

## **GLOBAL ANALYSIS OF DIFFERENTIAL EXPRESSION IN Coffea arabica SEEDS DURING THE PROCESSING AND DRYING**

## SANTOS, F. C.\*, ROSA, S. D. V. F\*\*, GROTEWOLD, E.\*, PADILHA, L.\*\*, MUKUNDI, E.\*

\*The Ohio State University, Columbus, Ohio, USA, \*\*Brazilian Agricultural Research Corporation, Lavras, MG, Brazil. Supported by CNPq, Fapemig, Capes and INCT Café.



Coffee is one of the most important agricultural commodities and ensures the livelihood of more than 80 million people worldwide. The mode of coffee processing, whether wet or dry, determines the flavor characteristic and also the seed quality. Novel, high-throughput, deepsequencing technologies are making an impact on genomic research by providing new strategies to analyze the functional complexity of transcriptomes. The RNA-Seq approach produces millions of short cDNA reads that are mapped to a reference genome to obtain a genome-scale transcriptional map, which consists of the transcriptional structure and the expression level for each gene. In this study we applied Illumina sequencing technology and the Venn diagram analysis for differential expression. Coffee seeds were processed by three different methods, dry, semi-dry or wet, and were dried under shade or dryer conditions. The samples were used to prepare a cDNA library for sequence analysis via Illumina (HiSeq 2000). After removal of lowquality reads, a total of 125,829,579 high-quality 50 bp reads were identified. The reads were assembled using Trinity program using default parameters. It was compared our data against ESTs Brazilian Coffee Genome data and Coffee Colombian data. As can be seen in the Venn diagram representation, we've got more blast in ESTs Brazilian Coffee Genome data than in Coffee Colombian data, indicating an alignment of 20,509 common sequences between ESTs Brazilian Coffee Genome and "Assembly de Novo" sequence (our data).



Venn diagram showing different alignment against ESTs Brazilian Coffee Genome Project and Colombian Coffee data.