



Data in Brief

Transcriptome analysis of *Carica papaya* embryogenic callus upon De-etiolated 1 (*DET1*) gene suppression

Diyana Jamaluddin, Normah Mohd Noor, Hoe-Han Goh*

Institute of Systems Biology, Universiti Kebangsaan Malaysia, Malaysia

ARTICLE INFO

Keywords:
3' mRNA sequencing
DET1
Embryogenic callus
Photomorphogenesis
QuantSeq

ABSTRACT

Papaya is considered to be one of the most nutritional fruits. It is rich in vitamins, carotenoids, flavonoids and other phytonutrient which function as antioxidant in our body [1]. Previous studies revealed that the suppression of a negative regulator gene in photomorphogenesis, De-etiolated 1 (*DET1*) can improve the phytonutrient in tomato and canola without affecting the fruit quality [2,3]. This report contains the experimental data on high-throughput 3' mRNA sequencing of transformed papaya callus upon *DET1* gene suppression.

Specifications

Organism/cell line/tissue	<i>Carica papaya</i> var. Sekaki embryogenic calli
Sex	Not applicable
Sequencer or array type	Illumina HiSeq 2500 system
Data format	Raw (FASTQ) sequences
Experimental factors	RNA used for library preparation and sequencing was isolated from embryogenic calli of <i>C. papaya</i> upon the suppression of <i>DET1</i> gene using Agrobacterium-mediated transformation compared to control calli
Experimental features	RNA-seq data was obtained from 3' mRNA sequencing to estimate gene abundance in count per million (CPM) which represents the expression level of each transcript
Consent	Not applicable
Sample source location	Bangi, Malaysia

Direct link to deposited data

<http://www.ncbi.nlm.nih.gov/sra/SRX2060181>
<http://www.ncbi.nlm.nih.gov/sra/SRX2060182>
<http://www.ncbi.nlm.nih.gov/sra/SRX2060183>
<http://www.ncbi.nlm.nih.gov/sra/SRX1816058>

Value of the data

- Provide the first transcriptome data of *C. papaya* embryogenic callus upon *DET1* gene suppression.
- The present dataset is valuable for the identification of the genes which response to *DET1* gene suppression.
- This information will be useful for better understanding on the relationship between photomorphogenesis and secondary metabolite biosynthesis.

1. Data

Data reported here describes the comparison of *C. papaya* embryogenic callus obtained from four transformed calli (TC1, TC3, TC4 and TC5). *DET1* gene expression was successfully down-regulated in TC3 and TC5 (suppression line) while TC1 and TC4 had similar *DET1* gene expression as in control calli (escapes). This transcriptomic dataset was generated by QuantSeq 3' mRNA sequencing [4] of cDNA prepared from RNA extracts of Agrobacterium-transformed papaya callus for both suppression and escape samples. The raw sequence data for transcript quantification and DEG analysis were deposited into the NCBI SRA database and can be accessed according to the links provided above.

2. Experimental design, materials and methods

2.1. Plant materials

Agrobacterium-mediated transformation using pHELLSGATE vector

* Corresponding author.

E-mail address: gohhh@ukm.edu.my (H.-H. Goh).

Table 1
Statistics on *Carica papaya* QuantSeq sequencing.

Sample	Raw reads	Processed reads	% mapped to reference genome
TC1	6,853,408	3,931,664	43.3
TC3	4,448,583	3,222,940	58.8
TC4	7,281,986	5,566,775	63.6
TC5	6,542,239	4,879,749	58.3

containing hairpin *DET1* construct (*hpDET1* through Gateway cloning using PCR product of primers 5'-CACCGAGGTGCATTAAGAACAAAGC-3' and 5'-TGCCGGTCCAATGCAGAAAT-3') were conducted to suppress *DET1* gene expression in 4 weeks old embryogenic calli of *C. papaya*. RT-qPCR was performed to verify the expression of *DET1* gene upon *hpDET1* suppression showed that *DET1* gene expression was suppressed in TC3 (SRX2060182) and TC5 (SRX1816058) compared to TC1 (SRX2060181) and TC4 (SRX2060183) which showed normal *DET1* gene expression level similar to non-transformed callus.

2.2. RNA isolation

RNA from calli showing suppressed (TC3 & TC5) and normal (TC1 & TC4) *DET1* gene expression and was extracted using TRIzol (Invitrogen) according to manufacturer's instruction. RNA purity and integrity was measured using the ND-1000 Nanodrop spectrophotometer (Thermo Fisher Scientific Inc., USA) and Agilent 2100 bioanalyzer (Agilent Technologies, USA), respectively.

2.3. Library construction and QuantSeq 3' mRNA sequencing

RNA samples were cleaned using the DNase I kit according to the Rapid out removal DNA kit instruction (Thermoscientific) and converted into cDNA by using the QuantSeq 3' mRNA-Seq Reverse (REV) Library Prep Kit (Lexogen) according to manufacturer's instruction to generate compatible libraries for Illumina sequencing. cDNA libraries were assessed using a TapeStation (Agilent Technologies, USA) before 100 bp single end sequencing using a Illumina HiSeq 2500 system at the Australian Genome Research Facility (AGRF) based on standard protocols.

2.4. Transcriptome analysis

Raw sequencing reads (FASTQ) from calli of suppressed

(TC3 & TC5) and normal (TC1 & TC4) *DET1* expression were processed individually to remove low quality sequences (QV < 10 of 4-base sliding window) and unknown sequences with 'N' using Trimmomatic [5]. To quantify transcript abundance, the processed reads (FASTA) were mapped to papaya genome reference version Cpapaya_113 (<http://www.plantgdb.org/CpGDB/>). The mapping (Table 1) was performed using bowtie2 [6] with stringent "end-to-end" alignment and all other parameters were set to default values according to recommended data analysis workflow by Lexogen (<http://www.lexogen.com/quantseqdataanalysis>). Transcript abundance in count per million (CPM) were determined using eXpress [7] followed by differentially expressed gene (DEG) analysis using edgeR [8] based on the following cut-off parameters: P < 0.001, FDR < 0.05 and Log₂|Fold Change| > 2. This results in the identification of 439 DEGs.

Acknowledgements

This research was supported by the Malaysian Ministry of Science, Technology and Innovation (MOSTI) Sciencefund grant 02-01-02-SF0907 and Universiti Kebangsaan Malaysia Research University Grant (GGPM-2011-053).

References

- [1] M.M. Wall, Ascorbic acid, vitamin A, and mineral composition of banana (*Musa sp.*) and papaya (*Carica papaya*) cultivars grown in Hawaii, *J. Food Compos. Anal.* 19 (2006) 434–445.
- [2] G.R. Davuluri, A. van Tuinen, P.D. Fraser, A. Manfredonia, R. Newman, D. Burgess, D.A. Brummell, S.R. King, J. Palys, J. Uhlig, P.M. Bramley, H.M.J. Pennings, C. Bowler, Fruit-specific RNAi-mediated suppression of *DET1* enhances carotenoid and flavonoid content in tomatoes, *Nat. Biotechnol.* 23 (2005) 890–895.
- [3] S. Wei, X. Li, M.Y. Gruber, R. Li, R. Zhou, A. Zebarjadi, A. Hannoufa, RNAi-mediated suppression of *DET1* alters the levels of carotenoids and sinapate esters in seeds of *Brassica napus*, *J. Agric. Food Chem.* 57 (2009) 5326–5333.
- [4] P. Moll, M. Ante, A. Seitz, T. Reda, QuantSeq 3' mRNA sequencing for RNA quantification, *Nat. Methods* 11 (2014).
- [5] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for illumina sequence data, *Bioinformatics* (2014) (btu170).
- [6] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, *Nat. Methods* 9 (2012) 357–359.
- [7] A. Roberts, L. Pachter, Streaming fragment assignment for real-time analysis of sequencing experiments, *Nat. Methods* 10 (2013) 71–73.
- [8] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a bioconductor package for differential expression analysis of digital gene expression data, *Bioinformatics* 26 (2010) 139–140.