Comparative Relationships and Chromosome Evolution in Switchgrass (*Panicum virgatum*) and Its Genomic Model, Foxtail Millet (*Setaria italica*)

**Guillaume Daverdin · Bochra A. Bahri · Xiaomei Wu ·
Desalegn D. Serba · Christian Tobias · Malay C. Saha ·
Katrien M. Devos**

Abstract A switchgrass (*Panicum virgatum* L.) consensus map was developed that combined data from two mapping populations and integrated recombination data from both parents of this largely obligate outcrossing species. The consensus map consisted of 1,321 loci and spanned 2,122 cM. An analysis of the distribution of genic and genomic markers across the length of a linkage group showed that genic markers were relatively overrepresented in distal regions, while genomic markers were relatively overrepresented in pericentromeric regions. Furthermore, genic markers located in distal and pericentromeric regions identified orthologs in the genome of the closely related species foxtail millet (*Setaria italica* (L.) P. Beauv.) at a similar rate, but fewer orthologs were found for genomic markers that mapped to pericentromeric regions compared to distal regions. Because cross-species sequence conservation is largely limited to genes, our data suggest that genomic markers that map to distal regions are more likely to be derived from genic regions than those that map to pericentromeric regions. A comparative analysis between loci on the switchgrass consensus map and their orthologs in the foxtail millet genome showed that the two species varied by at least nine inversions and one reciprocal translocation. Extending the comparative analysis to sorghum demonstrated that the majority of the rearrangements occurred in the foxtail millet genome in the past 13 million years. The reduction in chromosome number from 10 in the Andropogoneae to 9 in the Paniceae was achieved through a complex set of rearrangements involving three ancestral chromosomes orthologous to sorghum chromosomes 6, 8, and 9 and chromosomes III and VII in the lineage leading to switchgrass and foxtail millet. Upon insertion of ancestral chromosome 9 into ancestral chromosome 8, the centromere of the recipient chromosome was lost, and this was accompanied by a loss of repetitive DNA. The switchgrass–foxtail millet comparative map provides a guide of the rearrangements that need to be taken into account when using foxtail millet as a surrogate for switchgrass in genetic analyses.

G. Daverdin · B. A. Bahri · X. Wu · K. M. Devos (✉)
Institute of Plant Breeding, Genetics and Genomics, and Department of Plant Biology, University of Georgia, Athens, GA 30602, USA
e-mail: kdevos@uga.edu

B. A. Bahri
Department of Plant Protection and Postharvest Diseases, National Agronomical Institute of Tunis, Tunis 1082, Tunisia

D. D. Serba · M. C. Saha
Samuel Roberts Noble Foundation Inc., Ardmore, OK 73401, USA

C. Tobias
Genomics and Gene Discovery Research Unit, Western Regional Research Center, Albany, CA 94710, USA

Present Address:
X. Wu
College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 310036, China

Keywords Comparative genetics · Switchgrass · *Panicum virgatum* · Foxtail millet · *Setaria italica* · Genome evolution · Consensus map

Introduction

Driven by the need for greater energy security and a reduced impact of fuel consumption on the environment, there has been a growing interest in recent years to produce biofuels from cellulosic biomass. Switchgrass (*Panicum virgatum* L.), a perennial warm season C4 grass native to North America, is

one of the crops that was identified in the 1990s as a potential feedstock for the production of cellulosic biofuels because of its high biomass yield, low nutrient and water requirements, and adaptation to marginal lands [1–3]. Switchgrass belongs to the family *Poaceae*, subfamily Panicoideae, tribe Paniceae and comprises two ecotypes, upland and lowland, that largely vary in their ploidy level, zone of adaptation, and plant architecture. Switchgrass has extensive genetic diversity and a large potential for genetic improvement [1]. The improvement of switchgrass for traits important for bioenergy production, including total biomass yield and biomass composition, is key to making switchgrass an integral part of a successful cellulosic biofuel production program.

To enhance the breeding process, genetic maps have been developed for switchgrass [4–8], and its genome has been sequenced (www.phytozome.net). However, genetic analyses have been hampered by the complexity of the switchgrass genome. In its lowest ploidy form, switchgrass is an allotetraploid ($2n=4x=36$) and has a haploid DNA content of 1.35 pg (1,300 Mb) [9]. It is also largely an obligate outcrossing species, and consequently, switchgrass genotypes are heterozygous at most loci. Most genetic maps published to date have been developed in pseudo-F1 populations resulting from crossing two heterozygous switchgrass genotypes [4–6, 8]. The exception is the genetic map generated by Liu et al. [7] who used an F2 population obtained by selfing a heterozygous self-compatible accession. When analyzing marker segregation in pseudo-F1 progeny, in essence, the amount of recombination that took place in each of the parental genomes is measured. The output is two linkage maps per chromosome, one that represents recombination in the female (egg) parent and one that represents recombination in the male (pollen) parent. While algorithms have been written to integrate parental maps (e.g., JoinMap [10] and OneMap [11]), such maps have not yet been produced in switchgrass.

An additional layer of integration can be achieved by combining maps that have been generated in different mapping populations but comprise common markers into a consensus map (e.g., maize- Falque et al. [12]; sugar cane- Ming et al. [13]; pearl millet- Qi et al. [14], Rajaram et al. [15]; rye-Gustafson et al. [16]; wheat- Marone et al. [17]). Integrated maps compensate for inconsistencies in individual maps, provide better genome coverage, allow integration of QTL data, and increase marker densities for practical map applications such as marker-assisted selection, map-based cloning, and cross-species comparative analyses.

The genome of the lowland switchgrass genotype AP13 has been sequenced to 15× using mainly the Roche 454 platform (www.phytozome.net). Sequence assembly, however, remains a challenge due to the heterozygosity and tetraploidy of AP13 and the repeat content (≥33 % [18]) of the switchgrass genome. Sequence contigs can be ordered by anchoring them to a high-density genetic map. However,

marker densities on the individual classically constructed switchgrass genetic maps are relatively low (average of 0.32 markers/cM) [5, 6]. Recently, the Joint Genome Institute (JGI) has sequenced the progeny of the AP13×VS16 mapping population [6] and used this data combined with information from the assembly of the genome sequence of *Panicum hallii*, a close relative of *P. virgatum*, to order the switchgrass sequence contigs (assembly v1.1; www.phytozome.net). A consensus map integrating three published genetics maps could be used as an independent tool to validate the sequence-based map. Lu et al. [8] used genotyping by sequencing (GBS) to generate a high-density map containing an average of 2,450 markers per linkage group. However, marker ordering in this map was achieved by alignment against the foxtail millet (*Setaria italica*) genome and thus does not take into account potential structural differences between the foxtail millet and switchgrass genomes.

Foxtail millet is a perennial C4 grass which, as switchgrass, belongs to the *Poaceae* family, subfamily Panicoideae, tribe Paniceae. The two species diverged around 13.1 million years ago (MYA) [19]. Because of its close phylogenetic relationship with biofuel grasses and relatively simple genome (C=500 Mb, $2n=2x=18$), the foxtail millet genome underwent whole-genome Sanger shotgun sequencing in January 2010 so that it could be used as a model for switchgrass genome analyses [20]. The *Setaria* genome was assembled (http://www.phytozome.net/) into 336 scaffolds and anchored to a 992 loci genetic map [20]. Comparative analyses with sorghum and rice identified at least seven chromosomal inversions in addition to a complex set of rearrangements involving three ancestral chromosomes that occurred in *Setaria* or in the lineage leading to *Setaria* since its divergence from a common ancestor with sorghum approximately 26 MYA. Although the foxtail millet genome has been used as a "surrogate" for switchgrass [8], the structural relationship between the two genomes has not yet been established.

The objectives of this research were (1) to integrate several of the published switchgrass genetic maps into a consensus map with an increased marker density that can be used to validate the sequence-based genetic map and in practical breeding applications and (2) to employ the consensus map in a comparative analysis with the foxtail millet genome sequence.

Material and Methods

Mapping Populations and Markers

The dataset used to build the consensus maps combined segregation data from Missaoui et al. [4], Okada et al. [5], and Serba et al. [6]. Serba and colleagues analyzed 191 F1 plants from a cross between the lowland genotype "AP13" (female parent), derived from cv. "Alamo," and the upland

genotype "VS16" (male parent), derived from cv. "Summer." Sixty-two of the progeny were also analyzed by Missaoui et al. [4]. The Okada et al. [5] mapping population consisted of 238 F1 plants obtained from a cross between selected genotypes from the lowland cvs. "Kanlow" (female parent) and Alamo (male parent). Because markers that are heterozygous in the female parent are typically analyzed separately from those that are heterozygous in the male parent in F1 populations, there were two datasets available for each of the mapping populations. For the purpose of generating the consensus maps, female and male datasets were considered separate populations. The Missaoui dataset comprised 110 loci that segregated in the female parent and 102 that segregated in the male parent, the Okada dataset consisted of 563 female and 542 male loci, and the Serba dataset consisted of 515 female and 363 male loci. Also included in the Serba datasets are 32 loci (18 in the female and 14 in the male) detected by 19 primer pairs developed against genes that were of interest to our group or to collaborating groups and that were newly mapped in the AP13×VS16 population. Primer sequences for these newly mapped loci are listed in Online Resource 1. The markers were amplified in a total volume of 20 μl consisting of 50 ng genomic DNA, 0.4 μM of forward and reverse primers, 0.8 U GoTaq DNA Polymerase (Promega, Madison, WI), 2.5 mM MgCl$_2$, and 0.2 mM dNTPs in 1× buffer. PCR conditions consisted of an initial denaturation of 3 min at 94 °C followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at temperatures listed in Online Resource 1 for 30 s, and extension at 72 °C for 1 min. The final extension was held at 72 °C for 15 min after which the samples were cooled to 4 °C. PCR products were separated on 0.5 % Mutation Detection Enhancement (MDE) gels (Lonza, Rockland, ME, USA) to reveal single-strand conformation polymorphisms that were visualized by silver staining. Segregation patterns were scored manually.

Identification of Putative Foxtail Millet Orthologs for the Switchgrass Markers

The nucleotide sequences of the markers included in the Okada et al. [5] and Serba et al. [6] maps were used as queries in a BLASTn search against the *P. virgatum* genome (www. phytozome.net). For each search, the region that aligned best to the query sequence (E value<1e-30) plus, when available, 1 kb of sequence on either side of the alignment was then used as a query in a BLASTn search against the foxtail millet genome (nonmasked; www.phytozome.net). Top hits (E value<1e-10) were considered as being putatively orthologous to the mapped *P. virgatum* markers. For the RFLP markers included in the Missaoui et al. [4] map, sequence information was available only for the RZ loci, which were detected by rice cDNA clones [21]. RZ clone end sequence was downloaded from GenBank, and the corresponding full-length gene sequence was obtained through

a BLASTn search of the end sequences against the rice genome sequence present in GenBank. The sequence of the full-length genes was then used in a BLASTn search against the foxtail millet genome, and top hits (E value<1e-10) were retained as putatively orthologous to the RZ markers.

Generation of a Consensus Map

In first instance, we used Mapmaker v.3.0 [22] to identify RFLP markers from the partial map by Missaoui et al. [4] that could be integrated into the Serba et al. [6] maps based on the sixty-two common progeny. RFLP markers that did not link at a logarithm (base 10) of odds (LOD) score of 2.0 using the "group" command, or that mapped to the ends of linkage groups using the "try" command, were removed from the dataset. The Serba female and male datasets were then merged with the Okada female and male datasets to represent one single large mapping population consisting of 858 plants. Markers were assigned genotypic scores in the populations in which they had been mapped and were given missing data scores for the other population(s). A separate data file was made for each linkage group (Online Resource 2).

Consensus maps were generated using a multistep process. First, we generated a framework consensus map for each switchgrass linkage group using only markers that mapped to a syntenic foxtail millet chromosome. We used JoinMap 4.0 [10] with the population type set as backcross and a regression mapping algorithm at a pairwise recombination frequency estimate <0.40, a LOD score ≥3, a goodness-of-fit jump threshold of 5, and a ripple value of 1 to determine marker orders. Although JoinMap assigned a specific location to each marker, most markers could only be placed within a bin because they were segregating in only a subset of the four populations. We therefore manually scrutinized recombination events and used comparative information from foxtail millet to reorder markers. The assumption was that marker orders were conserved between foxtail millet and switchgrass unless we had evidence from recombination data that this was not the case. We then used the try command of Mapmaker v. 3.0 to place markers that did not map to syntenic locations in foxtail millet and markers for which no putative foxtail millet ortholog had been identified relative to the framework map. Because Mapmaker considers markers that are scored in different phases as unlinked, use of Mapmaker required conversion of all marker scores to the same phase. The conversion was done in Excel using phase information provided by the JoinMap 4.0 data tabsheet obtained for each LG group node output. Map distances were calculated using Mapmaker, and the Kosambi mapping function was used to convert recombination units into genetic distances. Finally, we manually verified recombination events to indicate the range of locations that markers without comparative information could occupy on the map. Linkage group designations are based on the

names given in Okada et al. [5] and Serba et al. [6] and do not reflect a specific subgenome allocation.

Comparative Analyses

Relationships between marker positions in the switchgrass consensus map and their corresponding positions in foxtail millet were visualized in a circle diagram using the Circos software [23]. Comparative dot plots were drawn using the "R programming language" [24]. Putative centromere locations in switchgrass were inferred from the centromere locations in colinear foxtail millet chromosomes. A rearrangement between the foxtail millet genome and the switchgrass genetic map was defined as a minimum of three adjacent markers being present in inverted orientation (inversion) or on a nonsyntenic chromosome (translocation). To investigate whether overall marker densities and the relative number of colinear markers varied along the length of a switchgrass linkage group, we divided each linkage group in five regions. The "pericentromeric" region is the region of a linkage group that corresponds to 33 % of the syntenic foxtail millet chromosome surrounding each centromere. The remainder of each arm is then split into two so that the two regions in each arm have corresponding regions in foxtail millet that are equal in length. Since short and long arms have not been assigned in either switchgrass or foxtail millet, data from the distal regions on each arm were combined into a single distal region for further analysis. Similarly, data from the two proximal regions were combined into a single proximal region. Chi-square tests were conducted to identify significant differences in marker distributions. Information on comparative relationships between the foxtail millet, sorghum, and rice genomes was from Bennetzen et al. [20].

Fig. 1 Venn diagram showing the number of markers that are common between the different mapping populations. *Letters and letter combinations* indicate the population(s) with *P* AP13 (female) [6], *S* VS16 (male) [6], *K* Kanlow (female) [5], *A* Alamo (male) [5]. For each population or combination of populations, the total number of markers and, *in brackets*, the number of markers included in the consensus map framework is given

Results

Consensus Map

A total of 1,459 loci from three different maps were integrated into the consensus map. Fifty RFLP loci originated from Missaoui et al. [4], 599 simple sequence repeat (SSR) loci were from Okada et al. [5], 778 loci including 336 Diversity Array Technology (DArT) loci, 439 SSR loci, and 3 sequenced-tag-sites markers developed from expressed sequenced tags (EST-STSs) were from Serba et al. [6], and 32 loci were newly mapped in the AP13×VS16 mapping population. Of the 1,459 loci, 314 were shared by at least two of the four mapping populations (K Okada female and A Okada male, or P Serba female and S Serba male) that constituted the mega mapping dataset, and of these, 138 were shared between the Okada and Serba datasets (Fig. 1), bringing the number of unique integrated loci to 1,321. The 1,321 loci were organized in 9 sets of homoeologous linkage groups (Fig. 2). The total length of the consensus map was 2,122 cM with individual linkage groups spanning from 70 to 179 cM (Table 1). The average distance between markers was 1.6 cM. Nevertheless, a total of 13 gaps larger than 10 cM were present on LG I-a (1 gap), LG II-a (2), LG II-b (2), LG III-a (3), LG V-a (1), LG VI-a (1), LG VII-a (1), LG VII-b (1), and LG VIII-a (1) (Fig. 2).

Comparative Analysis

Of the 1,321 loci that were integrated into the switchgrass consensus map, 1,215 loci (92.0 %) identified hits in the foxtail millet genome at the *E* value threshold of 1e-10. Of those, 70.9 % mapped to syntenic linkage groups/chromosomes in switchgrass and foxtail millet. Plotting of the switchgrass genetic map locations for these loci against the position of their putative orthologs
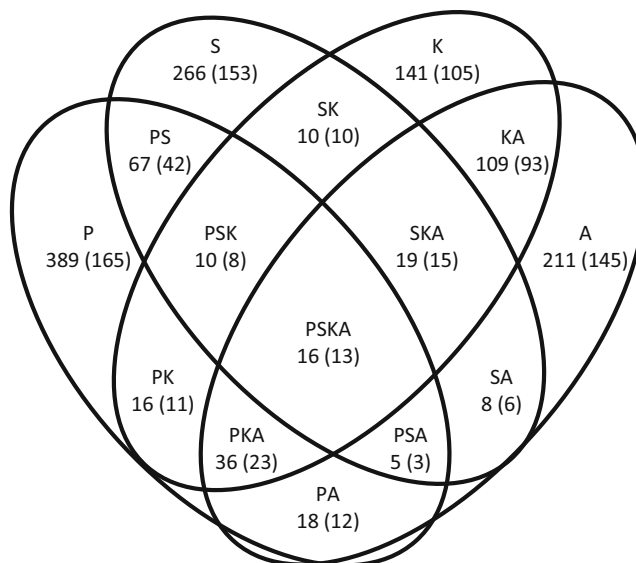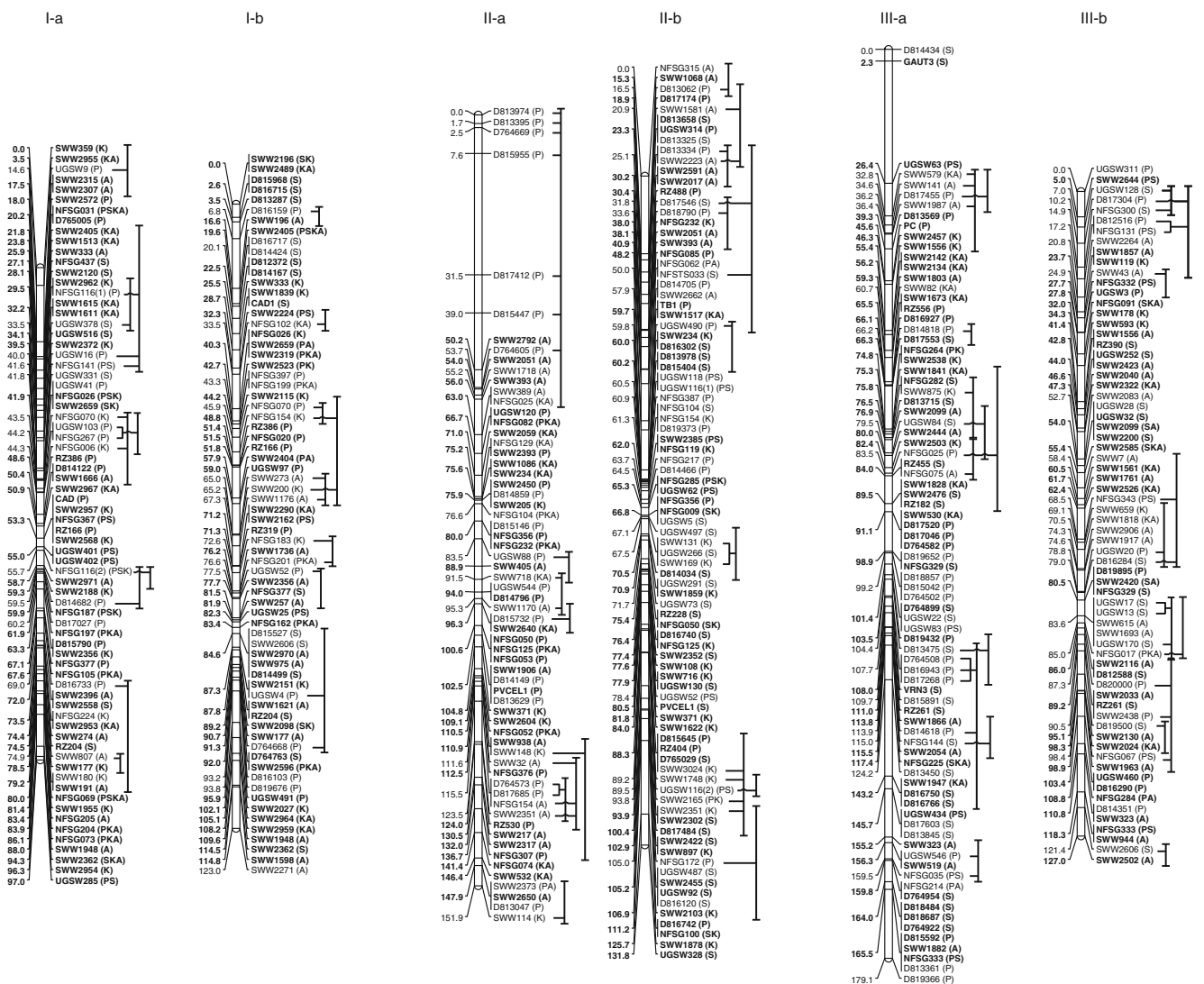
Fig. 2 Switchgrass consensus genetic maps. Loci with orthologs in foxtail millet are indicated in *bold*. The population(s) in which each marker was scored is indicated in *parentheses*: *(P)* AP13 (female) [6], *(S)* VS16 (male) [6], *(K)* Kanlow (female) [5], *(A)* Alamo (male) [5]. For markers that could not be mapped unambiguously, the range to which they can map is indicated by a *vertical line* on the *right-hand side* of linkage groups

on the foxtail millet genome sequence displayed a high level of colinearity between the switchgrass and foxtail millet genomes (Fig. 3, Online Resource 3) with most of the switchgrass linkage groups corresponding to a single foxtail millet chromosome. The exceptions were switchgrass LG III which largely corresponds to foxtail millet chromosome III but has a distal region composed of four loci that are orthologous to the distal region of foxtail millet chromosome VII, and switchgrass LG VII which largely corresponds to foxtail millet chromosome VII but has a distal region composed of seven loci that are orthologous to the distal region of foxtail millet chromosome III. In addition to this reciprocal translocation, nine inversions differentiate the switchgrass and foxtail millet

genomes (Fig. 3, Online Resource 3, Table 2). The inversions are located on switchgrass LGs I, III (2 inversions), IV (3 inversions), V, VI, and VII (Fig. 3, Online Resource 3). With the exception of two inversions on LG IV-b and one inversion on LG VII-a that could not be resolved on LGs IV-a and VII-b, respectively, all rearrangements were present on both switchgrass homoeologous chromosomes (Table 2). Fifty-eight markers (6.7 % of syntenic datapoints) mapped to non-colinear positions in the two species (Fig. 3, Online Resource 3). Overall, some 95 % of the foxtail millet genome is covered by the switchgrass consensus map with a minimum of 83.1 % for LGVIII-a and a maximum of 99.5 % for LGIII-a (Table 1).

IV-a

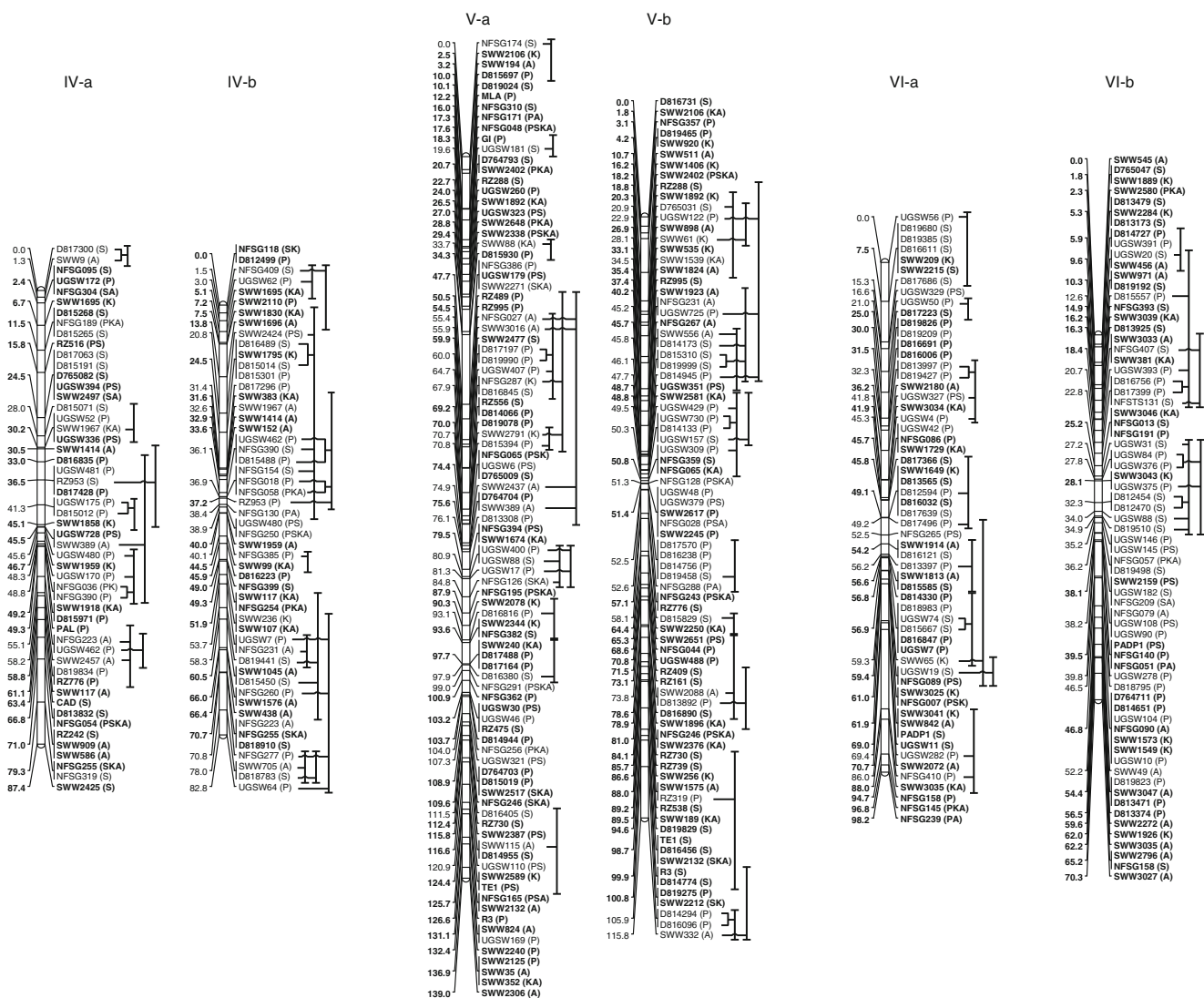| cM	Marker
0.0	D817300 (S)
1.3	SWW9 (A)
	NFSG095 (S)
2.4	UGSW172 (P)
	NFSG304 (SA)
6.7	SWW1695 (K)
	D815268 (S)
11.5	NFSG189 (PKA)
	D815265 (S)
15.8	RZ516 (PS)
	D817063 (S)
	D815191 (S)
24.5	D765082 (S)
	UGSW394 (PS)
	SWW2497 (SA)
28.0	D815071 (S)
	UGSW52 (P)
30.2	SWW1967 (KA)
	UGSW336 (PS)
30.5	SWW1414 (A)
33.0	D816835 (P)
	UGSW481 (A)
36.5	RZ953 (S)
	D817428 (P)
41.3	UGSW175 (P)
	D815012 (P)
45.1	SWW1858 (A)
45.5	UGSW728 (PS)
45.6	SWW389 (A)
46.7	SWW1959 (K)
48.3	UGSW170 (P)
	NFSG036 (PK)
48.8	SWW390 (A)
49.2	SWW1918 (KA)
49.3	D815971 (P)
	PAL (P)
55.1	NFSG223 (A)
	UGSW462 (A)
58.2	SWW2457 (A)
58.8	D819834 (P)
	RZ776 (P)
61.1	SWW117 (A)
63.4	CAD (S)
66.8	D813832 (A)
	NFSG054 (PSKA)
	RZ242 (P)
71.0	SWW909 (A)
	SWW586 (A)
79.3	NFSG255 (SKA)
	NFSG319 (S)
87.4	SWW2425 (S)

IV-b

cM	Marker
0.0	NFSG118 (SK)
	D812499 (P)
1.5	NFSG409 (A)
3.0	UGSW62 (P)
5.1	SWW1695 (KA)
7.2	SWW2110 (P)
7.5	SWW1830 (A)
13.8	SWW1696 (A)
20.8	SWW2424 (PS)
24.5	SWW1795 (K)
31.4	D816489 (S)
31.6	NFSG383 (KA)
32.6	SWW1967 (A)
32.9	SWW1414 (A)
33.6	SWW152 (A)
36.1	UGSW462 (P)
	NFSG390 (A)
	D815488 (A)
36.9	NFSG154 (S)
37.2	NFSG018 (P)
	NFSG058 (PKA)
38.4	RZ953 (P)
38.9	NFSG130 (PA)
40.0	UGSW480 (S)
40.1	NFSG250 (PSKA)
44.5	SWW1959 (A)
45.9	SWW99 (KA)
49.0	D816223 (P)
49.3	NFSG399 (S)
51.9	SWW117 (A)
	NFSG254 (PKA)
53.7	SWW236 (A)
58.3	SWW107 (KA)
60.5	NFSG7 (A)
	NFSG231 (A)
66.0	D819441 (S)
	UGSW1045 (A)
66.4	D815450 (S)
70.7	SWW1576 (A)
	SWW438 (A)
	NFSG255 (SKA)
70.8	NFSG256 (PKA)
	UGSW321 (PS)
78.0	SWW705 (A)
	D818783 (S)
82.8	UGSW64 (P)

V-a

cM	Marker
0.0	NFSG174 (S)
2.5	SWW2106 (K)
3.2	SWW194 (A)
10.0	D815697 (P)
10.1	D819024 (S)
12.2	MLA (P)
16.0	NFSG310 (S)
17.3	NFSG171 (PA)
17.6	NFSG048 (PSKA)
18.3	GI (P)
19.6	UGSW181 (S)
20.7	D764793 (S)
	SWW2402 (PKA)
22.7	RZ288 (S)
24.0	UGSW260 (P)
26.5	SWW1892 (KA)
27.0	SWW2648 (PKA)
28.8	SWW2338 (PSKA)
29.4	SWW88 (A)
33.7	D815930 (P)
34.3	RZ556 (S)
47.7	UGSW179 (PS)
	SWW2271 (SKA)
50.5	RZ489 (P)
54.5	RZ995 (P)
55.4	NFSG027 (A)
55.9	SWW3016 (A)
59.9	SWW2477 (S)
60.0	D817197 (P)
64.7	D819990 (P)
	UGSW407 (P)
67.9	NFSG287 (K)
	D816845 (S)
69.2	D814066 (P)
70.0	D819078 (P)
70.7	SWW2791 (K)
70.8	NFSG065 (PSK)
74.4	UGSW6 (PS)
74.9	D765009 (S)
75.6	UGSW2437 (A)
76.1	D764704 (S)
79.5	SWW389 (P)
	NFSG394 (PS)
	SWW1674 (KA)
80.9	UGSW400 (P)
81.3	UGSW88 (S)
84.8	NFSG126 (SKA)
87.9	NFSG195 (PSKA)
90.3	SWW2078 (K)
93.1	D816816 (A)
93.6	SWW2344 (K)
	NFSG382 (S)
97.7	SWW240 (KA)
	D817488 (P)
97.9	D817164 (P)
99.0	NFSG291 (PSKA)
100.9	NFSG362 (P)
	UGSW30 (PS)
103.2	UGSW46 (P)
103.7	RZ475 (S)
104.0	D814944 (P)
107.3	NFSG256 (PKA)
	UGSW321 (PS)
108.9	D764703 (P)
109.6	SWW2517 (SKA)
	NFSG246 (SKA)
111.5	D816405 (S)
112.4	RZ730 (S)
115.8	SWW2387 (PS)
116.6	SWW115 (A)
120.9	D814955 (S)
	UGSW110 (PS)
124.4	TE1 (PS)
125.7	NFSG165 (PSA)
	SWW2132 (A)
126.6	R3 (P)
131.1	SWW824 (A)
132.4	UGSW169 (P)
136.9	SWW2240 (P)
	SWW2125 (P)
	SWW35 (A)
	SWW352 (KA)
139.0	SWW2306 (A)

V-b

cM	Marker
0.0	D816731 (S)
1.8	SWW2106 (KA)
3.1	NFSG357 (P)
	D819465 (P)
4.2	SWW920 (K)
10.7	SWW511 (A)
16.2	SWW1406 (A)
18.2	SWW2402 (PSKA)
18.8	RZ288 (S)
20.3	SWW1892 (KA)
20.9	D765031 (S)
22.9	UGSW122 (A)
26.9	SWW898 (A)
28.1	SWW651 (A)
33.1	SWW535 (KA)
34.5	SWW1539 (KA)
35.4	SWW1824 (A)
37.4	RZ995 (P)
40.2	NFSG231 (A)
45.2	UGSW725 (P)
45.7	NFSG267 (A)
45.8	SWW556 (A)
	D814173 (S)
46.1	D815310 (S)
47.7	D819999 (S)
	UGSW407 (P)
48.7	UGSW351 (PS)
48.8	SWW2581 (KA)
49.5	UGSW429 (P)
50.3	UGSW730 (P)
	D814133 (P)
	UGSW157 (S)
	UGSW309 (P)
50.8	NFSG359 (KA)
	NFSG065 (KA)
51.3	NFSG128 (PSKA)
	UGSW48 (P)
	NFSG379 (PS)
51.4	SWW2617 (P)
	NFSG028 (PSA)
	SWW2245 (P)
52.5	SWW1914 (A)
	SWW1813 (A)
52.6	SWW243 (PS)
	RZ776 (S)
57.1	D815829 (S)
58.1	SWW2250 (KA)
64.4	SWW2651 (PS)
65.3	NFSG044 (P)
68.6	UGSW488 (P)
70.8	RZ409 (S)
71.5	RZ161 (S)
73.1	SWW2088 (A)
73.8	D816890 (S)
78.6	SWW1896 (KA)
78.9	NFSG246 (PKA)
81.0	SWW2376 (KA)
84.1	RZ730 (S)
85.7	RZ739 (S)
86.6	SWW256 (K)
88.0	SWW1575 (A)
89.2	RZ319 (P)
89.5	RZ538 (S)
94.6	SWW189 (A)
	D819829 (S)
98.7	TE1 (A)
	D816456 (S)
	SWW2132 (SKA)
99.9	R3 (S)
100.8	D814774 (P)
	D819275 (P)
	SWW2212 (SK)
105.9	D814294 (P)
115.8	D816096 (P)
	SWW332 (A)

VI-a

cM	Marker
0.0	UGSW56 (P)
	D819680 (S)
	D816611 (S)
7.5	
15.3	D817686 (S)
16.6	UGSW329 (PS)
21.0	UGSW50 (P)
25.0	D817223 (S)
30.0	D819826 (A)
	D819209 (P)
31.5	D816691 (P)
	D816006 (P)
32.3	D813997 (P)
	D819427 (P)
36.2	SWW2180 (A)
41.8	UGSW327 (PS)
41.9	SWW3034 (KA)
45.3	UGSW4 (P)
	UGSW42 (P)
45.7	NFSG086 (P)
	SWW1729 (KA)
45.8	D817366 (S)
	SWW1649 (K)
49.1	D813565 (S)
	D812594 (P)
	D816032 (S)
	D817639 (S)
49.2	D817494 (P)
52.5	NFSG265 (PS)
54.2	D817570 (P)
	D816238 (P)
	D814756 (P)
	D819458 (S)
56.2	NFSG288 (PA)
56.6	RZ2776 (S)
56.9	SWW74 (P)
	D815667 (S)
	D816847 (P)
	UGSW7 (P)
59.3	SWW65 (K)
59.4	SWW11 (A)
61.0	UGSW282 (P)
61.9	SWW2072 (A)
69.0	UGSW410 (P)
69.4	SWW3035 (KA)
70.7	NFSG158 (P)
86.0	NFSG145 (PKA)
88.0	NFSG239 (PA)
94.7	
96.8	
98.2	

VI-b

cM	Marker
0.0	SWW545 (A)
	D765047 (S)
	SWW1889 (K)
1.8	SWW2580 (PKA)
2.3	D813479 (S)
5.3	D813173 (S)
5.9	D814727 (P)
	UGSW391 (P)
9.6	SWW456 (A)
	SWW971 (A)
10.3	D819192 (S)
12.6	UGSW20 (S)
14.9	NFSG393 (S)
16.2	SWW3039 (KA)
16.3	D813925 (S)
18.4	SWW3033 (A)
	SWW381 (KA)
20.7	UGSW393 (P)
	D816756 (P)
22.8	NFST5131 (S)
25.2	SWW3046 (KA)
	NFSG013 (S)
27.2	NFSG191 (P)
	UGSW31 (S)
27.8	SWW84 (P)
	UGSW376 (P)
28.1	SWW3043 (K)
	UGSW375 (S)
32.3	D812454 (P)
34.0	D812470 (S)
34.9	D815010 (S)
35.2	UGSW146 (S)
	UGSW145 (PS)
36.2	NFSG057 (PKA)
38.1	SWW2159 (PS)
	UGSW182 (S)
	NFSG209 (SA)
38.2	NFSG079 (A)
	UGSW108 (PS)
39.5	PADP1 (PS)
	NFSG140 (P)
39.8	NFSG051 (PA)
46.5	UGSW278 (P)
	D818795 (P)
	D814651 (P)
46.8	UGSW104 (P)
	NFSG090 (A)
	SWW1573 (K)
	SWW1549 (K)
	UGSW10 (P)
52.2	SWW49 (A)
	D819823 (P)
54.4	SWW3047 (A)
	D813471 (P)
56.5	SWW2272 (A)
59.6	SWW1926 (K)
62.0	SWW3035 (A)
62.2	SWW2796 (A)
65.2	NFSG158 (S)
70.3	SWW3027 (A)

Fig. 2 (continued)

Marker Distribution

The positions of centromeres on the switchgrass linkage groups were extrapolated from the centromere locations in foxtail millet [20] (Online Resource 3). To evaluate marker distributions, each linkage group was divided into distal, proximal, and pericentromeric regions. Marker densities were calculated only for regions that could be clearly delineated (several regions with inversions were excluded) and that spanned at least 1 cM. The marker density across the pericentromeric regions was 1.29 markers per cM, across the proximal regions 0.75 markers per cM, and across the distal regions 0.49 markers per cM. The distribution of the different marker types (EST or gene-derived, PstI genomic SSRs, random genomic SSRs, and DArT markers) across the three chromosomal regions is given in Table 3. Genic markers made up a significantly higher percentage of the total number of markers in the distal chromosome regions compared to the pericentromeric regions ($\chi^2=21.37$; $p<0.001$), while the reverse was true for genomic SSR markers ($\chi^2=13.35$; $p<0.001$). There was no significant difference in the proportion of DArT markers in the distal compared to the pericentromeric regions ($\chi^2=2.12$; $p=0.15$). Different marker types varied significantly in their ability to identify orthologous sequences in foxtail millet ($\chi^2=513.80$; $p<0.001$). Colinear orthologs were found for 86 % of genic markers but only for 31 % of PstI-SSRs. Genes that mapped to pericentromeric regions were equally likely to identify colinear orthologs than genes that mapped to distal chromosome regions ($\chi^2=2.83$; $p=0.09$). While statistically, this was also true for genomic SSRs ($\chi^2=0.02$; $p=0.88$) and DArT markers ($\chi^2=3.18$; $p=0.07$), there was a trend for fewer orthologs to be identified for pericentromeric genomic markers than for proximal
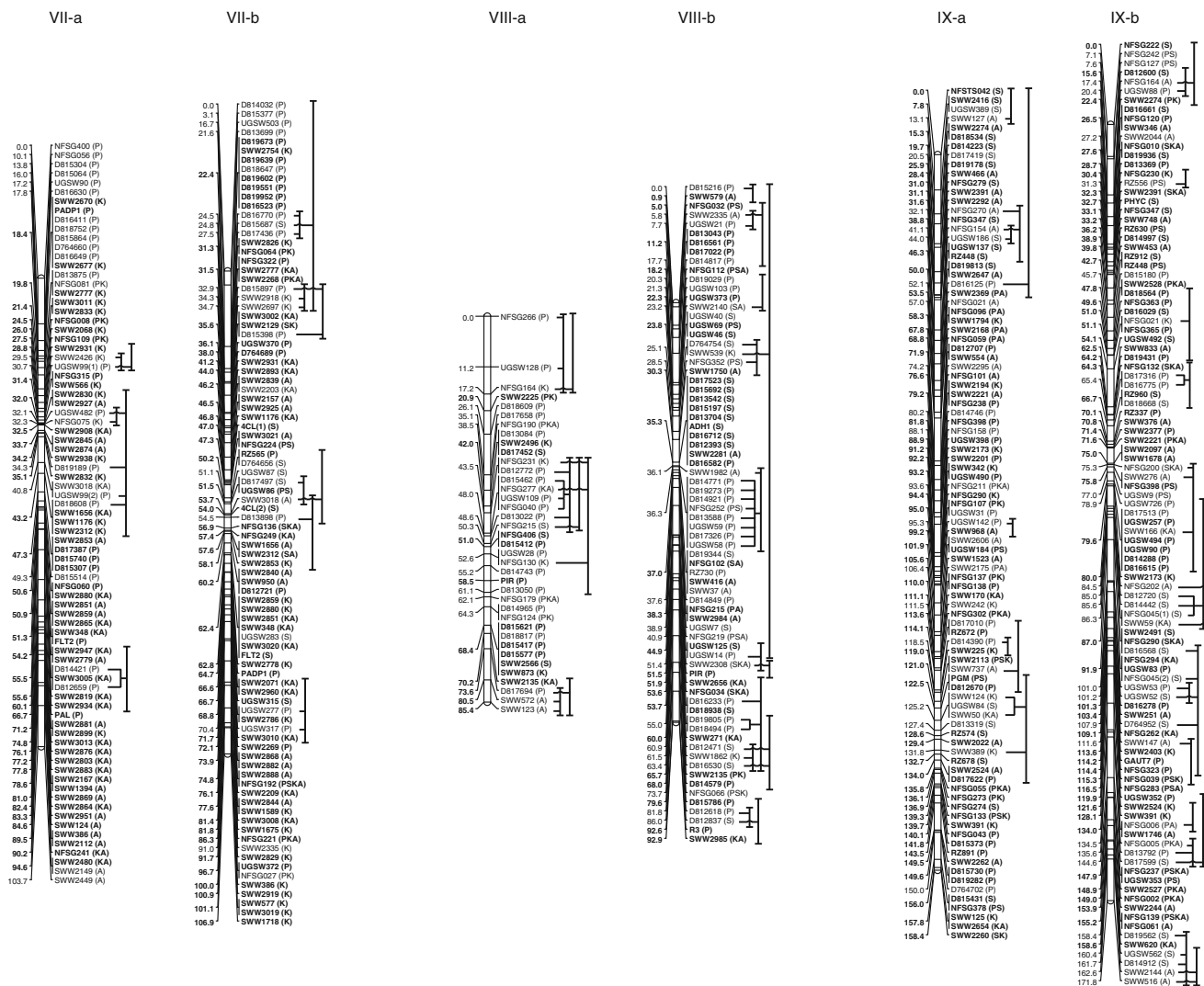
Fig. 2 (continued)

genomic markers than for distal genomic markers (Table 3).

Discussion

Consensus Map

The consensus map consisted of 1,321 loci and covered approximately 95 % of the switchgrass genome as estimated from the alignment of the switchgrass maps with the foxtail millet genome sequence. Our procedure for generating the consensus map combined the ease of use of JoinMap with the better error handling ability of Mapmaker. In our hands, merging all data (K Okada female, A Okada male, P Serba female, S Serba male) into a single large data file yielded consensus maps that were highly similar to the individual maps produced by Serba et al. [6] and Okada et al. [5]. Overall, 93 % of the markers mapped to the same relative position in the consensus map as in the individual published map. The two most likely explanations for the 7 % of discrepancies in map positions that we observed are scoring errors in one of the populations, which can lead to interstitially located markers to map to the end of a linkage group in an individual map, and the way Mapmaker and JoinMap handle scoring errors, which can lead to markers being ordered differently by different software programs [6]. The differences observed between the consensus and individual maps highlight the importance of the marker segregation quality and the software choice for developing a genetic map.

The distribution of loci is not uniform along the linkage groups. As expected, marker density (number of markers per cM) is highest in the pericentromeric regions of the genetic maps due to suppressed recombination. We also investigated whether genic and genomic markers were distributed

Table 1 The number of framework loci, the total number of loci, the length and percent coverage of switchgrass linkage groups

Linkage groups	I-a	I-b	II-a	II-b	III-a	III-b	IV-a	IV-b	V-a	V-b	VI-a	VI-b	VII-a	VII-b	VIII-a	VIII-b	IX-a	IX-b	Total
Number of loci (framework)	54	49	33	51	50	38	29	23	58	48	32	38	55	65	12	35	67	67	804
Number of loci (total)	72	70	63	87	82	68	52	52	91	80	58	69	81	90	38	72	93	103	1321
Total length (cM)	97	123	151.9	131.8	179.1	127	87.4	82.8	139	115.8	98.2	70.3	103.7	106.9	85.4	92.9	158.4	171.8	2122.4
Chromosome coverage (%)[a]	99.41	99.09	86.53	97.31	99.51	99.12	97.95	96.64	94.98	93.03	97.06	99.03	93.64	92.16	83.06	96.24	97.52	96.82	95.51

[a] The chromosome coverage represents the percent coverage of the syntenic foxtail millet chromosome by the most distal markers on the switchgrass map with orthologs in foxtail millet

differently along a linkage group. Gene-derived markers accounted for 44 % of the total marker number. However, they were overrepresented (52 %) in the distal regions of the linkage groups and underrepresented (30 %) in the pericentromeric regions (Table 3). This is consistent with the fact that genes are mostly located in distal chromosome regions [20, 25–28]. In contrast, genomic SSR markers and DArT markers made up a higher percentage of the total markers in pericentromeric regions than those in distal chromosome regions (Table 3), suggesting that the genomic markers that map to the pericentromeric regions likely represent repeats. Consequently, orthologous sequences were identified more frequently for genic markers than for genomic markers (Table 3). There was no significant difference between the chromosome regions in the percentage of genic markers with syntenic orthologs. Although more SSR markers that mapped to the distal chromosome regions had a foxtail millet ortholog compared to markers that mapped to pericentromeric regions (Table 3), this difference was not significant possibly because the power of the statistical analysis was reduced by the small number of markers in the pericentromeric regions. Nevertheless, the observed trend suggests that genomic SSRs that mapped to distal chromosome regions were more likely to be associated with genes, while genomic SSRs that mapped to pericentromeric regions were more likely to be located within more rapidly evolving repetitive DNA. The SSRs for which foxtail millet orthologs were identified were not necessarily located within genes. Because the SSR-containing switchgrass sequences were extended by 1 kb on either end by conducting a BLASTn search of the SSR clone sequences against the switchgrass genome sequence before them as queries in a BLASTn search against the foxtail millet genome sequence, the extended sequence might contain part of a switchgrass coding region if the SSR was located within ~1 kb of gene. Therefore, foxtail millet orthologs would likely be identified for any switchgrass SSR located within ~1 kb of a coding region.

Comparative Analysis with *Setaria*

The outcrossing nature and polyploidy of switchgrass complicate genetic analyses. Having a closely related diploid species with a well-assembled genome can be very helpful in unraveling traits into their genetic components and cloning the underlying genes. However, it is important to know the structural relationship between the model and target genomes. In 2010, the genome of foxtail millet, a species that diverged some 13.1 MYA from switchgrass, was sequenced as a model for switchgrass [20]. We therefore conducted a comparative analysis between the published foxtail millet genome and our switchgrass consensus map. Both switchgrass and foxtail millet have a basic chromosome number of nine, and the majority of loci that mapped to a single switchgrass linkage group identified a single foxtail
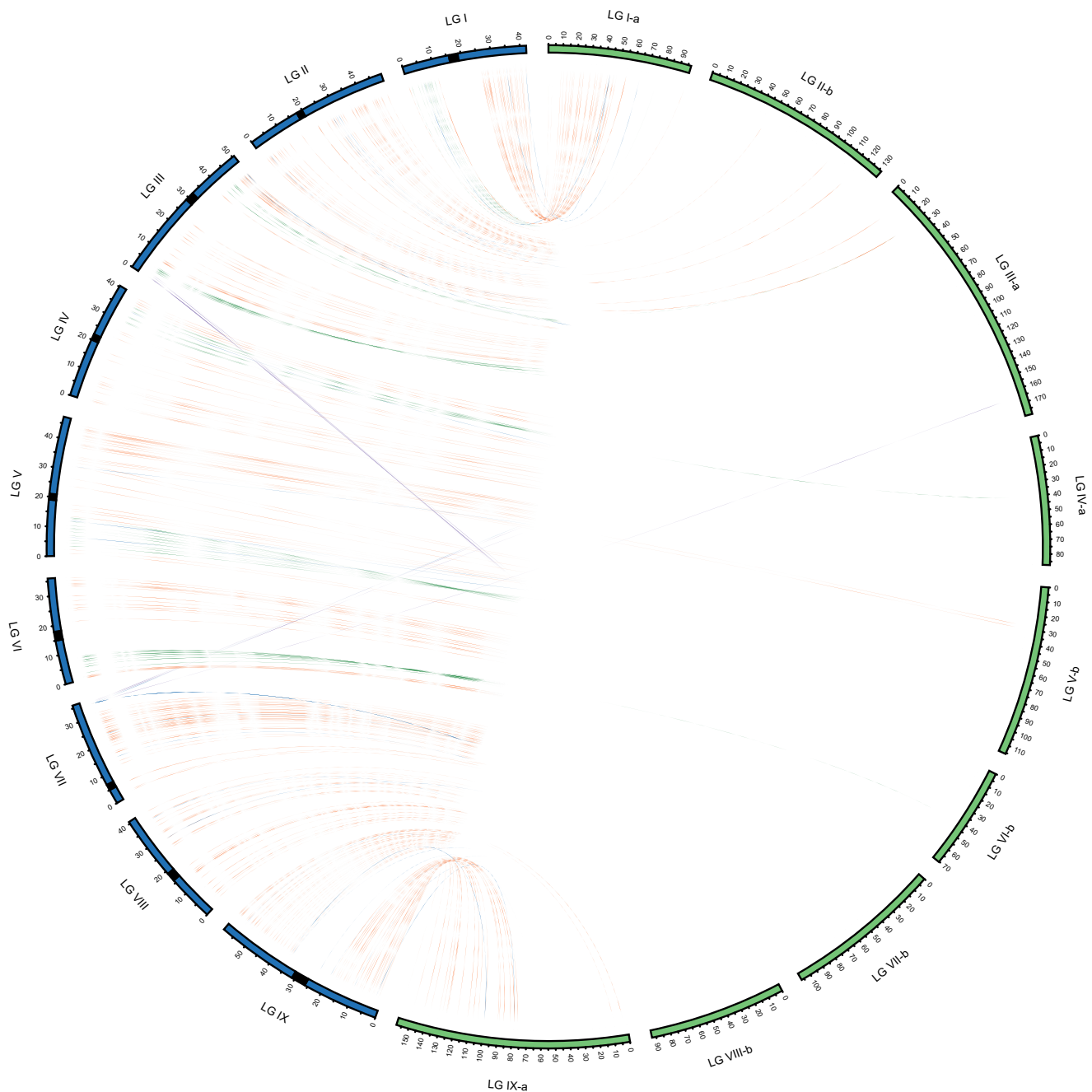
Fig. 3 Comparative relationship between loci on the switchgrass genetic map and their putative orthologs in the foxtail millet genome sequence. Foxtail millet chromosomes are in *blue* on the *left-hand side* of the *circle diagram*. Switchgrass linkage groups are in *green* on the *right-hand side* of the *circle diagram*. Because no rearrangements were identified between the homoeologous switchgrass chromosomes, only the linkage group (either *a* or *b*) with the highest number of comparative datapoints is represented. Markers that are colinear in switchgrass and foxtail millet chromosome indicating high levels of synteny between switchgrass and foxtail millet (Fig. 3, Online Resource 3). We therefore discarded the 29 % of BLASTn hits that mapped to nonsyntenic chromosomes, many of which represented repetitive sequences, as nonorthologous. The exceptions were four

are connected by *orange lines*, markers that characterize an inversion in switchgrass relative to foxtail millet are connected by *green lines*, markers that characterize an interchromosomal translocation in switchgrass relative to foxtail millet are connected by *purple lines*, and markers that are syntenic but non-colinear are connected by *blue lines*. Putative locations of centromeres on foxtail millet chromosomes are indicated with *black bars*

and three BLASTn hits to the region 33–35.2 Mb on foxtail millet chromosome VII that were identified when adjacent markers at the top of switchgrass LGs III-a and III-b, respectively, were used as queries, and seven BLASTn hits to the region 0.73–2 Mb on foxtail millet chromosome III identified using

Table 2 Breakpoints of rearrangements that differentiate the switchgrass linkage groups and foxtail millet chromosomes

Switchgrass LG	Type of rearrangement[a]	Breakpoint regions in switchgrass (cM)[a]	Foxtail millet chromosome	Breakpoint regions in foxtail millet (Mb)[a]	Region on ancestral chromosome[b]	Breakpoint regions of rearrangements that occurred in foxtail millet lineage[c]
I-a	Inversion	3.5-17.5 and 39.5-41.9	I	119,721-1,132,304 and 9,377,550-10,826,533		490,951-520,710 and 10,391,435-10,428,890
I-b	Inversion	0-2.6 and 32.3-40.3		398,765-5,004,140 and 10,069,278-10,826,533		
III-a	Translocation	0 and 45.6-46.3	VII	31,525,483-34,400,543 and 35,233,600-end	9.5inv, 8.4, 8.2, 8.3inv, 8.1	31,746,096-31,761,512 and 35,912,382-end
III-b	Translocation	0 and 23.7-27.7		30,888,778-32,999,598 and 34,284,539-end		
III-a	Inversion	45.6-46.3 and 66.1-66.3	III	1,993,865-2,939,364 and 4,746,666-7,269,995	9.4inv, 9.3, 9.2inv, 9.7, 8.5inv	2,117,275-2,185,638 and 6,425,689-6,447,471
III-b	Inversion	23.7-27.7 and 47.3-54	III	1,768,712-2,419,914 and 6,390,537-8,216,595		
III-a	Inversion	103.5-108 and 143.2	III	38,840,185-41,100,453 and 44,622,357-47,093,214	8.7inv	40,741,895-40,884,273 and 46,462,404 - 46,586,884
III-b	Inversion	86-89.2 and 103-109		39,261,617-40,995,879 and 45,941,660-47,634,375		
IV-a	ND		IV			
IV-b	Inversion	24.5-31.6 and 33.6-40		4,242,054-8,475,452 and 10,527,847-26,751,651		
IV-a	Inversion	46.7-49.2 and 66.8-71	IV	26,751,651-30,292,666 and 36,386,856-37,105,325		30,137,708-30,209,160 and 36,435,244-36,438,721
IV-b	Inversion	40-44.5 and 51.9-60.5		26,751,651-31,500,426 and 34,638,866-36,921,693		
IV-a	ND		IV			36,435,244-36,438,721 and 38,606,592-38,681,949
IV-b	Inversion	51.9-60.5 and 66.4-70.7		34,638,866-36,921,693 and 38,063,436-38,758,816		
V-a	Inversion	2.5-3.2 and 70-74.4	V	2,108,562-2,568,392 and 12,959,031-17,969,260		2,172,262-2,268,601 and 13,477,437-13,720,765
V-b	Inversion	1.8-3.1 and 40.2-45.7		2,108,562-3,381,952 and 13,404,935-13,872,566		
VI-a	Inversion	7.5-25 and 54.2-56.6	VI	580,711-1,762,329 and 6,560,799-18,269,576		890,822-893,586 and 8,064,531-8,086,597
VI-b	Inversion	5.3-5.9 and 28.1-38.1		759,895-1,951,810 and 8,053,014-20,443,680		
VII-a	Translocation	81-82.4 and 94.6-end	III	0-49,680 and 1,993,865-2,939,364	6.6	0-9,090 and 1,234,577-1,252,759
VII-b	Translocation	86.3-91.7 and 106.9-end		0-283,903 and 1,222,662-2,419,914		
VII-a	Inversion	81-82.4 and 84.6-89.5	III	530,626-1,504,982 and 1,993,865-2,939,364	6.5inv	1,234,577-1,252,759 and 2,010,660-2,185,638
VII-b	ND					

[a] Based on comparative analysis of an integrated switchgrass linkage map with the foxtail millet genome sequence (this study)

[b] Only given for chromosomes III and VII which underwent a series of rearrangements (see Fig. 4)

[c] Based on foxtail millet–sorghum–rice comparative analysis [20]

adjacent markers at the bottom of both switchgrass LGs VII-a and VII-b (Fig. 3, Online Resource 3). The close linkage of the putative orthologous loci on switchgrass LGs III and VII, and foxtail millet chromosomes VII and III, respectively, indicated the presence of a reciprocal translocation that took place in either the switchgrass or foxtail millet lineage. In addition to the reciprocal translocation, nine inversions differentiated the switchgrass and foxtail millet genomes. The breakpoints of these rearrangements are listed in Table 2. Centromere locations in switchgrass were extrapolated from the centromere locations in foxtail millet and mapped to the region with the highest ratio of physical distance (measured in foxtail millet) to genetic distance (measured in switchgrass) (Online Resource 3).

A comparative analysis of the foxtail millet genome with that of sorghum had previously revealed that sorghum chromosomes 6, 8, and 9 corresponded to foxtail millet

Table 3 Distribution of different marker types across distal, proximal, and pericentromeric chromosome regions

Marker type	Distal[a]			Proximal[a]			Peri-centromeric[a]			Across all regions		
	No. of markers[b]	% of markers	No. and (%) of markers with colinear orthologs	No. of markers[b]	% of markers	No. and (%) of markers with colinear orthologs	No. of markers[b]	% of markers	No. and (%) of markers with colinear orthologs	No. of markers[b]	% of markers	No. and (%) of markers with colinear orthologs
EST/gene-derived[c]	281 (266)	51.7 %	230 (86.5 %)	166 (157)	42.6 %	118 (75.2 %)	43 (42)	29.7 %	34 (81.0 %)	578 (550)	43.8 %	472 (85.8 %)
PstI-SSRs (ugsw)	45 (45)	8.3 %	16 (35.6 %)	59 (59)	15.1 %	16 (27.1 %)	24 (24)	16.6 %	2 (8.3 %)	166 (166)	12.6 %	51 (30.7 %)
Random SSRs (nfsg)	84 (76)	15.4 %	55 (72.4 %)	92 (85)	23.6 %	50 (58.8 %)	33 (29)	22.8 %	12 (41.4 %)	241 (220)	18.2 %	138 (62.7 %)
DArT[d]	134 (122)	24.6 %	62 (50.8 %)	73 (62)	18.7 %	23 (37.1 %)	45 (44)	31.0 %	18 (40.9 %)	336 (308)	25.4 %	143 (46.4 %)
Overall	544 (509)	100 %	363 (71.3 %)	390 (363)	100 %	207 (57.0 %)	145 (139)	100 %	66 (47.5 %)	1321 (1244)	100 %	804 (64.6 %)

[a] See "Materials and Methods" for the definition of distal, proximal, and pericentromeric chromosome regions

[b] The number in parenthesis is the number of markers for which sequence information is available

[c] Includes EST-SSRs (sww loci), EST-STS, and EST RFLP markers (RZ loci)

[d] DArT markers that cosegregated and had the same foxtail millet ortholog were counted as a single locus

chromosomes III and VII and that both sets of chromosomes were differentiated by a complex set of rearrangements [20]. The remaining sorghum and foxtail millet chromosomes showed one-to-one relationships, with inversions differentiating the syntenic foxtail millet (sorghum) chromosomes I (4) (2 inversions), IV (10) (4 inversions), V (3) (1 inversion), VI (7) (3 inversions), and VIII (5) (1 inversion) [20]. Of these, seven inversions also differentiated foxtail millet and rice, and hence occurred in the foxtail millet lineage after its divergence from a common ancestor with sorghum. Three inversions were specific to the sorghum lineage. For one inversion (on chromosome V), the evolutionary time point could not be established. Because the breakpoints of the inversions and translocations were not included in Bennetzen et al. [20], we have listed those that occurred in the foxtail millet lineage and that coincide with rearrangements that differentiate foxtail millet and switchgrass in Table 2. All rearrangements that differentiate switchgrass from foxtail millet, except one inversion on switchgrass LG IV-b, were rearrangements that had previously been shown to have occurred in foxtail millet after the divergence of foxtail millet and sorghum from a common ancestor (Table 2). The inversion on switchgrass LG IV-b included only three comparative data points and could not be confirmed on LG IV-a due to a lack of data in that region. Our analyses also showed that the inversion on foxtail millet chromosome V for which the timing had not previously been determined and part of the complex rearrangement that formed foxtail millet chromosomes III and VII were also specific to foxtail millet. Some of the foxtail millet lineage-specific rearrangements (Fig. 4) identified in comparative analyses with sorghum and rice were not covered by markers in the switchgrass consensus map, and hence, we could not determine their status in switchgrass.

The fact that part of the complex chromosome III–chromosome VII rearrangement that occurred in the foxtail millet lineage after its divergence from sorghum predates the divergence of switchgrass and foxtail millet and part occurred after the divergence of the two species allowed us to develop a model for the evolution of those chromosomes (Fig. 4). The model is based on the assumption that the most likely path of chromosome evolution from ancestral state to current-day state involved a minimal number of rearrangements. We considered the structure of current-day sorghum chromosomes 6, 8, and 9, bar an inversion on chromosome 8 in sorghum which postdates the divergence of the sorghum and foxtail millet lineages, as the ancestral chromosome configuration. The model hypothesizes a chromosome fusion in which ancestral chromosome 9 was inserted into the pericentromeric region of chromosome 8. This was followed by three inversions that led to the structure of current-day switchgrass chromosome III. The relative order of the rearrangements in the model is random, except when rearrangements were known to predate (rearrangements 1, 2, 3, and 4) or postdate (rearrangements 5,
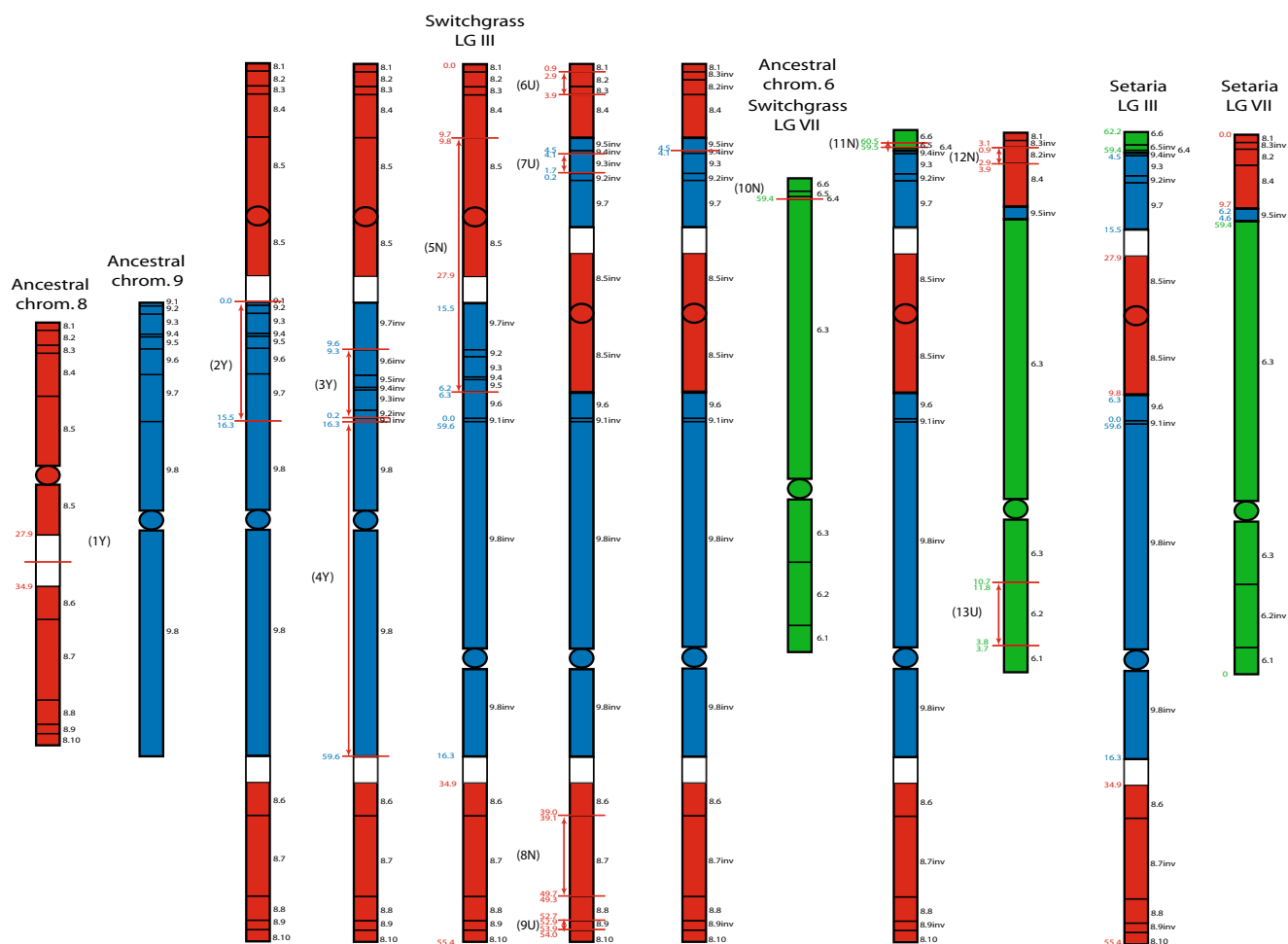
Fig. 4 Model showing the evolution from three ancestral chromosomes to current-day foxtail millet chromosomes III and VII. *Numbers on the right-hand side* of each chromosome give the ancestral chromosome number and regions. Chromosome regions are delineated by rearrangement breakpoints. *Numbers on the left-hand side* of the chromosome indicate the location of the chromosomal breakpoints in the sorghum genome, which was taken as a proxy for the chromosome structure of the common ancestor to sorghum and switchgrass/foxtail millet. Evolutionary events are *numbered* (*in parenthesis*), and each number is followed by *Y*, *N*, or *U* indicating whether this rearrangements took place (*Y*), did not take place (*N*), or could not be assessed (*U*) in switchgrass. Event (*1*) is a nested chromosome insertion, event (*10*) is a reciprocal translocation, and the remaining events are intrachromosomal inversions

8, 10, and 11) the divergence of switchgrass and foxtail millet, or when they involved overlapping chromosome segments (e.g., rearrangements 2, 3, and 5).

Our analysis could only identify chromosomal rearrangements that spanned, on average, at least 1.2 mega bases in foxtail millet due to the limited resolution of the switchgrass genetic map. Nevertheless, at the gross chromosomal level, chromosome evolution seems to have occurred at a faster pace in foxtail millet compared to switchgrass during the past 13.1 MY. An interesting question is whether chromosome level evolution is correlated with rearrangements at the gene level. Massa et al. [29] found that the rate of chromosomal rearrangements and the rate of single-gene deletions and insertions in the grass species *Aegilops tauschii*, *Brachypodium distachyon*, *Oryza sativa*, and *Sorghum bicolor* were highly correlated. However, single-gene inversions were not analyzed in the Massa et al. [29] study. An analysis of the orientation of genes in switchgrass BAC-end sequences

relative to both foxtail millet and sorghum has shown that single-gene inversions occur at a higher rate in foxtail millet compared to sorghum [20], suggesting that the same mechanism may underlie single-gene inversions and inversions of large gene blocks.

The accumulation of a higher number of rearrangements in foxtail millet compared to switchgrass was a somewhat surprising finding considering that switchgrass is a tetraploid with, presumably, a greater tolerance to rearrangements than the diploid foxtail millet. However, the divergence of the switchgrass A and B genomes and their subsequent hybridization might have been too recent for rearrangements to have accumulated in the switchgrass genome. Current phylogenies suggest that the two diploid genomes that formed tetraploid switchgrass diverged after the split of *Panicum* section *Panicum*, to which *P. hallii* belongs, and *Panicum* section *Virgata*, to which switchgrass belongs [30–32]. Because the divergence of

P. hallii and *P. virgatum* has been dated to approximately 5.3 MYA [33], tetraploid *P. virgatum* must have been formed in the past 5 MY. This number is in close agreement with divergence estimates based on 6,800 bp of sequence data of 6 MYA for the divergence of the switchgrasses A and B genomes and 3.8 MYA for the split between the closely related A genome of switchgrass and E genome of *Panicum rudgei* (B. Bahri and K.M. Devos, unpublished data).

Fate of Nonfunctional Centromeres after Chromosome Fusion

The insertion of ancestral chromosome 9 into ancestral chromosome 8 to form switchgrass and foxtail millet chromosomes III provided us with an opportunity to analyze the fate of the two centromeres in the fused chromosome. Luo and colleagues [34] noted that nested chromosome insertions during the evolution of *Ae. tauschii* chromosomes led to the loss of the centromere of the chromosome in which the insertion occurred (recipient chromosome) [34]. The same observation was made for nested chromosome insertions in *B. distachyon* and sorghum [34] and, in this study, in switchgrass/foxtail millet. In an attempt to uncover what happens after centromere inactivation, we plotted the ratio of the length of sorghum and foxtail millet synteny blocks, as determined by Bennetzen et al. [20], along each of the nine foxtail millet chromosomes (Fig. 5). The average ratio was 2.11, which corresponds to the approximately 2-fold larger genome size of sorghum compared to foxtail millet. However, as expected, the ratio was not constant across the length of the chromosomes but

peaked in the regions corresponding to the ten sorghum centromeres indicating that most of the genome expansion in sorghum occurred in the centromeric region. Interestingly, where the average ratio of the length of a sorghum block to the length of the syntenic foxtail millet block was 3.93 around the active foxtail millet centromeres, this ratio was 15.71 in the region of the "lost" foxtail millet centromere. A similar analysis of the relative length of syntenic rice and foxtail millet regions showed an average value of 1.08 along the length of the chromosome, a value of 0.84 around the active foxtail millet centromeres, and a value of 3.58 around the lost centromere on foxtail millet chromosome III. The <1 ratio around the foxtail millet centromeres indicates that centromeric regions in the 510-Mb foxtail millet genome have expanded more than the centromeric regions in the 430-Mb rice genome. Furthermore, the value of 3.58 at the position of the lost centromere in foxtail indicates that this region has been contracting in foxtail millet. We hypothesize that during a nested insertion, the centromere in the recipient chromosome is effectively moved to a more distal region where recombination rates are higher, which promotes removal of repeats around the inactive centromere through homologous recombination. It has been a long-standing question whether the accumulation of transposable elements in low recombinant regions is due to reduced recombination rates (ectopic recombination model [35]) or due to weaker selection against deleterious insertions in gene-poor regions [36]. Analysis of the repeat content around inactivated centromeres that have been moved to regions with higher recombination rates following a nested insertion event
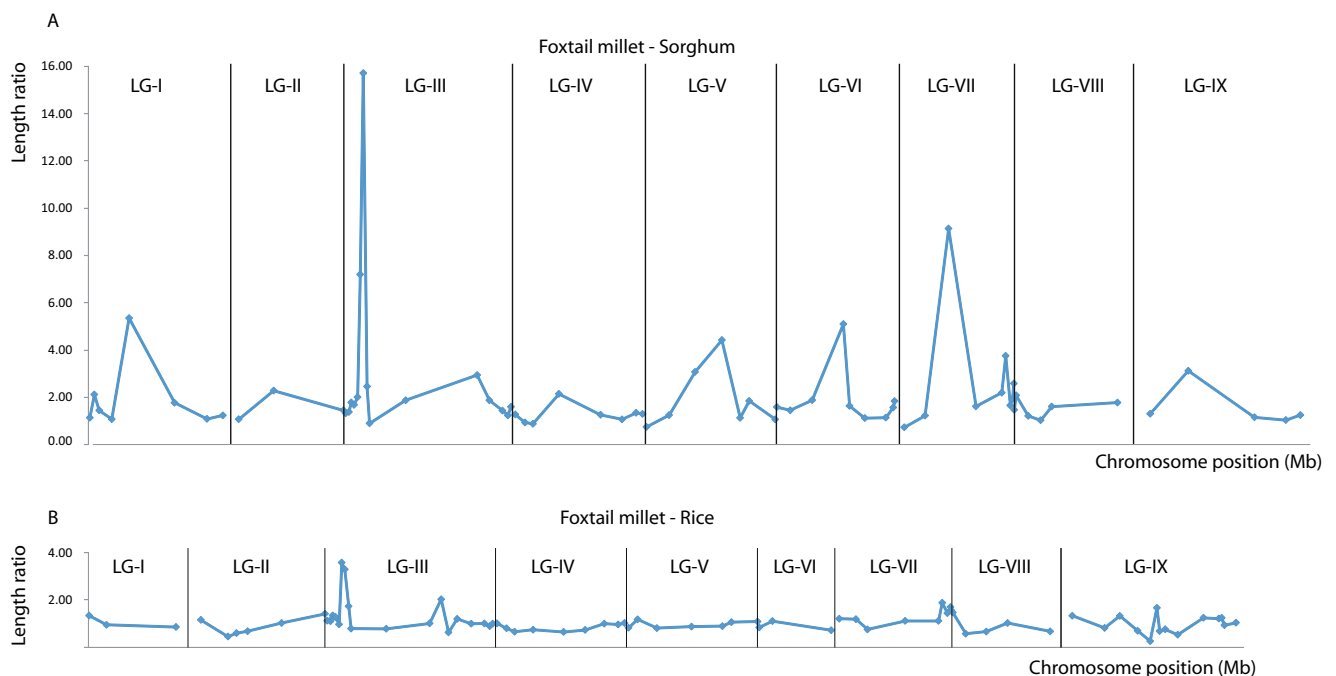

Fig. 5 Graph showing syntenic block length ratios along the length of the foxtail millet chromosomes. **a** Ratio of length of syntenic blocks in sorghum and foxtail millet. **b** Ratio of length of syntenic blocks in rice and foxtail millet. Foxtail millet chromosomes (labeled *LG I to LG IX*) are separated by *vertical lines*. For each set of syntenic blocks, the ratio is plotted against the location of the midpoint of the syntenic block on the foxtail millet chromosome

could help in uncoupling the effects of recombination and gene density.

Conclusion

Switchgrass consensus maps were developed as a tool for the switchgrass community to facilitate trait analyses and to provide the marker density needed to conduct comparative analyses between switchgrass and its genomic model, foxtail millet. The two species differed by one reciprocal translocation and at least nine inversions, all but one of which had occurred in the foxtail millet lineage during the past 13 MY. Considering that there are further rearrangements in the foxtail millet genome for which the status could not be determined in switchgrass due to a lack of markers in the relevant regions, it is likely that comparative analyses at the whole genome level will uncover additional differences between the switchgrass and foxtail millet genomes. The rearrangements need to be taken into account when using foxtail millet as a proxy for switchgrass in genetic and genomic analyses.

In addition to providing a resource for practical applications in breeding and genomic analyses, the newly gained comparative knowledge yielded insight into the evolutionary events that led to a reduction in chromosome number from 10 in the tribe Andropogoneae to 9 in the tribe Paniceae and, ultimately, to the formation of current-day switchgrass and foxtail millet chromosomes. The structure of the switchgrass genome, which is intermediate between the structures of the sorghum and foxtail millet genomes and appears to have been relatively stable within the past 13 MY, makes switchgrass an interesting species to include in evolutionary studies within the Panicoideae.

References

1. Sanderson MA, Reed RL, McLaughlin SB, Wullschleger SD, Conger BV, Parrish DJ, Wolf DD, Taliaferro C, Hopkins AA, Ocumpaugh WR, Hussey MA, Read JC, Tischler CR (1996) Switchgrass as a sustainable bioenergy crop. Bioresource Technol 56:83–93

2. Christian DC, Elberson HW (1998) Switchgrass (*Panicum virgatum* L.). In: El Bassam N (ed) Energy plant species: their use and impact on environment and development. James and James Ltd., London, pp 257–263

3. Alexopoulou E, Sharma N, Papatheohari Y, Christou M, Piscioneri I, Panoutsou C, Pignatelli V (2008) Biomass yields for upland and lowland switchgrass varieties grown in the Mediterranean region. Biomass Bioenerg 32:926–933

4. Missaoui AM, Paterson AH, Bouton JH (2005) Investigation of genomic organization in switchgrass (*Panicum virgatum* L.) using DNA markers. Theor Appl Genet 110:1372–1383

5. Okada M, Lanzatella C, Saha MC, Bouton J, Wu R, Tobias CM (2010) Complete switchgrass genetic maps reveal subgenome collinearity, preferential pairing and multilocus interactions. Genetics 185: 745–760

6. Serba D, Wu L, Daverdin G, Bahri BA, Wang X, Kilian A, Bouton JH, Brummer EC, Saha MC, Devos KM (2013) Linkage maps of lowland and upland tetraploid switchgrass ecotypes. Bioenergy Research 6:953–965

7. Liu L, Wu Y, Wang Y, Samuels T (2012) A high-density simple sequence repeat-based genetic linkage map of switchgrass. G3: Genes Genomes Genetics 2:357–370

8. Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, Buckler ES, Costich DE (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. PLoS Genet 9:e1003215

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

10. Van Ooijen JW (2006) JoinMap ® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V, Wageningen

11. Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: software for genetic mapping in outcrossing species. Hereditas 144:78–79

12. Falque M, Décousset L, Dervins D, Jacob AM, Joets J, Martinant JP, Raffoux X, Ribière N, Ridel C, Samson D, Charcosset A, Murigneux A (2005) Linkage mapping of 1454 new maize candidate gene loci. Genetics 170:1957–1966

13. Ming R, Liu SC, Bowers JE, Moore PH, Irvine JE, Paterson AH (2002) Construction of a *Saccharum* consensus genetic map from two interspecific crosses. Crop Sci 42:570–583

14. Qi X, Pittaway TS, Lindup S, Liu H, Waterman E, Padi FK, Hash CT, Zhu J, Gale MD, Devos KM (2004) An integrated genetic map and a new set of simple sequence repeat markers for pearl millet, *Pennisetum glaucum*. Theor Appl Genet 109:1485–1493

15. Rajaram V, Nepolean T, Senthilvel S, Varshney R, Vadez V, Srivastava R, Shah T, Supriya A, Kumar S, Kumari BR, Bhanuprakash A, Narasu ML, Riera-Lizarazu O, Hash CT (2013) Pearl millet [*Pennisetum glaucum* (L.) R. Br.] consensus linkage map constructed using four RIL mapping populations and newly developed EST-SSRs. BMC Genomics 14:159

16. Gustafson JP, Ma XF, Korzun V, Snape JW (2009) A consensus map of rye integrating mapping data from five mapping populations. Theor Appl Genet 118:793–800

17. Marone D, Laidò G, Gadaleta A, Colasuonno P, Ficco DB, Giancaspro A, Giove S, Panio G, Russo MA, Vita P, Cattivelli L, Papa R, Blanco A, Mastrangelo AM (2012) A high-density consensus map of A and B wheat genomes. Theor Appl Genet 125:1619–1638

18. Sharma MK, Sharma R, Cao P, Jenkins J, Bartley LE, Qualls M, Grimwood J, Schmutz J, Rokhsar D, Ronald PC (2012) A genome-wide survey of switchgrass genome structure and organization. PLoS ONE 7:e33892

19. Vicentini A, Barber JC, Giussani LM, Aliscioni SS, Kellogg EA (2008) Multiple coincident origins of C4 photosynthesis in the Mid-to Late Miocene. Global Change Biol 14:2963–2977

20. Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn J, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang X, Wu X, Mitros T, Triplett J, Yang X, Ye C-Y, Mauro-Herrera M, Wang L, Li P, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell T, Doust AN, Tuskan GA, Rokhsar D, Devos KM (2012) Reference genome sequence of the model plant Setaria. Nat Biotechnol 30:555–561

21. Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SE, Second G, McCouch SR, Tanksley SD (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138:1251–1274

22. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

23. Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645

24. Team RDC (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

25. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang HB, Wang XY, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang LF, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, Mccann MC, Ming R, Peterson DG, Mehboob uR, Ware D, Westhoff P, Mayer KFX, Messing J, Rokhsar DS (2009) The Sorghum bicolor genome and the diversification of grasses. Nature 457:551–556

26. Yu J, Hu SN, Wang J, Wong GKS, Li SG, Liu B, Deng YJ, Dai L, Zhou Y, Zhang XQ, Cao ML, Liu J, Sun JD, Tang JB, Chen YJ, Huang XB, Lin W, Ye C, Tong W, Cong LJ, Geng JN, Han YJ, Li L, Li W, Hu GQ, Huang XG, Li WJ, Li J, Liu ZW, Li L, Liu JP, Qi QH, Liu JS, Li L, Li T, Wang XG, Lu H, Wu TT, Zhu M, Ni PX, Han H, Dong W, Ren XY, Feng XL, Cui P, Li XR, Wang H, Xu X, Zhai WX, Xu Z, Zhang JS, He SJ, Zhang JG, Xu JC, Zhang KL, Zheng XW, Dong JH, Zeng WY, Tao L, Ye J, Tan J, Ren XD, Chen XW, He J, Liu DF, Tian W, Tian CG, Xia HG, Bao QY, Li G, Gao H, Cao T, Wang J, Zhao WM, Li P, Chen W, Wang XD, Zhang Y, Hu JF, Wang J, Liu S, Yang J, Zhang GY, Xiong YQ, Li ZJ, Mao L, Zhou CS, Zhu Z, Chen RS, Hao BL, Zheng WM, Chen SY, Guo W, Li GJ, Liu SQ, Tao M, Wang J, Zhu LH, Yuan LP, Yang HM (2002) A draft sequence of the rice genome (Oryza sativa L. ssp indica). Science 296:79–92

27. Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K, Lucas S, Harmon-Smith M, Lail K, Tice H, Grimwood J, McKenzie N, Huo NX, Gu YQ, Lazo GR, Anderson OD, You FM, Luo MC, Dvorak J, Wright J, Febrer M, Idziak D, Hasterok R, Lindquist E, Wang M, Fox SE, Priest HD, Filichkin SA, Givan SA, Bryant DW, Chang JH, Wu HY, Wu W, Hsia AP, Schnable PS, Kalyanaraman A, Barbazuk B, Michael TP, Hazen SP, Bragg JN, Laudencia-Chingcuanco D, Weng YQ, Haberer G, Spannagl M, Mayer K, Rattei T, Mitros T, Lee SJ, Rose JKC, Mueller LA, York TL, Wicker T, Buchmann JP, Tanskanen J, Schulman AH, Gundlach H, de Oliveira AC, Maia LD, Belknap W, Jiang N, Lai JS, Zhu LC, Ma JX, Sun C, Pritham E, Salse J, Murat F, Abrouk M, Bruggmann R, Messing J, Fahlgren N, Sullivan CM, Carrington JC, Chapman EJ, May GD, Zhai JX, Ganssmann M, Gurazada SGR, German M, Meyers BC, Green PJ, Tyler L, Wu JJ, Thomson J, Chen S, Scheller HV, Harholt J, Ulvskov P, Kimbrel JA, Bartley LE, Cao PJ, Jung KH, Sharma MK, Vega-Sanchez M, Ronald P, Dardick CD, De Bodt S, Verelst W, Inze D, Heese M, Schnittger A, Yang XH, Kalluri UC, Tuskan GA, Hua ZH, Vierstra RD, Cui Y, Ouyang SH, Sun QX, Liu ZY, Yilmaz A, Grotewold E, Sibout R, Hematy K, Mouille G, Hofte H, Pelloux J, O'Connor D, Schnable J, Rowe S, Harmon F, Cass CL, Sedbrook JC, Byrne ME, Walsh S, Higgins J, Li PH, Brutnell T, Unver T, Budak H, Belcram H, Charles M, Chalhoub B, Baxter I (2010) Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature 463:763–768

28. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA et al (2009) The B73 maize genome: complexity, diversity and dynamics. Science 326:1112–1115

29. Massa AN, Wanjugi H, Deal KR, O'Brien KO, You FM, Maiti R, Chan A, Gu YQ, Luo MC, Anderson OA, Rabinowicz PD, Dvorak J, Devos KM (2011) Gene space dynamics during the evolution of Aegilops tauschii, Brachypodium distachyon, Oryza sativa, and Sorghum bicolor genomes. Mol Biol Evol 28:2537–2547

30. Triplett JK, Wang Y, Zhong J, Kellogg EA (2012) Five nuclear loci resolve the polyploid history of switchgrass (Panicum virgatum L.) and relatives. PLoS ONE 7:e38702

31. Huang LK, Bughrara SS, Zhang XQ, 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

32. Aliscioni SS, Giussani LM, Zuloaga FO, Kellogg EA (2003) A molecular phylogeny of Panicum (Poaceae: Paniceae): tests of monophyly and phylogenetic placement within the Panicoideae. Am J Bot 90:796–821

33. 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

34. Luo MC, Gu YQ, You FM, Deal KR, Ma Y, Hu Y, Huo N, Wang Y, Wang J, Chen S, Jorgensen CM, Zhang Y, McGuire PE, Pasternak S, Stein JC, Ware D, Kramer M, McCombie WR, Kianian SF, Martis MM, Mayer KFX, Sehgal SK, Li W, Gill BS, Bevan MW, Šimková H, Doležel J, Weining S, Lazo GR, Anderson OD, Dvorak J (2013) A 4-gigabase physical map unlocks the structure and evolution of the complex genome of Aegilops tauschii, the wheat D-genome progenitor. Proc Natl Acad Sci 110:7940–7946

35. Langley CH, Montgomery EA, Hudson R, Kaplan N, Charlesworth B (1988) On the role of unequal exchange in the containment of transposable element copy number. GenetRes 52:223–235

36. Biémont C, Tsitrone A, Vieira C, Hoogland C (1997) Transposable element distributions in Drosophila. Genetics 147:1997–1999