Conservation Genetics Resources
https://doi.org/10.1007/s12686-023-01312-8

METHODS AND RESOURCES

Unveiling the demographic background and genetic diversity of *Urochloa mosambicensis* (Poaceae) through genome-wide identification of simple sequence repeats and molecular marker development

Ueslei S. Leão[1] · Luice G. Bueno[2] · Aline B. Negreiros[1] · Geice R. Silva[2] · Rodrigo Maggioni[3] · Fabio B. Britto[4] · Jose L. R. Sarmento[4] · Diego B. Galvani[2] · Fábio M. Diniz[2]

Received: 26 September 2022 / Accepted: 4 August 2023

Abstract

The first set of simple sequence repeat (SSR) markers for *Urochloa mosambicensis* (Hack.) Dandy is described in this study. SSRs were isolated from a genomic library with 3491 contigs obtained by high-throughput sequencing. Primers were designed for twenty *loci*, and all primer pairs were successfully amplified in 39 genotypes, yielding an average of 4.95 alleles per *locus*. The mean polymorphic information content (PIC) was 0.559, and the mean discriminating power (DP) was 0.520 for all *loci* considering also all *U. mosambicensis* genotypes. The non-Bayesian clustering analysis has confirmed clustering consistency of the individuals and provides enough support for the existence of the pattern that also emerged from PCoA and UPGMA. These markers demonstrated potential transferability to *U. advena*, *U. oligotricha*, *U. brachyura* and *U. xantholeuca*. Our findings showed that the set of molecular markers described here have strong potential for fingerprinting and characterizing genetic diversity in intra-specific plant collections, whose breeding programs could be accelerated by the effective use of these molecular tools, not only in *U. mosambicensis*, but also within the genus *Urochloa*.

Keywords Forage grass · Capim-corrente · Next-generation sequencing · Molecular markers · SSR · Transferability

✉ Fábio M. Diniz
fabio.diniz@embrapa.br

Ueslei S. Leão
ueslei.leao@srn.uespi.br

Luice G. Bueno
luice.bueno@embrapa.br

Aline B. Negreiros
aline.negreiros@ifpi.edu.br

Geice R. Silva
geiceamb_bio@yahoo.com.br

Rodrigo Maggioni
maggioni@ufc.br

Fabio B. Britto
fbbritto@ufpi.edu.br

Jose L. R. Sarmento
sarmento@ufpi.edu.br

Diego B. Galvani
diego.galvani@embrapa.br

[1] Northeast Biotechnology Network - RENORBIO, Universidade Federal do Piauí, Teresina, Piauí CEP: 64049-550, Brazil

[2] EMBRAPA Caprinos e Ovinos, Sobral, Ceará CEP: 62010-970, Brazil

[3] Institute of Marine Sciences, LABOMAR, Universidade Federal do Ceará, Fortaleza, Ceará CEP: 60165-081, Brazil

[4] Departamento de Zootecnia e Ciências Biológicas, Universidade Federal do Piauí, Universitário Ministro Petrônio Portella, Campus, Teresina, Piauí CEP: 64049-550, Brazil

Published online: 21 August 2023

Introduction

Urochloa (syn. *Brachiaria*) is a genus of grasses from Africa, which belong to the family Poaceae (Gramineae), known worldwide for their high tolerance to water stress. It is widely distributed in tropical and subtropical regions of the world, and has become naturalized in some areas (Jank et al. 2005).

Despite the fact that several commercial *Urochloa* cultivars are largely used as an animal forage crop in tropical and humid subtropical zones, many species of the genus remain poorly studied. Some of these forage grasses, however, have sparked great interest over the past few years in response to climatic changes and the need for improving livestock production in semiarid zones. Notably, the species *Urochloa mosambicensis* (Hack.) Dandy, also known as "capim-corrente" in Brazil, has demonstrated strong adaptability and persistence in dry environments (Bueno et al. 2019), and therefore, great potential for use in pastures within tropical semiarid regions. It is highly favored by animals as a forage grass, capable of enduring grazing near ground level, and has gained growing acceptance in the semiarid region of the Brazilian Northeast. Morphologically, it is very similar to Brachiaria and can be cultivated in various types of soil, offering forage potential that can be successfully utilized, including intercropping with other forage species (Bueno et al. 2022).

Traditional assessment methods of germplasm diversity for conservation purposes and for selecting most promising genotypes in breeding programs include morphological and agronomic characterizations, as the selection of more competitive and well-adapted forage grasses is essential for using the diversity available in the germplasm (Valle et al. 2009). However, new tools have become necessary to allow faster and reliable characterization and identification of genotypes with specific desirable traits.

Molecular markers provide not only a molecular characterization of genotypes at any stage of development, but also the construction of linkage and QTL maps, gene flow and mating system evaluations, thus producing a large number of genetic information to improve and accelerate the development of superior cultivars (Nadeem et al. 2018).

Among all marker systems, simple sequence repeats (SSRs) of two to six tandem nucleotides, or commonly known as microsatellite markers, have become increasingly important genetic markers to generate information for the selection of genotypes in breeding programs (Rakshit et al. 2010) and conservation genetics (Jungmann et al. 2009a, 2009b). Whereas some SSR markers for *Urochloa* species have been developed (*U. brizantha*, Jungmann et al. 2009a; *U. ruziziensis*, Silva et al. 2013; *U. humidicola*, Jungmann et al. 2009b; Vigna et al. 2011a; Santos et al. 2015; and *U. decumbens*, Ferreira et al. 2016), specific markers for *U. mosambicensis* have not been reported.

More recently, with the advance of genome sequencing techniques, the development of DNA-based markers has become, substantially, a less time consuming and labor-intensive procedure. Next-generation sequencing approaches have provided a reliable platform for the screening of the presence of simple sequence repeats, where species-specific sequence information is not available, as it generates large amounts of sequencing data, thus facilitating the development of molecular genetic markers (Abdelkrim et al. 2009).

In this study we present the first set of SSR markers developed and characterized for *Urochloa mosambicensis* potentially useful for phylogenetic evaluations, conservation and breeding applications using genomic data from a de novo partial genome assembly of paired-end Illumina reads, and tested the transferability of these markers to closely related *Urochloa* species (*U. oligotricha*, *U. brachyura*, *U. advena* and *U. xantholeuca*).

Material and methods

Urochloa genotypes (n = 39) were obtained from Embrapa Beef Cattle, but maintained at Embrapa Goats and Sheep, Sobral, CE, Brazil (3°45′03.55″ S, 40°20′37.45″ W). All genotypes are presented in Table S1. Genomic DNA from young leaves was isolated using ExtractME® Genomic DNA Kit (BLIRT DNA, Gdansk, Poland), according to the manufacturer's instructions.

A single individual *Urochloa mosambicensis* with the highest quality DNA yield was selected for whole genome sequencing using Illumina technology. One microgram was used to prepare the genomic library for sequencing, following the standard protocol of the Illumina Nextera XT Library preparation kit (Illumina, San Diego, CA, USA). DNA was tagged and fragmented by the Nextera XT transposome, followed by limited-cycle PCR amplification, AMPure XP magnetic-bead purification (Agencourt Bioscience Corporation, Beverly, MA, USA), and the Illumina Nextera XT bead-based normalization protocol. The DNA library was sequenced using a MiSeq Benchtop Sequencer (Illumina), targeting 500-bp fragments with 2 × 250-bp reads in a paired-end sequencing configuration. Illumina short-read sequencing data were assembled into contigs using CLC Genomics Workbench 7.0.4 (Qiagen, Carlsbad, CA, USA).Microsatellites were mined using Msatcommander 0.8.2 software (Faircloth 2008) by screening the assembled contigs for repeat motifs with at least six repeats of di-, tri-, tetra-, penta-, and hexanucleotides, according to Vigna et al. (2011a). We targeted 20 microsatellite *loci*, as there

appears to be no significant gain in increasing this number of *loci* to accurately estimate allele frequencies (Bernatchez and Duchesne 2000). Di-, tri- and tetranucleotides were randomly selected due to their inherent ability to generate a significant degree of polymorphism, thereby facilitating the discrimination between individuals and/or populations.

Forward and reverse PCR primers were designed for each simple sequence repeat at their conserved flanking regions. Primer design was performed using the Primer3Plus (https://www.primer3plus.com/).

Primer pairs transferability between the *Urochloa mosambicensis* and related species available in our germplasm collection (*U. brachyura*, *U. oligotricha*, *U. advena*, *U. xantholeuca* and an unknown *Urochloa* species) were also tested, which is important for reducing the cost of microsatellite isolation for these *Urochloa* species.

Genomic DNA from thirty-nine plant materials was used to validate all designed primer pairs using polymerase chain reactions (PCRs). Reactions were performed in a 10-µL total volume containing at least 10 ng of genomic DNA, with 1 × buffer, 1.5–2.5 mM $MgCl_2$, 0.8 mM dNTP mix, 0.5 µM of each primer and 0.5 U of Platinum *Taq* DNA polymerase (Invitrogen Life Technologies, Inc., Carlsbad, CA). All amplifications were run on a Veriti 96-well Thermal Cycler (Applied Biosystems, Waltham, MA, USA) using the PCR temperature profile indicated in Table 1. All amplification products were screened by silver nitrate detection on non-denatured 6% polyacrylamide gels. Due to the polyploidy nature of the genus *Urochloa* the molecular markers were treated as dominant markers according to previously published research (Abd El-Moghny et al. 2017). Each individual was scored based on the presence (1) or absence (0) of a clear, unambiguous and reproducible PCR band.

The polymorphic information content (*PIC*) was calculated based on a binary data matrix, according to the following equation: $1 - \sum_i^n pi^2 - \sum_i^{n-1} \sum_{i=i+1}^n 2pi^2pj^2$ (Botstein et al. 1980). Discriminatory power (DP) of each SSR *locus*, which represents the probability that two randomly chosen individuals are distinguishable from one another, was estimated according to Tessier et al. (1999) as $DP = 1 - C$, where C is the confusion probability, $C = \sum_{i=1}^{I} ci = \sum_{i=1}^{I} pi\frac{(Npi-1)}{N-1}$, where for N individuals, C is equal to the sum of all ci for all of the patterns generated by the SSR primer.

Genetic relationships among all genotyped species and among *U. mosambicensis* sampling localities were calculated using GenAlEx 6.5b3 software (Peakall and Smouse 2012), as well as the genetic diversity within the groups was calculated using Shannon's Information Index (SI). The frequency of all amplified bands in each group was used to estimate Nei´s distance and to create a dendrogram (*U. mosambicensis* genotypes) and a phylogram (including five *Urochloa* species) using the Neighbor-Joining and unweighted paired group method using arithmetic average (UPGMA) algorithms estimated by R package POPPR (Kamvar et al. 2014) and APE (Paradis et al. 2004), respectively. The confidence of branch support was then evaluated by a bootstrap analysis with 1,000 iterations.

The number of potential genetic populations or genetic clusters (*K*) was estimated using a maximum likelihood-based method in a model-free iterative reallocation approach implemented in the program FLOCK v3.1 (Duchesne and Turgeon 2012). *K* was searched as suggested in the manual of the software, with a random choice of samples for generating the initial partition, 50 iterations per run, 50 runs, and a log-likelihood minimum (LLOD) of 0.3. The number of clusters was determined using the plateau analysis.

Lastly, a principal coordinate analysis (PCoA) was also carried out in the R package ADEGENET (Jombart 2008) to explore associations among *U. mosambicensis* genotypes from different sampling localities using the Nei´s distance.

Results and discussion

Illumina MiSeq sequencing resulted in 57,170,592 paired-end reads, which were assembled into a total of 137,313 contig sequences of the *U. mosambicensis* genome. Contig size varied from 200 to 25,210 bases, with an average size of 411 bases. The software Msatcommander identified 3,491 contigs with simple sequence repeats of di- to hexanucleotides motifs with at least six repetitions in tandem. From these, 20 *loci* were randomly selected for primer design and validation in *Urochloa*. All sequences isolated and validated in this study were deposited in the GenBank database (NCBI) under accession numbers MH742936 to MH742955 (Table 1).

All 20 SSR primer pairs amplified consistently interpretable alleles in all investigated *Urochloa* species (Table 2). A total number of 99 alleles were revealed for all these genotypes (N = 39) with an average of 4.95 alleles/*locus*, ranging from 2 (*Umo*3, 5, 6, 9, 11, 13, 14, 17 and 19) to 16 (*Umo*16). From these, 83 alleles were present in *U. mosambicensis*, ranging from 1 (*Umo*9, and *Umo*11) to 11 (*Umo*16 and *Umo*20) with an average of 4.15 alleles/*locus*.

An average cross-species transferability success rate of at least 80% was found for *U. xantholeuca*, which means that 16 SSR primer pairs developed for *U. mosambicensis* successfully amplified at least a single *U. xantholeuca* genotype. However, the success rate across the other *Urochloa* species has increased to 100% in *U. oligotricha* with all newly developed SSR primers successfully amplifying at least a single genotype. These results suggest considerable sequence conservation within flanking SSR *loci* for all tested *Urochloa* related species available in the *Embrapa Caprinos e Ovinos* germplasm collection, where primers anneal to their specific

Table 1 Characteristics of 20 simple sequence repeat (SSR) *loci* developed from *Urochloa mosambicensis*

Locus	Primer sequences (5′–3′)	Repeat motif	Reaction profile	Allele size range (bp)	Ta (°C)	GenBank accession no
*Umo*1	F-CTTGTGGGTTTCATCTCTAGCC R-AACAAAGGTTTAGCTGCAATGG	$(TGT)_{11}$	PCR1	085–145	55	MH742936
*Umo*2	F-GTTCATCTTCATCACTACATTGCC R-ACCGCTTTAGCTCCTCACATAA	$(GAT)_{10}$	PCR2	139–184	55	MH742937
*Umo*3	F-GTTCACAGATCGAGACGTGG R-GTCTGGTTTCCTCTTGGTGC	$(CAA)_{10}$	PCR2	165–201	55	MH742938
*Umo*4	F-CCTTTAATGTCGCACGAAACA R-TGCTCTAACGTGAAGAAGCTAGG	$(GTA)_{12}$	PCR1	095–107	55	MH742939
*Umo*5	F-CCATGAGTGTTGTCGAAGAAGT R-GAGCACGATCAAGGAAGTTG	$(TGA)_{10}$	PCR1	087–093	55	MH742940
*Umo*6	F-CGCCCCACTTTTCAGGAT R-GGCCCGCATCTAAAATCACTAC	$(GTA)_{12}…(GTA)_{21}$	PCR2	073–076	55	MH742941
*Umo*7	F-GCTTTAGCTCCTCACATAATTCCT R-GTTCATCTTCATCACTACATTGCC	$(ATC)_{14}$	PCR2	219–225	55	MH742942
*Umo*8	F-TTGAGAAAGGGATGAGAAGGTG R-AAAACTGCAAAACCCGGC	$(CGAC)_7$	PCR2	090–122	50	MH742943
*Umo*9	F-CCGGGCTTAACTCATACACAA R-TTCCACTCTTTCTTCTACGGGA	$(GA)_{32}$	PCR3	138–152	57	MH742944
*Umo*10	F-TTCGTGTGTTTGTATGTGGTGA R-CGATAGCATTAGGGAAGGAATG	$(TG)_{29}$	PCR2	116–148	50	MH742945
*Umo*11	F-CTCCCCAGCACTCTCACCT R-GAGAAGAAGATCCAAGAAGCGA	$(CT)_{28}$	PCR3	079–083	57	MH742946
*Umo*12	F-CTTGAGATGAACGTCTTCGTTG R-TTCTCCACACAAACAGAACAGC	$(AG)_{28}$	PCR3	086–094	57	MH742947
*Umo*13	F-AGCCCACACATCACGTCC R-TCACTCGCCTCAACACGTC	$(CGGCT)_9$	PCR4	094–109	55	MH742948
*Umo*14	F-CGTAGGGCTCTCACAAACATA R-TTGTCATGGGTGTAGGTTCTAA	$(CT)_{18}$	PCR2	130–136	50	MH742949
*Umo*15	F-CATTGTCTCTGTTTATATCCCTTCC R-TTCGTGTGTTTGTATGTGGTGA	$(CA)_8CG(CA)_{41}$	PCR2	046–052	50	MH742950
*Umo*16	F-GGTTCCGTTGGAGGATTAGAC R-AGGAAGGAGCAAGGAGTGAAG	$(GA)_{31}$	PCR2	122–156	55	MH742951
*Umo*17	F-GATCCATGCATACATAAAATACTAG R-TGGAGACATATTTGAGAGGCAT	$(AT)_{22}$	PCR3	150–170	58	MH742952
*Umo*18	F-TTCGTGTGTTTGTATGTGGTGA R-TCTAGTCCACATGGTTTCTCGA	$(TG)_{25}$	PCR2	078–104	50	MH742953
*Umo*19	F-GGATCCATGCATACATAAAATACT R-TTGAGAGGCATGATATTGTTTG	$(TA)_{20}G(AG)_{15}$	PCR2	145–149	58	MH742954
*Umo*20	F-GGCTGAGTTCTCGGATAGAAAG R-ACCAACAAGAAAGCTAGGAACC	$(GT)_{29}$	PCR4	054–066	55	MH742955

Ta annealing temperature

PCR1: 94 °C-5 min, 32×(94 °C-1 min, Ta°C-1 min, 72 °C- 1 min), 72 °C-5 min; PCR2: 94 °C-5 min, 40×(94 °C-1 min, Ta°C-1 min, 72 °C-1 min), 72 °C-10 min; PCR3: 94 °C-5 min, 35×(94 °C-1 min, Ta°C-1 min, 72 °C- 1 min), 72 °C-7 min; PCR4: 95 °C-15 min, 40×(94 °C-30 s, Ta°C-90 s, 72 °C- 90 s), 72 °C-10 min

targets. Transferability success revealed by the cross-amplification amongst phylogenetically related species is a consequence of the homology of flanking regions of the DNA sequence in microsatellite *loci* (Primmer and Merilä, 2000). However, sequence conservation in these regions within genera at *loci* containing SSRs is not completely explained (Chen et al. 2002). Therefore, the extent of marker transferability across these *Urochloa* grass species was associated

Table 2 Characterization of simple sequence repeat markers developed for *Urochloa mosambicensis* and transferability to other *Urochloa* species

SSR *locus*	Marker power (N = 39)				Characteristics in *U. mosambicensis* (N = 22)				Transferability[a]				
	A	PIC	DP	I (SD)	A	PIC	DP	I (SD)	*Urochloa advena*	*Urochloa brachyura*	*Urochloa oligotricha*	*Urochloa xantholeuca*	*Urochloa* sp.
*Umo*1	11	0.232	0.276	0.393 (0.225)	10	0.129	0.199	0.420 (0.181)	2/2	3/4	5/5	1/1	5/5
*Umo*2	7	0.729	0.818	0.446 (0.129)	6	0.650	0.756	0.489 (0.143)	0/2	0/4	2/5	1/1	0/5
*Umo*3	2	Mono	Mono	Mono	2	Mono	Mono	Mono	2/2	4/4	5/5	0/1	5/5
*Umo*4	4	**0.063**	**0.074**	0.350 (0.282)	3	**0.099**	**0.043**	0.449 (0.233)	2/2	4/4	5/5	1/1	5/5
*Umo*5	2	Mono	Mono	Mono	2	Mono	Mono	Mono	2/2	4/4	5/5	1/1	5/5
*Umo*6	2	Mono	Mono	Mono	2	Mono	Mono	Mono	0/2	1/4	3/5	0/1	3/5
*Umo*7	3	0.253	0.304	0.313 (0.271)	3	0.320	0.343	0.334 (0.289)	2/2	4/4	5/5	1/1	5/5
*Umo*8	8	0.336	0.370	0.396 (0.219)	8	0.374	0.434	0.388 (0.196)	2/2	3/4	3/5	1/1	3/5
*Umo*9	2	Mono	Mono	Mono	1	Mono	Mono	Mono	2/2	4/4	5/5	1/1	5/5
*Umo*10	4	0.651	0.659	0.199 (0.324)	3	0.535	0.545	0.208 (0,361)	2/2	4/4	5/5	1/1	5/5
*Umo*11	2	0.132	0.146	0.271 (0.000)	1	Mono	Mono	Mono	2/2	4/4	5/5	1/1	5/5
*Umo*12	4	0.530	0.547	0.424 (0.328)	4	0.475	0.502	0.414 (0,327)	2/2	4/4	5/5	1/1	5/5
*Umo*13	2	0.589	0.595	0.326 (0.462)	2	0.595	0.606	0.328 (0.463)	2/2	4/4	5/5	1/1	5/5
*Umo*14	2	Mono	Mono	Mono	2	Mono	Mono	Mono	2/2	4/4	5/5	1/1	5/5
*Umo*15	4	**0.899**	**0.909**	0.304 (0.247)	3	**0.846**	**0.862**	0.326 (0.332)	2/2	4/4	5/5	1/1	4/5
*Umo*16	16	0.817	0.856	0.266 (0.140)	11	0.740	0.809	0.339 (0.154)	2/2	3/4	2/5	1/1	5/5
*Umo*17	2	Mono	Mono	Mono	2	Mono	Mono	Mono	2/2	4/4	5/5	1/1	5/5
*Umo*18	6	0.391	0.417	0.464 (0.228)	5	0.277	0.314	0.464 (0.201)	1/2	4/4	5/5	0/1	5/5
*Umo*19	2	Mono	Mono	Mono	2	Mono	Mono	Mono	2/2	4/4	5/5	1/1	5/5
*Umo*20	14	0.758	0.795	0.300 (0.168)	11	0.765	0.830	0.341 (0.147)	2/2	2/4	5/5	0/1	4/5

A allele number, *PIC* polymorphic information content; *DP* Discriminatory power; *mono* monomorphic. *I* Shannon's Information Index, *SD* standard deviation

[a]Number of successfully amplified genotypes/Number of tested genotypes

Bold indicates the highest and lowest values

with the phylogenetic proximity (Torres González and Morton 2005; Morrone and Zuloaga 1992) of the target species to the nonfocus species, as also seen in Fig. 1.

Seven SSRs were monomorphic for all evaluated genotypes (*Umo*3, 5, 6, 9, 14, 17, and 19), thus sharing the same number of repeat units. These *loci* may be useful in broader surveys in the future when additional accessions are sampled, as they may turn into polymorphic markers once new alleles are revealed in the population. If monomorphism is fully confirmed, these *loci* may become useful as a new source of data since their flanking regions may also be used as a novel tool, thereby serving as a source of sequence variation, in genetic and breeding studies (Nazareno and Reis 2011).

Unfortunately, the apparent lack of genetic variability of monomorphic microsatellites has excluded these potential markers from some genetic analyses. The display of monomorphism in a specific *loci* can be explained by the fact that microsatellites may be required for a specific metabolic function and are therefore conserved (Simpson 2012), or in part by some aspect of the species' reproductive biology influencing a higher frequency of homozygous alleles in the progeny (Engel et al. 1996).

As also shown in Table 2, PIC and DP values ranged from 0.063 (*Umo*4) to 0.899 (*Umo*15), and from 0.074 (*Umo*4) to 0.909 (*Umo*15), for the entire data set, respectively. For *Urochloa mosambicensis*, PIC and DP varied from 0.099 (*Umo*4) to 0.846 (*Umo*15), and from 0.043 (*Umo*4) to 0.862 (*Umo*15), respectively. The highest values for DP, among all examined *Urochloa* species, or even for all *U. mosambicensis* genotypes, were obtained for *locus Umo*15, which also showed the highest PIC values. The lowest PIC and DP values for *locus Umo*4 resulted from low levels of polymorphism, which suggests that this *locus* is located in a conserved region shared among the studied *Urochloa* species (Vigna et al. 2011b). The mean PIC value was 0.521 ± 0.275 for the entire data set and 0.559 ± 0.257 for *U. mosambicensis* genotypes. More than 50% of the polymorphic *loci* exhibited PIC values greater than 0.500, which characterizes most of the SSR *loci* set as highly informative markers (Botstein

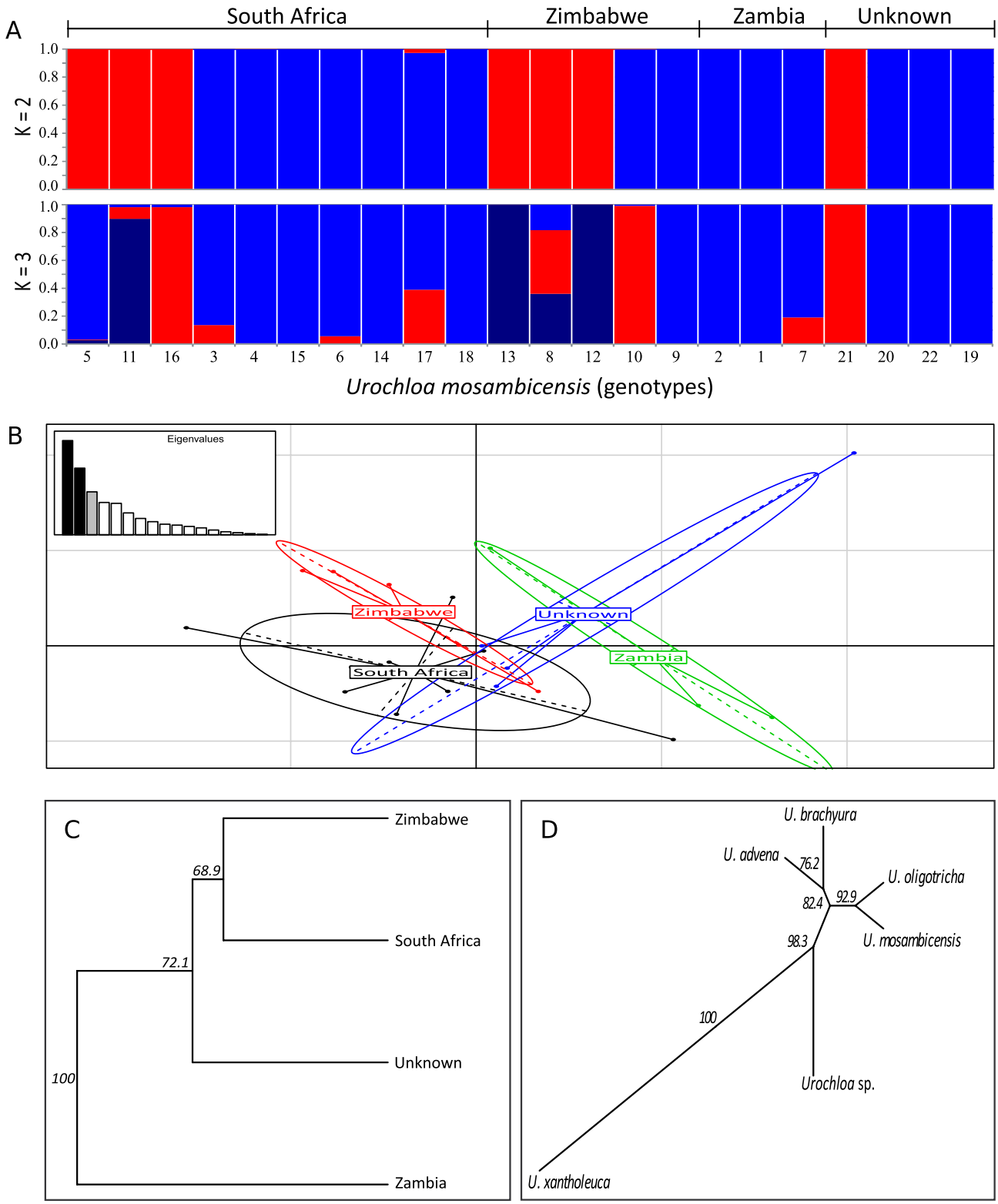

Fig. 1 **A** A grouping pattern revealed for *Urochloa mosambicensis* by a non-Bayesian recollection approach using FLOCK. Each bar represents each individual genotype. **B** Principal coordinate analysis (PCoA) plots of *U. mosambicensis* samples collected in Africa. Solid-colour circles represent genotypes and inertia ellipses predefined populations. The unweighted paired group method using arithmetic average (UPGMA) dendrogram summarizing genetic relationships among **C** *U. mosambicensis* populations and **D** *Urochloa* species based on Nei's distances. The numbers on the nodes indicate the percentage bootstrap values generated from 1000 resamplings

et al. 1980; Serrote et al. 2020), thus indicating the high level of polymorphism across *loci*, and suggesting that SSR markers had the necessary discriminating power to differentiate *Urochloa* genotypes.

High PIC values allow greater distinction between individuals and/or populations due to the greater number of different alleles with balanced frequencies, which increases the probability of unique combinations of alleles in different individuals (Serrote et al. 2020).

By assessing PIC as a measure of genetic diversity, it becomes apparent that the markers' polymorphism is primarily driven by variation within the population of *U. mosambicensis* genotypes rather than divergence between species. Despite the limited sample size, the results reinforce the existence of substantial genetic variability among genotypes. This also implies that the initial isolation between localities did not sufficiently reduce the number of effective alleles, indicating potential for future breeding programs (Fig. 1).

Notably, Ondabu et al. (2017) reported close PIC values of 0.582 for 79 *Urochloa* (formerly *Brachiaria*) ecotypes. These findings align with our study results, as the diversity level (PIC) obtained is likely influenced by the mixture of accessions from diverse origins without significant geographic isolation (Fst = 0.023). It is worth considering the impact of the apomictic mode of reproduction on reducing polymorphism, particularly in markers with low PIC values, in addition to the *locus* position in the genome.

Most polymorphic markers had greater power of discrimination (SD) among *U. mosambicensis* samples (60%) when considering values above 0.500, than the entire data set data (54%), which can be explained by the difference in allele frequency showed by each SSR *locus*. Tessier et al. (1999) points out that the closer the frequency between the bands, the lower the confusion probability and the greater the DP, which corroborates our results when observing Pearson's correlation between the standard deviation of the mean values of allele frequency for each *locus* and DP values (entire data set r = − 0.50, p value = 0.02; and *U. mosambicensis* r = − 0.75, p-value = 0.00012). Estimation of DP and PIC plays a crucial role in the selection of markers that are most effective in differentiating individuals and/or populations. By employing these estimations, researchers can obtain reliable results using a smaller set of analyzed *loci*, thereby optimizing their time and resource allocation.

The Shannon's Information Index ranged from 0.199 (*Umo*10) to 0.464 (*Umo*18) with an average value of 0.348 ± 0.078 for all examined *Urochloa* species. For *U. mosambicensis* separately, it ranged from 0.208 (*Umo*10) to 0.489 (*Umo*2), averaging 0.375 ± 0.078. These estimates showed the existence of genetic variation within all genotypes, but mainly within *U. mosambicensis*, and therefore, represent the effectiveness of microsatellite *loci* to reveal the variation, which was mirrored in the genetic diversity analysis depicted in Fig. 1. Indexes were higher than that reported for *U. ruziziensis* (0.12), *U. bovonei* (0.25), *U. jubata* (0.21); however, lower than that of *U. humidicola* (0.77) (Kuwi et al. 2018). These observations offer valuable insights into the potential factors influencing the level of genetic diversity observed in our study. Notably, we obtained a high gene flow index (Nm = 4.77, FST = 0.05), indicating a substantial historical gene flow between *U. mosambicensis* genotypes from Zimbabwe, South Africa, and Zambia. This finding reinforces the notion that gene exchange among these regions has been prevalent in the past.

The non-Bayesian clustering analysis, using FLOCK, did not determine an exact value of K, but indicated that there are two or three clusters in a complex genetic structure of *U. mosambicensis* among the geographical origin of the studied genotypes (Fig. 1A). The FLOCK assignment of genotypes confirms the clustering consistency of the individuals, one of which was restricted to Zambia, and provides enough support for the existence of the pattern that also emerged from PCoA and UPGMA (Fig. 1B, C). Therefore, the SSR markers developed in this work provide high-resolution power, even if the species seems to be panmictic occupying a broad range of habitats from tropical environments in Africa.

The PCoA analysis indicated that genetic similarity among South Africa and Zimbabwe genotypes was high and did not reveal any population structure reflecting the geographical origin of these samples (Fig. 1B). This result coincides with the results obtained by FLOCK assignments.

The dendrogram constructed by UPGMA based on Nei's genetic distance successfully discriminated geographical origin and species (Fig. 1C, D) with distinct relationships among tested genotypes, again as found by using FLOCK, although the two types of analyses used different statistical approaches. For *U. mosambicensis*, the resultant groups were strongly correlated with geographical origin (Fig. 1C). The tree in Fig. 1D indicated that *U. mosambicensis* and *U. oligotricha* are more closely related to one another than to the other species (*U. brachyura*, *U. advena* and *U. xantholeuca*). Indeed, *U. mosambicensis* and *U. oligotricha* exhibit a close relationship, which indicates their probable origin from a common ancestor (McIvor 1992). Despite the complexity of the interaction between genome and the environment, as well as the need for a more in-depth study in plant systematics, both species have several characteristics that may give clues to this genetic convergence. These include their perennial nature and clump growth habit, as well as stem heights (Umo = 20–150 cm and Uo = 60–120 cm). Additionally, their inflorescences are arranged in open panicles containing spikelets (http://www.worldfloraonline.org/).

The development of SSR markers from assembled Illumina pair-end DNA sequences using high throughput

technologies is indeed a very promising approach for species with limited genomic information, instead of creating a genomic library enriched for tandem repeats and subsequent molecular cloning and Sanger sequencing.

Our findings showed that the set of molecular markers described here, which are the first SSR *loci* for *Urochloa mosambicensis*, have strong potential for fingerprinting and characterizing genetic diversity in intra-specific plant collections, whose breeding programs could be accelerated by the effective use of these molecular tools, not only in *U. mosambicensis*, but also within the genus *Urochloa*. Most of the *U. mosambicensis* SSR markers (75%) have also amplified all tested species, and therefore, may benefit both genetic and breeding applications to facilitate improvement of also *U. oligotricha*, *U. brachyura*, *U. advena* and *U. xantholeuca*.

We expect that these markers may also be useful for further studies, including molecular mapping, species and hybrid identification, gene flow, seed purity and for other kinds of studies in ecology and evolution of *Urochloa*. Ongoing research is involved in applying these markers for the improvement and development of *U. mosambicensis* cultivars.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12686-023-01312-8.

Acknowledgements The authors would like to thank the Genomics and Bioinformatics Center of Drug Research and Development Center of Federal University of Ceará for technical support.

Authors' contributions USL, LGB, DBG, ABN, and GRS were involved in sample collection, methodology, laboratory work, data curation, formal analysis, manuscript writing. FBB and JLRS contributed to data curation, manuscript reviewing and editing. LGB, DBG, RM and FMD carried out fieldwork, partial funding acquisition and manuscript reviewing. USL and FMD supported laboratory work, LGB and DBG contributed to conceptualization, fieldwork, methodology, resources, supervision, writing original draft and editing. All authors read and approved the final manuscript.

Funding The Brazilian Agricultural Research Corporation–Embrapa and The Foundation for Scientific and Technological Development of Ceará–FUNCAP funded this study through the project grant no. MP 22.16.04.018.00 (MegFor) and BP5-0197-00214.01.00/22, respectively. FMD would like to thank FUNCAP for a Research Productivity Scholarship.

Data availability The datasets generated or analyzed during the current study are available in the GenBank repository (accession numbers MH742936–MH742955) but restrictions apply to the availability of these data, so they are not publicly available until article publication.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Abd El-Moghny AM, Santosh HB, Raghavendra KP, Sheeba JA, Singh SB, Kranthi KR (2017) Microsatellite marker based genetic diversity analysis among cotton (*Gossypium hirsutum*) accessions differing for their response to drought stress. J Plant Biochem Biotechnol 26(3):366–370

Abdelkrim J, Robersten BC, Stanton JA, Gemmell NJ (2009) Fast, cost-effective development of species-specific microsatellite markers by genome sequencing. Biotechniques 46:185–191

Bernatchez L, Duchesne P (2000) Individual-based genotype analysis in studies of parentage and population assignment; how many *loci*, how many alleles? Can J Fish Aquat Sci 57(1):1–12

Botstein D, White RL, Skolnick M (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Med Genet 32:314–331

Bueno LG, Alves MMA, Rocha JES, Cavalcante ACR, Galvani DB, Diniz FM, Valle CB, Cândido MJD (2019). Caracterização morfogênica de *Urochloa mosambicensis* para seleção de genótipos elite em programa de melhoramento de forrageiras. Boletim de Pesquisa e Desenvolvimento/Embrapa Caprinos e Ovinos 11, Sobral: Embrapa Caprinos e Ovinos, 2019. 19 p.

Bueno LG, Galvani DB, Voltolini TV, Diniz FM (2022) Capim-urocloa, alternativa para região semiárida. In: Santos MVF, Neiva JNM (Org.). Culturas forrageiras no Brasil: uso e perspectivas. Visconde do Rio Branco: Suprema Gráfica, p. 177–197.

Chen X, Cho Y, McCouch S (2002) Sequence divergence of rice microsatellites in *Oryza* and other plant species. Mol Genet Genomics 268(3):331–343

Duchesne P, Turgeon J (2012) FLOCK provides reliable solutions to the "number of populations" problem. J Hered 103:734–743

Engel SR, Linn RA, Taylor JF, Davis SK (1996) Conservation of microsatellite *loci* across species of Artiodactyls: Implications for population studies. J Mammalog 77(2):504–518

Faircloth B (2008) Msatcommander: Detection of microsatellite repeat arrays and automated, *locus*-specific primer design. Mol Ecol Resour 8:92–94

Ferreira RCU, Cançado LJ, Valle CB, Chiari L, Souza AP (2016) Microsatellite *loci* for *Urochloa decumbens* (Stapf) RD Webster and cross-amplification in other *Urochloa* species. BMC Res Notes 9:152

Jank L, Resende RMS, Valle CB (2005) Genética Em Pastagens. Rev USP 64:86–93

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24(11):1403–1405

Jungmann L, Sousa ACB, Paiva J, Francisco PM, Vigna BBZ, Valle CB, Zucchi MI, Souza AP (2009) Isolation and characterization of microsatellite markers for *Brachiaria brizantha* (Hochst. ex A. Rich.) Stap. Conserv Genet 10:1873

Jungmann L, Vigna BBZ, Paiva J, Sousa ACB, Valle CB, Laborda PR, Zucchi MI, Souza AP (2009b) Development of microsatellite markers for *Brachiaria humidicola* (Rendle) Schweick. Conserv Genet Resour 1:475

Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281

Kuwi SO, Kyalo M, Mutai CK, Mwilawa A, Hanson J, Djikeng A, Ghimire SR (2018) Genetic diversity and population structure of *Urochloa* grass accessions from Tanzania using simple sequence repeat (SSR) markers. Rev Bras Bot 41:699–709

McIvor JG (1992) *Urochloa mosambicensis* (Hack.) Dandy. In 'Plant Resources of South-East Asia No 4. Forages' (Eds L Mannetje and RM Jones) pp. 230–231. (Pudoc Scientific Publishers, Wageningen, The Netherlands).

Morrone O, Zuloaga FO (1992) Revision de las especies sudamericanas nativas e introducidas de los generos *Brachiaria* y *Urochloa* (Poacaeae:Panicoideae:Paniceae). Darwiniana 31(1–4):43–109

Nadeem MA, Nawaz MA, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, Hatipoğlu R, Ahmad F, Alsaleh A, Labhane N, Özkan H, Chung G, Baloch FS (2018) DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotechnol Biotechnol Equip 32(2):261–285

Nazareno AG, Reis MS (2011) The same but different: monomorphic microsatellite markers as a new tool for genetic analysis. Am J Bot 98(10):e265–e267

Ondabu N, Maina S, Kimani W, Njarui D, Djikeng A, Ghimire S (2017) Molecular characterizations of Kenyan *Brachiaria* grass ecotypes with microsatellite (SSR) markers. Agronomy 7(1):8

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290

Peakall R, Smouse PE (2012) GenAlEx 65: Genetic analysis in Excel population genetic software for teaching and research: an update. Bioinformatics 28:2537–2539

Primmer CR, Merilä J (2000) A low rate of cross-species microsatellite amplification success in Ranid frogs. Cons Gen 3:445–449

Rakshit A, Rakshit S, Singh J, Chopra SK, Balyan HS, Gupta PK, Bahat SR (2010) Association of AFLP and SSR markers with agronomic and fibre quality traits in *Gossypium hirsutum* L. J Genet 89(2):155–162

Santos JCS, Barreto MA, Oliveira FA, Vigna BBZ, Souza AP (2015) Microsatellite markers for *Urochloa humidicola* (Poaceae) and their transferability to other *Urochloa* species. BMC Res Notes 8:83

Serrote CML, Reiniger LRS, Silva KB, Rabaiolli SMDS, Stefanel CM (2020) Determining the Polymorphism Information Content of a molecular marker. Gene 726:144175

Silva PIT, Martins AM, Gouvea EG, Pessoa-Filho M, Ferreira ME (2013) Development and validation of microsatellite markers for *Brachiaria ruziziensis* obtained by partial genome assembly of Illumina single-end reads. BMC Genom 14:17

Simpson MC (2012) Microsatellite analysis of *Ceratocystis fimbriata*. Dissertation (M.Sc.), University of Pretoria, Pretoria, South Africa.

Tessier C, David J, This P, Boursiquot JM, Charrier A (1999) Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. Theor Appl Genet 98:171–177

Torres-González AM, Morton CM (2005) Molecular and morphological phylogenetic analysis of *Brachiaria* and *Urochloa* (Poaceae). Mol Phylogenet Evol 37:36–44

Valle CB, Jank L, Resende RMS (2009) O melhoramento de forrageiras tropicais no Brasil. Rev Ceres 56(4):460–472

Vigna BBZ, Alleoni GC, Jungmann L, Valle CB, Souza AP (2011a) New microsatellite markers developed from *Urochloa humidicola* (Poaceae) and cross amplification in different *Urochloa* species. BMC Res Notes 4:523

Vigna BBZ, Jungmann L, Francisco PM, Zucchi MI, Valle CB, Souza AP (2011b) Genetic diversity and population structure of the *Brachiaria brizantha* germplasm. Trop Plant Biol 4:157–169