Molecular Characterization of a Potato Gene Related to Plant Purple Acid Phosphatases

by

Allison B. Mackay

A thesis submitted to the Department of Biology

in conformity with the requirements for

the degree of Master of Science

Queen's University

Kingston, Ontario, Canada

January, 2001

copyright © Allison B. Mackay, 2001

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-55917-3

Canada

Abstract

Acid phosphatases (APases) catalyze the hydrolysis of orthophosphate esters at acid pH optima. APases have been detected in many plant tissues, including roots, tubers, leaves, and developing seedlings. APase induction is a major part of the plant phosphate-starvation response, and both cell-wall and secreted APases appear in the roots of phosphate-starved plants. Previous work identified an APase from potato tuber possessing unusually high activity. A genomic clone for potato tuber acid phosphatase (PTAP) was isolated, along with several other clones showing amino acid sequence homology to acid phosphatases. One of these, *pap7*, was sequenced and characterized in this study. Of the 18 kb clone, 6.6 kb of contiguous sequence was assembled, containing APase-homologous sequence. The sequence comprised seven predicted exons, with all the boundaries but one conforming to plant splice site consensus sequences. The predicted protein displayed 53% identity with PTAP, 63% identity with red kidney bean APase, and 61% identity with sweet potato APase. The amino acids believed to be critical for active site binding and catalysis in kidney bean and sweet potato APases were completely conserved in both PAP7 and PTAP. Southern blotting of genomic DNA indicates one copy of *pap7* is present in the potato genome. RNA from phosphate-fed and starved potato plants was probed with *pap7* but no expression was found in leaves, roots or stems. *Pap7* appears to be a member of a family of related plant acid phosphatase genes including the "purple" APases and a number of uncharacterized genes from diverse species. Although it does not appear to be phosphate-starvation induced, *pap7* could be expressed under other APase-inducing conditions, including senescence, salt stress, flowering, or seed germination.

Acknowledgements

I would first like to thank my supervisor, Dr. Daniel Lefebvre, for his continued support and advice on my research and the preparation of this thesis. Drs. Ken Ko and Keith Poole also deserve thanks for their service on my committee and the different perspectives they have often provided on the challenges this project has presented.

This thesis research is based on initial work performed by Dr. Kevin Gellatly in the course of his Ph.D. His continued advice and encouragement has been appreciated throughout this project. During my time in the Lefebvre laboratory, I have also been guided and assisted by others who have worked here: Ali Hannoufa, Ali Malboobi, Norman Njoroge, Linda Tremblay, and Peter Truesdell.

Of course, I have met many new friends during my years in the department. These individuals have helped make my graduate school experience a pleasant one, engaging me in discussions both professional and unprofessional. They include Stephanie Backman, Peter Chenaux, Marie-France Hetu, Erin Higgins, Richard Hodgson, David Law, Allison McDonald, Jennifer Moustgaard, Vanessa Murley, Fayek Negm, Lehli Pour, Sandy Reid, Christopher Smith, William Turner and many others.

Finally, my family deserve much credit for their encouragement and support, and for putting up with many grumpy rants about uncooperative experiments and writer's block.

Abbreviations

APase	acid phosphatase
Asp	aspartic acid
Asn	asparagine
bp	base pairs
CTAB	hexadecyltrimethylammonium bromide
DEPC	diethylpyrocarbonate
ddH_2O	distilled, deionized water
dNTP's	2'-deoxyribonucleoside 5'-phosphates
EDTA	ethylenediaminetetraacetic acid, disodium salt
HEPES	*N*-2-hydroxyethyl-piperazine-N'-2-ethanosufonic acid
His	histidine
IgG	immunoglobin G
IPTG	isopropyl β-D-thiogalactopyranoside
kb	kilobase pairs
KBPAP	kidney bean purple acid phosphatase
kDa	kilodalton
LB	Luria Broth (1% w/v tryptone, 0.5% w/v yeast extract, 1% NaCl; pH 7.5)
MOPS	3-(*N*-morpholino)propanesulfonic acid
MS	Murashige-Skoog medium
ORF	open reading frame
PAP	purple acid phosphatase
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque-forming units (bacteriophage λ)
P_i	orthophosphate
PMSF	phenylmethylsulfonyl fluoride
PSI	phosphate-starvation induced
PTAP	potato tuber acid phosphatase
PTPase	protein tyrosine phosphatase
PVP	polyvinylpyrrolidone
PVPP	polyvinylpolypyrrolidone
Ser	serine
SM	0.58% w/v NaCl, 0.2% w/v magnesium sulfate heptahydrate, 50 mM Tris pH 7.5, 0.01% gelatin
SDS	sodium dodecyl sulfate
SSC	3M NaCl, 0.3M sodium acetate, pH 7.0 (20x stock)
TE	10 mM Tris, 1mM EDTA, pH 8.0
Tris	tris(hydroxymethyl)aminoethane
Tyr	tyrosine
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Table of Contents

ABSTRACT....ii

ACKNOWLEDGEMENTS....iii

ABBREVIATIONS....iv

TABLE OF CONTENTS....v

LIST OF FIGURES....vii

CHAPTER 1: Introduction....1

CHAPTER 2: Literature Review

Phosphate and Plants....4

The Phosphate Starvation Response....5

Escherichia coli....5

Saccharomyces cerevisiae....6

Plants....8

Acid Phosphatases....15

Acid Phosphatase Characteristics....15

Specialized Acid Phosphatases....18

Purple Acid Phosphatases....20

CHAPTER 3: Materials and Methods....22

CHAPTER 4: Results....36

CHAPTER 5: Discussion

Coding Sequence Prediction....56

Polyadenylation Site Prediction....58

Pap7: Comparison with other APases....59

CHAPTER 5: DISCUSSION (cont.)

Non-APase Homologies....63

Expression Analysis....65

Conclusions and Future Work....67

LITERATURE CITED....69

APPENDICES....79

VITA....87

List of Figures

Figure 1 Diagrams of the *pap7* subclones isolated for sequencing......................37

Figure 2 Partial restriction map of the *pap7* genomic clone................................38

Figure 3 DNA sequence of the mapped portion of the genomic clone................39

Figure 4 Predicted amino acid sequence of Pap7..44

Figure 5 CLUSTAL alignment of Pap7 with other APase proteins.....................46

Figure 6 Phylogenetic relationship of PTAP and Pap7 to other APases.............49

Figure 7 CLUSTAL alignment of Pap7 and PTAP with PTPases.......................50

Figure 8 Southern blot of potato genomic DNA probed with *pap7*.....................53

Figure 9 Northern blot of potato RNA probed with *pap7*....................................55

Introduction

Acid phosphatases (APases) catalyze the dephosphorylation of orthophosphate monoesters. APases appear to be ubiquitous in plants and have been characterized in many species including *Arabidopsis thaliana* (del Pozo *et al.*, 1999), black mustard (Duff *et al.*, 1989a), maize (Miernyk, 1992), potato (Gellatly *et al.*, 1994), soybean (Ferriera *et al.*, 1998), and wheat (Van Etten and Waymack, 1991). Common features of these enzymes include pH optima between 5 and 6, a broad substrate specificity and inhibition by vanadate, molybdate and fluoride.

Purple acid phosphatases are a special class of APases with unique characteristics. Initially isolated from animals (eg. uteroferrin), they contain a binuclear metal center associated with the active site, and unlike regular APases, are not inhibited by tartrate. Purple acid phosphatases have been found in a number of plant species, but the one studied most extensively is from red kidney bean (Beck *et al.*, 1986). KBPAP is a 110 kDa dimeric glycoprotein, containing Fe(III) and Zn(II) at the active site. These ions are essential for activity via binding of the substrate's phosphate to Zn(II) and subsequent hydrolysis by hydroxide bound to Fe(III) (Klabunde *et al.*, 1996). Different PAP's can contain different metals – for example, sweet potato PAP contains Fe(III) and Mn(II) (Schenk *et al.*, 1999).

In addition to biochemical characterization, a number of plant APase genes have been cloned. APase-coding cDNAs have been isolated from *Arabidopsis* (Mullaney and Ullah, 1998; Utsugi *et al.*, 1998), potato (Gellatly, 1996), red kidney bean (Schenk *et al.*, 2000), rice (Gellatly, 1996), soybean (Mason *et al.*, 1988), sweet potato (Schenk *et al.*, 1999), and tomato (Erion *et al.*, 1991).

The acid phosphatase activity in potato has been studied for some time, using preparations of varying purity. Kubicz and colleagues (1974) co-purified three different isoforms from potato tuber. The isoforms differed in charge but not molecular weight. In another study (Sugawara *et al.*, 1981) six isozymes were isolated, with molecular weights of 250-400 kDa and pH optima between 4.5 and 6.0. It was not determined whether these were separate isozymes or just differentially glycosylated. Gellatly and colleagues (1994) isolated the major acid phosphatase from potato. The enzyme (PTAP) is a dimeric 100 kDa glycoprotein showing strong activity towards phosphotyrosine (1917 units/mg), pyrophosphate (1728 units/mg), and phospho*enol*pyruvate (764 units/mg), with a pH optimum of 5.8. It is most specific towards P-Tyr, with a specificity constant (V_{max}/K_m) of 1936, and was able to dephosphorylate potato phosphotyrosyl proteins *in vitro*. Cyanogen bromide cleavage of the subunits (57 and 55 kDa) showed that they consist of similar but not identical amino acid sequences. Although PTAP's role is not known, the enzyme was induced in fresh but not sprouting tubers, suggesting it is not involved in nutrient-mobilizing processes accompanying sprouting.

Two genomic clones and a cDNA encoding PTAP have been isolated (Gellatly, 1996). *PTAP-1* and *PTAP-2* (genomic) and *pPTAP* (cDNA) encode a 451 amino acid protein with a predicted molecular weight of 52 kDa. PTAP is strongly homologous to kidney bean purple acid phosphatase, with all of the residues (Asp, Asn, His, and Tyr) involved in the KBPAP active site fully conserved. Immunologically related proteins were detected in all potato tissues (dormant and sprouting tuber, tuber epidermis, stolon, root, stem and leaf) but *PTAP* mRNA was only detected in tuber and stolon.

Three genomic clones homologous to PTAP were also isolated (Gellatly, 1996). *Pap3*, *pap7*, and *pap11* all share deduced amino acid sequence homology with both PTAP and KBPAP. These clones were partially sequenced and the deduced amino acid sequences (from the portions sequenced) shared 54% (*pap3*), 65% (*pap7*), and 73% (*pap11*) identity with PTAP, with the KBPAP active site residues all conserved.

The purpose of this study was to sequence the *pap7* genomic clone, analyze the sequence, and determine whether it was expressed in potato plants during phosphate starvation. Sequence analysis indicates that this clone may contain part of a functional gene, encoding a protein highly homologous to PTAP and other acid phosphatases. Northern analysis did not indicate that *pap7* is induced under phosphate starvation in potato plants.

Literature Review

Phosphate and plants

The orthophosphate ion (PO_4^{2-}) is critical for plant (and other) cells, as it is incorporated into many important biomolecules including nucleic acids, phosphoproteins, phospholipids, and energy currency molecules such as ATP. The typical orthophosphate (P_i) concentration in soil is in the micromolar range, whereas in plant cytoplasm it is present in millimolar concentrations (Raghothama, 2000). Clearly, even under "normal" conditions, plants must maintain a steep P_i concentration gradient across the plasma membrane in order to ensure proper functioning of their metabolic and biosynthetic processes. In conditions of reduced environmental phosphate the effects on intracellular Pi supply are significant: the concentration of free P_i can be reduced to 1/40 of normal, and the concentrations of ATP and ADP can drop to 1/5 and 1/10 of normal, respectively (Duff *et al.*, 1989b).

Plants experiencing phosphate starvation exhibit a number of gross morphological changes. Compared to non-starved plants, they exhibit reduced growth, an increased root:shoot mass ratio (Lefebvre and Glass, 1982), reduced root diameter, and leaf purpling due to increased anthocyanin production (Carswell *et al.*, 1996). Phosphate-starved white lupin (*Lupinus albus*) and related species exhibit another visible alteration in root morphology: the growth of "proteoid" roots, short, densely clustered tertiary roots (Gilbert *et al.*, 2000).

The Phosphate Starvation Response

Escherichia coli

Microorganisms resemble plants in that they cannot relocate to a more favourable environment when they encounter conditions of nutrient deprivation. An effective biochemical system for surviving phosphate starvation would therefore be expected in both prokaryotic and eukaryotic microorganisms.

The phosphate-starvation response of prokaryotes has been studied extensively. largely using *Escherichia coli* as a model system. *E. coli* induces a number of phosphate uptake and metabolizing proteins during phosphate starvation, including alkaline phosphatase (PhoA; Shinagawa *et al.*, 1983), outer membrane porin (PhoE; Tommassen *et al.*, 1982), phosphate-binding protein (PhoS; Surin *et al.*, 1984), glutamate synthase, glycerol-3-phosphate transporter and glycerol-3-phosphate dehydrogenase (Metcalf *et al.*, 1990). Presumably, these proteins allow the cell to enhance its uptake and usage of phosphate and phosphorylated organic molecules from its surroundings under low-phosphate conditions.

The system by which the *E. coli* "pho regulon" is directly controlled has been elucidated. The PhoB protein is a transcriptional activator which binds to "pho box" consensus sequences in the promoters of *pho* genes such as *phoU* (outer membrane porin), *phoA* (periplasmic alkaline phosphatase), *pstS* (periplasmic phosphate-binding protein), and *pstA*, *B*, and *C* (inner membrane phosphate transporter complex; Rao and Torriani, 1990). PhoB is in turn activated (by phosphorylation) under low-phosphate conditions and deactivated under high-phosphate conditions by a second protein, PhoR, which somehow responds to external P_i concentration (Yamada *et al.*, 1989, Amemura *et*

al., 1995). This system is similar to other two-component genetic switches found in bacteria, and in fact the PhoB and PhoR proteins are homologous to a family of bacterial transcriptional activator/histidine protein kinase 2-component regulatory systems which respond to environmental stresses (Yamada *et al.*, 1989, Wanner, 1993).

The pho box is a sequence element found in varying copy numbers in the promoters of most *pho* genes. The consensus sequence for the *E. coli* pho box (as identified in the promoters of *phoA*, *phoE*, *pstS*, and *phoB*) is CT(G/T)TCATA(A/T)A(A/T)CTGTCA(C/T) (Makino *et al.*, 1986; Wanner, 1993). The PhoB protein has been shown to bind to two such adjacent sequences in the *pstS* promoter and to initiate transcription *in vitro* (Makino *et al.*, 1988). The pho box is actually a pair of sequences similar to CTGTCAT, and its influence in *pho* promoters is actually quite complex: the less similarity a promoter's pho boxes have to the consensus, the more are present in that promoter. As well, in one *pho* gene, *phoE*, the pho box compensates for a defective –10 promoter sequence ("Pribnow box") in initiating transcription (Scholten and Tommassen, 1994).

Saccharomyces cerevisiae

The phosphate-starvation response is significantly more complex in yeast than in bacteria. As a eukaryotic system, it is also of more relevance to the study of the plant phosphate-starvation response. Under phosphate-starvation conditions, *Saccharomyces cerevisiae* induces the production of three acid phosphatases (Pho5, Pho10, and Pho11), an alkaline phosphatase (Pho8), and a high-affinity inorganic phosphate transporter

(Pho84; Oshima *et al.*, 1996). A transcription factor, Pho4, activates expression of these genes, and is itself regulated at the post-translational level (Toh-e *et al.*, 1981).

Recently, a comprehensive model of the *S. cerevisiae* phosphate-starvation response has been constructed, based on extensive mutation studies and biochemical and genetic analysis. The transcription factor Pho4 is known to be repressed by phosphorylation at five serine residues under phosphate-fed conditions, while under phosphate-starved conditions this phosphorylation disappears and Pho4 (along with an additional factor, Pho2) binds to the promoters of *pho* genes (Oshima *et al.*, 1996). This phosphorylation is done by a cyclin-dependent protein kinase (CDK), Pho85, which is permanently complexed with the cyclin Pho80. The phosphorylating activity of Pho85/Pho80 is in turn regulated by another protein, Pho81, a CDK inhibitor whose transcription is induced under conditions of phosphate starvation (Lenburg and O'Shea, 1996). The end result of this system is that when the cell experiences phosphate starvation, Pho4 is dephosphorylated and the *pho* genes are expressed. The system incorporates a positive feedback loop: transcription of the Pho81 CDK inhibitor is positively regulated by Pho4 (Oshima *et al.*, 1996).

Like *E. coli*, *S. cerevisiae* has promoter sequence elements in common among phosphate-regulated genes. Promoter-deletion studies by Rudolph and Hinnen (1987) showed that the yeast *pho5* promoter contains four instances of the palindromic consensus sequence AYUTGYTAATTAUCAYUT, while the *pho11* promoter contained one copy of the sequence. Two of these elements were required for *pho5* expression: when they were removed, expression levels dropped to 0.2-0.1 times normal (Rudolph and Hinnen, 1987). These elements correspond to the sites of Pho4 binding to the *pho5*

promoter (Vogel and Hinnen, 1990). They are also necessary for *in vivo* nucleosome disassociation from the *pho5* promoter, which seems to be the critical event allowing transcription (Vogel and Hinnen, 1990). Other researchers (Ogawa *et al.*, 1994) found a 6-base pair sequence, CACGT(G/T), present in the promoters of *pho5*, *pho8*, *pho81*, and *pho84*, which bound to Pho4. Since these sequences overlap with three of the *pho5* sequences described previously, it is possible that the Pho4 binding site is slightly different in each instance, incorporating different combinations of sequence motifs and resulting in different Pho4 binding affinities.

It is still not clear how yeast cells sense phosphate starvation. It is possible that extracellular P_i concentration is sensed by the Pi-transport machinery: mutants in several transport-associated genes (including *pho84*) showed constitutive *pho5* expression (Lenburg and O'Shea, 1996). Of course, intracellular P_i concentration would probably be altered by a transporter mutation, affecting any putative intracellular Pi sensing system.

Plants

In addition to visible morphological changes, plants exhibit a well-characterized biochemical response to phosphate starvation, the components of which are similar across almost all species which have been studied. The two major cellular-level responses in plants are the induction of acid phosphatases, both intra- and extracellular, and the induction of high-affinity phosphate transporters.

Acid phosphatase (APase) induction was first noticed in aging suspension cell cultures which were depleting their supply of phosphate (Yamaoka *et al.*, 1969). In a subsequent study of tobacco suspension cells, Ueki and Sato (1971) found that secreted

acid phosphatase activity greatly increased when cells were placed in low-P_i cultures, and decreased again when the cells were re-fed P_i. It was also noted that inhibiting protein synthesis stopped the creation of new activity but not the release of pre-existing intracellular activity, suggesting that acid phosphatase was being synthesized *de novo* in response to phosphate starvation (Ueki and Sato, 1971).

Numerous studies on whole plants and suspension cells have confirmed this acid phosphatase induction. Studies of phosphate-starved plant roots (Dracup *et al.*, 1983, Szabo-Nagy *et al.*, 1987) found that in addition to secreted APases, much of the induced APase activity was bound to the root cell walls. Cell-wall APases could work similarly to secreted ones, by hydrolyzing phosphoesters in the immediate vicinity of the plasma membrane (in roots, the rhizosphere). However, results vary with respect to whether cell-wall APases are induced under phosphate starvation: Lefebvre and co-workers (1990) found that although secreted and intracellular APases were induced under phosphate starvation in *Brassica nigra* suspension cells, cell-wall APase activity did not correlate with medium [P_i]. Studies on the roots of clover (Dracup *et al.*, 1984) and winter wheat (Szabo-Nagy *et al.*, 1987), and on the proteoid roots of white lupin (Gilbert *et al.*, 1999) found that cell-wall APase activity was significantly higher in phosphate-starved plants. This discrepancy is possibly due to a tissue-specific expression pattern: only root tissue would require cell-wall phosphatases to hydrolyze phosphoesters from outside the plant, and suspension culture cells might not respond the same way differentiated root cells would.

Although APase induction is considered the hallmark of the phosphate-starvation response, plants also react to reduced phosphate supply by increasing their P_i intake

through membrane transporters. Increased phosphate uptake is a very rapid response to reduced external [Pi], more rapid in fact than APase induction. Studies of barley (Lefebvre and Glass, 1982) and *Brassica nigra* suspension cells (Lefebvre *et al.*, 1990) showed that increased influx occurs very quickly after the onset of phosphate starvation (approximately 1 day), while APase induction occurs somewhat later (3-6 days). Recently, cDNA library clones encoding phosphate transporters have been isolated from phosphate-starved *Arabidopsis thaliana* roots (Muchhal *et al.*, 1996), potato roots (Leggewie *et al.*, 1997), *Medicago truncatula* roots (Liu *et al.*, 1998b) and tomato roots (Muchhal and Raghothama, 1999). These transporters share amino acid sequence homology, and all display a similar 12-domain membrane-spanning structure. The *Arabidopsis* clones were able to complement a yeast high-affinity phosphate transporter mutant, while the potato clones partially restored high-affinity transport to a similar yeast mutant. All of these transporter genes showed increased expression under phosphate starvation, though at least in *M. truncatula* one of the transporters was much less sensitive to P_i status than the other (Liu *et al.*, 1998b). This appears to support enzymological studies showing that plants possess both constitutive and phosphate-starvation-induced phosphate-transport activities (Nandi *et al.*, 1987; Shimogawara *et al.*, 1999).

Of course, when phosphate becomes limiting, plant cells not only have to uptake and retain as much Pi as possible, but also maintain the cytoplasmic concentration at a level which allows crucial metabolic functions to continue. If the cytoplasmic [P_i] were directly responsive to the external [P_i], the cell could rapidly reach a state where even glycolysis was impossible. This is why the vacuolar P_i pool appears to act as a "buffer"

between cytoplasmic and external phosphate pools. Studies on sycamore cell-suspension cultures showed that under normal nutrition, cytoplasmic P_i concentration was about 6 mM, while in the vacuole it was 1.5 mM. Inside phosphate-starved cells, these P_i pools adjusted over time, with the vacuolar concentration decreasing rapidly, followed by a slower decrease in the cytoplasmic concentration (Rebeille *et al.* 1983). When the cells were subsequently fed phosphate, the cytoplasmic pool was replenished first. This system could serve to maintain an optimal cytoplasmic [P_i] needed for metabolism for as long as possible. Phosphate transporters in the tonoplast membrane could modulate this response by responding to changes in cytoplasmic [P_i] and moving P_i across the tonoplast membrane (Rebeille *et al.*, 1983).

Despite the cell's attempts to maintain cytoplasmic phosphate levels under starvation conditions, it is inevitable that under prolonged starvation, cytoplasmic [P_i], [ADP] and [ATP] will diminish (Duff *et al.*, 1989b). Hence, cells must be able to perform crucial metabolic reactions even as several of their fundamental substrates start to dwindle. In the glycolytic pathway, phosphofructokinase (PFK) requires ATP, NAD-glyceraldehyde-3-phosphate dehydrogenase (NAD-G3PDH) requires P_i, and both 3-phosphoglycerate kinase and pyruvate kinase require ADP in sufficient concentrations to function properly (Duff *et al.*, 1989b). It was found that under phosphate-starvation conditions, an alternative suite of enzymes were induced, which could effectively bypass the phosphate requirements of glycolysis. For example, pyrophosphate-dependent phosphofructokinase can substitute for PFK, using pyrophosphate instead of ATP, and non-phosphorylating NADP-glycerol-3-phosphate dehydrogenase could substitute for NAD-G3PDH and 3-PGA kinase using NADP as a phosphate source (Duff *et al.*, 1989b).

Another enzyme induced under phosphate starvation was an intracellular acid phosphatase with a strong substrate specificity for phospho*enol*pyruvate (PEP). This enzyme, which is strongly inhibited by P_i, ADP and ATP (and is thus unlikely to function well in phosphate-fed cells), could bypass the PK reaction without requiring scarce ADP (Duff *et al.*, 1989a). With a pH optimum of 5.6, it would probably function in the vacuole, producing pyruvate and P_i which could be transported to the cytoplasm (Theodorou and Plaxton, 1993).

The underlying question behind phosphate starvation research is, how do plants actually sense phosphate starvation and regulate the response? The most rapid response is increased P_i influx via the induction of high-affinity phosphate transporters, with APase activity exhibiting a similar increase significantly later (Lefebvre *et al.*, 1990). One possible way for plants to signal phosphate starvation would be with phytohormones. In fact, acid phosphatases are also induced under other conditions, with phytohormones appearing to be involved. APase activity was increased in barley seed aleurone during imbibition, and exogenous GA_3 appeared to increase the secretion of the activity (Bailey *et al.*, 1976). A similar result was found in seeds of wheat (Akiyama et al., 1981). IAA was found to induce APase secretion (but not *de novo* synthesis) in maize coleoptiles (Pfeiffer, 1996), and ABA induced expression of the *Arabidopsis* phosphate starvation-inducible APase AtACP5 (del Pozo *et al.*, 1999). Methyl jasmonate treatment has also been found to cause APase induction in detached leaves of rice (Yeh *et al.*, 1995). However, studies of phosphate starvation have not yielded similar links to phytohormones. Experiments on ABA-deficient and ABA-insensitive mutants of *Arabidopsis* by Trull and colleagues (1997) do not support a role for ABA in the

phosphate-starvation response: ABA mutants and wild-type plants displayed no differences in root:shoot ratio, intracellular [P_i], or APase induction during phosphate starvation. It is conceivable that other phytohormones may play a role in the phosphate-starvation response, but evidence is lacking.

The most promising insights into how plants may regulate the phosphate starvation response suggest that the phosphate status of the shoot may control phosphate acquisition by the roots. Dong and colleagues (1998) studied the *pho2* mutant of *Arabidopsis*, which accumulates excessive P_i in the shoots. They found that root phosphate influx and Pi translocation to the shoot was higher in *pho2* plants than wild type, and when the *pho2* shoots were removed, influx rates dropped to wild type levels. When phosphate-fed *pho2* plants were placed in –P medium, the shoot maintained a higher [Pi] and did not translocate as much Pi to the roots. This suggests that root-level phosphate-starvation adaptations (phosphate influx and possibly APase induction) are sensitive to the Pi flux between the shoot and root, and that the plant's sensing of the shoot's phosphate status governs that flux via shoot phosphate transporters either in the "sink" cells or the phloem. This is supported by divided-root experiments in which P-fed plants with half their roots placed in –P medium did not exhibit increased phosphate-transporter induction in those roots (Liu *et al.*, 1998a).

Although the regulation of phosphate starvation induced genes at the molecular level is still poorly understood, recent work has provided potential footholds for future discoveries. Malboobi and Lefebvre (1995) isolated a number of cDNA clones from phosphate-starved *Brassica nigra* cells, one of which shows homology to serine/threonine kinases. As phosphorylation is a common mechanism for control of

enzyme function, this could represent part of a signaling pathway activating gene expression under phosphate starvation. Wykoff and co-workers (1999) have cloned a gene, *psr1*, from the green alga *Chlamydomonas reinhardtii*, mutants of which are deficient in phosphate starvation-specific responses such as transporter induction and alkaline phosphatase secretion. The Psr1 protein is phosphate starvation-induced and localized to the nucleus. It contains sequence elements homologous to DNA binding, transcription factor dimerization, and trancriptional activation domains from plant TF's, and is homologous to several plant expressed sequence tags of unknown function. Psr1 could very well be a phosphate starvation-induced transcription factor controlling some or all of the genes induced under phosphate stress.

Recently, two discoveries have suggested important similarities in the promoters of plant phosphate-starvation induced genes. Burleigh and co-workers (1998) isolated *Mt4*, a gene upregulated under phosphate starvation in *Medicago truncatula*. *Mt4* transcript appeared only in roots and was dramatically reduced as medium P_i concentration increased. The promoter region of *Mt4* contained a CACGTG motif, similar to the CACGTG/T motif found in S. cerevisiae *Pho84* and tomato *TPSI1* promoters, both of which are phosphate-starvation induced (Burleigh *et al.*, 1998). Haran and co-workers (2000) cloned the promoter region of an *Arabidopsis* acid phosphatase gene induced only under phosphate starvation. The APase was expressed in leaves and roots and appeared to be secreted from cells. The promoter contained a CACTTG motif similar to the CACGTG from the *Mt4* promoter. If similar motifs are found in other phosphate-starvation inducible plant promoters, and are shown to be required for promoter function, it will be a significant advance in the understanding of plant PSR

control. If such elements are shown to interact with proteins such as the *Chlamydomonas* Psr1 (Wykoff *et al.*, 1999), it will be a step towards understanding plant PSR control as thoroughly as that in yeast.

Acid Phosphatases

Acid phosphatases (APases; E.C. 3.1.3.2) are enzymes which hydrolyze orthophosphate esters, releasing orthophosphate (P_i). APases typically have pH optima between 5.0 and 6.0 (Duff *et al.*, 1991b), although the pH optimum can vary between substrates and the enzymes can often function at neutral pH (Duff *et al.*, 1989a, Cashikar *et al.*, 1997). Plants also possess alkaline phosphatases, which have pH optima above 7 and, unlike APases, typically display absolute substrate specificity (Duff *et al.*, 1994). Alkaline phosphatases also hydrolyze phosphoesters via a different reaction mechanism than APases (Neuman, 1968; Vincent *et al.*, 1992).

Acid phosphatases appear to be ubiquitous in plants. They have been shown to occur in different tissues and developmental stages of the same species, in intracellular, secreted and cell-wall forms, and to appear in response to nutrient starvation (Duff *et al.*, 1994). Plant acid phosphatases can be divided into several different groups based on a number of physical and biochemical characteristics.

Acid Phosphatase Characteristics

The majority of APases characterized in plants appear to be substrate non-specific. Although they may display increased efficiency hydrolyzing certain substrates,

their ability to also hydrolyze other substrates prevents them being classified as "specific" in the same way that many enzymes (such as those of glycolysis) are.

Non-specific APases have been identified in numerous plant species, in various tissues under different environmental stresses and developmental stages. These APases are generally mono- or dimeric, with subunit molecular weights of 30-60 kDa (although molecular weights as high as 400 kDa have been observed; Sugawara *et al.*, 1981), and are glycoproteins.

APases have been characterized in many tissues, including roots, tubers, and leaves, and seeds and cotyledons of developing seedlings. APases have also been characterized from suspension cell cultures of black mustard (Duff et al., 1989a; Duff et al., 1991a), maize (Miernyk, 1992), rice (Igaue et al., 1976), and tomato (Paul and Williamson, 1987). Unfortunately, this strategy provides no information about any tissue-specific expression patterns of the isoforms characterized. A number of studies have reported APase activity in roots, although few have characterized purified enzyme preparations. Dracup and co-workers (1984) reported APase activity in the cell walls of roots of subterranean clover, and found that the activity increased as plants were starved for phosphate. Lee (1988) found "externally accessible" (cell-wall or secreted) acid phosphatase activity in the roots of phosphate-starved barley, and Pfeiffer (1998) reported an extracellular acid phosphatase from the growing root tips of maize. Gilbert and colleagues (1999) found APase activity in both primary and proteoid roots of white lupin, with secreted APase activity increasing under phosphate starvation. Several researchers have purified and characterized root APases. A 42 kDa soluble APase was characterized from phosphate-starved wheat roots (Szabo-Nagy *et al.*, 1987). It displayed non-absolute

substrate specificity, preferentially hydrolyzing ATP at pH 6.5 and pyrophosphate at pH 5.0 (Szabo-Nagy *et al.*, 1987). Panara and co-workers (1990) characterized the major soluble APase from barley roots: a 78 kDa homodimer with a broad pH optimum of 5.2-6.8, preferentially hydrolyzing ATP and phosphotyrosine.

Sugawara and co-workers (1981) reported six APase isozymes from potato tuber, ranging in size from 250 to 400 kDa. The isozymes were localized in the cell wall and showed a substrate preference for adenylates, though a limited number of substrates were tested. Gellatly and co-workers (1994) purified a 100 kDa dimeric potato tuber APase with strong phosphotyrosine phosphatase activity. APases have also been characterized from membrane and soluble fractions from yam tuber (Kamenan and Diopoh, 1983), and from sweet potato (Durmus *et al.*, 1999).

APases have been studied in seeds from a number of species, including kidney bean (Beck *et al.*, 1986), wheat (Van Etten and Waymack, 1991), *Vigna sinensis* (Biswas and Cundiff, 1991), poppy (Chung and Polya, 1992), soybean (Fujimoto *et al.*, 1977; Ferriera et al., 1998), lentil (Bose and Taneja, 1998), and castor bean (Granjeiro *et al.*, 1999). Although some of the enzymes appear to be nonspecific APases, the poppy, soybean, and lentil enzymes show significant activity towards phosphorylated amino acids (P-Tyr and P-Ser) which suggests they may have some regulatory functions as protein phosphatases (Chung and Polya, 1992; Bose and Taneja, 1998).

Different APases can have widely varying effectors. In general, vanadate, molybdate, and fluoride effectively inhibit APase activity (Randall and Tolbert, 1971; Gellatly, 1996), and several reports have shown lead and mercury cations to be strong inhibitors as well (Randall and Tolbert, 1971; Panara *et al.*, 1990). P_i itself is also an

inhibitor of APases, suggesting that feedback inhibition may be a common method of regulating phosphatase activity. In fact, ions such as tartrate, molybdate, and vanadate may act as inhibitors because of their structural similarity with P_i (Duff *et al.* 1994). The effects of the divalent metal cations such as Mg^{2+}, Zn^{2+}, Mn^{2+}, Ca^{2+}, and Cu^{2+} can vary between different APases: they can increase APase activity (Sugawara *et al.*, 1981; Duff et al., 1989a; Ferriera et al., 1998), partially inhibit it (Randall and Tolbert, 1971; Cheng and Tao, 1989; Panara *et al.*, 1990; Bose and Taneja, 1998; Granjeiro *et al.*, 1999), or both, depending on the specific enzyme. For example, yam tuber APases were activated by Mg^{2+}, but inhibited by Ca^{2+} (Kamenan and Diopoh, 1983). Gellatly and colleagues (1994) found that potato tuber APase was activated by Mg^{2+}, Ca^{2+}, Co^{2+} and Mn^{2+}, but inhibited by Zn^{2+}.

Specialized Acid Phosphatases

Although many plant APases do not appear to demonstrate specificity towards any particular substrate, some have been identified which show strong preference for a particular substrate. Examples of this are 3-phosphoglycerate phosphatase (Randall and Tolbert, 1971) and phospho*enol*pyruvate (PEP) phosphatase (Duff *et al.*, 1989a). These enzymes may be involved in specific steps of metabolism, and would require a degree of specificity.

There are two larger groups of specialized APases which bear closer examination, the phytases and the protein phosphatases. Phytases are enzymes which hydrolyze phytate (myo-inositol-hexakisphosphate), an organic polyphosphate compound found in plant seeds as well as soil. They typically have pH optima between 4.0 and 5.6 (Duff *et*

al., 1994), and are expressed in germinating seeds. Phytases are believed to have a role following germination, mobilizing P_i and mineral ions from phytate which is stored in the plant embryo (Maugenest *et al.*, 1999). Phytate is found in soil (sometimes making up 25% of the organic phosphoesters), but studies have shown that although phytase is strongly induced in roots of phosphate-starved plants, the plants are not able to sustain a normal phosphate nutritional status when phytate is the sole Pi source (Richardson *et al.*, 2000). However, phytate is known to be present in root tissues, and it is possible that the enzyme is involved in mobilizing root phytate reserves during phosphate starvation (Maugenest *et al.*, 1999).

Protein phosphatases catalyze the removal of phosphates from tyrosine, serine or threonine residues of proteins. As phosphorylation state is known to control the activity of many enzymes and transcription factors, protein phosphatases and kinases can be important regulatory proteins. Tyrosine phosphorylation is involved in cell growth regulation in animal cells (Pot and Dixon, 1992) and in regulation of the *Schizosaccharomyces pombe* cell cycle (via CDK phosphorylation; Charbonneau and Tonks, 1992). Although typically, proteins are activated by phosphorylation (such as by MAP kinases) the reverse can also occur, with dephosphorylation causing activation. One example in plants is sucrose phosphate synthase (Huber and Huber, 1996). Although tyrosine phosphatases acting in the cytosol would have to function at neutral pH, the reaction mechanisms of P-Tyr phosphatases and nonspecific acid phosphatases are similar, with an orthophosphate-enzyme intermediate being formed (Pot and Dixon, 1992; Vincent *et al.*, 1992). Several APases have been shown to exhibit significant P-Tyr

phosphatase activity (Cheng and Tao, 1989; Chung and Polya, 1992; Gellatly *et al.*, 1994; Ferriera *et al.*, 1998; Bose and Taneja, 1998).

Purple Acid Phosphatases

Purple acid phosphatases (PAP's) were first characterized in mammalian tissues, including pig uterine fluid (uteroferrin), bovine spleen and human osteoclasts. Animal PAP's are monomeric proteins of approximately 35 kDa, containing an Fe(III)-Fe(II) metal center (Klabunde *et al.*, 1995). Plant PAP's have been purified from many species including sweet potato (Uehara *et al.*, 1974), soybean (Fujimoto *et al.*, 1977), rice (Igaue *et al.*, 1976), and duckweed (Nakazato *et al.*, 1998), but the best-characterized plant PAP is from red kidney bean (*Phaseolus vulgaris*). The kidney bean enzyme (KBPAP) is a 110 kDa dimeric glycoprotein, containing an Fe(III)-Zn(II) metal center (Beck *et al.*, 1986). These metal ions are coordinated with several separate tyrosine, histidine, asparagine and aspartic acid residues, which are conserved between uteroferrin and KBPAP, even though the overall amino acid sequence similarity is only around 20% (Klabunde *et al.*, 1995). The *in vivo* substrate for KBPAP is not known, however it displays maximal activity against ATP, polyphosphate, and pyrophosphate. It appears to have no activity against phosphoproteins, phosphoglycerate, sugar-phosphates, or phytate (Cashikar *et al.*, 1997). Since KBPAP appears to be localized in the cell walls of peripheral cotyledon cells, a possible *in vivo* role could be scavenging phosphate out of surrounding soil to feed the developing embryo (Cashikar *et al.*, 1997). Several interesting but less substantiated alternate roles for KBPAP have been proposed: involvement in cell-wall breakdown during germination (Cashikar *et al.*, 1997), and the

removal of reactive oxygen species from the cell through reduction of the Fe(III) to Fe(II) (Klabunde *et al.*, 1995).

Materials and Methods

All reagents and chemicals were obtained from Sigma-Aldrich (Oakville, Ontario) or Fisher Scientific (Nepean, Ontario) if not otherwise indicated.

Genomic Library Clones

A λEMBL3 (SP6/T7) genomic library (purchased from Clontech, Palo Alto, California) had previously been screened by Dr. Kevin Gellatly using a degenerate oligonucleotide synthesized from PTAP N-terminal amino acid sequence. The *pap7* clone was one of five clones isolated from the library.

Bacteriophage screening and plaque lifts

In order to ensure the purity of the *pap7* genomic clone the λ bacteriophage were plated and screened with a known sequence from *pap7*. *Escherichia coli* cells (BB4) were cultured overnight in LB medium with 10 mM $MgSO_4$ and 0.2% w/v maltose. For each of three 100 mm plates 300 μl of cells were mixed with $2x10^4$ pfu of λ phage and incubated at room temperature for 20 minutes. The infected cells were mixed with 2.5 ml of top agar (0.7% LB-agar with 10 mM $MgSO_4$) and plated on LB-agar (1% agar w/v) plates containing 10 mM $MgSO_4$ and 0.2% w/v maltose and incubated at 37°C overnight. This resulted in approximately 30-40 well-separated phage plaques on each plate.

Circular Nytran membranes (Schleicher & Schuell, Keene, New Hampshire) were pre-soaked in 5x SSC, then placed on the plates' surface, marked for orientation, and incubated at room temperature for 5 minutes. The membranes were placed face-up on

Whatman 3M paper (Whatman, Clifton, New Jersey) soaked in 0.5 M NaOH / 1.5 M NaCl for 5 minutes, then neutralized with 1 M Tris-HCl (pH 8.0) for 5 minutes. The membranes were then incubated with 0.1 M Tris-HCl (pH 7.5) / 2x SSC for 5 minutes and briefly washed in 5x SSC. To immobilize phage DNA on the membranes they were air-dried and placed face-down on an ultraviolet transilluminator for 4 minutes.

Probe synthesis and colony hybridization

To probe the membranes, random-primed probes were made from a 1.7 kb *Eco*RI/*Bam*HI fragment of *pap7* previously shown (Gellatly, 1996) to be homologous to PTAP. Two micrograms of the fragment (digested out of the vector and gel-purified) were diluted with 17 µl of water and heated to 95°C for 4 minutes and chilled on ice. The Renaissance Random Primer labeling kit (NEN, Boston, Massachusetts) was used to generate fluorescein-labeled primers from the denatured DNA. The DNA was mixed with random primers and reaction buffer (5 µl), fluorescein-labeled nucleotides (5 µl), and DNA polymerase I Klenow fragment (1 µl; approx. 2 units). The synthesis reaction proceeded for 1 hour at 37°C and was stopped with 5 µl of 0.1 M EDTA (pH 8.0). Before use the probe solution was mixed with 300 µl of hybridization buffer (5x SSC, 0.1% w/v SDS, 0.5% Blocking Reagent; NEN, 5% w/v dextran sulfate) and heated to 95°C for 4 minutes, then cooled on ice.

Before beginning the probing steps, the membranes with bound bacteriophage DNA were wetted thoroughly in 2x SSC. The membranes were then incubated in a hybridization oven (Bio/Can Scientific, Mississauga, Ontario) with 15 ml of hybridization buffer for 1 hour at 65°C. For hybridization, the membranes were

incubated overnight at 65°C with 15 ml of hybridization buffer containing 50 μg/ml of salmon sperm DNA and approximately 20 ng/ml of the probe. After hybridization, the membranes were washed twice for 30 minutes at 65°C, first with 2x SSC, 1% SDS, then with 0.2x SSC, 0.1% SDS.

All subsequent probe detection steps were done at room temperature. Prior to probe detection, the membranes were agitated in Buffer 1 (0.15 M NaCl, 0.1 M Tris-HCl pH 7.5) for 5 minutes to remove SDS. The membranes were then blocked for 1 hour in Buffer 2 (Buffer 1 with 0.5% Blocking Reagent). The membranes were incubated for 1 hour in buffer 2 containing 1/1000 dilution of anti-fluorescein-HRP conjugate (NEN), then rinsed 4 times in Buffer 1. The membranes were then incubated in a 1:1 mixture of Enhanced Luminol Reagent and Oxidizing Reagent (NEN) for 1 minute, then blotted dry. The membranes were wrapped in plastic wrap and placed in an autoradiography cassette with X-ray film (Kodak, Rochester, New York) to detect the probe.

Growth of bacteriophage and DNA purification

Two plaques which hybridized to the probe were chosen as sources of phage DNA. Single plaques were excised from the plates and placed in 400 μl of SM buffer (0.58% w/v NaCl, 0.2% w/v $MgSO_4$ septahydrate, 50 mM Tris pH 7.5, 0.01% w/v gelatin) for 4 hours at 4°C to elute the phage particles. *E. coli* cells (BB4) were grown overnight in LB medium containing 0.2% maltose and 10 mM $MgSO_4$. 100 μl of cells were combined with 100 μl of eluted phage and 100 μl of 10 mM $MgCl_2$ / 10 mM $CaCl_2$ and incubated at 37°C for 15 minutes. The mixture was added to 50 ml of LB medium containing 10 mM $MgSO_4$ and incubated at 37°C until the culture contained stringy cell

debris indicating lysis. 2 drops of chloroform were added and the culture was centrifuged at 12,000 *g* for 10 minutes at 4°C. The supernatant (lysate) was retained and stored at 4°C in glass vials with a few drops of chloroform.

Pap7 bacteriophage clone DNA was purified from plate lysate using Lambdasorb Phage Adsorbent (Promega, Madison, Wisconsin). Prior to the procedure, *E. coli* cells (C600) were grown overnight in LB medium (with 10 mM $MgSO_4$, 0.2% maltose) and resuspended to an A_{600} of 0.5 in 10 mM $MgSO_4$. The *pap7* λ liquid lysate (see above) was diluted 1/5000 prior to use.

For a typical preparation, twelve 10 cm Petri dishes were each filled with 30 ml of 1% LB-agar. Two hundred microlitres of λ lysate and 300 μl of cells were mixed and incubated at 37°C for 30 minutes, then plated in 2.5 ml of top agarose (LB with 0.5% agarose and 10 mM $MgSO_4$ added) and incubated overnight at 37°C. When the plaques were almost confluent, 3 ml of SM was pipeted onto each plate, and the top agarose scraped with it into a 30 ml Corex glass centrifuge tube and incubated, shaking, for 30 minutes at room temperature. The tubes were then centrifuged at 10,000 *g* for 15 minutes at 4°C and the supernatant filtered through two layers of lint-free wipes (Kimberley-Clark, Dallas, Texas) and retained. One hundred microlitres of Lambdasorb were added to each 10 ml of lysate and incubated with shaking at room temperature for 30 minutes. The lysate was then centrifuged at 14,000 *g* for 15 minutes at 4°C and the supernatant carefully discarded. The pellet was twice resuspended in 1 ml of SM and spun down in a microcentrifuge, then resuspended in 500 μl of 10 mM Tris-HCl (pH 7.8), 10 mM EDTA. The suspension was heated to 67°C for 5 minutes to release the λ DNA, and the debris pelleted by centrifugation at 13,000 *g* for 5 minutes. Fifty microlitres of 5 M NaCl

were added per 500 μl of supernatant, then the supernatant was extracted twice with 25:24:1 phenol/chloroform/isoamyl alcohol, then with 24:1 chloroform/isoamyl alcohol. One-half volume of 5M NaCl and one-third volume of 30% w/v PEG 8000 were added and the supernatant was incubated on ice overnight.

The PEG precipitation was centrifuged at 13,000 *g* and the pellet retained and resuspended in 300 μl of TE buffer. The DNA solution was extracted once with phenol/chloroform/isoamyl alcohol and once with chloroform, then precipitated by adding 0.1 volume of 3M sodium acetate (pH 5.2) and 2 volumes of 100% ethanol and leaving at –80°C for 30 minutes. The DNA pellet was washed once with 75% ethanol, dried, and resuspended in 100 μl of TE buffer. The DNA was quantified by measuring the A_{260} with a spectrophotometer.

Restriction digestion of clone and subcloning of restriction fragments

All restriction enzymes used for cloning and mapping were obtained from Promega and MBI Fermentas (Burlington, Ontario). *Eco*RI and *Eco*RI/*Sal*I fragments of the *pap7* clone were subcloned into pBluescript SK+ (Stratagene, La Jolla, California) for subsequent manipulations. One microgram of phage DNA and 500 ng of plasmid DNA were each completely digested with 1 U of *Eco*RI or 1 U each of *Eco*RI and *Sal*I. The *pap7* phage DNA was mixed with approximately 0.1 μg of plasmid, 0.5 μl (0.5 units) of DNA ligase (Promega), and 0.5 μl of ligase buffer in 10 μl total volume and incubated at 16°C overnight.

Heat shock transformation was used to insert the ligated plasmid and fragments into *E. coli* cells. Fifty microlitres of frozen cells (pre-prepared competent DH-5α, Gibco

BRL/Life Technologies, Burlington, Ontario) were thawed slowly on ice. One microlitre of ligation mix was added to the cells and gently stirred. The cells were then incubated on ice for 30 minutes, followed by heat shocking at 37°C for 30 seconds and re-cooling on ice for 2 minutes. The cells were added to 1 ml of LB medium and incubated for 1 hour, shaking, at 37°C.

The cells were then plated on blue-white colour selection plates to select transformants with plasmid inserts. The plates contained 100 μg/ml of ampicillin and were coated with 40 μl each of 0.1M IPTG and 50 μg/μl X-gal (Promega). The plates were incubated at 37°C overnight. A total of 15 white colonies were selected from several plates and each was used to inoculate 2 ml of LB-amp and grown overnight. Plasmid DNA was purified from the liquid cultures using the Qiaprep Spin Column kit (Qiagen, Mississauga, Ontario) and quantitated spectrophotometrically. Aliquots of purified DNA were digested with *Eco*RI and *Sal*I and separated on an agarose gel to determine insert sizes.

Sequencing and analysis of subclones

Subclones were sequenced by MOBIX Inc. (Hamilton, Ontario). Sequence traces were manually transcribed into a computer and assembled into contigs using EditSeq and SeqMan, components of the Lasergene software package (DNAStar, Madison, Wisconsin). Sequence homology searches were done using BLAST (Altschul et al., 1990) on the NCBI web site at http://www.ncbi.nlm.nih.gov/BLAST/. Intron/exon boundary prediction was assisted by using the program SplicePredictor (Usuka and Brendel, 2000) on the web at http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi. Protein

sequence alignments and cladogram construction were performed with the CLUSTAL algorithm (using the software's default parameters: gap penalty=10, gap length penalty=10; Higgins and Sharp, 1988), using the Lasergene program MegAlign.

Probe pap7p1 synthesis and labeling

For the probing of Southern blots, a probe was designed corresponding to the end of the putative *pap7* coding sequence (the last exon) and the sequence immediately 3' to it. A 574 bp region was amplified by PCR, using genomic clone DNA as the template. The parameters were 94°C for 2 minutes, 55°C for 1 minute, and 72°C for 1 minute, for 30 cycles. PCR was performed using an MJ Research (Watertown, Maine) Minicycler.

The fragment was purified from an agarose gel using the Sephaglas BP kit (Amersham-Pharmacia, Uppsala, Sweden). Probe labeling was done using the Prime-a-Gene kit (Promega). Approximately 50 ng of gel-purified restriction fragment or PCR product was diluted to 10 µl volume and denatured by boiling for 2 minutes and cooling on ice. The denatured DNA was added to 10 µl of 5x kit labeling buffer, 2 µl of unlabeled dNTP's (G, A, and T; 500 µM each), 2 µl of BSA (10 mg/ml), 5 µl of ^{32}P-dCTP (approx. 50 µCi), and 7.5 units of DNA polymerase (Klenow fragment, Promega) in a reaction volume of 50 µl. The reaction proceeded for 90 minutes at room temperature and was stopped by boiling for 2 minutes. The probe was purified using a Nick size-fractionation column (Life Technologies, Burlington, Ontario). The column was equilibrated with TE buffer (pH 7.4) and the labeling reaction was loaded onto the column, followed by 350 µl of TE. Another 400 µl of TE was loaded and the second 400

μl eluate fraction was retained. Label incorporation was determined from 1 μl samples using a scintillation counter.

Southern blotting and hybridization

Southern blotting was used both in partial digestion mapping experiments and in determining genomic copy number of *pap7*. For potato genomic DNA Southern blots, 10 μg of DNA was loaded per lane. For plasmid and bacteriophage clones, an appropriate amount was loaded in each lane to ensure visibility on the gel.

Restriction-digested DNA was separated on an agarose gel containing 10 μg/ml ethidium bromide. The DNA was visualized on a UV transilluminator and photographed. The DNA was fragmented by soaking the gel in 0.25 M HCl for 30 minutes. Then the DNA was denatured by soaking in 1.5 M NaCl / 0.5 M NaOH twice for 20 minutes each. The gel was neutralized with 1.5 M NaCl / 0.5 M Tris pH 7.0 twice for 15 minutes. The gel was rinsed with distilled water between steps.

The gel was placed on a wick of Whatman 3M paper which was draped into a reservoir of 20x SSC (3M NaCl/0.3M sodium citrate, pH 7). A sheet of Hybond Plus nylon membrane (Schleicher and Schuell), cut to size and wetted with distilled water, was placed on top of the gel. Three sheets of Whatman 3M paper were soaked in 20X SSC and layered on top of the membrane. A stack of paper towels, cut to size, was placed on top of the filter papers and topped with a glass plate and a weight. The transfer was allowed to proceed overnight and the membrane was rinsed in 2X SSC and air-dried. The DNA was fixed to the membrane by exposure to UV light on a UV transilluminator for 4 minutes.

The membrane was pre-hybridized overnight at 65°C in hybridization buffer (5x SSC, 5x Denhardt's, 0.5% SDS) containing 40 μg/ml denatured herring sperm DNA. An appropriate amount of denatured radioactively labeled probe was added and the hybridization left overnight at 65°C. After hybridization, the membrane was washed twice for 15 minutes at room temperature with low-stringency wash (2x SSC, 0.1% SDS), twice for 15 minutes at 65°C with medium-stringency wash (1x SSC, 0.1% SDS), and twice for 15 minutes at 65° in high-stringency wash (0.1x SSC, 0.1% SDS). The membrane was rinsed in 2x SSC to remove SDS and results visualized with X-Omat Blue XB-1 autoradiography film (Kodak).

Mapping of subclones

The *pap7* subclones were mapped by partial *Eco*RI digestion. Twenty-five micrograms of λ clone DNA were digested to completion with *Sal*I. The DNA was then added to a 100 μl digestion containing 20 μl of 10x Tango buffer (MBI Fermentas). The reaction was divided among five microcentrifuge tubes, 30 μl in the first, 20 μl in the second, third, and fourth, and 10 μl in the fifth, and the tubes placed on ice. Ten units (1 μl) of *Eco*RI were added to the first tube and mixed, then 10 μl of the mixture added to the second tube. Ten microlitres from the second tube were added to the third, and so on, serially diluting the enzyme. The reactions were incubated at 37°C for 15 minutes, and stopped by heating to 65°C for 10 minutes. The reactions were run on a 1% agarose gel and photographed, then Southern blotted onto a Nytran N+ membrane.

To map the clone, radiolabeled probes were made of different restriction fragments and used to probe the blot. The sizes of the partial digest fragments they

hybridized to indicated which *Eco*RI and *Eco*RI/*Sal*I fragments were adjacent to one another. Combined with sequencing data this allowed mapping of the region of the clone containing the *pap7* sequence.

Growth of potato plants

Potato plants were derived from locally obtained sprouting tubers (Yukon Gold variety). Slices of tuber were planted in flats of Pearlite and grown in a growth chamber until the plants were approximately 10 cm tall. 24 plants were then transferred to a drip-culture hydroponics apparatus in a greenhouse. The plants were supported in a nutrient-free fragmented clay matrix in separate baskets, each irrigated by a drip spigot. The phosphate-fed (+P) plants were cultured in 0.1x MS medium (Murashige and Skoog, 1962) without vitamins. The phosphate-starved (-P) plants were grown in 0.1x MS medium with 0.13 mM KCl substituted for 0.13 mM K_2HPO_4. The apparatus was filled with fresh medium every 3-4 days to prevent depletion of the nutrients.

Tissue Protein Extraction and Phosphatase Assays

Tissues were harvested after 31 days. Leaf, stem and root tissues were separated and immediately frozen in liquid nitrogen and placed at -80°C. To confirm that plants were responding to phosphate starvation a phosphatase assay was done on leaf tissue. For each assay two leaf discs, excised with a #5 cork borer, were used. The assay was performed in triplicate for both +P and –P plants. 300 µl of extraction buffer (50 mM Tris pH 7.5, 3 mM $MgCl_2$, 20% w/v glycerol, 2% w/v PVPP, 1 µM benzamidine, 1 µM ε-aminocaproic acid) and 1.5 µl of 200 mM phenylmethylsulfonylfluoride (PMSF) were

added to each sample which was then ground with a Polytron homogenizer (Brinkmann Instruments, Westbury, New York) and small generator until homogenized. The samples were then centrifuged at 13,000 *g* for 2 minutes and the supernatant retained. Aliquots of supernatant (5 μl) were incubated with 800 μl of 50 mM sodium acetate (pH 5.0), 1.25 μM *p*-nitrophenol phosphate for 10 minutes at 37°C and the reactions were stopped with 200 μl of 1N NaOH. The A_{405} of the samples were then measured with a spectrophotometer to determine *p*-nitrophenol produced.

In order to normalize the phosphatase activities, protein concentrations in the samples were measured using a modified protocol of Bradford (1976). One microlitre of supernatant was added to 200 μl of dye concentrate (Bio-Rad, Mississauga, Ontario) and 800 μl of water. The reactions were incubated for 10 minutes at room temperature and the samples' A_{595} measured. Bovine IgG was used to produce a standard curve for the assay.

RNA isolation and Northern blotting

After 31 days in phosphate-free medium, the starved plants exhibited external signs of phosphate starvation (dark green foliage, smaller size than the control plants) as well as significantly greater phosphatase activity in their tissues (see Results). Leaf, stem and root tissues were harvested and immediately frozen in liquid nitrogen, then stored in sealed tubes at –80°C.

All RNA isolation and Northern blotting steps were done in pre-sterilized plasticware or DEPC-treated glassware and tubes. Solutions were treated with 0.2%

DEPC for 2 hours and autoclaved, or made with DEPC-treated ddH_2O if they contained ethanol, Tris-HCl or MOPS.

Total RNA was isolated from tissues using TRIzol reagent (Gibco BRL, Burlington, Ontario). 100-500 mg of frozen tissue was ground to powder with a pre-chilled mortar and pestle in liquid nitrogen. Samples were then mixed with 1 ml of TRIzol per 100 mg (all subsequent amounts are per 100 mg of sample), homogenized with a Polytron and small generator and incubated at room temperature for 5 minutes. Each sample was mixed with 1/5 volume of chloroform, shaken for 15 seconds, and incubated at room temperature for 3 minutes. The samples were centrifuged at 12,000 *g* for 15 minutes at 4°C and the clear upper phase retained. To remove carryover, the upper phase was loaded into 1.5 ml microcentrifuge tubes and centrifuged at 13,000 *g* for 2 minutes and the supernatant retained. The aqueous phase was vigorously mixed with 1 volume of isopropanol and incubated at room temperature for 10 minutes, then centrifuged at 13,000 *g* for 10 minutes at 4°C. The pellet was washed with 1 ml of 75% ethanol and air-dried briefly, then resuspended in 20-50 µl of water. The amount of RNA in each sample was quantitated spectrophotometrically.

RNA was electrophoresed in a 1% agarose gel containing 1x MOPS running buffer and 2.2 M formaldehyde. Samples were adjusted to 20 µl in 1x MOPS, 2.2 M formaldehyde, 50% formamide and incubated for 5 minutes at 55°C, and 2 µl of ethidium bromide RNA loading buffer added before loading. The RNA was visualized with a UV transilluminator. To determine band sizes, a DNA size marker (Mass Ladder, MBI Fermentas) was run on the gel. Prior to blotting, the gel was soaked in 0.05 M NaOH /

1.5 M NaCl for 30 minutes and 0.5 M Tris-HCl (pH 7.4) / 1.5 M NaCl for 30 minutes. The gel was then equilibrated in 20x SSC for 45 minutes.

The RNA was blotted onto a nylon membrane by capillary transfer. A large glass baking dish was half-filled with 20x SSC and a glass plate laid across it. A strip of Whatman 3M paper the same width as the gel was soaked in 20x SSC and laid across the glass plate with the ends hanging down into the buffer. The gel was laid on the paper, with cut strips of Parafilm placed around it. A Hybond Plus nylon membrane (Schleicher and Schuell), cut to the size of the gel and equilibrated in 20x SSC, was laid on top of the gel. Air bubbles were removed from between the layers by rolling with a glass pipet. Three sheets of Whatman 3M paper and a stack of paper towels, all cut to the size of the gel, were placed on top of the membrane. The transfer was left overnight at room temperature. Transfer efficiency was determined by staining the gel with ethidium bromide to determine if any RNA remained. The membrane was air-dried and placed on a UV transilluminator for 4 minutes to immobilize the RNA.

The membrane was pre-hybridized overnight at 42°C in hybridization buffer (5x SSC, 5x Denhardt's reagent, 50% formamide, 1% SDS) containing 40 mg/ml denatured herring sperm DNA. An appropriate amount of denatured radioactively labeled probe was added and the hybridization left overnight at 42°C. After hybridization, the membrane was washed twice for 5 minutes at room temperature with low-stringency wash (2x SSC, 0.1% SDS), twice for 5 minutes at room temperature with medium-stringency wash (0.2x SSC, 0.1% SDS), twice for 15 minutes at 42°C in medium-stringency wash and twice for 15 minutes at 68° in high-stringency wash (0.1x SSC, 0.1% SDS). The membrane's radioactivity was checked between each step with a hand-

held Geiger counter. The membrane was rinsed in 2x SSC and placed on film for visualization.

Potato genomic DNA extraction

Frozen tissues from the phosphate-fed potato plants were used as a source of genomic DNA. Genomic DNA was extracted from potato leaves by the CTAB (hexadecyltrimethylammonium bromide) method. One and a half grams of potato leaf was ground thoroughly in liquid nitrogen in a chilled mortar. The powdered tissue was put in a 30 ml polypropylene tube and 15 ml of CTAB buffer (100 mM Tris-HCl pH 7.5. 2M NaCl, 25 mM disodium-EDTA pH 8, 2.5% CTAB, 1.5% PVP 40) added, then vortexed. The tube was then incubated at 65°C for 2 hours, with slow agitation, then cooled to room temperature. An equal volume of 24:1 chloroform/2-octanol was added, and shaken for 5 minutes. The tube was centrifuged for 10 minutes in a JA-20 rotor at 12,000 *g* and the supernatant transferred to a 50 ml tube. Fifteen millilitres of isopropanol was added and the tube was incubated at room temperature for a few minutes, until strands of precipitate were visible. The DNA was recovered by hooking it out of the solution with the flame-bent tip of a Pasteur pipette, and transferred to a microcentrifuge tube. 1 ml of wash solution (70% ethanol, 100mM ammonium acetate, 0.2 µg/ml RNase A) was added and left at room temperature for 10 minutes, then spun down for 10 minutes in a microcentrifuge. The pellet was briefly air-dried and 1 ml of TE buffer added. The DNA was left to dissolve overnight and visually inspected by electrophoresis on a 1% agarose gel.

Results

Subclones

The *pap7* genomic library was constructed from a *Sau*3A1 partial digest, so the insert was separated from the λEMBL3 vector arms by cutting with *Sal*I, which only cut once inside the clone (cutting it into 9 and 9.4 kb fragments). "Shotgun" ligation and cloning into pBluescript yielded five *Eco*RI and two *Eco*RI/*Sal*I subclones from the *pap7* genomic clone. Figure 1 shows representations of the subclones: pap7-6 (5.6 kb), pap7-7 (800 bp), pap7-8 (3.1 kb), pap7-9 (3.3 kb), pap7-500 (464 bp; *Eco*RI/*Sal*I), pap7-10 (6 kb) and pap7-11 (1.1 kb; *Eco*RI/*Sal*I). Pap7-6 is a double clone, with two *Eco*RI fragments, one of which is the same as the insert in pap7-9.

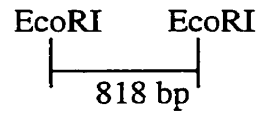
Restriction Mapping

A restriction map of the portion of the genomic clone containing the APase-homologous sequence was constructed. Partial digestion experiments revealed that the 3.3 and 2.3 kb fragments which were adjacent in subclone pap7-6 were in fact adjacent in the genomic clone, and the 1.1 kb *Eco*RI/*Sal*I fragment (subclone pap7-11) was adjacent to the 3.3 kb *Eco*RI fragment. The 1.1 kb fragment was adjacent to the large λ arm, indicating that it contained one end of the genomic clone. Figure 2 shows the arrangement of the three fragments, and the location of the APase-homologous sequence. The assembled sequence of these three fragments is shown in Figure 3.

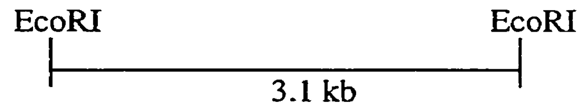
Figure 1. Schematic representations of the *pap7* subclones isolated for sequencing. All subclones are inserted into pBluescript SK+, with the T3 end of the multi-cloning site at the left of the figure and the T7 end at the right. The dashed line above subclone pap7-10 indicates the unsequenced region.

Pap7-6

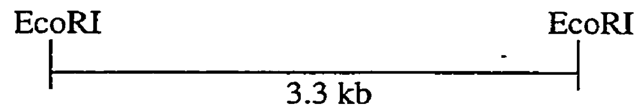
Pap7-7



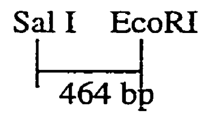
Pap7-8



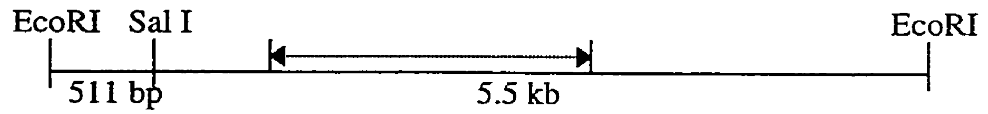
Pap7-9



Pap7-500



Pap7-10



Pap7-11

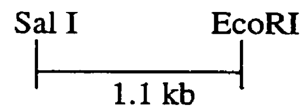


Figure 2. A restriction map of the three contiguous *pap7* clone sequences derived from partial digestion and sequence data. There are numerous *Sau*3A1 sites, only the one separating the genomic clone from the vector is shown. The *Bam*HI sites are 2.8, 4.4, and 5.0 kb from this *Sau*3A1 site, the *Hind*III site is 3.0 kb from the *Sau*3A1 site. The gray bar indicates the portion of the clone that shows homology to acid phosphatase sequences.

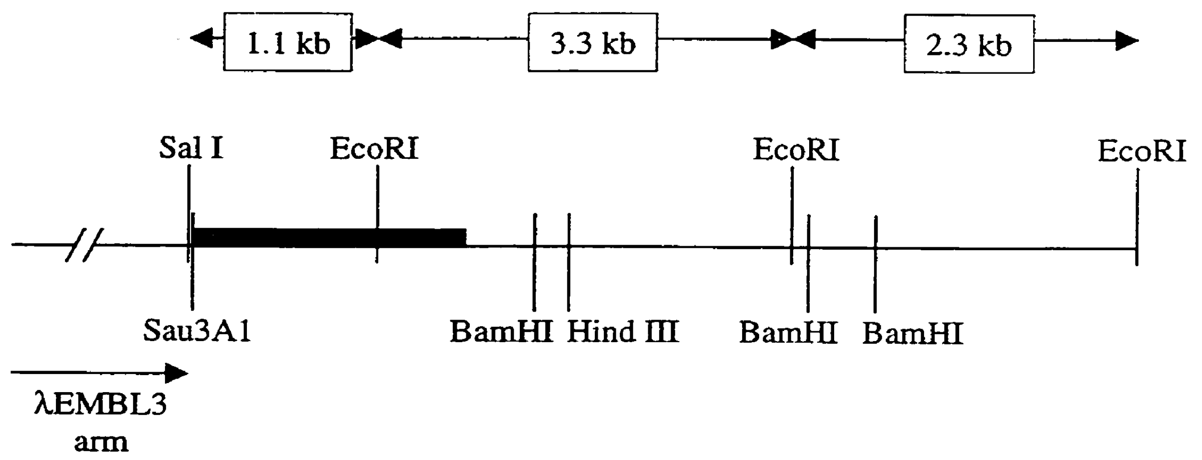

1.1 kb
3.3 kb
2.3 kb
Sal I
EcoRI
EcoRI
EcoRI
Sau3A1
BamHI
Hind III
BamHI
BamHI
λEMBL3
arm

Figure 3. The assembled DNA sequence of subclones pap7-6, pap7-9, and pap7-11, starting at the *Sau*3A1 cloning site. Protein-coding exons (underlined) were predicted using homology with known acid phosphatases and the intron/exon splice-site prediction program SplicePredictor.

GATCATATAGGAAGAAGTGTGATTGTTTCATGGGTAACACCACTAGAAAGCCAATCAAAT 60
TATGTGATCTTTTGGGAAGAAGGGGCCAAGCATCATCACAAACACAAAGCTCATGCTAAA 120
ACCACATCTTATAAATACTATAATTACACTTCTGGTTATATCCACCATGCCACAATCAAA 180
CGACTAAAGGTATATATATATATATATCTTGTGTTTTTTCTAGTTGCTTTTGATTCTTTG 240
AGATACATTGTTTCTACTCTTTTTTTTGGTCAAAATTCAGTATAATGTGACATACATATA 300
TCAACTGGGAGAACATAATTCTACTCGAAGATTCTCCTTCACAACTCCGCCAAAAGTTGG 360
ACCAGATGTTCCTTACACATTTGGCATTATCGGTGAGCACTGAGTACTAAACAAGTTGAT 420
TCATACCAAGGCATGATTGAAGAATTAACTTTAAGTTTAAATTCTATATCCGCCTCTGGA 480
TGCATGTCTGTTACAATGTGATATACTTGGATGCAGGGGATTTGGGGCAAACATTTGACT 540
CAAATCAAACGTTGGAGCAATATTTGTCGAACCCAAAAGGTCAAGCTATGTTGGTTGTTG 600
GTGATTTATCATATGCAGATCATTACCCGTTTCATGACAACGTGAGATGGGATACTTTTG 660
GCCGTTTTATTGAGAAGAGTGCTGCCTTTCAACCTTGGATTACTGCTGCTGGCAATCATG 720
AACTTGATTTTGCTCCTGAGATTGTAATGCTTCCTTTCAATGTATTATTAATAGTTCATA 780
AAACAACAACTCTTTTTCAGTGTGTTTTCGATGTTTTTAGGTGAAAATACTCCATTCAAG 840
CCATACACGCATAGGTATCATGTACCCTACAAGGCATCTCAGAGTACATCTCCTCTTTGG 900
TATTCAATCAAATGTGCATCGACTTATATAATAGTCCTATCCTCTTATTCAGCCTATGGT 960
AAGTACTACATTATAACAACATGTTTTGCTAGTAGTCTCAGTTTGTGTAGTAACTGACCT 1020
TATATAATGTTGTTCAGGTAAATATACTCCACAATATAGTTGGCTTGAACAACAACGTGA 1080
ATTCCCAAAGGTTAACAGAACTGAAACTCCGTGGCTAATTGTTCTGCTTCACTCTCCATG 1140
GTACAACAGTAACAACTATCATTACATGGAAGGTGAAAGCATGAGAGTGATGTTCGAGTC 1200
CTGGTTTGTTCAGAACAAGGTTGACATGGTGTTTGCAGGACATGTTCATTCTTATGAACG 1260
CTCGGTAGGCCCCTCAACTCCATCCTATAAAATTTCATCCTTGAACCAATGTGTATAGGT 1320
CTTGATTAATGGTATATATGTCACATGGACAATTTCGTAGGAACGAGTATCAAATGTTAT 1380
GTACAACATCACAAATGGACAGAGTACTCCAATTGAAGATCCTTCCGCGCCTATATACAT 1440
AACAATTGGGGATGGTGGAAATATTGAAGGCATTGCTAACAAGTTAGACTTCTATTCGTT 1500
CTGATCTCATTTTTATTTAATCTGATGTAATGAGAAGAGTATCTGATGGGAGTCTGTCTG 1560
GTTTTTGTTTGTGCAGTTTTACAGAACCACAGCCGAGCTACTCAGCTTATCGTGAAGCAA 1620
GTTTTGGTCACGCGATTCTTGAAATTAAGAACAGAACTCACGCCTATTATACTTGGCATC 1680
GTAACCAAGACAGTGAACGAGTTGCAGCAGATTCTTTGTGGATTTACAATAGACACTGGT 1740
ATCCTAAAAAGGAAACCAGCTCTATGGCTTGAAAGTAACATGTTTGGTAGGCAGTGCAGT 1800
TGTTATTTTCCTAGAAACCATGTGTTCTATTAGATTAGTTCGTCGGGTACTAAACCAGTT 1860
CCCATTTATGATTTCTTAAAGACCGTGTAAGAGGAAAATGTGACAACTAATATGGAACGA 1920
AGAGAGTATACCTTAACCCTTAGTAGGTGGAAGTTTATTTTCTTATGCCTTCTCCAATCT 1980
TAGTACTTATTTTTTTCTTTCAACTTCTTAATATATTTTTTTTCCGAACTTCAATTTTCA 2040
AATATTCTTTTTTTCTTTTTCATCTTTAACCAAATAAAACTCTCCTCCACTATTAAATAT 2100

AACTGGCAGTGTCACATCAAGTCATTATCTATAACACTAGAAATAGATAAATATAACATC 2160
AATTTTAGGTTTGAAGGTTCATTATCATCAACACTAATTCCATATAAATTCAGTATGTAT 2220
TACAGGTTACTTCTGTCCAATGCCATAAGATGTAATAAATATACATGAAAATGAACTTCA 2280
CTTATTTATGCCTTTAATTAATAGATTCTCTAAATACAAAGTACCTTCCAAAATATTAGC 2340
CAACAGAAAATTAAAATCTTTTGCCCCTCTTTGGCTCGTCTAGTCTTCACTATAAATTGA 2400
CCATTTGTAACTATAATCTTCATCCCCCCTCACTACTCTTTAGCTAAGTGTTCTTTCTTT 2460
TTATTCCTTTGTGTTTTTCTCAATCCTAAGAATTTGTCATTTCTTTATATTATTGGTTTT 2520
TAAGTTACATACACTGATATGAGTTGGGCTTCAAGAAAAGGCGAAAAATTTATATAGCCA 2580
ACCAGGCAACCACAACTCGTTTGTGACTTAGGTGTAGGTTGTTGTTGTTGTTGTTGTTAA 2640
GTTACATACATTAACAACATAAATATTTTTTTATGCTATCGGTATAGTTTATTTGATAAC 2700
GTAAATATTTACAAACAAGATTTGGTTGAGGATTCAAAGTGAGAAATTCTATTGACAGAA 2760
GAGAGAGAGTACAATCTGTCATCAGCAAAATGATATTGTTGTTTGTTGGAATGTACTCTC 2820
TCTGCTTCTGCCAATTGAAACTATAATGGATCCTCAACAAAGATTACACATTTTCCTATT 2880
CATAATGAAAATAGAATATGATGTTAAATGTTTTGTTGCTAGAGATTTTCTGTAAATTAT 2940
ATGAACATCTACTATTTACACTACACAAAAATGTTACAAAGTGTTTAAGAATGTGCCAGA 3000
ATGCTAAATAAGACTAGTGTTGGAAGTTTATTTAGTTTCCAGGCTTCAAGCTTGAGTTAA 3060
TGAGCAAAATGAAGATATAATATATCGTCCAAACGCCATGGTTTGCTCAAAGTCCCATCA 3120
GGTACATAGCTGAGAAGAAACAAGATTAAACGTTCCTCGGAGAAATTACTGAGTATTTCT 3180
GTACTCATTTTTGTTTTTCAAACCACTTCTGTTTGATCCTTCATTAGTTACACTGTACAT 3240
ATTTTTATTCAAAGGAAGAGAGATTTTCATGAACTTGATCTGTTTGCAACATAATTGCTG 3300
TTCTAATATTTCATACAAGTCTTGTACATCTTCTGTAACCAAAAATGATATTGGATGGCA 3360
ATGACATATTTACTCTGTTGAACAATGATGAAATGAAAGTCATTACATTAATATAGAGAT 3420
CAGTATTTCATTCTATTAGATGTTATGAAAAGGGAGGAGGCCTTAAACTCTCCAAATCAC 3480
TGCATGAACACAAGTGAAAAGTCATCTAATTCCGATTTCGAGTTCTGGGTATGCAAAAGT 3540
GTTTGATAGAAAGCGTTTCTCTCGAATGGACTATAAGTGGTGCGAATCAAAAAAAGTCAT 3600
CTAAATTACTAACGAGAAAGTCACTCTTATGTCCAGTATAATCTACCAGATAACACCAAA 3660
CTCAAATTTGAAACATATGTTTCGTGTTTCTTAAATAAAAATTCTTAATCATCAATTTTT 3720
CACAATAACAAATCTAACAAACTCAATTGACACTCTTTAATCTTTTTTCATATATCACTA 3780
TCGCTCAGTTATATATGATTGAAAGTTGAAACACATGTTAAAAATGACTAAAATTTTAGA 3840
TATATGCATTTTAATATTAAAACAAAATTTAAGAAAAGTACTAATTTTTCTTGAAAGCTA 3900
TCTATAGTGGCTTTTCGTTATAAAAACAAACATATAGGGTGTATTAATTAAATGGTAGAA 3960
CAAAAAGGACATACAATGTAATTTTTACAAAATATACCTATTGTGTACGTATTTGTTTGT 4020
TGGAAAAAATTTGGGGGTCTGAGCTTCCAACCCTAATCTCATCTGTTCACTCACTTCACC 4080
GAGACAGTGAGATAGAAAGTTTCGACAGACAAACAGAGTTCTGCACCGGTGCCACCGCCT 4140
CCGGCTGCCTGCGTCCCGTCTGTTTCCGCCTGCTCCCGTCGCGTTTCAAATTGTTTCATT 4200
TGAATAGTAAACAGCTTTTGAAATAAATTTGGAAATTGGATTTAATTAGTGATGCTGCGA 4260
ATGTTGTTATCCAAGCTACCAAATGAACTGATTTAGCCATTTGAAAACCTCAAAGTTGGC 4320
TTACATTTTTGTATTTTTATGTAAAATTAGGAGACCTCTCTACTTATTTGTCAATTGTCA 4380
ATGATGAAGAGGGTATGGAGAATTCATGATGCTTTCCAGAAGGAAGCAATTCTCCAAAAG 4440
TTCTCATCATGTTACACAATCGGACGAGGTTCAGCTAATACGTTGTGGATCCGAGGATTA 4500
GCAGGCAAGGCATCATCTTCTTCCCAACCCGCCGCCTGGCCACATGTATTTACTTTATTT 4560
GCAGATAGGTGGAGTCCTGCCGATTCGTTGGTCAACCAAGATATGAGAGAAAAAGTTTCA 4620
CATTTAAAGGATGAATTGTTAGCATACAGTGGTGATGCTGAAAAGTTTGAGAAAATATTA 4680
GCAGATAAAGGGGATTTGTTGTTTAGTAGGTATGCAGATGGTTCTGCGGTTGTTGAGCTC 4740
TTACAGCAGCTAAAATCTTCGCCCGGACTAGCGCTCCAGGTGATTCTCTTTCTCTATGCA 4800
ACATACGATAGCATCCTCGTTATAACTGAACCGATTTTGTTTTGCAATGATTGAAGTGAA 4860
TGTGATTATTCTTTTGTATGTTAGAAACTATATAAATCTCAATAGCCACCAATCTTTGAA 4920
CAAGAAAGATACCATAAAACAAATATATGAATGGAGTTGTGGTCAAAGAGGACTTGTTGA 4980
TTTTGCAAACAGAGGCGGATCCAGGATTTGAACTCTATGGGTTCAGTCTTTAAGGTTCTT 5040
AGCCTTTATCTGTTATGATTTTGATGTTATGGATTCATATCTAAAGGATGGATCTGCCCC 5100

TGCCTGCAAAGGTCAATTAAGTTGGGACATCATACAGTTATTGCGTGAGTACAACTTTTA 5160
TGATTAGATTATGTAAATTAGGCGGGGATAAGATGGAGCTTAGTTTTTGGATGAAAGAC 5220
GAGTAGTTGTCCTTCTTTTGGTAATATCAAACAGTCCTTTCAATTAATCAATAGTATTTT 5280
GGTGCCTTTCTACTAGCCACACTTGAAGAATTTTCTTGTATGCCTTTACTGTATTACGTA 5340
ATTTCTTCTCTATGCCTTTTACCTAAGAGATGATGTTAAGTCTATGATTACTGCCATTTT 5400
CCCTCATTGAGGAAGTAAACATAGTTTGACTATTTTCTATCAGGCTTTCGACTGGAGAAG 5460
AAGACAATTGGATTACCAGAATCCCATGACAGTTGAGGAGTATTCCAAGGCTATTGTTGT 5520
GGCTGGAAGGTTAAAGAATGTTGATCTTGCAGCCAAACTATTCAAGGAGGCTTCCAACAA 5580
GCAACTCAAGTCTACCTCTTTATATAATGCCCTTATGACTGCATATATGATCAATGGCTT 5640
GGCTGTTAAGTGTCAATCAGTTTTTCGGGACTTGAAGAGGGAAGCAACATGTACTCCAAC 5700
TATTGTAACATACAACATACTGATCTCTGTGTTTGGAAGATTAATGCTGATAGATCACAT 5760
GGAGGCAACTTTACGGGAAATAAATGATTTGGGTATCTGCCCAAATGTTGGTACGTACAA 5820
CTATTTAATTGCTGGGTATATCACAGCGTGGATGTGGGATGATGTGGAGAAGGCATATAG 5880
AATCATGAAGGCAGGAAGTATCAAACCGGATCTTACTACCCATTTATTGATGCTTCGAGG 5940
ATATGCACATTCTGGTAAGTTGGATAATATGGAAGAAATCTATGAGCTTGTCAAAGGTCA 6000
TGTTGATCGACATGGAATCCCATTAATTCGATCTATGATATGTGCTTATAGTAAAAGTTC 6060
TGATGTAAATAAAGTTCAAAAGATTGAGGAATTGATGAGGTTGATTCCTAAAGATGATTA 6120
TAGGCCTTGGTTGAATGTCATATTAATATGTTTGTATGCAAAGGAGGATCTATTAGATGA 6180
GATGGAAAATTCAATAAATGAGGCATTTAAATGTAACACATCCGTCACAACGGTTGGTGT 6240
CATGCATTGCATCATTTCAAGTTACTTCCGAAACAATGCAGTGGACAAACTTGCTAATTT 6300
TGTTAGCCGTGCAGAATGTGCTGGTTGGAAAATATGCAGGTCCCTTTACCACTGCAAGAT 6360
GGTTATGTATTCATCACAAAAGCGACTCATTGAAATGGAGAAAGTTCTTAGCGAGATGGA 6420
TAAAGTGAATTTGGACTTTTCAAAGAAAACACTTTGGATATTACTTAAAGCCTACGGAAC 6480
TTGGGGGGAAAAGGATAAACTTCATCAGGTCTTGGGTATGATGTGCAGGCATGGATATGG 6540
AATTC 6545

Sequencing and homology searches

All of the subclones were fully sequenced except pap7-10, and these sequences are shown in Figure 3 (7-6, 7-9 and 7-11 assembled) and appendices A-D (7-7, 7-8, 7-10 and 7-500). When the subclone sequences were analyzed by BLAST search (using BLASTN to compare nucleotide sequences, and BLASTX to compare all six reading frames to known protein sequences), only three subclones, pap7-6, pap7-9 and pap7-11, comprising three separate restriction fragments, showed homology to other plant APase genes (for details see "Sequence Analysis"). BLASTX analysis of subclone pap7-10 returned homology to plant transposable elements at its ends: approximately 350 bp at the T3 end and approximately 700 bp at the T7 end are homologous to *copia*-like and *gypsy*-like polyproteins respectively. The T3 sequence (the reverse complement of bases 153-494) displayed maximal protein identity (40%) with the polyprotein from *Phaseolus vulgaris* transposable element Tpv2-1c, a *copia*-type retrotransposon (accession # CAB42059). The T7 sequence (the reverse complement of bases 221-934) displayed maximal protein identity with a pineapple polyprotein-like sequence (40%; accession # T07863) and a *gypsy*-type polyprotein from *Sorghum bicolor* (39%; accession # AAD22153) The 2.3 kb *Eco*RI fragment in subclone pap7-6 displayed protein-level homology with a number of hypothetical *Arabidopsis* proteins of undetermined function (e.g. accession # AAC02749), as well as *Arabidopsis* clones similar to membrane-associated salt-inducible genes (23%; accession # BAB08985) and glutamine synthetase (22%; accession # CAB64220).

Sequence Analysis

Once the DNA sequence was mapped and assembled, the *pap7* deduced amino acid sequence showed the strongest homology (BLASTX) to an *Arabidopsis thaliana* purple acid phosphatase homologue (accession # T04599), kidney bean PAP (KBPAP, accession # P80366), and a white lupin (*Lupinus albus*) secreted acid phosphatase (accession # BAA97745). The sequence was analyzed to determine the extent of the putative *pap7* coding sequence and find intron/exon boundaries and other features. Exons were predicted by starting with the open reading frames showing APase homology, and adjusting the intron/exon boundaries to optimize both amino acid homology to potato tuber acid phosphatase (PTAP) and consensus mRNA splice site sequences (GT at the 5' end of the intron and AG at the 3' end; Egoavil *et al.*, 1997). The SplicePredictor intron/exon boundary prediction program was used to assist in this process, and although it predicted a number of unlikely splice sites it also predicted ones which flanked the highly-homologous ORF's. A consensus sequence-conforming site cound not be found for one acceptor splice site (at the 5' end of exon #4), so its location was assigned solely on the basis of ORF translation homology. Seven putative exons were mapped, coding for 404 amino acids (Fig. 4). Because the 5' end of the sequence is truncated by the boundary of the genomic clone, the predicted protein sequence is incomplete. Based on alignment with PTAP and purple acid phosphatases, PAP7 could contain between 33 and 71 additional residues at the amino terminus.

The predicted PAP7 sequence shows 53% identity with PTAP, 63% identity with KBPAP, and 61% identity with sweet potato PAP (expressed as percentages of the

Figure 4. Predicted amino acid sequence of PAP7.

HIGRSVIVSWVTPLESQSNYVIFWEEGAKHHHKHKAHAKT 40
TSYKYYNYTSGYIHHATIKRLKYNVTYIYQLGEHNSTRRF 80
SFTTPPKVGPDVPYTFGIIGDLGQTFDSNQTLEQYLSNPK 120
GQAMLVVGDLSYADHYPFHDNVRWDTFGRFIEKSAAFQPW 160
ITAAGNHELDFAPEIFLGENTPFKPYTHRYHVPYKASQST 200
SPLWYSIKCASTYIIVLSSYSAYGKYTPQYSWLEQQREFP 240
KVNRTETPWLIVLLHSPWYNSNNYHYMEGESMRVMFESWF 280
VQNKVDMVFAGHVHSYERSERVSNVMYNITNGQSTPIEDP 320
SAPIYITIGDGGNIEGIANNFTEPQPSYSAYREASFGHAI 360
LEIKNRTHAYYTWHRNQDSERVAADSLWIYNRHWYPKKET 400
SSMA 404

known Pap7 sequence). Figure 5 shows a CLUSTAL (Higgins and Sharp, 1988) alignment of the deduced PAP7 sequence with PTAP, KBPAP, spPAP, *Arabidopsis* PAP homologue, and lupin secreted APase. In the well-characterized kidney bean enzyme, the residues involved in metal binding at the active site have been determined by X-ray crystallography. Fe(III) is coordinated with Asp 135, Tyr 167, and His 325, while Zn(II) is coordinated with Asn 201, His 286, and His 323, and Asp 164 "bridges" the two metal ions (Klabunde *et al.*, 1995). These residues are all completely conserved among the APases in Figure 5, including PAP7.

Figure 6 shows a CLUSTAL phylogenetic tree of the APases aligned in Figure 5. as well as *Aspergillus niger* and *Aspergillus nidulans* pH 6.0 optimum APases (accession #'s U18553 and Z79750). The cladogram indicates that kidney bean, lupin, and sweet potato PAP's make up one subgroup with PTAP while pap7 and *Arabidopsis* PAP compose another. The *A. niger* and *A. nidulans* APases, which share the active-site residues with PAP's (Mullaney and Ullah, 1998), nevertheless are highly divergent from the plant PAP's.

Because PTAP showed strong tyrosine phosphatase activity, the PTAP and PAP7 amino acid sequences were compared with those of three plant protein tyrosine phosphatases, from pea (accession #AJ005589), soybean (accession #AJ006308), and *Arabidopsis* (accession #AF055635). The three PTPases showed strong homology with each other, however PTAP and PAP7 showed no homology to them (Fig. 7).

Figure 5. CLUSTAL alignment of the PAP7 protein sequence with PTAP, kidney bean PAP (accession # P80366), sweet potato PAP (AF200825), *Arabidopsis thaliana* PAP-homologue (T04599), and white lupin secreted APase precursor (BAA97745). Shaded residues indicate identity with PAP7. The residues marked with a dot above indicate the amino acids involved in active site metal ion-binding in KBPAP (Asp 135, Asp 164, Tyr 167, Asn 201, His 286, His 323, and His 325). The amino-terminus of the PAP7 sequence is incomplete due to the 5' truncation of the genomic clone.

```
pap7 protein- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  0
PTAP         T Q F P - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  4
KBPAP        - - - - - - - - - - - - - - - - - - - - - - - - - - - - F V R K - T N - - -  6
spPAP        M R L V V V G L W C L I L G L I L N P T K F C D A G V T S S Y V R K S L S A L P 40
AtPAP        M R M N K I - - - - L L V F V F L S I A T V I N S G T T S N F V R T A Q P S - -  34
Lupin APase  M K M G Y S S F - - V A I A L L M S V V V V C N G G K T S T Y V R N L I E - - -  35

pap7 protein- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  8
PTAP         - - - S V D I P L E N E V L S V P N G Y N A P Q Q V  T Q G D Y D  E A  I 41
KBPAP        - - K N R D M P L D S D V F R V P P G Y N A P Q Q V  T Q G D L V  A M  I 44
spPAP        N A E D V D M P W D S D V F A V P S G Y N A P Q Q V  T Q G D Y E  G  I 80
AtPAP        - - - - T E M S L E T - - F P S P A G H N A P E Q V  V Q G D Y N  G I  I 68
Lupin APase  - - K P V D M P L D S D A F A I P P G Y N A P Q Q V  T Q G D L V  Q A M  I 73

             10          20          30          40
pap7 protein  -  - -  45
PTAP          A - D E P G  S E  R Y G L - - - - S E G K Y D V T V E G T L N N  T F 76
KBPAP         M - D E P G  S A  R Y  S - - - - E  N G R  R I  K G  M S T  R F 79
spPAP         T  Y D K A G A  K  V Y  S - - - - E N S K S Q K R  M G T V V T  116
AtPAP         N L A G  V  T Y  K A V D  D V K P K  K R G  S  S  R F 108
Lupin APase   V - D E P G  Q  Y  S - - - - D S S L Q N F T  E G E V F T  T 108

             50          60          70          80
pap7 protein  85
PTAP          K  E  Q C L V T G  Q  D T K  Y  E I  K G D  A  K  W  E  116
KBPAP         F  S  F  T  R K  Y K  Y  E V  L R  T  I  119
spPAP         A F  C  D  E  D T K  Y  R  F G D A K  Q  W  V  156
AtPAP         D  F L  G  E  D T K  E V  T D G  V  Q  S  148
Lupin APase   F  T  T N  E F D T  Y  E V  I G  T  Q  W  I  148

             90          100         110         120
pap7 protein  -  124
PTAP          D  A S  K  Y N  L S  Q H  M A S G A  S - - V 154
KBPAP         Q T  L  L  S  T  S H  E L S  K  T V 159
spPAP         P  V  L  I  H  T  T H  E Q  S A  V 196
AtPAP         L A  E  Y H  M  -  V 187
Lupin APase   E  L  T  T H  Q N S - - N  T  L 186

             130         140         150         160
pap7 protein  -  163
PTAP          F  R  Q Y N  V G  L V  Q  T  Y  W S 194
KBPAP         F  R  N  -  W  T  R  V  Y  W T 198
spPAP         F M  S N R W  N  -  N  W  S  R  V  Y  W T 235
AtPAP         F P  D H  N  - Q R K  S W  V  P C  Y  T F  Y 226
Lupin APase   Y  D  Y  -  W  T  R  Y  W T 225
```

170 180 190 200

pap7 protein 203
PTAP I E Y F S M - - V V R S F L S P T R K S N 232
KBPAP I E - - N T E F S Y E F 236
spPAP I Y D - - Y Q V F N P T H E G G D 273
AtPAP I V N - - P H A I N A K I 264
Lupin APase I D L Q - - T Q F S T T E E F 263

210 220 230 240

pap7 protein 243
PTAP A R R A H N P F W H K - - K 270
KBPAP R A H H I R G T K K - - L R K 274
spPAP A R A H G F V S K F T S - - L E 311
AtPAP R R A H V - - L K 302
Lupin APase Y R G P A H V A T F Y S L K T A - - L 301

250 260 270 280

pap7 protein 283
PTAP E K M V I A A F S A Y R K Y 310
KBPAP S M L Y H F A T K A K Y 314
spPAP S V A L Y E A A A I P Y Y Y 351
AtPAP E M V A N S 342
Lupin APase S S M A Y N P I Y L L K Y 341

290 300 310 320

pap7 protein 323
PTAP V I A Y I I H V S G D A Y V P K A 350
KBPAP V A I A K D L C V K Q 354
spPAP I S A V A K C V S E 391
AtPAP L L S I K L Y V K 382
Lupin APase V A N K I C V K I T 381

330 340 350 360

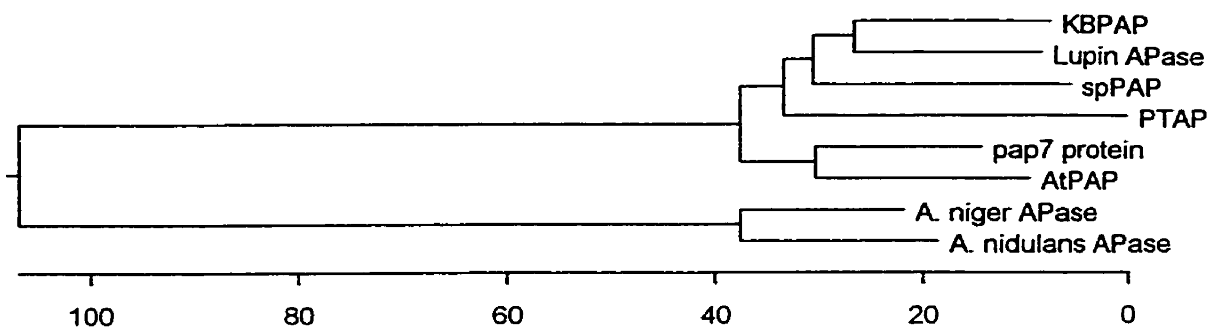
pap7 protein 363
PTAP V S L T S R R D E F Y 390
KBPAP V A Y G V D S M I Q E F G M F D 394
spPAP V S L S E M Q F G F D 431
AtPAP S D V 422
Lupin APase N L L - A T M K Q G F A 420

370 380 390 400

pap7 protein - - - 400
PTAP S N D G N A I T T F T L H Q 421
KBPAP H F S N G V A E V F F V D D - - S 432
spPAP H F S N G A S E L L Y A S E D A - - S S 469
AtPAP Y N P I M L H Y F F V E L E S G 462
Lupin APase H S N G Y A E K L F Y N L N D - - S 458

pap7 protein	[illegible] A	404
PTAP		421
KBPAP		432
spPAP	M [illegible] A M	473
AtPAP	N T R [illegible]	466
Lupin APase	I H I P	462

Figure 6. CLUSTAL cladogram of the PAP7 protein sequence, PTAP, kidney bean PAP (accession # P80366), sweet potato PAP (AF200825), *Arabidopsis thaliana* PAP-homologue (T04599), white lupin secreted APase precursor (BAA97745), *Aspergillus niger* APase (U18553) and *Aspergillus nidulans* APase (Z79750). The scale at the bottom indicates divergence as the number of residue substitutions between pairs.



KBPAP
Lupin APase
spPAP
PTAP
pap7 protein
AtPAP
A. niger APase
A. nidulans APase
100
80
60
40
20
0

Figure 7. CLUSTAL alignment of the PTAP and PAP7 protein sequences with pea (accession #AJ005589), soybean (accession #AJ006308), and *Arabidopsis* (accession #AF055635) protein tyrosine phosphatases. Shaded residues indicate identity with the consensus sequence.

PAP7 protein ------------------------------ ------GRS 7
PTAP TQFPSVDIPLENEVL VPNGYNA QQV QG YDGEA 40
P.sativum PTPase G---------SS PP TSSHNNS--- P F IP 28
G.max PTPase ---------NP TT SALS EK NFN P N IT 31
A.thaliana PTPas ---------T KTS AANLFTGSTR DLSSA SP KLS 32

10 20 30 40
PAP7 protein V WVTPLESQ NY FWEEGA H HK KAHAKTTSYKY N 47
PTAP I WVTADEPG SE RYGLSE YDVT-- EGTLNNYT K 78
P.sativum PTPase T IK E AL ---N L T S K----- F 60
G.max PTPase T V Q N ---E A NVIT ----- A 63
A.thaliana PTPas L HQ GVFR--- IQ DSIAH ----- TG 64

50 60 70 80
PAP7 protein YT GYIHHA IKR K V YI QL EHNST RFSFTTPPK 87
PTAP YE GYIHQCLVTG Q DTKYY EI KGDSA KFWFETPPK 118
P.sativum PTPase --- L TTKK H L A S --------- 88
G.max PTPase A--- MRRR DDV R --------- 91
A.thaliana PTPas A--- MW LL S M V VE --------- 92

90 100 110 120
PAP7 protein VGPDVPYTF II DLGQTFDSNQT EQ L NP QAMLVV 127
PTAP VDPDASYKF II DLGQTYNS ST QH MA G-AKSVLFV 157
P.sativum PTPase -------- S I SS E L ------ 114
G.max PTPase -------- T L ST PE Q ------ 117
A.thaliana PTPase -------- S V I NPCK -- ------ 116

130 140 150 160
PAP7 protein GDLS DHYPFHD-N RWD FGR EKSAAF PWITAAGN 166
PTAP GDLS DRYQYNDVG RWD FGRLVEQSTAY PWIWSAGN 197
P.sativum PTPase ---- I K P-- I E MP-------- 140
G.max PTPase ---- I V A-- N -------- 143
A.thaliana PTPase ---- V K ES--ESI P-------- 142

170 180 190 200
PAP7 protein ELDFAP IFLG NT FKPYTH YHVP SQSTSPLWYS 206
PTAP EI Y PS--MG VV FRSF S YPTP R S SSNPLWYA 235
P.sativum PTPase F M Y G T ----- 175
G.max PTPase Y I Y I MA ----- 178
A.thaliana PTPas M A V Q I NRT ----- 177

210 220 230 240
PAP7 protein IKCAST II LSSYSAY KYTPQYS LEQQREFPK R 246
PTAP IRRASAHII LSNYSP KYTPQWH LKQ--EFKK REK 273
P.sativum PTPase ------ S C ---------TK K 200
G.max PTPase ------ A V IIG ----------E 203
A.thaliana PTPase ------ D G T ----------IK D 202

```
                        250        260        270        280
PAP7 protein         PWLIV  HSPWYNSNNYHYMEGESMRVMFES FVQNKV   286
PTAP                 PWLIV MH PIYNSNA HFMEGESMRSA ER FVKYKV   313
P.sativum PTPase          H    R      T    ---    F       N   237
G.max PTPase              H    HR          ---            K   240
A.thaliana PTPase     M   N    Y  T   Q  M---               K   239

                        290        300        310        320
PAP7 protein         MV  GH ---HS RSERVSNVMYNITN QSTPIEDPSAP  323
PTAP                 VI  GH ---HA RSYRISN HYNV G DAYPVPDK AP  350
P.sativum PTPase     N L   A W    H                      T     277
G.max PTPase               E      H     F                T     280
A.thaliana PTPase      V   E      QV  S    I                   279

                        330        340        350        360
PAP7 protein         IY TIGDG NIEGIANNFTEPQPSY  Y EASF HAILEI  363
PTAP                 IY TVGDG NSEG TSRFR PQPEY    EASY HAILEI  390
P.sativum PTPase                   I-----   N    V        --     310
G.max PTPase               V        V-----      IAM       --     313
A.thaliana PTPase                   -----  L     AL  K    --     312

                        370        380        390        400
PAP7 protein         KNRTHAY TWHRNQDSER  A  LWIY R  YPKKETSSM  403
PTAP                 KNRTHAY SWNRNDDGNAITT  FTLH Q             421
P.sativum PTPase            E          S Q                     332
G.max PTPase                           S Q-------------Q---    337
A.thaliana PTPase    M  F       V      T GT--------------NAGT  339

PAP7 protein         A                                          404
PTAP                                                            421
P.sativum PTPase                                                332
G.max PTPase         E                                          338
A.thaliana PTPase    S                                          340
```

Genomic Copy Number

A Southern blot of potato genomic DNA was probed with pap7p1 (corresponding to the 3' end of the last *pap7* exon) to determine the number of copies of *pap7* in the genome. As seen in Figure 8, genomic DNA digested with *Eco*RI and *Xba*I each produced a single hybridizing band with sizes of 7 kb (*Eco*RI) and approximately 15 kb (*Xba*I). These results indicate there is a single copy of the *pap7* sequence in the potato genome. Although the *Eco*RI fragment size (7 kb) is not the same as the size expected from the genomic clone sequence (3.3 kb), the genomic DNA used was not from the same cultivar as the library clone.

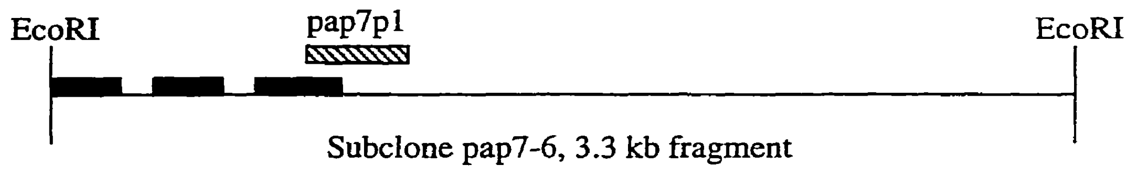
Expression of *pap7*

Northern blotting was used to determine if *pap7* is a gene expressed in phosphate-fed or starved potato plants. In order to confirm starved plants were experiencing Pi-deprivation, acid phosphatase activity was assayed. After 31 days of growth in zero-phosphate medium, assays showed significantly more acid phosphatase activity in the leaves of P_i-starved plants than in P_i-fed plants (one-tailed t-test, t=2.335, DF=4). Phosphatase activity in the Pi-starved plants was 164.0 ± 24.6 U/mg total protein while in the Pi-fed plants it was 118.1 ± 13.1 U/mg (means of three replicates each; U= µmol *p*-nitrophenol produced/min).

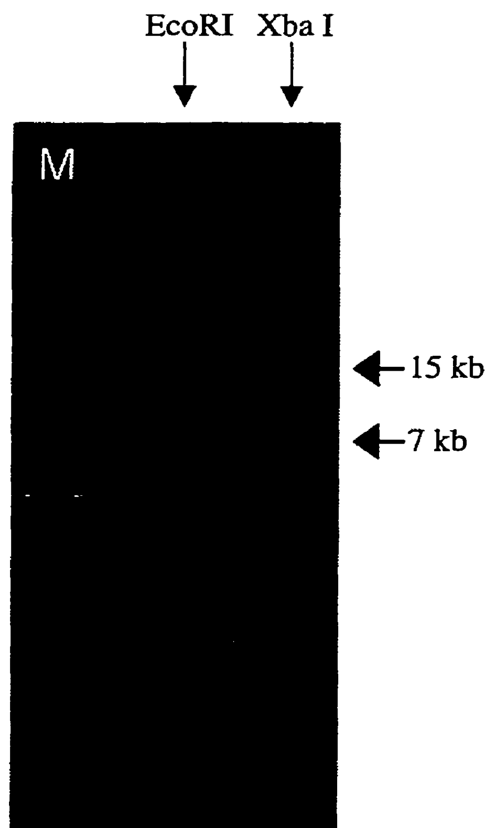
Total RNA from the leaves, stems, and roots of phosphate-fed and phosphate-starved potato plants was isolated and probed with a probe made from the 3.3 kb *Eco*RI fragment of *pap7* containing the last three predicted exons (subclone pap7-9). No signal was detected in any of the fed or starved tissues even after the autoradiogram was

Figure 8. a) The pap7p1 probe corresponds to the last predicted exon of *pap7* and a portion of the 3' region. The gray bars indicate the last three predicted exons of *pap7* and the hatched bar represents the probe. b) Southern blot of restriction-digested potato genomic DNA probed with pap7p1. The probe bound to a 7 kb *Eco*RI fragment and a 15 kb *Xba*I fragment. The smaller band is probe binding to the DNA size marker (lane marked "M").

a)



b)



exposed for 14 days (Fig. 9). Probing the same blot with a potato actin probe showed the presence of mRNA in all lanes in approximately equal amounts, except for the phosphate-fed stem, which contained a weaker but still distinct signal. In order to determine if the phosphate-starved tissues were expressing another, known phosphate-starvation induced gene, the blot was also probed with a probe made from the sequence of StPT2, a potato cDNA encoding a PSI root phosphate transporter gene (Leggewie *et al.*, 1997). The results appeared to show a correctly-sized (approximately 1.8 kb) band in the phosphate-starved roots, but as in the original StPT2 study, there was a large amount of nonspecific hybridization to the ribosomal RNA bands, making identification of the band difficult. When the blot was washed at increased stringency, the StPT2 band was reduced in intensity along with the nonspecific binding.

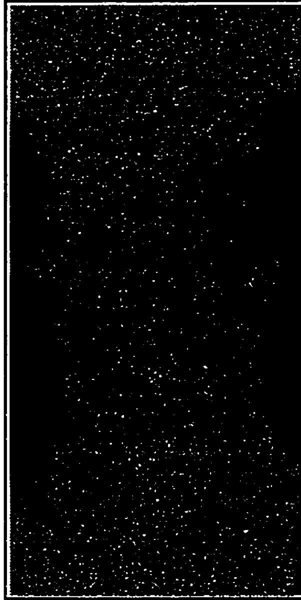
Figure 9. Northern blot to determine if *pap7* is expressed in phosphate starved potato plants. Lane identities: L = leaf, S = stem, R = root; +P = phosphate-fed (0.13 mM P_i), -P = phosphate starved (no P_i). a) The blot was probed with a potato actin probe to confirm RNA transfer and loading amounts. b) The blot probed with a *pap7* probe. The dark signal at the sides of the blot is from nonspecific hybridization with the size marker. c) The blot probed with a probe made from the sequence of StPT2, a phosphate-starvation induced root phosphate transporter gene from potato (Leggewie *et al.*, 1997). The arrow indicates the expected position of the 1.8 kb StPT2 transcript (size determined by comparison with the DNA size marker).

L S R

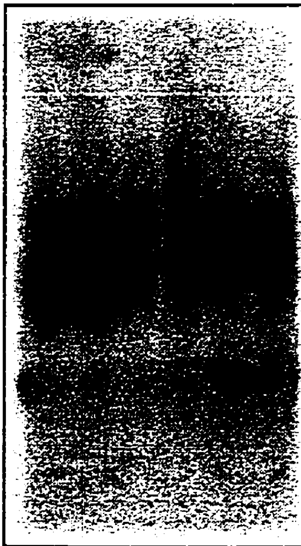
+P -P +P -P +P -P

a)

b)



c)



1.8 kb

Discussion

Previous work in our laboratory on potato acid phosphatases resulted in the cloning of the gene coding for the major tuber APase (PTAP) and the isolation of several potato genomic clones (*pap3*, *7*, and *11*) and a rice cDNA clone homologous to it (Gellatly, 1996). This was one of the most thorough studies of acid phosphatase genes in any species, and suggested that there exists a distinct family of plant acid phosphatases sharing sequence homology. The PTAP enzyme itself is a 100 kDa homodimer, showing 1917 U/mg activity towards phosphotyrosine and able to dephosphorylate phosphotyrosylated proteins. It cannot be classified as specific, however, as it is also able to hydrolyze other phosphoesters. PTAP shows 52% amino acid sequence identity with kidney bean purple acid phosphatase, especially at the C-terminus, and much of it in blocks of 5 to 12 completely conserved residues (Gellatly, 1996).

Coding sequence prediction

Prediction of a *pap7* protein sequence required locating exon-candidate ORF's which had functional exon characteristics as well as homology to other APases when translated to an amino acid sequence. Candidates were initially identified by examination of the report returned from a BLASTX search, showing ORF's highly homologous to other APases. This provided a crude estimation of the boundaries of any potential exons. The first 2500 bp of the assembled genomic sequence, containing all of the APase-homologous sequence and a substantial 3' region, was then analyzed by the SplicePredictor software. The program returned 14 potential donor sites (5' end of

intron) and 11 potential acceptor sites (3' end of intron). The donor and acceptor sites which corresponded to the ends of homologous sequences had total scores ranging from 13 ("highly likely") down to 4 ("doubtful"). The score represents the cumulative assessment of three factors: *P*, representing the likely *in vivo* efficiency of the site, *rho*, which factors in the *P* value of a possible complementary splice site, and *gamma*, which compares the site's *P* value to those of nearby sites of the same type (Brendel and Kleffe, 1998). One predicted AG acceptor site at the 5' end of exon #4 was rejected in favour of a nearby non-consensus TG because it would have resulted in a frameshift in the following exons, destroying the APase homology. It is possible that this problem is due to a sequencing error or a mutation introduced during cloning; alternately, *pap7* may be a non-expressed pseudogene.

The *pap7* genomic clone is approximately 18 kb in size. Of this, approximately 6.6 kb of contiguous sequence, comprising three subcloned restriction fragments, was sequenced and mapped. This sequence, adjacent to the EMBL3 vector large arm, includes seven open reading frames homologous to PTAP and purple acid phosphatases. The first predicted exon starts almost immediately at the *Sau*3AI cloning site, suggesting that it was truncated by cloning. The predicted protein sequence (Fig. 4) is almost certainly incomplete. At 404 amino acids, it is distinctly shorter than potato tuber APase (451 a.a.), kidney bean PAP (432 a.a.), sweet potato PAP (466 a.a.), and white lupin APase (462 a.a.). This suggests that anywhere from 84 to 186 bp of coding sequence is missing from the 5' end of the *pap7* clone, as well as a non-predictable amount of intron sequence. Of course, the 5' non-coding region is missing as well, preventing any

determination of whether *pap7* possesses a functional transcription start site, transcription complex binding sites, or recognizable promoter elements.

Polyadenylation site prediction

Assembly of the subcloned restriction fragments of *pap7* into 6.6 kb of uninterrupted genomic sequence, with the predicted coding sequence at one end, allows examination of the 3' region. The 3' regions of plant genes (and mRNA's) contain several elements, relating to endonucleolytic trimming of the transcript and subsequent polyadenylation. Polyadenylation sites in plant genes are preceded by near upstream elements (NUE's) and far upstream elements (FUE's). Unlike animals, plant mRNA's do not appear to require an AAUAAA consensus sequence in the NUE, although it does appear in many plant mRNA 3' regions (Rothnie, 1996). The pea *rbc*S-E9 gene, for example, has a functional NUE with the sequence AAAUGGAAA (Li and Hunt, 1995). The efficiency of recognition of this signal did improve when the AAUGGA was replaced with AAUAAA, but other, single base, changes only slightly affected the efficiency. Three additional NUE's were present, each controlling a separate poly-A site, but these additional NUE's were not characterizable except as U-rich regions. The maize 27 kDa *zein* gene has two NUE's, both AAUGAA, the downstream of which controls polyadenylation *in vivo* (Rothnie, 1996). FUE's are harder to analyze. The FUE of the pea *rbc*S-E9 gene is a 60 bp U-rich region ending 10 bp upstream of the first NUE, while the *zein* FUE is an 81 bp region (not U-rich) about 30 bp upstream from the NUE (Rothnie, 1996). The *pap7* sequence contains two AAUAAA motifs near the 3' end of the coding sequence, 300 bp and 481 bp downstream from the stop codon. There is also

an AAUGAA motif 497 bp downstream from the stop codon. Of these, the most likely candidate for a functional NUE is the one 300 bp downstream. It has a 120 bp region directly upstream of it which is T-rich (56%), as well as a number of (T/C)A doublets immediately downstream which could be candidates for mRNA cleavage/poly-A sites. However, this identification is not definitive, as polyadenylation sites and control sequences are difficult to identify from genomic sequence alone. In fact, comparative studies of yeast, plant, and animal suggest that multiple control sequences are involved in poly-A site determination, and that none of them are required to have any exact consensus sequence (Graber *et al.*, 1999).

Pap7: comparison with other APases

A number of plant acid phosphatase genes have recently been cloned and characterized. These genes fall into several categories of nonmetallic APases (or tartrate-inhibited), and purple acid phosphatases (metalloproteins, and tartrate-resistant). PAP7 and PTAP appear to fall within the purple acid phosphatases, based on their derived protein sequences.

Based on amino acid sequence, plant APases comprise three distinct groups: PhyA-like, bacterial class C-like, and purple acid phosphatases. PhyA is a phytase (myo-inositol hexakisphosphatase) from the fungus *Aspergillus niger*. It has dual pH optima at 2.5 and 5.0, and hydrolyses other phosphoesters besides phytate (Mullaney and Ullah, 1998). Another similar phytase, PhyB, and a nonspecific APase (Ullah and Dischinger, 1993) from *A. niger* have also been sequenced, and all share a common RHGxRxP active site motif. This motif also appears in the derived protein sequence of an *Arabidopsis*

gene which otherwise shares only 15% homology with PhyA (Mullaney and Ullah, 1988). The first plant phytase to be cloned was Phy S11 from maize, and although it does not maintain complete conservation of the RHGxRxP motif, it has a RHQLRLG motif which retains the Arg and His residues essential for PhyA phosphatase activity (Maugenest *et al.*, 1997).

Bacterial nonspecific acid phosphatases can be divided into three types based on sequence and characteristics. Class A APases are mono- or multimeric proteins with 25-27 kDa subunits, while Class B APases are homotetramers of 25 kDa subunits, with a metal cofactor requirement (Thaller *et al.*, 1998). Class C APases are membrane lipoproteins, and show limited homology to the other two classes. However, they do have two conserved sequence motifs with class B enzymes and some plant APases. These motifs, xDxDxL and GDxxxD, are found in *Lysopersicon esculentum* APS1, a soybean root APase, and an Arabidopsis vegetative storage protein (Thaller *et al.*, 1998). The same motifs are present in two soybean vegetative storage proteins (Vspα and Vspβ) which are homologous to *Arabidopsis* Vsp and which display acid phosphatase activity (DeWald *et al.*, 1992). In particular, the four Asp residues are also conserved across phosphatases from a wide variety of microbial species, suggesting they may be critical for these enzymes' phosphohydrolysing activity (Thaller *et al.*, 1998).

Purple acid phosphatases (PAP's) are unique among these groups of APases in that they require two metal cations bound within the active site for phosphatase activity. X-ray crystallography studies of kidney bean purple acid phosphatase showed that the cations are bound by specific amino acids: Fe(III) by Asp 135, Tyr 167, and His 325, and Zn(II) by Asn 201, His 286 and His 323, with Asp 164 binding both (Klabunde *et al.*,

1995). In mammalian PAP's the second ion is Fe(II) (Klabunde *et al.*, 1995), while in sweet potato PAP it is either Mn(II) (Schenk *et al.*, 1999) or Zn(II) (Durmus *et al.*, 1999). White lupin APase has not been analyzed for metal content, but when purified it displays the distinctive purple colour caused by the Fe(III)-Tyr charge transfer (Ozawa *et al.*, 1995). Although the overall sequence homology between mammalian PAP's and plant PAP's is poor, these metal-binding residues (and their relative positions within the polypeptide) are completely conserved (Klabunde *et al.*, 1995; see also Fig. 5). Both PAP7 and PTAP contain these conserved residues, as well as having a high degree of overall homology with plant PAP's.

The metals in PAP active sites appear to be integral to their reaction mechanism. In general, APases hydrolyze phosphoesters through the formation of an acid-labile phosphorylated amino acid intermediate (Vincent *et al.*, 1992). Studies of the crystal structure of kidney bean PAP have suggested a more detailed model of PAP-catalyzed hydrolysis (Klabunde *et al.*, 1996). Two oxygens in the phosphate group of the substrate hydrogen-bond to His 202 and His 296, and another coordinates with the Zn(II). A hydroxide ion coordinated with the Fe(III) then attacks the phosphate, producing a transition state where 5 oxygens are bound to the phosphorus atom (including the oxygen in the phosphoester bond). Then the phosphoester oxygen becomes bound to His 296, and finally electrons are transferred from the Fe(III) to its coordinated oxygen, and from the phosphorus atom to the phosphoester oxygen, which is protonated by His 296 and leaves with the dephosphorylated product, leaving the phosphate coordinated with the Zn(II) and His residues. This model is supported by stereochemical data and enzyme kinetics as well as crystallographic data on both the free and substrate-bound enzyme

(Klabunde *et al.*, 1996). The Fe(III) ion is especially important in this model, as it provides the strongly nucleophilic hydroxide which attacks the phosphate, releasing the dephosphorylated product (Vincent *et al.*, 1992).

In contrast to kidney bean PAP, PTAP is not known to be a metalloenzyme. In fact, purified PTAP does not display the characteristic purple colour, nor was it found to contain detectable amounts of iron or zinc, although some magnesium and calcium were detected (Gellatly, 1996). It seems unlikely that iron or zinc could have disassociated from the enzyme during the 3000-fold dialysis prior to the metals analysis, while other metals remained. Since the enzyme was active prior to dialysis, and other metals were found in the sample, it is possible that PTAP does not use the same metallic active site structure as other characterized PAP's. In addition, PTAP activity is eliminated by added Zn^{2+} (Gellatly, 1996), making it an unlikely active site component. Hence, it is possible that the *ptap* and *pap7* gene products are not purple acid phosphatases despite their strong sequence homology to plant PAP's. If this is true, it would pose the question of how the conserved residues are utilized, if not to bind the catalytically important metal ions.

Molecular modeling of the Cdc25 protein tyrosine phosphatase (Eckstein *et al.*, 1996) suggests that its catalysis proceeds via nucleophilic attack on the phosphate by cysteine, resulting in a temporary thiophosphate ester bond which is then hydrolyzed to release the phosphate. The dephosphorylated substrate is protonated by an aspartic acid residue as it is replaced by the bond to cysteine. Although PTAP is highly active against phosphotyrosine, it and PAP7 share little sequence homology with protein tyrosine phosphatases (Fig. 7). Significantly, the HCxxxxxR active-site motif shared by tyrosine

phosphatases (Eckstein *et al.*, 1996) is present in the PTPases in Fig. 7 but not in PTAP or PAP7. PTAP and PAP7 each have 1 cysteine residue but they are not in conserved positions. It is unlikely that these enzymes could share a reaction mechanism with PTPases. As PTAP is activated by Mg^{2+}, Mn^{2+}, Co^{2+}, and Ca^{2+} (Gellatly, 1996), it could be speculated that the enzyme utilizes a divalent cation to generate the nucleophilic hydroxide, while the leaving group is protonated by a histidine residue as in the proposed KBPAP mechanism.

The predicted *pap7* protein shows a higher homology with KBPAP (63%) than with PTAP (53%). If the potato and kidney bean enzymes represent two related groups of enzymes with subtly different mechanisms, this limited data would suggest that the PAP7 protein, if it exists *in vivo*, might be more "KBPAP-like" than "PTAP-like". Of course, the difference in homology could have no functional significance. The purification and characterization of the PAP7 protein from potato tissues is obviously a prerequisite to any serious analysis of its functional homology to other APases.

Non-APase Homologies

The 3' end of the contiguous mapped region of the genomic clone (corresponding to the 2.3 kb *Eco*RI fragment of subclone pap7-6) is homologous to an *Arabidopsis* clone coding for a glutamine synthetase-like protein. Glutamine synthetase (GS) catalyzes the addition of ammonia to glutamate in the first step of nitrogen incorporation into amino acid biosynthesis. It is expressed in non-green tissues like roots and legume nodules, where it serves to incorporate ammonia into usable metabolites, and in green tissues where it serves to metabolize ammonia produced by photorespiration (Canton *et al.*,

1999). When translated, the homologous open reading frame in the potato clone was missing the first 96 amino acids present in the *Arabidopsis* GS sequence. Although the ORF continued right up to the *Eco*RI site at the 3' end of the mapped region of the clone, it was missing only about six amino acids at the C-terminus when compared to the *Arabidopsis* GS sequence. The missing 5' end could be absent in vivo for this sequence; another possibility could be that it was mapped incorrectly on the genomic clone – since the entire clone was not mapped and sequenced the 5' end could be on another uncharacterized restriction fragment.

The other significant non-APase homology found in the genomic clone was at the protein level between regions of subclone pap7-10 and *copia*- and *gypsy*-like LTR (long terminal repeat) transposable elements. Retrotransposons are transposable elements that replicate themselves through the genome by producing an RNA molecule, and then reverse-transcribing it and inserting the resultant DNA into genomic DNA. LTR retrotransposons are bracketed by regions of repetitive sequence, and most fall into two categories: *copia*-like and *gypsy*-like. These are distinguished by the order of their protein-coding genes (protease, integrase, and reverse transcriptase/RNase H) in the *pol* ORF (Laten *et al.*, 1998). Subclone pap7-10 showed two separate regions of homology to LTR retrotransposons. In both cases the sequence only corresponded to a portion of the retroelement polyprotein. The 107 amino acids from the *copia*-like sequence were homologous to the integrase-coding portion of *Phaseolus vulgaris* Tpv2 (Garber *et al.*, 1999), while the 238 amino acids from the *gypsy*-like sequence were homologous to reverse transcriptase/RNase H-coding partial sequences from *Hordeum vulgare* and *Nicotiana tabacum* (Kumekawa *et al.*, 1999). In the case of the gypsy-like sequence,

there were ten stop codons in the homologous reading frame. Thus, these sequences do not appear to represent functional retrotransposons, or indeed, ones that were functional in the recent evolutionary past. Transposable element sequences can make up to 35% of human genome size (Labrador and Corces, 1997), and these sequences may make up much of the repetitive DNA that causes wide variations in plant genome sizes (Kumekawa *et al.*, 1999). In light of this, it is perhaps surprising that this genomic clone did not contain more sequence homologous to transposons.

Expression Analysis

Acid phosphatases are a ubiquitous component of the phosphate starvation response in plants. APase induction is found in roots in secreted forms (Wasaki *et al.*, 1997; Gilbert *et al.*, 2000), cell-wall bound forms (Dracup *et al.*, 1984, Wasaki *et al.*, 1997), and intracellularly (Szabo-Nagy *et al.*, 1987; Panara *et al.*, 1990;). There is also evidence for APase induction in leaves of phosphate-starved *Arabidopsis* (Haran *et al.*, 2000). Northern blot analysis of phosphate-fed and starved potato tissues did not reveal any expression of *pap7* under either condition in leaves, stem or roots. This suggests that *pap7* is not a phosphate starvation-induced gene, or since the Northern blot was of RNA from plants starved for 31 days, that it may be induced only in early stages of phosphate starvation. When the same RNA was probed with a known phosphate-starvation inducible gene (the potato phosphate transporter StPT2) a transcript was detected, though the signal was not as clear as in the original StPT2 paper (Leggewie *et al.*, 1997). This could indicate that the plants were not expressing PSI genes normally, possibly due to the extended starvation time. In the study reporting StPT2 expression (mainly in phosphate-

starved roots), plants were only starved for three days before analysis (Leggewie *et al.*, 1997).

Clearly, one of the most important continuations of this project will be the cloning and sequencing of the 5' end of the gene, containing the promoter region. Once this sequence is available, it can be examined for elements which suggest whether it is regulated by phosphate starvation, or possibly by some other conditions or developmental cues. This should increase the likelihood of determining *pap7*'s mode of regulation. Recently, the promoter region of a phosphate starvation-induced APase from Arabidopsis was cloned (Haran *et al.*, 2000). If *pap7* did turn out to be induced at an early stage of phosphate starvation, it would be interesting to compare the two genes' promoter sequences. Again, it is possible that *pap7* is a pseudogene, and does not possess a functional promoter at all.

Just because *pap7* does not appear to be expressed under phosphate starvation does not imply that it is never expressed *in vivo*. APases are induced under a number of other conditions, including seed germination (Biswas and Cundiff, 1991), senescence (De Leo and Sacher, 1970), flowering (Lal and Jaiswal, 1988), fruit ripening (Kanellis *et al.*, 1989), sink tissue removal ("depodding"; DeWald *et al.*, 1992), salt stress (Pan, 1987, del Pozo *et al.*, 1999), and oxidation stress (del Pozo *et al.*, 1999) as well as diverse tissues including seeds (Fujimoto *et al.*, 1977; Beck *et al.*, 1986; Van Etten and Waymack, 1991; Bose and Taneja, 1998; Ferriera *et al.*, 1998; Granjeiro *et al.*, 1999), legume root nodules (Penheiter *et al.*, 1997), and tubers (Kubicz *et al.*, 1974; Uehara *et al.*, 1974; Sugawara *et al.*, 1981; Kamenan *et al.*, 1983; Gellatly *et al.*, 1994). In some of these cases, such as senescing tissues and germinating seeds, the function of APases could be resource

mobilization: redistribution of P_i to other plant tissues in the case of senescence and senescence-causing stresses such as salinity, and retrieval of P_i from storage compounds in germinating seeds. In legume root nodules, APase was suggested to have a role in the synthesis of ureides, compounds involved in fixed nitrogen export from the nodule (Penheiter *et al.*, 1997). An intriguing possibility involves the role of APases in oxidative stress: at least one APase, AtACP5 (del Pozo *et al.*, 1999) also displays peroxidation activity and is induced by oxidative stress. Clearly, acid phosphatases are involved in many components of plant biochemistry other than the response to environmental phosphate starvation. Determination of the roles of PAP7 and PTAP will help to further elucidate some of the ways plants manage phosphate at the biochemical level.

Conclusions and Future Work

Pap7 appears to be a single-copy gene encoding an acid phosphatase homologous to potato tuber acid phosphatase and plant purple acid phosphatases. The *pap7* genomic clone did not contain the entire gene sequence, as the 5' end was interrupted at the clone boundary. The available sequence codes for 404 amino acids, which share 53% identity with PTAP and 63% identity with kidney bean PAP. Northern blotting did not detect *pap7* transcript in either phosphate-fed or phosphate-starved plant tissues, although the tissues studied may not have been expressing PSI genes normally due to extended starvation.

Clearly, the most important immediate continuation of this project is the cloning and sequencing of the 5' end of the *pap7* gene, including the beginning of the coding sequence and the promoter elements. Work is currently underway to isolate the upstream

genomic sequence using a PCR strategy. Restriction-digested potato genomic DNA will be ligated to an adapter DNA and amplified using an adapter-specific primer and a primer from the 5' end of the known *pap7* sequence. This should allow amplification of one or more upstream genomic clones overlapping the 5' end of the known *pap7* sequence.

Although promoter sequences often share homology with others activated under the same conditions, expression analysis is still required to confirm under which conditions or in which tissues a gene is expressed. If the *pap7* promoter contains homology to other known promoters, this process is simplified. However if it does not, it would be necessary to look for expression in plants under different nutrient stresses and developmental stages, using Northern blotting or RT-PCR screening to identify transcripts. Alternately, the promoter region could be combined with a reporter gene such as glutathione s-transferase (Smith and Johnson, 1988), β-glucuronidase (Jefferson *et al.*, 1987), or a fluorescent protein (Chalfie *et al.*, 1994) and used to transform plants. The transformants could be examined for reporter gene expression under different conditions.

If *pap7* expression was confirmed *in vivo*, attempts could be made to purify and characterize the gene product. Although the protein could be purified from potato tissues, isolation of a cDNA clone would allow it to be cloned into an expression vector and expressed in a bacterial or yeast system. An eukaryotic system would not only provide more protein to study but would also increase the likelihood that it would be folded and glycosylated correctly.

Literature Cited

Akiyama, T., Uchimiya, H., and Suzuki, H. (1981) Gibberellic acid-induced increase in activity of a particular isozyme of acid phosphatase in wheat half-seeds. Plant Cell Physiol. **22(6):** 1023-1028.

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

Amemura, M., Makino, K., Shinagawa, H., Kobayashi, A., and Nakata, A. (1985) Nucleotide sequence of the genes involved in phosphate transport and regulation of the phosphate regulon in *Escherichia coli*. J. Mol. Biol. **184:** 241-250.

Bailey, K.M., Phillips, D.J., and Pitt, D. (1976) Effects of gibberellic acid on the activation, synthesis and release of acid phosphatase in barley seed. J. Exp. Bot. **27(97):** 324-336.

Beck, J., McConachie, L., Summors, A., Arnold, W., De Jersey, J., and Zerner, B. (1986) Properties of a purple phosphatase from red kidney bean: a zinc-iron metalloenzyme. Biochim. Biophys. Acta **869:** 61-68.

Biswas, T.K. and Cundiff, C. (1991) Multiple forms of acid phosphatase in germinating seeds of *Vigna sinensis*. Phytochem. **30(7):** 2119-2125.

Bose, S.K. and Taneja, V. (1998) Induction of a germination-specific, low molecular weight, acid phosphatase isozyme with specific phosphotyrosine phosphatase activity in lentil (*Lens esculenta*) seeds. Biochem. Biophys. Res. Comm. **250(3):** 629-634.

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

Brendel, V. and Kleffe, J. (1998) Prediction of locally optimal splice sites in plant pre-mRNA with applications to gene identification in *Arabidopsis thaliana* genomic DNA. Nucl. Acids Res. **26(20):** 4748-4757.

Burleigh, S.M. and Harrison, M.J. (1998) Characterization of the Mt4 gene from *Medicago truncatula*. Gene **216:** 47-53.

Canton, F.R., Suarez, M.F., Jose, E.M., and Canovas, F.M. (1999) Expression analysis of a cytosolic glutamine synthetase gene in cotyledons of scots pine seedlings: developmental, light regulation and spatial distribution of specific transcripts. Plant Mol. Biol. **40(4):** 623-634.

Cashikar, A.G., Kumaresan, R., and Rao, N.M. (1997) Biochemical characterization and subcellular localization of the red kidney bean purple acid phosphatase. Plant Physiol. **114:** 907-915.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994) Green fluorescent protein as a marker for gene expression. Science **263(5148):** 802-805.

Cheng, H-F. and Tao, M. (1989) Purification and characterization of a phosphotyrosyl-protein phosphatase from wheat seedlings. Biochim. Biophys. Acta **998:** 271-276.

Chung, R.P-T. and Polya, G.M. (1992) Copurification and characterization of poppy seed phosphatase and phosphoprotein phosphatase activities. Plant Sci. **84:** 153-162.

De Leo, P. and Sacher, J.A. (1970) Control of ribonuclease and acid phosphatase by auxin and abscisic acid during senescence of *Rhoeo* leaf sections. Plant Physiol. **46:** 806-811.

del Pozo, J.C., Allona, I., Rubio, V., Leyva, A., de la Pena, A., Aragoncillo, C., and Paz-Ares, J. (1999) A type 5 acid phosphatase gene from *Arabidopsis thaliana* is induced by phosphate starvation and by some other types of phosphate mobilizing/oxidative stress conditions. Plant J. **19(5):** 579-589.

DeWald, D.B., Mason, H.S., and Mullet, J.E. (1992) The soybean storage proteins VSPα and VSPβ are acid phosphatases active on polyphosphates. J. Biol. Chem. **267(22):** 15958-15964.

Dong, B., Rengel, Z., and Delhaize, E. (1998) Uptake and translocation of phosphate by *pho2* mutant and wild-type seedlings of *Arabidopsis thaliana*. Planta **205:** 251-256.

Dracup, M.N.H., Barrett-Lennard, E.G., Greenway, H., and Robson, A.D. (1984) Effect of phosphorus deficiency on phosphatase activity of cell walls from roots of subterranean clover. J. Exp. Botany **35(153):** 466-480.

Duff, S.M.G., Sarath, G., and Plaxton, W.C. (1994) The role of acid phosphatases in plant phosphorus metabolism. Physiol. Plant. **90:** 791-800.

Duff, S.M.G., Lefebvre, D.D., and Plaxton, W.C. (1991) Purification, characterization, and subcellular localization of an acid phosphatase from black mustard cell-suspension cultures: comparison with phospho*enol*pyruvate phosphatase. Arch. Biochem. Biophys. **286:** 226-232.

Duff, S.M.G., Plaxton, W.C., and Lefebvre, D.D. (1991) Phosphate-starvation response in plant cells: *de novo* synthesis and degradation of acid phosphatases. Proc. Nat. Acad. Sci. **88:** 9538-9542.

Duff, S.M.G., Lefebvre, D.D., and Plaxton, W.C. (1989) Purification and characterization of phospho*enol*pyruvate phosphatase from *Brassica nigra* suspension cells. Plant Physiol. **90:** 734-741.

Duff, S.M.G., Moorhead, G.B.G., Lefebvre, D., and Plaxton, W.C. (1989) Phosphate starvation inducible "bypasses" of adenylate and phosphate dependent glycolytic enzymes in *Brassica nigra* suspension cells. Plant Physiol. **90:** 1275-1278.

Durmus, A., Eicken, C., Sift, B.H., Kratel, A., Kappl, R., Huttermann, J., and Krebs, B. (1999) The active site of purple acid phosphatase from sweet potatoes *(Ipomoea batatas)*. Eur. J. Biochem. **260:** 709-716.

Eckstein, J.W., Beer-Romero, P., and Berdo, I. (1996) Identification of an essential acidic residue in Cdc25 protein phosphatase and a general three-dimensional model for a core region in protein phosphatases. Prot. Sci. **5:** 5-12.

Egoavil, C., Marton, H.A., Baynton, C.E., McCullough, A.J., and Schuler, M.A. (1997) Structural analysis of elements contributing to 5'splice site selection in plant pre-mRNA transcripts. Plant J. **12(5):** 971-980.

Erion, J.L., Ballo, B., May, L., Bussell, J., Fox, T., and Thomas, S. (1991) Isolation and characterization of a tomato acid phosphatase complementary DNA associated with nematode resistance. Plant Physiol. **97:** 1462-1469.

Ferriera, C.V., Granjeiro, J.M., Taga, E.M., and Aoyama, H. (1998) Purification and characterization of multiple forms of soybean seed acid phosphatases. Plant Physiol. Biochem. **36(7):** 487-494.

Fujimoto, S., Nakagawa, T., and Ohara, A. (1977) Isolation of violet-coloured acid phosphatase from soybean. Agr. Biol. Chem. **41(3):** 599-600.

Gellatly, K.S., Moorhead, G.B.G., Duff, S.M.G., Lefebvre, D.D., and Plaxton, W.C. (1994) Purification and characterization of a potato tuber acid phosphatase having significant phosphotyrosine phosphatase activity. Plant Physiol. **106:** 223-232.

Gellatly, K. (1996) Biochemical and molecular characterization of potato tuber acid phosphatase and other members of the plant acid phosphatase family. Ph.D. thesis. Queen's University, Kingston, Ontario.

Gilbert, G.A., Knight, J.D., Vance, C.P., and Allan, D.L. (2000) Acid phosphatase activity in phosphorus-deficient white lupin roots. Plant Cell Env. **22:** 801-810.

Goldstein, A.H., Danon, A., Baertlein, D.A., and McDaniel, R.G. (1988) Phosphate starvation inducible metabolism in *Lysopersicon esculentum*. Plant Physiol. **87:** 716-720.

Graber, J.H., Cantor, C.R., Mohr, S.C., and Smith, T.F. (1999) *In silico* detection of control signals: mRNA 3'-end-processing sequences in diverse species. Proc. Nat. Acad. Sci. **96(24):** 14055-14060.

Granjeiro, P.A., Ferriera, C.V., Granjeiro, J.M., Taga, E.M., and Aoyama, H. (1999) Purification and kinetic properties of a castor bean seed acid phosphatase containing sulfhydryl groups. Physiol. Plant. **107:** 151-158.

Haran, S., Logendra, S., Seskar, M., Bratanova, M., and Raskin, I. (2000) Characterization of *Arabidopsis* acid phosphatase promoter and regulation of acid phosphatase expression. Plant Phys. **124:** 615-626.

Higgins, D.G. and Sharp, P.M. (1988) CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. Gene **73:** 237-244.

Huber, S.C. and Huber, J.L. (1996) Role and regulation of sucrose phosphate synthase in higher plants. Ann. Rev. Plant Physiol. Plant Mol. Biol. **47:** 431-444.

Igaue, I., Watabe, H., Takahashi, K., Takekoshi, M., and Morota, A. (1976) Violet-coloured acid phosphatase isozymes associated with cell wall preparations from rice plant cultured cells. Agr. Biol. Chem. **40(4):** 823-825.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987) GUS-fusions: ,β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J. **6:** 3901-3907.

Kamenan, A. and Diopoh, J. (1983) Properties of two membrane-bound acid phosphatases compared with those of a cytoplasmic acid phosphatase from *Dioscorea cayanensis rotundata*. Plant Sci. Letters **32:** 305-312.

Kanellis, A.K., Solomos, T., and Mattoo, A.K. (1989) Changes in sugars, enzymic activities and acid phosphatase isoenzyme profiles of bananas ripened in air or stored in 2.5% O_2 with and without ethylene. Plant Physiol. **90:** 251-258.

Klabunde, T., Strater, N., Krebs, B., and Witzel, H. (1995) Structural relationship between the mammalian Fe(III)-Fe(II) and the Fe(III)-Zn(II) plant purple acid phosphatases. FEBS Letters **367:** 56-60.

Klabunde, T., Strater, N., Frohlich, R., Witzel, H., and Krebs, B. (1996) Mechanism of Fe(III)-Zn(II) purple acid phosphatase based on crystal structure. J. Mol. Biol. **259:** 737-748.

Kubicz, A., Morawiecka, B., and Kruzel, M. (1974) Heterogeneity of acid phosphatase from potato tubers. Acta Biochim. Pol. **21(2):** 113-117.

Kumekawa, N., Ohtsubo, H., Horiuchi, T., and Ohtsubo, E. (1999) Identification and characterization of novel retrotransposons of the gypsy type in rice. Mol. Gen. Genet. **260:** 593-602.

Labrador, M., and Corces, V.G. (1997) Transposable element-host interactions: regulation of insertion and excision. Ann Rev. Genet. **31:** 381-404.

Lal, M., and Jaiswal, V.S. (1998) Modification of flower sex and acid phosphatase activity by phthalimides in female plants of *Morgus nigra* L. Plant Growth Reg. **7:** 29-37.

Laten, H.M., Majumdar, A., and Gaucher, E.A. (1998) SIRE-1, a *copia/Ty1*-like retroelement from soybean, encodes a retroviral envelope-like protein. Proc. Nat. Acad. Sci. USA **95:** 6897-6902.

Lee, R.B. (1988) Phosphate influx and extracellular phosphatase activity in barley roots and rose cells. New Phytol. **109:** 141-148.

Lefebvre, D., Duff, S.M.G., Fife, C.A., Julien-Inalsingh, C., and Plaxton, W.C. (1990) Response to phosphate deprivation in *Brassica nigra* suspension cells. Physiol. Plant. **93:** 504-511.

Lefebvre, D. and Glass, A. (1982) Regulation of phosphate influx in barley roots: effects of phosphate deprivation and reduction of influx with provision of orthophosphate. Physiol. Plant. **54:** 199-206.

Leggewie, G., Willmitzer, L., and Riesmeier, J.W. (1997) Two cDNA's from potato are able to complement a phosphate uptake-deficient yeast mutant: identification of phosphate transporters from higher plants. Plant Cell **9:** 381-392.

Lenburg, M.E., and O'Shea, E.K. (1996) Signaling phosphate starvation. Trends Biochem. Sci. **21:** 383-387.

Li, Q. and Hunt, A.G. (1995) A near-upstream element in a plant polyadenylation signal consists of more than six nucleotides. Plant Mol. Biol. **28:** 927-934.

Liu, C., Muchhal, U.S., Uthappa, M., Kononowicz, A.K., and Raghothama, K.G. (1998) Tomato phosphate transporter genes are differentially regulated in plant tissues by phosphorus. Plant Physiol. **116:** 91-99.

Liu, H., Trieu, A.T., Blaylock, L.A., and Harrison, M.J. (1998) Cloning and characterization of two phosphate transporters from *Medicago truncatula* roots:

regulation in response to phosphate and to colonization by Arbuscular Mycorrhizal (AM) fungi. Mol.-Plant Micr. Inter. **11(1):** 14-22.

Makino, K., Shinagawa, H., Amemura, M., Kimura, S., Nakata, A., and Ishihama, A. (1988) Regulation of the phosphate regulon of *Escherichia coli*: activation of *pstS* transcription by PhoB protein *in vitro*. J. Mol. Biol. **203:** 85-95.

Makino, K., Shinagawa, H., Amemura, M., and Nakata A. (1986) Nucleotide sequence of the *phoB* gene, the positive regulatory gene for the phosphate regulon of *Escherichia coli* K-12. J. Mol. Biol. **190:** 37-44.

Maugenest, S., Martinez, I., and Lescure, A.-M. (1997) Cloning and characterization of a cDNA encoding a maize seedling phytase. Biochem J. **322:** 511-517.

Maugenest, S., Martinez, I., Godin, B., Perez, P., and Lescure, A-M. (1999) Structure of two maize phytase genes and their spatio-temporal expression during seedling development. Plant Mol. Biol. **39:** 503-514.

Metcalf, W.W., Steed, P.M., and Wanner, B. (1990) Identification of phosphate starvation-inducible genes in *Escherichia coli* by DNA sequence analysis of *psi::lacZ*(Mu d1) transcriptional fusions. J. Bacteriol. **172(6):** 3191-3200.

Miernyk, J.A. (1992) Purification and characterization of the major acid phosphatase isozyme secreted by maize endosperm cultures. Phytochem. **31(8):** 2613-2616.

Muchhal, U.S., Pardo, J.M., and Raghothama, K.G. (1996) Phosphate transporters from the higher plant *Arabidopsis thaliana*. Proc. Nat. Acad. Sci. **93:** 10519-10523.

Muchhal, U.S. and Raghothama, K.G. (1999) Transcriptional regulation of plant phosphate transporters. Proc. Nat. Acad. Sci. **96:** 5868-5872.

Mullaney, E.J. and Ullah, A.H.J. (1998) Conservation of the active site motif in *Aspergillus niger (ficuum)* pH 6.0 optimum acid phosphatase and kidney bean purple acid phosphatase. Biochem. Biophys Res. Comm. **243:** 471-473.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol. Plant. **15:** 473-497.

Nakazato, H., Okamoto, T., Nishikoori, M., Washio, K., Morita, N., Haraguchi, K., Thompson, G.A., and Okuyama, H. (1998) The glycosylphosphatidylinositol-anchored phosphatase from *Spirodela oligorrhiza* is a purple acid phosphatase. Plant Physiol. **118:** 1015-1020.

Nandi, S.K., Pant, R.C., and Nissen, P. (1987) Multiphasic uptake of phosphate by corn roots. Plant Cell Env. **10:** 463-474

Neuman, H. (1968) Substrate selectivity in the action of alkaline and acid phosphatases. J. Biol. Chem. **243(18):** 4671-4676.

Ogawa, N., Hayashi, N., Saito, H., Noguchi, K., Yamashita, Y., and Oshima, Y. (1994) Regulatory circuit for phosphatase genes in *Saccharomyces cerevisiae*: specific *cis*-acting sites in *PHO* promoters for binding the positive regulator Pho4p. In: Torriani-Gorini, A., Yagil, E., and Silver, S. (Eds.), Phosphate in Microorganisms: Cellular and Molecular Biology. American Society for Microbiology, Washington, DC, pp. 56-62.

Oshima, Y., Ogawa, N., and Harashima, S. (1996) Regulation of phosphatase synthesis in *Saccharomyces cerevisiae* – a review. Gene **179:** 171-177.

Ozawa, K., Osaki, M., Matsui, H., Honma, M., and Tadano, T. (1995) Purification and properties of acid phosphatase secreted from lupin roots under phosphorus-deficiency conditions. Soil Sci. Plant Nutr. **41(3):** 461-469.

Pan, S.-m. (1987) Characterization of multiple acid phosphatases in salt-stressed spinach leaves. Aust. J. Plant Physiol. **14:** 117-124.

Panara, F., Pasqualini, S., and Antonielli, M. (1990) Multiple forms of barley root acid phosphatase: purification and some characteristics of the major cytoplasmic isoenzyme. Biochim. Biophys. Acta. **1037:** 73-80.

Paul, E.M. and Williamson, V.M. (1987) Purification and properties of acid phosphatase-1 from a nematode resistant tomato cultivar. Plant Physiol. **84:** 399-403.

Penheiter, A.R., Duff, S.M.G., and Sarath, G. (1997) Soybean root nodule acid phosphatase. Plant Physiol. **114:** 597-604.

Pfeiffer, W. (1996) Auxin induces exocytosis of acid phosphatase in coleoptiles from *Zea mays*. Physiol. Plant. **98:** 773-779.

Pfeiffer, W. (1998) Expression and low pH increase the activity of an apoplastic acid phosphatase in growing tissues from *Zea mays*. Physiol. Plant. **102:** 111-118.

Raghothama, K.G. (2000) Phosphate transport and signaling. Curr. Opin. Plant Biol. **3:** 182-187.

Rao, N.N. and Torriani, A. (1990) Molecular aspects of phosphate transport in *Escherichia coli*. Mol. Microbiol. **4(7):** 1083-1090.

Randall, D.D. and Tolbert, N.E. (1971) 3-Phosphoglycerate phosphatase in plants I. Isolation and characterization from sugarcane leaves. J. Biol. Chem. **246(17):** 5510-5517.

Rebeille, F., Bligny, R., Martin, J.-B., and Douce, R. (1983) Relationship between the cytoplasm and the vacuole phosphate pool in *Acer pseudoplatanus* cells. Arch. Biochem. Biophys. **225:** 143-148.

Richardson, A.E., Hadobas, P.A., and Hayes, J.E. (2000) Acid phosphomonoesterase and phytase activities of wheat (*Triticum aestivum* L.) roots and utilization of organic phosphorus substrates by seedlings grown in sterile culture. Plant Cell Env. **23:** 397-405.

Rothnie, H.M. (1996) Plant mRNA 3'-end formation. Plant Mol. Biol. **32:** 43-61.

Rudolph, H. and Hinnen, A. (1987) The yeast *PHO5* promoter: phosphate-control elements and sequences mediating mRNA start-site selection. Proc. Nat. Acad. Sci. **84:** 1340-1344.

Schenk, G., Ge, Y., Carrington, L.E., Wynne, C.J., Searle, I.R., Carroll, B.J., Hamilton, S., and de Jersey, J. (1999) Binuclear metal centers in plant purple acid phosphatases: Fe-Mn in sweet potato and Fe-Zn in soybean. Arch. Biochem. Biophys. **370(2):** 183-189.

Scholten, M. and Tommassen, J. (1994) Effect of mutations in the –10 region of the *phoE* promoter in *Escherichia coli* on regulation of gene expression. Mol. Gen. Genet. **245:** 218-223.

Shimogawara, K., Wykoff, D.D., Useda, H., and Grossman, A.R. (1999) *Chlamydomonas reinhardtii* mutants abnormal in their responses to phosphorus deprivation. Plant Phys. **120:** 685-693.

Shinagawa, H., Makino, K., and Nakata, A. (1983) Regulation of the *pho* regulon in *Escherichia coli* K-12: genetic and physiological regulation of the positive regulatory gene *phoB*. J. Mol. Biol. **168:** 477-488.

Smith, D.B., and Johnson, K.S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. Gene **67(1):** 31-40.

Sugawara, S., Inamoto, Y., and Ushijima, M. (1981) Resolution and some properties of acid phosphatase isozymes bound to the cell wall of potato tubers. Agr. Biol. Chem. **45(8):** 1767-1773.

Surin, B.P., Jans, D.A., Fimmel, A.L., Shaw, D.C., Cox, G.B., and Rosenberg, H. (1984) Structural gene for the phosphate-repressible phosphate-binding protein of *Escherichia coli* has its own promoter: complete nucleotide sequence of the *phoS* gene. J. Bacteriol. **157:** 772-778.

Szabo-Nagy, A., Olah, Z., and Erdei, L. (1987) Phosphatase induction in roots of winter wheat during adaptation to phosphorus deficiency. Physiol. Plant. **70:** 544-552.

Thaller, M.C., Schippa, S., and Rossolini, G.M. (1998) Conserved sequence motifs among bacterial, eukaryotic, and archaeal phosphatases that define a new phosphohydrolase superfamily. Prot. Sci. **7:** 1647-1652.

Toh-e, A., Inouye, S., and Oshima, Y. (1981) Structure and function of the *PHO82-pho4* locus controlling the synthesis of repressible acid phosphatase of *Saccharomyces cerevisiae*. J. Bacteriol. **145(1):** 221-232.

Tommassen, J., de Geus, P., and Lugtenberg, B. (1982) Regulation of the *pho* regulon of *Escherichia coli* K-12: cloning of the regulatory genes *phoB* and *phoR* and identification of their gene products. J. Mol. Biol. **157:** 265-274.

Trull, M.C., Guiltinan, M.J., Lynch, J.P., and Deikman, J. (1997) The responses of wild-type and ABA mutant *Arabidopsis thaliana* plants to phosphorus starvation. Plant Cell Env. **20:** 85-92.

Uehara, K., Fujimoto, S., Taniguchi, T., and Nakai, K. (1974) Studies on violet-coloured acid phosphatase of sweet potato II: enzymatic properties and amino acid composition. J. Biochem. **75:** 639-649.

Ueki, K. and Sato, S. (1971) Effect of inorganic phosphate on the extracellular acid phosphatase activity of tobacco cells cultured *in vitro*. Physiol. Plant. **24:** 506-511.

Ullah, A.H.J. and Dischinger, H.C. (1993) Identification of active-site residues in *Aspergillus ficuum* extracellular pH 2.5 optimum acid phosphatase. Biochem. Biophys Res. Comm. **192(2):** 754-759.

Usuka, J. and Brendel, V. (2000) Gene structure prediction by spliced alignment of genomic DNA with protein sequences: increased accuracy by differential splice site scoring. J. Mol. Biol. **297:** 1075-1085.

Utsugi, S. (1998) *Arabidopsis thaliana* vegetative storage protein (VSP) genes: gene organization and tissue-specific expression. Plant Mol. Biol. **38(4):** 565-576.

Van Etten, R. and Waymack, P.P. (1991) Substrate specificity and pH dependence of homogenous wheat germ acid phosphatase. Arch. Biochem. Biophys. **288(2):** 634-645.

Vincent, J.B., Crowder, M.W., and Averill, B.A. (1992) Hydrolysis of phosphate monoesters: a biological problem with multiple chemical solutions. Trends Biochem. Sci. **17:** 105-110.

Vogel, K., and Hinnen, A. (1990) The yeast phosphatase system. Mol. Microbiol. **4(12):** 2013-2017.

Wanner, B.L. (1993) Gene regulation by phosphate in enteric bacteria. J. Cell. Biochem. **51:** 47-54.

Wasaki, J., Ando, M., Ozawa, K., Omura, M., Osaki, M., Ito, H., Matsui, H., and Tadano, T. (1997) Properties of secretory acid phosphatases from lupin roots under phosphorus-deficient conditions. In: Ando, T. *et al.* (Eds.), Plant nutrition – for sustainable food production and environment. Kluwer Academic Publishers, Japan, pp. 295-300.

Wykoff, D.D., Grossman, A.R., Weeks, D.P., Usuda, H., and Shimogawara, K. (1999) Psr1, a nuclear localized protein that regulates phosphorus metabolism in *Chlamydomonas*. Proc. Nat. Acad. Sci. **96(26):** 15336-15341.

Yamada, M., Makino, K., Amemura, M., Shinagawa, H., and Nakata, A. (1989) Regulation of the phosphate regulon of *Escherichia coli*: mutant *phoB* and *phoR* genes causing different phenotypes. J. Bacteriol. **171(10):** 5601-5606.

Yeh, C-C., Tsay, H-S., Yeh, J-H., Tsai, F-Y., Shih, C-Y., and Kao, C-H. (1995). A comparative study of the effects of methyl jasmonate and abscisic acid on some rice physiological processes. J. Plant Growth Reg. **14:** 23-28.

Appendices

A Nucleotide sequence of *Eco*RI subclone pap7-7

B Nucleotide sequence of *Eco*RI subclone pap7-8

C Partial nucleotide sequence of *Eco*RI subclone pap7-10

D Nucleotide sequence of *Eco*RI/*Sal*I subclone pap7-500

E Sequences of oligonucleotides used

Appendix A. Nucleotide sequence of *Eco*RI subclone pap7-7.

```
GAATTCCTATAAATGGATGAAGTTGGTAGCTCATATTTCA 40
ATTTCATTTCCATCATAGTTCTATGACGGTTAACCCGATT 80
GTAGTTTTGTTGGAATTTGGGATTTTGATGAAGCAGCAGA 120
TGACACTCTATTTTGTTCTTCTCCTTTGCTTGTTCCCTCT 160
TGTTGCTGATAAATTACAGGCTAAAATCTCAAGTGATGAA 200
AGCATTTCTATGATATAGAAGAGTTAATTCAAGCATGCAT 240
GAACAACATGTCTGCTTATTGCCAATATAATTCAAATTGA 280
GGGTATTTTAGACTTCTCATAGTACCTAACTATATACCAT 320
GCCTTTATTTTGTTTGCCATGTATCACTTCTCTATTTTGT 360
TTTATTTTTGGCTCTAAATCTTCTACAGCCTTTTCGGTTC 400
ATCGCCATCCAGCTTGTCTGGGTCTTAGCATCATCTATCC 440
ACTATTAAGTTTGTTTGTATGTTCTTTCTTCTGCTATCCA 480
GTACCTTTTATAATACAAGTCTGAATCTGGCAATAACTAC 520
TGCTCTCTGTTCTTGGTTGTTCTTTCTAGATTCTGTCTCT 560
ACACTCGCTTGGAGATTTATGTTGTTCCACTGCCTCTGAT 600
CGGTGAGTTTCACTGATATTCTATCACGTCTTTTTCCTGC 640
CTGCTTTCTAAATTCAAATCTGGGTTTACGGTTTTCATCT 680
AAGGAAAAAACAAGTAATAATATAGATGATTCTGTTTCAT 720
GAATTAAGTATCTTAATGGCCTTCCAGTGGTTGTTGTTTG 760
ATGCATTTATTCATATGAAACATCCTGAATATCTGGAAGA 800
GTTACACTCCTTGAATTC 818
```

Appendix B. Nucleotide sequence of *Eco*RI subclone pap7-8.

GAATTCAAATAGAGTCACAACCCTTTACAACATAGTCTCAAAATGAAGGTACAAAGAAAG 60
AAAATAGAACAACATATTCCATCCTCGAACCATGAGGACTCATCGAAGTAACGATATTAG 120
TTGATCCAAGCTTTGACCGCAAAAGAAGCACACCTCTAGAACCGGGCCCTACACTTAGAG 180
AAAAGTGTCGAAAATATGGGTTAGTACAAATATGCACTAATTATGATAGTCATGCATAAA 240
GTCCTTCAAACATGCCAAAAAGGGACATTTTATTTGAAACCATGCACTTTATCAAGTGTA 300
AGATTCATCATAAACATTAGTGTAAAAACACAACATAAGTACTTAGATCTTCATATCATA 360
AGAGAACCATTGCATAAGCAATCCTAATCTATATTTGTTCAATGCATAGACAAGATCCCA 420
TATAATACCAAGTACCACCTTTAGATCACCACATTCATACTTATCATGCTTTGACCTTAT 480
CCATAGCTAATACTCGTAGACAAGGAAACATTCATGTACAATAACTCATGTCAACGAAAG 540
CCACTAATATAACCAATCCACATTAGGAATGCATAAATTCATATACTTCATACTTAAGAG 600
AAGAACACCATGACCACTCTGAAAGCGTACTTGTGCAATGCATGAATGAAATCCCATACC 660
CCCACTCACACTAAGTAGAACCTCTTGAGGAACCGTAGAACAAAAATATGTTTTTGTTCG 720
TAGTTCATGTTCTTGTTGTGGAGGGTAATAGCCTTAACCGACATAGACCATGTGAGATTA 780
CATGGAATCCTGTATCCTGCCCCCATACCGAAAAGAGGTGTCCTACTTGCCTAAGGAAAA 840
CTAAAGTGTAGCTAGGGGATCCCAACCAATCCTTGTAGTTCGACCACGACCTCAAAAGTA 900
GGTATATGCATACGATCACTATGCTGAAAATGGGGACATAATTGAATATTTATCAACCTC 960
CACAAAGACTTTGGGGACCACCCATACGAACCTGAGTATCGCCTCTAAACAGAGATGCTA 1020
CAACCCTGCGCCTAGGGCAATCTCTAGCATAGTGCCCAACCTCGCCGCAATCAAAGAAAC 1080
CTTGATCACCCCAAGGTATAGGTAGAAGACTGTCTAATCCACCCTAACCTCCACTGCCTA 1140
AAGTTCAAAGCGAGATCTGAATTTGTCAGCTAGACTGCCCCTAGGGATAATAAGCATGCT 1200
GACCTTTGTTTGAAAGCTCCTACCATGGAACTGAGAACCACTGTAGATACCCTGGAAGCA 1260
AGTCCTCTTATATCTACCCCACAGGCCTCACGATGGAGCTTCTCCATAGTACGAGTATGG 1320
TCAACTACATCTACGAAGGATCTATCTGCTTTCTCTTTAGGAGTTTCAACTTTCTTTCTT 1380
GTGTAGAATTGAATATTGAAGAAATCTCAAAAGAAATAAAAGATTTGATGATCTAATCAA 1440
ATAATAGTAGCTCTTCAATACTAGCCAAATCATTTGCAATCTTTTTCTTCTCTTCCCAAA 1500
TTATGTCTTTAGAAATTTTCACTTTCTATGGTGTATCAACTACCACCATATAATTATCGC 1560
ACATCCAGCCACTCTGAGTTGTAACAACTATACAATGACTCACATTACCTTCTTCAGCCT 1620
TCGAATTTTCCATAGTGTCGCTCAGAAGAGTGCCCTTTTGCCTTTGGTTATCAAGGCTAA 1680
CAGTTGGCTTAACTAACCTCTATTTGTTTAGTCGATGTGGATGAGAGTTTACCACTTAGT 1740
CTATGCTTGCTAGGTCACTATGTAGTTTTATACCATCATCTATAGATTAAAATTTCTTAA 1800
GGACACAAGATAACATATCATCTATTTTGTGCTCATGGGGTCACTCACGATCGAATGCTC 1860
AAAATTCTGGTGAGCACCTTGGGTGGAACATATGGATTGAAGCTACTATGGTTGAAGCCT 1920
TGATCCTAGTTGCCCTGGTTCCTATCTCACTAACCTATGTTCTGCTAATCTTACTCCCGA 1980
TTCCAACCTTAATTCCCTAGGCCTTGCTGCCAAAATCCCCGAGTGGAACCTTGATAGTTA 2040
GGTTGGGAACCTCTCAACTAATTGCTCACATAATGAGCCTCTTCTTCATACGACTAATAA 2100
ACTTCCTCATTTTTGTAGCCCATTATTAGAAGCCATAACATTCACCTCTTGGAATAGTGC 2160
ACCCATCACATCCTTTGCTAGGATTACGATCTATGTCATCAAGTATGTCATTTTTTCGTC 2220
TTAGGCCTTTTCTCTGAGTATCTCATCCTTTGATAGCTTTCTGATTGAAGCATTGTCCAT 2280
TACCTCAATATTTTTAGTGTGACATGCCCATGCCTTGCATGTCTTTGTTATGCAATCAAG 2340
TAGCGTGTTGGTCTCATTATAGGAATGCTTCACCAGTAAACCCTCTACAATATTGTTGGC 2400
CATATCTAAGGATCTGTAAAAGCTCTCTAGCAACACCCTATTTGATAATTTATAGTTTGA 2460
ACACTGAACCACCATAATTTTAATCATTAGCTAGTGTAACATCTCGCAACTAGAAAGAAC 2520
TAGAAGGCAGTTTAAAATCTGGAAATAGTTATTTTTGGAAAGTATAAGGAGATCTGGAAA 2580
ATTTTAAGCAAGTTAAAGTGAGTTTTTGGTCAACTTCAAACGACCATAACTCCTAGCTCA 2640
AGATGAGTTAGATGTATTTCCAGATATGGTTGGAAATTTTTTGGAAAGATCTTTCCAACA 2700
CCGCCAAGATTGCATGAATCTGAGTTTGTATGAGTGAGATACGCCTATTGGAAGTTGGGA 2760
TGTACAAATAAGGAAATCTCCAATCCGGATTTTTGAAGGGTATTTCAATCTTTTCCTTAT 2820
CCAATTAATTAAATTCGTTTTAGTGAATTAATTAAGGGTCTAAACTGAACCTTGTCAGTT 2880
TTACGCTTTTGGAAAAGAGTTAGGGTTTTAGAGAAGAGGAAAGAAGAGGAAAAAAGAGAA 2940
GGAGAAGCAAGGAAATCGTCAAGTTCGTCGAGTTTAGCTTGTGGATTTCGTCGAGGGTGA 3000
TCCCTACAAGGTATGTGAGTCTCACATAAAGTTGGTTTCATTCACCCATACACCAAACAT 3060
GTTTAGTTCAGCGAAATTCGTCCTAAAAGAGTTTGAAAGTTAATGTTTTTGACATGAATT 3120
C 3121

Appendix C. Nucleotide sequence of *Eco*RI subclone pap7-10 a) from T3 end of vector MCS, and b) from T7 end of vector MCS (following page).

a)

```
GAATTCCCATTTTGTCCCTCTTCATTGATGGAGATACTAATTGGTGTTTCCTCTATTCTA 60
GTGCTGGTTTACAAATATATCACTCTTTTTTTTCTACAATAAATGTTGCTCCTTTTCTCA 120
AATGACTCTGTTGCGGTAATTTAAGAGTAATGAAGCGACTTACTATTACAATTGCAAAGA 180
GTATATCTGGCCTTGTACAGGTCAATGTGATTTTTGAACAGATTGGAGTCAATCTTTTCT 240
CTATCATCTAACTTTGACAACTTTATTCCACTTTCCATTGGAATATTCATGAGGTTGCAA 300
TCGAACCTGTTGAACTTCAAAATCTCCTTTATATATCTTTCTTGAGATATGAAGATGCCT 360
TTCTCCATTTGCTTCACTTCTAAGCCCAAGTAATATGACATGAGACCTATGTCCGTCATC 420
TCGAACTCAAGAGACAGCTTTCTTGAAAGCTTCAAACAAACTTGGTTTATTAGTCCTAAG 480
CTGCTCAGACTATTATGGTGCCCCACCCGTGTCGACATGACATGGGTGTGGGTGTGAGAT 540
CTTTACCGGATCTGGTCAGTAACCAAACTCGGCAGAGAAATTCGGGAGAGATAAAAACGA 600
ATGCTAAAGGGAACTTGAAGAGGATGCAATGTCGGTTCCTTTTGCGGCCGGTTGGGATTG 660
CGGCGCTTAATCAACAGGCAGCTTAGGGTACCGTCTAAGTTTCCATATGATGTTTCATAA 720
TGCAGACATTTTTATAACTCTAGTTTTAGATATTTGAATTATTATCCATACCCCTGAATA 780
ATATAATTTAAATCATGAAGGATCTGCATCCGTATTGGACACCCACACCAGAATCCGAGC 840
AGTTTAGAGCTTACCAGATTTGCCCTTGTGATGTCGCATTGCTGATGTGTTGTTTGTAAT 900
AAGAAAGATGAGAGGTATAGTCTTCAAAATGTAGTTGATAGATTTACGCTGACCTTAATC 960
TACAGTGCAGTGGTTATATCATATATGTCACCTGAACAAGCAATGTCCACTTGTAACACG 1020
CCAAGATCTTTTATCCTTAATCAGTTCACAAGTCCAAATATGAGTTTTTTTGTTTACTAG 1080
ATTCACTGCATGTGTGTTGCACGTGTATGCCTAGTTTTGAGATTCATATAATCGTCCAAC 1140
TGTATCCATACTCGTGTTTAGAAAATTAGAGCACAATGATTCCAGATATACATATGCAGA 1200
TAAAAGAAAGCTACATCTGGTTCAATGGCAAAAGGTAATGATTGATAAGGAAGGAGGCTT 1260
CATAGGGTTACTCCTCCTTTAATTAAAAATCTCGAGTTTGAACTTGAGAATAAAGAAACT 1320
ATTGACTAGCTTACACTACATTCTCTACGCAGCTCGAATATAAGTTTGTGGAGTCAGTAT 1380
ATTTTTCATTATATACCAAATAATTTTAATAAATATGTATCCAACATTGAAAAATATATT 1440
ATAAATTGCATAATAAGATGGCAAAGTAAACTTGTCTTCGTGCAAAGCAGAAATGATTAG 1500
GCAACTTCCAACTCCAAGTCCATTTGTGACTTTGAAATTCTCTTCTTTTTTTCTAGAGAA 1560
AATGGTCTAAAACCCCTCCAATCTATACCCGAAATCTCAACTACACACTCCAACTTTACG 1620
GGTGTCCTATCATCCCCCTGAACTATTTAAAAATATAATAAATAACCCCTTAAAACGCAA 1680
TAACCACTCATGAGACTTGTGGTGAAACACACACCTGACACGTGTTTTTCACCAATTATT 1740
TTTTAAAATTATTTTTTCTCCTTCTTTTTTATACTTTCTTCCCCAATTACAACCCTAATT 1800
CTTTGTGCAAGTTTTGATTTTAAAAATATTGAGGGATTTGAATCTGACTTGGGTGATAAA 1860
AGGCTTGGGAGAATTA 1876
```

b)

```
GAATTCCTTAATTGTAATAGATGGCTAATTGAGACAAGTCTAAACTTCCATTTTCCTTTA 60
TAAAAATTTTGGACTTGAGAAATGCACTTTTACTCTTTATTGTTCTTTTATCTTATTTTA 120
TGATATACTAATGGCTTGTATGAGACCCCTTCGGGGTTGAGTACGCTGTGTTACATCCAG 180
GGGATACCCCTGGGTCGTGACAAACACTTCCTATAAACAACCTACTAATAACCTCCACTA 240
CCACATTAGCCTTGCAGGGATGATAAAGAACACTCGTGTCATAATCCTTCAATAATTGTA 300
GCCACCTTCTATGGCAAAGATTCAAGTCCCTTTGACTAAACACATATTGCAGGCTCTTAT 360
TATCCGTAAACACATCAATATGGACCCATACTGGTAGTGCCTTTAAATCATTAAGGCAAA 420
AACAACCTCCTCTAATTCAAGGTCATCGGTAGGATAATTCTTCTCATGGGCTTTAAGTTT 480
CCTTGAAGCATAGTCAATCACCTTACTATTTTGTGAAGCACACATCACAACGCAACTCTT 540
GAGGCATCACAATACACTACAAAACCACGGTACTTTCCGATAAAGTCAATACCGGGGCGA 600
AAGTAAGTCTATCTTTCAACTCTTGGAAACTTTTCTCACAAGCTTCCGACAAAATGAACT 660
TATGCTTCTTTTGAGTTAAAACCATCATTGGAGAGGCAATGGAAGAAAAACCATTAACAA 720
ATCTTCTATAATAACCAGCCAAACCCAAAAAGATTCTAATGTCCGAAGGAGTTAGAGGCC 780
TTGGCCAACTATTGACCGCATTTTTCTACTTACGATCAACCTCAATAACCTTACTAGACA 840
CAATATGACCGAGGAAATCCGTGGAACTTAACCAAAATTCACACTTGTTGTAACACTCGA 900
TAAACACTAGTAGACTAAGCTAGAGCCTAATGTGTAAACTAGTAACCTAACTGGGGATTA 960
ACATAAAGGTTATTGTCGTAAATAACCCCCGGTACTCAATATAGGTGTTTTTTGGTTTAG 1020
ACGTCAAGGTAAGACCCCCAAGGACTAACTAAGGGTCCTATAGAAGGACCCTTAAAACAA 1080
CCTAAAAAGCTATCCAAAACAGTGGCCTCCCACGAAGGGCATCCACAAGCCGTTAATTGA 1140
CCCACGCCCTGTGGGTCAGGGTCCTGGGTCAAGTCCTGACCTTCCCCTGTTTTCTTAGCT 1200
TCATACTATCACCCATGACAGACCAGTATGGTCCGTGGGTGAGGGCTGTTTTAGGTCTTA 1260
AGTTAATGAGTTGGGGGTTAAGGGAAGGTCATTGGGTTAATTAATTATTTAGTTAGTGTC 1320
ATTTTAATACTTATTAGACTACCTATATAAACCTAATAGACCTCTAAACTAAGTCAATTA 1380
AGTCATTCTTTCTAGACATCAAAATAAACTTATTCCCTCTCAAATATTTCTCTCTCAAAA 1440
ACTTCATAGAAGAAGAAAAATGTGGAGTCTAGGGTTGAATGGGCTGTCTTTTTCTCTAAA 1500
TAACTTGAGGATTCTTCATCAAGGCATGGTAGCTTTTCATCTATGGGTAGCTTTCACTCA 1560
TAGAGTCCCTTCAAAACTCAATTTCACAACTTGGATATCCCCCAAAAGCAAGGGTTTCAC 1620
TCCAAGTCATGGATTGCTTTTCAAAAATAATTTCAATGATTTAATTTCTTTATATTATGA 1680
TCTTATACATATTTTCATTAAAATTTCTTTTTAATTTACTTATTGAAAGCAAAAATGCCT 1740
AAGATGAAACATTCATCTTATTGATTTCATGTATATGAGTGTCTATGTCAATTTTTTGTG 1800
GGACTGAAAATGTTATTGTCTCTTCCTTCTAGGCCAATATAGTGAATTTTAAAGCTTTAT 1860
TGATAGAAAATTTAGTGATTGAAAGTTTAGTAATTTAAGTCATTCTAACTATTTTTCGTT -1920
TTGAATGAATTGTTGTCACTTTTCACATTTTAAATGAATTGTTTCTTTTATTTTTGTGTA 1980
TAGTTAATAACTGAATAACCGAATCAAACCAAACCGAAACCGATAATCACCAAATCGATA 2040
AATGTATATTATATTTGGTGATTTGGTTTTGATAATTTTAAAACCGATTAAGTTGATTTG 2100
GTTTTGATTTTGACCAATAACCGATCCAAACCAACCCGTGAACACCCCTAACCAAATGTA 2160
TAAAAAAAAAGTAATAAATTCACCAAAACCTATTTCATTCCTTATGAACAAATAAATTCC 2220
AACCATTAATGAGCTTTTAAGAAGCAATGTTGGGAGCAAAAGCAGAGGAAGAAGACATGC 2280
CAGCCTTATGTTCATGCACAAAAATTAGAACAATATTGCTATTGAAATAATCCAATAACT 2340
TCTTCCCCAAAAGTCACTTGTATAGATGTATTCACTAGTTCTTCAAATTCGGCGAAAAGG 2400
CAGTACGAGGTGGGTCAAATTCCACCAGAAAACAACAAGAAGAAAGAAACTAAAAAGTTC 2460
ATGCTGCTGCAA 2472
```

Appendix D. Nucleotide sequence of *Eco*RI/*Sal*I subclone pap7-500.

GTCGACTCTAGGCCTAAATGGCCATTTAGGTGACACTATA 40
GAAGAGCTCGAGGATCAACTAGTCTATGCAAAAGGATTCA 80
TCTAAAATGGATGACCCTTTTCTACCCATGGTGGCTAAAT 120
GATTTATGGAGACTTGAGTTAATGTGAACTCGCATCCCTA 160
TATCGGTGCTCAATACTACTCCCAAAAATATAATATTAGC 200
TCTTATGTTTAAAACATACTTCTTTTGTGATTTGAGATTA 240
TTGCTCAAAAACTTAGCTCAAAGGCTATCTTGGAAATCAA 280
AGTCTTCCCGTCTTGCTTAATTATGAAAGCATTTACTCTT 320
TTTTGAAAACTAGCCCGAAGGCTCTTTGGAAATCGAAGTT 360
TCCTTCCTTGTTTAAATGTGAAAACATTTTATACTCTTTG 400
GGAATACATAGTCCCCGTATGCTTTTGAAGAAATGAACTT 440
CAAGCTTTACTCTTTACTGAATTC 464

Appendix E. 5' to 3' sequences of oligonucleotides used as sequencing and PCR primers.

Primer Name	Template	Sequence
PCR Primers:		
pap7p1-upper	*pap7* genomic clone	tggcatcgtaaccaagacagtgaac
pap7p1-lower	*pap7* genomic clone	tatggcattggacagaagtaacctgtaatac
StPTupper	potato genomic DNA	ttttaggcattatggctgtggagg
StPTlower	potato genomic DNA	agtgagggcagctggcaacg
potato actin upper	potato genomic DNA	actcaccgaagcacctctcaatcc
potato actin lower	potato genomic DNA	ccatcaggcaactcgtaggtcttc
Sequencing Primers:		
T3	pBluescript	aattaaccctcactaaaggg
T7	pBluescript	gtaatacgactcactatagggc
M13 forward	pBluescript	gtaaaacgacggccagt
M13 reverse	pBluescript	ggaaacagctatgaccatg
7-6F2	pap7-6	tcgttctgatctcatttttattta
7-6F3	pap7-6	aagaccgtgtaagaggaaaatgtg
7-6F4	pap7-6	accttccaaaatattagccaaca
7-6R2	pap7-6	tctcatctaatagatcctcctttg
7-6R3	pap7-6	tcagcattaatcttccaaacaca
7-6R4	pap7-6	accaaaagaaggacaactactcg
7-6R5	pap7-6	gtccgggcgaagatttta
7-7F2	pap7-7	gtctgaatctggcaataactact
7-7R2	pap7-7	gatacatggcaaacaaaataaag
7-8F2	pap7-8	catgccaaaaagggacatt
7-8F3	pap7-8	gccttaaccgacatagaccat
7-8R2	pap7-8	tggaaatacatctaactcatcttg
7-8R3	pap7-8	gtaatggacaatgcttcaatcaga
7-8R4	pap7-8	tgaccccatgagcacaaaata
7-8R5	pap7-8	aaatgatttggctagtattgaagag
7-9R2	pap7-9	tgtatgtcctttttgttctaccattta
7-9R3	pap7-9	acttttgcatacccagaactcg
7-9R4	pap7-9	tatttagcattctggcacattct
7-10T72	pap7-10	atcggtaggataattcttctcat
7-10T73	pap7-10	cccggtactccatataggtgttt
7-10T74	pap7-10	aaaaatgtggagtctagggttga

Primer Name	Template	Sequence
7-10T75	pap7-10	cgaatcaaaccaaaccgaaac
7-10T32	pap7-10	aagctgctcagactattatggtg
7-10T33	pap7-10	gctgatgtgttgtttgtaataaga
7-10T34	pap7-10	tgtggagtcagtatatttttcatt
7-11T32	pap7-11	gccaaaagtatcccatctca
7-11T33	pap7-11	tatacctttagtcgtttgattgtg
7-11R1	pap7-11	atgacaacgtgagatgggatac