

HIGH-THROUGHPUT SEQUENCING AND DE NOVO ASSEMBLY OF TRANSCRIPTOME OF *Vigna unguiculata* UPON VIRAL INFECTION

Flavia Figueira Aburjaile¹, João Pacifico Bezerra Neto¹, Bruna Piereck Moura¹,
José Ribamar Costa Ferreira-Neto¹, Ana Maria Benko-Iseppon¹,

1 Federal University of Pernambuco, Center of Biological Sciences, Genetics Dept

Abstract

Plants are often submitted to adverse environmental conditions, such as abiotic and biotic stresses. According to the Food and Agriculture Organization of the United Nations, biotic stresses are responsible for diseases leading to 32 – 42% of productivity decrease. To neutralize infections, plants first recognize the invading pathogens by quickly and efficiently activating molecular mechanisms. During the infectious process, the plant response involves changes at physiological, biochemical and molecular level, through activation of specific gene expression programs. In this context, the study of the transcriptome is an excellent alternative for identification of genes involved in plant-pathogen interaction, mainly for species considered as non-models like cowpea bean (*V. unguiculata*). Our transcriptome assembly was generated from 12 libraries obtained for two different cowpea cultivars (IT85F-2687 and BR14-Mulato) submitted to Cowpea Severe Mosaic Virus (CpSMV) and Cowpea Aphid-borne Mosaic Virus (CABMV), respectively. The libraries were generated with three biological replicates for each treatment and control group, allowing us to decode its molecular behavior towards the pathogen in question. The total RNA from all samples described above was sequenced on the Illumina platform. The data were processed according to the following steps: (1) analysis of data quality, (2) assembly "de novo", (3) annotation and (4) statistical analysis of differentially expressed genes, according to the MUGQIC Pipeline protocol. These sequences were divided into structural or functional categories, according to the gene family of interest. After quality analysis, more than 350 millions were assembled, and returned 149,288 transcripts from 72,140 genes, presenting a mean of 1,289 bp of length and a N50 of 1,981. Our annotation aligned with Uniprot and Pfam sequences returned 54.8% and 40.9% of results, respectively representing 81,822 transcripts. Additionally, we performed gene ontology enrichment returning annotation for 76,089 transcripts (50.97%). This assembly will be an important resource to improve our understanding of main mechanisms that help cowpea cultivars to tolerate virus infection.

Funding: CAPES, CNPQ, FACEPE.