Proceedings X-Meeting 2018

Editor: AB^{3}C

October 2018

ANALYSIS OF PASPALUM NOTATUM TRANSCRIPTOME UNDER DROUGHT CONDITION

Joyce Etsuko Arakaki, Wilson Malagó Junior, MAURICIO DE ALVARENGA MUDADU, Patricia Menezes Santos, Alessandra Pereira Fávero, Ricardo Carneiro Borra, Bianca Baccili Zanotto Vigna

Abstract

Paspalum notatum Flüggé is a perennial grass widely used as forage and as turf in many countries as Argentina, Australia, United States and in the South of Brazil. The species present diploid and tetraploid citotypes with x=10 chromosomes. Transcriptomic and differential gene expression (DGE) studies in plants, under water deficit, are used to investigate the molecular mechanisms of transcriptional regulation and potential genes involved in the response to drought, making possible the genetic modification and/or gene editing of Paspalum and other plants of commercial interest, as already done in bahiagrass and maize. The objective of this study is to evaluate the gene expression profile of Paspalum notatum Flüggé in response to drought. The tetraploid accession P. notatum BGP 216 conserved at the germplasm bank of Embrapa Pecuária Sudeste (São Carlos, Brazil) was selected for a drought stress experiment based on the climate characteristics of its site of collection, on the importance of the species and on the ability of the species to tolerate drought. The experiment was conducted in greenhouse during Brazilian spring (Nov/2016), in triplicates, and leaf samples of each biological replicate were collected for mRNA extraction in two conditions, without drought stress and after five days under drought stress (4% of relative soil moisture). The samples were sequenced using Next Generation Sequencing (NGS) technology on Illumina HiSeq 2500 equipment with HiSeq-V4 kit and 2x100 bp paired-end reads. Bioinformatics tools were used to characterize the quality of the de novo transcriptome assembly, to analyze DGE between plants with and without drought stress and to perform the functional annotation of the genes. Around 80 millions of reads were sequenced per sample and Trinity software was used for de novo assembly, annotation and DGE studies of the transcriptome. Around 244 million bases were assembled, 168.509 transcripts were identified from 56.718 predicted genes and the transcriptome had a N50 stats of 2.226 bp with median contig length of 1.083 bp. The alignment of the reads to the assembly was performed by Bowtie2 and 91.61% were mapped as proper pairs which is better than the usually expected 70-80%. Diamond Blast determined that 17.381 of 47.514 proteins were covered by more than 90% of their protein lengths. The latter result, additionally to the results of ExN50 stats, FastQC and Transrate (score of 0.3701) softwares showed satisfactory quality of the reads and assembly. BUSCO3 was used to search a dataset of ortholog conserved genes and returned a good result of 86.9% of transcriptome completeness against the embryophyta dataset from OrthoDB v9 database. DGE analyses were performed using EdgeR software and identified 1.978 differentially expressed genes (DEG) (p-value < 0.001 and Fold Change > 2). The Trinotate software suite was used for automatic functional annotation of the transcriptome. Initial searches of the top ten DEG in UniProt databases and PubMed showed relation to biological process and molecular functions. The annotated genes ARF9, ARF24, DJC76 and THI1-2 have been related to plant and root development, biotic and abiotic stresses responses (including drought) in other organisms such as Arabidopsis thaliana, Gossypium barbadense, Solanum lycopersicum and Oryza sativa. The characterization of this accession as to their strategy of tolerance to drought and the better understanding of the molecular and physiological mechanisms involved may contribute for