IDENTIFICATION OF THE POKEWEED ANTIVIRAL PROTEIN INTERACTOME BY CO-IMMUNOPRECIPITATION-MASS SPECTROMETRY

Jennifer Chivers

A thesis submitted to the Faculty of Graduate Studies
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Graduate program in Biology
York University
Toronto, Ontario

August 2022

ABSTRACT

Ribosome-inactivating proteins (RIPs) are produced primarily by plants and are named for their enzymatic ability to depurinate ribosomal RNA. RIPs have been shown to have antiviral, antifungal, and antibacterial activity *in vitro* and when expressed transgenically. They are therefore of interest for their potential in human health, as both pathogenic agents and therapeutics, as well as in agriculture, to confer disease resistance in transgenic crops. However, little is known about the biological function of RIPs in their native context. *Phytolacca americana,* the American pokeweed, produces a RIP called pokeweed antiviral protein (PAP). The objective of this work is to investigate the role of PAP by mapping out the PAP-protein interactome; this will elucidate PAP's function by implicating the processes in which it is involved. Co-immunoprecipitation coupled with mass spectrometry was used to identify PAP protein interactors in pokeweed. Results identified protein interactions with diverse cellular functions in both the extracellular matrix, where PAP is primarily localized, and the cytoplasm, where the ribosomal target resides. One interactor was identified as a probable extracellular cysteine protease (paCP1); since this protein class has known roles in plant defense, paCP1 was chosen for further validation of its interaction with PAP using reverse co-IP. Differential expression and *in silico* promoter analysis demonstrated PAP and paCP1 co-expression in response to jasmonic acid, supporting the role of this interaction in defense. This work represents the first protein interactome mapping for a RIP; identification of PAP interactors in plant cells contributes to understanding PAP function and will aid in characterizing the biological role of RIPs in general.

ACKNOWLEDGMENTS

Firstly, I would like to thank my supervisor Dr. Katalin Hudak for her mentorship. Kathi has a tremendous enthusiasm for research and her attention to detail, calming influence, and confidence in her students has helped foster my development as a scientist over the past few years.

I would also like to thank my advisor Dr. Andy White. His friendliness, approachability, and especially his good advice have been of great help during my entire master's degree.

Thank you as well to Dr. Patricia Lakin-Thomas and Dr. Christopher Perry for being on my examination committee and taking the time to read my thesis.

Thank you to Hudak lab members, past and present, for your camaraderie, collaboration, and for making the lab somewhere to look forward to coming to every day. Thank you and good luck to current lab members Fernand, Tanya, and Kyra. A special thank you as well to past members Kira Neller, Alex Klenov, and Camille Diaz for making me feel welcome, supported, and for showing me the ropes.

Without everyone mentioned above, I wouldn't be graduating the scientist I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF COMMON ABBREVIATIONS

aa	Amino acid
ABA	Abscisic acid
AtXCP	*Arabidopsis thaliana* xylem cysteine protease
CA	Carbonic anhydrase
Co-IP	Co-immunoprecipitation
CoIP-MS	Co-immunoprecipitation coupled with mass spectrometry
CRE	Cis-regulatory element
CP	Cysteine protease
dH_2O	Distilled water
DTT	Dithiothreitol
ERAD	ER-associated degradation pathway
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
FDR	False discovery rate
FBA	Fructose bisphosphate aldolase
GO	Gene ontology
JA	Jasmonic acid
kDa	Kilodalton
LC/MS-MS	Liquid chromatography tandem mass spectrometry
paCP1	*Phytolacca americana* cysteine protease 1
pAB	Polyclonal antibody
PAP	Pokeweed antiviral protein
PR	Pathogenesis-related
PEG	Polyethylene glycol
PCD	Programmed cell death
PCR	Polymerase chain reaction
PLCP	Papain-like cysteine protease
POI	Protein-of-interest
PPI	Protein-protein interaction
ROS	Reactive oxygen species
RIP	Ribsome-inactivating protein
rRNA	Ribosomal RNA
SA	Salicylic acid
SAR	Systemic acquired resistance
TF	Transcription factor
UPR	Unfolded protein response
XCP	Xylem cysteine protease

1. INTRODUCTION

1.1. Ribosome-inactivating proteins

1.1.1 General characteristics

Ribosome-inactivating proteins (RIPs) are a family of proteins named for their ability to irreversibly disable the translation activity of ribosomes. RIPs are N-glycosidases that cleave the N-glycosidic bond linking a specific adenine to the sugar-phosphate backbone of the large ribosomal subunit rRNA, releasing the adenine base, and thus catalytically inactivating the ribosome (Endo et al. 1987; Stirpe et al. 1988). This depurination of the rRNA occurs in the universally conserved sarcin/ricin loop (SRL) and has been shown to interfere with the GTP-dependent binding of elongation factor 2/G to the ribosome, thus halting protein translation at the elongation step (Brigotti et al. 1989). Subsequently, evidence has also demonstrated RIP ability to depurinate polynucleotide substrates other than rRNA, such as supercoiled DNA, mRNA, viral RNA, poly(A), and naked rRNA (Li et al. 1991; Barbieri et al. 1994; Barbieri et al. 1997). In particular, the ability of RIPs to inhibit viral replication by depurination has been a popular topic of study, as various RIPs have demonstrated antiviral activity *in vitro* and *in vivo* against both plant and animal viruses (Citores et al. 2021). RIPs are therefore of scientific interest for their potential in human health as both toxic and therapeutic agents, and in agriculture to confer disease resistance in transgenic crops.

1.1.2 Classification of RIPs

To date, over 300 RIPs have been reported in over 100 different plant species, with overrepresentation in plant families found in the order Caryophyllales (Schrot et al. 2015). RIPs are classified into three groups, depending on the structure of the subunits (Figure 1). Type I RIPs are basic proteins of approximately 30 kDa, comprising a single polypeptide chain which contains the catalytic subunit. Many type-I RIPs encode an N-terminal signal sequence which co-translationally targets the protein to the endomembrane system (De Zaeytijd et al. 2019). From here, many type-I RIPs are secreted to the extracellular space while others are localized to the vacuole by additional C-terminal sequences. Type II RIPs are 60-65 kDa in size and comprise

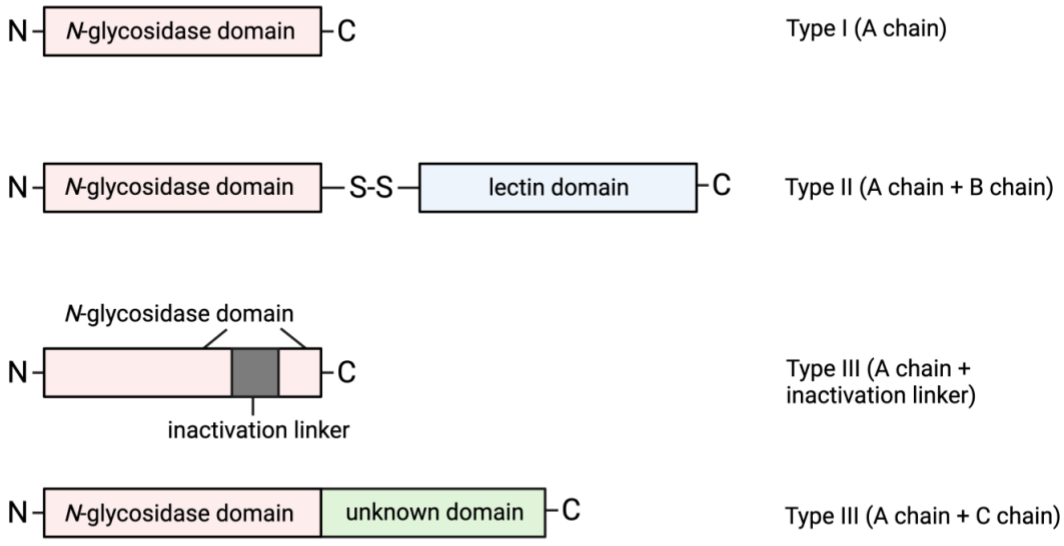


Figure 1. Schematic diagram of ribosome-inactivating protein (RIP) types according to mature protein domain organization. Type I RIPs consist solely of an A-chain possessing *N*-glycosidase activity, while type II RIPs are composed of a catalytic A-chain linked by disulfide bond to a B-chain with lectin-binding properties. Type III RIPs are irregularly structured RIPs, where the first subgroup consists of a catalytic A-chain fused to an additional domain with unknown function, and the second subgroup consists of an A-chain interrupted by an inactivation domain. For both type III subgroups, after the additional domains are cleaved, the processed RIP possesses similar function to type I RIPs. Figure created in BioRender.

two disulfide linked subunits: the A chain containing the catalytic subunit, and a B chain containing a lectin domain able to bind carbohydrates on the cell surface, allowing RIP uptake into the cytosol (Olsnes and Pihl 1973). Generally, type II RIPs display higher toxicity to cells than type I RIPs due to the ability to enter the cytoplasm; however, some type II RIPs show little to no toxicity (Stirpe et al. 1992; Barbieri et al. 2004). Finally, type III comprises RIPs with atypical organization and are further divided into two subgroups: one subgroup contains the enzymatic A chain fused to a C chain of unknown functionality, while the second subgroup contains an enzymatic A chain and a site for inactivation (Peumans et al. 2001).

1.1.3 Proposed roles of RIPs

The cytotoxicity and antiviral activity of RIPs has led to the hypothesis that RIPs function in plant defense against pathogens. RIPs have also been connected to this role through antifungal, antibacterial, and insecticidal properties, and expression of many RIPs is induced in response to various biotic and abiotic stresses (Song et al. 2000; Qin et al. 2005; Zhu et al. 2018). The mechanism by which RIPs exert these effects in their endogenous environments is not clear; however, several theories have been proposed. Type II RIPs are hypothesized to defend the plant against animal predators after ingestion of plant tissue or seeds produces a toxic effect; the lectin B chain provides a mechanism by which the RIP enters the animal cells (Polito et al. 2019). This theory however poses a problem for type I RIPs, which are sequestered to the vacuole or extracellular space and lack the lectin chain, leading to the question of how the depurination activity of RIPs could play a role in plant defense without ribosomal access. It has thus been hypothesized that damage to the cell membrane during a pathogen attack would also allow extracellular RIPs to re-enter the cytoplasm and access the host ribosomes (Ready et al. 1986; Kataoka et al. 1992). Inhibition of protein translation would result in the death of local cells, preventing pathogen spread. The above theories are supported by evidence that some type I RIPs can effectively act on plant ribosomes while type II RIPs are several-thousand-fold more active on animal ribosomes than on plant ribosomes (Prestle et al. 1992; Hartley et al. 1996).

More recently, a second mechanism has been proposed whereby RIPs are involved in signalling pathways that provide systemic protection against pathogens. When a RIP called α-momorcharin was applied to *Nicotiana benthamiana* leaves, expression of genes involved in modulating

reactive oxygen species (ROS) and pathogenesis-related defense proteins were enhanced, which correlated with resistance to tobacco mosaic virus (TMV), cucumber mosaic virus, turnip mosaic potyvirus, and chilli veinal mottle virus (Zhu et al. 2013; Zhu et al. 2020). A study by Yang et al. (2018) showed that when *N. tabacum* plants containing the TMV resistance gene N were sprayed with α-momorcharin and then infected with TMV, N gene expression increased as well as production of phytohormones salicylic acid (SA) and jasmonic acid (JA). SA and JA are phytohormones heavily implicated in plant defense: SA activates plant systemic acquired resistance (SAR), a broad spectrum and long-lasting defense response analogous to the innate immune system in animals, while JA mediates responses to pathogens and abiotic stress (Gao et al. 2015; Okada et al. 2015; Wang et al. 2020). Transcription and length of expression of JA and SA signalling pathway genes was increased, and JA and SA inhibitors were used to illustrate that α-momorcharin enhanced plant virus defense by manipulating SA-JA crosstalk (Yang et al. 2018). Taken together, these results indicate that RIPs may have a role in plant defense pathways, independent of their depurination activity.

It has also been proposed that RIPs could be polyfunctional proteins, with roles in biological processes in addition to defense. A high percentage of various RIPs in seeds and storage organs provides support for the hypothesis that RIPs could function as storage proteins. ME1 and ME2, type I RIPs isolated from *Mirabilis expansa* roots (a root crop), were found to constitute 20% of the storage-root proteins, but also exhibited antibacterial and antifungal properties (Vivanco et al. 1999). It has also been proposed that RIPs may play a role in cell senescence, as RIP levels have been reported to increase in senescent cells (Stirpe et al. 1996; Polito et al. 2013). Several RIPs have been described as having enzymatic activities in addition to depurination ability, suggesting bifunctionality: RIPs from *Trichosanthes kirilowii*, *Ricinus communis*, *Nicotiana tabacum*, and *Celosia plumosa* demonstrated chitinase (Remi Shih and McDonald 1997), phospholipase (Helmy et al. 1999), superoxide dismutase (Sharma et al. 2004), and antioxidant (Gholizadeh 2019) activities, respectively. However, as none of the aforementioned hypotheses have been thoroughly supported, the biological role of RIPs in plants remains unclear.

1.2 Pokeweed antiviral protein

1.2.1 Pokeweed antiviral protein 1

Pokeweed antiviral protein 1 (PAP-1) is a 29 kDa type-I RIP that is isolated from the American pokeweed plant (*Phytolacca americana*), belonging to the order Caryophyllales. Although considered a pest plant to agriculture, pokeweed has been of scientific interest since 1925 when Duggar and Armstrong reported that soluble extracts from pokeweed leaves reduced infectivity of TMV when applied to tobacco leaves (Duggar and Armstrong 1925). Since then, PAP-1 has been shown to depurinate animal viruses, including influenza (Tomlinson et al. 1974), poliovirus (Ussery et al. 1977), herpes simplex virus (Aron and Irvin 1980) and genomic HIV-1 RNA (Rajamohan et al. 1999). PAP-1 has also been investigated for antiviral activity against plant viruses that infect economically important crops; it has shown the ability to depurinate TMV (Chen et al. 1993), brome mosaic virus (Picard et al. 2005), and tobacco etch virus (Domashevskiy et al. 2012), and when expressed transgenically, PAP-1 confers resistance to potato virus X, potato virus Y, cucumber mosaic virus, and TMV (Moon et al. 1997; Wang et al. 1998). PAP-1 therefore remains an important research topic for agricultural as well as therapeutic purposes.

The PAP-1 gene encodes a 313 amino acid precursor, shown in Figure 2, which includes a 22 amino acid N-terminal signal peptide and a 29 amino acid C-terminal propeptide that are cleaved to yield the mature protein (Monzingo et al. 1993; Hur et al. 1995). The signal peptide directs co-translational insertion of PAP into the endoplasmic reticulum (ER) after which PAP is secreted primarily to the extracellular space (also known as the apoplast in plants), with other smaller populations potentially localized to the vacuole and cytoplasm (Ready et al. 1986). PAP has four amino acids (EAAR) in its active site which are highly conserved across RIPs, where glutamic acid (E176) and arginine 179 (R179) have been identified as critical for PAP depurination activity (Hur et al. 1995; Rajamohan et al. 2001). Two tyrosine residues (Y72 and Y123) also participate in binding the purine substrate (Monzingo et al. 1993).

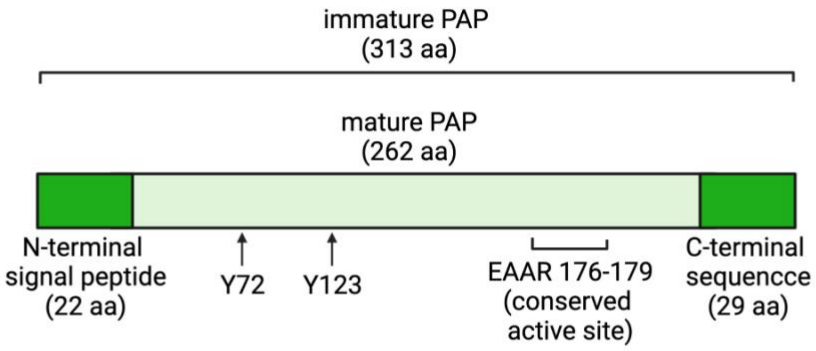


Figure 2. Schematic diagram of PAP-1 protein organization. The PAP-1 coding sequence encodes a 313 amino acid (aa) protein, with a 22 aa N-terminal signal peptide that is cleaved upon insertion into the endoplasmic reticulum. Processing of the mature PAP-1 protein also involves cleavage of 29 aa from the C-terminal end. The active site (EAAR), conserved across RIPs, spans residues 176-179, and residues Y72 and Y123 have been identified as critical for binding the purine base during depurination. Figure created in BioRender.

1.2.2 PAP isoforms

While PAP-1 remains the most studied pokeweed RIP, multiple PAP protein isoforms, encoded from separate genes, have been isolated: PAP-1 and PAP-2 have been purified from leaf tissue (Irvin 1975; Irvin et al. 1980), PAP-S1 and PAP-S2 from seeds (Honjo et al. 2002), and PAPα which is expressed in multiple tissues (Kataoka et al. 1992). PAP-1 shares 39%, 74%, 76% and homology with PAP-2, PAPα and PAP-S1, respectively. Active site residues are highly conserved among the PAP isoforms although they have been shown to differ in their ability to depurinate nucleic acid templates *in vitro* (Rajamohan et al. 1999; Honjo et al. 2002). Transcriptome analysis demonstrated that PAP isoforms are differentially expressed in pokeweed leaves, although all were upregulated with JA treatment, suggesting a potential common role in defense (Neller et al. 2016). However, their differences in expression and amino acid sequence also indicate potentially diversified roles for each.

1.2.3 Mechanisms of action

PAP-1 has the ability to depurinate ribosomes from yeast, plants, bacteria and animals (Rajamohan et al. 1999) which remains the most heavily studied feature of PAP. However, evidence exists to suggest PAP-1 mechanisms independent of enzymatic activity.

Firstly, PAP-1's cytotoxic effects can be decoupled from its depurination abilities. Site-directed mutagenesis produced seven mutations in the PAP-1 central domain that abolished the cytotoxic effects of PAP-1 when expressed in yeast, while maintaining depurination activity (Hur et al. 1995). PAP-1-expressing yeast cells also show nuclear fragmentation and production of ROS, which are features of apoptosis (Çakır et al. 2015). Yeast that co-expressed both PAP-1 and a plant anti-apoptotic protein, AtbI-1, demonstrated reduced cytotoxicity while maintaining PAP-1 ribosomal depurination and inhibition of translation. This indicates that PAP-1 may possess cytotoxic mechanisms independent of its enzymatic activity, including as a potential inducer of apoptosis, but these putative roles remain undefined.

Secondly, while the antiviral activity of PAP-1 *in vivo* is poorly understood, evidence suggests it is not fully dependent on ribosomal inactivation nor on viral RNA depurination. A study by Zhu

et al. (2016) showed that exogenous treatment of *Nicotiana benthamiana* leaves with PAP-1 was associated with increased antioxidant activity, notably superoxide dismutase, catalase and peroxidase. When PAP pre-treated leaves were inoculated with TMV, this effect was heightened and also associated with viral resistance and reduced ROS levels, suggesting PAP-1 may enhance viral resistance through antioxidant regulation of ROS (Zhu et al. 2016). A second study demonstrated that transgenic expression of a non-ribotoxic and non-cytotoxic PAP-1 mutant in tobacco resulted in overexpression of pathogenesis-related (PR) protein genes and genes associated with hypersensitive response, a type of plant-specific apoptosis associated with pathogen resistance (Zoubenko et al. 2000). PR proteins are induced in response to pathogen attack and also correlate with induction of SAR (Wang et al. 2005). These results indicate that PAP-1 may activate plant defense pathways independently of either ribosomal depurination or a cytotoxic effect. However, these lines of evidence are preliminary and the role of PAP-1 in plant defense, especially in its endogenous environment, remains unclear.

1.3 Protein interactions

1.3.1 Introduction

Proteins play the central role in fulfilling the biological processes in a cell, including gene expression, cell growth and proliferation, nutrient transport, intercellular communication, pathogen defense and apoptosis. Therefore, elucidation of the functional roles of proteins remains an important focus in research. However, proteins rarely act alone; instead, they function by interacting with other proteins, forming complex molecular networks that have effects at both cellular and organismal levels. Thus, a protein's function can only be understood in the context of its interaction network. A map of interactions between a protein of interest (POI) and other proteins in the cellular proteome is called the interactome. Mapping the protein-protein interactome can help elucidate a protein's function by implicating the processes in which the protein is involved (Singh et al. 2020).

Protein-protein interactions (PPIs) include not only direct, physical binding between proteins in a multi-subunit complex but also functional associations, such as between a transcriptional regulator and the pathway it controls. Correspondingly, PPIs can be classified according to the

stability of the complex as either stable or transient; the most stable interactions are those that exist only in their complexed form, while transient interactions associate and disassociate over time (Acuner Ozbabacan et al. 2011). Transient interactions can be further classified as either strong or weak, depending on the interaction affinity and the lifetime of the interaction. Weak transient interactions have a low affinity which results in continuous breaking and forming of the association and are characterized by a disassociation constant in the micromolar range. In contrast, strong transient interactions have a disassociation constant in the nanomolar range and include those interactions where association/disassociation is due to an external trigger such as conformational change, stabilization by an effector molecule, chemical modification, or co-localization (Nooren and Thornton 2003). However, the distinctions between these groups are not entirely discrete and all protein-protein interactions exist on a continuum, as illustrated in Figure 3. While the abundance of each type of association in cells is unknown, if it is assumed that protein partners in permanent complexes are strongly co-expressed, then permanent interactions comprise only a small fraction of interactions, with transient interactions comprising the vast majority (Jansen et al. 2002). The roles of transient complexes are essential in the regulation of biochemical pathways and signalling cascades which are indispensable to the cell.

1.3.2 Methods for studying PPIs

Many experimental approaches have been developed to investigate protein interactomes, and include *in vivo*, *in vitro*, and *in silico* techniques. Different experimental approaches result in different interactomes, due to the diverse physiochemical properties, abundance and localization of individual proteins, as well as the wide range of equilibrium disassociation constants for PPIs. Since one method will provide only a number of all true PPIs of the POI, use of *in silico* techniques based on literature curation, evolutionarily conserved interactions, and network integration of function can provide a more meaningful picture of the resulting interactome.

1.3.3 Co-immunoprecipitation

One strategy used to characterize protein interactomes is co-immunoprecipitation coupled with mass spectrometry (coIP-MS). Co-immunoprecipitation (co-IP) is a well-established method of purifying *in vivo* protein complexes. An antibody specific to the POI is immobilized to an

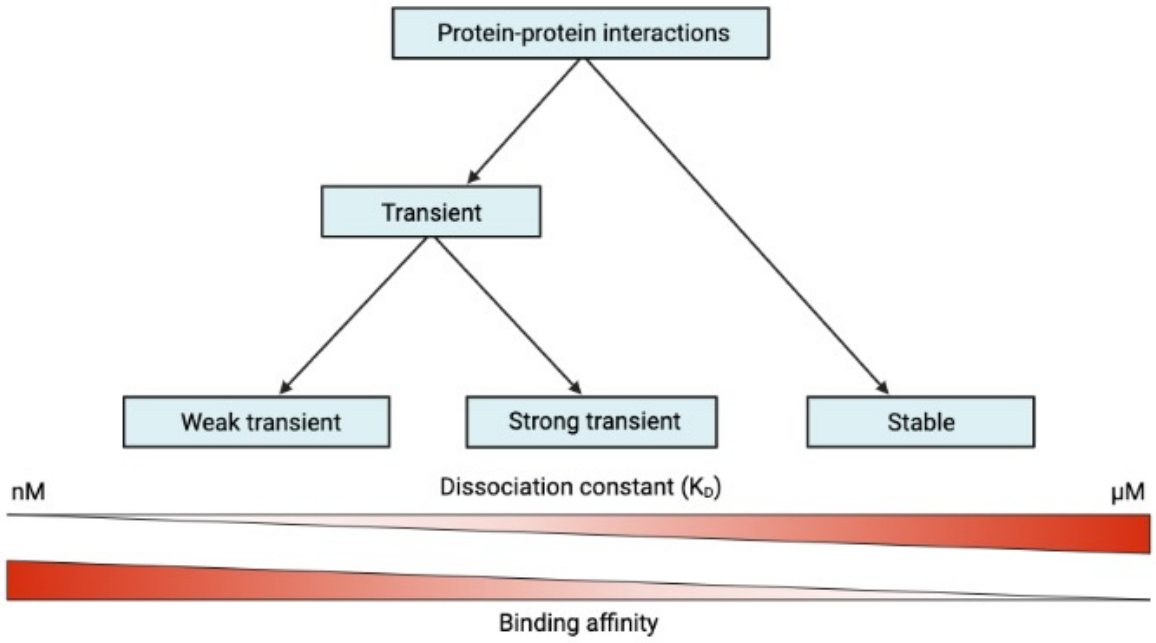


Figure 3. Types of protein-protein interactions (PPIs). PPIs can be considered either stable or transient, depending on the length of the interaction, and transient interactions can further be divided into weak and strong interactions. All PPIs can be placed along the binding affinity spectrum, which measures the attractive forces between two proteins in a complex. Binding affinity is quantified by the dissociation constant (K_D), an equilibrium constant that measures the propensity of the PPI to separate. Binding affinity is inversely proportional to $K_{D.}$ Figure created in BioRender.

affinity matrix and used to pull the POI from the cellular lysate along with any interactors it has bound (also known as the bait and prey proteins, respectively). After elution and purification, the unknown protein interactors can then be identified by mass spectrometry (Gingras et al. 2007; Zhang et al. 2019). The most commonly used mass spectrometry method involves digestion of complex protein mixtures into peptides, followed by peptide separation and identification using liquid chromatography coupled-tandem mass spectrometry (LC-MS/MS). Comparison of peptides to a protein database can identify proteins pulled down by co-IP and provide a map of the POI interactome (Aebersold and Mann 2003).

An advantage of coIP-MS is that it can identify interactions that occur in the native cellular environment, if it can be shown that proteins also co-localize and are temporally co-expressed. However, due to the possibility of technical false positives, validation of coIP-MS identified putative interactors must be performed using a second method to rule out non-specific interactors with the antibody-affinity matrix (Solstad et al. 2008).

1.3.4 Reverse co-IP

Reverse co-IP is considered a standard for validating protein interactions: immunoprecipitation of the putative interactor is followed by identification of the POI in the co-purified population, thus reversing the bait and prey proteins from the initial co-IP to demonstrate interaction. If no antibody exists for the putative interactor, then reverse co-IP can occur by expressing an affinity-tagged bait protein in *E. coli*, immobilizing the purified protein to beads followed by incubation of the immobilized bait protein with lysate containing endogenously expressed POI (Singh et al. 2020). Probing for the pulled-down POI by Western blot can then confirm that the protein interaction also occurs when bait and prey proteins are reversed, thus providing evidence that the initially identified interaction was specific.

1.4 Rationale and objective of thesis

The function of RIPs in their endogenous environment is unclear. Evidence points to a potential role for PAP to be involved in plant defense; however, the mechanism behind this potential role remains unknown. Since PAP has been reported to comprise 0.5% total soluble protein in

pokeweed leaf tissue (Bonness et al. 1994) and is primarily localized to the apoplastic space, this indicates a potential role for PAP independent of ribosomal access. Identification of a PAP-protein interactor may implicate PAP in a particular cellular process through guilt-by-association.

To date, no screening of a RIP-protein interactome has been performed; a significant roadblock in this work for any protein is a lack of genomic resources. The genome of the model plant organism *Arabidopsis thaliana* does not contain any putative RIP coding genes (Kaul et al. 2000), and even those plant species with the most heavily studied RIPs have either no genomic resources or only very preliminary resources. Pokeweed has only recently had a draft genome sequenced and published by our lab (Neller et al. 2019), which now allows for identification of any isolated protein interactors in a coIP-MS study.

The objective of my work was therefore to identify endogenous PAP-protein interactions. This was initially attempted using a protein proximity labelling method called BioID (see Appendix A). However, after this was unsuccessful, coIP-MS was utilized instead to identify PAP-protein interactions in pokeweed leaf tissue, with validation of a putative interaction by reverse co-IP. *In silico* work on PAP interactors was also performed to provide further support for the interaction through co-localization and co-expression. I hypothesized that PAP plays a role in plant defense while localized to the apoplastic space; therefore, I predicted coIP-MS results would show a putative protein interactor that co-localizes to the apoplast and has a potential role in plant defense. Previous work in the field of RIP interactions has focused primarily on RIP-nucleic acid interactions; this work will represent the first investigation of a RIP-protein interactome. Mapping the PAP-protein interactome will ultimately serve to further elucidate PAP's function and behaviour in the cell.

2. MATERIALS AND METHODS

2.1 Plant cultivation

2.1.1 Seeds

Pokeweed seeds were originally provided by the Tumer laboratory at Rutgers University, NJ (USA) and tobacco seeds (*Nicotiana tabacum* cv. Samsun) were originally obtained from the Canadian Tobacco Research Foundation, ON (Canada).

2.1.2 Sowing

To promote germination, pokeweed seeds were covered in sulphuric acid and rotated gently for 5 minutes, followed by rinsing in running water for 5 minutes. Seeds were then imbibed in fresh water for 4-5 days until over 90% of seed coats were visibly cracked. Cracked seeds were sown in soil mix (1-part garden soil: 1-part Pro-mix all-purpose soil: 0.5-part cattle manure: 0.5-part sand) and trays containing pots with sown seeds were placed on an electric heat pad to aid in germination speed. Trays were also covered with a plastic dome to maintain humidity until cotyledons had emerged fully.

Tobacco seeds were sown directly into the aforementioned soil mix and grown for 14 days under a plastic dome. Individual seedlings were then transplanted into small pots and grown for another 14 days under a plastic dome, after which the dome was removed.

2.1.3 Growth conditions

All pokeweed and tobacco plants were grown in growth chambers and given a 14-hour light/10-hour darkness light cycle; chamber lighting was comprised of 75% fluorescent and 25% incandescent bulbs (180 $\mu E/m^2/s$). Temperature was held at 24°C and 21°C during the light and darkness periods, respectively, and fan speed was set to 65%. Plants were watered approximately every 2 days and fertilized weekly with a 20-20-20 fertilizer (Plant Prod Classic).

Pokeweed and tobacco leaves were harvested at the 6-8 leaf stage, at the approximate age of 4 weeks for pokeweed and 6 weeks for tobacco. Leaf discs (1 cm diameter) were flash frozen in liquid nitrogen and stored at -80°C until use.

2.2 Co-immunoprecipitation of PAP for identification of protein interactors

2.2.1 Crosslinking antibody to beads

For each ~75 mg tissue sample, 40 µL of protein A magnetic beads (New England Biolabs #S1425S) was removed to a 1.7 mL microtube and washed three times in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4). Beads were then resuspended in 1 mL PBS and 2 µL of polyclonal PAP-specific antibody (Table 1) was added. Antibody and beads were incubated for 2 hours at 4°C with end-over-end rotation. Beads were then washed three times with 1 mL of 0.2 M sodium borate (pH 9.0). To crosslink bound PAP antibody to the beads, 20 mM dimethyl pimelimidate (DMP; Sigma-Aldrich #D8388) in 0.2 M sodium borate was added to the bead samples and rotated for 40 minutes at room temperature. Beads were washed once in 0.2 M ethanolamine, followed by resuspension in fresh 0.2 M ethanolamine and a 90 minute incubation with rotation at room temperature to quench residual DMP. Uncoupled IgGs were then removed by washing beads three times with 0.58% acetic acid in 150 mM NaCl. Finally, antibody-crosslinked beads were washed three times in cold PBS, followed by resuspension in PBS and storage at 4°C until use.

Table 1. List of antibodies used for co-immunoprecipitations and western blot analysis.

Antibody	Supplier and catalogue #	Final concentration
Rabbit PAP polyclonal	Gift of N. Tumer (Rutgers University)	1:5,000
Mouse 6xHis Tag monoclonal	Invitrogen MA1-21315	1:5,000
Anti-rabbit secondary	Sigma Aldrich A4914	1:10,000
Anti-mouse secondary	Sigma Aldrich A9044	1:10,000
Rabbit FLAG polyclonal	Sigma Aldrich F7425	n/a

2.2.2 Protein extraction from pokeweed

Pokeweed leaf tissue was homogenized by grinding in liquid nitrogen using a 2 mL tube and pestle. Four leaf discs (approximately 75 mg tissue) were homogenized in 300 µL cold protein extraction buffer (50 mM Tris-HCl pH 7.5, 1 mM EGTA, 1 mM dithiothreitol (DTT), 1x protease inhibitor (Thermo Scientific #A32963), 5% glycerol) and then incubated for 15 minutes on ice to allow solubilization of proteins. Lysates were centrifuged at 16,000 x g for 10 minutes at 4°C to pellet cellular debris, followed by transfer of supernatants to new tubes. A second centrifugation was performed at 16,000 x g and 4°C for 15 minutes and supernatant was again transferred to a new tube. Total protein concentration of lysate samples was quantified by Bradford assay using a 1:9 dilution of the lysate to prevent interference of the protein extraction buffer components.

2.2.3 PAP co-immunoprecipitation

Pokeweed lysate was added to either crosslinked PAP antibody- or crosslinked FLAG antibody-protein A magnetic beads; 1-1.5 mg total protein was added to each 40 µL bead sample with additional TBS (50 mM Tris-HCl pH 7.5, 150 mM NaCl) up to 1 mL. Samples were then rotated for 60 minutes at 4°C with end-over-end mixing. After incubation, beads were washed three times with TBS then co-immunoprecipitated proteins were eluted from beads in 30 µL 2x Laemmli buffer (125 mM Tris-HCl pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) b-mercaptoethanol, 0.02% (w/v) bromophenol blue).

2.2.4 In-gel reduction and alkylation

Beads were pelleted using a quick spin and then supernatant was loaded onto a 4-12% Bis-Tris precast gel (Millipore #MP41G12). Proteins were separated for 6 minutes at 150 V in MOPS/SDS running buffer (50 mM MOPS, 50 mM Tris, 3.5 mM SDS, 1 mM EDTA, pH 7.7). The gel was stained in Coomassie stain solution (0.1% (w/v) Coomassie blue R250, 50% (v/v) methanol, 10% (v/v) glacial acetic acid) for 15 minutes with rocking at room temperature, quickly rinsed two times with water, and then rinsed in water with rocking for 1 hour. Bands

were excised with a clean razor blade and then cut into cubes (approx. 1x1 mm). Gel pieces were transferred to a 1.7 mL microtube and briefly spun down. To destain gel pieces, 100 mM ammonium bicarbonate (ABC)/acetonitrile (1:1, v/v) was added to each tube and incubated at 37°C for 30 minutes at 300 rpm. Liquid was removed and the destaining step was repeated a second time. Acetonitrile (500 μL) was added to each tube and incubated for 10 minutes until gel pieces shrank and become opaque. Acetonitrile was removed and 10 mM DTT in 100 mM ABC was added to cover gel pieces. Gel pieces were then incubated for 30 minutes at 56°C for in-gel reduction of proteins. Tubes were cooled to room temperature and gel pieces were shrunk for 10 minutes in acetonitrile. Acetonitrile was removed followed by addition of 55 mM iodoacetamide in 100 mM ABC to cover gel pieces and incubated for 20 minutes at room temperature in the dark for alkylation of proteins. Acetonitrile was again added to shrink the gel pieces followed by removal of all liquid. Samples were finally washed in destaining solution again twice as described above.

2.2.5 In-gel digestion

Trypsin buffer (20 ng/μL trypsin in 10 mM ABC + 10% (v/v) acetonitrile) was added to cover gel pieces, and samples were incubated for 30 minutes on ice to allow for saturation with buffer. After 30 minutes, more trypsin buffer was added as necessary to cover gel pieces, followed by a 90-minute incubation on ice for diffusion of trypsin into the polyacrylamide. ABC (10-20 μL; 100 mM) was then added as needed to cover any exposed gel pieces due to buffer absorption, and tubes were incubated overnight at 37°C.

2.2.6 Peptide extraction

Liquid was removed from gel pieces to a new 1.7 mL microtube. To extract peptides from polyacrylamide matrix, 50 μL extraction buffer (1:2 (v/v) 5% formic acid/acetonitrile) was added and tubes were incubated at 37°C for 15 minutes at 300 rpm. Extraction buffer was then removed and added to the trypsin buffer fraction. A second extraction step using 50 μL extraction buffer was performed as described above and pooled with the first two fractions. Combined fractions were then dried down at 50°C in a vacuum centrifuge. Dried extracts were resuspended in 20 μL 0.1% formic acid in MS grade water and stored at -20°C until further analysis.

2.2.7 Mass spectrometry

MS analysis was performed by Maxime Rossato at the mass spectrometry centre (York University) using the Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) coupled to an Easy-nLC-1000 nanoflow LC system (Thermo Scientific). Separation of peptides (2 µL) was performed on an analytical column (75 µm x 20 cm) packed in-house with C18-AQ (Dr. Maisch GmbH, Ammerbuch, Germany, 3 µm particle size, pore size 120 Å) using a gradient of solvent A (0.1% formic acid in water) and solvent B (ACN with 0.1% formic acid) over 45 minutes from 1 to 35 % B at a flow rate of 300 nL/min. Two LC-MS/MS runs per sample lot were performed. Data-dependent MS acquisition in positive mode was achieved at a resolution of 60 000 for the top ten precursor ions with a scan range of 200– 2000 m/z. For MS/MS, top intense precursor ions were subjected to higher-energy collision dissociation with 35% normalized collision energy. The fragment ions were detected at a mass resolution of 30 000 at m/z 400.

2.2.8 Data analysis

The MS/MS searches were performed using SEQUEST HT search algorithm through Proteome Discoverer (version 2.3.0.523) software (Thermo Scientific) against a MAKER-annotated pokeweed protein database, assembled from the pokeweed genome, as described in Neller et al. 2019. The search workflow included spectrum selector, minora feature detector, spectrum files, Sequest HT, and Target decoy PSM validator. Carbamidomethylation of cysteine, acetylation of protein N-termini, and oxidation of methionine were set as variable modifications. MS and MS/MS mass tolerances were set to 10 ppm and 0.02 Da, respectively. The data was searched against the decoy database, and 1% false discovery rate (FDR) at both peptides and peptide spectral match levels was applied for the analysis.

2.3 PAP interactor analysis

2.3.1 Identification of homologous proteins

For identification of proteins with highest sequence similarity to co-immunoprecipitated pokeweed proteins, a BlastP search was performed using the pokeweed protein amino acid

sequence as the query against the Embryophyta clade in the SWISS-PROT database. The SWISS-PROT database was chosen to restrict results to annotated proteins. Only hits with E-values < 1 x 10^{-4} were considered to ensure homologous proteins were related to query sequence with an error rate < 0.01%. The protein hit with the highest total score were then used to annotate the pokeweed protein. Subcellular localizations and biological functions of pokeweed proteins were annotated using homologous protein information taken from the UNIPROT database (Bateman et al. 2021).

2.3.2 Interaction mapping and functional enrichment analysis

Protein interaction maps were generated with the STRING database (v. 11.5; Szklarczyk et al. 2021) using *Beta vulgaris* protein homologs. High confidence interaction scores (>0.7) were used for all types of evidence (experiments, databases, textmining, co-expression, neighbourhood, gene fusion, co-occurrence). Second shell proteins were expanded to include a maximum 10 interactions, to ensure ability to detect statistically significant enriched interaction networks. GO terms associated with proteins were assigned using Blast2GO (OmicsBox), classified into biological process, molecular function, and cellular localization. A one-sided Fisher's exact test (P<.01) was performed for each interaction cluster to detect associated GO term enrichment.

2.3.3 *In silico* analysis of paCP1 sequence

An Interproscan (v. 77; Blum et al. 2021) search was performed to identify protein domains, protein superfamily, and active site encoded within the paCP1 amino acid sequence. Global alignment of the paCP1 sequence to the AtXCP1 sequence was performed using EMBOSS needle (Madeira et al. 2019). SignalP (v. 5.0; Armenteros et al. 2019) was used to predict the presence of a signal peptide in the paCP1 amino acid sequence.

2.3.4 Co-expression analysis

Data used for differential gene expression in response to JA and SA treatment were obtained by previous lab members; methods for stress treatments, mRNA sequencing, transcriptome assembly, and differential gene expression analysis are described in Neller et al (2019).

For promoter analyses, a 1.3 kb sequence upstream of the MAKER-annotated +1 transcription start site (TSS) in the pokeweed genome was used for each gene. Annotation of JA-responsive cis-regulatory elements (CREs) was performed by inputting promoter sequences individually into the promoter analysis tool of PlantPan 3.0 (http://plantpan.itps.ncku.edu.tw/index.html) against the *Arabidopsis thaliana* transcription factor (TF) binding motifs database (Chow et al. 2019). Results were then filtered using locus identifiers for the TF subset of interest (MYC2, MYC3, MYC4, DREB2B, RVE2, ZAT18, TCP23, NAC3, WRKY51, and ZAT10) and only motifs with a similarity score ≥ 95% were considered. Determination of co-occurring TF binding motifs was performed by querying a pair of promoter sequences (paCP1 and each of the PAP isoform promoters in turn) using the gene group analysis function of PlantPan 3.0 against the *Arabidopsis thaliana* TF binding motifs database. To reduce false positives, only co-occurrent TFs with a threshold score of 100% were considered.

2.4 PAP-protein interaction validation by reverse co-IP

2.4.1 Generation of T7-His-XCP1 clone for protein expression in *E. coli*

The plasmid backbone for T7-His-XCP1 was a modified pET28a vector (MilliporeSigma) in which: (1) a TEV cleavage site was inserted in place of the original thrombin site, and (2) the multiple cloning site was replaced with a ligation independent cloning site. Both modifications were done by previous members of the Hudak lab. The pET28a vector is a bacterial expression vector where inserted genes are under control of the T7 lac promoter and can be expressed with either an N-terminal or C-terminal 6x-His tag. The plasmid also encodes the lacI repressor protein gene for the lac operon system and expresses a kanamycin resistance gene.

The pET28a vector was linearized by PCR amplification (see **2.4.2**) using forward primer pET-Vec-For and reverse primer pET-Vec-Rev with sequences listed in Table 2; pET-Vec-For was designed to remove the C-terminally expressed 6x-His tag from the backbone during PCR amplification.

Template for PCR amplification of the XCP1 insert was a gene fragment synthesized by Twist Bioscience (USA). Gene-specific primers (Table 2) used to amplify the XCP1 insert were designed with overhangs homologous to the ends of the linear pET28a vector sequence to allow for DNA assembly by sequence- and ligation independent-cloning. Cloning of the T7-His-XCP1 plasmid used for protein expression in *E. coli* was therefore a two-fragment cloning reaction (described in **2.4.6**). Correct insertion of the XCP1 sequence was verified by Sanger sequencing by The Centre for Applied Genomics at Sick Kids (Canada) using the universal T7 primers listed in Table 2 (T7-For and T7-Rev).

Table 2. Primer sequences used for cloning of T7-His-XCP1.

Primer name	Sequence (5'–3')	Annealing temperature used (°C)
pET-Vec-For	GATCCGGCTGCTAACAAAG	55-65°C
pET-Vec-Rev	GGATTGGAAGTACAGGTTTTCA	55-65°C
XCP1-Ins-For	CATCACCATGAAAACCTGTACTTCCAATC CCAGGATTTCTCCGTCGTC	60°C
XCP1-Ins-Rev	TGTCGACGGCGCTCGAATTCGGATCCGTT AGTATTTCTTGATAGGATAAACGG	60°C
T7-For	TAATACGACTCACTATAGGG	55°C
T7-Rev	GCTAGTTATTGCTCAGCGG	55°C

2.4.2 Polymerase chain reaction (PCR)

PCR was used for amplification of DNA fragments for cloning. Either pET28a plasmid or XCP1 gene fragment were used as template (0.5 ng or 1 ng, respectively). All 25 µl PCR reactions contained the following components: 5 µl of 5X Q5 buffer (New England Biolabs #B9027), 1.25

µl of 10 µM forward primer, 1.25 µl 10 µM reverse primer, 0.5 µl 10 mM dNTPs, 0.25 µl Q5 High-Fidelity DNA Polymerase (New England Biolabs #M0491), DNA template, and ddH_2O. PCR amplification was carried out using an ep gradient S thermocycler (Eppendorf AG) with the following settings: a 60 second initial denaturation at 98°C, followed by 40 cycles with a 15 second denaturation at 98°C, a 30 second annealing step at a primer specific temperature (Table 2), and an extension step at 72°C for 30 seconds per kb of DNA. The cycles were followed by a final 2-minute extension step at 72°C. PCR products (4 µl) were visualized on a 1-2% agarose gel (1% gel for products ≥ 1 kb, 2% gel for products < 1 kb) to check for correct size of product amplification.

2.4.3 DpnI digest

To remove template DNA from PCR reactions, 0.5 µL DpnI (NEB #R0176) and Cutsmart buffer (NEB# B7204; 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/mL BSA) were added to 25 µL PCR reaction and tubes were incubated at 37°C for 2 hours.

2.4.4 PCR clean-up

Following digest, reactions were PCR purified for removal of reaction components. Buffer PB (5 M guanidine hydrochloride, 20 mM Tris-HCl pH 6.6, 30% (v/v) ethanol) was added to DpnI digested reactions at a 5:1 volume ratio, then vortexed briefly and loaded onto a silica-membrane column (Enzymax #EZC101). Columns were centrifuged at 10,000 x g for 2 minutes followed by two washes with 700 µL PE buffer (16 mM NaCl, 1.6 mM Tris-HCl pH 7.5, 80% (v/v) 100% ethanol) where columns were centrifuged for 2 minutes at 10,000 x g. Columns were then dried by centrifugation for 1 minute at 10,000 x g before elution (1 minute at 10,000 x g) using 50 µL of 5 mM Tris-HCl (pH 8) into a new microtube.

2.4.5 Low-melt agarose gel extraction

Low-melt agarose gel extraction was used to purify PCR-generated fragments showing non-specific bands when separated on an agarose gel. DNA fragments were loaded on an 1% low-melt agarose precast with Midori Green dye (1 µL/mL) and separated for 90 minutes at 70 V.

Desired fragments were excised using a handheld blue light and a 510/520 nm bandpass filter and then extracted by adding 300 µL QG buffer (5.5 M guanidine thiocyanate, 20 mM Tris-HCl pH 6.6) per 100 mg of agarose and vortexed to dissolve. This mixture was loaded onto a silica column (Enzymax #EZC101) and centrifuged at 10,000 x g for 2 minutes. Flow-through was discarded. The column was washed twice with 700 µL PE buffer (16 mM NaCl, 1.6 mM Tris-HCl pH 7.5, 80% (v/v) 100% ethanol) where columns were centrifuged for 2 minutes at 10,000 x g and flow-through discarded. Columns were then dried by centrifugation for 1 minute at 10,000 x g and finally DNA was eluted by adding 50 µL of 5 mM Tris-HCl (pH 8) followed by centrifugation at 10,000 x g for 2 minutes into a new microtube. DNA concentration was quantified by A_{260} using a spectrophotometer.

2.4.6 TEDA cloning

For the two fragment assembly to ligate the XCP1 gene sequence into the modified pET28a backbone, T5 exonuclease-dependent assembly (TEDA) was used. TEDA cloning utilizes T5 exonuclease to create short homologous ends on the fragments, which anneal and then are further repaired in *E. coli* after transformation (Xia et al. 2015). For TEDA cloning of T7-His-XCP1, linear pET28a backbone was combined with XCP1 gene insert at a 1:3 vector to insert molar ratio (100 fmol of backbone: 300 fmol insert). The cloning reaction also included 3 µL of 5X TEDA mastermix (10 U/µL T5 exonuclease (New England Biolabs #M0363), 0.5 M Tris-HCl (pH 7.5), 50 mM $MgCl_2$, 50 mM DTT, 0.25 g PEG 8000) and ddH_2O to 15 µL. The reaction was incubated for 40 minutes at 30°C in a thermocycler. TEDA reaction product was then transformed into chemically competent DH5a (see **2.4.7**).

2.4.7 Plasmid transformation

Plasmid (2 µL TEDA reaction product, or 50-100 ng miniprepped plasmid) was added to thawed chemically competent DH5a cells (50 µL or 100 µL for plasmid and TEDA transformations, respectively) and incubated for 30 minutes on ice. Cells were then heat shocked at 42°C for 45 seconds. After cooling on ice for 2 minutes, 900 µL of Super Optimal Broth (SOB) was added and cells were recovered at 37°C for 60 minutes at 250 rpm. Cells were pelleted by centrifugation at 10,000 x g for 60 seconds and 900 µL of supernatant was discarded. The pellet

was resuspended in the remaining liquid followed by plating on LB plates with selective antibiotic. Plates were incubated overnight at 37°C for colony growth.

2.4.8 Colony PCR

Colony PCR was used to check individual *E. coli* colonies transformed with TEDA cloning reactions for correct fragment size insertion. Single colonies were resuspended in 20 µL ddH_2O of which 1 µL which was used to streak new selective LB plates as well as inoculate 5 mL cultures for plasmid isolation as described below in **2.4.9**. The remaining liquid was then boiled at 95°C for 5 minutes to lyse cells, followed by vortexing for 20 seconds and centrifugation at 16,000 x g for 10 minutes to pellet cellular debris. Supernatant was used for colony PCR template. Amplification was performed as described in **2.4.2**. using primer pair T7-For and T7-Rev (Table 2). PCR products (6 µL) were separated on a 1.5% agarose gel to check size of desired product amplification.

2.4.9 Small-scale plasmid isolation (miniprep)

A single *E. coli* colony transformed with T7-His-XCP1 was grown in 5 mL Lysogeny broth (LB) with kanamycin (50 µg/mL) at 37°C overnight with shaking at 250 rpm. Cultures were then centrifuged at 12,000 x g for 3 minutes and supernatant was removed completely. Pellets were resuspended in 500 µL buffer P1 (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 100 µg/mL RNAse A) and incubated for 1 minute. To lyse cells, 500 µL P2 buffer (200 mM NaOH, 1% (v/v) SDS) was added and tubes were gently inverted 4-6 times during a 60 second period. To neutralize lysis buffer for DNA binding, 750 µL N3 buffer (4 M guanidine hydrochloride, 0.5 M potassium acetate, pH 4.2) was added and tubes were incubated for 1 minute. Tubes were centrifuged at 12,000 x g for 5 minutes and then supernatant was applied to a silica-membrane column (Enzymax #EZC101) and incubated for 2 minutes. Column with lysate was then centrifuged at 10,000 x g for 2 minutes and flow-through was discarded. Column was washed twice by adding 700 µL PE buffer (16 mM NaCl, 1.6 mM Tris-HCl pH 7.5, 80% (vol/vol) 100% ethanol) and centrifuging for 2 minutes at 10,000 x g, discarding flow-through. Finally, column was spun-dry for 1 minute at 10,000 x g. To elute the DNA, the column was placed into a new microtube and 50 µL of 10 mM Tris-HCl pH 8.0 was applied to the centre of the column,

followed by centrifugation at 10,000 x g for 2 minutes. The DNA was quantified using a spectrophotometer (A_{260} nm) and plasmid purity was visualized by separating 0.5 µg DNA on a 1% agarose gel.

2.4.10 Growth and induction of Rosetta cultures

Rosetta strain *E. coli* (Sigma-Aldrich #70954) were used for production and purification of paXCP1 protein for use in reverse co-IP. Rosetta cells are a BL21 derivative *E. coli* strain used to enhance eukaryotic protein expression that requires codons used rarely by prokaryotes; they carry a chloramphenicol-resistant plasmid encoding the rare tRNA genes. Transformation of Rosetta cells with T7-His-XCP1 plasmid was performed as described in **2.4.7**. Single colonies were grown in 2 mL LB with chloramphenicol (35 µg/mL) and kanamycin (50 µg/mL) until the OD_{600} reached 0.5-0.7 when measured by spectrophotometer (approximately 4 hours). The small cultures were then transferred to 250 mL LB cultures with 35 µg/mL chloramphenicol and 50 µg/mL kanamycin and further incubated at 37°C and 275 rpm until OD_{600} of large culture reached 0.6-1.0 (approx. 4 hours). Protein production by Rosetta cells was then induced by adding 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) to the cultures. Induction proceeded for 4 hours at 37°C and 275 rpm after which cells were transferred to 250 mL Nalgene polycarbonate centrifuge bottles and pelleted by centrifuging at 4000 x g for 15 minutes at 4°C. Cells were washed once in 20 mM Tris-HCl (pH 8.0) and then pelleted a second time at 4000 x g and 4°C for 15 minutes. Supernatant was removed completely and cells were frozen at -20°C.

2.4.11 Cell lysis and protein purification

Pellets were thawed on ice and each gram of pellet was resuspended in 7 mL cold lysis buffer (50 mM Tris HCl pH 8.0, 300 mM NaCl, 5% (v/v) glycerol). Resuspended pellets were pooled into a 50 mL conical tube and sonicated for 15 seconds at ~18 micron amplitude, followed by 2 minutes on ice in between sonication cycles. Pellets were sonicated for approximately 10-15 sonication cycles or until foaming was seen. Supernatant was transferred to a pre-cooled 30 mL Nalgene centrifuge tube and centrifuged at 10,000 x g for 15 minutes at 4°C. Supernatant was then transferred to a fresh 50 mL conical tube with 30 µL reserved for analysis by sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE). The remaining supernatant

was loaded onto 250 µL Ni Sepharose 6 Fast Flow resin (Cytiva #17531806) in a 5 mL polypropylene column (Thermo Scientific #29922) which had been previously washed with 3 column bed volumes of lysis buffer. Loaded beads were then washed once with 30 mL 1mM imidazole in lysis buffer and once with 30 mL 10 mM imidazole in lysis buffer; 30 µL of each flow-through was reserved for SDS-PAGE. Bound protein was then eluted with 15 mL lysis buffer containing 250 mM imidazole. Reserved fractions including an elution fraction sample (30 µL) were mixed with equal volume 2X Laemmli buffer then analyzed by SDS-PAGE (see **2.4.14**) to determine successful purification of target protein before continuing with concentration.

2.4.12 Protein concentration

Purified protein fractions were concentrated using an Amicon Ultra-15 centrifugal filter unit with a 10 kDa molecular weight cut-off (Millipore #UFC9010). Filter units were first washed by centrifuging 10 mL ddH_2O for 5 minutes at 3275 x g in a swinging bucket centrifuge. Eluted fractions containing the protein of interest were then concentrated by centrifugation of 15 mL aliquots at 3275 x g for 15 minutes at 4°C until approximately 500 µL of liquid remained in the upper chamber of unit. The membrane was then washed with 40 mL of column wash solution (20 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM DTT, 100 mM NH_4Cl) in the same way until approximately 500 µL of liquid remains. The remaining 500 µL was then pipetted into a new 1.7 mL microtubule and 50 µL 100% glycerol was added. Purified and concentrated protein was quantified by spectrophotometer (A_{280} nm) and then frozen at -80°C until further use.

2.4.13 Reverse co-immunoprecipitation of His-tagged protein

For each reverse co-immunoprecipitation sample, 2 µL 6x-His tag monoclonal antibody (Table 1) was added to 5 µg of purified His-tagged protein with additional PBS to 1 mL in a 1.7 µL microtube. Antibody and protein were incubated for 90 minutes at 4°C with end-over-end rotation. Protein G magnetic beads (40 µL; New England Biolabs #S1430S) which had been washed three times with 1 mL PBS were added to the tube and incubated a further 90 minutes at 4°C with rotation. Pokeweed and tobacco lysates were prepared as described in **2.2.2** except using a bead beater (BioSpec 3110BX MiniBeadBeater) instead of a mortar and pestle to

homogenize the tissue. For this method, approximately 200 μL 1.0 mm glass beads were placed in a 2 mL microtube with 300 μL protein extraction buffer and four leaf discs, then tubes were placed in the bead beater for 1 minute at setting 48. Tubes were cooled on ice for 5 minutes followed by a second cycle in the bead beater. Samples were incubated for 15 minutes on ice to allow protein solubilization then lysates were cleared and proteins quantified as previously described. Pokeweed lysate or tobacco lysate (total protein = 1.5 mg) was added to protein-conjugated antibody bead samples with TBS added to 1 mL, followed by a 1-hour incubation at 4°C with end-over-end rotation. Samples were then washed 3 times in TBS and proteins were eluted from beads using 30 μL 2x Laemmli buffer.

2.4.14 SDS-PAGE and Western blotting

Protein samples to be analysed by SDS-PAGE were denatured by heating at 80°C for 5 minutes. Tubes were briefly centrifuged to bring down insoluble matter and supernatants were loaded onto a 12% gel. Proteins were separated at 200 V for 50 minutes in running buffer (192 mM glycine, 25 mM Tris-HCl pH 8.3, 0.1% SDS). Proteins resolved by SDS-PAGE were then transferred to a nitrocellulose membrane using a wet transfer system run for 45 minutes at 120 V in transfer buffer (160 mM glycine, 25 mM Tris-HCl pH 8.3, 20% (v/v) methanol, 0.02% SDS). The membrane was then blocked in 5% bovine serum albumin (BSA) in TBS for 1.5 hours at room temperature with rocking. After blocking the blot was incubated overnight with rocking at 4°C with the appropriate primary antibody (Table 1) in 5% BSA in TBS-T (TBS with 0.1% Tween 20). The following day the blot was washed 3 times in TBS-T for 5 minutes each with rocking at room temperature. Then, 5% BSA in TBS-T with the appropriate horseradish peroxidase-conjugated secondary antibody (Table 1) was added to the blot and incubated for 1 hour at room temperature with rocking. The blot was washed twice for 5 minutes each in TBT-T, once for 5 minutes in TBS, and once for 5 minutes in ddH_2O, all at room temperature with rocking. Finally, 1 mL of luminol reagent and 1 mL of peroxide reagent (Thermo Scientific #32106) were applied to cover the blot and incubated for 5 minutes before imaging using a MicroChemi imager (DNR Bio-imaging systems) with a 5-minute exposure.

3. RESULTS

3.1 Identification of PAP-protein interactions

CoIP-MS was used to identify PAP-protein interactors for this work and was performed using pokeweed lysate for mapping the PAP interactome (Figure 4A). Leaves from pokeweed plants at the six-leaf stage were chosen as our lab has previously shown that PAP expression is low before this stage (unpublished data). To verify successful immunoprecipitation of PAP, proteins from a PAP co-IP sample were eluted from the beads with Laemmli buffer, followed by Western blot using a PAP polyclonal antibody. Successful pull-down of PAP was demonstrated by a ~29 kDa band in PAP co-IP samples (Figure 4B). A band at ~58 kDa also corresponds to the previously identified PAP homodimer (Tourlakis et al. 2010).

Since pokeweed is resistant to transformation (our unpublished data; Kanzaki et al. 1999), transgenic pokeweed PAP knockout lines do not exist to serve as a control without PAP expression for non-specific binding to the antibody chains and protein A-magnetic bead support. Instead, a rabbit IgG FLAG antibody was used as a co-IP control for non-specific binding to the support matrix; this represents an antibody of the same species and isotype as the PAP polyclonal antibody, where the FLAG tag is not present endogenously in pokeweed.

Four biological replicates were sent for analysis by mass spectrometry, comprising four PAP co-IP samples and three FLAG co-IP samples. A total of 77 proteins were identified, using an applied FDR of < 1% at both the peptide and the protein level to ensure high confidence in protein identification (Table S1; Appendix B). Thirty-seven proteins were identified in PAP co-IP samples only, 9 proteins in FLAG co-IP samples only, and 31 proteins in both PAP and FLAG co-IP samples (Figure 4C). Four PAP isoforms were identified in the pull-down using the polyclonal PAP antibody (Table 3): PAP-1, PAP-2, K-PAP, and a novel PAP isoform, which were each identified in at least two replicates. The remaining 33 proteins identified only in PAP co-IP samples were considered PAP-specific interactors; however, to further minimize the chance of false positives, only those identified in at least two biological replicates were accepted as high confidence interactors. These six PAP interactors are organized by pokeweed gene ID and presented in Table 4. As pokeweed proteins other than PAP have not previously been of

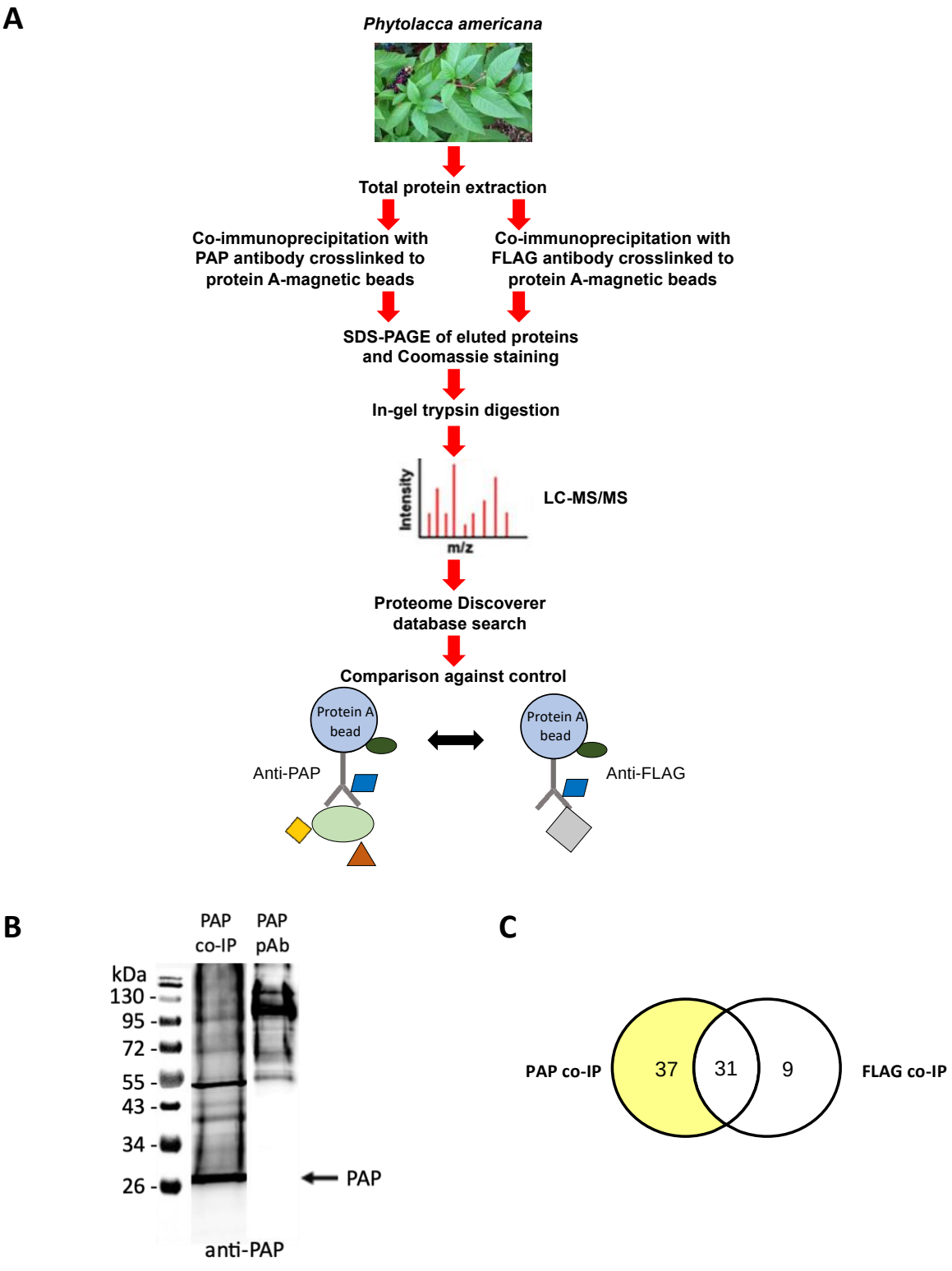

A
Phytolacca americana
Total protein extraction
Co-immunoprecipitation with PAP antibody crosslinked to protein A-magnetic beads
Co-immunoprecipitation with FLAG antibody crosslinked to protein A-magnetic beads
SDS-PAGE of eluted proteins and Coomassie staining
In-gel trypsin digestion
Intensity
m/z
LC-MS/MS
Proteome Discoverer database search
Comparison against control
Protein A bead
Anti-PAP
Protein A bead
Anti-FLAG
B
PAP co-IP
PAP pAb
kDa
130
95
72
55
43
34
26
PAP
anti-PAP
C
PAP co-IP
37
31
9
FLAG co-IP

Figure 4. Co-immunoprecipitation coupled with mass spectrometry for mapping of the PAP-protein interactome. (A) Co-IP/MS workflow. Total protein extraction was performed on 6-leaf *Phytolacca americana* plants and lysate was incubated with either PAP or FLAG rabbit IgG antibody crosslinked to protein A magnetic beads for co-immunoprecipitation. Antibody bound protein complexes were eluted in 2x Laemmli buffer and separated via SDS-PAGE, followed by Coomassie staining and in-gel trypsin digestion. Extracted peptides were analyzed by LC-MS/MS and obtained mass spectra were compared to a pokeweed protein database using Proteome Discoverer (ThermoFisher). Proteins identified in both control and experimental samples were removed from the list of putative PAP interactors. **(B)** Eluted proteins from a PAP co-IP sample were separated on a 12% Bis-Tris gel then visualized by immunoblot using a polyclonal PAP antibody and an antirabbit HRP-conjugated secondary antibody. PAP antibody-conjugated protein A-magnetic beads were separated as a control for binding of secondary antibody to co-eluted PAP antibody chains in the co-IP sample. **(C)** Venn diagram showing number of proteins pulled-down by PAP co-IP compared to the FLAG control co-IP. Specific PAP interactors are shown shaded in yellow.

Table 3. **PAP isoforms identified from PAP co-immunoprecipitation, with subcellular localization and biological function, ordered by gene ID.**

Pokeweed gene ID	Protein name	Subcellular localization	Biological function
PHYAM_010465	Novel PAP	Unknown	Unknown
PHYAM_012451	Antiviral protein alpha-like (K-PAP)	Apoplast[1]	Antiviral defense, negative regulation of translation
PHYAM_020596	Antiviral protein I (PAP-1)	Apoplast, cytoplasm[2]	Antiviral defense, negative regulation of translation
PHYAM_028184	Antiviral protein 2 (PAP-2)	Apoplast[1]	Antiviral defense, negative regulation of translation

[1]= presumed due to predicted N-terminal signal peptide (unpublished data), [2]= determined by immunohistochemistry (Ready et al. 1986)

Table 4. High confidence PAP specific interactors ordered by pokeweed gene ID, with homologous protein name, subcellular localization and biological function as obtained from Uniprot database. Protein score is the mean of all SEQUEST HT scores for each protein across replicates. The protein interactor chosen for further interaction validation is highlighted in yellow.

Pokeweed gene ID	Homologous protein name	Homologous protein species	Subcellular localization	Biological function	Protein score
PHYAM_000754	Fructose-bisphosphate aldolase (ALDP)	*Oryza sativa subsp. japonica*	Chloroplast, cytosol	Fructose 1,6-bisphosphate metabolic process, gluconeogenesis, glycolytic process	17.03
PHYAM_002561	Carbonic anhydrase (CA)	*Spinacia oleracea*	Chloroplast	Carbon utilization	22.49
PHYAM_003326	Elongation factor 1-alpha (EF-1-alpha)	*Solanum lycopersicum*	Cytoplasm	Translational elongation	71.78
PHYAM_018582	Glyceraldehyde-3-phosphate dehydrogenase B (GAPB)	*Spinacia oleracea*	Chloroplast	Glucose metabolism, reductive pentose-phosphate cycle	126.25
PHYAM_026164	Xylem cysteine protease (XCP1)	*Arabidopsis thaliana*	Extracellular space, lysosome, nucleus, plasma membrane, vacuole	Programmed cell death, proteolysis in protein catabolism	94.07
PHYAM_027772	40S ribosomal protein S26-1 (RPS26A)	*Arabidopsis thaliana*	Cytosol	Translation	11.96

scientific interest and the pokeweed genome was only recently sequenced to allow for this type of work (Neller et al. 2019), little is known regarding the pokeweed proteins identified in this work. Therefore, to infer protein identity, subcellular localization, and biological function of the novel proteins, a blastP search of the PAP interactor amino acid sequences was performed against the Embryophyta (land plants) clade in the SWISS-PROT database, to identify annotated plant proteins with the highest degree of homology. Only reference sequences with E-values < $1x10^4$ were considered (indicating that the result sequence is related to the query sequence with an error rate < 0.01%). The reference sequence with the highest total score was used to annotate the corresponding pokeweed protein (Table 4). Subcellular localizations and biological functions were then hand annotated using information taken from the UNIPROT database (Bateman et al. 2021). Results showed potential subcellular localization of three PAP protein interactors to the chloroplast, two to the cytoplasm, and one to the extracellular space (Table 4). Annotation of biological function indicated involvement of two proteins (PHYAM_003326 and PHYAM_027772) in translation; however, the remaining four protein interactors participated in a range of biological processes, including gluconeogenesis, carbon utilization, glucose metabolism, programmed cell death and cellular development.

3.2 Interaction mapping and functional enrichment analysis

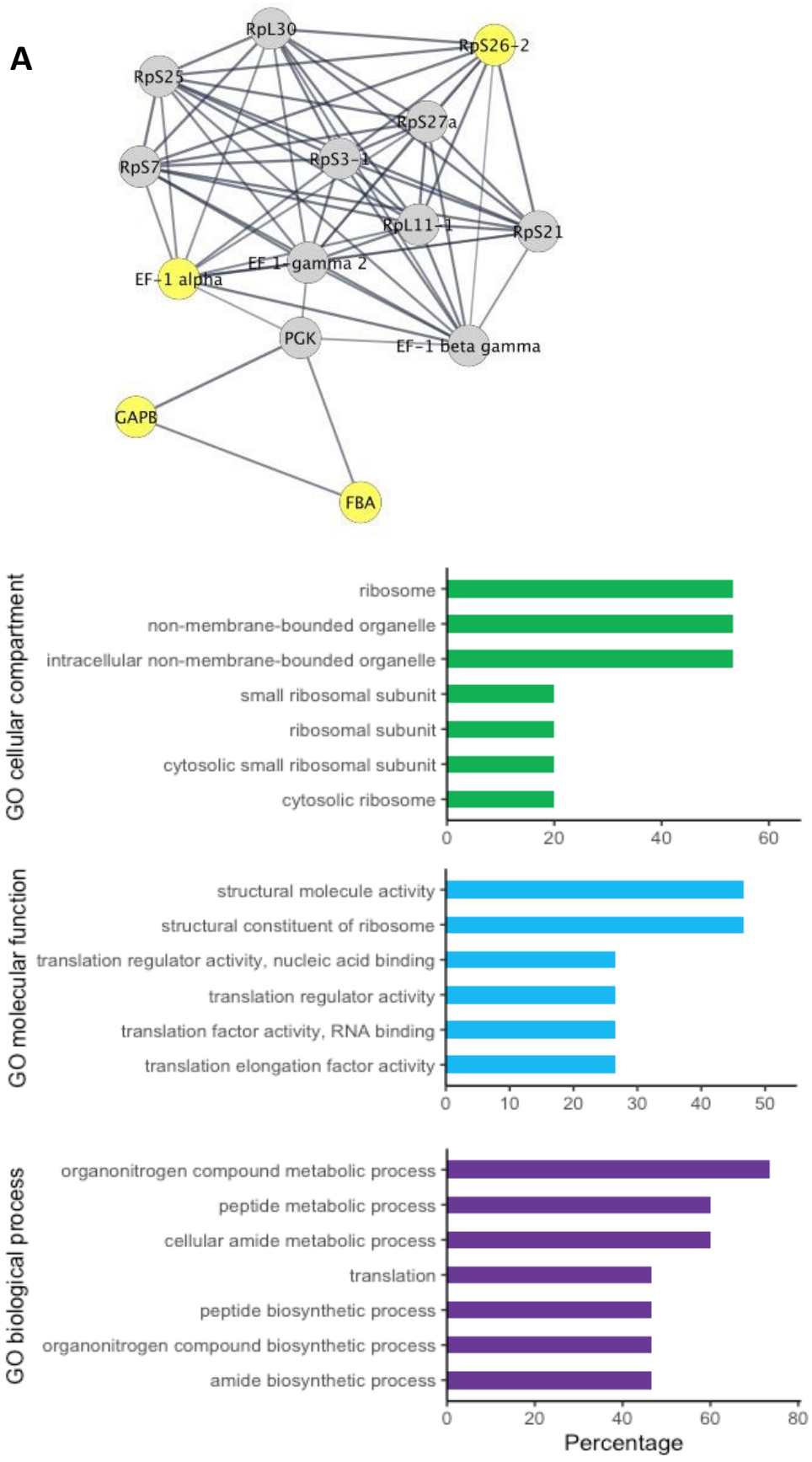
To identify whether the six high confidence PAP interactors were involved in common protein networks, the STRING database was used to generate protein interaction maps (Szklarczyk et al. 2021). As pokeweed is not an organism available in the STRING interaction database, *Beta vulgaris* homologs were chosen, as *B. vulgaris* is known to express extracellularly localized type 1 RIPs (Iglesias et al. 2015). This provided a more biologically relevant example for pokeweed RIP interactome maps than model organisms like *A. thaliana*, which have been found not to express RIPs (Kaul et al. 2000), or *Ricin communis*, which expresses the well-studied RIP ricin, a type II RIP that does not localize to the apoplast (Jolliffe et al. 2004). Additionally, to allow for statistical significance of interaction maps, a maximum of 10 proteins reported to interact with the PAP interactor proteins but not PAP itself, called second shell interactors, were included (Table S2; Appendix B). The predicted protein interactome map showed that out of the six high confidence PAP interactors identified in this study, four proteins (~67%) were connected to each

other in the STRING database (Figure 5A; P=1.08 x 10^{-6}). The remaining two proteins, carbonic anhydrase and xylem cysteine protease 2, were not connected to this interaction cluster or to each other, indicating that they were likely involved in separate functional clusters. Therefore, each of the two lone proteins were inputted separately to produce maps of their own clusters (Figures 5B and 5C; P=1.0 x 10^{-16} and P=1.56 x 10^{-2}, respectively).

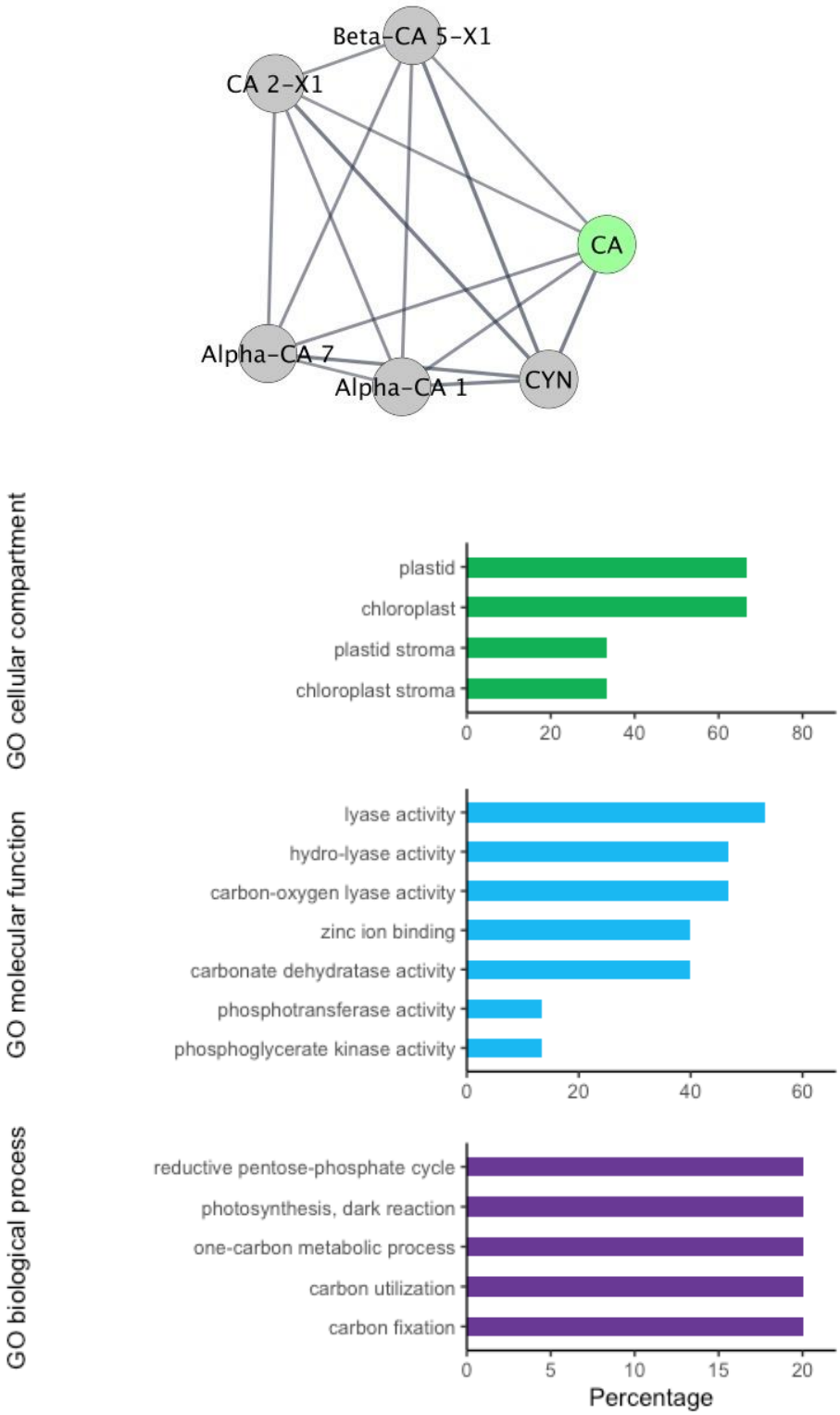
These three marked clusters indicate that PAP-1 interacting proteins are involved in distinct biological processes. Gene Ontology (GO) was therefore used to identify biological functions, molecular functions and cellular compartments associated with each of the three clusters. A one-sided Fisher's exact test identified GO terms associated with each of the three clusters (Figure 5A, 5B, and 5C; P<.01). The majority of cluster 1 proteins show a preferred subcellular localization to the ribosome (~53%) and involvement in the organonitrogen metabolic process (~73%) and the peptide metabolic process (~60%); this indicates that cluster 1 proteins are involved in protein synthesis by the ribosome. Cluster 2 proteins, which are those that interact with carbonic anhydrase, are associated with the chloroplast (~67%) and lyase activity (53%). Cluster 3 proteins, those interacting with XCP2, demonstrate a strong association with the endoplasmic reticulum (~75%) and endomembrane system (~58%), and exactly 50% of proteins in this cluster are involved in protein folding as a biological process. Taken together, these results indicate that PAP may have additional roles in the cell other than its known negative regulation of translation, influencing diverse cellular processes such as photosynthesis and protein folding.

3.3 *Phytolacca americana* cysteine protease I

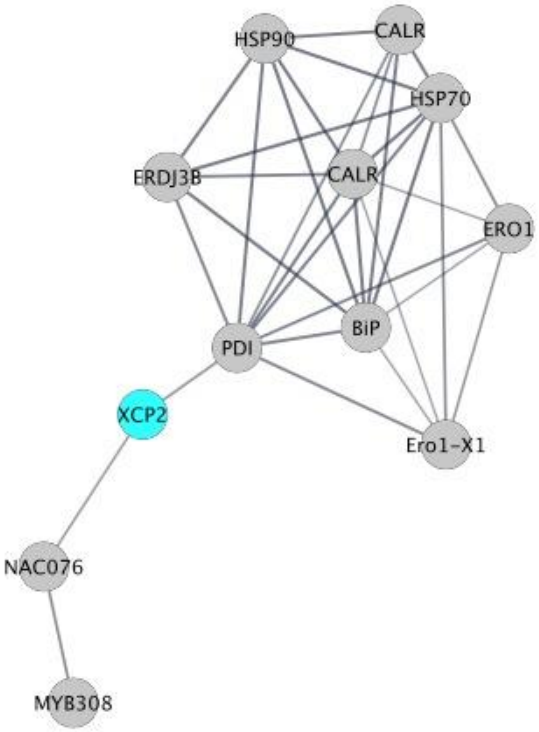
The pokeweed protein PHYAM_026164 was chosen for validation of its interaction with PAP to further confirm accuracy of the coIP-MS results. This protein was chosen due to potential extracellular localization, inferred from homology with *Arabidopsis thaliana* xylem cysteine protease I (AtXCP1; Table 4). Colocalization of PAP and an interactor to the extracellular space would provide insight into PAP's role in the apoplast. Since cysteine proteases have been implicated as important hubs in plant immunity (Misas-Villamil et al. 2016), validation of a PAP-cysteine protease interaction would also allow insight into the function of PAP in plant


A
RpL30
RpS26+2
RpS25
RpS27a
RpS3-1
RpS7
RpL11-1
RpS21
EF-1 gamma 2
EF-1 alpha
PGK
EF-1 beta gamma
GAPB
FBA
GO cellular compartment
ribosome
non-membrane-bounded organelle
intracellular non-membrane-bounded organelle
small ribosomal subunit
ribosomal subunit
cytosolic small ribosomal subunit
cytosolic ribosome
0
20
40
60
GO molecular function
structural molecule activity
structural constituent of ribosome
translation regulator activity, nucleic acid binding
translation regulator activity
translation factor activity, RNA binding
translation elongation factor activity
0
10
20
30
40
50
GO biological process
organonitrogen compound metabolic process
peptide metabolic process
cellular amide metabolic process
translation
peptide biosynthetic process
organonitrogen compound biosynthetic process
amide biosynthetic process
0
20
40
60
80
Percentage

B


Beta-CA 5-X1
CA 2-X1
CA
Alpha-CA 7
Alpha-CA 1
CYN
GO cellular compartment
plastid
chloroplast
plastid stroma
chloroplast stroma
0
20
40
60
80
GO molecular function
lyase activity
hydro-lyase activity
carbon-oxygen lyase activity
zinc ion binding
carbonate dehydratase activity
phosphotransferase activity
phosphoglycerate kinase activity
0
20
40
60
GO biological process
reductive pentose-phosphate cycle
photosynthesis, dark reaction
one-carbon metabolic process
carbon utilization
carbon fixation
0
5
10
15
20
Percentage

C


HSP90
CALR
HSP70
ERDJ3B
CALR
ERO1
BiP
PDI
XCP2
Ero1-X1
NAC076
MYB308

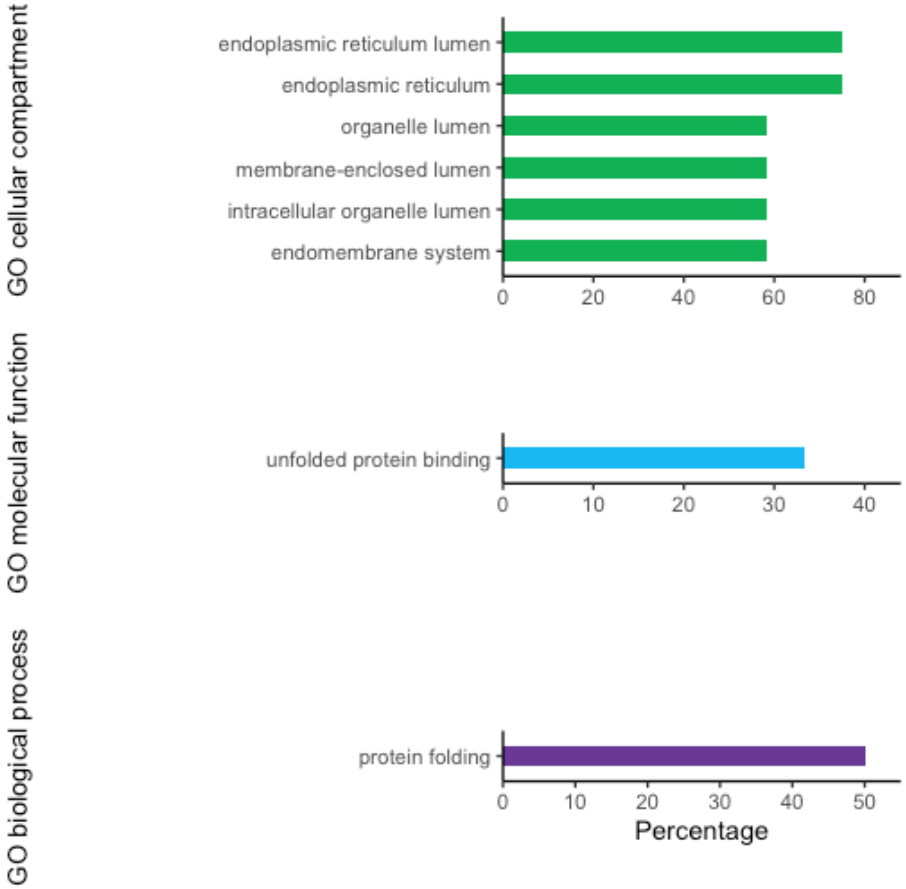

GO cellular compartment
endoplasmic reticulum lumen
endoplasmic reticulum
organelle lumen
membrane-enclosed lumen
intracellular organelle lumen
endomembrane system
0
20
40
60
80
GO molecular function
unfolded protein binding
0
10
20
30
40
GO biological process
protein folding
0
10
20
30
40
50
Percentage

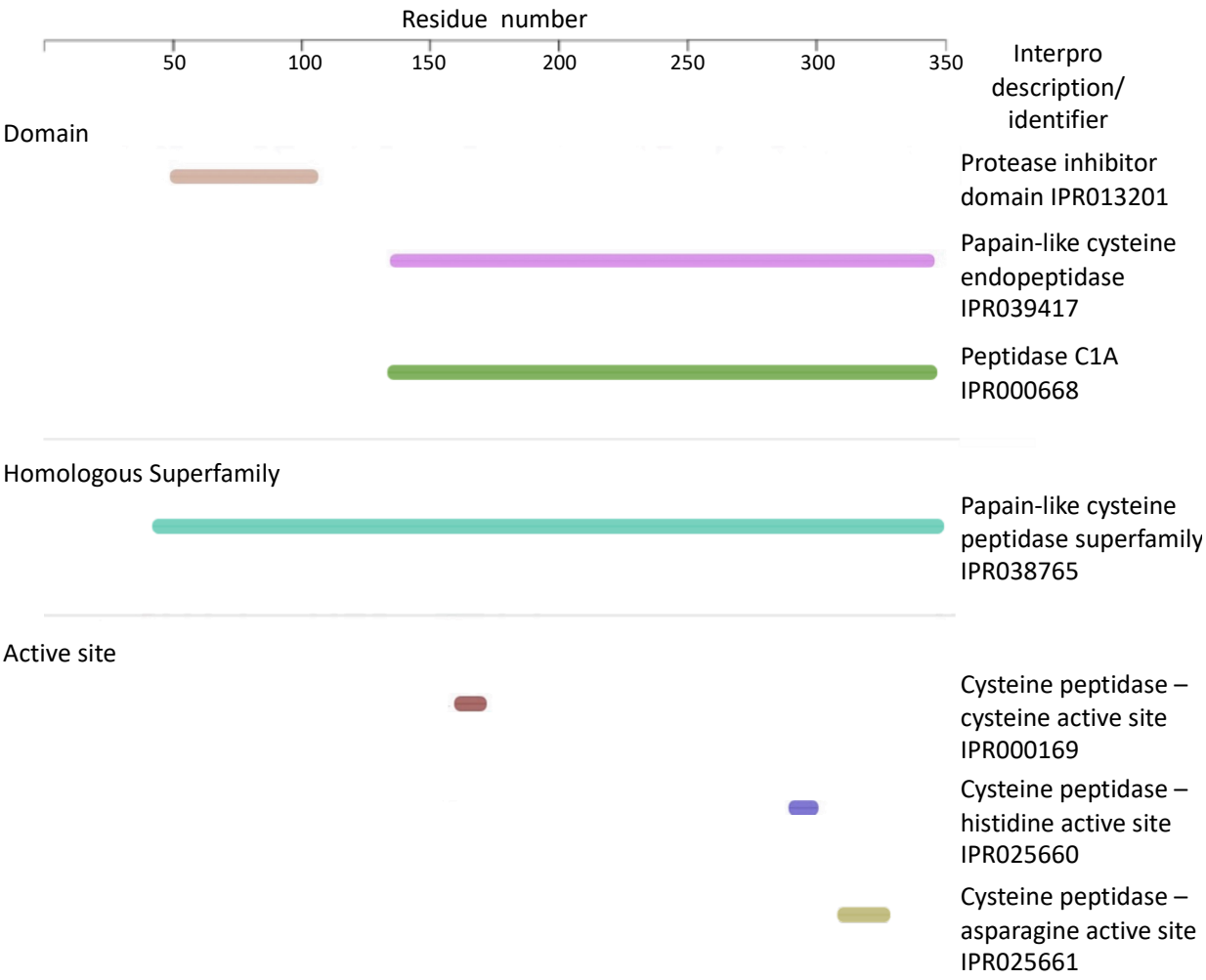
Figure 5. PAP interaction network arranged by functional cluster. The PAP interactome was visualized using Cytoscape v3.3.0 with coIP-MS identified protein interactors and indirect protein interactors as predicted using the STRING database. Coloured nodes represent PAP protein interactors identified by coIP-MS where yellow, green and blue nodes are those interactors belonging to **(A)** cluster 1, **(B)** cluster 2, or **(C)** cluster 3, respectively. Grey nodes represent indirect protein interactors identified in the STRING database (for full protein names see Table S2; Appendix B) and grey lines represent predicted functional interactions between proteins with interaction scores > 0.7. Enrichment analysis of GO terms associated with each cluster was performed using Fisher's Exact Test. Enriched GO terms for each cluster (p-value < 0.01) are shown by percent of proteins enriched for this term and organized by GO category: biological function, molecular function, and cellular component.

immunity. PHYAM_026164 is hereafter called pokeweed cysteine protease I (paCPI), where pa = *Phytolacca americana*, followed by its theorized protein class, cysteine protease (CP), inferred from homology with Xylem Cysteine Protease 1 in *Arabidopsis* (AtXCP1).

When the paCP1 amino acid sequence was queried in a blastP search against the Embryophyta clade in the SWISS-PROT database to identify annotated homologous plant proteins, this yielded a list of 50 homologous plant proteins with E-values < $1x10^4$, in addition to AtXCP1 (Table S3; Appendix B). Upon manual examination, it was determined that 100% of these returned homologous proteins belong to the cysteine protease family, providing additional evidence that paCP1 is a cysteine protease.

An Interproscan search was also performed using the paCP1 amino acid sequence to identify conserved protein regions, for further support of the inferred identity of paCP1 as a cysteine protease (Blum et al. 2021). The pokeweed protein was identified as belonging to the papain-like cysteine peptidase superfamily (Figure 6A); peptidases are alternately known as proteases, proteinases, or proteolytic enzymes. Members of this family have been categorized as subfamily C1A, in accordance with the protease-specific database MEROPS (Rawlings and Bateman 2021). The identified active site comprises a cysteine-histidine (159/290); this catalytic dyad is characteristic of cysteine proteases (Cstorer and Ménard 1994). A conserved asparagine is also identified; while the asparagine is also characteristic of C1A proteases, it has been reported as non-essential for catalytic action (Ménard et al. 1990). Additionally, amino acids 135-347 were identified as a peptidase C1A domain, and amino acids 48-104 were identified as a protease inhibitor domain (Figure 6A). Inhibitor domains are conserved features of cysteine proteases, which are synthesized as inactive precursors called zymogens. The inhibitor prodomain blocks substrate access to the active site, preventing inappropriate proteolytic activity and unwanted protein degradation, until the zymogen is activated upon cleavage of the prodomain (Coulombe et al. 1996; Pandey et al. 2009). Taken together, these results demonstrate that the newly identified pokeweed protein paCP1 is a probable cysteine protease.

A


Residue number
50
100
150
200
250
300
350
Interpro description/ identifier
Domain
Protease inhibitor domain IPR013201
Papain-like cysteine endopeptidase IPR039417
Peptidase C1A IPR000668
Homologous Superfamily
Papain-like cysteine peptidase superfamily IPR038765
Active site
Cysteine peptidase – cysteine active site IPR000169
Cysteine peptidase – histidine active site IPR025660
Cysteine peptidase – asparagine active site IPR025661

B

```
Length: 364
Identity:     188/364 (51.6%)
Similarity:   251/364 (69.0%)
Gaps:          23/364 ( 6.3%)
Score: 994.5
```

```
atXCP1            1 MAFSAPSLSKFSLLVAISASALLCC----AFARDFSIVGYTPEHLTNTDK     46
                      ::..||....|.:.|:.|    |||:|||:|||:||.|.:.::
paCP1             1 -------MASLSLFKIFSFTLLVFCVVYPAFAQDFSVVGYSPEDLNSDER     43

atXCP1           47 LLELFESWMSEHSKAYKSVEEKVHRFEVFRENLMHIDQRNNEINSYWLGL     96
                    |..||.||||:|.|:|.|:||::.|||||::||.:|.:||.:|::|||.|
paCP1            44 LYNLFVSWMSKHQKSYGSLEERLQRFEVFKDNLKYIHERNQKISNYWLSL     93

atXCP1           97 NEFADLTHEEFKGRYLG---LAKPQFSRKRQPSANFRYRDITDLPKSVDW    143
                    |.||||||||:|.:||.   :..|:.  :...|:||.:.|:.:||..|||
paCP1            94 NVFADLTHEEYKSKYLSSSIIVPPEI--REVSSSNFTHVDVKNLPGYVDW    141

atXCP1          144 RKKGAVAPVKDQGQCGSCWAFSTVAAVEGINQITTGNLSSLSEQELIDCD    193
                    |...||..||:||.||||||||..|||||||:|.||||.|||.|:|:||
paCP1           142 RNSYAVTTVKNQGTCGSCWAFSAAAAVEGINRIVTGNLVSLSPQQLLDC-    190

atXCP1          194 TTFNSGCNGGLMDYAFQYIISTGGLHKEDDYPYLMEEGICQEQKEDVERV    243
                     |....|.||.|..||:|||:.|||||.:||||:.::..||:...:|  |
paCP1           191 -TGRGNCKGGWMHSAFEYIINNGGLHKWNDYPYVGQQQYCQQAGGEV--V    237

atXCP1          244 TISGYEDVPENDDESLVKALAHQPVSVAIEASGRDFQFYKGGVFNGKCGT    293
                    ||.||:.||:| :.||::|:|.|||||.:|...:|||||.||||.|.|||
paCP1           238 TIQGYQHVPQN-ERSLLQAIAQQPVSVGLEGYSQDFQFYGGGVFRGNCGT    286

atXCP1          294 DLDHGVAAVGYG--SSKGSDYVIVKNSWGPRWGEKGFIRMKRNTGKPEGL    341
                    :|:|.|.|||||  |....||:::||||..||||:|::|:..|||:|||:
paCP1           287 NLNHAVLAVGYGRTSDTNEDYILLKNSWSERWGERGYMRILANTGRPEGM    336

atXCP1          342 CGINKMASYPTKTK    355
                    ||:|..|.||.|..
paCP1           337 CGVNMAAVYPIKKY    350
```

C

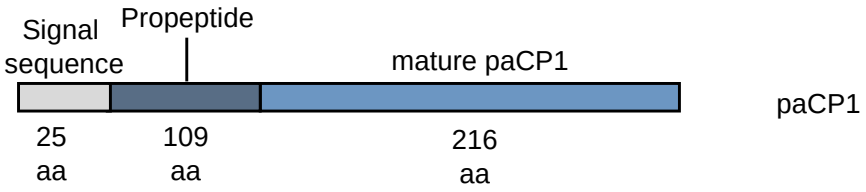


Figure 6. Protein sequence characterization of pokeweed cysteine protease paCP1. **(A)** Amino acid sequence for paCP1 was searched against the InterPro database to identify protein domains, superfamily, and active sites. Identified features are represented by coloured bars to indicate which amino acids they span, with corresponding Interproscan description and identifier on right hand side. **(B)** MAKER-annotated protein sequence for paCP1 was globally aligned to the plant protein sequence with highest homology (AtXCP1) from the SwissProt database using EMBOSS Needle. (|) indicates residues are identical, (:) indicates conservative mutation, (.) indicates semi-conservative mutation, and () indicates non-conservative mutation. Potential signal peptide cleavage site as predicted by Signal-P 5.0 for both sequences is indicated by a red line with a star. A green line and star marks the known propeptide cleavage site of AtXCP1. **(C)** Schematic of putative protein sequence organization for paCP1.

Additional characterization of the paCP1 sequence was performed by globally aligning the paCP1 amino acid sequence to the AtXCP1 sequence using EMBOSS needle (Madeira et al. 2019). Alignment results demonstrated 69% similarity between the two protein sequences, with 251/364 amino acids conserved (Figure 6B). To confirm that paCP1 and AtXCP1 enter the secretory pathway, amino acid sequences were inputted into Signal-P 5.0 to predict presence and cleavage site of a signal sequence (Almagro Armenteros et al. 2019). Results show that both AtXCP1 and paCP1 encode Sec/SPI signal peptides, with likelihoods of 0.8944 and 0.9842, respectively. The predicted cleavage site of paCP1 is between A25 and Q26 (Figure 6B), which aligns with predicted AtXCP1 cleavage site A28-R29. The surrounding amino acid sequences in this protein region are strongly conserved (AFA-QD and AFA-RD; paCP1 and AtXCP1, respectively). Additionally, the AtXCP1 cleavage site between the prodomain and mature AtXCP1 protein is marked (Zhao et al. 2000) and as adjacent amino acids to this site are identical for paCP1, prodomain cleavage for paCP1 is predicted to be after residue 135. The predicted location of the signal peptide, propeptide inhibitor domain, and mature protein locations inferred in Fig. 6B also correspond with Interproscan results. Therefore, paCP1 amino acid sequence organization is predicted as shown in Figure 6C.

3.4 Validation of PAP-paCP1 interaction

The interaction between PAP and paCP1 was then validated by reverse co-IP, where paCP1 was cloned and expressed fused to a 6x His tag and used to pull down PAP. Immunoprecipitation of His-paCP1 using an anti-His antibody, followed by incubation with pokeweed lysate resulted in a ~29 kDa band, visible only in immunoprecipitation from pokeweed lysate and not that from tobacco lysate (Figure 7, lane 5). This band corresponds in size to that of the purified PAP standard used as a positive control, and to the similarly sized band seen in the pokeweed lysate input. A weak band of the same size is also seen in the tobacco lysate, which does not express PAP; however, this weak band is due to overloading of pokeweed lysate and consequent slight leakage into the adjacent lane. The presence of the ~29 kDa band in the co-immunoprecipitation from pokeweed lysate corroborates a PAP-paCP1 interaction and provides support for the accuracy of the coIP-MS method.

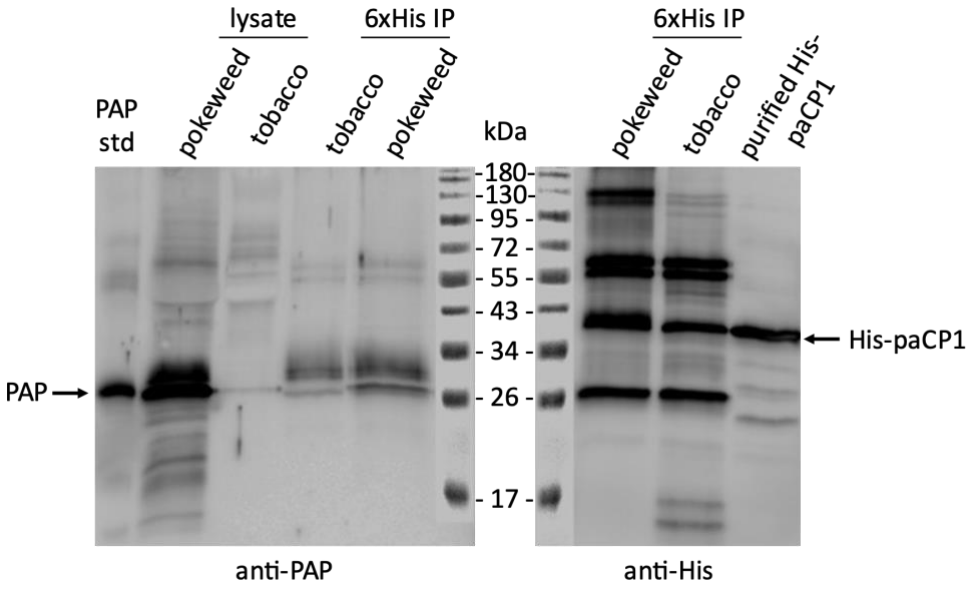


Figure 7. Validation of paCP1-PAP interaction by co-immunoprecipitation (co-IP). His-tagged paCP1 was purified from *E. coli* (Rosetta) and immobilized to His antibody-coupled magnetic beads, followed by incubation with protein extracts from *Phytolacca americana* or *Nicotiana tabacum* leaf tissue (6xHis IP pokeweed and 6xHis IP tobacco, respectively). Pokeweed and tobacco lysates were run to demonstrate IP input, and purified PAP and His-tagged paCP1 were run as protein standards (PAP std and purified His-paCP1, respectively). Co-IP of PAP was detected by Western blot using an anti-PAP polyclonal antibody. Expected protein sizes are indicated by arrows. Blot is representative of two independent experiments.

3.5 PAP-paCP1 co-expression analysis

Proteases from multiple families have been implicated in almost every stage of plant immunity, with cysteine proteases known specifically to play important roles in programmed cell death (PCD) and in response to biotic and abiotic stresses (Zamyatnin 2015; Balakireva and Zamyatnin 2018). It was therefore hypothesized that paCP1 and PAP may be co-involved in a plant defense response and that these proteins would be co-expressed upon defense response activation. Plants have two main branches of induced pathogen defense, known as systemic-acquired resistance (SAR) and induced systemic resistance, which are activated by phytohormones salicylic acid (SA) and jasmonic acid (JA), respectively (Nomoto et al. 2021). RNA-seq was previously performed by our lab on pokeweed leaves sprayed with either JA, SA, or with ethanol as a control, to examine global gene expression in response to stress treatments (Neller et al. 2019). These data were utilized for the current work to investigate whether paCP1 and PAP are co-regulated in either SA or JA-dependent responses. Transcript fold change across four replicates was calculated for paCP1, PAP-1, PAP-2, K-PAP and novel PAP, where differential gene expression was considered significant for FDR < 0.05 (Table 5). Three of the PAP isoforms (PAP-1, PAP-2, and novel PAP) and paCP1 were identified as JA-responsive genes; K-PAP was the only gene that demonstrated a lack of JA-responsiveness. In response to SA treatment, PAP-1 and PAP-2 expression was upregulated in contrast to paCP1, K-PAP, and novel PAP which showed no significant change in expression. The lack of SA-responsive paCP1 upregulation suggests that the paCP1-PAP interaction is unlikely to play a significant role in the initiation of the SAR response. However, the JA-responsive upregulation of paCP1 and several PAP isoforms implicate both proteins as active in JA-mediated cellular processes. While a paCP1-PAP interaction may also be part of the baseline proteomic cellular landscape, their common differential co-expression indicates their interaction may contribute to a shared role in the JA-dependent defense response.

To gain insight into the JA co-regulation of paCP1 and the PAP isoforms, the promoters for each gene were annotated for JA-responsive cis-regulatory elements (CREs). Since MYC2 is known as the master regulatory TF of the JA pathway, the promoters were annotated for CREs known to bind MYC2 and its homologues MYC3 and MYC4. MYC2 strongly activates JA-responsive

Table 5. PAP isoform and paCP1 gene expression in response to jasmonic acid (JA) and salicylic acid (SA) treatment. Expression changes of PAP isoforms identified in PAP co-IP in response to pokeweed stress treatments. Green indicates significant up-regulation. Average abundance of each transcript is shown in transcripts per million (TPM).

Gene ID	Protein name	log_2 fold change JA	log_2 fold change SA	Average Expression (TPM)
PHYAM_026164	paCP1	2.64	-0.18	3204
PHYAM_020596	PAP-1	4.50	2.00	9911
PHYAM_028184	PAP-2	3.24	0.75	1834
PHYAM_012451	KPAP	0.21	0.19	1092
PHYAM_010465	Novel PAP	12.38	-	57

genes through a transcriptional cascade; therefore, TFs shown to be direct MYC2 or MYC3 targets and rapidly upregulated by JA treatment (DREB2B, RVE2, ZAT18, TCP23), as well as TFs with known roles in JA signalling (NAC3, WRKY51, and ZAT10) were also considered when searching the promoters for JA-responsive TF binding motifs (Zander et al. 2020). Analysis of the 1.3 kb sequence upstream of each +1 TSS as determined by the MAKER-annotated pokeweed genome revealed MYC2, MYC3, and MYC4 binding sites in paCP1 and all PAP isoform promoters (Table 6). The PAP-1 and novel PAP promoters contained three unique MYC-binding motifs and a single MYC2-binding motif, respectively, while paCP1 and PAP-2 promoters contained a single core MYC-binding motif able to bind all three MYC homologues using varying flanking sequences. The K-PAP promoter was found to encode two potential MYC binding sequences, with one potentially able to bind MYC2, MYC3, and MYC4 through the flanking sequence. However, as K-PAP was not shown to be significantly upregulated in response to JA (Table 5), some or all these motifs may not be active, or may bind MYC transcription factors regulated by stimuli other than JA (Chen et al. 2019). PAP-1, PAP-2, K-PAP, novel PAP and paCP1 promoters contained 10, 14, 12, 20, and 11 binding motifs for the MYC-activated TF set, respectively, indicating that differential JA-responsive regulation of these genes may also occur through MYC-mediated transcriptional reprogramming. In addition to JA-responsive upregulation of paCP1 and three out of four of the examined PAP isoforms, the presence of JA-responsive CREs throughout the paCP1 and PAP promoters supports their co-expression in a JA-mediated plant response.

To further explore whether paCP1 and PAP isoforms were co-expressed, their 1.3 kb promoter sequences were also analysed using PlantPan 3.0 (Chow et al. 2019) to identify any CREs common to paCP1 and one or more PAP promoters, beyond the presence of JA responsive CREs. Twenty CREs (51%) were identified in all five promoters, with a further six CREs common to only paCP1 and PAP-1, and three common to paCP1 and each of the PAP-2, K-PAP, and novel PAP promoters (Table 7). The majority of CREs identified (74%) were associated with diverse biotic and abiotic stresses, including ABRE/MYB-binding motifs (abscisic acid (ABA)- and dehydration-response), WRKY-binding motifs (SA- and pathogen-response), gibberellin (GA)-responsive motifs, sulfur- and phosphate-responsive motifs, and the UPRMOTIFIIAT motif (unfolded protein response). Multiple light-responsive motifs were also

Table 6. Putative jasmonic acid-responsive cis-regulatory elements (CREs) identified in paCP1 and PAP promoters. Core binding motifs are represented by uppercase letters and flanking sequences are represented by lowercase letters.

Promoter	Transcription factor	Binding motif	Position	Strand
paCP1	MYC2	gtCACGTgat	663	+
		gtcACGTGat	663	-
	MYC3	ttgtCACGTg	661	-
		cACGTGataa	665	+
	MYC4	ttgtCACGTg	663	+
		gtcACGTGat	663	-
	DREB2B	GTCGGc	498	-
		GTCGGg	607	-
		aCCGAC	1048	+
	RVE2	caATATCtt	246	+
	ZAT18	gAGTGTg	586	-
	TCP23	gggCCCACca	406	+
		tggGGCCCac	404	-
	NAC3	ggGTCAAat	394	+
	WRKY51	gggGTCAAat	393	+
	ZAT10	cAGTGAtggtg	540	-
		gAGTGTg	586	-
PAP1	MYC2	CACTTg	370	+
		cAAGTG	512	-
	MYC4	aACGTGc	1075	+
	DREB2B	GTCGGa	822	-
	RVE2	caGATATga	178	-
		taATATCca	749	+
	ZAT18	aagcAGTGTgttt	1235	-
	NAC3	gaGTCAAgt	508	+
		taGTCAAcc	691	+
	WRKY51	gtaGTCAAcc	690	+
	ZAT10	gAGTGTt	132	-
		cACACTa	212	+
		cAGTGTg	1238	-
PAP2	MYC2	gtCACGTgtt	1195	+
		gtcACGTGtt	1195	-
	MYC3	acgtCACGTg	1193	-
		cACGTGttcc	1197	+
	MYC4	gtCACGTgtt	1195	+
		gtcACGTGtt	1195	-
	DREB2B	cCCGAC	885	+
	RVE2	aaGATATtt	378	-
		caGATATta	638	-
		tgGATATta	896	-

Table 6 (continued).

Promoter	Transcription factor	Binding motif	Position	Strand
PAP2 (cont.)	TCP23	gacCCCAC	1167	+
	NAC3	taGTCAAtt	18	+
		atTTGACct	273	-
		taGTCAAtt	356	+
		caGTCAAtg	589	+
		ttTTGACcc	1224	-
	WRKY51	atTTGACcta	273	-
		ttTTGACccg	1224	-
	ZAT10	caattTCACTa	360	+
		attTCACT	362	+
KPAP	MYC2	caCACGTgct	37	+
		cacACGTGct	37	-
		CACTTg	135	+
	MYC3	tgcaCACGTg	35	-
		cACGTGcttt	39	+
	MYC4	caCACGTgct	37	+
		cacACGTGct	37	-
	DREB2B	GTCGGt	718	-
		cCCGAC	765	+
		GTCGGa	1194	-
	NAC3	ccTTGACta	145	-
		gaGTCAAcc	825	+
		acTTGACct	837	-
		tgTTGACcc	845	-
		ggTTGACcc	1118	-
	WRKY51	cgaGTCAAcc	824	+
		acTTGACctg	837	-
		tgTTGACcca	845	-
		ggTTGACcca	1118	-
Novel PAP	MYC2	cATGTG	637	-
	DREB2B	GTCGGg	309	-
		GTCGGg	318	-
		GTCGGa	1038	-
		GTCGGg	1206	-
		GTCGGg	1230	-
	RVE2	aaGATATta	173	-
		aaGATATga	931	-

Table 6 (continued).

Promoter	Transcription factor	Binding motif	Position	Strand
Novel PAP (cont.)	NAC3	atTTGACcg	152	-
		agGTCAAcc	421	+
		CATGTg	637	+
		ggGTCAAtg	1083	+
		agGTCAAat	1137	+
		ggGTCAAct	1233	+
	WRKY51	atTTGACcga	152	-
		cagGTCAAcc	420	+
		tggGTCAAtg	1082	+
		tagGTCAAat	1136	+
		cggGTCAAct	1232	+
	ZAT10	aACACTa	395	+
		aAGTGTc	990	-

Table 7. Additional putative cis-regulatory elements (CREs) common to PAP and paCP1 promoters.

Cis-regulatory Element	Element sequence	Annotation	Copy number in promoter					References
			paCP1	PAP-1	PAP-2	KPAP	Novel PAP	
ABREAT-CONSENSUS	YACGTGGC	ABF-binding element; ABA responsive; stress-responsive	1	-	1	-	-	Choi et al. 2000
ABRELATERD1	ACGTG	ABA responsive; dehydration-responsive	20	25	24	46	16	Simpson et al. 2003 Nakashima et al. 2006
ACGTABRE-MOTIFA2OSEM	ACGTGKC	ABA-responsive; osmotic stress-responsive	1	2	2	1	-	Hattori et al. 2002 Kim et al. 2011
ANAERO1-CONSENSUS	AAACAAA	oxygen-responsive; involved in fermentative pathway	2	4	3	3	3	Mohanty et al. 2005
ARR1AT	NGATT	ARR1-binding element; cytokinin-responsive	11	12	21	9	14	Sakai et al. 2000
CARGCW8GAT	CWWWWWWWWG	AGL1-binding motif; found in homeotic genes	2	-	-	2	2	Tang et al. 2003
CDA1ATCAB2	CAAAACGC	light-responsive; involved in photomorphogenesis	1	-	-	-	2	Maxwell et al. 2003 Lau et al. 2011
E2FANTRNR	TTTCCCGC	E2F-binding element; cell cycle-responsive	1	-	-	2	2	Chaboute et al. 2000 de Jagr et al. 2001
E2FCONSENSUS	WTTSSCSS	E2F-DP-binding element; cell cycle-responsive	4	3	1	1	4	Vandepoele et al. 2005
GADOWNAT	ACGTGTC	GA-responsive; ABA-responsive; dehydration-responsive	1	1	3	1	-	Nakashima et al. 2006 Ogawa et al. 2003
GAREAT	TAACAAR	GA-responsive	2	3	2	5	1	Ogawa et al. 2004
GBOXLERBCS	MCACGTGGC	GBF-binding motif; light-responsive; ABA-responsive	1	-	1	-	-	Giuliano et al. 1988 Vasil et al. 1995
GT1CONSENSUS	GRWAAW	GT-1 binding motif; light-responsive; involved in defense response	9	17	17	11	17	Buchel et al. 1999 Gourrierec et al. 1999
L1BOXATPDF1	TAAATGYA	homeodomain protein-binding element; GA-responsive	1	1	4	4	-	Abe et al. 2003 Rombola-Caldenty et al. 2014
LEAFYATAG	CCAATGT	LEAFY-binding element; found in homeotic genes	4	1	1	1	1	Kamiya et al. 2003
LS5ATPR1	TCTACGTCAC	TGA2-binding site; SA-responsive; involved in defense response	2	3	-	-	-	Despres et al. 2000 Niggeweg et al. 2000

Table 7 (continued).

Cis-regulatory Element	Element sequence	Annotation	Copy number in promoter					References
			paCP1	PAP-1	PAP-2	KPAP	Novel PAP	
LS7ATPR1	ACGTCATAGA	TGA2-binding element; SA-responsive; defense-responsive	8	1	4	7	3	Johnson et al. 2003 Despres et al. 2000
MYB1AT	WAACCA	MYB-binding motif; ABA-responsive; dehydration-responsive	4	9	2	4	3	Abe et al. 2003
MYB2-CONSENSUSAT	YAACKG	MYB-binding site; ABA-responsive; dehydration-responsive	2	-	-	-	3	Abe et al. 2003
MYBCORE-ATCYCB1	AACGG	MYB-binding element; ABA-responsive; dehydration/salt-responsive	22	24	22	33	19	Planchais et al. 2002 Abe et al. 2003
MYBPLANT	MACCWAMC	MYB-binding element; involved in growth and metabolism regulation	1	1	1	1	1	Tamagnone et al. 1998
P1BS	GNATATNC	Phosphate-responsive; involved in phosphate starvation response	6	3	7	6	1	Rubio et al. 2001 Schuenmann et al. 2004
PALINDROMIC-CBOXGM	TGACGTCA	bZIP-binding site element	5	3	3	7	7	Cheong et al. 1998
PIATGAPB	GTGATCAC	light-responsive; found in GAPB gene	1	-	-	-	1	Chan et al. 2001
PROXBBNNAPA	CAAACACC	ABI3-binding motif; ABA-responsive	2	-	1	-	-	Ezcurra et al. 1999 Ezcurra et al. 2000
RHERPATEXPA7	KCACGW	Found in root-hair specific genes	2	1	2	4	2	Kim et al. 2006
SBOXATRBCS	CACCTCCA	ABA-responsive; sugar responsive; light-responsive	1	1	-	-	-	Acevedo-Hernandez et al. 2005
SITEIIATCYTC	TGGGCY	TCP-binding motif; found in cytochrome c and nuclear oxidative phosphorylation machinery genes	4	3	-	-	1	Welchen et al. 2005 Welchen et al. 2006
SORLIP1AT	GCCAC	light-responsive; phytochrome A-induced element	38	33	10	15	9	Hudson and Quail 2003
SORLIP2AT	GGGCC	light-responsive; phytochrome A-induced element	22	31	10	13	9	Hudson and Quail 2003
SORLIP5AT	GAGTGAG	light-responsive; found in phytochrome A-regulated genes	3	-	-	1	-	Hudson et al. 2003 Jiao et al. 2005
SURECORE-ATSULTR11	GAGAC	Sulfur-responsive; auxin-responsive	37	45	22	21	20	Maruyama-Nakashita et al. 2005
TBOXATGAPB	ACTTTG	light-responsive; found in GAPB gene	1	-	2	1	-	Chan et al. 2001

Table 7 (continued).

Cis-regulatory Element	Element sequence	Annotation	Copy number in promoter					References
			paCP1	PAP-1	PAP-2	KPAP	Novel PAP	
UP1ATMSD	GGCCCAWWW	Involved in growth regulation	2	1	-	-	-	Tatematsu et al. 2005
UP2ATMSD	AAACCCTA	TCP-binding element; wounding-responsive; involved in growth regulation	2	-	3	2	3	Tatematsu et al. 2005
UPRMOTIFIIAT	CCNNNNNNNNN NNNCCACG	unfolded protein response element; stress-responsive	1	1	-	-	-	Oh et al. 2003
WBOXATNPR1	TTGAC	WRKY-binding element; SA-responsive; pathogen-induced	41	31	42	62	58	Maleck et al. 2000 Yu et al. 2001
WBBOXPCWRKY1	TTTGACY	WRKY-binding element; induced by fungal elicitors	4	2	9	2	9	Rushton et al. 1996
XYLAT	ACAAAGAA	involved in xylem development	2	1	2	1	1	Ko et al. 2006

present (CDA1ATCAB2, GBOXLERBCS, GT1CONSENSUS, PIATGAPB, TBOXATGAPB), including several phytochrome A-mediated binding motifs (SORLIP1AT, SORLIP2AT, SORLIP5AT). Phytochrome A plays a critical role in gene upregulation in response to not only light signals, but also in defense and stress responses (Franklin and Quail 2010; Wang et al. 2016). The presence of these stress responsive CREs common to paCP1 and PAP promoters supports the potential co-expression of these proteins in defense responses in pokeweed. Additionally, the presence of ten CREs involved in growth regulation suggests a mechanism whereby these genes may be differentially regulated under normal versus stress conditions, reflecting the critical balance in plants whereby defense activation is necessarily a trade-off with plant growth (Major et al. 2020).

4. DISCUSSION

This work sought to chart the PAP interactome to gain insight into PAP's role in its endogenous environment. PAP and other RIPs have been heavily connected with conferral of pathogen resistance when applied to or expressed in heterologous systems (Zhu et al. 2018). Mechanisms proposed for this action include inhibition of protein synthesis, or participation in defense pathways; however, neither have been definitively linked to the pathogen resistance seen when expressed in transgenic plants. The pokeweed genome has recently been sequenced (Neller et al. 2019), allowing for use of the coIP-MS method to map out the PAP-protein interactome. Identification of PAP-protein interactions from pokeweed tissue would thus allow for inference of PAP's function in its endogenous environment. Since PAP is known to localize primarily to the apoplast (Ready et al. 1986), it was hypothesized that a coIP-MS-generated PAP interactome would include proteins that colocalize to the apoplast and participate in plant defense. This hypothesis was partially supported by the coIP-MS results, which included a cysteine protease-PAP interaction that was further supported by reverse coIP and co-expression data. However, interestingly, the interactome also comprised proteins with diverse functions and localizations, hinting at a more complex role for PAP in its endogenous environment than defense alone.

4.1 Proteins interacting with PAP are involved in diverse biological functions

Previous to this work, only three PAP-protein interactions have been reported in the literature: ribosomal protein L3 from yeast, eukaryotic translation initiation factor eIF4G from wheat germ, and AtBI-1, an anti-apoptotic protein expressed in *Arabidopsis* (Hudak et al. 1999; Wang and Hudak 2006; Çakır et al. 2015). Pokeweed homologs for these proteins were not among the mass spectrometry-identified proteins returned in this study. This discrepancy can be interpreted in view of experimental differences as well as technical biases associated with each. Identification of the aforementioned interactors was shown by co-IP either *in vitro* using purified proteins or yeast overexpression lines, both of which reflect much higher POI concentrations than those expected in a native expression context such as in wild-type plant tissues. Co-immunoprecipitation as a proteomic screening technique is known to be biased towards interaction partners with high abundances (Smaczniak et al. 2012). While maintenance of physiological protein levels is advantageous for identification of biologically relevant

interactions, a resultant disadvantage is the chance for missed lower abundance interactions. Identification of interactors is also linked to binding affinity, as interaction strength will affect its stability in response to buffer components in protein extraction and washing steps, which can disrupt unstable or transient bonds (Conlon et al. 2012). Furthermore, the interactome depends on cellular activity changes in response to internal and external shifts, thus the identified interactome at a point in time depends on the particular state of the cells.

Biases inherent in screening methods are illustrated by comparison of protein-protein interaction networks obtained for enzymes CYP83A1 and CYP83B1 in *Arabidopsis* by both untargeted yeast 2-hybrid screens and coIP-MS (Nintemann et al. 2017). Comparison of the generated interaction datasets revealed little overlap in the detected proteins: only one interactor was common to both datasets, out of 131 identified proteins. And, although there were 16 previously reported interactions for the bait enzymes, only one was identified by either screening assay. Therefore, the sensitivity, experimental design, and limitations and biases of different methods make it hard to draw comparisons between datasets and explain the lack of the three previously reported interactions for PAP in this dataset. An interactomic data set arises from the results of hundreds of experiments (Alvarez-Ponce 2017). As this represents a single screen of the PAP-protein interactome, it provides a functional window on PAP-protein interactions but cannot be considered exhaustive.

The majority of protein interactors identified in cluster 1 demonstrate GO term enrichment associated with protein synthesis, and subcellular localization to the proximity of the ribosome. These results are consistent with what is characterized about PAP: it participates in negative regulation of translation through its N-glycosidase activity, directed towards the cytoplasmic ribosomal target. Enrichment for ribosomal and translation-related GO terms in an interaction cluster was therefore expected.

GO term enrichment of interaction clusters for the other two PAP protein interactors, carbonic anhydrase and paCP1, indicate biological process GO term enrichment in carbon utilization and the unfolded protein response, respectively. This suggests that PAP may participate in biological functions in addition to its proposed role in plant defense. Moonlighting proteins are single-chain

polypeptide proteins that fulfill two or more physiologically relevant and distinct functions, not occurring through alternative splicing, isoforms, or post-translational processing (Espinosa-Cantú et al. 2018). As moonlighting proteins are increasingly viewed as widely occurring (Krantz and Klipp 2020; Turek and Irving 2021), dual functions have also been described for several RIPs. High concentrations of certain RIPs in seeds and storage organs support the hypothesis that some RIPs play a role in energy storage. Trichosanthin, a type I RIP produced by the Chinese cucumber *Trichosanthes kirilowii*, is isolated from both the leaves and the roots of the plant. In fibrous roots, whose purpose is absorption of nutrients and water from the soil, trichosanthin comprises only 0.5% total soluble protein; upon secondary root growth leading to storage root formation, trichosanthin accumulates as the major soluble protein, representing over 25% total soluble protein in fully developed tuberous roots (Savary and Flores 1994). This situation is mirrored in ME1 and ME2, two RIPs isolated from the Andean root crop *Mirabilis expansa*, where they constitute approximately 20% total protein in the storage organs (Vivanco et al. 1999). Storage proteins with dual roles in defense are not uncommon. For example, sporamin, which accounts for 60-80% of total soluble protein in sweet potato tubers, is induced by wounding and herbivore attack, shares amino acid sequence identity with trypsin inhibitors, and possesses strong trypsin inhibitory action *in vitro* (Yeh et al. 1997; Rajendran et al. 2014). That RIPs may have evolved this dual function is therefore not unexpected for storage proteins.

Dual functions for some RIPs in abiotic stress tolerance is suggested by differential RIP expression in response to abiotic stressors (Stirpe et al. 1996; Jiang et al. 2008). For example, expression of OsRIP1, a type I RIP isolated from rice, is upregulated in response to drought, salt, cadmium, and cold stress (Wytynck et al. 2021). Maize RIP2 expression is both upregulated in response to ABA treatment, a phytohormone involved in abiotic stress response, and plays a role in biotic stress response, resulting in decreased biomass of armyworm larvae fed on RIP2-expressing leaves (Bass et al. 2004; Chuang et al. 2014). Overexpression of some RIPs has also been shown to increase plant resistance to abiotic stresses (Jiang et al. 2012; Wytynck et al. 2021).

Some RIPs have also been found to possess enzymatic activities in addition to N-glycosidase activity. A RIP purified from *Cucurbita moschata*, and be27, a type I RIP from *Beta vulgaris*,

have been found to possess superoxide dismutase activity, conferring a potential antioxidant function to these RIPs (Barbieri et al. 2006; Iglesias et al. 2015). Since the vast majority of RIP investigations focus on RIP N-glycosidase activity and RIP role in defense, little exists in the literature investigating additional physiological or cellular roles of these proteins; however, based on these examples, it is speculated that PAP may have novel moonlighting functions which are thus far unknown but suggested by the cellular functions of the identified PAP interactors.

Biological process GO terms associated with the paCP1 cluster indicate the potential for PAP to participate in the unfolded protein response (UPR). The A-chain of ricin, a type II RIP, has been shown to inhibit the UPR in yeast when localized to the ER (Parikh et al. 2008). Although the mechanism is unknown, it was shown to occur upstream of transcriptional activation of chaperones and other UPR components that direct misfolded proteins to the cytosol for degradation (Ron and Walter 2007; Parikh et al. 2008). As type II RIP A-chains are functionally equivalent to type I RIPs, it is reasoned that PAP could have similar involvement in the UPR. Type II RIPs ricin and Shiga toxin have also been shown to undergo retrograde transport from the ER back into the cytosol using the ER-associated degradation (ERAD) pathway (Simpson et al. 1999; Yu and Haslam 2005). This pathway targets misfolded proteins to retrotranslocate from the ER into the cytosol for proteasomal degradation, and has a direct role in triggering and maintaining the UPR if there is an imbalance between protein-folding demand and folding capacity that results in accumulation of misfolded proteins in the ER (Hwang and Qi 2018). The UPR functions to return homeostasis by increasing protein folding and degradation, as well as decreasing protein translation. Ricin retrotranslocation by ERAD occurs to the catalytic A-chain of ricin only after cleavage from the B-chain, or when it is expressed alone (Frigerio et al. 1998; Di Cola et al. 2001), again indicating the potential for ERAD-mediated translocation for a type I RIP. Correspondingly, indirect evidence for PAP retrotranslocation into the cytosol has been demonstrated in yeast (Parikh et al. 2005), signalling PAP may have interactions involved with ERAD and by implication the UPR. Since retrotranslocated PAP escapes proteasomal degradation in the cytosol (Parikh et al. 2005), therefore accessing ribosomes, PAP may play a role in the UPR by decreasing translation rates to restore homeostasis.

For the remaining cluster, however, as only 20% of secondary shell proteins are associated with each GO biological process term, this suggests that proteins in this cluster could also have other unannotated cellular functions associated with them. For carbonic anhydrase (CA), the identified PAP interactor in this cluster, the enriched GO terms "carbon utilization" and "reductive pentose-phosphate cycle" point to its potential role in photosynthesis, where CA is hypothesized to play a role in carbon dioxide assimilation through catalysis the interconversion of CO2 and HCO3 (DiMario et al. 2017). Recent evidence however using CA tobacco knockout lines contradicts this as mutants showed normal photosystem II activity and CO_2 assimilation (Hines et al. 2021). CAs have also been implicated in crucial roles in stomatal closure in leaves, and in defense responses induced upon pathogen attack (Hu et al. 2015). Since stomata are the main entry point into the leaves, stomatal closure is related to plant defense, and evidence indicates it triggers further defense strategies to limit pathogen spread and aid in stress recovery (Gahir et al. 2021). CAs have also been identified as salicylic acid-binding proteins which play a role in the anti-viral response through antioxidant activity, and as a molecular scaffold facilitating interaction between other proteins in the defense response (Slaymaker et al. 2002; Wang et al. 2009; Medina-Puche et al. 2017). *Arabidopsis* lines resistant to the insect herbivore *Plutella xylostella* had higher than 2-fold expression of CA transcript levels (Collins et al. 2010). This highlights the strong possibility that some of the PAP-protein interactions identified here may function in unannotated cellular roles, with secondary linkages to plant defense.

4.2 Proteins interacting with PAP localize to diverse subcellular compartments

Proteins found to be interacting with PAP were inferred by homology to localize to multiple cellular compartments: three proteins have probable chloroplast-localization (FBA, CA, GAPB), two proteins cytoplasmic (EF-1-alpha, RPS26A), and one protein (paCP1) is likely secreted. Since PAP-1 has been shown by immunohistochemistry to localize primarily to the extracellular space (Ready et al. 1986), identification of an extracellular interactor was predicted. The identification of cytoplasmic-localized protein interactors is also consistent with known information about PAP-1; the immunohistochemistry data also reported probable subpopulations of PAP in the cytosol (Ready et al. 1986). Evidence for PAP cytoplasmic populations have also been reported using transgenic PAP expression in yeast, where PAP has been shown to

retrotranslocate into the cytosol (Parikh et al. 2005; Baykal and Tumer 2007), as well as in the cytosol of pokeweed protoplasts (Tourlakis et al. 2010). Investigation of PAP's ribosomal association has previously revealed two cytoplasm-localized protein interactors, ribosomal protein L3 and elongation factor eIF4G (Hudak et al. 1999; Wang and Hudak 2006), thus the identification of further interactors in the cytoplasm was also considered probable.

However, the identification of three proteins which likely localize to the chloroplast is surprising, as PAP involvement with the chloroplast is heretofore undocumented in the literature. Several possibilities arise for these interactions: (i) PAP may interact with these proteins before they localize to the chloroplast (ie. in the endoplasmic reticulum or the cytosol), (ii) protein interactors may be located in the outer chloroplast envelope membrane and interact with soluble PAP present in the cytosol, (iii) either PAP or the protein interactors are relocalized upon an external stimulus, or (iv) cellular localization of the PAP interactors differs from than the annotated chloroplast localization.

While chloroplasts synthesize some of their own proteins, these represent a small minority, comprising the photosynthetic machinery and factors involved with its assembly (Zybailov et al. 2008). The remaining ~95% of proteins found in the chloroplast are encoded by nuclear genes, translated in the cytoplasm and either imported into the chloroplast post-translationally from the cytosol directed by N-terminal targeting peptides or, in a small minority of cases, co-translationally via the secretory system (Villarejo et al. 2005; Nanjo et al. 2006; Baslam et al. 2016). Post-translationally imported proteins would therefore have the opportunity to interact with PAP in the cytosol before import, while those that are trafficked via a co-translational mechanism could interact in the endoplasmic reticulum. Only 0.6% of chloroplast-localized proteins are predicted to have an N-terminal signal sequence for insertion into the ER (Zybailov et al. 2008); however, a carbonic anhydrase in *Arabidopsis* was shown to possess an N-terminal signal peptide, enter the ER, and following cleavage of the signal sequence, reach the chloroplast through an unknown mechanism (Villarejo et al. 2005).

An important consideration is that protein localization annotation taken from the Uniprot database may not be correct, as these annotations are frequently predicted. Even when

subcellular localization is indicated by experimental evidence, the annotation is based on a consensus reflecting sometimes conflicting results. This can be due to protein subpopulations localizing to multiple compartments, or differential localization and relocalization upon certain cellular events or stimuli. Many proteins lack proper annotation, even in model species such as *Arabidopsis*, where still 32.4% of genes remain unannotated (Berardini et al. 2015; Depuydt and Vandepoele 2021). Therefore, interpretation of the inferred subcellular localization can be bolstered with additional manual annotation from the literature.

The homologous carbonic anhydrase in spinach belongs to the beta-class carbonic anhydrases (βCA), which have differing protein sequence and structure from alpha and gamma CAs, although all are zinc-containing metalloenzymes that catalyze the interconversion of CO2 and HCO3-. Plants have been found to encode between four and seven beta CAs (DiMario et al. 2017). Classically, CAs have been thought to localize to the chloroplast as research on these enzymes has focused on linking them to photosynthesis. However, subsequent studies using fluorescent protein fusion constructs demonstrated that two out of six βCA isoforms in *Arabidopsis* are in fact cytosolic (Fabre et al. 2007; DiMario et al. 2017), while a third isoform localizes to both the chloroplast and to the plasma membrane, depending on length of mRNA encoded from alternate transcription start sites (Hu et al. 2010). The latter case is also predicted for two more of the six isoforms, with RNA-seq data indicating that alternative splicing may produce multiple mRNA forms, resulting in differing N-termini and thus different cellular destinations (Aubry et al. 2014; Oh et al. 2014). Therefore, the assumption that a PAP-CA interaction occurs in chloroplast may be misleading; it seems more likely that PAP would interact with CA in either the cytoplasm or in the case of CA-localization to the plasma membrane, perhaps in the extracellular space.

Fructose bisphosphate aldolase (FBA) plays a role in the regeneration of the ribulose 1,5-bisphosphate molecules used to produce glyceraldehyde-3-phosphate in the Calvin-Benson cycle (Calvin 1962). Multiple FBA genes are found in plant genomes with eight found in both Arabidopsis and tomato (Lu et al. 2012; Cai et al. 2016). In Arabidopsis, three isoforms have been shown to localize to the chloroplast and one to the cytosol, with the remaining four predicted to localize to the cytosol (Vidi et al. 2006; Ytterberg et al. 2006; Lu et al. 2012;

Garagounis et al. 2017). Upon further analysis, the blastP results identifying a chloroplastic form as having highest sequence similarity to the pokeweed sequence were further analyzed showed that the pokeweed protein sequence did not encode any region homologous to the 60 amino acid transit peptide of chloroplastic FBA1 (results not shown). Thus, the FBA identified as a PAP protein interactor is more likely cytosolic, where the interaction is more consistent with known PAP localization.

Despite its initial characterization as a housekeeping glycolytic protein that catalyzes the conversion of glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate, GAPDH is also a known moonlighting protein, with isoforms localizing to multiple subcellular compartments to perform additional functions in stress and defense responses in plants (Henry et al. 2015; Kim et al. 2020). Classically, the GAPB isoform is described as a nuclear encoded chloroplast-localizing isoform which has been studied for its participation in the Calvin-Benson cycle (Graciet et al. 2004); GAPB has been identified as part of the proteomic landscape of both the chloroplast stroma and envelope membranes (Ferro et al. 2003; Kleffmann et al. 2004). Interestingly, GAPB was also identified in a proteomic screening of the extracellular space in Arabidopsis using hydroponic isotope labelling coupled with mass spectrometry, demonstrating that it is also secreted; levels of secreted GAPB were enriched after oxidative stress treatment (Bindschedler et al. 2008). This indicates the potential for a PAP-GAPB interaction in the apoplast.

4.3 PAP interacts with a cysteine protease paCP1

The pokeweed protein PHYAM_ 026164 was characterized *in silico* as a likely papain-like cysteine protease (PLCP) belonging to the C1A family. PLCPs have been implicated in diverse physiological roles in plants, including seed germination, development, senescence, programmed cell death, abiotic stress responses, and immunity (Diaz-Mendoza et al. 2016; Szewińska et al. 2016; Pružinská et al. 2017; Bárány et al. 2018; Cai et al. 2018; Paulus et al. 2020). The C1A family consists of PLCP members targeted to the classical secretory pathway via N-terminal signal sequences (Richau et al. 2012). Since direct plant-pathogen interactions mostly occur extracellularly, apoplastic PLCPs are strongly linked to involvement in plant defense and immunity through microbe and pathogen perception, signalling cascade initiation, and as direct defense molecules with antipathogenic activity. This involvement is reflected in alterations in

pathogen resistance when secreted PLCP levels are changed through overexpression and knockouts. For example, tomato Pip1 mutants are hypersusceptible to apoplastic colonizers *Cladosporium fulvum*, *Phytophthora infestans*, and *Pseudomonas syringae*, representing fungal, oomycete, and bacterial pathogen classes, respectively, demonstrating the importance of a secreted PLCP as a broad-range immune protease (Ilyas et al. 2015).

PaCP1 homologs AtXCP1 and AtXCP2 from *Arabidopsis* belong to the apoplastic PLCP C1A family. PLCPs are further classified into 9 subfamilies based on phylogenetic analysis; AtXCP1 and AtXCP2 are members of subfamily III, the XCP-like subfamily, which also includes papain, the most well-known PLCP, also known as CpXCP5 (Richau et al. 2012; Liu et al. 2018). XCPs are named after the initial characterization of AtXCP1 and AtXCP2 from *Arabidopsis* root xylem cDNA libraries (Zhao et al. 2000); the xylem is considered part of the apoplast and provides a water transport system necessary for growth and development of plants (Farvardin et al. 2020). Immunofluorescence and GUS analysis of the XCP1 and XCP2 promoters in transgenic plants showed that expression was associated with xylem cells called tracheary elements, where XCP1 localized to the vacuoles and the cell walls of these cells (Funk et al. 2002). XCP1 and XCP2 were further shown to participate in tracheary element autolysis, a type of PCD that forms mature xylem; knockout mutant lines *xcp1* and *xcp1 xcp2* demonstrated delayed autolysis progression during xylem formation, after non-degraded cellular components were cleared at a slower rate in the loss-of-function plants (Avci et al. 2008). However, even in loss-of-function mutants, vessels have normal anatomy and autolysis proceeded to completion (Avci et al. 2008; Pérez-López et al. 2021), demonstrating that XCPs are not essential in growth and development. However, as these studies implicate XCPs in developmental programmed cell death, a role for the PAP-paCP1 interaction in programmed cell death should be considered.

More recent research into *Arabidopsis* XCP1/2 has revealed roles in pathogen resistance and plant defense responses (summarized in Table 8). Arv2 is a pathogen-secreted effector protein from *Cladosporium fulvum*, a biotrophic fungal pathogen that grows in the plant host apoplast. Heterologous expression of Arv2 in *Arabidopsis* showed that Avr2 binds and inactivates both XCP1 and XCP2, suppressing plant resistance to multiple pathogens and thus suggesting a role for XCP1 and XCP2 as basal defense proteins (Van Esse et al. 2008). Similarly, XCP1 is a target

Table 8. Summary of XCP1/XCP2 results in pathogen resistance.

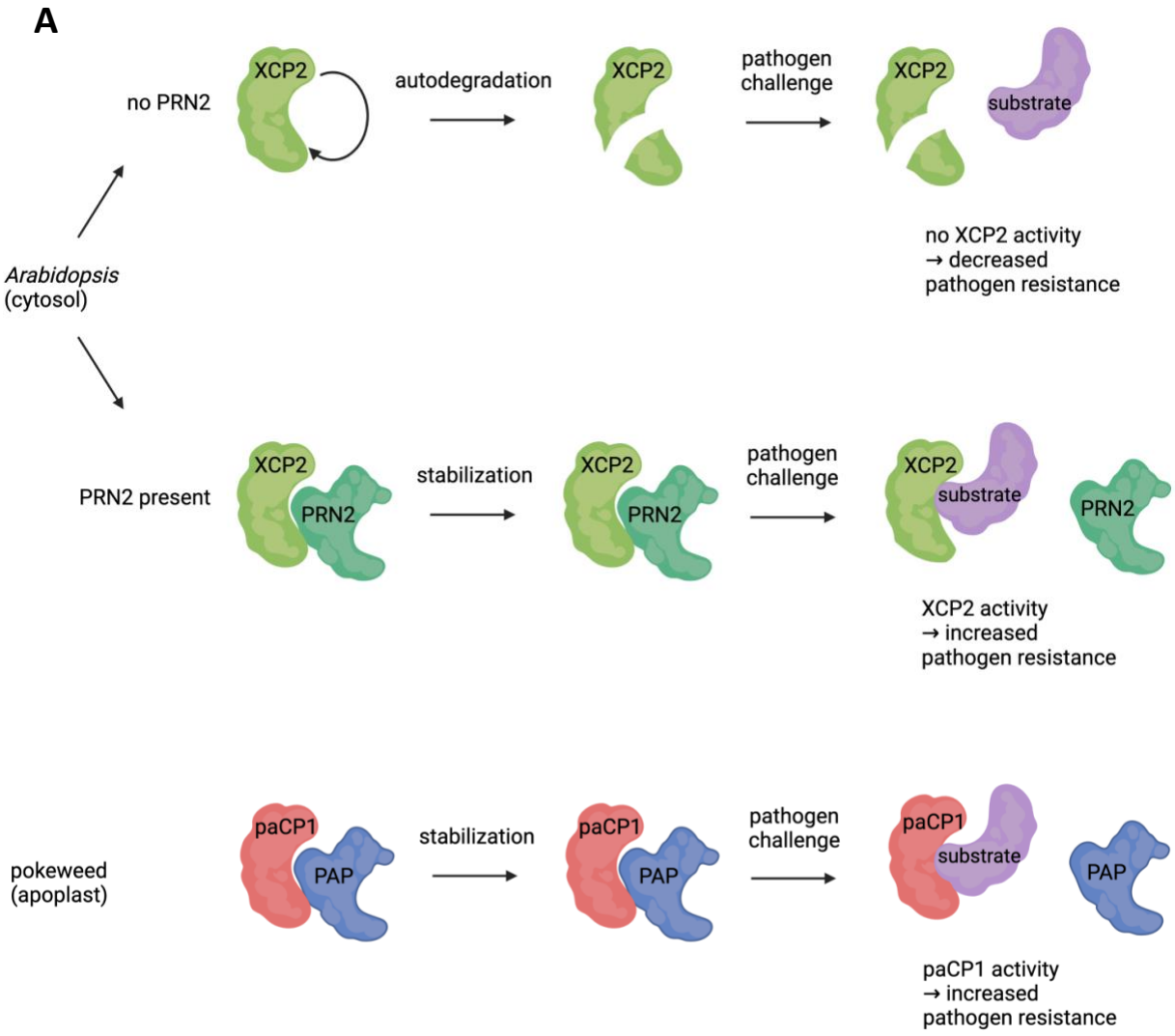
Plant species	Protease studied	Pathogen species and class	Effects after pathogen inoculation	Mechanism investigated	Reference
Arabidopsis thaliana	XCP1, XCP2	*Cladosporium fulvum* (biotrophic fungus)	XCP1/XCP2 inhibition decreases resistance to infection	XCP1/XCP2 inhibition by Avr2 (pathogen secreted)	Van Esse et al. 2008
Arabidopsis thaliana	XCP1	*Plasmodiphora brassicae* (protist)	XCP1 null mutants show less severe symptoms, increased resistance to infection	XCP1 inhibition by SSPbP53 (pathogen secreted)	Perez-Lopez et al. 2021
Arabidopsis thaliana	XCP2	*Ralstonia solanacearum* (bacteria)	XCP2 null mutants show increased resistance to infection	-	Zhang et al. 2014
Arabidopsis thaliana	XCP1	*Pseudomonas syringae* (bacteria)	XCP1 mutant shows decreased resistance to infection, unable to induce SAR	XCP1 cleaves PR1 (host protein) to produce CAPE9 peptide, induces SA biosynthesis	Chen et al. 2021
Zea mays	XCP2	*Xanthomonas oryzae* (bacteria)	XCP2 expression upregulated 30 mins after pathogen inoculation, XCP2 knockout shows increased disease symptoms, XCP2 overexpression shows decreased disease symptoms and induced expression of immune signalling and biotic stress resistance genes	-	Nino et al. 2020
Gossypium hirsutism	XCP1, XCP2	*Verticullium dahlia* (hemibiotrophic fungus)	XCP1 and XCP2 expression upregulated 6 hours post inoculation	-	Zhang et al. 2019
Arabidopsis thaliana	XCP2	*Ralstonia solanacearum* (bacteria)	XCP2 expression upregulated 9 days post-inoculation, PRN2 mutant shows no upregulation of XCP2 expression	PRN2 (host protein) reversibly binds XCP2, prevents XCP2 degradation and increases XCP2 activity	Zhang et al. 2014

of SSPbP53, an apoplastic cystatin-like protein effector secreted by the parasitic protist *Plasmodiphora brassicae* (Pérez-López et al. 2021). Not only was SSPbP53 shown to directly bind and inhibit XCP1 enzymatic activity, but downregulation of AtXCP1 expression was seen in infected *Arabidopsis* roots compared with mock-inoculated controls. Intriguingly, null mutant *At∆xcp1* plants showed reduced symptom severity and susceptibility to infection when compared with knockout mutants for other PLCPs rd19 and rd21, suggesting that the AtXCP1-effector interaction is a specific requirement for *P. brassicae* pathogenicity (Perez-Lopez et al. 2021). Similarly, *xcp2* null mutants are less susceptible to infection by bacterial pathogen *Ralstonia solanacearum* (Zhang et al. 2014). Recently, a direct mechanism for AtXCP1 participation in immunity has been demonstrated, linking it to defense signalling. AtXCP1 was shown to cleave a Pathogenesis-related protein 1 (AtPR1)-expressed propeptide and proteolytically activate AtCAPE9, an immunomodulatory peptide that induces antipathogen and antiherbivore responses (Chen et al. 2014; Chen et al. 2021). Since PR protein expression is critical in activating the SAR response in *Arabidopsis* (Wang et al. 2005) and AtCAPE9 induces production of SA, this implicates AtXCP1 in initiation and maintenance of systemic immunity through participation in a positive feedback loop. Additionally, AtXCP1 cleavage of AtPR1 was induced by both SA and bacterial challenge from *Pseudomonas syringae*, where its activity provided increased resistance to infection; bacterial pathogen-associated molecular pattern Flg22 was unable to elicit SAR in XCP1 double mutants (Chen et al. 2021). While the above studies clearly implicate XCP1 and XCP2 in immunity, its role appears to be pathogen dependent. The contrasting results highlight the contextual dependency of host proteins in pathosystem interactions: in the coevolutionary arms race between plants and pathogens, individual pathogens constantly evolve novel and unique mechanisms to hijack plant immunity. Thus, a single protein can both contribute to plant immunity but also play a role in susceptibility for specific pathogens. Therefore, to elucidate the PAP-paCP1 interaction involvement in immunity, multiple pathogen challenges should be considered as the interaction may have varying significance for each.

XCP homologues from species other than *Arabidopsis* demonstrate the conservation of the XCP PLCP subgroup throughout both monocots and dicots and its implication in pathogen response. In maize (*Zea mays*), biotrophic fungus *Ustilago maydis* inhibits ZmXCP2 by inducing expression of a host cysteine protease inhibitor, CC9, during early infection; CC9 localizes to the

apoplast where it inhibits endogenous ZmXCP2 protease activity (van der Linde et al. 2012). In rice (*Oryza sativa*), upregulation of OsXCP2 was seen 30 minutes after inoculation with bacterial pathogen *Xanthomonas oryzae* (Niño et al. 2020). OsXCP2 knockout and overexpression lines showed augmented and attenuated symptoms of bacterial blight, respectively, while transcriptome analysis of these lines indicated contrasting patterns of transcriptional modulation. DEGs identified in the overexpression lines included receptor-like kinases, calcium signalling, G-proteins, and diverse transcription factors, indicating a role for OsXCP2 in immune signalling. Activation of SA-dependent defense genes, up-regulation of PR proteins and biosynthesis of secondary metabolites were also seen, indicating resistance to biotic stress. Furthermore, in a dicot species, cotton, GhXCP1 and GhXCP2 were upregulated within 6 hours in response to fungal challenge with *Verticillium dahlia* (S. Zhang et al. 2019). It therefore seems reasonable to extend the XCP function in defense response, which has been conserved across diverse plant species, to pokeweed XCP-like PLCPs such as paCP1.

Pirin2 (PRN2) has been identified as an XCP2 interactor in *Arabidopsis*. Although PRN proteins are not well-understood in plants, studies indicate that these genes function in both developmental and pathogen induced PCD. A PRN2 homolog in *Zinnia* was identified as a potential regulator of PCD (Pesquet et al. 2013). In *Arabidopsis*, PRN2 binds to XCP2, inhibiting its enzymatic activity (Zhang et al. 2014). In this way, PRN2 prevents XCP autolytic degradation and ultimately increases its protease activity, as this interaction was demonstrated to be reversible. PRN2 is critical for upregulated AtXCP2 expression in response to *Ralstonia solanacerum* infection, which was not seen in *prn2* mutants. A member of the cupin domain-containing superfamily, PRN2 does not encode any known protease inhibitor domains (ie. those found in cystatins and serpins). Similarly, PAP does not possess any known inhibitor domains but could function to stabilize paCP1 through the interaction, or perhaps serve as a kind of scaffolding protein as proposed by the authors for PRN2 in *Arabidopsis*. The PRN2-AtXCP2 interaction was shown to occur in the cytosol; PAP may be able to provide a similar function in the pokeweed apoplast for paCP1 (Figure 8).


A
no PRN2
XCP2
autodegradation
XCP2
pathogen challenge
XCP2
substrate
no XCP2 activity
→ decreased
pathogen resistance
Arabidopsis
(cytosol)
PRN2 present
XCP2
PRN2
stabilization
XCP2
PRN2
pathogen challenge
XCP2
substrate
PRN2
XCP2 activity
→ increased
pathogen resistance
pokeweed
(apoplast)
paCP1
PAP
stabilization
paCP1
PAP
pathogen challenge
paCP1
substrate
PAP
paCP1 activity
→ increased
pathogen resistance

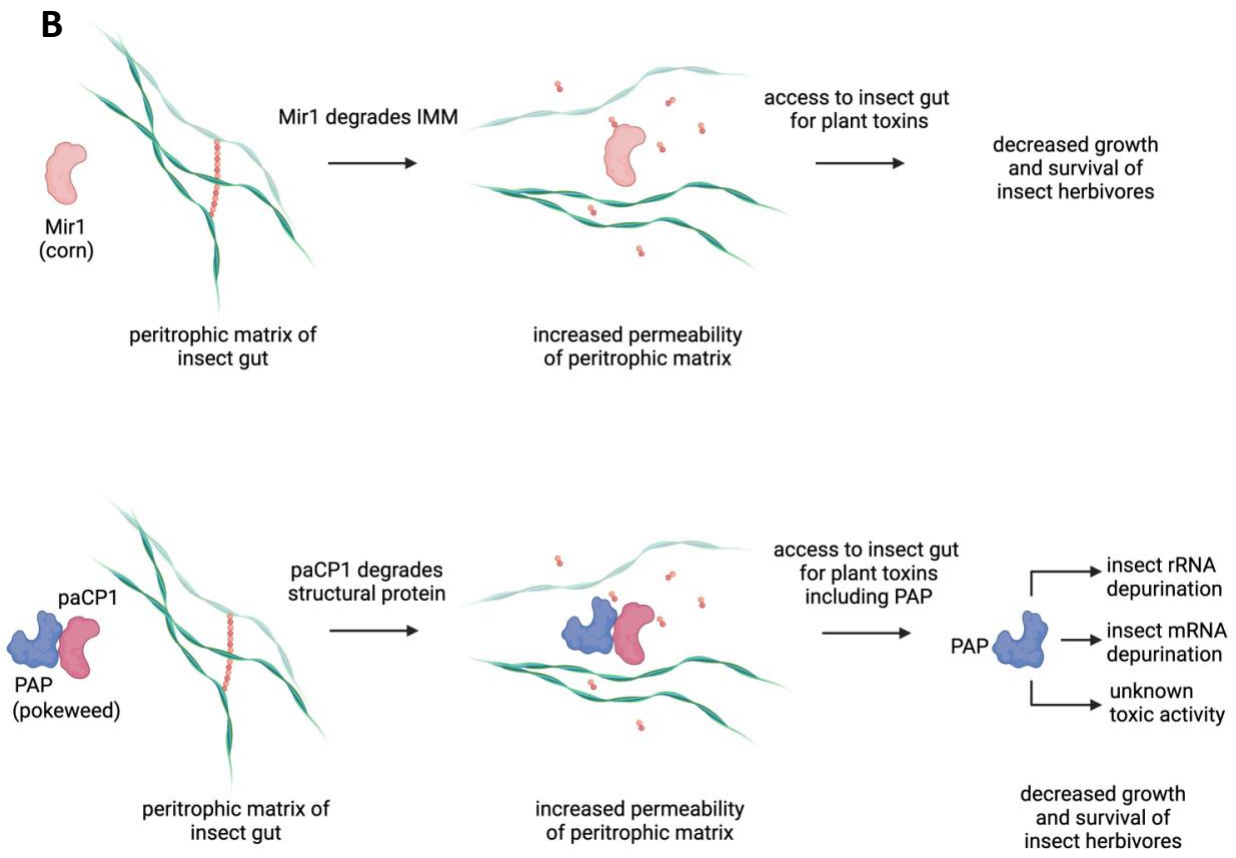


Figure 8. Potential mechanisms of the PAP-paCP1 interaction functioning in defense. (A) In *Arabidopsis,* PRN2 binds and inhibits XCP2, blocking its ability to autodegrade and resulting in increased XCP2 activity in response to pathogen challenge. In the absence of PRN2, XCP2 is autolytically degraded, resulting in low XCP2 activity and decreased pathogen resistance. In pokeweed, the PAP-paCP1 interaction may similarly stabilize the paCP1 protein, ultimately increasing paCP1 activity and limiting pathogen spread in the apoplast. **(B)** In corn, a mechanism by which secreted papain-like cysteine protease Mir1 increases resistance to insect herbivory has been elucidated. After ingestion by an herbivore, Mir1 cleaves insect intestinal mucin (IIM), a structural protein of the peritrophic matrix (PM), which provides a protective barrier to the insect gut. Mir1's action increases permeability of the gut to plant secondary metabolites, interfering with insect growth and survival. In pokeweed, paCP1 could serve a similar function, cleaving structural proteins of the PM; the paCP1-PAP interaction would result in PAP proximity to its target, potentially insect rRNA, mRNA or another unknown target. The toxicity of the PAP-paCP1 interaction would therefore provide additional resistance to insect herbivory. Figure created in BioRender.

A second possible function for the paCP1-PAP interaction may be as a direct defense molecule. Mir1 from corn is a C1A PLCP with anti-insect activity against pests including tobacco budworm, corn leaf aphids, and western corn rootworm (Mohan et al. 2008; Louis et al. 2015; Varsani et al. 2016). The mechanism through which Mir1 negatively affects growth of fall armyworm larvae (*Spodoptea frugiperda*) is through degradation of the peritrophic matrix of the insect gut, a semi-permeable structure that aids in digestion and protects against toxic plant secondary metabolites consumed by the insect. Mir1 was shown *in vivo* and *in vitro* to degrade insect intestinal mucin (IIM), the main structural protein of the midgut (Fescemyer et al. 2013). Peritrophic matrix degradation not only impairs digestion and interferes with uptake of nutrients, but also induces upregulation of genes for IIM replacement components as well as cysteine protease inhibitors in the midgut cells (Li et al. 2009; Fescemyer et al. 2013). It has been suggested that without this compensatory response complete peritrophic matrix degradation would occur (Fescemyer et al. 2013). Interestingly, feeding of PAP-S and other type I RIPs lychnin, momordin, gelonin, and saporin was shown to inhibit growth and decrease the survival rate of fall armyworm larvae by approximately 40% and 30%, respectively, in bioassays (Bertholdo-Vargas et al. 2009). It is therefore hypothesized that the paCP1-PAP interaction could play an antiherbivore role in pokeweed. The specificity of paCP1 towards a component of the insect digestive tract could position PAP closer to its target; PAP may interfere through mRNA depurination of upregulated transcripts of the compensatory response from the insect gut, through rRNA depurination of insect cells, or perhaps through a yet unknown mechanism (Figure 8).

A function for the paCP1-PAP interaction in an anti-herbivore response is also supported by co-expression of paCP1 and PAP genes in response to JA (see section 4.4). It is also important to note that as individual proteases can have multiple substrates and thus multiple functions (Paulus et al. 2020; Wang et al. 2021), the above hypotheses for the paCP1-PAP interaction are not mutually exclusive but could reflect context dependency for this interaction.

4.4 paCP1 and PAP isoforms are co-expressed upon defense response activation

Comparison of gene expression profiles for PAP and paCP1 show common upregulation in response to JA treatment. The jasmonate pathway comprises the core machinery that regulates

plant responses to wounding, insect herbivory, and necrotrophic pathogens (Degenhardt et al. 2010). Wounding alone is sufficient to induce accumulation of JA and its conjugate JA-isoleucine (JA-Ile) through plant perception of damage-associated molecular patterns (Heil et al. 2012). JA/JA-Ile production is then amplified if leaves are exposed to effector molecules present in insect oral secretions that suppress plant defenses, termed herbivore-associated molecular patterns (Schmelz et al. 2009). The JA-induced signal transduction pathway then activates transcription of plant defense genes, producing antibiotic and antixenotic molecules which adversely influence herbivore development and behaviour (Bleeker et al. 2009; Chen et al. 2012; Luan et al. 2013). Furthermore, herbivore challenges frequently involve additional pathogens, such as insect vector-transmitted viruses, or the threat of bacterial and fungal pathogens able to access the tissue through the wound (Liu et al. 2017). Therefore, proteins showing jasmonate-responsive upregulation are specialized for powerful deterrence of plant enemies.

Despite the focus implicating RIPs in plant defense, relatively little work exists into specific RIP involvement in defense against insects; however, several RIPs have been shown to enhance plant resistance to insect herbivory (Zhu et al. 2018). When fed transgenic RIP-expressing leaves or RIP-supplemented diets, biomass and survival of larvae decreased as well as adult fecundity, demonstrating the insecticidal activity of many RIPs on invertebrate growth and development. Type I RIPs that have demonstrated insecticidal activity using these methods include saporin, gelonin, PAP-S, lychnin, α-momocharin and a type I RIP isolated from apple (Bertholdo-Vargas et al. 2009; Hamshou et al. 2016; Hamshou et al. 2017). Therefore, combined with previously discussed examples implicating PLCPs in defense against herbivory (see section 4.3), a role for the paCP1-PAP interaction in an anti-insect defense response should be considered.

Interestingly, paCP1 did not show SA-responsiveness from the transcriptome data, in contrast with the most abundant PAP transcripts, PAP-1 and PAP-2, which demonstrated SA-responsive upregulation. This result indicates that the PAP-paCP1 interaction is not expected to act as a defense molecule produced in the SAR response, which comprises long-distance signalling mechanisms resulting in distal tissue disease resistance. SA-responsiveness is also linked to resistance against biotrophic pathogens, so taken together with JA-induced co-expression, this indicates that the paCP1-PAP interaction more likely plays a role in the JA-activated responses

to necrotrophic pathogens and herbivorous insects. However, these results do not completely exclude the possibility that the paCP1-PAP interaction is involved in SA-induced defense. paCP1 may in fact be SA-responsive but this expression was not captured by the single time point of the pokeweed transcriptome data, which was collected 24 hours post SA treatment. In rice, expression of OsXCP1 increased 5-fold 12 hours after treatment with SA; by 24 hours transcript levels had decreased to pre-treatment levels (Niño et al. 2020). Further work in rice showed that upregulation of direct defense-related genes in response to bacterial challenge was not seen to occur until 72-120 hours post-inoculation (Gan et al. 2011). Investigation of phytohormone-induced PAP and paCP1 expression patterns using a time course would provide more precise information regarding in which defense responses and at what stage the paCP1-PAP interaction could play a role.

4.5 Future work

Four PAP isoforms were identified in the pull-downs using the polyclonal PAP antibody: PAP-1, PAP-2, K-PAP, and a novel PAP isoform. As the reverse co-IP was performed from pokeweed leaf lysate and not purified protein, it is unknown which PAP isoform binds paCP1; the possibility that paCP1 binds multiple isoforms is also possible. Individual expression and isolation of PAP isoforms from *E. coli* followed by co-immunoprecipitation of purified PAP protein with paCP1 as the bait would identify which isoform(s) interact(s); alternately, paCP1-PAP isoform interactions could be tested using methods such as targeted yeast two-hybrid (Y2H), bimolecular fluorescence complementation (BiFC), or fluorescence resonance energy transfer (FRET).

Verification of paCP1 as a cysteine protease should also be performed using an *in vitro* functional assay for confirmation of well-known structural and functional PLCP characteristics. Catalytic activity of recombinant paCP1 could be tested using general protease substrates gelatin or casein, followed by inhibition with irreversible PLCP inhibitor E-64 to identify cysteine protease specificity.

Determining location of the paCP1-PAP interaction would provide initial information as to the nature of its function. PAP and paCP1 could be detected from isolated apoplastic fluid to confirm common extracellular localization, or by using confocal microscopy for visualization of fusion fluorescent proteins would be an important first step; however, as aforementioned, proteins are frequently localized to multiple locations within and outside the cell, and/or have dynamic localizations. Therefore, localization studies of the individual proteins may not provide enough information to conclude in which subcellular compartment the interaction occurs. Using BiFC or FRET microscopy would provide complementary information on subcellular location of the paCP1-PAP interaction, while controlling for dynamic and multiple protein localization.

Co-expression analysis implicated paCP1 and PAP in the JA-induced defense response. The biological role of the PAP-paCP1 interaction could be further investigated by generation of stable transgenic *Arabidopsis* plant lines expressing paCP1 and PAP. Pathogen challenge from various pathogen classes would demonstrate whether the interaction provides increased resistance for the plant. However, as overexpression lines may not illustrate endogenous role of the proteins, ideal experiments would involve generation of PAP and paCP1 knockdown lines in pokeweed to verify whether the interaction is required for immunity. Additionally, further proteomic screens using paCP1 as bait could provide mechanistic clues for how the PAP-paCP1 interaction may function *in planta*. Together, these experiments would investigate the preliminary hypothesis that the paCP1-PAP interaction has a biological role in immunity and pathogen defense.

The current study also revealed that in addition to paCP1, PAP putatively interacts with fructose-bisphosphate aldolase, beta carbonic anhydrase, elongation factor 1-alpha, glyceraldehyde-3-phosphate dehydrogenase B, and ribosomal protein S26-1. Validation of these interactions should be performed using reverse co-IP similarly to what was performed for paCP1-PAP. Exploring these interactions through overexpression and knockdown experiments, as well as co-expression analysis of induced transcripts and/or protein levels would provide additional information as to PAP's potential dynamic functions in the cell.

5. REFERENCES

Abe M, Takahashi T, Komeda Y. 2001. Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. Plant J. 26(5):487–494.

Acevedo-Hernández GJ, León P, Herrera-Estrella LR. 2005. Sugar and ABA responsiveness of a minimal RBCS light-responsive unit is mediated by direct binding of ABI4. Plant J. 43(4):506–519.

Acuner Ozbabacan SE, Engin HB, Gursoy A, Keskin O. 2011. Transient protein-protein interactions. Protein Eng Des Sel. 24(9):635–648.

Aebersold R, Mann M. 2003. Mass spectrometry-based proteomics. Nature. 422(6928):198–207.

Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 37(4):420–423.

Alvarez-Ponce D. 2017. Recording negative results of protein–protein interaction assays: an easy way to deal with the biases and errors of interactomic data sets. Brief Bioinform. 18(6):1017–1020.

Aron GM, Irvin JD. 1980. Inhibition of herpes simplex virus multiplication by the pokeweed antiviral protein. Antimicrob Agents Chemother. 17(6):1032–1033.

Aubry S, Smith-Unna RD, Boursnell CM, Kopriva S, Hibberd JM. 2014. Transcript residency on ribosomes reveals a key role for the Arabidopsis thaliana bundle sheath in sulfur and glucosinolate metabolism. Plant J. 78(4):659–673.

Avci U, Earl Petzold H, Ismail IO, Beers EP, Haigler CH. 2008. Cysteine proteases XCP1 and XCP2 aid micro-autolysis within the intact central vacuole during xylogenesis in Arabidopsis roots. Plant J. 56(2):303–315.

Balakireva AV, Zamyatnin AA. 2018. Indispensable role of proteases in plant innate immunity. Int J Mol Sci. 19(2):629.

Bárány I, Berenguer E, Solís MT, Pérez-Pérez Y, Santamaría ME, Crespo JL, Risueño MC, Díaz I, Testillano PS. 2018. Autophagy is activated and involved in cell death with participation of cathepsins during stress-induced microspore embryogenesis in barley. J Exp Bot. 69(6):1387–1402.

Barbieri L, Ciani M, Girbés T, Liu WY, Van Damme EJM, Peumans WJ, Stirpe F. 2004. Enzymatic activity of toxic and non-toxic type 2 ribosome-inactivating proteins. FEBS Lett. 563(1–3):219–222.

Barbieri L, Gorini P, Valbonesi P, Castiglioni P, Stirpe F. 1994. Unexpected activity of saporins. Nature. 372(6507):624.

Barbieri L, Polito L, Bolognesi A, Ciani M, Pelosi E, Farini V, Jha AK, Sharma N, Vivanco JM, Chambery A, et al. 2006. Ribosome-inactivating proteins in edible plants and purification and characterization of a new ribosome-inactivating protein from Cucurbita moschata. Biochim Biophys Acta. 1760(5):783–792.

Barbieri L, Valbonesi P, Bonora E, Gorini P, Bolognesi A, Stirpe F. 1997. Polynucleotide:adenosine glycosidase activity of ribosome-inactivating proteins: effect on DNA, RNA and poly(A). Nucleic Acids Res. 25(3):518–522.

Baslam M, Oikawa K, Kitajima-Koga A, Kaneko K, Mitsui T. 2016. Golgi-to-plastid trafficking of proteins through secretory pathway: Insights into vesicle-mediated import toward the plastids. Plant Signal Behav. 11(9):e1221558.

Bass HW, Krawetz JE, Obrian GR, Zinselmeier C, Habben JE, Boston RS. 2004. Maize ribosome-inactivating proteins (RIPs) with distinct expression patterns have similar requirements for proenzyme activation. J Exp Bot. 55(406):2219–2233.

Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Bursteinas B, et al. 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 49(D1):D480–D489.

Baykal U, Tumer NE. 2007. The C-terminus of pokeweed antiviral protein has distinct roles in transport to the cytosol, ribosome depurination and cytotoxicity. Plant J. 49(6):995–1007.

Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E. 2015. The Arabidopsis information resource: Making and mining the “gold standard” annotated reference plant genome. Genesis. 53(8):474–485.

Bertholdo-Vargas LR, Martins JN, Bordin D, Salvador M, Schafer AE, Barros NM de, Barbieri L, Stirpe F, Carlini CR. 2009. Type 1 ribosome-inactivating proteins - entomotoxic, oxidative and genotoxic action on Anticarsia gemmatalis (Hübner) and Spodoptera frugiperda (J.E. Smith) (Lepidoptera: Noctuidae). J Insect Physiol. 55(1):51–58.

Bindschedler L V., Palmblad M, Cramer R. 2008. Hydroponic isotope labelling of entire plants (HILEP) for quantitative plant proteomics; an oxidative stress case study. Phytochemistry. 69(10):1962–1972.

Bleeker PM, Diergaarde PJ, Ament K, Guerra J, Weidner M, Schütz S, de Both MTJ, Haring MA, Schuurink RC. 2009. The role of specific tomato volatiles in tomato-whitefly interaction. Plant Physiol. 151(2):925–935.

Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S, et al. 2021. The InterPro protein families and domains database: 20

years on. Nucleic Acids Res. 49(D1):D344.

Bonness MS, Ready MP, Irvin JD, Mabry TJ. 1994. Pokeweed antiviral protein inactivates pokeweed ribosomes; implications for the antiviral mechanism. Plant J. 5(2):173–183.

Brigotti M, Rambelli F, Zamboni M, Montanaro L, Sperti S. 1989. Effect of alpha-sarcin and ribosome-inactivating proteins on the interaction of elongation factors with ribosomes. Biochem J. 257(3):723.

Buchel AS, Brederode FT, Bol JF, Linthorst HJM. 1999. Mutation of GT-1 binding sites in the Pr-1A promoter influences the level of inducible gene expression in vivo. Plant Mol Biol. 40(3):387–396.

Cai B, Li Q, Xu Y, Yang L, Bi H, Ai X. 2016. Genome-wide analysis of the fructose 1,6-bisphosphate aldolase (FBA) gene family and functional characterization of FBA7 in tomato. Plant Physiol Biochem. 108:251–265.

Cai YM, Yu J, Ge Y, Mironov A, Gallois P. 2018. Two proteases with caspase-3-like activity, cathepsin B and proteasome, antagonistically control ER-stress-induced programmed cell death in Arabidopsis. New Phytol. 218(3):1143–1155.

Çakır B, Tumer NE, Çakir B. 2015. Arabidopsis Bax Inhibitor-1 inhibits cell death induced by pokeweed antiviral protein in Saccharomyces cerevisiae. Microb Cell. 2(2):43-56.

Calvin M. 1962. The path of carbon in photosynthesis. Science. 135(3507):879–889.

Chaboute ME, Clement B, Sekine M, Philipps G, Chaubet-Gigot N. 2000. Cell cycle regulation of the tobacco ribonucleotide reductase small subunit gene is mediated by E2F-like elements. Plant Cell. 12(10):1987–1999.

Chan CS, Guo L, Shih MC. 2001. Promoter analysis of the nuclear gene encoding the chloroplast glyceraldehyde-3-phosphate dehydrogenase B subunit of Arabidopsis thaliana. Plant Mol Biol. 46(2):131–141.

Chen R, Jiang H, Li L, Zhai Q, Qi L, Zhou W, Liu X, Li H, Zheng W, Sun J, et al. 2012. The Arabidopsis mediator subunit MED25 differentially regulates jasmonate and abscisic acid signaling through interacting with the MYC2 and ABI5 transcription factors. Plant Cell. 24(7):2898–2916.

Chen S, Zhao H, Luo T, Liu Y, Nie X, Li H. 2019. Characteristics and Expression Pattern of MYC Genes in Triticum aestivum, Oryza sativa, and Brachypodium distachyon. Plants (Basel). 8(8):274.

Chen Y, Lin F, Wang H. 2021. XCP1 is a caspase that proteolyzes Pathogenesis-related protein 1 to produce the cytokine CAPE9 for systemic immunity in Arabidopsis. Res Square. 10.21203/rs.3.rs-155784/v1

Chen YL, Lee CY, Cheng KT, Chang WH, Huang RN, Nam HG, Chen YR. 2014. Quantitative peptidomics study reveals that a wound-induced peptide from PR-1 regulates immune signaling in tomato. Plant Cell. 26(10):4135–4148.

Chen Z, Antoniw JF, Lin Q, White RF. 1993. Expression of pokeweed (Phytolacca americana) antiviral protein cDNA in Escherichia coli and its antiviral activity. Physiol Mol Plant Pathol. 42(4):237–247.

Cheong YH, Yoo CM, Park JM, Ryu GR, Goekjian VH, Nagao RT, Key JL, Cho MJ, Hong JC. 1998. STF1 is a novel TGACG-binding factor with a zinc-finger motif and a bZIP domain which heterodimerizes with GBF proteins. Plant J. 15(2):199–209.

Choi HI, Hong JH, Ha JO, Kang JY, Kim SY. 2000. ABFs, a family of ABA-responsive element binding factors. J Biol Chem. 275(3):1723–1730.

Chow CN, Lee TY, Hung YC, Li GZ, Tseng KC, Liu YH, Kuo PL, Zheng HQ, Chang WC. 2019. PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. Nucleic Acids Res. 47(D1):D1155–D1163.

Chuang WP, Herde M, Ray S, Castano-Duque L, Howe GA, Luthe DS. 2014. Caterpillar attack triggers accumulation of the toxic maize protein RIP2. New Phytol. 201(3):928–939.

Citores L, Iglesias R, Ferreras JM. 2021. Antiviral Activity of Ribosome-Inactivating Proteins. Toxins (Basel). 13(2):80.

Di Cola A, Frigerio L, Lord JM, Ceriotti A, Roberts LM. 2001. Ricin A chain without its partner B chain is degraded after retrotranslocation from the endoplasmic reticulum to the cytosol in plant cells. Proc Natl Acad Sci USA. 98(25):14726–14731.

Collins RM, Afzal M, Ward DA, Prescott MC, Sait SM, Rees HH, Brian Tomsett A. 2010. Differential proteomic analysis of Arabidopsis thaliana genotypes exhibiting resistance or susceptibility to the insect herbivore, Plutella xylostella. PLoS One. 5(4):e10103.

Conlon FL, Miteva Y, Kaltenbrun E, Waldron L, Greco TM, Cristea IM. 2012. Immunoisolation of protein complexes from Xenopus. Methods Mol Biol. 917:369–390.

Coulombe R, Grochulski P, Sivaraman J, Ménard R, Mort JS, Cygler M. 1996. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. EMBO J. 15(20):5492–5503.

Cstorer A, Ménard R. 1994. Catalytic mechanism in papain family of cysteine peptidases. Methods Enzymol. 244(C):486–500.

De Jager SM, Menges M, Bauer UM, Murray JAH. 2001. Arabidopsis E2F1 binds a sequence

present in the promoter of S-phase-regulated gene AtCDC6 and is a member of a multigene family with differential activities. Plant Mol Biol. 47(4):555–568.

De Zaeytijd J, Rougé P, Smagghe G, Van Damme EJM. 2019. Structure and activity of a cytosolic ribosome-inactivating protein from rice. Toxins (Basel). 11(6):325.

Degenhardt DC, Refi-Hind S, Stratmann JW, Lincoln DE. 2010. Systemin and jasmonic acid regulate constitutive and herbivore-induced systemic volatile emissions in tomato, Solanum lycopersicum. Phytochemistry. 71(17–18):2024–2037.

Depuydt T, Vandepoele K. 2021. Multi-omics network-based functional annotation of unknown Arabidopsis genes. Plant J. 108(4):1193–1212.

Després C, DeLong C, Glaze S, Liu E, Fobert PR. 2000. The Arabidopsis NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. Plant Cell. 12(2):279–290.

Diaz-Mendoza M, Dominguez-Figueroa JD, Velasco-Arroyo B, Cambra I, Gonzalez-Melendi P, Lopez-Gonzalvez A, Garcia A, Hensel G, Kumlehn J, Diaz I, et al. 2016. HvPAP-1 C1A protease and HvCPI-2 cystatin contribute to barley grain filling and germination. Plant Physiol. 170(4):2511–2524.

DiMario RJ, Clayton H, Mukherjee A, Ludwig M, Moroney JV. 2017. Plant carbonic anhydrases: structures, locations, evolution, and physiological roles. Mol Plant. 10(1):30–46.

Domashevskiy AV, Miyoshi H, Goss DJ. 2012. Inhibition of Pokeweed Antiviral Protein (PAP) by turnip mosaic virus genome-linked protein (VPg). J Biol Chem. 287(35):29729–29738.

Duggar BM, Armstrong JK. 1925. The Effect of Treating the Virus of Tobacco Mosaic With the Juices of Various Plants. Ann Missouri Bot Gard. 12(4):359.

Endo Y, Mitsui K, Motizuki M, Tsurugi K. 1987. The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. The site and the characteristics of the modifications in 28S ribosomal RNA caused by the toxins. J Biol Chem. 262(12):5908–5912.

Espinosa-Cantú A, Ascencio D, Herrera-Basurto S, Xu J, Roguev A, Krogan NJ, DeLuna A. 2018. Protein moonlighting revealed by noncatalytic phenotypes of yeast enzymes. Genetics. 208(1):419–431.

Ezcurra I, Ellerström M, Wycliffe P, Stålberg K, Rask L. 1999. Interaction between composite elements in the napA promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. Plant Mol Biol. 40(4):699–709.

Fabre N, Reiter IM, Becuwe-Linka N, Genty B, Rumeau D. 2007. Characterization and expression analysis of genes encoding alpha and beta carbonic anhydrases in Arabidopsis. Plant Cell Environ. 30(5):617–629.

Farvardin A, González-hernández AI, Llorens E, García-agustín P, Scalschi L, Vicedo B. 2020. The Apoplast: A Key Player in Plant Survival. Antioxidants (Basel). 9(7):1–26.

Ferro M, Salvi D, Brugière S, Miras S, Kowalski S, Louwagie M, Garin J, Joyard J, Rolland N. 2003. Proteomics of the chloroplast envelope membranes from Arabidopsis thaliana. Mol Cell Proteomics. 2(5):325–345.

Fescemyer HW, Sandoya GV, Gill TA, Ozkan S, Marden JH, Luthe DS. 2013. Maize toxin degrades peritrophic matrix proteins and stimulates compensatory transcriptome responses in fall armyworm midgut. Insect Biochem Mol Biol. 43(3):280–291.

Franklin KA, Quail PH. 2010. Phytochrome functions in Arabidopsis development. J Exp Bot. 61(1):11–24.

Frigerio L, Vitale A, Lord JM, Ceriotti A, Roberts LM. 1998. Free ricin A chain, proricin, and native toxin have different cellular fates when expressed in tobacco protoplasts. J Biol Chem. 273(23):14194–14199.

Funk V, Kositsup B, Zhao C, Beers EP. 2002. The Arabidopsis xylem peptidase XCP1 is a tracheary element vacuolar protein that may be a papain ortholog. Plant Physiol. 128(1):84–94.

Gahir S, Bharath P, Raghavendra AS. 2021. Stomatal closure sets in motion long-term strategies of plant defense against microbial pathogens. Front Plant Sci. 12:761952.

Gan Q, Bai H, Zhao X, Tao Y, Zeng H, Han Y, Song W, Zhu L, Liu G. 2011. Transcriptional characteristics of Xa21-mediated defense responses in rice. J Integr Plant Biol. 53(4):300–311.

Gao QM, Zhu S, Kachroo P, Kachroo A. 2015. Signal regulators of systemic acquired resistance. Front Plant Sci. 6:228.

Garagounis C, Kostaki KI, Hawkins TJ, Cummins I, Fricker MD, Hussey PJ, Hetherington AM, Sweetlove LJ. 2017. Microcompartmentation of cytosolic aldolase by interaction with the actin cytoskeleton in Arabidopsis. J Exp Bot. 68(5):885–898.

Gholizadeh A. 2019. Purification of a ribosome-inactivating protein with antioxidation and root developer potencies from Celosia plumosa. Physiol Mol Biol Plants. 25(1):243–251.

Gingras AC, Gstaiger M, Raught B, Aebersold R. 2007. Analysis of protein complexes using mass spectrometry. Nat Rev Mol Cell Biol. 8(8):645–654.

Giuliano G, Pichersky E, Malik VS, Timko MP, Scolnik PA, Cashmore AR. 1988. An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. Proc Natl Acad Sci USA. 85(19):7089–7093.

Gookin TE, Assmann SM. 2014. Significant reduction of BiFC non-specific assembly facilitates

in planta assessment of heterotrimeric G-protein interactors. Plant J. 280(3):553-67.

Graciet E, Lebreton S, Gontero B. 2004. Emergence of new regulatory mechanisms in the Benson-Calvin pathway via protein-protein interactions: a glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase complex. J Exp Bot. 55(400):1245–1254.

Hamshou M, Shang C, Smagghe G, Van Damme EJM. 2016. Ribosome-inactivating proteins from apple have strong aphicidal activity in artificial diet and in planta. Crop Prot. 87:19–24.

Hamshou M, Shang C, De Zaeytijd J, Van Damme EJM, Smagghe G. 2017. Expression of ribosome-inactivating proteins from apple in tobacco plants results in enhanced resistance to Spodoptera exigua. J Asia Pac Entomol. 20(1):1–5.

Hartley MR, Chacidock JA, Bonness MS. 1996. The structure and function of ribosome-inactivating proteins. Trends Plant Sci. 1(8):252.

Hattori T, Totsuka M, Hobo T, Kagaya Y, Yamamoto-Toyoda A. 2002. Experimentally determined sequence requirement of ACGT-containing abscisic acid response element. Plant Cell Physiol. 43(1):136–140.

Heil M, Ibarra-Laclette E, Adame-Álvarez RM, Martínez O, Ramirez-Chávez E, Molina-Torres J, Herrera-Estrella L. 2012. How plants sense wounds: damaged-self recognition is based on plant-derived elicitors and induces octadecanoid signaling. PLoS One. 7(2):e30537.

Helmy M, Lombard S, Piéroni G. 1999. Ricin RCA60: evidence of its phospholipase activity. Biochem Biophys Res Commun. 258(2):252–255.

Henry E, Fung N, Liu J, Drakakaki G, Coaker G. 2015. Beyond glycolysis: GAPDHs are multi-functional enzymes involved in regulation of ROS, autophagy, and plant immune responses. PLoS Genet. 11(4):e1005199.

Hines KM, Chaudhari V, Edgeworth KN, Owens TG, Hanson MR. 2021. Absence of carbonic anhydrase in chloroplasts affects C3 plant development but not photosynthesis. Proc Natl Acad Sci USA. 118(33).

Honjo E, Dong D, Motoshima H, Watanabe K. 2002. Genomic clones encoding two isoforms of pokeweed antiviral protein in seeds (PAP-S1 and S2) and the N-glycosidase activities of their recombinant proteins on ribosomes and DNA in comparison with other isoforms. J Biochem. 131(2):225–231.

Hu H, Rappel WJ, Occhipinti R, Ries A, Böhmer M, You L, Xiao C, Engineer CB, Boron WF, Schroeder JI. 2015. Distinct Cellular Locations of Carbonic Anhydrases Mediate Carbon Dioxide Control of Stomatal Movements. Plant Physiol. 169(2):1168–1178.

Hudak KA, Dinman JD, Tumer NE. 1999. Pokeweed antiviral protein accesses ribosomes by binding to L3. J Biol Chem. 274(6):3859–3864.

Hudson ME, Quail PH. 2003. Identification of Promoter Motifs Involved in the Network of Phytochrome A-Regulated Gene Expression by Combined Analysis of Genomic Sequence and Microarray Data. Plant Physiol. 133(4):1605.

Hudson ME, Quail PH. 2003. Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. Plant Physiol. 133(4):1605–1616.

Hur Y, Hwang DJ, Zoubenko O, Coetzer C, Uckun FM, Tumer NE. 1995. Isolation and characterization of pokeweed antiviral protein mutations in Saccharomyces cerevisiae: identification of residues important for toxicity. Proc Natl Acad Sci USA. 92(18):8448–8452.

Hwang J, Qi L. 2018. Quality Control in the Endoplasmic Reticulum: Crosstalk between ERAD and UPR pathways. Trends Biochem Sci. 43(8):593–605.

Iglesias R, Citores L, Di Maro A, Ferreras JM. 2015. Biological activities of the antiviral protein BE27 from sugar beet (Beta vulgaris L.). Planta. 241(2):421–433.

Iglesias R, Pérez Y, De Torre C, Ferreras JM, Antolín P, Jiménez P, Rojo MÁ, Méndez E, Girbés T. 2005. Molecular characterization and systemic induction of single-chain ribosome-inactivating proteins (RIPs) in sugar beet (Beta vulgaris) leaves. J Exp Bot. 56(416):1675–1684.

Ilyas M, Hörger AC, Bozkurt TO, Van Den Burg HA, Kaschani F, Kaiser M, Belhaj K, Smoker M, Joosten MHAJ, Kamoun S, et al. 2015. Functional divergence of two secreted immune proteases of tomato. Curr Biol. 25(17):2300–2306.

Irvin JD. 1975. Purification and partial characterization of the antiviral protein from Phytolacca americana which inhibits eukaryotic protein synthesis. Arch Biochem Biophys. 169(2):522–528.

Irvin JD, Kelly T, Robertus JD. 1980. Purification and properties of a second antiviral protein from Phytolacca americana which inactivates eukaryotic ribosomes. Arch Biochem Biophys. 200(2):418–425.

Jansen R, Greenbaum D, Gerstein M. 2002. Relating whole-genome expression data with protein-protein interactions. Genome Res. 12(1):37–46.

Jiang SY, Bhalla R, Ramamoorthy R, Luan HF, Venkatesh PN, Cai M, Ramachandran S. 2012. Over-expression of OSRIP18 increases drought and salt tolerance in transgenic rice plants. Transgenic Res. 21(4):785–795.

Jiang SY, Ramamoorthy R, Bhalla R, Luan HF, Venkatesh PN, Cai M, Ramachandran S. 2008. Genome-wide survey of the RIP domain family in Oryza sativa and their expression profiles under various abiotic and biotic stresses. Plant Mol Biol. 67(6):603–614.

Jiao Y, Ma L, Strickland E, Deng XW. 2005. Conservation and divergence of light-regulated

genome expression patterns during seedling development in rice and Arabidopsis. Plant Cell. 17(12):3239–3256.

Johnson C, Boden E, Arias J. 2003. Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. Plant Cell. 15(8):1846–1858.

Jolliffe NA, Brown JC, Neumann U, Vicré M, Bachi A, Hawes C, Ceriotti A, Roberts LM, Frigerio L. 2004. Transport of ricin and 2S albumin precursors to the storage vacuoles of Ricinus communis endosperm involves the Golgi and VSR-like receptors. Plant J. 39(6):821–833.

Kamiya N, Nagasaki H, Morikami A, Sato Y, Matsuoka M. 2003. Isolation and characterization of a rice WUSCHEL-type homeobox gene that is specifically expressed in the central cells of a quiescent center in the root apical meristem. Plant J. 35(4):429–441.

Kanzaki H, Kagemori T, Yamachika Y, Nitoda T, Kawazu K. 1999. Inhibition of Plant Transformation by Phytolaccoside B from Phytolacca americana Callus. Biosci Biotechnol Biochem. 63(9):1657–1659.

Kataoka J, Habuka N, Masuta C, Miyano M, Koiwai A. 1992. Isolation and analysis of a genomic clone encoding a pokeweed antiviral protein. Plant Mol Biol. 20(5):879–886.

Kataoka J, Habuka N, Miyano M, Masuta C, Koiwai A. 1992. Adenine depurination and inactivation of plant ribosomes by an antiviral protein of Mirabilis jalapa (MAP). Plant Mol Biol. 20(6):1111–1119.

Kaul S, Koo HL, Jenkins J, Rizzo M, Rooney T, Tallon LJ, Feldblyum T, Nierman W, Benito MI, Lin X, et al. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature. 408(6814):796–815.

Kim DW, Lee SH, Choi SB, Won SK, Heo YK, Cho M, Park Y Il, Cho HT. 2006. Functional conservation of a root hair cell-specific cis-element in angiosperms with different root hair distribution patterns. Plant Cell. 18(11):2959–2970.

Kim JS, Mizoi J, Yoshida T, Fujita Y, Nakajima J, Ohori T, Todaka D, Nakashima K, Hirayama T, Shinozaki K, et al. 2011. An ABRE promoter sequence is involved in osmotic stress-responsive expression of the DREB2A gene, which encodes a transcription factor regulating drought-inducible genes in Arabidopsis. Plant Cell Physiol. 52(12):2136–2146.

Kim SC, Guo L, Wang X. 2020. Nuclear moonlighting of cytosolic glyceraldehyde-3-phosphate dehydrogenase regulates Arabidopsis response to heat stress. Nat Commun. 11(1):3439.

Kleffmann T, Russenberger D, Von Zychlinski A, Christopher W, Sjölander K, Gruissem W, Baginsky S. 2004. The Arabidopsis thaliana chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol. 14(5):354–362.

Ko JH, Beers EP, Han KH. 2006. Global comparative transcriptome analysis identifies gene

network regulating secondary xylem development in Arabidopsis thaliana. Mol Genet Genomics. 276(6):517–531.

Krantz M, Klipp E. 2020. Moonlighting proteins - an approach to systematize the concept. In Silico Biol. 14(1–2):71-83.

Lau OS, Huang X, Charron JB, Lee JH, Li G, Deng XW. 2011. Interaction of Arabidopsis DET1 with CCA1 and LHY in mediating transcriptional repression in the plant circadian clock. Mol Cell. 43(5):703.

Le Gourrierec J, Li YF, Zhou DX. 1999. Transcriptional activation by Arabidopsis GT-1 may be through interaction with TFIIA-TBP-TATA complex. Plant J. 18(6):663–668.

Li C, Song X, Li G, Wang P. 2009. Midgut cysteine protease-inhibiting activity in Trichoplusia ni protects the peritrophic membrane from degradation by plant cysteine proteases. Insect Biochem Mol Biol. 39(10):726–734.

Li MX, Yeung HW, Pan LP, Chan SI. 1991. Trichosanthin, a potent HIV-1 inhibitor, can cleave supercoiled DNA in vitro. Nucleic Acids Res. 19(22):6309.

Liu B, Preisser EL, Shi X, Wu H, Li C, Xie W, Wang S, Wu Q, Zhang Y. 2017. Plant defence negates pathogen manipulation of vector behaviour. Funct Ecol. 31(8):1574–1581.

Liu J, Sharma A, Niewiara MJ, Singh R, Ming R, Yu Q. 2018. Papain-like cysteine proteases in Carica papaya: lineage-specific gene duplication and expansion. BMC Genomics. 19(1):26.

Louis J, Basu S, Varsani S, Castano-Duque L, Jiang V, Paul Williams W, Felton GW, Luthe DS. 2015. Ethylene contributes to maize insect resistance1-mediated maize defense against the phloem sap-sucking corn leaf aphid. Plant Physiol. 169(1):313–324.

Lu W, Tang X, Huo Y, Xu R, Qi S, Huang J, Zheng C, Wu C. 2012. Identification and characterization of fructose 1,6-bisphosphate aldolase genes in Arabidopsis reveal a gene family with diverse responses to abiotic stresses. Gene. 503(1):65–74.

Luan JB, Yao DM, Zhang T, Walling LL, Yang M, Wang YJ, Liu SS. 2013. Suppression of terpenoid synthesis in plants by a virus promotes its mutualism with vectors. Ecol Lett. 16(3):390–398.

Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, et al. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47(W1):W636–W641.

Major IT, Guo Q, Zhai J, Kapali G, Kramer DM, Howea GA. 2020. A Phytochrome B-independent pathway restricts growth at high levels of jasmonate defense. Plant Physiol. 183(2):733–749.

Maleck K, Levine A, Eulgem T, Morgan A, Schmid J, Lawton KA, Dangl JL, Dietrich RA. 2000. The transcriptome of Arabidopsis thaliana during systemic acquired resistance. Nat Genet. 26(4):403–410.

Maruyama-Nakashita A, Nakamura Y, Watanabe-Takahashi A, Inoue E, Yamaya T, Takahashi H. 2005. Identification of a novel cis-acting element conferring sulfur deficiency response in Arabidopsis roots. Plant J. 42(3):305–314.

Maxwell BB, Andersson CR, Poole DS, Kay SA, Chory J. 2003. HY5, Circadian Clock-Associated 1, and a cis-element, DET1 dark response element, mediate DET1 regulation of chlorophyll a/b-binding protein 2 expression. Plant Physiol. 133(4):1565–1577.

Medina-Puche L, Castelló MJ, Canet JV, Lamilla J, Colombo ML, Tornero P. 2017. β-carbonic anhydrases play a role in salicylic acid perception in Arabidopsis. PLoS One. 12(7):e0181820.

Ménard R, Khouri HE, Plouffe C, Dupras R, Ripoll D, Vernet T, Tessier DC, Laliberté F, Thomas DY, Storer AC. 1990. A protein engineering study of the role of aspartate 158 in the catalytic mechanism of papain. Biochemistry. 29(28):6706–6713.

Misas-Villamil JC, van der Hoorn RAL, Doehlemann G. 2016. Papain-like cysteine proteases as hubs in plant immunity. New Phytol. 212(4):902–907.

Mohan S, Ma PWK, Williams WP, Luthe DS. 2008. A naturally occurring plant cysteine protease possesses remarkable toxicity against insect pests and synergizes Bacillus thuringiensis toxin. PLoS One. 3(3):e1786.

Mohanty B, Krishnan SPT, Swarup S, Bajic VB. 2005. Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. Ann Bot. 96(4):669–681.

Monzingo AF, Collins EJ, Ernst SR, Irvin JD, Robertus JD. 1993. The 2.5 A structure of pokeweed antiviral protein. J Mol Biol. 233(4):705–715.

Moon YH, Song SK, Choi KW, Lee JS. 1997. Expression of a cDNA encoding Phytolacca insularis antiviral protein confers virus resistance on transgenic potato plants. Mol Cells. 7(6):807–815.

Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2006. Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of Arabidopsis. Plant Mol Biol. 60(1):51–68.

Nanjo Y, Oka H, Ikarashi N, Kaneko K, Kitajima A, Mitsui T, Muñoz FJ, Rodríguez-López M, Baroja-Fernandez E, Pozueta-Romero J. 2006. Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-golgi to the chloroplast through the secretory pathway. Plant Cell. 18(10):2582–2592.

Neller KCM, Diaz CA, Platts AE, Hudak KA. 2019. De novo assembly of the pokeweed genome provides insight into pokeweed antiviral protein (PAP) gene expression. Front Plant Sci. 10:1002.

Neller KCM, Klenov A, Hudak KA. 2016. The pokeweed leaf mRNA transcriptome and its regulation by jasmonic acid. Front Plant Sci. 7:283.

Niggeweg R, Thurow C, Weigel R, Pfitzner U, Gatz C. 2000. Tobacco TGA factors differ with respect to interaction with NPR1, activation potential and DNA-binding properties. Plant Mol Biol. 42(5):775–788.

Niño MC, Kang KK, Cho YG. 2020. Genome-wide transcriptional response of papain-like cysteine protease-mediated resistance against Xanthomonas oryzae pv. oryzae in rice. Plant Cell Rep. 39(4):457–472.

Nintemann SJ, Vik D, Svozil J, Bak M, Baerenfaller K, Burow M, Halkier BA. 2017. Unravelling protein-protein interaction networks linked to aliphatic and indole glucosinolate biosynthetic pathways in Arabidopsis. Front Plant Sci. 8:2028.

Nomoto M, Skelly MJ, Itaya T, Mori T, Suzuki T, Matsushita T, Tokizawa M, Kuwata K, Mori H, Yamamoto YY, et al. 2021. Suppression of MYC transcription activators by the immune cofactor NPR1 fine-tunes plant immune responses. Cell Rep. 37(11):110125.

Nooren IMA, Thornton JM. 2003. Diversity of protein-protein interactions. EMBO J. 22(14):3486–3492.

Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S. 2003. Gibberellin biosynthesis and response during Arabidopsis seed germination. Plant Cell. 15(7):1591–1604.

Oh DH, Kwon CS, Sano H, Chung W Il, Koizumi N. 2003. Conservation between animals and plants of the cis-acting element involved in the unfolded protein response. Biochem Biophys Res Commun. 301(1):225–230.

Oh DH, Hong H, Lee SY, Yun DJ, Bohnert HJ, Dassanayake M. 2014. Genome structures and transcriptomes signify niche adaptation for the multiple-ion-tolerant extremophyte Schrenkiella parvula. Plant Physiol. 164(4):2123–2138.

Okada K, Abe H, Arimura GI. 2015. Jasmonates induce both defense responses and communication in monocotyledonous and dicotyledonous plants. Plant Cell Physiol. 56(1):16–27.

Olsnes S, Pihl A. 1973. Different biological properties of the two constituent peptide chains of ricin, a toxic protein inhibiting protein synthesis. Biochemistry. 12(16):3121–3126.

Pandey KC, Barkan DT, Sali A, Rosenthal PJ. 2009. Regulatory elements within the prodomain

of Falcipain-2, a cysteine protease of the malaria parasite Plasmodium falciparum. PLoS One. 4(5):e5694.

Parikh BA, Baykal U, Di R, Tumer NE. 2005. Evidence for retro-translocation of pokeweed antiviral protein from endoplasmic reticulum into cytosol and separation of its activity on ribosomes from its activity on capped RNA. Biochemistry. 44(7):2478–2490.

Parikh BA, Tortora A, Li XP, Tumer NE. 2008. Ricin inhibits activation of the unfolded protein response by preventing splicing of the HAC1 mRNA. J Biol Chem. 283(10):6145–6153.

Paulus JK, Kourelis J, Ramasubramanian S, Homma F, Godson A, Hörger AC, Hong TN, Krahn D, Carballo LO, Wang S, et al. 2020. Extracellular proteolytic cascade in tomato activates immune protease Rcr3. Proc Natl Acad Sci USA. 117(29):17409–17417.

Pérez-López E, Hossain MM, Wei Y, Todd CD, Bonham-Smith PC. 2021. A clubroot pathogen effector targets cruciferous cysteine proteases to suppress plant immunity. Virulence. 12(1):2327–2340.

Pesquet E, Zhang B, Gorzsás A, Puhakainen T, Serk H, Escamez S, Barbier O, Gerber L, Courtois-Moreau C, Alatalo E, et al. 2013. Non-cell-autonomous postmortem lignification of tracheary elements in Zinnia elegans. Plant Cell. 25(4):1314–1328.

Peumans WJ, Hao Q, Van Damme EJM. 2001. Ribosome-inactivating proteins from plants: more than RNA N-glycosidases? FASEB J. 15(9):1493–1506.

Picard D, Kao CC, Hudak KA. 2005. Pokeweed antiviral protein inhibits brome mosaic virus replication in plant cells. J Biol Chem. 280(20):20069–20075.

Planchais S, Perennes C, Glab N, Mironov V, Inzé D, Bergounioux C. 2002. Characterization of cis-acting element involved in cell cycle phase-independent activation of Arath;CycB1;1 transcription and identification of putative regulatory proteins. Plant Mol Biol. 50(1):109–125.

Polito L, Bortolotti M, Battelli MG, Calafato G, Bolognesi A. 2019. Ricin: An ancient story for a timeless plant toxin. Toxins (Basel). 11(6).

Polito L, Bortolotti M, Mercatelli D, Mancuso R, Baruzzi G, Faedi W, Bolognesi A. 2013. Protein synthesis inhibition activity by strawberry tissue protein extracts during plant life cycle and under biotic and abiotic stresses. Int J Mol Sci. 14(8):15532–15545.

Prestle J, Schönfelder M, Adam G, Mundry KW. 1992. Type 1 ribosome-inactivating proteins depurinate plant 25S rRNA without species specificity. Nucleic Acids Res. 20(12):3179–3182.

Pružinská A, Shindo T, Niessen S, Kaschani F, Tóth R, Millar AH, van der Hoorn RAL. 2017. Major Cys protease activities are not essential for senescence in individually darkened Arabidopsis leaves. BMC Plant Biol. 17(1):4.

Qin W, Ming-Xing H, Ying X, Xin-Shen Z, Fang C. 2005. Expression of a ribosome inactivating protein (curcin 2) in Jatropha curcas is induced by stress. J Biosci. 30(3):351–357.

Rajamohan F, Mao C, Uckun FM. 2001. Binding interactions between the active center cleft of recombinant pokeweed antiviral protein and the alpha-sarcin/ricin stem loop of ribosomal RNA. J Biol Chem. 276(26):24075–24081.

Rajamohan F, Venkatachalam TK, Irvin JD, Uckun FM. 1999. Pokeweed antiviral protein isoforms PAP-I, PAP-II, and PAP-III depurinate RNA of human immunodeficiency virus (HIV)-1. Biochem Biophys Res Commun. 260(2):453–458.

Rajendran SK, Lin IW, Chen MJ, Chen CY, Yeh KW. 2014. Differential activation of sporamin expression in response to abiotic mechanical wounding and biotic herbivore attack in the sweet potato. BMC Plant Biol. 14(1):112.

Rawlings ND, Bateman A. 2021. How to use the MEROPS database and website to help understand peptidase specificity. Protein Sci. 30(1):83–92.

Ready MP, Brownt DT, Robertus JD. 1986. Extracellular localization of pokeweed antiviral protein. Proc Natl Acad Sci USA. 83:5053-5056.

Remi Shih NJ, McDonald KA. 1997. Purification and characterization of chitinases from transformed callus suspension cultures of Trichosanthes kirilowii Maxim. J Ferment Bioeng. 84(1):28–34.

Richau KH, Kaschani F, Verdoes M, Pansuriya TC, Niessen S, Stüber K, Colby T, Overkleeft HS, Bogyo M, van der Hoorn RAL. 2012. Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics. Plant Physiol. 158(4):1583–1599.

Rombolá-Caldentey B, Rueda-Romero P, Iglesias-Fernández R, Carbonero P, Oñate-Sánchez L. 2014. Arabidopsis DELLA and two HD-ZIP transcription factors regulate GA signaling in the epidermis through the L1 Box cis-element. Plant Cell. 26(7):2905–2919.

Ron D, Walter P. 2007. Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol Cell Biol. 8(7):519–529.

Roux KJ, Kim DI, Raida M, Burke B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol. 196(6):801-10.

Rubio V, Linhares F, Solano R, Martín AC, Iglesias J, Leyva A, Paz-Ares J. 2001. A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. Genes Dev. 15(16):2122–2133.

Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, Somssich IE. 1996. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of

parsley PR1 genes. EMBO J. 15(20):5690–5700.

Sakai H, Aoyama T, Oka A. 2000. Arabidopsis ARR1 and ARR2 response regulators operate as transcriptional activators. Plant J. 24(6):703–711.

Savary BJ, Flores HE. 1994. Biosynthesis of defense-related proteins in transformed root cultures of Trichosanthes kirilowii Maxim. var japonicum (Kitam.). Plant Physiol. 106(3):1195–1204.

Schmelz EA, Engelberth J, Alborn HT, Tumlinson JH, Teal PEA. 2009. Phytohormone-based activity mapping of insect herbivore-produced elicitors. Proc Natl Acad Sci USA. 106(2):653–657.

Schrot J, Weng A, Melzig MF. 2015. Ribosome-inactivating and related proteins. Toxins (Basel). 7(5):1556–1615.

Schünmann PHD, Richardson AE, Vickers CE, Delhaize E. 2004. Promoter analysis of the barley Pht1;1 phosphate transporter gene identifies regions controlling root expression and responsiveness to phosphate deprivation. Plant Physiol. 136(4):4205–4214.

Sharma N, Park SW, Vepachedu R, Barbieri L, Ciani M, Stirpe F, Savary BJ, Vivanco JM. 2004. Isolation and Characterization of an RIP (Ribosome-Inactivating Protein)-Like Protein from Tobacco with Dual Enzymatic Activity. Plant Physiol. 134(1):171.

Simpson JC, Roberts LM, Römisch K, Davey J, Wolf DH, Lord JM. 1999. Ricin A chain utilises the endoplasmic reticulum-associated protein degradation pathway to enter the cytosol of yeast. FEBS Lett. 459(1):80–84.

Simpson SD, Nakashima K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2003. Two different novel cis-acting elements of erd1, a clpA homologous Arabidopsis gene function in induction by dehydration stress and dark-induced senescence. Plant J. 33(2):259–270.

Singh R, Liyanage R, Gupta C, Lay JO, Pereira A, Rojas CM. 2020. The Arabidopsis proteins AtNHR2A and AtNHR2B are multi-functional proteins integrating plant immunity with other biological processes. Front Plant Sci. 11:232.

Slaymaker DH, Navarre DA, Clark D, Del Pozo O, Martin GB, Klessig DF. 2002. The tobacco salicylic acid-binding protein 3 (SABP3) is the chloroplast carbonic anhydrase, which exhibits antioxidant activity and plays a role in the hypersensitive defense response. Proc Natl Acad Sci USA. 99(18):11640–11645.

Smaczniak C, Li N, Boeren S, America T, Van Dongen W, Goerdayal SS, De Vries S, Angenent GC, Kaufmann K. 2012. Proteomics-based identification of low-abundance signaling and regulatory protein complexes in native plant tissues. Nat Protoc. 7(12):2144–2158.

Solstad T, Hervik Bull V, Breivik L, Fladmark KE, Bull VH. 2008. Identification of protein

interaction partners by immunoprecipitation: Possible pitfalls and false positives. J Clin Exp Hepatol. 4(1):21-27.

Song SK, Choi Yeonhee, Moon YH, Kim SG, Choi Yang Do, Lee JS. 2000. Systemic induction of a Phytolacca insularis antiviral protein gene by mechanical wounding, jasmonic acid, and abscisic acid. Plant Mol Biol. 43(4):439–450.

Stirpe F, Bailey S, Miller SP, Bodley JW. 1988. Modification of ribosomal RNA by ribosome-inactivating proteins from plants. Nucleic Acids Res. 16(4):1349–1357.

Stirpe F, Barbieri L, Battelli MG, Soria M, Lappi DA. 1992. Ribosome-inactivating proteins from plants: present status and future prospects. Biotechnology (N Y). 10(4):405–412.

Stirpe F, Barbieri L, Gorini P, Valbonesi P, Bolognesi A, Polito L. 1996. Activities associated with the presence of ribosome-inactivating proteins increase in senescent and stressed leaves. FEBS Lett. 382(3):309–312.

Szewińska J, Simińska J, Bielawski W. 2016. The roles of cysteine proteases and phytocystatins in development and germination of cereal seeds. J Plant Physiol. 207:10–21.

Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, et al. 2021. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 49(D1):D605–D612.

Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Macia FA, Roberts K, Martin C. 1998. The AmMYB308 and AmMYB330 transcription factors from antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. Plant Cell. 10(2):135–154.

Tang W, Perry SE. 2003. Binding site selection for the plant MADS domain protein AGL15: an in vitro and in vivo study. J Biol Chem. 278(30):28154–28159.

Tatematsu K, Ward S, Leyser O, Kamiya Y, Nambara E. 2005. Identification of cis-elements that regulate gene expression during initiation of axillary bud outgrowth in Arabidopsis. Plant Physiol. 138(2):757–766.

Teplova AD, Serebryakova MV, Galiullina RA, Chichkova NV, Vartapetian AB. 2021. Identification of Phytaspase Interactors via the Proximity-Dependent Biotin-Based Identification Approach. Int J Mol Sci. 22(23):13123.

Tomlinson JA, Walker VM, Flewett TH, Barclay GR. 1974. The inhibition of infection by cucumber mosaic virus and influenza virus by extracts from Phytolacca americana. J Gen Virol. 22(2):225–232.

Tourlakis ME, Karran RA, Desouza L, Siu KWM, Hudak KA. 2010. Homodimerization of pokeweed antiviral protein as a mechanism to limit depurination of pokeweed ribosomes. Mol

Plant Pathol. 11(6):757–767.

Turek I, Irving H. 2021. Moonlighting proteins shine new light on molecular signaling niches. Int J Mol Sci. 22(3):1–22.

Ussery MA, Irvin JD, Hardesty B. 1977. Inhibition of poliovirus replication by a plant antiviral peptide. Ann N Y Acad Sci. 284(1):431–440.

van der Linde K, Mueller AN, Hemetsberger C, Kashani F, van der Hoorn RAL, Doehlemann G. 2012. The maize cystatin CC9 interacts with apoplastic cysteine proteases. Plant Signal Behav. 7(11):1397–1401.

Van Esse HP, Van't Klooster JW, Bolton MD, Yadeta KA, Van Baarlen P, Boeren S, Vervoort J, Dewit PJGM, Thomma BPHJ. 2008. The Cladosporium fulvum virulence protein Avr2 inhibits host proteases required for basal defense. Plant Cell. 20(7):1948–1963.

Vandepoele K, Vlieghe K, Florquin K, Hennig L, Beemster GTS, Gruissem W, Van De Peer Y, Inzé D, De Veylder L. 2005. Genome-wide identification of potential plant E2F target genes. Plant Physiol. 139(1):316–328

Varsani S, Basu S, Williams WP, Felton GW, Luthe DS, Louis J. 2016. Intraplant communication in maize contributes to defense against insects. Plant Signal Behav. 11(8):e1212800.

Vasil V, Marcotte WR, Rosenkrans L, Cocciolone SM, Vasil IK, Quatrano RS, McCarty DR. 1995. Overlap of Viviparous1 (VP1) and abscisic acid response elements in the Em promoter: G-box elements are sufficient but not necessary for VP1 transactivation. Plant Cell. 7(9):1511–1518.

Vidi PA, Kanwischer M, Baginsky S, Austin JR, Csucs G, Dörmann P, Kessler F, Bréhélin C. 2006. Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles. J Biol Chem. 281(16):11225–11234.

Villarejo A, Burén S, Larsson S, Déjardin A, Monné M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, et al. 2005. Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. Nat Cell Biol. 7(12):1124–1131.

Vivanco JM, Savary BJ, Flores HE. 1999. Characterization of two novel type I ribosome-inactivating proteins from the storage roots of the andean crop Mirabilis expansa. Plant Physiol. 119(4):1447–1456.

Wang D, Weaver ND, Kesarwani M, Dong X. 2005. Induction of protein secretory pathway is required for systemic acquired resistance. Science. 308(5724):1036–1040.

Wang F, Guo Z, Li H, Wang M, Onac E, Zhou J, Xia X, Shi K, Yu J, Zhou Y. 2016. Phytochrome A and B function antagonistically to regulate cold tolerance via abscisic acid-

dependent jasmonate signaling. Plant Physiol. 170(1):459–471.

Wang J, Song L, Gong X, Xu J, Li M. 2020. Functions of jasmonic acid in plant regulation and response to abiotic stress. Int J Mol Sci. 21(4):1446.

Wang M, Hudak KA. 2006. A novel interaction of pokeweed antiviral protein with translation initiation factors 4G and iso4G: a potential indirect mechanism to access viral RNAs. Nucleic Acids Res. 34(4):1174–1181.

Wang P, Zoubenko O, Tumer NE. 1998. Reduced toxicity and broad spectrum resistance to viral and fungal infection in transgenic plants expressing pokeweed antiviral protein II. Plant Mol Biol. 38(6):957–964.

Wang S, Xing R, Wang Yan, Shu H, Fu S, Huang J, Paulus JK, Schuster M, Saunders DGO, Win J, et al. 2021. Cleavage of a pathogen apoplastic protein by plant subtilases activates host immunity. New Phytol. 229(6):3424–3439.

Wang YQ, Feechan A, Yun BW, Shafiei R, Hofmann A, Taylor P, Xue P, Yang FQ, Xie ZS, Pallas JA, et al. 2009. S-nitrosylation of AtSABP3 antagonizes the expression of plant immunity. J Biol Chem. 284(4):2131–2137.

Welchen E, Gonzalez DH. 2006. Overrepresentation of elements recognized by TCP-domain transcription factors in the upstream regions of nuclear genes encoding components of the mitochondrial oxidative phosphorylation machinery. Plant Physiol. 141(2):540.

Wytynck P, Lambin J, Chen S, Asci SD, Verbeke I, De Zaeytijd J, Subramanyam K, Van Damme EJM. 2021. Effect of RIP overexpression on abiotic stress tolerance and development of rice. Int J Mol Sci. 22(3):1–23.

Yang T, Zhu L, Meng Y, Lv R, Zhou Z, Zhu L, Lin H, Xi D. 2018. Alpha-momorcharin enhances Tobacco mosaic virus resistance in tobacco NN by manipulating jasmonic acid-salicylic acid crosstalk. J Plant Physiol. 223:116–126.

Yeh KW, Chen JC, Lin MI, Chen YM, Lin CY. 1997. Functional activity of sporamin from sweet potato (Ipomoea batatas Lam.): a tuber storage protein with trypsin inhibitory activity. Plant Mol Biol. 33(3):565–570.

Ytterberg AJ, Peltier JB, Van Wijk KJ. 2006. Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. Plant Physiol. 140(3):984–997.

Yu D, Chen C, Chen Z. 2001. Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. Plant Cell. 13(7):1527–1540.

Yu M, Haslam DB. 2005. Shiga toxin is transported from the endoplasmic reticulum following interaction with the luminal chaperone HEDJ/ERdj3. Infect Immun. 73(4):2524–2532.

Zamyatnin AA. 2015. Plant proteases involved in regulated cell death. Biochemistry (Mosc). 80(13):1701–1715.

Zander M, Lewsey MG, Clark NM, Yin L, Bartlett A, Saldierna Guzmán JP, Hann E, Langford AE, Jow B, Wise A, et al. 2020. Integrated multi-omics framework of the plant response to jasmonic acid. Nat plants. 6(3):290–302.

Zhang B, Tremousaygue D, Denancé N, Van Esse HP, Hörger AC, Dabos P, Goffner D, Thomma BPHJ, Van Der Hoorn RAL, Tuominen H. 2014. PIRIN2 stabilizes cysteine protease XCP2 and increases susceptibility to the vascular pathogen Ralstonia solanacearum in Arabidopsis. Plant J. 79(6):1009–1019.

Zhang S, Xu Z, Sun H, Sun L, Shaban M, Yang X, Zhu L. 2019. Genome-wide identification of papain-like cysteine proteases in Gossypium hirsutum and functional characterization in response to Verticillium dahliae. Front Plant Sci. 10:134.

Zhang Y, Natale R, Domingues AP, Toleco MR, Siemiatkowska B, Fàbregas N, Fernie AR. 2019. Rapid identification of protein-protein interactions in plants. Curr Protoc Plant Biol. 4(4):e20099.

Zhao C, Johnson BJ, Kositsup B, Beers EP. 2000. Exploiting secondary growth in Arabidopsis. Construction of xylem and bark cDNA libraries and cloning of three xylem endopeptidases. Plant Physiol. 123(3):1185–1196.

Zhu F, Zhang P, Meng YF, Xu F, Zhang DW, Cheng J, Lin HH, Xi DH. 2013. Alpha-momorcharin, a RIP produced by bitter melon, enhances defense response in tobacco plants against diverse plant viruses and shows antifungal activity in vitro. Planta. 237(1):77–88.

Zhu F, Yuan S, Zhang ZW, Qian K, Feng JG, Yang YZ. 2016. Pokeweed antiviral protein (PAP) increases plant systemic resistance to Tobacco mosaic virus infection in Nicotiana benthamiana. Eur J Plant Pathol. 146(3):541–549.

Zhu F, Zhou YK, Ji ZL, Chen XR. 2018. The plant ribosome-inactivating proteins play important roles in defense against pathogens and insect pest attacks. Front Plant Sci. 9:146.

Zhu F, Zhu PX, Xu F, Che YP, Ma YM, Ji ZL. 2020. Alpha-momorcharin enhances Nicotiana benthamiana resistance to tobacco mosaic virus infection through modulation of reactive oxygen species. Mol Plant Pathol. 21(9):1212–1226.

Zoubenko O, Hudak K, Tumer NE. 2000. A non-toxic pokeweed antiviral protein mutant inhibits pathogen infection via a novel salicylic acid-independent pathway. Plant Mol Biol. 44(2):219–229.

Zybailov B, Rutschow H, Friso G, Rudella A, Sun Q, van Wijk KJ. 2008. Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. PLoS One. 3(4):e1994.

APPENDIX A: BIOID ATTEMPT

An initial attempt was made to map out the PAP interactome using a method called proximity-dependent biotin identification (BioID). BioID utilizes a biotin ligase called BirA that is mutated in its active site, resulting in the transfer of a biotin molecule to side chains of proximal proteins (Roux et al. 2012). When mutated BirA is expressed fused to a protein of interest in living cells, it will biotinylate any protein that comes into proximity of the POI, allowing pull down of the biotin-tagged potential protein interactors by streptavidin (SA) and identification using mass spectrometry.

Since BioID of the PAP interactome requires expression of a BirA-PAP fusion protein, constructs were cloned to express a BirA-PAP fusion protein under control of the PAP promoter, as well as BirA and PAP alone as controls (Figure S1). Both fusion protein orientations (PAP-BirA/BirA-PAP) were cloned to account for the possibility that BirA could interfere with PAP interactions, localization, or functionality, depending on the terminal end of PAP to which it is fused. Two sets of constructs were cloned: one ("-apo") was created to localize experimental proteins to the apoplast, by flanking coding sequences with the PAP N- and C- terminal sequences (Figure 2), and a second ("-cyto") without the N or C-terminal sequences, for localization of proteins to the cytosol. After agroinfiltration of constructs into *Nicotiana tabacum* plants, only PAP-apo and BirA-cyto protein expression was detectable by Western blot (results not shown); attempts to express a BirA-PAP fusion protein by agroinfiltration into tobacco were unsuccessful for both orientations. Strategies employed for capturing BirA-PAP/PAP-BirA fusion protein included increasing the agroinfiltration cell OD (0.5-1.5), an expression time course (1-5 dpi), probing with various antibodies (anti-PAP, anti-BirA, HRP-SA), replacing the native PAP promoter with the constitutive 35S promoter, a biotin gradient to increase biotinylation for HRP-SA detection (0.5-2 mM), and spraying plants with 5 mM jasmonic acid (results not shown), which has been shown to upregulate expression of the PAP promoter (Neller et al. 2019).

Suspected reasons for lack of BirA-apo expression include instability of the BirA protein in the apoplast; Teplova et al. (2021) reported instability of TurboID, a mutated BioID moiety, when expressed in the apoplast compared with the intracellular fraction. A second potential problem

for apoplast-localizing constructs was the presence of the PAP C-terminal propeptide sequence; as this sequence is known to be cleaved from the mature PAP protein, it was placed on the C-terminal end of recombinant protein sequences to prevent protein cleavage. However, function of the C-terminal sequence of the PAP CDS is unknown, therefore its presence may disrupt expression or protein folding in a non-native context (ie. attached to the C-terminus of BirA). The potential for removing this sequence and using only the PAP N-terminal signal sequence to target recombinant proteins to the apoplast was explored by creating a construct coding for only the N-terminal signal sequence and mature PAP coding sequence, leaving out the 29 amino acids which comprise the C-terminal sequence. However, agroinfiltration of this construct showed extremely low levels of PAP protein expression when compared with expression using the full PAP sequence (PAP-apo; results not shown), demonstrating the necessity of the C-terminal sequence for PAP expression and precluding its removal.

For cytosolic-localizing constructs, a toxic yellow leaf phenotype was seen in plants agroinfiltrated with either PAP-cyto or BirA-PAP-cyto constructs when compared with controls (results not shown). As previously published work has shown that transgenic PAP expression in *N. tabacum* is toxic, resulting in leaves with a mottled phenotype (Lodge et al. 1993), this indicated the potential that an active BirA-PAP fusion protein was being expressed in the cytosol, where ribosomal depurination resulted in cellular death before protein expression could be visualized by Western blot. Therefore, an agroinfiltration OD gradient with low ODs was performed (0.002-0.02) for BirA-PAP-cyto and PAP-cyto, a strategy used to capture expression of certain proteins before their sustained activity becomes toxic to the cell (Gookin and Assmann 2014). A band of ~30 kDa at OD 0.02 was detected using an anti-PAP antibody in both PAP and BirA-PAP infiltrated samples (results not shown), indicating that the BirA-PAP samples were expressing PAP alone. Since re-sequencing of the BirA-PAP construct confirmed expected gene insert in the BirA-PAP clone, this raised the possibility that BirA was being cleaved after translation of the fusion protein. Therefore, due to difficulties with expressing the BirA-PAP fusion protein and a corresponding set of controls, BioID was replaced with a co-IP assay in order to collect data for mapping of the PAP protein interactome.

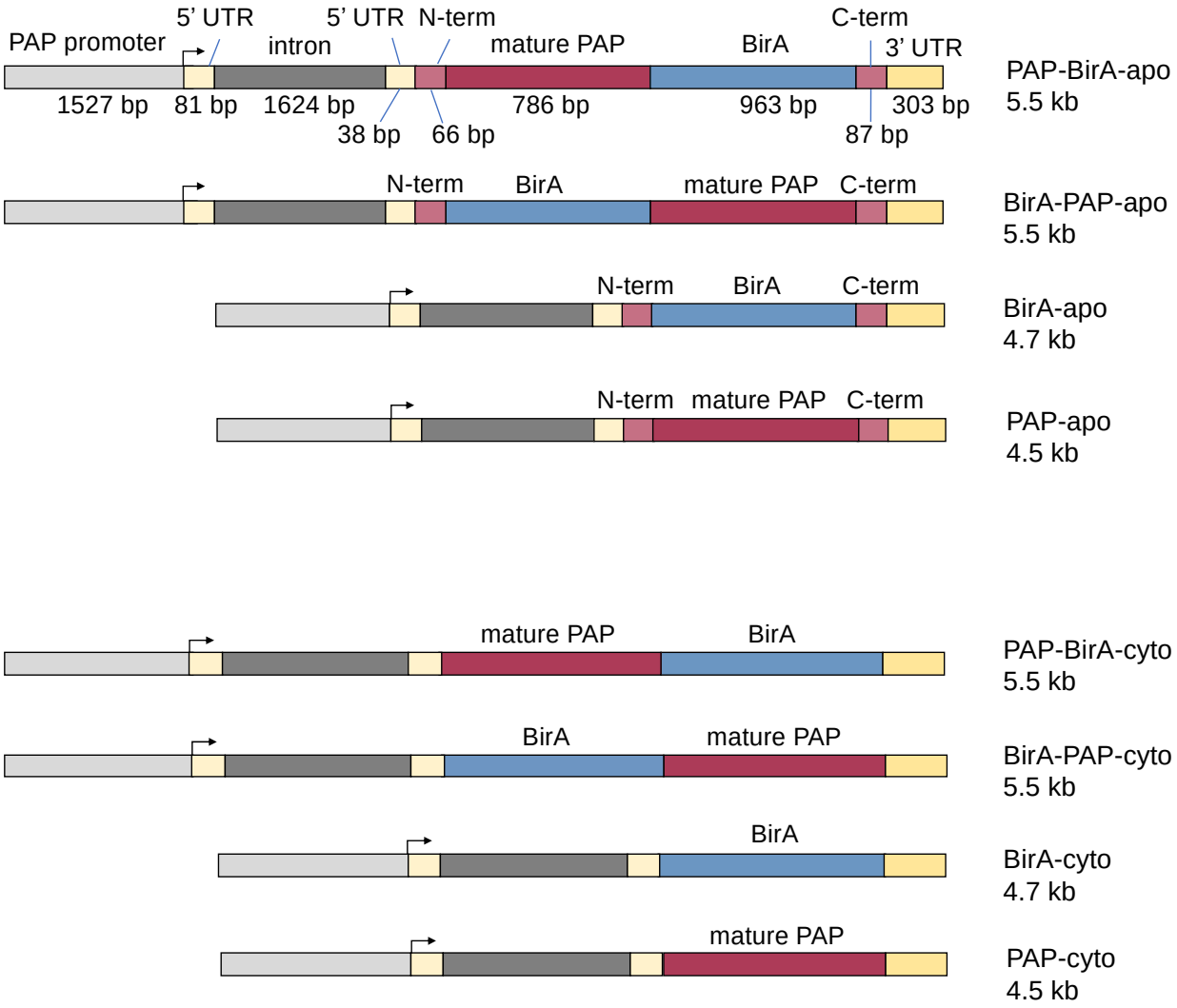


Figure S1. Schematic of constructs for agroinfiltration into tobacco plants for BioID assay. The constructs were inserted in place of the 35S promoter and GUSPlus reporter gene in the pCambia 0305.2 vector backbone. The PAP coding sequence (CDS) is shown in shades of pink. The mature PAP CDS is shown in darker pink, while the lighter pink sections represent the N-terminal and C-terminal signal sequences. These signal sequences are cleaved from the mature PAP protein and are included only for constructs expressing apoplast-localizing proteins. The +1 transcription start site is indicated by a bent arrow.

APPENDIX B: SUPPLEMENTARY TABLES

Table S1. List of pokeweed proteins identified by LC-MS/MS with associated protein scores (Score Sequest HT).

Gene ID/protein name	PAP IP replicate				FLAG IP replicate		
	1	2	3	4	1	2	3
PHYAM_020596-RA PAP-1 Antiviral protein I	2102.71	169.55	614.47	752.91			
PHYAM_006407-RC RCA2 Ribulose bisphosphate carboxylase/oxygenase activase 2, chloroplastic	2068.23	73.2	9.93	4.11	51.03		13.6
PHYAM_010842-RB GAPC Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	648.64	92.75		62.86	27.46		
PHYAM_013847-RA RBCS-1 Ribulose bisphosphate carboxylase small chain 1, chloroplastic	306.66			54.79			11.82
PHYAM_000095-RA GAPA Glyceraldehyde-3-phosphate dehydrogenase A, chloroplastic	549.62	105.01	28.22	58.36	105.68	11.41	43.96
PHYAM_003256-RC rps5 30S ribosomal protein S5, chloroplastic	216.08			157.87		15.66	165.29
PHYAM_025849-RA RPS17 30S ribosomal protein S17, chloroplastic	94.84			19.38			20.28
PHYAM_025507-RB 40S ribosomal protein S14	85.91	2.46		13.71			14.66
PHYAM_020806-RA RBCS-1 Ribulose bisphosphate carboxylase small chain 1, chloroplastic	157.61	6.58		52.84		5.98	24.28
PHYAM_018446-RA RBCS-1 Ribulose bisphosphate carboxylase small chain 1, chloroplastic	208.5						
PHYAM_010841-RB GAPC Glyceraldehyde-3-phosphate dehydrogenase, cytosolic	332.26						54.5
PHYAM_013566-RA RPL23A 60S ribosomal protein L23	46.13						
PHYAM_009893-RA TUBB6 Tubulin beta-6 chain	142.51						
PHYAM_026164-RA XCP1 Cysteine protease XCP1	179.92			8.21			
PHYAM_007276-RA IWF1' Non-specific lipid-transfer protein	37.01			24.99		1.88	3.74
PHYAM_018630-RA ACT7 Actin-7	150.19			16.58	5.6		33.14
PHYAM_025788-RA atpB ATP synthase subunit beta, chloroplastic	83.43						
PHYAM_003326-RB REFA1 Elongation factor 1-alpha	118.43						
PHYAM_010495-RA atpB ATP synthase subunit beta, chloroplastic	100.19	9.95		14.23		4.18	22.85
PHYAM_023306-RA RPL9 60S ribosomal protein L9	34.84						
PHYAM_010893-RA RPL9 60S ribosomal protein L9	36.89						
PHYAM_026431-RA LHCB5 Chlorophyll a-b binding protein CP26, chloroplastic	20.67			7.57			6.98
PHYAM_025787-RA rbcL Ribulose bisphosphate carboxylase large chain	52.15	3.02		13.4		9.7	18.61

PHYAM_004949-RA atpB ATP synthase subunit beta, chloroplastic	32.59	2.06		6.5			4.16
PHYAM_005922-RC GAPB Glyceraldehyde-3-phosphate dehydrogenase B, chloroplastic	167.49						
PHYAM_018582-RB GAPB Glyceraldehyde-3-phosphate dehydrogenase B, chloroplastic	164.42	88.07					
PHYAM_012532-RA RPS5A 40S ribosomal protein S5-1	26.76						2.01
PHYAM_015113-RA Glycine-rich RNA-binding protein	17.71			11.49		8.76	4.56
PHYAM_000542-RB Peroxisomal (S)-2-hydroxy-acid oxidase	30.77			19.38			10.74
PHYAM_028828-RA rps2 30S ribosomal protein S2, chloroplastic	57.84						
PHYAM_006846-RA TUBA3 Tubulin alpha-3 chain	52.91						
PHYAM_007306-RA RPS5A 40S ribosomal protein S5-1	54.16						
PHYAM_028184-RA PAP-2 Antiviral protein 2	50.84		12.37	12.51			
PHYAM_026783-RA PAP6 Plastid-lipid-associated protein 6, chloroplastic	25.88						15.08
PHYAM_027772-RB RPS26A 40S ribosomal protein S26-1	16.53	7.38					
PHYAM_002561-RC Carbonic anhydrase, chloroplastic	38.91			6.07			
PHYAM_026698-RA ATPC ATP synthase gamma chain, chloroplastic	37.38						
PHYAM_012933-RA CAB36 Chlorophyll a-b binding protein 36, chloroplastic	50.31		2.55	12.55			2.11
PHYAM_012451-RA Antiviral protein alpha (K-PAP)	16.28			26.27			
PHYAM_024627-RA rps21 30S ribosomal protein S21, chloroplastic	29.44			16.55			28.18
PHYAM_006623-RA RuBisCO large subunit-binding protein subunit beta, chloroplastic	35.24	5.57		25.37	3.7	5.75	31.34
PHYAM_023348-RA PGIP Polygalacturonase inhibitor	20.34						
PHYAM_020136-RA THI1 Thiamine thiazole synthase, chloroplastic	6.4						
PHYAM_006163-RB CHLD Magnesium-chelatase subunit ChlD, chloroplastic	34.25						
PHYAM_022377-RC RPL19 50S ribosomal protein L19, chloroplastic	7.39						2.37
PHYAM_000754-RA ALDP Fructose-bisphosphate aldolase, chloroplastic	13.24			20.81			
PHYAM_002072-RC FFC Signal recognition particle 54 kDa protein, chloroplastic	10.84						
PHYAM_008210-RA Elongation factor 1-alpha	125.42	18.13					
PHYAM_011950-RB AC97 Actin-97		4.75				5.74	
PHYAM_014095-RA RPS18A 40S ribosomal protein S18		2.67					
PHYAM_010902-RB RPL28A 60S ribosomal protein L28-1		2.87					

PHYAM_009658-RA rbcL Ribulose bisphosphate carboxylase large chain		7.55	15.38	26.16			14.78
PHYAM_013597-RA GAPC Glyceraldehyde-3-phosphate dehydrogenase, cytosolic		54.25					
PHYAM_010465-RA PAP-1 Antiviral protein I (novel PAP)		10.1	39.42	80.67			
PHYAM_012854-RA At3g47520 Malate dehydrogenase, chloroplastic			12.68				
PHYAM_017089-RB At2g35920 DExH-box ATP-dependent RNA helicase DExH1			4.81				
PHYAM_008460-RA PA200 Proteasome activator subunit 4			7.76				
PHYAM_023415-RA Protein of unknown function			5.2				
PHYAM_003082-RA RPM1 Disease resistance protein RPM1				10.84			
PHYAM_007955-RB RAD4 DNA repair protein RAD4				5.2			
PHYAM_011424-RA Glutamine synthetase leaf isozyme, chloroplastic				9.94		1.84	6.43
PHYAM_014929-RA MTERF1 Transcription termination factor MTEF1, chloroplastic				10.78			
PHYAM_017175-RA ATPA ATP synthase subunit alpha, mitochondrial				9.47		1.81	11.73
PHYAM_019896-RB PSBO Oxygen-evolving enhancer protein 1, chloroplastic				20.6			8.25
PHYAM_008582-RB RuBisCO large subunit-binding protein subunit alpha, chloroplastic				14.55		1.85	37.07
PHYAM_023921-RB RPL31 60S ribosomal protein L31				4.27			2.03
PHYAM_001854-RA Phosphoribulokinase, chloroplastic				8.32			6.06
PHYAM_016968-RA AZI1 pEARLI1-like lipid transfer protein 1							9.9
PHYAM_012619-RC FBA2 Fructose-bisphosphate aldolase 2, chloroplastic							9.48
PHYAM_008051-RA RPL22B 60S ribosomal protein L22-2							6.92
PHYAM_004942-RA psbA Photosystem II protein D1							2
PHYAM_024261-RC CSP41A Chloroplast stem-loop binding protein of 41 kDa a, chloroplastic							2.09
PHYAM_004546-RA PSAN Photosystem I reaction center subunit N, chloroplastic							2.55
PHYAM_004362-RC Protein of unknown function					20.3		
PHYAM_019463-RC REM4.1 Remorin 4.1					6.18	3.62	3.7

Table S2. Protein interactome map abbreviations.

	Abbreviation	Protein name	Accession number
Cluster 1	FBA	Fructose-bisphosphate aldolase 1, chloroplastic	XP_010690763.1
	GAPB	Glyceraldehyde-3-phosphate dehydrogenase B	XP_010686856.1
	RpS26-2	40S ribosomal protein S26-2-like	XP_010670630.1
	EF-1 alpha	Elongation factor 1-alpha	XP_010669662.1
	RpS3-1	40S ribosomal protein S3-1-like	XP_010667658.1
	PGK	Phosphoglycerate kinase, chloroplastic	XP_010669994.1
	RpS21	40S ribosomal protein S21-like	XP_010674310.1
	RpS27a	ubiquitin-40S ribosomal protein S27a	XP_010675161.1
	EF-1 beta gamma	Elongation factor 1-gamma 2-like	XP_010683294.1
	RpL30	60S ribosomal protein L30	XP_010685518.1
	EF-1 gamma 2	Elongation factor 1-gamma-like	XP_010686883.1
	RpL11-1	60S ribosomal protein L11-1 isoform X1	XP_010689741.1
	RpS25	40S ribosomal protein S25	XP_010690848.1
	RpS7	40S ribosomal protein S7	XP_010693181.1
Cluster 2	CA	Carbonic anhydrase, chloroplastic-like	XP_010687996.1
	CYN	Cyanate hydratase	XP_010670973.1
	Alpha-CA 1	Alpha carbonic anhydrase 1, chloroplastic	XP_010683723.1
	Alpha-CA 7	Alpha carbonic anhydrase 7-like	XP_010696305.1
	Beta-CA 5-X1	Beta carbonic anhydrase 5, chloroplastic isoform X1	XP_010691408.1
	CA 2-X1	Carbonic anhydrase 2 isoform X1	XP_010682554.1
Cluster 3	XCP2	Xylem cysteine proteinase 2	XP_010690568.1
	PDI	Protein disulfide isomerase-like 2-3	XP_010688926.1
	NAC076	Nac domain-containing protein 30; NAC domain-containing protein 76	XP_010678128.1
	CALR	Calreticulin isoform X1	XP_010674177.1
	MYB308	Transcription factor myb, plant; Myb-related protein 308	XP_010676421.1
	Ero1-X1	Endoplasmic reticulum oxidoreductin-1-like isoform X1	XP_010672064.1
	BiP	Endoplasmic reticulum chaperone bip	XP_010674036.1
	ERDJ3B	Dnaj homolog subfamily b member 11; dnaJ protein ERDJ3B	XP_010679350.1

	HSP70	Heat shock 70 kDa protein 17	XP_010679465.1
	HSP90	Heat shock protein 90kda beta; Endoplasmin homolog	XP_010691883.1
	Ero1	Endoplasmic reticulum oxidoreductin-1	XP_010694837.1
	CALR	Calnexin homolog	XP_010695743.1

Table S3. BlastP results for proteins with high homology to paCP1 amino acid sequence.

	Description	Scientific Name	Total Score	Query Cover	E-value	Per. ident	Acc. Len	Accession
1	Cysteine protease XCP1	*Arabidopsis thaliana*	387	99%	2.00E-134	52.96	355	O65493.1
2	Cysteine protease XCP2	*Arabidopsis thaliana*	369	92%	2.00E-127	54.27	356	Q9LM66.2
3	Chymopapain	*Carica papaya*	320	100%	4.00E-108	47.74	352	P14080.2
4	Papain	*Carica papaya*	310	98%	2.00E-104	48.56	345	P00784.1
5	Caricain	*Carica papaya*	309	98%	5.00E-104	48.28	348	P10056.2
6	Papaya proteinase 4	*Carica papaya*	306	97%	1.00E-102	49	348	P05994.3
7	Probable cysteine protease RD21B	*Arabidopsis thaliana*	300	98%	5.00E-99	45.4	463	Q9FMH8.1
8	KDEL-tailed cysteine endopeptidase CEP2	*Arabidopsis thaliana*	291	98%	6.00E-97	45.45	361	Q9STL4.1
9	KDEL-tailed cysteine endopeptidase CEP1	*Arabidopsis thaliana*	288	96%	2.00E-95	44.48	361	Q9FGR9.1
10	KDEL-tailed cysteine endopeptidase CEP3	*Arabidopsis thaliana*	287	97%	3.00E-95	44.57	364	Q9STL5.1
11	Probable cysteine protease RDL6	*Arabidopsis thaliana*	285	96%	2.00E-94	45.51	356	F4JNL3.1
12	Vignain	*Phaseolus vulgaris*	285	98%	4.00E-94	44.29	362	P25803.2
13	Vignain	*Vigna mungo*	284	90%	5.00E-94	46.13	362	P12412.1
14	Senescence-specific cysteine protease SAG39	*Oryza sativa Indica Group*	283	98%	7.00E-94	40.86	339	A2XQE8.1
15	Senescence-specific cysteine protease SAG39	*Oryza sativa Japonica Group*	283	98%	1.00E-93	40.86	339	Q7XWK5.2
16	Oryzain alpha chain	*Oryza sativa Japonica Group*	286	92%	1.00E-93	45.18	458	P25776.2
17	Cysteine proteinase RD21A	*Arabidopsis thaliana*	286	96%	3.00E-93	43.3	462	P43297.1
18	Probable cysteine protease RDL4	*Arabidopsis thaliana*	279	95%	7.00E-92	44.35	364	Q9SUT0.1
19	Cysteine proteinase COT44	*Brassica napus*	276	86%	2.00E-91	45.89	328	P25251.1
20	Thiol protease SEN102	*Hemerocallis hybrid cultivar*	277	95%	3.00E-91	44.77	360	P43156.1
21	Probable cysteine protease RDL5	*Arabidopsis thaliana*	276	95%	6.00E-91	42.9	371	Q9SUS9.1

22	Senescence-specific cysteine protease SAG12	*Arabidopsis thaliana*	268	95%	6.00E-88	41.52	346	Q9FJ47.1
23	Probable cysteine protease RD21C	*Arabidopsis thaliana*	271	97%	1.00E-87	42.78	452	Q9LT78.1
24	Probable cysteine protease RDL2	*Arabidopsis thaliana*	266	86%	8.00E-87	44.55	362	Q9LT77.1
25	Germination-specific cysteine protease 1	*Arabidopsis thaliana*	266	88%	1.00E-86	44.14	376	Q94B08.2
26	Cysteine proteinase EP-B 1	*Hordeum vulgare*	263	92%	2.00E-85	43.37	371	P25249.1
27	Vignain	*Ricinus communis*	261	87%	3.00E-85	45.51	360	O65039.1
28	Cysteine proteinase EP-B 2	*Hordeum vulgare*	262	92%	3.00E-85	43.37	373	P25250.1
29	Oryzain beta chain	*Oryza sativa Japonica Group*	261	81%	7.00E-84	44.9	466	P25777.2
30	Actinidain	*Actinidia deliciosa*	258	88%	1.00E-83	44.58	380	A5HII1.1
31	Actinidain	*Actinidia chinensis var. chinensis*	254	88%	3.00E-82	44.27	380	P00785.5
32	Cysteine protease 1	*Oryza sativa Japonica Group*	256	81%	3.00E-81	46.62	490	Q7XR52.2
33	Fruit bromelain	*Ananas comosus*	249	89%	2.00E-80	39.81	351	O23791.1
34	Cysteine protease Amb a 11.0101	*Ambrosia artemisiifolia*	248	96%	2.00E-79	40.11	386	V5LU01.1
35	Ananain	*Ananas comosus*	246	95%	2.00E-79	38.42	345	P80884.2
36	Zingipain-2	*Zingiber officinale*	233	61%	5.00E-76	51.6	221	P82474.1
37	Mexicain	*Jacaratia mexicana*	221	61%	2.00E-71	51.6	214	P84346.1
38	Zingipain-1	*Zingiber officinale*	221	61%	3.00E-71	50.46	221	P82473.1
39	Probable cysteine protease RDL3	*Arabidopsis thaliana*	225	90%	1.00E-70	42.42	376	Q9LXW3.1
40	Chymomexicain	*Jacaratia mexicana*	210	61%	3.00E-67	49.32	215	P84347.1
41	Ervatamin-B	*Tabernaemontana divaricata*	210	61%	4.00E-67	49.54	215	P60994.1
42	P34 probable thiol protease	*Glycine max*	209	91%	9.00E-65	36.69	379	P22895.1
43	Low-temperature-induced cysteine proteinase	*Solanum lycopersicum*	203	61%	9.00E-63	46.36	346	P20721.1
44	Ervatamin-C	*Tabernaemontana divaricata*	195	61%	2.00E-61	49.07	208	P83654.1
45	Stem bromelain	*Ananas comosus*	187	60%	5.00E-58	43.93	212	P14518.1
46	Cysteine proteinase 3	*Solanum lycopersicum*	191	99%	8.00E-58	37.5	356	Q40143.1
47	Macrodontain-1	*Ananas macrodontes*	183	60%	1.00E-56	42.52	213	P83443.1

48	Thiol protease aleurain-like	*Arabidopsis thaliana*	183	85%	7.00E-55	39.35	358	Q8RWQ9.1
49	Pro-cathepsin H	*Medicago truncatula*	180	96%	1.00E-53	36.49	350	A0A072UTP9.1
50	Thiol protease aleurain	*Arabidopsis thaliana*	177	85%	2.00E-52	37.3	358	Q8H166.2