Gene Duplication in the Carotenoid Biosynthetic Pathway Preceded Evolution of the Grasses[1]

Cynthia E. Gallagher, Paul D. Matthews[2], Faqiang Li, and Eleanore T. Wurtzel*

Department of Biological Sciences, Lehman College, The City University of New York (CUNY), Bronx, New York 10468 (C.E.G., P.D.M., F.L., E.T.W.); and The Graduate School and University Center-CUNY, New York, New York 10016 (P.D.M., F.L., E.T.W.)

Despite ongoing research on carotenoid biosynthesis in model organisms, there is a paucity of information on pathway regulation operating in the grasses (Poaceae), which include plants of world-wide agronomic importance. As a result, efforts to either breed for or metabolically engineer improvements in carotenoid content or composition in cereal crops have led to unexpected results. In comparison to maize (*Zea mays*), rice (*Oryza sativa*) accumulates no endosperm carotenoids, despite having a functional pathway in chloroplasts. To better understand why these two related grasses differ in endosperm carotenoid content, we began to characterize genes encoding phytoene synthase (PSY), since this nuclear-encoded enzyme appeared to catalyze a rate-controlling step in the plastid-localized biosynthetic pathway. The enzyme had been previously associated with the maize *Y1* locus thought to be the only functional gene controlling PSY accumulation, though function of the *Y1* gene product had never been demonstrated. We show that both maize and rice possess and express products from duplicate PSY genes, *PSY1* (*Y1*) and *PSY2*; *PSY1* transcript accumulation correlates with carotenoid-containing endosperm. Using a heterologous bacterial system, we demonstrate enzyme function of PSY1 and PSY2 that are largely conserved in sequence except for N- and C-terminal domains. By database mining and use of ortholog-specific universal PCR primers, we found that the *PSY* duplication is prevalent in at least eight subfamilies of the Poaceae, suggesting that this duplication event preceded evolution of the Poaceae. These findings will impact study of grass phylogeny and breeding of enhanced carotenoid content in an entire taxonomic group of plant crops critical for global food security.

Carotenoids, a class of over 600 structures derived from isoprenoids, are synthesized by all photosynthetic organisms, some bacteria, and fungi. In plants, carotenoids are essential for plant growth and development; mutations blocking carotenoid accumulation have pleiotropic effects on chloroplast biogenesis and seed development (Robertson et al., 1978; Wurtzel, 1992). Carotenoids function as accessory pigments in photosynthesis, as photoprotectors preventing photooxidative damage, and as precursors to the plant hormone, abscisic acid (Hirschberg, 2001). The presence of carotenoids in plant endosperm tissue adds nutritional value; in humans and animals, dietary carotenoids are essential precursors to vitamin A and to retinoid compounds needed in development (Lee et al., 1981; Bendich and Olson, 1989). Nonprovitamin A carotenoids, such as lycopene, lutein, zeaxanthin, and others, also play beneficial roles in human health (Giovannucci et al., 1995; Kohlmeier et al., 1997; Sommerburg et al., 1998; Krinsky et al., 2003). The various roles of carotenoids affecting plant yield and nutritional potential has made them targets for breeding and metabolic engineering (Shewmaker et al., 1999; Matthews and Wurtzel, 2000; Ye et al., 2000; Davison, 2002; Blott et al., 2003; Gallagher et al., 2003).

In plants, the biosynthesis of carotenoids occurs on membranes of chloroplasts, chromoplasts, and amyloplasts, genetically identical plastids of very different internal membrane architecture. The plant enzymes, which are for the most part well established, are encoded in the nucleus and targeted to the plastids. Despite ongoing research on carotenoid biosynthesis in model organisms or carotenoid accumulating flowers, there is a paucity of information on pathway regulation operating in plants of world-wide agronomic importance, most specifically in the grasses (Poaceae). As a result, efforts to either breed for or metabolically engineer improvements in carotenoid content or composition in cereal crops have led to unexpected results because of the insufficient understanding of how metabolon assembly and function are controlled in plastids of different membrane architectures (Ye et al., 2000).

The biosynthesis of all carotenoids begins with the formation of the 40-carbon backbone, phytoene, a step mediated by phytoene synthase (PSY; Cunningham and Gantt, 1998; Hirschberg, 2001). In maize (*Zea mays*) endosperm, carotenoid content positively correlates with the dosage of the PSY structural gene, *Y1* (Randolph and Hand, 1940; Buckner et al., 1996;

[1] This work was supported by the NIH (grant no. S06–GM08225), by PSCUNY, and by the Rockefeller Foundation International Rice Biotechnology Program.

[2] Present address: S.S. Steiner, Inc., 655 Madison Avenue, New York, NY 10021–8078.

* Corresponding author; e-mail etwlc@cunyvm.cuny.edu or wurtzel@lehman.cuny.edu; fax 718–960–7348, 718–960–8236.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.104.039818.

Palaisa et al., 2003). In comparison to maize, rice (*Oryza sativa*) accumulates no endosperm carotenoids, despite having a functional PSY in green tissue. To better understand why these two grasses, representatives of two different subfamilies of the Poaceae, differ in endosperm carotenoid content, we began to characterize the genes encoding PSY, since this enzyme appeared to catalyze a rate-controlling step in the pathway (Bird et al., 1991; Bramley et al., 1992; Fray and Grierson, 1993; Giuliano et al., 1993; Kumagai et al., 1995; von Lintig et al., 1997). Previous cloning of the maize *Y1* locus, established this gene to encode PSY on the basis of sequence homology with other known phytoene synthase genes, though function of the gene product had never been demonstrated (Buckner and Robertson, 1993). Most plants have single genes encoding PSY and this was long thought to be true for maize and as a corollary, true for rice. We present evidence that both maize and rice possess duplicate *PSY* genes encoding structurally unique enzymes that function when tested in a heterologous bacterial system. Furthermore, the *PSY* duplication is prevalent throughout the grasses (Poaceae), suggesting that this genetic event preceded the evolution of the Poaceae.

RESULTS AND DISCUSSION

Isolation and Characterization of the Duplicate PSY Genes in Maize and Rice

In an effort to isolate the rice ortholog of maize *Y1*, a homologous rice EST (AY024350) was identified and used as a hybridization probe to isolate several rice genomic DNA bacterial artificial chromosome (BAC) clones. In Southern blots against rice genomic DNA, all of these clones shared the same pattern of hybridizing fragments, indicating they were the same gene. One genomic clone was sequenced and deposited as GenBank AY024351. Phylogenetic comparison of the deduced rice PSY against deduced peptides from all available PSY expressed sequence tags (ESTs), indicated that the rice PSY did not cluster with the deduced maize Y1 product, while another rice EST, AU082986, was the closest relative of maize *Y1* (Matthews, 2001). Although the novel rice gene encoded an apparently complete PSY protein and had an exon structure identical to maize *Y1*, cluster analyses together with genomic DNA hybridization patterns showed that the novel rice *PSY* was not the *Y1* ortholog. Either the rice gene shared sequence homology but did not encode a functional PSY, or as confirmed below, this gene represented a second but different, functional rice *PSY* gene, *PSY2*. With availability of the published rice genomic DNA sequence (http://portal.tmri.org/rice/), we identified the two different *PSY* genes, designated *PSY1* and *PSY2* on chromosomes 6 and 12, respectively, and through comparison to available cDNAs, annotated their exon/intron structures, as shown in Figure 1.

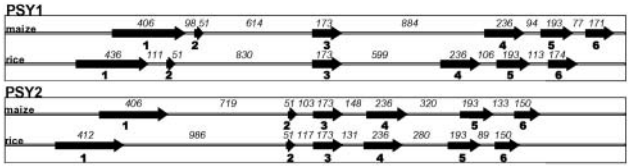


Figure 1. Gene structures of maize and rice *PSY1* and *PSY2*. Maize *PSY1* (GenBank AY324431); rice *PSY1* (GenBank AP005750); maize *PSY2* (GenBank AY325302); rice *PSY2* (GenBank AL831803). Arrows and thin bars indicate exons (bold numbers) and introns, respectively, sizes for which are in bp.

Was the PSY Gene Duplication Unique to Rice or Was It Also Present in Maize?

Prior maize mapping results identified both the *Y1* locus on chromosome 6 L (6.01) and a second locus, termed *psy2*, on chromosome 8 L (8.07; http://www.maizegdb.org/), suggesting that maize also contained duplicate *PSY* genes. Though there was no evidence that the second locus encoded a functional PSY enzyme, evidence for transcripts originating from two loci was obtained when we found maize ESTs in GenBank that showed homology to either rice *PSY1* or *PSY2*. We then screened a maize B73 genomic DNA BAC library and identified and sequenced both a *Y1* ortholog (denoted *PSY1*) and a maize ortholog for rice *PSY2*, for which gene structures are shown in Figure 1. Gene structures showed conservation across species; maize *PSY1* was more similar to rice *PSY1* than maize *PSY1* was to maize *PSY2* and the same relationship was seen for the rice genes. All four genes contain six exons and five introns. While the size of each of the six exons is conserved across all four genes, intron size is conserved only between orthologous pairs; *PSY1* genes have small first introns (approximately 100 bp), longer second and third introns (greater than 600 bp), while *PSY2* genes have long first introns (greater than 700 bp) and short second and third introns (approximately 100 bp). The possibility of the second gene, *PSY2*, being a pseudogene seemed less likely as suggested by the extensive gene structure conservation spanning two different Poaceae subfamilies. Moreover, there was evidence that all four genes were transcribed as indicated by the presence of ESTs in GenBank. At this point, we could not rule out the possibility that the transcripts might not all encode functional products.

Comparison of the Deduced Sequences for the PSY1 and PSY2 Proteins

The deduced protein sequences for PSY1 and PSY2 of maize and rice were determined and aligned (data not shown). Rice and maize PSY1 proteins shared 84.3% similar residues compared with 71.4% similar residues shared between maize PSY1 and maize PSY2. What most distinguished PSY1 from PSY2 were the distinct domains found at the carboxy termini; PSY1 proteins had a 15-residue domain distinct from the

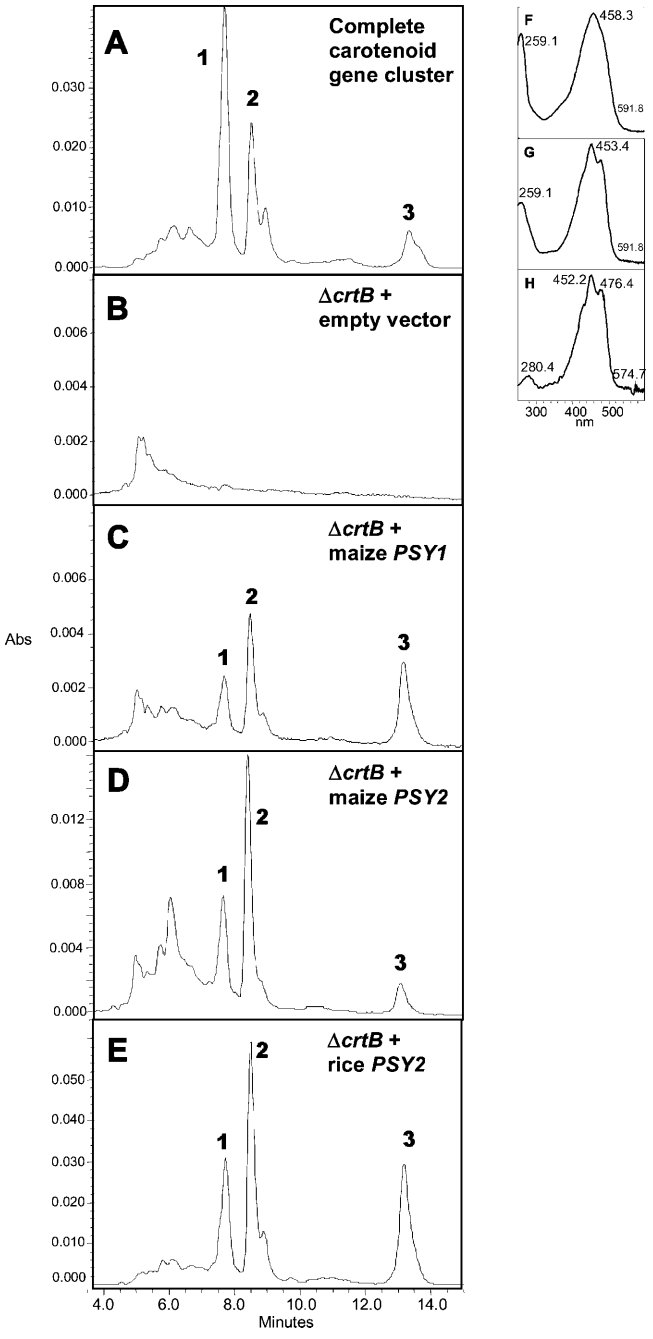


Figure 2. Functional complementation of PSY1 and PSY2. *E. coli* cells were transformed with: (A) pACCAR25; (B) pACCAR25ΔcrtB + pET23a (empty vector); (C) pACCAR25ΔcrtB + pEMPSY1-1; (D) pACCAR25ΔcrtB + pEMPSY2-1; and (E) pACCAR25ΔcrtB + pERPSY2-1 and extracted pigments analyzed by HPLC chromatograms shown at 450 nm. The spectral fine spectrum for the pathway end products, zeaxanthin β-D diglucoside (peak 1), zeaxanthin mono glucoside (peak 2), and zeaxanthin (peak 3) are shown in sections F, G, and H with retention times of 7.2 min, 8.1 min, and 13.1 min, respectively.

7-residue C-terminal domain in PSY2. Using the ChoroP Transit Peptide Predictor, all but rice PSY1 were predicted to have transit peptides (Emanuelsson et al., 1999). Maize PSY1 was predicted to be 46.5 kD (420 residues) having a 66-residue transit peptide and processed to a 39.8-kD (348 residues) mature plastid protein; rice PSY1 was predicted as 47.6 kD. Maize PSY2 was predicted as 45.2 kD (403 residues) with a 54-residue transit peptide and processed to a 39.5-kD (349 residues) mature protein; rice PSY2 was predicted as 44.7 kD (398 residues) with an 80-residue transit peptide and processed to a 36.2-kD (318 residues) mature protein. The N-terminal sequences, though not highly conserved, were more similar among the orthologs than between the paralogs.

Functional Testing of PSY1 and PSY2

To test whether the duplicated genes encoded potentially functional enzymes, we used a common tool for testing functionality of carotenoid biosynthetic enzymes (Matthews et al., 2003). Expression constructs were produced and cDNA gene products were transformed into *Escherichia coli* cells carrying a bacterial gene cluster for the entire pathway except for the gene encoding the bacterial counterpart of PSY (CrtB). Such cells produce pathway end products, zeaxanthin and its glycosylated derivates (Fig. 2, section A: peaks 3, 1, and 2, respectively), only when a functional PSY enzyme is present; these peaks are absent in the PSY deletion strain transformed with empty vector (section B) or with a truncated PSY2 construct (data not shown). When cDNAs encoding either maize PSY1 (section C), maize PSY2 (section D), or rice PSY2 (section E) were cotransformed along with the *crtB* deletion gene cluster, the expected products (peaks 1, 2, and 3) and matching spectra (sections F, G, and H) and retention times were observed, indicating that the *PSY1* and *PSY2* cDNAs tested encoded enzymes that were functional in the bacterial system.

Reverse Transcription-PCR to Test Expression of the PSY1 and PSY2 Transcripts in Maize and Rice Tissues

If either *PSY1* or *PSY2* transcript accumulation correlates with carotenoid accumulation, we would expect that maize and rice would vary specifically in endosperm transcript levels for one or both genes, given that maize endosperm accumulates carotenoids and rice endosperm does not. To test this possibility, RNA was extracted from young seedlings or endosperm from carotenoid-containing yellow maize, carotenoid-deficient white maize, and rice. Gene-specific primers, pretested on cDNAs to confirm specificity (Fig. 3A), were used to amplify the transcripts corresponding to the two PSY genes. As seen in Figure 3B, transcripts for both PSY genes were present in RNA extracted from either tissue only in the yellow maize endosperm line; for the white endosperm line,

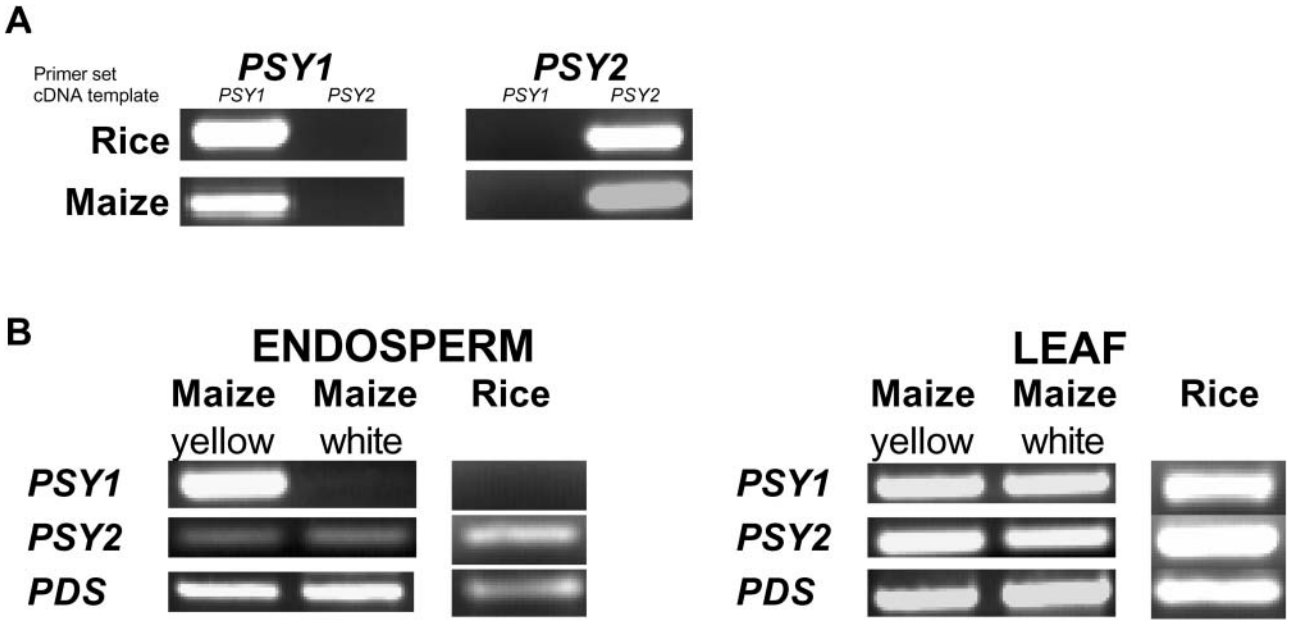


Figure 3. Leaf and endosperm transcript profiles for *PSY1* and *PSY2* in maize and rice tested by RT-PCR. A, Specificity of species- and gene-specific primers tested for both maize and rice. Left column, *PSY1* primers amplify only *PSY1* template and not *PSY2* template; right column, *PSY2* primers amplify only PSY2 template and not *PSY1* template. B, Specific transcripts for genes indicated were amplified by RT-PCR from leaf or endosperm mRNA of maize or rice. Yellow and white correspond to the endosperm phenotype of maize B73 and the maize *y1* mutant, respectively.

PSY1 transcripts were absent in endosperm, but both *PSY1* and *PSY2* transcripts were found in leaves. Similar to the white maize, *PSY1* transcripts were only present in rice leaves, but not in endosperm, while PSY2 transcripts were present in both tissues. In comparison, PDS transcripts were amplified and detected in both tissues for both plants regardless of endosperm phenotype (Li et al., 1996). These results indicate that carotenoid accumulation in endosperm correlates with expression of *PSY1* but not *PSY2* transcripts; expression of *PSY2* and *PDS* transcripts in rice endosperm is insufficient for carotenoid accumulation.

Duplication of the *PSY* Genes Preceded Evolution of the Poaceae

Since maize and rice belong to different subfamilies of the Poaceae, Panicoideae and Ehrhartoideae, respectively, it was likely that the gene duplication was more widespread in the Poaceae, a phenomenon proven correct by further GenBank database searching. *PSY* ESTs were identified for Triticum and Hordeum, species in another Poaceae subfamily, Pooideae. Together with the deduced protein sequences for the maize and rice PSY proteins, these additional sequences were compared with deduced PSY proteins of representative dicots and another monocot, Narcissus. The resulting phylogenetic tree seen in Figure 4 shows that for each of the grasses, sequences either cluster into PSY1-like or PSY2-like groups; the grass duplication appeared to have evolved from a common ancestor prior to the evolution of the grasses. In contrast, the dicot duplication seen for the tomato (Lycopersicon) PSY gene is not found in Arabidopsis, which has been fully sequenced, and generally not the rule for dicot taxa, the only other known exception being tobacco (*Nicotiana tabacum*; Bartley and Scolnik, 1993; Busch et al., 2002).

How Widespread within the Poaceae Is the Gene Duplication of *PSY*?

To test further for the distribution of the duplicated *PSY* genes among other Poaceae subfamilies, ortholog-specific universal primers were designed and tested for specificity. As seen in Figure 5A, we observed the expected *PSY1* products of 1,123 bp (maize), and not 387 bp that would be obtained if the corresponding region of *PSY2* was nonspecifically amplified; and 838 bp (rice), and not a nonspecific *PSY2* product of 370 bp. Similarly for *PSY2* universal primers, the expected products for maize and rice, 434 bp and 394 bp, respectively, were observed, and not the nonspecific amplification of *PSY1*, predicted to be 208 bp and 127 bp, respectively. These universal primers were then used to amplify DNA from representative taxa of the Poaceae. DNA sequences of the PCR amplification products, which have been deposited into GenBank, revealed that some amplified genes contained introns, while others did not. Therefore, we used only the exonic regions in a cluster analysis along with corresponding regions of either *PSY1* or *PSY2* of maize and rice to confirm that the amplified sequence clustered

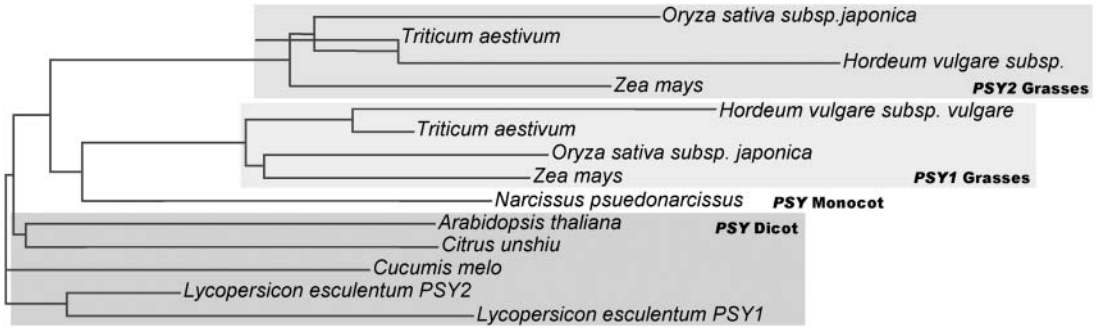
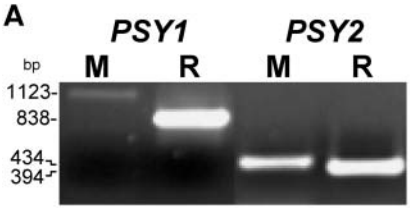


Figure 4. Phylogenic analysis of PSY amino acid sequences. SwissProt numbers, in bold, and GenBank accessions are in parentheses. *Lycopersicon esculentum* (**AAA34187**, L23424.1; J. Hirschberg, unpublished data), *L. esculentum* (**AAA34153.1**, M84744), *Cucumis melo* (**CAA85775.1**, Z37543), Arabidopsis (**AAAN17427**, BT000450.1), *Citrus unshiu* (**BAB18514**, AB037975.1), *Hordeum vulgare* subsp. *vulgare* (BI955682), *Triticum aestivum* (BM137086), *Z. mays* (AY325302), *O. sativa* (japonica, AL831803), *H. vulgare* (BE421261), *O. sativa* (japonica, AP005750), *Z. mays* (AY324431), *T. aestivum* (CD862515), and *Narcissus pseudonarcissus* (**CAA55391**, X78814).

with either *PSY1* or *PSY2* (data not shown). Using this simple PCR assay, we detected duplicate genes in 12 taxa representing 8 subfamilies in the Poaceae, as shown in Figure 5B. The use of universal *PSY* PCR primers will be valuable in assessing the distribution of the PSY duplication among the monocots. These tools will also be useful in phylogenetic analyses within the Poaceae subfamilies to offer improved resolution of evolutionary relationships.



B

Poaceae (Grass Family)			
Subfamily	**Tribe**	**Subtribe**	**Species**
Pharoideae	Phareae		*Pharus lappulaceus*
Bambusoideae	Bambuseae	Bambusinae	*Bambusa vulgaris*
Ehrhartoideae	Oryzeae	Zizaniinae	*Oryza sativa*
Pooideae	Triticeae	Hordeinae	*Hordeum muticum*
		Triticinae	*Secale cereale*
Aristidoideae	Aristideae		*Aristida adscensionis*
Arundinoideae	Arundineae		*Phragmites australis*
Centothecoideae	Centotheceae		*Zeugites pittieri*
Panicoideae	Paniceae	Cenchrinae	*Pennisetum tristachyum*
	Andropogoneae	Sorghinae	*Sorghastrum nutans*
		Tripsacinae	*Tripsacum zopilotense*
			Zea mays

Figure 5. Poaceae subfamily genomes tested and found to possess the PSY gene duplication. A, Testing of the ortholog-specific universal primers indicated at the top using as template, M (maize B73) or R (rice IR36) genomic DNA. B, Species representing 8 of the 12 subfamilies of the Poaceae family found to have both the *PSY1* and *PSY2* genes based on amplification using the ortholog-specific universal primers tested in A and further confirmed by alignment of the amplified exonic DNA sequences.

CONCLUSION

We have found that throughout the Poaceae, the gene for PSY is duplicated, suggesting that this duplication occurred prior to evolution of the grasses. Without evidence of gene product function, Palaisa et al. (2003) previously used associative genetics to correlate endosperm carotenoids with allelic states of maize *PSY1* (*Y1*) but not *PSY2* loci. Our data support that study and show that *PSY1* but not *PSY2* transcripts in endosperm correlate with endosperm carotenoid accumulation. Whereas prior studies did not address whether the genes encoded functional enzymes, we demonstrated that for maize, both *PSY1* and *PSY2* encode functional enzymes as tested in a bacterial system; rice PSY2 is also functional using this heterologous platform. However, in planta, function requires not only the potential for enzyme activity as demonstrated in the bacterial milieu, but also that the enzyme must localize to a plastid membrane where it gains access to substrate produced by an upstream enzyme. Similarly in Narcissus, where only one PSY gene has been described, PSY was found as an inactive soluble plastid stromal form and as an active plastid membrane-bound enzyme (Schledz et al., 1996). We provide data that suggest that only PSY1 seems to have both demonstrated activity in *E. coli* and to function in endosperm. While rice PSY2 is functional in *E. coli*, it is apparently not functional in rice endosperm. The transcript is translated in rice endosperm (data not shown) but the enzyme is not functional given the absence of carotenoids or carotenoid intermediates in rice endosperm (Burkhardt et al., 1997). Therefore, our data suggest that PSY1 and PSY2 are not functionally equivalent in planta and that endosperm carotenoid accumulation requires expression of *PSY1*. The duplicate grass genes are predicted to encode enzymes with variant N and C termini, suggesting that the grass PSYs may target to different plastid membranes. The difference in membrane architecture between endosperm amyloplasts and leaf chloroplasts may offer a possible explanation of why

different PSY isoforms may be associated with presence or absence of endosperm carotenoid accumulation. Further characterization of the two grass enzymes will be needed to define their roles in carotenogenesis in the variety of plastid architectures found in cereals and provide insight into the potential for endosperm carotenoid accumulation throughout the grasses.

The *PSY* duplication in the grasses predisposed evolution of tissue-specific pathway control, providing a mechanism to modify gene expression in the seed without deleterious effects on photosynthetic organs. Had there been only a single *PSY* gene, its overexpression would have interfered with the photosynthetic complex, most likely causing photosensitivity in the plant (Busch et al., 2002). The occurrence and persistence of PSY duplications suggests that recruitment of primary carotenoids as secondary metabolites has been adaptive in many species. The existence of parallel (convergent) PSY duplications among the monocots and the dicots, taken together with evidence of altered spatial expression of the gene product of one locus among photosynthetic and nonphotosynthetic organs, supports this supposition. While rice does not accumulate endosperm carotenoids, there are other grasses besides maize that do accumulate seed carotenoids, including sorghum, millet, and wheat (FAO, 1995). The existence of duplicate PSY factors in the grasses offers novel opportunities to use conventional breeding or biotechnology to select for enhanced endosperm carotenoids in grass species that are of agronomic importance.

MATERIALS AND METHODS

Plant Materials

Maize (*Zea mays*; Maize Genetic Stock Center, University of Illinois) was field-grown in Bronx, NY; rice (*Oryza sativa* indica variety) IR36 was greenhouse-grown with supplemental lighting. Maize endosperm dissected at 20 d after pollination, dissected mature rice endosperm and leaf samples from maize and rice were frozen in liquid nitrogen and stored at −80°C prior to use. For Poaceae subfamily *PSY* gene amplifications, DNA (Dr. Lynn Clark, Iowa State University) was obtained for *Bambusa vulgaris*, *Pharus lappulaceus*, and *Zeugites pittieri*, or prepared from dried leaves (Dr. Paul Peterson, Smithsonian Institution) for *Hordeum muticum* J. Presl. collected in Ayacucho, Peru, sample identification (ID): Peterson, P. M., Refulio-Rodriguez, N. 16440, *Secale cereale* collected in Maryland, ID: Pennington, S. J. 1200, *Aristida adscensionis* collected in Cajamarca, Peru, ID: Peterson, P.M., Refulio-Rodriguez, N. 15059, *Phragmites australis* (Cav.) Trin. ex Steud., ID: Peterson, P.M. 17519, *Pennisetum tristachyum* collected in Cajamarca, Peru, ID: Peterson, P.M. Refulio-Rodriguez, N. 15031, *Sorghastrum nutans* (L.) Nash collected in Mexico, ID: Peterson, P.M., Gonzalez-Elizondo, S., Brothers, L. E. 16684 and *Tripsacum zopilotense* Hernández & Randolph collected in Tamaulipas, Mexico, ID: Peterson, P. M., Valdes-Reyna, J. 15903.

Genomic DNA Isolation and Sequence Analysis

A maize B73 genomic BAC library containing 92,160 clones in pECBAC1 and representing 5.2× genome equivalents (Dr. H. Zhang, Texas A & M University) was probed with PSY cDNAs (maize *PSY1*; GenBank ZMU32636); rice *PSY2*, GenBank AY024350; Zhang et al., 1996). Five *PSY1* and three *PSY2* BAC clones were obtained and representatives chosen for further sequencing of both strands by primer walking (DNA Sequencing Facility, University of Chicago Research Center). Sequence assembly and analysis of these and all other DNA samples were performed using Vector NTI Suite, Version 7.0 (InforMax, North Bethesda, MD), and BLAST 2.1 (Altschul et al., 1997). Maize *PSY1* and *PSY2* genomic sequences were deposited as GenBank AY324431 and AY325302, respectively. For comparison, rice (japonica) *PSY1* and *PSY2* used in Figure 1 were AP005750 and AL831803, respectively.

Plasmids and Functional Complementation

Plasmid pACCAR25 (Misawa et al., 1990) contains the *Erwinia uredovora* gene cluster conferring accumulation of glycosylated zeaxanthin when transformed into *Escherichia coli* and was used as a positive control; pACCAR25Δ*crtB* (Chamovitz et al., 1992), containing a frame-shift mutation in *crtB* (bacterial PSY), was used for heterologous complementation to test function of *PSY1* and *PSY2* cDNAs subcloned as in-frame translational fusions as follows. A maize *PSY2* cDNA (nt no. 3 to nt no. 1,348) was amplified from pAY450646 (GenBank no. AY450646) using forward primer (no. 622) 5′-AAGAGATC**GAATTC**GGCACGG-3′ with an *Eco*RI site (bolded) and reverse primer (no. 635) 5′-TCCTTGTAA**CTCGAG**CTGATTGAG-3′, with an *Xho*I site (bolded), digested with *Eco*RI and *Xho*I, subcloned into the corresponding sites of pET23a (Novagen, Madison, WI), and renamed pEMPSY2-1. The rice *PSY2* cDNA pAY452768 (GenBank no. AY452768) was inserted as an *Eco*RI/*Xho*I fragment into corresponding sites of pET23b (Novagen, Madison, WI) and renamed pERPSY2-1. A maize PSY1 cDNA was cloned in-frame in pBluescript, p12A33A, (GenBank no. ZMU32636; Dr. Brent Buckner, Truman State University). *E. coli* BL21 (DE3) cells (Novagen) were transformed (Sambrook et al., 1989) with combinations of pACCAR25Δ*crtB* and the expression constructs p12A33A, pEMPSY2-1, or pERPSY2-1 or with pACCAR25 and pET23b. Transformants were grown in liquid Luria-Bertani medium with appropriate antibiotics overnight at 37°C with aeration and held at room temperature for 2 d in the dark, centrifuged at 2,000*g* for 30 min, and pellets extracted twice with acetone. Combined extracts were dried over Na_2SO_4, concentrated to dryness under a stream of nitrogen, resuspended in injection solvent (acetonitrile, 85%; methanol, 10%; dichloromethane, 2.5%; and hexane, 2.5%), and filtered through a 0.45-μ nylon filter (Phenomenex, Torrance, CA) into a 300-μL glass insert in a 2-dram amber vial and subjected to HPLC analysis.

HPLC Analysis

Carotenoids were separated on a Waters (Millipore, Franklin, MA) HPLC system with 2690 separation module, Millenium version 2.0 software (Waters, Franklin, MA), 996 photodiode array detector (Waters), 717 autosampler, using a Nucleosil 5 C_{18} (5μ, 250 × 4.6 mm) column (Phenomenex, Torrance, CA) with a Nucleosil C_{18} (5μ, 4 × 3.0 mm) guard column (Phenomenex). Solvent mixtures used for mobile phases were A, acetonitrile:methanol (9:1, v/v) and B, hexanes:methylene chloride:methanol (4.5:4.5:1, v/v/v; Khachik et al., 1999). Sample injection was followed by 10 min of isocratic conditions using 95% A:5% B, followed by a linear gradient to 45% A:55% B over 30 min. Between samples, columns were reequilibrated for 10 min using 95% A:5% B. All solvent flow rates were 0.7 mL/min. Carotenoids were identified by comparison of retention times and absorption pattern spectra with those of authentic standards.

Isolation of RNA and RT-PCR

Total RNA was isolated from maize leaves and 20 d after pollination endosperm using B73 (a yellow endosperm line) or *y1* (a white endosperm line), and from rice IR36 leaves and mature endosperm (RNeasy Plant Mini kit for total RNA isolation, Qiagen, Valencia, CA), concentrations measured spectrophotometrically, and 100 ng of RNA used as template for first strand cDNA synthesis (SuperScript First-Strand Synthesis system for RT-PCR, Invitrogen, Carlsbad, CA). A 2-μL aliquot of the first-strand reaction, 1 to 5 ng of cDNA, was used for PCR amplification under conditions pretested for linearity. Gene-specific primers were designed to flank introns and were tested for specificity using maize *PSY1* (GenBank ZMU32636), maize *PSY2* (GenBank AY450646), rice *PSY1* (pTRPSY1-1, GenBank AY445521), and rice *PSY2* (GenBank AY024350). PCR reactions contained 20 mM Tris-HCl pH 8.4, 50 mM KCl, 0.2 mM each dNTP, 1.5 mM $MgCl_2$, 0.4 μM each primer, 0.025 units/μL *Taq* DNA Polymerase (Invitrogen) and were carried out for one cycle of 3 min at 94°C; followed by 35 cycles of (30 s at 94°C; 30 s at annealing temperature, 45 s at 72°C); and one cycle of 10 min at 72°C. Gene-specific

primers, annealing temperatures, and expected products were: maize *PSY1*, forward (no. 509) 5′-GCATTGCTCAAACGCCAG-3′; reverse (no. 519) 5′-CAGAGAGAGCGGCATCAAG-3′, 54°C, 300 bp; maize *PSY2*, forward (no. 532) 5′-GCGGCAAGTTCCACCACCTGT-3′; reverse (no. 529) 5′-CGAGGTCTGCGCCGAGTA-3′, 58°C, 150 bp; rice *PSY2*, forward (no. 145) 5′-CCTGAAAGGCGCAAAGCTG-3′; reverse (no. 146) 5′-CGATAGCATCAAGGATCTGCC-3′, 65°C, 682 bp; rice *PDS*, forward (no. 151) 5′-GACCATGTTCGCTCTTTGGGTGG-3′; reverse (no. 152) 5′-CGATGATTTCAGTGTCACTCCGTCC-3′, 61°C, 430 bp; and maize PDS, as described previously (Matthews et al., 2003).

DNA Extraction and PCR Amplification of *PSY1* and *PSY2* from Poaceae Subfamilies

For PCR, genomic DNA from dried grass samples was extracted (REDExtract-N-Amp Plant PCR kit, Sigma, St. Louis) and 4 μL added to 10 μL of REDExtract-N-Amp PCR Ready Mix and 0.5-μM final concentration of each primer in a 20-μL reaction. Universal gene-specific were designed based on conserved sequences between maize and rice: *PSY1*, forward (no. 530) 5′-TTTGGACCGGTGGGAGAA-3′ and reverse (no. 520) 5′-GCCCATCACAGGTACGCTCATT-3′ (annealing temperature, 54°C); *PSY2*, forward (no. 672) 5′-GACGAATATTCTCAGAGACG-3′ and reverse (no. 673) 5′-ACTTTCCCTCTGAATATGTC-3′ (annealing temperature, 50°C). All reactions were as follows: one cycle of 3 min at 94°C; followed by 40 cycles of (30 s at 94°C; 30 s at annealing temperature, 1 min at 72°C); and one cycle of 10 min at 72°C and products purified using the Qiagen MinElute PCR Purification kit (Qiagen) prior to sequencing. The PSY1 GenBank accession numbers were CG892534 (*Phragmites australis*), CG892535 (*Aristida adscensionis*), CG892537 (*Hordeum muticum*), CG892538 (*Pennisetum tristachyum*), CG892539 (*Secale cereale*), CG892540 (*Sorghastrum nutans*), CG892541 (*Tripsacum zopilotense*), CG892543 (*Pharus lappulaceus*), CG892544 (*Zeugites pittieri*) and CG892545 (*Bambusa vulgaris*). The PSY2 GenBank accession numbers were CG892547 (*Phragmites australis*), CG892548 (*Aristida adscensionis*), CG892549 (*Hordeum muticum*), CG892550 (*Secale cereale*), CG892551 (*Sorghastrum nutans*), CG892552 (*Tripsacum zopilotense*), CG892553 (*Bambusa vulgaris*), CG892554 (*Pharus lappulaceus*), CG892555 (*Zeugites pittieri*), and CG892559 (*Pennisetum tristachyum*).

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under accession numbers AY024351, ZMU32636, AY024350, AY324431, AY325302, AP005750, AL831803, AY450646, AY452768, ZMU32636, AY450646, AY445521, AY024350, CG892534, CG892535, CG892537, CG892538, CG892539, CG892540, CG892541, CG892543, CG892544, CG892545, CG892547, CG892548, CG892549, CG892550, CG892551, CG892552, CG892553, CG892554, CG892555, CG892559, L23424.1, M84744, Z37543, BT000450.1, AB0379751, BI955682, BM137086, AY325302, AL831803, BE421261, CD862515, AU082986, and X78814.

ACKNOWLEDGMENTS

We thank Drs. A. Vigneswaran and V. Upasani for assistance in some of the earlier work, Dr. F. Khachik (University of Maryland) for valuable advice on carotenoid analysis, Drs. B. Buckner (Truman State University) for the PSY1 cDNA, N. Misawa for Erwinia constructs, S. McCouch (Cornell University) for rice, P. Peterson (Smithsonian Institution), and L. Clark (Iowa State University) for Poaceae subfamily plant and DNA samples, respectively.

Received January 27, 2004; returned for revision April 19, 2004; accepted April 28, 2004.

LITERATURE CITED

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389–3402

Bartley GE, Scolnik PA (1993) cDNA cloning, expression during development, and genome mapping of Psy2, a second tomato gene encoding phytoene synthase. J Biol Chem **268:** 25718–25721

Bendich A, Olson J (1989) Biological actions of carotenoids. FASEB J **3:** 1927–1932

Bird CR, Ray JA, Fletcher JD, Boniwell JM, Bird AS, Teulieres C, Blain I, Bramley PM, Schuch W (1991) Using antisense RNA to study gene function: inhibition of carotenoid biosynthesis in transgenic tomatoes. Biotechnology **9:** 635–639

Blott S, Kim J-J, Moisio S, Schmidt-Kuntzel A, Cornet A, Berzi P, Cambisano N, Ford C, Grisart B, Johnson D, et al (2003) Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. Genetics **163:** 253–266

Bramley P, Teulieres C, Blain I, Bird C, Schuch W (1992) Biochemical characterization of transgenic tomato plants in which carotenoid synthesis has been inhibited through the expression of antisense RNA to pTOM5. Plant J **2:** 343–349

Buckner B, Robertson DS (1993) Cloning of carotenoid biosynthetic genes from maize. Methods Enzymol **214:** 311–323

Buckner B, San Miguel P, Bennetzen JL (1996) The *y1* gene of maize codes for phytoene synthase. Genetics **143:** 479–488

Burkhardt PK, Beyer P, Wünn J, Klöti A, Armstrong GA, Schledz M, von Lintig J, Potrykus I (1997) Transgenic rice (*Oryza sativa*) endosperm expressing daffodil (*Narcissus pseudonarcissus*) phytoene synthase accumulates phytoene, a key intermediate of provitamin A biosynthesis. Plant J **11:** 1071–1078

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

Chamovitz D, Misawa N, Sandmann G, Hirschberg J (1992) Molecular cloning and expression in *Escherichia coli* of a cyanobacterial gene coding for phytoene synthase, a carotenoid biosynthesis enzyme. FEBS Lett **296:** 305–310

Cunningham FX, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol **49:** 557–583

Davison PA (2002) Overexpression of beta-carotene hydroxylase enhances stress tolerance in arabidopsis. Nature **418:** 203–206

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci **8:** 978–984

FAO (1995) Sorghum and Millets in Human Nutrition. Vol 27, FAO, Rome

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

Gallagher CE, Cervantes-Cervantes M, Wurtzel ET (2003) Surrogate biochemistry: use of *Escherichia coli* to identify plant cDNAs that impact metabolic engineering of carotenoid accumulation. Appl Microbiol Biotechnol **60:** 713–719

Giovannucci E, Ascherio A, Rimm EB, Stampfer MJ, Colditz GA, Willett WC (1995) Intake of carotenoids and retinol in relation to risk of prostate cancer. J Natl Cancer Inst **87:** 1767–1776

Giuliano G, Bartley GE, Scolnik PA (1993) Regulation of carotenoid biosynthesis during tomato development. Plant Cell **5:** 379–387

Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr Opin Plant Biol **4:** 210–218

Khachik F, Steck A, Pfander H (1999) Isolation and structural elucidation of (13Z,13′Z,3R,3′R,6′R)-lutein from marigold flowers, kale, and human plasma. J Agric Food Chem **47:** 455–461

Kohlmeier L, Kark JD, Gomez-Gracia E, Martin BC, Steck SE, Kardinaal AF, Ringstad J, Thamm M, Masaev V, Riemersma R, et al (1997) Lycopene and myocardial infarction risk in the EURAMIC Study. Am J Epidemiol **146:** 618–626

Krinsky NI, Landrum JT, Bone RA (2003) Biologic mechanisms of the protective role of lutein and zeaxanthin in the eye. Annu Rev Nutr **23:** 171–201

Kumagai MH, Donson J, della-Cioppa G, Harvey D, Hanley K, Grill LK (1995) Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. Proc Natl Acad Sci USA **92:** 1679–1683

Lee C, McCoon P, LeBowitz J (1981) Vitamin A value of sweet corn. J Agric Food Chem **29:** 1294–1295

Li ZH, Matthews PD, Burr B, Wurtzel ET (1996) Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway. Plant Mol Biol **30:** 269–279

Matthews PD (2001) Carotenogenesis in maize and rice. PhD thesis. Graduate School and University Center, The City University of New York, New York

Matthews PD, Luo R, Wurtzel ET (2003) Maize phytoene desaturase and

zetacarotene desaturase catalyze a poly-Z desaturation pathway: implications for genetic engineering of carotenoid content among cereal crops. J Exp Bot **54:** 2215–2230

Matthews PD, Wurtzel ET (2000) Metabolic engineering of carotenoid accumulation in *Escherichia coli* by modulation of the isoprenoid precursor pool with expression of deoxyxylulose phosphate synthase. Appl Microbiol Biotechnol **53:** 396–400

Misawa N, Nakagawa M, Kobayashi K, Yamano S, Izawa Y, Nakamura K, Harashima K (1990) Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functional analysis of gene products expressed in Escherichia coli. J Bacteriol **172:** 6704–6712

Palaisa KA, Morgante M, Williams M, Rafalski A (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell **15:** 1795–1806

Randolph LF, Hand DB (1940) Relation between carotenoid content and number of genes per cell in diploid and tetraploid corn. J Agric Res **60:** 51–64

Robertson D, Anderson I, Bachmann M (1978) Pigment-deficient mutants: genetic, biochemical and developmental studies. *In* D Walden, ed, Maize Breeding and Genetics. John Wiley & Sons, New York, pp 461–494

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Schledz M, Al-Babili S, von Lintig J, Haubruck H, Rabbani S, Kleinig H, Beyer P (1996) Phytoene synthase from *Narcissus pseudonarcissus*: functional expression, galactolipid requirement, topological distribution in chromoplasts and induction during flowering. Plant J **10:** 781–792

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

Sommerburg O, Keunen JE, Bird AC, van Kuijk FJ (1998) Fruits and vegetables that are sources for lutein and zeaxanthin: the macular pigment in human eyes. Br J Ophthalmol **82:** 907–910

von Lintig J, Welsch R, Bonk M, Giuliano G, Batschauer A, Kleinig H (1997) Light-dependent regulation of carotenoid biosynthesis occurs at the level of phytoene synthase expression and is mediated by phytochrome in *Sinapis alba* and *Arabidopsis thaliana* seedlings. Plant J **12:** 625–634

Wurtzel ET (1992) Use of a *Ds* chromosome breaking element to examine maize *Vp5* expression. J Hered **83:** 109–113

Ye X, Al-Babili S, 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

Zhang H-B, Choi S, Woo S-S, Li Z, Wing RA (1996) Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. Mol Breed **2:** 11–24