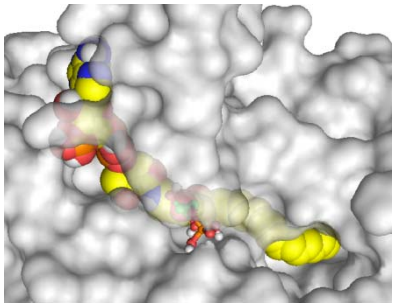
Cloning, heterologous expression and biochemical characterization of plastidial *sn*-glycerol-3-phosphate acyltransferase from *Helianthus annuus.*

Miriam Payá-Milans, Mónica Venegas-Calerón*, Joaquín J. Salas, Rafael Garcés and Enrique Martínez-Force.

Biochemical characterization of recombinant GPAT showed an oleate specific isoform in sunflower plastids, concomitant with plant resistance to chilling temperatures.



Cloning, heterologous expression and biochemical characterization of plastidial *sn*-glycerol-3-phosphate acyltransferase from *Helianthus annuus.*

Miriam Payá-Milans, Mónica Venegas-Calerón*, Joaquín J. Salas, Rafael Garcés and Enrique Martínez-Force.

Instituto de la Grasa, CSIC, Carretera de Utrera Km1, 41013-Sevilla, Spain.

*To whom correspondence should be addressed

Address Correspondence to: Mónica Venegas-Calerón, Instituto de la Grasa (CSIC), Carretera de Utrera Km1, 41013, Sevilla, Spain.

Tlf: +34 954611550-272

Fax: +34 954616790

Email: mvc@ig.csic.es

Abstract

The acyl-[acyl carrier protein]:*sn*-1-glycerol-3-phosphate acyltransferase (GPAT; E.C. 2.3.1.15) catalyzes the first step of glycerolipid assembly within the stroma of the chloroplast. In the present study, the sunflower (*Helianthus annuus,* L.) stromal GPAT was cloned, sequenced and characterized. We identified a single ORF of 1344 base pairs that encoded a GPAT sharing strong sequence homology with the plastidial GPAT from *Arabidopsis thaliana* (ATS1, At1g32200). Gene expression studies showed that the highest transcript levels occurred in green tissues in which chloroplasts are abundant. The corresponding mature protein was heterologously overexpressed in *Escherichia coli* for purification and biochemical characterization. In vitro assays using radiolabelled acyl-ACPs and glycerol-3-phosphate as substrates revealed a strong preference for oleic versus palmitic acid, and weak activity towards stearic acid. The positional fatty acid composition of relevant chloroplast phospholipids from sunflower leaves did not reflect the in vitro GPAT specificity, suggesting a more complex scenario with mixed substrates at different concentrations, competition with other acyl-ACP consuming enzymatic reactions, etc. In summary, this study has confirmed the affinity of this enzyme ~~affinity~~ which would partly explain the resistance to cold temperatures observed in sunflower plants.

Keywords:

Helianthus annuus; sunflower; substrate specificity; glycerol-3-phosphate acyltransferase; plastids.

Abbreviations:

ACP: acyl-carrier protein

DAF: days after flowering

DAG: days after germination

DGDG: digalactosyldiacylglycerol

ER: endoplasmic reticulum

EST: expressed sequence tag

G3P: glycerol-3-phosphate

GPAT: *sn*-1 glycerol-3-phosphate acyltransferase

HaPLSB: Helianthus annuus chloroplast GPAT gene

LPA: lysophosphatidic acid

MGDG: monogalactosyldiacylglycerol

PA: phosphatidic acid

PC: phosphatidylcholine

PG: phosphatidylglycerol

SQDG: sulfoquinovosyldiacylglycerol

1. Introduction

Plants are characterized for being photosynthetic organisms with most of their photosynthetic apparatus embedded in a lipid bilayer. The photosynthetic membranes are largely composed of monogalactosyldiacylglycerol (MGDG), digalactosyldiacylglycerol (DGDG), sulfoquinovosyldiacylglycerol (SQDG) and phosphatidylglycerol (PG) (Moreau et al., 1998). The diacylglycerol backbones required for the biosynthesis of these lipids are provided by two pathways: the eukaryotic pathway, through which glycerolipids are imported from the endoplasmic reticulum (ER) to chloroplasts, mainly in the form of phosphatidylcholine (PC); and the prokaryotic pathway, through which glycerolipids are synthesized de novo within chloroplasts (Ohlrogge and Browse, 1995). Acyltransferases (EC 2.3.1.x) catalyze essential reactions in the biosynthesis of phospholipids and triacylglycerols. Even though the mechanism of catalysis is similar for these acyltransferases, they differ in target substrate and specificity. Glycerol-3-phosphate acyltransferases (GPATs) and lysophosphatidic acid acyltransferases (LPAATs), with *sn*-1 and *sn*-2 positional specificities respectively, form the first step in the synthesis de novo of glycerolipids, yielding their precursor phosphatidic acid. The high degree of fatty acid specificity of *sn*-2 acyltransferases for 16:0 in chloroplasts and 18:1 in the ER may serve as a useful indicator of the molecular species proceeding from each of these two pathways (Ohlrogge and Browse, 1995). Plants can be classified based on the presence or absence of *cis*-7,10,13-hexadecatrienoic acid (16:3) in the *sn*-2 position of MGDG (Mongrand et al., 1998). In 16:3 plants, both pathways are functional and produce 18:3/16:3-MGDG in chloroplasts, while in 18:3 plants the biosynthesis of galactolipids and SQDG by the prokaryotic pathway is truncated, and the backbones are transported from the ER-localized PC moiety probably in the form of phosphatidic acid (Roston et al., 2012). Galactolipids and SQDG are derived from both ER and plastid sources, while plastid PG is exclusively synthesized by the prokaryotic pathway (Kunst et al., 1988; Xu et al., 2006) and constitutes the main phospholipid in thylakoid membranes (Moreau et al., 1998).

Glycerol-3-phosphate acyltransferases (GPATs; E.C. 2.3.1.15; PlsB) are enzymes that catalyze the acylation of a molecule of glycerol-3-phosphate (G3P) at the *sn*-1 position to produce lysophosphatidic acid (LPA). They are found in both soluble and membrane-bound forms in higher plants (Bertrams and Heinz, 1976). Membrane-bound GPATs are mainly involved in the synthesis of cutin, suberin and storage lipids, and they are localized in the ER and mitochondrial membranes (Yang et al., 2012). The soluble forms are found in chloroplasts (green tissues) and plastids (sink tissues), and they participate in the prokaryotic biosynthesis of phospholipids, sulfolipids and galactolipids (Ohlrogge and Browse, 1995). The fatty acid substrates of these enzymes are synthesized within the chloroplast, yielding 16:0-ACP, 18:0-ACP and 18:1-ACP, which can be used by either the soluble GPATs or hydrolyzed by acyl-ACP thioesterases. Following such hydrolysis, the acyl moieties are exported out of the plastids in the form of acyl-CoAs, which are then incorporated into glycerolipids via ER-based pathways (Sanchez-Garcia et al., 2010).

Several different approaches have been used to characterize stromal GPAT, the earliest using chloroplasts or leaves from pea (Pisum sativum), spinach (Spinacea oleracea) and squash (Cucurbita moschata) to obtain enzyme-enriched fractions by ion-exchange chromatography (Bertrams and Heinz, 1981; Frentzen et al., 1983; Frentzen et al., 1987). Both acyl-CoAs and acyl-ACPs were identified as substrates of these enzymes, which were generally more specific for their physiological substrates (i.e.: acyl-ACPs). The first cDNA nucleotide sequence of a plastidial GPAT was published in 1988 (Ishizaki et al., 1988); it was cloned from squash leaves and later used for the first crystallographic three-dimensional structure description of a recombinant GPAT (Turnbull et al., 2001). In enzymatic studies two different forms of GPATs were depicted: selective GPAT, targeting 18:1 over 16:0 fatty acids (Frentzen et al., 1983); and non-selective GPAT, targeting 16:0 and 18:1 fatty acids at comparable rates (Cronan and Roughan, 1987). Research suggests a correlation between the ability of some plants to tolerate cold temperatures and the presence of a certain form of GPAT. Thus, those carrying the nonselective form of the enzyme are usually chill-sensitive plants, whereas those with selective

forms of the enzyme are usually chill-tolerant (Tamada et al., 2004; Zhu et al., 2009). This is due to the higher capacity of cis unsaturated phosphatidylglycerol (PG) to prevent membrane phase transitions that may take place with saturated PG species at low temperatures (Nishida and Murata, 1996; Sun et al., 2011). An analogous mechanism has been described for high temperature stress, with increased thermotolerance associated with an increase in the level of saturation of thylakoid membrane lipids (Sui et al., 2007a; Yan et al., 2008).

Here, we describe the identification of a cDNA sequence encoding the plastidial GPAT from sunflower and the properties of this recombinant sunflower GPAT when expressed in *E. coli*, which includes a high degree of specificity for oleate. Based on our findings, we discuss the role of this enzyme in the fatty-acid profile of polar lipids in sunflower leaves.

2. Results and discussion

2.1. Identification of the plastidial GPAT from sunflower

A BLAST search with the plastidial GPAT sequence from *Arabidopsis (*ATS1) returned one relevant EST corresponding to its sunflower homologue. Based on this result, an internal fragment was cloned by PCR and sequenced. Primers for PCR were designed within this region and using the RACE technique, a cDNA clone of 1751 bp was obtained that contained an ORF of 1344 bp and UTR regions 74 bp upstream and 333 bp downstream. The gene was designated *HaPLSB,* given its homology with the *Escherichia coli* GPAT encoded by the *plsB* gene.

GPAT is synthesized as a 447 amino acid precursor protein, yet once inside the chloroplast the mature protein presumably lacks the transit peptide located at the N-terminus. Previous studies predicted the cleavage site of GPATs in various plants (Murata and Los, 1997; Slabas et al., 2002), although producing no clear consensus. However, on the basis of TargetP results and previous findings (Slabas et al., 2002), a transit peptide of 57 amino acids was predicted for sunflower GPAT. An alignment of the plastidial GPAT protein sequences from sunflower and representatives of distinct photosynthetic taxa (dicot, monocot, moss and green algae) highlighted the strong conservation of these proteins during evolution (Figure 1). For example, the acyltransferase from the green alga *Chlamydomonas reinhardtii* shares 33% identity and 50% sequence similarity with *Helianthus annuus* stromal GPAT. The amino terminal regions correspond to the transit peptides and present major changes compared with the rest of the sequences to the point that the sequences are totally dissimilar. The cores of the enzymes are highly conserved indicating essential motifs that are required to maintain its activity. Some of these conserved residues determine the substrate binding sites or the catalytic center and will be discussed in the next subsection.

A phylogenetic tree was constructed for the deduced amino acid sequence of *Ha*PLSB and for GPAT protein sequences from chloroplasts in a variety of eukaryotic organisms, including

diverse taxa from algae to higher plants (Supplementary data, Figure 1). Since no identifiable homologues exist in cyanobacteria, GPATs from *Chlamydiae* species were used as an outgroup to root the tree. The short distance at the root of the tree suggests a very similar ancestral form of the protein in both prokaryotes and eukaryotes, probably resulting from horizontal gene transfer. The branching pattern of flowering plant families, with very close nodes, is indicative of their rapid evolution. Despite the long period of evolutionary time since the divergence of green algae and higher plants, stromal GPAT has not undergone substantial evolutionary changes, also viewed in the alignment.

2.2. Analysis of enzyme structure based on the GPAT sequence

The tertiary structure of recombinant GPAT from squash chloroplasts, with which *Ha*PLSB showed a high degree of similarity (67.6% identity), has been determined (Turnbull et al., 2001). Based on this structure, we modeled the tertiary structure of *Ha*PLSB in silico and we predicted sunflower GPAT to be composed mainly of alternating α/β secondary structures. The *Ha*PLSB model contains 15 α-helices and 9 β-strands, which account for 46% and 13% of the sequence, respectively (the relative positions of these secondary structure elements are shown in the polypeptide sequence: Figure 2).

The modeled 3D structure is composed of two domains: a helical domain forming a four-helix bundle (residues 82-154) and an alternating α/β domain (residues 162-445) linked by a loop region (residues 155-161: Figure 3A). The α/β domain is folded into parallel/antiparallel β-strands surrounded by the remaining 11 α-helices. Pockets and tunnels that may bind substrates and regulators are visible on the protein surface (Figure 3B and C). One pocket (surrounded by the blue patch in Figure 3B) is highly conserved among chloroplast GPATs and it contains a positively charged surface patch formed by residues His217, Lys271, His272, Arg313 and Arg315. This pocket is predicted to contain the G3P binding site, which would attract the phosphate moiety to the basic surface patch and then accommodate the molecule in the pocket at position *sn*-1, located near the catalytic residues His217 and Asp222. These residues, which

belong to the H(X_4)D motif (found in most G3P acyltransferases and that is important for catalytic activity), are located near a predicted hydrophobic patch (highlighted in yellow in Figure 3B) and the positive cluster that bind the acyl chain and G3P molecule, respectively. Considering that a L261P mutation in squash (Slabas et al., 2002; homolog L339 in sunflower shown in brown) in this hydrophobic patch changed enzyme selectivity, specificity may be related to flexibility of the acyl chain when binding to the enzyme surface. Docking experiments using G3P and oleyl-CoA (18:1-CoA) as acyl donor (Figure 3C and D) corroborates the positioning of the G3P molecule as described, while the acyl group fits in a long tunnel interacting with residues A221, P223, Y243, V244, V249, C255, P258, S259, L264, W306 and M360. Docking experiments of palmitoyl-CoA (16:0-CoA) and stearoyl-CoA (18:0-CoA) gave similar results to oleyl-CoA (data not shown), concordant to the suggestion that selectivity may occur prior to entering the tunnel.

2.3. HaPLSB expression

Quantitative RT-PCR was used to study the expression of *HaPLSB*, normalizing the data to that of actin as a housekeeping gene. In accordance with the role of its protein product in chloroplast lipid synthesis, the *HaPLSB* gene was predominantly expressed in green tissues with little expression in sink tissues (Figure 4A). Its expression was enhanced during cotyledon development, consistent with the elevated rate of de novo membrane lipid biosynthesis in the chloroplast during the early stages of plant growth, and was maintained at high levels in mature leaves. The expression profile obtained was similar to that of the corresponding plastidial GPAT from Arabidopsis (ATS1; Nishida et al., 1993), with the highest levels of expression found in cotyledons and early leaves (Figure 4B; data from Schmid et al., 2005). One significant difference in the expression of this gene was observed during the development of *Arabidopsis* and sunflower seeds, reflecting the maintenance of active chloroplasts during seed development in *Arabidopsis* and their continuous decay in sunflower in favor of oil biosynthesis and accumulation. This difference was expectable due to *Arabidopsis* producing green seeds that are

able to do photosynthesis during most of their developmental period, whereas sunflower yield white seeds that are sink organs and lack any photosynthetic machinery.

2.4. Heterologous expression in E. coli

A heterologous expression system was prepared by cloning the predicted *Ha*PLSB mature protein coding region of 1140 nucleotides into a pQE-80L inducible expression vector (primers shown in Table 1). This recombinant pQE80::*Ha*PLSB expression vector was used to transform the competent *E. coli* XL1-Blue strain. Cultures of transformed *E. coli* cells were grown with IPTG induction for the over-expression of the recombinant *HaPLSB* gene and without induction as control. The fatty acid composition of the host in both uninduced and induced cultures was analyzed (Table 2).

The differences encountered in the fatty acid composition of both cultures were an indicative of the successful production of an active recombinant enzyme that was altering the fatty acid profile in the host bacteria. Induced bacteria showed increased unsaturated fatty acid proportion (54% compared to 50% in control) concomitant with a decrease of cyclopropane fatty acids (11% compared to 14% in control), which reflects the higher membrane production due to a more rapid growth rate. In this case, the fatty acid species with 16 carbon atoms experienced more relevant changes, varying their proportion of saturated and unsaturated, by diminishing and increasing 9% and 10% respectively, taking into account that the 17:0Δ derives from 16:1. The flux of intermediates into each branch of saturated and unsaturated fatty acids biosynthesis in bacteria is regulated by their long-chain products, and then the increase of unsaturated fatty acids in induced cultures may respond to their consumption by *Ha*PLSB (Rock and Jackowski, 2002). The lipid content of induced cultures was also slightly lower than that of uninduced cultures, due to the elongation of the exponential phase in GPAT overproducing cultures instead of the decrease in GPAT activity and accumulation of intracellular acyl-ACP when reaching the stationary phase (Heath et al., 1994). Thus, the over-production of the exogenous *Ha*PLSB in *E.*

coli cells alters the composition and levels of the pool of acyl-ACP, leading to a variation in cell fatty acid profile.

2.5. In vitro substrate specificity of recombinant GPAT

Specificity assays were performed in vitro with the recombinant *Ha*PLSB enzyme purified from *E. coli* (Figure 5) to identify the form (selective or nonselective) of the enzyme expressed by this sunflower line. Radioactive 16:0, 18:0 and 18:1 acyl-ACPs were used as separate substrates, and activity was measured in function of the production of radiolabelled lysophosphatidic acid (LPA) from unlabelled G3P. The reaction mixture contained varying concentrations of acyl-ACP and G3P (with 5 mg/mL BSA), at a pH of 7.4, that of the plastid stroma (Shen et al., 2013). According to previous calculations in spinach chloroplasts, the physiological concentration of acyl-ACP is around 1-2 µM (Soll and Roughan, 1982) while the stromal G3P concentration in *Amaranthus* is estimated at 450-620 µM (Cronan and Roughan, 1987).

At identical acyl-ACP concentrations, oleoyl-ACP was the most active acyl donor to sunflower GPAT (Figure 6A), and at the predicted physiological acyl-ACP concentrations, the rates of 18:1-LPA, 16:0-LPA and 18:0-LPA synthesis were 0.2, 0.1 and 0.01 nkat per mg protein, respectively. In these conditions, the rate of 18:1-ACP incorporation was twice that of 16:0-ACP, suggesting that the products contained mainly 18:1, some 16:0 and traces of 18:0 in position *sn*-1.

GPAT displayed a lower K_M value for 16:0-ACP than for 18:1-ACP at the predicted physiological concentration, and a 9-fold lower V_{max} (Table 3). By contrast, the K_M for 18:1-ACP was 4-fold higher than the predicted stromal 18:1-ACP concentration. The k_{cat}/K_M ratio was used to compare the catalytic efficiencies of *Ha*PLSB towards its different substrates, revealing greater specificity and catalytic efficiency of *Ha*PLSB towards 18:1 than 16:0, as expected for a chilling-resistant plant (Zhu et al., 2009). Moreover, the low activity levels

towards 18:0-ACP suggest limited incorporation of this fatty acid into LPA via the intraplastidial GPAT enzyme.

Due to the better availability and stability, as well as not leading to saturating levels during the assays, palmitoyl-ACP was used as an acyl donor to study the kinetic parameters of G3P (Figure 6B). The K_M value for this substrate was 659 μM, which was in the same order of magnitude reported for *Amaranthus* (Cronan and Roughan, 1987), suggesting similar conditions in sunflower chloroplasts. As expected, the turnover under saturating conditions of G3P was similar to that obtained under saturating conditions of 16:0-ACP.

2.6. Effect of pH on HaPLSB activity

The pH inside the chloroplast ranges from 7.4 at night to 8.0 during the day and it regulates the activity of many enzymes. We studied the response of the *Ha*PLSB enzyme to pH using palmitoyl-ACP as an acyl donor, observing maximum *Ha*PLSB activity at pH 6.9 that decreased significantly up to pH 8.0. This pH curve (Figure 6C) is similar to that reported for the pea chloroplast GPAT (Bertrams and Heinz, 1981), for which the optimal pH value is 7.4. A recent study reported that the activity of the plastidial GPAT from eight different plant species at pH 7.4 is lower than at pH 8.0 (Zhu et al., 2009), showing that plastidial GPAT activity responds to pH changes.

2.7. Lipid analysis of sunflower leaves

After characterizing sunflower GPAT, we analyzed the species and fatty acid composition from sunflower leaf chloroplast lipids, paying special attention to MGDG, DGDG and PG. Unlike other organelles, chloroplast membranes are rich in galactolipids and contain low levels of phospholipids (Moreau et al., 1998). Thus, some phospholipid species like phosphatidyl ethanolamine (PE) , which accounts for 20 to 40% of the total species present in extraplastidial membranes, are absent in chloroplast ones. The polar lipid extract prepared from sunflower leaves contained 4% of PE (Table 4A) and 7% of phosphatidic acid (PA), which are probably

derived from the ER and the plasma membranes (with an estimated PA level of 5-20%). As expected, almost half of the lipids analyzed corresponded to MGDG, being the levels of sulfolipids below the 6% estimated for chloroplast membranes.

The sunflower belongs to a plant group known as 18:3 plants (Mongrand et al., 1998) in which only PG is exclusively synthesized inside the chloroplast, all other plastid lipids being synthesized via the eukaryotic pathway (see Introduction). To confirm this lipid profile, MGDG, DGDG and PG were isolated and their positions analyzed. These lipids were treated with lipozyme from *Mucor miehei*, which specifically liberates the acyl group in position *sn*-1. The PG purified, mainly produced via the prokaryotic pathway, provides an indication of the specificity of the plastid acyltransferase. Assuming that 18:2 and 18:3 are derived from an 18:1 precursor, and 16:1 from a 16:0 precursor, this phospholipid had 52% 18:1, 29% 16:0 and 19% 18:0 at position *sn*-1, and 18% 18:1, 77% 16:0 and 5% 18:0 in position *sn*-2. The *sn*-1 proportions do not appear to tally with the specificity of the *Ha*PLSB enzyme. Previous reports usually measured very low levels of stearic acid at *sn*-1 position in leaf lipids with the exception of *Cucumis sativus* with 12% 18:0 (Murata et al., 1982; Whitaker, 1986). Regarding these data, the proportion of stearic acid in sunflower is relatively high, but the final PG composition must be influenced by the effect of intraplastidial substrate concentrations in the enzymatic activity, as well as other possible regulatory mechanisms. Moreover, differences in the degree of unsaturation of 16 carbon atoms molecules were detected between the two positions of the PG molecule. A similar finding was reported previously in tomato (Sui et al., 2007), where 16:0 desaturation occurs asymmetrically between positions *sn*-1 and -2, enriching the *sn*-1 position in saturated fatty acids and the *sn*-2 in delta 3-trans-hexadecenoyl (16:1(3t)).

Galactolipids are major components of chloroplast membranes and play an active role during cold acclimation (Moellering et al., 2010). In sunflower, both MGDG and DGDG are synthesized using phosphatidic acid precursors imported from the ER (Roston et al., 2012) and we found that MGDG molecules contained mainly 18:1 in position *sn*-1, while DGDG also contained 16:0 and 18:0 (Table 4B). This asymmetry suggests a possible selection against the

galactosylation of 18:1/18:1 MGDG by DGDG synthase in favor of molecular species with other fatty acids. In both galactolipids, the proportion of 18:1 derivatives in *sn*-2 was over 90%, mainly in the 18:3 form. Being the galactolipids the most abundant membrane lipids in plastids, we suggest that the different fatty acid distribution in both MGDG and DGDG moieties facilitates membrane remodelling, as MGDG molecules are consumed during freezing tolerance to form oligogalactolipids and diacylglycerol (Moellering et al., 2010).

3. Conclusions

G3P acyltransferases play an essential role in membrane synthesis by catalyzing the first step in glycerolipid assembly. Enzymatic specificity determines membrane fluidity and in chloroplasts, the degree of unsaturation of PG molecules determines their ability to protect the photosynthetic apparatus from damage caused by high or low temperatures. Sunflower chloroplast GPAT displays a specificity profile that confers resistance to cold temperatures due to its specificity for unsaturated substrates as oleate. This characteristic is reflected in the profile of fatty acids in position *sn*-1 in leaf PG, the majority of which are oleoyl-derived. Concerning the nature of this enzyme, the sequence of this selective GPAT form is now available for future studies seeking to transform sensitive plants to overcome temperature sensitivity.

4. Experimental

4.1. Biological material and growth conditions

Plants from the standard sunflower line CAS-6 were cultivated in growth chambers at 25/15ºC (day/night) with a 16 h photoperiod (300 $\mu E\ m^{-2}\ s^{-1}$). RNA was extracted from 0.5 g of mature functional leaves from two-month-old plants to produce cDNA, as described previously (Sanchez-Garcia et al., 2010). The *E. coli* strain XL1-Blue (Stratagene, California) was used as a host for the expression of sunflower plastidial GPAT (*HaPLSB*), growing the bacteria at 37ºC with vigorous shaking in LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl). For plasmid selection, ampicillin (100 μg/mL) was added to the medium.

4.2. Cloning and sequencing of the sunflower GPAT gene

The sunflower EST database was searched against the plastid-located GPAT protein sequence from *Arabidopsis thaliana* (L.) Heynh. (ATS1, accession no. NP_174499) using the tBLASTn search tool (Altschul et al., 1990) of the NCBI. The EST with GenBank accession no. CD847532.1 was used to design primers. Using cDNAs from sunflower leaves as a template, an

internal fragment of 665 bp was amplified using the specific HaPLSB-F1 and HaPLSB-R1 primers (Table 1). All primers were synthesized by Eurofins MWG Operon (Germany). The 5' cDNA end was amplified using the BD SMART™ RACE cDNA Amplification Kit (Clontech, Japan) and the specific internal primers HaPLSB-R1 and HaPLSB-R2 (Table 1). The 3' end was amplified using the external oligonucleotide FA2Z (Table 1), complementary to the sequences incorporated during the initial cDNA synthesis, and the specific internal primers HaPLSB-F1 and HaPLSB-F2 (Table 1). The amplification products were cloned into the pMBL-T cloning vector (Dominion MBL, Spain), sequenced (Secugen, Spain) and the sequences were verified by comparison with other GPATs. Finally, using the specific HaPLSB-F and HaPLSB-R PCR primers (Table 1), the full *HaPLSB* gene (1344 bp) was amplified, cloned and sequenced, and the sequence was deposited at GenBank under accession no. HM490309.

4.3. Overexpression and purification of the HaPLSB protein in E. coli

The coding region of the putative mature *Ha*PLSB protein was amplified with the SphIHaPLSB-F and KpnIHaPLSB-R primers (Table 1) and subcloned into the *Sph*I-*Kpn*I sites of the pQE-80L expression vector (Qiagen, Germany) to produce a fusion protein with a 6xHis tag at the amino terminus. Correct in frame insertion was verified by sequencing and the resulting construct was designated pQE::*Ha*PLSB. *E. coli* XL1-Blue cells were transformed with pQE::*Ha*PLSB or the empty vector (as a control) and grown in LB medium at 37ºC with shaking. Protein over-production and cell lysis were performed as described previously (Sanchez-Garcia et al., 2010). The *Ha*PLSB protein was purified from the resulting soluble fraction using the His SpinTrap Kit (GE Healthcare, UK) and the enzyme was stored at -20ºC in glycerol-containing buffer. The concentration of the purified protein was quantified by the Bradford method (Bradford, 1976) using the Bio-Rad Protein Assay (Bio-Rad, California) and the proteins were visualized by SDS-PAGE (Sanchez-Garcia et al., 2010).

4.4. Synthesis of radioactive acyl-ACPs

Radiolabelled acyl-ACP substrates were synthesized from commercial [1-^{14}C] palmitic, [1-^{14}C] stearic and [1-^{14}C] oleic acids (ARC, Missouri; 55 Ci/mol), as described elsewhere (Sanchez-Garcia et al., 2010).

4.5. Determination of GPAT substrate specificity

Substrate specificity of GPAT was determined as described previously (Frentzen et al., 1987; Zhu et al., 2009) with some minor modifications. The reaction was performed in a total volume of 100 µL containing 0.25 M HEPES buffer (pH 7.4), 5.0 mg/mL BSA and 0.4 mM G3P, with variable amounts of radioactive acyl-ACP. During the acyl-ACP specificity assays, substrate concentrations ranged from 0.5 to 24 µM, while the concentration of G3P ranged from 50 to 800 µM, with 3 µM 16:0-ACP. To study the influence of pH, the system was buffered with a mixture of 0.1 M Tris, 0.05 M MES and 0.05 M acetic acid, with 2 µM 16:0-ACP. Assays began with the addition of 200 ng of the purified protein and after a 5 min incubation at 25ºC, the reaction was stopped by adding 100 µL acetic acid (0.15 M).

4.6. Lipid extraction and analysis

Lipids formed in the reactions were extracted as described previously (Ruiz-Lopez et al., 2010). Polar lipids were analyzed on TLC plates, previously dried and activated by heating in an oven at 120ºC for 1 h, using chloroform-methanol-acetic acid-water (85:15:10:3.5) as the solvent. The radioactive product was visualized by autoradiography on a Packard Instant Imager (Packard Instrument Co, Connecticut). LPA was scraped from the plate and the radioactivity incorporated was determined using on a LS-6500 Multipurpose Scintillation Counter (Beckman Coulter, California) Ecoscint scintillant (National Diagnostics, Georgia).

4.7. Total lipid extraction from leaves

The total lipid fraction was extracted using modified conditions of the Hara and Radin (1978) method, from mature sunflower leaves (0.5 g) maintained at 80ºC in 6 mL isopropanol for 15 min. The tissue was then ground with sea sand and the lipids extracted after the addition of 9

mL of hexane and 7.5 mL of 6.7% Na_2SO_4. The organic phase was transferred to a fresh tube and the aqueous phase was re-extracted with 10 mL hexane-isopropanol (7:2). After combining the organic phases, the solvent was evaporated under vacuum in a rotary evaporator (VWR by IKA®, Germany). The lipid residue was dissolved in 3 mL of chloroform and different aliquots of this extract were used for to assay the polar lipids by HPLC and fatty acid analysis.

4.8. Isolation and lipolysis of polar lipids from chloroplasts

Total lipids were loaded onto a TLC plate coated with 0.15 M $(NH_4)_2SO_4$ and the polar lipids were separated with a mixture of acetone-toluene-water (70:23:6). The lipids were visualized by exposure to iodine vapor and identified using standards. MGDG, DGDG and PG were scraped off the plate and eluted from the silica with chloroform-methanol (1:1). The lipids were then dried under nitrogen gas and dissolved in acetonitrile. Lipolysis was performed in a final volume of 200 µL by adding an equal volume of 40 mM HEPES (pH 7.4) and 10 U of immobilized lipozyme from *Mucor miehei* (Sigma-Aldrich, Missouri) and incubating the mixture overnight at room temperature. Next, 1.4 mL chloroform-methanol (1:1) and 560 µL H_2PO_4 (0.2 M) in 1 M KCl was added to the reaction, mixed vigorously and centrifuged to separate the phases. The lower phase (containing the lipids) was removed and dried under nitrogen gas. The lipolytic products were separated by chromatography in two-stages: first in hexane/diethyl ether-acetic acid (40:60:1) to recover the non-polar lipids and then with chloroform-methanol-acetic acid-water (85:15:10:3.5) to obtain the polar lipids. The lipids were recovered from the TLC plate and used to prepare fatty acid methyl esters.

4.9. Preparation and analysis of fatty acid methyl esters

The fatty acids in transformed *E. coli* cells expressing the recombinant *Ha*PLSB protein were analyzed. Cells taken from 50 mL of induced and non-induced cultures were harvested at OD_{600} ~ 2, washed with distilled water and 3 mL of the methylation mixture (~1.25 M hydrogen chloride in methanol; Fluka, Germany), and 0.15 mg of the internal standard heptadecanoic acid (17:0) was added to the pellet. The samples were methylated for 1 h at 80ºC and heptane (1 mL)

was added after cooling to extract the methyl esters. The upper phase was transferred to a fresh tube and washed with 2 mL Na_2SO_4 (6.7%) and the solvent was then evaporated under nitrogen gas. The methyl esters were finally re-suspended in 200 µL heptane.

Fatty acid methyl esters from lipolysis were obtained by heating the lipid residues dissolved in 2 mL methanol-toluene-sulfuric acid (170:30:5) at 80ºC for 1 h. After cooling, 2 mL Na_2SO_4 (6.7%) and 3 mL hexane were added, the solution was mixed vigorously and the upper layer was transferred to a fresh tube and the solvent was evaporated under nitrogen gas. The residues containing methyl esters were dissolved in 300 µL heptane.

Methyl esters were analyzed by gas-liquid chromatography (GLC) using a Hewlett–Packard 6890 gas chromatograph (California) with a Supelco SP-2380 fused-silica capillary column (length, 30 m; i.d., 0.25 mm; film thickness, 0.20 µm: Supelco, Pennsylvania). Hydrogen was used as the carrier gas at 28 cm/s, while the temperature of the flame ionization injector and detector was set at 200ºC, the oven temperature at 170ºC, and the split ratio was 1:50.

4.10. Quantitative PCR

Sunflower cDNAs from vegetative tissues and seeds at different developmental stages were subjected to quantitative real time PCR (qRT-PCR) using a pair of specific primers (QHaPLSB-F and QHaPLSB-R, Table 1) and SYBR Green I (QuantiTect® SYBR® Green PCR Kit: Qiagen, Crawley, UK) using a MiniOpticon system (Bio-Rad, California). The reaction mixture was heated to 95ºC for 15 min before subjecting it to 40 PCR cycles of 94ºC for 15 s, 58ºC for 30 s and 72ºC for 30 s, while monitoring the resulting fluorescence in each cycle. The calibration curve was drawn up using sequential dilutions of pQE::*Ha*PLSB and the comparative expression between samples was calculated using the Livak method (Livak and Schmittgen, 2001). The sunflower actin gene *HaACT1* (GenBank accession no. FJ487620) was used as an internal reference to normalize the amount of cDNA in each sample relative to the specific primer (QHaAct-F4 and QHaAct-R4, Table 1).

4.11. Bioinformatics

GPAT protein sequences from public databases were aligned to identify regions of homology using the ClustalX v2.1 program (Larkin et al., 2007). The sequences were aligned using BioEdit (Hall, 1999) and a phylogenetic tree was constructed using MEGA 5.0 software (Tamura et al., 2011). The amino terminal transit peptide was predicted using TargetP 1.1 (Emanuelsson et al., 2000). The tertiary structure of sunflower GPAT was modelled using the SWISS-MODEL server (Guex and Peitsch, 1997) based on the structure of *Cucurbita moschata* (squash) GPAT (accession no. BAB17755) from residues 82 to 445. Molecular docking experiment was performed using Vina with AutoDock Tools and PyMOL (Trott and Olson, 2010).

Acknowledgements

Thanks are due to A. Gonzalez-Callejas and B. Lopez-Cordero for skillful technical assistance. This work was supported by the MINECO and FEDER [project AGL2011-23187].

Figure legends and Tables

Figure 1. Alignment of deduced *Ha*PLSB amino acid sequence with PLSB proteins from various species. Identical amino acids are shaded in black. The reference sequences are: *Ha*PLSB (HM490309), *Helianthus annuus*; ATS1 (GI: 15222600), *Arabidopsis thaliana*; *Cm*PLSB (BAB17755), *Cucurbita moschata*; *Os*PLSB (GI: 115483650), *Oryza sativa*; *Pp*PLSB (GI: 168037616), *Physcomitrella patens* subsp. *patens*; *Cr*PLSB (GI: 159473711), *Chlamydomonas reinhardtii*.

Figure 2. Alignment of *Ha*PLSB with squash (*Cucurbita moschata*) GPAT. The secondary structure elements from modeled sunflower GPAT are represented by cylinders (α helices) and arrows (β strands) above the aligned sequences. The black arrow indicates the putative processing site of the transit peptide. The $H(X)_4D$ motif conserved among glycerol-3-phosphate acyltransferase enzymes is underlined. Identical amino acids are shaded in black.

Figure 3. Modelled tertiary structure of *Ha*PLSB. (A) Ribbon diagram. Helices and strands are colored in succession from blue at the amino terminus to red at the carboxyl terminus. (B) Surface computing. Predicted patches to attract and bind both the G3P and acyl molecules are highlighted in blue and yellow respectively; L339 marked in brown may play a role in acyl selectivity. Substrate binding site is framed. (C) Framed binding site with 20% transparency showing the best scored positions for ligands G3P (sticks) and oleyl-CoA (spheres) predicted by docking experiments. (D) Rotated view of the cluster interacting with the docked G3P molecule at maximum distance of 4 Å. The atom type color scheme is the following: carbon, white; oxygen, red; nitrogen, blue; phosphate, orange.

Figure 4. Expression of plastid GPAT gene in *Helianthus annuus* and *Arabidopsis thaliana* tissues at different developmental stages. (A) *HaPLSB* gene expression data normalized with *HaACT1*. (B) *Arabidopsis thaliana ATS1* gene data from microarray [33]. DAF: days after flowering; DAG: days after germination; Stage 3: mid-globular to early heart embryos; Stage 4: early to late heart embryos; Stage 5: late heart to mid-torpedo embryos; Stage 6: mid to late torpedo embryos; Stage 7: late torpedo to early walking-stick embryos; Stage 8: walking-stick

to early curled cotyledons embryos; Stage 9: curled cotyledons to early green cotyledon embryos. Leaves 1 + 2 are aged 7 DAG. The data in panel A represent the mean values ± SD of three independent seedling samples.

Figure 5. Coomasie Brilliant Blue stained SDS-PAGE of recombinant sunflower GPAT purified by affinity column after heterologous expression. Lane 1, soluble fraction of pQE80::*Ha*PLSB transformed *E. coli* lysate. Lane 2, purified His-tagged *Ha*PLSB protein. *Ha*PLSB represents 6.9% of the soluble fraction and 84.0% after purification, as estimated using ImageJ (http://imagej.nih.gov/ij/; Schneider et al., 2012).

Figure 6. *Ha*PLSB was assayed to determine the specificity and kinetic parameters using 16:0 (○), 18:0 (□) and 18:1 (△) acyl-ACPs as substrates. Enzymatic activity is expressed as the amount of radiolabelled LPA formed in reactions as a function of the acyl-ACP concentration (A), G3P concentration (B) or pH (C). Values are means ± S.D.

Table 1. Oligonucleotides used in this study. [a] Restriction sites are indicated in bold.

Primer name	Sequence 5' → 3'
HaPLSB-F1	AGAGAGGTAGAAGCTGGAGCAC
HaPLSB-R1	TGATTTGTGACGGGATCAGG
FA2Z	AACTGGAAGAATTCGCGG
HaPLSB-F2	TGAACGATGTTCCTGAGCTTGC
HaPLSB-R2	TGCTTCGGATTGGTGGTTTGAG
HaPLSB-F	TGTCGATTCTCCCGTCTTCTTCTCC
HaPLSB-R	TCAACAGGTTGCGACAATGAGACAC
pQE_HaPLSB-F [a]	ATTTT**GCATGC**GCGGAGACGTTTGAAGGC
pQE_HaPLSB-R [a]	T**GGTACC**CCTGCTGATATACTAGTCAACAGG
QHaPLSB-F	TGAACGATGTTCCTGAGCTTG
QHaPLSB-R	TGATTTGTGACGGGATCAGG
QHaAct-F4	GCTAACAGGGAAAAGATGACTC
QHaAct-R4	ACTGGCATAAAGAGAAAGCACG

Table 2. Fatty acid composition (mol%) of bacterial cells that do or do not produce the GPAT protein. Cells were harvested at OD = 2. The data correspond to average values (<10% variation). 16:0, palmitic acid; 16:1, palmitoleic acid; 17:0Δ, *cis*-9,10-methylenehexadecanoic acid; 18:0, stearic acid; 18:1, cis-vaccenate; 19:0Δ, *cis*-9,10-methyleneoctadecanoic acid.

Fatty acid composition (mol%)							
	16:0	16:1	17:0Δ	18:0	18:1	19:0Δ	µg FA/mg cel
Uninduced	48.17	19.07	13.03	1.07	17.57	1.09	4.79
Induced	44.14	24.83	10.70	1.57	18.26	0.50	4.23

Table 3. Apparent kinetic parameters calculated for recombinant sunflower chloroplast GPAT. [a] Values were calculated using the Hill equation (adjusted R-squared > 0.95). Assays were performed in triplicate as described in the Materials and Methods.

Substrate	K_M [a]	V_{max} [a]	k_{cat}	k_{cat}/K_M
	μM	nkat per mg GPAT	s^{-1}	$M^{-1} s^{-1}$
16:0-ACP	1.96	0.22	9.44×10^{-3}	4.82×10^{3}
18:0-ACP	6.12	0.16	6.74×10^{-3}	1.10×10^{3}
18:1-ACP	7.80	2.12	9.18×10^{-2}	1.18×10^{4}
G3P	659	0.25	1.10×10^{-2}	17

Table 4. Polar lipid composition of mature sunflower leaves. The data represent the mean of three separate measurements (<10% variation).

A. Glycerolipids (% of Total)

	PA	PC	PE	PG	PI	MGDG	DGDG	SQVDG
%	7.53	14.19	4.09	5.14	0.87	47.03	19.24	1.91

B. Positional distribution of fatty acids in galactolipids and phosphatidylglycerol (mol %) [a] 16:1(3t)

		16:0	16:1	18:0	18:1	18:2	18:3
MGDG	*sn*-1	7.8	0.1	8.8	2.0	7.7	73.7
	sn-2	2.1	0.2	3.6	0.3	10.2	83.6
DGDG	*sn*-1	25.8	0.1	15.4	3.4	5.2	50.1
	sn-2	6.1	0.2	4.0	0.4	7.5	81.8
PG	*sn*-1	27.8	1.3	19.4	7.4	17.6	26.5
	sn-2	48.1	28.7[a]	5.4	0.4	9.6	7.7

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., 1990. Basic local alignment search tool. J Mol Biol 215, 403-410.

Bertrams, M., Heinz, E., 1976. Experiments on enzymatic acylation of sn-glycerol 3-phosphate with enzyme preparations from pea and spinach leaves Planta (Berl.) 132, 161-168.

Bertrams, M., Heinz, E., 1981. Positional specificity and fatty acid selectivity of purified sn-glycerol 3-phosphate acyltransferases from chloroplasts. Plant physiology 68, 653-657.

Bradford, M. M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72, 248-254.

Cronan, J. E., Roughan, P. G., 1987. Fatty acid specificity and selectivity of the chloroplast sn-glycerol 3-phosphate acyltransferase of the chilling sensitive plant, Amaranthus lividus. Plant physiology 83, 676-680.

Emanuelsson, O., Nielsen, H., Brunak, S., von Heijne, G., 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300, 1005-1016.

Frentzen, M., Nishida, I., Murata, N., 1987. Properties of the plastidial acyl-(acyl-carrier-protein): glycerol-3-phosphate acyltransferase from the chilling-sensitive plant squash (Cucurbita moschata). Plant Cell Physiol. 28, 1195–1201.

Hall, T. A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41, 95-98.

Hara, A., Radin, N. S., 1978. Lipid extraction of tissues with a low-toxicity solvent. Analytical biochemistry 90, 420-426.

Heath, R. J., Jackowski, S., Rock, C. O., 1994. Guanosine tetraphosphate inhibition of fatty acid and phospholipid synthesis in Escherichia coli is relieved by overexpression of glycerol-3-phosphate acyltransferase (plsB). J Biol Chem 269, 26584-26590.

Kunst, L., Browse, J., Somerville, C., 1988. Altered regulation of lipid biosynthesis in a mutant of Arabidopsis deficient in chloroplast glycerol-3-phosphate acyltransferase activity. Proc Natl Acad Sci U S A 85, 4143-4147.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., Higgins, D. G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947-2948.

Livak, K. J., Schmittgen, T. D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-408.

Moellering, E. R., Muthan, B., Benning, C., 2010. Freezing tolerance in plants requires lipid remodeling at the outer chloroplast membrane. Science 330, 226-228.

Mongrand, S., Bessoule, J.-J., Cabantous, F., Cassagne, C., 1998. The C16:3\C18:3 fatty acid balance in photosynthetic tissues from 468 plant species. Phytochemistry 49, 1049-1064.

Moreau, P., Bessoule, J. J., Mongrand, S., Testet, E., Vincent, P., Cassagne, C., 1998. Lipid trafficking in plant cells. Prog Lipid Res 37, 371-391.

Murata, N., Los, D. A., 1997. Membrane fluidity and temperature perception. Plant physiology 115, 875-879.

Murata, N., Sato, N., Takahashi, N., Hamazaki, Y., 1982. Compositions and positional distributions of fatty acids in phospholipids from leaves of chilling-sensitive and chilling-resistant plants. Plant and Cell Physiology 23, 1071-1079.

Nishida, I., Tasaka, Y., Shiraishi, H., Murata, N., 1993. The gene and the RNA for the precursor to the plastid-located glycerol-3-phosphate acyltransferase of Arabidopsis thaliana. Plant Mol Biol 21, 267-277.

Ohlrogge, J., Browse, J., 1995. Lipid biosynthesis. Plant Cell 7, 957-970.

Roston, R. L., Gao, J., Murcha, M. W., Whelan, J., Benning, C., 2012. TGD1, -2, and -3 proteins involved in lipid trafficking form ATP-binding cassette (ABC) transporter with multiple substrate-binding proteins. The Journal of biological chemistry 287, 21406-21415.

Rock, C. O., Jackowski, S., 2002. Forty years of bacterial fatty acid synthesis. Biochemical and biophysical research communications 292, 1155-1166.

Ruiz-Lopez, N., Garces, R., Harwood, J. L., Martinez-Force, E., 2010. Characterization and partial purification of acyl-CoA:glycerol 3-phosphate acyltransferase from sunflower (Helianthus annuus L.) developing seeds. Plant Physiol Biochem 48, 73-80.

Sanchez-Garcia, A., Moreno-Perez, A. J., Muro-Pastor, A. M., Salas, J. J., Garces, R., Martinez-Force, E., 2010. Acyl-ACP thioesterases from castor (Ricinus communis L.): an enzymatic system appropriate for high rates of oil synthesis and accumulation. Phytochemistry 71, 860-869.

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., Lohmann, J. U., 2005. A gene expression map of Arabidopsis thaliana development. Nat Genet 37, 501-506.

Schneider, C. A., Rasband, W. S., Eliceiri, K. W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nature methods 9, 671-675.

Shen, J., Zeng, Y., Zhuang, X., Sun, L., Yao, X., Pimpl, P., Jiang, L., 2013. Organelle pH in the Arabidopsis endomembrane system. Mol Plant.

Slabas, A. R., Kroon, J. T., Scheirer, T. P., Gilroy, J. S., Hayman, M., Rice, D. W., Turnbull, A. P., Rafferty, J. B., Fawcett, T., Simon, W. J., 2002. Squash glycerol-3-phosphate (1)-acyltransferase. Alteration of substrate selectivity and identification of arginine and lysine residues important in catalytic activity. J Biol Chem 277, 43918-43923.

Soll, J., Roughan, G., 1982. Acyl-acyl carrier protein and pool sizes during steady-state fatty acid synthesis by isolated spinach chloroplasts. FEBS Letters 146, 189-192.

Sui, N., Li, M., Zhao, S. J., Li, F., Liang, H., Meng, Q. W., 2007. Overexpression of glycerol-3-phosphate acyltransferase gene improves chilling tolerance in tomato. Planta 226, 1097-1108.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28, 2731-2739.

Trott, O., Olson, A. J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry 31, 455-461.

Turnbull, A. P., Rafferty, J. B., Sedelnikova, S. E., Slabas, A. R., Schierer, T. P., Kroon, J. T., Simon, J. W., Fawcett, T., Nishida, I., Murata, N., Rice, D. W., 2001. Analysis of the structure,

substrate specificity, and mechanism of squash glycerol-3-phosphate (1)-acyltransferase. Structure 9, 347-353.

Whitaker, B. D., 1986. Fatty-acid composition of polar lipids in fruit and leaf chloroplasts of "16:3"- and "18:3"-plant species. Planta 169, 313-319.

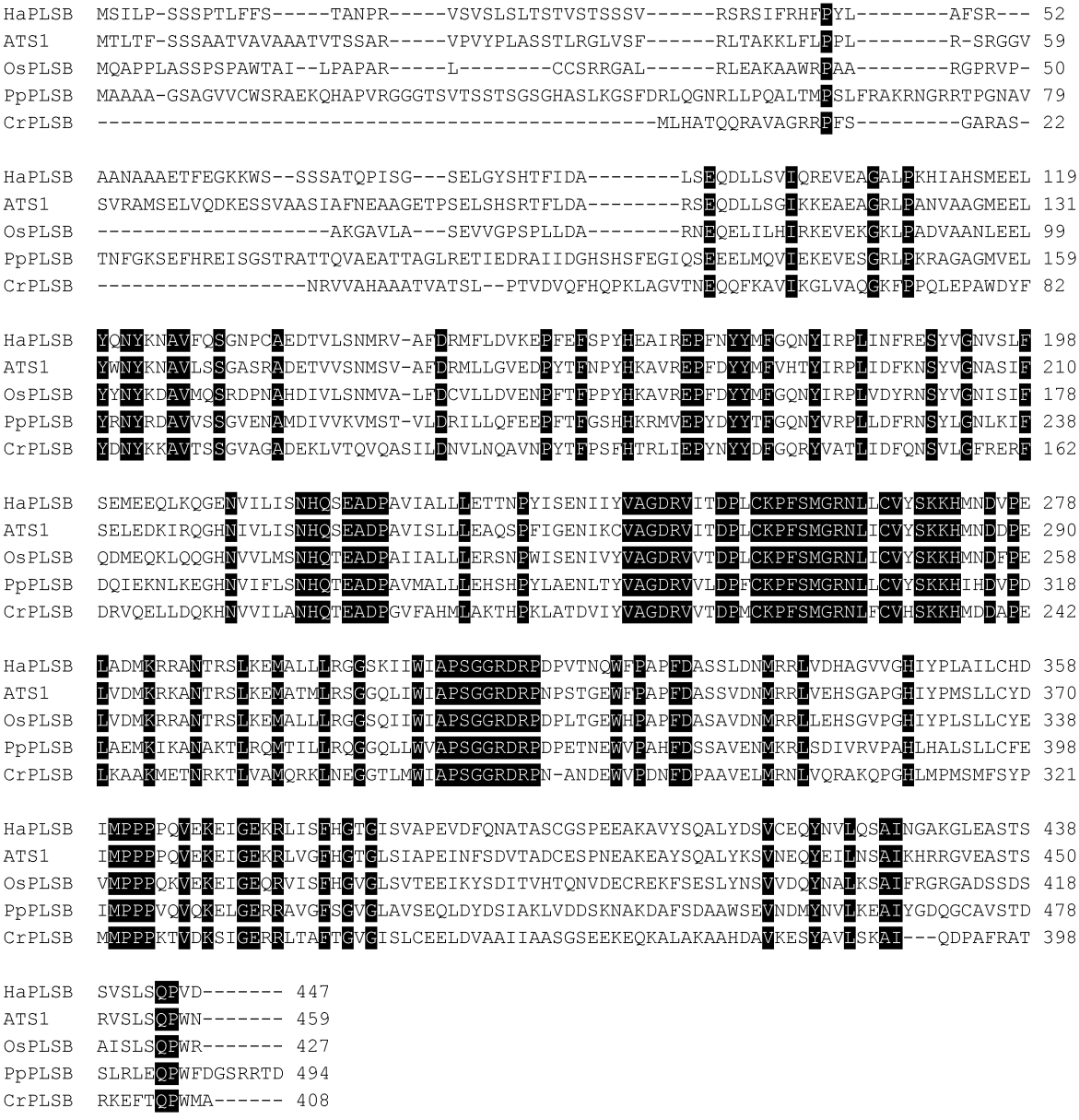
Xu, C., Yu, B., Cornish, A. J., Froehlich, J. E., Benning, C., 2006. Phosphatidylglycerol biosynthesis in chloroplasts of Arabidopsis mutants deficient in acyl-ACP glycerol-3- phosphate acyltransferase. Plant J 47, 296-309.

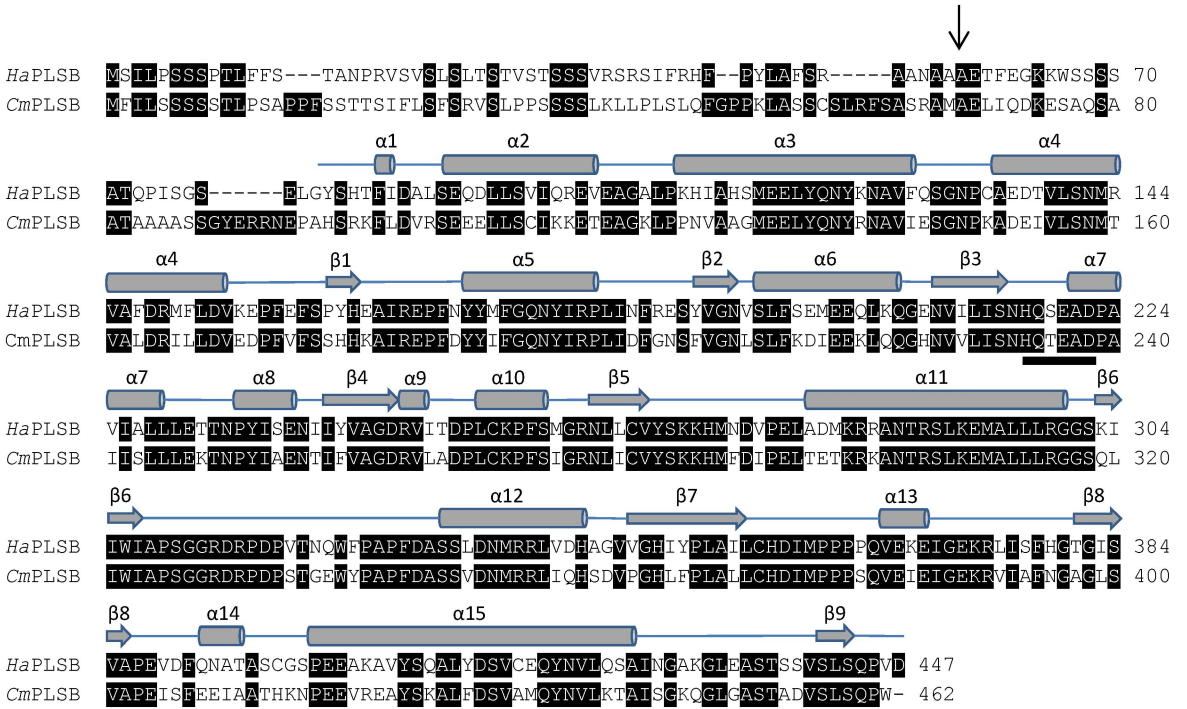
Yang, W., Simpson, J. P., Li-Beisson, Y., Beisson, F., Pollard, M., Ohlrogge, J. B., 2012. A land-plant-specific glycerol-3-phosphate acyltransferase family in Arabidopsis: substrate specificity, sn-2 preference, and evolution. Plant physiology 160, 638-652.

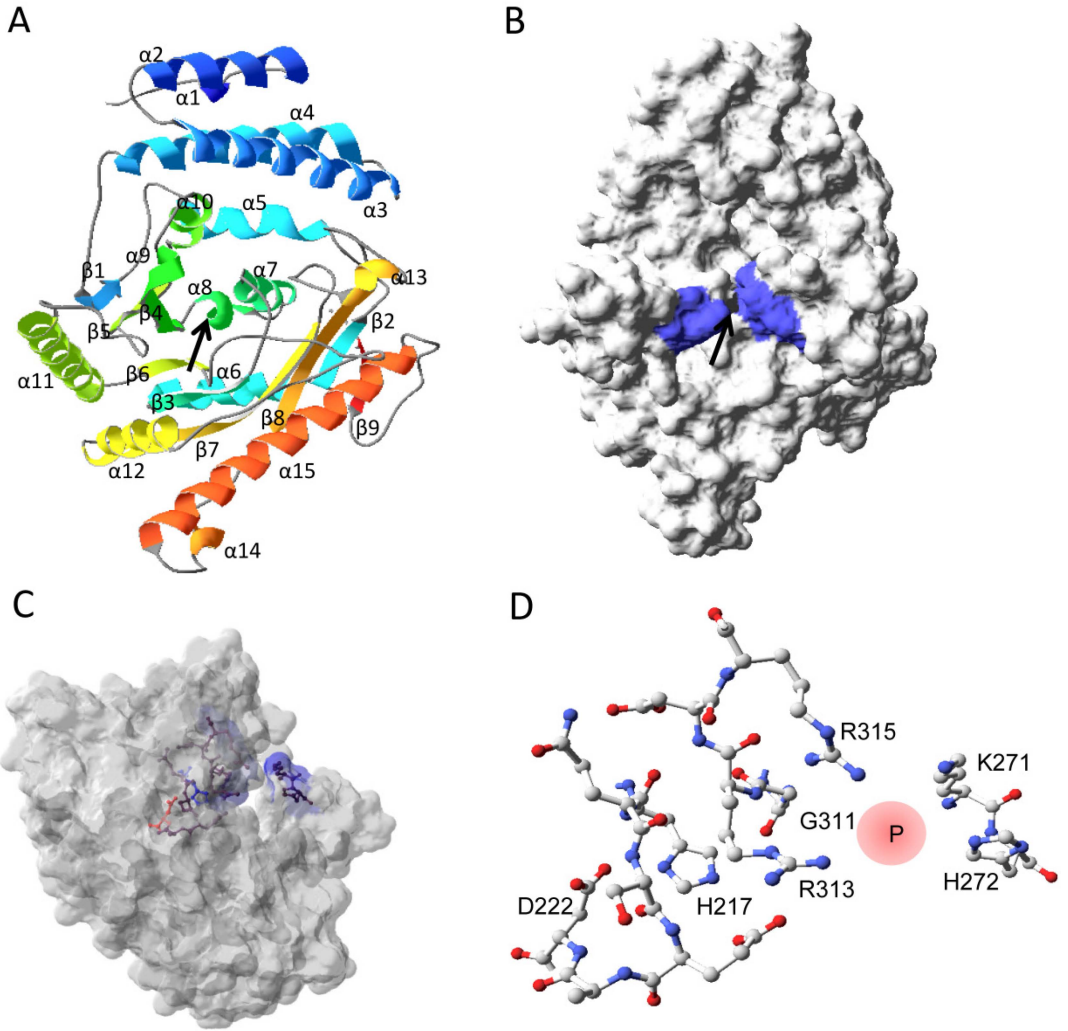
Zhu, S. Q., Zhao, H., Zhou, R., Ji, B. H., Dan, X. Y., 2009. Substrate selectivity of glycerol-3-phosphate acyl transferase in rice. J Integr Plant Biol 51, 1040-1049.

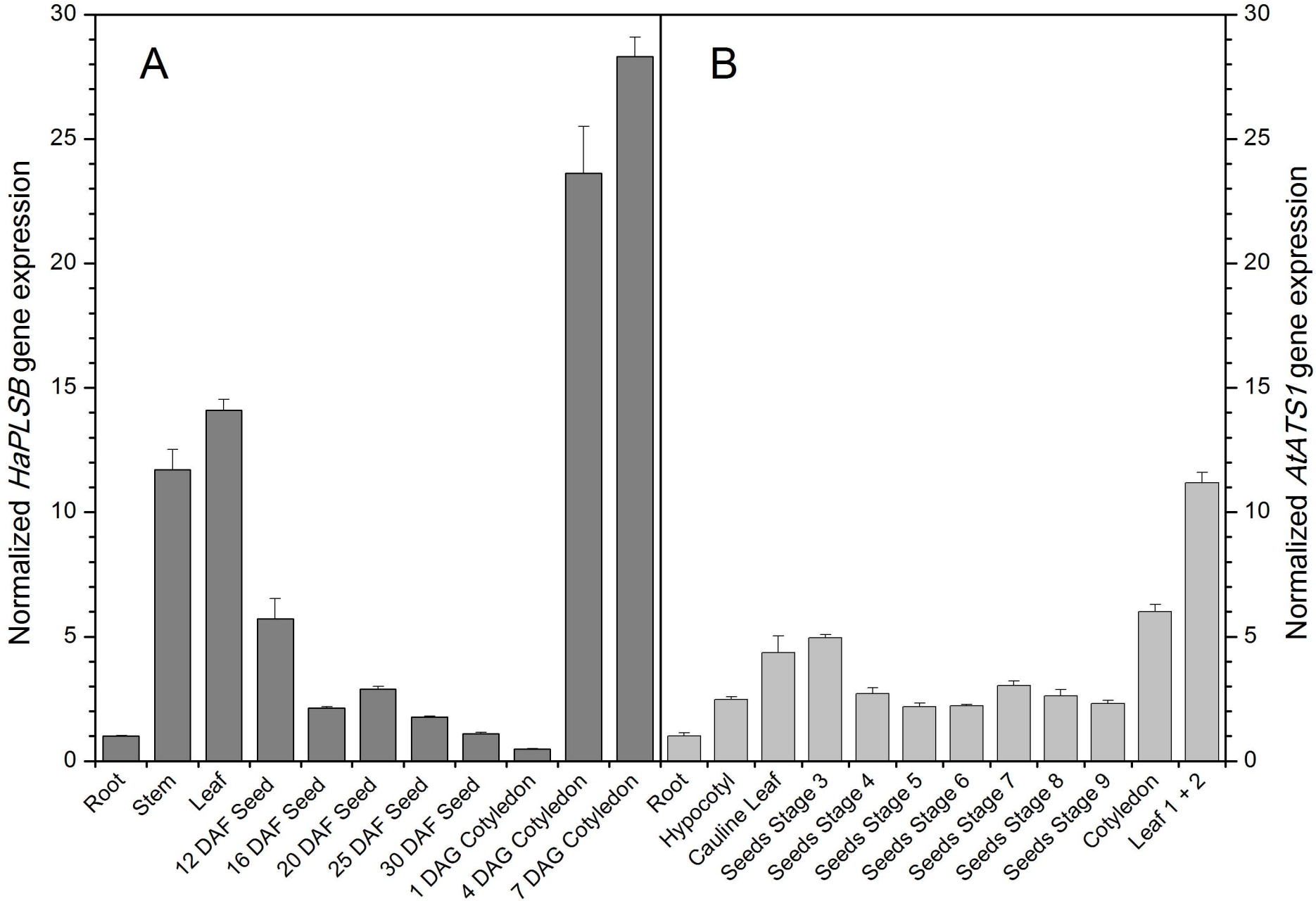
Appendix A. Supplem**entary data**

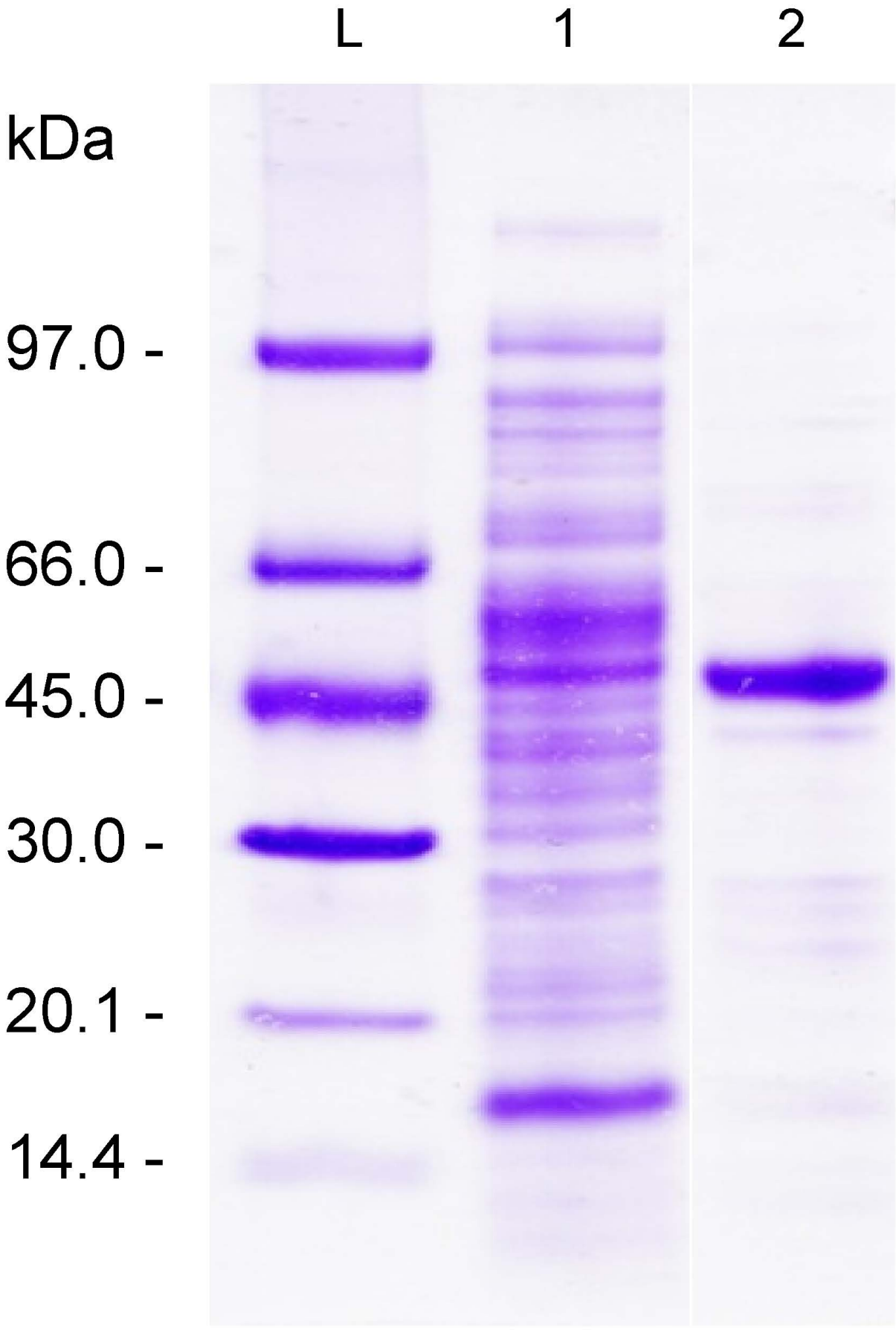
Supplementary Figure 1. Phylogenetic study of chloroplast GPAT enzymes using GPATs from *Chlamydia*-like bacteria as outgroup to root the tree. The tree was constructed using the neighbor-joining method.


HaPLSB MSILP-SSSPTLFFS-----TANPR-----VSVSLSLTSTVSTSSSV------RSRSIFRHFPYL--------AFSR--- 52
ATS1 MTLTF-SSSAATVAVAAATVTSSAR-----VPVYPLASSTLRGLVSF------RLTAKKLFLPPL--------R-SRGGV 59
OsPLSB MQAPPLASSPSPAWTAI--LPAPAR-----L--------CCSRRGAL------RLEAKAAWRPAA--------RGPRVP- 50
PpPLSB MAAAA-GSAGVVCWSRAEKQHAPVRGGGTSVTSSTSGSGHASLKGSFDRLQGNRLLPQALTMPSLFRAKRNGRRTPGNAV 79
CrPLSB ----------MLHATQQRAVAGRRPFS---------GARAS- 22
HaPLSB AANAAAETFEGKKWS--SSSATQPISG---SELGYSHTFIDA--------LSEQDLLSVIQREVEAGALPKHIAHSMEEL 119
ATS1 SVRAMSELVQDKESSVAASIAFNEAAGETPSELSHSRTFLDA--------RSEQDLLSGIKKEAEAGRLPANVAAGMEEL 131
OsPLSB --------------------AKGAVLA---SEVVGPSPLLDA--------RNEQEILHIRKEVEKGKLPADVAANLEEL 99
PpPLSB TNFGKSEFHREISGSTRATTQVAEATTAGLRETIEDRAIIDGHSHSFEGIQSEEELMQVTEKEVESGRLPKRAGAGMVEL 159
CrPLSB ------------------NRVVAHAAATVATSL--PTVDVQFHQPKLAGVTNEQQFKAVIKGLVAQGKFPPQLEPAWDYF 82
HaPLSB YQNYKNAVFQSGNPCAEDTVLSNMRV-APDRMFLDVKEPFEFSPYHEAIREPFNYYMFGQNYIRPLINFRESYVGNVSLF 198
ATS1 YWNYKNAVLSSGASRADETVVSNMSV-APDRMLLGVEDPYTFNPYHKAVREPFDYYMFVHTYIRPLIDFKNSYVGNASIF 210
OsPLSB YYNYKDAVMQSRDPNAHDIVLSNMVA-LPDCVLLDVENPFTFPPYHKAVREPFDYYMFGQNYIRPLVDYRNSYVGNISIF 178
PpPLSB YRNYRDAVVSSGVENAMDIVVKVMST-VLDRILLQFEEPFTFGSHHKRMVEPYDYYTFGQNYVRPLLDFRNSYLGNLKIF 238
CrPLSB YDNYKKAVTSSGVAGADEKLVTQVQASILDNVLNQAVNPYTFPSFHTRLIEPYNYYDFGQRYVATLIDFQNSVLGFRERF 162
HaPLSB SEMEEQLKQGENVILISNHQSEADPAVIALLLETTNPYISENIIYVAGDRVITDPLCKPFSMGRNLLCVYSKKHMNDVPE 278
ATS1 SELEDKIRQGHNIVLISNHQSEADPAVISLLLEAQSPFIGENIKCVAGDRVITDPLCKPFSMGRNLICVYSKKHMNDDPE 290
OsPLSB QDMEQKLQQGHNVVLMSNHQTEADPAIIALLLERSNPWISENIVYVAGDRVVTDPLCKPFSMGRNLICVYSKKHMNDFPE 258
PpPLSB DQIEKNLKEGHNVIFLSNHQTEADPAVMALLLEHSHPYLAENLTYVAGDRVVLDPFCKPFSMGRNLLCVYSKKHIHDVPD 318
CrPLSB DRVQELLDQKHNVVILANHQTEADPGVFAHMLAKTHPKLATDVIYVAGDRVVTDPMCKPFSMGRNLFCVHSKKHMDDAPE 242
HaPLSB LADMKRRANTRSLKEMALLLRGGSKIIWIAPSGGRDRPDPVTNQWFPAPFDASSLDNMRRLVDHAGVVGHIYPLAILCHD 358
ATS1 LVDMKRKANTRSLKEMATMLRSGGQLIWIAPSGGRDRPNPSTGEWFPAPFDASSVDNMRRLVEHSGAPGHIYPMSLLCYD 370
OsPLSB LVDMKRRANTRSLKEMALLLRGGSQIIWIAPSGGRDRPDPLTGEWHPAPFDASAVDNMRRLLEHSGVPGHIYPLSLLCYE 338
PpPLSB LAEMKIKANAKTLRQMTILLRQGGQLLWVAPSGGRDRPDPETNEWVPAHFDSSAVENMKRLSDIVRVPAHLHALSLLCFE 398
CrPLSB LKAAKMETNRKTLVAMQRKLNEGGTLMWIAPSGGRDRPN-ANDEWVPDNFDPAAVELMRNLVQRAKQPGHLMPMSMFSYP 321
HaPLSB IMPPPPQVEKEIGEKRLISEHGTGISVAPEVDFQNATASCGSPEEAKAVYSQALYDSVCEQYNVLQSAINGAKGLEASTS 438
ATS1 IMPPPPQVEKEIGEKRLVGEHGTGLSIAPEINFSDVTADCESPNEAKEAYSQALYKSVNEQYEILNSAIKHRRGVEASTS 450
OsPLSB VMPPPQRVEKEIGEQRVISEHGVGLSVTEEIKYSDITVHTQNVDECREKFSESLYNSVVDQYNALKSAIFRGRGADSSDS 418
PpPLSB IMPPPVQVQKELGERRAVGESGVGLAVSEQLDYDSIAKLVDDSKNAKDAFSDAAWSEVNDMYNVLKEAIYGDQGCAVSTD 478
CrPLSB MMPPPKTVDKSIGERRLTAETGVGISLCEELDVAAIIAASGSEEKEQKALAKAAHDAVKESYAVLSKAI---QDPAFRAT 398
HaPLSB SVSLSQPVD------- 447
ATS1 RVSLSQPWN------- 459
OsPLSB AISLSQPWR------- 427
PpPLSB SLRLEQPWFDGSRRTD 494
CrPLSB RKEFTQPWMA------ 408

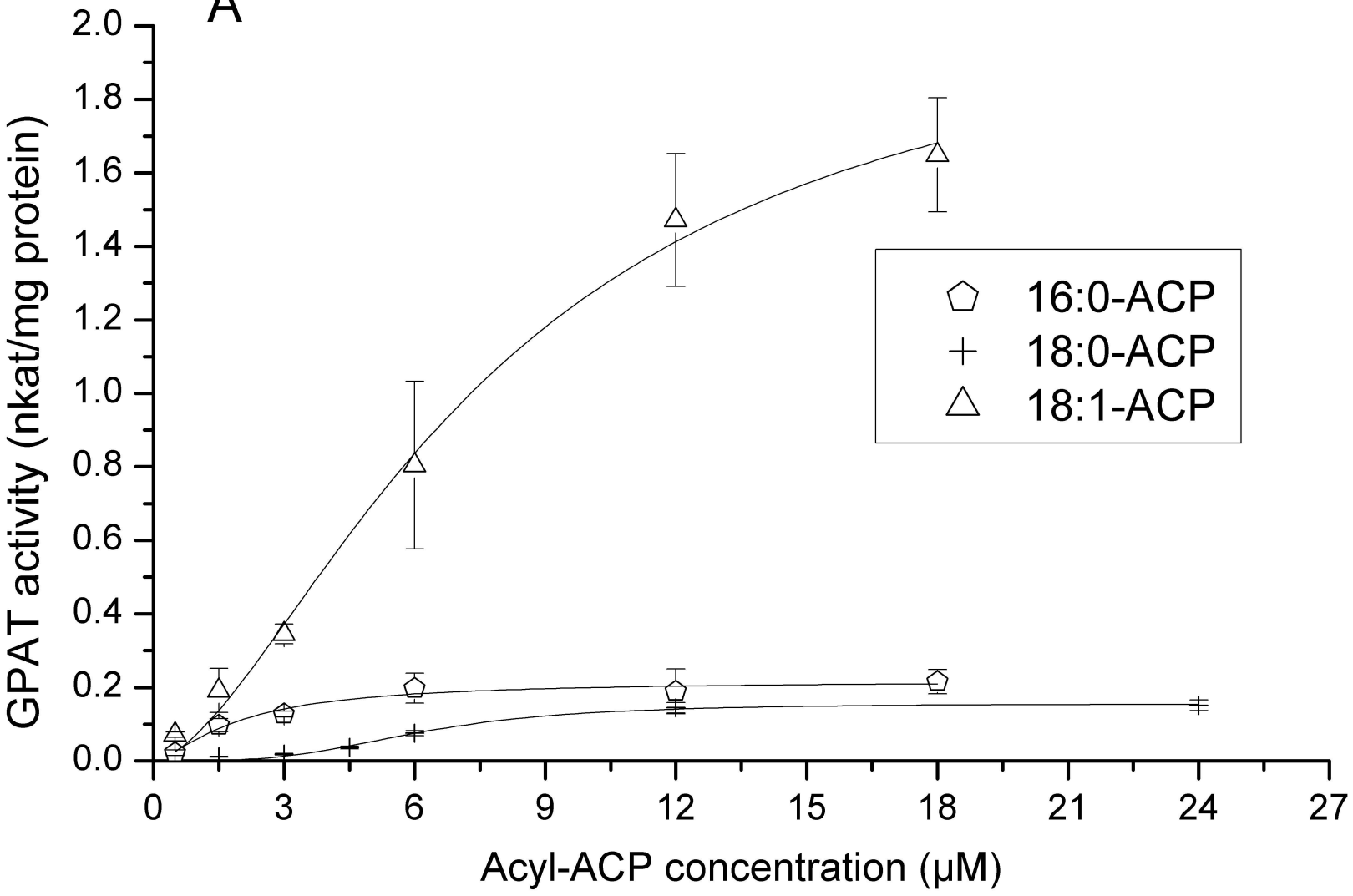

HaPLSB MSIIPSSSPTLFFS---TANPRVSVSLSLTSTVSTSSSVRSRSIFRHE--PYLAFSR-----AANAAAETFEGKKWSSSS 70
CmPLSB MFILSSSSSTLPSAPPFSSTTSIFLSFSRVSLPPSSSSLKLLPLSLQFGPPKLASSCSLRFSASRAMAELIQDKESAQSA 80
α1 α2 α3 α4
HaPLSB ATQPISGS------ELGYSHTFIDALSEQDLLSVIQREVEAGALPKHIAHSMEELYQNYKNAVFQSGNPCAEDTVLSNMR 144
CmPLSB ATAAAASSGYERRNEPAHSRKFLDVRSEEELLSCIKKETEAGKLPPNVAAGMEELYQNYRNAVIESGNPKADEIVLSNMT 160
α4 β1 α5 β2 α6 β3 α7
HaPLSB VAFDRMFLDVKEPFEFSPYHEAIREPFNYYMFGQNYIRPLINFRESYVGNVSLFSEMEEQLKQGENVILISNHQSEADPA 224
CmPLSB VALDRILLDVEDPFVFSSHHKAIREPFDYYIFGQNYIRPLIDFGNSFVGNLSLFKDIEEKLQQGHNVVLISNHQTEADPA 240
α7 α8 β4 α9 α10 β5 α11 β6
HaPLSB VIALLLETTNPYISENIIYVAGDRVITDPLCKPFSMGRNLLCVYSKKHMNDVPELADMKRRANTRSLKEMALLLRGGSKI 304
CmPLSB IISLLLEKTNPYIAENTIFVAGDRVLADPLCKPFSIGRNLICVYSKKHMFDIPELTETKRKANTRSLKEMALLLRGGSQL 320
β6 α12 β7 α13 β8
HaPLSB IWIAPSGGRDRPDPVTNQWFPAPFDASSLDNMRRLVDHAGVVGHIYPLAILCHDIMPPPPQVEKEIGEKRLISFHGTGIS 384
CmPLSB IWIAPSGGRDRPDPSTGEWYPAPFDASSVDNMRRLIQHSDVPGHLFPLALLCHDIMPPPSQVEIEIGEKRVIAFNGAGLS 400
β8 α14 α15 β9
HaPLSB VAPEVDFQNATASCGSPEEAKAVYSQALYDSVCEQYNVLQSAINGAKGLEASTSSVSLSQPVD 447
CmPLSB VAPEISFEEIAATHKNPEEVREAYSKALFDSVAMQYNVLKTAISGKQGLGASTADVSLSQPW- 462


A
α2
α1
α4
α3
α10
α5
β1
α9
α7
α13
α8
β4
β2
β5
α11
β6
α6
β3
β8
β9
β7
α12
α15
α14
B
C
D
R315
K271
G311
P
R313
H272
D222
H217

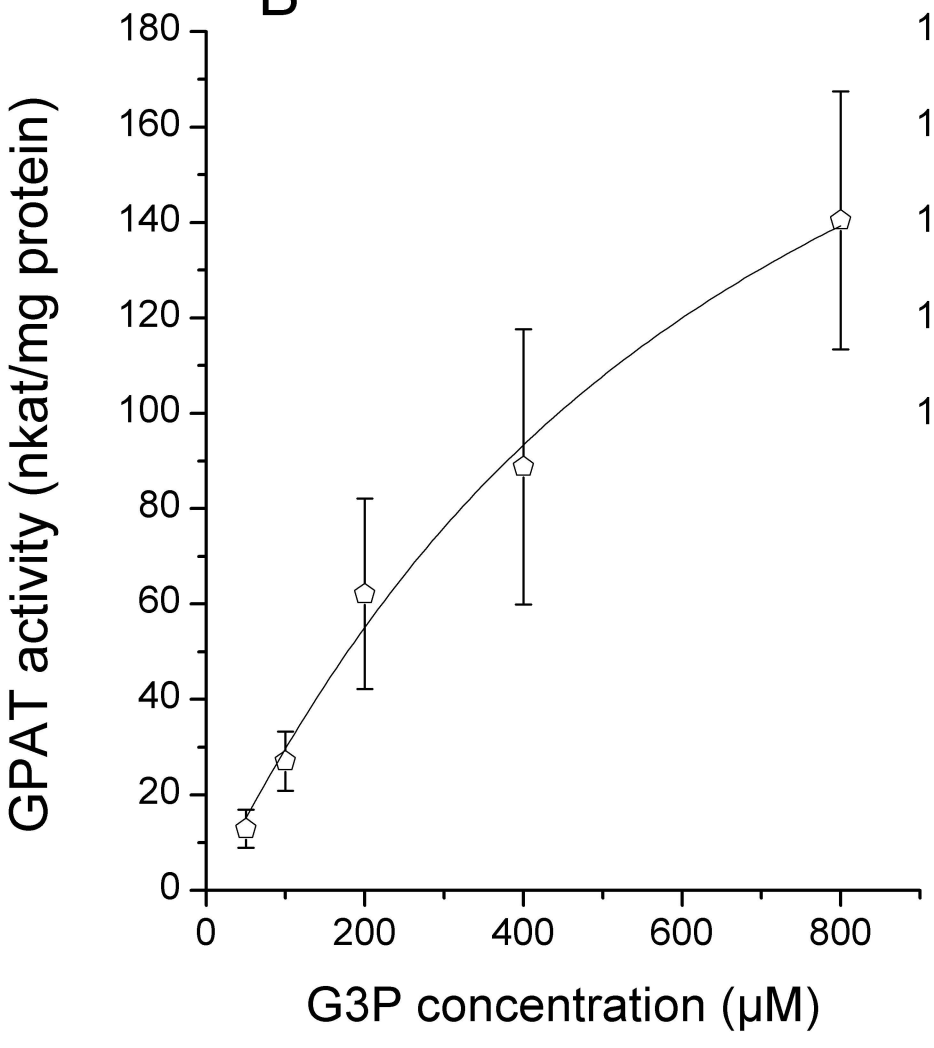

A
B
Normalized *HaPLSB* gene expression
Normalized *AtATS1* gene expression
0
5
10
15
20
25
30
Root
Stem
Leaf
12 DAF Seed
16 DAF Seed
20 DAF Seed
25 DAF Seed
30 DAF Seed
1 DAG Cotyledon
4 DAG Cotyledon
7 DAG Cotyledon
Root
Hypocotyl
Cauline Leaf
Seeds Stage 3
Seeds Stage 4
Seeds Stage 5
Seeds Stage 6
Seeds Stage 7
Seeds Stage 8
Seeds Stage 9
Cotyledon
Leaf 1 + 2


L
1
2
kDa
97.0 -
66.0 -
45.0 -
30.0 -
20.1 -
14.4 -

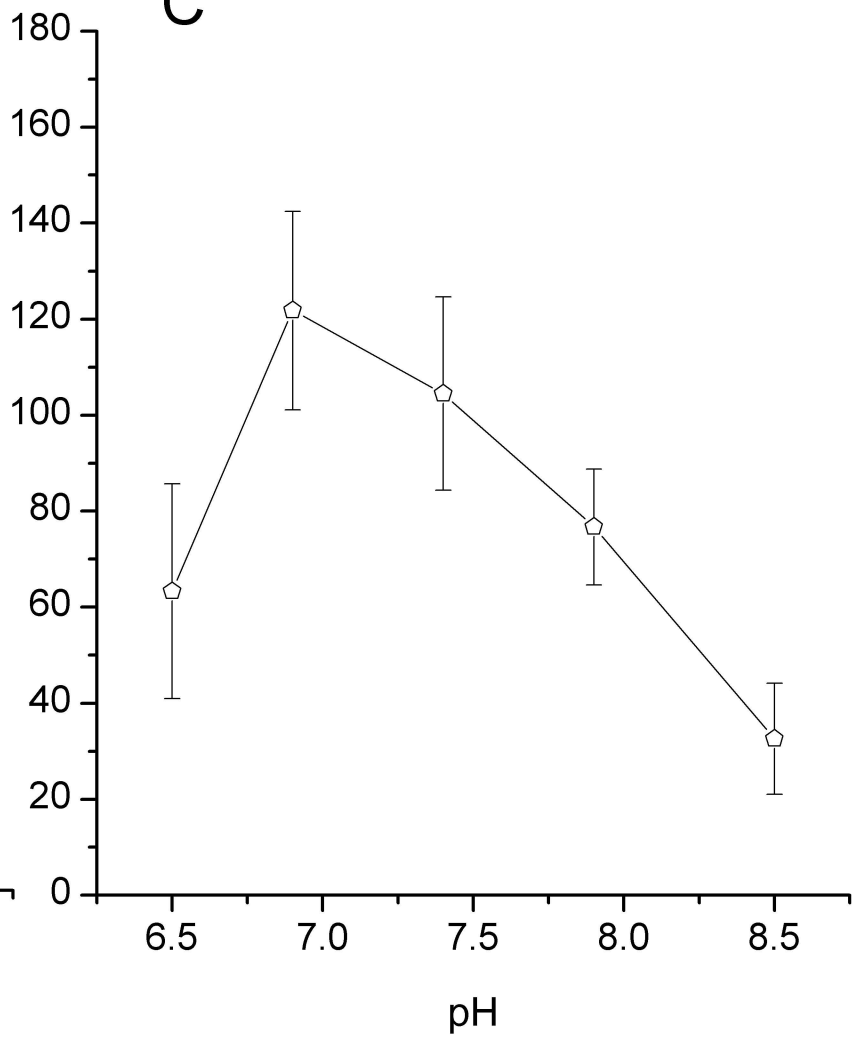
A

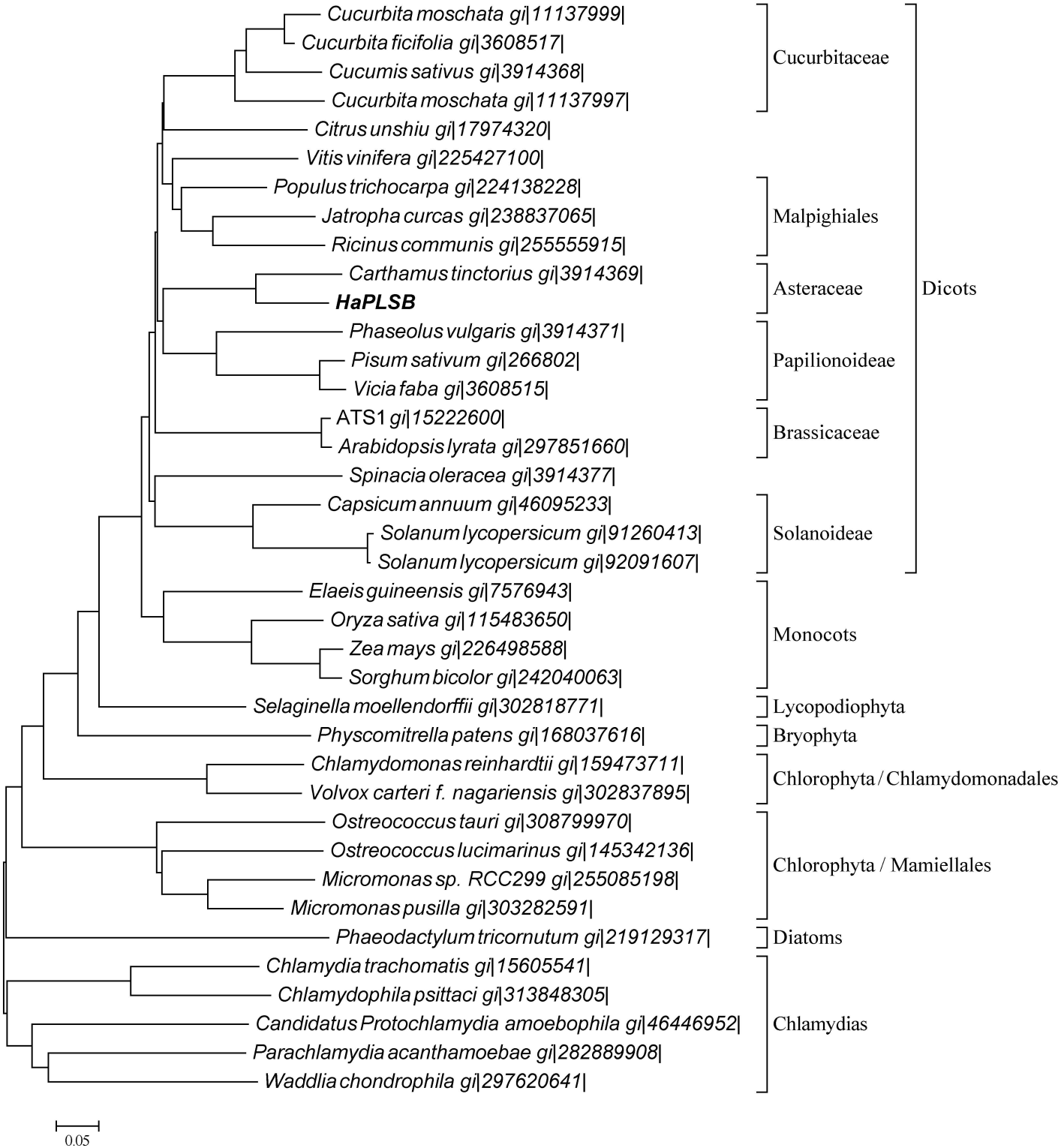

GPAT activity (nkat/mg protein)
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
3
6
9
12
15
18
21
24
27
Acyl-ACP concentration (μM)
16:0-ACP
18:0-ACP
18:1-ACP

B


GPAT activity (nkat/mg protein)
180
160
140
120
100
80
60
40
20
0
0
200
400
600
800
G3P concentration (μM)

C


180
160
140
120
100
80
60
40
20
0
6.5
7.0
7.5
8.0
8.5
pH


Cucurbita moschata gi|11137999|
Cucurbita ficifolia gi|3608517|
Cucumis sativus gi|3914368|
Cucurbita moschata gi|11137997|
Citrus unshiu gi|17974320|
Vitis vinifera gi|225427100|
Populus trichocarpa gi|224138228|
Jatropha curcas gi|238837065|
Ricinus communis gi|255555915|
Carthamus tinctorius gi|3914369|
HaPLSB
Phaseolus vulgaris gi|3914371|
Pisum sativum gi|266802|
Vicia faba gi|3608515|
ATS1 gi|15222600|
Arabidopsis lyrata gi|297851660|
Spinacia oleracea gi|3914377|
Capsicum annuum gi|46095233|
Solanum lycopersicum gi|91260413|
Solanum lycopersicum gi|92091607|
Elaeis guineensis gi|7576943|
Oryza sativa gi|115483650|
Zea mays gi|226498588|
Sorghum bicolor gi|242040063|
Selaginella moellendorffii gi|302818771|
Physcomitrella patens gi|168037616|
Chlamydomonas reinhardtii gi|159473711|
Volvox carteri f. nagariensis gi|302837895|
Ostreococcus tauri gi|308799970|
Ostreococcus lucimarinus gi|145342136|
Micromonas sp. RCC299 gi|255085198|
Micromonas pusilla gi|303282591|
Phaeodactylum tricornutum gi|219129317|
Chlamydia trachomatis gi|15605541|
Chlamydophila psittaci gi|313848305|
Candidatus Protochlamydia amoebophila gi|46446952|
Parachlamydia acanthamoebae gi|282889908|
Waddlia chondrophila gi|297620641|
Cucurbitaceae
Malpighiales
Asteraceae
Papilionoideae
Brassicaceae
Solanoideae
Dicots
Monocots
Lycopodiophyta
Bryophyta
Chlorophyta / Chlamydomonadales
Chlorophyta / Mamiellales
Diatoms
Chlamydias
0.05