RESEARCH ARTICLE https://doi.org/10.7235/HORT.20200064

The Fusion Transcript of *Phytoene Synthase 1* Controls Yellow Fruit in Tomato

Eun Sol Kang and Je Min Lee*

Department of Horticultural Science, Kyungpook National University, Daegu 41566, Korea

*Corresponding author: jemin@knu.ac.kr

Abstract

Tomato (*Solanum lycopersicum*) fruit color is an important aspect of quality in terms of appearance and nutrient content. Carotenoids, chlorophyll, and flavonoids are the main components responsible for tomato fruit color. Phytoene synthase (PSY) is the key regulatory enzyme involved in the first committed step of carotenoid biosynthesis. Here, we found that the yellow fruit of *S. lycopersicum* 'YF2359', which lacks carotenoids, was regulated by a single recessive gene and cosegregated with *PSY1* in the F_2 population. Two different *PSY1* transcripts, including the wild-type and mutant (fusion-type), were discovered in YF2359. The mutant was generated by the fusion of exons from two different DNA strands, namely, *PSY1* and the antisense strand of *CoA ligase*, by *trans*-splicing. This *yellow flesh* phenotype, which was caused by a mutation of *PSY1* and was undetected in *S. lycopersicum* 'LA4044', was predicted to nullify PSY because of amino acid substitutions in the conserved *trans*-isoprenyl diphosphate synthase domain. Additionally, the *trans*-splicing of *PSY1* reduced wild-type *PSY1* expression, suggesting that the metabolic flux in the carotenoid pathway was suppressed in the presence of this mutation. Carotenoid biosynthetic genes were expressed significantly less in 'YF2359' than in the red-fruited wild-type tomato 'M82'. Information on this unusual mutation in *PSY1* will improve the understanding of genetic variations in plants and enable novel strategies for improving fruit quality.

Additional key words: candidate gene, carotenoid, fruit color, molecular breeding, *trans*-splicing

Received: April 20, 2020
Revised: July 30, 2020
Accepted: July 31, 2020

HORTICULTURAL SCIENCE and TECHNOLOGY
38(5):705-716, 2020
URL: http://www.hst-j.org

pISSN : 1226-8763
eISSN : 2465-8588

This research was supported by the National Research Foundation of Korea (2018R1A2B6002620); and the Vegetable Breeding Research Center, the Ministry of Agriculture, Food and Rural Affairs, Republic of Korea.

Introduction

Tomato (*Solanum lycopersicum*) is an important crop that provides various nutrients, including vitamins, minerals, and carotenoids (Guil-Guerrero and Rebolloso-Fuentes, 2009). It is also the most studied model system for fleshy fruit biology. In tomato, fruit color is a crucial quality trait that affects consumer preferences. During fruit ripening, the chlorophyll-containing chloroplasts are converted to chromoplasts, which gradually accumulate more carotenoids. Consequently, fruit color successively changes from green to yellow, orange, or red (Egea et al., 2010).

Carotenoids are isoprenoids with critical roles in photoprotection and light absorption (Enfissi et al., 2017). In addition, carotenoids are the precursors of various plant hormones, such as abscisic acid, and represent one of the largest classes of natural pigments, conferring the bright-red, orange, and yellow hues of fruits, flowers, and seeds (Dorais et al., 2008; Ahrazem et al., 2016). Carotenoids

provide various human health benefits, including prevention of chronic diseases, cancers, and visual impairment. Humans cannot synthesize carotenoids *de novo* and depend exclusively on the consumption of carotenoid-containing foods for these nutrients (Fraser and Bramley, 2004).

Carotenoids are synthesized by the methylerythritol 4-phosphate pathway. The isopentenyl diphosphate (Shinozaki et al., 2018) and dimethylallyl diphosphate produced in the methylerythritol 4-phosphate pathway undergo a condensation reaction to form the phytoene precursor geranylgeranyl diphosphate (GGPP). Two GGPP molecules are then condensed by phytoene synthase (PSY) to produce phytoene, which is further converted to phytofluene by phytoene desaturase. Subsequently, ζ-carotene isomerase and ζ-carotene desaturase consecutively synthesize *cis*-lycopene from phytofluene. Carotenoid isomerase then converts *cis*-lycopene to *trans*-lycopene. The cyclization of lycopene by lycopene cyclases is a critical step in directing the metabolic flux of the bifurcated carotenoid pathway. Both open ends of lycopene can be cyclized. Cyclization by lycopene ε-cyclase or β-cyclase (LCYB) generates ε- or β-ionone rings, respectively. The combination of β- and ε-ionone rings produces δ-carotene and α-carotene and its derivate lutein. Conversely, when two β-ionone rings are formed, β-carotene and its derivatives, for example, neoxanthin and zeaxanthin, are produced (Cunningham, 2002; Fraser and Bramley, 2004). γ-Carotene has one β-ring and one uncyclized end, and δ-carotene has one ε-ring and one uncyclized end. The accumulation of *trans*-lycopene increases during fruit ripening as a result of reduced lycopene cyclization and up-regulation of *PSY1* (Fraser et al., 2002).

Characterization of several mutants in tomato has contributed to the understanding of carotenoid biosynthesis and metabolism. For example, the *yellow flesh* (*r*) mutant has ripe yellow fruit because of a loss-of-function mutation in *PSY1*, leading to a lack of carotenoids during fruit ripening (Fray and Grierson, 1993). The *tangerine* (*t*) mutant has ripe orange fruit resulting from the accumulation of prolycopene instead of *trans*-lycopene because of the deletion of *carotenoid isomerase* (Isaacson et al., 2002). The *Beta* (*B*) mutant also has ripe orange fruit, but this phenotype is caused by a high accumulation of β-carotene because of *LCYB* over-expression (Ronen et al., 2000). The *Delta* mutant, which shows *lycopene-* ε*-cyclase* up-regulation, produces ripe orange fruit as a result of increased accumulation of δ-carotene and decreased accumulation of the substrate *trans*-lycopene (Ronen et al., 1999; Yoo et al., 2017). Additionally, mutation in genes that control the ripening process influence the fruit color. For example, the *ripening inhibitor* (*rin*) mutant produces yellow fruit that do not ripen because of the deletion of the intergenic region between *LeMADS-RIN* and *LeMADS-MC*, and part of *LeMADS-RIN* (Vrebalov et al., 2002). The *Colorless non-ripening* (*Cnr*) mutant also produces yellow fruit, but this phenotype results from the methylation of the promoter of the SBP-box (SQUAMOSA promoter binding protein-like) gene (Manning et al., 2006).

PSY is the most studied gene family in the carotenoid biosynthesis pathway. To date, three *PSY* homologs have been identified in tomato. Among them, *PSY1* is highly expressed during fruit ripening, whereas *PSY2* mainly functions in vegetative tissues. *PSY3* is reported to regulate carotenoid biosynthesis in the root (Kang et al., 2014; Liu et al., 2015). During fruit ripening, the transcript level and enzyme activity of PSY1 are increased (Ronen et al., 1999). In addition, PSY1 has a higher flux control coefficient, which is a quantitative indicator of the influence of an enzyme on carotenoid biosynthesis. Therefore, PSY1 is the critical enzyme involved in the first committed step of the carotenoid biosynthesis pathway in fruit (Fraser et al., 2002). Various mutations related to *PSY1* have been reported. For example, a transcriptional lesion resulting from a mutation of *Rider* insertion in *PSY1* results in the *r* mutant (Cheng et al., 2009). The r^y mutant results from a small truncation of the 3′ end of *PSY1* (Jiang et al., 2012). Sugar Yellow tomato has a single T/A substitution in the sixth exon of *PSY1* that results in a premature stop codon (Shin et al., 2019). Galina, lidi, and T1039

produce yellow-colored cherry tomato fruit because of a single G deletion in *PSY1*. This deletion results in an amino acid substitution from Lys_{389} to Ser, leading to a premature stop codon (Aflitos et al., 2014). Here, we identified two *PSY1* transcripts in the yellow-fruited tomato YF2359 by linkage and sequence analysis. Both *PSY1* transcripts were predicted to be generated by *trans*-splicing. Information on this unusual mutation in *PSY1* caused by *trans*-splicing will elucidate a wide spectrum of genetic variations in plants and can be used to improve the fruit quality of tomato.

Materials and Methods

Plant Materials

Solanum lycopersicum 'M82', *S. lycopersicum* 'E6203', and *S. pimpinellifolium* 'LA1589' show ripe red fruit. *Solanum lycopersicum* 'LA4044', known as *yellow flesh* (*r*), shows ripe yellow fruit. Only cultivar names are used hereafter. Seeds were obtained from the Tomato Genetic Resource Center (University of California, Davis, CA, USA). Seeds of *S. lycopersicum* 'YF2359' were obtained from Nongwoo Bio, Korea (Fig. 1A). YF2359 was crossed with LA1589 to produce F_2 populations. All the plants were grown in the greenhouse and field at Kyungpook National University, Daegu, Korea. The pericarps of ripe fruits were collected, immediately frozen in liquid nitrogen, and ground for extraction of carotenoids and total RNA.

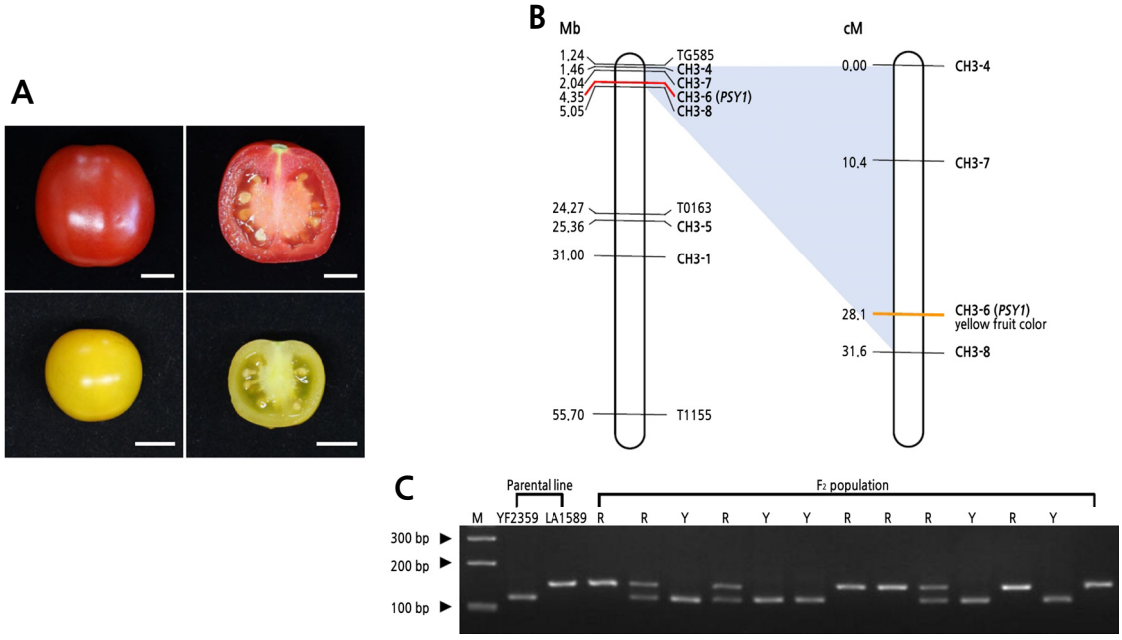


Fig. 1. Linkage analysis of yellow fruit color. (A) Ripe fruit of *Solanum lycopersicum* 'M82' (top) and YF2359 (bottom). Ripe fruit was harvested 10 days after the breaker stage. The scale bars indicate 1 cm. (B) Physical (left) and linkage maps (right) for genetic mapping of the yellow fruit color on chromosome 3. A total of 235 F_2 individuals (YF2359 × LA1589) were genotyped with five InDel markers and one dCAPS marker (CH3-6 marker). Two markers (CH3-7 and CH3-8) were described previously (Liu et al., 2017). The yellow fruit color was mapped to a 28.1-cM region with *PSY1*. (C) Co-segregation analysis of yellow fruit color and *PSY1*. The dCAPS marker was based on a substitution in the promoter region of *PSY1* in LA1589. The PCR products amplified by CH3-6 marker (*PSY1*) were digested with *Mnl*I and separated on 3% agarose gels. Numbers on the left indicate the DNA ladder size; "M" indicates the 1-kb DNA ladder; "R" and "Y" indicate red and yellow fruit color in the F_2 population, respectively.

cDNA Synthesis and Extraction of DNA and RNA

Genomic DNA was extracted from leaves using the cetyltrimethylammonium bromide method (Murray and Thompson, 1980). After total RNA was isolated using a Ribospin™ Seed/Fruit Kit (GeneAll Biotechnology Co., Ltd., Seoul, Korea), cDNA was synthesized using a DiaStar™ RT Kit (SolGent Co., Ltd., Daejeon, Korea).

Linkage and Marker Analysis

A total of 235 F_2 individuals (YF2359 × LA1589) were utilized to conduct a linkage analysis of the yellow fruit. The fruit color phenotype and genotype data of the F_2 population were used to construct the linkage map using MAPMAKER/EXP 3.0 (Lander et al., 1987). Polymerase chain reaction (PCR) was performed using a T100™ Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules, CA, USA) under the following conditions: 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, annealing temperature for 30 s, 72°C for 1 min, and 72°C for 5 min. The reaction mixture was composed of 18.875 μL of ddH_2O, 1 μL of DNA (200 ng), 0.5 μL of 10 mM dNTP mix, 2.5 μL of 10× e-*Taq* DNA reaction buffer, 1 μL of 10 pmol forward and reverse primer each, and 0.125 μL of Solg™ e-*Taq* DNA polymerase (SolGent Co., Ltd.).

Cloning and Sequence Analysis of *PSY1*

Isolation of total RNA from the ripe fruit pericarps of M82 and YF2359 and synthesis of cDNA were undertaken using commercial kits, as mentioned above. The *PSY1* promoter and coding region in M82 and YF2359 were cloned with two and one primer set, respectively (Suppl. Table 1s.). PCR amplification was conducted using Solg™ e-*Taq* DNA polymerase (SolGent Co., Ltd.) according to the manufacturer's instructions. Extension of the 3′-untranslated region (3'-UTR) in YF2359 was carried out using classic rapid amplification of cDNA ends, as previously described (Scotto‐Lavino et al., 2006). All amplified PCR products were cloned using a T-blunt™ PCR Cloning Kit (SolGent Co., Ltd.) and the *Escherichia* coli strain DH5 α. Plasmid DNA was isolated by an MG Plasmid DNA MiNi Kit (MGmed Co., Ltd., Seoul, Korea), and three clones per cultivar were sequenced at SolGent Co., Ltd. ClustalX 2.0 (https://clustalx.software.informer.com/2.1/) and GeneDoc (https://genedoc.software.informer.com/2.7/) programs were used for sequence alignment.

Carotenoid Extraction and High-Performance Liquid Chromatography (HPLC) Analysis

Approximately 0.1 g of frozen pericarp powder from ripe fruit was added to a 2-mL screw-cap tube containing two 6-mm glass beads for carotenoid extraction, as previously described (Yoo et al., 2017). Extracted carotenoids were analyzed using a 1260 Infinity HPLC (Agilent Technologies, Inc., Santa Clara, CA, USA) equipped with a YMC Carotenoid C_{30} S-5 column (4.6 × 250 mm) and Agilent ChemStation software (Agilent Technologies, Inc.). A two-solvent mobile phase system consisting of MeOH with 0.1% ammonium acetate (solvent A) and methyl *tert*-butyl ether (solvent B) was used, with detection at 286, 348, 434, and 450 nm. Each carotenoid was identified by the absorption maxima spectrum (Gupta et al., 2015).

Expression Analysis of the Carotenoid Biosynthesis Genes

Total RNA of ripe fruit and the corresponding cDNA were obtained for quantitative real-time (qRT)-PCR using commercial kits by following the manufacturers' instructions, as described above. qRT-PCR was performed with 100 ng of the cDNA and gene-specific primers (Suppl. Table 1s.) using the Power SYBR® Green PCR Master Mix (Applied Biosystems, Waltham, MA, USA) according to the manufacturer's instructions. At least three biological and two technical replicates were set up for analysis by qRT-PCR. Amplifications were conducted on a CFX Connect™ Real-Time PCR Detection System (Bio-Rad Laboratories, Inc.) under the following conditions: 95°C for 10 min, and 40 cycles of 95°C for 15 s and 60°C for 1 min, followed by a melting curve of 95°C for 10 s, 60°C for 5 s, 95°C for 5 s, and 22°C for 5 s. *ACTIN* was used as an internal control for normalization.

Results

Carotenoid Analysis of YF2359 by HPLC

The HPLC carotenoid profiles of M82 (red fruit), E6203 (red fruit), and YF2359 (yellow fruit) revealed that YF2359 had a different carotenoid profile from M82 (Fig. 1A). Lutein, β-carotene, γ-carotene, and *trans*-lycopene were detected in YF2359, whereas phytoene, phytofluene, and *cis*-lycopene were not. Although the β- and γ-carotene levels in YF2359 were only 2-fold lower than those in M82, the *trans*-lycopene and lutein levels were 120- and 6-fold lower, respectively. The carotenoid levels in YF2359 were also found to be lower than those in E6203 (Table 1) but similar to those in LA4044, known as *yellow flesh* (*r*) mutant, as described previously (Shin et al., 2019). Phytofluene and *cis*-lycopene were not detected in LA4044. The total carotenoid level in YF2359 and LA4044 was 6.61 ± 2.10 and 8.54 ± 3.12 $\mu g \cdot g^{-1}$ fresh weight, respectively. Therefore, YF2359 was predicted to have a *PSY1* mutation, leading to a lack of carotenoids in ripe fruit. The yellow fruit of the *rin* mutant is reported to have enhanced firmness (Vrebalov et al., 2002). To identify whether the YF2359 fruit had the *yellow flesh* phenotype because of a ripening defect, we compared YF2359 fruit firmness with that of M82 tomato fruit. However, there was no significant difference in the firmness between YF2359 and M82 (Suppl. Fig. 1s.). From the combined fruit firmness data and carotenoid profiles, YF2359 was predicted to have a defect of carotenoid accumulation rather than a defect of ripening. We hypothesized that there could be a *PSY1* sequence variation in YF2359 and so we analyzed the sequences of the *PSY1* promoter and coding regions in YF2359. No sequence variation was found (Suppl. Fig. 2s.); thus, we hypothesized that YF2359 might have an unusual mutation.

Table 1. Carotenoid contents ($\mu g \cdot g^{-1}$ fresh weight) of the ripe fruit of M82, YF2359, and E6203 quantified by HPLC analysis

Acession	Phytoene	Phytofluene	Lutein	β-carotene	γ-carotene	*cis*-lycopene	*trans*-lycopene	Total carotenoid
M82	14.56 ± 1.13[z]	10.55 ± 1.08	4.28 ± 0.38	9.62 ± 1.29	21.61 ± 1.80	3.86 ± 0.65	144.44 ± 11.73	219.47 ± 19.14
YF2359	ND[y]	ND	0.67 ± 0.14**	4.60 ± 0.94*	0.14 ± 0.14**	ND	1.20 ± 0.89**	6.61 ± 2.10**
E6203	33.81 ± 6.26**[x]	26.77 ± 4.76**	3.92 ± 0.79	7.99 ± 2.09	41.26 ± 4.53**	9.59 ± 2.11**	248.27 ± 21.91**	398.38 ± 47.21

[z]Mean ± standard error (n = 4).

[y]ND: not detected.

[x]Significant difference by Dunnett's test (* $p < 0.05$ and ** $p < 0.01$).

Linkage Analysis of the Yellow Fruit in YF2359

To identify the linkage between the yellow fruit and genes in YF2359, we constructed a linkage map for 235 F_2 individuals (YF2359 × LA1589). To construct the linkage map of the yellow fruit, 30 InDel and dCAPS markers that spanned the tomato genome were developed between *S. lycopersicum* and *S. pimpinellifolium*. Among them, two InDel markers (CH3-7 and CH3-8) have been described (Liu et al., 2017) (Suppl. Table 1s.). The yellow fruit of YF2359 was found to be closely linked with the CH3-8 marker on chromosome 3. However, the CH3-8 marker was linked with *PSY1* (Fig. 1B), so we developed a gene-specific dCAPS marker (CH3-6 marker) in *PSY1*. We constructed the additional linkage map using six markers, including the CH3-6 marker on chromosome 3. The yellow fruit was mapped to a 28.1-cM region and cosegregated with *PSY1* (Fig. 1B). Using the CH3-6 marker, we identified 69 and 110 homozygous and heterozygous red-fruited progeny, respectively, and 56 homozygous yellow-fruited progeny corresponding to the expected ratio of 1:2:1 (Fig. 1C and Table 2). In total, 179 and 56 progenies had red and ripe yellow fruit in the F_2 population, respectively. The segregation ratio of fruit color phenotypes in the F_2 population did not deviate from the expected 3:1 (red:yellow) ratio (Table 2). Thus, inconsistent with our initial hypothesis, the yellow fruit of YF2359 was cosegregated with *PSY1*, so we decided to analyze the *PSY1* sequence in YF2359 comprehensively.

Analysis of the *PSY1* Transcripts in YF2359

We sequenced the 3′-UTR of *PSY1* cDNAs from YF2359 fruit to assess whether they had a sequence variation. Two different *PSY1* transcripts were detected in the fruit of YF2359. One transcript was identical to the wild-type *PSY1* transcript in M82 and consisted of a 545-bp 5′-UTR, a 1239-bp coding region, and a 742-bp 3′-UTR. The other transcript was a fusion-type consisting of a 545-bp 5′-UTR, a 1299-bp coding region, and a 514-bp 3′-UTR (Fig. 2B). This transcript was produced by the fusion of exons from two different DNA strands, namely, the *PSY1* strand and the antisense strand of *CoA ligase*. The fusion transcript had a portion of the sixth exon (76 bp) and deletion of the entire 3′-UTR (742 bp), with a new 645-bp segment fused at the deletion site of the sixth exon (Fig. 2B). The new 645-bp segment was predicted to be produced from the antisense strand of *CoA ligase* and was complementary to the 5′-UTR (445 bp), first exon (177 bp), and second exon (23 bp) of *CoA ligase* (Fig. 2A). An additional five nucleotides (TATGC) were also inserted at the deletion site. However, their origin is unclear because of the lack of a corresponding sequence in the complementary strand (Fig. 2B). The coding region of the wild-type and fusion-type encoded 412 and 432 amino acids, respectively. We observed that 25 amino acids in the fusion-type were substituted by 45 amino acids at the C terminus. Two conserved lysine residues in the *trans*-isoprenyl diphosphate (*trans*-IPP-HH) domain of PSY were substituted by asparagine and methionine in the fusion-type (Fig. 3A). Given that these two lysine residues are phylogenetically conserved (Fig. 3B), they are presumably essential for the function of PSY1, so the fusion mutation probably nullifies PSY1 activity.

Table 2. Segregation of the *PSY1* genotype and ripe fruit color in 235 F2 individuals derived from a cross between YF2359 (*Solanum lycopersicum*) and LA1589 (*Solanum pimpinellifolium*)

	Expected ratio	Observed frequency	x^2-test	*p*-value
Phenotype	3:1	179 (red):56 (yellow)	0.17^{nsz}	0.9178
Genotype	1:2:1	69 (-/-):110 (-/*r*):56 (*r*/*r*)	2.40^{ns}	0.3018

[z]ns: not significant.

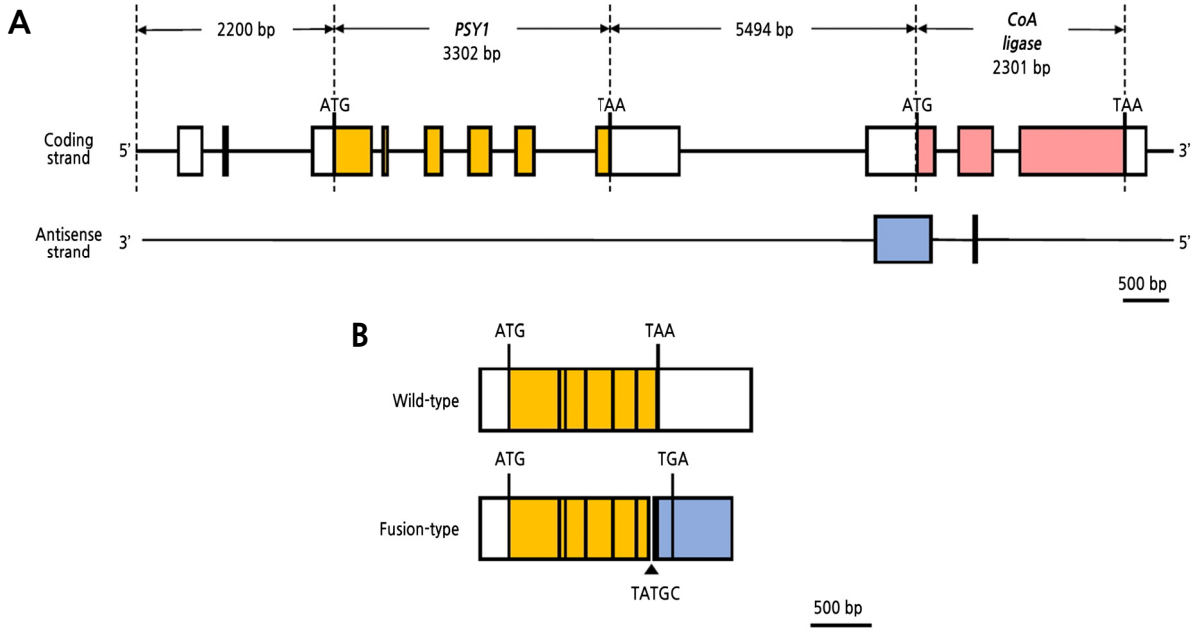


Fig. 2. Schematic representation of the genomic structure of *PSY1* in YF2359. (A) Schematic representation of the genomic structure of tomato *PSY1* and *CoA ligase* in chromosome 3. Exons are shown by orange, pink, and blue boxes, and introns are represented by solid lines. White boxes represent the 5′-UTR and 3′-UTR. *PSY1* consisted of 3302 nucleotides, including six exons and five introns. *CoA ligase* was composed of 2301 nucleotides, with three exons and two introns. The antisense nucleotides of *CoA ligase* were complementary to 445 bp of 5′-UTR, 177 bp of the first exon, and 23 bp of the second exon of *CoA ligase*. (B) Schematic representation of the full-length cDNA structure of the two *PSY1* transcripts in YF2359. The wild-type consisted of a 545-bp 5′-UTR, a 1239-bp coding region, and a 770-bp 3′-UTR. The fusion-type was composed of a 545-bp 5′-UTR, a 1299-bp coding region, and a 514-bp 3′-UTR. In the fusion-type, a portion of the sixth exon (76 bp) and the entire 3′-UTR (742 bp) were deleted, and 645 nucleotides from the antisense strand in *CoA ligase* were fused at the deletion site with an additional five nucleotides (TATGC). The stop codon was altered from TAA to TGA in the fusion-type.

Expression Analysis of the Carotenoid Biosynthesis Genes in YF2359

The expression level of the two *PSY1* transcripts, that is, the wild-type and fusion-type, was analyzed by qRT-PCR in red (M82 and E6203) and yellow fruit (LA4044 and YF2359), respectively. The expression level of the wild-type transcript in YF2359 was significantly lower (1000-fold) than that in M82. The level of the fusion-type transcript in YF2359 was comparable to that in wild-type M82, which did not express a detectable level of the fusion-type. In addition, the expression of the fusion-type was higher than the wild-type in YF2359 (Fig. 4A). These results indicate that the fusion mutation renders the wild-type *PSY1* expression in YF2359 much lower than that in M82. Furthermore, the fusion-type was not detected in E6203 and LA4044 (Fig. 4A) even though the LA4044 has been known as *yellow flesh* (*r*) with *PSY1* mutation (Tomato Genetic Resource Center, USA). These findings suggest that the fusion-type was a *PSY1* transcript detectable only in YF2359.

We also evaluated the expression level of the genes involved in the carotenoid biosynthesis pathway in M82 and YF2359. Except for *ß-carotene hydroxylase 2* (*CrtRb2*), the expression level of the carotenoid biosynthesis genes was significantly lower in YF2359 than in M82 (Fig. 4B). Therefore, the nonfunctional fusion mutant presumably renders a lower expression of wild-type *PSY1*, in turn, perturbing the carotenoid biosynthetic flux, resulting in ripe yellow fruit with low carotenoid contents in YF2359 (Fig. 1A and Table 1).

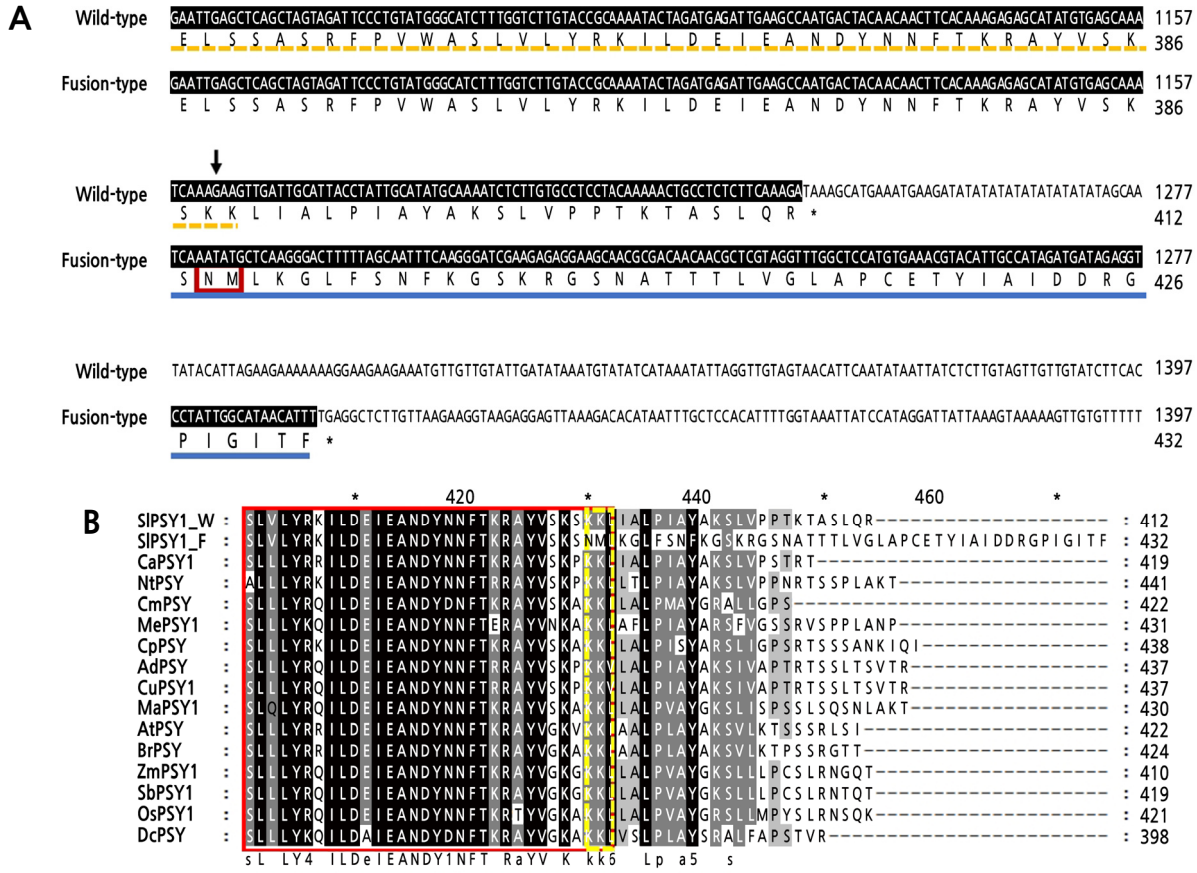


Fig. 3. Comparison of cDNA and protein sequences between the two *PSY1* transcripts. (A) Comparison of the 3′ end of the cDNAs and altered amino acid sequences between the two transcripts of *PSY1*, namely, wild-type and fusion-type, in YF2359. cDNA sequences were different from the arrow point. Different amino acids are *underlined* in *blue* in the fusion-type. The stop codons are shown in *red* font. The *trans*-IPP-HH domain is underlined by a *yellow dotted line* in the wild-type, and the substitution of amino acids in the *trans*-IPP-HH domain is *boxed* in *red* in the fusion-type. (B) Multiple sequence alignment of the *PSY1* protein in various species. Protein sequences were aligned using ClustalX 2.1. The *red box* indicates the part of the *trans*-isoprenyl diphosphate synthase (*trans*-IPP-HH) domain. The *yellow dashed line box* indicates the different sequences between fusion-type and various species in the *trans*-IPP-HH domain. *SlPSY1*, both wild-type (W) and fusion-type (F), was sequenced in this study. The *PSY1* sequences are as follows: *CaPSY1* (ACE78189.1) from *Capsicum annuum*; *NtPSY* (ADK25054.1) from *Nicotiana tabacum*; *CmPSY* (AEH03200.1) from *Cucumis melo*; *MePSY1* (ACY42664.1) from *Manihot esculenta*; *CpPSY* (ABG72805.1) from *Carica papaya*; *AdPSY* (ACO53104.1) from *Actinidia deliciosa*; *CuPSY1* (AF220218) from *Citrus unshiu*; *MaPSY1* (JX195664) from *Musa acuminata*; *AtPSY* (NP_001031895) from *Arabidopsis thaliana*; *ZmPSY1* (AAR08445) from *Zea mays*; *SbPSY1* (AY705389) from *Sorghum bicolor*; *OsPSY1* (AAS18307) from *Oryza sativa*; *DcPSY* (DQ192186) from *Daucus carota*.

Discussion

To determine the genetic factor underlying the yellow fruit of YF2359, we conducted a linkage analysis of this trait. The recessive trait was mapped to a 28.1-cM region covering *PSY1* on chromosome 3 (Fig. 1B and Table 2) and the yellow fruit trait cosegregated with *PSY1*.

The fruit color of tomato is associated with accumulation of carotenoids encoded by the *PSY1* gene during ripening (Lado et al., 2016). PSY1 is a crucial enzyme involved in the first committed step of the carotenoid biosynthetic flux to produce phytoene for synthesizing carotenoids in tomato (Liu et al., 2015; Sun et al., 2018) Overexpression of *PSY1* gene

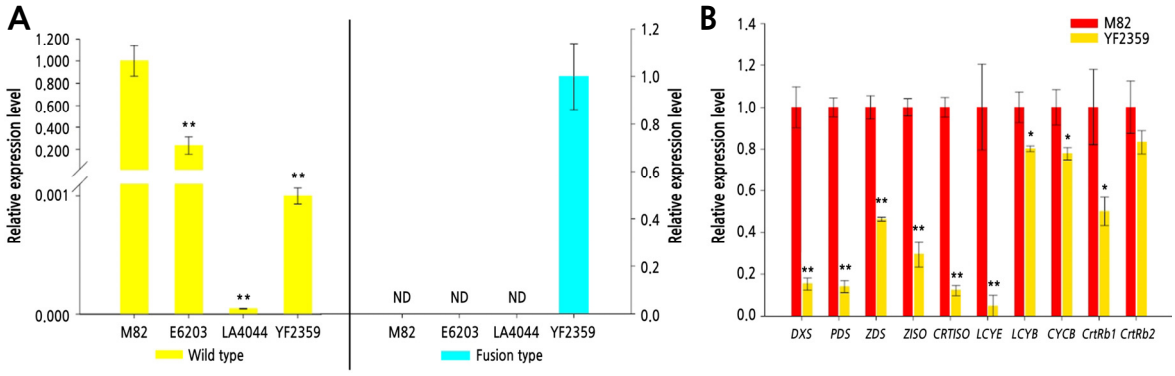


Fig. 4. Expression analysis of the carotenoid biosynthetic genes by qRT-PCR. Expression level of carotenoid biosynthetic genes was analyzed in ripe fruit (10 days after the breaker stage). (A) The relative expression levels of the two transcripts of *PSY1* in M82, E6203, LA4044, and YF2359. There was no detectable fusion transcript in M82, E6203, and LA4044. The expression level of wild-type *PSY1* transcript in YF2359 was significantly lower than that in M82, whereas the fusion-type expression in YF2359 exhibited a similar level as the wild-type expression level in M82. * and ** indicate significant difference at $p < 0.05$ and $p < 0.01$ by Dunnett's test ($n \geq 3$). (B) The relative expression levels of carotenoid biosynthetic genes in M82 and YF2359. *DXS, 1-deoxy-D-xylulose 5-phosphate synthase; CYCB, chromoplast-specific lycopene β-cyclase; CRTISO, carotenoid isomerase; CrtRb1, chloroplast-specific β-ring hydroxylase; CrtRb2, chromoplast-specific β-ring hydroxylase; LCY-B, chloroplast-specific lycopene β-cyclase; LCY-E, lycopene ε-cyclase; PDS, phytoene desaturase; PSY1, phytoene synthase; ZDS, ζ-carotene desaturase; ZISO, ζ-carotene isomerase. ACTIN* was used for normalizing the expression level. * and ** indicate significant difference at $p < 0.05$ and $p < 0.01$ by t-test ($n \geq 3$).

enhanced carotenoid content in fruit, whereas there is no carotenoid accumulation in the ripe fruit with knock-out of the *PSY1* gene in tomato (Gady et al., 2012). In YF2359, which has the *yellow flesh* phenotype, carotenoid accumulation was much less than in the red-fruited tomato M82 (Table 1). In addition, the carotenoid profile of YF2359 was similar to that of the *yellow flesh* (*r*) mutant (Shin et al., 2019). This result implies that YF2359 may have a defect in carotenoid accumulation, resulting in the yellow fruit.

The fruit-specific expression of *PSY1* was substantially increased during fruit ripening (Cao et al., 2019). A previous study also reported a concomitant increased transcriptional level of *PSY1* and carotenoid contents during fruit ripening (Pandurangaiah et al., 2016). Here, we identified that there were two *PSY1* transcripts, namely, wild-type and fusion-type, in YF2359 (Fig. 2B). The expression level of the fusion-type was higher than that of the wild-type in the ripe fruit of YF2359. In addition, although the expression of the fusion-type in YF2359 showed a similar level as the wild-type *PSY1* in M82, the wild-type *PSY1* level was much lower in YF2359 than in M82 (Fig. 4A). This result suggests that the fusion-type impaired the accumulation of the wild-type transcript because of a shared region of *PSY1* between the two transcripts (Fig. 2B). Therefore, the reduced wild-type transcript level may result in deficient production of phytoene synthase required for the normal level of carotenoid biosynthesis, leading to the ripe yellow fruit of YF2359 (Fig. 1A and Table 1).

Several mechanisms regulate the carotenoid biosynthesis pathway, such as the control of gene expression at the transcriptional level, post-transcriptional regulation, the substrate specificity of enzymes, and feedback regulation. In the feedback regulation, a lack of β-carotene causes increased activity of the up-stream enzymes in the carotenoid biosynthesis pathway (Diretto et al., 2020). Accordingly, a high level of lycopene has been detected in the *old-gold crimson* tomato mutant, which is deficient in β-carotene because of a mutation in *LCYB* (Bramley, 2002). Prolycopene

and neurosporene produced by ζ-carotene desaturase are predicted to affect *PSY1* transcription by feedback regulation. In t^{3002}/r^{2997} double mutant tomato, *PSY1* transcription is partially restored by prolycopene, so sufficient amounts of phytoene and down-stream carotenoids are produced (Kachanovsky et al., 2012). We predicted that down-stream carotenoids of the carotenoid biosynthetic flux, such as β-carotene, may affect the *PSY1* transcription in YF2359 by negative feedback regulation. However, the underlying molecular mechanism is not well established and requires further research.

PSY is the major rate-limiting factor responsible for the pool size of carotenoids in plant (Sun et al., 2018). Highly conserved residues are important for PSY activity. For example, mutations of the key residues in PSY structures significantly decreased PSY activity (Cao et al., 2019). PSY1 contains a *trans*-IPP-HH synthase domain between the Asp_{122} and Lys_{389} residues and is predicted to affect catalytic activity and substrate recognition (Supathaweewat and Klanrit, 2013). We identified that the *PSY1* fusion mutant had amino acid substitutions in the *trans*-IPP-HH domain (Fig. 3A). Two conserved lysine residues were substituted for asparagine and methionine at the C-terminal end of the fusion-type transcript (Fig. 3B). Given their phylogenetic conservation, these two amino acids may be crucial for functional PSY1, which could explain the fusion mutation that may nullify the PSY1 function in YF2359. These data indicate that in addition to the PSY1 catalytic activity, the fusion-type may not recognize GGPP as the substrate required for the production of phytoene. Therefore, the reduced phytoene synthase level hinders the carotenoid biosynthesis and results in yellow fruit.

Through sequencing analysis, the wild-type was determined to be identical to the *PSY1* transcript of M82, whereas the fusion-type consisted of *PSY1* and the antisense strand of *CoA ligase*, suggesting the post-transcriptional regulation of *PSY1* in YF2359 (Fig. 2).The sequence difference between the two transcripts was expected due to *trans*-splicing in YF2359. Genic *trans*-splicing is a type of *trans*-splicing that results from the combination of parts of two different mRNA transcripts joined at the splice site (Lasda and Blumenthal, 2011). This type of *trans*-splicing involves exons produced from different pre-mRNAs of the same gene or transcripts of different genes or intergenic regions. Such *trans*-splicing events have been found in the *modifier of mdg4* of *Drosophila* (Labrador et al., 2001). However, *trans*-splicing in plants has not been well studied. Here, we identified that the fusion-type transcript was generated by the joining of exons from two different strands,*PSY1* and the antisense strand of *CoA ligase* (Fig. 2). Therefore, the fusion-type transcript in YF2359 appears to be produced by genic *trans*-splicing. Furthermore, the *trans*-splicing in *PSY1* led to reduced expression of wild-type *PSY1*, implicating perturbation of the metabolic flux in the carotenoid pathway and resulted in the *yellow flesh* phenotype of YF2359. The results of studying this unusual mutation in *PSY1* by *trans*-splicing that alters the fruit color of tomato will be helpful to understand other genetic variations in plants.

Literature Cited

Aflitos S, Schijlen E, de Jong H, de Ridder D, Smit S, Finkers R, Wang J, Zhang G, Li N, et al. (2014) Exploring genetic variation in the tomato (*Solanum section Lycopersicon*) clade by whole-genome sequencing. Plant J 80:136-148. doi:10.1111/tpj.12616

Ahrazem O, Gómez-Gómez L, Rodrigo MJ, Avalos J, Limón MC (2016) Carotenoid Cleavage Oxygenases from Microbes and Photosynthetic Organisms: Features and Functions. Int J Mol Sci 17:1781. doi:10.3390/ijms17111781

Bramley PM (2002) Regulation of carotenoid formation during tomato fruit ripening and development. J Exp Bot 53:2107-2113. doi:10.1093/jxb/erf059

Cao H, Luo H, Yuan H, Eissa MA, Thannhauser TW, Welsch R, Hao Y-J, Cheng L, Li L (2019) A Neighboring Aromatic-Aromatic Amino

Acid Combination Governs Activity Divergence between Tomato Phytoene Synthases. Plant Physiol 180:1988-2003. doi:10.1104/pp.19.00384

Cheng X, Zhang D, Cheng Z, Keller B, Ling HQ (2009) A new family of Ty1-copia-like retrotransposons originated in the tomato genome by a recent horizontal transfer event. Genetics 181:1183-1193. doi:10.1534/genetics.108.099150

Cunningham FX (2002) Regulation of carotenoid synthesis and accumulation in plants. Pure Appl Chem 74:1409-1417. doi:10.1351/pac200274081409

Diretto G, Frusciante S, Fabbri C, Schauer N, Busta L, Wang Z, Matas AJ, Fiore A, K.C. Rose J, et al. (2020) Manipulation of β-carotene levels in tomato fruits results in increased ABA content and extended shelf life. Plant Biotechnol J 18:1185-1199. doi:10.1111/pbi.13283

Dorais M, Ehret DL, Papadopoulos AP (2008) Tomato (*Solanum lycopersicum*) health components: from the seed to the consumer. Phytochem Rev 7:231-250. doi:10.1007/s11101-007-9085-x

Egea I, Barsan C, Bian W, Purgatto E, Latché A, Chervin C, Bouzayen M, Pech JC (2010) Chromoplast differentiation: current status and perspectives. Plant Cell Physiol 51:1601-1611. doi:10.1093/pcp/pcq136

Enfissi EM, Nogueira M, Bramley PM, Fraser PD (2017) The regulation of carotenoid formation in tomato fruit. Plant J 89:774-788. doi:10.1111/tpj.13428

Fraser PD, Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. Prog Lipid Res 43:228-265. doi:10.1016/j.plipres.2003.10.002

Fraser PD, Romer S, Shipton CA, Mills PB, Kiano JW, Misawa N, Drake RG, Schuch W, Bramley PM (2002) Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. Proc Natl Acad Sci U S A 99:1092-1097. doi:10.1073/pnas.241374598

Fray RG, Grierson D (1993) Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression. Plant Mol Biol 22:589-602. doi:10.1007/BF00047400

Gady AL, Vriezen WH, Van de Wal MH, Huang P, Bovy AG, Visser RG, Bachem CW (2012) Induced point mutations in the phytoene synthase 1 gene cause differences in carotenoid content during tomato fruit ripening. Mol Breed 29:801-812. doi:10.1007/s11032-011-9591-9

Guil-Guerrero JL, Rebolloso-Fuentes MM (2009) Nutrient composition and antioxidant activity of eight tomato (*Lycopersicon esculentum*) varieties. J Food Compost Anal 22:123-129. doi:10.1016/j.jfca.2008.10.012

Gupta P, Sreelakshmi Y, Sharma R (2015) A rapid and sensitive method for determination of carotenoids in plant tissues by high performance liquid chromatography. Plant methods 11:5. doi:10.1186/s13007-015-0051-0

Isaacson T, Ronen G, Zamir D, Hirschberg J (2002) Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of β-carotene and xanthophylls in plants. Plant Cell 14:333-342. doi:10.1105/tpc.010303

Jiang N, Visa S, Wu S, van der Knaap E (2012) *Rider* Transposon Insertion and Phenotypic Change in Tomato. *In* M-A Grandbastien, JM Casacuberta, eds, Plant Transposable Elements. Springer-Verlag, Berlin, Heidelberg, pp 297-312. doi:10.1007/978-3-642-31842-9_15

Kachanovsky DE, Filler S, Isaacson T, Hirschberg J (2012) Epistasis in tomato color mutations involves regulation of *phytoene synthase 1* expression by *cis*-carotenoids. Proc Natl Acad Sci USA 109:19021-19026. doi:10.1073/pnas.1214808109

Kang B, Gu Q, Tian P, Xiao L, Cao H, Yang W (2014) A chimeric transcript containing *Psy1* and a potential mRNA is associated with *yellow flesh* color in tomato accession PI 114490. Planta 240:1011-1021. doi:10.1007/s00425-014-2052-z

Labrador M, Mongelard F, Plata-Rengifo P, Baxter EM, Corces VG, Gerasimova TI (2001) Protein encoding by both DNA strands. Nature 409:1000. doi:10.1038/35059000

Lado J, Zacarías L, Rodrigo MJ (2016) Regulation of Carotenoid Biosynthesis During Fruit Development. *In* C Stange, ed, Carotenoids in Nature: Biosynthesis, Regulation and Function. Springer International Publishing, Subcell Biochem., pp 161-198. doi:10.1007/978-3-319-39126-7_6

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174-181. doi:10.1016/0888-7543(87)90010-3

Lasda EL, Blumenthal T (2011) *Trans*-splicing. Wiley Interdiscip Rev RNA 2:417-434. doi:10.1002/wrna.71

Liu L, Shao Z, Zhang M, Wang Q (2015) Regulation of carotenoid metabolism in tomato. Mol Plant 8:28-39. doi:10.1016/j.molp.2014.11.006

Liu X, Geng X, Zhang H, Shen H, Yang W (2017) Association and Genetic Identification of Loci for Four Fruit Traits in Tomato Using InDel Markers. Front Plant Sci 8:1269. doi:10.3389/fpls.2017.01269

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB (2006) A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat Genet 38:948-952. doi:10.1038/ng1841

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321-4325. doi:10.1093/nar/8.19.4321

Pandurangaiah S, Ravishankar KV, Shivashankar KS, Sadashiva AT, Pillakenchappa K, Narayanan SK (2016) Differential expression of carotenoid biosynthetic pathway genes in two contrasting tomato genotypes for lycopene content. J Biosci 41:257-264. doi:10.1007/s12038-016-9602-4

Ronen G, Carmel-Goren L, Zamir D, Hirschberg J (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 97:11102-11107. doi:10.1073/pnas.190177497

Ronen G, Cohen M, Zamir D, Hirschberg J (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. doi:10.1046/j.1365-313x.1999.00381.x

Scotto-Lavino E, Du G, Frohman MA (2006) 3'End cDNA amplification using classic RACE. Nat Protoc 1:2742-2745. doi:10.1038/nprot.2006.481

Shin JH, Yoo HJ, Yeam I, Lee JM (2019) Distinguishing two genetic factors that control yellow fruit color in tomato. Hortic Environ Biotechnol 60:59-67. doi:10.1007/s13580-018-0093-0

Shinozaki Y, Nicolas P, Fernandez-Pozo N, Ma Q, Evanich DJ, Shi Y, Xu Y, Zheng Y, Snyder SI, et al. (2018) High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening. Nature Communications 9:364. doi:10.1038/s41467-017-02782-9

Sun T, Yuan H, Cao H, Yazdani M, Tadmor Y, Li L (2018) Carotenoid Metabolism in Plants: The Role of Plastids. Mol Plant 11:58-74. doi:10.1016/j.molp.2017.09.010

Supathaweewat K, Klanrit P (2013) cDNA cloning and expression analyses of phytoene synthase 1, phytoene desaturase and ζ-carotene desaturase genes from *Solanum lycopersicum* KKU-T34003. Songklanakarin J Sci Technol 35:517-527

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni J (2002) A MADS-box gene necessary for fruit ripening at the tomato *ripening-inhibitor* (*rin*) locus. Science 296:343-346. doi:10.1126/science.1068181

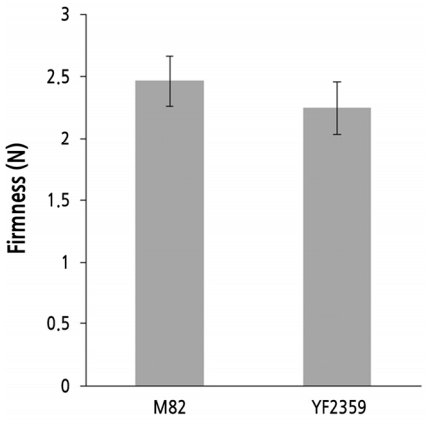
Yoo HJ, Park WJ, Lee GM, Oh CS, Yeam I, Won DC, Kim CK, Lee JM (2017) Inferring the Genetic Determinants of Fruit Colors in Tomato by Carotenoid Profiling. Molecules 22:764. doi:10.3390/molecules22050764

Supplementary Table 1s. Primers used in this study

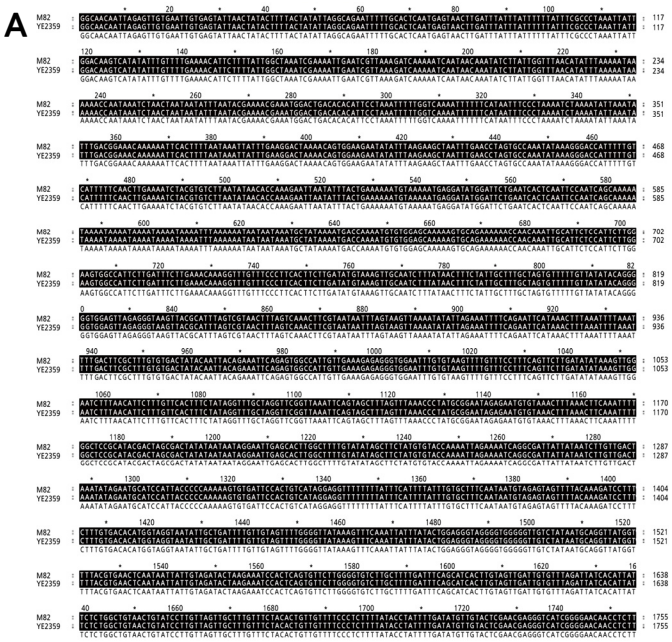
Primer ID	Primer sequence (5'-3')	Purpose
PSY1-promoter-1F	GGCAACAATTAGAGTTGTGA	Sequence analysis of *PSY1* promoter
PSY1-promoter-1R	AAAACCTCCTATGACAGTGGAA	
PSY1-promoter-2F	TCAAATTTTGGCTCCGCATA	
PSY1-promoter-2R	TGATTCCATGAAACTTGTCCCA	
SlPSY1F	ATGTCTGTTGCCTTGTTATGGGTT	Sequence analysis of *PSY1* coding region
SlPSY1R	TCATGCTTTATCTTTGAAGAGAGG	
Qtotal[z]	CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGCTTTTTTTTTTTTTTTT	3'RACE for 3'-UTR of *PSY1*
Qouter[z]	CCAGTGAGCAGAGTGACG	
Qinner[z]	GAGGACTCGAGCTCAAGC	
GSP(PSY1)1	CCAGAAGAGGAAGAGTCTAC	
GSP(PSY1)2	CAGAATTGAGCTCAGCTAGTAGAT	
CH3-1F	CCTGT TGCAATATTC CACTTATCA	Linkage analysis
CH3-1R	ATCTAGTTTCCCTAGTTCTAGG	
CH3-4F	CTCTTTTTGT TTCCTTGACT CGT	
CH3-4R	GAAATCTCTGCTTATAATCATACAC	
CH3-5F	TTAGAAACCCCCTGCTGTTG	
CH3-5R	ACCACGTCCACTGACAATGA	
CH3-6F	GGCAACAATTAGAGTTGTGAATTG	
CH3-6R	TTCAAAACAAATATATGACTTGTCCAATAATTGAG	
CH3-7F[y]	AAACTAGCAGGAGGATTCCCA	
CH3-7R[y]	GGCATCATCTTGGTTTGGAT	
CH3-8F[y]	TCATTGTGAGGATTGTGCAAC	
CH3-8R[y]	CGTCCTGCATACAAAGGACAT	
qRT-DXS-F	GCGGAGCTATTTCACATGGT	qRT-PCR of the carotenoid biosynthesis genes
qRT-DXS-R	CTGCTGAGCATCCCAAT	
qRT-PSY1-F	TGGCCCAAACGCATCATATA	
qRT-PSY1-R	CACCATCGAGCATGTCAAATG	
qRT-PDS-F	GTGCATTTTGATCATCGCATTGAAC	
qRT-PDS-R	GCAAAGTCTCTCAGGATTACC	
qRT-ZDS-F	TTGGAGCGTTCGAGGCAAT	
qRT-ZDS-R	AGAAATCTGCATCTGGCGTATAGA	
qRT-ZISO-F	TATGAGGATTACCAGGCATCC	
qRT-ZISO-R	ACAGCGTGTGAGCTAAGCAC	
qRT-CRTISO-F	TTTTGGCGGAATCAACTACC	
qRT-CRTISO-R	GAAAGCTTCACTCCCACAGC	
qRT-LCYE-F	GGATCAGCAGTCTAAGCTTTCTG	
qRT-LCYE-R	CGATTACCACTAAATCCAGTACG	
qRT-LCYB-F	TCGTTGGAATCGGTGGTACAG	
qRT-LCYB-R	AGCTAGTGTCCTTGCCACCAT	
qRT-CYCB-F	TGTTATTGAGGAAGAGAAATGTGTGAT	
qRT-CYCB-R	TCCCACCAATAGCCATAACATTTT	
qRT-CrtRb1-F	CACTTCTTATCAATGAGGAAGGG	
qRT-CrtRb1-R	CTTACTGAATAAGAAGAGCCTAC	
qRT-CrtRb2-F	GCGATTGTATTTAGGCTGTAGA	
qRT-CrtRb2-R	GGTTTGTCAACAAGTACTAGGAC	
qRT-ACTIN-F	CCTCAGCACATTCCAGCAG	qRT-PCR for ACTIN
qRT-ACTIN-R	CCACCAAACTTCTCCATCCC	

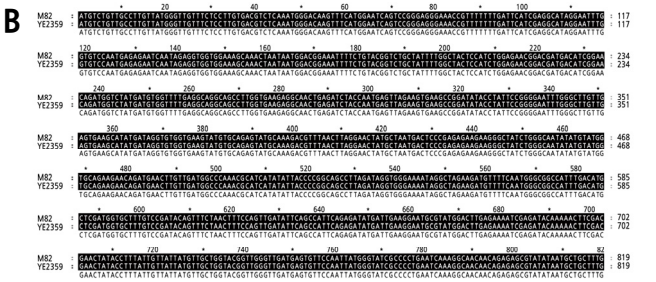
[z]described previously (Scotto－Lavino et al., 2006).

[y]described previously (Liu et al., 2017).



Supplementary Fig. 1s. Fruit firmness of M82 and YF2359 in ripe fruits. The firmness was measured using a SUN RHEO METER COMPAC-100 II (Sun Scientific Co., Ltd.) fitted with a 3 mm plunger. Each fruit with eliminated peel was compressed to 10 mm on the equatorial zone. The data present the mean with standard error (n = 14).





Supplementary Fig. 2s. Alignment for promoter and coding regions of *PSY1* between M82 and YF2359. (A) Comparison of the promoter region sequences between the Heinz1706 and YF2359. (B) Comparison of the coding region sequences between the M82 and YF2359.