Genetic mapping with allele dosage information in tetraploid *Urochloa decumbens* (Stapf) R.D. Webster reveals insights into spittlebug (*Notozulia entreriana* Berg) resistance

Rebecca Caroline Ulbricht Ferreira*, Letícia Aparecida de Castro Lara[†], Lucimara Chiari[‡], Sanzio Carvalho Lima Barrios[‡]; Cacilda Borges do Valle[‡], José Raul Valério[‡], Fabrícia Zimermann Vilela Torres[‡], Antonio Augusto Franco Garcia[‡]; Anete Pereira de Souza[*,§]

[*] Center for Molecular Biology and Genetic Engineering (CBMEG), University of Campinas (UNICAMP), 13083-875, Campinas, SP, Brazil

[†] Genetics Department, ESALQ, University of São Paulo (USP), 13418-900, Piracicaba, SP, Brazil

[‡] Embrapa Beef Cattle, Brazilian Agricultural Research Corporation, 79106-550, Campo Grande, MS, Brazil

[§] Plant Biology Department, Biology Institute, UNICAMP, 13083-862, Campinas, SP, Brazil

The GBS sequence data files supporting the conclusions of this article are in the process of submission to the NCBI Sequence Read Archive (SRA), with accession number SRP092493

Short running title

Genetic mapping in *Urochloa decumbens*

Corresponding author

Anete Pereira de Souza[*,§]

Plant Biology Department, Biology Institute, UNICAMP, 13083-862, Campinas, SP, Brazil

Telephone number: +55 19-3521-1132

Fax number: +55 19-3521-1089

E-mail: anete@unicamp.br

Abstract

Urochloa decumbens (Stapf) R.D. Webster is one of the most important forage grasses in Brazilian beef production; therefore, genomic studies such as constructing molecular maps and defining the genetic architecture of relevant agronomic traits, are essential to understanding *U. decumbens* genetic structure and advancing breeding programs. The study objectives were to (i) obtain an integrated *U. decumbens* genetic map using molecular markers with allele dosage information, (ii) detect QTLs for spittlebug (*Notozulia entreriana*) resistance, and (iii) seek putative candidate genes involved in resistance/defense against pathogens. We used five pseudo-genome references to align GBS reads and selected 4,240 high-quality SNP markers identified in the *Setaria viridis* genome. Of these markers, 1,000 were distributed throughout nine homologous groups with a cumulative map length of 1,335.09 cM and an average marker density of 1.33 cM. The map contains SNP markers with different configurations for a tetraploid species. For the first time, this study detected QTLs for resistance to spittlebug, an

important pasture insect pest, explaining between 4.66% and 6.24% of the variation. These QTLs had markers linked to putative candidate genes related to resistance/defense against pathogens. Because this is the first *U. decumbens* map to include SNP dosage data and QTL detection, it will be useful for future evolutionary studies, genome assembly, and other quantitative trait locus analyses in *Urochloa* spp. Moreover, the results might facilitate the isolation of spittlebug-related genes and help clarify the mechanism of spittlebug resistance. These approaches will improve selection efficiency and accuracy in *U. decumbens* molecular breeding.

Keywords: Allele dosage, Linkage map, Polyploidy, Signalgrass, SNPs, Quantitative traits.

Introduction

Brazilian cultivated pastures are the basis for the production of beef cattle, and they cover extensive areas, most of which are populated by grasses of the genus *Urochloa* (syn. *Brachiaria*) (Jank *et al.* 2011; ABIEC 2016). This genus belongs to the Poaceae family and comprises species of plants distributed in tropical and subtropical regions, mainly in the African continent (Renvoize *et al.* 1996; Valle *et al.* 2009).

One of the most widely cultivated species of genus is *Urochloa decumbens* (Stapf) R.D. Webster, also known as signalgrass. This forage grass has relevant agronomic attributes, including exceptional adaptation to the poor and acidic soils that are typical of the tropics, leading to good animal performance (Valle *et al.* 2010). However, the species is susceptible to several types of spittlebug, including *Notozula entreriana* Berg (Hemiptera: Cercopidae), the most damaging pest for tropical pastures (Gusmão *et al.* 2016).

Spittlebugs reduce biomass production, decrease palatability, and reduce the carrying capacity of pastures (Valério and Nakano 1988). Therefore, breeding programs aim to develop new cultivars with reasonable or high resistance to spittlebugs to reduce forage loss. To achieve this advance, it is essential to understand the genetic mechanisms involved in the spittlebug resistance response in forage grass and to identify valuable markers for selection in breeding schemes. Currently, it is known that spittlebug resistance in *Brachiaria* grasses is under relatively simple genetic control and is highly heritable (Miles *et al*. 1995). Therefore, despite being a trait with little genetic information, spittlebug resistance can probably be easily manipulated in a plant breeding program.

Genetic breeding of *U. decumbens* is recent and has proven challenging because this grass is predominantly tetraploid (2n = 4x = 36) and apomictic (Naumova *et al.* 1999; do Valle *et al.* 2008). In this scenario, the number of currently available molecular genetic resources is limited. Previous molecular studies have increased knowledge of *U. decumbens* genetics, including the development of sets of microsatellite markers (Ferreira *et al.* 2016; Souza *et al*. 2018) and the first transcriptome (Salgado *et al.* 2017) of the species as well as linkage maps from the interspecific progeny of *U. decumbens* x *U. ruziziensis* (Worthington *et al.* 2016). Other studies with molecular markers have analyzed the genetic relationships of this species to other species in the genus *Urochloa* (Almeida *et al.* 2011; Pessoa-Filho *et al.* 2017; Triviño *et al.* 2017).

Until recently, it was not possible to perform intraspecific crossings due to the ploidy barriers between apomictic (polyploid) and sexual (diploid) accessions, but this restriction changed when diploid accessions were artificially tetraploidized using colchicine (Simioni and Valle 2009). This advance allowed the generation of a base

population and the exploration of a genetic variability previously conserved by apomixis, favoring new molecular studies, such as construction of specific genetic maps, to increase genetic knowledge of the species.

Genetic maps are essential tools for genetics and genomics research, and they are becoming increasingly common in polyploid species, mainly due to the development of next-generation-sequencing (NGS) technologies and advances in genetic and statistical methods (Garcia *et al.* 2013). In this investigation, we utilized the genotyping-by-sequencing (GBS) technique, which allows the discovery of hundreds of thousands of SNPs, abundant markers in the plant genome, at low cost (Elshire *et al.* 2011; Poland *et al.* 2012). Allied to this, robust software has been developed for polyploid organisms, such as programs that can estimate the allele dosage of SNPs and provide more genetic information to generate linkage maps with very high marker densities (Serang *et al.* 2012; Garcia *et al.* 2013; Mollinari and Serang 2015; Bourke *et al.* 2016).

The mapping process is more complex in polyploid species, mainly because of the larger number of possible genotypes (Ripol *et al.* 1999; Vigna *et al.* 2016). Most polyploid genetic maps have been limited to the use of single-dose markers (SDMs). In this case, either each marker is considered as a single-allele copy from only one of the parents of the cross, with a 1:1 segregation ratio, or the SDMs in both parents segregate in a 3:1 ratio (Wu *et al.* 1992). This method has been developed and successfully adopted for various forage grass species, such as *Panicum maximum* (Ebina *et al.* 2005), *Paspalum notatum* (Stein *et al.* 2007), *Panicum virgatum* (Okada *et al.* 2010), and *Urochloa humidicola* (Vigna *et al.* 2016). However, the use of only SDMs represents a loss of genetic information, because in tetraploid species, such as *U. decumbens*, five possible dosage classes exist: nulliplex (aaaa), simplex (Aaaa), duplex (AAaa), triplex

(AAAa), and quadruplex (AAAA). Thus, many other marker segregations can be considered. Knowledge of the dosage of a SNP is essential for genetic studies in polyploid species and can significantly increase the information imparted by each locus, providing greater genome coverage (Voorrips *et al.* 2011; Garcia *et al.* 2013; Bourke *et al.* 2016).

High-resolution genetic linkage mapping for polyploid species can identify beneficial trait loci and allow genome-based breeding (Shirasawa *et al.* 2017). Molecular markers have been widely used to locate quantitative trait loci (QTL) associated with quantitative resistance to insects in many crop plants; however, genomic loci related to resistance to spittlebugs have not been determined in any plant species. Particularly in *U. decumbens*, the identification of loci involved in this trait is a promising tool for characterizing the genetic architecture and can assist in the design of strategies to be introduced into breeding programs to increase the efficiency of the selection processes and accelerate the release of new cultivars (Valle *et al.* 2009).

This study is the first to report on the development and application of GBS for mapping studies in the intraspecific progeny of *U. decumbens*. To the best of our knowledge, there have been no reports of QTL mapping in signalgrass. Our goals were to (i) construct the first GBS-based integrated genetic map using tetraploid allele dosage information in a bi-parental progeny, (ii) identify QTLs related to spittlebug resistance on the integrated genetic map, and (iii) search for putative candidate genes that may be involved in spittlebug resistance.

Materials and methods

Plant material and DNA extraction

At the Brazilian Agricultural Research Corporation (Embrapa) Beef Cattle (EBC), Campo Grande/MS, an intraspecific cross was made between *U. decumbens* D24/27 (sexual accession tetraploidized by colchicine) and *U. decumbens* cv. Basilisk (tetraploid, apomictic cultivar used as the pollen donor); therefore, both parents were tetraploid. The full-sib progeny of 239 F_1 individuals was analyzed, from which 217 hybrids were identified using SSR markers obtained from Ferreira *et al*. (2016).

Leaf samples from each hybrid and both parents were collected, and genomic DNA was extracted using the DNeasy 96 Plant Kit (Qiagen GMbH, Germany). DNA concentrations were determined using a Qubit fluorometer (Invitrogen, Carlsbad, CA).

GBS library construction and sequencing

The GBS library was prepared at the University of Campinas following the protocol described by Elshire *et al.* (2011). Samples from both parents of the progeny were replicated five times for sequencing. Each individual within a library was part of a 96-plex reaction. Each DNA sample (300 ng in a volume of 10 μL) was digested with the restriction enzyme NsiI (NEB) and then ligated to a unique barcoded adapter plus a common adapter. The 96-plex libraries were sequenced as 150-bp single-end reads on the Illumina NextSeq 500 platform. Finally, the quality of the generated sequencing data was evaluated (Patel and Jain 2012).

SNP calling

SNP discovery and genotype calling were performed with the Tassel-GBS pipeline (Glaubitz *et al.* 2014), which was modified to obtain the original count of the

number of reads for each SNP allele (Pereira *et al*. 2017). Because this pipeline requires a reference genome, and the *U. decumbens* genome has not yet been sequenced, we aligned the GBS tags using five alternative pseudo-genomes: (i) the *Panicum hallii* genome (v 2.0; ~554 Mb arranged in 9 chromosomes and 8,405 scaffolds; diploid forage), (ii) the *Panicum virgatum* genome (v 1.0; ~1,230 Mb arranged in 18 chromosomes and 220,646 contigs; tetraploid forage), (iii) the *Setaria viridis* genome (v 1.0; ~394.9 Mb arranged in 9 chromosomes and 724 scaffolds; diploid forage), (iv) the *Setaria italica* genome (v 2.2; ~405 Mb arranged in 336 scaffolds; diploid forage), and (v) the transcriptome of *U. decumbens* roots exposed to aluminum and obtained by Salgado *et al.* (2017) (164,920 sequences). These genomes are available from the Phytozome website (http://www.phytozome.net/) (Goodstein *et al.* 2012). *S. viridis*, *P. hallii* and *P. virgatum* sequences data were produced by the US Department of Energy Joint Genome Institute. The Bowtie2 algorithm (Langmead and Salzberg 2012) was used to align tags against each reference with -D and -R parameters defined as 20 and 4, respectively, and with the very-sensitive-local argument.

Allele dosage estimation and data filtering

SuperMASSA software (Serang *et al.* 2012; Pereira *et al*. 2017) was used to estimate the correct tetraploid allele dosage (number of copies of each allele) of each individual. The minimum overall depth considered was 25 reads, and the model used was the F_1 Population Model. Markers were fitted and filtered to ploidy 4. Posteriorly, markers were selected manually (using R software) with up to 25% missing data and monomorphic markers were removed. A circular plot was used to summarize the redundant loci within and between references using the R *circlize* package (Gu *et al*. 2014).

Linkage map construction

The linkage map was constructed using the TetraploidSNPMap (TPM, BioSS) software (Hackett *et al.* 2017), following the methodology described in Hackett *et al*. (2013, 2014). Linkage groups were constructed separately using the SNPs identified on each relative *S. viridis* chromosome. Markers with significance of the χ^2 goodness-of-fit statistic less than 0.001 for simplex and 0.01 for higher segregations were considered distorted and were removed. A total of 2,745 distorted SNPs were excluded from the cluster analyses. The remaining 1,515 non-distorted SNPs were clustered into 9 homologous groups using a Chi-square test for independent segregation as described by Hackett *et al*. (2013).

Considering the ordination of markers, a 2-point analysis was used to calculate the recombination frequency and LOD score for each pair of SNPs in each possible phase by an expectation-maximization (EM) algorithm to maximize the likelihood and to infer the most likely phase as the one with the highest likelihood among those with a recombination frequency < 0.5. Duplicate and near-duplicate SNPs were removed in this step. The recombination frequencies were converted to map distances using Haldane's mapping function, and a multi-dimensional scaling analysis (MDS) was then performed to calculate the best order for the SNPs in the linkage group. After this step, we excluded some SNPs that showed low LOD scores or that were distant from the rest of the linkage group (Hackett *et al.* 2017).

Some phases of the ordered SNPs were inferred by the automated phase analysis in TetraploidSNPMap and completed manually based on the most likely phase for previous informative pairs prior to carrying out QTL analysis.

A genetic linkage map was drawn using MapChart 2.32 (Voorrips 2002), and SNP configurations were identified with different colors.

Phenotypic evaluation

The resistance of *U. decumbens* to spittlebugs was evaluated in greenhouse experiments at Embrapa Beef Cattle, Brazil, according to the methodology described by Lapointe *et al.* (1992) and Valério *et al.* (1997). From 2011 to 2013, a total of 12 experiments were conducted; each experiment had a randomized complete block design (RCBD) with common treatments repeated in all experiments and used as checks in the statistical analysis. A total of 349 individuals were evaluated for spittlebug resistance, of which 157 were hybrids of the mapping population.

During 2011 and 2012, five RCBD experiments were conducted with ten replicates and 114 hybrids evaluated. The best hybrids were selected and included in the experiments conducted in 2012 and 2013, when seven RCBD experiments were conducted, with four replicates and 259 individuals evaluated. The common checks for all experiments were *U. decumbens* cv. Basilisk, *U. brizantha* cv. Marandu, *Urochloa* spp. cv. BRS Ipyporã and *Urochloa* spp. MulatoII. The variable evaluated was the percentage of nymphal survival (number of nymphs that reached the adult stage, as a percentage). The phenotypic model is:

$$y_{ijlk} = \mu + t_i + b_{j(l)} + p_l + a_k + ta_{ik} + \varepsilon_{ijlk}$$

where y_{ijlk} is the phenotype of the *i*-th individual, at *j*-th block, *l*-th experiment, and *k*-th year; μ is the overall mean; t_i is the random effect of the *i*-th individual ($t_i \sim N(0, \sigma_T^2)$); $b_{j(l)}$ is the random effect of the *j*-th block within the *l*-th experiment ($b_{j(l)} \sim N(0, \sigma_B^2)$); p_l is the random effect of the *l*-th experiment ($p_l \sim N(0, \sigma_P^2)$); a_k is the fixed effect of the *k*-th year; ta_{ik} is the random effect of the *i*-th individual by *k*-th year interaction ($ta_{ik} \sim N(0, \sigma_{TA}^2)$); and ε_{ijlk} is the residual error ($\varepsilon_{ijlk} \sim N(0, \sigma_E^2)$).

The analysis was performed in the R package ASReml (Butler *et al.* 2009), and the heritability between genotype means was calculated using the index proposed by Cullis *et al.* (2006):

$$H_C^2 = 1 - \frac{PEV}{2\sigma_G^2}$$

where PEV is the prediction error variance, and σ_G^2 is the genetic variance.

QTL mapping

The set adjusted phenotypic means for the trait were used for the QTL mapping, which was performed with TetraploidSNPMap software (Hacket *et al*. 2017). The QTL analysis was performed for each linkage group separately using the SNP data for each linkage group, the integrated linkage map with phase information and the phenotypic trait. The mapping method was applied using Interval mapping (IM) based on tetraploid allele dosage information. In brief, the initial interval mapping was fit to a model of additive effects using a weighted regression approach. The results were displayed as a LOD profile along the chromosome and included the summary statistics on the LOD score, the percentage variation explained and the QTL effects for each homologous chromosome. A significant QTL was declared if the LOD score was above the 95% threshold obtained from a permutation test with 1000 permutations. From a significant QTL, simple models such as a simplex, duplex or double-simplex QTL models were then tested using the Schwarz Information Criteria (SIC) (Schwarz 1978) to determine if the estimated genotype corresponded to the most likely QTL location. The SIC was calculated in TPM as follows:

$$\text{SIC} = -2 \log L + p \log m_o,$$

where L is the likelihood for the simple model, p is the number of parameters in the simple model, and m_o is the number of observations (the 36 genotype means). Models with the lowest value for the SIC were considered the best models (Hackett *et al.* 2014).

If two closely linked significant QTLs were identified on the same chromosome, TetraploidSNPMap could not estimate the individual QTL LOD and parameter values, yielding data only pertaining to the largest effect.

Search for potential genes integrated in detected QTL regions

We investigated the functional annotation of the potential genes adjacent to the SNP markers linked to the mapped QTLs (Figure 3; Table S2) through comparative analysis with the *S. viridis* genome. First, we selected *S. viridis* genome 2.1 and determined the regions of approximately 5 kb spanning the markers adjacent to the mapped QTLs (Table S2). Then, we performed protein alignment using BLASTX with an e-value cutoff of 1e-0.5 using the JBrowse tool (http://www.phytozome.org/), against the *Arabidopsis thaliana* and *Orysa sativa* databases. Finally, from this alignment, we selected only genes reported in the literature as associated with resistance/defense against pathogens that were present in *U. decumbens* transcriptome (Salgado et al. 2017).

Data availability

The GBS sequence data files supporting the conclusions of this article are in the process of submission to the NCBI Sequence Read Archive (SRA), http://www.ncbi.nlm.nih.gov/sra, with accession number SRP148665.

Table S1 - Bowtie2 alignment results.

Table S2 - Functional description of mapped QTLs.

Figure S1 - Histogram of the distance between adjacent markers.

Figure S2 - Interval mapping (IM) for spittlebug resistance.

Results

SNP calling

A total of 1,183,089,925 sequence reads were obtained from four NextSeq 500 lanes for the *U. decumbens* mapping population, of which 94% contained high-quality bases (phred score 20). The alignment results using *P. virgatum*, *P. hallii*, *S. viridis*, *S. italica*, and the transcriptome of *U. decumbens* showed that the overall alignment ranged from 11.84% to 18.96% of the tags. The transcriptome of *U. decumbens* was the pseudo-reference genome that resulted in the highest overall alignment rate and the highest unique alignment rate (Table S1).

Considering all pseudo-references, 262,449 markers were obtained by the Tassel-GBS pipeline with a minimum minor allele frequency of 0.01 and minimum minor allele count of 1000. The alignment that identified the most markers in this step was the *P. virgatum* pseudo-reference genome, while the alignment with the transcriptome of *U. decumbens* identified the fewest SNPs (Table 1).

The output of the Tassel-GBS pipeline was used as the input for the SuperMASSA software, using the F_1 Population Model and fitting and filtering for ploidy 4. A total of 41,470 markers were selected with a minimum overall depth of 25 reads (Table 1).

Table 1. The number of markers obtained with the Tassel-GBS pipeline, estimation of allele dosage with SuperMASSA software and data filtration analysis.

Reference genomes	Tassel-GBS	SuperMASSA	Excluded missing data and filtered polymorphic sites
Panicum virgatum	60,683	8,303	3,501
Panicum hallii	49,603	7,748	3,153
Setaria viridis	58,370	9,345	4,240
Setaria italica	56,713	8,960	4,049
Transcriptome	37,080	7,114	3,043
Total	262,449	41,470	17,986

Using three different SNP markers, Figure 1 shows how the SuperMASSA software uses the ratio of the intensity of two alleles to classify individuals according to their genotypes using a probabilistic graphical model (Serang *et al.* 2012). The markers were named with the allele dosage segregation, the chromosome number, and the position of the SNP in the chromosome. In this step, the *S. viridis* alignment continued to reveal the highest number of markers, followed by that of *S. italica*. Afterward, markers with up to 25% missing data and monomorphic markers were removed. The final number of markers was 17,986, including all the pseudo-reference genomes (Table 1). The highest number of markers, 4,240, was obtained by alignment with the *S. viridis* genome (Table 1). These markers were used in the subsequent steps.

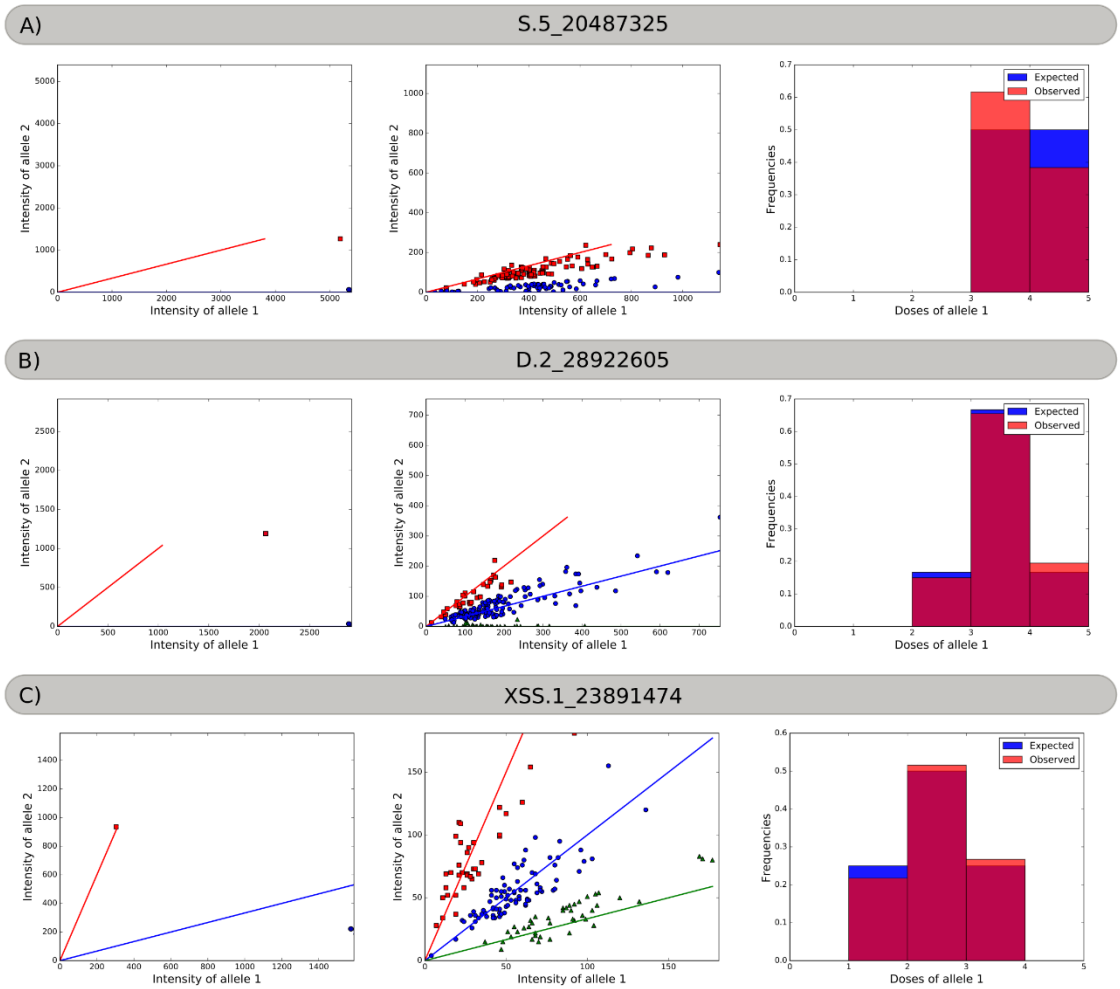

Figure 1. Allele dosage in the parents and progeny, and the frequency histograms. A) Marker S.5_20487325. Red squares represent the Aaaa parent and offspring, and blue circles represent the aaaa parent and offspring. B) Marker D.2_28922605. Red squares represent the AAaa parent and offspring, blue circles represent the aaaa parent and the Aaaa offspring, and green triangles represent the aaaa offspring. C) Marker XSS.1_23891474. Red squares represent AAAa parent and offspring, blue circles represent the Aaaa parent and the AAaa offspring, and green triangles represent the Aaaa offspring.

The complete redundancy among markers was inspected after quality and ploidy filtration (Figure 2). The five pseudo-references (*Panicum hallii* genome, *P. virgatum*

genome, *Setaria viridis* genome, *S. italica* genome, and *U. decumbens* transcriptome) showed some similarities, especially between *S. italica* and *S. viridis*. This result is expected due to the phylogenetic proximity of these species. The highest number of markers classified as unique were found for the *Panicum* genomes and *U. decumbens* transcriptome pseudo-references. These species remain largely undomesticated, and this fact is reflected in their greater genomic variability, favoring breeding programs.

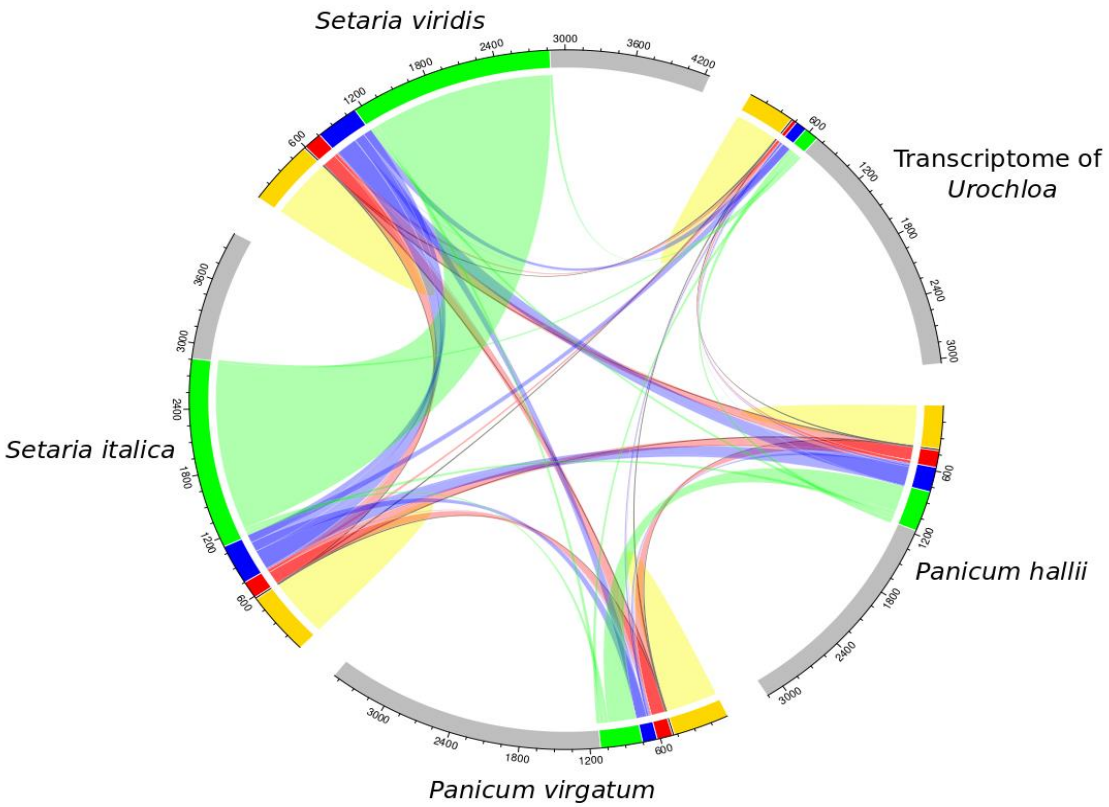

Figure 2. Circular plot showing the amount of complete redundancy among the markers (Gu *et al.* 2014) in the five pseudo-references (*Panicum hallii* genome, *P. virgatum* genome, *Setaria viridis* genome, *S. italica* genome, and *Urochloa decumbens* transcriptome) that were used to align the GBS signalgrass tags. The yellow regions represent redundancy within each pseudo-reference, whereas the black, red, blue, and green regions represent redundancy between five, four, three, and two pseudo-references,

respectively. The remaining grey regions represent the markers with unique information in each reference.

Genetic map

An integrated genetic map was constructed using 217 full sibs obtained from a cross between *U. decumbens* cv. Basilisk (apomictic cultivar) and *U. decumbens* D24/27 (sexual accession). Of the 4,240 markers used for linkage analysis, 1,515 showed significance of the χ^2 goodness-of-fit statistic, and 1,000 were placed in the linkage map (Table 2).

The markers were distributed throughout nine chromosomes (homology groups, HGs), with a cumulative map length of 1,335.09 cM and an average marker density of 1.33 cM. The length of each group ranged from 129.70 cM (HG7) to 166.57 cM (HG9), with an average of 148 cM (Table 2; Figure 3). The number of markers varied from 63 to 187 among the nine homologous groups (Table 2). The distances between adjacent markers on the genetic map were plotted as a histogram (Figure S1). The majority of the markers are within 1 cM of each other. The two largest distances were located in HG8 and had lengths of 11.2 cM and 12.4 cM. HG9 was the most saturated and the one with the shortest distances between adjacent markers (Figure 3).

Table 2. The number of GBS-based markers after significance testing, the number of mapped markers within each homology group (HG), and the length and marker density of each HG (cM) in the genetic map.

HG	No. GBS-based markers*	No. mapped markers	Length of HG (cM)	Marker density (cM)
1	132	100	136.74	1.37
2	181	121	164.95	1.36

3	219	138	151.56	1.10
4	169	101	139.70	1.38
5	183	123	129.70	1.05
6	100	63	161.72	2.57
7	154	99	140.08	1.41
8	82	68	144.07	2.12
9	295	187	166.57	0.89
Total	1,515	1,000	1,335.09	1.33

*Markers were selected if the χ^2 goodness-of-fit statistic was higher than 0.001 for simplex and 0.01 for higher segregations.

In Figure 3, the possible genotype configuration of each marker is indicated by the marker name prefix and the different colors. For example, markers with the prefix S and the color black represent the simplex configuration (Table 3). Four homologous groups showed all configurations (HG1, HG3, HG5 and HG9), and simplex markers (S) were the most frequent. The X-double-simplex (XSS) configuration was less frequent and appeared in only HG3 and HG5 (Table 3; Figure 3).

Table 3. Distribution of SNPs into genotype classes.

Type	**Parent 1**	**Parent 2**	**P1 dosage**	**P2 dosage**	**SNP initial number**	**Mapped SNP**
Null (N)	AAAA	BBBB	4	0	8	0
Simplex (S)	AAAA	AAAB	4	3	1,332	591
Simplex (S)	ABBB	BBBB	1	0		
Triplex (T)	AAAA	ABBB	4	1	126	0
Triplex (T)	AAAB	BBBB	3	0		
Duplex (D)	AAAA	AABB	4	2	720	172

Duplex (D)	AABB	BBBB	2	0		
Double-simplex (SS)	AAAB	AAAB	3	3		
Double-simplex (SS)	ABBB	ABBB	1	1	1,013	78
X-double-simplex (XSS)	AAAB	ABBB	3	1	65	8
Simplex-duplex (SD)	AAAB	AABB	3	2	11	0
Duplex-simplex (DS)	AABB	ABBB	2	1	773	116
Double-duplex (DD)	AABB	AABB	2	2	192	35
Total					4,240	1,000

QTL mapping

Based on the phenotypic data, a multi-allelic QTL analysis was performed for resistance to spittlebugs (*Notozulia entreriana*) by applying an IM model to the integrated genetic map. Thus, we detected three significant QTLs for the studied trait on homologous groups 1, 2, and 6. The male genitor of the mapping population, *U. decumbens* D62 (cv. Basilisk), contributed to the resistance alleles for these QTLs. The statistical analysis showed that the heritability between genotype means for the trait was high (H_C^2 = 0.79).

On chromosome 1, the QTL had a LOD score of 4.68 and explained 4.80 of the trait variation for resistance to spittlebugs. This LOD score was above the upper 95% LOD permutations threshold of 3.11, and this first QTL was located at position 95 cM (Figure S2). Analyses of different simple genetic models were performed with TetraploidSNPMap to determine the best simple fitting model for the studied trait. For this QTL, the best model was a duplex genotype (AAAA x BBAA), with the allele B

present on parent 2 associated with spittlebug resistance. This model had the lowest SIC (SIC = 85.34) compared with the full model (SIC = 102.99) (Table 4).

On chromosome 2, the maximum LOD score was 4.39, and this QTL explained 4.66 of the phenotypic variation for spittlebug resistance. The LOD peak was located at position 60 cM, and its score was above the upper 95% LOD permutation threshold of 3.29 (Figure S3). Again, the analysis with the simpler models estimated a duplex genotype (AAAA x BBAA), with the resistance allele present on parent 2. This model had the lowest SIC (SIC = 93.03) in comparison with the full model (SIC = 89.55) (Table 4).

The last detected QTL on chromosome 6 had the maximum LOD score of 5.49, explained 6.24 of the trait variation and was above the 95% LOD permutation upper threshold of 2.79 (Figure S4). The QTL peak was located at 23 cM, and an analysis of different genetic models indicated that an allele explains the trait variation. This QTL is linked to a simplex SNP (AAAA x ABAA), with the B allele present on parent 2 associated with spittlebug resistance. This model had the lowest SIC of 99.07 compared to the full model with SIC of 115.47 (Table 4).

Table 4. QTL information for the spittlebug resistance trait analyzed in the *U. decumbens* progeny.

Trait	Chr	QTL Position (cM)	LOD	Variance Explained (%)	SIC	Configuration	Homologous Chr
Spittlebug	1	95	4.68	4.80	85.33	Duplex (D)	h5 + h6
Spittlebug	2	60	4.39	4.66	93.03	Duplex (D)	h5 + h6
Spittlebug	6	23	5.49	6.24	99.07	Simplex (S)	h6

Chr, chromosome; QTL, quantitative trait loci; LOD, logarithm of the odds; SIC, Schwarz Information Criterion.

Search for potential resistance genes

We found sequence similarity in the regions of all adjacent markers of the three mapped QTLs, with homologies for *Arabidopsis thaliana* and *Oryza sativa*. Functional annotation of the sequences showed possible candidate genes involved in the spittlebug resistance response (Table S2) that were transcribed in *U. decumbens* (Salgado et al. 2017).

The first identified QTL, located in HG1, had two adjacent markers that were each identified in the *S. viridis* genome reference. The region of the GBS-based marker D.1_40431816 showed similarity to pathogenesis-related gene 1 (*PR1*). The region of the second adjacent GBS-based marker of this QTL, S.1_21570712, showed similarity to the leucine-rich repeat (LRR) family protein and the *WRKY11* gene, which is involved in pathogen defense (Table S2).

The regions of the GBS-based adjacent markers to QTL2, DS.2_39902899 and S.2_399022936, showed similarity to NBS-LRR and NB-ARC protein domains and with the pentatricopeptide repeat (PPR) superfamily proteins (Table S2).

QTL3, located in HG6, also had two adjacent markers that were each identified in the *S. viridis* genome reference. The region of the GBS-based marker D.6_25761630 showed similarity to a WRKY DNA-binding protein, a kinase superfamily protein, and an F-box family protein. The region of the other adjacent GBS-based marker, S.6_35357677, showed similarity to NB-ARC protein domains, kinase superfamily proteins and F-box family proteins. This region also showed similarity to the thaumatin-like protein superfamily (TLPs), which are classified within the pathogenesis-related (PR) protein family. Both regions of these markers adjacent to QTL3 showed similarity to LRR protein domains (Table S2).

Discussion

The development of GBS technology allows the identification of thousands of SNP markers across the genome at a relatively low cost. This advance represents an evolution in the genetic studies of several species, since a larger part of the genome can be represented, favoring approaches such as the construction of genetic maps. The map reported here is the first map developed for an intraspecific progeny of *U. decumbens* that includes SNPs with tetraploid allele dosages, an approach never previously employed in a forage grass species. This map provides important information about the genome of this species and allows the detection of QTLs involved in spittlebug (*N. entreriana* Berg) resistance. QTL mapping is a useful tool to evaluate genotypes of putative relevance for breeding programs.

In addition, we reported the first application of GBS in an intraspecific progeny of *U. decumbens*. The total number of sequence reads obtained in this study using the GBS protocol of Elshire *et al.* (2011) with the NsiI (NEB) restriction enzyme was greater than that obtained in Worthington *et al.* (2016), in which the authors considered an interspecific progeny between *U. ruziziensis* and *U. decumbens* and used the GBS protocol of Heffelfinger *et al.* (2014) with the HincII restriction enzyme. However, both protocols have provided satisfactory results for the construction of robust genetic maps in complex polyploid species.

We used five pseudo-references from species genetically close to *U. decumbens* to discover GBS-based markers (Table 1). The highest final number of markers (4,240) was detected in the alignment with the genome of *S. viridis* (green foxtail) (Table 1), a member of the Poaceae, or grass, family. The reason that the transcriptome of *U. decumbens* did not show the greatest number may be that it covers only the markers found in the expressed regions of the genome, even though some of these may be exclusive to *U. decumbens*. In a previous analysis, we tested the pipeline UNEAK

(Universal Network-Enabled Analysis Kit) as implemented in the TASSEL-GBS software program (Glaubitz *et al.* 2014) for identifying SNPs, but this method detected fewer markers (1,210) than alignment with the pseudo-references did (Table 1). After filtering, Worthington *et al.* (2016) detected 3,912 SNP markers using the UNEAK pipeline, that is, a number of markers equivalent to that detected by us using a pseudo-reference genome (Table 1). Most likely, the cross between two different species (*U. decumbens* x *U. ruziziensis*) favored the discovery of more SNPs with UNEAK, in addition to the different methodology used for GBS library construction (Worthington *et al.* 2016).

Previous polyploid genetic maps have been developed with markers scored as presence/absence of alleles, but this method does not make use of all the information available in the SNP at multiple doses. The GBS data obtained here were sufficient in terms of read depth coverage to call allele dosage in 41,470 markers, considering all pseudo-references (Table 1). The SuperMASSA software provided the distribution of alleles in the progeny and the relative intensities of each allele (Serang *et al.* 2012; Garcia *et al.* 2013; Pereira *et al*. 2017).

Approximately 55% (2,345) of the markers identified followed the parental genotype configurations simplex (AAAA x AAAB, ABBB x BBBB) and double-simplex (AAAB x AAAB, ABBB x ABBB). Slightly less than 1% of the identified markers followed the nulliplex (AAAA x BBBB) and simplex-duplex (AAAB x AABB) configurations (Table 3). Our results were satisfactory, but a GBS library construction methodology with more than one restriction enzyme might provide a greater depth of coverage and, accordingly, increase the accuracy of the allele dosage estimation and the number of markers detected.

The map contains 1000 SNP markers distributed on nine homologous groups, with a total length of 1,335.09 cM and an average map density of 1.33 cM (Table 2; Figure 3). Three previous genetic maps have been published in the genus *Urochloa* (Thaikua *et al*. 2016; Vigna *et al*. 2016; Worthington *et al*. 2016). Vigna *et al.* (2016) constructed a genetic map for hexaploid *U. humidicola* using 89 SSR markers based on simplex and double-simplex markers. The map consisted of 49 linkage groups (LGs) and had a total length of 1,702.82 cM with an average map density of 10.6 cM. Our map is more saturated and informative than this previous map because we identified SNPs covering much of the genome and with allele dosage information. In another study, genetic maps for *U. decumbens* x *U. ruziziensis* were constructed with GBS-based markers (Worthington *et al.* 2016), and 1,916 SDA (single-dose allele) markers were distributed across the parental linkage maps. The genetic map developed in our study has a higher marker density than those of other genetic maps for tetraploid species (Stein *et al.* 2007; Rajput *et al.* 2016; Thaikua *et al.* 2016).

The density of markers per HG ranged from 63 (HG6) to 187 (HG9) (Table 2). The greater density of markers for certain HGs may correspond to a greater recombination frequency. On the other hand, less saturated groups may have fewer SNPs in these regions and/or correspond to highly homozygous regions that have lower recombination frequencies (Bai *et al.* 2016).

The largest distance between adjacent markers, 12.4 cM, was observed on chromosome eight (Figure 3), one of the chromosomes with the fewest mapped markers. These regions in the genetic map may be homozygous in both parents, and hence, no recombination can be detected in this case. Additionally, it is possible that the number of hybrids in the mapping population was not sufficient to observe recombination in these regions (Ma *et al.* 2014). One possibility to fill these distances is

to use different combinations of enzymes to enlarge the sequencing pools and thereby enable the capture of other genomic regions and obtain a density close to saturation. Furthermore, we mapped only markers identified in the alignment with the *S. viridis* genome. Therefore, these regions could contain markers that are exclusive to the *U. decumbens* genome and thus could not be represented with the methodology used.

Among the 1000 mapped SNPs, 59% (591) followed the simplex parental genotype configuration (Table 3; Figure 3). This configuration was also the most common in other studies (Hackett *et al.* 2013; da Silva *et al.* 2017). The least common configuration was X-double-simplex (XSS) (0.8%) (Table 3), followed by double-duplex (DD), which is the one that allows the representation of all types of segregation for a tetraploid species. Overall, the map displays most of the possible configurations, and these data significantly increase the information about each locus and provide several advantages for genetic analysis (Garcia *et al.* 2013). Therefore, the use of markers with allele dosage information allows the development of highly saturated and robust genetic maps facilitating the detection of significant QTLs (Hackett *et al.* 2013; da Silva *et al.* 2017).

The observed percentage of unlinked markers (34%) is greater than those obtained in other *Urochloa* maps (Vigna *et al.* 2016; Worthington *et al.* 2016), and this result may reflect the different methodologies used. An alternative for this question is to associate other types of molecular markers, such as microsatellites, since these are similar to anchors, linking previously unlinked SNPs (Ball et al. 2010). In addition, the association between these two types of markers may enable the generation of more robust maps (Ball *et al.* 2010). Moreover, we mapped the markers according to the *S. viridis* chromosomes, and translocation events may have occurred in signalgrass, altering the chromosome arrangement (Aitken *et al.* 2014). Therefore, advances in the

assembly of complex polyploid genomes are expected to enable the use of the full signalgrass genome as a reference for future studies (Margarido and Heckerman 2015).

The integrated genetic map associated with phenotypic data allow QTL mapping studies. Until recently, there was no intraspecific progeny available for *U. decumbens*; therefore, no QTL mapping studies have been performed to date. One of the most important agronomic traits to advance the *U. decumbens* breeding program is spittlebug (*N. entreriana*) resistance because spittlebugs severely damage this forage grass (Valle *et al.* 2009).

There is only one report of quantitative trait loci mapping for agronomic traits in *Urochloa* hybrids from the cross between cv. 'Mulato' and cv. 'Miyaokikoku'. In that report, QTLs for leaf width, leaf shape ratio (width/length), stem diameter and percentage of filled seeds using simple interval mapping and composite interval mapping were identified (Thaikua *et al.* 2016). In our study, QTL mapping was performed by applying the statistical model IM (Interval Mapping) for a tetraploid progeny (Hackett *et al.* 2017). There are more accurate models; however, these models consider the few possible segregations usually applied to diploids, and this approach represents a loss of genetic information in polyploid mapping studies (Hackett *et al.* 2014). In this study, we report the detection of QTLs for resistance to spittlebugs as well as all possible segregations for tetraploid *U. decumbens*.

The heritability for the spittlebug resistance trait was high (H_C^2 = 0.79), and this reflects a larger genetic effect that contributes to the total phenotypic variance. Thus, loci related to the trait can be detected more easily (Huang *et al*. 2007). For the multi-allelic QTL analysis, we performed a permutation test to obtain the threshold for declaring significant QTLs. Three QTLs were identified, and their LODs ranged from 4.39 to 5.49 (Table 4). The number of QTLs found was low, but they represent a great

gain since it is the first genetic architecture study with *U. decumbens* progeny. These QTLs were located in HGs 1, 2, and 6, at 95 cM, 60 cM and 23 cM, respectively (Figure 3). In HG1, the SNP D.1_40431816 was located at 0.6 cM from the first QTL. The second QTL was located in HG2 at 0.4 cM from the SNP DS.2_39902899. In HG6, the QTL was located between the SNPs D.6_25761630 and S.6_35357677, at 0.8 cM and 1.8 cM, respectively (Figure 3; Table 4). The regions that gave rise to these markers could be evaluated for future applications by signalgrass breeding programs.

Although *U. decumbens* is considered a susceptible species to the pasture spittlebug, these results show that a combination of different alleles in each parent segregates and contributes to the observed polymorphism in progeny, allowing the detection of significant QTLs. Mateus *et al*. (2015) estimated genetic parameters for resistance to spittlebugs in *U. decumbens* progenies and observed genetic variability among the hybrids for this trait, allowing the selection of superior hybrids. In addition, our results showed that it is possible to identify the alleles that contribute to the increased values of the phenotypic trait in the male parent. This allows genotypes that are less suitable for spittlebug development and survival to be found.

The percentage of phenotypic variance explained by each QTL ranged from 4.66 to 6.24 (Table 4). These percentages were relatively low, which indicates that no major loci were detected and the resistance to spittlebugs trait is regulated by some genes, each with a small effect. Miles et al. (1995) reported that spittlebug resistance in *Brachiaria* grasses is probably not genetically complex, but likely involves more than a single major resistance gene. The values of phenotypic variation and the small number of detected QTLs may be due to factors such as low genetic variability for spittlebug resistance in the progeny and the methodology used for QTL detection and/or map saturation.

The genome of *U. decumbens* has not been sequenced, and the genomic research is still at the beginning. Therefore, the detection of these QTLs is the first step in the advancement of genomic studies involving resistance to the spittlebugs. The annotation of sequences that originated markers with mapped QTLs is important to identify candidate genes involved in spittlebug resistance in *U. decumbens* and may contribute to new insights and new areas of research into these genes.

N. entreriana spittlebug nymphs constantly suck the sap of roots, causing yellowing of the plant, and the saliva of the adults induces phytotoxemia that causes plant death (Valério and Nakano 1992; Gusmão *et al*. 2016). Thus, to respond mechanical damage due to these insect attacks, plants possess multiple molecular defense mechanisms (War *et al*. 2012). The QTL regions of this study showed similarity to genes containing protein domains reported to be involved in disease resistance and defense responses to pathogens that attack plants (Table S2).

For the spittlebug resistance trait, we can highlight the similarity of the region of the marker D.1_40431816, adjacent to QTL1, with pathogenesis-related gene 1 (*PR1*). PR proteins are induced in response to attack by various types of pathogens, being that the *PR1* gene is involved in biological processes related to defense responses and systemic acquired resistance (Uknes *et al*. 1992; Metzler *et al*. 1991; van Loon *et al*. 1999; van Loon *et al*. 2006; Ali *et al*. 2018). The region of the marker S.1_21570712, also adjacent to QTL1, showed similarity to the *WRKY11* gene, which acts as a negative regulator of basal resistance to bacterial pathogens (Journot-Catalino *et al*. 2006) and plays a crucial role in cellular defense responses and defense-related gene transcriptional regulation (Jiang *et al*. 2016) (Table S2).

For QTL3, we can highlight the similarity of the region to the adjacent marker D.6_25761630 with WRKY DNA binding domains. These transcription factors are

involved in plant defense responses (Du and Chein 2000). Moreover, F-box and WRKY DNA binding domains have orthologs in *A. thaliana*, with genes that develop a function in plant defense. The second GBS-based marker of this QTL, S.6_35357677, showed similarity to TLPs, which play a variety of roles including plant protection against pathogen attacks (Liu et al. 2010; Mukherjee et al. 2010; Misra *et al*. 2016).

The regions of the markers adjacent to the mapped QTLs 2 and 3 showed similarity to NBS-LRR protein domains, which are widely distributed in plants and thought to respond to pathogen attacks, including viruses, bacteria, fungi, and even insects (Song *et al.* 2017). NBS-LRRs are the most prevalent class of resistance proteins (Gupta *et al*. 2012). F-box proteins form one of the largest multigene superfamilies and control many important biological functions such as pathogen resistance (Lechner *et al*. 2006), and the PPR superfamily proteins (Barkan and Small 2014) have been implicated in plant defense mechanisms. Moreover, the NB-ARC protein domain is a functional ATPase domain containing a nucleotide-binding site that is proposed to regulate plant disease resistance (van Ooijen *et al*. 2008).

These results demonstrate the potential of these QTLs to in fact be related to resistance to spittlebug and provide a starting point, indicating possible candidate genes for this trait in signalgrass. However further studies should be conducted to validate these genomic regions as well as the role of the candidate genes in signalgrass and their effects on phenotypic expression.

The sequences obtained by the GBS libraries are dynamic, allowing the raw data to be reanalyzed with the evolution of bioinformatics techniques (Poland and Rife 2012). The possibility of estimating the tetraploid dosage of each SNP greatly increases genetic information and represents an important evolution for polyploid genetic maps, including those of important crops (Bourke *et al.* 2016). Furthermore, the molecular

data developed here can be applied in future studies, such as evolutionary history, genome assembly, and other QTL analyses in *U. decumbens* and related species. These studies are the first step towards possible marker-assisted selection (MAS), a subject of great interest in genetic improvement programs for economically important plant species.

Figure 3. A linkage map for *U. decumbens*: homologous groups from 1 to 4. The genotype configuration of each marker is indicated by the marker name prefix and color (Table 3). The QTLs are identified in HG1 and HG2.

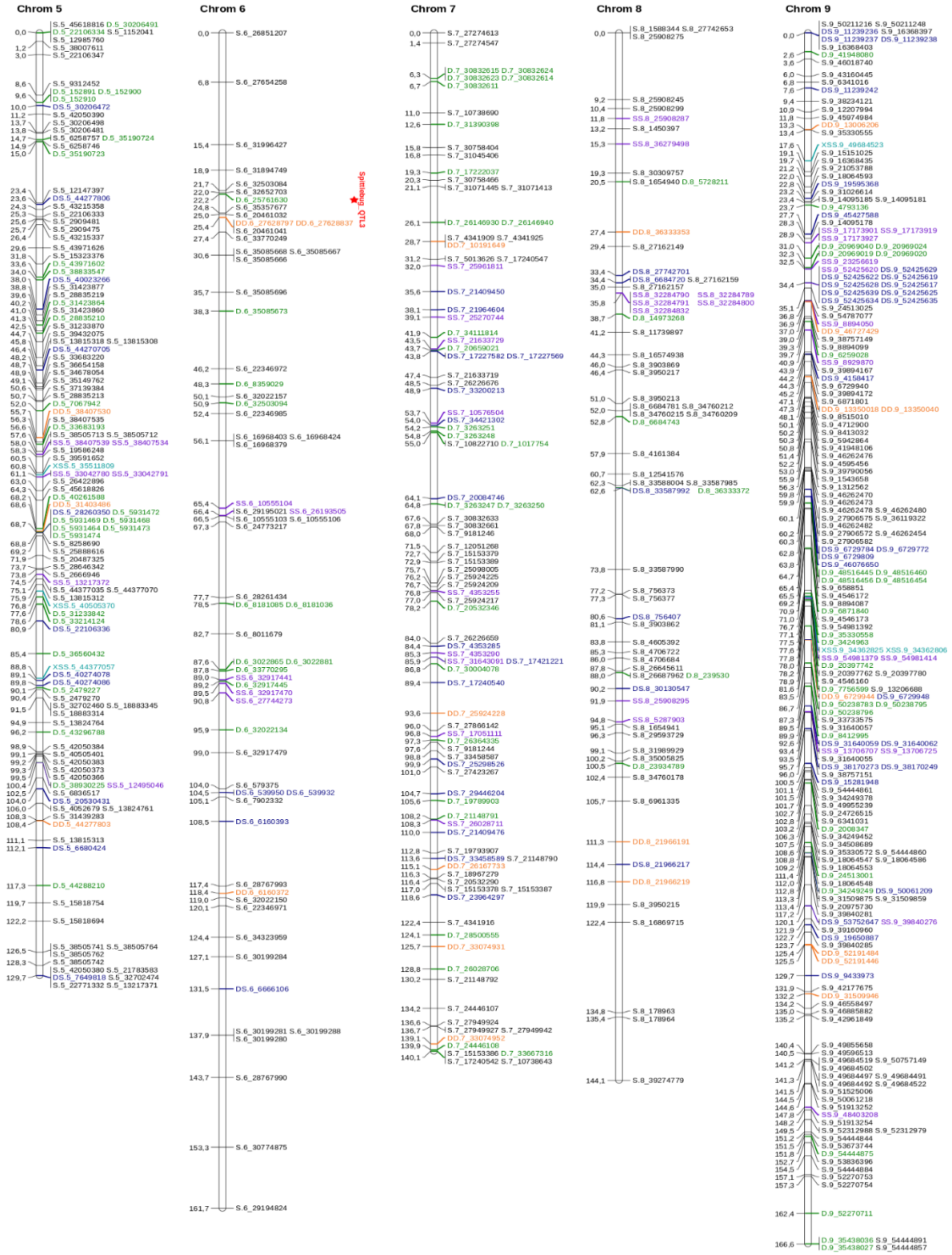

Figure 3, cont. Linkage map for *U. decumbens*: homologous groups from 5 to 9. The genotype configuration of each marker is indicated by the marker name prefix and color (Table 3). A QTL is identified in HG6.

Conclusions

This study presents the first genetic map in an *U. decumbens* intraspecific progeny using GBS-based SNPs with tetraploid allele dosage. These markers showed great potential for QTL mapping and allowed, for the first time, the detection of three QTLs involved in resistance to spittlebugs (*N. entreriana*), thereby making a valuable contribution to our understanding of the genetic basis underlying forage grass resistance to this pest. The verification of possible candidate genes for the mapped QTLs as a preliminary analysis provided relevant insights into relations between phenotypes and genotypes, complementing our results. The inclusion of microsatellite markers and map saturation are further steps to map other characteristics of interest, such as apomixis.

Acknowledgments

We thank the Brazilian Agricultural Research Corporation (EMBRAPA Beef Cattle) for providing the *Urochloa decumbens* progeny used in this study. We thank the Aline da Costa Lima Moraes for the assistance in the GBS library construction and performing DNA sequencing. This work was supported by grants from the Fundação de Amparo à Pesquisa de do Estado de São Paulo (FAPESP 08/52197-4) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES - Computational Biology Programme). RCUF received PhD fellowships from the CAPES-EMBRAPA Programme and the CAPES Computational Biology Programme, and LACL received a PD fellowship from the CAPES Computational Biology Programme. AAFG and APS were recipient of a research fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

Aitken, K. S., M. D. McNeil, P. J. Berkman, S. Hermann, A. Kilian *et al.*, 2014 Comparative mapping in the Poaceae family reveals translocations in the complex polyploid genome of sugarcane. BMC Plant Biol. 14: 190. DOI: 10.1186/s12870-014-0190-x.

Ali, S., Z. A. Mir, J. A. Bhat, A. Tyagi, N. Chandrashekar *et al*., 2018 Isolation and characterization of systemic acquired resistance marker gene PR1 and its promoter from *Brassica juncea*. 3Biotech 8: 10.1007/s13205-017-1027-8.

Almeida, M. CdC., L. Chiari, L. Jank, and C. Bd Valle, 2011 Diversidade genética molecular entre cultivares e híbridos de *Brachiaria* spp. e *Panicum maximum*. Cienc. Rural 41: 1998-2003.

Associação Brasileira das Indústrias Exportadoras de Carne (ABIEC), 2016 *Perfil da Pecuária do Brasil: Relatório anual.* [Online]. <http://abiec.siteoficial.ws/images/upload/sumario-pt-010217.pdf> Accessed 15 Nov 2017.

Bai, Z. Y., X. K. Han, X. J. Liu, Q. Q. Li, and J. L. Li, 2016 Construction of a high-density genetic map and QTL mapping for pearl quality-related traits in *Hyriopsis cumingii*. Sci. Rep. 6: 32608. DOI: 10.1038/srep32608.

Ball, A. D., J. Stapley, D. A. Dawson, T. R. Birkhead, T. Burke *et al.*, 2010 A comparison of SNPs and microsatellites as linkage mapping markers: Lessons from the

zebra finch (*Taeniopygia guttata*). BMC Genomics 11: 218. DOI: 10.1186/1471-2164-11-218.

Barkan, A., and I. Small 2014 Pentatricopeptide repeat proteins in plants. Annu. Rev. Plant Biol. 65: 415–442. doi: 10.1146/annurev-arplant-050213-040159.

Bourke, P. M., R. E. Voorrips, T. Kranenburg, J. Jansen, R. G. F. Visser *et al.*, 2016 Integrating haplotype-specific linkage maps in tetraploid species using SNP markers. Theor. Appl. Genet. 129: 2211-2226. DOI: 10.1007/s00122-016-2768-1.

Butler, D. G., Cullis, B. R., Gilmour A. R., and B. J. Gogel, 2009 *Mixed Models for S Language Environments: ASReml-R Reference Manual*, 149p.

Cullis, B. R., A. B. Smith, and N. E. Coombes, 2006 On the design of early generation variety trials with correlated data. J. Agric. Biol. Environ. Stat. 11: 381–393.

da Silva, W. L., J. Ingram, C. A. Hackett, J. J. Coombs, D. Douches *et al.*, 2017 Mapping loci that control tuber and foliar symptoms caused by PVY in Autotetraploid potato (*Solanum tuberosum* L.). G3 7: 3587–3595. DOI: 10.1534/g3.117.300264.

Du, L., and Z. Chen 2000 Identification of genes encoding receptorlike protein kinases as possible targets of pathogen- and salicylic acidinduced WRKY DNA-binding proteins in Arabidopsis. Plant. J. 24: 837–847. doi: 10.1046/j.1365-313x.2000.00923.x.

Ebina, M., H. Nakagawa, T. Yamamoto, H. Araya, S. Tsuruta *et al.*, 2005 Co-

segregation of AFLP and RAPD markers to apospory in Guineagrass (*Panicum maximum* Jacq.). Grassl. Sci. 51: 71–78. DOI: 10.1111/j.1744-697X.2005.00011.x.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto *et al.*, 2011 A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6: e19379. DOI: 10.1371/journal.pone.0019379.

Ferreira, R. C. U., L. J. Cançado, C. B. doValle, L. Chiari, and A. P. deSouza, 2016 Microsatellite loci for *Urochloa decumbens* (Stapf) R.D. Webster and cross-amplification in other *Urochloa* species. BMC Res. Notes 9: 152. DOI:10.1186/s13104-016-1967-9.

Garcia, A. A. F., M. Mollinari, T. G. Marconi, O. R. Serang, R. R. Silva *et al.*, 2013 SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. Sci. Rep. 3: 3399. DOI: 10.1038/srep03399.

Glaubitz, J. C., T. M. Casstevens, F. Lu, J. Harriman, R. J. Elshire *et al.*, 2014 TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS OnE 9: e90346. DOI: 10.1371/journal.pone.0090346.

Goodstein, D. M., S. Shu, R. Howson, R. Neupane, R. D. Hayes *et al.*, 2012 Phytozome: A comparative platform for green plant genomics. Nucleic Acids Res. 40: D1178–D1186. DOI: 10.1093/nar/gkr944.

Gu, Z., L. Gu, R. Eils, M. Schlesner, and B. Brors, 2014 Circlize implements and

enhances circular visualization in R. Bioinformatics 30: 2811–2812. DOI: 10.1093/bioinformatics/btu393.

Gupta, SK., A.K. Rai, S. S. Kanwar and T. R. Sharma 2012 Comparative analysis of zinc finger proteins involved in plant disease resistance. PLoS One 7: e42578. doi: 10.1371/journal.pone.0042578.

Gusmão, M. R., J. R. Valério, F. P. Matta, F. H. D. Souza, B. B. Z. Vigna *et al.*, 2016 Warm-season (C4) turfgrass genotypes resistant to spittlebugs (Hemiptera: Cercopidae). J. Econ. Entomol. 109: 1914–1921 DOI: 10.1093/jee/tow135.

Hackett, C. A., B. Boskamp, A. Vogogias, K. F. Preedy, and I. Milne, 2017 TetraploidSNPMap: Software for linkage analysis and QTL mapping in autotetraploid populations using SNP dosage data. J. Hered. 108: 438–442. DOI: 10.1093/jhered/esx022.

Hackett, C. A., J. E. Bradshaw, and G. J. Bryan, 2014 QTL mapping in autotetraploids using SNP dosage information. Theor. Appl. Genet. 127: 1885-1904. DOI: 10.1007/s00122-014-2347-2.

Hackett, C. A., K. McLean, and G. J. Bryan, 2013 Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. PLoS One 8: e63939. DOI: 10.1371/journal.pone.0063939.

Heffelfinger, C., C. A. Fragoso, M. A. Moreno, J. D. Overton, J. P. Mottinger *et al.*,

2014 Flexible and scalable genotyping-by-sequencing strategies for population studies. BMC Genomics 15: 979. DOI: 10.1186/1471-2164-15-979.

Huang, S., Ballard, D, and H. Zhao, 2007 The role of heritability in mapping expression quantitative trait loci. BMC Proceedings 1(Suppl 1): S86.

Jank, L., C. B. Valle, and R. M. S. Resende, 2011 Breeding tropical forages. Crop Breed. Appl. Biotechnol. 11: 27-34. DOI: 10.1590/S1984-70332011000500005.

Jiang, C., Z. Huang, P. Xie, C. Gu, K. Li et al., 2016 Transcription factors WRKY70 and WRKY11 served as regulators in rhizobacterium *Bacillus cereus* AR156-induced systemic resistance to *Pseudomonas syringae* pv. tomato DC3000 in Arabidopsis. J. Exp. Bot. 67: 157-174. doi:10.1093/jxb/erv445.

Journot-Catalino, N., I. E. Somssich, D. Roby and T, Kroj, 2006 The Transcription Factors WRKY11 and WRKY17 Act as Negative Regulators of Basal Resistance in Arabidopsis thaliana. The Plant Cell 18: 3289-3302.

Langmead, B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. Nat. Methods 9: 357–359. DOI: 10.1038/nmeth.1923.

Lapointe, S. L., M. S. Serrano, G. L. Arango, G. Sotelo, and F. Cordoba, 1992 Antibiosis to spittlebugs (Homoptera: Cercopidae) in accessions of *Brachiaria* spp. J. Econ. Entomol. 85: 1485-1490. DOI: 10.1093/jee/85.4.1485.

Lechner, E., P. Achard, A. Vansiri, T, Potuschak and P. Genschik 2006 F-box proteins everywhere. Curr. Opin. Plant Biol. 9:631–638.

Liu, J.J., R. Sturrock and A. K. Ekramoddoullah 2010 The superfamily of thaumatin-like-proteins: its origin, evolution, and expression towards biological function. Plant Cell Rep. 29: 419-436. doi: 10.1007/s00299-010-0826-8.

Ma, J. Q., M. Z. Yao, C. L. Ma, X. C. Wang, J. Q. Jin *et al.*, 2014 Construction of a SSR-based genetic map and identification of QTLs for catechins content in tea plant (*Camellia sinensis*). PLoS One 9: e93131. DOI: 10.1371/journal.pone.0093131.

Margarido, G. R. A., and D. Heckerman, 2015 ConPade: Genome assembly ploidy estimation from next-generation sequencing data. PLoS Comput. Biol. 11: e1004229. DOI: 10.1371/journal.pcbi.1004229.

Mateus, R. G., S. C. L. Barrios, C. B. Valle, J. R. Valério, F. Z. V. Torres *et al*., 2015 Genetic parameters and selection of *Brachiaria decumbens* hybrids for agronomic traits and resistance to spttlebugs. Crop Breed. Appl. Biotechnol. 15: 227-234.

Metzler, M. C., J. R. Cutt and D. F. Klessig (1991) Isolation and characterization of a gene encoding a PR-1-Like Protein from *Arabidopsis thaliana*. Plant Physiol. 96: 346-348.

Miles, J. W., S. L. Lapointe, M. L. Escando´n, and G. Sotelo (1995) Inheritance of resistance to spittlebug (Homoptera: Cercopidae) in interspecific *Brachiaria* spp. hybrids. J. Econ. Entomol. 88: 1477Ð1481.

Misra, R.C., Sandeep, M. Kamthan, S. Kumar and S, Ghosh 2016 A thaumatin-like protein of *Ocimum basilicum* confers tolerance to fungal pathogen and abiotic stress in transgenic Arabidopsis. Sci. Rep. 6:25340. DOI: 10.1038/srep25340.

Mollinari, M., and O. Serang, 2015 Quantitative SNP genotyping of polyploids with MassARRAY and other platforms. Methods Mol. Biol. 1245: 215–241. DOI: 10.1007/978-1-4939-1966-6_17.

Mukherjee AK., M. J. Carp, R. Zuchman, T. Ziv, B. A. Horwitz *et al*., 2010 Proteomics of the response of Arabidopsis thaliana to infection with *Alternaria brassicicola*. J Proteomics 73:709–720. DOI: 10.1016/j.jprot.2009.10.005.

Naumova, T. N., M. D. Hayward, and M. Wagenvoort, 1999 Apomixis and sexuality in diploid and tetraploid accessions of *Brachiaria decumbens*. Sex. Plant Reprod. 12: 43-52. DOI: 10.1007/s004970050170.

Okada, M., C. Lanzatella, M. C. Saha, J. Bouton, R. Wu *et al.*, 2010 Complete switchgrass genetic maps reveal subgenome collinearity, preferential pairing and multilocus interactions. Genetics 185: 745–760. DOI: 10.1534/genetics.110.113910.

Patel, R. K., and M. Jain, 2012 NGS QC toolkit: A toolkit for quality control of next

generation sequencing data. PLoS One 7: e30619. DOI: 10.1371/journal.pone.0030619.

Pereira, G. S., A. A. F. Garcia, and G. R. A. Margarido, 2017 A fully automated pipeline for quantitative genotype calling from next generation sequencing data in polyploids. BMC Bioinformatics (submitted).

Pessoa-Filho, M., A. M. Martins, and M. E. Ferreira, 2017 Molecular dating of phylogenetic divergence between *Urochloa* species based on complete chloroplast genomes. BMC Genomics 18: 516. DOI: 10.1186/s12864-017-3904-2.

Poland, J. A., P. J. Brown, M. E. Sorrells, and J. L. Jannink, 2012 Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7: e32253. DOI: 10.1371/journal.pone.0032253.

Poland, J. A., and T. W. Rife, 2012 Genotyping-by-sequencing for plant breeding and genetics. Plant Genome 5: 92.

Rajput, S. G., D. K. Santra, and J. Schnable, 2016 Mapping QTLs for morpho-agronomic traits in proso millet (*Panicum miliaceum* L.). Mol. Breed. 36: 37.

Renvoize, S. A., W. D. Clayton, and C. H. S. Kabuye, 1996 Morphology, taxonomy, and natural distribution of *Brachiaria* (Trin.). Griseb, pp. 1-15 in *Biology, Agronomy and Improvement, Centro Nacional de Pesquisa de Gado de Corte (CNPGC) (CIAT publication no. 259)*, edited by J. W. Miles, B. L. Maass, C. B. do Valle, and V. Krumble. Centro Internacional de Agricultura Tropical (CIAT), Empresa Brasileira de

Pesquisa Agropecuaria (EMBRAPA), Campo Grande, BR.

Ripol, M. I., G. A. Churchill, J. A. G. Da Silva, and M. Sorrells, 1999 Statistical aspects of genetic mapping in autopolyploids. Gene 235: 31–41. DOI: 10.1016/S0378-1119(99)00218-8.

Salgado, L. R., R. Lima, B. Fd Santos, K. T. Shirakawa, Md. A. Vilela *et al.*, 2017 De novo RNA sequencing and analysis of the transcriptome of signalgrass (*Urochloa decumbens*) roots exposed to aluminum. Plant Growth Regul. 83: 157–170. DOI: 10.1007/s10725-017-0291-2.

Schwarz, G., 1978 Estimating the dimension of a model. Ann. Statist. 6: 461-464.

Serang, O., M. Mollinari, and A. A. F. Garcia, 2012 Efficient exact maximum a posteriori computation for Bayesian SNP genotyping in polyploids. PLoS One 7: e30906. DOI: 10.1371/journal.pone.0030906.

Shirasawa, K., M. Tanaka, Y. Takahata, D. Ma, Q. Cao *et al.*, 2017 A high-density SNP genetic map consisting of a complete set of homologous groups in autohexaploid sweetpotato (Ipomoea batatas). Sci. Rep. 7: 44207. DOI: 10.1038/srep44207.

Simioni, C., and C. B. Valle, 2009 Chromosome duplication in *Brachiaria* (A.Rich.) Stapf allows intraspecific crosses. Crop Breed. Appl. Biotechnol. 9: 328-334.

Song, H., P. Wang, C. Li, S. Han, C. Zhao *et al*., 2017 Comparative analysis of NBS-LRR genes and their response to *Aspergillus flavus* in *Arachis.* PLoS One 12: e0171181.doi:10.1371/journal.pone.0171181.

Souza, J. S, L. Chiari, R. M. Simeão, M. deM. Vilela and L. R. Salgado (2018) Development, validation and characterization of genic microsatellite markers in *Urochloa* species. Am. J. Plant Sci. 9: 281-295.

Stein, J., S. C. Pessino, E. J. Martínez, M. P. Rodriguez, L. A. Siena *et al.*, 2007 A genetic map of tetraploid *Paspalum notatum* Flügge (bahiagrass) based on single-dose molecular markers. Mol. Breed. 20: 153–166. DOI: 10.1007/s11032-007-9083-0.

Thaikua, S., M. Ebina, N. Yamanaka, K. Shimoda, K. Suenaga *et al.*, 2016 Tightly clustered markers linked to an apospory-related gene region and quantitative trait loci mapping for agronomic traits in *Brachiaria* hybrids. Grassl. Sci. 62: 69–80. DOI: 10.1111/grs.12115.

Triviño, N. J., J. G. Perez, M. E. Recio, M. Ebina, N. Yamanaka *et al.*, 2017 Genetic diversity and population structure of *Brachiaria* species and breeding populations. Crop Sci. 57: 1-12.

Uknes, S., B. Mauch-Mani, M. Moyer, S. Potter, S. Williams *et al*., 1992 Acquired resistance in Arabidopsis. The Plant Cell 4: 645-656.

Valério, J. R., and O. Nakano, 1992. Sintomatologia dos danos causados pelo adulto da cigarrinha *Zulia entreriana* (Berg, 1879) (Homoptera:Cercopidae) em *Brachiaria*

decumbens. An. Soc. Entomol. Bras. 21: 95–100.

Valério, J. R., H. Jeller, and J. Peixer, 1997 Seleção de introduções do gênero *Brachiaria* (Griseb) resistentes à cigarrinha *Zulia entreriana* (Berg) (Homoptera: Cercopidae). An. Soc. Entomol. Bras. 26: 383–387. DOI: 10.1590/S0301-80591997000200023.

Valério, J. R., and O. Nakano, 1988 Influência do adulto de Zulia entreriana (Berg, 1879) (Homoptera: Cercopidae) na produção e qualidade de *Brachiaria decumbens*. Pesqui. Agropecuaria Bras. 23: 447–453.

Valle, C. B. do, C. Simioni, R. M. S. Resende, and L. Jank, 2008 Melhoramento genético de *Brachiaria*, pp. 13-54 in *Melhoramento de Forrageiras Tropicais,* edited by R. M. S. Resende, C. B. do Valle and L. Jank. Embrapa Gado de Corte, Campo Grande, MS.

Valle, C. B., L. Jank, and R. M. S. Resende, 2009 O melhoramento de forrageiras tropicais no Brasil. Ceres 460-472.

Valle, C. B., M. C. M. Macedo, V. P. B. Euclides, L. Jank, and R. Resende, 2010 Gnero *Brachiaria*, pp. 30–77 in *Plantas Forrageiras*, edited by D.M.D.F. Martuscello, and J. Azevedo. Editora UFV, Viçosa.

van Loon, L. C., and E. A. Van Strien 1999 The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol. Mol. *Plant* Pathol. 55: 85-97.

van Loon, L. C., M. Rep and C. M. Pieterse 2006 Significance of inducible defende-related proteins in infected plants. Annu. Rev. Phytopathol. 44: 135-162.

van Ooijen, G., G. Mayr ,M. M. Kasiem, M. Albrecht, B. J. Cornelissen *et al*., 2008 Structure- function analysis of the NB-ARC domain of plant disease resistance proteins. J. Exp. Bot. 59: 1383-1397. doi: 10.1093/jxb/ern045.

Vigna, B. B. Z., J. C. S. Santos, L. Jungmann, C. B. do Valle, M. Mollinari *et al.*, 2016 Evidence of allopolyploidy in *Urochloa humidicola* based on cytological analysis and genetic linkage mapping. PLoS One 11: e0153764. DOI: 10.1371/journal.pone.0153764.

Voorrips, R. E., 2002 MapChart: Software for the graphical presentation of linkage maps and QTLs. J. Hered. 93: 77–78. DOI: 10.1093/jhered/93.1.77.

Voorrips, R. E., G. Gort, and B. Vosman, 2011 Genotype calling in tetraploid species from bi-allelic marker data using mixture models. BMC Bioinformatics 12: 172. DOI: 10.1186/1471-2105-12-172.

War, A. R., M. G. Paulraj, T. Ahmad, A. A. Buhroo, B. Hussain *et al*., 2012 Mechanisms of plant defense against insect herbivores. *Plant Signal.* Behav. 7:1306-20.

Worthington, M., C. Heffelfinger, D. Bernal, C. Quintero, Y. P. Zapata *et al.*, 2016 A parthenogenesis gene candidate and evidence for segmental allopolyploidy in apomictic *Brachiaria decumbens*. Genetics 203: 1117–1132. DOI: 10.1534/genetics.116.190314.

Wu, K. K., W. Burnquist, M. E. Sorrells, T. L. Tew, P. H. Moore *et al.*, 1992 The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor. Appl. Genet. 83: 294–300. DOI: 10.1007/BF00224274.