



Data in Brief

RNA-seq analysis of mangosteen (*Garcinia mangostana* L.) fruit ripening

Azhani Abdul-Rahman, Hoe-Han Goh, Kok-Keong Loke, Normah Mohd Noor, Wan Mohd Aizat*

Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

ARTICLE INFO

Keywords:

Fruit ripening
Garcinia mangostana L.
 Illumina sequencing
 Transcriptome

ABSTRACT

Mangosteen (*Garcinia mangostana* L.) is known for its delectable taste and contains high amount of xanthenes which have been reported to possess anti-cancer, anti-inflammatory and other bioactive properties. However, stage-specific regulation of mangosteen fruit ripening has never been studied in detail. We have performed a comparative transcriptomic analysis of three ripening stages (Stage 0, 2 and 6) of mangosteen. We have obtained a raw data from six libraries through Illumina HiSeq 4000. A total of ~40 Gb of raw data were generated. Clean reads of 650,887,650 (bp) were obtained from 656,913,570 (bp) raw reads. The raw transcriptome data were deposited to SRA database, with the BioProject accession number of PRJNA339916. These data will be beneficial for transcriptome profiling in order to study the regulation of mangosteen fruit ripening. The lack of a complete sequence database from this species impedes protein identification. These data sets provide a reference data for the exploration of novel genes or proteins to understand mangosteen fruit ripening behaviour.

Specifications

Organism/cell line/tissue	<i>Garcinia mangostana</i> L. (Pericarps)
Sex	N/A
Sequencer or array type	Illumina HiSeq™ 4000
Data format	Raw data: FASTQ file
Experimental factors	Ripening stages of mangosteen, stage 0 pericarp/ aril, stage 2 pericarp, stage 6 pericarp
Experimental features	Two biological replicates for each stage
Consent	N/A
Sample source location	UKM Bangi, Malaysia (2°55′09.0″N 101°47′04.8″E)

1. Direct link to deposited data

The data is accessible via the following link <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA339916> and individual link for each sample are provided as follows:

2. Introduction

Mangosteen (*Garcinia mangostana* L.) is a tropical climacteric fruit

from the family of Clusiaceae (Guttiferae) cultivated in Southeast Asian countries such as Malaysia, Thailand, Indonesia and Philippines [1,2]. Its extract is rich in beneficial metabolites particularly xanthenes which are known to have anti-cancer, anti-inflammatory, anti-oxidant, anti-bacteria and anti-viral properties, among others [1,2]. The ripening of mangosteen fruit can be divided into seven stages, Stage 0 until Stage 6 [3]. Despite being classified as climacteric (ripening with a surge of ethylene production), mangosteen will only be fully ripen if harvested at the middle of ripening process (Stage 2 onwards), not at mature green stage (Stage 0). In contrast, other climacteric fruits such as tomato and banana will ripen off the plant once they reach mature green stage [4]. A recent study has been conducted which employed an ion proton sequencer for the *de novo* transcriptome analysis of mangosteen [5] yet this is only specific to one fully ripened stage. In this study, three mangosteen ripening stages (Stages 0, 2 and 6 corresponding to early, middle and late ripening stages) were analysed using RNA-seq, specifically Illumina HiSeq 4000 platform, to understand the ripening regulation. The short reads were trimmed, processed, assembled and analysed as described below. Raw reads for this project were deposited in the NCBI SRA database (Table 1). These data sets would be beneficial to reveal any novel genes or proteins in the mangosteen ripening process.

* Corresponding author.

E-mail address: wma@ukm.edu.my (W.M. Aizat).

Table 1
SRA accession links for mangosteen raw data.

Stages	Biological replicates	Accession number	Accession links
Stage 1	GmS0-1	SRX2066746	http://www.ncbi.nlm.nih.gov/sra/SRX2066746
		SRX2066747	http://www.ncbi.nlm.nih.gov/sra/SRX2066747
	GmS0-3	SRX2066753	http://www.ncbi.nlm.nih.gov/sra/SRX2066753
Stage 2	GmS2-1	SRX2066754	http://www.ncbi.nlm.nih.gov/sra/SRX2066754
		SRX2066757	http://www.ncbi.nlm.nih.gov/sra/SRX2066757
	GmS2-2	SRX2066758	http://www.ncbi.nlm.nih.gov/sra/SRX2066758
		SRX2066759	http://www.ncbi.nlm.nih.gov/sra/SRX2066759
Stage 6	GmS6-1	SRX2066761	http://www.ncbi.nlm.nih.gov/sra/SRX2066761
		SRX2066763	http://www.ncbi.nlm.nih.gov/sra/SRX2066763
	GmS6-4	SRX2066764	http://www.ncbi.nlm.nih.gov/sra/SRX2066764
		SRX2066751	http://www.ncbi.nlm.nih.gov/sra/SRX2066751
		SRX2066752	http://www.ncbi.nlm.nih.gov/sra/SRX2066752

Table 2
Statistics of mangosteen fruit transcriptome assembly.

Attributes	Value
Pre-assembly	
Total raw reads	656,913,570
Total processed reads	650,887,650
Post-assembly	
Number of unigene	181,646
Number of unique transcript	250,682
N50 (bp)	1158

3. Experimental design, materials and methods

3.1. Plant materials

Garcinia mangostana L. fruits were obtained from experimental plot at Universiti Kebangsaan Malaysia, Bangi (2°55'09.0"N 101°47'04.8"E). Mangosteen were harvested according to different stages (Stage 0, 2 and 6) of ripening process (May until September 2014) and their pericarps were separated and grounded in liquid nitrogen before being stored at −80 °C for analysis. Whole fruit of Stage 0 was used because pericarp and aril were inseparable at this stage [3]. Five biological replicates from each stage were used for RNA extraction and two replicates from each stage with the highest RIN number were chosen for RNA-seq analysis.

3.2. Total RNA extraction and quality control, library preparation and RNA-seq

Modified CTAB method was used to isolate pure RNA from the mangosteen fruit [6]. NanoDrop spectrophotometer (Thermo Fisher Scientific Inc., USA) and Agilent 2100 Bioanalyzer (Agilent Technologies, USA) were used to confirm the total RNA quantity and integrity (RIN > 8). Purified samples were then prepared using the standard polyA-enriched library preparation protocol implemented by Macrogen, South Korea. Sequencing was performed using the Illumina HiSeq 4000 platform that generates paired end reads of 100 bp.

3.3. Transcriptome de novo assembly

Adapter sequences were removed from the raw reads with

Trimmomatic program [7]. Only high quality reads with phred score ≥ 25 were retained for *de novo* assembly using Trinity (V2.2.0) [8]. We obtained 250,682 and 181,646 of unique transcripts and unigenes respectively. Statistics of the assembly is shown in Table 2.

Conflict of interest

All the authors have approved the submissions and there are no conflicts of interest.

Acknowledgements

This research was supported by Fundamental Research Grant Scheme (FRGS/2/2014/SG05/UKM/02/2) from the Malaysian Ministry of Higher Education (MOHE) and Geran Galakan Penyelidik Muda (GGPM-2014-008) from Universiti Kebangsaan Malaysia (UKM).

References

- [1] T. Shan, Q. Ma, K. Guo, J. Liu, W. Li, F. Wang, E. Wu, Xanthones from mangosteen extracts as natural chemopreventive agents: potential anticancer drugs, *Curr. Mol. Med.* 11 (2011) 666–677.
- [2] J. Pedraza-Chaverri, N. Cárdenas-Rodríguez, M. Orozco-Ibarra, J.M. Pérez-Rojas, Medicinal properties of mangosteen (*Garcinia mangostana*), *Food Chem. Toxicol.* 46 (2008) 3227–3239.
- [3] M. Osman, A.R. Milan, *Mangosteen: Garcinia mangostana* L. University of Southampton, Southampton Centre for Underutilised Crops, 2006.
- [4] G.B. Seymour, N.H. Chapman, B.L. Chew, J.K.C. Rose, Regulation of ripening and opportunities for control in tomato and other fruits, *Plant Biotechnol. J.* 11 (2013) 269–278.
- [5] D.D. Matra, T. Kozaki, K. Ishii, R. Poerwanto, E. Inoue, *De novo* transcriptome assembly of mangosteen (*Garcinia mangostana* L.) fruit, *Genomics Data.* 10 (2016) 35–37.
- [6] S. Kim, T. Hamada, Rapid and Reliable Method of Extracting DNA and RNA from Sweetpotato, *Ipomoea batatas* (L). *Lam. Biotechnol. Lett.* 27 (2005) 1841–1845.
- [7] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics* 30 (2014) 2114–2120.
- [8] B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D. Eccles, B. Li, M. Lieber, M.D. Macmanes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C.N. Dewey, R. Henschel, R.D. Leduc, N. Friedman, A. Regev, *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis, *Nat. Protoc.* 8 (2013) 1494–1512.