

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Characterization of an AtPAP26-like
protein (TrPAP26) from white clover
(*Trifolium repens* L.)**

A thesis presented in partial fulfillment of the requirements for the degree of

Master of Science in Plant Biology

At Massey University, Palmerston North, New Zealand

Jennifer Y. Huang

2013

Abstract

Phosphate levels in soils are often in deficit in New Zealand agriculture systems, resulting in the need for phosphate supplements in the form of fertilizers. Plants are able to adapt to many environmental stresses and display a wide range of responses designed to cope with phosphate-deficiency, and the study of these may lead to the production of crop and pasture plants that can utilize added P more efficiently. One adaptive mechanism is to express purple acid phosphatase (PAP) genes, the protein products of which are able to generate, transport, and recycle inorganic phosphates from phosphate-rich compounds both intracellularly and extracellularly. Their general mechanism of action is to hydrolyze phosphate-rich esters that are found within cells, the cell wall or in the rhizosphere. One PAP, *AtPAP26*, has been extensively characterized in *Arabidopsis thaliana* and displays high levels of acid phosphatase activity during phosphate-starvation. *AtPAP26* has been found to be the predominantly expressed PAP during phosphate-starvation and the enzyme plays a key role in supplying inorganic phosphate to the plant by hydrolyzing the organic phosphates present in the rhizosphere. An *AtPAP26*-like sequence has been identified previously in white clover and so this project firstly cloned the full-length *TrPAP26* and then examined expression in response to phosphate-starvation. The protein product (TrPAP26) was also characterized and compared to *AtPAP26* in terms of its putative biochemical functions.

TrPAP26 was predicted to be a 55 kDa protein with three N-glycosylation sites, a signal peptide of 21 amino acid residues, and a metal-ligating motif typical of PAPs. Its observed mass was closer to 45 kDa, and preliminary experiments, using recombinant TrPAP26 partially purified from transgenic tobacco, suggested that it hydrolyzed a wide range of phosphate-rich esters including adenosine triphosphate (ATP), phosphoenolpyruvate (PEP), and pyrophosphate (PPi), but not inositol hexakisphosphate (phytate). *TrPAP26* transcript levels were found to be constitutive in the roots of white clover, but correlated positively with phosphate supply in other tissues. The protein and activity levels were not directly correlated with the transcript levels suggesting other methods of regulation such as post-

translational modifications, including N-glycosylation. TrPAP26 accumulated more in the mature leaves of white clover plants grown with a full supply of phosphates. Taken together, these results suggest that TrPAP26 may play a role in internal P remobilization, rather than P scavenging directly.

Acknowledgements

I gratefully acknowledge the following people for all their help, without which I could not have completed this project:

- My supervisor Dr. Michael T. McManus for his excellent supervision throughout the course of this project,
- Susanna Leung for all her technical help,
- The New Zealand Pastoral and Agricultural Research Institute for providing a research grant, and
- All my fellow members of lab C 5.19 for their camaraderie.

Table of contents

Abstract	i
Acknowledgments	iii
List of figures	ix
List of tables	xii
Abbreviations	xiii
I. Introduction	1
1.1 Overview.....	1
1.2 P in soils and agriculture.....	2
1.3 Pi uptake mechanisms by plants.....	4
1.4 Responses of plants to low Pi.....	5
1.5 Acid phosphatases (APases).....	9
1.6 Purple acid phosphatases (PAPs)	10
1.7 Genetic strategies to improve phosphate use efficiency	13
1.8 White clover.....	13
1.9 White clover and PUE.....	14
1.10 Project and aims.....	16
2. Materials and methods	17
2.1 Plant material.....	17
2.1.1 Maintenance of <i>Trifolium repens</i> (white clover) stock plants..	17
2.1.2 Excision and treatment of white clover cuttings	17
2.2 Chemicals.....	18
2.3 Biochemical methods.....	22
2.3.1 Protein extraction.....	22
2.3.1.1 Soluble protein and cell wall protein extraction.....	22
2.3.1.2 Glutathione S-transferase GST protein purification.....	22

2.3.1.3 Ammonium sulfate precipitation.....	23
2.3.1.4 Dialysis.....	24
2.3.1.5 Immobilized metal ion affinity chromatography.....	24
2.3.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).....	25
2.3.3 Coomassie brilliant blue staining.....	26
2.3.4 Western analysis.....	27
2.3.5 Enzyme assays.....	30
2.3.5.1 Protein quantification.....	30
2.3.5.2 BSA standard curve.....	30
2.3.5.3 pNP standard curve	31
2.3.5.4 Inorganic phosphate standard curve	31
2.3.5.5 Acid phosphatase activity assay A	34
2.3.5.6 Modification to acid phosphatase activity assay A...	34
2.3.5.7 Acid phosphatase activity assay B	35
2.3.5.8 Root surface acid phosphatase activity assay.....	36
2.4 Molecular methods.....	37
2.4.1 Genomic DNA isolation.....	37
2.4.2 RNA isolation.....	37
2.4.2.1 Generation of cDNA.....	37
2.4.2.1.1 DNase treatment of RNA.....	37
2.4.2.1.2 Reverse transcription.....	38
2.4.3 Quantification and purity assessment of nucleic acid.....	39
2.4.4 Primer design.....	39
2.4.5 Polymerase chain reaction (PCR) amplification.....	41
2.4.5.1 Semi-quantitative reverse-transcriptase PCR (sqPCR).....	42
2.4.5.2 Quantitative RT-PCR (qPCR).....	42
2.4.6 Agarose gel electrophoresis.....	42
2.4.7 Gel slice extraction and purification.....	43
2.4.8 DNA sequencing.....	44

2.4.9	Vector ligation.....	45
2.4.9.1	cDNA preparation.....	45
2.4.9.2	Vector preparation.....	46
2.4.9.3	Ligation.....	47
2.4.10	Transformation of bacterial cells.....	48
2.4.10.1	Preparation of competent <i>E. coli</i> cells.....	48
2.4.10.2	Transformation of <i>E. coli</i> with pGEM-T vector.....	49
2.4.10.3	Plasmid DNA purification.....	50
2.4.10.4	Transformation of <i>E. coli</i> with pGEX-6P-3 vector...	50
2.4.10.5	Transformation of <i>Agrobacterium tumefaciens</i> with pART27 vector.....	51
2.4.10.6	<i>A. tumefaciens</i> plasmid miniprep.....	52
2.4.11	Transformation of <i>Nicotiana tabacum</i> with <i>A. tumefaciens</i> plasmid.....	53
2.5	Bioinformatic techniques.....	55
2.5.1	Basic local alignment search tool (BLAST).....	55
2.5.2	Clustal omega.....	56
2.6	Statistical analysis.....	56
3.	Results.....	57
3.1	Isolation of a <i>PAP26-like</i> gene from white clover.....	57
3.1.1	Obtaining the full-length <i>PAP26-like</i> cDNA sequence.....	57
3.1.2	Bioinformatic analysis.....	64
3.2	Biochemical characterization of recombinant TrPAP26.....	67
3.2.1	Protein expression with the construct <i>TrPAP26::GST</i>	67
3.2.2	Protein expression with the construct Δ sp- <i>TrPAP26::GST</i>	73
3.2.3	Identification of recombinant Δ sp- <i>TrPAP26::GST</i> using MALDI- TOF/TOF.....	77
3.3	Biochemical characterization of purified TrPAP26.....	80
3.3.1	Transforming tobacco with the construct <i>35S::TrPAP26::6xHis</i>	80

3.3.2	Integration of <i>35S::TrPAP26::6xHis</i> and expression of TrPAP26::6xHis in tobacco.....	81
3.3.3	Assessment of TrPAP26::6xHis accumulated in transgenic tobacco lines.....	85
3.3.4	Affinity-purification of TrPAP26::6xHis protein from tobacco..	88
3.3.5	Activity of partially purified TrPAP26::6xHis.....	89
3.4	Expression of <i>TrPAP26</i> and accumulation of TrPAP26 in response to changes in P supply <i>in planta</i>	93
3.4.1	White clover grown in short-term phosphate-deficient (-Pi) conditions.....	93
3.4.2	White clover grown in long-term phosphate-deficient (-Pi) conditions.....	94
4.	Discussion.....	100
4.1	Overview.....	100
4.2	Isolation of an <i>AtPAP26-like</i> gene sequence from <i>Trifolium repens</i> (white clover)	101
4.3	Recombinant TrPAP26 expressed in <i>E. coli</i>	102
4.3.1	Protein expression with the construct <i>TrPAP26::GST</i>	103
4.3.2	Protein expression with the construct Δsp - <i>TrPAP26::GST</i>	104
4.4	Recombinant TrPAP26 expressed in tobacco.....	107
4.4.1	<i>35S::TrPAP26::6xHis</i> transcript levels.....	107
4.4.2	TrPAP26::6xHis accumulation.....	108
4.4.3	TrPAP26::6xHis affinity-purification.....	109
4.4.4	TrPAP26::6xHis activity.....	110
4.5	Expression of <i>TrPAP26</i> and accumulation of TrPAP26 in response to changes in Pi supply <i>in planta</i>	111
4.5.1	<i>TrPAP26</i> transcript levels.....	112
4.5.2	Acid phosphatase activity in white clover tissues exposed to the +Pi/-Pi treatments.....	114
4.5.3	TrPAP26 protein levels.....	115
4.5.4	Possible physiological role of TrPAP26.....	117

4.6 Summary.....	118
4.7 Future work.....	119
5. Appendices.....	121
5.1 Nucleotide sequences of a partial EST (897 bp) of an <i>AtPAP26</i> -like gene gifted by the New Zealand Pastoral Agricultural Research Institute (Palmerston North, NZ) (Section 3.1.1).....	121
5.2 Consensus nucleotide sequence from 5 positive colonies harbouring the pGEM vector with the 1 kb fragment of the putative <i>Tr-PAP26</i> sequence (Section 3.1.1).....	121
5.3 Consensus nucleotide sequence from 10 colonies harbouring the pGEM vector with the 0.4 kb fragment of the putative <i>Tr-PAP26</i> sequence (Section 3.1.1).....	122
5.4.1 Full list of BLAST hits of the 0.4 kb fragment of the putative <i>Tr-PAP26</i> sequence (Section 3.1.1).....	123
5.4.2 BLAST hits of the 1 kb fragment of the putative <i>Tr-PAP26</i> sequence (Section 3.1.1).....	127
5.5 Full contig nucleotide sequence of the partial EST sequence obtained from AgResearch (Appendix 5.1) and 1 kb fragment of the putative <i>Tr-PAP26</i> sequence Appendix 5.2) (Section 3.1.1).....	132
5.6 Full-length nucleotide sequence of <i>TrPAP26</i> with putative start codon and stop codon bolded (section 3.1.1).....	133
5.7 BLAST matches after MALDI-TOF/TOF for the ~72 kDa putative recombinant Δ sp-TrPAP26::GST fusion protein (Fig. 3.8B, lane 11) (section 3.2.3).....	134
5.8 BLAST match after MALDI-TOF/TOF for the ~55 kDa unknown recombinant protein (Fig. 3.8B, lanes 11 and 12) (section 3.2.3).....	134
5.9 BLAST match after MALDI-TOF/TOF for the ~50 kDa putative recombinant Δ sp-TrPAP26 protein (Fig. 3.8B, lane 12) (section 3.2.3).....	134
5.10 Primer design and rationale (section 2.4.4 and table 2.1).....	135
5.11 qPCR protocol (section 2.4.5.2).....	137
6. References.....	138

List of figures

Figure 1.1	Schematic representation of the major components of the soil-plant P cycle in agricultural systems.....	3
Figure 1.2	A model suggesting various adaptive metabolic processes (indicated by asterisks) that are believed to help plants acclimate to Pi-deficiency.....	8
Figure 1.3	A classification scheme for <i>Arabidopsis</i> PAPs based on clustering analysis of amino acid sequences.....	12
Figure 2.1	Drawing of a main stolon (MS) of white clover, showing axillary buds (AB), lateral branches (LB), and a lateral stolon (LS).....	19
Figure 2.2	Four-nodal stolon cutting prior to planting in vermiculite to induce roots at nodes 3 and 4 (arrowed)..	20
Figure 2.3	A. Example of white clover stolon cutting grown in -Pi media for 21 days. B. Schematic of the stolon cutting to indicate the harvested tissues.....	21
Figure 2.4	A diagrammatic set up for the electrophoretic transfer of proteins from an acrylamide gel to a PVDF membrane.....	29
Figure 2.5	BSA standard curve used for protein quantification.....	32
Figure 2.6	pNP standard curve used for measuring acid phosphatase activity	33
Figure 2.7	Inorganic phosphate standard curve used for measuring acid phosphatase activity.....	23
Figure 3.1	PCR products separated using 1% (w/v) agarose gel electrophoresis.....	59
Figure 3.2	Separation using 1% (w/v) agarose gel electrophoresis of a PCR product of <i>T. repens</i> cDNA amplified using the gene-specific forward and reverse primers	59
Figure 3.3	Clustal multiple sequence alignment of the 0.4 kb and 1 kb fragment.....	62

Figure 3.4	The translated sequence of <i>TrPAP26</i> from the putative initiating methionine (M) to the stop codon (*).	65
Figure 3.5	Coomassie blue stain of the induced expression of TrPAP26::GST in <i>E. coli</i> after separation using 12% SDS-PAGE.	70
Figure 3.6	Coomassie blue stain of the induced (+IPTG) or non-induced (-IPTG) expression of TrPAP26::GST in <i>E. coli</i> strain BL21 after separation using 12% (w/v) SDS-PAGE	71
Figure 3.7	Separation of TrPAP26::GST expressed in <i>E. coli</i> origami using 12% SDS-PAGE followed by western blotting using the anti-AtPAP26 antibody (A) or anti-AtPAP12 antibody (B).	72
Figure 3.8	Acid phosphatase assay on TrPAP26::GST, TrPAP26, and a crude plant extract activity, as indicated.	73
Figure 3.9	Separation of Δ sp-TrPAP26::GST expressed in <i>E. coli</i> origami using 12% SDS-PAGE followed by western blotting using the anti-AtPAP26 antibody (A) or anti-AtPAP12 antibody (B).	76
Figure 3.10	Alignment of the tryptic peptide sequences generated from the 72 kDa band (Fig. 3.8B, lane 11).	78
Figure 3.11	Alignment of the tryptic peptide sequences generated from the 57 kDa band (Fig. 3.8B, lanes 11 and 12).	78
Figure 3.12	Alignment of the tryptic peptide sequences generated from the 45 kDa band (Fig. 3.8B, lane 12).	79
Figure 3.13	Acid phosphatase assay of Δ sp-TrPAP26::GST and Δ sp-TrPAP26 as glutathione-sepharose purified proteins, as indicated	79
Figure 3.14	Development of tobacco leaf pieces 21 days after inoculation with <i>Agrobacterium tumefaciens</i> .	82
Figure 3.15	Separation of PCR products amplified from the genomic DNA isolated from putative transgene lines or control lines of tobacco.	83
Figure 3.16	Separation of PCR products amplified from the cDNA isolated from putative transgene lines or control lines of tobacco.	84
Figure 3.17	Analysis of <i>TrPAP26</i> expression using semi-quantitative PCR.	84

Figure 3.18	Acid phosphatase determined in the wall fraction (A) or soluble fraction (B) in the transgenic and control (C2-C10) tobacco lines...	86
Figure 3.19	Separation of soluble (top panel) or cell wall (bottom panel) protein extracts from putative transgenic plants (lanes 1-8) or control plants (C6, C8, C10) using 12% (w/v) SDS-PAGE, following by western blotting using the anti-AtPAP12 antibody.....	87
Figure 3.20	Acid phosphatase activity assay of the soluble protein extract and affinity-purified protein.....	90
Figure 3.21	Separated proteins from transgenic plant line 8 or control plant line C8, as indicated, that were either bound to the chelating sepharose or remained unbound in the supernatant	90
Figure 3.22	Separation of protein extracts from transgenic plant line 8 and control plant line C8, as indicated, using 12% (w/v) SDS-PAGE and western blotting using the anti-AtPAP12 antibody	91
Figure 3.23	Acid phosphatase activity assay with or without MgCl ₂	91
Figure 3.24	<i>TrPAP26</i> expression in the roots of plants grown in +Pi or -Pi media, as indicated, for 24 hours.....	94
Figure 3.25	<i>TrPAP26</i> expression in the roots of plants grown in +Pi or -Pi media, as indicated, for 7 days.....	94
Figure 3.26	Acid phosphatase activity of whole elongation zone roots of plants grown in +Pi or -Pi media, as indicated, for 16 or 18 days.....	97
Figure 3.27	qPCR of <i>TrPAP26</i> expression in various tissues from white clover grown in +Pi or -Pi for 21 days.....	97
Figure 3.28	Acid phosphatase activity assay of soluble (A) and cell wall protein extracts (B) from the tissues, as indicated, and subjected to +Pi or -Pi treatment, as indicated.....	98
Figure 3.29	Separation of protein extracts from white clover grown in +Pi or -Pi media, as indicated by "+" or "-", respectively, using 12% (w/v) SDS-PAGE and western blotting using the anti-AtPAP12 antibody.....	99

List of tables

Table 2.1	List of custom-made primers used.....	40
Table 2.2	List of pre-made primers used.....	40
Table 3.1	BLAST alignments against the consensus sequence of the 0.4 kb fragment generated using the “TrPAP26 F1” and “TrPAP26 F2” primers.....	60
Table 3.2	BLAST alignments against the consensus sequence of the 1 kb fragment generated using the “TrPAP26 F1” and “TrPAP26 F2” primers.....	61
Table 3.3	Predicted features of the translated TrPAP26 sequence.....	66
Table 3.4	Substrate specificity assay using the affinity-purified protein from transgenic line 8 and control line C8.....	92

Abbreviations

-Pi	inorganic phosphate-deficient
+Pi	inorganic phosphate-sufficient
6xHis	six histidine residues
amp	ampicillin
APS	ammonium persulfate
ATP	adenosine triphosphate
AtPAP12	<i>Arabidopsis thaliana</i> purple acid phosphatase 12
AtPAP26	<i>Arabidopsis thaliana</i> purple acid phosphatase 26
BAP	6-benzyl aminopurine
BCIP	5-bromo-4-chloro-3'-indolyphosphate
BLAST	basic local alignment search tool
bp	base pairs
BSA	bovine serum albumen
cDNA	complementary deoxyribonucleic acid
cm	centimetre
Da	Dalton
DEPC	diethylpyrocarbonate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EST	expressed sequence tag
FFE	first fully expanded leaf
g	gram
<i>g</i>	gravity
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GC	guanidine and cytosine
GDP	gross domestic product
GS4B	glutathione sepharose 4B

GST	glutathione S-transferase
GUS	β -glucuronidase
ha	hectare
I	internodes
IPTG	isopropyl thiogalactoside
kan	kanamycin
kb	kilobase pairs
kDa	kiloDalton
kg	kilogram
KOH	potassium hydroxide
L	litre
LB	Luria-Bertani
M	molar
m	metre
MALDI-TOF/TOF	mass spectrometer for amino acid sequence of peptides
mg	milligram
min	minute
mL	millilitre
mM	millimolar
mRNA	messenger ribonucleic acid
NAA	1-naphthaleneacetic acid
ng	nanogram
nm	nanometre
NZ	New Zealand
P	phosphorus
PAGE	polyacrylamide gel electrophoresis
PAP	purple acid phosphatase
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEP	phosphoenolpyruvate
Pi	inorganic phosphate

pNPP	<i>p</i> -nitrophenol phosphate
pNP	<i>p</i> -nitrophenol
Po	organic phosphate
PUE	phosphate use efficiency
PVDF	polyvinylidene difluoride
qPCR	quantitative polymerase chain reaction
R	roots
RACE	rapid amplification of cDNA ends
rpm	rotations per minute
SDS	sodium dodecyl sulfate
SEM	standard error of the mean
spec	spectinomycin
TAE	Tris-acetate-EDTA
TEMED	tetramethylethylenediamine
T _m	melting temperature
TrPAP26	<i>Trifolium repens</i> purple acid phosphatase 26
uL	microlitre
ug	microgram
V	volt
v/v	volume by volume
w/v	weight by volume
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside

1. Introduction

1.1 Overview

Inorganic phosphate (Pi) is an essential element for plant growth and development, making its availability a necessity for productive farming (Richardson, 2009). Worldwide soil concentrations of phosphates (Pi) are in deficit and are lower than what is required for optimum plant growth. In addition to this problem of extremely low levels in soils, most of the Pi that is present in soils is in a form not usable by plants. Both aluminum and iron ions, present in significant amounts in soil, will bind strongly with P to create aluminum-phosphate (AlPO_4) and iron-phosphate (FePO_4), and metal-bound P is in a form that is unavailable to be taken up by roots.

To accommodate for this paucity of available soil Pi, the nutrient must be applied in the form of fertilizers. Because Pi will bind easily to aluminum and iron, an excess of fertilizer is generally applied to ensure a good crop yield. However, most (up to 80%) of the applied Pi will become metal-bound, and any that is not taken up by plants leaches out of the soil in runoff, and can cause eutrophication in nearby bodies of water (Raghothama, 1999). Thus the over-application of Pi-based fertilizer is clearly unsustainable and harmful to the environment.

In response to this urgent problem, many researchers are trying to produce plants that have improved phosphate use efficiency (PUE) (USGS, 2000). There are breeding programs in place to try to screen for germplasm with better PUE in various crop plants. In addition to traditional breeding, there are also molecular approaches being used. These molecular approaches target genes involved in the uptake and use of phosphates, including phosphate transporters, transcription factors, and acid phosphatases, and over-expression studies have shown that many of these gene candidates can confer some increased PUE (Richardson,

2009; Yuan and Liu, 2008).

White clover (*Trifolium repens* L.) is an agronomically important plant that is widely used in pastures in New Zealand and underpins New Zealand's pastoral industries. This project aims to characterize a white clover purple acid phosphatase enzyme (TrPAP26), the utilization of which may improve its PUE under phosphate-deficient conditions. Ultimately, such knowledge may provide a breeding trait for a constitutive increase in PUE in this important species.

1.2 P in soils and agriculture

Inorganic phosphate (Pi) is generally a mixture of HPO_4^{2-} and H_2PO_4^- ions. In most soils worldwide, the concentration of Pi is very low (less than 10 μM), which is far lower than what is generally found in plant tissues (5-20 mM) (Raghothama, 1999). In addition to the naturally low levels of phosphorus (P) in soils, much of the soil P present (20-80%) is organic and must be made inorganic before becoming bioavailable (Fig. 1.1) (Richardson et al., 2009). This conversion is achieved through acid phosphatases secreted by plant roots or by the mycorrhizae associated with those roots or by a population of soil bacteria in this rhizosphere. However, to remove Pi as the limiting nutrient in plant growth, farmers and growers often apply Pi-based fertilizers.

Phosphate rock, a nonrenewable resource, is the most significant source of Pi used in fertilizers. At the current mining rates, all of the readily accessible reserves of phosphate rock is predicted to be depleted very soon--within the century by most estimates (Filippelli, 2011; Jasinski, 2009). If this is allowed to occur, agricultural industries around the world will decrease in production rates, resulting in increased food prices and, in the worst-case scenario, a food deficit because production is unable to support the world's growing population.

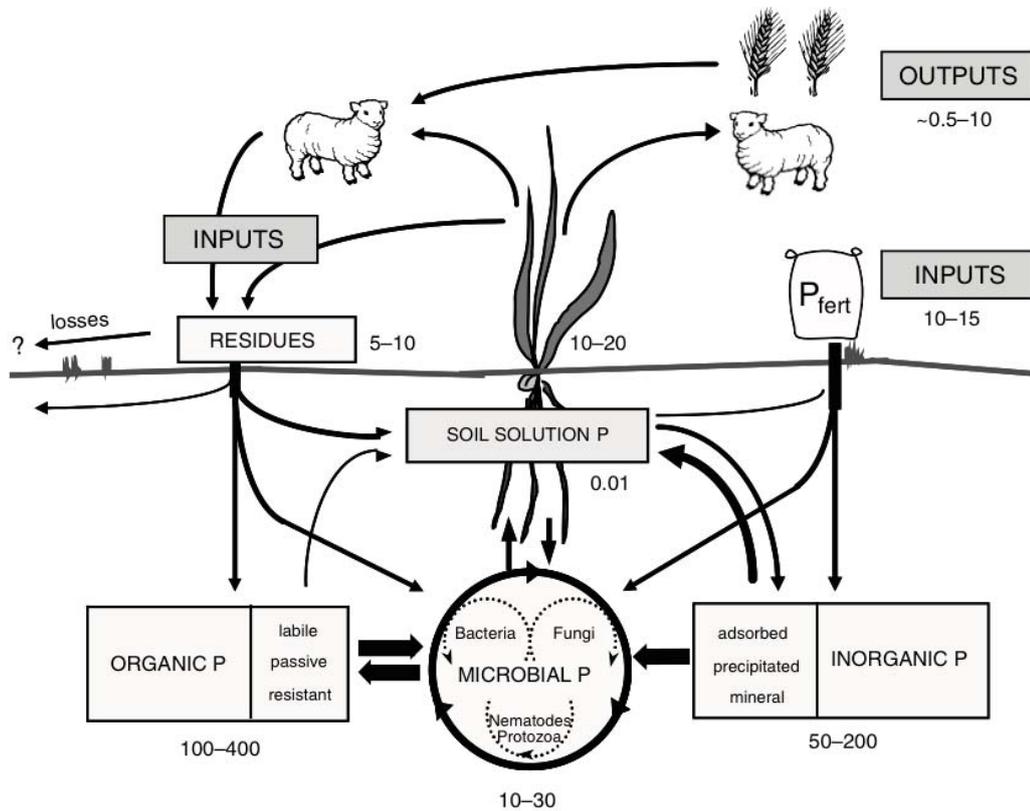


Figure 1.1. Schematic representation of the major components of the soil-plant P cycle in agricultural systems. Major P fractions and the amounts of P typically associated with each component (kg/ha, 0-0.10 m) are indicated. Major P transformation processes and the passage of P between components are indicated by arrows. Redrawn from Jakobsen et al. (2005).

1.3 Pi uptake mechanisms by plants

To uptake Pi, the ion is transported against a concentration gradient by roots, so plants require active transport processes to take in sufficient Pi. The epidermis, cortex, and associated intercellular spaces of the root make up the apoplast. It is here that the Pi is taken up from the rhizosphere. In younger roots, the apoplast contains interlaced fibers that filter the soil. Carboxyl groups present in the fibers create a negative charge in the apoplast. Various root excretions increase the negative charge to further modify the flow of anions such as Pi (Miller et al., 2001; Raghothama, 1999). The effect is to restrict the movement of Pi within the apoplast so it can be concentrated there and subsequently taken up through the plasmalemma, the semi-permeable boundary between the apoplast and symplast (Raghothama, 1999; Smith, 2002).

Embedded in the plasmalemma are high affinity phosphate transporters. Two families of these transporters have been identified in the model plant *Arabidopsis*, *Pht1* and *Pht2*. They function as cotransporters of H_2PO_4^- or HPO_4^{2-} and protons and rely on an ATP-driven proton pump to maintain a high proton gradient across the membrane (Yuan and Liu, 2008). The rate at which Pi is transported is determined by the number of transporter proteins embedded in the membrane and by the rate of protein activity. The number of transport proteins is regulated by negative feedback by Pi itself. That is, accumulation of Pi within the root will inhibit the transcription and translation of *Pht1* and *Pht2*. The opposite is also true: low amounts of Pi within the root will induce transcription and translation of the transporter protein-encoding genes (Shin et al., 2004).

Studies on the over-expression of *Arabidopsis* high affinity phosphate transporters in tobacco have shown that in cell suspension cultures, Pi uptake was increased (101-203 nmol Pi per h per g fresh weight range in the control lines versus a range of 490-874 nmol Pi per h per g fresh weight in the transgenic lines) (Mitsukawa et al., 1997). However, in whole transgenic tobacco plants grown in culture or in soil

there was no change in Pi uptake (Rae et al., 2004). Therefore Pi uptake is not determined solely by the amount of transporter proteins. Further studies on what transcription factors determine the rate of Pi uptake may be more beneficial in developing varieties of plants that are able to utilize Pi more efficiently.

1.4 Responses of plants to low Pi

As the soil Pi content becomes limiting, roots will extend into areas of soil where there is more Pi available, and hairs will grow along the root axis to further increase the area of soil explored. Other adaptations include: high root/shoot ratio, increased root elongation rate, longer root length, and, in some species, the formation of proteoid roots (Tran et al., 2010a).

Further, the majority of plant species have evolved symbiotic relationships with mycorrhizal fungi, and arbuscular mycorrhizal fungi in particular. These fungi will colonize the apoplast of the roots to obtain sugars from the plant. In return, the plant benefits from the Pi and other nutrients that the long hyphae of the fungi are able to scavenge very effectively. The reliance on these fungi differs among plant species and is dependent on the soil Pi concentration. However, this beneficial relationship is often disrupted with the application of fertilizer and to soil tilling. While the fungi are able to aid greatly in soils where Pi levels are low, in soils where there is sufficient Pi, the fungal associations may deter plant growth by consuming plant sugars without conferring any benefits to the plant (Bolan, 1991).

Plants that do not form such mycorrhizal associations have evolved other adaptations to cope with low soil Pi concentrations. For example, some plant species will form proteoid or cluster roots (Johnson et al., 1996). This is a modified root system comprising of single clusters of densely packed lateral roots with a huge abundance of root hairs that are able to absorb Pi at a much faster rate when compared with normal roots due to the increase in surface area (Miller et al., 2001). The formation of these roots is regulated by the Pi status of the shoots

rather than roots. Representatives from many plant families will form proteoid roots, including the *Proteaceae*, *Cheonpodiaceae*, *Cruciferae*, *Cyperaceae*, and *Junaceae*. Such plants are often slow-growing shrubs and trees that live in soils that are severely Pi-deficient (Vance et al., 2003). It would be worthwhile to study what controls these root architecture modifications. If the changes are driven by the same factors that control biochemical adaptations to Pi-starvation, then it may be possible to engineer any plant species to form such root systems in soils at normal Pi concentrations and thus reduce the need for added P inputs.

Much research has been done to elucidate both the morphological and biochemical changes that occur in a plant undergoing Pi-starvation (Richardson, 2009; Richardson et al., 2009; Tran et al., 2010a; Yuan and Liu, 2008). The morphological changes are the modifications to root architecture as mentioned previously. The biochemical changes are numerous and affect the entire plant.

Under Pi-starvation, the transcriptome, proteome, and secretome of plants change dramatically to maintain homeostasis (Misson, 2005; Tran et al., 2010a). The changes are very precisely coordinated and are regulated spatially and temporally, differing greatly between the shoots and roots (Wu, 2003). In *Arabidopsis*, well over 1000 genes are affected, where 600-1800 genes are induced while 250-700 genes are repressed (Morcuende et al., 2007; Muller et al., 2004). In addition to the change in gene expression, in many plant species, carbohydrates, amino acids, and anthocyanins accumulate in the leaves, high affinity Pi transporters are induced, and glycolytic by-pass pathways are induced to exclude Pi as a substrate (fig. 1.2) (Duff et al., 1989; Wu, 2003). The glycolytic by-pass pathways are extremely useful during periods of Pi-starvation because they are able to conserve ATP and recycle Pi while still allowing essential functions such as maintaining vacuolar pH by using pyrophosphate (fig 1.2).

In *Arabidopsis*, those genes directly involved in Pi metabolism were found to revert to normal expression levels immediately (within 3 hours) upon Pi-resupply. The

resupply also leads to consumption of the accumulated carbohydrates and amino acids (Bozzo et al., 2004; Morcuende et al., 2007).

Various enzymes and compounds are secreted both into the intercellular space and into the rhizosphere to aid in Pi scavenging. Secreted enzymes include ribonucleases, phosphodiesterases, and acid phosphatases, all of which catalyze their substrates to release a free Pi group (Tran et al, 2010).

In *Arabidopsis*, an important post-transcriptional mechanism is the inorganic Pi signaling pathway consisting of PHR1, *miR399*, PHO2, and *At4*. During Pi-sufficiency, PHO2 represses genes that are induced during Pi-starvation. During Pi-starvation, PHR1 is induced, which then induces expression of *miR399*, which binds and destroys the mRNA of PHO2, leading to the expression of Pi-starvation inducible genes. After prolonged periods of Pi-starvation, *At4* is induced, which inhibits *miR399*, allowing once again PHO2 to be expressed (Bari et al., 2006). Future studies should focus on elucidating the mechanisms and substrates of this pathway in other plant species as it may provide new targets for developing plants that are more Pi use efficient.

Other biochemical changes that have been observed to occur include post-translational modifications to proteins involved in Pi transport, scavenging, and recycling. Phosphorylation and glycosylation regulate the activity and target of various proteins, including acid phosphatases, mentioned previously, which play a role in the intra- and intercellular metabolism of Pi and in scavenging Pi from the rhizosphere (Lung et al., 2008; Zimmermann et al., 2004).

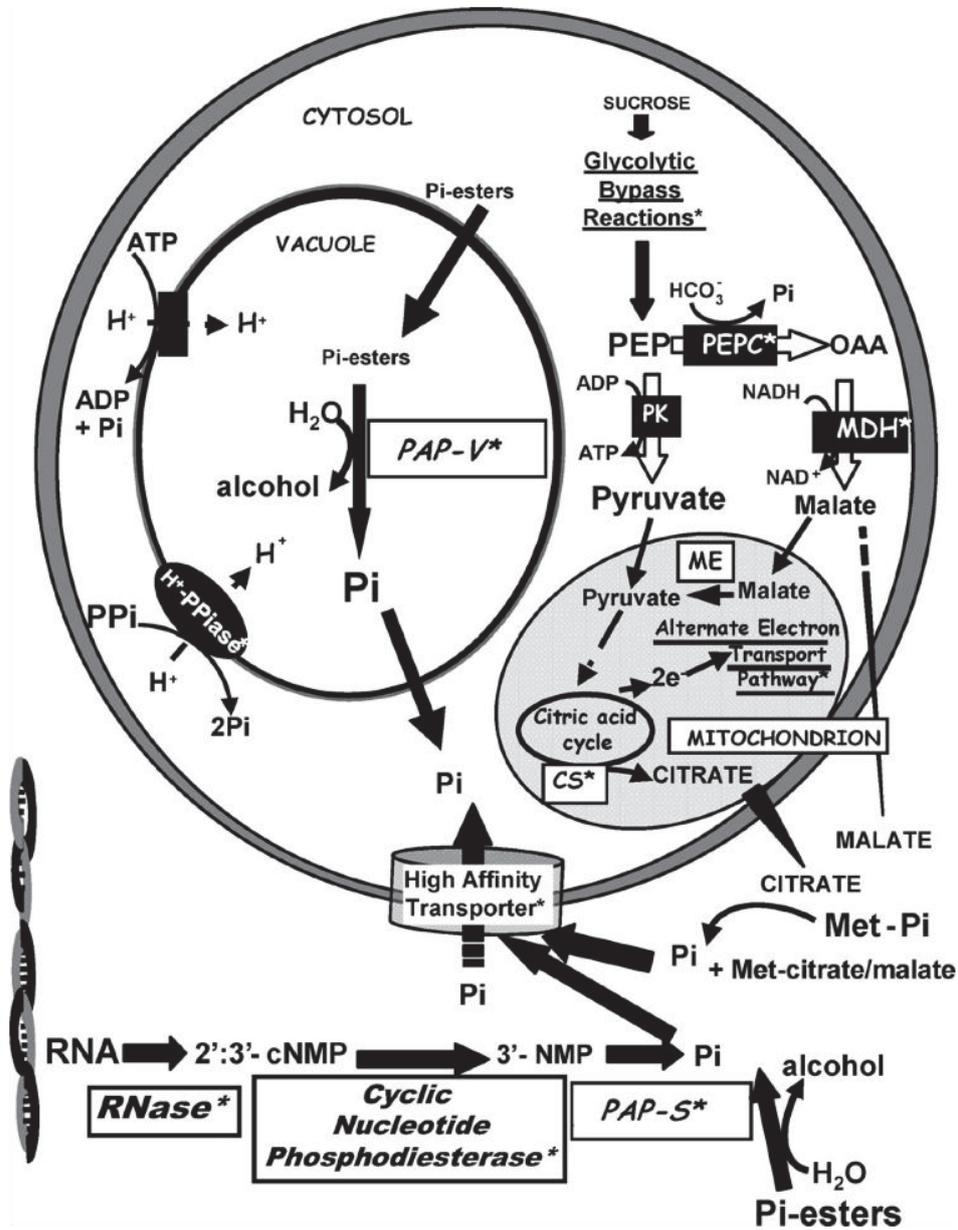


Fig. 1.2. A model suggesting various adaptive metabolic processes (indicated by asterisks) that are believed to help plants acclimate to Pi-deficiency. (Redrawn from Tran et al, 2010).

1.5 Acid phosphatases (APases)

APases are enzymes that catalyze the hydrolysis of orthophosphate monoesters to release Pi under acidic conditions (Duff et al., 1994). Their activity is a biochemical marker for Pi-deficiency (Vance et al., 2003). This class of enzymes encompasses a large number of proteins that are heterogenous in native molecular mass, subunit structure, and pI. In addition to having varying physical and kinetic properties, there are many tissue- and cellular compartment-specific isozymes (Tran et al., 2010a). Some generalizations can be made, however.

APases are ubiquitous enzymes that have been found in all higher organisms and tissues examined, and they have been detected throughout plant development (Vance et al., 2003). They function in releasing free Pi from esterified substrates, and in the transport and recycling of Pi (Tran et al., 2010a). The majority of plant APases are monomeric or dimeric glycoproteins and are non-specific enzymes that display high activity with the following key metabolites: ATP, PPi, 3-PGA, and PEP (Duff et al., 1994). However, the APases involved in carbon metabolism do display stricter substrate specificity (Sun et al., 2012; Vance et al., 2003). There has also been a soybean APase identified that does not appear to have any known metabolic functions, but comprises 40% of the total protein content in the leaves of depodded plants (Leelapon et al., 2004).

There are two main types of APases: extracellular and intracellular. The extracellular proteins are found in roots and plant cell cultures and are phytases or root-secreted APases that have a broad substrate specificity (Vance et al., 2003). They are localized in the cell wall and/or are secreted into the rhizosphere or culture media. The intracellular proteins are ubiquitous in all plant tissues. They are localized mostly in the vacuole but can also be found in the cytoplasm (Duff et al., 1994). The vacuolar APases function in remobilizing and recycling intracellular Pi and in bypassing P-requiring steps in carbon metabolism (Tran et al., 2010a; Vance et al., 2003)

1.6 Purple acid phosphatases (PAPs)

A particular class of APases is the PAPs. All PAPs are resistant to inhibition by tartrate, contain a bimetallic nucleus, and exhibit a pink/purple color in solution when sufficiently concentrated (Vincent and Averill, 1990). These enzymes are crucial for plant growth during periods of phosphate-starvation because they enable intracellular and extracellular phosphates to be recycled.

PAPs are very widespread and have been found in all higher plants that have been studied thus far, including red kidney bean, (Cashikar et al., 1997), sweet potato (Durmus et al., 1999; Waratrujiwong et al., 2006), duckweed (Nishikoori et al., 2001), tomato (Bozzo et al., 2002), *Arabidopsis thaliana* (Li et al., 2002), tobacco (Kaida et al., 2003; Lung et al., 2008), soybean (Li et al., 2012; Liao et al., 2003), potato (Zimmermann et al., 2004), lupin (Olczak and Olczak, 2007), common bean (Liang et al., 2010; Olczak and Olczak, 2007), oilseed rape (Lu et al., 2008), rice (Hur et al., 2010; Zhang et al., 2011), wheat, barley, and maize (Dionisio et al., 2011), latex (Pintus et al., 2011), and alfalfa (Ma et al., 2012). The roles attributed to the discovered PAPs range widely for the individual proteins, and include generating Pi for the developing seed, scavenging Pi during phosphate-deficiency, participating in cell wall regeneration, mitigating salinity stress, and remobilizing internal Pi.

Schenk et al (2000) showed that although there was low sequence homology for PAPs across the kingdoms, there were seven residues involved in metal-ligation that did not vary. These seven residues (in bold) occur in five amino acid blocks: **DXG/GDXXY/GNH(E,D)/VXXH/GHXH** (Schenk et al., 2000). By using these conserved motifs, PAP genes have been identified from a range of species.

A genome-wide search for the five amino acid block motifs turned up 29 genes for *Arabidopsis thaliana* (Li et al., 2002), 26 for *Oryza sativa* (rice) (Zhang et al., 2011), and 35 for *Glycine max* (soybean) (Li et al., 2012). The white clover

genome has not yet been sequenced in its entirety, which precludes the possibility of doing a genome-wide search, but precedence would suggest it is highly likely that the white clover genome would contain a complement of PAPs.

Plant PAPs in general can be split into two main groups based on their known characteristics (Olczak et al., 2003). The first group is composed of large dimeric plant PAPs (~55 kDa), which can be further divided into two subgroups. Proteins in the first subgroup contain cysteine residues around position 340-370, which form disulfide bridges that bind the two subunits together covalently. The second subgroup lacks these cysteine residues, suggesting that the subunits of the PAPs within this group are held together non-covalently. The third group is composed of small monomeric plant PAPs (~35 kDa), which are structurally similar to mammalian PAPs, but this group has not been well studied in plants (Olczak, Morawiecka et al. 2003).

The 29 putative *Arabidopsis* PAP amino acid sequences cluster into three main groups, each with subgroups to give a total of eight (Fig. 1.3). As predicted, those in groups I and II have a higher molecular weight (~55 kDa) while those in group III have a lower molecular weight (~35 kDa). PAPs in each group not only have a similar molecular weight but are also predicted to have similar biochemical properties as well. However, because most *Arabidopsis* PAPs have not yet been characterized, the possibility of the AtPAPs within each group having similar properties is only speculation based on sequence homology. However, all the PAPs within one subgroup, Ia-2, have been studied. This group consists of AtPAP10, AtPAP12, and AtPAP26, all of which have been shown to be induced during Pi starvation and to aid *Arabidopsis* in acclimating to low external Pi levels (Hurley et al., 2010; Robinson et al., 2012; Tran et al., 2010b; Veljanovski et al., 2006; Wang et al., 2011).

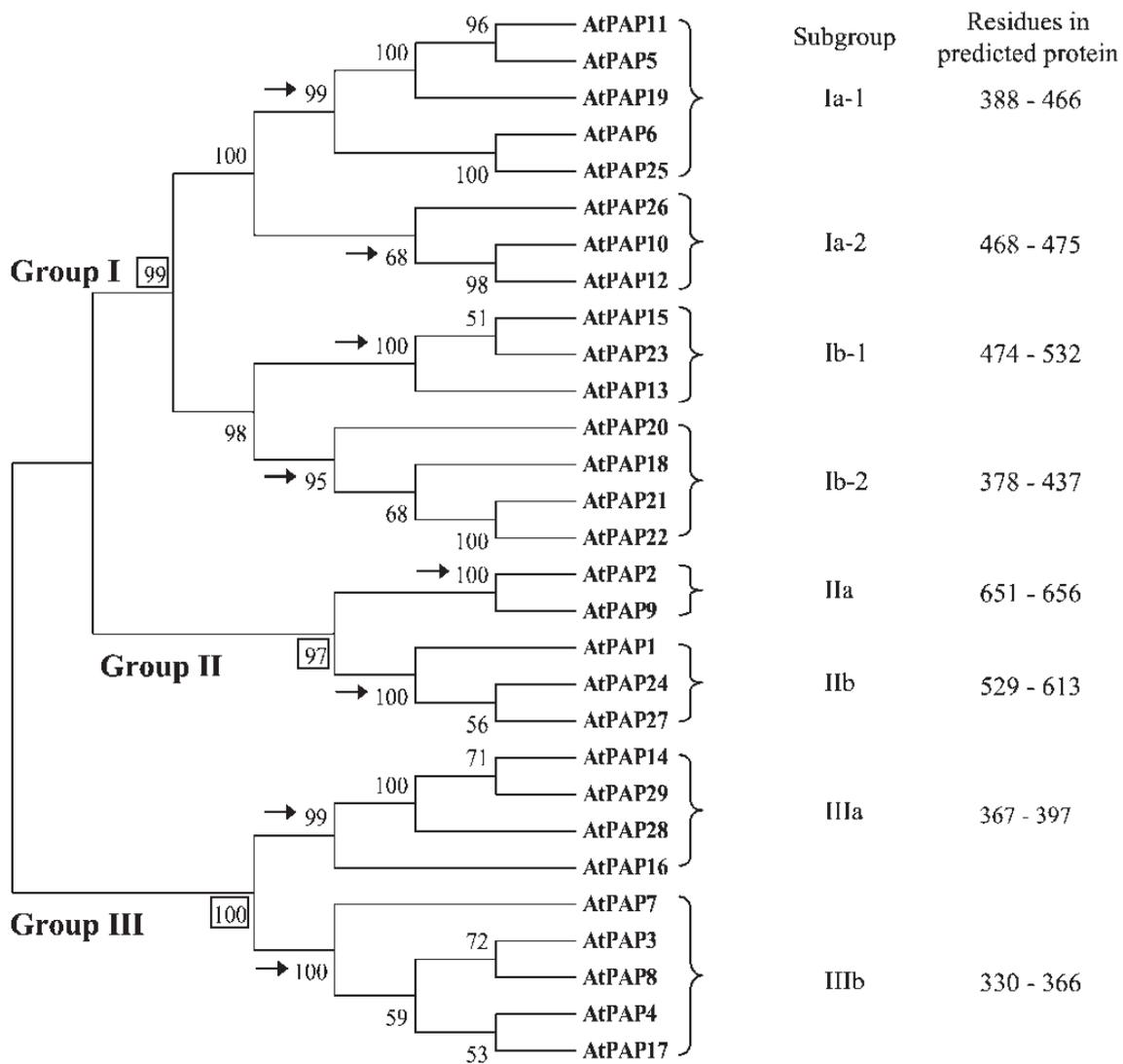


Figure 1.3. A classification scheme for *Arabidopsis* PAPs based on clustering analysis of amino acid sequences. Redrawn from Li et al. 2002.

1.7 Genetic strategies to improve phosphate use efficiency

A series of studies have been reported by which specific genes, including those coding for transcription factors, have been over-expressed in homologous and heterologous backgrounds to improve the PUE of the transformants (Veneklaas et al., 2012). For example, over-expression of the bHLH transcription factor *ZmPTF1* in *Zea mays* resulted in improved root development and higher biomass yield when the plants were grown in Pi-deficient soil (Li et al., 2011).

Some studies have also tried and were successful in over-expressing plant PAPs either in a homologous or heterologous genetic background to generate transgenic lines that are more efficient at utilizing the available Pi supply (Hurley et al., 2010; Ma et al., 2012; Tian et al., 2012; Zhang et al., 2012). For example, over-expression of *AtPAP2* in *Arabidopsis* resulted in earlier bolting and a higher seed yield (Sun et al., 2012). These successes are encouraging because they mean that there is a possible solution to the problem of the world's dwindling natural P resources. Crop plants might be able to be engineered to be more efficient at using and recycling available Pi, and therefore require less P inputs.

Of relevance to this study is that over-expression of *AtPAP26* in *Arabidopsis* has been shown to increase the efficiency of Pi usage (Hurley et al., 2010). This means that if an *AtPAP26* homolog exists in crop species (e.g. white clover), then overexpressing the homolog in those species could be a means of improving PUE.

1.8 White clover

White clover is a prostrate legume and a short-lived perennial that can behave as an annual under water stress. It is stoloniferous, and each node on the stolon consists of a trifoliate leaf, two root primordia, and an axillary bud. If the node comes into contact with moist soil, then roots will develop from the root primordia. The primary root is a shallow taproot that may grow to 1 meter in length. The

flowers of white clover are white globular racemes, which are produced from active apical buds (OGTR, 2008).

White clover reproduces asexually through the generation of stolons. It can also reproduce sexually through flowering and seed production. White clover is a tetraploid ($2n = 4x = 32$) in which self-pollination is infrequent (OGTR, 2008). Because white clover will predominantly outbreed, populations are composed of a heterogenous mixture of highly heterozygous individuals, which translates to a high level of genetic variation allowing for ease of adaptation to various environmental conditions.

Of the pasture crops, white clover is the most important because it has nitrogen-fixing abilities and because its stolons allow it to survive well in closely grazed pastures (Parsons et al., 2012).

Agriculture is very important to New Zealand's economy with dairy exports contributing around 3% of the country's GDP (Schilling et al., 2010). As a result of this huge industry, a very large portion of NZ's land is covered by pastures, and is devoted to grazing livestock that include dairy cattle, beef, sheep, and deer. In 2002, the pasture cover was 39% of NZ's total land and so it is clear that the maintenance of pastures and the grazing crops grown on them is of high priority in NZ (MfE, 2010).

Because white clover is an agronomically important crop, it is an excellent example with which to study and characterize, with an aim to improve the efficiency of Pi uptake and usage.

1.9 White clover and PUE

It's well recognized that the shortage of Pi in soils is a problem, and that current fertilizer application rates are unsustainable. In the past, there have been

programs set up to breed and identify cultivars of white clover that have improved PUE (Dunlop et al., 1990). Some research has also been conducted on the biochemical changes that occur within white clover under phosphate-stress.

Because of the high genetic variability of white clover, there is significant variation in how different genotypes respond to Pi-supply (Hunter and McManus, 1999). Several studies have found that APase activity increases greatly in the roots under Pi-deficient conditions (Almeida et al., 1999; Caradus and Snaydon, 1987; Hunter and McManus, 1999; Hunter et al., 1999; Zhang and McManus, 2000). Not only does APase activity increase in the roots, but the root morphology will also change under Pi-deficient conditions. In particular, both the primary and lateral root length increase in response to low Pi supply (Dinh et al., 2012).

More than one white clover APase has been implicated in the increased APase activity, although whether or not the different studies were looking at the same proteins as each other is unclear (Hunter and McManus, 1999; Hunter et al., 1999; Zhang and McManus, 2000).

Transgenic white clover lines have also been generated. Here, the APase genes *MtPHY1* and *MtPAP1* from *Medicago truncatula* were expressed in white clover with the 35S CaMV promoter and successfully created a transgenic plant that was 2.6-fold more efficient at using organic P when there was a lack of Pi (Ma et al., 2009).

However, prior to this project, there had been no reports of the isolation of any PAP genes from white clover. Thus this project aims to elucidate one possible mechanism by which white clover is able to adapt to Pi-deficiency.

1.10 Project and aims

This project began with a partial EST of a white clover *AtPAP26-like* gene gifted by AgResearch (Palmerston North, NZ). In order to take a close look at the protein itself, it is necessary to discover the full-length gene sequence first. Following discovery of the entire sequence, the protein can then be over-expressed and purified for assessment. The role of the gene and protein *in planta* under varying conditions is also of interest, so it is necessary to grow white clover plants in phosphate-sufficient and phosphate-deficient media. While it is possible to analyze the plant tissues for expression of the gene with only the partial EST sequence, it is not possible to analyze the plant tissues for accumulation and activity of the protein, which makes discovery of the full-length gene sequence especially important.

The project aims, therefore, were based around cloning the full-length *TrPAP26* and then characterizing the substrate specificity of the protein (TrPAP26), and regulation of the transcription of *TrPAP26* in white clover. Specifically:

1. Determine the characteristics of TrPAP26
 - a. Express TrPAP26 in *Escherichia coli* and *Nicotiana tabacum*
 - b. Evaluate the substrate specificity of TrPAP26
2. Determine the role of TrPAP26 in white clover
 - a. Evaluate the expression of *TrPAP26* under +Pi and –Pi conditions
 - b. Evaluate the accumulation and activity of TrPAP26 under +Pi and –Pi conditions

2. Materials and Methods

2.1 Plant material

2.1.1 Maintenance of *Trifolium repens* (white clover) stock plants

White clover of the genotype designated “10F” from the Grasslands Challenge cultivar (AgResearch Grasslands, Palmerston North, New Zealand) was used in this project. The Grasslands 10F genotype was used as a model to study the biochemical responses to Pi-deficiency. The stock plants of this genotype were grown and propagated in potting mix (Oderings, Palmerston North, New Zealand) in a glasshouse at the Plant Growth Unit at Massey University, Palmerston North. In the glasshouse, under natural light conditions, the temperature was controlled to keep a minimum of 12°C at night and cooled by ventilation during the day when the temperature inside the glasshouse reached 25 °C. The plants were automatically watered twice a day for 5 min, at 10:00 and 17:00.

2.1.2 Excision and treatment of white clover cuttings

Four-node stolon cuttings were excised from established wild-type white clover plants (as shown in Fig. 2.1) and all leaves, except the first emerged leaf, were removed (Fig. 2.2). Immediately after excision from the stock plants, the 3rd and 4th nodes were first submerged in vermiculite supplemented with Hoagland solution for 7 days, during which time roots subtended from the 3rd and 4th nodes.

The stolons were then transferred into 50 mL-capacity dark tubes and maintained in complete media for a further 7 days. After this second week, the stolons were transferred into 500 mL of either complete Hoagland solution (1 mM Pi; +Pi) or Hoagland solution with potassium dihydrogen phosphate content adjusted to give a concentration of 10 mM (-Pi). The stolons were maintained in either of these media for 21 days, with the media refreshed every other day.

After 16 days of treatment, 3 plants were removed for the determination of root phosphatase activity. For this, the elongation zone from the root subtending from the 3rd node was used (Fig. 2.3 A,B). After 19 days, another 3 plants were removed from the experiment for a repeat root surface acid phosphatase activity assay (Section 2.3.5.8).

On the 21st day (35 days post-excision), various tissues were harvested from all of the remaining 12 plants in the experiment (Fig. 2.3). The tissues harvested were the first fully expanded leaf (F), a mature leaf (M), the internodes between the first fully expanded leaf and the mature leaf (I), and the primary root from the 3rd node (R), as depicted schematically in Fig. 2.3.

2.2 Chemicals

Unless otherwise stated, the chemical reagents used were analytical grade, obtained from BDH Laboratory Supplies (Poole, BH15 1 TD, England), Sigma Chemicals Company (St. Louis, Mo., USA), and Bio-Rad Laboratories (Richmond, CA, USA).

The purified water used for making solutions was produced by reverse-osmosis, followed by passage through a microfiltration system containing ion exchange, solvent exchange, organic and inorganic removal cartridges (Milli-Q, Millipore Corp., Bedford, Massachusetts, USA).

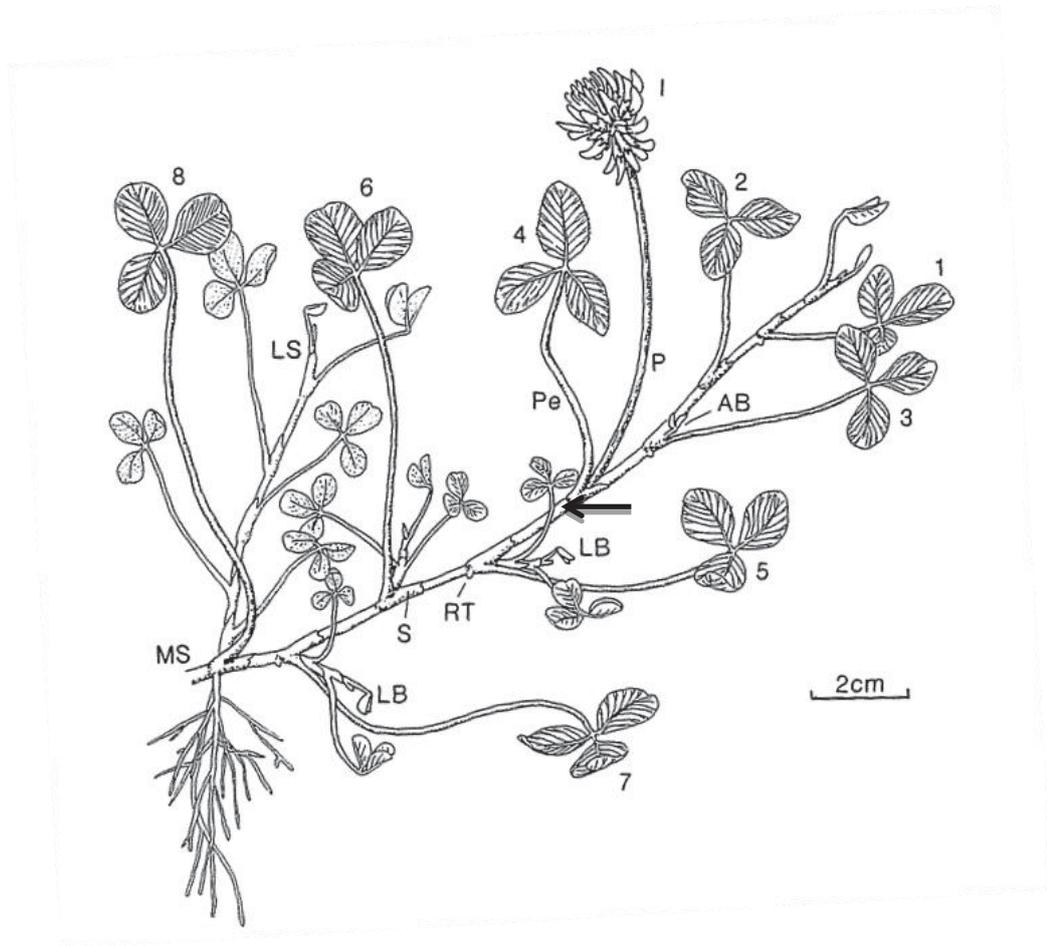


Figure 2.1. Drawing of a main stolon (MS) of white clover, showing axillary buds (AB), lateral branches (LB), and a lateral stolon (LS). S = stipule; Pe = petiole; RT = nodal root primordium; I = inflorescence; P = peduncle. Emerged leaves on the main stolon, and the nodes bearing them, are numbered 1 to 8. The arrow denotes the point of excision of the four-noded explants.

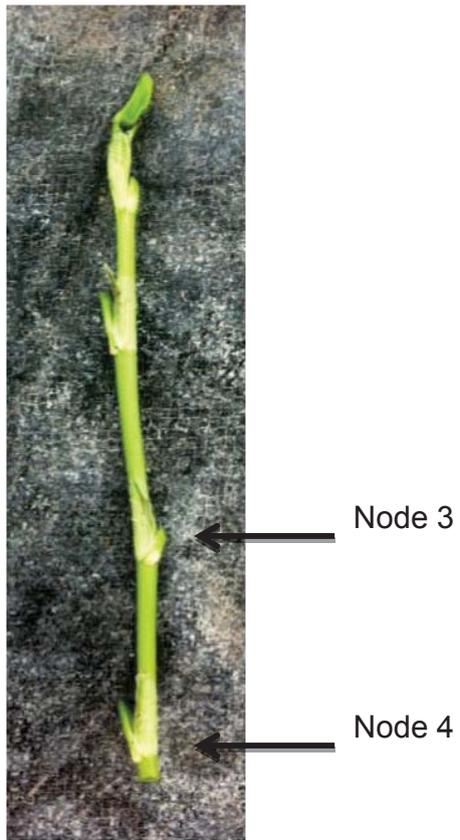


Figure 2.2. Four-nodal stolon cutting prior to planting in vermiculite to induce roots at nodes 3 and 4 (arrowed).

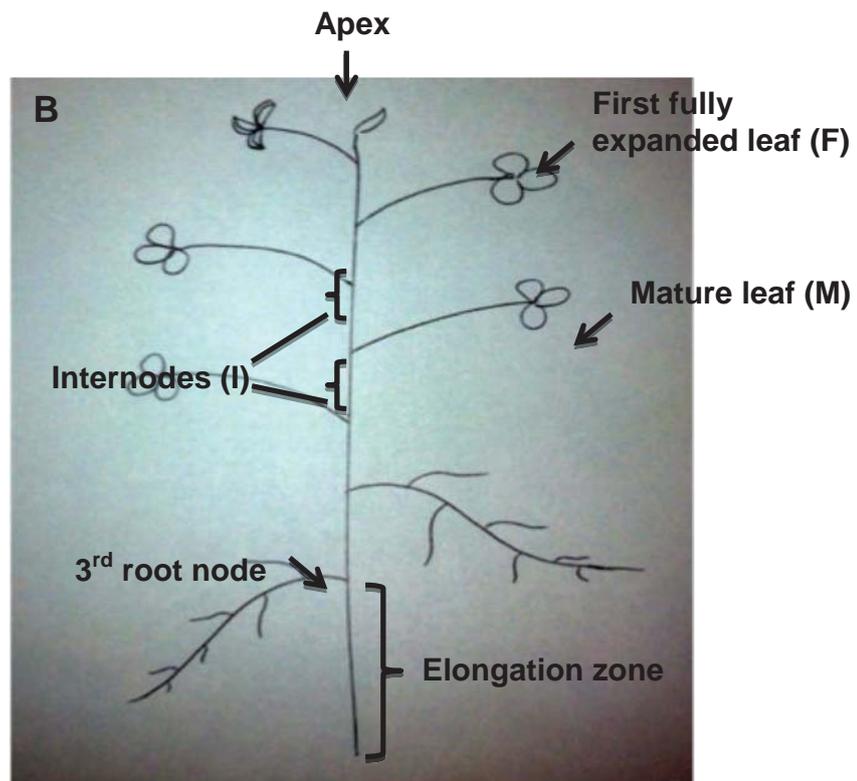


Figure 2.3 A. Example of white clover stolon cutting grown in -Pi media for 21 days. B. Schematic of the stolon cutting to indicate the harvested tissues.

2.3 Biochemical methods

2.3.1 Protein extraction of plant tissue

2.3.1.1 Soluble protein and cell wall protein extraction

Reagents:

1 mM dithiothreitol (DTT)

1 M NaCl

With a mortar and pestle sterilized with 70% ethanol, plant tissue that had been frozen with liquid nitrogen was ground to a powder, resuspended in 1 mM DTT in a ratio of 1 g plant tissue to 3 mL of 1 mM DTT, and the slurry was centrifuged at 6000xg for 10 min at 4°C. The supernatant was transferred to a fresh tube and designated the soluble fraction. The pellet was washed with 1 mM DTT three times, with centrifuging at 6000xg for 5 min at 4°C and discarding the supernatant for each wash. The pellet was then washed with water three times, as described above. Sufficient 1 M NaCl was added to cover the pellet, which was then resuspended, and the mixture incubated at 37°C for 1 h. The extraction was then centrifuged at 6000xg for 5 min at 4°C, and the supernatant was transferred to a fresh tube, and the pellet was again covered and resuspended with 1 M NaCl, and the mixture incubated at 4°C overnight to extract as much of the cell wall protein as possible. The extraction was then centrifuged at 6000xg for 5 min at 4°C, and the supernatant pooled with the first cell wall protein extraction, and collectively designated the cell wall fraction.

2.3.1.2 Glutathione S-transferase (GST) protein purification

Reagents:

LB^{Amp100} broth

100 mM isopropyl β-D-thiogalactopyranoside (IPTG)

Glutathione sepharose 4B beads (GS4B)

Phosphate buffered saline (PBS) (50 mM sodium phosphate, pH 7.4, 0.25 M NaCl)

BugBuster (Novagen, San Diego, CA, USA)

PreScission Protease (GenScript Corporation, USA)

Ten mL of LB^{Amp100} broth was inoculated with the appropriate *E. coli* strain (BL21 or Origami™ 2(DE3)pLysS) transformed with the appropriate vector (section 2.4.10) and incubated overnight at 37°C with shaking at 200 rpm. One mL of the overnight culture was used to inoculate 100 mL of LB^{Amp100} broth, and this new culture was incubated for 5 h at 37°C with shaking at 200 rpm. The accumulation of the GST fusion protein in *E. coli* was induced with the addition of 400 µL of 100 mM IPTG to give a final concentration of 4 mM.

The GST fusion protein was then purified using the GST-fused protein purification protocol (GenScript Corporation, USA). No changes were made to the protocol.

2.3.1.3 Ammonium sulfate precipitation

Reagents:

100 mM Tris-HCl, pH 7.5

Ammonium sulfate

Binding buffer (section 2.3.1.5)

One hundred mM Tris-HCl, pH 7.5, was added to the protein extract (section 2.3.1.1) in a 1:10 buffer : extract ratio in a beaker, and the total volume was noted. The beaker was placed in a container of ice and stirred very slowly. Ammonium sulfate was then added to the beaker slowly to a final concentration of 8 g ammonium sulfate : 10 mL buffered extract. The solution was stirred continuously on ice for 1 h. The solution was then centrifuged at 10000xg at 4°C for 15 min, the supernatant discarded, and the pellet resuspended in binding buffer, typically to

achieve a 5-fold increase in concentration over the initial extract.

2.3.1.4 Dialysis

Reagents:

Binding buffer (section 2.3.1.5)

About 15 cm lengths of dialysis tubing were soaked in binding buffer. Two knots were tied at one end of the tubing, and the resuspended protein which had been concentrated through ammonium sulfate precipitation (section 2.3.1.3), transferred into the dialysis tube, and then a knot was tied at the open end of the tube to seal it off. The sealed dialysis tube was then placed into a beaker filled with binding buffer (typically a 100-fold greater volume than the protein solution in the dialysis tube) so that the protein was submerged in the buffer. The beaker and tube were then stirred continuously at 4°C overnight.

2.3.1.5 Immobilized metal ion affinity chromatography

Reagents:

Chelating Sepharose Fast Flow beads (GE Healthcare, Auckland, NZ)

Acidic buffer (20 mM sodium acetate, pH 4.0, 0.5 M NaCl)

Binding buffer (phosphate buffered saline (PBS) [3.2 mM Na₂HPO₄, 0.5 mM KH₂PO₄, 1.3 mM KCl, 135 mM NaCl, pH 7.4])

Elution buffer (20 mM sodium phosphate, pH 7.4, 0.5 M NaCl, 50 mM imidazole)
0.2 M NiSO₄

To charge the beads, 1 mL was aliquoted into a 15 mL-capacity tube, and the beads washed three times with water, with centrifugation at 500xg for 5 min and then discarding the supernatant after each wash. Then 5 mL of 0.2 M NiSO₄ was added to the beads, followed by incubation at room temperature with end-over-end rotation. The beads were collected by centrifugation at 500xg for 5 min, and the

supernatant was discarded. The beads were then washed five times with water, five times with acidic buffer (or until the effluent was pH 4.0), and finally twice with binding buffer.

To bind histidine-tagged protein to the charged beads, the protein sample was added to the beads and the mixture incubated at room temperature for 2 h with end-over-end rotation. The beads were then collected by centrifugation at 500xg for 5 min, the supernatant discarded, and finally the beads were washed three times with binding buffer to remove any remaining unbound proteins.

To elute the bound protein from the beads, elution buffer was added in increments of 50 uL until the solution was no longer viscous. The beads were then incubated at room temperature for 1 h with end-over-end rotation, collected by centrifugation, as before, and the supernatant was transferred to a fresh tube, and designated the eluted protein.

2.3.2 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Reagents:

12% (w/v) acrylamide separating gel

- 1x separating gel resolving buffer (375 mM Tris-HCl, pH 8.8, 0.1% (w/v) SDS)
- 12% (w/v) acrylamide
- 1% (v/v) 10% (w/v) ammonium persulfate (APS)
- 0.1% (v/v) tetramethylethylenediamine (TEMED)

4% (w/v) acrylamide stacking gel

- 1x stacking buffer (125 mM Tris-HCl, pH 6.8, 0.1% (w/v) SDS)
- 4% (w/v) acrylamide
- 1% (v/v) 10% (w/v) APS
- 0.1% (v/v) TEMED

Isopropanol

Electrophoresis buffer

- 25 mM Tris
- 192 mM glycine
- 0.1% (w/v) sodium dodecyl sulfate (SDS)

Sample loading buffer

- 100 mM tris (pH 6.8)
- 4% (w/v) SDS
- 0.2% (w/v) bromophenol blue
- 20% (v/v) glycerol
- 10% (v/v) 2-mercaptoethanol

Protein samples were mixed with SDS loading buffer in a 1:1 ratio and boiled for 10 min.

The 12% (w/v) acrylamide gel solution was loaded into a gel apparatus before the solution polymerized. Isopropanol was loaded on top of the gel solution to remove air bubbles and to make the top of the gel smooth. Once the gel polymerized, the isopropanol was poured out, and the 4% (w/v) stacking gel solution was loaded on top of the 12% (w/v) acrylamide gel. A well comb was inserted into the stacking gel solution, and once the gel had polymerized, the comb was removed, and the gel was put into a chamber filled with electrophoresis buffer. The prepared protein samples were then loaded into the wells in the gel. The gel was electrophoresced at 150 V for 75 min.

2.3.3 Coomassie brilliant blue staining

Reagents:

Coomassie brilliant blue solution (Meyer and Lamberts, 1965)

30% (v/v) ethanol

Following SDS-PAGE, the gel was stained with Coomassie brilliant blue solution for total protein visualization. To do this, the gel was immersed in Coomassie brilliant blue solution and incubated at room temperature with shaking at 100 rpm for 30 min. The gel was then destained by immersing the gel in 30% (v/v) ethanol, microwaving until bubbles formed in the ethanol, and then incubating at room temperature with shaking at 100 rpm for 15 min.

2.3.4 Western analysis

Reagents:

Primary antibody (anti-PAP12 or anti-PAP26 purified from rabbit) (gifted by the Plaxton lab, Department of Biology, Queen's University, Kingston, Ontario, Canada)

Secondary antibody (anti-rabbit alkaline phosphatase purified from goat) (Sigma-Aldrich, St. Louis, MO, USA)

Transfer buffer [25 mM Tris, pH 8.3, 192 mM glycine, 10% (v/v) methanol]

PBS-Tween (PBS, pH 7.4)(section 2.3.12), 0.05% (v/v) Tween 20)

Blocking buffer (12% (w/v) skimmed milk powder in PBS-Tween)

Substrate buffer (0.1 M Tris, pH 9.6, 0.1 M NaCl)

Developing solution (to be made fresh before use)

- Substrate buffer
- 10 mM MgCl₂
- 1% (w/v) 5-bromo-4-chloro-3'-indolyphosphate (BCIP)
- 1% (w/v) nitro-blue tetrazolium (NBT)

The primary antibody was diluted 1000-fold with PBS-Tween. The secondary antibody was diluted 10000-fold with PBS-Tween.

Following SDS-PAGE, the gel was removed from the gel apparatus and placed into cold (4 °C) transfer buffer. The polyvinylidene fluoride (PVDF) membrane was incubated in 100% methanol for 15 sec, and then incubated in cold (4 °C) transfer

buffer for 5 min. Whatman papers (3 mm) and Scotch pads were also soaked in cold (4 °C) transfer buffer. The gel, membrane, Whatman papers, and Scotch pads were then sandwiched together in a cassette as shown in figure 2.4. The cassette was then placed into the transfer device filled with cold transfer buffer and an ice block. The transfer buffer was continuously stirred, and the cassette was electrophoresced at 100 V for 60 min.

After the protein transfer onto PVDF membrane was completed, the membrane was incubated in blocking buffer at 4°C with shaking at 100 rpm overnight.

On the following day, the membrane was rinsed 3 times with PBS-Tween and then washed 5 times with PBS-Tween for 5 min with shaking at room temperature. The membrane was then incubated with the primary antibody at 37°C with shaking at 100 rpm for 1 h. Following incubation with the primary antibody, the membrane was washed with PBS-Tween as described previously. The membrane was then incubated with the secondary antibody at room temperature with shaking at 100 rpm for 2 h. After incubation, the membrane was washed with PBS-Tween as described.

The membrane was then washed with substrate buffer twice, and incubated with developing buffer in complete darkness at room temperature with shaking at 100 rpm for 1-3 min until bands could be visualized on the membrane. Once bands appeared, the reaction was stopped by rinsing the membrane with water.

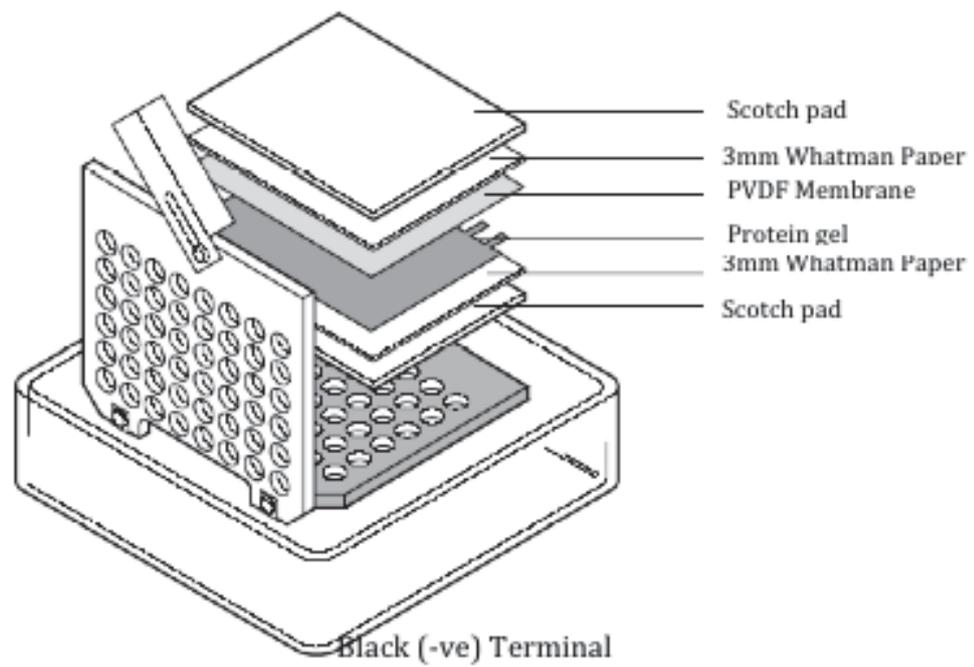


Figure 2.4. A diagrammatic set up for the electrophoretic transfer of proteins from an acrylamide gel to a PVDF membrane

2.3.5 Enzyme assays

2.3.5.1 Protein quantification

Reagents:

Bio-Rad protein assay dye reagent concentrate

Protein extract

Ingredients per reaction:

40 uL dye reagent concentrate

150 uL water

10 uL protein extract

The ingredients were combined in a single well on a microtiter plate. The absorbance at 595 nm was then read using the PowerWave XS spectrophotometer (BioTek Instruments, Winooski, VT, USA). Amount of protein was determined using a standard curve made with bovine serum albumen (BSA) (Fig. 2.3).

2.3.5.2 BSA standard curve

Reagents:

Bio-Rad protein assay dye reagent concentrate

100 mM bovine serum albumen (BSA)

Ingredients per reaction:

40 uL dye reagent concentrate

X uL 100 mM BSA

160 – X uL water

An equivalent of 0 - 3 ug of BSA in a volume of 160 uL was added to microtiter plate wells and 40 uL of dye reagent concentrate was then added. The absorbance at 595 nm was measured using the PowerWave XS spectrophotometer (BioTek Instruments) (Fig. 2.5).

2.3.5.3 pNP standard curve

Reagents:

100 mM *p*-nitrophenol (pNP)

A range of concentrations (0 - 0.5 umoles pNP in a volume of 200 uL) were added to microtiter plate wells. The absorbance at 405 nm was measured using the PowerWave XS spectrophotometer (BioTek, USA) (Fig. 2.6).

2.3.5.4 Inorganic phosphate standard curve

Reagents:

Phosphate assay reagent (section 2.3.5.7)

0.2 mM KH_2PO_4

KH_2PO_4 (0 - 200 uM), in a volume of 50 uL, was added to 200 uL of working phosphate assay reagent in microtiter plate wells to provide a range of concentrations. The absorbance at 620 nm was measured using the PowerWave XS spectrophotometer (BioTek) (Fig. 2.7).

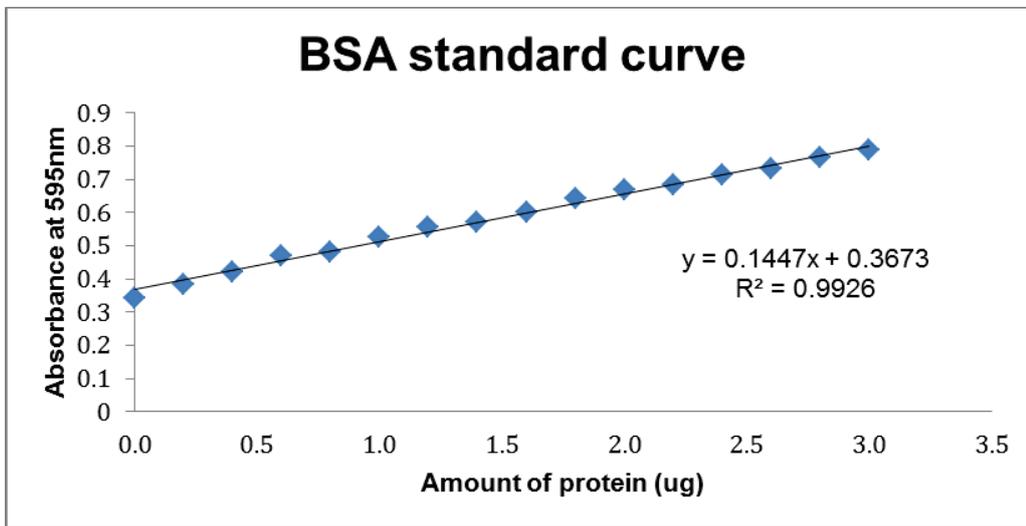


Figure 2.5. BSA standard curve used for protein quantification. Each point represents the mean of three replicates.

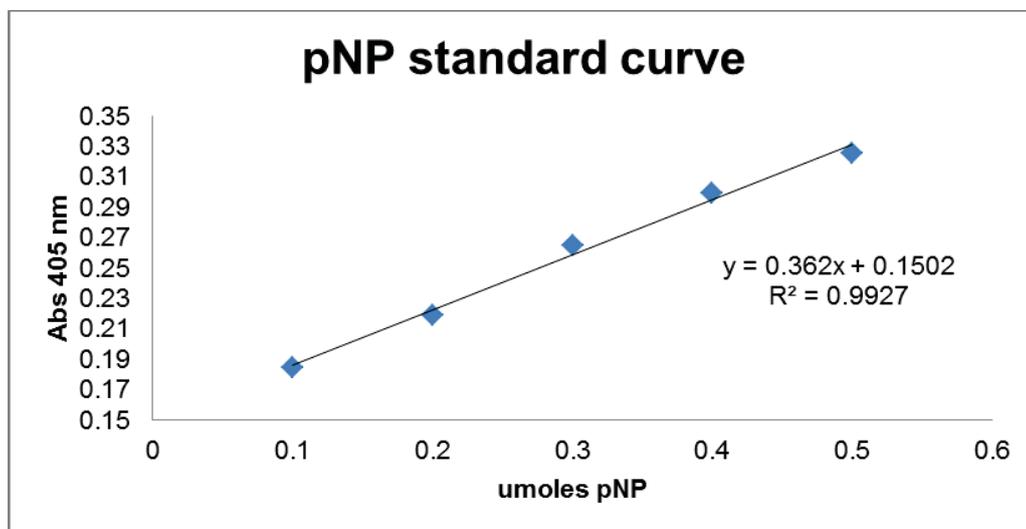


Figure 2.6. pNP standard curve used for measuring acid phosphatase activity. Each point represents the mean of three replicates.

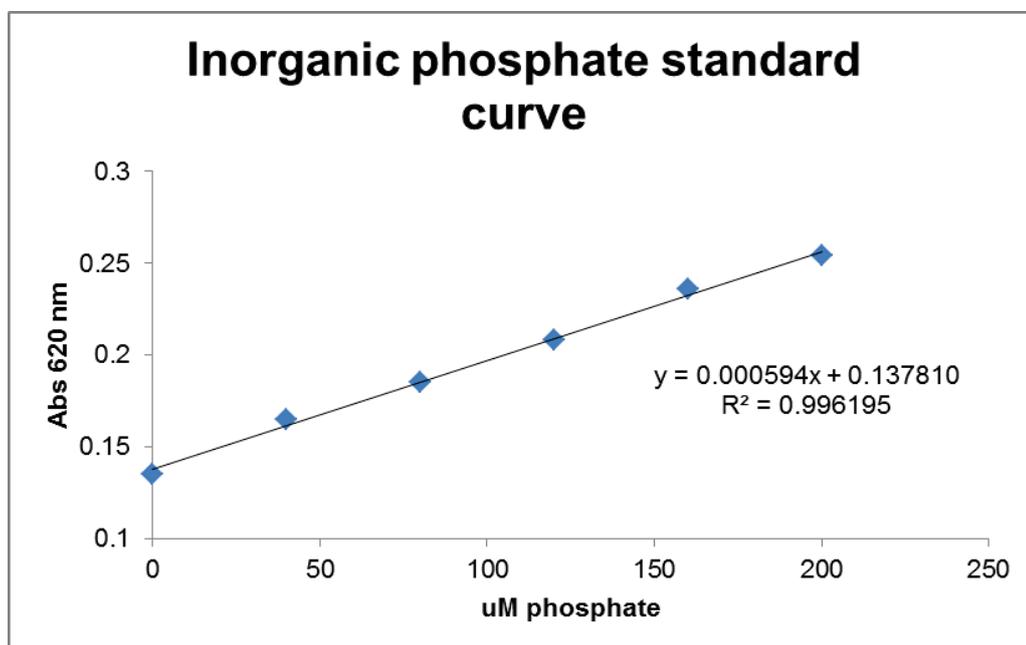


Figure 2.7. Inorganic phosphate standard curve used for measuring acid phosphatase activity. Each point represents the mean of three replicates.

2.3.5.5 Acid phosphatase activity assay A

For routine acid phosphatase activity assays, pNPP was used as the substrate.

Reagents:

p-nitrophenol phosphate (pNPP)

50 mM sodium citrate, pH 5.6

Protein extract

Ingredients per reaction:

200 μ L of 8 mM pNPP in 50 mM sodium citrate, pH 5.6

X μ L protein (5 μ g protein)

50 – X μ L water

The reaction mixture (250 μ L total volume) was added to microtiter plate wells and read at 405 nm every 5 min in the PowerWave XS spectrophotometer (BioTek Instruments), which had been preheated to 37°C. The activity of the protein was measured by the amount of pNP released in the reaction, which was determined using a pNP standard curve (Fig. 2.6).

2.3.5.6 Modification to acid phosphatase activity assay A

Reagents:

p-nitrophenol phosphate (pNPP)

MgCl₂

50 mM sodium citrate, pH 5.6

Protein extract

Ingredients per reaction:

200 μ L of 8 mM pNPP, 10 mM MgCl₂ in 50 mM sodium citrate, pH 5.6

X μ L protein (5 μ g protein)

50 – X uL water

The assay was carried out using the same procedure as described in section 2.15.1.

2.3.5.7 Acid phosphatase activity assay B

For acid phosphatase activity assays using substrates other than pNPP, this protocol was used.

Reagents:

Ammonium molybdate

H₂SO₄

Antimony potassium tartrate (oxide)

Ascorbic acid

Adenosine triphosphate (ATP)

Phospho-enol-pyruvate (PEP)

Phytic acid

Sodium pyrophosphate

50 mM Tris-maleate buffer, pH 6.3 (Gomori, 1955)

Concentrated phosphate assay reagent:

16 mM ammonium molybdate

2.25 mM H₂SO₄

0.15 mM antimony potassium tartrate (oxide)

Phosphate assay reagent (to be made fresh before use):

0.87 g ascorbic acid dissolved in 25 mL of water

6.25 mL concentrated phosphate assay reagent

Adjusted to 50 mL with water

Substrate solution:

8 mM substrate in 50 mM Tris-maleate buffer, pH 6.3

Ingredients per reaction:

245 uL substrate solution, preheated to 37°C

X uL protein (5 ug protein)

5 – X uL water

The ingredients were mixed in eppendorf tubes and incubated at 37°C. Aliquots (50 uL) of the reaction mixture were removed every 5 min and added to 200 uL of the phosphate assay reagent in a microtiter plate well. The absorbance at 620 nm was read using the PowerWave XS spectrophotometer (BioTek). The activity of the protein was measured by the amount of inorganic phosphate released in the reaction, which was determined using an inorganic phosphate standard curve (Fig. 2.7).

2.3.5.8 Root surface acid phosphatase activity assay

Reagents:

Reaction buffer (50 mM sodium citrate, pH 5.6, 10 mM MgCl₂)

Substrate buffer (10 mg/mL pNPP in reaction buffer)

The elongation zone of the root subtending from the 3rd node was suspended in 180 uL of reaction buffer and 20 uL of substrate buffer was then added for a total of 200 uL. The reaction was kept in the dark and incubated at 37°C with shaking for 30 min. The absorbance was then read at 405 nm.

2.4 Molecular methods

2.4.1 Genomic DNA isolation

Reagents:

Genomic DNA Mini Kit (Geneaid Biotech Ltd., New Taipei City, Taiwan)

Plant tissue (50 – 100 mg) was frozen in liquid nitrogen and ground to a powder using mortar and pestle. The Genomic DNA Mini Kit was then used to extract genomic DNA from the tissue according to the manufacturer's protocol. No changes were made to the protocol.

2.4.2 RNA isolation

Reagents:

Trizol (Invitrogen)

Chloroform

Isopropanol

75% (v/v) ethanol

DEPC-treated Milli-Q water

Plant tissue (50 – 100 mg) was frozen in liquid nitrogen and ground to a powder using mortar and pestle. The Trizol reagent was then used to extract RNA from the tissue according to the manufacturer's protocol. No changes were made to the protocol.

2.4.2.1 Generation of cDNA

2.4.2.1.1 DNase Treatment of RNA

Two µg of RNA was used to remove any genomic DNA contamination using recombinant RNase free DNAase I (Roche) in the following reaction mixture:

RNA (2 µg)	x µL
10 x DNase I buffer	5 µL
DNase I	1 µL
DEPC-treated water	44 - x µL

TOTAL	50 µL

The solution was incubated at 37°C for 15 min, 8 mM EDTA was then added and the DNase enzyme heat-inactivated at 75°C for 5 min. DEPC-treated water (450 µL) was added to each sample followed by 400 µL of cold (4°C) 4 M LiCl and the RNA precipitated at 4°C overnight. The precipitate was collected by centrifuging at 20,800 xg for 30 min at 4°C, the pellet was washed with 200 µL of 80% (v/v) ethanol and the RNA pellet air-dried for 10 min before being re-suspended in 10 µL of DEPC-treated water. The RNA was checked to ensure that it was free of DNA using PCR (section 2.4.5) with *NtGAPDH1*.

2.4.2.1.2 Reverse Transcription

Reverse transcription of RNA to generate cDNA was performed using Transcriptor reverse transcriptase (Roche) using the following reactions:

(DNA-free) RNA (1 µg)	5 µL
50 µM Anchor Oligo dT	1 µL
DEPC-treated water	7 µL

	13 µL

Heat denatured at 65°C for 10 min.

The following reaction mixture was then added:

5 x Transcriptor Buffer	4 μ L
20 mM dNTP	2 μ L
PROTECTOR RNase Inhibitor	0.5 μ L
Transcriptor Reverse transcriptase	0.5 μ L

	20 μ L

The reaction was incubated at 55°C for 30 min and then heat inactivated at 85°C for 5 min.

The resulting cDNA was used for both sqRT-PCR (section 2.4.5.1) and qPCR (section 2.4.5.2).

2.4.3 Quantification and purity assessment of nucleic acid

Nucleic acid purity and quantity were assessed using a Nanodrop spectrophotometer (Thermo Fisher Scientific Inc, MA, USA) according to the manufacturer's protocol.

2.4.4 Primer design

Primers in this thesis were designed based on the following characteristics:

- melting temperature (T_m) between 50°C and 60°C
- approximately 50% guanine and cytosine (GC) content
- approximately 20 base pairs (bp) in length

Primers were obtained from Sigma-Aldrich (St. Louis, MO, USA), and are listed in Table 2.1 and Table 2.2. A summary of the rationale and of all of the primers used in this project specifically is given in Appendix 5.10.

Primer name	Sequence (5' to 3')
<i>TrPAP26 F1</i>	CATGGGAGAAGTTGTTCC
<i>TrPAP26 F2</i>	ACTACTCCTTATTTGGCGTC
<i>ATG F</i>	ATGCAGGGSTTGTTGTTT
<i>TAA R</i>	CCTATTAAATGCTGACAACCTC
<i>pGEX F</i>	GACGACGAATTCAATGCAGGGSTTGTTGTTT
<i>pGEX R</i>	GACGACGTCGACTTAAATGCTGACAACCTTCATCG
<i>pGEX Δsp F</i>	GACGACGAATTCAGGGATCACTAGTTCTTTGTTAGGT
<i>PAP26 int seq F</i>	GAAACACCTCCTAAAGTTG
<i>PAP26::6xHis F</i>	GACGACTCTAGAATGCAGGGGTTGTTGTT
<i>PAP26::6xHis R</i>	GACGACTCTAGATTAGTGGTGGTGGTGGTGGTGAATG CTGACAACCTTCATCG
<i>NtGAPDH qPCR F</i>	TGCTGCTGTGAGGAGTCTGT
<i>NtGAPDH qPCR R</i>	GACTGGGTCTCGGAATGTGT

Primer name	Sequence (5' to 3')
3' RACE adapter	GACTCGAGTCGACATCG
Autoseq pGEM13F	GGTTTCCCAGTCACGAC
Autoseq pGEM13R	TCACACAGGAAACACTCTATGAC
Autoseq pGEXF	AAACGTATTGAAGCTATCC
Autoseq pGEXR	AGAATTATACACTCCGCTAT

2.4.5 Polymerase chain reaction (PCR) amplification

Reagents:

10 μ M forward primer

10 μ M reverse primer

Filter-sterilized, autoclaved Milli-Q water (PCR water)

2x PCR master mix (Promega Corporation, Madison, WI, USA)

PCR ingredients (per reaction):

1 μ L 10 μ M forward primer

1 μ L 10 μ M reverse primer

10 μ L 2x PCR master mix

7 μ L PCR water

1 μ L genomic DNA (section 2.4.1), RNA (section 2.4.2.1.1), or cDNA (section 2.4.2.1.2)

All the ingredients were mixed together in a 200 μ L clear Eppendorf tube, and centrifuged briefly to collect at the bottom of the tube. PCR was performed using the Maxygene thermal cycler (Axygen Inc., Union Bay, CA, USA) using the following protocol:

1 cycle: 95°C for 3 min

35 cycles:

- 95°C for 30 sec
- 50-60°C for 30 sec (temperature depends on T_m of primers)
- 72°C for X sec (add 15 sec for every 200 bp to be annealed)

1 cycle: 72°C for 10 min

Finish and hold at 10°C

2.4.5.1 Semi-quantitative reverse-transcriptase PCR (sqPCR)

To determine the expression of *TrPAP26* in T₀ lines of transgenic tobacco, sqRT-PCR was used with expression of *TrPAP26* standardised against the reference gene *NtGAPDH*. To do this PCR was performed with cDNA (2.4.2.1.2) as described in section 2.4.5. At the conclusion of PCR, the products were separated using gel electrophoresis and the gel bands of interest were quantified using Image Lab Software version 3.0 (BioRad). The relative *TrPAP26* expression = the volume of *TrPAP26* band / the volume of *NtGAPDH* band for each treatment.

2.4.5.2 Quantitative RT-PCR (qPCR)

The qPCR method was used to detect the expression of *TrPAP26* in the white clover background. These procedures were performed by Ms Susanna Leung (Institute of Fundamental Sciences, Massey University), and the protocol is reproduced in Appendix 5.11.

2.4.6 Agarose gel electrophoresis

Reagents:

Agarose

TAE buffer (40 mM Tris acetate, pH 8.3, 1 mM EDTA)

Gel loading dye

- 40% (w/v) sucrose
- 0.25% (w/v) bromophenol blue
- 0.25% (w/v) xylene cyanol

HyperLadder I (Bioline USA Inc., Randolph, MA, USA)

Ethidium bromide

A 1% (w/v) agarose gel solution was prepared by combining 1 g of agarose with 100 mL of TAE buffer. This solution was microwaved on high for about 2 min or

until all of the agarose had dissolved. Once dissolved, the hot solution was poured into a gel mould until the mould was filled. A well comb was inserted approximately 1 cm from one end of the mould. The solution was then allowed to cool to room temperature, and once solidified into a gel, TAE buffer was poured over it until the buffer filled the tray and covered the gel. The comb was gently removed to prevent ripping of the wells.

Five μL of the HyperLadder I was loaded into one well, and 5 μL of the appropriate DNA sample was mixed with 2 μL of gel loading dye and then loaded carefully into the remaining wells. Electrophoresis was then conducted at 100 V for 45 min.

A destaining solution was made containing TAE buffer and a few drops of ethidium bromide. Following electrophoresis, the gel was incubated in the destaining solution for 10 min.

The gel was then exposed to ultraviolet light for either photography or gel extraction. For photography, the Bio-Rad Transilluminator (Bio-Rad) was used.

2.4.7 Gel slice extraction and purification

Reagents:

1% (w/v) agarose gel

TAE buffer

Ethidium bromide

QIAquick gel extraction purification kit (Qiagen)

3 M sodium acetate, pH 5.0

Three to four replicates (each 20 μL) from PCR amplification were pooled into one 1.5 mL Eppendorf tube, and the pooled sample was then concentrated down to approximately 30 μL using the Speed-Vac (Savant Instrument, Holbrook, NY)) set at medium speed. At this speed, there is a decrease of *ca.* 1 μL per minute.

Thirty μL of the concentrated sample were then mixed with 5 μL of gel loading dye and then loaded into a single well in a 1% (w/v) agarose gel for electrophoresis followed by destaining as described in section 2.4.6.

The resulting band(s) of DNA was visualized under ultraviolet light and excised from the gel using a clean scalpel. The gel slice was placed in a 1.5 mL clear Eppendorf tube and weighed. The DNA was purified from the gel using the QIAquick kit according to the manufacturer's protocol. No changes were made to the protocol.

2.4.8 DNA sequencing

Reagents:

10 μM primer (the primer used depends on the DNA to be sequenced)

Sequence buffer (Solexa Genome Sequencing)

Big Dye Terminator (Solexa Genome Sequencing)

PCR-grade water

DNA

Sequencing ingredients (per reaction):

2 μL 10 μM primer

3 μL sequence buffer

2 μL Big Dye Terminator

X μL DNA

13 – X μL PCR water

(20 μL total)

Purified DNA was first quantified using a Nanodrop spectrophotometer (Thermo Fisher Scientific Inc.). The DNA was amplified through PCR using the Maxygene thermal cycler (Axygen Inc.) with the following protocol:

27 cycles of

- 96°C for 30 sec
- 50°C for 15 sec
- 60°C for 4 min.

The amplified DNA was then sent to the Solexa Genome Sequencing Service, Massey University for reaction clean-up and capillary separation using in-house protocols.

2.4.9 Vector ligation

2.4.9.1 cDNA preparation

Reagents:

White clover cDNA (section 2.4.2.1.1)

10 uM forward primer

10 uM reverse primer

PCR water

2x PCR master mix (Promega Corporation, Madison, WI, USA)

QIAquick gel extraction purification kit (Qiagen)

*Xba*I

10x buffer H

The cDNA was modified to be flanked by restriction enzyme digestion sites through PCR amplification (see section 2.4.5). The PCR product was then separated using 1% (w/v) agarose gel electrophoresis, and the resulting band was extracted from the gel using the QIAquick gel extraction purification kit (Qiagen) (see section 2.4.7). This produced pure cDNA, which was then digested with the appropriate restriction enzyme.

cDNA digestion ingredients (per reaction):

10 uL modified cDNA

1.5 uL Buffer H

1 uL restriction enzyme

2.5 uL water

(15 uL total)

The ingredients were combined into a tube and incubated overnight at 37°C, and then at 65°C for 15 min to deactivate the restriction enzyme. The cDNA was then quantified using a Nanodrop spectrophotometer (Thermo Fisher Scientific Inc.).

2.4.9.2 Vector preparation

Reagents:

Vector

10x buffer H

Restriction enzyme

rAPid alkaline phosphatase (Roche Applied Science)

10x buffer (Roche Applied Science)

sterile Milli-Q water

Vector digestion ingredients (per reaction):

2 uL vector

1 uL 10x buffer H

1 uL *Xba*I

6 uL water

(10 uL total)

The ingredients were combined into a tube and incubated overnight at 37°C. The digested vector was then dephosphorylated to prevent the vector from religation.

Dephosphorylation ingredients (per reaction):

10 uL vector digestion reaction

2 uL 10x buffer

1 uL rAPid alkaline phosphatase

7 uL sterile Milli-Q water

(20 uL total)

The ingredients were mixed together and incubated at 37°C for 15 min, then at 75°C for 4 min to deactivate the alkaline phosphatase. The vector concentration was then quantified using Nanodrop spectrophotometry (Thermo Fisher Scientific Inc.).

2.4.9.3 Ligation

Reagents:

Digested cDNA

Digested and dephosphorylated vector

T4 DNA ligase (5 u/uL) (Promega Corporation, Madison, WI, USA)

10x buffer (Promega Corporation, Madison, WI, USA)

Ligation ingredients (per reaction):

2 uL 10x buffer

1 uL T4 DNA ligase

X uL cDNA

Y uL pDAH2

17 – (X + Y) uL Milli-Q water (to give a 20 uL reaction volume)

To determine the ratio of vector-to-insert, the following equation was used:

$$\text{ng cDNA} = (50 \text{ ng vector} / \text{vector size}) * \text{insert size} * 3$$

The ingredients were mixed together in a tube and incubated at 25°C overnight.

2.4.10 Transformation of bacterial cells

2.4.10.1 Preparation of competent *E. coli* cells

Reagents:

E. coli DH5 α or BL21

Luria-Bertani (LB) broth

- 1% (w/v) peptone
- 0.5% (w/v) granulated yeast extract
- 0.17 M NaCl

LB agar

- LB broth
- 1.5% (w/v) agar

Ampicillin (100 mg/mL)

Kanamycin (100 mg/mL)

60 mM CaCl₂

Glycerol

Ampicillin was added to liquid LB agar to a concentration of 100 ug/mL (LB^{Amp100}), and kanamycin was added to a concentration of 50 ug/mL (LB^{Kan50}). The resulting LB agar containing either ampicillin or kanamycin was poured into petri dishes and allowed to set. The *E. coli* cells were streaked on both types of media to check if they contained antibiotic-resistant genes. Once it was confirmed that they were not resistant, 10 mL of LB broth was inoculated with the cells in a flask, which was then stopped with a cotton ball. The cells were incubated overnight at 37°C with shaking at 180 rpm. Forty mL of LB broth was then inoculated with 400 uL of the overnight culture, this new inoculation was then incubated at 37°C with shaking at 180 rpm. The culture was sampled occasionally to check the absorbance at 600 nm, and once the culture reached a reading of 0.7 (about 2-2.5 h), it was poured into chilled centrifuge tubes on ice. The cells were centrifuged at 400xg at 4°C for 5 min, the supernatant discarded, and 10 mL of 60 mM CaCl₂ added to resuspend

the cell pellet. Once resuspended, another 10 mL of 60 mM CaCl₂ was added, the mixture incubated on ice for 30 min, and then centrifuged at 400xg at 4°C for 5 min. The supernatant was once again discarded, 4 mL of 60 mM CaCl₂ in 15% (w/v) glycerol added to resuspend the pellet, and then 100 uL aliquots were aliquoted into individual eppendorf tubes and either stored at 4°C or frozen in liquid nitrogen and stored at -80°C.

Competency was checked by transforming the cells with a pure plasmid containing either a kanamycin- or ampicillin-resistant gene, and then streaking the cells onto an LB agar plate containing the appropriate antibiotic. If colonies formed on the plate after an overnight incubation at 37°C, then the cells were deemed to be competent.

2.4.10.2 Transformation of *E. coli* with pGEM-T vector

Reagents:

Competent *E. coli* DH5α

LB broth

100 mM IPTG

50 mg/mL 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal)

pGEM-T ligation reaction

LB^{Amp100} plates

Competent *E. coli* DH5α cells were thawed on ice, half of the ligation reaction (10 uL) was mixed with 50 uL of cells, and the mixture incubated on ice for 10 min. The cells were then incubated at 42°C, after which 500 µl of LB broth was immediately added, and the mixture incubated at 37°C for 1 h. An aliquot (100 uL) of 100 mM IPTG and 20 uL of 50 mg/mL X-gal were spread onto LB^{Amp100} plates, the plates then incubated for 30 min at 37°C, before 200 µl of the culture was spread onto the plates before incubation overnight at 37°C.

Blue colonies indicated those cells that had taken up an empty vector. White colonies indicated those cells that had taken up a vector containing the gene of interest. Using a toothpick, single white colonies were picked from the plates and dipped into the PCR reaction mixture containing the appropriate primers. After dipping into the PCR reaction mixture, the toothpicks were streaked on LB^{Amp100} plates. The plates were incubated at 37°C overnight. The PCR product was confirmed using agarose gel electrophoresis to determine which colonies contained the gene of interest (see section 2.4.5 and 2.4.6).

2.4.10.3 Plasmid DNA purification

Reagents:

E. coli culture

LB^{Amp100} broth

High pure plasmid isolation kit (Roche Applied Science)

The *E. coli* colonies found to contain the gene of interest (section 2.4.10.2) were used to inoculate 8 mL of LB^{Amp100} broth in a 15 mL-capacity tube and the cells incubated overnight at 37°C at 200 rpm. After incubation, the cells were collected by centrifugation at 13000xg for 10 min, and the supernatant was discarded. The plasmid DNA was extracted from the cells using the high pure plasmid isolation kit according to the manufacturer's protocol. No changes were made to the protocol.

2.4.10.4 Transformation of *E. coli* with pGEX-6P-3 vector

Reagents:

Competent *E. coli* BL21 or origami

LB broth

pGEX-TrPAP26

LB^{Amp100} plates

Transformation into *E. coli* was undertaken using the same protocol as described for pGEM-T vector (see section 2.4.10.2) except the LB^{Amp100} plates did not contain IPTG and X-gal. Colony screening was also carried out as described in section 2.4.10.2.

2.4.10.5 Transformation of *Agrobacterium tumefaciens* with pART27 vector

Reagents:

Competent *A. tumefaciens*

YEB media, pH 7.3

- 0.5% (w/v) beef extract
- 0.1% (w/v) yeast extract
- 0.5% (w/v) peptone
- 0.05% (w/v) MgSO₄*7H₂O

YEB agar

- YEB media
- 1.5% (w/v) agar

YEB^{Spec100}

- YEB agar
- 0.1 mg/mL spectinomycin

Vector DNA (pART27 containing the full-length *TrPAP26* insert)

Five uL of the plant transformation vector, pART27 containing the full-length *TrPAP26* as insert (pART27-*TrPAP26*) was added to 40 uL of competent *A. tumefaciens* cells. This mixture was placed into an electroporation cuvette that had been cooled on ice. The mixture was electroporated using the MicroPulser Electroporation Apparatus (Bio-Rad) on the “Agr” setting, before 1 mL of YEB media was immediately added to the cuvette after electroporation. The mixture was then transferred to a 15 mL-capacity tube and incubated at 30°C for 3 h at 180 rpm. After incubation, the culture was centrifuged at 7500xg for 5 min and 500 uL of the supernatant was removed from the tube. The pellet was resuspended in

what remained of the supernatant, and YEB^{Spec100} plates were then spread with 50 uL of the concentrated cell culture. The plates were incubated at 25°C for 72 h.

White colonies were picked for PCR screening using the “35s F” and “PAP26::6xHis R” primers (see section 2.4.5) (Table 2.1).

2.4.10.6 A. tumefaciens plasmid miniprep

Reagents:

YEB^{Spec100} liquid media

Solution A, pH 8.0

- 50 mM glucose
- 10 mM EDTA
- 25 mM Tris

Solution B (to be prepared fresh before use)

- 0.2 M NaOH
- 1% (w/v) SDS

Solution C, pH 5.2

- 3 M potassium acetate

1 M NaCl

Isopropanol

Ethanol

10 mM Tris, pH 8.0

A. tumefaciens colonies found to contain the insert were used to inoculate 10 mL of YEB^{spec 100} liquid media, and the inoculated broth incubated at 25°C for 48 h. The culture was then centrifuged at 700xg for 10 min, the supernatant discarded, and the pellet resuspended in 200 uL of 1 M NaCl. The mixture was then centrifuged at 7500xg for 1 min, the supernatant discarded, and the pellet was resuspended in 200 uL of solution A. Solution B (200 uL) was then added, and the mixture incubated at room temperature for 5 min, before 350 uL of solution C was

added, and the tube was inverted to mix the contents. This mixture was incubated on ice for 5 min, centrifuged at 7500xg for 10 min, and the supernatant was transferred to a fresh tube. Four hundred uL of isopropanol was then added, the mixture incubated at room temperature for 15 min, and then centrifuged at 7500xg for 10 min. The supernatant was discarded, 500 uL of ethanol was added, but the contents were not mixed. This was then centrifuged at 7500xg for 2 min. The supernatant was discarded. The pellet was air-dried, and then resuspended in 30 uL of 10 mM Tris, pH 8.0.

2.4.11 Transformation of *Nicotiana tabacum* with transformed *A. tumefaciens*

Reagents:

LB broth

100 mg/mL kanamycin

Sterile wild-type tobacco plants

10 mM MgSO₄

Nic I media, pH 5.8

- 1 mL/L Murashige and Skoog powder with vitamins
- 3% (w/v) sucrose
- 1 ug/mL 6-benzyl aminopurine (BAP)
- 0.1 ug/mL 1-napthaleneacetic acid (NAA)
- 0.8% (w/v) phytoagar
- adjust pH using KOH

Nic II media, pH 5.8

- Nic 1 media
- 0.1 mg/mL kanamycin

Nic III media, pH 5.8

- 1 mL/L Murashige and Skoog powder with vitamins
- 3% (w/v) sucrose
- 0.8% (w/v) phytoagar

- 0.1 mg/mL cefotaxime

Nic IV media, pH 5.8

- 1 mL/L Murashige and Skoog powder with vitamins
- 3% (w/v) sucrose
- 0.8% (w/v) phytoagar
- 0.1 mg/mL kanamycin

0.1 mg/mL cefotaxime

The purified *A. tumefaciens* plasmid (section 2.4.10.6) was confirmed to have the insert through an *E. coli* transformation. To do this, *E. coli* DH5 α was transformed with the plasmid (see section 2.10) and plated on LB spec 100 plates. Clone screening was done through PCR, followed by plasmid miniprep and DNA sequencing (see sections 2.4.5, 2.4.10.2, and 2.4.8, respectively). The *A. tumefaciens* culture that the purified plasmid was identified from was used to transform *N. tabacum* (tobacco).

To do this, kanamycin was added to 15 mL of LB broth to a concentration of 25 μ g/mL. The broth was inoculated with *A. tumefaciens* containing the pART27-TrPAP26 plasmid (see section 2.4.10.3), and then incubated at 25°C for 48 h at 180 rpm. After incubation, the culture was centrifuged at 1000xg for 10 min, the supernatant was discarded, and the pellet resuspended in 10 mL of 10 mM MgSO₄.

A healthy wild-type tobacco leaf was cut into 1 cm² pieces avoiding the midrib. The pieces were placed in the suspension of MgSO₄ and *Agrobacterium tumefaciens* containing the gene of interest and allowed to sit for 3 min. The pieces were then blotted dry on sterile filter paper and placed on Nic II plates with the upper epidermis facing down. The control was leaf pieces not dipped in the MgSO₄ suspension and placed on Nic I and Nic II plates. Putting untransformed tobacco on Nic II plates was to ensure that the kanamycin did kill untransformed tobacco.

Tobacco leaf pieces that were transformed (transgenic) or not (wild-type control) were transferred onto fresh Nic I (control) or Nic II (transgenic) media once a week until shoot formation was observed. One shoot from each callus was cut off and transferred onto Nic III (control) or Nic IV (transgenic) media to allow for root development. The plantlets were transferred onto fresh media on a weekly basis to maintain a high supply of nutrients and to ensure rigorous selection for the transgenics.

Plants that had developed adequate root systems in tissue culture were then transferred to individual PB1 horticultural pots containing sterile potting mix (Daltons, Tauranga, NZ). Each pot was covered with a transparent plastic bag to generate a humidified atmosphere and then transferred to a GMO Containment Glasshouse and left to harden off for two weeks. Once hardened off, the plants were removed from the bag and maintained under natural light at 23⁰C day/15⁰C night temperatures, with 70% relative humidity.

2.4 Bioinformatic techniques

2.4.1 Basic local alignment search tool (BLAST)

The nucleotide BLAST program from the National Library of Medicine found at <http://blast.ncbi.nlm.nih.gov/Blast.cgi> was used to perform sequence searches. The *TrPAP26* partial sequence was searched against the "Others (nr etc.)" database, optimized for "somewhat similar sequences (blastn)."

Results from the BLAST search were analyzed. Hits with an E value lower than 1e-32 were considered, especially hits to other legume species (*Medicago truncatula*, *Lotus japonicus*, *Lupinus luteus*, *Glycine max*, *Phaseolus vulgaris*) and to *Arabidopsis thaliana*.

2.4.2 Clustal Omega

The multiple sequence alignment program from the European Bioinformatics Institute found at <http://www.ebi.ac.uk/Tools/msa/clustalo/> was used to align sequences obtained from BLAST searches.

2.5 Statistical analysis

A two-tailed Student's paired t-test was used in Microsoft Excel to analyze statistical significance of the data (Student, 1908).

3. Results

3.1 Isolation of a *PAP26-like* gene from white clover

3.1.1 Obtaining the full-length *PAP26-like* cDNA sequence

A partial EST sequence (897 bp) of an *AtPAP26*-like gene (designated “*TrPAP26*”) was gifted by the New Zealand Pastoral Agricultural Research Institute (Palmerston North, NZ) (Appendix 5.1).

To clone the full-length gene sequence from *T. repens* cDNA, two gene-specific forward primers, designated “*TrPAP26* F1” and “*TrPAP26* F2,” were designed based on the partial sequence, and a degenerative reverse primer (3' RACE adapter) was used (Table 2.1). 3' RACE-PCR was used rather than 5' RACE-PCR because the partial EST sequence contained the 5' end of the gene sequence, but not the 3' end. Following PCR amplification of white clover cDNA, gel electrophoresis was performed on the PCR product. Two bands were detected at ~1 kb and ~0.4 kb in size and were excised from the gel (Fig. 3.1A). The purified fragments (Fig. 3.1B) were then cloned into the pGEM vector, and transformed into *E. coli* strain DH5 α .

E. coli colonies were screened for plasmids containing the inserted gene fragment of either 1 kb or 0.4 kb. To do this, plasmid DNA minipreps were performed on cultures from five colonies putatively containing the 1 kb insert and ten colonies putatively containing the 0.4 kb insert, and the resulting purified plasmid DNA sequenced. The consensus sequence from the five sequences of the 1 kb fragment is shown as Appendix 5.2, and the consensus sequence from the ten 0.4 kb fragment sequences is shown as Appendix 5.3.

A BLAST search of the consensus sequences returned 110 hits against the 0.4 kb fragment, many of which had sequence identities of >80% and were identified as acid phosphatase genes. However, the query coverage was only 56% at most

(Table 3.1). For the 1 kb fragment, the BLAST search returned 120 hits, many of which had sequence identities of >80%. These were identified as acid phosphatase genes from various plant species, with a query coverage of up to 75% (Table 3.2). A full list of the BLAST hits is shown as Appendix 5.4.1 for the 0.4 kb fragment and Appendix 5.4.2 for the 1 kb fragment.

A Clustal Omega alignment between the two sequences showed that the 0.4 kb fragment is contained within the 1 kb fragment (Fig. 3.3). There were numerous differences between them including base changes and gaps. However, the sequences are very similar where they do align. A comparison of the 3' RACE primer sequence to the 3' end of the 0.4 kb fragment showed that the primer was able to bind because the sequences were similar enough, producing a false 3' end. Because of this mis-binding, the 0.4 kb fragment was dropped from further use, and focus was placed solely on further characterization of the 1 kb fragment.

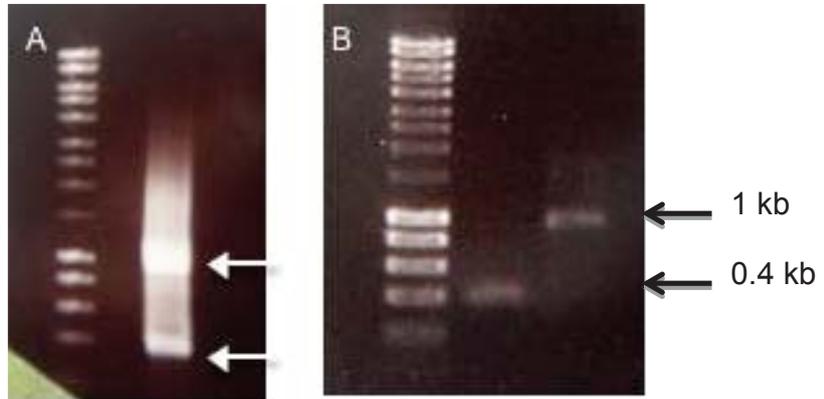


Figure 3.1. PCR products separated using 1% (w/v) agarose gel electrophoresis. A: PCR products from primers, “TrPAP26 F1” and “TrPAP26 F2.” Arrows indicate the bands to be excised from the gel as 1 kb and 0.4 kb. B: Purified bands (arrowed) from A separated using 1% (w/v) agarose gel electrophoresis.

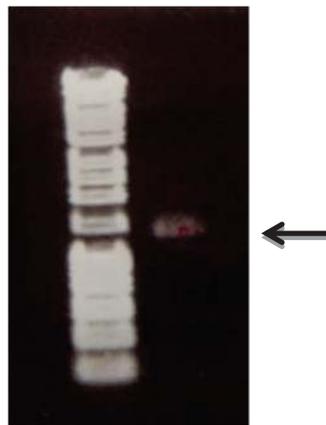


Figure 3.2. Separation using 1% (w/v) agarose gel electrophoresis of a PCR product of *T. repens* cDNA amplified using the gene-specific forward and reverse primers (Table 2.1). The band at ~1.4 kb (arrowed) represents the putative full-length sequence of *TrPAP26*.

Table 3.1. BLAST alignments against the consensus sequence of the 0.4 kb fragment generated using the “TrPAP26 F1” and “TrPAP26 F2” primers (Table 2.1)

Description	Max score	Total score	Query cover	E value	Max ident	Accession
Medicago truncatula chromosome 8 clone mth2-13911, complete sequence	289	289	56%	1e-74	84%	AC157349.7
Medicago truncatula chromosome 8 clone mth2-16a6, complete sequence	289	289	56%	1e-74	84%	AC144657.6
Medicago truncatula clone JCVI-FLMt-20A11 unknown mRNA	145	145	21%	3e-31	92%	BT147925.1
Medicago truncatula Purple acid phosphatase-like protein (MTR_4g115670) mRNA, complete cds	145	145	21%	3e-31	92%	XM_003609387.1
Phaseolus vulgaris gene for acid phosphatase, complete cds	129	129	20%	2e-26	89%	AB116720.1
Phaseolus vulgaris mRNA for acid phosphatase, complete cds	129	129	20%	2e-26	89%	AB116719.1
Lupinus luteus mRNA for acid phosphatase (acPase1 gene)	123	123	20%	1e-24	88%	AJ458943.2
Glycine max cDNA, clone: GMFL01-02-N14	122	122	21%	4e-24	86%	AK244377.1
Glycine max cDNA, clone: GMFL01-36-E23	122	122	21%	4e-24	86%	AK286765.1
Glycine max purple acid phosphatase-like protein (PAP3), mRNA >gb AY151272.1 Glycine max cultivar Union purple acid phosphatase-like protein (Pap3) mRNA, complete cds >gb GQ422778.1 Glycine max clone GmPhy12 phytase mRNA, complete cds	122	122	21%	4e-24	86%	NM_001249748.1
Glycine max bifunctional purple acid phosphatase 26-like (LOC100818438), mRNA >gb GQ422779.1 Glycine max clone GmPhy13 phytase mRNA, complete cds	120	120	20%	1e-23	87%	NM_001252721.1
Glycine max cDNA, clone: GMFL01-08-K22	120	120	20%	1e-23	87%	AK285394.1
Glycine max cultivar Jixian11 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	120	120	20%	1e-23	87%	AY151274.1
Ricinus communis Iron(III)-zinc(II) purple acid phosphatase precursor, putative, mRNA	118	118	20%	4e-23	87%	XM_002530719.1
Populus trichocarpa predicted protein, mRNA	118	118	20%	4e-23	87%	XM_002324582.1
Vitis vinifera, whole genome shotgun sequence, contig VV78X201381.7, clone ENTAV 115	118	118	23%	4e-23	83%	AM466178.1
Citrus maxima acid phosphatase mRNA, partial cds	116	116	20%	2e-22	86%	JN580586.1
Citrus sinensis acid phosphatase mRNA, partial cds	116	116	20%	2e-22	86%	JN580569.1
Populus trichocarpa predicted protein, mRNA	116	116	19%	2e-22	88%	XM_002330986.1
Vitis vinifera contig VV78X026477.7, whole genome shotgun sequence	116	116	21%	2e-22	85%	AM431847.2
Glycine max cultivar Wenfeng7 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	116	116	20%	2e-22	86%	AY151271.1
Glycine max cultivar Mengjin1 purple acid phosphatase-like protein (Pap3) mRNA,	116	116	20%	2e-22	86%	AY151275.1

Gene sequences isolated from other legume species are highlighted in yellow.

Table 3.2. BLAST alignments against the consensus sequence of the 1 kb fragment generated using the “TrPAP26 F1” and “TrPAP26 F2” primers (Table 2.1)

Description	Max score	Total score	Query cover	E value	Max ident	Accession
Medicago truncatula clone JCVI-FLMt-20A11 unknown mRNA	1141	1141	90%	0.0	89%	BT147925.1
Medicago truncatula Purple acid phosphatase-like protein (MTR_4g115670) mRNA, complete cds	1124	1124	86%	0.0	89%	XM_003609387.1
Glycine max cDNA, clone: GMFL01-02-N14	814	814	75%	0.0	85%	AK244377.1
Glycine max cDNA, clone: GMFL01-36-E23	810	810	75%	0.0	85%	AK286765.1
Lupinus luteus mRNA for acid phosphatase (acPase1 gene)	776	776	69%	0.0	86%	AJ458943.2
Glycine max purple acid phosphatase-like protein (PAP3), mRNA >gb AY151272.1 Glycine max cultivar Union purple acid phosphatase-like protein (Pap3) mRNA, complete cds >gb GQ422778.1 Glycine max clone GmPhy12 phytase mRNA, complete cds	704	704	61%	0.0	86%	NM_001249748.1
Glycine max cultivar Wenfeng7 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	699	699	61%	0.0	86%	AY151271.1
Glycine max cultivar Mengjin1 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	699	699	61%	0.0	86%	AY151275.1
Glycine max cultivar Jixian11 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	699	699	61%	0.0	86%	AY151274.1
Glycine max cultivar Zaoshu6 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	699	699	61%	0.0	86%	AY151273.1
Phaseolus vulgaris mRNA for acid phosphatase, complete cds	637	637	53%	4e-179	87%	AB116719.1
Glycine max cDNA, clone: GMFL01-08-K22	600	600	74%	1e-167	78%	AK285394.1
Glycine max bifunctional purple acid phosphatase 26-like (LOC100818438), mRNA >gb GQ422779.1 Glycine max clone GmPhy13 phytase mRNA, complete cds	538	538	53%	3e-149	83%	NM_001252721.1
Populus trichocarpa predicted protein, mRNA	488	488	56%	5e-134	80%	XM_002324582.1
Ricinus communis Iron(III)-zinc(II) purple acid phosphatase precursor, putative, mRNA	484	484	56%	6e-133	79%	XM_002530719.1
Populus trichocarpa predicted protein, mRNA	464	464	56%	6e-127	79%	XM_002330986.1
Jatropha curcas CL2382.Contig2_JC-CK_1A transcribed RNA sequence	452	452	56%	4e-123	78%	GAHK01006622.1
PREDICTED: Fragaria vesca subsp. vesca bifunctional purple acid phosphatase 26-like (LOC101306162), mRNA	452	452	56%	4e-123	78%	XM_004287219.1

Gene sequences isolated from other legume species are highlighted in yellow.


```

0.4 kb -----
1 kb  AGCAAAGATCACGTCAACCTTGTACTTGATGAACCAGCTCTCGAAAACCACTCGCATGCT

0.4 kb -----
1 kb  TTCACCTTCCATATAGTGAGCTTCATTAGTGTTTGTAGAGCGGCACGTGCATGAGAACAA

0.4 kb -----
1 kb  TGAGCCAAGGTGTCTTCTCCCTATCAACCCTCGTCAGCTCTTCACAGAGCCATTTATATT

0.4 kb -----
1 kb  GTGGTGTGTACTTTACGTAGGGGAATAGCTGGGATA

```

Figure 3.3 Clustal multiple sequence alignment of the 0.4 kb and 1 kb fragment.

* = matching base

. = purine to pyrimidine or vice versa base change

: = purine to purine or pyrimidine to pyrimidine base change

- = gap

To clone the full-length *TrPAP26* cDNA sequence, a contig of 1422 bp was made between the 897 bp partial sequence from AgResearch and the newly obtained 1 kb sequence by PCR. A BLAST search of the contig revealed the putative start and stop codons based on the positions of these codons in matches with high sequence identity (not shown). The full contig sequence is shown as Appendix 5.5.

From the contig, gene-specific forward (“ATG F”) and reverse (“TAA R”) primers containing the putative start and stop codons, respectively, were designed (Table 2.1). PCR amplification of *T. repens* cDNA from a mixture of tissues using these primers produced a single band at ~1.4 kb (Fig. 3.2). This band was purified and sequenced and was found to contain the full-length sequence of *TrPAP26* from start codon to stop codon (Appendix 5.6).

3.1.2 Bioinformatic analysis

The full sequence was analyzed using various proteomics tools from www.expasy.org. First, the nucleotide sequence was translated into an amino acid sequence containing 473 amino acids (Fig. 3.4). Using the ProtParam tool, the mass of the protein was found to be 54671 Da (55 kDa) and the theoretical pI was found to be 6.32. Using the MyHits tool to scan for motifs, the protein was revealed to contain a region belonging to the metallophos family and to contain several sites that allow for post-translational modification (i.e. N-glycosylation [three sites]). A predictive phosphorylation tool (Musite) identified a T₇₄ residue (on the mature protein) with a 97% chance of phosphorylation (Table 3.3).

MQGLLFFCSFVFFISIRDGYAGITSSFVRSSEYPSVDIPLDHQ
VFAVPKGHNAPQQVHITQGDYEGKAVIISWV(T)PDEPGSSR
VQFGTSENKFEASAEGTVS((NYT))FGEYKSGYIHHCLVEGL
EHNTKYYYRIGSGDSSREFWFETPPKVDADAPYKFGIIGDM
GQTFNSLSTLEHYIESKAQTVLFLVGDLSYADRYKYTDVGVR
WDSWARFVEKSTAYQPWIWSAGNHEIEYFPYMGEVVPFN
NYLRRYTTPYLASISSSPLWYAIRRASAHIIVLSSYSPYVKYT
PQYKWLSEELTRVDREKTPWLIVLMHVPLYNTNEAHYMEG
ESMRVVFESWFIKYKVDVIFAGHVHAYERSYRFSNVDY((NI
T))GGNRYPVADKSAPVYITVGDGGNQEGLASRFMDPQPEY
SAFREASYGHSTLEIK((NRT))HAVYHWNRNDDGKKVTTDS
*FVLHNQYWGNNRRRRRKLKHYILSVIDEVVSI**

Figure 3.4. The translated sequence of *TrPAP26* from the putative initiating methionine (M) to the stop codon (*). The putative signal peptide is italicized. Potential N-glycosylation sites (N-X-S/T) are indicated with double parentheses, and the potential phosphorylation site (T) predicted using the Musite tool is indicated with parentheses. The metallophos protein domain is underlined. The five blocks of amino acid sequences conserved among PAPs are highlighted.

Table 3.3. Predicted features of the translated TrPAP26 sequence

Tool	Results
ProtParam	Mass: 54671 Da (55kDa) Theoretical pI: 6.32
MyHits	Metallophos family Three N-glycosylation sites
Musite	Phosphorylation at T74

3.2 Biochemical characterization of recombinant TrPAP26

Some studies fused GST with various *Arabidopsis* purple acid phosphatase proteins and expressed them in *E. coli* (Kuang et al., 2009; Wang et al., 2011; Zhang et al., 2008; Zhu et al., 2005). Two of the studies were successful in producing active recombinant proteins using this method, and were able to study the characteristics of their respective proteins without much change in activity between the recombinant form and the native form of the protein (Wang et al., 2011; Zhu et al., 2005). All the studies mentioned above used the full-length protein, including the signal peptide. Therefore, the full-length protein was used first in this project to assay the characteristics of recombinant TrPAP26.

3.2.1 Protein expression with the construct *TrPAP26::GST*

To clone the cDNA sequence including the signal sequence into the pGEX expression plasmid using PCR, a short linker sequence (GACGAC) and an EcoRI site (GAATCA) were added immediately upstream of the start codon. Similarly, a Sall (GTCGAC) site and a short linker sequence (GACGAC) were added to the sequence immediately downstream of the stop codon (the primers used to accomplish this were “pGEX F” and “pGEX R”, Table 1). After PCR amplification, the *TrPAP26* cDNA with flanking short linker sequences and digestion sites was ligated into the pGEX-6P-3 vector and the plasmid transformed into *E. coli* BL21 for expression. IPTG was added to induce expression of TrPAP26::GST, and transformed *E.coli* incubated without IPTG was used as the control.

After induction, cells were lysed and centrifuged. The resulting supernatant contains soluble protein (designated the soluble fraction), while the pellet contains recombinant protein that has been incorporated into membrane-based inclusion bodies (designated the insoluble fraction). Any recombinant protein in the soluble fraction was purified using glutathione-sepharose and either eluted with glutathione to yield the TrPAP26::GST fusion protein (predicted mass ~80 kDa) or

by treatment of the resin with PreScission Protease® to yield TrPAP26 (predicted mass 55 kDa). Separation of these fractions using SDS-PAGE and subsequent Coomassie blue staining revealed that a prominent band of ~72 kDa was evident in the insoluble fraction only (Fig. 3.5, lane 2).

No bands of similar mass could be detected by Coomassie blue staining in the soluble fraction (Fig 3.5, lane 1) or after glutathione-sepharose purification (Fig. 3.5, lanes 3-5). However, more evidence that the induced 72 kDa protein in the insoluble fraction was TrPAP26::GST was obtained by performing SDS-PAGE on the insoluble and soluble fractions of *E. coli* that had been incubated either with or without IPTG (Fig. 3.6). The 72 kDa protein was most prominent in extracts from cells treated with IPTG.

A different strain of *E. coli*, Origami (DE3), which allows for the formation of disulfide bonds in eukaryotic proteins, was then used to determine if this background would result in the fusion protein in the soluble fraction, and therefore make it extractable. Unfortunately, the same results as with *E. coli* strain BL21 were observed, i.e. the fusion protein was found only in the insoluble fraction (data not shown).

There was the possibility that there was a low concentration of the TrPAP26::GST fusion protein in the soluble fraction that is below the limit of detection using Coomassie blue staining. To determine this, western blots were done using a primary antibody raised in rabbits against AtPAP26 (i.e. raised against glycosylated protein; Veljanovski et al 2006) (Fig. 3.7A) and another primary antibody raised in rabbits against recombinant AtPAP12 (i.e. raised against non-glycosylated protein; described in Tran et al., 2010) (Fig. 3.7B). In *Arabidopsis*, the anti-AtPAP12 antibody is known to cross-react with both the 60 kDa AtPAP12 subunits and 55 kDa AtPAP26 subunits while the anti-AtPAP26 antibody is more specific for AtPAP26. Before applying the anti-AtPAP26 antibody, the polyvinylidene fluoride (PVDF) membrane with proteins transferred onto it was

treated with sodium periodate to eliminate any cross-reactivity of the glycosyl groups with the antibody. This treatment is not necessary with the anti-AtPAP12 antibody since that antibody was raised against a recombinant protein and therefore will not cross-react with glycosyl groups (Tran, 2010). However, in agreement with the Coomassie blue stains, the western blots showed a prominent band of 72 kDa in the insoluble fraction and no bands in the soluble extracts with both antibodies (Fig. 3.7A and 3.7B).

The anti-AtPAP12 antibody seemed to cross-react with many proteins in the insoluble and soluble fractions, whereas the anti-AtPAP26 antibody displayed much less cross-reactivity to other proteins. The anti-AtPAP26 antibody also seemed to react strongly with a protein of ~25 kDa, which is most likely to be an *E. coli* protein since it occurs in the samples both with and without the IPTG treatment.

Finally, to determine if only a small amount of TrPAP26 was in the soluble fraction (but undetectable by antibody), an acid phosphatase activity assay was also performed. No phosphatase activity was detectable, further confirming that TrPAP26 had not been extracted from the insoluble fraction (Fig. 3.8).

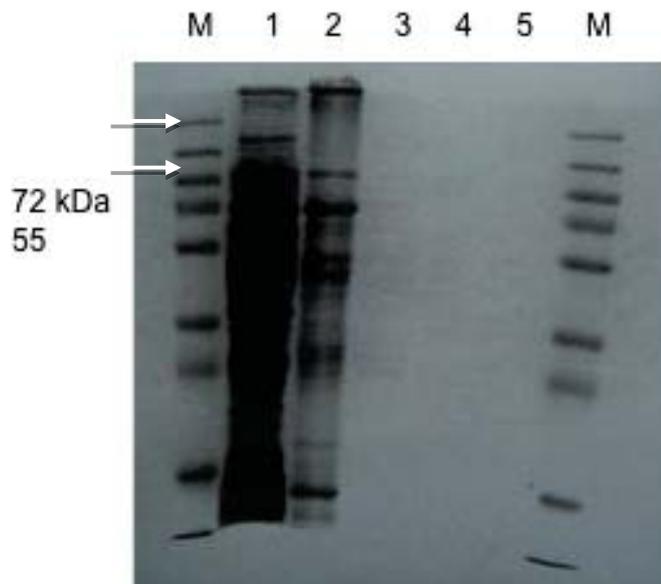


Figure 3.5. Coomassie blue stain of the induced expression of TrPAP26 in *E. coli* strain BL21 after separation using 12% SDS-PAGE. Lane M- marker; 1- soluble fraction; 2- insoluble fraction; 3- purified GST-TrPAP26; 4- purified GST-TrPAP26; 5- purified TrPAP26. Bands were expected at ~55 kDa (TrPAP26) and ~80 kDa (TrPAP26::GST).

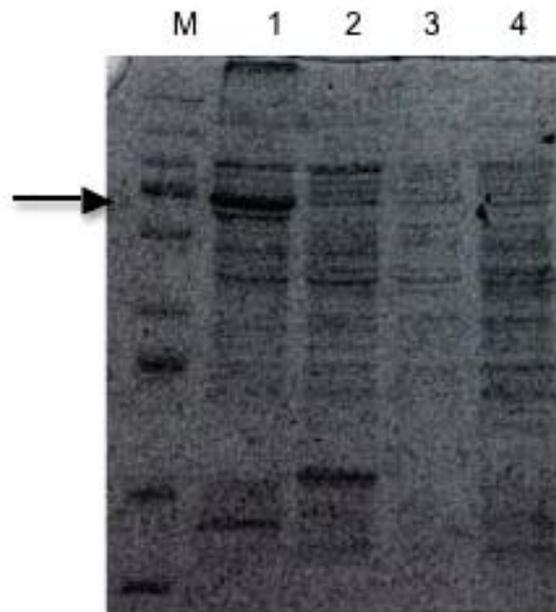


Figure 3.6. Coomassie blue stain of the induced (+IPTG) or non-induced (-IPTG) expression of TrPAP26::GST in *E. coli* strain BL21 after separation using 12% (w/v) SDS-PAGE. Lane M- marker; 1- insoluble fraction with IPTG; 2- insoluble fraction without IPTG; 3- soluble fraction with IPTG; 4- soluble fraction without IPTG. Arrow indicates the predicted mass of TrPAP26::GST.

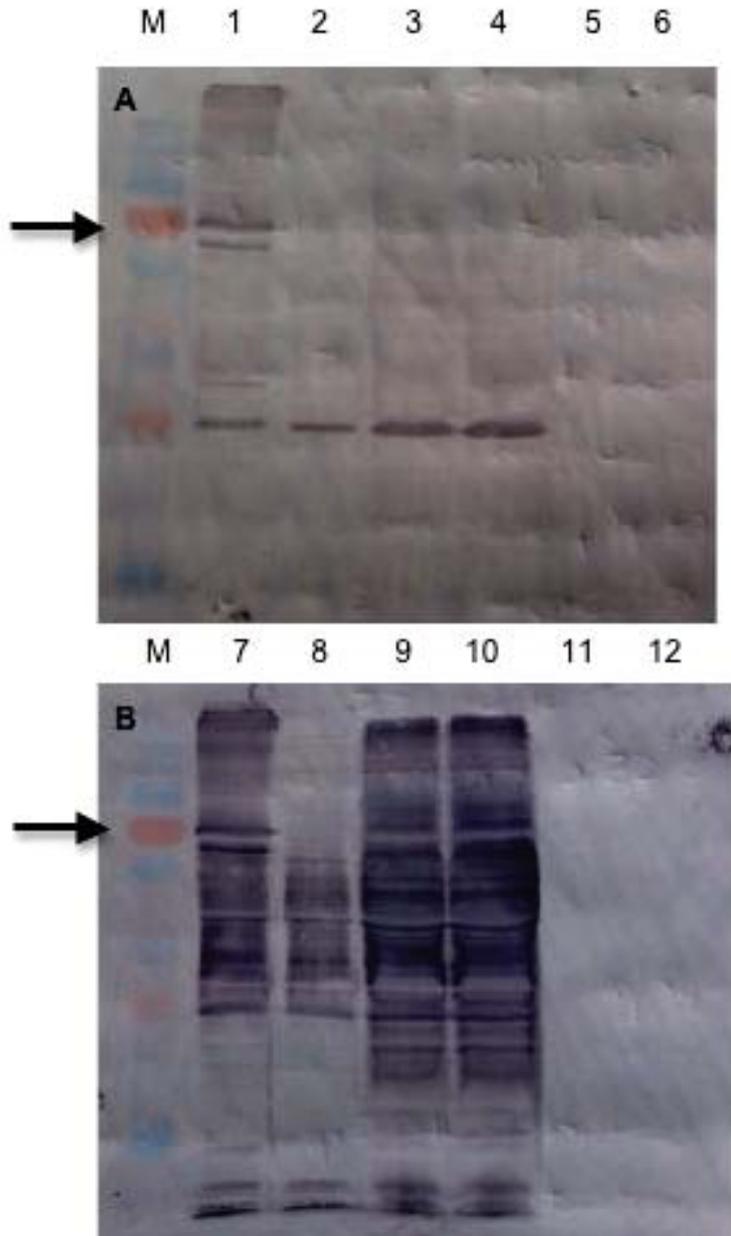


Figure 3.7. Separation of TrPAP26::GST expressed in *E. coli* strain origami using 12% SDS-PAGE followed by western blotting using the anti-AtPAP26 antibody (A) or anti-AtPAP12 antibody (B). Antibody detection was visualized using colorimetric detection.

A: Lane M- marker; 1- insoluble fraction (+IPTG); 2- insoluble fraction (-IPTG); 3- soluble fraction (+IPTG); 4- soluble fraction (-IPTG); 5- TrPAP26::GST; 6- TrPAP26. Arrow indicates TrPAP26::GST.

B: M- marker; 7- insoluble fraction (+IPTG); 8- insoluble fraction (-IPTG); 9- soluble fraction (+IPTG); 10- soluble fraction (-IPTG); 11- TrPAP26::GST; 12- TrPAP26. Arrow indicates TrPAP26::GST.

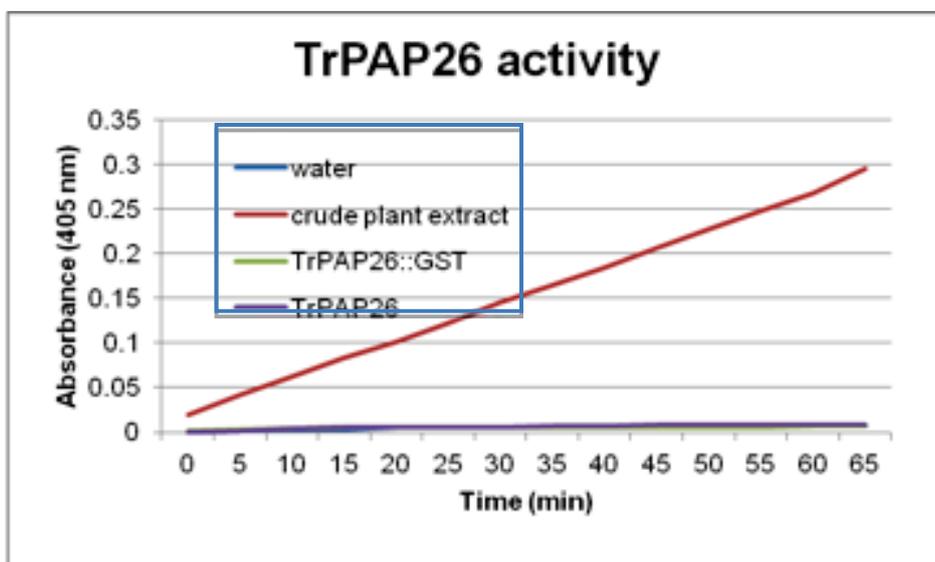


Figure 3.8. Acid phosphatase assay on TrPAP26::GST, TrPAP26, and a crude plant extract activity, as indicated. Activity was measured using pNPP as the phosphate source.

3.2.2 Protein expression with the construct, Δ sp-TrPAP26::GST

Since the previous construct, which made no modifications to the gene sequence itself, proved unfruitful in terms of yielding recombinant proteins that could be assayed, a second construct was used that contained a deleted signal peptide portion (incorporating the first 63 bp of the gene sequence). This was done because recombinant proteins without a signal peptide will more likely accumulate in the cytoplasm rather than in inclusion bodies (de Marco, 2009). It has also been shown that removing the signal peptide does not affect the activity of the recombinant proteins (Campos et al., 2008).

To do this, the sequence was flanked with a linker DNA sequence (GACGAC), an EcoRI site at the 5' end, and a Sall site at the 3' end. The primers including these sequences are "pGEX Δ sp F" and "pGEX R" (Table 2.1). After PCR, the modified cDNA was then ligated into the pGEX-6P-3 vector and transformed into the *E. coli* strain Origami. Following induction with IPTG, protein purification of the soluble fraction was again carried out using glutathione-sepharose and either elution with glutathione to yield Δ sp-TrPAP26::GST or treatment with the Precision protease to release Δ sp-TrPAP26. Western blots were then carried out using both the anti-AtPAP26 and anti-AtPAP12 antibodies (Fig. 3.9A and 3.9B, respectively) and both blots did show bands of the expected mass (~70 kDa for Δ sp-TrPAP26::GST and ~45 kDa for Δ sp-TrPAP26) in the glutathione-sepharose purified fractions (Fig. 3.9A, lanes 5, 6; Fig. 3.9B, lanes 11, 12). A third band of ~57 kDa was also recognised in the extracts both with and without GST. Additionally, as with the full-length construct, there was also a band of ~25 kDa that showed prominently in the western blots with both antibodies. This band appears in both the soluble and insoluble fractions in samples both treated with IPTG and without. However it was not detectable in the purified extracts.

Though the anti-AtPAP12 antibody cross-reacts with many other proteins (as seen in Fig. 3.7A and 3.7B), it also cross-reacts much more strongly with TrPAP26

when compared with the anti-AtPAP26 antibody. Therefore, the anti-AtPAP12 antibody was the preferred one for use in all other western blots from this point on in the project.

To confirm that the ~ 70 kDa protein was Δ sp-TrPAP26::GST and the ~ 45 kDa protein was Δ sp-TrPAP26 (and to identify the ~57 kDa protein), SDS-PAGE followed by Coomassie blue staining of the glutathione-sepharose purified fractions was performed. The three stained bands were excised from the gel and sent to the Centre for Protein Research (University of Otago, Dunedin, NZ) for sequencing by MALDI-TOF/TOF.

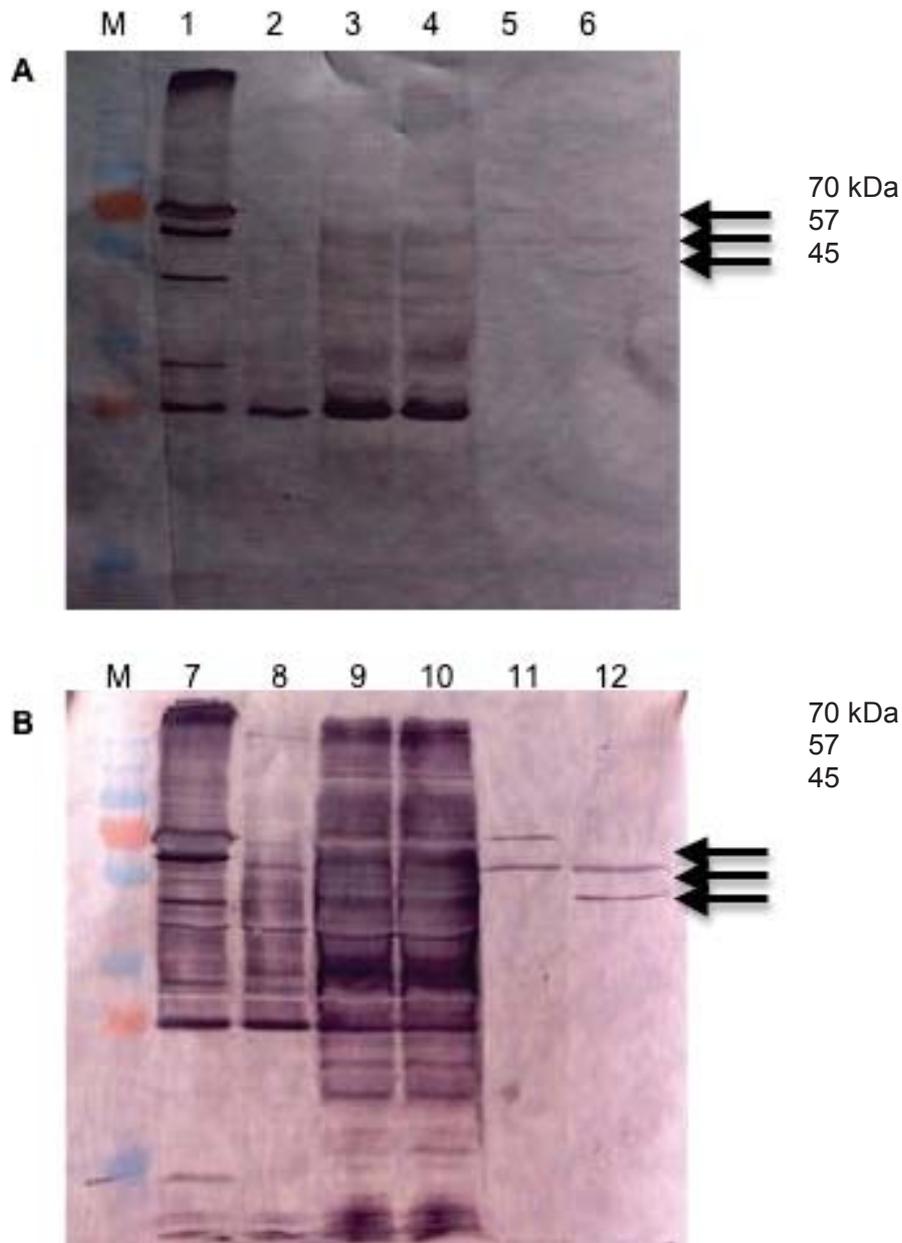


Figure 3.9. Separation of Δ sp-TrPAP26 expressed in *E. coli* origami using 12% SDS-PAGE followed by western blotting using the anti-AtPAP26 antibody (A) or anti-AtPAP12 antibody (B). Antibody detection was visualized using colorimetric detection.

A: Lane M- marker; 1- insoluble fraction with IPTG; 2- insoluble fraction without IPTG; 3- soluble fraction with IPTG; 4- soluble fraction without IPTG; 5- Δ sp-TrPAP26::GST; 6- Δ sp-TrPAP26. Arrows indicate the bands to be sequenced.

B: Lane M- marker; 7- insoluble fraction with IPTG; 8- insoluble fraction without IPTG; 9- soluble fraction with IPTG; 10- soluble fraction without IPTG; 11- Δ sp-TrPAP26::GST; 12- Δ sp-TrPAP26. Arrows indicate the bands to be sequenced.

3.2.3 Identification of recombinant Δ sp-TrPAP26 using MALDI-TOF/TOF

The ~72 kDa band that occurs only in the purified extract eluted with glutathione (the putative TrPAP26::GST fusion protein) (Fig. 3.9B, lane 11) was matched to two proteins: glutathione S-transferase (from an unidentified cloning vector) and purple acid phosphatase (from *Medicago truncatula*). There was 47% sequence coverage of the translated sequence of *TrPAP26* (Fig. 3.10).

The ~57 kDa band that occurs in both extracts (Fig. 3.9B, lanes 11 and 12) was matched to the GroEL-GroES-(ADP)₇ chaperonin complex from *E. coli*. There was 52% sequence coverage for that complex (Fig. 3.11).

The ~45 kDa band that occurs only in the purified extract eluted with the protease (Fig. 3.9B, lane 12) was matched to purple acid phosphatase. There was 52% sequence coverage of the translated sequence of *TrPAP26* (Fig. 3.12).

The BLAST alignments for all three sequences are shown in Appendices 5.7, 5.8, and 5.9, respectively.

However, despite the occurrence of the Δ sp-TrPAP26 protein in the purified soluble fraction, no acid phosphatase activity was associated with this protein under the assay conditions used (Fig. 3.13).

```

1  MQXLLFFCSF  VFFISIRDGY  AGITSSFVRS  EYPSXDIPLD  HQVFAVXKGH
51  NAPQQXHITQ  GDYEGKAVII  SWVTPDEPGS  SRXQFGTSEN  KFEASAEGTV
101 SNYTFGEYKX  GYIHHCLVEG  LEHNTKYYR  IGSGDSSREF  WFETPPKVDA
151 DAPYKFGIIG  DMGQTFNLS  TLEHYIESKA  QTVLFGDLS  YADRYKYTDV
201 GVRWDSWARF  VEKSTAYQPW  IWTGNHEIE  YFPYMGEVVP  FNNYLRRYTT
251 PYLASNSSSP  LWYAIRRASA  HIIVLSSYSP  YVKYTPQYKW  LSEELTRVDR
301 EKTPWLIVLM  HVPLYNTNEA  HMEGESMRV  VFESWFIKYK  VDVIFAGHVH
351 AYERSYRFSN  VDYNITGGNR  YPVADKSAPV  YITVGDGGNQ  EGLASRFMDP
401 QPEYSAFREA  SYGHSTLEIK  NRTHAVYHWN  RNDDGKKVTT  DSFVLHNQYW
451 GNNRRRRRKLK  HYILSVIDEV  VSI

```

Figure 3.10. Alignment of the tryptic peptide sequences generated from the purified ~72 kDa band (Fig. 3.8B, lane 11). The peptide sequence shown is the *TrPAP26* translated sequence (Fig. 3.3), and the matching tryptic peptides are shown in bold.

```

1  AAKDVKFGND  ARVKMLRGVN  VLADAVKVTL  GPKGRNVVLD  KSGGAPTITK
51  DGVSVAREIE  LEDKFENMGA  QMVKEVASKA  NDAAGDGTIT  ATVLAQAIIT
101 EGLKAVAAGM  NPMDLKRIGD  KAVTAAVEEL  KALSVPCSDS  KAIAQVGTIS
151 ANSDETVGKL  IAEAMDKVGK  EGVITVEDGT  GLQDEL DVVE  GMQFDRGYLS
201 PYFINKPETG  AVELESPFIL  LADKKISNIR  EMLPVLEAVA  KAGKPLLIIA
251 EDVEGEALAT  LVVNTMRGIV  KVAAVKAPGF  GDRRKAMLQD  IATLTGGTVI
301 SEEIGMELEK  ATLEDLGQAK  RVVINKDTT  IIDGVGEEAA  IQGRVAQIRQ
351 QIEEATSDYD  REKLQERVAK  LAGGVAVIKV  GAATEVEMKE  KVARVEDALH
401 ATRAAVEEGV  VAGGGVALIR  VASKLADLRG  QNEDQNVGIK  VALRAMEAPL
451 RQIVLNCGEE  PSVVANTVKG  GDGNYGYNAA  TEEYGNMIDM  GILDPTKVTR
501 SALQYAASVA  GLMITTECMV  TDLPKNDAAD  LGAAGGMGGM  GGMGMM

```

Figure 3.11. Alignment of the tryptic peptide sequences generated from the purified ~57 kDa band (Fig. 3.8B, lanes 11 and 12). The peptide sequence shown is the *GroEL* sequence and the matching tryptic peptides are shown in bold.

```

1  MQXLLFFCSF VFFISIRDGY AGITSSFVRS EYPSXDIPLD HQVFAVXKGH
51  NAPQQXHITQ GDYEGKAVII SWVTPDEPGS SRXQFGTSEN KFEASAEGTV
101 SNYTFGEYKX GYIHHCLVEG LEHNTKYYYR IGSGDSSREF WFETPPKVDA
151 DAPYKFGIIG DMGQTFNSLS TLEHYIESKA QTVLFVGDLS YADRYKYTDV
201 GVRWDSWARF VEKSTAYQPW IWSGNEHEIE YFPYMGEVVP FNNYLRRYTT
251 PYLASNSSSP LWYAIRRASA HIIVLSSYSP YVKYTPQYKW LSEELTRVDR
301 EKTPWLIVLM HVPLYNTNEA HMEGESMRV VFESWFIKYK VDVIFAGHVH
351 AYERSYRFSN VDYNITGGNR YPVADKSAPV YITVGDGGNQ EGLASRFMDP
401 QPEYSAFREA SYGHSTLEIK NRTHAVYHWN RNDDGKKVTT DSFVLEHQYW
451 GNNRRRRRKLK HYILSVIDEV VSI

```

Figure 3.12. Alignment of the tryptic peptide sequences generated from the 45 kDa band (Fig. 3.8B, lane 12). The peptide sequence shown is the *TrPAP26* translated sequence (Fig. 3.3), and the matching tryptic peptides are shown in bold.

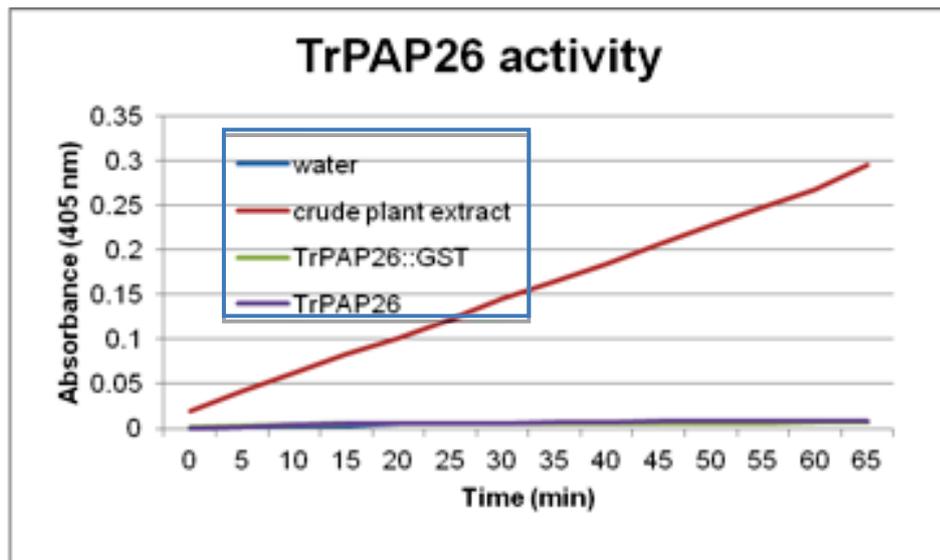


Figure 3.13. Acid phosphatase assay of Δ sp-TrPAP26::GST and Δ sp-TrPAP26 as glutathione-sepharose purified proteins, as indicated.

3.3 Biochemical characterization of purified TrPAP26

Since the recombinant protein produced in a prokaryotic background (*E. coli*) was inactive and mostly insoluble, a eukaryotic background (*Nicotiana tabacum* [tobacco]) was then used. Expression of the gene was driven by the cauliflower mosaic virus (CaMV) 35S promoter. A six residue histidine tag was added to the C-terminus of the protein to allow for affinity purification.

3.3.1 Transforming tobacco with the construct, *35S::TrPAP26::6xHis*

Using PCR, an *Xba*I site (TCTAGA) was added immediately upstream of the start codon and immediately downstream of the stop codon of *TrPAP26* (the primers used to accomplish this were designated “6xHis-PAP26 F” and “6xHis-PAP26 R,” Table 2.1). The modified sequence was then digested with *Xba*I and ligated into the pDAH2 vector, which had also been digested with *Xba*I. The *pDAH2-TrPAP26* vector was then digested with *Not*I, producing *35S::TrPAP26*, which was then ligated into the pART27 vector, which had also been digested with *Not*I. The final vector, *pART27-TrPAP26*, housed the CaMV 35S promoter upstream of the full length *TrPAP26* gene with six histidine residues downstream (*35S::TrPAP26::6xHis*). This vector was transformed into *Agrobacterium tumefaciens* and subsequently into wild type tobacco (sections 2.4.10 and 2.4.11).

Inoculated tobacco leaf discs were grown on media containing kanamycin to select for those that had taken up the vector and are expressing the *neomycin-kanamycin phosphotransferase* gene (Fig. 3.14A), while other discs (non-inoculated) were grown on media that did not contain kanamycin to provide a set of control plants (Fig. 3.14 B2-B4). One plate of control leaf discs was grown on media with kanamycin to ensure that the antibiotic did kill any discs that did not have the inserted construct (Fig. 3.14 B1). One shoot from each callus was subcultured and allowed to grow roots to produce a total of 14 plants comprising 7 transformed and 7 (non-transformed) controls. Once the new plants were well

established in soil, a leaf was excised from each, and DNA, RNA, and protein (soluble and cell wall fractions) were extracted.

3.3.2 Integration of *35S::TrPAP26::6xHis* and expression of *TrPAP26::6xHis* in tobacco

DNA was isolated from seven putative transformed lines and seven control lines and genomic PCR was performed (section 2.4.5) with an equal amount (20 ng) of DNA per sample using the “35s F” and “6xHis PAP26 R” primers (Table 2.1), and the PCR products were subjected to agarose gel electrophoresis. All the putative transgenic plants were deemed to be positive because they had a DNA insert of ~1.8 kb, although the band line 1 difficult to see in the digital image selected (Fig. 3.15, lane 2).

RNA was isolated from seven putative transformed lines and three control lines, reverse-transcribed to cDNA, and then semi-quantitative PCR used to determine the relative level of expression of *TrPAP26::6xHis* (section 2.4.5.1). To do this, “PAP26 int seq F” and “6xHis PAP26 R” primers were used, with *GAPDH* used as the housekeeping gene, (primers “NtGAPDH qPCR F” and “NtGAPDH qPCR R”; Table 2.1). Following gel electrophoresis of the PCR products, a nearly opposite pattern to that seen in the genomic DNA PCR product was observed (Fig. 3.16). The intensity of the bands observed was quantified and plotted against the quantified intensity of *GAPDH* expression for each sample (Fig. 3.17). Line 1 of the transgenic plants had the highest relative expression level, showing a 3.8-fold difference to the *GAPDH* expression. However, the use of qPCR would more accurately distinguish between differences in expression.

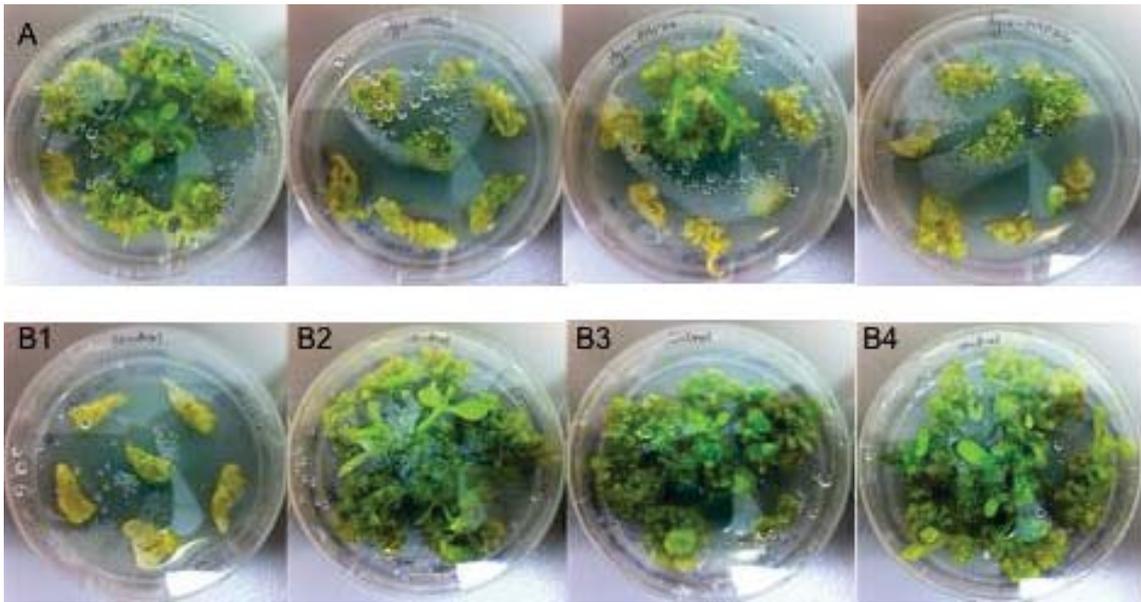


Figure 3.14. Development of tobacco leaf pieces 21 days after inoculation with *Agrobacterium tumefaciens*. A) Tobacco transformed with *A. tumefaciens* containing *35S::TrPAP26::6xHis* growing on media supplemented with kanamycin. B1) Untransformed tobacco growing on media supplemented with kanamycin; B2-4) Untransformed tobacco growing on media without kanamycin.



Figure 3.15. Separation of PCR products amplified from genomic DNA isolated from putative transgene lines or control lines of tobacco, as indicated. Lanes 1-8 are the putative transgenic plants. Lanes C2-C10 are the control plants. "+" is the *pART27-TrPAP26* plasmid diluted 100-fold. "-" is water. The arrow indicates bands ~1.8 kb in size representing *35S::TrPAP26::6xHis*.



Figure 3.16. Separation of PCR products amplified from cDNA isolated from putative transgene lines or control lines of tobacco, as indicated. Lanes 1-8 are the putative transgenic plants; Lanes C6-C10 are the control plants. *TrPAP26* denotes products amplified using primers specific for *TrPAP26*, while *GAPDH* denotes products amplified with *GAPDH*-specific primers.

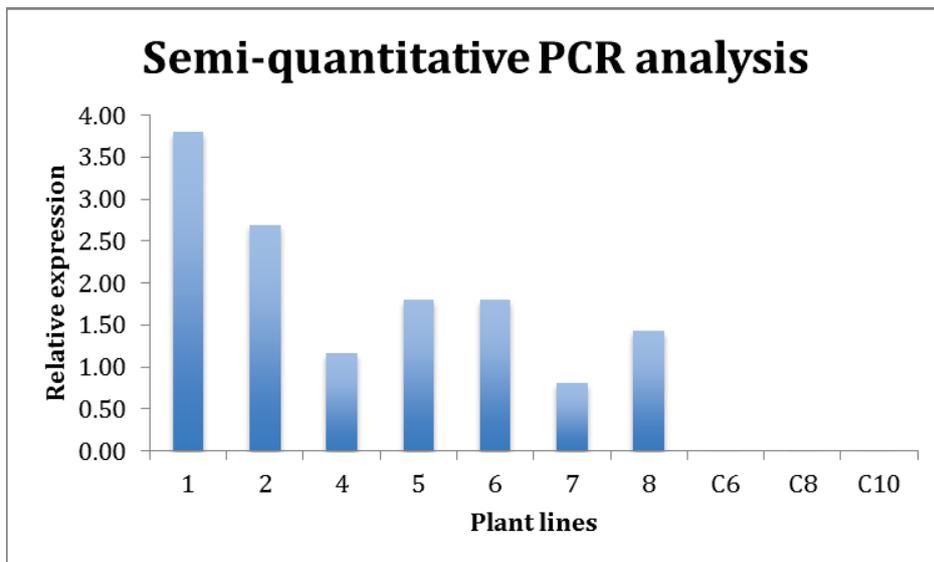


Figure 3.17. Analysis of *TrPAP26* expression using semi-quantitative PCR. Expression of the transgene is relative to the expression of the housekeeping gene, *GAPDH*, as determined from the PCR shown in Figure 3.15.

3.3.3 Assessment of TrPAP26::6xHis accumulated in transgenic tobacco lines

Using whole leaf protein extracts, an acid phosphatase activity assay was carried out (Fig. 3.18). Activity in the extracts isolated from the cell wall (wall fraction) had an average of 0.237 umoles pNP per ug protein from the transgenic plants against an average of 0.116 umoles pNP per ug protein from the control plants (Fig. 3.18A). Activity in the protein extracts isolated from the cytoplasm (soluble fraction) had an average of 0.171 umoles pNP per ug protein from the transgenic plants versus an average of 0.017 umoles pNP per ug protein from the control plants (Fig. 3.18B). This assay showed that there was a difference in activity between the transgenic and control plants, indicating that the transgenic plants were putatively accumulating TrPAP26::6xHis.

Western blots using the anti-PAP12 antibody were also performed on these same protein extracts (Fig. 3.19). The blots showed two bands in the transgenics: one at ~49 kDa and one at ~44 kDa, with the ~49 kDa band being more prominent. Interestingly, faint bands of the same mass were also seen in control lines C6 and C10.

Based on the results from the acid phosphatase activity assay and western blots, it was decided to purify the TrPAP26::6xHis from the soluble fraction of transgenic line 8 because it showed higher acid phosphatase activity and produced the most prominent bands in the western blot. The control line C8 was used as the control because it showed low acid phosphatase activity and didn't produce any bands in the western blot.

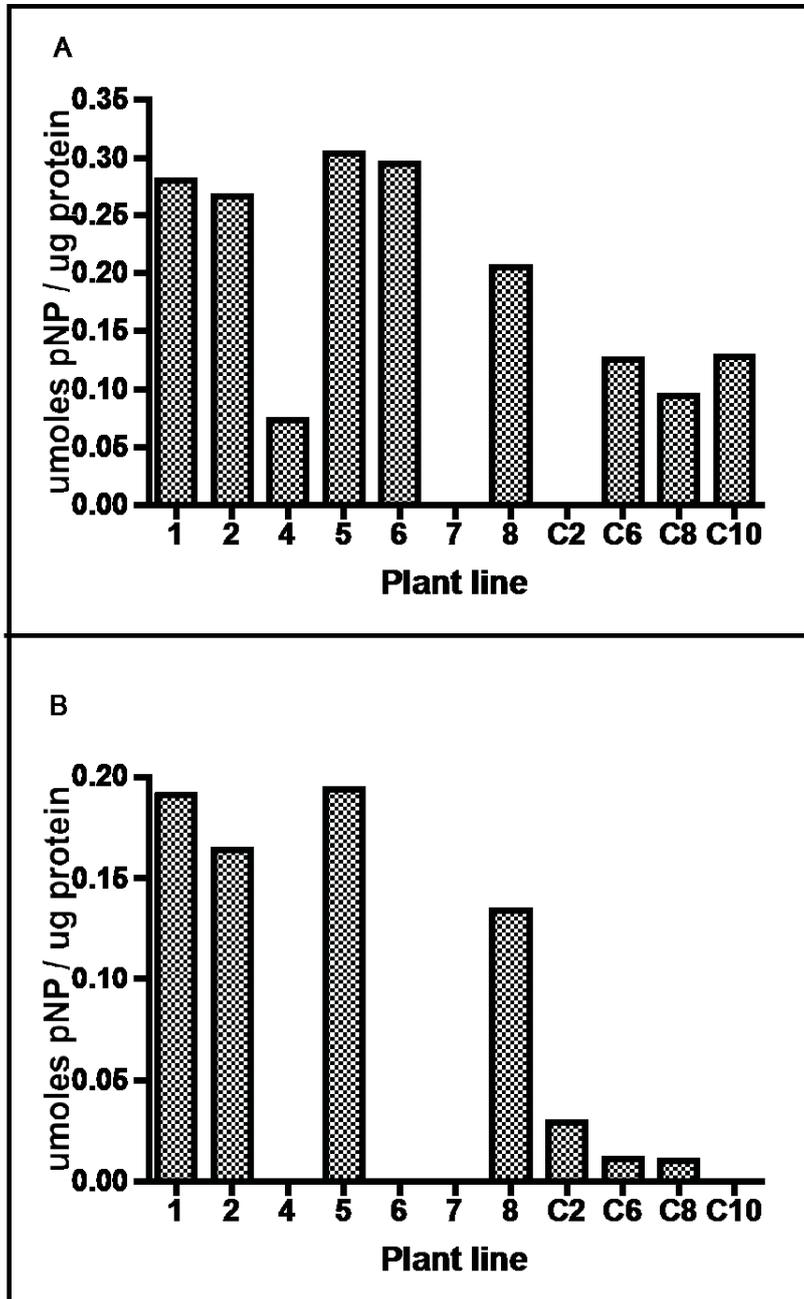


Figure 3.18. Acid phosphatase determined in the wall fraction (A) or soluble fraction (B) in the transgenic and control (C2-C10) tobacco lines, as indicated. Values expressed are the mean of technical duplicates.

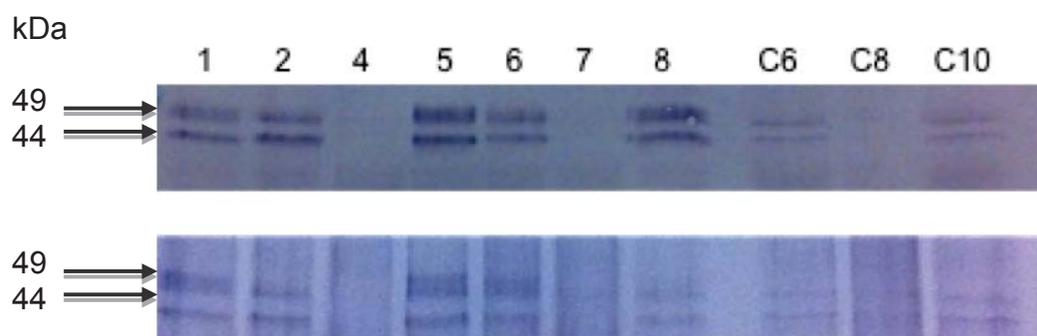


Figure 3.19. Separation of soluble (top panel) or cell wall (bottom panel) protein extracts from putative transgenic plants (lanes 1-8) or control plants (C6, C8, C10) using 12% (w/v) SDS-PAGE, following by western blotting using the anti-AtPAP12 antibody. The arrows indicate bands at ~49 and ~44 kDa. Antibody binding was visualized using colorimetric detection.

3.3.4 Affinity-purification of TrPAP26::6xHis protein from tobacco

Leaf tissue excised from both transgenic line 8 and control line C8 was subjected to the same purification treatment: soluble protein extraction via 1 mM DTT, ammonium sulfate precipitation to separate protein from other cellular components, dialysis to remove salts and ammonium sulfate, and finally affinity purification of histidine-rich protein via nickel-bound Chelating Sepharose Fast Flow beads.

A comparative acid phosphatase activity assay of the soluble protein extract and affinity-purified protein extract showed that the activity had been enriched in both the transgenic and the control line when compared with the activity in the crude extract (Fig. 3.20).

To monitor the purification, SDS-PAGE and Coomassie staining were used to visualize proteins bound or unbound to the Chelating Sepharose Fast Flow beads (Fig. 3.21). The SDS-PAGE revealed that proteins, other than that expected at ~50 kDa, had been eluted from the beads in both the transgenic line 8 and the control line C8.

A western blot was performed to determine if TrPAP26::6xHis was present before and after purification (Fig. 3.22). Further, an aliquot of the Chelating Sepharose Fast Flow beads was also included in the western blot to confirm that TrPAP26::6xHis was actually bound to the beads, as well as an aliquot of the supernatant following protein binding to see if more TrPAP26::6xHis was unbound than bound to the beads. As seen previously (Fig. 3.19), there were two bands, one at ~49 kDa and one at ~44 kDa. Both bands were seen in the protein extracts from both transgenic and control lines. However, the ~49 kDa band was more prominent in the transgenic line 8, and it was very faint in the control line C8. This observation suggests that TrPAP26::6xHis had been enriched in the transgenic tobacco background, but also extracted along with other tobacco proteins

including tobacco acid phosphatases that may have been similar enough in sequence to cross-react with the anti-AtPAP12 antibody.

3.3.5 Activity of partially purified TrPAP26::6xHis

A study on AtPAP26 (Plaxton et al, 2006) showed that the protein's activity was stimulated with the addition of $MgCl_2$. Therefore, the partially purified TrPAP26::6xHis was also incubated in pNPP either with or without $MgCl_2$ to see if there was any change in activity (Fig. 3.21). Contrary to what was expected, the addition of $MgCl_2$ caused a slight decrease in TrPAP26::6xHis activity. The same was also seen for the acid phosphatase activity in the control.

The partially purified TrPAP26::6xHis protein after affinity chromatography was then incubated with various substrates to ascertain whether TrPAP26::6xHis has a broad or a narrow substrate specificity (Table 3.4). The results of this assay suggest that TrPAP26::6xHis has a broad substrate specificity because there was a high amount of phosphates released with four out of the five phosphate substrates tested as compared to the control. The exception was phytic acid.

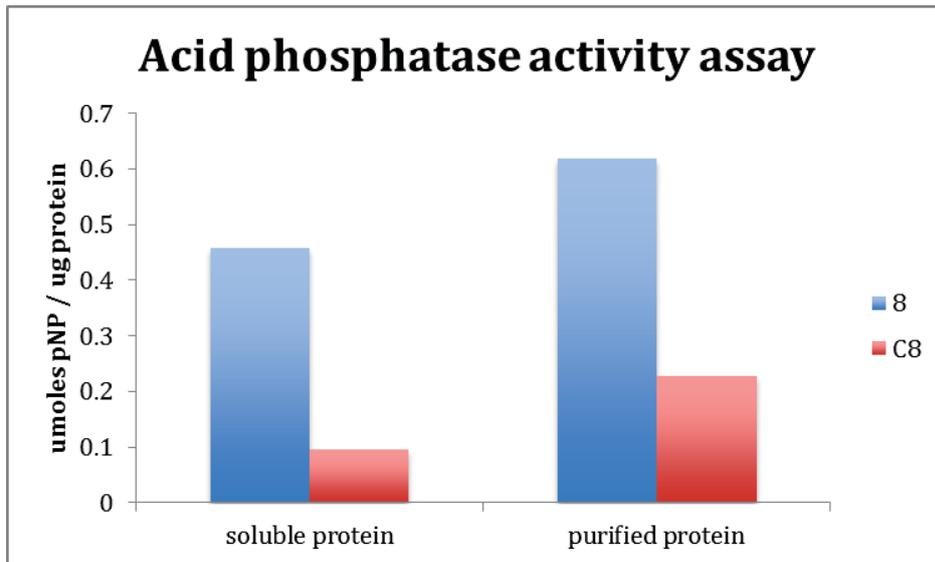


Figure 3.20. Acid phosphatase activity assay of the soluble protein and affinity-purified protein extracts, as indicated, from either the transgenic line 8 or the control line 8, as indicated. Each value represents the mean of technical duplicates.

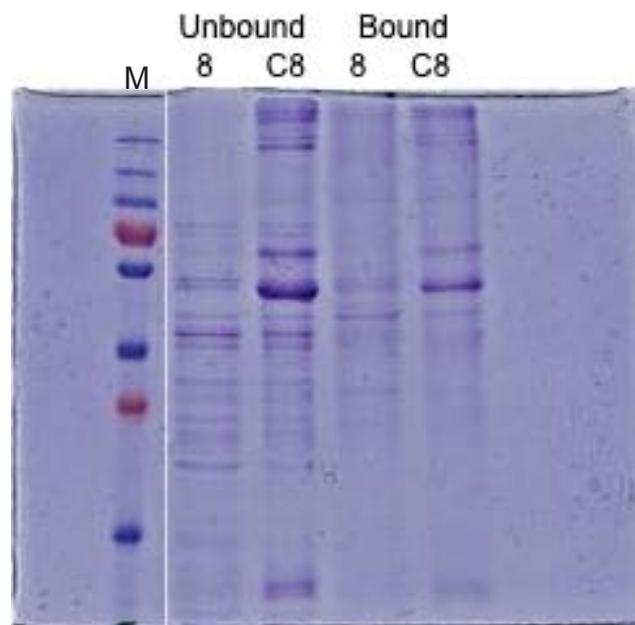


Figure 3.21. Separated proteins from transgenic plant line 8 or control plant line C8, as indicated, that were either bound to the chelating sepharose or remained unbound in the supernatant. Proteins were separated using 12% (w/v) SDS-PAGE. The separated proteins were visualized using Coomassie blue staining. Lane M – molecular mass ladder.

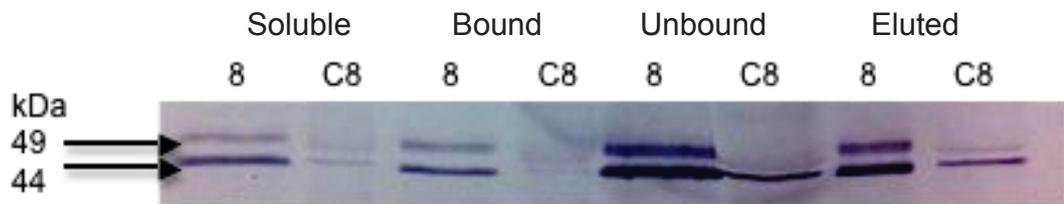


Figure 3.22. Separation of protein extracts from transgenic plant line 8 and control plant line C8, as indicated, using 12% (w/v) SDS-PAGE and western blotting using the anti-AtPAP12 antibody. Antibody binding was visualized using colorimetric detection.

“Soluble” denotes leaf protein extracted with 1 mM DTT.

“Bound” denotes soluble protein extracts partially purified with ammonium sulfate and desalting that had bound to the chelating sepharose resin.

“Unbound” denotes protein that did not bind to the column.

“Eluted” denotes protein that had been eluted from the chelating sepharose using elution buffer containing 50 mM imidazole.

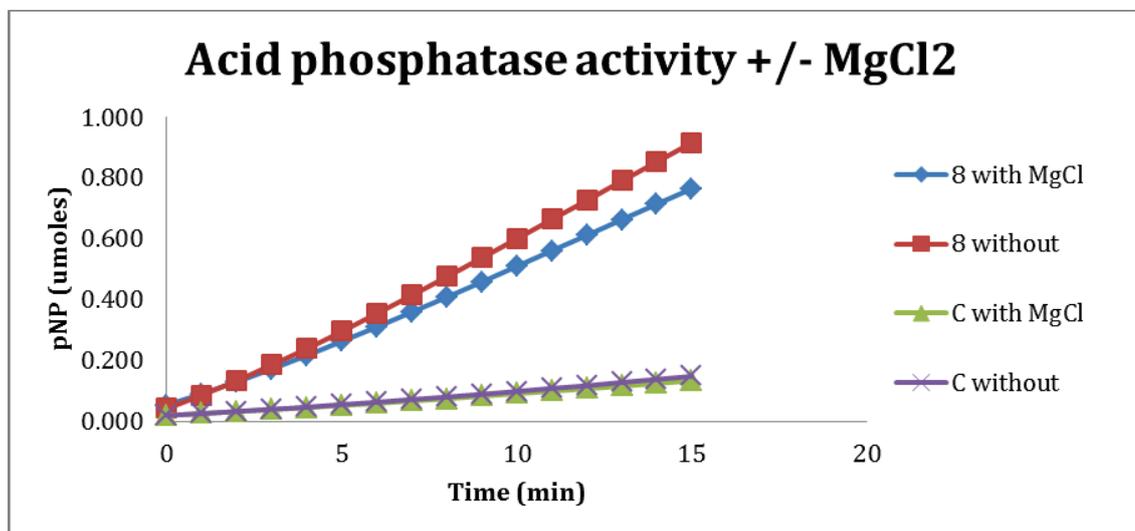


Figure 3.23. Acid phosphatase activity assay with or without MgCl₂. Affinity-purified transgenic line 8 and control line C8 were incubated with pNPP either with or without MgCl₂. Each point represents the mean of two technical replicates.

Table 3.4 Substrate specificity assay using the affinity-purified protein from transgenic line 8 and control line C8.

Sample	Substrate	uM phosphate/ug protein
8	ATP	538.9
C	ATP	83.8
8	PEP	466.3
C	PEP	84.7
8	Phytic acid	5.2
C	Phytic acid	7.3
8	Na pyrophosphate	342.3
C	Na pyrophosphate	49.4
8	pNPP	388.8
C	pNPP	64.6

Values represent the mean of three technical replicates.

3.4 Expression of *TrPAP26* and accumulation of TrPAP26 in response to changes in Pi supply *in planta*

To extend the characterization of the *PAP26-like* gene in white clover, the expression of *TrPAP26* and accumulation of TrPAP26 was studied, *in planta*, in response to changes in phosphate (Pi) supply.

3.4.1 White clover grown in short-term phosphate-deficient (-Pi) conditions

Four-node white clover cuttings, where two nodes were submerged in liquid nutrient media to induce root formation, were grown in Pi-sufficient (1 mM Pi; +Pi) or low Pi (10 μ M Pi; -Pi) for either 24 hours or 7 days. Total RNA was then extracted from the root subtending from the third node and analyzed using qPCR with the primers “TrPAP26 qPCR F1” and “TrPAP26 qPCR R1” (Table 2.1). No significant change occurred in *TrPAP26* expression over the 24-hour treatment in plants grown in the -Pi media as compared to plants grown in complete media was observed (Fig. 3.24). Over the 7-day treatment, the *TrPAP26* expression was significantly ($P < 0.05$) down-regulated from day 0 in the +Pi treatment at day 1, 3, and 5, and in the -Pi treatment at days 3, 5 and 7. Significant ($P < 0.05$) differences between treatments occurred at day 3 and at day 7, where the expression of *TrPAP26* in the +Pi treatment was also significantly higher than day 0 (Fig. 3.25).

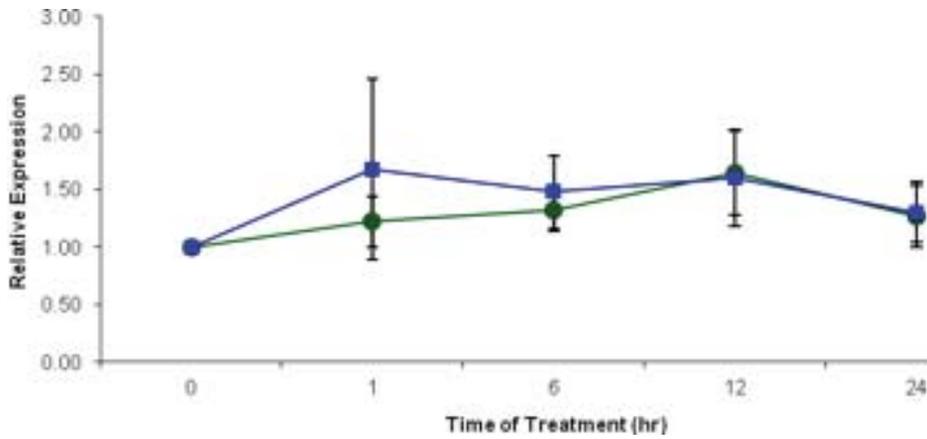


Figure 3.24. *TrPAP26* expression in the roots of plants grown in +Pi or -Pi media, as indicated, for 24 hours. n = 3. ±SEM. Green circle= +Pi; blue square= -Pi

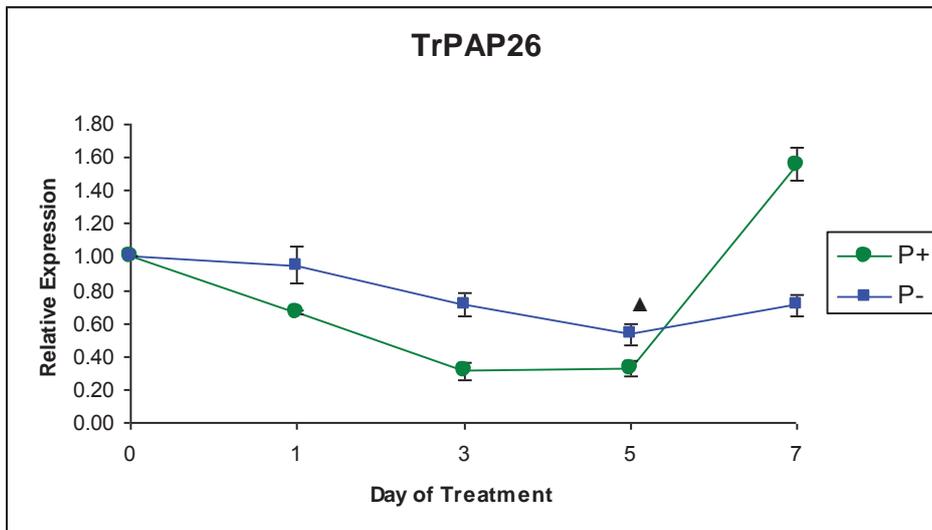


Figure 3.25. *TrPAP26* expression in the roots of plants grown in +Pi or -Pi media, as indicated, for 7 days. n = 3. ±SEM. Green circle= +Pi; blue square = -Pi
 * denotes a significant difference between treatments at P<0.05; ▲ denotes a significant difference for that treatment and day from day 0 at P<0.05.

3.4.2 White clover grown in long-term phosphate-deficient (-Pi) conditions

Experiments with four-node white clover stolon cuttings, maintained in the +Pi or – Pi media were continued for 20 days to examine longer-term changes in *TrPAP26* expression and TrPAP26 accumulation. On the 16th and 18th days, a root surface acid phosphatase activity assay was performed on the elongation zone of the primary root subtending from the third node. These assays, at the two different time points, showed a significantly higher activity in plants grown in –Pi than in those grown in +Pi (Fig. 3.26), and thus indicating that the –Pi treatment was having an effect.

On the 21st day, the following tissues were harvested: the first fully expanded leaf (F), the mature leaf at the third node from the apex (M), the internodes between the first fully expanded leaf and the mature leaf (I), and the primary root subtending from the original third node (R). RNA was extracted from all the tissues and reverse-transcribed to cDNA. Protein was extracted in two fractions, soluble and cell wall, from the F, M, and R. There was insufficient internode (I) tissue to conduct any protein analysis.

Using the cDNA, qPCR was performed to measure the expression levels of *TrPAP26* in the various tissues, with *actin* and *GAPDH* used as the housekeeping genes. The results showed no difference in *TrPAP26* expression in the first fully expanded leaf, a significantly ($P < 0.05$) higher expression in the +Pi plants in the mature leaf, a significantly ($P < 0.05$) higher expression in the –Pi plants in the internodes, and no difference in expression in the root (Fig. 3.27).

Using protein extracts from these tissues, two assays were performed. In terms of acid phosphatase activity, there was little to no difference for most of the extracts (Fig. 3.28). In the soluble extract for the root and in the cell wall extract for the first fully expanded leaf, there was no detectable activity in the –Pi samples, the opposite of what was observed when the elongation zone was assayed using the root surface assay (Fig. 3.28).

A western blot using the anti-AtPAP12 antibody did not recognize any proteins in the cell wall extract (Fig. 3.29, bottom panel). For the soluble extract, accumulation of TrPAP26 was observed in the –Pi plants for the first fully expanded leaf, in the +Pi plants for the mature leaf, and no difference in accumulation in the internodes and the root (Fig. 3.29, top panel).

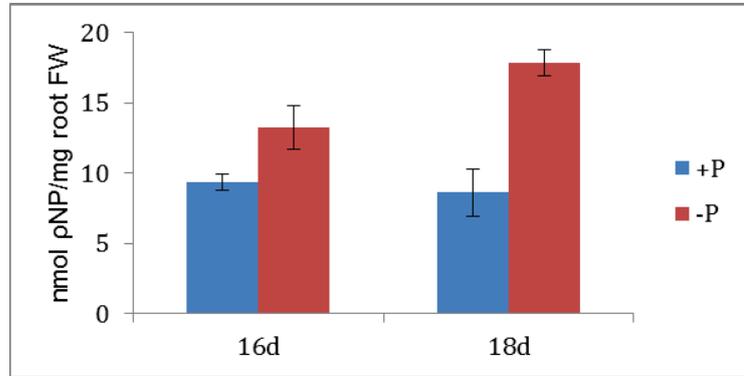


Figure 3.26. Acid phosphatase activity of whole elongation zone roots of plants grown in +Pi or –Pi media, as indicated, for 16 or 18 days. n = 3, ±SEM.

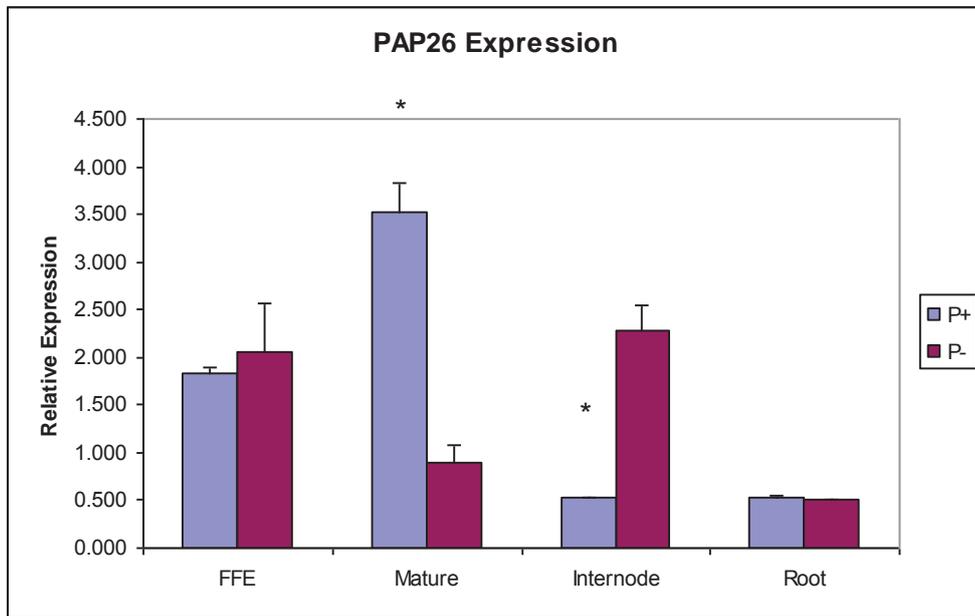


Figure 3.27. qPCR of *TrPAP26* expression in various white clover tissues grown in +Pi or –Pi media, as indicated, for 21 days. F- first fully expanded leaf, M- mature leaf, I- internodes, and R- root. Expression levels are relative to *actin* and *GAPDH*. n = 3, ±SEM. * denote a significant difference at P<0.05 between the two treatments in the respective tissue.

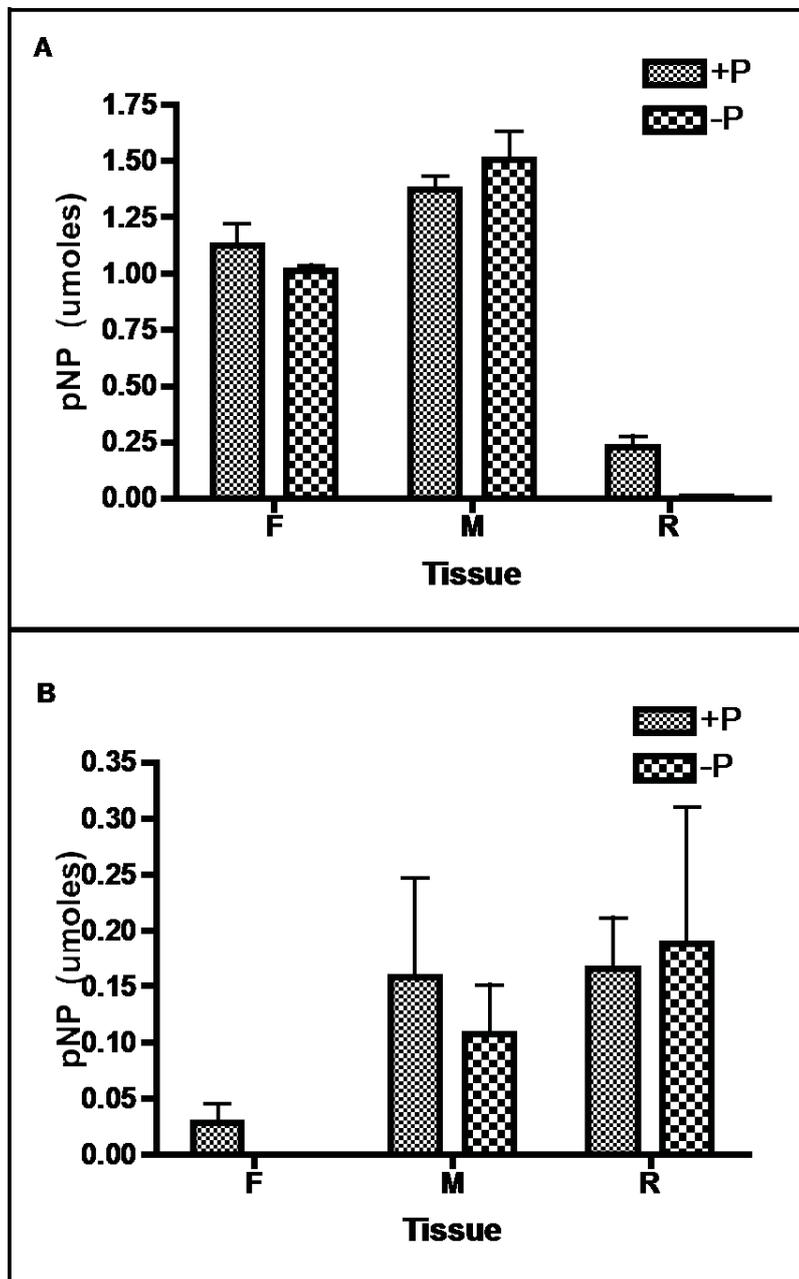


Figure 3.28. Acid phosphatase activity assay of soluble (A) and cell wall protein extracts (B) from the tissues, as indicated, and subjected to +Pi or -Pi treatment, as indicated. F) first fully expanded leaf; M) mature leaf; and R) root. n = 6. Values are the means of technical triplicates.

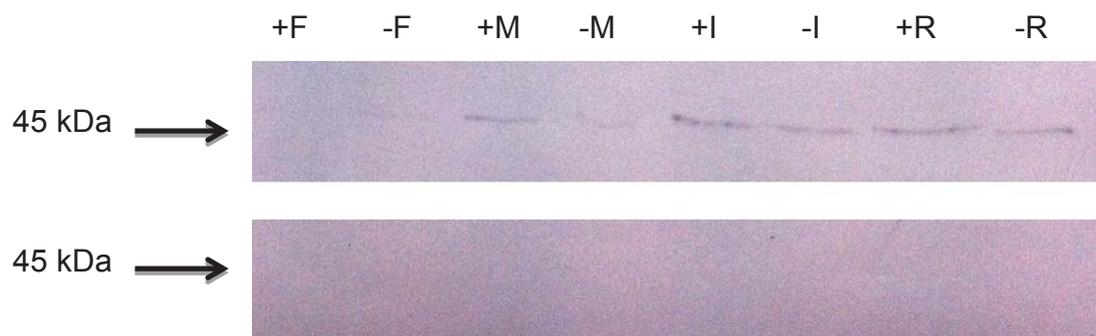


Figure 3.29. Separation of protein extracts from white clover grown in +Pi or -Pi media, as indicated by “+” or “-”, respectively, using 12% (w/v) SDS-PAGE and western blotting using the anti-AtPAP12 antibody. Antibody binding was visualized using colorimetric detection. Top panel: soluble extract. Bottom panel: cell wall extract.

“F” denotes soluble protein extracted from the first fully expanded leaf.

“M” denotes soluble protein extracted from the mature leaf.

“I” denotes soluble protein extracted from the internodes.

“R” denotes soluble protein extracted from the root.

4. Discussion

4.1 Overview

Purple acid phosphatases (PAPs) are important proteins and have been identified in mammals, plants, fungi, and bacteria. While there are sequence similarities across the kingdoms, their physiological functions differ. For example, in mammals, PAPs have been found to have a role in transporting iron, immunological functions, and in bone metabolism (Bresciani and Von Figura, 1996; Flanagan et al., 2006; Hayman and Cox, 1994; Valizadeh, 2004; Vincent and Averill, 1990). In plants, PAPs function in intracellular and extracellular phosphorus acquisition by hydrolyzing phosphate-rich esterified compounds to release a free phosphate group (Olczak et al., 2003).

Within the plant kingdom, there is a high degree of conservation in both sequence and function of these proteins between genera and species. Li et al (2002) identified at least 29 PAPs in *Arabidopsis thaliana* by screening the *Arabidopsis* genome database for the conserved sequence motifs that had been previously defined for PAPs (Schenk et al., 2000). The fact that these proteins are highly conserved suggested that the white clover (*Trifolium repens* L.) genome would also contain PAPs, although no white clover PAP gene sequences have been documented yet.

Using the sequence from a well-characterized *Arabidopsis* PAP, *AtPAP26*, a partial EST sequence from a white clover EST database at AgResearch Grasslands EST database that had high identity to *AtPAP26* was gifted to the project. Initially, research in this thesis focused on cloning the full-length *TrPAP26* gene.

Once cloned, the aims of the project were to determine the role of *TrPAP26* in

white clover. The steps taken to achieve this were:

- clone the gene to determine the full-length sequence,
- express the recombinant protein in *Escherichia coli*,
- express the recombinant protein in *Nicotiana tabacum* (tobacco),
- analyze the purified recombinant protein, and
- analyze the expression pattern of the gene and activity of the protein in white clover supplied with different levels of inorganic phosphate (Pi).

4.2 Isolation of an *AtPAP26*-like gene sequence from *Trifolium repens* (white clover)

The *TrPAP26* full-length sequence translates to a 473 amino acid polypeptide with all five conserved motifs, three N-glycosylation sites, and a predicted mass of 55 kDa, which together identify the protein as a purple acid phosphatase belonging to group I (Li et al., 2002). Furthermore, a comparison between the *TrPAP26* amino acid sequence and *AtPAP26* showed a 78% identity. The signal peptides, comprising the N-terminal 21 amino acid residues in *TrPAP26* and 22 residues in *AtPAP26*, had no significant similarities, indicating that these two proteins may be targeted to different cellular compartments. Aside from the signal peptide, the rest of the sequences were highly conserved (81% identity). The metallophos motif was conserved, as were the five blocks of amino acid residues that are characteristic of all PAPs. This high degree of similarity suggests that *TrPAP26* is likely to be an *AtPAP26* ortholog with similar functions and characteristics. Since *AtPAP26* was found to be glycosylated, it is very likely that *TrPAP26* will also undergo similar post-translational modifications that have been reported for the *Arabidopsis* protein (Veljanovski et al., 2006).

While comparisons with *AtPAP26* are revealing, *Medicago truncatula* is a legume whose genome is very similar to that of white clover, and so comparisons with this species may be more valid. Only two PAPs have thus far been found in *Medicago*, *MtPAP1* and *MtPHY1* (Ma et al., 2009). A comparison between the amino acid

sequences showed that MtPAP1 has a 61% identity with TrPAP26, while MtPHY1 has only a 29% identity with TrPAP26. MtPAP1 has a 78% identity with AtPAP10, which is also a group Ia-2 PAP in common with AtPAP26, and was also found to aid in phosphorus acquisition during phosphate starvation (Xiao et al., 2006). MtPHY1 was found to have high phytase activity, so its low sequence identity with TrPAP26 suggests that TrPAP26 doesn't function as a phytase (Ma, Wright et al. 2009).

Overall, the white clover gene, *TrPAP26*, is very much *AtPAP26*-like and all characteristics of the sequence, as a 473 amino acid polypeptide, containing all five PAP conserved amino acid blocks, and N-glycosylation sites, point to the conclusion that it is a PAP. The recombinant protein was therefore expressed in *Escherichia coli* to determine if TrPAP26 also functioned as an acid phosphatase.

4.3 Recombinant TrPAP26 expressed in *E. coli*

E. coli is very commonly used as a genetic background to produce heterologous protein due to its short reproduction cycle, ability to grow on cheap growth medium, extensively characterized genetics, and a large number of vectors and strains available to researchers. However, it has known problems in producing properly folded heterologous proteins that are soluble and active (Baneyx, 1999). Despite these shortcomings and with no guarantee of success, *E. coli* is still used because it is such a versatile means of protein production.

A few studies have been successful in producing active plant PAPs in an *E. coli* background (Wang et al., 2011; Zhang et al., 2008; Zhu et al., 2005). Other studies have attempted to express PAPs in *E. coli*, but reported either inactive enzymes and/or insolubility (Dionisio et al., 2011; Kuang et al., 2009). Therefore, *E. coli* was tried initially in this project to produce TrPAP26.

4.3.1 Protein expression with the construct *TrPAP26::GST*

This project used two strains of *E. coli*, BL21 and Origami, and induced expression with IPTG. BL21 is a standard strain for use in expressing heterologous protein. Origami (DE3) is a strain engineered to produce a background that produces proteins with disulfide bridges. Although AtPAP26 is a monomer (Tran et al., 2010), AtPAP10 and AtPAP12 have both been found to be homodimers whose subunits are connected by a disulfide bridge (Tran et al., 2010; Wang et al., 2011). All three are group Ia-2 PAPs, which suggests that there is a possibility that TrPAP26 is also a homodimer that requires a disulfide bridge to assemble.

Nozach et al (2013) showed that it was possible to produce soluble, active disulfide-rich proteins in both BL21 and Origami, although BL21 was more efficient despite its reducing cytoplasm. Lehmann et al (2003) were able to express a disulfide-rich protein of plant origin (peanut) in origami. However, to date, there are no publications in which the expression of a PAP of any origin has been attempted in Origami.

With BL21, it could be seen that *TrPAP26* expression was induced but the protein (~72 kDa) was insoluble (Fig. 3.6). This could be due to the failure of the disulfide bridge formation to connect the subunits if TrPAP26 is a homodimer in the *E. coli* BL21 background (Baneyx, 1999). With this in mind, the *E. coli* Origami strain was used instead. However, the same result occurred: protein expression was induced but the protein was insoluble (Fig. 3.7).

Following expression in Origami, the anti-AtPAP26 and anti-AtPAP12 antibodies were used to detect TrPAP26 accumulation. Both antibodies have been found to cross-react with AtPAP26 so it was likely that they would cross-react with TrPAP26 as well. AtPAP12 is a 60 kDa protein that is also up-regulated during Pi starvation, and the anti-AtPAP12 antibody is reported to cross-react with an AtPAP26 subunit (Tran et al., 2010). Both antibodies detected proteins of ~72 kDa

in the insoluble fraction, just as they did when BL21 was used as the background for expression. While the anti-AtPAP12 antibody cross-reacted with many more proteins than the anti-AtPAP26 antibody, it recognized TrPAP26 more strongly than the anti-AtPAP26 antibody did. The fact that the antibodies detected the protein at ~72 kDa even in a strain of *E. coli* that is able to add disulfide bridges (origami) implies that TrPAP26 is a monomer in common with AtPAP26.

The insoluble protein recognized by the antibodies was ~72 kDa, and given the mass of GST is 26 kDa, this suggests that the observed mass of TrPAP26 is ~46 kDa, which is smaller than its predicted mass of 55 kDa. The SDS-PAGE has been repeated many times, with the same results each time suggesting that the size discrepancy was not due to experimental error. A possible explanation for this could be that TrPAP26 is first synthesized as a preprotein and then cleaved to produce the active form. This discrepancy between the predicted and observed mass has also been noted in some mammalian PAPs, which had been found to require proteolysis before becoming active (Lång and Andersson, 2005; Mitić et al., 2005).

It is not clear why the protein remained insoluble even in Origami, but a reason could be incorrect cellular targeting. TrPAP26 has a signal sequence and eukaryotic signal sequences are not always interpreted correctly in *E. coli* and so the heterologous proteins are inserted into inclusion bodies (Baneyx, 1999). Therefore, a second attempt was made to express *TrPAP26* in *E. coli* to accumulate soluble protein.

4.3.2 Protein expression with the construct Δ sp-*TrPAP26::GST*

The protein that accumulated from the expression of this new construct lacking a signal sequence was poorly soluble, but enough could be purified using glutathione sepharose 4B beads. Following purification, the anti-AtPAP26 and anti-AtPAP12 antibodies both detected proteins of three masses: ~70, ~57, and

~45 kDa.

MALDI-TOF/TOF analysis showed that the ~70 kDa band had an amino acid sequence that matched with GST and TrPAP26, the sequence of the ~57 kDa band matched with GroEL-GroES-(ADP)₇ chaperonin complex from *E. coli*, and the sequence of the ~45 kDa band matched with TrPAP26.

The identity of the ~70 and ~45 kDa bands as PAPs were expected are the sizes that would correspond to Δ sp-TrPAP26::GST and Δ sp-TrPAP26 cleaved from GST, respectively, as either glutathione (to elute the Δ sp-TrPAP26::GST) or the protease (to elute TrPAP26) was used to treat the glutathione sepharose. This further suggests that the observed mass of TrPAP26 is ~45 kDa as opposed to its predicted mass of 55 kDa. Because Δ sp-TrPAP26::GST could be purified, the ~10 kDa loss is likely to be from the C-terminal of the protein, since GST is fused to the N-terminal and required for purification using glutathione sepharose. The mechanism by which this cleavage may occur was not investigated as part of this thesis.

As for the ~57 kDa band, it seems that both antibodies used were sufficiently non-specific to recognize the 57 kDa subunit of GroEL, an *E. coli* protein that is not closely related to any PAP at all (Braig et al., 1994). However, the GroEL-GroES complex plays an important role in the folding of heterologous proteins in *E. coli*, and it has even been recommended to overexpress this chaperonin to aid in heterologous protein folding and solubility (Baneyx, 1999). Thus the GroEL chaperonin may not have completely disassociated from the Δ sp-TrPAP26::GST that it was folding and so was co-purified in the process but did subsequently disassociate using SDS-PAGE. Co-purification of chaperonins is a common occurrence in protein purification, and the GroEL chaperonin has been found to be co-purified with a plant protein fused to GST previously (Joseph and Andreotti, 2008; Keresztessy et al., 1996).

An alternative suggestion is that because the anti-AtPAP12 antibody was raised against recombinant protein, some chaperone proteins were also purified as background contaminants when AtPAP12 was prepared as an antigen. However, the AtPAP26 antibody was raised against purified enzyme from plants and so free from any chaperones, further confirming that the chaperones have formed a complex in *E. coli*.

Performing an acid phosphatase assay on the purified protein using pNPP as the substrate revealed no detectable activity for either Δ sp-TrPAP26::GST or Δ sp-TrPAP26. This was a disappointing result, but not entirely unexpected since previous studies have also reported inactive protein after purification (Dionisio et al., 2011; Kuang et al., 2009). The protein may be inactive because proteins are not glycosylated in the prokaryotic *E. coli* background (Baneyx, 1999), and it has been shown that N-glycosylation contributes to PAP stability and activity (Olczak and Olczak, 2007; Wang et al., 2005). Olczak and Olczak (2007) showed that if the plant PAP could be successfully secreted, then its activity was unaffected even if it lacked glycosyl groups. The studies that were successful in producing soluble, active PAPs may have been fortunate in having chosen PAPs where N-glycosylation was not as important to the stability of the enzyme. However, none of those studies have tried mutating the N-glycosylation sites in their respective PAPs to show definitively that N-glycosylation is not necessary for enzyme stability and activity.

In summary, over the attempts made as part of this thesis, active recombinant TrPAP26 did not accumulate in the *E. coli* backgrounds tested. If this project is continued, Sf9 cells (from the insect *Spodoptera frugiperda*) could be a good alternative to producing TrPAP26 because several studies have successfully used insect Sf9 cells for the expression of active PAPs (Hayman and Cox, 1994; Hur et al., 2010; Kaija et al., 2002; Mitić et al., 2005; Olczak and Olczak, 2007; Wang et al., 2005; Waratrujiwong et al., 2006). As eukaryotic organisms themselves, insects have all the necessary machinery to produce the post-translational

modifications specific to eukaryotic proteins. In addition, no PAPs have been found in *S. frugiperda*, which makes these cells an excellent background for PAP expression because the purified PAP cannot be confused for a native insect PAP.

4.4 Recombinant TrPAP26 expressed in tobacco

Since the *E. coli* prokaryotic protein production system was unsuccessful, tobacco was used instead because this species is very easily transformed, produces a large amount of biomass, and as a eukaryote would be expected to perform any necessary post-translational modifications. Additionally, there should be no trouble in expressing active PAPs in tobacco, since endogenous PAP genes have been cloned from this species (Kaida et al., 2008; Kaija et al., 2002; Lung et al., 2008). The 35S cauliflower mosaic virus promoter was added upstream of the full-length reading frame to drive constitutive expression, and to aid in purification, a six-histidine residue tag was added to the C-terminal just prior to the stop codon. In other studies, a 6xHis tag has been added to PAPs with no detrimental effect to enzyme activity (Dionisio et al., 2011; Kuang et al., 2009; Zhang et al., 2008). Kuang et al (2009) expressed AtPAP15::6xHis in tobacco successfully (i.e. obtained a soluble, active protein), so there was a high expectation that TrPAP26::6xHis would also accumulate in tobacco as an active protein.

4.4.1 35S::TrPAP26::6xHis transcript levels

Seven transgenic tobacco lines (1, 2, 4, 5, 6, 7, and 8) and four wild-type control tobacco lines (C2, C6, C8, and C10) were analyzed for the insertion of 35S::TrPAP26::6xHis into the genome. All seven lines were confirmed as containing the 35S::TrPAP26::6xHis gene. Semi-quantitative PCR revealed that line 1 had the highest transcript levels, while line 7 had the lowest (Fig. 3.16). However, semi-quantitative PCR requires rigorous linearity checks to ensure that band intensity is reliable and show individual true differences in expression. In this thesis, the technique was used simply to confirm mRNA expression in the

transgenic lines. Alternatively, qPCR should be used if a detailed comparison of any consistent differences in expression between lines is to be undertaken.

4.4.2 TrPAP26::6xHis accumulation

Whole leaf protein extracts showed that the transgenic lines had approximately 2-fold higher acid phosphatase activity than the control lines in the cell wall fraction, and ten-fold more in the soluble fraction. This indicates that TrPAP26::6xHis had accumulated in both protein fractions, but more so in the soluble fraction, which may mean that it is preferentially targeted to the vacuole. It is known that AtPAP26 is targeted towards both the vacuole and the extracellular matrix (Hurley et al., 2010). However, the sequence of the AtPAP26 signalling peptide is significantly different from that of TrPAP26, so while AtPAP26 is secreted, TrPAP26 may not be.

The anti-AtPAP12 antibody detected proteins of two masses, ~49 and ~44 kDa, in both the soluble and the cell wall protein extracts (Fig. 3.19). When *TrPAP26* was expressed in *E. coli*, the antibody detected one protein at ~45 kDa. The 4 kDa size discrepancy between the ~45 kDa protein that accumulated in *E. coli* and the ~49 kDa protein that accumulated in tobacco is most likely due to the addition of glycosyl groups in the tobacco background. The glycosyl groups are likely to be complex-type oligosaccharides containing xylose, fucose, and/or mannose as these are the ones previously identified for a plant PAP (Olczak and Watorek, 1998). Each oligosaccharide can have a mass of up to ~1 kDa (Stahl et al., 1994). TrPAP26 is predicted to have up to 3 N-glycosylation sites and so it is possible that glycosylation can provide the extra 4 kDa. Note too that the masses determined from the western blots are not very accurate, so the detected proteins are not exactly 49 and 44 kDa in mass but they are close to those values.

The ~44 kDa protein could be an isoform or a splice variant of TrPAP26 encoded by the same gene. Li et al (2002) found splice variants for AtPAP10 and AtPAP13

(based on the presence or lack of introns in the cDNA), so this is a possibility for TrPAP26 as well. Alternatively, the ~44 kDa protein could be a tobacco PAP. If the anti-AtPAP12 antibody can cross-react with a white clover PAP, then it is possible that it also cross-reacts with an orthologous tobacco PAP(s).

For all the transgenic and control lines, the same amount of protein had been loaded into each well for SDS-PAGE. After incubation with the antibody, two bands at ~49 and ~44 kDa showed up in each lane, though recognition was much more prominent in the transgenic lines (except in lines 4 and 7) when compared with the control lines. However, because both bands were also detected, although faintly, in the control lines (except in control line C8), it is possible that either 1) TrPAP26 is the ~49 kDa protein with a ~44 kDa splice variant, 2) tobacco has an orthologous PAP that also has a splice variant, and both of these have similar masses to TrPAP26 and its variant, or 3) tobacco has a PAP12 that cross-reacts with the anti-AtPAP12 antibody. The third option is unlikely though because AtPAP12 has a mass of 60 kDa (Tran et al., 2010). However, if the project is continued, these two proteins should be sequenced to confirm their identity.

The same two bands were seen in both the soluble and cell wall fractions, but the bands were much more prominent in the soluble fraction. Therefore, TrPAP26::6xHis was purified from the soluble fraction even though the wall fraction had a higher total acid phosphatase activity. This may be due to a higher number of endogenous tobacco wall-associated acid phosphatases when compared with endogenous cytoplasmic acid phosphatases.

4.4.3 TrPAP26::6xHis affinity-purification

Affinity-purification was undertaken using Chelating Sepharose Fast Flow beads charged with Ni²⁺. A Coomassie stain of the protein that was eluted from the beads showed multiple bands, though only two (at ~49 and ~44 kDa) were expected (Fig. 3.20). Unfortunately, this purification system requires further work

to optimize purity because the concentration of imidazole determines which proteins are eluted. However, the time constraints in the project did not allow for a detailed investigation. In the future, the process could be optimized *via* a step-wise gradient of imidazole concentration in the elution buffer, addition of a protease inhibitor, and/or addition of detergent, glycerol, or 2-mercaptoethanol. All of these optimization steps are suggested by the product manufacturer as determinants that can improve specific protein elution.

Nevertheless, TrPAP26::6xHis was extracted and concentrated, though not completely pure, and the partially purified enzyme tested for its substrate specificity. It is noted therefore that some contaminating endogenous tobacco PAPs may also be contributing to the substrate analysis.

4.4.4 TrPAP26::6xHis activity

PAPs hydrolyze a variety of phosphate esters, which can be found in abundance in cells in the forms of ATP, PEP, sugar phosphates, and mononucleotides (Olczak et al., 2003). In common with AtPAP26, TrPAP26::6xHis displayed a broad substrate range with little to no phytase activity (i.e. it cannot hydrolyze phytic acid or phytate) (Table 3.4). This means that it may have a role in scavenging Pi because Pi exists in many different compounds in the environment and in the plant itself, so a broad substrate range allows for the enzyme to make use of more of the available Pi. The lack of phytase activity further confirms that it belongs to group Ia-2 PAPs because none of the other members of that group have phytase activity.

If this project is continued, other substrates should also be tested with TrPAP26. Mammalian PAPs can dephosphorylate phosphoproteins, but only two plant PAPs have been found to exhibit similar activity (Bozzo et al., 2004; Zhu et al., 2005). If TrPAP26 can use phosphoproteins as substrates then it would show that it has similarities with mammalian PAPs as well. Therefore, phosphotyrosine,

phosphothreonine, and phosphoserine should also be tested.

Robinson, Carson, et al (2012) found that AtPAP26 accumulated greatly in senescing leaves where ribosomal RNA accounts for 40-60% of the total organic P. Since only the signal peptide was dissimilar between AtPAP26 and TrPAP26, it is very likely that TrPAP26 can also utilize deoxyribonucleotides and ribonucleotides as substrates, and so these should be tested.

Some studies have found that in an alkaline environment (pH 8-9), some PAPs, including AtPAP26, display alkaline peroxidase activity as a possible response to pathogen attack (Bozzo et al., 2002; Bozzo et al., 2004; Kaija et al., 2002; Veljanovski et al., 2006). Again, because TrPAP26 and AtPAP26 are very similar in sequence, it is very likely that TrPAP26 will also display alkaline peroxidase activity, and so hydrogen peroxide should also be tested as a substrate.

TrPAP26::6xHis has not yet undergone rigorous biochemical analysis due to time constraints. In addition to further substrate testing, the pH optima need to be determined. The pH optima will give clues to where in the plant cell TrPAP26 resides because different cellular compartments have different pH, when/how it functions, and which substrates are preferred. If different substrates at different pH optima are preferred, then the pH could be one way by which the activity of the protein is regulated. If the enzyme does have alkaline peroxidase activity, then it should display this activity at a high pH (pH 8-9), but not at a lower pH.

4.5 Expression of *TrPAP26* and accumulation of TrPAP26 in response to changes in Pi supply *in planta*

In *Arabidopsis*, there are many PAPs performing different functions at various stages in the development of the plant and in response to various environmental stimuli (Tran and Plaxton, 2008). This has been found to hold for other plant

species as well (Bozzo et al., 2006). PAPs appear to have wider functions than plant acclimation during times of Pi starvation, and each one of these functions is different. For example, AtPAP2 in *Arabidopsis* regulates carbon metabolism (Sun et al., 2012), and NtPAP12 in tobacco aids in regenerating cell walls (Kaida et al., 2009). However, many PAPs specialize in Pi scavenging, and so most studies have looked at the differing expression and synthesis patterns of PAPs under Pi-sufficient (+Pi) and Pi-deficient (-Pi) conditions.

In these studies, the transcript, protein, and activity levels of PAPs in plants subjected to +/-Pi treatment is measured, and commonly the transcript levels reflect the protein and acid phosphatase activity levels (Bozzo et al., 2006; Hur et al., 2010; Kaida et al., 2003; Liang et al., 2012; Liang et al., 2010; Liao et al., 2003; Nishikoori et al., 2001; Tian et al., 2012; Tran et al., 2010; Wang et al., 2011; Zimmermann et al., 2004).

However, some studies have found that the transcript levels remained constant regardless of the Pi supply, but the protein (measured by probing with antibodies) and activity levels increased or decreased depending on the Pi supply (Liang et al., 2012; Veljanovski et al., 2006). AtPAP26 is one of the proteins in this category.

The protein AtPAP26 is a major responder to -Pi, so it's possible that TrPAP26 is also inducible under -Pi. To determine how TrPAP26 might function in white clover, the expression of *TrPAP26* and the accumulation and activity of TrPAP26 in white clover was studied in plants exposed to either Pi-sufficient or Pi-deficient media.

4.5.1 *TrPAP26* transcript levels

In the first experiment, white clover was grown in +Pi or -Pi for 24 hours, but there was no significant change in the transcript levels of *TrPAP26* in the roots of white clover grown in either treatment (Fig. 3.24). The second experiment was

conducted over 7 days, and while there was evidence of a significant decrease in transcript abundance over the time course, the trend was broadly similar in both treatments (Fig. 3.25). At day 7 there was a significant increase in *TrPAP26* expression in the +Pi treatment, but the time course would need to be extended to determine if this is a consistent trend. In a third experiment, where the plants were grown in +Pi or –Pi for 21 days, there was no significant difference in the transcript levels in the roots (Fig. 3.27). The results from these three experiments seem to suggest that *TrPAP26* transcription in the roots may be largely insensitive to Pi supply and is more constitutively expressed, and this agrees with the behavior of *AtPAP26* transcription observed in *Arabidopsis* roots in response to changes in Pi supply (Veljanovski et al., 2006).

While the first two experiments only looked at the expression level of *TrPAP26* in the roots, the third experiment also analyzed the expression level in other tissues: the first fully expanded leaf (F), a mature leaf (M), and the internodes between the first fully expanded leaf and the mature leaf (I) (Fig. 3.27). The transcription level in F was double what was seen in the roots, but there was no difference between the +Pi and –Pi treatments. However, the results in M and I did differ significantly. In M, the +Pi plants showed a 4.6-fold higher expression when compared with the –Pi plants. In I, the opposite was seen: a 4.5-fold higher expression was observed in the –Pi plants when compared with the +Pi plants. It should be noted that in these analyses, qPCR was used to examine *TrPAP26* expression, and so more confidence can be attributed to the fold-changes in expression. Thus these differences suggest that *TrPAP26* expression is being regulated differently in various tissues by Pi supply, which could suggest that the protein plays different roles according to the tissue in which it accumulates.

In contrast to roots, Robinson et al (2012) found that *AtPAP26* expression and *AtPAP26* protein levels were both up-regulated in senescing *Arabidopsis* leaves and played a pivotal role in P remobilization. They observed a 3-fold increase in transcript levels in fully senesced leaves. While the mature white clover leaves

analyzed in this project were not yellowing yet, the senescing process had most likely already begun, which could explain the 4.6-fold increase in *TrPAP26* expression in the plants grown in the +Pi treatment.

The fact that the *TrPAP26* expression level is 4.6-fold lower in mature leaves from plants grown in the –Pi media further supports the possibility that *TrPAP26* is involved in Pi remobilization. Since phosphorus is a highly mobile nutrient in plants, it follows that plants have evolved mechanisms of recycling the phosphorus stored in different tissues (Chapin, 1980). A few studies have found that plants, when exposed to P-deficiency, preferentially accumulate Pi in the leaves at the expense of the roots as an adaptive strategy to maximize photosynthesis rates (Bozzo et al., 2006; Mimura et al., 1996; Theodorou et al., 1992).

The 4.5-fold higher expression in internodes of plants exposed to Pi-deficiency supports this role in the remobilization of Pi. There is a well-developed vascular system in the stolons of white clover, so it would be advantageous to the plant to have *TrPAP26* scavenging free intercellular Pi where it could be then directly transported through the vascular system to sink tissues.

4.5.2 Acid phosphatase activity in white clover tissues exposed to the +Pi/–Pi treatments

From the third experiment where white clover plants were grown in either the +Pi or –Pi treatment for 21 days, root surface acid phosphatase activity assays were conducted. After 16 or 18 days, some plants were taken from the experiment. The elongation zone of the main root from the third root node of these plants was used intact in an acid phosphatase assay to detect activity at the root surface. There was a 1.4-fold increase in activity in the –Pi plants after 16 days, and a 2.1-fold increase after 18 days, which means that the Pi-starvation had induced the root surface acid phosphatase activity (Fig. 3.26). No tests for significance were carried out, but the differences are consistent with many other time courses carried out by

others using this experimental system. Further, these differences suggested that after 21 days of the treatment, there should have been clear increases in the APase activity in the $-Pi$ plants. However, the results of the assay on the 21-day treated plants were not as expected.

The 16-day and 18-day results were from a surface acid phosphatase activity assay using the whole root elongation zone, while the 21-day results used total soluble and cell wall protein extracted from ground tissues (Fig. 3.28). However, it was anticipated that there should have been higher acid phosphatase activity in the $-Pi$ tissues, especially in the root protein extracts, (even though in this assay the main root was used, including the elongation zone but excluding lateral roots). Instead, there was no detectable activity in the soluble fraction from roots of plants grown in $-Pi$ media, and no difference in activity between the $+Pi$ and $-Pi$ roots in the wall fraction. The acid phosphatase assay with the first fully expanded leaf (F) showed negligible difference in the soluble fraction and no activity could be detected in the $-Pi$ tissues in the cell wall fraction. The acid phosphatase assay with the mature leaf (M) showed no difference between the $+Pi$ and $-Pi$ tissues. The lack of an increase in acid phosphatase activity in the $-Pi$ tissues could be due to the enzymes being diluted in the extract, and so some concentration on extracts may be required so the assay used can more reliably measure any differences in activity.

Further, when whole roots were used, only surface proteins were exposed to the assay medium. The surface enzymes may be very sensitive to changes in Pi supply while changes in the soluble enzymes may not yet be induced over the same time frame because internal Pi levels may buffer any detection of a Pi deficit (Bozzo et al., 2006). Furthermore, pNPP acts as a good substrate for most phosphatases and doesn't discriminate between the induced and constitutively expressed phosphatases.

Since the total acid phosphatase activity assays showed conflicting results, the

TrPAP26 protein levels were then examined through antibody probing.

4.5.3 TrPAP26 protein levels

TrPAP26 accumulation was determined through antibody probing with the anti-AtPAP12 antibody. Tissues from white clover grown in the +Pi or –Pi treatments for 21 days were ground, and the soluble and cell wall proteins were extracted. The extracts from six biological replicates that had been assayed individually in the acid phosphatase activity assay (section 4.5.2) were pooled into one sample to eliminate individual differences. No proteins were detected by the antibody in the cell wall fraction, but a single protein at ~45 kDa was detected in the soluble fraction (Fig. 3.29).

For F, there was faint recognition of the ~45 kDa protein that is barely observable in both +Pi and –Pi plants. For M, there is a darker band in the +Pi extract when compared with the –Pi extract, but it is not a hugely prominent band. For I, recognition of the ~45 kDa protein in the +Pi plants seems to be just slightly more prominent than the band for the –Pi plants. For R, recognition was at about the same intensity in both +Pi and –Pi.

Comparing the intensity of antibody recognition with the corresponding mRNA levels shows that transcriptional control is not the only factor determining how much protein accumulates. The amount of protein accumulated does not necessarily reflect the differences in mRNA levels. There was a higher abundance of the *TrPAP26* transcripts in F than in R for both treatments, but there was more protein in R than in F, by 3.7-fold (+Pi) and 2.2-fold (-Pi). Though the –Pi internodes had a 4.5-fold higher level of *TrPAP26* transcripts, the +Pi internodes had a 1.8-fold higher level of protein accumulation. Only M showed a direct correlation between the mRNA levels and protein levels: both were determined to be >4-fold higher in the +Pi treatment.

It should be noted that in these protein accumulation studies, the AtPAP12

antibody is being used to detect TrPAP26 in the white clover background. From molecular weight comparisons, particularly after expressing *TrPAP26* in *E. coli*, some confidence in the antibody recognition in whole plants is valid. Ideally though, it should be confirmed that the antibody is recognising TrPAP26. However, more confidence in the transcriptional data for *TRPAP26* is possible as the full-length gene has been cloned from white clover and qPCR procedures can confirm that the abundance of only a single transcript is being measured. Finally, in the time available for this thesis, this experiment, while replicated, was only carried out once and so at least one repeat is necessary to be more certain of any conclusions.

4.5.4 Possible physiological role of TrPAP26

Given the data that was gathered in this thesis, based on the mRNA expression and protein accumulation patterns in white clover plants exposed to a change in Pi supply (+Pi/-Pi), only a tentative physiological role for TrPAP26 can be proposed. While the transcription in the roots seems to be constitutively expressed, TrPAP26 does not accumulate greatly under Pi starvation under the conditions used in this project. This differs from the observations made for *AtPAP26*. However, transcript levels in other tissues in the plant seem to be more Pi-supply-sensitive, or at least more temporally sensitive, as in the case of the mature leaves. This is similar to changes in the levels of *AtPAP26* transcripts. In white clover, protein levels overall seem to be more or less constant without much variation between the tissues despite Pi supply status. So while TrPAP26 has many similarities to *AtPAP26*, the physiological functions of the protein may differ.

Many plants, including white clover, have evolved to have symbiotic relationships with mycorrhizal fungi and arbuscular mycorrhizal fungi in particular (Lim and Cole, 1984). These fungi will colonize the apoplast of the roots to obtain sugars from the plant. In return, the plant benefits from the Pi and other nutrients that the long hyphae of the fungi are able to scavenge. The amount of reliance on these fungi

differs from species to species and is also dependent on the soil Pi concentration (Bolan, 1991). While *Arabidopsis* is a good model plant species, it doesn't form mycorrhizal associations. So even though white clover was grown free of mycorrhizae in this project, the complement of Pi scavenging proteins that white clover accumulates may still behave differently from that of plants without natural mycorrhizae (*Arabidopsis*).

More experiments are thus needed to determine the physiological role of TrPAP26 (see 4.7 future work).

4.6 Summary

A 1422 bp sequence was isolated from white clover (*Trifolium repens*) that had a 75% identity with *AtPAP26*. This gene was named *TrPAP26*, accordingly. It translates to a 473 amino acid sequence with a predicted mass of 55 kDa and contains motifs that identify it as a purple acid phosphatase. The protein has three putative N-glycosylation sites and one putative phosphorylation site.

Expression studies of the recombinant Δ sp-TrPAP26::GST were inconclusive because the protein was poorly soluble and inactive. However, these experiments did give some indication of the observed mass of TrPAP26, which is ~45 kDa. Accumulation studies of the recombinant TrPAP26::6xHis in tobacco suggests that the protein gains 3-4 kDa from the addition of glycosyl groups and may have a splice variant to generate a protein of ~44 kDa. These studies also showed that TrPAP26::6xHis has a broad substrate range and can use ATP, PEP, pyrophosphate, and pNPP very efficiently as substrates. However, these results should be viewed with caution, as the enzyme was not purified to homogeneity.

In planta studies showed that *TrPAP26* transcription was regulated in a manner similar to *AtPAP26* in that the gene was constitutively expressed in roots, but transcription was sensitive to Pi supply in other tissues. These studies also

showed that the protein accumulation and activity did not precisely match transcriptional abundance and may suggest that some post-translational modification of the protein to influence the stability and activity of the protein.

4.7 Future work

The biochemical characterization of TrPAP26 is incomplete. Purple acid phosphatases have some key characteristics that hold true for all the proteins within this classification: 1) the enzymes bind two metal ions, Fe(III) and M(II), where M is iron, zinc, or manganese (Kimura, 2000; Olczak et al., 2003), 2) they exhibit a purple (inactive enzyme) or pink (active enzyme) color when sufficiently concentrated in solution (Vincent and Averill, 1990), 3) they are insensitive to inhibition by tartrate (Vincent and Averill, 1990), and 4) they are N-glycosylated (Olczak et al., 2003).

Within its time frame, this project was unable to purify enough TrPAP26 to be able to carry out the analyses necessary to confirm that TrPAP26 is a PAP. In the future, TrPAP26 should be overexpressed in a eukaryotic background, purified, and concentrated so that: a metal content analysis can be done, its color in solution can be observed, its sensitivity to inhibition by tartrate can be assayed, and its post-translational modifications can be determined. In addition to these basic analyses, TrPAP26 should be tested to find its pH optima with various substrates and to find what other substrates can be catalyzed including: phosphotyrosine, phosphothreonine, phosphoserine, ribonucleotides, deoxyribonucleotides, and hydrogen peroxide (see section 4.4.4 for discussion as to why these substrates should be tested).

If TrPAP26 is able to use hydrogen peroxide as a substrate and therefore displays alkaline peroxidase activity, white clover plants should be exposed to pathogen attack and/or elicitor treatment to see if *TrPAP26* expression is induced and if more TrPAP26 accumulates at the site of attack/treatment. Other studies that have

shown alkaline peroxidase activity associated with a PAP hypothesized that this activity was induced in response to reactive oxygen species (ROS) generated by pathogen attack, but this has yet to be shown conclusively (Bozzo et al., 2002; Bozzo et al., 2004; Kaija et al., 2002; Veljanovski et al., 2006).

Two studies have found that PAPs become more active after proteolytic cleavage (Lång and Andersson, 2005; Mitić et al., 2005). It would be interesting to see if TrPAP26 behaves similarly after cleavage with trypsin. If its activity does increase with cleavage, then it would be a point of similarity with mammalian PAPs.

This project looked at the transcript levels of *TrPAP26* in young and mature leaves, but not in senescing leaves. Robinson et al (2012) found that *AtPAP26* expression was up-regulated in senescing leaves and played a major role in Pi remobilization, so *TrPAP26* might have a similar role since its transcripts were found to be up-regulated in the mature leaves of white clover.

For a full characterization of TrPAP26, immunohistochemical testing should be done either with GUS staining or with a fluorescent protein tag to see where and when the protein accumulates in the development of white clover and how its accumulation is affected by the level of Pi supplied at the cellular level. This would give an indication to its physiological role.

Finally, the EST database from AgResearch could be mined further for additional PAP candidates as the family has been shown in *A. thaliana* to be a large one. Some of these other PAPs may respond more readily to P supply.

5. Appendices

Appendix 5.1

Nucleotide sequences of a partial EST (897 bp) of an *AtPAP26*-like gene gifted by the New Zealand Pastoral Agricultural Research Institute (Palmerston North, NZ) (start codon highlighted) (Section 3.1.1):

```
TGTTTATGGATASAAAGATGCAGGGSTTGTGTTTTTTTGTTCCTTTGTGTTCTTCATC
TCTATCAGAGATGGATATGCAGGGATCACTAGTTCTTTTGTTAGGTCAGAGTATCCAT
CTRRTGATATCCCACCTTGATCATCAAGTATTTGCTGTTCCWAAGGGTCATAATGCACC
TCAACAAGTRCATATCACACAAGGTGATTATGAGGGAAAAGCAGTAATCATCTCATGG
GTGACCCCGAGATGAACCAGGATCCAGCCGTGTRCAATTTGGCACATCAGAGAATAAA
TTTGAAGCTAGTGCAGAAGGCACAGTTTCCAATTACACTTTTCGGCGAATACAAGTCY
GGTTACATTCATCATTGTCTTGTGTTGAAGGCCCTTGAGCACAATACTAAATACTACTACA
GAATTGGAAGCGGTGATTCTTCTCGAGAATTTTGGTTTGAACACCTCCTAAAGTTGA
TGCAGATGCCCTTACAAATTTGGGATCATTGGTGATATGGGACAAACATTTAATTCT
CTTTCCACTCTTGAGCACTATATAGAGAGTAAAGCACAGACTGTTCTATTTGTTGGAG
ATCTGTCTTATGCTGATAGGTATAAGTACACTGATGTTGGTGTAAGGTGGGATTCCTG
GGCCCGATTTGTTGAAAAGAGTACAGCATATCAGCCATGGATATGGTCTACAGGAAA
TCATGAAATAGAGTACTTTCCCTACATGGGAGAAGTTGTTCCCTTCAATAACTATCTTC
GACGCTATACTACTCCTTATTTGGCGTCCAATAGCAGCAGCCCTCTTTGGTATGCAAT
CAGGCGTGCATCTGCTCATATAATTGTACTIONTCCAGCTATTCACCCTACGTAAAGTAC
ACACCACAATATAAATGGCTCAGTGAAG
```

Appendix 5.2

Consensus nucleotide sequence from 5 positive colonies harbouring the pGEM vector with the 1 kb fragment of the putative *Tr-PAP26* sequence (Section 3.1.1):

```
CGACTCACTATAGGGCGAATTGGGCCCGACGTCGCATGCTCCCGGCCCGCCATGGC
GGCCGCGGGAATTCGATTGACTCGAGTCGACATCGATTTTTTTTTTTTTTTTTAAGAG
GTAACCTTGAGTACTTGACCACCATAATCCTTAATAATAATGAATATCTTGTCCTGAAAC
ATTTTGCTCCGACTACAAGCATATATGTTCAATATTTATTACTAACACACCGAAGGAAA
CATAGATGCTACAAGCCAAGTACATAGAACCCTTCGGACAACCATCATTCACTCAC
TCGGAGACTATTGAGCAGCATATATCGATCAGGAGGCACATTTTTATTCTTTTCACAA
CATCaAATACAAGATTGCTTCAAATTCCTATTAATGCTGACAACCTTCATCGATAACTG
ATAATATATAATGCTTCAGTTTTCTTCTCTCCTATTGTTTCCCAATACTGGTTATGCA
ATACGAATGAGTCAGTTGTCACTTTTTTGCCGTCATCATTGCGGTTCCAGTGGTAGAC
TGCATGGGTTCTATTTTTTATCTCCAATGTGGAGTGCCATAGCTTGCTTCACGAAAT
GCAGAGTATTCTGGCTGAGGATCCATAAACCTAGAAGCAAGACCTTCTTGATTTCCCTC
CATCTCCGACTGTTATGTACACAGGTGCCGATTTGTTCGGCTACGGGGTACCGGTTTTC
CGCCTGTTATGTTGTAATCTACATTGGAGAATCGATACGATCTTTCATAAGCATGGAC
ATGGCCAGCAAAGATCACGTCAACCTTGTACTTGATGAACCAGCTCTCGAAAACCACT
TCGCATGCTTTACCTTCATATAGTGAGCTTCATTAGTGTGTTGTAGAGCGGCACGTG
CATGAGAACAATGAGCCAAGGTGTCTTCTCCCTATCAACCCTCGTCAGCTCTTCACA
GAGCCATTTATATTGTGGTGTGTACTTTACGTAGGGGGAATAGCTGGGATA
```

Appendix 5.3

Consensus nucleotide sequence from 10 colonies harbouring the pGEM vector with the 0.4 kb fragment of the putative *Tr-PAP26* sequence (Section 3.1.1):

```
CGGCTCCCCCGGCCGCCATGGCGGCCGCGGGAATTCGATTGACTCGAGTCGACAT
CATTTTTTTTTTTTTTTTTTAAAATTAATGGTCAATATTCAATAGCACATGTTTTCAATGA
AGAATAAATAGCGCTACTTAGTTGCTGTAAATCTTGCATAGTTTCGTCAAATATGAGA
AAAAGATTTTACTCACCAACTAAAACCTTTAAAGTGAGCAAACAAGCTTTTTTACACCT
ATGGAACACTAAACCTTATTTATGTAAAATAATTTAATTGTCAATCAACATGTAAGAA
AGGTGAATAATTGTATAGAGAGGAGATGAGTTACCGTATGGCGAATAGCTGGATAGT
ACAATTATATGAGCAGATGCACGCCTGATTGCATACCAAAGAGGGCTGCTGCTATTG
GACGCCRAATAAGGAGTAGTAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGAC
CATATG
```

Appendix 5.4.1: Full list of BLAST hits of the 0.4 kb fragment of the putative *Tr-PAP26* sequence (Section 3.1.1):

Description	Max score	Total score	Query cover	E value	Max ident	Accession
Medicago truncatula chromosome 8 clone mth2-13911, complete sequence	289	289	56%	1e-74	84%	AC157349.7
Medicago truncatula chromosome 8 clone mth2-16a6, complete sequence	289	289	56%	1e-74	84%	AC144657.6
Medicago truncatula clone JCVI-FLMt-20A11 unknown mRNA	145	145	21%	3e-31	92%	BT147925.1
Medicago truncatula Purple acid phosphatase-like protein (MTR_4g115670) mRNA, complete cds	145	145	21%	3e-31	92%	XM_003609387.1
Phaseolus vulgaris gene for acid phosphatase, complete cds	129	129	20%	2e-26	89%	AB116720.1
Phaseolus vulgaris mRNA for acid phosphatase, complete cds	129	129	20%	2e-26	89%	AB116719.1
Lupinus luteus mRNA for acid phosphatase (acPase1 gene)	123	123	20%	1e-24	88%	AJ458943.2
Glycine max cDNA, clone: GMFL01-02-N14	122	122	21%	4e-24	86%	AK244377.1
Glycine max cDNA, clone: GMFL01-36-E23	122	122	21%	4e-24	86%	AK286765.1
Glycine max purple acid phosphatase-like protein (PAP3), mRNA >gb AY151272.1 Glycine max cultivar Union purple acid phosphatase-like protein (Pap3) mRNA, complete cds >gb GQ422778.1 Glycine max clone GmPhy12 phytase mRNA, complete cds	122	122	21%	4e-24	86%	NM_001249748.1
Glycine max bifunctional purple acid phosphatase 26-like (LOC100818438), mRNA >gb GQ422779.1 Glycine max clone GmPhy13 phytase mRNA, complete cds	120	120	20%	1e-23	87%	NM_001252721.1
Glycine max cDNA, clone: GMFL01-08-K22	120	120	20%	1e-23	87%	AK285394.1
Glycine max cultivar Jixian11 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	120	120	20%	1e-23	87%	AY151274.1
Ricinus communis Iron(III)-zinc(II) purple acid phosphatase precursor, putative, mRNA	118	118	20%	4e-23	87%	XM_002530719.1
Populus trichocarpa predicted protein, mRNA	118	118	20%	4e-23	87%	XM_002324582.1
Vitis vinifera, whole genome shotgun sequence, contig VV78X201381.7, clone ENTAV 115	118	118	23%	4e-23	83%	AM466178.1
Citrus maxima acid phosphatase mRNA, partial cds	116	116	20%	2e-22	86%	JN580586.1
Citrus sinensis acid phosphatase mRNA, partial cds	116	116	20%	2e-22	86%	JN580569.1
Populus trichocarpa predicted protein, mRNA	116	116	19%	2e-22	88%	XM_002330986.1
Vitis vinifera contig VV78X026477.7, whole genome shotgun sequence	116	116	21%	2e-22	85%	AM431847.2
Glycine max cultivar Wenfeng7 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	116	116	20%	2e-22	86%	AY151271.1
Glycine max cultivar Mengjin1 purple acid phosphatase-like protein (Pap3) mRNA,	116	116	20%	2e-22	86%	AY151275.1

complete cds Glycine max cultivar Zaoshu6 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	116	116	20%	2e-22	86%	AY151273.1
PREDICTED: Vitis vinifera bifunctional purple acid phosphatase 26-like (LOC100252214), mRNA	113	113	20%	2e-21	86%	XM_002264644.2
Vitis vinifera clone SS0ACG73YK21	113	113	20%	2e-21	86%	FQ380291.1
Jatropha curcas CL2382.Contig1_JC-CK_1A transcribed RNA sequence	109	109	19%	2e-20	86%	GAHK01006621.1
Jatropha curcas CL2382.Contig2_JC-CK_1A transcribed RNA sequence	109	109	19%	2e-20	86%	GAHK01006622.1
PREDICTED: Vitis vinifera bifunctional purple acid phosphatase 26-like (LOC100252468), mRNA	104	104	20%	1e-18	83%	XM_002285460.2
Arabidopsis thaliana chromosome 5, complete sequence	104	104	22%	1e-18	82%	CP002688.1
Genomic Sequence For Arabidopsis thaliana, Clone T5E15, Chromosome V, complete sequence	104	104	22%	1e-18	82%	AC019013.2
Arabidopsis lyrata subsp. lyrata ATPAP26/PAP26, mRNA	102	102	20%	3e-18	83%	XM_002870362.1
Arabidopsis thaliana purple acid phosphatase 26 (PAP26) mRNA, complete cds	102	102	20%	3e-18	83%	NM_122874.3
Arabidopsis thaliana putative purple acid phosphatase (PAP26) mRNA, complete cds	102	102	20%	3e-18	83%	AY842026.1
Arabidopsis thaliana putative acid phosphatase (At5g34850) mRNA, complete cds	102	102	20%	3e-18	83%	AY091415.1
Arabidopsis thaliana putative acid phosphatase (At5g34850) mRNA, complete cds	102	102	20%	3e-18	83%	AY050812.1
Ricinus communis Purple acid phosphatase precursor, putative, mRNA	100	100	20%	1e-17	83%	XM_002530721.1
Oryza sativa Japonica Group genomic DNA, chromosome 6, PAC clone:P0017B12	98.7	98.7	21%	4e-17	81%	AP003568.3
Oryza sativa Japonica Group genomic DNA, chromosome 6, PAC clone:P0416A11	98.7	98.7	21%	4e-17	81%	AP003523.3
PREDICTED: Fragaria vesca subsp. vesca bifunctional purple acid phosphatase 26-like (LOC101306162), mRNA	96.9	96.9	20%	1e-16	82%	XM_004287219.1
Oryza sativa Japonica Group cDNA clone:001-036-A09, full insert sequence	96.9	96.9	20%	1e-16	81%	AK061635.1
Gossypium hirsutum PAP26 mRNA, complete cds	91.5	91.5	20%	6e-15	81%	JN835293.1
Triticum aestivum cDNA, clone: WT012_E06, cultivar: Chinese Spring	91.5	91.5	19%	6e-15	82%	AK335179.1
Triticum aestivum mitochondrial acid phosphatase (ACP) mRNA, complete cds; nuclear gene for mitochondrial product	91.5	91.5	19%	6e-15	82%	FJ974002.1
Oryza sativa Japonica Group Os06g0643900 (Os06g0643900) mRNA, complete cds	91.5	91.5	20%	6e-15	80%	NM_001064717.1
Oryza sativa ossap1 mRNA for putative secretory acid phosphatase precursor, complete cds	91.5	91.5	20%	6e-15	80%	AB066560.1
Oryza sativa acid phosphatase mRNA, complete cds	91.5	91.5	20%	6e-15	80%	AF356352.1
PREDICTED: Solanum lycopersicum						

bifunctional purple acid phosphatase 26-like, transcript variant 3 (LOC101261339), mRNA	86.0	86.0	19%	3e-13	80%	XM_004242695.1
PREDICTED: Solanum lycopersicum bifunctional purple acid phosphatase 26-like, transcript variant 2 (LOC101261339), mRNA	86.0	86.0	19%	3e-13	80%	XM_004242694.1
PREDICTED: Solanum lycopersicum bifunctional purple acid phosphatase 26-like, transcript variant 1 (LOC101261339), mRNA	86.0	86.0	19%	3e-13	80%	XM_004242693.1
Ricinus communis Iron(III)-zinc(II) purple acid phosphatase precursor, putative, mRNA	86.0	86.0	20%	3e-13	79%	XM_002515439.1
Solanum lycopersicum cDNA, clone: LEFL1095BG01, HTC in leaf	86.0	86.0	19%	3e-13	80%	AK325339.1
Solanum lycopersicum cDNA, clone: LEFL1029DF04, HTC in leaf	86.0	86.0	19%	3e-13	80%	AK321824.1
Solanum lycopersicum cDNA, clone: LEFL1023AF10, HTC in leaf	86.0	86.0	19%	3e-13	80%	AK321315.1
Lycopersicon esculentum clone 133554R, mRNA sequence	86.0	86.0	19%	3e-13	80%	BT014303.1
Pennisetum glaucum AFLP fragment RQVTKKA2-10	84.2	143	19%	9e-13	97%	FN552488.1
Camelina sativa comp58787_c0_seq1 transcribed RNA sequence	82.4	82.4	20%	3e-12	78%	GABL01063206.1
Camelina sativa comp60101_c0_seq1 transcribed RNA sequence	82.4	82.4	20%	3e-12	78%	GABL01068428.1
PREDICTED: Cucumis sativus bifunctional purple acid phosphatase 26-like (LOC101223803), mRNA	82.4	82.4	20%	3e-12	79%	XM_004165597.1
PREDICTED: Cucumis sativus bifunctional purple acid phosphatase 26-like (LOC101207864), mRNA	82.4	82.4	20%	3e-12	79%	XM_004150479.1
PREDICTED: Brachypodium distachyon bifunctional purple acid phosphatase 26-like (LOC100831930), mRNA	82.4	82.4	20%	3e-12	78%	XM_003563256.1
Uncultured bacterium partial 16S rRNA gene, isolate lagoon october 2007, clone Contigseq_G06_plate1_laguna_07_gamma	82.4	141	17%	3e-12	100%	FN435233.1
Uncultured Firmicutes bacterium partial 16S rRNA gene, clone NL2BD-01-E06	82.4	143	16%	3e-12	100%	FM252597.1
Uncultured Firmicutes bacterium partial 16S rRNA gene, clone NL2BD-01-G10	82.4	143	16%	3e-12	100%	FM252596.1
Populus trichocarpa predicted protein, mRNA	82.4	82.4	20%	3e-12	79%	XM_002324840.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone:OTU-103	82.4	82.4	10%	3e-12	98%	AB239607.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone:OTU-26	82.4	82.4	10%	3e-12	98%	AB239606.1
Meloidogyne izabethae partial 18S rRNA gene	80.6	80.6	10%	1e-11	98%	HE667743.1
Uncultured Chloroflexi bacterium partial 16S rRNA gene, clone NL10BD-03-B08	80.6	134	16%	1e-11	98%	FM252667.1
Uncultured Chloroflexi bacterium partial 16S rRNA gene, clone NL10BD-02-A05	80.6	141	17%	1e-11	100%	FM252666.1
Uncultured beta proteobacterium partial 16S rRNA gene, clone NL10BD-05-G05	80.6	80.6	9%	1e-11	100%	FM252568.1
Uncultured beta proteobacterium partial 16S rRNA gene, clone NL10BD-04-C01	80.6	128	14%	1e-11	100%	FM252567.1

Uncultured beta proteobacterium partial 16S rRNA gene, clone NL5BD-01-F06	80.6	80.6	9%	1e-11	100%	FM252565.1
Uncultured alpha proteobacterium partial 16S rRNA gene, clone NL2BD-03-A05	80.6	141	16%	1e-11	100%	FM252530.1
Uncultured gamma proteobacterium partial 16S rRNA gene, clone NL1BD-02-E03	80.6	80.6	10%	1e-11	98%	FM252178.1
Pectobacterium chrysanthemi partial 16S rRNA, isolate SCH-01	80.6	80.6	9%	1e-11	100%	FM946179.1
Uncultured bacterium partial 16S rRNA gene, clone Crozet_d_189	80.6	80.6	10%	1e-11	98%	FM214900.1
Populus x canadensis mRNA for ferulate-5-hydroxylase (f5h gene)	80.6	80.6	9%	1e-11	100%	AM921699.1
Uncultured Firmicutes bacterium partial 16S rRNA gene, clone firmicutes_Raunefjorden 11	80.6	80.6	9%	1e-11	100%	AM706663.1
Uncultured Firmicutes bacterium partial 16S rRNA gene, clone firmicutes_Raunefjorden 04	80.6	80.6	9%	1e-11	100%	AM706659.1
Uncultured cyanobacterium partial 16S rRNA gene, clone cyanobacteria-or-chloroplast_Raunefjorden 02	80.6	80.6	10%	1e-11	98%	AM706638.1
Uncultured cyanobacterium partial 16S rRNA gene, clone cyanobacteria-or-chloroplast_North Atlantic gyre 35	80.6	180	17%	1e-11	100%	AM706905.1
Uncultured gamma proteobacterium partial 16S rRNA gene, clone betaproteobacteria_North Atlantic gyre 45	80.6	80.6	9%	1e-11	100%	AM706765.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone:OTU-24	80.6	80.6	9%	1e-11	100%	AB239609.2
Uncultured bacterium gene for 16S rRNA, partial sequence, clone:OTU-3	80.6	80.6	9%	1e-11	100%	AB239605.2
Physcomitrella patens transcriptional factor DBF1 mRNA, complete cds	80.6	80.6	9%	1e-11	100%	DQ202211.2
Solanum tuberosum purple acid phosphatase 3 (PAP3) mRNA, complete cds	80.6	80.6	19%	1e-11	79%	AY598342.1
Bradyrhizobium sp. ISLU256 partial nodC gene for N-acetylglucosaminyltransferase, strain ISLU256	80.6	80.6	9%	1e-11	100%	AJ560651.1
Bactrocera dorsalis clone F19 cytochrome P450 family 4 (CYP4) mRNA, complete cds	78.8	78.8	9%	4e-11	100%	GU292427.2
Uncultured bacterium partial 16S rRNA gene, clone Furnas-R-21	78.8	78.8	9%	4e-11	100%	HE795803.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone: CN41	78.8	78.8	9%	4e-11	100%	AB698048.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone: CN43	78.8	78.8	9%	4e-11	100%	AB698046.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone: CN42	78.8	78.8	9%	4e-11	100%	AB698045.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone: CN14	78.8	78.8	9%	4e-11	100%	AB698043.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone: CN15	78.8	78.8	9%	4e-11	100%	AB698044.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone: SK1	78.8	78.8	9%	4e-11	100%	AB698042.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone: SK211	78.8	78.8	9%	4e-11	100%	AB698041.1
Uncultured bacterium gene for 16S rRNA, partial sequence, clone: SK44	78.8	78.8	9%	4e-11	100%	AB698040.1

Appendix 5.4.2: BLAST hits of the 1 kb fragment of the putative *Tr-PAP26* sequence (Section 3.1.1):

Description	Max score	Total score	Query cover	E value	Max ident	Accession
Medicago truncatula clone JCVI-FLMt-20A11 unknown mRNA	1141	1141	90%	0.0	89%	BT147925.1
Medicago truncatula Purple acid phosphatase-like protein (MTR_4g115670) mRNA, complete cds	1124	1124	86%	0.0	89%	XM_003609387.1
Glycine max cDNA, clone: GMFL01-02-N14	814	814	75%	0.0	85%	AK244377.1
Glycine max cDNA, clone: GMFL01-36-E23	810	810	75%	0.0	85%	AK286765.1
Lupinus luteus mRNA for acid phosphatase (acPase1 gene)	776	776	69%	0.0	86%	AJ458943.2
Glycine max purple acid phosphatase-like protein (PAP3), mRNA >gb AY151272.1 Glycine max cultivar Union purple acid phosphatase-like protein (Pap3) mRNA, complete cds >gb GQ422778.1 Glycine max clone GmPhy12 phytase mRNA, complete cds	704	704	61%	0.0	86%	NM_001249748.1
Glycine max cultivar Wenfeng7 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	699	699	61%	0.0	86%	AY151271.1
Glycine max cultivar Mengjin1 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	699	699	61%	0.0	86%	AY151275.1
Glycine max cultivar Jixian11 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	699	699	61%	0.0	86%	AY151274.1
Glycine max cultivar Zaoshu6 purple acid phosphatase-like protein (Pap3) mRNA, complete cds	699	699	61%	0.0	86%	AY151273.1
Phaseolus vulgaris mRNA for acid phosphatase, complete cds	637	637	53%	4e-179	87%	AB116719.1
Glycine max cDNA, clone: GMFL01-08-K22	600	600	74%	1e-167	78%	AK285394.1
Glycine max bifunctional purple acid phosphatase 26-like (LOC100818438), mRNA >gb GQ422779.1 Glycine max clone GmPhy13 phytase mRNA, complete cds	538	538	53%	3e-149	83%	NM_001252721.1
Populus trichocarpa predicted protein, mRNA	488	488	56%	5e-134	80%	XM_002324582.1
Ricinus communis Iron(III)-zinc(II) purple acid phosphatase precursor, putative, mRNA	484	484	56%	6e-133	79%	XM_002530719.1
Populus trichocarpa predicted protein, mRNA	464	464	56%	6e-127	79%	XM_002330986.1
Jatropha curcas CL2382.Contig2_JC-CK_1A transcribed RNA sequence	452	452	56%	4e-123	78%	GAHK01006622.1
PREDICTED: Fragaria vesca subsp. vesca bifunctional purple acid phosphatase 26-like (LOC101306162), mRNA	452	452	56%	4e-123	78%	XM_004287219.1

Camelina sativa comp58787_c0_seq1 transcribed RNA sequence	439	439	57%	2e-119	77%	GABL01063206.1
Camelina sativa comp60101_c0_seq1 transcribed RNA sequence	439	439	57%	2e-119	77%	GABL01068428.1
Arabidopsis thaliana purple acid phosphatase 26 (PAP26) mRNA, complete cds	439	439	57%	2e-119	77%	NM_122874.3
Arabidopsis thaliana putative acid phosphatase (At5g34850) mRNA, complete cds	439	439	57%	2e-119	77%	AY091415.1
Arabidopsis thaliana putative acid phosphatase (At5g34850) mRNA, complete cds	439	439	57%	2e-119	77%	AY050812.1
Jatropha curcas CL2382.Contig1_JC-CK_1A transcribed RNA sequence	430	430	53%	1e-116	78%	GAHK01006621.1
Arabidopsis thaliana putative purple acid phosphatase (PAP26) mRNA, complete cds	430	430	57%	1e-116	77%	AY842026.1
PREDICTED: Vitis vinifera bifunctional purple acid phosphatase 26-like (LOC100252468), mRNA	428	428	53%	4e-116	78%	XM_002285460.2
Gossypium hirsutum PAP26 mRNA, complete cds	412	412	57%	3e-111	77%	JN835293.1
Arabidopsis lyrata subsp. lyrata ATPAP26/PAP26, mRNA	407	407	57%	1e-109	76%	XM_002870362.1
Ricinus communis Iron(III)-zinc(II) purple acid phosphatase precursor, putative, mRNA	401	401	54%	6e-108	77%	XM_002515439.1
PREDICTED: Vitis vinifera bifunctional purple acid phosphatase 26-like (LOC100252214), mRNA	398	398	56%	7e-107	76%	XM_002264644.2
Vitis vinifera clone SS0ACG73YK21	398	398	56%	7e-107	76%	FQ380291.1
Populus trichocarpa predicted protein, mRNA	398	398	56%	7e-107	76%	XM_002324840.1
Citrus maxima acid phosphatase mRNA, partial cds	389	389	47%	4e-104	79%	JN580586.1
Citrus sinensis acid phosphatase mRNA, partial cds	387	387	47%	1e-103	79%	JN580569.1
Solanum lycopersicum cDNA, clone: LEFL1095BG01, HTC in leaf	383	383	53%	2e-102	76%	AK325339.1
Solanum lycopersicum cDNA, clone: LEFL1029DF04, HTC in leaf	383	383	53%	2e-102	76%	AK321824.1
Solanum tuberosum purple acid phosphatase 3 (PAP3) mRNA, complete cds	379	379	53%	2e-101	76%	AY598342.1
Ricinus communis Purple acid phosphatase precursor, putative, mRNA	374	374	54%	8e-100	76%	XM_002530721.1
PREDICTED: Solanum lycopersicum bifunctional purple acid phosphatase 26-like, transcript variant 3 (LOC101261339), mRNA	370	370	53%	1e-98	76%	XM_004242695.1
PREDICTED: Solanum lycopersicum bifunctional purple acid phosphatase 26-like, transcript variant 2 (LOC101261339), mRNA	370	370	53%	1e-98	76%	XM_004242694.1
PREDICTED: Solanum lycopersicum bifunctional purple acid phosphatase 26-like, transcript variant 1 (LOC101261339), mRNA	370	370	53%	1e-98	76%	XM_004242693.1

Solanum lycopersicum cDNA, clone: LEFL1023AF10, HTC in leaf	370	370	53%	1e-98	76%	AK321315.1
Lycopersicon esculentum clone 133554R, mRNA sequence	365	365	53%	4e-97	76%	BT014303.1
Allium cepa ACPEPP mRNA for PEP phosphatase, complete cds	365	365	49%	4e-97	77%	AB052619.1
PREDICTED: Cucumis sativus bifunctional purple acid phosphatase 26-like (LOC101223803), mRNA	356	356	57%	2e-94	74%	XM_004165597.1
PREDICTED: Cucumis sativus bifunctional purple acid phosphatase 26-like (LOC101207864), mRNA	356	356	57%	2e-94	74%	XM_004150479.1
Medicago truncatula chromosome 8 clone mth2-13911, complete sequence	347	1123	85%	1e-91	93%	AC157349.7
Medicago truncatula chromosome 8 clone mth2-16a6, complete sequence	347	1123	85%	1e-91	93%	AC144657.6
PREDICTED: Brachypodium distachyon bifunctional purple acid phosphatase 26-like (LOC100831930), mRNA	336	336	54%	2e-88	74%	XM_003563256.1
PREDICTED: Solanum lycopersicum bifunctional purple acid phosphatase 26-like (LOC101262785), mRNA	334	334	54%	7e-88	74%	XM_004251686.1
Solanum lycopersicum cDNA, clone: LEFL1079CC03, HTC in leaf	334	334	54%	7e-88	74%	AK324572.1
Solanum lycopersicum cDNA, clone: LEFL1057BD08, HTC in leaf	334	334	54%	7e-88	74%	AK323467.1
Oryza sativa Japonica Group Os06g0643900 (Os06g0643900) mRNA, complete cds	327	327	54%	1e-85	74%	NM_001064717.1
Oryza sativa ossap1 mRNA for putative secretory acid phosphatase precursor, complete cds	327	327	54%	1e-85	74%	AB066560.1
Oryza sativa acid phosphatase mRNA, complete cds	327	327	54%	1e-85	74%	AF356352.1
Oryza sativa Japonica Group cDNA clone:001-036-A09, full insert sequence	327	327	54%	1e-85	74%	AK061635.1
Arabidopsis thaliana mRNA for L-ascorbate peroxidase, complete cds, clone: RAFL22-74-M19	324	324	43%	1e-84	77%	AK230096.1
Hordeum vulgare subsp. vulgare mRNA for predicted protein, complete cds, clone: NIASHv1130A07	315	315	53%	6e-82	73%	AK360979.1
Hordeum vulgare subsp. vulgare cDNA clone: FLbaf11i04, mRNA sequence	315	315	53%	6e-82	74%	AK248861.1
Triticum aestivum cDNA, clone: WT012_E06, cultivar: Chinese Spring	300	300	53%	1e-77	73%	AK335179.1
Triticum aestivum mitochondrial acid phosphatase (ACP) mRNA, complete cds; nuclear gene for mitochondrial product	297	297	53%	2e-76	73%	FJ974002.1
Phaseolus vulgaris gene for acid phosphatase, complete cds	284	651	53%	1e-72	90%	AB116720.1
Zea mays purple acid phosphatase (LOC100281588), mRNA >gb EU957194.1 Zea mays clone 1585432 purple acid phosphatase precursor, mRNA, complete cds	277	277	53%	2e-70	72%	NM_001154507.1
Zea mays full-length cDNA clone						

ZM_BFc0034G15 mRNA, complete cds	277	277	53%	2e-70	72%	BT039610.1
Picea glauca clone GQ03010_I09 mRNA sequence	206	206	42%	2e-49	71%	BT106786.1
Picea sitchensis clone WS0293_I04 unknown mRNA	206	206	42%	2e-49	71%	EF085421.1
Arabidopsis thaliana chromosome 5, complete sequence	183	418	51%	3e-42	80%	CP002688.1
Nicotiana tabacum PAP mRNA for purple acid phosphatase, complete cds, isolate:NtPAP21	183	183	43%	3e-42	70%	AB084124.1
Genomic Sequence For Arabidopsis thaliana, Clone T5E15, Chromosome V, complete sequence	183	418	51%	3e-42	80%	AC019013.2
Vitis vinifera contig VV78X026477.7, whole genome shotgun sequence	176	393	51%	4e-40	80%	AM431847.2
Ipomoea batatas mRNA for purple acid phosphatase	174	174	46%	1e-39	69%	AJ006224.1
Vitis vinifera, whole genome shotgun sequence, contig VV78X201381.7, clone ENTAV 115	172	396	45%	5e-39	86%	AM466178.1
Populus EST from severe drought-stressed opposite wood	165	165	15%	7e-37	83%	CU233121.1
Jatropha curcas Unigene37_JC-CK_1A transcribed RNA sequence	163	163	47%	3e-36	68%	GAHK01012586.1
Spirodela punctata mRNA for purple acid phosphatase, complete cds	161	161	45%	9e-36	69%	AB039746.1
Ipomoea batatas purple acid phosphatase precursor (PAP2) mRNA, complete cds	158	158	46%	1e-34	68%	AF200826.1
PREDICTED: Vitis vinifera purple acid phosphatase 2-like (LOC100247127), mRNA	156	156	43%	4e-34	69%	XM_002280837.2
Vitis vinifera clone SS0AFA12YI08	156	156	43%	4e-34	69%	FQ396020.1
Populus trichocarpa predicted protein, mRNA	156	156	46%	4e-34	68%	XM_002306090.1
PREDICTED: Vitis vinifera purple acid phosphatase 2-like, transcript variant 3 (LOC100246877), mRNA	150	150	46%	2e-32	68%	XM_002264014.2
PREDICTED: Vitis vinifera purple acid phosphatase 2-like, transcript variant 1 (LOC100246877), mRNA	150	150	46%	2e-32	68%	XM_002263901.2
PREDICTED: Vitis vinifera purple acid phosphatase 2-like, transcript variant 2 (LOC100246877), mRNA	150	150	46%	2e-32	68%	XM_002263935.1
PREDICTED: Solanum lycopersicum purple acid phosphatase 2-like (LOC101246629), misc_RNA	147	147	37%	2e-31	69%	XR_182703.1
PREDICTED: Fragaria vesca subsp. vesca purple acid phosphatase 2-like (LOC101292469), mRNA	145	145	34%	7e-31	70%	XM_004290234.1
Arabidopsis thaliana purple acid phosphatase 5 (PAP5) mRNA, complete cds	145	145	43%	7e-31	68%	NM_104172.2
PREDICTED: Fragaria vesca subsp. vesca purple acid phosphatase 2-like (LOC101294821), mRNA	143	143	45%	2e-30	68%	XM_004291971.1
Oryza sativa Japonica Group genomic DNA, chromosome 6, PAC clone:P0017B12	141	324	47%	8e-30	78%	AP003568.3

Oryza sativa Japonica Group genomic DNA, chromosome 6, PAC clone:P0416A11	141	324	47%	8e-30	78%	AP003523.3
PREDICTED: Fragaria vesca subsp. vesca purple acid phosphatase 2-like (LOC101293549), mRNA	140	140	32%	3e-29	70%	XM_004291353.1
PREDICTED: Vitis vinifera purple acid phosphatase 2-like (LOC100260651), mRNA	138	138	47%	1e-28	67%	XM_002274356.1
Saccharomyces cerevisiae x Saccharomyces kudriavzevii ALD6-SK gene for aldehyde dehydrogenase 6, strain Eg8/136, allele S. kudriavzevii	136	136	7%	4e-28	100%	HE585130.1
Durio zibethinus DNA, microsatellite locus DzMTa005	136	136	7%	4e-28	100%	AB292165.1
Uncultured Trebouxia photobiont 18S rRNA gene (partial), ITS1, 5.8S rRNA gene, ITS2 and 26S rRNA gene (partial), isolate L-121-1dd	136	136	7%	4e-28	100%	AJ969546.1
Serranus cabrilla microsatellite DNA, clone 31	136	136	7%	4e-28	100%	AM049419.1
Meloidogyne cruciani partial 18S rRNA gene	134	134	7%	1e-27	100%	HE667740.1
Prunus kansuensis pseudo PamSFB2 gene, putative pseudo S locus-linked F-box protein, strain: 0582. B	134	134	7%	1e-27	100%	AB725591.1
PREDICTED: Vitis vinifera purple acid phosphatase 2-like, transcript variant 3 (LOC100241752), mRNA	134	134	37%	1e-27	69%	XM_002264188.2
PREDICTED: Vitis vinifera purple acid phosphatase 2-like, transcript variant 1 (LOC100241752), mRNA	134	134	37%	1e-27	69%	XM_002264077.2
Uncultured bacterium partial 16S rRNA gene, isolate lagoon june 2008, clone FRA_38_K13_2008-07-07	134	134	7%	1e-27	100%	FN434773.1
Uncultured archaeon partial 16S rRNA gene, isolate AEXO15n.4	134	134	7%	1e-27	100%	FM242736.1

Appendix 5.5

Full contig nucleotide sequence of the partial EST sequence obtained from AgResearch (Appendix 5.1) and 1 kb fragment of the putative *Tr-PAP26* sequence Appendix 5.2) (Section 3.1.1):

```
ATGCAGGGGTTGTTGTTTTTTTGTTCCTTTGTGTTCTTCATCTCTATCAGAGATGGATA
TGCAGGGATCACTAGTTCTTTTGTAGGTGAGATATCCATCTGTTGATATCCCCTT
GATCATCAAGTATTTGCTGTTCCAAAGGGTCATAATGCACCTCAACAAGTACATATCA
CACAAGGTGATTATGAGGGAAAAGCAGTAATCATCTCATGGGTGACCCCAGATGAAC
CAGGATCCAGCCGTGTGCAATTTGGCACATCAGAGAATAAATTTGAAGCTAGTGCAG
AAGGCACAGTTTCCAATTACACTTTTCGGCGAATAACAAGTCTGGTTACATTCATTG
TCTTGTGAAGGCCTTGAGCACAATACTAAATACTACTACAGAATTGGAAGCGGTGAT
TCTTCTCGAGAATTTTGGTTTGAACACCTCCTAAAGTTGATGCAGATGCCCTTACA
AATTTGGGATCATTGGTGATATGGGACAAACATTTAATTCTTTTCCACTCTTGAGCA
CTATATAGAGAGTAAAGCACAGACTGTTCTATTTGTTGGAGATCTGTCTTATGCTGAT
AGGTATAAGTACACTGATGTTGGTGTAAGGTGGGATTCCTGGGCCCGATTTGTTGAA
AAGAGTACAGCATATCAGCCATGGATATGGTCTGCAGGAAATCATGAAATAGAGTAC
TTCCCTACATGGGAGAAGTTGTTTCTTTCAATAACTATCTTCGACGCTATACTACTC
CTTATTTGGCGTCCATTAGCAGCAGCCCTCTTTGGTATGCAATCAGGCGTGATCTG
CTCATATAATTGTACTATCCAGCTATTCACCCTACGTAAAGTACACACCACAATATAAA
TGGCTCAGTGAAGAGCTGACGCGGGTTGATAGGGAGAAGACACCTTGGCTCATTGT
TCTCATGCACGTGCCGCTCTACAACACTAATGAAGCTCACTATATGGAAGGTGAAAG
CATGCGAGTGGTTTTTCGAGAGCTGGTTCATCAAGTACAAGGTTGACGTGATCTTTGC
TGGTCATGTCCATGCTTATGAAAGATCGTATCGATTCTCCAATGTAGATTACAACATA
ACAGGCGGAAACCGGTACCCTGTAGCTGACAAATCAGCACCTGTGTACATAACAGTC
GGAGATGGAGGAAATCAAGAAGGTCTTGCTTCTAGGTTTATGGATCCTCAGCCAGAA
TACTCTGCATTTTCGTGAAGCAAGCTATGGACACTCCACATTGGAGATAAAAAATAGAA
CCCATGCAGTCTACCACTGGAACCGCAATGATGACGGCAAGAAAGTGACAACTGACT
CATTCGTATTGCATAATCAGTATTGGGGAAACAATAGGAGAAGAAGAAAAGTGAAGC
ATTAtATTATCAGTTATCGATGAAGTTGTCAGCATTTAATAGG
```

Appendix 5.6

Full-length nucleotide sequence of *TrPAP26* with putative start codon and stop codon bolded (section 3.1.1):

ATGCAGGGGTTGTTGTTTTTTGTTCTTTGTGTTCTTCATCTCTATCAGAGATGGATA
TGCAGGGATCACTAGTTCTTTTGTAGGTCAGAGTATCCATCTGTTGATATCCCCTT
GATCATCAAGTATTTGCAGTTCCAAAGGGTCATAATGCACCTCAACAAGTACATATCA
CACAAGGTGATTATGAGGGAAAAGCAGTAATCATCTCATGGGTGACCCAGATGAAC
CAGGATCCAGCCGTGTGCAATTTGGCACATCAGAGAATAAGTTTGAAGCTAGTGCAG
AAGGCACAGTTTCCAATTACACTTTTGGCGAATACAAGTCTGGTTACATTCATCATTG
TCACGTTGAAGGCCTTGAGCATAATACTAAATACTATTACAGAATTGGAAGCGGTGAT
TCTTCTCGAGAATTTTGGTTTGAACACCTCCTAAAGTTGATGCAGATGCCCTTGCA
AATTTGGGATCATTGGTGATATGGGACAAACATTTAATTCTCTTTCCACTCTTGAGCA
CTATATAGAGAGTAAAGCACAGACTGTTCTATTTGTTGGAGATCTGTCTTATGCTGAT
AGGTATAAGTACACTGATGTTGGTGTAAAGGTGGGATTCTGGGCCCGATTTGTTGAA
AAGAGTACAGCATATCAGCCATGGATATGGTCTGCAGGAATCATGAAATAGAGTACT
TTCCCTACATGGGAGAAGTTGTTCTTTCAATAACTATCTTCGACGCTATACTACTCC
TTATTTGGCGTCCAATAGCAGCAGCCCTCTTTGGTATGCAATCAGGCGTGCATCTGC
TCATATAATTGTAATCCAGCTATTCGCCATACGTAAAGTACACACCACAATATAAAT
GGCTCAGTGAAGAGCTGACGCGGGTTGATAGGGAGAAGACGCCTTGGCTCATTGTT
CTCATGCACGTGCCGCTCTACAACACTAATGAAGCTCACTATATGGAAGGTGAAAGC
ATGCGAGTGGTTTTCGAGAGCTGGTTCATCAAGTACAAGGTTGACGTGATCTTTGCT
GGTCATGTCCATGCTTATGAAAGATCGTATCGATTCTCCAATGTAGATTACAACATAA
CAGGCGGAAACCGGTACCCTGTAGCTGACAAATCAGCACCTGTGTACATAACAGTCCG
GAGATGGAGGAAATCAAGAAGGTCTTGCTTCTAGGTTTATGGATCCTCAGCCAGAAT
ACTCTGCATTTTGTGAAGCAAGCTATGGACACTCCACATTGGAGATAAAAAATAGAAC
CCATGCAGTCTACCACTGGAACCGCAATGATGACGGCAAGAAAGTGACAACCTGACTC
ATTCGTATTGCATAATCAGTATTGGGGAAACAATAGGAGAAGAAGAAAACCTGAAGCAT
TATATACTATCAGTTATCGATGAAGTTGTCAGCATT**TAA**

Appendix 5.7

BLAST matches after MALDI-TOF/TOF for the ~70 kDa putative recombinant Δ sp-TrPAP26::GST fusion protein (Fig. 3.8B, lane 11) (section 3.2.3):

1. [12109 Seq1](#) Mass: 54765 Score: 1165 Matches: 20
(20) Sequences: 16(16)
(Seq1_12109) TrPAP26
2. [Seq1_10175](#) Mass: 69057 Score: 453 Matches: 5(
5) Sequences: 5(5)
(10175_DnaK) E coli chaperone protein mutant
3. [GSTPep_ORF1](#) Mass: 28725 Score: 399 Matches: 6
(6) Sequences: 6(6)
(GSTPep_ORF1) 08004 GST_Peptide Orf 1

Appendix 5.8

BLAST match after MALDI-TOF/TOF for the ~57 kDa unknown recombinant protein (Fig. 3.8B, lanes 11 and 12) (section 3.2.3):

1. [gi|2624772](#) Mass: 57162 Score: 1837 Matches: 19
(19) Sequences: 13(13)
Chain A, Crystal Structure Of The Asymmetric
Chaperonin Complex GroelGROES(ADP)7

Appendix 5.9

BLAST match after MALDI-TOF/TOF for the ~45 kDa putative recombinant Δ sp-TrPAP26 protein (Fig. 3.8B, lane 12) (section 3.2.3):

1. [12109 Seq1](#) Mass: 54765 Score: 1325 Matches: 22
(22) Sequences: 20(20)
(Seq1_12109) TrPAP26

Appendix 5.10

Primer design and rationale (section 2.4.4 and Table 2.1):

TrPAP26 F1

A specific forward primer which hybridized to a sequence close to the 3' end of the partial EST sequence of *TrPAP26*.

TrPAP26 F2

Another specific forward primer that hybridized to a sequence close after "TrPAP26 F1." This primer was used in a second polymerase chain reaction (PCR) run to ensure that the amplified cDNA is actually the desired sequence.

ATG F

Using the contig made from the partial sequence and the sequence obtained from the PCR run using "TrPAP26 F2" as the forward primer and a 3' RACE adapter as the reverse primer, this forward primer was designed. It includes the putative start codon.

TAA R

Using the contig made from the partial sequence and the sequence obtained from the PCR run using "TrPAP26 F2" as the forward primer and a 3' RACE adapter as the reverse primer, this specific reverse primer was designed. It includes the putative stop codon

pGEX F

For *Escherichia coli* (*E. coli*) transformation, three primers were designed to clone *TrPAP26* into the pGEX-6P-3 vector (Smith et al, 1988). This first forward primer included some non-specific bases (GACGAC), an EcoRI digestion site (GAATTC), and the beginning of the coding region of *TrPAP26*.

pGEX R

This reverse primer designed for pGEX-6P-3 cloning included some non-specific bases (GACGAC), a Sall digestion site (GTCGAC), and the end of the coding region of *TrPAP26*.

pGEX Δsp F

This second forward primer designed for pGEX-6P-3 cloning excludes the sequence coding for the signal peptide of *TrPAP26*. It included some non-specific bases (GACGAC), an EcoRI digestion site (GAATTC), and the beginning of the coding region of *TrPAP26* after the signal peptide.

PAP26 int seq F

To check integration and/or expression of *TrPAP26*, this forward primer was designed using an internal sequence located 426 bp from the 5' end of the gene, ensuring that it was downstream of the signal peptide portion.

PAP26::6xHis F

For *Agrobacterium tumefaciens* transformation, two primers were designed to clone *TrPAP26* into the pART27 vector (Gleave, 1992). This forward primer included a short linker sequence (GACGAC), an XbaI digestion site (TCTAGA), and the 5' end of the coding region of *TrPAP26*.

PAP26::6xHis R

This reverse primer designed for pART27 cloning included a short linker sequence (GACGAC), an XbaI digestion site (TCTAGA), 6 histidine residues, and the end of the coding region of *TrPAP26*.

NtGAPDH qPCR F and R

For qPCR of mRNA levels of *TrPAP26::6xHis* in tobacco, tobacco GAPDH was used as the housekeeping gene. The forward primer was named "NtGAPDH qPCR F" and the reverse primer was named "NtGAPDH qPCR R."

Appendix 5.11

qPCR protocol (section 2.4.5.2):

qRT-PCR was performed using the LightCycler® 480 Real-Time PCR (Roche) and system series software 1.7, with three technical replicates of each cDNA sample (20-fold dilution). SYBR green I was used to monitor efficient DNA synthesis.

Reaction setup for qRT-PCR:

Forward primer (10 µM)	0.5 µL
Reverse primer (10 µM)	0.5 µL
2X LightCycler® 480 SYBR Green I Master Mix	5 µL
cDNA	2.5 µL
Sterile water	1.5 µL
<hr/>	
Total volume	10 µL

Master mixture and cDNA templates were dispensed into 96 well plates. The following program was used:

Steps	Temperature	Time	Cycle
Preincubation	95 ⁰ C	5 min	1
Amplification	Denaturation	95°C	10 sec
	Annealing	60°C	10 sec
	Extension	72°C	10 sec
Melting curve	95°C	5 min	1
Cooling	40°C		1

Relative transcript abundance was determined by comparative quantification to the geometric mean of the two reference genes. Fluorescence measurements were performed at 72⁰C for each cycle and continuously during final melting.

6. References

Almeida, J.P.F., Luscher, A., Frehner, M., Oberson, A., and Nosberger, J. (1999). Partitioning of P and the activity of root acid phosphatase in white clover (*Trifolium repens* L.) are modified by increased atmospheric CO₂ and P fertilisation. *Plant and Soil* 210, 159-166.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology* 10, 411-421.

Bari, R., Datt Pant, B., Stitt, M., and Scheible, W.R. (2006). PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiology* 141, 988-999.

Bolan, N. (1991). A critical review on the role of mycorrhizal fungi in the uptake of phosphorus by plants. *Plant and Soil* 134, 187-207.

Bozzo, G.G., Dunn, E.L., and Plaxton, W.C. (2006). Differential synthesis of phosphate-starvation inducible purple acid phosphatase isozymes in tomato (*Lycopersicon esculentum*) suspension cells and seedlings. *Plant, Cell & Environment* 29, 303-313.

Bozzo, G.G., Raghothama, K.G., and Plaxton, W.C. (2002). Purification and characterization of two secreted purple acid phosphatase isozymes from phosphate-starved tomato (*Lycopersicon esculentum*) cell cultures. *European Journal of Biochemistry* 269, 6278-6286.

Bozzo, G.G., Raghothama, K.G., and Plaxton, W.C. (2004a). Structural and kinetic properties of a novel purple acid phosphatase from phosphate-starved tomato (*Lycopersicon esculentum*) cell cultures. *Biochemical Journal* 377, 419-428.

Bozzo, G.G., Singh, V.K., and Plaxton, W.C. (2004b). Phosphate or phosphite addition promotes the proteolytic turnover of phosphate-starvation inducible tomato purple acid phosphatase isozymes. *FEBS Letters* 573, 51-54.

Braig, K., Otwinowski, Z., Hedgde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., and Sigler, P.B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371, 578-586.

Bresciani, R., and Von Figura, K. (1996). Dephosphorylation of the mannose-6-phosphate recognition marker is localized in late compartments of the endocytic route. Identification of purple acid phosphatase (uteroferrin) as the candidate phosphatase. *European Journal of Biochemistry* 238, 669-674.

Campos, M.d.A., Silva, M.S., Magalhães, C.P., Ribeiro, S.G., Sarto, R.P.D., Vieira, E.A., and Grossi de Sá, M.F. (2008). Expression in *Escherichia coli*, purification, refolding and antifungal activity of an osmotin from *Solanum nigrum*. *Microbial Cell Factories* 7, 7.

Caradus, J.R., and Snaydon, R.W. (1987). Aspects of the phosphorus nutrition of white clover populations II. Root exocellular acid phosphatase activity. *Journal of Plant Nutrition* 10, 287-301.

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

Chapin, F.S. (1980). The mineral nutrition of wild plants. *Annual Reviews of Ecology and Systematics* 11, 233-260.

de Marco, A. (2009). Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microbial Cell Factories* 8, 26.

Dinh, P.T.Y., Roldan, M.B., Leung, S., and McManus, M.T. (2012). Regulation of root growth by auxin and ethylene is influenced by phosphate supply in white clover (*Trifolium repens* L.). *Plant Growth Regulation* 66, 179-190.

Dionisio, G., Madsen, C.K., Holm, P.B., Welinder, K.G., Jorgensen, M., Stoger, E., Arcalis, E., and Brinch-Pedersen, H. (2011). Cloning and Characterization of Purple Acid Phosphatase Phytases from Wheat, Barley, Maize, and Rice. *Plant Physiology* 156, 1087-1100.

Duff, S.M., Moorhead, G.B., Lefebvre, D.D., and Plaxton, W.C. (1989). Phosphate Starvation Inducible 'Bypasses' of Adenylate and Phosphate Dependent Glycolytic Enzymes in *Brassica nigra* Suspension Cells. *Plant Physiology* 90, 1275-1278.

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

Dunlop, J., Lambert, M.G., Bosch, J., Caradus, J.R., Hart, A.L., Wewala, G.S., Mackay, A.D., and Hay, M.J.M. (1990). A programme to breed a cultivar of *Trifolium repens* L. for more efficient use of phosphate. In *Genetic Aspects of Plant Mineral Nutrition*, N. Bassam, M. Dambroth, and B.C. Loughman, eds. (Springer Netherlands), pp. 547-552.

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*). Metal content and spectroscopic characterization. *European Journal of Biochemistry* 260, 709-716.

Filippelli, G.M. (2011). Phosphate rock formation and marine phosphorus geochemistry: The deep time perspective. *Chemosphere* 84, 759-766.

Flanagan, J.U., Cassady, A.I., Schenk, G., Guddat, L.W., and Hume, D.A. (2006). Identification and molecular modeling of a novel, plant-like, human purple acid phosphatase. *Gene* 377, 12-20.

Hayman, A.R., and Cox, T.M. (1994). Purple acid phosphatase of the human macrophage and osteoclast. *Journal of Biological Chemistry* 269, 1294-1300.

Hunter, D.A., and McManus, M.T. (1999). Comparison of acid phosphatases in two genotypes of white clover with different responses to applied phosphate. *Journal of Plant Nutrition* 22, 679-692.

Hunter, D.A., Watson, L.M., and McManus, M.T. (1999). Cell wall proteins in white clover: influence of plant phosphate status. *Plant Physiology and Biochemistry* 37, 25-32.

Hur, Y.J., Jin, B.R., Nam, J., Chung, Y.S., Lee, J.H., Choi, H.K., Yun, D.J., Yi, G., Kim, Y.H., and Kim, D.H. (2010). Molecular characterization of OsPAP2: transgenic expression of a purple acid phosphatase up-regulated in phosphate-deprived rice suspension cells. *Biotechnology Letters* 32, 163-170.

Hurley, B.A., Tran, H.T., Marty, N.J., Park, J., Snedden, W.A., Mullen, R.T., and Plaxton, W.C. (2010). The dual-targeted purple acid phosphatase isozyme AtPAP26 is essential for efficient acclimation of Arabidopsis to nutritional phosphate deprivation. *Plant Physiology* 153, 1112-1122.

Jakobsen, I., Leggett, M.E. and Richardson, A.E. (2005). Rhizosphere microorganisms and plant phosphorus uptake. In "Phosphorus, agriculture and the environment". (Eds J.T. Sims, A.N. Sharpley). pp. 437-494. (American Society for Agronomy: Madison, WI).

Jasinski, S.M. (2009). Phosphate Rock. U.S.D.o.t. Interior, and U.S.G. Survey, eds., p. 10.

Johnson, J.F., Vance, C.P., and Allan, D.L. (1996). Phosphorus deficiency in *Lupinus albus*. Altered lateral root development and enhanced expression of phosphoenolpyruvate carboxylase. *Plant Physiology* 112, 31-41.

Joseph, R.E., and Andreotti, A.H. (2008). Bacterial expression and purification of interleukin-2 tyrosine kinase: single step separation of the chaperonin impurity. *Protein Expression and Purification* 60, 194-197.

Kaida, R., Hayashi, T., and Kaneko, T.S. (2008). Purple acid phosphatase in the walls of tobacco cells. *Phytochemistry* 69, 2546-2551.

Kaida, R., Sage-Ono, K., Kamada, H., Okuyama, H., Syono, K., and Kaneko, T.S. (2003). Isolation and characterization of four cell wall purple acid phosphatase genes from tobacco cells. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* 1625, 134-140.

- Kaida, R., Satoh, Y., Bulone, V., Yamada, Y., Kaku, T., Hayashi, T., and Kaneko, T.S. (2009). Activation of beta-glucan synthases by wall-bound purple acid phosphatase in tobacco cells. *Plant Physiology* 150, 1822-1830.
- Kaija, H., Alatalo, S.L., Halleen, J.M., Lindqvist, Y., Schneider, G., Kalervo Väänänen, H., and Vihko, P. (2002). Phosphatase and Oxygen Radical-Generating Activities of Mammalian Purple Acid Phosphatase Are Functionally Independent. *Biochemical and Biophysical Research Communications* 292, 128-132.
- Keresztessy, Z., Hughes, J., Kiss, L., and Hughes, M.A. (1996). Co-purification from *Escherichia coli* of a plant B-glucosidase-glutathion S-transferase fusion protein and the bacterial chaperonin GroEL. *Biochemical Journal* 314, 41-47.
- Kimura, E. (2000). Dimetallic hydrolases and their models. *Current Opinion in Chemical Biology* 4, 207-213.
- Kuang, R., Chan, K.H., Yeung, E., and Lim, B.L. (2009). Molecular and Biochemical Characterization of AtPAP15, a Purple Acid Phosphatase with Phytase Activity, in *Arabidopsis*. *Plant Physiology* 151, 199-209.
- Lång, P., and Andersson, G. (2005). Differential expression of monomeric and proteolytically processed forms of tartrate-resistant acid phosphatase in rat tissues. *CMLS Cellular and Molecular Life Sciences* 62, 905-918.
- Leelapon, O., Sarath, G., and Staswick, P.E. (2004). A single amino acid substitution in soybean VSP alpha increases its acid phosphatase activity nearly 20-fold. *Planta* 219, 1071-1079.
- Lehmann, K., Hoffman, S., Neudecker, P., Suhr, M., Becker, W.M., and Rosch, P. (2003). High-yield expression in *Escherichia coli*, purification, and characterization of properly folded major peanut allergen Ara h 2. *Protein Expression and Purification* 31, 250-259.
- Li, C., Gui, S., Yang, T., Walk, T., Wang, X., and Liao, H. (2012). Identification of soybean purple acid phosphatase genes and their expression responses to phosphorus availability and symbiosis. *Annals of Botany* 109, 275-285.
- Li, D., Zhu, H., Liu, K., Liu, X., Leggewie, G., Udvardi, M., and Wang, D. (2002). Purple acid phosphatases of *Arabidopsis thaliana*. Comparative analysis and differential regulation by phosphate deprivation. *Journal of Biological Chemistry* 277, 27772-27781.
- Li, Z., Gao, Q., Liu, Y., He, C., Zhang, X., and Zhang, J. (2011). Overexpression of transcription factor ZmPTF1 improves low phosphate tolerance of maize by regulating carbon metabolism and root growth. *Planta* 233, 1129-1143.

- Liang, C., Sun, L., Yao, Z., Liao, H., and Tian, J. (2012). Comparative analysis of PvPAP gene family and their functions in response to phosphorus deficiency in common bean. *PLoS One* 7.
- Liang, C., Tian, J., Lam, H.M., Lim, B.L., Yan, X., and Liao, H. (2010). Biochemical and Molecular Characterization of PvPAP3, a Novel Purple Acid Phosphatase Isolated from Common Bean Enhancing Extracellular ATP Utilization. *Plant Physiology* 152, 854-865.
- Liao, H., Wong, F.-L., Phang, T.-H., Cheung, M.-Y., Li, W.-Y.F., Shao, G., Yan, X., and Lam, H.-M. (2003). GmPAP3, a novel purple acid phosphatase-like gene in soybean induced by NaCl stress but not phosphorus deficiency. *Gene* 318, 103-111.
- Lim, L.L., and Cole, A.L.J. (1984). Growth response of white clover to vesicular-arbuscular mycorrhizal infection with different levels of applied phosphorus. *New Zealand Journal of Agricultural Research* 27, 587-592.
- Lu, K., Chai, Y.-R., Zhang, K., Wang, R., Chen, L., Lei, B., Lu, J., Xu, X.-F., and Li, J.-N. (2008). Cloning and characterization of phosphorus starvation inducible *Brassica napus* PURPLE ACID PHOSPHATASE 12 gene family, and imprinting of a recently evolved MITE-minisatellite twin structure. *Theoretical and Applied Genetics* 117, 963-975.
- Lung, S.C., Leung, A., Kuang, R., Wang, Y., Leung, P., and Lim, B.L. (2008). Phytase activity in tobacco (*Nicotiana tabacum*) root exudates is exhibited by a purple acid phosphatase. *Phytochemistry* 69, 365-373.
- Ma, X.-F., Tudor, S., Butler, T., Ge, Y., Xi, Y., Bouton, J., Harrison, M., and Wang, Z.-Y. (2012). Transgenic expression of phytase and acid phosphatase genes in alfalfa (*Medicago sativa*) leads to improved phosphate uptake in natural soils. *Molecular Breeding* 30, 377-391.
- Ma, X.-F., Wright, E., Ge, Y., Bell, J., Xi, Y., Bouton, J.H., and Wang, Z.-Y. (2009). Improving phosphorus acquisition of white clover (*Trifolium repens* L.) by transgenic expression of plant-derived phytase and acid phosphatase genes. *Plant Science* 176, 479-488.
- MfE (2010). *Environment New Zealand: Chapter 9 Land*. M.f.t. Environment, ed. (New Zealand), pp. 1-6.
- Miller, S.S., Liu, J., Allan, D.L., Menzhuber, C.J., Fedorova, M., and Vance, C.P. (2001). Molecular control of acid phosphatase secretion into the rhizosphere of proteoid roots from phosphorus-stressed white lupin. *Plant Physiology* 127, 594-606.

Mimura, T., Sakano, K., and Shimmen, T. (1996). Studies on the distribution, re-translocation and homeostasis of inorganic phosphate in barley leaves. *Plant, Cell & Environment* 19, 311-320.

Misson, J. (2005). A genome-wide transcriptional analysis using *Arabidopsis thaliana* Affymetrix gene chips determined plant responses to phosphate deprivation. *Proceedings of the National Academy of Sciences* 102, 11934-11939.

Mitić, N., Valizadeh, M., Leung, E.W.W., de Jersey, J., Hamilton, S., Hume, D.A., Cassady, A.I., and Schenk, G. (2005). Human tartrate-resistant acid phosphatase becomes an effective ATPase upon proteolytic activation. *Archives of Biochemistry and Biophysics* 439, 154-164.

Mitsukawa, N., Okumura, S., Shirano, Y., Sato, S., Kato, T., Harashima, S., and Shibata, D. (1997). Overexpression of an *Arabidopsis thaliana* high-affinity phosphate transporter gene in tobacco cultured cells enhances cell growth under phosphate-limited conditions. *Proceedings of the National Academy of Sciences* 94, 7098-7102.

Morcuende, R., Bari, R., Gibon, Y., Zheng, W., Pant, B.D., Bläsing, O., Usadel, B., Czechowski, T., Udvardi, M.K., Stitt, M., *et al.* (2007). Genome-wide reprogramming of metabolism and regulatory networks of *Arabidopsis* in response to phosphorus. *Plant, Cell & Environment* 30, 85-112.

Muller, R., Nilsson, L., Krintel, C., and Hamborg Nielsen, T. (2004). Gene expression during recovery from phosphate starvation in roots and shoots of *Arabidopsis thaliana*. *Physiologia Plantarum* 122, 233-243.

Nishikoori, M., Washio, K., Hase, A., Morita, N., and Okuyama, H. (2001). Cloning and characterization of cDNA of the GPI-anchored purple acid phosphatase and its root tissue distribution in *Spirodela oligorrhiza*. *Physiologia Plantarum* 113, 241-248.

Nozach, H., Fruchart-Gaillard, C., Fenaille, F., Beau, F., Ramos, O.H.P., Douzi, B., Saez, N.J., Moutiez, M., Servent, D., Gondry, M., *et al.* (2013). High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in *E. coli*. *Microbial Cell Factories* 12.

Olczak, M., Morawiecka, B., and Watorek, W. (2003). Plant purple acid phosphatases - genes, structures and biological function. *Acta Biochimie Polanska* 50, 1245-1256.

Olczak, M., and Olczak, T. (2007). N-Glycosylation sites of plant purple acid phosphatases important for protein expression and secretion in insect cells. *Archives of Biochemistry and Biophysics* 461, 247-254.

- Olczak, M., and Watorek, W. (1998). Oligosaccharide and polypeptide homology of lupin (*Lupinus luteus* L.) acid phosphatase subunits. *Archives of Biochemistry and Biophysics* 360, 85-92.
- Parsons, A.J., Edwards, G.R. Newton, P.C.D., Chapman, D.F., Caradus, J.R., Rasmussen, S. and Rowarth, J.R. (2012) Past lessons and future prospects: plant breeding for yield and persistence in cool-temperate pastures. *Grass and Forage Science* 66,153-172.
- Pintus, F., Spano, D., Corongiu, S., Floris, G., and Medda, R. (2011). Purification, primary structure, and properties of *Euphorbia characias* latex purple acid phosphatase. *Biochemistry (Moscow)* 76, 694-701.
- Rae, A.L., Jarmey, J.M., Mudge, S.R., and Smith, F.W. (2004). Over-expression of a high-affinity phosphate transporter in transgenic barley plants does not enhance phosphate uptake rates. *Functional Plant Biology* 31, 141-148.
- Raghothama, K.G. (1999). Phosphate Acquisition. *Annual Review of Plant Physiology and Plant Molecular Biology* 50, 665-693.
- Richardson, A.E. (2009). Regulating the phosphorus nutrition of plants: molecular biology meeting agronomic needs. *Plant and Soil* 322, 17-24.
- Richardson, A.E., Hocking, P.J., Simpson, R.J., and George, T.S. (2009). Plant mechanisms to optimise access to soil phosphorus. *Crop & Pasture Science* 60, 124-143.
- Robinson, W.D., Carson, I., Ying, S., Ellis, K., and Plaxton, W.C. (2012a). Eliminating the purple acid phosphatase AtPAP26 in *Arabidopsis thaliana* delays leaf senescence and impairs phosphorus remobilization. *New Phytologist* 196, 1024-1029.
- Robinson, W.D., Park, J., Tran, H.T., Del Vecchio, H.A., Ying, S., Zins, J.L., Patel, K., McKnight, T.D., and Plaxton, W.C. (2012b). The secreted purple acid phosphatase isozymes AtPAP12 and AtPAP26 play a pivotal role in extracellular phosphate-scavenging by *Arabidopsis thaliana*. *Journal of Experimental Botany* 63, 6531-6542.
- Schenk, G., Guddat, L.W., Carrington, L.E., Hume, D., Hamilton, S., and de Jersey, J. (2000). Identification of mammalian-like purple acid phosphatase in a wide range of plants. *Gene* 250, 117-125.
- Schilling, C., Zuccollo, J., and Nixon, C. (2010). Dairy's role in sustaining New Zealand. (New Zealand, NZIER), pp. 1-39.

- Shin, H., Shin, H.S., Dewbre, G.R., and Harrison, M.J. (2004). Phosphate transport in *Arabidopsis*: Pht1;1 and Pht1;4 play a major role in phosphate acquisition from both low- and high-phosphate environments. *Plant Journal* 39, 629-642.
- Smith, F.W. (2002). The phosphate uptake mechanism. *Plant and Soil* 245, 105-114.
- Stahl, B., Klabunde, T., Witzel, H., Krebs, B., Steup, M., Karas, M., and Hillenkamp, F. (1994). The oligosaccharides of the Fe(III)-Zn(II) purple acid phosphatase of the red kidney bean. *European Journal of Biochemistry* 220, 321-330.
- Student (1908). The probable error of a mean. *Biometrika* 6, 1-25.
- Sun, F., Suen, P.K., Zhang, Y., Liang, C., Carrie, C., Whelan, J., Ward, J.L., Hawkins, N.D., Jiang, L., and Lim, B.L. (2012). A dual-targeted purple acid phosphatase in *Arabidopsis thaliana* moderates carbon metabolism and its overexpression leads to faster plant growth and higher seed yield. *New Phytologist* 194, 206-219.
- Theodorou, M.E., Cornel, F.A., Duff, S.M., and Plaxton, W.C. (1992). Phosphate starvation inducible synthesis of the α -subunit of pyrophosphate-dependent phosphofructokinase in black mustard cell-suspension cultures. *Journal of Biological Chemistry* 267, 21901-21905.
- Thomas, R.G. (1987) The structure of the mature plant. In: *White clover*. Baker, M.J. and Williams W.M (eds) C.A.B. International, Wallingford, UK. 534 pp.
- Tian, J., Wang, C., Zhang, Q., He, X., Whelan, J., and Shou, H. (2012). Overexpression of OsPAP10a, A Root-Associated Acid Phosphatase, Increased Extracellular Organic Phosphorus Utilization in Rice. *Journal of Integrative Plant Biology* 54, 631-639.
- Tran, H.T., Hurley, B.A., and Plaxton, W.C. (2010a). Feeding hungry plants: The role of purple acid phosphatases in phosphate nutrition. *Plant Science* 179, 14-27.
- Tran, H.T., and Plaxton, W.C. (2008). Proteomic analysis of alterations in the secretome of *Arabidopsis thaliana* suspension cells subjected to nutritional phosphate deficiency. *Proteomics* 8, 4317-4326.
- Tran, H.T., Qian, W., Hurley, B.A., She, Y.M., Wang, D., and Plaxton, W.C. (2010b). Biochemical and molecular characterization of AtPAP12 and AtPAP26: the predominant purple acid phosphatase isozymes secreted by phosphate-starved *Arabidopsis thaliana*. *Plant Cell & Environment* 33, 1789-1803.

USGS (2000). The future of plant earth: scientific challenges in the coming century. (US Geographical Survey).

Valizadeh, M. (2004). Phosphotyrosyl peptides and analogues as substrates and inhibitors of purple acid phosphatases. *Archives of Biochemistry and Biophysics* 424, 154-162.

Vance, C.P., Uhde-Stone, C., and Allan, D.L. (2003). Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. *New Phytologist* 157, 423-447.

Veljanovski, V., Vanderbeld, B., Knowles, V.L., Snedden, W.A., and Plaxton, W.C. (2006). Biochemical and molecular characterization of AtPAP26, a vacuolar purple acid phosphatase up-regulated in phosphate-deprived Arabidopsis suspension cells and seedlings. *Plant Physiology* 142, 1282-1293.

Veneklaas, E.J., Lambers, H., Bragg, J., Finnegan, P.M., Lovelock, C.E., Plaxton, W.C., Price, C.A., Scheible, W.R., Shane, M.W., White, P.J., *et al.* (2012). Opportunities for improving phosphorus-use efficiency in crop plants. *New Phytologist* 195, 306-320.

Vincent, J.B., and Averill, B.A. (1990). An enzyme with a double identity: purple acid phosphatase and tartrate-resistant acid phosphatase. *FASEB Journal* 4, 3009-3014.

Wang, L., Li, Z., Qian, W., Guo, W., Gao, X., Huang, L., Wang, H., Zhu, H., Wu, J.W., Wang, D., *et al.* (2011). The Arabidopsis Purple Acid Phosphatase AtPAP10 Is Predominantly Associated with the Root Surface and Plays an Important Role in Plant Tolerance to Phosphate Limitation. *Plant Physiology* 157, 1283-1299.

Wang, Y., Norgård, M., and Andersson, G. (2005). N-glycosylation influences the latency and catalytic properties of mammalian purple acid phosphatase. *Archives of Biochemistry and Biophysics* 435, 147-156.

Waratrujiwong, T., Krebs, B., Spener, F., and Visoottiviseth, P. (2006). Recombinant purple acid phosphatase isoform 3 from sweet potato is an enzyme with a diiron metal center. *FEBS Journal* 273, 1649-1659.

Wu, P. (2003). Phosphate Starvation Triggers Distinct Alterations of Genome Expression in Arabidopsis Roots and Leaves. *Plant Physiology* 132, 1260-1271.

Xiao, K., Harrison, M., and Wang, Z.Y. (2006). Cloning and characterization of a novel purple acid phosphatase gene (MtPAP1) from *Medicago truncatula* Barrel Medic. *Journal of Integrative Plant Biology* 48, 204-211.

Yuan, H., and Liu, D. (2008). Signaling Components Involved in Plant Responses to Phosphate Starvation. *Journal of Integrative Plant Biology* 50, 849-859.

Zhang, C., and McManus, M.T. (2000). Identification and characterisation of two distinct acid phosphatases in cell walls of roots of white clover. *Plant Physiology and Biochemistry* 38, 259-270.

Zhang, Q., Wang, C., Tian, J., Li, K., and Shou, H. (2011). Identification of rice purple acid phosphatases related to phosphate starvation signalling. *Plant Biology* 13, 7-15.

Zhang, W., Gruszewski, H.A., Chevone, B.I., and Nessler, C.L. (2008). An Arabidopsis Purple Acid Phosphatase with Phytase Activity Increases Foliar Ascorbate. *Plant Physiology* 146, 431-440.

Zhang, Y., Yu, L., Yung, K.-F., Leung, D.Y.C., Sun, F., and Lim, B.L. (2012). Over-expression of AtPAP2 in *Camelina sativa* leads to faster plant growth and higher seed yield. *Biotechnology for Biofuels* 5, 19.

Zhu, H., Qian, W., Lu, X., Li, D., Liu, X., Liu, K., and Wang, D. (2005). Expression Patterns of Purple Acid Phosphatase Genes in Arabidopsis Organs and Functional Analysis of AtPAP23 Predominantly Transcribed in Flower. *Plant Molecular Biology* 59, 581-594.

Zimmermann, P., Regierer, B., Kossmann, J., Frossard, E., Amrhein, N., and Bucher, M. (2004). Differential expression of three purple acid phosphatases from potato. *Plant Biology* 6, 519-528.