Pak. J. Bot., 52(1): 129-137, 2020. DOI: http://dx.doi.org/10.30848/PJB2020-1(19)

FUNCTIONAL ANALYSIS OF *LYCOPENE β-CYCLASE GENES* FROM TWO WATERMELON CULTIVARS LSW-177 (RED-FLESH) AND COS (PALE-YELLOW FLESH)

CHUAN WU[1, 2], PENG GAO[1, 2], SHI LIU[1, 2], HUILING LV[1, 2] AND FEISHI LUAN[1, 2*]

[1]*Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (Northeast Region), Ministry of Agriculture, Northeast Agricultural University, 600 Changjiang Road, Harbin 150030, China*
[2]*Horticultural College, Northeast Agricultural University, 600 Changjiang Road, Harbin 150030, China*
**Correspondence author's email: luanfeishi@neau.edu.cn*

Abstract

Carotenoid pigments are extensively found in plant tissues with multifunction. The specific carotenoids have been classified in plant photosynthetic centers and most of them are exclusively abundant in colorful fruits and vegetables. Previous results show that lycopene cyclases are fundamental for determining the proportion of carotenoids in fruit, leaves, flowers and roots. A quantitative trait locus (QTL) analysis exhibited that an enzyme is responsible for red versus non-red flesh in watermelon (*Citrullus lanatus*) and also correlated with the function of lycopene β-cyclase (lcyb) with flesh colour. The current study was conducted to characterize the two polymorphic genes encoding lcyb enzymes, from two watermelon cultivars (LSW-177and COS) containin the red and pale-yellow flesh, respectively. Results had confirmed the cyclase function of the two lcyb proteins in an *Escherichia coli* by heterologous complementation. These twolcyb proteins were differential in the amino acids at the 226th and 435th positions and had the effective ability to cyclase lycopene into bicyclic β-carotene, showing that these amino acid variations did not influence the functionality of lcyb enzymes in *E. coli*. In addition, the conversion of Ala168 to Glu168 by site-directed mutagenesis which inactivated the enzyme, suggested that this residue was part of the core region for lcyb function in watermelon. The results presented here will help more to clarify the mechanism by which lycopene β-cyclase performs its function.

Key words: Carotenoids; HPLC; Lycopene β-cyclase; Site-directed mutagenesis; Watermelon.

Introduction

Carotenoids are the foremosttetraterpenoid pigments and there are more than 750 naturally occurring pigment molecules which are already catalogued. These pigments confer various colours (red, orange, pink and yellow) to the fruit flesh of tomato, pepper, melon, watermelon and other plants (Britton *et al*., 2004). In addition to fleshy fruits, carotenoids accumulate in endosperm tissue, kernels as well as roots, leaves and flowers (Bartley & Scolnik, 1995). Acting as widely existed pigments, carotenoids can protect the photosynthetic apparatus from photo damages (Frank & Cogdell, 1996; Robert *et al*., 2004). Up to date, more and more studies on the effect of carotenoids (especially for lycopene) in human health were focused, which exhibitedcertain benefits against chronic diseases and also decreased the probability of cancer and cardiovascular diseases (Fraser & Bramley, 2004; Nisar *et al*., 2015). Due to the strong antioxidant property, lycopene can protect the body against oxidative damage (Agarwal & Rao, 2000). The carotenoids containing two β-rings serve as the precursor and source of vitamin A due to its importance as fundamental nutrient for those humans who cannot directly intake it from green plants (Fraser & Bramley, 2004; Nisar *et al*., 2015). Watermelon (*Citrullus lanatus*) is an important horticultural and economic crop, which hold a large amount of lycopene and can meet people's daily supplement needs. Actually, watermelon fruit is an alternative of tomato juice with regard to lycopene contents and bioavailability. Therefore, the topics of carotenoid biosynthesis, functions and their metabolism in watermelon are receiving increasing attention.

The regulation of the biosynthesis, metabolism and stable storage of carotenoids controlled their composition and content in different plant organs and tissues (Cazzonelli & Pogson, 2010; Lu & Li, 2008). The carotenoid biosynthetic pathway exists in all photosynthetic species, including green algal, cyanobacteria and some other bacteria (Paniagua-Michel *et al*., 2012), and nearly all the proteins engaged in the pathway have been characterized. Additionally, the complex network had been discussed in tomato and citrus, which reported that the dramatic colour changes occurred during fruit development and ripening stage (Kato *et al*., 2004; Liu *et al*., 2015). Two molecules of geranylgeranyl diphosphate were condensed into a molecule of phytoene, and the step was catalyzed by phytoene synthase (PSY), which performed as the key enzyme for controlling the carotenoid biosynthesis flux (Fraser *et al*., 2002). Later, lycopene is produced from phytoene through a couples of dehydrogenation and isomerization steps catalyzed by phytoenedesaturase (PDS), ζ-carotene desaturase (ZDS) and carotenoid isomerase (CRTISO). Though fruit flesh turned into red due to lycopene accumulation, some reports indicated that the enzymatic ability of lycopene cyclases were the most critical factor determining fleshy colour in tomato, papaya and Star Ruby grapefruit (Pecker *et al*., 1996; Ronen *et al*., 1999). Using the lycopene as substrate, β-carotene and α-carotene produced by lycopene cyclases (lcyb and lcye) is the key point that enriches the profiles of carotenoids during fruit maturation (Harjes *et al*., 2008). A metabolic pathway adds one β- and one ε-ring at each end of

lycopene to produce α-carotene and then lutein; the final pigment serves as a predominant carotenoid in the photosynthesis system. On the other hand, two β- rings are added at each end of lycopene to form β-carotene and its derivatives, such as zeaxanthin, antheraxanthin, violaxanthin, neoxanthinand finally abscisic acid (ABA). Overall, the relative cyclase activities regulated by lcye and lcyb may modulate the conversion rate from lycopene to α- or β-carotene and thus controlling the pathway flux (Diretto *et al*., 2007).

Due to their high homology in amino acid sequences and similar functions, lcyb and lcyepossibly originated from the same ancestor *via* gene replication. Most of the higher plants (tomato, papaya, citrus, *et al*.,) contain two types of lcyb protein (Alquézar *et al*., 2009; Ampomah-Dwamena *et al*., 2009; Blas *et al*., 2010; Ronen *et al*., 2000), which can be classified as chloroplast and chromoplast isoforms, whereas *Arabidopsis* possesses single *lcyb* gene (Lange & Ghassemian, 2003). Conversely, it has been confirmed that only one copy of the *lcye* gene exists in all higher plants. During tomato maturation, the expression of *lcyb* gene was decreased along with lycopene accumulation, which supported the evidence that the transcriptional regulation play a predominant role in carotenogenesis(Pecker *et al*., 1996). At the developmental stage of muskmelon fruit, the variation trend of β-carotene content was consistent with the*lcyb*gene expression, indicated that β-carotene accumulation was closely related with the transcript level of *Cmlcyb1* (Hao *et al*., 2012). The phenotype and carotenoid profiles of tomato mutants (such as *Delta*, *Beta*, and *old-gold*) had also emphasized the pivotal role of lycopene cyclases in the carotenogenesis. Moreover, a frameshift mutation was found in *CpCYCb*and caused a premature protein, which was responsible for the red flesh in papaya (*Carica papaya*) (Blas *et al*., 2010). In red papaya fruit, the large accumulation of lycopene was caused by an inactivated lycopene β-cyclase 2, which contained a mutation in the gene (Devitt *et al*., 2010).

The various concentrations and compositions of carotenoids in the flesh of watermelon fruits are the result of genetic fluctuations and highly regulated processes. For example, Bang *et al*., (2007, 2010) reported that there were two loci (*C and i-C*) responsible for canary yellow and red watermelon fruit colour, and regarded *lcyb* as the colour-determinant gene for canary yellow (C). Additionally, Liu *et al*., (2016) developed a new cleaved amplified polymorphic sequence (CAPS) marker, *WII04EBsaHI-6*, for distinguishing red and non-red fleshy fruit. The marker was properly located in the *lycopene β-cyclase (lcyb)* gene (*Cla005011)* coding region. In order to make sure that whether the differential enzymatic ability of lcyb existed in the fruit development of red and pale-yellow watermelon, the *lcyb*genes were cloned and compared in this study, and their functions were also investigated.

Materials and Methods

Plant materials: All the watermelon cultivars were planted in Xiangfang Farm, Northeastern Agriculture University (44.04 N, 125.42 E), haerbin, China (Table S1). The seeds of LSW-177 (red-flesh) and COS (pale-yellow flesh) were kindly provided by Angela R. Davis. Fruit samples of the two representative watermelon cultivars were harvested at full ripenin stageand stored at -80°C until RNA preparation. All cultivars in this study were *C. lanatus*, and only the cultivar names were provided as mentioned in the supplementary table S1.

Cloning and bioinformatics analysis of the *lycopene β-cyclase (lcyb)* gene: The coding regions of the *lcyb* gene from watermelon were amplified using gene-specific primers (5′- ATGGATACTTTACTTAAAATCAATA -3′ and 5′- AATCTCTATCCTTTACCAG -3′), which were designed according to the sequence from *C. lanatus* (GenBank accession number EF183522) submitted by Bang *et al*., (2007). The fragments were inserted into the pMD18-T Easy vector (TaKaRa, Japan) and the ppositive clones were selected and sequenced, individually. The putative molecular mass and isoelectric points were calculated from the deduced amino acid sequence using DNAMAN 9.0 software (LynnonBiosoft, Canada). The signal peptide and glycosylation sites were predicted by DTU Bioinformatics (http://www.cbs.dtu.dk/). Protein secondary structures and models were computed by SABLE (http://sable.cchmc.org) (Porollo *et al*., 2004) and SWISS-MODEL. The l*cyb* genes from watermelon cultivars with other flesh colours in our lab were cloned and sequenced according to the same procedure.

Phylogenetic analysis: The lcyb proteins were retrieved from the NCBI protein database for assessing the phylogenetic relationships. In order to analyse amino acid diversity and relationships with flesh colour of wild, semi-wild and cultivated watermelons, the *lcyb* genes were obtained from NCBI sequence read archive (SRA) (Guo *et al*., 2013); partial genes were cloned in this experiment. The phylogenetic analysis was performed using MEGA 4.0 software with Neighbor-Joining method (Tamura *et al*., 2007).

Construction of a watermelon lcyb vector and site-directed mutagenesis: A heterologous complementation assay was carried out to evaluate whether the lcyb proteins from the two watermelon cultivars with different fruit flesh colours possess the same cyclase activity. For bacterial expression, the *Cllcyb*coding sequences were tagged with *Bam* H I and *Xho*I sites at the 5′and 3′end, and then ligated into the pET-28a (+) (Novagen, Germany). The recombinant plasmid (named pET-Cllcyb) and the negative control pET-28a (+) (empty vector) were transformed into *Escherichia coli* BL21 (*DE3*) cells containing expression plasmid pACCRT-EIB for the lycopene biosynthesis, respectively. Moreover, pET-Cllcyb-177 was used as the template for site-directed mutagenesis by Sangon biotech, and two mutants (named Mut1 and Mut2) were obtained and confirmed by sequencing. Transformation of the plasmids carrying the mutant sequences were performed according to a described protocol.

Table S1. All the watermelon lycopene β-cyclases included in this study are listed in a table.

Watermelon Name	Run_s	Chromosome position	Flesh colour	Group
JXF	SRR494424	chr4:8886138-8887652	red	cultivated watermelon
JX-2	SRR494422	chr4:8886138-8887652	red	cultivated watermelon
XHBFGM	SRR494427	chr4:8886138-8887652	red	cultivated watermelon
Black_Diamond	SRR494428	chr4:8886138-8887652	red	cultivated watermelon
Calhoun_Gray	SRR494429	chr4:8886138-8887652	red	cultivated watermelon
Sugarlee	SRR494430	chr4:8886138-8887652	red	cultivated watermelon
Sy-904304	SRR494431	chr4:8886138-8887652	red	cultivated watermelon
RZ-900	SRR494432	chr4:8886138-8887652	red	cultivated watermelon
LSW-177	cloned in this study		red	cultivated watermelon
PI97103	cloned in this study		red	cultivated watermelon
Dumara	cloned in this study		red	cultivated watermelon
ZXG00055	cloned in this study		red	cultivated watermelon
ZXG00079	cloned in this study		red	cultivated watermelon
ZXG01452	cloned in this study		red	cultivated watermelon
ZXG01594	cloned in this study		red	cultivated watermelon
RZ-901	SRR494425	chr4:8886138-8887652	red	cultivated watermelon
PI482271	SRR494434	chr4:8886138-8887652	white	semi-wild watermelon
PI189317	SRR494441	chr4:8886138-8887652	white	semi-wild watermelon
PI500301	SRR494444	chr4:8886138-8887652	white	semi-wild watermelon
PI249010	SRR494443	chr4:8886138-8887652	white	semi-wild watermelon
PI248178	SRR494446	chr4:8886138-8887652	white	semi-wild watermelon
PI186490	cloned in this study		white	semi-wild watermelon
ZXG0000	cloned in this study		white	cultivated watermelon
ZXG01555	cloned in this study		green	cultivated watermelon
PI482276b	SRR494437	chr4:8886138-8887652	green	wild watermelon
PI482303	SRR494442	chr4:8886138-8887652	green	wild watermelon
PI296341-FR	SRR494435	chr4:8886138-8887652	green	wild watermelon
PI482326	SRR494445	chr4:8886138-8887652	yellow	wild watermelon
COS	cloned in this study		pale yellow	cultivated watermelon
ZXG00077	cloned in this study		yellow	cultivated watermelon

High-performance liquid chromatography (HPLC) profile of carotenoids in *E. coli:* Transformed *E. coli* cells containing plasmid pET-Cllcyb were cultured at 37°C on Luria-Bertani (LB) solid medium supplemented with 100 μg mL^{-1} ampicillin and 35μg mL^{-1} chloramphenicol. Single colony was grown into LB liquid medium with corresponding antibiotics and shaken at 37°C for 2 d in darkness. The cells were pelleted by centrifugation (5000×g for 15 min at 4°C) and washed with sterile distilled water to remove medium. After adding 6 ml of methanol containing 0.1% butylatedhydroxytoluene (BHT), the lysate was extracted five times with 3 ml of petroleum ether containing 2% dichloromethane and centrifuged into separate phases (5000×g for 3 min at 4°C). The supernatant extract were collected together, and then dried in a centrifugal evaporator at 30°C. The carotenoids were dissolved with 2 ml of methanol/acetonitrile/ dichloromethane (20/50/30, v/v/v) containing 0.1% BHT, and filtered through a 0.22 μm syringe filter (Millipore, Germany) into HPLC vials. Finally, the liquids were stored at -80°C prior to HPLC analysis. The same procedure was used to extract the carotenoids from bacteria containing pACCRT-EIB and pACCAR16-ΔcrtX plasmids that enabled them to biosynthesize lycopene and β-carotene, respectively. The standards were purchased from Sigma-Aldrich. The amount of carotenoids was calculated by integrating the total peak area of various carotenoids detected by HPLC at a wavelength of 450 nm, concerning the standard curves of lycopene and β-carotene.

Results

Cloning and characterization of Cllcyb: According to the available information, a pair of specific primers for the *lcyb* gene was designed, and cDNA fragments from LSW-177 and COS were amplified by RT-PCR. After sequence confirmation, 1515-bp cDNA clones for LSW-177 and COS, each containing an ORF, were isolated and named *Cllcyb-177* and *Cllcyb-COS,* respectively. The deduced Cllcyb protein contains 504 amino acids, with a theoretical molecular weight (MW) of 56.46 kDa and a predicted isoelectric point (*pI*) of 8.46. No signal peptide was found by SignalP 4.1 Server. Meanwhile, we used the SABLE server to predict the secondary structures of the protein and transmembrane domains (Fig. 1). Glycosylation prediction revealed eight positive O-GlcNAc sites in Cllcyb-177 and seven sites in Cllcyb-COS, and most of the sites were located at the N terminus (Table 1).

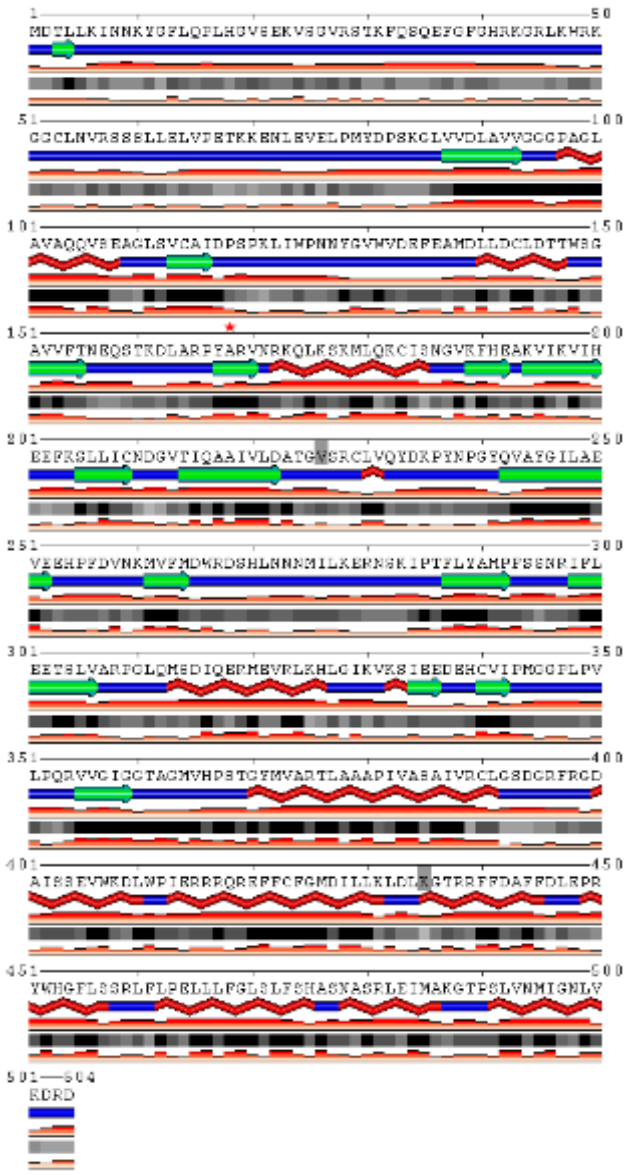

Fig. 1. The secondary structure of the deduced protein Cllcyb-177 according to the SABLE website (http://sable.cchmc.org). The upper line shows the amino acid sequence. The second line indicates the secondary structure (α helices with red colour, β strands with green for and blue for random coils). The relative solvent accessibility, physicochemical properties and confidence level of the prediction are arranged in order. The different amino acids between Cllcyb-177 and -COS were indicated with gray shadow and the key amino acid for lcyb function in watermelon was marked with a red pentagram.

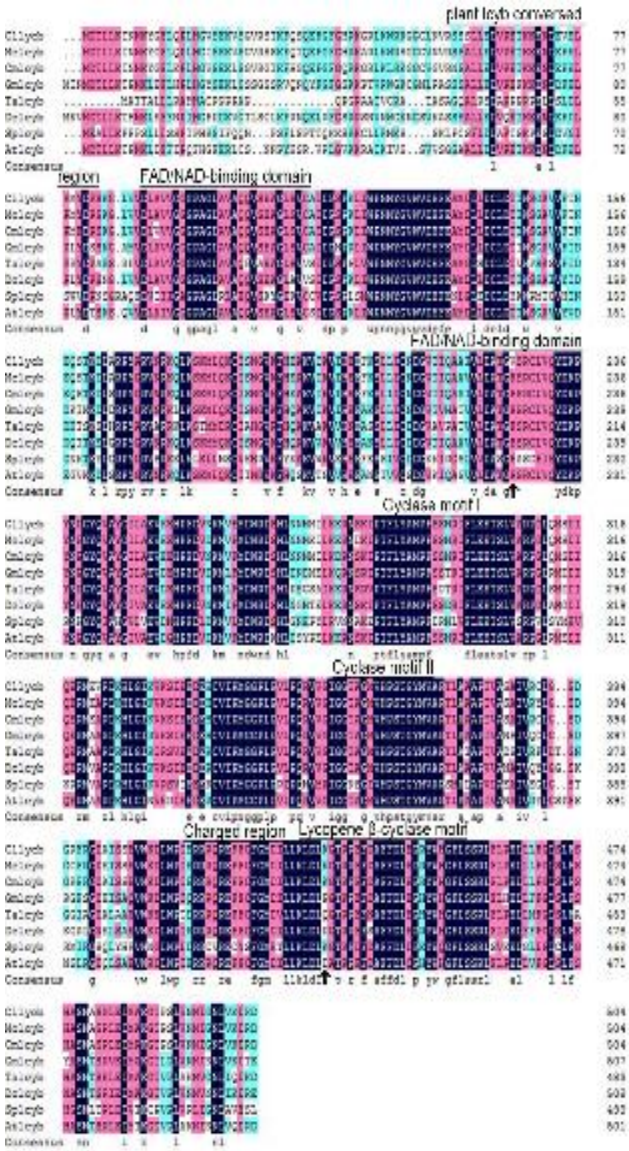

Fig. 2. Analysis of conserved domains in plant lcyb proteins. The GenBank accession numbers for the plant lcyb sequences are below: Atlcyb, *Arabidopsis thaliana,* NP_187634.1; Mclcyb, *Momordica charantia*, XP 022153903.1; Cmlcyb, *Cucumis melo*, ADN34000.1; Gmlcyb, *Glycine max*, KRH66793.1; Dclcyb, *Daucus carota,* ABB52071.1; Talcyb, *Triticum aestivum,* AFA35121.1; Splcyb, *Solanum pennellii*, ADZ24717.1. Conserved residues are shaded in black and consensus residues in below. Red, blue and white colours indicate homology from high to low. The dark arrows indicate the amino acids that differ between Cllcyb-177 and Cllcyb-COS.

Several conserved motifs considered as characteristics of plant lycopene cyclases were also observed in the watermelon lcyb protein (Fig. 2): a plant LCYB conserved region; a dinucleotide FAD/NAD-binding domain ($DX_4GXGXAX_4A$); cyclase motif I and II and a charged region (Bouvier, Florence *et al*., 1996; Chen, 1995; Cunningham *et al*., 1996). Comparison of the *Cllcyb-177* and *Cllcyb-COS* ORFs revealed that three nucleotide variances resulted in two amino acid mutations. The first amino acid change occurred in the dinucleotide FAD/NAD-binging signature and the second in the lycopene β-cyclase motif, which indicated that these changes might be correlated with enzyme activity (dark arrows showed in Fig. 2.). Alignment analysis by the NCBI web site showed that Cllcyb shares 96% identity with *Momordica charantia* lcyb (GenBank accession: XP_022153903.1), 96% with *Cucumis melo* lcyb (GenBank accession: ADN34000.1), 82% with *Glycine max* lcyb (GenBank accession: KRH66793.1), 80% with *Daucus carota* lcyb (GenBank accession: ABB52071.1), 76% with *Arabidopsis thaliana* lcyb (GenBank accession: NP_187634.1), 74% with *Triticum aestivum* lcyb (GenBank accession: AFA35121.1), and 57% with *Solanum pennellii* lcyb (GenBank accession: ADZ24717.1).

Table 1. Summary of predicted glycosylation sites in Cllcyb-177 and -COS.

Residue	O-GlcNac result	Potential	Cllcyb-177	Cllcyb-COS
Ser21	+	0.5072	Y	Y
Ser25	++	0.6208	Y	Y
Thr30	+	0.4611	Y	Y
Ser107	++	0.5973	Y	Y
Ser149	+	0.5322	Y	Y
Thr155	+	0.5221	Y	Y
Ser227	++	0.6052	Y	N
Ser294	+	0.565	Y	Y

Y indicates a glycosylation site; N indicates a non-glycosylation site

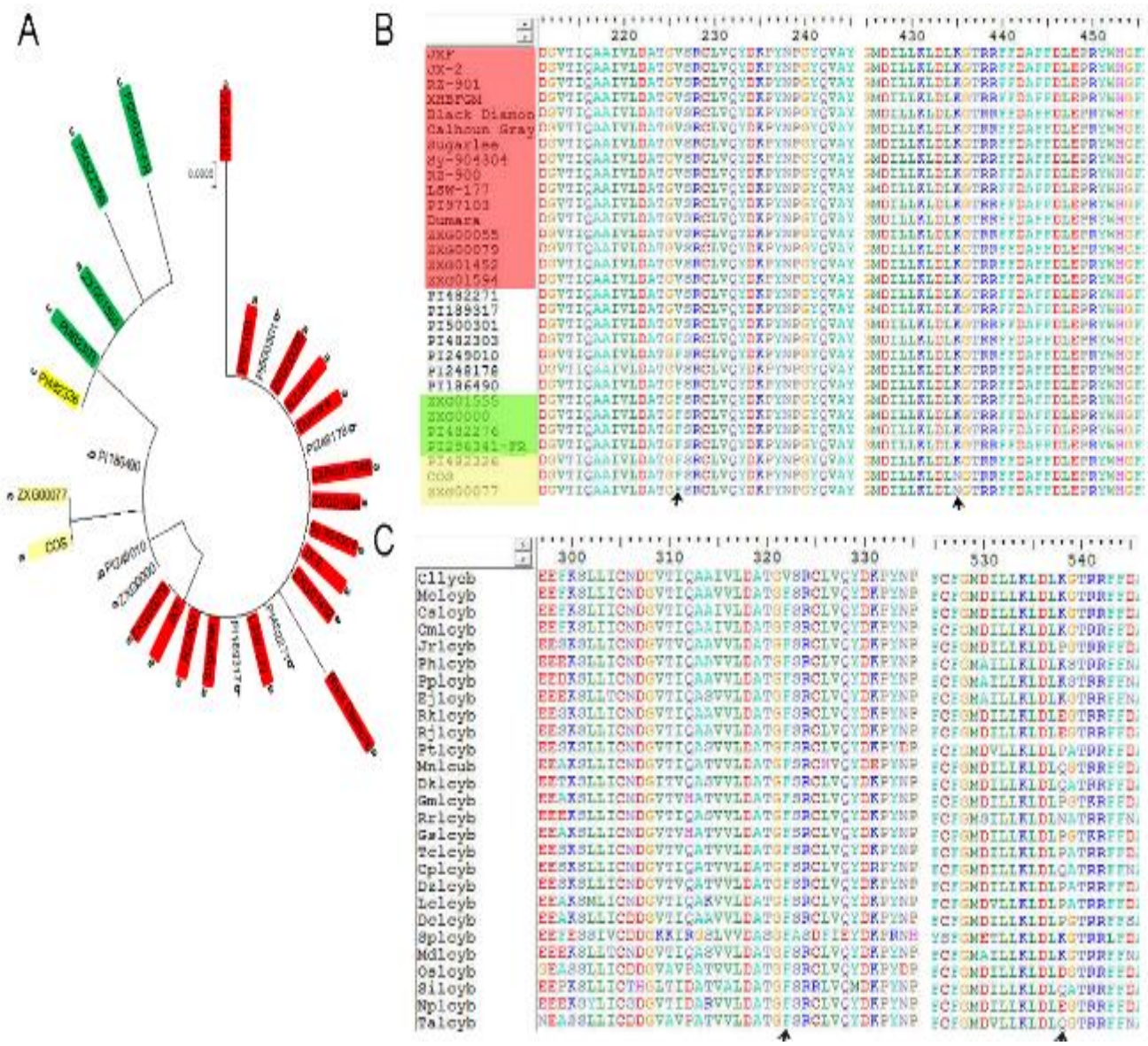

Fig. 4. Relationships among the watermelon lcyb protein sequences, flesh colour and cultivars. A, phylogenetic tree constructed with the alignment of deduced amino acid sequences from 30 watermelon cultivars. Red colour indicates red flesh, green indicates green flesh, yellow indicates yellow flesh, and white indicates white flesh. The 30 watermelon cultivars are also classified into cultivated (a), semi-wild (b) and wild groups (c). B, sequence comparison of lcyb proteins from 30 watermelon cultivars. C, sequence comparison of plant lcyb proteins. The dark arrowsdenote the different amino acids discussed in the text.

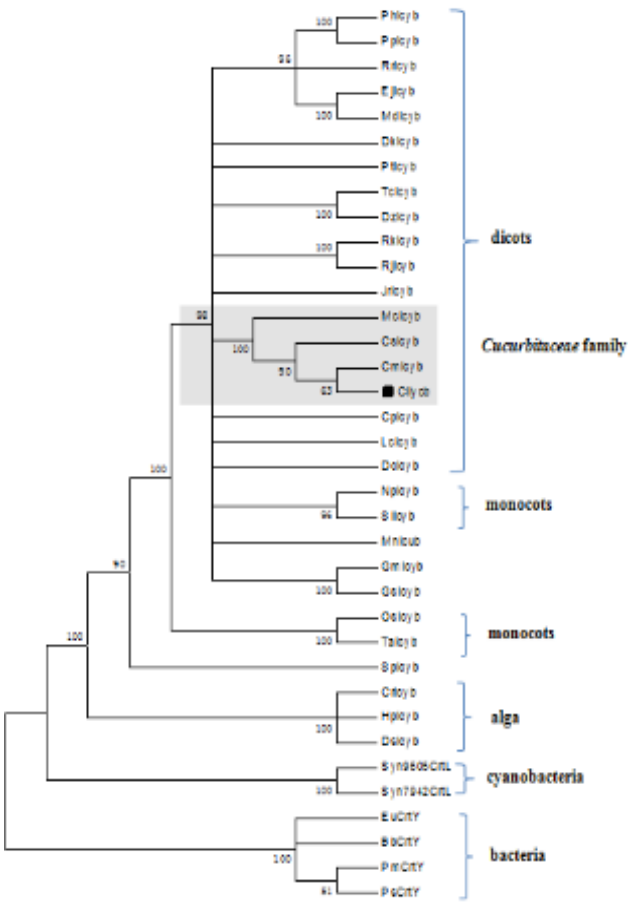

Fig. 3. The phylogenetic tree of plant, cyanobacteria, algal and bacterial lycopene β-cyclase proteins. The tree was constructed upon the basis of the neighbor-joining method by the program MEGA 4.0. Accession numbers for the β-cyclases are as follows: *Citrullus lanatus* (Cllycb, ABM90917.1); *Momordica charantia* (Mclcyb, XP_022153903.1); *Cucumis sativus* (Cslcyb, KGN53976.1); *Cucumis melo* subsp. Melo (Cmlcyb, ADN34000.1); *Juglans regia* (Jrlcyb, XP_018817188.1); *Prunus humilis* (Phlcyb, ANH11477.1); *Prunus persica* (Pplcyb, ONH95007.1); *Eriobotrya japonica* (Ejlcyb, AFP43698.1); *Rhododendron kiusianum* x Rhododendron indicum (Rklcyb, BAS69576.1); *Rhododendron japonicum* f. flavum (Rjlcyb, BAS69435.1); *Populus trichocarpa* (Ptlcyb, EEF05021.1); *Morus notabilis* (Mnlcub, EXC00005.1); *Diospyros kaki* (Dklcyb, ACR25158.1); *Glycine max* (Gmlcyb, KRH66793.1); *Rosa rugosa* (Rrlcyb, AKT74331.1); *Glycine soja* (Gslcyb, KHN30373.1); *Theobroma cacao* (Tclcyb, EOY09059.1); *Carica papaya* (Cplcyb, ABD91578.1); *Durio zibethinus* (Dzlcyb, ARH12814.1); *Lycium chinense* (Lclcyb, AKO73675.1); *Daucus carota* (Dclcyb, ABB52071.1); *Solanum pennellii* (Splcyb, ADZ24717.1); *Malus domestica* (Mdlcyb, ANS58096.1); *Oryza sativa* Japonica Group (Oslcyb, BAD16478.1); *Setaria italic* (Silcyb, AAQ02668.1); *Narcissus pseudonarcissus* (Nplcyb, ACT78995.1); *Triticuma estivum* (Talcyb, AFA35121.1); *Chlamydomonas reinhardtii* (Crlcyb, AAX54906); *Haematococcus pluvialis* (Hplcyb, AAO64977); *Dunaliella salina* (Dslcyb, ACA34344); *Synechococcus* sp. CC9605 (Syn9605CrtL, ABB35682); *Synechococcus elongatus* PCC7942 (Syn7942CrtL, CAA52677); *Paracoccus marcusii* (PmCrtY, CAB56061); *Pantoea stewartii* (PsCrtY, AAN85598); *Pantoea ananatis* (EuCrtY, BAA14126); *Bdellovibrio bacteriovorus* (BbCrtY, CAE79599).The grey shadow marks the *Cucurbitaceae* family.

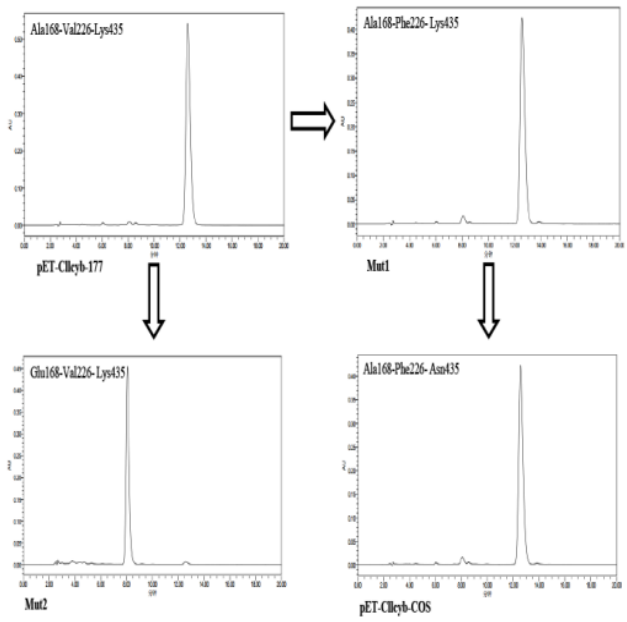

Fig. 5. Functional analysis of Cllcyb-177, Cllcyb-COS and two mutants expressed in *E. coli*. HPLC peaks of carotenoids extracted from cultures of *E. coli* carrying plasmids pET-Cllcyb-177 (Ala168-Val226-Lys435), Cllcyb-COS (Ala168-Phe226-Asn435), Mut1 (Ala168-Phe226-Lys435) and Mut2 (Glu168-Val226-Lys435). The retention time of 12.5 min corresponds to the β-carotene peak, and the one at 8.0 min corresponds to the lycopene peak. Mut1 and Mut2 were obtained by site-directed mutagenesis using pET-Cllcyb-177 as a template.

Supplemental Materials

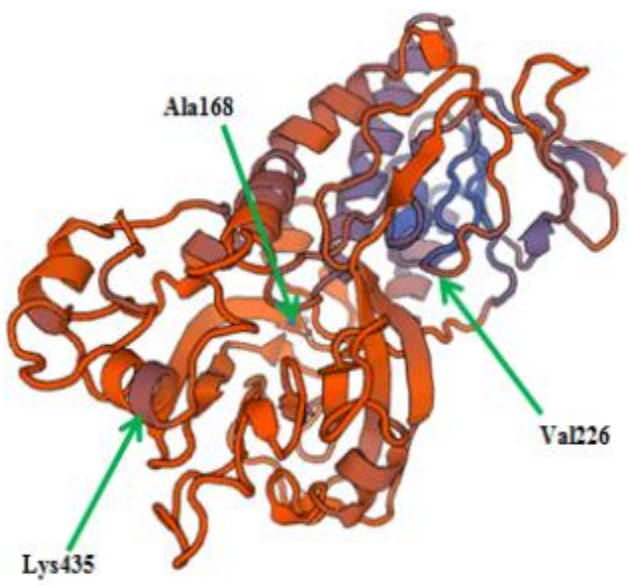

Fig. S1. The prediction of 3-D structure of the active site of lycopene β-cyclase from watermelon. The figure was generated by SWISS-MODEL. The green arrows indicate the important amino acids mentioned in the text.

Phylogenetic analysis of lcyb proteins: The lycopene cyclization acts as the crucial node in carotenoid biosynthesis pathway, and lycopene β-cyclases had been identified in high plant, cyanobacteria, green algal and bacteria. In this study, we examined the evolution of lcybs in different species, in different plant groups and

also in watermelon with different fruit flesh colours. Phylogenetic trees were generated by the NJ method using MEGA4.0. The *O. sativa* japonica group and *Z. mays* were considered as monocot representatives of lcybs(Garg *et al*., 2010; Suwarno & Pixley, 2015), and lcybs from *V. vinifera* and *S. lycopersicum* were utilized as dicot representatives (Jaillon *et al*., 2007; Liu *et al*., 2015; Pecker *et al*., 1996).

Initially, the lcybsfrom high plant, cyanobacteria, green algal and bacterial were evaluated (Fig. 3) and the constructed phylogenetic tree revealed that these lcybs formed their own clusters. In higher plants, all lcybs clustered into two subgroups, with monocots and dicots following different evolutionary paths. In addition, the proximity of the higher plant domain supported the existence of an ancient lcyb, which may be the common ancestor of monocots and dicots lcyb. The lcyb protein from watermelon exhibited high homology with the proteins from *M. charantia*, *C. sativus*, and *C. melo* subsp. melo, which are specifically grouped together in the *Cucurbitaceae* family.

Different watermelon groups including wild, semi-wild and cultivated species were also selected for analysis of evolutionary relationships. A total of 18 lcyb proteins were obtained from the NCBI sequence read archive (SRA) stated by Guo *et al*., (2013), and 12 proteins were cloned and sequenced in our lab (Table S1). After the final confirmation, phylogenetic trees were constructed. All watermelon cultivars clustered into two branches: red flesh and non-red flesh (Fig. 4A). The results depicted that *lcyb* alleles are co-segregated perfectly with colour phenotype and the CAPS marker located in the gene indicating that the gene plays a crucial function for the formation of flesh colour. Therefore, multiple alignments of all lcybs from wild, semi-wild and cultivated varieties with red, white, green, yellow, and orange flesh colours were performed (Fig. 4B). The residue Asn435 was only found in COS (pale-yellow flesh) and ZXG0077 (yellow flesh) but a Lys435 existed at the same position in red-fleshed LSW-177 and other watermelon cultivars. Another contrasted amino acid in LSW-177 is Val226 and oppositeto Phe226 in COS. Most of the red-fleshed watermelon cultivars contain the same amino acid as found in LSW-177, however the presence of Val226 was also found in white-fleshed watermelon (PI482271, PI189317, PI500301, PI248178), indicating that Val226 is not strictly consistent with flesh colour in watermelon.

Heterologous complementation and site-directed mutagenesis of watermelon lcybs in *E. coli*: The well-engineered strain of *E. coli* support various carotenogenic substrates for characterization of genes in the carotenoids biosynthetic pathway, which is a convenient, rapid and efficient method. To investigate whether the observed variances in amino acids lead to changes in catalytic function in LSW-177 and- COS, the ORFs were inserted into the expression vector pET-28a (+) and later these recombinants were transformed into an *E. coli* strain harbouring the plasmid pACCRT-EIB, which was engineered to cause cells to accumulate the reddish-pink pigment lycopene. An intermediate mutant containing Phe226/Lys435 (named Mut1) was generated based on Val226/Lys435 in 177 and Phe226/Asn435 in COS. All three vectors were applied to assess the cyclase ability and the accumulated carotenoids in each cell line were prepared for HPLC analysis (Fig. 5). The bacteria containing pACCRT-EIB and pACCAR16-ΔcrtX produced lycopene and β-carotene, respectively, and these two were used as a control. The results showed that cells harboringpET-Cllcyb-COS, pET-Cllcyb-177 and pET-Mut1 could fully converted lycopene into β-carotene in 48 h and had no difference in catalytic activity among the three bacterial lines. These results indicated that the presence of Val versus Phe at position 226th and the presence of Lys versus Asn at position 435th do not change the function of the Cllcyb enzyme. Interestingly, only the lycopene peak was observed by HPLC and the amount of β-carotene was traced in another site-directed mutation changed Ala168 in Cllcyb-177 to Glu168 (Mut2 in Fig. 5). These results indicated that lycopene was not cyclased into β-carotene. Thus, Ala168 in watermelon lcyb determined the catalytic activity towards lycopene.

Discussion

Numerous studies have supported the hypothesis that the activity of lycopene cyclases determines the carotenoid contents and profiles in fruit, flower and root, *et al*., (Bouvier *et al*., 2005; Cunningham, Francis, 2002). Regarding the same function of lcyb in canary yellow and red watermelon fruit (Bang *et al*., 2007), novel CAPS markers (named *WII04E08-38* and *WII04EBsaHI-6*) were developed that can effectively distinguish the red colourfruits from others in the F_2 segregated population (Liu *et al*., 2016). These results indicated that *lcyb* may function as the red colour dominant gene in watermelon.

According to the results from Zhu *et al*., (2017), the genetic diversity of *lcyb* gene sequences was calculated by the values of Pi, θ, and Tajima's D. The data decreased in the population of wild and semi-wild watermelon accessions from 0.0044, 0.0039, and 0.6273 to 0.0004, 0.0007, and -1.5622, respectively, compared to cultivated watermelon, suggesting that the diversity of the *lcyb* gene had been decreasing during evolution. Nucleotide diversity leading to differences in amino acid sequence may change the enzymatic cability. In our study, we successfully cloned the *lcyb* gene from watermelon cultivars (LSW-177 and COS) with distinct red and pale-yellow flesh. Three nucleotide differences had been identified and two amino acid substitutions were present in conserved motifs, which may be related to enzyme activity. Meanwhile, two phylogenetic branches were generated in correlating amino acids with various watermelon flesh colours, and all the red-fleshed lines grouped together. Interestingly, some white-flesh watermelon (PI482271, PI189317, PI500301, PI248178) were found to be grouped with red ones. The white flesh of these watermeloncultivars may be unrelated to lcyb function and may instead be caused by inactivation of carotenoid biosynthesis genes. Asn435 was found only in COS and ZXG0077, whereas other cultivars (including LSW-177) contained Lys at the same position; however, the residue at position 435th was strictly conserved in the included watermelon cultivars.

All red-fleshed watermelon displayed a Val at position 226th; some white-coloured watermelon cultivars have the same amino acid as red-flesh cultivars, and these cultivars belong to the semi-wild group. The two substitutions were also checked in other plant lcybs (Fig. 4C). Our results showed position 435th to be variable in plant lcybs. By contrast, the amino acid at position 226th is strictly conserved, and Phe is found in all included proteins, emphasizing its vital role in enzyme function. The amino acid Val226 was changed to Phe226 by site-directed mutagenesis, and HPLC profiles of extracts showed that the three proteins were able to convert lycopene with high efficiency. Although processing of the putative watermelon *lcyb*gene in a prokaryotic system most likely differs from what happens in the plant system, the expressed protein obtained performed the function efficiently. It is speculated that the primary structure of Cllcyb in *E. coli* was sufficient for activity towards lycopene, with no need for other modifications. However, these results do not explain whether the observed amino acid differences are responsible for red and non-red flesh in watermelon. Comparative transcriptome and real-time PCR analysis of LSW-177 and COS revealed that the *lcyb* gene expression had no significantly difference in fruit development and ripening, and the results were also confirmed by RNA sequencing for 97103 and Dumara (Guo, Shaogui *et al*., 2011; Guo *et al*., 2015; Zhu *et al*., 2017). Regulation of post-translational modification of the lcyb protein may be responsible for the difference in flesh colour between the two genotypes. Compared with the predicted glycosylation sites, only one site, at the position 227th, differed between Cllcyb-177 and Cllcyb-COS (Table 1). Val226 in Cllcyb-177 and Ser227 can be predicted to be glycosylated, and Phe226 abolishes this glycosylation site. Regardless, this speculation requires further confirmation. Additionally, it needs to be further explored whether the two divergences at the positions 226th and 435th influence the lcyb three-dimensional structure or the binding of FAD or NAD (P).

Despite the separation of monocot and dicot plants in the phylogenetic tree, the functional domains confirmed in plant lcybs showed highly conserved, which are involved in many reactions in β-ionone catalysis. It has been reported that two-nucleotide insertion resulted in a prematurely truncated lcyb2 protein, causing an nonfunctionalprotein in red papaya (Devitt *et al*., 2010). In citrus, the differential efficiency of Cslcyb1 and Cslcyb2 activity towards lycopene maybe related to the 23-nt sequence variation, resulting in 11-amino acid substitutions (Zhang *et al*., 2013). However, there are no reports to date demonstrating the amino acids or motifs that are critical for lycopene cyclase function.The present study significantly exposed a nonfunctional protein due to the change of Ala168 in watermelon lcyb to Glu168 *via* site-directed mutagenesis. Secondary structure analysis showed that Ala168 was located in the middle of a β-strand (red pentagram in Fig. 1), which was a part of the catalytic core region by protein modelling (Fig. S1). Therefore, *Cllcyb* genes can be manipulated to increase or decrease β-carotene contents in watermelon fruit.

Conclusions

The current results expressed the understanding of the relationship betweenlcyb protein sequences and watermelon flesh colour, and the characterization of *Cllcyb*genes from two different watermelon cultivars (red-flesh LSW-177 and pale-yellow flesh COS). Glycosylation of Cllcybmay have a pivotal role in lycopene cyclization. The critical amino acid Ala168 in the Cllcyb protein may be located in the catalytic core region, conferring lycopene cyclase capacity.

Acknowledgments

The research was supported by the National Nature Science Foundation of China (31572144) and by the China Agriculture Research System (CARS-25). We are extremely grateful to Dr Angela R. Davis (South Central Agricultural Research Laboratory, Agricultural Research Service, United States Department of Agriculture) for the provisions of the LSW-177 and COS seeds. Furthermore, we are grateful to Professor Misawa (Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Japan) for kindly providing the plasmids pACCRT-EIB and pACCAR16-ΔcrtX.

References

Agarwal, S. and A.V. Rao. 2000. Tomato lycopene and its role in human health and chronic diseases. *CMAJ*., 163: 739.

Alquézar, B., L. Zacarías and M.J. Rodrigo. 2009. Molecular and functional characterization of a novel chromoplast-specific lycopene beta-cyclase from Citrus and its relation to lycopene accumulation. *J. Exp. Bot*., 60: 1783.

Ampomah-Dwamena, C., T. McGhie, R. Wibisono, M. Montefiori, R.P. Hellens and A.C. Allan. 2009. The kiwifruit lycopene beta-cyclase plays a significant role in carotenoid accumulation in fruit. *J. Exp. Bot*., 60: 3765-3779.

Bang, H., S. Kim and L.S. King. 2007. Development of a codominant CAPS marker for allelic selection between canary yellow and red watermelon based on SNP in lycopene β-cyclase (LCYB) gene. *Mol. Breed*., 20: 63-72.

Bang, H.J., A.R. Davis, S.G. Kim, D.I. Leskovar and S.R. King. 2010. Flesh color inheritance and gene interactions among canary yellow, pale yellow, and red watermelon. *Infect. Immun*., 65: 4620-4623.

Bartley, G.E. and P.A. Scolnik. 1995. Plant carotenoids: pigments for photoprotection, visual attraction, and human health. *Plant Cell*, 7: 1027.

Blas, A.L., R. Ming, Z. Liu, O.J. Veatch, R.E. Paull, P.H. Moore and Q. Yu. 2010. Cloning of the papaya chromoplast-specific lycopene beta-cyclase, CpCYC-b, controlling fruit flesh color reveals conserved microsynteny and a recombination hot spot. *Plant Physiol*., 152: 2013-2022.

Bouvier, F., A. D'Harlingue, P. Hugueney, E. Marin, A. Marionpoll and B. Camara. 1996. Xanthophyll biosynthesis cloning, expression, functional reconstitution, and regulation of β-cyclohexenyl carotenoid epoxidase from pepper (*Capsicum annuum*). *J. Biol. Chem*., 271: 28861-28867.

Bouvier, F., A. Rahier and B. Camara. 2005. Biogenesis, molecular regulation and function of plant isoprenoids. *Prog. Lipid Res*., 44: 357.

Britton, G., S. Liaaen-Jensen, H. Pfander, A.Z. Mercadante and E.S. Egeland. 2004. Carotenoids: Handbook. Springer. Basel.

Cazzonelli, C.I. and B.J. Pogson. 2010. Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci*., 15: 266-274.

Chen, H.C. 1995. Metabolism of cyclic carotenoids: a model for the alteration of this biosynthetic pathway in *Capsicum annuum* chromoplasts. *Plant J*., 8: 417-424.

Cunningham, F.X. 2002. Regulation of carotenoid synthesis and accumulation in plants. *Pure Appl. Chem*., 74: 1409-1417.

Cunningham, F.X., B. Pogson, Z. Sun, K.A. Mcdonald, D. Dellapenna and E. Gantt. 1996. Functional analysis of the beta and epsilon lycopene cyclase enzymes of Arabidopsis reveals a mechanism for control of cyclic carotenoid formation. *Plant Cell*, 8: 1613.

Devitt, L.C., K. Fanning, R.G. Dietzgen and T.A. Holton. 2010. Isolation and functional characterization of a lycopene β-cyclase gene that controls fruit colour of papaya (*Carica papaya* L.). *J. Exp. Bot*., 61: 33.

Diretto, G., S. Al-Babili, R. Tavazza, V. Papacchioli, P. Beyer and G. Giuliano. 2007. Metabolic engineering of potato carotenoid content through tuber-specific overexpression of a bacterial mini-pathway. *Plos One*, 2: e350.

Frank, H.A. and R.J. Cogdell. 1996. Carotenoids in photosynthesis. *Photochem. Photobiol*., 63: 257-264.

Fraser, P.D. and P.M. Bramley. 2004. The biosynthesis and nutritional uses of carotenoids. *Prog. Lipid Res*., 43: 228.

Fraser, P.D., S. Romer, C.A. Shipton, P.B. Mills, J.W. Kiano, N. Misawa, R.G. Drake, W. Schuch and P.M. Bramley. 2002. Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proc. Natl. Acad. Sci*., USA., 99: 1092-1097.

Garg, R., S. Jhanwar, A.K. Tyagi and M. Jain. 2010. Genome-wide survey and expression analysis suggest diverse roles of glutaredoxin gene family members during development and response to various stimuli in rice. *DNA Res*., 17: 353-367.

Guo, S., H. Sun, H. Zhang, J. Liu, Y. Ren, G. Gong, C. Jiao, Y. Zheng, W. Yang, Z. Fei and Y. Xu. 2015. Comparative transcriptome analysis of cultivated and wild watermelon during fruit development. *PLoS One*, 10: e0130267.

Guo, S., J. Liu, Y. Zheng, M. Huang, H. Zhang, G. Gong, H. He, Y. Ren, S. Zhong and Z. Fei. 2011. Characterization of transcriptome dynamics during watermelon fruit development: sequencing, assembly, annotation and gene expression profiles. *BMC Genom.*, 12: 454.

Guo, S., J. Zhang, H. Sun, J. Salse, W.J. Lucas, H. Zhang, Y. Zheng, L. Mao, Y. Ren, Z. Wang, J. Min, X. Guo, F. Murat, B.K. Ham, Z. Zhang, S. Gao, M. Huang, Y. Xu, S. Zhong, A. Bombarely, L.A. Mueller, H. Zhao, H. He, Y. Zhang, Z. Zhang, S. Huang, T. Tan, E. Pang, K. Lin, Q. Hu, H. Kuang, P. Ni, B. Wang, J. Liu, Q. Kou, W. Hou, X. Zou, J. Jiang, G. Gong, K. Klee, H. Schoof, Y. Huang, X. Hu, S. Dong, D. Liang, J. Wang, K. Wu, Y. Xia, X. Zhao, Z. Zheng, M. Xing, X. Liang, B. Huang, T. Lv, J. Wang, Y. Yin, H. Yi, R. Li, M. Wu, A. Levi, X. Zhang, J.J. Giovannoni, J. Wang, Y. Li, Z. Fei and Y. Xu. 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet*., 45: 51-58.

Hao, H., L. Ma, H. Cong, Q. Li and X. Yu. 2012. Isolation and characterization of a muskmelon cDNA encoding Lycopene Beta-cyclase. *Gene*, 503: 147-151.

Harjes, C.E., T.R. Rocheford, L. Bai, T.P. Brutnell, C.B. Kandianis, S.G. Sowinski, A.E. Stapleton, R. Vallabhaneni, M. Williams and E.T. Wurtzel. 2008. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science,* 319: 330.

Jaillon, O., J.M. Aury, B. Noel, A. Policriti, C. Clepet, A. Casagrande, N. Choisne, S. Aubourg, N. Vitulo, C. Jubin, A. Vezzi, F. Legeai, P. Hugueney, C. Dasilva, D. Horner, E. Mica, D. Jublot, J. Poulain, C. Bruyere, A. Billault, B. Segurens, M. Gouyvenoux, E. Ugarte, F. Cattonaro, V. Anthouard, V. Vico, C. Del Fabbro, M. Alaux, G. Di Gaspero, V. Dumas, N. Felice, S. Paillard, I. Juman, M. Moroldo, S. Scalabrin, A. Canaguier, I. Le Clainche, G. Malacrida, E. Durand, G. Pesole, V. Laucou, P. Chatelet, D. Merdinoglu, M. Delledonne, M. Pezzotti, A. Lecharny, C. Scarpelli, F. Artiguenave, M.E. Pe, G. Valle, M. Morgante, M. Caboche, A.F. Adam-Blondon, J. Weissenbach, F. Quetier, P. Wincker and C. French-Italian Public Consortium for Grapevine Genome. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature,* 449: 463-467.

Kato, M., Y. Ikoma, H. Matsumoto, M. Sugiura, H. Hyodo and M. Yano. 2004. Accumulation of carotenoids and expression of carotenoid biosynthetic genes during maturation in citrus fruit. *Plant Physiol*., 134: 824.

Lange, B.M. and M. Ghassemian. 2003. Genome organization in *Arabidopsis thaliana*: a survey for genes involved in isoprenoid and chlorophyll metabolism. *Plant Mol. Biol*., 51: 925-948.

Liu, L., Z. Shao, M. Zhang and Q. Wang. 2015. Regulation of carotenoid metabolism in tomato. *Mol. Plant*, 8: 28-39.

Liu, S., P. Gao, Q. Zhu, F. Luan, A.R. Davis and X. Wang. 2016. Development of cleaved amplified polymorphic sequence markers and a CAPS-based genetic linkage map in watermelon (*Citrullus lanatus* [Thunb.] Matsum. and Nakai) constructed using whole-genome re-sequencing data. *Breed. Sci*., 66: 244-259.

Lu, S. and L. Li. 2008. Carotenoid metabolism: biosynthesis, regulation, and beyond. *J. Integr. Plant. Biol.*, 50: 778-785.

Nisar, N., L. Li, S. Lu, Nay C. Khin and Barry J. Pogson. 2015. Carotenoid metabolism in plants. *Mol. Plant*, 8: 68-82.

Paniagua-Michel, J., J. Olmos-Soto and M.A. Ruiz. 2012. Pathways of carotenoid biosynthesis in bacteria and microalgae. *Methods Mol. Biol*., 892: 1-12.

Pecker, I., R. Gabbay, C.F. Jr and J. Hirschberg. 1996. Cloning and characterization of the cDNA for lycopene beta-cyclase from tomato reveals decrease in its expression during fruit ripening. *Plant Mol. Biol*., 30: 807.

Porollo, A.A., R. Adamczak and J. Meller. 2004. POLYVIEW: a flexible visualization tool for structural and functional annotations of proteins. *Bioinformatics,* 20: 2460-2462.

Robert, B., P. Horton, A.A. Pascal and A.V. Ruban. 2004. Insights into the molecular dynamics of plant light-harvesting proteins in vivo. *Trends Plant Sci*., 9: 385-390.

Ronen, G., L. Carmelgoren, D. Zamir and J. Hirschberg. 2000. An alternative pathway to β-carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. *Proc. Natl. Acad. Sci*., USA 97: 11102.

Ronen, G.M., D. Zamir and J. Hirschberg. 1999. Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant Delta. *Plant J*., 17: 341-351.

Suwarno, W.B. and K.V. Pixley. 2015. Genome-wide association analysis reveals new targets for carotenoid biofortification in maize. *Theor. Appl. Genet*., 128: 851-864.

Tamura, K., J. Dudley, M. Nei and S. Kumar. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol*., 24: 1596-1599.

Zhang, J.C., W.J. Zhou, Q. Xu, N.G. Tao, J.L. Ye, F. Guo, J. Xu and X.X. Deng. 2013. Two lycopene β-cyclases genes from sweet orange (*Citrus sinensis* L. Osbeck) encode enzymes with different functional efficiency during the conversion of lycopene-to-provitamin A. *J. Integr. Agri*., 12: 1731-1747.

Zhu, Q., P. Gao, S. Liu, Z. Zhu, S. Amanullah, A.R. Davis and F. Luan. 2017. Comparative transcriptome analysis of two contrasting watermelon genotypes during fruit development and ripening. *BMC Genom*., 18: 3.

(Received for publication 2 April 2018)