HIGH-THROUGHPUT POLYMORPHISM DETECTION AND GENOTYPING IN POLYPLOID *Paspalum plicatulum* USING NEXT-GENERATION RAD SEQUENCING

de Oliveira FA[1], Vigna BBZ[2], Fávero AP[2], Hohenlohe PA[3], de Souza AP[1,4].

[1]Center for Molecular Biology and Genetic Engineering (CBMEG), University of Campinas (UNICAMP), Campinas, SP, Brazil. [2]Brazilian Agricultural Research Corporation (Embrapa) Southeast Livestock, São Carlos, SP, Brazil. [3]Institute for Bioinformatics and Evolutionary Studies, Department of Biological Sciences, University of Idaho, Moscow, USA. [4]Plant Biology Department, Biology Institute, UNICAMP, Campinas, SP, Brazil.

Paspalum plicatulum Michx. is an American native grass, originating in Brazil and widely distributed from the south of the U.S.A. to the south of Argentina. Is is important to central and south Brazil as natural grasslands and presents adaptation to drought, cold, flooding, fire and defoliation. Is a polyploid species that presents high morphological variation and difficulties in species classification. *P. plicatulum* belongs to the botanical group Plicatula, which presents high genetic and morphological diversity and forms an agamic complex. We aim to contribute to the genetic knowledge of *P. plicatulum*, setting a first set of DNA polymorphisms based on RAD sequencing. We applied this approach in 20 individuals from the Active Bank of Germplasm, maintained at EMBRAPA Southeast Livestock. RAD libraries were prepared according to Miller et al. (2007), using the restriction enzyme SbfI and each individual was barcoded with a unique 8bp sequence. We modified the standard protocol to target DNA fragments of 600–700 bp during shearing size selection. All samples were sequenced on an Illumina MiSeq v3 with paired-end length of 300bp. We conducted initial processing of the sequence data using several modules from the Stacks software. First using *process_radtags*, we sorted read pairs by barcode and remove any pairs in which the forward read did not contain both a correct barcode and the remaining six bases of the SbfI recognition sequence. As a pre-processor to assembly, we used FLASH software to merge paired-end reads, resulting longer reads. Then we grouped the forward and reverse reads from all individuals into a separate file for each RAD locus, using the *sort_read_pairs.pl*. We assembled the reads in each file separately to produce a set of RAD contigs, using CAP3 assembly software. After filtering, approximately 29 million RAD sequence read pairs were generated across 20 individuals. We identified a total of 225 327 RAD tags and 287 642 SNP alleles in Stacks using the forward reads of each pair across all individuals. Of these, a total of 55000 RAD contigs were generated from CAP3 and will be use to genotype SNP loci.Financial support: FAPESP, CAPES.

Corresponding author: Fernanda A. Oliveira, f.ancelmo.o@gmail.com