Plant Biotechnology 32, 193–203 (2015)
DOI: 10.5511/plantbiotechnology.15.0526a

Original Paper

Biochemical and molecular characterization of orange- and tangerine-colored rice calli

Atsushi Ishihara[1], Kazuhiko Ohishi[2], Tetsuya Yamada[3], Mari Shibata-Hatta[4], Yuko Arai-Kichise[4], Satoru Watanabe[2], Hirofumi Yoshikawa[2,4], Kyo Wakasa[2,4,*]

[1] Faculty of Agriculture, Tottori University, Tottori, Tottori 680-8553, Japan; [2] Department of Bioscience, Tokyo University of Agriculture, Tokyo 156-8502, Japan; [3] Graduate School of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan; [4] Genome Research Center, NODAI Research Institute, Tokyo University of Agriculture, Tokyo 156-8502, Japan
* E-mail: k3wakasa@nodai.ac.jp Tel: +81-3-5477-2769 Fax: +81-3-5477-2377

Received May 8, 2015; accepted May 26, 2015 (Edited by Y. Hoshino)

Abstract We established an orange (L) and a tangerine (D) colored callus line by repeatedly subculturing rice callus line (N). The L line accumulated lycopene, β-carotene, and two lycopene isomers, while the D line accumulated much higher concentration of lycopene as well as a lycopene isomer than the L line. Genomic sequencing revealed the presence of SNP that results in the incomplete lycopene β-cyclase (β LCY) enzyme in the D line. We analyzed expression levels of carotenoid biosynthetic genes by qRT-PCR. Regarding the 2-*C*-methyl-D-erythritol-4-phosphate pathway genes, expression of the gene encoding geranylgeranyl diphosphate synthase 1 (GGPS1) was enhanced in both L and D lines as compared to the N line. Regarding the carotenoid pathway genes, the transcript amounts of the genes encoding phytoene synthase 2, phytoene desaturase, and ζ-carotene desaturase, were larger in the L line than in the N line; however, this pattern was not found in the D line. The expression levels of *β LCY* and lycopene ε-cyclase (*ε LCY*) gene in the D line and *ε LCY* in the L line were lower than in the N line. The transcript amounts of two carotenoid oxygenase genes were decreased both in L and D lines. We concluded that increased expression of *GGPS1* and reduced expression of *ε LCY*, *β LCY*, and carotenoids oxygenase genes might allow an excess accumulation of carotenoids in L and D lines, and that the mutation in *β LCY* is responsible for the conspicuous phenotype in the D line as compared to the L line.

Key words: Callus, β-carotene, carotenoid, lycopene, lycopene β-cyclase, rice.

Carotenoids are commonly present in most organisms living in aerobic conditions and exposed to sunlight. In plants, carotenoids play an important role in photosynthesis where they work as accessory pigments, harvesting the energy from the sun and transferring it to chlorophyll. These pigments also protect organisms from oxidative damage by reacting with reactive oxygen species generated during photoinhibition processes. In addition, carotenoids accumulate in specific organs such as fruits, roots, and flowers, providing them with characteristic colors. These colors are useful ecological signals, used and recognized by a large range of organisms.

Carotenoids produced by plants are transferred to animals through the food chain and metabolized into functional molecules. Their role as a precursor of vitamin A is particularly important for human health. Vitamin A deficiency can lead to preventable but severe visual impairment and blindness, and significantly increases the chances to contract and develop severe infections, including diarrhea and measles (Underwood and Arthur 1996). Vitamin A and its active metabolite, retinoic acid, are essential for the normal development and function of the hematopoietic and immune systems (Ross et al. 2011). Carotenoids are also important for their own physiological functions. For example, their antioxidant activity may prevent the oxidation of lipids (Di Mascio 1991). Accumulation of carotenoids in the corpus luteum and testicles at a high concentration also suggests its function during animal reproduction (Stahl et al. 1992).

Carotenoid biosynthetic pathway has been well studied in plants. They are tetraterpene molecules derived from the 2-*C*-methyl-D-erythritol-4-phosphate (MEP) pathway, which generates geranylgeranyl diphosphate (GGPP) (Phillips et al. 2008). Among the enzymes involved in this pathway, 1-deoxy-D-xylulose-5-phosphate synthase (DXS), 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR), and 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR) have been suggested to play regulatory roles (Botella-Pavia et al. 2004; Lois et al. 2000). GGPP is synthesized by GGPP synthase (GGPS) from a molecule of dimethylallyl diphosphate and three molecules of isopentenyl diphosphate. GGPP is funneled into the carotenoid

This article can be found at http://www.jspcmb.jp/
Published online July 18, 2015

pathway by phytoene synthase (PSY), which catalyzes the condensation of two molecules of GGPP into phytoene. The conversion of phytoene to lycopene is a four-step desaturation reaction catalyzed by phytoene desaturase (PDS), 15-*cis*-ζ-carotene isomerase (Z-ISO), ζ-carotene desaturase (ZDS), and carotene isomerase (CRTISO). The cyclization of lycopene ends catalyzed by lycopene β-cyclase (β LCY) and lycopene ε-cyclase (ε LCY) forms β-carotene (β, β-carotene) or α-carotene (β, ε-carotene). Further, xanthophylls are formed by hydroxylation of carotenes by enzymes such as carotene β-hydroxylases (β OHase), CYP97A, CYP97B and CYP97C. Carotenoids are further converted to apocarotenoids, which include representative phytohormones, abscisic acid (ABA) and strigolactones. In plants, apocarotenoids are biosynthesized by cleavage reactions of carotenoids mediated by dioxygenases such as 9-*cis*-epoxycarotenoid cleavage dioxygenases (NCEDs) and carotenoid cleavage dioxygenases (CCDs). An ectopic expression of CCDs has generated various cleaved molecules from carotenoids including lycopene and β-carotene (Auldridge et al. 2006).

Regulation of carotenoid biosynthesis is a complex process. The supply of substrates from the MEP pathway affects carotenoid biosynthesis. The expression of genes encoding DXS and HDR on the MEP pathway correlates with carotenoid accumulation (Carretero-Paulet et al. 2006; Lois et al. 2000), and the overexpression of these genes is known to increase carotenoid content (Carretero-Paulet et al. 2006; Estévez et al. 2001). PSY is generally considered to be an important regulatory point in the control of carotenoid biosynthesis (Cazzonelli and Pogson 2010). The expression of *PSY* is regulated by ABA, light intensity, salinity, drought, temperature, photoperiod, and developmental cues (Li et al. 2008a, 2008b; Welsch et al. 2008). The expression of *CRTISO* is also transcriptionally controlled. A chromatin-modifying histone methyltransferase enzyme, SDG8, has been shown to be required for the expression of *CRTISO* (Cazzonelli et al. 2009, 2010). β LCY and ε LCY activities control the ratio of the most abundant carotenoid, lutein, to other β-carotenoids. The *β LCY* gene from the eubacterium *Erwinia herbicola* and daffodil (*Narcissus pseudonarcissus* L.) flowers were introduced into the tomato (*Solanum lycopersicum* L.) plastid genome, and lycopene was channeled into the beta-branch, resulting in an increased accumulation of xanthophylls in leaves and predominantly β-carotene in fruits (Apel and Bock 2009). In the absence of *β LCY*, several unusual carotenes have been shown to accumulate in maize (*Zea mays* L.; Bai et al. 2009). The suppression of *ε LCY* expression results in an increase in carotenoid content and the ratio of β-carotene to lutein in *Brassica napus* seed (Yu et al. 2008). On the other hand, tuber-specific silencing of the *ε LCY* gene, by introducing an antisense fragment of the gene, increases β-carotene and zeaxanthin contents in potato tuber (*Solanum tuberosum* L.; Diretto et al. 2006). The capacity of a metabolic sink is also known to increase carotenoid content in plants. The transformation of *Or* gene, promoting the formation of a large membranous chromoplast, leads to an orange color, associated with increased levels of β-carotene, in cauliflower curds (*Brassica oleracea* L. var. *botrytis*) (Li et al. 2001; Lu et al. 2006).

In an effort to provide long-term and sustainable prevention of vitamin A deficiency, scientists have genetically engineered a variety of rice known as the Golden Rice (Ye et al. 2000). The first Golden Rice was generated by introducing three genes: phytoene synthase (*psy*) gene from daffodil, carotene desaturase (*crtI*) gene from the bacterium *Pantoea ananatis*, and lycopene β cyclase (*lcy*) from *Erwinia uredovora*. The genes *psy* and *lcy* were expressed under the control of an endosperm-specific rice glutelin promoter, while *crtI* was under the control of a CaMV *35S* promoter. The *aphIV* expressed under the control of CaMV *35S* promoter was used as selective marker. However, it was found that the introduction of *lcy* was not essential to increase the levels of β-carotene. To increase the level of β-carotene, the variety Golden Rice 2 was generated by introducing maize *psy* together with *Erwinia uredovora crtI* (Paine et al. 2005). The total carotenoid in Golden Rice 2 was up to 23 times higher (maximum 37 μg/g) than that of the first Golden Rice, with a preferential accumulation of β-carotene.

Wakasa et al. (1984) established in rice anther culture, the NR^{-}120-2 callus line (N), derived from pollen and deficient in nitrate reductase. After repeatedly subculturing the calli of this line, we established two callus lines showing orange (L) and tangerine (D) colors. These colors suggested that the pigments accumulating in L and D lines might be carotenoids. The calli derived from rice pollen and scutellum typically show white or faint yellow colors; however, to the best of our knowledge, orange rice calli have never been reported. Thus, the biosynthesis of pigments is preferentially suppressed in rice calli generated from these organs. Alternatively, carotenoids may not be accumulated in those calli because of fast flux of carotenoid metabolism. The emergence of calli with orange and tangerine colors may be attributable to the changes in the biosynthesis and/or catabolism of carotenoids through repeated subculturing. Elucidating the underlying mechanism may provide an insight into the regulation of carotenoid metabolism, and may promote the development of new technologies to generate carotenoid-rich rice varieties.

Here, we identified the pigments accumulated in these callus lines and performed genomic analysis by using next generation sequencing (NGS) techniques and gene expression analysis by quantitative RT-PCR,

to understand the mechanisms associated with the accumulation of these pigments.

Materials and methods

Callus lines

Three rice callus lines (*Oryza sativa* L. cultivar Norin 8) were used in this study. Callus line N is the line 120-2, deficient in nitrate reductase, established by anther culture by Wakasa et al. (1984). The other two callus lines (L and D) were derived from the callus line N. Line L is an orange-colored callus that was found after repeatedly subculturing line N and subsequently isolated. Line D, which shows a tangerine color, was isolated from L after repeated subcultures. All calli were maintained on AA medium (Müller and Grafe 1978) as previously described (Wakasa et al. 1984).

Protoplast isolation

Protoplasts were isolated from each callus line (N, L, and D) by treating the calli with Cellulase Onozuka RS (Wako Pure Chemical Industries, Ltd., Osaka, Japan) and Macerozyme R-10 (Yakult Pharmaceutical Co., Ltd., Tokyo, Japan), following a previously described protocol (Wakasa et al. 1984). The suspension obtained was passed through a layer of cell strainer (100 μm, BD Bioscience, San Jose, CA, USA) and inspected using an Olympus DP71 digital camera (Olympus, Tokyo, Japan) at 150×.

Extraction of pigments and measurement of absorption spectra

The calli (0.1 g) were frozen in liquid N_2 and homogenized using a multi-bead shocker (Yasui Kikai, Osaka, Japan) at 1800 rpm for 8 s, 28 days after subculturing. Acetone (1 ml) was added to the homogenate and mixed well. After centrifugation (13,000×g, 5 min, 4°C), the supernatant was collected. This procedure was repeated, and a second supernatant was combined with the first supernatant and dried up. The residue obtained was dissolved in 500 μl of acetone. This solution was diluted 20 times in acetone, and the absorption spectrum was measured using a Hitachi U-3310 spectrophotometer (Hitachi, Tokyo, Japan).

Chromatographic analyses of pigments

Callus extracts were analyzed by reversed phase TLC (TLC silica gel 60 RP-18 $F_{254}S$, Merck, Darmstadt, Germany) and LC-MS. TLC was carried out in an acetone-water (9 : 1, v/v). LC analyses were carried out with a Waters Accquity UPLC coupled with a Quattro Micro mass spectrometer (Waters, Milford, MA, USA). An Accquity UPC BEH C18 1.7-μm column (2.1×50 mm, Waters) was used for the LC-MS analyses. The solvents used were 0.1% formic acid in water (solvent A) and acetonitrile-acetone (1 : 1 ratio (v/v), solvent B). The analyses were carried out with a 90–99.9% B/(A+B) (0–5 min) and 99.9% B/(A+B) (5–10 min) gradient. For the LC-PDA analysis, THF-acetonitrile-methanol (7-58-35, v/v) was used as solvent.

DNA extraction

Genomic DNA was isolated from each callus following the CTAB method (Porebski et al. 1997) with minor modifications. Calli (approx. 30 mg) were homogenized using a multi-beads shocker in the presence of liquid N_2 28 days after subculture. The homogenate was suspended in 400 μl of CTAB buffer (2% CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl [pH 8.0], and 0.3% β-mercaptoethanol). The supernatant (200 μl) was subsequently extracted with 200 μl of chloroform and centrifuged for 5 min at 13,000×g. Each upper aqueous phase was transferred into a new tube, and the DNA was precipitated with 500 μl of isopropanol and corrected by immediately spinning the mixture at 13,000×g for 10 min. The DNA pellet was washed with 70% ethanol, dried, and re-suspended in 50 μl of TE solution containing 0.1% RNaseA (Takara, Ohtsu, Japan).

Genome sequence analysis

Genomic DNA samples, extracted from calli by using the CTAB method (Porebski et al. 1997), were used for the preparation of sequencing libraries, according to the manufacturer's protocol (Illumina, San Diego, CA, USA). The fragments of the libraries were paired-end sequenced using HiSeq2000 and HiSeq2500 (Illumina). The length of all sequences generated was 100 nucleotides. The NGS reads of these genomes have been submitted to DDBJ (http://www.ddbj.nig.ac.jp/index-e.html) and deposited under the accession number DRA002209.

Sequence mapping and SNP/InDels identification were done according to previously described methods (Arai-Kichise et al. 2014). These data are available at GBrowse, NGRC Orange-colored Rice (http://www.nodai-genome.org/oryza_sativa_en.html). Detected SNPs and InDels were annotated as gene structures in the Nipponbare genome. When SNPs and InDels were annotated in multiple gene-structures, we selected top three results in ascending order of gene ID.

For direct sequencing of β *LCY*, ε *LCY*, and *DXS3* around the SNPs called by the mapping programs, genomic regions were amplified by PCR using the primer sets (β LCY_f: 5′-CGA CGA GTT CGA CGC CAT GG-3′, β LCY_r: 5′-GGA CGC GTA TCC CGA GGT-3′, ε LCY_f: 5′-GGC TGG CTT TGA TAA TCC AAC TGA A-3′ and ε LCY_r: 5′-GCC AAA TTC ATG CAT AAA ATT CTT GCA GTG-3′, DXS3_f: 5′-GGA TGC ACG GCG TGG TC-3′ and DXS3_r: 5′-CCT CGA GAG GTG TGC CCT T-3′) and sequenced on a capillary sequencer with the Sanger method using the commercial sequence service of MACROGEN (Tokyo, Japan).

RNA extraction from callus lines

Each callus (100 mg) was frozen in liquid N_2 and ground into powder using a mortar and pestle 28 days after subculturing. Total RNA was extracted using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. After determination of the RNA concentration, 10 μg of RNA was treated with DNase (TURBO DNase, Toyobo,

Osaka, Japan). Total RNA was purified by an RNeasy MiniElute Cleanup Kit (Qiagen, Venlo, Netherlands).

First-strand cDNA was synthesized from 1 μg of purified RNA by reverse transcription using a SuperScript III First-Strand Synthesis System (Invitrogen) with oligo (dT) primers, according to the manufacturer's instructions. After reverse transcription reaction, RNase H was added to the reaction mixture and incubated at 37°C for 20 min, and the mixture was used as PCR template after 40 times dilution. Quantitative real-time PCR was performed using a StepOnePlus Real-Time PCR System (Applied Biosystems, Foster, CA, USA) with the specific set of primers (Supplemental Table 1). The reaction mixture (final volume 20 μl) contained 2 μl of cDNA solution, 0.4 μl of forward and reverse primer solutions (10 μM), 10 μl of Fast SYBR Green Master Mix (×2) (Applied Biosystems) and 7.2 μl of DNAase free water. The template cDNA was denatured at 95°C for 20 s, followed by 40 amplification cycles as follows: 95°C for 3 s, 60°C for 30 s, and 95°C for 15 s. The reaction mixture was kept at 60°C for 1 min and 95°C for 15 min for drawing melt curve. For the primer sets used, amplification rates were in the range of 80–120%. The relative abundance of transcripts was expressed by the ΔΔCT method.

Results

Identification of orange pigments

We established two callus lines presenting orange (L) and tangerine (D) colors (Figure 1). The L line was isolated from the mutant callus line 120-2 (N), deficient in nitrate reductase (Wakasa et al. 1984) after repeated subculturing. The D line was isolated from a callus that emerged from L following subsequent repeated subculturing.

To examine the intracellular localization of the orange and tangerine pigments, we prepared protoplasts from N, L, and D calli to be observed under a microscope (Figure 2A–2C). The accumulation of pigment was not detected microscopically in protoplasts from line N (Figure 2A). On the other hand, almost all protoplasts from line L contained intracellular orange granules and some protoplasts contained red crystalline needle-like structures (Figure 2B). The color of these pigments resembled those observed in red pepper (*Capsicum annuum* L.) and carrot (*Daucus carota* L.) protoplasts (Figure 2D, 2E). On the other hand, almost all protoplasts form the D line contained needle-like red crystals (Figure 2C), which were similar to the crystals observed in tomato protoplasts (Figure 2F). A small portion of the protoplasts from the D line contained orange granules.

We extracted pigments from N, L, and D calli with acetone and measured the absorption spectra of each extract (Figure 3). The extract from N calli showed a weak absorption, with λ_{max} at 450 nm. The extract from L showed a level of absorption higher than the N extract, with λ_{max} at 445, 470, and 524 nm. The extract from D also showed a strong absorption, with λ_{max} at 445, 470 and 524 nm, but the shape of the spectrum of the D extract was different from that of the L extract. In fact, the relative absorbance of peaks at 470 and 524 nm was larger in D extracts. These spectra were subsequently compared with the spectrum of extracts from tomato fruity flesh and carrot root, whose main pigments are lycopene and β-carotene, respectively. The spectrum of D extracts was almost identical to that of tomato extract, suggesting that the main pigment in the D line was lycopene. In addition, the spectrum of L extract showed the same λ_{max} as tomato extract, suggesting that the L line also contains large amounts of lycopene as a major pigment. Relatively smaller absorbance at 470 and 524 nm of L can probably be attributed to the presence of pigments with small λ_{max} values.

Thin layer chromatography (TLC) was used to identify the nature of these pigments (Figure 4). The TLC analysis showed that the main color spots, obtained from the L and D extracts, had Rf values of 0.17 identical with that of authentic lycopene (Figure 4A).

The extracts were also analyzed by liquid chromatography coupled with mass spectrometry (LC-MS) and liquid chromatography coupled with photodiode array detection (LC-PDA). The L extract showed four major peaks at 4.15, 4.20, 4.46, and 5.28 min (Figure 4B). The measured retention times of 4.15 and 5.28 min accurately matched those of purified lycopene and β-carotene, respectively. ESI-MS and absorption spectra of the 4.15 and 5.28 min peaks (Supplemental Figures 1C, 1F, 2C, and 2F) were also identical to those of lycopene and β-carotene, respectively (Supplemental Figures 1A, 1B, 2A, and 2B). ESI- MS of the 4.20 and 4.46 min peaks (Supplemental Figure 1D, 1E) were closely similar to those of lycopene and β-carotene, respectively, indicating that they are isomers of these carotenoids. The absorption spectra of the peaks at 4.20 and 4.46 min showed a λ_{max} at 468 nm and 440 nm, respectively (Supplemental Figure 2D, 2E).

The D extract showed two major peaks at 4.13 and 4.61 min (Figure 4B). The peak at 4.13 also showed an identical retention time, similar ESI-MS and absorption spectrum with lycopene (Supplemental Figures 1G and 2G). The ESI-MS of the peak at 4.61 min (Supplemental Figure 1H) was very similar to those of lycopene and β-carotene, suggesting that this compound is also an isomer of these two carotenoids. Its absorption spectrum showed three λ_{max} at 488, 457, and 432 nm (Supplemental Figure 2H).

The accumulated amounts of lycopene and β-carotene were determined by HPLC analysis with standard curves generated with authentic compounds. The amounts of unidentified carotenoids eluted at 4.13, 4.20, and 4.61 min were estimated on the basis of the

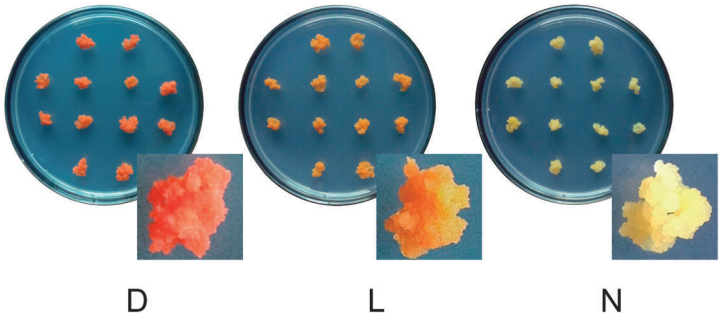


Figure 1. Callus lines D (left), L (middle), and N (right). Pictures were taken 28 days after subculturing.

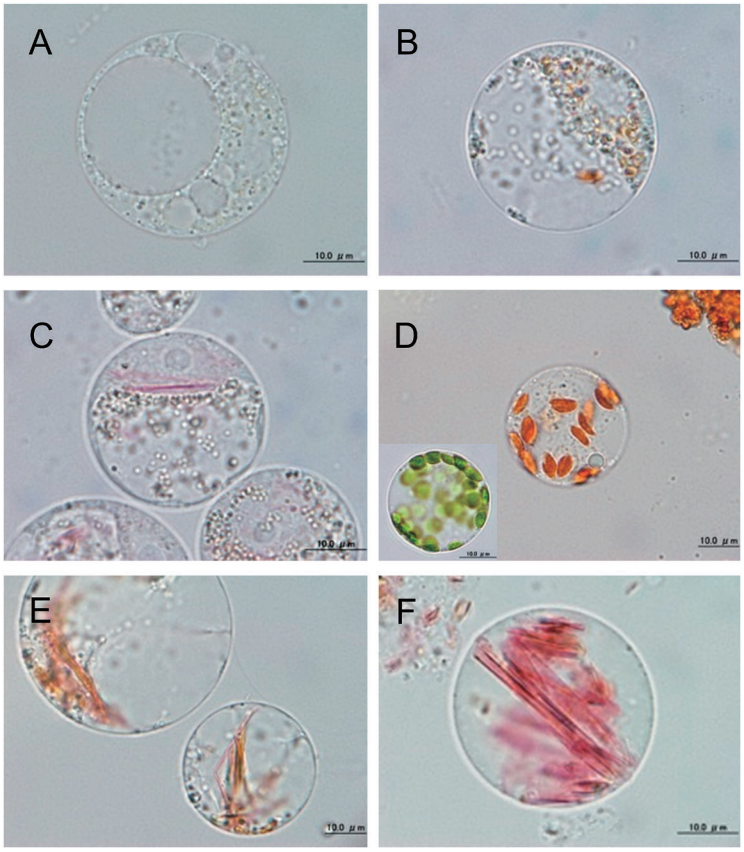


Figure 2. Protoplast isolated from N (A), L (B), and D (C) calli, red and green pepper (D), carrot (E), and tomato (F).

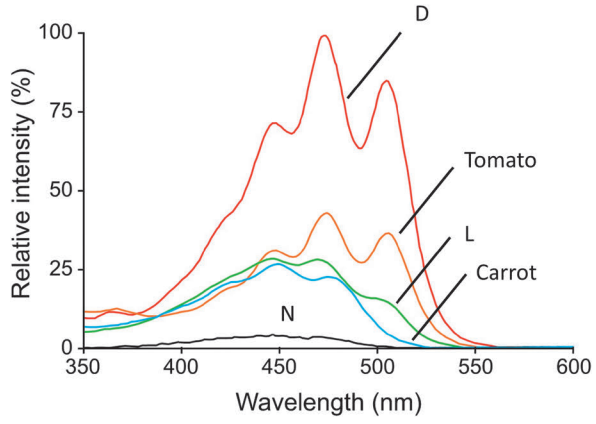


Figure 3. Absorption spectra of extracts from N (black), L (green) and D (red) callus lines, tomato (orange), and carrot (blue).

standard curve for β-carotene. The L line accumulated lycopene, β-carotene, and two unidentified carotenoids, while the D line accumulated mainly lycopene with smaller amounts of another unidentified carotenoid (Table 1). The total amount of carotenoids in D line was approximately ten times larger than that in L line.

Analysis of genome sequences from the N, L, and D lines

Changes in color as a consequence of subculturing can be attributed to carotenoid accumulation in both L and D lines. The D line was derived from the L line and showed an even more pronounced color phenotype. We analyzed genomic sequences of N, L, and D lines using NGS techniques. More than 365 million short-read

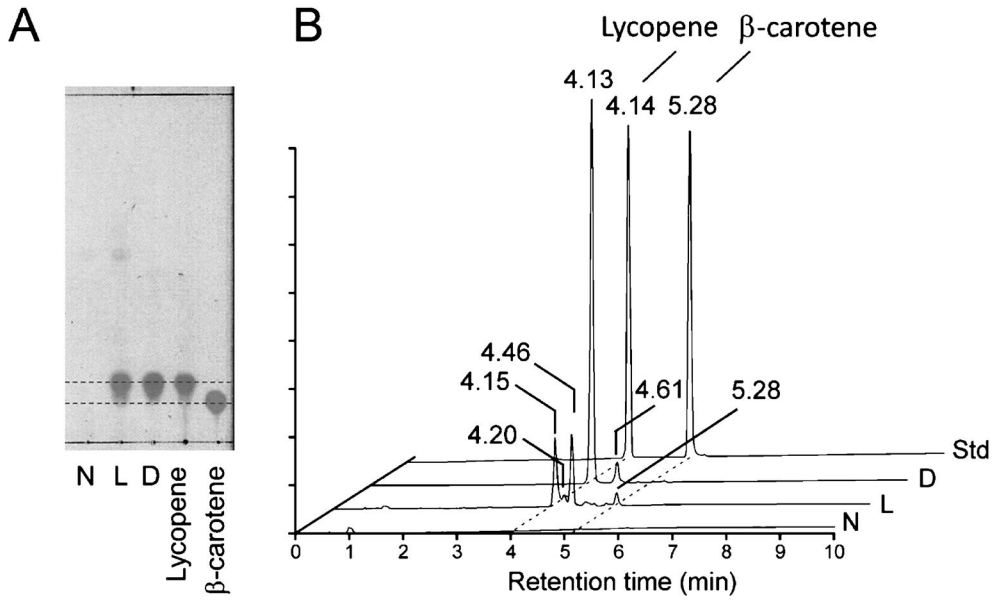


Figure 4. Chromatographic analyses of the pigments isolated from colored callus. The pigments from N, L, and D calli were extracted in acetone and analyzed using TLC (A) with authentic lycopene and *β*-carotene. The extracts were also analyzed by UPLC. Standard contains lycopene (retention time: 4.14 min) and *β*-carotene (retention time: 5.28 min). Traces are max plots obtained with detection range 330–600 nm.

Table 1. The accumulation of carotenoids in N, L, and D callus lines.

Line	Lycopene[a]	Rt 4.20[a,b]	Rt. 4.46[a,b]	Rt. 4.61[a,b]	*β*-carotene[a]	Total[a]
	*μ*g/gFW	*μ*g/gFW	*μ*g/gFW	*μ*g/gFW	*μ*g/gFW	*μ*g/gFW
N	0.829±0.333	n.d.[c]	n.d.[c]	0.103±0.054	0.220±0.102	1.15±0.33
L	30.7±8.80	1.10±0.36	6.09±2.27	n.d.[c]	1.18±0.20	39.1±11.0
D	406.1±76.0	n.d.[c]	n.d.[c]	20.7±3.56	n.d.[c]	426.8±78.6

[a] Data represent means±SD from three independent replicates. [b] Amounts of compounds were calculated by using standard curve generated for *β*-carotene. [c] n.d.: not detected.

Table 2. Coverage and sequencing depth of mapped reads with reference to the Nipponbare chromosomal genome.

	Mapped reads		Uniquely mapped reads		
	Number of nucleotides (bp)	Genome coverage with sequencing depth ≧5 (%)	Number of nucleotides (bp)	Genome coverage with sequencing depth ≧5 (%)	Average of sequencing depth (fold)
N	63,185,727,677	98.3	31,816,310,259	91.6	93.1
L	20,334,420,655	94.1	16,360,570,104	87.5	50.1
D	84,345,207,907	98.4	63,828,944,914	92.0	185.9

sequences from each callus line were generated. Short-reads from the genome of each line were mapped against the Nipponbare reference genome, and 55.9–75.0% and 21.5–35.6% of the total short–reads were successfully assigned to chromosomal and organelle genomes, respectively (Supplemental Table 2). Genomic short-reads were classified into 2 groups: unique or multiple locations. The average sequencing depth of the uniquely mapped reads in each line ranged from 50× to 185× across the entire genome, and covered 87.5–92.0% of the Nipponbare reference genome IRGSP1.0 (Table 2). Thus, all three lines were covered with a sufficient sequencing depth for further genomic analyses (Table 2).

Single-nucleotide polymorphism (SNP) sites were obtained from the three callus lines and annotated using the IRGSP1.0 gene set. All 3 lines presented over 77,000 SNPs and 20,000 insertions and deletions (InDels). From these, more than 5,000 SNPs and approximately 700 InDels were located within a specific gene region (data not shown). These data are available at a GBrowse, NGRC_Orange-colored Rices (http://www.nodai-genome.org/oryza_sativa_en.html).

If carotenoid accumulation is the result of a specific mutation, this mutation would have occurred in the genome of L and/or D lines. Although the total number of L-specific SNPs and InDels were 4,314 and 3,906, respectively, only 248 SNPs and 134 InDels were located in coding sequences (Supplemental Table 3). D specific SNPs and InDels were 10,947 and 5,919, respectively, and among them, 590 SNPs and 152 InDels were located

Table 3. Number of genes containing non-synonymous SNPs and InDels specific to L and/or D line.

	Number of genes with non-synonymous SNP	Number of genes with InDels
L specific	138	116
D specific	242	142
LD common	49	21

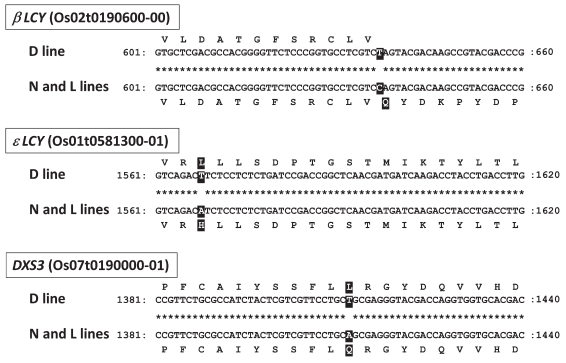


Figure 5. Mutations in *β LCY*, *ε LCY* and *DXS3*. The nucleotide sequences around the mutation sites in the ORFs of *β LCY*, *ε LCY* and *DXS3* genes were aligned. Asterisks indicate identical nucleotides, and numbers indicated the positions relative to the translational initiation site in the cDNA sequences. Predicted amino acid sequences were indicated upper and bottom sides of nucleotide sequences. Nucleotides and amino acid residues at the mutation sites were highlighted.

in coding sequences. Among line- specific SNPs, non-synonymous SNPs were detected in 138 genes in L line and in 242 genes in D line, respectively (Table 3). Only 85 SNPs were found to be common to both L and D lines and located in 49 genes. The common InDels between L and D were as few as 29 and were located in 21 genes.

Focusing on the genes encoding enzymes involved in carotenoid metabolism, one homozygous mutation was detected in the D lines. The SNP resulted in a non-synonymous substitution in the *β LCY* gene (locus name: Os02t0190600) of the D line (Figure 5 and Supplemental Figure 3). A direct sequencing analysis confirmed a single base substitution at nucleotide position 637 (relative to the first ATG codon) in the D line. This base substitution resulted in the replacement of glutamine by the stop codon at the putative amino acid position 213. Non-synonymous SNPs, which were heterozygous mutations, were also observed in *ε LCY* and *DXS3* genes in D and L lines, respectively. The other SNPs of carotenoid-biosynthetic genes detected in both L and D lines were silent mutations.

Analysis of expression of carotenoid pathway genes

DNA polymorphism analysis in L and D callus lines revealed the mutations in genes involved in carotenoid biosynthesis. To ensure expression of these genes, and to explore additional, possible causes leading to carotenoid accumulation in L and D lines from a different perspective, we determined the level of transcription of genes involved in both biosynthesis and catabolism of carotenoids by quantitative real-time PCR analyses (Figure 6). We selected genes for the PCR analysis based on the annotation in the RAP-DB, rice annotation project (http://rapdb.dna.affrc.go.jp/).

Regarding the MEP pathway, we analyzed the expression of genes encoding the enzymes DXS, DXR, HDR, and GGPS (Figure 6A). Enhanced expression of these genes would potentially cause the accumulation of lycopene and other carotenoids. *DXS1* was transcribed more intensively in L, whereas the amount of *DXS3* transcripts was higher in D. The amount of *GGPS1* transcripts was higher in both D and L lines than in the control (N) line.

Regarding the pathway from GGPP to *α*- and *β*-carotene, we analyzed the amount of transcripts for *PSY*, *PDS*, *ZDS*, *CRTISO*, *ε LCY*, and *β LCY* (Figure 6B). The amount of transcripts detected for *PSY2*, *PDS*, and *ZDS* was higher in L, whereas none of the genes upstream of lycopene showed enhanced expression in D. Furthermore, the transcription level of *PSY1*, *PSY3*, and *PDS* was lower in the D line in comparison with the N line. The amount of *ε LCY* transcripts was lower in both D and L lines, with their respective amounts being 3.4% and 39% of that of the N line. The amount of *β LCY* transcript in the D line was 34% of that of the N line.

Among the genes involved in the biosynthesis of xanthophylls from carotenes, we analyzed the expression of genes encoding *β*-hydroxylase (*β*-OHase) and CYP97s. The amount of *β-OHase1* transcripts was lower in both D and L lines than in the N line. The amount of *β-OHase2*, *β-OHase3*, and *CYP97A* transcripts was lower in the L line, while those of *CYP97C* were lower in the D line.

We further analyzed the changes of expression of genes annotated as carotenoid dioxygenases based on RAP-DB, because a cleavage reaction of carotenoids mediated by CCDs or NCEDs potentially affect for carotenoid contents. Total five genes, Os02t0704000, Os04t0550600, Os06t0162550, Os08t0371608, and Os12t0640600, were listed as putative carotenoid dioxygenase genes. Os02t0704000, Os04t0550600 and Os12t0640600 were corresponding to CCD4a, CCD7 and CCD1 (Vallabhaneni et al. 2010), respectively, while Os06t0162550 and Os08t0371608 were annotated as genes encoding carotenoid oxygenase domain containing proteins by RAP-DB. Among the analyzed genes, the transcript amounts of Os06t0162550 and Os08t0371608 markedly decreased in L and D lines compared with N line (Figure 7).

Discussion

We found that L and D callus lines accumulated high

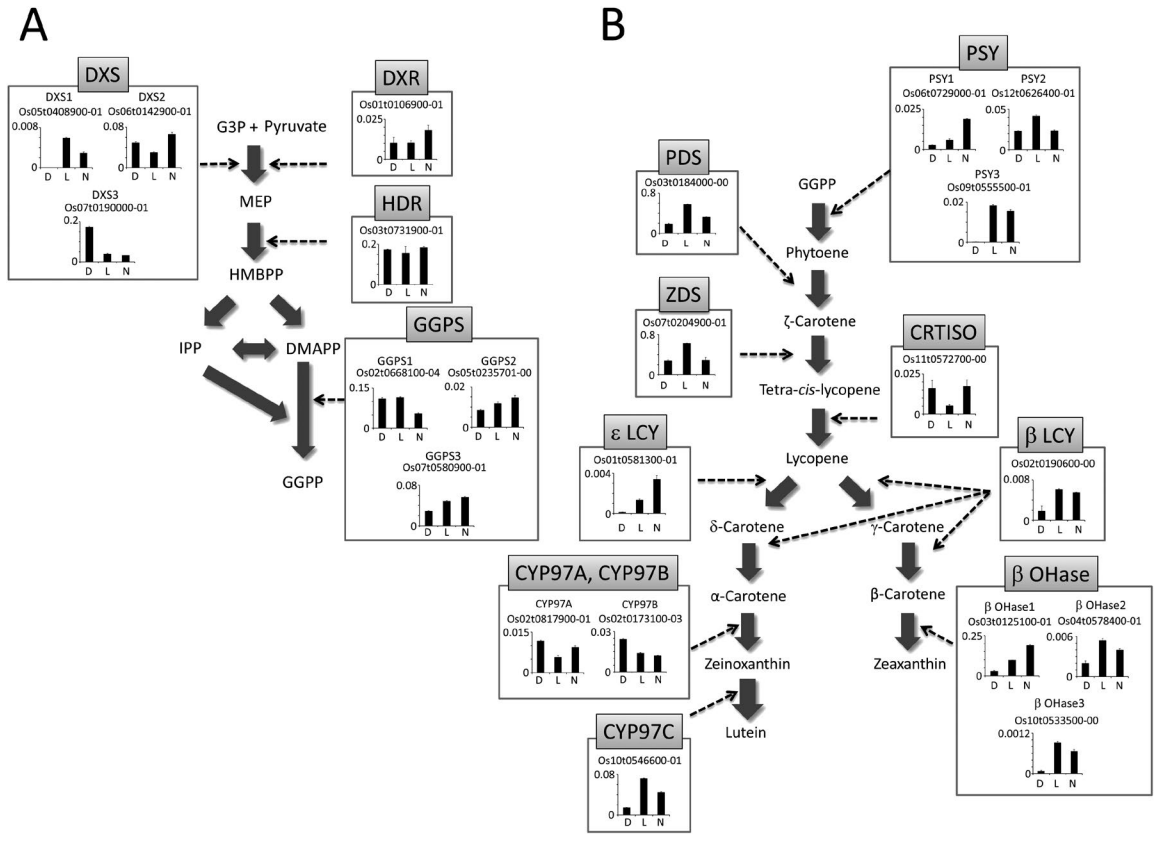


Figure 6. Changes in gene expression in MEP pathway (A) and carotenoid pathway (B). Y-axes of graphs indicate the relative transcript amounts expressed by the $2^{\Delta CT}$ values determined by qRT-PCR analyses by using *actin1* as internal control. Data represent means±SD from three independent replicates.

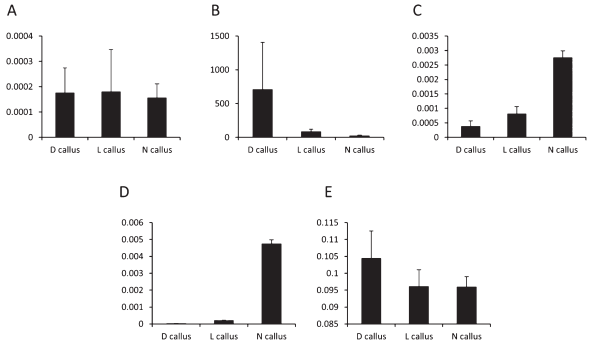


Figure 7. Changes in gene expression in putative carotenoid oxygenases. The changes in the transcript amounts of Os02t0704000 (A), Os04t0550600 (B), Os06t0162550 (C), Os08t0371608 (D), and Os12t0640600 (E) were analyzed by quantitative RT-PCR. Y-axes of graphs indicate the relative transcript amounts expressed by the $2^{\Delta CT}$ values determined by qRT-PCR analyses by using *actin1* as internal control. Data represent means±SD from three independent replicates.

concentrations of carotenoids. Interestingly, L and D callus lines showed the different carotenoid composition; the L line accumulated lycopene, β-carotene, and two unidentified carotenoids, while the D line accumulated lycopene at an extremely high concentration with a minor contribution of an unidentified carotenoid. Microscopic inspection of protoplasts prepared from D line indicated the presence of red needle-like structures that were similar to lycopene crystals in tomato protoplasts. Considering the high concentration of lycopene in D line, the formation of lycopene crystals in the cells is quite reasonable. The L line was isolated through repeated subculturing of the N line (callus line 120-2), which was a mutant line deficient in nitrate reductase (Wakasa et al. 1984) and did not show enhanced accumulation of carotenoids. The D line was emerged by subsequent repeated subculturing of the L line. Thus, it is likely that the metabolic changes occurred in N line led to orange color phenotype, and additional metabolic changes in L line further led to the more pronounced phenotype in the D line.

Carotenoids are essential pigments for photosynthetic organisms. In addition, the accumulation of carotenoids in plants gives colors to flowers and fruits, which attract insects and other animals serving as pollinators or seed dispersers. Therefore, carotenoids accumulate in specific organs at specific times in the plant life cycle. Accordingly, carotenoid accumulation is regulated by several different mechanisms. For example, during the ripening of tomatoes, carotenoid accumulation is strongly correlated with the enhanced expression of genes encoding biosynthetic enzymes (Ronen et al. 1999). In citrus species, carotenoid accumulation is

also correlated with similar changes in gene expression, although the types of carotenoids accumulated in citrus species differ from those in tomato (Kato et al. 2004). The variation in size and number of plastids have also been suggested to affect carotenoid accumulation on the basis of the enhanced accumulation of carotenoids observed in *hp2* and *hp3* mutants in tomato (Galpaz et al. 2008; Kolotilin et al. 2007). Yellow-colored petals of chrysanthemums accumulate carotenoids, while white-colored petals lack them because of the presence of carotenoid cleavage dioxygenase (CCD), which degrades carotenoids (Ohmiya et al. 2006). Cassava (*Manihot esculenta* Crányz) cultivars accumulate carotenoids in their yellow tubers (Welsch et al. 2010). The expression level of genes encoding carotenoid biosynthetic enzymes did not differ from other cultivars. However, in cultivars with yellow tubers, the PSY enzyme, which is considered the rate-limiting enzyme of carotenoid biosynthesis, showed higher enzymatic activity than those in other cultivars because of an amino acid substitution in the enzyme.

Since the color of L and D lines emerged after long-term culturing, it is likely that somatic mutations of the regulatory or biosynthetic genes cause the accumulation of carotenoids in these lines. Toward this end, we sequenced the genome of the D, L, and N lines by NGS, comparatively analyzed them, and found a mutation in *β LCY* gene in D line. This mutation generates a stop codon and results in the formation of a polypeptide with 214 amino acids, which is shorter than the original *β* LCY by 275 amino acid residues. Since this incomplete *β* LCY lacks the protic activation motif (FLEET) that is conserved in plant LCYs (Yu and Beyer 2012), the enzyme is considered inactive. Besides, 50% of the genomic sequences of *ε LCY* in the D line had a SNP leading to an amino acid substitution. This mutation may also negatively affect the enzymatic function of *ε LCY* because, in general, rarely a mutation improves enzyme function. A similar mutation was found in *DXS3* in the L line.

The qRT-PCR revealed the differences in expression level of carotenoid biosynthetic pathway genes among three lines. Several genes in the MEP pathway were upregulated in the L or D lines, but the *GGPS1* gene was the only gene whose expression was upregulated in both lines. This upregulation of *GGPS1* may support the increased carotenoid biosynthesis in L and D lines. Rice genome contains two additional *GGPS* homologs, *GGPS2* and *GGPS3*. Their expression was not enhanced in L and D lines. The estimation of the contribution of each GGPS isozyme to the metabolic flux of the MEP pathway is difficult, but, since the expression level of *GGPS1* was much higher than those of other *GGPS* genes, *GGPS1* may greatly contribute to flux of the MEP pathway. In addition, the increased accumulation of *DXS3* transcripts in D line may also affect the metabolic flux of the MEP pathway, as DXS is a key step in the MEP pathway. The enhanced expression of *DXS* has been shown to increase the rate of biosynthesis of carotenoid in tomato (Enfissi et al. 2005) and potato (Morris et al. 2006). The *DXS3* transcripts were the most abundant among all *DXS* transcripts.

Among the genes encoding the carotenoid metabolic pathway enzymes, the amounts of transcripts of the *PSY2*, *PDS*, and *ZDS* genes were increased in the L line. On the other hand, the transcription of *ε LCY* gene was downregulated both in L and D lines, with the degree of decrease in D line being larger than that in the L line. In addition, the transcriptional level of *β LCY* gene decreased in D line. The level of transcription of *β-OHase1* gene encoding the enzyme that initiates the metabolic pathway for the production of xanthophylls, is also lower both in L and D lines, and it was lower in the D line than in the L line. In addition, we detected decreases in the transcript amounts of two genes involved in carotenoid oxidation both in D and L lines with the degree of decrease being larger in the D line. Some of CCD enzymes have an activity to catalyze the cleavage of lycopene and carotene (Auldridge et al. 2006; Bouvier et al. 2005; Giuliano et al. 2003a). However, it remains that other gene products might also be involved in direct catabolism of lycopene and carotenes.

Considering the changes in the gene expression on the whole, in L line, the increased flux of carotenoid biosynthesis by upregulation of *PSY2*, *PDS*, and *ZDS* and suppression of conversion of lycopene into *α*-carotene by downregulation of *ε LCY*, as well as the suppression of carotenoid catabolism by carotenoid oxygenases, may have caused the accumulation of lycopene, *β*-carotene and other unidentified carotenoids. In D line, it is highly probable that the mutation of *β LCY* causes the accumulation of lycopene at an extremely high concentration without the accumulation of *β*-carotene. The markedly downregulated expression of *ε LCY* in D line is considered to interfere with funneling the metabolic flux into the other branch in the carotenoid biosynthetic pathway. The strong suppression of the expression of the *β-OHase 3* gene may also be involved in the accumulation of lycopene in the D line. Another possible factor that causes the lycopene accumulation in L and D line may be the suppressed conversion of carotenoids to appocarotenoids because the expression of two genes encoding carotenoid oxygenase domain containing proteins were lowered in these lines.

It has been reported that the suppression of *LCY* gene expression greatly affects carotenoid composition in plants. The embryos of maize *lcyB* mutant have been shown to accumulate lycopene and *δ*-carotene, whereas the endosperm accumulates lycopene, *δ*-carotene, *γ*-carotene, and *ε*-carotene (Bai et al. 2009); however,

these carotenoids are almost absent in the wild type. The suppression of ε *LCY* expression by RNAi techniques in *Brassica napus* resulted in the increase of carotenoids and in the ratio of β-carotene to lutein (Yu et al. 2008). In potato, tuber-specific silencing of ε *LCY* gene by introducing an antisense fragment of the gene caused an increase in β-carotene and zeaxanthin, although not in lutein (Diretto et al. 2006). Differently from these examples, the suppressed expression of both ε *LCY* and β *LCY* and mutation in β *LCY* may have caused an extreme increase in lycopene in the D line.

The enhanced accumulation of *PSY2*, *PDS*, and *ZDS* transcripts was not detected in line D, and their amounts were similar to those in the N line. In addition, the amount of *PSY1* and *PSY3* transcripts in the D line were much smaller than that in the N line. Maize *lcyB* mutant accumulates carotenoids, including lycopene, at high concentrations in embryos and endosperm, and it shows a suppressed accumulation of *Psy1* transcript, suggesting a feedback regulation mechanism of carotenoid biosynthesis by accumulation of downstream carotenoids (Bai et al. 2009). Lycopene or unidentified carotenoids accumulated at extremely high concentrations in line D, possibly leading to a feedback mechanism responsible for the observed lack of enhanced expression of lycopene biosynthetic genes.

In the present study, we detected a mutation in β *LCY* gene in D line, leading to formation of a truncated polypeptide, when we extracted the genes with the mutation in 90% of read depth. We performed direct sequencing β *LCY* gene, and confirmed that the mutation is homozygous and specific to D line. All genes of the N line are theoretically homozygous, because it was derived from pollen through anther culture. Thus, the heterozygosity generated by a somaclonal mutation in β *LCY* is considered to be lost during repeated subculturing. Loss of heterozygosity (LOH) has been well characterized in human cancers (Lasko et al. 1991). Tumor cells exhibit various forms of genetic instability including chromosomal deletion, mitotic non-disjunction, and recombination between homologous chromosomes, and these genetic instability causes LOH. Since enhanced genetic instability was also indicated in plant tissue cultures (Phillips et al. 1994), the LOH of the mutated β *LCY* allele in the D line during subculturing is not surprising.

Here, we report the establishment of two callus lines that accumulated carotenoids at high concentrations. The qRT-PCR analysis of the lines suggested that the activation of carotenoid biosynthesis and the suppression of the conversion of accumulated carotenoids are one of the causes of carotenoid accumulation. The changes in expression of multiple genes in the carotenoid biosynthetic pathway suggest the mutation of transcriptional factor(s) regulating the gene expression, but it remains unsolved. In addition, we could not exclude the possibility of any epigenetic changes such as DNA methylation and chromatin modifications (Lopez et al. 2010; Miguel and Marum 2011). Elucidating the regulatory mechanisms underlying the changes in the expression of the carotenoid biosynthetic pathway genes should help promoting the expansion of new technologies for the development of carotenoid-biofortified rice cultivars by manipulating mutated rice genes.

Acknowledgements

This study was supported by a MEXT-Supported Program for the Strategic Research Foundation at Private Universities (S0801025 and S1311017).

References

Apel W, Bock R (2009) Enhancement of carotenoid biosynthesis in transplastomic tomatoes by induced lycopene-to-provitamin A conversion. *Plant Physiol* 151: 59–66

Arai-Kichise Y, Shiwa Y, Ebana K, Shibata-Hatta M, Yoshikawa H, Yano M, Wakasa K (2014) Genome-wide DNA polymorphisms in seven rice cultivars of temperate and tropical japonica groups. *PLoS ONE* 9: e86312

Auldridge ME, McCarty DR, Klee HJ (2006) Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Curr Opin Plant Biol* 9: 315–321

Bai L, Kim E-H, DellaPenna D, Brutnell TP (2009) Novel lycopene epsilon cyclase activities in maize revealed through perturbation of carotenoid biosynthesis. *Plant J* 59: 588–599

Botella-Pavía P, Besumbes Ó, Phillips MA, Carretero-Paulet L, Boronat A, Rodríguez-Concepción M (2004) Regulation of carotenoid biosynthesis in plants: evidence for a key role of hydroxymethybutenyl diphosphate reductase in controlling the supply of plastidial isoprenoid precursors. *Plant J* 40: 188–199

Bouvier F, Isner JC, Dogbo O, Camara B (2005) Oxidative tailoring of carotenoids: a prospect towards novel functions in plants. *Trends Plant Sci* 10: 187–194

Carretero-Paulet L, Cairo A, Botella-Pavía P, Besumbes O, Campos N, Boronat A, Rodríguez-Concepción M (2006) Enhanced flux through the methylerythritol 4-phosphate pathway in Arabidopsis plants overexpressing deoxyxylulose 5-phosphate reductoisomerase. *Plant Mol Biol* 62: 683–695

Cazzonelli CI, Cuttriss AJ, Cossetto SB, Pye W, Crisp P, Whelan J, Finnegan EJ, Turnbull C, Pogson BJ (2009) Regulation of carotenoid composition and shoot branching in Arabidopsis by a chromatin modifying histone methyltransferase, SDG8. *Plant Cell* 21: 39–53

Cazzonelli CI, Pogson BJ (2010) Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci* 15: 266–274

Cazzonelli CI, Roberts AC, Carmody ME, Pogson BJ (2010) Transcriptional control of SET DOMAIN GROUP8 and CAROTENOID ISOMERASE during Arabidopsis development. *Mol Plant* 3: 174–191

Di Mascio P, Murphy ME, Sies H (1991) Antioxidant defense systems: the role of carotenoids, tocopherols, and thiols. *Am J Clin Nutr* 97(Suppl): 194S–200S

Diretto G, Tavazza R, Welsch R, Pizzichini D, Mourgues F, Papacchioli V, Beyer P, Giuliano G (2006) Metabolic engineering

of potato tuber carotenoids through tuber-specific silencing of lycopene epsilon cyclase. *BMC Plant Biol* 6: 13

Estévez JM, Cantero A, Reind A, Reichler S, León P (2001) 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J Biol Chem* 276: 22901–22909

Enfissi EMA, Fraser PD, Lois LM, Boronat A, Schuch W, Bramley PM (2005) Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. *Plant Biotechnol J* 3: 17–27

Galpaz N, Wang Q, Menda N, Zamir D, Hirschberg J (2008) Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. *Plant J* 53: 717–730

Giuliano G, Al-Babili S, von Lintig J (2003) Carotenoid oxygenases: cleave it or leave it. *Trends Plant Sci* 8: 145–148

Kato M, Ikoma Y, Matsumoto H, Sugiura M, Hyodo H, Yano M (2004) Accumulation of carotenoids and expression of carotenoid biosynthetic genes during maturation in citrus fruit. *Plant Physiol* 134: 824–837

Kolotilin I, Koltai H, Tadmor Y, Bar-Or C, Reuveni M, Meir A, Nahon S, Shlomo H, Chen L, Levin I (2007) Transcriptional profiling of *high pigment-2*^{dg} tomato mutant links early fruit plastid biogenesis with its overproduction of phytonutrients. *Plant Physiol* 145: 389–401

Lasko D, Cavenee W, Nordenskjöld M (1991) Loss of constitutional heterozygosity in human cancer. *Annu Rev Genet* 25: 281–314

Li F, Vallabhaneni R, Wurtzel ET (2008a) PSY3, a new member of the phytoene synthase gene family conserved in the Poaceae and regulator of abiotic stress-induced root cartenogenesis. *Plant Physiol* 146: 1333–1345

Li F, Vallabhaneni R, Yu J, Rocheford T, Wurtzel ET (2008b) The maize phytoene synthase gene family: overlapping roles for carotenogenesis in endosperm, photomorophogenesis, and thermal stress tolerance. *Plant Physiol* 147: 1334–1346

Li L, Paolillo DJ, Parthasarathy MV, Dimuzio EM, Garvin DF (2001) A novel gene mutation that confers abnormal patterns of beta-carotene accumulation in cauliflower (*Brassica oleracea* var. *botrytis*). *Plant J* 26: 59–67

Lois LM, Rodríguez-Concepción M, Gallego F, Campos N, Boronat A (2000) Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant J* 22: 503–513

Lopez CMR, Wetten AC, Wilkinson MJ (2010) Progressive erosion of genetic and epigenetic variation in callus-derived cocoa (*Theobroma cacao*) plants. *New Phytol* 186: 856–868

Lu S, Eck Van J, Zhou X, Lopez AB, O'Halloran DM, Cosman KM, Conlin BJ, Paolillo DJ, Garvin DF, Vrebalov J, et al. (2006) The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of beta-carotene accumulation. *Plant Cell* 18: 3594–3605

Miguel C, Marum L (2011) An epigenetic view of plant cells cultured in vitro: somaclonal variation and beyond. *J Exp Bot* 62: 3713–3725

Morris WL, Ducreux LJ, Hedden P, Millam S, Taylor MA (2006) Overexpression of a bacterial 1-deoxy-D-xylulose 5-phosphate synthase gene in potato tubers perturbs the isoprenoid metabolic network: implications for the control of the tuber life cycle. *J Exp Bot* 57: 3007–3018

Müller AJ, Grafe R (1978) Isolation and characterization of cell lines of *Nicotiana tubacum* laking nitrate reductase. *Mol Gen Genet* 177: 145–153

Ohmiya A, Kishimoto S, Aida R, Yoshioka S, Sumitomo K (2006) Carotenoid cleavage dioxygenase (CmCCD4a) contributes to white color formation in chrysanthemum petals. *Plant Physiol* 142: 1193–1201

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchliffe E, Adams JL, Silverstone AL, et al. (2005) Improving the nutritional valueof Golden Rice through increased pro-vitamin A content. *Nat Biotechnol* 23: 482–487

Phillips MA, León P, Boronat A, Rodríguez-Concepción M (2008) The plastidial MEP pathway: unified nomenclature and resources. *Trends Plant Sci* 13: 619–623

Phillips RL, Kaepplert SM, Olhoft P (1994) Genetic instability of plant tissue cultures: Breakdown of normal controls. *Proc Natl Acad Sci USA* 91: 5222–5226

Porebski S, Bailey LG, Baum BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Rep* 15: 8–15

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

Ross AC, Chen Q, Ma Y (2011) Vitamin A and retinoic acid in the regulation of B-cell development and antibody production. *Vitam Horm* 86: 103–126

Stahl W, Schwarz W, Sundquist AR, Sies H (1992) *cis*-trans isomers of lycopene and β-carotene in human serum and tissues. *Arch Biochem Biophys* 294: 173–177

Underwood BA, Arthur P (1996) The contribution of vitamin A to public health. *FASEB J* 10: 1040–1048

Vallabhaneni R, Bradbury LM, Wurtzel ET (2010) The carotenoid dioxygenase gene family in maize, sorghum, and rice. *Arch Biochem Biophys* 504: 104–111

Wakasa K, Kobayashi M, Kamada H (1984) Colony formation from protoplasts of nitrate reductase-deficient rice cell lines. *J Plant Physiol* 117: 223–231

Welsch R, Arango J, Bär C, Salazar B, Al-Babili S, Beltrán J, Chavarriaga P, Ceballos H, Tohme J, Beyer P (2010) Provitamin A accumulation in cassava (*Manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. *Plant Cell* 22: 3348–3356

Welsch R, Wust F, Bar C, Al-Babili S, Beyer P (2008) A third phytoene synthase is devoted to abiotic stress-induced ABA formation in rice and defines functional diversficationof PYSs. *Plant Physiol* 147: 367–380

Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, Beyer P, Potrykus I (2000) Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* 287: 303–305

Yu B, Lydiate DJ, Young LW, Schäfer UA, Hannoufa A (2008) Enhancing the carotenoid content of *Brassica napus* seeds by downregulating lycopene epsilon cyclase. *Transgenic Res* 17: 573–585

Yu Q, Beyer P (2012) Reaction specificities of the ε-ionone-forming lycopene cyclase from rice (*Oryza sativa*) elucidated in vitro. *FEBS Lett* 586: 3415–3420

Supplemental Table 1 Primer sets used for qRT-PCR

Gene name	Forward (5' to 3')	Reverse (5' to 3')
DXS1	AGG CAC GTG AAT GAA AAA GAA AA	TTT GGC GCA TTT CGT AGT AGT TAT
DXS2	GGA ACG GTT GGT GGT TTT GG	CAA GCA ACC CAT CGA GTG AAA
DXS3	GCT CCT CGG AAA AGC AGT GAA	TCT ATA TAC CGC GAC GCA GCA A
DXR	CAC CTG GCC CCG ACT AGA TC	AGC ACT CAG AAC TCC TGT CAT GGT
HDR	ACA AGC GAT GTC TGT ACA GTA GTG AA	ATT TGC CCC CAA CAT ACT ACG T
GPS	CCG TGC AAT CAA AGC CAT AGA	AGA GCA CGC CTT GAT GTC AGA
GGPS1	GTC GGG CAA CGG CAA GAT	ACG AAG CTG TAC AAG GGA AGC A
GGPS2	GAT TGA ACG TTT GCG CAT GT	AGT CCT TCC CCG CTG TCT TG
GGPS3	TCT TGT GAT ACG TGC TAG TTG TAG TTT C	TAG GGC AGA GAG GCT GAT TTT C
PSY1	AGG CGG GAA GAG GAG ATA AAG G	TCG GAA GGC AGA TGC TGT CT
PSY2	GGT TCT GGC ATC TTT GTG GTT AT	TTT ACA TAC GCG CGC TTG GT
PSY3	CGA GGC CAA CGA TTA CAA CA	GCC ACA ATC TTC TTT GCC TTT G
PDS	TGA AGA TGG AGA TTG GTA TGA AAC TG	GCT CGC CAA ACA AGT TCT GTA TG
ZDS	TGC TGG AGT GAG TGG ATT GCT A	ATG AGA GTC GCT GCG GAA A
CRTISO	TTT TGC CTG CTG ATA CTG ATT GC	TGG GTC AAG CAC TGT AGG GAT ACT
ε LCY	GAC GCT TTC TTC AGT GGA TCT CA	TGG TTC GGC GCA ATT GT
β LCY	CCA GGA GAT GAA GTG TGT GAG TTG	CAG TAC ATT GAG AGG CAC ATC ATC A
β OHase1	AGG AGG TGG GTG GGA TTG A	ATC GCA TCT AAG GTC TCT TTC CTC TT
β OHase2	GAG GAG GTT GGT GGT CTG GAA	CCG CCG TCA AAA TCA CAA G
β OHase3	GAG GAG GTG GGT GGG ACT GA	AAA GCG ATC GAA AAT GGT AGC A
CYP97A	GGC GGA CCA AGG AAA TGT GTA G	GAA AAT CAA AGC GCC TGA CAA G
CYP97B	CCG GGT TCC ATG CCT TAT TC	CCC ATG TTC GCC CCT TTC
CYP97C	CGG TGA TGG TTG ACA GAG TTT TT	CAA GGA CAA ACC AAT CAC ATC TAA AG
Os02t0704000	AAA GGA TTT GCC CTG CTC ACT	GCG TTC TTC TTC CTG CCA TAG C
Os04t0550600	GTC CGG CTG CAC ATG TCA A	GCT TCT TGT GGT GCC AAT TGT
Os06t0162550	GCT GAA GGC GAG CCA TGT G	AGG CTT TCC AAA TCT CAC GGT AT
Os08t0371608	CAG CAG CAG CTT ACC ACC AA	CCA CAA ACG CAT CCA GCA A
Os12t0640600	TGA TCC GGT GGC AGT TGT T	GTT TTG CCA GTT GTT CCT CGT T
actin1	TGG CAT CTC TCA GCA CAT TCC	TGC ACA ATG GAT GGG CCA GA

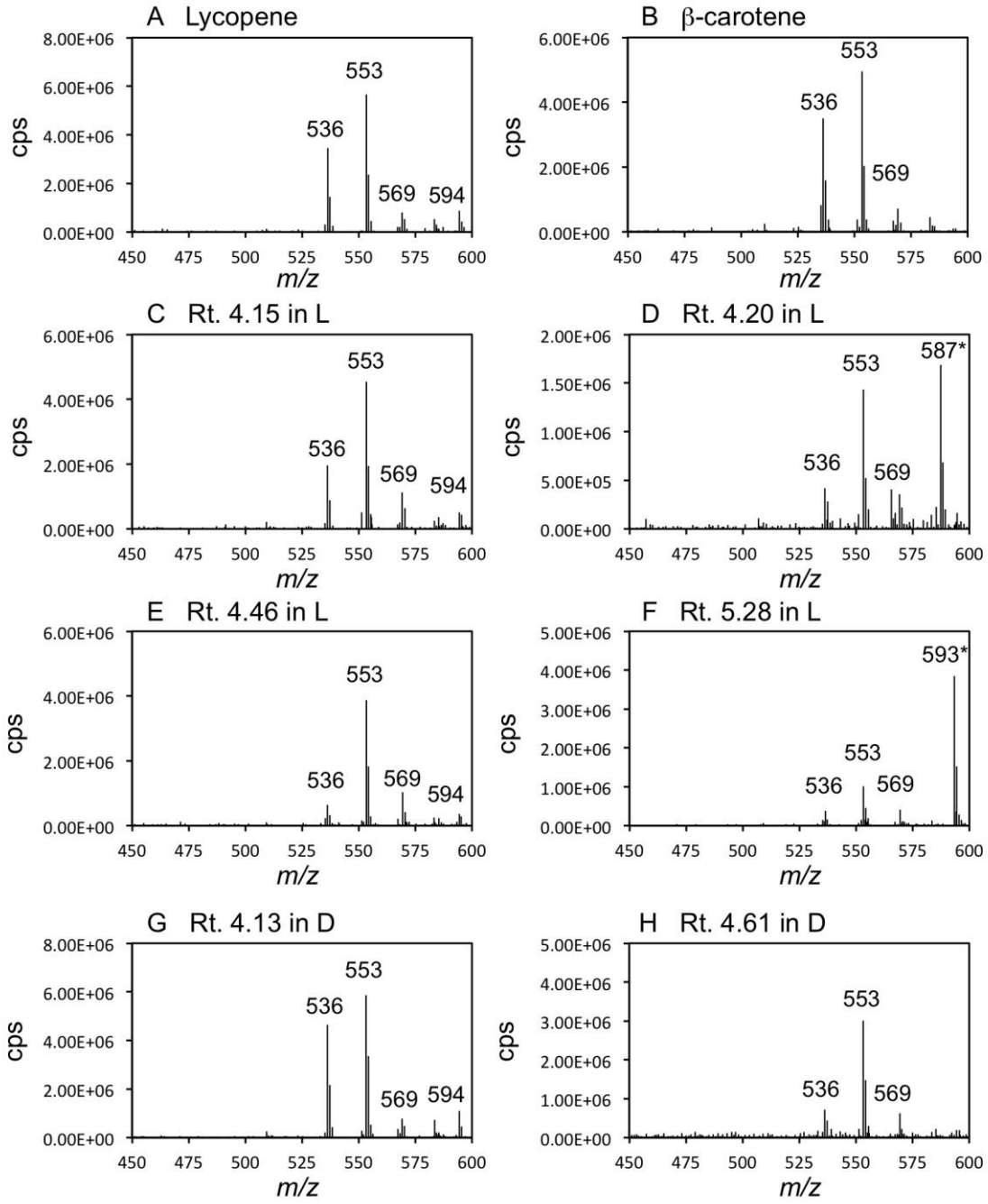
Supplemental Table 2. Classification of short-reads mapped onto mitochondrial, plastid, and chromosomal genome IRGSP1.0, as well as unmapped reads

	Mapped reads												Unmapped reads	
	Mitochondrion				Plastid				Chromosome					
	Unique		Multi		Unique		Multi		Unique		Multi			
	No. of reads	%	No. of reads	%	No. of reads	%	No. of reads	%	No. of reads	%	No. of reads	%	No. of reads	%
N	136,932,866	14.0	85,897,393	8.8	31,086,872	3.2	48,926,742	5.0	554,520,219	56.5	86,390,809	8.8	37,542,464	3.8
L	48,731,948	13.3	33,563,568	9.2	11,896,058	3.3	35,964,227	9.8	176,133,059	48.2	28,183,691	7.7	31,050,054	8.5
D	106,737,809	9.3	66,758,486	5.8	71,918,205	6.3	806,730	0.1	748,146,854	65.2	112,992,618	9.8	40,508,256	3.5

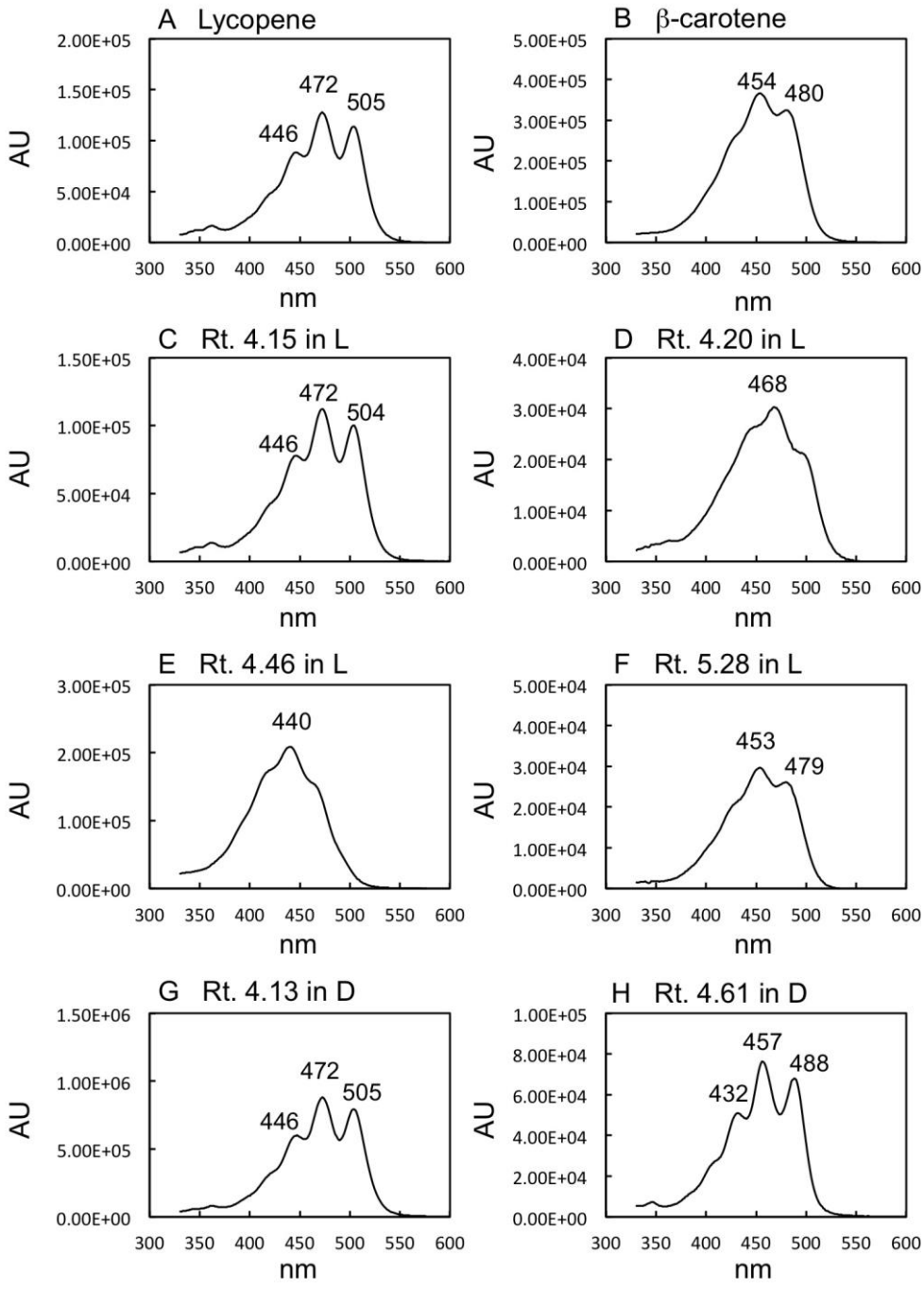
Supplemental Table 3 Annotation of SNPs, insertions and deletions specific to L and/or D line

	L specific	D specific	LD common
SNP			
Total	4,314	10,947	1,385
Intergenic	3,500	8,992	1,089
Genic	814	1,955	296
Intron	434	958	112
UTRs	132	407	58
CDS	248	590	126
Synonymous	78	187	41
Nonsynonymous	170	403	85
Insertion, Deletion			
Total	3,906	5,919	775
Intergenic	3,176	4,974	631
Genic	730	945	144
Intron	446	575	70
UTRs	150	218	45
CDS	134	152	29

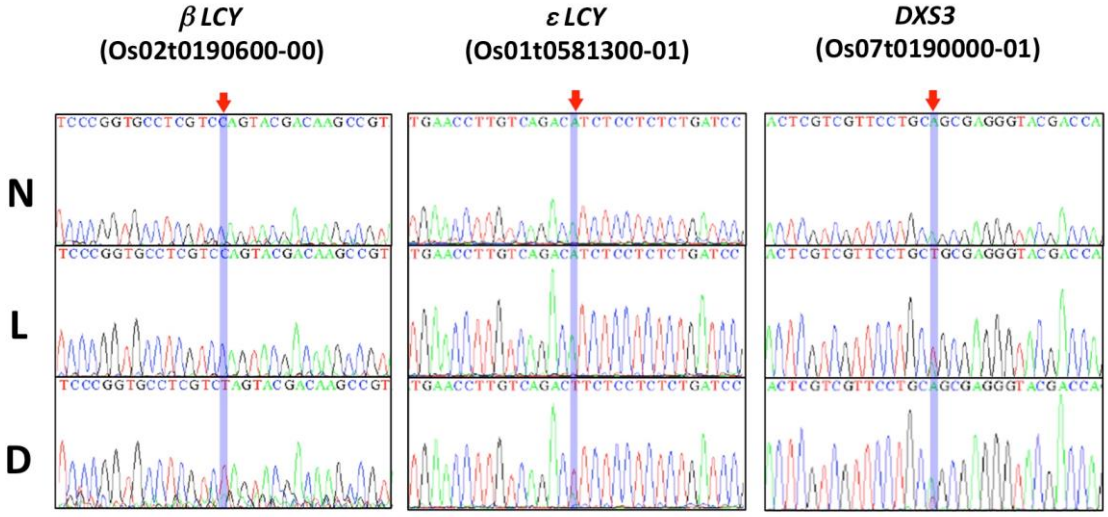
SNPs, insertions, and deletions specific to the L or D line on the IRGSP rice pseudomolecules were classified as genic and intergenic, and locations within gene models were annotated. The number of SNPs, insertions, and deletions in each class is shown.



Supplemental Figure 1 ESI-MS of pigments. ESI-MS of authentic lycopene (A), β-carotene (B), Rt. 4.15 (C), 4.20 (D), 4.46 (E) and 5.28 (F) min in line L, Rt. 4.13 (G) and 4.61 (H) min in line D were recorded. Peaks with asterisks are ions derived from impurities.



Supplemental Figure 2 Absorption spectra of the pigments analyzed. Absorption spectra of authentic lycopene (A), β-carotene (B), Rt. 4.15 (C), 4.20 (D), 4.46 (E) and 5.28 (F) min in line L, Rt. 4.13 (G) and 4.61 (H) min in line D were recorded.



Supplemental Figure 3 Mutations detected in *βLCY*, *εLCY* and *DXS3*.