Subcellular localization and developmental regulation of cytosolic, soluble pyrophosphatase isoforms in *Arabidopsis thaliana*

Zahide Neslihan ÖZTÜRK[1,2,*], **Steffen GREINER**[1], **Thomas RAUSCH**[1]

[1]Centre for Organismal Studies, Plant Molecular Physiology Group, Ruprecht Karl University of Heidelberg, Heidelberg, Germany
[2]Department of Agricultural Genetic Engineering, Faculty of Agricultural Science and Technologies, Niğde University, Central Campus, Niğde, Turkey

Abstract: Pyrophosphate (PP_i) is the by-product of several reversible key reactions of primary metabolism. Thus, generated PP_i should be removed by hydrolysis via pyrophosphatases to prevent accumulation and to drive anabolism. In plastids, a plastidic, soluble pyrophosphatase splits PP_i released by ADPG pyrophosphorylase. The cytosolic PP_i pool is believed to be hydrolyzed by tonoplast- and/or Golgi-integral H^+-translocating pyrophosphatases. The *Arabidopsis thaliana* (L.) Heynh. genome encodes 6 soluble pyrophosphatase isoforms (PPas), 1 of which was shown to be localized in plastids. The remaining 5 of those PPas (PPas 1–5) are more similar to each other with highly conserved protein sequences. They all lack known targeting sequences and are thought to reside in the cytosol. To address their role and redundancy in plants, we performed (i) subcellular targeting of C-terminal PPa:EGFP fusions, (ii) analysis of transgenic promoter β-glucuronidase (GUS) lines to depict differential expression during plant development, and (iii) transcript quantification via real-time PCR to corroborate promoter-GUS data. The results revealed pronounced tissue specificity and developmental regulation for each *PPa* isoform, but also partial redundancy with coexpression of several *PPa* isoforms in all plant tissues.

Key words: Brassicaceae, Cruciferae, soluble pyrophosphatase family, in silico subcellular localization, EGFP chimeras, histochemical GUS staining, real-time PCR

1. Introduction

Polymer-forming anabolic reactions for DNA, RNA, proteins, and polysaccharides generate inorganic pyrophosphate (PP_i) as a by-product (Stitt, 1998). The removal of excess PP_i is performed by ubiquitous inorganic pyrophosphatases (EC 3.6.1.1). Since PP_i coupled reactions are reversible in nature, the hydrolysis of PP_i is essential to provide a pull for biosynthetic reactions by making them thermodynamically irreversible (Sivula et al., 1999). Soluble pyrophosphatases (sPPases) and membrane-bound proton-pumping inorganic pyrophosphatases are 2 structurally and functionally different types of pyrophosphatases (du Jardin et al., 1995; Perez-Castineria et al., 2001). Among the membrane-bound proton-pumping inorganic pyrophosphatases, vacuolar pyrophosphatase (vPPase) is ubiquitous in the plant kingdom (Maeshima et al., 1994; Perez-Castineria et al., 2001).

Plant vPPases are relatively well studied, both structurally and functionally (Maeshima, 2000; Ferjani et al., 2011). However, there are almost no data available in the literature on plant sPPases. Previous studies showed that plant cytosol contains a high concentration of PP_i and low sPPase activity, whereas the plastids contain low PP_i concentrations and high pyrophosphatase activity (Weiner et al., 1987). The removal of the cytoplasmic PP_i pool with overexpression of *Escherichia coli* sPPase impaired plant growth and development, indicating the importance of large cytoplasmic PP_i concentrations (Jelitto et al., 1992; Sonnewald, 1992). These observations and the presence of a PP_i-degrading, proton-pumping, vPPase-coupling PP_i hydrolysis with proton motive force generation led to a generally well-accepted hypothesis that plant cytosol lacks sPPase activity and that vPPase is the sole enzyme responsible for the removal of excess cytoplasmic PP_i (Stitt, 1998).

There are 6 sPPase isoforms in the *Arabidopsis* genome, only 1 of which was shown to be localized in the plastids (Schulze et al., 2004). Recently, the chloroplast-localized isoform was proven to be essential for plastidial metabolism through virus-induced gene silencing (George et al., 2010). Ubiquitous expression of 2 *Arabidopsis thaliana* (L.) Heynh. sPPase isoforms (At1g01050 and At3g53629)

* Correspondence: zahideneslihan_ozturk@nigde.edu.tr

were reported based on the GENEVESTIGATOR database (Navarro-De la Sancha et al., 2007), and the importance of these 2 isoforms in the mobilization of sucrose during seed-filling has recently been shown (Meyer et al., 2012). Current data suggest that cytosolic PP_i concentration regulated by differential expression of sPPase isoforms might function as a regulator of plant metabolism (Navarro-De la Sancha et al., 2007; Meyer et al., 2012).

In this study, we aimed to understand the subcellular localization and tissue- and developmental-stage dependent expression of the *Arabidopsis thaliana* (L.) Heynh. soluble pyrophosphatase (PPa) family to get insight into their importance in plant metabolism.

2. Materials and methods

2.1. Plant growth
Arabidopsis thaliana (L.) Heynh. ecotype Columbia was used for all experiments. The growth of *A. thaliana* in hydroponic culture (25 μM H_3BO_3, 0.06 μM $CuSO_4$, 0.14 μM $MnSO_4$, 0.03 μM Na_2MoO_4, 0.001 μM $CoCl_2$, 0.1 μM $ZnSO_4$, 20 μM Fe-EDTA, 2 mM KNO_3, 1 mM $Ca(NO_3)_2$, 1 mM KH_2PO_4, and 1 mM $MgSO_4$) was performed according to Noren et al. (2004). *A. thaliana* seeds were surface-sterilized and inoculated onto pipette tips filled with hydroponic culture solidified with 1% plant agar. The pipette boxes were sealed and transferred to climate chambers [14 h light (24 °C), 8 h dark (18 °C), light intensity 125 μmol $m^{-2} s^{-1}$, 50% relative humidity] after 2 days of incubation at 4 °C (Gharari et al., 2014). The seals were opened 10 days after transfer and seedlings were grown until they reached the fourth leaf growth stage. The plants were grown for up to 4 weeks with weekly renewal of the nutrient solution (6 L) and no aeration. Whole leaves and roots were then collected, directly frozen in liquid nitrogen, and stored at –80 °C.

2.2. Real-time PCR
Total RNA was isolated using the RNeasy Plant Kit (QIAGEN, USA) and reverse transcription of 2 μg of total RNA was performed according to the manufacturer's instructions (Omniscript Reverse Transcriptase, QIAGEN). In order to prevent degradation of RNA, RNaseOUT (Invitrogen, USA) was added to the reaction mixture. The real-time PCR was prepared in a 25-μL volume and performed using the iCycler (Bio-Rad Laboratories, USA). The annealing temperature was optimized for each gene-specific primer pair (*PPa1*, *PPa2*, *PPa3*, and *PPa5* at 60 °C and *PPa4* and *PPa6* at 58 °C). Gene-specific primer pairs spanning one intron sequence were used for amplification of *Arabidopsis thaliana* (L.) Heynh. soluble pyrophosphatases (Table 1). The data were normalized according to Muller et al. (2002) using actin expression as the reference (Aydın et al., 2014).

Table 1. Primer sets used for PCR.

Primer name (accession no.)	Sequence (5´ to 3´)
For real-time PCR	
PPa1 (At1g01050)	ACA ATC GGC TGT TTC GTT TC TTC CTT TAG TGA TCT CAA CAA CCA C
PPa2 (At2g18230)	GAT TCT CTG CTT CGG TTT CG CAG TAG GAG CTT CTG GAC CAA TC
PPa3 (At2g46860)	CAA ATG CTC TGT TTT CTT CTG C CCT TTG TGA TCT CAA CCA CCA C
PPa4 (At3g53620)	TGA GAT CTG TGC TTG CGT TT TGG GGC TTC AGG TCC TAT C
PPa5 (At4g01480)	CTC CAC ACT TTC CGC AAG AT ACT GGA GCT CCA GGT CCG
PPa6 (At5g09650)	GAG ACA AAC CAG CAA ACA AAG AC AAA CAA AAT CCA AAT CCC AAT G
Actin (At3g18780)	GGT AAC ATT GTG CTC AGT GGT GG CTC GGC CTT GGA GAT CCA CAT C

Table 1. (Continued).

For Gateway cloning procedure, first amplification step

PPa1	ACA ATC GGC TGT TTC GTT TC
(At1g01050)	CAG AAG GAC CAA ATG ATA AGG A
PPa2	GAT TCT CTG CTT CGG TTT CG
(At2g18230)	TAG AAG GCG GCT TCA GTG AT
PPa3	TCT TTC CAA ATG CTC TGT TTT C
(At2g46860)	GGA TGC TGC GAA GAG AAT CA
PPa4	TGA GAT CTG TGC TTG CGT TT
(At3g53620)	TCA TTC CAT TGC TTG AGG TG
PPa5	CTC CAC ACT TTC CGC AAG AT
(At4g01480)	ACC CAA AAA TTG GCA GGA AT
PPa6	CCA GTC GAG GGT AAA ACG AA
(At5g09650)	CTC ATA AGT CGA CAT TGC AAA A

For Gateway cloning procedure, second amplification step (containing att recombination sequences)

PPa1	AAA AAA GCA GGC TTC ATG AGT GAA GAA ACT AAA GAT AAC C
(At1g01050)	AGA AAG CTG GGT CAC GCC TCA GGG TGT GGA G
PPa2	AAA AAA GCA GGC TTC ATG GCT GAA ATC AAG GAT GAA G
(At2g18230)	AGA AAG CTG GGT CGC GTT GCA GGC CAG CTT TG
PPa3	AAA AAA GCA GGC TTC ATG AGT GAA GAA GCA TAT GAA G
(At2g46860)	AGA AAG CTG GGT CTC TCC TCA ACG TGT GGA GAA TAT AC
PPa4	AAA AAA GCA GGC TTC ATG GCG CCA CCG ATT GAG
(At3g53620)	AGA AAG CTG GGT CAC GTC TTA GGT TCT CCA CGA CG
PPa5	AAA AAA GCA GGC TTC ATG AAT GGA GAA GAA GTG AAA AC
(At4g01480)	AGA AAG CTG GGT CTC TCC TCA GGG TGT GAA GAA TG
PPa6	AAA AAA GCA GGC TTC ATG GCG GCT ACT AGA GTG
(At5g09650)	AGA AAG CTG GGT CGT AAA GTG AAA GGT CTC CAG

For generation of Promoter:GUS constructs (containing Gateway-compatible ends)

PPa1	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TGT GGA TAT TGA TTG ATG AGA
(At1g01050)	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TCA TCT GTC AAA ATC AGA ACG
PPa2	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTT TTG GTA CAT CAC TGT CGT
(At2g18230)	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TCA TAT CGG ATG TTC ACT ATA A
PPa3	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TCA CCA ACA TTT CTT CTG GTA
(At2g46860)	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TCA TCT ATT GTC ACA TCA ATC G
PPa4	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TCC GTC TAT GTC TCT TTT TCC
(At3g53620)	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TCA TTG CTG CAA AAC AAT GC
PPa5	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TCA CTC CAT ATG ATA GCG AAA
(At4g01480)	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TCA TCT AAT AAA ACA GAG CAA AAC
PPa6	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TAG CTT CAC ACC TCA AGA AAT
(At5g09650)	GGGG ACC ACT TTG TAC AAG AAA AGC TGG GTC ATT ATT GTG GAA AGA TTG A

2.3. Generation of subcellular localization constructs and transient expression in tobacco

The full-length coding sequences of *Arabidopsis thaliana* (L.) Heynh. soluble pyrophosphatases were amplified from leaf (*PPa1* and *PPa2*), etiolated seedling (*PPa3*), flower (*PPa4*), root (*PPa5*), or seedling (*PPa6*) cDNAs using gene-specific primers designed for 2-step PCR for the generation of Gateway-compatible overhangs (Table 1). PCR products were cloned into the entry vector pDONR201 (Invitrogen) via the BP reaction (Invitrogen). The reactions were transformed into *Escherichia* XL1-Blue strain by electroporation and inserts were verified by sequencing. The LR reaction (Invitrogen) was used to transfer inserts from entry clone pDONR201 into destination vector pK7WGF2 (Kamiri et al., 2002) according to the manufacturer's instructions. The products of LR reactions were mobilized in the *E. coli* DH5α strain and inserts were confirmed by restriction digestion. The constructs were then mobilized in *Agrobacterium tumefaciens* strain C58C1 containing Ti plasmid pGV2260 by electroporation.

The *Agrobacterium tumefaciens* cells containing the vector were grown in YEB medium supplemented with 100 μg/mL rifampicin, 50 μg/mL carbenicillin, and 100 μg/mL spectinomycin at 28 °C for 24 h. Cells were collected by centrifugation at $4000 \times g$ for 10 min and resuspended in deionized water to achieve $OD_{600} = 1$. The *A. tumefaciens* suspension was infiltrated into the abaxial side of *Nicotiana tabacum* 'Petite Havanna SNN' leaves with a needleless 10-mL syringe (Wroblewski et al., 2005). The fluorescence was detected 48 h after infiltration using Zeiss META LSM 510 confocal laser scanning microscopy (Carl Zeiss AG, Germany).

2.4. Generation of promoter: β-glucuronidase constructs and stable transformation of *Arabidopsis thaliana* (L.) Heynh.

Leaves (100 mg) were ground, suspended in 500 μL of extraction buffer (200 mM Tris-HCl, pH 9, 400 mM LiCl, 25 mM EDTA, 1% SDS), and extracted twice with phenol :chloroform:isoamylalcohol (25:24:1) by centrifugation at $15,000 \times g$ for 5 min at 4 °C for genomic DNA isolation. The supernatant was precipitated with isopropanol and genomic DNA was collected by centrifugation at 15,000 $\times g$ for 10 min at 4 °C. After drying, the pellet was resuspended in 500 μL of TNE (10 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, pH 8) and total RNA was removed by RNaseA digestion. Genomic DNA was then precipitated and dissolved in TE buffer.

Promoter regions of *PPa* isoforms were amplified using primer pairs containing Gateway-compatible ends (Table 1). PCR products were cloned to the entry vector pDONR201 (Invitrogen) via the BP reaction (Invitrogen) according to manufacturer's instructions. The reactions

were mobilized in the *Escherichia coli* XL1-Blue strain by electroporation and inserts were verified by DNA sequencing. The LR reaction (Invitrogen) was used to transfer inserts from entry clone pDONR201 into destination vector pBGWFS7 (Kamiri et al., 2002). The products of LR reactions were mobilized in *E. coli* strain DH5α and inserts were confirmed by restriction digestion. The constructs were then mobilized in *Agrobacterium tumefaciens* strain C58C1 containing Ti plasmid pGV2260 by electroporation.

The *Agrobacterium tumefaciens* cells were grown in YEB medium supplemented with 100 μg/mL rifampicin, 50 μg/mL carbenicillin, and 100 μg/mL spectinomycin at 28 °C. The cells were collected by centrifugation at $4000 \times g$ for 10 min and suspended in DIP medium [½ MS, 5% sucrose, pH 5.8, 10 μL/L BAP, 0.05% Vac-In Stuff (Silwet L-77, Lehle Seeds)] to achieve $OD_{600} = 0.9$. The *Arabidopsis thaliana* (L.) Heynh. plants were transformed by floral dip according to Clough and Bent (1998). Positive plants from the T_1 and T_2 generations were selected by BASTA screening. Only T_2 generation plants were used in histochemical β-glucuronidase (GUS) staining experiments

2.5. Histochemical GUS staining

The tissues were incubated in GUS staining buffer (100 mM sodium phosphate buffer, pH 7, 10 mM $Na_2(EDTA)$, 0.5 mM $K_3[Fe(CN)_6]$, 0.5 mM $K_4[Fe(CN)_6]$, 0.08% X-GlcA) overnight at 37 °C with slight shaking. The visualization was performed after decolorizing the tissues in 70% ethanol.

3. Results

3.1. Alignment of *Arabidopsis thaliana* (L.) Heynh. PPa family

The *Arabidopsis thaliana* (L.) Heynh. genome encodes 6 sPPase isoforms (Table 2). The lengths of the amino acid chains of 5 PPa isoforms (PPas 1–5) are very close to each other (approximately 215 amino acids), whereas the plastidial isoform PPa6 is slightly larger (244 amino acids without transit peptide) (Table 2). The percentage of homology is high within the 5 paralogous PPa isoforms (75% to 91%), whereas the plastidial isoform (PPa6) shares much less homology with the others (about 35%) (Figure 1).

3.2. In silico and in vivo subcellular localization of PPa isoforms

The in silico predictions for the subcellular localization of PPa isoforms (Table 3) indicated that PPa1 and PPa6 could localize in mitochondria and chloroplasts, respectively, whereas all other isoforms may localize to the cytoplasm. Note that none of the isoforms contain known nuclear localization signals (Table 3).

Table 2. Properties of the *Arabidopsis thaliana* (L.) Heynh. soluble pyrophosphatase protein family.

Abbreviation	Accession number	Amino acids	Predicted MW (kDa)	Predicted pI
PPa1	At1g01050	212	24.5	5.97
PPa2	At2g18230	218	24.6	5.97
PPa3	At2g46860	216	24.9	5.63
PPa4	At3g53620	216	24.5	5.15
PPa5	At4g01480	216	24.7	6.09
PPa6	At5g09650	300 [1] 244 [2]	26.4 [2]	4.75 [2]

Prediction of molecular weight and isoelectric point was performed using the ExPASy Compute pI/MW tool (ExPASy:SIB Bioinformatics Resource Portal). For the plastidic PPa6 isoform: [1]with transit peptide; [2]without transit peptide.

```
PPa1        ----MSE--ETKDNQRLQRPAPRLNERILSSLSRRSVAAHPWHDLEIGPGAPQIFNVVVE       54
PPa2        MAEIKDEGSAKGYAFPLRNPNVTLNERNFAAFTHRSAAAHPWHDLEIGPEAPTVFNCVVE       60
PPa3        MSEEAYE--ETQESSQSPRPVPKLNERILSTLSRRSVAAHPWHDLEIGPEAPLVFNVVVE       58
PPa4        MAPPIEV--STKSYVEKHVSLPTLNERILSSMSHRSVAAHPWHDLEIGPEAPIIFNCVVE       58
PPa5        MNGEEVK--TSQPQKKLQNPTPRLNERILSSLSKRSVAAHPWHDLEIGPGAPVIFNVVIE       58
PPa6        --------------VQEEGPAESLDYRVF-FLDGSGKKVSPWHDIPLTLG-DGVFNFIVE       44
St_sPPase   -----MS--NENDDLSPQRRAPRLNERILSSISRRSVAAHPWHDLEIGPEAPSVFNVVIE       53
                          *:  *  :   :    .  . ****:  :      :** ::*

PPa1        ITKGSKVKYELDKKTGLIK------VDRILYSSVVYPHNYGFVPRTL-------------       95
PPa2        ISKGGKVKYELDKNSGLIK------VDRVLYSSIVYPHNYGFIPRTI-------------      101
PPa3        ITKGSKVKYELDKKTGLIK------VDRILYSSVVYPHNYGFIPRTL-------------       99
PPa4        IGKGSKVKYELDKTTGLIK------VDRILYSSVVYPHNYGFIPRTL-------------       99
PPa5        ISKGSKVKYELDKKTGLIK------VDRILYSSVVYPHNYGFVPRTL-------------       99
PPa6        IPKESKAKMEVATDEDFTPIKQDTKKGKLRYYPYNINWNYGLLPQTWEDPSHANSEVEGC      104
St_sPPase   ISKGSKVKYELDKKTGLIK------VDRILYSSVVYPQNYGFIPRTL-------------       94
            *  *  .*.*  *:  .    :             ::  *       ***::*:*

PPa1        CEDNDPIDVLVIMQEPVLPGCFLRARAIGLMPMIDQGEKDDKIIAVCVDDPEYKHYTDIK      155
PPa2        CEDSDPMDVLVLMQEPVLTGSFLRARAIGLMPMIDQGEKDDKIIAVCADDPEFRHYRDIK      161
PPa3        CEDNDPLDVLVLMQEPVLPGCFLRARAIGLMPMIDQGEKDDKIIAVCADDPEYKHFTDIK      159
PPa4        CEDSDPIDVLVIMQEPVIPGCFLRAKAIGLMPMIDQGEKDDKIIAVCADDPEYRHYNDIS      159
PPa5        CEDNDPIDVLVIMQEPVLPGCFLRARAIGLMPMIDQGEKDDKIIAVCVDDPEYKHITNIN      159
PPa6        FGDNDPVDVVEIGETQRKIGDILKIKPLAALAMIDEGELDWKIVAISLDDPKAHLVNDVE      164
St_sPPase   CEDNDPMDVLVLMQEPVLPGCFLRARAIGLMPMIDQGEKDDKIIAVCADDPEYRHYTDIK      154
             *.**:**: : :       * :*: : :. : ***:** * **:*:. ***: :   ::.

PPa1        EL---PPHRLSEIRRFFEDYKKNENKEVAVNDFLPS----ESAVEAIQYSMDLYAEYILH      208
PPa2        EL---PPHRLAEIRRFFEDYKKNENKKVDVEAFLPA----QAAIDAIKDSMDLYAAYIKA      214
PPa3        QL---APHRLQEIRRFFEDYKKNKVNKVAVNDFLPS----ESAHEAIQYSMDLYAEYILH      212
PPa4        EL---PPHRMAEIRRFFEDYKKNENKEVAVNDFLPA----TAAYDAVQHSMDLYADYVVE      212
PPa5        EL---PPHRLSEIRRFFEDYKKNENKEVAVNDFLQP----GPAIEAIQYSMDLYAEYILH      212
PPa6        DVEKHFPGTLTAIRDWFRDYKIPDGKPANRFGLGDKPANKDYALKIIQETNESWAKLVKR      224
St_sPPase   QL---PPHRLAEIRRFFEDYKKNENKDVAVDDFLPP----NSAVNAIQYSMDLYAEYILH      207
            ::      *   :  **  :*.***  :  *   .    :        *  .:: : : :*  :

PPa1        TLRR------    212
PPa2        GLQR------    218
PPa3        TLRR------    216
PPa4        NLRR------    216
PPa5        TLRR------    216
PPa6        SVDAGDLSLY    234
St_sPPase   SLRK------    211
                :
```

Figure 1. Protein sequences within the PPa family of cytosolic soluble pyrophosphatases. Box alignment of amino acid sequences of PPa isoforms performed according to Sievers et al. (2011). The pyrophosphatase signature referring to the conserved active site motif (EX[7-8]KXE) of soluble PPases (Baltscheffsky et al., 1999) extends from positions 60 to 70. Note that the plastidial isoform (PPa6) sequence used in the alignment does not include the transit peptide. For comparison, the sequence of a soluble PPase from potato (*Solanum tuberosum* L.) is included (St_sPPase). Accession numbers are At1g01050 (PPa1), At2g18230 (PPa2), At2g46860 (PPa3), At3g53620 (PPa4), At4g01480 (PPa5), At5g09650 (PPa6, plastidic), and Z36894 (St_sPPase).

Table 3. In silico prediction of subcellular localization of PPa isoforms.

	PSORT (Nakai and Horton, 1999)	TargetP (Emanuelsson et al., 2000)	LOCTree (Nair and Rost, 2005)	PredictNLS (Cokol et al., 2000)
PPa1	Cytoplasm	Mitochondria	Cytoplasm	No NLS
PPa2	Cytoplasm	Other compartments	Cytoplasm	No NLS
PPa3	Cytoplasm	Other compartments	Cytoplasm	No NLS
PPa4	Cytoplasm	Other compartments	Cytoplasm	No NLS
PPa5	Cytoplasm	Other compartments	Cytoplasm	No NLS
PPa6	Vacuole	Chloroplast	Chloroplast	No NLS

NLS: Nuclear localization signal.

Figure 2 shows the results of in vivo subcellular localization studies of PPa-enhanced green fluorescence protein (EGFP) chimeras (C-terminal fusion constructs) after transient transformation of tobacco epidermal cells. PPa1:EGFP (Figure 2a), PPa2:EGFP (Figure 2b), PPa3:EGFP (Figure 2c), PPa4:EGFP (Figure 2d), and PPa5:EGFP (Figure 2e) were localized in the cytoplasm and nucleus, whereas overlay of EGFP fluorescence with the chlorophyll autofluorescence (data not shown) indicated exclusive localization of PPa6 to the plastids (Figure 2f).

3.3. Histochemical staining of transgenic PPa promoter-GUS lines

The *Arabidopsis thaliana* (L.) Heynh. sPPase genes are located on different chromosomes (Table 2) but are usually found in heavily gene-rich areas (The Arabidopsis Information Resource, 2013). This is especially outstanding for the *PPa3*, *PPa5*, and *PPa6* genes, whose adjacent promoter regions comprise only about 500 base pairs. The upstream promoter region of *PPa1* is approximately 1 kb long, whereas those of *PPa2* and *PPa4* are longer than 3 kb. Nevertheless, for comparative analysis of promoter activities, the 1 kb region upstream of the start codon of each *PPa* promoter was fused to GUS and was introduced into *A. thaliana* plants. The histochemical GUS analysis of the T_2 generation of transgenic plants revealed a partial redundancy of promoter activities with some isoform-specific expression patterns (Table 4; Figures 3 and 4). For example, the dot-shaped staining patterns of source leaves were observed exclusively for *PPa1* (Figure 3a). Detailed analysis revealed that GUS-derived staining was specific to the spongy mesophyll cells (data not shown). Other examples for isoform-specific expressions were observed in *PPa2* where the vascular tissues in the internode regions were stained (Figure 3b), and in *PPa5* where the transcriptional activity was solely detected in the hydathodes of sink rosette leaves (Figure 3c). The staining of funiculi was detected for *PPa1* (Figure 3d) and

PPa5 (Figure 3e), whereas trichome staining was observed in stably transformed *A. thaliana* with *PPa3* or *PPa4* promoter region-driven GUS expression (Figures 3f and 3g). The activities of all *PPa* promoters were detected in vascular bundles of source leaves (Figures 3a and 3i).

The promoter activities of all *PPa* isoforms increased during floral development (Figure 3h). Only the *PPa4* promoter drove GUS activity in very early stages of flower development (data not shown). On the contrary, the promoter activity of *PPa6* (plastidial isoform) was detected only in fully developed flowers (Figure 3h). The histochemical GUS staining of other *PPa* promoter lines indicated that the activity starts at a certain stage during flower development when flower buds are still closed but pollen formation has started (data not shown). No seed staining was detected in any of the transgenic lines.

The promoter activities of *PPa* in root tips (Figure 4) indicated the high specificity of isoforms. The *PPa1* (Figure 4a) and *PPa6* (Figure 4f) promoters were found to be active in the cell division zone, *PPa3* only in the elongating region (Figure 4c), and *PPa2* (Figure 4b), *PPa4* (Figure 4d), and *PPa5* (Figure 4e) in the overall root tip. Only the promoter activity of *PPa5* was observed in the whole root (data not shown), whereas the activities of other isoforms were restricted to defined root domains.

3.4. Real-time PCR analyses of *PPa* isoforms in different tissues

The mean normalized expression levels of each *PPa* isoform transcript are given in Figure 5. The expression of the plastidial isoform (*PPa6*) is considerably higher than those of other *PPa* isoforms. Interestingly, the *PPa6* transcript was also detected in heterotrophic tissues like roots and internodes, although at relatively low levels compared to the expression in leaves.

Other *PPa* isoforms (which were shown to localize in the cytoplasm and nucleus; Figure 2) had very low expression levels compared to the actin gene; however,
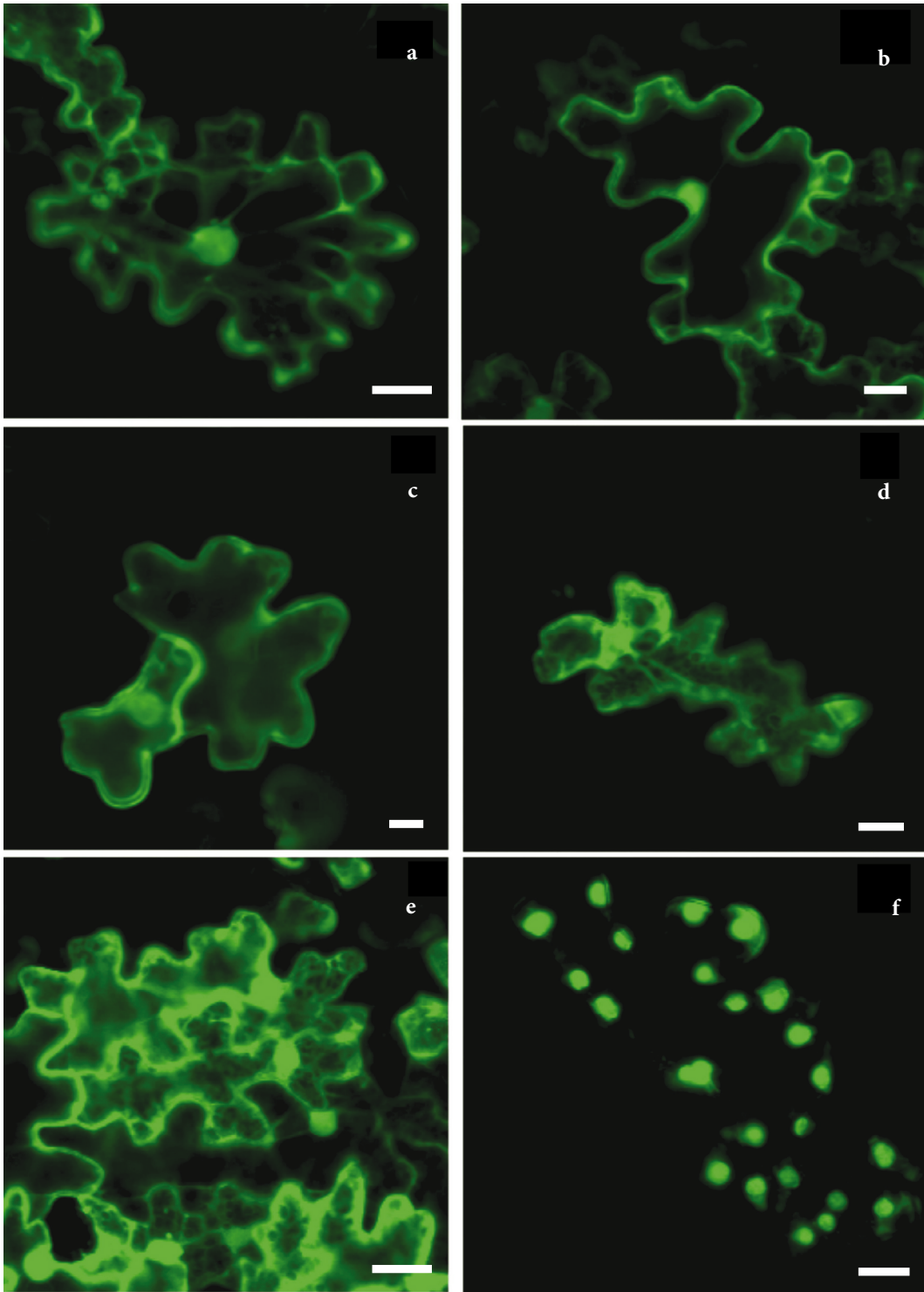
Figure 2. Subcellular localizations of PPa isoforms. Confocal laser scanning microscopy images reveal the targeting of C-terminal PPa:EGFP fusion proteins after transient expression in tobacco leaves. Tobacco leaves were analyzed 48 h after infiltration with *Agrobacterium tumefaciens* bearing the different PPa:EGFP constructs. (a) PPa1:EGFP, (b) PPa2:EGFP, (c) PPa3:EGFP, (d) PPa4:EGFP, (e) PPa5:EGFP, and (f) PPa6:EGFP. Scale bar = 10 μm.

Table 4. Distribution of GUS-derived staining of T_2 generation of $PPa_{promoter}$:GUS transgenic plants. Sink and source leaves were collected from 4-week-old plants after germination. Note that flower staining patterns are given only for the fully developed flowers.

Tissue	PPa1	PPa2	PPa3	PPa4	PPa5	PPa6
Sink leaves						
Whole blade		+		+		+
Vascular bundles		+	+	+		+
Hydathodes	+	+	+	+	+	+
Source leaves						
Whole blade			+	+	+	+
Vascular bundles	+	+	+	+		+
Hydathodes	+	+	+	+	+	+
Trichomes		+		+		
Internode			+	+		
Flowers						
Sepals	+	+		+	+	
Stigma	+	+	+	+	+	+
Stamen	+	+	+	+	+	+
Petal	+		+	+	+	
Pollen	+	+	+	+	+	+
Receptacle	+	+	+	+	+	+
Siliques						
Whole silique	+			+		
Funiculus	+				+	

their mRNAs were present in all tissues studied (Figure 5). *PPa1* and *PPa5* have higher expression rates in whole plant compared to other isoforms, whereas the transcripts of *PPa2* and *PPa3* are almost undetectable.

4. Discussion

4.1. PPa isoforms share high homology
The *Arabidopsis thaliana* (L.) Heynh. sPPase isoforms shared high homology with each other (Figure 1), except the plastidial isoform (PPa6; Schulze et al., 2004), which is slightly larger (Table 2). The alignment of known plant sPPases with animal and fungal sPPases revealed the presence of 2 stretches of amino acid sequences that are present in the plants but missing in animal and fungal sPPases. Interestingly, plant plastidial sPPases do not contain these deletions (Sivula et al., 1999) and, consequently, are found to be slightly larger. Therefore,

they are said to be more 'eukaryotic-like' compared to other isoforms, which are more 'prokaryotic-like' (Perez-Castineira et al., 2001).

4.2. PPa isoforms localize in the cytoplasm, plastids, and possibly nucleus, but not in mitochondria
The in silico analysis of the *Arabidopsis thaliana* (L.) Heynh. PPa family for subcellular localizations demonstrated that most of the enzyme isoforms are located in the cytosol, except PPa6, which is known to be plastidial (Schulze et al., 2004). On the other hand, PPa1 is known to contain a possible mitochondrial localization signal. Perez-Castineria et al. (2001) referred to PPa1 as a 'mitochondrial polypeptide precursor' in their report. It is noteworthy that this putative mitochondrial localization was detected only by TargetP (Emanuelsson et al., 2000); however, other programs like PSORT (Nakai and Horton, 1999) and LOCTree (Nair and Rost, 2005) predicted
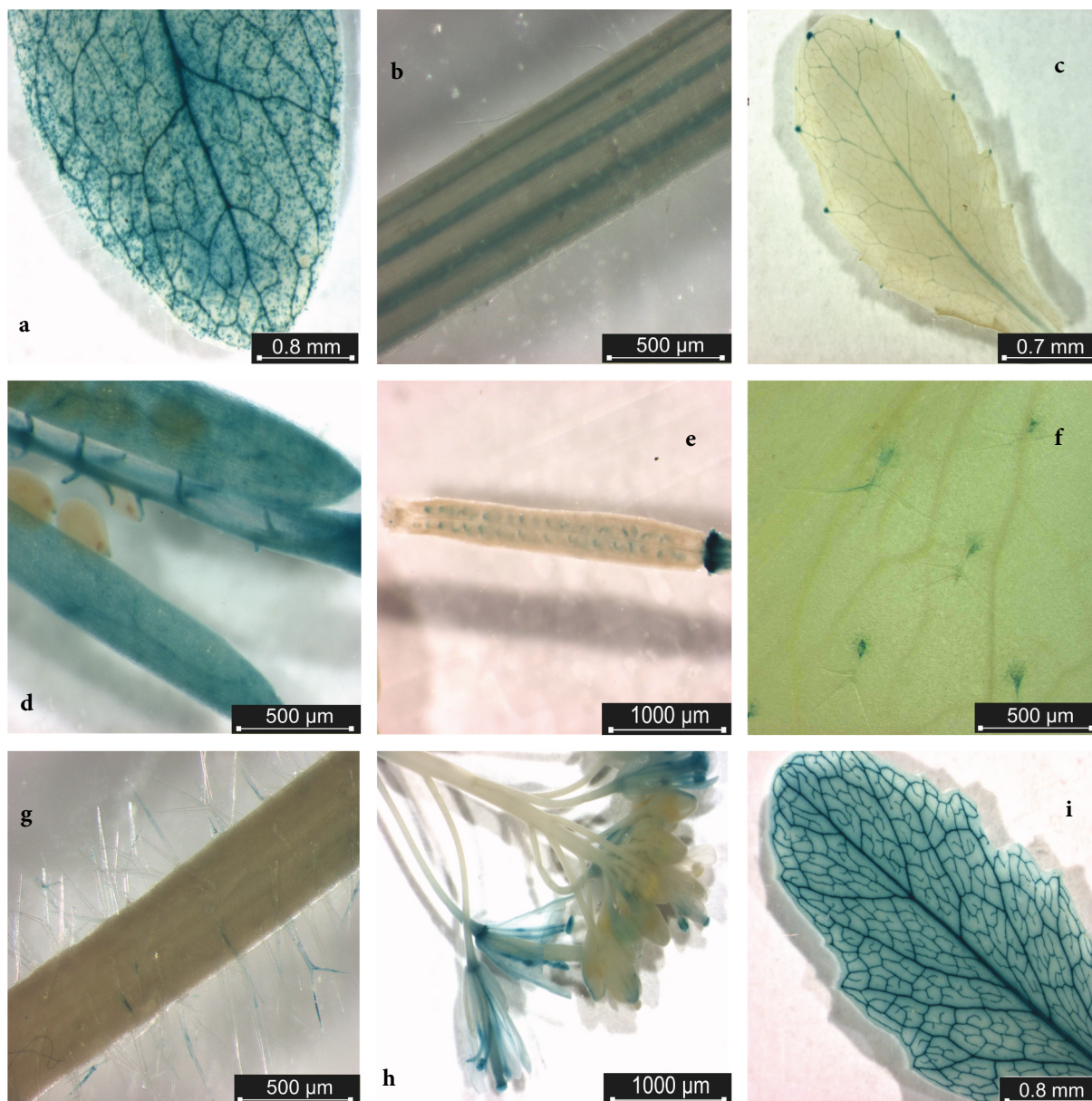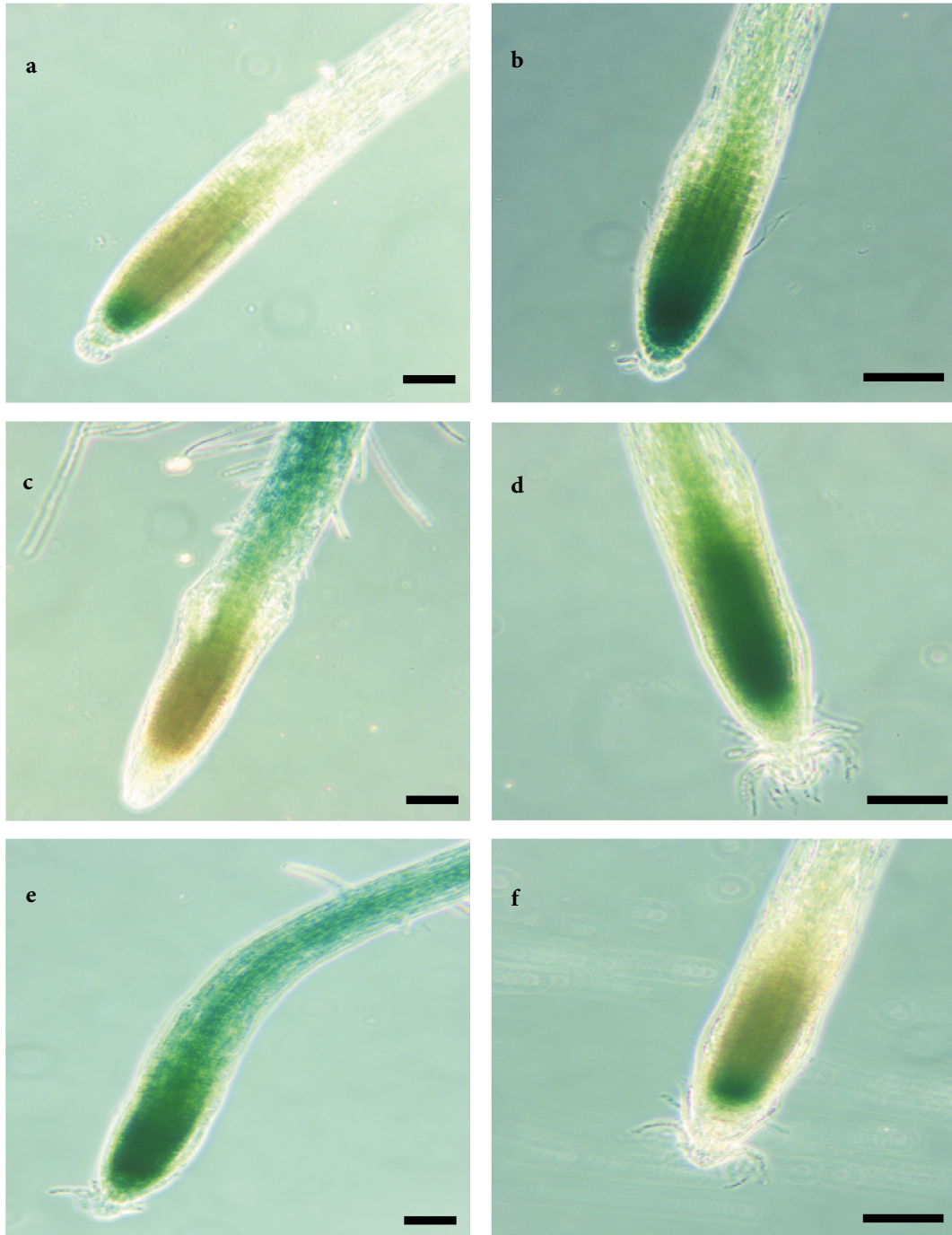
Figure 3. Histochemical staining of T$_2$ generation *PPa*$_{promoter}$:GUS transformants. Histochemical analysis of *Arabidopsis thaliana* (L.) Heynh. transformants expressing *PPa1*$_{promoter}$:GUS constructs in source leaves (a) and in siliques (d). *PPa2*$_{promoter}$:GUS staining was observed in vascular tissues of internodes (b). *PPa3*$_{promoter}$:GUS (f) and *PPa4*$_{promoter}$:GUS (g) activities were observed in trichomes and *PPa5*$_{promoter}$:GUS in hydathodes of sink leaves (c) and in funiculi of siliques (e). Histochemical analysis of *PPa6*$_{promoter}$:GUS in flower (h) and source leaves (i).

cytoplasmic localization (Table 3). A further analysis of the PPa1 amino acid sequence with MitoProt (Claros and Vincens, 1996) also failed to detect any mitochondrial localization.

The results of transient overexpression of PPa-EGFP chimeras after transformation to tobacco epidermal cells localized all PPa isoforms (except PPa6) into the cytosol and nucleus. The nuclear localization of PPa isoforms is not conclusive, as the nuclear pore complex has an exclusion limit of 60 kDa (Haasen et al., 1999; Koroleva et al., 2005) and the fusion proteins have approximate sizes of 55 kDa. Hence, nuclear localization can be the result of free diffusion between the cytoplasm and nucleus. The lack of any nuclear localization signal in PPa isoforms (Table 3) supported this hypothesis.

Among the sPPases revealed from the *Arabidopsis thaliana* (L.) Heynh. genome database (Table 2), none of them localized in the mitochondria under the studied

Figure 4. GUS expression in roots of *PPa*_{promoter}:GUS transformants. Histochemical analysis of *Arabidopsis thaliana* (L.) Heynh. transformants ectopically expressing *PPa*_{promoter}:GUS constructs: (a) *PPa1*, (b) *PPa2*, (c) *PPa3*, (d) *PPa4*, (e) *PPa5*, and (f) *PPa6*. Note that isoforms *PPa1* and *PPa6* are exclusively expressed in the root meristem, whereas expression of isoforms *PPa2* and *PPa4* extends into the elongation zone. Isoform *PPa3* expression begins only in the elongation zone, whereas isoform *PPa5* appears to be uniformly expressed along the entire root axis. Scale bar = 1 mm.

conditions (Figure 2). There are a number of reports discussing the possible presence of mitochondrial sPPases in plants, all of which were based on crude mitochondrial preparations (Zancani et al., 1995; Vianello and Macri, 1999; Casolo et al., 2002). Since there are biosynthetic pathways active in the plant mitochondria where PP$_i$ is
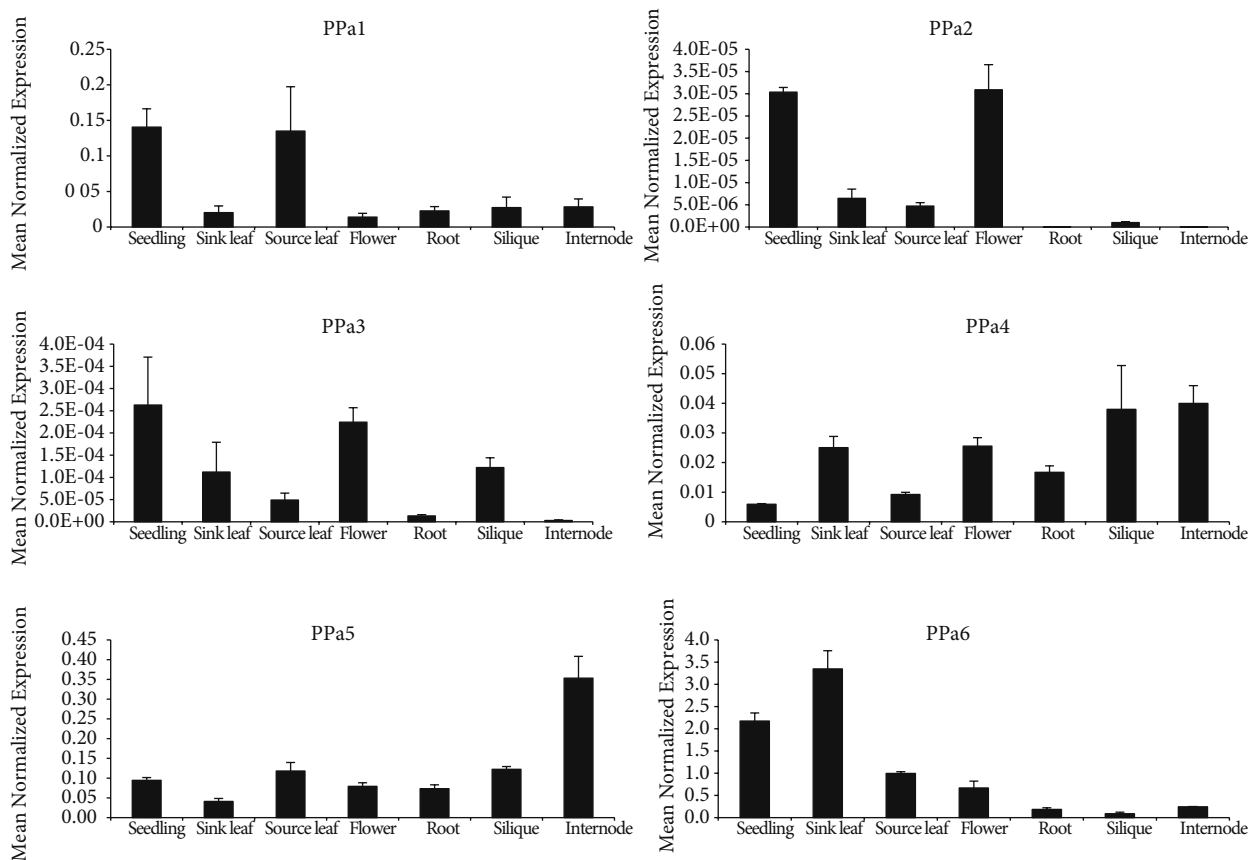
Figure 5. Isoform-specific real-time PCR analysis of *PPa* expression in several tissues. Data were normalized (Muller et al., 2002) for actin (At3g18780) as the reference gene. Presented are the means of 3 independent experiments; error bars indicate ±SD. Whole seedlings were analyzed at the 4-leaf stage. In hydroponically grown, rosette-stage nonflowering plants, leaves were divided into 'sink' (growing) and 'source' (fully expanded) leaves.

generated, either the action of sPPase or translocation of excess PP$_i$ to the cytoplasm has to be postulated for the removal of the excess PP$_i$ pool. Since none of the PPa:EGFP constructs localized to mitochondria in our experiments, the targeting of soluble pyrophosphatase to mitochondria may require specific environmental conditions, or another possibility is that the mitochondrial localization can be limited to certain tissues (heterotrophic tissues like stem cells; Zancani et al., 1995). Additionally, a mitochondrial sPPase from plants was shown to be associated with a protein complex at the inner membrane (Vianello and Macri, 1999); thus, it is also possible that the addition of EGFP may interfere with complex formation and thereby causes mistargeting of the enzyme to the cytoplasm. Although it is likely that transient overexpression itself can cause an ambiguous localization of the target protein, one cannot rule out the possibility that plants do not possess mitochondrial sPPase activity and excess PP$_i$ is exported out of this organelle to the cytoplasm by an as-of-yet unknown translocator protein.

In summary, the results of in vivo subcellular localization studies showed that *Arabidopsis thaliana* (L.) Heynh. sPPase isoforms PPa1, PPa2, PPa3, PPa4, and PPa5 are localized in the cytosol and possibly the nucleus, whereas PPa6 is localized exclusively in the plastids (Figure 2), confirming the expectations of in silico subcellular localization predictions (Table 3). To our knowledge, this is the first in vivo proof for the localization of sPPases in the cytoplasm of plant cells.

4.3. Promoter activities of *PPa* isoforms in different tissues indicate isoform specificity and developmental stage dependency

The histochemical analysis of the T$_2$ generation of *PPa$_{promoter}$*:GUS transgenic plants revealed some isoform-specific expression patterns of *Arabidopsis thaliana* (L.) Heynh. sPPase isoforms. For example, *PPa1* showed some dot-shaped staining patterns in spongy mesophyll cells of source leaves (Figure 3a). Although the reasons for specific activity of *PPa1* in some, but not all, mesophyll cells are not known, a similar expression pattern in the leaf area

was previously reported for some specialized cell types like idioblasts (Webb, 1999; Franceschi and Nakata, 2005). The funiculi staining observed for *PPa1* (Figure 3d) and *PPa5* (Figure 3e) promoter region-driven GUS expression may imply direct roles of these enzymes in seed development. Although the changes in promoter activities were not analyzed according to precise stages in flower development, histochemical GUS staining of stigma, filaments, anther, and especially pollen (Figure 3h) suggested the possible roles of sPPases during pollen maturation.

The staining of trichomes in *PPa3* and *PPa4* transgenic plants (Figures 3f and 3g) is another example of isoform-specific promoter activity. Trichomes are expanded single-celled structures, and, for that reason, they require higher amounts of new membrane lipids, cell wall components, and proteins (Smart et al., 1998). The hydrolysis of sucrose is required to produce UDP-glucose, which is the entering form of carbon for cell wall biosynthesis (Dennis and Blakeley, 2000); therefore, the promoter activities of *PPa3* and *PPa4* in trichomes may elucidate their specific role during sucrose hydrolysis.

The histochemical staining of all *PPa* promoters in the vascular bundles of source leaves corresponds to the in vivo function of sPPases. Lerchl et al. (1995) proved that cytosolic PP_i is essential for long-distance sucrose transport and that excessive removal of PP_i from the phloem by overexpression of *Escherichia coli* sPPase impairs sucrose translocation. Although the promoter activities of *E. coli* sPPase transgenes (CaMV 35S; Lerchl et al., 1995) and the endogenous sPPase genes are different and therefore not comparable, the transcriptional activity in vascular bundles may imply that sPPase activity is essential for phloem loading and/or cleavage of sucrose to achieve sufficient sink strength. A strictly controlled pyrophosphate concentration in the cytoplasm may be required for the regulation of carbohydrate metabolism and/or translocation of sucrose to phloem.

The promoter activities of *Arabidopsis thaliana* (L.) Heynh. sPPases in root tips (Figure 4) indicated a possible involvement of sPPases in root growth. The differences in the staining patterns, on the other hand, clearly showed the high specificity of sPPase isoforms in function.

The histochemical analysis of stably transformed *Arabidopsis thaliana* (L.) Heynh. plants expressing *PPa* promoter-driven GUS expression demonstrated tissue specificity and developmental stage dependency of transcription of *A. thaliana* soluble pyrophosphatases.

4.4. Real-time PCR confirms gene expression of most *Arabidopsis thaliana* (L.) Heynh. sPPases in several plant tissues

Supporting the results derived from the microarray data obtained from the GENEVESTIGATOR database (Navarro-De la Sancha et al., 2007), the mean normalized expression levels of *PPa* isoforms (Figure 5) revealed that the isoform having the highest transcript amount was the plastidial isoform *PPa6*; the rest of the isoforms had very low expression levels. For example, the expressions of *PPa2* and *PPa3* were almost at undetectable levels. Since promoter activities of *PPa2* and *PPa3* were observed by histochemical GUS staining of several plant tissues (Figures 3 and 4), the extremely low levels of their RNA may indicate the presence of a posttranscriptional regulation or a low mRNA stability of these *PPa* isoforms.

It is important to note that the cytoplasmic (and/or nuclear) soluble pyrophosphatase (*PPa1–5*) mRNAs were present not only in growing tissues like sink leaves, but also in later developmental tissues (Figure 5). Perez-Castineria et al. (2001) claimed that the plant sPPases missing a leader peptide should be expressed in nonphotosynthetic tissues, like roots, based on their similarity to previously cloned sPPase genes from potato tuber (du Jardin et al., 1995). The results obtained by both promoter-driven GUS expression and by real-time PCR analysis are clearly in opposition to this hypothesis, showing the expression of several *PPa* isoforms in both photosynthetically active and heterotrophic tissues.

In conclusion, subcellular localization studies of PPa isoforms indicated that 5 isoforms are localized in the cytoplasm and/or nucleus and the last 1 exclusively in the plastids. Since the presence of mitochondrial sPPases was shown to be essential for the function of this organelle in animal and fungal cells, the first interesting finding here was the lack of mitochondrial localization for *Arabidopsis thaliana* (L.) Heynh. sPPase:EGFP fusion proteins. This may imply that plant mitochondria lack sPPase activity and accumulated PP_i is translocated to the cytosol for removal by cytoplasmic sPPases. The promoter activity analyses and real-time PCR data indicated that *A. thaliana* sPPase isoforms are differentially expressed in several plant tissues. The in vivo functions of sPPases have not been proven yet, and therefore the possibility of another substrate cannot be ruled out; however, the isoform-dependent expression may imply the specificity of functions of plant soluble pyrophosphatases. In addition, the observation that *PPa* transcripts are present not only in young and growing but also in fully developed tissues may imply the function of plant soluble pyrophosphatases throughout development.

References

Aydın S, Büyük İ, Aras ES (2014). Expression of SOD gene and evaluating its role in stress tolerance in NaCl and PEG stressed *Lycopersicum esculentum*. Turk J Bot 38: 89–98.

Baltscheffsky M, Schultz A, Baltscheffsky H (1999). H(+)-PPases: a tightly membrane-bound family. FEBS Lett 457: 527–533.

Casolo V, Micolini S, Macri F, Vianello A (2002). Pyrophosphate import and synthesis by plant mitochondria. Physiol Plantarum 114: 516–523.

Claros MG, Vincens P (1996). Computational method to predict mitochondrially imported proteins and their target sequences. Eur J Biochem 241: 779–786.

Clough SJ, Bent AF (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J 16: 735–743.

Cokol M, Nair R, Rost B (2000). Finding nuclear localization signals. EMBO Rep 1: 411–415.

Dennis DT, Blakeley SD (2000). Carbohydrate metabolism. In: Buchanan BB, Gruissem W, Jones RL, editors. Biochemistry and Molecular Biology of Plants. Rockville, MD, USA: American Society of Plant Physiologists, pp. 630–675.

du Jardin P, Rojas-Beltran J, Gebhardt C, Brasseur R (1995). Molecular cloning and characterization of a soluble inorganic pyrophosphatase in potato. Plant Physiol 109: 853–860.

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.

ExPASy:SIB Bioinformatics Resource Portal. Website: http://web. expasy.org/compute_pi/ [accessed 01 December 2013].

Ferjani A, Segami S, Horiguchi G, Muto Y, Maeshima M, Tsukaya H (2011). Keep an eye on PPi: the vacuolar-type H+-pyrophosphatase regulates postgerminative development in *Arabidopsis*. Plant Cell 23: 2895–2908.

Franceschi VR, Nakata PA (2005). Calcium oxalate in plants: formation and function. Annu Rev Plant Biol 56: 41–71.

George GM, van der Merwe MJ, Nunes-Nesi A, Bauer R, Fernie AR, Kossmann J, Lloyd JR (2010). Virus-induced gene silencing of plastidial soluble inorganic pyrophosphatase impairs essential leaf anabolic pathways and reduced drought stress tolerance in *Nicotiana benthamiana*. Plant Physiol 154: 55–66.

Gharari Z, Khavari Nejad R, Shekaste Band R, Najafi F, Nabiuni M, Irian S (2014). The role of *Mn-SOD* and *Fe-SOD* genes in the response to low temperature in *chs* mutants of *Arabidopsis*. Turk J Bot 38: 80–88.

Haasen D, Kohler C, Neuhaus G, Merkle T (1999). Nuclear export of proteins in plants: AtXPO1 is the export receptor for leucine-rich nuclear export signals in *Arabidopsis thaliana*. Plant J 20: 695–705.

Jelitto T, Sonnewald U, Willmitzer L, Hajirezaei M, Stitt M (1992). Inorganic pyrophosphate content and metabolites in potato and tobacco plants expressing *E. coli* pyrophosphatase in their cytosol. Planta 188: 238–244.

Kamiri M, Inzé D, Depicker A (2002). GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. Trends Plant Sci 7: 193–195.

Koroleva OA, Tomlinson ML, Leader D, Shaw P, Doonan JH (2005). High-throughput protein localization in Arabidopsis using *Agrobacterium*-mediated transient expression of GFP-ORF fusions. Plant J 41: 162–174.

Lerchl J, Geigenberger P, Stitt M, Sonnewald U (1995). Impaired photoassimilate partitioning caused by phloem-specific removal of pyrophosphate can be complemented by a phloem-specific cytosolic yeast-derived invertase in transgenic plants. Plant Cell 7: 259–270.

Maeshima M (2000). Vacuolar H(+)-pyrophosphatase. Biochim Biophys Acta 1465: 37–51.

Maeshima M, Mimura T, Sato T (1994). Distribution of vacuolar H+-pyrophosphatase and a membrane integral protein in a variety of green plants. Plant Cell Physiol 35: 323–328.

Meyer K, Stecca KL, Ewell-Hicks K, Allen SM, Everard JD (2012). Oil and protein accumulation in developing seeds is influenced by the expression of a cytosolic pyrophosphatase in Arabidopsis. Plant Physiol 159: 1221–1234.

Muller PY, Janovjak H, Miserez AR, Dobbie Z (2002). Processing of gene expression data generated by quantitative real-time RT-PCR. Biotechniques 32: 1372–1379.

Nair R, Rost B (2005). Mimicking cellular sorting improves prediction of subcellular localization. J Mol Biol 348: 85–100.

Nakai K, Horton P (1999). PSORT: a program for detecting the sorting signals of proteins and predicting their subcellular localization. Trends Biochem Sci 24: 34–35.

Navarro-De la Sancha E, Coello-Coutino MP, Valencia-Turcotte LG, Hernandez-Dominguez EE, Trejo-Yepes G, Rodriguez-Sotres R (2007). Characterization of two soluble inorganic pyrophosphatases from *Arabidopsis thaliana*. Plant Sci 172: 796–807.

Noren H, Suensson P, Anderson B (2004). A convenient and versatile hydroponic cultivation system for *Arabidopsis thaliana*. Physiol Plantarum 121: 343–348.

Perez-Castineira JR, Gomez-Garcia R, Lopez-Marques RL, Losada M, Serrano A (2001). Enzymatic systems of inorganic pyrophosphate bioenergetics in photosynthetic and heterotrophic protists: remnants or metabolic cornerstones? Int Microbiol 4: 135–142.

Schulze S, Mant A, Kossmann J, Lloyd JR (2004). Identification of an Arabidopsis inorganic pyrophosphatase capable of being imported into chloroplasts. FEBS Lett 565: 101–105.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J et al (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7: 539.

Sivula T, Salminen A, Parfenyev AN, Pohjanjoki P, Goldman A, Cooperman BS, Baykov AA, Lahti R (1999). Evolutionary aspects of inorganic pyrophosphatase. FEBS Lett 454: 75–80.

Smart LB, Vojdani F, Maeshima M, Wilkins TA (1998). Genes involved in osmoregulation during turgor-driven cell expansion of developing cotton fibers are differentially regulated. Plant Physiol 116: 1539–1549.

Sonnewald U (1992). Expression of E. coli inorganic pyrophosphatase in transgenic plants alters photoassimilate partitioning. Plant J 2: 571–581.

Stitt M (1998). Pyrophosphate as an energy donor in the cytosol of plant cells: an enigmatic alternative to ATP. Bot Acta 111: 167–175.

The Arabidopsis Information Resource (TAIR). Website: http://www.arabidopsis.org [accessed 01 December 2013].

Vianello A, Macri F (1999). Proton pumping pyrophosphatase from higher plant mitochondria. Physiol Plantarum 105: 763–768.

Webb MA (1999). Cell-mediated crystallization of calcium oxalate in plant. Plant Cell 11: 751–761.

Weiner H, Stitt M, Heldt HW (1987). Subcellular compartmentation of pyrophosphate and alkaline pyrophosphatase in leaves. Biochim Biophys Acta 893: 13–21.

Wroblewski T, Tomczak A, Michelmore R (2005). Optimization of Agrobacterium-mediated transient assays of gene expression in lettuce, tomato and Arabidopsis. Plant Biotechnol J 3: 259–273.

Zancani M, Macri F, Dal Belin Peruffo A, Vianello A (1995). Isolation of the catalytic subunit of a membrane-bound H+-pyrophosphatase from pea stem mitochondria. Eur J Biochem 228: 138–143.