Schnabel 1985; Boshier et al. 1995; Wu and Vankat 1995; Rajanikanth

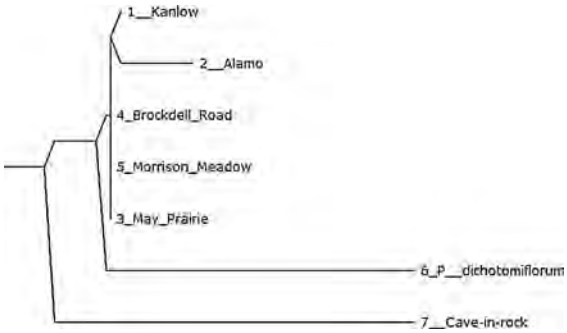

Fig. 5 Clustering analysis of agronomic switchgrass cultivars (1_'Kanlow', 2_'Alamo'), natural (3_May Prairie, 4_Brockdell Road, 5_Morrison Meadow), and 6_*Panicum dichotomiflorum* Michx, 7_'Cave-in-Rock' samples based on all five chloroplast primer pair (see Table 2) nucleotide sequencing

et al. 2010). The tendency of certain alleles to persist in a given population suggests that there has been independent evolution of populations into small fragments. Our results revealed that the natural switchgrass populations, although small in their size, add significantly to the diversity of switchgrass genetics.

The total genetic diversity value residing among switchgrass populations ($G_{ST} = 0.33$, Table 4) indicated that most of the variability (67 %) was observed in individuals within populations. This result corroborates research reported in other switchgrass germplasms (e.g., Gunter et al. 1996; Casler et al. 2007; Narasimhamoorthy et al. 2008) and in other cross-pollinated, perennial plant species (Hamrick and Godt 1989). We observed that the natural switchgrass populations were genetically more diverse and had less mean genetic similarity as compared to agronomic switchgrass cultivars (Table 5). Similar observations of a decrease in genetic similarity with growing geographical distance have been reported in wild perennial ryegrass (*L. perenne* L.) populations (Monestiez et al. 1994). Studies carried out on cultivated and wild perennial ryegrass populations as well as *Agrostis curtisii* populations also showed the wild populations to be more diverse than the cultivated populations (Warren et al. 1998). Such studies highlight the importance of maintaining the genetic diversity of natural populations and some potential effects of the introgression of natural populations with genes from improved varieties.

The microsatellite-PCR and the chloroplast nucleotide sequence diversity data analysis consistently discriminated the individuals from different natural and agronomic switchgrass populations into their expected ecotypes. Using chloroplast DNA restriction fragment length polymorphism, Hultquist et al. (1996) were able to separate switchgrass populations into upland and lowland ecotypes. However, that study contrasted with our findings in which no genetic variation was associated with ecotype (Fig. 5). Studies have reported clustering of switchgrass populations based on ecotype by utilizing samples obtained mainly from the USDA Germplasm Resources Information Network. Gunter et al. (1996) were able to group switchgrass populations into upland and lowland ecotypes based on RAPD genetic coefficients, while Cortese et al. (2010) reported the clustering of populations into separate clades based on both their geographic occurrence as well as their ecotypes using EST-SSRs. Casler et al. (2007) also reported the switchgrass genotype differentiation based on their hardiness zones (upland varieties grow in the USDA hardiness zones 3–8, whereas lowland varieties grow in zones 5–10). Distinction of ecotypes using SSRs was also possible but not by utilizing only chloroplast markers (Zalapa et al. 2011). In this study, using both microsatellite as well as chloroplast nucleotide sequences, all the methods analyzed (genetic diversity, similarity, principal component analysis, dendrogram and STRUCTURE clustering as well as chloroplast UPGMA neighbor-joining analysis) were able to discriminate the natural and agronomic switchgrass populations.

Anthropogenic pressures such as expansion of agriculture, deforestation, and commercial developmental activities pose a challenge to the sustainable conservation of natural grasslands in the mid-southern region of the United States. The growing demand to meet the United States government-mandated advanced biofuels targets has further intensified these challenges (McDonald et al. 2009; Kwit et al. 2014). A scenario in which new areas of monoculture of high biomass-yielding agronomic cultivars of switchgrass in this region might affect the genetics of native wild stands, or introgression or dispersal of agronomic switchgrass genes into wild populations (Zhang et al. 2011; Kwit and Stewart 2012; Nageswara-Rao et al. 2013; Ridley et al. 2013). Dispersal of seeds or hybrids resulting from introgression could alter the genetic structure of native populations (Raghu et al. 2006; Kwit et al. 2011; Kwit and Stewart 2012; Lewis 2013; Ridley et al. 2013).

Future plantings of agronomic switchgrass cultivars for bioenergy would likely involve use of selected cultivars of limited geographic origin (e.g., 'Alamo' from Texas, 'Kanlow' from Oklahoma), and possibly transgenic switchgrass with increased potential for biofuel yield. The risk of transgene or agronomic gene introgression into natural switchgrass populations, and their regulatory concerns, have been raised by many researchers (Kausch et al. 2010; Kwit and Stewart 2012; Stottlemyer 2012; Lewis 2013). The native-agronomic gene exchange may be possible and discernible, especially when its population ploidy levels are the same (Martinez-Reyna and Vogel 2002). In restoration of grasslands, if the origin of seed mixtures do not match that of the remnant population, it could lead to mixed ploidy populations, which may result in lower reproductive success for the remnant population (Delaney and Baack 2012). For introgression to occur there must be proximity of improved agronomic plantings to natural populations and the spatio-temporal flowering phenologies must overlap.

A clear understanding of the population genetic variability of natural switchgrass populations before significant landscape alteration takes place will allow the assessment and detection of future introgression. The majority of the genetic variation in mid-southern switchgrass populations exist among individuals within populations. Our study also noted that the genetic diversity of improved agronomic cultivars, 'Alamo' and 'Kanlow', differs notably from those typifying natural switchgrass populations. Such differences may be instrumental in documenting future changes in genetic diversity as a function of introgression. Further, our results will help in addressing the issues of bioenergy sustainability which would, in turn, help regional, state, and federal policy-makers as and when imminent decisions are made on how to best change our landscape to incorporate bioenergy feedstocks. Lignocellulosic systems, that involve crops in a landscape containing wild relatives, should include conservation measures for the native populations. In situ conservation of crop wild relatives can further assist with crop sustainability in the future.

Acknowledgments We thank numerous people who facilitated and assisted with field work, including B. Black, S. Bobzin, T. Crabtree, S. Jackson. We also thank G. Wein, X. Yang, D. Hadziabdic, R. Govindarajulu and P. A. Wadl for their assistance with laboratory and logistical assistance. Permits to collect switchgrass tissue from Tennessee State Natural Areas were obtained through the Tennessee Department of Environment and Conservation. This project was supported by Biotechnology Risk Assessment Grant Program competitive Grant No. 2010-39211-21699 from the USDA National Institute of Food and Agriculture (NIFA) as well as by Building Research Interest and Developing Global Engagement (BRIDGE), an Inter-College Undergraduate Research Funding Opportunity, awarded to M. Hanson. Neal Stewart Jr. also received support from the BioEnergy Science Center, a Bioenergy Research Center, supported by the Office of Biological and Environmental Research in the US Department of Energy Office of Science and funding from NIFA to the University of Tennessee Integrated Biomass Supply Systems (IBSS) Center.

References

Barnett FL, Carver RF (1967) Meiosis and pollen stainability in switchgrass, *Panicum virgatum* L. Crop Sci 7:301–304

Boshier DH, Chase MR, Bawa KS (1995) Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 3. Gene flow, neighborhood, and population structure. Am J Bot 82:484–490

Burton G (1942) A cytological study of some species in the tribe Paniceae. Am J Bot 29:355–359

Casler MD, Stendal CA, Kapich L, Vogel KP (2007) Genetic diversity, plant adaptation regions, and gene pools for switchgrass. Crop Sci 47:2261–2273

Church GL (1940) Cytotaxonomic studies in the gramineae *Spartina*, *Andropogon*, and *Panicum*. Am J Bot 27:267–271

Clauss MJ, Cobban H, Mitchell-Olds T (2002) Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaeae). Mol Ecol 11:591–601

Cortese LM, Honig J, Miller C, Bonos SA (2010) Genetic diversity of twelve switchgrass populations using molecular and morphological markers. Bioenergy Res 3:262–271

Costich DE, Friebe B, Sheehan MJ, Casler MD, Buckler ES (2010) Genome-size variation in switchgrass (*Panicum virgatum*): flow cytometry and cytology reveal rampant aneuploidy. Plant Genome 3:130–141

Delaney JT, Baack EJ (2012) Intraspecific chromosome number variation and prairie restoration—a case study in Northeast Iowa, U.S.A. Restor Ecol 20:576–583

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620

Gascon C, Williamson GB, da Fonseca GAB (2000) Receding forest edges and vanishing reserves. Science 288:1356–1358

Gunter LE, Tuskan GA, Wullschleger SD (1996) Diversity among populations of switchgrass based on RAPD markers. Crop Sci 36:1017–1022

Hadziabdic D, Fitzpatrick BM, Wang X, Wadl PA, Rinehart TA, Ownley BH, Windham MT, Trigiano RN (2010) Analysis of genetic diversity in flowering dogwood natural stands using microsatellites: the effects of dogwood anthracnose. Genetica 138:1047–1057

Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding and genetic resources. Sinauer Associates, Sunderland, pp 43–63

Hamrick JL, Schnabel A (1985) Understanding the genetic structure of plant populations: some old problems and new approaches. In: Gregorius HR (ed) Population genetics in forestry. Springer, Berlin, pp 50–70

Honnay O, Jacquemyn H, Bossuyt B, Hermy M (2005) Forest fragmentation effects on patch occupancy and population viability of herbaceous plant species. New Phytol 166:723–736

Hopkins AA, Taliaferro CM, Murphy CD, Christian D (1996) Chromosome number and nuclear DNA content of several switchgrass populations. Crop Sci 36:1192–1195

Hu T, Li H, Li D, Sun J, Fu J (2011) Assessing genetic diversity of perennial ryegrass (*Lolium perenne* L.) from four continents by inter simple sequence repeat (ISSR) markers. Afr J Biotechnol 10:19365–19374

Huang L-K, Bughrara SS, Zhang X-Q, Bales-Arcelo CJ, Bin X (2011) Genetic diversity of switchgrass and its relative species in *Panicum* genus using molecular markers. Biochem Syst Ecol 39:685–693

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

Kausch AP, Hague J, Oliver M, Li Y, Daniell H, Maschia P, Watrud LS, Stewart CN Jr (2010) Transgenic biofuel feedstocks and strategies for bioconfinement. Biofuels 1:163–176

Kwit C, Stewart CN Jr (2012) Gene flow matters in switchgrass (*Panicum virgatum* L.), a potential widespread biofuel feedstock. Ecol Appl 22:3–7

Kwit C, Moon HS, Warwick SI, Stewart CN Jr (2011) Transgene introgression in crop relatives: molecular evidence and mitigation strategies. Trends Biotechnol 29:284–293

Kwit C, Nageswara-Rao M, Stewart CN Jr (2014) Switchgrass: concerns, compliance and future prospects. In: Hong L, Wu Y (eds) Compendium of bioenergy plants: switchgrass. CRC Press, Taylor and Francis Group, Florida, pp 403–421

Lannér C, Bryngelsson T, Gustafsson M (1997) Relationships of wild *Brassica* species with chromosome number 2n = 18, based on RFLP studies. Genome 40:302–308

Laurance WF (2000) Do edge effects occur over large spatial scales? Trends Ecol Evol 15:134–135

Lewis E (2013) The rate and characterization of hybridization between wild-type and cultivated switchgrass (*Panicum virgatum* L.) for biofuel use. A senior honors dissertation, The Ohio State University

Martinez-Reyna JM, Vogel KP (2002) Incompatibility systems in switchgrass. Crop Sci 42:1800–1805

McDonald RI, Fargione J, Kiesecker J, Miller WM, Powell J (2009) Energy sprawl or energy efficiency: climate policy impacts on natural habitat for the United States of America. PLoS One 4:e6802. doi:10.1371/journal.pone.0006802

McLaughlin SB, Kszos LA (2005) Development of switchgrass (*Panicum virgatum*) as a bioenergy feedstock in the United States. Biomass Bioenergy 28:515–535

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 Crop Evol 53:1291–1302

Monestiez P, Goulard M, Charmet G (1994) Geostatistics for spatial genetic structures: study of wild populations of perennial ryegrass. Theor Appl Genet 88:33–41

Morgante M, Olivieri AM (1993) PCR-amplified microsatellite as markers in plant genetics. Plant J 3:175–182

Moser LE, Vogel KP (1995) Switchgrass, big bluestem, and Indiangrass. In: Barnes RF, Miller DA, Nelson CJ (eds) Forages, vol 1., An introduction to grassland agriculture, Iowa State University Press, Ames, pp 409–420

Nageswara-Rao M, Ganeshaiah KN, Uma Shaanker R (2007) Assessing threats and mapping sandal (*Santalum album* L.) resources in peninsular India: identification of genetic hotspot for in situ conservation. Conserv Genet 8:925–935

Nageswara-Rao M, Soneji JR, Chen C, Huang S, Gmitter FG Jr (2008) Characterization of zygotic and nucellar seedlings from citrus rootstock candidates using RAPD and EST-SSR markers. Tree Genet Genomes 4:113–124

Nageswara-Rao M, Soneji JR, Kwit C, Stewart CN Jr (2013) Advances in biotechnology and genomics of switchgrass. Biotechnol Biofuels 6:77. doi:10.1186/1754-6834-6-77

Narasimhamoorthy B, Saha MC, Swaller T, Bouton JH (2008) Genetic diversity in switchgrass collections assessed by EST-SSR markers. Bioenergy Res 1:136–146

Nei M (1972) Genetic distance between populations. Am Nat 106:283–292

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Nielsen EL (1944) Analysis of variation in *Panicum virgatum* L. J Agric Res 69:327–353

Pimm SL, Raven P (2000) Biodiversity: extinction by numbers. Nature 403:843–845

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Raghu S, Anderson RC, Daehler CC, Davis AS, Wiedenmann RN, Simberloff D, Mack RN (2006) Adding biofuels to the invasive species fire? Science 313:1742

Rajanikanth G, Nageswara-Rao M, Tambat B, Uma Shaanker R, Ganeshaiah KN, Kushalappa CG (2010) Are small forest fragments more heterogeneous among themselves than the large fragments? Biorem Biodivers Bioavailab 4:42–46

Ravikanth G, Ganeshaiah KN, Uma Shaanker R (2002) Identification of hotspots of species richness and genetic variability in rattans: an approach using geographical information systems (GIS) and molecular tools. Plant Genet Resour Newsl 132:17–21

Ridley CE, Jager HI, Clark CM, Efroymson RA, Kwit C, Landis DA, Leggett ZH, Miller DA (2013) Debate: can bioenergy be produced in a sustainable manner that protects biodiversity and avoids the risk of invaders? Bull Ecol Soc Am 94:277–290

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). Am J Bot 84:1120–1136

Selbo SM, Snow AA (2005) Flowering phenology and genetic similarity among local and recently introduced populations of *Andropogon gerardii* in Ohio. Restor Ecol 13:441–447

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose

noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am J Bot 94:275–288
Sneath PHA, Sokal RR (1972) Numerical taxonomy: the principles and practice of numerical classification. WH Freeman, San Francisco
Statsoft (1993) Statistica ver 4.5. Statsoft Inc, Tulsa
Stottlemyer AL (2012) Investigating hybridization potential, components of fitness, and volunteerism in wild and cultivated *Panicum virgatum* L. (switchgrass). M.S. dissertation, The Ohio State University
Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three noncoding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109
Talbert LE, Timothy DH, Burns JC, Rawlings JO, Moll RH (1983) Estimates of genetic parameters in switchgrass. Crop Sci 23:725–728
Uma Shaanker R, Ganeshaiah KN, Nageswara-Rao M (2001) Genetic diversity of medicinal plant species in deciduous plants of India: impacts of harvesting and other anthropogenic pressures. J Plant Biol 28(1):91–97
Vogel KP (2004) Switchgrass. In: Moser LE, Burson BL, Sollenberger LE (eds) Warm-season (C_4) grasses. Agronomy monograph 45. ASA, CSSA, and SSSA, Madison, pp 561–588
Wang X, Rinehart TA, Wadl PA, Spiers JM, Hadziabdic D, Windham MT, Trigiano RN (2009) A new electrophoresis technique to separate microsatellite alleles. Afr J Biotechnol 8:2432–2436
Wang Z, Kenworthy KE, Wu Y (2010) Genetic diversity of common carpetgrass revealed by amplified fragment length polymorphism markers. Crop Sci 50:1366–1374
Wang YW, Samuels TD, Wu YQ (2011) Development of 1,030 genomic SSR markers in switchgrass. Theor Appl Genet 122:677–686
Warren JM, Raybould AF, Ball T, Gray AJ, Hayward MD (1998) Genetic structure in the perennial grasses *Lolium perenne* and *Agrostis curtisii*. Heredity 81:556–562
Wright L, Turhollow A (2010) Switchgrass selection as a "model" bioenergy crop: a history of the process. Biomass Bioenergy 34:851–868
Wu J, Vankat JL (1995) Island biogeography: theory and applications. In: Nierenberg WA (ed) Encyclopedia of environmental biology 2. Academic Press, San Diego, pp 371–379
Wullschleger SD, Davis EB, Borsuk ME, Gunderson CA, Lynd LR (2010) Biomass production in switchgrass across the United States: database description and determinants of yield. Agron J 102:1158–1168
Yeh FC, Boyle TJB (1997) *POPGENE* Version 1.2 Microsoft window-based software for population genetics analysis. University of Alberta
Young AG, Boyle TJB (2000) Forest fragmentation. In: Young A, Boshier D, Boyle TJB (eds) Forest conservation and genetics—principles and practice. CABI, CSIRO, pp 123–134
Zalapa JE, Price DL, Kaeppler SM, Tobias CM, Okada M, Casler MD (2011) Hierarchical classification of switchgrass using SSR and chloroplast sequences: ecotypes, ploidies, gene pools, and cultivars. Theor Appl Genet 122:805–817
Zeng C-L, Wang G-Y, Wang J-B, Yan G-X, Chen B-Y et al (2012) High-throughput discovery of chloroplast and mitochondrial DNA polymorphisms in Brassicaceae species by ORG-EcoTILLING. PLoS One 7(11):e47284. doi:10.1371/journal.pone.0047284
Zhang Y, Zalapa JE, Jakubowski AR, Price DL, Acharya A, Wei Y, Brummer EC, Kaeppler SM, Casler MD (2011) Post-glacial evolution of *Panicum virgatum*: centers of diversity and gene pools revealed by SSR markers and cpDNA sequences. Genetica 139:933–948