

## **De novo assembly, characterization and functional annotation of pineapple fruit transcriptome through massively parallel sequencing**

### **Abstract**

Background: Pineapple (*Ananas comosus* var. *comosus*), is an important tropical non-climacteric fruit with high commercial potential. Understanding the mechanism and processes underlying fruit ripening would enable scientists to enhance the improvement of quality traits such as, flavor, texture, appearance and fruit sweetness. Although, the pineapple is an important fruit, there is insufficient transcriptomic or genomic information that is available in public databases. Application of high throughput transcriptome sequencing to profile the pineapple fruit transcripts is therefore needed.

Methodology/Principal Findings: To facilitate this, we have performed transcriptome sequencing of ripe yellow pineapple fruit flesh using Illumina technology. About 4.7 millions Illumina paired-end reads were generated and assembled using the Velvet de novo assembler. The assembly produced 28,728 unique transcripts with a mean length of approximately 200 bp. Sequence similarity search against non-redundant NCBI database identified a total of 16,932 unique transcripts (58.93%) with significant hits. Out of these, 15,507 unique transcripts were assigned to gene ontology terms. Functional annotation against Kyoto Encyclopedia of Genes and Genomes pathway database identified 13,598 unique transcripts (47.33%) which were mapped to 126 pathways. The assembly revealed many transcripts that were previously unknown.

Conclusions: The unique transcripts derived from this work have rapidly increased of the number of the pineapple fruit mRNA transcripts as it is now available in public databases. This information can be further utilized in gene expression, genomics and other functional genomics studies in pineapple.