2012

Characterization of PhPRP1, an extracellular arabinogalactan protein from Petunia hybrida pistils

Megan Twomey
Western Washington University

CHARACTERIZATION OF *Ph*PRP1, AN EXTRACELLULAR ARABINOGALACTAN PROTEIN FROM *PETUNIA HYBRIDA* PISTILS

By

Megan Twomey

Accepted in Partial Completion
Of the Requirements for the Degree
Master of Science

Dr. Kathleen L. Kitto, Dean of the Graduate School

ADVISORY COMMITTEE

Dr. Anu Singh-Cundy

Dr. Marion Brodhagen

Dr. Jeffery Young

MASTER'S THESIS

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the non-exclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Megan Twomey

May 2012

CHARACTERIZATION OF *Ph*PRP1, AN EXTRACELLULAR ARABINOGALACTAN PROTEIN FROM *PETUNIA HYBRIDA* PISTILS

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By

Megan Twomey
April 2012

ABSTRACT

An extracellular arabinogalactan protein, *Ph*PRP1, was identified from *Petunia hybrida* pistils on the basis of nucleotide sequence similarity with *Na*TTS from *Nicotiana alata* and TTS-1 from *Nicotiana tabacum*. *PhPRP1* exhibits 83% and 81% nucleotide identity with *NaTTS* and *TTS-1* cDNAs, respectively. The amino acid sequence deduced from the cDNA predicts a 27.4 kDa polypeptide backbone with a 25 amino acid signal sequence. Sequence alignments of *Ph*PRP1 and the TTS proteins reveals two hypervariable regions, including a proline-rich domain with noncontiguous KPP repeats. A short, highly-conserved histidine-rich domain separates the two hypervariable regions. The C-terminal segment of the protein shows significant sequence similarity to the Pollen Ole e 1 superfamily of arabinogalactan proteins, with perfect conservation of the six cysteine residues characteristic of the family. This multidomain architecture--especially the histidine-rich conserved domain and the proline-rich hypervariable region with KPP motifs--is a distinctive feature of TTS proteins and their orthologs in *Petunia hybrida* (*Ph*PRP1) and *Capsicum annuum* (*Ca*PRP1). An antibody specific to *Na*TTS detected *Ph*PRP1, and purified *Ph*PRP1 also reacted with *β*-glucosyl Yariv reagent, a dye that is diagnostic for arabinogalactans. BLAST searches identified sequences highly similar to *Ph*PRP1 in expressed sequence tag (EST) libraries prepared from *Petunia axillaris* pistils and *P. axillaris* primary roots, which suggests that *Ph*PRP1 is derived from the *P. axillaris* progenitor of *P. hybrida*, and that the PRP/TTS family of arabinogalactan proteins is not unique to pistil tissues. Phylogenetic analyses cluster the vegetative tissue ESTs with the pistil-expressed PRPs/TTS proteins at polytomies, indicating little differentiation among them. The gene geneology of PRPs/TTS proteins is

mostly concordant with recent interpretations of the phylogeny of the Solanaceae based on concatenation of chloroplast and nuclear genes, with some differences in topology depending on the evolutionary models employed to construct the gene tree. As key mediators of pollen-pistil interactions, PRP/TTS proteins may have an important role in generating reproductive barriers between sister species.

ACKNOWLEDGEMENTS

I thank Dr. Bruce McClure, University of Missouri, for providing the anti-TTS antibody. I thank Dr. Craig Moyer for help with DNA sequencing, and Dr. Sandra Schulze for help with immunoblot analysis. I am grateful to Dr. Jeffery Young for the EST identification of *Ph*PRP1, extensive guidance on bioinformatics and PCR troubleshooting, and the use of his thermocyclers. Dr. Marion Brodhagen offered valuable advice and much help, including kit reagents and lab space for the full length cloning of *PhPRP1*. Undergraduate researcher Mimosa Chau maintained plant stocks and helped with protein purification. This work was supported by a grant from the Office of Research and Sponsored Programs, Western Washington University.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

More than 70% of the calories consumed by humans worldwide come ultimately from fruit and seed, the products of plant reproduction. My thesis research focuses on understanding how the male and female structures--the pollen and the pistil--interact to bring about fertilization, resulting in seed. I have investigated pollen-pistil interactions in *Petunia hybrida*, a member of the Solanaceae (tomato and capsicum peppers also belong to this large and important plant family).

Previous work has identified a protein, called the transmitting tissue- specific (TTS) protein, from the pistils of the tobacco plant (*Nicotiana tabacum*) and demonstrated that the protein facilitates pollen tube growth (Cheung, 2000). I cloned and sequenced a cDNA encoding an ortholog of the protein from *Petunia hybrida* pistils and named it *PhPRP1*. I have delineated the biochemical hallmarks of this subfamily of arabinogalactan proteins and identified several distinctive domains. On the basis of domain architecture, I propose that a cDNA cloned from *Capsicum annuum* roots (Mang et al, 2004) and another from *Arabidopsis thaliana* vegetative tissues (Liu and Mehdy, 2007) are both orthologs of *PhPRP1*, as are several ESTs identified in GenBank based on sequence homology. Bioinformatics tools have identified EST sequences nearly identical to *PhPRP1* from the roots of *P. hybrida*, and based on this observation I propose that this subfamily of arabinogalactan proteins is not unique to reproductive tissue and that it has some unknown role in vegetative tissues.

Overview

The objectives of my research center on identifying and characterizing an ortholog of the TTS proteins, which have been described only from *Nicotiana* species (Cheung, 2000). Questions have been raised in the literature as to whether TTS proteins have any significance beyond the *Nicotiana* genus (reviewed by Hiscock and Allen, 2010). I chose to search for homologs of TTS proteins in another member of the Solanaceae, *Petunia hybrida*. *P. hybrida* is a diploid (2n=14) that is easily cultivated, produces numerous large flowers, and can be readily maintained through vegetative propagation. *Petunia* species have been used as model organisms for studying pollen-pistil interactions since the time of Charles Darwin (Darwin, 1876).

Using molecular tools and bioinformatics, a cDNA was isolated from *P. hybrida* pistils that is highly homologous to a cDNA encoding the transmitting tissue-specific (TTS) protein previously identified in *Nicotiana tabacum* pistils by Cheung *et al*. (1993). I began this research with a partial cDNA sequence from an expressed sequence tag (EST) library prepared from *P. hybrida* flowers by Shibuya *et al*. (2004). Aligning this partial sequence with the cDNA encoding the *Na*TTS protein from *Nicotiana alata* enabled me to identify conserved regions. I used these conserved sites to design RT-PCR (reverse transcription-polymerase chain reaction) primers. Through RT-PCR using total RNA extracted from mature *P. hybrida* pistils, I isolated a full length cDNA homologous to *NaPRP4* (encoding *Na*TTS) and named it *PhPRP1*.

Immunoblot analysis, using a primary antibody directed to an epitope that is almost perfectly conserved between *Na*TTS and *Ph*PRP1, revealed that *Ph*PRP1 has a molecular

mass ranging from about 55-100 kDa when protein extracts from the pistil are resolved by SDS-PAGE. *β*-glucosyl Yariv reagent, a diagnostic dye for arabinogalactans, stains the purified protein, confirming that *Ph*PRP1 is an arabinogalactan.

A multiple alignment of TTS proteins from *Nicotiana alata (Na*TTS) and *Nicotiana tabacum* (TTS), and *Ca*PRP1, a putative ortholog from *Capsicum annuum* (Mang et al., 2004), with *Ph*PRP1 revealed sequence similarity of 73%, 69%, and to 58%, respectively, and a distinctive domain architecture. All four proteins have a histidine-rich domain flanked by two hypervariable domains, HV1 and HV2. HV1 is short and lacks proline residues, while HV2 is rich in prolines and includes multiple KPP (lysine-proline-proline) motifs that differ in number and location in each pairwise comparison. All four proteins have a C-terminal *Ole e* I domain found in many proline-rich/hydroxyproline-rich glycoproteins classified as "nonclassical" arabinogalactans (Siefert and Roberts, 2007).

Phylogenetic analysis including cDNAs for *Ph*PRP1, *Nicotiana* TTS cDNAs, *CaPRP1,* and several ESTs in the Solanaceae mined from GenBank based on sequence homology were used to generate Maximum Likelihood and Neighbor Joining trees. The gene geneology of these PRP/TTS proteins was similar to phylogenies inferred from the chloroplast *ndhF* gene (Wu *et al.*, 2006) and *SAMT*, a nuclear gene (Martins *et al*., 2005). The main difference in the topology of this gene tree and the published molecular phylogenies of the Solanaceae is that *ndhF* and *SAMT*-based phylogenies cluster *Nicotiana* (subfamily Nicotianoidea) and the Solanoideae subfamily (tomato, *Capsicum*) into sister lineages, whereas my Maximum Likelihood tree of the known PRP/TTS proteins places the

Nicotiana and *Petunia* (subfamily Petunioideae) lineages at the same node as sister clades. The Neighbor Joining tree however is consistent with the published phylogenies.

Cheung at al. (1993) described the TTS proteins of *Nicotiana* as specific to the transmitting tissue, which forms the pollen tube pathway in the center of the pistil. However, BLAST searches of EST databases show that sequences virtually identical to *Ph*PRP1 have been found in the corolla tube and roots of *P. hybrida* (Breullin *et al.*, 2010). This is congruent with the report by Mang et al. (2004), who isolated *Ca*PRP1 from young roots of *Capsicum annuum* and found that the cDNA hybridizes to a sequence of the same size in RNA isolated from young leaves and pistils on RNA blots washed under stringent conditions. Thus, data mining suggests that the orthologs of *Ph*PRP1/TTS proteins are widespread in solanacous subfamilies (occurring in the Petunioideae and Solanoideae, in addition to the Nicotianoidea) and their expression is not confined to reproductive tissue. Clarifying the precise role of *Ph*PRP1 in facilitating pollen tube growth, its function in vegetative tissues, and its evolutionary history, must be left to future studies.

Pollen-pistil interactions

Composed of approximately 260,000 species in 450 families, angiosperms are the most diverse group and make up approximately 80% of all plants (Barnes-Svarney, *et al.*, 1999). The closed carpel, eight-celled female gametophyte, double fertilization, and monophyletic heritage are all defining characteristics of the angiosperm phylum (Molnar, 2001).

The angiosperm pollen grain is shed as a dormant, relatively dehydrated, two-celled or three-celled haploid structure (male gametophyte). A pollen grain alighting on the receptive surface of the pistil--the stigma--imbibes water, swells in size, and produces a cytoplasmic extension called the pollen tube. The pollen tube grows intercellularly through the transmitting tissue, which consists of vertical files of secretory cells in the central core of the pistil that serves as the pollen tube pathway (reviewed by Hiscock and Allen, 2008). Fertilization occurs when pollen tubes arriving in the ovules (protected within the ovary at the base of the pistil) burst open and the released sperm cells fuse with the egg cells.

Successful fertilization requires species-specific recognition between pollen and the tissues of the stigma and the style, which is the columnar portion of the pistil below the stigma and above the ovary (Figure 1). The growth of the pollen tube through these highly specialized pistil tissues requires extensive cross talk to ensure delivery of the fittest conspecific gametes to the eggs. Highly effective modes of pollen-pistil interactions, including mechanisms that ensure accelerated growth of the most competitive conspecific pollen tubes, may have contributed to the burst of adaptive radiation in angiosperms (Williams, 2008).

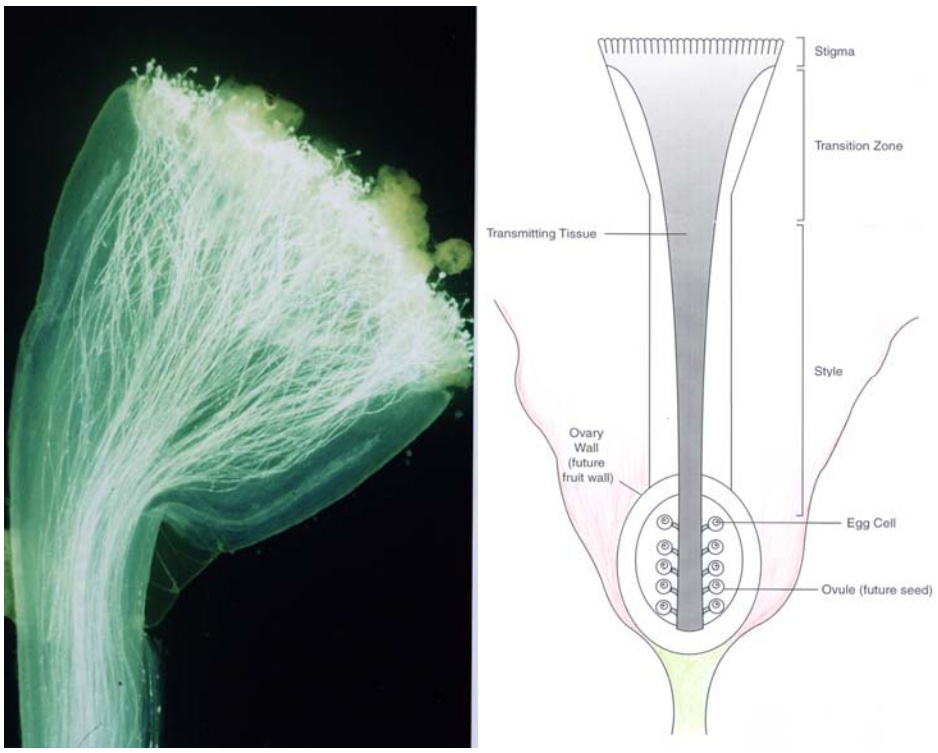


Figure 1. Fluorescence microscopic image of pollen tubes growing through a *Petunia* pistil. The pollen tubes were stained with aniline blue, a chromophore that fluoresces on binding to the callose (β-1,3-glucan) that lines pollen tube cell walls. A schematic sketch of the main parts of the solanaceous pistil is shown to the right of the photo (courtesy: Singh-Cundy lab).

Pollen Germination and Tube Growth

Pollination refers to the delivery of pollen to the stigmatic surface of a pistil. Pollination is followed by the germination of pollen grains on the stigma. In the great majority of flowering plants, mature pollen is shed as a bicellular gametophyte containing a tube nucleus and a generative cell. The generative cell divides to yield two sperm cells as pollen tubes grow through the pistil. In some families of plants, including the Brassicaceae to which the model plant *Arabidopsis* belongs, pollen is released from the anther in a tricellular state, already containing two sperm cells along with the tube nucleus.

In *Petunia hybrida*, as in all the known Solanaceae, bicellular pollen grains imbibe water and germinate on the copious exudates that cover the wet-type stigma (Heslop-Harrison and Shivanna, 1977). Pollen grains germinate in about 30 minutes to produce pollen tubes, which extend via tip growth. This polarized growth, localized exclusively to the apical dome at the tip of the tube, is facilitated by exocytotic delivery of membrane and cell wall materials to the apex, accompanied by endocytotic recovery of "excess" membrane (Zonia *et al.,* 2009). Pollen tubes enter the stigma and grow intercellularly within the transmitting tissue, reaching the transition zone of the style (see Figure 1) in about 6-8 hours.

Pollen tubes grow in the transmitting tissue through an extracellular matrix (ECM) that is rich in pectins, proteins, and nutrients such as sugars and amino acids (reviewed by Hiscock and Allen*,* 2008). Pollen tube growth in solanaceous pistils is biphasic: initially, pollen tubes grow through the transmitting tissue in an "autotrophic" manner, fueled mainly by storage reserves held in the pollen grain. Between six to eight hours after germination, pollen tubes enter the transition zone between the stigma and style (Lubliner *et al.,* 2003,

Mulcahy *et al*., 1982). The generative cell within the pollen tube divides, producing two sperm cells, and pollen tube growth rates increase as much as two to five fold (Herrero *et al.*, 1980; Lubliner *et al.,* 2003, Mulcahy *et al*., 1982). The pollen tube continues to elongate via tip growth, at a rate of about 1mm/h in *P. hybrida*, to traverse the length of the pistil to the ovule, where it enters the embryo sac and bursts, releasing the male gametes for double fertilization. The rapid second phase of growth appears to be heavily dependent on factors within the transmitting tissue, as suggested by the observation that these rates of growth cannot be matched *in vitro*. The most efficient *in vitro* culture media (containing PEG-6000, copper salts, and sucrose) reportedly produced no more than 40% of the *in vivo* growth rates (Read *et al*., 1993). Pollen tubes cultured *in vitro* grow randomly, in contrast to the strong directional growth displayed by *in vivo* pollen tubes. They also grow significantly shorter distances compared to pollen tubes growing within the pistil (reviewed by Higashiyama and Inatsugi, 2006).

Self-Incompatibility

Self-incompatibility is an outbreeding system that causes the pistil to reject pollen from the same plant or one that is genetically similar. The pistils of most solanaceous species exhibit gametophytic self-incompatibility (GSI) in addition to favoring the growth of conspecific over heterospecific pollen, and screening against less fit pollen genotypes. In GSI, the genotype of the haploid pollen and the genotype of the diploid female plant determine whether or not the growth of conspecific pollen tubes is blocked and fertilization thereby prevented.

GSI in the Solanaceae is controlled by a complex multi-gene locus, known as the S-locus (reviewed by Marshall *et al.,* 2011). Pollen bearing the same S-allele as the pistil is incompatible with that pistil, with the result that the pollen tubes it produces stop growing shortly after the transition to accelerated growth. On the pistil side, the S-locus encodes an extracellular ribonuclease, known as the S-RNase, with different alleles of the S-locus producing variant forms of the ribonuclease (Hua *et al.,* 2008). The S-RNase is endocytosed by pollen tubes as they grow through the ECM of the transmitting tissue. According to the cytotoxic model of GSI, the endocyotosed S-RNase destroys cytoplasmic RNA in pollen tubes that bear the same S-locus (have an identical S-haplotype), but not in pollen tubes with a different S-haplotype (Cruz-Garcia *et al.*, 2003; McCubbin *et al.*, 2000). S-RNases act in concert with other ECM proteins, including the 120 kDa glycoprotein and the HT family of extracellular proteins (reviewed by Kumar and McClure, 2010). Cruz-Garcia et al. (2005) have shown that RNases form complexes with other pollen and pistil proteins, including *Na*TTS, in tobacco.

The Pistil Extracellular Matrix (ECM)

The continuum of extracellular surfaces within the transmitting tissue of the ECM provides nutrition, guidance, structural support, protection, and mediation of cell-cell interactions between the pollen tube and the pistil (Roberts, 1994; Wu *et al*., 2001). These chemical and structural contributions are facilitated by an enrichment of the transmitting tissue ECM with secreted products, such as sugars, polysaccharides, amino acids,

glycoproteins, and glycolipids (Cassab, 1998; Lind *et al.,* 1994; Showalter, 1993; Wu *et al.,* 2001). The most extensively characterized group of ECM proteins is an assortment of proline-rich and hydroxyproline-rich glycoproteins (Cheung, *et al.*, 2001).

Traditionally, proline-rich proteins of plants have been grouped into four major classes. These include the extensins, hydroxyproline-rich glycoproteins (HRGPs), arabinogalactan proteins (AGPs), and solanaceous lectins. Each protein class is characterized by distinctive amino acid motifs that serve as targets for post-translational hydroxylation of proline residues followed by O-glycosylation. Proline residues that lie within these motifs are hydroxylated by prolyl 4-hydroxylases, rendering them targets for O-glycosyltransferases that attach specific sugar sidechains at the hydroxyl site (reviewed by Siefert and Roberts, 2007).

It is worth noting that numerous primary structure analyses have been performed on purified proline-rich proteins and they reveal that many of these proteins have domains and characteristics found in more than one of the classes mentioned above (Cheung *et al.,* 2001). For example, a 120kDa glycoprotein reported in *Nicotiana alata* has biochemical characteristics of both extensins and AGPs (Cassab, 1998; Lind *et al.,* 1994). In fact, Kieliszewski *et al.* (1992) suggest that the extensins, hydroxyproline-rich proteins, and AGPs form a continuum of related proline-rich molecules rather than discreet classes. With an increase in the discovery and characterization of cell wall and ECM proteins, variations and intermediate characteristics are blurring the distinction between these classes. This point is underscored in Table 1, in which I have summarized the characteristics of four well-characterized pistil ECM glycoproteins from *Nicotiana*.

Table 1. Summary of characteristics of four pistil arabinogalactan proteins (AGPs) found in *N. tabacum* and *N. alata*. Note the high degree of similarity in both backbone pIs and amino acid composition. Hydroxyproline content is determined through amino acid analysis, and because the direct information is not available in the literature for *Na*TTS and TTS-1 and TTS-2, I derived the values computationally.

Protein class and species	**Gene(s) (cDNA)**	**Pre-pollination Localization**	**Post-pollination Localization (Compatible)**	**Backbone properties**	**Most abundant amino acids (>7%)**	**Apparent Molecular Mass from SDS-PAGE**	**Carbohydrate Composition and Linkage (mol %)**	**Class Properties (Yariv reactivity) and structural features if known**	**Sources**
TTS; (*Nicotiana tabacum*)	*TTS-1* *TTS-2*	Pistil transmitting tissue, basipetal increase in glycosylation -Intercellular space of the TT	Pollen tube wall, absent from pollen tube cytoplasm	-Basic -Backbone pI ~10 -65% weight of total protein	Pro (17.2 %) Lys (12.1 %) Ser (8.6 %) Leu (8.2 %) Val (7.8 %) Thr (7.8 %) Ala (7.3 %)*	-Chemically deglycosylated: 28 kDa -Native: 45-105 kDa	Galactose (70%) 32% terminal 6% 1,3-linked 28% 1,6-linked 15% 1,2,3-linked 13% 1,3,6-linked Glucose (17%) 6-linked Galact. acid(6.5%) Arabinose (6%) 68% terminal, 15% 1,2-linked 15% 1,5-linked	-Arabinogalactan (AGP) (Yariv reactive) 11 Lys-Pro-Pro rpts TTS-2, 8 in TTS-1 6 C-term Cys res -Native pI ~7.5-9	Cheung, 1993,1995 Wu, 1995, 2000, 2001
NaTTS; (*Nicotiana alata)*	Na*PRP4*	Pistil transmitting tissue, basipetal increase in glycosylation	Pollen tube wall, absent from pollen tube cytoplasm	-Basic -Estimated backbone pI ~9.9*	Pro (18.5 %) Lys (13.3 %) Val (8.1 %) Ser (7.3 %)	-Predicted backbone: 27.5 kDa* -Native: 55-120 kDa	Uncharacterized	-AGP (Yariv reactive)	Chen, 1993 Wu, 2000 Hancock, 2005
120 kDa; (*Nicotiana alata*)	Na*PRP5*	8.8.% TT fluid, Style specific Distributed evenly over the extracellular matrix	Accumulation in cytoplasm and inner walls of pollen tubes, also in ECM concentrated adjacent to pollen tube surfaces	-Basic -Estimated backbone pI ~10* -65% weight of total protein	Hyp (21.4) Pro (12.0) Ser (7.5) Lys (7.3) Ala (6.9)	-Backbone: 78 kDa -Native: 80-120 kDa	Arabinose (55 %) 27% terminal 36% 1,2-linked Galactose (45%) 17% terminal, 8.5% 1,6-linked 3.2 % 1,3-linked 8.5% 1,3,6-linked	-Chimeric AGP: extensin and AGP characteristics and linkages (not Yariv reactive) (low Tyr, ~1.5%) -Native pI ~9	Lind, 1994, 1996 Shultz, 1997
PELPIII; (*Nicotiana tabacum*)	(p)*MG14* (p)*MG15*	Pistil transmitting tissue, Basipetal increase in glycosylation	Pollen tube wall, absent from pollen tube cytoplasm. Callosic layer and Callose plugs of pollen tubes	-Basic -35-50% weight of total protein	Hyp (18.9 mol%) Pro (13.5 mol%) Leu (9.0 mol%) Lys (8.8 mol%) Ser (7.6 mol%)	-Predicted backbone: 42-47 kDa -C-deglycosylated backbone: 2 at 50 and 60 kDa -Native: 110-140 kDa	Arabinose (48 %) 23% terminal 17% 1,2-linked Galactose (50 %) 16% terminal 6% 1,3-linked 12% 1,6-linked 16% 1,3,6-linked	-Chimeric AGP: (Yariv reactive) -Extensin-like (1,2 -linked Ara) on Ser -Pro moieties (low Tyr%, low Ser-Pro rpts)	Bosch, 2001 de Graff, 2003

*These values were computationally derived from reported cDNA and amino acid sequences when not provided in literature.

Plant Proline-Rich and Arabinogalactan Proteins

Hydroxyproline-Rich Glycoproteins

Hydroxyproline-rich glycoproteins (HRGPs) encompass a broad category of extracellular proteins containing variations of Pro-Pro repeats contained within a range of other larger repeat units (Cassab, 1998; Cheung *et al.*, 2001; Showalter, 1993). They are lightly glycosylated at hydroxylated proline residues, and they have been implicated in various aspects of plant growth and development. In addition, it was found that in some tissues wounding and pathogen attack induced HRGP expression. It is suspected that HRGPs are insolubilized in the cell wall over time, or in a rapid response to the above mentioned stresses (Fowler *et al.*, 1999; Showalter, 1993). This is indicated by a relatively high content of tyrosine that could form isodityrosine cross-links (Brady *et al.,* 1997; Showalter, 1993). It is believed that HRGPs may have important roles in normal development, in nodule formation, and in cell wall lignification (Showalter, 1993; Wu *et al.,* 2001; Ye *et al.*, 1991).

Arabinogalactan Proteins

Arabinogalactan proteins (AGPs) are proline rich proteins constituting a major and complex class of extracellular plant proteins characterized by extensive glycosylation with arabinose, galactose, or both. Their cell surface localization suggests involvement in functions such as cell-cell communication and interactions with the environment, as well as cell expansion (Siefert and Roberts, 2007). AGPs are highly glycosylated, with the protein backbone accounting for as little as 10-50% of the total molecular mass of the glycoprotein (Clarke *et al.*, 1979; Fincher *et al.*, 1983; Lord *et al.*, 1992).

AGPs tend to be rich in Ser, Ala, Thr, Gly, and post-translational Hyp (Cassab, 1998; Showalter, 1993). The carbohydrate groups are predominantly linked by O-glycosylation to the hydroxyl group of Ser and Hyp (Cassab, 1998). D-galactose and L-arabinose are the predominant carbohydrate moieties found in AGPs. The carbohydrate decorations are highly branched polysaccharides, comprising a (1-3) *β*-D-galactan backbone substituted with (1-6) *β*-D-galactan side chains at some galactosyl residues. Another characteristic feature is *β*-glucosyl Yariv reagent reactivity, a synthetic phenylglycoside that binds to and aggregates AGPs, though the mechanism by which *β*-glucosyl Yariv reagent selectively binds to AGPs non-covalently is not fully understood (Cassab, 1998; Cheung *et al.*, 1999; Showalter, 1993; Tang *et al.,* 2006). AGPs are chemically stable and proteolysis resistant in their native state, presumably a result of the presence of the bulky carbohydrate decorations (reviewed by Fincher *et al.,* 1983). Many AGPs contain a glycosylphosphatidylinositol (GPI) anchor that tethers them to the plasma membrane (Shultz *et al*., 2004).

AGPs are abundant in the stigmatic and transmitting tissue ECM (Cheung *et al.*, 1999; Ellis *et al.,* 2010; Showalter, 1993). As shown in Table 1, pistil AGPs display a great deal of diversity in the degree and pattern of O-glycosylation, stemming from variability in the type and length of the glycan chains attached to the protein backbone. Many of them, including all four shown in Table 1, possess a C-terminal *Ole e* 1 domain (Ellis *et al.,* 2008; Kieliszewski *et al.,* 2011; Lee *et al.,* 2008). The pollen Ole e I superfamily of arabinogalactans are major pollen allergens, named after a highly abundant secreted protein isolated from *Olea europeae* (olive tree) pollen by Villaba *et al*. (1993). Members of this AGP superfamily possess a C-terminal domain, known variably as the *Ole e* 1 domain (Jiang et al., 2005) or the PAC domain (Baldwin *et al*., 2001), and they lack a GPI anchor. As noted earlier, S-RNases bind to a number of pollen

and pistil glycoproteins to function in GSI, and Lee *et al*. (2008) have demonstrated that the 120kDa glycoprotein, a chimeric AGP-and-extensin glycoprotein, interacts with pollen proteins via its *Ole e* 1 domain.

Pistil AGPs such as the PELPIII proteins of *Nicotiana* are believed to play redundant architectural functions (Bosch *et al.*, 2002), while other AGPs, such as the TTS proteins (Cheung, 2000) and the 120kDa glycoprotein (Hancock *et al.*, 2005) are proposed to play an important role in pollen-pistil interactions. Antisense gene knockdowns have revealed that TTS proteins are critical for optimal pollen tube growth, and the proteins may exert that role by acting as cell surface adhesives and as substrates for pollen tube metabolism (Wu *et al,*, 2000).

The four pistil AGPs listed in Table 1 are considered "non-classical" AGPs because they lack the low pI value (~3) and high sugar content (up to 90% mol. mass) characteristic of "classical" AGPs (reviewed by Wu *et al*., 2001). These non-classical pistil AGPs can be further distinguished on the basis of their biochemical properties, such as reactivity with *β*-glucosyl Yariv reagent and carbohydrate linkages. Of the four types of pistil AGPs shown in Table 1, only the 120kDa glycoproteins fail to react with Yariv reagent.

TTS Proteins and Other Pistil Arabinogalactan Proteins

A few of the pistil-specific arabinogalactans have been shown to be specifically involved in the pollen-pistil crosstalk that takes place during germination and fertilization. These pistil AGPs can be split into two groups: those implicated in a role in supporting pollen tube growth and those that have an inhibitory role such as involvement in S-RNase-specific pollen rejection. Two widely studied AGPs have been clearly shown to influence pollen tube growth through

either support or inhibition, the TTS proteins and the 120 kDa glycoproteins, respectively (Lee, *et al*., 2008).

Cheung and colleagues (Cheung, *et al.*, 1993, 1999, Wu, *et al*., 2000, 2001) characterized AGPs expressed in the transmitting tissue of *Nicotiana tabacum* and named them TTS proteins, for transmitting-tissue specific proteins, because an immunoblot analysis failed to detect the proteins in extracts prepared from leaf, root, ovary, and petal tissue. They also described an ortholog, *Na*TTS, from *Nicotiana alata* (encoded by the cDNA *NaPRP4*). All three *Nicotiana* TTS proteins are highly soluble, Yariv-reactive arabinogalactans that are secreted into the ECM of stylar transmitting tissue. *Nicotiana tabacum* contains two cDNAs encoding TTS proteins, presumably because of its allotetraploid origins from two diploid ancestors, *N. sylvestris* and *N. tomentosiformis* (Fulnecek *et al.,* 2002).

In *Nicotiana* styles, there is an apical to basal increase in glycosylation of TTS, *Na*TTS, and another TTS homologue in *Nicotiana sylvestris,* with more highly glycosylated forms being found closer to the ovary (Cheung, *et al.*, 1999; Wu, *et al.*, 2000). The glycosylation of TTS decreases the backbone pI from ~10 to a range of 7.5-9, with a correlation between higher molecular weights and more acidic pIs (Wu, *et al.*, 1999, 2001). The highly adhesive nature of TTS proteins is evidenced by a tendency to oligomerize (Cheung *et al.*, 1995).

In *N. tabacum* and *N. alata*, TTS proteins have been shown to enhance pollen tube growth rates and attract pollen tubes from a distance when added to an *in vitro* medium lacking sucrose (Cheung, *et al.*, 1995; Wu, *et al.*, 2001). In addition, transgenic *N. tabacum* plants containing an antisense knockdown resulting in greatly reduced levels of TTS proteins had diminished pollen tube growth rates and reduced female fertility (Cheung *et al.*, 1995). TTS proteins have been shown to adhere to growing pollen tube tips and undergo deglycosylation by

pollen tube cell wall or membrane-bound enzymes. The adhesive nature of TTS proteins combined with the physical and biochemical interactions observed between these proteins and pollen tubes has led to the speculation that TTS proteins serve as both a nutrient source and surface adhesive for growing pollen tubes (Wu *et al.*, 2001). The gradient of increasing glycosylation through the length of the style may also imply a role in pollen tube growth directionality (Cheung *et al.*, 1995).

TTS proteins contain the conserved C-terminal *Ole e* 1 domain shown to interact with pollen proteins and speculated to contribute to their ability to interact with pollen tubes (Lee *et al*., 2008). TTS proteins have also been shown to form complexes with S-RNase glycoproteins and with pollen specific S- binding proteins, which bind to endocytosed S-RNase within the pollen tube (Cruz-Garcia *et al.*, 2003; Hancock, *et al.*, 2005; Hua, *et al.,* 2008; Lee *et al.*, 2008). Yeast two-hybrid analysis showed an interaction of *Na*TTS with three pollen-expressed proteins, including *Na*PCCP, a pollen-specific C2 domain-containing protein. The interaction with the pollen proteins is mediated by the non-glycosylated, highly conserved *Ole e* 1 domain in the C-terminal region of TTS proteins (Lee et al., 2008).While the exact role for TTS proteins in GSI is unknown, it has been hypothesized that TTS proteins are non-S-RNase factors that contribute to S-specific pollen rejection (Hancock, *et al.*, 2005; Lee *et al.*, 2008). S-specific pollen rejection requires the uptake of pistil S-RNase glycoproteins into the pollen tube, and TTS proteins have been shown to localize only on the pollen tube tip (Cheung, 1995, Wu *et al.*, 1995, 2001), suggesting an indirect role, if any, in GSI.

The 120 kDa glycoprotein from *Nicotiana* species is also highly soluble and found to exist in complexes with S-RNase, though in contrast to TTS proteins, the 120 kDa glycoprotein is taken up into the pollen tube cytoplasm (Cruz-Garcia *et al*., 2003; McClure *et al.,* 2011; Lind

et al., 1996). First described by Lind *et al.* 1994, the 120 kDa glycoprotein is a style-specific glycoprotein that has carbohydrate moieties and linkages typical of both extensins and AGPs (Table 1). Comprising approximately 8.8% of the protein content in the transmitting tissue, the 120 kDa glycoprotein, like TTS proteins, is developmentally regulated and widespread in the Solanaceae (Lind *et al.*, 1994). The 120 kDa glycoprotein also contains the conserved C-terminal *Ole e* 1 domain implicated in pollen-pistil interactions (Lee *et al.*, 2008). Unlike TTS proteins, the 120 kDa glycoprotein is uniformly glycosylated and, among Solanaceous species, shows a high degree of polymorphism arising mainly from a variable proline-rich region in the N-terminal portion of the protein. No basipetal gradient in the degree of glycosylation in the pistil ECM has been reported to date (Hancock *et al.*, 2005, Lind *et al.*, 1996).

The carbohydrate content of the 120 kDa glycoprotein is slightly greater in arabinose (55%) than galactose (45%). Interestingly, although this glycoprotein shares carbohydrate linkages with classical AGP (terminal arabinose linkages, 1,3,6; 1,6; 1,3; and terminal galactose linkages), it does not react with the AGP diagnostic β-glucosyl Yariv reagent (Lind *et al.*, 1994). The 120 kDa glycoprotein has a pI and protein:carbohydrate ratio similar to that of the extensins and solanaceous lectins, and it contains 1,2 and 1,4 arabinose linkages not found in other AGPs (Lind *et al.*, 1994).

The sequence variation reported in the 120 kDa glycoproteins of several *Nicotiana* species, and the interaction of this protein with S-RNases after endocystosis by growing pollen tubes, has made it a prime candidate for a non-S-RNase factor involved in S-locus-specific pollen rejection. RNAi experiments in which the expression of the120 kDa glycoprotein was suppressed showed a loss of GSI (Hancock *et al.*, 2005). Yeast two-hybrid analysis also showed that this glycoprotein interacts with pollen specific S-locus encoded F-box protein (the S-locus

component expressed in pollen tubes and believed to interact with S-RNases) via the C-terminal domain (Lee *et al.*, 2008). The authors speculate the glycoprotein may also be involved in metabolic pathways as pollen tubes need to metabolize pistil ECM materials for pollen tube growth.

Similar to TTS proteins and the 120 kDa glycoprotein, the Class III pistil-specific extensin-like protein (PELPIII) in *N. tabacum* also contain the conserved C-terminal binding domain in place of a GPI anchor. They display characteristics of both AGPs and extensins and are developmentally regulated (Bosch *et al.*, 2001; de Graaf *et al.*, 2003; Hancock *et al*., 2005; Mariani *et al.,*1992). Rich in proline/hydroxyproline, PELPIII has a basic protein backbone, 35-50% protein content, the presence of large AGP-like glucan moieties, and is *β*-glucosyl Yariv reactive, characterizing it as an AGP-extensin chimera (Table 1). Like TTS proteins and 120 kDa glycoproteins, PELPIII is developmentally regulated. Like TTS proteins, PELPIII displays an apical-basal glycosylation gradient and does not show a large amount of polymorphism across the Solanaceae (de Graaf *et al.*, 2003).

PELPIII is incorporated into the pollen tube wall and later deposited into the walls of growing pollen tubes. Unlike TTS proteins, the antisense suppression of PELPIII produced no change in pollen tube growth and pollination success, suggesting a likely structural role rather than a nutritive one (Bosch *et al.*, 2001; de Graaf *et al*., 2003). In spite of the significant similarities that TTS proteins, 120 kDa glycoproteins, and PELPIII share at the amino acid level, they have quite different types and patterns of post-translational glycosylation (de Graaf *et al.*, 2003), which indicates different functions for these presumed paralogs. A goal of this thesis project is to delineate the similarities and differences between these pistil-expressed arabinogalactans and a putative ortholog in *Petunia hybrida*.

METHODS

Plant material

P. hybrida (var. Tidal Wave Silver) plants were grown from seed under full spectrum lights at 23^{o}C and constant illumination. Flowers were emasculated at the late bud stage prior to anthesis to ensure that pistils harvested for analysis were unpollinated. Pistils were harvested from mature emasculated flowers that had reached the post-anthesis stage but before they showed any signs of floral senescence (approximately 1 to 4 days past anthesis). Harvested pistils were flash frozen in liquid nitrogen and stored at -80^{o}C prior to RNA extraction.

Synthesis of cDNA by reverse transcription

RNA was extracted from 40 pistils (approximately 144 mg fresh weight) according to the protocol from Ambion's TōTALLY RNA kit. RNA quality was determined using the denaturing gel electrophoresis protocol recommended by the kit manufacturers. The gel was stained with ethidium bromide to visualize the 28S and 18S bands as an indication of good RNA quality.

cDNA synthesis was performed using Invitrogen's ThermoScript RT-PCR System Plus Platinum Taq DNA Polymerase. The manufacturer's protocol was followed, using 2µg of RNA as template to extend an oligo(dT)$_{20}$ primer in each of six tubes yielding 20µl of product per tube. The reverse transcription reaction was carried out at 50^{o}C for 60 minutes, followed by enzyme inactivation at 85^{o}C for 5 minutes, in a Perkin Elmer GeneAmp PCR System 2400. RNA was hydrolyzed by incubating each reaction with 1µl RNase H at 37 ^{o}C for 20 minutes.

PCR amplification of cDNA

To amplify complete *Ph*PRP1 cDNA, a gene-specific forward primer 3Pr2 (5'-GGTTAATGGGTTGGTCATTGACCAAACATG-3') was designed using Primer3 (Rozen and Skaletsky, 2000) on the basis of a partial cDNA clone from an existing EST library of *P. hybrida* floral organs generated by Shibuya *et al.* 2004. Aligning the partial EST clone (GenBank Acc: CV295550; dbEST Id: 25722268) with cDNAs encoding *Na*TTS (*Nicotiana alata*) and TTS-1 (*Nicotiana tabacum)* revealed 83% and 81% sequence identity, respectively, for the sequence coverage available. This enabled the design of a forward primer expected to land upstream of the translational start site and encompassing some of the 5' noncoding sequence of the *P. hybrida* ortholog I was seeking. In designing this primer, parameters were modified to favor specificity and optimized for concentration and melting temperature (Tm). The reverse primer used was a 39-base polyT oligonucleotide with a degenerate base at the 3' position. Control reactions were performed in tandem using a sequence-specific reverse primer DnCtl (5'-GTTCAACTAGTGTCTTCTTTGT GTTGTTG-3') targeted against a conserved downstream region. Primers were ordered from Integrated DNA Technologies and delivered at a concentration of 100µM in IDTE Buffer pH 8.0 and diluted to a working concentration of 10µM.

The target cDNA encoding putative *Ph*PRP1 was amplified with Invitrogen Platinum Taq DNA Polymerase. Each 50 µl PCR reaction contained 2µl template, 1.5mM $MgCl_2$, 10µM dNTPs, and 0.1µM primers. Cycling parameters are a modified touchdown protocol outlined in Table 2. Two 5 cycle annealing steps at 70ºC and 65ºC, respectively, were found to increase specificity and decrease smearing due to nonspecific amplification.

Table 2. PCR cycling parameters

Temperature	Time	Cycles
94°C	2 minutes	1
94°C	30 seconds	5
70°C	1 minute	
94°C	30 seconds	5
65°C	1 minute	
94°C	30 seconds	25
60°C	30 seconds	
72°C	1 minute	
72°C	5 minutes	1
4°C	∞	

DNA purification and sequencing

PCR reactions were separated on a 1.2% agarose gel (FisherBiotech molecular biology grade) with tris-acetate EDTA running buffer. Gels were viewed using an Ultra-Lum Electronic UV Transilluminator at maximum UV intensity. Images were captured with a 10 megapixel Canon PowerShot A640 camera set to remote capture with macros on. Images were viewed, optimized for contrast and brightness, and stored in Canon Utilities ZoomBrowser EX version 5.7.

PCR reactions were purified prior to sequencing using a Millipore YM-50 Centrifugal Filter Unit followed by a modified protocol from Montage PCR Cleanup Kit (Millipore). To each ultrafiltered PCR reaction, 200 µl Tris (with 0.1 mM EDTA) was added, and then centrifuged for 5 min at 1000 x *g*. Samples were washed again with the addition of 500 µl TE (containing 0.1mM EDTA). The columns were inverted, rinsed with 50 µl ultrapure Optima H_2O (Fisher) and briefly centrifuged. DNA concentration in the samples was quantified with a NanoDrop ND-1000.

The cleaned-up cDNA template was prepared according to the ABI BigDye Terminator v3.1 Cycle Sequencing Kit. Two sequencing reactions were set up for each sample, one using 3Pr2 primer and the other using nested primer SeqP (5'-AACATGAAGACCACCTTCCAC-3'). Each 10µl sequencing reaction comprised 2µl Ready Reaction premix, 1 µl BigDye Sequencing Buffer, 1.5 µl of 3Pr2/SeqP primer (10 pmoles/µl), 2.5 µl and 5 µl of samples α and β respectively, with the final volume brought to 10 µl with Optima H_2O (Fisher). Cycle sequencing parameters are outlined in Table 3.

Table 3. Cycle sequencing parameters (ABI BigDye Terminator v3.1 Cycle Sequencing Kit). Rapid thermal ramp is 1°C/second.

Temperature	Time	Cycles
96°C	1 minute	1
96°C	10 seconds	25
50°C	5 seconds	
60°C	4 minutes	
4°C	∞	1

Extension products were purified using a Sephadex G-50 column and centrifuged at 750 x *g* for 5 min. Samples were dried in a 96-well plate under vacuum and resuspended in 15 µl formamide. Samples were denatured by heating to 95^{o}C for 5 minutes and were then sequenced by capillary electrophoresis using an ABI PRISM 3130xl Genetic Analyzer with a 36-cm capillary array using POP 6 from Applied Biosystems (courtesy of Dr. Craig Moyer). The sequence obtained covered the complete translated region.

PCR cloning

To obtain a full length sequence, PCR was repeated according to parameters previously outlined with the addition of 20 µg of T4 Gene 32 Protein from New England Biolabs to each reaction. T4 is a single-stranded DNA binding protein thought to stabilize DNA and increase amplification products (Jefferies *et al.,* 2002). PCR products were purified with the Montage kit, ligated into pCR4-TOPO vector (TOPO TA Cloning Kit for Sequencing, Invitrogen) and mobilized into DH5α strain of *E coli*. Plasmids were purified from single colonies inoculated on 'patch plates' using the QIAprep Spin Miniprep Kit (Qiagen). Plasmid preparations were quantified on the NanoDrop ND-1000 and screened for the presence of the cDNA insert via PCR using 3Pr2 and M13R primers, and restriction digestion with *Eco*RI (restriction sites adjacent to the insert). Restriction digests were set up in a 20 µl volume as follows:

- 5 µl plasmid DNA
- 2 µl 10x Buffer H (New England Biolabs)
- 2 µl *Eco*RI (12U/µl, New England Biolabs)
- 11 µl H_2O

Sequencing of the cDNA clone, named *PhPRP1*, was carried out at Nevada Genomics using the ABI BigDye Terminator v3.1 Cycle Sequencing Ready Reaction Kit on an ABI 3730 DNA Analyzer with 3Pr2 forward and M13 reverse primers.

Bioinformatics analysis

Electropherograms were imported into the sequence analysis software BioEdit Sequence Alignment Editor version 7.0.5.3 (Hall, 1999) for manual base calling and annotation. Several web-based softwares were used for further analysis and these are summarized in Table 4.

The full length sequence, *PhPRP1,* was submitted to GenBank (accession FJ719032). A cDNA identified in a library prepared from very young roots of *Capsicum annuum* (hot peppers) by Mang et al. (2004) and reported to share a high degree of sequence similarity to known TTS proteins was translated and included in the multiple sequence alignment I constructed.

Table 4. Web based programs used for bioinformatics and sequence analysis.

Web Based Program	Analysis Purpose	Citation
NCBI-Basic Local Alignment Search Tool (BLAST): blastn, blastp, blastx, tblastn	Nucleotide and amino acid sequence homology	Zhang *et al.*, 2000
ClustalW- Multiple Sequence Alignment	Sequence alignments	Thompson *et al.*, 1997
ExPASy translate tool	Nucleotide→amino acid, compute pI/MW	Hulo *et al.,* 2006
ExPASy ProtParam	Amino acid composition, compute pI/MW, hydropathicity	Gasteiger *et al.*, 2005
NetNGlyc 1.0 NetOGlyc 3.1	N-linked glycosylation site prediction O-linked glycosylation site prediction	Gupta *et al.*, 2004
SignalP 3.0	Signal peptide cleavage site identification	Bendtsen *et al.*, 2004
PATTINPROT	Scans peptide sequence for patterns	Sapay *et al.*, 2006
Simple Modular Architectural Research Tool (SMART)	Pfam protein domain search	Letunic *et al.*, 2006; Shultz *et al.*, 1998
TAIR- The Arabidopsis Information Resource	Putative gene function based on homology to the model *Arabidopsis*.	

Protein Analysis and Purification of *Ph*PRP1

Pistil extracts were prepared by bisecting 40 mature *P. hybrida* var. 'Tidal Wave Silver' pistils longitudinally and floating them in 500µl isotonic buffer containing 300 mM sucrose and 100mM NaCl in 40mM Tris-HCl, at pH 7.5 (Wang et al., 1993) for 30 min. This pistil eluate was clarified by centrifugation, then frozen at -20^{0}C or used immediately for protein analysis. Proteins in the pistil eluate (40 µl eluate per lane, with approximately 2 µg total protein) were separated on a discontinuous SDS-PAGE gel (8% stacking gel, 12% separating gel) and the protein bands visualized by silver staining.

For immunoblot analysis, 55 unpollinated pistils were bisected longitudinally and added to 500 µl of buffer in a 2ml spin column tube (Ambion catalog# 10065). 100 µl additional buffer was added after the bisections, and the spin columns were inverted gently several times and the pistil elution was incubated at room temperature for 40 minutes. The spin columns were centrifuged at 6000 rpm for one minute and the eluate that passed through the polypropylene mesh screen into the collection tube was stored at 4^{o}C or at -20^{o}C if not used immediately.

Protein concentrations were estimated using the Bradford assay (Coomassie Plus reagents from Pierce/ThermoScientific). An aliquot of pistil eluate, pistil extract, or purified *Ph*PRP1 (concentrations ranging from 50 ng-4 µg) was boiled for 10 min with denaturing loading buffer (containing 20mM mercaptoethanol and 4% sodium dodecyl sulfate) and separated by discontinuous SDS-PAGE on an 8% resolving gel. The proteins were visualized by silver staining or transferred to nitrocellulose membranes for staining with Yariv reagent or immunoblotting.

Immunoblot supplies, reagents, instruction, and lab space were kindly provided by Dr. Sandra Schultz. Antibodies specific to *Na*TTS were generously provided by Dr. Bruce McClure

at University of Missouri. Proteins were separated on a 12% polyacrylamide gel using a BioRad Mini-PROTEAN Tetra Cell gel running and transfer apparatus. Sample was loaded in a dilution series from undiluted, 1:1, 1:9, and 1:19. Proteins were transferred to a Whatman Protran nitrocellulose transfer membrane. Immunoblot protocol was followed according to Schultz, 2005. The membrane was incubated with rabbit IgG anti-*Na*TTS antibodies at a 1:6000 dilution. A stabilized goat anti-rabbit IgG horseradish peroxidase conjugated secondary antibody was used at a 1:1000 dilution (Pierce #32460). Pierce Super Signal West Pico Chemiluminescent Substrate (#34080) was used for detection and the signal was viewed using an Alpha Innotech Molecular Imager Fluor Chem HD2 and the AlphaEase FC Fluor Chem HD2 AIC imaging software. A 30 second exposure at high/medium setting was used to capture the image followed by a 5 minute exposure at super speed (lowest resolution, highest sensitivity) to ensure no cross reactivity. Bright field and chemiluminescent images were overlayed using Adobe Photoshop CS4.

*Ph*PRP1 was purified essentially as described by Wu et al. (2000), except for these modifications: we used batch processing instead of gravity-flow columns; we used cation exchange chromatography first, followed by affinity chromatography on Talon his-affinity beads (Clontech); 0.01% Tween-20 was used in the wash buffers to reduce contamination with S-RNases; proteins bound to affinity beads were eluted with 300mM imidazole (instead of EDTA); and, we used a desalting spin column (Pierce/Thermo Scientific) instead of size exclusion chromatography on Sephadex G100. We also used Dynabeads, a magnetized matrix with cobalt-based surface chemistry for binding polyhistidine (Dynal/Invitrogen), and found that these 1.1µm supraparamagnetic particles gave considerably higher yields of the purified protein than Talon beads did.

To visualize pistil arabinogalactan proteins and purified *Ph*PRP1, aliquots of pistil eluates (4 µg total protein) and purified *Ph*PRP1 (2 µg protein) were resolved by SDS-PAGE on an 8% resolving gel, and then transferred to a nitrocellulose membrane (0.2 µm) as described earlier. The protein blot was stained with 0.2% *β*-glucosyl Yariv reagent in 1% NaCl, and then destained by washing in 1% NaCl briefly.

Phylogenetic analysis

EST sequences with significant homology to *PhPRP1* were identified in GenBank. The accession numbers of the sequences used in analysis are listed in Table 5. An *Arabidopsis* cDNA encoding AGP31 (Liu and Mehdy, 2007) was identified as a homolog based on significant sequence similarity with *Ph*PRP1 and used as an outgroup. A Maximum Likelihood tree was generated using PAUP v 4.0 (Swofford, D.L. 2012). A full heuristic search was conducted with 100 bootstrap replicates. Empirical state frequencies were used, no assumed proportion of invariable sites, and equal distribution of rates at variable sites. The molecular clock was not enforced. Gaps were treated as missing, starting trees were obtained via stepwise addition and the TBR branch swapping algorithm was used. Branches were collapsed if the maximum branch length was zero. A strict consensus tree was generated from these results. Both the published *ndhF* phylogeny (Wu *et al.* 2006) and my consensus tree were re-drawn by hand using Mesquite version 2.75+ (Maddison and Maddison, 2011).

For Neighbor Joining analysis, sequences were realigned using ClustalX2 (Larkin *et al.*, 2007) and a Neighbor Joining tree was generated using default parameters (which included the Jukes-Cantor model) with 1000 bootstrap replicates.

Table 5. List of species and GenBank Accession numbers used in Parsimony.

Species	Tissue	Sequence Source	Sequence size (nt)	Gene or clone	Accession ID
Petunia hybrida	Pistil	cDNA	964	*PhPRP1*	FJ719032
Petunia hybrida	Root	EST	552	dr001P0011N09	FN003773
Petunia axillaris	Corolla	EST	749	dr004P0024C09	FN015061
Nicotiana tabacum	Pistil	cDNA	935	*tts-1*	Z16403
Nicotiana alata	Pistil	cDNA	915	*NaPRP4*	X70441
Capsicum annum	Root	cDNA	1096	*CaPRP1*	AY533017
Solanum lycopersicum	Ovary	EST	628	cLED15J14	AI489839
Solanum lycopersicum	Shoot	EST	749	cTOF5F24 5-	BG124526
Solanum lycopersicum	Leaf	EST	741	LEFL1012AF12 5-	DB681959
Solanum tuberosum	Tuber	EST	722	21243 5-	DN906833
Solanum melogena	Ovary	EST	666	OVS01L05 3-	FS029950
Arabidopsis thaliana	Seedling, pistil	cDNA	1186	At1g28290	NM_001084147

RESULTS

The cDNA and deduced amino acid sequences of *Ph*PRP1

A 964 bp cDNA product, named *P. hybrida* Proline-Rich Protein 1, or *PhPRP1*, was cloned and sequenced. The sequence analysis revealed a 27 bp 5'-untranslated region, a 768 bp coding region, and a 169 bp 3'-untranslated region. The 3' end terminates at the poly-A tail confirming that the complete cDNA was obtained.

The deduced amino acid sequence of *Ph*PRP1 reveals a consensus signal peptide sequence identified in the open reading frame at nucleic acid positions 24 to 32 (underlined, Fig.2). This signal sequence is nearly identical to the consensus sequence identified in plants (Lutcke *et al.*, 1987), with the AACAAUGGC consensus appearing as AATAAUGGC in *Ph*PRP1.

Biochemical characteristics of *Ph*PRP1

The open reading frame deduced from the *PhPRP1* cDNA encodes a protein 256 amino acids in length. A 12-amino acid motif of *Ph*PRP1 (AsnAsnThrLysLysThrLeuValGluGln GlyThr) that is strongly conserved in all the published TTS proteins, and to a lesser degree in *Ca*PRP1, is marked in Figure 2. A signal sequence prediction and a hydropathy plot (Figure 3, A and B respectively), and sequence comparison with known TTS proteins, collectively reveal a 25-amino acid putative signal peptide that is nearly identical to the secretion peptide of the TTS proteins from *N. tabacum* (Cheung *et al*., 1993)*, Na*TTS from *N. alata* (Wu *et al*., 1995), and *Ca*PRP1 from *Capsicum annuum* (Mang *et al*., 2004). The calculated molecular mass of *Ph*PRP1 is 27.4 kDa. The predicted pI is 9.96.

```
    5         15        25        35        45        55        65        75        85
                           MetAlaLysAlaPheValLeuPheHisLeuSerValLeuLeuLeuSerSerPheThrValLeu
GTTCAGCACAATTAGTACTTAGCAATAATGGCAAAGGCCTTTGTTCTTTTTCATCTTTCAGTTTTATTACTCAGCTCATTCACAGTTCTT

    95        105       115       125       135       145       155       165       175
SerHisGlyGluGlyLeuMetGlyTrpSerLeuThrLysHisGluAspHisLeuProProAlaGlnAlaProLysProHisLysGlyHis
AGCCATGGTGAAGGGTTAATGGGTTGGTCATTGACCAAACATGAAGACCACCTTCCACCAGCTCAAGCCCCAAAGCCTCACAAGGGCCAC

   185       195       205       215       225       235       245       255       265
HisHisProLysHisSerProAlaProSerProAlaThrProProProAlaTyrSerProSerLysProProValLysProProThrPro
CACCATCCAAAACATTCCCCAGCCCCTTCACCAGCAACCCCACCACCAGCTTATAGCCCATCAAAACCACCAGTTAAACCACCTACCCCC

   275       285       295       305       315       325       335       345       355
SerValLysProProAlaLysProProValLysProProThrProSerValLysProProThrProSerValLysProProThrProSer
TCAGTTAAACCACCAGCTAAGCCACCAGTTAAACCACCTACCCCATCAGTTAAACCACCTACCCCATCAGTTAAGCCACCAACACCGTCA

   365       375       385       395       405       415       425       435       445
ProTyrTyrProSerArgLysProValAlaValArgGlyLeuValTyrCysLysProCysLysTyrArgGlyValGluThrLeuAsnLeu
CCTTATTACCCTTCTAGGAAACCTGTAGCTGTTCGTGGCCTTGTTTACTGCAAACCTTGCAAGTATAGAGGGGTTGAAACTTTAAACCTG

   455       465       475       485       495       505       515       525       535
AlaThrProLeuGlnGlyAlaIleValLysLeuAlaCysAsnAsnThrLysLysThrLeuValGluGlnGlyThrThrAspLysAsnGly
GCTACCCCACTCCAGGGAGCGATAGTGAAACTAGCGTGCAACAACACAAAGAAGACACTAGTTGAACAGGGCACAACAGACAAGAATGGA

   545       555       565       575       585       595       605       615       625
PhePheLeuIleLeuProLysMetLeuSerSerGlyAlaTyrHisLysCysLysValPheLeuValSerSerLysAsnThrHisCysAsp
TTCTTCTTGATCTTGCCCAAAATGTTGTCCTCAGGGGCCTACCACAAATGCAAGGTGTTCTTAGTCTCATCAAAGAATACTCACTGCGAT

   635       645       655       665       675       685       695       705       715
ValProThrAsnPheAsnGlyGlyLysSerGlyAlaLeuLeuLysTyrThrProLeuProLysProProAlaThrSerHisLeuProVal
GTCCCAACAAATTTCAATGGTGGAAAATCTGGTGCTCTCTTAAAATACACCCCACTTCCCAAACCACCAGCGACTAGTCATCTCCCTGTT

   725       735       745       755       765       775       785       795        805
LysProProThrPheAspValPheThrValGlyProPheGlyPheGluAlaSerSerLysValProCysLysLys
AAACCCCCAACATTTGATGTATTCACTGTGGGGCCTTTTGGATTCGAAGCCTCAAGCAAGGTGCCTTGCAAAAAATAGTTGATTGGGGAA

   815       825       835       845       855       865       875       885       895

AGAAAGTTAGGAAGAGGAGAAGTAACAAAGGAAAAAAATGTGGAAGGAGTTTTAACTGTTGAAGAATGAGAATTTGTTACCTGTGTTTCG

   905       915       925       935       945       955       965       975       985

TGTATCCTTGTTATTCGAATAATAAATGAGATAGAAGGAAAGTGCCTAGTGTTCTTCTTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Figure 2. Nucleotide and deduced amino acid sequence of the *PhPRP1* cDNA clone. The translation start site consensus sequence is underlined and the polyadenylation signal double underlined. The predicted signal peptide cleavage site is indicated with a red arrow. Two 14 amino acid peptide repeats are colored blue; three internal and external repeats are underlined dark orange. Six conserved cysteine residues are highlighted in orange. The 12-amino acid epitope that is strongly conserved between *Ph*PRP1 and the TTS proteins, and that is targeted by the primary antibody used in later immunoblot analysis, is shown with a dotted underline.

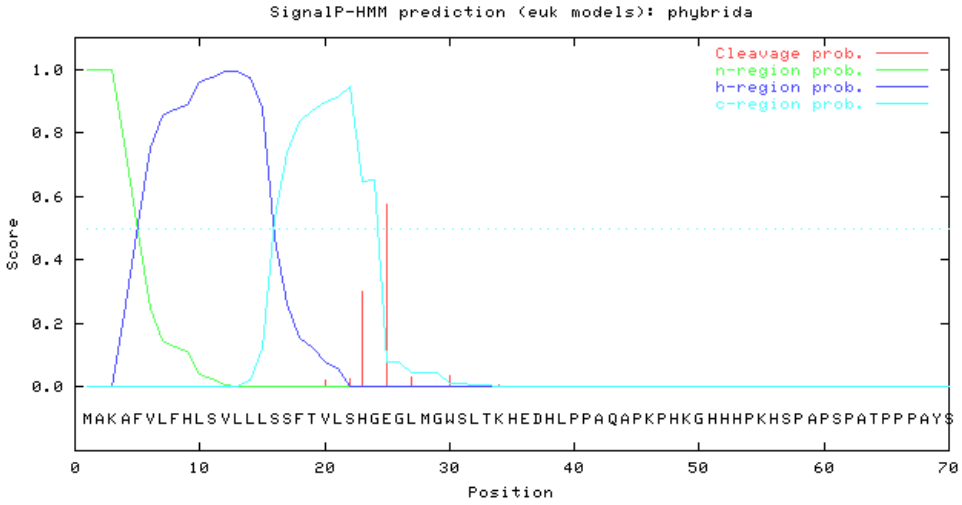
In silico amino acid composition analysis of *Ph*PRP1 (summarized in Table 5) shows that the protein is conspicuously rich in lysine (18%) and proline (11.7%). Tyrosine, abundant in some extracellular extensins, makes up only 2.7% of the protein. There are 18 potential O-glycosylation sites and one potential N-glycosylation site.

Except for the signal peptide, the first 100 residues at the N-terminal end of *Ph*PRP1 are strongly hydrophilic (Figure 3). The N-terminal segment (up to and including residue 120) contains two hypervariable (HV) domains and a histidine rich domain (Figure 4). There are 9 repeats of the lys-pro-pro (KPP) triplet, 6 of them within larger repeats confined to the second hypervariable proline-rich domain in the N-terminal segment (Figures 2 & 4). The C-terminal segment, starting at residue 121, contains interspersed hydrophobic and hydrophilic sections (Figure 3B) and has similarity with the Ole e 1 superfamily of proteins that are known to be potent pollen allergens (Huecas *et al*., 2001). There are six conserved C-terminal cysteines in *Ph*PRP1 specific to the Pollen *Ole e* I domain within the superfamily (Figure 5).

Electrophoretic and immunoblot analysis of pistil eluate proteins

SDS-PAGE analysis of eluates from *P. hybrida* pistils reveals a number of protein bands on silver-stained gels (Figure 6A). Based on the known molecular mass and electrophoretic profile of *P. hybrida* S-RNases (Broothaerts *et al*., 1991; Ai *et al*., 1992), the most abundant proteins in the eluates appear to be S-RNases with a molecular mass between 30-35kDa. A sharply-defined protein band with molecular mass of about 120kDa is also abundant and may be the *Petunia* ortholog of the 120 kDa glycoproteins described from *Nicotiana alata* (Lind *et al*., 1996).

A)



B)

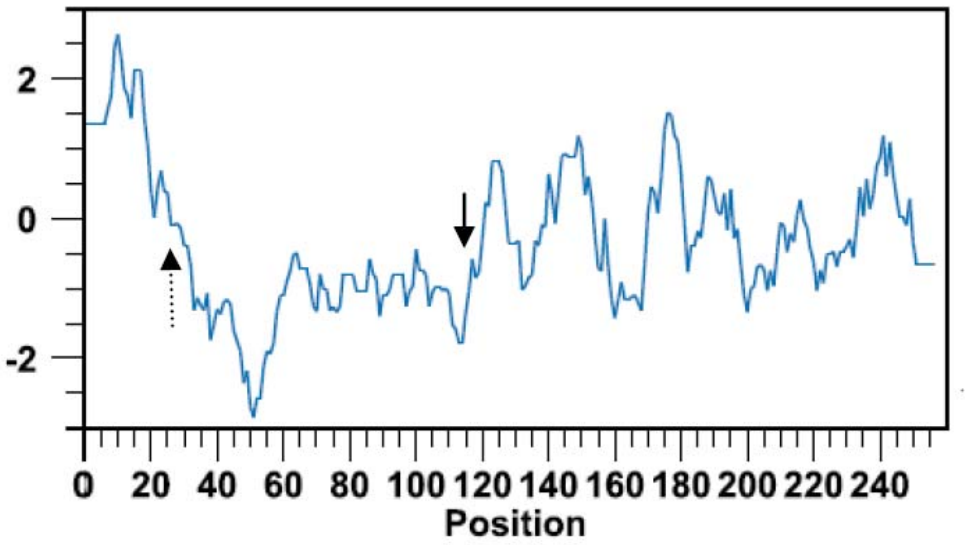


Figure 3. A) Signal peptide prediction by SignalP 3.0 shows peptide cleavage at amino acid 25 (glutamic acid). **B)** Hydropathy analysis of the amino acid sequence of *Ph*PRP1. The end of the signal peptide is denoted with the dotted arrow, and the end of the N-terminal region is noted with the second arrow. The signal peptide is hydrophobic (<0), whereas the rest of the N-terminal half of the protein is highly hydrophilic. Hydrophobic segments alternate with hydrophilic stretches in the C-terminal region of the protein.

Table 6. Amino acid composition of *Ph*PRP1. The N-terminal region (25-113) and C-terminal region (114-256) were analyzed separately, as well as collectively (total composition).

Amino acids	Total Composition	N-Terminal Composition	C-Terminal Composition
Ala (A)	6.2 %	6.7%	5.6 %
Arg (R)	1.2 %	0.0 %	2.1 %
Asn (N)	2.7 %	0.0 %	4.9 %
Asp (D)	1.6 %	1.1 %	2.1 %
Cys (C)	2.3 %	0.0 %	4.2 %
Gln (Q)	1.2 %	1.1 %	1.4 %
Glu (E)	2.0 %	2.2 %	2.1 %
Gly (G)	5.9 %	3.4 %	7.7 %
His (H)	4.7 %	7.9 %	2.1 %
Ile (I)	0.8 %	0.0 %	1.4 %
Leu (L)	9.0 %	3.4 %	9.8 %
Lys (K)	11.7 %	12.4 %	12.6 %
Met (M)	1.2 %	1.1 %	0.7 %
Phe (F)	4.3 %	0.0 %	5.6 %
Pro (P)	18.0 %	34.8 %	10.5 %
Ser (S)	8.6 %	10.1 %	6.3 %
Thr (T)	7.4 %	6.7 %	8.4 %
Trp (W)	0.4 %	1.1 %	0.0 %
Tyr (Y)	2.7 %	2.2 %	3.5 %
Val (V)	8.2 %	5.6 %	9.1 %
Estimated pI: 9.96		**10.07**	**9.79**

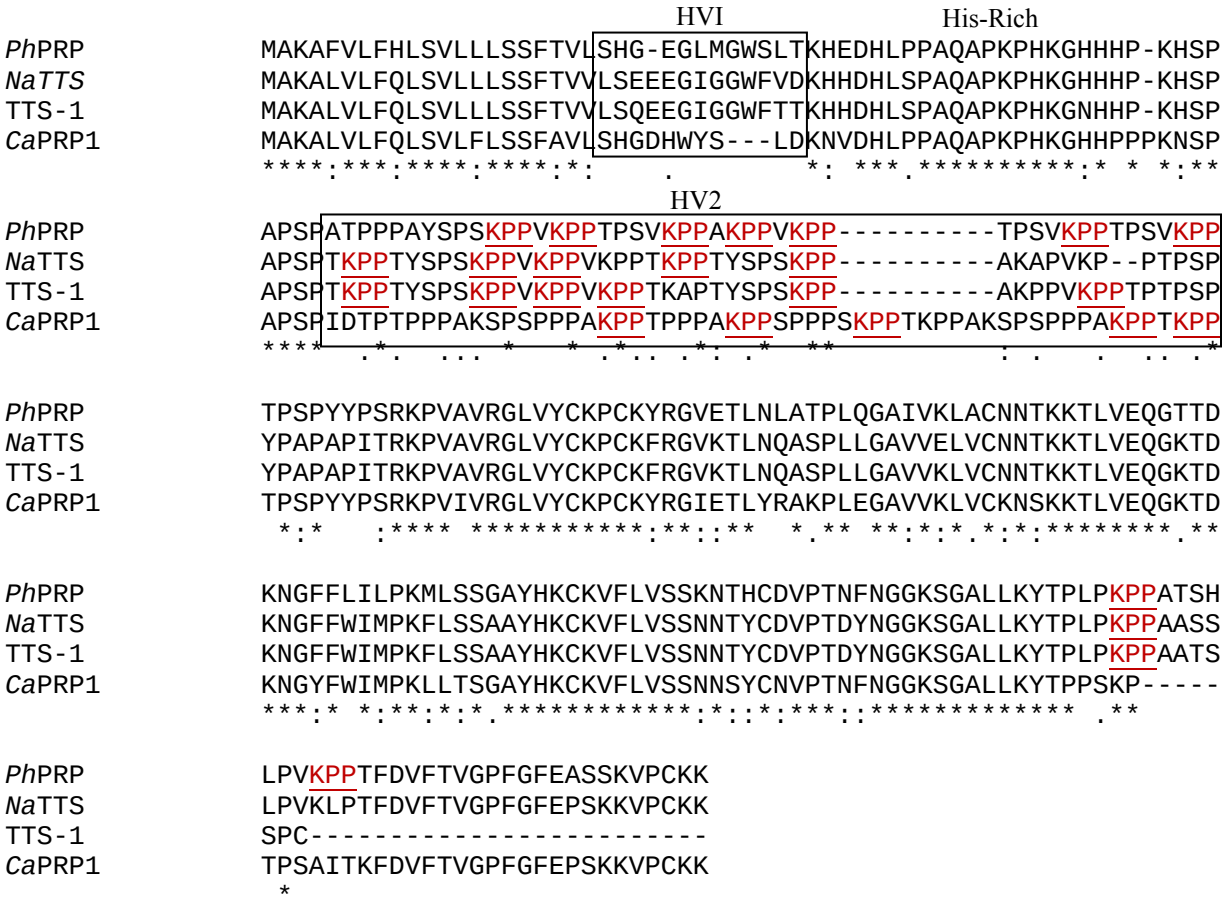


Figure 4. Amino acid alignment with *Ph*PRP1 and three other well characterized proline-rich AGPs expressed in the pistils of solanaceous species. The three proteins other than *Ph*PRP1 are: *Na*TTS from *Nicotiana alata*, TTS-1 from *Nicotiana tabacum*, and *Ca*PRP1 from *Capsicum annuum*. The conserved histidine residues in the histidine-rich domain are shown in purple, and the KPP motifs are shown in red.

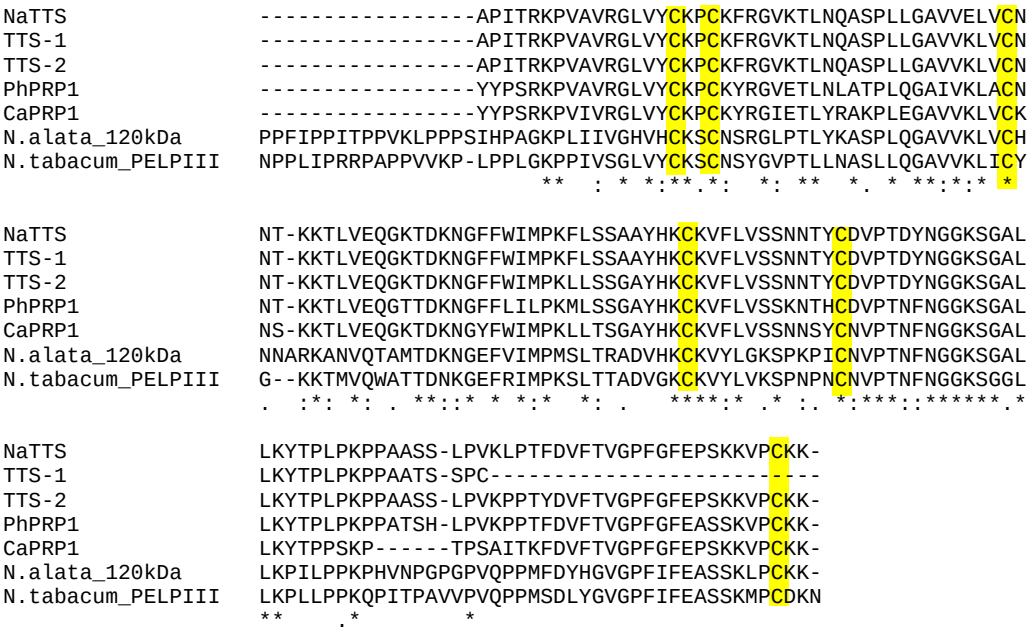


Figure 5. Alignment of the C-terminal domains of *Ph*PRP1 and other AGP/extensin chimeric proteins described from solanaceous pistils. Highlighted are six conserved cysteine residues specific to the Pollen Ole e I domain. The proteins in this alignment are: *Na*TTS from *Nicotiana alata*; TTS-1 and TTS-2 from *Nicotiana tabacum*; PhPRP1 from *Petunia hybrida*; *Ca*PRP1 from *Capsicum annum*; 120kDa from *Nicotiana alata*; PELPIII from *Nicotiana tabacum*.

Immunoblot analysis of pistil eluates revealed a band of heterogenous size migrating at approximately 50-80 kDa (Figure 6B). The cross-reactivity of this band with the anti-*Na*TTS antibody identifies it as the protein encoded by *PhPRP1*.

*Ph*PRP1 purified via ion exchange and His-affinity chromatography migrates as a broad band from 55 kDa to about 80kDa. The identity of this band was confirmed by its cross-reactivity with the anti-TTS antibody. *β*-glucosyl Yariv staining of purified *Ph*PRP1 shows a somewhat diffuse band that ranges from about 55 to 80kDa (Figure 6C).

Phylogenetic analysis of *PhPRP1* and homologs from the Solanaceae and other taxa

A Maximum Likelihood gene tree of the chloroplast *ndhF* (Figure 7A, Wu *et al.,* 2006) is in agreement with gene trees published previously using *ndhF* and the *SAMT* nuclear gene (Bohs & Olmstead, 1997; Martins *et al.,* 2005, respectively). A comparison of these published species trees with my Maximum Likelihood consensus tree generated with *PhPRP1* orthologs identified in GenBank (Figure 7B) shows differences in the grouping of the *Petunia*, *Nicotiana*, and *Capsicum* taxa. Specifically, where the *ndhF* tree shows *P. axillaris* as a sister lineage to *N. tomentosiformis*, which was also a sister lineage to *C. bacatum*, the *PhPRP1* phylogeny groups *Petunia* spp. and *Nicotiana* spp. in a sister clade, and shows *Capsicum annum* as a sister lineage to the *Nicotiana* clade. Polytomies resulting from collapsed branches when the length was 0 appeared in the *Petunia* and *Solanum lycopersicum* clades, which was expected based on sequence homology. In contrast, the Neighbor Joining tree was in agreement with the published phylogeny with respect to the placement of the clades (Figure 7C).

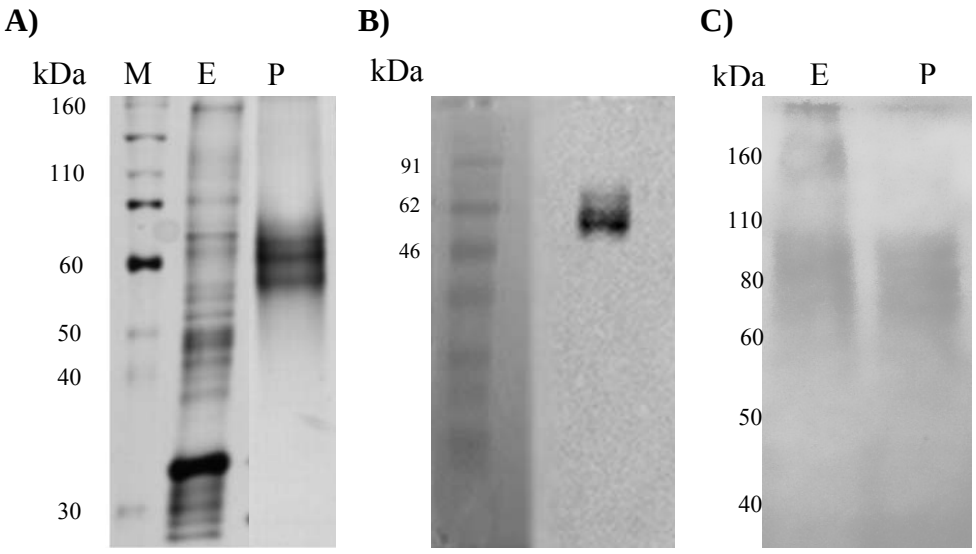


Figure 6. A) Silver stain of pistil eluate (E) and purified *Ph*PRP1 (P) shown alongside the marker lane (M). **B)** Immunoblot of *Ph*PRP1 detected in pistil eluates using an anti-TTS protein antibody. The chemiluminescence image is overlayed with a bright field image to show the protein size marker. **C)** Yariv staining of pistil eluate (E) and purified *Ph*PRP1 (P).

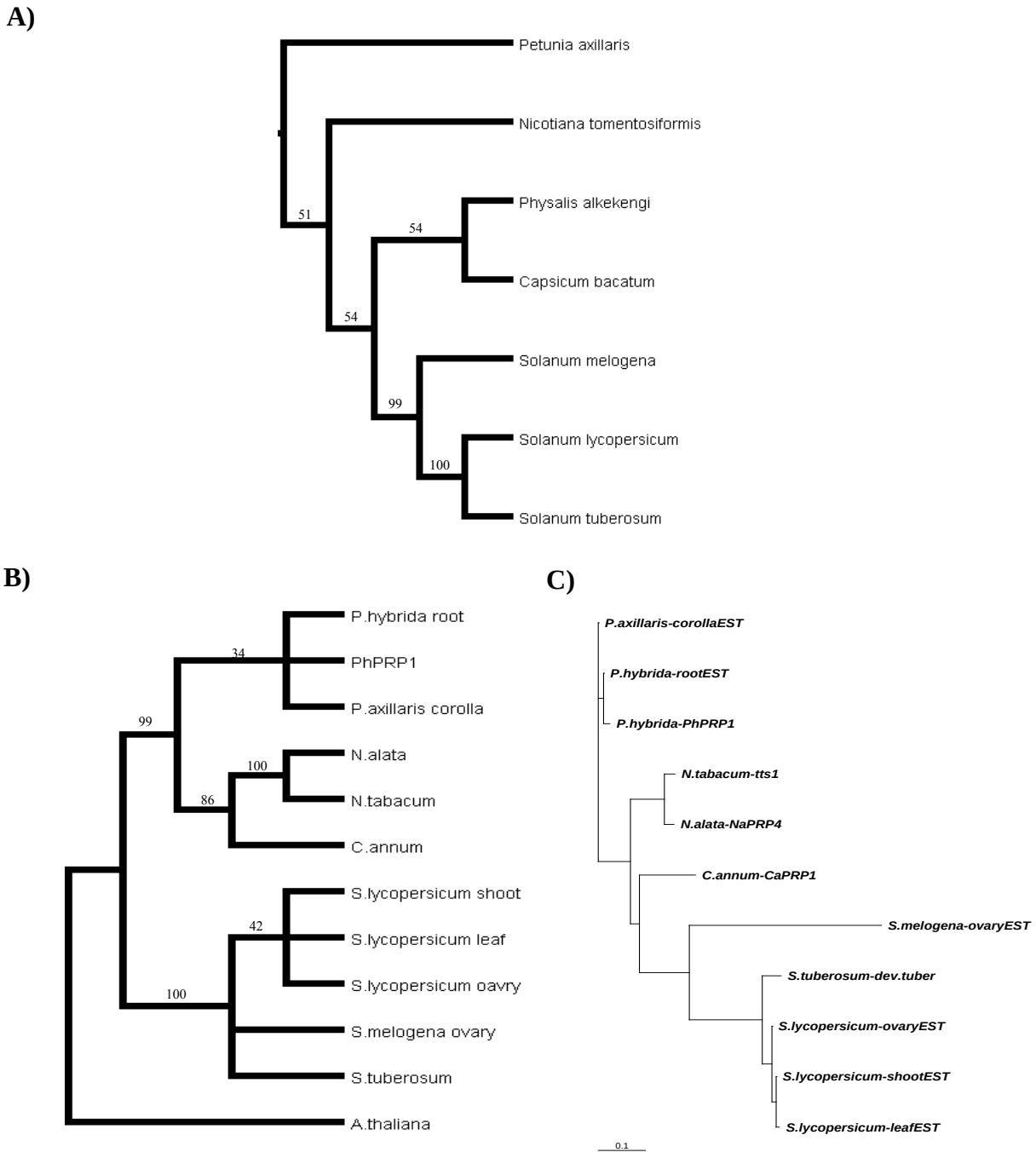


Figure 7. A) Phylogenetic tree showing relationships of solanaceous taxa, re-drawn from the published phylogeny based on chloroplast *ndhF* sequences (Wu *et al.,* 2006). *Arabidopsis thaliana* was set as an outgroup (not shown). **B)** Maximum Likelihood consensus tree showing phylogeny of *PhPRP1* and its homologs, including EST sequences from GenBank. *A. thaliana* was set as an outgroup. **C)** Neighbor Joining tree generated with homologous sequences and no defined outgroup.

DISCUSSION

Arabinogalactan proteins are abundant in the ECM of solanaceous pistils, and some--most notably the TTS proteins and the 120kDa proteins from *Nicotiana*--have been shown to be critical for pollen tube growth (Cheung and Wu, 1999) and the self-incompatibility response (McClure *et al*., 2011), respectively. I obtained the full length cDNA sequence of *PhPRP1* and bioinformatics analysis shows that the encoded protein is an ortholog of *Na*TTS from *Nicotiana alata*, TTS-1 and TTS-2 from *Nicotiana tabacum*, and *Ca*PRP1 from *Capsicum annuum*. Multiple alignments of the known orthologs reveal several distinctive domains, including a histidine-rich conserved domain flanked by two hypervariable domains. I have found transcripts virtually identical to *PhPRP1* in EST libraries made from *P. hybrida* roots (Breullin *et al*., 2010), which suggests that *Ph*PRP1 may be expressed in some vegetative tissues as well.

*Ph*PRP1 is an extracellular arabinogalactan protein expressed in *P. hybrida* pistils

Sequence analysis (Figures 2 and 4) shows strong nucleotide and amino acid similarity between *Ph*PRP1 and the published TTS protein sequences. Based on high sequence similarity and the preservation of domain architecture, I suggest that a clone (*CaPRP*1) identified from a cDNA library prepared from very young roots of *C. annuum* by Mang et al. (2004) is also an ortholog of TTS proteins. As described shortly, my studies indicate that *PhPRP1* expression may not be unique to the transmitting tissue of the pistil. Consequently, I refer to these orthologs collectively as PRPs/TTS proteins, for Proline-Rich Proteins/Transmitting tissue-specific proteins.

The signal sequence is almost perfectly conserved among the three *Nicotiana* TTS proteins, *Ca*PRP1, and *Ph*PRP1, which is one feature that sets these proteins apart from the two other well-known pistil-expressed arabinogalactan proteins, PELPIII and the 120kDa glycoprotein (Figure 8). The variation in signal sequence between these two proteins and the PRPs/TTS proteins is notable considering that they have a common destination: the extracellular matrix (ECM) of the transmitting tissue in the pistil. It is possible that the different signal sequences direct the nascent polypeptide to different subdomains of the endoplasmic reticulum, which results in different patterns of post translational processing between these presumed paralogs.

The histidine-rich domain in the N-terminal segment is another distinctive feature of the PRP/TTS proteins. Histidine-rich domains (HDs) are rare in plant and animal proteins (Hara *et al*., 2005; Poon *et al*., 2011; Zhao and Waite, 2006), as evidenced by the widespread use of 6X His-fusion tags for separating recombinant proteins from cell lysates. The run of three histidine residues, along with four others in the vicinity within the HD, enabled us to use His-affinity beads as part of the strategy for purifying *Ph*PRP1.

Conservation of the histidine-rich domain in PRPs/TTS proteins suggests functional importance. The domain may exert its effects through the metal coordinating ability of the histidine residues, as is known for HDs found in citrus dehydrins (Hara et al., 2005), mussel byssus proteins (Zhao and Waite, 2006) and the multifunctional histidine-rich glycoproteins (HRGs) in human plasma (Poon *et al*., 2011). It has also been suggested that the phosphporylation of histidine residues in eukaryotic cells may be an important regulatory mechanism in signal transduction pathways (reviewed by Bowler and Chua, 1994).

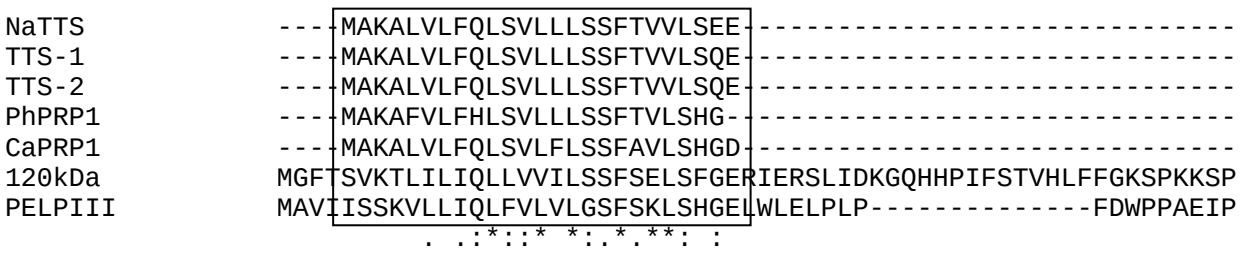


Figure 8. Alignment of the signal sequence portion (boxed) of PRPs/TTS proteins as compared to *Nicotiana alata* 120kDa and *Nicotiana tabacum* PELPIII. Note the high degree of similarity between the PRP/TTS proteins and the marked difference between 120 kDa and PELPIII, both of which are incorporated into the growing pollen tube.

Within the second, much longer proline-rich hypervariable domain, HV2, several lys-pro-pro (KPP) triplets are present. The number and location of these KPP motifs varies even among TTS proteins from closely-related taxa, such as the two species of *Nicotiana* (subfamily Nicotianoidea). With greater evolutionary distance, for example in comparing *Na*TTS to *Ph*PRP1 from *P. hybrida* (subfamily Petunioideae) and to *Ca*PRP1 from *C. annuum* (subfamily Solanoideae), the HV2 region shows yet greater divergence, especially with respect to the position of the KPP motifs. The HV1 and HV2 domains of these proteins appear to have been under positive selection over the phylogenetic history of the Solanaceae, while the constant regions have been under strong purifying selection. It can be speculated that mutations within the highly repetitive HV regions block conspecific pollination and thereby spur speciation between populations of protospecies. Further research is warranted to further elucidate the evolutionary history of these domains and is presently underway (Singh-Cundy, personal communication).

*Ph*PRP1 is highly soluble and readily eluted from bisected pistils in a low-salt buffer and in standard pollen culture media such as the classic medium described by Brewbaker and Kwack (1961). This, along with the absence of a GPI anchor indicates that *Ph*PRP1 is secreted into the extracellular matrix (ECM). Cross-reactivity with the *Na*TTS antibody and heterogeneity in molecular mass on SDS-PAGE gels (molecular mass ranging from 55-100kDa) are further indications that *Ph*PRP1 is a *P. hybrida* ortholog of *Nicotiana* TTS proteins. Purified *Ph*PRP1 transferred to nitrocellulose by electroblotting stains readily with *β*-glucosyl Yariv reagent (Figure 6), further confirming our hypothesis that *Ph*PRP1 is an arabinogalactan like the TTS proteins (Cheung et al., 1993).

Role of PRPs/TTS proteins in enhancing pollen tube growth

Cheung et al. (1995) have demonstrated that supplementing a sucrose-free culture medium with purified TTS protein stimulates the *in vitro* growth of tobacco pollen tubes at least in the first six hours of culture. Wu et al. (1995) observed a basipetal gradient in glycosylation of TTS proteins in *Nicotiana* pistils. In contrast, studies in the Singh-Cundy lab show that in *P. hybrida* pistils the most highly-glycosylated (80-100kDa) forms of *Ph*PRP1 accumulate in both the upper and lower segments of the style. However, a less glycosylated subpopulation (55-80kDa) is abundant in the upper one-third but not in the lower two-thirds of the style (Singh-Cundy, personal communication and Figure 9).

The heterogeneity in the glycosylation PRPs/TTS proteins may reflect the structural and functional differentiation of tissue types along the pollen tube pathway from the stigma to the base of the style. For example, the cells of the transition zone--which is located below the stigma and is a major component of the tissue mass in the upper segment of the pistil--are distinctly different from the cells in the lower style in terms of their ultrastructure and their response to pollination (Herrero and Dickinson, 1977; Lubliner *et al*., 2003). Furthermore, glycosylation of PELPIII proteins exhibits a basipetal pattern, with the most glycosylated forms accumulating preferentially in the lower segment of the style, yet the proteins seem to have no direct role in pollen tube guidance (de Graaf *et al*., 2003).

It has been suggested that the TTS proteins may have an adhesive role in addition to a nutritive one, or rather, that they aid in pollen tube growth by adhering the growing tip to the ECM (Wu *et al.,* 2000). This may be the case, given the sticky nature of the TTS proteins (and likely *Ph*PRP1), their tendency to oligomerize, and the observation that they tend to exist in complexes with other proteins (Wu *et al.,* 2000; Hancock *et al*., 2005).

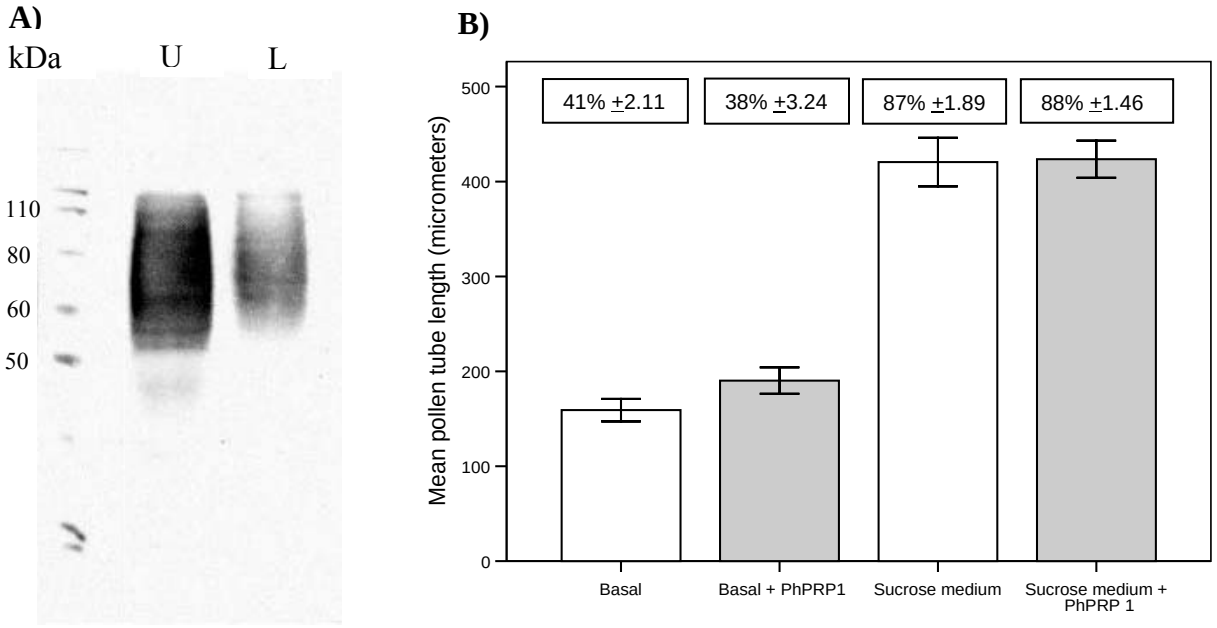


Figure 9. A) Immunoblot analysis of *Ph*PRP1. Total protein extracts made from the upper (U) and lower (L) segments of the stigma/style. Each lane contains 40 µL of a crude extract made by homogenizing 20 pistils segments (upper or lower) in 500 µL of low-salt pistil extraction buffer. **B)** Effect of purified *Ph*PRP1 on *in vitro* pollen germination and tube growth after six hours of culture**.** Mean germination rate for each treatment is shown in the box above each bar. Mean germination rate and pollen tube length are shown ± standard error of the mean at the 95% confidence interval. At least 200 pollen grains or pollen tubes were measured in each replicate. One-way ANOVA was used to compare the various treatments. The approximately 20% difference in mean pollen tube length in basal medium, and basal medium supplemented with purified *Ph*PRP, was significant at $p < 0.001$ (df, 1-214; $F=11.748$). (Bar graph and legend courtesy of Singh-Cundy lab.)

PRPs/TTS proteins may be expressed in rapidly expanding tissues other than the pistil

TTS proteins and *Na*TTS from *Nicotiana* species were originally described as unique to the transmitting tissue within the pistil (reviewed by Cheung *et al.,* 2000). However, Mang et al. (2004) isolated a cDNA clone (*CaPRP1*) from seedling roots of *C. annuum* that has high sequence similarity to the TTS proteins. CaPRP1 also shows strong sequence similarity (78% at the nucleic acid level) to *PhPRP1*. The deduced amino acids sequence of this root cDNA clone from *Capsicum* shows a signal sequence that is nearly identical to that of *Ph*PRP1 (Figure 8), and it also displays the distinctive domain architecture that unifies all known PRPs/TTS proteins (Figure 4). RNA blot analysis conducted under highly stringent conditions shows that the *Ca*PRP1 probe hybridizes to transcripts from young leaves and from pistil tissue (Mang *et al*., 2004). This raises the possibility that PRPs/TTS proteins are not unique to the pistil transmitting tissue, but are expressed in young, rapidly growing vegetative tissues as well. At the very least, this class of proteins can be inferred to have a function in vegetative tissues that are at an early stage of development. Different paralogs of *Ca*PRP1 could be expressed in these diverse vegetative and reproductive tissues. However, the possibility that a single *Ca*PRP1 gene is expressed in all these tissues is strengthened by the discovery of sequences highly similar to the pistil-expressed PRPs/TTS in other solanaceous taxa as well (Table 5).

Furthermore, I have found a clone with high homology to *PhPRP1* in an EST library made from *P. axillaris* corolla tubes (Breullin *et al.*, 2010). Thus, it may be that *PhPRP* is expressed in other rapidly developing tissues within *P. hybrida* as well. The spatial and temporal expression of PRPs/TTS proteins is an issue that needs further investigation. The sequence similarity between *PhPRP1* and the translated *P. axillaris* corolla tube EST clone is striking, differing in only one amino acid in the available sequence (Figure 10).

```
PhPRP1          MAKAFVLFHLSVLLLSSFTVLSHGEGLMGWSLTKHEDHLPPAQAPKPHKGHHHPKHSPAP
P.axillaris     MAKAFVLFHLSVLLLSSFTVLSHGEGLMGWSLTKHEDHLPPAQAPKPHKGHHHPKHSPAP
                ************************************************************

PhPRP1          SPATPPPAYSPSKPPVKPPTPSVKPPAKPPVKPPTPSVKPPTPSVKPPTPSPYYPSRKPV
P.axillaris     SPATPPPAYSPSKPPVKPPTPSVKPPAKPPVKPPSPSVKPPTPSVKPPTPSPYYPSRKPV
                **********************************.*************************
```

Figure 10. Multiple alignment of *Ph*PRP1 and the translated *P. axillaris* EST. The available translated sequence is identical to *Ph*PRP1 with the exception of a threonine-serine residue (highlighted), both of which are polar. The polar nature of both of these residues would likely not impact folding or function of the native protein.

Given that *P. hybrida* originated as an amphidiploid from hybridization of *P. axillaris* and *P. integrifolia*, it is possible that *PhPRP1* represents the gene copy inherited from the *P. axillaris* parent. There may be another copy of *Ph*PRP1 that is highly similar that would likely react with the *Na*TTS antibody and co-migrate on PAGE gels. Identifying the ortholog from *P. integrifolia* in future investigations will help resolve this issue with greater confidence.

Phylogeny of PRPs/TTS Proteins

The hypothesis that *PhPRP1* was inherited by *P. hybrida* from its *P. axillaris* progenitor is supported by the gene tree, which places both *PhPRP1* and the virtually-identical (3 nucleotide difference) *P. axillaris* corolla tube EST clone in a polytomy (Figure 7B). Also striking is the polytomy between *PhPRP1* and the *P. hybrida* root EST (Figure 7B). In this tree, the root EST is too similar to *PhPRP1* to be placed on a separate branch (branch length =0). The lack of differentiation between the two sequences supports the hypothesis that the same gene is expressed in pistil and root tissue of *P. hybrida*, instead of *Ph*PRP1 being paralagous to a root-specific PRP/TTS protein (Figure 10).

The grouping of the *Solanum* species is similar to the *ndhF* gene tree (Figure 7A) with the exception of the grouping of *S. melogena* and *S. tuberosum* on a polytomy rather than grouping *S. tuberosum* with *S. lycopersicum* within the same clade, and the polytomy between the EST sequences from different tissue types is similar to the pattern described above for the *Petunia* clade, in that homologs isolated from different vegetative and reproductive tissues cannot be separated in this phylogenetic analysis. These observations strengthen our hypothesis that PRPs/TTS proteins are not unique to pistil tissues and are in fact expressed in certain other vegetative and reproductive tissues, especially tissues with rapidly-expanding cell types (such as

young roots or expanding corolla tubes in floral tubes). It would be interesting to further characterize these non-pistil-specific homologs to compare and contrast glycosylation patterns and cellular localization. A homolog from *A. thaliana*, AGP31 (Liu and Mehdy, 2010), has all the structural characteristics--including all hypervariable and conserved domains--that I have identified as diagnostic features of PRPs/TTS proteins. Interestingly, this arabinogalactan protein localizes to vascular tissue in roots and leaves of seedlings and is most abundant in pistil tissue (Liu and Mehdy, 2010). A BLAST search of the Arabidopsis Information Resource database (TAIR) using *Ph*PRP1 revealed 19 base pair homologies to response regulator genes involved in cytokinin mediated response pathways expressed in all tissues and during shoot development.

The Maximum Likelihood cladogram of PRP/TTS proteins has much the same topology as the published Solanaceous phylogenies, such as the ones based on the *rbcL* and *ndhF* chloroplast sequences (Olmstead et al., 1997; Wu et al.,2006) and the one based on *SAMT* (salicylic acid methyltransferase), a nuclear gene (Martins and Barkman 2005; Wu *et al.,* 2006). However, there are some differences. For example, published molecular phylogenies place the Petunioideae (the subfamily to which *Petunia* belongs) in a separate clade, which includes the Nicotianoidea (*Nicotiana*) and Solanoideae (*Capsicum*, *Solanum*). The *PhPRP1* gene tree groups *Nicotiana* with *Petunia*, instead of clustering it with the *Solanum* species. On the other hand, the sub-genera groups are completely congruent with the established phylogenies of these taxa, including the separation of *Solanum melongena* (eggplant) from the other solanoideae (*S. lycopersicon*, tomato, and *C. annuum*, capsicum pepper).

In comparison, the Neighbor-Joining (N-J) tree is in complete agreement with the published phylogenies. This could be explained by the fact that the construction of the N-J tree uses an algorithm to calculate the distance of two taxa without making any assumptions about

the rate or pattern of evolution for these sequences. It is likely that reproductive genes, especially those involved in species-specific pollen-pistil interactions, are under very different selective pressures compared to chloroplast genes and genes involved in synthesis of secondary metabolites (e.g. *SMAT*). The rate and pattern of evolution in such reproductive genes may be significantly different from genes that accumulate change at a constant rate over time in different lineages. In addition, assuming a uniform rate of evolution over all domains of the protein would likely not give an accurate representation of phylogeny, as the highly variable HV domains appear to be under a different selection pressure from the His-rich and C-terminal domains. Further investigation into the evolutionary history of the domains of *PhPRP1* and other PRPs/TTS proteins is warranted.

Overall, my research sought to describe the characteristics of the gene *PhPRP1* and compare and contrast them with putative orthologs previously characterized in *Nicotiana* and other species. From my data, I can conclude that *Ph*PRP1 is an extracellular arabinogalactan protein expressed in *P. hybrida* pistils. Based on the inferred homology and phylogeny of *PhPRP1* with sequences found in tissues other than the pistil and within other species, it is likely that *PhPRP1* is not solely expressed in the pistil and likely has a role in developing tissues and rapidly-expanding cell types. The exact function of AGPs in rapidly expanding tissues is unknown, but identification of AGPs in lipid rafts associated with cortical microtubules imply a role in cell morphology (Driouich *et al.,* 2008; Ellis *et al.,* 2010), especially when a disruption to the array of microtubules is seen with AGP interference. Future research will determine the prevalence and relatedness of *PhPRP1* to other species in Petuniodieae and within other solanaceous taxa, as well as possible identification of a *PhPRP1* ortholog in *P. hybrida* that is derived from its *P. integrifolia* progenitor. The functional importance of PRP/TTS proteins

should be assessed in transgenic *Petunia* plants expressing RNAi constructs that suppress *PhPRP1*. Elucidating the functional and developmental importance of *PhPRP1* should further our understanding of pollen-pistil interactions in the Solanaceae and plant reproduction generally.

LITERATURE CITED

Ai Y, Tsai DS, Kao TH. 1992. Cloning and sequencing of cDNAs encoding two S proteins of a self-compatible cultivar of *Petunia hybrida*. Plant Mol. Biol. 19(3): 523-8.

Anderson MA, Cornish EC, Mau S-L, Hoggart WR, Atkinson A, Bonig I, Grego B, Simpson R, Roche PJ, Haley JD, Penschow JD, Niall HD, Treagear GW, Coghlan JP, Crawford RJ, Clarke AE. 1986. Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature. 231: 483-491.

Ando T, Kokubun K, Watanabe H, et al. 2005. Phylogenetic analysis of *Petunia sensu* Jussieu (Solanaceae) using chloroplast DNA RFLP. Ann. Bot. 96: 289-297.

Bacic A, Gell AC, Clark AE. 1988. Arabinogalactan proteins from stigmas of *Nicotiana alata*. Phytochemistry. 27: 679-684.

Baldwin TC, Domingo C, Schindler T, Seetharaman G, Stacey N, Roberts K. 2001. DcAGP1, a secreted arabinogalactan protein, is related to a family of basic proline-rich proteins. Plant Molecular Biology. 45(4): 421-435.

Barnes-Svarney PL, Svarney TE. 1999. The oryx guide to natural history: the earth and all its inhabitants. Phoenix, Arizona: Oryx Press.

Bause E. 1983. Structural requirements for N-glycosylation of proteins. Biochem. Journ. 209: 331-336.

Bendtsen JD, Henrik N, Heijne G von, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. Journ. Molec. Biol. 340: 783-795.

Bosch M, Sommer-Knudsen J, Derksen J, Mariani C. 2001. Class III pistil-specific extensin-like proteins from tobacco have characteristics of arabinogalactan proteins. Amer. Soc. of Plant Phys. 125: 2180-2188.

Bowler C, Hua NH. 1994. Emerging themes of plant signal transduction. The Plant Cell. 6: 1529-1541.

Brady JD, Fry SC. 1997. Formation of di-isodityrosine and loss of isodityrosine in the cell walls of tomato cell-suspension cultures treated with fungal elicitors or H202. Plant Physiology. 115(1):87-92.

Breuillin F, Schramm J, Hajirezaei, M, et al. 2010. Phosphate systemically inhibits development of arbuscular mycorrhiza in *Petunia hybrida* and represses genes involved in mycorrhizal functioning. Plant J 64: 1002-1017.

Broothaerts W, Vanvickenroye P, Decock B, Vandamme J, Vendrig JC. 1991. *Petunia hybrida* S- proteins, ribonuclease activity and the role of their glycan side-chains in self-incompatibility. Sex. Plant Rep. 4(4): 258-266.

Cassab GI. Plant cell wall proteins. 1998. Ann. Rev. Plant Physiol. Plant Mol. Biol. 49: 281-309.

Chae K, Lord EM. 2011. Pollen tube growth and guidance: roles of small, secreted proteins. Ann. Bot. 108(4): 627-636.

Chen C-G, Mau S-L, Clarke AE. 1993. Nucleotide sequence and style-specific expression of a novel proline-rich protein gene from *Nicotiana alata.* Plant Molec. Biol. 21: 391-395.

Cheung AY. 1995. Pollen-pistil interactions in compatible pollination. Proc. Natl. Acad. Sci. 92: 3077-3080.

Cheung AY, Wang H, Wu H-M. 1995. A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. Cell. 82: 383-393.

Clarke AE, Gleeson P, Harrison S, Knox RB. 1979. Form and function of arabinogalactans and arabinogalactan-proteins. Phytochemistry. 18: 521-540.

Clark KR, Okuley JJ, Collins PD, Sims TL. 1990. Sequence variability and developmental expression of *S*-alleles in self-incompatible and pseudo-self-incompatible *Petunia*. The Plant Cell. 2: 815-826.

Cruz-Garcia F, Hancock CN, Kim D, McClure B. 2005. Stylar glycoproteins bind to S-RNase *in vitro*. The Plant Journal. 42: 295-304.

Cruz-Garcia F, Hancock CN, McClure B. 2003. S-RNAse complexes and pollen rejection. Journ. of Exp. Botany. 54(380): 123-130.

Dalgic O, Dane F. 2005. Some of the molecular mechanisms of pollen tube growth and guidance. Asian Journal of Plant Sciences. 4: 702-710.

Driouich A, Baskin TI. 2008. Intercourse between cell wall and cytoplasm exemplified by arabinogalactgan proteins and cortical microtubules. Amer. Journ. Bot. 95(12): 1491-1497.

Fincher GB, Stone BA, Clarke AE. 1983. Arabinogalactan-proteins: Structure, biosynthesis, and function. Ann. Rev. Plant Physiol. 34: 47-70.

Fincher GB, Stone BA, Clarke AE. 1980. Arabinogalactans of sexual and somatic tissues of *Gladiolus* and *Lilium*. Phytochemistry. 19: 1777-1782.

Fulnecek J, Lim KY, Leitch AR, Kovaik A, Matyasek R. 2002. Evolution and structure of 5s rDNA in allotetraploid *Nicotiana tabacum* and its putative parental species. Heredity. 88:19-25.

Fowler TJ, Bernhardt C, Tierney ML. 1999. Characterization and expression of four proline-rich cell wall protein genes in Arabidopsis encoding two distinct subsets of multiple domain proteins. Plant Physiol. 121: 1081-1091.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. (In) John M. Walker (ed): The Proteomics Protocols Handbook. Humana Press. 571-607.

Goldman MHS, Pezzotti M, Seurinck J, Mariane C. 1992. Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins. The Plant Cell. 4: 1041-1051.

Graaf BHJ de, Derksen JWM, Mariani C. 2001. Pollen and pistil in the programic phase. Sex. Plant Reprod. 14: 41-55.

Graaf, BHJ de, Knuiman BA, Derksen J, Mariani C. 2003. Characterization and localization of the transmitting tissue-specific PELPIII proteins of *Nicotiana tabacum*. Journ. Exp. Botany. 54: 55-63.

Gupta R, Jung E, Brunak S. 2004. Prediction of N-glycosylation sites in human proteins. In preparation.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

Hancock CN, Kent L, McClure BA. 2005. The stylar 120 kDa glycoprotein is required for *S*-specific pollen rejection in *Nicotiana*. The Plant Journal. 43: 716-723.

Hara M, Fuginaga M, Kuboi T. 2005. Metal binding by citrus dehydrin with histidine-rich domains. Journ. Exp. Bot. 56(420): 2695-2703.

Herrero M, Dickinson HG. 1980. Pollen tube growth following compatible and incompatible intraspecific pollination in *Petunia hybrida*. Planta. 148: 217-221.

Herrero M, Dickinson HG. 1981. Pollen tube development in *Petunia hybrida* following compatible and incompatible intraspecific matings. J. Cell. Sci. 47: 365-383.

Herrero M, Hormaza JI. 1996. Pistil strategies controlling pollen tube growth. Sex. Plant Reprod. 9: 343-347.

Heslop-Harrison J. 1987. Pollen germination and pollen tube growth. Int. Rev. Cytol. 107: 1-78.

Higashiyama T, Inatsugi R. 2006. The Pollen Tube. Comparative analysis of biological models used in the study of pollen tube growth. Plant Cell Monographs. 3: 265-286.

Hiscock SJ, Allen AM. 2008. Diverse cell signalling pathways regulate pollen-stigma interactions: the search for consensus. New Phytol 179: 286-317.

Hua Z-H, Fields A, Kao T-H. 2008. Biochcemical models for S-RNase based self-incompatibility. Molecular Plant. 1(4): 575-585.

Jefferies D, Farquharson C. 2002. Effects of choice of reverse-transcriptase enzyme and use of T4 gene 32 protein on banding patterns in agarose gel differential display. Analytical Biochem. 308: 192-194.

Jiang S-Y, Jasmin PXH, Ting Y-Y, Ramachandran S. 2005. Genome-wide identification and molecular characterization of Ole_e_I, Allerg_I, and Allerg_2 domain-containing pollen-allergen-like genes in *Oryza sativa*. DNA Research. 12: 167-179.

Kieliszewski M, Kamyab A, Leykam JF, Lamport DTA. 1992. A histidine-rich extensin from *Zea mays* is an arabinogalactan protein. Plant Physiol. 99: 538-547.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics. 23:2947-2948.

Lee, CB, Swatek KN, McClure B. 2008. Pollen proteins bind to the C-terminal domain of *Nicotiana alata* pistil arabinogalactan proteins. Journ. Biol. Chem. 283(40): 26964-26973.

Li Q, Hunt AG. 1995. A near-upstream element in a plant polyadenylation signal consists of more than six nucleotides. Plant Molec. Biol. 28(5): 927-934.

Lind JL, Bacic A, Clarke AE, Anderson MA. 1994. A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan proteins. The Plant Journal. 6: 491-502.

Lind JL, Bonig I, Clarke AE, Anderson MA. 1996. A style-specific 120-kDa glycoprotein enters pollen tubes of *Nicotiana alata* in vivo. Sex. Plant Reprod. 9: 75-86.

Lord EM, Sanders LC. 1992. Roles for the extracellular matrix in plant development and pollination: a special case of cell movement in plants. Developmental Biology. 153: 16-28.

Letunic I, Copley RR, Pils B, Pinkert S, Schultz J, Bork P. 2006. SMART 5: domains in the context of genomes and networks. Nucleic Acids Res. 34: D257-D260.

Lubliner N, Singh-Cundy D, Singh-Cundy A. 2003. Characterization of the pollen growth transition in self-incompatible *Petunia inflate*. Sex. Plant Reprod. 15: 243-253.

Lutcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA. 1987. Selection of AUG initiation codons differs in plants and animals. EMBO Journ. 6(1): 43-48.

Marchler-Bauer A et al. 2007. CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Res. 35: 237-240.

Marshall E, Costa LM, Gutierrez-Marcos J. 2011. Cystein-Rich Peptides (CRPs) mediate diverse aspects of cell-cell communication in plant reproduction and development. Journal of Experimental Botany. 62(5): 1677-1686.

Matsui M, Toyosawa I, Fukuda M. 1995. Purification and characterization of a glycine-rich protein from the aleurone layer of soybean seeds. Biosce. Biotechnol. Biochem. 59: 2231-2234.

McClure, B, Cruz-Garcia, F, Romero, C. 2011. Compatibility and incompatibility in S-RNase-based systems. Ann. Bot. 108(4): 647-658.

McCubbin AG, Kao T-H. 2000. Molecular recognition and response in pollen-pistil interactions. Annu. Rev. of Cell and Dev. Biol. 16: 333-364.

Martins TR, Barkman TJ. 2005. Reconstruction of Solanaceae phylogeny using the nuclear gene SAMT. Syst. Bot. 30(2): 435-447.

Nettancourt D de. 2001. Incompatibility and incongruity in wild and cultivated plants. Berlin: Springer-Verlag.

Poon I, Patel K, Davis D, Parish C, Hulett M. 2011. Histidine-rich glycoprotein: the Swiss Army knife of mammalian plasma. Blood. 117(7): 2093-2101.

Quiapim AC, Brito MS, Bernardes LA, Dasilva I, Malavazi I, DePaoli HC, Molfetta-Machado JB, Giuliatti S, Goldman GH, Goldman MH. 2009. Analysis of the *Nicotiana tabacum* stigma/style transcriptome reveals gene expression differences between wet and dry stigma species. Plant Phys 149:1211-30.

Read SM, Clark AE, Bacic A. 1993. Stimulation of growth of cultured *Nicotiana tabacum* W 38 pollen tubes by poly (ethylene glycol) and Cu (II) salts. Protoplasma. 117: 1-14.

Roberts K. 1994. The plant extracellular matrix: in a new expansive mood. Curr. Opin. Cell Biol. 6: 688-694.

Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawets S, Misener S. (eds) Bioinformatics Methods and protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ. 365-386.

Sanders LC, Lord EM. 1989. Directed movement of latex particles in the gynoecia of three species of flowering plants. Science. 243: 1606-1608.

Sapay N, Guermeur Y, Deleage G. 2006. Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier. BMC Bioinformatics. 7(1): 255.

Shibuya K, Underwood B, Loucas H, Farmerie W, Jones M, Clark D. 2004. Petunia x hybrida EST collection. Unpublished.

Schultz CJ, Ferguson KL, Lahnstein J, Bacic A. 2004. Post-translational modifications of arabinogalactan-peptides of *Arabidopsis thaliana*. Journ. Biol. Chem. 279: 302-305.

Schultz CJ, Hauser K, Lind JL, Atkinson AH, Pu Z, Anderson MA, Clarke AE. 1997. Molecular characterization of a cDNA sequence encoding the backbone of a style-specific 120 kDa glycoprotein which has features of both extensins and arabinogalactan proteins. Plant Molec. Biol. 35: 833-845.

Schultz SR, Curio-Penny B, Li Y, Imani RA, Rydberg L, Geyer PK, Wallrath LL. 2005. Molecular genetic analysis of the nested Drosophila melanogaster lamin C gene. Genetics. 171: 185-196.

Shultz J, Milpetz F, Bork P, Pontig CP. 1998. SMART, a simple modular architecture research tool: Identification of signaling domains. Proc. Natl. Acad. Sci. USA 95: 5857-5864.

Showalter AM. 1993. Structure and function of plant cell wall proteins. Plant Cell. 5: 9-23.

Steer MW, Steer JM. 1989. Pollen tube tip growth. New Phytol. 111: 323-358.

Sommer-Knudsen J, Clarke AE, Bacic A. 1996. A galactose-rich, cell-wall glycoprotein from styles of *Nicotiana alata*. The Plant Journal. 9: 71-83.

Sommer-Knudsen J, Lush WM, Bacic A, Clarke AE. 1998. Re-evaluation of the role of a transmitting tract-specific glycoprotein on pollen tube growth. The Plant Journal. 13: 529-535.

Tang XC, He YQ, Wang Y, Sun MX. 2006. The role of Arabinogalactan proteins binding to Yariv reagents in the initiation, cell developmental fate, and maintenance of microspore embryogenesis in *Brassica napus* L. cv. Topas. Journ. Exp. Bot. 57(11): 2639-2650.

Taylor LP, Hepler PK. 1997. Pollen germination and tube growth. Annu. Rev. Plant Physiol. 48: 461-491.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research. 25(24): 4876-4882.

Villalba M, Batanero E, Monsalve RI, Gonzalez MA, Lahoz C, Rodriguez R. 1994. Cloning and expression of Ole e I, the major allergen from olive tree pollen. Polymorphism analysis and tissue specificity. J. Biol. Chem. 269(21): 15217-15222.

Williams JH. 2008. Novelties of the flowering plant pollen tube underlie diversification of a key life history stage. PNAS. 105(32): 11259-11263.

Wu H-M, Wang H, Cheung AY. 1995. A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. Cell. 82: 393-403.

Wu H-M, Wong E, Ogdahl J, Cheung AY. 2000. A pollen tube growth-promoting arabinogalactan protein from *Nicotiana alata* is similar to the tobacco TTS protein. The Plant Journal. 22: 165-176.

Wu H-M, de Graaf B, Mariani C, Cheung AY. 2001. Hydroxyproline-rich glycoproteins in plant reproductive tissues: structure, functions and regulation. Cell. Mol. Life Sci. 58: 1418-1429.

Ye Z-H, Song Y-R, Marcus A, Varner JE. 1991. Comparative localization of three classes of cell wall proteins. The Plant Journal. 1: 175-183.

Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. Journ. Comput Biol. 7(1): 203-214.

Zhao H, Waite JH. 2006. Proteins in load bearing junctions: The histidine-rich metal-binding protein of mussel byssus. Biochemistry. 45(47): 14223-14231.