Differential accumulation of β -carotene and tissue specific expression of phytoene synthase (*MaPsy*) gene in banana (*Musa* sp) cultivars

R. Dhandapani¹ · V. P. Singh¹ · A. Arora¹ · R. C. Bhattacharya² · Ambika Rajendran³

Revised: 25 September 2017 / Accepted: 29 September 2017 / Published online: 16 October 2017
© Association of Food Scientists & Technologists (India) 2017

Abstract An experiment was conducted with twelve major Indian banana cultivars to investigate the molecular relationship between the differential accumulation of β -carotene in peel and pulp of the banana fruit and carotenoid biosynthetic pathway genes. The high performance liquid chromatography showed that all banana cultivars accumulated two–three fold more β -carotene in non-edible portion of the banana fruit. However, *Nendran*, a famous orange fleshed cultivar of South India, had high β -carotene content (1362 $\mu\text{g}/100\text{ g}$) in edible pulp. The gene encoding *Musa accuminata* phytoene synthase (*MaPsy*) was successfully amplified using a pair of degenerate primers designed from *Oncidium* orchid. The deduced amino acid sequences shared a high level of identity to *phytoene*

synthase gene from other plants. Gene expression analysis confirmed the presence of two isoforms (*MaPsy1* and *MaPsy2*) of *MaPsy* gene in banana fruits. Presence of two isoforms of *MaPsy* gene in peel and one in pulp confirmed the differential accumulation of β -carotene in banana fruits. However, *Nendran* accumulated more β -carotene in edible pulp due to presence of both the isoforms of *MaPsy* gene. Thus, carotenoid accumulation is a tissue specific process strongly dependent on differential expression pattern of two isoforms of *MaPsy* gene in banana.

Keywords *Musa accuminata* · *Nendran* · Vitamin A · Carotenoids · β -Carotene · Phytoene synthase

Electronic supplementary material The online version of this article (doi:10.1007/s13197-017-2918-8) contains supplementary material, which is available to authorized users.

✉ R. Dhandapani
dandyman2k3@yahoo.co.in

V. P. Singh
vpsingh@iari.res.in

A. Arora
ajayarora@iari.res.in

R. C. Bhattacharya
bhattacharyarc@rediffmail.com

Ambika Rajendran
rambikarajendran@gmail.com

¹ Division of Plant Physiology, ICAR-Indian Agricultural Research Institute, New Delhi, India

² ICAR-National Research Centre for Plant Biotechnology, New Delhi, India

³ Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Introduction

Carotenoid is a group of isoprenoid derived natural pigments produced in plants for various physiological processes starting from germination to senescence (Cazzonelli 2011). Carotenoids play an important role in acquisition of light energy during photosynthesis, protecting photosynthetic apparatus from excessive light energy as accessory pigment, development of colorful flowers and fruits to attract pollinators and agents involved in seed dispersal. These functional pigments act as antioxidant which protects the membranes against damaging free radicals and increases seed longevity (Smolikova and Medvedev 2015). These lipid soluble pigments serve as a precursor for biosynthesis of hormones like abscisic acid (ABA), gibberellic acid (GA) in plants and vitamin A in human (Cazzonelli 2011).

Nutritional deficiency like vitamin deficiency and chronic malnutrition are common among the poorest segments of Indian population. Among these, Vitamin A

Deficiency (VAD) is a serious preventable health problem affecting about 78 million preschool age children and 4 million pregnant women in India (Akhtar et al. 2013). Vitamin A, a natural antioxidant protects from oxidative stress that leads to nutritional related chronic diseases like diabetes and age related muscular degeneration, anaemia, mortality, morbidity and vision related eye problems (Pinazo-Duran et al. 2014). Human body cannot synthesize vitamin A, thus the diet based on consumption of carotenoid (lutein, zeaxanthin, lycopene and β -carotene) rich fruits and vegetables is important for prevention of chronic diseases (Boeing et al. 2012). Out of 50 or so carotenoids having provitamin A activity, β -carotene is the most potent precursor of vitamin A. The main dietary source of β -carotene in plant foods is commonly characterized by the yellow or orange colour in fruits and vegetables such as mango, banana, papaya, carrot and beetroot. Breeding or genetic engineering for increasing nutritional value of a common food like banana ought to have positive effects on health and well-being of millions of people.

Earlier investigation suggested that carotenoids are synthesized and accumulated in plastids by nuclear-encoded enzymes. The first committed step in carotenoids biosynthesis is a head-to-head condensation of two molecules of geranyl geranyl pyrophosphate to form colorless phytoene (Nisar et al. 2015). This step was catalyzed by an enzyme called phytoene synthase (PSY) and is characterized as a rate limiting enzyme for carotenoid biosynthesis. Genes (*Psy*) encoding phytoene synthase enzyme have been identified and characterized from numerous higher plants including *Arabidopsis*, daffodil, tomato, tobacco, maize, melon, citrus, sunflower and banana *etc* (Mlalazi et al. 2012). In most plants, including *Arabidopsis*, daffodil and citrus, only one *Psy* gene has been found. However, two differentially expressed genes, *Psy1* and *Psy2*, have been found in tomato, bell pepper, banana, tobacco, muskmelon, cassava and carrot (Fraser et al. 1999; Romer et al. 1993). *Psy1* has been predominantly identified as a fruit ripening related gene, whereas *Psy2* was related to carotenoid biosynthesis in mature leaves (Yuan et al. 2015). Recently, *Psy3* a new paralogue of *Psy* gene was reported to play a key role in controlling flux to carotenoids in maize and rice in response to abiotic stress (Li et al. 2008). The best model system for carotenoid formation has been well established in tomato. At colour break stage in tomato fruit, with the accumulation of lycopene and β -carotene, mRNA levels of *Psy1* increased significantly, while the mRNAs of both lycopene cyclases (*Lcy-b* and *-e*) disappeared (Fraser et al. 1999). Although similar reports have been demonstrated in other fruit producing plants, there is no research report available on isolation of banana *Psy* gene and its expression pattern in peel and pulp of banana fruits from Indian cultivars.

India is the largest producer of banana with 24% of world production, followed by China and Philippines (FAOSTAT, 2016). The most edible bananas and plantains are descended from a wild ancestor, *Musa acuminata* (AA) and *Musa balbisiana* (BB). Diploids (AA, AB), triploids (AAB, ABB) and tetraploids (AAAA, ABBB) arising from the genomes of these two genera gives rise to various types of edible bananas and plantains with different fruit composition. Previous research in banana indicated a positive significant correlation between carotenoid content and deeper yellow or orange flesh colour (Englberger et al. 2003). The cultivars with the deepest yellow to orange colour flesh (*Asupina*, a Fe'i banana) contain the highest level of *trans* β -carotene ($1412 \mu\text{g } 100 \text{ g}^{-1}$). The level of total carotenes in the yellow-fleshed cultivars ranged from 400 to 1000 μg per 100 g. Bananas of the cream flesh contains low levels of provitamin A carotenoids, ranging from 21 to 70 μg of β -carotene per 100 g of edible portion (Englberger et al. 2006). Literature also suggested that the level of β -carotene content in banana fruits is highly specific to genotype, ploidy level and geographic locations (Amorim et al. 2009).

Our earlier research work on banana fruit with highest β -carotene content identified red banana as carotenoid rich banana fruit in India (Arora et al. 2008). Literature also suggested that non edible peel of the banana fruits contain more β -carotene than the edible pulp (Borah et al. 2006; Mohapatra et al. 2010). Some tribal communities in India prepare a folk Indian medicine (“*Kol Khar*”) from plantain banana fruit (“*bheem kol*”) peel for the treatment of various diseases (Purkayastha and Nath 2006). These research reports led us to report quantification of β -carotene, isolation of cDNA clone encoding banana *phytoene synthase* gene (*MaPsy*) and its expression profile in peel and pulp of major Indian banana cultivar fruits. The major aim of this experiment was to investigate the rationale of high β -carotene content in banana cultivars with orange yellow colour flesh over other cream-fleshed cultivars and understand the molecular basis of differential expression of β -carotene in peel and pulp of banana fruit.

Materials and method

Plant material

Fruits of twelve major Indian banana (*Musa accuminata*) cultivars viz., *Kunnan*, *Njali Poovan*, *Padalai Moongil*, *Cavendish*, *Morish*, *Red Banana*, *YKM-5*, *Karpuravalli*, *Monthan*, *Nendran*, *Poovan*, and *FHIA-3* were collected at green ripe stage (GR-2) from the orchard of Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India. The peel and pulp samples from fruits of all banana

cultivars was separated using a sharp knife and stored in refrigerator at -20°C till further analysis.

High performance liquid chromatography quantification of β -carotene

β -Carotene was separated and purified by column chromatography as described by Peach and Tracy (1995) and then quantified by high performance liquid chromatography (HPLC) system. Banana samples (1 g peel and 5 g pulp) were repeatedly extracted with 5 ml of 80% acetone using a pestle and mortar until the residue was colourless. This acetone extract was filtered using Whatman filter paper 1 and transferred into a separating funnel containing twenty ml of stabilised petroleum ether with 0.1% butylated hydroxytoluene (BHT). After mixing gently, 20 ml of 5.0% sodium sulphate solution was added. The two phases were separated and the aqueous phase was re-extracted with petroleum ether until aqueous phase turned colourless. The pooled extracts were evaporated in vacuum at 30°C and residues were dissolved in petroleum ether. The petroleum ether extract was passed through a $10\text{ cm} \times 1\text{ cm}$ column of alumina (neutral charge) with 1 cm layer of sodium sulphate at the top and bottom forming a sandwich bilayer. Acetone and hexane (2:1) solvent mixture was used to elute pure β -carotene fraction. An aliquot (10 μl) of this elute was injected in to the normal phase silica column ($125 \times 4\text{ mm}$ I.D: particle size, 5 μm) of the HPLC system along with hexane as mobile phase at the rate of 1 ml/min. 0.5, 1.0 and 2.0 ppm of β -carotene standards were used as reference. The total carotenoid content was estimated using the primary acetone extract in visible spectrophotometer and the standard formula.

Total DNA isolation

Genomic DNA was isolated from banana fruit using a DNeasy Plant mini kit according to the manufacturer's instructions.

Total RNA isolation

RNA from peel and pulp of four selected banana fruits was isolated using Asif et al. (2000) method with suitable modification. 1.0 g banana fruit tissue were frozen and ground in liquid nitrogen using pestle and mortar. The frozen powder was transferred to a polypropylene centrifuge tube containing pre heated 10 ml extraction buffer and freshly added 10 μl of β -mercaptoethanol. After vortexing, the tubes were incubated at 65°C for 1 h with gentle mixing every 15 min. The tubes were cooled down to room temperature in running tap water and an equal volume of chloroform: isoamyl alcohol (24:1) was added.

Tubes were shaken vigorously until the two phases form an emulsion. Tubes were centrifuged at 12,000 rpm for 15 min at room temperature to pellet insoluble cell debris. The aqueous phase was collected and re-extracted with an equal volume of chloroform: isoamyl alcohol (24:1). Tubes were centrifuged once again at 12,000 rpm for 15 min (4°C). The aqueous phase was separated and 10 M LiCl was added to a final concentration of 3 M. RNA was allowed to precipitate at 4°C overnight. The precipitated RNA was recovered by centrifugation at 12,000 RPM at 4°C for 20 min. The upper aqueous phase was collected and 1/30th volume of 3 M sodium acetate (pH 5.2) was added along with 0.1 volume of 100% ethanol. After mixing the solution was kept on ice for 30 min. After centrifuge in a microfuge at 4°C for 25 min, the pellets were discarded. To the clear supernatant 3 M Na acetate pH 5.2 to a final concentration of 0.3 M and 3 volumes of 100% ethanol was added. The RNA was allowed to precipitate at -80°C for overnight, and then it was recovered by centrifugation in a microfuge at 4°C for 20 min. The pellets were washed twice with equal volume of 70% ethanol. The pellets were vacuum dried and re-suspended in 10 μl of DEPC treated water and store at -80°C until use. The purity and concentration of RNA was checked using UV-Vis spectrophotometer by taking the absorbance at 260 and 280 nm. The pure RNA with good yield was selected for further analysis based on the absorption ratio of ~ 1.8 –2.0 at 260/280.

Degenerate primer designing

The degenerate primers were synthesized based on the conserved coding regions of *Psy* genes from *Oncidium excavatum*, *Zea mays*, *Sorghum bicolor* and *Triticum aestivum*. The Coding DNA Sequence (CDS) were collected and the conserved regions were predetermined using alignment of *Psy* sequences from closely related plants available at NCBI database. The major bioinformatics website used for collecting protein and CDS sequences was <http://www.ncbi.nlm.nih.gov/>. The gathered protein sequences were aligned using multiple sequence alignment software (Clustal Omega) available at <http://www.ebi.ac.uk/Tools/msa/clustalo/>. The most conserved region of nucleotide marked with asterisks (*) were selected for making degenerate primers. Both forward and reverse primers.

(F-5'ACAAGAATTTTCGAYGAGCTCTATC3' and R-5'AWGCCMACACCGGCCACCTA3') were derived from the conserved amino acid sequences. The relative position of the primers to the coding region of a known plant PSY was towards 3' end and expected fragment size was 390 bp. The degenerate primers designed were blast searched twice to get probability of amplification from

different plant and animal relatives using <http://www.ncbi.nlm.nih.gov/blast>. The nucleotide sequences of designed primers were characterised to study the quality parameter like possibility of forming primer dimer, hairpin loop, GC content and melting temperature using oligoanalyzer software available at <http://eu.idtdna.com/calc/analyzer>.

Semi-quantitative reverse transcription PCR

Reverse transcription to synthesis cDNA was carried out using PrimeScript^{TM} first strand cDNA synthesis Kit from TAKARA BIO INC. According to the manufactures protocol, 5 µg of RNA was added along with 1 µl of dNTP (10 mM) and 1 µl of Oligo (dT)₁₈ primer (50 µM). The volume was made up to 10 µl using sterile water and the mixture was placed at 65 °C for 5 min. After the brief spin of vial, 4 µl of RT buffer (5X), 0.5 µl (20 units) of RNase inhibitor (40U/µl), 0.5 µl (200 units) of Reverse Transcriptase (200U/µl) and 5 µl (made up to 20 µl) of sterile water were added. The solution was mixed well and incubated at 42 °C for 60 min to denature RNA-cDNA hybrids. The enzymes was inactivated by incubate at 95 °C for 5 min and quickly placed on ice. This product was directly used as cDNA template for PCR expression analysis.

PCR conditions

PCR amplifications was done using degenerate primers in 25 µl reactions consisting of 0.2 mM dNTPs, 2 mM MgSO₄, 0.5 mM of both forward and reverse primers, 0.5U of Taq polymerase (NEB) and 200 ng of genomic DNA as template. Thermo cycler conditions consisted of 3 min initial denaturation at 94 °C, followed by 35 cycles of 94 °C for 1 min, 55 °C for 45 s, and 72 °C for 50 s and a final extension of 72 °C for 10 min.

Cloning and sequencing of PCR products

PCR amplification products were size separated on a 1.0% TAE agarose gel containing 0.5 µg/ ml intercalating dye (Ethidium bromide). The amplicon of interest was excised and purified using the DNA, E. Z. N. A. Gel Extraction Kit (OMEGA BIO-TEK) prior to ligating into pCR4-TOPO cloning vector (Invitrogen). Purified plasmid DNA was sequenced using universal M13 reverse and T7 sequencing primers. Sequencing was performed using the UDSC (University of Delhi South Campus) DNA sequencing facility funded by Department of Biotechnology, New Delhi, India.

Sequencing and bioinformatics tools

From sequenced nucleotide data, the amino acid sequence and its open reading frame were predicted using ORF

Finder (<https://www.ncbi.nlm.nih.gov/orffinder/>). EMBOSS transeq function in EMBL-EBI database were used for in-silico translation of nucleotide sequence of *MaPsy* cDNA fragments in to amino acids. Amino acid sequence comparison with other related plants (Maize, Sorghum and Oncidium) were analysed using multiple alignments of deduced amino acid sequences in Clustal Omega bioinformatics tool (<http://www.ebi.ac.uk/clustalo/mega/index.html>) with a gap open penalty of 10.00 and a gap extension penalty of 0.01. Phylogenetic tree were constructed using neighbour joining statistical method and executed in MEGA program package (Version 6.06). The branch lengths are proportional to number of amino acid substitutions and represented the changing evolutionary lineages over time. Gonnet protein weight matrix was used to generate the distances between consensus coding regions of *Psy* genes to compare with other selected monocot and dicot sequences. The tree was bootstrapped to 1000 replicates to assess the support of individual branches and bootstrapping values (in percentage) are shown in each nodes of the rooted rectangular cladogram (Fig. 5). The sequence structure, in-vivo biological function and conserved domains present in *MaPsy* amino acid sequences were predicted using NCBI-CDD Database (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). The localisation signals present in N and C terminal of the deduced peptide fragments were analysed using TargetP 1.1 from CBS bioinformatics tools (<http://www.cbs.dtu.dk/services/TargetP/>). The location assignment is based on the presence of any predicted N-terminal pre-sequences such as chloroplast transit peptide (cTP), mitochondrial targeting peptide (mTP) and secretory pathway signal peptide (SP).

Results

Differential accumulation of provitamin A carotenoid in banana fruits

The total carotenoid content present in peel and pulp of the banana fruits from selected Indian cultivars is presented in Fig. 1. To our surprise, all twelve cultivars were found to contain more total carotenoid in non-edible portion of the banana fruits i.e. peel. Total carotenoid content ranged between 3.81 and 14.82 µg g⁻¹ in peel and 1.59–14.15 µg g⁻¹ in pulp of the banana fruits (Table S1). Among cultivars, *Monthan* (ABB) showed the maximum amount (14.8 µg g⁻¹) of total carotenoid content in non-edible peel. However, South India's famous cultivar *Nendran* (AAB) accumulated maximum total carotenoid (14.1 µg g⁻¹) in edible pulp portion of the banana fruit. Cultivar *Nendran* also had fairly high amount of total carotenoid content (13.3 µg g⁻¹) in non edible peel.

Fig. 1 Total carotenoid content ($\mu\text{g g}^{-1}$) in peel and pulp of selected Indian banana fruits

Quantification of β -carotene using HPLC revealed that except *Nendran* cultivar, all other cultivars were found to contain 2–3 times more β -carotene in peel than in pulp of the banana fruits (Fig. 2). *Nendran*, an orange fleshed cultivar alone showed more β -carotene content in edible pulp ($13.6 \mu\text{g g}^{-1}$) of the banana fruit than in non-edible peel ($12.5 \mu\text{g g}^{-1}$). Next to *Nendran*, the popular plantain cultivar *Monthan* recorded the second highest β -carotene content ($11.2 \mu\text{g g}^{-1}$) in peel. Among the yellow fleshed cultivars, *Red banana* and *Karpuravalli* showed 7.8 and $7.3 \mu\text{g g}^{-1}$ β -carotene content in peel, respectively. However, β -carotene content in their pulp was quite low (3.05 ; $0.19 \mu\text{g g}^{-1}$). The white fleshed cultivars *Kunnan* and *FHIA-3* contained the lowest β -carotene ($0.09 \mu\text{g g}^{-1}$) content in edible pulp.

Relationship between *MaPsy* gene and carotenoid accumulation

Isolation and molecular characterisation of *MaPsy* gene

Phytoene synthase (Psy) is a carotenoid biosynthesis enzyme, which catalyzes the condensation reaction of two molecules of geranylgeranyl diphosphate to produce a colorless phytoene, a precursor of β -carotene. The DNA fragment of expected size 390 base pair (bp) encoding Psy enzyme was successfully amplified from Cavendish banana fruit using degenerate primers designed from *Oncidium* orchid Psy orthologs (Fig. 3). Then this was the first report of *Musa accuminata* Phytoene synthase gene (*MaPsy*) that encoded the rate limiting enzyme of carotenoid biosyn-

Fig. 2 β -carotene content ($\mu\text{g g}^{-1}$) in peel and pulp of selected Indian banana fruits

Fig. 3 Genomic DNA clone of *Musa accuminata* *Phytoene synthase* (*MaPsy*) gene from Cavendish banana fruit. K-1 Kb ladder, L-100 bp ladder

thetic pathway in banana fruits. The 192 bp coding exon sequence from *MaPsy* gene has been submitted to NCBI data base (gene accession no: ACT83530.1 dated August 21, 2009). The amplified fragment was subsequently eluted, purified and sequenced. The sequencing results showed that the cloned 390 bp partial DNA sequence shared a high level of similarity to *Psy* gene from other plants including Daffodil, Wheat, Sorghum, Oncidium and Maize (Fig. S1). The predicted amino acid sequence revealed that open reading frame contained 66 amino acid without any stop codons from 190 to 389 bp. Comparison of amino acid sequences with other related plants inferred that, about 64 amino acids were conserved (79–86%) and belonged to the isoprenoid biosynthesis super family gene called *phytoene synthase* (Fig. 4).

The phylogenetic tree constructed using neighbour-joining statistical method to study the evolutionary relationship of *Psy* amino acid sequences from banana and other crop plants is represented in Fig. 5. The phylogenetic tree analysis showed that, carrot *Psy* formed a root cluster as ultimate common ancestor of all *Psy* analysed and all other cluster branches built upon with internal and external nodes. The average evolutionary distance calculated using poisson correction method was 0.02 predicting an average of 2 amino acid substitution per 100 nucleotide sites analysed. Cladogram showed that monocot and dicot *Psy*

sequences are clustered separately with sum of branch length of 1.31 revealing the greater distances between them. Within these two subclades, *Psy* from banana clustered together along with Oncidium and Daffodil suggesting that banana *Psy* exists as ancestor prior to all other monocot *Psy*. The branch support of the tree was high as re-sampling test values ranged up to 99% of replicate trees in which the associated taxa clustered together.

Expression analysis of MaPsy gene using semi quantitative reverse-transcription PCR

We investigated the expression pattern of *MaPsy* gene in peel and pulp of four selected Indian banana cultivar (Red banana, *Monthan*, Cavendish and *Nendran*) fruit. The transcripts of banana *Psy* were detected in all the investigated cultivars (Fig. 6). The results clearly revealed that the non-edible peel contain two cDNA fragments (445 and 365 bp) encoding two paralogues of *MaPsy* gene whereas, in edible pulp a single cDNA fragment (445 bp) was amplified in Cavendish cultivar. Red Banana and *Monthan* also showed similar differential expression pattern in peel and pulp of the banana fruit. In contrast, *Nendran* banana cultivar was found to amplify two cDNA fragments (445 and 365 bp) both in peel and pulp of the banana fruits. The level of *MaPsy* transcripts present in peel and pulp also differed as peel contains three folds higher *MaPsy* expression than in edible pulp (Fig. 6).

Molecular characterisation of MaPsy transcript variants

The sequencing results of two cDNA fragments amplified from Cavendish banana peel clearly showed that these amplicons were two transcript variants differing by eighty nucleotides. Figure S2 depicted that the eighty nucleotide base pairs in 445 bp fragments were present between those two conserved domains. A similar result by ClustalW bioinformatics tool reconfirmed the presence of eighty base pair in 445 bp fragments and absence of the same in 365 bp fragment. NCBI-Blast results revealed that nucleotide sequence of both these partial cDNA fragments shared the query coverage of 77% with high level of homology (97%; Fig. S3). The global NCBI-Blastn analysis showed that the cloned 445 bp cDNA sequence shared 96% of homology to complete CDS sequences of *Musa accuminata PsyI* gene (Fig. S4). The local NCBI-Blastn analysis revealed that *MaPsyI* gene fragment shared a high level of nucleotide homology to other closely related plants including Daffodil (CAA55391.1), Oncidium (ACP27623.1), Maize (ACY70880.1), Wheat (ABS83385.1) and Sorghum (ACY70871.1) with 81, 79, 80, 79 and 78%, respectively (Fig. S5). The 365 bp

		250	260	270	280	290	300
Rice_ACY70856.1	229	GMRSDLRKTR	YKNFDELYMY	CYYVAGTVGL	MSVPVMGIAP	ESKATTESVY	SAALALGIAN
Wheat_ABS83385.1	236	GMRIDLKKAR	YKNFDELYMY	CYYVAGTVGL	MSVPVMGIAP	DSKATAESVY	GAALALGLAN
Maize_ACY70880.1	218	GMRSDLRKTR	YNNFDELYMY	CYYVAGTVGL	MSVPVMGIAT	ESKATTESVY	SAALALGIAN
Sorghum_ACY70871.1	227	GMRSDLRKTR	YNNFDELYMY	CYYVAGTVGL	MSVPVMGIAP	ESKATTESVY	SAALALGIAN
Oncidium_ACP27623.1	218	GMRMDLKKSR	YKNFDELYLY	CYYVAGTVGL	MSVPVMGIDP	ESDATTESVY	NAALSLGIAN
Daffodil_CAA55391.1	228	GMRMDLKKSR	YKNFDELYLY	CYYVAGTVGL	MSVPVMGIAP	ESLAEAESVY	NAALALGIAN
Banana_ACT83530.1	1	-----	-----	-----	-----	-----	-----
		310	320	330	340	350	360
Rice_ACY70856.1	289	QLTNILRDVG	EDARRGRIVL	PQDELAQAGL	SDEDIFNKGVV	TNKRWSFMKR	QIKRARMFFE
Wheat_ABS83385.1	296	QLTNILRDVG	EDARRGRIVL	PQDELAQAGL	SDEDIFNKGVV	TDKWRKFMKR	QIKRARMFFE
Maize_ACY70880.1	278	QLTNILRDVG	EDARRGRIVL	PQDELAQAGL	SDEDIFNKGVV	TNRWRNFMKR	QIKRARMFFE
Sorghum_ACY70871.1	287	QLTNILRDVG	EDARRGRIVL	PQDELAQAGL	SDEDIFNKGVV	TNRWRNFMKR	QIKRARMFFE
Oncidium_ACP27623.1	278	QLTNILRDVG	EDARRGRIVL	PQDELAQAGL	SDEDIFNKGVV	TDRWRNFMKN	QIKRARTFFE
Daffodil_CAA55391.1	288	QLTNILRDVG	EDARRGRIVL	PQDELAQAGL	SDEDIFNKGVV	TDKWRKFMKR	QIKRARTFFE
Banana_ACT83530.1	1	-----	-----	-----	-----	-----	-----
		370	380	390	400	410	420
Rice_ACY70856.1	349	EAERGVTELS	QASRWPVWAS	LLLYRQILDE	IEANDYNNFT	KRAYVGKAKK	LLALPVAYGR
Wheat_ABS83385.1	356	EAERGVTELR	KESRWPVWAS	LLLYRQILDE	IEANDYNNFT	KRAYVGKAKK	VLALPVAYGR
Maize_ACY70880.1	338	EAERGVTELS	QASRWPVWAS	LLLYRQILDE	IEANDYNNFT	KRAYVGKAKK	LLALPVAYGR
Sorghum_ACY70871.1	347	EAERGVTELS	QASRWPVWAS	LLLYRQILDE	IEANDYNNFT	KRAYVGKAKK	LLALPVAYGR
Oncidium_ACP27623.1	338	EAERGVTELS	QASRWPVWAS	LLLYRQILDE	IEANDYNNFT	KRAYVGKAKK	LLALPVAYGR
Daffodil_CAA55391.1	348	EAERGVTELS	QASRWPVWAS	LLLYRQILDE	IEANDYNNFT	KRAYVGKAKK	LLALPVAYGR
Banana_ACT83530.1	49	EAERGVTELS	QASRWPVWAS	LLLYRQILDE	IEANDYNNFT	KRAYVGKAKK	LLALPVAYGR
		430					
Rice_ACY70856.1	409	SLLMPYSLRN	SQ---				
Wheat_ABS83385.1	416	SLLLPYSLRN	NQT---				
Maize_ACY70880.1	398	SLLLPYSLRN	GQT---				
Sorghum_ACY70871.1	407	SLLLPYSLRN	TQT---				
Oncidium_ACP27623.1	398	SLLIP-SMKK	PSLVKP				
Daffodil_CAA55391.1	408	SLLIPSLRP	PSLSKA				
Banana_ACT83530.1	64	-----	-----				

Fig. 4 Predicted amino acid sequence of the *Musa accuminata* *Phytoene synthase* (*MaPsy*) gene from Cavendish banana fruit. Encircled boxes showed the highly conserved amino acids sequences

present within the conserved domains like ARM-2 (Aspartate rich motifs-2) and Trans-Isoprenyl Diphosphate Synthases (*Trans_IPPS*)

Fig. 5 Phylogenetic analysis of *Musa accuminata* *Phytoene synthase* (*MaPsy*) gene from Cavendish banana fruit. The bootstrapping values (%) are shown at each node of the rooted rectangular cladogram. The average evolutionary distance is depicted in left bottom of the tree

Fig. 6 Semi quantitative RT-PCR analysis of *Musa accuminata* *Phytoene synthase* (*MaPsy*) gene from four selected Indian banana cultivars. Lane L. 1 Kb ladder, 1. Cavendish peel, 2. Cavendish pulp, 3. Red banana peel, 4. Red banana pulp, 5. Monthan peel, 6. Monthan pulp, 7. Nendran peel, 8. Nendran pulp

amplicon showed a 98% sequence homology to complete CDS sequences of *Musa accuminata* *Psy2* gene, which were constantly absent from edible portion of all the banana fruits analysed, except in *Nendran* cultivar (Fig. S6). This amplicon also shared 81, 79, 80, 80 and 80% nucleotide homology to Daffodil (CAA55391.1), Oncidium (ACP27623.1), Maize (AY324431), Wheat (ABS83385.1) and Sorghum (ACY70871.1), respectively (Fig. S7).

The EMBOSS in-silico transeq analysis showed that the open reading frame contained 118 and 70 translated amino acid sequence in 445 and 365 bp *MaPsy* transcripts, respectively (Figs. S8 and S9). The NCBI-Blastp analysis revealed that *MaPsy1* amino acid sequence shared a homology of 59–70% with other closely related plants species including *Arabidopsis* (Fig. S10). The results of *MaPsy2* amino acid sequence shared a homology of 82–96% with other plants including *chrysanthemum* (Fig. S11). The localisation signals present in N and C terminal of the deduced *MaPsy* peptide fragments predicted the presence of mitochondrial signal peptide in *MaPsy1* that expressed both in peel and pulp of the banana fruits (Fig. S12). The conserved domain search using NCBI-CDD Database revealed the presence of aspartate rich motif-2 (ARM-2) in both *MaPsy1* and *MaPsy2* translated amino acid sequences (Figs. S13 and S14). The results also showed the presence of a highly conserved domain known as Trans-Isoprenyl Disphosphate Synthases (Trans_IPPS) and squalene/phytoene synthase involved in carotenoid biosynthetic pathway.

Discussion

Among the major Indian banana cultivars, the edible portion of orange-fleshed banana cultivar *Nendran* (AAB) was reported to contain the highest β -carotene content. Earlier

studies, therefore, have focused mainly on the edible portion and showed a significant correlation between flesh colour (Orange > Yellow > White) and β -carotene content (Englberger et al. 2003, 2006). Arora et al. (2008) were the first to report the differential accumulation pattern and presence of more β -carotene in non-edible portion of banana fruits. Similar results observed in our experiment suggested that carotenoids play a versatile role as accessory pigments in protecting photosynthetic pigments and provide attractive colour to peel of the banana fruits. Generally, carotenoid content in banana fruit is influenced by the degree of maturation, type of soil, climatic conditions, storage conditions and genotype among other factors. Since, all the cultivars in this experiment were collected from the same germplasm conservation field at mature green stage, the variations in carotenoid content appear to be unaffected by the differences in the soil, climate and developmental status of the banana fruit. Among the genotypes, triploid cultivars (AAB) contained higher total carotenoid content compared to AAA and ABB triploids (Table S1). Diversity array analysis and mean carotenoid content of 42 banana accessions in Brazil also proved that the variation in carotenoid content among the banana cultivars was purely genotype-specific (Amorim et al. 2009). Analysis of the within-fruit, within-hand and within-plant as well as the between plant variations under standardized field conditions demonstrated that carotenoid content varied significantly across all genotypic groups (Darvin et al. 2011).

Our attempt to characterize the gene(s) controlling banana fruit flesh colour led to the cloning of *Psy* gene from Indian banana cultivars. Earlier, the cDNA encoding *Psy* enzyme was isolated and characterized from various monocot and dicot plants including *Arabidopsis*, maize, melon, carrot, citrus, daffodil, rice, pepper, marigold, soybean etc. (Li et al. 2008). The *MaPsy* gene shared a

high (up to 81%) nucleotide sequence identity with *Psy* from other crop plants confirming their association in carotenoid biosynthesis pathway. Phylogenetic tree analysis showed that banana, *Oncidium* and daffodil had common ancestors such as *Helianthus annuus* and *Tagetes erecta* (Salvini et al. 2005), *Arabidopsis* spp and *Brassica* spp (Fig. 5). Dendrogram analysis clearly showed that banana has evolved along with daffodil, closer in timeline with *Oncidium*, whereas monocots like wheat, maize and sorghum belong to a later time. Hence the distance between banana and these monocots were predicted to be wide. These results indicated that *MaPsy* gene can be potentially exploited for development of next generation Golden rice replacing *Psy* from daffodil and maize.

Further, working towards the elucidation of the mechanism behind higher accumulation of β -carotene in the peel led to cloning of two isoforms (*MaPsy1* and *MaPsy2*) of *MaPsy* gene. Analysis of gene structure and presence of **NILRDVGED** amino acid sequence in conserved domains like ARM-2 (Aspartate rich motifs-2) and Trans-Isoprenyl Diphosphate Synthases (Trans_IPPS) confirmed that these two *MaPsy* paralogs participate in carotenoids biosynthesis pathway (Fig. S13 and S14). Evidently, the Trans_IPPS motif present in phytoene synthases enzyme catalyze a head-to-head (HH) ($1'-1$) condensation reaction of two molecules of 20-carbon geranylgeranyl diphosphates. The catalytic site consists of a large central cavity formed by mostly anti-parallel α helices with two aspartate-rich regions (DXXXD) located on opposite walls. These residues are an in vivo substrate binding site for both Mg^{2+} and phosphate groups of prenyl donor via magnesium salt bridge. The differential expression of two isoforms in banana fruit indicated that *MaPsy* gene is responsible for tissue-specific accumulation of β -carotene. The two–three fold increase in the expression of the *MaPsy* gene in peel than in pulp showed that transcriptional regulation is involved in carotenoid biosynthesis of banana fruits. Interestingly, these two *MaPsy* paralogs exist in all four banana cultivars, which have been selected for preliminary gene analysis with different A, B, T genome and fruit provitamin A content (pVAC). This suggested that the mRNA abundance profile of both genes were similar in all the cultivars. Similar results in melon showed that *Psy* transcript was present at low quantities in unripe (green) fruit and reached its highest levels when the fruit turned orange (Karvouni et al. 1995). In citrus, the transcripts of *Psy* were markedly increased in both peel and juice sacs (edible part) during fruit ripening (Ikoma et al. 2016). *LePsy1*, a tomato fruit specific phytoene synthase was shown to be significantly increased at the breaker stage (Fraser et al. 1999).

Both isoforms of *MaPsy* gene shared a very high sequence homology (> 96%) with recently characterized

MaPsy1 and *MaPsy2* paralogs of Asupina banana cultivars in Australia (Mlalazi et al. 2012). Pairwise sequence alignment reconfirmed that amplicon present both in peel and pulp of the banana fruits were highly similar to *Psy1* from banana and tomato (Fig. S15). The peel specific fragment (365 bp) was highly similar to *Psy2* from tomato and banana (Fig. S15). The presence of mitochondrial signal peptide in *MaPsy1* indicated its sub cellular localization and role in fruit ripening (Fig. S12). The deduced localization of *MaPsy2* is predicted to be chloroplast and involved in carotenoid biosynthesis in leaves. Thus, the increased accumulation of β -carotene in non edible portion of the banana fruits were confirmed due to constant expression of two isoforms of *MaPsy* gene viz., a fruit ripening related *MaPsy1* gene and a carotenoid biosynthesis related *MaPsy2* gene. The constant absence of *MaPsy2* gene expression in edible portion of the cream fleshed Cavendish banana fruit suggested that this gene might have been down-regulated during developmental stage, similar to the situation in tomato (Giuliano et al. 1993), pepper (Romer et al. 1993), muskmelon (Karvouni et al. 1995) and papaya (Skelton et al. 2006) *Psy* ortholog.

Conversely, orange-fleshed *Nendran* banana fruit contains high β -carotene in edible pulp since both *MaPsy1* and *MaPsy2* genes are expressed. Thus this study proved that the variation between the orange-fleshed and cream-fleshed banana cultivars depended on the expression of the carotenoid biosynthesis related *MaPsy2* gene in the edible portion of the banana fruits. These results showed that carotenoids biosynthesis in banana fruits are under both transcriptional and post-transcriptional regulation involving more than one *MaPsy* paralogs. Li et al. (2008) reported the presence of three *Psy* paralogs (*Psy1*, 2, 3) in the Poaceae family with tissue specificity of expression for sub-functionalization of cellular activities. To elucidate the mechanism of gene regulation at the post-transcriptional level, full length cDNA sequencing and quantitative expression analysis of both *MaPsy1* and *MaPsy2* genes during both fruit developmental and ripening stages is required.

The isolation and further characterization of the two *MaPsy* genes from banana has potential applications in enhancing banana breeding and genetic improvement. Functional characterisation of the *MaPsy* genes will provide us insights into the intricacies of the carotenoid biosynthetic pathway in banana. The *Psy* gene has been considered to be one of the primary targets for genetic modification of plant carotenoids. Attempts to genetically modify the carotenoid content in transgenic plants have produced mixed results in the past. The development of genetically modified Golden Rice 2 was reported (Paine et al. 2005) and the carotene content in this second generation plants has been much improved compared to the first generation of Golden Rice (Ye et al. 2000). Therefore, *MaPsy* can also be tested as a

source of β -carotene production in generating transgenics, which could lead to higher accumulation of carotene. Also, *MaPsy2* can be introgressed/re-engineered into fruits of banana cultivars with poor β -carotenes content in edible pulp such as Cavendish banana, which are predominant in the Indian markets.

It is also worth noting, highest provitamin A activity in peel (*Monthan*) and pulp (*Nendran*) were observed in plantain cooking banana cultivars. As these fruits are consumed oil fried and its consumption increases the bioavailability of fat-soluble β -carotene (Ghavami et al. 2012), it would be far more advantageous if plantain banana fruits are promoted as cost-effective source of vitamin A. It is highly imperative here to record the trend of consuming oil fried *Nendran* banana chips/bujji in daily cuisine of people from Kerala, where nil occurrences of bitot spot (VAD) were reported by National Institute of Nutrition, Hyderabad. Nevertheless, the orange fleshed *Nendran* cultivar could be potentially exploited as best source of β -carotene for children suffering from VAD under the age of 5 years in developing countries like India.

Conclusion

MaPsy transcripts were detected both in the peel and pulp of the banana fruit. The expression of *MaPsy* showed a differential pattern in the peel and pulp suggesting *MaPsy* to be a regulatory enzyme in carotenoid accumulation in banana fruit. Based on the sequence comparison both at nucleic acid and deduced amino acid level, the clones obtained indeed code for banana *Psy* as evidenced by their high identity to other plant *Psy* genes. These findings clearly demonstrate the prerequisite of *MaPsy2* gene expression for increased accumulation of β -carotenes in orange fleshed *Nendran* banana cultivar. Further, *MaPsy2* paralogue gene is expressed in a tissue specific differential manner in edible and non edible portion of the banana fruits. Since peel contains both (*MaPsy1* and *MaPsy2*) the isoforms of *MaPsy* gene, it accumulated more β -carotene.

References

- Akhtar S, Ahmed A, Randhawa MA, Atukorala S, Arlappa N, Ismail T, Ali Z (2013) Prevalence of vitamin A deficiency in South Asia: causes, outcomes, and possible remedies. *J Health Popul Nutr* 31(4):413–423
- Amorim EP, Vilarinhos AD, Cohen KO, Amorim VBO, Santos-Serejo JA, Silva SO, Pestana KN, Santos VJ, Paes NS, Monte DC, Reis RV (2009) Genetic diversity of carotenoid-rich bananas evaluated by diversity arrays technology (DART). *Genet Mol Biol* 32:234–245
- Arora A, Choudhary D, Agarwal G, Singh VP (2008) Compositional variation in β -carotene content, carbohydrate and antioxidant enzymes in selected banana cultivars. *Int J Food Sci Technol* 43(11):1913–1921
- Asif MH, Dhawan P, Nath P (2000) A simple procedure for the isolation of high quality RNA from ripening banana fruit. *Plant Mol Biol Report* 18:109–115
- Boeing H, Bechthold A, Bub A, Ellinger S, Haller D, Kroke A, Leschik-Bonnet E, Müller MJ, Oberritter H, Schulze M, Stehle P, Watzl B (2012) Critical review: vegetables and fruit in the prevention of chronic diseases. *Eur J Nutr* 51:637–663
- Borah PK, Gogoi P, Phukan AC, Mahanta J (2006) Traditional medicine in the treatment of gastrointestinal diseases in upper Assam. *Indian J Tradit Know* 5(4):510–512
- Cazzonelli CI (2011) Carotenoids in nature: insights from plants and beyond. *Funct Plant Biol* 8:833–847
- Darvin ME, Albrecht SW, Lademann J, Vergou T (2011) The role of carotenoids in human skin. *Molecules* 16:10491–10506
- Englberger L, Darnton-Hill I, Coyne T, Fitzgerald MH, Marks GC (2003) Carotenoid-rich bananas: a potential food source for alleviating vitamin A deficiency. *Food Nutr Bull* 24:303–318
- Englberger L, Wills RB, Blades B, Dufficy L, Daniells JW, Coyne T (2006) Carotenoid content and flesh colour of selected banana cultivars growing in Australia. *Food Nutr Bull* 27(4):281–291
- FAOSTAT (2016) Food and Agriculture Organization statistical databases, Available at <http://faostat.fao.org/site/339/default.aspx>. Accessed 26 May 2016
- Fraser PD, Kiano JW, Truesdale MR, Schuch W, Bramley PM (1999) Phytoene synthase-2 enzyme activity in tomato does not contribute to carotenoid biosynthesis in ripening fruit. *Plant Mol Biol* 40:687–698
- Ghavami A, Coward WA, Bluck LJ (2012) The effect of food preparation on the bioavailability of carotenoids from carrots using intrinsic labelling. *Br J Nutr* 107(9):1350–1366
- Giuliano G, Bartley GE, Scolnik PA (1993) Regulation of carotenoid biosynthesis during tomato development. *Plant Cell* 5:379–387
- Ikoma Y, Matsumoto H, Kato M (2016) Diversity in the carotenoid profiles and the expression of genes related to carotenoid accumulation among citrus genotypes. *Breed Sci* 66:139–147
- Karvouni Z, John I, Taylor JE, Watson CF, Turner AJ, Grierson D (1995) Isolation and characterisation of a melon cDNA clone encoding phytoene synthase. *Plant Mol Biol* 27:1153–1162
- Li F, Vallabhaneni R, Wurtzel ET (2008) PSY3, A new member of the phytoene synthase gene family conserved in the Poaceae and regulator of abiotic stress-induced root carotenogenesis. *Plant Physiol* 146:1333–1345
- Mlalazi B, Welsch R, Namanya P, Khanna H, Geijskes RJ, Harrison MD, Harding R, Dale JL, Bateson M (2012) Isolation and functional characterisation of banana phytoene synthase genes as potential cisgenes. *Planta* 236:1585–1598
- Mohapatra D, Mishra S, Sutar N (2010) Banana and its by-product utilisation: an overview. *J Sci Ind Res* 69:323–329
- Nisar N, Li L, Lu S, Khin NC, Pogson BJ (2015) Carotenoid metabolism in plants. *Plant Metab Synth Biol* 8(1):68–82
- Paine JA, Shipton A, Chaggar S, Howells RM, Kennedy MJ, Vernon G, Wright SY, Hinchliffe E, Adams JL, Silverstones AL, Drake R (2005) Improving the nutritional value of golden rice through increased provitamin A content. *Nat Biotechnol* 23:482–487
- Peach K, Tracy M (1995) Rapid and accurate method for estimation of beta-carotene. *Methods of plant analysis*. Springer, Berlin, pp 308–314
- Pinazo-Duran MD, Gallego-Pinazo R, Garcia-Medina JJ, Zanon-Moreno V, Nucci C, Dolz-Marco R, Martinez-Castillo S, Galbis-Estrada C, Marco-Ramirez C, Lopez-Galvez MI, Galarreta DJ, Diaz-Llopis M (2014) Oxidative stress and its downstream signalling in aging eyes. *Clin Interv Aging* 9:637–652

- Purkayastha J, Nath SC (2006) Biological activities of ethanomedicinal claims of some plant species of Assam. *Indian J Tradit Know* 5(2):229–236
- Romer S, Huguency P, Bouvier F, Camara B, Kuntz M (1993) Expression of the genes encoding the early carotenoid biosynthetic enzymes in *Capsicum annuum*. *Biochem Biophys Res Commun* 196:1414–1421
- Salvini M, Bernini A, Fambrini M, Pugliesi C (2005) cDNA cloning and expression of the phytoene synthase gene in sunflower. *J Plant Physiol* 162:479–484
- Skelton RL, Yu Q, Srinivasan R, Manshardt R, Moore PH, Ming R (2006) Tissue differential expression of lycopene β -cyclase gene in papaya. *Cell Res* 16:731–739
- Smolikova GN, Medvedev SS (2015) Seed carotenoids: synthesis, diversity, and functions. *Russ J Plant Physiol* 62(1):1–13
- Ye X, Babili SA, Kloti 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
- Yuan H, Zhang J, Nageswaran D, Li L (2015) Carotenoid metabolism and regulation in horticultural crops. *Hort Res* 2:15036. doi:[10.1038/hortres.2015.36](https://doi.org/10.1038/hortres.2015.36)