Journal of Food Science & Technology (ISSN: 2472-6419)

# Putative Allergens Identified in Mango (*Mangifera indica Linn*) Leaf and Fruit with Transcriptome Analysis

DOI: 10.25177/JFST.5.2.RA.10636

Research

**Accepted Date: 24th April 2020; Published Date:30th April 2020**



**Xiaolong Huang [1,2,3], Huiqing Yan [3]*, Zongmin Wu[1,2] , Yin Yi[1,2]***

[1] *Key Laboratory of Plant Physiology and Development Regulation, Guizhou Normal University, Guiyang 550001, China.*
[2] *Key Laboratory of State Forestry Administration on Biodiversity Conservation in Mountainous Karst Area of Southwestern China, Guizhou Normal University, Guiyang 550001, China.*
[3] *School of Life Sciences, Guizhou Normal University, Guiyang 550001, China.*

**CORRESPONDING AUTHOR**
Yin Yi (Email: gzklppdr@gznu.edu.cn)
Huiqing Yan (Email: yanhuiqing@gznu.edu.cn)

**CITATION**
Xiaolong Huang, Huiqing Yan, Zongmin Wu, Yin Yi, Putative Allergens Identified in Mango (Mangifera indica Linn) Leaf and Fruit with Transcriptome Analysis(2020) Journal of Food Science & Technology 5(2) pp: 98-110

**ABSTRACT**
Some allergens have previously been identified in mango (*Mangifera indica* Linn), including profilins, Bet v 1-like proteins and chitinase. In this paper, we identified the deepest investigation of mango potential allergens using high-throughput Illumina sequencing. RNA-Seq generated 11,751,123 contigs that were assembled into 99,328 unigenes with 16,848 unigenes of >1000 bp. A total of 230,242 unigenes were annotated using public protein databases, with a cut-off E-value above $10^{-5}$, of which 27,295, 46,030, 24,227 and 14,023 unigenes were assigned to gene ontology terms, Nr, Swiss-Prot and clusters of orthologous groups, respectively. A total of 66 potential allergen genes were identified, and their relative expressions were evaluated using Illumina RNA-Seq technology. Allergens mainly belonged to pollen allergen, pathogenesis-related protein Bet v I family and NADPH-dependent FMN reductase. We selected c61327.graph_c0, highly expressed in fruit and annotated as Pollen Ole e 1, was used as a template to obtain homologous protein structure in the RCSB PDB bank (PDB: 4z8w). We over-expressed and purified c61327.graph_c0, which could bind to human IgE by immunoblotting analysis.
The epitope (74-SFRQEVKTEKHGEFKVHLPFSVSEHV-99) was speculated to confer allergic reactions. Therefore, this study provided a comprehensive systemic view of the transcriptome between mango leaf and fruit allergens endowed with biological activities, which will be useful for further genomic research studies and breeding of lower allergenic mango cultivars.

**Keywords**: *Mangifera indica* Linn; transcriptome analysis; allergens; protein structure

## 1. INTRODUCTION

Mango (*Mangifera indica* Linn) belongs to the *Anacardiaceae* family and is the most important tropical fruit crop in China, distributed in Hainan, Guangxi, Guizhou and other provinces (Wu *et al.*, 2014). Mango is popularly regarded as "the king of fruits". Their fruits are rich in antioxidant vitamins A and C, B6 (pyridoxine), folate, potassium and omega-3 and -6 polyunsaturated fatty acids which all benefit human health (Dautt-Castro *et al.*, 2015; Srivastava *et al.*, 2016). Mango has also been reported to have antibacterial and anti-carcinogenic action (Noratto *et al.*, 2010). Mango can be processed to make juices, ice creams, fruit bars, smoothies and spicy chilli paste (Fasoli and Righetti, 2013).

Despite the massive consumption of mango worldwide, hypersensitivity reactions caused by fruits cannot be ignored (Paschke *et al.*, 2001; Hassan and Venkatesh, 2015). Allergy to mango has a range of symptoms of varying levels (Weinstein *et al.*, 2004). Some people show delayed hypersensitivity reaction, presenting with erythema, urticaria, dyspnea, anaphylaxis or angioedema, and others are entirely debilitated with immediate hypersensitivity reactions, showing oral allergy syndrome and manifesting in facial angioedema, hoarseness, pruritus of palms and respiratory distress (Wu *et al.*, 2012). Moreover, conventional mango processing into products does not allow the complete elimination of allergenic potency (Dube *et al.*, 2004).

Allergy cross-reactivity is frequently observed in daily life (Diaz-Perales *et al.*, 1999). Cross-reactions between mango fruit and various other foods have been reported due to the typical structure and properties of each protein family over a wide range of plant species, genera and even families (Wellhausen *et al.*, 1996; Oka *et al.*, 2004). These proteins could be recognized by the immune system and the ingestion of pollen, which can trigger an allergic reaction in a susceptible individual (Vargas Correa *et al.*, 1991; Renner *et al.*, 2008). Mango allergens were reported to cross-react with birch pollen, celery, citrus, pistachio nut, *Artemisia* pollen and papaya (Song *et al.*, 2008). Several allergens in mango fruits were identified and purified using an immunoglobulin (IgE) detection system and were also identified on the international allergen list (http://www.allergome.org/index.php). In mango, they are mainly denoted as Man i1, Man i2 and Man i3, attributed to ribosomal protein, NADH-plastoquinone oxidoreductase subunit and cytochrome c heme attachment protein, respectively. Recombinant Man i1 was purified in *Escherichia coli*, with potential for use in immunotherapy against mango allergy (Tsai *et al.*, 2017).

Previous research has revealed mango physiological results, including volatile composition, postharvest management and fruit quality during the ripening process (Hoang *et al.*, 2015). Recently, genomic information about mango development has received more attention (Luria *et al.*, 2014). Large-scale Illumina sequencing has provided a comprehensive gateway to determine the new transcripts, gene expression and more accurate profiles of the transcriptome, and has become a powerful technology for species that lack reference genome information (Hong *et al.*, 2016; Liu *et al.*, 2016). The assembled data can be analyzed to evaluate a wide variety of genetic characteristics and metabolic pathways in mango fruit, and much mango transcriptome data have been reported. More than 13,500 unigenes of mango related to expression in leaf tissues and the chloroplast genome were annotated to 293 KEGG pathways (Dautt-Castro *et al.,* 2015). Transcriptomic and proteomic analysis of a mixed mango sample with flesh and peel of mango variety "Zill" were reported, with 54,000 transcripts assembled and 2754 proteins matched to mango transcripts. This revealed critical pathways during fruit ripening. Comparative transcriptome analysis of unripe and mid-ripe mango fruit determined to ripen associated genes. Overall, there were 74,312 unique transcripts obtained and 127 pathways identified in the mango transcriptome by KEGG analysis, which were mainly involved in detoxification, carbon metabolism, ethylene biosynthesis and aromatic amino acid degradation. This study also revealed differences in softening associated genes and other nutritional characteristics.

In order to better understand the different transcriptomes of mango leaf and fruit, their individual expression profiles were determined in our research. Notwithstanding previous reports showing allergen sensitization to mango fruit, a more comprehensive assessment of allergen genes in mango leaf and fruit was undertaken in our study. We report a larger dataset of the transcriptome profiles and elucidate the significant allergens involved in different pathways. This information provides an essential platform for further allergen studies in mango. Producing lower allergenic cultivars through molecular biology should be useful in new mango breeding programs.

## 2. MATERIAL AND METHODS

### 2.1. RNA extraction

Mango leaf and fruit were collected and immediately frozen in liquid nitrogen, then grounded mechanically into the fine powder and stored at −80 °C for RNA extraction. Total RNA was isolated using an RNA Isolation Kit (Takara, Japan), according to the manufacturer's guidelines. The total RNA was suspended in RNase-free water, and RNA integrity and quality were assessed using an Agilent 2100 (Agilent, Santa Clara, CA, USA).

### 2.2. Library construction and illumina sequencing

The Magnetic Oligo (dT) beads (Invitrogen, USA) was performed to isolate poly (A) mRNA from total RNA. Then mRNA was randomly fragmented by the fragmentation buffer. Used these fragments as templates, cDNA was synthesized and purified. After purified cDNA, adapters were then connected. Suitable fragments were selected for PCR amplification and the library was then identified. Agilent 2100 Bioanalyzer was applied for the quantification and qualification of the library. Finally, high-throughput sequencing was conducted through the Illumina HiSeq 4000 (Illumina, San Diego, CA, USA) to generate 150-bp paired-end reads. The process of *de-nova* transcriptome sequence was shown in Figure 1. The raw sequence reads are available at the NCBI GEO (Gene Expression Omnibus) database under the accession number (GEO No. GSE142427).

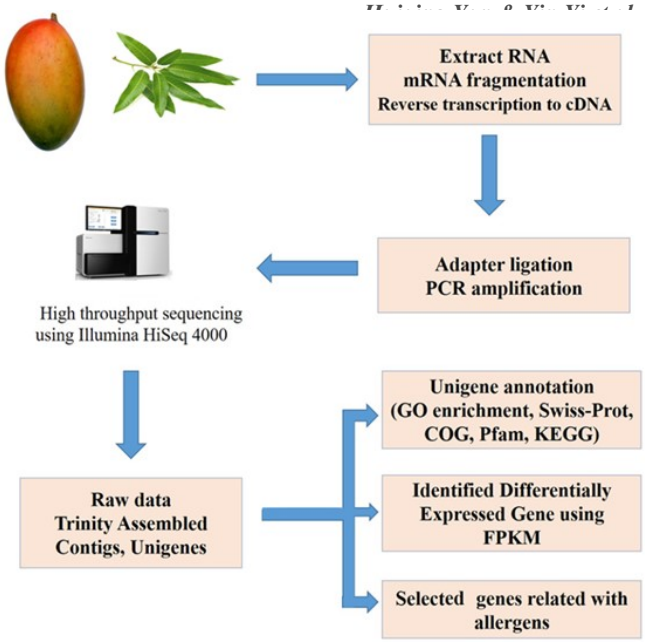


**Figure 1.** Transcriptome sequence of mango fruit and leaf and biological analysis.

### 2.3. Data processing, assembly and annotation of unigenes

The clean reads were selected from raw data by filtering out adaptor-only reads, reads containing more than 5% N bases unknown, and low-quality reads (reads containing more than 50% bases with Q-value ≤10) and used in the following analysis. Trinity assembly program was used to obtain data (Grabherr *et al.*, 2011). All unigenes were respectively compared with different databases. Clusters of Orthologous Groups of proteins database (COG, http://www.ncbi.nlm.nih.gov/COG/) (Tatusov *et al.*, 2000), NCBI non-redundant protein database (Nr, http://www.ncbi.nlm.nih.gov/), Swiss-Prot (http://www.expasy.ch/sprot) (UniProt Consortium, 2018), Gene Ontology (GO, http://www.geneontology.org/) (Ashburner *et al.*, 2000), Protein family (Pfam, http://pfam.xfam.org/) (Finn *et al.*, 2014), and KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/) with an E-value cut off of $10^{-5}$ (Kanehisa *et al.*, 2004). The best-aligning results were used to identify the sequence direction of the unigenes. If different databases conflicted, the results were prioritized in the order: Nr, Pfam, Swiss-Prot, KEGG, COG and GO. When transcripts did not align with any of the databases, EST Scan (http://myhits.isb-sib.ch/cgi-bin/estscan) was performed to xdecide its sequence direction.

**Table 1.** Primers used for real-time PCR analysis

| Gene ID | Forward primers | Reverse primers |
|---|---|---|
| *c61327.graph_c0* | CGACAAGAAGTGAAGACAGAGA | AGCCTGAGTGAAGATGAAGTTG |
| *c49299.graph_c0* | CTCTTCTCTTCACTCGCTTTCC | CCTTGGTGTTGCCGTCAGA |
| *c44502.graph_c0*, | GCTGAGTTATGGCGGTCAAG | TGGCTTCACAATTCAAGGCATT |
| *c64434.graph_c0* | CGCACGCTCCTTGAACTTC | ACGACGCCGACATTGACA |
| *c61297.graph_c0* | CCACACCACACCACAACAAC | CCAGTGATGACGACGACCTT |
| *c17261.graph_c0* | CCTGTCAACCACTGGAACTCA | AGACACAGCACTGCCATACC |
| *c64821.graph_c0* | ACACCGAAGAGATTGACAAGTC | GTTCCACAGGCACCGTAGT |
| *c47476.graph_c0* | ATGGTACTCGGCGATCTTGA | CACTCTGGCGGTCCTTCTAT |
| *c64956.graph_c0* | TCCGCAGCCAGTTCCATT | GCCAGCATTGTGTTACTCTCA |
| *c52048.graph_c0* | GAAGTGTTGGCGAGGAGGAT | AGCGTTGTTCAATAGCGGTTC |
| *β-actin* | ATCGCTGAGCACCTTCCAACA | CCAATCCTGACCTCTGACACTTCT |

### 2.4. Allergen analysis in mango leaf and fruit

All genes related to allergen were selected and clustered based on Nr annotation and made descriptions. The gene lengths of allergens were also listed. The expression level of all allergen genes in mango leaf and fruit were evaluated using FPKM (fragments per kilobase of transcript per million mapped reads).

### 2.5. Real-time RT-PCR analysis

The total RNA of mango flower, leaf and fruit were extracted using TAKARA Trizol Reagent according to the protocol of manufacturer. The extracted RNA was reversely transcribed using an RT-PCR Kit® with an oligo dT-adaptor primer. *β-actin* as an internal standard gene and genes were amplified. The primers for each allergic gene and *β-actin* were shown in Table 1. The fragments were separated on 1 % (w/v) agarose gel electrophoresis. Quantitative real-time PCR was performed in a LightCycler 480 instrument (Roche) with the FastStart DNA Master SYBR Green I kit. Amplification was performed for 30 cycles: denatured at 95°C for 30 s, annealed at 60 °C for 30 s, and extended at 74 °C for 1 min. The allergen expression levels relative to the control were estimated by calculating △△Ct and subsequently analyzed using $2^{-\triangle\triangle Ct}$ method.

### 2.6. Protein structure analysis

The amino acid sequence of allergens was obtained with ORF (https://www.ncbi.nlm.nih.gov/orffinder/) and used as templates by homology models with protein data bank (http://www.rcsb.org/pdb/home/home.do). Cluspro software was used for protein-protein docking. The assembly 3D structure was selected based on cluster scores that yield large clusters of docked structures with relatively low energies ($E=0.40E_{rep}+-0.40E_{att}+600E_{elec}+1.00E_{DARS}$) (Kozakov *et al.*, 2013). Discovery studio 4.5 software was used to obtain the sequence of docked proteins (Swellmeen *et al.*, 2017). The proteins were aligned with ESPript 3.0 webserver (http://espript.ibcp.fr/ESPript/ESPript/) based on the crystal structure (https://swissmodel.expasy.org/).

### 2.7. Immunodetection assays

The identified allergen was constructed with recombinant prokaryotic expression vector pET28a (+). Then, we transformed the recombinant into *E.coli* BL21(DE3) to over-expression the allergen and purified the recombinant protein by nickel chelating chromatography. Finally, a total of 30μg protein were determined and separated on 10% SDS-PAGE then transferred onto polyvinylidene difluoride membranes (Millipore, MA). After blocked with 5% non-fat milk in TBST (0.1%) for 1 h, then the membranes were respectively incubated overnight with IgE (SouthernBiothech, Cat. No. B312E8, 1:200 dilution) for 2 h at 25°C. After washed six times with TBST, the membranes were hybridized with goat anti-human IgE conjugated with horseradish peroxidase (SouthernBiothech, Cat.No.1110-05) antibody 1: 5, 000 dilution for 45 min at 25 °C. After washed six times with TBST. The membrane finally was incubated with ECL detection kits (Thermo Scientific

Pierce) for 2 min and exposed to X-ray film then IgE-binding components were revealed by enhanced chemiluminescence. Prestained protein molecular weight marker (Thermo Scientific Fermentas, SM0671) was used.

## 3. RESULTS

### 3.1. Characterization of Mango Transcriptome and *de novo* Assembly

Mango, as a member of the family *Anacardiaceae*, is an allotetraploid fruit tree with a small genome size of about 450 Mbp (Dautt-Castro *et al.*, 2015). A new mango transcriptome was assembled from 8.36 Gbp of sequence data using Trinity software, which generated 11,751,123 contigs that were assembled into 99,328 unigenes with 16,848 unigenes above 1000 bp and an average length of 1357 bases. A total of 230,242 transcripts were annotated using public protein databases, with a cut-off E-value above $10^{-5}$. The length ranges of 200–300, 300–500, 500–1000, 1000–2000 and >2000 bp represented 20.08, 15.39, 15.89, 23.10 and 25.54%, respectively. Transcripts were also analyzed in the KEGG database, and a total of 15,520 unigenes were assigned to 327 KEGG pathways. The numbers of unigenes annotated using Non-redundant (Nr), Swiss-Prot, GO and COG databases were respectively 46,030, 24,227, 27,295 and 14,023. Additionally, 6568 unigenes were annotated in all databases (Figure 2).

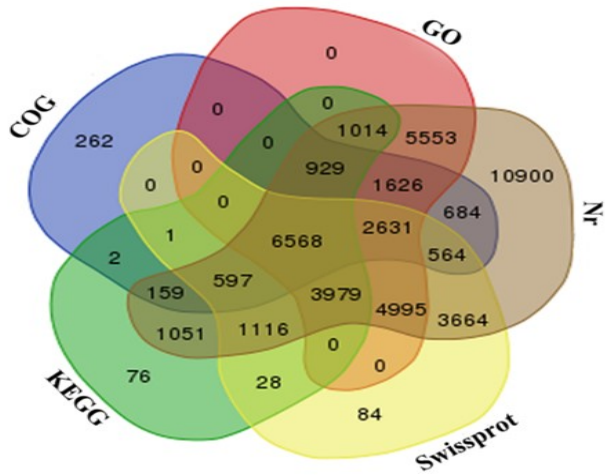


**Figure 2.** Venn diagram of mango unigenes annotated according to the NCBI Nr, Swiss-Prot, COG and KEGG databases and classified into GO. The overlapped unigenes are indicated in the intersections.

### 3.2. Expression Profile of Mango Fruit and Leaf

Using Trinity software, a total of 7369 genes were identified as differentially expressed between mango leaf and fruit, with a predicted FDR < 0.05: 4558 up-regulated and 2811 down-regulated. We determined if particular GO terms and KEGG pathways were enriched in the different genes compared with the complete transcriptome. The GO enrichment terms were classified into three categories: biological process, cellular components and molecular function. In terms of biological process, these different genes were mainly involved in the metabolic process, the cellular process, and the single-organism process based on sequence homology. Cell part, organelle, membrane, macromolecular complex and extracellular region were the most significant cellular components. In terms of molecular function, catalytic activity, binding, transporter activity and electron carrier activity were noted. Most genes were involved in photosynthesis, starch and sucrose metabolism, plant hormone signal transduction, carbon metabolism and ribosomes (Figure 3A).

As reported in fruit ripening, a wide range of genes were up-regulated in mango fruit, such as those involved in the oxidation-reduction process and ATP binding nucleus. These genes were generally involved in essential carbohydrate and secondary metabolite accumulation during fruit maturation with COG annotation function classification of the consensus sequence. Translation, ribosomal structure and biogenesis, carbohydrate transport and metabolism, energy production and conversion, secondary metabolite biosynthesis, transport and catabolism were significant processes in which genes were up-regulated (Figure 3B). There were also many other genes predicted to participate in general function only, which will be very useful in further research where their expression profiles will be assayed at the transcriptional level for different developmental stages.

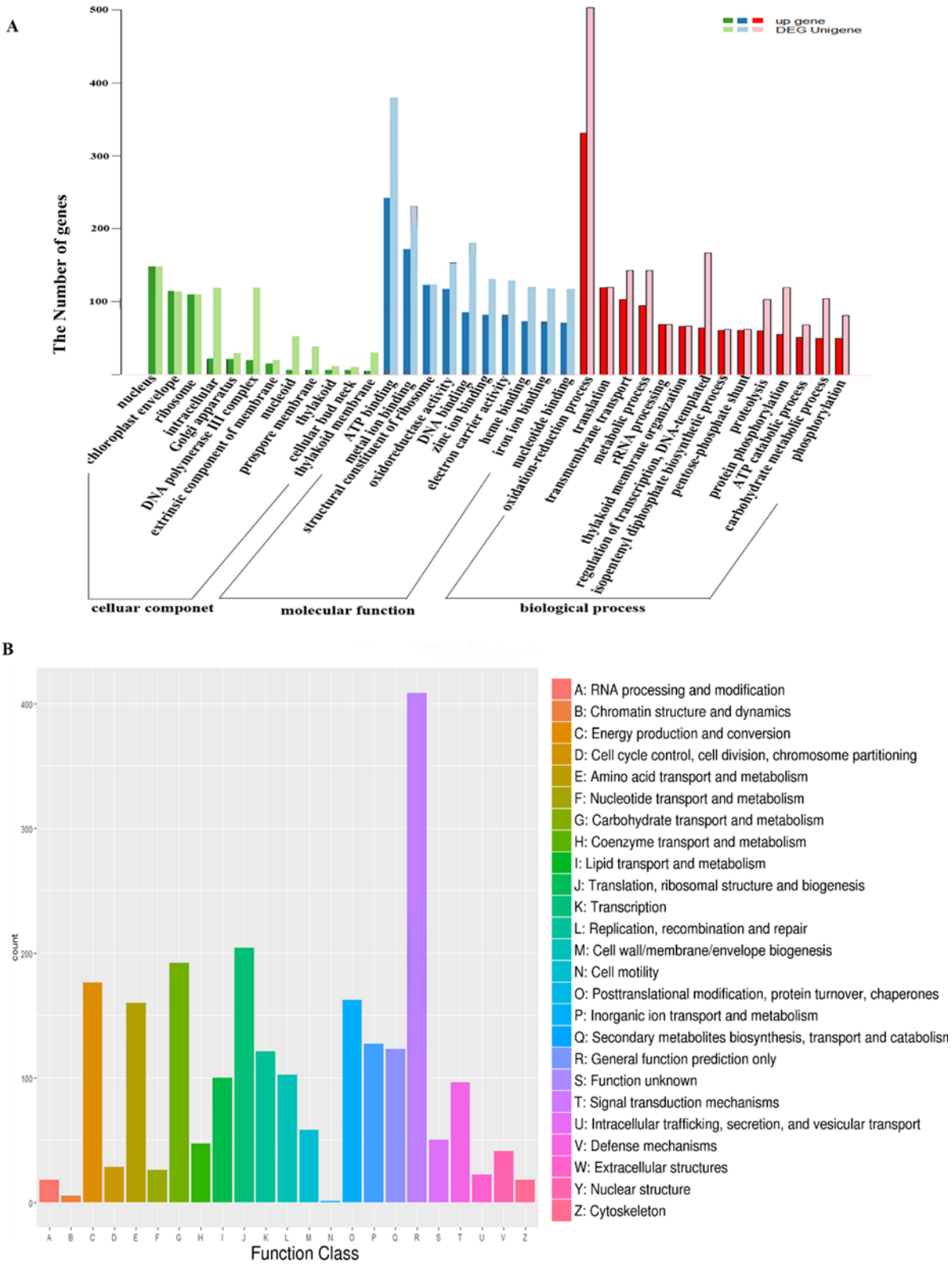


**Figure 3.** The GO (including biological process, cellular component and molecular function) and COG enrichment of up-regulated genes in mango fruit. The x-axis indicates the function class.

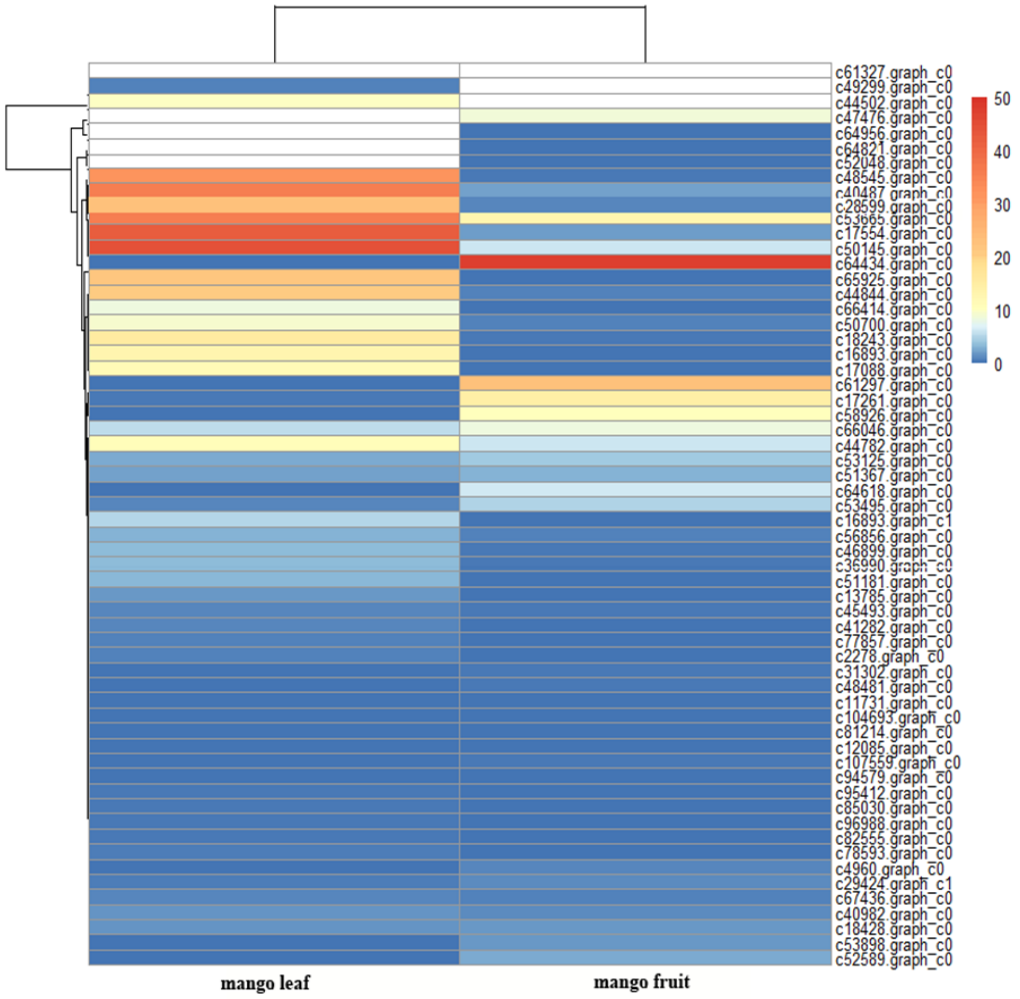


**Figure 4.** The relative expression heatmap of 60 genes related to allergens identified from the massive amount of transcriptome sequencing data in mango leaf and fruit with RNA-Seq (shown in white for the reason the expression value was higher than the figure legend).

### 3.3. Genes Associated with Allergens in Mango Fruit and Leaf

A total of sixty-six allergens were obtained at the transcription level using the databases (Table 2). Pollen allergen was the major component. The genes in mango leaf and fruit could be BLASTed in different species and encoded expansin-like proteins. Some other pollen allergens, Che a 1 and Ole e 10, were also determined in mango. Bet v type allergens were also present, including allergenic isoflavone reductase-like protein Bet v 6.0102, pathogenesis-related protein Bet v I and calcium-binding protein Bet v 3-like protein. Some allergens of Pru ar 1-like, including c45493.graph_c0, c51367.graph_c0, c94579.graph_c0 and c40982.graph_c0, also belonged to the pathogenesis-related protein Bet v I family. Allergen Alt was a type of NADPH-dependent FMN reductase. Some other kinds of allergens were also identified in mango-Ana o 2, chitinase, profilin, Can a 1, Mal d 1, Hsp90/Hsp, Pis v 2.0201, Cla h and aldehyde dehydrogenase, which are all potentially harmful to human health.

Sixty allergen genes were expressed in mango leaf and fruit, and their relative expression patterns are shown in Figure4. The genes *c61327.graph_c0*, *c49299.graph_c0* and *c44502.graph_c0* were highly expressed in mango fruit, followed by *c64434.graph_c0* and *c61297.graph_c0*. The three top genes were those for pollen allergen Che a 1, aldehyde dehydrogenase and the pathogenesis-related protein Bet v I family. There were 24 allergen genes mainly expressed in mango fruit and 36 others were highly expressed in leaf (Figure 4). Most of these genes encoded pollen allergen expansion-like proteins.

**Table 2.** The 66 genes related to allergens identified in mango leaf and fruit

| Description | GeneID | Gene length (bp) |
|---|---|---|
| Chitinase [*Mangifera indica*] | c28599.graph_c0; c64821.graph_c0 | 467;1404 |
| Pollen allergen expansin [*Citrus sinensis]* | c29424.graph_c1; c48545.graph_c0; c50700.graph_c0 ; c16893.graph_c1; c17261.graph_c0; | 208; 1299; 1261 ; 385; 1032 |
| Pollen allergen expansin [*Jatropha curcas*] | c44782.graph_c0; c17088.graph_c0 c107559.graph_c0; c56856.graph_c0; c18243.graph_c0 | 685; 769 931; 1066 |
| Pollen allergen expansin [*Gossypium hirsutum*] | c78593.graph_c0 | 202 |
| Pollen allergen expansin [*Theobroma cacao*] | c17554.graph_c0; c46399.graph_c0 | 1347;591 |
| Pollen allergen expansin [*Gossypium arboreum*] | c47476.graph_c0; c16893.graph_c0 | 1934; |
| Pollen allergen expansin [*Theobroma cacao*] | c50145.graph_c0 | 1178 |
| Pollen allergen expansin [*Populus trichocarpa*] | c64956.graph_c0 | 1229 |
| Pollen allergen expansin [*Populus euphratica*] | c53495.graph_c0 | 988 |
| Pollen allergen expansin [*Glycine max*] | c96988.graph_c0 | 211 |
| Pollen allergen expansin [*Sesamum indicum*] | c52048.graph_c0; c44844.graph_c0 | 851; |
| Pollen allergen expansin [*Ricinus communis*] | c2278.graph_c0; c66414.graph_c0 | 324; 351 |
| Pollen allergen expansin [*Morus notabilis*] | c40487.graph_c0 | 646 |
| Pollen allergen Che a 1 [*Citrus sinensis*] | c61327.graph_c0 | 963 |
| Pollen allergen Ole e 10 [*Citrus sinensis*] | c51181.graph_c0 | 662 |
| Pollen allergen Ole e 10 [*Fragaria vesca subsp. vesca*] | c65925.graph_c0 | 661 |
| Allergenic Bet v 6.0102 [*Betula pendula*] | c88876.graph_c0; c82555.graph_c0 | 247 |
| Pathogenesis-related Bet v I [*Theobroma cacao*] | c67436.graph_c0 | 405 |
| Pathogenesis-related Bet v I [*Alnus glutinosa*] | c85030.graph_c0 | 212 |
| Calcium-binding Bet v 3-like [*Citrus sinensis*] | c13785.graph_c0 | 592 |
| Allergen Pru ar 1-like [*Nelumbo nucifera*] | c45493.graph_c0 | 767 |
| Allergen Pru ar 1-like [*Eucalyptus grandis*] | c51367.graph_c0 | 790 |
| Allergen Pru ar 1-like [*Vitis vinifera*] | c94579.graph_c0 | 263 |
| Allergen Pru av 1 [*Vitis hybrid cultivar*] | c40982.graph_c0 | 796 |
| Allergen Alt a 7-like [*Citrus sinensis*] | c53665.graph_c0; c66046.graph_c0 | 1547; 374 |
| Allergen Alt a [*Jatropha curcas*] | c44502.graph_c0 | 877 |
| Allergen Alt a 7 [*Alternaria alternata*] | c91727.graph_c0 | 252 |
| Allergen Alt a [*Neofusicoccum parvum]* | c4960.graph_c0; c58926.graph_c0 | 594; 925 |
| Allergen Alt a [*Ricinus communis*] | c53125.graph_c0 | 1361; |
| Alternaria Alternata Allergen Alt A 1 | c5477.graph_c0 | 252 |
| Allergen Ana o 2 [*Anacardium occidentale*] | c77857.graph_c0 | 417 |
| Can a 1 allergen protein [*Sphaerulina musiva*] | c104693.graph_c0 | 247 |
| Allergen Hsp90/Hsp1 [*Sphaerulina musiva*] | c64618.graph_c0 | 2513 |
| Pis v 2.0201 allergen [*Pistacia vera*] | c41282.graph_c0 | 751 |
| Allergen Cla h, Aldehyde dehydrogenase [*Cladosporium herbarum*] | c81214.graph_c0; c64434.graph_c0; c31302.graph_c0 c14312.graph_c0; c12085.graph_c0; c69695.graph_c0 c53898.graph_c0;c49299.graph_c0; c52589.graph_c0 c11731.graph_c0; c48481.graph_c0; c23071.graph_c0 c61297.graph_c0; c18428.graph_c0 | 249; 1773; 247 228; 215; 203 248; 636; 673 252; 352; 333 1083; 390 |
| Major allergen Mal d 1 [*Malus domestica*] | c95412.graph_c0 | 232 |
| Profilin [*Glycine soja*] | c36990.graph_c0 | 366 |

### 3.4. The analysis of Allergen Expressions

Allergen genes that were highly expressed in mango leaf or fruit were selected. These allergens were determined with real-time PCR in three different mango tissues: flower, leaf and fruit (Figure 5). Their expressions basically corresponded with RNA-Seq results. Expressions of *c44502.graph_c0*, *c49299.graph_c0*, *c61297.graph_c0*, *c52048.graph_c0*, *c17261.graph_c0* and *c61327.graph_c0* were higher in fruit than the leaf. Other genes were expressed more highly in leaf than fruit. Most genes were highly expressed in flowers. Only *c64821.graph_c0* was highly expressed in leaf.

### 3.5. Protein Structure Analysis

The allergen c61327.graph_c0, which is highly expressed in fruit, was used as a template to obtain homologous protein structure in the RCSB PDB bank (PDB: 4z8w). When searched using BLAST in NCBI, this matched Pollen Ole e 1 (*Theobroma cacao*) with a high level of amino acid identity (71%). In the predicted 3D structure of c61327.graph_c0, the human IgE-binding simulation was at this epitope (74-SFRQEVKTEKHGEFKVHLPFSVSEHV-99) (Figure 6A). The immunodetection using IgE (Fig. 6C) disclosed a main binding band of around 35 KDa. The molecular weight was approximately in accordance with c61327.graph_c0, which were 33.53 KDa respectively. Thus, we concluded that c61327.graph_c0 showed the reaction to IgE, which could cause allergic responses in the human body. The docking protein sequences are shown in Figure 6B. c61327.graph_c0 was aligned with the Pollen Ole e 1 allergen/ extension protein of *T. cacao* (GenBank: EOY03810.1), *Cephalotus follicularis* (GenBank: GAV60102.1), *Corchorus olitorius* (GenBank: OMO89486.1) and *Corchorus capsularis* (GenBank: OMP03537.1) using ESPript 3.0. Based on the high similarity of protein sequences, cross-sensitivity reactions between mango and other species should be noticed.

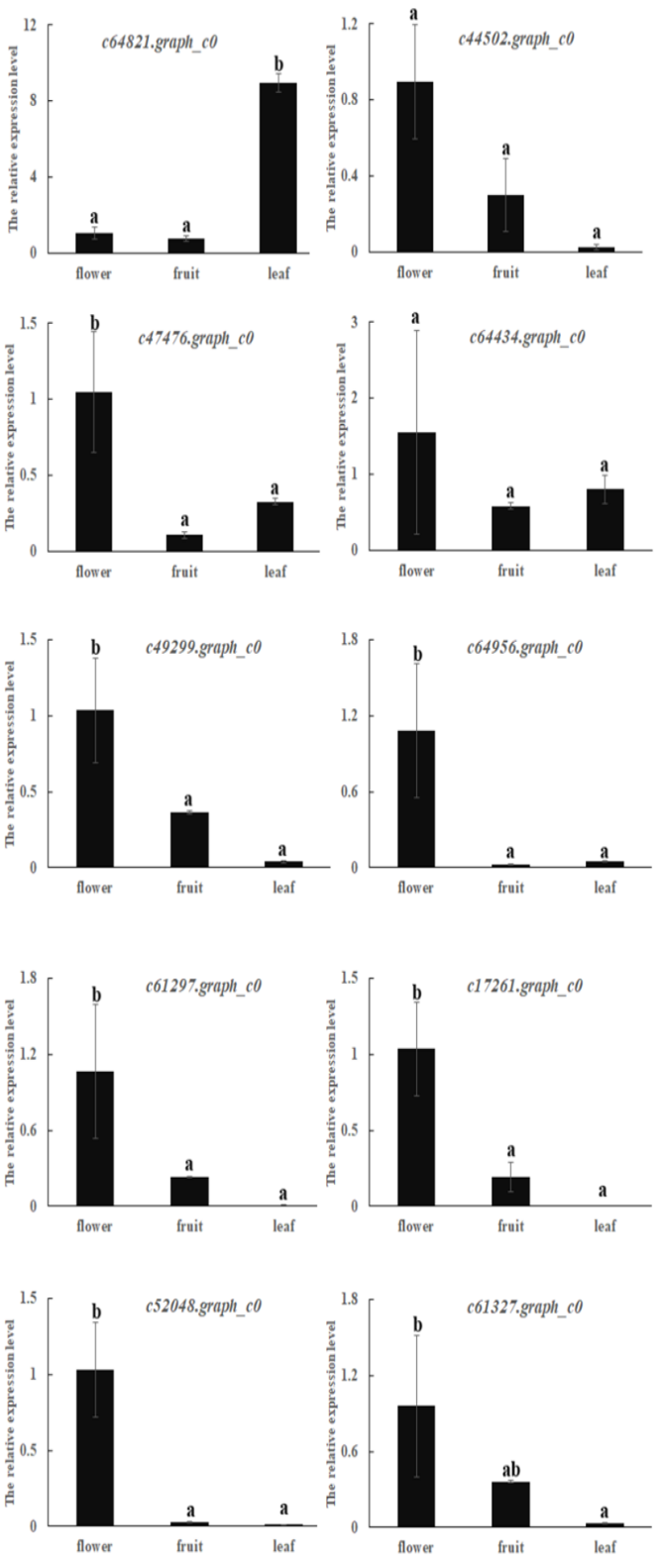


**Figure 5.** Relative expressions of ten allergens in mango flower, fruit and leaf using real-time PCR.

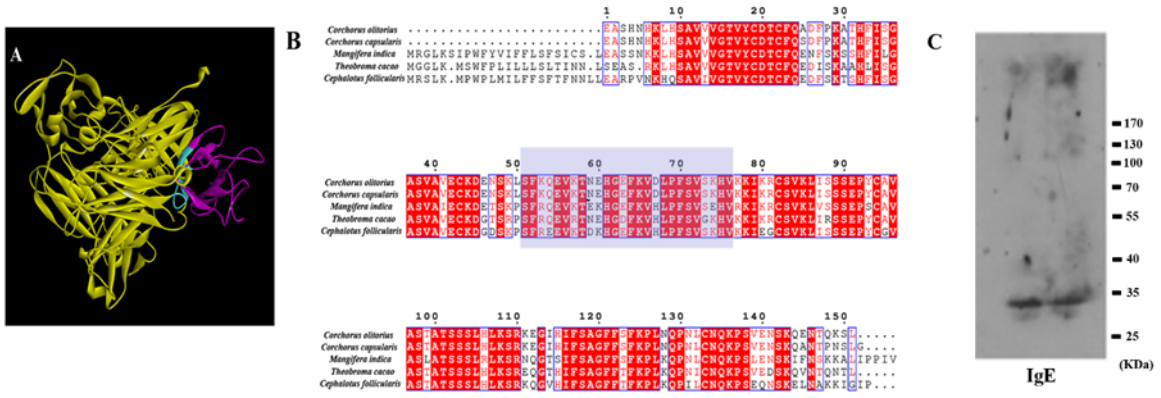


**Figure 6.** Prediction of human IgE-Fc (cyan) (PDB: 5 mol) bound to the epitopes of c61327.graph_c0 (PDB: 4z8w) allergen. (A) The allergen (purple) bound to IgE-Fc (yellow); image created using Discovery studio 4.5 software; (B) amino acid alignment among mango, *Theobroma cacao*, *Cephalotus follicularis*, *Corchorus olitorius* and *Corchorus capsularis;* image created using ESPript 3.0. The shadow is the putative IgE epitope (74-SFRQEVKTEKHGEFKVHLPFSVSEHV-99). (C) IgE-binding pattern with sensitization to c61327.graph_c0 (each was repeated twice) by western-blotting analysis.

## 4. DISCUSSION

Illumina mRNA sequencing technology is an efficient technology to characterize the transcriptome profile of mango (Sherman *et al.*, 2015), and has been used in studies on grape (Tu *et al.*, 2016), sweet orange (Hu *et al.*, 2016) and pineapple (Sharma *et al.*, 2017), producing data on differentially expressed genes or new genes annotated with potential or novel pathways. Therefore, Illumina sequencing of mRNA is a priority for gene function research in mango. There have been several recent studies on mango fruit development and fruit quantity (Pandit *et al.*, 2010). However, as far as we know, few reports on the use of the RNA-Seq technique to identify different transcriptome profiles of mango leaf and fruit. We obtained many genes involved in different metabolic pathways during mango leaf and fruit development. We generated 28 million sequence reads corresponding to 8.36 Gb of raw sequence data and obtained 99,328 unique sequences with 16,848 of >1 kb, of which 47,949 unigenes were annotated with different databases, which was relatively higher than those obtained for other fruits using transcriptome sequencing and assembly. Our results provide the most extensive published sequencing resource for mango.

Out of 47,949 transcripts, 46,303 (96.57%) were successfully aligned within the Nr database, 3.43% of transcripts could not be Blasted to known genes because of genome limitation and lack of EST information in mango. Of the annotated transcripts, some unigenes were classified as 'hypothetical protein', 'predicted protein' and 'putative', which did not receive a confirmative annotation. It was challenging to identify function and classification. Therefore, these genes should receive more attention and analysis in classical molecular biological experiments in order to determine their potential roles, which may be critical to allergen pathways.

Sixty-six allergen genes were obtained at the transcriptome level, and mainly included pollen allergens, the pathogenesis-related protein Bet v I family, profilins, chitinase class I and Allergen Alt a 1. Pollen allergen expansin-like proteins are significant allergens, exist in many species and are considered pan-allergens– they are responsible for cross-reactions between food and pollen. The pathogenesis-related protein Bet v I family and chitinase are involved in plant defense against fungi and bacteria, as well as hydrophobic protection of chitin in animals (Diaz-Perales *et al.*, 1999). Profilins are actin mono-

mer binding proteins and regulate the cytoskeleton in higher plants (Song *et al.*, 2008). Allergen Alt is a type of NADPH-dependent FMN reductase, which partially corresponds to the Man i allergen. The Alt a 1 is a species-specific molecular marker in citrus that has been strongly associated with allergenicity, and may have potential in immunotherapy against mango allergies (Moreno *et al.*, 2016; Gabriel *et al.*, 2017).
Despite many allergens existing in mango leaf and fruit, their expressions differed. Expressions of the ten allergenic genes in three different tissues were compared using qRT-PCR, which corresponded very well with RNA-Seq in mango and leaf. However, allergens in flower were remarkably higher than any other tissue. This indicates that people with strong, sensitive, allergic reactions after eating mango should avoid touching mango flowers.

Increasingly, low-allergenic cultivars are bred and selected using molecular approaches, which can be used to lower the expression of major allergens. Geneticists have identified low allergenic apple cultivars (Kootstra *et al.*, 2007). A similar approach was initiated in peach through collaboration between China and Europe (Brenna *et al.*, 2004). A total of 66 potential allergens in mango fruit were assessed in this study, and major allergens can be screened. So, this will aid the breeding of lower allergenic cultivars through molecular biology. Similarities in protein structure can be used to predict protein function; c61327.graph_c0 was assigned the function of Pollen Ole e 1 (*T. cacao*) based on similar amino acid sequences and proved to show interaction with human IgE by immunoblotting analysis. However, the IgE epitope of other allergens should be further investigated. These results are a step toward understanding the allergens of mango fruit and in developing new cultivars with enhanced health properties.

## CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships. All the authors declare that they have no conflict of interest.

## ACKNOWLEDGEMENTS

This work was supported by grants from the National Natural Science Foundation of China (Grant No. 31660554), Science Foundation of Guizhou Provinces (LH zi (2015) 7772), and Guizhou Normal University Dr. Scientific Research Fund (Grant No.0514157 and No.0514156).

## REFERENCES

[1] Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin and G. Sherlock, 2000. Gene ontology: Tool for the unification of biology. The gene ontology consortium. Nat Genet, 25(1): 25-29. PMid:10802651 View Article PubMed/NCBI

[2] Brenna, O.V., E.A. Pastorello, L. Farioli, V. Pravettoni and C. Pompei, 2004. Presence of allergenic proteins in different peach (prunus persica) cultivars and dependence of their content on fruit ripening. J Agric Food Chem, 52(26): 7997-8000. PMid:15612787 View Article PubMed/NCBI

[3] Dautt-Castro, M., A. Ochoa-Leyva, C.A. Contreras-Vergara, M.A. Pacheco-Sanchez, S. Casas-Flores, A. Sanchez-Flores, D.N. Kuhn and M.A. Islas-Osuna, 2015. Mango (mangifera indica L.) cv. Kent fruit mesocarp de novo transcriptome assembly identifies gene families important for ripening. Front Plant Sci, 6: 62. PMid:25741352 View Article PubMed/NCBI

[4] Diaz-Perales, A., C. Collada, C. Blanco, R. Sanchez-Monge, T. Carrillo, C. Aragoncillo and G. Salcedo, 1999. Cross-reactions in the latex-fruit syndrome: A relevant role of chitinases but not of complex asparagine-linked glycans. J Allergy Clin Immunol, 104(3 Pt 1): 681-687. 70342-8 View Article

[5] Dube, M., K. Zunker, S. Neidhart, R. Carle, H. Steinhart and A. Paschke, 2004. Effect of technological processing on the allergenicity of mangoes (mangifera indica L.). J Agric Food Chem, 52(12): 3938-3945. PMid:15186120 View Article PubMed/NCBI

[6] Fasoli, E. and P.G. Righetti, 2013. The peel and pulp of mango fruit: A proteomic samba. Biochim Biophys Acta, 1834(12): 2539-2545. PMid:24056186 View Article PubMed/NCBI

[7] Finn, R.D., A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S.R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E.L. Sonnhammer, J.

Tate and M. Punta, 2014. Pfam: The protein families database. Nucleic Acids Res, 42(Database issue): D222-230. PMid:24288371 View Article PubMed/NCBI

[8] Gabriel, M.F., N. Uriel, F. Teifoori, I. Postigo, E. Sunen and J. Martinez, 2017. The major alternaria alternata allergen, alt a 1: A reliable and specific marker of fungal contamination in citrus fruits. Int J Food Microbiol, 257: 26-30. PMid:28633053 View Article PubMed/NCBI

[9] Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev, 2011. Full-length transcriptome assembly from rna-seq data without a reference genome. Nat Biotechnol, 29(7): 644-652. PMid:21572440 View Article PubMed/NCBI

[10]Hassan, A.K. and Y.P. Venkatesh, 2015. An overview of fruit allergy and the causative allergens. Eur Ann Allergy Clin Immunol, 47(6): 180-187. View Article

[11]Hoang, V.L., D.J. Innes, P.N. Shaw, G.R. Monteith, M.J. Gidley and R.G. Dietzgen, 2015. Sequence diversity and differential expression of major phenylpropanoid-flavonoid biosynthetic genes among three mango varieties. BMC Genomics, 16: 561. Available from PMid:26220670 View Article PubMed/NCBI

[12]Hong, K., D. Gong, L. Zhang, H. Hu, Z. Jia, H. Gu and K. Song, 2016. Transcriptome characterization and expression profiles of the related defense genes in postharvest mango fruit against colletotrichum gloeosporioides. Gene, 576(1 Pt 2): 275-283. PMid:26496007 View Article PubMed/NCBI

[13]Hu, Y., S. Duan, Y. Zhang, D. Shantharaj, J.B. Jones and N. Wang, 2016. Temporal transcription profiling of sweet orange in response to ptha4-mediated xanthomonas citri subsp. Citri infection. Phytopathology, 106(5): 442-451. PMid:26780431 View Article PubMed/NCBI

[14]Kanehisa, M., S. Goto, S. Kawashima, Y. Okuno and M. Hattori, 2004. The kegg resource for deciphering the genome. Nucleic Acids Res, 32 (Database issue): D277-280. PMid:14681412 View Article PubMed/NCBI

[15]Kootstra, H.S., B.J. Vlieg-Boerstra and A.E. Dubois, 2007. Assessment of the reduced allergenic properties of the santana apple. Ann Allergy Asthma Immunol, 99(6): 522-525. 60381-X View Article

[16]Kozakov, D., D. Beglov, T. Bohnuud, S.E. Mottarella, B. Xia, D.R. Hall and S. Vajda, 2013. How good is automated protein docking? Proteins, 81(12): 2159-2166. PMid:23996272 View Article PubMed/NCBI

[17]Liu, F., J.B. Wu, R.L. Zhan and X.C. Ou, 2016. Transcription profiling analysis of mango-fusarium mangiferae interaction. Front Microbiol, 7: 1443. View Article

[18]Luria, N., N. Sela, M. Yaari, O. Feygenberg, I. Kobiler, A. Lers and D. Prusky, 2014. De-novo assembly of mango fruit peel transcriptome reveals mechanisms of mango response to hot water treatment. BMC Genomics, 15: 957. PMid:25373421 View Article PubMed/NCBI

[19]Moreno, A., F. Pineda, J. Alcover, D. Rodriguez, R. Palacios and E. Martinez-Naves, 2016. Orthologous allergens and diagnostic utility of major allergen alt a 1. Allergy Asthma Immunol Res, 8(5): 428-437. PMid:27334781 View Article PubMed/NCBI

[20]Noratto, G.D., M.C. Bertoldi, K. Krenek, S.T. Talcott, P.C. Stringheta and S.U. Mertens-Talcott, 2010. Anticarcinogenic effects of polyphenolics from mango (mangifera indica) varieties. J Agric Food Chem, 58(7): 4104-4112. PMid:20205391 View Article PubMed/NCBI

[21]Oka, K., F. Saito, T. Yasuhara and A. Sugimoto, 2004. A study of cross-reactions between mango contact allergens and urushiol. Contact Dermatitis, 51(5-6): 292-296. PMid:15606656 View Article PubMed/NCBI

[22]Pandit, S.S., R.S. Kulkarni, A.P. Giri, T.G. Kollner, J. Degenhardt, J. Gershenzon and V.S. Gupta, 2010. Expression profiling of various genes during the fruit development and ripening of mango. Plant Physiol Biochem, 48(6): 426-433. PMid:20363641 View Article PubMed/NCBI

[23]Paschke, A., H. Kinder, K. Zunker, M. Wigotzki, H. Steinhart, R. Wessbecher and I. Vieluf, 2001. Characterization of cross-reacting allergens in mango fruit. Allergy, 56(3): 237-242. PMid:11251404 View Article PubMed/NCBI

[24]Renner, R., C. Hipler, R. Treudler, W. Harth, A. Suss and J.C. Simon, 2008. Identification of a 27 kda protein in patients with anaphylactic reactions to mango. J Investig Allergol Clin Immunol, 18(6): 476-481. PMid:19123442 PubMed/NCBI

[25]Sharma, A., C.M. Wai, R. Ming and Q. Yu, 2017. Diurnal cycling transcription factors of pineapple revealed by genome-wide annotation and global transcriptomic analysis. Genome Biol Evol, 9(9): 2170-2190. PMid:28922793 View

Article PubMed/NCBI

[26]Sherman, A., M. Rubinstein, R. Eshed, M. Benita, M. Ish-Shalom, M. Sharabi-Schwager, A. Rozen, D. Saada, Y. Cohen and R. Ophir, 2015. Mango (mangifera indica L.) germplasm diversity based on single nucleotide polymorphisms derived from the transcriptome. BMC Plant Biol, 15: 277. PMid:26573148 View Article PubMed/NCBI

[27]Song, J., H. Zhang, Z. Liu and P. Ran, 2008. Mango profilin: Cloning, expression and cross-reactivity with birch pollen profilin bet v 2. Mol Biol Rep, 35(2): 231-237. PMid:17417721 View Article PubMed/NCBI

[28]Srivastava, S., R.K. Singh, G. Pathak, R. Goel, M.H. Asif, A.P. Sane and V.A. Sane, 2016. Comparative transcriptome analysis of unripe and mid-ripe fruit of mangifera indica (var. "Dashehari") unravels ripening associated genes. Scientific Reports, 6: 32557. PMid:27586495 View Article PubMed/NCBI

[29]Swellmeen, L., R. Shahin, Y. Al-Hiari, A. Alamiri, A. Hasan and O. Shaheen, 2017. Structure based drug design of pim-1 kinase followed by pharmacophore guided synthesis of quinolone-based inhibitors. Bioorg Med Chem, 25(17): 4855-4875. PMid:28760531 View Article PubMed/NCBI

[30]Tatusov, R.L., M.Y. Galperin, D.A. Natale and E.V. Koonin, 2000. The cog database: A tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res, 28(1): 33-36. PMid:10592175 View Article PubMed/NCBI

[31]Tsai, W.C., T.C. Wu, B.L. Chiang and H.W. Wen, 2017. Cloning, expression, and purification of recombinant major mango allergen man i 1 in escherichia coli. Protein Expr Purif, 130: 35-43. PMid:27350535 View Article PubMed/NCBI

[32]Tu, M., X. Wang, T. Feng, X. Sun, Y. Wang, L. Huang, M. Gao, Y. Wang and X. Wang, 2016. Expression of a grape (vitis vinifera) bzip transcription factor, vlbzip36, in arabidopsis thaliana confers tolerance of drought stress during seed germination and seedling establishment. Plant Sci, 252: 311-323. PMid:27717468 View Article PubMed/NCBI

[33]UniProt Consortium, T., 2018. Uniprot: The universal protein knowledgebase. Nucleic Acids Res, 46(5): 2699. PMid:29425356 View Article PubMed/NCBI

[34]Vargas Correa, J.B., L. Sanchez Solis, J.A. Farfan Ale, H. Noguchi, M.T. Moguel Banos and M.I. Vargas de la Pena, 1991. Allergological study of pollen of mango (magnifera indica) and cross reactivity with pollen of piru (schinus molle). Rev Alerg, 38(5): 134-138. PMid:1792479 PubMed/NCBI

[35]Weinstein, S., S. Bassiri-Tehrani and D.E. Cohen, 2004. Allergic contact dermatitis to mango flesh. Int J Dermatol, 43(3): 195-196. PMid:15009389 View Article PubMed/NCBI

[36]Wellhausen, A., B. Schoning, A. Petersen and S. Vieths, 1996. Ige binding to a new cross-reactive structure: A 35 kda protein in birch pollen, exotic fruit and other plant foods. Z Ernahrungswiss, 35 (4): 348-355. PMid:9000332 View Article PubMed/NCBI

[37]Wu, H.X., H.M. Jia, X.W. Ma, S.B. Wang, Q.S. Yao, W.T. Xu, Y.G. Zhou, Z.S. Gao and R.L. Zhan, 2014. Transcriptome and proteomic analysis of mango (mangifera indica linn) fruits. J Proteomics, 105: 19-30. PMid:24704857 View Article PubMed/NCBI

[38]Wu, T.C., T.C. Tsai, C.F. Huang, F.Y. Chang, C.C. Lin, I.F. Huang, C.H. Chu, B.H. Lau, L. Wu, H.J. Peng and R.B. Tang, 2012. Prevalence of food allergy in taiwan: A questionnaire-based survey. Intern Med J, 42(12): 1310-1315. PMid:22530688 View Article PubMed/NCBI