Genetic Linkage Mapping and Transmission Ratio Distortion in a Three-Generation Four-Founder Population of *Panicum virgatum* (L.)


*Samuel Roberts Noble Foundation, Forage Improvement Division, Ardmore, OK 73401 USA and †USDA-Agricultural Research Service, Western Regional Research Center, Albany, CA 94710 USA
Running Title
Four Founder Switchgrass Genetic Map

Key Words
Biomass, Feedstock, Biofuel, Allopolyploid

Corresponding Author
Christian Tobias, USDA-ARS, 800 Buchanan Street, Albany CA, 510-559-6172, christian.tobias@ars.usda.gov
ABSTRACT

Switchgrass (*Panicum virgatum* L.), a warm season, C4, perennial grass, is one of the predominant grass species of the North American tall grass prairies. It is viewed as a high-potential bioenergy feedstock species because it can produce large amounts of lignocellulosic material with relatively few inputs. The objectives of this project were to develop an advanced switchgrass population and use it for the construction of genetic linkage maps and trait characterization. A three-generation, four-founder population was created and a total of 182 progeny of this advanced population were genotyped that included a mixture of self-pollinated and hybrid individuals. The female map integrated both sub-populations and covered 1,629 cM of the switchgrass genome with an average map length of 91 cM per linkage group. The male map of the hybrid progeny covered 1,462 cM with an average map length of 81 cM per linkage group. Average marker density of the female and male maps was 3.9 and 3.5 cM per marker interval, respectively. Based on the parental maps the genome length of switchgrass was estimated to be 1,776 cM and 1,596 cM for the female and male map, respectively. The proportion of the genome within 5 cM of a mapped locus was estimated to be 92% and 93% for the female and male maps, respectively. Thus, the linkage maps have covered most of the switchgrass genome. The assessment of marker transmission ratio distortion found that 26% of the genotyped markers were distorted from either 1:1 or 3:1 ratios expected for segregation of single dose markers in one or both parents respectively. Several regions affected by transmission ratio distortion were found with linkage groups *Ib-m* and *VIIIa-f* most affected.
INTRODUCTION

Switchgrass (Panicum virgatum L.) is a morphologically diverse, warm season, C4, perennial grass, and a predominant species of the tall grass prairies of North America (Bouton, 2008). It has long been used for forage production, for soil conservation plantings due to its large fibrous root system, and is now being developed as a bioenergy crop for second generation biofuel production (Sanderson et al. 2006). Its suitability as a biomass feedstock is due to its high biomass yield potential, relatively simple growth requirements, and wide adaptation (Barney et al. 2009; Elbersen et al. 2001; Evanylo et al. 2005; Fike et al. 2006; Moser and Vogel., 1995). Efforts to improve yield, feedstock quality, and stress tolerance have recently been initiated and some of these efforts have focused on lowland ecotypes with high yield potential. These are tetraploid (2n=4x=36) while upland populations are both tetraploids and octoploids (2n=8x=72). Upland and lowland ecotypes can be distinguished genetically, based on morphology, and also by ecological preference (Casler, 2005; Casler et al. 2004; Das et al. 2004). Lowland ecotypes are adapted to heavier soils, are tolerant of flooding and are generally found in the warmer and wetter regions of the southern USA while upland ecotypes are adapted to lighter soils, dry and cold zones of the middle and northern latitudes of the USA (Hopkins et al. 1996; Hultquist et al. 1996). Upland and lowland ecotypes occupy overlapping distributions, with significant gene flow between groups (Zhang et al. 2011).

To hasten efforts at efficient breeding marker-based methods need to be adopted. Hence, the development of genomic resources, and analysis of both structured and unstructured populations for genetic dissection of quantitative traits is a priority research area for switchgrass. One main goal of genetic mapping is to identify simply inherited markers that are closely linked to quantitative trait loci (QTL) (Jannink and Walsh, 2002). A high density genetic map of switchgrass is essential for effective application of marker-assisted
selection (MAS) in breeding to increase the efficiency of selection and for characterization of specific regions in the genome that hold a special interest for map-based cloning. Mapping of molecular markers distributed throughout the genome of a plant species or at least around the gene of interest is a prelude for the application of molecular approaches in breeding switchgrass for improved yield and tolerance to various biological and environmental stresses.

The segregation patterns of markers in controlled biparental crosses of switchgrass are predicted to behave as a full-sib family derived from two heterozygous parents. The segregation may have different forms as the number of alleles or QTL segregating at a given locus may reach up to four and vary across loci in their informativeness (Wu et al. 2007; Wu et al. 2002). This introduces complexity to mapping and, in addition, linkage phase between QTL and nearby markers may be unknown which can introduce serious biases in estimations of QTL size and affect. A practical approach to this is to treat single-dose alleles segregating in each parent as dominant markers for the purposes of mapping in a two-way pseudo-testcross design (Grattapaglia and Sederoff 1994). Single-dose alleles in common to both parents, though less informative, can also be mapped with this approach. This is advantageous when there are insufficient numbers of fully informative markers (ab x cd) to integrate the parental maps. Transmission ratio distortion (TRD), deviation from the expected Mendelian ratio of segregating alleles at a locus (Sandler et al. 1959), may also affect estimates of recombination fraction. This phenomenon has been commonly encountered in mapping populations (Garcia-Dorado and Gallego, 1992; Lorieux et al. 1995a; Lorieux et al. 1995b) and is increasingly recognized as a potentially powerful evolutionary force which also affects the construction of genetic linkage maps (Lu et al. 2002).

Previous linkage mapping in switchgrass has been performed on a full-sib population of two lowland individuals (Okada et al. 2010), on a lowland x upland full-sib population
Missaoui et al. 2005; Serba et al. 2013) and on the selfed progeny of a heterozygous individual (Liu et al. 2012). Though self incompatibility systems are present in switchgrass Martinez-Reyna and Vogel 2002), these incompatibility systems are not 100% effective. Up to 61% self-fertilized individuals were reported in one controlled cross of a heterozygous northern lowland genotype which allowed linkage map construction (Liu and Wu, 2011; Liu et al. 2012). These studies have consistently demonstrated the nine base chromosomes of switchgrass are inherited in a disomic manner and have clearly demonstrated synteny with related grass species, however there has not yet been definitive assignment of individual chromosomes to the two subgenomes. Here we describe joint linkage analysis in a cross that produced both self-pollinated and hybrid individuals from a population of four founders derived from the two linkage populations above with two parents and 182 progeny. The structure of this population enables marker phase to be unambiguously determined, particularly for those markers with alleles shared in common to both parents. The objectives of this project were (1) to construct a genetic linkage map from this population with markers distributed throughout the switchgrass genome; (2) jointly map both self-pollinated and hybrid progeny as a prelude to analysis of loci underlying TRD and the analysis of QTL present in the parents.

**MATERIALS AND METHODS**

**Mapping population development**

This mapping population, designated AL-NF, was developed from four founders in three generations, and as such is unique from previously reported mapping populations in switchgrass. First, selected genotypes from Kanlow (K5) and Alamo (A4) were crossed at Albany, CA to generate K5 (female) x A4 (male) (ALB) population (Okada et al. 2010). The F1 pseudo-testcross population was field evaluated at Athens, GA and four distinct genotypes
were selected. A second population (NF-UGA) was developed by crossing a lowland selection, AP13, (female) to an upland selection, VS16, (male) at the University of Georgia (Missaoui et al. 2005). The population was further expanded at the Noble Foundation and 251 F1 progenies have been field evaluated at three locations: Ardmore (34° 11' 32" N and 97° 5' 21" W) and Burneyville (33° 53' 20" N & 97° 16' 36" W), OK and Watkinsville (33° 52' 19" N & 83° 24' 20" W), GA. Four distinct genotypes were selected. Genotypes selected from both populations were grown in greenhouses at the Noble Foundation and reciprocal crosses were made in the spring of 2008. A panicle from each of the female and male parent of the selected genotypes was put together in a paper bag. To enhance pollen shedding, the bags were manually tapped every morning. The tillers were tied to a supporting bamboo stick until seed maturity. The seeds were harvested from both the parents to obtain cross and reciprocal seed. The cross with the most seed set e.g. PV281 (female) (selection from ALB) × NFGA472 (male) (selection from NF-UGA), was selected to develop a three-generation and four-founder population (AL-NF). The crossing scheme for population development is illustrated in Figure 1.

Seeds collected from the cross were scarified in diluted sulphuric acid and chilled at 4°C for seven days. The seeds were germinated at 24°C with a 16h photoperiod. Seedlings were grown in a greenhouse at 32°C day and 21°C night temperature. The young plants were propagated by splitting the tillers and then set out in the field following a R-256 (251 progeny, two sets of parents and one Alamo check) honeycomb design. The population has been field evaluated for three years at two locations in Oklahoma: Red River Farm (Burneyville) and Ardmore. A set of 188 plants were randomly selected and genotyped for construction of genetic linkage maps.

**DNA Extraction and Quantification**
Fresh tissue from young leaves of greenhouse-grown plants was collected and frozen in liquid nitrogen. The frozen samples were ground in 2 ml tubes with two zinc plated ultra smooth ball bearings (Daisy Outdoor Products, Rogers, AZ 72757) using a Tissuelyzer (Qiagen Inc., Valencia, CA). The genomic DNA was extracted using DNeasy Plant Mini Kit (QIAGEN Inc., Valencia, CA) following the manufacturer’s recommendation.

**Marker screening**

Several sequence based markers have been developed for switchgrass. In this study, 342 pairs of eSSR (expressed sequence tag – simple sequence repeat) primers developed from publicly available EST resources, 512 pairs of gSSR (genomic SSR) primers developed from (GA/CT)n enriched genomic libraries, and 144 pairs of STS primers developed from switchgrass cellulase and transferase gene sequences were screened for polymorphism. Primers were synthesized by Sigma-Aldrich Corporation (Woodlands, TX) with 18 nucleotides of M13 (5’- TGTAAAACGACGGCCAGT -3’) universal primer sequence added onto the 5’ end of the forward primer (Schuelke, 2000). The M13 (-21) universal primer sequence labeled with PET, VIC, 6-FAM, or NED fluorescent dyes were synthesized by Sigma-Aldrich (St. Louis, MO). All of the 998 primer pairs were prescreened using a panel of two parents and six randomly selected F1 individual plants. The primers that were polymorphic in two parents and/or segregate in the progeny were used for genotyping the whole population.

**PCR protocol**

Approximately 20 ng template genomic DNA was amplified in a total volume of 10 µl. PCR reactions consisted of 1X PCR buffer, 150 µM dNTPs, 0.25 µM forward primer and 1 µM reverse primer, 1 µM of M13 florescent dye, and 0.5 U GoTaq DNA polymerase (Promega, Madison, WI), and brought to volume by adding ddH2O. The PCR were carried out on a GeneAmp PCR System 9700 thermocycler (Life Technologies, Carlsbad, CA) following a
touchdown PCR program. After an initial 3 min denaturation at 95 °C the thermal profile consisted of three phases. The first 6 cycles consisted of a 45 sec denaturation step at 94 °C and 5 min annealing/extension at 68 °C which was reduced by 2 °C per cycle. The next 8 cycles consisted of a 45 sec denaturation at 94 °C, followed by 2 min annealing step at 58 °C that was reduced by 1 °C per cycle, and extension step at 72 °C for 1 min. The final 25 cycles consisted of a 45 sec denaturation step at 94 °C, a 2 min annealing step at 50 °C, and a 1 min extension step at 72 °C. A final 7 min extension step at 72 °C was then performed. This method was observed to enrich the correct product over any non-specific products amplified by the primers.

The amplification products were diluted in formamide and analyzed by capillary electrophoresis with labeled size standards using an ABI3730 and GeneMapper v.3.7 software (Life Technologies). All markers were scored as dominant loci. Markers with more than 15% missing data were removed from analysis.

**Identification of hybrid and self-fertilized populations**

Switchgrass is highly cross-pollinated and known to be self-incompatible. However, varying degrees of self-fertilization have been observed (Okada *et al.* 2010; Liu and Wu, 2011). To identify the selfed plants among the hybrids, the genomic DNA of the four founders, two parents and 188 progeny were genotyped using 12 polymorphic SSR markers segregating in the male parent. The amplified product was visualized and scored as indicated above. The amplicons of each individual were compared with the two parents. Inbred individuals were identified as those that yielded female-specific but lacked male-specific amplicons.

**Linkage analysis**

The linkage analysis was conducted using polymorphic amplicons referred to henceforth as “markers” that were single-dose (SD) in either parent and expected to segregate 1:1, or were shared by both parents (SD x SD) and expected to segregate 3:1. The SD x SD markers were
placed on the map after the marker order created using only SD markers had been fixed. JoinMap 4.0 software (VanOoijen, 2006) was used following out breeding full-sib family (CP) model or an F2 model for self-fertilized individuals using the regression mapping algorithm (Stam, 1993). For the female map, mean recombination fractions and combined likelihoods of odds (LOD) values of markers from self and hybrid groups were used to integrate the separate linkage groups (LG) and produce the final maps. The SD markers were grouped into linkage groups at independence test LOD threshold of 7.0. The calculations of the linkage map were done using all pair-wise recombination frequency estimate of less than 0.40 and a LOD score larger than 1.0 (ripple value = 1, goodness-of-fit jump threshold = 5, and a triple threshold = 5). The male and female hybrid maps were analyzed following the two-way pseudo-testcross strategy (Grattapaglia and Sederoff, 1994). Ungrouped markers or misgrouped markers were assigned based on their strongest cross link (SCL). If the SCL was greater than 7.0 ungrouped or suspect markers were reassigned for both female and male segregating markers.

Genotyping data from six individuals was excluded from map construction due to either large amounts of missing data or inconsistent genotyping. Map construction for the NF472 map was conducted using marker data from 108 hybrid genotypes while genotypic data from the entire population of 182 (74 self-fertilized, and 108 hybrid) individuals was used for the PV281 map. The combined LOD values and mean recombination fractions were used to integrate the two maps using the regression mapping algorithm. The Kosambi mapping function (Kosambi, 1944) was used to convert recombination units into genetic distances.

Switchgrass genome length was estimated by the formulae \( \sum_{i} L_i \left( \frac{k_i+1}{k_i-1} \right) \) as proposed by Chakravarti et al. (1991) and \( L + \frac{2tL}{n} \) (Fishman et al. 2001), where, \( L \) is the total length of the genetic map (cM), \( n = k-t \) is the number of marker intervals, \( k \) is the number
of marker loci, $k_i$ is the number of marker loci on the $i$th linkage group, where $i=1, 2,..., t$, and $t$ is the number of linkage groups. With the assumption of random distribution of marker loci in the genome, the proportion of the genome within $d$ cM of a marker locus was estimated to be $1- e^{-2dk/G}$, where $k$ is the number of mapped loci and $G$ is the estimated genome length.

For determining homologies between linkage groups and deriving consistent nomenclature of subgenomes with previous linkage analysis (Okada et al. 2010), the populations were genotyped with subsets of previously mapped markers. Resolving homoeologous relationships and alignment of parental LG was performed by: (1) alignment of multiple markers derived from the same primer set present on different LG; (2) assignment of LG based on identity with the founder’s linkage map order determined by Okada et al. 2010; (3) Placement of SD x SD markers. For those few LG with no markers in common, a sequence similarity-based strategy was employed similar to one used with oat (Gutierrez-Gonzalez and Garvin, 2011). Thus EST-STS markers were used to identify collinear stretches of foxtail millet ($Setaria italica$) genome as the criteria for homoeologue formation. EST-STS marker sequences were compared using blastn to the foxtail millet genome (www.phytozome.net; Bennetzen et al. 2012) and 1.0x10^{-5} E-value threshold was used to determine synteny. When switchgrass marker sequences from ungrouped LG aligned to the same $S. italica$ scaffold, irrespective of the position, these LG were considered homoeologous/homologous. Following construction of the LG, map graphics were constructed using MapChart 2.2 (Voorrips, 2002).

**Analysis of Transmission Ratio Distortion**

All markers were scored as dominant loci and those with TRD were identified on the basis of significant deviations from 1:1, 3:1, or 5:1 ratios expected based on allele dosage and di- or tetrasomic inheritance using the $\chi^2$-square test ($p < 0.05$). Markers displaying significant TRD
were excluded from initial map construction and placed at their most likely position after the initial mapping step. To assess the extent of multilocus interactions between unlinked loci associated with single-locus segregation ratio distortion, all mapped TRD–markers were tested for independence against all markers in other LG in both male and female maps using the two-locus genotypes in the two-by-two contingency chi-square test. Since comparisons were made among unlinked SD alleles, the significance was evaluated with a correction for 31,871 comparisons of 28 LG with at least one TRD-marker with 35 other LG ($p < 1.57 \times 10^{-6}$). The distribution of the TRD markers was also assessed in the female and male maps and when three or more TRD markers were clustered together we considered as a candidate TRD locus (Li et al. 2011).

**Comparative Mapping**

Individual maps across studies were compared based on their correlation coefficients for marker interval distance and for marker order based on their longest common subsequence (LCS) using the qualV package (Jachner et al. 2007). Results were expressed as markers present in the LCS as a fraction of the number of markers in common between groups. EST-SSR markers were compared to v. 2.1 of the *Setaria italica* genome ([www.phytozome.net](http://www.phytozome.net)) using blast.

**RESULTS**

**Population development**

Four F1 genotypes from each of ALB and NF-UGA populations were selected with distinct morphologies and a total of 10 successful crosses and reciprocal crosses were made between them. The highest number of seeds (850) was obtained from the cross PV281 × NFGA472 and the reciprocal cross NFGA472 × PV281 provided the second highest seed (746). Poor seed set was observed in the cross between the parents PV204 and NFGA472 where 25 seed
were obtained when NFGA472 served as the pollen donor and 33 were obtained when PV204 served as the pollen donor. The cross with highest seed set (PV281 × NFGA472), was developed as the mapping population ALBxNF. In this cross PV281 was the female parent and NFGA472 was the male parent.

Markers and their inheritance

Among the 998 primer pairs, a total of 447 were genotyped on the entire population. These consisted of 158 (46.2%) eSSR, 272 (53.1%) gSSR, and 17 (11.8%) STS primer pairs that produced a total of 1,203 amplified size polymorphisms (referred to henceforth as markers) between the 2 parents (Table 1). The average number of polymorphic markers per primer set was 2.97, 2.03, and 1.97 respectively for the gSSR, eSSR and STS. Of the total markers, 484 segregated in the female (PV281) and 460 segregated in the male (NF472) parent. The remaining 259 markers were monomorphic in the parents but segregated in the progeny population (Table 2). This result indicated that there was high molecular marker polymorphism between the two parents.

In the hybrid population single-dose markers in either parent are expected to have 1:1 segregation in the hybrid progeny while double-dose markers in only one parent will be expected to segregate 3:1 under strict disomic inheritance or 5:1 under strict tetrasomic inheritance. Chi-square tests of markers segregating in the hybrid population indicated that 390 of the polymorphic markers present in the female parent and 274 markers in the male parent fit a segregation ratio of of 1:1 in hybrid or 3:1 in self-pollinated progeny. In the hybrid population, there were 8 markers that fit 3:1 and 7 markers that fit 5:1 ratios among those that segregated in the female parent; while among those that segregated in the male parent there were 26 markers that fit a 3:1 and 6 markers that fit a 5:1 ratio. The remaining 232 markers (79 in the female and 153 in the male parent) did not fit any expected
segregation ratio. Among 259 markers monomorphic in the parents that segregated in the progeny, 138 fit a 3:1 ratio expected of a single dose marker heterozygous in both parents under either di- or tetrasomic inheritance (Table 2). The remaining 90 markers did not fit any expected segregation ratio. In total 26% of the markers in the hybrids were significantly distorted from expected segregation patterns.

In the AL-NF-self subpopulation, markers affected by TRD were not necessarily identical to those affected in the hybrid subpopulations. Overall 39% of the mapped markers in this population did not fit an expected 3:1 ratio of these 200 markers, only 42 were also affected by TRD in the hybrid subpopulation (data not shown).

**Construction of parental linkage maps**

The molecular markers used to produce the parental linkage maps and the genome coverage is summarized in Table 3. The linkage analysis of 527 polymorphic markers for the female parent map produced from both AL-NF and AL-NF-self subpopulations consisted of 375 SD markers and 111 SD x SD markers that included data from 255 different primer pairs. These were formed into 18 linkage groups. Among the 445 mapped markers in the female map, 318, 116, and 10 were gSSR, eSSR and STS markers, respectively. A total of 82 SD markers remained ungrouped. The female map spanned 1,629 cM. Average length of each LG was 91 cM. An average of 25 polymorphic markers per LG with an average density of 3.8 cM per marker were mapped. The range of polymorphic markers mapped was from 13-43 per LG. The longest linkage group was $V_b-f$ with a total length of 131 cM, while the shortest was $IVa-f$ that had a total length of 65.6 cM (Table 4 and Fig. 2).

Similarly, linkage analysis of 453 polymorphic markers segregating in the male parent map allowed mapping of 315 SD markers and 36 SD x SD markers, the resulting 18 linkage groups included data from 284 different primer pairs. The number of gSSR, eSSR and STS
markers was 306, 102, and 9 respectively. The male map spanned a total length of 1,462 cM, with the average length of 81 cM and an average of 23 polymorphic markers per LG. The number of polymorphic markers mapped per LG ranged from 16 to 32, with LG lengths ranging from 57 cM (VIIIb-m) to 109 cM (IXa-m). Marker density ranged from 2.6-4.2 with an average of 1.46 mapped markers per primer pair.

The distribution of the three types of markers among the linkage groups was random and no aggregation of genomic or EST markers were observed. The STS markers were mapped in seven LG of both the female and male maps. In the female map, STS markers were placed in LG IIIb-f, IVa-f, IVb-f, Va-f, IXa-f, and IXb-f, while in the male map they were placed in LG IIb-m, IIIb-m, IVb-m, Va-m, IXa-m, and IXb-m.

**Genome length and coverage**

As estimated from the information of the parental maps, total genome length of switchgrass was 1,776 cM and 1,596 cM for the female and male maps, respectively. Two methods of estimation were employed and consistent genome length estimates were obtained using both Fishman *et al.* (2001) and Chakravarti *et al.* (1991) methods. The proportion of the genome within 5 cM of a marker locus was estimated to be 92% for the female map and 93 % for the male map.

The average of two genome length estimates (1,686 cM) was taken as the expected genome length of switchgrass. Based on an estimated genome size of ~1600 Mbp for tetraploid switchgrass (Saski *et al.* 2011), the average recombination rate across all LG was approximated at 1.05 cM per Mbp.

**Map integration**

The linkage groups were integrated with one another based on shared SDxSD markers, or SD markers co-amplified by the same primer set that represented either homoeologous loci or
alternate alleles heterozygous in either both subgenomes or both parents respectively (Fig. 2). Accordingly, nine homology groups of four linkage groups each were formed. All groups were then assigned names consistent with previous nomenclature based on marker alleles shared in common with the linkage map of Okada et al. 2010 and synteny of selected eSSR markers sequences with the foxtail millet (Setaria italica) pseudomolecules (http://www.phytozome.net v7.0). The nine homology groups formed represent the base chromosome number for switchgrass.

**Transmission ratio distortion**

We found that 26% of the genotyped markers were affected by TRD, however only 17% of the markers included in the final maps were distorted. The rest could not be positioned with certainty. TRD markers that affected the AL-NF (PV281 x NF472) map were compared with the two related switchgrass pseudo-testcross populations that it was developed from: ALB (A4 x K5) and NF-UGA (AP13 x VS16). As compared to the ALB population where 8.7% of the markers were affected by TRD, the AL-NF and NF-UGA populations showed a higher frequency of TRD markers (Table 5) with the highest rates (25.1 %) found in the NF-UGA population. A total of 143 TRD markers were recorded in the AL-NF population. Of these, 69 segregated in the female and 70 in male parent while 4 were distorted SD x SD markers that segregated in both parents. In the male map 44% of the SD TRD markers were scored as present in excess while in the female map the percentage of SD TRD markers present in excess was 54%.

Distorter regions consisting of clusters of three or more consecutive TRD markers were observed on LG IIa-f, Vb-f, VIIIa-f, IXa-f, IXb-f, and VIIIa-f in the PV281 map indicating a biological phenomenon as opposed to genotyping error. Similarly, in the NF472 map LG Ia-m, Ib-m, IIIb-m, and VIIb-m were affected by distorter regions that are indicated in Figure 2. In three cases genotypic information from the founders in these distorter regions
allowed determination of which haplotype was present in excess. Regions affected by TRD in LG VIIb-m and IIIb-m were skewed toward the AP13 haplotype, while the TRD region on Ia-m was skewed toward the VS16 haplotype. Significant interactions between loci on corresponding LG in the male and female maps in groups IXa-m/f and VIa-m/f between NFSG012-268 on LG IXa-m and the three adjacent markers NFSG238-195_ds, NFSG238-146, and NFSG328-142_ds on LG IXa-f as well as between NFSG007-357 on LG VIa-m and NFSG203-256_ds on LG VIa-f were observed. These interactions are indicated in Figure 2. There were no significant interactions between TRD markers and markers on different homology groups. Comparison of TRD with the related F1 population of Okada et al. demonstrated that LG Ia-m and Ib-m were affected by distortion in both maps as was LG VIIb-m.

Map Comparisons

eSSR, gSSR, and STS primer pairs used in several other switchgrass linkage studies with the ALB population (Okada et al. 2010), the NF population (Serba et al. 2013) and the unrelated NL94 population (Liu et al. 2012) were compared using a string order comparison based on the longest common subsequence (LCS) (Agarwala et al. 2000). This method identified the largest number of sequences or primer sets that had the same relative order between two different linkage groups. The LCS numbers were summed across all LG by subgenome and male or female linkage maps where appropriate and then expressed as a fraction of the total number of SSR shared in common. These results are presented in Table 6. In 75% of the cases it was not possible to compare a and b sub-genomes due to too few SSR shared in common. The percent of SSR in the LCS ranged from 47-77% while the number of SSR mapped in common ranged from 3 to 103. Within the AL-NF population the LCS ranged between 0.74 and 0.57 with a maximum of 159 and a minimum of 24 SSR shared in common between subgenomes and male or female linkage maps. The correlation coefficient
between AL-NF marker distances intervals was 0.77 for the NF population and 0.40 for the NF94 population.

Similar comparisons to the foxtail millet genome sequence were performed based on the top scoring blast hits of EST-SSR sequences to the foxtail millet genome sequence. The e-values of these results were less than $1 \times 10^{-40}$ except for sww2209 which was $2 \times 10^{-12}$. Comparisons were limited to EST-SSR markers and thus there were fewer markers present in the LCS, which contained between 76% and 91% of the total number (Table 6).

**DISCUSSION**

**Population development**

The mapping population was developed in three generations using four switchgrass genotypes selected from four founders. Two of the genotypes, A4 and AP13 were selected from Alamo. The other genotypes, K5 and VS16 were selected from Kanlow and Summer, respectively. Alamo and Kanlow are lowland cultivars that have robust growth and high biomass yield potential, while Summer is a northern-adapted upland cultivar with limited biomass yield but cold tolerance.

These populations combine the variability from different founder populations and will be useful for detecting QTL for yield and other relevant traits. Knowledge of linkage phase is critical for QTL mapping and other applications. In an outcrossing pedigree the linkage phase of the parents is not known. Genotyping the grandparents allows determination of the linkage phase of the parents using identity by descent and yields improved results where there is limited marker coverage over inferring phase based on independence LOD scores of marker pairs.

**Selfed Individuals**
The presence of 74 self-pollinated individuals among the genotyped population required a slightly modified mapping approach, but also provides an opportunity to compare traits potentially impacted by inbreeding depression in related populations with different degrees of heterozygosity. The true rate of outcrossing in this population was not directly determined as differences in fitness at early life stages biased establishment and genotyping of the population. However, those individuals that were genotyped consisted of 59% hybrid and 41% selfed individuals. This indicates that the female parent of the cross (PV281) was effectively self-compatible. There have been several studies examining reproductive systems in switchgrass. These have found that both pre- and post-zygotic barriers to selfing are active, and that bagged inflorescences produce far fewer seed than open pollinated inflorescences (Martinez-Reyna and Vogel, 2002; Talbert et al. 1983). Estimates of 0.35% self-compatibility were obtained from 17 upland and lowland tetraploid genotypes by Martinez-Reyna and Vogel, 2002.

The discovery of conditionally-compatible genotypes NL94 LYE 16×13 and SL93 7×15 that are self-compatible under controlled conditions but effectively self-incompatible in the field (Liu et al., 2013) as well as the high self-compatibility found in PV281 demonstrates the existence of both reproductive modes in switchgrass. However, very little information is available on the genetic basis of incompatibility in switchgrass or other polyploid species. Loss of self-incompatibility due to modifiers of major SI loci, non-functional incompatibility alleles, or decreased competition from unrelated pollen at the stigma surface or stylar transmitting tissue under controlled conditions may play roles in conditional self-fertility in switchgrass.

**Marker diversity and segregation**

The marker systems utilized in this study were SSR and STS markers. The SSR markers were developed from both genomic and EST sequences. The gSSR markers were developed from
(GA/CT)$_n$ enriched libraries. The eSSR markers were developed from publicly available switchgrass EST sequences. SSR markers have become quite useful in various aspects of molecular genetic studies including assessment of genetic diversity (Amsellem et al. 2001; Ashley et al. 2003), ecological-genetic studies (Muraya et al. 2011), marker-assisted selection (Fazio et al. 2003), and genetic linkage mapping (Akkaya et al. 1995; Broun and Tanksley, 1996). SSR have an additional evolutionary role in creating and maintaining quantitative genetic variation (Kashi et al. 1997).

Both gSSR and eSSR markers were used in the construction of this linkage map to cover genic as well as the intergenic regions of switchgrass. The gSSR markers proved to be evenly distributed throughout the genome. Sequence-tagged markers such as eSSR and STS markers have special importance in covering the euchromatic regions of genome (Mun et al. 2006; Ren et al. 2009) and reveal variation in transcribed genes among individuals (Ramchiary et al. 2011). A total of 136 and 130 eSSR markers were placed on the female and male maps, respectively. In addition, 12 and 9 STS markers were also mapped in the female and male maps, respectively. These STS markers were developed from cellulase and transferase gene sequences and have a special importance in mapping the genome region and genes involved in cellulose accumulation in switchgrass. It will be interesting to see cosegregation of any of these markers with QTL associated with biomass quality traits.

Switchgrass is an outcrossing plant, therefore the parental genotypes are heterozygous. In a cross between such heterozygous parents, many single-dose polymorphic markers will amplify in one parent, be absent in the other and therefore are expected to segregate in a 1:1 ratio in their progeny (Grattapaglia and Sederoff, 1994). Other markers are heterozygous in both the parents and expected to segregate 3:1 under disomic inheritance, or 5:1 under tetrasomic inheritance models. The great majority of the SSR and STS markers assayed in this population were found to follow the expected 1:1 ratio for the SD markers and
3:1 ratio for the SD x SD cross. This agrees with previous findings that switchgrass inheritance is disomic.

**Linkage map and genome coverage**

The maps lengths produced by this study were similar to those reported in the NF population in which the female linkage map was 13% larger than the male map. In this study the female linkage map was 11% larger than the male’s. Similarly, the overall map lengths were comparable with this genetic linkage map being 90% of the length of that reported by Serba *et al.* 2013. These map lengths were both significantly smaller than that reported for the NL94 population. The relative lengths of the male and female linkage maps in this study were the opposite of the ALB population where the male map was found to be relatively shorter. These results suggest that meiosis in the NF and AL-NF populations was behaving in a consistent manner, and that inclusion of increasing percentages of TRD markers in the individual switchgrass linkage studies did not appear to greatly inflate map length, consistent with theory (Hackett and Broadfoot, 2003).

**Transmission ratio distortion**

Distorter loci affecting the transmission of closely linked markers are commonly present in mapping populations and can be linked to pre- or postzygotic mechanisms including pollen tube competition (Mangelsdorf and Jones 1926), pollen lethals (Rick 1966), preferential fertilization (Gadish and Zamir 1987), and selective elimination of zygotes (Rick, 1963). Selfish genetic elements are also be selected for in the asymmetric female meiosis of most plants and animals as meiotic drive. In the AL-NF population, TRD reached 17.1% in mapped markers. These rates of distortion are similar to that observed in the NF-GA mapping population (Serba *et al.* 2013). However, Okada *et al.* (2010) observed a SD of only 9% for the single dose markers of genomic and eSSR markers. This lower SD in the ALB population is attributed to the omission of highly distorted markers from the attempted mapping rather
than significant biological differences. The distorted markers were distributed throughout the maps. There were several clusters of three or more loci in both the female and male maps. In the female map TRD regions were observed on six LG while TRD regions were observed on 11 LG in the male map. Significant interactions between markers in male and female LG VI and IX in the AL-NF population indicates possible post-zygotic interactions. The presence of TRD loci may lead to bias in estimation of the recombination fraction when two or more TRD loci is present on a single chromosome (Lorieux et al. 1995a; Lorieux et al. 1995b). However in our study we did not detect this situation, leading us to believe that our maps were not biased due to TRD.

Conclusions
A three-generation and four-founder mapping population was developed in tetraploid switchgrass. We have constructed genetic linkage map using a combination of SSR and STS markers, with marker density close to saturation. The LG were arranged in nine homoeologs corresponding to the haploid chromosome number of tetraploid switchgrass. Clustering of TRD markers was observed in some genomic regions. The use of parental maps to identify QTLs underlying agronomic traits for MAS is in progress. Due to the long establishment time of switchgrass and the need for multiple year yield data, this approach has high potential to speed up selection.

ACKNOWLEDGEMENTS
This research was supported with funds provided by the Department of Energy and United States Department of Agriculture. Our sincere appreciation goes to Yuhong Tang and her group for genotyping and the greenhouse group of the Noble Foundation for greenhouse maintenance of the population. We are very thankful to Brian Motes and his group members for taking care of field experiments. The US Department of Agriculture, Agricultural
Research Service, is an equal opportunity/affirmative action employer and all agency services are available without discrimination. Mention of commercial products and organizations in this manuscript is solely to provide specific information. It does not constitute endorsement by USDA-ARS over other products and organizations not mentioned.

**LITERATURE CITED**


Akkaya M. S., Shoemaker R. C., Specht J. E., Bhagwat A. A., P. B. Cregan, 1995 Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci. 35: 1439.


Broun P., S. D. Tanksley, 1996 Characterization and genetic mapping of simple repeat

Bouton, J., 2008 Improvement of switchgrass as a bioenergy crop, in: W. Vermerris (Ed.),
Genetic improvement of bioenergy crops, Springer Science+Business Media, LLC, New
York.

Brummer, E. C., J. H. Bouton, and G. Kochert, 1993 Development of an RFLP map in

Casler, M. D., 2005 Ecotypic variation among switchgrass populations from the northern
USA. Crop Sci. 45: 388-398.

Casler, M. D., K. P. Vogel, C. M. Taliaferro, and R. L. Wynia, 2004 Latitudinal adaptation of
switchgrass populations. Crop Sci. 44: 293-303.


Evanylo, G. K., A. O. Abaye, C. Dundas, C. E. Zipper, R. Lemus, et al., 2005 Herbaceous
vegetation productivity, persistence, and metals uptake on a biosolids-amended mine soil.
J. Environ. Qual. 34: 1811-1819.

Fazio G., Staub J. E., M. R. Stevens, 2003 Genetic mapping and QTL analysis of
horticultural traits in cucumber (Cucumis sativus L.) using recombinant inbred lines

Fike, J. H., D. J. Parrish, D. D. Wolf, J. A. Balasko, J.T. Green Jr., et al., 2006 Long-term


Liu L., Thames S. L., Y. Wu, 2013 Lowland switchgrass plants in populations set completely outcrossed seeds under field conditions as assessed with SSR markers.


Voorrips, R. E., 2002 MapChart: Software for the graphical presentation of linkage maps and QTLs. J. Hered. 93: 77-78.


Table 1: Molecular marker categories used, their amplification and polymorphism in the AL-NF mapping population

<table>
<thead>
<tr>
<th>Marker categories</th>
<th>Total primers screened</th>
<th>Polymorphic primers</th>
<th>Average polymorphic markers</th>
<th>Polymorphism (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>eSSR</td>
<td>342</td>
<td>158</td>
<td>2.03</td>
<td>46.2</td>
</tr>
<tr>
<td>gSSR</td>
<td>512</td>
<td>272</td>
<td>2.97</td>
<td>53.1</td>
</tr>
<tr>
<td>STS</td>
<td>144</td>
<td>17</td>
<td>1.94</td>
<td>11.8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>998</strong></td>
<td><strong>447</strong></td>
<td><strong>2.69</strong></td>
<td><strong>44.8</strong></td>
</tr>
</tbody>
</table>
Table 2: Segregation of molecular markers genotyped in AL-NF and AL-NF-self populations

<table>
<thead>
<tr>
<th>Marker</th>
<th>Female (PV281) parent</th>
<th>Male (NF472) parent</th>
<th>Both parents</th>
</tr>
</thead>
<tbody>
<tr>
<td>eSSR</td>
<td>136</td>
<td>109</td>
<td>1(1)*</td>
</tr>
<tr>
<td>gSSR</td>
<td>336</td>
<td>270</td>
<td>7(5)</td>
</tr>
<tr>
<td>STS</td>
<td>12</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>484</td>
<td>390</td>
<td>8(6)</td>
</tr>
</tbody>
</table>

*Numbers in parentheses indicate number of markers fitting into both 3:1 and 5:1 segregation ratios.
<table>
<thead>
<tr>
<th>Mapping features</th>
<th>Female (PV281) map</th>
<th>Male (NF472) map</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of linkage groups</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Total map length (cM)</td>
<td>1,629</td>
<td>1,462</td>
</tr>
<tr>
<td>Average linkage group length (cM)</td>
<td>91</td>
<td>81</td>
</tr>
<tr>
<td>Total number of loci mapped</td>
<td>445</td>
<td>417</td>
</tr>
<tr>
<td>Average loci per LG (#)</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>SD markers mapped</td>
<td>293</td>
<td>315</td>
</tr>
<tr>
<td>SD x SD mapped</td>
<td>111</td>
<td>36</td>
</tr>
<tr>
<td>Unmapped SD markers</td>
<td>82</td>
<td>51</td>
</tr>
<tr>
<td>Estimated genome length (cM)</td>
<td>1,766</td>
<td>1,543</td>
</tr>
<tr>
<td>Percent genome coverage within 5 cM mapped loci</td>
<td>91.8</td>
<td>92.7</td>
</tr>
</tbody>
</table>
Table 4: Number of markers, linkage group length, and marker density comparisons between female (PV281) and male (NF472) parental linkage groups

<table>
<thead>
<tr>
<th></th>
<th>LG</th>
<th>mapped markers</th>
<th>Total Length (cM)</th>
<th>Average interlocus distance</th>
<th>LG</th>
<th>mapped polymorphic markers</th>
<th>Total Length (cM)</th>
<th>Average interlocus distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PV21</td>
<td></td>
<td></td>
<td></td>
<td>NF472</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ia-f</td>
<td>30</td>
<td>95</td>
<td>3.2</td>
<td></td>
<td>Ia-m</td>
<td>23</td>
<td>89</td>
<td>3.9</td>
</tr>
<tr>
<td>Ib-f</td>
<td>13</td>
<td>59</td>
<td>4.5</td>
<td></td>
<td>Ib-m</td>
<td>20</td>
<td>74</td>
<td>3.7</td>
</tr>
<tr>
<td>IIa-f</td>
<td>43</td>
<td>115</td>
<td>2.7</td>
<td></td>
<td>IIa-m</td>
<td>27</td>
<td>78</td>
<td>2.9</td>
</tr>
<tr>
<td>IIb-f</td>
<td>33</td>
<td>86</td>
<td>2.6</td>
<td></td>
<td>IIb-m</td>
<td>32</td>
<td>98</td>
<td>3.1</td>
</tr>
<tr>
<td>IIIa-f</td>
<td>25</td>
<td>99</td>
<td>4.0</td>
<td></td>
<td>IIIa-m</td>
<td>23</td>
<td>78</td>
<td>3.4</td>
</tr>
<tr>
<td>IIIb-f</td>
<td>24</td>
<td>117</td>
<td>4.9</td>
<td></td>
<td>IIIb-m</td>
<td>23</td>
<td>107</td>
<td>4.6</td>
</tr>
<tr>
<td>IVa-f</td>
<td>20</td>
<td>86</td>
<td>4.3</td>
<td></td>
<td>IVa-m</td>
<td>21</td>
<td>83</td>
<td>4.0</td>
</tr>
<tr>
<td>IVb-f</td>
<td>21</td>
<td>89</td>
<td>4.2</td>
<td></td>
<td>IVb-m</td>
<td>19</td>
<td>77</td>
<td>4.1</td>
</tr>
<tr>
<td>IXa-f</td>
<td>24</td>
<td>74</td>
<td>3.1</td>
<td></td>
<td>IXa-m</td>
<td>31</td>
<td>109</td>
<td>3.5</td>
</tr>
<tr>
<td>IXb-f</td>
<td>37</td>
<td>107</td>
<td>2.9</td>
<td></td>
<td>IXb-m</td>
<td>26</td>
<td>73</td>
<td>2.8</td>
</tr>
<tr>
<td>Va-f</td>
<td>34</td>
<td>115</td>
<td>3.4</td>
<td></td>
<td>Va-m</td>
<td>23</td>
<td>83</td>
<td>3.6</td>
</tr>
<tr>
<td>Vb-f</td>
<td>27</td>
<td>131</td>
<td>4.8</td>
<td></td>
<td>Vb-m</td>
<td>26</td>
<td>93</td>
<td>3.6</td>
</tr>
<tr>
<td>VIa-f</td>
<td>20</td>
<td>65</td>
<td>3.2</td>
<td></td>
<td>VIa-m</td>
<td>24</td>
<td>83</td>
<td>3.5</td>
</tr>
<tr>
<td>VIb-f</td>
<td>20</td>
<td>74</td>
<td>3.7</td>
<td></td>
<td>VIb-m</td>
<td>18</td>
<td>62</td>
<td>3.4</td>
</tr>
<tr>
<td>VIIa-f</td>
<td>17</td>
<td>114</td>
<td>6.7</td>
<td></td>
<td>VIIa-m</td>
<td>16</td>
<td>65</td>
<td>4.1</td>
</tr>
<tr>
<td>VIIb-f</td>
<td>20</td>
<td>57</td>
<td>2.9</td>
<td></td>
<td>VIIb-m</td>
<td>24</td>
<td>74</td>
<td>3.1</td>
</tr>
<tr>
<td>VIIIa-f</td>
<td>13</td>
<td>73</td>
<td>5.6</td>
<td></td>
<td>VIIIa-m</td>
<td>19</td>
<td>80</td>
<td>4.2</td>
</tr>
<tr>
<td>VIIIb-f</td>
<td>24</td>
<td>72</td>
<td>3.0</td>
<td></td>
<td>VIIIb-m</td>
<td>22</td>
<td>57</td>
<td>2.6</td>
</tr>
<tr>
<td>Average</td>
<td>25</td>
<td>91</td>
<td>3.8</td>
<td></td>
<td>23</td>
<td>81</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>445</td>
<td>1629</td>
<td></td>
<td></td>
<td>417</td>
<td>1462</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5: Number and frequency of molecular markers with transmission ratio distortion in three switchgrass pseudo-testcross mapping populations

<table>
<thead>
<tr>
<th>Population</th>
<th>Type</th>
<th>Total</th>
<th>Distorted (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALB</td>
<td>gSSR</td>
<td>509</td>
<td>40 (7.86)</td>
</tr>
<tr>
<td></td>
<td>STS</td>
<td>55</td>
<td>4 (7.3)</td>
</tr>
<tr>
<td></td>
<td>eSSR</td>
<td>945</td>
<td>88 (9.31)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>1509</strong></td>
<td><strong>132 (8.7)</strong></td>
</tr>
<tr>
<td>NF</td>
<td>gSSR</td>
<td>874</td>
<td>222 (25.4)</td>
</tr>
<tr>
<td></td>
<td>STS</td>
<td>36</td>
<td>11 (30.6)</td>
</tr>
<tr>
<td></td>
<td>eSSR</td>
<td>168</td>
<td>38 (22.6)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>1078</strong></td>
<td><strong>271 (25.1)</strong></td>
</tr>
<tr>
<td>AL-NF</td>
<td>gSSR</td>
<td>602</td>
<td>94 (15.6)</td>
</tr>
<tr>
<td></td>
<td>STS</td>
<td>19</td>
<td>6 (31.6)</td>
</tr>
<tr>
<td></td>
<td>eSSR</td>
<td>214</td>
<td>43 (20.1)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>835</strong></td>
<td><strong>143 (17.1)</strong></td>
</tr>
</tbody>
</table>
Table 6. LCS map comparisons across all linkage groups<sup>a</sup>.

<table>
<thead>
<tr>
<th></th>
<th>AL-NFa-f</th>
<th>AL-NFa-m</th>
<th>AL-NFb-f</th>
<th>AL-NFb-m</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALBa-f&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.68 (98)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.72 (86)</td>
<td>-- (17)</td>
<td>0.59 (27)</td>
</tr>
<tr>
<td>ALBa-m</td>
<td>0.77 (96)</td>
<td>0.68 (89)</td>
<td>0.69 (23)</td>
<td>0.68 (25)</td>
</tr>
<tr>
<td>ALBb-f</td>
<td>-- (17)</td>
<td>0.53 (32)</td>
<td>0.58 (103)</td>
<td>0.58 (100)</td>
</tr>
<tr>
<td>ALBb-m</td>
<td>-- (15)</td>
<td>-- (16)</td>
<td>0.67 (94)</td>
<td>0.67 (90)</td>
</tr>
<tr>
<td>NFa-f</td>
<td>0.53 (99)</td>
<td>0.60 (88)</td>
<td>-- (10)</td>
<td>-- (13)</td>
</tr>
<tr>
<td>NFa-m</td>
<td>0.57 (75)</td>
<td>0.6 (55)</td>
<td>-- (13)</td>
<td>-- (10)</td>
</tr>
<tr>
<td>NFb-f</td>
<td>-- (11)</td>
<td>-- (7)</td>
<td>0.49 (81)</td>
<td>0.60 (84)</td>
</tr>
<tr>
<td>NFb-m</td>
<td>-- (18)</td>
<td>0.71 (21)</td>
<td>0.48 (77)</td>
<td>0.59 (83)</td>
</tr>
<tr>
<td>NL94a</td>
<td>0.52 (36)</td>
<td>0.56 (39)</td>
<td>-- (3)</td>
<td>-- (5)</td>
</tr>
<tr>
<td>NL94b</td>
<td>-- (16)</td>
<td>-- (13)</td>
<td>0.47 (46)</td>
<td>0.6 (35)</td>
</tr>
<tr>
<td>S. italica</td>
<td>0.88 (18)</td>
<td>0.84 (19)</td>
<td>0.76 (21)</td>
<td>0.91 (23)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Linkage mapping studies on populations ALB (Okada <i>et al.</i> 2010), NF (Serba <i>et al.</i> 2013, and NL94 (Liu <i>et al.</i> 2012) were compared with the AL-NF population using the LCS method and SSR primer sets employed in multiple studies. The <i>Setaria italica</i> genome was compared to the AL-NF population using the molecular coordinates on the scaffold sequences of the best blast hits to EST-SSR.

<sup>b</sup>Italicized letters (<i>a</i>) and (<i>b</i>) refer to individual subgenomes, while (<i>m</i>) and (<i>f</i>) refer to male or female linkage maps.

<sup>c</sup>The first number represents the fraction of SSR compared that were present in the LCS while the number in parentheses is the number of SSR shared in common between the two groups. LCS of groups with 18 or fewer markers in common were not reported.
Figure 1: Switchgrass AL-NF mapping population development scheme using four founders in three generations. Other crosses were performed, however the PV281 x NF472 cross resulted in sufficient seed for further evaluation. AL-NF-self population represents progeny derived from self pollination of PV281 that were initially identified by limited genotyping. AL-NF population represents hybrid individuals. The number \( n \) of individuals in each population is indicated.
Figure 2. Consensus linkage map of PV281 and NF472. PV281 linkage groups were constructed by integrating genotypic data from both self-fertilized individuals (n=74) and PV281xNF472 hybrid individuals (n=108) (see materials and methods). The Roman numeral designation of each homology group (I–IX) follows the foxtail millet chromosomes. NF472 linkage groups were constructed based only on hybrid individuals as it served as the male donor in the cross. Individual linkage groups are indicated by vertical bars. Numbers to the left of the bar indicated position in centiMorgans while marker names are indicated to the right of the bar. Markers ending in “ds” are double simplex markers present in both parents. Markers used to align the subgenomes were amplicons of different fragment length detected with the same primer set. These markers are set off by similar color labeling. The letters (f) and (m) at the end of the linkage group name denote female (PV281) and male (NF472) parental map, respectively. Severely distorted markers are noted with asterisks: (*) P < 0.01, (**) P < 0.001, and (***) P < 0.0001. Shaded regions of the individual linkage groups indicate the presence of three or more consecutive markers with TRD. Markers with significant interactions on different linkage groups are indicated by the pound sign (#).