

**Division-4-01**

**Status of Philippine Mango Genomics: Enriching Molecular Genomics Towards a Globally Competitive Philippine Mango Industry**

Eureka Teresa M. Ocampo<sup>1</sup>, Cris Q. Cortaga<sup>2</sup>, Jhun Laurence S. Rasco<sup>3</sup>, John Albert P. Lachica<sup>2</sup>, Darlon V. Lantican<sup>2</sup>

<sup>1</sup>Institute of Crop Science

<sup>2</sup>Institute of Plant Breeding

<sup>3</sup>Institute of Crop Science, College of Agriculture and Food Science, University of the Philippines Los Baños, College, Laguna, Philippines

**[Abstract]**

This paper presents the first genome assemblies of Philippine mangoes that provide valuable reference for varietal improvement and genomic studies on mango and related fruit crops. WE sequenced whole genomes of 3 species, *Mangifera odorata* (Huani), *Mangifera altissima* (Paho), and *Mangifera indica* ‘Carabao’ (Sweet Elena). ‘Carabao’ is the major export variety of the Philippines; Paho is identified as vulnerable by the IUCN Red List of Threatened Species; Huani has fruit sap acrid which is the primary defense mechanism against insects and birds. We used Falcon, a diploid aware *de novo* assembler to assemble SMRT generated long-read sequences. Falcon-unzip was employed to phase the output assembly producing larger contig sets (primary contigs) and shorter contigs corresponding to haplotypes (haplotigs). Assembly statistics were generated by comparing the assembly to a reference genome, Tommy Atkins, using Quality Assessment Tool (QUAST). Moreover, the extent of duplication and completeness of gene content was measured using Benchmarking Universal Single-Copy Orthologs (BUSCO). Draft assemblies with high duplications were processed using Purge Haplotigs and Purge Dups to lessen duplications with minimal impact on genome completeness. *De novo* assemblies of Huani, Paho and ‘Carabao’ were then generated with primary contig sizes of 463.64 Mb, 508.95 Mb and 401.51 Mb respectively. These draft assemblies of Huani, Paho and ‘Carabao’ showed 96.90%, 95.17% and 99.07% complete BUSCOs respectively which is comparable to ‘Tommy Atkins’ genome (98.6%). Using two mango transcriptome data (pooled RNA-seq from different mango varieties and tissues), 91-96% or 24-30 million reads were successfully mapped back for each generated assembly indicating high degree of completeness. The results obtained demonstrated the highly contiguous, phased, and near complete genome assembly of three Philippine mango species for structural and functional annotation of gene units, especially those with economic importance.

**[Acknowledgement]**

We would like to acknowledge the funding support from the Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development of the Philippines’ Department of Science and Technology for the project Full Genome Sequencing of Selected Philippine Mango Species.

\*Corresponding author: E-mail. emocampo1@up.edu.ph Tel. +63 9474421034