doi: 10.17221/43/2016-CJGPB

Comparative Transcriptomic Analysis Reveals a Series of Single Nucleotide Polymorphism between Red- and White-fleshed Loquats (*Eriobotrya japonica*)

*Shuxia SUN[1,2], Jing LI[1,2], Dong CHEN[1,2], Hongjiang XIE[1,2], Meiyan TU[1,2], Liu JIA[1,2] and Guoliang JIANG[1,2]**

[1]Horticulture Research Institute, Sichuan Academy of Agricultural Sciences, Chengdu, Sichuan Province, P.R. China; [2]Southwestern Key Laboratory of Horticultural Crops Biology and Germplasm Enhancement, Ministry of Agriculture, Chengdu, Sichuan Province, P.R. China

**Corresponding author: sshuxia@163.com*

Abstract

Sun S., Li J., Chen D., Xie H., Tu M., Jia L., Jiang G. (2017): Comparative transcriptomic analysis reveals a series of single nucleotide polymorphism between red- and white-fleshed loquats (*Eriobotrya japonica*). Czech J. Genet. Plant Breed., 53: 97–106.

Loquat (*Eriobotrya japonica*) is an economically important crop and red-fleshed cultivars have a much higher carotenoid content than white-fleshed cultivars. We used Illumina RNA-seq technology to gain a global overview of the loquat transcriptome from a mixture of fruit samples at different developmental stages for both red-fleshed and white-fleshed loquat. A total of 94.98 million paired-end short reads were obtained and 61 586 unigenes were generated from *de novo* assembly with an average length of 817 bp. Among these unigenes, 44 710 unigenes were annotated by blast against Nr, Swissprot, GO, COG and KEGG databases. For these annotated unigenes, 123 biosynthesis pathways were predicted by mapping these unigenes to the reference canonical pathways and 41 unigenes were predicted to be involved in carotenoid biosynthesis. RT-qPCR analysis showed that the expression level of the *LCYB* gene was higher in red-fleshed loquat and the *CRTRB* gene had a higher expression level in white-fleshed loquat. Comparative analysis of the two transcriptomes revealed 2396 single nucleotide polymorphisms (SNPs) between red- and white-fleshed loquats. The majority of SNPs identified between the two loquat cultivars were nonsense mutations and one out of eleven SNPs in candidate genes involved in carotenoid biosynthesis was a sense mutation. This suggests that the analysis based on transcriptomes can reveal key genes related to the carotenoid biosynthesis and more carotene in red-fleshed loquat cultivars may result from both more carotene produced by the higher expression of *LCYB* genes and less carotene converted because of the low expression of the *CRTRB* gene. All these results from the transcriptome analysis will be useful for the elucidation of genetic differences between red- and white-fleshed loquat fruits and further functional analysis for genes responsible for carotenoid accumulation.

Keywords: carotenoid biosynthesis; expression analysis; SNP; transcriptome

Loquat (*Eriobotrya japonica*, 2*n* = 34), belonging to the dicotyledonous family Rosaceae is an ancient fruit, native to China. Currently, loquat is grown commercially for its fruit in a broad range of climate zones, including Asia, Europe, South Africa, and Central and South America. Loquat fruit is considered as a delicacy and has a high carotenoid content in the red-fleshed loquat, while its leaf is also used

Shuxia Sun and Jing Li contributed equally to this work.

doi: 10.17221/43/2016-CJGPB

in herbal medicine therapy in China. There are two kinds of loquat fruits: red-fleshed and white-fleshed loquats which have significantly different carotenoid contents in fruits (Zhou *et al*. 2007). Beta-carotene can be converted to retinol and increasing carotenoid content in loquat fruits is important for the improvement of fruit quality. Loquat has a long life cycle and generally requires 8 to 10 years for inflorescence production. Therefore, conventional breeding schemes are time-consuming to improve important traits such as yield and fruit quality in loquat and molecular marker assisted breeding is more preferred.

Single nucleotide polymorphisms (SNPs) and simple sequence repeat (SSR) markers have been developed based on transcriptomic data of loquat (Li *et al*. 2015), and a few SSR markers have been developed to identify red-fleshed and white-fleshed loquat (Li *et al*. 2014). Beta-carotene and lutein are the major carotenoids in the loquat flesh (Zhou *et al*. 2007); the functional analysis of multiple phytoene synthase genes indicated that this gene family is involved in the accumulation of loquat carotenoids (Fu *et al*. 2014). However, no thorough transcriptome information for red-fleshed and white-fleshed fruits has been provided for these two loquat accessions, which may reveal more clues for genes involved in carotenoid accumulation. The marker discovery and development in functional genes linked to important agronomic traits should accelerate the loquat breeding scheme.

With the development of the next generation sequencing technology, RNA-seq has been widely used to do a comparative analysis of multiple transcriptomes in many different organisms (Wang *et al*. 2009). The usage of RNA-seq technology has greatly enhanced the understanding of transcriptome complexity in animal and plant systems, such as revealing novel transcripts, alternative splicing, transcript isoforms, new large intergenic noncoding RNAs and SNPs (Gulledge *et al*. 2012; Lei *et al*. 2014; Xia *et al*. 2014; Gupta *et al*. 2015; Xu *et al*. 2015). The next generation sequencing technology had been widely applied to the genome and transcriptome sequencing in the family Rosaceae. Presently, the whole-genome sequences of ten plant species in the family Rosaceae have been released, including the domesticated apple – *Malus* × *domestica* (Velasco *et al*. 2010), *Fragaria vesca* (Shulaev *et al*. 2011), *Pyrus bretschneideri* (Wu *et al*. 2013), and *Prunus persica*(Verde *et al*. 2013). Meanwhile, a large number of transcriptome datasets (1752) for the family Rosaceae is available in the National Centre for Biotechnology Information (NCBI) database. The studies conducted in the family Rosaceae using transcriptome sequencing have provided an overall insight into the gene expression profiles and candidate genes associated with targeted phenotype (Hyun *et al*. 2014; Shin *et al*. 2016). The transcriptome analysis for specific biological processes can provide basic information for the further molecular analysis.

In this study, we applied RNA-seq to obtain transcriptome data from the mixture samples for the four developmental stages of both red-fleshed and white-fleshed loquat fruits, and characterized gene expression profiles in the two accessions with detailed comparison of the transcripts related with carotenoid biosynthesis pathway. This study was applied to discover candidate genes involved in carotenoid biosynthesis and then detect potential SNP markers in these candidate genes in red- and white-flesh loquat cultivars. Our work will be useful for the elucidation of genetic differences between red- and white-flesh loquat fruits and further functional analysis of the genes responsible for carotenoid accumulation.

MATERIAL AND METHODS

Plant material. Two loquat accessions were used in the study: a wild type loquat with red-fleshed fruit and a natural mutant loquat with white-fleshed fruit. The loquat accessions used in the present experiment were grown in the Abei Wenchuan Mountain of Sichuan province in southwestern China. The fruit samples were collected from two 15-years-old loquat trees with close flowering time. The fruit samples of the two loquat accessions were separately collected from four developmental stages: 171, 175, 180, and 185 days after pollination. The fruit samples were immediately frozen in liquid nitrogen. Total RNA was extracted using TRIzol Reagent (Invitrogen, Waltham, USA) (0.1 g tissue per 1000 ml TRIzol reagent), and the RNA samples from the four developmental stages were equally mixed based on the RNA quantity for both loquat accessions.

Synthesis of cDNA and subsequent sequencing. The purified mRNA was fragmented with bivalent cations under a temperature gradient. These short fragments were used as templates to synthesize first-strand cDNA using random hexamer primers and Superscript™ III (Invitrogen™, Carlsbad, USA). Second-strand cDNA was then synthesized in a solution containing buffer, dNTPs, RNaseH and DNA

doi: 10.17221/43/2016-CJGPB

polymerase I. Second-strand cDNA was subsequently purified using a QiaQiuck PCR extraction kit (Qiagen, Hilden, Germany). EB (Elution Buffer) buffer was used to resolve these short fragments for end repair and poly(A) tail addition. The sequence adaptors were linked to two ends of short cDNA sequences, and suitably sized cDNA fragments were selected for polymerase chain reaction (PCR) amplification based on the agarose gel electrophoresis results. Finally, the library established was sequenced with an Illumina Hiseq^{TM} 2000 system. The paired-end library was developed according to the protocol for the Paired-End Sample Preparation kit (Illumina, San Diego, USA).

***De novo* assembly and annotation**. *De novo* assembly of the loquat transcriptome from the fruit tissues was performed using Trinity, a short-read assembly program, as described in Haas *et al*. (2013). The annotation for protein-coding unigene was done as follows: unigenes were compared using BLAST against the Nr database (NCBI non-redundant sequence database) using an E-value cut-off of 10^{-5} and then against other protein databases, such as Swiss-Prot, Clusters of Orthologous Groups (COG) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG), following the method used by Fan *et al*. (2013).

Evaluation of loquat assembled expressed sequence transcriptome unigenes by comparison with related species. A total of 383 expressed sequences for loquat were downloaded from the NCBI database. Gene sequences of the apple (*Malus domestica*) genome were downloaded from Phytozome version 8.0 (http://www.phytozome.net). Subsequently, these downloaded sequences were used for alignment with the assembled unigenes from *Eriobotrya japonica* transcriptome using an E-value cut-off of 10^{-5} (E value < 0.000 01).

Real-time qPCR assays. The primers used for RT-qPCR are listed in Table 1. Real-time PCR was performed following a standard SYBR Premix Ex TaqTM kit (TaKaRa, Otsu, Japan) protocol in 96-well optical plates (Axygen, Redwood, USA) using a final volume of 10 µl. The reactions were incubated in 0.2 ml tubes of a Mastercycler ep realplex4 (Eppendorf, Hamburg, Germany) machine as follows: 95°C for 5 s, 55°C for 15 s and 68°C for 20 s. The procedure ended by a melt-curve ramping from 60 to 95°C for 20 min to check the PCR specificity. All qPCR reactions were carried out in biological and technical triplicate. The final Ct values were the means of nine values. The comparative expression levels of candidate genes involving carotenoid biosynthesis were normalized to those of *ACTIN*.

Identification of single nucleotide polymorphisms between red- and white-fleshed accessions. The Genome Analysis Toolkit Unified Genotyper (GATK) was used to detect SNPs in the red- and white-fleshed loquat cultivars. Default parameters were used for the GATK software which assumed a heterozygosis ratio of one every 1000 bp. Variants detected with a quality less than 20 were deleted using the VCF filter (McKenna *et al*. 2010). The SNPs found in candidate genes for carotenoid biosynthesis were further validated by counting the percentage of the nucleotides in the SNP locus; if the coverage reads more than 30 and the variant nucleotide accounts for 80% of the reads, the SNP loci are considered as reliable SNP. The transcripts for these genes were translated into amino sequences to catalogue sense mutation and nonsense mutation.

RESULTS

RNA-seq and assembly quality statistics. The RNA-seq datasets in this study were derived from the two fruit mixtures for red-fleshed loquat and white-fleshed loquat. The mixture of fruit sam-

Table 1. Primer information for four genes for carotenoid biosynthesis and used for qPCR analysis

Gene name		5'-3' sequence	Tm (°C)	Product length (bp)
Unigene0028987	F	AGACAAATGAATATCACAAGACGCA	60.9	168
	R	GAGAGATGAGAAAGTTCCCAAGAGA	60.4	
Unigene0007369	F	CACTCGTGCATTTTCAGTGGAG	59.0	202
	R	TCTTCCCTCTCTTAATTTTTCTCAA	58.6	
Unigene0034788	F	AGTTGGACTCGTTAACTTCCTATCC	59.8	232
	R	CCAATTGTGTTCTTGTCACCCTT	60.7	
Unigene0024512	F	TTCAGACACCATCAATCTCTCTACTT	58.7	166
	R	ATCACAACACTTTCTTCTACTACTTCCA	59.9	

doi: 10.17221/43/2016-CJGPB

Table 2. Summary of RNA-seq and *de novo* assembly in loquat fruits

	No.	Mean size	N50 size	Total nucleotides
		(bp)		
Read	94 977 782	90	90	9 497 778 200
Unigene	61 587	817	1 282	50 317 628

ples included four developmental stages with an equal RNA quantity. In total, 94 million paired-end reads were produced, with an average read length of 90 bp (Table 2, Figure 1). All clean reads from two independent Illumina sequence runs for the two RNA mixtures were deposited in NCBI (Submission Number: SRS978982 and BioProject: PRJNA281556). *De novo* assembly for the two transcriptome datasets has produced 71 032 contigs with the average transcript size of 851 bp and N50 of 1282 bp and 61 586 unigenes with the average length of 817 bp (Table 2).

The assembled transcripts (61 587 unigenes) were aligned with available loquat expressed sequences and closely related species in the family Rosaceae downloaded from the NCBI database in July 2014 to assess the accuracy of these transcripts (Figure S1 in electronic supplementary material (ESM)). A total of 63 517 gene sequences of the apple genome were downloaded from Phytozome version 8.0 (http://www.phytozome.net) and used for the sequence alignments and subsequent estimation of the coverage of the loquat transcriptome. Using a cut-off E-value of 10^{-5}, 1276 unigenes (1.8%) were aligned to the downloaded loquat nucleotide sequences, and 76.5% of the unigenes were aligned to the gene sequences of the apple genome (Figure 2). Moreover, 86.4% of the expressed sequences from loquat and 90.6% of gene sequences in the apple genome were matched with unigenes with a percentage of 44% and 65% of the unigene sequences that had an E-value less than $1e^{-50}$, respectively (Figure 2).

Functional annotation and classification of unigenes in loquat. The unigenes obtained in this study were aligned with the five databases – Non-redundant protein sequences (Nr), Swissprot, GO, COG and KEGG databases using a cut-off E-value of $1e^{-5}$ for the gene function annotation. Of 61 587 unigenes, a total of 44 710 (72.6%) were matched to the five databases with the most unigenes (28 900 unigenes, 46.9%) annotated in the Swissprot database (Table 2). In the Nr database, 21 461 unigenes were matched to 404 species and the largest proportion of loquat unigenes (19 820 unigenes) exhibited similarity to the protein database for *Fragaria vesca* subsp. *vesca*. The third number of unigenes was annotated by the GO database.

Subsequently, GO terms were assigned to loquat unigenes according to the matching result from the GO database. A total of 19 344 unigenes (31.4%) were classified into at least one GO term, of which 13 059 unigenes were in the biological process category, 13 459 unigenes were in the cellular component category and 14 139 unigenes were in the molecular function category (Table 3 and Figure 3). Metabolic process (10 562, 14.9%), cell (13 227, 18.6%) and catalytic activity (8899, 12.5%) were the most abundant GO categories in the biological process, cellular component and molecular function categories, respectively. Moreover, a large number of unigenes were also assigned to the terms associated with biological regulation (2780), cellular component organization (2327), development (2769), establishment of localization (2707), localization (2837), multicellular

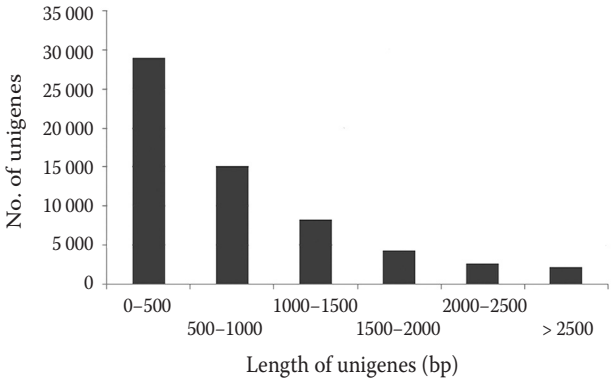

Figure 1. Length distribution of the assembled unigenes in red-fleshed and white fleshed loquat fruits

Table 3. Summary of the unigene annotation in loquat

Matched database	No. of unigenes	Percentage (%)
Nr	21 461	34.8
Swissprot	28 900	46.9
GO – biological process	13 059	21.2
Molecular function	14 139	23.0
Cellular component	13 459	21.9
GO – total	19 344	31.4
COG	15 354	24.9
KEGG	13 041	21.2
Total (cutoff evalue $1e^{-5}$)	44 710	72.6

doi: 10.17221/43/2016-CJGPB

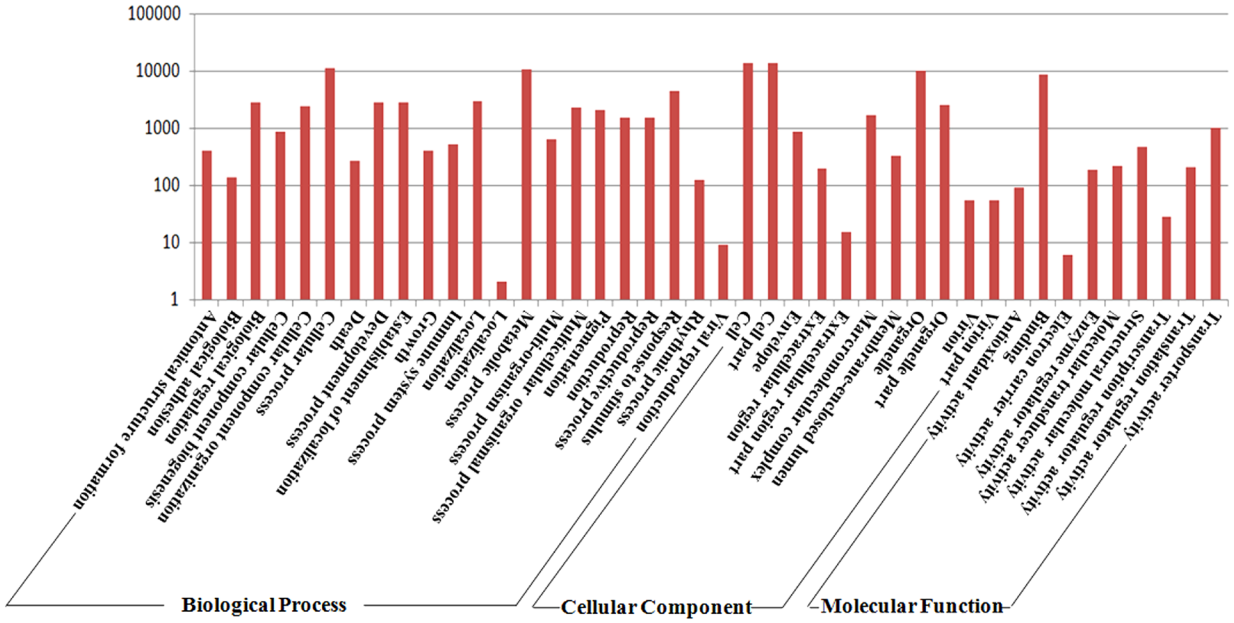

Figure 2. Histogram of GO classifications of the assembled loquat unigenes

organismal process (2184), pigmentation (2009), response to stimulus (4433), cell part (13 227), organelle (9961), organelle part (2518), binding (8322), and catalytic activity (8899).

These unigenes were also used as query sequences to BLAST against the COG protein database to further evaluate the functionality of the loquat transcriptome (Figure 3). For all unigenes, 15 354 (31.9%) were grouped into 25 COG categories and the largest proportion of unigenes (5017) was classified into general function prediction, and then followed by transcription (3114), post translation modification,

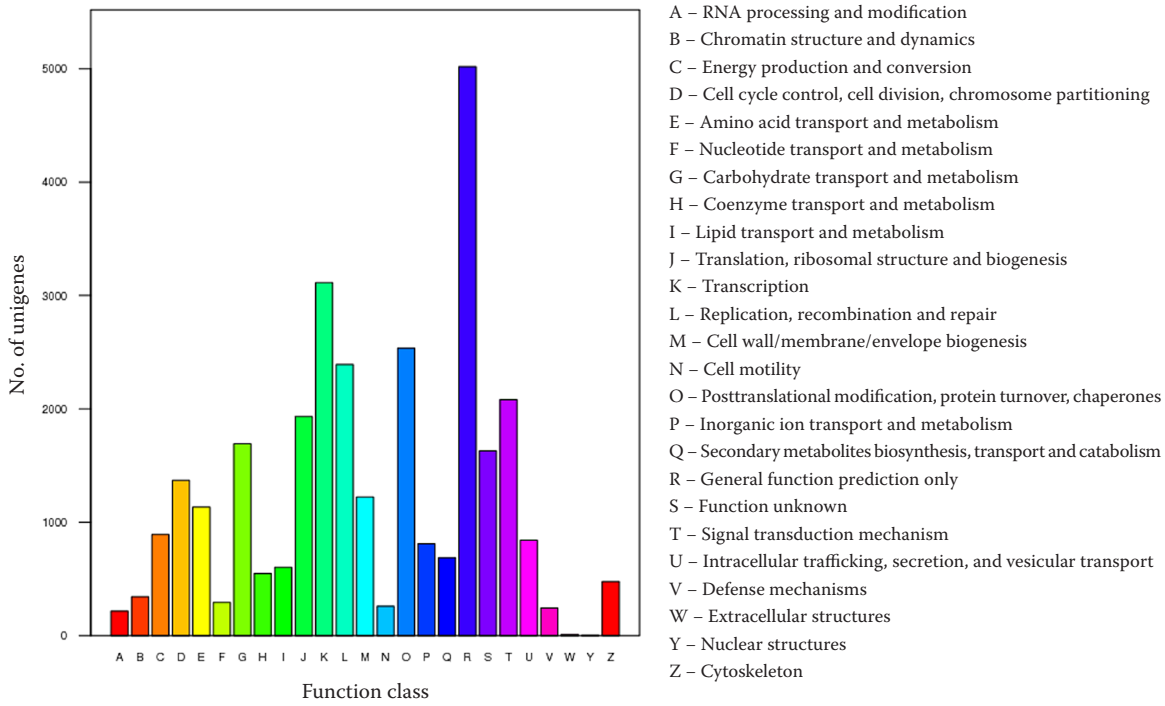

Figure 3. Histogram of COG classification of the expressed sequences in loquat; here, 22 689 assembled unigenes were clustered into 25 clusters

doi: 10.17221/43/2016-CJGPB

protein turnover, chaperones (2537), replication, recombination and repair (2392), signal transduction mechanisms (2081), translation, ribosomal structure and biogenesis (1932), carbohydrate transport and metabolism (1692), function unknown (1630), cell cycle control, cell division, chromosome portioning (1370), cell wall/membrane/envelope biogenesis (1223), and amino acid transport and metabolism (1134) (Figure 3).

The unigenes were further mapped to the reference canonical pathways in KEGG to elucidate the biochemical pathways in the loquat fruit development. A total of 123 biosynthesis pathways were predicted and the metabolic pathway was the largest group (3131 unigenes), followed by biosynthesis of secondary metabolic processes (1457 unigenes), ribosome (637 unigenes), protein processing in endoplasmic reticulum (484 unigenes), RNA transport (432 unigenes) and plant hormone signal transduction (406 unigenes).

Genes potentially involved in carotenoid biosynthesis. We characterized the assembled unigenes with regard to biochemical pathways related to carotenoid biosynthesis because of the significant difference in carotenoid contents between red-fleshed loquat and white-fleshed loquat. Eight unigenes exhibiting high similarity to phytoene synthase (*PSY*), which may participate in converting two molecules of GGPP to the colourless molecule phytoene, were identified from the loquat transcriptomes (Table S1 in ESM). Moreover, six unigenes and four unigenes have high similarity to known zeta-carotene desaturase genes (*ZDS*) and lycopene beta-cyclase genes (*LCYB*), respectively. Unigene 0006261 was annotated as lycopene ε-cyclase (*LCYE*), which synthesizes α-carotene. Subsequently, β-carotene can be hydroxylated to produce lutein and zeaxanthin catalyzed by β-ring hydroxylase (*BCH/CRTRB*) and zeaxanthin can be transformed into violaxanthin by zeaxanthin epoxidase (*ZEP1*). Based on the transcriptomic data, seven unigenes were annotated as *BCH* and fifteen unigenes were identified as *ZEP*.

We predicted the amino acid sequence of these expressed sequences using online software (http://www.fr33.net/translator.php) to confirm whether the identified unigenes involved in carotenoid biosynthesis are full-length or truncated cDNA. Then, these predicted amino sequences were aligned with Nr database (Table S2 in ESM). Based on the blastp result, the first amino acid of nine unigenes was matched to the start amino acid of reference genes and the full length of the amino acid sequences for the expressed sequences varied from 336 to 604, with an average length of 466. Almost all unigene amino sequences showed high similarity and full length with the amino acid sequences of the functional reference genes involved in carotenoid biosynthesis except for unigene 0034723 (ζ-carotene desaturase, *ZDS*) and unigene 0006068 (*ZDS*).

Gene expression analysis for genes related with carotenoid biosynthesis. Real-time quantitative PCR was applied to detect gene expression changes for four genes (unigene 0028987, unigene 0007369, unigene 0034788, and unigene 0024512) during different developmental stages of the red-fleshed and white-fleshed fruit of loquat (Figure 4). The end of the linear carotenoid lycopene can be cyclized by lycopene beta-cyclase. Unigene 0028987 (*LCYB*) had a higher expression level in red-fleshed loquat compared to that in white-fleshed loquat based on the RT-qPCR result. Alpha-carotene and beta-carotene are hydroxylated to produce lutein and zeaxanthin by the catalyzing of beta-carotene hydroxylase (*CRTRB*, unigene 0007369) and epsilon-ring hydroxylase (*CRTRB*). At the first developmental stage (171 days after pollination (DAP)), the expression level of unigene 0007369 in white-fleshed loquat was significantly higher ($P = 0.003$; $P < 0.05$) than that of red-fleshed loquat (Figure 4). While at the other three stages (175, 180 and 185 DAP), the expression levels were close between red-fleshed and white-fleshed loquat. Meanwhile, at the 171, 175, and 180 DAP stages, a low expression level of unigene 0034788 was detected in both red-fleshed and white-fleshed loquat. However, at the 185 DAP stage, unigene 0034788 (*ZEP1*) had a higher expression level in red-fleshed loquat. Zeaxanthin can be transformed into violaxanthin by zeaxanthin epoxidase and no significant expression level changes of unigene 0024512 (*ZEP1*) were detected in red-fleshed and white-fleshed loquat at the four developmental stages.

SNPs for candidate unigenes in red-fleshed and white-fleshed loquat accessions. A total of 2,396 SNPs were identified from 1,361 unigenes with an average of 1.76 SNPs per unigene based on transcriptome data from red-fleshed and white-fleshed loquat (Table S3 in ESM). The number of SNPs in a single unigene varied from 1 to 13 and 13 SNPs detected in unigene 0033613. Twelve SNP types were identified between the red-fleshed and white-fleshed loquat listed as A/G, G/A, C/T, T/C, T/G, G/T, G/C, C/G, A/T, T/A, A/C, and C/A. Among all SNPs (2396) detected in the unigenes,

doi: 10.17221/43/2016-CJGPB

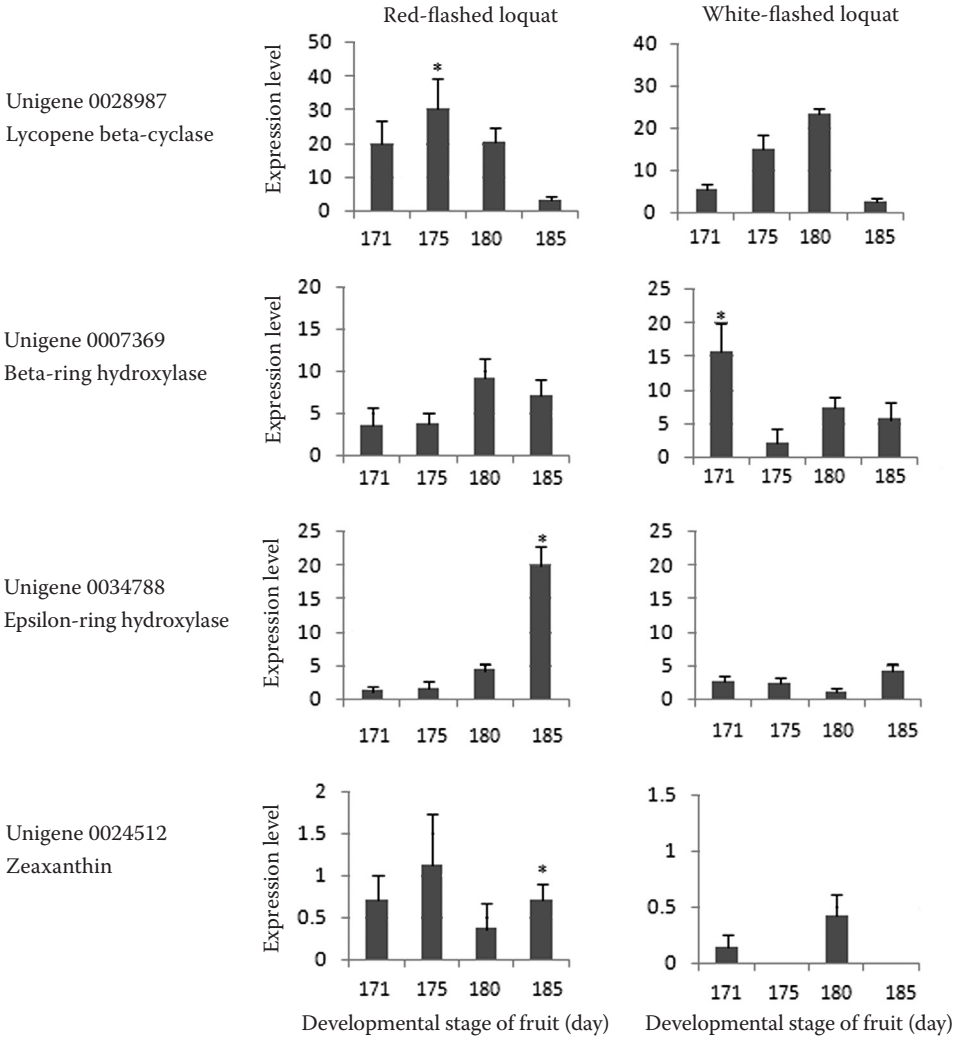

Figure 4. Expression changes of four identified full-length genes involved in carotenoid biosynthesis during different developmental stages of loquat fruits between red- and white-fleshed loquat; *represent a significant expression change in the same developmental stage

the majority of SNPs (87%) are nonsense mutations, whereas the remaining 311 SNPs are sense mutations.

Eleven SNPs were detected in three candidate genes involved in carotenoid biosynthesis. Ten SNPs are nonsense mutations and only one SNP (*ZEP1*, unigene 0024512) is a sense mutation that changed the translated amino acid (Table S3 in ESM).

DISCUSSION

Loquat is an evergreen fruit tree and the red-fleshed cultivars contain high vitamin A. There are two specific kinds of loquat cultivars (red-fleshed and white-fleshed loquat) which have natural variations in carotenoid contents of the fruit. In this research, two transcriptomes for fruit tissues of the two cultivars have been obtained and a comprehensive analysis has been conducted to reveal the expression profiles and the sequence polymorphisms for the candidate genes related with carotenoid biosynthesis. A total of 61 586 unigenes were assembled from the transcriptomes and 44 710 unigenes were annotated by blast against Nr, Swissprot, GO, COG and KEGG databases (Table 2). Among these annotated unigenes, six key gene families involved in the pathway of carotenoid biosynthesis were identified as a number of homologous genes in loquat: eight *PSY*, six *ZDS*, four *LCYB*, one *LCYE* gene, seven *BCH/CRTRB* and fifteen *ZEP* unigenes. The RT-qPCR for four candidate genes indicated different expression profiles

doi: 10.17221/43/2016-CJGPB

in red-fleshed and white fleshed loquat cultivars. These results suggest that the analysis based on transcriptomes can provide key genes related with the carotenoid biosynthesis in loquat.

The pathway of carotenoid biosynthesis has been extensively studied (Lu & Li 2008), all genes involved in this pathway have been isolated and analysed in many species (Bertrand 2010; Gomez-Garcia & Ochoa-Alejo 2013; Rodriguez-Concepción & Stange 2013). The researches to reveal the mechanism for the variation of carotenoid content have been done in maize and demonstrated that the genetic variation in the β-carotene hydroxylase gene – *CRTRB1* caused reduced transcript expression correlating with higher carotene concentrations (Yan *et al*. 2010). In previous research, the alternative splicing of *PSY* gene affected the carotenoid contents (Fu *et al*. 2014). In this study, the analysis of gene coding sequences showed that only two transcripts of *ZDS* genes were truncated. The expression analysis for key genes in the carotenoid biosynthesis pathway showed that unigene 0028987 (*LCYB*) and unigene 0034788 (*ZEP1*) had a higher expression level in red-fleshed loquat compared to that in white-fleshed loquat, while the expression level of unigene 0007369 (*CRTRB*) ($P = 0.003$, $P < 0.05$) was higher in white-fleshed loquat than that of red-fleshed loquat in the 171 DAP stage (Figure 4). It suggests that in red-fleshed loquat cultivars containing more carotene it may result from both more carotene produced by the higher expression in *LCYB* genes and less carotene converted because of the low the expression of *CRTRB* gene.

From the transcriptome comparison between red-fleshed and white-fleshed loquats, a total of 2396 SNPs (1361 unigenes) were identified from the fruit transcriptomes of red-fleshed and white-fleshed loquat cultivars and the majority of the SNPs (87%) caused nonsense changes in protein coding (Table S2 in ESM). For the genes in the carotenoid biosynthesis, 11 SNPs were detected and a sense mutation was found in *ZEP1*. Since the *ZEP1* gene is related with carotene further converted to other components and this SNP may related with the carotene accumulation in loquat. In further research, the catalysis efficiency for this gene should be analysed.

In this research, we characterized the variations of the expression level and transcript sequence for the genes related with carotenoid biosynthesis. The expression levels for genes in the pathway of carotenoid biosynthesis in red-fleshed loquat are higher than in white-fleshed loquat. The majority of SNPs identified between the two loquat cultivars were nonsense mutations and one out of eleven SNPs in candidate genes was a sense mutation. All these results will be useful for the elucidation of genetic differences between red- and white-fleshed loquat fruits and further functional analysis of the genes responsible for carotenoid accumulation.

Acknowledgements. This work was supported by Key Laboratory Program of Agriculture Ministry of China (2013JCYJ-004) and two subsidy programs (2015LWJJ-010): Program of Modern Agriculture Technology Innovation and Demonstration of Provincial Finance Department (2014CXSF-015).

References

Bertrand M. (2010): Carotenoid biosynthesis in diatoms. Photosynthesis Research, 106: 89–102.

Fan H., Xiao Y., Yang Y., Xia W., Mason A.S., Xia Z., Qiao F., Zhao S., Tang H. (2013): RNA-Seq analysis of *Cocos nucifera*: transcriptome sequencing and de novo assembly for subsequent functional genomics approaches. PLoS ONE, 8: e59997.

Fu X., Feng C., Wang C., Yin X., Lu P., Grierson D., Xu C., Chen K. (2014): Involvement of multiple phytoene synthase genes in tissue- and cultivar-specific accumulation of carotenoids in loquat. Journal of Experimental Botany, 65: 4679–4689.

Gomez-Garcia M., Ochoa-Alejo N. (2013): Biochemistry and molecular biology of carotenoid biosynthesis in chili peppers (*Capsicum* spp.). International Journal of Molecular Sciences, 14: 19025–19053.

Gulledge A.A., Roberts A.D., Vora H., Patel K., Loraine A.E. (2012): Mining *Arabidopsis thaliana* RNA-seq data with Integrated Genome Browser reveals stress-induced alternative splicing of the putative splicing regulator SR45a. American Journal of Botany, 99: 219–231.

Gupta V., Estrada A.D., Blakley I., Reid R., Patel K., Meyer M.D., Andersen S.U., Brown A.F., Lila M.A., Loraine A.E. (2015): RNA-Seq analysis and annotation of a draft blueberry genome assembly identifies candidate genes involved in fruit ripening, biosynthesis of bioactive compounds, and stage-specific alternative splicing. Gigascience, 4: 5.

Haas B.J., Papanicolaou A., Yassour M., Grabherr M., Blood P.D., Bowden J., Couger M.B., Eccles D., Li B., Lieber M., Macmanes M.D., Ott M., Orvis J., Pochet N., Strozzi F., Weeks N., Westerman R., William T., Dewey C.N., Henschel R., Leduc R.D., Friedman N., Regev A. (2013): *De novo* transcript sequence reconstruction from RNA-seq

doi: 10.17221/43/2016-CJGPB

using the Trinity platform for reference generation and analysis. Nature Protocols, 8: 1494–1512.

Hyun T.K., Lee S., Rim Y., Kumar R., Han X., Lee S.Y., Lee C.H., Kim J.Y. (2014): *De-novo* RNA sequencing and metabolite profiling to identify genes involved in anthocyanin biosynthesis in Korean black raspberry (*Rubus coreanus* Miquel). PLoS ONE, 9: e88292.

Lei X., Xiao Y., Xia W., Mason A.S., Yang Y., Ma Z., Peng M. (2014): RNA-seq analysis of oil palm under cold stress reveals a different C-repeat binding factor (CBF) mediated gene expression pattern in *Elaeis guineensis* compared to other species. PLoS ONE, 9: e114482.

Li X., Xu H., Feng J., Zhou X., Chen J. (2015): Mining of genic SNPs and diversity evaluation of landraces in loquat. Scientia Horticulturae, 195: 82–88.

Li X.Y., Xu H.X., Chen J.W. (2014): Rapid identification of red-flesh loquat cultivars using EST-SSR markers based on manual cultivar identification diagram strategy. Genetics and Molecular Research, 13: 3384–3394.

Lu S., Li L. (2008): Carotenoid metabolism: biosynthesis, regulation, and beyond. Journal of Integrative Plant Biology, 50: 778–785.

McKenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., DePristo M.A. (2010): The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 20: 1297–1303.

Rodriguez-Concepción M., Stange C. (2013): Biosynthesis of carotenoids in carrot: an underground story comes to light. Archives of Biochemistry and Biophysics, 539: 110–116.

Shin S., Zheng P., Fazio G., Mazzola M., Main D., Zhu Y. (2016): Transcriptome changes specifically associated with apple (*Malus domestica*) root defense response during *Pythium ultimum* infection. Physiological and Molecular Plant Pathology, 94: 16–26.

Shulaev V., Sargent D.J., Crowhurst R.N., Mockler T.C., Folkerts O., Delcher A.L., Jaiswal P., Mockaitis K., Liston A., Mane S.P., Burns P., Davis T.M., Slovin J.P., Bassil N., Hellens R.P., Evans C., Harkins T., Kodira C., Desany B., Crasta O.R., Jensen R.V., Allan A.C., Michael T.P., Setubal J.C., Celton J.M., Rees D.J., Williams K.P., Holt S.H., Ruiz Rojas J.J., Chatterjee M., Liu B., Silva H., Meisel L., Adato A., Filichkin S.A., Troggio M., Viola R., Ashman T.L., Wang H., Dharmawardhana P., Elser J., Raja R., Priest H.D., Bryant D.W., Fox S.E., Givan S.A., Wilhelm L.J., Naithani S., Christoffels A., Salama D.Y., Carter J., Lopez G.E., Zdepski A., Wang W., Kerstetter R.A., Schwab W., Korban S.S., Davik J., Monfort A., Denoyes-Rothan B., Arus P., Mittler R., Flinn B., Aharoni A., Bennetzen J.L., Salzberg S.L., Dickerman A.W., Velasco R., Borodovsky M., Veilleux R.E., Folta K.M. (2011): The genome of woodland strawberry (*Fragaria vesca*). Nature Genetics, 43: 109–116.

Velasco R., Zharkikh A., Affourtit J., Dhingra A., Cestaro A., Kalyanaraman A., Fontana P., Bhatnagar S.K., Troggio M., Pruss D., Salvi S., Pindo M., Baldi P., Castelletti S., Cavaiuolo M., Coppola G., Costa F., Cova V., Dal Ri A., Goremykin V., Komjanc M., Longhi S., Magnago P., Malacarne G., Malnoy M., Micheletti D., Moretto M., Perazzolli M., Si-Ammour A., Vezzulli S., Zini E., Eldredge G., Fitzgerald L.M., Gutin N., Lanchbury J., Macalma T., Mitchell J.T., Reid J., Wardell B., Kodira C., Chen Z., Desany B., Niazi F., Palmer M., Koepke T., Jiwan D., Schaeffer S., Krishnan V., Wu C., Chu V.T., King S.T., Vick J., Tao Q., Mraz A., Stormo A., Stormo K., Bogden R., Ederle D., Stella A., Vecchietti A., Kater M.M., Masiero S., Lasserre P., Lespinasse Y., Allan A.C., Bus V., Chagne D., Crowhurst R.N., Gleave A.P., Lavezzo E., Fawcett J.A., Proost S., Rouze P., Sterck L., Toppo S., Lazzari B., Hellens R.P., Durel C.E., Gutin A., Bumgarner R.E., Gardiner S.E., Skolnick M., Egholm M., Van de Peer Y., Salamini F., Viola R. (2010): The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nature Genetics, 42: 833–839.

Verde I., Abbott A.G., Scalabrin S., Jung S., Shu S., Marroni F., Zhebentyayeva T., Dettori M.T., Grimwood J., Cattonaro F., Zuccolo A., Rossini L., Jenkins J., Vendramin E., Meisel L.A., Decroocq V., Sosinski B., Prochnik S., Mitros T., Policriti A., Cipriani G., Dondini L., Ficklin S., Goodstein D.M., Xuan P., Del Fabbro C., Aramini V., Copetti D., Gonzalez S., Horner D.S., Falchi R., Lucas S., Mica E., Maldonado J., Lazzari B., Bielenberg D., Pirona R., Miculan M., Barakat A., Testolin R., Stella A., Tartarini S., Tonutti P., Arus P., Orellana A., Wells C., Main D., Vizzotto G., Silva H., Salamini F., Schmutz J., Morgante M., Rokhsar D.S. (2013): The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nature Genetics 45: 487–494.

Wang Z., Gerstein M., Snyder M. (2009): RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics, 10: 57–63.

Wu J., Wang Z., Shi Z., Zhang S., Ming R., Zhu S., Khan M.A., Tao S., Korban S.S., Wang H., Chen N.J., Nishio T., Xu X., Cong L., Qi K., Huang X., Wang Y., Zhao X., Deng C., Gou C., Zhou W., Yin H., Qin G., Sha Y., Tao Y., Chen H., Yang Y., Song Y., Zhan D., Wang J., Li L., Dai M., Gu C., Shi D., Wang X., Zhang H., Zeng L., Zheng D., Wang C., Chen M., Wang G., Xie L., Sovero V., Sha S., Huang W., Zhang M., Sun J., Xu L., Li Y., Liu X., Li Q., Shen J., Paull R.E., Bennetzen J.L. (2013): The genome of

doi: 10.17221/43/2016-CJGPB

the pear (*Pyrus bretschneideri* Rehd.). Genome Research, 23: 396–408.

Xia W., Mason A.S., Xiao Y., Liu Z., Yang Y., Lei X., Wu X., Ma Z., Peng M. (2014): Analysis of multiple transcriptomes of the African oil palm (*Elaeis guineensis*) to identify reference genes for RT-qPCR. Journal of Biotechnology, 184: 63–73.

Xu H.M., Kong X.D., Chen F., Huang J.X., Lou X.Y., Zhao J.Y. (2015): Transcriptome analysis of *Brassica napus* pod using RNA-Seq and identification of lipid-related candidate genes. BMC Genomics, 16: 858.

Yan J., Kandianis C.B., Harjes C.E., Bai L., Kim E.-H., Yang X., Skinner D.J., Fu Z., Mitchell S., Li Q., Fernandez M.G.S., Zaharieva M., Babu R., Fu Y., Palacios N., Li J., DellaPenna D., Brutnell T., Buckler E.S., Warburton M.L., Rocheford T. (2010): Rare genetic variation at *Zea mays crtRB1* increases β-carotene in maize grain. Nature Genetics, 42: 322–327.

Zhou C., Xu C., Sun C., Li X., Chen K. (2007): Carotenoids in white- and red-fleshed loquat fruits. Journal of Agricultural and Food Chemistry, 55: 7822–7830.

Received for publication March 27, 2016
Accepted after corrections March 9, 2017
Published online June 14, 2017