GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

In silico identification and analysis of phytoene synthase genes in plants

Y. Han[1]*, Q.S. Zheng[2]*, Y.P. Wei[3,4]*, J. Chen[1], R. Liu[2,4] and H.J. Wan[4]

[1]Suzhou Polytechnic Institute of Agriculture, Suzhou, China
[2]College of Resources and Environmental Science,
Nanjing Agricultural University, Nanjing, Jiangsu, China
[3]College of Horticulture, Nanjing Agricultural University, Nanjing, China
[4]Institute of Vegetables, Zhejiang Academy of Agricultural Sciences,
Hangzhou, China

*These authors contributed equally to this study.
Corresponding author: H.J. Wan
E-mail: wanhongjian@sina.com

Genet. Mol. Res. 14 (3): 9412-9422 (2015)
Received December 16, 2014
Accepted May 26, 2015
Published August 14, 2015
DOI http://dx.doi.org/10.4238/2015.August.14.5

ABSTRACT. In this study, we examined phytoene synthetase (PSY), the first key limiting enzyme in the synthesis of carotenoids and catalyzing the formation of geranylgeranyl pyrophosphate in terpenoid biosynthesis. We used known amino acid sequences of the PSY gene in tomato plants to conduct a genome-wide search and identify putative candidates in 34 sequenced plants. A total of 101 homologous genes were identified. Phylogenetic analysis revealed that PSY evolved independently in algae as well as monocotyledonous and dicotyledonous plants. Our results showed that the amino acid structures exhibited 5 motifs (motifs 1 to 5) in algae and those in higher plants were highly conserved. The PSY gene structures showed that the number of intron in algae varied widely, while the number of introns in higher plants was 4 to 5. Identification of PSY genes in plants and the analysis of the gene structure may provide a theoretical basis for studying evolutionary relationships in future analyses.

Key words: Carotenoids; Evolution; Phytoene synthase

INTRODUCTION

A carotenoid is a generic term for a type of natural pigment and is an important class of lipid soluble antioxidants in organisms. In nature, there are more than 750 different carotenoids, which have different structures and are distributed in photosynthetic nutritional organisms, including terrestrial plants, algae, cyanobacteria, and photosynthetic bacteria, and are present in non-photosynthetic nutritional organisms, including photosynthetic eubacteria, archaea, and animals (Takaichi, 2013). In animals, various structural types of carotenoids can form through different synthesis pathways and play an important biological role in organisms. In plants, carotenoids play important regulatory roles in optical system assembly, light harvesting and protection, photo-morphogenesis, non-photochemical quenching, lipid peroxidation, and seed dormancy and aging. Recent studies have found that the plant hormones abscisic acid and strigolactone are produced by the metabolic pathway of the carotenoids and may play an important role in the regulation of plant adversity stress and morphogenesis (Alder et al., 2008).

Phytoene synthase (PSY) is the first key enzyme in the biosynthesis pathway of carotenoids in plants. PSY catalyzes the condensation of 2 molecules of geranylgeranyl pyrophosphate, the first member of the carotenoid family-PSY, which regulates the synthesis and content of other members of the carotenoid family. Ray et al. (1988) first found the gene encoding the enzyme in tomato plant fruit. Subsequently, Kamoda and Saburi (1993) isolated the homologous gene of the PSY gene in tomato plants and referred to it as PSY2. Whole-genome sequencing in tomato plants showed that 3 genes encode PSY (gene IDs: Solyc03g031860, Solyc02g081330, and Solyc01g005940). In addition to tomato plants, the PSY gene has been cloned in other plants, including tobacco (Busch et al., 2002), *Arabidopsis thaliana* (Paine et al., 2005), rice (Welsch et al., 2008), corn, citrus (Zhang et al., 2009), carrot (Bowman et al., 2014), and unicellular algae (*Chlamydomonas reinhardtii*) (McCarthy et al., 2004). Recent studies have focused on obtaining specific carotenoids in some plant species or improving their contents in order to increase the economic benefits and nutritional value through genetic engineering. For example, researchers introduced the PSY gene of maize into the rice endosperm in order to increase the content of β-carotenoid to treat vitamin A deficiency in humans (Paine et al., 2005). Cai et al. (2014) found that the red phenotype was associated with the PSY gene in cotton plants.

In this study, the catalytic pathway of PSY, the first rate-limiting step of the synthesis of other members of the carotenoid family, was examined. Recent studies have mainly focused on the cloning and expression of PSY in different plants (Arango et al., 2010; Han et al., 2013; Bowman et al., 2014). While the genome sequencing of numerous plants has been completed, the identification of PSY genes and the analysis of the gene structure based on the whole genome has rarely been reported. In the present study, tBLASTn was performed using the protein coding sequences of the PSY genes from tomato plants as a query against the JGI database (http://genome.jgi-psf.org/). The phylogenetic relationships, conserved motifs, and gene structures of the obtained genes were analyzed.

MATERIAL AND METHODS

Retrieval and analysis of the PSY in plants

In this study, the amino acid sequences of the PSY gene (Solyc03g031860, Solyc02g081330, and Solyc01g005940) in tomato plants, which have been previously published, were used as target

sequences, and a BLASTp search (threshold values of E are set to 1e-10) of the sequenced plants was conducted on the whole genome level using the database of Phytozome v9.1 (http://www.phytozome.net/#url). Redundant genes were removed from the sequences obtained and structural domains were identified using the PFAM website (http://pfam.janelia.org/) to obtain candidate genes. The ExPASy online tool (http://web.expasy.org/protparam/) was utilized to predict the isoelectric points and molecular weights of the amino acid sequences of the candidate genes. The online tools Plant-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/), CELLO v.2.5 (http://cello.life.nctu.edu.tw/), and WoLF PSORT (http:www.genscript.com/psort/wolf_psort.html) were used to analyze the subcellular localization of all protein sequences. The results are listed in Table 1.

Sequence alignment and construction of the phylogenetic tree

Multiple alignments were conducted for the amino acid sequences obtained of PSY using Clustal X. The comparison results were analyzed using the MEGA 5.0 software, and the neighbor-joining method was used to construct a phylogenetic tree. Bootstrapping (1000 replicates) was used to evaluate the degree of support for a particular grouping pattern in the phylogenetic tree (Tamura et al., 2011).

Analysis of the conserved motif of plant PSY

The conserved motifs in the amino acid sequence of PSY were analyzed using the MEME online tools (http://meme-suite.org/tools/meme).

The parameters were set as follows: the maximum length of the conserved motif was 50; the smallest length of the conserved motif was 6; the highest number of the conserved motif was 10; and other parameters were set as default values.

Analysis of gene structure of plant PSY

The obtained coding sequences and genomic sequences of the PSY genes in the plants were introduced into the Gene Structure Display Server website (http://gsds.cbi.pku.edu.cn/) to construct the diagram of the exon-intron structure of the PSY genes.

RESULTS

Retrieval and basic characteristics of the phytoene gene

The amino acid sequences of PSY in tomato plants (gene ID: Solyc03g031860, Solyc02g081330, and Solyc01g005940) were used as target sequences, with the threshold set to 1e-10. Genome-wide searches were conducted for 34 plants that had been sequenced, including dicotyledon, monocotyledon, and algae. A total of 101 PSYs were obtained (Pfam: 00494), as shown in Table 1. Analysis of the subcellular localization demonstrated that the PSY enzymes were present in the chloroplasts of all species examined, and several enzymes were located in the mitochondria and pigment cells in plants. In addition, the protein sequences encoded by these genes contained 263 to 599 amino acids. Pavirv00047252m.g encoded 263 amino acids, which was the lowest number. The highest numbers of amino acids were encoded by Medtr5g090780 (599 amino acids). The molecular weight was 21.99 to 67.67 kDa.

The molecular weight of evm.TU.supercontig_119.76 was the lowest, with a value of 21.99 kDa. The molecular weight of Medtr3g083630 was the largest, with a value of 67.67 kDa. The isoelectric points ranged from 4.88 to 9.30. The isoelectric point of e_gw2.10.185.1 was the lowest, with a value of 4.88. The isoelectric point of Gorai.001G083700 was the highest, with a value of 9.30. Furthermore, the isoelectric point of nearly 90% of the PSY protein was greater than 7 in advanced plants, indicating their alkalinity. Additionally, more than 60% of the PSY proteins were found to be acidic in lower plants, including algae, as shown in Table 1.

Phylogenetic relationship analysis

Although PSY is a type of enzyme encoded by a small gene family, it is present in all plants and regulates the carotenoid biosynthetic pathway. To understand the phylogenetic relationships of PSY, a phylogenetic tree of 101 amino acid sequences of PSY from 34 plants, including algae, monocotyledons, and dicotyledons, was constructed (Figure 1). The results showed that all PSYs could be divided into 3 large groups (I, II, and III). Within these groups, branch I was further divided into 3 subgroups (A, B, and C). PSY from the monocotyledons and dicotyledons in the A and B subgroups were found to have evolved independently, and only the C subgroup contained 2 members of the monocotyledonous plants (Si029859 mg and Sobic.002G292600). In group II, the phylogenetic tree exhibited a multi-branched shape, and all members were from dicotyledons. Furthermore, most PSY genes contained orthologous genes in groups I and II. In dicotyledonous plants, multiple paralogous genes (cassava 4.1008121 mg and cassava4.1008056 mg, Eucgr.F02913 and Eucgr.F02914, Glyma02g40720 and Glyma14g39045, Lus10020729 g and Lus10029809 g, Lus10001416 g and Lus10001050 g, Glyma14g03500 and Glyma02g45270, Glyma18g13700 and Glyma08g41890, Potri.005G205800 and Potri.002G056800, Lus10039186 g and Lus10013753 g, Glyma18g00350 and Glyma11g36420, and Migut.O00865 and Migut.B03619) were observed. In group III, the PSY of *Physcomitrella patens* and algae were clustered together. Only 1 PSY gene was found in the algae, while multiple PSY genes were present in *P. patens*. Phylogenetic analysis showed that the PSY genes might have different evolutionary patterns in algae, monocotyledonous, and dicotyledonous plants.

Conserved motif analysis of PSY

In this study, we found that PSY proteins in plants contained the SQS-PSY domain (Pfam accession No. 00494). To further analyze the sequence conservation of the proteins in plants, conserved motif analysis was conducted on the obtained 101 PSY from 34 species using the MEME online tools, as shown in Figure 2. The results showed that PSY in the plants contained 10 types of conserved motifs. With the exception of PSY from papaya plants (evm.TU.superconting119.76), motifs 1 to 5 were found in all analyzed sequences of algae, moss, monocotyledons, and dicotyledons, suggesting that these motifs are highly conserved in the evolution of plants. In group I, motif 10 of both Potri.017G138900 and Eucgr.F02914 was absent, while motifs 6, 7, and 10 from LOC_Os06g51290 was also lost. The remaining members contained 10 conserved motifs. In branches II and III, a deletion of the N-terminal motif was noted.

Table 1. Identification and analysis of PSY gene family in plants.

Species	Gene ID	Protein localization	Protein length (AA)	(kDa)	pI
Solanum lycopersicum	Solyc03g031860	Chloroplast	412	46.62	8.91
	Solyc02g081330	Chloroplast	438	49.33	8.87
	Solyc01g005940	Chloroplast	384	44.58	8.18
Solanum tuberosum	PGSC0003DMG400024063	Chloroplast	412	46.48	8.44
	PGSC0003DMG400016721	Chloroplast	438	49.42	8.76
Mimulus guttatus	Migut.K00976	Chloroplast	416	46.89	8.86
	Migut.O00865	Chloroplast Mitochondria	387	44.52	8.30
	Migut.D02541	Chloroplast Mitochondria	387	44.54	8.64
Vitis vinifera	GSVIVG01035255001	Chloroplast	419	47.76	8.42
	GSVIVG01020828001	Chloroplast	303	34.71	8.20
	GSVIVG01025421001	Chloroplast	357	41.21	6.25
Zea mays	GRMZM2G300348	Chloroplast	414	46.92	9.07
	GRMZM2G149317	Chloroplast	402	45.07	8.75
Sorghum bicolor	Sobic.010G276400	Chloroplast	419	47.26	8.64
	Sobic.008G180800	Chloroplast	397	44.47	8.19
	Sobic.002G292600	Chloroplast	441	48.96	9.01
Setaria italica	Si006539m.g	Chloroplast	415	46.90	8.97
	Si021888m.g	Chloroplast	487	53.69	9.37
	Si029859m.g	Chloroplast	440	48.66	9.04
Panicum virgatum	Pavirv00023344m.g	Chloroplast	417	47.13	8.96
	Pavirv00037205m.g	Chloroplast	420	47.53	9.01
	Pavirv00047252m.g	Chloroplast	263	30.06	7.65
Oryza sativa	LOC_Os06g51290	Chloroplast	420	47.58	8.96
	LOC_Os12g43130	Chloroplast	398	44.73	8.64
Physcomitrella patens	Pp1s196_120V6	Chloroplast	432	48.25	8.43
	Pp1s134_62V6	Chloroplast	432	48.62	6.17
	Pp1s323_70V6	Chloroplast	437	48.91	5.97
Chlamydomonas reinhardtii	g1866	Chloroplast Mitochondria	382	43.53	8.59
Volvox carteri	Vocar20014501m.g	Chloroplast	393	44.84	9.20
Coccomyxa subellipsoidea C-171	estExt_fgenesh1_pm.C_10036	Chloroplast	347	39.48	7.61
Micromonas pusilla CCMP1547	MicpuC2.estExt_Genewise1Plus.C_30483	Chloroplast	433	48.47	6.23
*Micromonas pusilla CCMP155*1	MicpuC2.e_gw1.15.194.1	Chloroplast	356	36.89	5.86
Micromonas pusilla RCC301	e_gw2.10.185.1	Chloroplast Cytochrome	352	39.99	4.88
Micromonas pusilla RCC305	fgenesh2_pm.C_Chr_13000048	Chloroplast	342	37.10	6.56
Ostreococcus lucimarinus	fgenesh1_pm.C_Chr_5000056	Chloroplast Cytochrome	273	31.65	5.19
Manihot esculenta	cassava4.1_008121m.g	Chloroplast	429	48.16	9.07
	cassava4.1_008056m.g	Chloroplast	431	47.59	6.79
	cassava4.1_033101m.g	Chloroplast Mitochondria	391	44.99	7.05
Ricinus communis	28611.t000005	Chloroplast	428	48.03	8.86
	29835.t000015	Chloroplast Cytochrome	385	44.20	5.87
	29835.t000016	Chloroplast Cytochrome	370	42.69	6.60
Linum usitatissimum	Lus10001416.g	Chloroplast	407	45.55	8.59
	Lus10001050.g	Chloroplast	405	45.30	8.83
	Lus10020729.g	Chloroplast	377	42.50	9.16
	Lus10029809.g	Chloroplast	451	49.59	8.51
	Lus10013753.g	Chloroplast	388	44.56	6.07
	Lus10039186.g	Chloroplast	398	45.86	6.36
Populus trichocarpa	Potri.005G205800	Chloroplast	405	45.39	9.12
	Potri.002G056800	Chloroplast	406	45.46	9.10
	Potri.004G081500	Chloroplast	433	48.62	8.51
	Potri.017G138900	Chloroplast	301	34.36	8.92
	Potri.001G007700	Chloroplast Mitochondria Cytochrome	382	43.86	7.53

Continued on next page

Table 1. Continued.

Species	Gene ID	Protein localization	Protein length (AA)	(kDa)	pI
	Potri.003G218000	Chloroplast	384	44.26	8.33
Medicago truncatula	Medtr5g076620	Chloroplast	434	48.80	8.64
	Medtr3g083630	Chloroplast	387	44.69	7.00
	Medtr5g090780	Chloroplast Cytochrome	599	67.67	5.33
Phaseolus vulgaris	Phvul.008G241500	Chloroplast	435	49.27	8.59
	Phvul.006G024100	Chloroplast	395	44.32	8.99
	Phvul.008G195800	Chloroplast	399	44.86	8.60
	Phvul.001G268600	Chloroplast Mitochondria	388	44.83	9.23
Glycine max	Glyma02g40720	Chloroplast	436	48.88	6.79
	Glyma14g39045	Chloroplast	433	48.50	9.03
	Glyma18g13700	Chloroplast	396	44.55	8.91
	Glyma08g41890	Chloroplast	396	44.40	8.36
	Glyma14g03500	Chloroplast Mitochondria	400	45.05	8.56
	Glyma02g45270	Chloroplast	399	44.74	8.11
	Glyma18g00350	Chloroplast	384	44.37	9.26
	Glyma11g36420	Chloroplast	315	36.04	8.87
Cucumis sativus	Cucsa.173570	Chloroplast	421	47.50	6.79
	Cucsa.179860	Chloroplast	388	43.85	8.75
	Cucsa.153210	Chloroplast	300	35.07	5.50
Prunus persica	ppa005962m.g	Chloroplast	435	48.88	8.73
	ppa006596m.g	Chloroplast	404	45.70	8.35
	ppa014590m.g	Chloroplast	382	45.49	8.53
Fragaria vesca	gene31674-v1.0-hybrid	Chloroplast	428	48.02	5.86
	gene28765-v1.0-hybrid	Chloroplast	398	45.10	8.85
	gene24795-v1.0-hybrid	Chloroplast	380	43.66	6.18
Arabidopsis thaliana	AT5G17230	Chloroplast	437	49.22	9.25
Arabidopsis lyrata	941228	Chloroplast	421	47.34	9.07
Capsella rubella	Carubv10001034m.g	Chloroplast	423	47.54	9.09
Brassica rapa	Bra008569	Chloroplast	425	47.68	9.02
	Bra023603	Chloroplast	416	46.40	8.84
	Bra006391	Chloroplast	414	46.74	9.09
Thellungiella halophila	Thhalv10013608m.g	Chloroplast	431	48.55	9.35
Carica papaya	evm.TU.supercontig_119.76	Chloroplast	191	21.99	7.63
Gossypium raimondii	Gorai.001G083700	Chloroplast	418	47.19	9.30
	Gorai.006G009400	Chloroplast	389	44.16	8.69
	Gorai.012G039100	Chloroplast	356	40.72	8.71
	Gorai.010G126900	Chloroplast	380	43.62	7.96
Theobroma cacao	Thecc1EG017615	Chloroplast	544	61.56	9.12
	Thecc1EG001945	Chloroplast	394	44.52	8.76
	Thecc1EG040196	Chloroplast Mitochondria	380	43.59	8.85
Citrus sinensis	orange1.1g036368m.g	Chloroplast	378	43.42	5.85
	orange1.1g016696m.g	Chloroplast	384	43.90	6.39
Citrus clementina	Ciclev10011841m.g	Chloroplast	416	47.41	5.78
	Ciclev10018150m.g	Chloroplast	385	44.16	5.85
Eucalyptus grandis	Eucgr.F02913	Chloroplast	424	43.95	6.84
	Eucgr.A02030	Chloroplast	390	47.70	8.63
	Eucgr.F02914	Chloroplast	385	43.89	8.89
	Eucgr.B03619	Chloroplast	342	39.59	6.96

Structure analysis of the gene for PSY in plants

The genome and coding sequences of each PSY gene was obtained from the plant genome database Phytozome v9.1 (http://www.phytozome.net/#url) and used to draw the structural diagram of introns and exons using the Gene Structure Display Server (http://gsds.cbi.pku.edu.cn), as shown in Figure 3. The results showed that the Medtr5g090780 con-

tained 9 introns, which was the highest number. MicpuC2.estExt_Genewise1Plus.C_3048, e_gw2.10.185.1 and fgenesh1_pm.C_Chr_5000056 in the algae had the fewest introns. In addition, nearly 60% of the members (60) contained 5 introns in the 101 PSY genes analyzed, which were mainly distributed in the A subgroup and group II. Nearly 25% of the members contained 4 introns, which were mainly distributed in the B subgroup. Approximately 10% of the members contained 6 introns (AT5G17230, Thecc1EG017615, Potri.004G081500, GSVIVG0103525500, gene31674.v1.0-hybrid, Glyma11g36420, Eucgr.B036420, Eucgr.B03619, GSVIVG01025421001, and estExt_fgenesh1_pm.C_10036). Gene structure analysis of PSY in the plants showed that the number of introns in higher plants was typically 4 to 5, whereas the number of introns in the algae varied, with results mainly between 1 and 6.

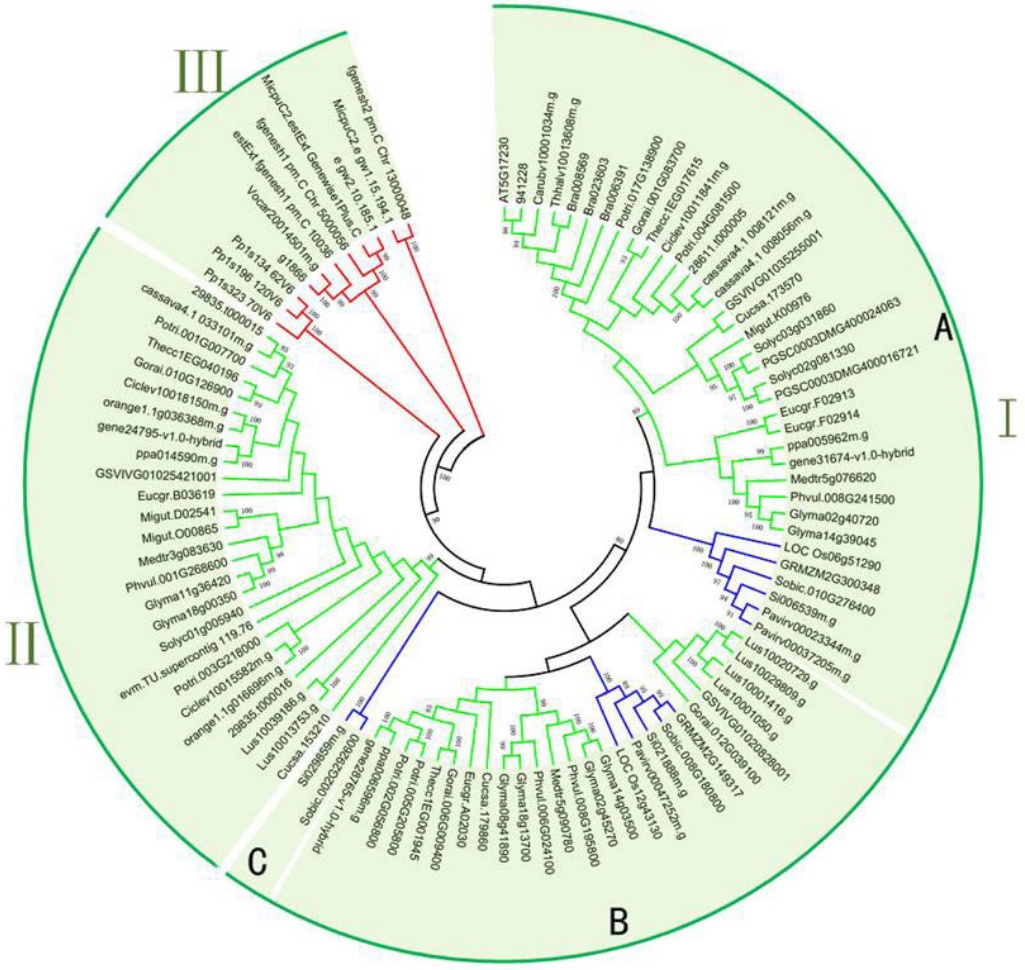

Figure 1. Phylogenetic tree of PSY proteins from 34 plant genomes.

Figure 2. Distribution of the conserved motif of phytoene synthase.

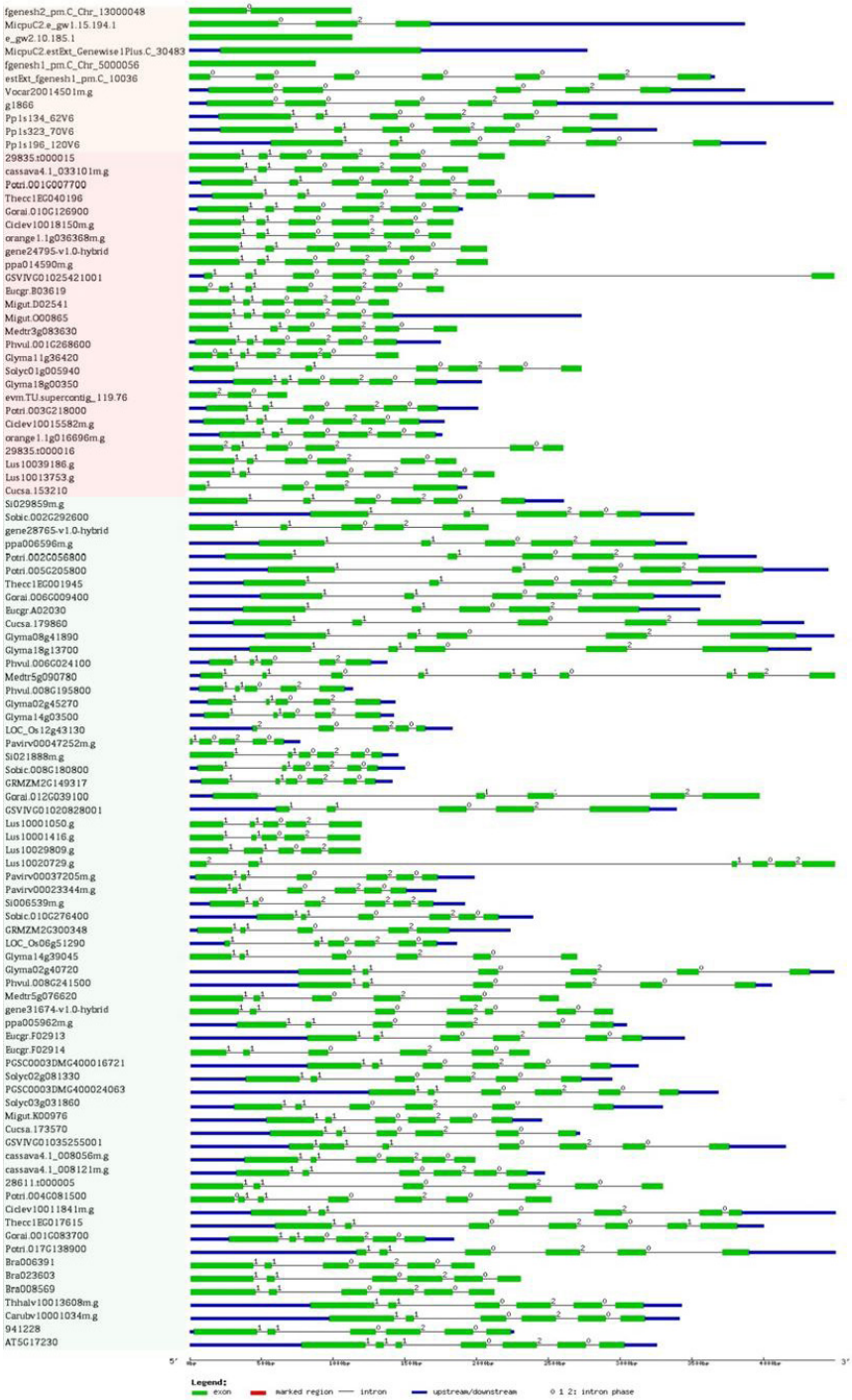

Figure 3. Structure diagram of intron/exon of the PSY gene in plants.

DISCUSSION

Carotenoids play an important role in the regulation of the human visual and immune systems, and carotenoids in the human body are mainly derived from food intake. Overexpression of the genes related to the carotenoid biosynthesis pathway in rapeseed (Shewmaker et al., 1999), rice endosperm (Ye et al., 2000), potato tuber (Ducreux et al., 2005), and soybean (Kim et al., 2012) can significantly increase the content of carotenoids in these crops. PSY is the first rate-limiting enzyme in the synthesis process of carotenoid production and the regulation metabolic of geranylgeranyl pyrophosphate. Previous studies demonstrated that the constitutive and specific expression of PSY could significantly improve the content of carotenoid in the tissues of higher plants, which has also become the basis of genetically modified carotenoid genes in plants (Shewmaker et al., 1999; Ye et al., 2000; Fraser et al., 2002; Ducreux et al., 2005; Kim et al., 2012).

In recent years, the development and utilization of new technologies has resulted in the generation of convenient molecular tools, including the widely used chip technology, high-throughput sequencing technology, and transgenic technology, for investigating the regulatory mechanisms and metabolic networks of the carotenoid biosynthesis pathway. Directional mutants produced by using targeted gene modified technology provide rich research materials for examining the function of PSY. However, recent studies have mainly focused on the cloning and expression of the PSY in a single plant strain or clone, using advanced tools and means of biological technology, while the analysis of PSY based on whole plants has been rarely reported. In addition to the gradual increase in the whole genome database of the plants that had been sequenced, it is important to analyze the evolutionary and genetic differences between different plants on the entire genome level.

In the present study, the retrieval of the genome database Phytozome v9.1 was conducted on the entire genome level in order to identify the homologous sequences. The phylogenetic relationship, conserved domain, and gene structure were analyzed. The results showed that although the plant PSY was encoded by a small gene family, it exists in a wide range of algae and higher plants. Of these, the gene copy numbers and gene structure were diverse. According to the phylogenetic tree, we found that the PSY had a separate evolutionary history in the monocotyledons, dicotyledons, and algae. Previous studies showed that the evolution of the PSY gene was independent. Although some algae possess 2 copies of the PSY gene (PSYI and PSYII), these were found to be missing in the evolutionary course of higher plants (Tran et al., 2009). Therefore, the varied copy numbers of the PSY genes in plants may have resulted from gene loss after duplication and the segmental duplication in the formation process of different species. In addition, the PSY enzymes were highly conserved in prokaryotes and eukaryotes, consisting of 6 conserved domains (substrate binding regions; the substrate-Mg^{2+} binding sites; catalytic residues; active site residues; and 2 aspartic acid rich regions) (Pandit et al., 2000; Marchler-Bauer et al., 2009). We predicted the conserved domains of the obtained amino acid sequence of the PSY genes, and the results showed that all sequences contained motifs 1 to 5, which were conserved domains of PSY.

In the plants studied, the PSY genes not only showed diversity in gene evolution and structure, but also tissue-specific expression and function evolution in different plants (Giorio et al., 2008; Stange et al., 2008; Arango et al., 2010; Fuentes et al., 2012). In the present study, we comprehensively analyzed the PSY gene family in plants on the entire genome level. Our results provide a foundation for understanding the evolutionary relationship and regulatory roles of PSY genes in the biosynthetic pathway of carotenoids.

ACKNOWLEDGMENTS

Research partially supported by the Program from the Natural Science Foundation of Suzhou City (#SYN201323), the Program 'San Xin Gong Cheng' from the Agriculture Committee of Jiangsu Province [#SXGC(2014)281], the Program from the Natural Science Foundation of Zhejiang Province (#LQ12C15004), and the Program from the National Natural Science Foundation of China (#31301774).

REFERENCES

Alder A, Holdermann I, Beyer P and Al-Babili S (2008). Carotenoid oxygenases involved in plant branching catalyse a highly specific conserved apocarotenoid cleavage reaction. *Biochem. J.* 416: 289-296.

Arango J, Wüst F, Beyer P and Welsch R (2010). Characterization of phytoene synthases from cassava and their involvement in abiotic stress-mediated responses. *Planta* 232: 1251-1262.

Bowman MJ, Willis DK and Simon PW (2014). Transcript abundance of phytoene synthase 1 and phytoene synthase 2 is associated with natural variation of storage root carotenoid pigmentation in carrot. *J. Amer. Soc. Horticul. Sci.* 139: 63-68.

Busch M, Seuter A and Hain R (2002). Functional analysis of the early steps of carotenoid biosynthesis in tobacco. *Plant Physiol.* 128: 439-453.

Cai C, Zhang X, Niu E, Zhao L, et al. (2014). GhPSY, a phytoene synthase gene, is related to the red plant phenotype in upland cotton (*Gossypium hirsutum* L.). *Mol. Biol. Rep.* 1-12.

Ducreux LJ, Morris WL, Hedley PE, Shepherd T, et al. (2005). Metabolic engineering of high carotenoid potato tubers containing enhanced levels of β-carotene and lutein. *J. Exp. Bot.* 56: 81-89.

Fraser PD, Romer S, Shipton CA, Mills PB, et al. (2002). Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proc. Nat. Acad. Sci. U. S. A.* 99: 1092-1097.

Fuentes P, Pizarro L, Moreno JC, Handford M, et al. (2012). Light-dependent changes in plastid differentiation influence carotenoid gene expression and accumulation in carrot roots. *Plant Mol. Biol.* 79:47-59.

Giorio G, Stigliani AL and D'Ambrosio C (2008). Phytoene synthase genes in tomato (*Solanumlycopersicum* L.) - new data on the structures, the deduced amino acid sequences and the expression patterns. *FEBS J.* 275: 527-535.

Han Y, Li L, Dong M, Yuan W, et al. (2013). cDNA cloning of the phytoene synthase (PSY) and expression analysis of PSY and carotenoid cleavage dioxygenase genes in *Osmanthus fragrans*. *Biologia* 68: 258-263.

Kamoda S and Saburi Y (1993). Cloning, expression and sequence analysis of a lignostilbene-a, b-dioxygenase gene from Pseudomonas paucimobilis TMY1009. *Biosci. Biotechnol. Biochem.* 57: 926-930.

Kim M-J, Kim JK, Kim HJ, Pak JH, et al. (2012). Genetic modification of the soybean to enhance the β-carotene content through seed-specific expression. *PloS One* 7: e48287.

Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, et al. (2009). CDD: specific functional annotation with the Conserved Domain Database. *Nucl. Acids Res.* 37: D205-D210.

McCarthy SS, Kobayashi MC and Niyogi KK (2004). White mutants of *Chlamydomonas reinhardtii* are defective in phytoene synthase. *Genetics* 168: 1249-1257.

Paine JA, Shipton CA, Chaggar S, Howells RM, et al. (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat. Biotechnol.* 23: 482-487.

Pandit J, Danley DE, Schulte GK, Mazzalupo S, et al. (2000). Crystal structure of human squalene synthase. A key enzyme in cholesterol biosynthesis. *J. Biol. Chem.* 275: 30610-30617.

Ray J, Bird C, Maunders M, Grierson D, et al. (1988). Sequence of pTOM5, a ripening related cDNA from tomato. *Nucl. Acids Res.* 15: 10587.

Shewmaker CK, Sheehy JA, Daley M, Colburn S, et al. (1999). Seed-specific overexpression of phytoene synthase: increase in carotenoids and other metabolic effects. *Plant J.* 20: 401-412.

Stange C, Fuentes P, Handford M and Pizarro L (2008). Daucus carota as a novel model to evaluate the effect of light on carotenogenic gene expression. *Biol. Res.* 41: 289-301.

Takaichi S (2013). Tetraterpenes: Carotenoids. *Nat. Prod.* 3251-3283.

Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.

Tran D, Haven J, Qiu W-G and Polle JE (2009). An update on carotenoid biosynthesis in algae: phylogenetic evidence for the existence of two classes of phytoene synthase. *Planta* 229: 723-729.

Welsch R, Wüst F, Bär C, Al-Babili S, et al. (2008). A third phytoene synthase is devoted to abiotic stress-induced abscisic acid formation in rice and defines functional diversification of phytoene synthase genes. *Plant Physiol.* 147: 367-380.

Ye X, Al-Babili S, Klöti A, Zhang J, et al. (2000). Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* 287: 303-305.

Zhang J, Tao N, Xu Q, Zhou W, et al. (2009). Functional characterization of *Citrus* PSY gene in Hongkong kumquat (*Fortunella hindsii* Swingle). *Plant Cell Rep.* 28: 1737-1746.