Allelic diversity of phytoene synthase gene influences the transcription level in citrus fruit among a citrus F1 hybrid population

Aiko Sugiyama1), Yoshinori Ikoma2), Hiroshi Fujii3), Tomoko Endo3), Hirohisa Nesumi2), Takehiko Shimada4),5) and Mitsuo Omura3)

1) The United Graduate School of Agriculture Science, Gifu University, Gifu 501-1193, Japan
2) National Agriculture and Food Research Organization Institute of Fruit Tree and Tea Science, Shimizu, Shizuoka 424-0292, Japan
3) Faculty of Agriculture, Shizuoka University, Suruga, Shizuoka 422-8529, Japan

Phytoene synthase (PSY) is one of the key regulatory enzymes on the biosynthesis and accumulation of carotenoids in citrus fruits. The transcriptional diversity of PSY is mainly attributed to the structural variation in promoter region among PSY alleles. In aim to clarify how this transcriptional diversity is regulated among them, PSY alleles responsible for carotenoid biosynthesis in the fruits are characterized and their promoter sequences were compared. Based on gene structure and expression pattern of PSY homologues on the clementine mandarin genome sequence, PSY alleles responsible for carotenoid biosynthesis are derived from a single locus in the scaffold 6. AG mapping population possessed four PSY alleles derived from parent lines of A255 and G434, and their F1 individuals with PSY-g2 allele tended to have low transcription level. From sequence comparison of their promoter regions, the cis-motif alternation from MYBPZM to RAV1AAT might be a candidate to influence the transcription level. Among the ancestral pedigree varieties of AG mapping population, the transcription level of PSY correlated with genotypes of MYBPZM and RAV1AAT motifs in the promoter region of PSY alleles, so that homozygous genotype of MYBPZM showed higher transcription level while heterozygous genotype of MYBPZM and RAV1AAT showed lower transcription level.

Key Words: citrus, carotenoid, allele, phytoene synthase, cis-regulatory motif.

Introduction

Most photosynthetic organisms produce carotenoids, which are essential for both plant and animal lives (Wong et al. 2004). Carotenoids are involved in various functions in plants, including phyto-hormone precursor action (Schwartz et al. 2003) and environmental adaptation through the modulation of the photosynthetic apparatus (Demmig-Adams and Adams 2002). Citrus fruit contains significant amounts of various carotenoids and more than 100 different kinds of carotenoids have been isolated in Citrus (Gross 1987). To date, many studies have described the isolation and characterization of the physiological and molecular features of carotenoid biosynthesis during fruit development and ripening in citrus (Kato et al. 2004, Rodrigo et al. 2004). The composition and content varied of carotenoids among citrus varieties rather than species, and the carotenoid diversity in cultivated citrus is highly influenced by genetic factors (Fancicullino et al. 2006). As a typical example, satsuma mandarin (Citrus unshiu Marc.) mainly accumulates β-cryptoxanthin (B-Cry) in the flavedo and juice sacs in mature fruit (Goodner et al. 2001, Ikoma et al. 2001), while sweet orange (Citrus sinensis Osbeck) fruit accumulates violaxanthins, predominantly 9-cis-violaxanthin (Lee and Castle 2001, Molnár and Szabolcs 1980). The carotenoid accumulation that occurs during citrus fruit ripening is highly regulated by the coordinated expression of carotenoid biosynthetic genes, and the differences in the variance of transcripts encoding biosynthetic enzymes is significantly associated with the carotenoid composition and content among varieties (Kato et al. 2004). In addition, several transcription factors, such as RAP2.2 (Welsch et al. 2007) and CubHLH1 (Endo et al. 2016), related to carotenoid metabolism were isolated and characterized in Arabidopsis thaliana and satsuma mandarin. However, the genetic information about how carotenoid composition and contents would extend the variation among citrus varieties at a molecular level has been quite limited.

In the Japanese citrus breeding program, the enrichment of carotenoids, especially B-Cry with health-promoting properties (Sugiura et al. 2011), is an important object.
with the aim of expanding citrus fruit consumption. To advance the molecular breeding for the enrichment of B-Cry, Sugiyama et al. (2011) examined quantitative trait locus (QTL) mapping for each carotenoid’s content level using the F1 hybrid population (AG population) between breeding lines ‘Okitsu 46 gou’ (A255) and ‘Kankitsu Chukanbohon Nou 5 gou’ (G434) with wide variation in carotenoid content and composition. A major QTL for B-Cry was detected on the Gm0005 locus in linkage group 6 of the G434-map and had a LOD value of 3.4. Various QTLs for the other carotenoid contents were also detected but most of them had low LOD values. This result was reasonable considering that carotenoid composition and content were affected by transcriptional amounts among various carotenoid biosynthetic genes. Recently, Shimada et al. (2014) constructed a citrus framework genetic map anchored by 708 gene-based markers using the same AG population. In this AG map, major carotenoid biosynthetic genes were mapped as follows: phytoene synthase (PSY), on linkage group (LG)-4, phytoene desaturase (PDS) on LG-3, ß-carotene desaturase (ZDS) on LG-9, ß-ring hydroxylase (HYb) on LG-3, zeaxanthin epoxidase (ZEP) on LG-2, and 9-cis-epoxycarotenoid dioxygenase (NCED) on LG-06. The other carotenoid biosynthetic genes could not be mapped owing to the low polymorphic features of genomic sequence and so on. Among the mapped loci of carotenoid biosynthetic genes, an eQTL contributing to ZEP expression level was located on ZEP locus of the genetic map for AG population (Sugiyama et al. 2014) and pairs of allelic ZEP gene with different sequences on ZEP locus were observed. In mandarin and orange, there are at least four ZEP alleles and the amounts and accumulation patterns of their transcripts were different during fruit development (Sugiyama et al. 2010). Therefore, the allelic variation among carotenoid biosynthetic genes would likely to extend wide variation in the carotenoid contents and composition among citrus varieties. In addition to ZEP expression, the eQTL analysis of PSY also suggested that the major regulation site for PSY expression might be tightly associated with PSY locus (Gm0009 locus) assigned by PsyG-CT marker (Sugiyama et al. 2014). PSY is a rate-limiting key enzyme in the carotenoid biosynthetic pathway and the flux of carotenoid pathway is generally controlled by multiple PSYs in plants. It is reported that Arabidopsis genome possesses one PSY gene while tomato and rice has two and three PSYs with different roles in plant development and tissues-specific expression or seasonal different expression (Bartley and ScoInik 1993, Ruiz-Sola et al. 2012, Welsch et al. 2008). In citrus, many cDNAs for PSY have been isolated and functionally characterized during fruit development (Kato et al. 2004, Maheswary et al. 2006, Rodrigo et al. 2004), however, little is known about how many loci or copies citrus PSYs comprise in the genome and how transcriptional variation of PSY is regulated to extend among citrus varieties.

In this study, we investigated the allelic diversity on genomic sequence and gene expression patterns of PSY homologues in clementine mandarin (C. clementine hort ex. Tanaka) genome, and inquired how transcriptional variation of PSYs in eQTL analysis was occurred in the AG population. The PSYs in the AG population comprised four alleles (PSY-a1 and PSY-a2 from A255, PSY-g1 and PSY-g2 from G434) which were derived from the parent lines, and F1 individuals with PSY-g2 tended to show low transcription level in the fruits. The sequence analysis was carried out to compare the promoter structures of four alleles and possible cis-regulatory elements to influence the transcription level were discussed.

Materials and Methods

Plant materials and preparation of DNA and RNA

All plants used in the experiments were cultivated in the research field of the National Agriculture and Food Research Organization Institute of Fruit Tree and Tea Science, Citrus Research Center, Okitsu, Shizuoka, Japan. Forty-eight F1 individuals bearing the fruits, which were obtained from crossing A255 and G434, were used for genetic analyses. The female parent of A255 was derived from ‘Sweet spring’ (‘Ueda unshiu’ (C. unshiu Marc.) × ‘Hassaku’ (C. hassaku hort ex. Tanaka)) × ‘Trovita’ orange, and the male parent of G434 was derived from ‘Lee’ (clementine mandarin × ‘Orlando’ tangelo (‘Duncan’ grapefruit (C. paradisi Macf.) × ‘Dancy’ tangerin (C. tangerina) × ‘Mukaku kishu’ (C. kinokuni hort. Tanaka). Genomic DNA was extracted from fresh and fully expanded leaves of F1 individuals and their parent varieties according to the method of Dellaporta et al. (1983). Total RNA was extracted from juice sacs on the middle of November (more than three fruits per individual) from the AG population using the method described by Ikoma et al. (1996).

In addition, various tissues were collected from clementine mandarin and immediately frozen by liquid nitrogen for and mRNA expression analysis, as follows: flower, leaf, stem, young whole fruit at DAF30, and juice sacs and peels at DAF60, DAF120 and DAF180. Total RNAs were also extracted from these samples.

Sequence comparison of PSYs among clementine mandarin, A255 and G434

Blastx search was carried out using a query nucleotide sequence of CitPSY (AF220201) against the clementine mandarin genome in the public databases (http://phytozome.jgi.doe.gov/pz/portal.html). Four nucleotide sequences show high homology against CitPSY and their nucleotide sequences were downloaded. Their deduced amino acid sequences were compared with those of PSY alleles detected in A255 and G434 and those of citrus PSYs published in the public DNA database. The alignment analysis and phylogenetic analysis were carried out using the computer program Genejex-Win Ver.11.0.3 (Software Development, Tokyo, Japan). Phylogenetic tree was constructed under the unweighted pair group maximum average method based on the amino acid sequences.
**Reverse transcriptase (RT)-PCR for various tissues of clementine mandarin**

The cDNA was synthesized from 1 μg of total RNA by Quantitect® Reverse Transcription (Qiagen, Hilden, Germany) using a poly-dT primer. Gene-specific primer sets for clementine mandarin PSY homologues were designed to prohibit the cross amplification between them. Primer sequences were described as follows: Ciclev10011841m.g: forward primer (5'-ATGTCTTGGTACATGGTATGGAAG-3') and reverse primer (5'-GGTTATCCAGATGAGCAAG-3'), Ciclev10015582m.g: forward primer (5'-TGCTCCTGCTTATATAGAGAAGG-3') and reverse primer (5'-TGCAGGTTTCTCTCAGGGATTTATGTTGTC3'). The primers used for Ciclev10018272m.g were not designed owing to the incomplete gene structure. Elongation factor 1 alpha (EF1-α) is used as an endogenous control gene and the primer sequences are referred to our previous report (Endo et al. 2006).

RT-PCR was conducted by the condition of 30 PCR cycles. Each cycle was composed of 94°C for 30 sec., 56°C for 1 min and 72°C for 1 min. The reaction mixture consisted of 100 mM Tris-HCl (pH 8.0), 50 mM of KCl, 1.5 mM of MgCl2, 0.2 mM of each primer, 5 ng of cDNA and 1 U of AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA) transformed into XL-1 Gold DNA strain. The thermal cycling conditions were 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 60 s.

**Sequence analysis for the promoter regions of four PSY alleles**

The 1.2 kbps of fragments for four PSY alleles were obtained by PCR amplification using AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA) protocols. The primers were designed based on the main database for the genome sequences of clementine mandarin and sweet orange (http://phytozome.jgi.doe.gov/pz/portal.html) as following: sense primer: 5'-TTTCCCAAGTTTGCACACGGTCAGTC-3' in the promoter region, antisense primer: 5'-AAAAAGTATTGAGAAAGGCA-3' in the primary exon. The amplified PCR products were cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA), transformed into Escherichia coli strain XL-1 Blue, and sequenced using a BigDye Terminator v. 3.1 Cycle Sequencing Kit (Applied Biosystems) with an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). These nucleotide sequences were registered in DDBJ (PSY-a1: LC169453, PSY-a2: LC169454, PSY-g1: LC169455, PSY-g2: LC169456). Genetyx-Win ver. 11.0.3 (Software Development) were used for the alignment analysis. The 1.2 kbps of promoter sequences were applied to a homology search using a plant cis-acting regulatory DNA element (PLACE, http://www.dna.aaffrc.go.jp/htdocs/PLACE/) database.


The following primer set was used to amplify the promoter region, antisense primer: 5'-AAAAAGTATTGAGAAAGGCA-3' in the promoter region, antisense primer: 5'-AAAAAGTATTGAGAAAGGCA-3'. The transcription level of PSY was also investigated using juice sac tissues on the middle of November by qRT-PCR. These analyses were carried out under above described methods.

**Results**

**Gene structures of PSY homologues on the clementine mandarin genome**

In the clementine mandarin genome sequence, there are four homologues with high homology with CitPSY as follows: Ciclev10018150m.g, Ciclev10018272m.g, Ciclev10015582m.g and Ciclev10011841m.g. Ciclev10011841m.g is annotated as chloroplastic phytoene
synthase and is located from 21,390,477 bps to 21,396,087 bps in the scaffold 6, which was equivalent to Gn0009 locus at 57.4 cM of linkage group 4 in AGI map (Table 1). Ciclev10018150m.g, Ciclev10018272m.g and Ciclev10015582m.g are annotated as squarene synthase, 15-cis-phytoene synthase, squarene/phytoene synthase, respectively and are located from 587,065 to 4,443,386 in the scaffold 2, which was equivalent to the position between Ks9005 and Lp0118 at 11.8–13.6 cM of linkage group 2 in AGI map. These three homologues were located tandemly in clementine mandarin genome and could be interpreted to derive from the same locus in the linkage map. The deduced amino acids from these homologues were aligned and their structures were compared (Fig. 1). PSY comprises two important functional domains of the chloroplastic transit peptide, which is essential for protein targeting to the plastid compartment (Giorio et al. 2008) and trans-isoprenyl diphosphate synthase (trans–IPP–HH) domain (Marchler-Bauer et al. 2009). In the deduced amino acids alignment of four clementine mandarin PSY homologues,

### Table 1. Sequence similarity and location among Clementine PSY homologues and PSY alleles of A255 and G434

<table>
<thead>
<tr>
<th>Gene annotation</th>
<th>Homology in amino acid sequence (%)</th>
<th>Location in Clementine genome sequence</th>
<th>Location in AGI map</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSYa-1</td>
<td>PSYa-2</td>
<td>PSYg-1</td>
</tr>
<tr>
<td>Ciclev10018150m.g</td>
<td>Squarene synthase</td>
<td>49.9</td>
<td>50.0</td>
</tr>
<tr>
<td>Ciclev10018272m.g</td>
<td>15-cis-phytoene synthase</td>
<td>34.1</td>
<td>34.2</td>
</tr>
<tr>
<td>Ciclev10015582m.g</td>
<td>Squarene/phytoene synthase</td>
<td>44.2</td>
<td>45.0</td>
</tr>
<tr>
<td>Ciclev1001841m.g</td>
<td>Chloroplastic phytoene synthase</td>
<td>84.7</td>
<td>85.1</td>
</tr>
</tbody>
</table>

* Incomplete coding region.

Fig. 1. Multiple alignment of deduced amino acid sequences of four clementine mandarin PSY homologues. Ciclev1001841m.g contains a predicted transient peptide at the N-terminal region (underlined in black) predicted by SignalP 4.1 Server. The putative active site (DXXXD) is boxed in black. The conserved amino acids are shown with black background and these conserved region among four PSY homologues is considered as trans-isoprenyl diphosphate synthase (trans–IPP–HH) domain (Marchler-Bauer et al. 2009). *: The ORF is properly re-predicted based on the structure of other citrus PSYs characterized by Sanger sequencing.
Allelic diversity influences transcription level of PSY in citrus fruit

Ciclev10011841m.g possessed two important functional domains, while other three homologues possessed trans–IPP–HH domain but lacked putative transit peptide predicted by SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/). In addition, Ciclev10018272m.g lacked N-terminal and C-terminal regions of trans–IPP–HH domain resulting to incomplete coding region. Among four clementine mandarin PSY homologues, their trans–IPP–HH domains were well conserved but other regions were divergent.

Phylogenetic analysis with the past reported citrus PSYs and PSYs in the AG population indicated that Ciclev10011841m.g clustered with them and revealed high homology with them (Fig. 2). In contrast, other three PSY homologues detected in clementine genome were not clustered with them.

Transcriptional changes of three PSYs in various tissues of clementine mandarin

To address the physiological role of clementine mandarin PSY homologues during the fruit development, RT-PCR was carried out using various tissues except for Ciclev10018272m.g owing to the incomplete gene structure (Fig. 3). The transcript of Ciclev10011841m.g was accumulated in all tissues of clementine mandarin and the transcription level was higher in flower, leaf and peel. The transcript of Ciclev10015582m.g was accumulated in flower, young whole fruit at DAF30 and peels at DAF60, DAF120 and DAF180. The transcript of Ciclev10018150m.g was not detected in all tissues even under the 35 PCR cycles and the transcription level would like to be extremely low or none. These results were agreed with the fact that Ciclev10011841m.g would play the important roles in carotenoid biosynthesis of the fruit.

Allelic genotype and gene expression analyses for PSYs in F1 individuals of AG population

In the AG population, there were four alleles with independent nucleotide sequences (PSY-a1 and PSY-a2 from A255, PSY-g1 and PSY-g2 from G434). Four PSY alleles in the AG population showed high identities in amino acid sequence ranging from 97.9% to 100% and the amino acid sequence of PSYa-1 and PSYG-1 was identical. The PSY sequence genotypes of the 48 F1 individuals of AG population bearing the fruits were determined using the TaqMan allele discrimination system. The predicted genotypes of the F1 individuals from the four parental alleles were PSY-a1/PSY-g1, PSY-a2/PSY-g1, PSY-a1/PSY-g2 and PSY-a2/PSY-g2. The numbers of F1 individuals with each genotype

Fig. 2. Phylogenetic tree of citrus phytoene synthases. The tree is constructed using the unweighted pair group maximum average (UPGMA) method based on their amino acid sequences. PSYs of satsuma mandarin (CitPSY), sweet orange (CsPSY) and pummelo (CmPSY) are used as reference clones for citrus PSY. The examined citrus PSYs are divided into two groups of “Chloroplastic phytoene synthases” and “Cytosolic squarene/phytoene synthases”. *: The ORF is properly re-predicted based on the structure of other citrus PSYs characterized by Sanger sequencing.
were 13, 12, 13 and 10, respectively. The segregation ratio of these genotypes fitted to the expected Mendelian proportions of 1:1:1:1 in the Chi-square test at the 0.05% level, indicating that these four alleles were derived from a single locus.

The transcription level of PSY in the juice sac tissues on middle of November was investigated for 48 F1 individuals using the TaqMan primers/probes reported by Kato et al. (2004) because quantification of the transcriptional amounts of each allele was difficult owing to the high similarity levels in the exon regions of the four alleles. The distribution of the transcription level in each genotype was described using box and whisker plots in Fig. 4. The transcription level of A255 with PSY-a1 and PSY-a2 was an average of 0.93 and that of G434 with PSY-g1 and PSY-g2 was an average of 0.22. The average expression level of F1 individuals with PSY-a2 and PSY-g1 was 0.57, followed by PSY-a1 and PSY-g1 individuals with an average of 0.49, PSY-a1 and PSY-g2 individuals with an average of 0.35, and PSY-a2 and PSY-g2 individuals with an average of 0.26. Thus, F1 individuals with PSY-g2 tended to have low transcription level, indicating that PSY allelic combination likely influences the transcription level of PSY.

**Promoter sequence of four PSY alleles in the AG population**

To understand the structural differences in the promoter regions of four PSY alleles, approximately 1.2 kbps of the sequences were compared among four alleles. Their sequences were relatively conserved with the high identities ranging from 99.8% to 93.2%. Among the four alleles, nucleotide substitutions and the indel mutations were observed. The frequency of indel mutation was less than that of nucleotide substitution. Most observed indels were 6–7 nucleotides in length. A homology search using a plant cis-acting regulatory DNA element (PLACE, http://www.dna.affrc.go.jp/htdocs/PLACE/) database was carried out against the 1.2 kbps of promoter regions. The representative cis-acting regulatory motifs related to carotenoid metabolism are summarized in Fig. 5. There are various types of cis-regulatory motifs corresponding to light, plant hormones and water stress including GATABOX, which is found in many light-regulated genes (Teakle et al. 2002), WRKY710S, which is a transcriptional repressor of the gibberellin signaling pathway (Zhang et al. 2004), and ABREATRD22 (Iwasaki et al. 1995), DPBFCOREDCDC3 (Kim et al. 1997) and ABREZMRAB28 (Guan et al. 2000), which are related to abscisic acid (ABA) and dehydration stresses. In addition, several MYB-binding motifs were found in the promoter regions, such as MYBPLANT found in phenylpropanoid biosynthetic genes of Antirrhinum majus (Sablowski et al. 1994), MYB1ST in Solanum tuberosum (Baranowski et al. 1994), and so on. Thus, various cis-regulatory motifs were scattered in their promoter regions. The organization of cis-regulatory motifs were different in a kind and a copy number among four alleles but the cis-regulatory motifs specific to PSY-g2 were very limited. The structural differences specific to PSY-g2 redundantly located in the promoter region from −59 bps to −350 bps. There were 15 single nucleotide substitutions and two indels in this region. Out of them, four mutations altered the cis-regulatory motifs. The single nucleotide substitution at approximately −340 bps generated a CPBCSPOR motif, which is a specific binding site of the cytokinin-dependent protein of the NADPH-protoclorophyllide oxidoreductase gene in cucumber (Cucumis sativus L.) (Fusada et al. 2005),
Allelic diversity influences transcription level of *PSY* in citrus fruit

![Diagram of the regions upstream of the *PSY* genes (−1260 bp to +1 from the translational start site). Nucleotide positions are shown relative to the putative translational start site. The box indicated the candidate alteration of *cis*-regulatory motif responsible for low expression of *PSY* in F1 individuals with *PSY*-2. The dot box indicated the other *cis*-regulatory motifs which were specifically altered in *PSY*-g-2.](Image)

while the single nucleotide substitution at approximately −310 bps generated a GTGANTGA motif that is responsible for the pectin lyase of pollen in *Nicotiana tabacum* (Rogers et al. 2001). The single nucleotide substitution at approximately −280 bps generated the SEF4MOTIFG7MS motif, which is responsible for the beta-conglycinin gene (Lessard et al. 1991). These *cis*-regulatory motifs would like to be not associated with low *PSY* expression because they generally work positively and some of them were located redundantly in the upstream regions of the other alleles. A 6-bps deletion in *PSY*-g2 altered the *cis*-regulatory motif from MYBPZM (Grotewold et al. 1994) to RAV1AAT (Kagaya et al. 1999) at −329 bps. MYBPZM is a core *cis*-regulatory motif for Maize P gene, and is responsible for red pigmentation of the kernel pericarp and flavonoid biosynthesis (Grotewold et al. 1994). The RAV1AAT motif is a *cis*-regulatory motif of the RAV1 protein, which is uniquely found in plants. It is involved in the immediate physiological responses and developmental adaptations to environmental stimuli (Kagaya and Hattori 2009).

**Genotype of MYBPZM and RAV1AAT motifs in the promoter region influences the transcription level of PSY**

To clarify whether the alteration of *cis*-regulatory motif from MYBPZM to RAV1AAT would be major factor to influence the transcription level of *PSY*, the genotype of these motifs in the promoter region of *PSY* was investigated for ten ancestral varieties and two parent lines of AG mapping population. Sequence analysis revealed that most ancestral varieties and G434 have heterozygous genotype of MYBPZM and RAV1AAT motifs in the promoter region, while ‘Mukaku kishu’, ‘Ueda unshu’ and A255 have homozygous genotype of MYBPZM motif (Fig. 6A). Except for ‘Orlando’, ‘Sweet spring’ and ‘Lee’, which of the fruits were not available, the transcription level of *PSY* gene in the juice sac tissues on the middle of November was investigated by qRT-PCR (Fig. 6B). Interestingly, the transcription level of *PSY* was higher in ‘Mukaku kishu’, ‘Ueda unshu’ and A255, which have homozygous genotype of MYBPZM motif in the promoter region of *PSY*. In contrast, the other varieties with heterozygous genotype of MYBPZM and RAV1AAT motifs showed lower transcription level of *PSY*. Thus, the positive relationship was observed between genotype of these motifs and the transcription level of *PSY* in the pedigree of AG mapping population, revealing 0.95 of Pearson’s correlation coefficient.

**Discussion**

*PSY* regulates the beginning step of the carotenoid metabolic pathway and its expression significantly affects the amount of subsequently produced carotenoids. In *Arabidopsis* genome, *PSY* is a single gene located on chromosome 5, although many *PSY* mRNA sequence entries have been registered. In tomato and rice, two and three *PSY* genes have been molecularly characterized and they played different roles and expression profiles during plant development and tissues-specific function or seasonal variation (Bartley and Scolnik 1993, Welsch et al. 2008). Based on the gene structure and expression pattern of *PSY* homologues in the clementine
mandarin genome sequence, there are two loci for chloro-
plastic phytoene synthase and cytosolic squarene/phytoene
synthase. The \( PSY \) locus in the scaffold 6, which is mapped by \( PsyG-CT \) marker on the linkage group 4 of the AGI map, is considered as a principal locus for carotenoid biosynthe-
sis in citrus. The alternative locus comprises three tandem
\( PSY \) homologues without putative transit peptide, which are members of cytosolic squarene/phytoene synthases, and
might not contributed to carotenoid biosynthesis. In addition,
it was confirmed that four \( PSY \)s in the AG population have high identities with the past reported citrus \( PSY \)s and they were derived from the single \( PSY \) locus in the scaffold 6 from the segregation analysis. Therefore, it is proposed that \( PSY \)s for carotenoid biosynthesis would be derived from a single locus with various sequence diversity working as multiallelism among citrus varieties.

The promoter structure was very similar among of the
examined four \( PSY \) alleles were but it was found that the
slight mutations could alter the organization of the \(cis\)-
regulatory motifs, resulting to influence gene expression
level of \( PSY \). In a precise comparison of promoter region,
we supposed that the motif substitution from \( MYBPZM \) to \( RA V1AAT \) might be associated with low expression level of \( PSY \) in AG progenies with \( PSY-g2 \). MYB family transcrip-
tion factors are involved in regulating the expression of fla-
vonoid and anthocyanin biosynthesis genes (Aharoni et al.
in \( Reduced Carotenoid Pigmentation 1 \) (\( RCP1 \)), an \( R2R3 \)
type of MYB transcription factor, lead to the downregula-
tion of all carotenoid biosynthesis genes and a reduced ca-
rotenoid content in \( Mimulus lewisii \) flowers (Sagawa et al.
2016). In contrast, \( RA V1 \) acts as a negative regulator of
\( ABA \) in seed germination and green seedling rates (Feng et
al. 2014). In pepper and cotton, \( RA V1 \) increases the toler-
ance to drought and salt stress (Li et al. 2015, Sohn et al.
2006). Based on these reports, these two \(cis\)-regulatory mo-
tifs are likely involved in the transcriptional regulation of
\( PSY \). In the plant kingdom, the antioxidant antagonism
between carotenoids and flavonoids is reported to avoid the
functional redundancy in the evolutionary mechanism (Han

---

Fig. 6. Pedigree of AG mapping population and their genotypes of MYBPZB and RA V1AAT in the promoter region of \( PSY \) alleles (A). M/R indicates heterozygous genotypes of these motifs in the promoter region and M/M indicates homozygous genotypes of MYBPZB motif in the promoter region. Transcription level of \( PSY \) in the juice sac tissues on the middle of November of seven ancestral varieties and two parent lines of AG mapping population (B). The transcription level of \( PSY \) in ‘Mukaku kishu’, ‘Ueda unshiu’ and A255 with M/M genotype, which are marked with an asterisk, show the significant differences at \( P < 0.05 \) against those of the other varieties and line with M/R genotype.
Allelic diversity influences transcription level of PSY in citrus fruit

et al. 2012). In Solanaceae plants, common transcription factors regulating carotenoid and anthocyanin pathways were found to act as initial sensors against various environmental stimuli (Dhar et al. 2014). Recently, MYBPZM is reported as one of the candidate cis-regulatory motif to influence gene expression of capsanthin-capsorubin synthase, which catalyzes the conversion of antheraxanthin and violaxanthin into capsanthin and capsorubin responsible for red-orange coloration in Pepper (Capsicum sp.) (Zheng et al. 2013). In citrus, the genetic diversity of ZEP alleles among various citrus varieties was characterized and similar results were obtained that the transcription level of ZEP-1m allele with MYBPZM was higher than that of ZEP-2m allele without it (Sugiyama et al. 2010). Therefore, MYBPZM could be one of the important cis-regulatory motifs to influence the carotenoid pathway in citrus fruit. A further analysis is required to obtain the direct demonstration whether the lack of MYBPZM in the promoter region of PSY-g2 is responsible for the low expression level of PSY by promoter assay.

Fanciullino et al. (2006) reported that the carotenoid composition and content varied among citrus varieties, rather than species, and the carotenoid diversity in cultivated citrus is highly influenced by genetic factors. Considering a single PSY locus in the citrus genome, the allelic combinations with different cis-regulatory motifs are one of the possible factors to cause a transcriptional variation of PSY. There have been several reports that allelic differences in PSY play critical roles in the modulation of carotenogenesis. In wheat grain (Triticum turgidum), the allelic divergence of PSY may be responsible for the grain’s yellow pigment content (Zhang and Dubcovsky 2008). In maize (Zea mays), sorghum (Sorghum bicolor) and rice (Oryza sativa), the three PSY genes have overlapping functions in modulating carotenogenesis in different tissues and in response to multiple developmental and/or stress signals (Gallagher et al. 2004, Li et al. 2008). Out of the two PSY alleles in maize, insertions in the Y1 phytoene synthase gene’s promoter increased the expression in endosperm and the carotenoid content of yellow maize (Palaisa et al. 2003). Thus, genetic variation within the cis-regulatory motifs could affect the transcription rate or tissue specificity of the associated allele and caused phenotypic differentiations through changes in gene expression. On one hand, transcription factors play essential roles controlling gene expression. Various endogenous and exogenous factors affecting carotenoid content and composition during fruit development have been characterized in citrus, such as plant hormones, temperature, light and nutritional factors (Alquézar et al. 2008), implicating that numerous transcription factors are involved in regulating carotenoid metabolism in response to environmental and endogenous factors. Therefore, the combination of these transcription factors derived from seven ancestral varieties would also promote the wide variation of the PSY transcription level in AG population.

In conclusion, the transcriptional diversity of PSY among citrus varieties is affected by PSY allelic combination derived from a single locus in the scaffold 6 of clementine genome sequence (linkage group 4 in the AGI map). The genomic sequences on the promoter region of the four PSY alleles were very similar at the nucleotide sequence level, but the alteration of cis-regulatory motifs would influence the gene expression level of PSY. Although the obtained new finding was limited on PSY locus in the AG population, it would be applied to the other carotenoid biosynthetic genes. Therefore, it is considered that the allelic diversity in carotenoid biosynthesis genes would be one of the explanations for wide variation in carotenoid composition and content among citrus breeding varieties. Allele mining of carotenoid biosynthetic genes is considered as a suitable approach to promote the molecular breeding for improvements of the carotenoids in citrus fruits through the genome editing and marker assisted selection. To promote the molecular breeding of a desirable carotenoid content in citrus, further research is required to understand the allelic diversity of all carotenoid biosynthesis genes among citrus breeding resources. Also, to characterize the transcription factors in various signaling networks related to carotenoid metabolism.

Acknowledgement

This work was partially supported by a grant from the Ministry of Agriculture, Forestry, and Fisheries of Japan (Genomics-based Technology for Agricultural Improvement, HOR-2003, DNA-marker breeding project).

Literature Cited


Allelic diversity influences transcription level of PSY in citrus fruit


