A dual-targeted purple acid phosphatase in *Arabidopsis thaliana* moderates carbon metabolism and its overexpression leads to faster plant growth and higher seed yield

Feng Sun¹, Pui Kit Suen¹, Youjun Zhang¹, Chao Liang¹, Chris Carrie², James Whelan², Jane L. Ward³, Nathaniel D. Hawkins³, Liwen Jiang⁴, Boon Leong Lim¹,*

¹School of Biological Sciences, the University of Hong Kong, Pokfulam, Hong Kong, China.
²Australian Research Council Centre of Excellence in Plant Energy Biology, University of Western Australia, Crawley WA 6009, Australia.
³National Centre for Plant and Microbial Metabolomics, Rothamsted Research, West Common, Harpenden, Herts, AL5 2JQ, United Kingdom
⁴School of Life Sciences, Centre for Cell and Developmental Biology, the Chinese University of Hong Kong.

Author for correspondence:
*Boon Leong Lim*

Tel: +852 22990319
Email: bllim@hku.hk

Word count for main body of the text: 6237
Word count for Introduction: 452
Word count for Materials and Methods: 2035
Word count for Results: 2384
Word count for Discussion: 1289
Word count for Acknowledgements: 77
The number of figures: 6
The number of tables: 2
The number of supporting figures: 11
The number of supporting tables: 3
Summary

- Overexpression of AtPAP2, a purple acid phosphatase (PAP) with a unique C-terminal hydrophobic motif in Arabidopsis, resulted in earlier bolting and a higher seed yield. Metabolite analysis showed that the shoots of AtPAP2 overexpression lines contained higher levels of sugars and tricarboxylic acid metabolites. Enzyme assays showed that sucrose phosphate synthase (SPS) activity was significantly upregulated in the overexpression lines. The higher SPS activity arose from a higher level of SPS protein, and was independent of SnRK1.

- AtPAP2 was found to be targeted to both plastids and mitochondria via its C-terminal hydrophobic motif. Ectopic expression of a truncated AtPAP2 without this C-terminal motif in Arabidopsis indicated that the subcellular localization of AtPAP2 is essential for its biological actions.

- Plant PAPs are generally considered to mediate phosphorus acquisition and redistribution. AtPAP2 is the first PAP shown to modulate carbon metabolism and the first shown to be dual-targeted to both plastids and mitochondria by a C-terminal targeting signal.

- One PAP-like sequence carrying a hydrophobic C-terminal motif could be identified in the genome of the smallest free-living photosynthetic eukaryote, Ostreococcus tauri. This might reflect a common ancestral function of AtPAP2-like sequences in the regulation of carbon metabolism.

**Key words:** purple acid phosphatase, sugars, flowering, sucrose phosphate synthase, TCA, chloroplasts, mitochondria
Introduction

Modification of carbon metabolism is a prime target for maximizing crop productivity (Sharma-Natu & Ghildiyal, 2005). Sugars not only represent an energy product from photosynthesis, but also play roles in complex cellular signaling pathways (Rolland et al., 2002; Gibson, 2005; Rolland et al., 2006), which are mediated by a variety of protein kinases (PKs), and protein phosphatases (PPs) (Baena-Gonzalez & Sheen, 2008; Smeekens et al., 2010; Zheng et al., 2010). In Arabidopsis thaliana, SNF1-related kinase (SnRK1) is a central resource regulator in response to energy deficiency (Baena-Gonzalez et al., 2007; Baena-Gonzalez & Sheen, 2008). By phosphorylating and inactivating sucrose phosphate synthase (SPS), nitrate reductase (NR), and a few other metabolic enzymes (Lunn & MacRae, 2003; Polge et al., 2008), energy and resources were reallocated from anabolic pathways (e.g. sucrose synthesis and nitrogen assimilation) to catabolic pathways (Smeekens et al., 2010). In Arabidopsis, geminivirus Rep interacting kinases (GRIK1 and GRIK2) activate SnRK1 by phosphorylating it at the T175 residue (Shen & Hanley-Bowdoin, 2006; Shen et al., 2009).

PPs regulate a variety of genes involved in sugar metabolism, such as those encoding the two major amylases of the tuberous root of sweet potato (Takeda et al., 1994), the ADP-glucose pyrophosphorylase subunit in sweet potato and Arabidopsis (Takeda et al., 1994; Siedlecka et al., 2003), and the enzymes for fructan synthesis (Martinez-Noel et al., 2009), the UDP-glucose pyrophosphorylase, and the sucrose synthase of Arabidopsis (Winter et al., 1997; Ciereszko et al., 2001; Martinez-Noel et al., 2009). Okadaic acid, a PP inhibitor, prevents activation of SPS and impairs sucrose accumulation in leaves (Huber et al., 1992). Okadaic acid also induces vacuolar acid invertase transcription, suggesting that PP activity plays a role in carbon partitioning.
Purple acid phosphatases (PAPs) represent a large group of nonspecific acid phosphatases (Schenk et al., 2000). In the Arabidopsis genome, 29 PAP genes have been identified based on sequence comparison (Zhu et al., 2005). Plant PAPs are considered to mediate phosphorus acquisition and redistribution based on their ability to hydrolyze phosphorus compounds (Cashikar et al., 1997; Lung et al., 2008; Kuang et al., 2009). Certain PAPs also exhibit peroxidase activity, such as GmPAP3 (Liao et al., 2003; Li et al., 2008) and AtPAP17 (del Pozo et al., 1999). Recent studies also showed that a tobacco PAP (NtPAP12) can modulate polysaccharide synthesis; it can upregulate β-glucan synthesis and cellulose deposition in the cell wall by dephosphorylating α-xylosidase and β-glucosidase (Kaida et al., 2009; Kaida et al., 2010).

In this study, we show that AtPAP2 is dual-targeted to mitochondria and plastids via a novel and unique C-terminal transmembrane targeting signal. Overexpression of AtPAP2 significantly enhanced the growth rate and seed yield of Arabidopsis, through increased sucrose phosphate synthase activity.
Materials and methods

Plant materials and growth conditions

WT Arabidopsis thaliana ecotype Columbia (Col-0) was used in this study. The T-DNA mutant of AT1G13900 (Salk_013567, ecotype Col-0) and AT2G03450 (Salk_129905, ecotype Col-0) were obtained from the Arabidopsis Biological Resource Center (http://abrc.osu.edu). Chilled Arabidopsis seeds were surface-sterilized with 20% (v/v) bleach for 15 minutes, washed, and plated on Murashige and Skoog (MS) medium supplemented with 2% (w/v) sucrose for 10 days. Seedlings with the same size were transferred to soil under a 16-hr light (22°C)/8-hr dark (18°C) regime (long day, LD) or an 8-hr light (22°C)/16-hr dark (18°C) regime (short day, SD) under a light intensity of 120–150 µmol m⁻² s⁻¹. The pots were placed in the growth chamber in a randomized design. Bolting time was measured when the primary inflorescence reached 1 cm above the rosette leaves. This observation of phenotype was repeated at least three times (n = 10–15).

Sequence alignment and phylogenetic analysis

Sequence data was retrieved from the TAIR website (http://www.arabidopsis.org) and the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov). Homology searches in GenBank were done using the Basic Local Alignment Search Tool server (http://www.ncbi.nlm.nih.gov/BLAST/). Multiple alignments of protein sequences were performed by MEGA 4.1 (Beta 3) software (Kumar et al., 2004) (http://www.megasoftware.net) using the Clustal X and N-J plot programs (Saitou & Nei, 1987). The amino acid sequences were aligned by CLC sequence viewer 6.3 software (http://www.clcbio.com). Protein expression
patterns were analyzed via the spot history microarray analysis tool (http://www.affymetrix.Arabidopsis.info/narrays/experimentbrowse.pl). Signal peptide and transmembrane motif were predicted by the SignalP and TMHMM programs (http://www.cbs.dtu.dk/services).

Generation of specific anti-AtPAP2 antiserum

A fragment of *AtPAP2* cDNA corresponding to the N-terminus of AtPAP2 (a.a. 25-144) was amplified by primers P2AbF and P2AbR. This region was selected to avoid cross-reactivity toward other AtPAPs. The PCR product was digested with *SacI* and *KpnI* and fused with the N-terminal His-tag of vector pRSET-A (Invitrogen, H.K.) and the resulting plasmids were transformed into *Escherichia coli* strain BL21 (DE3). The overexpressed fusion protein in the inclusion bodies was solubilized in extraction buffer A (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 8 M urea) and purified by a HisTrap FF column (GE Healthcare) using buffer A containing 100 mM imidazole. The eluted protein was dialyzed in 20 mM PBS, pH 7.2 to remove urea. The antigen was further gel-purified before being used for rabbit immunization. The specificity of this antiserum was confirmed by western blotting analysis on the proteins isolated from wild type (WT) and the T-DNA insertion line (Fig. S3d).

Expression of AtPAP2 under P-starvation and sugar treatments

Arabidopsis plants were grown in phosphate (Pi) starvation MS medium (0 mM Pi) for 5 days and then transferred to complete MS medium for 0, 2, 4, 6 or 8 days. Alternatively, plants were grown on MS agar for 7 days before transfer to Pi starvation medium for 0, 2, 4, 7 or 10 days (del Pozo et al., 1999). For sugar treatments, 5-day-old seedlings on MS agar plates were transferred to solid MS medium containing a sucrose,
glucose, sorbitol or mannitol gradient (% w/v) for 24 hr. Total protein isolated from seedlings was analyzed by western blot analysis using AtPAP2 antiserum.

Generation of AtPAP2 overexpression lines in Arabidopsis

The full-length coding region of the *AtPAP2* cDNA (AT1G13900) was amplified by Pfx DNA polymerase (Invitrogen) using primers P2YF and P2NR. The resulting product (1,971 bp) was initially cloned into vector pGEM-T Easy (Promega, H. K.) and then subcloned into the binary vector pBA002 downstream of the cauliflower mosaic virus (CaMV) 35S promoter (pBA002-CaMV35: *AtPAP2*). The vector was then introduced into *Agrobacterium tumefaciens* strain GV3101 and then transformed by the floral dip method (Clough & Bent, 1998) into WT plants. Homozygous CaMV35S: *AtPAP2* OE lines were selected on MS plates containing 5 mg/L glufosinate ammonium (Riedel-deHaen, Germany). Resistant lines were transferred to soil to grow to maturity, and their transgenetic status was confirmed by genomic PCR and western blot analyses. Homozygous T3 seeds of the OE plants were used for further analysis. A construct of a truncated *AtPAP2* fragment without its C-terminus (P2NC) was generated by PCR using primers P2YF and P2NCR and cloned in the pBA002 vector. *AtPAP2* with a deleted signal peptide (P2NS) was amplified with primers P2NSF and P2NR and cloned into the pCXSN vector (Chen *et al.*, 2009). All primers used are listed in Table S3.

Characterization of homozygous AtPAP2 T-DNA line

A T-DNA insertion line of *AtPAP2* gene (Salk_013567) in the Col ecotype was obtained from TAIR. The homozygous T-DNA line was verified by genomic PCR screening using the T-DNA left border primer LBa1 and *AtPAP2* gene-specific primer.
LP2 and reverse primer RP2 (SIGnAL database). The T-DNA insertion site was confirmed by DNA sequencing of the PCR product.

Subcellular localization of AtPAP2

The putative signal peptide (SP) and the transmembrane domain and cytoplasmic tail (TMD/CT) were PCR-amplified from full-length AtPAP2 cDNA and cloned into a GFP vector in pBI221 to generate SP-GFP, GFP-TMD/CT and SP-GFP-TMD/CT using full-length AtPAP2 cDNA as template and the primer sets listed in Table S3. Organelle markers were used to elucidate the subcellular localization of AtPAP2-GFP constructs.

Transient expression in WT Arabidopsis PSB-D cell culture was performed as described previously (Miao & Jiang, 2007). Generally, PSB-D cells at day 5 after subculture were collected for protoplast preparation. Plasmids were electroporated into protoplasts and incubated at 26°C for 13 hr before being observed under a confocal laser scanning microscope. Protoplasts were incubated for 5 min at room temperature with MitoTracker Orange (CMTMRos, Molecular Probes) before analysis. Confocal images were collected with a Fluoview FV1000 microscope (Olympus, Japan) with a 60x objective water lens. The settings for collecting confocal images within the linear range were described previously (Jiang & Rogers, 1998; Tse et al., 2004). Images were processed using Adobe Photoshop software (Jiang & Rogers, 1998).

Organelle isolation and western blot analysis

Mitochondria and chloroplasts were isolated from 10-day-old Arabidopsis seedlings as previously described (Lister et al., 2007; Kubis et al., 2008). Proteins (25 µg) were resolved by SDS-PAGE and transferred to Hybond-C nitrocellulose membranes, then immunodetected as previously described (Carrie et al., 2008). The antibodies used,
Tom40 (Carrie et al., 2009a; Carrie et al., 2009b) and Risp (Carrie et al., 2010), have been described previously. Antibodies to CoxII, Rubisco large subunit and Rubisco small subunit (SSU) were purchased from Agrisera (Sweden).

**Measurement of soluble sugars and starch in Arabidopsis**

Soluble sugars were extracted from Arabidopsis using the chloroform/methanol method (Antonio et al., 2007). Liquid chromatography-MS/MS spectrophotometry was employed for sugar measurement. Filtered samples (10 µl) were injected into an HP 1100 HPLC system (Agilent Technologies, Palo Alto, CA, USA) connected to a zwitterionic ZIC-HILIC column (3.5 µm, 150 mm x 2.1 mm i.d.; SeQuant, Umea, Sweden) (Antonio et al., 2008). Separation was performed using a solvent system of 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in methanol (B) with a linear gradient of 20–60% B over 25 minutes. Flow rate was maintained at 0.2 ml min⁻¹ and the elution was monitored by a diode-array detector (200–600 nm). MRM experiments were conducted at the electrospray ionization interface of an API 2000 QTRAP system (Applied Biosystems, H.K.) operating in negative mode. Starch measurement was carried out according to the method of Focks and Benning (1998) using a starch assay kit (Sigma).

**Enzyme assays**

Sucrose synthase, soluble invertase and insoluble cell wall invertase activities were measured as described (Xu, et al., 1996; Dejardin et al., 1997). SPS activity was assayed by the anthrone test (Lunn & Furbank, 1997; Baxter et al., 2003). Acid phosphatase activity was assayed by a colorimetric method (Lung et al., 2008) and the total NR activity assay was carried out as described (Yu et al., 1998).
Pull-down assays using recombinant AtSnRK1.1, AtSnRK1.2, AtGRIK1, AtGRIK2 and 14-3-3 protein

Vectors for expressing GST-GRIK1, GST-GRIK2, His-AtSnRK1.1 (a.a. 1–341) and His-AtSnRK1.2 (a.a. 1–342) were gifts from Dr. Shen Wei (Shen et al., 2009). The expression vector GST-BMH1 (containing a yeast 14-3-3 homolog) was a gift from Prof. Carol MacKintosh (University of Dundee, UK). The fusion protein His-AtSnRK1.1/His-AtSnRK1.2 was purified by Ni-NTA beads and the fusion protein GST-GRIK1/GST-GRIK2 by a GST-Trap column (Amersham Biosciences) (Moorhead et al., 1999). For His-AtSnRK1.1/His-AtSnRK1.2 pull-down assays, 20 µg of protein in 200 µl His binding buffer was coupled to 200 µl of Ni-NTA beads and incubated with 1 mg total soluble Arabidopsis protein for 2 hr at 4°C. The bound proteins were washed and eluted in binding buffer containing 100 mM imidazole. For the GST-GRIK1/GST-GRIK2 and GST-14-3-3 pull-down assays, 1 mg total plant protein extract was mixed with 5 µg purified GST-fusion protein on 50 µl GST beads with gentle agitation for 2 hr at 4°C. The beads were washed with washing buffer (50 mM HEPES-KOH, pH 7.5, 500 mM NaCl, 1 mM dithiothreitol) five times. Bound proteins were eluted in 50 mM Tris-HCl, pH 8.0, 10 mM reduced glutathione, 500 mM NaCl and subjected to immunoblot analysis.

Analysis of proteins in AtPAP2 transgenic Arabidopsis plants

Total soluble proteins were extracted according to the trichloroacetic acid method (Zhu et al., 2006). The samples were run in the pH 4–7 range and stained with SYPRO Ruby for total protein staining and Pro-Q Diamond for phosphoprotein staining. Images were analyzed using ImageMaster 2-D Platinum software (GE Healthcare). The target spots were run on a MALDI-TOF/TOF mass spectrometer and identified by peptide
mass fingerprinting with data searches carried out using MASCOT search tools (http://www.matrixscience.com) in the NCBI non-redundant public protein database. All of the peptide masses were assumed to be monoisotopic and corresponding to [M+H+]. Carboxymethyl was considered to be a fixed modification, while Oxidation (M), Phospho (ST) and Phospho (Y) were considered variable modifications. The peptide mass tolerance was set at ± 0.2 Da and the maximum number of missed cleavages was set at one. The identified proteins had more than 5 matched peptides, and the percentage of sequence coverage was greater than 20%. All of the positive protein identification scores were significant (MASCOT score > 60, P < 0.05).

Metabolomic analysis by $^1$H-NMR and GC-MS

Shoots of 20-day-old Arabidopsis were freeze-dried immediately after harvest. Three biological replicates were used in this study. For $^1$H-NMR analysis, dried plant tissues (15 mg) were extracted at 50°C with D$_2$O:CD$_3$OD (80:20, 1 ml) containing 0.05% (w/v) d$_4$-TSP (Ward et al., 2003; Beale et al., 2010). NMR data was collected using 128 scans on a Bruker 600 MHz Avance NMR spectrometer (Bruker BioSpin, Coventry, UK) with a water suppression pulse sequence. Data was scaled to an internal standard (d$_4$-TSP) and binned to regions of equal width (0.01 ppm). For GC-MS analysis, lyophilized plant material (1.00 mg) was mixed with methoxyamine hydrochloride in anhydrous pyridine (50 µL, 16 mg mL$^{-1}$, containing 40 µg mL$^{-1}$ ribitol) and the samples were heated at 30°C for 90 min with agitation at 550 rpm. MSTFA + 1% TMCS (70 µL) was then added and the sample was heated for 30 minutes at 37°C with agitation at 550 rpm. The derivatized samples were equilibrated at room temperature for 2 hr prior to analysis. GC-MS analysis was carried out on a Pegasus III time of flight mass spectrometer (Leco Corp., USA) coupled to an Agilent 6890N gas chromatograph.
(Agilent Technologies) system fitted with a Microseal septum (Merlin Instrument Co., USA), a FocusLiner port liner (SGE Analytical Science, UK) packed with deactivated quartz wool and a DB-5ms capillary column (15 m x 0.18 mm I.D. x 0.18 µm d.f. with 5 m integrated guard column). Data were exported from ChromaTOF software as netCDF files and converted to MassLynx format using the DataBridge module of MassLynx v4.0 software (Waters, UK) for semi-quantification in the QuanLynx module of MassLynx. A unique mass was assigned to each analyte observed in the dataset and its contribution (response factor) to the total ion current of its corresponding mass spectrum determined. For each analyte, peak areas for the unique mass were determined by integration of its corresponding extracted ion chromatogram, and the extracted ion chromatogram areas and response factors were used to calculate a reconstituted TIC peak area. Principal component analysis was carried out using Simca-P v.11 software (Umetrics, Umea) with unit variance scaling. Significance was determined using ANOVA.
Results

Identification of a conserved class of PAP, with a C-terminal hydrophobic motif, in photosynthetic eukaryotes

AtPAP2 (AT1G13900) was classified as a group IIa PAP, among the 29 AtPAPs (Li et al., 2002). The open reading frame of AtPAP2 encodes a polypeptide of 656 amino acids with a predicted molecular weight of 73.72 kDa and a pI of 6.11. Database searches revealed that the protein contains a metallophosphoesterase domain within amino acids (a.a.) 256–487. In the genomic DNA, AtPAP2 contains 2 exons. The program TMHMM predicts that AtPAP2 possesses a transmembrane motif at its C-terminus, suggesting that AtPAP2 is a membrane protein.

Phylogenetic analysis showed that PAPs with this C-terminal motif could only be found in plants, not in animals (Fig. S1). There are many more PAP-like sequences in the genomes of higher plants than in the genomes of primitive photosynthetic eukaryotes. For instance, there are 29 putative PAP genes in the genome of Arabidopsis (Li et al., 2002) and 2 in Ostreococcus tauri (Fig. S1). Unlike Arabidopsis, where two PAPs contain a C-terminal hydrophobic region (AtPAP2 and AtPAP9), only a single PAP-like sequence from other plant genomes was found to carry this additional hydrophobic motif at the C-terminus (Figs. S1, S2 and Table S1). Interestingly, in the genome of the smallest free-living photosynthetic eukaryote, O. tauri (Derelle et al., 2006), only 2 PAP-like sequences could be found among its 8166 ORFs, of which one (OtPAP2) carries a C-terminal hydrophobic motif (Table S1). In addition to this hydrophobic motif, all the sequences homologous to AtPAP2 carry an acidic a.a. doublet (e.g. EE, DD, ED) in the last 2–8 residues of the polypeptides (Table S1). These findings indicate that PAPs with a C-terminal hydrophobic motif are only found in
photosynthetic eukaryotes, and have been conserved during the evolution of green plants.

Expression of AtPAP2 is not affected by phosphorus and sugar treatments

To examine the expression pattern of the AtPAP2 protein in various plant organs and at various developmental stages, a polyclonal antiserum was raised against recombinant AtPAP2 protein. A 74 kDa protein was detected in WT Arabidopsis but was absent in the T-DNA insertion line (Fig. S3). This indicates that the antiserum recognizes AtPAP2 specifically. AtPAP2 is highly expressed in siliques, stems, flowers, roots and senescing leaves, but its expression is relatively lower in leaves and mature seeds (Fig. S4). These results corroborate the results from RT-PCR experiments (Zhu et al., 2005).

Since AtPAP2 was predicted to be a constitutively expressed protein by spot history analysis (data not shown), the level of AtPAP2 protein expression under various stimuli was examined. As shown in Fig. S5, inorganic phosphate starvation, sucrose, glucose, sorbitol and mannitol treatments did not affect the level of AtPAP2, confirming that its expression is constitutive.

AtPAP2 overexpression accelerates plant growth and produces higher seed yield

A homozygous T-DNA line of AtPAP2 that carries a T-DNA insertion site in the second exon (1864 bp/1971 bp) was identified (Fig. S3a). The insertion resulted in the loss of AtPAP2 expression, as indicated by the absence of detectable transcript and protein (Fig. S3). Under normal germination conditions on MS agar, the length of hypocotyls of the AtPAP2 T-DNA insertion line was shorter (~20%–30%) than WT. To confirm whether the disruption of AtPAP2 was responsible for the phenotype, the
T-DNA insertion line was complemented by introducing the full-length coding sequence of \textit{AtPAP2} into Arabidopsis under the control of the CaMV 35S promoter. Four homozygous lines were verified by western blot analysis and their hypocotyls were similar to WT seedlings (data not shown). Therefore, \textit{AtPAP2} expression indeed rescued the short hypocotyl phenotype of the T-DNA line.

Overexpression studies were used to dissect the roles of \textit{AtPAP2} \textit{in vivo}. T3 homozygous \textit{AtPAP2} overexpression (OE) transgenic lines under the control of CaMV 35S promoter were generated. Multiple independent lines with increased levels of \textit{AtPAP2} protein, as confirmed by western blot analysis, were obtained. Two homozygous lines, OE7 and OE21, were randomly selected for subsequent studies. Compared to WT, both \textit{AtPAP2} OE lines displayed earlier bolting (Fig. 1) and produced more lateral branches and inflorescences. Typically, \textit{AtPAP2} OE plants started bolting earlier than both WT and the T-DNA line, \(~6\) days earlier under LD growth conditions and \(~14\) days earlier under SD conditions (Table 1a). At day 28 (LD) and day 43 (SD), when WT first bolted, \textit{AtPAP2} OE lines had produced significantly more cauline leaves and inflorescences than WT (Fig. S6). Conversely, the number of rosette leaves of \textit{AtPAP2} OE lines was fewer than WT (Fig. 1b, Table 1a), although the leaf area and total biomass of OE lines were similar to WT at 20-day-old, at day 40, the total leaf area of \textit{AtPAP2} OE lines was only 65\% that of WT, but the total biomass was 2-fold higher than WT (Fig. 1c). \textit{AtPAP2} OE plants not only matured earlier but also produced higher seed yields. In two independent experiments, the total seed yield of \textit{AtPAP2} OE7 and OE21 plants increased remarkably by 40–57\% over WT, which was due to higher numbers of auxiliary branches and a higher silique density (Table 1b). No significant differences in plant growth or seed yield relative to WT were observed in the \textit{AtPAP2} T-DNA line, possibly due to redundancy of gene function.
AtPAP2 is dual-targeted to chloroplasts and mitochondria

The protein sequence of AtPAP2 carries a hydrophobic C-terminal motif (TMD/CT) predicted to carry a trans-membrane motif and it has already been identified as a chloroplast protein by proteomics (Kleffmann et al., 2004). In contrast, the SignalP server predicts that AtPAP2 has a 24 a.a. SP at its N-terminus to direct the protein to the secretory pathway. To investigate the targeting of AtPAP2, 30 a.a. residues from the N-terminus and 69 a.a. residues from the C-terminus were fused to green fluorescent protein (GFP) under the control of the CaMV 35S promoter (Fig. 2a). Red fluorescence protein (RFP) was fused to either the mitochondrial presequence of the ATP synthase F$_1$-γ subunit from *Nicotiana plumbaginifolia* (F$_1$–RFP) (Duby et al., 2001) as a mitochondrial targeting marker or to the targeting sequence (a.a. 1–79) of the small subunit of tobacco Rubisco (plastid-mCherry) (Dabney-Smith et al., 1999) as a chloroplast targeting marker. Transient expression of the GFP fusion constructs was performed in WT Arabidopsis PSB-D cell culture (Fig. 2a).

Confocal microscopy analysis showed that both GFP (Fig. 2b and c, green) and GFP fused to the N-terminal SP (SP-GFP) (Fig. 2d and e, green) localized to the cytosol, and did not co-localize with the mitochondrial F$_1$-RFP (Fig. 2b and d, red) or plastid-mCherry (Fig. 2c and e, red) markers. However, GFP constructs containing the TMD/CT region of AtPAP2, including SP-GFP-TMD/CT (Fig. 2f and g) and GFP-TMD/CT (Fig. 2h and i), were targeted to both plastids and to mitochondria, which were also labeled by the MitoTracker Orange (Molecular Probes, USA) (Fig. S7). In addition, the patterns of expression of all GFP constructs were distinct from that of the Golgi marker (RFP-ManI) (Cai et al., 2011), ER marker (RFP-HDEL) (De Caroli et al., 2011) (Fig. S8) and peroxisome marker (peroxisome-mCherry) (Nelson et al., 2007) (Fig. S9). These results confirmed the importance of the TMD/CT in dual targeting of
the reporter GFP protein to mitochondria and plastids, and the TMD/CT alone is sufficient to direct the GFP protein to both organelles.

To confirm the subcellular localization of AtPAP2, protein extracts from isolated chloroplasts and mitochondria were probed with the antibody against AtPAP2. A protein band with an apparent molecular mass of 74 kD was detected in both mitochondria and chloroplasts. The purity of the mitochondria and chloroplasts was verified using antibodies raised against a variety of mitochondrial and chloroplast specific proteins (Fig. 2j).

C-terminal targeting is crucial for AtPAP2 function

To further study the function of the C-terminal motif and the putative signal peptide of AtPAP2, OE lines with AtPAP2 truncation of the C-terminal motif (a.a. 588–656) or its predicted SP (a.a. 1–24) were generated in Arabidopsis by Agrobacterium-mediated plant transformation (Fig. 3). Western blot analysis was employed to verify its overexpression in homozygous transgenic lines.

As shown in Fig. 3a, transgenic Arabidopsis lines expressing SP-deleted AtPAP2 (P2NS, a.a. 25–656) exhibited phenotypes similar to the full-length AtPAP2 OE lines, such as earlier bolting and higher biomass. In contrast, when the C-terminal motif was deleted, the OE lines (P2NC, a.a. 1–588) grew like the WT Arabidopsis (Fig. 3a). These results suggest that the C-terminal motif of AtPAP2 is critical for its function in Arabidopsis. In contrast, the putative N-terminal signal peptide of AtPAP2 does not seem to be essential.

In AtPAP2 OE lines (P2OE and P2NS), two protein bands could be seen in membrane fractions but only the low molecular weight band appeared in the soluble fraction. Furthermore, the high molecular weight protein band was not seen in the
membrane fraction from the P2NC lines (with no C-terminal motif). Therefore, it is
likely that the high molecular weight protein band is a membrane-associated full-length
AtPAP2 and the cleavage of its C-terminal motif releases it to the soluble fraction.
Correlated with the growth phenotype, the presence of the membrane-associated high
molecular weight form is essential to the robust growth phenotypes of the OE lines (Fig.
3b).

AtPAP2 overexpression changes the carbon metabolism of Arabidopsis

Sugars supplied by the shoot are the driving force for plant growth and anabolism. We analyzed the carbohydrate contents of leaves from 20-day-old plants halfway through the light period. The size and morphology of leaves from various lines were still similar at day 20 and therefore this age was chosen for further analysis. Both sucrose and hexose sugar contents were greatly elevated in the shoots of AtPAP2 OE plants (Table 1c), in which sucrose was 1.3- to 1.8-fold and hexose sugars were 1.9- to 2.6-fold that of WT ($P < 0.01$). However, the starch level did not differ significantly between the various lines (Table 1c).

The increased sucrose and hexose sugars (Table 1c) in the AtPAP2 OE plants may be due to altered activity of enzymes involved in sugar metabolism. SPS activity was significantly enhanced in the 20-day-old AtPAP2 OE lines under both optimal $V_{\text{max}}$ and limiting $V_{\text{limit}}$ conditions, and was not affected by light/dark transitions, as in wild-type Arabidopsis (Huber et al., 1989). In addition, the activities of enzymes involved in sucrose catabolism, including sucrose synthase, acid and alkaline soluble invertase, insoluble cell wall invertases, were unaffected (Table 2). The enzyme activity of acid phosphatase was elevated in both OE lines but the total NR activity was only significantly elevated in one OE line (OE21, Table 2). The higher SPS activity could
result from either a higher level of protein expression or a higher activity per enzyme molecule. Western blotting indicated that the SPS protein levels were significantly elevated in the OE lines during both day and night. The levels of the other enzymes, such as fructose-1,6-bisphosphatase (FBPase) and FBP aldolase in the sucrose synthetic pathway did not change significantly (Fig. 4).

To investigate whether enhanced SPS activity was due to a change in the relative amounts of phosphorylated SPS versus total SPS, a 14-3-3 capture experiment was carried out. The recombinant 14-3-3 captured similar amounts of phosphorylated SPS in all four lines (Fig. 4). Hence, the increase in SPS activity in the OE lines was due to an increase of the unphosphorylated active form of enzyme. Theoretically, a suppression of SnRK1 kinase activity could enhance SPS activity. Hence, the expression levels of SnRK1 protein subunits and the phosphorylation status of the Thr175 residue on its activation loop were examined by western blotting (Fig. S10); the levels of the AtSnRK1.1, AtSnRK β1 and AtSnRK β2 subunits were unaltered. Blots using antiserum against phosphorylated-AMPK (T172), which cross-reacts with phosphorylated Thr175 on AtSnRK1 in Arabidopsis (Sugden et al., 1999), indicated that the phosphorylation status of T175 was unaltered. This suggested that the higher SPS activity in the OE lines was not mediated by the SnRK1 pathway.

Overexpression of AtPAP2 results in alteration of soluble plastid proteins

Two-dimensional gel analysis of total leaf soluble proteins revealed that two proteins involved in photosynthesis exhibited changes in the OE lines; first, PSBP-1, an oxygen-evolving enhancer protein (AT1G06680), was induced, and secondly, Rubisco activase (RCA, AT2G39730) showed a mobility shift (Fig. 5). PSBP-1 is required for high levels of oxygen evolution (Mayfield et al., 1987). RCA enhances photosynthetic
rate by freeing Rubisco for the next round of carbon fixation (Portis, 2003; Portis et al., 2008). In Arabidopsis, there are 2 RCA genes (At1g73110 and At2g39730), and alternative splicing of At2g39730 produced 3 isoforms. Deletion of the short thermolabile isoform of At2g39730 (RCA1) caused decreased seed yield while transgenic Arabidopsis overexpressing thermostable RCA1 exhibited higher biomass and photosynthetic rate (Kurek et al., 2007). According to the PhosphAT database (phosphat.mpimp-golm.mpg.de), there is at least one phosphorylation site in both gene products (GLAYDpTSDDQ on At2G39730.1; AGGMEpTLGKV and VPLIVpTGNDF on AT1G73110). A change in phosphorylation status of RCA conceivably would affect pI, but whether this occurred was not determined.

Metabolomic analysis

Metabolomic analysis was carried out using a combination of NMR-MS and GC-MS. Shoot samples were collected from 20-day-old plants before bolting. Principal component analysis of 1H-NMR and GS-MS datasets (Fig. S11) was carried out to discern any global changes in the metabolomes of these lines. There was a very clear separation of the OE lines from WT and the T-DNA insertion line by each analytical technique. 1H-NMR was also able to discriminate the T-DNA line from WT samples, although this separation was much smaller. The main changes that could be discerned by 1H-NMR between the T-DNA samples and WT included a reduction in fumarate and citrate levels and several amino acids.

Combining the results of metabolomics analysis and LC-MS-MS measurements, metabolites which were increased in the OE lines included alanine, proline, several tricarboxylic acid (TCA) metabolites (citrate, fumarate, malate), sugars (glucose and sucrose) and a metabolite putatively identified as indole acetonitrile, a breakdown...
For Peer Review

product of indole glucosinolate. Reduced metabolites in the OE lines included certain amino acids (asparagine, aspartate, glutamine, lysine, methionine, phenylalanine, threonine, tyrosine), GABA, ethanolamine, 4-hydroxybenzoic acid, ribose, phytol, linoleic acid, and linolenic acid (Table S2). In addition, OE lines contained significantly higher levels of putrescine and significantly lower levels of spermidine. The metabolites detected in the glycolysis, TCA and amino acid synthesis pathways are presented in Fig. 6.

Discussion

The genomes of plant species generally contain multiple PAP-like sequences, among which, except Arabidopsis, usually only a single PAP gene carries a C-terminal hydrophobic motif. Amino acid sequence alignment of the 29 PAP-like sequences from the Arabidopsis genome showed that only two AtPAPs (AtPAP2, AT1G13900 and AtPAP9, AT2G03450) carry an additional hydrophobic motif (amino acids 614 to 636 of AtPAP2 and amino acids 604 to 626 of AtPAP9) at their C-termini (Table S1), which was predicted as a trans-membrane motif by the TMHMM program. Although AtPAP9 shares up to 72% sequence identity with AtPAP2 at protein level, AtPAP9 did not exhibit the biological function of AtPAP2. A homozygous T-DNA line of AtPAP9 (Salk_129905) and several homozygous AtPAP9 OE lines were identified by genomic PCR and RT-PCR analysis. All of them exhibited normal growth phenotypes, which were very different from that of AtPAP2 OE lines (data not shown). A double T-DNA insertion line of AtPAP2 and AtPAP9 was also generated through crossing the pollens of AtPAP2 insertion line and the pistil of AtPAP9 insertion line and the interruption of both genes were confirmed by genomic PCR analysis (Fig.S3). The double insertion line exhibited similar phenotypes with that of the AtPAP2 T-DNA line (data not shown). In addition, two PAP-like sequences could be identified in the genomes of the small
free-living photosynthetic eukaryotes *O. tauri* (Derelle et al., 2006) and *Micromonas pusilla* (Worden et al., 2009), of which only one carries a hydrophobic C-terminal motif. *O. tauri*, belonging to the Prasinophyceae, is the smallest (size < 1 µm) free-living photosynthetic eukaryote known. This unicellular organism lacks a cell wall and each cell carries one chloroplast, one mitochondrion, one Golgi body and a large nucleus. The Prasinophyceae are a primitive group of marine green algae, which suggests an ancient ancestral function of AtPAP2-like proteins in the regulation of carbon partitioning in photosynthetic organisms.

In our study, AtPAP2 was shown to be dual-targeted to both chloroplasts and mitochondria, key organelles for two important energy processes in plant cells: photosynthesis and respiration (Nunes-Nesi et al., 2007; Nunes-Nesi et al., 2011). An increasing number of proteins in mitochondria and plastids have been found to be encoded by single nuclear genes, and are referred to as dual-targeted proteins (Peeters & Small, 2001). To date, as many as 47 different proteins from seven plant species have been reported to be dual-targeted (Carrie et al., 2009a). In Arabidopsis, 11 PKs and 10 PPs have been shown to be targeted to the chloroplast and all are nuclear encoded (Schliebner et al., 2008; Shapiguzov et al., 2010). In mitochondria, 10 PKs and one PP were identified in Arabidopsis (Heazlewood et al., 2004). All of these PKs and PPs are targeted by N-terminal targeting signals. The identification of a C-terminal dual-targeting signal in AtPAP2 is novel and AtPAP2 is the only PP that is dual-targeted to both organelles.

Overexpression of AtPAP2 in Arabidopsis leads to a faster growth rate. The higher biomass but lower total leaf area of 40-day-old transgenic plants provided indirect evidence of a higher photoassimilation rate per leaf area in the OE lines. The changes in RCA and PSBP-1 in the OE lines might also affect the photosynthetic activity of the OE.
lines. Whether the observed modification of RCA affects the activity of Rubisco and thus results in the enhanced growth rate of AtPAP2 OE lines is an interesting topic for further study.

Sucrose is the fuel and building block for sink tissues. AtPAP2 overexpression drastically enhanced the sucrose and hexose levels in rosette leaves while the starch level did not change significantly. SPS activity shows a high correlation with sucrose production in leaves (Baxter et al., 2003; Haigler et al., 2007) and overexpression of SPS promotes earlier flowering, fruit development, and increased biomass in tomato (Micallef et al., 1995) and tobacco (Park et al., 2008). In the two AtPAP2 OE lines, SPS was upregulated at the protein level, which resulted in a significant enhancement of its activity. SnRK1 can phosphorylate and inhibit SPS activity (Halford et al., 2003) under high sucrose levels. Since overexpression of AtSnRK1.1 resulted in phenotypes opposite to that of the AtPAP2 OE plants, such as delayed flowering and retarded growth (Baena-Gonzalez et al., 2007), it may be worth asking whether AtPAP2 is an antagonistic phosphatase of SnRK’s kinase activity. Under energy deficiency, activated SnRK1 phosphorylates SPS hence suppresses activity posttranslationally (Toroiser et al., 1998; Moorhead et al., 1999). In contrast, overexpression of AtPAP2 resulted in higher SPS activity. Theoretically, a phosphatase (Sugden et al., 1999) that specifically dephosphorylates SnRK1 or its upstream kinases (e.g. GRIK) in vivo might suppress their biological activity and thus boost the activity of SPS. However, it is unlikely that AtPAP2 mediated its actions through GRIK or SnRK1 because the phosphorylation status of the T175 residue was unaltered in the OE lines; also, pull-down experiments only found an interaction between GRIK and SnRK (data not shown), but not between AtPAP2 and GRIK or SnRK1 (Fig. S10). Unlike SnRK1, which inactivates SPS by posttranslational modification, AtPAP2 overexpression enhances sucrose synthesis by
increasing the expression of SPS.

The mechanism of how AtPAP2 stimulates plant growth remains unclear. As the acid phosphatase activity was significantly increased in the AtPAP2 OE lines (Table 2), AtPAP2 may act as a PP or as a phosphatase for inorganic P compounds. In the former scenario, AtPAP2 might exhibit its effects as a PP like PPH1, a chloroplast phosphatase that dephosphorylates light-harvesting complex II during state transition (Finazzi et al., 2002; Shapiguzov et al., 2010). In the latter scenario, AtPAP2 might increase the acid phosphatase activity in the chloroplasts and mitochondria, thus releasing more free inorganic phosphate (Pi). In chloroplasts, as Pi is an activator of Rubisco activity, production of triose phosphates might be enhanced by a higher Pi supply (Giersch & Robinson, 1987), and in turn provide more triose phosphate for sucrose synthesis. Similarly, the possibly higher Pi supply in the mitochondrial matrix could also modulate TCA flux and activate oxidative phosphorylation and ATP synthesis (Wu et al., 2007). The higher concentrations of fumarate and malate in OE lines (Fig. 6) might reflect a higher flux from succinyl-CoA to succinate (Johnson et al., 1998), another step in the TCA cycle that produces ATP (Meyer et al., 2010). ATP generated from the TCA cycle could be transported out of the mitochondria to sustain the high rate of sucrose synthesis (Kromer et al., 1993). The higher levels of organic acids of the TCA cycle in the OE lines did not result in a higher amino acid level of the aspartate family, indicating that the regulation and the flux of the TCA cycle is a complicated process. Alteration of key enzymes and their activities in the TCA cycle affect photosynthesis: downregulation of aconitase, the iron-sulfur subunit of succinate dehydrogenase and the mitochondrial malate dehydrogenase enhance the rate of photosynthesis (Carrari et al., 2003; Nunes-Nesi et al., 2005; Araujo et al., 2011), whereas inhibition of citrate synthase, succinyl CoA ligase, or isocitrate dehydrogenase show little effect on the rate
All these lines displayed changes in expression levels of organic acids and amino acids. Thus, the mechanism by which AtPAP2 influences both photosynthesis and the TCA cycle is likely to be complex and may be involved in a number of pathways or factors.

In summary, we have shown that ectopic overexpression of AtPAP2 in Arabidopsis causes dramatic effects on plant growth and carbon metabolism. Its biological activity is dependent on its dual targeting to both plastids and mitochondria. Whether overexpression of AtPAP can be used as a tool to boost the yield of important crop plants will be an interesting subject of future studies.

**Supporting Information**

**Fig. S1** Phylogenetic analysis of full length AtPAP2 protein sequence with related proteins in the other plant species.

**Fig. S2** Amino acid sequence alignment of AtPAP2 with its homologous sequences from other plants.

**Fig. S3** Verification of AtPAP2 T-DNA mutant line.

**Fig. S4** Western blot analysis of AtPAP2 in Arabidopsis tissues.

**Fig. S5** Expression of AtPAP2 under various treatments.

**Fig. S6** Growth phenotypes of genetically modified Arabidopsis.

**Fig. S7** The TMD/CT region of AtPAP2 contains targeting signal to mitochondria.

**Fig. S8.** AtPAP2-GFP constructs did not co-localize with ER or Golgi markers.

**Fig. S9.** AtPAP2-GFP constructs did not co-localize with peroxisome marker.

**Fig. S10.** Analysis of SnRK1 proteins.

**Fig. S11.** Principal Component Analysis of $^1$H NMR data and GC-MS data.
Table S1. Signature motifs of AtPAP2-like proteins.

Table S2. Relative metabolite content of shoots of 20-day-old AtPAP2 OE lines.

Table S3. Primers used in this study.

Acknowledgements

This project was supported by the General Research Fund (HKU772710M) and the Innovation and Technology Fund (ITS158/09) of the HKSAR, China. We thank Dr. Yip’s lab at HKU for kindly providing plant vectors and technical support for plant cultures. We thank Aimee Llewellyn and Delia Corol at Rothamsted Research for preparation of the analytical extracts and for $^1$H NMR data collection, Dr. David Secco at the University of Western Australia for his comments on this manuscript.

References


Carrie C, Giraud E, Duncan O, Xu L, Wang Y, Huang SB, Clifton R, Murcha M,


Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-


For Peer Review

and function specificities; physical interaction with nitrate reductase specifically occurs via AKINbeta1-subunit. Plant Physiology 148(3):
1570-1582.


1597-1604.

675-709.


America 107(10): 4782-4787.


274-279.


Sugden C, Crawford RM, Halford NG, Hardie DG. 1999. Regulation of spinach SNF1-related (SnRK1) kinases by protein kinases and phosphatases is associated with phosphorylation of the T loop and is regulated by 5'-AMP. *Plant Journal* 19(4): 433-439.


**Figure Legends**

Fig. 1 Growth phenotypes of AtPAP2 overexpression (OE) and T-DNA inactivation Arabidopsis lines. (a) 4-week-old plants grown under LD conditions. (b) Leaves at day 27 under LD conditions. (c) Dry weight and total leaf area of 40-day-old plants. AtPAP2 OE lines exhibited higher biomass but much lower leaf area at 40 days of age. Leaves were cut (including leaves on inflorescences, if any) and scanned. The leaf area was calculated using Metamorph (www.image1.com). The dry weights of individual plant samples (including leaves) were measured after water evaporation by heat. The experiments were repeated at least 3 times.
Fig. 2 AtPAP2 protein is targeted to both mitochondria and chloroplasts. (a) Targeting of GFP fusion proteins in Arabidopsis PSB-D protoplasts. Transient expression in Arabidopsis PSB-D protoplasts showed that GFP (GFP only) (b, c) and GFP fused with the putative SP region of AtPAP2 (SP-GFP) (d, e) were directed to the cytosol, whereas GFP constructs fused with the TMD/CT region of AtPAP2, including SP-GFP-TMD/CT (f, g) and GFP-TMD/CT (h, i), co-localized with mitochondrial (F1-RFP) and chloroplastic (plastid-mCherry) markers. Bar = 50 µm. Western blot of AtPAP2 protein in mitochondria and chloroplasts. (j) Isolated mitochondria and chloroplasts from 10-day-old seedlings were analyzed by western blotting for the presence of AtPAP2. Purity of the isolated organelles was determined by blotting with antibodies specific to each organelle. Mitochondria: Cytochrome oxidase subunit II (CoxII), the Rieske iron sulfur cluster binding protein (Risp) and translocase of the outer membrane protein of 40 kDa (Tom40). Chloroplasts: Rubisco large subunit and Rubisco small subunit (SSU).

Fig. 3 AtPAP2 sublocalization in transgenic plants. (a) Phenotype of 40-day-old plants (n = 12) of WT, P2NS (a.a. 25–656), OE7 (a.a. 1–656), P2NC (a.a. 1–588) and AtPAP2 T-DNA plants. (b) AtPAP2 protein band identification in WT, P2NS, OE7, P2NC and AtPAP2 T-DNA plants by western blot analysis.

Fig. 4 Western blotting of enzymes involved in sucrose biosynthesis. Total soluble protein (30 µg) extracted from the shoots of 20-day-old Arabidopsis was loaded to each lane and was analyzed by western blotting using anti-SPS, anti-FBP aldolase and anti-cFBPase antibodies (Agrisera, Sweden). The pulled-down eluates of GST-14-3-3 were diluted 1:1 with 40 µl SDS sample buffer and half was loaded into each lane.
Fig. 5 Analysis of protein from AtPAP2 OE lines. An equal volume of total soluble protein (300 µg) extracted from 20-day-old Arabidopsis was analyzed by 2-D gel electrophoresis. Gels were stained with Pro-Q Diamond and SYPRO Ruby. The pI range of the gel was 4-7. Targeted spots are identified by the arrows. Mobility shifts of Rubisco activase (RCA) and oxygen-evolving enhancer protein (PSBP-1) are indicated by the arrows. Three biological replicates were used in this study.

Fig. 6 Altered metabolites in 20-day-old AtPAP2 OE lines. Metabolites in red and blue indicate increased or decreased levels, respectively, in AtPAP2 OE plants. Metabolites in black were not detected and those in green were unaltered. The dashed blue circles represent the Asp-family and aromatic amino acid network. DHAP, dihydroxyacetone phosphate; GABA, γ-aminobutyric acid; SA, salicylic acid; DC-SAM, decarboxylated S-adenosylmethionine.

Fig. S1 Phylogenetic analysis of the full-length AtPAP2 protein sequence with related proteins in other plant species.

The phylogenetic tree was constructed by MEGA 4.1 software based on the neighbor-joining method (Kumar et al., 2004). Scale bar represents 0.2 substitutions per site. The first two letters of each protein represent the abbreviated species name from which they are derived: At: Arabidopsis thaliana (AtPAP9, AAM15910; 72%); Br: Brassica rapa subsp. Pekinensis (AC176823; 84%); Hv: Hordeum vulgare (PUT-161a-Hordeum_vulgare-1024109806, 56%); Mt: Medicago truncatula (AC202582; 60%); Os: Oriza sativa (NM_001065273; 57%); Pp: Physcomitrella patens subsp. patens (XM_001768668.1, 46%); Pt: Poplar trichocarpa (NC_008476.1;
57%); St: *Solanum tuberosum* (PUT-157a-Solanum_tuberosum-29226, 60%); Zm: *Zea mays* (ACG47621; 58%); MpPAP1: *Micromonas pusilla* (XP_003057348); MpPAP2: *M. pusilla* (EEH53803, 25%); OtPAP1: *Ostreococcus tauri* (XM_003079445); OtPAP2: *O.s tauri* (XP_003075106; 29%).

**Fig. S2** Amino acid sequence alignment of AtPAP2 with homologous sequences from other plants. The full-length amino acid sequence of AtPAP2 was aligned to homologous sequences from *Arabidopsis thaliana*, *Brassica rapa*, *Ostreococcus tauri* and AtPAP15, which does not carry the C-terminal hydrophobic motif (Kuang *et al.*, 2009), with CLC Sequence Viewer 6 software. All five of the conserved domains (in blue) and the seven invariable residues (indicated by an asterisk) of the PAP metal-ligating residues were conserved among PAPs. The pink box indicates the predicted signal peptide of AtPAP2. Hydrophobic motifs at the C-termini of these polypeptides are underlined in red.

**Fig. S3** Verification of AtPAP2 T-DNA mutant line.
(a) Schematic map showing the location of the T-DNA insertion in AtPAP2 (second exon, Col ecotype, Salk_013567). (b) Genomic PCR screening for homozygous AtPAP2 T-DNA line. Primers (LBa and RP2) in the T-DNA and its flanking sequence are indicated. (c) RT-PCR of the AtPAP2 T-DNA line using full-length AtPAP2 primers P2YF and P2NR. Elongation factor (*EF*) was employed as a control. (d) Western blot analysis of AtPAP2 T-DNA line using AtPAP2-specific antibody. Coomassie Blue protein staining was used as a loading control. (e) Genomic PCR screening for AtPAP2 and AtPAP9 double mutant. Primers employed for amplifying wild type AtPAP2 and its T-DNA insertion were LP2 + RP2 (A) and LBa + RP2 (B), respectively. Primers
employed for amplifying wild type *AtPAP9* and its T-DNA insertion were LP9 + RP9 (C) and LBa + RP9 (D), respectively.

**Fig. S4** Western blot analysis of AtPAP2 in Arabidopsis tissues. Ten-day-old seedlings were germinated and grown on MS agar plates. Mature tissues and senescent leaves were collected from 5- to 6-week-old plants. Equal amounts of total protein (30 µg per lane) were loaded into each lane. Se, dry seeds; SL, senescent leaves; L mature leaves; St, stems; Si, siliques; Fl, flowers; R, roots; YS: young seedlings. Two independent experiments generated the same results.

**Fig. S5** Expression of AtPAP2 following various treatments. Five-day-old seedlings of WT Arabidopsis were transferred to Pi starvation conditions and MS media containing different concentrations of sugars. Total soluble protein (30 µg) was extracted and loaded for western blot analysis using anti-AtPAP2 antibody. (a) Seedlings germinated in Pi starvation medium were transferred to Pi-sufficient MS medium for 2, 4, 7, or 10 days. (b) Seedlings germinated under Pi-sufficient MS medium were transferred to Pi starvation medium for 0, 2, 4, 6, or 8 days. Cytosolic marker protein FBPase was selected as the internal control. Seedlings were treated with sucrose (c), glucose (e), sorbitol (d) and mannitol (f) gradients (% w/v) for 24 hr. At least 3 independent experiments were performed.

**Fig. S6** Growth phenotypes of genetically modified Arabidopsis. Bolting statistics of 4 lines in LD and SD growth conditions. Number of rosette leaves (a, d), inflorescences (b, e) and cauline leaves (c, f). (n = 10–15).
**Fig. S7** The TMD/CT region of AtPAP2 contains a targeting signal to mitochondria. Transient expression in Arabidopsis PSB-D protoplasts showed that GFP without any targeting sequence (GFP only; a) and GFP with the putative SP region of AtPAP2 (SP-GFP; b) were localized in the cytosol and were distinct from mitochondria. GFP constructs containing the TMD/CT region of AtPAP2, including SP-GFP-TMD/CT (c) and GFP-TMD/CT (d), were both targeted to mitochondria and co-localized with mitochondrial marker MitoTracker Orange (red). Bar = 50 µm.

**Fig. S8** AtPAP2-GFP constructs did not co-localize with ER or Golgi markers. Transient expression in Arabidopsis PSB-D protoplasts showed that SP-GFP (a, d), SP-GFP-TMD/CT (b, e) and GFP-TMD/CT (c, f) did not co-localize with an ER marker (RFP-HDEL; a–c) or Golgi marker (RFP-ManI; d–f). Bar = 50 µm.

**Fig. S9** AtPAP2-GFP constructs did not co-localize with a peroxisome marker. Transient expression in Arabidopsis PSB-D protoplasts showed that GFP only (a), SP-GFP (b), SP-GFP-TMD/CT (c) and GFP-TMD/CT (d) did not co-localize with a peroxisome marker (peroxisome-mCherry, red). Bar = 50 µm.

**Fig. S10** Analysis of SnRK1 proteins. (a) Western blotting analysis. Samples were collected from 20-day-old Arabidopsis in the middle of the day and the middle of the night. The total soluble protein was separated by SDS-PAGE and immunodetected with anti-SnRK1.1, anti-AMPKT\textsubscript{172} (Cell Signaling Technology, USA), anti-SnRK\textbeta{}1 (anti-AKIN\textbeta{}1, Agrisera, Sweden) and anti-SnRK\textbeta{}2 (anti-AKIN\textbeta{}2, Agrisera, Sweden) antibodies. Anti-cFBPase was used as a marker for loading controls. (b) *In vitro* pull-down assays. AtSnRK1/AtSnRK1.2 (SnRK1.1/SnRK1.2) and its upstream kinase,
GST-AtGRIK1/GST-AtGRIK2 (GRIK1/ GRIK2), did not interact with AtPAP2.

**Fig. S11** Principal component analysis of $^1\text{H}$ NMR data (a) and GC-MS data (b). The shoots of 20-day-old Arabidopsis were used for metabolomic analysis. Three biological replicates per line were used in this study.
Table 1. Growth phenotypes of transgenic Arabidopsis. Seeds of WT, T-DNA line and two AtPAP2 OE lines (OE7 and OE21) were germinated in the MS agar for 10 days, seedlings with the same size were transferred to a growth chamber. Statistically differences ($p < 0.05$) in the same row for each line were based on one-way analysis of variance (ANOVA) analysis followed by Tukey's Honestly Significant Differences (HSD) test using statistical program SPSS 11.5. There are 8-15 replicates for each line. The data were reproducible in at least 2 independent experiments.

(a) Bolting time

<table>
<thead>
<tr>
<th>Lines</th>
<th>AEI 16h/8h (days)</th>
<th>NRL 16h/8h</th>
<th>AEI 8h/16h</th>
<th>NRL 8h/16h</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>26.9 ± 1.2a</td>
<td>13.0 ± 0.8a</td>
<td>41.0 ± 4.7a</td>
<td>18.0 ± 3.0a</td>
</tr>
<tr>
<td>T-DNA</td>
<td>25.7 ± 0.7a</td>
<td>11.6 ± 1.1a</td>
<td>40.7 ± 4.9a</td>
<td>15.0 ± 3.0a</td>
</tr>
<tr>
<td>OE7</td>
<td>20.0 ± 1.1b</td>
<td>6.4 ± 0.5b</td>
<td>25.6 ± 1.3b</td>
<td>5.3 ± 0.5b</td>
</tr>
<tr>
<td>OE21</td>
<td>20.8 ± 0.6b</td>
<td>6.5 ± 0.7b</td>
<td>26.0 ± 1.1b</td>
<td>5.4 ± 0.5b</td>
</tr>
</tbody>
</table>

AEI: Average date of emergence of inflorescence
NRL: No. of rosette leaves at the first appearance of inflorescence

(b) Seed yield at maturity

<table>
<thead>
<tr>
<th>Lines</th>
<th>No. of Siliques/plant</th>
<th>Seed Yield (g/plant)</th>
<th>Seed weight (mg/100 seeds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>396 ± 89a</td>
<td>0.225 ± 0.058a</td>
<td>1.96 ± 0.14a</td>
</tr>
<tr>
<td>T-DNA</td>
<td>386 ± 70a</td>
<td>0.240 ± 0.049a</td>
<td>2.02 ± 0.16a</td>
</tr>
<tr>
<td>OE7</td>
<td>621 ± 76b</td>
<td>0.351 ± 0.050b</td>
<td>1.91 ± 0.10a</td>
</tr>
<tr>
<td>OE21</td>
<td>625 ± 94b</td>
<td>0.355 ± 0.066b</td>
<td>1.98 ± 0.03b</td>
</tr>
</tbody>
</table>

All seeds and siliques from a single plant were harvested after the plant was completely dried. The plants were grown under LD regime.

(c) Carbohydrate contents in 20-day-old Arabidopsis shoots.

<table>
<thead>
<tr>
<th>Lines</th>
<th>Sucrose (µg/g FW)</th>
<th>Hexose (µg/g FW)</th>
<th>Starch (mg/g FW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>110.1 ± 33.3a</td>
<td>16.7 ± 7.1a</td>
<td>3.3 ± 0.4a</td>
</tr>
<tr>
<td>T-DNA</td>
<td>116.1 ± 15.6a</td>
<td>24.9 ± 7.3a</td>
<td>4.4 ± 1.1a</td>
</tr>
<tr>
<td>OE7</td>
<td>207.1 ± 22.5b</td>
<td>43.9 ± 9.2b</td>
<td>3.8 ± 1.0a</td>
</tr>
<tr>
<td>OE21</td>
<td>147.7 ± 17.3c</td>
<td>31.7 ± 4.8ac</td>
<td>3.5 ± 0.8a</td>
</tr>
</tbody>
</table>

Sucruses were extracted 8 hours after illumination during the LD regime from 20-day-old Arabidopsis.
Table 2. Enzyme assays. Samples were collected 8 hours after the light period (Day) and 4 hours after the dark period (Night). Values set in bold type marked with a different letter were determined by the *t* test to be significantly from WT and were determined by the ANOVA-Turkey test to be significantly different (*p* < 0.05) from each other. There are 3 biological replicates in this study.

<table>
<thead>
<tr>
<th>Plant line</th>
<th>WT</th>
<th>T-DNA</th>
<th>OE7</th>
<th>OE21</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sucrose phosphate synthase (µmole sucrose/mg total soluble protein /hour)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vmax (Day)</td>
<td>5.44 ± 0.90(^a)</td>
<td>5.80 ± 0.59(^a)</td>
<td>7.89 ± 1.08(^b)</td>
<td>7.97 ± 0.95(^b)</td>
</tr>
<tr>
<td>Vlimit (Day)</td>
<td>3.18 ± 0.32(^a)</td>
<td>3.47 ± 0.12(^a)</td>
<td>4.54 ± 0.92(^b)</td>
<td>3.99 ± 0.68(^a)</td>
</tr>
<tr>
<td>Vmax (Night)</td>
<td>5.91 ± 0.57(^a)</td>
<td>5.24 ± 0.72(^a)</td>
<td>7.54 ± 0.97(^b)</td>
<td>6.82 ± 0.75(^a)</td>
</tr>
<tr>
<td>Vlimit (Night)</td>
<td>3.74 ± 0.31(^a)</td>
<td>2.83 ± 0.16(^b)</td>
<td>4.67 ± 0.18(^c)</td>
<td>4.86 ± 0.54(^c)</td>
</tr>
<tr>
<td><strong>Sucrose synthase (µmole glucose /mg total soluble protein /hour)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day</td>
<td>12.46 ± 0.23(^a)</td>
<td>12.35 ± 1.20(^a)</td>
<td>12.43 ± 0.22(^a)</td>
<td>12.76 ± 0.27(^a)</td>
</tr>
<tr>
<td>Night</td>
<td>12.47 ± 0.36(^a)</td>
<td>12.63 ± 0.44(^a)</td>
<td>12.54 ± 0.43(^a)</td>
<td>12.94 ± 0.27(^a)</td>
</tr>
<tr>
<td><strong>Soluble invertase (µmole glucose /mg total soluble protein /hour)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day (Acid)</td>
<td>0.71 ± 0.11(^a)</td>
<td>0.60 ± 0.27(^a)</td>
<td>0.91 ± 0.45(^a)</td>
<td>0.90 ± 0.04(^a)</td>
</tr>
<tr>
<td>Night (Acid)</td>
<td>0.70 ± 0.34(^a)</td>
<td>1.30 ± 0.5(^a)</td>
<td>0.88 ± 0.21(^a)</td>
<td>0.69 ± 0.04(^a)</td>
</tr>
<tr>
<td>Day (Alkaline)</td>
<td>8.48 ± 0.49(^a)</td>
<td>8.06 ± 1.61(^a)</td>
<td>8.04 ± 1.39(^a)</td>
<td>8.94 ± 0.84(^a)</td>
</tr>
<tr>
<td>Night (Alkaline)</td>
<td>6.50 ± 0.41(^a)</td>
<td>5.25 ± 0.63(^a)</td>
<td>6.80 ± 0.68(^a)</td>
<td>6.83 ± 0.07(^a)</td>
</tr>
<tr>
<td><strong>Insoluble cell wall invertase (µmole sucrose/mg total soluble protein /hour)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day (Acid)</td>
<td>1.11 ± 0.19(^a)</td>
<td>0.90 ± 0.24(^a)</td>
<td>1.17 ± 0.31(^a)</td>
<td>1.20 ± 0.33(^a)</td>
</tr>
<tr>
<td>Night (Acid)</td>
<td>1.08 ± 0.50(^a)</td>
<td>0.85 ± 0.21(^a)</td>
<td>1.41 ± 0.21(^a)</td>
<td>1.11 ± 0.16(^a)</td>
</tr>
<tr>
<td>Day (Alkaline)</td>
<td>6.08 ± 0.14(^a)</td>
<td>6.35 ± 0.11(^a)</td>
<td>5.09 ± 0.32(^a)</td>
<td>5.27 ± 0.10(^a)</td>
</tr>
<tr>
<td>Night (Alkaline)</td>
<td>6.94 ± 0.31(^a)</td>
<td>7.57 ± 0.43(^a)</td>
<td>6.17 ± 0.28(^a)</td>
<td>6.77 ± 0.73(^a)</td>
</tr>
<tr>
<td><strong>Nitrate reductase (µmole nitrite/g fresh weight /hour)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle of the day</td>
<td>0.14 ± 0.00(^a)</td>
<td>0.15 ± 0.01(^a)</td>
<td>0.17 ± 0.00(^a)</td>
<td>0.24 ± 0.02(^b)</td>
</tr>
<tr>
<td>Middle of the night</td>
<td>0.13 ± 0.01(^a)</td>
<td>0.13 ± 0.00(^a)</td>
<td>0.15 ± 0.04(^ab)</td>
<td>0.20 ± 0.01(^c)</td>
</tr>
<tr>
<td><strong>Acid phosphatase activity (µmole Pi/mg total soluble protein /hour)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle of the day</td>
<td>8.33 ± 1.38(^a)</td>
<td>8.05 ± 1.14(^a)</td>
<td>18.46 ± 1.91(^b)</td>
<td>16.59 ± 1.04(^b)</td>
</tr>
<tr>
<td>Middle of the night</td>
<td>6.04 ± 0.11(^a)</td>
<td>6.00 ± 0.81(^a)</td>
<td>10.71 ± 0.12(^b)</td>
<td>13.71 ± 0.56(^c)</td>
</tr>
</tbody>
</table>
Fig. 1 Growth phenotypes of AtPAP2 overexpression (OE) and T-DNA inactivation Arabidopsis lines. (a) 4-week-old plants grown under LD conditions. (b) Leaves at day 27 under LD conditions. (c) Dry weight and total leaf area of 40-day-old plants. AtPAP2 OE lines exhibited higher biomass but much lower leaf area at 40 days of age. Leaves were cut (including leaves on inflorescences, if any) and scanned. The leaf area was calculated using Metamorph (www.image1.com). The dry weights of individual plant samples (including leaves) were measured after water evaporation by heat. The experiments were repeated at least 3 times.
**Fig. 2** AtPAP2 protein is targeted to both mitochondria and chloroplasts. (a) Targeting of the GFP fusion proteins in Arabidopsis PSB-D protoplasts. Transient expression in Arabidopsis PSB-D protoplasts showed that GFP (GFP only) (b, c) and GFP fused with putative SP region of AtPAP2 (SP-GFP) (d, e) were directed to the cytosol, whereas GFP constructs fused with the TMD/CT region of AtPAP2, including SP-GFP-TMD/CT (f, g) and GFP-TMD/CT (h, i), were co-localized with the mitochondrial (F1-RFP) and chloroplastic (plastid-mCherry) markers. Bar = 50 μm. (j). AtPAP2 protein located in both mitochondria and chloroplasts. Western blot of AtPAP2 protein in mitochondria and chloroplasts. (j) Isolated mitochondria and chloroplasts from 10-day-old seedlings were analyzed by western blotting for the presence of
AtPAP2. Purity of the isolated organelles was determined by blotting with antibodies specific to each organelle. Mitochondria: Cytochrome oxidase subunit II (CoxII), the Rieske iron sulfur cluster binding protein (Risp) and translocase of the outer membrane protein of 40 kDa (Tom40). Chloroplasts: Rubisco large subunit and Rubisco small subunit (SSU).

Fig. 3 AtPAP2 sublocalization in transgenic plants. (a) Phenotype of 40-day-old plants (n=12) of WT, P2NS (a.a 25-656), OE7 (a.a 1-656), P2NC (a.a 1-588) and AtPAP2 T-DNA plants. (b) AtPAP2 protein band identification in WT, P2NS, OE7, P2NC and AtPAP2 T-DNA plants by western blot analysis.
Fig. 4 Western blotting of enzymes involved in sucrose biosynthesis. Total soluble protein (30 µg) extracted from the shoots of 20-day-old Arabidopsis was loaded to each lane and was analyzed by western blot using anti-SPS, anti-FBP aldolase and anti-cFBPase antibodies (Agrisera, Sweden). The pulled-down eluates of GST-14-3-3 were diluted 1:1 with 40 µl SDS-sample buffer and half was loaded into each lane.
Fig. 5 Analysis of protein from AtPAP2 OE lines. An equal volume of total soluble proteins (300 µg) extracted from 20-day-old Arabidopsis were analyzed by 2-D gel electrophoresis. Gels were stained with Pro-Q Diamond and SYPRO Ruby. The pI range of the gel was 4-7. Targeted spots are identified by the arrows. Mobility shifts of Rubisco activase (RCA) and oxygen-evolving enhancer protein (PSBP-1) are indicated by the arrows. Three biological replicates were used in this study.
Fig. 6 Altered metabolites in the 20-day-old AtPAP2 OE lines. Metabolites in red and blue colors indicate increased or decreased levels in the AtPAP2 OE lines. Metabolites in black were not detected and those in green were unaltered. The dashed blue circles represent the Asp-family and aromatic amino acid network. DHAP, dihydroxyacetone phosphate; GABA, γ -Aminobutyric acid; SA, salicylic acid; DC-SAM, decarboxylated S-adenosyl-methionine.